Journal of Uncertain Systems
Vol.6, No.3, pp.163-185, 2012
Online at: www.jus.org.uk

Uncertainty Statistics

Renkuan Guo[1,*], Yanhong Cui[1], Danni Guo[2]
[1] *Department of Statistical Sciences, University of Cape Town, Private Bag Rondebosch 7701, Cape Town, South Africa*
[2] *Kirstenbosch Research Center, South African National Biodiversity Institute, Private Bag X7 Claremont 7735, Cape Town, South Africa*

Received 30 October 2010; Revised 30 November 2011

Abstract

In this paper, we will investigate the uncertainty statistics from the aspects of the potential scope, the spirit or soul, the starting ground, the foundations of uncertainty statistics, etc. We also, explore the non-parametric and parametric estimation of an empirical uncertainty distribution. Furthermore, we explore the concept of uncertainty statistics and the distribution, particularly, uncertainty χ^2 statistic and distribution, uncertainty T statistic and distribution and uncertainty F statistic and distribution in order to break into uncertainty hypothesis testing and estimation doctrine. The argument style of this paper is by comparison to the classical statistics via intensive literature search.

Keywords: classical statistics, randomness study, uncertain measure, uncertainty data, uncertainty empirical distribution, uncertainty statistic, uncertainty chi-square distribution, uncertainty T distribution, uncertainty F distribution

1 Introduction

As a mathematical branch, Liu's uncertainty theory includes the uncertainty calculus, uncertainty processes, uncertainty logic and inference and has gained wide applications already, for example, uncertainty programming in system reliability, equipment allocation, and uncertainty finance, etc. For details, see [11, 12, 13, 14, 15, 16, 17]. The current research focus in uncertainty theory is the *uncertainty statistics*.

Ancient Chinese educationists Confucius said, it is the best way to learn new knowledge by reviewing the old ones. Liu's axiomatic uncertain theory [16] is new. Particularly, the σ-sub-additivity axiom brings an intrinsic feature into uncertain measure and uncertainty distribution theory [18]. Therefore, in order to avoid mixing unconsciously with the probabilistic thinking into the developments of uncertainty statistics and sharply recognize the new features of the uncertainty statistics, we will investigate the concepts and their connotations in the uncertainty statistics in comparative manner with its existing probabilistic counterpart, statistics, throughout this paper.

The statistics is the subject with respect to how to collect, organize, and interpret data, which commonly take numerical form but other forms such as symbols, qualitative, and relationships between entities may be possible [5, 24, 25, 26, 27, 28, 29].

General speaking, the contents of the statistics are abstractions of the real world. The abstraction layers in statistics are defined with clarity: population, sampling, data, data processing, presentation and inferences and drawing conclusions on the ground of the sampling data. It is evident that data are the starting ground of statistics and the whole data collection and analysis has to be guided by probability measure.

Measure defines an event measuring grade system for abstracting a conceptual uncertainty environment. Without measure specification, there is no scientific ground or consensus language to discuss any individual form of uncertainty. In other words, a measure specification is a *prerequisite* for exploring any form of uncertainty and collecting and analyzing information from real world with mathematical rigor.

[*] Corresponding author. Email: renkuan.guo@gmail.com (R. Guo).

Probability measure, proposed by Kolmogorov [10], is a "completely additive measure", i.e., the measure of union of disjoint events is just the sum of the individual measures. The statistics cannot make sense without probability measure [30, 31, 32].

Without any doubts, the future construction of an uncertainty statistics must solidly build on the uncertain measure foundation with normality, monotonicity, self-duality, and σ -sub-additivitivity and product measure five axioms established by Liu [16]. One advantage of the uncertain measure is the sound consistency with the law of contradiction and law of excluded middle. Table 1 offers a comparison between the two measure systems.

Table 1: Axiomatic probability measure and uncertain measure

	Probability Measure	Uncertain Measure
	Axiom 1: (Normality) $P\{\Omega\}=1$.	**Axiom 1:** (Normality) $\mathcal{M}\{\Xi\}=1$.
	Axiom 2: (Boundedness) $P\{\cdot\}$ is bounded by 0 and 1, i.e., for any event A, $0\le P\{A\}\le 1$.	**Axiom 2:** (Monotonicity) $\mathcal{M}\{\cdot\}$ is non-decreasing, i.e., whenever $A\subset B$, $\mathcal{M}\{A\}\le \mathcal{M}\{B\}$.
Axioms		**Axiom 3:** (Self-Duality) $\mathcal{M}\{\cdot\}$ is self-dual, i.e., for any $A\in\mathfrak{A}(\Xi), \mathcal{M}\{A\}+\mathcal{M}\{A^c\}=1$.
	Axiom 3: (σ- additivity) $P\left\{\bigcup_{i=1}^{\infty}A_i\right\}=\sum_{i=1}^{\infty}P\{A_i\}$ for any countable event sequence $\{A_i\}$, where $A_i\bigcap A_j=\varnothing,\ i\ne j,\ i,j=1,2,\cdots$	**Axiom 4:** (σ- Subadditivity) $\mathcal{M}\left\{\bigcup_{i=1}^{\infty}A_i\right\}\le\sum_{i=1}^{\infty}\mathcal{M}\{A_i\}$ for any countable event sequence $\{A_i\}$.
		Axiom 5: (Product Measure Axiom) Let Ξ_k be nonempty sets on which the uncertainty measures $\mathcal{M}_k\{\cdot\}$ are defined, $k=1,2,\cdots,d$, respectively. Then the product measure $\mathcal{M}\{\cdot\}$ on the product σ -algebra $\mathfrak{A}(\Xi)$, where $\Xi=\Xi_1\times\Xi_2\times\cdots\times\Xi_d$, i.e., $\mathfrak{A}(\Xi)=\mathfrak{A}(\Xi_1)\times\mathfrak{A}(\Xi_2)\times\cdots\times\mathfrak{A}(\Xi_d)$, is an uncertain measure. In other words, for any measurable rectangle $\Lambda=\Lambda_1\times\Lambda_2\times\cdots\times\Lambda_d$, where $\Lambda_k\in\mathfrak{A}(\Xi_k),\ k=1,2,\cdots,d$, $\mathcal{M}\{\Lambda\}=\min_{1\le k\le d}\mathcal{M}_k\{\Lambda_k\}$, i.e., $\mathcal{M}\{\Lambda\}=\begin{cases}\sup\limits_{\Lambda_1\times\Lambda_2\times\cdots\times\Lambda_d\subset\Lambda}\min\limits_{1\le k\le d}\mathcal{M}_k\{\Lambda_k\} & \sup\limits_{\Lambda_1\times\Lambda_2\times\cdots\times\Lambda_d\subset\Lambda}\min\limits_{1\le k\le d}\mathcal{M}_k\{\Lambda_k\}>0.5\\ 1-\sup\limits_{\Lambda_1\times\Lambda_2\times\cdots\times\Lambda_d\subset\Lambda^c}\min\limits_{1\le k\le d}\mathcal{M}_k\{\Lambda_k\} & \sup\limits_{\Lambda_1\times\Lambda_2\times\cdots\times\Lambda_d\subset\Lambda^c}\min\limits_{1\le k\le d}\mathcal{M}_k\{\Lambda_k\}>0.5\\ 0.5 & \text{otherwise}\end{cases}$

The structure of the remaining sections is stated as follows. Section II will be used to discuss the overall picture on classical statistics. In Section III, we will discuss the potential scope of the uncertainty statistics in order to guide the future developments of this new mathematical branch. Section IV we discuss the estimation problem of empirical uncertainty distribution. In the Section V we investigate the quadratic form related three uncertainty statistics, i.e., χ^2, T and F and derive their uncertainty distributions. Section VI concludes this paper.

2 The Soul of Statistics

The classical statistics is a collection of mathematical developments and methodologies. Statistics is a mathematical branch but it is different from axiomatic foundational mathematical branch because statistics is application-oriented.

Whenever and wherever a real world problem appears in front of scientific community, it is always possible to develop an appropriate statistical methodology to associate with it.

However, it must be emphasized that statistics is not a subject without soul or spirit. As matter of fact in its hundreds' developments, it has gradually formed its soul or spirit. Statistical spirit guided thinking is referred to statistical thinking. Guided by statistical thinking, we can grasp the spine of the skeleton of the statistics and consequently to address the practical real world problems. Without soul, it is impossible to solidly understand the nature of statistics. Figure 1 gives an overview of shell or scope of statistics.

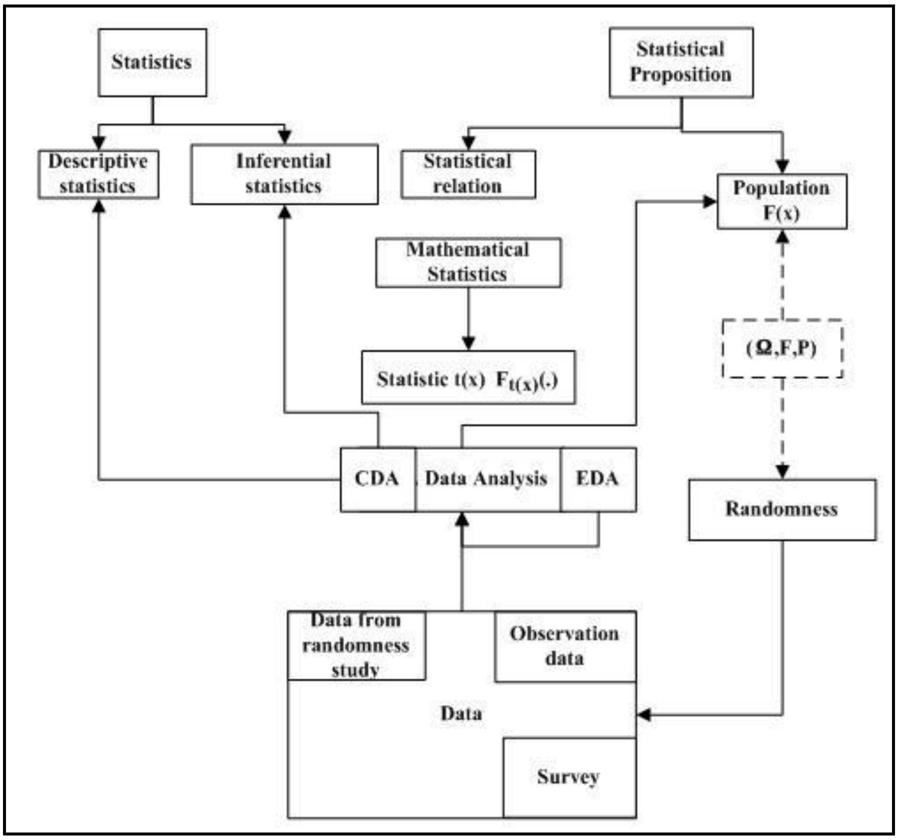

Figure 1: Shell or scope of classical statistics

From the scope of statistics, it is immediate to conclude that "statistics deals with techniques for collecting, analyzing, and drawing conclusions from data" [27]. To gain further inside of statistics, i.e., the fundamental spirit of statistics let further dig out the details.

2.1 The Population

The so-called population is in general referred to the collection of all the elements pertaining to a given feature or property.

