

Synopsis of the Ph.D. Thesis

Title: A Statistical Approach to Gödel's Theorems With Reference to Central Limit Property

Research Scholar: Binoy Jacob

Research Guide: Dr.K.K.Jose

1 Introduction

In this research work, we are making a statistical appropriation of Gödel's theorems with specific reference to Gödel's incompleteness theorem in mathematics and the Central Limit Theorem in statistics. Gödel's theorems, comprised of the completeness theorem and the incompleteness theorem, are named after Kurt Friedrich Gödel (1906-1978), an Austrian logician, mathematician and philosopher. The question of incompleteness and uncertainty has been an area of research for several centuries. It was in the context of science claiming absolute certainty and exact predictability that Gödel introduced his famous incompleteness theorems. Till his times Mathematics was considered to be error free and the queen of sciences. Gödel's theorems came as a bolt from the blue breaking the till then solid foundations of mathematics. The result of Gödel's theorems had a stunning effect upon logicians, mathematicians and philosophers interested in the foundations of mathematics. The discovery and the proof of the incompleteness theorem led the world to very many interesting results in logic, computer science, mathematics, statistics and other related areas in science.

In view of setting a platform for interfacing the Central Limit Theorem (CLT) in statistics and Gödel's incompleteness theorem in mathematics, we explore the historical developments in statistics with reference to the Central Limit Theorem. The reason for choosing CLT for the statistical appropriation of the Gödel's theorems becomes clear from the historical analysis of CLT and its definitions. It raises several concerns. Will it be possible to make a statistical appropriation of Gödel's incompleteness theorem adequately using CLT? Is 'statistical completeness' possible? What is the limiting behavior of CLT? Is there any fuzziness in Central Limit Property? How does fuzziness differ from randomness? From a critical scrutiny of the Central Limit property we understand that there is deficiency in statistical certainty and precision. To enhance its effectiveness, probability theory draws techniques from fuzzy logic.

Every theoretical system is built upon axioms. Apart from the ontological similarity of CLT's axiomatic basis, the independent nature of the identically distributed random variables in the Central Limit Theorem and their convergence to normal distribution draw parallel to *consistency* in Gödel's theorems. *Consistency* means, given the axioms and the derivation rules, one can never derive two contradictory propositions/conclusions. The 'stable' and consistent features of the distribution are substantiated by the Law of large Numbers. Similarly, *completeness* refers to provability of a true statement within the system. We try to examine whether the identically distributed independent random variables converge to a limit (if there exists a limit) satisfying the given conditions of the CLT, if done, we say that the theorem is *complete*, otherwise *incomplete*. Though there is no guarantee of converging to a limit, the paper explores the possibility of statistical completeness as an application of Gödel's theorems that it is inherently necessary to leave a system to prove a true statement within the system or to make the system complete. The hybrid concept of fuzzy random variable is a powerful impetus in our searching for statistical 'completeness'. Bringing fuzziness and randomness under one roof, we introduce fuzzy random variable to attain statistical 'incompleteness' in CLT. We find that statistical 'completeness' is a form of 'lesser incompleteness,' not fully complete. Ultimately the combination of fuzzy logic and

many valued logic opens up the possibility of a new statistical system (*fuzzy-many valued statistical system*) which takes into account the intermediary truth values along with the absolutes and makes CLT “‘more’ lesser incomplete”. We consider two applications of our findings. The first application introduces a new kind of statistical science which is interdisciplinary in nature and the second application portrays the richer possibility of interfacing statistics, philosophy and religion for a better enhancement of the world and humanity.

2 Review

In tracing the historical track of mathematics, specifically of Gödel's theorems, we take into account different stages in the development of mathematics. A detailed historical account is given in Odifreddi (2005). We shall also discuss the contemporary schools of philosophy for drawing the philosophical implications of Gödel's theorems. Jacquette (2002) and Aichtner (2011) provide with the different phases in the history of the philosophy of mathematics. The possibility of interdisciplinarity in mathematics is highlighted in Kozhamthadam (2003), Freedman (2005), Brechner (2006), Boas (2006), and Bovell (2010). Gödel's theorems consist of completeness theorem and incompleteness theorem. Nagel and Newman (2002) give a comprehensive explanation of Gödel's theorems with proofs and definitions. For an authentic understanding of Gödel's theorems we define logic, different kinds of logic, and the properties of logic referring to Hawking (2005) and Borchert (2006). The completeness theorem is first proved by Gödel in 1929. According to Gödel, a deductive system of first-order predicate calculus is “complete” in the sense that no additional inference rules are required to prove all the logically valid formulas. It is realized that Gödel's completeness theorem is a fundamental theorem in mathematical logic that establishes a correspondence between semantic truth and syntactic provability in first-order logic. A detailed explanation is given in Benacerraf (1967). Gödel's incompleteness theorem deals with axiomatic systems. The idea behind Gödel's proof is descriptively portrayed in Srivastava (2001, 2007) and Jacob (2005). The incompleteness theorem claims that if ‘ T ’ is a consistent system, then ‘ T ’ cannot prove its own consistency. It means that any adequate axiomatizable theory is incomplete. In any consistent axiomatizable theory, the consistency of the system is not provable in the system. For details one can refer Jose (2006, 2007, 2008, 2011) and Jacob (2008 a).

