Research Report

KSTS/RR-06/005
April 19, 2006

Regression analysis, Kalman filter and measurement error model in measurement theory

by

Kohshi Kikuchi
Shiro Ishikawa

Kohshi Kikuchi
Keio University
Shiro Ishikawa
Keio University

Department of Mathematics
Faculty of Science and Technology
Keio University

3-14-1 Hiyoshi, Kohoku-ku, Yokohama, 223-8522 Japan

KSTS/RR-06/005
April 19, 2006

Regression analysis, Kalman filter and measurement error model in measurement theory

Kohshi KIKUCHI
Shiro ISHIKAWA
Keio University, Yokohama, Japan

Abstract

Recently we propose "measurement theory", which includes measurements in classical and quantum systems and is constructed in terms of a C^*-algebra. The purpose of this paper is to study (1): "regression analysis" (2): "Kalman filter" (3): "measurement error model" in measurement theory. And, we show that this approach is applicable to very general situations.

1 Introduction

Recently in Ishikawa (1997), Ishikawa (2000), (or see Ishikawa's papers in the references in Ishikawa (2001)), one of the authors proposed "measurement theory", which includes measurements in classical and quantum systems and is constructed within the framework of a C^*-algebra. This theory is characterized as a kind of generalization of von Neumann's theory proposed in his book: "*The mathematical foundation of quantum mechanics*"(cf. von Neumann (1932)), in which quantum mechanics is completely described in terms of mathematics (i.e., the theory of Hilbert space (cf. Prugovečki (1981))). Therefore, the quantum part of measurement theory is essentially the same as von Neumann's theory, and thus, it is well-authorized. On the other hand, the classical part has not been developed yet. Therefore, we are interested in the classical part of this measurement theory rather. Of course, we believe that the classical part is as profound as the quantum part (i.e., von Neumann's theory).

In Section 2, we introduce measurement theory (with Axioms 1 and 2, Proclaim 1), which includes measurements in classical and quantum systems and is constructed within the framework of a C^*-algebra. The purpose of this paper is to study (1): "regression analysis" in Section 3, (2): "Kalman filter" in Section 5, (3): "measurement error model" in Section 6. And, we show that this approach is applicable to very general situations.

2 Measurement theory

Measurement theory (=MT) can be classified two subjects, i.e., "(pure) measurement theory (= PMT)" and "statistical measurement theory (= SMT)". That is,

$$\text{MT (=“measurement theory”)} \begin{cases} \text{PMT (=“(pure) measurement theory”)}, \\ \\ \text{SMT (=“statistical measurement theory”)}. \end{cases}$$

PMT is essential and it is represented in the framework of the mathematical theory of C^*-algebra and is summarized in the following scheme:

$$\underset{\text{((pure) measurement theory)}}{\text{PMT}} = \underset{\text{(Axiom 1)}}{\text{measurement}} + \underset{\text{(Axiom 2)}}{\text{the relation among systems}}, \tag{2.1}$$

which includes classical and quantum measurements. PMT is introduced as a kind of generalization of quantum mechanics, i.e.,

$$\text{quantum mechanics} = \text{Born's quantum measurement} + \underset{\text{(= Heisenberg kinetic equation)}}{\text{Schrödinger equation}}. \tag{2.2}$$

Also, it should be noted that the classical part of MT includes the following conventional dynamical system theory (= DST):

$$\text{DST} = \begin{cases} \frac{dx(t)}{dt} = f(x(t), u_1(t), t), \quad x(0) = x_0 & \cdots \quad \text{(state equation)}, \\ y(t) = g(x(t), u_2(t), t) & \cdots \quad \text{(measurement equation)}, \end{cases} \tag{2.3}$$

where u_1 and u_2 are external forces.

In PMT, the initial state (e.g., x_0 in (2.3)) is composed of one point (of the state space) and not distributed on the state space. However, if an initial distributed state is permitted, we can propose SMT (= statistical measurement theory) as follows.

$$\text{SMT} = \underset{\text{(Axioms 1 and 2)}}{\text{PMT}} + \underset{\text{(the probabilistic interpretation of distributed state)}}{\text{“statistical state”}} \text{ in } C^*\text{-algebra}. \tag{2.4}$$

Thus, if we define Proclaim 1 by

$$\text{“Proclaim 1”} = \text{“Axiom 1”} + \underset{\text{(the probabilistic interpretation of distributed state)}}{\text{“statistical state”}}. \tag{2.5}$$

we see that SMT is formulated as follows.

$$\text{SMT} = \underset{\text{(Proclaim 1)}}{\text{statistical measurement}} + \underset{\text{(Axiom 2)}}{\text{the relation among systems}} \quad \text{in } C^*\text{-algebra}. \tag{2.6}$$

Thus, we say that PMT is more fundamental than SMT. That is, we see that *there is no SMT without PMT*.

It is generally considered that the system theory (= DST) is a kind of epistemology called "the mechanical world view", namely, an epistemology to understand and analyze every phenomenon (that appears in our usual life) — economics, psychology, engineering and so on — by an analogy of mechanics. Therefore, we also consider that MT is just "*the mathematical representation of the mechanical world view*". Since MT is regarded as a kind of generalization of "DST (2.3)", MT is also called "general system theory (= GST)", i.e., GST = MT. In this paper we use the term "measurement theory (= MT)".

2.1 Measurements (Axiom 1 in PMT (2.1) and Proclaim 1 in SMT(2.6))

We intend that this paper is essentially self-contained. And further, the mathematical deep knowledge will never be required. Let $\mathcal{A}$ be a C^*-algebra (cf. Sakai (1971)), that is, a Banach algebra with the involution "$*$" and the norm $\|\cdot\|$ satisfying the C^*-condition: $\|F^*F\| = \|F\|^2$ ($\forall F \in \mathcal{A}$). For simplicity, in this paper we always assume that $\mathcal{A}$ is *unital*, i.e., $\mathcal{A}$ has the identity $1_{\mathcal{A}}$. The complex-valued linear functional $\rho(F)$ on $\mathcal{A}$ is denoted by ${}_{\mathcal{A}^*}\langle \rho, F\rangle_{\mathcal{A}}$, where $\rho \in \mathcal{A}^*$ $\big($the dual Banach space, i.e., $\mathcal{A}^* = \{\rho \mid \rho$ is a continuous linear functional on $\mathcal{A}\}\big)$ and $F \in \mathcal{A}$. An element F ($\in \mathcal{A}$) is said to be *self-adjoint* if it holds that $F = F^*$. Also, a self-adjoint element F is called a *positive element*, denoted by $F \geq 0$, if $F = F_0^* F_0$ holds for some F_0 ($\in \mathcal{A}$). Define the *mixed state class* (or, *distributed state class*) $\mathfrak{S}^m(\mathcal{A}^*)$ by $\{\rho \in \mathcal{A}^* \mid \|\rho\|_{\mathcal{A}^*}(\equiv \sup_{F\in\mathcal{A},\|F\|_{\mathcal{A}}\leq 1} |\rho(F)|) = 1$ and $\rho(F) \geq 0$ for all $F \geq 0\}$. A mixed state ρ ($\in \mathfrak{S}^m(\mathcal{A}^*)$) is called a *pure state* if it satisfies that "$\rho = \theta\rho_1 + (1-\theta)\rho_2$ for some ρ_1, ρ_2 ($\in \mathfrak{S}^m(\mathcal{A}^*)$) and $0 < \theta < 1$" implies "$\rho = \rho_1 = \rho_2$". Define $\mathfrak{S}^p(\mathcal{A}^*) \equiv \{\rho^p \in \mathfrak{S}^m(\mathcal{A}^*) \mid \rho^p$ is a pure state $\}$, which is called a *state space* (or, *pure state space*). Note that $\mathfrak{S}^p(\mathcal{A}^*)$ and $\mathfrak{S}^m(\mathcal{A}^*)$ are compact Hausdorff spaces in the sense of the weak*-topology $\tau(\mathcal{A}^*, \mathcal{A})$.

A C^*-algebra $\mathcal{A}$ is said to be *commutative* if it holds that $F_1F_2 = F_2F_1$ for all $F_1, F_2 \in \mathcal{A}$. Gelfand theorem (cf. Sakai (1971)) says that any commutative C^*-algebra $\mathcal{A}$ can be identified with some $C(\Omega)$, the algebra composed of all complex valued continuous functions f on a compact Hausdorff space Ω. Here, the norm is defined by $\|f\| = \sup_{\omega\in\Omega} |f(\omega)|$. Riesz representation theorem (cf. Yosida (1980)) reads that $C(\Omega)^*$, the dual Banach space of $C(\Omega)$, can be identified with $\mathcal{M}(\Omega)$, where $\mathcal{M}(\Omega)$ is the Banach space composed of all regular complex-valued measures ρ on Ω with the norm $\|\rho\|_{\mathcal{M}(\Omega)} = \sup_{\|f\|_{C(\Omega)}\leq 1} \int_\Omega f(\omega)\rho(d\omega)$. The identification of $C(\Omega)^*$ with $\mathcal{M}(\Omega)$ is prescribed in the following: there exists an isometric, linear, and bijective operator Ψ : $\mathcal{M}(\Omega) \to C(\Omega)^*$ such that $\Psi(\rho)(f) = \int_\Omega f(\omega)\rho(d\omega)$ $(\forall f \in C(\Omega), \forall \rho \in \mathcal{M}(\Omega))$. We see and denote that $\mathfrak{S}^m(C(\Omega)^*) = \{\rho^m \in \mathcal{M}(\Omega) \mid \rho^m \geq 0, \|\rho^m\|_{\mathcal{M}(\Omega)} = 1\} \equiv \mathcal{M}^m_{+1}(\Omega)$, and $\mathfrak{S}^p(C(\Omega)^*) = \big\{\delta_\omega \in \mathcal{M}(\Omega) \mid \delta_\omega$ is a point measure at $\omega \in \Omega$, i.e., ${}_{\mathcal{M}(\Omega)}\langle \delta_\omega, f\rangle_{C(\Omega)} = f(\omega)$ $(\forall f \in C(\Omega), \forall \omega \in \Omega)\big\} \equiv \mathcal{M}^p_{+1}(\Omega)$. Under the identification: $\Omega \ni \omega \longleftrightarrow \delta_\omega \in \mathcal{M}^p_{+1}(\Omega)$ (that is, $\Omega \approx \mathcal{M}^p_{+1}(\Omega)$), the Ω is also called a *state space*. Also, note that the state space Ω is called a *parameter space* in the conventional formulation of statistics.

As a typical non-commutative C^*-algebra, we know the $B(V)$, that is,

$$B(V) = \{T \mid T \text{ is a continuous linear operator from a Hilbert space } V \text{ into itself}\}. \tag{2.7}$$

Although this $B(V)$ is essential to quantum mechanics, we omit to mention the elementary knowledge of the $B(V)$. That is because our concern is concentrated on classical measurements in this paper.

As a natural generalization of Davies' idea (cf. Davies (1976) and Holevo (1973)) in quantum mechanics, we define "observable" as follows. A triple $\mathbf{O} \equiv (X, \mathcal{F}, F)$ is called an *observable* (or precisely, *C^*-observable*) formulated in a C^*-algebra $\mathcal{A}$, if it satisfies that

(i) [field]. *X is a set (called a "measured value set" or "label set"), and $\mathcal{F}$ is the subfield of the power set $\mathcal{P}(X)(\equiv \{\Xi \mid \Xi \subseteq X\})$,*

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

(iii) *for any countable decomposition $\{\Xi_1, \Xi_2, \Xi_3, ...\}$ of Ξ, $\Big($i.e., $\Xi, \Xi_i \in \mathcal{F}$ $(i = 1, 2, 3, ...)$, $\cup_{n=1}^\infty \Xi_n = \Xi$, $\Xi_n \cap \Xi_m = \emptyset$ (if $n \neq m$)$\Big)$, it holds that*

$$\rho\big(F(\Xi)\big) = \lim_{N\to\infty} \rho\big(\sum_{n=1}^N F(\Xi_n)\big) \qquad (\forall \rho \in \mathfrak{S}^m(\mathcal{A}^*)).$$

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

Remark 2.1. When we want to stress that X is finite, we write $(X, 2^X, F)$ instead of $(X, \mathcal{F}, F)$. In this paper we usually assume that X is finite, even when we can do well without the assumption that X is finite.

EXAMPLE 2.2. ((i). Gaussian observable in $C(\Omega)$). Put $\Omega = [-L, L]$ ($\subseteq \mathbf{R}$, the real line). And let σ be a fixed positive real number. Define the *normal observable* (or, *Gaussian observable*) $\mathbf{O}_{G^\sigma} \equiv (\mathbf{R}, \mathcal{B}_{\mathbf{R}}, G^\sigma)$ in $C(\Omega)$ such that

$$[G^\sigma(\Xi)](\omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_\Xi e^{-\frac{(x-\omega)^2}{2\sigma^2}} dx \quad (\forall \Xi \in \mathcal{B}_{\mathbf{R}}, \forall \omega \in \Omega).$$

[(ii).Discrete Gaussian observable] Under the condition that X is finite, the definition of "Gaussian observable" is somewhat complicated as follows. $\Big($For the ordinary Gaussian observable (i.e., X $(= \mathbf{R})$ is infinite), see Ishikawa (2000).$\Big)$ Put $\Omega \equiv [a, b]$ ($\subseteq \mathbf{R}$, the real line), the closed interval. Let σ^2 be a variance. And let N be a sufficiently large fixed positive integer. Put $X_N \equiv \{\frac{k}{N} \mid k = 0, \pm 1, \pm 2, ..., \pm N^2\}$. And define a *discrete Gaussian observable* $\mathbf{O}_{\sigma,N} \equiv (X_N, 2^{X_N}, F_{\sigma,N})$ in the commutative C^*-algebra $C([a, b])$ such that

$$
\begin{aligned}
&[F_{\sigma,N}(\{k/N\})](\omega) \\
&= \begin{cases}
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{N-\frac{1}{2N}}^{\infty} \exp[-\frac{(x-\omega)^2}{2\sigma^2}]dx & (k = N^2, \forall \omega \in [a, b]), \\
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{\frac{k}{N}-\frac{1}{2N}}^{\frac{k}{N}+\frac{1}{2N}} \exp[-\frac{(x-\omega)^2}{2\sigma^2}]dx & (\forall k = 0, \pm 1, \pm 2, ..., \pm(N^2 - 1),\ \forall \omega \in [a, b]), \\
\frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{-N+\frac{1}{2N}} \exp[-\frac{(x-\omega)^2}{2\sigma^2}]dx & (k = -N^2, \forall \omega \in [a, b]).
\end{cases}
\end{aligned}
\tag{2.7}
$$

And thus, for any Ξ $(\in 2^{X_N})$, we define $[F_{\sigma,N}(\Xi)](\omega) = \sum_{\frac{k}{N} \in \Xi} [F_{\sigma,N}(\{k/N\})](\omega)$. This $\mathbf{O}_{\sigma,N}$ is the most useful observable in classical measurements.

With any *system* S, a C^*-algebra $\mathcal{A}$ can be associated in which measurement theory of that system can be formulated. A *state* of the system S is represented by a pure state ρ^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$, an *observable* is represented by a C^*-observable $\mathbf{O} \equiv (X, \mathcal{F}, F)$ formulated in the C^*-algebra $\mathcal{A}$. Also, a *measurement of the observable* $\mathbf{O}$ *for the system* S *with the state* ρ^p is denoted by $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]})$ $\big($or in short, $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})\big)$. We can obtain a *measured value* x $(\in X)$ by the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$.

The axiom presented below is analogous to (or, a kind of generalization of) Born's probabilistic interpretation of quantum mechanics (cf. von Neumann (1932)).

AXIOM 1. (Measurement). *Consider a measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]})$ *formulated in a* C^**-algebra* $\mathcal{A}$*. Then, the probability that a measured value* x $(\in X)$ *obtained by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ *belongs to a set* Ξ $(\in \mathcal{F})$ *is given by* $\rho^p(F(\Xi)) = {}_{\mathcal{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathcal{A}}$.

Thus, the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]})$ induces the probability space $(X, \mathcal{F}, \rho^p(F(\cdot)))$, which is called a *sample space*.

We introduce the following classification in measurement theory:

$$\text{measurement theory} \begin{cases} \text{classical measurement theory (for classical systems)} \\ \\ \text{quantum measurement theory (for quantum systems)} \\ \qquad\qquad \text{(i.e., von Neumann's theory, cf. (1932))} \end{cases} \tag{2.8}$$

where a C^*-algebra $\mathcal{A}$ is commutative or non-commutative.

For each $k = 1, 2, ..., n$, we consider an observable $\mathbf{O}_k \equiv (X_k, 2^{X_k}, F_k)$ in a C^*-algebra $\mathcal{A}$. An observable $\mathbf{O} \equiv (\prod_{k=1}^n X_k, 2^{\Pi_{k=1}^n X_k}, F)$ in $\mathcal{A}$ is called a *quasi-product observable of* $\{\mathbf{O}_k \mid k = 1, 2, ..., n\}$ if it satisfies (i) $\prod_{k=1}^n X_k$ is the product set, (ii) it holds that

$$F(X_1 \times \cdots \times X_{k-1} \times \Xi_k \times X_{k+1} \times \cdots \times X_n) = F_k(\Xi_k) \quad (\forall \Xi_k \in 2^{X_k}, \forall k = 1, ..., n). \tag{2.9}$$

The quasi-product observable $\mathbf{O}$ [resp. the F] is denoted by $\overset{\text{qp}}{\times}_{k\in\{1,2,...,n\}} \mathbf{O}_k$ [resp. $\overset{\text{qp}}{\times}_{k\in\{1,2,...,n\}} F_k$]. Note that the existence and the uniqueness of the quasi-product observable $\mathbf{O} \equiv (\prod_{k=1}^n X_k,\ 2^{\Pi_{k=1}^n X_k}, F)$ of $\{\mathbf{O}_k \mid k = 1, 2, ..., n\}$ are not guaranteed in general. If $\mathbf{O}_k$ ($k = 1, 2, ..., n$) is commutative, i.e., $F_k(\Xi_k)F_{k'}(\Xi_{k'}) = F_{k'}(\Xi_{k'})F_k(\Xi_k)$ ($\forall \Xi_k \in 2^{X_k}, \forall \Xi'_k \in 2^{X'_k}, k \neq k'$), the quasi-product observable of $\{\mathbf{O}_k \mid k = 1, 2, ..., n\}$ always exists. For example, it suffices to define F such that

$$F(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_n) = F_1(\Xi_1)F_2(\Xi_2) \cdots F_n(\Xi_n) \quad (\forall \Xi_k \in 2^{X_k}, \forall k = 1, 2, ..., n). \tag{2.10}$$

The quasi-product observable which satisfies (2.10) is called a *product observable* of $\{\mathbf{O}_k \mid k = 1, 2, ..., n\}$ and denoted by $\times_{k=1}^n \mathbf{O}_k$ [resp. $\times_{k=1}^n F_k$].

Example 2.3. (1. Product discrete Gaussian observable. (Continued from Example 2.2)). The product discrete Gaussian observable $\mathbf{O}_{\sigma,N} \times \mathbf{O}_{\sigma,N} = (X_N \times X_N, 2^{X_N \times X_N}, F_{\sigma,N} \times F_{\sigma,N}) = \mathbf{O}^2_{\sigma,N} = (X_N^2, 2^{X_N^2}, F^2_{\sigma,N})$ in the commutative C^*-algebra $C([a, b])$ is defined by

$$[F^2_{\sigma,N}(\Xi_1 \times \Xi_2)](\omega) = \sum_{\frac{k_1}{N} \in \Xi_1, \frac{k_2}{N} \in \Xi_2} [F_{\sigma,N}(\{\frac{k_1}{N}\})](\omega) \cdot [F_{\sigma,N}(\{\frac{k_2}{N}\})](\omega) \qquad (\forall \omega \in [a, b]).$$

Let $\omega_0 \in [a, b]$. Then, Axiom 1 says that the probability that the measured value $(\frac{k_1}{N}, \frac{k_2}{N})$ ($\in X_N^2$) is obtained by the measurement $\mathbf{M}_{C([a,b])}(\mathbf{O}^2_{\sigma,N}, S_{[\delta_{\omega_0}]})$ is given by $[F^2_{\sigma,N}(\{(\frac{k_1}{N}, \frac{k_2}{N})\})](\omega_0)$.

(2. Null observable). Define the observable $\mathbf{O}^{(\text{nl})} \equiv (\{0, 1\}, 2^{\{0,1\}}, F^{(\text{nl})})$ in $\mathcal{A}$ such that

$$F^{(\text{nl})}(\emptyset) \equiv 0, \quad F^{(\text{nl})}(\{0\}) \equiv 0, \quad F^{(\text{nl})}(\{1\}) \equiv 1_{\mathcal{A}}, \quad F^{(\text{nl})}(\{0, 1\}) \equiv 1_{\mathcal{A}},$$

which may be called a *null observable* (or, *existence observable*). Then, we have a measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^{(\text{nl})}, S_{[\rho^p]})$. Note that

(♯) the probability that the measured value by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^{(\mathrm{nl})}, S_{[\rho^p]})$ is equal to 1 ($\in \{0,1\}$) is given by 1. That is, the measured value is always equal to 1 ($\in \{0,1\}$).

Thus, we consider that "to take the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^{(\mathrm{nl})}, S_{[\rho^p]})$" is the same as "to take no measurement", or more precisely, "to assure the existence of the system".

