ELSEVIER

Fuzzy Sets and Systems 116 (2000) 141–154

FUZZY sets and systems

www.elsevier.com/locate/fss

Statistics in measurements

Shiro Ishikawa

Department of Mathematics, Faculty of Science and Technology, Keio University, 3-14-1, Hiyoshi, Kohoku-ku, Yokohama, 223-8522, Japan

Received June 1997; received in revised form May 1998

Abstract

Recently, we proposed “measurement theory”, i.e., the foundation of measurements. This is a general measurement theory for both classical and quantum systems. The purpose of this paper is to give a foundation to statistics in terms of measurements, that is, to characterize or understand statistics as one of the aspects of measurement theory. Studying several statistical examples in the light of measurement theory, we show that Fisher’s and Bayes’s methods are described in terms of measurements. Also, we characterize “estimation under loss function in statistics” in measurement theory. Therefore, we may conclude that statistics is a certain aspect of measurements. This viewpoint is important since it clarifies the relation between statistics and the other aspects of measurements.

Keywords: Probability; Data analysis methods; Statistical methods; Measurements; Fuzzy logic

1. Introduction

Recently in [6,7] we proposed the foundation of measurements, which was called “fuzzy measurement theory” or, in short, “measurement theory”. This theory is a general measurement theory for both classical and quantum systems. Also, it is composed of two parts, that is, “objective measurement theory” and “subjective measurement theory”. The former is fundamental, and the latter is rather methodological.

Most statisticians consider that statistics is closely related to “measurements”, or, statistics is the study to analyze measured data for some purpose. Therefore, if a foundation of measurements has been proposed, the proposal should be immediately examined in comparison with statistics. The purpose of this paper is to execute it, in other words, to propose a measurement theoretical formulation of statistics. In Section 2 we review the measurement theory that was proposed in [6,7], and add an example concerning “at random”. In Section 3 we study “statistical inferences for states” in objective (resp. subjective) measurement theory, which should be compared to Fisher’s method (resp. Bayes’s method) in statistics. In Section 4 we study “approximate measurements for quantities” in measurement theory, which corresponds to “estimation under loss function” in statistics. Since two main topics in statistics, i.e., “Fisher’s and Bayes’s methods” and “estimation under loss function”, can be described in terms of measurements, we may conclude that statistics is a certain aspect of measurement theory. This viewpoint is important since it bridges the gap between statistics and the other aspects of measurements. For example, from the viewpoint, we can

E-mail address: ishikawa@math.keio.ac.jp (S. Ishikawa).

0165-0114/00/$ - see front matter
PII: S0165-0114(98)00280-2

understand the relation between "fuzzy logic" and statistics. (Throughout this paper, the word "fuzzy logic" is used in the meaning of Remark 4.6 in Section 4.)

For completeness, we again note that our purpose is "to formulate statistics in measurement theory" and not "to apply statistics to measurements". We believe that "measurement" is the most fundamental concept in science. Therefore, we always start from "measurement".

2. Measurements and "at random"

According to [6,7], in this section we introduce the foundation of measurements. Kolmogorov's probability theory is of course a useful mathematical theory for analyzing measured data. However, we are concerned with measurements as well as measured data. Therefore, we must prepare the theory of operator algebras, which is indispensable for a mathematical formulation of measurements, or more generally, "system theory". However, it should be noted that "mathematics" is only used as a tool to represent the concept of "measurement". In fact, this paper does not include theorems but only examples.

Let $\mathscr{A}$ be a C^*-algebra, i.e., a Banach $*$-algebra satisfying the C^*-condition, cf. [6,7,15]. Throughout this paper, we always assume, for simplicity, that $\mathscr{A}$ has the identity I. An element T in $\mathscr{A}$ is called *self-adjoint* if $T=T^*$ holds. Also, a self-adjoint element T in $\mathscr{A}$ is called *positive* (and denoted by $T\geqslant 0$) if there exists an element T_0 in $\mathscr{A}$ such that $T=T_0^*T_0$ where T_0^* is the adjoint element of T_0. A positive element T is called a *projection* if $T=T^2$ holds. Let $\mathscr{A}^*$ be the dual Banach space of $\mathscr{A}$. That is, $\mathscr{A}^*=\{\rho$: ρ is a continuous linear functional on $\mathscr{A}\}$ with the norm $\|\cdot\|_{\mathscr{A}^*}(\equiv \sup\{|\rho(T)|:\ \|T\|_{\mathscr{A}}\leqslant 1\})$. (The linear functional $\rho(T)$ is sometimes denoted by ${}_{\mathscr{A}^*}\langle\rho,T\rangle_{\mathscr{A}}$.) Define the *mixed state class* $\mathfrak{S}^m(\mathscr{A}^*)$ such that $\mathfrak{S}^m(\mathscr{A}^*)=\{\rho\in\mathscr{A}^*:\ \|\rho\|_{\mathscr{A}^*}=1$ and $\rho(T)\geqslant 0$ for all $T\geqslant 0\}$. A mixed state ρ $(\in\mathfrak{S}^m(\mathscr{A}^*))$ is called a *pure state* if it satisfies that "$\rho=\lambda\rho_1+(1-\lambda)\rho_2$ for some $\rho_1,\rho_2\in\mathfrak{S}^m(\mathscr{A}^*)$ and $0<\lambda<1$" implies "$\rho=\rho_1=\rho_2$". Define $\mathfrak{S}^p(\mathscr{A}^*)\equiv\{\rho^p\in\mathfrak{S}^m(\mathscr{A}^*)$: ρ^p is a pure state$\}$, which is called a *state space*.

When $\mathscr{A}$ is a commutative C^*-algebra, that is, $T_1\cdot T_2=T_2\cdot T_1$ holds for all $T_1,T_2\in\mathscr{A}$, by Gelfand theorem (cf. [15]) we can put $\mathscr{A}=C(\Omega)$, the algebra composed of all continuous complex-valued functions on a compact Hausdorff space Ω. It is well known that $C(\Omega)^*=\mathscr{M}(\Omega)$, i.e., the Banach space composed of all regular complex-valued measures on Ω. And therefore, $\mathfrak{S}^m(\mathscr{M}(\Omega))=\{\rho\in\mathscr{M}(\Omega):\ \rho\geqslant 0, \|\rho\|_{\mathscr{M}(\Omega)}=1\}$, which is denoted by $\mathscr{M}_{+1}(\Omega)$. Also, it is clear that $\mathfrak{S}^p(\mathscr{M}(\Omega))=\{\delta_\omega\in\mathscr{M}(\Omega)$: δ_ω is a point measure at $\omega\in\Omega$, i.e., ${}_{\mathscr{M}(\Omega)}\langle\delta_\omega,f\rangle_{C(\Omega)}=f(\omega)$ $(\forall f\in C(\Omega),\forall\omega\in\Omega)\}$, which is denoted by $\mathscr{M}_{+1}^p(\Omega)$. And therefore, we have the identification: $\Omega\ni\omega\leftrightarrow\delta_\omega\in\mathscr{M}_{+1}^p(\Omega)$. Thus, the compact Hausdorff space Ω may be also called a *state space*.

As a natural generalization of Davies' idea in quantum mechanics (cf. [2]), a C^*-*observable* (or in short, *observable*, *fuzzy observable*) $\boldsymbol{O}\equiv(X,\mathscr{F},F)$ in a C^*-algebra $\mathscr{A}$ is defined such that it satisfies

(i) X is a set, and $\mathscr{F}$ is the subfield of the power set $\mathscr{P}(X)\,(\equiv\{\Xi:\ \Xi\subseteq X\})$,

(ii) for every $\Xi\in\mathscr{F}$, $F(\Xi)$ is a positive element in $\mathscr{A}$ such that $F(\emptyset)=0$ and $F(X)=I$ (where 0 is the 0-element in $\mathscr{A}$),

(iii) for any countable decomposition $\{\Xi_1,\Xi_2,\ldots,\Xi_n,\ldots\}$ of Ξ, $(\Xi,\Xi_n\in\mathscr{F})$, it holds that $\rho^m(F(\Xi))=\lim_{N\to\infty}\rho^m(\sum_{n=1}^N F(\Xi_n))$ $(\forall\rho^m\in\mathfrak{S}^m(\mathscr{A}^*))$.

Also, if $F(\Xi)$ is a projection for every Ξ $(\in\mathscr{F})$, a C^*-observable $(X,\mathscr{F},F)$ is called a *crisp* C^*-*observable*.

Remark 2.1 (*Sample space*)**.** Let ρ^m be a mixed state, i.e., $\rho^m\in\mathfrak{S}^m(\mathscr{A}^*)$. Applying Hopf extension theorem, we can get the measure space $(X,\overline{\mathscr{F}},\overline{\rho^m(F(\cdot))})$ such that $\overline{\rho^m(F(\Xi))}=\rho^m(F(\Xi))$ for all $\Xi\in\mathscr{F}$, where $\overline{\mathscr{F}}$ is the smallest σ-field that contains $\mathscr{F}$. For simplicity, the $\overline{\rho^m(F(\cdot))}$ is also denoted by $\rho^m(F(\cdot))$ or ${}_{\mathscr{A}^*}\langle\rho^p,F(\cdot)\rangle_{\mathscr{A}}$. Also Axiom 1 (or Method 1) proposed later makes us call the measure space $(X,\overline{\mathscr{F}},\overline{\rho^m(F(\cdot))})$ a *sample space* or a *probability space*.

Let $\boldsymbol{O}\equiv(X,\mathscr{F},F)$ be an observable in a commutative C^*-algebra $\mathscr{A}$ $(\equiv C(\Omega))$. Note that, for any fixed Ξ $(\in\mathscr{F})$, the $F(\Xi)$ is a membership function on Ω, i.e., a continuous function on Ω such that $0\leqslant[F(\Xi)](\omega)\leqslant 1$ $(\forall\omega\in\Omega)$. Thus, the $F(\cdot)$ will be usually denoted by $F_{(\cdot)}$, that is, $[F(\Xi)](\omega)=F_\Xi(\omega)$ $(\forall\Xi,\forall\omega)$. And therefore, $(X,\mathscr{F},F)$ is often denoted by $(X,\mathscr{F},F_{(\cdot)})$ in a commutative C^*-algebra $C(\Omega)$.

