PROBABILITY SPACES

The One Most Important Definition

Probability space is $(\Omega, \mathcal{B}, P)$

- Ω is the sample space containing elementary outcomes ω,
- $\mathcal{B}$ is the $\sigma-$algebra of subsets (events) of Ω,
- P is a probability measure, a function $\mathcal{B} \ni A \to P(A) \in [0, 1]$

(i) $P(A) \geq 0$, all $A \in \mathcal{B}$

(ii) P is σ-additive,

$$\{A_n,\ n \geq 1, \text{disjoint}\} \Longrightarrow P(\cup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$$

(iii) $P(\Omega) = 1$

Basic properties of probabilistic measures

1. $P(A^c) = 1 - P(A)$ *prove ...*
2. $P(\emptyset) = 0$ *prove ...*
3. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ *prove ...*
4. $\{A_n,\, n \geq 1\}$ arbitrary events, the *exclusion-inclusion* formula holds

$$\begin{aligned} P(\cup_{j=1}^{n} A_j) &= \sum_{j=1}^{n} P(A_j) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) \\ &+ \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) - \ldots \\ &+ (-1)^{n+1} P(A_1 \cap A_2 \cap \cdots \cap A_n) \end{aligned}$$

moreover, the Bonferroni inequalities hold

$$
\begin{aligned}
P(\cup_{j=1}^{n} A_j) &\leq \sum_{j=1}^{n} P(A_j) \\
P(\cup_{j=1}^{n} A_j) &\geq \sum_{j=1}^{n} P(A_j) - \sum_{1\leq i<j\leq n} P(A_i \cap A_j) \\
& \quad etc.
\end{aligned}
$$

First part proved by induction, starting $n = 2$ (see 3), second part trickier.

5. *Monotonicity*

$$A \subseteq B \Longrightarrow P(A) \leq P(B) \ \ \textit{prove ...}$$

6. *Subadditivity*, $\{A_n, n \geq 1\}$ arbitrary events

$$P(\cup_{j=1}^{\infty} A_j) \leq \sum_{j=1}^{\infty} P(A_j) \ \ \textit{prove ...}$$

7. *Continuity*, if $A_n \uparrow A \Longrightarrow P(A_n) \uparrow P(A)$; if $A_n \downarrow A \Longrightarrow P(A_n) \downarrow P(A)$.

$$
\begin{aligned}
A_1 &\subset A_2 \subset A_3 \subset \cdots \subset A_n \subset \cdots \\
B_1 &= A_1,\ B_2 = A_2 \backslash A_1, \ldots,\ B_n = A_n \backslash A_{n-1}, \ldots \text{are disjoint} \\
\cup_{i=1}^n B_i &= A_n,\ \cup_{i=1}^\infty B_i = \cup_i A_i = A \\
\Longrightarrow & P(A) = P(\cup_{i=1}^\infty B_i) = \sum_{i=1}^\infty P(B_j) = \lim_{n\to\infty} \uparrow \sum_{i=1}^n P(B_j) \\
&= \lim_{n\to\infty} \uparrow P(\cup_{i=1}^n B_i) = \lim_{n\to\infty} \uparrow P(A_n)
\end{aligned}
$$

and the other part by complementarity of the measure ...

8. *Fatou's lemma,* $\{A_n,\, n \geq 1\}$ arbitrary events

$$\begin{aligned}\text{(i) } P(\liminf_{n\to\infty} A_n) &\leq \liminf_{n\to\infty} P(A_n) \\ &\leq \limsup_{n\to\infty} P(A_n) \leq P(\limsup_{n\to\infty} A_n)\end{aligned}$$

(ii) also if $A_n \to A \Longrightarrow P(A_n) \to P(A)$.

