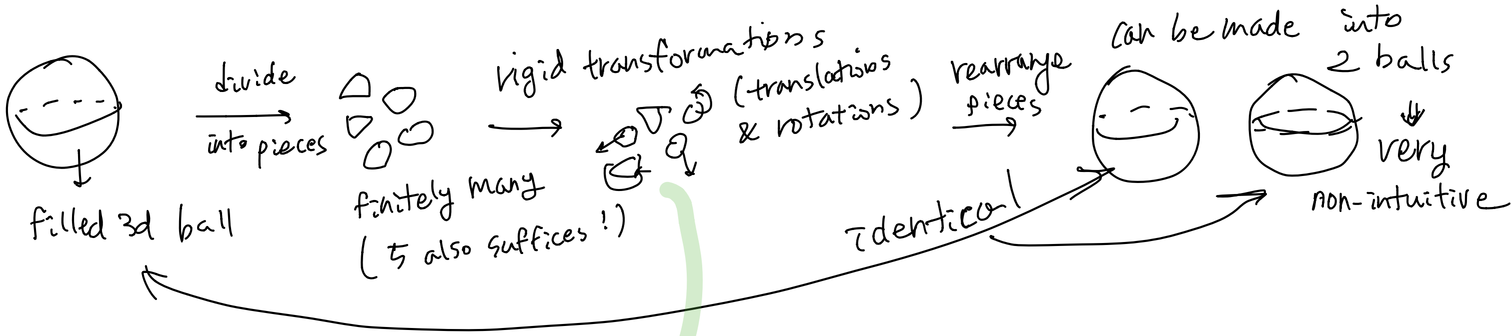
1.1 Measure Theory: Motivation

* The Banach-Tarski Paradox



<u>Assumption</u> ZFC axioms (Zermelo-Franco set theory with the axiom of choice)

<u>Resolution</u> i) Reject axiom of choice

or ii) Embrace the concept of non-measurable sets.

These individual sets cannot be assigned a measure in any meaningful way.

1.2 Measure Theory: Sigma-algebras.

<u>Def</u> Given a set Ω, (can be countable, or uncountable ... any arbitrary set!)

a σ-algebra on Ω

is a collection $\mathcal{A} \subset 2^{\Omega}$

ex) $\Omega = \{0,1\}$. $2^{\Omega} = \{\phi, \{0\}, \{1\}, \{0,1\}\}$
Power set : collection of all the subsets.

such that $\mathcal{A}$ is non-empty

and $\mathcal{A}$ is closed under complements ($E \in \mathcal{A} \Rightarrow E^c \in \mathcal{A}$)

ex) $\phi \in \mathcal{A} \Rightarrow \{0,1\} \in \mathcal{A}$.
$\{1\} \in \mathcal{A} \Rightarrow \{0\} \in \mathcal{A}$.

and $\mathcal{A}$ is closed under countable unions

$$\left(E_1, E_2, \cdots \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{\infty} E_i \in \mathcal{A} \right)$$

↳ this also covers finite i

ex) $\forall i \geq 3, E_i = E_1 \Rightarrow \bigcup_{i=1}^{\infty} E_i = E_1 \cup E_2$

Remarks i) $\Omega \in \mathcal{A}$ since $E \in \mathcal{A}$ and $E^c \in \mathcal{A}$ $\Rightarrow E \cup E^c = \Omega \in \mathcal{A}$.

ii) $\phi \in \mathcal{A}$ since $\Omega \in \mathcal{A} \Rightarrow \Omega^c = \phi \in \mathcal{A}$.

iii) $\mathcal{A}$ is closed under countable intersections

pf) Suppose $E_1, E_2, \cdots \in \mathcal{A}$.

$$\bigcap_{i=1}^{\infty} E_i = \bigcap_{i=1}^{\infty} (E_i^c)^c \underset{\text{De Morgan's Law}}{=} \left(\bigcup_{i=1}^{\infty} E_i^c \right)^c \in \mathcal{A}.$$

1.3 Measure Theory : Measures

Def Given $\mathcal{C} \subset 2^{\Omega}$, the σ-algebra generated $\mathcal{C}$, written $\sigma(\mathcal{C})$, is the "smallest" σ-algebra containing $\mathcal{C}$

That is, $\sigma(\mathcal{C}) = \bigcap_{\mathcal{A} \supset \mathcal{C}} \mathcal{A}$ $\to$ every existing σ-algebra $\mathcal{A}$ containing $\mathcal{C}$.

Remarks $\sigma(\mathcal{C})$ always exists, because

i) 2^{Ω} is a σ-algebra $\Rightarrow \mathcal{A}$ always exists

ii) Any intersection of σ-algebras is a σ-algebra.

$\hookrightarrow \sigma(\mathcal{C})$ is an intersection of σ-algebra(s).

