Some Exercises on Coherent Lower Previsions

Gert de Cooman, Erik Quaeghebeur and Matthias C. M. Troffaes

4th August 2004

1 Probability Measures and Linear Previsions

Suppose we have a probability measure μ defined on the power set $\wp(\mathcal{X})$ of a finite set $\mathcal{X}$. Mathematically, this means that μ is a $\wp(\mathcal{X}) - \mathbb{R}$ map satisfying $\mu(\emptyset) = 0$, $\mu(\mathcal{X}) = 1$, $\mu(A) \geq 0$ for all $A \subseteq \mathcal{X}$, and $\mu(A \cup B) = \mu(A) + \mu(B)$ for any $A, B \subseteq \mathcal{X}$ such that $A \cap B = \emptyset$.

We may interpret these values $\mu(A)$ as fair prices for indicator gambles I_A. This corresponds to the lower prevision $\underline{P}$ defined on $\{I_A : A \subseteq \mathcal{X}\} \cup \{-I_A : A \subseteq \mathcal{X}\}$ by

$$\underline{P}(I_A) = \overline{P}(I_A) := \mu(A) \tag{1.1}$$

for all $x \in \mathcal{X}$. Recall that we denote $\underline{P}(I_A)$ by $\underline{P}(A)$ and $-\underline{P}(-I_A) = \overline{P}(I_A)$ by $\overline{P}(A)$.

In this exercise, we shall show that natural extension of the lower prevision $\underline{P}$ representing μ coincides with integration with respect to μ.

(a) Preparatory exercise. Show that the integral with respect to μ of gambles $f \in \mathcal{L}(\mathcal{X})$,

$$P_\mu(f) := \int f \, \mathrm{d}\mu = \sum_{x \in \mathcal{X}} \mu(\{x\}) f(x), \tag{1.2}$$

defines a linear prevision on $\mathcal{L}(\mathcal{X})$.

Answer. The domain of P_μ is a linear space (its the set of all gambles on $\mathcal{X}$), so it suffices to check that

(i) $P_\mu(f) \geq \inf_{x \in \mathcal{X}} f(x)$ for any $f \in \mathcal{L}(\mathcal{X})$,

(ii) $P_\mu(f + g) = P_\mu(f) + P_\mu(g)$ for any $f, g \in \mathcal{L}(\mathcal{X})$.

□

(b) Show that $\underline{P}$ avoids sure loss.

Answer: Dual Approach. P_μ is a linear prevision that dominates $\underline{P}$:

$$P_\mu(I_A) = \int I_A \, \mathrm{d}\mu = \mu(A) \geq \underline{P}(I_A), \tag{1.3}$$

$$P_\mu(-I_A) = \int -I_A \, \mathrm{d}\mu = -\mu(A) \geq \underline{P}(-I_A). \tag{1.4}$$

Hence, $\mathcal{M}(\underline{P}) \neq \emptyset$ since $P_\mu \in \mathcal{M}(\underline{P})$, and therefore $\underline{P}$ avoids sure loss. □

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

Answer: Think Before You Act. $\underline{P}$ is self-conjugate: $\underline{P}(I_A) = -\underline{P}(-I_A)$ for any $A \subseteq \mathcal{X}$. It also avoids sure loss. These two conditions are sufficient for coherence. □

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

$$\underline{E}(f) = \int f \, \mathrm{d}\mu = \sum_{x \in \mathcal{X}} \mu(\{x\}) f(x). \tag{1.5}$$

Answer 1: Primal Approach (Masochist). By definition, the natural extension $\underline{E}(f)$ of a gamble $f \in \mathcal{L}(\mathcal{X})$ is equal to

$$\sup\Big\{\gamma \in \mathbb{R}\colon (\forall A \subseteq \mathcal{X})(\exists \lambda_A \in \mathbb{R})(\forall x \in \mathcal{X})\Big(f(x) - \gamma \geq \sum_{A \subseteq \mathcal{X}} \lambda_A(I_A(x) - \mu(A))\Big)\Big\} \tag{1.6}$$

This is a linear program with free variables γ and λ_A (for all $A \subseteq \mathcal{X}$), and a linear inequality for each $x \in \mathcal{X}$.

First, observe that the system of inequalities in Eq. (1.6) is satisfied for $\gamma = \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$ by choosing $\lambda_{\{x\}} = f(x)$ and $\lambda_A = 0$ for all $A \subseteq \mathcal{X}$ that are not singletons. Hence, $\underline{E}(f) \geq \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$.

Suppose γ is another solution of the system of inequalities in Eq. (1.6). We show that $\gamma \leq \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$. Indeed, taking a convex combination of the inequalities, with coefficient $\alpha_x \geq 0$ for the inequality that corresponds to $x \in \mathcal{X}$ (with $\sum_{x\in\mathcal{X}} \alpha_x = 1$), we find that

$$\sum_{x\in\mathcal{X}} \alpha_x f(x) - \gamma \geq \sum_{A\subseteq\mathcal{X}} \Big(\lambda_A\Big(\sum_{x\in\mathcal{X}} \alpha_x I_A(x)\Big) - \mu(A)\Big). \tag{1.7}$$

It is clear that $\alpha_x = \mu(\{x\})$ is the combination we are looking for. Use the additivity of μ to see that the right hand side is zero. Hence, $\gamma \leq \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$, and therefore also $\underline{E}(f) \leq \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$. We already proved that $\underline{E}(f) \geq \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$, hence, $\underline{E}(f) = \sum_{x\in\mathcal{X}} \mu(\{x\})f(x)$, which establishes the proof. □

Answer 2: Dual Approach. First, observe that the natural extension $\underline{E}$ of $\underline{P}$ is a linear prevision. Indeed, for any non-negative gamble $f \in \mathcal{L}(\mathcal{X})$ it holds that

$$\overline{E}(f) \leq \sum_{x\in\mathcal{X}} f(x)\overline{E}(\{x\}) \tag{1.8}$$

$$= \sum_{x\in\mathcal{X}} f(x)\overline{P}(\{x\}) \tag{1.9}$$

$$= \sum_{x\in\mathcal{X}} f(x)\underline{P}(\{x\}) \tag{1.10}$$

$$= \sum_{x\in\mathcal{X}} f(x)\underline{E}(\{x\}) \tag{1.11}$$

$$\leq \underline{E}(f), \tag{1.12}$$

where we used the coherence of $\underline{P}$, the coherence of $\underline{E}$ and the self-conjugacy of $\underline{P}$. By coherence of $\underline{E}$ also $\underline{E}(f) \leq \overline{E}(f)$, and hence, $\underline{E}(f) = \overline{E}(f)$ for all non-negative gambles f. But then, by coherence of $\underline{E}$,

$$\underline{E}(g) = \underline{E}\big(g - \underline{P}_{\mathcal{X}}(g)\big) + \underline{P}_{\mathcal{X}}(g) = \overline{E}\big(g - \underline{P}_{\mathcal{X}}(g)\big) + \underline{P}_{\mathcal{X}}(g) = \overline{E}(g) \tag{1.13}$$

for any gamble g. Therefore, $\underline{E}$ is self-conjugate. Since it is also coherent, it must be linear.