The historical developments in statistics can be divided into three important stages: empirical stage, comparative stage and modern stage which contributed various theories and methods in statistics. These stages are explained in Nair et al. (1999). The historical phases in the development of statistics reveal that statistics is not an autonomous discipline, but has close association with the philosophy of science. For details one may refer Stigler (1992) and Hald (1998). The interdisciplinary nature of statistics, especially its association with mathematics and philosophy, substantiates our attempt to make a statistical approach to Gödel's theorems and also their implications to science in general and statistics in particular. The CLT is one of the most remarkable results in the theory of probability. It states that the sum of a sufficiently large number of identically distributed independent random variables each with finite mean and variance will be approximately normally distributed. If X_1, X_2, \dots, X_n be a sequence of n independent and identically distributed (i.i.d.) random variables each having finite mean and variance., then the Central Limit theorem states that as the sample size n increases, the distribution of the sample average of these random variables approaches the normal distribution irrespective of the shape of the original distribution. The proof and definition of CLT is given in Adams (1955) and Armoe (1966). In its simplest form, the theorem states that the sum of a large number of independent observations from the same distribution has, under certain general conditions, an approximate

normal distribution. Moreover, the approximation steadily improves as the number of observations increases. The central limit property is shown for both independent sequences and dependent sequences. Coulon-Prieur (2000) and Fisher (2011) explain different types of Central Limit Theorem. The proof of CLT and its limiting behavior is also portrayed using Fourier's Transform and Taylor's Series. These proofs can be seen in Jumarie (1999). Statisticians have further extended the concept of CLT to domains of attraction which will further extend normality to any statistical distribution provided certain conditions are satisfied. One may get a vivid idea of it from the work of Geluk and Hann (2000).

The convergence of the central limit theorem emerges from probability theory. The deficiency of the probability theory in statistical precision and accuracy makes the CLT 'incomplete'. First of all, CLT justifies the approximation of large sample statistics to the normal distribution only in controlled experiments. Secondly, since real-world quantities are often the balanced sum of many unobserved random events, this theorem provides a partial explanation for the prevalence of the normal probability distribution. In that sense, the theorem is inherently vague and fuzzy. Most often we sideline or do away with fuzziness. In fact both randomness and fuzziness together make the Central Limit theorem incomplete, to say the theorem is statistically incomplete. Statistical probabilities take into account the provability of what is absolutely true and absolutely false, but the truth values in between these absolutes are often neglected. Very often the real life situations are between absolutes. The deficiency of CLT leads us to draw insights from fuzzy logic for statistical completeness. Fuzzy logic differs in a significant way from classical logic by way of extensions, deviations, and variations. In 1965 Lofti A. Zadeh introduced fuzzy subsets with boundaries that are not precise; also introduced membership in a fuzzy subset not as a matter of affirmation or denial but as a matter of degree. A detailed explanation is given in Zadeh (1965, 1975) and Nozer et al. (2004), Mori et al. (2005). The main idea behind a fuzzy set is that of gradual membership to a set without sharp boundary. This idea fits for a better human representation of reality. In a fuzzy set, the degree of membership of an element is expressed by any number from 0 to 1 rather than only the limiting extremes. A fuzzy number in fuzzy logic is a fuzzy set of R with its membership function. Nathan (2004) brings these ideas in fuzzy logic in a systematic and coherent manner.

