International Journal of Advanced Education and Research
ISSN: 2455-5746, Impact Factor: RJIF 5.34
www.alleducationjournal.com
Volume 2; Issue 4; July 2017; Page No. 32-34

A beautiful normal distribution and its model

Dr. Kumud Gore – Kherdekar

Principal & Associate Professor in Statistics, Govt. College of Arts & Science, Kile Ark, Aurangabad, Maharashtra, India

Abstract

Statistics is one of the applied sciences or helper science as it is called. In addition to the students of statistics, researchers, students with no statistics background also visit the department for learning statistics. Normal Distribution is the most widely used phenomenon in theory as well as in applications of Statistics. Hence, in 1992 I prepared a Wooden Model of Normal Distribution for better understanding of the concepts of the same. Since then students are taking benefit of the model.

In 2016, in Pune University. I presented this model in National Level Competition on Innovative Models for Effective Teaching, organized by The National Academy of Sciences, India (Pune Chapter) & Dept. of Chemistry, SPP University, Pune, which was appreciated by many students and faculty members.

The normal distribution is the topic of interest to scientists as well as social scientists.

In this paper we have discussed Random Variable, Probability Distribution and also Normal Distribution, its importance and properties as depicted by The Wooden Model.

This will enable to understand the concepts of random variable and probability distribution especially to non-statisticians. This will further enable them to learn conceptually the Normal Distribution, its importance and properties.

Keywords: random variable, normal distribution, model for normal distribution

Introduction

Uncertainty is the basis of life and probability is one of the bases of statistics. We define a variable behaving randomly, having different values at different points that are unpredictable, then the variable is a random variable. Though its values are unpredictable in the light of uncertainty, we can assign some probability to the values of the random variable.

For example train arrival time at a particular station is 7:30AM. But it does not arrive exactly at 7:30AM every day at the station. Then the train arrival time at the station is a random variable, whose values may be 7:35AM, 8:00AM, 7:28AM, 7:25AM, 9:00AM… Taking this into consideration this variation in the values, one cannot say at what exact time the train is going to arrive tomorrow at the station. But we can make a probabilistic statement, "The probability that the train will arrive at the right time at the station is, say, 0.90". In simple terms this statement is just similar to, "There is 90% possibility that the train will arrive at the right time at the station".

Fig 1: Wooden Model of Normal Distribution

As we all know, we are making such statements in our day to day life without having any prior knowledge of probability theory. To every value of the random variable we can assign some probability.

The value of random variable along with corresponding probabilities constitutes a probability distribution. For example X is a random variable defined as, X is the grade given to a student in tenth class of 50 students. Then X takes values $0, 1, 2, \dots, 10$ and we have

X_i :	**0**	**1**	**2**	**3**	…	**10**
$P(X_i)$:	$P(X=0)$	$P(X=1)$	$P(X=2)$	$P(x=3)$	…	$P(X=10)$

Constitutes probability distribution of X.

Many times $P(X_i)$ follows some mathematical law. That is the probabilities can be defined in general terms, by some mathematical expression. For example $P(X) = \frac{1}{x}$ then by substituting the possible values of one can obtain probabilities corresponding to different values of random variable X.

Accordingly there are some standard probability laws, which are called as standard probability distributions. Some of the known standard probability distributions are, Binomial Probability Distribution, Poisson Probability Distribution, Normal Probability Distribution etc.

Normal probability distribution or Normal probability law or normal distribution has its own place in the world of theory of statistics as well as in applied statistics.

The normal distribution is discovered by De Moivre (1667 – 1754) and is also attributed to Carl Friedrich Gauss (1777 – 1855). It is often called as Gaussian distribution in the honor of Carl Friedrich Gauss.

Quetelet (1796 – 1874) found that the body lengths of soldiers of an age group apparently follow normal distribution. For him it was the distribution of error that nature made in the reproduction of an ideal average man. Hence it is also termed as an error law.

Importance of the distribution

Normal distribution is called as cornerstone of statistical analysis and drawing of inferences. It is the boon to statistics theory and applications.

W.J. Youden of national Bureau of Standards describe the importance of normal distribution artistically as follows:

THE NORMAL
LAW OF ORDER
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALISATIONS OF NATURAL
PHILOSOPHY. IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES,
IN THE PHYSICAL AND SOCIAL SCIENCES
AND IN MEDICINE, AGRICULTURE AND
ENGINEERING. IT IS AN INDESPENSABLE TOOL FOR
THE ANALYSIS AND THE INTERPRETATION OF THE
BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT.

Fig 2: Artistic description of Normal Distribution

Lipman says, "Everybody believes in the law of errors that is the normal curve. The experimenters believe because they think it is the mathematical theorem and mathematicians believe because they think it is an experimental fact".

