Probability Definitions

Probability Space (or **Sample Space**): ordered triple $(\Omega, \mathcal{F}, P)$.

- Ω is a set (possible outcomes); elements are ω called elementary outcomes.

- $\mathcal{F}$ is a family of subsets (**events**) of Ω with the property that $\mathcal{F}$ is a σ-field (or Borel field or σ-algebra):

  1. Empty set $\emptyset$ and Ω are members of $\mathcal{F}$.

  2. $A \in \mathcal{F}$ implies $A^c = \{\omega \in \Omega : \omega \notin A\} \in \mathcal{F}$.

  3. $A_1, A_2, \cdots$ in $\mathcal{F}$ implies $A = \cup_{i=1}^{\infty} A_i \in \mathcal{F}$.

- P a function, domain $\mathcal{F}$, range a subset of $[0,1]$ satisfying:

  1. $P(\emptyset) = 0$ and $P(\Omega) = 1$.

  2. **Countable additivity**: $A_1, A_2, \cdots$ **pair-wise disjoint** $(j \neq k \ A_j \cap A_k = \emptyset)$

$$P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

Axioms guarantee can compute probabilities by usual rules, including approximation.

Consequences of axioms:

$$A_i \in \mathcal{F}; i = 1, 2, \cdots \text{ implies } \cap_i A_i \in \mathcal{F}$$

$$A_1 \subset A_2 \subset \cdots \text{ implies } P(\cup A_i) = \lim_{n\to\infty} P(A_n)$$

$$A_1 \supset A_2 \supset \cdots \text{ implies } P(\cap A_i) = \lim_{n\to\infty} P(A_n)$$

Vector valued random variable: function $X : \Omega \mapsto R^p$ such that, writing $X = (X_1, \ldots, X_p)$,

$$P(X_1 \le x_1, \ldots, X_p \le x_p)$$

is defined for any constants $(x_1, \ldots, x_p)$. Formally the notation

$$X_1 \le x_1, \ldots, X_p \le x_p$$

is a subset of Ω or **event**:

$$\{\omega \in \Omega : X_1(\omega) \le x_1, \ldots, X_p(\omega) \le x_p\} \, .$$

Remember X is a function on Ω so X_1 is also a function on Ω.

In almost all of probability and statistics the dependence of a random variable on a point in the probability space is hidden! You almost always see X not $X(\omega)$.

Borel σ-field in R^p: smallest σ-field in R^p containing every open ball.

Every common set is a Borel set, that is, in the Borel σ-field.

An R^p valued **random variable** is a map $X : \Omega \mapsto R^p$ such that when A is Borel then $\{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{F}$.

Fact: this is equivalent to

$$\{\omega \in \Omega : X_1(\omega) \le x_1, \ldots, X_p(\omega) \le x_p\} \in \mathcal{F}$$

for all $(x_1, \ldots, x_p) \in R^p$.

Jargon and notation: we write $P(X \in A)$ for $P(\{\omega \in \Omega : X(\omega) \in A\})$ and define the **distribution** of X to be the map

$$A \mapsto P(X \in A)$$

which is a probability on the set R^p with the Borel σ-field rather than the original Ω and $\mathcal{F}$.

Cumulative Distribution Function (CDF) of X: function F_X on R^p defined by

$$F_X(x_1,\ldots,x_p) = P(X_1 \le x_1,\ldots,X_p \le x_p)\,.$$

Properties of F_X (usually just F) for $p = 1$:

1. $0 \le F(x) \le 1$.

2. $x > y \Rightarrow F(x) \ge F(y)$ (monotone non-decreasing).

3. $\lim_{x\to-\infty} F(x) = 0$ and $\lim_{x\to\infty} F(x) = 1$.

4. $\lim_{x\searrow y} F(x) = F(y)$ (right continuous).

5. $\lim_{x\nearrow y} F(x) \equiv F(y-)$ exists.

6. $F(x) - F(x-) = P(X = x)$.

7. $F_X(t) = F_Y(t)$ for all t implies that X and Y have the same distribution, that is, $P(X \in A) = P(Y \in A)$ for any (Borel) set A.

Defn: Distribution of rv X is **discrete** (also call X discrete) if $\exists$ countable set $x_1, x_2, \cdots$ such that

$$P(X \in \{x_1, x_2 \cdots\}) = 1 = \sum_i P(X = x_i)\,.$$

In this case the **discrete density** or **probability mass function** of X is

$$f_X(x) = P(X = x)\,.$$

Defn: Distribution of rv X is **absolutely continuous** if there is a function f such that

$$P(X \in A) = \int_A f(x)dx \qquad (1)$$

for any (Borel) set A. This is a p dimensional integral in general. Equivalently

$$F(x) = \int_{-\infty}^{x} f(y)\,dy\,.$$

Defn: Any f satisfying (**??**) is a **density** of X.

For most x F is differentiable at x and

$$F'(x) = f(x)\,.$$

Example: X is Uniform[0,1].

$$F(x) = \begin{cases} 0 & x \leq 0 \\ x & 0 < x < 1 \\ 1 & x \geq 1\,. \end{cases}$$

$$f(x) = \begin{cases} 1 & 0 < x < 1 \\ \text{undefined} & x \in \{0, 1\} \\ 0 & \text{otherwise}\,. \end{cases}$$

Example: X is exponential.

$$F(x) = \begin{cases} 1 - e^{-x} & x > 0 \\ 0 & x \leq 0\,. \end{cases}$$

$$f(x) = \begin{cases} e^{-x} & x > 0 \\ \text{undefined} & x = 0 \\ 0 & x < 0\,. \end{cases}$$