Some exercises for the 5th SIPTA Summer School

July 16-20, 2012 – Pescara, Italy

1 Properties of coherent lower & upper previsions

The questions in this section look at a coherent lower prevision $\underline{P}$ on a linear space of gambles $\mathcal{K}$. The aim is to prove a number of properties starting from the basic axioms (which hold for all f and g in $\mathcal{K}$ and all $\lambda > 0$):

(P1) $\underline{P}(f) \geq \inf f$ (accepting sure gains)

(P2) $\underline{P}(\lambda f) = \lambda \underline{P}(f)$ (positive homogeneity)

(P3) $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$ (superlinearity)

The order of the exercises is such that you can build on earlier results. It is sometimes useful to use the shorthand $\overline{P}(f) = -\underline{P}(-f)$ provided by conjugacy.

1. Prove that $\underline{P}(0) = 0$.

   Let $f := 0$ in (P2), then $\underline{P}(0) = \lambda \underline{P}(0)$ for any $\lambda > 0$, which implies $\underline{P}(0) = 0$.

2. Prove that $\inf h \leq \underline{P}(h) \leq \overline{P}(h) \leq \sup h$.

   The first inequality follows from (P1) with $f := h$. The second follows from letting $(f, g) := (h, -h)$ in (P3): $0 = \underline{P}(0) \geq \underline{P}(h) + \underline{P}(-h) = \underline{P}(h) - \overline{P}(-h)$, where we used the result of Exercise 1. The last inequality follows from applying (P1) to $f := -h$: $\overline{P}(h) = -\underline{P}(-h) \leq -\inf(-h) = \sup h$.

3. For $\mathcal{K}$ that contain constant gambles, prove that $\underline{P}(\mu) = \mu$ for all μ in $\mathbb{R}$.

   Apply the result of Exercise 2 with $h = \mu$.

4. Prove that $\underline{P}(h_1 + h_2) \leq \underline{P}h_1 + \overline{P}h_2$ (mixed subadditivity).

   Apply (P3) with $f := h_1 + h_2$ and $g := -h_2$:
   $\underline{P}(h_1) = \underline{P}(h_1 + h_2 - h_2) \geq \underline{P}(h_1 + h_2) + \underline{P}(-h_2) = \underline{P}(h_1 + h_2) - \overline{P}(h_2)$.

5. For $\mathcal{K}$ that contain constant gambles, prove that $\underline{P}(h + \nu) = \underline{P}(h) + \nu$ for all ν in $\mathbb{R}$ (constant additivity).

   From (P3) with $f := h$ and $g := \nu$ and the result of Exercise 4 with $h_1 := h$ and $h_2 := \nu$, we get $\underline{P}(h) + \underline{P}(\nu) \leq \underline{P}(h + \nu) \leq \underline{P}(h) + \overline{P}(\nu)$. From the result of Exercise 3 with $\mu := \nu$ and $\mu := -\nu$, we get $\underline{P}(\nu) = \nu$ and $\overline{P}(\nu) = -\underline{P}(-\nu) = \nu$, so that the result follows.

6. For $h_1 \geq h_2$, prove that $\underline{P}(h_1) \geq \underline{P}(h_2)$ and $\overline{P}(h_1) \geq \overline{P}(h_2)$ (monotonicity).

   If $h_1 \geq h_2$, then $h_1 - h_2 \geq 0$ and so $\underline{P}(h_1 - h_2) \geq 0$ by (P1) and $\overline{P}(h_2 - h_1) \leq 0$ by the result of Exercise 2. Then by the result of Exercise 4 (and its conjugate version), we get $\underline{P}(h_1) + \overline{P}(-h_2) \geq 0$ and $\overline{P}(h_2) + \underline{P}(-h_1) \leq 0$, from which the result follows by conjugacy.

7. Prove that $\underline{P}(|h|) \geq \underline{P}(h)$ and $\overline{P}(|h|) \geq \overline{P}(h)$.

   As $|h| \geq h$, the result follows from the result of Exercise 6 with $h_1 := |h|$ and $h_2 := h$.

8. Prove that $\overline{P}(|h_1 - h_2|) \geq \max\big\{|\underline{P}(h_1) - \underline{P}(h_2)|, |\overline{P}(h_1) - \overline{P}(h_2)|\big\}$.

   We get the inequality $\overline{P}(|h_1 - h_2|) \geq \max\{\overline{P}(h_1 - h_2), \overline{P}(h_2 - h_1)\}$ from the result of Exercise 7 with the choice $h := h_1 - h_2$. By applying the result of Exercise 4 (and its conjugate version) twice to each of the maximands—the second time switching the roles of h_1 and h_2—, the right-hand side becomes $\max\big\{\max\{\overline{P}(h_1) - \overline{P}(h_2), \underline{P}(h_1) - \underline{P}(h_2)\}, \max\{\overline{P}(h_2) - \overline{P}(h_1), \underline{P}(h_2) - \underline{P}(h_1)\}\big\}$. By reorganizing the inner maxima, we obtain the result.

9. Let $(g_n : n \in \mathbb{Z}_{\geq 0})$ be a sequence of gambles. Prove that $\lim_{n\to\infty} \underline{P}(g_n) = \underline{P}(f)$ and $\lim_{n\to\infty} \overline{P}(g_n) = \overline{P}(f)$ if $\lim_{n\to\infty} \overline{P}(|g_n - f|) = 0$.

   As a consequence of the result of Exercise 8 $\lim_{n\to\infty}|\underline{P}(g_n) - \underline{P}(f)| = 0$ and $\lim_{n\to\infty}|\overline{P}(g_n) - \overline{P}(f)| = 0$, which is equivalent to the result.

10. Given, $\underline{P}_1$ and $\underline{P}_2$, two coherent lower previsions on $\mathcal{K}$, prove that any convex combination is also coherent.

    Let $0 \leq \mu \leq 1$ and $\underline{Q} := \mu \underline{P}_1 + (1-\mu)\underline{P}_2$. First (P1):

    $$\underline{Q}(f) = \mu \underline{P}_1(f) + (1-\mu)\underline{P}_2(f) \geq \mu \inf f + (1-\mu) \inf f = \inf f.$$

    Next (P2):

    $$\underline{Q}(\lambda f) = \mu \underline{P}_1(\lambda f) + (1-\mu)\underline{P}_2(\lambda f) = \lambda \mu \underline{P}_1(f) + \lambda(1-\mu)\underline{P}_2(f) = \lambda \underline{Q}(f).$$

Finally (P3):

$$\begin{aligned}\underline{Q}(f+g) &= \mu\underline{P}_1(f+g) + (1-\mu)\underline{P}_2(f+g)\\ &\geq \mu\left(\underline{P}_1(f)+\underline{P}_1(g)\right) + (1-\mu)\left(\underline{P}_2(f)+\underline{P}_2(g)\right) = \underline{Q}(f)+\underline{Q}(g).\end{aligned}$$

2 Few gambles, many lower previsions

The questions in this section essentially deal with avoiding sure loss, coherence, and natural extension on finite spaces. At the end, credal sets and their barycentric representation make an appearance.

Consider a possibility space $\mathcal{X} := \{a,b,c\}$ and a set of gambles $\mathcal{K} := \{g_1, g_2\}$, whose values are given as column vectors in the table on the right. The so-called vacuous lower prevision $\underline{P}_A$ relative to a subset A of $\mathcal{X}$ is defined by $\underline{P}_A(f) := \min_{x\in\mathcal{X}} f(x)$ for any gamble f on $\mathcal{X}$.

	g_1	g_2
a	1	0
b	$1/2$	1
c	0	$1/2$

11. We consider a number of lower previsions on $\mathcal{K}$ defined by

$$\underline{P}_1 := \underline{P}_{\mathcal{X}}, \qquad \underline{P}_2 := \underline{P}_{\{a,b\}}, \qquad \underline{P}_3 := \underline{P}_{\{a\}}, \qquad \underline{P}_4 := \tfrac{1}{3}(\underline{P}_{\{a\}} + \underline{P}_{\{b\}} + \underline{P}_{\{c\}}),$$
$$\underline{P}_5 := \tfrac{3}{4}\underline{P}_{\{c\}} + \tfrac{1}{2}\underline{P}_{\{b\}}, \qquad \underline{P}_6 := \tfrac{1}{3}\underline{P}_{\{c\}} + \tfrac{2}{3}\underline{P}_{\{a,b\}}, \qquad \underline{P}_7 := \tfrac{3}{4}\underline{P}_{\{b\}} + \tfrac{1}{2}\underline{P}_{\{a\}}.$$

Write the values attained by these lower previsions in a table (so a value for g_1 and g_2 in each column).

	$\underline{P}_1$	$\underline{P}_2$	$\underline{P}_3$	$\underline{P}_4$	$\underline{P}_5$	$\underline{P}_6$	$\underline{P}_7$
g_1	0	$1/2$	1	$1/2$	$1/4$	$1/3$	$7/8$
g_2	0	0	0	$1/2$	$7/8$	$1/6$	$3/4$

12. Given a lower prevision $\underline{P}$, then $G_{\underline{P}}(f) := f - \underline{P}(f)$ is the marginal gamble corresponding to the gamble f; here $\underline{P}(f)$ is seen as a constant gamble (vector). Write the matrices $(G_{\underline{P}})_{\mathcal{K}}$ for each of the lower previsions in Exercise 11.

$G_{\underline{P}_1}$	g_1	g_2
a	1	0
b	$1/2$	1
c	0	$1/2$

$G_{\underline{P}_2}$	g_1	g_2
a	$1/2$	0
b	0	1
c	$-1/2$	$1/2$

$G_{\underline{P}_3}$	g_1	g_2
a	0	0
b	$-1/2$	1
c	-1	$1/2$

$G_{\underline{P}_4}$	g_1	g_2
a	$1/2$	$-1/2$
b	0	$1/2$
c	$-1/2$	0

$G_{\underline{P}_5}$	g_1	g_2
a	$3/4$	$-7/8$
b	$1/4$	$1/8$
c	$-1/4$	$-3/8$

$G_{\underline{P}_6}$	g_1	g_2
a	$2/3$	$-1/6$
b	$1/6$	$5/6$
c	$-1/3$	$1/3$

$G_{\underline{P}_7}$	g_1	g_2
a	$1/8$	$-3/4$
b	$-3/8$	$1/4$
c	$-7/8$	$-1/4$

13. (a) Verify that a lower prevision $\underline{P}$ incurs sure loss iff the matrix inequality $(G_{\underline{P}})_{\mathcal{K}}\lambda \leq -1$ holds (row per row) for some column vector λ in $(\mathbb{R}_{\geq 0})^{\mathcal{K}}$.

    $\underline{P}$ incurs sure loss iff $\sum_{g\in\mathcal{K}} \lambda_g G_{\underline{P}}(g) < 0$ for some λ in $(\mathbb{R}_{\geq 0})^{\mathcal{K}}$. In that case $\sum_{g\in\mathcal{K}} \lambda_g G_{\underline{P}}(g) < -\varepsilon$ for some ε in $\mathbb{R}_{>0}$. So then $\sum_{g\in\mathcal{K}} (\lambda_g/\varepsilon) G_{\underline{P}}(g) < -1$ and λ/ε is an element of $(\mathbb{R}_{\geq 0})^{\mathcal{K}}$.
    (The reason for using '≤ -1' instead of '< 0' is that the former is in LP form.)

    (b) Check whether the lower previsions of Exercise 11 avoid sure loss or incur sure loss. For those that incur sure loss: give a λ such that the right-hand side is attained, i.e. such that $\max\big((G_{\underline{P}})_{\mathcal{K}}\lambda\big) = -1$.

    Avoid sure loss: $\underline{P}_1$, $\underline{P}_2$, $\underline{P}_3$, $\underline{P}_4$, $\underline{P}_5$, $\underline{P}_6$.
    Incur sure loss: $\underline{P}_7$, e.g., $\max\big(G_{\underline{P}_7}(g_1) + G_{\underline{P}_7}(g_2)\big) = -1/8$, so take λ to be the constant vector 8.

14. (a) Verify that a lower prevision $\underline{P}$ that avoids sure loss is incoherent iff the matrix inequality $(G_{\underline{P}})_{\mathcal{K}\setminus\{f\}}\lambda - (G_{\underline{P}})_{\{f\}}\mu \leq -1$ holds (row per row) for some f in $\mathcal{K}$, λ in $(\mathbb{R}_{\geq 0})^{\mathcal{K}\setminus\{f\}}$ and μ in $\mathbb{R}_{\geq 0}$.

    Repeat the reasoning from the case of Exercise 13, but now with $\sum_{g\in\mathcal{K}} \lambda_g G_{\underline{P}}(g)$ replaced by $\sum_{g\in\mathcal{K}\setminus\{f\}} \lambda_g G_{\underline{P}}(g) - \mu G_{\underline{P}}(f)$.

    (b) Check whether the lower previsions of that were found to avoid sure loss in Exercise 13 are coherent or incoherent. For those that are incoherent: give an (f,λ,μ)-triple such that the right-hand side is attained, i.e. such that $\max\big((G_{\underline{P}})_{\mathcal{K}\setminus\{f\}}\lambda - (G_{\underline{P}})_{\{f\}}\mu\big) = -1$.

    Coherent: $\underline{P}_1$, $\underline{P}_2$, $\underline{P}_3$, $\underline{P}_4$, $\underline{P}_6$.
    Incoherent: $\underline{P}_5$, e.g., $\max\big(-G_{\underline{P}_5}(g_1) + G_{\underline{P}_5}(g_2)\big) = -1/8$, so take $f := g_1$ and take (λ,μ) to be the constant vector 8.

15. (a) For those lower previsions that were found to be incoherent in Exercise 14, find the smallest dominating coherent lower prevision using natural extension. (From Exercise 14, you already know which—of the two—lower prevision values is too low; write down the linear program and solve it on sight.)

    We only need to find $\underline{E}_{\underline{P}_5}(g_1)$:

$$\begin{aligned}\underline{E}_{\underline{P}_5}(g_1) &= \max\left\{\alpha\in\mathbb{R} \colon \alpha\in\mathbb{R}, g_1 - \alpha \leq \lambda G_{\underline{P}_5}(g_2), \lambda\in\mathbb{R}_{\geq 0}\right\}\\ &= \max\left\{\alpha \colon \begin{pmatrix}-7/8\\ 1/8\\ -3/8\end{pmatrix}\lambda + \begin{pmatrix}\alpha\\ \alpha\\ \alpha\end{pmatrix} \leq \begin{pmatrix}1\\ 1/2\\ 0\end{pmatrix}, \lambda\in\mathbb{R}_{\geq 0}\right\} = 3/8.\end{aligned}$$

    (b) The lower prevision $\underline{P}_6$ is coherent (as you have showed in Exercise 14). Recall from Exercise 11 how it was defined on $\mathcal{K}$. Given the gamble $f := (1/2, 0, 1)$, show that $\underline{E}_{\underline{P}_6}(f) = 0 < 1/3 = \frac{1}{3}\underline{P}_{\{c\}}(f) + \frac{2}{3}\underline{P}_{\{a,b\}}(f)$.

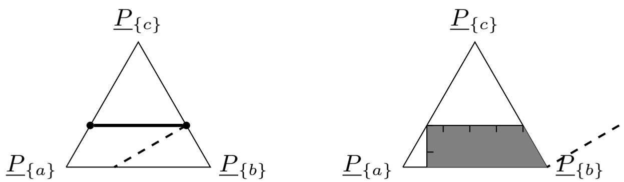
$$\tfrac{1}{3}\underline{P}_{\{c\}}(f) + \tfrac{2}{3}\underline{P}_{\{a,b\}}(f) = \tfrac{1}{3}1 + \tfrac{2}{3}0 = \tfrac{1}{3}$$

$$\underline{E}_{\underline{P}_6}(f) = \max\left\{\alpha \in \mathbb{R} \colon \begin{pmatrix} 2/3 & -1/6 \\ 1/6 & 5/6 \\ -1/3 & 1/3 \end{pmatrix}\lambda + \begin{pmatrix} \alpha \\ \alpha \\ \alpha \end{pmatrix} \leq \begin{pmatrix} 1/2 \\ 0 \\ 1 \end{pmatrix}, \lambda \in (\mathbb{R}_{\geq 0})^{\mathcal{K}}\right\} = 0.$$

The second equality follows immediately from the fact that the b-column of $(G_{\underline{P}})_{\mathcal{K}}$ is strictly positive, weherеas $f(b) = 0$; so we must have $\lambda = 0$ for nonnegative α.

(c) Get additional insight into why the strict inequality of the previous question occurs by constructing (and sketching) the credal sets $\mathcal{M}(\underline{P}_6)$ and $\mathcal{M}(\frac{1}{3}\underline{P}_{\{c\}} + \frac{2}{3}\underline{P}_{\{a,b\}})$. These credal sets are subsets of the simplex of all probability mass functions (column vectors) $p = (p_a, p_b, p_c)$ on $\mathcal{X}$. Add the lines $p^T f = 1/3$ and $p^T f = 0$ to your sketches.