Let $\boldsymbol{O} \equiv (X, \mathscr{F}, F)$ be an observable in a C^*-algebra $\mathscr{A}$. Let Y be a set with the field $\mathscr{G}$. Consider a measurable map $h: X \to Y$, i.e., $h^{-1}(\Gamma) \in \mathscr{F}$ $(\forall \Gamma \in \mathscr{G})$. Then we have the observable $\boldsymbol{O}_h \equiv (Y, \mathscr{G}, F \circ h^{-1})$ in $\mathscr{A}$, where $(F \circ h^{-1})(\Gamma) = F(h^{-1}(\Gamma))$ $(\forall \Gamma \in \mathscr{G})$. This $\boldsymbol{O}_h$ is called the *image observable of* $\boldsymbol{O}$ *concerning the map* h.

For each $k = 1, 2, \ldots, n$, consider an observable $\boldsymbol{O}_k \equiv (X_k, \mathscr{F}_k, F_k)$ in a C^*-algebra $\mathscr{A}$. Define the $\times_{k=1}^n \mathscr{F}_k$ such as the smallest field (on $\times_{k=1}^n X_k$) that contains $\times_{k=1}^n \Xi_k$, $\Xi_k \in \mathscr{F}_k$. An observable $\boldsymbol{O} \equiv (\times_{k=1}^n X_k, \times_{k=1}^n \mathscr{F}_k, F)$ in $\mathscr{A}$ is called the *quasi-product observable of* $\{\boldsymbol{O}_k: k = 1, 2, \ldots, n\}$ if it holds that $\boldsymbol{O}_{h_k} = \boldsymbol{O}_k$ $(\forall k = 1, \ldots, n)$ where $\boldsymbol{O}_{h_k}$ is the image observable concerning the kth coordinate map, i.e., $\times_{j=1}^n X_j \ni (x_j)_{j=1}^n \overset{h_k}{\mapsto} x_k \in X_k$. Note that the existence and the uniqueness of the quasi-product observable of $\{\boldsymbol{O}_k: k = 1, 2, \ldots, n\}$ are not guaranteed in general. However, when $\boldsymbol{O}_k$, $k = 1, 2, \ldots, n$, commute, i.e., $F_k(\Xi_k) F_{k'}(\Xi_{k'}) = F_{k'}(\Xi_{k'}) F_k(\Xi_k)$ for all $\Xi_k \in \mathscr{F}_k$, $\Xi_{k'} \in \mathscr{F}_{k'}$ such that $k \neq k'$, we can construct F such that $F(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1) F_2(\Xi_2) \ldots F_n(\Xi_n)$. This kind of quasi-product observable is called a *direct product observable* (or in short, *product observable*), and denoted by $\times_{k=1}^n \boldsymbol{O}_k$ (or $(\times_{k=1}^n X_k, \times_{k=1}^n \mathscr{F}, \times_{k=1}^n F_k)$). In this paper we always deal with direct product observables. However, it should be noted that various quasi-product observables play important roles in measurement theory (cf. [6–8]).

With any *system* S, a C^*-algebra $\mathscr{A}$ can be associated in which the fuzzy measurement theory (or more generally, the system theory) of that system can be formulated. A *state* of the system S is represented by a pure state ρ^p $(\in \mathfrak{S}^p(\mathscr{A}^*))$, a *quantity* is represented by a self-adjoint element Q in the C^*-algebra $\mathscr{A}$. Also, an *observable* is represented by a C^*-observable $\boldsymbol{O} \equiv (X, \mathscr{F}, F)$ in the C^*-algebra $\mathscr{A}$. The *measurement of the observable* $\boldsymbol{O}$ *for the system* S *with the state* ρ^p is represented by $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$ in the C^*-algebra $\mathscr{A}$. In particular, the measurement of a quasi-product observable (resp. direct product observable) is called a *simultaneous measurement* (resp. *iterated measurement*).

The axiom presented below is analogous to (or, a kind of generalization of) Born's probabilistic interpretation of quantum mechanics. We of course assert that the axiom is a principle for all measurements, i.e., classical and quantum measurements (cf. [6,7]).

Axiom 1 (*Measurement axiom*)**.** Consider a measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F), S_{[\rho^p]})$ formulated in a C^*-algebra $\mathscr{A}$. Assume that the measured value x $(\in X)$ is obtained by the measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$. Then, the probability that the x $(\in X)$ belongs to a set Ξ $(\in \overline{\mathscr{F}})$ is given by $\rho^p(F(\Xi))$ $(\equiv {}_{\mathscr{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathscr{A}})$ (cf. Remark 2.1).

We introduce the following classification in measurement theory (cf. [6,7]):

$$\text{measurement theory} \begin{cases} \text{classical measurement theory} \\ \quad \text{(for classical systems)}, \\ \text{quantum measurement theory} \\ \quad \text{(for quantum systems)}, \end{cases}$$

where a C^*-algebra $\mathscr{A}$ is commutative or non-commutative. Note that quantum measurement theory is well known as a principle of quantum mechanics (cf. [2,3]). Our interest in this paper is mainly concentrated to classical systems. Therefore, in most cases, it suffices to consider that $\mathscr{A} = C(\Omega)$.

Remark 2.2 (*Several results derived from Axiom* 1, cf. [6–8]). We believe that this axiom dominates all measurements, i.e., classical and quantum measurements. In fact, as consequences of Axiom 1 (or, Methods 1 and 2 mentioned later), in [6,7] we clarified several fundamental facts, for example, the justification of "standard syllogism", ergodic problem (i.e., the principle of equal weight in statistical mechanics), the subjective foundation of Shannon's entropy, the errors in Heisenberg's uncertainty relation and so on. And furthermore, the relation between Kolmogorov's probability theory and Axiom 1 (or, Methods 1 and 2 mentioned later) was well discussed in [7]. The probability space $(X, \overline{\mathscr{F}}, \rho^p(F(\cdot)))$, for the first time, acquires a reality under Axiom 1. Also, in [8] we asserted that measurement theory, i.e., Axiom 1, had great power of expression. Therefore, a good translation from "natural language" into "system theoretical language" can be expected. We believe that this is the essence of "fuzzy logic" (cf. Remark 4.6 later).

Let $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F), S_{[\rho^p]})$ be a measurement formulated in a C^*-algebra $\mathscr{A}$. Assume that

(♯) we get the measured value x_0 $(\in X)$ by the measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$.

Then, we may also say that "x_0 *is the value of an observable* $\boldsymbol{O}$ *for the system S with a state* ρ^p". Let $\boldsymbol{O}' \equiv (X, \mathscr{F}', F')$ be an observable in $\mathscr{A}$ such that $\mathscr{F}' \subseteq \mathscr{F}$ and $F(\Xi) = F'(\Xi)$ $(\forall \Xi \in \mathscr{F}')$. That is, $\boldsymbol{O}$ is finer than $\boldsymbol{O}'$. Then, under the assumption (♯), we may also say that "x_0 *is the value of an observable* $\boldsymbol{O}'$ *for the system S with a state* ρ^p". Define the equivalence relation $x_1 \sim_{\mathscr{F}'} x_2$ $(x_1, x_2 \in X)$ such that $x_1 \in \Xi \Leftrightarrow x_2 \in \Xi$ $(\forall \Xi \in \mathscr{F}')$. When $x_0 \sim_{\mathscr{F}'} x_1$, it is natural to consider that "x_0 *is the value of the observable* $\boldsymbol{O}'$ *for the state* ρ^p" $\Leftrightarrow$ "x_1 *is the value of the observable* $\boldsymbol{O}'$ *for the state* ρ^p". For any $x \in X$, define $[x]_{\mathscr{F}'}$ $(\subseteq X)$ such that $[x]_{\mathscr{F}'} = \{x' \in X : x \sim_{\mathscr{F}'} x'\}$. Then we may also say that "$[x_0]_{\mathscr{F}'}$ *is the value of the observable* $\boldsymbol{O}'$ *for the state* ρ^p". Also consider a measurable map from X into Y (with the field $\mathscr{G}$). Thus, we get the field $\mathscr{F}' \equiv \{h^{-1}(\Gamma) : \Gamma \in \mathscr{G}\}$ $(\subseteq \mathscr{F})$. Note that "$h(x_1) = h(x_2)$" implies that "$x_1 \sim_{\mathscr{F}'} x_2$". Therefore, under the above assumption (♯), we may say that "$h(x_0)$ *is the value of the image observable* $\boldsymbol{O}_h$ $(\equiv (Y, \mathscr{G}, F(h^{-1}(\cdot)))$ *for the state* ρ^p".

Remark 2.3 (*Measurement theory and statistics*). As mentioned in Remark 2.1, a measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F), S_{[\rho^p]})$ always determines the sample space $(X, \overline{\mathscr{F}}, {}_{\mathscr{A}^*}\langle \rho^p, F(\cdot) \rangle_{\mathscr{A}})$. Here note that the mathematical structure of the sample space $\{{}_{\mathscr{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathscr{A}}\}_{\rho^p \in \mathfrak{S}^p(\mathscr{A}^*), \Xi \in \overline{\mathscr{F}}}$ is the same as that of the conventional formulation of statistics (i.e. $\{P(\Xi, \theta)\}_{\theta \in \Theta, \Xi \in \overline{\mathscr{F}}}$, where, for each θ in a parameter space Θ, $P(\cdot, \theta)$ is a probability measure on a measurable space $(X, \overline{\mathscr{F}})$, cf. [16]). Therefore, there is a good hope that statistics can be described in terms of measurements. Also, this is precisely our motivation in this paper. Note that measurement theory has a principle, i.e., Axiom 1, in which the relation between "measurement" and "probability" is declared. On the other hand, the meaning of "probability" is not clear in the conventional formulation of statistics since "mathematics" can be always interpreted by various ways. (Also, see the arguments appearing below Method 1 later.) Following the common knowledge of quantum mechanics, we believe that any scientific statement including the term "probability" is not meaningful without the concept of "measurement".

Remark 2.4 (*Possibility*). In most cases of measurements, we do not have the information concerning the state ρ^p of the systems S. Note that one of the main topics in statistics is to infer the unknown state from the measured value (cf. Section 3). Therefore, the $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$ is sometimes denoted by $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]})$ when we want to stress the situation that the state ρ^p is unknown. The following statement (i) is clearly equivalent to Axiom 1.

