Stochastic Optimization
Recalls on probability

V. Leclère

December 1st 2021

École des Ponts ParisTech

UNIVERSITÉ PARIS-EST

Presentation Outline

Presentation Outline

Probability space

- Let Ω be a set.
- A σ-algebra $\mathcal{F}$ of Ω is a collection of subset of Ω such that
  - $\Omega \in \mathcal{F}$
  - $\mathcal{F}$ is closed under complementation
  - $\mathcal{F}$ is closed under countable union
- A measure $\mathbb{P} : \mathcal{F} \to [0,1]$ is a probability if
  - $\mathbb{P}(\Omega) = 1$
  - $\mathbb{P}(\cup_{i\in\mathbb{N}} A_i) = \sum_{i\in\mathbb{N}} \mathbb{P}(A_i)$ where $\{A_i\}_{i\in\mathbb{N}}$ is a collection of pairwise disjoint sets of $\mathcal{F}$
- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space.
- $A \in \mathcal{F}$ is $\mathbb{P}$-almost-sure if $\mathbb{P}(A) = 1$, and negligible if $\mathbb{P}(A) = 0$.
- $(\Omega, \mathcal{F}, \mathbb{P})$ is complete if all subset of a negligible set is measurable.

Measurability and representation

- Let $\mathcal{F}$ be a σ-algebra on Ω.
- A σ-algebra is generated by a collection of sets if it is the smallest containing the collection.
- A function $X : \Omega \to \mathbb{R}^n$ is $\mathcal{F}$-measurable if $X^{-1}(I) \in \mathcal{F}$ for all boxes I of $\mathbb{R}^n$, we note $X \preceq \mathcal{F}$.
- A σ-algebra $\sigma(X)$ is generated by a function $X : \Omega \to \mathbb{R}^n$ sets if it is generated by $\left\{X^{-1}(I) \mid I \text{ boxes of } \mathbb{R}^n\right\}$.
- The σ-algebra generated by all boxes is called the Borel σ-algebra.

Theorem (Doob-Dynkin)

Let $X : \Omega \to \mathbb{R}^n$, $Y : \Omega \to \mathbb{R}^p$ be two $\mathcal{F}$-measurable functions. Then $Y \preceq \sigma(X)$ iff there exists a Borel measurable function $f : \mathbb{R}^n \to \mathbb{R}^p$ such that $Y = f(X)$.

Random variables

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space.
- Define the equivalence class over the $\mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$

$$X \sim Y \quad \iff \quad \mathbb{P}\Big(\{\omega \in \Omega \mid X(\omega) = Y(\omega)\}\Big) = 1$$

- A random variable $\boldsymbol{X}$ is an element of $L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n) := \mathcal{L}^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)/\sim$.
- In other word a random variable is a measurable function from Ω to $\mathbb{R}^n$ defined up to negligeable set.

Expectation and variance

- We recall that $\mathbb{E}\left[\boldsymbol{X}\right] := \int_\Omega \boldsymbol{X}(\omega)\mathbb{P}(d\omega)$.
- If $\mathbb{P}$ is discrete, we have $\mathbb{E}\left[\boldsymbol{X}\right] = \sum_{\omega=1}^{|\Omega|} X(\omega)p_\omega$.
- If $\boldsymbol{X}$ admit a density function f we have $\mathbb{E}\left[\boldsymbol{X}\right] = \int_\mathbb{R} xf(x)dx$.
- We define the variance of $\boldsymbol{X}$

$$var(\boldsymbol{X}) := \mathbb{E}\left[\left(\boldsymbol{X} - \mathbb{E}\left[\boldsymbol{X}\right]\right)^2\right] = \mathbb{E}\left[\boldsymbol{X}^2\right] - \left(\mathbb{E}\left[\boldsymbol{X}\right]\right)^2$$

