Computing expectations over p-boxes : two views (LP and RS) of the same problem

L. Utkin[1] and S. Destercke[2]

[1]State Forest Technical Academy, St. Petresburg, Russia
[2]IRSN (Institut de Radioprotection et de Sûreté Nucléaire), Cadarache, France

Comp. Sc. Dpt.

IRSN INSTITUT DE RADIOPROTECTION ET DE SÛRETÉ NUCLÉAIRE

DPAM/SEMIC/ LIMSI

Problem statement

- information on r.v. X modeled by a p-box $[\underline{F},\overline{F}]$
- lower ($\underline{\mathbb{E}}$) and upper ($\overline{\mathbb{E}}$) exp. on continuous function $h(X)$:

$$\underline{\mathbb{E}}(h)=\inf_{\underline{F}\leq F\leq\overline{F}}\int_{\mathbb{R}}h(x)\mathrm{d}F,\quad \overline{\mathbb{E}}(h)=\sup_{\underline{F}\leq F\leq\overline{F}}\int_{\mathbb{R}}h(x)\mathrm{d}F \tag{1}$$

- Find optimal distribution F ($\underline{F}(x)\leq F(x)\leq\overline{F}(x)$) for which $\underline{\mathbb{E}},\overline{\mathbb{E}}$ are reached.

Linear programming (LP) general view

- Approximate F by N points $F(x_i)$, $i=1,...,N$ and solve

$$\underline{\mathbb{E}}^*(h)=\inf\sum_{k=1}^{N}h(x_k)z_k \;\text{ or }\; \overline{\mathbb{E}}^*(h)=\sup\sum_{k=1}^{N}h(x_k)z_k$$

subject to

$$z_i\geq 0,\quad i=1,...,N,\ \sum_{k=1}^{N}z_k=1,$$
$$\sum_{k=1}^{i}z_k\leq\overline{F}(x_i),\ \sum_{k=1}^{i}z_k\geq\underline{F}(x_i),\ i=1,...,N.$$

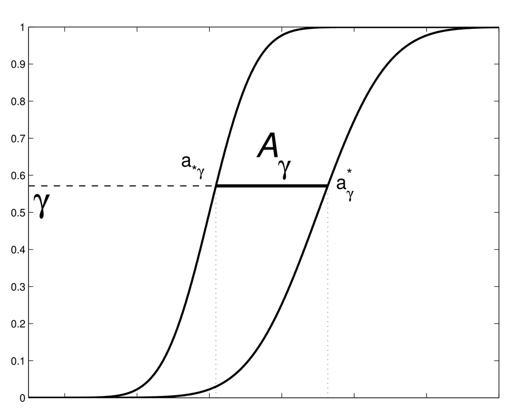
- $\underline{\mathbb{E}}^*,\overline{\mathbb{E}}^*$ are approximations of $\underline{\mathbb{E}},\overline{\mathbb{E}}$
- If N high, computation costs increase, and if N low, approximations can be bad ones

Random set (RS) general view

- Mapping Γ from prob. space to power set $\wp(X)$ of a space X
- Here, mapping from $[0,1]$ with Lebesgue measure to measurable subsets of $\mathbb{R}$.

Given continuous p-box $[\underline{F},\overline{F}]$ $A_\gamma=[a_{*\gamma},a^*_\gamma]$ is the set s.t.

$$a_{*\gamma}:=\sup\{\overline{F}(x)<\gamma\}=\overline{F}^{-1}(\gamma),$$
$$a^*_\gamma:=\inf\{\underline{F}(x)>\gamma\}=\underline{F}^{-1}(\gamma),$$



- P-box $[\underline{F},\overline{F}]$ equivalent continuous random set with unif. density on $[0,1]$ and

$$\Gamma(\gamma)=A_\gamma=[a_{*\gamma},a^*_\gamma]\ \gamma\in[0,1].$$

- Computing $\underline{\mathbb{E}},\overline{\mathbb{E}}$ of h can be reformulated

$$\underline{\mathbb{E}}(h)=\int_0^1\inf_{x\in A_\gamma}h(x)\,d\gamma,\qquad \overline{\mathbb{E}}(h)=\int_0^1\sup_{x\in A_\gamma}h(x)\,d\gamma. \tag{2}$$

