BHARATHIDASAN UNIVERSITY
Tiruchirappalli- 620024
Tamil Nadu, India.

Programme: M.Sc. Statistics

Course Title: Distribution Theory

Course Code: 23ST02CC

Unit-I

Distribution Theory

Dr. T. Jai Sankar
Associate Professor and Head
Department of Statistics

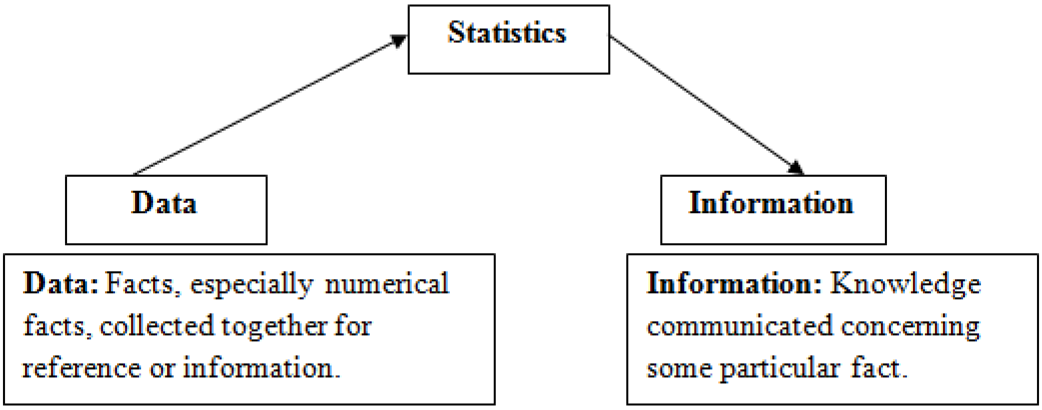
Ms. J. Jenitta Edal Queen
Guest Faculty
Department of Statistics

Distribution Theory

UNIT - I

What is Statistics?

Statistics is a way to get information from data.



Probability Distributions

Probabilistic Experiment

A probabilistic experiment is some occurrence such as the tossing of coins, rolling dice, or observation of rainfall on a particular day where a complex natural background leads to a chance outcome.

Sample Space

The set of possible outcomes of a probabilistic experiment is called the sample, event, or possibility space. For example, if two coins are tossed, the sample space is the set of possible results HH, HT, TH, and TT, where H indicates a head and T a tail.

Random Variable

A random variable is a function that maps events defined on a sample space into a set of values. Several different random variables may be defined in relation to a given experiment. Thus, in the case of tossing two coins the number of heads observed is one random variable, the number of tails is another, and the number of double heads is another. The random variable "number of heads" associates the number 0 with the event TT, the number 1 with the events TH and HT, and the number 2 with the event HH.

In other words, a random variable is a function $X:S \rightarrow R$, where S is the sample space of the random experiment under consideration.

NOTE: By convention, we use a capital letter, say X, to denote a random variable, and use the corresponding lower-case letter x to denote the realization (values) of the random variable.

Discrete Random Variable

If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are whole numbers (countable), it is called a discrete sample space. A random variable is called a discrete random variable if its set of possible outcomes is countable.

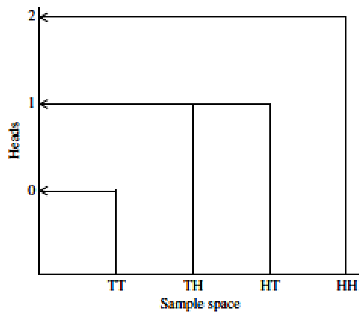
Continuous Random Variable

If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a continuous sample space. When a random variable can take on values on a continuous scale, it is called a continuous random variable.

Variate

A variate is a generalization of the idea of a random variable and has similar probabilistic properties but is defined without reference to a particular type of probabilistic experiment. A *variate* is the set of all random variables that obey a given probabilistic law. The number of heads and the number of tails observed in independent coin tossing experiments are elements of the same variate since the probabilistic factors governing the numerical part of their outcome are identical.

A *multivariate* is a vector or a set of elements, each of which is a variate. A *matrix variate* is a matrix or two-dimensional array of elements, each of which is a variate. In general, dependencies may exist between these elements.



Random Number

A *random number* associated with a given variate is a number generated at a realization of any random variable that is an element of that variate.

Range

Let ***X*** denote a variate and let _***X*** be the set of all (real number) values that the variate can take. The set _***X*** is the *range* of ***X***. As an illustration (illustrations are in terms of random variables) consider the experiment of tossing two coins and noting the number of heads. The range of this random variable is the set {0, 1, 2} heads, since the result may show zero, one, or two heads. (An alternative common usage of the term *range* refers to the largest minus the smallest of a set of variate values.)

Quantile

For a general variate ***X*** let *x* (a real number) denote a general element of the range _***X***. We refer to *x* as the *quantile* of ***X***. In the coin tossing experiment referred to previously, $x \in \{0, 1, 2\}$ heads; that is, *x* is a member of the set {0, 1, 2} heads.

Probability Statement

Let ***X*** = *x* mean "the value realized by the variate ***X*** is *x*." Let Pr[***X*** ≥ *x*] mean "the probability that the value realized by the variate ***X*** is less than or equal to *x*."

Probability Domain

Let *α* (a real number between 0 and 1) denote probability. Let _*α* ***X*** be the set of all values (of probability) that Pr[***X*** ≥ *x*] can take. For a continuous variate, _*α* ***X*** is the line segment [0, 1]; for a discrete variate it will be a subset of that segment. Thus _*α* ***X*** is the *probability domain* of the variate ***X***.

In examples we shall use the symbol ***X*** to denote a random variable. Let ***X*** be the number of heads observed when two coins are tossed. We then have

$$\Pr[X \le 0] = \frac{1}{4}$$

$$\Pr[X \le 1] = \frac{3}{4}$$

$$\Pr[X \le 2] = 1$$

$$\text{and hence } \Re_X^{\alpha} = \{\tfrac{1}{4}, \tfrac{3}{4}, 1\}.$$

Discrete Probability Distributions

The probability distribution of a discrete random variable *X* lists the values and their probabilities.

