cairo.thws
Center for Artificial Intelligence
Technical University of
Applied Sciences
Würzburg-Schweinfurt

Magda Gregorová
magda.gregorova@thws.de

Meausure-theory view of probability

handwavy and informal

November 25, 2025

Outline

cairo.thws
Center for Artificial Intelligence
Technical University of Applied Sciences Würzburg-Schweinfurt

Probability space definition

Positive measure and push-forward

Random variables and their distribution

Null sets and equivalence

1. σ-algebra $\mathscr{S}$ - non-empty collection of subsets closed under complement and countable unions
   - if $A \in \mathscr{S}$, then $A^c \in \mathscr{S}$
   - if $A_i \in \mathscr{S}$ for $i \in I$ (countable index set), then $\bigcup_{i \in I} A_i \in \mathscr{S}$
   - if $A \in \mathscr{S}$ and $A \in \mathscr{S}$, then $A \cup B \in \mathscr{S}$
   - in consequence $\emptyset \in \mathscr{S}$ and $S \in \mathscr{S}$
2. $(S, \mathscr{S})$ forms a measurable space in this context called the sample space.
3. Probability measure is the same as probability distribution or probability law
4. More generally a positive measure on $(S, \mathscr{S})$ is a function $\mu : \mathscr{S} \to [0, \infty]$ satisfying non-negativity and countable additivity. A probability measure is a positive measure with total measure equal to 1.
5. The triplet $(S, \mathscr{S}, \mu)$ is a measure space. Probability space is a special case of a measure space where the total measure is 1.
6. Any finite positive measure μ on the sample space $(S, \mathscr{S})$ can be re-scaled into a probability measure as $\mathbb{P}(A) = \mu(A)/\mu(S)$, $A \in \mathscr{S} \Rightarrow$ link to energy models.

Probability space

cairo.thws
Center for Artificial Intelligence
Technical University of Applied Sciences Würzburg-Schweinfurt

Probability space $(S, \mathscr{S}, \mathbb{P})$:

- measurable space $(S, \mathscr{S})$
  - S - sample space
  - $\mathscr{S}$ - σ-algebra on S - collection of subsets
- probability measure $\mathbb{P}$ - real-valued function on sample space $(S, \mathscr{S})$ s.t.:
  - non-negativity: $\mathbb{P}(A) \geq 0$ for all $A \in \mathscr{S}$
  - countable additivity: countable disjoint $\{A_i : i \in I\} \in \mathscr{S} \Rightarrow \mathbb{P}(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mathbb{P}(A_i)$
  - normalization: $\mathbb{P}(S) = 1$

Note: any finite positive measure μ on $(S, \mathscr{S}) \Rightarrow$ prob. measure $\mathbb{P}(A) = \mu(A)/\mu(S)$.

1. Definition of discrete and continuous space is a bit tricky and depends on the definition of topology but simply speaking in discrete space the set S is countable.
2. For d-dimensional Eucledean space $(\mathbb{R}^d, \mathscr{R}^d)$, $\mathbf{A} = A_1 \times A_2 \times \ldots \times A_d \in \mathscr{R}^d$, $A_1, A_2, \ldots, A_d \in \mathscr{R}$, $\lambda(\mathbf{A}) = \lambda(A_1) \times \lambda(A_2) \times \ldots \times \lambda(A_d)$

Positive measure

cairo.thws
Center for Artificial Intelligence
Technical University of Applied Sciences Würzburg-Schweinfurt

Positive measure on $(S, \mathscr{S})$ - function $\mu : \mathscr{S} \to [0, \infty]$ s.t.:

- $\mu(\emptyset) = 0$
- countable additivity: countable disjoint $\{A_i : i \in I\} \in \mathscr{S} \Rightarrow \mu(\bigcup_{i \in I} A_i) = \sum_{i \in I} \mu(A_i)$
- $\Rightarrow$ measure space $(S, \mathscr{S}, \mu)$

Note: if $\mu(S) < \infty \Rightarrow (S, \mathscr{S}, \mu)$ finite measure space.