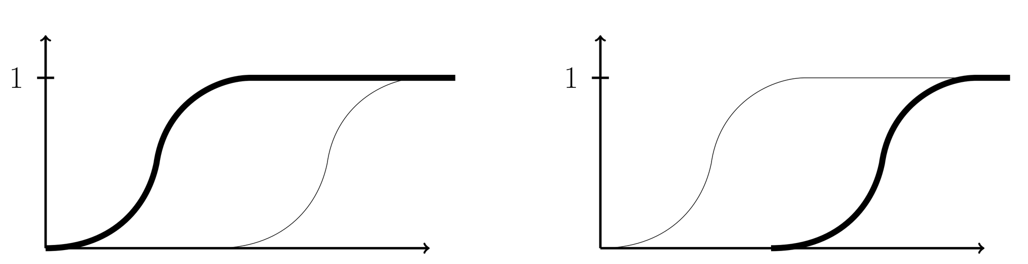
- Solution easily approximated by discretizing p-box on finite number of levels γ_i. Finding inf(sup) on many levels can be difficult, and choosing too few γ_i or poor heuristics can again lead to bad approximations.

> - For both approaches, need to find efficient AND reliable algorithms to compute $\underline{\mathbb{E}},\overline{\mathbb{E}}$.
> - Here, we interest ourselves to the case where h behavior is partially known

The easy case of monotonic functions

- If h is non-decreasing in $\mathbb{R}$, then we have :

$$\underline{\mathbb{E}}(h)=\int_{\mathbb{R}}h(x)\mathrm{d}\overline{F}\ ,\qquad \overline{\mathbb{E}}(h)=\int_{\mathbb{R}}h(x)\mathrm{d}\underline{F} \tag{3}$$



Optimal F for non-decreasing h in bold

One dimension, One maximum

- h has one maximum at point a and is increasing (decreasing) in $(-\infty,a]$ ($[a,\infty)$).

Unconditional expectations

- upper and lower expectations of $h(X)$ on $[\underline{F},\overline{F}]$ are

$$\overline{\mathbb{E}}(h)=\int_{-\infty}^{a}h(x)\mathrm{d}\underline{F}+h(a)\left[\overline{F}(a)-\underline{F}(a)\right]+\int_{a}^{\infty}h(x)\mathrm{d}\overline{F} \tag{4}$$

$$\underline{\mathbb{E}}(h)=\int_{-\infty}^{\overline{F}^{-1}(\alpha)}h(x)\mathrm{d}\overline{F}+\int_{\underline{F}^{-1}(\alpha)}^{\infty}h(x)\mathrm{d}\underline{F} \tag{5}$$

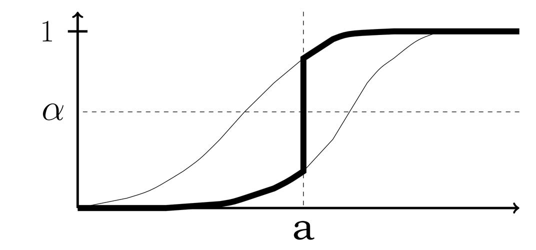
or, equivalently

$$\overline{\mathbb{E}}(h)=\int_0^{\underline{F}(a)}h(a^*_\gamma)d\gamma+[\overline{F}(a)-\underline{F}(a)]h(a)+\int_{\overline{F}(a)}^{1}h(a_{*\gamma})d\gamma \tag{6}$$

