The many interpretations of 'probability'

- there are multiple concepts of 'probability' as they emerged over the centuries
- on the mathematical level, these can be unified in the sense of symbolic manipulations of those 'probabilities' always taking the same form; this unification is achieved by measure theory.
- Afterwards, you can still assign a specific interpretation to your probabilities
- Most problems occur when scientists mixed up interpretations : the computations are then still symbolically correct (which is why we will leave 'probability' an abstract concept in our treatment of measure theory), but the meaning got lost, leading to „statistical misinterpretation" and/or statistical paradoxes
- Famous interpretations are the frequentist and the Bayesian viewpoint, but first let's turn to the abstraction and unification brought about by measure theory.

Measure Theory and Lebesgue integration

- Let there be a set S.

  Examples: $S = \{T, F\}$ true and false of boolean logic

  $S = \{1, 2, 3, 4, 5, 6\}$ outcomes of dice throws

- Denote the „set of all subsets" of S by $\mathcal{S}$, the „power set"

  For boolean logic, the power set is

$$\mathcal{S} = \{\underbrace{\emptyset, S}_{\text{always starts like this}}, \{T\}, \{F\}\}$$

- Def: „σ-Algebra", σ-Algebras indicate countability or measurability.

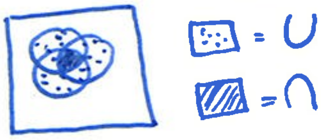
  σ: small sigma: indicates „countability", like Σ for the sum symbol

$\rightsquigarrow$

Continuation: definition of σ-Algebra.

- Let $\mathcal{A} \subseteq \mathfrak{P}(S)$ be some subset of the power set of S.
- $\mathcal{A}$ is an σ-Algebra if
  (i) $\emptyset$ and $S \in \mathcal{A}$
  (ii) for all $A \in \mathcal{A} \Rightarrow \bar{A} := S/A \in \mathcal{A}$
  (iii) for $A_i \in \mathcal{A}$ with $i \in \mathbb{N}$, then $\left(\bigcup_{i=1}^{\infty} A_i \right) \in \mathcal{A}$

- (i) indicates a notion of „nothing" and „all"
- (ii) $\bar{A}$ is the complement of A, indicating a form of closure or self-consistency
- (iii) indicates countability: conjunctions of elements are again „countable"



⇝ If $\mathcal{A}$ is a σ-algebra, then all $A_i \in \mathcal{A}$ are called „$\mathcal{A}$-measurable" subsets of S

⇝ $(S, \mathcal{A})$ is then called „a <u>measurable</u> space" [the „measure" still needs to be specified]

$\uparrow$ set $\quad$ $\uparrow$ a σ-Algebra on it

It follows: If S some set, I some range of integers, $I = [1, 2, \dots, N]$ and many $\mathcal{A}_i$ are ~~..~~ σ-Algebrae on S, then

$$\left(\bigcap_{i=1}^{N} \mathcal{A}_i \right) \text{ is again a } \sigma\text{-Algebra on } S$$

(⇝ combined posteriors of many experiments are again a 'probability')

Bayesian „updating"

or data collection and joint analys

• Def: „a measure" [sth that behaves like a generalized notion of „volume"]

If $(S, \mathcal{A})$ a measurable space using set S and σ-Algebra $\mathcal{A}$, then

$$\mu : \mathcal{A} \to [0, \infty]$$

is called a „measure", if $[0,\infty]$ is $\mathbb{R}^+$, including ∞, and if μ satisfies

(0) $\mu(A_i) \geq 0 \;\forall A_i \in \mathcal{A}$

(i) $\mu(\emptyset) = 0$ [Empty set has measure zero / sth needs to happen]

(ii) σ-additivity of the measure:

when $A_i \cap A_j = \emptyset \quad \forall i \neq j$, then

$$\boxed{\sum_{i=1}^{\infty} \mu(A_i) = \mu\left(\bigcup_{i=1}^{\infty} A_i\right)}$$

σ-additivity

($A_i \in \mathcal{A}$ of course)

$\Rightarrow$ we can „parcel-ate" the entire space

— all the rest follows —

Noteworthy points and examples:

• If $(S, \mathcal{A})$ is a measur<u>able</u> space, then $(S, \mathcal{A}, \mu)$ is a measu<u>re</u> space.

• many measures exist.