Value of X	x_1	x_2	x_3	. . .	x_k
Probability	p_1	p_2	p_3	. . .	p_k

where, $0 \le p_i \le 1$ and $p_1 + p_2 + \cdots + p_k = 1$

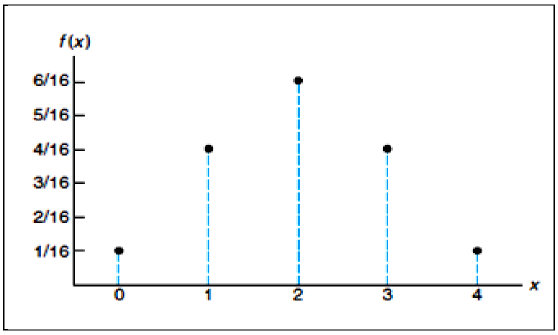
Probability Mass Function (PMF)

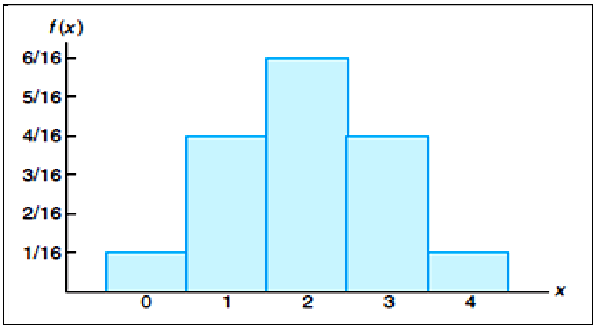
The set of ordered pairs *(x, f (x))* is a probability function, probability mass function, or probability distribution of the discrete random variable *X* if, for each possible outcome *x*,

i). $f(x) \ge 0$,

ii). $\sum_x f(x) = 1$,

iii). $P(X = x) = f(x)$.

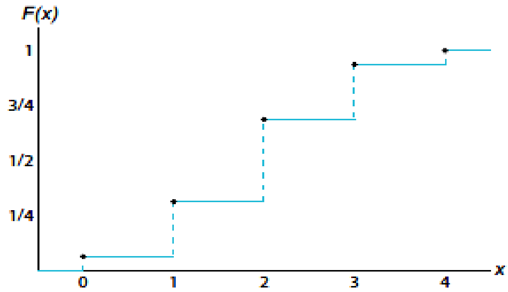




Cumulative Distribution Function (CDF) of a Discrete Random Variable

The cumulative distribution function *F(x)* of a discrete random variable *X* with probability mass function *f(x)* is

$$F(x) = P(X \le x) = \sum_{t \le x} f(t), \qquad \text{for } -\infty < x < \infty.$$



Continuous Probability Distributions

Probability Density Function (PDF)

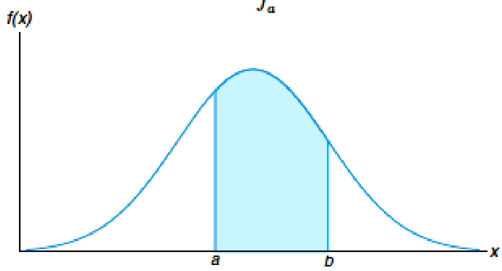
The function *f (x)* is a probability density function (pdf) for the continuous random variable *X*, defined over the set of real numbers, if

i). $f(x) \ge 0$ for all $x \in R$.

ii). $\int_{-\infty}^{\infty} f(x) = 1.$

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

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



NOTE. If X is a continuous random variable, then

$$\begin{aligned} P(a < X < b) &= P(a \le X < b) \\ &= P(a < X \le b) \\ &= P(a \le X \le b). \end{aligned}$$

This is NOT the case for the *discrete* situation.

Cumulative Distribution Function (CDF) of a Continuous Random Variable

The cumulative distribution function *F(x)* of a continuous random variable *X* with probability density function *f (x)* is

$$\begin{aligned} F(x) &= P(X \le x) \\ &= \int_{-\infty}^{x} f(t)\,dt, \quad \text{for } -\infty < x < \infty. \end{aligned}$$

NOTE: A random variable is continuous if and only if its CDF is an everywhere continuous function.

Problem

If the r.v. X takes values 1, 2, 3 and 4 such that 2P(x=1) = 3P(x=2) = P(x=3)=5P(x=4). Find the probability distribution and cdf of x.

Solution

Given X is a discrete random variable (i.e., the values are X= x = 1, 2, 3, 4).

Let 2P(x=1) = 3P(x=2) = P(x=3) = 5P(x=4) = k

$$\Rightarrow \; 2P(x=1) = k \Rightarrow P(x=1) = \frac{k}{2}$$

$$\Rightarrow \; 3P(x=2) = k \Rightarrow P(x=2) = \frac{k}{3}$$

$$\Rightarrow \; P(x=3) = k \Rightarrow P(x=3) = k$$

$$\Rightarrow \; 5P(x=4) = k \Rightarrow P(x=4) = \frac{k}{5}.$$

Since, the total probability is 1,

$$\frac{k}{2}+\frac{k}{3}+k+\frac{k}{5}=1$$

Simplifies the above equation we get,

$$k=\frac{30}{61}$$

$$\therefore P(x=1)=\frac{k}{2}=\frac{30}{61\times 2}=\frac{15}{61}$$

$$P(x=2)=\frac{k}{3}=\frac{30}{61\times 3}=\frac{10}{61}$$

$$P(x=3)=k=\frac{30}{61}$$

$$P(x=4)=\frac{k}{5}=\frac{30}{61\times 5}=\frac{6}{61}$$

So the probability distribution is

X = x	P(x)
1	15/61
2	10/61
3	30/61
4	6/61

Which is the required probability distribution?

To find cdf:-

when $x<1$, $F(x)=0$

when $x\leq 1$, $F(x)=15/61$

when $x\leq 2$, $F(x)=10/61+15/61=25/61$

when $x\leq 3$, $F(x)=30/61+10/61+15/61=55/61$

when $x\leq 4$, $F(x)=61/61+61/61=1$.

Joint Probability Distributions

We frequently need to examine two more (discrete or continuous) random variables simultaneously.

Joint Probability Mass Function (Joint PMF)

The function *f (x, y)* is a joint probability function, or probability mass function of the discrete random variables *X* and *Y* if

i). $f(x,y) \geq 0$ for all (x,y),

ii). $\sum_x \sum_y f(x,y) = 1$,

iii). $\mathrm{P}(X = x, Y = y) = f(x,y)$.

For any region A in the xy plane,

$$\mathrm{P}[(X,Y) \in A] = \sum\sum_A f(x,y).$$

Marginal Probability Mass Function (Marginal PMF)

The marginal distributions of *X* alone and *Y* along are respectively.

$$g(x) = \sum_y f(x,y) \quad \text{and} \quad h(y) = \sum_x f(x,y).$$

Conditional Probability Mass Function (Conditional PMF)

Let *X* and *Y* be two discrete random variables. The conditional distribution of *Y* given that *X* = *x* is

$$\begin{aligned} f(y|x) &= \mathrm{P}(Y = y | X = x) \\ &= \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(X = x)} \\ &= \frac{f(x,y)}{g(x)}, \end{aligned}$$

provided $\mathrm{P}(X = x) = g(x) > 0$.

Similarly, the conditional distribution of X given that $Y = y$ is

$$\begin{aligned} f(x|y) &= \mathrm{P}(X = x | Y = y) \\ &= \frac{\mathrm{P}(X = x, Y = y)}{\mathrm{P}(Y = y)} \\ &= \frac{f(x,y)}{h(y)}, \end{aligned}$$

provided $\mathrm{P}(Y = y) = h(y) > 0$.