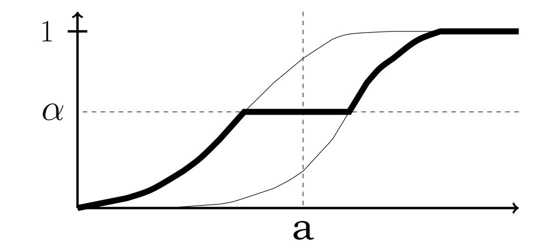
$$\underline{\mathbb{E}}(h)=\int_0^{\alpha}h(a_{*\gamma})d\gamma+\int_{\alpha}^{1}h(a^*_\gamma)d\gamma, \tag{7}$$

where α is one of the solution of the equation

$$h\left(\overline{F}^{-1}(\alpha)\right)=h\left(\underline{F}^{-1}(\alpha)\right). \tag{8}$$



Optimal F for $\overline{\mathbb{E}}(h)$ (vert. jump)



Optimal F for $\underline{\mathbb{E}}(h)$ (hor. jump)

- LP approach suggest to analytically find the level α, or to approximate solution by scanning different values of α.
- Following formula derived with the RS approach

$$\underline{\mathbb{E}}h=\int_0^{\underline{F}(a)}h(a_{*\gamma})d\gamma+\int_{\underline{F}(a)}^{\overline{F}(a)}\min(h(a_{*\gamma}),h(a^*_\gamma))d\gamma+\int_{\overline{F}(a)}^{1}h(a^*_\gamma)d\gamma$$

shows that approximation (either outer or inner) by discretization requires at most 2 computations per discretized levels, if α is unknown.

Conditional expectations

- We suppose that the event $B=[b_0,b_1]$ has been observed. Lower and upper conditional expectations under B are computed as follows :

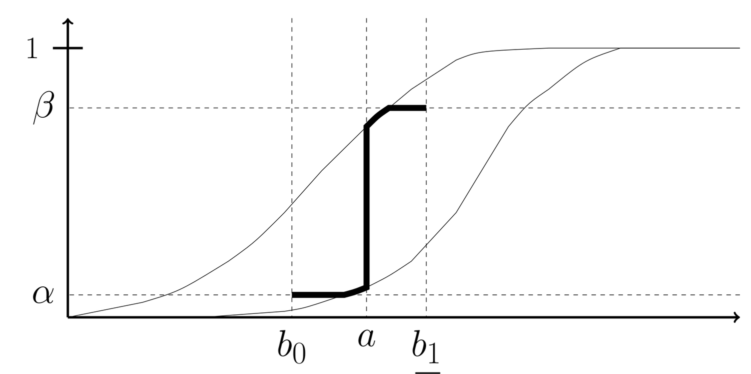
$$\underline{\mathbb{E}}(h|B)=\inf_{\underline{F}\leq F\leq\overline{F}}\frac{\int_{\mathbb{R}}h(x)I_B(x)\mathrm{d}F}{\int_{\mathbb{R}}I_B(x)\mathrm{d}F},\quad \overline{\mathbb{E}}(h|B)=\sup_{\underline{F}\leq F\leq\overline{F}}\frac{\int_{\mathbb{R}}h(x)I_B(x)\mathrm{d}F}{\int_{\mathbb{R}}I_B(x)\mathrm{d}F}.$$

In the case of h having one maximum, these formulas become

$$\overline{\mathbb{E}}(h|B)=\sup_{\substack{\underline{F}(b_0)\leq\alpha\leq\overline{F}(b_0)\\ \underline{F}(b_1)\leq\beta\leq\overline{F}(b_1)}}\frac{1}{\beta-\alpha}\Psi(\alpha,\beta),\quad \underline{\mathbb{E}}(h|B)=\inf_{\substack{\underline{F}(b_0)\leq\alpha\leq\overline{F}(b_0)\\ \underline{F}(b_1)\leq\beta\leq\overline{F}(b_1)}}\frac{1}{\beta-\alpha}\Phi(\alpha,\beta),$$

with

$$\Psi(\alpha,\beta)=\int_{\alpha}^{\beta}\sup_{x\in(A_\gamma\cap B)}h(x)\mathrm{d}\gamma,\quad \Phi(\alpha,\beta)=\int_{\alpha}^{\beta}\inf_{x\in(A_\gamma\cap B)}h(x)\mathrm{d}\gamma.$$



Optimal F for $\overline{\mathbb{E}}(h|B)$

- Numerator and denominator play opposite role in the evolution of expectations (e.g. for the upper one, both increase with the value of $\beta-\alpha$).
- The main problem is to find the couple (α,β) for which extremal expectations are reached. One possibility is to start with $(\alpha,\beta)=(\underline{F}(b_0),\overline{F}(b_1))$ and then to shrink this interval.