- and the standard deviation

$$std(\boldsymbol{X}) := \sqrt{var(\boldsymbol{X})}$$

- the covariance is given by

$$cov(\boldsymbol{X}, \boldsymbol{Y}) = \mathbb{E}\left[\boldsymbol{X}\boldsymbol{Y}\right] - \mathbb{E}\left[\boldsymbol{X}\right]\mathbb{E}\left[\boldsymbol{Y}\right]$$

Random variables spaces

- $L^0(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is the set of rv
- $L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is the set of rv such that $\mathbb{E}\big[|\boldsymbol{X}|\big] < +\infty$
- $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is the set of rv such that $\mathbb{E}\big[|\boldsymbol{X}|^p\big] < +\infty$
- $L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is the set of rv that is almost surely bounded
- $L^p(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$, for $p \in]1, +\infty[$ is a reflexive Banach space, with dual L^q, where $\dfrac{1}{p} + \dfrac{1}{q} = 1$
- $L^1(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is a non-reflexive Banach space with dual L^∞
- $L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is a Hilbert space
- $L^\infty(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$ is a non-reflexive Banach space

Independence

- The cumulative distribution function (cdf) of a random variable $\boldsymbol{X}$ is

$$F_X(x) := \mathbb{P}(\boldsymbol{X} \leq x)$$

- Two random variables $\boldsymbol{X}$ and $\boldsymbol{Y}$ are independent iff (one of the following)
    - $F_{X,Y}(a,b) = F_X(a)F_Y(b)$ for all a, b
    - $\mathbb{P}(\boldsymbol{X} \in A, \boldsymbol{Y} \in B) = \mathbb{P}(\boldsymbol{X} \in A)\mathbb{P}(\boldsymbol{Y} \in B)$ for all Borel sets A and B
    - $\mathbb{E}\big[f(\boldsymbol{X})g(\boldsymbol{Y})\big] = \mathbb{E}\big[f(\boldsymbol{X})\big]\mathbb{E}\big[g(\boldsymbol{Y})\big]$ for all Borel functions f and g
- A sequence of identically distributed indenpendent variables is denoted iid.

Inequalities

- (Markov) $\mathbb{P}(|\boldsymbol{X}| \geq a) \leq \frac{\mathbb{E}\left[|\boldsymbol{X}|\right]}{a}$, for $a > 0$.
- (Chernoff) $\mathbb{P}(\boldsymbol{X} \geq a) \leq \frac{\mathbb{E}\left[e^{t\boldsymbol{X}}\right]}{e^{ta}}$, for $t, a > 0$.
- (Chebyshev) $\mathbb{P}(|\boldsymbol{X} - \mathbb{E}\left[X\right]| \geq a) \leq \frac{var(\boldsymbol{X})}{a^2}$, for $a > 0$.
- (Jensen) $\mathbb{E}\left[f(\boldsymbol{X})\right] \geq f(\mathbb{E}\left[\boldsymbol{X}\right])$ for f convex
- (Cauchy-Schwartz) $\mathbb{E}\left[|\boldsymbol{X}\boldsymbol{Y}|\right] \leq \|\boldsymbol{X}\|_2 \|\boldsymbol{Y}\|_2$
- (Hölder) $\mathbb{E}\left[|\boldsymbol{X}\boldsymbol{Y}|\right] \leq \|\boldsymbol{X}\|_p \|\boldsymbol{Y}\|_q$ for $\frac{1}{p} + \frac{1}{q} = 1$
- (Hoeffding) $\mathbb{P}\Big(\boldsymbol{M}_n - \mathbb{E}\left[\boldsymbol{M}_n\right] \geq t\Big) \leq \exp\left(\frac{-2n^2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right)$ where $\{\boldsymbol{X}_i\}_{i\in\mathbb{N}}$ is a sequence of bounded independent rv with $a_i \leq \boldsymbol{X}_i \leq b_i$.

Limits of random variable

Let $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ be a sequence of random variables.