Joint Probability Density Function (Joint PDF)

The function *f (x, y)* is a joint probability density function of the continuous random variables *X* and *Y* if

i). $f(x,y) \geq 0$ for all (x,y),

ii). $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\, dx\, dy = 1$,

iii). $\mathrm{P}[(X,Y) \in A] = \int\int_A f(x,y)\, dx\, dy$, for any region A in the xy plane.

Marginal Probability Density Function (Marginal PDF)

The marginal distributions of *X* alone and *Y* along are respectively,

$$g(x) = \int_{-\infty}^{\infty} f(x,y)\, dy$$

and

$$h(y) = \int_{-\infty}^{\infty} f(x,y)\, dx.$$

Conditional Probability Density Function (Conditional PDF)

Let *X* and *Y* be two continuous random variables. The conditional distribution of $Y|X = x$ is

$$f(y|x) = \frac{f(x,y)}{g(x)}, \quad \text{provided } g(x) > 0.$$

Similarly, the conditional distribution of $X|Y = y$ is

$$f(x|y) = \frac{f(x,y)}{h(y)}, \quad \text{provided } h(y) > 0.$$

Problem

The joint pmf of (X, Y) is given by P(x,y) = k(2x+3y), x = 0, 1, 2; y = 1, 2, 3. Find all the marginal and conditional probability distributions. Also find the probability distribution of X + Y.

Solution

P(x,y) = k(2x+3y), x = 0, 1, 2; y= 1, 2, 3

The probability distribution of given function is

x \ y	1	2	3
0	3k	6k	9k
1	5k	8k	11k
2	7k	10k	13k

Since $\Sigma\Sigma P(x, y) = 1$.

$$\Rightarrow 3k + 6k + 9k + 5k + 8k + 11k + 7k + 10k + 13k = 1 \;\Rightarrow k=1/72$$

Therefore, the probability distribution is

x \ y	1	2	3	P_{i*}
0	$\frac{3}{72}$	$\frac{6}{72}$	$\frac{9}{72}$	$\frac{18}{72}$
1	$\frac{5}{72}$	$\frac{8}{72}$	$\frac{11}{72}$	$\frac{24}{72}$
2	$\frac{7}{72}$	$\frac{10}{72}$	$\frac{13}{72}$	$\frac{30}{72}$
P_{*j}	$\frac{15}{72}$	$\frac{24}{72}$	$\frac{33}{72}$	

The marginal probability distribution of X is

X	P_{i*}
0	18/72
1	24/72
2	30/72

The marginal probability distribution of Y is

Y	P_{i*}
1	15/72
2	24/72
3	33/72

The conditional distribution of X given Y = 1 is

$$\frac{P_{ij}}{P_{*j}} = \frac{P_{i1}}{P_{*1}}$$

$$\frac{P_{01}}{P_{*1}} = \frac{3/72}{15/72} = \frac{3}{72} \times \frac{72}{15} = \frac{1}{5}$$

$$\frac{P_{11}}{P_{*1}} = \frac{5/72}{15/72} = \frac{5}{72} \times \frac{72}{15} = \frac{1}{3}$$

$$\frac{P_{21}}{P_{*1}} = \frac{7/72}{15/72} = \frac{7}{72} \times \frac{72}{15} = \frac{7}{15}$$

X	0	1	2
$\frac{P_{ij}}{P_{*j}} = \frac{P_{i1}}{P_{.1}}$	$\frac{1}{5}$	$\frac{1}{3}$	$\frac{7}{15}$

Similarly we can find the conditional distribution of X given Y = 2

X	$\frac{P_{ij}}{P_{*j}}$
0	$\frac{1}{4}$
1	$\frac{1}{3}$
2	$\frac{5}{12}$

The conditional distribution of X given Y = 3

X	$\frac{P_{ij}}{P_{*j}}$
0	$\frac{9}{33}$
1	$\frac{11}{33}$
2	$\frac{13}{33}$

The conditional distribution of Y given X = 0

Y	$\frac{P_{ij}}{P_{i*}}$
1	$\frac{1}{6}$
2	$\frac{1}{3}$
3	$\frac{1}{2}$

The conditional distribution of Y given X = 1

Y	$\frac{P_{ij}}{P_{i*}}$
1	$\frac{5}{24}$
2	$\frac{1}{3}$
3	$\frac{11}{24}$

The conditional distribution of Y given X = 2

Y	$\frac{P_{ij}}{P_{i*}}$
1	$\frac{7}{30}$
2	$\frac{10}{30}$
3	$\frac{13}{30}$

$$P(X+Y) = P(X+Y = 1,2,3,4,5)$$

$$= \frac{3}{72} + \left(\frac{6}{72}+\frac{5}{72}\right) + \left(\frac{9}{72}+\frac{8}{72}+\frac{7}{72}\right) + \left(\frac{10}{72}+\frac{11}{72}\right) + \frac{13}{72}$$

$$P(X+Y) = 1.$$

Problem 3:

The joint pdf of a two dimensional random variable (X, Y) is given by

$$f(X,Y) = xy^2 + \frac{x^2}{8} \;,\; 0 \le x \le 2, \quad 0 \le y \le 1.$$

Compute

$$\text{(i) } P(X > 1) \text{ (ii) } P\left(Y < \frac{1}{2}\right) \text{ (iii) } P\left(X > 1 \mid Y < \frac{1}{2}\right) \text{(iv) } P\left(Y < \frac{1}{2} \mid X > 1\right)$$

$$\text{(v) } P(X < Y) \quad \text{(vi) } P(X + Y \le 1).$$

Also

(a) Are X and Y independent?
(b) Find the conditional pdf of X given Y.
(c) Find the condition pdf of Y given X.