Many dimensions, One global maximum

- We assume $h(X,Y)$ is a function from $\mathbb{R}^2\to\mathbb{R}$.

Our uncertainty model about X,Y becomes a bivariate p-box

$$\underline{F}(x,y)\leq F(x,y)\leq\overline{F}(x,y),\ \forall(x,y)\in\mathbb{R}^2.$$

- h has one global maximum at point (x_0,y_0) and is non-increasing in every direction from (x_0,y_0).
- We study how upper/lower expectations can be computed under various assumptions of independence.

Random set corresponding to the marginal p-box of Y given by sets $B_\kappa=[b_{*\kappa},b^*_\kappa]$ s.t.

$$b_{*\kappa}:=\sup\{y\in[b_{inf},b_{sup}]:\overline{F}(y)<\kappa\}=\overline{F}^{-1}(\kappa),$$
$$b^*_\kappa:=\inf\{y\in[b_{inf},b_{sup}]:\underline{F}(y)>\kappa\}=\underline{F}^{-1}(\kappa).$$

Case of strong independence (LP)

- If h separable (i.e. $h(X,Y)=h_1(X)h_2(Y)$), then under strong independence, $\underline{\mathbb{E}}(h)=\underline{\mathbb{E}}(h_1)\cdot\underline{\mathbb{E}}(h_2)$ and $\overline{\mathbb{E}}(h)=\overline{\mathbb{E}}(h_1)\cdot\overline{\mathbb{E}}(h_2)$.
- If h not separable, then, under our assumptions and by LP approach, we get the formula

$$\overline{\mathbb{E}}(h(X,Y))=\sup_{\underline{F}_2\leq F_2\leq\overline{F}_2}\int_{\mathbb{R}}\overline{\mathbb{E}}(h(X,z))\mathrm{d}F_2(z)=\sup\xi(y_0)\left[\overline{F}_2(y_0)-\underline{F}_2(y_0)\right]$$
$$+\int_{-\infty}^{y_0}\sup\xi(z)\mathrm{d}\underline{F}_2(z)+\int_{y_0}^{\infty}\sup\xi(z)\mathrm{d}\overline{F}_2(z)$$

where

$$\sup_{\underline{F}_1\leq F_1\leq\overline{F}_1}\xi(z)=h(x_0,z)\left[\overline{F}_1(x_0)-\underline{F}_1(x_0)\right]+\int_{-\infty}^{x_0}h(x,z)\mathrm{d}\underline{F}_1+\int_{x_0}^{\infty}h(x,z)\mathrm{d}\overline{F}_1.$$

- This explicit formula comes down to concentrate probability mass on (x_0,y_0) and is similar to the one obtained for the univariate case.
- Formula obtained for lower expectation is

$$\underline{\mathbb{E}}(h(X,Y))=\inf_{\underline{F}_2\leq F_2\leq\overline{F}_2}\int_{\mathbb{R}}\underline{\mathbb{E}}(h(X,z))\mathrm{d}F_2(z)$$
$$=\int_{-\infty}^{\overline{F}_2^{-1}(\beta)}\int_{-\infty}^{\overline{F}_1^{-1}(\alpha_z)}h(x,z)\mathrm{d}\overline{F}_1\mathrm{d}\overline{F}_2+\int_{-\infty}^{\overline{F}_2^{-1}(\beta)}\int_{\underline{F}_1^{-1}(\alpha_z)}^{\infty}h(x,z)\mathrm{d}\underline{F}_1\mathrm{d}\overline{F}_2$$
$$+\int_{\underline{F}_2^{-1}(\beta)}^{\infty}\int_{-\infty}^{\overline{F}_1^{-1}(\alpha_z)}h(x,z)\mathrm{d}\overline{F}_1\mathrm{d}\underline{F}_2+\int_{\underline{F}_2^{-1}(\beta)}^{\infty}\int_{\underline{F}_1^{-1}(\alpha_z)}^{\infty}h(x,z)\mathrm{d}\underline{F}_1\mathrm{d}\underline{F}_2.$$