Fuzzy random variables are introduced as random variables whose values are not real numbers but fuzzy numbers, and subsequently redefined as a particular kind of fuzzy set. Luhandjula (2004 a), Feng et al. (2001) provide an authentic description of fuzzy random variables. A fuzzy random variable is nothing but a fuzzy-valued function subject to some measurability conditions which accounts adequately for both randomness and fuzziness. The blend of randomness and fuzziness is beautifully summarized by Luhandjula (2004 b). It is shown in Proske and Puri (2001) that a sum or mean of independent fuzzy random variables converges in the limit to a fuzzy Gaussian random variable. Since it provides a fuzzy analogue of the Central Limit theorem of classical probability theory, we say it attains statistical 'completeness' as we envisaged. However, it raises further questions and leads to the inherent limitations in fuzzy logic. One can refer the deficiencies of fuzzy logic in Haack (1996) and Mori et al. (2005). Taking these limitations in fuzzy logic along with the deficiencies in probability theory, we still try for attaining completeness. Now, we rely on many valued logic, specifically four valued logic. A detailed account of multi-valued logic is seen in Priest (2008) and Ord (2012). The new statistical system, based on fuzzy-many valued random variables, will be 'more close' to completeness, yet not fully complete because we cannot do away with the inherent fuzziness/incompleteness in any statistical system. We conclude that statistical 'completeness' is a form of 'lesser incompleteness.'

The applications of our findings call for integral and interdisciplinary understanding of science in general and statistics in particular. The historical synthesis of the roots of science reveals how science has lost its integral nature, especially in the modern era, corresponding to the emergence of the dualistic world-view in philosophy. It is scientifically and meticulously narrated by Capra (1992). However, contemporary science, very specially Gödel's incompleteness theorem, the theory of relativity, the theory of chaos, quantum mechanics etc., takes us to a realm in which the dichotomized world-views are blurred. For details one may refer Ziaeddin and Abram (1999), Pandikkattu (2004) and Jacob (2008 b). In this view, mathematics, statistics, physics, biology, quantum mechanics etc. are not isolated entities in themselves rather relational and complementary. It affirms the interdisciplinary paradigm in contemporary science and the emergence of such a paradigm substantiates the statistical appropriation of Gödel's theorems. One would get a clear picture of the interdisciplinarity dimension of statistics in Hayashi (2009). The second application seeks for interdisciplinarity in science-religion interface. Applying the findings of the statistical appropriation of Gödel's incompleteness theorem we show that neither science nor religion arrives at certainty/completeness. Both science and religion provide a partial description of a part of reality. This concept is brilliantly exposed by Wolfram (2002), Pamplany (2005), and Russell (2008), Taotao (2011). The reality of incompleteness in both science and religion should be a launching pad for both these disciplines to come together to forge a more dynamic partnership in interfacing the empirical world of science with the metaphysical world of religion and also to work for the betterment and enrichment of the world and the whole of the cosmos.

3 Objectives of the Work

The main objective of this work is to study the possibility of making a statistical appropriation of Gödel's theorems with reference to central limit property. It opens up wider possibilities in different areas of statistics, mathematics, philosophy and religion. The present study has been undertaken with the following specific objectives:

- To make a critical enquiry into the possibility of statistical completeness in central limit property appropriating Gödel's theorems and using fuzzy logic and many valued logic.
- To affirm the interdisciplinary features and possibilities of statistical sciences.
- To explore the concrete possibilities of interfacing statistics with mathematics, philosophy and religion.
- To delineate the deficiency of probability theory (Central Limit Theorem) in accuracy and precision and to address the fuzziness inherent in statistical theorems.
- To foster science-religion dialogue on a statistical platform.
- To articulate the foundations for a new kind of statistical science that is integral in nature.
- To introduce a new statistical system which is the combination of fuzzy logic and many valued logic in attaining more 'completeness.'

4 Summary

The thesis is divided into 7 chapters. Chapter 1 presents the introduction and summary of the thesis. Chapter 2 is a brief survey of the history of mathematics in view of placing Gödel's

Theorems within the historical trajectory of mathematics. The historical analysis of the development of mathematics and the philosophy of mathematics leads to the possibility of applying mathematics into various disciplines and the interdisciplinary dimension of statistics substantiates our attempt to make a statistical approach to Gödel's incompleteness theorem. Having placed Gödel's theorems within the historical and philosophical trajectory of mathematical sciences, in chapter 3 we deal with the statement and proof of the theorems. The completeness theorem is a central property of first-order logic that does not hold for higher-order logic. It is possible to produce sound deductive systems for higher-order logic, but no such system can be complete. Through his incompleteness theorem, Gödel exhibited the inherent limitations of axiomatic method and the impossibility of proving the logical consistency of a very large class of deductive systems. For him, an axiomatic formalized consistent system that consists of true statements is essentially incomplete. It means, any adequate consistent mathematical logic is incomplete. Hermeneutically, consistency can be achieved only at the cost of completeness. We can have certainty and accuracy only within a frame work but can not have absolute accuracy. An epistemological appropriation of Gödel's incompleteness theorem shows that science in general and mathematics in particular can not provide us absolute certainty. The mathematical notions of completeness and incompleteness together provide us with insights for a better statistical appropriation of Gödel's theorems.