Definition 2.1 Population in statistics is referred to the distribution (function) of a random variable X , denoted by F_X .

In elemental statistics, it often is referred to a random variable or equivalently, the probability distribution function of the random variable. Examining the roles played by population, it is the original or primary object under statistical investigation. The aim to investigate a population is to reveal its intrinsic feature, the probability distribution, or some statistical relation (or mechanism) governing the population.

Probability distribution fully characterizes a population. In other words, the full knowledge is contained in the distribution, or some of its equivalent functions, for example, moment generating function, characteristic function, and etc. In summary, the population is the working focus throughout the statistics.

2.2 The Data

Data is the primary object of statistics, which gains information from data. Without data, statisticians have no working base. It is true that data comes with noise, which is assumed to be random in classical statistics and thus relies probability theory [25].

Definition 2.2 The data is the collection of the qualitative or quantitative attributes, recorded as numbers, or symbols, or characters, or images, or statements of a random variable or set of random variables.

It is true that a population is the primary research object. However, it is seldom to scrutiny a population in one by one manner. The statistical methodology is to investigate a small group of the elements selected from the population and thus infer the relevant characteristic of the population. Intuitively speaking, such "a small group of the elements" is called as a data set, which should be a representative of the population and thus reveal the unbiased, efficient, and intrinsic feature(s) of the population.

Data is not direct copy of the real world but an abstraction. Inevitably, the abstraction level determines the data level. "The terms information and knowledge are frequently used for overlapping concepts". "Data is the lowest level of abstraction, information is the next level, and finally, knowledge is the highest level among all three". Figure 2 offers a general view on generating and data levels [25, 26, 27].

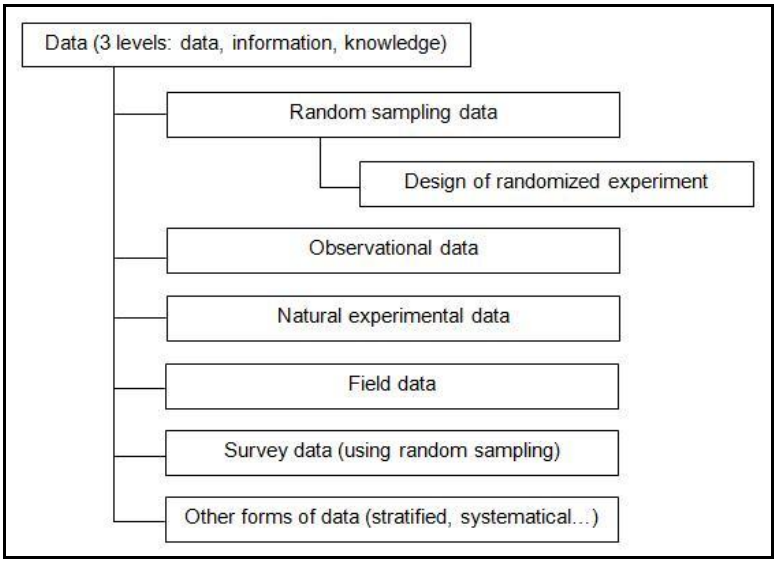

Figure 2: Data generation and abstraction level

The data collection or selection of the elements is critical. What collection scheme pursued by statisticians is that the resulted dada can be direct analyzed and the inference on the population without any bias. Such a scheme is called as the data resulted from "randomness study".

Definition 2.3 Randomness is a term referred to a characteristic by which no outcomes with describable pattern can be generated. The occurrence of each outcome is governed by a probability law.

2.3 The Statistical Data Analysis

Once a data set is generated, the next task is to gain the inside information from the data in order to extract knowledge on the population. Such a process is referred to the statistical data analysis. Without the statistical thinking the processing is aimless and no spirit. In scope, statistical data analysis can be divided into sub-stages: exploratory data analysis (EDA) and confirmatory data analysis (CDA).

2.3.1 The Exploratory Data Analysis

Once a data set is generated, the next task is to gain the information from the data in order to extract knowledge on the population, which is called as the exploratory data analysis. Turkey [28, 29] sharply pointed out that the exploratory data analysis (EDA) of statistics "is an approach to analyzing data for the purpose of formulating hypotheses worth testing, complementing the tools of conventional statistics for testing hypotheses.

2.3.2 The Confirmatory Data Analysis

Confirmatory Data Analysis (CDA), the term used for the set of ideas about hypothesis testing, p-values, confidence intervals etc. which formed the key tools in the arsenal of practicing statisticians at the time. In short, the two parts of statistics: EDA and CDA are all regarded hypothesis testing as the final destination [5, 24, 27].

Turkey [28, 29] initially suggested the distinction between exploratory and confirmatory data analysis. The first consisting in "finding patterns" in data, the second one in attempting to validate them, making sure that the perceived association are "real" and not due to random chance.

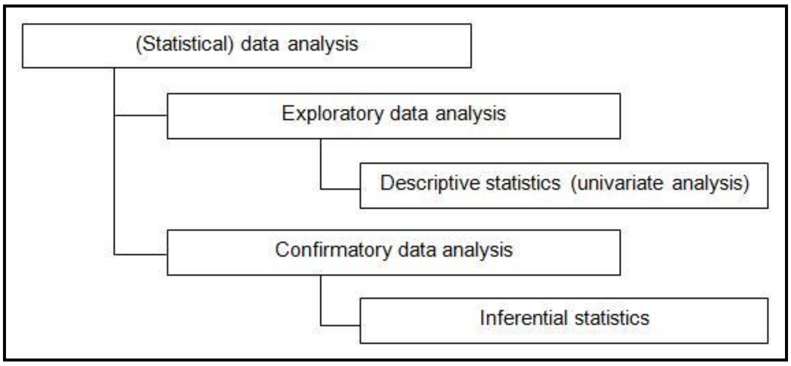

Figure 3: Statistical data analysis and scope of statistics

2.4 The Descriptive Statistics

Descriptive statistics serves double-fold purpose: (1) get organized information on the sample data in terms of statistical spirit; (2) get some indication of the population.

Descriptive statistics covers the central tendency and the dispersion about the central tendency of a set of data, which reveals the key summary of the data set quantitatively without probability measure but paves a way toward the inferential statements about the population, see [27].

Typically, the central tendency utilizes mean or median calculated from the data set, while the dispersion is presented by the standard error or inter-quantile. Today, standard statistical software offers descriptive statistical summary table, which lists the mean, the median, the mode, standard deviation, inter-quantile, skewness and kurtosis etc. of the data under investigation. Figure 4 gives the basics of the descriptive statics.

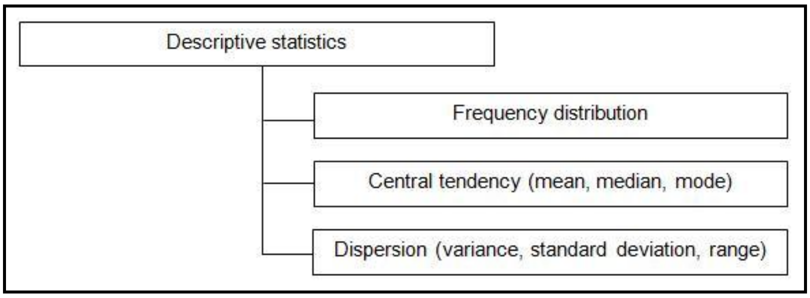

Figure 4: Descriptive statistics

2.5 The Inferential Statistics

Statistical inference is a subfield in statistics including statistical estimation, hypothesis testing and confidence interval. Typically, some statements on the population appeared as propositions, which reflect some relational functions or some mathematical characteristics, such as the expectation, variance, or linear correlations etc. will be treated as the evidence to support or deny the population related statements. Figure 5 summarizes the scope of inferential statistics.

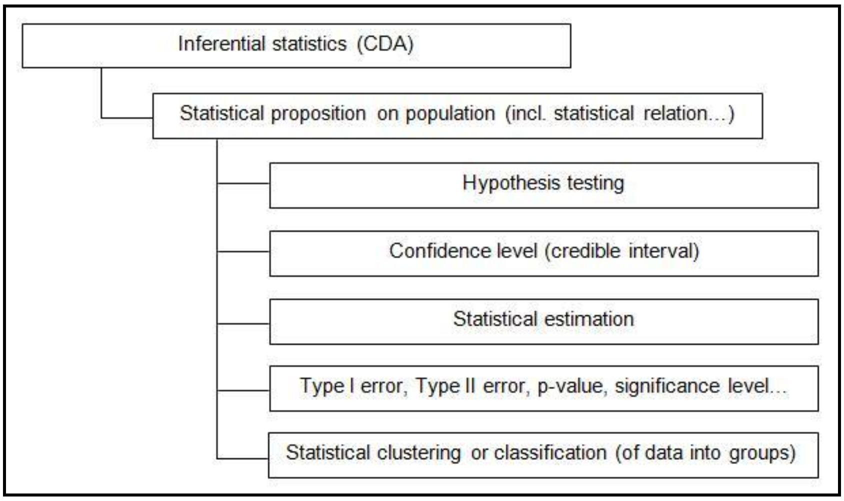

Figure 5: Inferential statistics (including scope of CDA)

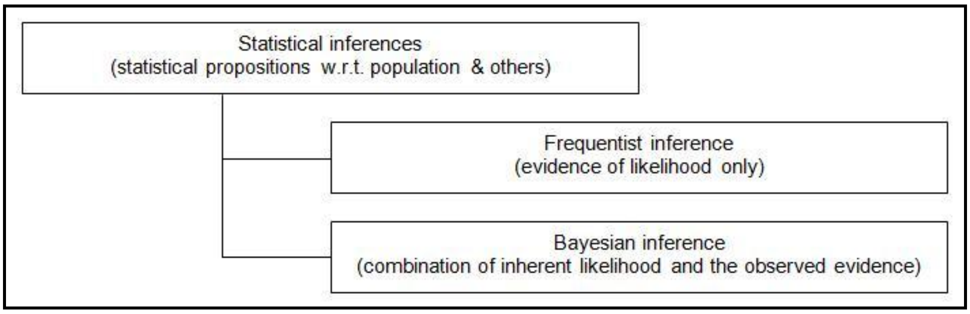

Figure 6: Statistical inferences

Statistical inference is classified into two important schools: the Frequentist inference school and the Bayesian inference school. Figure 6 shows the classification. The Frequentist inference school emphasizes the data drawn from the population is the only "source" to perform the inference on the population via the likelihood function, while the Bayesian inference stresses the modeler's experiences on the population, stated as a prior distribution as well as the data evidence stated as likelihood function together to contribute toward the inference on the population [1, 2, 4, 5]. It is worthwhile to mention that Bayesian inference is computationally intensive and becomes one of the frontier in modern statistics.

2.6 The Statistical Proposition and Statistic

At the inferential stage, the statistical proposition is the target or the focus.

Definition 2.4 A statistical proposition is a statistical statement on the population under investigation or a statement on the statistical relation rooted in the population.

In the definition above, we use adjective "statistical" repeatedly because a "statistical statement" or "statistical relation" is typically describing via statistical hypothesis testing or modeling component. Particularly, a statistic is specified to accurately express the statistical connotation of a "statistical proposition".