Remark 2.4. (Simultaneous measurement). The quasi-product observable (or, the product observable) is used to represent *the measurement of more than one observables* as follows. For example, consider *a measurement of* $\mathbf{O}_1$ *and* $\mathbf{O}_2$ *for the system with the state* ρ^p ($\in \mathfrak{S}^p(\mathcal{A}^*)$). If the quasi-product observable $\mathbf{O}_1 \overset{\mathrm{qp}}{\times} \mathbf{O}_2$ of $\mathbf{O}_1$ and $\mathbf{O}_2$ exists, the measurement is represented by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1 \overset{\mathrm{qp}}{\times} \mathbf{O}_2, S_{[\rho^p]})$ $\Big($and not "$\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1, S_{[\rho^p]}) + \mathbf{M}_{\mathcal{A}}(\mathbf{O}_2, S_{[\rho^p]})$"$\Big)$. If the quasi-product observable $\mathbf{O}_1 \overset{\mathrm{qp}}{\times} \mathbf{O}_2$ does not exist, the measurement also does not exist. That is, the symbol "$\mathbf{M}_{\mathcal{A}}(\mathbf{O}_1, S_{[\rho^p]}) + \mathbf{M}_{\mathcal{A}}(\mathbf{O}_2, S_{[\rho^p]})$" is non-sense. Thus we can say that

(♯) only one measurement is permitted to be conducted even in the classical measurement theory.

which is the well-known fact in quantum mechanics. The measurement $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O}_1 \overset{\mathrm{qp}}{\times} \mathbf{O}_2, S_{[\rho^p]}\big)$ is sometimes called a *simultaneous measurement* of two observables $\mathbf{O}_1$ and $\mathbf{O}_2$.

The following example will promote the better understanding of Axiom 1.

Example 2.5. (The urn problem). There are two urns ω_1 and ω_2. The urn ω_1 [resp. ω_2] contains 8 red and 2 blue balls [resp. 4 red and 6 blue balls]. That is,

	red balls	blue balls
urn ω_1	8	2
urn ω_2	4	6

(2.11)

Here, consider the following measurement M_1:

M_1 := "Pick out one ball from the urn ω_1, and recognize the color of the ball".

The measurement M_1 is formulated as follows. Put $\Omega = \{\omega_1, \omega_2\}$. And define the observable $\mathbf{O} = (\{r,b\}, 2^{\{r,b\}}, F)$ in $C(\Omega)$ such that

$$F(\emptyset)(\omega_1) = 0, \quad F(\{r\})(\omega_1) = 0.8, \quad F(\{b\})(\omega_1) = 0.2, \quad F(\{r,b\})(\omega_1) = 1.0,$$
$$F(\emptyset)(\omega_2) = 0, \quad F(\{r\})(\omega_2) = 0.4, \quad F(\{b\})(\omega_2) = 0.6, \quad F(\{r,b\})(\omega_2) = 1.0. \tag{2.12}$$

Then, we see that

$$M_1 = \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\delta_{\omega_1}]}). \tag{2.13}$$

The probability that a measured value r [resp. b] is obtained is, by Axiom 1, given by

$$F(\{r\})(\omega_1) = 0.8, \qquad [\text{resp. } F(\{b\})(\omega_1) = 0.2]. \tag{2.14}$$

The following example will promote the better understanding of Proclaim 1, mentioned latter.

Example 2.6. (Coin-tossing and urn problem). Under the same situation of Example 2.5, consider the following procedures (P_1) and (P_2).

(P_1) One of the two urns (i.e., ω_1 or ω_2) is chosen by an unfair tossed coin ($C_{p,1-p}$), i.e.,

$$\text{Head } (100p\%) \to \omega_1, \ \text{Tail } (100(1-p)\%) \to \omega_2 \quad (0 \le p \le 1). \tag{2.15}$$

The chosen urn is denoted by $[*]$ ($\in \{\omega_1, \omega_2\}$). Note, for completeness, that we do not know whether $[*]$ is ω_1 or ω_2 since the two can not be distinguished in appearance. Here define the mixed state ν_0 ($\in \mathcal{M}_{+1}^m(\Omega)$) such that $\nu_0(\{\omega_1\}) = p$, $\nu_0(\{\omega_2\}) = 1-p$, which is considered to be "the distribution of $[*]$". Thus we call the ν_0 a *statistical state.*

(P_2) Take one ball, at random, out of the urn chosen by the procedure (P_1). That is, we take the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$.

Note that

(i) the probability that $[*] = \delta_{\omega_1}$ $\big[$resp. $[*] = \delta_{\omega_2}\big]$ is given by p $\big[$resp. $1-p\big]$.

(ii) If $[*] = \delta_{\omega_1}$ $\big[$resp. if $[*] = \delta_{\omega_2}\big]$, the probability that the measured value obtained by $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$ is equal to x ($\in \{r, b\}$) is, by Axiom 1, given by

$$\begin{aligned} &{}_{\mathcal{M}(\Omega)}\langle \delta_{\omega_1}, F(\{x\}) \rangle_{C(\Omega)} = 0.8 \ \ (\text{if } x = r), \quad = 0.2 \ \ (\text{if } x = b), \\ \big[\text{resp. } &{}_{\mathcal{M}(\Omega)}\langle \delta_{\omega_2}, F(\{x\}) \rangle_{C(\Omega)} = 0.4 \ \ (\text{if } x = r), \quad = 0.6 \ \ (\text{if } x = b)\big]. \end{aligned} \tag{2.16}$$

Thus, under the condition (P_1), the probability that the measured value obtained by the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$ is equal to x ($\in \{r, b\}$) is given by

$$P(\{x\}) = \int_\Omega {}_{\mathcal{M}(\Omega)}\langle \delta_\omega, F(\{x\}) \rangle_{C(\Omega)} \nu_0(d\omega) = {}_{\mathcal{M}(\Omega)}\langle \nu_0, F(\{x\}) \rangle_{C(\Omega)}$$
$$= \begin{cases} 0.8p + 0.4(1-p) & \text{(if } x = r) \\ 0.2p + 0.6(1-p) & \text{(if } x = b) \end{cases}$$

Therefore, we see;

($\sharp$) There is a reason that the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$ in (P_2) under the condition (P_1) is denoted by $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]}(\nu_0))$, and called a *"statistical measurement"*. Here the mixed state ν_0 ($\in \mathcal{M}^m_{+1}(\Omega)$) is called a *"statistical state"*, which represents the distribution of [$*$]. And, the probability that the measured value x ($\in \{r, b\}$) is obtained by the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]}(\nu_0))$, is given by

$$_{C(\Omega)^*}\langle \nu_0, F(\{x\}) \rangle_{C(\Omega)} \equiv \int_\Omega {}_{C(\Omega)^*}\langle \delta_\omega, F(\{x\}) \rangle_{C(\Omega)} \nu_0(d\omega).$$

That is, the statistical state ν_0 is *the mixed state with probabilistic interpretation.*

Summing up, we have the following proclaim:

PROCLAIM 1. (The probabilistic interpretation of mixed states). *Consider a statistical measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[*]}(\rho^m))$ formulated in a C^*-algebra $\mathcal{A}$. Then, the probability that x ($\in X$), the measured value obtained by the statistical measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]}(\rho^m))$, belongs to a set Ξ ($\in \mathcal{F}$) is given by*

$$\rho^m(F(\Xi)) \equiv {}_{\mathcal{A}^*}\langle \rho^m, F(\Xi) \rangle_{\mathcal{A}}.$$

The statistical measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]}(\rho^m))$ is sometimes denoted by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S(\rho^m))$.

Thus, we see that Proclaim 1 is characterized as follows.

$$\text{"Proclaim 1"} = \text{"Axiom 1"} + \underset{\text{(the probabilistic interpretation of distributed state)}}{\text{"statistical state"}} \tag{2.17}$$

2.2 The relation among systems (Axiom 2 in PMT (2.1) and SMT (2.6))

In this section we devote ourselves to the "relation among systems (i.e., Axiom 2)" in PMT (2.1) and SMT (2.6).

Let $\mathcal{A}_1$ and $\mathcal{A}_2$ be C^*-algebras. A continuous linear operator $\Phi_{1,2} : \mathcal{A}_2 \to \mathcal{A}_1$ is called a *Markov operator*, if it satisfies

(i) $\Phi_{1,2}(F_2) \geq 0$ for any positive element F_2 in $\mathcal{A}_2$,

(ii) $\Phi_{1,2}(1_2) = 1_1$, where 1_k is the identity in $\mathcal{A}_k$ $(k = 1, 2)$.

Let $\mathbf{O}_2 = (X_2, 2^{X_2}, F_2)$ be an observable in $\mathcal{A}_2$. Put $(\Phi_{1,2}F_2)(\Xi_2) = \Phi_{1,2}(F_2(\Xi_2))$ $(\forall \Xi_2 \in 2^{X_2})$. Then, the $(X_2, 2^{X_2}, \Phi_{1,2}F_2)$ is an observable in $\mathcal{A}_1$, which is denoted by $\Phi_{1,2}\mathbf{O}_2$.

A Markov operator $\Phi_{1,2} : \mathcal{A}_2 \to \mathcal{A}_1$ is called a *homomorphism* (or precisely, a *C^*-algebraic homomorphism*), if it holds that

(i) $\Phi_{1,2}(F_2)\Phi_{1,2}(G_2) = \Phi_{1,2}(F_2 G_2)$ for any F_2 and G_2 in $\mathcal{A}_2$,

(ii) $(\Phi_{1,2}F_2)^* = \Phi_{1,2}(F_2^*)$ for any F_2 in $\mathcal{A}_2$ (where $*$ is the involution in $\mathcal{A}$).

Let $\Phi_{1,2}^* : \mathcal{A}_1^* \to \mathcal{A}_2^*$ be the dual operator of a Markov operator $\Phi_{1,2} : \mathcal{A}_2 \to \mathcal{A}_1$. Then, the following mathematical results are well known (cf. Sakai (1971)):

(i) $\Phi_{1,2}^*(\mathfrak{S}^m(\mathcal{A}_1^*)) \subseteq \mathfrak{S}^m(\mathcal{A}_2^*)$,

(ii) $\Phi_{1,2}^*(\mathfrak{S}^p(\mathcal{A}_1^*)) \subseteq \mathfrak{S}^p(\mathcal{A}_2^*)$, if $\Phi_{1,2} : \mathcal{A}_2 \to \mathcal{A}_1$ is homomorphic.

Suppose that $\mathcal{A}_1$ and $\mathcal{A}_2$ are commutative C^*-algebras, i.e., $\mathcal{A}_1 = C(\Omega_1)$ and $\mathcal{A}_2 = C(\Omega_2)$ with compact Hausdorff spaces Ω_i $(i = 1, 2)$. Under the identification that $\mathfrak{S}^p(\mathcal{A}_1^*) = \mathcal{M}_{+1}^p(\Omega_1) \approx \Omega_1$ and $\mathfrak{S}^m(\mathcal{A}_2^*) = \mathcal{M}_{+1}^m(\Omega_2)$, the above property (i) implies that the dual operator $\Phi_{1,2}^*$ of a Markov operator $\Phi_{1,2}$ can be identified with a *transition probability rule* $P(\omega_1, B_2)$ $(\omega_1 \in \Omega_1, B_2 \in \mathcal{B}_{\Omega_2}$; Borel field on $\Omega_2)$ such that $M(\omega_1, B_2) = (\Phi_{1,2}^*(\delta_{\omega_1}))(B_2)$. Also, under the identification that $\mathcal{M}_{+1}^p(\Omega_1) \approx \Omega_1$ and $\mathcal{M}_{+1}^p(\Omega_2) \approx \Omega_2$, the above property (ii) implies that the dual operator $\Phi_{1,2}^*$ of a homomorphism $\Phi_{1,2}$ is identified with a continuous map $\phi_{1,2}$ from Ω_1 into Ω_2 defined by $(\Phi_{1,2}f_2)(\omega_1) = f_2(\phi_{1,2}(\omega_1))$ $(\forall \omega_1 \in \Omega_1, \forall f_2 \in C(\Omega_2))$ in the following sense:

$$\Phi_{1,2}^*(\delta_{\omega_1}) = \delta_{\phi_{1,2}(\omega_1)} \quad (\forall \omega_1 \in \Omega_1). \tag{2.18}$$

Let $(T, \leq)$ be a tree-like partial ordered set, i.e., a partial ordered set such that "$t_1 \leq t_3$ and $t_2 \leq t_3$" implies "$t_1 \leq t_2$ or $t_2 \leq t_1$". Put $T_{\leq}^2 = \{(t_1, t_2) \in T \times T \mid t_1 \leq t_2\}$. An element $t_0 \in T$ is called a *root* if $t_0 \leq t$ $(\forall t \in T)$ holds. In this paper, we always assume, for simplicity, that T is finite (cf. Remark 2.1).

DEFINITION 2.7. (General system). *The pair* $\mathbf{S}_{[\rho^p_{t_0}]} \equiv [S_{[\rho^p_{t_0}]}; \{\mathcal{A}_{t_2} \overset{\Phi_{t_1,t_2}}{\to} \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}]$ $\big[$*resp.* $\mathbf{S}(\rho^m_{t_0}) \equiv [S(\rho^m_{t_0}); \{\mathcal{A}_{t_2} \overset{\Phi_{t_1,t_2}}{\to} \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}]\big]$ *is called a general system with an initial state* $\rho^p_{t_0}$ $(\in \mathfrak{S}^p(\mathcal{A}^*_{t_0}))$ $\big[$*resp. general system with an initial statistical state* $\rho^m_{t_0}$ $(\in \mathfrak{S}^m(\mathcal{A}^*_{t_0}))\big]$ *if it satisfies the following conditions* (i)$\sim$(iii).

(i) *With each* t $(\in T)$*, a* C^**-algebra* $\mathcal{A}_t$ *is associated.*

(ii) *The* t_0 $(\in T)$ *is the root of* T*. And, assume that a system* S *has the state* $\rho^p_{t_0}$ *at* t_0 $\big[$*resp. the statistical state* $\rho^m_{t_0}$ *at* $t_0\big]$*, that is, the initial state is equal to* $\rho^p_{t_0}$ $\big[$*resp.* $\rho^m_{t_0}\big]$

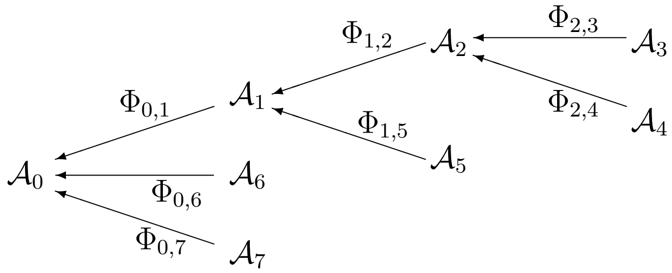
(iii) *For every* $(t_1, t_2) \in T^2_{\leq}$*, a Markov operator* $\Phi_{t_1,t_2} : \mathcal{A}_{t_2} \to \mathcal{A}_{t_1}$ *is defined such that* $\Phi_{t_1,t_2}\Phi_{t_2,t_3} = \Phi_{t_1,t_3} : \mathcal{A}_{t_3} \to \mathcal{A}_{t_1}$ *holds for all* $(t_1, t_2), (t_2, t_3) \in T^2_{\leq}$*, where* $\Phi_{t,t} : \mathcal{A}_t \to \mathcal{A}_t$ *is the identity map.*

The family $\{\mathcal{A}_{t_2} \overset{\Phi_{t_1,t_2}}{\to} \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}$ *is also called a Markov relation among systems. Let an observable* $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ *in a* C^**-algebra* $\mathcal{A}_t$ *be given for each* $t \in T$*. The pair* $[\{\mathbf{O}_t\}_{t\in T}; \{\mathcal{A}_{t_2} \overset{\Phi_{t_1,t_2}}{\to} \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}]$ *is called a sequential observable.*

Before we propose Axiom 2, we make some preparations. Let $T = \{0, 1, ..., N\}$ be a tree with the root 0. Define the *parent map* $\pi : T \setminus \{0\} \to T$ such that $\pi(t) = \max\{s \in T \mid s < t\}$. It is clear that the tree $(T \equiv \{0, 1, ..., N\}, \leq)$ can be identified with the pair $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$. The Markov relation $\{\mathcal{A}_{t_2} \overset{\Phi_{t_1,t_2}}{\to} \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}$ is also denoted by $\{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}$.

The following example will promote the better understanding of Axiom 2 mentioned later.

EXAMPLE 2.8. (A simple general system, Heisenberg picture). Suppose that a tree $(T \equiv \{0, 1, ..., 7\}, \pi)$ has an ordered structure such that $\pi(1) = \pi(6) = \pi(7) = 0$, $\pi(2) = \pi(5) = 1$, $\pi(3) = \pi(4) = 2$. $\big($See the figure (2.19).$\big)$ Consider a general system $\mathbf{S}_{[\rho^p_0]} \equiv [S_{[\rho^p_0]}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]$ with the initial system $S_{[\rho^p_0]}$ $\big[$resp. a general statistical system $\mathbf{S}(\rho^m_0) \equiv [S(\rho^m_0); \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]$ with the initial system $S(\rho^m_0)\big]$.



(2.19)

Also, for each t ($\in \{0, 1, ..., 6, 7\}$), consider an observable $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ in a C^*-algebra $\mathcal{A}_t$. Now we want to consider the following "measurement",

($\sharp$) for a system $S_{[\rho_0^p]}$ $\big[$resp. a statistical system $S(\rho_0^m)\big]$, take a measurement of a "*sequential observable* $[\{\mathbf{O}_t\}_{t\in T}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\rightarrow} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]$", i.e., take a measurement of an observable $\mathbf{O}_0$ at 0 ($\in T$), and next, take a measurement of an observable $\mathbf{O}_1$ at 1 ($\in T$), $\cdots\cdots$, and finally take a measurement of an observable $\mathbf{O}_7$ at 7 ($\in T$),

which is symbolized by $\mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}_{[\rho_0^p]})$ $\big[$resp. $\mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}(\rho_0^m))\big]$. Note that the $\mathfrak{M}$ $(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}_{[\rho_0^p]})$ $\big[$resp. $\mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}(\rho_0^m))\big]$ is merely a symbol since the above ($\sharp$) is a rough statement which seems to include "many measurements" $\big($in spite of the spirit that only one measurement is permitted in measurement theory (cf. Remark 2.4)$\big)$. In what follows let us describe the above ($\sharp$) ($= \mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}_{[\rho_0^p]})$) $\big[$resp. ($= \mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}(\rho_0^m))\big]$ precisely. Put

$$\widetilde{\mathbf{O}}_t = \mathbf{O}_t \ \text{ and thus } \ \widetilde{F}_t = F_t \quad (t = 3, 4, 5, 6, 7).$$

First we construct the quasi-product observable $\widetilde{\mathbf{O}}_2$ in $\mathcal{A}_2$ such as

$$\widetilde{\mathbf{O}}_2 = (X_2 \times X_3 \times X_4, 2^{X_2\times X_3\times X_4}, \widetilde{F}_2) \ \text{ where } \ \widetilde{F}_2 = F_2 \overset{\text{qp}}{\times} (\overset{\text{qp}}{\times}_{t=3,4} \Phi_{2,t}\widetilde{F}_t), \tag{2.20}$$

if it exists. Iteratively, we construct the following:

$$\begin{array}{ccccc}
\mathcal{A}_0 & \xleftarrow{\Phi_{0,1}} & \mathcal{A}_1 & \xleftarrow{\Phi_{1,2}} & \mathcal{A}_2 \\
F_0 \overset{\text{qp}}{\times} \Phi_{0,6}\widetilde{F}_6 \overset{\text{qp}}{\times} \Phi_{0,7}\widetilde{F}_7 & & F_1 \overset{\text{qp}}{\times} \Phi_{1,5}\widetilde{F}_5 & & \\
\downarrow & & \downarrow & & \\
\widetilde{F}_0 & \xleftarrow{\Phi_{0,1}} & \widetilde{F}_1 & \xleftarrow{\Phi_{1,2}} & \widetilde{F}_2 \\
{\scriptstyle (F_0 \overset{\text{qp}}{\times} \Phi_{0,6}\widetilde{F}_6 \overset{\text{qp}}{\times} \Phi_{0,7}\widetilde{F}_7 \overset{\text{qp}}{\times} \Phi_{0,1}\widetilde{F}_1)} & & {\scriptstyle (F_1 \overset{\text{qp}}{\times} \Phi_{1,5}\widetilde{F}_5 \overset{\text{qp}}{\times} \Phi_{1,2}\widetilde{F}_2)} & & {\scriptstyle (F_2 \overset{\text{qp}}{\times} \Phi_{2,3}\widetilde{F}_3 \overset{\text{qp}}{\times} \Phi_{2,4}\widetilde{F}_4)}
\end{array} \tag{2.21}$$

That is, we get the quasi-product observable $\widetilde{\mathbf{O}}_1 \equiv (\prod_{t=1}^5 X_t, 2^{\Pi_{t=1}^5 X_t}, \widetilde{F}_1)$ of $\mathbf{O}_1$, $\Phi_{1,2}\widetilde{\mathbf{O}}_2$ and $\Phi_{1,5}\widetilde{\mathbf{O}}_5$, and finally, the quasi-product observable $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t=0}^7 X_t, 2^{\Pi_{t=0}^7 X_t}, \widetilde{F}_0)$ of $\mathbf{O}_0$, $\Phi_{0,1}\widetilde{\mathbf{O}}_1$, $\Phi_{0,6}\widetilde{\mathbf{O}}_6$ and $\Phi_{0,7}\widetilde{\mathbf{O}}_7$, if it exists. Here, $\widetilde{\mathbf{O}}_0$ is called the *realization* (or, *Heisenberg picture representation*) of a *sequential observable* $[\{\mathbf{O}_t\}_{t\in T}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\rightarrow} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]$. Then, we have the measurement $\big[$resp. the statistical measurement$\big]$:

$$\mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0, S_{[\rho_0^p]}) \quad \big[\text{resp. } \mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0, S(\rho_0^m))\big] \tag{2.22}$$

which is called the *realization* (or, *Heisenberg picture representation*) *of the symbol* $\mathfrak{M}$ $(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}_{[\rho_0^p]})$ $\big[$resp. $\mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}(\rho_0^m))\big]$.

