International Journal of Research in Engineering, Science and Management
Volume-3, Issue-2, February-2020
www.ijresm.com | ISSN (Online): 2581-5792

A Study on Theoretical Distributions

E. S. Sasikala[1], Binu George[2]
[1]Assistant Professor, Department of Mathematics, Christ College, Kattappana, India
[2]Assistant Professor, Department of Statistics, Christ College, Kattappana, India

***Abstract*: "A problem well stated is a problem half solved" Charles F. Kettering. This paper deals with the study on Theoretical distributions and different types. It is written to enhance the academic knowledge of students about the different ideologies of Probability distribution in Statistics. Special attention given to properties and applications of Binomial, Poisson and Normal distribution.**

***Keywords*: Binomial, Normal, Poisson, Probability distribution, Random variable.**

1. Introduction

Probabilities can be estimated either practically or mathematically. The probability distributions which are not obtained by actual observations or experiments but are mathematically deduced on certain assumptions are called Theoretical Probability Distributions. Binomial, Poisson and Normal are the three popular theoretical probability distributions. Binomial and Poisson deals with discrete random variables while Normal distribution is used to estimate the probability of continuous variables.

A. Random variable

A real valued function defined over the sample space of a random experiment is called a random variable. Random variable X is a mapping from sample space to real line. Random variables can be classified into two groups, namely discrete random variables and continuous random variables. A discrete random variable can assume only discrete number of point from a real line. While a continuous random variable can take any values from an interval.

B. Probability Distribution

Probability distribution is a schedule that shows the various values of the random variable and probability associated with each value. Probability distribution is distributing the total probability to the different intervals or points.

C. Binomial distribution

The Binomial distribution is an important discrete probability distribution in probability theory. The distribution was first proposed by James Bernoulli in 1700 to deal with random experiments having two possible outcomes namely success and failure.

Consider a random experiment with two possible outcomes only-one is called success and other is called failure. Let the probability of success be p and that of failure is 1-p=q(say). Such an experiment is called a Bernoulli experiment.

Suppose we repeat a Bernoulli experiment n times. Let X denote the number of times the success occurs. Clearly X can take any of the values 0, 1, 2,……….,n. Now,

P[X=x]=P[getting x success and (n-x) failures in the n repetitions of the Bernoulli experiment]

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

This distribution is called the Binomial distribution.

D. Definition

A discrete random variable X is said to follow the binomial distribution with parameters 'n' and 'p' if its probability density function is given by,

$$f(x) = {}^nC_x\, p^x\, q^{n-x}, \qquad x = 0, 1, 2, \ldots, n$$
$$0<p<1$$
$$p+q=1$$

The distribution is usually denoted by the symbol B(x, n, p) or B(n,p)

The various probabilities of the Binomial random variable are ${}^nC_0\ p0\ q^{n-0}$, ${}^nC1\ p^1\ q^{n-1}$, ${}^nC2\ p2\ q^{n-2}$, ….. ${}^nC_r\, p^r\ q^{n-r}$, ………… ${}^nC_n\ p^n\ q^{n-n}$.

We can see that these probabilities are the respective terms of the binomial expansion of $(q+p)^n$. This is the reason for calling the above distribution as Binomial distribution.

E. Assumptions

1. The experiment can be repeated any number of times under uniform conditions.
2. An experiment consists of 'n' trials. Each trial is independent. It means that outcome of any one trial does not affect the outcome of any other trial.
3. In each trial there are only two outcomes namely success and failure.
4. Probability of success is denoted as 'p' and probability of failure is denoted as 'q'. Probability for success in each trial remains the same.

F. Properties

1. Binomial distribution is a discrete probability distribution.
2. The parameters of the distribution are 'n' and 'p'.
3. Arithmetic mean of Binomial distribution is the product of the parameters 'n' and 'p' i. e μ=np.
4. Standard deviation of the distribution is given as,

$$\sigma = \sqrt{npq}$$
$$\text{Variance} = npq$$

5. First four central moments of the distribution are estimated as,
   - $\mu_1 = 0$
   - $\mu_2 = npq$
   - $\mu_3 = npq(q-p)$
   - $\mu_4 = 3n^2p^2q^2 + npq(1-6pq)$
6. Measure of Skewness $\beta_1 = \frac{(q-p)2}{npq}$

The binomial distribution is a positively skewed distribution.

7. Measure of Kurtosis , $\beta_2 = 3+\frac{1-6pq}{npq}$
8. It has one or two modal values. When (n+1)p is an integer there are two modes. They are, (n+1)p and {(n+1)p}-1. When (n+1)p is not an integer, mode is the integral part of (n+1)p .
9. Mean of the Binomial distribution increases as 'n' increases with 'p' remaining constant.
10. The shape and location of binomial distribution changes as 'p' changes for a given 'n'.
11. Moment Generating Function of the Binomial distribution is $(q+pe^t)^n$
12. If X is a Binomially distributed random variable with parameters n_1 and p and Y is an independent random variable following the Binomial distribution with parameters n_2 and p. Then the additive property states that their sum X+Y is also distributed as binomial with parameters n_1+n_2 and p.

G. *Applications*

- Binomial distribution is often used in Quality Control for estimating probabilities of defective items.
- Used in Radar detection
- Used in estimation of Reliability of systems.
- Used to estimate the number of rounds fire from gun hitting a target.
- Used in problems related to disease of people work in industry.