Definition 2.5 A statistic $t(x)$ is a measurable mapping $t:(\mathfrak{X},\mathfrak{B}_{\mathfrak{X}})\to(\mathfrak{T},\mathfrak{B}_{\mathfrak{T}})$.

Remark 2.6 A statistic is typically represented by a value of sample function. However, we should be fully aware that in mathematical statistics, statistic is a function of certain observations (i.e., a random sample typically) from a population. Therefore, a statistic is supposed to be described by a value of function calculated by the given sample and the sampling distribution of this function.

An observational value, in front of the eyes of statisticians, never is an isolated real number form observed data. It is a representative of the observed population. In mathematical statistics, a population is the distribution function of a random variable. Therefore, if denote a value in a sample as x and population as the distribution F_X of a random variable X, then x is a realization or a representative of F_X. A random sample, denoted by $(x_1, x_2, \cdots, x_n)$, where x_i is called the i^{th} sample value or sampling point, $i = 1, 2, \cdots, n$, number n is called the size of a sample. Obviously, a sample $(x_1, x_2, \cdots, x_n)$ is not just a group of real numbers, but the sample $(x_1, x_2, \cdots, x_n)$ is a series of n realizations or representatives of population F_X, thus a joint distribution $F_{X_1, X_2, \cdots, X_n}$ underlying the random sample $(X_1, X_2, \cdots, X_n)$. A statistic, as sample function, denote by $g(x_1, x_2, \cdots, x_n)$, is a sample function value and without any doubt, a distribution $F_{g(X_1, X_2, \cdots, X_n)}$ is standing behind sample function value $g(x_1, x_2, \cdots, x_n)$. In other words, whenever a statistic appears, its sampling distribution must be available (or derivable) to characterize the specific statistic value under investigation. Without the sampling distribution of a statistic development, there is no information to evaluate the efficiency of the statistic interested.

2.7 The Mathematical Statistics

Mathematical statistics is the subfield or collection of all relevant mathematical theories for supporting statistical inferences, for example, linear models, probability distributions, and etc. [4, 5, 24, 25].

The scope of mathematical statistics includes (1) distributional foundation; (2) point estimation theory; (3) hypothesis testing theory; (4) interval estimation theory; (5) linear models and generalized models; (6) non-parametric statistical theory. Figure 7 describes the scope of mathematical statistics and Figure 8 offers the modeling level impacts to mathematical statistics.

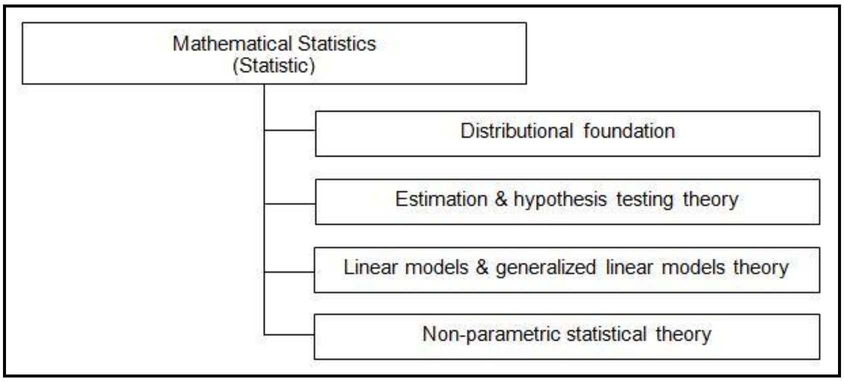

Figure 7: Scope of mathematical statistics

Statisticians distinguish between three levels of modeling assumptions.

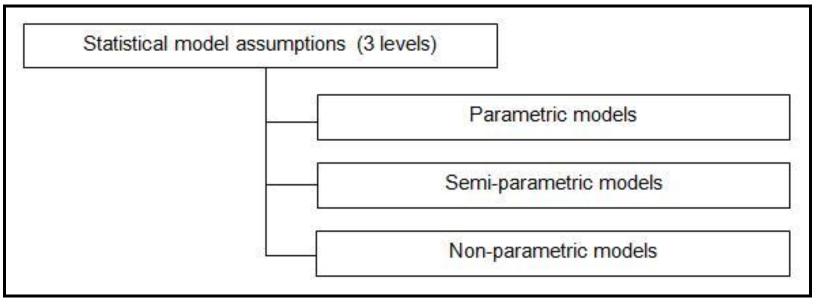

Figure 8: Level of statistical modeling assumptions

As the convention of the mathematical statistics, the connotation of statistic requires a statistic must be estimable and a sampling distribution at least an asymptotic distribution must available or derivable. In other words, the basic requirement of “statistic” is that the value of statistic is calculable. However, calculable or estimable quantity calculated from a group of observations is in no way automatically stands for a "statistic". For example, in engineering practices, assuming that we have three points (x_i, y_i), $i = 1,2,3,$ can we get a linear relationship between y and x? The answer is yes. Or even further, we can ask if we can fit a quadratic curve to pass the three points because engineers see the least-square criterion fitted straight line only passes the points in between, the answer is yes, in terms of spline function we can calculate and plot the quadratic curve. It is often seen some commercial software to do the slope calculation and provide the plot to show what the fitted straight line looks. Do the calculable straight line (interception and slope are estimable) or the spline calculated quadratic curve (three coefficients: (a,b,c) in the fitted spline-quadratic curve, $y = ax^2 + bx + c$) brings you into statistics? No, not at all! Using least-square criterion or other optimization criterion, for example, minimax criterion does not warrant a statistic to be obtained.

When talking statistic in probabilistic statistics, a statistic does not only require the calculated *value of statistic* but also require *the sampling distribution* as a part of statistic [5], at least the estimated variance of the statistic. We can understand the connotation is just containing these two aspects, which in a sense is guiding rule for facilitating a statistic. As a matter of fact, sampling is an integral part underlying the modern statistics [26].

The statistical decision theory proposed by Wald [5, 24], is an important doctrine, which is trying to unify the statistical estimation and hypothesis testing etc. into one theory. Although statistical decision theory is still a shell, however, the spirit of the statistical decision theory has merged into many sub-branches of statistical science.

3 The Potential Scope of Uncertainty Statistics

The uncertainty statistics may be defined as a subject of investigating the uncertainty data collection, analyzing the observations, and drawing conclusion from the observational data based on the uncertain measure theoretic foundation. It is obvious that the definition of uncertainty statistics is similar to the probabilistic counterpart – statistics. However, there is fundamental difference between the classical statistics and the uncertainty statistics to be developed. Figure 9 offers a potential scope of the uncertainty statistics.

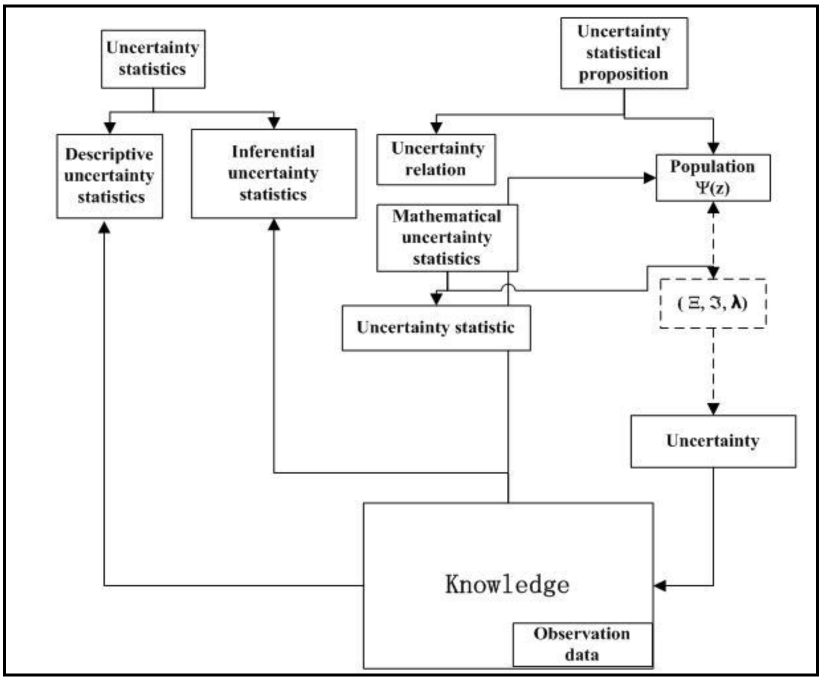

Figure 9: Potential scope of the uncertainty statistics

3.1 The Soul of Uncertainty Statistics

Parallel to statistics, the uncertainty statistics should be inevitably set up the uncertainty hypothesis testing as the cornerstone (uncertainty statistical proposition), i.e., the "soul" or the "nerve" of the uncertainty statistics is to carry on hypothesis testing on the uncertainty population based on observational data.

Whenever we talk about inference methodology, i.e., hypothesis testing, it does not mean a verification methodology but does mean the inference on population from the representative information, i.e., observational data. Verification of claim or statement is an important task and an integrated part of mathematical science, physical science, chemical science or other scientific branches. It is well known fact that a mathematical conjecture needs a formal rigorous proof to be a mathematical theorem. Any individual counter example to a mathematical theorem can lead to the full rejection of the theorem. The verification methodology is a deterministic one in nature. An uncertainty "observation" provided by a scientific experimentation or an expert's experimentation may play a role in verifying the claim or the statement in certain sense. However, the uncertainty character of the random sample or uncertainty observation serves the intrinsic role to inference on the probability distribution or the uncertainty distribution respectively, not for the verification of a deterministic claim or a guess purpose. Table 2 gives a comparison between two statistics.

Table 2: Concept comparisons between statistics and uncertainty statistics

Concept	Probabilistic Statistics	Uncertainty Statistics
Measure space	Probability space $(\Omega, \mathfrak{F}, P)$	Uncertainty space $(\Xi, \mathfrak{A}(\Xi), \hat{\lambda})$
Population	Probability distribution of a random variable F_X	Uncertainty distribution Ψ_ξ of an uncertainty variable ξ
Collection of data on population	Sampling data by a probability measure, P	Uncertainty data–expert's experimentation, observation data, i.e., data at knowledge level.
Data	A realization from the population, $(x_1, x_2, \cdots, x_n)$	Knowledge data sequence $(z_1, z_2, \cdots, z_n)$ in terms of expert's opinion of the uncertainty population
Feature of data	Observable, sampling repeatedly, and objectively	Subjective judgment or phenomenological data in the form of knowledge
Inference	Inference on probability distribution F_X from a sample $(x_1, x_2, \cdots, x_n)$	Inference on uncertainty distribution Ψ_ξ from an knowledge data $(z_1, z_2, \cdots, z_n)$

Now, let us illustrate the similarity and dissimilarity between the two statistics and the related concepts by the following example.

