M A Economics
Semester II

Paper: Statistical Methods (CC 09)
Topic: Probability (Part iii) (Module 4)
Content By: Smita Dubey, Department of Economics
E-mail: smitadubey78@icloud.com

Probability: Theoretical Distributions

Theoretical Distribution or Probability Distributions are distributions which are expected on the basis of previous experience or theoretical considerations. In other words, we can theoretically list outcomes and their probabilities from mathematical models representing any phenomenon of interest.

A probability distribution for a discrete random variable is a mutually exclusive listing of all possible numerical outcomes for that random variable such that a particular probability ofoccurance is associated with each outcome.

If a coin is tossed 2 times, then sample space, S = [TT,TH,HT,HH]

Replacing T by 0 and H by 1, the number of heads in the 2 throws will be,

(TT)= 0, (TH) = 1, (HT)= 1,(HH)= 2.

So we may write S =[0,1,2] and probability of each event will be

P(X=0)= P(T,T)=1/4 P(X=1)= P[(T,H), (H,T)]=1/2 P(X=2)= P(H,H)=1/4

Here, $\sum P(X)=1/4+1/2+1/4=1$

Such a function P(X) is called a probability function of the random variable X. If random variable X is a discrete one, the probability function P(X) is called *probability mass function* and its distribution as *discrete probability distribution.* If the random variable X is of continuous type then the probability function P(X) is called *probability density function* and its distribution as *continuous probability distribution.*

Examples of discrete probability functions are

i. Binomial Distribution
ii. Multinomial distribution
iii. Negative binomial distribution
iv. Poisson distribution
v. Hypergeometric distribution

Example of continuous probability function is

i. Normal distribution

1. Binomial Distribution

Also known as Bernoulli distribution after James Bernoulli (1654-1705) variously called Jacques or Jakob Bernoulli. Binomial distribution is a probability distribution expressing the probability of one set of dichotomous alternatives, i.e. success or failure. Here the frequencies are proportional to the successive terms of the binomial expansion.

Assumptions:

i. n is finite
ii. Trials are repeated n number of times under the same conditions.
iii. Each trial results in either a success or a failure.
iv. Probability of success p in each trial is constant. It does not change from trial to trial.
v. The trials are independent.

The general form of the Binomial distribution is given by

P(r)= $\boldsymbol{{}^{n}C_r\, p^r\, q^{n-r}}$
Where, p= probability of success in a single trial
Probability of failure, denoted by q= 1-p
n= number of trials
r= number of successes in n trials

The probabilities of various possible events are given by the successive terms of the binomial expansion $(q+p)^n$ which is

$$(q+p)^n = q^n + {}^{n}C_1 q^{n-1} p + {}^{n}C_2\, q^{n-2}\, p^2 + \ldots..+ {}^{n}C_r\, p^r\, q^{n-r} + \ldots..+ p^n$$

If the independent trials are repeated k times, the expected frequency of r successes is given by

$$k(q+p\,)^n = k\,{}^{n}C_r\, p^r\, q^{n-r}\,.$$

Obtaining the coefficients of the Binomial

i. The first term is q^n .
ii. Second term is ${}^{n}C_1\, p\, q^{n-1}$
iii. In each succeeding term the power of q is reduced by 1 and power of p is increased by 1.
iv. The coefficient of any term is found by multiplying the coefficient of the preceding term by the power of q in that preceding term and dividing the products by one more than the power of p in that preceding term. So expansion of $(q+p)^n$ will give

$$(q+p)^n = q^n + {}^{n}C_1 q^{n-1} p + {}^{n}C_2\, q^{n-2}\, p^2 + \ldots..+ {}^{n}C_r\, p^r\, q^{n-r} + \ldots..+ p^n$$

Where 1, ${}^{n}C_1$, ${}^{n}C_2$,…… are called coefficients of the binomial.

General Relationships

i. Number of terms in a binomial expansion is n+1.
ii. Exponents of p and q for any single term, when added always sum to 1.
iii. Exponents of q are n, n-1, n-2,….., 1,0 and the exponents of p are 0,1, 2,…, n-1, n respectively.

iv. Coefficients of the n+1terms are always symmetrical, ascending to the middle and then descending. When n is odd, n+1 is even and the coefficients of the two central terms are identical. Coefficients are conveniently obtained from the Pascal's Triangle.

Constants of the Binomial Distribution

i. Mean of the binomial distribution is *np*.

By definition arithmetic mean is = $\frac{\sum x.p(x)}{\sum p(x)}$

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

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

Taking np common

$= np\,[q^{n-1} + (n-1)q^{n-2}p + \frac{(n-1)(n-2)}{2!}q^{n-3}p^2 + \ldots..+ p^{n-1}] = np\,(q+p)^{n-1} = np.1^{n-1} = np$

Also, $\sum p(x) = 1$. So, mean = np

ii. Variance is given by *npq* and Standard deviation is given by $\sqrt{npq}$.