The line $p^T f = 1/3$ goes through $(0, 2/3, 1/3)$ and $(2/3, 1/3, 0)$. The line $p^T f = 0$ goes through $\underline{P}_{\{b\}} = (0, 1, 0)$ and is parallel to the line $p^T f = 1/3$. The lower prevision $\frac{1}{3}\underline{P}_{\{c\}} + \frac{2}{3}\underline{P}_{\{a,b\}}$ is a mixture of two lower previsions of which we know the extreme points of their respective credal sets, so the extreme points of its credal set are among the mixture of these extreme points, i.e., $\frac{1}{3}\underline{P}_{\{c\}} + \frac{2}{3}\underline{P}_{\{a\}}$ and $\frac{1}{3}\underline{P}_{\{c\}} + \frac{2}{3}\underline{P}_{\{b\}}$. To find the lines bounding $\mathcal{M}(\underline{P}_6)$, we just need to find (on sight) two points on the lines $p^T g_1 = \underline{P}_6(g_1) = 1/3$ and $p^T g_2 = \underline{P}_6(g_2) = 1/6$: $(1/3, 0, 2/3)$ and $(0, 2/3, 1/3)$, respectively $(5/6, 1/6, 0)$ and $(2/3, 0, 1/3)$.



16. Consider the lower prevision given by:

	$f(a)$	$f(b)$	$f(c)$	$\underline{P}(f)$
f_1	2	1	0	0.5
f_2	0	1	2	1
f_3	0	1	0	1

(a) Does it avoid sure loss? (Hint: Try to find a dominating linear prevision.)

Yes. It suffices to see that there is a linear prevision that dominates $\underline{P}$ on its domain. The prevision P given by $P(f) = \underline{P}_{\{b\}}(f) = f(b)$ satisfies this.

(b) Is it coherent?
(Hint: Calculate the lower envelope of the set of linear previsions dominating P.)

No. Since $\underline{P}(f_3) = 1 = \max\{f_3(a), f_3(b), f_3(c)\}$, the only linear prevision that dominates $\underline{P}$ for all f is precisely $P = \underline{P}_{\{b\}}$. But P does not coincide with $\underline{P}$ on all gambles: $P(f_1) = 1 > 0.5 = \underline{P}(f_1)$. Since $\underline{P}$ is not the lower envelope of the set $\mathcal{M}(\underline{P})$, we deduce that it is not coherent.

17. Let $\underline{P}$ be the lower prevision on a linear space of gambles $\mathcal{L}(\mathcal{X})$ given by

$$\underline{P}(f) \coloneqq \frac{1}{2}(\min f + \max f)$$

for all f on $\mathcal{X}$. Is it coherent? (Hint: test superadditivity for functions that sum up to a constant.)

No. Since $\underline{P}$ is defined on a linear space, a necessary condition for coherence is that it is supper-additive, meaning that $\underline{P}(f+g) \geq \underline{P}(f) + \underline{P}(g)$ for any pair of gambles f, g. To see that this does not hold, consider $f \coloneqq -I_A$ and $g \coloneqq -I_{\mathcal{X}\setminus A}$. Then $\underline{P}(f) = \underline{P}(g) = -1/2$, while $\underline{P}(f+g) = -1$.

18. Let $\underline{P}$ be a coherent lower prevision on $\mathcal{L}(\mathcal{X})$, where $\mathcal{X} = \{0, 1\}$. Prove that $\underline{P}$ is a *linear-vacuous* mixture, i.e., that there is some $\alpha \in [0, 1]$ and a linear prevision P on $\mathcal{X}$ such that $\underline{P} = \alpha P + (1-\alpha)\underline{P}_{\mathcal{X}}$.

It follows from the coherence of $\underline{P}$ that for every gamble f on $\mathcal{X}$,

$$\underline{P}(f) = \underline{P}(f - \min f) + \min f.$$

Assume that $\underline{P}$ is non-vacuous (otherwise the result is trivial). Then it follows from the above equation that it must be either $\underline{P}(0) > 0$ or $\underline{P}(1) > 0$. Let $\alpha = \underline{P}(0) + \underline{P}(1)$, and let P be the linear prevision associated to the probability $P(0) = \frac{\underline{P}(0)}{\alpha}$ and $P(1) = \frac{\underline{P}(1)}{\alpha}$. Then for every gamble f it holds that

$$\alpha P(f) + (1-\alpha)\underline{P}_{\mathcal{X}}(f) = \underline{P}(0)f(0) + \underline{P}(1)f(1) + (1 - \underline{P}(0) - \underline{P}(1))\min f.$$

Assume for instance that $\min f = f(0)$ (the other case is similar). Then

$$\begin{aligned}\alpha P(f) + (1-\alpha)\underline{P}_{\mathcal{X}}(f) &= \underline{P}(0)f(0) + \underline{P}(1)f(1) + (1 - \underline{P}(0) - \underline{P}(1))f(0) \\ &= \underline{P}(1)(f(1) - f(0)) + f(0) = \underline{P}(f - f(0)) + f(0) = \underline{P}(f).\end{aligned}$$

19. Consider an urn with 10 balls, of which 3 are red, and the other 7 are either blue or yellow.

(a) Determine the set $\mathcal{M}$ of linear previsions that represent the possible compositions of the urn.

The set of possible compositions of the urn is given by the table below left. It produces the set of linear previsions given in the table below right (we give the probability mass functions that are their restrictions to events).

Red	Blue	Yellow
3	0	7
3	1	6
3	2	5
3	3	4
3	4	3
3	5	2
3	6	1
3	7	0

P_i	Red	Blue	Yellow
P_1	$3/10$	0	$7/10$
P_2	$3/10$	$1/10$	$6/10$
P_3	$3/10$	$2/10$	$5/10$
P_4	$3/10$	$3/10$	$4/10$
P_5	$3/10$	$4/10$	$3/10$
P_6	$3/10$	$5/10$	$2/10$
P_7	$3/10$	$6/10$	$1/10$
P_8	$3/10$	$7/10$	0

(b) Let f be a gamble given by $f(\text{blue}) = 2, f(\text{red}) = 1, f(\text{yellow}) = -1$. What is the lower prevision of f?

The lower prevision of f is the lower envelope of the set $\{P_1(f), P_2(f), \ldots, P_8(f)\}$, which in this case is equal to $\underline{P}(f) = 0.3 \cdot 1 - 0.7 \cdot 1 = 0.4$.

(c) Do the same for an arbitrary gamble g.

Again, we have $\underline{P}(f) = \min\{P_1(f), P_2(f), \ldots, P_8(f)\}$; since $P_2, \ldots, P_7$ are convex combinations of P_1, P_8, it follows that

$$\begin{aligned}\underline{P}(f) &= \min\{P_1(f), P_8(f)\}\\ &= \min\{0.3f(\text{blue}) + 0.7f(\text{yellow}), 0.3f(\text{blue}) + 0.7f(\text{red})\}\\ &= 0.3f(\text{blue}) + 0.7\min\{f(\text{yellow}), f(\text{red})\}.\end{aligned}$$

20. Let $\mathcal{X} = \{1, 2, 3\}$, and consider the following sets of desirable gambles:

$$\begin{aligned}\mathcal{K}_1 &:= \{f : f(1) + f(2) + f(3) > 0\}\\ \mathcal{K}_2 &:= \{f : \max\{f(1), f(2), f(3)\} > 0\}.\end{aligned}$$

(a) Are $\mathcal{K}_1, \mathcal{K}_2$ coherent?

To see that the set of gambles $\mathcal{K}_1$ is coherent, we check that it verifies axioms (D1)–(D4) (note that because we are dealing with finite spaces, we can use maximum instead of supremum and minimum instead of infimum):

(D1) Trivially $0 \notin \mathcal{K}_1$.

(D2) If $f \geq 0$ then $f(1) + f(2) + f(3) > 0$, whence $f \in \mathcal{K}_1$.

(D3) If $f, g \in \mathcal{K}_1$, then $f(1) + f(2) + f(3) > 0$ and $g(1) + g(2) + g(3) > 0$. As a consequence, $(f+g)(1) + (f+g)(2) + (f+g)(3) = f(1) + f(2) + f(3) + g(1) + g(2) + g(3) > 0$, and this means that $f + g \in \mathcal{K}_1$.

(D4) Consider $f \in \mathcal{K}_1$ and $\lambda > 0$; then since $f(1) + f(2) + f(3) > 0$, we deduce that $(\lambda f)(1) + (\lambda f)(2) + (\lambda f)(3) = \lambda(f(1) + f(2) + f(3)) > 0$. This implies that $\lambda f \in \mathcal{K}_1$.

To see that the set $\mathcal{K}_2$ is not coherent, it suffices to note that it does not satisfy axiom (D3): consider the gambles f, g given by

$$f(1) = 1, f(2) = f(3) - 2; \quad g(2) = 1, g(1) = g(3) = -2;$$

then it holds that $\max f = \max g = 1$, whence both f and g belong to $\mathcal{K}_2$; however, $(f+g)(1) = (f+g)(2) = -1, (f+g)(3) = -4$, whence $f + g \notin \mathcal{K}_2$.

(b) If they are, what is the lower prevision they induce on the gamble f given by $f(1) = 2, f(2) = 3, f(3) = -1$?

We only need to verify it for $\mathcal{K}_1$. Taking into account the correspondence between sets of desirable gambles and coherent lower previsions, we have that

$$\begin{aligned}\underline{P}(f) &= \sup\{\mu \colon f - \mu \in \mathcal{K}_1\}\\ &= \sup\{\mu \colon (f-\mu)(1) + (f-\mu)(2) + (f-\mu)(3) > 0\}\\ &= \sup\{\mu \colon f(1) + f(2) + f(3) - 3\mu > 0\} = \sup\{\mu \colon 4 - 3\mu > 0\} = \frac{4}{3}.\end{aligned}$$

3 The Choquet integral

Suppose we have a monotone set function μ defined on the power set $2^{\mathcal{X}}$ of a finite set $\mathcal{X}$, that is,

(i) $\mu(\emptyset) = 0$, $\mu(\mathcal{X}) = 1$ and

(ii) $\mu(A) \geq \mu(B)$ for all A, $B \subseteq \mathcal{X}$ such that $A \supseteq B$.

The Choquet integral relative to μ of a gamble f on a set $\mathcal{X}$ is defined as

$$(\mathrm{C})\int f\mathrm{d}\mu \coloneqq \inf f + \int_{\inf f}^{\sup f} \mu(\{x \in \mathcal{X} \colon f(x) \geq t\})\mathrm{d}t = \alpha_0 + \sum_{i=1}^{n} \alpha_i \mu(A_i),$$

where the latter equality holds when $\mathcal{X}$ is finite, as without loss of generality we can write f as a linear combination of indicator functions of nested events as $f = \alpha_0 + \sum_{i=1}^n \alpha_i I_{A_i}$, with $\alpha_0 \in \mathbb{R}$, $\alpha_1 > 0$, $\alpha_2 > 0$, $\dots$, $\alpha_n > 0$ and $A_1 \supset A_2 \supset \cdots \supset A_n$ (where $A \supset B$ means $A \supseteq B$ and $A \neq B$).

If μ is moreover 2-monotone, i.e., additionally satisfies

(iii) $\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B)$ for any A, $B \subseteq \mathcal{X}$,

then we may interpret the values $\mu(A)$ as supremum buying prices for indicator gambles I_A. This corresponds to the lower prevision $\underline{P}$ defined on $\{I_A \colon A \subseteq \mathcal{X}\}$ by $\underline{P}(I_A) \coloneqq \mu(A)$ for all $x \in \mathcal{X}$.

Assume for these exercises that $\mathcal{X}$ is finite.

21. Show that the Choquet integral defines a coherent lower prevision. Use (i.e., do not prove) the super-additivity theorem, which says that $(\mathrm{C})\int (f + g)\mathrm{d}\mu \geq (\mathrm{C})\int f\mathrm{d}\mu + (\mathrm{C})\int g\mathrm{d}\mu$ for all gambles f and g on $\mathcal{X}$.

   $\alpha_0 = \inf f$ by construction and hence $(\mathrm{C})\int f\mathrm{d}\mu \geq \inf f$. Next, it is easily seen that for any $\lambda > 0$, $(\mathrm{C})\int \lambda f\mathrm{d}\mu = \lambda\alpha_0 + \sum_{i=1}^n \lambda\alpha_i\mu(A_i) = \lambda\left((\mathrm{C})\int f\mathrm{d}\mu\right)$, because $\lambda\alpha_0 \in \mathbb{R}$, $\lambda\alpha_1 > 0$, $\lambda\alpha_2 > 0$, $\dots$, $\lambda\alpha_n > 0$, $A_1 \supset A_2 \supset \cdots \supset A_n$, and $\lambda f = \lambda\alpha_0 + \sum_{i=1}^n \lambda\alpha_i I_{A_i}$. The super-additivity theorem gives us the last property that the Choquet integral needs to satisfy for coherence.

22. Show that the lower prevision $\underline{P}$ representing μ is coherent.
[Hint: use the result of Exrcise 21.]

   By 21, $(\mathrm{C})\int \cdot\mathrm{d}\mu$ is coherent. If we establish that $\underline{P}$ is a restriction of $(\mathrm{C})\int \cdot\mathrm{d}\mu$, then $\underline{P}$ must be coherent too. Indeed, we can write the indicator of any event $A \subseteq \mathcal{X}$ as $I_A = 0 + 1I_A$, hence, $(\mathrm{C})\int I_A\mathrm{d}\mu = 0 + 1\mu(A) = \mu(A) = \underline{P}(I_A)$, which shows that $\underline{P}$ is indeed a restriction of the Choquet integral $(\mathrm{C})\int \cdot\mathrm{d}\mu$.

23. Prove that the natural extension of $\underline{P}$ coincides with the Choquet integral with respect to μ. [Hint: show that $(\mathrm{C})\int \cdot\mathrm{d}\mu$ is the point-wise smallest coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$ on its domain; i.e., show that any other such $\underline{P}$-dominating coherent lower prevision $\underline{Q}$ dominates it.]

   Suppose $\underline{Q}$ is another coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$ on its domain, that is, $\underline{Q}(I_A) \geq \mu(A)$ for all $A \subseteq \mathcal{X}$. Let $f \in \mathcal{L}(\mathcal{X})$ and consider its decomposition as in this section's introduction. Since $\underline{Q}$ is coherent, we find by coherence (superlinearity) and $\underline{Q}$-dominance of $\underline{Q}$ that

   $$\underline{Q}(f) = \underline{Q}\big(\alpha_0 + \textstyle\sum_{i=1}^n \alpha_i I_{A_i}\big) \geq \alpha_0 + \sum_{i=1}^n \alpha_i \underline{Q}(I_{A_i}) \geq \alpha_0 + \sum_{i=1}^n \alpha_i\mu(A_i) = (\mathrm{C})\int f\mathrm{d}\mu$$

   since $\underline{Q}(I_{A_i}) \geq \mu(A_i)$ for all $i \in \{1, \dots, n\}$. This establishes the proof.

24. Given are a possibility space $\mathcal{X} \coloneqq \{a, b, c, d\}$, the two gambles below right and the monotone set function μ defined below:

$$\begin{cases} \mu(\mathcal{X}) = 1, & \\ \mu(A) = 1/2, & A \subset \mathcal{X} \text{ and } A \supseteq \{a, b\}, \\ \mu(A) = 0, & A \in \big\{\{b, d\}, \{c, d\}\big\} \text{ or } |A| \leq 1, \\ \mu(A) = 1/4, & \text{otherwise.} \end{cases}$$

	g_1	g_2
a	0	−1
b	1	2
c	1	3
d	2	−2

Calculate $(\mathrm{C})\int g_1\mathrm{d}\mu$ and $(\mathrm{C})\int g_2\mathrm{d}\mu$.

   For each of the two gambles, we first its nested-event decomposition, from which the integral value follows by replacing indicators with the corresponding set function value:

   $$g_1 = 0I_{\mathcal{X}} + 1I_{\{b,c,d\}} + 1I_{\{d\}} \qquad g_2 = -2I_{\mathcal{X}} + 1I_{\{a,b,c\}} + 3I_{\{b,c\}} + 1I_{\{c\}}$$

   $$(\mathrm{C})\int g_1\mathrm{d}\mu = 0\,1 + 1\frac{1}{4} + 1\,0 = \frac{1}{4}, \qquad (\mathrm{C})\int g_2\mathrm{d}\mu = -2\,1 + 1\frac{1}{2} + 3\frac{1}{4} + 1\,0 = -\frac{3}{4}.$$

4 The Möbius transform

Sometimes, a lower probability $\underline{P}$ on $2^{\mathcal{X}}$ is specified using a basic belief assignment m on $2^{\mathcal{X}}$. They are related to each other using Möbius inversion and (recursive) Möbius transformation:

$$\underline{P}(A) = \textstyle\sum_{B\subseteq A} m(A), \quad m(A) = \underline{P}(A) - \sum_{B\subset A} m(A) = \sum_{B\subseteq A}(-1)^{|A\setminus B|}\underline{P}(B).$$

(When m is a probability mass function on $2^{\mathcal{X}}$, $\underline{P}$ is a belief function and completely monotone; for other coherent lower probabilities, some values of m may be negative.)