Proof

Part (ii) follows from part (i) *prove ...*

To prove (i) let us note that

$$\begin{aligned} P(\liminf_{n\to\infty} A_n) &= P(\lim_{n\to\infty} \uparrow \cap_{k\geq n} A_k) \overset{7}{=} \lim_{n\to\infty} \uparrow P(\cap_{k\geq n} A_k) \\ &= \liminf_{n\to\infty} \uparrow P(\cap_{k\geq n} A_k) \leq \liminf_{n\to\infty} P(A_n) \\ &\quad \text{attention! } \lim_{n\to\infty} P(A_n) \text{ may not exist!} \end{aligned}$$

The other inequalities are analogous ...

Probability distribution function

Suppose $\Omega = \mathbb{R}$ and P be a probability measure on $\mathbb{R}$

$$F(x) = P[(-\infty, x]], \; x \in \mathbb{R}$$

is called the distribution function (df), $F : \mathbb{R} \to [0, 1]$.

Properties of the distribution function

(i) F is right-continuous

(ii) F is monotone non-decreasing

(iii)

$$\begin{aligned} F(\infty) &= \lim_{x \uparrow \infty} \uparrow F(x) = 1 \\ F(-\infty) &= \lim_{x \downarrow -\infty} \downarrow F(x) = 0 \end{aligned}$$

Proof

(ii) *prove ...*

(i) consider $x_n \to \infty$

$$\begin{aligned} F(\infty) &= \lim_{x_n \to \infty} F(x_n) = \lim_{x_n \to \infty} \uparrow P[(-\infty, x_n]] \\ &= P[\lim_{x_n \to \infty} \uparrow (-\infty, x_n]] = P[\cup_n (-\infty, x_n]] \\ &= P(\mathbb{R}) = P(\Omega) = 1 \end{aligned}$$

etc.

(iii) $x_n \downarrow x \Longrightarrow (-\infty, x_n] \downarrow (-\infty, x] \Longrightarrow P[(-\infty, x_n]] \downarrow P[(-\infty, x]] \Longrightarrow$
$F(x_n) \downarrow F(x)$

$\square$

Example 2.1.2 Self-study

Dynkin's Theorem and constructions of Probability Spaces

Philosophy: Define probability on simpler structures and then extend to a σ-field ($= \sigma$-algebra).

Two important classes of structures

π-system: $\mathcal{P} \subset 2^\Omega$ is a π-system if it is closed under finite intersections: $A, B \in \mathcal{P} \Longrightarrow A \cap B \in \mathcal{P}$.

$\boldsymbol{\lambda}$-system (Dynkin's system, $\boldsymbol{\sigma}$-additive class):

$\mathcal{L} \subset 2^{\Omega}$ is a λ-system if

(old definition) (i) $\Omega \in \mathcal{L}$; (ii) $A, B \in \mathcal{L}$, $A \subseteq B \Longrightarrow B \setminus A \in \mathcal{L}$; (iii) $A_n \uparrow$, $A_n \in \mathcal{L} \Longrightarrow \cup_n A_n \in \mathcal{L}$.

equivalent to *(new definition)* (i) $\Omega \in \mathcal{L}$; (ii) $A \in \mathcal{L} \Longrightarrow A^c \in \mathcal{L}$; (iii) $A_n \in \mathcal{L}$, $A_n A_m = \emptyset$, $n \neq m \Longrightarrow \cup_n A_n \in \mathcal{L}$.

Example: A σ-field is always a λ-system, since the *new definition* is satisfied

Example: If class C is a λ-system and a π-system it is a σ-field. *(prove ...)*

Definition 2.2.1. The minimal structure $\mathcal{S}(C)$ generated by class C is a non-empty structure satisfying

(i) $\mathcal{S}(C) \supset C$

(ii)If $\mathcal{S}'$ is any other structure containing class C, then $\mathcal{S}' \supset \mathcal{S}(C)$

Proposition 2.2.1. The minimal structure $\mathcal{S}(C)$ exists and is unique.

