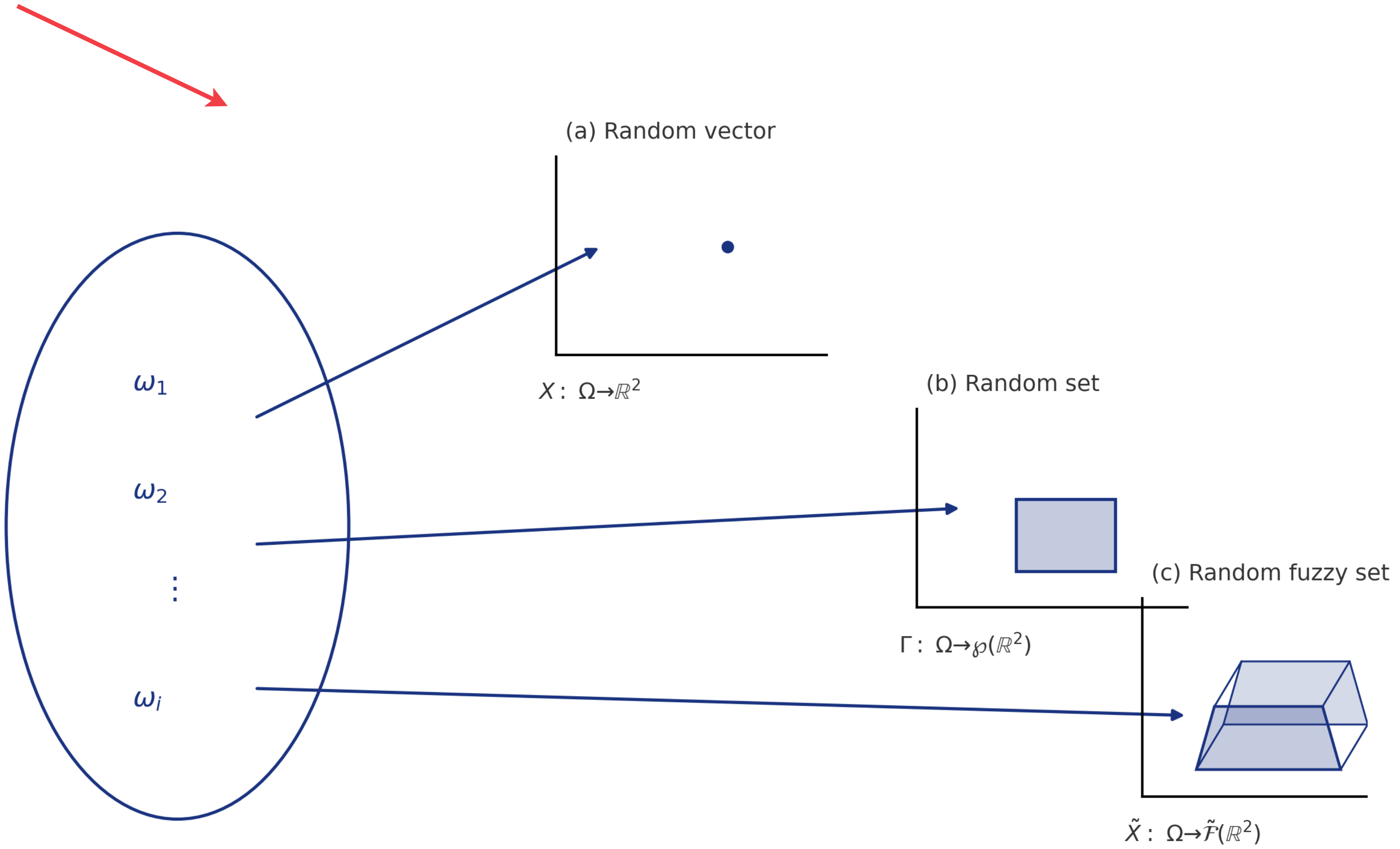
RANDOM SETS AND RANDOM FUZZY SETS

Inés Couso
Universidad de Oviedo

BFTA 2025
(7th School on Belief Functions and their Applications)
Oct 21, Granada, Spain

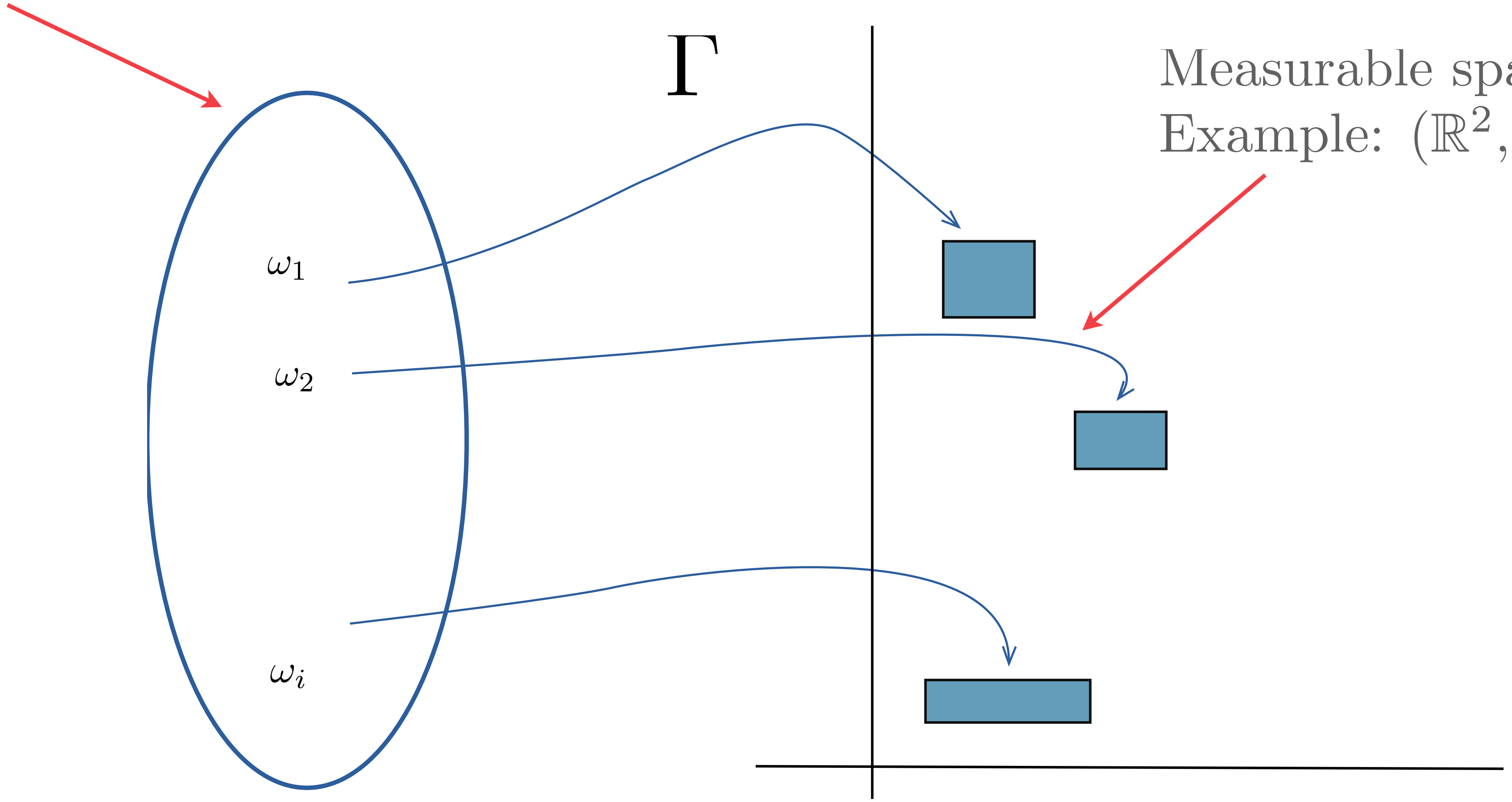
RANDOM VECTORS, RANDOM SETS, RANDOM FUZZY SETS

Prob. space $(\Omega, \mathcal{A}, P)$



Different measurability conditions lead to different formal definitions

RANDOM SETS: FORMAL DEFINITIONS

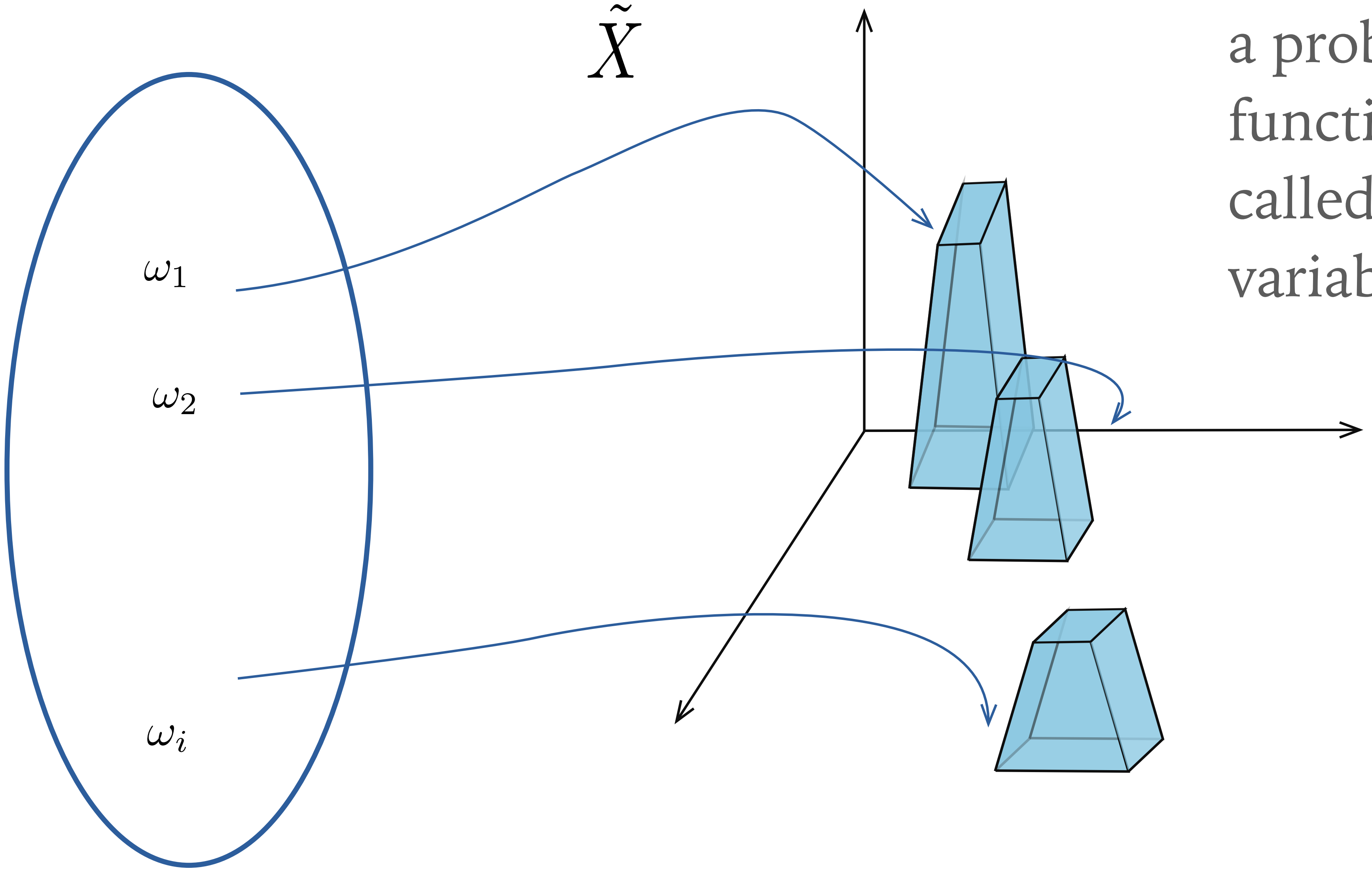


Prob. space $(\Omega, \mathcal{A}, P)$

Measurable space.
Example: $(\mathbb{R}^2, \beta)$.

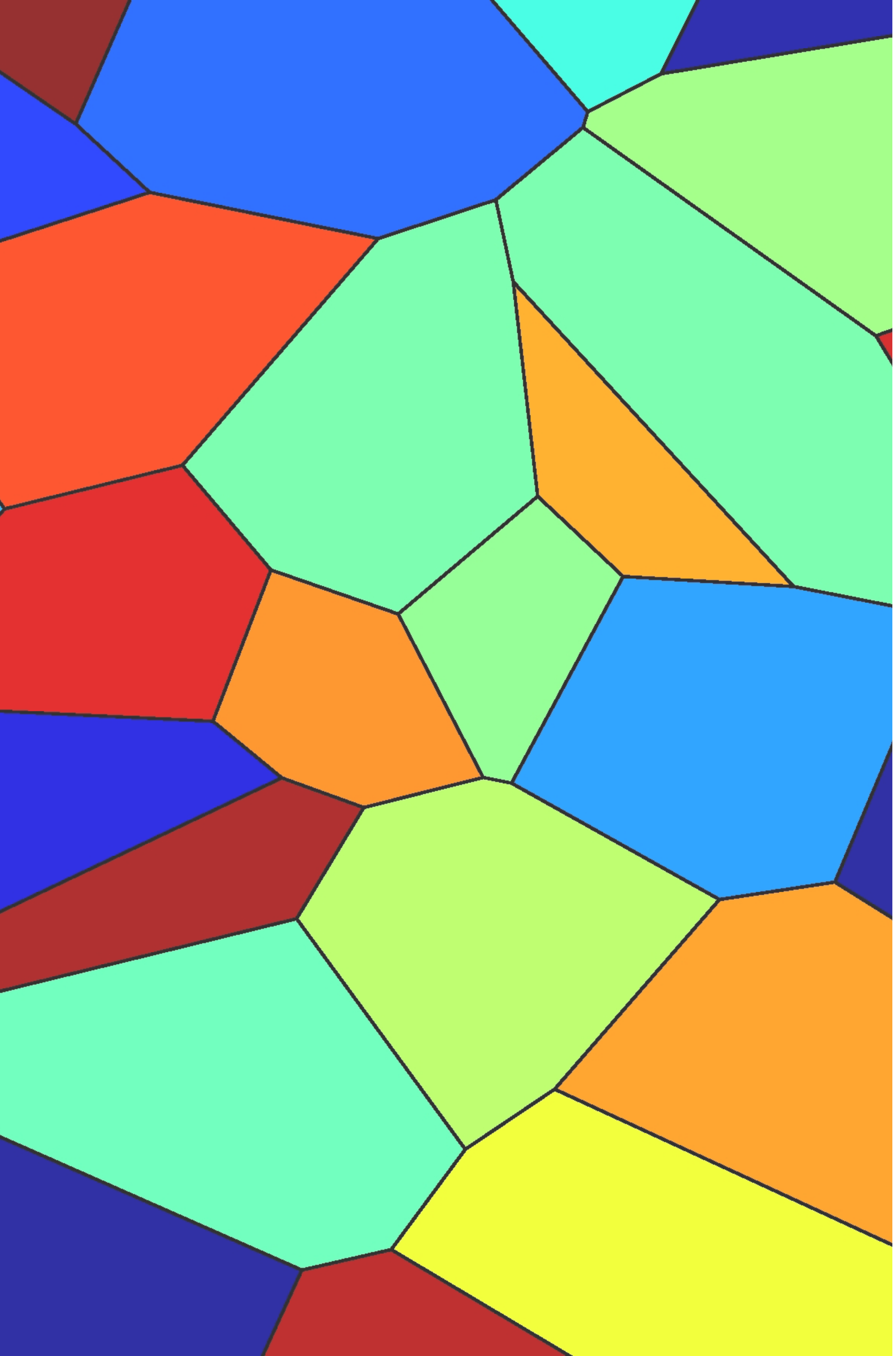
Aumann, Arstein, Debreu, Hess, Hildebrand, Himmelberg, Mathéron, Molchanov, Novikov…

RANDOM FUZZY SET: FORMAL DEFINITIONS



Measurable function from a probability space to a functional space. Also called ‘fuzzy random variable’

Féron, Krätschmer, Kruse, Kwakernaak, Puri & D. Ralescu …



RANDOM SETS

- Shafer’s Evidence Theory and random sets.
- Conjunctive vs disjunctive approaches to random sets
- Induced probability, expectation, variance, independence (conjunctive vs disjunctive interpretations)
- Statistics with coarse data (application of disjunctive random sets)

SHAFER'S EVIDENCE THEORY RECAP

- **Mass assignment:** a mapping $m : \wp(\Theta) \to [0,1]$ such that

$$m(\varnothing) = 0 \qquad \text{and} \qquad \sum_{A \subseteq \Theta} m(A) = 1.$$

- **Belief function:** given m, the belief function $Bel : \wp(\Theta) \to [0,1]$ is

$$Bel(A) = \sum_{B \subseteq A} m(B), \qquad \text{for all } A \subseteq \Theta.$$

- **Plausibility function:** the plausibility function $Pl : \wp(\Theta) \to [0,1]$ is

$$Pl(A) = \sum_{B \cap A \neq \varnothing} m(B) = 1 - Bel(A^c), \qquad \text{for all } A \subseteq \Theta,$$

where A^c denotes the complement of A in Θ.

RANDOM SETS AND SHAFER'S EVIDENCE THEORY

- Dempster, 1967: Upper and lower probabilities induced by a multi-valued mapping, Ann. Math. Statist. 38(2): 325-339.
- Shafer, 1973: Allocations of Probability: Theory of Partial Belief, PhD thesis, Princeton University.
- Shafer, 1976: A Mathematical Theory of Evidence. Princeton University Press.

BELIEF/PLAUSIBILITY FUNCTIONS INDUCED BY RANDOM SETS

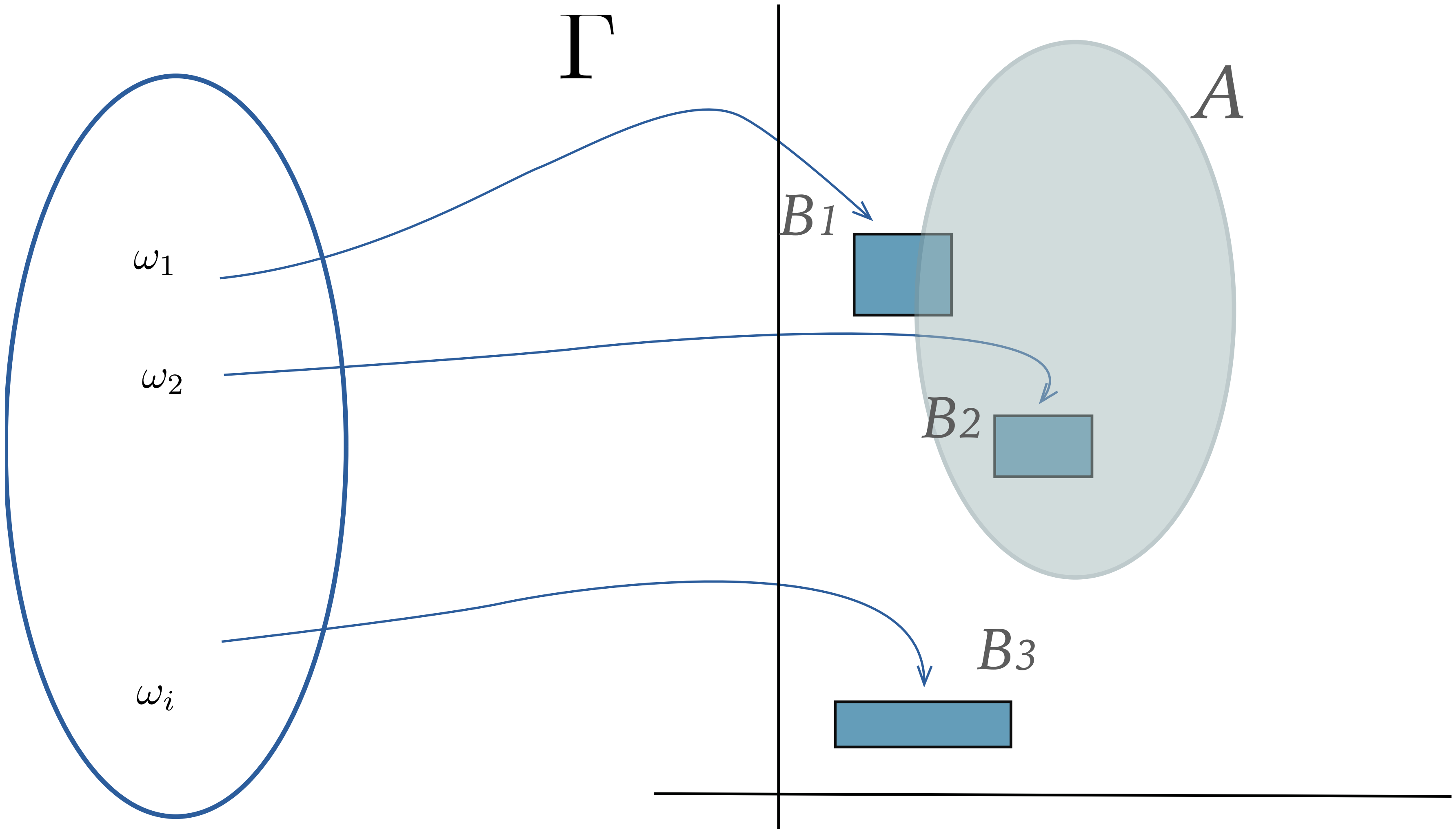
- Probability space $(\Omega, \mathcal{A}, P)$, Θ finite.
- Non-empty multivalued mapping: $\Gamma : \Omega \to \wp(\Theta)$, with $\Gamma(\omega) \neq \varnothing$, $\forall\, \omega \in \Omega$.
- Suppose $\Gamma^{-1}(\{A\}) = \{\omega \in \Omega : \Gamma(\omega) = A\} \in \mathcal{A}$.

Define

- $m(A) = P(\Gamma = A) = P\big(\{\omega \in \Omega : \Gamma(\omega) = A\}\big)$.
- $\underline{P}(A) = P\big(\{\omega \in \Omega : \Gamma(\omega) \subset A\}\big) = \displaystyle\sum_{B \subset A} m(B)$.
- $\overline{P}(A) = P\big(\{\omega \in \Omega : \Gamma(\omega) \cap A \neq \varnothing\}\big) = \displaystyle\sum_{B \cap A \neq \varnothing} m(B)$.

  Then m, $\underline{P}$ and $\overline{P}$ are a basic mass assignment, a belief and a plausibility measure, respectively.

THE FOCAL SETS ARE THE IMAGES OF THE RANDOM SET



- $m(B_2)$ =“that portion of one’s belief which one exactly commits to B_2 and to nothing more specific than B_2” (Shafer)

- $P(\Gamma = B_2) = P(\{\omega_2\})$ The proportion of times that the outcome of the random experiment is committed to B_2 and nothing more specific that B_2.

CHOQUET/MATHÉRON THEOREM

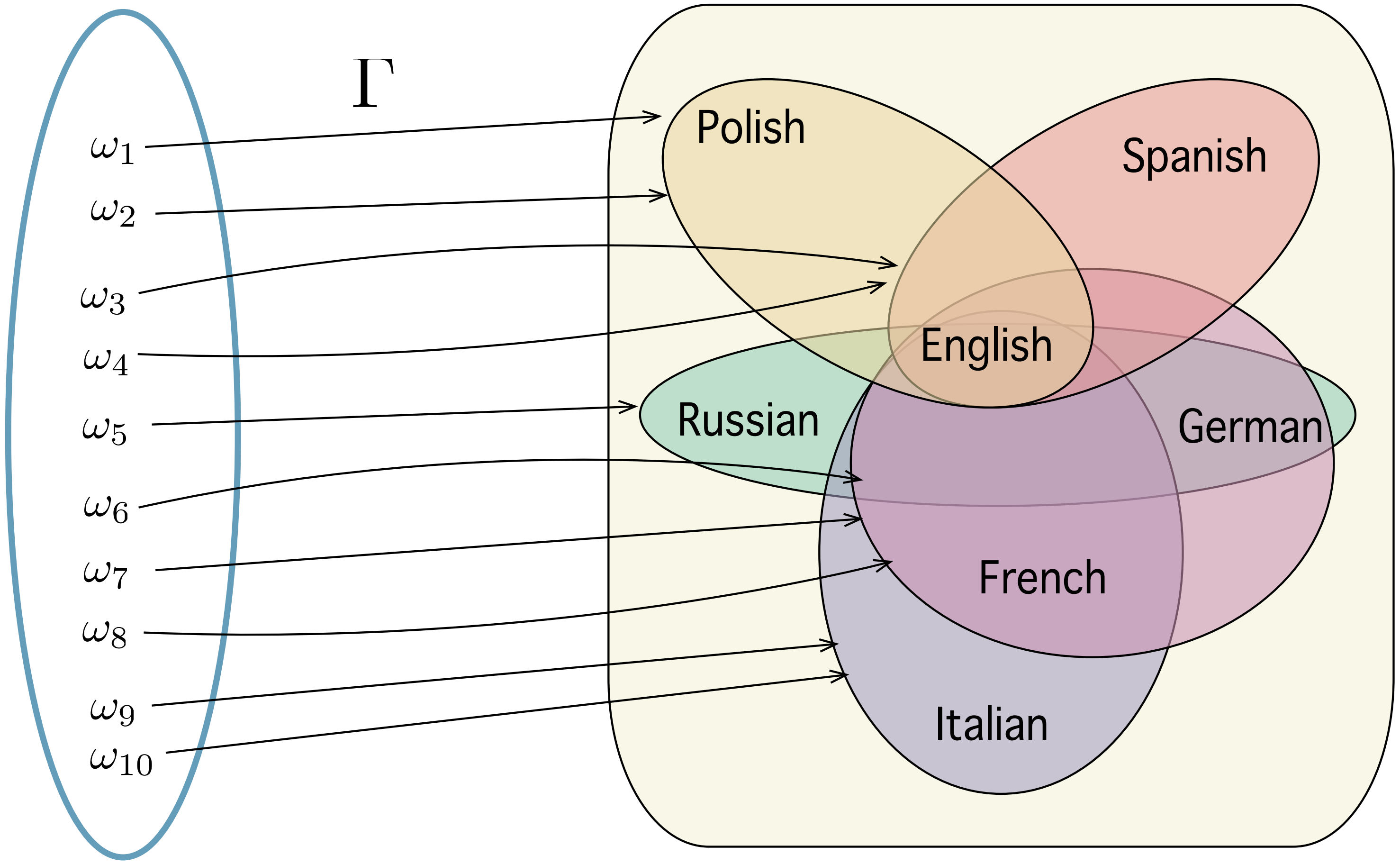
- The upper probability of a random set is completely alternating. Conversely, any infinite alternating capacity is the upper probability of random set.
- Recall: plausibility measures are special cases of infinite-alternating Choquet capacities (finite universe).
- Consequence of Choquet theorem: any plausibility/belief function induces a mass assignment (Möbius transform) and is induced by a multivalued mapping.