Examples of measures:

- counting measure: for discrete $(S, \mathscr{S})$, $\#(A)$ = number of elements in $A \in \mathscr{S}$
- Lebesgue measure: for Eucledean $(\mathbb{R}, \mathscr{R})$, interval $I = [a, b] \in \mathscr{R}$, $\lambda(I) = b - a$ - length
  Eucledean $(\mathbb{R}^d, \mathscr{R}^d)$, $\mathbf{A} \in \mathscr{R}^d$, $\lambda(\mathbf{A}) = \lambda(A_1) \times \lambda(A_2) \times \ldots \times \lambda(A_d)$ - area, volume
- probability measure: positive finite measure on $(S, \mathscr{S})$ s.t. $\mathbb{P}(S) = 1$

1. Careful, though the notation is the same, the inverse image does not have to be a function (the inverse function may not exist).
2. A **measurable function** is a function $f : S \to T$ where $(S, \mathscr{S})$ and $(T, \mathscr{T})$ are measurable spaces and $f^{-1}(A) \in \mathscr{S}$ for any $A \in \mathscr{T}$.
3. A continuous function $f : S \to T$ is measurable.
4. assume $(R, \mathscr{R}) \subset (S, \mathscr{S})$ then μ restricted to $\mathscr{R}$ is measure on $(R, \mathscr{R})$
5. measure space $(S, \mathscr{S}, \mu)$ func $f : S \to T$ and measure ν on $(T, \mathscr{T}) \Rightarrow$ pullback measure μ on $(S, \mathscr{S})$ defined as $\mu(A) = \nu(f(A))$ for $A \in \mathscr{S}$
6. if μ is positive measure on $(S, \mathscr{S})$ then $c\mu$ for $c \in (0, \infty)$ is also positive measure on $(S, \mathscr{S})$
7. Assume a measure space $(S, \mathscr{S}, \mu)$ a measurable space $(T, \mathscr{T})$ and a measurable function $f : S \to T$. Then ν defined as below is a positive measure on $(T, \mathscr{T})$

Pre-image and push-forward

Forward (direct) image

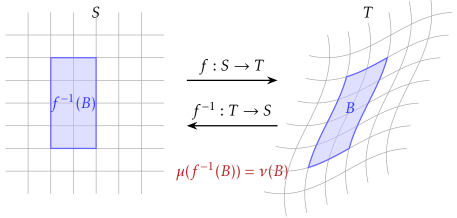
Assume sets S and T, func $f : S \to T$, and $A \subseteq S$.
Forward (direct) image of A under f is subset of T: $f(A) = \{f(x) \in T : x \in A\}$

Pre-image (ivnerse image)

Assume sets S and T, func $f : S \to T$, and $B \subseteq T$.
Pre-image of B under f is subset of S: $f^{-1}(B) = \{x \in S : f(x) \in B\}$

Push-forward measure

Assume $(S, \mathscr{S}, \mu)$, $(T, \mathscr{T})$ and $f : S \to T$
push-forward of μ by f is measure ν on $(T, \mathscr{T})$: $\nu(B) = \mu(f^{-1}(B)), \quad B \in \mathscr{T}$



1. Here S is still the event set in the sense of abstract outcomes of experiments.

Random variables

$(S, \mathscr{S}, \mathbb{P})$ probability space, $(T, \mathscr{T})$ another measurable space

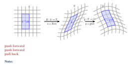
Random variable $X : S \to T$ - measurable function S to T

- for outcome $s \in S$, X takes value $x = X(s) \in T$ - realization of r.v. X
- pre-image of $x \in T$: $\{X = x\} = X^{-1}(x) = \{s \in S : X(s) = x\} \in \mathscr{S}$
- pre-image of $B \in T$: $\{X \in B\} = X^{-1}(B) = \{s \in S : X(s) \in B\} \in \mathscr{S}$

