CS 4450

Coding and Information Theory

Overview of Probability (B)

Dr. Xuejun Liang

Overview of Probability

1. Fundamentals of Probability
2. Random Variables and its Characteristics
3. Statistical Averages

Summary of Fundamentals of Probability

1. Operations of Events
2. Axioms of Probability
3. Properties of Probability
4. Conditional Probability
5. Independent Events
6. Total Probability

1.6 Total Probability

Let $A_1, A_2, \ldots, A_n$ be mutually exclusive ($A_i \cap A_j = \emptyset$, for $i \neq j$) and exhaustive ($\bigcup_{i=1}^{n} A_n = S$). Now B is any event in S. Then

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(B/A_i)P(A_i) \qquad (1.16)$$

This is known as the total probability of event B.

If $A = A_i$ in Equation (1.12), then using (1.16) we have the following important theorem (Bayes theorem):

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)} \qquad (1.17)$$

Recall $P(A/B) = \frac{P(B/A)P(A)}{P(B)}$ (1.12)

Example: Cancer Test

- What we know
  - P(C) = 0.001
  - P(¬C) = 0.999
  - P(POS | C) = 0.8
  - P(POS | ¬C) = 0.1
- Compute probability of Cancer when having a positive cancer test
  - P(C | POS) = ?

$$P(A_i/B) = \frac{P(B/A_i)P(A_i)}{\sum_{i=1}^{n} P(B/A_i)P(A_i)}$$

$$P(C|POS) = \frac{P(POS|C)P(C)}{P(POS)}$$

$$P(C|POS) = \frac{0.8 \times 0.001}{0.1007} \approx 0.008$$

$$P(POS) = P(POS|C)P(C) + P(POS|\neg C)P(\neg C)$$

$$P(POS) = 0.8 \times 0.001 + 0.1 \times 0.999 = 0.1007$$

2. Random Variable and its Characteristics

- Random Variable (RV)
  - A function X: S → R, where S = $\{s_1, s_2, s_3, \ldots\}$ is a sample space and R is the set of real number.
- Discrete Random Variable:
  - A random variable whose range (number of values) is finite or countably infinite.
  - An example of a discrete RV is the number of cars passing through a street in a finite time.
- Continuous Random Variable
  - A random variable whose range is one or more intervals on the real line.
  - An example of this type of variable is the measurement of noise voltage across the terminals of some electronic device.

Example

Example 1.11: Two unbiased coins are tossed. Suppose that RV X represents the number of heads that can come up. Find the values taken by X.

Solution: The sample space S contains four sample points. Thus, S = {HH, HT, TH, TT}. Table 1.1 illustrates the sample points and the number of heads that can appear (i.e., the values of X)

Table 1.1 *Random Variable and Its Values*

Outcome	Value of *X* (Number of Heads)
HH	2
HT	1
TH	1
TT	0

Discrete Random Variable and Probability Mass Function

Consider a discrete RV X that can assume the values $x_1, x_2, x_3, \ldots$. Suppose that these values are assumed with probabilities given by

$$P(X = x_i) = f(x_i), i = 1, 2, 3 \ldots \qquad (1.18)$$

This function $f(x)$ is called the probability mass function (PMF), discrete probability distribution or probability function of the discrete RV X.

$f(x)$ satisfies the following properties

1. $0 \leq f(x_i) \leq 1, i = 1, 2, 3, \quad .$
2. $f(x) = 0$, if $x \neq x_i$ $(i = 1, 2, 3, \quad)$.
3. $\sum_i f(x_i) = 1$.

Example

Example 1.12: Find the PMF corresponding to the RV X of Example 1.11.

