# BINOMIAL DISTRIBUTION

# BINOMIAL DISTRIBUTION

- This was discovered by James Bernoulli's in 1700 and it expresses probabilities of events of dichotomous (dicho + tomy = Two parts) nature i.e, which results in only two ways, success or failure.

# DEFINITION

- A random variable is said to follow the Binomial distribution if the probability of $x$ is given by

$$P(X = x) = {}^{n}C_{x} p^{x} q^{n-x},$$

$$x = 0,1,2,3 \ldots n \ and \ q = 1 - p$$

The two constants $n$ and $p$ are called the parameters of the distribution.

# REMARKS:

- The distribution is called "Binomial Distribution" because the probabilities

$$^{n}C_{x}p^{x}q^{n-x}, x = 01,2, \dots, n$$

are the successive terms of the expansion of the binomial expression $(q + p)^{n}$

- If $X$ is a binomial variate with parameters $n \; and \; p$, it is denoted as $b(X, n, p)$

- The sum of the probabilities is 1.

$$\sum_{x=0}^{n} p(x) = \sum_{x=0}^{n} {}^{n}C_x p^x q^{n-x}$$
$$= (p+q)^n = 1$$

- Let the experiment of n trials be repeated N times. Then we except x successes to occur

$$N.\left({}^{n}C_x p^x q^{n-x}\right) \text{ times.}$$

This is called **frequency function.**

# WHEN DO WE GET BINOMIAL DISTRIBUTION

- We get a binomial distribution when the following conditions are satisfied:

**(i)** A trial is repeated n times where n is a finite number.

**(ii)** Each trial results only in two ways success or failure.

(iii) These possibilities are mutually exclusive, exhaustive but not necessarily equally likely.

**(iv)** If p and q are the probabilities of success and failure then $p + q = 1$

**(v)** The events are independent, i.e the probability p of success in each trial remains constant in all trials.

# USES

Binomial distribution is used in problems involving

**(i)** The tossing of a coin – heads or tails,

**(ii)** The results of an examination success or failure,

**(iii)** The result of an election – success or failure,

**(iv)** The results of inspection of an article defective or non defective,

**(v)** Habit of a person smoker or non smoker etc.

# MEAN , VARIANCE AND MODE

- $Mean = np$
- $variance = npq$
- If $(n + 1)p$ is an integer, say, k then there are two modes $k$ $and$ $k - 1$.
- If $(n + 1)p$ is not an integer then the mode is the integral part of $(n + 1)p$.

# ADDITIVE PROPERTY OF BINOMIAL DISTRIBUTION

- If $X_1$ is a Binomial variate with parameter $n_1$ $and$ $p_1$ and $X_2$ is another Binomial variate with parameter $n_2$ $and$ $p_2$ then $X_1 + X_2$ in general is not a Binomial variate.
- If $X_1$ $and$ $X_2$ are two Binomial variates with parameters $n_1, p$ and $n_2, p$ then $X_1 + X_2$ is a Binomial variate with parameters

$(n_1 + n_2), p$

# RECURRENCE RELATION FOR THE PROBABILITY OF BINOMIAL DISTRIBUTION

- $$\frac{p(x+1)}{p(x)} = \frac{{}^{n}C_{x+1}p^{x+1}q^{n-x-1}}{{}^{n}C_{x}p^{x}q^{n-x}}$$

$$= \frac{n(n-1).....(n-x)}{(x+1)!} \cdot \frac{x!}{n(n-1)....(n-x+1)} \cdot \frac{p}{q}$$

