International Journal of Applied Mathematics
& Statistical Sciences (IJAMSS)
ISSN(P): 2319-3972; ISSN(E): 2319-3980
Vol. 7, Issue 6, Oct - Nov 2018; 1-10
© IASET

International Academy of Science, Engineering and Technology
Connecting Researchers; Nurturing Innovations

JACOB BERNOULLI AND HIS WORKS ON PROBABILITY AND LAW OF LARGE NUMBERS: A HISTORICAL SEARCH

AkhilGoswami[1] & Gautam Choudhury[2]

[1]Research Scholar, Department of Mathematics, Gauhati University, Assam, India

[2]Instituteof Advanced Studies for Science and Technology, Boragaon, Assam, India

ABSTRACT

Jacob Bernoulli is a very well-known Swiss Mathematician. Around 328 years ago, Jacob Bernoulli worked out on the game of chances (from 1684 to 1689) which forms the background of the theory of probability. His book 'Ars Conjectandi' was published by his nephew, another great Mathematician Nicolaus I Bernoulli and brother Johann Bernoulli in 1713. The aim of this article is too high light Jacob Bernoulli's initial contributions to the theory of probability and Law of Large Numbers.

KEYWORDS: *Mathematician, Jacob Bernoulli, Nicolaus*

Article History

Received:* *11 Sep2018* | *Revised:* *17 Sep 2018* | *Accepted:* *22 Sep 2018

1. INTRODUCTION

Before Jacob Bernoulli started his work, the word "*probability*" did not belong to the mathematical theory. Jacob had to make the situation for the existence and meaningfulness of numerical probabilities. Probability belonged to the world of practical problems and arguments. The Game of chances was on the one side and the probability was on the other side. Jacob wanted to apply the *theory of equity* in the game of chances which ultimately led to the development of probability theory [1, 3, 7, and 10]. According to Jacob Bernoulli, "*if an experiment is repeated a number of times, then the relative frequency with which an event occurs equal to the probability of the event*"[1,2,4].

Jacob Bernoulli was born in Basel, Switzer Landon the 27th of December of 1654 [1]. He was famous as a Mathematician, died on 16 August 1705 AD at Basel, Switzerland. His father Nicolaus Bernoulli was a member of the town council and a magistrate, his mother was MargarethaSchönauer came from a family of bankers and councilors. Jacob was also known as James Bernoulli or Jacques Bernoulli. His siblings were: Nicolaus I Bernoulli (1662-1716) was a painter and alderman of Basel and another one Johann Bernoulli (1667-1748, also known as Jean Bernoulli) was a great mathematician and early adopter of infinitesimal calculus [2, 5,9]. Jacob Bernoulli's parents forced him to study Philosophy and Theology during his early days. Though he fulfilled their demands, Jacob did not reduce his interest in Mathematics. So, he studied the subject of his interest simultaneously. He graduated from the University of Basel.In1671 he received master's degree in Philosophy and five years hence, in 1676, he was awarded a *licentiate* in Theology. From 1676 to 1682, Jacob traveled all around Europe and collected the latest information on discoveries made by Scientists in the field of Mathematics and Astronomy. His first halt was at Geneva where he took up a job of a tutor. Later on, he moved to France, where he spent a couple of years studying with the followers of Descartes. In 1681 Jacob's next stop was at the Netherlands where he met a number of mathematicians including Hudde. Jacob also traveled to England where he met Boyle and Hooke. These travels widened his knowledge [2, 7,9, 13].

In 1683, Jacob joined the University of Basel in Switzerland as a lecturer in Mechanics. During his teaching job, he also studied the works of Descartes '*Géométrie'*. Jacob started publishing his writings in the journal '*Act an Eruditorum*' during that time[2, 5, 12].

Jacob Bernoulli may be called the initiator of the Bernoulli family's Mathematical dynasty [1, 2, 5]. Jacob was bright and intelligent right from his early childhood. Research concepts and works of Bernoulli family brought a revolution in the field of Mathematics.