Now, we can propose Axiom 2, which corresponds to "the rule of the relation among systems" in PMT (2.1) and SMT(2.6).

Examining Example 2.8, we see as follows. Let ($T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T$) be a tree with root 0 and let $\mathbf{S}_{[\rho_0^p]} \equiv [S_{[\rho_0^p]}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t \in T\setminus\{0\}}]$ be a general system with the initial system $S_{[\rho_0^p]}$. And, let an observable $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ in a C^*-algebra $\mathcal{A}_t$ be given for each $t \in T$. And further, for any s $(\in T)$, put $T_s \equiv \{t \in T \mid s \leq t\}$. For each s $(\in T)$, define the observable $\widetilde{\mathbf{O}}_s \equiv (\prod_{t \in T_s} X_t, 2^{\Pi_{t\in T_s} X_t}, \widetilde{F}_s)$ in $\mathcal{A}_s$ such that

$$\widetilde{\mathbf{O}}_s = \begin{cases} \mathbf{O}_s & (\text{if } s \in T \setminus \pi(T)) \\ \mathbf{O}_s \overset{\text{qp}}{\times} (\overset{\text{qp}}{\times}_{t \in \pi^{-1}(\{s\})} \Phi_{\pi(t),t} \widetilde{\mathbf{O}}_t) & (\text{if } s \in \pi(T)), \end{cases} \tag{2.23}$$

if possible (i.e., if the existence of the quasi-product observable $\widetilde{\mathbf{O}}_s$ is guaranteed). Then, if an observable $\widetilde{\mathbf{O}}_0$ (i.e., the Heisenberg picture representation of the sequential observable $[\{\mathbf{O}_t\}_{t\in T}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t \in T\setminus\{0\}}]$) in $\mathcal{A}_0$ exists (such as in Example 2.8), we have the measurement

$$\mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0, S_{[\rho_0^p]}), \quad [\text{resp. } \mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0, S(\rho_0^m))] \tag{2.24}$$

which is called the *Heisenberg picture representation of the symbol* $\mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}_{[\rho_0^p]})$ $[\text{resp. } \mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}(\rho_0^m))]$.

Summing up the essential part of the above argument, we can propose the following axiom, which corresponds to "the rule of the relation among systems" in MT.

AXIOM 2. (The Markov relation among systems, the Heisenberg picture). *The relation among systems is represented by a Markov relation* $\{\Phi_{t_1,t_2} : \mathcal{A}_{t_2} \to \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}$. *Let* $\mathbf{O}_t$ $(\equiv (X_t, 2^{X_t}, F_t))$ *be an observable in* $\mathcal{A}_t$ *for each* t $(\in T)$. *If the procedure (2.23) is possible, a sequential observable* $[\mathbf{O}_T] \equiv [\{\mathbf{O}_t\}_{t\in T}; \{\Phi_{t_1,t_2} : \mathcal{A}_{t_2} \to \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}]$ *can be realized as the observable* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\Pi_{t\in T} X_t}, \widetilde{F}_0)$ *in* $\mathcal{A}_0$.

Also, we must add the following statement, which explains the relation between Axiom 1 [resp. Proclaim 1] and Axiom 2:

- Let $\mathbf{S}_{[\rho_0^p]} \equiv [S_{[\rho_0^p]}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]$ $[\text{resp. } \mathbf{S}(\rho_0^m) \equiv [S(\rho_0^m); \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\to} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]]$ be a general system with an initial state ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ [resp. with an initial statistical state ρ_0^m $(\in \mathfrak{S}^m(\mathcal{A}^*))$]. And, a measurement represented by the symbol $\mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}_{[\rho_0^p]})$ $[\text{resp. } \mathfrak{M}(\{\mathbf{O}_t\}_{t\in T}, \mathbf{S}(\rho_0^m))]$ can be realized by $\mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\Pi_{t\in T} X_t}, \widetilde{F}_0), S_{[\rho_0^p]})$ $[\text{resp. } \mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\Pi_{t\in T} X_t}, \widetilde{F}_0), S(\rho_0^m))]$, if $\widetilde{\mathbf{O}}_0$ exists.

In physics, T always represents the time axis, and therefore, $\{\mathcal{A}_{t_2} \overset{\Phi_{t_1,t_2}}{\longrightarrow} \mathcal{A}_{t_1}\}_{(t_1,t_2)\in T^2_{\leq}}$ represents the time revolution of the mechanical system.

Now we have two measurement theories (i.e., PMT and SMT) as follows.

$$\begin{cases} \text{PMT} = \underset{\text{(Axiom 1)}}{\text{measurement}} + \underset{\text{(Axiom 2)}}{\text{the relation among systems}}, \\ \\ \text{SMT} = \underset{\text{(Proclaim 1)}}{\text{statistical measurement}} + \underset{\text{(Axiom 2)}}{\text{the relation among systems}} \end{cases}$$

Here, it should be noted that Axiom 2 is common to PMT(2.1) and SMT(2.6).

Summing up the above argument in Example 2.8, we can mention the following theorem.

THEOREM 2.9. (The measurability of a general system). *Let* $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$ *be a tree with root* 0 *and let* $\mathbf{S}_{[\rho_0^p]} \equiv [S_{[\rho_0^p]}; \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\longrightarrow} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]$ $\big[$*resp.* $\mathbf{S}(\rho_0^m) \equiv [S(\rho_0^m); \{\mathcal{A}_t \overset{\Phi_{\pi(t),t}}{\longrightarrow} \mathcal{A}_{\pi(t)}\}_{t\in T\setminus\{0\}}]\big]$ *be a general system with the initial system* $S_{[\rho_0^p]}$ $\big[$*resp.* $S(\rho_0^m)\big]$. *And, let an observable* $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ *in a* C^**-algebra* $\mathcal{A}_t$ *be given for each* $t \in T$. *Then, if an observable* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\Pi_{t\in T} X_t}, \widetilde{F}_0)$ *in* $\mathcal{A}_0$ *exists (cf. the formulation (2.23)), we have the measurement*

$$\mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0, S_{[\rho_0^p]}) \quad \big[\textit{resp. } \mathbf{M}_{\mathcal{A}_0}(\widetilde{\mathbf{O}}_0, S(\rho_0^m))\big]. \tag{2.25}$$

If the system is classical, i.e., $\mathcal{A}_t \equiv C(\Omega_t)$ $(\forall t \in T)$, *then the measurement (2.25) always exists, while the uniqueness is not always guaranteed. Also, it should be noted, by (2.23), that, for each* s $(\in T \setminus \{0\})$, *it holds that* $\Phi_{\pi(s),s}\widetilde{F}_s(\prod_{t\in T_s} \Xi_t) = \widetilde{F}_{\pi(s)}\big((\prod_{t\in T_{\pi(s)}\setminus T_s} X_t) \times (\prod_{t\in T_s} \Xi_t)\big)$ $(\forall \Xi_t \in 2^{X_t}$ $(\forall t \in T_s))$.

PROOF. It suffices to prove it in classical measurements. However, it is clear since, in classical measurements, the product observable of any observables always exists (cf. the formula (2.10)). Therefore the construction mentioned in Example 2.8 is always possible in classical systems. □

3 Regression analysis in PMT (2.1)

In this section, we study regression analysis in PMT, i.e., Axioms 1 and 2.

3.1 Our motivation

In order to explain our main assertion, let us begin with the following example (the conventional argument of regression analysis in Fisher's maximum likelihood method), which is not only well known but also located in the central point of statistics.

Example 3.1. (The conventional argument of regression analysis in Fisher's method). We have a water tank of rectangular shape filled with some water. Assume that the height of water at time t is given by the following function $h(t)$:

$$h(t) = \alpha_0 + \beta_0 t, \tag{3.1}$$

where α_0 and β_0 are unknown fixed parameters such that α_0 is the height of water filling the tank at the beginning and β_0 is the increasing height of water per a unit time. The measured height $h_m(t)$ of water at time t is assumed to be represented by

$$h_m(t) = \alpha_0 + \beta_0 t + e(t), \tag{3.2}$$

where $e(t)$ represents a noise (or more precisely, a measurement error) with some suitable conditions. And assume that we obtained the measured data of the heights of water at $t = 1, 2, 3$ as follows:

$$h_m(1) = 1.9, \quad h_m(2) = 3.0, \quad h_m(3) = 4.7. \tag{3.3}$$

Under this setting, we consider the following problem.

(i) Infer the true value $h(2)$ of the water height at $t = 2$ from the measured data (3.3).

This problem (i) is usually solved as follows. From the theoretical point of view, we can infer, by Fisher's maximum likelihood method and regression analysis, that

$$(\alpha_0, \beta_0) = (0.4, 1.4). \tag{3.4}$$

$\Big($For the derivation of (3.4) from (3.3), see Example 3.6 later.$\Big)$ And next, we can infer that

$$h(2) = 3.2, \tag{3.5}$$

by the calculation: $h(2) = 0.4 + 1.4 \times 2 = 3.2$. This is the answer to the problem (i).

The above argument in Example 3.1 is, of course, well known and adopted as the usual regression analysis. Thus all statisticians may think that there is no serious problem in regression analysis. However, it is not true. For example, we have the basic problem in the argument of Example 3.1 as follows.

(ii) What kinds of axioms are hidden behind the argument in Example 3.1? And moreover, justify the argument in Example 3.1 under the axioms.

It is important. If we have no answer to the question: "What kinds of rules are permitted to be used in statistics?", we can not prove (or, justify) that the argument in Example 3.1 is true (or not). That is because there is no justification without an axiomatic formulation. In this sense, we believe that the above (ii) is the most important problem in theoretical statistics. Also, if someone knows the great success of the axiomatic formulation in physics (e.g., the three laws in Newtonian mechanics, or von Neumann's formulation of quantum mechanics, cf. von Neumann (1932)), it is a matter of course that he wants to understand statistics axiomatically.

Trying to solve the problem (ii), some may consider as follows.

(iii) Firstly, Fisher's maximum likelihood method should be declared as an axiom. Also, the derivation of (3.5) from (3.4) should be justified under some axioms. That is, it must not be accepted as a common sense.

This opinion (iii) may not be far from our assertion proposed in this paper. However, in order to describe the above (iii) precisely, we must make vast preparations.

It should be noted that theoretical statistics already has the mathematical formulation, called "*Kolmogorov's probability theoretical formulation*" (cf. Kolmogorov (1950)). Nevertheless, the problem (ii) has not be solved yet. This is, of course, due to the fact that "the mathematical formulation" does not always mean "the axiomatic formulation". However, we may expect that the reverse is true, that is, "the axiomatic formulation" always implies "the mathematical formulation". That is because, in the great history of physics, we always see that the true axiomatic formulation (of physics) is not only described in terms of mathematics but also accepted as the true mathematical formulation.

3.2 Fisher's maximum likelihood method

Consider a measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, 2^X, F),\ S_{[\rho^p]})$ formulated in a C^*-algebra $\mathcal{A}$.

In most measurements, it is usual to consider that the state ρ^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ is unknown. That is because the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ may be taken in order to know the state ρ^p. Thus, under the condition that we do not know the state ρ^p, the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$ is often denoted by $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$. By using this notation, we can say our present problem as follows:

(I) Infer the unknown state $[*]$ $(\in \mathfrak{S}^p(\mathcal{A}^*))$ from the measured data obtained by the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, 2^X, F), S_{[*]})$.

In order to answer this problem, in Ishikawa (2000) we introduced Fisher's method (precisely, Fisher's maximum likelihood method) as follows. $\Big($Strictly speaking, Theorem 3.2 should not be called "theorem" but "assertion", since it is not a purely mathematical result but a consequence of Axiom 1.$\Big)$

THEOREM 3.2. (Fisher's maximum likelihood method in classical and quantum measurements). *Consider a measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, 2^X, F), S_{[*]})$ *in* $\mathcal{A}$. *When we know that the measured value obtained by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$ *belongs to* Ξ $(\in 2^X)$, *there is a reason to infer that the state* $[*]$ *of the system* S *is equal to* ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ *such that*

$$ {}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \rangle_{\mathcal{A}} = \max_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)} {}_{\mathcal{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathcal{A}}. \tag{3.6} $$

Here, note, for completeness, that the state $[*]$ *(in* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$*) is the state before the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$. $\Big($*Cf. Corollary 3.4 later.*$\Big)$

PROOF. See Ishikawa (2000). To make it self-contained, we add the proof (presented in Ishikawa (2000)) as follows. Let ρ_1^p and ρ_2^p be elements in $\mathfrak{S}^p(\mathcal{A}^*)$. Assume that $\rho_1^p(F(\Xi)) < \rho_2^p(F(\Xi))$. Then, Axiom 1 says that the fact that the measured value obtained by the $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho_1^p]})$ belongs to Ξ happens more rarely than the fact that the measured value obtained by the $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho_2^p]})$ belongs to Ξ happens. Since $\rho^p(F(\Xi)) \le \rho_0^p(F(\Xi))$ $(\forall \rho^p \in \mathfrak{S}^p(\mathcal{A}^*))$, there is a reason to regard the unknown state $[*]$ as the state ρ_0^p. Examining this proof, we can easily see that the state $[*]$ (in $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$) is the state before the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$. This completes the proof. $\square$

The following corollary is a direct consequence of Theorem 3.2.

COROLLARY 3.3. (The conditional probability representation of Fisher's method). *Let* $\mathbf{O} \equiv (X, 2^X, F)$ *and* $\mathbf{O}' \equiv (Y, 2^Y, G)$ *be observables in* $\mathcal{A}$. *Let* $\widehat{\mathbf{O}}$ *be a quasi-product observable of* $\mathbf{O}$ *and* $\mathbf{O}'$, *that is,* $\widehat{\mathbf{O}} \equiv \mathbf{O} \overset{\text{qp}}{\times} \mathbf{O}' = (X \times Y, 2^{X \times Y}, F \overset{\text{qp}}{\times} G)$. *Assume that we know that the measured value* (x, y) $(\in X \times Y)$ *obtained by a measurement* $\mathbf{M}_{\mathcal{A}}(\widehat{\mathbf{O}}, S_{[*]})$ *belongs to* $\Xi \times Y$ $(\in 2^{X \times Y})$. *Then, there is a reason to infer that the unknown measured value* y $(\in Y)$ *is distributed under the conditional probability* $P_{\Xi}(\cdot)$:

$$P_{\Xi}(\Gamma) = \frac{{}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \overset{\text{qp}}{\times} G(\Gamma) \rangle_{\mathcal{A}}}{{}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \rangle_{\mathcal{A}}} \left(\equiv \frac{\rho_0^p(F(\Xi) \overset{\text{qp}}{\times} G(\Gamma))}{\rho_0^p(F(\Xi))} \right) \qquad (\forall \Gamma \in 2^Y), \tag{3.7}$$

where ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ *is defined by*

$${}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \rangle_{\mathcal{A}} = \max_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)} {}_{\mathcal{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathcal{A}}. \tag{3.8}$$

PROOF. Since we know that the measured value (x, y) $(\in X \times Y)$ obtained by a measurement $\mathbf{M}_{\mathcal{A}}(\widehat{\mathbf{O}}, S_{[*]})$ belongs to $\Xi \times Y$ $(\in 2^{X \times Y})$, we can infer, by Theorem 3.2 (Fisher's method) and the equality $F(\Xi) = F(\Xi) \overset{\text{qp}}{\times} G(Y)$, that the unknown state $[*]$ (in $\mathbf{M}_{\mathcal{A}}(\widehat{\mathbf{O}}, S_{[*]})$) is equal to ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$. Thus, the conditional probability $P_{\Xi}(\cdot)$ under the condition that we know that $(x, y) \in \Xi \times Y$ is given by

$$P_{\Xi}(\Gamma) = \frac{\rho_0^p(F(\Xi) \overset{\text{qp}}{\times} G(\Gamma))}{\rho_0^p(F(\Xi) \overset{\text{qp}}{\times} G(Y))} = \frac{\rho_0^p(F(\Xi) \overset{\text{qp}}{\times} G(\Gamma))}{\rho_0^p(F(\Xi))} \qquad (\forall \Gamma \in 2^Y). \tag{3.10}$$

This completes the proof. □

The following corollary is essential in classical measurements. That is because what we want to infer is usually "the state after the measurement" (or precisely, "the S-state after the measurement", cf. Definition 3.9) and not "the state before the measurement".

COROLLARY 3.4. (Fisher's maximum likelihood method in classical measurements). *Let* $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, 2^X, F), S_{[*]})$ *be a measurement formulated in a commutative* C^*-*algebra* $C(\Omega)$. *Assume that we know that the measured value obtained by the measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$ *belongs to* Ξ $(\in 2^X)$. *Then, we see that*

(i) *there is a reason to infer that the state* $[*]$ *of the system* S $\Big($*i.e., "the state before the measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$*"*$\Big)$ *is equal to* δ_{ω_0} $(\in \mathcal{M}_{+1}^p(\Omega))$, *where*

$$[F(\Xi)](\omega_0) = \max_{\omega \in \Omega} [F(\Xi)](\omega), \tag{3.10}$$

and,

(ii) *there is a reason to infer that the state after the measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$ *is also regarded as the same* δ_{ω_0} $(\in \mathcal{M}_{+1}^p(\Omega))$.

Summing up the above (i) and (ii), we see that

(iii) *there is a reason to infer that*

$$[*] = \text{"the state after the measurement } \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})\text{"} = \delta_{\omega_0}. \tag{3.11}$$

Proof. (i) is the special case of Theorem 3.2 (Fisher's method), i.e., $\mathcal{A} = C(\Omega)$. Thus it suffices to prove (ii) as follows. $\Big($This (ii) will be again proved in Remark 3.14 as a special case of Lemma 3.13 later. Thus, the proof presented below may be somewhat temporary.$\Big)$ Let $\mathbf{O}' \equiv (Y, 2^Y, G)$ be any observable in $C(\Omega)$. Let $\widehat{\mathbf{O}}$ be the product observable of $\mathbf{O}$ and $\mathbf{O}'$, that is, $\widehat{\mathbf{O}} \equiv \mathbf{O} \times \mathbf{O}' = (X \times Y, 2^{X \times Y}, F \times G)$. Consider a measurement $\mathbf{M}_{C(\Omega)}(\widehat{\mathbf{O}} \equiv (X \times Y, 2^{X \times Y}, F \times G), S_{[*]})$. And assume

(A) we know that the measured value (x, y) $(\in X \times Y)$ obtained by the measurement $\mathbf{M}_{C(\Omega)}(\widehat{\mathbf{O}} \equiv (X \times Y, 2^{X \times Y}, F \times G), S_{[*]})$ belongs to $\Xi \times Y$.

Then, Corollary 3.3 says that there is a reason to infer that the unknown measured value y $(\in Y)$ is distributed under the conditional probability $P_\Xi(\cdot)$, where

$$P_\Xi(\Gamma) = \frac{[F(\Xi) \times G(\Gamma)](\omega_0)}{[F(\Xi)](\omega_0)} = [G(\Gamma)](\omega_0) \qquad (\forall \Gamma \in 2^Y), \tag{3.12}$$

where ω_0 $(\in \Omega)$ is defined in (3.10). Also note that the above (A) can be represented by the following two steps (A_1) and (A_2) $\Big($i.e., $(\mathrm{A}) = (\mathrm{A}_1) + (\mathrm{A}_2)\Big)$:

(A_1) we know that the measured value by a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, 2^X, F), S_{[*]})$ belongs to Ξ $(\in 2^X)$,

and

(A_2) And successively, we take a measurement of the observable $\mathbf{O}' \equiv (Y, 2^Y, G)$, and get a measured value y $(\in Y)$.

$\Big($The above $\big($i.e., (A) = (A$_1$) + (A$_2$)$\big)$ is somewhat metaphorical since "two measurements" seem to appear in (A$_1$) and (A$_2$). (*Cf.* Remarks 2.4 and 3.14.)$\Big)$ Comparing (A) and "(A$_1$) + (A$_2$)", we see, by (3.12), that

$$\text{“the probability that the measured value } y \text{ belongs to } \Gamma \ (\in 2^Y) \text{ in (A}_2\text{)”} = [G(\Gamma)](\omega_0) \tag{3.13}$$

That is, we get the sample space $(Y, 2^Y, [G(\cdot)](\omega_0))$ in (A$_2$). Since $\mathbf{O}' \equiv (Y, 2^Y, G)$ is arbitrary, we say that

(B) the state after (A$_1$) $\Big($i.e., the state after the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})\Big)$ is equal to δ_{ω_0} $\Big($since the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}', S_{[\delta_{\omega_0}]})$ induces the sample space $(Y, 2^Y, [G(\cdot)]$ $(\omega_0))\Big)$.

This completes the proof. $\Big($This corollary does not hold in quantum measurements, since the product observable $\widehat{\mathbf{O}} \equiv \mathbf{O} \times \mathbf{O}' = (X \times Y, 2^{X \times Y}, F \times G)$ does not always exist. That is, the concept of "the state after a measurement" is not always meaningful in quantum theory.$\Big)$ □

The "Bayes operator (in the following remark)" is hidden in the above proof. This will be more completely clarified in Remark 3.14 later. Although Corollary 3.4 and Remark 3.5 may be temporary, we believe that they promote a better understanding of Remark 3.14.