Proof exactly like in the minimal σ-algebra case:

$$\mathcal{S}(C) = \cap\{\mathcal{G}\colon \mathcal{G} \text{ is a structure}, \mathcal{G} \supset C\}$$

Theorem 2.2.2 (Dynkin's Theorem)

(a) If $\mathcal{P} \subset 2^\Omega$ is a π-system and $\mathcal{L} \subset 2^\Omega$ is a λ-system, such that $\mathcal{P} \subset \mathcal{L}$, then

$$\sigma(\mathcal{P}) \subset \mathcal{L}$$

(b) If $\mathcal{P} \subset 2^\Omega$ is a π-system, then

$$\sigma(\mathcal{P}) = \mathcal{L}(\mathcal{P})$$

i.e., the minimal σ-algebra over $\mathcal{P}$ equals the minimal λ-system over $\mathcal{P}$.

Applications of Dynkin's Theorem

Proposition 2.2.3. Let P_1, P_2 be two probability measures on $(\Omega, \mathcal{B})$, then the class

$$\mathcal{L} := \{A \in \mathcal{B} : P_1(A) = P_2(A)\}$$

is a λ-system.

Corollary 2.2.1. Let P_1, P_2 be two probability measures on $(\Omega, \mathcal{B})$ and $\mathcal{P}$ is a π-system such that

$$\forall A \in \mathcal{P} : P_1(A) = P_2(A)$$

Then

$$\forall B \in \sigma(\mathcal{P}) : P_1(B) = P_2(B)$$

Proof. By Prop. 2.2.3, $\mathcal{L} := \{A \in \mathcal{B} : P_1(A) = P_2(A)\}$ is a λ-system. But $\mathcal{P} \subset \mathcal{L}$ and so by Dynkin's Theorem $\sigma(\mathcal{P}) \subset \mathcal{L}$.

$\square$

Corollary 2.2.2. **(Uniqueness of probability measure defined by its distribution function)**. Let $\Omega = \mathbb{R}$. Let P_1, P_2 be two probability measures on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\forall x \in \mathbb{R} : F_1(x) = F_2(x).$$

Then $P_1(A) = P_2(A)$, $A \in \mathcal{B}(\mathbb{R})$.

Proof. Define the π-system

$$\mathcal{P} = \{(-\infty, x] : x \in \mathbb{R}\}$$

We have $\sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$ (why?). So, by Prop. 2.2.3,

$$\forall x \in \mathbb{R} : F_1(x) = F_2(x) \iff P_1 = P_2 \text{ on } \sigma(\mathcal{P}) = \mathcal{B}(\mathbb{R})$$

□

Proof of Proposition 2.2.3. We show that the **new** definition is satisfied. Conditions (i) and (ii) are satisfied (*prove ...*).

Now take

$$\{A_n \in \mathcal{L} : A_n A_m = \emptyset,\ n \neq m\}$$

Then

$$\begin{aligned} \forall j \quad &: \quad P_1(A_j) = P_2(A_j) \\ &\Longrightarrow \quad P_1(\cup_j A_j) = \sum_j P_1(A_j) = \sum_j P_2(A_j) = P_2(\cup_j A_j) \\ &\Longrightarrow \quad \cup_j A_j \in \mathcal{L} \end{aligned}$$

□

Attention:

Subsection 2.3 (Two constructions of discrete spaces) is self-study

Semialgebras

Definition 2.4.1. Class $\mathcal{S} \subseteq 2^{\Omega}$ is a semialgebra if

(i) $\emptyset, \Omega \in \mathcal{S}$

(ii) $\mathcal{S}$ is a π-system

(iii) If $A \in \mathcal{S}$, then there exist finite n and disjoint $C_1, \dots C_n \in \mathcal{S}$ such that

$$A^c = \sum_{i=1}^{n} C_i$$

Extension Theorems (no proof)

Suppose $\mathcal{G}_1 \subset \mathcal{G}_2 \subset 2^\Omega$ are two structures of subsets of Ω. Consider two set functions

$$P_i : \mathcal{G}_i \longrightarrow [0, 1], \ i = 1, 2$$

If

$$P_{2|\mathcal{G}_1} = P_1, \text{ i.e., } P_2(A_1) = P_1(A_1), \ A_1 \in \mathcal{G}_1$$

then

P_1 is a *restriction* of P_2 to $\mathcal{G}_1$

P_2 is an *extension* of P_1 to $\mathcal{G}_2$ (this latter generally not necessarily unique!)

