# **Evaluating Imprecise Forecasts**

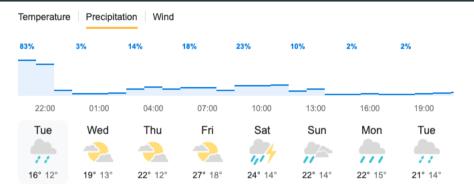
Jason Konek

11 July 2023

Department of Philosophy, Bristol ISIPTA 2023, Oviedo

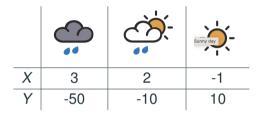
# **Motivation**

## **Probabilistic Forecasting**



Penalise "inaccurate forecasts" using loss functions

- Ex Ante: incentivise "careful" and "honest" forecasts
- Ex Post: sift accurate/inaccurate forecasters; how to improve



Is X desirable? Is Y desirable?

Ex Ante Ex Post

## **Ex Ante: Prediction Markets**



Will Russia and Ukraine sign or announce an agreement to end the current conflict in Ukraine?

A Before 1 June 2022	1%	<b>#</b> 1
Between 1 June 2022 and 31 July 2022	4%	<b>#</b> 1
Between 1 August 2022 and 30 September 2022	7%	-
Between 1 October 2022 and 30 November 2022	9%	-
E Not before 1 December 2022	79%	<b>e</b> 2

Probabilities as of 11 May 2022

#### When will Russia and Ukraine sign or announce an agreement to end the current conflict in Ukraine?

Opened: 14 Oct 2022, Suspends 30 Sep 2024	Current 1-Week Forecast Change
A Before 1 April 2023	1% 🖶 1
B Between 1 April 2023 and 30 September 2023	7% —
C Between 1 October 2023 and 31 March 2024	13% 🖶 1
D Between 1 April 2024 and 30 September 2024	18% —
E Not before 1 October 2024	61% 🔶 2
Probabilities as of 6 January 2023	Implied Median: 18 Dec 2024

 Traders can change market IP forecasts; pay score of status quo forecast in exchange for score of updated forecast

 Incentivise reporting of "best" IP model from class of admissible models

Source: Good Judgment Inc

#### **Ex Post: Machine Learning**



Ankle boot





Ankle boot

Trouser

Pullover



T-shirt/top



Ankle boot







Dress



Pullover

T-shirt/top



Sandal



T-shirt/top





Sandal

Sandal

Shirt

Bag





Sandal



Sneaker



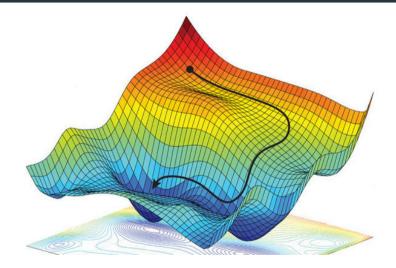




Coat

Label Description T-shirt/top 0 Trouser Pullover 2 3 Dress Coat 4 5 Sandal Shirt 6 7 Sneaker 8 Bag 9 Ankle boot

## **Ex Post: Machine Learning**



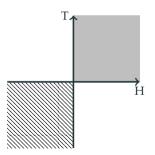
# **IP Loss Functions**

Let  $\Omega$  be a finite possibility space.

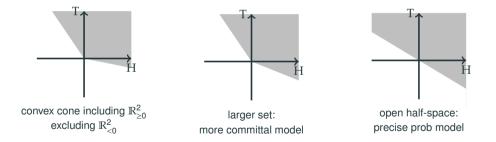
A gamble  $X : \Omega \to \mathbb{R}$  is an uncertain reward. We will treat them as elements  $X = \langle x_1, \dots, x_n \rangle$  of  $\mathbb{R}^n$ .