where α_z is a solution of equation $h(\overline{F}_1^{-1}(\alpha),z)=h(\underline{F}_1^{-1}(\alpha),z)$ and β solution of $\underline{\mathbb{E}}(h(X,\underline{F}_2^{-1}(\beta)))=\underline{\mathbb{E}}(h(X,\overline{F}_2^{-1}(\beta))$.

- Again, "transitions" levels α_z,β have to be found, most of the time by numerical approximations.
- For a n dimensional function with one global maximum, n such levels must be found to compute lower expectation.

Case of random set independence (RS)

Given marginal random sets, we have

$$\underline{\mathbb{E}}(h)=\int_0^1\int_0^1\inf_{(x,y)\in[B_\kappa\times A_\gamma]}h(x,y)\mathrm{d}\kappa\mathrm{d}\gamma,\quad \overline{\mathbb{E}}(h)=\int_0^1\int_0^1\sup_{(x,y)\in[B_\kappa\times A_\gamma]}h(x,y)\mathrm{d}\kappa\mathrm{d}\gamma,$$

- Again, solution can be (outer or inner) approximated by discretized levels, the main difficulty being to find the $\inf,\sup$ (here, at most 4 computations are needed per descretized levels)
- From a numerical standpoint, RS ind. equivalent to 1^{st} order Monte-carlo sim. where A_γ,B_κ are randomly sampled.
- Interest : random set independence computationally attractive, while result is an outer approximation of results in case of strong and epistemic independence.

Case of unknown interaction

- Given random set marginals, unknown interaction is qeuivalent to consider every possible joint random sets having those for marginals.
- Method : approximate $[\underline{F},\overline{F}]_X,[\underline{F},\overline{F}]_Y$ with sets $A_{\gamma_i},B_{\kappa_j}$ ($i,j=1,\dots,n$) and where all sets have equal weights. Then compute (for an approximation of lower expectation)

$$\underline{\mathbb{E}}^*(h)=\inf_{\Gamma_{\gamma,\kappa}\in\Gamma^*_{\gamma,\kappa}}\sum\inf_{\substack{x\in A_{\gamma_i}\\ y\in B_{\kappa_j}}}h(x,y)m_{\Gamma_{\gamma,\kappa}}(A_{\gamma_i}\times B_{\kappa_j})$$