*Lemma 2.4.1. The field generated by a semialgebra (**study the proof and compare Problem 1.20!**).* Suppose $\mathcal{S}$ is a semialgebra of subsets of Ω. Then

$$\mathcal{A}(\mathcal{S}) = \left\{ \sum_{i \in I} S_i : I \text{ finite}, \{S_i, i \in I\} \text{ disjoint}, S_i \in \mathcal{S} \right\}$$

is the family of all sums of finite families of mutually disjoint subsets of Ω in $\mathcal{S}$.

Theorem 2.4.1. (First Extension Theorem). Suppose $\mathcal{S}$ is a semialgebra of subsets of Ω and $P : \mathcal{S} \longrightarrow [0, 1]$ is σ-additive on $\mathcal{S}$ and satisfies $P(\Omega) = 1$. There is a unique extension P' of P to $\mathcal{A}(\mathcal{S})$, defined by

$$P'(\sum_{i \in I} S_i) = \sum_{i \in I} P(S_i)$$

which is a probability measure on $\mathcal{A}(\mathcal{S})$; that is $P'(\Omega) = 1$ and P' is σ-additive on $\mathcal{A}(\mathcal{S})$.

Theorem 2.4.2. (Second Extension Theorem). A probability measure P defined on a field $\mathcal{A}$ of subsets has a unique extension to a probability measure on $\sigma(\mathcal{A})$,

Combine these to obtain *Thm 2.4.3 ...*

Lebesgue Measure on $[0, 1]$

$\Omega = 1$

$\mathcal{B} = \mathcal{B}((0, 1])$

$\mathcal{S} = \{(a, b] : 0 \leq a \leq b \leq 1\}$, it is a *semi*algebra!

Define $\lambda : \mathcal{S} \longrightarrow [0, 1]$ by

$$\lambda(\emptyset) = 0, \;\; \lambda(a, b] = b - a$$

Note $\lambda(A) \geq 0$.

To show that λ has a unique extension, we have to show λ is σ-additive

Finite additivity

Take

$$a = a_1, b_1 = a_2, \ldots, b_{k-1} = a_k, b_k = b$$

such that

$$(a, b] = \sum_{i=1}^{k} (a_i, b_i] \in \mathcal{S}$$

We see that (*prove ...*)

$$\lambda(a, b] = \sum_{i=1}^{k} \lambda(a_i, b_i]$$

Denumerable additivity

Compactness:

- A set is called *compact*, if from any open cover of the set, we can select a finite open subcover.
- A set in metric space is called *sequentially compact* if from any sequence of its elements, it is possible to select a convergent subsequence with a limit in this set.
- Therefore compact sets have to be *closed.*
- In *finite-dimensional* metric spaces, *bounded closed sets are compact*.

Proof of σ-additivity

Let

$$(a, b] \subset \bigcup_{i=1}^{\infty} (a_i, b_i]$$

We first prove

$$b - a \leq \sum_{i=1}^{\infty} (b_i - a_i).$$

We have

$$[a + \varepsilon, b] \subset \bigcup_{i=1}^{\infty} (a_i, b_i + \frac{\varepsilon}{2^i})$$

The infinite sum on the right is an *infinite open cover* of the *compact* set on the left.

Therefore we can choose a *finite* subcover so that

$$\exists N : [a + \varepsilon, b] \subset \bigcup_{i=1}^{N} (a_i, b_i + \frac{\varepsilon}{2^i})$$

It is enough to prove

$$b - a - \varepsilon \leq \sum_{i=1}^{N} (b_i - a_i + \frac{\varepsilon}{2^i}) \qquad (*)$$

which implies

$$b - a \leq 2\varepsilon + \sum_{i=1}^{\infty} (b_i - a_i)$$

but since $\varepsilon > 0$ is arbitrary (and so it can be arbitrarily small),

$$b - a \leq \sum_{i=1}^{\infty} b_i - a_i \iff \lambda(a, b] \leq \sum_{i=1}^{\infty} \lambda(a_i, b_i]$$

Instead of proving (*), it is sufficient to prove that

$$[a,b] \subset \bigcup_{i=1}^{N} (a_i, b_i) \implies b - a \leq \sum_{i=1}^{N} (b_i - a_i) \qquad (**)$$

Induction proof:

(**) is satisfied for $N = 1$.

