Probability and Statistical Inference

Tianqi Zhang
Emory University

Apr 17th 2025

3 Discrete Probability Measures

3.1 Discrete Probability Measures

We start with discrete, countable latent space Ω.

Definition 3.1.1 (Discrete Probability Measure). A discrete probability measure on sample space Ω, finite or countable, is a sequence of $\{p_\omega\}_{\omega\in\Omega}$ of non-negative real numbers such that:

1. $p_\omega \geq 0$, $\forall \omega \in \Omega$,
2. $\sum_{\omega\in\Omega} p_\omega = 1$

A general definition that works not only for finite Ω but also for countable Ω since it allows in both cases to compute for any random event $A \subseteq \Omega$.

$$P(A) = \sum_{\omega\in A} P_\omega$$

Definition 3.1.2 (Measure/Distribution).
A measure P on Ω is a mapping from the power set of the latent space Ω.

$$P : \mathscr{P}(\Omega) \to [0, \infty]$$

such that the following two axioms are satisfied:

1. **Non-negativity:** $\forall A \subseteq \Omega, P(A) \geq 0$

2. **Countable additivity:** (Or sigma additivity) For disjoint $A_n \subset \Omega$,

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_k P(A_n)$$

3. **Empty set measurability:** $P(\varnothing) = 0$.

Note that for P to be a probability measure: $P(A) \in [0,1]\ \forall A \in \mathscr{P}(\Omega)$, with $P(\Omega) = 1$.

Theorem 3.1.3 Komolgrov Axioms
$P : \mathscr{P}(\Omega) \to [0,1]$ is a probability measure if the above conditions are true.

Remark 3.1 *inclusion of infinity: In most measure-theoretic contexts, it is permissible for certain subsets $\mathcal{A} \subset \Omega$ to have infinite measure. Consequently, the codomain of a measure is typically extended to include infinity. This is commonly represented as the set of positive real numbers together with infinity, denoted by $[0,\infty) \cup \{\infty\}$. For simplicity, this notation is often abbreviated as $[0,\infty]$.*

Example 3.1.4 Common Measures

- **Counting Measure:** $\mu(A) = |A|$.
- **Dirac Measure at $p \in \Omega$:**

$$\delta_p(A) = \begin{cases} 1, & \text{if } p \in A, \\ 0, & \text{otherwise.} \end{cases}$$

- **Lebesgue Measure on $\mathbb{R}$:** For simple intervals $[a,b) \subset \mathbb{R}$ with $a \leq b$. Lebesgue measure μ is defined to be $b - a$.

Proposition 3.1.5 Properties of Komolgrov axioms:

1. $P(\overline{A}) = 1 - P(A)$ where $\overline{A} \equiv \Omega \backslash A$.
2. Define $A + B$ as the disjoint union of A, B, i.e., $A \cap B = \varnothing$, then $P(A+B) = P(A) + P(B)$.
3. If $B \subset A$, define $A - B \equiv A \cap \overline{B}$ then $P(A - B) = P(A) - P(B)$.
4. Partition Ω into disjoint $\{H_i\}_{i\in\mathbb{N}}$, i.e. $H_i \cap H_j = \varnothing\ \forall i,j$ and $\bigcup_i H_i = \Omega$. Then $\forall A \subset \Omega, P(A) = \sum_i P(A \cap H_i)$.

Corollary 3.1.6 Monotonicity:
If $A \subseteq B$, then $0 \leq P(A) \leq P(B) \leq 1$.

Corollary 3.1.7 Sylvester Formula:
For any collection of subsets $\{A_i\}$ of Ω

$$P\left(\bigcup_{i=1}^{n}\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) + \cdots + (-1)^{n-1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

Or a more intuitive version, for any $A, B \subseteq \Omega$,

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Remark 3.2 *Banach-Tarski Paradox: If Ω is countable, we can assign a finite measure to all subsets of Ω satisfying the Kolmogorov axioms. However, if Ω is uncountable, these axioms can lead to paradoxes.*

To address these issues, we may need to restrict our measure $P : \mathscr{P}(\Omega) \to [0, \infty]$ to a carefully chosen collection of subsets $\mathcal{F} \subset \Omega$. This restriction sets the foundation for further discussion on the role of σ-algebras in measure theory.

3.2 Results and Properties

Proposition 3.2.1 Boole's Inequality:
For $A_i \subseteq \Omega$, not necessarily disjoint:

$$P\left(\bigcup_{i} A_i\right) \leq \sum_{i} P(A_i).$$

Proof: **??**

Proposition 3.2.2 Bonferroni's Inequality:
For $A_i \subseteq \Omega$:

$$P\left(\bigcap_{i} A_i\right) \geq \sum_{i} P(A_i) - (n-1).$$

This result is useful for multiple hypothesis testing.
Proof: **??**

Proposition 3.2.3 De Morgan's Laws:

$$P\left(\bigcap_{i} A_i\right) = P\left(\bigcup_{i} \overline{A_i}\right).$$

Some notations for limits of sets:

- **Increasing Sequence:** (A_n) is called an increasing sequence if $(A_n \subseteq A_{n+1})$, and

$$\lim_{n\to\infty} \uparrow A_n \equiv \bigcup_{k=1}^{\infty} A_n$$

- **Decreasing Sequence:** (B_n) is called an increasing sequence if $(B_n \supseteq B_{n+1})$, and

$$\lim_{n\to\infty} \downarrow B_n \equiv \bigcap_{k=1}^{\infty} B_n$$

Proposition 3.2.4 Continuity of Measures:

The continuity of measures is preserved under increasing and decreasing set limits.

Let (A_n) be an increasing sequence of sets:
Increasing continuity:

$$\lim_{n\to\infty} P(A_n) = P\Big(\lim_{n\to\infty} \uparrow A_n\Big) \equiv P\left(\bigcup_{n=1}^{\infty} A_n\right).$$

Let (B_n) be a decreasing sequence of sets:
Decreasing Sequence:

$$\lim_{n\to\infty} P(B_n) = P\Big(\lim_{n\to\infty} \downarrow B_n\Big) \equiv P\left(\bigcap_{n=1}^{\infty} B_n\right).$$