But this implies that $\mathcal{M}(\underline{P}) = \{\underline{E}\}$. Indeed, suppose that $Q \in \mathcal{M}(\underline{P})$. Then $Q(f) \geq \underline{E}(f)$ for any gamble f. But also, $Q(f) \leq \overline{E}(f) = \underline{E}(f)$ for any gamble f. Hence indeed, $Q = \underline{E}$.

We have already established that $P_\mu \in \mathcal{M}(\underline{P})$ in our proof for avoiding sure loss. Therefore, it can only be that $P_\mu = \underline{E}$. □

Answer 3: Think Before You Act. The statement is established if we show that P_μ is the point-wise smallest coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$ on its domain.

Suppose $\underline{Q}$ is another coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$ on its domain, that is, $\underline{Q}(I_A) \geq \mu(A)$ and $\underline{Q}(-I_A) \geq -\mu(A)$. Let $f \in \mathcal{L}(\mathcal{X})$. Since $\underline{Q}$ is coherent, we easily find that

$$\underline{Q}(f) = \underline{Q}\left(\sum_{x\in\mathcal{X}} f(x)I_{\{x\}}\right) \tag{1.14}$$

$$\geq \sum_{\substack{x\in\mathcal{X}\\ f(x)\geq 0}} f(x)\underline{Q}(I_{\{x\}}) + \sum_{\substack{x\in\mathcal{X}\\ f(x)<0}} [-f(x)]\underline{Q}(-I_{\{x\}}) \tag{1.15}$$

$$\geq \sum_{x\in\mathcal{X}} f(x)\mu(\{x\}) \tag{1.16}$$

$$= P_\mu(f), \tag{1.17}$$

since $\underline{Q}(I_{\{x\}}) \geq \mu(\{x\}) \geq -\underline{Q}(-I_{\{x\}})$. This establishes the proof. □

(e) Extra exercise. Show that natural extension defines a one-to-one correspondence between probability measures μ on $\wp(\mathcal{X})$ and linear previsions Q on $\mathcal{L}(\mathcal{X})$. Hence, through coherence, linear previsions are uniquely determined by their values on singletons whenever $\mathcal{X}$ is a finite set.

(f) Another extra exercise. Show that even when $\mathcal{X}$ is not a finite set, there still is a one-to-one correspondence between finitely additive probability measures μ on $\wp(\mathcal{X})$ and linear previsions Q on $\mathcal{L}(\mathcal{X})$. Hint: first stick to simple gambles, then use continuity of linear previsions with respect to the topology of uniform convergence. Hence, a linear prevision on $\mathcal{L}(\mathcal{X})$ is uniquely determined by its value on events (subsets of $\mathcal{X}$).

(g) Yet another extra exercise (non-additive measures). Suppose we have a 2-monotone measure μ defined on the power set $\wp(\mathcal{X})$ of a finite set $\mathcal{X}$, that is, $\mu(\emptyset) = 0$, $\mu(\mathcal{X}) = 1$, $\mu(A) \geq 0$ for all $A \subseteq \mathcal{X}$, and

$$\mu(A \cup B) + \mu(A \cap B) \geq \mu(A) + \mu(B) \tag{1.18}$$

for any $A, B \subseteq \mathcal{X}$. We may interpret these 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 $\underline{P}(I_A) := \mu(A)$ for all $x \in \mathcal{X}$. Show that the lower prevision $\underline{P}$ representing μ is coherent, and that the natural extension of $\underline{P}$ coincides with $\mathrm{C}\int \cdot \,\mathrm{d}\mu$, the Choquet integral with respect to μ. Hint: first show that the Choquet integral defines a coherent lower prevision (use the sub-additivity theorem, which says that $\mathrm{C}\int(f+g)\,\mathrm{d}\mu \geq \mathrm{C}\int f\,\mathrm{d}\mu + \mathrm{C}\int g\,\mathrm{d}\mu$).

Note that the Choquet integral of a gamble f on a finite set $\mathcal{X}$ can be constructed as follows. Since $\mathcal{X}$ is finite, without loss of generality we can write f as

$$f = \alpha_0 + \sum_{i=1}^{n} \alpha_i I_{A_i} \tag{1.19}$$

with $\alpha_0 \in \mathbb{R}$, $\alpha_1 > 0$, $\alpha_2 > 0, \ldots, \alpha_n > 0$ and $A_1 \supset A_2 \supset \cdots \supset A_n$ (where $A \supset B$ means $A \supseteq B$ and $A \neq B$). In terms of α_i's and A_i's, the Choquet integral of f is simply given by

$$\mathrm{C}\int f\,\mathrm{d}\mu := \alpha_0 + \sum_{i=1}^{n} \alpha_i \mu(A_i). \tag{1.20}$$

2 Vacuous Lower Previsions

Let A be a non-empty subset of a (not necessarily finite) set $\mathcal{X}$. Say we only know that the lower probability of A is equal to 1. This assessment is embodied through the lower prevision $\underline{P}$ defined on the singleton $\{I_A\}$ by $\underline{P}(A) = 1$ (again, recall that we denote $\underline{P}(I_A)$ by $\underline{P}(A)$).

(a) Preparatory exercise. Show that the vacuous lower prevision relative to A, defined by

$$\underline{P}_A(f) := \inf_{x \in A} f(x) \tag{2.1}$$

for any $f \in \mathcal{L}(\mathcal{X})$, is a coherent lower prevision on $\mathcal{L}(\mathcal{X})$.

