Conditional Probability

Probability that one event will happen *given that* another action or event happens.

$$P(X, given\ Y) \text{ or } P(X|Y)$$

This is to say that the chance of one event happening is conditional on another event happening.

For example, from a deck of cards, the probability that you get a six, given that you drew a red card is $P(6 | \text{red}) = 2/26 = 1/13$, since there are two sixes out of 26 red cards.

Formula for conditional probability

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

Law of Total Probability

Statement: Let S be the sample space and $E_1, E_2, \ldots, E_n$ be n mutually exclusive and exhaustive events with $P(E_i) \neq 0$; $i = 1, 2, \ldots, n$. Let A be any event which is a sub-set of $E_1 \cup E_2 \cup \ldots \cup E_n$ (i.e. at least one of the events $E_1, E_2, \ldots, E_n$) with $P(A) > 0$, then $P(A) = P(E_1)\, P(A|E_1) + P(E_2)\, P(A|E_2) + \ldots + P(E_n)\, P(A|E_n)$

Bayes Theorem

Prior Probability: It is the probability of an event occurring before new data is collected or before experiment is performed.

Example: Probability of occurrence of head in a coin toss - 0.5 (Given coin is unbiased)

Formula for Bayes Theorem

Determines the conditional probability of an event A given that event B has already occurred.

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

$P(A|B)$ is called posterior probability.

1. the prior probability $P(A)$;
2. the likelihood (or conditional probability) $P(B|A)$;
3. the marginal probability $P(B)$.

$$P(A|B) = P(A \cap B) / P(B)$$

Similarly, $P(B|A) = P(A \cap B) / P(A)$

It follows that $P(A \cap B) = P(A|B) * P(B) = P(B|A) * P(A)$

Thus, $P(A|B) = P(B|A)*P(A) / P(B)$

Generalization of Bayes Theorem

https://online.stat.psu.edu/stat414/lesson/6/6.2

Refer Basic Concepts in Probability
http://www.ignouhelp.in/ignou-mst-03-study-material/

Random Variable: A random variable is a real-valued function whose domain is a set of possible outcomes of a random experiment and range is a subset of the set of real numbers.

Note: We shall denote random variables by capital letters like X, Y, Z, etc. and write r.v. for random variable.

Discrete Random Variable

A random variable X is a discrete random variable if there are a finite number of possible outcomes of X or there are a countably infinite number of possible outcomes of X.

Recall that a countably infinite number of possible outcomes means that there is a one-to-one correspondence between the outcomes and the set of integers.

Homework Question:

Give Example of two discrete variables.

Probability Mass Function

Let X be a r.v. which takes the values $x_1, x_2, \ldots$ and let $P[X = x_i] = p(x_i)$. This function $p(x_i)$, $i = 1, 2, \ldots$ defined for the values $x_1, x_2, \ldots$ assumed by X is called probability mass function of X satisfying $p(x_i) \geq 0$ and $\sum_i p(x_i) = 1$.

The set $\{(x_1, p(x_1)), (x_2, p(x_2)), \ldots\}$ specifies the probability distribution of a discrete random variable X .

Example 1: State, giving reasons, which of the following are not probability distributions:

X	0	1	2
$p(x)$	$\frac{3}{4}$	$-\frac{1}{2}$	$\frac{3}{4}$

Example 2: For the following probability distribution of a discrete r.v. X, find the constant c.

X	0	1	2	3	4	5
$p(x)$	0	c	c	2c	3c	c

Example 3: Find the probability distribution of the number of heads when three fair coins are tossed simultaneously.

Let X be the number of heads in the toss of three fair coins.
As the random variable, "the number of heads" in a toss of three coins may be 0 or 1 or 2 or 3 associated with the sample space {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}, X can take the values 0, 1, 2, 3, with

P[X = 0] = P[TTT] = 1
P[X = 1] = P[HTT, THT, TTH] = 3
P[X = 2] = P[HHT, HTH, THH] = 3
P[X = 3] = P [HHH] = 1

Homework Question: Write down probability distribution for above solution.

Continuous Random Variable

A random variable is said to be continuous if it can take all possible real (i.e. integer as well as fractional) values between two certain limits.

Homework Question: Give example of 2 continuous random variables.