REMARK 3.5. (1. Bayes operator). Let $\mathbf{O} \equiv (X, 2^X, F)$ be an observable in $C(\Omega)$. For each Ξ $(\in 2^X)$, define the continuous linear operator $B_{\Xi}^{(0,0)} : C(\Omega) \to C(\Omega)$ such that

$$B_{\Xi}^{(0,0)}(g) = F(\Xi) \cdot g \qquad (\forall g \in C(\Omega)), \tag{3.14}$$

which is called the *Bayes operator* (or, *simplest Bayes operator*). Note that it clearly holds that

(i) for any observable $\mathbf{O}' \equiv (Y, 2^Y, G)$, there exists an observable $\widehat{\mathbf{O}} \equiv (X \times Y, 2^{X \times Y}, \widehat{F})$ in $C(\Omega)$ such that

$$\widehat{F}(\Xi \times \Gamma) = B_{\Xi}^{(0,0)}(G(\Gamma)) \qquad (\forall \Xi \in 2^X, \forall \Gamma \in 2^Y). \tag{3.15}$$

That is because it suffices to define the $\widehat{\mathbf{O}}$ by the product observable $\mathbf{O} \times \mathbf{O}'$ in $C(\Omega)$. Define the map $R_{\Xi}^{(0,0)} : \mathcal{M}_{+1}^m(\Omega) \to \mathcal{M}_{+1}^m(\Omega)$ (which may be called "*normalized dual Bayes operator*") such that

$$R_{\Xi}^{(0,0)}(\nu) = \frac{(B_{\Xi}^{(0,0)})^*(\nu)}{\|(B_{\Xi}^{(0,0)})^*(\nu)\|_{\mathcal{M}(\Omega)}} \qquad (\forall \nu \in \mathcal{M}_{+1}^m(\Omega)), \tag{3.16}$$

where $(B_{\Xi}^{(0,0)})^* : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ $\Big($i.e., $C(\Omega)^* \to C(\Omega)^*\Big)$ is the adjoint operator of $B_{\Xi}^{(0,0)}$, that is, it holds that ${}_{C(\Omega)^*}\langle (B_{\Xi}^{(0,0)})^*(\nu), g \rangle_{C(\Omega)} = {}_{C(\Omega)^*}\langle \nu, B_{\Xi}^{(0,0)}(g) \rangle_{C(\Omega)}$ $(\forall \nu \in C(\Omega)^* \equiv \mathcal{M}(\Omega), \forall g \in C(\Omega))$. Using it, we can describe the well known Bayes theorem (cf. Ishikawa (2000)) such as

$$\mathcal{M}_{+1}^m(\Omega) \ni \nu \ (= \text{a priori state}) \mapsto (\text{posterior state} =) \ R_{\Xi}^{(0,0)}(\nu) \in \mathcal{M}_{+1}^m(\Omega). \tag{3.17}$$

Note that (3.17) says that (i)⇒(ii) in Corollary 3.4, since a simple calculation shows that $R_{\Xi}^{(0,0)}(\delta_{\omega_0}) = \delta_{\omega_0}$ in the case of Corollary 3.4. In Section 3.4, readers will again study the Bayes operator in more general situations.

(2. "Before" and "after"). The term "before" [or, "after"] in "the state before the measurement" [or, "the state after the measurement"] is somewhat metaphorical. Note that the concept of "time" is not included in Axioms 1 and 2. The tree T does not always represent "time" in MT.

3.3 Regression analysis I (the conventional form)

From here onward, we always devote ourselves to the classical cases, that is, $\mathcal{A} = C(\Omega)$. Under the preparations in the previous sections, we can propose that

"Regression Analysis I (the conventional regression analysis)"

= "Theorem 2.9 (measurability)" + "Corollary 3.4 (classical Fisher's method)" .
(in the case: $\mathcal{A} = C(\Omega)$)

(3.18)

That is, we can assert:

REGRESSION ANALYSIS I. (The conventional form). *Let* $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$ *be a tree with root* 0, *and let* $\mathbf{S}_{[*]} \equiv [S_{[*]}; \{\Phi_{\pi(t),t} : C(\Omega_t) \to C(\Omega_{\pi(t)})\}_{t \in T \setminus \{0\}}]$ *be a general system with the initial system* $S_{[*]}$. *And, let an observable* $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ *in* $C(\Omega_t)$ *be given for each* $t \in T$. *Then, we have a measurement*

$$\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t \in T} X_t, 2^{\prod_{t \in T} X_t}, \widetilde{F}_0), S_{[*]}). \qquad (\textit{Cf. Theorem 2.9}). \tag{3.19}$$

Assume that the measured value by the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$ *belongs to* $\prod_{t \in T} \Xi_t$ $(\in 2^{\prod_{t\in T} X_t})$. *Then, there is a reason to infer that the state* $[*]$ *of the system* S $\Big($*i.e., the state before the measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\Big)$, *the state after the measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$ *and the* δ_{ω_0} $(\in \mathcal{M}_{+1}^p(\Omega))$ *(defined by (3.21)) are equal. That is, Corollary 3.4 says that there is a reason to infer that*

$$[*] = \text{"the state after the measurement } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{"} = \delta_{\omega_0}. \tag{3.20}$$

Here the δ_{ω_0} $(\in \mathcal{M}_{+1}^p(\Omega_0))$ *is defined by*

$$[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega_0) = \max_{\omega \in \Omega_0} [\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega). \tag{3.21}$$

Now we shall review Example 3.1 in the light of Regression Analysis I.

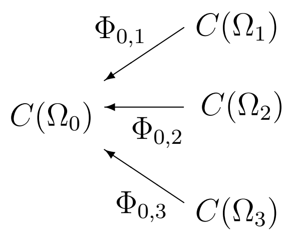
Example 3.6. (Continued from Example 3.1, the conventional argument of regression analysis in Fisher's method). Put $\Omega_0 = [0, 1.0] \times [0, 2.0]$, and put $\Omega_1 = \Omega_2 = \Omega_3 = [0, 10.0]$. For each t $(\in \{1, 2, 3\})$, define a continuous map $\phi_{0,t} : \Omega_0 \to \Omega_t$ such that

$$\Omega_0(\equiv [0, 1.0] \times [0, 2.0]) \ni \omega \equiv (\alpha, \beta) \overset{\phi_{0,t}}{\mapsto} \alpha + \beta t \in \Omega_t(\equiv [0, 10.0]). \tag{3.22}$$

Thus, for each t $(\in \{1, 2, 3\})$, we have a homomorphism $\Phi_{0,t} : C(\Omega_t) \to C(\Omega_0)$ such that

$$[\Phi_{0,t} f_t](\omega) = f_t(\phi_{0,t}(\omega)) \qquad (\forall \omega \in \Omega_0, \forall f_t \in C(\Omega_t)). \tag{3.23}$$

It is usual to assume that regression analysis is applied to the system with a parallel structure such as in the figure (3.24). $\Big($From the peculiarity of this problem, we can also assume that this system has a series structure. However, we are not concerned with it.$\Big)$



(3.24)

For each $t \in \{1, 2, 3\}$, consider the discrete Gaussian observable $\mathbf{O}_{\sigma,N} \equiv (X_N, 2^{X_N}, F_{\sigma,N})$ in $C(\Omega_t)$, cf. Example 2.2. Here, we define the observable $\widetilde{\mathbf{O}}_0 \equiv (X_N^3, 2^{X_N^3}, \widetilde{F}_0)$ in $C(\Omega_0)$

such that

$$
\begin{aligned}
&[\widetilde{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\omega) = [\Phi_{0,1} F_{\sigma,N}(\Xi_1)](\omega) \cdot [\Phi_{0,2} F_{\sigma,N}(\Xi_2)](\omega) \cdot [\Phi_{0,3} F_{\sigma,N}(\Xi_3)](\omega) \\
=&[F_{\sigma,N}(\Xi_1)](\phi_{0,1}(\omega)) \cdot [F_{\sigma,N}(\Xi_2)](\phi_{0,2}(\omega)) \cdot [F_{\sigma,N}(\Xi_3)](\phi_{0,3}(\omega)) \\
&\qquad (\forall \Xi_1, \Xi_2, \Xi_3 \in 2^{X_N},\ \forall \omega = (\alpha, \beta) \in \Omega_0 = [0, 1.0] \times [0, 2.0]).
\end{aligned} \tag{3.25}
$$

Then, we have the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$. The data (3.3) says that the measured value obtained by the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$ is equal to

$$
(1.9,\ 3.0,\ 4.7)\ (\in X_N^3). \tag{3.26}
$$

Here, Fisher's method (Corollary 3.4) says that it suffices to solve the problem

$$
\text{"Find } (\alpha_0, \beta_0) \text{ such as } \max_{(\alpha,\beta)\in\Omega_0} [\widetilde{F}_0(\{1.9\} \times \{3.0\} \times \{4.7\})](\alpha, \beta)\text{"}. \tag{3.27}
$$

Putting

$$
\Xi_1 = [1.9 - \frac{1}{2N}, 1.9 + \frac{1}{2N}],\ \Xi_2 = [3.0 - \frac{1}{2N}, 3.0 + \frac{1}{2N}],\ \Xi_3 = [4.7 - \frac{1}{2N}, 4.7 + \frac{1}{2N}],
$$

we see, under the assumption that N is sufficiently large, that

$$
\begin{aligned}
&(3.27) \Rightarrow \max_{(\alpha,\beta)\in\Omega_0} \frac{1}{\sqrt{2\pi\sigma^2}^3} \iiint_{\Xi_1\times\Xi_2\times\Xi_3} e^{[-\frac{(x_1-(\alpha+\beta))^2+(x_2-(\alpha+2\beta))^2+(x_3-(\alpha+3\beta))^2}{2\sigma^2}]} dx_1 dx_2 dx_3 \\
&\Rightarrow \max_{(\alpha,\beta)\in\Omega_0} \exp\Big[-[(1.9-(\alpha+\beta))^2 + (3.0-(\alpha+2\beta))^2 + (4.7-(\alpha+3\beta))^2]/(2\sigma^2)\Big] \\
&\Rightarrow \min_{(\alpha,\beta)\in\Omega_0} [(1.9-(\alpha+\beta))^2 + (3.0-(\alpha+2\beta))^2 + (4.7-(\alpha+3\beta))^2] \\
&\qquad\qquad \text{(by the least squares method)} \\
&\Rightarrow \begin{cases} (1.9-(\alpha+\beta)) + (3.0-(\alpha+2\beta)) + (4.7-(\alpha+3\beta)) = 0 \\ (1.9-(\alpha+\beta)) + 2(3.0-(\alpha+2\beta)) + 3(4.7-(\alpha+3\beta)) = 0 \end{cases} \\
&\Rightarrow \quad (\alpha_0, \beta_0) = (0.4, 1.4).
\end{aligned} \tag{3.28}
$$

This is the conclusion of Regression Analysis I. Also, using the notations in Regression Analysis I, we remark that

(R) the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t,\ 2^{\prod_{t\in T} X_t},\ \widetilde{F}_0), S_{[*]})$ is hidden behind the inference (3.28) $\Big(= (3.4) \text{ in Example 3.1}\Big)$.

This fact will be important in Section 3.5.

The above may be the standard argument of the conventional regression analysis in measurement theory. However, our problem (i) in Example 3.1 is not to infer the (α_0, β_0) but $h(2)$. In this sense the above regression analysis I is not sufficient. As the answer of the problem (i) in Example 3.1, we usually consider that it suffices to calculate $h(2)$ $\Big(\equiv \phi_{0,2}(0.4, 1.4)\Big)$ in the following:

$$h(2) = 0.4 + 1.4 \times 2 = 3.2. \tag{3.29}$$

However, this is doubtful. $\Big($In fact, this (3.29) is not always true in general situations. (*Cf.* Regression analysis II (3.59) and (3.60) later).$\Big)$ Recall that our purpose of this paper is to propose "an axiomatic understanding of statistics". Thus we should not rely on "a common sense" but Axioms 1 and 2. That is, we must solve the problem:

- How can the above (3.29) $\Big(= (3.5)$ in Example 3.1$\Big)$ be deduced from Axioms 1 and 2?

In order to do this, we will make some preparations in the next section.

3.4 Bayes operator, Schrödinger picture and S-states

In order to improve Regression Analysis I (introduced in Section 3.3), in this section we make some preparations (i.e., Bayes operator, Schrödinger picture, S-state, etc.). Our main assertion (Regression Analysis II) will be proposed in Section 3.5. We begin with the following definition, which is a general form of "Bayes operator" in Remark 3.5.

Definition 3.7. (Bayes operator). Let $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$ be a tree with root 0 and let $\mathbf{S}_{[*]} \equiv [S_{[*]}; \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t \in T\setminus\{0\}}]$ be a general system with the initial system $S_{[*]}$. And, let an observable $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ in $C(\Omega_t)$ be given for each $t \in T$. Let $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0)$ be as in Theorem 2.9 in the case $\mathcal{A}_t = C(\Omega_t)$ $(\forall t \in T)$. That is, $\widetilde{\mathbf{O}}_0$ is the Heisenberg picture representation of the sequential observable $[\{\mathbf{O}_t\}_{t\in T}; \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T\setminus\{0\}}]$. Let τ be any element in T. If a positive bounded linear operator $B^{(0,\tau)}_{\prod_{t\in T}\Xi_t} : C(\Omega_\tau) \to C(\Omega_0)$ satisfies the following condition (BO), we call $\{B^{(0,\tau)}_{\prod_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ [resp. $B^{(0,\tau)}_{\prod_{t\in T}\Xi_t}$] a *family of Bayes operators* [resp. a *Bayes operator*]:

(BO) for any observable $\mathbf{O}'_\tau \equiv (Y_\tau, 2^{Y_\tau}, G_\tau)$ in $C(\Omega_\tau)$, there exists an observable $\widehat{\mathbf{O}}_0 \equiv \Big((\prod_{t\in T} X_t) \times Y_\tau, 2^{(\prod_{t\in T} X_t)\times Y_\tau}, \widehat{F}_0\Big)$ in $C(\Omega_0)$ such that

(i) $\widehat{\mathbf{O}}_0$ is the Heisenberg picture representation (cf. Theorem 2.9) of $[\{\overline{\mathbf{O}}_t\}_{t\in T}; \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{\ t\in T\setminus\{0\}}]$, where $\overline{\mathbf{O}}_t = \mathbf{O}_t$ (if $t \neq \tau$), $= \mathbf{O}_\tau \times \mathbf{O}'_\tau$ (if $t=\tau$),

(ii) $\widehat{F}_0((\prod_{t\in T}\Xi_t)\times\Gamma_\tau) = B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau)) \quad (\forall \Xi_t \in 2^{X_t}\ (\forall t \in T), \forall\Gamma_\tau \in 2^{Y_\tau})$,

(iii) $\widehat{F}_0((\prod_{t\in T}\Xi_t)\times Y_\tau) = \widetilde{F}_0(\prod_{t\in T}\Xi_t)\Big(\equiv B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)\Big)$, $(\forall \Xi_t \in 2^{X_t}\ (\forall t\in T))$, where 1_τ is the identity in $C(\Omega_\tau)$.

Also, define the map $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} : \mathcal{M}^m_{+1}(\Omega_0) \to \mathcal{M}^m_{+1}(\Omega_\tau)$ such that

$$R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\nu) = \frac{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\nu)}{\|(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\nu)\|_{\mathcal{M}(\Omega_\tau)}} \qquad (\forall\nu\in\mathcal{M}^m_{+1}(\Omega_0)) \tag{3.30}$$

where $(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^* : C(\Omega_0)^* \to C(\Omega_\tau)^*$ is the adjoint operator of $B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} : C(\Omega_\tau) \to C(\Omega_0)$. The map $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}$ may be called a *"normalized dual Bayes operator"*.

It holds that

$$B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(g_\tau) \leq \Phi_{0,\tau} g_\tau \qquad (\forall g_\tau \in C(\Omega_\tau) \text{ such that } g_\tau \geq 0), \tag{3.31}$$

because it holds, for any observable $\mathbf{O}'_\tau \equiv (Y_\tau, 2^{Y_\tau}, G_\tau)$ in $C(\Omega_\tau)$,

$$\begin{aligned} B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau)) &= \widehat{F}_0((\prod_{t\in T}\Xi_t)\times\Gamma_\tau) \leq \widehat{F}_0((\prod_{t\in T}X_t)\times\Gamma_\tau) \\ &= \Phi_{0,\tau}G_\tau(\Gamma_\tau)\Big(= B^{(0,\tau)}_{\Pi_{t\in T}X_t}(G_\tau(\Gamma_\tau))\Big) \qquad (\forall\Gamma_\tau\in 2^{X_\tau}). \end{aligned} \tag{3.32}$$

The following theorem is essential to Regression Analysis II later.

THEOREM 3.8. (The existence theorem of the Bayes operator (cf. Ishikawa (2001))). *Let $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T}X_t,\ 2^{\Pi_{t\in T}X_t},\ \widetilde{F}_0)$ be as in Theorem 2.9 in the case $\mathcal{A}_t = C(\Omega_t)$ $(\forall t\in T)$. And, for any s $(\in T)$, put $T_s \equiv \{t\in T \mid s\leq t\}$. Assume that, for each s $(\in T)$, there exists an observable $\widetilde{\mathbf{O}}_s \equiv (\prod_{t\in T_s}X_t, 2^{\Pi_{t\in T_s}X_t}, \widetilde{F}_s)$ in $C(\Omega_s)$ such that $\Phi_{\pi(s),s}\widetilde{F}_s(\prod_{t\in T_s}\Xi_t) = \widetilde{F}_{\pi(s)}\Big((\prod_{t\in T_{\pi(s)}\setminus T_s}X_t)\times(\prod_{t\in T_s}\Xi_t)\Big)$ $(\forall\Xi_t\in 2^{X_t}\ (\forall t\in T))$, (cf. Theorem 2.9). Let τ be any element in T. Then, there exists a family of Bayes operators $\{B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t\in T)\}$.*

PROOF. See Theorem 3.4 in Ishikawa (2001). The proof in Ishikawa (2001) is essentially true, but it is not complete. That is because the definition of "Bayes operator" (i.e., Definition 3.7) was not mentioned in Ishikawa (2001). Thus, we add the complete proof in Section 6 (Appendix). □

Let $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0)$, $\mathbf{O}'_\tau \equiv (Y_\tau, 2^{Y_\tau}, G_\tau)$, $\{B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$, $\widehat{\mathbf{O}}_0 \equiv ((\prod_{t\in T} X_t) \times Y_\tau, 2^{(\prod_{t\in T} X_t)\times Y_\tau}, \widehat{F}_0)$ and $\{R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ be as in Definition 3.7. Assume that

(C$_1$) we know that the measured value $(x_t)_{t\in T}$ $(\in (\prod_{t\in T} X_t))$ obtained by $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$ belongs to $\prod_{t\in T} \Xi_t$.

Note that this (C$_1$) is the same as the following (C$_2$).

(C$_2$) we know that the measured value $((x_t)_{t\in T}, y)$ $(\in (\prod_{t\in T} X_t)\times Y_\tau)$ obtained by $\mathbf{M}_{C(\Omega_0)}$ $(\widehat{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$ belongs to $(\prod_{t\in T} \Xi_t) \times Y_\tau$.

Thus we see that

(C$_3$) the probability distribution of unknown y $\Big($under the assumption (C$_2$) $\big($=(C$_1$)$\big)\Big)$, i.e., the probability that y $(\in Y_\tau)$ belongs to Γ_τ, is represented by

$$\frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widehat{F}_0((\prod_{t\in T} \Xi_t) \times \Gamma_\tau)\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widehat{F}_0((\prod_{t\in T} \Xi_t) \times Y_\tau)\rangle_{C(\Omega_0)}} \Bigg(\equiv \frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau))\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)\rangle_{C(\Omega_0)}} \Bigg). \tag{3.33}$$

A simple calculation shows:

$$(3.33) = {}_{C(\Omega_\tau)^*}\langle \frac{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})}{\|(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}}, G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)} = {}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}), G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)}.$$

Therefore, we say that

(C$_4$) the probability distribution of unknown y $\Big($under (C$_2$) $\big($=(C$_1$)$\big)\Big)$ is represented by

$${}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}), G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)}. \tag{3.34}$$

Let this (C$_4$) be, as an abbreviation, denoted by

(C$_5$) the *S-state* $\Big($after the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})\Big)$ at τ in T is equal to $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}$ (δ_{ω_0}).

For completeness, again note that (C$_4$) = (C$_5$), i.e., (C$_5$) is an abbreviation for (C$_4$). Note that the concept of "S-state" and that of "state" are completely different. In measurement theory, as seen in Axiom 1, the state always appears as the ρ^p in $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[\rho^p]})$. That is, the state is always fixed and never moves. In this sense it may be called a "real state". On the other hand, the "S-state" is used in the abbreviation (C$_5$) of (C$_4$).

Summing up the above argument, we have the following definition.

DEFINITION 3.9. (S-state, Schrödinger picture). Assume the above situation. If the above statement (C_4) holds, then we say "(C_5) *holds*", i.e., "*the S-state* $\Big($*after the measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})\Big)$ *at* τ $(\in T)$ *is equal to* $R^{(0,\tau)}_{\prod_{t\in T}\Xi_t}(\delta_{\omega_0})$". The representation using "S-state" is called the *Schrödinger picture representation.* The S-state is also called a *Schrödinger state* or *imaginary state.*

As seen in the above argument, we must note that the Bayes operator is always hidden behind the Scrödinger picture representation.

We sum up the above argument $\big($i.e., (C_1)$\Rightarrow$(C_5)$\big)$ as the following lemma.

LEMMA 3.10. (S-state). *Let* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0)$, $\{B^{(0,\tau)}_{\prod_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ *and* $\{R^{(0,\tau)}_{\prod_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ *be as in Definition 3.7. Assume that*

($\sharp$) *we know that the measured value* $(x_t)_{t\in T}$ $(\in \prod_{t\in T} X_t)$ *obtained by* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$ *belongs to* $\prod_{t\in T}\Xi_t$.

Then, we can say

($\flat$) *the S-state* $\big($*after the measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})\big)$ *at* τ *in* T *is equal to* $R^{(0,\tau)}_{\prod_{t\in T}\Xi_t}$ (δ_{ω_0}).

The following lemma will be used in Section 3.5 (Theorem 3.15).

LEMMA 3.11. (Inference and S-state). *Let* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0)$, $\{B^{(0,\tau)}_{\prod_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ *and* $\{R^{(0,\tau)}_{\prod_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ *be as in Definition 3.7. Assume that*

($\sharp$) *we know that the measured value* $(x_t)_{t\in T}$ $(\in \prod_{t\in T} X_t)$ *obtained by* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$ *belongs to* $\prod_{t\in T}\Xi_t$.