A set  $\mathcal{D} \subseteq \mathbb{R}^n$  is a **coherent set of almost desirable gambles** if and only if it satisfies the following five axioms:

AD1. If X < 0 then  $X \notin \mathcal{D}$  (where  $X < 0 \Leftrightarrow x_i < 0$  for all  $i \le n$ ) AD2. If  $X \ge 0$  then  $X \in \mathcal{D}$  (where  $X \ge 0 \Leftrightarrow x_i \ge 0$  for all  $i \le n$ )



AD3. If  $X \in \mathcal{D}$  and  $\lambda > 0$  then  $\lambda X \in \mathcal{D}$ AD4. If  $X, Y \in \mathcal{D}$  then  $X + Y \in \mathcal{D}$ AD5. If  $X + \epsilon \in \mathcal{D}$  for all  $\epsilon > 0$  then  $X \in \mathcal{D}$ 

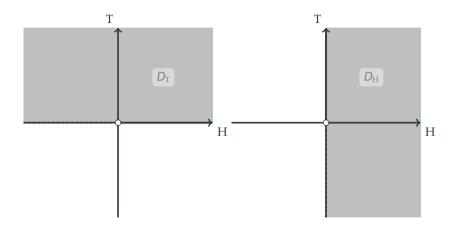


#### **Ideal Sets of Gambles**

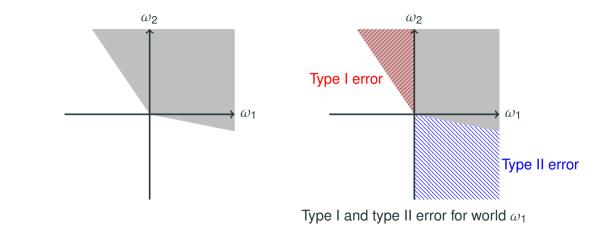
The ideal set of almost desirable gambles if  $\omega_i$  is true is given by

 $\mathcal{D}_i = \{X \mid x_i \ge 0\}$ 

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



## Type 1 and 2 Error



#### $\mathcal{E}_i = \mathbf{T1} \bigcup \mathbf{T2}$ is $\mathcal{D}$ 's error set at $\omega_i$

• The set of gambles that  $\mathcal{D}$  mischaracterizes at  $\omega_i$ .

Inaccuracy is a measure of error.

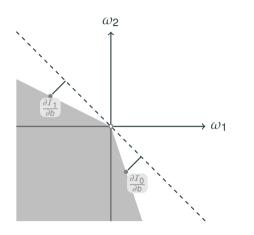
The inaccuracy of  $\mathcal{D}$  at  $\omega_i$ ,  $\mathcal{I}(\mathcal{D}, \omega_i)$ , is the measure of  $\mathcal{E}_i$  according to an appropriate measure  $v_i$ :

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

Assume that  $v_i$  is finite and absolutely continuous with respect to the product Lebesgue measure  $\mu$ . In that case

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

#### **Main Result**



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

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

$$\delta \mathcal{I}_{i}(b,h) = \int_{\mathbb{R}^{n-1}} \frac{\partial \mathcal{I}_{i}}{\partial b} h \, \mathrm{d}\lambda < 0$$

First variation—calculus of variations analogue of directional derivative

2. 
$$0 \notin \text{posi}\left(\left\{\phi_i(\cdot, b(\cdot)) \mid i \leq n\right\}\right)$$

Suppose that for all  $i \leq n$ 

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

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

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

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

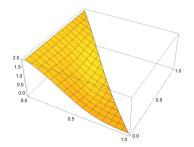
I is a strictly proper scoring rule.

#### Loss functions for (infinitely many) precise previsions

**Example.** Let  $\rho$  be the normal distribution on the Borel  $\sigma$ -algebra  $\mathfrak{B}(\mathbb{R})$  with mean 0 and standard deviation 5. Let  $\mu$  be the product measure  $\rho \times \rho \times \rho$  on  $\mathfrak{B}(\mathbb{R}^3)$ . In that case

$${\mathcal I}_i({\mathcal D}_{
ho}) = rac{5 \left(1 - rac{p_i}{\sqrt{p_1^2 + p_2^2 + p_3^2}}
ight)}{2\pi}$$

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



#### Loss functions for non-maximal sets of almost desirable gambles

Suppose that there are  $\lambda \ge \gamma > 0$  such that

$$I_{i}(\mathcal{D}) = \int_{\mathcal{E}_{i}} \begin{cases} -\lambda x_{i} & \text{if } x_{i} < 0 \\ \gamma x_{i} & \text{if } x_{i} \geq 0 \end{cases} d\mu$$

for all  $i \leq n$ . Then  $\mathcal{D}_b = \{\langle g_1, g_2, g_3 \rangle | g_3 \geq b(g_1, g_2)\} \subseteq \mathbb{R}^3$  is admissible, *i.e.*, not dominated by some  $\mathcal{D} \neq \mathcal{D}_b$ , iff there are  $\alpha, \beta \geq 0$  s.t.