Solution

$$\text{(i) } P(X>1) = \int_0^1\int_1^2 f(x,y)\,dx\,dy$$

$$= \int_0^1\int_1^2 \left(xy^2 + \frac{x^2}{8}\right) dx\,dy$$

$$= \int_0^1 \left[y^2.\frac{x^2}{2} + \frac{x^3}{24}\right]_1^2 dy$$

$$= \int \left(y^2\left(\frac{4}{2}\right) + \frac{8}{24} - \frac{y^2}{2} - \frac{1}{24}\right) dy$$

$$= \int_0^1 \left(\frac{3y^2}{2} + \frac{7}{24} \right) dy$$

$$= \left[\frac{3}{2} \times \frac{y^3}{3} + \frac{7}{24} \times y \right]_0^1$$

$$P(X > 1) = \frac{19}{24}.$$

(ii) $P(Y < \frac{1}{2}) = \int_0^{\frac{1}{2}} \int_0^2 f(x, y)\, dx\, dy$

$$= \int_0^{1/2} \int_0^2 \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$$

$$= \int_0^{1/2} \left[y^2 . \frac{x^2}{2} + \frac{x^3}{24} \right]_0^2 dy$$

$$= \int_0^{1/2} \left(2y^2 + \frac{1}{3} \right) dy$$

$$= \left[\frac{2y^3}{3} + \frac{1}{3} y \right]_0^{1/2}$$

$$P\left(Y < \frac{1}{2} \right) = \frac{1}{4}.$$

(iii) $P\left(X > 1 | Y < \frac{1}{2} \right) = \dfrac{P\left(X > 1, Y < \frac{1}{2} \right)}{P\left(Y < \frac{1}{2} \right)}$

$$P\left(X > 1 . Y < \frac{1}{2} \right) = \int_0^{\frac{1}{2}} \int_1^2 \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$$

$$= \int_0^{\frac{1}{2}} \int_1^2 \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$$

$$= \int_0^{\frac{1}{2}} \left[y^2 . \frac{x^2}{2} + \frac{x^3}{24} \right]_1^2 dy$$

$$= \int_0^{\frac{1}{2}} \left(\frac{3y^2}{2} + \frac{7}{24} \right) dy$$

$$= \left[\frac{3y^3}{6} + \frac{7}{24} \right]_0^{\frac{1}{2}} = \frac{5}{24}.$$

$$\text{But } P\left(Y < \frac{1}{2} \right) = \frac{1}{4}$$

$$\therefore P\left(X > 1 \mid Y < \frac{1}{2} \right) = \frac{5/24}{1/4} = \frac{5}{6}.$$

(iv) $P\left(Y < \frac{1}{2} \mid X > 1 \right) = P\left(Y < \frac{1}{2}, X > 1 \right) | P(X > 1)$

$$\frac{5/24}{24/19} = \frac{5}{24} \times \frac{24}{19}$$

$$P\left(Y < \frac{1}{2} \mid X > 1 \right) = \frac{5}{19}.$$

(v) $P(X < Y) = \int_0^1 \int_0^y f(x, y)\, dx\, dy$

$$= \int_0^1 \int_0^y \left(xy^2 + \frac{x^2}{8} \right) dx\, dy$$

$$= \int_0^1 \left[y^2 . \frac{x^2}{2} + \frac{x^3}{24} \right]_0^y dy$$

$$= \int_0^1 \left(\frac{y^4}{2} + \frac{y^3}{24} \right) dy$$

$$= \left(\frac{y^5}{10} + \frac{y^4}{96} \right)_0^1$$

$$P(X < Y) = \frac{53}{480}.$$

(vi) $P(X+Y \le 1) = \iint f(x,y)\,dx\,dy$

$$P(X+Y \le 1) = \int_0^1 \int_0^{1-y} f(x,y)\,dx\,dy$$

$$= \int_0^1 \int_0^{1-y} \left(xy^2 + \frac{x^2}{8} \right) dx\,dy$$

$$= \int_0^1 \left[y^2 \frac{x^2}{2} + \frac{x^3}{24} \right]_0^{1-y} dy$$

$$= \int_0^1 \left[\frac{1}{2}\left(\frac{y^3}{3} + \frac{y^5}{5} - \frac{2y^4}{4} \right) + \frac{1}{24}\left(1 - y^3 - 3y + 3y^2\right) \right] dy$$

$$= \frac{1}{2}\left(+ \frac{1}{5} - \frac{2}{4} \right) + \frac{1}{24}\left(1 - \frac{1}{4} - \frac{3}{2} + \frac{3}{3} \right)$$

$$P(X+Y \le 1) = \frac{13}{480}.$$

Also,

(a) In order to prove X and Y are independent, we prove

$f(x,y) = f_x(x) \,.\, f_y(y)$

$f(X,Y) = xy^2 + \frac{x^2}{8}$; x= 0 to 2; y= 0 to 1

$$f_x(x) = \int_{-\infty}^{\infty} f(x,y)\,dy$$

$$= \int_0^1 \left(xy^2 + \frac{x^2}{8} \right) dy$$

$$= \left(x.\frac{y^3}{3} + \frac{x^2}{8} y \right)_0^1$$

$$f_x(x) = \frac{x}{3} + \frac{x^2}{8}, \; 0 \le x \le 2$$

Similarly, $f_y(y) = \int_{-\infty}^{\infty} f(x,y)\,dx$

$$= \int_0^2 \left(xy^2 + \frac{x^2}{8} \right) dx$$

$$= \left[y^2 . \frac{x^2}{2} + \frac{x^3}{24} \right]_0^2$$

$$f_y(y) = 2y^2 + \frac{1}{3}$$

$$f_x(x) \times f_y(y) = \left(\frac{x}{3} + \frac{x^2}{8} \right) \times \left(2y^2 + \frac{1}{3} \right)$$

$$\neq f(x, y)$$

$\therefore$ X and Y are not independent random variables.

(b) Conditional pdf of X given Y is given by

$$f(x \mid y) = \frac{f(x, y)}{f_y(y)}$$

$$= \frac{xy^2 + \frac{x^2}{8}}{2y^2 + \frac{1}{3}}$$

$$f(x \mid y) = \frac{3}{8} . \frac{8xy^2 + x^2}{6y^2 + 1} .$$

(c) Conditional pdf of Y given X is given by

$$f(y \mid x) = \frac{f(x, y)}{f_x(x)}$$

$$= \frac{xy^2 + \frac{x^2}{8}}{\frac{x}{3} + \frac{x^2}{8}}$$

$$= \frac{3}{8} . \frac{8xy^2 + x^2}{\frac{x}{3} + \frac{x^2}{8}}$$

$$f(y \mid x) = 3 . \frac{8xy^2 + x^2}{8x + 3x^2} .$$

Standard Distributions

Binomial Distribution - James Bernoulli in 1700

Definition

A random variable X which takes two values 0 and 1 with probabilities q and p respectively. i.e., P(X=1) = p; P(X=0) = q is called a Bernoulli variate and its said have a Bernoulli distribution.

If the experiment is repeated n-times independently with two possible outcomes, then they are called Bernoulli trials.

An experiment consisting of a repeated n number of Bernoulli trails is called Bernoulli experiment.

Binomial Experiment

A binomial distribution can be used under the following condition:

(i) Any trail with two possible outcomes that is any trail result in a success or failure.
(ii) The number of trials n is finite and independent, when n is number of trial.
(iii) a probability of success is the same in each trial. i.e., p is the constant.

