DMS625: Introduction to stochastic processes and their applications

Sourav Majumdar

Probability Review

Probability axioms

We define a probability space with the triple $(\Omega, \mathcal{A}, \mathbb{P})$. Ω is referred to as the outcome space, $\mathcal{A}$ as the event space and $\mathbb{P}$ as the probability measure, where $\mathbb{P} : \mathcal{A} \to \mathbb{R}$. Elements of Ω are referred to as outcomes, and elements of $\mathcal{A}$ are referred to as events.

Kolmogorov axiomatised probability theory by giving a certain set of axioms that a well behaved probability space must possess.

1. $\forall A \in \mathcal{A}, \mathbb{P}(A) \geq 0$
2. $\mathbb{P}(\Omega) = 1$
3. If $A_1, \dots, A_n$ are disjoint sets (also referred to as mutually exclusive events), then,

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

We have not yet defined what the event space is, or why it may be needed. First let's define the event space and give an example of a probability space.

We say $\mathcal{A}$ to be the event space if,

1. $\Omega \in \mathcal{A}$
2. If $A \in \mathcal{A}$, then $A' \in \mathcal{A}$
3. If $A_1, A_2, \dots, A_n \in \mathcal{A}$, then, $A = A_1 \cup A_2 \cup \dots A_n \in \mathcal{A}$

Example 0.1. *Consider the random experiment of tossing a fair coin. The two possible outcomes are H and T. Therefore, $\Omega = \{H, T\}$. The possible events whose probabilities one may be interested in are i) getting a head, ii) getting a tail, iii) getting a head or tail. There are no other possible events. We will denote set of all impossible events by ϕ, such as getting a head and tail, getting neither a head or tail, etc. Therefore the elements of the event space are $\mathcal{A} = \{\{H, T\}, \{H\}, \{T\}, \phi\}$. Since the coin is fair, $\mathbb{P}(H) = \mathbb{P}(T) = 1/2$. Verify that this forms a probability space.*

Exercises

1. Consider a random experiment of throw of a fair die. Identify the probability space.
2. Show that for any probability space, $(\Omega, \mathcal{A}, \mathbb{P})$, $\mathbb{P}(\phi) = 0$.

Remark 1. *One can verify that the event space $\mathcal{A}$ in the example, follows the mathematical definition of event space we outlined above. However, we didn't use the definition to arrive at the event space, but rather we considered the set of all subsets of Ω to be the elements of the event space. For finite sets, one can form a probability space with the event space containing all subsets of the outcome space. However, consider $\Omega = \mathbb{R}^+$, the set of positive real numbers, this is of interest when considering say the prices of stocks. It turns out that for sets such as $\mathbb{R}^+$, one will not be able to define a probability space with the event space containing the set of all subsets of $\mathbb{R}^+$. Instead we work with event spaces following the definition above, this will exclude certain subsets of $\mathbb{R}^+$. This is a rather deep result and we will not be delving into the details of this, but as students of stochastic processes this is good to know.*

Exercises

1. Using probability axioms, show the following,
   1. For any two events A_1, A_2, $\mathbb{P}(A_1 \cup A_2) = \mathbb{P}(A_1) + \mathbb{P}(A_2) - \mathbb{P}(A_1 \cap A_2)$
   2. For any three events A_1, A_2, A_3,$\mathbb{P}(A_1 \cup B \cup C) = \mathbb{P}(A_1) + \mathbb{P}(A_2) + \mathbb{P}(A_3) - \mathbb{P}(A_1 \cap A_2) - \mathbb{P}(A_1 \cap A_3) - \mathbb{P}(A_2 \cap A_3) + \mathbb{P}(A_1 \cap A_2 \cap A_3)$
   3. For events $A_1, \dots, A_n$, $\mathbb{P}(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n \mathbb{P}(A_i)$

Conditional probability

We define $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$ to be the probability of A given B has occurred, given $\mathbb{P}(B) > 0$.