- We say that $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ converges almost surely toward $\boldsymbol{X}$ if

$$\mathbb{P}\Big(\lim_n (\boldsymbol{X}_n - \boldsymbol{X}) = 0\Big) = 1.$$

- We say that $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ converges in probability toward $\boldsymbol{X}$ if

$$\forall \varepsilon > 0, \qquad \mathbb{P}(|\boldsymbol{X}_n - \boldsymbol{X}| > \varepsilon) \to 0.$$

- We say that $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ converges in L^p toward $\boldsymbol{X}$ if

$$\|\boldsymbol{X}_n - \boldsymbol{X}\|_p = \mathbb{E}\Big[|\boldsymbol{X}_n - \boldsymbol{X}|^p\Big] \to 0.$$

- We say that $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ converges in law toward $\boldsymbol{X}$ if

$$\mathbb{E}\big[f(\boldsymbol{X}_n)\big] \to \mathbb{E}\big[f(\boldsymbol{X})\big] \quad \text{for all bounded Lipschitz } f$$

Conditional expectation

- $\mathbb{P}(A|B) = \mathbb{P}(A \cap B)/\mathbb{P}(B)$
- If $(\boldsymbol{X}, \boldsymbol{Y})$ has density $f_{X,Y}$, then the conditional law $(\boldsymbol{X}|\boldsymbol{Y})$ has density $f_{X|Y}(x|y) = f_{X,Y}(x,y)/f_Y(y)$.
- In the continuous case we have

$$\mathbb{E}\left[\boldsymbol{X}|\boldsymbol{Y} = y\right] = \int_{\mathbb{R}} x f_{X|Y}(x|y) dx.$$

- More generally if $\mathcal{G}$ is a sub-sigma-algebra of $\mathcal{F}$, the conditional expectation of $\boldsymbol{X} \in L^1(\Omega, \mathcal{F}, \mathbb{P})$ w.r.t $\mathcal{G}$ is the $\mathcal{G}$-measurable random variable $\boldsymbol{Y}$ satisfying

$$\mathbb{E}\left[\boldsymbol{Y}\mathbb{1}_G\right] = \mathbb{E}\left[\boldsymbol{X}\mathbb{1}_G\right], \qquad \forall G \in \mathcal{G}$$

- Finally, we always have

$$\mathbb{E}\Big[\mathbb{E}\left[\boldsymbol{X}|\boldsymbol{Y}\right]\Big] = \mathbb{E}\left[\boldsymbol{X}\right]$$

Presentation Outline

Monotone and dominated convergence

Theorem (Monotone convergence)

Let $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ *be a sequence of random variables such that*

- $\boldsymbol{X}_{n+1} \geq \boldsymbol{X}_n$ $\mathbb{P}$-*a.s.*
- $\boldsymbol{X}_n \to \boldsymbol{X}_\infty$ $\mathbb{P}$-*a.s.*

then $\lim_{n\to\infty} \mathbb{E}\big[\boldsymbol{X}_n\big] = \mathbb{E}\big[\lim_n \boldsymbol{X}_n\big]$

Theorem (Dominated convergence)

Let $\{\boldsymbol{X}_n\}_{n\in\mathbb{N}}$ *be a sequence of random variables, and* $\boldsymbol{Y}$ *such that*

- $|\boldsymbol{X}_n| \leq \boldsymbol{Y}$ $\mathbb{P}$-*a.s. with* $\mathbb{E}\big[|\boldsymbol{Y}|\big] < +\infty$
- $\boldsymbol{X}_n \to \boldsymbol{X}_\infty$ $\mathbb{P}$-*a.s.*

then $\lim_{n\to\infty} \mathbb{E}\big[\boldsymbol{X}_n\big] = \mathbb{E}\big[\lim_n \boldsymbol{X}_n\big]$

Measurability of multi-valued function

Consider a measurable space $(\Omega, \mathcal{F})$.