Definition

A random variable X is said to have a binomial distribution, if its pmf is given by

$$P(X=x)=\begin{cases} nC_x P^x q^{n-x}, & x=0,1,2,..n \\ o & ,\text{otherwise} \end{cases} \quad \text{where } q = 1-p$$

It is denoted by B(n, p), where n and p are parameters

Applications of Binomial Distribution

1. The quality control measures and sampling process in industries to classify the items are defective or non-defective.
2. Medical applications as a success or failure of a surgery and cure or non cure of a patient.
3. Military application as a hit a target or miss a target

Derivation of mean and variance of B (n, p)

By the definition of mathematical expectation,

$$E(X) = \sum_{x=0}^{n} x P(x) = \sum_{x=0}^{n} x\, nC_x p^x q^{n-x}$$

$$= np \sum_{x=1}^{n} n-1C_x\, p^{x-1} q^{n-x}$$

$$= np\ (\ q + p\ )^{n-1} \quad \text{(by binomial expansion)}$$

$$= np(1) \quad (q+p=1)$$

Mean = E(x) = np

$\mathrm{Var}(x) = E(x^2) - [E(x)]^2$

$$E(x^2) = \sum_{x=0}^{n} x^2 P(x)$$

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

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

$$= \sum_{x=0}^{n} x(x-1).nC_x p^x q^{n-x} + np \text{ (From (1))}$$

$$= \sum_{x=0}^{n} x(x-1).\frac{n(n-1)}{x(x-1)} n-2C_{x-2} p^2 .p^{x-2} q^{n-x} + np.$$

$$= n(n-1)p^2 \sum_{x=0}^{n} n-2C_{x-2} p^{x-2} q^{n-x} + np$$

$$= n\ (n-1)\ p\ (q+p)^{n-2} + np$$

$$= n\ (n - 1)\ p^2 + np$$

$$E(x^2) = np\ (np + q)$$

$$\mathrm{Var}\ (x) = E(x^2) - [E(x)]^2$$

$$= np\ (\ np + q) - (np)^2$$

$$= n^2p^2 + npq - n^2p^2$$

$$\mathrm{Var}\ (x) = npq.$$

MGF of mean and variance

By the definition of MGF,

$$M_x(t) = E[e^{tx}]$$

$$= \sum_{x=0}^{n} e^{tx} p(x)$$

$$= \sum_{x=0}^{n} e^{tx} nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} nC_x \left(pe^t\right)^x q^{n-x}$$

$$= nC_0 \left(pe^t\right)^0 q^n + nC_1 \left(pe^t\right)^1 q^{n-1} + \ldots + nC_n \left(pe^t\right)^n q^{n-n}$$

$$= q^n + nC_1 \left(pe^t\right) q^{n-1} + \ldots + \left(pe^t\right)^n$$

$$M_x(t) = \left(q + pe^t\right)^n$$

Differentiate with respect to t, we get

$$\frac{d}{dt} M_x(t) = n(q + pe^t)^{n-1} . pe^t$$

$$\text{Put } t = 0,\ \frac{d}{dt} M_x(t) = n(q + p)^{n-1} . pe^0$$

$$\text{Mean} = np = \mu_1'$$

$$\frac{d}{dt} M_x(t) = n(q + pe^t)^{n-1} . pe^t$$

$$= np\left(q + pe^t\right)^{n-1} e^t$$

$$\frac{d^2}{dt^2} M_x(t) = np\left\{(q + pe^t)^{n-1} e^t + e^t (n-1).\left(q + pe^t\right)^{n-2} . pe^t\right\}$$

$$\frac{d^2}{dt^2} M_x(t)\Big|_{t=0} = np\{1 + (n-1)p\}$$

$$np + n^2p^2 - np^2 = \mu_2'$$

$$\therefore \text{var}(x) = \mu_2' - (\mu_1')^2$$

$$= np + n^2p^2 - np^2 - (np)^2$$

$$\text{Var}\ (x) = npq$$

Definition of Moments

Moments about origin μ_r' is defined as the expectations of the powers of the r.v X. That is $\mu_r' = E(x^r)$. Similarly, the central momentsabout mean is defined as $\mu_r = E(x-\mu)^r$.

Recurrence relation for the central moments of a B(n, p)

By the definition of k^{th} order central moment μ_k is given by

$$\mu_k = E(x-\mu)^k = E(x-np)^k$$

$$= \sum_{x=0}^{n} (x-np)^k nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} (x-np)^k nC_x p^x (1-p)^{n-x}$$

$$= \sum_{x=0}^{n} nC_x (x-np)^k p^x (1-p)^{n-x}$$

Differentiate with respect to p, we get

$$\frac{d}{dp}\mu_k = \sum_{x=0}^{n} nC_x \left\{ (x-np)^k \left(p^x (n-x)(1-p)^{n-x-1}(-1) + (1-p)^{n-x}.(xp^{x-1}) + p^x (1-p)^{n-x}.k(x-np)^{k-1}(-n)\right\}$$

After simplification, we get,

$$\frac{d\mu_k}{dp} = -nk\mu_{k-1} + \frac{1}{pq}\mu_{k+1}$$

$$\mu_{k+1} = pq\left[\frac{d\mu_k}{dp} + nk\mu_{k-1}\right]$$

The first four raw moments (or) moment about origin of B(n, P)

By the definition of moments about origin $\mu_r' = E(x^r)$

To find the first four raw moments:

Put r = 1

$$\mu_1' = E(x^1)$$

$$= \sum_{x=0}^{n} xp(x)$$

$$= \sum_{x=0}^{n} x\, nC_x p^x q^{n-x}$$

$$= np\sum_{x=0}^{n} n-1C_x p^{x-1} q^{n-x}$$

$$= np\ (q+p)^{n-1}$$

$$\mu_1' = np$$

$$\mu_2' = E(x^2)$$

$$= \sum_{x=0}^{n} x^2 p(x)$$

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

$$= n(n-1)p^2 \sum_{x=2}^{n} n-2C_{x-2} p^{x-2} q^{n-x} + np$$

$$= n\ (n-1)P^2\ (q+p)^{n-2} + np$$

$$\mu_3' = E(x^3)$$

$$= \sum_{x=0}^{n} x^3 p(x)$$

$$= \sum_{x=0}^{n} [x(x-1)(x-2) + 3x(x-1) + x] nC_x p^x q^{n-x}$$

$$= n(n-1)(n-2)p^3 \sum_{x=0}^{n} n-3C_{x-3} p^{x-3} q^{n-x} + 3n(n-1)p^2 \sum_{x=0}^{n} n-2C_{x-2} p^{x-2} q^{n-x} + np$$

$$\mu_3' = n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np$$

$$\mu_4' = E(x^4)$$

$$= \sum_{x=0}^{n} x^4 p(x)$$

$$= \sum_{x=0}^{n} [x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x] \ nC_x p^x q^{n-x}$$

$$= \sum_{x=0}^{n} x(x-1)(x-2)(x-3) nC_x p^x q^{n-x} + 6\sum_{x=0}^{n} x(x-1)(x-2) nC_x p^x q^{n-x}$$

$$+ 7\sum_{x=0}^{n} x(x-1) nC_x p^x q^{n-x} + \sum_{x=0}^{n} x \ nC_x p^x q^{n-x}$$

$$= n(n\text{-}1)(n\text{-}2)(n\text{-}3)p^4(p+q)^{n-4} + 6n(n\text{-}1)(n\text{-}2)p^3(p+q)^{n-3} + 7n(n\text{-}1)p^2(p+q)^{n-2} + np$$

$$\mu_4' = n(n-1)(n-2)(n-3)p^4 + 6n(n-1)(n-2)p^3 + 7n(n-1)p^2 + np.$$

Additive property of B(n, p) or Reproductive property

Statement:

If $X \sim B(n_1, p)$ and $Y \sim B(n_2, p)$, then $X+Y \sim B(n_1+n_2, p)$ where X and Y are independent.