25. Prove that m is normed (its values sum up to one) if $\underline{P}$ is coherent.

   By coherence, $\underline{P}(\mathcal{X}) = 1$, so by the definition of Möbius inversion, $\sum_{A\in 2^{\mathcal{X}}} m(A) = \sum_{B\subseteq A} m(A) = 1$.

26. Prove that $m(\emptyset) = 0$ if $\underline{P}$ is coherent.

By coherence, $\underline{P}(\emptyset) = 0$, so by the definition of the (recursive) Möbius transformation, $m(\emptyset) = \underline{P}(\emptyset) - 0 = 0$.

27. Prove that $m(\{x\}) \geq 0$ for every x in $\mathcal{X}$ if $\underline{P}$ is coherent.

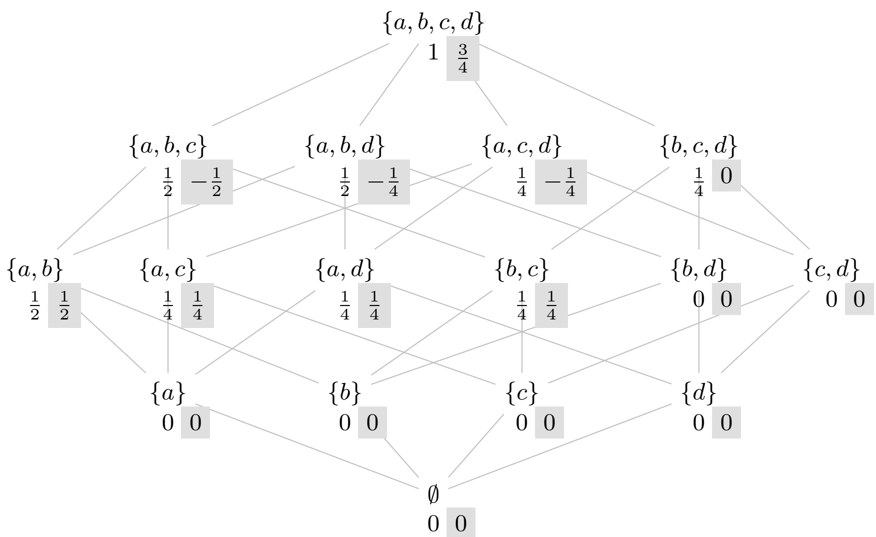
By coherence, $\underline{P}(\{x\}) \geq 0$, so by the definition of the (recursive) Möbius transformation, $m(\{x\}) = \underline{P}(\{x\}) - m(\emptyset) = \underline{P}(\{x\}) - 0 \geq 0$.

28. Given are a possibility space $\mathcal{X} := \{a, b, c, d\}$ and the lower probability $\underline{P}$ defined below:

$$\begin{cases} \underline{P}(\mathcal{X}) = 1, & \\ \underline{P}(A) = 1/2, & A \subset \mathcal{X} \text{ and } A \supseteq \{a, b\}, \\ \underline{P}(A) = 0, & A \in \big\{\{b, d\}, \{c, d\}\big\} \text{ or } |A| \leq 1, \\ \underline{P}(A) = 1/4, & \text{otherwise.} \end{cases}$$

Calculate the corresponding basic belief assignment m corresponding to $\underline{P}$. Do this both recursively and non-recursively. [Hint: to do this efficiently, draw the partial order $(2^{\mathcal{X}}, \subseteq)$ and write down the values $\underline{P}(A)$ below the events A in the order; leave room to write $m(A)$ next to it (in a different color).]

The partial order, $\underline{P}$, and m (gray background) are given below.



The advantage of the recursive transform is that there is only one multiplication with -1 and no set cardinalities need to be determined.

29. Let $\mathcal{X} := \{a, b, c\}$. Given is a set of probability intervals, i.e., lower and upper probabilities $[\underline{P}(\{x\}), \overline{P}(\{x\})]$ for all x in $\mathcal{X}$. Assume the given lower and upper probability values are coherent.

(a) What are the extensions of $\underline{P}$ and $\overline{P}$ to the whole of $2^{\mathcal{X}}$?

By coherence, we have $\underline{P}(\emptyset) = \overline{P}(\emptyset) = 0$ and $\underline{P}(\mathcal{X}) = \overline{P}(\mathcal{X}) = 1$. By conjugacy, we have $\underline{P}(\mathcal{X} \setminus \{x\}) = 1 - \overline{P}(\{x\})$ and $\overline{P}(\mathcal{X} \setminus \{x\}) = 1 - \underline{P}(\{x\})$ for all x in $\mathcal{X}$.

(b) What is the Möbius transform m of $\underline{P}$, as defined on $2^{\mathcal{X}}$ just above?

Using the recursive Möbius transform:

$$\begin{aligned} m(\emptyset) &= \underline{P}(\emptyset) = 0, \\ \forall x \in \mathcal{X} : m(\{x\}) &= \underline{P}(\{x\}), \\ \forall x \in \mathcal{X} : m(\mathcal{X} \setminus \{x\}) &= \underline{P}(\mathcal{X} \setminus \{x\}) - \textstyle\sum_{y \in \mathcal{X} \setminus \{x\}} \underline{P}(\{y\}) \\ &= 1 - \overline{P}(\{x\}) - \textstyle\sum_{y \in \mathcal{X} \setminus \{x\}} \underline{P}(\{y\}), \\ m(\mathcal{X}) &= 1 - \textstyle\sum_{x \in \mathcal{X}} \big(1 - \overline{P}(\{x\}) - \sum_{y \in \mathcal{X} \setminus \{x\}} \underline{P}(\{y\})\big) \\ &\quad - \textstyle\sum_{x \in \mathcal{X}} \underline{P}(\{x\}) \\ &= \textstyle\sum_{x \in \mathcal{X}} \big(\underline{P}(\{x\}) + \overline{P}(\{x\})\big) - 2 \end{aligned}$$

(c) Show that $m(\mathcal{X} \setminus \{x\}) \geq 0$ for all x in $\mathcal{X}$ is necessary for coherence of $\underline{P}$.

From the expression for $m(\mathcal{X} \setminus \{x\})$ in terms of values of $\underline{P}$ encountered above, we see that $m(\mathcal{X} \setminus \{x\}) \geq 0$ is equivalent to $\underline{P}(\mathcal{X} \setminus \{x\}) \geq \sum_{y \in \mathcal{X} \setminus \{x\}} \underline{P}(\{y\})$, which expresses superlinearity, a property necessary for coherence.

(d) Find a coherent set of probability intervals such that $\underline{P}(\mathcal{X}) < 0$.
[Hint: take uniform lower and upper probability mass functions; draw the corresponding credal set to ensure the values you choose are coherent.]

Take $[\underline{P}(\{x\}), \overline{P}(\{x\})] := [0, \beta]$ for all x in $\mathcal{X}$. We get $\underline{P}(\mathcal{X}) = 3\beta - 2$, which is negative if $\beta < 2/3$, so β. From our credal set drawing, we see that moreover $\beta \geq 1/2$ is required for coherence. So for $1/2 \leq \beta < 2/3$ we have $m(\mathcal{X}) < 0$.

(e) Consider the coherent set of probability intervals $[0.2, 0.4]$, $[0.3, 0.5]$, and $[0.1, 0.4]$ for a, b, and c, respectively. (Coherence can be verified by drawing the corresponding credal set.) Use the results you derived above to show that

- $m(\mathcal{X}) = -0.1$,
  Using the formula derived above, we have:

$$m(\mathcal{X}) = \textstyle\sum_{x \in \mathcal{X}} \big(\underline{P}(\{x\}) + \overline{P}(\{x\})\big) - 2 = 1.9 - 2 = -0.1.$$

- $\underline{P}$ is not 3-monotone, hence is not a belief function.

3-monotonicity is equivalent to

$$\underline{P}(A \cup B \cup C) + \underline{P}(A \cap B) + \underline{P}(A \cap C) + \underline{P}(B \cap C) \geq \underline{P}(A) + \underline{P}(B) + \underline{P}(C) + \underline{P}(A \cap B \cap C)$$

for all $A, B, C \subseteq \mathcal{X}$. We need only find one such triple for which this inequality does not hold to give a proof. Take $A := \{a, b\}$, $B := \{a, c\}$, and $C := \{b, c\}$, then the inequality becomes $\underline{P}(\mathcal{X}) + \sum_{x \in \mathcal{X}} \underline{P}(\{x\}) \geq \sum_{x \in \mathcal{X}} \underline{P}(\mathcal{X} \setminus \{x\}) + \underline{P}(\emptyset)$. This is equivalent to $\sum_{x \in \mathcal{X}} \big(\underline{P}(\{x\}) + \overline{P}(\{x\})\big) - 2 \geq 0$ and thus $m(\mathcal{X}) \geq 0$.

5 Belief functions & Dempster's rule of combination

30. Paul wakes up in a room with shuttered windows. Is it day or night? He sees a wall clock indicating 3 o'clock (up to 5 minutes). Paul assesses that his probability that this wall clock works correctly (elementary event a) is 0.8. He knows that this type of wall clock can run slow up to half an hour (elementary event b, with probability 0.05) or run fast up to a quarter hour (elementary event c, with probability 0.1). Also, it is also possible that the wall clock is defective (elementary event d, with probability 0.05).

(a) The elementary events a, b, c, and d can be identified with subsets A, B, C, and D of the interval $(0, 24]$ that are the focal elements of a basic belief assignment m_0 for the current time for Paul. Give these focal elements and the value m_0 attained in them.

If the wall clock is defective, it can be any time of day, so $D = (0, 24]$ and $m_0(D) = 0.05$. For the other events, it is possible that it is both night or day, so each of the corresponding sets will the union of two intervals. If the wall clock is running correctly, we have a five minute accuracy, i.e., a ten minute interval around three, so $A = [2{:}55, 3{:}05] \cup [14{:}55, 15{:}05]$ and $m_0(A) = 0.8$. When the wall clock is running slow, it can run slow any number of minutes from 0 to 30, and the accuracy is still five minutes, so $B = [2{:}55, 3{:}35] \cup [14{:}55, 15{:}35]$ and $m_0(B) = 0.05$. Similarly when the wall clock is running fast: $C = [2{:}40, 3{:}05] \cup [14{:}40, 15{:}05]$ and $m_0(D) = 0.1$.

(b) Paul manages to open the windows a little bit and sees it is day, but not the position of the sun; he knows the amount of light he perceives can occur between 8:00 and 19:00. The elementary events a, b, c, and d can now be identified with subsets A', B', C', and D' of the interval $(0, 24]$ that are the focal elements of a basic belief assignment m_1 for the current time for Paul. Give these focal elements and the value m_1 attains in them.

Now $D = [8, 19]$ and $m_1(D') = 0.05$, $A' = [14{:}55, 15{:}05]$ and $m_1(A') = 0.8$, $B = [14{:}55, 15{:}35]$ and $m_1(B') = 0.05$, $C = [14{:}40, 15{:}05]$ and $m_1(D') = 0.1$.

(c) Calculate the value of the belief function (inner set function) $\underline{P}_{m_i}$ and plausibility function (outer set function) $\overline{P}_{m_i}$ in $[14{:}50, 15{:}15]$, $[14{:}45, 15{:}45]$, and $[14{:}00, 15{:}00]$ for $i \in \{0, 1\}$.

- $[14{:}55, 15{:}35]$ contains A' and intersects all focal elements. So

$$[\underline{P}_{m_0}([14{:}50, 15{:}15]), \overline{P}_{m_0}([14{:}50, 15{:}15])] = [0, 1],$$
$$[\underline{P}_{m_1}([14{:}50, 15{:}15]), \overline{P}_{m_1}([14{:}50, 15{:}15])] = [0.8, 1].$$

- $[14{:}45, 15{:}45]$ contains A' and B' and intersects all focal elements. So

$$[\underline{P}_{m_0}([14{:}45, 15{:}45]), \overline{P}_{m_0}([14{:}45, 15{:}45])] = [0, 1],$$
$$[\underline{P}_{m_1}([14{:}45, 15{:}45]), \overline{P}_{m_1}([14{:}45, 15{:}45])] = [0.85, 1].$$

- $[14{:}00, 15{:}00]$ contains no focal element and intersects all focal elements. So

$$[\underline{P}_{m_0}([14{:}00, 15{:}00]), \overline{P}_{m_0}([14{:}00, 15{:}00])] = [0, 1],$$
$$[\underline{P}_{m_1}([14{:}00, 15{:}00]), \overline{P}_{m_1}([14{:}00, 15{:}00])] = [0, 1].$$

(d) Paul finds his watch at the bottom of his pocket. It indicates a quarter past 3 o'clock with an accuracy of 1 minute. His assessed probability for it functioning correctly (elementary event e) is 0.9. Given that he knows it is day, and only basing himself on his watch, what are the focal elements of the basic belief assignment m_2 for the current time for Paul.

To e there corresponds $E = [15{:}1415{:}16]$ with $m_2(E) = 0.9$. The rest of the probability mass is assigned to D': $m_2(D') = 0.1$.

(e) Assume the information provided by the wall clock is independent of the information provided by the watch. Use Dempster's rule of combination to combine m_1 and m_2 into a basic belief assignment μ and give its value in its focal elements (use a table to write down all the focal element intersections and basic belief mass products necessary). Are the sources conflicting: What is the degree of conflict.

Dempster's rule says that $\mu(\emptyset) = 0$ and that for any nonempty intersection of focal elements F we have

$$\mu(F) = \frac{1}{1-K} \sum_{G \in \{A', B', C', D'\}, H \in \{E, D'\}, G \cap H = F} m_1(G) m_2(H),$$

where

$$K = \sum_{G \in \{A', B', C', D'\}, H \in \{E, D'\}, G \cap H = \emptyset} m_1(G) m_2(H)$$

is the degree of conflict. The intersection/product table is given below:

m_1 \ m_2	E 0.9	D' 0.1
A' 0.8	$\emptyset$ 0.72	A' 0.08
B' 0.05	E 0.045	B' 0.005
C' 0.1	$\emptyset$ 0.09	C' 0.01
D' 0.05	E 0.045	D' 0.005

From this table, we can see that $K = 0.81$, so there is a large conflict, and (up to rounding)

$$m(A') = 0.421, \quad m(B') = m(D') = 0.026, \quad m(C') = 0.053, \quad m(E) = 0.474.$$

31. A crime has been committed and Sherlock Holmes becomes convinced that the guilty person (supposedly acting alone) is one of the following three individuals: Peter, Paul, or Mary. He has also learned that the one who ordered the murder tossed a fair coin to decide whether the murder should be committed by a man or a woman. Multiple days of investigation pass and Sherlock Holmes finds out that Peter is not the murderer, because he has an alibi: a policeman was talking with him at the time of the crime.

We are analysing this problem in two ways, with the goal of representing the evolution of the beliefs of Sherlock Holmes.

The Bayesian approach. (i) Use the Principle of Indifference (equiprobability of all possibilities in case of ignorance) to give the prior probability of each suspect beng guilty—so before learning about Peter's alibi.

The possibility space is $\{\text{Peter}, \text{Paul}, \text{Mary}\}$. From the coin-tossing information, we know that $P(\text{man}) = P(\{\text{Peter}, \text{Paul}\}) = P(\{\text{Mary}\}) = 1/2$. So, by the Principle of Indifference, we furthermore infer that $P(\text{Peter}) = P(\text{Paul}) = 1/4$.

(ii) By using Bayes' conditioning rule, calculate the posterior probabilities of the remaining suspects being guilty—so after learning about Peter's alibi.

$$P(\text{Mary}|\text{not Peter}) = \frac{P(\text{Mary})}{P(\text{Mary}) + P(\text{Paul})} = \frac{1/2}{1/2 + 1/4} = \frac{2}{3},$$
$$P(\text{Paul}|\text{not Peter}) = \frac{P(\text{Paul})}{P(\text{Mary}) + P(\text{Paul})} = \frac{1/4}{1/2 + 1/4} = \frac{1}{3}.$$

The Dempster–Shafer approach. (i) Give the prior basic belief assignment on the subsets of the set of suspects describing the presumed culpability of the three suspects, without using the Principle of Indifference.