$$b(x_1, x_2) = \begin{cases} \frac{-\gamma(\alpha x_1 + \beta x_2)}{\lambda} & \text{if } x_1 \ge 0, x_2 \ge 0\\ \frac{-\lambda(\alpha x_1 + \beta x_2)}{\gamma} & \text{if } x_1 < 0, x_2 < 0\\ \frac{-(\alpha \lambda x_1 + \beta \gamma x_2)}{\gamma} & \text{if } x_1 < 0, x_2 \ge 0, \alpha \lambda x_1 + \beta \gamma x_2 < 0\\ \frac{-(\alpha \lambda x_1 + \beta \gamma x_2)}{\lambda} & \text{if } x_1 < 0, x_2 \ge 0, \alpha \lambda x_1 + \beta \gamma x_2 \ge 0\\ \frac{-(\alpha \gamma x_1 + \beta \lambda x_2)}{\gamma} & \text{if } x_1 \ge 0, x_2 < 0, \alpha \gamma x_1 + \beta \lambda x_2 < 0\\ \frac{-(\alpha \gamma x_1 + \beta \lambda x_2)}{\lambda} & \text{otherwise} \end{cases}$$

0

0

## Loss functions for non-maximal sets of almost desirable gambles

#### Come to the poster!

#### **Evaluating Imprecise Forecasts** lason Konek Exporte: Scoring rules, loss functions, forecasting, lower previsions, sets of almost desirable gambles Background Scoring imprecise forecasts And the second data and the second The black set of almost desirable gambles $T \sim h$ the $-I_0 - I_1^2 \cup I_2^2 h$ (7) error set at $c_1$ —the ball set of angeneres. Italiair anigeraris of precise hereasis by a sender of the bar bar bar of the sender of t Py = LEW 2 Py 5 W. Ty contains all and only the gambles that are to fact dense describe at 15 Taperis amounter probabilitate herecasis for A scoring rule Z : C = D == R. .... is strictly proper $\sum p(\alpha | f(p, \omega) < \sum p(\alpha | f(x, \omega))$ 100 To any probability function $p \in e$ and any $e \neq p$ , in Belier Score: $T(e, p) = \sum {(i_1 j_2 p) - i_1^2 k_2^2}$ Once an extended at 2 C F<sup>2</sup> of draw decide. Assure that to b forty and decideds continuous with $\label{eq:log_state} \begin{array}{l} & \log \ \mathrm{house} \\ T(r, s) = \sum_{l \geq 0} - \log (3 - 1_{ll}(s) - r(b)) \end{array}$ $\mathcal{J}_1^2 = \mathcal{D} \setminus \mathcal{D}_1$ in $\mathcal{D}_2$ not of type 1 senses at $\omega_2$ $T_{n}(P) = \int_{0}^{1} |\psi_{n}| dx$ Adomatic constraints: $T_{1}(z_{0}) = \sum_{t \in U} \left(1 - \frac{3 - 1_{t}(z_{0}) - s(t)}{\sqrt{s(t)^{2} + 3 - s(t)}}\right)$ Tate I ave And a state of the \* Accepting a bigger lines is a bigger tope I error Strictly proper scoring rules: admissibility Sectors. Accepting a bigger loss is a bigger type 2 area Leaving more atility on the table is a bigger type 2 area Read and here & some her world as As an engenerated the bore set $x : F \rightarrow \mathbb{R}$ is a shadowiddly relative to a scoring rule $\mathbb{T}$ if and only if it is real-modulations: the Proofs 2015 ob. If (x, y) is not indicated downloaded by some $h \leq z$ in the same the Linear previsions and non-additivity for all $\omega \in \Omega$ . Let a be a produc forward for event E and $\mu$ be a produc forward for -E. The following is a straightforward conservation of Diracley Mill Lemma 10. Example 1. Let $\hat{U} = \{\omega_1, \omega_2, \omega_3\}$ and let P be the set of all probability mass functions of U. Conver, $\mu = (\mu, \mu_2, \mu_3) \in P$ . Let $\mu$ be the mesonic distributions on $\mu_1 = (\mu_1, \mu_2, \mu_3) \in P$ . Caradian $\{T_{2j}(n, z) = u_{2}(z) + u_{2}(z) \text{ and } \tilde{x}_{2}(x, z) = u_{2}(z) + u_{2}(z)$ is a constraintially differentiable stated proper strategy rate. Here following draw