Chapter 4 proves the validity of using the Central Limit theorem in making a statistical appropriation of the Gödel's theorems. In probability theory, CLT states conditions under which the mean of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed. Its common form requires the random variables to be identically distributed. The law of large numbers as well as the Central Limit theorem is a partial solution to a general problem as far as the limit behavior is concerned. The proof of the Law of Large Numbers affirms the limiting behavior and the stability of the normal distribution. The 'stable' feature of the normal distribution of CLT with its independent and identically distributed random variables can be associated with the 'consistency' dimension in Gödel's theorems. Moreover the Central Limit property in general does not confine to normal distribution alone. It can be extended to any statistical distribution (or any statistical law) such that the sum of the identically distributed random variables or their mean converges to a general domain (or a general Law). Let X_1, X_2, \dots be independent random variables all of them from the same probability distribution with distribution function F . Consider the sequence $S_n = X_1 + X_2 + \dots + X_n$; $n = 1, 2, \dots$ and suppose that for some sequences of norming constants $a_n > 0$ and b_n ($n = 1, 2, \dots$) the sequence $(S_n - a_n)/b_n$ has a non-degenerate limit distribution. It is possible to find the general form of all the possible distributions by way of giving necessary and sufficient conditions on the distribution function F for each of these limit distributions, in order that $(S_n - a_n)/b_n$ converges to that particular distribution function. The limit distributions are called stable distributions and the set of distribution functions F such that $(S_n - a_n)/b_n$ converges to a particular stable distribution is called its domain of attraction. The stable feature of the normal distribution of CLT and its capacity to extend to a general domain of attraction irrespective of any distribution, prove the validity of using the Central Limit theorem in making a statistical appropriation of the Gödel's theorems.

Chapter 5 aims at attaining statistical 'completeness' using Fuzzy Random Variable. Analyzing the Central limit property we understand that randomness and fuzziness co-occur in the central limit theorem. The fact that the Central Limit theorem (in its common form) requires identically distributed random variables affirms the randomness of the theorem. Fuzziness and randomness differ conceptually and theoretically. The sense of uncertainty in probability theory

revolves around making a prediction or expectation of a future event based on some known facts. The sense of fuzziness is not the uncertainty of expectation (the event occurrence), but uncertainty resulting from the imprecision of meaning of a concept (the very meaning of the event itself/event ambiguity). This is a serious deficiency since much of human knowledge consists of propositions that in one way or another are partially certain and /or partially possible and /or partially true. Probability theory by itself is not sufficient for dealing with uncertainty and imprecision in the real world. To enhance its effectiveness, probability theory needs an infusion of concepts and techniques drawn from fuzzy logic. It is shown that a sum or mean of independent fuzzy random variables converges in the limit to a fuzzy Gaussian random variable, thus providing a fuzzy analogue of the Central Limit theorem of classical probability theory. According to the fuzzy central limit theorem the data which are influenced by many small and unrelated random events are normally distributed and hence we say it attains statistical 'completeness' provided the identically distributed independent fuzzy random variables converge to a limit satisfying the given conditions of the Central Limit Theorem.

According to our hypothesis, when the identically distributed independent fuzzy random variables converge to a limit satisfying the given conditions of CLT, we say that the theorem is complete. It raises two questions: Is statistical completeness really possible? Does the concept of 'completeness' completely do away with the notion of incompleteness? In fact the convergence of the identically distributed independent fuzzy random variables to a limit satisfying the given conditions (if and only if the conditions are satisfied) of the Central Limit theorem is only a theoretical plausibility, not a certainty because of the deficiencies in fuzzy logic. First of all fuzzy truth values are much more certain than numerical truth values, but to take truth-values as all fuzzy subsets of $(0,1)$ would be 'much too rich and much too difficult to manipulate'. Naturally it leads to extreme complexities. Secondly fuzzy logic imposes artificial precision. Because it only postpones, and does not eliminate, the need to introduce arbitrary boundaries and in the process one has arbitrarily to fix and define the degree to which a given numerical true value should belong to the 'truthness' of the linguistic truth-value. Thirdly fuzzy logic suffers from linguistic incorrectness. The view of truth on which the motivation for fuzzy logic relies is mistaken. Zadeh's claim that truth is a matter of degree rests on the thesis that the adverbial modifiers that apply to 'true' are those which typically apply to fuzzy predicates. Hence becomes his linguistic evidence wrong. Consequently statistical completeness is a form of 'lesser incompleteness' which is still fuzzy/incomplete in nature.