From the coin-tossing information, we know that $m(\{\text{Peter}, \text{Paul}\}) = m(\{\text{Mary}\}) = 1/2$ (and zero elsewhere).

(ii) Give the basic belief assignment resulting from applying Dempster's conditioning rule.

Because of his alibi, Peter is removed from the focal element $\{\text{Peter}, \text{Paul}\}$, so now $m'(\{\text{Paul}\}) = m'(\{\text{Mary}\}) = 1/2$ (and zero elsewhere), which corresponds to $P(\text{Mary}|\text{not Peter}) = P(\text{Paul}|\text{not Peter}) = 1/2$.

Comparison (i) Do both approaches give the same result? If not, what is the (interpretational) difference between both approaches.

The approaches differ: the Bayesian approach gives more weight to Mary's guilt, whereas the Dempster–Shafer approach is more neutral in that it puts a uniform distribution over $\{\text{Paul}, \text{Mary}\}$. The alibi information is more than just discovering the (conditioning) event, it is incompatible with the assessment resulting from the initial application of the Principle of Indifference, in fact, it is incompatible with any probability mass function assigning positive mass to Peter.

(ii) Treat the problem with both approaches again, but now starting from a situation in which there are five suspects left: a woman, Debbie, and a man, John, in addition to the three of the original problem. Now Debbie, John, and Peter turn out to have an alibi.

The possibility space is $\{\text{Peter}, \text{Paul}, \text{John}, \text{Mary}, \text{Debbie}\}$. From the coin-tossing information, we know that $P(\text{man}) = P(\{\text{Peter}, \text{Paul}, \text{John}\}) = P(\text{woman}) = P(\{\text{Mary}, \text{Debbie}\}) = 1/2$ and also $m(\{\text{Peter}, \text{Paul}, \text{John}\}) = m(\{\text{Mary}, \text{Debbie}\}) = 1/2$ (and zero elsewhere). using the Principle of Indifference, we furthermore infer that $P(\text{Peter}) = P(\text{Paul}) = P(\text{John}) = 1/6$ and $P(\text{Mary}) = P(\text{Debbie}) = 1/4$.

Because of the alibis, using the Bayesian approach, we condition on $\{\text{Paul}, \text{Mary}\}$ and get $P(\text{Mary}|\text{alibis}) = 3/5$ and $P(\text{Paul}|\text{alibis}) = 2/5$. In the Dempster-Shafer approach, there is no change as compared to the original problem, the modification of the focal elements leads to $m'(\{\text{Paul}\}) = m'(\{\text{Mary}\}) = 1/2$ (and zero elsewhere), which corresponds to $P(\text{Mary}|\text{not Peter}) = P(\text{Paul}|\text{not Peter}) = 1/2$.

(iii) Compare the resulting prior and posterior probabilities and belief assignments with the ones of the unmodified problem. What do you conclude about the relative applicability of the Bayesian and Dempster–Shafer approaches to this kind of problems.

The Bayesian posterior result depends on the prior number and configuration of suspects, even though the posterior suspects are the same. The Dempster–Shafer approach is unaffected and is better suited for dealing with the type of information that is learned because of the alibis.

(iv) What happens if the Bayesian approach is applied without the Principle of Indifference, i.e., by applying the GBR instead of Dempster's rule to the basic belief assignment model seen as a lower prevision specification.

More imprecision…

6 Extra exercises on non-additive measures

32. Let Ω be an infinite set, and define μ on $\mathcal{P}(\Omega)$ by

$$\mu(A) = \begin{cases} 1 & \text{if } A^c \text{ finite;} \\ 0 & \text{otherwise.} \end{cases}$$

Show that μ is 2-monotone.

Consider two sets A, B. There are a number of possibilities:

- $\mu(A) = 0 = \mu(B)$, whence trivially $\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B)$.
- $\mu(A) = 1, \mu(B) = 0$; this means that A^c is finite, whence $A^c \cap B^c = (A \cup B)^c$ is finite and therefore $\mu(A \cup B) = 1$. Hence, $\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B)$.
- $\mu(A) = 0, \mu(B) = 1$; this case is analogous to the previous one.
- $\mu(A) = 1 = \mu(B)$; this means that A^c, B^c are finite, whence so are $A^c \cap B^c = (A \cup B)^c$ and $A^c \cup B^c = (A \cap B)^c$. Hence, $\mu(A \cup B) = 1 = \mu(A \cap B)$ and therefore $\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B)$.

33. Show that any coherent lower probability μ on $\mathcal{P}(\Omega)$ is 2-monotone when $|\Omega| = 2$.

We have showed elsewhere that μ is a linear-vacuous mixture. Since both probabilities and vacuous previsions are 2-monotone and 2-monotonicity is preserved under convex combinations, we deduce that μ is necessarily 2-monotone.

34. Let μ a 2-monotone capacity defined on a field of sets $\mathcal{A}$, and let us extend it to $\mathcal{P}(\Omega)$ by

$$\mu_*(A) = \sup\{\mu(B) : B \subseteq A\}.$$

Show that μ_* is also 2-monotone.

Consider two events A, B, and $\epsilon > 0$. Then there are $C, D \in \mathcal{A}$ such that $C \subseteq A, \mu^*(A) \leq \mu(C) + \epsilon$ and $D \subseteq B, \mu^*(D) \leq \mu(B) + \epsilon$. As a consequence, $C \cap D \in \mathcal{A}, C \cap D \subseteq A \cap B$ and similarly $C \cup D \in \mathcal{A}, C \cup D \subseteq A \cup B$. Hence,

$$\mu^*(A \cup B) + \mu^*(A \cap B) \geq \mu(C \cup D) + \mu(C \cap D) \geq \mu(C) + \mu(D) \geq \mu^*(A) + \mu^*(B) - 2\epsilon,$$

and since we can do this for every $\epsilon > 0$ we deduce that $\mu^*(A \cup B) + \mu^*(A \cap B) \geq \mu^*(A) + \mu^*(B)$, meaning that μ^* is 2-monotone.

35. Consider $\Omega = \{1, 2, 3, 4\}$, and let μ be the lower envelope of the probabilities P_1, P_2 given by

$$P_1(1) = P_1(2) = 0.5, P_1(3) = P_1(4) = 0$$
$$P_2(1) = P_2(2) = P_2(3) = P_2(4) = 0.25.$$

Show that μ is not 2-monotone.

Take $A = \{1, 3\}, B = \{1, 4\}$. Then

$$\mu(A \cup B) + \mu(A \cap B) = 0.5 + 0.25 = 0.75 < \mu(A) + \mu(B) = 0.5 + 0.5 = 1$$

36. Consider $\Omega = \{1, 2, 3\}$.

(a) Let m be the basic probability assignment given by $m(\{1, 2\}) = 0.5, m(\{3\}) = 0.2, m(\{2, 3\}) = 0.3$. Determine the belief function associated to m.

Using the relationship between belief functions and their lower inverses, we obtain

$$\mu(\{1\}) = 0, \quad \mu(\{2\}) = 0, \quad \mu(\{3\}) = 0.2,$$
$$\mu(\{1,2\}) = 0.5, \quad \mu(\{1,3\}) = 0, \quad \mu(\{2,3\}) = 0.5$$

and trivially $\mu(\emptyset) = 0, \mu(\{1, 2, 3\}) = 1$.

(b) Consider the belief function $\underline{P}$ given by $\underline{P}(A) = \frac{|A|}{3}$ for every $A \subseteq \Omega$. Determine its basic probability assignment.

Using the formula, we obtain

$$m(\{1\}) = m(\{2\}) = m(\{3\}) = \frac{1}{3},$$
$$m(\{1, 2\}) = m(\{1, 3\}) = m(\{2, 3\}) = 0$$
$$m(\{1, 2, 3\}) = 0,$$

so it corresponds to the uniform probability measure.

37. Consider $\Omega = \{1, 2, 3\}$, and let $\underline{P}$ be the belief function on $\mathcal{P}(\Omega)$ associated to the basic probability assignment

$$m(\{1, 2\}) = 0.1, m(\{1, 3\}) = 0.2, m(\{2, 3\}) = 0.3, m(\{1, 2, 3\}) = 0.4.$$

Let f be the gamble on Ω given by $f(1) = 4, f(2) = 0, f(3) = 2$. Calculate $(C)\int f d\underline{P}$ and $(C)\int (f + \mathbb{I}_1 - \mathbb{I}_2) d\underline{P}$, where $\mathbb{I}_A$ denotes the indicator function of A.

Using the formula,

$$(C)\int f d\mu = 4\mu(\{1\}) + 2(\mu\{1, 3\} - \mu(\{1\}) + 0(\{\mu(\{1, 2, 3\} - \mu(\{1, 3\})$$
$$= 4 \cdot 0 + 2 \cdot 0.2 + 0 = 0.4.$$

On the other hand, the gamble $g = f + \mathbb{I}_1 - \mathbb{I}_2$ satisfies $g(1) = 5, g(2) = -1, g(3) = 2$, so applying again the formula,

$$(C)\int g d\mu = 5\mu(\{1\}) + 2(\mu\{1, 3\} - \mu(\{1\}) - 1(\{\mu(\{1, 2, 3\} - \mu(\{1, 3\})$$
$$= 5 \cdot 0 + 2 \cdot 0.2 - 1(1 - 0.2) = -0.4.$$

This second part can also be solved using that the gambles f and $\mathbb{I}_1 - \mathbb{I}_2$ are comonotone, so

$$(C)\int g d\mu = (C)\int f d\mu + (C)\int I_{\{1\}} - I_{\{2\}} d\mu = 0.4 - 0.8 = -0.4$$

because

$$(C)\int I_{\{1\}} - I_{\{2\}} d\mu = 1\mu(\{1\}) + 0(\mu(\{1,3\}) - \mu(\{1\}) - 1(\mu(\{1,2,3\} - \mu(\{1,3\}) = -0.8$$

38. Consider $\mu : \beta_{[0,1]} \to [0,1]$ given by $\mu(A) = 1$ if A is uncountable, and $\mu(A) = 0$ otherwise.

(a) Show that μ is maxitive.

It suffices to take into account that the union of two sets is uncountable if and only if at least one of them is uncountable.

(b) Show that μ is lower continuous.

Let $(A_n)_n$ be an increasing sequence of sets. Then if all of them are countable, their union is also countable, and $\mu(\cup_n A_n) = \lim_n \mu(A_n) = 0$; and if one of them is uncountable then so is their union, and $\mu(\cup_n A_n) = \lim_n \mu(A_n) = 1$.

(c) Show that μ is not a possibility measure.

It suffices to consider that $\mu[0,1] = 1 > \sup_{x\in[0,1]} \mu(\{x\})$.

39. Show that for a non-additive measure μ, its maxitivity neither implies or is implied by the property $\mu(A) = \sup_{K\subseteq A, K compact} \mu(K)$ for every $A \in \beta_X$. Use the following:

(a) $\mu : \beta_{[0,1]} \to [0,1]$ equal to the Lebesgue measure.

The Lebesgue measure satisfies that $\mu(A) = \sup_{K\subseteq A, K compact} \mu(K)$ for every $A \in \beta_X$, but it is not maxitive: given $A = (0, 0.5)$ and $B = (0.5, 1)$, $\mu(A \cup B) = 1 > \max\{\mu(A), \mu(B)\}$.

(b) $\mu : \mathcal{P}(\mathbb{N}) \to \{0,1\}$ given by $\mu(A) = 1$ if and only if $|A| = \infty$.

This function is maxitive because the union of two sets is infinite if and only if at least one of them is infinite. To see that it does not satisfy the other property, note that in $\mathcal{P}(\mathbb{N})$ a set is compact if and only if it is finite, and as a consequence $\mu(\mathbb{N}) = 1 > \sup_{A\subseteq\mathbb{N} \text{ compact}} \mu(A)$.

40. Consider $\Omega = \{1,2,3,4\}$.

(a) Let Π be the possibility distribution associated to the possibility distribution $\pi(1) = 0.3, \pi(2) = 0.5, \pi(3) = 1, \pi(4) = 0.7$. Determine its focal elements and its basic probability assignment.

We can do this using the formula that determines the Möbius inverse of an upper probability; however, if we take into account that the focal elements must be nested and that $\Pi(\{1\}) = 0.3, \Pi(\{1,2\}) = 0.5, \Pi(\{1,2,4\}) = 0.7$ and $\Pi(\{1,2,3,4\}) = 1$, we can deduce that it must be $m(\{3\}) = 0.3, m(\{3,4\}) = 0.2, m(\{2,3,4\}) = 0.2$ and $m(\{1,2,3,4\}) = 0.3$.

(b) Given the basic probability assignment $m(\{1\}) = 0.2, m(\{1,3\}) = 0.1, m(\{1,2,3\}) = 0.4, m(\{1,2,3,4\}) = 0.3$, determine the associated possibility measure and its possibility distribution.

It suffices to determine the possibility distribution. Using the formula that relates upper probabilities with their Möbius inverse, we obtain

$$\pi(\{1\}) = \sum_{1\in B} m(B) = 1 \qquad \pi(\{2\}) = \sum_{2\in B} m(B) = 0.7$$

$$\pi(\{3\}) = \sum_{3\in B} m(B) = 0.8 \qquad \pi(\{4\}) = \sum_{4\in B} m(B) = 0.3$$

41. Consider a finite space $\mathcal{X}$.

(a) What are the only possibility measures which are at the same time a probability measure?

Since it must be $1 = \Pi(\mathcal{X}) = \max_{x\in\mathcal{X}} \Pi(\{x\}) = \sum_{x\in\mathcal{X}} \Pi(\{x\})$, the only possibility is that there is some x with $\Pi(\{x\}) = 1$ and $\Pi(\{x'\}) = 0\ \forall x' \neq x$. Hence, the only possibility measures which are at the same time a probability measure are the degenerate distributions.

(b) What is their basic probability assignment?

It would be $m(\{x\}) = 1$ for the associated x.

(c) And their associated Choquet integral?

$(C)\int f d\Pi = f(x)\ \forall f$.

42. **Belief functions and random sets.** Suppose we have a probability measure μ defined on the power set $\wp(\mathcal{Y})$ of a finite set $\mathcal{Y}$, and a multi-valued mapping Γ from $\mathcal{Y}$ into a finite set $\mathcal{X}$. As Dempster (1967) puts it: "if the uncertain outcome y is known to correspond to an uncertain outcome $x \in \Gamma(y)$, what probability judgements may be made about the uncertain outcome $x \in \mathcal{X}$?" For the sake of simplicity, we shall assume that $\Gamma(y) \neq \emptyset$ for all $y \in \mathcal{Y}$.

Consider the set function ν on $\wp(\mathcal{X})$ defined as

$$\nu(A) := \mu(\{y \in \mathcal{Y} \colon \Gamma(y) \subseteq A\})$$

for every $A \subseteq \mathcal{X}$.

We may interpret the values $\nu(A)$ as supremum buying prices for indicator gambles I_A. This corresponds to the lower prevision $\underline{P}$ defined on $\{I_A\colon A \subseteq \mathcal{Y}\}$ by $\underline{P}(I_A) := \nu(A)$ for all $y \in \mathcal{Y}$.

(a) Preparatory exercise. Show that

$$\underline{P}(I_A) = \sum_{y\in\mathcal{Y}} \mu(\{y\})\underline{P}_{\Gamma(y)}(I_A).$$

Use the additivity of μ, and observe that $\Gamma(y) \subseteq A$ if and only if $\underline{P}_{\Gamma(y)}(I_A) = 1$, and $\Gamma(y) \nsubseteq A$ if and only if $\underline{P}_{\Gamma(y)}(I_A) = 0$.

(b) Prove that ν is a 2-monotone set function, as defined in Section 3. [Hint: first show that for all $y \in \mathcal{Y}$, $\underline{P}_{\Gamma(y)}$ is 2-monotone as a set function restricted to events, and then use ((a)).]