Answer. The domain of $\underline{P}_A$ is a linear space, so it suffices to check that

(i) $\underline{P}_A(f) \geq \inf_{x \in \mathcal{X}} f(x)$ for any $f \in \mathcal{L}(\mathcal{X})$,

(ii) $\underline{P}_A(\lambda f) = \lambda \underline{P}_A(f)$ for any $f \in \mathcal{L}(\mathcal{X})$ and any $\lambda > 0$, and

(iii) $\underline{P}_A(f + g) \geq \underline{P}_A(f) + \underline{P}_A(g)$ for any $f, g \in \mathcal{L}(\mathcal{X})$.

□

(b) Show that $\underline{P}$ avoids sure loss.

Answer 1: Primal Approach. For any $\lambda \geq 0$ it holds that $\sup\{\lambda(I_A - 1)\} = 0$ (recall that $A \neq \emptyset$), which is non-negative, and hence, by definition, $\underline{P}$ avoids sure loss. □

Answer 2: Dual Approach. First, observe that

$$\mathcal{M}(\underline{P}) = \{Q \in \mathcal{P}(\mathcal{X}) \colon Q(A) = 1\}. \tag{2.2}$$

Indeed, if $Q(A) = 1$ then $Q(A) \geq \underline{P}(A)$, and hence, Q dominates $\underline{P}$. Therefore, $Q \in \mathcal{M}(\underline{P})$. Conversely, if $Q \in \mathcal{M}(\underline{P})$ then $Q(A) \geq \underline{P}(A) \geq 1$. Since Q is coherent, also $Q(A) \leq 1$. So it must be that $Q(A) = 1$.

Take any $x \in A$. Observe that by Eq. (2.2) the vacuous lower prevision $\underline{P}_x$ relative to the singleton $\{x\}$ belongs to $\mathcal{M}(\underline{P})$ since $\underline{P}_x$ is a linear prevision and

$$\underline{P}_x(A) = I_A(x) = 1. \tag{2.3}$$

Hence, $\mathcal{M}(\underline{P})$ is non-empty. Therefore $\underline{P}$ avoids sure loss. □

Answer 3: A Variation on Answer 2. Observe that the vacuous lower prevision relative to A is a coherent lower prevision which dominates $\underline{P}$ on $\{I_A\}$:

$$\underline{P}_A(A) = \inf_{x \in A} I_A(x) = 1 \geq \underline{P}(A). \tag{2.4}$$

So $\underline{P}$ avoids sure loss. □

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

Answer 1: Primal Approach. We already demonstrated that $\underline{P}$ avoids sure loss. We are left to show that for any $\lambda < 0$ it holds that $\sup\{\lambda(I_A - 1)\} \geq 0$. If $A = \mathcal{X}$ then this supremum is zero. Otherwise, the supremum is equal to $-\lambda$, which is strictly positive. In all cases the supremum is non-negative, and hence, $\underline{P}$ is coherent. □

Answer 2: Dual Approach. Consider the set $\mathcal{M}(\underline{P})$ of linear previsions on $\mathcal{L}(\mathcal{X})$ that dominate $\underline{P}$. Since $Q(A) = 1$ for any $Q \in \mathcal{M}(\underline{P})$, it follows that

$$\inf_{Q\in\mathcal{M}(\underline{P})} Q(A) = \inf_{Q\in\mathcal{M}(\underline{P})} 1 = 1 = \underline{P}(A), \tag{2.5}$$

and hence, $\underline{P}$ must be coherent, since it coincides on its domain with the lower envelope of its set of dominating linear previsions. □

Answer 3: Think Before You Act. Observe that $\underline{P}$ is the restriction to $\{I_A\}$ of $\underline{P}_A$, which is coherent:

$$\underline{P}_A(A) = \inf_{x\in A} I_A(x) = 1 = \underline{P}(A). \tag{2.6}$$

Hence, $\underline{P}$ is coherent too. □

(d) Prove that the natural extension $\underline{E}$ of $\underline{P}$ is equal to the vacuous lower prevision relative to A:

$$\underline{E}(f) = \underline{P}_A(f) = \inf_{x\in A} f(x), \tag{2.7}$$

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

Answer 1: Primal Approach. $\underline{E}(f)$ is equal to the supremum achieved by the free variable γ subject to the constraint

$$f - \gamma \geq \lambda(I_A - 1) \tag{2.8}$$

with variable $\lambda \geq 0$. Note that $\gamma = \inf_{x\in A} f(x)$ and $\lambda = \inf_{x\in A} f(x) - \inf_{x\in\mathcal{X}} f(x)$ yields a feasible solution of Eq. (2.8). Therefore, $\underline{E}(f) \geq \inf_{x\in A} f(x)$.

Let γ and λ constitute any feasible solution of Eq. (2.8). Then, since $\inf_{x\in A}$ is monotone, we find in particular that

$$\inf_{x\in A} f(x) - \gamma \geq \inf_{x\in A} \lambda(I_A(x) - 1) \tag{2.9}$$

Note that the right hand side is zero. Hence, $\gamma \leq \inf_{x\in A} f(x)$. Therefore, also $\underline{E}(f) \leq \inf_{x\in A} f(x)$. But, we already proved that $\underline{E}(f) \geq \inf_{x\in A} f(x)$, and hence, $\underline{E}(f) = \inf_{x\in A} f(x)$. □

Answer 2: Dual Approach. Again consider the set $\mathcal{M}(\underline{P})$ of linear previsions on $\mathcal{L}(\mathcal{X})$ that dominate $\underline{P}$. Now, if we can show that

$$\operatorname{ext}\mathcal{M}(\underline{P}) = \{\underline{P}_x : x \in A\}. \tag{2.10}$$

then the claim is established, since in that case

$$\underline{E}(f) = \inf_{Q\in\mathcal{M}(\underline{P})} Q(f) = \inf_{Q\in\operatorname{ext}\mathcal{M}(\underline{P})} Q(f) = \inf_{x\in A} \underline{P}_x(f) = \inf_{x\in A} f(x). \tag{2.11}$$

We shall prove Eq. (2.10) in case that $\mathcal{X}$ is a finite set (it holds in general case, but the proof becomes a bit more complex).