- A function $f : \Omega \to \mathbb{R}$ is $\mathcal{F}$-measurable if $f^{-1}(I) \in \mathcal{F}$ for all interval I of $\mathbb{R}$.
- A multi-function $\mathcal{G} : \Omega \rightrightarrows \mathbb{R}^n$ is $\mathcal{F}$-measurable if

$$\forall A \subset \mathbb{R}^n \text{closed}, \quad \mathcal{G}^{-1}(A) := \{\omega \in \Omega \mid \mathcal{G}(\omega) \cap A \neq \emptyset\} \in \mathcal{F}.$$

- A closed valued multi-function $\mathcal{G} : \Omega \rightrightarrows \mathbb{R}^n$ is $\mathcal{F}$-measurable iff $d_x(\omega) := \mathrm{dist}(x, \mathcal{G}(\omega))$ is $\mathcal{F}$-measurable.

Theorem (Measurable selection theorem)

If $\mathcal{G} : \Omega \rightrightarrows \mathbb{R}^n$ is a closed valued measurable multifunction, then there exists a measurable selection of $\mathcal{G}$, that is a measurable function $\pi : \mathrm{dom}(\mathcal{G}) \subset \Omega \to \mathbb{R}^n$ such that $\pi(\omega) \in \mathcal{G}(\omega)$ for all $\omega \in \mathrm{dom}(\mathcal{G})$.

Normal integrand

Assume that $\mathcal{F}$ is $\mathbb{P}$-complete.

Definition (Caratheodory function)

$f : \mathbb{R}^n \times \Omega \to \mathbb{R}$ is a Carathéodory function if

- $f(\cdot, \omega)$ is continuous for a.a. $\omega \in \Omega$
- $f(x, \cdot)$ is measurable for all $x \in \mathbb{R}^n$

Definition (Normal integrand)

$f : \mathbb{R}^n \times \Omega \to \bar{\mathbb{R}}$ is a normal integrand (aka random lowersemicontinuous function) if

- $f(\cdot, \omega)$ is lsc for a.a. $\omega \in \Omega$
- $f(\cdot, \cdot)$ is measurable

f is a convex normal integrand if in addition it is convex in x for a.a. $\omega \in \Omega$.

Measurability of minimum and argmin

Theorem (Measurability of minimum)

Let $f : \mathbb{R}^n \times \Omega \to \bar{\mathbb{R}}$ *be a normal integrand and define*

$$\vartheta(\omega) := \inf_x f(x,\omega) \qquad X^*(\omega) := \arg\min_x f(x,\omega).$$

Then, ϑ *and* X^* *are measurable.*

Theorem (Pointwise minimization)

Let $f : \mathbb{R}^n \times \Omega \to \bar{\mathbb{R}}$ *be a normal convex integrand then*

$$\inf_{\boldsymbol{U}\in L^0, \boldsymbol{U}\in U} \mathbb{E}\Big[f(\boldsymbol{U}(\omega),\omega)\Big] = \mathbb{E}\Big[\inf_{u\in U(\omega)} f(u,\omega)\Big]$$

Continuity and derivation under expectation

Let $f : \mathbb{R}^n \times \Omega$ be a random function (i.e. measurable in ω for all x). We say that f is dominated on X if, for all $x \in X$, there exists an integrable random variable $\boldsymbol{Y}$ such that $f(x,\cdot) \leq \boldsymbol{Y}$ almost surely. If f is dominated on $X \subset \mathbb{R}^n$, we define $F(x) := \mathbb{E}\big[f(x,\omega)\big]$.