Note: continuous random variable is represented by different representation known as probability density function unlike the discrete random variable which is represented by probability mass function.

Probability Density Function

Let X be a continuous random variable and f(x) be a continuous function of x . Suppose $(x, x + \delta x)$ be an interval of length δx . Then f(x) defined by $\lim_{\delta x \to 0} \frac{P[x \le X \le x + \delta x]}{\delta x} = f(x)$.

is called probability density function of X.

Probability density function has the same properties as that of probability mass function i.e. $f(x) \geq 0$ and $\int_R f(x)\,dx = 1$, where integral has been taken over the entire range R of values of X.

$$P[a < X < b] = \int_a^b f(x)\,dx.$$

Example: The p.d.f. of the different weights of a “1 litre pure ghee pack” of a company is given by:

$$f(x) = \begin{cases} 200(x-1) & \text{for } 1 \leq x \leq 1.1 \\ 0, & \text{otherwise} \end{cases}$$

Examine whether the given p.d.f. is a valid one. If yes, find the probability that the weight of any pack will lie between 1.01 and 1.02.

DISTRIBUTION FUNCTION

A function F defined for all values of a random variable X by F(x) = P[X <= x] is called the distribution function. It is also known as the cumulative distribution function (c.d.f.) of X since it is the cumulative probability of X up to and including the value x.

Discrete Distribution Function

Distribution function of a discrete random variable is said to be discrete distribution function or cumulative distribution function (c.d.f.). Let X be a discrete random variable taking the values x1, x2, x3,... with respective probabilities p1, p2, p3,...

$$\text{Then } F(x_i) = P[X \le x_i] = P[X = x_1] + P[X = x_2] + \ldots + P[X = x_i]$$

$$= p_1 + p_2 + \ldots + p_i .$$

For example, Let X be a random variable having the following probability distribution:

X	0	1	2
p(x)	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

Determine the distribution function of X.

Continuous Distribution Function

Let X be a continuous random variable having the probability density function f(x), as defined in the last section of this unit, then the distribution function F(x) is given by

$$F(x) = P[X \le x] = \int_{-\infty}^{x} f(x)\,dx.$$

Homework Question

Example 8: The diameter 'X' of a cable is assumed to be a continuous random variable with p.d.f. $f(x) = \begin{cases} 6x(1-x), & 0 \le x \le 1 \\ 0, & \text{elsewhere} \end{cases}$.

Obtain the c.d.f. of X.

http://www.ignouhelp.in/ignou-mst-03-study-material/

List of Available MST-03 Study Materials:

MST-03 Probability Theory		
Block-1	Basic Concepts in Probability	Download
Block-2	Random Variables and Expectation	Download
Block-3	Discrete Probability Distributions	Download
Block-4	Continuous Probability Distributions	Download

Discrete Probability Distribution

BERNOULLI DISTRIBUTION

Suppose a piece of a product is tested which may be defective (failure) or non-defective (a success). Let p the probability that it found non-defective and q = 1 – p be the probability that it is defective. Let X be a random variable such that it takes value 1 when success occurs and 0 if failure occurs.

The above experiment is a Bernoulli trial, the r.v. X defined in the above experiment is a Bernoulli variate and the probability distribution of X as specified above is called the Bernoulli distribution.

Definition 3.4

A random variable X is said to be a *Bernoulli* random variable with *parameter* p, shown as $X \sim Bernoulli(p)$, if its PMF is given by