(i) Assume the fact that the measured value obtained by $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F), S_{[*]})$ belongs to Ξ $(\in \overline{\mathscr{F}})$. Then we can assert the following statement: if $[*] = \rho^p$ (i.e., $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]}) = \boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$), then the probability that the fact (i.e., the measured value belongs to Ξ) occurred is given by ${}_{\mathscr{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathscr{A}}$.

This statement (i) is usually represented as follows.

(ii) Assume the fact that the measured value obtained by $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]})$ belongs to Ξ $(\in \overline{\mathscr{F}})$. Then the possibility that $[*] = \rho^p$ is given by ${}_{\mathscr{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathscr{A}}$.

This (ii) should be read as the abbreviation of the above (i). In other words, the definition of "possibility" is given in the above (i). Therefore, three statements (i.e., Axiom 1, the statements (i) and (ii)) are equivalent. Also, we must be careful for quantum measurements since the "reduction of wave packet" is usually considered to occur after the quantum measurement. That is, the unknown state $[*]$ in the above statement (i) (and therefore, (ii)) is the state before the measurement (cf. Remark 3.1 later).

Remark 2.5 (*Possibility and likelihood*). Assume that there exists a measure ν on $(X, \overline{\mathscr{F}})$ and $f(\cdot, \rho^p) \in L^1(\Omega, \nu)$ $(\forall \rho^p \in \mathfrak{S}^p(\mathscr{A}^*))$ such that $\rho^p(F(\Xi)) = \int_\Xi f(x, \rho^p) \nu(\mathrm{d}x)$ for all $\Xi \in \overline{\mathscr{F}}$ and all ρ^p in $\mathfrak{S}^p(\mathscr{A}^*)$. Then, even if $\Xi = \{x\}$ and $\rho^p(F(\{x\}) = 0)$ $(\forall \rho^p \in \mathfrak{S}^p(\mathscr{A}^*))$ in the statement (ii) of Remark 2.4, we may calculate as follows:

$$\frac{\text{the possibility that } [*] = \rho_1^p}{\text{the possibility that } [*] = \rho_2^p} = \frac{\rho_1^p(F(\{x\}))}{\rho_2^p(F(\{x\}))}$$

$$= \lim_{\Xi \to \{x\}} \frac{\rho_1^p(F(\Xi))}{\rho_2^p(F(\Xi))} = \frac{f(x, \rho_1^p)}{f(x, \rho_2^p)}.$$

In this sense (or, in the sense of "Radon–Nikodým derivative"), we can use "likelihood function $f(x,\cdot)$" instead of "possibility function ${}_{\mathscr{A}^*}\langle\cdot, F(\Xi)\rangle_{\mathscr{A}}$". In this paper we are not concerned with "likelihood" since we consider that basic ideas should be first described in terms of "possibility".

Next we introduce "subjective measurement" $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S(\rho^m))$, which is a mathematical symbol such that ρ^m is a mixed state, i.e., $\rho^m \in \mathfrak{S}^m(\mathscr{A}^*)$. The ρ^m is called a *subjective state* (or, *statistical state*, *weight*, *prior*). Note that "subjective measurement" has no reality in itself since the state of a system S is always represented by a pure state ρ^p and not a mixed state ρ^m (cf. Axiom 1). That is, we have no experiment that tests Method 1 (presented below) directly. Therefore, Method 1, i.e., "subjective measurement theory", is meaningless without a proper interpretation (cf. [7]).

Method 1 (Subjective measurement). *Consider a subjective measurement* $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F), S(\rho^m))$ *formulated in a* C^**-algebra* $\mathscr{A}$. *Then, we consider that*

(♯) *the "subjective probability" that* x $(\in X)$, *the measured value by the subjective measurement* $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S(\rho^m))$, *belongs to a set* Ξ $(\in \overline{\mathscr{F}})$ *is given by* $\rho^m(F(\Xi))$ $(\equiv {}_{\mathscr{A}^*}\langle \rho^m, F(\Xi)\rangle_{\mathscr{A}})$.

Though a subjective measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S(\rho^m))$ is merely a mathematical symbol, it must not be underestimated. In fact statistical mechanics is based on Method 1 with a proper interpretation (i.e., "the principle of equal weight in statistical mechanics"), cf. [7]. Also, in [7] we showed that some objective interpretation (based on Axiom 1) could always be added to the subjective measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S(\rho^m))$ if it was needed. In other words, the "subjective probability" in Method 1 can be characterized as a frequency probability of the objective measurement $\boldsymbol{M}_{\otimes_{k=1}^n \mathscr{A}}(\bigotimes_{k=1}^n \boldsymbol{O}, S_{[\otimes_{k=1}^n \rho_k^p]})$ in a tensor C^*-algebra $\bigotimes_{k=1}^n \mathscr{A}$, where $\rho^m \approx (1/n)\sum_{k=1}^n \rho_k^p$ for sufficiently large n. Note that "probability space $(X, \overline{\mathscr{F}}, P)$" in Kolmogorov's probability theory (and consequently, $\{P(\Xi,\theta)\}_{\theta\in\Theta,\Xi\in\overline{\mathscr{F}}}$ in the conventional formulation of statistics) is also a mathematical symbol. As Kolmogorov himself said so in his famous book [11], the statement S_K: "*The probability that an event* Ξ $(\in \overline{\mathscr{F}})$ *occurs is given by* $P(\Xi)$" is meaningless without a proper interpretation. In this sense, the S_K and Method 1 are similar. To be compared with the statement S_K, Method 1 has a merit such that it has a form like Axiom 1.

We also consider a mathematical symbol $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]}(\rho^m))$, which is called an *objective and subjective measurement* in $\mathscr{A}$. That is, we consider that $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]}(\rho^m)) = \boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$ from the objective point of view, and $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]}(\rho^m)) = \boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S(\rho^m))$ from the subjective point of view. Therefore, the phrase "*measured value obtained by* $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]}(\rho^m))$" in Method 1 is meaningful, that is, it is interpreted as "*measured value obtained by* $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$". By the same reason mentioned in Remark 2.4, the $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]}(\rho^m))$ is also written by $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]}(\rho^m))$. In this paper, a subjective measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S(\rho^m))$ is chiefly used as the subjective part of $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]}(\rho^m))$. However, we must again note that it is merely one of the interpretations of Method 1.

The difference between "objectivity" and "subjectivity" is rather delicate. The following example will promote a better understanding of our theory.

Example 2.6 (*Objective and subjective aspects of* "*at random*"). Suppose we have an urn that contains 10 balls; six blue and four red. Now we consider the following two measurements:

(I) Consider the following measurement: that is, "choose a ball at random from the urn, and uninterruptedly look at the ball".

(II) Choose a ball at random from the urn. Assume that the information of the ball is unknown since it is held in one's fist. Here, consider the following measurement: that is, "look at the ball".

Our present problem is to formulate these two measurements (I) and (II).

The above example is solved in what follows. Consider the state space $\overline{\Omega} \approx \mathscr{M}_{+1}^p(\overline{\Omega})$, in which the state of a ball B_j $(j = 1,\ldots,10)$ is represented. Therefore, we have the correspondence: $\boldsymbol{B} \equiv \{B_1,\ldots,B_{10}\} \ni B_j \mapsto \omega_j^0 \in \overline{\Omega}$. Since our interest is concentrated to the set $\boldsymbol{B}$, it suffices to consider the restricted state space $\Omega \equiv \{\omega_1^0,\ldots,\omega_{10}^0\}$ $(\subset \overline{\Omega})$. That is, we have the identification: $\boldsymbol{B} \ni B_j \leftrightarrow \omega_j^0 \in \Omega$. Also assume that B_j is blue $(j = 1,2,\ldots,6)$ and the others are red. Therefore, we can define the observable

$\boldsymbol{O} \equiv (X=\{b,r\}, \mathscr{P}(X), F_{(\cdot)})$ in $C(\Omega)$ such that

$$F_{\{b\}}(\omega_j^0) = \begin{cases} 1 & \text{if } 1 \leqslant j \leqslant 6, \\ 0 & \text{if } 7 \leqslant j \leqslant 10, \end{cases}$$

and $F_{\{r\}}(\omega_j^0) = 1 - F_{\{b\}}(\omega_j^0)$. Here we of course consider that "observing that the ball B_j is blue" $\Leftrightarrow$ "getting the measured value 'b' by the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[\delta_{\omega_j^0}]})$". Thus, Axiom 1 says, for example, that the probability that the measured value 'b' is obtained by the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[\delta_{\omega_5^0}]})$ is given by 1 $(=\delta_{\omega_5^0}(F_{\{b\}}) = F_{\{b\}}(\omega_5^0))$.

(I) The system S (or precisely, $S_{(\mathrm{I})}$) composed of 10 balls is formulated in a tensor commutative C^*-algebra $\bigotimes_{j=1}^{10} C(\Omega) = C(\Omega^{10})$. (Recall many particles system in classical mechanics.) The state of the system S is of course represented by a point measure $\bigotimes_{j=1}^{10} \delta_{\omega_j^0} = \delta_{(\omega_1^0, \omega_2^0, \ldots, \omega_{10}^0)}$ in $\mathscr{M}_{+1}^p(\Omega^{10}) \approx \Omega^{10}$. Define the observable $\widehat{\boldsymbol{O}} \equiv (X=\{b,r\}, \mathscr{P}(X), \widehat{F}_{(\cdot)})$ in $C(\Omega^{10})$ such that

$$\begin{aligned} &\widehat{F}_{\{b\}}(\omega_1, \omega_2, \ldots, \omega_{10}) \\ &\quad = \tfrac{1}{10}(F_{\{b\}}(\omega_1) + F_{\{b\}}(\omega_2) + \cdots + F_{\{b\}}(\omega_{10})) \\ &\qquad \forall (\omega_1, \ldots, \omega_{10}) \in \Omega^{10}, \end{aligned}$$

and $\widehat{F}_{\{r\}} = 1 - \widehat{F}_{\{b\}}$. This $\widehat{\boldsymbol{O}}$ may be called "*average observable*" *in* $C(\Omega^{10})$. Now, we have the measurement $\boldsymbol{M}_{C(\Omega^{10})}(\widehat{\boldsymbol{O}}, S_{[\delta_{(\omega_1^0, \omega_2^0, \ldots, \omega_{10}^0)}]})$, which represents the situation (I), i.e., the objective view of "at random". Clearly, the (objective) probability that a measured value 'b' [resp. 'r'] is obtained by the measurement $\boldsymbol{M}_{C(\Omega^{10})}(\widehat{\boldsymbol{O}}, S_{[\delta_{(\omega_1^0, \omega_2^0, \ldots, \omega_{10}^0)}]})$ is given by $\delta_{(\omega_1^0, \omega_2^0, \ldots, \omega_{10}^0)}(\widehat{F}_{\{b\}}) = \widehat{F}_{\{b\}}(\omega_1^0, \omega_2^0, \ldots, \omega_{10}^0) = \frac{6}{10}$ [resp. $\delta_{(\omega_1^0, \omega_2^0, \ldots, \omega_{10}^0)}(\widehat{F}_{\{r\}}) = \frac{4}{10}$].