Exercises

1. A family has two children. What is the probability that both are boys given that atleast one of them is a boy?

2. Show that if D_i are disjoint and $\mathbb{P}(C|D_i) = p$ independently of i, then $\mathbb{P}(C| \cup_i D_i) = p$.

3. Show that if C_i are disjoint, then $\mathbb{P}(\cup_i C_i|D) = \sum \mathbb{P}(C_i|D)$.

4. Show that if E_i are disjoint and $\cup_i E_i = \Omega$, then,

$$\mathbb{P}(C|D) = \sum_i \mathbb{P}(E_i|D)\mathbb{P}(C|E_i \cap D)$$

Independence

Two events A and B are independent if $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. The consequence of independence is that $\mathbb{P}(A|B) = \mathbb{P}(A)$, i.e., two events are independent if the occurrence of one doesn't influence the probability of another.

Exercises

1. Suppose we toss two fair dies. Let A denote the event that the sum of the faces of die is 6 and B denote the event that first die is 2. Are A and B independent?

2. Are two events being disjoint equivalent to them being independent?

We will call A, B, C to be mutually independent events, if $\mathbb{P}(ABC) = \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$, $\mathbb{P}(AB) = \mathbb{P}(A)\mathbb{P}(B)$, $\mathbb{P}(AC) = \mathbb{P}(A)\mathbb{P}(C), \mathbb{P}(BC) = \mathbb{P}(B)\mathbb{P}(C)$.

Formally $A_1, A_2, \ldots, A_n$ are said to be mutually independent iff,

$P(\cap_{j=1}^{k} A_{i_j}) = \prod_{i=1}^{k} \mathbb{P}(A_{i_j}), \forall k \leq n$, for every set of indices $1 \leq i_1 \leq \ldots i_k \leq n$.

It turns out that pairwise independence is not sufficient to ensure mutual independence.

Example 0.2. *Define a probability space with* $\Omega = \{1, 2, 3, 4\}$, *where* $\mathbb{P}(\{1\}) = \mathbb{P}(\{2\}) = \mathbb{P}(\{3\}) = \mathbb{P}(\{4\}) = 1/4$. *Now consider the events* $A = \{1, 2\}, B = \{1, 3\}, C = \{1, 4\}$. *It follows that* $\mathbb{P}(A) = \mathbb{P}(\{1\}) + \mathbb{P}(\{2\}) = 1/2$. *Similarly,* $\mathbb{P}(B) = \mathbb{P}(C) = 1/2$.

$$\mathbb{P}(AB) = \mathbb{P}(\{1\}) = \mathbb{P}(A)\mathbb{P}(B) = 1/4$$
$$\mathbb{P}(AC) = \mathbb{P}(\{1\}) = \mathbb{P}(A)\mathbb{P}(C) = 1/4$$
$$\mathbb{P}(BC) = \mathbb{P}(\{1\}) = \mathbb{P}(B)\mathbb{P}(C) = 1/4$$

Therefore, A, B, C *are pairwise independent events. Now consider* $\mathbb{P}(ABC) = \mathbb{P}(\{1\}) = 1/4$. *However,* $\mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C) = 1/8$ *and hence,* $\mathbb{P}(ABC) \neq \mathbb{P}(A)\mathbb{P}(B)\mathbb{P}(C)$. *Therefore, they aren't mutually independent events.*

Exercises

1. Show that there are $2^n - n - 1$ conditions to be satisfied for n mutually independent random variables.

Random variable

A random variable X is a function mapping the outcome space to the real numbers, $X : \Omega \to \mathbb{R}$. Since the underlying outcome space could be abstract such as heads or tails in case of a coin toss, for quantitative assessments on the outcomes we may be interested in random variables. So far we have defined how to evaluate the probability of an outcome (i.e. element of Ω). To evaluate the probability that a random variable takes a certain value, we shall evaluate the probability of the events for which the random variable takes that value. Formally,