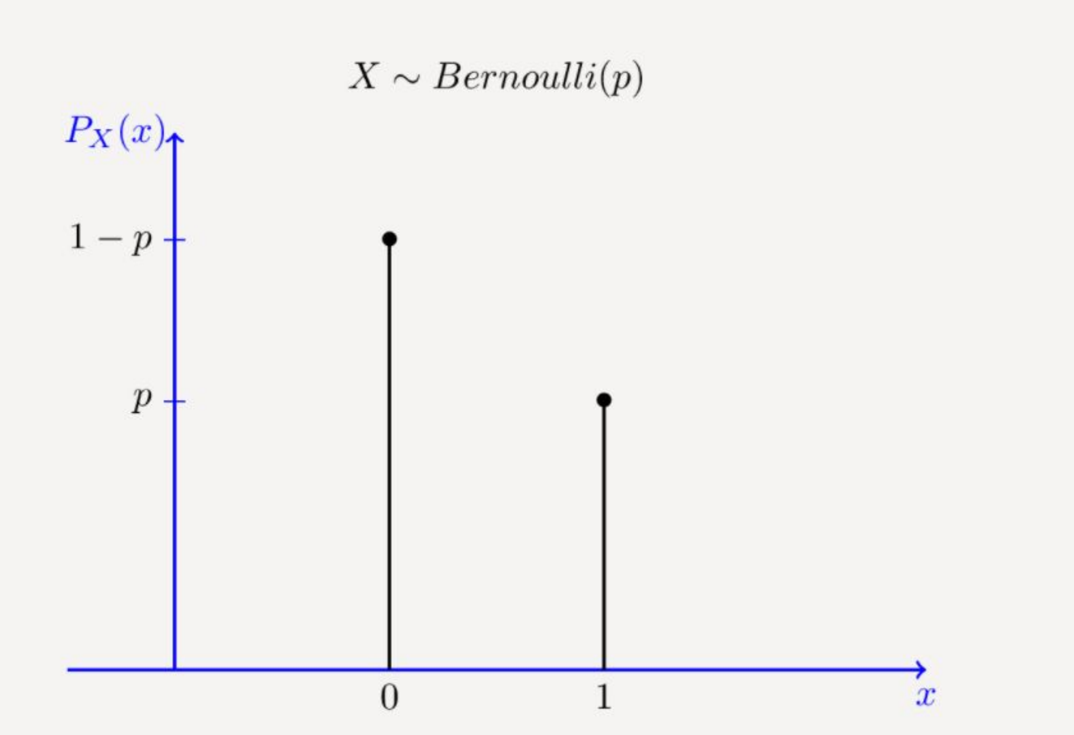
$$P_X(x) = \begin{cases} p & \text{for } x = 1 \\ 1 - p & \text{for } x = 0 \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$.

Remark 1: The Bernoulli distribution is useful whenever a random experiment has only two possible outcomes, which may be labelled as success and failure.

Examples:

1. Student will pass or fail an exam
2. Team will win a championship or not



Expected Value and Variance

The definitions of the expected value and the variance for a continuous variation are the same as those in the discrete case, except the summations are replaced by integrals.

$$\mu = E(X) = \int_{-\infty}^{\infty} x\, f(x)dx$$

$$\sigma^2 = E(\,(x-\mu)^2\,) = \int_{-\infty}^{\infty} (x-\mu)^2\, f(x)dx$$

Mean of Bernoulli = p

Variance of Bernoulli = p(1-p)

Binomial Distribution

The random experiment behind the binomial distribution is as follows. Suppose that I have a coin with $P(H) = p$. I toss the coin n times and define X to be the total number of heads that I observe. Then X is binomial with parameter n and p, and we write $X \sim Binomial(n, p)$. The range of X in this case is $R_X = \{0, 1, 2, \ldots, n\}$. As we have seen in Section 2.1.3, the PMF of X in this case is given by binomial formula

