Evaluating Imprecise Forecasts

Jason Konek

University of Bristol

EPIMP

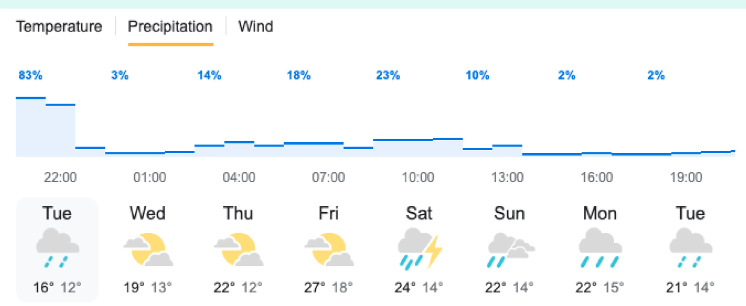
Keywords: **Scoring rules, loss functions, forecasting, lower previsions, sets of almost desirable gambles**

Background

$\Omega = \{\omega_1, \ldots, \omega_n\}$ is a **finite possibility space**.

$\mathcal{F}$ is the power set of Ω. Elements E of $\mathcal{F}$ are **events**.

Experts announce **probabilistic forecasts** for events.



Forecasts are **accurate** insofar as they are "close" to the indicators of the events being forecast.

- If it does not rain on Wednesday or Thursday, then Wednesday's forecast (0.03) is more accurate (closer to 0) than Thursday's (0.14).

$c : \mathcal{F} \to \mathbb{R}$ is an assignment of precise forecasts to events in $\mathcal{F}$. $\mathcal{C}$ is the space of all such assignments.

Evaluate assignments of precise forecasts by a **scoring rule** or **loss function** $\mathcal{I} : \mathcal{C} \times \Omega \to \mathbb{R}_{\geq 0}$.

A scoring rule $\mathcal{I} : \mathcal{C} \times \Omega \to \mathbb{R}_{\geq 0}$ is **strictly proper** iff

$$\sum_{\omega \in \Omega} p(\omega)\mathcal{I}(p, \omega) < \sum_{\omega \in \Omega} p(\omega)\mathcal{I}(c, \omega)$$

for any probability function $p \in \mathcal{C}$ and any $c \neq p$.

- **Brier Score**: $\mathcal{I}(c, \omega) = \sum_{X \in \mathcal{F}} (\mathbb{1}_X(\omega) - c(X))^2$
- **Log Score**: $\mathcal{I}(c, \omega) = \sum_{X \in \mathcal{F}} -log(|1 - \mathbb{1}_X(\omega) - c(X)|)$
- **Spherical Score**:

$$\mathcal{I}(c, \omega) = \sum_{X \in \mathcal{F}} \left(1 - \frac{|1 - \mathbb{1}_X(\omega) - c(X)|}{\sqrt{c(X)^2 + (1 - c(X))^2}}\right)$$

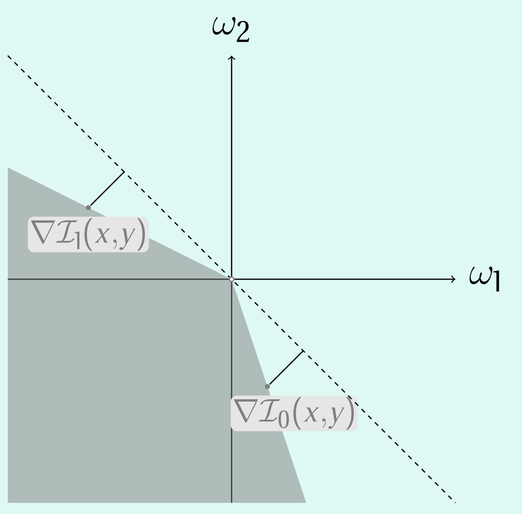
Strictly proper scoring rules: admissibility

An assignment of forecasts $c : \mathcal{F} \to \mathbb{R}$ is **admissible** relative to a scoring rule $\mathcal{I}$ if and only if it is not **incoherent**$_2$ [de Finetti, 1974, ch. 3], *i.e.*, it is not (uniformly) dominated by some $b \neq c$ in the sense that

$$\mathcal{I}(b, \omega) < \mathcal{I}(c, \omega)$$

for all $\omega \in \Omega$.

