MA2202: Probability I

Random variables

Spring 2021

Satvik Saha
19MS154

*Indian Institute of Science Education and Research, Kolkata,
Mohanpur, West Bengal, 741246, India.*

Definition 3.1 (Random variable)**.** Given a probability space $(\Omega, \mathcal{E}, P)$, a function $X\colon \Omega \to \mathbb{R}$ is called a random variable if $X^{-1}(r, \infty) \in \mathcal{E}$ for all $r \in \mathbb{R}$.

Remark. For some $S \subseteq \mathbb{R}$, we denote

$$P(X \in S) = P(\{\omega \in \Omega \colon X(\omega) \in S\}).$$

Definition 3.2 (Discrete random variable)**.** A random variable which can assume only a countably infinite number of values is called a discrete random variable.

Example. Let $X\colon \Omega \to \mathbb{R}$ denote the number of heads obtained when a fair coin is tossed thrice. Note that $\Omega = \{0, 1, 2, 3\}$. Thus, $P(X = 0) = P(X = 4) = 1/8$ and $P(X = 1) = P(X = 2) = 3/8$.

Definition 3.3 (Probability distribution)**.** The probability distribution of a random variable X is the set of pairs $(X(A), P(A))$ for all $A \in \mathcal{E}$.

Definition 3.4 (Probability mass function)**.** Let X be a discrete random variable. The probability mass function of X is the function $p_X\colon \mathbb{R} \to [0, 1]$,

$$p_X(\alpha) = P(X = \alpha).$$

Remark. Since X is a discrete random variable, the set $S = \{x \in \mathbb{R}\colon p_X(\alpha) > 0\}$ is countable, and

$$\sum_{x \in S} p_X(x) = 1.$$

Definition 3.5 (Expectation)**.** The expectation of $g(X)$, for $g\colon \mathbb{R} \to \mathbb{R}$ and a discrete random variable X is defined as

$$E[g(X)] = \sum_{x \in S} g(x)\, p_X(x),$$

if the series converges absolutely.

Example. The n^{th} moment of a discrete random variable $E[X^n]$ is defined as

$$E[X^n] = \sum_{x \in S} x^n\, p_X(x),$$

if the series converges absolutely.

The first moment $\mu = E[X]$ is called the mean. The second moment $\sigma^2 = E[(X - \mu)^2]$ is called the variance. Note that

$$\sigma^2 = \sum (x - \mu)^2\, p(x) = \sum x^2\, p(x) - 2\mu x\, p(x) + \mu^2 p(x).$$

Simplifying,

$$\sigma^2 = E[X^2] - E[X]^2.$$

Definition 3.6 (Cumulative distribution function)**.** The cumulative distribution function of a random variable X is defined as the function $F_X\colon \mathbb{R} \to [0,1]$,

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

Definition 3.7 (Continuous random variable)**.** A continuous random variable X is such that its cumulative distribution function F_X is continuous.

Definition 3.8 (Probability density function)**.** Let X be a continuous random variable with a cumulative distribution function F_X. If we write

$$F_X(\alpha) = \int_{-\infty}^{\alpha} f_X(x)\, dx$$

for all $\alpha \in \mathbb{R}$, then f_X is a probability density function. If f_X is continuous, then the Fundamental Theorem of Calculus guarantees that $f_X(x) = F'_X(x)$.

Remark. Note that we can write

$$P(\alpha \leq X \leq \beta) = \int_{\alpha}^{\beta} f_X(x)\, dx.$$

We also demand

$$\int_{-\infty}^{+\infty} f_X(x)\, dx.$$

Example. The uniform distribution on an interval $[a, b] \subset \mathbb{R}$ is defined using the probability density function

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

Definition 3.9 (Expectation)**.** The expectation of $g(X)$, for $g \colon \mathbb{R} \to \mathbb{R}$ and a continuous random variable X is defined as

$$E[g(X)] = \int_{-\infty}^{+\infty} g(X)\, f_X(x)\, dx$$

if the integral converges absolutely.

Definition 3.10 (Mixed random variable)**.** A random variable whose cumulative distribution function F_X is discontinuous at countably many points, with F_X being continuous and strictly increasing in at least one interval is called a mixed random variable.

Definition 3.11 (Conditional probability distributions)**.** Let X and Y be two random variables, and let $A, B \subseteq \mathbb{R}$. If $P(Y \in B) > 0$, then

$$P(X \in A \,|\, Y \in B)\, P(Y \in B) = P(X \in A, Y \in B).$$

Definition 3.12 (Independent random variables)**.** We say that X and Y are independent if

$$P(X \in A, Y \in B) = P(X \in A)\, P(Y \in B)$$

for all $A, B \subseteq \mathbb{R}$.

Example. Consider an experiment where a fair die is rolled until a 6 is obtained, and let X be the random variable denoting the number of throws. Also, let Y be a random variable which is 1 if the first even outcome is a 6, and 0 otherwise. Observe that $P(X = 1) = 1/6$ and $P(Y = 1) = 1/3$. Also, $P(X = 1, Y = 1) = 1/6 \neq P(X = 1)\, P(Y = 1)$, hence X and Y are not independent random variables.

It can be shown that

$$P(X = n \,|\, Y = 1) = \frac{1}{2^n}.$$

Bernoulli distribution

Consider an experiment with one trial, which has two possible outcomes; the probability of a success is given by p and the probability of a failure is $q = 1 - p$. The discrete random variable such that $X = 1$ on success and $X = 0$ on failure is said to follow the Bernoulli distribution.

Binomial distribution

Consider an experiment with n Bernoulli trials. We could let X_i be a random variable denoting the outcome of the i^{th} trial, so we demand that $\{X_i\}$ are independent and identically distributed. The distribution of the sum $X = X_1 + \cdots + X_n$ follows the Binomial distribution $B(n, p)$, where

$$P(X = k) = \binom{n}{k} p^k q^{n-k}.$$

Geometric distribution

Consider an experiment where Bernoulli trials are repeated until success. The random variable denoting the number of trials required is said to follow the geometric distribution, where the probability that n trials were required is given by

$$P(X = n) = pq^{n-1}.$$

Pascal distribution

Consider an experiment where Bernoulli trials are repeated until k successes. The random variable denoting the number of trials required follows the Pascal, or negative binomial distribution. The probability that $n \geq k$ trials were required is given by

$$P(X = n) = \binom{n-1}{k-1} p^k q^{n-k}.$$

Note that we choose $k - 1$ successes from $n - 1$ trials, since the last trial is by definition a success.

Hypergeometric distribution

Consider an experiment where we choose n balls randomly from a population of N distinct balls, of which m are red. The random variable denoting the number of red balls follows the hypergeometric distribution. The probability of obtaining $k \leq m$ red balls is given by

$$P(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}.$$

Poisson distribution

The probability mass function of the Poisson distribution is given by

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda},$$

where $k \in \mathbb{N}$. This is generally used to model the number of times an event occurs in a given interval. This is a limiting form of the binomial distribution.

Exponential distribution

The probability density function of the exponential distribution is given by

$$f_X(x) = \lambda e^{-\lambda x},$$

where $x \geq 0$ and 0 elsewhere. This is generally used to model the waiting time between successive events.

Normal distribution

The probability density function of the normal distribution is given by

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}.$$