(II) Assume that the ball chosen in the statement (II) is B_{j_0}. Also note that the information of the j_0 is unknown. Therefore, there is a very good reason (cf. [7]) to consider that the subjective state of the system S (or precisely, $S_{(\mathrm{II})}$) is represented by the uniform weight ρ_{uni}^m $(\in \mathscr{M}_{+1}(\Omega))$, that is, $\rho_{\mathrm{uni}}^m = \frac{1}{10}\sum_{j=1}^{10} \delta_{\omega_j^0}$. Thus, we have the objective and subjective measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[*]}(\rho_{\mathrm{uni}}^m))$. (Here the observer does not know the fact that $[*] = \delta_{\omega_{j_0}^0}$.) Then the subjective probability that a measured value 'b' [resp. 'r'] is obtained by the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[*]}(\rho_{\mathrm{uni}}^m))$ is given by $\rho_{\mathrm{uni}}^m(F_{\{b\}}) = \int_\Omega F_{\{b\}}(\omega)\rho_{\mathrm{uni}}^m(\mathrm{d}\omega) = \frac{6}{10}$ [resp. $\rho_{\mathrm{uni}}^m(F_{\{r\}}) = \frac{4}{10}$].

3. Statistical inferences for states

In this section we study "statistical inferences for states" in measurement theory. In other words, we focus on the following problem:

($\sharp$) how to infer the unknown state $[*]$ $(\in \mathfrak{S}^p(\mathscr{A}^*))$ from the measured data obtained by a measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]})$ or $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]}(\rho^m))$.

Also this should be of course compared to Fisher's method or Bayes's method in statistics.

Let us begin with Fisher's method. The statement (ii) in Remark 2.4 gives the justification to "maximum likelihood function method" as follows (cf. Remark 2.5).

(M) When we know that the measured value by a measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F),\ S_{[*]})$ belongs to Ξ, there is a very good reason to consider that the state $[*]$ of the system S is equal to ρ_0^p $(\in \mathfrak{S}^p(\mathscr{A}^*))$ such that $\rho_0^p(F(\Xi)) = \max_{\rho^p \in \mathfrak{S}^p(\mathscr{A}^*)} \rho^p(F(\Xi))$.

Also, by the statement (i) in Remark 2.4 we get the following test (T).

(T) Assume that Ξ_0 $(\in \mathscr{F})$, disjoint sets H_0 and H_1 $(\subseteq \mathfrak{S}^p(\mathscr{A}^*))$ satisfy that $0 < \rho^p(F(\Xi_0)) < \varepsilon \ll 1$ $(\forall \rho^p \in H_0)$ and $0 \ll 1 - \varepsilon' < \rho^p(F(\Xi_0)) \leqslant 1$ $(\forall \rho^p \in H_1)$ for some sufficiently small ε and ε'. And assume the fact that the measured value by a measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[*]})$ belongs to Ξ_0. Here we see that, if $[*] \in H_0$, then the probability that the fact occurred is less than ε, that is, the fact is a rare occurrence. Therefore, there is a very good reason to consider that $[*] \notin H_0$.

Thus we consider that Fisher's spirit is described in terms of Axiom 1. The statement (M) (or (T)) is of course valid for quantum systems as well as classical systems. However, as stated in Remark 2.4, we must note that the unknown state $[*]$ in the statement (M) (or (T)) is the state before the measurement (cf. Remark 3.1 below). Though it may be one of the topics in quantum measurement theory, our concern in this paper is concentrated to classical systems.

Remark 3.1 (*The "collapse of the wave pocket" in quantum mechanics*). The "collapse of the wave pocket" is the most significant and unsolved problem in quantum mechanics. That is, some physicists consider that the state $[*]$ $(\in \mathfrak{S}^p(\mathscr{A}^*))$ of the system S changes to some state $[*']$ $(\in \mathfrak{S}^p(\mathscr{A}^*))$ after we know the measured value by the measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O} \equiv (X, \mathscr{F}, F), S_{[*]})$. Though there is a very good reason that they consider so, this produces several famous paradoxes such as Schrödinger's cat. Note that the above (M) also includes a paradox if we consider that the ρ_0^p in the above (M) approximates $[*']$. However, in spite of the paradox, some physicists may assert so. Their opinion must not be denied since it is unsolved in quantum mechanics.

Next, we review Bayes's method (proposed in [7], i.e., the formula (5.2) in [7]) in measurements. Let $\boldsymbol{O} \equiv (X, \mathscr{F}, F_{(\cdot)})$ be an observable in a commutative C^*-algebra $C(\Omega)$. Consider an objective and subjective measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[*]}(\rho^m))$. Then the following statement is justified by Method 1 (or Method 2 presented in the next Section 4).

(B) Assume the fact that the measured value by the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[*]}(\rho^m))$ belongs to Ξ $(\in \overline{\mathscr{F}})$. Then, from the subjective point of view, we consider that the new subjective state of the system S is given by ρ^m_{new} $(\in \mathscr{M}_{+1}(\Omega))$ such that

$$\rho^m_{\text{new}}(B) = \frac{\int_B F_\Xi(\omega)\rho^m(\mathrm{d}\omega)}{\int_\Omega F_\Xi(\omega)\rho^m(\mathrm{d}\omega)}$$
$$(\forall B \in \mathscr{B}_\Omega, \text{ the Borel } \sigma\text{-field of } \Omega).$$

That is, there is a very good reason to consider that the state $[*]$ is approximated by ρ^m_{new}. This (B) of course corresponds to Bayes's method in statistics. Therefore, we see that Bayes's spirit is described in terms of Method 1.

For completeness, we add the outline of the measurement theoretical justification of (B) in what follows. Let $\boldsymbol{O}_1 \equiv (Y, \mathscr{G}, G_{(\cdot)})$ be any observable in $C(\Omega)$. And consider the iterated measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O} \times \boldsymbol{O}_1, S_{[*]}(\rho^m))$. Applying Method 1 to $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O} \times \boldsymbol{O}_1, S_{[*]}(\rho^m))$ and $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}_1, S_{[*]}(\rho^m_{\text{new}}))$ respectively, we can expect that the new subjectiv state ρ^m_{new} $(\in \mathscr{M}_{+1}(\Omega))$ satisfies the following condition:

$$\frac{\int_\Omega F_\Xi(\omega) G_\Gamma(\omega)\rho^m(\mathrm{d}\omega)}{\int_\Omega F_\Xi(\omega)\rho^m(\mathrm{d}\omega)} = \int_\Omega G_\Gamma(\omega)\rho^m_{\text{new}}(\mathrm{d}\omega) \quad (\forall \Gamma \in \mathscr{G}).$$

(For the arguments about "conditional probability", see [7].) Thus we get (B) since $\boldsymbol{O}_1$ is arbitrary.

Now we have Fisher's method and Bayes's method in measurement theory. Thus we now study some statistical examples in terms of measurements. Though these examples are quite simple (i.e., X and Ω are supposed to be finite sets), we believe that these do not miss the essence of statistics.

Example 3.2 (*Urn problem*). Let U_j, $j = 1, 2, 3$, be urns that contain sufficiently many colored balls as follows:

	"blue"	"green"	"red"	"yellow"
U_1:	50%	30%	10%	10%
U_2:	30%	30%	30%	10%
U_3:	20%	20%	40%	20%

Put $\boldsymbol{U} = \{U_1, U_2, U_3\}$. By the same argument in Example 2.6, we consider the state space Ω $(\equiv \{\omega_1, \omega_2, \omega_3\})$ with the discrete topology, which is identified with $\boldsymbol{U}$, that is, $\boldsymbol{U} \ni U_j \leftrightarrow \omega_j \in \Omega \approx \mathscr{M}^p_{+1}(\Omega)$. Define the observable $\boldsymbol{O} \equiv (X = \{b, g, r, y\}, \mathscr{P}(X), F_{(\cdot)})$ in $C(\Omega)$ by the usual way. That is, $F_{\{b\}}(\omega_1) = \frac{5}{10}$, $F_{\{b\}}(\omega_2) = \frac{3}{10}$, $F_{\{y\}}(\omega_3) = \frac{2}{10}$ and so on. (Recall the "average observable" $\widehat{\boldsymbol{O}}$ in Example 2.6 (I), i.e., the objective view of "at random".) Then, we have the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[*]})$. We of course consider that, for example,

(♯) "choosing a ball at random from the urn U_j, and observing that the ball is blue" ⇔ "getting the measured value 'b' by the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[\delta_{\omega_j}]})$".

Next consider the iterated measurement $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^2 \boldsymbol{O} \equiv (X^2, \mathscr{P}(X^2), \times_{k=1}^2 F), S_{[*]})$ where $(\times_{k=1}^2 F)_{\Xi_1 \times \Xi_2}(\omega) = F_{\Xi_1}(\omega) \cdot F_{\Xi_2}(\omega)$. Also, assume that the measured value (r, b) is obtained by the iterated measurement $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^2 \boldsymbol{O}, S_{[*]})$. Applying Fisher's method (M), we get the conclusion as follows. Put $E(\omega) = F_{\{r\}}(\omega) F_{\{b\}}(\omega)$. Clearly it holds that $E(\omega_1) = 1 \times 5/10^2 = 0.05$, $E(\omega_2) = 3 \times$

$3/10^2=0.09$ and $E(\omega_3)=4\times 2/10^2=0.08$. Therefore, there is a very good reason to consider that $[*]=\delta_{\omega_2}$, that is, the unknown urn is U_2.