Example 3.1: To address the problem "the distance from Beijing to Tianjin", there are three approaches available: (1) Random sampling; (2) Experts' evaluations; (3) verification. Firstly, assuming that "the distance from Beijing to Tianjin" is a population, i.e., normal distribution, denoted by F_X of random quantity–distance, denoted as $X \overset{d}{\sim} N(\mu, \sigma^2)$. This population by can be constituted of all the measurements by every possible devices available, from ruler, electronic to laser instruments. Form the population $F_X(x) = \Phi((x-\mu)/\sigma)$, we take a random sample of 5 points $(105, 99, 101, 95, 87)$ (unit: KM) from satellite remote sensing image measurements. Now, we can answer the claim "the distance from Beijing to Tianjin is about 100 Km". For justification of this claim, a hypothesis testing: $H_0{:}\mu{=}100$ may be perused. The test statistic is $t = (\bar{x} - \mu)/\hat{\sigma}_{\bar{x}} = (97.4 - 100)/3.0594 = -0.8498$ with a critical points -2.776 or $+2.776$ at 5% significant level. Therefore, we say, at 5 percent significant level, there is no reason to reject the null hypothesis $H_0{:}\mu{=}100$. Consequently, we can confirm the claim "the distance from Beijing to Tianjin is about 100 Km" statistically. Secondly, assuming that "the distance from Beijing to Tianjin" is an uncertainty population, which is constituted of all experts' opinion on "the distance from Beijing to Tianjin". This population, i.e., Liu's normal Ψ_ξ of an normal uncertainty variable with standard deviation $\sigma_0 = 3.0$. Assuming further an uncertainty observational data of size 5 is taken from the experts' distance population $(98, 99, 100, 102, 101)$. From the uncertainty sample, bear the hypothesized uncertainty distribution $\Psi_\xi(z){=}1\Big/\left(1+\exp\left(-\pi(z-100)\big/\sqrt{3}\sigma_0\right)\right)$. We can construct an uncertainty chi-square statistic $\chi^2_{(5)} = \sum_{i=1}^{5}\left((z_i - 100)/3.0\right)^2 = 11.1111$, while the uncertainty statistic $\Psi_{\chi^2_{(5)}}(1.256667) = 0.71286$ does indicate that the uncertainty observational evidence rejects the hypothesized population. Thirdly, for verifying the claim: "the distance from Beijing to Tianjin is about 100 Km". We either take a device measurement or an expert's opinion to get a number, say, 98.8 Km. Once the distance 98.8 Km is available, the verification task is done. We can accept the claim, since the relative error is merely 0.012 .

The uncertainty statistical inference is just to utilize the information contained in the observational data - knowledge to inference on the population as accurate as possible. In other words, uncertainty statistical inference is the spine or nerve system throughout the uncertainty statistics. Due to the fact that an uncertainty observational data do not contain full information of the population, the uncertainty statistical inference can not be 100% precise.

3.2 Uncertainty Population

As the starting or grounding point of uncertainty statistics, it is necessary to address is the concept of uncertainty population. The so-called "population" is just the distribution of an uncertainty variable, denoted by Ψ_ξ, while the "observation" is the realization or the representative (denoted by z) of the uncertainty variable, ξ. The uncertainty distribution Ψ_ξ is unknown or partially unknown. If Ψ_ξ is fully known, there is no need to carry on any inference at the standing point of the uncertainty statistics. The uncertainty observational data-knowledge reveals some information with respect to the uncertainty population in certain degree. As a matter of fact, an uncertainty observational data–knowledge is the starting point or ground for inference on the uncertainty population, where the knowledge comes from.

Definition 3.2 An uncertainty population in uncertainty statistics is the uncertainty distribution Ψ_ξ of an uncertainty variable ξ.

Definition 3.3 An uncertain observational data is referred to the knowledge level data from observing the uncertainty population in terms of characteristic aspect, recording the observed feature in real value format, and collecting these values into knowledge data set.

Remark 3.4 Different from sampling data in statistics, there is no uncertainty sampling data term being defined and thus available in the uncertainty statistics although an uncertainty population in uncertainty statistics is defined parallel to the probabilistic population in statistics.

3.3 Uncertainty Data-Knowledge

Liu [16] pointed out, Uncertainty statistics is a group of "methodology for collecting and interpreting expert's experimental data by uncertainty theory".

Conventional statistical experimentation in its own standing is a structured procedure guided by statistical theory and methodology to generate experimental outcomes optimally for pursuing certain scientific truth.

In contrast, the connotation of "expert's experimental data by uncertainty theory" deserves an exploration. In certain degree, Liu's uncertainty statistics intends to concentrate on expert's opinions and therefore he regards the expert's opinions as the basic data resources for uncertainty statistics.

On the other hand, just as Liu emphasized in his book [16] uncertainty theory is an abstraction of human uncertainty and thus can guide the modeling efforts of human behavior and activities. Human is not an isolated living being, therefore human uncertainty comes from its basic individual life, social life, economic life, the interactions between human and its living environments, the Earth, the Solar System, and the Universal, etc. Just as the fields to which Liu's uncertainty theory is applied are far beyond the scale of expert's opinion.

Remark 3.5 Parallel to revelation of the connotation of randomness. Impreciseness occupies an fundamental position in uncertainty statistics. Expert's knowledge is inevitably imprecise and imperfect. Impreciseness is referred to a term with an intrinsic property governed by an uncertain measure or an uncertainty distribution for each of the actual or hypothetical members of an uncertainty population (i.e., collection of expert's knowledge). An uncertainty process is a repeating process whose outcomes follow no describable deterministic pattern, but follow an uncertainty distribution, such that the uncertain measure of the occurrence of each outcome can be only approximated or calculated.

Definition 3.6 Impreciseness is an intrinsic property of a variable or an expert's knowledge being specified by an uncertainty measure.

Remark 3.7 Impreciseness exists in engineering, business and research practices due to measurement imperfections, or due to more fundamental reasons, such as insufficient available information, ... , or due to a linguistic nature, because it is an unarguable fact that impreciseness exists intrinsically in expert's knowledge on the real world.

Definition 3.8 Let ξ be a uncertainty quantity of impreciseness on an uncertainty measure space $(\Xi,\mathfrak{A}(\Xi),\tilde{\lambda})$. The uncertainty distribution of ξ is $\Psi_{\xi}(x)=\tilde{\lambda}\{\tau\in\Xi \mid \xi(\tau)\le x\}$.

Remark 3.9 An imprecise variable ξ is an uncertainty variable and thus is a measurable mapping, i.e., $\xi:\mathbb{D}\to\mathbb{R},\ \mathbb{D}\subseteq\mathbb{R}$. An observation of an imprecise variable is a real number, (or more broadly, a symbol, or an interval, or a real-valued vector, a statement, etc), which is a representative of the population or equivalently of an uncertainty distribution $\Psi_{\xi}(\cdot)$ under a given scheme comprising set and σ-algebra. The single value of a variable with impreciseness should not be understood as an isolated real number rather an interval or a set.

In conclusion, data in uncertainty statistics should include both statistical knowledge, which are observable objectively and repeatedly, and expert's experimental data, which may be subjective, even phenomenological, as long as the data collection, interpretation, and inference on the population are all with uncertainty statistical spirit.

3.4 Uncertainty Statistic

In classical statistics, the elementary function of sampling data is called a statistic, which implies two inter-related aspects: (1) a summary of sample data into *a characteristic value*; and (2) the *distribution* governs the statistic.

Therefore, when we engage the developments of *uncertainty statistics*, we must keep the soul or spirit of statistics, *an uncertainty statistic must have two aspects: value of the uncertainty statistic and the uncertainty distribution of the uncertainty statistic*. Any part missing will prevent us from approaching the truth underlying the uncertainty observational data and therefore alienate us away from conventional statistics communities.

Definition 3.10 An uncertainty statistic $t(z)$ is a measurable mapping $t:(\mathfrak{Z},\mathfrak{B}_{\mathfrak{Z}})\to(\mathfrak{T},\mathfrak{B}_{\mathfrak{T}})$, where $\mathfrak{Z}\subset\mathbb{R}$ is the real valued set of experts' knowledge, $\mathfrak{B}_{\mathfrak{Z}}$ is the Borel σ-algebra of $\mathfrak{Z}$, $\mathfrak{T}$ is the real valued set of the function t, and $\mathfrak{B}_{\mathfrak{T}}$ is the Borel σ-algebra of $\mathfrak{T}$.

Remark 3.11 The measurability of an uncertainty statistic warrants the two intrinsic sides of it: estimable and distribution extractability. Such a feature of an uncertainty statistic should be a lifeline throughout the whole uncertainty statistics. Table 3 gives a comparison between the two statistics.

Table 3: Statistic and uncertainty statistic

Item	Probabilistic statistic	Uncertainty statistic
statistic	Statistic: A measurable mapping $t:(\mathfrak{X},\mathfrak{B}_{\mathfrak{X}})\to(\mathfrak{T},\mathfrak{B}_{\mathfrak{T}})$ - $t(x)$	Uncertainty statistic: A measurable mapping $t:(\mathfrak{Z},\mathfrak{B}_{\mathfrak{Z}})\to(\mathfrak{T},\mathfrak{B}_{\mathfrak{T}})$ - $t(z)$
Connotation of a statistic	(Sample) value of the statistic	value of an uncertain statistic
	(Sampling) probability distribution of the statistic	The uncertainty distribution of the uncertainty statistic
Inferential role	Inference on probability distribution F_X from a sample $(x_1,x_2,\cdots,x_n)$ in terms of a statistic, $t(x)$	Inference on uncertainty distribution Ψ_{ξ} from a observational data $(z_1,z_2,\cdots,z_n)$ in terms of an uncertainty statistic $t(z)$

3.5 Uncertainty Statistical Decision Theory

In the paper of Guo et al. [7], the basic elements of decision analysis oriented to observational data arising from a general uncertainty environment, so that a shell for Bayesian uncertainty decision doctrine is established. Further, Guo et al propose a mechanism, which paves the way towards the establishment of a posterior uncertainty distribution of the parameter vector given the observational data, based on uncertain measure Axiom 5. The significance of this paper is to establish for the first time a Bayesian uncertainty data inference and decision framework, which constitutes a critical step towards the establishment of uncertainty statistics and a Bayesian uncertainty decision theory.

4 Estimating an Empirical Uncertainty Distribution

An uncertainty distribution is just the population, i.e., the target of the uncertainty statistical inference. Hence it is vital to develop the estimation schemes of uncertainty distribution.

Classical statistics put a data-constructed *frequency distribution* as the important step guiding the data analysts' entrance into statistical analysis [5, 27]. It can be also parallel to argue the critical step to obtain an uncertainty observational data based *empirical uncertainty distribution*. Although different from probability and statistics, a distribution can determine or induce the probability measure uniquely, an uncertainty distribution could not. Nevertheless, in the sense of equivalent class, uncertainty distribution is still the ladder towards an uncertain measure.

Gao [6] and Wang et al. [30] utilized the Delphi method to summarize experts' opinion for establishing an empirical uncertainty distribution. The Delphi method was initiated by the Rand Corporation for the military projects during 1950-1960. Then it is developed as experts' opinion-oriented methodology for forecasting technology, socio-economic and business trends. The Delphi method is a structural survey and utilizes the intuitive available information of the (independently acting) participants, who are mainly experts. The "standard" process is that (1) set up the topic; (2) a facilitator collects opinions from individual panel member (i.e., expert); (3) The facilitator produces a synthesized and consistent feedback to all participants; (4) Panel members resubmit their opinions in the light of the first round feedback information provided; (5) Repeat the resubmit-feedback iterations until a consensus is reached [3, 9, 21].