TWO DIFFERENT INTERPRETATIONS OF SETS

- **What does a conjunctive set represent?**
  - Precise complex information. For instance, the collection of languages that someone else is able to speak is: {Spanish, Portuguese, Italian}.

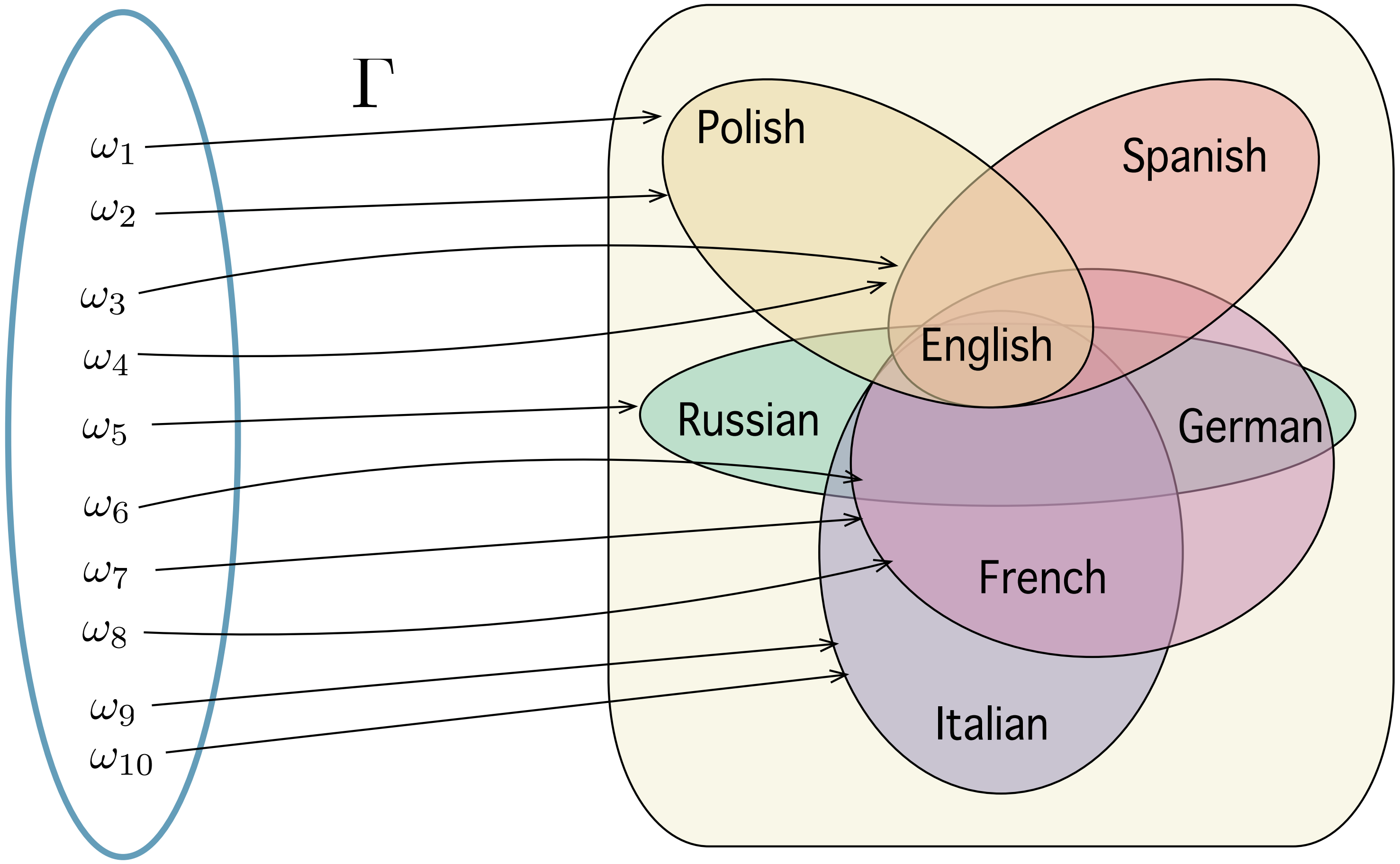
- **What does a disjunctive set represent?**
  - Incomplete information about an element of the universe. For instance, our incomplete information about someone's nationality is that it belongs to {Spanish, Portuguese, Italian}.

CONJUNCTIVE VS DISJUNCTIVE INTERPRETATIONS OF RANDOM SETS: EXAMPLE



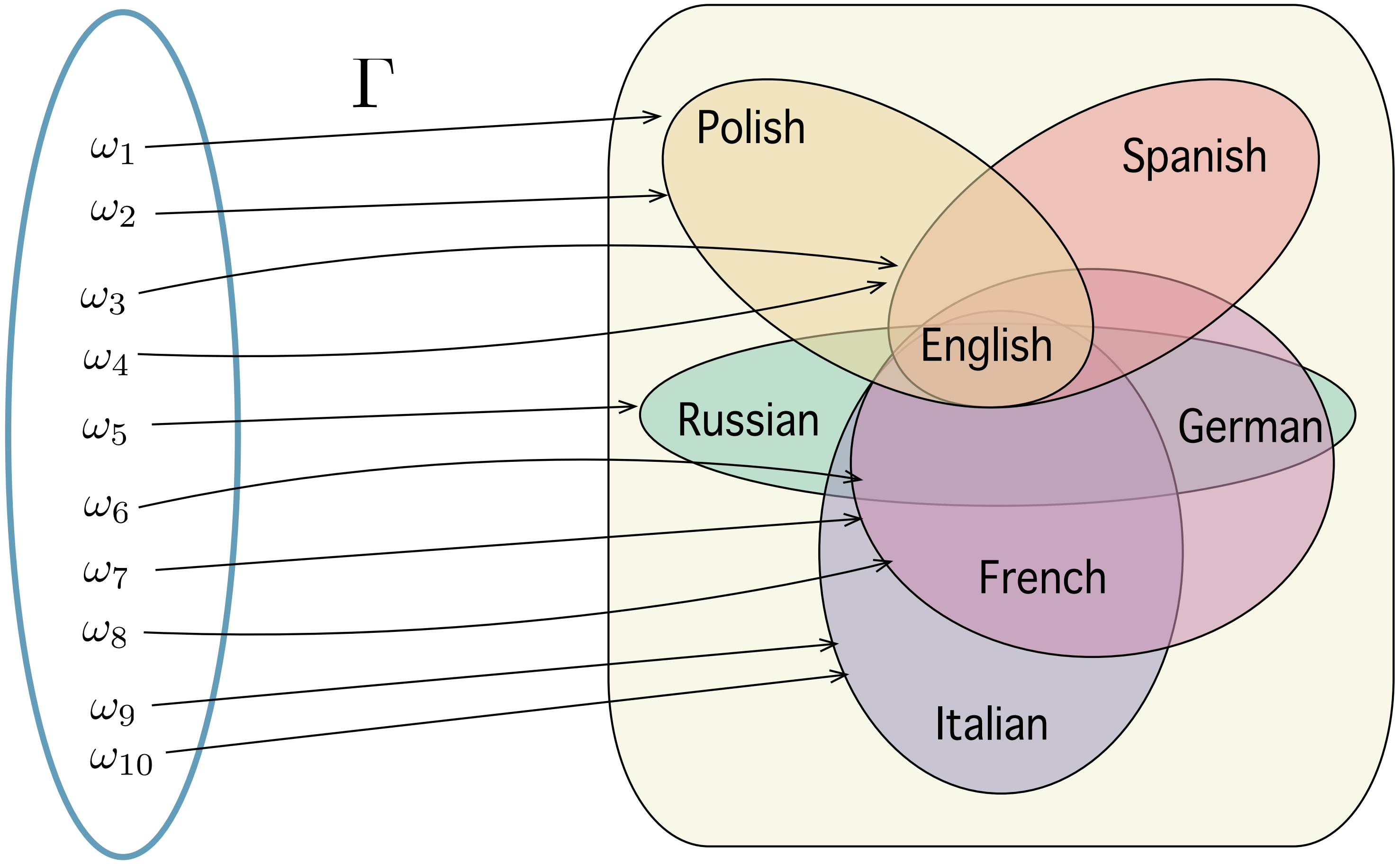
- **Situation 1:** Set of spoken languages (conjunctive). Proportion of people speaking at least French: P({e,f,i}, {e,f,g})=0.5.
- **Situation 2:** imprecise info about nationality (disjunctive). Proportion of French people: 0≤P({f})≤0.5.

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE CONJUNCTIVE SETTING

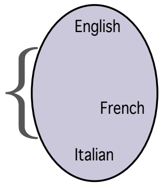
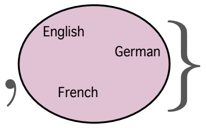


- **Situation 1:** Proportion of French speakers?

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE CONJUNCTIVE SETTING

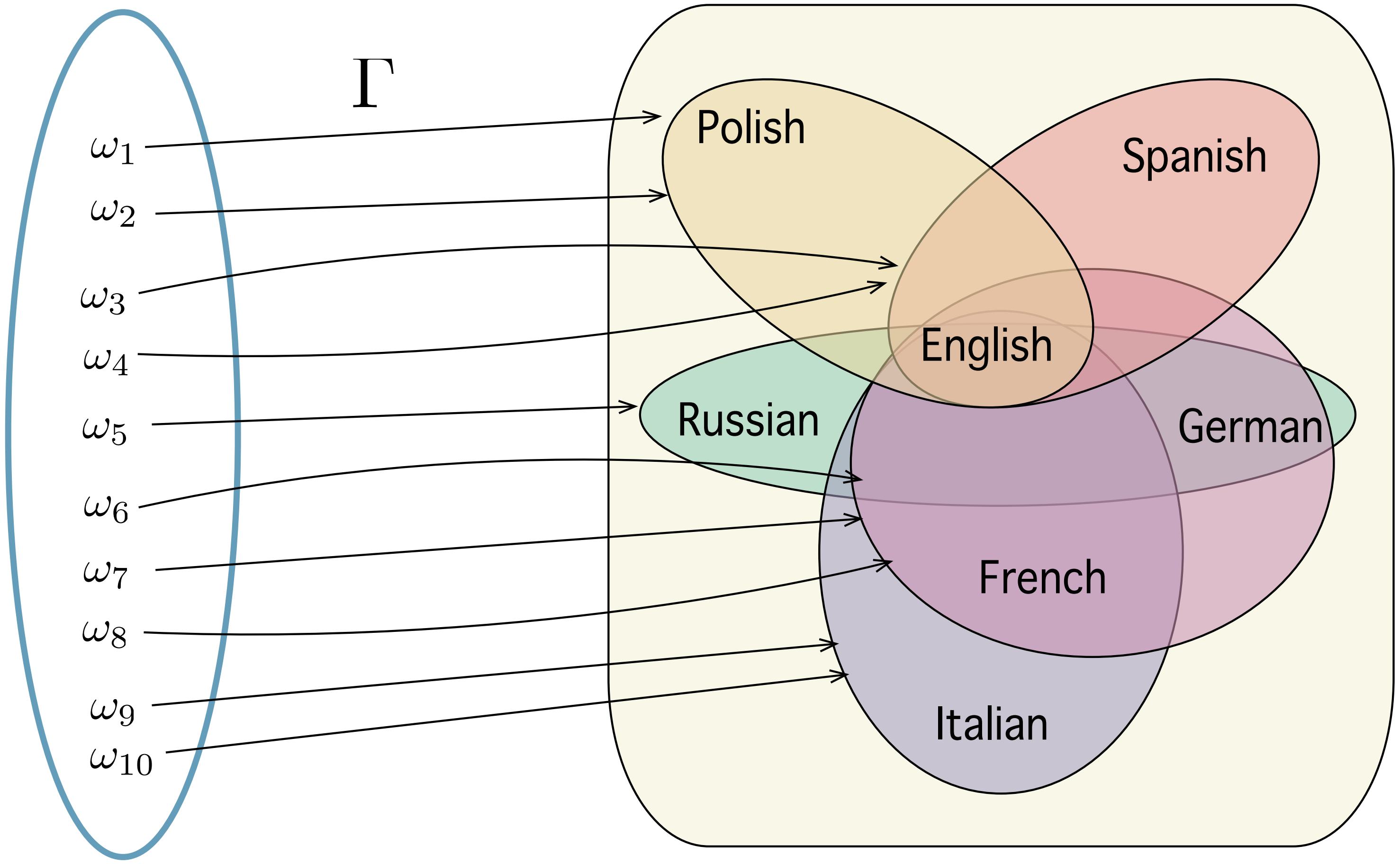


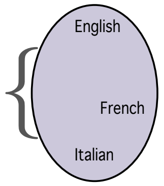
- **Situation 1:** Proportion of French speakers?

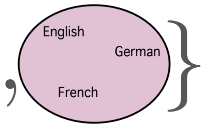
$P_\Gamma(\{$  $,$  $\}) = P(\{\omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\})$

$$\overline{P}(\{French\}) = P(\{\omega \in \Omega\} : \Gamma(\omega) \cap \{French\} \neq \varnothing)$$

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE CONJUNCTIVE SETTING



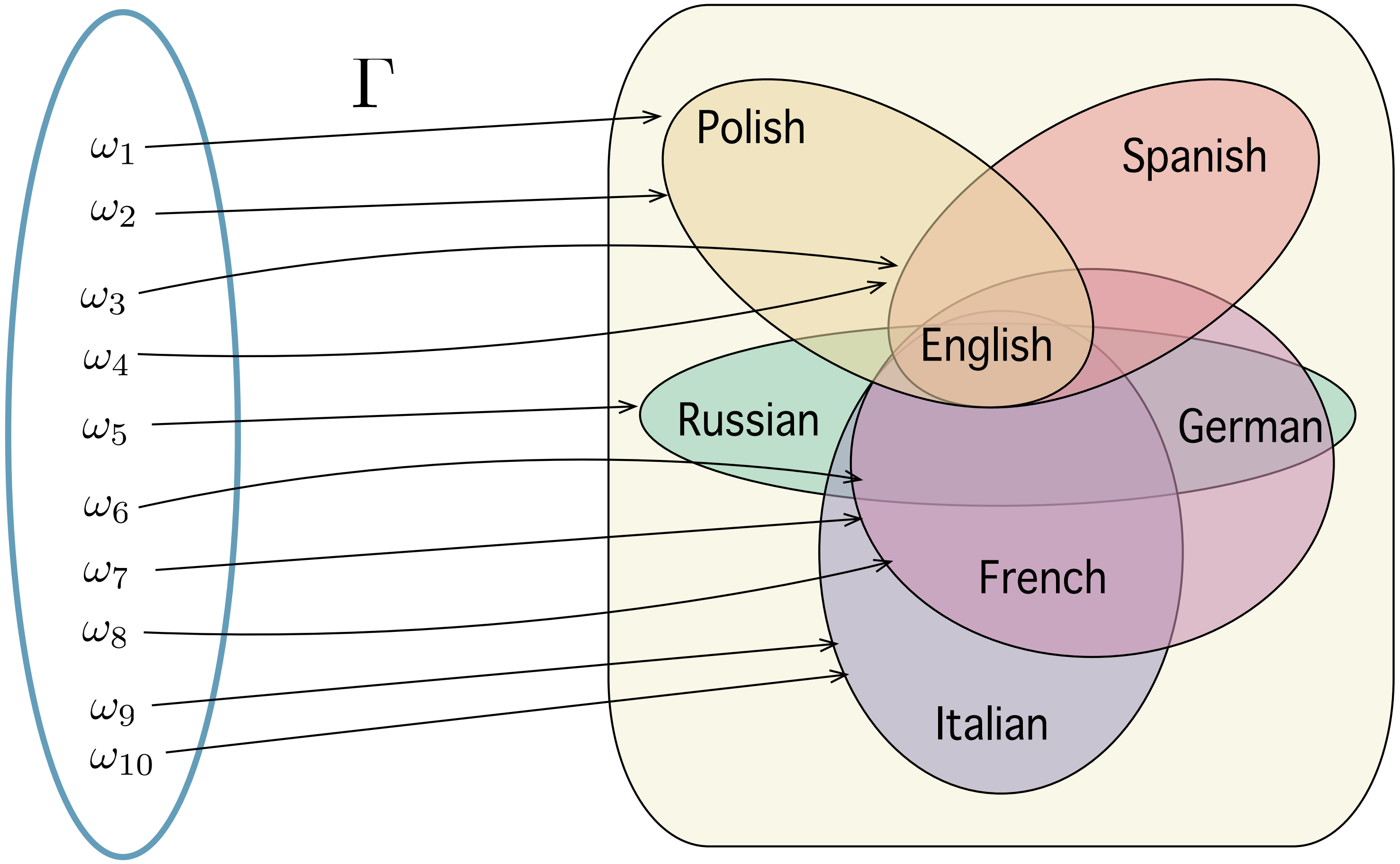
$P_\Gamma(\{$ 

, 

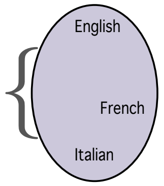
$\}) = P(\{\omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\})$

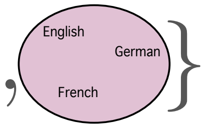
$$\overline{P}(\{\mathit{French}\}) = P(\{\omega \in \Omega\} : \Gamma(\omega) \cap \{\mathit{French}\} \neq \emptyset)$$

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE CONJUNCTIVE SETTING



- **Situation 1:** Proportion of non French speakers?

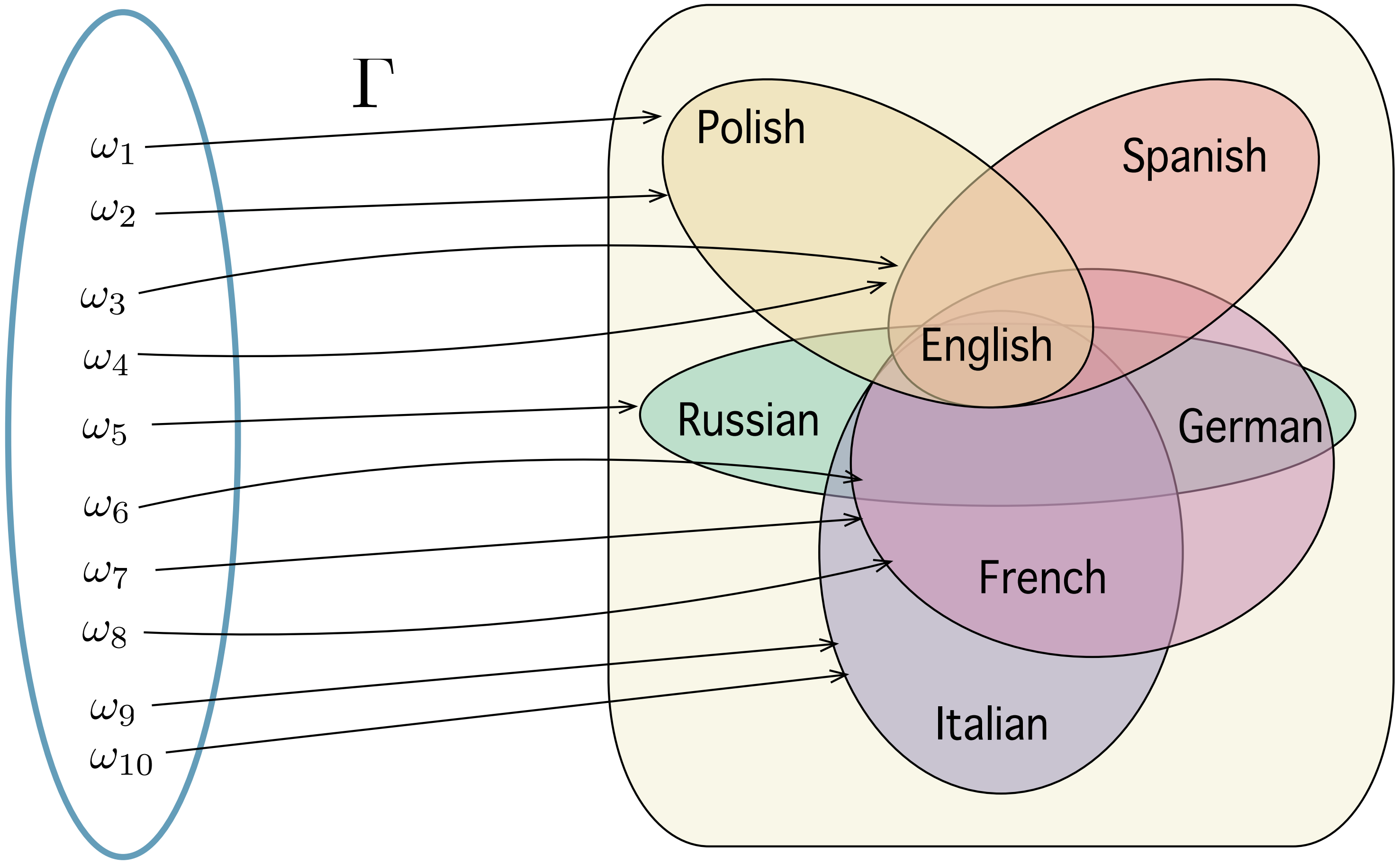
$P_\Gamma(\{$ 

, 

$\}) = P(\{\omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\})$

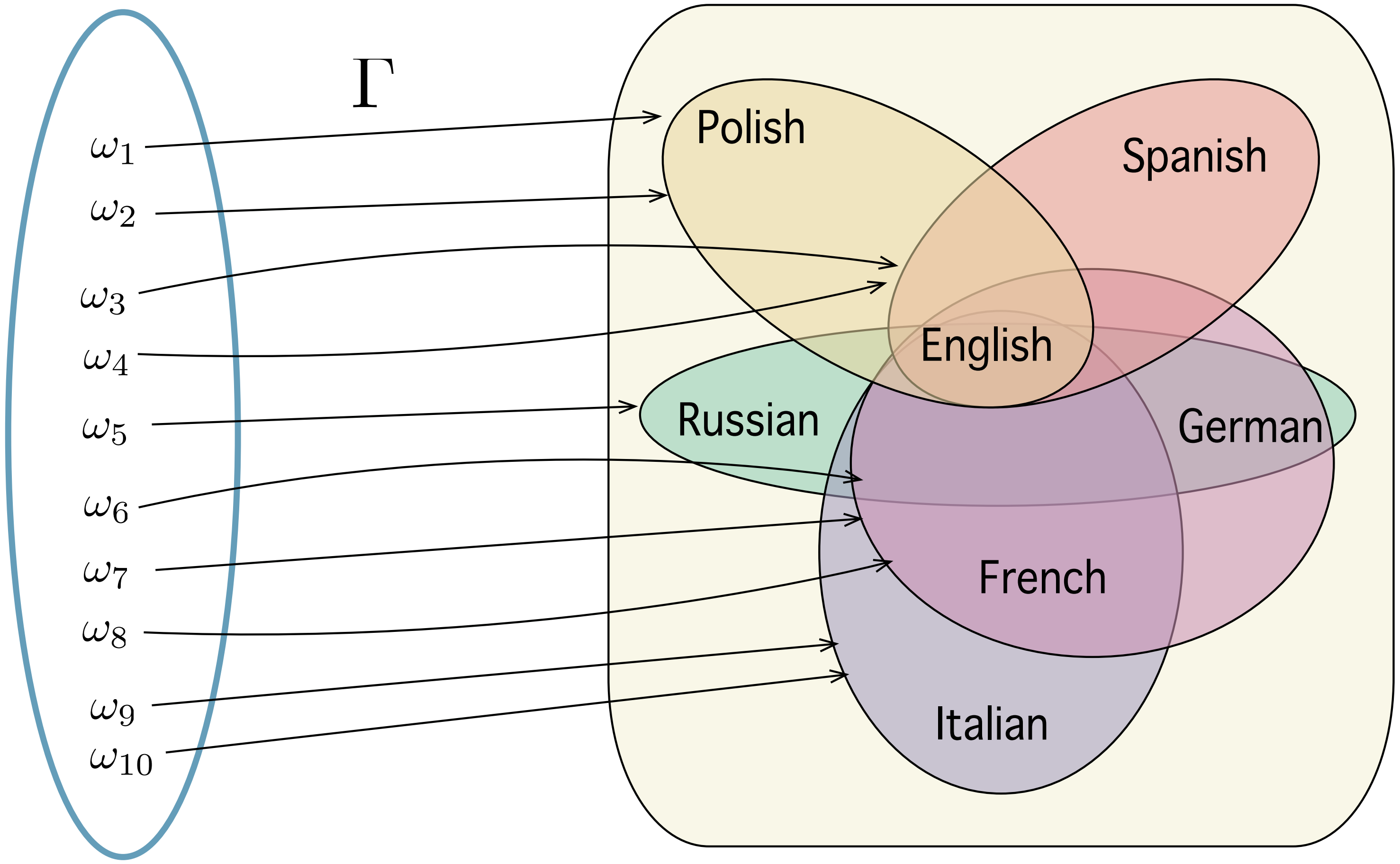
$$\overline{P}(\{French\}) = P(\{\omega \in \Omega\} : \Gamma(\omega) \cap \{French\} \neq \emptyset)$$

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE CONJUNCTIVE SETTING



- **Situation 1:** Proportion of non French speakers?

