Probability And Statistics

AIMS Rwanda

August 17, 2025

Definition 1

Probability will be define on a collection of events that should be a σ-field. A collection $\mathcal{F}$ of subsets of Ω is called a σ-**field** if it satisfies the following conditions:

1. $\emptyset \in \mathcal{F}$ or $\Omega \in \mathcal{F}$,
2. If $A \in \mathcal{F}$ then $\bar{A} \in \mathcal{F}$. Every element A of $\mathcal{F}$ is called an event.
3. If $A_1, A_2, \cdots \in \mathcal{F}$ then $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Two events A and B are said to be independent if and only if $A \cap B = \emptyset$. For every event $A \in \mathcal{F}$,

$$A \cap \emptyset = \emptyset, A \cup \emptyset = A, A \cup \bar{A} = \Omega, A \cap \bar{A} = \emptyset.$$

Definition 2

A **probability measure** P on $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \to [0, 1]$ satisfying:

1. $P(A) > 0$, for all $A \in \mathcal{F}$,
2. $P(\Omega) = 1$,
3. If $A_1, A_2, \cdots$ is a collection of disjoint members of $\mathcal{F}$

$$\left(A_i \cap A_j = \emptyset \text{ for all pairs satisfying } i \neq j\right),$$

then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

The triplet $(\Omega, \mathcal{F}, P)$ is called a **probability space**.

Properties

1. $P(\emptyset) = 0$,
2. For every A, $P(\bar{A}) = 1 - P(A)$,
3. Let A and B be two events such that $A \subseteq B$, then
$$P(A) \leq P(B)$$
4. Let A and B two events,
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
5. For every A, $P(A) \leq 1$.
6. **Finite additivity**: Let $A_1, \cdots, A_n$ be disjoint events, then
$$P(A_1 \cup \cdots \cup A_n) = P(A_1) + \cdots + P(A_n).$$
7. **Boole's inequality**: For any sequence $A_1, A_2, A_3, \cdots$ of events,
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P(A_i)$$

Exercise 1

Let A, B, C be 3 events. Consider $E_1 = A \cap \bar{B} \cap \bar{C}$ and $E_2 = A \cap (B \cup C)$.

1. Are E_1 and E_2 disjoints? Calculate $E_1 \cup E_2$.
2. Suppose that $P(A) = 0.6$, $P(B) = 0.4$, $P(C) = 0.3$, $P(A \cap B) = 0.2$, $P(B \cap C) = 0.1$, $P(A \cap C) = 0.1$ and

$$P(A \cap B \cap C) = 0.05.$$

   Calculate $P(E_1)$ and $P(E_2)$.

Solution

1. E_1 and E_2 are disjoints. Indeed,

$$\begin{aligned} E_1 \cap E_2 &= (A \cap \bar{B} \cap \bar{C}) \cap (A \cap (B \cup C)) \\ &= (A \cap \overline{B \cup C}) \cap (A \cap (B \cup C)) \\ &= A \cap (\overline{B \cup C} \cap B \cup C) \\ &= A \cap \emptyset \\ &= \emptyset. \end{aligned}$$

On the other hand:

$$\begin{aligned} E_1 \cup E_2 &= (A \cap \bar{B} \cap \bar{C}) \cup (A \cap (B \cup C)) \\ &= A \cup \emptyset \\ &= A. \end{aligned}$$

2. $P(E_2) = 0.25$ and $P(E_1) = 0.35$.

Definition 3

Consider two events A and B. Suppose that B occurs.
What can we say about the chance of occurrence for event A?
What is the probability of A given B?
If $P(B) > 0$ **the conditional probability** of A given B is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Lemma: For any events A and B such that $0 < P(B) < 1$,
$P(A) = P(A|B)P(B) + P(A|\bar{B})P(\bar{B})$.
More generally,
Let $B_1, \cdots, B_n$ be a partition of Ω. Then, for any event A,

$$P(A) = \sum_{i=1}^{n} P(A|B_i)P(B_i).$$

Bayes' Theorem

Let $B_1, \cdots, B_n$ be a partition of Ω such that $P(B_i) > 0$ for each i. If $P(A) > 0$ then, for each $i = 1, \cdots, n$,

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{j=1}^{n} P(A|B_j)P(B_j)}.$$

Exercise 2

Larry divides his email into three categories
$A_1 =$ "*spam*", $A_2 =$ "*lowpriority*" and $A_3 =$ "*highpriority*".
From previous experience, he finds that $P(A_1) = 0.7$, $P(A_2) = 0.2$ and $P(A_3) = 0.1$.
Let B be the event that the email contains the word "free".
From previous experience, Larry knows that
$P(B|A_1) = 0.9$, $P(B|A_2) = 0.01$ and $P(B|A_3) = 0.01$. Larry receives an email with the word "free".
What is the probability that it is "spam"?