σ^2 or $\mu_2 = \nu_2 - \nu_1^2$

$\nu_2 = \sum x^2.p(x) = np + n^2p^2 - np^2$

$\nu_1 = np$

So, $\mu_2 = \nu_2 - \nu_1^2 = np + n^2p^2 - np^2 - n^2p^2 = np - np^2 = np(1-p) = npq$

$\therefore \sigma^2 = \mu_2 = \nu_2 - \nu_1^2 = npq$

And $\sigma = \sqrt{\mu_2} = \sqrt{npq}$

2. Poisson distribution

Poisson distribution is a discrete probability distribution developed by Simeon Denis Poisson (1781-1840) in 1837. It is used in cases where the chance of an individual event being a success is very small so it is also called the "law of improbable events". Generally used for describing the behavior of rare events like floods, accidental deaths, release of radiation from nuclear reactor etc.

The Poisson distribution is defined as:

$P(r) = \frac{e^{-m}m^r}{r!}$ Where r =0, 1,2, 3,4…. e= 2.7183 (base of natural logarithm) m = the mean of the Poisson distribution = *np* or the average number of occurrences of an event

It is a discrete distribution with single parameter *m*. the probabilities of 0, 1, 2…. successes are given by the successive terms of the expansion

$$e^{-m}\left(1 + m + \frac{m^2}{2!} + \frac{m^3}{3!} + \ldots + \frac{m^r}{r!} + \ldots\right)$$

To know the expected number of occurrences for different successes, we have to multiply each term by the total number of observations, N.

Constants of the Poisson distribution

Since p is very small, the value of q is almost equal to 1. The constants of the Poisson distribution can be found by putting 1 in place of q in the Binomial distribution. Thus,

Mean of the Poisson distribution = m

Standard deviation = $\sqrt{m}$ and Variance = $\mu_2 = m$

For the fitting of Poisson distribution, obtain value of *m* the average occurance and calculate the frequency of 0 successes. The other frequencies can be calculated from

$N(P_0) = Ne^{-m}$, $N(P_1) = N(P_0) \times \frac{m}{1}$, $N(P_2) = N(P_1) \times \frac{m}{2}$, $N(P_3) = N(P_2) \times \frac{m}{3}$ etc.

3. Normal Distribution

The Normal distribution or the Normal probability distribution is the most useful theoretical distribution for continuous variables. It was first described by Abraham De Moivre (1667-1754) as the limiting form of the binomial model in 1733. It was later rediscovered by Gauss in 1809 and Laplace in 1812.

Normal distribution is an approximation to Binomial distribution which tends to the form of continuous curve when *p* is approximately equal to *q* and *n* becomes large. The limiting frequency curve obtained as N becomes large is called the normal frequency curve or the normal curve.

The Normal distribution is given by,

$$P(X) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)}{2\sigma^2}}$$

Where, X = values of the continuous variable μ = mean of the normal random variable

$e = 2.7183 \qquad \pi = 3.1416 \qquad \sqrt{2\pi} = 2.5066$

The equation to a normal curve corresponding to a particular distribution is given by,

$$y = \frac{N}{\sigma\sqrt{2\pi}} e^{\frac{-x^2}{2\sigma^2}}$$

General remarks

i. There is only one normal distribution for any pair of values for μ and σ.
ii. It is a limiting case of binomial distribution when *n* tends to infinity and neither *p* nor *q* is very small.
iii. Normal distribution is limiting case of Poisson distribution when its mean *m* is large.
iv. The mean of a normally distributed population lies at the centre of its normal curve.
v. The two tails of a normal distribution extend indefinitely and never touch the horizontal axis.

Properties of Normal distribution

i. The normal curve is bell-shaped and symmetrical in its appearance.
ii. The height of the curve is maximum at the mean. Hence, the mean and mode coincide and mean, median and mode are all equal.
iii. There is on maximum point of the curve which occurs at the mean. The curve is asymptotic to the base on either side.
iv. It is unimodal.
v. The points of inflexion, the points where the changes in curvature occur, are $\mu \pm \sigma$.
vi. Unlike in Binomial and Poisson distribution, where the variable is discrete, the variable distributed according to the Normal curve is continuous.
vii. The first and third quartiles are equidistant from the median.
viii. The mean deviation is 4^{th} or more precisely 0.7979 of the standard deviation.
ix. The area under the normal curve is distributed as under:
   a. Mean $\pm \sigma$ covers 68.27% of area, 34.135% will lie on either side of mean.
   b. $\mu \pm 2\sigma$ covers 95.45% area
   c. $\mu \pm 3\sigma$ covers 99.73% area.

Conditions for normality

i. Causal forces must be numerous and of approximately equal weight.
ii. Under the condition of homogeneity, the forces must be the same over the universe from which observations are drawn.
iii. Forces must be independent of one another.
iv. Under the condition of symmetry, the operation of causal forces must be such that deviation above the population mean are balanced by the deviations below thw population mean.

Constants of normal distribution

i. The mean of the normal distribution is given by X-bar.
ii. Standard deviation of the normal distribution is σ.
$\mu_2 = \sigma^2$, $\mu_3 = 0$, $\mu_4 = 3\sigma^4$

Significance of Normal distribution

i. To approximate or fit a distribution of measurement under certain circumstances.
ii. To approximate the binomial distribution and other discrete or continuous distributions.
iii. To approximate the distribution of means and certain other quantities calculated from samples, especially large samples.
iv. A normal curve can be converted to standard normal distribution by change of scale and origin. The formula is
$z = \frac{x}{\sigma}$ where x= X- mean