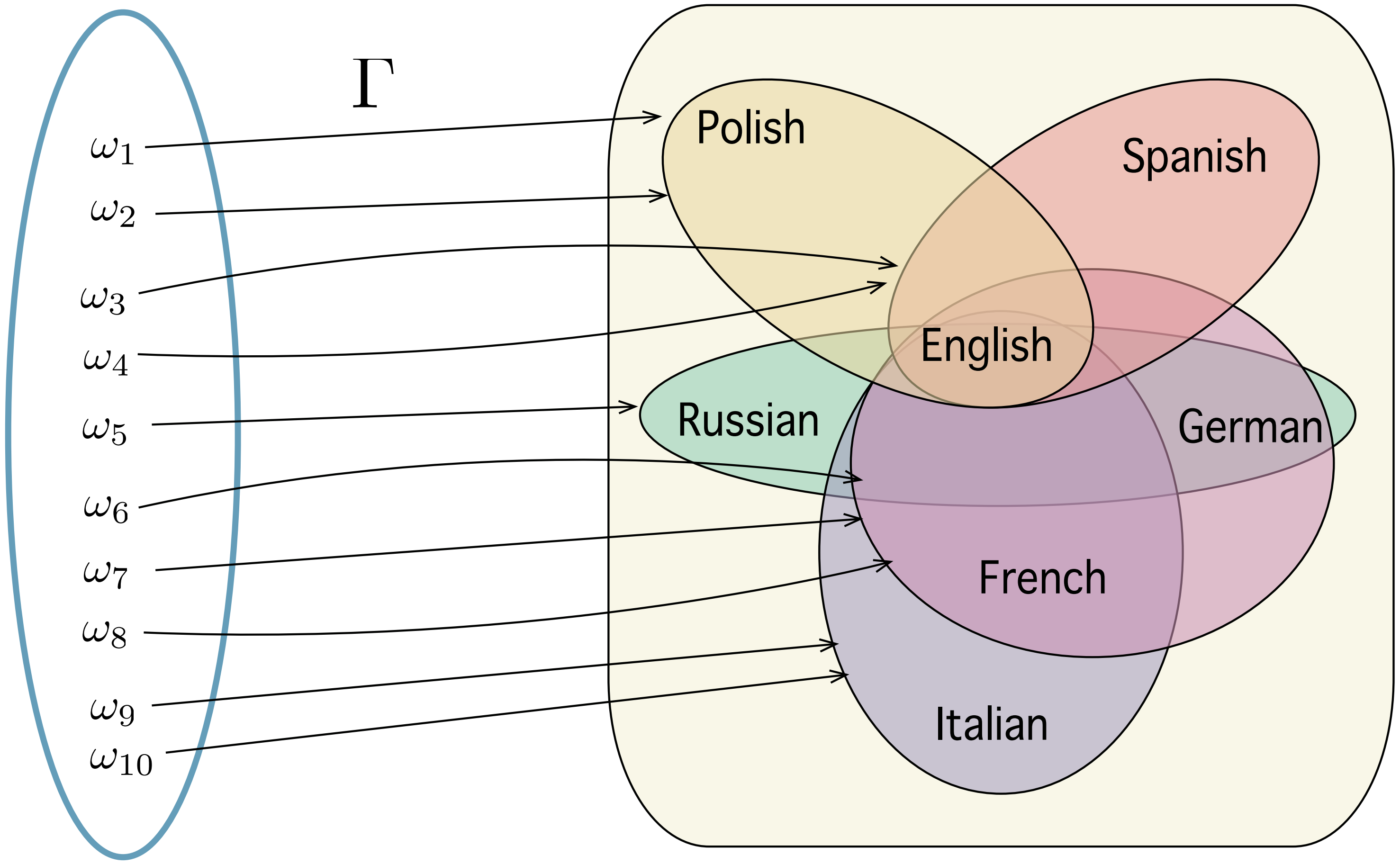
UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE CONJUNCTIVE SETTING



- **Situation 1:** Proportion of non French speakers?

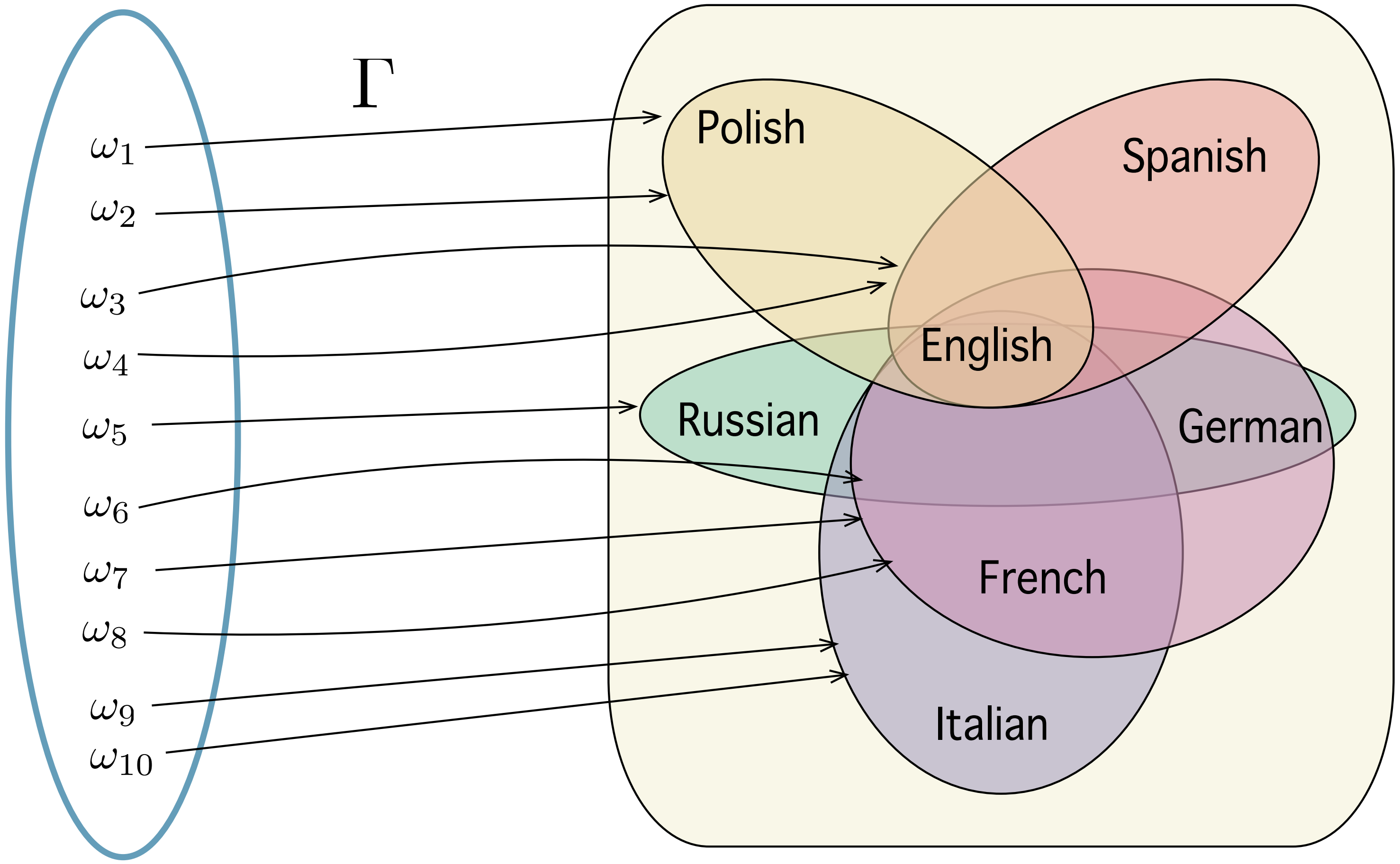
$$\underline{P}(\{French\}^c) = 1 - \overline{P}(\{French\})$$

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE DISJUNCTIVE SETTING



- **Situation 2:** Proportion of French people: upper and lower bounds?

UPPER AND LOWER PROBABILITIES INDUCED BY RANDOM SETS IN THE DISJUNCTIVE SETTING



- **Situation 2:** Proportion of French people: upper and lower bounds?

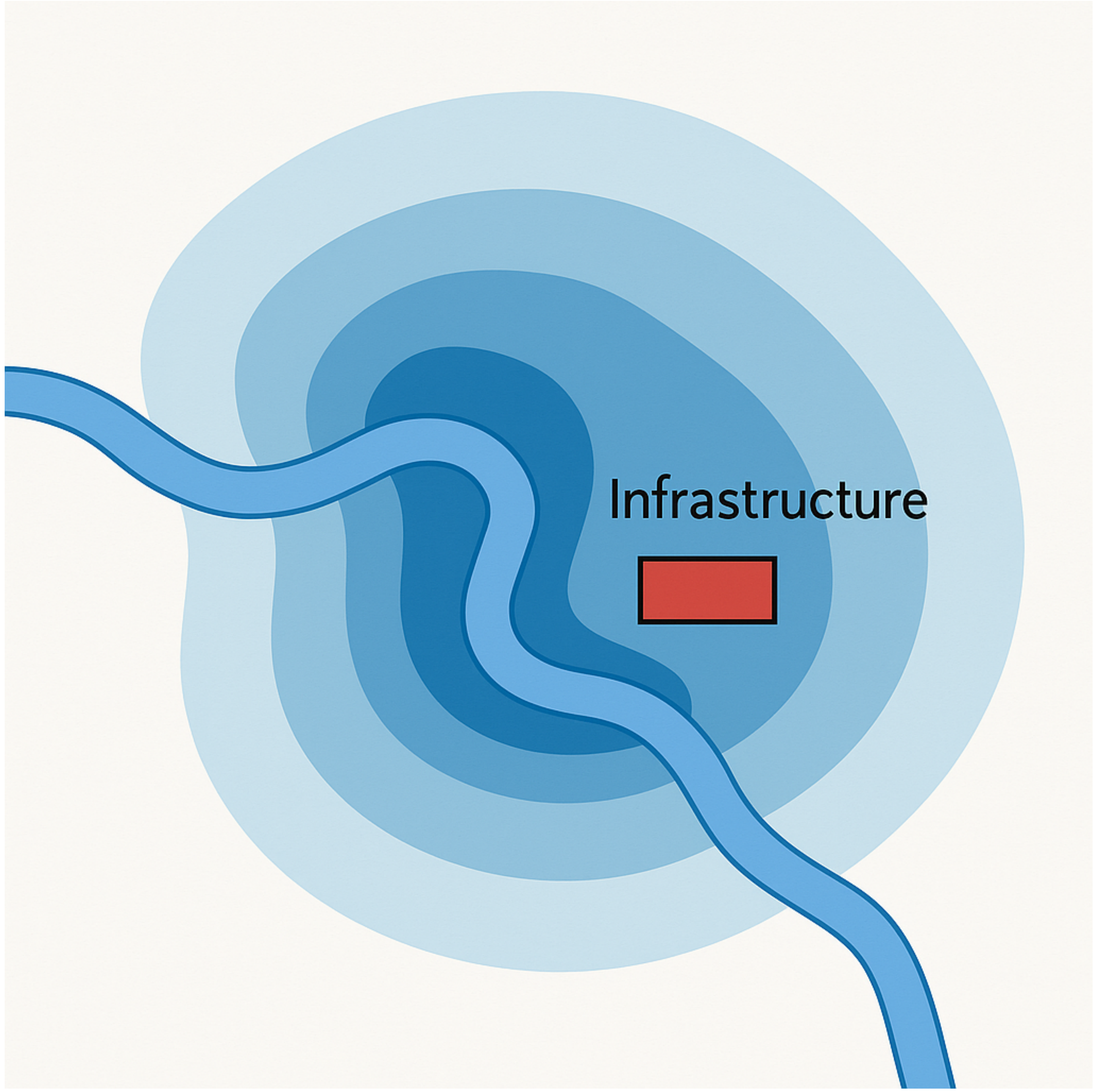
$$\overline{P}(\{French\}) = P(\{\omega \in \Omega : \Gamma(\omega) \cap \{French\} \neq \emptyset\}) = P(\{\omega_6, \ldots, \omega_{10}\})$$

$$\underline{P}(\{French\}) = P(\{\omega \in \Omega : \Gamma(\omega) \subset \{French\}\}) = 0$$

SOME APPLICATIONS OF RANDOM SETS

- Conjunctive view:
  - Stochastic geometry and spatial data: analyze random geometric objects (regions, boundaries, coverings).
  - Microeconomics choice theory.
- Disjunctive view: Statistics with coarse data.

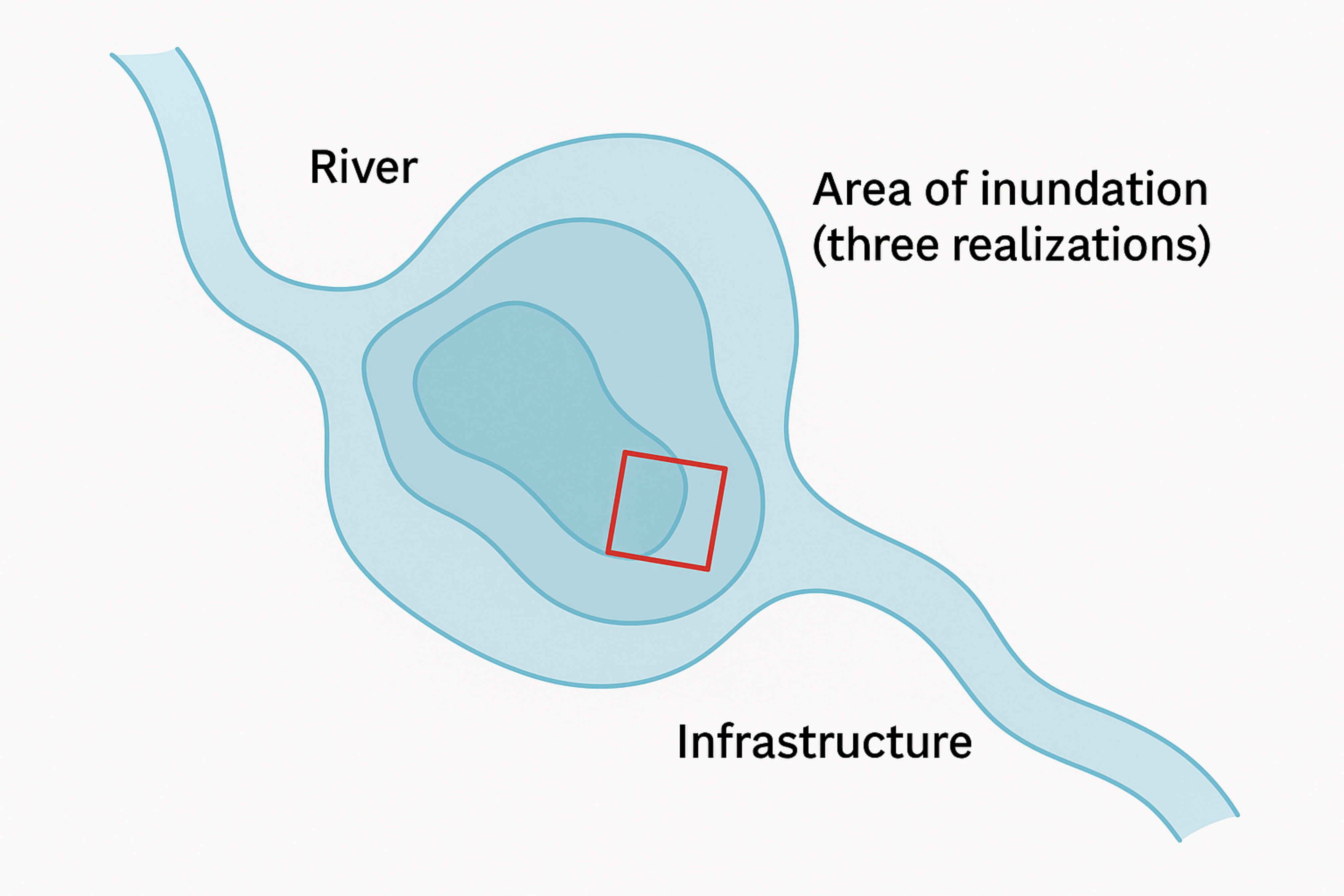
CONJUNCTIVE VIEW: SPATIAL DATA



Random regions

- A random set associates to each time instant the region flooded at the moment.
- Each realisation represents a possible flooding configuration.
- This representation allows studying the expected proportion of times a given facility is flooded.

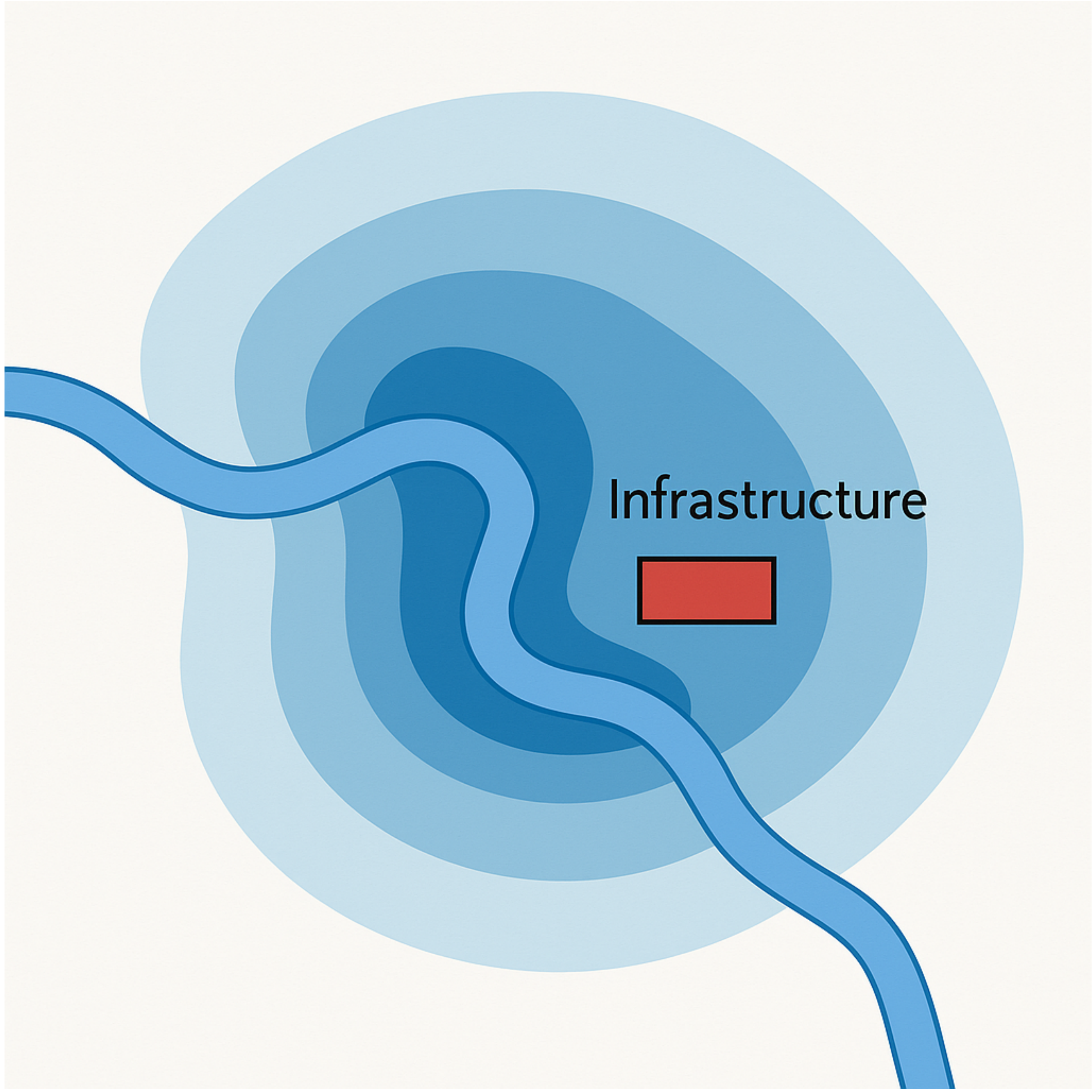
MY DISAGREEMENT WITH CHATGPT



It floods outward, not inward…

MY DISAGREEMENT WITH CHATGPT

MY DISAGREEMENT WITH CHATGPT

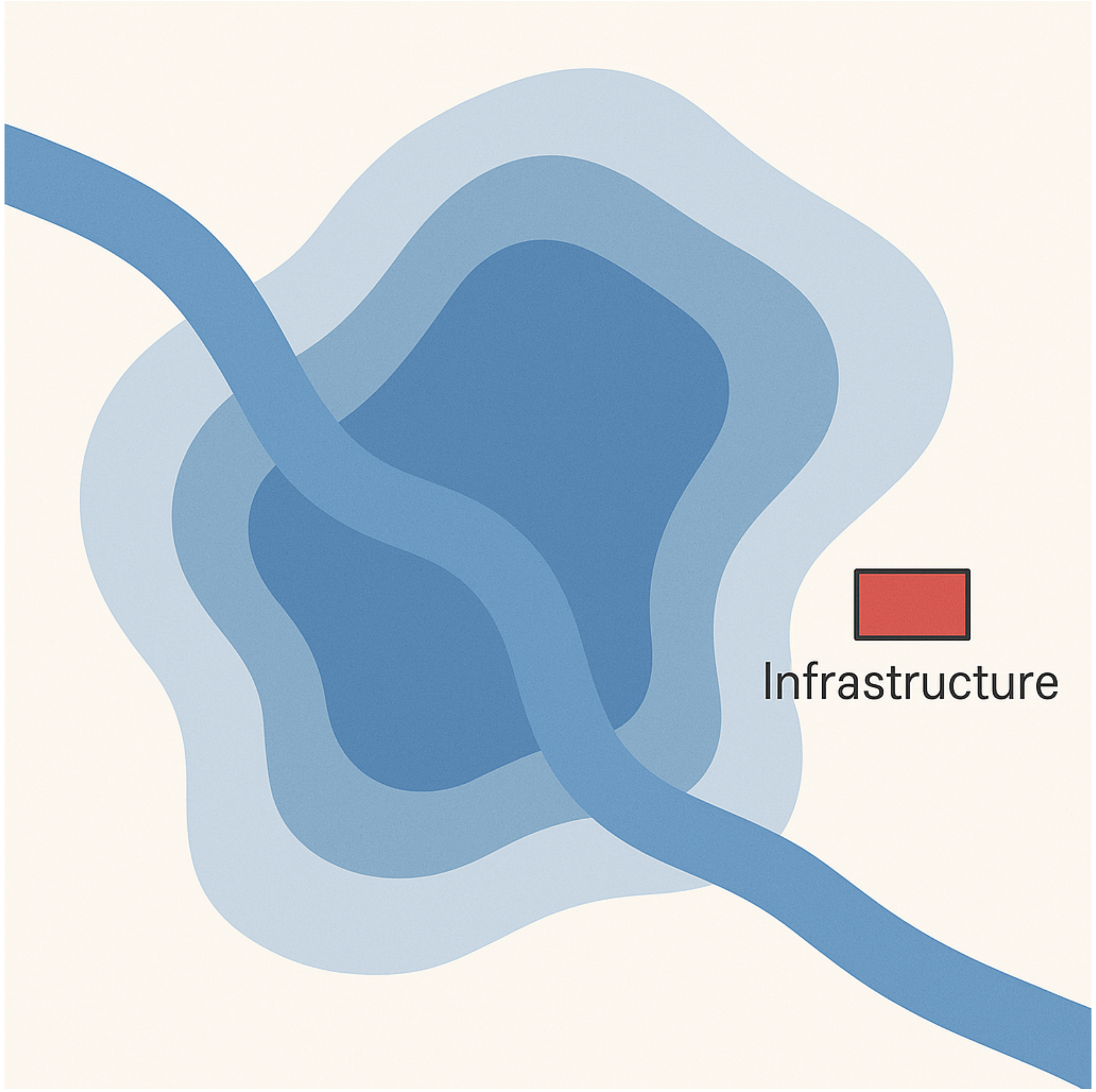


But don’t build infrastructure that’s always underwater,

please!

MY DISAGREEMENT WITH CHATGPT

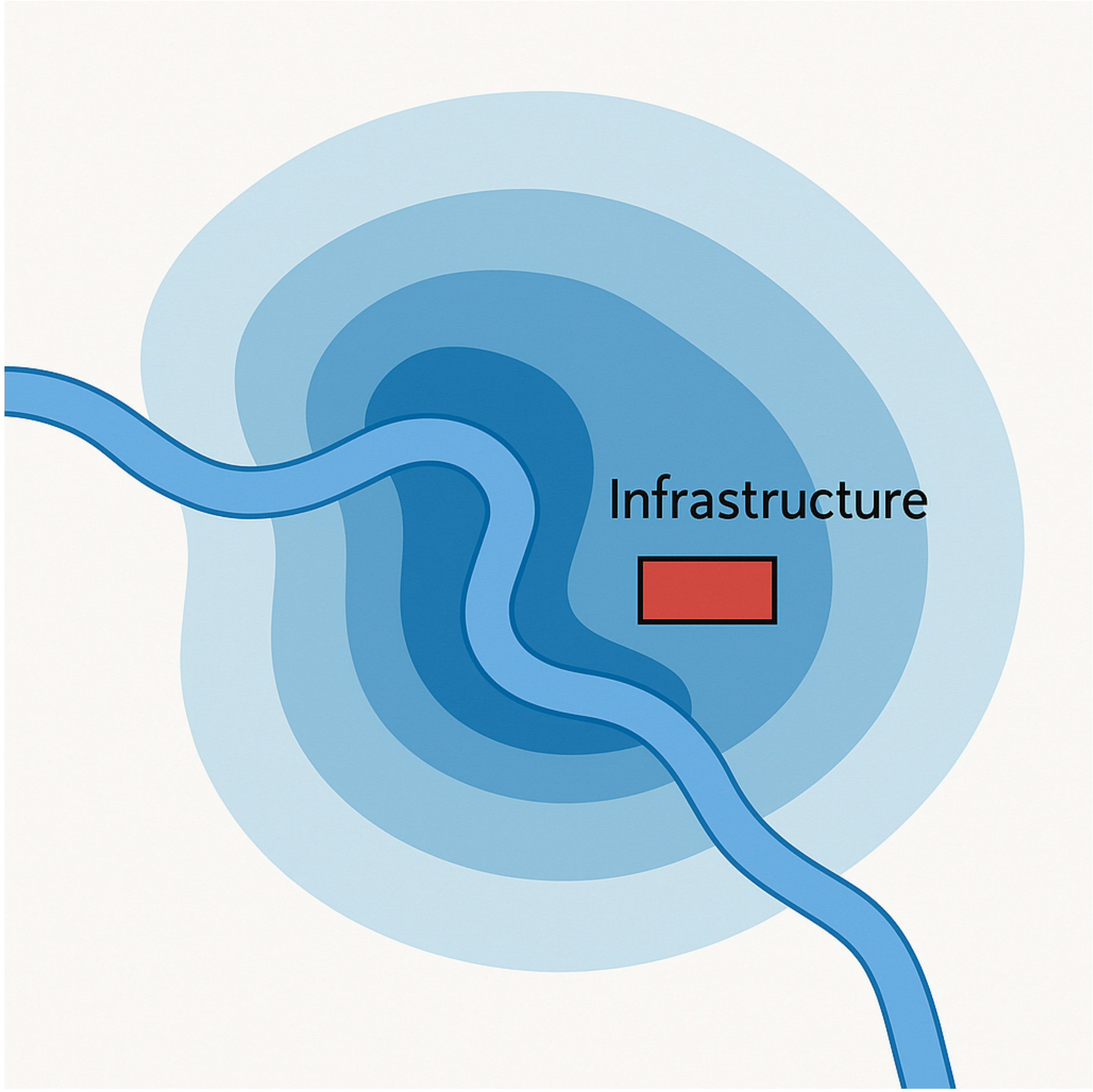
MY DISAGREEMENT WITH CHATGPT



But don't make it so it never floods, either. Try to strike a middle ground.