$$= \frac{n-x}{x+1} \cdot \frac{p}{q}$$

$$\therefore p(x+1) = \frac{n-x}{x+1} \cdot \frac{p}{q} p(x)$$

- Further since the expected frequency of x

  i.e $f(x) = Np(x)$,

  we have $f(x+1) = N.p(x+1)$

$$= N.\left(\frac{n-x}{x+1} \cdot \frac{p}{q} p(x)\right)$$

$$= \left(\frac{n-x}{x+1} . \frac{p}{q}\right) . N.p(x)$$

$$\therefore f(x+1) = \frac{n-x}{x+1} . \frac{p}{q} . f(x)$$

**EX 1.** The mean and variance of a Binomial variate are 3 and 1.2 . Find '$n$', '$p$' and $\mathrm{P}(X < 4)$ .

ANS: given $np = 3$ and $npq = 1.2$

$\therefore q = \frac{npq}{np} = \frac{1.2}{3} = 0.4$

$\therefore p = 1 - q = 0.6$

$np = 3 \Rightarrow n \times 0.6 = 3$

$\therefore n = 5$

$P(X = x) = {}^5C_x 0.6^x 0.4^{5-x}, \; x = 0,1,2,3 \ldots 5$

$$\mathrm{P}(X < 4) = 1 - P(X \geq 4) = 1 - P(X = 4) - P(X = 5)$$

$$= 1 - {}^5C_4 0.6^4 0.4^1 - {}^5C_5 0.6^5 0.4^0$$

$= \; 0.66304$

**EX 2.** Find the Binomial distribution if the mean is 4 and variance is 3. Find also its mode.

**Solution:**

We have mean $= np = 4$ and variance $= npq = 4$

$$\therefore \frac{np}{npq} = \frac{4}{3} \qquad \therefore \frac{1}{q} = \frac{4}{3} \qquad \therefore q = \frac{3}{4}$$

$$\therefore p = 1 - q = 1 - \frac{3}{4} = \frac{1}{4}$$

$$\therefore np = 4 \qquad \therefore n \cdot \frac{1}{4} = 4 \qquad \therefore n = 16$$

$$\therefore P(X = x) = {}^{n}C_x \, p^x \, q^{n-x} = {}^{16}C_x \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{16-x}$$

$$(n+1)p = (16+1) \cdot \frac{1}{4} = \frac{17}{4}$$

$\therefore$ Mode $=$ integral part of $\frac{17}{4} = 4$

**EX 3.** A Binomial variate X satisfies the relation $9P(X=4)=P(X=2)$ when $n=6$. Find the value of the parameter $'p'$.

**Solution:**

We have $P(X=x)={}^{n}C_{x}\,p^{x}\,q^{n-x}={}^{6}C_{x}\,p^{x}\,q^{6-x}$

Since, $9\,P(X=4)=\ P(X=2)$

$\therefore 9\,{}^{6}C_{4}\,p^{4}\,q^{6-4}={}^{6}C_{2}\,p^{2}\,q^{6-2}$

$\therefore 9\ {}^{6}C_{4}\,p^{4}\,q^{2}={}^{6}C_{2}\,p^{2}\,q^{4}$

$\because {}^{6}C_{4}={}^{6}C_{2},$

$9p^{2}=\ q^{2}$ $\quad\quad \therefore 9\,p^{2}=(1-p)^{2}$

$\therefore 9p^{2}=1-2p+p^{2}$ $\quad\quad \therefore 8p^{2}+2p-1=0$

$\therefore (4p-1)(2p+1)=0$ $\quad\quad \therefore p=-\frac{1}{2}$ or $\frac{1}{4}$

Since $p$ cannot be negative, $p=1/4$

**EX 4.** If 10 fair coins are tossed simultaneously, what is the chance of getting atleast 7 heads?

**Solution:** Probability of $x$ successes in a trial $= {}^{n}C_x\, p^x\, q^{n-x}$

Here $n = 10, p = \frac{1}{2}, q = \frac{1}{2}$

**(i)** $x = 7, P(7 \text{ heads}) = {}^{10}C_7 \left(\frac{1}{2}\right)^{10-7} \left(\frac{1}{2}\right)^{7} = 120 \times \left(\frac{1}{2}\right)^{10}$

**(ii)** $x = 8, P(8 \text{ heads}) = {}^{10}C_8 \left(\frac{1}{2}\right)^{10-8} \left(\frac{1}{2}\right)^{8} = 45 \times \left(\frac{1}{2}\right)^{10}$

**(iii)** $x = 9, P(9 \text{ heads}) = {}^{10}C_9 \left(\frac{1}{2}\right)^{10-9} \left(\frac{1}{2}\right)^{9} = 10 \times \left(\frac{1}{2}\right)^{10}$

**(iv)** $x = 10, P(10 \text{ heads}) = {}^{10}C_{10} \left(\frac{1}{2}\right)^{10} \left(\frac{1}{2}\right)^{0} = 1 \times \left(\frac{1}{2}\right)^{10}$

$$P(X \geq 7) = P(X = 7) + P(X = 8) + P(X = 9) + P(X = 10)$$

$$= [120 + 45 + 10 + 1]\frac{1}{2^{10}} = \frac{176}{1024} = \frac{11}{64}$$

**EX 5.** The odds in favour of $X$'$s$ winning a game against $Y$ are 4:3. Find probability of $Y$'$s$ winning 3 games out of 7 played.