$(U, \mathscr{U})$ yet another measurable space

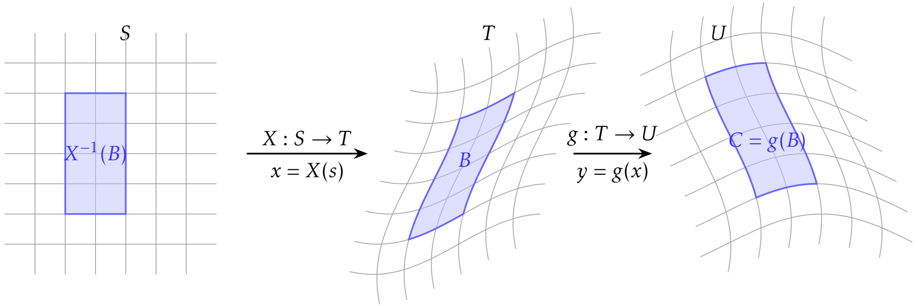
Random variable $Y = g(X)$ with measureable func $g : T \to U$

- for outcome $s \in S$, Y takes value $y = g(x) = g(X(s)) \in U$ - realization of r.v. Y
- pre-image of $C \in U$: $\{Y \in C\} = \{s \in S : X(s) \in g^{-1}(C)\} \in \mathscr{S}$



Probability distribution of random variable

$$(S, \mathscr{S}, \mathbb{P}),\ (T, \mathscr{T}, P_X),\ (U, \mathscr{U}, P_Y)$$



push forward of $\mathbb{P}$ by X: $P_X(B) = \mathbb{P}(X \in B) = \mathbb{P}(\{s \in S : X(s) \in B\})$
push forward of P_X by g: $P_Y(C) = \mathbb{P}(Y \in C) = \mathbb{P}(\{s \in S : g(X(s)) \in C\}) = \mathbb{P}(X \in B) = P_X(B)$
pull back of P_Y by g: $P_X(B) = P_Y(\{u \in U : g^1(u) \in B\})$

Note: R.v. directly as results of experiment $\Rightarrow$ prob. space $(\Omega, \mathscr{F}, \mathbb{P})$ (common notation)

Null sets and equivalence

1. sets $A, B \in \mathscr{S}$ are equivalent if $\mu(A \Delta B) = 0$, where $A \Delta B = (A \setminus B) \cup (B \setminus A)$
2. In probability space the almoste everywhere (a.e.) is equivalent to almost surely (a.s.) with respect to the probability measure $\mathbb{P}$.
3. The collection of essentially deterministic events $\mathcal{D}$ is a sub σ-algebra of $\mathscr{S}$.

Null sets and equivalence

cairo.thws
Center for Artificial Intelligence
Technical University of Applied Sciences Würzburg-Schweinfurt

Measure space $(S, \mathscr{S}, \mu)$

- set $A \in \mathscr{S}$ s.t. $\mu(A) = 0$ is called null set
- if statement holds for all $s \in S$ except for null set, it holds almost everywhere (a.e.)
- sets $A, B \in \mathscr{S}$ are equivalent if $\mu(A \Delta B) = 0$, where $A \Delta B = (A \setminus B) \cup (B \setminus A)$
- measureble funcs $f, g : S \to T$ are equivalent if $\mu\{s \in S : f(s) \neq g(s)\} = 0$

Probability space $(S, \mathscr{S}, \mathbb{P})$

- $\mathcal{N} = \{A \in \mathscr{S} : \mathbb{P}(A) = 0\}$ collection of null events
- $\mathcal{M} = \{A \in \mathscr{S} : \mathbb{P}(A) = 1\}$ collection of almost sure events
- $\mathcal{D} = \mathcal{N} \bigcup \mathcal{M} = \{A \in \mathscr{S} : \mathbb{P}(A) = 0 \text{ or } \mathbb{P}(A) = 1\}$ essentially deterministic events

Equivalence of r.v. - $X \equiv Y$ **iff** $\mathbb{P}(X = Y) = 1$

- X and Y have the same distribution: $P_X = P_Y$
- $\{X \in B\}$ and $\{Y \in B\}$ are equivalent events for any $B \in \mathscr{T}$
- equivalence class of r.v. X: $[X] = \{Y : S \to T \,|\, \mathbb{P}(X = Y) = 1\}$