MY DISAGREEMENT WITH CHATGPT

MY DISAGREEMENT WITH CHATGPT



OK, that's more or less fine…

CONJUNCTIVE VIEW: MICRECONOMICS CHOICE THEORY

- $(\Omega, \mathcal{A}, \mathbb{P})$ prob. space, $X \subset \mathbb{R}_+^L$ compact, $\mathcal{W} \subset (0, \infty)$.
- $x \in X$: consumption bundle (quantities of L goods).
- $u : X \to \mathbb{R}$: continous utility (representing preferences over bundles).
- $p = (p_1, \ldots, p_L) \in (0, \infty)^L$: unit price vector;
- $w \in \mathcal{W}$: income.
- Random environment: $\mathbf{p} : \Omega \to \mathcal{P}$, $\mathbf{w} : \Omega \to \mathcal{W}$ (meas.).
- Budget set $B(p, w) = \{x \in X : \ p \cdot x \leq w\}$ (compact).
- Demand correspondence: $D(p, w) := \arg\max_{x \in B(p,w)} u(x)$

$$\boxed{X(\omega) = D\big(\mathbf{p}(\omega),\ \mathbf{w}(\omega)\big) \subset X} \quad \text{is a random closed set.}$$

DISJUNCTIVE VIEW: COARSE DATA

APARTMENTS RENTAL

id	Neighborhood	Rent €	Area m2	Beds	Move in Date	Min Metro	Deposit €	Energy	Pets
R-101	Center	[1250, 1400]	[48, 52]	1	[2025-11-01, 2025-12-01]	[3, 5]	[1200, 1600]	C	unknown
R-108	Riverside	1350	NA	2	>=2025-12-15	[8, 12]	2x rent	NA	no
R-115	Old Town	[950, 1050]	[35, 40]	studio	[2025-10-20, 2025-11-05]	<=2	[900, 1000]	D	yes
R-121	Campus	[1100, 1200]	[45, 47]	1	NA	[5, 7]	[1100, 1500]	B	no
R-132	Harbor	[1600, 1850]	[55, 60]	2	[2026-01-01, 2026-02-01]	>15	[1600, 2000]	A	yes
R-141	North	[800, 900]	<30	studio	[2025-11-10, 2025-11-30]	[6, 9]	[800, 900]	NA	unknown
R-156	Hillview	[1450, 1500]	[50, 55]	2	<=2025-11-01	[10, 14]	[1400, 1800]	C	no
R-162	Market	[1200, 1300]	[42, 45]	1	[2025-12-01, 2026-01-15]	[4, 6]	[1200, 1500]	B	yes
R-177	Parkside	[1000, 1100]	NA	1	>=2025-11-20	<=5	[1000, 1200]	D	no
R-184	Lakeshore	[1700, 1900]	[60, 68]	3	NA	[12, 18]	[1700, 2200]	A	yes

COARSE DATA

- Incomplete data due to:
    - Limited reliability of measuring instruments. Significant digits.
    - Intermittent measurements.
    - Censoring.
    - Partially categorized data.
    - Missing data.
    - etc.

- R. J. A. Little, D. B. Rubin (1987), Statistical Analysis with Missing Data. Wiley, New York.
- D. F. Heitjan, D. B. Rubin, Ignorability and coarse data, The Annals of Statistics 19 (1991) 2244-2253.
- S. Ferson et al., Experimental Uncertainty Estimation and Statistics for Data Having Interval Uncertainty, SAND2007-0939, 2007.

DIFFERENT CONTEXTS REQUIRE DIFFERENT TOOLS

- **Conjunctive view**
  - Goal: study the probability distribution of a random object with set-valued images. (Probability on the power set -we consider sets of sets-).
  - Tools: A numerical distance between sets is defined. The images of r.s. are elements of a metric space. (Classical) probability measure on metric spaces.
- **Disjunctive view**
  - Goal: study the distribution of the random variable (or vector) imprecisely described by the random set.
  - Tools: (non-necessarily convex) sets of probabilities and set-valued moments and sample statistics (mean, variance, etc). (Extended) imprecise probability.

- I. Couso, D. Dubois, L. Sánchez (2014) Random sets and random fuzzy sets as ill-perceived random variables, Springer.
- I. Couso, D. Dubois (2014) Statistical reasoning with set-valued information: Ontic vs epistemic views, IJAR 55, 1502-1518.
- I. Couso, L. Sánchez (2014) Harnessing the information contained in low-quality data sources, Special Issue in IJAR (position papers by I.Couso & D. Dubois, T. Denoeux, E. Hüllermeier, S. Moral, SMIRE Group, discussions and rejoinders).

CONJUNCTIVE RANDOM SETS: MAIN NOTIONS

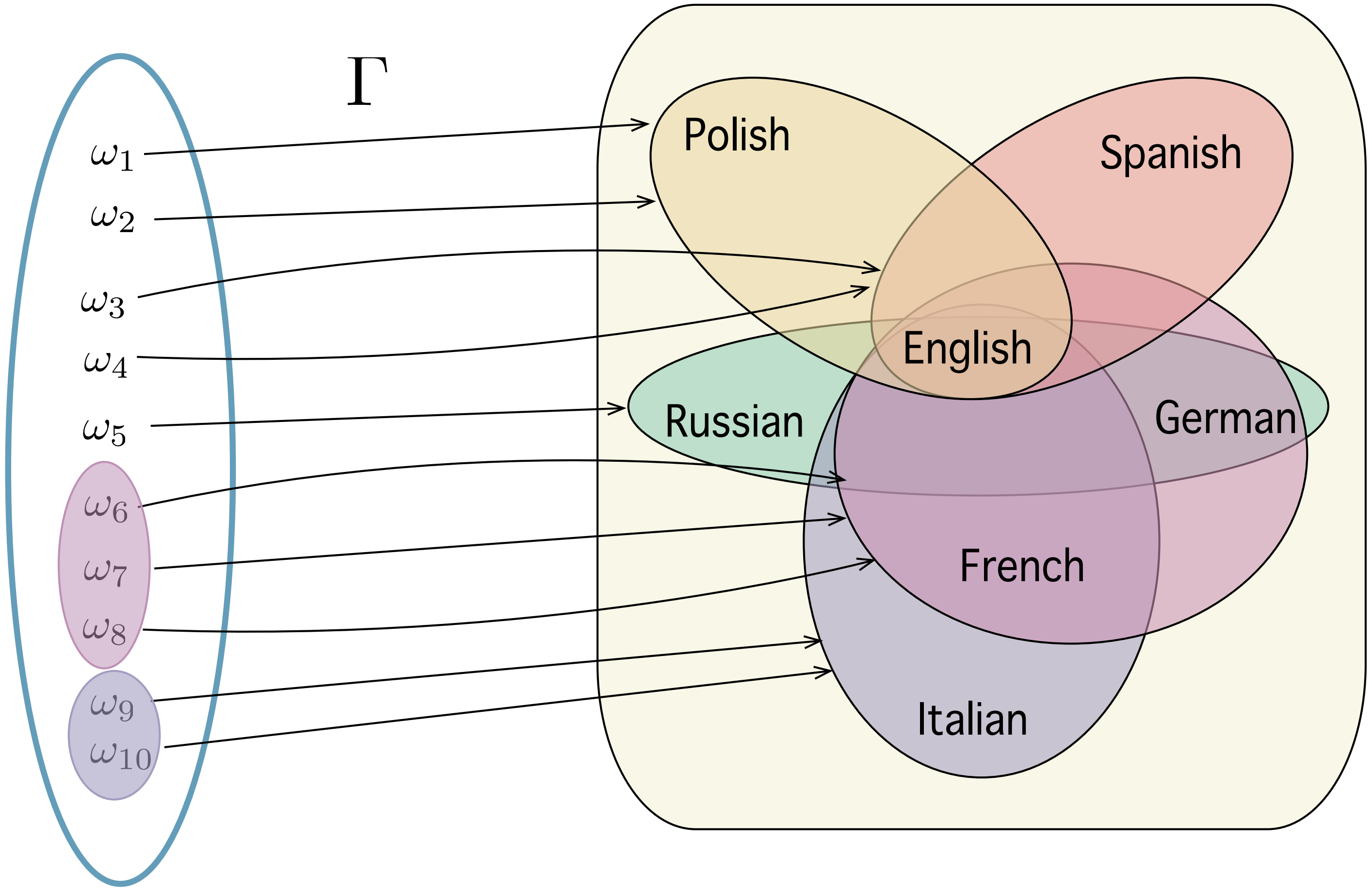
Induced probability

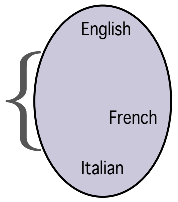
Expectation

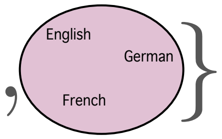
Variance

Independence

CONJUNCTIVE RANDOM SETS: INDUCED PROBABILITY



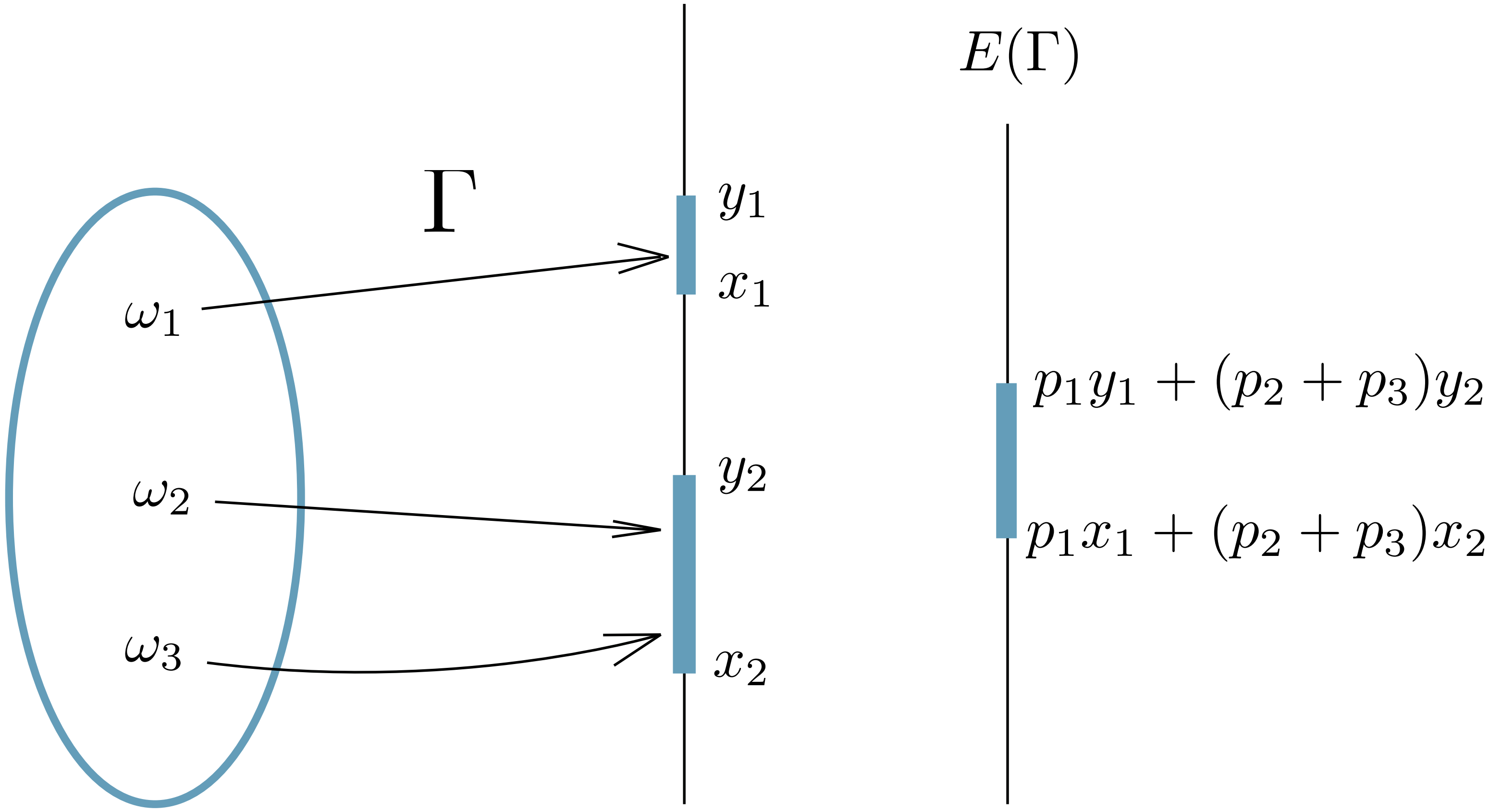
$P_\Gamma(\{$ 

, 

$\}) = P(\{\omega_6, \omega_7, \omega_8, \omega_9, \omega_{10}\})$

$$\boxed{P_\Gamma(\mathcal{F}) = P(\{\omega \in \Omega \,:\, \Gamma(\omega) \in \mathcal{F}\})}$$

CONJUNCTIVE RANDOM SETS: EXPECTATION



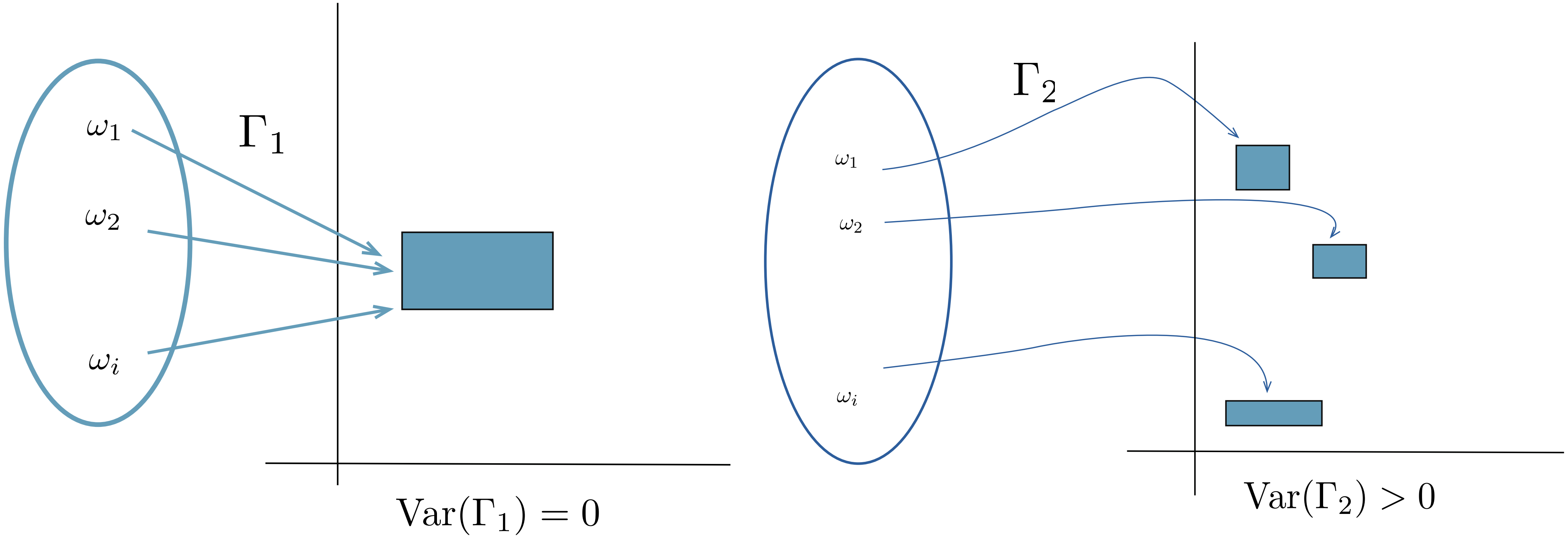
Two different formal approaches:

- Lebesgue-like approach: Set-valued arithmetic for simple r.s. plus limite calculus for non-simple r.s.

- Fréchet approach: the set minimising the square distance wrt to the images of the random set.

- A. Baddeley & I. Molchanov, S. Bochner, G. Debreu, V. Krätschmer, Näther, Stojakovic, etc…

CONJUNCTIVE RANDOM SETS: SCALAR VARIANCE



Formal assumptions:

- d is a metric.
- $f(\omega) = d(\Gamma(\omega), E(\Gamma))$ is integrable.

$$\mathrm{Var}(\Gamma) = \int [d(\Gamma(\omega), E(\Gamma))]^2 \, dP(\omega)$$

- R. Körner (FSS 1997), Lubiano et al. (FSS 2000), Näther (2001), Y. Feng et al. (FSS 2001), I. Couso & D. Dubois (IEEE-TFS 2009).

CONJUNCTIVE RANDOM SETS: INDEPENDENCE

Γ_1 and Γ_2 independent iff

$$P(\Gamma_1 \in \mathcal{F}_1, \Gamma_2 \in \mathcal{F}_2) = P(\Gamma_1 \in \mathcal{F}_1) \cdot P(\Gamma_1 \in \mathcal{F}_2).$$

(Ordinary definition in probability theory).

(Connected to Dempster’s rule of combination -mass assignment associated to the intersection of two independent random sets-)

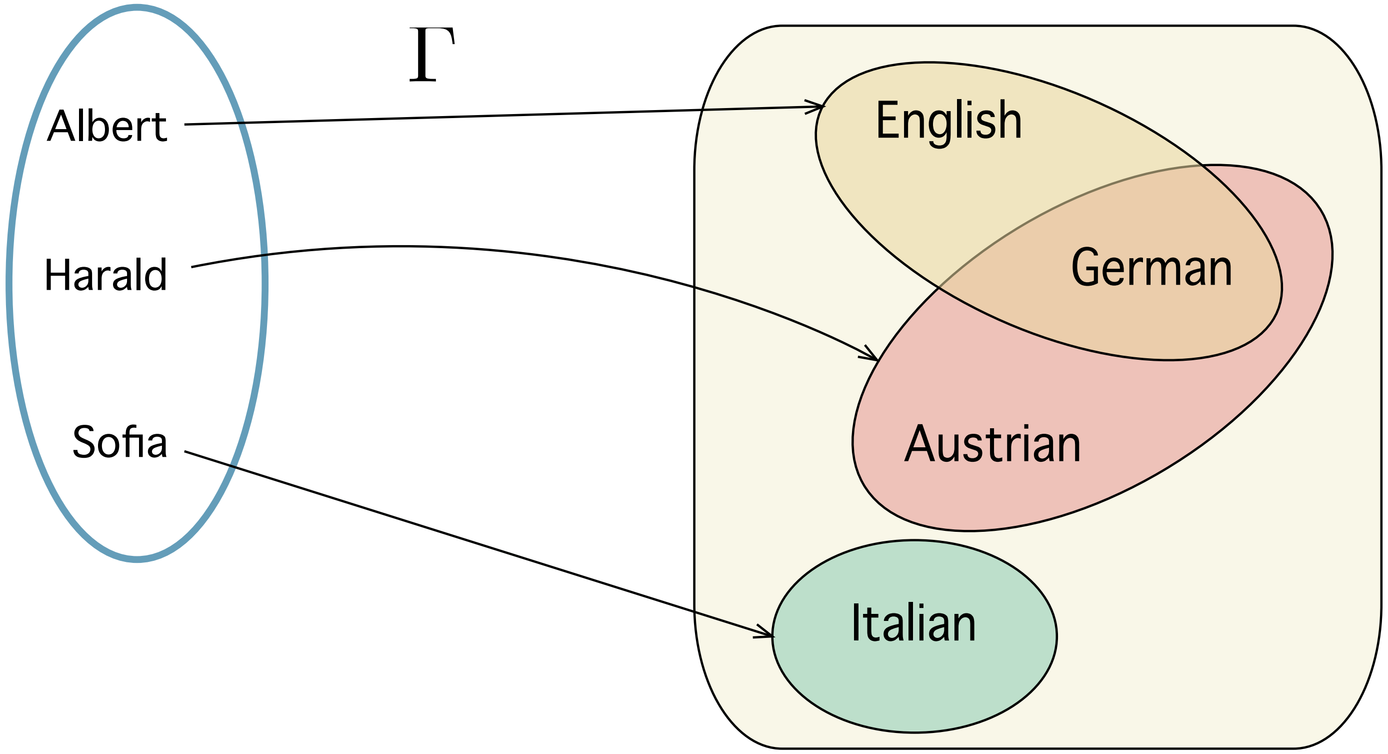
DISJUNCTIVE RANDOM SETS: MAIN NOTIONS

Family of (measurable) selections

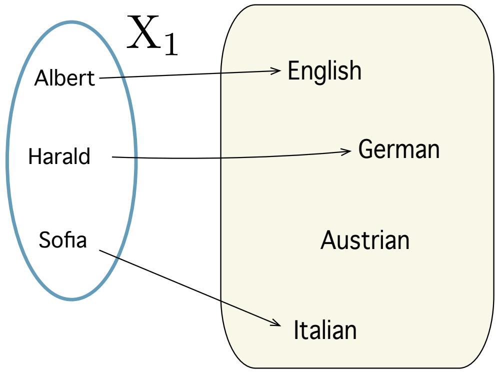
Imprecise info about parameters: family of feasible probabilities, expectations, variances..

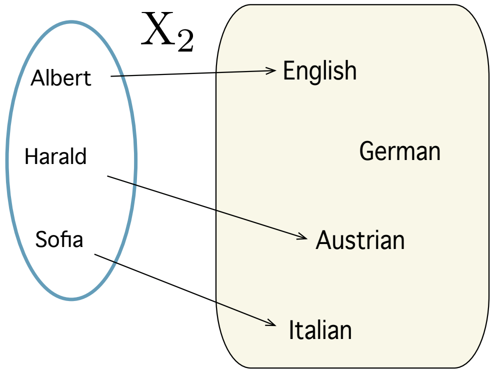
And the notion of independence?

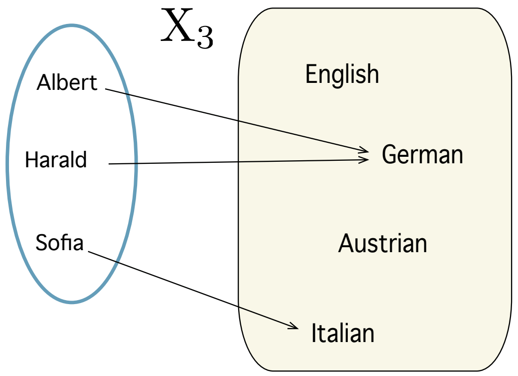
DISJUNCTIVE RANDOM SETS: FAMILY OF MEASURABLE SELECTIONS

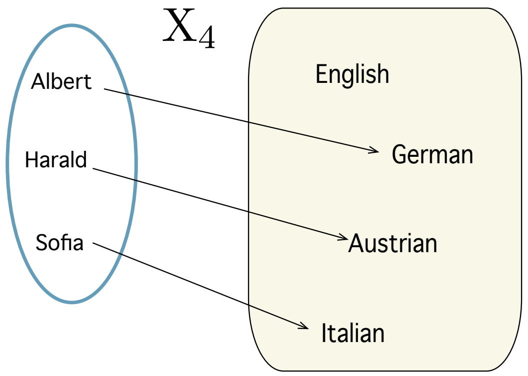


$$S(\Gamma) = \{X_1, X_2, X_3, X_4\}$$





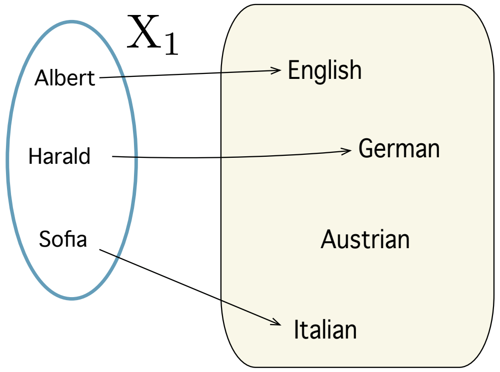




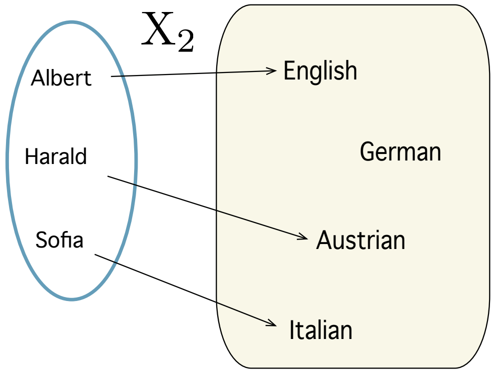
DISJUNCTIVE RANDOM SETS: IMPRECISE INFO ABOUT PARAMETERS

$$\theta_\Gamma = \{\theta_X \, : \, X \in S(\Gamma)\}$$

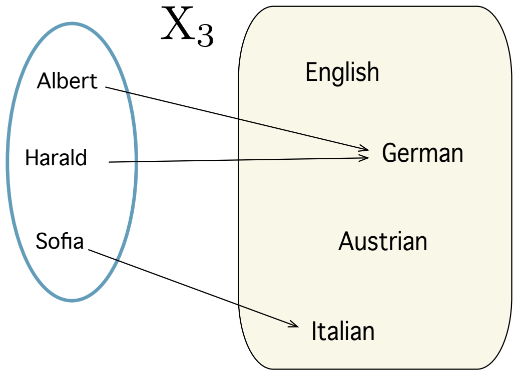
Examples: Aumann expectation, Kruse variance, etc.