**ANS:** 0.293

**EX 6.** If X is the random variable showing the number of boys in a family with 4 children, construct a table showing the probability distribution of $X$

**EX 7.** In a multiple choice examination there are 20 questions. Each question has 4 alternative answers following it-and the student must select one correct answer. 4 marks are given for correct answer and 1 mark is deducted for wrong answer. A student must secure at least 50% of maximum possible marks to pass the examination. Suppose a student has not studied at all, so that he answers the questions by guessing only. What is the probability that he will pass the examination?

**Solution:** Since there are 20 questions and each carries 4 marks, maximum marks are 80.
If the student solves 12 questions correctly and 8 questions wrongly, he gets more than 40 marks
Now, probability of getting a correct answer is $p = 1/4$, of wrong answer $q = 3/4$

$$P(X = x) = {}^{n}C_{x}\, p^{x}\, q^{n-x}$$

$$n = 20, p = 1/4, q = 3/4$$

$\therefore$ Probability of passing $= P(X = 12, 13, \ldots., 20)$

$$= \sum_{x=12}^{20} {}^{20}C_{x} \left(\frac{1}{4}\right)^{x} \left(\frac{3}{4}\right)^{20-x}$$

**EX 8.** In a Binomial distribution consisting of 5 independent trials, probability of 1 and 2 successes are $0.4096 \quad and \; 0.2048$ respectively. Find the parameter '$p$' of the distribution.

**HINT : similar to EX 3**

We have $P(X = x) = {}^{n}C_x \, p^x \, q^{n-x} = {}^{5}C_x \, p^x \, q^{5-x}$

Since, $P(X = 1) = 2P(X = 2)$

Solve further….

**EX 9.** If hens of a certain breed lay eggs on 5 days a week on an average; find on how many days during a season of 100 days, a poultry keeper with 5 hens of this breed, will expect to receive at least 4 eggs?

**Solution:** Probability of an hen laying an egg, $p = 5/7$ and probability of not laying an egg, $q = 1 - p = 2/7$

$$P(X = x) = {}^{n}C_{x}\, p^{x}\, q^{n-x} \text{ and } n = 5, p = \frac{5}{7}, q = \frac{2}{7}$$

$$\therefore P(X \geq 4) = P(X = 4) + P(X = 5)$$

$$= {}^{5}C_{4}\left(\frac{5}{7}\right)^{4}\left(\frac{2}{7}\right)^{1} + {}^{5}C_{5}\left(\frac{5}{7}\right)^{5}\left(\frac{2}{7}\right)^{0}$$

$$= 0.5578$$

$$\text{Expectation} = Np = 100 \times 0.5578 = 55.78 = 56$$

**EX 10.** A communication system consists of n components, each of which functions independently with probability $p$. The total system will be able to function effectively if at least one – half of its components are functioning. For what value of p is a 5 – component system more likely to operate effectively than a 3 – component system ?

**Solution:** Here, we have a Binomial distribution with parameters $n$ and $p$

$$\therefore P(X = x) = {}^{n}C_{x}\, p^{x}\, q^{n-x}, x = 0, 1, 2, \dots \dots \dots, n$$

$$P(\text{5 component system will work effectively})$$
$$= P(X = 3, \text{or } 4, \text{or } 5)$$
$$= P(X = 3) + P(X = 4) + P(X = 5)$$
$$= \sum_{x=3}^{5} {}^{3}C_{x}\, p^{x}\, q^{5-x} \qquad (\because n = 5)$$

$$P(\text{3 component system will work effectively}) = P(X = 2 \text{ or } 3)$$
$$= \sum_{x=2}^{3} {}^{3}C_{x}\, p^{x}\, q^{3-x} \qquad (\because n = 3)$$