Jacob Bernoulli was asked by his brother Johann Bernoullitoteach him Mathematics. In 1684, they started studying calculus presented by Leibniz in his paper published in '*ActaEruditorum*'. The cordial relationship between the two brothers did not last long as both got transformed from colleagues to competitors. Jacob and Johann made great contributions towards Mathematics which are extremely vital and important. The jealousy and enmity between the two brothers resulted in a rift in the relationship and finally, in 1695, Johann moved to Holland [2, 5, 12].

Bernoulli family produced a number of famous Mathematicians and Scientists in the 18thCentury [1, 2, 5]. Nicolaus II Bernoulli (1687- 1759) was a son of Nicolaus I Bernoulli and he was a Mathematician. Nicolaus III Bernoulli (1695-1726) was the son of Johann Bernoulli, who worked on curves, differential equations, probability and originator of the St. Petersburg paradox. Daniel Bernoulli (1700- 1782) was a son of Johann Bernoulli and he was the developer of Bernoulli's principle and originator of the concept of *expected utility* for resolving the St. Petersburg paradox. Johann II Bernoulli(1710-1790, also known as Jean) was the son of Johann Bernoulli and he was a Mathematician and Physicist. Johann III Bernoulli(1746-1807, also known as Jean) was a son of Johann II Bernoulli and he was an Astronomer, Geographer, and a Mathematician. Jacob II Bernoulli (1759-1789, also known as Jacques) was a son of Johann II Bernoulli and he was a Physicist and a Mathematician. The family tree of Bernoulli's is shown below :

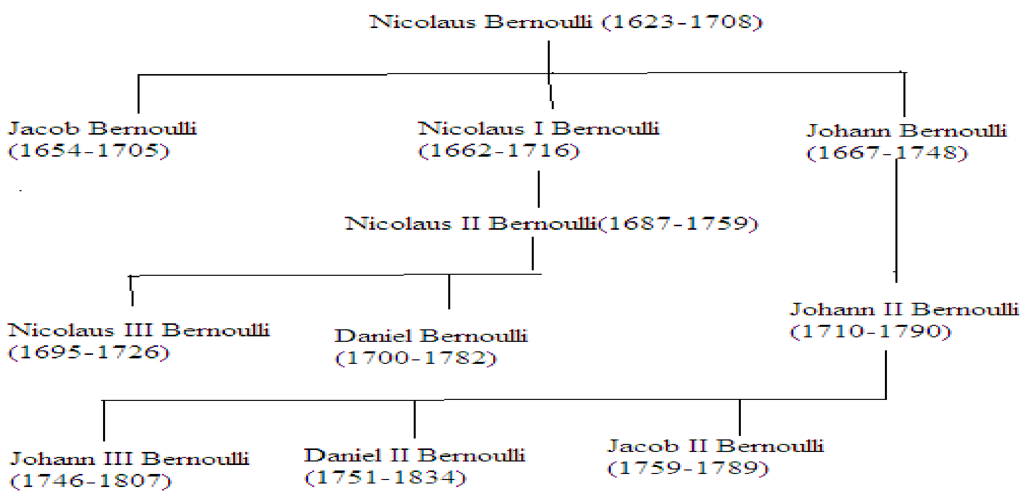

Figure 1

In 1684 Jacob Bernoulli married Judith Stupnas. The couple was blessed with a son and a daughter but both the children did not go to the field of Mathematics.

The Remaining Portion of the Paper Proceeds as Follows

In section 2, we have given a brief review of the terminologies required for further discussion of this paper. Section 3 is a discussion about Jacob Bernoulli's contribution towards the theory of Probability. In section 4 we have given a review about how Jacob Bernoulli contributed towards the development of Law of Large Numbers. In Section 5we have shown the importance of Bernoulli's Law of Large Numbers and section 6 consists of the conclusion of our study in this regard.