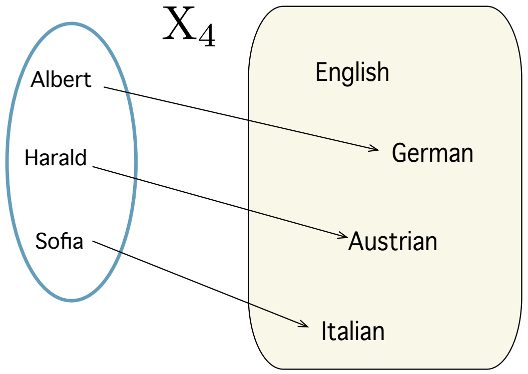
$$\mathrm{P}_{X_1}(\{g\}) = 1/3$$



$$\mathrm{P}_{X_2}(\{g\}) = 0$$



$$\mathrm{P}_{X_3}(\{g\}) = 2/3$$



$$\mathrm{P}_{X_4}(\{g\}) = 1/3$$

$$\mathrm{P}_\Gamma = \{P_{X_1}, P_{X_2}, P_{X_3}, P_{X_4}\}$$

$$P_\Gamma(\{g\}) = \{P_{X_1}(\{g\}), P_{X_2}(\{g\}), P_{X_3}(\{g\}), P_{X_4}(\{g\})\} = \{0, 1/3, 2/3\}$$

- R. Aumann (1965), R. Kruse (1987), De Meyer & Kruse (1990), I. Couso & L. Sánchez (2008).

SETS OF PROBABILITIES ASSOCIATED TO A RANDOM SET

- Set of probability measures induced by the selections:

$$\mathcal{P}(\Gamma) = \{P_X : X \in S(\Gamma)\},$$

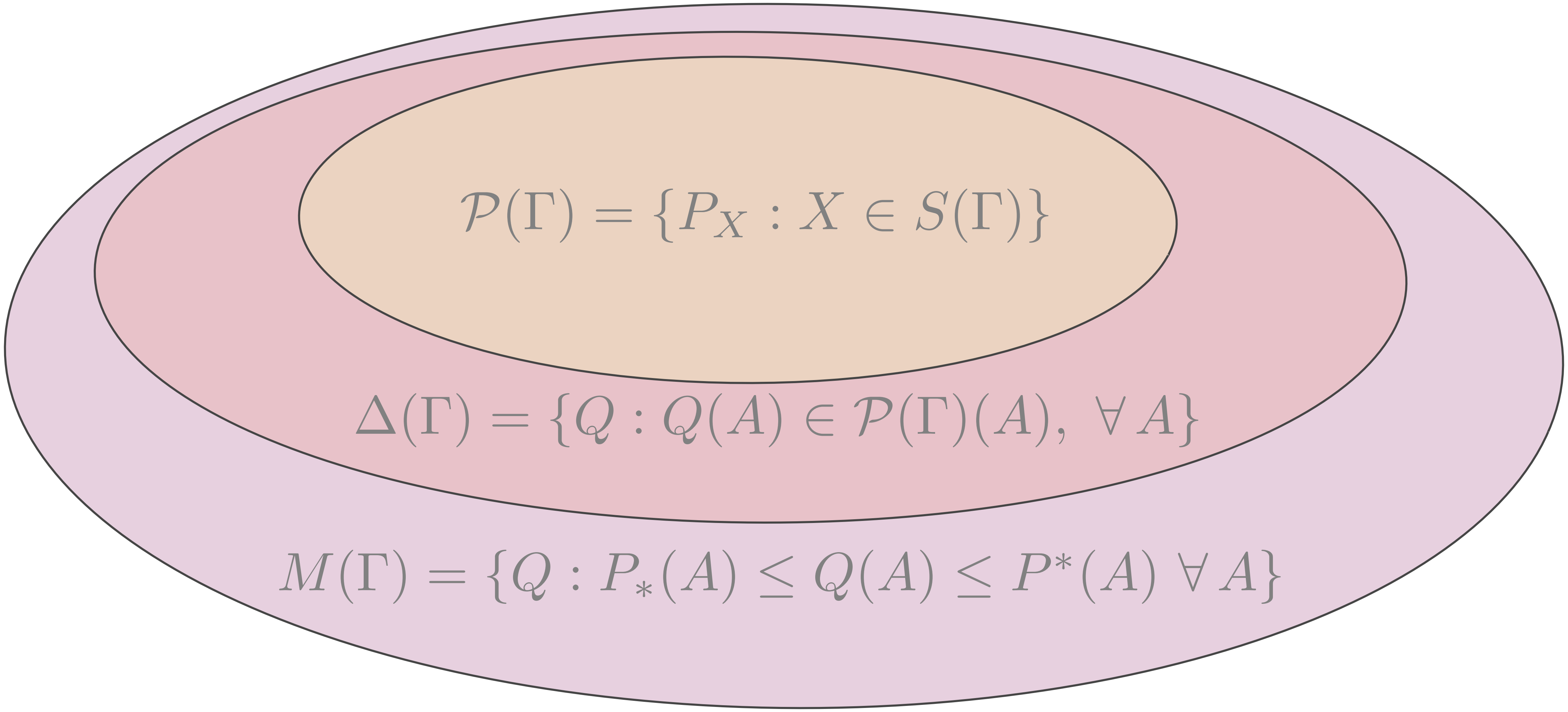
where $S(\Gamma)$ denotes the set of measurable selections of Γ.

- Credal set:

$$\mathcal{M}(P^*) = \{P : P \text{ is dominated by } P^*\} = \{P : P \text{ dominates } P_*\}$$

- $\mathcal{P}(\Gamma) \subseteq \mathcal{M}(P^*)$.

DISJUNCTIVE RANDOM SETS: ASSOCIATED SETS OF PROBABILITIES

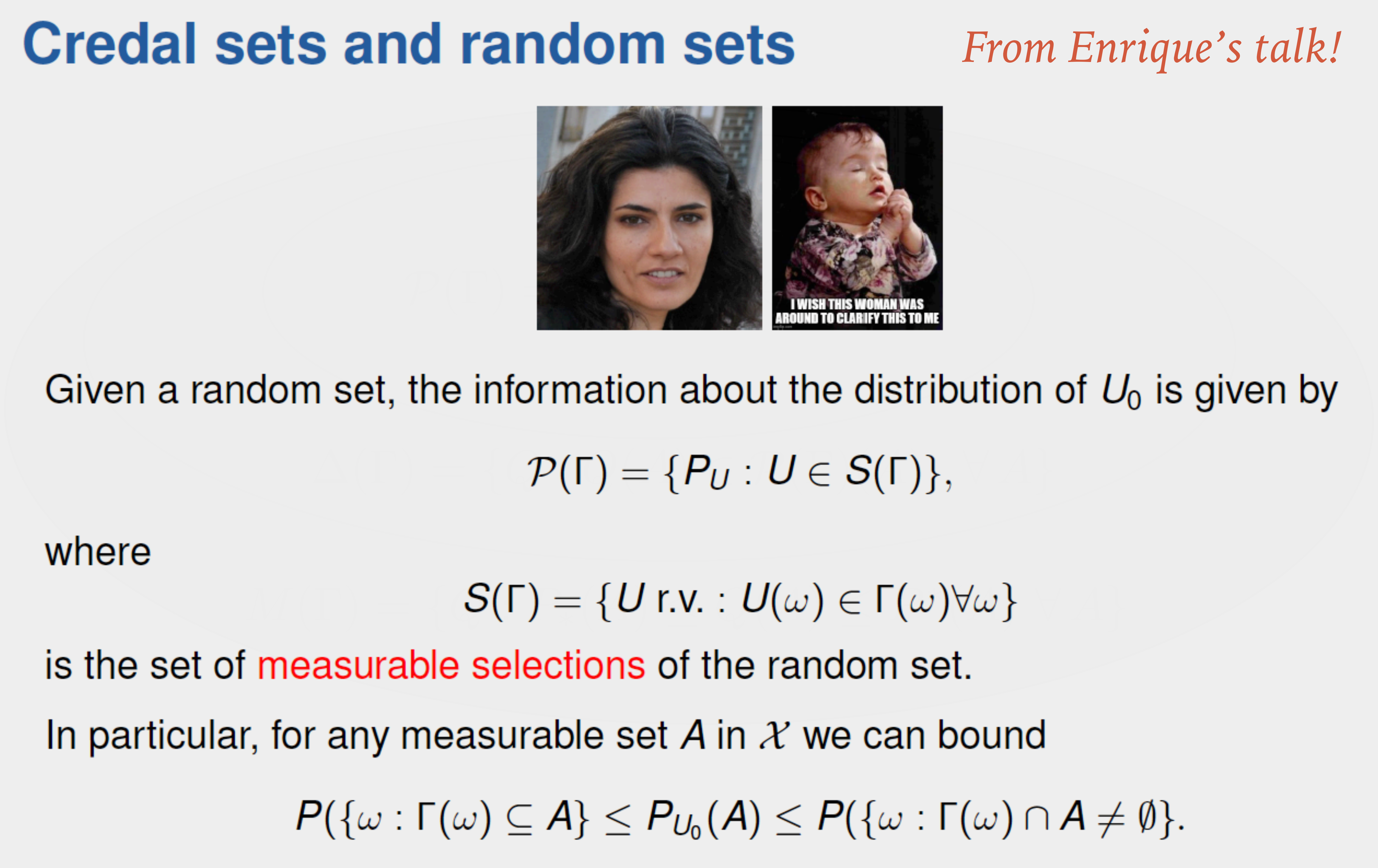


$$\mathcal{P}(\Gamma) = \{P_X : X \in S(\Gamma)\}$$

$$\Delta(\Gamma) = \{Q : Q(A) \in \mathcal{P}(\Gamma)(A),\ \forall A\}$$

$$M(\Gamma) = \{Q : P_*(A) \leq Q(A) \leq P^*(A)\ \forall A\}$$

DISJUNCTIVE RANDOM SETS: ASSOCIATED SETS OF PROBABILITIES



From Enrique's talk!

Credal sets and random sets

Given a random set, the information about the distribution of U_0 is given by

$$\mathcal{P}(\Gamma) = \{P_U : U \in S(\Gamma)\},$$

where

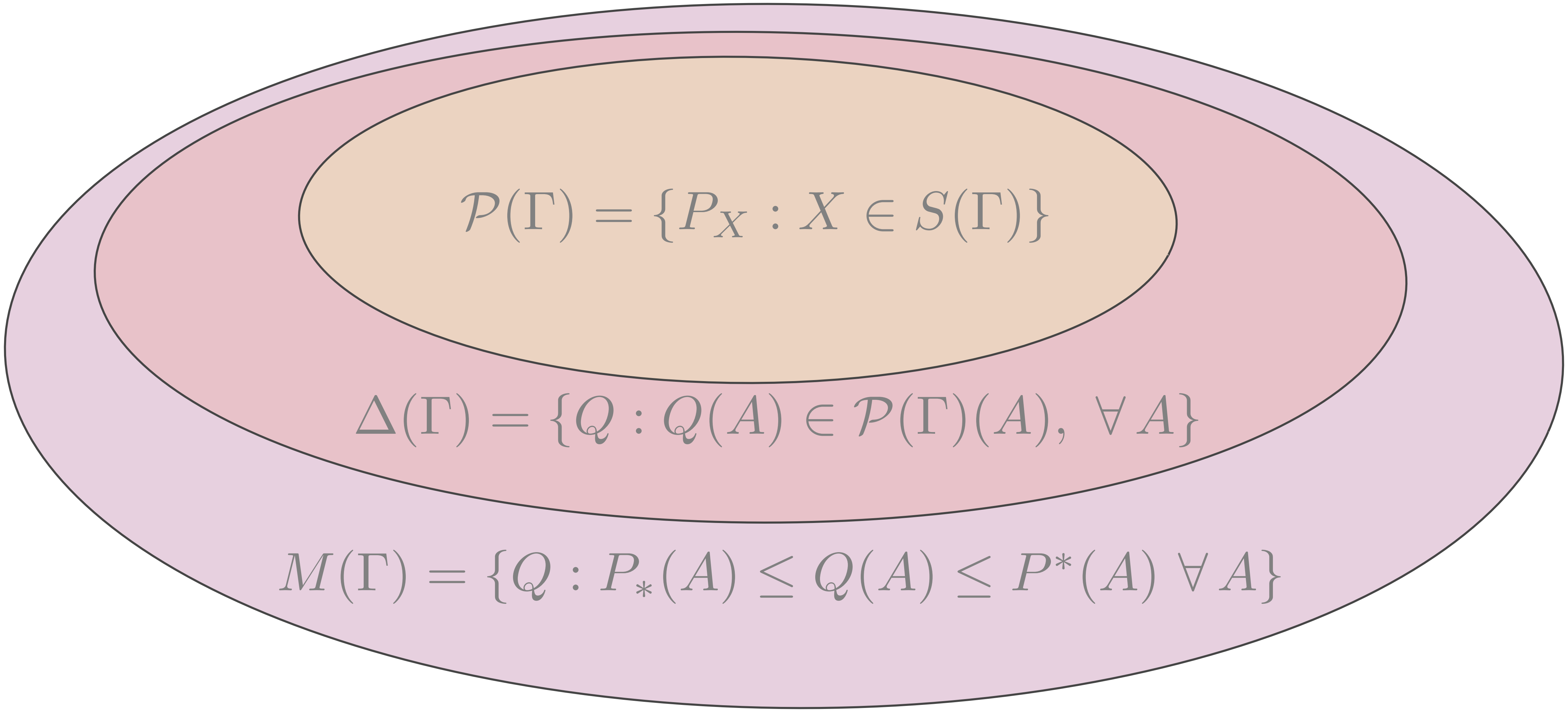
$$S(\Gamma) = \{U \text{ r.v.} : U(\omega) \in \Gamma(\omega) \forall \omega\}$$

is the set of measurable selections of the random set.

In particular, for any measurable set A in $\mathcal{X}$ we can bound

$$P(\{\omega : \Gamma(\omega) \subseteq A\} \leq P_{U_0}(A) \leq P(\{\omega : \Gamma(\omega) \cap A \neq \emptyset\}.$$

DISJUNCTIVE RANDOM SETS: ASSOCIATED SETS OF PROBABILITIES



$$\mathcal{P}(\Gamma) = \{P_X : X \in S(\Gamma)\}$$

$$\Delta(\Gamma) = \{Q : Q(A) \in \mathcal{P}(\Gamma)(A),\ \forall A\}$$

$$M(\Gamma) = \{Q : P_*(A) \leq Q(A) \leq P^*(A)\ \forall A\}$$

DIFFERENCES BETWEEN THE CONJUNCTIVE AND THE DISJUNCTIVE APPROACHES

- A random set induces a probability distribution over the power set of the final universe.
- No remarkable practical differences between the conjunctive and the disjunctive expectations of a random set (both are set valued).
- Important differences between the conjunctive (scalar, i.e. numerical) and the disjunctive (set-valued) variance.
- Important differences between “random set independence” and “strong independence”.

DIFFERENCES BETWEEN THE CONJUNCTIVE AND THE DISJUNCTIVE APPROACHES: THE VARIANCE

- **Situation 1:** There is a check inside a box. You are told it is between €20 and €60. You open the box and take the check.
- **Situation 2:** You flip a coin. If it lands on “heads”, you receive the money that is inside Box A. Otherwise, you receive the money which is inside Box B. You are told that both quantities are between €20 and €60.

Both random sets induce the same (degenerate) distribution. Their respective scalar variances are null. But our knowledge about the variance of the price of the gift is different in each case.

EXAMPLE. CALCULATION OF SCALAR VARIANCE

- The set $\Omega = \{\omega_1, \ldots, \omega_4\}$ comprises four objects, whose actual weights are $X_0(\omega_1) = 10.2$, $X_0(\omega_2) = 10.0$, $X_1(\omega_3) = 10.4$, $X_1(\omega_1) = 9.7$
- We sense the weights with a digital device that rounds the measure to the nearest integer, and displays the value ‘10’ in all of these cases.

- The variance of the above random variable is strictly positive.
- The random set representing our information about the weights is deterministic (the constant set [0.5, 10.5]).
- What is the scalar variable of the random set? And what do we know about the variance of the variable if we just take into account the information from the digital scale?

EXAMPLE. CALCULATION OF SCALAR VARIANCE

- Suppose that four objects $\omega_1, \ldots, \omega_4$ weigh the same:
- $X_0(\omega_1) = X_0(\omega_2) = X_0(\omega_3) = X_0(\omega_4) = 9.8g.$
- For some reason, the weight of the fourth object was imprecisely measured, and we only know that it is between the values 9.5 and 10.5. The rest are measured with precision.

- The random variable X_0 (true weight) is a constant.
- The multivalued mapping indicating our information about the weight is not a constant.
- What is the (true) variance of X_0? What is the scalar variance of the random set? What do we know about the variance of X_0 if we just have the information given by the random set?

BUT WHAT IS THE PROBLEM WITH THE VARIANCE?

- Nomenclature: $\mathbf{l} = (l_1, \ldots, l_n)$ and $\mathbf{u} = (u_1, \ldots, u_n)$.
- We can easily determine bounds for $\overline{\mathbf{x}}$ and median($\mathbf{x}$).
  - Mean: $\bar{\mathbf{l}} \leq \overline{\mathbf{x}} \leq \overline{\mathbf{u}}$.
  - Median: $\text{median}(\mathbf{l}) \leq \text{median}(\mathbf{x}) \leq \text{median}(\mathbf{u})$.
  - Variance: $\min\{s^2_{\mathbf{l}}, s^2_{\mathbf{u}}\} \leq s^2_{\mathbf{x}} \leq \max\{s^2_{\mathbf{l}}, s^2_{\mathbf{u}}\}$?

(The mean and the median are comonotonic operators, while the variance is not.)

THE VARIANCE IN AN EPISTEMIC SETTING

- The upper and lower bounds of the variance cannot be written in terms of the respective variances of **l** and **u** in general.
- We need to solve the problem:

$$\text{Calculate } \max[\overline{y^2} - (\overline{y})^2] \text{ and } \min[\overline{y^2} - (\overline{y})^2]$$

$$\text{Subject to: } l_i \leq y_i \leq u_i,\ i = 1, \ldots, n.$$

- Ferson, S., Ginzburg, L., Kreinovich, V., Longpré, L., Aviles, M. (2002). Computing variance for interval data is NP-hard. ACM SIGACT News, 33(2), 108-118.
- Dubois, D., Fargier, H., & Fortin, J. (2005, May). The empirical variance of a set of fuzzy intervals. In The 14th IEEE International Conference on Fuzzy Systems, 2005. FUZZ'05. (pp. 885-890).
- Salamanca, J. J., Couso, I. (2022). The minimum variance of a random set on a Euclidean space. Fuzzy Sets and Systems, 443, 106-126.

DIFFERENCES BETWEEN THE CONJUNCTIVE AND THE DISJUNCTIVE APPROACHES: THE NOTION OF INDEPENDENCE

- It may happen that two independent random sets represent incomplete information about two dependent random variables.
- It may also happen that two dependent random sets represent incomplete information about two independent random variables.

- I. Couso, S. Moral, P. Walley (2000) A survey of concepts of independence for imprecise probabilities, Risk Decision and Policy 5, 165-181.
- I. Couso, S. Moral (2010) Independence concepts in Evidence Theory, IJAR 51, 748-758.

INDEPENDENT RANDOM SETS AND DEPENDENT RANDOM VARIABLES

- X and Y respectively represent the temperature (in °C) of an ill person taken at random in a hospital just before taking an antipyretic (X) and 3 hours later (Y).
- A random set represents the information about X provided by a thermometer with +/-0.1°C of precision.
- Another random set represents the information about Y using a very crude measure (it reports always the same interval [37,40.5]).
    - Are X and Y stochastically independent?
    - And what about both random sets?

INDEPENDENT RANDOM VARIABLES AND DEPENDENT RANDOM SETS

- We have a light sensor that displays numbers between 0 and 255.
- We take 10 measurements per second. When the brightness is higher than a threshold (255), the sensor displays the value 255 during 3/10 seconds, regardless the actual brightness value.
- Below we provide data for six measurements:
  - The actual values of brightness represent a realization of a simple random sample of size n = 6.
  - But what about the displayed quantities and our interval-valued information? Are them independent?

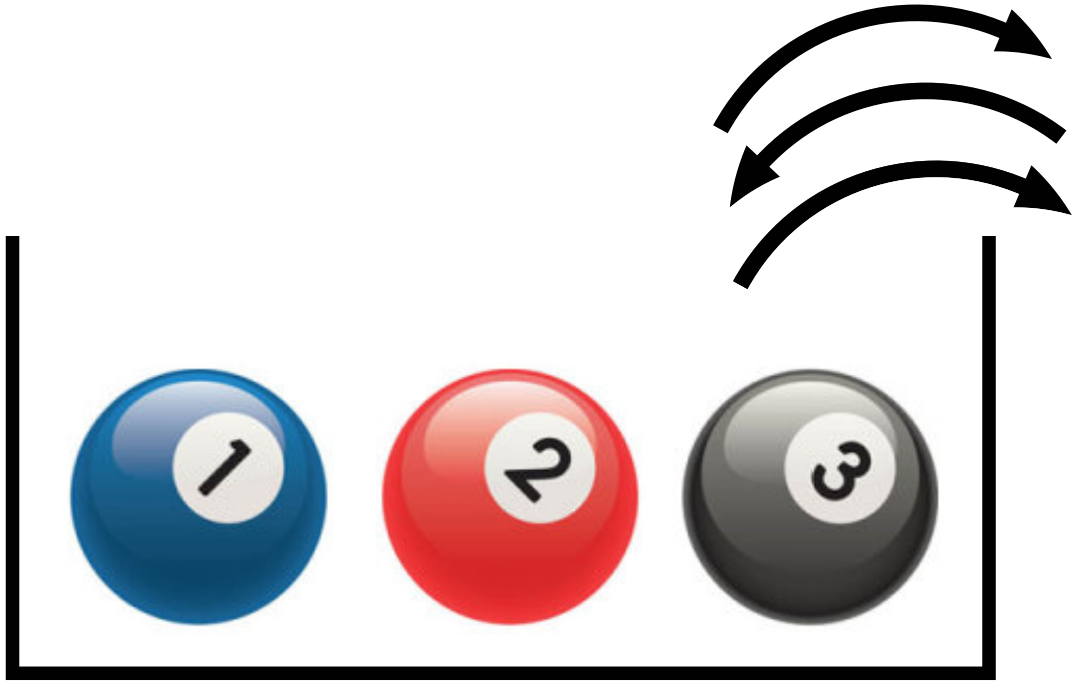
actual values	215	150	200	300	210	280
displayed quantities	215	150	200	255	255	255
set-valued information	$\{215\}$	$\{150\}$	$\{200\}$	$[255,\infty)$	$[0,\infty)$	$[0,\infty)$.

INDEPENDENT RANDOM VARIABLES AND DEPENDENT RANDOM SETS (CONT.)

- The sample of the “true” values of brightness: vector of 6 iid random variables.
- The sequence of random sets representing our incomplete information about it does not satisfy the condition of independence. In fact:

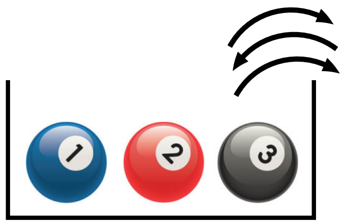
$$P(\Gamma_i \supseteq [255, \infty) | \Gamma_{i-1} \supseteq [255, \infty), \Gamma_{i-2} \not\supseteq [255, \infty)) = 1, \ \forall i \geq 3.$$

INDEPENDENCE UNDER THE EPISTEMIC SETTING: MASS FUNCTIONS ARE NOT ENOUGH



- An urn contains 3 balls numbered from 1 to 3. The first one is blue, and the second one is red. We do not know the color of the third one (blue or red).
- Let us define a random set on {1,2,3} representing our knowledge about the color of a ball drawn at random from the urn.
- We take two balls with replacement. Probability of (blue, red)?

URN: RANDOM SET AND SELECTIONS



- $\Gamma(1)=\{b\ (\text{blue})\}\quad \Gamma(2)=\{r\ (\text{red})\}\quad \Gamma(3)=\{b\ (\text{blue}), r\ (\text{red})\}.$

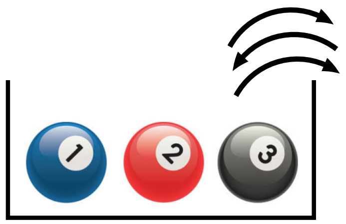
  Set of probabilities associated to the color of one ball drawn from the urn:

$$\mathcal{P}(\Gamma)=\left\{\left(\frac{1}{3},\frac{2}{3}\right),\left(\frac{2}{3},\frac{1}{3}\right)\right\}.$$

- Set of probabilities associated to the colors of two balls independently drawn from the urn (with replacement):

$$\{P\otimes P\ :\ P\in\mathcal{P}(\Gamma)\}=\left\{\left(\frac{1}{9},\boxed{\frac{2}{9}},\frac{2}{9},\frac{4}{9}\right),\left(\frac{4}{9},\boxed{\frac{2}{9}},\frac{2}{9},\frac{1}{9}\right)\right\}.$$

URN: MASS ASSIGNMENT AND PRODUCT MEASURES



- $m(\{b\}) = 1/3$, $m(\{r\} = 1/3$, $m(\{b, r\}) = 1/3$.
- Credal set induced by m:

$$\mathcal{M} = \{P \,:\, 1/3 \leq P(\{b\}) \leq 2/3\}.$$

- What do we know about $(P \times P)(b, r)$ if we just know that $P \in \mathcal{M}$?

$$2/9 \leq \{p^2 \,:\, 1/3 \leq p \leq 2/3\} \leq 1/4$$

Conclusion: We lose relevant information about $P \otimes P$ if we replace $\mathcal{P}(\Gamma)$ by $\mathcal{M}$.