Assume (**) is satisfied for $N - 1$.

Suppose

$$a_N = \max_{1 \leq i \leq N} a_i$$

and

$$a_N \leq b \leq b_N$$

Case 1: $a_N \leq a$

$$b - a \leq b_N - a_N \leq \sum_{i=1}^{N}(b_i - a_i)$$

Case 2: $a_N > a$

$$[a, b] \subset \bigcup_{i=1}^{N}(a_i, b_i) \implies [a, a_N] \subset \bigcup_{i=1}^{N-1}(a_i, b_i)$$

$$\overset{induction}{\implies} a_N - a \leq \sum_{i=1}^{N-1}(b_i - a_i)$$

Furthermore

$$
\begin{aligned}
b - a &= b - a_N + a_N - a \\
&\leq b - a_N + \sum_{i=1}^{N-1}(b_i - a_i) \leq b_N - a_N + \sum_{i=1}^{N-1}(b_i - a_i) \\
&= \sum_{i=1}^{N}(b_i - a_i)
\end{aligned}
$$

Now the other way around. Suppose

$$(a, b] = \sum_{i=1}^{\infty} (a_i, b_i]$$

Claim:

$$\exists n : \lambda(a, b] = b - a \geq \sum_{i=1}^{n} \lambda(a_i, b_i] = \sum_{i=1}^{n} (b_i - a_i)$$

Indeed,

$$(a, b] \setminus \sum_{i=1}^{n} (a_i, b_i] = \sum_{j=1}^{m} I_j$$

and by finite additivity

$$\lambda(a, b] = \sum_{i=1}^{n} \lambda(a_i, b_i] + \sum_{j=1}^{m} \lambda(I_j) \geq \sum_{i=1}^{n} \lambda(a_i, b_i]$$

If $n \longrightarrow \infty$,

$$\lambda(a, b] \geq \sum_{i=1}^{\infty} \lambda(a_i, b_i]$$

$\square$

Construction of probability measure on $\mathbb{R}$ with given distribution function $F(x)$

$$P_F((-\infty, x]) = F(x).$$

Define the left-continuous inverse of F

$$F^{\longleftarrow}(y) = \inf\{s : F(s) \geq y\},\ 0 < y \leq 1$$

and set

$$A(y) = \{s : F(s) \geq y\}$$

Properties:

(a) $A(y)$ is closed

(b) Therefore,

$$\inf A(y) \in A(y) \Longleftrightarrow F(F^{\leftarrow}(y)) \geq y$$

(c) Furthermore

$$[F^{\leftarrow}(y) > t \iff y > F(t)] \iff [F^{\leftarrow}(y) \leq t \iff y \leq F(t)]$$

Lemma 2.5.1.

If $A \in \mathcal{B}(\mathbb{R})$, then

$$\xi_F(A) = \{x \in (0,1] : F^{\longleftarrow}(x) \in A\} \in \mathcal{B}((0,1])$$

Proof, self-study

Construction of P_F

Define, for $A \in \mathcal{B}(\mathbb{R})$,

$$P_F(A) = \lambda(\xi_F(A))$$

Verify it is a probability measure ... Also,

$$\begin{aligned} P_F((-\infty, x]) &= \lambda(\xi_F((-\infty, x])) \\ &= \lambda\{y \in (0,1] : F^{\leftarrow}(y) \leq x\} \\ &= \lambda\{y \in (0,1] : y \leq F(x)\} \\ &= \lambda((0, F(x)]) = F(x) \end{aligned}$$

□