2. TERMINOLOGIES FOR FURTHER DISCUSSION

2.1. Mathematical Model of Probability

The *Mathematical model* for probability has three major components and they are

The *sample space S, the family of events E, and the probability measure P.*

Example: Suppose a coin is tossed thrice then

$S = \{HHH, THH, HTH, HHT, TTH, .THT, HTT, TTT\}$

Suppose we are looking for the event of occurrence of three heads in three tosses of the coin, then $E = \{HHH\}$ and $P(\text{Three Heads}) = \frac{1}{8}$

2.2. Bernoulli Trial and Experiment:

A *Bernoulli trial* ends with one of the two outcomes i.e. either a success or a failure. If success occurs with probability p then failure occurs with probability $q = 1 - p$.*Bernoulli experiment* satisfies the following conditions:

- The number of trials is *finite* and *fixed.*
- In each trial, there are two possible outcomes: *success or failure*.
- Trials are *independent* i.e. the probability of the outcome of one trial does not affect the outcome of the other trial.
- The probability of *success* in each trial is equal to p (say) and the probability of *failure* in each trial is equal to $q = 1 - p$.

Example: A card is drawn from a pack of 52 cards. If we consider getting of an ace is a *success* and getting any other card as the *failure* then the probability of getting a success is $\frac{4}{52}$ and meeting with a failure is $\frac{48}{52}$ as the trial is a *Bernoulli trial*. Before a second card is drawn, the card drawn in the first draw is replaced back. If the card is not replaced, we cannot have a *Bernoulli experiment.*

2.3. Empirical Probability

The *empirical probability*, *relative frequency* or *experimental probability* of an event is the ratio of the number of outcomes in which a specified event occurs to the total number of trials, *not in a theoretical sample space but in an actual experiment*.

Example: Suppose a coin was tossed 4 times and it results in a head only once. Then the *empirical probability* of getting ahead if the coin is tossed 4 times is $\frac{1}{4}$.

2.4. Theoretical Probability

The ratio of a number of favorable outcomes to the total number of possible outcomes is called *theoretical probability*.

Example: Suppose a coin was tossed 4 times and it results in a head only once. Then the *empirical probability* of getting ahead if the coin is tossed again is $\frac{1}{4}$ and the *theoretical probability* is $\frac{1}{2}$ because in this case, we are not concerned about how many times the coin is already tossed but refer to the probability what would happen in case of the single toss of a coin.

2.5. Statistical Experiment

A *statistical experiment* is an experiment which has more than one possible outcome. A coin toss has all the attributes of a *statistical experiment*. We can specify each possible outcome in advance i.e. head or tail and there is an element of chance. We cannot know the outcome until we actually flip the coin.

2.6. Relative Frequency

A relative frequency is the fraction of times an answer occurs. The relative frequencies, divide each frequency by the total number of outcomes in the sample.

If A is an event and if A occurs r times in n experiments, then the ratio $\frac{r}{n}$ is called the *relative frequency or the frequency ratio* of the event A. The ratio will change with the change in the number n. If the frequency ratio $\frac{r}{n}$ for the event A is observed for increasing values of n, it is generally found that there is a tendency for the ratio to be more or less constant for large values of n. this tendency towards regularity is called a *statistical regularity* [15].

Example: Suppose a total number of pets in a firm is 100 as given below**:**

Pet	Number	Relative Frequency
Cat	15	0.15
Cow	45	0.45
Dog	15	0.15
Horse	25	0.25

2.7. Statistical Definition of Probability

If $\frac{r}{n}$ is the relative frequency or frequency ratio of an event A connected with a random experiment, then the limiting value of the ratio $\frac{r}{n}$ as n increases indefinitely is called the probability of the event A [15].