Example 3.3 (*Continued from Example* 3.2)**.** Let Q be a quantity in $C(\Omega)$, i.e., $Q:\Omega\ (\approx \mathscr{M}_{+1}^p(\Omega))\to \boldsymbol{R}$ is a real valued continuous function on Ω. (For the relation between quantities and observables, see Section 4.) For example, we may consider what follows. Assume that the weight of a blue ball is given by 10 (gram), and green 20, red 30 and yellow 10. (Thus, we can define the map $w:X\to\boldsymbol{R}$ such that $w(b)=10, w(g)=20, w(r)=30$ and $w(y)=10$.) Therefore, the average weight $Q(\omega_1)$ of the balls in the urn U_1 is given by 15 $(=(10\times 50+20\times 30+30\times 10+10\times 10)/100)$, and similarly, $Q(\omega_2)=19$ and $Q(\omega_3)=20$. Now we have the following problem.

(♯) *How do we infer* $Q(*)$ *from the measured value* (r,b) *obtained by the iterated measurement* $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^{2}\boldsymbol{O}, S_{[*]})$?

This problem is easily solved as follows. Since we inferred that $[*]=\delta_{\omega_2}(\leftrightarrow\omega_2)$ in Example 3.2, we can immediately conclude that $Q(*)=Q(\omega_2)=19$. Also note that the map $w:X\to\boldsymbol{R}$ is not essential in this argument, that is, it is the preparation for Example 4.3 later.

Remark 3.4 (*Parameter space*)**.** In the above example, some may use a parameter space $\Theta\equiv\{15,19,20\}$ instead of the state space Ω. This is not wrong if the parameter space Θ is regarded as a state space under the identification: $\Omega\ni\omega_j \overset{Q}{\mapsto} Q(\omega_j)\in\Theta$, that is, $C(\Theta)$ is C^*-isomorphic to $C(\Omega)$. As stated in Remark 4.6 later, measurement theory is a part of system theory. Therefore, it has the system theoretical concepts, for example, "state space", "observable", "quantity", "measurement" and so on. Recall the statements appearing above Axiom 1, and note that these concepts were described in terms of a C^*-algebra $\mathscr{A}$, in which the system is formulated. Thus, if we want to use the word "parameter space" in order to represent a certain non-mathematical concept, it should be defined in terms of the C^*-algebra $\mathscr{A}$. We consider that the word "parameter space", as well as "probability space", is not clear in the conventional formulation of statistics (cf. Remark 2.3).

Example 3.5 (*Bayes's method*)**.** Next study the problem (♯) in Example 3.3 from the subjective point of view. Consider an objective and subjective measurement $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^{2}\boldsymbol{O}, S_{[*]}(\rho_0^m))$. For example, assume that $\rho_0^m=\rho_{\text{uni}}^m$, i.e., $\rho_{\text{uni}}^m=\frac{1}{3}\sum_{j=1}^{3}\delta_{\omega_j}$ on Ω. When we get the measured value (r,b) by the measurement $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^{2}\boldsymbol{O}, S_{[*]}(\rho_0^m))$, we infer, by Bayes's method (B), that the new state is $\rho_{\text{new}}^m=1/(5+9+8)(5\delta_{\omega_1}+9\delta_{\omega_2}+8\delta_{\omega_3})$. Thus there is a very good reason to consider that $Q(*)$ is approximated by

$$\int_\Omega Q(\omega)\rho_{\text{new}}^m\,(\mathrm{d}\omega) = \frac{15.5+19.9+20.8}{5+9+8}=18.45\ldots.$$

Now let us provide another example, which is essentially the same as Example 3.2. In order to appreciate measurement theory, we must practice a lot of examples.

Example 3.6 (*At a gun shop*)**.** Let $\boldsymbol{G}\equiv\{G_1,\ldots,G_{50}\}$ be a set of guns in a gun shop. Assume that

$$\text{the percentage of "hits of a gun } G_j\text{"} = \begin{cases} 80\% & \text{if } 1\leqslant j\leqslant 30,\\ 70\% & \text{if } 31\leqslant j\leqslant 40,\\ 10\% & \text{if } 41\leqslant j\leqslant 50.\end{cases}$$

Assume the following situation (i) + (ii):

(i) Some one picks up a certain gun G_{j_0} from $\boldsymbol{G}$. He does not know the information concerning the j_0.
(ii) He shoots the gun G_{j_0} three times. First and second he hits the mark, and third he misses the mark.

Our present problem is to formulate the measurement (i) + (ii).

The above example is solved in what follows. Let Ω be a state space, which is identified with the set $\boldsymbol{G}$. That is, we have the identification: $\boldsymbol{G}\ni G_j\leftrightarrow\omega_j\in\Omega$. Define the observable $\boldsymbol{O}\equiv(X=\{0,1\},\mathscr{P}(X),F_{(\cdot)})$ in $C(\Omega)$ such that

$$F_{\{1\}}(\omega_j)=\begin{cases} 0.8 & \text{if } 1\leqslant j\leqslant 30,\\ 0.7 & \text{if } 31\leqslant j\leqslant 40,\\ 0.1 & \text{if } 41\leqslant j\leqslant 50\end{cases}$$

and $F_{\{0\}}(\omega_j) = 1 - F_{\{1\}}(\omega_j)$. Of course we consider that

(♯) "hit the mark by a gun G_{j_0}" ⇔ "get the measured value 1 by the measurement $\boldsymbol{M}_{C(\Omega)}(\boldsymbol{O}, S_{[\delta_{\omega_{j_0}}]})$".

Here, consider the (three times) iterated measurement $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^{3}\boldsymbol{O} = (X^3, \mathscr{P}(X^3), \times_{k=1}^{3}F), S_{[\delta_{\omega_{j_0}}]})$ in $C(\Omega)$ such that

$$(\times_{k=1}^{3}F)_{\Xi_1\times\Xi_2\times\Xi_3}(\omega) = F_{\Xi_1}(\omega)F_{\Xi_2}(\omega)F_{\Xi_3}(\omega)$$
$$(\forall \Xi_1\times\Xi_2\times\Xi_3 \in \mathscr{P}(X^3),\ \forall\omega\in\Omega).$$

Clearly, the above statement (ii) in Example 3.6 implies that the measured value $(1,1,0)$ is obtained by $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^{3}\boldsymbol{O}, S_{[*]})$. (The observer does not know that $[*] = \delta_{\omega_{j_0}}$.)

By a simple calculation, we see

$$F_{\{1\}}(\omega_j)F_{\{1\}}(\omega_j)F_{\{0\}}(\omega_j) = \begin{cases} 0.128 & \text{if } 1\leqslant j\leqslant 30,\\ 0.147 & \text{if } 31\leqslant j\leqslant 40,\\ 0.009 & \text{if } 41\leqslant j\leqslant 50.\end{cases}$$

Therefore, by Fisher's method (M), there is a very good reason to consider that $31 \leqslant j_0 \leqslant 40$.

Remark 3.7 (*Continued from Example* 3.6. *Test*). Let $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^{3}\boldsymbol{O}, S_{[*]})$ be as in the above arguments. Define the map $T : X^3 \equiv \{0,1\}^3 \to \{0,1\}$ such that

$$T(x_1,x_2,x_3) = \begin{cases} 1 & \text{if } x_1+x_2+x_3 \geqslant 2,\\ 0 & \text{if } x_1+x_2+x_3 < 2.\end{cases}$$

Hence, we get the image observable $\boldsymbol{O}_T = (\{0,1\}, \mathscr{P}(\{0,1\}), (\times_{k=1}^{3}F)_{T^{-1}(\cdot)})$ in $C(\Omega)$. Put $\Xi_0 = T^{-1}(\{1\})$, $H_0 = \{\delta_{\omega_j}: 41\leqslant j\leqslant 50\}$ and $H_1 = \{\delta_{\omega_j}: 1\leqslant j\leqslant 30\}$. Then, we see that

$$_{\mathscr{M}(\Omega)}\langle \delta_{\omega_j}, (\times_{k=1}^{3}F)_{T^{-1}(\{1\})}\rangle_{C(\Omega)} = \begin{cases} (0.1)^3 + 3(0.1)^2(0.9) = 0.028 & \text{if } \omega_j \in H_0,\\ (0.8)^3 + 3(0.8)^2(0.2) = 0.896 & \text{if } \omega_j \in H_1.\end{cases}$$

Clearly, the hypothesis (ii) in Example 3.6 implies that the fact "$(1,1,0) \in T^{-1}(\{1\}) \equiv \Xi_0$" occurs, that is, the value of the observable $\boldsymbol{O}_T$ for the system with the state $[*]$ is equal to 1. Then we can say, by Fisher's method (T), that

(♯) if $[*] \in H_0$, the probability that the fact (i.e., "$(1,1,0) \in T^{-1}(\{1\})$") occurred is given by 0.0271.

That is, if $[*] \in H_0$, we can say "A very rare case occurred". Therefore, there is a very good reason to consider that $[*] \notin H_0$, that is, $1 \leqslant j_0 \leqslant 40$.

Remark 3.8 (*Continued from Remark* 2.3). All examples (except Examples 2.6 and 4.2) in this paper may be easy for statisticians. That is because the mathematical structure $\{_{\mathscr{A}^*}\langle \rho^p, F(\Xi)\rangle_{\mathscr{A}}\}$ is the same as that of statistics, i.e., $\{P(\Xi,\theta)\}$. However, it should be noted that measurement theory has other aspects (cf. Remark 2.2). Also, we believe that statistics must not be one of the fields of mathematics. Thus we do not start from "Kolmogorov's probability theory" but "measurement". All results in this section are consequences of Axiom 1 or Method 1. Therefore, we conclude that Fisher's and Bayes's spirits are described in terms of measurements.

4. Approximate measurements for quantities

In this section we study "approximate measurements for quantities" in measurement theory, which corresponds to "estimation under loss function" in statistics. For this, we must study the concept of "measurement error", which was first introduced in the formulation of "Heisenberg's uncertainty relation" (cf. [7]).