RANDOM SETS ARE NOT JUST MASS ASSIGNMENTS (EVEN IN THE FINITE CASE)

- Consider a random set representing coarse data observations.
- Credal set induced by the lower probability ≈ probability induced on power set ≈ mass assignment (in the finite case).
- None of them characterises the set of probabilities associated to the measurable selections.
- In other words, no one preserves all the information encompassed by the random set about the “original distribution" (i.e. about the true underlying distribution).
- There is additional information included in the initial space that is not preserved.

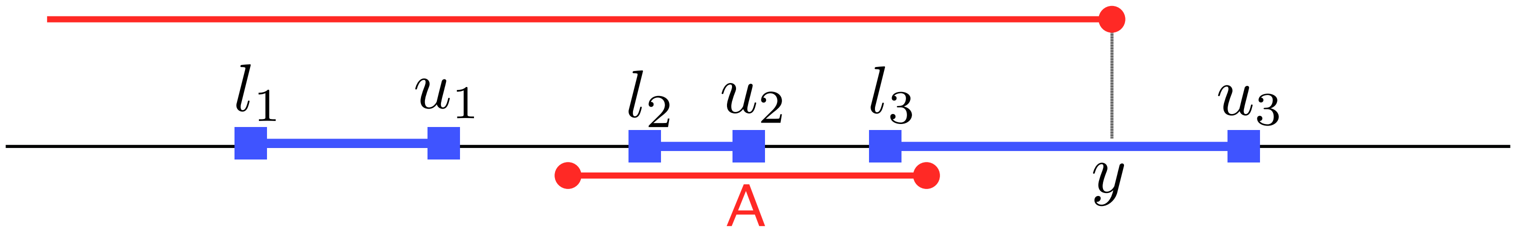
CONJUNCTIVE VS DISJUNCTIVE RANDOM SETS: SUMMARY

	CONJUNCTIVE	DISJUNCTIVE
Probability	Prob. measure on a sigma- field of families of sets	family of (feasible) ordinary probabilities
Expectation	set	set (of feasible expectations)
Variance	number	set (of feasible variances)
Independence	classical notion	strong independence

PAIR OF UPPER AND LOWER CDFS

Upper and lower cumulative distribution functions:

- $\underline{F}(x) = \underline{P}\big((-\infty, x]\big) = P(\Gamma \subset (-\infty, x]),\ \forall\, x \in \mathbb{R}$
- $\overline{F}(x) = \overline{P}\big((-\infty, x]\big) = P(\Gamma \cap (-\infty, x] \neq \emptyset)\ \forall\, x \in \mathbb{R}.$



PAIR OF UPPER AND LOWER CDFS

Example: *p*-boxes

Assume $\mathcal{X}$ is ordered, and let $\underline{F}, \overline{F} : \mathcal{X} \to [0,1]$ be distribution functions, i.e.,

$$x_i \leq x_j \Rightarrow \underline{F}(x_i) \leq \underline{F}(x_j) \text{ and } \overline{F}(x_i) \leq \overline{F}(x_j), \underline{F}(x_n) = \overline{F}(x_n) = 1.$$

The pair $(\underline{F}, \overline{F})$ is called a *p*-box, and the credal set

$$\mathcal{M}(\underline{F}, \overline{F}) = \{P : \underline{F} \leq F_P \leq \overline{F}\}$$

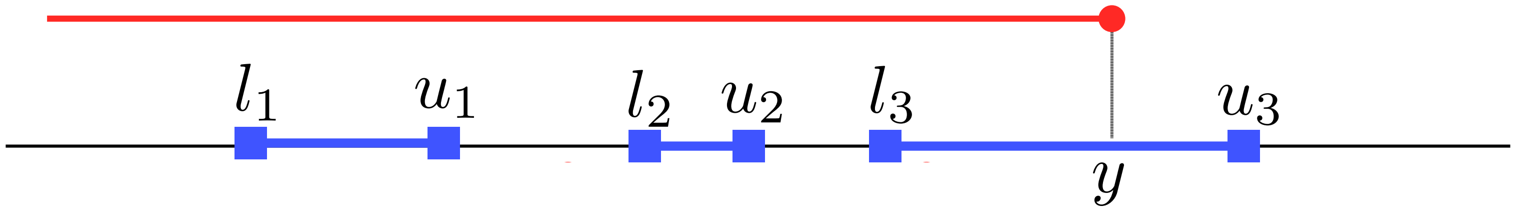
determines coherent lower and upper probabilities $\underline{P}, \overline{P}$ taking lower and upper envelopes.

$\underline{P}$ is a belief function and $\overline{P}$ is a plausibility function.

PAIR OF UPPER AND LOWER CDFS (P-BOX)

Upper and lower cumulative distribution functions:

- $\underline{F}(x) = \underline{P}\big((-\infty, x]\big) = P(\Gamma \subset (-\infty, x]),\ \forall x \in \mathbb{R}$
- $\overline{F}(x) = \overline{P}\big((-\infty, x]\big) = P(\Gamma \cap (-\infty, x] \neq \emptyset)\ \forall x \in \mathbb{R}.$



- I. Couso, S. Montes, P. Gil, Función de distribución y mediana de variables aleatorias difusas, in: Proceedings of ESTYLF'98, Pamplona, Spain, 1998.
- Ferson, S., V. Kreinovich, L. Ginzburg, K. Sentz and D.S. Myers. 2003. Constructing probability boxes and Dempster-Shafer structures. Sandia National Laboratories, SAND2002-4015, Albuquerque, New Mexico.
- Couso, I., Sánchez, L., & Gil, P. (2004). Imprecise distribution function associated to a random set. Information Sciences, 159(1-2), 109-123.
- Destercke, S., Dubois, D., & Chojnacki, E. (2008). Unifying practical uncertainty representations–I: Generalized p-boxes. International Journal of Approximate Reasoning, 49(3), 649-663

THE P-BOX DOES NOT CHARACTERISE THE CREDAL SET OF A RANDOM SET

- The pair of upper and lower probabilities induced by the random set determines a p-box.
- This p-box does not determine the pair of upper and lower probabilities induced by the random set.
- The pair of CDFs does not determine the prob. distribution induced by it.

- Destercke, S., Dubois, D., & Chojnacki, E. (2007). On the relationships between random sets, possibility distributions, p-boxes and clouds. 28th Linz Seminar on Fuzzy Set Theory.

- Destercke, S., Dubois, D., & Chojnacki, E. (2008). Unifying practical uncertainty representations–I: Generalized p-boxes. International Journal of Approximate Reasoning, 49(3), 649-663.

- Destercke, S., & Dubois, D. (2009). The role of generalised p-boxes in imprecise probability models. ISIPTA'09 (pp. 179-188).

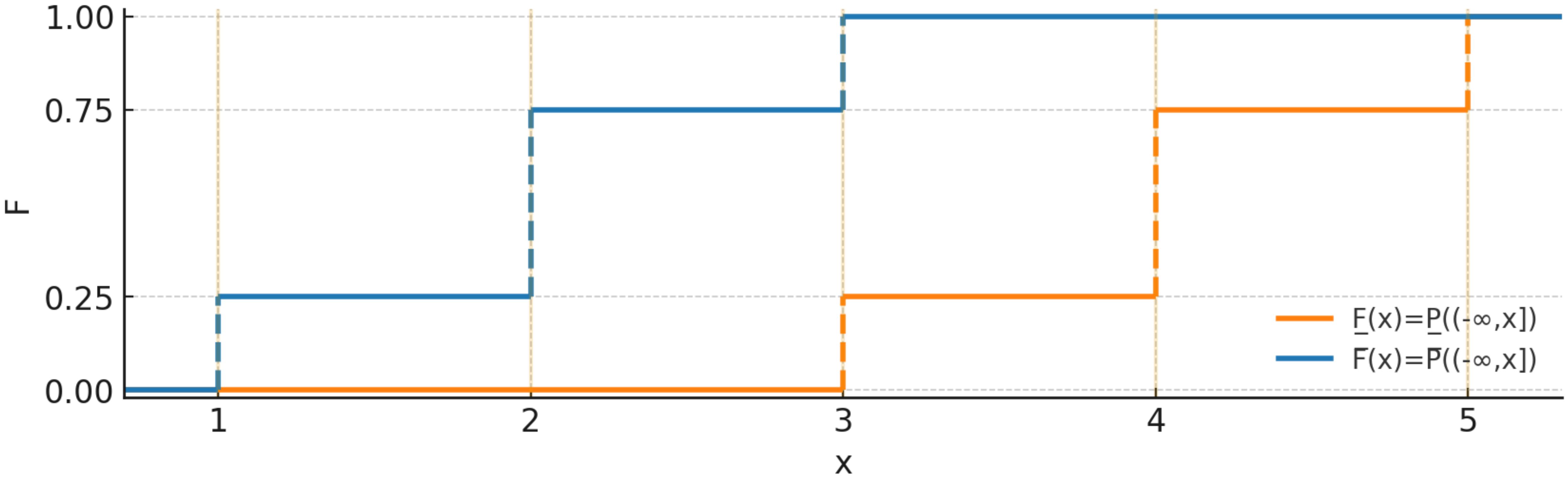
- Additional recent papers by S.Destercke together with M. Troffaes, E. Miranda, I. Montes.

EXAMPLE: P-BOX VS CREDAL SET INDUCED BY A R.S.

Three random intervals inducing the same p-box (same pair $\underline{F}$, $\overline{F}$) :

Ω	$P(\{\omega_i\})$	Γ_1	Γ_2	Γ_3
ω_1	0.25	$[1,5]$	$[3,5]$	$[2,5]$
ω_2	0.25	$[2,4]$	$[2,4]$	$[3,4]$
ω_3	0.25	$[2,4]$	$[2,4]$	$[2,4]$
ω_4	0.25	$\{3\}$	$[1,3]$	$[1,3]$

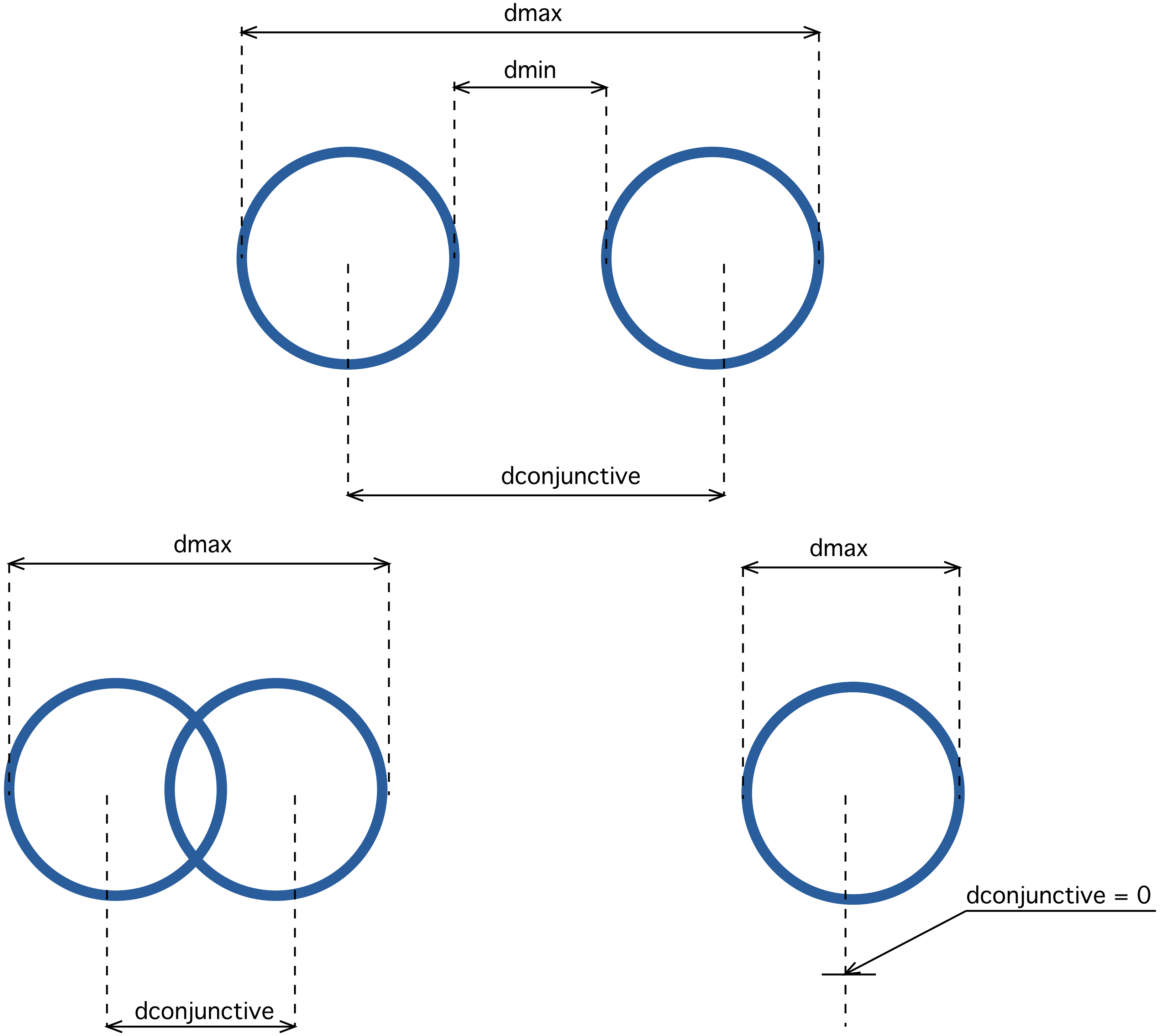
- $\underline{F} = F_{\max \Gamma_i}, \ i = 1, 2, 3.$
- $\overline{F} = F_{\min \Gamma_i}, \ i = 1, 2, 3.$



A LOOK AT STATISTICAL INFERENCE: HYPOTHESIS TESTING

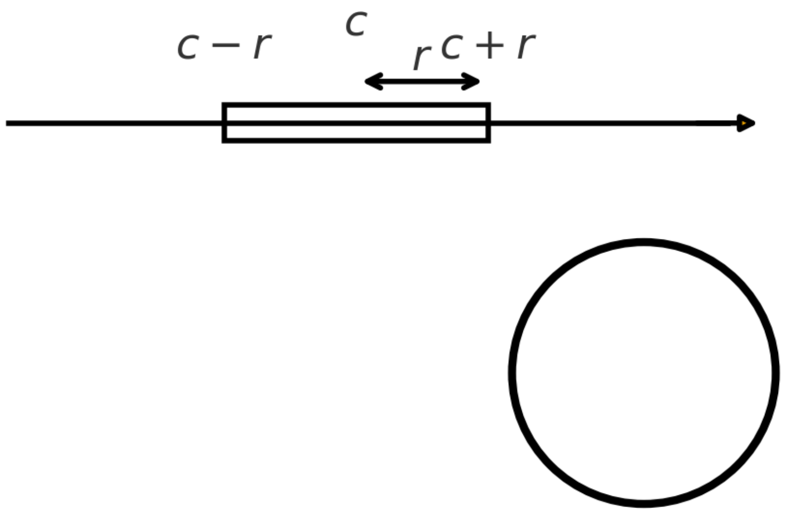
- **Conjunctive view:** classical tests for set-valued parameters. Decisions usually based on numerical distances between sets.
- **Epistemic view:** Tests for point-valued parameters. Incomplete knowledge about the p-value represented by a subset of [0,1]. No sharp decisions.

DISTANCE BTW SETS: CONJUNCTIVE VS DISJUNCTIVE VIEW



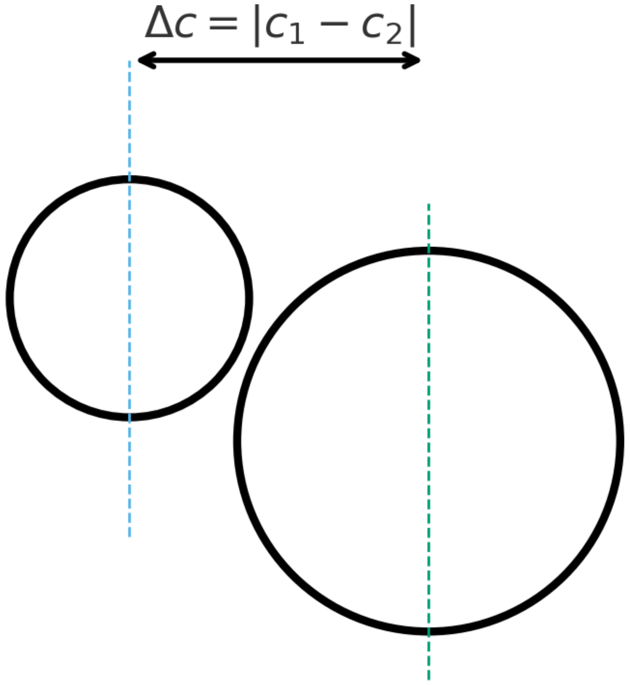
DISTANCE BTW SETS: CONJUNCTIVE VS DISJUNCTIVE VIEW

Interval → Circle



$[c-r,\, c+r] \;\mapsto\; \mathrm{circle}(c, r)$

Two Circles & Distance

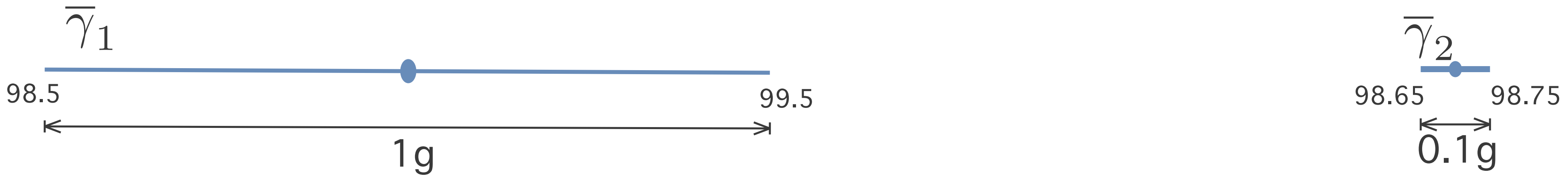


$\Delta r = |r_1 - r_2|$

$$d_{\text{conjunctive}} = \sqrt{(\Delta c)^2 + (\Delta r)^2}$$

CLASSICAL/CONJUNCTIVE HYPOTHESIS TESTING DOES NOW WORK WITH IMPRECISE DATA

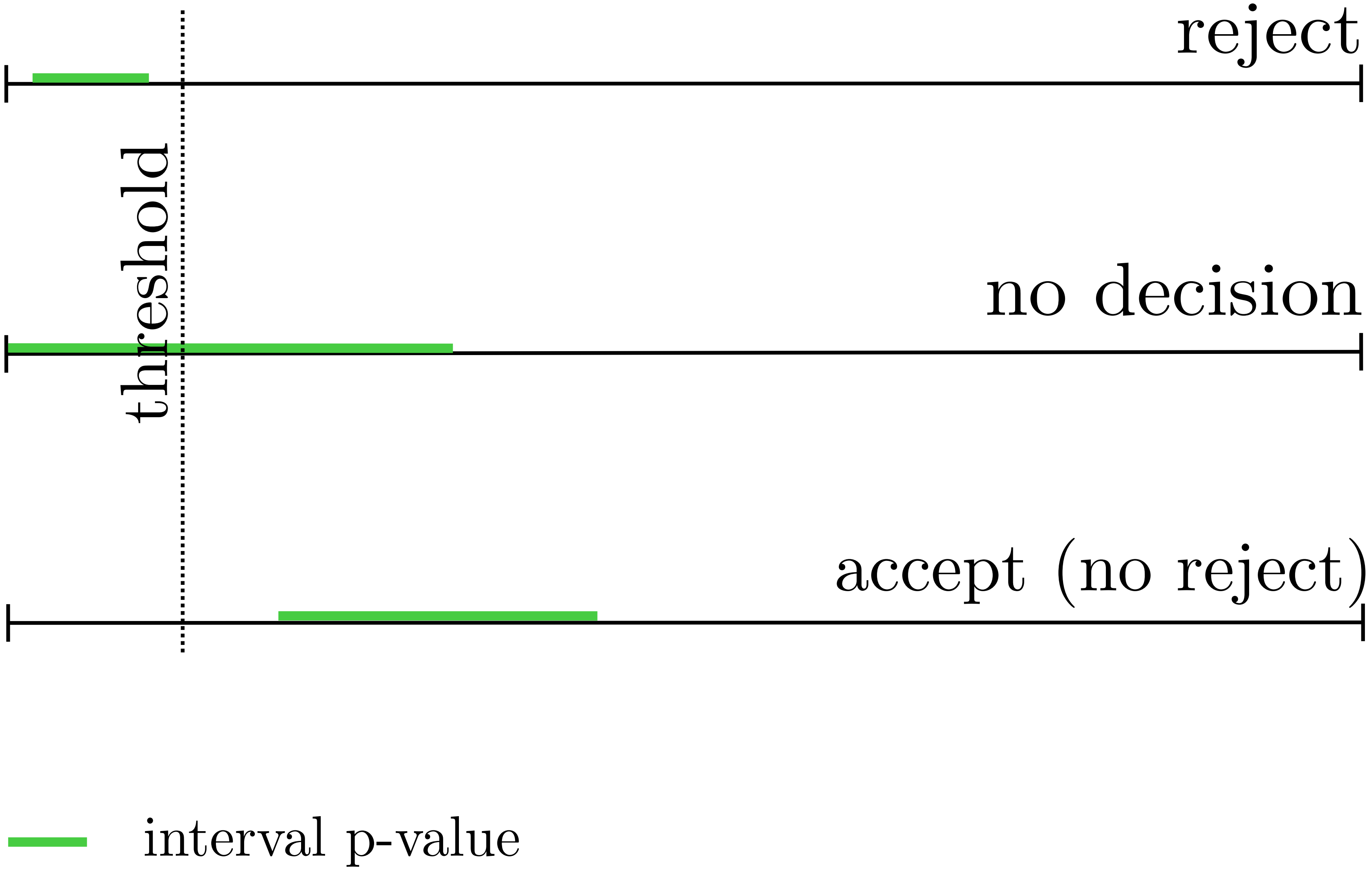
	Scale 1	Scale 2
apple 1:	[98.5, 99.5]	[99.15, 99.25]
apple 2:	[96.5, 97.5]	[96.75, 97.85]
apple 3:	[102.5, 103.5]	[102.45, 103.55]
apple 4:	[99.5, 100.5]	[99.65, 100.75]
	...	...



(Ontic) statistic based on the distance between both set-valued sample means, and the dispersion of set-valued differences. Big distance, small dispersion.

Decision: we reject $H_0 : E(\Gamma_1) = E(\Gamma_2)$.
Same population of apples. What happened?

HYPOTHESIS TESTING WITH COARSE DATA: SET-VALUED P-VALUES AND IMPRECISE TESTS


reject

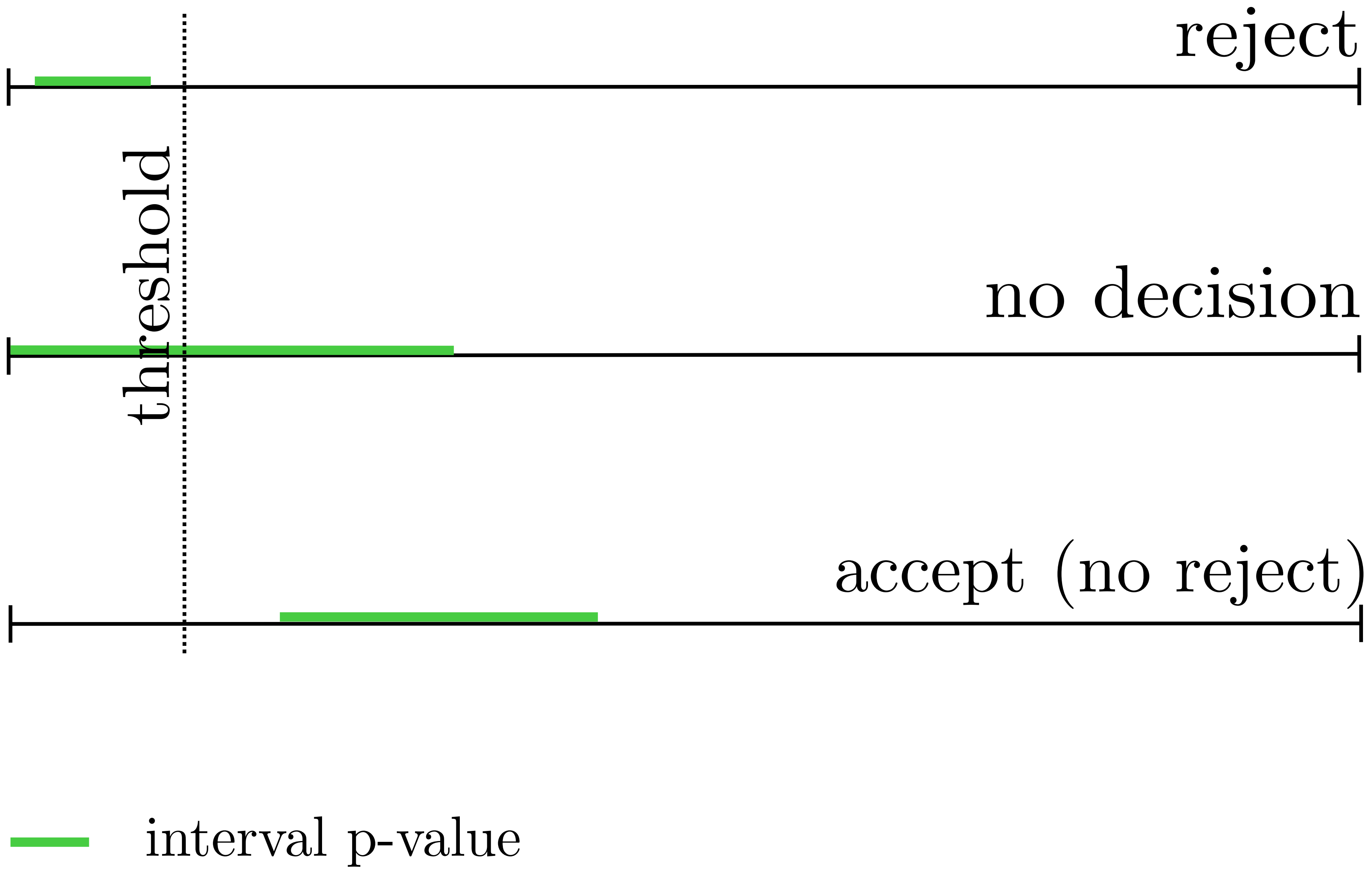
threshold

no decision

accept (no reject)

🟩 interval p-value

HYPOTHESIS TESTING WITH COARSE DATA: SET-VALUED P-VALUES AND IMPRECISE TESTS



Marina Iturrate's poster (U. Oviedo) yesterday was about this!