2.8. Frequency Interpretation of Probability

The *frequency interpretation of probability* is fairly self-explanatory as it defines probability as a limiting frequency. If we throw a fair die often enough, then we will eventually find that each number comes up about a sixth of the total time. The longer we continue to throw the die, the nearer the result will come to the ideal value of $\frac{1}{6}$ for each number, the probability of an event is then defined by the relative frequency, as the throws continue to an idealized infinity.

3. PRELIMINARY WORKS OF JACOB BERNOULLI ON PROBABILITY

Jacob Bernoulli discussed probability and the nature of probability in his book entitled *'Ars Conjectandi'* as:

"*...probability is a degree of certainty and differs from the absolute certainty as a part differs from the whole. For example, the whole and absolute certainty which we denote by the letter 'a' or by the unity symbol '1', is supposed to consist of five probabilities or parts, three of which stand for the existence or future existence of some event, the remaining two standing against its existence or future existence, this event is said to have* $\frac{3}{5}a$ *or* $\frac{3}{5}$ *certainty*[9]."

Bernoulli's above statement in simple language can be viewed as that if we consider an urn containing 5 balls out of which 3 balls are red and 2 balls are blue, then the *existence of the event of drawing a red ball* will be $\frac{3}{5}$ and the *existence of the event of drawing a blue ball* will be $\frac{2}{5}$.

Bernoulli viewed that probability is a consequence of uncertainty which is evident from his following statement :

"*…those data which are supposed to determine later events have nevertheless not been learned well enough by us*."[9].

For illustration, let us consider a coin tossing experiment where there are only two possible outcomes viz. Head and Tail. Each outcome of such an experiment is a Bernoulli trial. Now if a coin is tossed an infinite number of times until we get a Head, then we can illustrate it with the help of the following figure:1

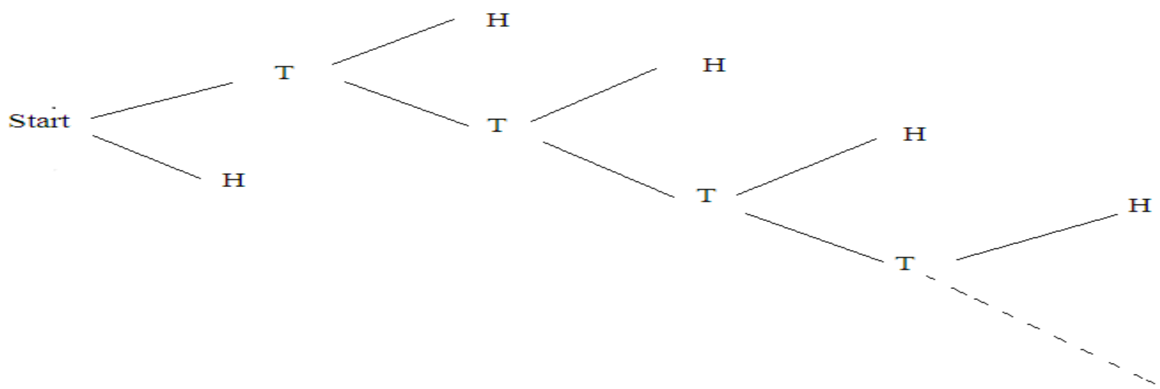

Figure 2

The above incident which he expressed in his statement as "…*those data which are supposed to determine later events have nevertheless not been learned well enough by us*. "motivated Jacob Bernoulli to think in the line which at present times we know as *Law of Large Numbers* as it becomes evident from his subsequent statement which is mentioned below.