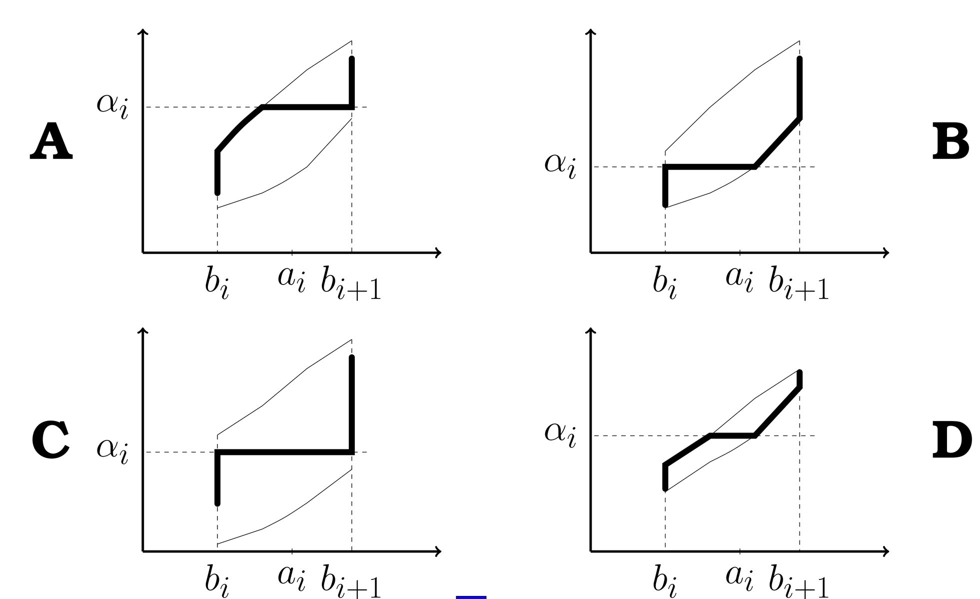
subject to

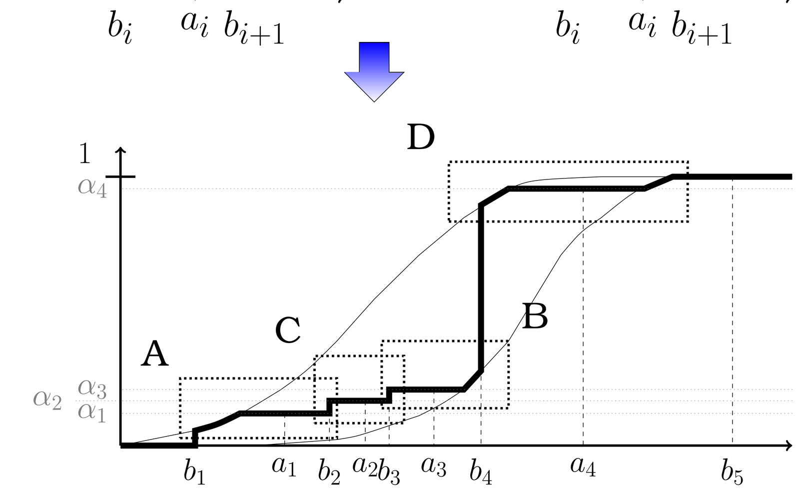
$$\sum_{j=1}^{n}m_{\Gamma_{\gamma,\kappa}}(A_{\gamma_i}\times B_{\kappa_j})=m_{\Gamma_\gamma}(A_{\gamma_i}),\quad \sum_{i=1}^{n}m_{\Gamma_{\gamma,\kappa}}(A_{\gamma_i}\times B_{\kappa_j})=m_{\Gamma_\gamma}(B_{\kappa_j}),$$

where $\Gamma^*_{\gamma,\kappa}$ is the set of joint random sets. $\overline{\mathbb{E}}^*(h)$ is computed by replacing inf with sup.

One dimension, many extrema

- h has alternate local maxima (a_i) and local minima (b_i).
- LP approach shows that optimal F reaching $\underline{\mathbb{E}}(h)$ is a combination of four different local subcases that are part of a large LP problem.





Example of Optimal F with general h which extrema are known

- These four subcases can be found back in the following formula using random sets

$$\underline{\mathbb{E}}(h)=\int_0^{\underline{F}(b_n)}\min_{b_i\in A_\gamma}(h(a_{*\gamma}),h(b_i),h(a^*_\gamma))d\gamma+\int_{\underline{F}(b_n)}^{1}h(a_{*\gamma})d\gamma,$$

- Optimal distribution F is a succession of vertical jumps (prob. mass concentrated on b_i) and of horizontal jumps (to avoid highest values of h)

perspectives

- Pursue investigations on multivariate case, by generalizing existing results to more general functions and to n dimensional case and by exploring the case of epistemic independence
- Design efficient algorithms to make good approximations (i.e. how to find good values for levels α_i with functions having many extrema ?)
- Study various ways to integrate information about dependencies, e.g. by using copulas or adding constraints to LP problems.