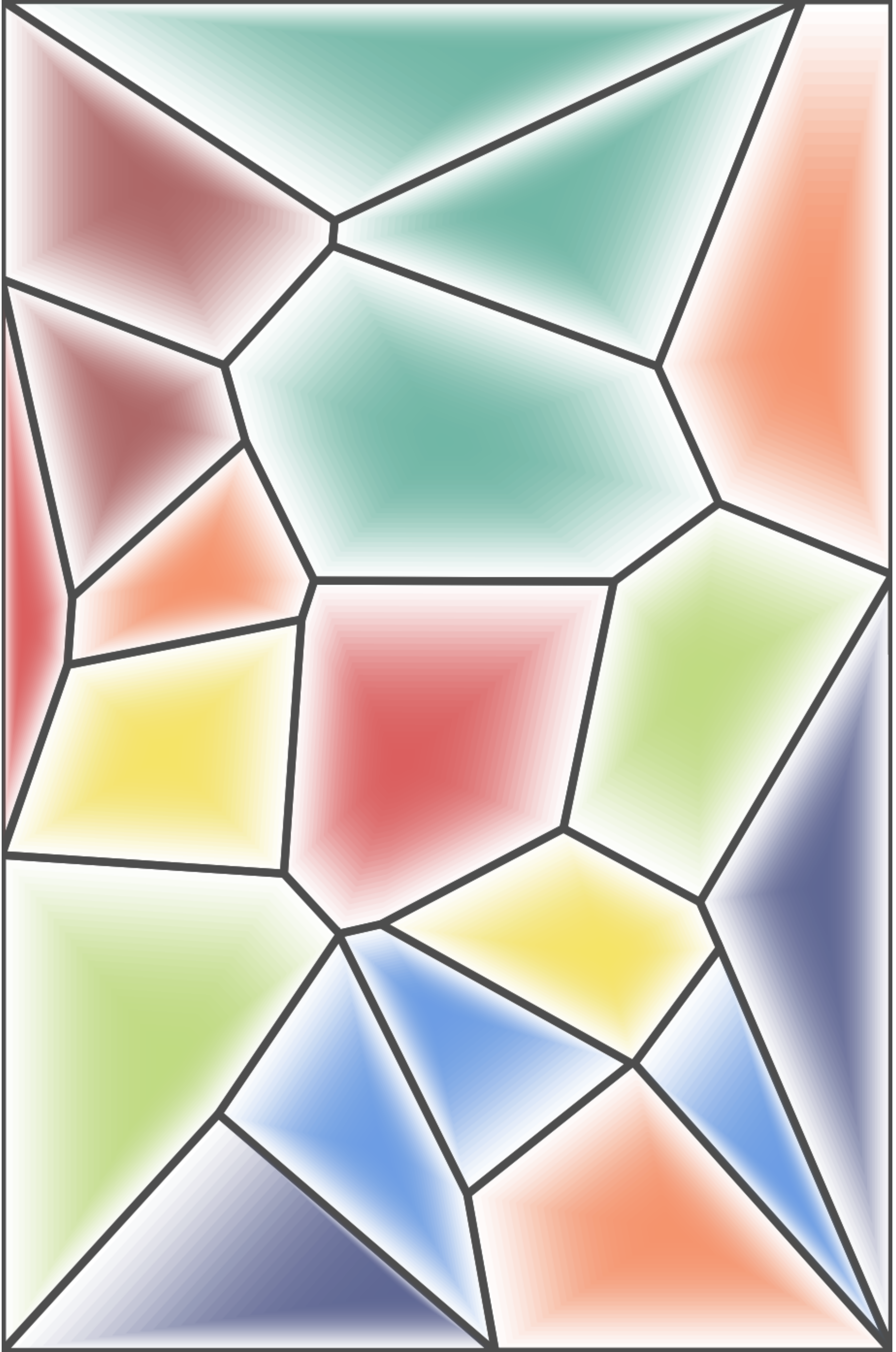
(Different normality tests from interval data)

CLOSED EXPRESSIONS FOR THE INTERVAL P-VALUE

……

Solve an optimization problem to calculate the minimum and maximum p-values associated to the set-valued observations.

- Destercke, S., & Strauss, O. (2014). Kolmogorov-Smirnov test for interval data. In International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems (pp. 416-425).
- Perolat, J., Couso, I., Loquin, K., & Strauss, O. (2015). Generalizing the Wilcoxon rank-sum test for interval data. International Journal of Approximate Reasoning, 56, 108-121.
- Couso, I., Strauss, O., & Saulnier, H. (2018). Kendall's rank correlation on quantized data: An interval-valued approach. Fuzzy Sets and Systems, 343.



FUZZY RANDOM VARIABLES

- Ontic vs epistemic interpretation of fuzzy sets
- Fuzzy random variables as possibility distributions over the set of random variables: The acceptability function
- Incomplete information about parameters, statistics, induced probability…
- Statistics with epistemic fuzzy random variables

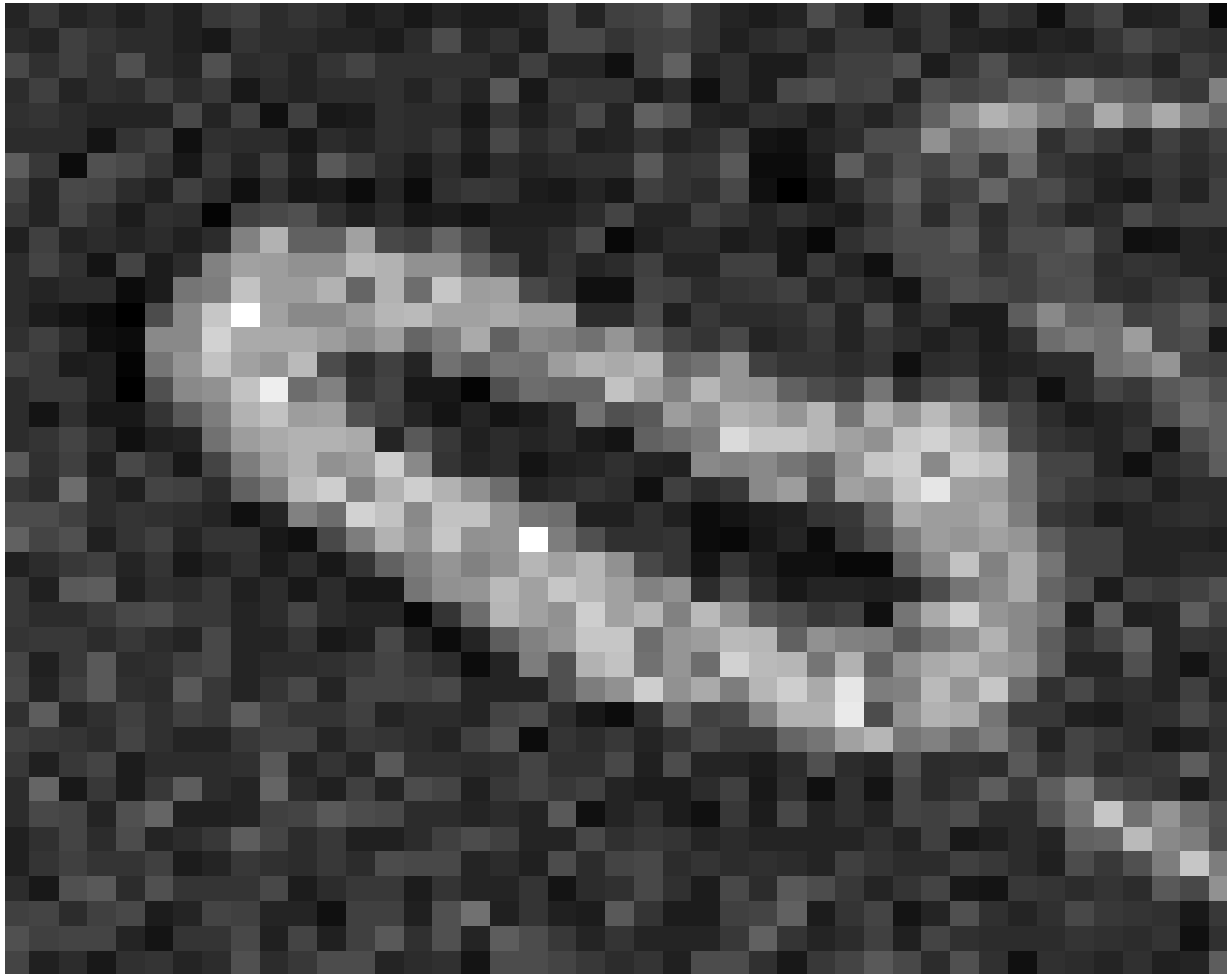
ONTIC VS EPISTEMIC FUZZY SETS

- **What is a fuzzy set?**
  - It is just a mapping between the universe and [0,1].
- **What does it represent?**
  - Whatever that can be represented by such a mapping.
- **But more concretely?**
  - We may distinguish two main interpretations: the "ontic" and the "epistemic".

- D. Dubois, H. Prade (2012), Gradualness, uncertainty and bipolarity: making sense of fuzzy sets FSS 192, 3-24.
- D. Dubois, H. Prade (1997), The three semantics of fuzzy sets, FSS 90, 141-150.

THE ONTIC INTERPRETATION: IMAGE EXAMPLE

The fuzzy set represents precise information about a complex entity.

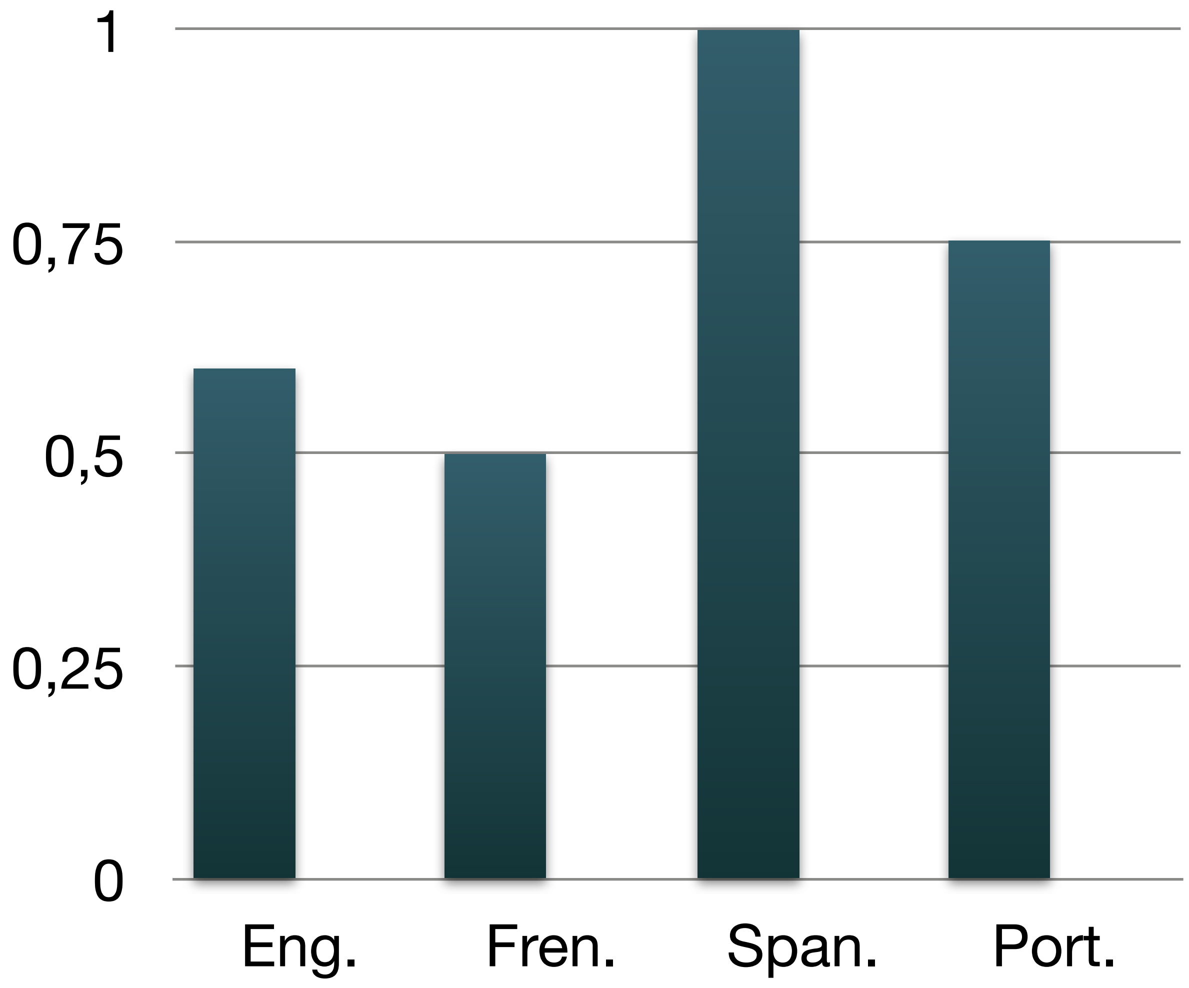


A fuzzy set represents this image. Each pixel location is assigned a level of grey.

Image taken from N. Sladoje, "On Analysis of Discrete Spatial Fuzzy Sets in 2 a 3 Dimensions", PhD Thesis, 2005.

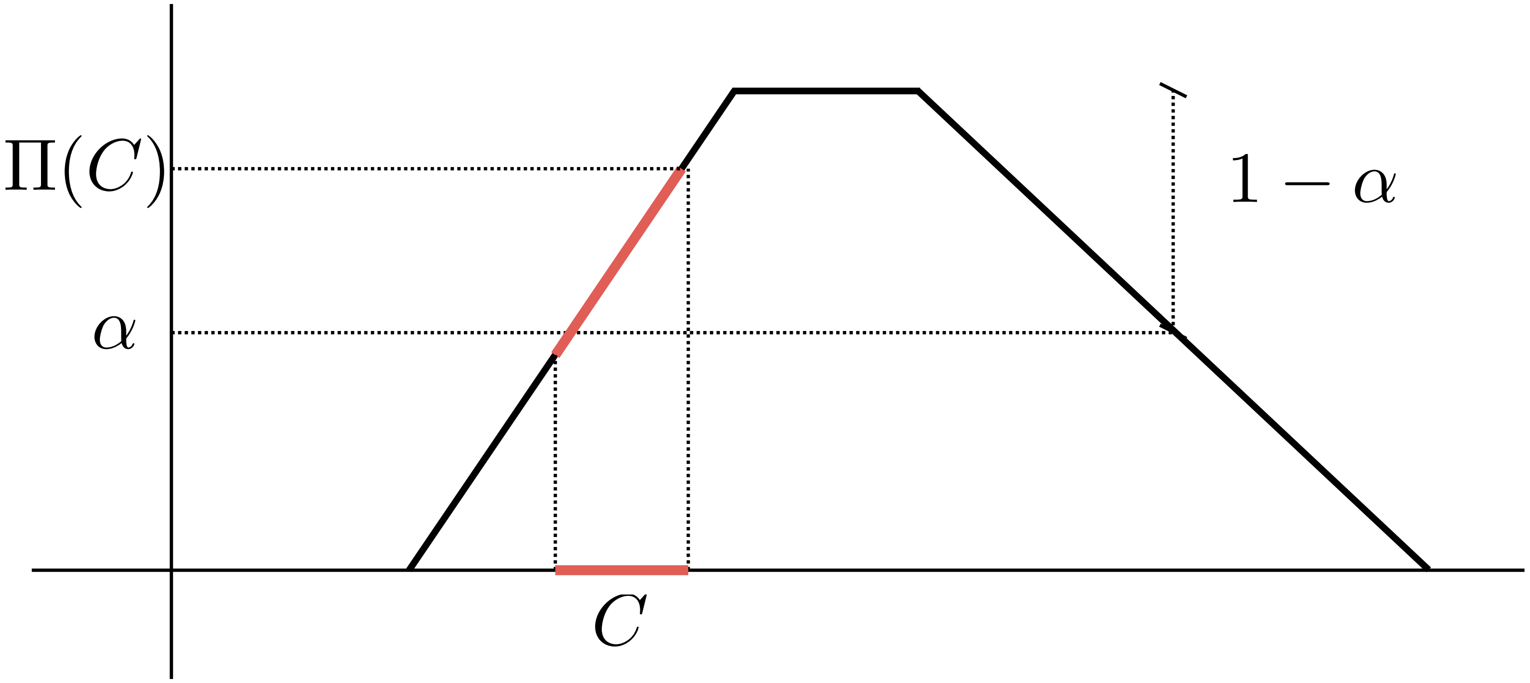
THE ONTIC INTERPRETATION: LANGUAGES EXAMPLE

The fuzzy set represents precise information about a complex entity.



A fuzzy set represents my (optimistic) communication skills in different languages

EPISTEMIC VIEW: POSSIBILITY MEASURE



$$\Pi(C) = \sup_{x\in C} \mu(x)$$

$$P(A_\alpha) \geq 1-\alpha,\ \forall\, \alpha$$

$$P(C) \leq \Pi(C),\ \forall\, C$$

Couso et al. (2001). The necessity of strong alpha-cuts, IJUFKS 9, 249-262.

D. Dubois et al. (2004). Probability-possibility transformations, Rel. Comput. 10, 273-297.

Ryan Martin Talk from yesterday

- Imprecise prob corresponds to sets of probabilities
- This (non-empty) set is called the *credal set*

$$\mathscr{C}(\overline{\Pi}_y) = \{\mathsf{Q}_y \in \mathsf{probs}(\mathbb{T}) : \mathsf{Q}_y(\cdot) \leq \overline{\Pi}_y(\cdot)\}$$

- For possibility measures $\overline{\Pi}_y$, there's a characterization:[10]

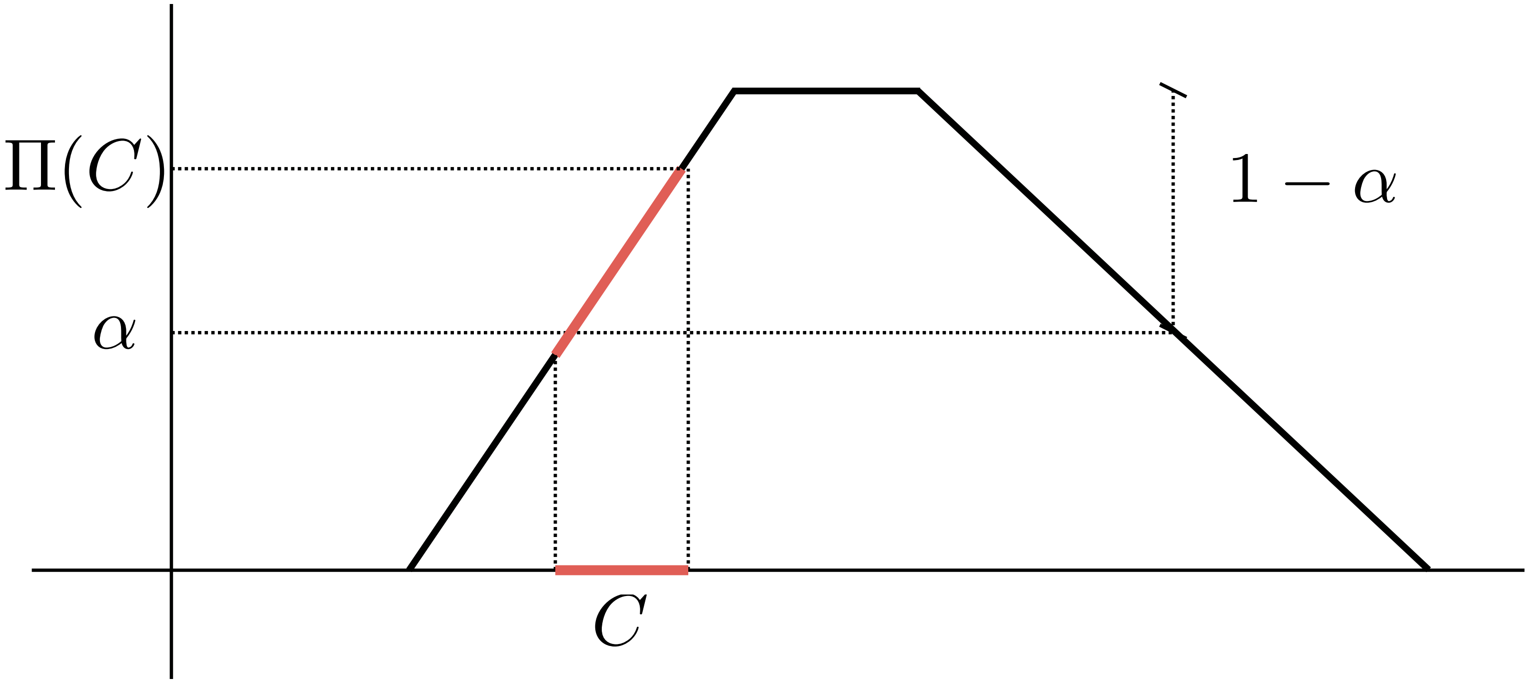
$$\mathsf{Q}_y \in \mathscr{C}(\overline{\Pi}_y) \iff \mathsf{Q}_y(\, \underbrace{\{\theta : \pi_y(\theta) > \alpha\}}_{100(1-\alpha)\%\text{ pl region}} \,) \geq 1 - \alpha$$

- Elements of $\mathscr{C}(\overline{\Pi}_y)$ are confidence distributions — they assign $1 - \alpha$ probability to $100(1 - \alpha)\%$ confidence sets

---

[10] e.g., Destercke & Dubois, Ch. 4 of *Intro to IP*

EPISTEMIC VIEW: POSSIBILITY MEASURE



$$\Pi(C) = \sup_{x \in C} \mu(x)$$

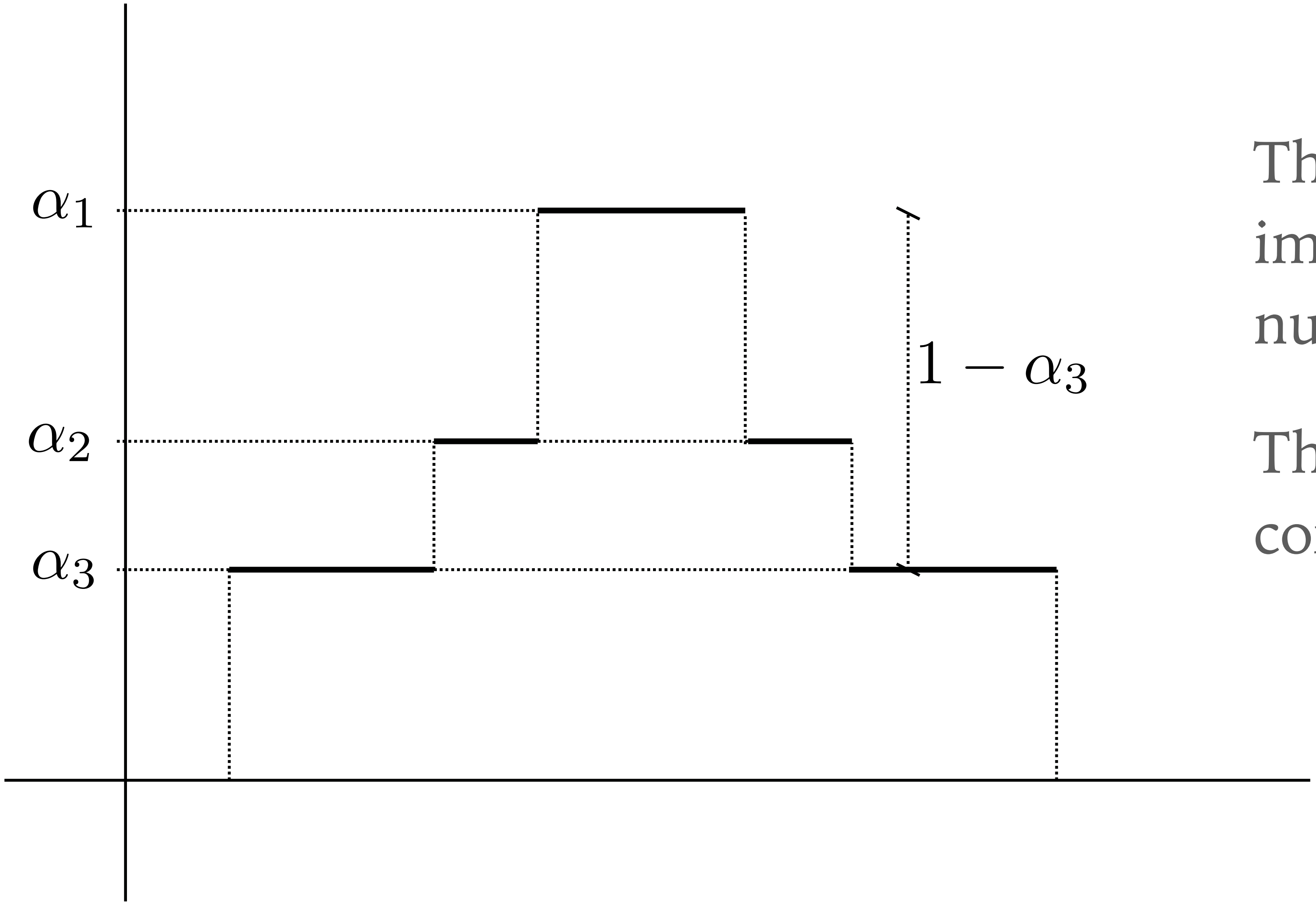
$$P(A_\alpha) \geq 1 - \alpha,\ \forall \alpha$$

$$P(C) \leq \Pi(C),\ \forall C$$

Couso et al. (2001). The necessity of strong alpha-cuts, IJUFKS 9, 249-262.