In Chapter IV of Part IV of *Arts Conjectandi*, Bernoulli wrote: "*Something further must be contemplated here which perhaps no one has thought of about till now. It certainly remains to be inquired whether after the number of observations has been increased, the probability is increased of attaining the true ratio between the numbers of cases in which some event can happen and in which it cannot happen, so that the probability finally exceeds any given degree of certainty*..."[9].To overcome this problem Jacob Bernoulli made the following statement :

"*Let the number of favorable cases to the number of unfavorable cases be exactly or nearly* $\frac{a}{b}$, *therefore to all the cases as* $\frac{a}{a+b}=\frac{a}{t}$ *where* $a+b=t$. *This ratio is between* $\frac{a+1}{t}$ *and* $\frac{a-1}{t}$. *As many observations are taken, it becomes more probable arbitrarily often that the ratio of favorable to all observations lies in the range with boundaries* $\frac{a+1}{t}$ *and* $\frac{a-1}{t}$."

i.e. $\frac{a-1}{t}<\frac{a}{a+b}<\frac{a+1}{t}$ (1)

An explanation of the Bernoulli's above proposition can be clarified with the help of the following example:

Suppose a box contains 5 red balls and 3 blue balls and drawing of a red ball is a favorable case for us. Now, the ratio of red balls to all types of balls are $\frac{5}{8}=0.625$. In accordance with Bernoulli's symbolism, we can take $a=5, b=3$ and $t=5+3=8$ then

$$\frac{a}{a+b}=\frac{5}{8}=0.625,\ \frac{a+1}{t}=\frac{6}{8}=0.75,\quad \frac{a-1}{t}=\frac{4}{8}=\frac{1}{2}=0.5$$

which verifies that $\frac{a}{a+b}$ lies between $\frac{a-1}{t}$ and $\frac{a+1}{t}$

i.e. $\frac{a-1}{t}<\frac{a}{a+b}<\frac{a+1}{t}$

In the above example, it is important to be note that

$\frac{a+1}{t}$ - $\frac{a}{a+b}$ = 0.75-0.625 = 0.125,and

$\frac{a}{a+b}$ - $\frac{a-1}{t}$ = 0.625-0.5 = 0.125.

which shows that $\frac{a}{a+b}$ not only lie between $\frac{a-1}{t}$ and $\frac{a+1}{t}$ but in fact it is the mean of the two boundaries.

Review of the literature shows that Bernoulli established the following proposition without applying any mathematical analysis [6,8].

4.1. Jacob Bernoulli's Proposition and Law of Large Numbers:

In 1682 Jacob Bernoulli started to give lectures privately in Physics at different universities. In 1687 he was appointed as professor of Mathematics at the University of Basel. During that period he started to apply Leibniz's infinitesimal calculus and carried out further investigation on the Law of Large Number.

Basically, the Law of Large Number states that independent repetition of an experiment average over long time horizons in arithmetic mean which is not generated randomly but is a well-specified deterministic value. This reflects the intuition that a random experiment averages if it is repeated sufficiently often.

The Law of Large Numbers formulated in modern mathematical language reads as:

Let us suppose that $X_1, X_2, \ldots$ is a sequence of uncorrelated and identically distributed random variables having finite mean $E(X_i) = \mu$.

If we now define, $S_n = \sum_{k=1}^{n} X_k$, then for every $\varepsilon > 0$ we have

$$\lim_{n\to\infty} P_r\left(\left|\frac{S_n}{n} - \mu\right| \geq \varepsilon\right) = 0. \tag{1}$$

Bernoulli proved (1) for independent and identically distributed(i.i.d) random variables $X_1, X_2, \ldots$ taking only values in $\{0,1\}$, having probability mass functions

$$P_r\{X_i = k\} = p^k (1-p)^{1-k} \text{ for } k = 0, 1. \tag{2}$$

We call it a Bernoulli experiment which we already mentioned in 2.2 and in this case we have $\mu = p$.

Nowadays, *Bernoulli Law Of Large Numbers*, (1) is also known as *Weak Law of Large Numbers.*

Historical search shows that in establishing the law of large numbers Bernoulli depended on direct calculations with binomial coefficients only.

Now for simplicity, let us consider the symmetric Bernoulli's case with $p = \frac{1}{2}$ considering '*n*' drawings with replacement, with no further restriction on $\left(0, \frac{1}{2}\right)$.