Since the convex combination of a 2-monotone set function is again 2-monotone, by ((a)) it suffices to show that each vacuous lower prevision $\underline{P}_{\Gamma(y)}$ is 2-monotone as a set function restricted to events.
Clearly, $\underline{P}_{\Gamma(y)}(\emptyset) = 0$, $\underline{P}_{\Gamma(y)}(\mathcal{X}) = 1$, and $\underline{P}_{\Gamma(y)}(A) \geq 0$ for all $A \subseteq \mathcal{X}$. We are left to show that $\underline{P}_{\Gamma(y)}(A\cup B) + \underline{P}_{\Gamma(y)}(A\cap B) \geq \underline{P}_{\Gamma(y)}(A) + \underline{P}_{\Gamma(y)}(B)$ for any A, $B \subseteq \mathcal{X}$.
Indeed,

$$\underline{P}_{\Gamma(y)}(A\cup B) + \underline{P}_{\Gamma(y)}(A\cap B) = \begin{cases} 2, & \text{if } \Gamma(y) \subseteq A\cap B,\\ 1, & \text{if } \Gamma(y) \subseteq A\cup B \text{ but } \Gamma(y) \nsubseteq A\cap B,\\ 0, & \text{if } \Gamma(y) \nsubseteq A\cup B.\end{cases}$$

If $\Gamma(y) \subseteq A\cap B$, then also $\Gamma(y) \subseteq A$ and $\Gamma(y) \subseteq B$, and hence, $\underline{P}_{\Gamma(y)}(A) + \underline{P}_{\Gamma(y)}(B) = 2$. Therefore, in this case indeed

$$\underline{P}_{\Gamma(y)}(A\cup B) + \underline{P}_{\Gamma(y)}(A\cap B) \geq \underline{P}_{\Gamma(y)}(A) + \underline{P}_{\Gamma(y)}(B). \tag{1}$$

Next, note that if $\Gamma(y) \nsubseteq A\cap B$ then it cannot hold that $\Gamma(y) \subseteq A$ and $\Gamma(y) \subseteq B$. Hence, in case $\Gamma(y) \subseteq A\cup B$ but $\Gamma(y) \nsubseteq A\cap B$, we find that $\underline{P}_{\Gamma(y)}(A)+\underline{P}_{\Gamma(y)}(B) \leq 1$. In this case, again the condition for 2-monotonicity, Eq. (1), holds.
Finally, if $\Gamma(y) \nsubseteq A\cup B$, then $\Gamma(y) \nsubseteq A$ and $\Gamma(y) \nsubseteq B$, and hence, $\underline{P}_{\Gamma(y)}(A) + \underline{P}_{\Gamma(y)}(B) = 0$. We conclude that Eq. (1) holds in all cases.

(c) Show that $\underline{P}$ is coherent.

The convex combination of coherent lower previsions is again coherent. Alternatively, 2-monotonicity implies coherence.

(d) Show that the natural extension of $\underline{P}$ is given by

$$\underline{E}(f) = \sum_{y\in\mathcal{Y}} \mu(\{y\})\underline{P}_{\Gamma(y)}(f).$$

Briefly, by 2-monotonicity, the natural extension of the convex combination is the convex combination of the natural extensions.

(e) Prove that $\underline{E}$ is the $\mathcal{X}$-marginal of the marginal extension of P_μ and $\underline{P}_\Gamma(\cdot|\mathcal{Y})$, where

$$P_\mu(f) := \int f \mathrm{d}\mu = \sum_{y\in\mathcal{Y}} \mu(\{y\})f(y)$$

for all $f \in \mathcal{L}(\mathcal{Y})$, and

$$\underline{P}_\Gamma(f|y) := \underline{P}_{\Gamma(y)}(f)$$

for all $f \in \mathcal{L}(\mathcal{X})$ and $y \in \mathcal{Y}$. Hence, ν is indeed the (least committal) lower probability following from the premises.

Immediate, from the definition of marginal extension: for any $f \in \mathcal{L}(\mathcal{X})$, it holds that

$$P_\mu(\underline{P}_\Gamma(f|\mathcal{Y})) = \sum_{y\in\mathcal{Y}} \mu(\{y\})\underline{P}_\Gamma(f|y) = \sum_{y\in\mathcal{Y}} \mu(\{y\})\underline{P}_{\Gamma(y)}(f).$$

(f) Show that, for all gambles f on $\mathcal{X}$,

$$\mathrm{C}\int f \mathrm{d}\nu = \sum_{y\in\mathcal{Y}} \mu(\{y\})\underline{P}_{\Gamma(y)}(f).$$

] By ((b)), it follows from Section 3 that the natural extension of $\underline{P}$ is the Choquet integral with respect to ν. Now use ((d)).

43. **Possibility and necessity measures.** Suppose we have a minitive set function ν defined on the power set $\wp(\mathcal{X})$ of a finite set $\mathcal{X}$, that is,

(i) $\nu(\emptyset) = 0$, $\nu(\mathcal{X}) = 1$,
(ii) $\nu(A) \geq 0$ for all $A \subseteq \mathcal{X}$, and
(iii) $\nu(A\cap B) = \min\{\nu(A), \nu(B)\}$ for any A, $B \subseteq \mathcal{X}$.

We may interpret the values $\nu(A)$ as supremum buying prices for indicator gambles I_A. This corresponds to the lower prevision $\underline{P}$ defined on $\{I_A\colon A \subseteq \mathcal{X}\}$ by $\underline{P}(I_A) := \nu(A)$ for all $x \in \mathcal{X}$.

(a) Show that ν is a necessity measure.

Because $\mathcal{X}$ is finite, minitivity is sufficient.

(b) Show that, for every $A \subseteq \mathcal{X}$, $A \neq \mathcal{X}$,

$$\nu(A) = \min_{x\in A^c} \nu(\{x\}^c)$$

Use

$$A = \bigcap_{x\in A^c} \{x\}^c$$

and minitivity.

(c) Show that ν is 2-monotone.

Since ν is minitive, it follows that $\min\{\nu(A \cup B), \nu(A)\} = \nu(A)$, or in other words, $\nu(A \cup B) \geq \nu(A)$. Similarly, $\nu(A \cup B) \geq \nu(B)$. We conclude that $\nu(A \cup B) \geq \max\{\nu(A), \nu(B)\}$. Hence,

$$\begin{aligned}\nu(A \cup B) + \nu(A \cap B) &\geq \max\{\nu(A), \nu(B)\} + \min\{\nu(A), \nu(B)\} \\ &= \nu(A) + \nu(B),\end{aligned}$$

which means that ν is 2-monotone.

(d) Show that ν is a belief function. Start with defining $n(x) := \nu(\{x\}^c)$. Let $y_1, \ldots, y_m$ be an enumeration of the values of n with $y_1 < y_2 < \cdots < y_m$ (note that $y_1 = 0$ because $\nu(\emptyset) = 0$). Now let $\mathcal{Y} = \{y_1, \ldots, y_m\}$ and define the multi-valued mapping

$$\Gamma(y_i) := A_i \text{ where } A_i := \{x \in \mathcal{X} : n(x) \leq y_i\}.$$

Find a probability measure μ on $\mathcal{Y}$ such that

$$\nu(A) = \min_{x \in A^c} n(x) = \sum_{i=1}^{n} \mu(\{y_i\}) \underline{P}_{\Gamma(y_i)}(I_A). \tag{2}$$

Indeed, if $i < m$ then

$$\begin{aligned}\nu(A_i) &= \min_{x \in A_i^c} n(x) \\ &= \min\{n(x) : x \in \mathcal{X}, x > y_i\} \\ &= y_{i+1},\end{aligned}$$

and $\nu(A_m) = 1$ because $A_m = \mathcal{X}$. Hence, for Eq. (2) to hold, we need that

$$\begin{aligned}&y_2 = \nu(A_1) = \mu(\{y_1\}) = y_1 + \mu(\{y_1\}) \\ &y_3 = \nu(A_2) = \mu(\{y_1\}) + \mu(\{y_2\}) = y_2 + \mu(\{y_2\}) \\ &y_4 = \nu(A_3) = \mu(\{y_1\}) + \mu(\{y_2\}) + \mu(\{y_3\}) = y_3 + \mu(\{y_3\}) \\ &\quad\vdots \\ &1 = \nu(A_m) = y_m + \mu(\{y_m\})\end{aligned}$$

We conclude that

$$\mu(\{y_i\}) = y_{i+1} - y_i$$

for all $i \in \{1, \ldots, m\}$, if we define $y_{m+1} = 1$.
We are left to show that Eq. (2) holds for all events $A \subseteq \mathcal{X}$:

$$\begin{aligned}\nu(A) &= \min_{x \in A^c} n(x) \\ &= \min\{y_i : A_i \cap A^c \neq \emptyset\} \\ &= \min\{y_i : A_i \not\subseteq A\}\end{aligned}$$

and with i^* the smallest index i such that $A_i \not\subseteq A$ (take $i^* = m+1$ if $A = \mathcal{X}$), we have

$$\begin{aligned}&= y_{i^*} \\ &= \sum_{i=1}^{i^*-1} \mu(\{y_i\})\end{aligned}$$

and since $\underline{P}_{A_i}(I_A) = 1$ for all $i < i^*$, and $\underline{P}_{A_i}(I_A) = 0$ for all $i \geq i^*$,

$$\begin{aligned}&= \sum_{i=1}^{n} \mu(\{y_i\}) \underline{P}_{A_i}(I_A) \\ &= \sum_{i=1}^{n} \mu(\{y_i\}) \underline{P}_{\Gamma(y_i)}(I_A)\end{aligned}$$

44. **P-boxes.** Let $\mathcal{X} = \mathbb{R}$. Let $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2$. Consider the linear previsions $\underline{P}_{x_1}$ and $\underline{P}_{x_2}$ defined by

$$\begin{aligned}\underline{P}_{x_1}(f) &:= f(x_1), \\ \underline{P}_{x_2}(f) &:= f(x_2),\end{aligned}$$

for all $f \in \mathcal{L}(\mathcal{X})$. Note that these linear previsions are vacuous lower previsions relative to singletons. The lower envelope $\underline{P}$ of $\underline{P}_{x_1}$ and $\underline{P}_{x_2}$ is nothing but the vacuous lower prevision relative to the pair $\{x_1, x_2\}$:

$$\underline{P}(f) = \min\{f(x_1), f(x_2)\}.$$

Note that $\underline{P}$ is coherent.

(a) Draw the p-box that corresponds to $\underline{P}$.

The cumulative distribution functions for $\underline{P}_{x_1}$ and $\underline{P}_{x_2}$ are step functions:

$$F_{x_1}(x) = \underline{P}_{x_1}(\{y \in \mathcal{X} : y \leq x\}) = \begin{cases} 0, & \text{if } x < x_1, \\ 1, & \text{if } x \geq x_1, \end{cases}$$

and similar for the cumulative distribution function F_{x_2} of $\underline{P}_{x_2}$. The p-box corresponding to $\underline{P}$ is now simply the "rectangle" between F_{x_1} and F_{x_2}.

(b) Prove that the "natural extension" of this p-box, that is, the lower envelope $\underline{E}$ of all linear previsions Q on $\mathcal{X}$ whose cumulative distribution function

$$F_Q(x) = Q(\{y \in \mathcal{X} : y \leq x\})$$

belongs to this p-box, is dominated by the vacuous lower prevision relative to the *interval* $[x_1, x_2]$, that is,

$$\underline{E}(f) \leq \underline{P}_{[x_1,x_2]}(f) \text{ for any gamble } f \in \mathcal{L}(\mathcal{X}).$$

What does this mean?

Let $\mathcal{M}$ denote the set of linear previsions whose cumulative distribution function belongs to the p-box corresponding to $\underline{P}$. A linear prevision Q belongs to $\mathcal{M}$ if and only if its cumulative distribution function lies between F_{x_1} and F_{x_2}:

$$F_{x_1}(x) \geq Q(\{y \in \mathcal{X} : y \leq x\}) \geq F_{x_2}(x)$$

for all $x \in \mathcal{X}$. Since F_{x_1} and F_{x_2} are simple step functions, these conditions reduce to

$$\begin{cases} Q(\{y \in \mathcal{X} : y \leq x\}) = 0, & \text{if } x < x_1, \\ Q(\{y \in \mathcal{X} : y \leq x\}) = 1, & \text{if } x \geq x_2. \end{cases} \tag{3}$$

Observe that Eq. (3) is satisfied for $Q = \underline{P}_x$, the vacuous lower prevision relative to the singleton $\{x\}$, whenever $x \in [x_1, x_2]$. Hence,

$$\{\underline{P}_x : x \in [x_1, x_2]\} \subseteq \mathcal{M}.$$

Therefore,

$$\underline{E}(f) = \inf_{Q \in \mathcal{M}} Q(f) \leq \inf_{x \in [x_1, x_2]} \underline{P}_x(f) = \underline{P}_{[x_1, x_2]}(f),$$

for any gamble $f \in \mathcal{L}(\mathcal{X})$.

(c) Extra exercise. If you are fond of ϵ's, show that

$$\underline{E}(f) = \sup_{\epsilon > 0} \underline{P}_{(x_1 - \epsilon, x_2]}(f) \text{ for any gamble } f \in \mathcal{L}(\mathcal{X}).$$

Let's give just the essential steps.
Define the lower prevision $\underline{P}$ as

$$\underline{P}(-I_{(-\infty, x_1 - \epsilon]}) = 0 \text{ for all } \epsilon > 0,$$
$$\underline{P}(I_{(-\infty, x_2]}) = 1.$$

By (3) it follows that $Q \in \mathcal{M}$ if and only if $Q \in \mathcal{M}(\underline{P})$. Therefore, it suffices to show that $\underline{E}$ is the least committal extension of $\underline{P}$.
Suppose $\underline{Q}$ is another coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$. Show that

$$\underline{Q}(f) \geq \underline{P}_{(x_1 - \epsilon, x_2]}(f), \tag{4}$$

for all $\epsilon > 0$, and hence,

$$\underline{Q}(f) \geq \lim_{\substack{\epsilon \to 0 \\ >}} \underline{P}_{(x_1 - \epsilon, x_2]}(f) = \sup_{\epsilon > 0} \underline{P}_{(x_1 - \epsilon, x_2]}(f) = \underline{E}(f).$$

To prove Eq. (4), use the fact that

$$f \geq \underline{P}_{(x_1 - \epsilon, x_2]}(f) + \left[\underline{P}_{(x_1 - \epsilon, x_2]}(f) - \inf[f]\right] \left(I_{(x_1 - \epsilon, x_2]} - 1\right)$$

Apply $\underline{Q}$ on both sides of the above inequality, and use

$$I_{(x_1 - \epsilon, x_2]} - 1 = -I_{(-\infty, x_1 - \epsilon]} - I_{(x_2, +\infty)}$$
$$\underline{Q}\left(-I_{(-\infty, x_1 - \epsilon]}\right) = 0$$
$$\underline{Q}\left(-I_{(x_2, +\infty)}\right) = \underline{Q}\left(I_{(-\infty, x_2])}\right) - 1 = 0,$$

by which

$$\underline{Q}\left(I_{(x_1 - \epsilon, x_2]} - 1\right) = \underline{Q}\left(-I_{(-\infty, x_1 - \epsilon]} - I_{(x_2, +\infty)}\right) \geq 0,$$

By monotonicity of $\underline{Q}$, we deduce that $\underline{Q}\left(I_{(x_1 - \epsilon, x_2]} - 1\right) \leq \underline{Q}(0) = 0$. Hence,

$$\begin{aligned}\underline{Q}(f) &\geq \underline{P}_{(x_1 - \epsilon, x_2]}(f) + \left[\underline{P}_{(x_1 - \epsilon, x_2]}(f) - \inf[f]\right] \underline{Q}\left(I_{(x_1 - \epsilon, x_2]} - 1\right) \\ &= \underline{P}_{(x_1 - \epsilon, x_2]}(f).\end{aligned}$$

This establishes Eq. (4).

7 Conditional lower previsions

45. (a) Let $\mathcal{B}$ be a partition of $\mathcal{X}$, and let $\{P_\gamma(\cdot|\mathcal{B}) : \gamma \in \Gamma\}$ be a set of conditional linear previsions. Show that their lower envelope $\underline{P}(\cdot|\mathcal{B})$ is separately coherent.

Since the domain of $\underline{P}(\cdot|\mathcal{B})$ is the linear space of all gambles, we need to show the following:

- $\underline{P}(f|B) \geq \inf_{\omega\in B} f(\omega)$: it suffices to take into account that $P_\gamma(f|B) \geq \inf_{\omega\in B} f(\omega)$ for every $\gamma \in \Gamma$.
- $\underline{P}(f+g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$: for every $\gamma \in \Gamma$ it holds that $P_\gamma(f+g|B) = P_\gamma(f|B) + P_\gamma(g|B)$; since the infimum of the sum is greater than or equal to the sum of the infima, we deduce that

$$\begin{aligned}\underline{P}(f+g|B) &= \inf_{\gamma\in\Gamma} P_\gamma(f+g|B)\\ &\geq \inf_{\gamma\in\Gamma} P_\gamma(f|B) + \inf_{\gamma\in\Gamma} P_\gamma(g|B) = \underline{P}(f|B) + \underline{P}(g|B).\end{aligned}$$

- $\underline{P}(\lambda f|B) = \lambda\underline{P}(f|B)$: for every $\gamma \in \Gamma$ it holds that $P_\gamma(\lambda f|B) = \lambda P_\gamma(f|B)$; as a consequence,

$$\begin{aligned}\underline{P}(\lambda f|B) &= \inf_{\gamma\in\Gamma} P_\gamma(\lambda f|B)\\ &= \inf_{\gamma\in\Gamma} \lambda P_\gamma(f|B)\\ &= \lambda \inf_{\gamma\in\Gamma} P_\gamma(f|B) = \lambda\underline{P}(f|B).\end{aligned}$$

(b) Conversely, show that any separately coherent $\underline{P}(\cdot|\mathcal{B})$ on $\mathcal{L}(\mathcal{X})$ is the lower envelope of a family of conditional linear previsions.