Example Examples of σ-algebra.

i) $\mathcal{A} = \{\phi, \Omega\}$

ii) $\mathcal{A} = \{\phi, E, E^c, \Omega\}$.

iii) (Def of Borel σ-algebra)

If $\Omega = \mathbb{R}$, the Borel σ-algebra is

$\mathcal{B} = \sigma(\tau)$ where $\tau = \{$ open sets of $\mathbb{R}$ $\}$.

ex) (a, b) $a < x < b$

Any topological space is fine

Def A measure μ on Ω with σ-algebra $\mathcal{A}$
is a function $\mu : \mathcal{A} \to [0, \infty]$
such that i) $\mu(\emptyset) = 0$
and ii) Countable Additivity.

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

for any $E_1, E_2, \cdots \in \mathcal{A}$ of pairwise disjoint sets.

This also covers finite i.
ex) $\forall i \geq 4$, $E_i = \emptyset$. $\mu(E_i) = 0$.
$\mu\left(\bigcup_{i=1}^{3} E_i\right) = \sum_{i=1}^{3} \mu(E_i)$.

Def A probability measure is a measure P
such that $P(\Omega) = 1$.

* All these conditions, which is specified for probability measure, is called Kolmogorov's Axioms.

1.4 Measure Theory: Examples of Probability Measures.

i) Uniform Distribution

Finite Set. $\Omega = \{1, 2, \cdots, n\}$, $\mathcal{A} = 2^{\Omega}$.

$P(\{k\}) = P(k) = \frac{1}{n} \quad \forall k \in \Omega$. ← Shorthand notation.

Note that we have to define P for all sets of $\mathcal{A}$, but defining P on each every element is sufficient for inducing in the whole space.

(Claim: There exists a unique probability measure on all the sets of $\mathcal{A}$ that is consistent with the definition.)

ex) $P(\{1, 2, 4\}) = P(\{1\} \cup \{2\} \cup \{4\}) = P(1) + P(2) + P(4)$

"Decomposed Uniquely".   Pairwise disjoint.

$P(\Omega) = P\left(\bigcup_{i=1}^{n} \{i\}\right) = \sum_{i=1}^{n} P(\{i\}) = 1$. $\Rightarrow$ This is a probability measure.

ii) Geometric Distribution

Countably infinite set $\Omega = \{1, 2, 3, \cdots\}$, $\mathcal{A} = 2^{\Omega}$.

$P(k)$ = Probability it takes k coinflips to get heads

$= \alpha(1-\alpha)^{k-1} = 1/2^k$ for fair coin.

α: probability of getting heads.

Similar to i), it can be decomposed to single-element sets in a unique way to sum up the probabilities to get the probability of the whole set.

So, P is uniquely defined probability measure on $\mathcal{A}$.

$$P(\Omega) = P\left(\bigcup_{i=1}^{\infty} \{i\}\right) = \sum_{i=1}^{\infty} P(\{i\}) = \sum_{i=1}^{\infty} \alpha(1-\alpha)^{i-1} = \frac{\alpha}{1-(1-\alpha)} = 1.$$

iii) Exponential Distribution.

Uncountable Set $\Omega = [0, \infty)$, $\mathcal{A} = \mathcal{B}([0,\infty))$

$P([0, x)) := 1 - e^{-x} \quad \forall x > 0$ — it is automatically defined when $x=0$, because $P([0,0)) = P(\phi) = 0$.

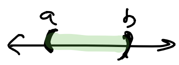
Sets of this form also "generates" Borel σ-algebra.

So, in fact, defining a probability measure on $\mathcal{B}([0,x))$ alone uniquely induces a probability measure on the whole σ-algebra.

Note that $P(\{x\}) = 0 \quad \forall x \geq 0$.

$$P(\Omega) = P([0,\infty)) = \lim_{x\to\infty} P([0,x)) = \lim_{x\to\infty} 1 - e^{-x} = 1.$$

* Lebesgue measure (on $\mathbb{R}$).



"length of interval". This is not a probability measure.

"most natural way to define the measure".

$\Omega = \mathbb{R}$, $\mathcal{A} = \mathcal{B}(\mathbb{R})$.

$\mu((a,b)) = b - a$ for any $a, b \in \mathbb{R}$, $a < b$.

cf. This is used for Lebesgue integral. idea: $dx \longrightarrow dm$ ($m = \mu$).

Riemann — Lebesgue.

1.5-6 Measure Theory: Basic Properties of Measures

Thm Basic Properties of measures.

Let $(\Omega, \mathcal{A}, \mu)$ be a measure space.

i) Monotonicity

If $E, F \in \mathcal{A}$ and $E \subset F$, then $\mu(E) \le \mu(F)$

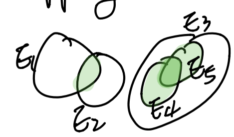
pf) $\mu(F) = \mu(E \cup (E^c \cap F)) = \mu(E) + \mu(E^c \cap F) \ge \mu(E)$.