In measurement theory, every measurement is supposed to be exact, that is, it does not have the concept of "error" in itself. Assume that we hope to know the value x_0 of an observable $\boldsymbol{O} \equiv (X, \mathscr{F}, F)$ for a system with the state ρ^p, but we cannot conduct the measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$. Therefore, we may take another measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}' \equiv (X, \mathscr{F}, F'), S_{[\rho^p]})$ instead of $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$. When we get the measured value x_1 by the measurement $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}', S_{[\rho^p]})$, we may regard the x_1 as the value x_0 of the observable $\boldsymbol{O}$ for the state ρ^p. Under this situation, there is a reason to consider that the "distance" between x_0 and x_1 can be regarded as the measurement error. Also, $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}', S_{[\rho^p]})$ may be called an approximate measurement of $\boldsymbol{M}_{\mathscr{A}}(\boldsymbol{O}, S_{[\rho^p]})$. Also, in this situation, we may be concerned with the subjective measurement rather than the (objective) measurement since the state ρ^p is unknown in general.

In the above arguments, we usually assume that the observable $\boldsymbol{O}$ is a crisp one, or a quantity. Note that a C^*-algebra does not have sufficient projections in general. For example, the $C([0,1])$ has only two projections, i.e., the constant functions 0 and 1. Therefore, we now introduce the W^*-algebraic formulation of measurements (cf. [7]). If we are concerned with classical systems, we can say that the C^*-algebraic formulation is topological, on the other hand, the W^*-algebraic formulation is measure theoretical.

Let $\mathcal{N}$ be a W^*-algebra (i.e., von Neumann algebra), that is, $\mathcal{N}$ is a C^*-algebra with the predual Banach space $\mathcal{N}_*$ (i.e., $\mathcal{N}=(\mathcal{N}_*)^*$). The linear functional $\overline{\rho}(T)$ (i.e., a continuous linear functional on $\mathcal{N}$ in the sense of weak*-topology $\sigma(\mathcal{N};\mathcal{N}_*)$) is also denoted by ${}_{\mathcal{N}_*}\langle\overline{\rho},T\rangle_{\mathcal{N}}$. Then, we can define the *normal state-class* $\mathfrak{S}^n(\mathcal{N}_*)$ such as

$$\mathfrak{S}^n(\mathcal{N}_*)\equiv\{\overline{\rho}\in\mathcal{N}_*\colon \|\overline{\rho}\|_{\mathcal{N}_*}=1 \text{ and } \overline{\rho}\geqslant 0 \\ \text{(i.e., } \overline{\rho}(T^*T)\geqslant 0 \ \text{ for all } T\in\mathcal{N})\}.$$

The element $\overline{\rho}$ of $\mathfrak{S}^n(\mathcal{N}_*)$ is called a *normal state*. The $B(V)$, the space of bounded linear operators on a Hilbert space V, is a typical noncommutative W^*-algebra with the predual space $\mathrm{Tr}(V)$, the space of trace operators. And we see that $\mathfrak{S}^n(B(V)_*)=\mathrm{Tr}_{+1}(V)$, i.e., the space of density operators. Also, any commutative W^*-algebra $\mathcal{N}$ is represented by $L^\infty(\Omega,\nu)$, cf. [15]. Of course its predual space is $L^1(\Omega,\nu)$. Therefore, $\mathfrak{S}^n(L^\infty(\Omega,\nu)_*)=L^1_{+1}(\Omega,\nu)\equiv\{\overline{\rho}\in L^1(\Omega,\nu)\colon\ \overline{\rho}\geqslant 0,\int_\Omega\overline{\rho}(\omega)\nu(\mathrm{d}\omega)=1\}$, i.e., the space of density functions. The Ω is also called a *state space*.

Let $\mathcal{N}$ be a W^*-algebra. A W^*-*observable* (or in short, *observable*) $\overline{\boldsymbol{O}}\equiv(X,\overline{\mathcal{F}},F)$ in $\mathcal{N}$ is defined such that it satisfies that

(i) $(X,\overline{\mathcal{F}})$ is a measurable space, that is, $\overline{\mathcal{F}}$ is a σ-field on X,

(ii) for every $\Xi\in\overline{\mathcal{F}}, F(\Xi)$ is a positive element in $\mathcal{N}$ (i.e., $0\leqslant F(\Xi)\in\mathcal{N}$) such that $F(\emptyset)=0$ and $F(X)=I$, where 0 is the 0-element and I is the identity element in $\mathcal{N}$, and

(iii) for any countable decomposition $\{\Xi_j\}_{j=1}^\infty$ of Ξ, $(\Xi_j,\Xi\in\overline{\mathcal{F}})$, $F(\Xi)=\sum_{j=1}^\infty F(\Xi_j)$ holds in the sense of the weak*-topology $\sigma(\mathcal{N};\mathcal{N}_*)$.

If $F(\Xi)$ is a projection for every Ξ $(\in\overline{\mathcal{F}})$, a W^*-observable $(X,\overline{\mathcal{F}},F)$ in $\mathcal{N}$ is called a *crisp* W^*-*observable*.

Let $\boldsymbol{R}$ and $\mathcal{B}$ be the real line and the Borel σ-field respectively. A crisp W^*-observable $\overline{\boldsymbol{Q}}\equiv(\boldsymbol{R},\mathcal{B},G)$ in a W^*-algebra $\mathcal{N}$ is called a *quantity* in $\mathcal{N}$. Consider quantum systems, that is, assume that $\mathcal{N}=B(V)$. Then, by the spectral representation theorem, we have the identification: $\overline{\boldsymbol{Q}}\equiv(\boldsymbol{R},\mathcal{B},G)\leftrightarrow\int_{\boldsymbol{R}}\lambda G(\mathrm{d}\lambda)$, i.e., (unbounded) self-adjoint operator on a Hilbert space V, that is, "quantity" = "self-adjoint operator". Next, consider classical systems, that is, $\mathcal{N}=L^\infty(\Omega,\nu)$. Then, we have the identification: $\overline{\boldsymbol{Q}}\equiv(\boldsymbol{R},\mathcal{B},G_{(\cdot)})\leftrightarrow\int_{\boldsymbol{R}}\lambda G_{\mathrm{d}\lambda}(\omega)$. Note that the $\int_{\boldsymbol{R}}\lambda G_{\mathrm{d}\lambda}(\cdot)$ is equal to a measurable function $Q:\Omega\to\boldsymbol{R}$ such that $G_\Xi(\omega)=\chi_{\{\omega'\in\Omega\colon\, Q(\omega')\in\Xi\}}(\omega)$ $(\forall\omega\in\Omega,\forall\Xi\in\mathcal{B})$, where χ_B, B $(\subseteq\Omega)$, is a characteristic function on Ω, i.e., $\chi_B(\omega)=1$, $(\omega\in B)$, $=0$, $(\omega\notin B)$. Therefore, we can see that "quantity" = "real valued measurable function on Ω" in classical measurement theory.

Let $\overline{\boldsymbol{O}}\equiv(X,\overline{\mathcal{F}},F)$ be a W^*-observable in $\mathcal{N}$ and let $\overline{\rho}\in S^n(\mathcal{N}_*)$. Then, the symbol $\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{O}},S(\overline{\rho}))$ is called a *subjective W^*-measurement* (or in short, *W^*-measurement*) in $\mathcal{N}$. And the normal state $\overline{\rho}$ is called a *subjective state*. Also, the measure space $(X,\overline{\mathcal{F}},\overline{\rho}(F(\cdot)))$ is called a *sample space*.

The following "method" is a W^*-algebraic version of Method 1, cf. [7]. Therefore, it should be used like Method 1.

Method 2 [W^*-measurement]. *Consider a W^*-measurement $\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{O}}\equiv(X,\overline{\mathcal{F}},F),S(\overline{\rho}))$ in a W^*-algebra $\mathcal{N}$. Then, we consider that*

(♯) *"subjective probability" that x $(\in X)$, the measured value obtained by the W^*-measurement $\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{O}},S(\overline{\rho}))$, belongs to a set Ξ $(\in\overline{\mathcal{F}})$ is given by $\overline{\rho}(F(\Xi))(\equiv{}_{\mathcal{N}_*}\langle\overline{\rho},F(\Xi)\rangle_{\mathcal{N}})$.*

Now let us define "measurement error" in what follows. Let $\overline{\boldsymbol{Q}}\equiv(\boldsymbol{R},\mathcal{B},G)$ be a crisp W^*-observable (i.e., quantity) in $\mathcal{N}$. Let $\overline{\boldsymbol{O}}\equiv(\boldsymbol{R},\mathcal{B},F)$ be a W^*-observable in $\mathcal{N}$ such that $\overline{\boldsymbol{Q}}$ and $\overline{\boldsymbol{O}}$ commute. Let $\overline{\boldsymbol{Q}}\times\overline{\boldsymbol{O}}\equiv(\boldsymbol{R}^2,\mathcal{B}^2,G\times F)$ be the product observable of $\overline{\boldsymbol{Q}}$ and $\overline{\boldsymbol{O}}$. Consider the iterated measurement $\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{Q}}\times\overline{\boldsymbol{O}},\ S(\overline{\rho}))$. According to Method 2, the (subjective) probability that the measured value (λ_1,λ_2) $(\in\boldsymbol{R}^2)$ belong to $\Xi_1\times\Xi_2$ $(\in\mathcal{B}^2)$ is given by

$\overline{\rho}((G \times F)(\Xi_1 \times \Xi_2))$. Then we have the following definition.

Definition 4.1 (*Measurement error*, cf. Ishikawa [7]). Assume the above notations. And assume the situation that we hope to approximate the value λ_1 of the quantity $\overline{\boldsymbol{Q}}$ by the value λ_2 of the observable $\overline{\boldsymbol{O}}$, that is, $\overline{\boldsymbol{O}}$ is the approximation of $\overline{\boldsymbol{Q}}$. Then the measurement error, $\Delta(\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}, S(\overline{\rho})))$, is defined by

$$\Delta(\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}, S(\overline{\rho}))) = \left[\int\int_{\boldsymbol{R}^2} |\lambda_1 - \lambda_2|^2 \, \overline{\rho}((G \times F)(\mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2)) \right]^{1/2}.$$

This is also called the distance between $\overline{\boldsymbol{Q}}$ and $\overline{\boldsymbol{O}}$ concerning $\overline{\rho}$.