- If f is lsc in x and dominated on X, then F is lsc.
- If f is continuous in x and dominated on X, then F is continuous.
- If f is Lispchitz in x, with $\mathbb{E}\big[\mathrm{lip}(f(\cdot,\omega))\big] < +\infty$, then F in Lipschitz continous. Moreover if f is differentiable in x, we have

$$\nabla F(x) = \mathbb{E}\big[\nabla_x f(x,\omega)\big].$$

- If f is a convex normal integrand, and $x_0 \in \mathrm{int}(\mathrm{dom}(F))$, then

$$\partial F(x_0) = \mathbb{E}\big[\partial f(x_0,\omega)\big]$$

Presentation Outline

Strong Law of large number

- We consider a function $f : \mathbb{R}^n \times \Xi \to \mathbb{R}$, and a random variable ξ which takes values in Ξ, and define $F(x) := \mathbb{E}\big[f(x, \xi)\big]$.
- We consider a sequence of random variables $\{\xi_i\}_{i \in \mathbb{N}}$.
- We define the average function

$$\hat{F}_N(x) := \frac{1}{N} \sum_{i=1}^{N} f(x, \xi_i)$$

- We say that we have a Law of Large Number (LLN) if,

$$\forall x \in \mathbb{R}^n, \qquad \mathbb{P}\Big(\lim_n \hat{F}_n(x) = F(x)\Big) = 1$$

- The strong LLN state that LLN holds if $f(x, \xi)$ is integrable, and $\{\xi_i\}_{i \in \mathbb{N}}$ is a iid (with same law as ξ).

Uniform Law of large number

- Having LLN means that, for all $\varepsilon > 0$ (and almost all sample),

$$\forall x, \quad \exists N_\varepsilon \in \mathbb{N}, \qquad n \geq N \quad \Longrightarrow \quad |\hat{F}_N(x) - F(x)| \leq \varepsilon$$

- We say that we have ULLN if for all $\varepsilon > 0$ (and almost all sample),

$$\exists N_\varepsilon \in \mathbb{N}, \quad \forall x, \qquad n \geq N \quad \Longrightarrow \quad |\hat{F}_N(x) - F(x)| \leq \varepsilon$$

or equivalently

$$\exists N \in \mathbb{N} \qquad n \geq N \quad \Longrightarrow \quad \sup_x |\hat{F}_N(x) - F(x)| \leq \varepsilon$$

Theorem

If f is a dominated Caratheodory function on X compact and the sample is iid then we have ULLN on X.

Central Limit Theorem

Theorem

Let $\{\boldsymbol{X}_i\}_{i\in\mathbb{N}}$ *be a sequence of rv iid, with finite second order moments. Then we have*

$$\sqrt{n}\Big(\underbrace{\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{X}_i}_{M_n} - \mathbb{E}\big[\boldsymbol{X}\big]\Big) \to \mathcal{N}(0, std(\boldsymbol{X}))$$

where the convergence is in law.

Monte-Carlo method

- Let $\{\boldsymbol{X}_i\}_{i\in\mathbb{N}}$ be a sequence of rv iid with finite variance.
- We have $\mathbb{P}\Big(M_N \in \Big[\mathbb{E}\big[\boldsymbol{X}\big] \pm \frac{\Phi^{-1}(p) std(\boldsymbol{X})}{\sqrt{N}}\Big]\Big) \approx p$
- In order to estimate the expectation $\mathbb{E}\big[\boldsymbol{X}\big]$, we can
  - sample N independent realizations of $\boldsymbol{X}$, $\{X_i\}_{i\in[\![1,N]\!]}$
  - compute the empirical mean $M_N = \frac{\sum_{i=1}^N X_i}{N}$, and standard-deviation s_N
  - choose an error level p (e.g. 5%) and compute $\Phi^{-1}(1-p/2)$ (1.96)
  - and we know that, asymptotically, the expectation $\mathbb{E}\big[\boldsymbol{X}\big]$ is in $\Big[M_N \pm \frac{\Phi^{-1}(p) s_N}{\sqrt{N}}\Big]$ with probability (on the sample) $1-p$
- In the case of bounded independent variable we can use Hoeffding

$$\mathbb{P}\Big(\mathbb{E}\big[\boldsymbol{X}\big] \in [M_n \pm t]\Big) \geq 2e^{-\frac{2nt^2}{b-a}}$$

Presentation Outline

The (deterministic) newsboy problem

In the 50's a boy would buy a stock u of newspapers each morning at a cost c, and sell them all day long for a price p. The number of people interested in buying a paper during the day is d. We assume that $0 < c < p$.