We still try for attaining completeness relying on four-valued logic and later on a combination of fuzzy logic and four-valued logic. First, our aim is to make it possible to construct different models of logical truth which takes into account the uncertain and the undecidable truth values along with absolutes. Later, we introduce a new system combining fuzzy logic and four-valued logic for comprehending different levels of truth. To this end, we consider truth value gaps (*neither true nor false* sentences) and truth value gluts (both *true and false* sentences). Four valued Logic allows sentences to be *true*, *false*, *both* or *neither*, but not able to represent the absolutes along with truth value gaps and truth value gluts. A more promising approach, which is the combination of fuzzy logic and many-valued logic, is to consider degrees of truth alongside degrees of falsity. Let each truth value be represented by a pair of real numbers from 0 to 1. Suppose, in a truth value (p, q) , p is the degree of truth and q the degree of falsity. Then the point $(1,0)$ represents absolute truth while the point $(0,1)$ represents absolute falsity. The point $(0,0)$ represents the most extreme absence of truth and falsity, whilst the point $(1,1)$ represents the most extreme excess of truth and falsity. This new space of truth values provides us with a great degree of flexibility such as over-defined (truth value glut: both true and false),

under-defined (truth value gap: neither true nor false), well-defined (values in fuzzy logic), more true than false, more false than true, equally true and false, more than half true, and more than half false. These truth values take us beyond those present in fuzzy logic and four valued logic, but the extension is very natural. It should be noted that the combination of fuzzy logic and four-valued logic can address adequately the issues related to statistical incompleteness to a great extent, provided it is applied to fuzzy random variables and consequently to fuzzy Central Limit Theorem. However, we cannot completely do away with incompleteness/fuzziness that is inherently present in any theoretical system, but can attain a 'more lesser incompleteness.'

In chapter 6, we consider two applications of our findings. The first application brings to light that our inability to attain perfect completeness is not a failure of statistical science but a striving force to launch a dynamic and interdisciplinary paradigm in science. On a closer scrutiny, we learn that the fractal geometry, the chaos theory, the uncertainty paradigm of quantum mechanics etc. are mathematically supported and substantiated by Gödelian incompleteness and that paves the way for an integral and interdisciplinary paradigm in science. The interdisciplinary nature of statistical science is further affirmed by the quantum Central Limit theorem. The second application affirms the interdisciplinary paradigm of contemporary science by way of fostering science-religion interface. The statistical appropriation of Gödel's theorems reveal that no scientific system can completely do away with 'incompleteness/fuzziness' inherent in that system. Because truth exists at a meta-level beyond the system and it is necessary to go out of the system to prove the 'truthness' of a true statement or to make the system complete. A metaphorical/analogical appropriation of the incompleteness theorem and the fuzzy Central limit Theorem implies that religion based on axiomatic beliefs is also incomplete. It is further substantiated by the 'Limit Theory' of David Tracy. We show that the experience of limit is common to both scientific and religious language and the experience of limit/incompleteness both in science and religion underline the need to foster science - religion interface for a better understanding of reality/truth. Neither science nor religion/theology arrives at certainties. Both science and religion provide with a partial description of a part of reality. From that point of view the apparent contradictions between science and religion are only complementary and both together provide an authentic understanding of the absolute truth that each aims for. This awareness would pave the way for teaming up and collaborating science and religion to serve the world and humanity better. Chapter 7 is the general conclusion of the thesis.

5 Scope for Further Research

This research opens the door for applying a philosophically interdisciplinary method in statistical explorations. The statistical appropriation of Gödel's theorems with specific reference to Central Limit theorem would enhance richer possibilities in probability theory, fuzzy logic, philosophy and religion. Though many possibilities are opened up, our study is restricted to the Central Limit Property and its religious and philosophical implications. The application of fuzzy random variable is used only as a mathematical tool to combine randomness and fuzziness in the process of attaining statistical certainty and precision.