Then, there is a reason to infer that

($\flat$) *the S-state* $\big($*after the measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\big)$ *at* τ *in* T *is equal to* $R^{(0,\tau)}_{\prod_{t\in T}\Xi_t}(\delta_{\omega_0})$.

Here the δ_{ω_0} $(\in \mathcal{M}^p_{+1}(\Omega_0))$ *is defined by*

$$[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0) = \max_{\omega\in\Omega_0}[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega). \tag{3.35}$$

PROOF. The proof is similar to that of Corollary 3.3. Let $(Y_\tau, 2^{Y_\tau}, G_\tau)$ be any observable in $C(\Omega_\tau)$. Note that the above ($\sharp$) is the same as the following statement:

(♯)' we know the measured value $((x_t)_{t\in T}, y)$ $(\in (\prod_{t\in T} X_t)\times Y_\tau)$ obtained by $\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}, S_{[*]})$ belongs to $(\prod_{t\in T} \Xi_t) \times Y_\tau$ (where $\widehat{\mathbf{O}}_0$ is as in Definition 3.7).

Thus we can infer, by Theorem 3.2 (Fisher's method) and the equality $\widetilde{F}_0(\prod_{t\in T} \Xi_t) = \widehat{F}_0((\prod_{t\in T} \Xi_t)\times Y_\tau)$, that the unknown state $[*]$ $\Big($in $\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}, S_{[*]})\Big)$ is equal to δ_{ω_0} $\Big($defined by (3.35)$\Big)$. Thus the conditional probability $P_{\Pi_{t\in T}\Xi_t}(\cdot)$ under the condition that we know $((x_t)_{t\in T}, y) \in (\prod_{t\in T} \Xi_t) \times Y_\tau$ is given by

$$
\begin{aligned}
P_{\Pi_{t\in T}\Xi_t}(\Gamma_\tau) &= \frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widehat{F}_0((\prod_{t\in T} \Xi_t) \times \Gamma_\tau)\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widehat{F}_0((\prod_{t\in T} \Xi_t) \times Y_\tau)\rangle_{C(\Omega_0)}} = \frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau))\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)\rangle_{C(\Omega_0)}} \\
&= {}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}), G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)} \qquad (\forall \Gamma_\tau \in 2^{Y_\tau}) \qquad \text{(cf. (3.30))}.
\end{aligned}
$$

From the equivalence of (C_4) and (C_5), we can conclude (♭). □

Now we consider the simplest case such that $T \equiv \{0, \tau\}$ and $\mathbf{S}_{[\delta_{\omega_0}]} \equiv [S_{[\delta_{\omega_0}]}; C(\Omega_\tau) \overset{\Phi_{0,\tau}}{\to} C(\Omega_0)]$. For each $k = 0, \tau$, consider the null observable $\mathbf{O}^{(\mathrm{nl})}_k \equiv (\{0,1\}, 2^{\{0,1\}}, F^{(\mathrm{nl})}_k)$ in $C(\Omega_k)$ (cf. Example 2.3 (ii)). Then, we have the measurement

$$
\mathbf{M}_{C(\Omega_0)}\Big(\widetilde{\mathbf{O}}_0 \equiv (\{0,1\}^2, 2^{\{0,1\}^2}, F^{(\mathrm{nl})}_0 \times \Phi_{0,\tau} F^{(\mathrm{nl})}_\tau), S_{[\delta_{\omega_0}]}\Big). \tag{3.36}
$$

Note that

(i) the probability that the measured value (by $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$) is equal to $(1,1)$ is given by 1. That is, the measured value is always (or surely) equal to $(1,1)$.

Thus,

(ii) the measured value obtained by $\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$ has always the form $((1,1), y)$ $(\in \{0,1\}^2 \times Y_\tau)$. Here $\widehat{\mathbf{O}}_0$ is defined by

$$
(\{0,1\}^2 \times Y_\tau, 2^{\{0,1\}^2\times Y_\tau}, F^{(\mathrm{nl})}_0 \times \Phi_{0,\tau} F^{(\mathrm{nl})}_\tau \times \Phi_{0,\tau} G_\tau) \tag{3.37}
$$

for any any observable $(Y_\tau, 2^{Y_\tau}, G_\tau)$ in $C(\Omega_\tau)$.

Note that $\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$ and $\mathbf{M}_{C(\Omega_0)}((Y_\tau, 2^{Y_\tau}, \Phi_{0,\tau} G_\tau), S_{[\delta_{\omega_0}]})$ are essentially the same. That is because "to take $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[\delta_{\omega 0}]})$" is essentially the same as "to take no measurement" (cf. Example 2.3 (ii)). Thus, the above (ii) implies that

(iii) the probability distribution of unknown y $\Big($under (ii) $(= \text{(i)})\Big)$, i.e., the probability that y $(\in Y_\tau)$ belongs to Γ_τ, is represented by

$$
{}_{C(\Omega_\tau)^*}\langle \Phi^*_{0,\tau}(\delta_{\omega_0}), G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)}
$$

for any observable $(Y_\tau, 2^{Y_\tau}, G_\tau)$ in $C(\Omega_\tau)$ and any Γ_τ $(\in 2^{Y_\tau})$.

That is because it holds that

$$
\frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, (F_0^{(\mathrm{nl})} \times \Phi_{0,\tau}F_\tau^{(\mathrm{nl})} \times \Phi_{0,\tau}G_\tau)(\{(1,1)\} \times \Gamma_\tau)\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, (F_0^{(\mathrm{nl})} \times \Phi_{0,\tau}F_\tau^{(\mathrm{nl})} \times \Phi_{0,\tau}G_\tau)(\{(1,1)\} \times Y_\tau)\rangle_{C(\Omega_0)}}
$$
$$
= {}_{C(\Omega_\tau)^*}\langle \Phi^*_{0,\tau}(\delta_{\omega_0}), G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)}.
$$

Thus,we get the following (iv), which is short for (iii).

(iv) the S-state at τ $(\in T \equiv \{0,\tau\})$ is equal to $\Phi^*_{0,\tau}(\delta_{\omega_0})$.

Thus we conclude that (i) $\Rightarrow$ (iv). However, note that (i) always holds. Therefore, we consider that (iv) always holds.

From the above argument, we have the following lemma. This will be used in the statement (3.41) later.

Lemma 3.12. (The Schrödinger picture representation). *Put $T = \{0,\tau\}$. Let $\mathbf{S}_{[\delta_{\omega_0}]} \equiv [S_{[\delta_{\omega_0}]}; \{C(\Omega_\tau) \overset{\Phi_{0,\tau}}{\to} C(\Omega_0)\}]$ be a general system with an initial state $S_{[\delta_{\omega_0}]}$. Then we see that*

($\sharp$) *the S-state at τ $(\in T \equiv \{0,\tau\}$ $)$ is $\Phi^*_{0,\tau}(\delta_{\omega_0})$.*

Here it should be noted that the measurement $\mathbf{M}_{C(\Omega_0)}((Y_\tau, 2^{Y_\tau}, \Phi_{0,\tau}G_\tau), S_{[\delta_{\omega_0}]})$ $\Big($or, $\mathbf{M}_{C(\Omega_0)}$ $(\widehat{\mathbf{O}}_0, S_{[\delta_{\omega_0}]})$, cf. (3.37)$\Big)$ is hidden behind the assertion ($\sharp$).

Also, the following lemma is the formal representation of Corollary 3.4 (ii). $\Big($*Cf.* Remark 3.14.$\Big)$

Lemma 3.13. (Inference and the Schrödinger picture representation). *Put $T = \{0,\tau\}$. Let $\mathbf{S}_{[*]} \equiv [S_{[*]}; \{\Phi_{0,\tau} : C(\Omega_\tau) \to C(\Omega_0)\}]$ be a general system with an initial state $S_{[*]}$. Let $\mathbf{O}_0 = (X_0, 2^{X_0}, F_0)$ be an observable in $C(\Omega_0)$. And, let $\mathbf{O}_\tau^{(\mathrm{nl})} = (\{0,1\}, 2^{\{0,1\}}, F_\tau^{(\mathrm{nl})})$ be the null observable in $C(\Omega_\tau)$ (cf. Example 2.3). Consider a measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0(\equiv \mathbf{O}_0 \times \Phi_{0,\tau}\mathbf{O}_\tau^{(\mathrm{nl})}), S_{[*]})$, which is essentially the same as $\mathbf{M}_{C(\Omega_0)}(\mathbf{O}_0, S_{[*]})$. Assume that*

($\sharp$) *we know that the measured value obtained by* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv \mathbf{O}_0 \times \Phi_{0,\tau}\mathbf{O}_\tau^{(\mathrm{nl})}, S_{[*]})$ *belongs to* $\Xi_0 \times \{1\}$ $(\in 2^{X_0 \times \{0,1\}})$.

Then we see that

($\flat$) *there is a reason to infer that the S-state (after the measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$*) at* τ $(\in T \equiv \{0, \tau\})$ *is* $\Phi_{0,\tau}^*(\delta_{\omega_0})$,

where δ_{ω_0} $(\in \mathcal{M}_{+1}^p(\Omega_0))$ *is defined by*

$$[F_0(\Xi_0)](\omega_0) = \max_{\omega \in \Omega_0}[F_0(\Xi_0)](\omega). \tag{3.38}$$

Proof. Let $B_{\Xi_0 \times \{1\}}^{(0,\tau)} : C(\Omega_\tau) \to C(\Omega_0)$ and $R_{\Xi_0 \times \{1\}}^{(0,\tau)} : \mathcal{M}_{+1}^m(\Omega_0) \to \mathcal{M}_{+1}^m(\Omega_\tau)$ be as in Definition 3.7. Here, note that, from the property of null observable, it holds that $F_0(\Xi_0) \times \Phi_{0,\tau}F_\tau^{(\mathrm{nl})}(\{1\}) = F_0(\Xi_0)$. Thus we see that $B_{\Xi_0 \times \{1\}}^{(0,\tau)}(g_\tau) = F_0(\Xi_0) \times \Phi_{0,\tau}g_\tau$ for any g_τ $(\in C(\Omega_\tau))$. By Lemma 3.11, it suffices to prove $R_{\Xi_0}^{(0,\tau)}(\delta_{\omega_0}) = \Phi_{0,\tau}^*(\delta_{\omega_0})$. This is shown as follows.

$$
\begin{aligned}
&{}_{C(\Omega_\tau)^*}\langle R_{\Xi_0 \times \{1\}}^{(0,\tau)}(\delta_{\omega_0}), g_\tau \rangle_{C(\Omega_\tau)} = {}_{C(\Omega_\tau)^*}\langle \frac{(B_{\Xi_0 \times \{1\}}^{(0,\tau)})^*(\delta_{\omega_0})}{\|(B_{\Xi_0 \times \{1\}}^{(0,\tau)})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}}, g_\tau \rangle_{C(\Omega_\tau)} \\
=&\frac{1}{\|(B_{\Xi_0 \times \{1\}}^{(0,\tau)})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}} {}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B_{\Xi_0 \times \{1\}}^{(0,\tau)}(g_\tau) \rangle_{C(\Omega_0)} = \frac{[F_0(\Xi_0)](\omega_0) \times [\Phi_{0,\tau}g_\tau](\omega_0)}{[F_0(\Xi_0)](\omega_0)} \\
=&{}_{C(\Omega_\tau)^*}\langle \Phi_{0,\tau}^*(\delta_{\omega_0}), g_\tau \rangle_{C(\Omega_\tau)} \qquad (\forall g_\tau \in C(\Omega_\tau)).
\end{aligned} \tag{3.39}
$$

Then, we see that $R_{\Xi_0 \times \{1\}}^{(0,\tau)}(\delta_{\omega_0}) = \Phi_{0,\tau}^*(\delta_{\omega_0})$. This completes the proof. □

The following remark shows that Corollary 3.4 (ii) is a direct consequence of Lemma 3.13.

Remark 3.14. (Continued from Corollary 3.4). As mentioned before, the proof of Corollary 3.4 is temporary. Corollary 3.4 should be understood as a corollary of Lemma 3.13 as follows. In Lemma 3.13, put $\Omega_0 = \Omega_\tau = \Omega_{+0}$. And let $\Phi_{0,\tau} : C(\Omega_{+0}) \to C(\Omega_0)$ be the identity map. Since "the S-state (after the measurement $\mathbf{M}_{C(\Omega_0)}(\mathbf{O}_0, S_{[*]})$) at $\tau(= +0)$" $= \Phi_{0,\tau}(\delta_{\omega_0}) = \delta_{\omega_0}$, we easily see that Corollary 3.4 is a consequence of Lemma 3.13. This should be regarded as the formal proof of Corollary 3.4.

3.5 Regression analysis in measurements

Now let us explain the reason why we consider:

(♯) it is worth while to doubting the derivation of (3.5) $\big(= (3.29)\big)$ from (3.4) $\big(= (3.28)\big)$, i.e., the formula $h(2) = 0.4 + 1.4 \times 2 = 3.2$.

Using the notations in Regression Analysis I, as the statement (R) of Example 3.6, we say that

- the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t,\ 2^{\prod_{t\in T} X_t},\ \widetilde{F}_0), S_{[*]})$ is hidden behind the inference (3.4) $\big(=(3.28)\big)$.

And we conclude, by Corollary 3.4 (or Remark 3.14), that

$$
\begin{aligned}
[*] = &\text{ "the S-state after the measurement } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{"} \\
= &\ \delta_{\omega_0}.
\end{aligned} \tag{3.40}
$$

Here the δ_{ω_0} $(\in \mathcal{M}_{+1}^p(\Omega_0))$ is defined by $[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega_0) = \max_{\omega\in\Omega_0}[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega)$. On the other hand,

- the map "$\delta_{\omega_0} \mapsto \Phi_{0,\tau}^*(\delta_{\omega_0})$" $\Big($i.e., the derivation of (3.5) $\big(= (3.29)\big)$ from (3.4) $\big(= (3.28)\big)\Big)$ is due to the Schrödinger picture, behind which the measurement $\mathbf{M}_{C(\Omega_0)}(\Phi_{0,\tau}\mathbf{O}'_\tau \equiv (Y_\tau, 2^{Y_\tau}, \Phi_{0,\tau}G_\tau),\ S_{[\delta_{\omega_0}]})$ is hidden. *Cf.* Lemma 3.12. (3.41)

Thus, in order to conclude the assertion (3.5) $\big(= (3.29)\big)$, we need the above "two measurements", that is,

$$
\text{"}\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{" and "}\mathbf{M}_{C(\Omega_0)}(\Phi_{0,\tau}\mathbf{O}'_\tau, S_{[\delta_{\omega_0}]})\text{"}. \tag{3.42}
$$

However, note that it is forbidden to conduct "two measurements" (cf. Remark 2.4). This is the reason that we consider that it is worth while to doubting (3.5) $\big(= (3.29)\big)$. In order to avoid this confusion, it suffices to consider the "simultaneous" measurement:

$$
\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}_0 \equiv ((\prod_{t\in T} X_t) \times Y_\tau, 2^{(\prod_{t\in T} X_t)\times Y_\tau}, \widehat{F}_0), S_{[*]}), \ \ \text{(where } \widehat{\mathbf{O}}_0 \text{ is as in Definition 3.7)}, \tag{3.43}
$$

instead of (3.42).

Then, we rewrite Lemma 3.11 as our main theorem as follows.

THEOREM 3.15. (= Lemma 3.11, Inference in Markov relation). *Let* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t \in T} X_t, 2^{\prod_{t \in T} X_t}, \widetilde{F}_0)$ *be as in Theorem 2.9 in the case* $\mathcal{A}_t = C(\Omega_t)$ $(\forall t \in T)$*. And consider a measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$*. Let* τ *be any element in* T*. Let* $\{R^{(0,\tau)}_{\prod_{t \in T} \Xi_t} \mid \Xi_t \in 2^{X_t} \ (\forall t \in T)\}$ *be as in Definition 3.7. Assume that we know that the measured value (obtained by* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$*) belongs to* $\prod_{t \in T} \Xi_t$*. Then, there is a reason to infer that*

$$(\sharp) \qquad \text{“the } S\text{-state at } \tau \ (\in T) \text{ after } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = R^{(0,\tau)}_{\prod_{t \in T} \Xi_t}(\delta_{\omega_0}). \tag{3.44}$$

Here δ_{ω_0} $(\in \mathcal{M}^p_{+1}(\Omega))$ *is defined by*

$$[\widetilde{F}_0(\prod_{t \in T} \Xi_t)](\omega_0) = \max_{\omega \in \Omega_0} [\widetilde{F}_0(\prod_{t \in T} \Xi_t)](\omega). \tag{3.45}$$

Lastly, we prove the following lemma, which justify the inference (3.5).

LEMMA 3.16. (Some property of homomorphic relation). *Let* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t \in T} X_t, 2^{\prod_{t \in T} X_t}, \widetilde{F}_0)$ *be as in Theorem 2.9 in the case* $\mathcal{A}_t = C(\Omega_t)$ $(\forall t \in T)$*. Consider the family of Bayes operators* $\{B^{(0,\tau)}_{\prod_{t \in T} \Xi_t} \mid \Xi_t \in 2^{X_t} \ (t \in T)\}$ *such as in Definition 3.7. Let* τ *be any element in* T*. Assume that* $\Phi_{\pi(t),t} : C(\Omega_t) \to C(\Omega_{\pi(t)})$ $(\forall t \in T$ *such that* $0 < t \le \tau)$ *is homomorphic. Then, it holds that:*

$$B^{(0,\tau)}_{\prod_{t \in T} \Xi_t}(G_\tau(\Gamma_\tau)) = \widetilde{F}_0(\prod_{t \in T} \Xi_t) \times \Phi_{0,\tau} G_\tau(\Gamma_\tau) \qquad (\forall \Xi_t \in 2^{X_t} \ (\forall t \in T), \forall \Gamma_\tau \in 2^{Y_\tau}), \tag{3.46}$$

for any observable $(Y_\tau, 2^{Y_\tau}, G_\tau)$ *in* $C(\Omega_\tau)$*. Thus we see that the Bayes operator* $B^{(0,\tau)}_{\prod_{t \in T} \Xi_t} : C(\Omega_\tau) \to C(\Omega_0)$ *is determined uniquely under the homomorphic condition.*

PROOF. The proof is shown in the following three steps.
[Step 1]. Let ω_0 be any element in Ω_0. And let g_τ and h_τ be in $C(\Omega_\tau)$ such that

$$0 \le g_\tau \le 1, \ g_\tau(\phi_{0,\tau}(\omega_0)) = 0, \ 0 \le h_\tau \le 1, \text{ and } h_\tau(\phi_{0,\tau}(\omega_0)) = 1. \tag{3.47}$$

where $\phi_{0,\tau} : \Omega_0 \to \Omega_\tau$ is defined by (2.18). Then we see, by (3.32), that

$$0 \le [B^{(0,\tau)}_{\prod_{t \in T} \Xi_t}(g_\tau)](\omega) \le (\Phi_{0,\tau} g_\tau)(\omega) = g_\tau(\phi_{0,\tau}(\omega)) \qquad (\forall \omega \in \Omega_0). \tag{3.48}$$

Putting $\omega = \omega_0$ in (3.48), we get, by (3.47), that

$$[B^{(0,\tau)}_{\prod_{t \in T} \Xi_t}(g_\tau)](\omega_0) = 0. \tag{3.49}$$

Also, from the linearity of Bayes operator and the condition (iii) of Definition 3.7, we get

$$[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau - h_\tau)](\omega) = [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)](\omega) - [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(h_\tau)](\omega)$$
$$= [\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega) - [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(h_\tau)](\omega) \qquad (\forall\omega\in\Omega_0). \tag{3.50}$$

Thus, putting $\omega = \omega_0$ in (3.50), we get, by (3.47), that

$$0 \le [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau - h_\tau)](\omega_0)$$
$$\le [(\Phi_{0,\tau}(1_\tau - h_\tau))](\omega_0) = 1_\tau(\phi_{0,\tau}(\omega_0)) - h_\tau(\phi_{0,\tau}(\omega_0)) = 1 - 1 = 0. \tag{3.51}$$

Then, we obtain

$$[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(h_\tau)](\omega_0) = [\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0). \tag{3.52}$$

[Step 2]. Let ω_0 be any element in Ω_0. Fix any f $(\in C(\Omega_\tau))$ such that $0 \le f \le 1$. Define g_τ, h_τ $(\in C(\Omega_\tau))$ such that

$$g_\tau(\omega_\tau) = \max\{0, f(\omega_\tau) - f(\phi_{0,\tau}(\omega_0))\} \qquad (\forall\omega_\tau\in\Omega_\tau),$$
$$h_\tau(\omega_\tau) = \min\{\frac{f(\omega_\tau)}{f(\phi_{0,\tau}(\omega_0))}, 1\} \qquad (\forall\omega_\tau\in\Omega_\tau). \tag{3.53}$$

The g_τ and the h_τ clearly satisfy (3.47). And moreover, we see, for any $\omega_\tau \in \Omega_\tau$, that

$$\begin{aligned} & g_\tau(\omega_\tau) + f(\phi_{0,\tau}(\omega_0))h_\tau(\omega_\tau) \\ &= \max\{0, f(\omega_\tau) - f(\phi_{0,\tau}(\omega_0))\} + \min\{f(\omega_\tau), f(\phi_{0,\tau}(\omega_0))\} \\ &= \begin{cases} (f(\omega_\tau) - f(\phi_{0,\tau}(\omega_0))) + f(\phi_{0,\tau}(\omega_0)), & \text{if } f(\omega_\tau) \ge f(\phi_{0,\tau}(\omega_0)) \\ 0 + f(\omega_\tau), & \text{if } f(\omega_\tau) \le f(\phi_{0,\tau}(\omega_0)) \end{cases} \\ &= f(\omega_\tau). \end{aligned} \tag{3.54}$$

[Step 3]. Let ω_0 be any element in Ω_0. Let Γ_τ be any element in 2^{Y_τ}. From the [Step 2], we see that there exist $\widehat{g}_\tau$ $(\in C(\Omega_\tau))$ and $\widehat{h}_\tau$ $(\in C(\Omega_\tau))$ such that $G_\tau(\Gamma_\tau) = \widehat{g}_\tau + [G_\tau(\Gamma_\tau)](\phi_{0,\tau}(\omega_0))\widehat{h}_\tau$, $\widehat{g}_\tau(\phi_{0,\tau}(\omega_0)) = 0$, $\widehat{h}_\tau(\phi_{0,\tau}(\omega_0)) = 1$. Then we see

$$[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau))](\omega) = \Big[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}\Big(\widehat{g}_\tau + [G_\tau(\Gamma_\tau)](\phi_{0,\tau}(\omega_0))\widehat{h}_\tau\Big)\Big](\omega)$$
$$= [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{g}_\tau)](\omega) + [G_\tau(\Gamma_\tau)](\phi_{0,\tau}(\omega_0)) \times [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{h}_\tau)](\omega) \qquad (\forall\omega\in\Omega_0). \tag{3.55}$$

Putting $\omega = \omega_0$, we see, by (3.49) and (3.42), that $[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{g}_\tau)](\omega_0) = 0$ and $[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{h}_\tau)](\omega_0)$ $= [\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0)$. Thus, we see, by (3.55), that

$$[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau))](\omega_0) = [G_\tau(\Gamma_\tau)](\phi_{0,\tau}(\omega_0)) \times [\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0)$$
$$= [\Phi_{0,\tau}G_\tau(\Gamma_\tau)](\omega_0) \times [\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0). \tag{3.56}$$

Since ω_0 $(\in \Omega_0)$ is arbitrary, we obtain (3.46). This completes the proof. □

Now we can propose our main assertion as follows.