Proof

We know that, the MGF of $B(n, p) = (q+pe^t)^n$.

$\therefore$ The MGF of $X \sim B(n_1, p) = (q+pe^t)^{n_1}$.

Also the MGF of $Y \sim B(n_2, P) = (q+pe^t)^{n_2}$.

We know that, If X and Y are independent r.vs, then

$$M_{X+Y}(t) = M_x(t) \cdot M_x(t)$$

$$= (q+pe^t)^{n_1} \cdot (q+pe^t)^{n_2}$$

$$= (q+pe^t)^{n_1+n_2}$$

$$\therefore M_{X+Y}(t) = (q+pe^t)^{n_1+n_2}$$

Which is the MGF of $B(n_1+n_2, p)$

Note:

If $X_1, X_2, \ldots, X_k$ are independent binomial variates with parameters $(n_1,p), (n_2,p), \ldots, (n_k,p)$ respectively, then $X_1+X_2+\ldots+X_k$ is also a binomial variate with parameter $(n_1+n_2+\ldots+n_k, p)$.

Mode of Binomial distribution

Definition

The value of x at which p(x) obtains maximum is called mode of the distribution.

Let X be a binomial random variable. Then

$$P(X=x)=p(x)=nC_x p^x q^{n-x};\ x = 0, 1, 2,\ldots n$$

The mode of the binomial distribution is defined by m_0 and it is given by

$$p(m_0-1) \le p(m_0) \ge p(m_0+1)$$

Consider,

$$p(m_0-1) \le p(m_0)$$

$$nC_{m_0-1}p^{m_0-1}.q^{n-(m_0-1)} \le nC_{m_0}p^{m_0}q^{n-m_0}$$

$$\Rightarrow \frac{(n-m_0)!m_0!}{(n-m_0+1)!(m_0-1)!}.\frac{q}{p} \le 1$$

$$\frac{m_0}{n-m_0+1} \le \frac{p}{q}$$

$$m_0 \le p(n+1)$$

Consider,

$$P(m_0) \ge p(m_0+1)$$

$$nC_{m_0}p^{m_0}q^{n-m_0} \ge nC_{m_0+1}p^{m_0+1}.q^{n-(m_0+1)}$$

$$\Rightarrow \frac{(n-m_0-1)!(m_0+1)!}{(n-m_0)!(m_0)!} \ge \frac{p}{q}$$

$$\frac{m_0+1}{n-m_0} \ge \frac{p}{q}$$

$$m_0 \ge np-q$$

from (1) and (2)

$np - q \le m_0 \le p\,(n+1)$

For checking:

when n = 10, p=1/2 , q = ½

$4.5 \le m_0 \le 5.5.$

Characteristic function and Cumulative function or cumulative generating function

The characteristic function is defined

$$\varphi_x(t) = E[e^{itx}]$$

Cumulative generating function is defined by

$$\kappa_x(t) = \log\ M_x(t)$$

Characteristic function of B(n,p)

By the definition of characteristic function,

$$\varphi_x(t) = E[e^{itx}]$$

$$= \sum_{x=0}^{n} e^{itx} p(x)$$

$$= \sum_{x=0}^{n} e^{itx} nC_x p^x q^{n-x}$$

$$\varphi_x(t) = (q + pe^{it})^n .$$

Poisson Distribution - Simen Denis Poisson

Definition

A random variable X is said to follow the Poisson distribution if its probability mass function is given by,

$$p(X = x) = p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0,1,2,\ldots\infty$$

Here the λ is the parameter and λ > 0.

Poisson distribution as a limiting case of Binomial distribution:

Poisson distribution as a limiting case of Binomial distribution under the following condition:

i) The number of trial n is infinitely large. i.e, n→∞.
ii) The constant probability of success p in each trail is vary small. i.e, p→0
iii) np = λ is finite, where λ is a positive real number.

Proof:

In the case of Binomial distribution, the probability of x success is given by,

$$p(X = x) = p(x) = nC_x p^x q^{n-x}$$

$$= \frac{n(n-1)(n-2)\ldots[n-(x-1)]}{x!} p^x q^{n-x}$$

Put $np = \lambda$; $p = \lambda/n$

$$q = 1 - \frac{\lambda}{n}$$

$$\Rightarrow p(x) = \frac{n(n-1)(n-2)\ldots[n-(x-1)]}{x!}\left(\frac{\lambda}{n}\right)^{x}\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{\lambda^x}{x!}.\frac{n}{n}.\frac{n-1}{n}.\frac{n-2}{n}\ldots\frac{n-(x-1)}{n}.\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}$$

$$= \frac{\lambda^x}{x!}\left[1.\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\left(1-\frac{3}{n}\right)..\left(1-\frac{x-1}{n}\right)\right]\left(1-\frac{\lambda}{n}\right)^{n}\left(1-\frac{\lambda}{n}\right)^{-x}$$

Taking limit $n \to \infty$, we get

$$p(X = x) = p(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}, x = 0,1,2,\ldots\infty$$

which is the pmf of Poisson distribution.

∴ Poisson distribution is the limiting case of binomial distribution.