Let $\overline{\boldsymbol{Q}} \equiv (\boldsymbol{R}, \mathcal{B}, G)$ and $\overline{\boldsymbol{O}} \equiv (X, \overline{\mathcal{F}}, F)$ be a quantity and a W^*-observable in a W^*-algebra $\mathcal{N}$, respectively. Consider the measurable map $h : X \to \boldsymbol{R}$, and the image observable $\overline{\boldsymbol{O}}_h \equiv (\boldsymbol{R}, \mathcal{B}, F(h^{-1}(\cdot)))$ in $\mathcal{N}$. Also assume that $\overline{\boldsymbol{Q}}$ and $\overline{\boldsymbol{O}}_h$ commute. Thus, the distance between $\overline{\boldsymbol{Q}}$ and $\overline{\boldsymbol{O}}_h$ (concerning $\overline{\rho} \in \mathfrak{S}^n(\mathcal{N}_*)$) is defined by $\Delta(\boldsymbol{M}_{\mathcal{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}_h, S(\overline{\rho})))$ as in the above definition. Now, we have the following problem:

(♯) how to choose a proper image observable $\overline{\boldsymbol{O}}_h$ (i.e., $\overline{\boldsymbol{O}}$ and h) as the approximation of $\overline{\boldsymbol{Q}}$.

Our interest in this section is concentrated to the problem (♯). Note that this (♯) is entirely different from Fisher's and Bayes's spirits in Section 3, that is, how to infer the unknown state from the measured data obtained by a measurement.

Concerning the above problem (♯), we can state Heisenberg's uncertainty relation in what follows.

Example 4.2 (*Heisenberg's uncertainty relation*, cf. [4,5,7]). Let Q_1 and Q_2 be a position quantity, and a momentum quantity, respectively (i.e. Q_1 and Q_2 are self-adjoint operators on a Hilbert space V satisfying that $Q_1Q_2 - Q_2Q_1 = i\hbar$, $\hbar$ is the Plank constant). As mentioned before, we identify Q_i with the spectral measure $\overline{\boldsymbol{Q}}_i \equiv (\boldsymbol{R}, \mathcal{B}, G_i)$ in $B(V)$, i.e., $Q_i = \int_{\boldsymbol{R}} \lambda G_i(\mathrm{d}\lambda)$. Since Q_1 and Q_2 do not commute, the quasi-product observable does not exist. Therefore, consider an observable $\overline{\boldsymbol{O}} \equiv (X, \overline{\mathcal{F}}, F)$ in $B(V)$ and measurable maps $h_i : X \to \boldsymbol{R}$, $(i = 1, 2)$, and define the image observables $\overline{\boldsymbol{O}}_{h_i} \equiv (\boldsymbol{R}, \mathcal{B}, F(h_i^{-1}(\cdot)))$ in $B(V)$. And furthermore, assume the conditions: (i) $\int_{\boldsymbol{R}} \lambda \langle \psi, \; G_i(\mathrm{d}\lambda)\psi \rangle_V = \int_{\boldsymbol{R}} \lambda \langle \psi, \; F(h_i^{-1}(\mathrm{d}\lambda))\psi \rangle_V$ ($\forall \psi \in \bigcap_{i=1}^2$"the domain of Q_i"), (ii) $\overline{\boldsymbol{Q}}_i$ and $\overline{\boldsymbol{O}}_{h_i}$ commute. Then we get the following inequality:

$$\Delta(\boldsymbol{M}_{B(V)}(\overline{\boldsymbol{Q}}_1 \times \overline{\boldsymbol{O}}_{h_1}, S(\overline{\rho}))) \times \Delta(\boldsymbol{M}_{B(V)}(\overline{\boldsymbol{Q}}_2 \times \overline{\boldsymbol{O}}_{h_2}, S(\overline{\rho}))) \geqslant \hbar/2$$

$$\text{for all } \overline{\rho} \in \mathrm{Tr}_{+1}(V).$$

This is just Heisenberg's uncertainty relation, which was discovered by Heisenberg in the famous thought experiment of γ-rays microscope (cf. [14]).

The following example is a main part of this section. The reader should find "estimation under loss function in statistics" in the following example.

Example 4.3 (*Continued from Examples* 3.2, 3.3 *and* 3.5 "*Urn problem*"). Let $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^2 \boldsymbol{O}, S_{[*]}(\rho_0^m))$ and $Q : \Omega \to \boldsymbol{R}$ be as in Example 3.5. That is, $\boldsymbol{O} = (X = \{b, r, w, y\}, \mathcal{P}(X), F_{(\cdot)})$ in $C(\Omega)$ ($\equiv C(\{\omega_1, \omega_2, \omega_3\})$) and $\rho_0^m \in \mathcal{M}_{+1}(\Omega)$. Consider a measure v on Ω, for example, $v(\{\omega_j\}) = 1$ $(j = 1, 2, 3)$. Define the W^*-observable $\overline{\boldsymbol{O}}$ in $L^\infty(\Omega, v)$ such that $\overline{\boldsymbol{O}} = \boldsymbol{O}$, and define the normal state $\overline{\rho}$ $(\in L^1_{+1}(\Omega, v))$ such that $\rho_0^m(B) = \int_B \overline{\rho}(\omega) v(\mathrm{d}\omega)$ for all B $(\subseteq \Omega)$. Then, we can identify $\boldsymbol{M}_{C(\Omega)}(\times_{k=1}^2 \boldsymbol{O}, S_{[*]}(\rho_0^m))$ with $\boldsymbol{M}_{L^\infty(\Omega, v)}(\times_{k=1}^2 \overline{\boldsymbol{O}}, S(\overline{\rho}))$. Note that Q is equivalent to the crisp observable $\overline{\boldsymbol{Q}} \equiv (\boldsymbol{R}, \mathcal{B}, G^Q)$ in $L^\infty(\Omega, v)$ such that $G^Q_\Xi(\omega) = \chi_{\{\omega' \in \Omega : \, Q(\omega') \in \Xi\}}(\omega)$ for all $\Xi \in \mathcal{B}$ and all $\omega \in \Omega$. Define the map $h : X^2 \to \boldsymbol{R}$ such that

$$h(x_1, x_2) = \tfrac{1}{2}(w(x_1) + w(x_2)) \quad (\forall (x_1, x_2) \in X^2 \equiv \{b, r, w, y\}^2), \tag{4.1}$$

where $w(b) = 10$, $w(g) = 20$, $w(r) = 30$ and $w(y) = 10$ (cf. Example 3.3). Consider the image observable $(\times_{k=1}^2 \overline{\boldsymbol{O}})_h \equiv (\boldsymbol{R}, \mathcal{B}, \widehat{F} = (\times_{k=1}^2 F)_{h^{-1}(\cdot)})$. Then, $\Delta(\boldsymbol{M}_{L^\infty(\Omega, v)}(\overline{\boldsymbol{Q}} \times (\times_{k=1}^2 \overline{\boldsymbol{O}})_h, S(\overline{\rho})))$, the distance between $\overline{\boldsymbol{Q}}$ and $(\times_{k=1}^2 \overline{\boldsymbol{O}})_h$ concerning $\overline{\rho}$, is calculated as

$$\Delta(\boldsymbol{M}_{L^\infty(\Omega, v)}(\overline{\boldsymbol{Q}} \times (\times_{k=1}^2 \overline{\boldsymbol{O}})_h, S(\overline{\rho}))) = \left[\int\int_{\boldsymbol{R}^2} |\lambda_1 - \lambda_2|^2 \, \overline{\rho}((G^Q \times \widehat{F})(\mathrm{d}\lambda_1 \, \mathrm{d}\lambda_2)) \right]^{1/2}$$

$$= \left[\sum_{j=1}^{3} \sum_{(x_1,x_2)\in X^2} \overline{\rho}(\omega_j)|Q(\omega_j) - h(x_1,x_2)|^2 \times F_{\{x_1\}}(\omega_j) F_{\{x_2\}}(\omega_j) \right]^{1/2}$$

$$= [22.5\overline{\rho}(\omega_1) + 34.5\overline{\rho}(\omega_2) + 40\overline{\rho}(\omega_3)]^{1/2}. \quad (4.2)$$

Therefore, we see that $(4.2) \leqslant \sqrt{40} \approx 6.32$ for all $\overline{\rho} \in L^1_{+1}(\Omega, v)$. Now we can also answer the question (♯) in Example 3.3. That is, $Q(*) = \frac{1}{2}(w(r) + w(b)) = (30+10)/2 = 20$, though it of course includes the error 6.32.

The map $h: X^n \to \boldsymbol{R}$, $(n=2)$, in (4.1) may be chosen by the hint of "the law of large numbers". That is, if n is sufficiently large, the map $h: X^n \to \boldsymbol{R}$ (defined by $h(x_1,\ldots,x_n) = (1/n)\sum_{k=1}^{n} w(x_k)$) has a proper property, i.e., $\lim_{n\to\infty} \Delta(\boldsymbol{M}_{L^\infty(\Omega,v)}(\overline{\boldsymbol{Q}} \times (\times_{k=1}^{n} \overline{\boldsymbol{O}})_h, S(\overline{\rho}))) = 0$ for all $\overline{\rho} \in L^1_{+1}(\Omega, v)$. However, there are several ideas for the choice of h. Let $\overline{\boldsymbol{Q}} \equiv (\boldsymbol{R}, \mathscr{B}, G)$ and $\overline{\boldsymbol{O}} \equiv (X, \overline{\mathscr{F}}, F)$ be a quantity and W^*-observable in a W^*-algebra $\mathscr{N}$, respectively. For each $i = 1,2$, consider a measurable map $h_i: X \to \boldsymbol{R}$, and the image observable $\overline{\boldsymbol{O}}_{h_i} \equiv (\boldsymbol{R}, \mathscr{B}, F(h_i^{-1}(\cdot)))$ in $\mathscr{N}$. Also assume that $\overline{\boldsymbol{Q}}$ and $\overline{\boldsymbol{O}}_{h_i}$ commute. When it holds that

$$\Delta(\boldsymbol{M}_{\mathscr{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}_{h_1}, S(\overline{\rho}))) \leqslant \Delta(\boldsymbol{M}_{\mathscr{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}_{h_2}, S(\overline{\rho}))) \quad \forall \overline{\rho} \in \mathfrak{S}^n(\mathscr{N}_*), \quad (4.3)$$

we say that $\overline{\boldsymbol{O}}_{h_1}$ is *better* than $\overline{\boldsymbol{O}}_{h_2}$ as the approximation of $\overline{\boldsymbol{Q}}$. Also, $\overline{\boldsymbol{O}}_{h_2}$ is called *admissible* as the approximation of $\overline{\boldsymbol{Q}}$, if there exists no h_1 that satisfies (4.3) and the following condition:

$$\Delta(\boldsymbol{M}_{\mathscr{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}_{h_1}, S(\overline{\rho}_0))) < \Delta(\boldsymbol{M}_{\mathscr{N}}(\overline{\boldsymbol{Q}} \times \overline{\boldsymbol{O}}_{h_2}, S(\overline{\rho}_0))) \quad \text{for some } \overline{\rho}_0 \in \mathfrak{S}^n(\mathscr{N}_*).$$

As a well-known result concerning "admissibility", we mention the following example, which is also the preparation of Remark 4.5 later.