The Delphi method is not of statistical nature, although the variation of Delphi method does add some statistical approaches to improve its efficiency. To be fair, Delphi method provides a successful methodology to expert's knowledge processing until a consensus is reached. Recall that in (probabilistic) statistics, any data point, even generated from an expert's opinion, is not regarded as an isolated qualitative index or a quantity number but a "representative" from a population, i.e., the uncertainty distribution of an uncertainty variable. Delphi method reaches a consensus, which is either a predicted qualitative index or a forecasted quantity number without a distribution underlying the future prediction and therefore disqualifies itself from a statistical method. In fact, no matter how popular the Delphi method in the circle dealing with experts' opinions, it is not a part of statistics, including Bayesian statistics. While Gao [6] and Wang et al. [30] initiated an innovative uncertainty statistical scheme for utilizing experts' opinion to facilitating an empirical uncertainty distribution via the Delphi-like procedure, it should not be attributed to Delphi method. Subsection 4.1 will further carry on the discussion.

4.1 A Non-parametric Estimation Scheme of Empirical Uncertainty Distribution

To fully understand the non-parametric estimation of an empirical uncertainty distribution, let us examine the methodology pursued in statistics. In [27], the frequency distribution is described and the construction method is detailed. Data are collected from a population, and every data point represents the population in certain degree. The frequency distribution is purely data based under the assumption that every data point comes with equal likelihood. Except equal likelihood assumption, there is no hypothesized distribution in curve shape or parameters underlying the data. Assuming that 130 data points are collected, denoted as $\{x_1, x_2, \cdots, x_{130}\}$, total number $n = 130$, let us construct a frequency distribution from the 130 data points.

The construction can be divided into ten steps:

(1) Determine the domain for the distribution, $D = [a, b]$, typically, $a = \min_{1\le i\le 130}\{x_i\} - c$, $b = \max_{1\le i\le 130}\{x_i\} + c$, where $c > 0$, a small constant;

(2) Determine the number of group or number of sub-intervals over $D = [a, b]$, say, $m = 10$;

(3) Calculate the endpoints of all sub-intervals, denoted by $I_i = [a_i, b_i), i = 1, 2, \cdots, 10$. It is obvious that $a_1 = a$,and $b_{10} = b$, the length of sub-interval is $l = (b-a)/10$, then $I_i = [a + (i-1)l, a + il)$, $i = 1, 2, \cdots, 10$;

(4) Grouping original data $\{x_1, x_2, \cdots, x_{130}\}$ into 10 groups according to a criterion that data point $x_j^{(i)}$ in the i^{th} sub-group $\{x_1^{(i)}, x_2^{(i)}, x_{n_i}^{(i)}\}$, satisfy the inequality $a_i \le x_j^{(i)} < b_i$, $i = 1, 2, \cdots, 10$;

(5) Collect the number of data points within each sub-interval, n_i, $i = 1, 2, \cdots, 10$, named as the frequency of sub-interval $I_i = [a + (i-1)l, a + il)$;

(6) Calculating the relative frequency, $f_i = n_i / n$, where $n = \sum_{i=1}^{10} n_i$;

(7) Calculating accumulative relative frequency, denoted as $F_i = \sum_{j=1}^{i} f_i$, $i = 1,2,\cdots,10$;

(8) Collecting relative frequencies $\{F_0 \triangleq 0, F_1, \cdots, F_0 \triangleq 1.0\}$;

(9) Calculate the mid-points of all sub-intervals, denoted by, $x_{(i)}$, $x_{(i)} = a + il/2$, $i = 1,2,\cdots,10$;

(10) Connecting the m pairs $\left(x_{(i)}, F_i\right)$, $i = 1,2,\cdots,10$, the frequency distribution F is obtained.

In contrast, let us examine the "Delphi method" created by Wang et al. [30]. For a given statement, a group of m experts is invited. A delicate questionnaire is distributed among them. The i^{th} expert is requested to evaluate the likelihood α_i of a specific event, denoted by z_i . It is assumed that the facilitator pre-arrange the knowledge sequence: $z_1 < z_2 < \cdots < z_m$ and the likelihood evaluations are independent. Experts submit their evaluations $\{(z_i, \alpha_i), i = 1,2,\cdots,m\}$. When the facilitator receive the $\{(z_i, \alpha_i), i = 1,2,\cdots,m\}$, s/he check if the monotonic property of likelihoods $\alpha_1 \le \alpha_2 \le \cdots \le \alpha_m$ is satisfied. The facilitator can either discard a few α_i , the use $\hat{\alpha}_i = (\alpha_{i-1} + \alpha_{i+1})/2$ to replace α_i or keep original α_i as $\hat{\alpha}_i$ the utilize least-squares criterion fitting a monotonic curve $F : 0 \le F(z) \le 1.0$ such that

$$\sum_{i=1}^{m} \left(F(z_i) - \alpha_i\right)^2 \tag{1}$$

is minimized, then $(z_i, \hat{\alpha}_i)$, $i = 1,2,\cdots,m$ is obtained. If the facilitator satisfies with the fitted empirical uncertainty distribution, the procedure stops, otherwise, feedback the first round results, i.e., $(z_i, \hat{\alpha}_i)$, $i = 1,2,\cdots,m$ to experts for the second round likelihood evaluations; each expert bases on feedback, re-evaluate and submit new pair $\left(z_i, \alpha_i^{(2)}\right)$, $i = 1,2,\cdots,m$. The facilitator will re-examine these $\left(z_i, \alpha_i^{(2)}\right)$, $i = 1,2,\cdots,m$, re-fit the empirical uncertainty distribution, re-feedback, if necessary, otherwise, repeat until the consensus is reached.

As to the arbitrary point z , the empirical uncertainty distribution may utilize a linear extrapolation approach

$$F(x) = \begin{cases} 0 & \text{if } z < z_1 \\ \alpha_i + \dfrac{z - z_i}{z_{i+1} - z_i}(\alpha_{i+1} - \alpha_i) & \text{if } z_i \le z \le z_{i+1} \\ 1 & \text{if } z > z_m \end{cases} \tag{2}$$

which is referred to Liu's empirical uncertainty distribution [16].

Remark 4.1 In viewing Wang et al.'s [30] fitting an empirical uncertainty distribution with comparison with Snedecor and Cochran's [27] construction for a frequency distribution, it is obvious that Step 1 to Step 9 in Snedecor and Cochran's construction of a frequency distribution is now replaced by experts' knowledge processing and evaluation in Wang et al.'s [30] fitting an empirical uncertainty distribution. Whether the quality is high or not depends upon the knowledge level and processing skill of the experts as well as their psychological qualifications.

Remark 4.2 Gao's [6] fitting an empirical uncertainty distribution is more delicate than that of Wang et al.'s [30]. Both methods are of non-parametric empirical uncertainty distribution fitting.

4.2 A Parametric Estimation of Uncertainty Distribution

Wang [31] develops a moment method to estimate an empirical uncertainty distribution. For an uncertainty variable ξ , the k^{th} moment is defined $m_k(\xi) = \mathrm{E}\lfloor \xi^k \rfloor$, see [16].

Wang has shown that for an empirical uncertainty distribution with form as Eq.(2), the theoretical k^{th} moment

$$m_k(\xi) = \mathrm{E}\left[\xi^k\right] = \alpha_1 z_1^k + \frac{1}{k+1} \sum_{i=1}^{m-1} \sum_{j=0}^{k} (\alpha_{i+1} - \alpha_i) z_i^j z_{i+1}^{k-j} + (1 - \alpha_m) z_m^k . \tag{3}$$

Accordingly, Wang defines an empirical moment $\hat{m}_k(\xi)$ in [31] if the expert's knowledge on the empirical uncertainty distribution are given by $(\hat{z}_i, \hat{\alpha}_i)$, $i = 1,2,\cdots,n$, such that $(\hat{z}_i, \hat{\alpha}_i) \le (\hat{z}_{i+1}, \hat{\alpha}_{i+1})$, $i = 1,2,\cdots,n$,

$$\hat{m}_k(\xi) = \hat{\alpha}_1 \hat{z}_1^k + \frac{1}{k+1} \sum_{i=1}^{n-1} \sum_{j=0}^{k} (\hat{\alpha}_{i+1} - \hat{\alpha}_i) \hat{z}_i^j \hat{z}_{i+1}^{k-j} + (1 - \hat{\alpha}_n) \hat{z}_n^k ,\ k = 1,2,\cdots . \tag{4}$$

The moment method is assuming that an uncertainty variable, ξ, has an uncertainty distribution $\Psi_\xi\left(z;\theta_1,\theta_2,\cdots,\theta_p\right)$, where $\theta_1,\theta_2,\cdots,\theta_p$ are p unknown parameters. Furthermore, the expert's knowledge on the empirical uncertainty distribution are given by $\left(\hat{z}_i,\hat{\alpha}_i\right)$, $i=1,2,\cdots,n$, such that $\left(\hat{z}_i,\hat{\alpha}_i\right)\le\left(\hat{z}_{i+1},\hat{\alpha}_{i+1}\right)$, $i=1,2,\cdots,n$, such that the empirical uncertainty distribution Ψ and empirical moments $\hat{m}_k\left(\xi\right)$ are estimable, $k=1,2,\cdots,p$. Then a non-linear equations system

$$\hat{m}_k\left(\xi\right)=k\int_0^{+\infty}z^{k-1}\left(1-\Psi\left(z;\theta_1,\theta_2,\cdots,\theta_p\right)\right)dz,\ k=1,2,\cdots,p. \tag{5}$$

Substitute the solution to equation system, $\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_p$ into the theoretical uncertainty distribution $\Psi_\xi\left(z;\theta_1,\theta_2,\cdots,\theta_p\right)$, the moment-method estimated uncertainty distribution is $\Psi_\xi\left(z;\hat{\theta}_1,\hat{\theta}_2,\cdots,\hat{\theta}_p\right)$. For details and examples, see [31] and [32].

Remark 4.3 Many of Wang's terms used in [30] had changed. For example, we change Wang's "sample moment" into "empirical moment", change "sample data" into "expert's knowledge" etc, because in the uncertainty statistics, the uncertainty sample is undefined.

Remark 4.4 The semi-parametric approach can be constructed, say, in terms of spline function, cross entropy, or other methodology. However, integration form should not apply to specify the empirical uncertainty distribution.

5 Uncertainty Distributions for Quadratic Form Related Quantities

Classical mathematical statistics addresses the inferential analysis under random uncertainty environment, where Gaussian distributional doctrine facilitates the basic sampling statistics and their sampling distributions. Under an uncertainty environment, the uncertainty mathematical statistics will address uncertainty, which differs from randomness. Gaussian distribution is disqualified in uncertainty statistics since that the integration involved in Gaussian specification violates the sub-σ-additivity axiom. It is expected that Liu's uncertainty normal distributional theory will play the similar roles of those of Gaussian in probabilistic statistics [11, 12, 13, 14, 15, 16, 17, 18].