Assume that $\mathcal{X}$ is a finite set. To see that Eq. (2.10) holds, let $Q \in \mathcal{M}(\underline{P})$. Observe that

$$Q(\{x\}) = 1 - Q(\mathcal{X} \setminus \{x\}) \leq 1 - Q(A) = 0 \tag{2.12}$$

for any $x \notin A$, and hence

$$\sum_{x\in A} Q(\{x\}) = \sum_{x\in\mathcal{X}} Q(\{x\}) = 1. \tag{2.13}$$

This implies that any $Q \in \mathcal{M}(\underline{P})$ is a convex mixture of $\underline{P}_x$ for $x \in A$ (see exercise on linear previsions):

$$Q(f) = \sum_{x \in \mathcal{X}} Q(\{x\}) f(x) = \sum_{x \in A} Q(\{x\}) f(x) = \sum_{x \in A} Q(\{x\}) \underline{P}_x(f). \tag{2.14}$$

Since all $\underline{P}_x$ are linearly independent, Eq. (2.10) follows. □

Answer 3: Think Before You Act. The statement is established if we show that $\underline{P}_A$ is the point-wise smallest coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$ on $\{I_A\}$.

Suppose $\underline{Q}$ is another coherent lower prevision on $\mathcal{L}(\mathcal{X})$ which dominates $\underline{P}$ on $\{I_A\}$, that is, $\underline{Q}(A) = 1$. Let $f \in \mathcal{L}(\mathcal{X})$. It is easy to check that

$$f \geq \underline{P}_A(f) + [\underline{P}_A(f) - \underline{P}_{\mathcal{X}}(f)](I_A - 1), \tag{2.15}$$

Where $\underline{P}_{\mathcal{X}}(f) = \inf_{x \in \mathcal{X}} f(x)$. Since $\underline{Q}$ is coherent, this implies that

$$\underline{Q}(f) \geq \underline{Q}\big(\underline{P}_A(f) + [\underline{P}_A(f) - \underline{P}_{\mathcal{X}}(f)](I_A - 1)\big) \tag{2.16}$$

$$= \underline{P}_A(f) + [\underline{P}_A(f) - \underline{P}_{\mathcal{X}}(f)]\underline{Q}(I_A - 1)\big) \tag{2.17}$$

$$= \underline{P}_A(f), \tag{2.18}$$

since $\underline{Q}(I_A - 1) = \underline{Q}(A) - 1 = 0$. This establishes the proof. □

(e) Extra exercise. Each one of the questions (b), (c) and (d) can be solved in three different ways, either using

- (i) the primal form—combinations of desirable gambles,
- (ii) the dual form—sets of probability measures, or
- (iii) the properties of coherence and natural extension, invoking the result proven in the preparatory exercise (a).

Invoke each one of the methods (i), (ii) and (iii) to answer each one of the questions (b), (c) and (d). You may cheat when solving question (d) using method (ii): it is much easier if you assume that $\mathcal{X}$ is finite.

3 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

$$\underline{P}_{x_1}(f) := f(x_1), \tag{3.1}$$

$$\underline{P}_{x_2}(f) := f(x_2), \tag{3.2}$$

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 $\{x_1, x_2\}$:

$$\underline{P}(f) = \min\{f(x_1), f(x_2)\}. \tag{3.3}$$

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

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

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

$$F_{x_1}(x) = \underline{P}_{x_1}(\{y \in \mathcal{X} \colon y \le x\}) = \begin{cases} 0, & \text{if } x < x_1, \\ 1, & \text{if } x \ge x_1, \end{cases} \tag{3.4}$$

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 $Q \in \mathcal{P}(\mathcal{X})$ whose cumulative distribution function

$$F_Q(x) = Q(\{y \in \mathcal{X} \colon y \le x\}) \tag{3.5}$$

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

$$\underline{E}(f) \le \underline{P}_{[x_1,x_2]}(f) \text{ for any gamble } f \in \mathcal{L}(\mathcal{X}). \tag{3.6}$$

What does this mean?

Answer. 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}:

$$F_{x_1}(x) \ge Q(\{y \in \mathcal{X} \colon y \le x\}) \ge F_{x_2}(x) \tag{3.7}$$

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

$$\begin{cases} Q(\{y \in \mathcal{X} \colon y \le x\}) = 0, & \text{if } x < x_1, \\ Q(\{y \in \mathcal{X} \colon y \le x\}) = 1, & \text{if } x \ge x_2. \end{cases} \tag{3.8}$$

Observe that Eq. (3.8) is satisfied for $Q = \underline{P}_x$, the vacuous lower prevision relative to the singleton $\{x\}$, whenever $x \in [x_1, x_2]$. Hence,

$$\{\underline{P}_x \colon x \in [x_1, x_2]\} \subseteq \mathcal{M}. \tag{3.9}$$

Therefore,

$$\underline{E}(f) = \inf_{Q \in \mathcal{M}} Q(f) \le \inf_{x \in [x_1,x_2]} \underline{P}_x(f) = \underline{P}_{[x_1,x_2]}(f), \tag{3.10}$$

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

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

$$\underline{E}(f) = \sup_{\epsilon > 0} \underline{P}_{[x_1-\epsilon,x_2]}(f) \text{ for any gamble } f \in \mathcal{L}(\mathcal{X}). \tag{3.11}$$

4 The Fréchet Bounds

Assume A and $B \subseteq \mathcal{X}$ are logically independent events: $A \cap B$, $A^c \cap B$, $A \cap B^c$ and $A^c \cap B^c$ are non-empty ($\cdot^c$ denotes complementation in $\mathcal{X}$). We assess lower and upper probabilities for A and B, embodied through a lower prevision $\underline{P}$ defined on the set of gambles

$$\mathcal{K} = \{I_A, -I_A, I_B, -I_B\}. \tag{4.1}$$

Recall that we denote $\underline{P}(I_A)$ by $\underline{P}(A)$ and $-\underline{P}(-I_A) = \overline{P}(I_A)$ by $\overline{P}(A)$, and similar for B. Also recall that $\underline{P}(A^c) = 1 - \overline{P}(A)$, because $\underline{P}(1 - I_A) = 1 - \overline{P}(I_A)$.