If $\underline{P}(\cdot|\mathcal{B})$ is a separately coherent conditional lower prevision, it follows that for every $B \in \mathcal{B}$ $\underline{P}(\cdot|B)$ is a coherent lower prevision and moreover $\underline{P}(B|B) = 1$; as a consequence, for every $B \in \mathcal{B}$ there exists a class $\mathcal{M}_B$ of linear previsions such that $\underline{P}(f|B) = \min_{P_B\in\mathcal{M}_B} P_B(f)$; it follows in particular that $P_B(B) = 1$ for every $P_B \in \mathcal{M}_B$.
Let us define $\mathcal{M} := \{P(\cdot|\mathcal{B}) : P(\cdot|B) \in \mathcal{M}_B \ \forall B\}$. It follows that the previsions thus defined are separately coherent conditional linear previsions that dominate $\underline{P}(\cdot|\mathcal{B})$, and moreover $\underline{P}(\cdot|\mathcal{B})$ is the lower envelope of $\mathcal{M}$.

46. Let $\mathcal{B}$ be a partition of $\mathcal{X}$, and let $\mathcal{K}$ be a coherent set of desirable gambles. Define $\underline{P}(\cdot|\mathcal{B})$ by

$$\underline{P}(f|B) = \sup\{\mu : B(f-\mu) \in \mathcal{K}\}.$$

Show that $\underline{P}(\cdot|\mathcal{B})$ is separately coherent.

Since the domain of $\underline{P}(\cdot|\mathcal{B})$ is the linear space of all gambles, we need to show the following:

- $\underline{P}(f|B) \geq \inf_{\omega\in B} f(\omega)$: given $\mu < \inf_{\omega\in B} f(\omega)$, the gamble $B(f-\mu)$ is non-negative and as a consequence it belongs to $\mathcal{K}$.
- $\underline{P}(f+g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$: fix $\epsilon > 0$. Then $B(f - \underline{P}(f|B) + \frac{\epsilon}{2})$ and $B(g - \underline{P}(g|B) + \frac{\epsilon}{2})$ belong to $\mathcal{K}$, taking into account that a coherent set of desirable gambles is always closed under dominance. Applying (D3), $B(f+g-\underline{P}(f|B)-\underline{P}(g|B)+\epsilon)$ belongs to $\mathcal{K}$, and therefore $\underline{P}(f+g|B) \geq \underline{P}(f|B) + \underline{P}(g|B) - \epsilon$. Since we can do this for every $\epsilon > 0$, we conclude that $\underline{P}(f+g|B) \geq \underline{P}(f|B) + \underline{P}(g|B)$.
- $\underline{P}(\lambda f|B) = \lambda\underline{P}(f|B)$: this follows taking into account that, from axiom (D4) of the coherence of gambles, $B(f-\mu)$ belongs to $\mathcal{K}$ if and only if $B(\lambda f - \lambda\mu)$ belongs to $\mathcal{K}$.

47. Let $\underline{P}(\cdot|\mathcal{B})$ be the vacuous conditional lower prevision, given by $\underline{P}(f|B) = \inf_{x\in B} f(x)$ for every $B \in \mathcal{B}$, $f \in \mathcal{K}(\mathcal{X})$.

(a) Show that $\underline{P}(\cdot|\mathcal{B})$ is coherent with the vacuous lower prevision $\underline{P}$.

The vacuous conditional lower prevision $\underline{P}(\cdot|\mathcal{B})$ satisfies $\inf_{x\in B} G(f|B)(\omega) = 0$ for every $B \in \mathcal{B}$. As a consequence,

$$\underline{P}(G(f|B)) = \inf_{x\in\mathcal{X}} G(f|B)(x) = \min\{0, \inf_{x\in B} G(f|B)(x)\} = 0,$$

from which we conclude that $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are coherent.

(b) Show that any coherent lower prevision $\underline{P}$ on $\mathcal{K}(\mathcal{X})$ which satisfies $\underline{P}(B) = 0$ for all $B \in \mathcal{B}$ is coherent with $\underline{P}(\cdot|\mathcal{B})$.

For any gamble f, if $\underline{P}(B) = 0$ then

$$\begin{aligned}\underline{P}(G(f|B)) = \underline{P}(B(f - \underline{P}(f|B)) &\leq \underline{P}(B(\sup_{x\in B} f(x) - \underline{P}(f|B)))\\ &\leq (\sup_{x\in B} f(x) - \underline{P}(f|B)))\underline{P}(B) = 0;\end{aligned}$$

on the other hand, since $G(f|B) \geq 0$ for all B, it follows from the coherence of $\underline{P}$ that $\underline{P}(G(f|B)) \geq 0$. Hence, $\underline{P}(G(f|B)) = 0$ and we deduce that $\underline{P}, \underline{P}(\cdot|\mathcal{B})$ are coherent.

48. Three horses (a,b and c) take part in a race. Our a priori lower probability for each horse being the winner is

$$\begin{aligned}&\underline{P}(\{a\}) = 0.1, && \underline{P}(\{b\}) = 0.25, && \underline{P}(\{c\}) = 0.3,\\ &\underline{P}(\{a,b\}) = 0.4, && \underline{P}(\{a,c\}) = 0.6, && \underline{P}(\{b,c\}) = 0.7.\end{aligned}$$

There are rumors that c is not going to take part in the race due to some injury. What are the updated lower probabilities for a, b?

Taking into account that we are dealing with finite spaces and that the conditioning event has positive lower probability, applying the Generalised Bayes Rule is equivalent to taking the lower envelope of the linear conditional previsions that we obtain applying Bayes's rule on the elements of $\mathcal{M}(\underline{P})$. Thus, we obtain:

$$\underline{P}(\{a\}|\{a,b\}) = \inf\left\{\frac{P(\{a\})}{P(\{a,b\})} : P \in \mathcal{M}(\underline{P})\right\} = 0.1/0.5 = 0.2,$$

$$\underline{P}(\{b\}|\{a,b\}) = \inf\left\{\frac{P(\{b\})}{P(\{a,b\})} : P \in \mathcal{M}(\underline{P})\right\} = 0.25/0.55 = 0.45.$$

49. Consider two binary random variables X_1, X_2, and let $\underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ be given by:

$$\underline{P}(f|X_2 = 0) = \min\left\{\frac{f(0,0)+f(1,0)}{2}, f(0,0)\right\}$$
$$\underline{P}(f|X_2 = 1) = \min\left\{\frac{f(0,1)+f(1,1)}{2}, f(1,1)\right\}$$
$$\underline{P}(f|X_1 = 0) = \min\left\{\frac{f(0,0)+f(0,1)}{2}, f(0,0)\right\}$$
$$\underline{P}(f|X_1 = 1) = \min\left\{\frac{f(1,0)+f(1,1)}{2}, f(1,1)\right\}$$

for any gamble f on $\{0,1\}^2$. Are these conditional lower previsions coherent?

Yes. Let P_1 be the linear prevision associated to the uniform probability mass function on $\{0,1\}\times\{0,1\}$, and let P_2 be the linear prevision associated to the uniform mass function on $\{(0,0),(1,1)\}$. Let $\underline{P}$ be the coherent lower prevision given by the lower envelope of P_1, P_2. Then $\underline{P}(X_1 = i), \underline{P}(X_2 = i)$ are greater than 0 for $i = 0,1$; moreover, $\underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ are the conditional lower previsions defined from $\underline{P}$ by means of the Generalised Bayes' Rule. As a consequence, $\underline{P}, \underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ are weakly coherent, and since all the conditioning events have positive probability, this implies that they are also coherent.

50. Consider $\mathcal{X}_1 = \mathcal{X}_2 = \{1,2,3\}$, and let $\mathcal{M}$ be the set of probability mass functions on $\mathcal{X}_1 \times \mathcal{X}_2$ satisfying $P(1,2) = P(2,2) = P(3,1) = 0, P(1,1) = P(2,1), P(1,1) \geq P(1,3), P(2,1) \leq P(2,3)$, where the first index denotes the value of X_1 and the second the value of X_2. Let $\underline{P}$ be the lower envelope of the set $\mathcal{M}$.

(a) Compute the regular extensions $\underline{R}(X_1|X_2), \underline{R}(X_2|X_1)$.

The regular extensions are given by:

$$\underline{P}(f|X_2 = 1) = \frac{f(1,1)+f(2,1)}{2}$$
$$\underline{P}(f|X_2 = 2) = f(3,2)$$
$$\underline{P}(f|X_2 = 3) = \min\left\{f(3,3), f(1,3), \frac{f(1,3)+f(2,3)}{2}\right\}$$
$$\underline{P}(f|X_1 = 1) = \min\left\{f(1,1), \frac{f(1,1)+f(1,3)}{2}\right\}$$
$$\underline{P}(f|X_1 = 2) = \min\left\{f(2,3), \frac{f(2,1)+f(2,3)}{2}\right\}$$
$$\underline{P}(f|X_1 = 3) = \min\{f(3,2), f(3,3)\}$$

(b) Compute the natural extensions $\underline{E}(X_1|X_2), \underline{E}(X_2|X_1)$.

The natural extensions are vacuous, because all the conditioning events have lower probability equal to 0.

(c) Define $\underline{P}(X_2|X_1)$ from $\underline{P}$ using regular extension, and let $\underline{P}(X_1|X_2 = x)$ be defined from $\underline{P}$ by natural extension if $x = 3$ and by regular extension otherwise. Are $\underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ weakly coherent with $\underline{P}$?

Yes. Since each of the conditionals is bounded between the natural and the regular extensions, they are coherent with $\underline{P}$. Applying the characterisation of weak coherence, this implies that $\underline{P}(X_1|X_2), \underline{P}(X_2|X_1)$ weakly coherent with $\underline{P}$.

(d) Are they coherent?

No. Consider the gambles f_1, f_2, f_3 on $\mathcal{X}_1 \times \mathcal{X}_2$ given by the following table:

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
f_1	1	0	0	0	0	3	0	0	0
f_2	0	0	0	2	0	0	0	0	0
f_3	0	0	0	0	0	2	0	0	2

Then $G(f_1|X_1), G(f_2|X_2)$ and $-G(f_3|X_2 = 3)$ are given by

	(1,1)	(1,2)	(1,3)	(2,1)	(2,2)	(2,3)	(3,1)	(3,2)	(3,3)
$G(f_1\|X_1)$	0.5	-0.5	-0.5	-1.5	-1.5	1.5	0	0	0
$G(f_2\|X_2)$	-1	0	0	1	0	0	-1	0	0
$-G(f_3\|X_2 = 3)$	0	0	0	0	0	-2	0	0	-2

As a consequence, $[G(f_1|X_1) + G(f_2|X_2) - G(f_3|X_2 = 3)](x) < 0$ for any $x \in \pi_2^{-1}(3) \cup A_{f_1,f_2} = (\mathcal{X}_1 \times \mathcal{X}_2) \setminus \{(3,2)\}$.

51. **The three prisoners problem.** Three men, a, b and c, are in jail. Prisoner a knows that only two of the three prisoners will be executed, but he doesn't know who will be spared. He only knows that all three prisoners have equal probability $\frac{1}{3}$ of being spared. To the warden who knows which prisoner will be spared, a says, "Since two out of the three will be executed, it is certain that either b or c will be. You will give

me no information about my own chances if you give me the name of one man, b or c, who is going to be executed." Accepting this argument after some thinking, the warden says, "Prisoner b will be executed."

Does the warden's statement truly provide no information about the chance of a to be executed? We try to solve this problem using the theory of lower previsions.

(a) Let the variable X denote the prisoner that will be spared. Since all three prisoners have equal probability $\frac{1}{3}$ of being spared, we have a prior prevision specified by $\underline{P}_0(\{a\}) = \underline{P}_0(\{b\}) = \underline{P}_0(\{c\}) = \overline{P}_0(\{a\}) = \overline{P}_0(\{b\}) = \overline{P}_0(\{c\}) = \frac{1}{3}$. In a previous exercise, we have shown that the natural extension of $\underline{P}_0$ is given by

$$\underline{E}_0(f) = \frac{1}{3}\big(f(a) + f(b) + f(c)\big). \tag{5}$$

for any $f \in \mathcal{L}(X)$.

(b) Let the variable Y denote the prisoner named by the warden. Since the warden will not name a, we know that if $X = a$, then Y will be b or c, if $X = b$ then $Y = c$ and if $X = c$ then $Y = b$. Such information is modelled by vacuous conditional lower previsions, again, as described in one of the previous exercises:

$$\underline{P}(g|X = a) = \min\{g(b), g(c)\} \tag{6}$$
$$\underline{P}(g|X = b) = g(c) \tag{7}$$
$$\underline{P}(g|X = c) = g(b) \tag{8}$$

for any gamble $g \in \mathcal{L}(Y)$. Note that in case $X = a$, we do not know the mechanism by which the warden names either b or c for Y. Therefore, it seems appropriate to model this situation through a vacuous lower prevision relative to $\{b, c\}$.

(c) Combine the lower previsions $\underline{E}_0(\cdot)$ on $\mathcal{L}(X)$ and $\underline{P}(\cdot|X)$ on $\mathcal{L}(Y)$, using the marginal extension theorem, to a coherent lower prevision $\underline{E}$ on $\mathcal{L}(X \times Y)$.

From the marginal extension theorem, $\underline{E}(h) = \underline{E}_0(\underline{P}(h|X))$ for all $h \in \mathcal{L}(X \times Y)$. Using Eqs. (5) and (6)-(8), we find that

$$\underline{E}(h) = \frac{1}{3}\left(\min\{h(a, b), h(a, c)\} + h(b, c) + h(c, b)\right). \tag{9}$$

(d) Apply the generalised Bayes rule to calculate $\underline{E}(X = a|Y = b)$, $\overline{E}(X = a|Y = b)$ and $\underline{E}(X \neq a|Y = b)$, $\overline{E}(X \neq a|Y = b)$.

[Answer 1] Because $\underline{E}(Y = b) = \frac{1}{3}$, we can use the GBR to find the unique coherent conditional previsions $\underline{E}(f|Y = b)$, i.e., we have to solve

$$\underline{E}(I_{\{b\}}(f - \underline{E}(f|Y = b))) = 0.$$

Using Eq. (9), we find

$$\frac{1}{3}\left(\min\{f(a) - \underline{E}(f|Y = b), 0\} + 0 + f(c) - \underline{E}(f|Y = b)\right) = 0.$$

By putting $f = I_{\{a\}}$ respectively $f = 1 - I_{\{a\}}$, we find that $\underline{E}(X = a|Y = b) = 0$ respectively $\underline{E}(X \neq a|Y = b) = \frac{1}{2}$. Calculating the corresponding conjugate previsions gives $\overline{E}(X \neq a|Y = b) = 1$ respectively $\overline{E}(X = a|Y = b) = \frac{1}{2}$

[Answer 2: Alternative Approach] The same answer can be found more intuitively as follows.
First, suppose the warden decided beforehand to name c when a is spared (when c is spared he must name b). Because the warden actually names b, a is not spared and thus $P(a|b) = 0$. Secondly, suppose the opposite: the warden decided beforehand to name b when a is spared. As he names b, there are two equally likely possibilities: a or c is spared, so then $P(a|b) = \frac{1}{2}$.
We have given two extreme (deterministic) ways the warden can determine how two name a prisoner. He can also randomise between these two options: use the first with probability $1 - \lambda$ and the second with probability λ. This results in $P(a|b) = \frac{\lambda}{2}$, which can vary between 0 and $\frac{1}{2}$. These bounds correspond to the lower and upper previsions found previously.

(e) Extra exercise. After naming prisoner b as one of the prisoners to be executed, the warden thinks a little more and decides to play the following slightly sadistic game with prisoner a. The warden continues: "Are you really sure that I have given you no information at all by naming b? If you want to, for a reasonable fee I can arrange your fate to be switched with the fate of prisoner c. Of course, since I have not given you any information at all, you might not care about such arrangement. On the other hand, switching with prisoner c might just save your life... It's up to you to decide!"
Assume the utility of your life is equal to 25,000,000 Cuban Peso and the bribe requested by the warden is 25,000 Cuban Peso. Assuming that the warden really tells the truth about being able to arrange the switch, what would you do if you were prisoner a? (If the value of the bribe is zero, this game is isomorphic to the Monty Hall puzzle, as for instance described in de Cooman & Zaffalon, "Updating beliefs with incomplete observations", Artificial Intelligence, 2004, 159, pp.75-125.)