The real beauty of Normal distribution is due to central limit theorem. The central limit theorem says that for relatively large sample size, the variable $\bar{x}$ is approximately normally distributed, regardless of the distribution of the variable under consideration. The approximation becomes better with increasing sample size. Generally sample size, n, 30 or greater $(n \geq 30)$ is considered to be large enough.

Eventually central limit theorem permits to apply Normal Distribution to sample tests which are widely used in application, in life sciences, in mathematical sciences, social sciences and anywhere else.

In addition, normal distribution has its applications in industry in quality control and other departments. Most of the standard probability distributions, Binomial, Poisson etc tend to normal distribution for large n.

The complete theory of small sample tests or exact sample tests viz. $t, F, \chi^2 etc. tests$ is based on the assumptions that the parent population from which the samples have been drawnfollows Normal distribution.

Properties of Normal Distribution as depicted by the Wooden Model:

1. The normal curve is bell shaped. These are some examples of bell shaped curve which can be seen in wooden model. An interesting example is, let dry sand run through a funnel into the space between two parallel vertically placed glass walls, an approximately normal distribution will appear in the glass panes.
2. The curve is symmetric about mean. This property also can be viewed in Wooden Model. The curve is symmetric about mean implies that the shape of the curve that is observed on one side (say left) is the mirror image of the shape of the curve on the other side (say right) of the mean.
3. Since the curve is symmetric all odd ordered moments are zero. This is the property of symmetric distribution. Similarly since the curve is symmetric, the skewness, which is lack of symmetry, is zero.
4. Mean, Mode and Median of normal distribution are equal.
5. The curve is normal or mesokurtic having kurtosis equal to 3.

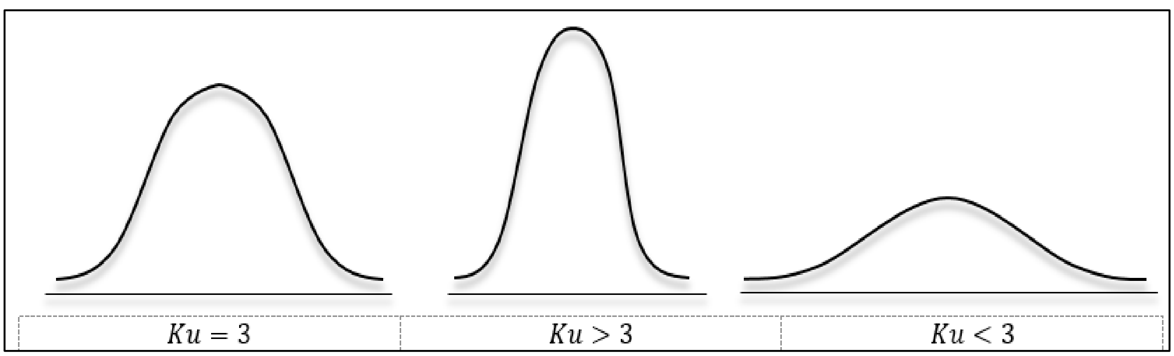

Fig 3: Graphs showing kurtosis or peakedness of the Normal Curves

6. X-axis is asymptote to the curve, meaning that the normal curve, though it tapers on both sides & goes very near to the X-axis, never touches the X-axis.
7. Theoretically, the range of normal curve is∞ $to + \infty$, but practically the range is $X - 3\sigma \; to \; X + 3\sigma$ and the magnitude of range is 6σ.
8. The normal distribution is unimodal.
9. The most important area property of the Normal Distribution can be well depicted by the Wooden Model.

The area under the curve is shown in red color is $0.6826 \; or$ 68.26% within the limits$(X - \sigma) \; to \; (X + \sigma)$. The area under curve shown in red and green colors added together is $0.9544 \; or$ 95.44% within 2σ limits or from $(X - 2\sigma) \; to \; (X + 2\sigma)$

The total area under the curve (three colors together) is $0.9973 \; or$ 99.73% within 3σ limits or from$(X - 3\sigma) \; to \; (X + 3\sigma)$. This means the probability that the value of normal variable falls outside 3σ limits is 0.0027 which is very low. On the contrary nearly 100% items fall under 3σ limits.

The value of random variable X increases, computation of $P(X)$ becomes very difficult in all probability distributions. But due to central limit theorem, the distributions tend to normal distribution and we can find normal probabilities which are readily available in Biometrika Tables.

These are some of the properties which can be explained through the Wooden Model for the benefit of the learners. There are many more properties which are useful in theory and application of normal distribution, which are mathematical and cannot be explained through the model and hence are beyond the scope of this paper.
In short due to the applications in theory and practical and the support gained by central limit theorem, normal distribution has become a sensation in statistics.

References

1. Applied Statistics A handbook of techniques II edition, Springer – Verlag publication
2. Elhance DN, Veena Elhance BM. Agarwal, Kitab Mahal Publishers
3. Neil Weiss, Introductory Statistics, VII Edition, Pearsons Education.
4. Gupta SC. Fundamentals of Statistics, VII Edition, Himalaya Publishing House.