Example 4.4 (*Gaussian observable*). Let $\overline{\boldsymbol{O}} \equiv (\boldsymbol{R}, \mathscr{B}, F_{(\cdot)})$ be a "Gaussian observable" in $\mathscr{N} \equiv L^\infty(\boldsymbol{R} \times \boldsymbol{R}^+, \mathrm{d}\mu\,\mathrm{d}\sigma)$, that is,

$$F_\Xi(\mu,\sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi \exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right) \mathrm{d}u \quad (\forall(\mu,\sigma) \in \boldsymbol{R} \times \boldsymbol{R}^+ \equiv \boldsymbol{R} \times \{\sigma: \sigma > 0\}, \forall \Xi \in \mathscr{B}).$$

Consider the quantity $Q: \boldsymbol{R} \times \boldsymbol{R}^+ \to \boldsymbol{R}$ such that $Q(\mu,\sigma) = \mu$ $(\forall(\mu,\sigma) \in \boldsymbol{R} \times \boldsymbol{R}^+)$, which is identified with the observable $\overline{\boldsymbol{Q}} \equiv (\boldsymbol{R}, \mathscr{B}, G^Q_{(\cdot)})$, where $G^Q_\Xi(\mu,\sigma) = \chi_\Xi(\mu)$. Consider the product observable $\times_{k=1}^{n} \overline{\boldsymbol{O}} \equiv (\boldsymbol{R}^n, \mathscr{B}^n, \times_{k=1}^{n} F_{(\cdot)})$ in $L^\infty(\boldsymbol{R} \times \boldsymbol{R}^+, \mathrm{d}\mu\,\mathrm{d}\sigma)$. Define the map $h: \boldsymbol{R}^n \to \boldsymbol{R}$ such that $\boldsymbol{R}^n \ni (\lambda_1,\ldots,\lambda_n) \overset{h}{\mapsto} ((\lambda_1 + \cdots + \lambda_n)/n) \in \boldsymbol{R}$. Then, it is well known (cf. [12]) that $(\times_{k=1}^{n} \overline{\boldsymbol{O}})_h$ is admissible as the approximation of $\overline{\boldsymbol{Q}}$. In comparison with Remark 3.4, the state space $\boldsymbol{R} \times \boldsymbol{R}^+$ may be identified with the (parameterized) set of uncountable infinite urns, which contain sufficiently many balls with various weights.

Here let us add the following remark, which will also promote a better understanding of our assertion.

Remark 4.5 (*Fundamental observables in statistics*). In measurement theory, the discoveries of fundamental (or, useful) observables should be estimated very much. For example, "position quantity" and "momentum quantity" are fundamental in both classical and quantum mechanics. Also, we have Glauber–Sudarshan observable (i.e., observable on phase space) in semi-classical mechanics, cf. [2,3]. And also, "fuzzy logic" has fuzzy numbers observables (e.g., Lukasiewicz observable, cf. [8]). Therefore, if statistics is a certain aspect of measurements, we may find "fundamental observables" in statistics. However this has been already solved. That is, as mentioned in Remark 3.2 and Example 4.4, "Gaussian distribution" is induced by "Gaussian observable" Therefore, statistics already has useful observables, i.e., "Gaussian observable", "Poisson observable" and so on.

Though we focused on only "statistical inferences" in this paper, we are convinced that all other methods in statistics can be formulated in terms of measurements. For example, see [10], in which "factor analysis" is formulated in measurement theory. Also, see the lecture note [9], in which our recent and new results are summarized.

Lastly let us mention "fuzzy logic" since our original motivation of this paper is to clarify the confusion between "fuzzy logic" and statistics. Though there may be other opinions, our opinion for "fuzzy logic" is presented below.

Remark 4.6 (*Fuzzy logic*, cf. Ishikawa [8]). In this paper we were not concerned with the dynamical properties of systems. This is of course important. Motivated by quantum mechanics, i.e., "quantum mechanics" = "Born's measurement axiom" + "Heisenberg's kinetic equation", in [7] we proposed the viewpoint of "mechanics" such as "mechanics" = "measurement" + "kinetic equation". System theory is usually considered to be modeled on mechanics. Therefore, if we use the terms of system theory, this viewpoint is represented as follows:

"(dynamical) system theory"

= "measurement" + "the rule of time evolution". (4.4)

Here, "the rule of time evolution" may be extended to "the rule of the relation among systems" if we consider "general system theory" rather than "dynamical system theory" The system theory (4.4) is regarded as a mechanical approach to an understanding of (non-physical) phenomena. In this sense, the system theory (4.4) is an epistemology, or a philosophy, which may be called "*mechanical world view*". In [8], we asserted that the system theory (4.4) had great power of expression. Namely, a good translation from "natural language" into "system theoretical language" can be expected. We believe that this fact is the essence of "fuzzy logic" since the translation changes "fuzzy (or, loose) statements" to "system theoretical statements". That is, we consider that "fuzzy logic" is mainly related to the following aspect of measurements:

($\sharp$) how to translate a statement in a natural language into a statement in measurement theory, or more generally, in the system theory (4.4),

which is a kind of modeling problem in a broad sense (cf. [1] or [17], in which similar spirits can be found). Since statements in a natural language are rather "logical" or "qualitative", this ($\sharp$) may be also considered as the logical aspect of measurements. For example, several "syllogisms" were shown in [6]. Note that the above ($\sharp$) is not mathematical but system theoretical. As emphasized here and there in this paper, we believe that "measurement" is the most fundamental concept in science. Therefore, we do not start from "mathematics" (e.g., "Kolmogorov's probability theory", "mathematical logic", etc.) but "measurement". From the mathematical point of view, it is of course desirable that some mathematicians make "mathematical fuzzy logic" motivated by the logical aspect of measurements. In fact, "quantum logic" is a good mathematical theory, which was created by the hint of "Born's quantum measurement axiom". Also, in general we consider that "fuzzy system theory" is characterized as the study concerning grade quantities (i.e., membership functions) in the system theory (4.4).

5. Conclusions

The purpose of this paper was to propose a measurement theoretical formulation of statistics, that is, to understand or characterize statistics as one of the aspects of measurement theory. In Section 3 we showed that Fisher's and Bayes's spirit was, respectively, described in objective and subjective measurement theory. Also, in Section 4 we characterized "estimation under loss function in statistics" as "approximate measurements for quantities" in measurement theory. Since two main topics in statistics, i.e., "Fisher's and Bayes's methods" and "estimation under loss function", could be described in terms of measurements, we may conclude that statistics is a certain aspect of measurement theory. That is, we consider that statistics is mainly related to the following aspect of measurement theory:

($\sharp$) how to derive some useful information from the measured data obtained by a measurement.

This viewpoint for statistics seems to be important. That is because it bridges the gap between statistics and the other aspects of measurements (cf. Remark 2.2). For example, there seems to be some confusion between statistics and "fuzzy logic" (cf. [13]). It is clear that this confusion cannot be solved by comparing Kolmogorov's probability theory (or, the conventional formulation of statistics) with "mathematical fuzzy logic". Comparing the above ($\sharp$) with the ($\sharp$) in Remark 4.6, we can immediately clarify the confusion.

As seen in Remark 4.6, measurement theory is indispensable for “fuzzy logic”. On the other hand, from the practical point of view there may be some reason to consider that measured data can be analyzed without the knowledge of “measurements”. In fact, statistics has been developing without the concept “measurements”. However, as mentioned in Remarks 2.3 and 3.4, we consider that the conventional formulation of statistics is not sufficient. From the scientific point of view, we believe that measurement theory promotes a deep understanding of statistics.

Acknowledgements

I am grateful to Professors M. Sibuya, N. Shinozaki, R. Shibata and T. Iida in Keio University. From statistician’s standing point, they gave me valuable suggestions.

References

[1] M. Black, Reasoning with loose concepts, Dialogue 2 (1963) 1–12.
[2] E.B. Davies, Quantum Theory of Open Systems, Academic Press, New York, 1976.
[3] A.S. Holevo, Probabilistic and Statistical Aspects of Quantum Theory, North-Holland, Amsterdam, 1982.
[4] S. Ishikawa, Uncertainty relations in simultaneous measurements for arbitrary observables, Rep. Math. Phys. 29 (3) (1991) 257–273.
[5] S. Ishikawa, Uncertainties and an interpretation of nonrelativistic quantum theory, Internat. J. Theoret. Phys. 30 (4) (1991) 401–417.
[6] S. Ishikawa, Fuzzy inferences by algebraic method, Fuzzy Sets and Systems 87 (1997) 181–200.
[7] S. Ishikawa, A quantum mechanical approach to a fuzzy theory, Fuzzy Sets and Systems 90 (1997) 277–306.
[8] S. Ishikawa, Fuzzy logic in measurements, Fuzzy Sets and Systems 100 (1998) 291–300.
[9] S. Ishikawa, Fundamentals of system theory (in preparation). This will be submitted to Keio Science and Technology Reports. For further information, see http://www.math.keio.ac.jp/~ishikawa.
[10] S. Ishikawa, T. Iida, A system theoretical characterization of factor analysis, EUFIT’98 in Aachen (1998), to appear.
[11] A. Kolmogorov, Foundations of Probability (transl.), Chelsea, New York, 1950.
[12] E.K. Lehmann, Theory of Point Estimation, Wiley, New York, 1983.
[13] D. McNeill, P. Preiberger, Fuzzy Logic, Simon and Schuster, New York, 1993.
[14] J. von Neumann, Die Mathematischen Grundlagen Der Quantenmechanik, Springer, Berlin, 1932.
[15] S. Sakai, C^*-algebras and W^*-algebras, Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 60, Springer, Berlin, 1971.
[16] A. Wald, Statistical Decision Functions, Wiley, New York, 1950.
[17] L.A. Zadeh, Fuzzy sets, Inform. and Control 8 (1965) 338–353.