[Answer 1: Maximality, Rule 1] The decision "don't switch" corresponds to the following gamble $h_0 \in \mathcal{L}(X \times Y)$:

$$\begin{cases} h_0(a, b) = 25,000,000 \\ h_0(a, c) = 25,000,000 \\ h_0(b, c) = 0 \\ h_0(c, b) = 0 \end{cases}$$

The decision "switch" corresponds to the gamble $h_1 \in \mathcal{L}(X \times Y)$ specified by

$$\begin{cases} h_1(a, b) = -25,000 \\ h_1(a, c) = -25,000 \\ h_1(b, c) = 24,975,000 \\ h_1(c, b) = 24,975,000 \end{cases}$$

After the warden has named prisoner b, the conditional lower prevision $\underline{E}(\cdot|Y=b)$ describes our buying prices regarding $\mathcal{L}(X \times Y)$. In particular, we should prefer h_0 over h_1 if $\underline{E}(h_0 - h_1|Y=b) > 0$, and conversely, we should prefer h_1 over h_0 if $\underline{E}(h_1 - h_0|Y=b) > 0$. We find $\underline{E}(h_0 - h_1|Y=b)$ equal to

$$\begin{aligned}&\tfrac{1}{2}[h_0(c,b) - h_1(c,b)] + \tfrac{1}{2}\min\{h_0(a,b) - h_1(a,b); h_0(c,b) - h_1(c,b)\}\\ &= \tfrac{1}{2}[-24,975,000] + \tfrac{1}{2}\min\{25,025,000; -24,975,000\}\\ &= -24,975,000,\end{aligned}$$

and $\underline{E}(h_1 - h_0|Y=b)$ equal to

$$\begin{aligned}&\tfrac{1}{2}[h_1(c,b) - h_0(c,b)] + \tfrac{1}{2}\min\{h_1(a,b) - h_0(a,b); h_1(c,b) - h_0(c,b)\}\\ &= \tfrac{1}{2}[+24,975,000] + \tfrac{1}{2}\min\{-25,025,000, +24,975,000\}\\ &= -25,000.\end{aligned}$$

Both results are negative, this means that we have no preference at all: we have insufficient information in order to decide whether it is profitable to bribe the warden.

[Answer 2: Maximality, Rule 2] The decision criterion used above is based on conditional lower previsions. It has been criticised on the ground that it might assign a negative value to cost-free information. Using the unconditional lower prevision $\underline{E}$ (the "static" model) instead of $\underline{E}(\cdot|Y=b)$ (the "dynamic" model), we come up with a decision rule that never assigns negative value to information we should prefer h_0 over h_1 if $\underline{E}(h_0 - h_1) > 0$, and conversely, we should prefer h_1 over h_0 if $\underline{E}(h_1 - h_0) > 0$. (On the other hand, can we still use the unconditional lower prevision in case we have already observed $Y = b$ at the time of decision?) Using this rule, we find that $\underline{E}(h_1 - h_0)$ is equal to

$$\begin{aligned}&\tfrac{1}{3}\Big(\min\{h_1(a,b) - h_0(a,b); h_1(a,c) - h_0(a,c)\}\\ &\qquad + h_1(b,c) - h_0(b,c) + h_1(c,b) - h_0(c,b)\Big)\\ &= \tfrac{1}{3}\left(\min\{-25,025,000; -25,025,000\} + 24,975,000 + 24,975,000\right)\\ &= 8,308,333\end{aligned}$$

and $\underline{E}(h_0 - h_1)$ equal to

$$\begin{aligned}&\tfrac{1}{3}\Big(\min\{h_0(a,b) - h_1(a,b); h_0(a,c) - h_1(a,c)\}\\ &\qquad + h_0(b,c) - h_1(b,c) + h_0(c,b) - h_1(c,b)\Big)\\ &= \tfrac{1}{3}\left(\min\{25,025,000; 25,025,000\} - 24,975,000 - 24,975,000\right)\\ &= -8,308,333\end{aligned}$$

hence, we should strictly prefer h_1 over h_0: bribe the guard.
Note that the same two answers are obtained when taking E-admissibility instead of maximality as a decision rule. Indeed, E-admissibility coincides with maximality on pair-wise comparisons and only two actions h_0 and h_1 are involved. The result for Γ-maximin is left to the reader. The bribe was introduced to emphasize that the problem is not in the marginal preferences. In the Monty Hall puzzle (i.e., the case with bribe zero), under maximality rule 1, one of the preferences is marginal.

8 Extra exercises on conditional lower previsions

52. **The two envelopes problem.** The aim of this exercise is to demonstrate how mixing can annihilate imprecision, and how extra information does not necessarily lead to extra precision, when updating using Bayes rule. This latter phenomenon is called *dilation*.

I have two sealed envelopes, both containing money. One of them contains twice as much money as the other. You are free to pick one of them. You open it and find 100 Euro inside. You are provided the choice of either keeping the 100 Euro, or switching with whatever amount there is in the other envelope, which you know to be either 50 Euro or 200 Euro. Should you switch or not?

Let's introduce a few random variables. Let X be the amount in your envelope. Let Y be the smallest amount in the envelopes. Let Z be 1 if your envelope has the lowest value and 2 if the your envelope has the highest value. So,

$$X = YZ.$$

Since you pick at random, you know that the probability of $Z = 1$ is $\frac{1}{2}$, as is the probability of $Z = 2$. Moreover, since your choice is independent of how the money was distributed in the envelopes, Y is irrelevant to Z. A priori, we know nothing about Y.

It is evident that once $X = 100$ has been observed, if $Z = 1$ then $Y = 100$ and we should switch, and if $Z = 2$ then $Y = 50$ and we should not switch. However, we do not know Z; we only know that Z is uniformly distributed over $\{1, 2\}$. What should we do?

(a) Let f be a gamble on $\mathcal{Z}$ (with $\mathcal{Z} = \{1, 2\}$). For any $y \in \mathcal{Y}$ (with $\mathcal{Y} = \mathbb{R}^+$), what is the conditional lower prevision $\underline{P}(f|Y=y)$?

$\frac{1}{2}(f(1) + f(2))$.

(b) Let f be a gamble on $\mathcal{Y}$. What is the marginal lower prevision $\underline{P}(f)$?

$\inf_{y\in\mathcal{Y}} f(y)$

(c) Let f be a gamble on $\mathcal{Y} \times \mathcal{Z}$. Use ((a)), ((b)), and marginal extension, to arrive at an expression for $\underline{P}(f)$.

$\underline{P}(f) = \underline{P}(\underline{P}(f|\mathcal{Y})) = \inf_{y\in\mathcal{Y}} \frac{1}{2}(f(y,1) + f(y,2))$.

(d) Calculate $\mathcal{M}(\underline{P})$: show that $Q \in \mathcal{M}(\underline{P})$ if and only if there is a linear prevision R on $\mathcal{L}(Y)$ such that $Q(f) = \frac{1}{2}R(f(\cdot,1) + f(\cdot,2))$ for all gambles f on $\mathcal{Y} \times \mathcal{Z}$. [Hint: use ((c)), and the fact that there is a one-to-one correspondence between convex and compact sets of linear previsions and coherent lower previsions. (You do not need to prove compactness.)]

By ((c)), it is evident that the lower envelope of all linear previsions of the form $\frac{1}{2}R(f(\cdot,1)+f(\cdot,2))$ equals $\underline{P}$. Because the set $\{\frac{1}{2}R(f(\cdot,1)+f(\cdot,2))\colon R\in\mathbb{P}(\mathcal{Y})\}$ is clearly convex and compact, it follows that $\mathcal{M}(\underline{P})$ can only be this set.

(e) Conclude that $Q\in\mathcal{M}(\underline{P})$ if and only if $Q(f)=\frac{1}{2}Q(f(\cdot,1)+f(\cdot,2))$ for all gambles f on $\mathcal{Y}\times\mathcal{Z}$.

(f) What are the *prior* lower and upper previsions—before you open the envelope that you picked—of the amount in the other envelope minus the amount in your envelope? [Hint: consider every linear prevision Q in $\mathcal{M}(\underline{P})$ and invoke ((e)).]

By ((e)), we have for every Q in $\mathcal{M}(\underline{P})$,

$$Q(Y(3-Z)-YZ))=\frac{1}{2}Q(2Y-Y+Y-2Y)=0.$$

Hence the lower and upper prevision are zero: we are indifferent to switching.

(g) What can you say about the posterior distribution—after observing $X=100$—of Z? That is, for every Q in $\mathcal{M}(\underline{P})$ such that $Q(X=100)>0$, what are $Q(Z=1|X=100)$ and $Q(Z=2|X=100)$? (Note that the envelope of these probabilities is called *regular extension*.) Express these probabilities in terms of the distribution of Y under Q conditional on the event ($Y=100$ or $Y=50$).

First,

$$\begin{aligned}Q(Z=1|X=100)&=Q(Z=1\text{ and }X=100)/Q(X=100)\\&=Q(Z=1\text{ and }Y=100)/Q(X=100)\\&=\tfrac{1}{2}Q(Y=100)/Q(X=100)\end{aligned}$$

$$\begin{aligned}Q(Z=2|X=100)&=Q(Z=2\text{ and }X=100)/Q(X=100)\\&=Q(Z=2\text{ and }Y=50)/Q(X=100)\\&=\tfrac{1}{2}Q(Y=50)/Q(X=100)\end{aligned}$$

Next, $Q(X=100)$?

$$\begin{aligned}Q(X=100)&=Q(ZY=100)\\&=Q((Z=1\text{ and }Y=100)\text{ or }(Z=2\text{ and }Y=50))\\&=\tfrac{1}{2}Q(Y=100+Y=50)\\&=\tfrac{1}{2}(Q(Y=100)+Q(Y=50))\end{aligned}$$

Hence,

$$\begin{aligned}Q(Z=1|X=100)&=\frac{Q(Y=100)}{Q(Y=100)+Q(Y=50)}\\&=Q(Y=100|Y=100\text{ or }Y=50),\end{aligned}$$

$$\begin{aligned}Q(Z=2|X=100)&=\frac{Q(Y=50)}{Q(Y=100)+Q(Y=50)}\\&=Q(Y=50|Y=100\text{ or }Y=50).\end{aligned}$$

(h) What can you say about the posterior prevision—after observing $X=100$—of the amount in the other envelope minus the amount in your envelope? Again consider every Q in $\mathcal{M}(\underline{P})$ such that $Q(X=100)>0$.

It can range anywhere between -50 and $+100$, because

$$\begin{aligned}&Q(Y(3-Z)-YZ|X=100)\\&=Q(Y(3-Z)|X=100)-Q(YZ|X=100)\end{aligned}$$

and conditioning on X, $YZ=X$ is constant, and $Y(3-Z)$ is a function of Z only, whence

$$\begin{aligned}&=200\times Q(Z=1|X=100)+50\times Q(Z=2|X=100)-100\\&=200\times Q(Y=100|Y=100\text{ or }Y=50)\\&\qquad+50\times Q(Y=50|Y=100\text{ or }Y=50)-100\end{aligned}$$

9 Independent products

In these exercises, the different independent products will be made concrete. Consider two binary variables, X and Y, both for example representing coin flips, so taking values $\{H, T\}$. A coherent lower prevision for a binary variable is always a linear-vacuous lower prevision and therefore completely determined by its lower probability mass function. We here assume $\underline{P}_X(H) = \underline{P}_X(T) = \frac{1-\varepsilon}{2}$ and $\underline{P}_Y(H) = \underline{P}_Y(T) = \frac{1-\delta}{2}$, or in other words

$$\underline{P}_X\big(g(X)\big) = (1-\varepsilon)\frac{g(H)+g(T)}{2} + \varepsilon \min\{g(H), g(T)\}$$
$$\underline{P}_Y\big(h(Y)\big) = (1-\delta)\frac{h(H)+h(T)}{2} + \delta \min\{h(H), h(T)\}$$

We are interested in the event $HH \vee TT$, i.e., $(X, Y) \in \{(H, H), (T, T)\}$.

53. Calculate $(\underline{P}_X \boxtimes \underline{P}_Y)(HH \vee TT)$. Start by finding the extreme points of $\mathcal{M}(\underline{P}_X)$ and $\mathcal{M}(\underline{P}_Y)$. Use these to find a set with the extreme points of $\mathcal{M}(\underline{P}_X \boxtimes \underline{P}_Y)$. Then find the requested lower probability using the lower envelope theorem.

 The extreme points of the marginal credal sets are found by applying normalisation to the given lower probability values (let $\varepsilon^+ := 1+\varepsilon$, $\varepsilon^- := 1-\varepsilon$, $\delta^+ := 1+\delta$, and $\delta^- := 1-\delta$):

$$p := \begin{pmatrix} p_H \\ p_T \end{pmatrix}, \qquad \operatorname{ext}\mathcal{M}(\underline{P}_X) = \left\{ \frac{1}{2}\begin{pmatrix} \varepsilon^- \\ \varepsilon^+ \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \varepsilon^+ \\ \varepsilon^- \end{pmatrix} \right\},$$
$$\operatorname{ext}\mathcal{M}(\underline{P}_X) = \left\{ \frac{1}{2}\begin{pmatrix} \delta^- \\ \delta^+ \end{pmatrix}, \frac{1}{2}\begin{pmatrix} \delta^+ \\ \delta^- \end{pmatrix} \right\}.$$

 The extreme points of $\mathcal{M}(\underline{P}_X \boxtimes \underline{P}_Y)$ are then formed by pairwise multiplications of elements of $\mathcal{M}(\underline{P}_X)$ and $\mathcal{M}(\underline{P}_Y)$:

$$p := \begin{pmatrix} p_{HH} \\ p_{HT} \\ p_{TH} \\ p_{TT} \end{pmatrix},$$
$$\operatorname{ext}\mathcal{M}(\underline{P}_X \boxtimes \underline{P}_Y) = \left\{ \frac{1}{4}\begin{pmatrix} \varepsilon^-\delta^- \\ \varepsilon^-\delta^+ \\ \varepsilon^+\delta^- \\ \varepsilon^+\delta^+ \end{pmatrix}, \frac{1}{4}\begin{pmatrix} \varepsilon^-\delta^+ \\ \varepsilon^-\delta^- \\ \varepsilon^+\delta^+ \\ \varepsilon^+\delta^- \end{pmatrix}, \frac{1}{4}\begin{pmatrix} \varepsilon^+\delta^- \\ \varepsilon^+\delta^+ \\ \varepsilon^-\delta^- \\ \varepsilon^-\delta^+ \end{pmatrix}, \frac{1}{4}\begin{pmatrix} \varepsilon^+\delta^+ \\ \varepsilon^+\delta^- \\ \varepsilon^-\delta^+ \\ \varepsilon^-\delta^- \end{pmatrix} \right\}.$$

 By the lower envelope theorem,

$$\begin{aligned}(\underline{P}_X \boxtimes \underline{P}_Y)(HH \vee TT) &= \min\nolimits_{p \in \operatorname{ext}\mathcal{M}(\underline{P}_X \boxtimes \underline{P}_Y)}(p_{HH} + p_{TT}) \\ &= \frac{1}{4}\min\{\varepsilon^-\delta^- + \varepsilon^+\delta^+, \varepsilon^-\delta^+ + \varepsilon^+\delta^-\} \\ &= \frac{1}{2}\min\{1+\varepsilon\delta, 1-\varepsilon\delta\} = \frac{1-\varepsilon\delta}{2}.\end{aligned}$$

54. Calculate $(\underline{P}_X \circledcirc \underline{P}_Y)(HH \vee TT)$. Start by applying the definition of epistemic irrelevance to find $\underline{P}_Y(\cdot|X)$. Then apply the marginal extension theorem.

 By epistemic irrelevance of X to Y, we know that $\underline{P}_Y(\cdot|H) = \underline{P}_Y(\cdot|T) = \underline{P}_Y$. By the marginal extension theorem,

$$\begin{aligned}(\underline{P}_X \circledcirc \underline{P}_Y)(HH \vee TT) &= \underline{P}_X\big(\underline{P}_Y(HH \vee TT|X)\big) \\ &= \underline{P}_X\big(\underline{P}_Y(HH \vee TT|H)I_H + \underline{P}_Y(HH \vee TT|T)I_T\big) \\ &= \underline{P}_X\big(\underline{P}_Y(H)I_H + \underline{P}_Y(T)I_T\big) \\ &= \underline{P}_X\big(\frac{1-\delta}{2}I_H + \frac{1-\delta}{2}I_T\big) = \frac{1-\delta}{2}\underline{P}_X(1) = \frac{1-\delta}{2}.\end{aligned}$$

55. Calculate $(\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT)$. Start by finding the marginal gambles defining the sets of almost desirable gambles corresponding to $\underline{P}_X$ and $\underline{P}_Y$ as their positive linear hull. Use these to write down the marginal gambles whose positive linear hull is the set of almost desirable gambles corresponding to $\underline{P}_X \otimes \underline{P}_Y$. Then write down the linear program for calculating the natural extension to $I_{HH \vee TT}$ and solve it.