Then $S_n = X_1 + X_2 + \ldots + X_n$ and the total number of successes in '*n*' i.i.d. Bernoulli experiment having the probability of success $p = \frac{1}{2}$ follows a Binomial distribution with p.m.f. (probability mass functions)

$$P_r\{S_n = k\} = \binom{n}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{n-k}$$ where $k = 0,1,2,\ldots n$

$$Or\, B\left(n,\ k\ ;\ \frac{1}{2}\right) = \binom{n}{k}\left(\frac{1}{2}\right)^n \tag{3}$$

Now for $\varepsilon \in \left(0, \frac{1}{2}\right)$, we have

$$P_r\left(\ \left|\frac{S_n}{n} - p\right| \geq \varepsilon\ \right) = P_r\left(\ \left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \varepsilon\ \right) = 2P_r\left(\ S_n \geq \left(\frac{1}{2} + \varepsilon\right) n\right) \tag{4}$$

By considering the following quotients for $k < n$, Bernoulli obtained

$$\frac{B\left(n, k; \frac{1}{2}\right)}{B\left(n, k+1; \frac{1}{2}\right)} = \frac{k+1}{n-k}$$

from which he concluded that

$$Max_k\ B\left(n, k\,; \frac{1}{2}\right) = B\left(n,\ \frac{n}{2}; \frac{1}{2}\right)$$

The general formulation on Law of a Large Number of expression (1) is taken from *Khinch in*(1929) [16] of which a particular case was derived by Bernoulli as mentioned earlier. In introductory courses of probability theory and statistics, one often proves (1) under a more restrictive assumption of $\{X_i\}$ having finite variance. In this case, the proof easily follows from *Chebyshev's* inequality.

Let S_n be the observed number of successes in '*n*' independent repeated Bernoulli trials with probability '*p*' of success in each trial.

Let $f_n = \frac{S_n}{n}$ denote the relative frequency of success in the '*n*' trials. Then for any positive number ε, no matter how small, it follows that

$$\lim_{n\to\infty} P_r\left(\left|\ f_n - p\right| \geq \varepsilon\right) = 0 \tag{5}$$

Bernoulli proved (5) by a process of tedious calculations of probabilities. However, using *Chebyshev's inequality* we can provide a very simple proof of (1). It is well known to us from the *Chebyshev's inequality,* for every $\varepsilon > 0$

$$P_r\left(\left|\ f_n - \mu\right| \geq \varepsilon\right) \leq \frac{Var(f_n)}{\varepsilon^2} \tag{6}$$

Now by using the fact that the probability law of f_n has the mean $\mu = p$ and variance $Var(f_n) = \frac{p(1-p)}{n}$ consequently, from (6), we have

$$P_r(|f_n - p| \geq \varepsilon) \leq \frac{p(1-p)}{n\varepsilon^2}$$

Now any value of 'p' in the interval $0 \leq p \leq 1$, for symmetric Bernoulli's case

$$p(1-p) \leq \frac{1}{4} \Rightarrow 4p(1-p) - 1 \leq 0$$

Consequently, for $\varepsilon > 0$,

$$P_r(|f_n - p| \geq \varepsilon) \leq \frac{1}{4n\varepsilon^2}$$

And therefore we have $\lim_{n\to\infty} P_r(|f_n - p| \geq \varepsilon) = 0$

4.2 Remark

The Bernoulli Law of Large Number states that to estimate the unknown value of 'p', as an estimate of 'p', the observed relative frequency 'f_n' of successes in 'n' trials can be employed; this estimate becomes perfectly correct as the number of trials becomes infinitely large.

In practice, a finite number of trials are performed. Consequently, the number of trials must be determined, so that with high probability, the observed relative frequency be within a pre assigned distance ε from 'p'.