As matter of fact, the mathematical statistics starts with the "sampling distributions", e.g., chi-square distribution (including Gamma distribution), *t*-distribution, *F*-distributions, and etc, [5, 24, 25]. Without the sampling distributional theory, the inferential statistics, say, hypothesis testing or confidence intervals have no foundation. Therefore, we will investigate the uncertainty distributions of certain important uncertainty statistic related to uncertainty quadratic form. Now, let us investigate the uncertainty statistic of uncertainty observational data from an uncertainty population, i.e., an uncertainty normal distribution.

5.1 Uncertainty Normal Population or Uncertainty Normal Distribution

As previous section stated, an uncertainty statistic is a function of observations from a uncertainty population. Thus, let us discuss the connotation of an uncertainty normal population.

Definition 5.1 [11, 12, 13, 14, 15, 16, 17] An uncertainty variable ξ on $\left(\Xi,\mathfrak{A}\left(\Xi\right),\lambda\right)$ is called normal if its uncertain distribution takes the form

$$\Psi_\xi\left(x\right)=\frac{1}{1+e^{-\frac{\pi}{\sqrt{3}\sigma}(x-\theta)}},\quad x\in\mathbb{R}. \tag{6}$$

Definition 5.2 An uncertain normal distribution is standard if its uncertain distribution takes the form

$$\Psi_{\xi_0}\left(x\right)=\frac{1}{1+e^{-\frac{\pi}{\sqrt{3}}x}},\quad x\in\mathbb{R}. \tag{7}$$

The standard uncertain normal distribution function ($\theta=0,\sigma=1$) is plotted in Figure 10.

Table 4 gives an initial comparison between Gaussian and Liu's uncertainty normal variable.

Finally, we should emphasize that the functional form of an uncertainty normal distribution can be used in probability theory to play a role of cumulative distribution, while the probabilistic counterpart, Gaussian distribution cannot play any distributional roles in uncertainty theory because Gaussian distribution is expressed by an integration, which in nature is σ-additive.

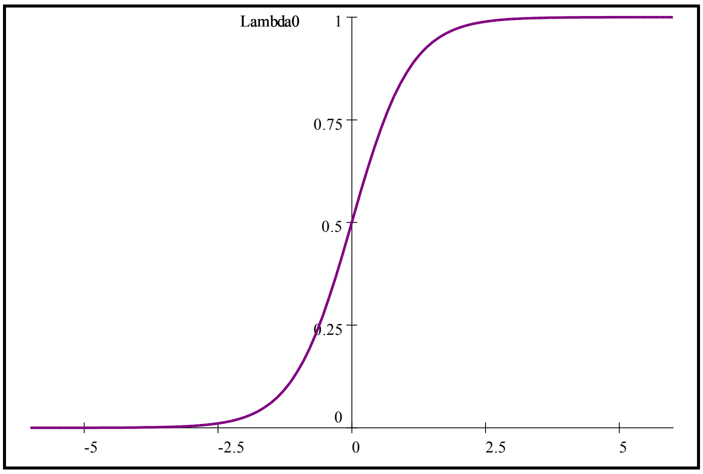

Figure 10: The standard uncertain normal distribution ($\theta=0,\sigma=1$)

Table 4: Basic comparisons between uncertainty normal variable and Gaussian random variable (mean-standard deviation parametric form)

	Uncertainty normal Variable	Gaussian Random Variable
Standard	ξ_0	Z
	$\Psi_{\xi_0}(x)=\dfrac{1}{1+e^{-\frac{\pi}{\sqrt{3}}x}}$	$\Phi(x)=\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}}e^{-s^2/2}ds$
General	ξ	X
	$\Psi_{\xi}(x)=\dfrac{1}{1+e^{-\frac{\pi}{\sqrt{3}\sigma}(x-\theta)}}$	$F(x)=\int_{-\infty}^{x}\dfrac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}ds$
Link	$\xi=\theta+\sigma\xi_0$	$X=\mu+\sigma Z$

5.2 Uncertainty Chi-square Statistic

Definition 5.3 Let ξ_0 be single observation from a standard uncertainty normal population, define $\chi^2_{(1)}=\xi_0^2$. Then we call it a chi-square statistic (with degree of freedom one).

Definition 5.4 The uncertainty distribution of an uncertainty statistic with degree of freedom one, $\chi^2_{(1)}$, is defined by

$$\Psi_{\chi^2_{(1)}}(x)=\lambda\left\{\chi^2_{(1)}\le x\right\}=\begin{cases}0 & x\le 0\\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3}}\sqrt{x}\right)} & x>0.\end{cases} \tag{8}$$

Theorem 5.5 Let $\xi_{0,1},\xi_{0,2},\cdots,\xi_{0,n}$ be *i.i.d.* observations from a standard uncertainty normal population. Then the uncertainty chi-square statistic $\chi^2_{(n)}=\sum_{i=1}^{n}\xi^2_{0,i}$ has an uncertainty distribution taking a form

$$\Psi_{\chi^2_{(n)}}(x)=\lambda\left\{\chi^2_{(n)}\le x\right\}=\begin{cases}0 & x\le 0\\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\sqrt{x}\right)} & x>0.\end{cases} \tag{9}$$

Proof: Based on Definition 5.4, $\xi^2_{0,i}=\chi^2_{(1),i}$ has uncertainty distribution

$$\Psi_{\chi^2_{(1),i}}(x) = \mathcal{M}\left\{\chi^2_{(1),i} \le x\right\} = \frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}}\sqrt{x}\right)},\ x > 0. \tag{10}$$

Also, notice that function $f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} x_i$ is monotonic increasing. Hence, Theorem 1.19 [27] operational law is applicable. Further notice that the summand $\chi^2_{(1),i}$ follows distribution function

$$\Psi_{\chi^2_{(1),i}}(x) = \frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}}\sqrt{x}\right)},\ x > 0. \tag{11}$$

Its inverse takes the form

$$\Psi^{-1}_{\xi_0^2}(\alpha) = \left[\frac{\sqrt{3}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2. \tag{12}$$

Hence

$$\Psi^{-1}_{\chi^2_{(n)}}(\alpha) = n\left[\frac{\sqrt{3}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2 = \left[\frac{\sqrt{3n}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2, \tag{13}$$

i.e.,

$$\Psi_{\chi^2_{(n)}}(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3n}}\sqrt{x}\right)} & x > 0. \end{cases} \tag{14}$$

which concludes the proof of the theorem.

Remark 5.6 If ξ_0 is assumed to be a random variable with probability distribution

$$F_{\xi_0}(x) = \mu\{\xi_0 \le x\} = \frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}}x\right)},\ x \in \mathbb{R}, \tag{15}$$

in [8], the probability distribution of the random variable version $\chi^2_{pseudo,n} = \sum_{i=1}^{n}\xi^2_{0,i}$ where $\xi_{0,1}$, $\xi_{0,2}, \cdots, \xi_{0,n}$ are assumed to be *i.i.d.* from probability distribution taking a form shown in Eq.(3), $\chi^2_{pseudo,n}$ is called *pseudo-chi-square* random variable with degree of freedom *n*. The random variable ξ_0^2 has a probability distribution [8] takes a form

$$F_{\xi_0^2}(x) = \mu\left\{\xi_0^2 \le x\right\} = \frac{\sinh\left(\frac{\pi}{\sqrt{3}}\sqrt{x}\right)}{1+\cosh\left(\frac{\pi}{\sqrt{3}}\sqrt{x}\right)},\ 0 < x < +\infty. \tag{16}$$

Also, in probability theory, the square of a standard normal random variable, Z^2, follows chi-square distribution with degree-of-freedom one and its density takes a form

$$f(t) = \frac{1}{\Gamma(1/2)2} t^{\frac{1}{2}-1} e^{-\frac{t}{2}},\ t > 0. \tag{17}$$

Actually, $F_{\xi_0^2}(x)$ is very close to chi-square distribution with degree of one. The following plot gives a visual justification. The pseudo-chi-square distribution function of ξ_0^2 is plotted in Figure 11.

Figure 12 shows the overlay plot of the probability distribution of the pseudo-chi-square variable, ξ_0^2, and the chi-square distribution with degree-of-freedom one, Z^2.

The plot of uncertainty chi-square distribution with degree of freedom one is shown in Figure 13, which clearly demonstrates the sharp differences between the probability distribution of the random variable $\chi^2_{pseudo,1}$ and the uncertainty distribution of the uncertainty variable $\chi^2_{(1)}$.

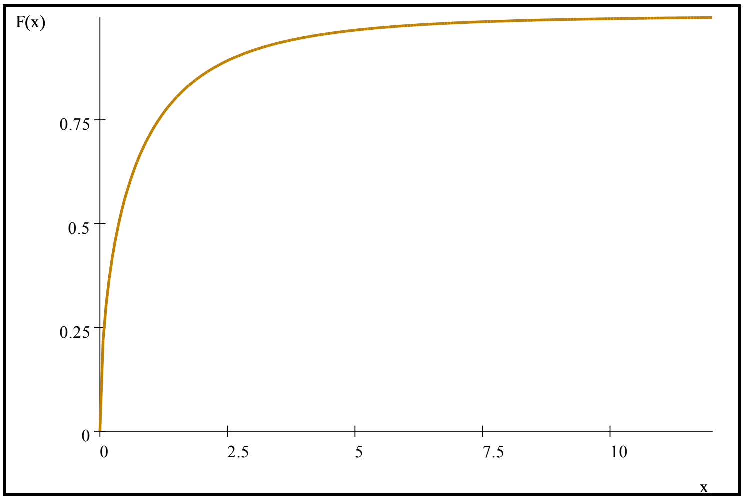

Figure 11: The probability distribution of the pseudo-chi-square variable ξ_0^2

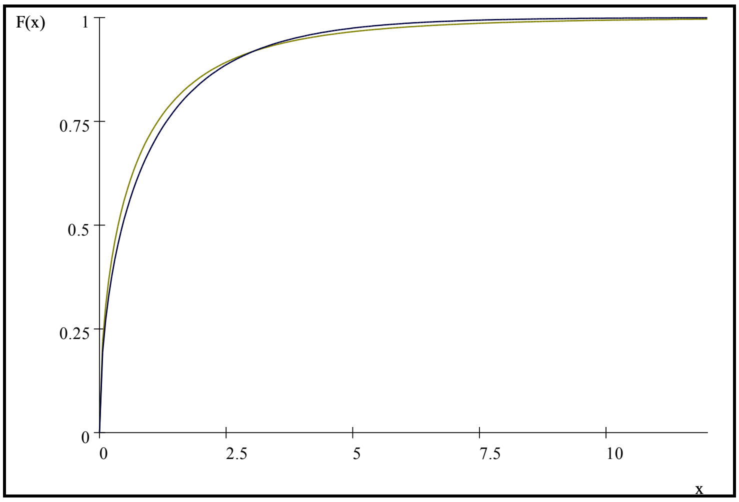

Figure 12: The overlay plot of probability distribution of ξ_0^2 (brown-colored curve) and $Z^2 = \chi_1^2$ (navy-colored curve)