 (The argument order used here for the gambles is the same as the one used in the solution of Exercise 53 for mass functions.)

 Because by convention $\underline{P}_X(H) = \underline{P}_X(I_H)$, and likewise for Y instead of X and T instead of H, we have

$$G_{\underline{P}_X}(I_H) = G_{\underline{P}_X}(I_T) = \frac{1}{2}\begin{bmatrix} \varepsilon^+ \\ -\varepsilon^- \end{bmatrix}, \qquad G_{\underline{P}_Y}(I_H) = G_{\underline{P}_Y}(I_T) = \frac{1}{2}\begin{bmatrix} \delta^+ \\ -\delta^- \end{bmatrix}.$$

 The marginal gambles determining $\underline{P}_X \otimes \underline{P}_Y$ are then $G_{\underline{P}_X}(I_H)(X)I_H(Y)$, $G_{\underline{P}_X}(I_H)(X)I_T(Y)$, $G_{\underline{P}_X}(I_T)(X)I_H(Y)$, $G_{\underline{P}_X}(I_T)(X)I_T(Y)$, $I_H(X)G_{\underline{P}_Y}(I_H)(Y)$, $I_T(X)G_{\underline{P}_Y}(I_H)(Y)$, $I_H(X)G_{\underline{P}_Y}(I_T)(Y)$, and $I_T(X)G_{\underline{P}_Y}(I_T)(Y)$. So the linear program becomes (we can scale the marginal gambles at-will):

$$(\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT) = \max\left\{\alpha \in \mathbb{R} : \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} \lambda \right.$$

$$+ \begin{pmatrix} \varepsilon^+ & -\varepsilon^- & & \\ & & \varepsilon^+ & -\varepsilon^- \\ -\varepsilon^- & \varepsilon^+ & & \\ & & -\varepsilon^- & \varepsilon^+ \end{pmatrix} \mu$$

$$+ \begin{pmatrix} \delta^+ & -\delta^- & & \\ -\delta^- & \delta^+ & & \\ & & \delta^+ & -\delta^- \\ & & -\delta^- & \delta^+ \end{pmatrix} \nu$$

$$\left. + \begin{pmatrix} \alpha \\ \alpha \\ \alpha \\ \alpha \end{pmatrix} \leq \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, (\lambda, \mu, \nu) \in \left((\mathbb{R}_{\geq 0})^{\{H,T\}^2}\right)^3 \right\}.$$

As compared to $\lambda = 0$, λ with positive components can only decrease α, so we may take $\lambda = 0$. Also, because of the symmetry of and the position of the zeros and ones in the right-hand side gamble, we may take $\mu_{HH} = \mu_{TT}$, $\mu_{HT} = \mu_{HT} = 0$, $\nu_{HH} = \nu_{TT}$, and $\nu_{HT} = \nu_{HT} = 0$. So we can write:

$$(\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT) = \max\left\{\alpha \in \mathbb{R} : \begin{pmatrix} \varepsilon^+/\varepsilon^- \\ -1 \end{pmatrix} \mu + \begin{pmatrix} \delta^+/\delta^- \\ -1 \end{pmatrix} \nu \right.$$

$$\left. + \begin{pmatrix} \alpha \\ \alpha \end{pmatrix} \leq \begin{pmatrix} 1 \\ 0 \end{pmatrix}, (\mu, \nu) \in (\mathbb{R}_{\geq 0})^2 \right\}.$$

Unless $\epsilon = \delta$, either $\varepsilon^+/\varepsilon^-$ or δ^+/δ^- will be smaller, and therefore $\nu =$ or $\mu = 0$. So if $\varepsilon^+/\varepsilon^- \geq \delta^+/\delta^-$, then $(\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT) = (\underline{P}_X \circledcirc \underline{P}_Y)(HH \vee TT)$ and if $\varepsilon^+/\varepsilon^- \leq \delta^+/\delta^-$, then $(\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT) = (\underline{P}_X \circledcirc \underline{P}_Y)(HH \vee TT)$. So using the result of Exercise 54, we find

$$(\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT) = \frac{1 - \min\{\varepsilon, \delta\}}{2}.$$

56. Calculate $\underline{E}_{\underline{P}_x, \underline{P}_Y}(HH \vee TT)$. Start by writing down the cylindrical extensions of the marginal gambles corresponding to $\underline{P}_X$ and $\underline{P}_Y$, whose positive linear hull determines the set of almost desirable gambles corresponding to $\underline{E}_{\underline{P}_x, \underline{P}_Y}$. Then write down the linear program for calculating the natural extension to $I_{HH \vee TT}$ and solve it.

(The argument order used here for the gambles is the same as the one used in the solution of Exercise 53 for mass functions.)

$$\underline{E}_{\underline{P}_x, \underline{P}_Y}(HH \vee TT) = \max\left\{\alpha \in \mathbb{R} : \begin{pmatrix} \varepsilon^+ & -\varepsilon^- \\ \varepsilon^+ & -\varepsilon^- \\ -\varepsilon^- & \varepsilon^+ \\ -\varepsilon^- & \varepsilon^+ \end{pmatrix} \mu + \begin{pmatrix} \delta^+ & -\delta^- \\ -\delta^- & \delta^+ \\ \delta^+ & -\delta^- \\ -\delta^- & \delta^+ \end{pmatrix} \nu \right.$$

$$\left. + \begin{pmatrix} \alpha \\ \alpha \\ \alpha \\ \alpha \end{pmatrix} \leq \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, (\mu, \nu) \in \left((\mathbb{R}_{\geq 0})^2\right)^2 \right\}.$$

If $(\mu, \nu) = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$ is an optimal solution, then by symmetry of the gambles involved so are $(\lambda_2, \lambda_1, \lambda_3, \lambda_4)$ and $(\lambda_1, \lambda_2, \lambda_4, \lambda_3)$ and therefore by convexity of the solution set, we may take $\mu_1 = \mu_2$ and $\nu_1 = \nu_2$. But then the right-hand side becomes the vector with constant value $(\varepsilon^+ - \varepsilon^-)\mu + (\delta^+ - \delta^-)\nu + \alpha$. Because $\varepsilon^+ - \varepsilon^- = 2\varepsilon \geq 0$ and $\delta^+ - \delta^- = 2\delta \geq 0$, we must take $\mu = \nu = 0$ to get our optimal solution, which must then be $\alpha = 0$. So $\underline{E}_{\underline{P}_x, \underline{P}_Y}(HH \vee TT) = 0$, and by a similar reasoning $\underline{E}_{\underline{P}_x, \underline{P}_Y}(HT \vee TH) = 0$, so in the absence of any structural assumptions, the marginals contain no information about $HH \vee TT$.

57. Verify that in terms of commitments added about gambles on $HH \vee TT$, strong independence dominates epistemic independence, which dominates epistemic irrelevance, which dominates no structural assumptions.

Indeed, from the previous exercises, we see that

$$\begin{aligned}(\underline{P}_X \boxtimes \underline{P}_Y)(HH \vee TT) &\geq (\underline{P}_X \otimes \underline{P}_Y)(HH \vee TT) \\ &\geq (\underline{P}_X \circledcirc \underline{P}_Y)(HH \vee TT) \geq \underline{E}_{\underline{P}_x, \underline{P}_Y}(HH \vee TT),\end{aligned}$$

independently of the values of ε and δ.

58. Calculate the value of the
   - strong product,
   - independent natural extension, and
   - forward irrelevant $(X \to Y)$ natural extension

   for the gamble $f := I_{HH} - I_{HT} + 2I_{TH} - 2I_{TT}$. Start by factorizing the gamble as $f(X,Y) = g(X)h(Y)$. Then apply properties of the products to simplify the calculations.

$$f(X,Y) = \underbrace{\big(I_H(X) + 2I_T(X)\big)}_{g(X)} \underbrace{\big(I_H(Y) - I_T(Y)\big)}_{h(Y)}.$$

Because $g \geq 0$ and because the strong product and the independent natural extension are factorizing, we can write:

$$(\underline{P}_X \boxtimes \underline{P}_Y)(f) = (\underline{P}_X \otimes \underline{P}_Y)(f) = \underline{P}_X\big(g(X)\underline{P}_Y(h(Y))\big).$$

Now,

$$\underline{P}_Y\big(h(Y)\big) = (1-\delta)\frac{1-1}{2} + \delta\min\{1,-1\} = -\delta.$$

So then

$$\begin{aligned}\underline{P}_X\big(g(X)\underline{P}_Y(h(Y))\big) &= \underline{P}_X\big(-\delta g(X))\big) \\ &= -\delta\overline{P}_X\big(g(X))\big) \\ &= -\delta\big((1-\varepsilon)\frac{1+2}{2} + \varepsilon\max\{1,2\}\big) = -\frac{\delta(3+\varepsilon)}{2}.\end{aligned}$$

For the forward irrelevant natural extension, we need to calculate a marginal extension, which also simplifies because f factorizes and g is nonnegative:

$$(\underline{P}_X \circledS \underline{P}_Y)(f) = \underline{P}_X\big(g(X)\underline{P}_Y(h(Y))\big),$$

so the same value as for the other extensions.

59. Calculate the value of the
    - strong product,
    - independent natural extension, and
    - forward irrelevant $(X \to Y)$ natural extension

    for the gamble that returns the number of heads. Start by writing this gamble using indicator functions. Then apply properties of the products to simplify the calculations.

$$f(X,Y) \coloneqq I_H(X) + I_H(Y).$$

Because f is the sum of gambles that are the cylindrical extension of gambles on the marginal possibility spaces, we can use external additivity to calculate both the strong product and the independent natural extension:

$$(\underline{P}_X \boxtimes \underline{P}_Y)(f) = (\underline{P}_X \otimes \underline{P}_Y)(f) = \underline{P}_X\big(H\big) + \underline{P}_Y\big(H\big) = \frac{1-\varepsilon}{2} + \frac{1-\delta}{2} = 1 - \frac{\varepsilon+\delta}{2}.$$

For the forward irrelevant natural extension, we need to calculate a marginal extension, which simplifies thanks to constant additivity (cf. Exercise 5):

$$\begin{aligned}(\underline{P}_X \circledS \underline{P}_Y)(f) &= \underline{P}_X\big(\underline{P}_Y(I_H(X) + I_H(Y))\big) \\ &= \underline{P}_X\big(I_H(X) + \underline{P}_Y(H)\big) = \underline{P}_X(H) + \underline{P}_Y(H),\end{aligned}$$

so again the same value as for the other extensions.

10 Choquet integral with respect to Hausdorff measures

The following exercises allow you to practice calculating the Choquet integral with respect to Hausdorff measures. Let $\Omega \coloneqq [0,1]$ and let (Ω, d) be the Euclidean metric space.

60. Determine $P([0,1/2)|\Omega) = P(I_{[0,1/2)}|\Omega)$ and $P([0,1/2)|\{\varpi\}) = P(I_{[0,1/2)}|\{\varpi\})$ for any ϖ in Ω.

$$\begin{aligned}P([0,1/2)|\Omega) &= (\mathrm{C})\int I_{[0,1/2)}\mathrm{d}h^1 \\ &= \inf I_{[0,1/2)} + \int_{\inf I_{[0,1/2)}}^{\sup I_{[0,1/2)}} h^1(\{\omega\in\Omega\colon I_{[0,1/2)}(\omega) > x\})\mathrm{d}x \\ &= 0 + \int_0^1 h^1([0,1/2))\mathrm{d}x = \int_0^1 \frac{1}{2}\mathrm{d}x = \frac{1}{2}.\end{aligned}$$

$$\begin{aligned}P([0,1/2)|\{\varpi\}) &= (\mathrm{C})\int I_{[0,1/2)\cap\{\varpi\}}\mathrm{d}h^0 \\ &= \inf I_{[0,1/2)\cap\{\varpi\}} + \int_{\inf I_{[0,1/2)\cap\{\varpi\}}}^{\sup I_{[0,1/2)\cap\{\varpi\}}} h^0(\{\omega\in\{\varpi\}\colon I_{[0,1/2)}(\omega) > x\})\mathrm{d}x \\ &= I_{[0,1/2)}(\varpi) + 0 = I_{[0,1/2)}(\varpi).\end{aligned}$$

61. Determine $P(\mathbb{Q}|\Omega) = P(I_{\mathbb{Q}}|\Omega)$ and $P(\mathbb{Q}|\{\varpi\}) = P(I_{\mathbb{Q}}|\{\varpi\})$ for any ϖ in Ω.

$$\begin{aligned}P(\mathbb{Q}|\Omega) &= (\mathrm{C})\int I_{\mathbb{Q}\cap\Omega}\mathrm{d}h^1 \\ &= \inf I_{\mathbb{Q}\cap\Omega} + \int_{\inf I_{\mathbb{Q}\cap\Omega}}^{\sup I_{\mathbb{Q}\cap\Omega}} h^1(\{\omega\in\Omega\colon I_{\mathbb{Q}}(\omega) > x\})\mathrm{d}x \\ &= 0 + \int_0^1 h^1(\mathbb{Q}\cap\Omega)\mathrm{d}x = \int_0^1 0\mathrm{d}x = 0.\end{aligned}$$

$$\begin{aligned}P(\mathbb{Q}|\{\varpi\}) &= (\mathrm{C})\int I_{\mathbb{Q}\cap\{\varpi\}}\mathrm{d}h^0 \\ &= \inf I_{\mathbb{Q}\cap\{\varpi\}} + \int_{\inf I_{\mathbb{Q}\cap\{\varpi\}}}^{\sup I_{\mathbb{Q}\cap\{\varpi\}}} h^0(\{\omega\in\{\varpi\}\colon I_{\mathbb{Q}}(\omega) > x\})\mathrm{d}x \\ &= I_{\mathbb{Q}}(\varpi) + 0 = I_{\mathbb{Q}}(\varpi).\end{aligned}$$

A remark about Exercises 60 and 61: Coherent conditional and unconditional prevision are linear because the conditioning events are Borel-measurable and the random variables are Borel-measurable.

62. Determine $P(A|\Omega) = P(I_A|\Omega)$ and $P(A|\{\varpi\}) = P(I_A|\{\varpi\})$ for any ϖ in Ω and $A \subseteq \mathbb{R}$.

$$\begin{aligned}
P(A|\Omega) &= (\mathrm{C})\int I_{A\cap\Omega}\mathrm{d}h^1\\
&= \inf I_{A\cap\Omega} + \int_{\inf I_{A\cap\Omega}}^{\sup I_{A\cap\Omega}} h^1(\{\omega\in\Omega\colon I_A(\omega) > x\})\mathrm{d}x\\
&= 0 + \int_0^1 h^1(A\cap\Omega)\mathrm{d}x = h^1(A\cap\Omega).
\end{aligned}$$

$$\begin{aligned}
P(A|\{\varpi\}) &= (\mathrm{C})\int I_{A\cap\{\varpi\}}\mathrm{d}h^0\\
&= \inf I_{A\cap\{\varpi\}} + \int_{\inf I_{A\cap\{\varpi\}}}^{\sup I_{A\cap\{\varpi\}}} h^0(\{\omega\in\{\varpi\}\colon I_A(\omega) > x\})\mathrm{d}x\\
&= I_A(\varpi) + 0 = I_A(\varpi).
\end{aligned}$$

A remark about Exercise 62: If the set A is Lebesgue measurable then its indicator function is h^1-measurable and the unconditional prevision is linear. The conditional prevision is linear because the indicator function of any set A is h^0-measurable

63. Given $X : \Omega \to \mathbb{R} : \omega \mapsto \omega$ and $Y : \Omega \to \mathbb{R} : \omega \mapsto \omega^2$ Determine $P(X|\Omega)$ and $P(Y|\Omega)$.

$$\begin{aligned}
P(X|\Omega) = (\mathrm{C})\int X\mathrm{d}h^1 &= \inf X + \int_{\inf X}^{\sup X} h^1(\{\omega\in\Omega\colon X(\omega) > x\})\mathrm{d}x\\
&= 0 + \int_0^1 h^1(\{\omega\in\Omega\colon \omega > x\})\mathrm{d}x\\
&= \int_0^1 (1-x)\mathrm{d}x = \frac{1}{2}.
\end{aligned}$$

$$\begin{aligned}
P(Y|\Omega) = (\mathrm{C})\int Y\mathrm{d}h^1 &= \inf Y + \int_{\inf Y}^{\sup Y} h^1(\{\omega\in\Omega\colon Y(\omega) > x\})\mathrm{d}x\\
&= 0 + \int_0^1 h^1(\{\omega\in\Omega\colon \omega^2 > x\})\mathrm{d}x\\
&= \int_0^1 (1-\sqrt{x})\mathrm{d}x = 1 - \frac{2}{3}\left[\sqrt[3]{x}\right]_{x=0}^{x=1} = \frac{1}{3}.
\end{aligned}$$