How shall we model this ?

The (deterministic) newsboy problem

In the 50's a boy would buy a stock u of newspapers each morning at a cost c, and sell them all day long for a price p. The number of people interested in buying a paper during the day is d. We assume that $0 < c < p$.

How shall we model this ?

- Control $u \in \mathbb{R}^+$
- Cost $L(u) = cu - p\min(u, d)$

Leading to

$$\begin{aligned} \min_u \quad & cu - p\min(u, d) \\ s.t. \quad & u \geq 0 \end{aligned}$$

The (stochastic) newsboy problem

Demand d is unknown at time of purchasing. We model it as a random variable $\boldsymbol{d}$ with known law. Note that

- the control $u \in \mathbb{R}^+$ is deterministic
- the cost is a random variable (depending of $\boldsymbol{d}$). We choose to minimize its expectation.

The (stochastic) newsboy problem

Demand d is unknown at time of purchasing. We model it as a random variable $\boldsymbol{d}$ with known law. Note that

- the control $u \in \mathbb{R}^+$ is deterministic
- the cost is a random variable (depending of $\boldsymbol{d}$). We choose to minimize its expectation.

We consider the following problem

$$\begin{aligned} \min_u \quad & \mathbb{E}\big[cu - p\min(u, \boldsymbol{d})\big] \\ s.t. \quad & u \geq 0 \end{aligned}$$

How can we justify the expectation ?

The (stochastic) newsboy problem

Demand d is unknown at time of purchasing. We model it as a random variable $\boldsymbol{d}$ with known law. Note that

- the control $u \in \mathbb{R}^+$ is deterministic
- the cost is a random variable (depending of $\boldsymbol{d}$). We choose to minimize its expectation.

We consider the following problem

$$\begin{aligned} \min_u \quad & \mathbb{E}\big[cu - p\min(u, \boldsymbol{d})\big] \\ s.t. \quad & u \geq 0 \end{aligned}$$

How can we justify the expectation ?
By law of large number: the Newsboy is going to sell newspaper again and again. Then optimizing the sum over time of its gains is closely related to optimizing the expected gains.

Solving the stochastic newsboy problem

For simplicity assume that the demand $\boldsymbol{d}$ has a continuous density f. Define $J(u)$ the expected "loss" of the newsboy if he bought u newspaper. We have

$$\begin{aligned}
J(u) &= \mathbb{E}\Big[cu - p\min(u, \boldsymbol{d})\Big] \\
&= (c-p)u - p\mathbb{E}\Big[\min(0, \boldsymbol{d} - u)\Big] \\
&= (c-p)u - p\int_{-\infty}^{u} (x-u)f(x)dx \\
&= (c-p)u - p\Big(\int_{-\infty}^{u} xf(x)dx - u\int_{-\infty}^{u} f(x)dx\Big)
\end{aligned}$$

Solving the stochastic newsboy problem

For simplicity assume that the demand $\boldsymbol{d}$ has a continuous density f. Define $J(u)$ the expected "loss" of the newsboy if he bought u newspaper. We have

$$\begin{aligned} J(u) &= \mathbb{E}\Big[cu - p\min(u, \boldsymbol{d})\Big] \\ &= (c-p)u - p\Big(\int_{-\infty}^{u} xf(x)dx - u\int_{-\infty}^{u} f(x)dx\Big) \end{aligned}$$

Thus,

$$\begin{aligned} J'(u) &= (c-p) - p\Big(uf(u) - \int_{-\infty}^{u} f(x)dx - uf(u)\Big) \\ &= c - p + pF(u) \end{aligned}$$

where F is the cumulative distribution function (cdf) of $\boldsymbol{d}$. F being non decreasing, the optimum control u^* is such that $J'(u^*) = 0$, which is

$$u^* \in F^{-1}\Big(\frac{p-c}{p}\Big)$$

Newsvendor problem (continued)

We assume that the demand can take value $\{d_i\}_{i\in[\![1,n]\!]}$ with probabilities $\{p_i\}_{i\in[\![1,n]\!]}$.