REGRESSION ANALYSIS II. (The new proposal). *Let* $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$ *be a tree with root* 0, *and let* $\mathbf{S}_{[*]} \equiv [S_{[*]}; \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T\setminus\{0\}}]$ *be a general system with the initial system* $S_{[*]}$. *And, let an observable* $\mathbf{O}_t \equiv (X_t, 2^{X_t}, F_t)$ *in a* C^**-algebra* $C(\Omega_t)$ *be given for each* $t \in T$. *Then, we have a measurement*

$$\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\Pi_{t\in T} X_t}, \widetilde{F}_0), S_{[*]}) \qquad (\textit{cf. Theorem 2.9}). \tag{3.57}$$

Assume that the measured value by the measurement $\mathbf{M}_{C(\Omega)}(\widetilde{\mathbf{O}}_0, S_{[*]})$ *belongs to* $\prod_{t\in T}\Xi_t$ $(\in 2^{\Pi_{t\in T} X_t})$. *Also define* δ_{ω_0} $(\in \mathcal{M}^p_{+1}(\Omega_0))$ *such that*

$$[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0) = \max_{\omega\in\Omega_0}[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega). \tag{3.58}$$

Let τ *be any element in* T. *Let* $\{R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ *be as in Definition 3.7.* $\Big($*The existence of* $\{R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t \in T)\}$ *is assumed by Theorem 3.8.*$\Big)$ *Then, we see:*

(i). *[The S-state at* τ $(\in T)$*]. There is a reason to infer that*

$$(\sharp) \qquad \text{“The S-state at } \tau\ (\in T) \text{ after } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}). \tag{3.59}$$

Also

(ii). [*The S-state at* τ $(\in T)$ *for the homomorphism* $\Phi_{0,\tau}$]. *Assume that* $\Phi_{0,\tau} : C(\Omega_\tau) \to C(\Omega_0)$ *is homomorphic* $\Big($*i.e.,* $\Phi_{\pi(t),t} : C(\Omega_t) \to C(\Omega_{\pi(t)})$ $(\forall t \in T$ *such that* $0 < t \leq \tau)$ *is homomorphic*$\Big)$. *Then there is a reason to infer that*

$$\text{“the S-state at } \tau\ (\in T) \text{ after } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = \Phi^*_{0,\tau}(\delta_{\omega_0}). \tag{3.60}$$

Here note that $\Phi^*_{0,\tau}(\delta_{\omega_0}) = \delta_{\phi_{0,\tau}(\omega_0)}$ *where* $\phi_{0,\tau} : \Omega_0 \to \Omega_\tau$ *is defined by (2.18).*

PROOF. (i). See Theorem 3.15 (= Lemma 3.11).
(ii). We see, by Lemma 3.16, that

$$
\begin{aligned}
& {}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}), g_\tau\rangle_{C(\Omega_\tau)} = {}_{C(\Omega_\tau)^*}\langle \frac{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})}{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})}, g_\tau\rangle_{C(\Omega_\tau)} \\
= & \frac{1}{\|(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}} {}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(g_\tau)\rangle_{C(\Omega_0)} \\
= & \frac{1}{[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0)} {}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widetilde{F}_0(\prod_{t\in T}\Xi_t) \times \Phi_{0,\tau} g_\tau\rangle_{C(\Omega_0)} \qquad \text{(by Lemma 3.10)} \\
= & {}_{C(\Omega_\tau)^*}\langle \Phi^*_{0,\tau}(\delta_{\omega_0}), g_\tau\rangle_{C(\Omega_\tau)} \qquad (\forall g_\tau \in C(\Omega_\tau)).
\end{aligned}
$$

Then, we see that $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}) = \Phi^*_{0,\tau}(\delta_{\omega_0})$. □

REMARK 3.17. (1. Continued from Examples 3.1 and 3.6). Note that our problem (i) in Example 3.1 was to infer the $h(2)$ and not (α_0, β_0). Regression analysis II (3.60) is applicable to our problem, that is, the above (3.60) says that there is a reason to calculate $h(2)$ in the following:

$$
h(2) = \phi_{0,2}(0.4, 1.4) = 0.4 + 1.4 \times 2 = 3.2. \tag{3.61}
$$

(2. Interesting logic). It should be noted that, when $\tau = 0$, the Regression Analysis II is the same as the Regression Analysis I. Thus, we also conclude (3.4), i.e., $(\alpha_0, \beta_0) = (0.4, 1.4)$. After all, the Regression Analysis II says that

(M_1) as the result in the case that $\tau = 0$, the conclusion (3.4) in Example 3.1 is reasonable,

or

(M_2) as the result in the case that $\tau \neq 0$, the conclusion (3.5) in Example 3.1 is reasonable,

However, it should be noted that the Regression Analysis II does not guarantee that

(M_3) both (3.4) and (3.5) in Example 3.1 are (simultaneously) reasonable.

That is because two measurements (i.e., the measurement $\mathbf{M}_1$ behind (M_1) and the measurement $\mathbf{M}_2$ behind (M_2)) are included in (M_1) and (M_2). If we want to conclude this (M_3), we must consider the simultaneous measurement of "measurement $\mathbf{M}_1$" and "measurement $\mathbf{M}_2$", that is, we must generalize Definition 3.7 (Bayes operator) such as $B^{(0,(0,\tau))}_{\Pi_{t\in T}\Xi_t} : C(\Omega_0) \times C(\Omega_\tau) \to C(\Omega_0)$ satisfying similar conditions since only one measurement is permitted (cf. Remark 2.4). This is, of course, interesting, though it is not discussed in this paper.

3.6 Conclusion

It is too optimistic to claim that an axiomatic approach is always possible and powerful to every field in science. We can easily check it if we, for example, examine the standard description of chemistry, psycohlogy or botany, etc.. That is, we consider that the success of the axiomatic approach to physics (e.g., "Newtonian mechanics is based on the three laws" etc.) is a quite rare case in the history of science. And further, we believe that only the most fundamental theories can be completely formulated under some axioms. Here, we have the important question *"Which category does statistics belong to?".* This is precisely our motivation of this paper. For the axiomatic approach to statistics, we start from the two axioms in measurement theory:

$$\underset{\text{(measurement theory)}}{\text{MT}} = \underset{\text{(Axiom 1)}}{\text{measurement}} + \underset{\text{(Axiom 2)}}{\text{the relation among systems}} \tag{3.62}$$

which includes classical and quantum measurements.

In this paper we show that regression analysis can be completely understood in MT as follows.

$$\text{measurement theory} \Longrightarrow \begin{cases} \text{Axiom 1} \Rightarrow \underset{\text{(Fisher's method)}}{\text{Theorem 3.2}} \Rightarrow \begin{cases} \text{Corollary 3.3 (conditional probability)} \\ \text{Corollary 3.4 (classical Fisher's method)} \end{cases} \\ \\ \text{Axiom 2} \Rightarrow \begin{cases} \text{Theorem 2.9 (measurability)} \\ \text{Theorem 3.8 (the existence of Bayes operator)} \\ \text{Lemma 3.10 (some property of homomorphic relation).} \end{cases} \end{cases}$$

And, using these results, we derive "regression analysis" as follows.

(i) : "Theorem 2.9" + "Corollary 3.4" ⇒ "Regression Analysis I ",

$$\text{(ii)}: \quad \text{"Corollary 3.3"} + \text{"Theorem 3.8"} \Rightarrow \left.\begin{matrix} \text{"Theorem 2.9"} \\ \underset{\text{(Markov inference)}}{\text{"Theorem 3.15"}} \\ \text{"Lemma 3.10"} \end{matrix}\right\} \Rightarrow \text{"Regression Analysis II".}$$

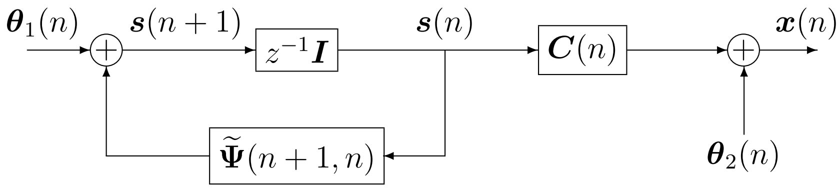
We believe that Regression Analysis II is the best (i.e., precise, wide, deep etc.) in all conventional proposals of regression analysis (though it should be generalized as mentioned in Remark 3,17.). It is surprising that both statistics and quantum mechanics can be understood in the same theory, i.e., measurement theory (3.62).

We believe that every statistician wants to know the complete justification of (3.4) and (3.5) in Example 3.1. Thus we expect that many statisticians will be interested in our axiomatic approach. That is because there is no justification without axioms.

We hope that our theory will be generally accepted as a standard theory of an axiomatic formulation of statistics.

4 Kalman filter in Noise

In this section we formulate "Kalman filter" (cf. Kalman (1960)) in SMT. Consider the following conventional "Kalman" filter as follows.



(4.1)

where $\boldsymbol{s}(n)$: L-dimensional state vector at time n, $\boldsymbol{x}(n)$: M-dimensional measured data vector. And $\boldsymbol{s}(n)$ and $\boldsymbol{x}(n)$ are described by the following equations:

$$\begin{cases} \boldsymbol{s}(n+1) = \widetilde{\boldsymbol{\Psi}}(n+1,n)\boldsymbol{s}(n) + \boldsymbol{\theta}_1(n) & \text{: stochastic difference state equation} \\ & \qquad (n = 0, 1, \cdots, N-1) \\ \boldsymbol{x}(n) = \boldsymbol{C}(n)\boldsymbol{s}(n) + \boldsymbol{\theta}_2(n) & \text{: measurement equation} \end{cases} \tag{4.2}$$

Here, it is assumed that $\widetilde{\boldsymbol{\Psi}}(n+1, n)$, $\boldsymbol{C}(n)$, $\boldsymbol{\theta}_1(n)$ (and its initial distribution) and $\boldsymbol{\theta}_2(n)$ are known, where $\widetilde{\boldsymbol{\Psi}}(n+1, n)$: $K \times K$-dimensional transition matrix, $\boldsymbol{\theta}_1(n)$: L-dimensional input vector which represents a white noise, $\boldsymbol{C}(n)$: $L \times K$-dimensional measurement matrix, $\boldsymbol{\theta}_2(n)$: L-dimensional vector which represents a measurement error. Here, our problem is as follows.

($\flat$) Infer the state vector $\overline{\boldsymbol{s}}(\tau)$ $(0 \le \tau \le N)$ from the measured data $\boldsymbol{x}(1), \boldsymbol{x}(2), \cdots, \boldsymbol{x}(N)$

Also, note the original equation of the stochastic difference equation (4.2) is the following equation:

$$\overline{\boldsymbol{s}}(n+1) = \widetilde{\boldsymbol{\Psi}}(n+1,n)\overline{\boldsymbol{s}}(n) \quad (n = 0, 1, \cdots, N-1). \tag{4.3}$$

In this section, we consider this problem ($\flat$) in SMT.

The following theorem is an analogy of Theorem 3.15. This theorem (= Theorem 4.1) is also called “Bayes’ method“.

THEOREM 4.1. (Generalized Bayes theorem, Bayes’ method or Bayes-Kalman filter). *Let* $(T \equiv \{0, 1, ..., N\}, \pi : T \setminus \{0\} \to T)$ *be a tree with the root* 0 *and let* $\mathbf{S}(\nu_0) \equiv [S(\nu_0), C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)}) \ (t \in T \setminus \{0\})]$ *be a general system with the initial system* $S(\nu_0)$*. And, let an observable* $\mathbf{O}_t \equiv (X_t, \mathcal{F}_t, F_t)$ *in a* C^**-algebra* $C(\Omega_t)$ *be given for each* $t \in T$*. Then, we have a statistical measurement*

$$\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t \in T} X_t, \bigotimes_{t \in T} \mathcal{F}_t, \widetilde{F}_0), S(\nu_0)). \qquad \textit{(cf. Theorem 3.15)}$$

Assume that the measured value by the statistical measurement $\mathbf{M}_{C(\Omega)}(\widetilde{\mathbf{O}}_0, S(\nu_0))$ *belongs to* $\prod_{t \in T} \Xi_t$ $(\in \bigotimes_{t \in T} \mathcal{F}_t)$*. Let* τ *be any element in* T*. Then, we see*

(*a*) *“the (statistical) S-state at* $\tau (\in T)$ *after* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S(\nu_0))$*”* $= R^{(0,\tau)}_{\prod_{t \in T} \Xi_t}(\nu_0)$.

Proof. Since the sequential observable $[\{\mathbf{O}_t\}_{t \in T}, \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t \in T \setminus \{0\}}]$ is common to PMT and SMT, Theorem 3.15 is applicable. □

4.1 The measurement theoretical formulation of the (4.2)

Firstly, we formulate the (4.2) in SMT. Assume, for simplicity, that $T (\equiv \{0, 1, ..., N\})$ is a tree with a series structure (though this assumption is not needed). For each $t \ (\in T)$, consider compact Hausdorff spaces $\mathcal{S}_t$ and Θ_t. Note that the $\mathcal{S}_t$ [resp. Θ_t] $(t \ (\in T))$ is the state space [resp. “the set of the white noise”]. Although, it is natural to assume that $\mathcal{S}_0 = \mathcal{S}_1 = \cdots = \mathcal{S}_N$ and $\Theta_0 = \Theta_1 = \cdots = \Theta_N$. In this papaer, we can do well without this assumption. Now, consider the following two Markov relations among systems: $[\{\Psi_{t_1,t_2} : C(\mathcal{S}_{t_2}) \to C(\mathcal{S}_{t_1})\}_{(t_1,t_2) \in T^2_{\leq}}]$ and $[\{\Upsilon_{t_1,t_2} : C(\Theta_{t_2}) \to C(\Theta_{t_1})\}_{(t_1,t_2) \in T^2_{\leq}}]$ such as

$$[C(\mathcal{S}_0)] \xleftarrow{\Psi_{0,1}} [C(\mathcal{S}_1)] \xleftarrow{\Psi_{1,2}} \cdots \xleftarrow{\Psi_{N-2,N-1}} [C(\mathcal{S}_{N-1})] \xleftarrow{\Psi_{N-1,N}} [C(\mathcal{S}_N)] \tag{4.4}$$

where the initial state δ_{s_0} $(\in \mathcal{M}^p_{+1}(\mathcal{S}_0))$ is assumed to be unknown, and

$$[C(\Theta_0)] \xleftarrow{\Upsilon_{0,1}} [C(\Theta_1)] \xleftarrow{\Upsilon_{1,2}} \cdots \xleftarrow{\Upsilon_{N-2,N-1}} [C(\Theta_{N-1})] \xleftarrow{\Upsilon_{N-1,N}} [C(\Theta_N)]$$
$$\big(\text{with the initial state } \nu_0^{\Theta} \ (\in \mathcal{M}^m_{+1}(\Theta_0))\big). \tag{4.5}$$

Here, it should be noted that the above (4.4) [resp. (4.5)] is the measurement theoretical formulation of (4.3) [resp. the $\boldsymbol{\theta}_1$ in (4.1)]. Also, note that the above (4.4) is equivalent to

$$[\mathcal{M}_{+1}^m(\mathcal{S}_0)] \xrightarrow{\Psi_{0,1}^*} [\mathcal{M}_{+1}^m(\mathcal{S}_1)] \xrightarrow{\Psi_{1,2}^*} \cdots \xrightarrow{\Psi_{N-2,N-1}^*} [\mathcal{M}_{+1}^m(\mathcal{S}_{N-1})] \xrightarrow{\Psi_{N-1,N}^*} [\mathcal{M}_{+1}^m(\mathcal{S}_N)]$$

where $\Psi_{n,n+1}^* : \mathcal{M}_{+1}^m(\mathcal{S}_n) \to \mathcal{M}_{+1}^m(\mathcal{S}_{n+1})]$ is the dual operator of $\Psi_{n,n+1} : C(\mathcal{S}_{n+1}) \to C(\mathcal{S}_n)$. Since the (4.4) corresponds to the conventional (4.3), it is natural to assume that the (4.4) is deterministic, i.e., $\Psi_{n,n+1}$ is homomorphic. Thus, for each $n = 0, 1, 2, \cdots, N-1$, there exists a continuous map $\psi_{n,n+1} : \mathcal{S}_n \to \mathcal{S}_{n+1}$, i.e.,

$$[\mathcal{S}_0] \xrightarrow{\psi_{0,1}} [\mathcal{S}_1] \xrightarrow{\psi_{1,2}} \cdots \xrightarrow{\psi_{N-2,N-1}} [\mathcal{S}_{N-1}] \xrightarrow{\psi_{N-1,N}} [\mathcal{S}_N]$$

where

$$f_{n+1}(\psi_{n,n+1}(s_n)) = [\Psi_{n,n+1}(f_{n+1})](s_n) \quad (\forall f_{n+1} \in C(\mathcal{S}_{n+1}), \forall s_n \in \mathcal{S}_n).$$

Next, consider a continuous map $\lambda_n : \mathcal{S}_n \times \Theta_n \to \mathcal{S}_n$, that is,

$$\mathcal{S}_n \times \Theta_n \ni (s_n, \theta_n) \mapsto \lambda_n(s_n, \theta_n) \in \mathcal{S}_n \quad (n = 0, 1, \cdots, N) \tag{4.6}$$

which corresponds to the left $\oplus$ in (4.1). The continuous map $\lambda_n : \mathcal{S}_n \times \Theta_n \to \mathcal{S}_n$ induces the continuous map $\Lambda_n : \mathcal{M}_{+1}^m(\mathcal{S}_n \times \Theta_n) \to \mathcal{M}_{+1}^m(\mathcal{S}_n)$ such that

$$\begin{aligned}\Lambda_n(\nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta})](B_n) = (\nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta})(\lambda_n^{-1}(B_n))\\ (\forall \nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta} \in \mathcal{M}_{+1}^m(\mathcal{S}_n \times \Theta_n), \forall B_n \subseteq \mathcal{S}_n : \text{open}).\end{aligned} \tag{4.7}$$

Further, define the continuous map $\widehat{\Phi}_{n,n+1}^* : \mathcal{M}_{+1}^m(\mathcal{S}_n \times \Theta_n) \to \mathcal{M}_{+1}^m(\mathcal{S}_{n+1} \times \Theta_{n+1})$, such that

$$\begin{aligned}\mathcal{M}_{+1}^m(\mathcal{S}_n \times \Theta_n) \ni \nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta} \mapsto & \widehat{\Phi}_{n,n+1}^*(\nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta})\\ & \equiv [\Lambda_{n+1}(\Psi_{n,n+1}^* \nu_n^{\mathcal{S}} \otimes \Upsilon_{n,n+1}^* \nu_n^{\Theta})] \otimes \Upsilon_{n,n+1}^* \nu_n^{\Theta} \in \mathcal{M}_{+1}^m(\mathcal{S}_{n+1} \times \Theta_{n+1})\end{aligned}$$

where $\Upsilon_{n,n+1}^* : \mathcal{M}_{+1}^m(\Theta_n) \to \mathcal{M}_{+1}^m(\Theta_{n+1})$ is a dual operator of $\Upsilon_{n,n+1} : C(\Omega_{n+1}) \to C(\Omega_n)$. That is,

$$\begin{aligned}& \nu_{n+1}^{\mathcal{S}} \otimes \nu_{n+1}^{\Theta} \big(\equiv \widehat{\Phi}_{n,n+1}^*(\nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta}) \big)\\ = & [\Lambda_{n+1}(\Psi_{n,n+1}^* \nu_n^{\mathcal{S}} \otimes \Upsilon_{n,n+1}^* \nu_n^{\Theta})] \otimes \Upsilon_{n,n+1}^* \nu_n^{\Theta} \quad (n = 0, 1, \cdots, N-1)\end{aligned} \tag{4.8}$$

which (or, the following (4.9)) corresponds to the state equation (4.2). Thus, we have the following Markov relation $[\{\widehat{\Phi}_{n,n+1} : C(\mathcal{S}_{n+1} \times \Theta_{n+1}) \to C(\mathcal{S}_n \times \Theta_n)\}_{n=0}^{N-1}]$:

$$[C(\mathcal{S}_0 \times \Theta_0)] \xleftarrow{\widehat{\Phi}_{0,1}} [C(\mathcal{S}_1 \times \Theta_1)] \xleftarrow{\widehat{\Phi}_{1,2}} \cdots \xleftarrow{\widehat{\Phi}_{N-2,N-1}} [C(\mathcal{S}_{N-1} \times \Theta_{N-1})] \xleftarrow{\widehat{\Phi}_{N-1,N}} [C(\mathcal{S}_N \times \Theta_N)] \tag{4.9}$$

where $\widehat{\Phi}_{n,n+1}$ is the pre-dual operator of $\widehat{\Phi}^*_{n,n+1}$ (i.e., $(\widehat{\Phi}_{n,n+1})^* = \widehat{\Phi}^*_{n,n+1}$).