Mean and variance of Poisson distribution

$$\text{Mean, } E(x) = \sum_{x=0}^{\infty} x\, p(x)$$

$$= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda\lambda^{x-1}}{x(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$= \lambda e^{-\lambda} e^{\lambda}$$

$$\therefore \text{Mean } E(x) = \lambda$$

$$\text{Variance } (x) = E(x^2) - [E(x)]^2$$

$$E(x^2) = \sum_{x=0}^{\infty} x^2\, p(x)$$

$$= \sum_{x=0}^{\infty} [x(x-1) + x] \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^{x}}{x!} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^{x}}{x!}$$

$$E(x^2) = \lambda^2 + \lambda$$

$$\text{Var}(x) = E(x^2) - [E(x)]^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$\text{Var}(x) = \lambda$$

$$\therefore \text{Mean} = \text{Variance} = \lambda.$$

MGF of mean and variance

By the definition of MGF,

$$M_x(t) = E[e^{tx}]$$

$$= \sum_{x=0}^{n} e^{tx} p(x)$$

$$= \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda et} = e^{\lambda(e^t - 1)}$$

$$M_x(t) = e^{\lambda(e^t - 1)}$$

To find mean and variance

By the property of MGF,

$$M_x'(t) = e^{\lambda(e^t - 1)}.\lambda(e^t)$$

$$M_x'(t)\big|_{t=0} = e^{\lambda(1-1)}.\lambda(e^0) = \lambda$$

$$M_x'(t) = \lambda$$

$$\because M_X'(t) = \lambda e^t e^{\lambda(e^t - 1)}$$

$$M_x''(t) = \lambda\left[e^t.e^{\lambda(e^t - 1)}.\lambda e^t + e^{\lambda(e^t - 1)}.e^t\right]$$

$$M_x''(t)\big|_{t=0} = \lambda[\lambda + 1] = \lambda^2 + \lambda = \mu_2'$$

$$\text{Var}(x) = \mu_2 = \mu_2' - \left(\mu_1'\right)^2$$

$$= \lambda^2 + \lambda - \lambda^2$$

$$\text{Var}(x) = \lambda$$

$$\therefore \text{Mean} = \text{Variance} = \lambda.$$

Recurrence formula for the central moments of the Poisson distribution:

For Poisson distribution with parameter λ; the recurrence formula is,

$$\mu_{r+1} = \lambda\left[\frac{d\mu_r}{d\lambda} + r.\mu_{r-1}\right]$$

Proof

By definition of r^{th} order central moment is given by

$$\mu_r = E(x-\mu)^r$$

$$= E(x-\lambda)^r \qquad (\because E(x) = \lambda)$$

$$= \sum_{x=0}^{\infty}(x-\lambda)^r.p(x)$$

$$\mu_r = \sum_{x=0}^{\infty}(x-\lambda)^r \frac{e^{-\lambda}\lambda^x}{x!}$$

Differentiate with respect to λ, we get,

$$\frac{d}{d\lambda}\mu_r = \sum_{x=0}^{\infty}\frac{1}{x!}\left[(x-\lambda)^r.(e^{-\lambda}x\lambda^{x-1} + \lambda^x e^{-\lambda}(-1)) + (e^{-\lambda}\lambda^x).r(x-\lambda)^{r-1}(-1)\right]$$

$$\Rightarrow \lambda\frac{d\mu_r}{d\lambda} = \mu_{r+1} - \lambda r\ \mu_{r-1}$$

$$\Rightarrow \mu_{r+1} = \lambda\frac{d\mu_r}{d\lambda} + \lambda r\ \mu_{r-1}$$

$$\Rightarrow \mu_{r+1} = \lambda\left[\frac{d\mu_r}{d\lambda} + r\mu_{r-1}\right].$$

The central moment's μ_1, μ_2, μ_3 and μ_4:

The recurrence formula for central moments of Poisson distribution is,

$$\mu_{r+1} = \lambda \frac{d\mu_r}{d\lambda} + \lambda r \, \mu_{r-1} \qquad (*)$$

Also, we know that, $\mu_0 = 1$

$$\mu_1 = 0.$$

In order to get μ_2, put r=1 in (*),

$$\therefore \mu_2 = \lambda \frac{d\mu_1}{d\lambda} + \lambda \mu_0$$

$$= \lambda x_0 + \lambda x_1$$

$$\mu_2 = \lambda .$$

In order to get μ_3,Put r = 2 in (*),

$$\therefore \mu_3 = \lambda \frac{d\mu_2}{d\lambda} + 2\lambda \, \mu_{2-1}$$

$$= \lambda.1 + 2\lambda(0)$$

$$\mu_3 = \lambda$$

In order to get μ_4,Put r = 3 in (*),

$$\therefore \mu_4 = \lambda \frac{d\mu_3}{d\lambda} + 3\lambda \, \mu_2$$

$$= \lambda.1 + 3\lambda.\lambda$$

$$\mu_4 = \lambda + 3\lambda^2$$

$\therefore$ $\mu_1 = 0$, $\mu_2 = \lambda$, $\mu_3 = \lambda$, $\mu_4 = \lambda + 3\lambda^2$ are the first four central moments.

The first four moments about origin

By the definition of r[th] order raw moments,

$$\mu_r' = E[x^r]$$

$$\therefore \mu_1' = E(x) = E(x)$$

$$= \sum_{x=0}^{\infty} x \, . \, p(x)$$

$$= \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} x \frac{e^{-\lambda}\lambda\lambda^{x-1}}{x(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!}$$

$$\mu_1' = \lambda$$

Also, $\mu_2' = E(x^2)$

$$\mu_2' = \sum_{x=0}^{\infty} x^2 \, p(x)$$

$$= \sum_{x=0}^{\infty} [x(x-1) + x] \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$\mu_2' = \lambda^2 + \lambda$$

Also, $\mu_3{}' = E(x^3)$

$$\mu_3{}' = \sum_{x=0}^{\infty} x^3\, p(x)$$

$$= \sum_{x=0}^{\infty} x^3 \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} [x(x-1)(x-2) + 3x(x-1) + x] \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2) \frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} 3x(x-1) \frac{e^{-\lambda}\lambda^x}{x!} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda}\lambda^3 e^{\lambda} + 3e^{-\lambda}\lambda^2 e^{\lambda} + \lambda$$

$$\mu_3{}' = \lambda^3 + 3\lambda^2 + \lambda$$

Also $\mu_4{}' = E(x^4)$

$$\mu_4{}' = \sum_{x=0}^{\infty} x^4\, p(x)$$

$$= \sum_{x=0}^{\infty} x^4 \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} [x(x-1)(x-2)(x-3) + 6x(x-1)(x-2) + 7x(x-1) + x] \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= \sum_{x=0}^{\infty} x(x-1)(x-2)(x-3) \frac{e^{-\lambda}\lambda^4\lambda^{x-4}}{x(x-1)(x-2)(x-3)(x-4)!}$$

$$+ 6\sum_{x=0}^{\infty} x(x-1)(x-2) \frac{e^{-\lambda}\lambda^3\lambda^{x-3}}{x(x-1)(x-2)(x-3)!} + 7\sum_{x=0}^{\infty} x(x-1) \frac{e^{-\lambda}\lambda^2\lambda^{x-2}}{x(x-1)(x-2)!} + \sum_{x=0}^{\infty} x \frac{e^{-\lambda}\lambda^x}{x!}$$

$$\mu_4{}' = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda.$$

Additive property:

The sum of independent Poisson variates is also a Poisson variate. i.e., X_1, X_2, …. X_n are n independent Poisson variates with parameter λ_1, λ_2, ….λ_n. Then $X_1 + X_2 + \ldots + X_n$ is also a Poisson variate with parameter $\lambda_1 + \lambda_2 + \ldots + \lambda_n$.