5 −component system will work more effectively than 3 −component system if

$$\sum_{x=3}^{5} {}^{3}C_x \, p^x \, q^{5-x} \geq \sum_{x=2}^{3} {}^{3}C_x \, p^x \, q^{3-x}$$

$$\therefore ({}^{5}C_3 \, p^3 \, q^2 + {}^{5}C_4 \, p^4 \, q + 5c_5 p^5) \geq ({}^{3}C_2 \, p^2 \, q + {}^{3}C_3 \, p^3)$$

$$\therefore (10 \, p^3(1-p)^2 + 5 \, p^4(1-p) + \ p^5) \geq$$
$$1(3 \, p^2(1-p) + \ p^3)$$

$$\therefore 10p^3 - 20p^4 + 10p^5 + 5p^4 - 5p^5 + p^5 - 3p^2 + 3p^3 -$$
$$p^3 \geq 0$$

$$6p^5 - 15p^4 + 12p^3 - 3p^2 \geq 0$$

$$\therefore 3p^2(2p^3 - 5p^2 + 4p - 1) \geq 0$$

$$3p^2(p-1)^2(2p-1) \geq 0$$

$$\therefore 2p - 1 \geq 0 \qquad [\because p^2 \geq 0, (p-1)^2 \geq 0]$$

$\therefore p \geq \frac{1}{2}$ is the required value

**EX 11.** An irregular six faced die is thrown and the probability that in 20 throws it will give 5 even numbers is twice the probability that it will give 5 odd numbers. How many times in 10,000 sets of 10 throws would you expect it to give no even number?

**Solution:** Let the probability of getting an even number in a single throw by $p$ then, the probability of getting an odd number is $q = 1 - p$

Now, the probability of getting 5 even numbers in 20 throws $= {}^{20}C_5 q^{15}\, p^5$

And the probability of getting 5 odd numbers in 20 throws $= {}^{20}C_5 q^5\, p^{15}$

By data ${}^{20}C_5 p^5\, q^{15} = 2\,{}^{20}C_5 q^5\, p^{15}$

$\therefore \frac{q^{15}}{q^5} = \frac{2p^{15}}{p^2}$ $\quad\quad \therefore q^{10} = 2p^{10}$ $\quad\quad \therefore \left(\frac{q}{p}\right)^{10} = 2$

$$\therefore \left[\frac{q}{p}\right] = 2^{1/10} \qquad \therefore \log\left[\frac{q}{p}\right] = \frac{1}{10}\log 2$$

$$\therefore \log\left[\frac{q}{p}\right] = \frac{1}{10}\log(0.3010) = 0.0301$$

$$\therefore \frac{q}{p} = \text{antilog}\, 0.0301 = 1.072$$

$$\therefore q = p(1.072) = (1-q)(1.072) = 1.072 - 1.072q$$

$$\therefore 2.072q = 1.072 \qquad \therefore q = \frac{1.072}{2.072} = 0.517$$

$\therefore$ Probability of getting no even number in 10 throws $= {}^{10}C_0 q^{10} = q^{10} = (0.517)^{10} = 0.001365$

The expected number $= Np = 10{,}000(0.001365) = 13.65$

**EX 12.** Out of 800 families with 5 children each how many would you expect to have

**(i)** 3 Boys and 2 Girls, **(ii)** 5 girls, **(iii)** 5 boys?

**Solution:** Here $P(\text{Boy}) = p = \frac{1}{2}, P(\text{Girl}) = q = \frac{1}{2}, n = 5,$

(i)$P(3\text{ boys and }2\text{ girl}) = {}^5C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{5}{16}$

$\therefore$ Expected number of families $= Np = 800 \times \frac{5}{16} = 250$

(ii)$P(5\text{ girls}) = {}^5C_0 \left(\frac{1}{2}\right)^5 = \frac{1}{32}$

$\therefore$ Expected number of families $= Np = 800 \times \frac{1}{32} = 25$

(ii)$P(5\text{ boys }) = {}^5C_5 \left(\frac{1}{2}\right)^5 = \frac{1}{32}$

$\therefore$ Expected number of families $= Np = 800 \times \frac{1}{32} = 25$

**EX 13.** Let X, Y be two independent binomial variates with parameters $(n_1 = 6, p = 1/2)$ and $(n_2 = 4, p = 1/2)$ respectively.