Next, we consider the measurement theoretical characterization of the measurement equation (4.2). That is, consider the following Markov relation:

$$[C(\Theta'_0)] \xleftarrow{\Upsilon'_{0,1}} [C(\Theta'_1)] \xleftarrow{\Upsilon'_{1,2}} \cdots \xleftarrow{\Upsilon'_{N-2,N-1}} [C(\Theta'_{N-1})] \xleftarrow{\Upsilon'_{N-1,N}} [C(\Theta'_N)]$$
$$\big(\text{with the initial state } \nu_0^{\Theta'} (\in \mathcal{M}^m_{+1}(\Theta'_0))\big)$$

which corresponds to the $\boldsymbol{\theta_2}$ in (4.2). Also, for each n ($\in T$), consider an observable $\mathbf{O}_n = (X_n, 2^{X_n}, F_n)$ in $C(\mathcal{S}_n \times \Theta_n \times \Theta'_n)$, which corresponds to the measurement equation (4.2). Thus, we see that the (4.2) corresponds to the following (4.10):

$$\begin{matrix} [C(\mathcal{S}_0 \times \Theta_0 \times \Theta'_0)] & \xleftarrow{\widehat{\Phi}_{0,1}} & [C(\mathcal{S}_1 \times \Theta_1 \times \Theta'_1)] & \xleftarrow{\widehat{\Phi}_{1,2}} & \cdots & \xleftarrow{\widehat{\Phi}_{N-1,N}} & [C(\mathcal{S}_N \times \Theta_N \times \Theta'_N)] \\ (X_0, 2^{X_0}, F_0) & & (X_1, 2^{X_1}, F_1) & & \cdots & & (X_N, 2^{X_N}, F_N) \end{matrix} \tag{4.10}$$

with the initial state $\delta_{s_0} \otimes \nu_0^{\Theta} \otimes \nu_0^{\Theta'}$, where ν_0^{Θ} ($\in \mathcal{M}^m_{+1}(\Theta_0)$) and $\nu_0^{\Theta'}$ ($\in \mathcal{M}^m_{+1}(\Theta'_0)$) are known, but δ_{s_0} ($\in \mathcal{M}^p_{+1}(\mathcal{S}_0)$) is unknown.

Now, we can skip to the next section. However, in what follows we mention the concrete form of the family $\{\mathbf{O}_n = (X_n, 2^{X_n}, F_n)\}_{n=0}^N$, which corresponds to the measurement equation (4.2).

For each $n(= 0, 1, ..., N)$, consider an observable $\mathbf{O}'_n = (X_n, 2^{X_n}, F'_n)$ in $C(\mathcal{S}'_n)$, where $\mathcal{S}'_n$ is a compact space. And consider a continuous map $\lambda'_n : \mathcal{S}_n \times \Theta'_n \to \mathcal{S}'_n$, which induces the continuous map $\Lambda'_n : \mathcal{M}^m_{+1}(\mathcal{S}_n \times \Theta_n \times \Theta'_n) \to \mathcal{M}^m_{+1}(\mathcal{S}'_n)$ such that

$$[\Lambda'_n(\nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta} \otimes \nu_n^{\Theta'})](B_n) = (\nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta'})((\lambda'_n)^{-1}(B'_n))$$
$$(\forall \nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta} \otimes \nu_n^{\Theta'} \in \mathcal{M}^m_{+1}(\mathcal{S}_n \times \Theta_n \times \Theta'_n), \forall B'_n \subseteq \mathcal{S}'_n : \text{open}).$$

Thus, for each n ($\in T$), we can define the observable $\mathbf{O}_n = (X_n, 2^{X_n}, F_n)$ (in (4.10)) in $C(\mathcal{S}_n \times \Theta_n \times \Theta'_n)$ such that

$${}_{C(\mathcal{S}_n \times \Theta_n \times \Theta'_n)^*}\langle \nu_n^{\mathcal{S}_n} \otimes \nu_n^{\Theta} \otimes \nu_n^{\Theta'}, F_n(\Xi_n)\rangle_{C(\mathcal{S}_n \times \Theta_n \times \Theta'_n)} = {}_{C(\mathcal{S}'_n)^*}\langle \Lambda'_n(\nu_n^{\mathcal{S}_n} \otimes \nu_n^{\Theta} \otimes \nu_n^{\Theta'}), F'_n(\Xi_n)\rangle_{C(\mathcal{S}'_n)}$$
$$(\forall \nu_n^{\mathcal{S}} \otimes \nu_n^{\Theta} \otimes \nu_n^{\Theta'} \in \mathcal{M}^m_{+1}(\mathcal{S}_n \times \Theta_n \times \Theta'_n)).$$

4.2 Kalman filter in Noise

For simplicity, put $\widehat{\Theta}_n = \Theta_n \times \Theta'_n$ and $\nu_0^{\widehat{\Theta}} = \nu_0^{\Theta} \otimes \nu_0^{\Theta'}$. Then, we can rewrite the (4.10) as follows.

$$
\begin{array}{ccccccccc}
[C(\mathcal{S}_0 \times \widehat{\Theta}_0)] & \xleftarrow{\widehat{\Phi}_{0,1}} & [C(\mathcal{S}_1 \times \widehat{\Theta}_1)] & \xleftarrow{\widehat{\Phi}_{1,2}} & \cdots \xleftarrow{\widehat{\Phi}_{N-2,N-1}} & [C(\mathcal{S}_{N-1} \times \widehat{\Theta}_{N-1})] & \xleftarrow{\widehat{\Phi}_{N-1,N}} & [C(\mathcal{S}_N \times \widehat{\Theta}_N)] \\
(X_0, 2^{X_0}, F_0) & & (X_1, 2^{X_1}, F_1) & & \cdots & (X_{N-1}, 2^{X_{N-1}}, F_{N-1}) & & (X_N, 2^{X_N}, F_N)
\end{array}
$$

with the initial state $\delta_{s_0} \otimes \nu_0^{\widehat{\Theta}}$, where $\nu_0^{\widehat{\Theta}}$ $(\in \mathcal{M}_{+1}^m(\widehat{\Theta}_0))$ is known $\big($that is, ν_0^{Θ} $(\in \mathcal{M}_{+1}^m(\Theta_0))$ and $\nu_0^{\Theta'}$ $(\in \mathcal{M}_{+1}^m(\Theta'_0))$ are known$\big)$, but δ_{s_0} $(\in \mathcal{M}_{+1}^p(\mathcal{S}_0))$ is unknown.

Now, we get the sequential observable $[\mathbf{O}_T] \equiv [\{\mathbf{O}_t\}_{t \in T}, \{\widehat{\Phi}_{t_1,t_2} : C(\mathcal{S}_{t_2} \times \widehat{\Theta}_{t_2}) \to C(\mathcal{S}_{t_1} \times \widehat{\Theta}_{t_1})\}_{(t_1,t_2) \in T^2_{\leq}}]$. Then, we can construct the observable $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t \in T} X_t, 2^{\Pi_{t \in T} X_t}, \widetilde{F}_0)$ in $C(\mathcal{S}_0 \times \widehat{\Theta}_0)$, which is the realization of the sequential observable $[\mathbf{O}_T]$, such as

$$
\begin{array}{ccccccccc}
[C(\mathcal{S}_0 \times \widehat{\Theta}_0)] & \xleftarrow{\widehat{\Phi}_{0,1}} & [C(\mathcal{S}_1 \times \widehat{\Theta}_1)] & \xleftarrow{\widehat{\Phi}_{1,2}} & \cdots \xleftarrow{\widehat{\Phi}_{N-2,N-1}} & [C(\mathcal{S}_{N-1} \times \widehat{\Theta}_{N-1})] & \xleftarrow{\widehat{\Phi}_{N-1,N}} & [C(\mathcal{S}_N \times \widehat{\Theta}_N)] \\
F_0 & & F_1 & & \cdots & F_{N-1} & & F_N \\
\downarrow & & \downarrow & & & \downarrow & & \downarrow \\
\boxed{\begin{matrix}(F_0 \overset{\text{qp}}{\times} \widehat{\Phi} \widetilde{F}_1) \\ = \widetilde{F}_0\end{matrix}} & \xleftarrow{\widehat{\Phi}_{0,1}} & \begin{matrix}(F_1 \overset{\text{qp}}{\times} \widehat{\Phi} \widetilde{F}_2) \\ = \widetilde{F}_1\end{matrix} & \xleftarrow{\widehat{\Phi}_{1,2}} & \cdots \xleftarrow{\widehat{\Phi}_{N-2,N-1}} & \begin{matrix}(F_{N-1} \overset{\text{qp}}{\times} \widehat{\Phi} \widetilde{F}_N) \\ = \widetilde{F}_{N-1}\end{matrix} & \xleftarrow{\widehat{\Phi}_{N-1,N}} & \begin{matrix}(F_N) \\ = \widetilde{F}_N.\end{matrix}
\end{array}
\tag{4.11}
$$

$\big($The existence of the $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t \in T} X_t, 2^{\Pi_{t \in T} X_t}, \widetilde{F}_0)$ is assured by Theorem 2.9.$\big)$

Now, we can represent the "measurement" $\mathfrak{M}(\{\mathbf{O}_t\}_{t \in T}, \mathbf{S}(\delta_{s_0} \otimes \nu_0^{\widehat{\Theta}}))$ such as

$$
\mathfrak{M}(\{\mathbf{O}_t\}_{t \in T}, \mathbf{S}(\delta_{s_0} \otimes \nu_0^{\widehat{\Theta}})) = \mathbf{M}_{C(\mathcal{S}_0 \times \widehat{\Theta}_0)}(\widetilde{\mathbf{O}}_0, S(\delta_{s_0} \otimes \nu_0^{\widehat{\Theta}})).
$$

Assume we know that the measured value $(x_t)_{t \in T}$ $(\in \prod_{t \in T} X_t)$, obtained by the measurement $\mathbf{M}_{C(\mathcal{S}_0 \times \widehat{\Theta}_0)}(\widetilde{\mathbf{O}}_0, S(\delta_{s_0} \otimes \nu_0^{\widehat{\Theta}}))$, belongs to $\prod_{t \in T} \Xi_t$. Thus, Fisher's maximum likelihood method (cf. Theorem 3.8, Corollary 4.1) says that there is a reason to infer that the unknown s_0 $(\in \mathcal{S}_0)$ is determined by

$$
{}_{C(\mathcal{S}_0 \times \widehat{\Theta}_0)^*}\langle \delta_{s_0} \otimes \nu_0^{\widehat{\Theta}}, \widetilde{F}_0(\prod_{t \in T} \Xi_t) \rangle_{C(\mathcal{S}_0 \times \widehat{\Theta}_0)} = \max_{s \in \mathcal{S}_0} {}_{C(\mathcal{S}_0 \times \widehat{\Theta}_0)^*}\langle \delta_s \otimes \nu_0^{\widehat{\Theta}}, \widetilde{F}_0(\prod_{t \in T} \Xi_t) \rangle_{C(\mathcal{S}_0 \times \widehat{\Theta}_0)}.
$$

Let $\tau \in T$, and let $\{B^{(0,\tau)}_{\Pi_{t \in T} \Xi_t} \mid \prod_{t \in T} \Xi_t \in 2^{\Pi_{t \in T} X_t}\}$ be a family of Bayes operators. $\big($The existence is assured by Theorem 3.2.$\big)$ Then, we see, by Theorem 3.4, that the new S-state $\nu_{new}^{\mathcal{S}_\tau \times \widehat{\Theta}_\tau}$ $(\in \mathcal{M}_{+1}^m(\mathcal{S}_\tau \times \widehat{\Theta}_\tau))$ is defined by

$$
\nu_{new}^{\mathcal{S}_\tau \times \widehat{\Theta}_\tau} = R^{(0,\tau)}_{\Pi_{t \in T} \Xi_t}(\delta_{s_0} \otimes \nu_0^{\widehat{\Theta}}),
$$

where $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} : \mathcal{M}^m_{+1}(\mathcal{S}_0 \times \widehat{\Theta}_0) \to \mathcal{M}^m_{+1}(\mathcal{S}_\tau \times \widehat{\Theta}_\tau)$ is a normalized dual Bayes operator, i.e., $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\nu) = \frac{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\nu)}{\|(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\nu)\|}$ $(\forall \nu \in \mathcal{M}^m_{+1}(\mathcal{S}_0 \times \widehat{\Theta}_0))$. Thus there is a reason to consider that the new S-state (in $\mathcal{M}^m_{+1}(\mathcal{S}_\tau)$) is equal to $\nu^{\mathcal{S}_\tau}_{new}$ such that

$$\nu^{\mathcal{S}_\tau}_{new}(D_\tau) \equiv \nu^{\mathcal{S}_\tau \times \widehat{\Theta}_\tau}_{new}(D_\tau \times \widehat{\Theta}_\tau) \quad (\forall D_\tau (\subseteq \mathcal{S}_\tau) : \text{open set}). \tag{4.12}$$

4.3 Conclusion

In this paper, we studied “Kalman filter” in SMT (= statistical measurement theory).

$$\begin{cases} \boldsymbol{s}(n+1) = \lambda(\psi(n+1,n)\boldsymbol{s}(n), \boldsymbol{\theta}_1(n)) & : \text{stochastic difference state equation} \\ \qquad \text{(where } \lambda \text{ is the additive operation)} & \\ & \qquad (n = 0, 1, \cdots, N-1) \\ \boldsymbol{x}(n) = \lambda'(\boldsymbol{C}(n)\boldsymbol{s}(n), \boldsymbol{\theta}_2(n)) & : \text{measurement equation} \\ \qquad \text{(where } \lambda' \text{ is the additive operation)} & \end{cases} \tag{4.13}$$

Here, it is assumed that $\widetilde{\boldsymbol{\Psi}}(n+1,n)$, $\boldsymbol{C}(n)$, $\boldsymbol{\theta}_1(n)$ (and its initial distribution) and $\boldsymbol{\theta}_2(n)$ are known, where $\widetilde{\boldsymbol{\Psi}}(n+1,n)$: $K \times K$-dimensional transition matrix, $\boldsymbol{\theta}_1(n)$: L-dimensional input vector which represents a white noise, $\boldsymbol{C}(n)$: $L \times K$-dimensional measurement matrix, $\boldsymbol{\theta}_2(n)$: L-dimensional vector which represents a measurement error.

5 Measurement error model

Although we have several kinds of measurement error models in statistics (cf. Fuller), the following may be the simplest one:

$$\begin{aligned} &\widetilde{y}_n = \theta_0 + \theta_1 x_n + e_n, \quad \widetilde{x}_n = x_n + u_n \quad (n = 1, 2, ..., N), \\ &(e_n, u_n) \sim \text{NI}[\text{average}(0,0), \text{variance}(\sigma^2_{ee}, \sigma^2_{uu})]. \end{aligned} \tag{5.1}$$

The first equation is a classical regression specification, but the true explanatory variable x_n is not observed directly. The observed measure of x_n, denoted by $\widetilde{x}_n$, may be obtained by a certain measurement. Our present concern is how to infer the unknown parameters θ_0 and θ_1 from the measured data $\{(\widetilde{x}_n, \widetilde{y}_n)\}_{n=1}^N$. Precisely speaking, our purpose of this paper is to study this problem in general situations (i.e., without the assumption of normal distributions). We show that the above problem is naturally formulated in measurement theory, and assert that the method of measurement error model is valid for more general situations (i.e., the abstract form of (5.1) without the assumption of normal distributions).

5.1 Measuement error model in measurement theory

From here and onwards, we, for simplicity, devote ourselves to classical systems (i.e., the case that $\mathcal{A} = C(\Omega)$), and not quantum systems (cf. the last statement in this section). Put $\mathcal{A}_0 \equiv C(\Omega_0)$ and $\mathcal{A}_1 \equiv C(\Omega_1)$. Let Θ be a compact space, which may be called a *parameter state space*. Consider a parametrized continuous map $\psi^\theta : \Omega_0 \to \Omega_1$, $\theta \in \Theta$, which induces the parametrized Markov operator $\Psi^\theta : C(\Omega_1) \to C(\Omega_0)$ such that

$$(\Psi^\theta f_1)(\omega) = f_1(\psi^\theta(\omega)) \quad (\forall f_1 \in C(\Omega_1), \forall \omega \in \Omega_0). \tag{5.2}$$

Consider observables $\mathbf{O}_0 \equiv (X, \mathcal{F}, F)$ in $C(\Omega_0)$ and $\mathbf{O}_1 \equiv (Y, \mathcal{G}, G)$ in $C(\Omega_1)$. Recall that $\Psi^\theta \mathbf{O}_1$ can be identified with the observable in $C(\Omega_0)$ (cf. Axiom (ii)). Thus, we can consider the product observable $\widetilde{\mathbf{O}}^\theta = (X \times Y, \mathcal{F} \times \mathcal{G}, F \times \Psi^\theta G)$ in $C(\Omega_0)$. And, we get the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}^\theta, S_{[\delta_\omega]})$ $(\omega \in \Omega_0)$. Consider the N times repeated measurement of $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}^\theta, S_{[\delta_\omega]})$, which is represented by $\mathbf{M}_{C(\Omega_0^N)}(\bigotimes_{n=1}^N \widetilde{\mathbf{O}}^\theta, S_{[\otimes_{n=1}^N \delta_{\omega_n}]})$. Here, $\otimes_{n=1}^N \delta_{\omega_n} = \delta_{(\omega_1,\omega_2,\ldots,\omega_N)}$ $(\in \mathcal{M}_{+1}^p(\Omega_0^N))$ and $\bigotimes_{n=1}^N \widetilde{\mathbf{O}}^\theta = (X^N \times Y^N, \mathcal{F}^N \times \mathcal{G}^N, \bigotimes_{n=1}^N (F \times \Psi^\theta G))$ in $\bigotimes_{n=1}^N C(\Omega_0) \equiv C(\Omega_0^N)$, that is,

$$\begin{aligned}
&[(\bigotimes_{n=1}^N (F \times \Psi^\theta G))(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_N \times \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_N)](\omega_1, \omega_2, ..., \omega_N) \\
=&[F \times \Psi^\theta G(\Xi_1 \times \Gamma_1)](\omega_1) \cdot [F \times \Psi^\theta G(\Xi_2 \times \Gamma_2)](\omega_2) \cdots [F \times \Psi^\theta G(\Xi_N \times \Gamma_N)](\omega_N) \\
&(\forall \Xi_n \in \mathcal{F}, \forall \Gamma_n \in \mathcal{G}, \forall(\omega_1, \omega_2, ..., \omega_N) \in \Omega_0^N, \forall \theta \in \Theta).
\end{aligned} \tag{5.3}$$

Our present problem is as follows.

($\sharp$) Consider the measurement $\mathbf{M}_{C(\Omega_0^N)}(\bigotimes_{n=1}^N \widetilde{\mathbf{O}}^{\bar\theta}, S_{[\otimes_{n=1}^N \delta_{\bar\omega_n}]})$ where it is assumed that $(\bar\omega_1, \bar\omega_2, ..., \bar\omega_N)$ $(\in \Omega_0^N)$ and $\bar\theta$ $(\in \Theta)$ are unknown. Assume that we know that the measured value $(\widetilde{x}_1, ..., \widetilde{x}_N, \widetilde{y}_1, ..., \widetilde{y}_N)$ $(\in X^N \times Y^N)$ obtained by the measurement $\mathbf{M}_{C(\Omega_0^N)}(\bigotimes_{n=1}^N \widetilde{\mathbf{O}}^{\bar\theta}, S_{[\otimes_{n=1}^N \delta_{\bar\omega_n}]})$ belongs to $\prod_{n=1}^N (\Xi_n \times \Gamma_n)$. Then, infer the unknown $\bar\omega_1, \bar\omega_2, ..., \bar\omega_N$ and $\bar\theta$.

This problem is solved as follows. Define the observable $\widehat{\mathbf{O}} \equiv (X^N \times Y^N, \mathcal{F}^N \times \mathcal{G}^N, \widehat{H})$ in $C(\Omega_0^N \times \Theta)$ such that $[\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, ..., \omega_N, \theta) =$ (5.4). Note that we have the following identification:

$$\mathbf{M}_{C(\Omega_0^N \times \Theta)}(\widehat{\mathbf{O}}, S_{[(\otimes_{n=1}^N \delta_{\omega_n}) \otimes \delta_\theta]}) = \mathbf{M}_{C(\Omega_0^N)}(\bigotimes_{n=1}^N \widetilde{\mathbf{O}}^\theta, S_{[\otimes_{n=1}^N \delta_{\omega_n}]}). \tag{5.4}$$

Consider the measurement $\mathbf{M}_{C(\Omega_0^N \times \Theta)}(\widehat{\mathbf{O}}, S_{[(\otimes_{n=1}^N \delta_{\bar{\omega}_n}) \otimes \delta_{\bar{\theta}}]})$ where it is assumed that we do not know $\bar{\omega}_1, \bar{\omega}_2, ..., \bar{\omega}_N, \bar{\theta}$. Then, we can, by Fisher's maximum likelihood method (Corollary 3.4), infer the unknown state $(\otimes_{n=1}^N \delta_{\bar{\omega}_n}) \otimes \delta_{\bar{\theta}}$ such that

$$
\begin{aligned}
&[\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\bar{\omega}_1, ..., \bar{\omega}_N, \bar{\theta}) \\
= &\max_{(\omega_1,...,\omega_N,\theta) \in \Omega_0^N \times \Theta} [\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, ..., \omega_N, \theta). \qquad (5.5)
\end{aligned}
$$

This is the answer to the above problem (♯). Note that Ω_0, Ω_1, X and Y are not necessarily the real lines $\mathbf{R}$. Also, if readers are familiar with the theory of tensor product C^*-algebras, they can easily see that the assertion (5.5) is valid for even quantum systems under the condition that $\mathbf{O}_0$ and $\Psi^\theta \mathbf{O}_1$ commute. Therefore, the answer (5.5) is stated under the very general situations.

