Some important distributions

A. Discrete distributions

1. Bernoulli distribution
2. Binomial distribution
3. Poisson distribution

B. Continuous distributions

1. Uniform distribution
2. Normal distribution
3. Exponential distribution
4. Gamma distribution

Bernoulli distribution

Bernoulli trial

A trial that has two possible outcomes is known as a Bernoulli trial. The two outcomes are called 'success' (probability p) and 'failure' (probability $1-p$).

Examples

1. A fair coin is tossed. Here, head (or tail) is a success. $P(H) = p = 0.5$.

2. An unfair coin is tossed. $P(H) = p = 0.7$ (say).

3. A fair dice is tossed. $P(6) = p = 1/6$.

4. A patient is tested for COVID-19. $P(+) = p = 0.05$ (say).

Bernoulli random variable

Let a Bernoulli trial be performed <u>once</u>. Then, $S = \{s,\ f\}$. Let X be the number of successes obtained. Then, X is a Bernoulli random variable with possible values '0' and '1'. It has the following pmf:

x	0	1
$P(X = x)$	$1-p$	p

We can write the above pmf as follows:

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

Note

Here, p is the 'parameter' of the probability distribution. It characterizes the probability distribution (population). We write,

$X \sim \text{Bernoulli}(p)$.

That is, X follows Bernoulli distribution with parameter p.

Mean and variance

$E(X) = 0 \times (1 - p) + 1 \times p = p$

$E(X^2) = 0^2 \times (1 - p) + 1^2 \times p = p$

$V(X) = p - p^2 = p(1 - p)$

Note that, mean $>$ variance.

Also, $V(X) \leq 0.25$. (Why?)

Example

In a multiple-choice question (MCQ), there are four choices. A student does not know the correct answer and selects one of the choices at random. Let X be the number of correct answers (out of 1 question). Then

$X \sim \text{Bernoulli}(p = 0.25)$.

Binomial distribution

Binomial experiment

A binomial experiment consists of n independent Bernoulli trials, where n is a positive integer. The probability of success (p) is same in each trial.

Example

1. A fair coin is tossed 5 times. Here, $n = 5$ and $p = 0.5$.

2. A coin (probability of head 0.7) is tossed 10 times. Here, $n = 10$ and $p = 0.7$.

Binomial random variable

Let n independent Bernoulli trials be performed, each having probability of success p. Let X be the number of successes obtained. Then, X is a binomial random variable with possible values $0, 1, 2, \cdots, n$. It has the following pmf:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \,; x = 0, 1, 2, \cdots, n.$$

Note

The distribution has two parameters n and p. We write

$X \sim \text{binomial}(n,\ p)$.

That is, X follows binomial distribution with parameters n and p.

Explanation of the pmf

Let 5 independent Bernoulli trials be performed, each having probability of success p. The sample space is as follows:

$$S = \{sssss, ssssf, sssfs, sssff, \cdots, sffss, \cdots, ffsss, \cdots, fffff\}$$

Since the trials are independent, we have

$$P(sssss) = p \times p \times p \times p \times p = p^5$$
$$\vdots$$
$$P(sssff) = p \times p \times p \times (1-p) \times (1-p) = p^3(1-p)^2$$
$$\vdots$$
$$P(sffss) = p \times (1-p) \times (1-p) \times p \times p = p^3(1-p)^2$$
$$\vdots$$
$$P(fffff) = (1-p) \times \cdots \times (1-p) = (1-p)^5$$

Probability that we will get 3 successes:

$$\begin{aligned} P(X = 3) &= P(sssff \text{ or } ssffs \text{ or } \cdots \text{or } ffsss) \\ &= p^3(1-p)^2 + p^3(1-p)^2 + \cdots + p^3(1-p)^2 \\ &= \binom{5}{3} p^3(1-p)^2 \end{aligned}$$

Note

$(p + (1 - p))^n = 1$

If we expand the left-hand-side of the above equation (binomial expansion), we get the sum of all the probabilities of binomial distribution, which equals one.

Note

Let $X \sim \text{binomial}(n,\ p)$. Then

$X = Y_1 + Y_2 + \cdots + Y_n$

where $Y_i =$ number of success in the ith trial, $i = 1, 2, \cdots, n$. That is, $Y_i \sim \text{Bernoulli}(p)$. Thus, a $\text{binomial}(n,\ p)$ random variable is the sum of n independent $\text{Bernoulli}(p)$ random variables.

Mean and variance

$E(X) = np$

$V(X) = np(1 - p)$

Variance is smaller than the mean.

Exercise

In an MCQ test, there are 10 questions each with 4 choices. Smith chooses all the answers at random. What is the probability that the number of correct answers will be (i) exactly 2 (ii) less than 2 (iii) 2 or less?