consistence are equivalence. B. Nove we at $b \in \mathbb{R}$ so d. (27) $r_1(\ell_1) = r_2(\ell_1^2)$ where $\ell_1^2$ is the rands of permuting $I_{\text{AUV}} = \frac{8\left(1 - \frac{4}{\sqrt{6(-6) + 4}}\right)}{1}$ This is a non-solding analogue of the Spherical Barren Emphasism 4 and all increases a soil a fee 5 and -5 mercelarity are adminished and and 4 zzri - Limbia in that case, for any probability mass function $p: \mathbb{R} \to \mathbb{R} \text{ and any } \mathbb{P} \neq \mathbb{D}_{\mathbf{Y}}$ Sets of almost desirable gambles A purely $g : E \rightarrow \mathbb{R}$ is an uncertain reward which peak out is linear utility. We will trust from as observes $g = \langle g_1, ..., g_N \rangle$ of $\mathbb{R}^n$ . fulful as a function of a to axid and as to axid. unitor both D' D, and D. D are arts of measure An alternative to the strictly proper additive scoting rules of the strictly proper additive scoting rules 1 м. A set $\mathcal{D} \subseteq \mathbb{R}^n$ is a coherent set of above description and the set of a set of equilibrium of a set of equilibrium. And the part of th IP scoring rules: admissibility (c) if g < 0 then $g \notin D$ takens g < 0 or $g_i < 0$ for. The **replaceph** of a function $b : \mathbb{R}^{n-1} \to [-\infty, \infty]$ of J < 0. $\lim_{d \neq i \neq m} F_{d} \ge 0 \text{ from } g \ge 0 \text{ solver } g \ge 0 \text{ so } g_i \ge 0 \text{ for } D_k = \{0, \dots, d_i \mid i_1 \ge 0, \dots, d_{i-1}\} \le \mathbb{R}^k$ a fram obsent at of about durable particles Every otherwet net of abnum-destrable gand Ω is oppropriated. Many systemptical sets of abnust devirable gandutes are not roboverst. $\delta \Sigma_i(0, 0) = \int \frac{\partial U_i}{\partial x} \delta dx = 3m_i^{-1} \left[ \Sigma_i(0+iA) - \Sigma_i(0) \right] < 0$ Jg=12b +14+ Piew variation-valuable of variations analogue of directional derivative Challenges = Suppose that the rid $|| \le n$ , $n_j$ and the P( PI and super additivity $\partial_i |I + g| \ge d_i |I|$ for all $\int_{-\infty}^{1} \frac{d_i |I|}{d_i |I|} = \frac{d_i |I|}{d_$ Example 3. Suppose that $\Omega = l_{12}, q_{2}, q_{3}$ and that for some $\lambda \ge q > 0$ . $f_{1}(\underline{a},\underline{a},\underline{a}) = \begin{cases} \lambda_{\underline{a}} & \text{if } \underline{a} < 1 \\ \gamma_{\underline{a}} & \text{if } \underline{a} \geq 1 \end{cases}$ Then $\mathbf{I} \in posit[(A_i], \mathbf{i}(\{i \in S\})]$ if there are $\alpha, \beta \ge \mathbf{I} \times \mathbf{I}$ $= \frac{1}{2} \frac{|\mathbf{I}| |\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}| + |\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}| + |\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2} \frac{|\mathbf{I}|}{|\mathbf{I}|} = \frac{1}{2$ Code sufficient in results a diable providend class of To from regarate 2 photoslife. -Distal Fackackslathor Base do Frenz. Houry of Puddidly: 4 (2004) Annula: International Statistical Basics, 505–50, 500 tray Transmiss observe 2 also Billey Is Jans, Osheninis Diri Kalak J. Kakenahi, Kalak Sakabalik, and Jacoph R. English meetitesise of Dir Frenz Bills. S. Ladore: Antonic parameters and detaily additor expensions. D. S. Ladore: Antonic parameters and strategy addition of parameters. It is an in order to be a based of the second of the law on advantation in the second of the second

erc

References