$$\mathbb{P}(X \in S) = \mathbb{P}(\{\omega \in \Omega : X(\omega) \in S\})$$

Example 0.3. *Consider the toss of a fair coin. Define a random variable,* X, *such that* $X(\{H\}) = 1$ *and* $X(\{T\}) = -1$. *Then,* $\mathbb{P}(X = 1) = \mathbb{P}(\{H\}) = 1/2$. *Likewise,* $\mathbb{P}(X^2 = 1) = \mathbb{P}(\{H\} \cup \{T\}) = \mathbb{P}(\{H\}) + \mathbb{P}(\{T\}) = 1$. *Since,* $X^2(\{H\}) = 1$ *and* $X^2(\{T\}) = 1$.

The cumulative distribution function (CDF) of a random variable X is defined for any real number, by,

$$F(x) = \mathbb{P}(X \leq x)$$

1. $F(x)$ is a non-decreasing function in x.

2. $\lim_{x\to\infty} F(x) = 1$

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

For a discrete random variable, we define the probability mass function of X, $p(x) = \mathbb{P}(X = x)$. Some examples of discrete random variables,

1. Bernoulli random variable: For a random experiment whose outcomes are success($X = 1$) and failure($X = 0$), the probability mass function of X is given by, $p(1) = p$ and $p(0) = 1 - p$, then X is said to be a Bernoulli random variable. $X \sim \text{Bernoulli}(p)$

2. Binomial random variable: In a sequence of n independent Bernoulli trials, let X denote the number of successes in these n trials. X is said to be a Binomial random variable. Then, $p(k) = \mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$. $X \sim \text{Bin}(n, p)$

3. Geometric random variable: Suppose in a sequence of independent Bernoulli trials, let X denote the number of trials to observe the first success, then, $p(k) = \mathbb{P}(X = k) = (1-p)^{k-1} p^k$. X is said to be a geometric random variable. $X \sim \text{Geo}(k, p)$.

For a continuous random variable, we can define the probability density function $f(x)$ of a random variable, if it exists. The density function of a random variable needn't always exist, however for most common cases it exists.

To evaluate probability of a random variable X in an interval (a, b) with density $f(x)$, we perform,

$$\mathbb{P}(a \leq X \leq b) = \int_a^b f(x)dx$$

It also follows from the above that $F'(x) = f(x)$. (Verify!)

Some examples of continuous random variables,

1. Uniform random variable: We say that X is uniformly distributed over the interval (a, b) if,

$$f(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

   $X \sim \text{Unif}(a, b)$

2. Normal random variable: We say X is normally distributed over $\mathbb{R}$ if, $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$. $X \sim N(\mu, \sigma)$.

We define the expectation of a discrete random variable X by, $\mathbb{E}[X] = \sum xp(x)$. We define the expectation of a continuous random variable X by, $\mathbb{E}[X] = \int xf(x)dx$.

Exercises

1. Show that $\mathbb{E}[X] = np$, $X \sim \text{Bin}(n, p)$
2. Show that $\mathbb{E}[X] = \frac{b+a}{2}$, $X \sim \text{Unif}(a, b)$

Consider two random variables, X and Y, the joint CDF of X and Y is given by,

$$F(x, y) = \mathbb{P}(X \leq, Y \leq y)$$

If X, Y are discrete, then, $p(x, y) = \mathbb{P}(X = x, Y = y)$, here $p(x, y)$ is the joint probability mass function of X and Y.

If X, Y are continuous, then,

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

here $f(x, y)$ is the joint probability density function of X and Y if it exists.

Two random variables X and Y are said to be independent if, $\mathbb{P}(X \leq x, Y \leq y) = \mathbb{P}(X \leq x)\mathbb{P}(Y \leq y)$. Notice some consequence of this definition,

1. If X, Y are discrete, then $p(x, y) = p(x)p(y)$.
2. If X, Y are continuous, then $F(x, y) = F(x)F(y)$ and $f(x, y) = f(x)f(y)$.

References

1. Sheldon Ross, Introduction to Probability Models, Academic Press, 2024.