Solution

Let $X =$ number of correct answers. Then

$X \sim \text{binomial}(n = 10,\ p = 0.25)$.

(i) $P(X = 2) = P(2) = \binom{10}{2} 0.25^2 (1 - 0.25)^{10-2} = 0.2816$

(ii) $P(X < 2) = P(0) + P(1) = 0.2440$

(iii) $P(X \leq 2) = P(0) + P(1) + P(2) = 0.5256$

Exercise

In an MCQ test, there are 10 questions each with 4 choices. Neither Smith nor Jones studied for the test. They independently chose all the answers at random. What is the probability that their total number of correct answers will be 5?

Solution

Let X_1 = number of correct answers by Smith.

and X_2 = number of correct answers by Jones.

Here, X_1 and X_2 are independent.

Let $Y = X_1 + X_2$. Then

$Y \sim$ binomial$(n = 20,\ p = 0.25)$.

$$P(Y = 5) = P(5) = \binom{20}{5} 0.25^5 (1 - 0.25)^{20-5} = 0.2023$$

Note

The above problem is an example of 'additive property' of binomial distribution. If $X_1 \sim$binomial$(n_1,\ p)$ and $X_2 \sim$binomial$(n_2,\ p)$, and if X_1 and X_2 are independent, then

$Y = X_1 + X_2 \sim$ binomial$(n_1 + n_2,\ p)$.

Poisson distribution

A Poisson distribution is appropriate for 'count data' which are obtained by counting the number of occurrences in an interval of time or in an area.

Examples

1. Number of accidents at a particular spot in a week
   $(X = 0, 1, 2, \cdots)$
2. Number of pens required in a month
   $(X = 0, 1, 2, \cdots)$
3. Number of jackfruit trees in every square kilometer
   $(X = 0, 1, 2, \cdots)$

Poisson pmf

A Poisson random variable X has the following pmf:

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \;;\;\; x = 0, 1, 2, \cdots; \;\; \lambda > 0.$$

Here, λ is the parameter of the distribution. We write

$X \sim \text{Poisson}(\lambda)$.

Mean and variance

$E(X) = \lambda$

$V(X) = \lambda$

Note that mean = variance.

Exercise

The average number of accidents at a particular spot in a week is 3.5. What is the probability that the number of accidents in the next week will be (i) zero (ii) exactly one (iii) more than one?

Solution

Let $X =$ number of accidents in the next week. Then

$X \sim \text{Poisson}(\lambda = 3.5)$.

(i) $P(X = 0) = P(0) = \dfrac{e^{-3.5}3.5^{0}}{0!} = 0.0302$

(ii) $P(X = 1) = P(1) = 0.1057$

(iii) $P(X > 1) = 1 - P(0) - P(1) = 0.8641$

Exercise

On an average, 3.5 customers enter a shop between 10:00-11:00 and 2.5 customers enter the shop between 11:00-12:00. What is the probability that exactly 5 customers will enter the shop between 10:00-12:00 tomorrow?

Solution

Let X_1 = number of customers between 10:00-11:00.

and X_2 = number of customers between 11:00-12:00.

Here, X_1 and X_2 are independent.

Let $Y = X_1 + X_2$. Then

$Y \sim \text{Poisson}(\lambda = 3.5 + 2.5 = 6)$.

$$P(Y = 5) = P(5) = \frac{e^{-6}6^5}{5!} = 0.1606$$

Note

The above problem is an example of 'additive property' of Poisson distribution. If $X_1 \sim \text{Poisson}(\lambda_1)$ and $X_2 \sim \text{Poisson}(\lambda_2)$, and if X_1 and X_2 are independent, then

$Y = X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Binomial-Poisson relationship

Theorem: Let $X \sim \text{binomial}(n,\ p)$. If $n \to \infty$ and $p \to 0$ so that np is finite, then

$X \sim \text{Poisson}(\lambda = np)$, approximately.

Explanation

Let X be a binomial variable. If n is very large, p is very small and np is neither large nor small, then X follows Poisson distribution approximately. In that case, we can use either binomial or Poisson distribution to solve the problem.

Exercise

Batteries of a particular brand malfunctions with probability 0.001. Calculate the probability that at least one out of 1000 batteries will malfunction.

Solution

Method 1

Let X = number of batteries that will malfunction. Then

$X \sim$ binomial$(n = 1000,\ p = 0.001)$.

$P(X \geq 1) = 1 - P(0) = 0.6323$

Method 2

Let $X =$ number of batteries that will malfunction. Then

$X \sim$ binomial$(n = 1000,\ p = 0.001)$.

Here, n is large and p is small. Also, $np = 1$ (neither large nor small).

$X \sim$ Poisson$(\lambda = np = 1)$, approximately.

$P(X \geq 1) = 1 - P(0) = 0.6321$