5.2 A simple example with normal distributions

In this section, we apply our main result (5.5) to the simple measurement error model (5.1) with normal distributions.

Let L be a sufficiently large number. Put $\Omega_0 = [-L, L], \Omega_1 = [-L^2 - L, L^2 + L], \Theta = [-L, L]^2$ and define the map $\psi^{(\theta_0,\theta_1)} : \Omega_0 \to \Omega_1$ such that

$$
\psi^{(\theta_0,\theta_1)}(\omega) = \theta_1 \omega + \theta_0 \quad (\forall \omega \in \Omega_0, \forall (\theta_0, \theta_1) \in \Theta).
$$

Also, put $(X, \mathcal{F}, F) = (\mathbf{R}, \mathcal{B}_{\mathbf{R}}, G^{\sigma_1})$ in $C(\Omega_0)$ and $(Y, \mathcal{G}, G) = (\mathbf{R}, \mathcal{B}_{\mathbf{R}}, G^{\sigma_2})$ in $C(\Omega_1)$ (cf. Example 2.2). Thus, we define the product observable $\widetilde{\mathbf{O}}^{(\theta_0,\theta_1)} = (X \times Y, \mathcal{F} \times \mathcal{G}, F \times \Psi^\theta G) = (\mathbf{R}^2, \mathcal{B}_{\mathbf{R}^2}, H^{(\theta_0,\theta_1)})$, where $H^{(\theta_0,\theta_1)} \equiv F \times \Psi^\theta G$, in $C(\Omega_0)$ such that

$$
[H^{(\theta_0,\theta_1)}(\Xi \times \Gamma)](\omega) = (\frac{1}{\sqrt{2\pi\sigma_1\sigma_2}})^2 \iint_{\Xi \times \Gamma} \exp[-\frac{(x-\omega)^2}{2\sigma_1^2} - \frac{(y - (\theta_1\omega + \theta_0))^2}{2\sigma_2^2}] dxdy
$$
$$
(\forall \Xi \in \mathcal{B}_{\mathbf{R}}, \forall \Gamma \in \mathcal{B}_{\mathbf{R}}, \forall \omega \in \Omega_0).
$$

Thus, we have the observable $\widehat{\mathbf{O}} = (\mathbf{R}^{2N}, \mathcal{B}_{\mathbf{R}^{2N}}, \widehat{H})$ in $C(\Omega_0^N \times \Theta)$ such that

$$
\begin{aligned}
&[\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, ..., \omega_N, \theta_0, \theta_1) = \prod_{n=1}^N [H^{(\theta_0,\theta_1)}(\Xi_n \times \Gamma_n)](\omega_n) \\
=&(\frac{1}{\sqrt{2\pi\sigma_1\sigma_2}})^{2N} \int \cdots \int_{\Pi_{n=1}^N (\Xi_n \times \Gamma_n)} e^{-\frac{\sum_{n=1}^N (x_n - \omega_n)^2}{2\sigma_1^2} - \frac{\sum_{n=1}^N (y_n - (\theta_1\omega_n + \theta_0))^2}{2\sigma_2^2}} dx_1 dy_1 \cdots dx_N dy_N \\
&(\forall \Xi_n \in \mathcal{B}_{\mathbf{R}}, \forall \Gamma_n \in \mathcal{B}_{\mathbf{R}}, \forall (\omega_1, \omega_2, ..., \omega_N) \in \Omega_0^N, \forall (\theta_0, \theta_1) \in \Theta).
\end{aligned}
$$

Assume the conditions in the problem (♯) in Section 5.1, and further add that

$$\Xi_n^\varepsilon = [\widetilde{x}_n - \varepsilon, \widetilde{x}_n + \varepsilon], \quad \Gamma_n^\varepsilon = [\widetilde{y}_n - \varepsilon, \widetilde{y}_n + \varepsilon] \quad \text{(for sufficiently small positive } \varepsilon\text{)}.$$

Then, our main result (5.5) says that

$$\max_{(\omega_1,...,\omega_N,\theta_0,\theta_1)\in\Omega_0^N\times\Theta} [\widehat{H}^{(\theta_0,\theta_1)}(\Xi_1^\varepsilon \times \cdots \times \Xi_N^\varepsilon \times \Gamma_1^\varepsilon \times \cdots \times \Gamma_N^\varepsilon)](\omega_1, ..., \omega_N, \theta_0, \theta_1)$$

$$\Longleftrightarrow \min_{(\omega_1,...,\omega_N,\theta_0,\theta_1)\in\Omega_0^N\times\Theta} \Big[\sum_{n=1}^N (\frac{\widetilde{x}_n}{\sigma_1} - \frac{\omega_n}{\sigma_1})^2 + \sum_{n=1}^N (\frac{\widetilde{y}_n}{\sigma_2} - (\frac{\theta_1\sigma_1}{\sigma_2}\frac{\omega_n}{\sigma_1} + \frac{\theta_0}{\sigma_2}))^2\Big] \quad \text{(since } \varepsilon \text{ is small)}$$

(Here, note that the distance between a point $(\frac{\widetilde{x}_n}{\sigma_1}, \frac{\widetilde{y}_n}{\sigma_2})$ and a line $y = \frac{\theta_1\sigma_1}{\sigma_2}x + \frac{\theta_0}{\sigma_2}$ is equal to $\frac{|\widetilde{y}_n - \theta_1\widetilde{x}_n - \theta_0|}{\sqrt{\sigma_2^2 + \sigma_1^2\theta_1^2}}$. Then, we see)

$$\Longleftrightarrow \min_{(\theta_0,\theta_1)\in\Theta} \frac{\sum_{n=1}^N (\widetilde{y}_n - \theta_1\widetilde{x}_n - \theta_0)^2}{\sigma_2^2 + \sigma_1^2\theta_1^2} \tag{5.6}$$

$$\Longleftrightarrow \begin{cases} \sum_{n=1}^N (\widetilde{y}_n - \bar{\theta}_1\widetilde{x}_n - \bar{\theta}_0) = 0 & (\leftarrow \ \frac{\partial}{\partial\theta_0}(5.6) = 0), \\ \sum_{n=1}^N (\bar{\theta}_1\widetilde{y}_n\sigma_1^2 + \widetilde{x}_n\sigma_2^2 - \bar{\theta}_0\bar{\theta}_1\sigma_1^2)(\widetilde{y}_n - \bar{\theta}_1\widetilde{x}_n - \bar{\theta}_0) = 0 & (\leftarrow \ \frac{\partial}{\partial\theta_1}(5.6) = 0). \end{cases} \tag{5.7}$$

Thus, the unknown parameters $\bar{\theta}_0$ and $\bar{\theta}_1$ are inferred by the solution of this equation (5.7). Note that this is a direct consequence of our main result (5.5), which is the general assertion applicable to both classical and quantum systems (cf. the last statement in Section 5.1). Also as mentioned in Section 1, we can see that, if $\sigma_1 = 0$, the (5.6) (or the (5.7)) is the same as the result of ordinary regression analysis.

References

[1] Davies,E.B. *Quantum theory of open systems*, Academic Press 1976

[2] Fuller, W.A. *Measurement error models* , John Wiley & Sons (1987)

[3] Holevo, A.S. *Probabilistic and statistical aspects of quantum theory* North-Holland publishing company (1982)

[4] Ishikawa,S. *Fuzzy inferences by algebraic method*, Fuzzy Sets and Systems **87**, 181-200 (1997)

[5] Ishikawa,S. *Statistics in measurements*, Fuzzy Sets and Systems **116**, 141-154 (2000)

[6] Ishikawa,S. *The existence theorem of Kalman filter in measurement theory* Far east journal of dynamical system **3**, 175-204 (2001)

[7] Ishikawa,S. *Fundamentals of measurements (Pleminary lecture)*, The 2002 annual conference of the Japan Society for Industrial and Applied mathematics (20th Sep. 2002)
http://www.jstage.jst.go.jp/article/pjsiam/2002/0/2002_16/

[8] Kalman, R. E. *A new approach to linear filtering and prediction problems*, Trans. ASME, J. Basic Eng., **82**, 35 (1960)

[9] Kolmogorov,A. *Foundations of probability (translation)*, Chelsea Publishing Co. 1950

[10] Sakai,S. *C^*-algebras and W^*-algebras*, Ergebnisse der Mathematik und ihrer Grenzgebiete (Band 60), Springer-Verlag, Berlin, Heidelberg, New York 1971

[11] Takesaki, M. *Theory of operator algebras*, New York : Springer-Verlag , 1979

[12] von Neumann, J. *Mathematical foundations of quantum mechanics*, Springer Verlag, Berlin (1932)

[13] Yosida,Y. *Functional analysis*, Springer-Verlag (Sixth Edition) 1980

6 Appendix (The proof of Theorem 3.8)

PROOF. It will be proved by induction. Let $\mathbf{O}'_\tau = (Y_\tau, 2^{Y_\tau}, G_\tau)$ be any observable in $C(\Omega_\tau)$.
[Step 1]. First, define the positive bounded linear operator $\widehat{B}^{(\tau,\tau)}_{\Pi_{t\in T_\tau}\Xi_t} : C(\Omega_\tau) \to C(\Omega_\tau)$ such that

$$\widehat{B}^{(\tau,\tau)}_{\Pi_{t\in T_\tau}\Xi_t}(g_\tau) = \widetilde{F}_\tau(\prod_{t\in T_\tau} \Xi_t) \times g_\tau \qquad (\forall g_\tau \in C(\Omega_\tau)), \tag{6.1}$$

and define the observable $\widehat{\mathbf{O}}_\tau \Big(\equiv ((\prod_{t\in T_\tau} X_t) \times Y_\tau, 2^{(\Pi_{t\in T_\tau} X_t)\times Y_\tau}, \widehat{F}_\tau) \Big)$ in $C(\Omega_\tau)$ such that

$$\widehat{F}_\tau(\prod_{t\in T_\tau} \Xi_t \times \Gamma_\tau) = \widehat{B}^{(\tau,\tau)}_{\Pi_{t\in T_\tau}\Xi_t}(G_\tau(\Gamma_\tau)) \qquad (\forall \Gamma_\tau \in 2^{Y_\tau}), \tag{6.2}$$

which is clearly the Heisenberg picture representation of the *sequential observable* $[\{\overline{\mathbf{O}}_t\}_{t\in T_\tau}; \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T_\tau\setminus\{\tau\}}]$, where $\overline{\mathbf{O}}_t = \mathbf{O}_t$ (if $t \neq \tau$), $= \mathbf{O}_\tau \times \mathbf{O}'_\tau$ (if $t = \tau$). Thus, the operator $\widehat{B}^{(\tau,\tau)}_{\Pi_{t\in T_\tau}\Xi_t} : C(\Omega_\tau) \to C(\Omega_\tau)$ is the Bayes operator induced from the $\widetilde{\mathbf{O}}_\tau$ $\Big(\equiv (\prod_{t\in T_\tau} X_t, 2^{\Pi_{t\in T_\tau} X_t}, \widetilde{F}_\tau) \Big)$.
[Step 2 (Assumption)]. Let s be any element in $T \setminus \{0\}$ such that $s \leq \tau$. Here, assume that $\widehat{B}^{(s,\tau)}_{\Pi_{t\in T_s}\Xi_t} : C(\Omega_\tau) \to C(\Omega_s)$ is the Bayes operator induced from the $\widetilde{\mathbf{O}}_s$ $\Big(\equiv (\prod_{t\in T_s} X_t, 2^{\Pi_{t\in T_s} X_t}, \widetilde{F}_s) \Big)$. That is, there exists an observable $\widehat{\mathbf{O}}_s \Big(\equiv ((\prod_{t\in T_s} X_t) \times Y_\tau, 2^{(\Pi_{t\in T_s} X_t)\times Y_\tau}, \widehat{F}_s) \Big)$ in $C(\Omega_s)$ such that

(i) $\widehat{\mathbf{O}}_s$ is the Heisenberg picture representation (cf. Theorem 2.9) of the *sequential observable* $[\{\widehat{\mathbf{O}}_t\}_{t\in T_s};\ \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T_s\setminus\{s\}}]$, where $\overline{\mathbf{O}}_t = \mathbf{O}_t$ (if $t \neq \tau$), $= \mathbf{O}_\tau \times \mathbf{O}'_\tau$ (if $t = \tau$).

(ii) $\widehat{F}_s((\prod_{t\in T_s}\Xi_t)\times\Gamma_\tau) = \widehat{B}^{(s,\tau)}_{\prod_{t\in T_s}\Xi_t}(G_\tau(\Gamma_\tau)) \quad (\forall \Xi_t \in 2^{X_t}\ (\forall t\in T_s), \forall \Gamma_\tau \in 2^{Y_\tau})$,

(iii) $\widehat{F}_s((\prod_{t\in T_s}\Xi_t)\times Y_\tau) = \widetilde{F}_s(\prod_{t\in T_s}\Xi_t) \quad (\forall \Xi_t \in 2^{X_t}\ (\forall t\in T_s))$.

[Step 3]. Let $(x_t)_{t\in T_{\pi(s)}}$ be any element in $\prod_{t\in T_{\pi(s)}} X_t$. Note that $\{(x_t)_{t\in T_{\pi(s)}}\} = \prod_{t\in T_{\pi(s)}}\{x_t\}$. Define the positive bounded linear operator $\widehat{B}^{(\pi(s),\tau)}_{\prod_{t\in T_{\pi(s)}}\{x_t\}} : C(\Omega_\tau)\to C(\Omega_{\pi(s)})$ by

$$[\widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\{x_t\}}(g_\tau)](\omega_{\pi(s)}) = \frac{[\widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\})](\omega_{\pi(s)}) \times [\Phi_{\pi(s),s}\widehat{B}^{(s,\tau)}_{\Pi_{t\in T_s}\{x_t\}}(g_\tau)](\omega_{\pi(s)})}{[\widetilde{F}_{\pi(s)}((\Pi_{t\in T_{\pi(s)}\setminus T_s}X_t)\times\prod_{t\in T_s}\{x_t\})](\omega_{\pi(s)})} \tag{6.3}$$

$$(\forall g_\tau \in C(\Omega_\tau), \forall \omega_{\pi(s)} \in \Omega_{\pi(s)}).$$

Here, the above is assumed to be equal to 0 if the denominator of (6.3) is equal to 0 $\Big($i.e., $[\widetilde{F}_{\pi(s)}\ ((\Pi_{t\in T_{\pi(s)}\setminus T_s}X_t)\times\prod_{t\in T_s}\{x_t\})](\omega_{\pi(s)}) = 0\Big)$. And thus, we can define the positive bounded linear operator $\widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\Xi_t} : C(\Omega_\tau)\to C(\Omega_{\pi(s)})$ by

$$\widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\Xi_t} = \sum_{(x_t)_{t\in T_{\pi(s)}}\in\prod_{t\in T_{\pi(s)}}\Xi_t} \widehat{B}^{(\pi(s),\tau)}_{\{(x_t)_{t\in T_{\pi(s)}}\}}.$$

Define the observable $\widehat{\mathbf{O}}_{\pi(s)} \equiv ((\prod_{t\in T_{\pi(s)}} X_t)\times Y_\tau,\ 2^{(\Pi_{t\in T_{\pi(s)}} X_t)\times Y_\tau}, \widehat{F}_{\pi(s)})$ in $C(\Omega_{\pi(s)})$ such that

$$\widehat{F}_{\pi(s)}((\prod_{t\in T_{\pi(s)}}\Xi_t)\times\Gamma_\tau) = \widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\Xi_t}(G_\tau(\Gamma_\tau)) \qquad (\forall \Xi_t\in 2^{X_t}\ (\forall t\in T_{\pi(s)}), \forall\Gamma_\tau\in 2^{Y_\tau}),$$

which is clearly the Heisenberg picture representation of the *sequential observable* $[\{\overline{\mathbf{O}}_t\}_{t\in T_{\pi(s)}};$ $\{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T_{\pi(s)}\setminus\{\pi(s)\}}]$, where $\overline{\mathbf{O}}_t = \mathbf{O}_t$ (if $t\neq\tau$), $= \mathbf{O}_\tau\times\mathbf{O}'_\tau$ (if $t=\tau$). Also, we see that (cf. Theorem 2.9) of

$$\widehat{F}_{\pi(s)}((\prod_{t\in T_{\pi(s)}}\Xi_t)\times Y_\tau) = \widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\Xi_t) \qquad (\forall\Xi_t\in 2^{X_t}\ (\forall t\in T_{\pi(s)})).$$

That is because we see

$$
\begin{aligned}
&\widehat{F}_{\pi(s)}((\prod_{t\in T_{\pi(s)}} \Xi_t) \times Y_\tau) = \widehat{B}^{(\pi(s),\tau)}_{\prod_{t\in T_{\pi(s)}}\Xi_t}(1_\tau) = \sum_{(x_t)_{t\in T_{\pi(s)}}\in\prod_{t\in T_{\pi(s)}}\Xi_t} \widehat{B}^{(\pi(s),\tau)}_{\prod_{t\in T_{\pi(s)}}\{x_t\}}(1_\tau) \\
&= \sum_{(x_t)_{t\in T_{\pi(s)}}\in\prod_{t\in T_{\pi(s)}}\Xi_t} \frac{\widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\}) \times \Phi_{\pi(s),s}\widehat{B}^{(s,\tau)}_{\prod_{t\in T_s}\{x_t\}}(1_\tau)}{\widetilde{F}_{\pi(s)}((\prod_{t\in T_{\pi(s)}\setminus T_s} X_t) \times \prod_{t\in T_s}\{x_t\})} \\
&= \sum_{(x_t)_{t\in T_{\pi(s)}}\in\prod_{t\in T_{\pi(s)}}\Xi_t} \frac{\widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\}) \times \widetilde{F}_{\pi(s)}((\prod_{t\in T_{\pi(s)}\setminus T_s} X_t) \times \prod_{t\in T_s}\{x_t\})}{\widetilde{F}_{\pi(s)}((\prod_{t\in T_{\pi(s)}\setminus T_s} X_t) \times \prod_{t\in T_s}\{x_t\})} \\
&= \sum_{(x_t)_{t\in T_{\pi(s)}}\in\prod_{t\in T_{\pi(s)}}\Xi_t} \widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\}) = \widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\Xi_t). \qquad (6.4)
\end{aligned}
$$

Therefore, we see that $\widehat{B}^{(\pi(s),\tau)}_{\prod_{t\in T_{\pi(s)}}\Xi_t} : C(\Omega_\tau) \to C(\Omega_{\pi(s)})$ is the Bayes operator induced from the $\widetilde{\mathbf{O}}_{\pi(s)}\Big(= (\prod_{t\in T_{\pi(s)}} X_t, 2^{\prod_{t\in T_{\pi(s)}} X_t}, \widetilde{F}_{\pi(s)})\Big)$. Thus, we can, by induction, finish the proof since it suffices to put $B^{(0,\tau)}_{\prod_{t\in T}\Xi_t} = \widehat{B}^{(0,\tau)}_{\prod_{t\in T_0}\Xi_t}$. □

Department of Mathematics
Faculty of Science and Technology
Keio University

Research Report

2005

[05/001] T. Kawazoe, *On the Littlewood-Paley g function and the Lusin area function on real rank 1 semisimple Lie groups*, KSTS/RR-05/001, March 28, 2005

[05/002] A. Sako, T. Suzuki, *Partition functions of supersymmetric gauge theories in noncommutative* $\mathbb{R}^{2D}$ *and their unified perspective,* KSTS/RR-05/002, April 25, 2005

[05/003] K. Aoyama, K. Shimizu, *A generalization of the inverse trinomial,* KSTS/RR-05/003, June 23, 2005

[05/004] T. Tanaka, *Algebraic independence of modified reciprocal sums of products of Fibonacci numbers,* KSTS/RR-05/004, July 27, 2005

[05/005] Y. Aoki, T. Kato, *New method of ground reflection coefficient reconstruction and shape recovery by another SAR signal modeling,* KSTS/RR-05/005, October 12, 2005

[05/006] A. Sako, T. Suzuki, *Dimensional reduction of Seiberg-Witten monopole equations, N=2 noncommutative supersymmetric field theories and Young diagrams,* KSTS/RR-05/006, November 28, 2005

2006

[06/001] N. Kumasaka, R. Shibata, *High dimensional data visualisation: Textile plot*, KSTS/RR-06/001, February 13, 2006

[06/002] H. Omori, Y. Maeda, N. Miyazaki, A. Yoshioka, *Geometric objects in an approach to quantum geometry*, KSTS/RR-06/002, March 15, 2006

[06/003] M. Nakasuji, *Prime geodesic theorem for higher dimensional hyperbolic manifold*, KSTS/RR-06/003, March 15, 2006

[06/004] T. Kawazoe, *Uncertainty principle for the Fourier-Jacobi transform*, KSTS/RR-06/004, April 11, 2006

[06/005] K. Kikuchi, S. Ishikawa, *Regression analysis, Kalman filter and measurement error model in measurement theory*, KSTS/RR-06/005, April 19, 2006