(a) Preparatory exercise. Consider the case $\underline{P}(A) = \overline{P}(A) = p \in [0,1]$ and $\underline{P}(B) = \overline{P}(B) = q \in [0,1]$. Find a one-dimensional parametrisation of the set of probability measures, defined on the algebra generated by A and B, which are compatible with $\underline{P}$. Recall that a probability measure μ is compatible with a lower prevision if for all gambles f in the domain of $\underline{P}$ it holds that f is integrable with respect to μ, and

$$\int f \,\mathrm{d}\mu \geq \underline{P}(f). \tag{4.2}$$

Answer. First note that a probability measure on the algebra generated by A and B is compatible with $\underline{P}$ exactly when $\mu(A) = p$ and $\mu(B) = q$, since it must satisfy

$$\mu(A) = \int I_A \,\mathrm{d}\mu \geq \underline{P}(I_A) = p, \tag{4.3}$$

$$-\mu(A) = \int -I_A \,\mathrm{d}\mu \geq -\underline{P}(-I_A) = -p, \tag{4.4}$$

$$\mu(B) = \int I_B \,\mathrm{d}\mu \geq \underline{P}(I_B) = q, \tag{4.5}$$

$$-\mu(B) = \int -I_B \,\mathrm{d}\mu \geq -\underline{P}(-I_B) = -q. \tag{4.6}$$

Since A and B are logically independent, any probability measure on the algebra generated by A and B can be parameterised by the values of $\mu(A \cap B)$, $\mu(A \cap B^c)$ and $\mu(A^c \cap B)$. From the properties of probability measures, it follows that these parameters must be non-negative, and their sum must be less than or equal to one. This gives us a three-dimensional parametrisation of the set of probability measures defined on the algebra generated by A and B.

Probability measures defined on the algebra generated by A and B that are compatible with $\underline{P}$ additionally satisfy $\mu(A) = p$ and $\mu(B) = q$, or, in terms of our parametrisation,

$$\mu(A \cap B) + \mu(A \cap B^c) = p, \tag{4.7}$$

$$\mu(A \cap B) + \mu(A^c \cap B) = q. \tag{4.8}$$

Therefore, any such measure is uniquely characterised by the value of $\mu(A \cap B)$ only. From the properties of probability measures and the equalities above, it follows that $\mu(A \cap B)$ must satisfy the following inequalities:

$$\mu(A \cap B) \geq 0, \tag{4.9}$$

$$p - \mu(A \cap B) \geq 0, \tag{4.10}$$

$$q - \mu(A \cap B) \geq 0, \tag{4.11}$$

$$p + q - \mu(A \cap B) \leq 1. \tag{4.12}$$

This set of inequalities simply reduces to,

$$\max\{0, p+q-1\} \leq \mu(A \cap B) \leq \min\{p, q\}. \tag{4.13}$$

In summary, the set of probability measures on the algebra generated by A and B which are compatible with $\underline{P}$ is given by (on its atoms)

$$\mu(A \cap B) = \alpha, \tag{4.14}$$
$$\mu(A \cap B^c) = p - \alpha, \tag{4.15}$$
$$\mu(A^c \cap B) = q - \alpha, \tag{4.16}$$
$$\mu(A^c \cap B^c) = 1 - p - q + \alpha, \tag{4.17}$$

where α may vary between $\max\{0, p+q-1\}$ and $\min\{p, q\}$. □

(b) Preparatory exercise (continued). Again consider the case $\underline{P}(A) = \overline{P}(A) = p \in [0,1]$ and $\underline{P}(B) = \overline{P}(B) = q \in [0,1]$. Using the one-dimensional parametrisation of the above exercise, characterise the set $\mathcal{M}(\underline{P})$ of linear previsions on $\mathcal{L}(\mathcal{X})$ that dominate $\underline{P}$. Show that $\mathcal{M}(\underline{P})$ is non-empty, and hence, that $\underline{P}$ avoids sure loss. Finally, show that $\underline{P}$ is coherent.

Answer. From the above parametrisation, we infer that

$$\begin{aligned}\mathcal{M}(\underline{P}) = \Big\{Q \in \mathcal{P}(\mathcal{X})\colon\ & Q(A \cap B) = \alpha, \\ & Q(A \cap B^c) = p - \alpha, \\ & Q(A^c \cap B) = q - \alpha, \\ & Q(A^c \cap B^c) = 1 - p - q + \alpha \\ & \text{for some } \alpha \in [\max\{0, p+q-1\}, \min\{p, q\}]\Big\}\end{aligned} \tag{4.18}$$

Since $p \in [0,1]$ and $q \in [0,1]$, it follows that $\max\{0, p+q-1\} \leq \min\{p, q\}$. Therefore $\mathcal{M}(\underline{P})$ is non-empty, and hence, $\underline{P}$ avoids sure loss. In the case under consideration, $\underline{P}$ is also self-conjugate. Avoiding sure loss and self-conjugacy are sufficient for coherence. □

(c) Now we move on to the general case. Show that $\underline{P}$ avoids sure loss if and only if $\underline{P}(A) \leq 1$, $\underline{P}(B) \leq 1$, $\overline{P}(A) \geq 0$, $\overline{P}(B) \geq 0$, $\underline{P}(A) \leq \overline{P}(A)$ and $\underline{P}(B) \leq \overline{P}(B)$.

Answer 1: Dual Approach. "only if". The inequalities are standard properties of avoiding sure loss, that is, $\underline{P}(f) \leq \sup_{x \in \mathcal{X}} f(x)$ and $\underline{P}(f) \leq \overline{P}(f)$ for any gamble f in the domain of $\underline{P}$.

"if". For this problem, a simple way of proving that $\underline{P}$ avoids sure loss consists of showing the existence of a finitely additive probability measure on $\wp(\mathcal{X})$ that dominates $\underline{P}$. First, observe that the given inequalities imply that there are real numbers p and $q \in [0,1]$ such that $\underline{P}(A) \leq p \leq \overline{P}(A)$ and $\underline{P}(B) \leq q \leq \overline{P}(B)$. The lower prevision $\underline{Q}$ defined on $\{I_A, -I_A, I_B, -I_B\}$ by $\underline{Q}(A) = \overline{Q}(A) = p$ and $\underline{Q}(B) = \overline{Q}(B) = q$ dominates $\underline{P}$ and satisfies the conditions of our preparatory exercise above. In particular, $\mathcal{M}(\underline{Q})$ is non-empty. But, since $\underline{Q}$ dominates $\underline{P}$, $\mathcal{M}(\underline{Q}) \subseteq \mathcal{M}(\underline{P})$, and hence, $\mathcal{M}(\underline{P})$ is non-empty too. We conclude that $\underline{P}$ avoids sure loss. □

Answer 2: Primal Approach (Masochist). "only if". Assume $\underline{P}$ avoids sure loss. Then the above inequalities simply follow from choosing the right coefficients in the expression for avoiding sure loss:

$$\sup[I_A - \underline{P}(A)] \geq 0, \tag{4.19}$$
$$\sup[\overline{P}(A) - I_A] \geq 0, \tag{4.20}$$
$$\sup[(I_A - \underline{P}(A)) + (\overline{P}(A) - I_A)] \geq 0 \tag{4.21}$$

(and similar expressions for the event B). The argument of each supremum is a sum of almost desirable gambles. Therefore, each supremum must be non-negative for $\underline{P}$ to avoid sure loss.