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, 1, 2, \ldots, n.$$

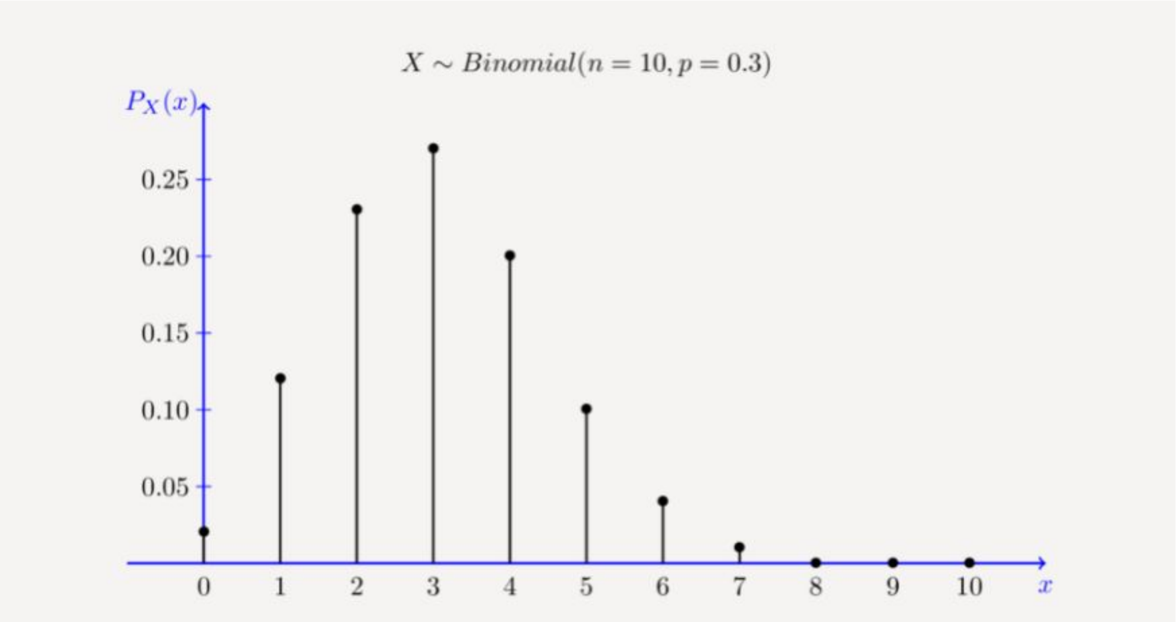
Definition

A random variable X is said to be a *binomial* random variable with parameters n and p, shown as $X \sim Binomial(n, p)$, if its PMF is given by

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, 2, \cdots, n \\ 0 & \text{otherwise} \end{cases}$$

where $0 < p < 1$.

The binomial distribution is the probability distribution of sum of n independent Bernoulli variates.



Mean = np

Variance = npq

Example: 80% of people who purchase pet insurance are women. If 9 pet insurance owners are randomly selected, find the probability that exactly 6 are women.

Example: 60% of people who purchase sports cars are men. If 10 sports car owners are randomly selected, find the probability that exactly 7 are men.