↳ measure is non-negative.

ii) Subadditivity

inequality due to "overlapping"

If $E_1, E_2, \cdots \in \mathcal{A}$, then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \le \sum_{i=1}^{\infty} \mu(E_i)$



↳ arbitrary sets. not necessarily pairwise disjoint.

pf). The disjointization trick.

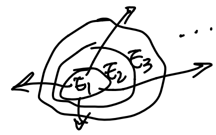
Sets F_k defined by $F_1 = E_1$, $F_2 = E_2 - E_1$, $F_3 = E_3 - (E_1 \cup E_2)$ ...

are disjoints, belong to $\bigcup_{i=1}^{\infty} E_i$, and satisfy $\bigcup_{i=1}^{\infty} E_i = \bigcup_{i=1}^{\infty} F_i$.

Using this trick, $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \mu\left(\bigcup_{i=1}^{\infty} F_i\right) = \sum_{i=1}^{\infty} \mu(F_i) \le \sum_{i=1}^{\infty} \mu(E_i)$

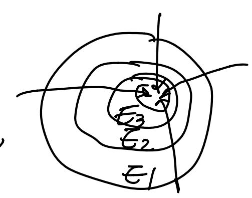
↳ F_i are disjoint.

iii) Continuity from below



If $E_1, E_2, \cdots \in \mathcal{A}$ and $E_1 \subset E_2 \subset \cdots$, then $\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \lim_{i \to \infty} \mu(E_i)$.

iv) Continuity from above



If $E_1, E_2, \cdots \in \mathcal{A}$ and $E_1 \supset E_2 \supset \cdots$ and $\mu(E_1) < \infty$,

then $\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = \lim_{i \to \infty} \mu(E_i)$.

Note that it holds for every probability measure.

ex) Lebesgue. Let $E_i = [i, \infty)$. then $\mu\left(\bigcap_{i=1}^{\infty} E_i\right) = 0 \ne \lim_{i \to \infty} \mu(E_i)$.

↳ violates $\mu(E_1) < \infty$.

[1.7] Measure Theory: More properties of Probability Measures.

Let $(\Omega, \mathcal{A}, P)$ be a probability measure space, with $E, F, E_i \in \mathcal{A}$.

i) $P(E \cup F) = P(E) + P(F)$ if $E \cap F = \phi$.

ii) $P(E \cup F) = P(E) + P(F) - P(E \cap F)$.

iii) $P(E) = 1 - P(E^c)$

iv) $P(E \cap F^c) = P(E) - P(E \cap F)$

v) Inclusion-Exclusion Formula.

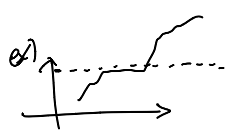
$$P\left(\bigcup_{i=1}^{n} E_i\right) = \sum_i P(E_i) - \sum_{i<j} P(E_i \cap E_j) + \sum_{i<j<k} P(E_i \cap E_j \cap E_k) - \cdots$$

$$+ (-1)^{n+1} P(E_1 \cap E_2 \cap \cdots \cap E_n).$$

[1.8] Measure Theory: CDFs and Borel Probability Measures

Def. A Borel (Probability) Measure on $\mathbb{R}$ is a (probability) measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Def. A CDF (Cumulative Distribution Function)



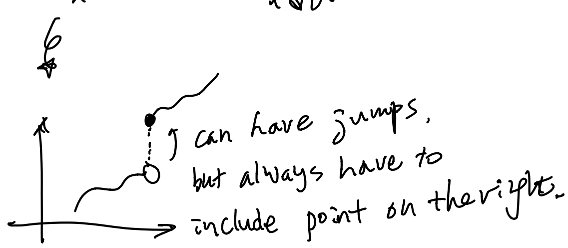
is a function $F: \mathbb{R} \to \mathbb{R}$

such that i) F is nondecreasing $(x \le y \Rightarrow f(x) \le f(y))$

ii) F is right-continuous $\left(\lim_{x \to a^+} F(x) = \lim_{x \downarrow a} F(x) = F(a)\right)$

iii) $\lim_{x \to \infty} F(x) = 1$.

iv) $\lim_{x \to -\infty} F(x) = 0$.



<u>Thm</u> i) If F is a CDF,

then there is a unique Borel probability measure on $\mathbb{R}$

such that $P((-\infty, x]) = F(x)$ $\forall x \in \mathbb{R}$.

ii) If P is a Borel probability measure on $\mathbb{R}$,

then there is a unique CDF F

such that $F(x) = P((-\infty, x])$ $\forall x \in \mathbb{R}$.

That is, there is an equivalence between CDF and Borel probability measure.