2. Poisson distribution

The Poisson distribution is an important discrete probability distribution used for modelling natural phenomenon. This distribution was first proposed by the French Mathematician Simon Denis Poisson in 1837.

A. *Definition*

A discrete random variable X is said to follow the Poisson distribution with parameter λ if its probability density function (p. d. f) is given by

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

This distribution is usually denoted by the symbol P (x, λ) or simply P(λ).

3. Result

The Binomial distribution tends to the Poisson distribution when $n\rightarrow\infty$ (number of trial is very large) and $P\rightarrow 0$ (probability of success is very small) such that $np=\lambda$, say is finite.

A. *Properties*

1. Poisson distribution is a discrete probability distribution.
2. It has a single parameter, λ. If we know λ, all the terms of the Poisson distribution can be obtained.
3. Mean and variance of the Poisson distribution are same and is the parameter λ.
4. First four central moments of the distribution are estimated as,
   - $\mu_1 = 0$
   - $\mu_2 = \lambda$
   - $\mu_3 = \lambda$
   - $\mu_4 = 3\lambda+\lambda$
5. Measure of Skewness, $\beta_1=\frac{1}{\lambda}$. Since $\beta_1>0$, Poisson distribution is positively skewed.
6. Measure of Kurtosis, $\beta_2=3+\frac{1}{\lambda}$. Since $\beta_2>3$, Poisson distribution is leptokurtic.
7. Moment generating function of the Poisson distribution is $e^{\lambda(e^t-1)}$.
8. If λ is an integer value, there are two modes for the Poisson distribution- λ and λ-1. Also if λ is not an integer, there is only a single mode for the Poisson distribution i. e. the integral part of λ.
9. If X and Y are two independent Poisson Random variables with parameter λ_1 and λ_2 respectively, then their sum X+Y follows a Poisson distribution with parameter $\lambda_1+\lambda_2$.

4. Applications

Generally, the Poisson distribution is used to describe the number of occurrences of a rare event in a short period. Some examples are given below.

- To count the number of wrong telephone calls receiving in a house in a particular day.
- To count the number of suicides reported in a municipal area in a week.
- To count the number of accidents at a traffic junction in a particular city.
- To count the number of bacteria per unit.

5. Normal distribution

Normal distribution is a theoretical probability distribution that deals with continuous variables. It was first introduced by Abraham De-Moivre in 1733 as limiting case of the Binomial distribution.

A. *Definition*

A continuous random variable X is said to be normally

International Journal of Research in Engineering, Science and Management
Volume-3, Issue-2, February-2020
www.ijresm.com | ISSN (Online): 2581-5792

distributed if it satisfies the following probability density function,

$$f(x)= \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \qquad \text{for } -\infty \leq x \leq \infty$$

6. Result

When n is very large, p and q not very small Binomial distribution tends to Normal with mean np and variance npq. Also even when n is not very large p = q = $\frac{1}{2}$, Binomial can be approximated to Normal.

A. Properties

1. The normal curve is bell shaped, unimodal, symmetric about the ordinate at mean and approaches X-axis at $\pm\infty$.
2. Mean, median and mode coincide at x = μ and σ is the S. D.
3. It has points of inflection at $x = \mu \pm \sigma$.
4. Mean deviation about mean is $\sqrt{\frac{2}{\pi}}\sigma \approx \frac{4}{5}\sigma$ and Quartile deviation is $0.6745\sigma \approx \frac{2}{3}\sigma$.
5. The interval (μ-2σ, μ+2σ) contains 95% of the area and (μ-3σ, μ+3σ) contains 99% of the area under the curve.
6. Moment Generating Function of Normal distribution is $e^{\mu t+\frac{t^2\sigma^2}{2}}$.
7. The ordinate at mean divides the whole area into two equal parts.
8. The total area under the Normal curve is 1. Hence, the area to the right and left of the ordinate x=μ is 0.5 each.
9. No portion of the curve below the X-axis.
10. The normal curve is asymptotic to the base.
11. All odd central moments of Normal distribution is zero.
12. Measure of skewness is $\beta_1=0$.
13. Measure of Kurtosis, $\beta_2=3$. The curve is mesokurtic.

B. Applications

- In sampling theory, Normal distribution is used to estimate the values of parameters of sample statistic.
- In Inferential Statistics, the test of significance is based on the assumption that the parent population from which samples are drawn is Normal distribution.
- Most of the Theoretical distributions can be approximated to Normal distribution.
- Normal distribution is widely used in Statistical Quality Control (S. Q. C) and Industrial experiments.
- Many distributions in social and economic data are considered to be normally distributed. Eg: Age. Height of adult persons, the intelligent test scores of school children etc. follows Normal distribution.

7. Conclusion

This paper presented an overview on theoretical distributions.

References

[1] S. P Gupta & V. K Kapoor, Fundamentals of Mathematical Statistics, Sultan Chand & sons, India, Revised edition 1989.
[2] K. Ramakrisna Pillai, “Standard Distributions”, in Subsidiary Statistics, Sai Publishers, India.
[3] L. R. Potti, “Theoretical Distributions” in Basic Statistics, Yamuna Publications, India.