• Examples:

1) Lebesgue-measure:
n-dimensional Euclidean space E^n, then

$$\boxed{\mu_L([0,1]^n) = 1}$$

$\rightsquigarrow$ length, area, volume, hypervolume

2) Cardinality // Counting measure:

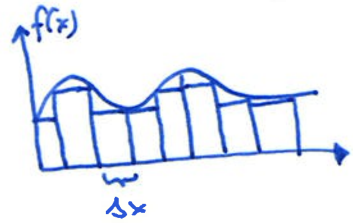
$$\mu_c(A) = \begin{cases} \text{cardinality}(A) & \text{if } A \text{ finite} \\ \infty & \text{else} \end{cases}$$

3) Probability measures, including Dirac-measure
(we'll come to that in a second)

• Lebesgue-Integration

= integration of a function against (or „with respect to") a measure

• Reminder: $\int f(x)\,dx$ = a „Riemann-Integral" when derived as limit of a Riemann sum:

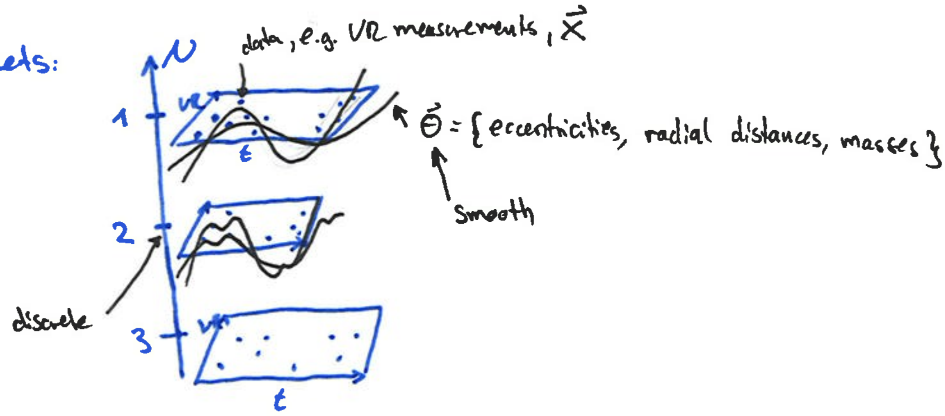


• Riemann-integral:
- good if f(x) is „the intesting" part, and $\mathbb{R}^n$ the „boring part"
- good if f(x) „nice" ← physicists

• Statistics:
- E^n, or $\mathbb{R}^n$ insufficient: „data space" includes discrete events and ~~...~~ continuous events alike,
- more data ⇝ dim(dataspace) constantly changed
- parameters $\vec{\Theta}$ span a „manifold" if non-linear
- functions f(x) often „not nice": degenerate distributions / non-existant densities

⇝ Examples:

Exoplanets:



⇝ What type of „space" is spanned by $\vec{\Theta} \otimes N \otimes \vec{x}$?

⇝ Integrals go over a „domain", but our domains are highly complicated: our domains come from spaces which capture the joint variability of continuous and discrete physical parameters $(\vec{\Theta}, N)$ and the variability of the data, which might again be composed of many sub-measurements (radial velocity data + transit curves + traces of the telescope + ...)

⇝ Henri Lebesgue (1875–1941) introduces the „Lebesgue Integral" in 1904, so a fairly modern integral that can handle „funky spaces"!

- $\underbrace{\text{Lebesgue Integral}}_{\text{integration w.r.t. a measure}} \neq \underbrace{\text{Lebesgue measure}}_{\text{area, volume ...}}$

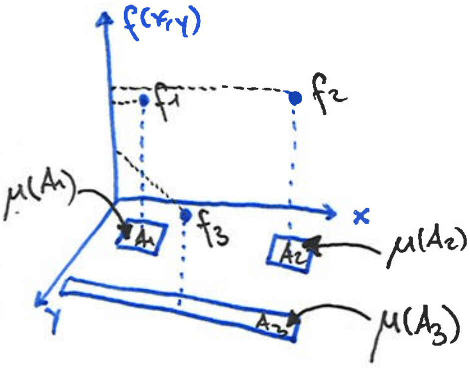
- Lebesgue Integral with respect to the Lebesgue measure is a special case.

- May $(X, \mathcal{A}, \mu)$ be a measure space, using set X, σ-Algebra $\mathcal{A}$, and measure μ.

  X: the set we work on

  $\mathcal{A}$: "extended counting"

  $\mu: \mathcal{A} \to [0, \infty]$



Integral

$$\rightsquigarrow \quad \mathcal{I}[f(x,y)] = \sum_{i=1}^{3} f_i \, \mu(A_i)$$

- We write the Lebesgue integral as

$$\int_X f \, d\mu = \int_X f(x) \, d\mu(x) \overset{\text{if discrete}}{=} \sum_{i=1}^{N} f_i \, \mu(A_i)$$

X: integral runs over this set

f: of this function

$d\mu$: with respect to this measure

- Example: data stream:

assigned measure

(Blanking of high noise)

⇓

cleaned data