Let x be a precise forecast for event E and y be a precise forecast for $\neg E$. The following is a straightforward consequence of [Lindley, 1982, Lemma 2]:



Corollary

If $\mathcal{I}_0(x, y) = s_0(x) + s_1(y)$ and $\mathcal{I}_1(x, y) = s_1(x) + s_0(y)$ is a continuously differentiable strictly proper scoring rule, then following three conditions are equivalent:

- *There are $a, b \in \mathbb{R}$ s.t.*

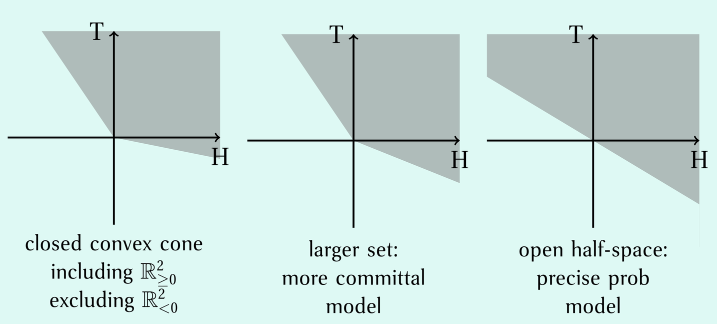
$$\nabla_{\langle a,b \rangle}\mathcal{I}_0(x, y) < 0$$
$$\nabla_{\langle a,b \rangle}\mathcal{I}_1(x, y) < 0$$

- $0 \notin \text{posi}\left(\{\nabla\mathcal{I}_0(x, y), \nabla\mathcal{I}_1(x, y)\}\right)$
- $y \neq 1 - x$

Conclusion: A pair of forecasts, x and y, for E and $\neg E$ respectively, are admissible if and only if probabilistic.

Sets of almost desirable gambles

A gamble $g : \Omega \to \mathbb{R}$ is an uncertain reward which pays out in linear utility. We will treat them as elements $g = \langle g_1, \ldots, g_n \rangle$ of $\mathbb{R}^n$.



A set $\mathcal{D} \subseteq \mathbb{R}^n$ is a **coherent** set of almost desirable gambles iff it satisfies:

- AD1. If $g < 0$ then $g \notin \mathcal{D}$ (where $g < 0 \Leftrightarrow g_i < 0$ for all $i \leq n$)
- AD2. If $g \geq 0$ then $g \in \mathcal{D}$ (where $g \geq 0 \Leftrightarrow g_i \geq 0$ for all $i \leq n$)
- AD3. If $g \in \mathcal{D}$ and $\lambda > 0$ then $\lambda g \in \mathcal{D}$
- AD4. If $f, g \in \mathcal{D}$ then $f + g \in \mathcal{D}$
- AD5. If $g + \epsilon \in \mathcal{D}$ for all $\epsilon > 0$ then $g \in \mathcal{D}$

The **epigraph** of a function $b : \mathbb{R}^{n-1} \to [-\infty, \infty]$ is

$$\mathcal{D}_b = \{\langle g_1, \ldots, g_n \rangle \,|g_n \geq b(g_1, \ldots, g_{n-1})\} \subseteq \mathbb{R}^n$$

- Every coherent set of almost-desirable gambles $\mathcal{D}$ is epigraphical.
- Many epigraphical sets of almost desirable gambles are not coherent.