Continuous random variables

- X : Lap time for a given operation,
- $\Omega = \mathbb{R}_+$.
- A possible outcome is: $1h, 27\,\min., 30\,\sec., 5/10\text{msec.}, \cdots$
- In this case, it is not possible to associate an exact value to the outcome.
- Event: $X^{-1}(]a,b[) = \{\omega \in \Omega | a < X(\omega) < b\}$.
- The distribution is characterized by:

$$P(X^{-1}(]a,b[) = P(a < X(\omega) < b) = P(a < X < b).$$

- When a random variable is continuous, we consider:

$$P(x < X < x+h) = P(X \leq x+h) - P(X \leq x)$$

Definition 4

A **cumulative distribution function** (c.d.f.) is a function $F_X : \mathbb{R} \to [0, 1]$ defined by:

$$F_X(x) = P(X \leq x).$$

$F : \mathbb{R} \to [0, 1]$ is a cumulative distribution function for some probability P if and only if F satisfies the following three conditions:

1. F is non-decreasing: $x_1 < x_2$ implies that $F(x_1) < F(x_2)$.
2. F is normalized:
$$\lim_{x \to -\infty} F(x) = 0 \text{ and } \lim_{x \to +\infty} F(x) = 1.$$
3. F is right-continuous:
$$F(x) = \lim_{y \to x, y > x} F(y).$$

- We define $f_X(x)$, the **probability density function** (p.d.f.) of X as:

$$f_X(x) = \lim_{h \to 0} \frac{1}{h} \left(F_X(x+h) - F_X(x) \right)$$

- The p.d.f. is the derivative of F_X, the c.d.f. i.e.

$$f_X(x) = F'_X(x).$$

- We have:

$$F_X(x) = P(X \leq x) = \int_{y \leq x} f_X(y)dy \text{ and } \int_V f_X(y)dy = 1$$

Remark: Discrete random variable

$$F_X(x) = P(X \leq x) = \sum_{y \in V | y \leq x} P(X = y)$$

Some classical rv's

1. A r.v. X is said to be a **Bernoulli's** r.v. with parameter p iff:

$$P(X = x) = p^x(1-p)^{1-x},\, x \in \{0,1\},\, 0 < p < 1.$$

Notation: $X \sim B(1,p)$

2. A r.v. X follows a **binomial** distribution with parameters n and p iff:

$$P(X = x) = \binom{n}{x} p^x(1-p)^{n-x},\, x \in \{0,\cdots,n\},\, 0 < p < 1.$$

Notation: $X \sim B(n,p)$

3. A r.v. X has an exponential distribution with parameter $\theta > 0$ iff its p.d.f has the form:

$$f(x|\theta) = \frac{1}{\theta}\exp\{-\frac{x}{\theta}\},\, x \in \mathbb{R}_+.$$

Notation: $X \sim \mathcal{E}(\theta)$

Exercise 3

For each of the preceding rv's,

- Check that they are well-defined
- Compute the cdf of an the exponential rv.

Definition 6

1. The expected value or mean or first moment of a random variable X is defined as:

$$E(X) = \begin{cases} \sum_{x \in V} xP(X = x) & \text{if } X \text{ is discrete,} \\ \int_V x f_X(x)\, dx & \text{if } X \text{ is continuous.} \end{cases}$$

2. Let g be a regular function. Then:

$$E(g(X)) = \begin{cases} \sum_{x \in V} g(x)P(X = x) & \text{if } X \text{ is discrete,} \\ \int_V g(x) f_X(x)\, dx & \text{if } X \text{ is continuous.} \end{cases}$$

3. q-th moments

$$E(X^q) = \begin{cases} \sum_{x \in V} x^q P(X = x) & \text{if } X \text{ is discrete,} \\ \int_V x^q f_X(x)\, dx & \text{if } X \text{ is continuous.} \end{cases}$$

Moment generating function (m.g.f): $M_X(t) = E[e^{tX}]$.

$$\forall b \in \mathbb{R},\ E(X+b) = E(X) + b.$$

$$\forall a \in \mathbb{R},\ E(aX) = aE(X).$$

Exercise: A fair coin is tossed n times. Let X be the number of tails. X follows a Binomial distribution with parameters (n, p) where $p = \frac{1}{2}$.

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x},\ x = 0, 1, \ldots, n.$$

Let a be a real number. Consider the random variable $Y = \frac{a^x}{2^n}$. Compute $E(Y)$.

Definition

The variance of a random variable is defined as:

$$Var(X) = E\left((X - E(X))^2\right).$$

$$Var(X) = \begin{cases} \sum_{x \in V}(x - E(X))^2 P(X = x) & \text{if } X \text{ is discrete,} \\ \int_V (x - E(X))^2 f_X(x)\, dx & \text{if } X \text{ is continuous.} \end{cases}$$

Properties

$$\text{Var}(X) = E(X^2) - (E(X))^2.$$

$$\forall b \in \mathbb{R}, \ \text{Var}(X + b) = \text{Var}(X).$$

The variance is quadratic:

$$\forall a, \ \text{Var}(aX) = a^2 \cdot \text{Var}(X).$$

Exercise: Compute the variance for:

$$X \sim B(1, p).$$

$$X \sim \mathcal{E}(\theta).$$

Exercises

Compute the m.g.f. of a Bernoulli's with parameter p and a binomial r, with parameters n and p. Using the moment generating function,

- Compute the expectation and the variance of $B(n, p)$.
- Prove the following: Let $X_1, X_2, \cdots, X_n$ be n independent Bernoulli $r.v'$s with parameter p. Then,

$$X = \sum_{i=1}^{n} X_i$$

  has a binomial distribution with parameter n and p