"if". Suppose $\underline{P}$ satisfies the above inequalities. We need to prove that $\underline{P}$ avoids sure loss. Does the inequality

$$\sup[\lambda_1(I_A - \underline{P}(A)) + \lambda_2(\overline{P}(A) - I_A) + \lambda_3(I_B - \underline{P}(B)) + \lambda_4(\overline{P}(B) - I_B)] \geq 0 \tag{4.22}$$

hold for all λ_1, λ_2, λ_3 and $\lambda_4 \geq 0$? Since A and B are logically independent, Eq. (4.22) is equivalent to

$$\begin{aligned}
\max\{ &- \lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) && - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B), \\
&\lambda_1 - \lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) - \lambda_2 && - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B), \\
&- \lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) && +\lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4, \\
&\lambda_1 - \lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) - \lambda_2 && +\lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4\} \geq 0
\end{aligned} \tag{4.23}$$

The given inequalities imply that there are real numbers p and $q \in [0, 1]$ such that $\underline{P}(A) \leq p \leq \overline{P}(A)$ and $\underline{P}(B) \leq q \leq \overline{P}(B)$. In particular, it holds that

$$-\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) \geq (-\lambda_1 + \lambda_2)p, \tag{4.24}$$
$$-\lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) \geq (-\lambda_3 + \lambda_4)q. \tag{4.25}$$

But, this means that Eq. (4.23) is implied by the inequality

$$\begin{aligned}
\max\{&(-\lambda_1 + \lambda_2)p && +(-\lambda_3 + \lambda_4)q, \\
&(-\lambda_1 + \lambda_2)(p - 1) && +(-\lambda_3 + \lambda_4)q, \\
&(-\lambda_1 + \lambda_2)p && +(-\lambda_3 + \lambda_4)(q - 1), \\
&(-\lambda_1 + \lambda_2)(p - 1) && +(-\lambda_3 + \lambda_4)(q - 1)\} \geq 0,
\end{aligned} \tag{4.26}$$

which trivially holds. Indeed, if $-\lambda_1 + \lambda_2 \geq 0$ and $-\lambda_3 + \lambda_4 \geq 0$ then the first expression is non-negative, if $-\lambda_1 + \lambda_2 < 0$ and $-\lambda_3 + \lambda_4 \geq 0$ then the second expression is non-negative, if $-\lambda_1 + \lambda_2 \geq 0$ and $-\lambda_3 + \lambda_4 < 0$ then the third expression is non-negative, and finally, if $-\lambda_1 + \lambda_2 < 0$ and $-\lambda_3 + \lambda_4 < 0$ then the fourth expression is non-negative. □

(d) Show that $\underline{P}$ is coherent if and only if

$$\begin{aligned}
&0 \leq \underline{P}(A) \leq \overline{P}(A) \leq 1 \quad \text{and} \\
&0 \leq \underline{P}(B) \leq \overline{P}(B) \leq 1.
\end{aligned} \tag{4.27}$$

Answer 1: Dual Approach. "only if". If $\underline{P}$ is coherent then the above inequalities follow from the properties of coherence (namely, $\underline{P}(f) \geq \inf_{x\in\mathcal{X}} f(x)$ and $\underline{P}(f) \leq \overline{P}(f)$ for any gamble f in the domain of $\underline{P}$).

"if". Suppose Eq. (4.27) is satisfied. From the preparatory exercise, it follows easily that

$$\begin{aligned}\mathcal{M}(\underline{P}) = \Big\{Q \in \mathcal{P}(\mathcal{X})\colon\ & Q(A \cap B) = \alpha,\\ & Q(A \cap B^c) = p - \alpha,\\ & Q(A^c \cap B) = q - \alpha,\\ & Q(A^c \cap B^c) = 1 - p - q + \alpha\\ & \qquad \text{for some } p \in [\underline{P}(A), \overline{P}(A)],\\ & \qquad\qquad q \in [\underline{P}(B), \overline{P}(B)],\\ & \qquad \text{and } \alpha \in [\max\{0, p+q-1\}, \min\{p, q\}]\Big\}\end{aligned} \tag{4.28}$$

Since $\underline{P}$ avoids sure loss, the natural extension $\underline{E}$ of $\underline{P}$ is the lower envelope of $\mathcal{M}(\underline{P})$. For example,

$$\underline{E}(A) = \inf_{Q \in \mathcal{M}(\underline{P})} Q(A) = \inf_{Q \in \mathcal{M}(\underline{P})} \Big(Q(A \cap B) + Q(A \cap B^c)\Big) \tag{4.29}$$

$$\begin{aligned} = \inf\Big\{\alpha + p - \alpha\colon\ & \underline{P}(A) \le p \le \overline{P}(A),\\ & \underline{P}(B) \le q \le \overline{P}(B),\\ & \max\{0, p+q-1\} \le \alpha \le \min\{p, q\}\Big\}\end{aligned} \tag{4.30}$$

$$= \underline{P}(A). \tag{4.31}$$

In a similar way, it is easily shown that $\overline{E}(A) = \overline{P}(A)$, $\underline{E}(B) = \underline{P}(B)$ and $\overline{E}(B) = \overline{P}(B)$. Hence, $\underline{P}$ is coherent. □

Answer 2: Primal Approach (Masochist). "only if". Assume $\underline{P}$ is coherent. Then the above inequalities simply follow from choosing the right coefficients in the (primal) expression for coherence.