Challenges

- Suppose that for all $i \leq n$, ϕ_i satisfies P1, P2 and super-additivity: $\phi_i(f + g) \geq \phi_i(f) + \phi_i(g)$, for all $f, g \in \mathbb{R}^n$.
  - Super-additivity is useful for ensuring that admissible $\mathcal{D}$ satisfy AD4.
- **Triviality Result (Van Camp):** ϕ_i satisfies P1, P2 and super-additivity for all $i \leq n$ if and only if ϕ_i is *represented* by a function $\gamma_i : \mathbb{R} \to \mathbb{R}$ that satisfies Properties P1, P2 and super-additivity, in the sense that $\phi_i(g) = \gamma_i(g_i)$ for every $g \in \mathbb{R}^n$.
- Only sufficient to render a slightly generalised class of $\mathcal{D}_f$ from example 2 admissible.

References

Bruno de Finetti. *Theory of Probability: A Critical Introductory Treatment*, volume 1. John Wiley & Sons, Chichester, 1974. English translation of De Finetti (1970).

D. V. Lindley. Scoring rules and the inevitability of probability. *International Statistical Review*, 50:1–26, 1982.

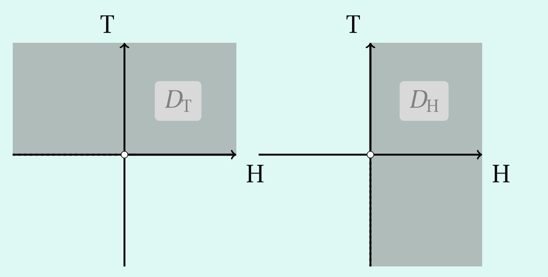
Mark J. Schervish, Teddy Seidenfeld, and Joseph B. Kadane. Infinite previsions and finitely additive expectations. *arXiv:1308.6761*, 2013.

Scoring imprecise forecasts

The **ideal set of almost desirable gambles** if ω_i is the true state of the world is given by

$$\mathcal{D}_i = \{g | g_i \geq 0\} \subseteq \mathbb{R}^n$$

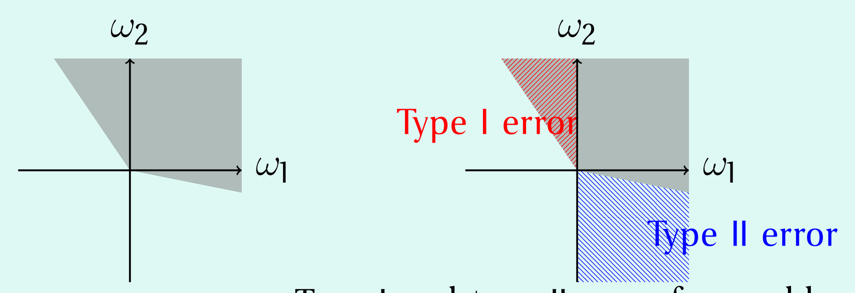
$\mathcal{D}_i$ contains all and only the gambles that are *in fact* almost desirable at ω_i.



Choose an epigraphical set $\mathcal{D} \subseteq \mathbb{R}^n$ of almost desirable gambles (coherent or not).

$\mathcal{E}_i^1 = \mathcal{D} \setminus \mathcal{D}_i$ is $\mathcal{D}$'s set of **type 1** errors at ω_i.

$\mathcal{E}_i^2 = \mathcal{D}_i \setminus \mathcal{D}$ is $\mathcal{D}$'s set of **type 2 errors** at ω_i.



Type I and type II error for world ω_1

$\mathcal{E}_i = \mathcal{E}_i^1 \bigcup \mathcal{E}_i^2$ is $\mathcal{D}$'s **error set** at ω_i—the total set of gambles that $\mathcal{D}$ mischaracterizes at ω_i.

Inaccuracy is a measure of error.

The inaccuracy of $\mathcal{D}$ at ω_i, $\mathcal{I}(\mathcal{D}, \omega_i)$, is the measure of $\mathcal{E}_i$ according to an appropriate measure ν_i:

$$\mathcal{I}(\mathcal{D}, \omega_i) = \mathcal{I}_i(\mathcal{D}) = \nu_i(\mathcal{E}_i)$$

- $\nu_i(\mathcal{E}_i)$ is something like the "size" of the error set $\mathcal{E}_i$.