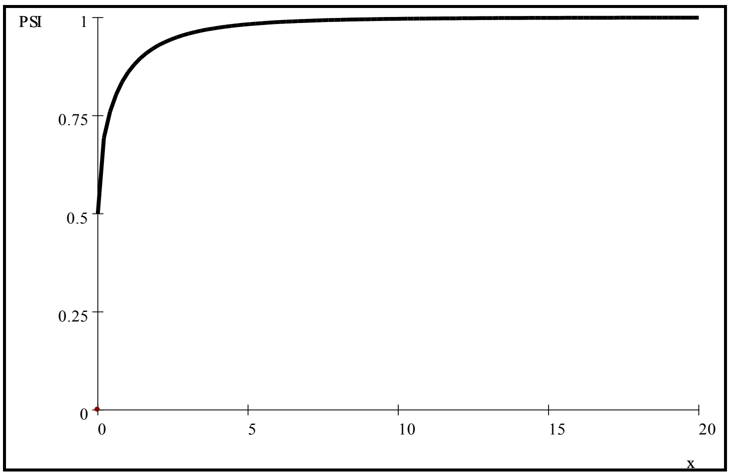

Figure 13: The plot of the uncertainty chi-square distribution with degree of freedom one

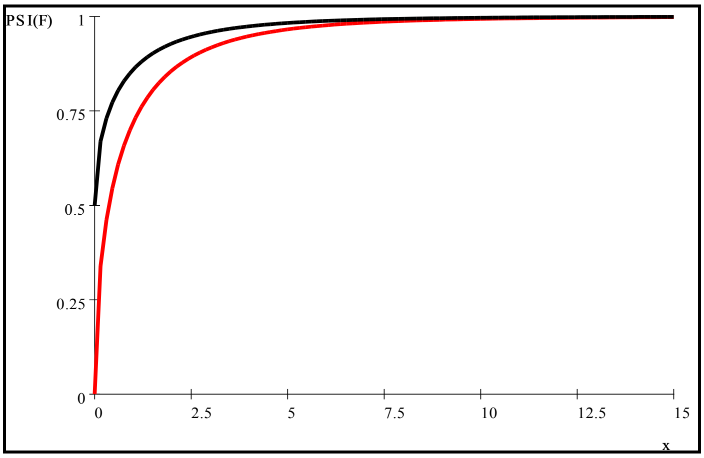

Figure 14: The overlay plot of pseudo-chi-square of probability distribution (red-colored curve) and the uncertainty chi-square distribution (black-colored curve)

The uncertainty distribution function $\Psi_{\chi^2_{(1)}}(x)$ of the uncertainty chi-square variable $\chi^2_{(1)}$ is a left horizontal shifted distributional curve of the pseudo-chi-square random variable, see Figure 14, Therefore, giving a name of uncertainty chi-square distribution seems inappropriate at the first place. However, considering about the composition of an uncertainty chi-square distribution with degree of freedom one, it seems similar to that of the probabilistic chi-square distribution.

The following table offers a systematic comparison of the probabilistic chi-square, the probabilistic pseudo-chi-square, and the uncertainty chi-square distributions.

Table 5: Comparison of three chi-square variables

	Chi-square χ_n^2	Pseudo-chi-square $\chi^2_{pseudo,n}$	Uncertainty chi-square $\chi^2_{(n)}$
Component	Z	Z_{pseudo}	ξ_0
Component distribution	$\Phi(x)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{s^2}{2}\right)ds$	$\Lambda(x)=\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}}x\right)}$	$\Psi_{\xi_0}(x)=\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}}x\right)}$
Construction	$\chi_n^2=\sum_{i=1}^{n}Z_i$	$\chi^2_{pseudo,n}=\sum_{i=1}^{n}Z_{pseudo,i}$	$\chi^2_{(n)}=\sum_{i=1}^{n}\xi^2_{0,i}$
Distribution	$\int_0^x\frac{1}{\Gamma(n/2)2^n}s^{\frac{n}{2}-1}e^{-\frac{s}{2}}ds$	$\frac{\sinh\left(\frac{\pi}{\sqrt{3n}}\sqrt{x}\right)}{1+\cosh\left(\frac{\pi}{\sqrt{3n}}\sqrt{x}\right)}$	$\begin{cases}0 & x\le 0\\ \frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3n}}\sqrt{x}\right)} & x>0\end{cases}$
Density	$\frac{1}{\Gamma(n/2)2^n}x^{\frac{n}{2}-1}e^{-\frac{x}{2}},x>0$	$\frac{\pi}{2\sqrt{3}}\frac{1}{1+\mathrm{ch}\left(\frac{\pi}{\sqrt{3n}}x\right)}$	Undefined

5.3 Uncertainty *T* Statistic

The next important uncertainty statistic is *T* statistic with *n* degree of freedom.

Definition 5.7 Let ξ_0 be a single observation from a standard uncertainty normal population and $\chi^2_{(n)}$ be an uncertainty chi-square statistic constructed from *n i.i.d.* observations $\xi^*_{0,1},\xi^*_{0,2},\cdots,\xi^*_{0,n}$ from another standard normal population, ξ^*_0. Furthermore, ξ_0 and ξ^*_0 are assumed to be mutually independent. Then the function

$$T_n = \frac{\xi_0}{\sqrt{\chi^2_{(n)}/n}} \tag{18}$$

is called an uncertainty student statistic with degree of freedom *n*, where

$$\chi^2_{(n)} = \sum_{i=1}^{n} \xi^*_{0,i} \ . \tag{19}$$

The uncertainty distribution of the uncertainty *T* statistic is called the uncertainty *t* distribution function. In order to derive uncertainty *t* distribution a series of theorems will be stated and proved as preparations.

Theorem 5.8 Let $\xi_{0,1}, \xi_{0,2}, \cdots, \xi_{0,n}$ be *i.i.d.* standard uncertainty normal variables. Then the uncertainty average chi-square variable $\bar{\chi}^2_{(n)} = \chi^2_{(n)}/n = \frac{1}{n}\sum_{i=1}^{n}\xi^2_{0,i}$ has an uncertainty distribution

$$\Psi_{\chi^2_{(n)}}(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\sqrt{x}\right)} & x > 0. \end{cases} \tag{20}$$

Proof: Notice that we have shown that the uncertainty chi-square variable has a distribution function

$$\Psi_{\chi^2_{(n)}}(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\sqrt{x}\right)} & x > 0. \end{cases} \tag{21}$$

which has an inverse

$$\Psi^{-1}_{\chi^2_{(n)}}(\alpha) = n\left[\frac{\sqrt{3}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2 = \left[\frac{\sqrt{3n}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2 . \tag{22}$$

Thus, for the uncertainty variable $\bar{\chi}^2_{(n)} = (1/n)\chi^2_{(n)}$, its inverse of the uncertainty distribution is

$$\Psi^{-1}_{\bar{\chi}^2_{(n)}}(\alpha) = n\Psi^{-1}_{\chi^2_{(n)}} = n^2\left[\frac{\sqrt{3}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2 = \left[\frac{n\sqrt{3}}{\pi}\left(\ln(\alpha) - \ln(1-\alpha)\right)\right]^2 \tag{23}$$

which leads to the conclusion:

$$\Psi_{\bar{\chi}^2_{(n)}}(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\sqrt{x}\right)} & x > 0. \end{cases} \tag{24}$$

Theorem 5.9 Let $R_n = \sqrt{\bar{\chi}^2_{(n)}}$. Then the uncertainty distribution of R_n is

$$\Psi_{R_n}(r) = \begin{cases} 0 & r \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}r\right)} & r > 0. \end{cases} \tag{25}$$

Proof: Notice that we have shown that the uncertainty variable $\bar{\chi}^2_{(n)}$ has an uncertainty distribution

$$\Psi_{\bar{\chi}^2_{(n)}}(x) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\sqrt{x}\right)} & x > 0. \end{cases} \tag{26}$$

Further notice that $f(x) = \sqrt{x}$ is monotonic increasing in x for $x > 0$, the inverse of the uncertainty distribution of R_n

$$\Psi_{R_n}^{-1}(\alpha) = f\left(\Psi_{\bar{\chi}^2_{(n)}}^{-1}(\alpha)\right) \tag{27}$$

which implies

$$\Psi_{R_n}(r) = \Psi_{\bar{\chi}^2_{(n)}}\left(f^{-1}(r)\right) = \Psi_{\bar{\chi}^2_{(n)}}\left(r^2\right) = \begin{cases} 0 & r \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}r\right)} & r > 0, \end{cases} \tag{28}$$

which concludes the proof.

Corollary 5.10 Let $S_n = 1\Big/\sqrt{\bar{\chi}^2_{(n)}}$. Then the uncertainty distribution of S_n is

$$\Psi_{S_n}(s) = \begin{cases} 0 & s \le 0 \\ 1 - \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\dfrac{1}{s}\right)} & s > 0. \end{cases} \tag{29}$$

Proof: Notice that the uncertainty distribution of $S_n = \dfrac{1}{R_n} = \dfrac{1}{\sqrt{\chi^2_{(n)}/n}}$ is

$$\Psi_{S_n}(s) = 1 - \Psi_{R_n}\left(f^{-1}(s)\right) = 1 - \Psi_{R_n}\left(\frac{1}{s}\right) = \begin{cases} 0 & s \le 0 \\ 1 - \dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\dfrac{1}{s}\right)} & s > 0. \end{cases} \tag{30}$$

Corollary 5.11 Let $W_n = 1\big/\bar{\chi}^2_{(n)}$. Then the uncertainty distribution of W_n is

$$\Psi_{W_n}(w) = \begin{cases} 0 & w \le 0 \\ 1 - \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\dfrac{1}{\sqrt{w}}\right)} & w > 0. \end{cases} \tag{31}$$

Proof: Notice that the uncertainty distribution of $W_n = 1\big/\bar{\chi}^2_{(n)}$ is

$$\Psi_{W_n}(w) = 1 - \Psi_{\bar{\chi}^2_{(n)}}\left(f^{-1}(w)\right) = 1 - \Psi_{\bar{\chi}^2_{(n)}}\left(\frac{1}{w}\right) = \begin{cases} 0 & x \le 0 \\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\dfrac{1}{\sqrt{x}}\right)} & x > 0. \end{cases} \tag{32}$$

Theorem 5.12 The uncertainty statistic

$$T_n = \frac{\xi_0}{\sqrt{\chi^2_{(n)}/n}} \tag{33}$$

has an uncertainty t distribution function

$$\Psi_{T_n}(t) = \sup_{x_1/x_2 = t}\left(\min\left(\frac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3}}x_1\right)}, \frac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\dfrac{1}{x_2}\right)}\right)\right). \tag{34}$$

Proof: Notice that

$$T_n = \frac{\xi_0}{\sqrt{\chi^2_{(n)}/n}} = \frac{\xi_0}{\sqrt{\chi^2_{(n)}}} = \xi_0 \cdot S_n\,. \tag{35}$$

According to Corollary 5.8, the uncertainty variable S_n follows an uncertainty distribution

$$\Psi_{S_n}(s)=1-\Psi_{R_n}\left(f^{-1}(s)\right)=1-\Psi_{R_n}\left(\frac{1}{s}\right)=\begin{cases}0 & s\le 0\\ 1-\dfrac{1}{1+\exp\left(-\dfrac{\pi}{n\sqrt{3}}\dfrac{1}{s}\right)} & s>0.\end{cases} \tag{36}$$

Furthermore, T_n are monotonic increasing function of ξ_0 and S_n, thus

$$\Psi_{T_n}(t)=\sup_{x_1 s=x}\min\left(\Psi_{\xi_0}(x_1),\Psi_{S_n}(s)\right) \tag{37}$$

which implies

$$\Psi_{T_n}(t)=\sup_{x_1/x_2=t}\left(\min\left(\frac{1}{1+\exp\left(-\frac{\pi}{\sqrt{3}}x_1\right)},\frac{1}{1+\exp\left(-\frac{\pi}{n\sqrt{3}}\frac{1}{x_2}\right)}\right)\right). \tag{38}$$

5.4 Uncertainty *F* Statistic

The third important uncertainty statistic is the uncertainty *F* statistic.