"if". To prove that $\underline{P}$ is coherent, we again need to check Eq. (4.22), and the inequalities

$$\sup[-\lambda_1(I_A - \underline{P}(A)) + \lambda_2(\overline{P}(A) - I_A) + \lambda_3(I_B - \underline{P}(B)) + \lambda_4(\overline{P}(B) - I_B)] \ge 0 \tag{4.32}$$

$$\sup[\lambda_1(I_A - \underline{P}(A)) - \lambda_2(\overline{P}(A) - I_A) + \lambda_3(I_B - \underline{P}(B)) + \lambda_4(\overline{P}(B) - I_B)] \ge 0 \tag{4.33}$$

$$\sup[\lambda_1(I_A - \underline{P}(A)) + \lambda_2(\overline{P}(A) - I_A) - \lambda_3(I_B - \underline{P}(B)) + \lambda_4(\overline{P}(B) - I_B)] \ge 0 \tag{4.34}$$

$$\sup[\lambda_1(I_A - \underline{P}(A)) + \lambda_2(\overline{P}(A) - I_A) + \lambda_3(I_B - \underline{P}(B)) - \lambda_4(\overline{P}(B) - I_B)] \ge 0, \tag{4.35}$$

for all λ_1, λ_2, λ_3 and $\lambda_4 \ge 0$. Eq. (4.22) was established above. Since A and B are logically independent, Eq. (4.32) is established if we can show that

$$\begin{aligned}\max\{&\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) && - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B),\\ -\lambda_1 + &\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) - \lambda_2 && - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B),\\ &\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) && + \lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4,\\ -\lambda_1 + &\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) - \lambda_2 && + \lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4\} \ge 0,\end{aligned} \tag{4.36}$$

From Eq. (4.27) we easily see that $\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A)$ is non-negative. So, Eq. (4.36) is implied by

$$\begin{aligned}\max\{ & -\lambda_3\underline{P}(B) + \lambda_4\overline{P}(B),\\ & +\lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4\} \ge 0.\end{aligned} \tag{4.37}$$

Now, again replace $\underline{P}(B)$ and $\overline{P}(B)$ with a $q \in [0,1]$ such that $\underline{P}(B) \le q \le \overline{P}(B)$ (which is possible by Eq. (4.27)) to see that this maximum is non-negative.

Regarding Eq. (4.33), we need to show that

$$\begin{aligned}
\max\{ &-\lambda_1\underline{P}(A) - \lambda_2\overline{P}(A) && - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B), \\
&\lambda_1 - \lambda_1\underline{P}(A) - \lambda_2\overline{P}(A) + \lambda_2 && - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B), \\
&-\lambda_1\underline{P}(A) - \lambda_2\overline{P}(A) && +\lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4, \\
&\lambda_1 - \lambda_1\underline{P}(A) - \lambda_2\overline{P}(A) + \lambda_2 && +\lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4\} \ge 0,
\end{aligned} \tag{4.38}$$

From Eq. (4.27) we easily see that $\lambda_1 - \lambda_1\underline{P}(A) - \lambda_2\overline{P}(A) + \lambda_2$ is non-negative. So, Eq. (4.38) is implied by

$$\begin{aligned}
\max\{ &-\lambda_3\underline{P}(B) + \lambda_4\overline{P}(B), \\
&+\lambda_3 - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B) - \lambda_4\} \ge 0,
\end{aligned} \tag{4.39}$$

which was already established above.

Eqs. (4.34) and (4.35) can be proven in a similar way by reversing the roles of A and B. □

(e) Assume $\underline{P}$ satisfies Eq. (4.27). Derive the Fréchet bounds by showing that the natural extension $\underline{E}$ of $\underline{P}$ satisfies

$$\begin{aligned}
\underline{E}(A \cap B) &= \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}, & \underline{E}(A \cup B) &= \max\{\underline{P}(A), \underline{P}(B)\}, \\
\overline{E}(A \cap B) &= \min\{\overline{P}(A), \overline{P}(B)\}, & \overline{E}(A \cup B) &= \min\{\overline{P}(A) + \overline{P}(B), 1\}.
\end{aligned} \tag{4.40}$$

Answer 1: Dual Approach. Since $\underline{P}$ avoids sure loss, the natural extension $\underline{E}$ of $\underline{P}$ is the lower envelope of $\mathcal{M}(\underline{P})$. Let's again use Eq. (4.28):

$$\underline{E}(A \cap B) = \inf_{Q \in \mathcal{M}(\underline{P})} Q(A \cap B) \tag{4.41}$$

$$\begin{aligned}
= \inf\Big\{\alpha \colon &\underline{P}(A) \le p \le \overline{P}(A), \\
&\underline{P}(B) \le q \le \overline{P}(B), \\
&\max\{0, p+q-1\} \le \alpha \le \min\{p, q\}\Big\}
\end{aligned} \tag{4.42}$$

$$= \max\{0, \underline{P}(A) + \underline{P}(B) - 1\}, \tag{4.43}$$

and similarly, $\overline{E}(A \cap B) = \min\{\overline{P}(A), \overline{P}(B)\}$. For the union, we have

$$\underline{E}(A \cup B) = \inf_{Q \in \mathcal{M}(\underline{P})} Q(A \cup B) = \inf_{Q \in \mathcal{M}(\underline{P})} \big(1 - Q(A^c \cap B^c)\big) \tag{4.44}$$

$$\begin{aligned}
= \inf\Big\{p + q - \alpha \colon &\underline{P}(A) \le p \le \overline{P}(A), \\
&\underline{P}(B) \le q \le \overline{P}(B), \\
&\max\{0, p+q-1\} \le \alpha \le \min\{p, q\}\Big\}
\end{aligned} \tag{4.45}$$

$$= \underline{P}(A) + \underline{P}(B) - \min\{\underline{P}(A), \underline{P}(B)\} = \max\{\underline{P}(A), \underline{P}(B)\}, \tag{4.46}$$

and similarly,

$$\overline{E}(A \cup B) = \overline{P}(A) + \overline{P}(B) - \max\{0, \overline{P}(A) + \overline{P}(B) - 1\} \tag{4.47}$$

$$= \min\{\overline{P}(A) + \overline{P}(B), 1\}. \tag{4.48}$$

□

Answer 2: Primal Approach (Masochist). We must find the solution of the linear program

$$\underline{E}(A \cap B) = \sup\Big\{\gamma \in \mathbb{R} \colon \Big(\exists \lambda_1, \dots, \lambda_4 \geq 0\Big) \\ \Big(I_{A\cap B} - \gamma \geq \lambda_1(I_A - \underline{P}(A)) + \lambda_2(\overline{P}(A) - I_A) + \lambda_3(I_B - \underline{P}(B)) + \lambda_4(\overline{P}(B) - I_B)\Big)\Big\}. \tag{4.49}$$