Newsvendor problem (continued)

We assume that the demand can take value $\{d_i\}_{i \in [\![1,n]\!]}$ with probabilities $\{p_i\}_{i \in [\![1,n]\!]}$.
In this case the stochastic newsvendor problem reads

$$\begin{aligned} \min_{u} \quad & \sum_{i=1}^{n} p_i \Big(cu - p \min(u, d_i) \Big) \\ s.t. \quad & u \geq 0 \end{aligned}$$

Two-stage newsvendor problem

We can represent the newsvendor problem in a 2-stage framework.

- Let u_0 be the number of newspaper bought in the morning.
- let u_1 be the number of newspaper sold during the day.

Two-stage newsvendor problem I

We can represent the newsvendor problem in a 2-stage framework.

- Let u_0 be the number of newspaper bought in the morning.
  $\rightsquigarrow$ first stage control
- let u_1 be the number of newspaper sold during the day.
  $\rightsquigarrow$ second stage control

Two-stage newsvendor problem

We can represent the newsvendor problem in a 2-stage framework.

- Let u_0 be the number of newspaper bought in the morning.
  $\rightsquigarrow$ first stage control
- let u_1 be the number of newspaper sold during the day.
  $\rightsquigarrow$ second stage control

The problem reads

$$
\begin{aligned}
\min_{u_0, \boldsymbol{u}_1} \quad & \mathbb{E}\Big[cu_0 - p\boldsymbol{u}_1\Big] & \\
s.t. \quad & u_0 \geq 0 & \\
& \boldsymbol{u}_1 \leq u_0 & \mathbb{P} - as \\
& \boldsymbol{u}_1 \leq \boldsymbol{d} & \mathbb{P} - as \\
& \boldsymbol{u}_1 \preceq \boldsymbol{d} &
\end{aligned}
$$

Two-stage newsvendor problem II

In extensive formulation the problem reads

$$\begin{aligned} \min_{u_0, \{u_1^i\}_{i \in [\![1,n]\!]}} \quad & \sum_{i=1}^{n} p_i \big(c u_0 - p u_1^i \big) \\ s.t. \quad & u_0 \geq 0 \\ & u_1^i \leq u_0 & \forall i \in [\![1,n]\!] \\ & u_1^i \leq d_i & \forall i \in [\![1,n]\!] \end{aligned}$$

Note that there are as many second-stage control u_1^i as there are possible realization of the demand $\boldsymbol{d}$, but only one first-stage control u_0.

Two-stage newsvendor problem II

In extensive formulation the problem reads

$$\begin{aligned}
\min_{u_0, \{u_1^i\}_{i \in [\![1,n]\!]}} \quad & \sum_{i=1}^{n} p_i \big(c u_0 - p u_1^i \big) \\
s.t. \quad & u_0 \geq 0 \\
& u_1^i \leq u_0 && \forall i \in [\![1,n]\!] \\
& u_1^i \leq d_i && \forall i \in [\![1,n]\!]
\end{aligned}$$

Note that there are as many second-stage control u_1^i as there are possible realization of the demand $\boldsymbol{d}$, but only one first-stage control u_0.

Practical work

- Using julia we are going to model and work around the Newsvendor problem
- Download the files at `https://github.com/leclere/TP-Saclay`
- Start working on the "Newsvendor Problem" up to question 3.