Definition 5.13 Let $\xi_{0,1},\xi_{0,2},\dots,\xi_{0,m}$ be m *i.i.d* observations from an uncertainty standard normal, ξ_0, and $\xi^*_{0,1},\xi^*_{0,2},\cdots,\xi^*_{0,n}$ be n *i.i.d.* observations from another uncertainty standard normal population, ξ^*_0, respectively. Define $\bar{\chi}^2_{(m)}=\chi^2_{(m)}\big/m=\sum_{i=1}^m\xi_{0,i}\big/m$ and $\bar{\chi}^2_{(n)}=\chi^2_{(n)}\big/n=\sum_{i=1}^m\xi^*_{0,j}\big/n$, then

$$F_{(m,n)}(v)=\frac{\chi^2_{(m)}\big/m}{\chi^2_{(n)}\big/n}=\frac{\bar{\chi}^2_{(n)}}{\bar{\chi}^2_{(n)}} \tag{39}$$

is called uncertainty *F* statistic with degree of freedom m and n.

Theorem 5.14 An uncertainty F statistic follows an uncertainty distribution taking a form

$$\Psi_{F_{(m,n)}}(v)=\begin{cases}0 & x_1\le 0, x_2\le 0\\ \sup\limits_{x_1/x_2=v}\left(\min\left(\dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3m}}\sqrt{x_1}\right)},\dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\dfrac{1}{\sqrt{x_2}}\right)}\right)\right) & x_1>0, x_2>0.\end{cases} \tag{40}$$

Proof: Notice that

$$F_{(m,n)}(v)=\frac{\chi^2_{(m)}\big/m}{\chi^2_{(n)}\big/n}=\frac{\bar{\chi}^2_{(n)}}{\bar{\chi}^2_{(n)}}. \tag{41}$$

According to Theorem 5.7 and Corollary 5.9, the uncertainty variable $\bar{\chi}^2_{(m)}$ follows an uncertainty distribution

$$\Psi_{\bar{\chi}^2_{(m)}}(x)=\begin{cases}0 & x\le 0\\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3m}}\sqrt{x}\right)} & x>0.\end{cases} \tag{42}$$

And the uncertainty variable $W_n=1\big/\bar{\chi}^2_{(n)}$ follows an uncertainty distribution

$$\Psi_{W_n}(w)=1-\Psi_{\bar{\chi}^2_{(n)}}\left(f^{-1}(w)\right)=1-\Psi_{\bar{\chi}^2_{(n)}}\left(\frac{1}{w}\right)=\begin{cases}0 & w\le 0\\ \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\dfrac{1}{\sqrt{w}}\right)} & x>0.\end{cases} \tag{43}$$

Furthermore, $F_{(m,n)}$ are monotonic increasing function of $\bar{\chi}^2_{(m)}$ and W_n, thus

$$\Psi_{F_{(m,n)}}(v) = \sup_{x_1 w = v} \min\left(\Psi_{\bar{\chi}^2_{(m)}}(x_1), \Psi_{W_n}(w)\right), \tag{44}$$

which implies

$$\Psi_{F_{(m,n)}}(v) = \begin{cases} 0 & x_1 \le 0, x_2 \le 0 \\ \sup\limits_{x_1/x_2 = v}\left(\min\left(\dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3m}}\sqrt{x_1}\right)}, \dfrac{1}{1+\exp\left(-\dfrac{\pi}{\sqrt{3n}}\dfrac{1}{\sqrt{x_2}}\right)}\right)\right) & x_1 > 0, x_2 > 0. \end{cases} \tag{45}$$

6 Conclusion

In this paper, we systemically survey the classical statistics from the scope, the spirit or soul, the frequency distribution, the concept of statistic, and the hypothesis testing and interval theory, the scope of mathematical statistics, the statistical decision theory, for a solid understanding of the nature of statistics. Then we similarly study the potential scope and spirit of the uncertainty statistics, the starting ground - uncertainty data – knowledge, the empirical uncertainty distribution, the uncertainty statistic and its uncertainty distribution. For example, uncertainty χ^2 statistic, uncertainty T statistic and uncertainty F statistic and their uncertainty distributions are developed. It should be aware that the uncertainty data – expert's knowledge is established on the ground of specification of an uncertainty measure space $(\Xi, \mathfrak{A}(\Xi), \tilde{\lambda})$, i.e., the uncertainty knowledge must be the representatives of the uncertainty population (or alternatively, the uncertainty distribution of an uncertainty variable) according to the uncertain measure (of the events). It is also should be emphasized that impreciseness is an intrinsic feature to expert's knowledge (data), which is inevitably governed by an uncertain measure or by an uncertainty distribution.

Acknowledgements

This research is supported by VC's Incentive for NRF Rated Researchers, University of Cape Town, and South African National Research Foundation Grant IFR2010042200062, and IFR2009090800013 as well.

The research is also partially supported by Tsinghua University during Professor Guo's visit to the Uncertainty Theory Laboratory, Department of Mathematical Science, Tsinghua University from July 18 to October 14, 2010.

The authors sincerely thank Professor B.D. Liu and his PhD students: X.W. Chen, W. Dai, K. Yao, and Z. X. Peng for their invaluable debates and inputs.

References

[1] Berger, J.O., *Statistical Decision Theory and Bayesian Analysis* (Springer Series in Statistics), 2nd Edition, Springer-Verlag, 1985.

[2] Berger, J.O., and R.L. Wolpert, *The Likelihood Principle*, 2nd Edition, The Institute of Mathematical Statistics, Haywood, CA, 1988.

[3] Blind, K., K. Cuhls, and H. Grupp, Personal attitudes in the assessment of the future of science and technology: a factor analysis approach, *Technological Forecasting & Social Change*, vol.68, pp.131–149, 2001.

[4] Casella, G., and R.L. Berger, *Statistical Inference*, Duxbury Press, 2001.

[5] Cheng, X.R., *Introductory Mathematical Statistical Theory*, The Scientific Publishing House, Beijing, China (in Chinese), 1981.

[6] Gao, J.W., Determine uncertainty distribution via Delphi method, *Proceedings of the First International Conference on Uncertainty Theory*, pp.291–297, 2010.

[7] Guo, R., Y.H., Cui, and D. Guo, Uncertainty decision analysis, *Proceedings of the Eighth Annual Conference on Uncertainty*, pp.250–261, 2010.

[8] Guo, R., D. Guo, and C. Thiart, Liu's uncertainty normal distribution, *Proceedings of the First International Conference on Uncertainty Theory*, pp.191–207, 2010.

[9] Kaplan, A., A.L., Skogstad, and M.A. Girshick, The prediction of social and technological events, *The Public Opinion Quarterly*, vol.14, no.1, pp.93–110, 1950.

[10] Kolmogorov, A.N., *Foundations of the Theory of Probability* (translated by Nathan Morrison), Chelsea, New York, 1950.

[11] Liu, B.D., *Uncertainty Theory*, 2nd Edition, Springer-Verlag Heidelberg, Berlin, 2007.

[12] Liu, B.D., Some research problems in uncertainty theory, *Journal of Uncertain Systems*, vol.3, no.1, pp.3–10, 2009.

[13] Liu, B.D., Uncertain risk analysis and uncertain reliability analysis, *Journal of Uncertain Systems*, vol.4, no.3, pp.163–170, 2010.

[14] Liu, B.D., Uncertain entailment and modus ponens in the framework of uncertain logic, *Journal of Uncertain Systems*, vol.3, no.4, pp.243–251, 2009.

[15] Liu, B.D., Uncertain set theory and uncertain inference rule with application to uncertain control, *Journal of Uncertain Systems*, vol.4, no.2, pp.83–98, 2010.

[16] Liu, B.D., *Uncertainty Theory: A Branch of Mathematics of Modelling Human Uncertainty*, Springer-Verlag, Berlin, 2010.

[17] Liu, B.D., Extreme value theorems of uncertain process with application to insurance risk models, http://orsc.edu.cn/online/100722.pdf, 2010.

[18] Liu, Y.H., How to generate uncertain measures, *Proceedings of Tenth National Youth Conference on Information and Management Sciences*, pp.23–26, 2008.

[19] Peng, J., and K. Yao, A stock model for uncertain markets, http://orsc.edu.cn/online/090209.pdf, 2012.

[20] Peng, Z.X., and K. Iwamura, A sufficient and necessary condition of uncertainty distribution, *Journal of Interdisciplinary Mathematics*, to be published.

[21] Pill, J., The Delphi method: substance, context, a critique and an annotated bibliography, *Socio-Economic Planning Science*, vol.5, no.1, pp.57–71, 1971.

[22] Primas, H., Basic elements and problems of probability theory, *Journal of Scientific Exploration*, vol.13, no.4, pp.579–613, 1999.

[23] Qin, Z.F., and X. Gao, Fractional Liu process with application to finance, *Mathematical and Computer Modelling*, vol.50, nos.9-10, pp.1538–1543, 2009.

[24] Rao, C.R., *Linear Statistical Inference and Its Applications*, 2nd Edition, John Willey & Sons, New York, 1973.

[25] Ross, S.M., *Introduction to Probability and Statistics for Engineers and Scientists*, 3rd Edition, Elsevier, Academic Press, Burlington, MA, U.S.A, 2004.

[26] Smith, T.M.F., Biometrika centenary: sample surveys, *Biometrika*, vol.88, no.1, pp.167–243, 2001.

[27] Snedecor, G.W., and W.G. Cochran, *Statistical Methods*, 8th Edition, Iowa State University Press, Ames, Iowa, USA, 1989.

[28] Turkey, J.W., Exploratory data analysis is an attitude, a flexibility, and a reliance on display, not a bundle of techniques, and should be so taught, *The American Statistician*, vol.34, no.1, pp.23–25, 1980.

[29] Turkey, J.W., *Exploratory Data Analysis*, Addison-Wesley, 1977.

[30] Wang, X.S., Z.C. Gao, and H.Y. Guo, Delphi method for estimating uncertainty distributions, http://orsc.edu.cn/online/100830.pdf, 2010.

[31] Wang, X.S., Moment estimation of uncertainty distribution, *Proceedings of the Eighth Annual Conference on Uncertainty*, pp.168–172, 2010.

[32] Wang, X.S., and Z.X. Peng, Method of moments for estimating uncertainty distributions, http://orsc.edu.cn/online/100408.pdf, 2010.