Since A and B are logically independent, the constraints of the linear program are equivalent to

$$\begin{aligned} 1 - \gamma &\geq \lambda_1(1 - \underline{P}(A)) + \lambda_2(\overline{P}(A) - 1) + \lambda_3(1 - \underline{P}(B)) + \lambda_4(\overline{P}(B) - 1), \\ -\gamma &\geq -\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) + \lambda_3(1 - \underline{P}(B)) + \lambda_4(\overline{P}(B) - 1), \\ -\gamma &\geq \lambda_1(1 - \underline{P}(A)) + \lambda_2(\overline{P}(A) - 1) - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B), \\ -\gamma &\geq -\lambda_1\underline{P}(A) + \lambda_2\overline{P}(A) - \lambda_3\underline{P}(B) + \lambda_4\overline{P}(B). \end{aligned} \tag{4.50}$$

First, observe that $\gamma = \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}$ is a feasible solution of the linear program, by taking $\lambda_2 = \lambda_4 = 0$ and $\lambda_1 = \lambda_3 = 1$ if $\underline{P}(A) + \underline{P}(B) \geq 1$, and $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ otherwise. Indeed, for $\lambda_2 = \lambda_4 = 0$ and $\lambda_1 = \lambda_3 = 1$ Eq. (4.50) becomes

$$\begin{aligned} 1 - \gamma &\geq 1 - \underline{P}(A) + 1 - \underline{P}(B), \\ -\gamma &\geq -\underline{P}(A) + 1 - \underline{P}(B), \\ -\gamma &\geq 1 - \underline{P}(A) - \underline{P}(B), \\ -\gamma &\geq -\underline{P}(A) - \underline{P}(B). \end{aligned} \tag{4.51}$$

These are satisfied for $\gamma = \underline{P}(A) + \underline{P}(B) - 1$. If $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_4 = 0$ Eq. (4.50) is trivially satisfied for $\gamma = 0$. Hence, $\underline{E}(A \cap B) \geq \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}$.

If we can show that $\gamma \leq \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}$ for any other feasible solution γ of Eq. (4.50), then we have established that $\underline{E}(A \cap B) = \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}$. Taking a convex combination with coefficients $\alpha_1, \dots, \alpha_4 \geq 0$ ($\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4 = 1$) of the constraints in Eq. (4.50), we find that

$$\begin{aligned} \alpha_1 - \gamma \geq &\lambda_1[-\underline{P}(A) + (\alpha_1 + \alpha_3)] \\ &+\lambda_2[\overline{P}(A) - (\alpha_1 + \alpha_3)] \\ &+\lambda_3[-\underline{P}(B) + (\alpha_1 + \alpha_2)] \\ &+\lambda_4[\overline{P}(B) - (\alpha_1 + \alpha_2)]. \end{aligned} \tag{4.52}$$

Take $\alpha_1 = \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}$, $\alpha_2 = \underline{P}(B) - \alpha_1$ and $\alpha_3 = \underline{P}(A) - \alpha_1$ (it is easy to check that $\alpha_1 \geq 0$, $\alpha_2 \geq 0$, $\alpha_3 \geq 0$ and $1 - \alpha_1 - \alpha_2 - \alpha_3 \geq 0$; so this is indeed a convex combination). We find that

$$\begin{aligned} \gamma &\leq \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\} - \lambda_2(\overline{P}(A) - \underline{P}(A)) - \lambda_4(\overline{P}(B) - \underline{P}(B)) \\ &\leq \max\{\underline{P}(A) + \underline{P}(B) - 1, 0\}, \end{aligned} \tag{4.53}$$

which establishes the proof. □

5 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

$$\underline{E}_0(f) = \frac{1}{3}\Big(f(a) + f(b) + f(c)\Big). \tag{5.1}$$

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

(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:

$$\underline{P}(g|X = a) = \min\{g(b), g(c)\} \tag{5.2}$$
$$\underline{P}(g|X = b) = g(c) \tag{5.3}$$
$$\underline{P}(g|X = c) = g(b) \tag{5.4}$$

for any gamble $g \in \mathcal{L}(\mathcal{Y})$. 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 $\mathcal{L}(\mathcal{X})$ and $\underline{P}(\cdot|X)$ on $\mathcal{L}(\mathcal{Y})$, using the marginal extension theorem, to a coherent lower prevision $\underline{E}$ on $\mathcal{L}(\mathcal{X} \times \mathcal{Y})$.

Answer. The marginal extension theorem tells us that $\underline{E}(h) = \underline{E}_0(\underline{P}(h|X))$ for all $h \in \mathcal{L}(\mathcal{X} \times \mathcal{Y})$. Using Eqs. (5.1) and (5.2)-(5.4), we find that

$$\underline{E}(h) = \frac{1}{3}\left(\min\{h(a, b), h(a, c)\} + h(b, c) + h(c, b)\right). \tag{5.5}$$

□

(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

$$\underline{E}(I_{\{b\}}(f - \underline{E}(f|Y = b))) = 0. \tag{5.6}$$

Using Eq. (5.5), we find

$$\frac{1}{3}\Big(\min\{f(a) - \underline{E}(f|Y = b), 0\} + 0 + f(c) - \underline{E}(f|Y = b)\Big) = 0. \tag{5.7}$$

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 (in press).)

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

$$\begin{cases} h_0(a,b) = 25,000,000 \\ h_0(a,c) = 25,000,000 \\ h_0(b,c) = 0 \\ h_0(c,b) = 0 \end{cases} \tag{5.8}$$

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

$$\begin{cases} h_1(a,b) = -25,000 \\ h_1(a,c) = -25,000 \\ h_1(b,c) = 24,975,000 \\ h_1(c,b) = 24,975,000 \end{cases} \tag{5.9}$$

After the warden has named prisoner b, the conditional lower prevision $\underline{E}(\cdot|Y = b)$ describes our buying prices regarding $\mathcal{L}(\mathcal{X} \times \mathcal{Y})$. 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

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

and,

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

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, based on conditional lower previsions, 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 (as explained in Thomas Augustin's summerschool lecture notes): 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

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

and

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

hence, we should strictly prefer h_1 over h_0: bribe the guard. □

Note. 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.

The correct answer to the Monty Hall puzzle has been the subject of many debates during the summerschool breakfasts, coffee breaks, lunches, diners and long sleepless nights. Here we have only described two possible answers. Thanks to everyone participating in these debates! □