5. IMPORTANCE OF BERNOULLI'S LAW OF LARGE NUMBERS:

Bernoulli's law of large numbers can explain why and how insurance works [14]. The main argument is that a collection of similar and uncorrelated risks $X_1, X_2, X_3, \ldots$ in an insurance, portfolio $S_n = \sum_{i=1}^{n} X_i$ provides an equal balance within the portfolio that makes the outcome more predictable when the portfolio size 'n' is large. Basically it means that for sufficiently large n there is only a small probability 'p' that the total claim 'S_n' exceeds the threshold $n(\mu + \varepsilon) = n(p + \varepsilon)$ and thus, the bigger the portfolio the smaller the required security margin ε (per risk X_i),so that the total claim S_n remains below $n(\mu + \varepsilon)$,with a very high probability. In fact, this is the foundation of the functioning of insurance system all over the world where the aim of every insurance company is to build sufficiently large and sufficiently homogeneous portfolios which makes the claim "predictable" up to a small shortfall probability.

6. CONCLUSIONS

In this paper, we have tried to unfurl Jacob Bernoulli's initial contribution towards the theory of probability which has now become a very important topic in itself having applications in many scientific disciplines. We have further discussed his initial contributions towards the development of the theory of Law of Large Numbers. Further research regarding the contributions of Jacob Bernoulli will definitely help us to evaluate what a great contribution Jacob Bernoulli made towards the development of probability theory in particular and scientific field in general.

REFERENCES

1. *The Significance of Jacob Bernoulli's ArsConjectandifor the Philosophy of Probability Today, Glenn Shafer, Rutgers University.*
2. *The Bernoullis and the Origin of Probability Theory: Looking back after 300 years Wolfgang Polasek, 26 RESONANCE, August 2000,GENERAL, ARTICLE Anders Hald.*
3. *Direct and indirect influences of Jakob Bernoulli's Ars Conjectandi in 18th century Great Britain, Electronic Journ@l for History of Probability and Statistics, vol 2, $n^0$1b; November/2006 and $n^0$1,june/2006.*
4. *A Tricentenary history of the Law of Large Numbers by EUGENE SENETA, School of Mathematics and Statistics FO7, University of Sydney, NSW 2006, Australia. ArXiv: 1309.6488v1 [math.ST] 25 Sep 2013.*
5. *Tercentenary of Ars Conjectandi(1713),Jacob Bernoulli and the Founding of Mathematical Probability, Edith Dudley Sylla, North Carolina State University, Raleigh, North Carolina, USA.*
6. *Jakob Bernoulli, On the Law of Large Numbers, Translated into English by Oscar Sheyn in Berlin 2005. ISBN 3-938417-14-5.*
7. *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713 to 1935, by Anders Hald. Department of applied mathematics and statistics university of Copenhagen universitetsparken 5 dk-2100 copenhagen/ 2004 /ISBN 87-7834-628-2.*
8. *Theory of Probability, A Historical Essay by Oscar Sheyn in Second revised and enlarged edition\ Berlin, 2009,© Oscar Sheynin 2009ISBN 3-938417-88-9.*
9. *An Historical Survey of the Development of Probability and Statistics based on the Appearance of Fundamental Concepts, Submitted by Daniel McFadyen, September 8, 2003.*
10. *The Early Development of Mathematical Probability by Glenn Shafer.*
11. *The Bernoulli process and discrete distributions Math 217 Probability and Statistics, Prof. D. Joyce, fall 2014.*
12. *The Famous people, Society for recognition of famous people. (From Internet)*
13. *Introduction to Probability by Charles M. Grinstead, Swarthmore College and J. Laurie Snell, Dartmouth College.*
14. *Bernoulli's Law of Large Numbers by Erwin Bolthausen and Mario V. Wuthrich.*
15. *Statistical Methods, an Introductory Text, by J. Medhi, Emeritus Professor of Statistics, Gauhati University, ISBN: 978-81-224-1957-3.*
16. *Khinchin, A. Sur la loi des grandsnombres. Complete rendus de I'Acaé me des Sciences 189 (1929), 477-479.*