D. Dubois et al. (2004). Probability-possibility transformations, Rel. Comput. 10, 273-297.

THE EPISTEMIC INTERPRETATION: NESTED CONFIDENCE REGIONS



This fuzzy set represents imprecise info about a number.

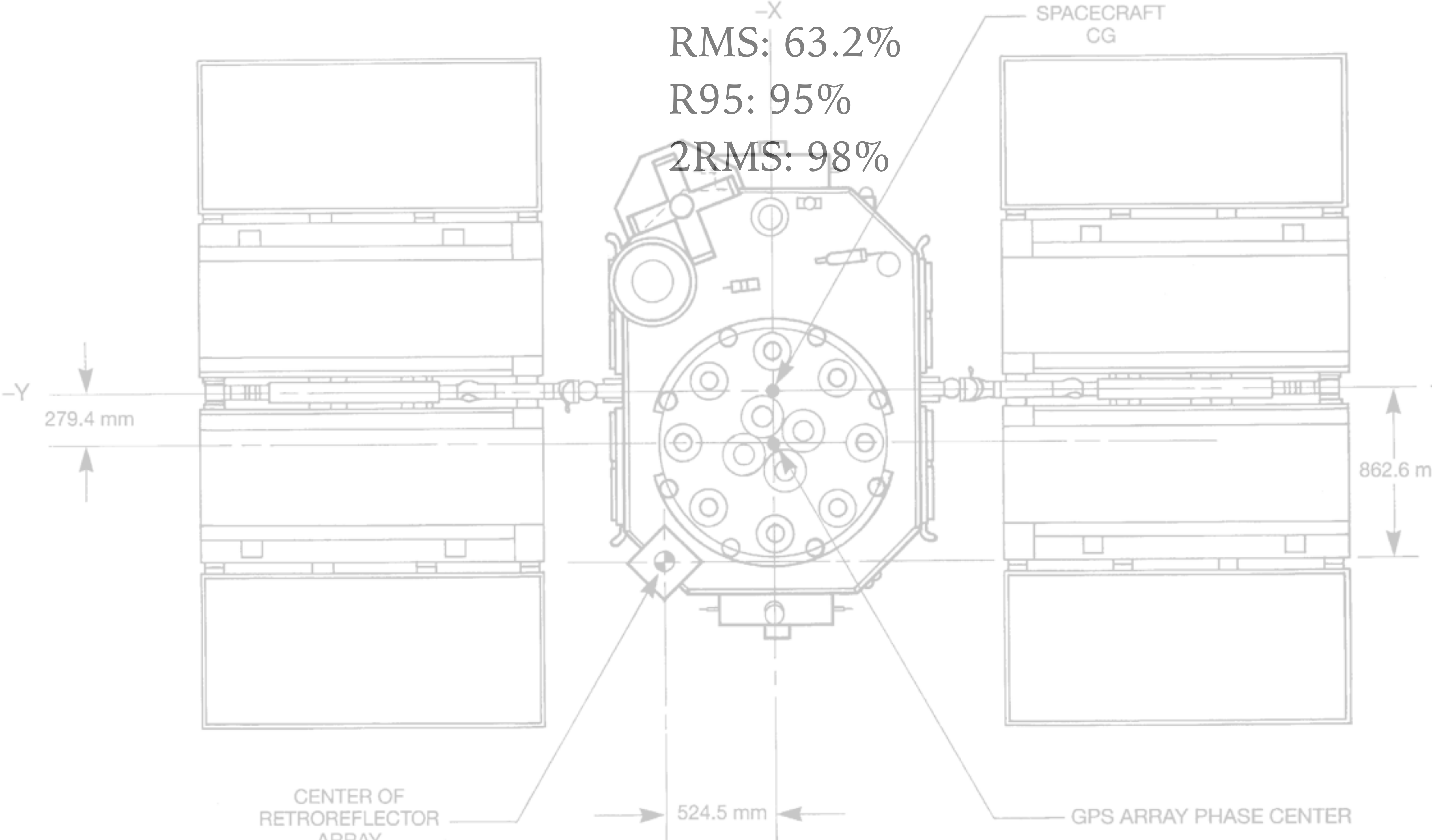
The strong cuts determine confidence regions.

The fuzzy set represents imprecise information about a point.

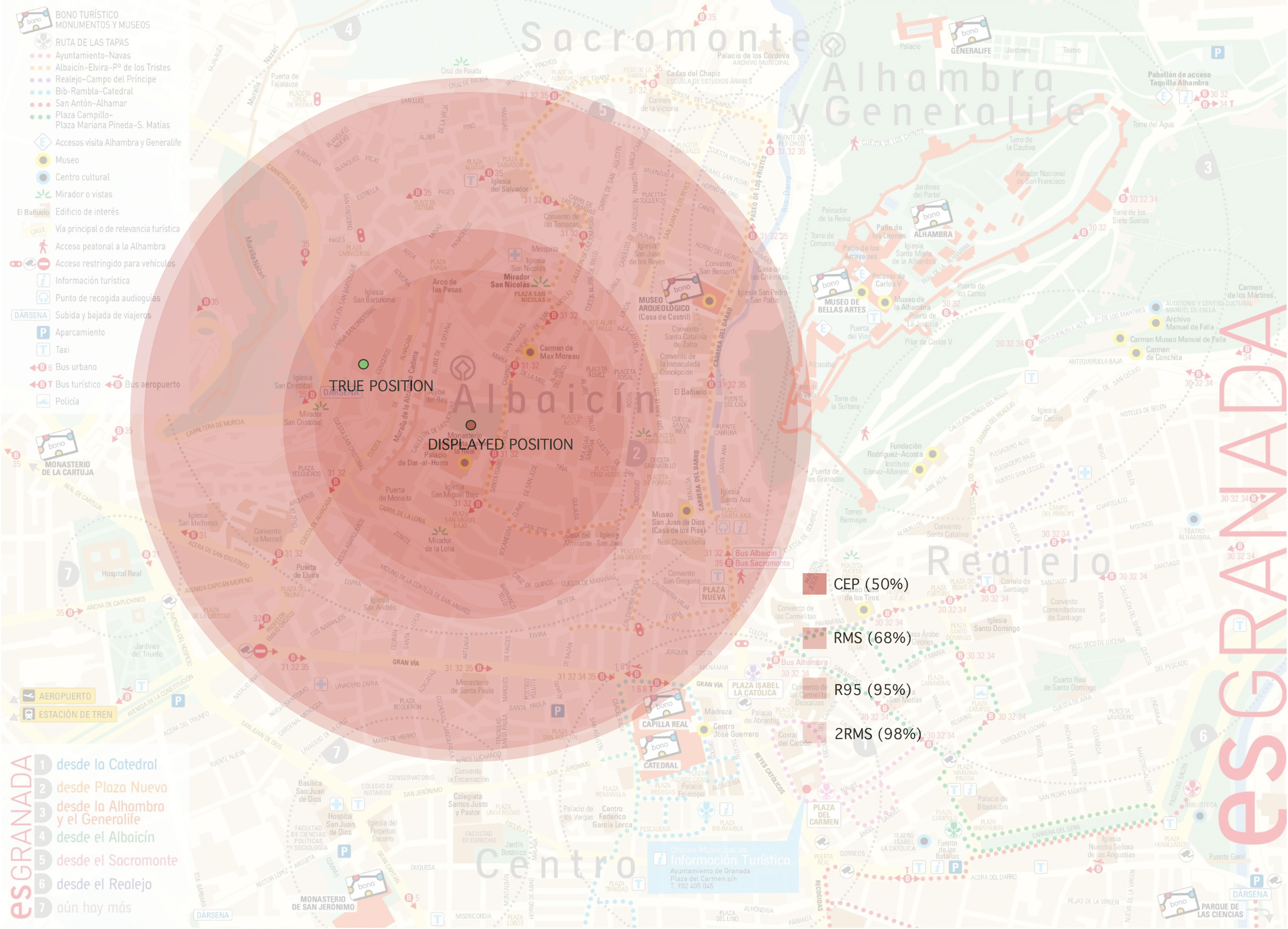
THE EPISTEMIC INTERPRETATION: GPS EXAMPLE

Specifications using a GPS device:

CEP: 50%
RMS: 63.2%
R95: 95%
2RMS: 98%



THE EPISTEMIC INTERPRETATION

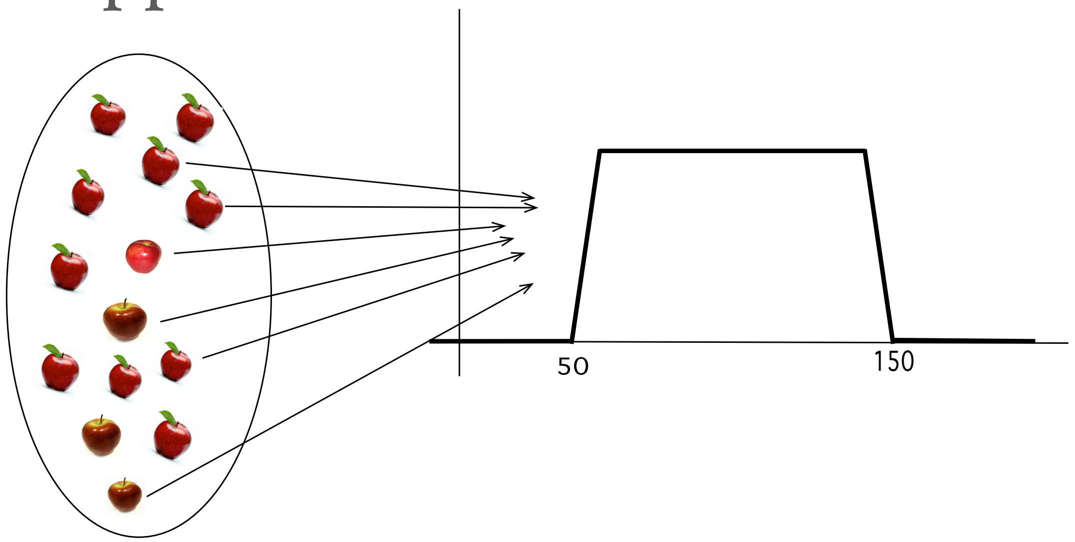


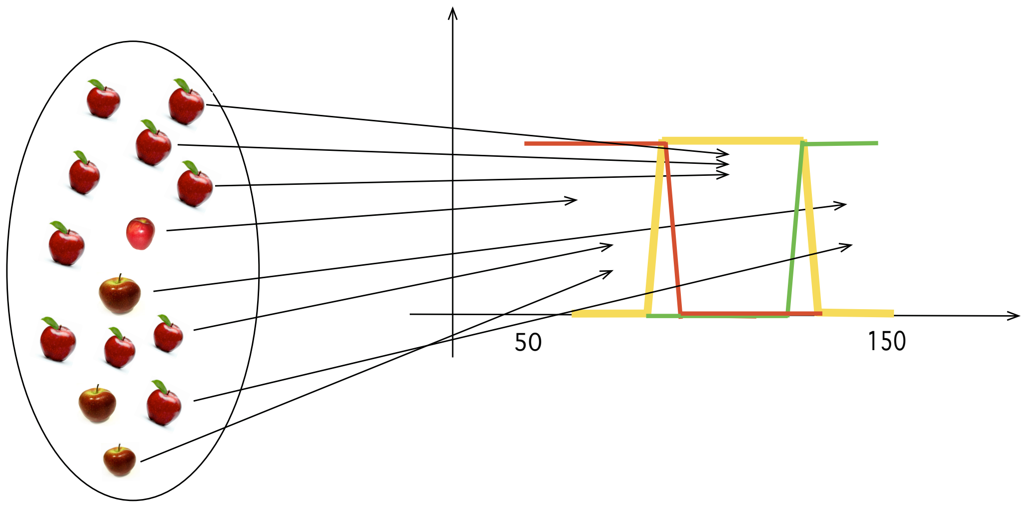
DIFFERENT CONTEXTS, DIFFERENT TOOLS: ONTIC VIEW

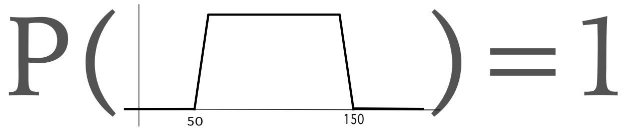
- A probability distribution over the family of fuzzy sets (generalisation of fuzzy mass function to non-necessarily finite case).
- A specific numerical distance is defined over a family of fuzzy sets. Many distances extending the usual distance can be defined.
- Each distance determines a metric space. Thus we can use (classical) probability on metric spaces.
- The expectation is a fuzzy set. Two options to define it:
  - Fréchet-type expectation. (Fuzzy set that minimises square distances)
  - Lebesgue-type expectation. (Fuzzy arithmetic for “simple” f.rv. + limit.
- The variance is a number. (Expected squared distance between the fuzzy images and the Fréchet fuzzy expectation).

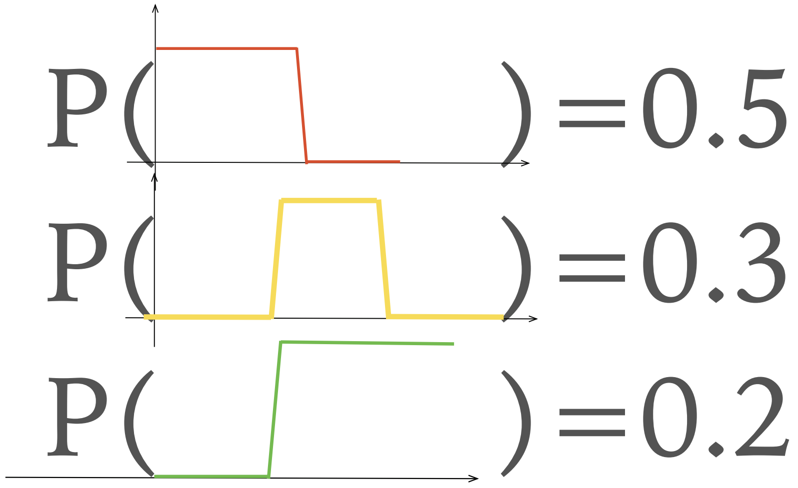
ONTIC RANDOM FUZZY SETS: MAIN IDEAS

- Their (fuzzy) outcomes do exist, and do not represent imprecise info about points. We are interested on the probability of appearance of each outcome.
- Example: Peter and Paul are asked about the weight of the same bunch of apples.

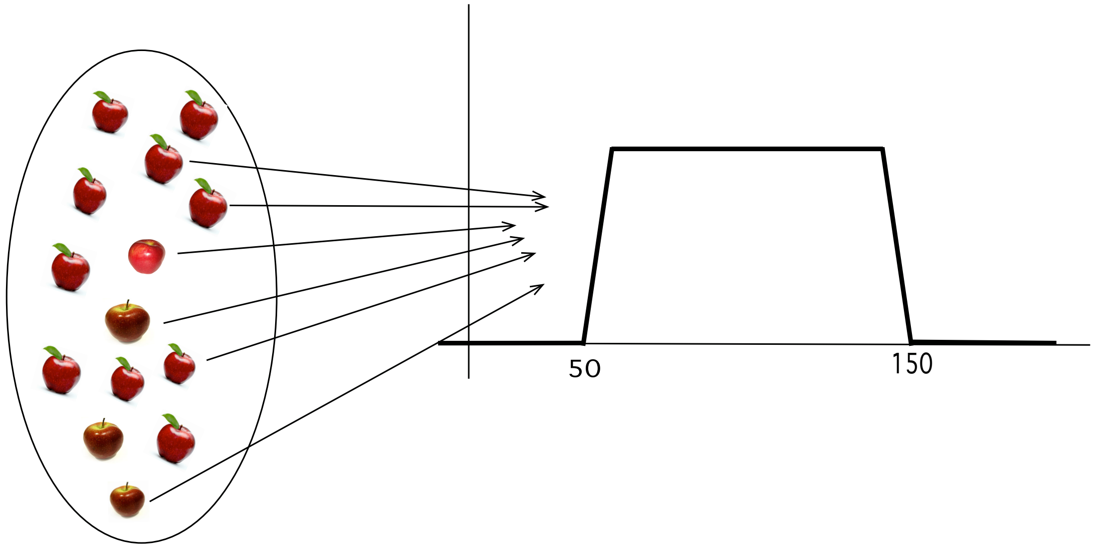


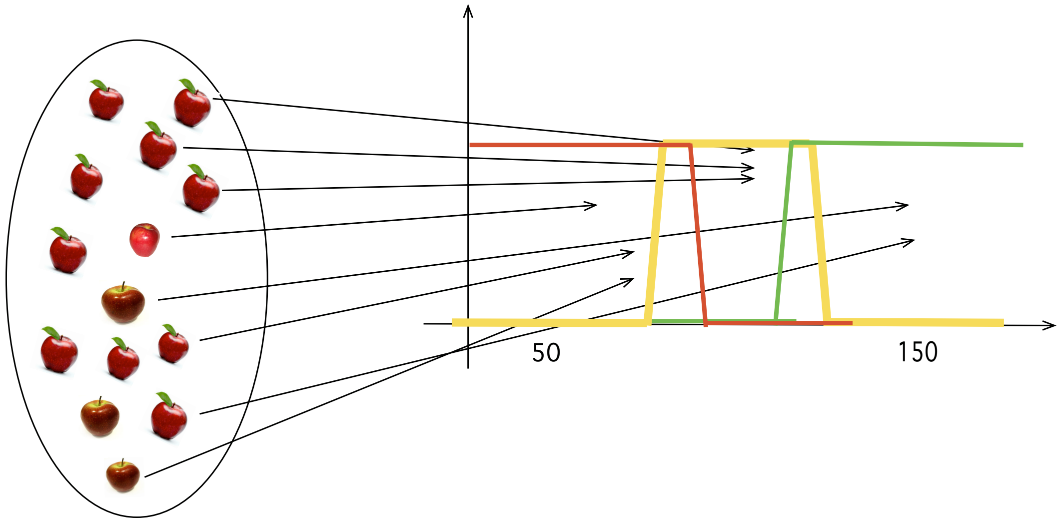




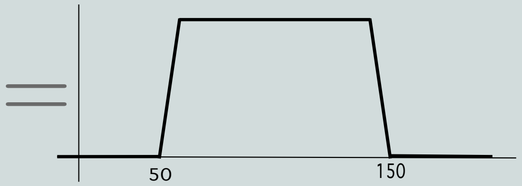


EXPECTATION AND VARIANCE OF ONTIC RANDOM FUZZY SETS



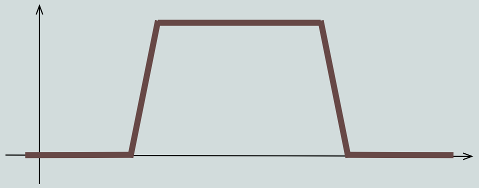


$$E(\tilde{X}_{\text{Peter}}) =$$



$$V(\tilde{X}_{\text{Peter}}) = E[d^2(\tilde{X}_{\text{Peter}}, E(\tilde{X}_{\text{Peter}}))] = 0$$

$$E(\tilde{X}_{\text{Paul}}) =$$



$$V(\tilde{X}_{\text{Paul}}) = E[d^2(\tilde{X}_{\text{Paul}}, E(\tilde{X}_{\text{Paul}}))] > 0$$

DIFFERENT CONTEXTS, DIFFERENT TOOLS: EPISTEMIC VIEW

Goal: study a probability distribution incompletely described by the fuzzy random variable. Two interpretations model two different situations:

- Second-order approach: $\tilde{X}(\omega)$ represents a possibility distribution indicating our knowledge about the true outcome $X_0(\omega) \in \Theta$.
- First-order approach: A two-step random experiment, on Ω and Θ. $\tilde{X}(\omega)$ represents a conditional possibility given ω.
  Both determine a belief-plausibility model on Θ.

- Couso, I., & Sánchez, L. (2008). Higher order models for fuzzy random variables. Fuzzy Sets and Systems, 159(3), 237-258.
- Couso, I., & Sánchez, L. (2011). Upper and lower probabilities induced by a fuzzy random variable. Fuzzy Sets and Systems, 165(1), 1-23.

SECOND-ORDER APPROACH

- Given $\omega \in \Omega$, $\tilde{X}(\omega) : \Theta \to [0,1]$ represents our incomplete knowledge about $X_0(\omega)$.
- $\tilde{X}(\omega)(x) =$“degree of possibility that $X_0(\omega)$ coincides with x”.

- Kruse, R., & Meyer, K. D. (1987). Statistics with vague data. D. Reidel Publishing Company.
- Couso, I., & Sánchez, L. (2008). Higher order models for fuzzy random variables. Fuzzy Sets and Systems, 159(3), 237-258.

POSSIBILITY DISTRIBUTION OVER THE COLLECTION OF RANDOM VARIABLES

$$\mathrm{acc}_{\tilde{X}}(X) = \inf_{\omega \in \Omega} \tilde{X}(\omega)\big(X(\omega)\big).$$

According to Kruse and Meyer, $\mathrm{acc}_{\tilde{X}}(X)$ represents the grade of possibility that X is the “true” random variable that models the studied experiment. In other words, it represents the grade of possibility of the assertion

$$X_0(\omega) = X(\omega), \qquad \forall\, \omega \in \Omega.$$

SECOND-ORDER POSSIBILITY MEASURE

For every $\alpha \in (0,1]$ define

$$S(\tilde{X}_\alpha) = \Big\{ X : \Omega \to \mathbb{R} \text{ measurable } \ : \ X(\omega) \in \tilde{X}_\alpha(\omega), \ \forall\, \omega \in \Omega \Big\}.$$

Now consider the set probability measures $\mathbb{P}$ satisfying the inequalities:

$$\mathbb{P}\Big(S(\tilde{X}_\alpha) \Big) \geq 1 - \alpha, \qquad \forall\, \alpha \in [0,1].$$

They are the ones dominated by a the possibility measure generated by "acc". This possibility measure over the set of random variables induces a 2nd-order possibility generated by:

$$\mathbb{\Pi}(Q) = \sup\{\mathrm{acc}_{\tilde{X}}(X) \, : \, P_X = Q\}.$$

FIRST-ORDER APPROACH

- We model a two-step random experiment.
- First step modelled by P on Ω. (An outcome ω is selected).
- Second step: an outcome from Θ.
- $\tilde{X} : \Omega \to \mathcal{F}(\Omega)$ links both steps.

  $\tilde{X}(\omega)$ represents a possibility distribution conditional to the outcome ω.
- Our knowledge about the probability on Θ is characterised by $\overline{P}$):

$$\overline{P}(A) = \int \Pi_{\tilde{X}(\omega)}(A)\, dP(\omega) = \int_0^1 \overline{P}_{\Gamma_\alpha}(A)\, d\alpha$$

Couso, I., & Sánchez, L. (2011). Upper and lower probabilities induced by a fuzzy random variable. Fuzzy Sets and Systems, 165(1), 1-23.

FIRST-ORDER APPROACH

- We model a two-step random experiment.
- First step modelled by P on Ω. (An outcome ω is selected).
- Second step: an outcome from Θ.
- $\tilde{X} : \Omega \to \mathcal{F}(\Omega)$ links both steps.

  $\tilde{X}(\omega)$ represents a possibility distribution conditional to the outcome ω.
- Our knowledge about the probability on Θ is characterised by $\overline{P}$:

$$\overline{P}(A) = \int \Pi_{\tilde{X}(\omega)}(A)\, dP(\omega) = \int_0^1 \overline{P}_{\Gamma_\alpha}(A)\, d\alpha$$

- $\overline{P}$, upper probability associated to $\Gamma(\omega, \alpha) = \tilde{X}_\alpha(\omega),\ \forall\, (\omega, \alpha) \in \Omega \times [0, 1]$.

Couso, I., & Sánchez, L. (2011). Upper and lower probabilities induced by a fuzzy random variable. Fuzzy Sets and Systems, 165(1), 1-23.

LITERATURE ON EPISTEMIC FUZZY RANDOM VARIABLES

- Kruse, R., & Meyer, K. D. (1987). Statistics with vague data. D. Reidel Publishing Company.
- Couso, I., & Sánchez, L. (2008). Higher order models for fuzzy random variables. Fuzzy Sets and Systems, 159(3), 237-258.
- Couso, I., & Sánchez, L. (2011). Upper and lower probabilities induced by a fuzzy random variable. Fuzzy Sets and Systems, 165(1), 1-23.
- I. Couso, D. Dubois, L. Sánchez (2014) Random sets and random fuzzy sets as ill-perceived random variables, Springer.
- I. Couso, I., Borgelt, C., Hüllermeier, E., & Kruse, R. (2019). Fuzzy sets in data analysis: From statistical foundations to machine learning. IEEE Computational Intelligence Magazine, 14(1), 31-44.
- Denœux, T. (2023). Reasoning with fuzzy and uncertain evidence using epistemic random fuzzy sets: General framework and practical models. Fuzzy Sets and Systems, 453, 1-36.
- Denœux, T. (2025). Uncertainty Measures in a Generalized Theory of Evidence.

KEY TAKEAWAYS

Formal notion	**Induced distribution (finite setting)**
Random variable/vector	Bayesian mass assignment, $m(\{b\})$
Random set	Mass assignment, $m(B)$
Random fuzzy sets	Fuzzy mass assignment, $m(\tilde{B})$

Formal notion	**Info. about X_0**	**Info. about P on Θ**
Random variable	X_0	P_{X_0}
Epistemic random set	$S(\Gamma)$	$\mathcal{P}(\Gamma)$
Epist. r.f.s. (2^{nd} order appr.)	$acc_{\tilde{X}}$	2^{nd}-order possib. measure
Epist. r.f.s. (1^{nd} order appr.)	$\ldots$	$\infty-$order alt. prob.

Even when Θ is finite, m does not characterize $\mathcal{P}(\Gamma)$ or the 2nd order possibility measure. It does not encompass information about the nature of Ω (initial space) and the behaviour of the r.(f.)s. on it.