Proof

We know that the MGF of Poisson distribution is,

$$M_x(t) = e^{\lambda(e^t - 1)}$$

Also we know that,

$$M_{x_1+x_2+\ldots+x_n}(t) = M_{x_1}(t).M_{x_2}(t)..M_{x_n}(t)$$

$$= e^{\lambda_1(e^t-1)} + e^{\lambda_2(e^t-1)} + \ldots\ldots + e^{\lambda_n(e^t-1)}$$

$$\therefore M_{X_1+X_2+\ldots+X_n}(t) = e^{(\lambda_1+\lambda_2+\ldots+\lambda_n)(e^t-1)}.$$

which is the MGF of $X_1 + X_2 + \ldots + X_n$ with parameter $\lambda_1 + \lambda_2 + \ldots + \lambda_n$.

$\therefore$ $X_1 + X_2 + \ldots + X_n$ is also Poisson variate.

Examples of a Poisson distribution (Real life Problems)

1. Number of printing mistakes at each page of a book.
2. The number of road accident reported in a city per day.
3. The number of death in a district due to rare disease.
4. The number of defective articles in a pocket of 200.
5. The number of cars passing through a time interval t.

Multinomial Distribution

Definition

Multinomial distribution is the generalization of binomial distribution. Consider k events E_1, E_2,…, E_k. The event E_1 occurs X_1 times, E_2 occurs X_2 times and so on, with the corresponding probability p_1, p_2, …p_k respectively.

Let us assume that the probability of getting i^{th} event in x_i times is $p_i^{x_i}$, i = 1, 2, …, k.

Then the joint probability function of k events is given by,

$$p(x_1, x_2, \ldots x_k) = \frac{n!}{x_1! x_2! \ldots x_k!} . p_1^{x_1} . p_2^{x_2} \ldots p_k^{x_k}$$

$$= \frac{n!}{\prod_{i=1}^{k} x_i!} \prod_{i=1}^{k} p_i^{x_i}$$

This distribution is called multinomial distribution, where $(p_1+p_2+ \ldots+p_k) = 1$, $n = x_1+x_2+ \ldots+ x_k$.

For example, if a fair die is tossed twelve times, the probability of getting 1, 2, 3, 4, 5 and 6 points exactly twice each is given by,

$$p(x_1 = 2, x_2 = 2, x_3 = 2, x_4 = 2, x_5 = 2, x_6 = 2) = \frac{12!}{2!\,2!\,2!\,2!\,2!\,2!} \times \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^2$$

$$= 0.00344.$$

MGF of Multinomial Distribution

To derive MGF, first let us consider a trail which has two outcomes A_1, A_2.

Assume the outcome A_1 occurs x_1 times and A_2 occurs x_2 times then the probability of getting A_1, x_1 times and A_2, x_2 times is given by the function,

$$p(x_1, x_2) = \frac{n!}{x_1! x_2!} p_1^{x_1} p_2^{x_2} \qquad \text{where } p_1 + p_2 = 1 \text{ and } n = x_1 + x_2$$

$$M_{x_1, x_2}(t) = E(e^{t_1 x_1 + t_2 x_2})$$

$$= \sum_x e^{t_1 x_1 + t_2 x_2} . p(x_1, x_2)$$

$$= \sum_x e^{t_1 x_1 + t_2 x_2} \frac{n!}{x_1! x_2!} p_1^{x_1} . p_2^{x_2}$$

$$= \left(p_1 e^{t_1} + p_2 e^{t_2}\right)^n$$

Which is the MGF of $p(x_1, x_2)$. By simply extending this result the mgf for multinomial distribution can be written as,

$$M_{x_1,x_2,\dots,x_k}(t) = \left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^n$$

From this MGF, we can find mean and variance as follows:

$$\therefore M_x(t) = \left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^n$$

$$\frac{d}{dt_i}M_x(t) = n\left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^{n-1} \times p_ie^{t_i}$$

$$\frac{d}{dt_i}M_x(t)\Big|_{t_i=0} = n(p_1 + p_2 + \dots + p_k)^{n-1} \times p_i$$

$$= np_i$$

$$= \mu_1'$$

$$\text{Mean } = \mu' = np_i$$

$$\because \frac{d}{dt_i}M_X(t) = np_i\left[e^{t_i}\left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)\right]$$

$$\frac{d^2}{dt_i^{\,2}}M_x(t) = np_i\left[e^{t_i}(n-1)\left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^{n-2} \times p_ie^{t_i}\right]$$

$$+\left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^{n-1} \times e^{t_i}\Big]$$

$$\frac{d^2}{dt_i^{\,2}}M_x(t) = np_i\left[e^{t_i}(n-1)\left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^{n-2} \times p_ie^{t_i}\right]$$

$$+\left(p_1e^{t_1} + p_2e^{t_2} + \dots + p_ke^{t_k}\right)^{n-1} \times e^{t_i}\Big]$$

$$\frac{d^2}{dt_i^{\,2}}M_x(t)\Big|_{t_i=0} = np_i\left[(n-1).(p_1 + p_2 + \dots + p_k)^{n-2} \times p_i + (p_1 + p_2 + \dots + p_k)^{n-1}\right.$$

$$= np_i[(n-1)p_i + 1]$$

$$= n(n-1)p_i^{\,2} + np_i$$

$$\mu_2' = n(n-1)p_i^{\,2} + np_i$$

$$\text{var}(x) = \mu_2' - (\mu_1')^2$$

$$= n(n-1)p_i^{\,2} + np_i + (np_i)^2$$

$$= np_i[(n-1)p_i + 1 - np_i]$$

$$= np_i(1 - P_i)$$

$$= np_i q_i$$

$$V(x) = np_i q_i \,.$$

Normal Distribution or Gaussian distribution

A random variable X is said to follow a normal distribution if its pdf is given by,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}; \qquad -\infty < x < \infty$$

$$-\infty < \mu < \infty$$

$$\sigma > 0$$

Here, f(x) is a legitimate density function as the total area under the normal curve is unity.

To prove that total probability is one,

$$\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\left(\frac{x-\mu}{\sqrt{2}\sigma}\right)^2} dx$$

$$\text{put } t = \frac{x-\mu}{\sqrt{2}\sigma}$$

$$dt = \frac{1}{\sqrt{2}\sigma} dx$$

$$\Rightarrow dx = \sqrt{2}\sigma \, dt$$

$$= \int_{\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2} \sqrt{2} \, \sigma \, dt$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-t^2} dt$$

$$= \frac{1}{\sqrt{\pi}} \int_{\infty}^{\infty} e^{-t^2} dt$$

$$= \frac{2}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2} dt$$