Assume that ν_i is finite and absolutely continuous with respect to the product Lebesgue measure μ. In that case

$$\mathcal{I}_i(\mathcal{D}) = \int_{\mathcal{E}_i} |\phi_i| \, d\mu$$

Axiomatic constraints:

- P1. $\phi_i(g_1, \ldots, g_n)$ is (at least weakly) increasing in g_i
- P2. $\phi_i(g_1, \ldots, g_{i-1}, 0, g_{i+1}, \ldots, g_n) = 0$

- Accepting a bigger loss is a bigger type 1 error
- Leaving more utility on the table is a bigger type 2 error.

Linear previsions and non-additivity

Precision-inducing constraints:

- SP1. $\phi_i(\lambda g) = \lambda\phi_i(g)$ for any $\lambda > 0$
- SP2. $\nu_i(\mathcal{E}_i) = \nu_i(\mathcal{E}_i^*)$ for any $\mathcal{E}_i^*$ s.t.

$$\mathcal{E}_i^* = \{\langle x_1, \ldots, x_{i-i}, g_i, x_{i+1}, \ldots, x_n \rangle \,|g \in \mathcal{E}_i, x_1, \ldots, x_n \in \mathbb{R}\}$$

- SP3. $\nu_i(\mathcal{E}_i) = \nu_j(\mathcal{E}_j^\dagger)$ where $\mathcal{E}_j^\dagger$ is the result of permuting the i^{th} and j^{th} component of any $g \in \mathcal{E}_i$, *i.e.*,

$$\mathcal{E}_j^\dagger = \left\{g^\dagger | g \in \mathcal{E}_i, g_i^\dagger = g_j, g_j^\dagger = g_i, g_k^\dagger = g_k \textit{ for all } k \neq i, j\right\}$$

Theorem

If $\mathcal{I}$ satisfies P1-P2 and SP1-SP3, then there is some $c > 0$ such that for all $i \leq n$

$$\mathcal{I}_i(\mathcal{D}) = \int_{\mathcal{E}_i} |cg_i| \, d\mu$$

In that case, for any probability mass function $p : \Omega \to \mathbb{R}$ and any $\mathcal{D} \neq \mathcal{D}_p$

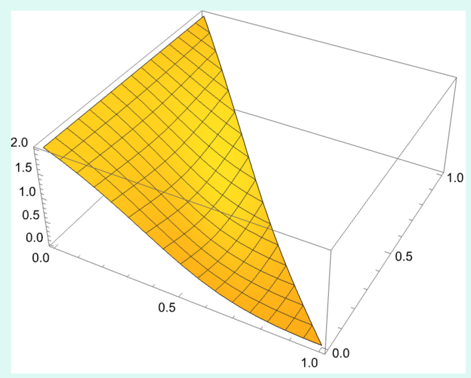
$$\sum_{i \leq n} p_i\mathcal{I}_i(\mathcal{D}_p) < \sum_{i \leq n} p_i\mathcal{I}_i(\mathcal{D})$$

unless both $\mathcal{D} \setminus \mathcal{D}_p$ and $\mathcal{D}_p \setminus \mathcal{D}$ are sets of measure zero.

Example 1. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and let $\mathcal{P}$ be the set of all probability mass functions of Ω. Choose $p = \langle p_1, p_2, p_3 \rangle \in \mathcal{P}$. Let ρ be the normal distribution on the Borel σ-algebra $\mathfrak{B}(\mathbb{R})$ with mean 0 and standard deviation 5. Let μ be the product measure $\rho \times \rho \times \rho$ on $\mathfrak{B}(\mathbb{R}^3)$. In that case

$$\mathcal{I}_i(\mathcal{D}_p) = \frac{5\left(1 - \frac{p_i}{\sqrt{p_1^2 + p_2^2 + p_3^2}}\right)}{2\pi}$$