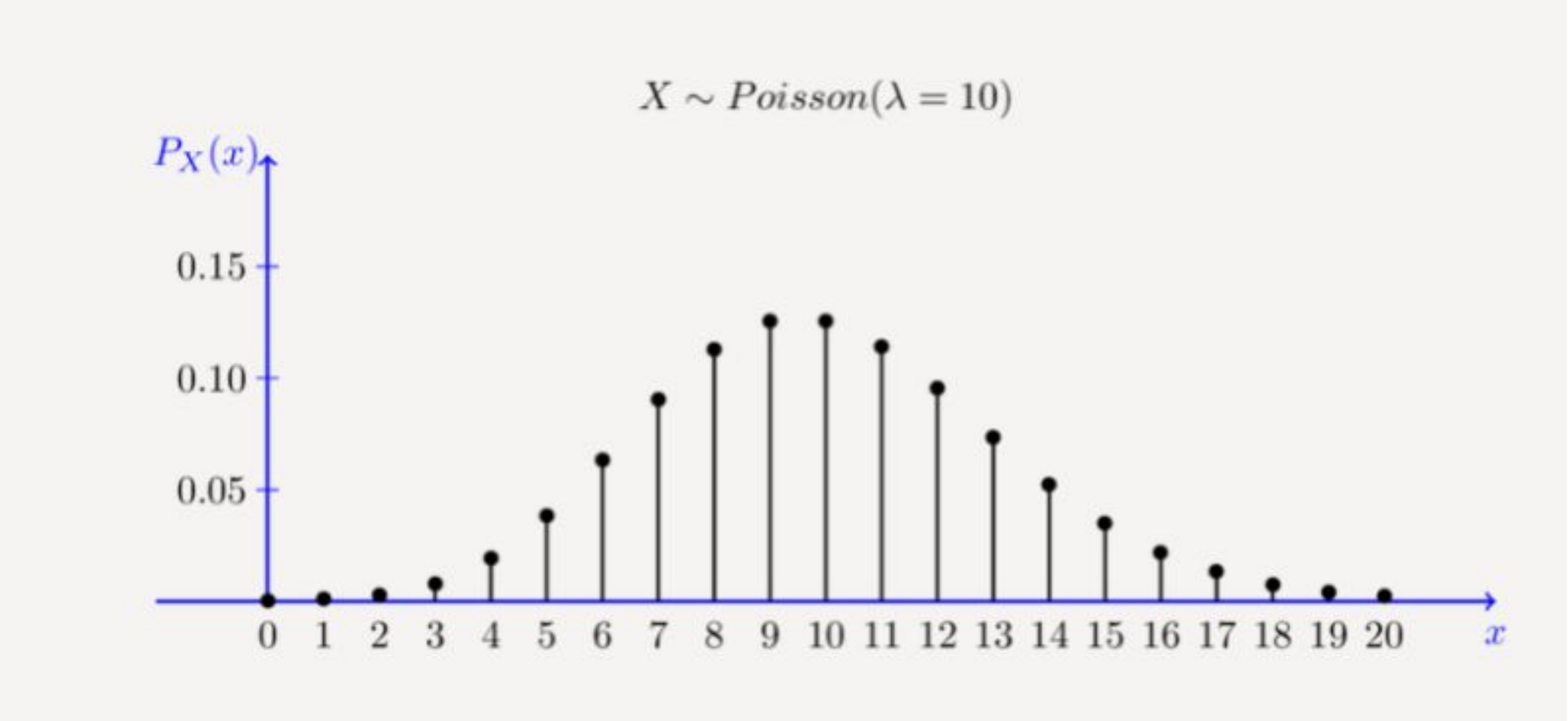
POISSON DISTRIBUTION

We have studied binomial distribution which is applied in the cases where the probability of success and that of failure do not differ much from each other and the number of trials in a random experiment is finite. However, there may be practical situations where the probability of success is very small, that is, there may be situations where the event occurs rarely and the number of trials may not be known. For instance, the number of accidents occurring at a particular spot on a road everyday is a rare event. For such rare events, we cannot apply the binomial distribution. To these situations, we apply Poisson distribution.

A random variable X is said to be a *Poisson* random variable with parameter λ, shown as $X \sim Poisson(\lambda)$, if its range is $R_X = \{0, 1, 2, 3, \ldots\}$, and its PMF is given by

$$P_X(k) = \begin{cases} \frac{e^{-\lambda}\lambda^k}{k!} & \text{for } k \in R_X \\ 0 & \text{otherwise} \end{cases}$$

Mean = Variance = λ


$X \sim Poisson(\lambda = 10)$
$P_X(x)$
0.15
0.10
0.05
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
x

Poisson as an approximation for binomial

The Poisson distribution can be viewed as the limit of binomial distribution. Suppose $X \sim Binomial(n,p)$ where n is very large and p is very small. In particular, assume that $\lambda = np$ is a positive constant. We show that the PMF of X can be approximated by the PMF of a $Poisson(\lambda)$ random variable. The importance of this is that Poisson PMF is much easier to compute than the binomial. Let us state this as a theorem.

Theorem 3.1

Let $X \sim Binomial(n, p = \frac{\lambda}{n})$, where $\lambda > 0$ is fixed. Then for any $k \in \{0, 1, 2, \ldots\}$, we have

$$\lim_{n \to \infty} P_X(k) = \frac{e^{-\lambda}\lambda^k}{k!}.$$

https://www.probabilitycourse.com/chapter3/3_1_5_special_discrete_distr.php

Refer for proof

Homework Question: Proof of Theorem

Examples

1. Calls per Hour at a Call Center
2. Number of Arrivals at a Restaurant
3. Number of Website Visitors per Hour
4. Number of Network Failures per Week

HomeWork Question

Example 1: It is known that the number of heavy trucks arriving at a railway station follows the Poisson distribution. If the average number of truck arrivals during a specified period of an hour is 2, find the probabilities that during a given hour
a) no heavy truck arrive,
b) at least two trucks will arrive.

http://www.ignouhelp.in/ignou-mst-03-study-material/

Refer: Discrete Probability Distribution Unit 10

Discrete Uniform Probability Distribution

Discrete uniform probability function is given by

$$f(x) = \frac{1}{n}$$

Where n is the number of values the random variable may take

Example: Probability distribution of events while rolling a dice

x	f(x)
1	1/6
2	1/6
3	1/6
4	1/6
5	1/6
6	1/6