Table 1.1 *Random Variable and Its Values*

Outcome	**Value of *X* (Number of Heads)**
HH	2
HT	1
TH	1
TT	0

x_i	0	1	2
$f(x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Cumulative Distribution Function

The cumulative distribution function (CDF) or briefly the distribution function of a continuous or discrete RV X is given by.

$$F(x) = P(X \leq x), \ -\infty < x < \infty \qquad (1.19)$$

The CDF $F(x)$ has the following properties:

1. $0 \leq F(x) \leq 1.$
2. $F(x)$ is a monotonic non-decreasing function, i.e.,

   $$F(x_1) \leq F(x_2) \text{ if } x_1 \leq x_2.$$
3. $F(-\infty) = 0.$
4. $F(\infty) = 1.$
5. F(x) is continuous from the right, i.e.,

   $$\lim_{h \to 0^+} F(x + h) = F(x) \text{ for all } x.$$

Distribution Function for Discrete Random Variable

The CDF of a discrete RV X is given by

$$F(x) = P(X \leq x) = \sum_{u \leq x} f(u), \quad -\infty < x < \infty \qquad (1.20)$$

If X assumes only a finite number of values $x_1, x_2, x_3, \ldots$, ($x_1 \leq x_2 \leq x_3 \leq \cdots$)then the distribution function can be expressed as follows:

$$F(x) = \begin{cases} 0, & -\infty < x < x_1 \\ f(x_1), & x_1 \leq x < x_2 \\ f(x_1) + f(x_2), & x_2 \leq x < x_3 \\ \vdots & \\ f(x_1) + \cdots + f(x_n), & x_n \leq x < \infty \end{cases}$$

Example 1.13

(a) Find the CDF for the RV X of Example 1.12.

(b) Obtain its graph.

PMF

x_i	0	1	2
$f(x_i)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Continuous Random Variable and Probability Density Function

The distribution function of a continuous RV is represented as follows:

$$F(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,du, \quad -\infty < x < \infty \qquad (1.21)$$

where $f(x)$ satisfies the following conditions:

1. $f(x) \ge 0.$
2. $\int_{-\infty}^{\infty} f(x)\,dx = 1.$
3. $P(a < X < b) = \int_{a}^{b} f(x)\,dx$

$f(x)$ is known as the probability density function (PDF) or simply density function.

3. Statistical Averages

Expectation: The expectation or mean of an RV X is defined as follows:

$$\mu = E(X) = \begin{cases} \Sigma_i x_i \, f(x_i), & X\text{: discrete} \\ \int_{-\infty}^{\infty} x f(x)\, dx, & X\text{: continuous} \end{cases} \qquad (1.22)$$

Variance: The variance of an RV X is expressed as follows:

$$\mathrm{Var}(X) = \sigma^2 = E[(X-\mu)^2] = \begin{cases} \Sigma_i (x_i-\mu)^2 \, f(x_i), & X\text{: discrete} \\ \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx, & X\text{: continuous} \end{cases} \qquad (1.23)$$

Eq. (1.23) is simplified as follows:

$$\sigma^2 = E[X-\mu)^2] = E[X^2] - \mu^2 = E[X^2] - (E[X])^2 \qquad (1.24)$$

Standard Deviation: The positive square root of the variance (σ)

Example: Expectation and Variance

A discrete random variable X assumes 5 values: 2, 3, 4, 5, 6 with the probability mass function:

x_i	2	3	4	5	6
$P(x_i)$	0.2	0.15	0.3	0.15	0.2

Find the mean (expectation) and variance of X.

$$E(X) = \sum_i x_i p(x_i) = 2 \times 0.2 + 3 \times 0.15 + 4 \times 0.3 + 5 \times 0.15 + 6 \times 0.2$$
$$= 8 \times 0.2 + 8 \times 0.15 + 4 \times 0.3 = 1.6 + 1.2 + 1.2 = 4$$

$$E(X^2) = \sum_i {x_i}^2 p(x_i) = 4 \times 0.2 + 9 \times 0.15 + 16 \times 0.3 + 25 \times 0.15 + 36 \times 0.2$$
$$= 40 \times 0.2 + 24 \times 0.15 + 16 \times 0.3 = 8 + 3.6 + 4.8 = 16.4$$

$$V(X) = E(X^2) - E(X)^2 = 16.4 - 16 = 0.4$$