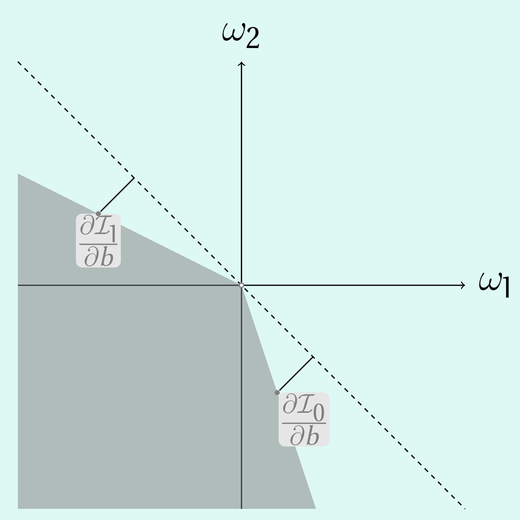
This is a non-additive analogue of the Spherical score.



$\mathcal{I}_1(\mathcal{D}_p)$ as a function of p_1 (x-axis) and p_2 (y-axis).

An alternative to the strictly proper additive scoring rules for linear previsions considered by Schervish et al. [2013].

IP scoring rules: admissibility



Theorem

If ν_i is finite and absolutely continuous with respect to μ, for all $i \leq n$, then the following two conditions are equivalent:

- *There is some $h : \mathbb{R}^{n-1} \to \mathbb{R}$ s.t. for all $i \leq n$*

$$\delta\mathcal{I}_i(b, h) = \int_{\mathbb{R}^{n-1}} \frac{\partial\mathcal{I}_i}{\partial b} h \, d\lambda = \lim_{\epsilon \to 0} \frac{1}{\epsilon}[\mathcal{I}_i(b + \epsilon h) - \mathcal{I}_i(b)] < 0$$

  First variation—calculus of variations analogue of directional derivative

- $0 \notin \text{posi}\left(\{\phi_i(\cdot, b(\cdot)) \,|\, i \leq n\}\right)$

Example 2. Suppose that $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and that for some $\lambda \geq \gamma > 0$:

$$\phi_i(g_1, g_2, g_3) = \begin{cases} \lambda g_i & \text{if } g_i < 0 \\ \gamma g_i & \text{if } g_i \geq 0 \end{cases}$$

Then $0 \in \text{posi}\left(\{\phi_i(\cdot, b(\cdot)) \,|\, i \leq 3\}\right)$ iff there are $\alpha, \beta \geq 0$ s.t.

$$b(g_1, g_2) = \begin{cases} \frac{-\gamma(\alpha g_1 + \beta g_2)}{\lambda} & \text{if } g_1 \geq 0, g_2 \geq 0 \\ \frac{-\lambda(\alpha g_1 + \beta g_2)}{\gamma} & \text{if } g_1 < 0, g_2 < 0 \\ \frac{-(\alpha\lambda g_1 + \beta\gamma g_2)}{\gamma} & \text{if } g_1 < 0, g_2 \geq 0, \alpha\lambda g_1 + \beta\gamma g_2 < 0 \\ \frac{-(\alpha\lambda g_1 + \beta\gamma g_2)}{\lambda} & \text{if } g_1 < 0, g_2 \geq 0, \alpha\lambda g_1 + \beta\gamma g_2 \geq 0 \\ \frac{-(\alpha\gamma g_1 + \beta\lambda g_2)}{\gamma} & \text{if } g_1 \geq 0, g_2 < 0, \alpha\gamma g_1 + \beta\lambda g_2 < 0 \\ \frac{-(\alpha\gamma g_1 + \beta\lambda g_2)}{\lambda} & \text{otherwise} \end{cases}$$

It is easy to verify that $\mathcal{D}_b$ is coherent. Only coherent $\mathcal{D}_b$ of this form are admissible relative to $\mathcal{I}$.

University of Bristol

erc