

Magda Gregorová  
[magda.gregorova@thws.de](mailto:magda.gregorova@thws.de)

# Measure-theory view of probability

handwavy and informal

November 25, 2025

Licensed according to [CC-BY-SA 4.0](https://creativecommons.org/licenses/by-sa/4.0/)

# Outline

## Probability space definition

## Positive measure



# Probability space

1.  $\sigma$ -algebra  $\mathcal{S}$  - non-empty collection of subsets closed under complement and countable unions
  - if  $A \in \mathcal{S}$ , then  $A^c \in \mathcal{S}$
  - if  $A_i \in \mathcal{S}$  for  $i \in I$  (countable index set), then  $\bigcup_{i \in I} A_i \in \mathcal{S}$
  - if  $A \in \mathcal{S}$  and  $B \in \mathcal{S}$ , then  $A \cup B \in \mathcal{S}$
  - in consequence  $\emptyset \in \mathcal{S}$  and  $S \in \mathcal{S}$
2.  $(S, \mathcal{S})$  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  $(S, \mathcal{S})$  is a function  $\mu: \mathcal{S} \rightarrow [0, \infty]$  satisfying non-negativity and countable additivity. A probability measure is a positive measure with total measure equal to 1.
5. The triplet  $(S, \mathcal{S}, \mu)$  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  $\mu$  on the sample space  $(S, \mathcal{S})$  can be re-scaled into a probability measure as  $\mathbb{P}(A) = \mu(A)/\mu(S)$ ,  $A \in \mathcal{S} \Rightarrow$  link to energy models.

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

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

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

# Positive measure

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 Euclidean space  $(\mathbb{R}^d, \mathcal{R}^d)$ ,  $\mathbf{A} = A_1 \times A_2 \times \dots \times A_d \in \mathcal{R}^d$ ,  $A_1, A_2, \dots, A_d \in \mathcal{R}$ ,  
 $\lambda(\mathbf{A}) = \lambda(A_1) \times \lambda(A_2) \times \dots \times \lambda(A_d)$

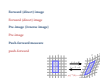
**Positive measure on  $(S, \mathcal{S})$  - function  $\mu : \mathcal{S} \rightarrow [0, \infty]$  s.t.:**

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

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

**Examples of measures:**

- counting measure: for discrete  $(S, \mathcal{S})$ ,  $\#(A)$  = number of elements in  $A \in \mathcal{S}$
- Lebesgue measure: for Euclidean  $(\mathbb{R}, \mathcal{R})$ , interval  $I = [a, b] \in \mathcal{R}$ ,  $\lambda(I) = b - a$  - **length**  
 Euclidean  $(\mathbb{R}^d, \mathcal{R}^d)$ ,  $\mathbf{A} \in \mathcal{R}^d$ ,  $\lambda(\mathbf{A}) = \lambda(A_1) \times \lambda(A_2) \times \dots \times \lambda(A_d)$  - **area, volume**
- probability measure: positive finite measure on  $(S, \mathcal{S})$  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  $f : S \rightarrow T$  where  $(S, \mathcal{S})$  and  $(T, \mathcal{T})$  are measurable spaces and  $f^{-1}(A) \in \mathcal{S}$  for any  $A \in \mathcal{T}$ .
3. A continuous function  $f : S \rightarrow T$  is measurable.
4. assume  $(R, \mathcal{R}) \subset (S, \mathcal{S})$  then  $\mu$  restricted to  $\mathcal{R}$  is measure on  $(R, \mathcal{R})$
5. measure space  $(S, \mathcal{S}, \mu)$  func  $f : S \rightarrow T$  and measure  $\nu$  on  $(T, \mathcal{T}) \Rightarrow$  **pullback measure**  $\mu$  on  $(S, \mathcal{S})$  defined as  $\mu(A) = \nu(f(A))$  for  $A \in \mathcal{S}$
6. if  $\mu$  is positive measure on  $(S, \mathcal{S})$  then  $c\mu$  for  $c \in (0, \infty)$  is also positive measure on  $(S, \mathcal{S})$
7. Assume a measure space  $(S, \mathcal{S}, \mu)$  a measurable space  $(T, \mathcal{T})$  and a measurable function  $f : S \rightarrow T$ . Then  $\nu$  defined as below is a positive measure on  $(T, \mathcal{T})$

## Pre-image and push-forward

### Forward (direct) image

Assume sets  $S$  and  $T$ , func  $f : S \rightarrow T$ , 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)

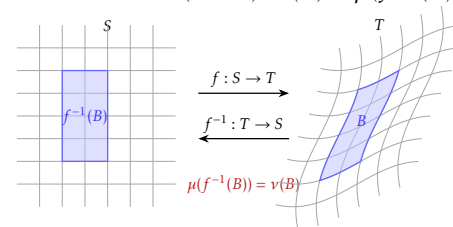
Assume sets  $S$  and  $T$ , func  $f : S \rightarrow T$ , and  $B \subseteq T$ .

**Pre-image** of  $B$  under  $f$  is subset of  $S$ :  $f^{-1}(B) = \{x \in S : f(x) \in B\}$

### Push-forward measure

Assume  $(S, \mathcal{S}, \mu)$ ,  $(T, \mathcal{T})$  and  $f : S \rightarrow T$

**push-forward** of  $\mu$  by  $f$  is measure  $\nu$  on  $(T, \mathcal{T})$ :  $\nu(B) = \mu(f^{-1}(B))$ ,  $B \in \mathcal{T}$



# Null sets and equivalence

1. sets  $A, B \in \mathcal{S}$  are **equivalent** if  $\mu(A \Delta B) = 0$ , where  $A \Delta B = (A \setminus B) \cup (B \setminus A)$
2. In probability space the almost 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  $\sigma$ -algebra of  $\mathcal{S}$ .

## Measure space $(S, \mathcal{S}, \mu)$

- set  $A \in \mathcal{S}$  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  $A, B \in \mathcal{S}$  are **equivalent** if  $\mu(A \Delta B) = 0$ , where  $A \Delta B = (A \setminus B) \cup (B \setminus A)$
- measurable funcs  $f, g : S \rightarrow T$  are **equivalent** if  $\mu\{s \in S : f(s) \neq g(s)\} = 0$

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

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

# Random variables

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

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

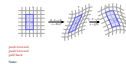
**Random variable  $X: S \rightarrow T$  - measurable function  $S$  to  $T$**

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

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

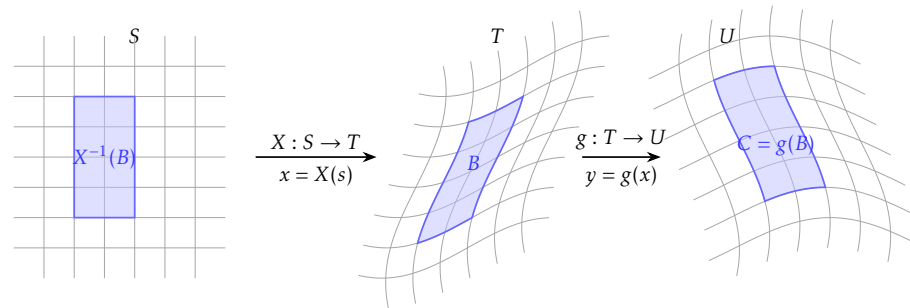
**Random variable  $Y = g(X)$  with measurable func  $g: T \rightarrow U$**

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



# Probability distribution of random variable

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



**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  $(\Omega, \mathcal{F}, \mathbb{P})$  (common notation)