Find $P(X + Y = 3)$ , $P(X + Y \geq 3)$

**Solution:** By the additive property of Binomial variates $Z = X + Y$ is a Binomial variate with parameters

$n = n_1 + n_2 = 6 + 4 = 10$ and $p = 1/2$

$$\therefore P(Z) = {}^{n}C_z\, p^z \cdot q^{n-z}$$

$$\therefore P(Z = 3) = {}^{10}C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^7 = \frac{15}{128} = 0.1172$$

$$P(Z \geq 3) = 1 - [P(Z = 0) + P(Z = 1) + P(Z = 2)]$$

$$= 1 - \left[{}^{10}C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^{10} + {}^{10}C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^9 + {}^{10}C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^8\right]$$

$$= 1 - \left[({}^{10}C_0 + {}^{10}C_1 + {}^{10}C_2) \left(\frac{1}{2}\right)^{10}\right]$$

$$= 0.945$$

**EX 14.** A lot contains 1% defective items. What should be the number of items in a lot so that the probability of finding at least one defective item in it, is at least 0.95?

**EX 15.** Five fair coins are tossed 3200 times, find the frequency distribution of number of heads obtained. Also find mean and standard deviation.

**Solution:** We have $p = \frac{1}{2}, q = \frac{1}{2}, n = 5$

$\therefore P(X = x) = {}^{n}C_x \, p^x \, q^{n-x} = {}^{5}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x}$

Putting $x = 0, 1, 2, 3,4,5$ we get

$P(X = 0) = {}^{5}C_0 \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^5 = \frac{1}{32}$

$P(X = 1) = {}^{5}C_1 \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^4 = \frac{5}{32}$

$P(X = 2) = {}^{5}C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^3 = \frac{10}{32}$

$P(X = 3) = {}^{5}C_3 \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 = \frac{10}{32}$

$P(X = 4) = {}^{5}C_4 \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^1 = \frac{5}{32}$

$P(X = 5) = {}^{5}C_5 \left(\frac{1}{2}\right)^5 \left(\frac{1}{2}\right)^0 = \frac{1}{32}$

The corresponding frequencies are obtained by multiplying these probabilities by 3000

| No of Heads | 0 | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|---|
| Frequency | 100 | 500 | 1000 | 1000 | 500 | 100 |

Now the mean of the binomial distribution $= np = 5 \times \frac{1}{2} = 2.5$

The standard deviation of the Binomial distribution $= \sqrt{npq} = \sqrt{5 \times \frac{1}{2} \times \frac{1}{2}} = \frac{\sqrt{5}}{2}$

Ex 16 Seven coins are tossed and the number of heads obtained is noted. The experiment is repeated 128 times and the following distribution

| No. of heads | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | Total |
|---|---|---|---|---|---|---|---|---|---|
| Frequency | 7 | 6 | 19 | 35 | 30 | 23 | 7 | 1 | 128 |

Fit a Binomial distribution if **(i)** the coins are unbiased, **(ii)** if the nature of the coins is not known.

**Solution:** To fit a distribution to given data means to find the constants of the distribution which will adequately describe the given situation

**(i) When the coins are unbiased**

$p = \frac{1}{2}, q = \frac{1}{2}$ and by data $n = 7$

$\therefore P(X = x) = {}^{7}C_x \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{7-x}$

Putting $x = 0, 1, 2, 3, \dots\dots, 7$, we get,

$P(0) = \frac{1}{2^7}, \quad P(1) = \frac{7}{2^7}, \quad P(2) = \frac{21}{2^7}, \dots\dots\dots.$

Expected frequency $= Np$ and $N = 128$

Multiplying the above probabilities by 128 i.e. by $2^7$ we get the expected frequencies as

$1, 7, 21, 35, 35, 21, 7, 1$

## (ii) When the nature of the coins is not known

We have $\bar{x} = \frac{\sum f_i x_i}{N} = \frac{0\times7+1\times6+2\times19+\cdots+7\times1}{128} = \frac{433}{128} = 3.38$

But $\bar{x} = np$,

$\therefore p = \frac{\bar{x}}{n} = \frac{3.38}{7} = 0.48$ $\quad\quad\quad\quad$ $\therefore q = 1 - p = 0.52$

$\therefore P(X = x) = {}^7C_x (0.48)^x\,(0.52)^{7-x}$

Putting $x = 0, 1, 2, 3, \dots\dots, 7$ we get

$P(0) = 0.01, P(1) = 0.066, P(2) = 0.184, \dots\dots$

Multiply these probabilities by 128 we get the expected frequencies as $1, 8, 23, 36, 33, 18, 6, 3$

(Last term $= 128 -$ sum of other terms)

Write the table again with these values