Title	Chapter 6 : Fisher's statistics II (related to Axioms 1 and 2)
Sub Title	
Author	石川, 史郎(Ishikawa, Shiro)
Publisher	Keio University Press Inc.
Publication year	2006
Jtitle	Mathematical Foundations of Measurement Theory (測定理論の数学的基礎). (2006.) ,p.127- 148
JaLC DOI	
Abstract	
Notes	
Genre	Book
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=KO52003001-00000000-0127

Chapter 6

Fisher's statistics II (related to Axioms 1 and 2)

As mentioned in Chapters 2 and 3, measurement theory is formulated as follows:

$$\underset{\text{[Axiom 1 (2.37)]}}{\text{PMT} = \text{measurement}} + \underset{\text{[Axiom 2 (3.26)]}}{\text{the relation among systems}} \quad \text{in } C^*\text{-algebra} . \qquad \underset{(=(1.4))}{(6.1)}$$

In this chapter we study the relation between Fisher's statistics (mentioned in the previous chapter) and Axiom 2. Particularly we show that regression analysis can be completely understood within the framework of Axioms 1 and 2. We expect that our result will make the readers notice that regression analysis is more profound than they usually think. As mentioned in Chapter 1 (*cf.* Declaration (1.11)), we assert that the results in Chapters 5 and 6 guarantee that "Fisher's statistics is theoretically true (in PMT)". [1]

6.1 Regression analysis I

6.1.1 Introduction

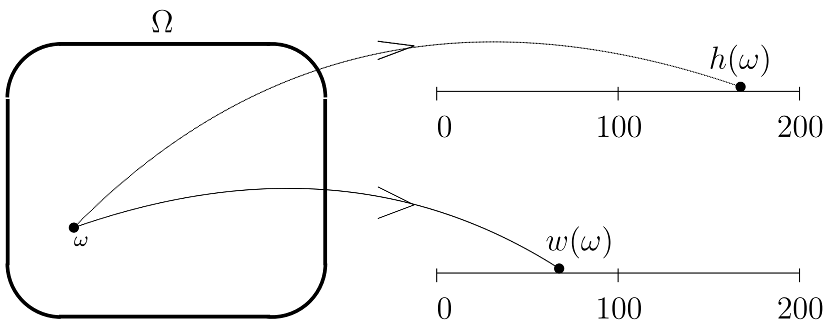
The purpose of this chapter is to study and understand "regression analysis" completely under Axiom 1 and 2 (of measurement theory). The following Example 6.1 is the most typical in all examples of "regression analysis".

Example 6.1. [A typical example of regression analysis]. Let $\Omega \equiv \{\omega_1, \omega_2, ..., \omega_{100}\}$ be a set of all students of a certain high school. Define $h : \Omega \to [0, 200]$ [resp. $w : \Omega \to [0, 200]$]

[1] We believe that only "Fisher's maximum likelihood method" and "regression analysis" are most essential in statistics. Thus we believe that, in order to justify statistics, it suffices to show that the two (i.e., "Fisher's maximum likelihood method" and "regression analysis") are formulated in PMT.

such that:

$$h(\omega_n) = \text{"the height of a student } \omega_n\text{"} \quad (n = 1, 2, ..., 100)$$
$$\Big[\text{resp. } w(\omega_n) = \text{"the weight of a student } \omega_n\text{"} \quad (n = 1, 2, ..., 100)\Big]$$



(Note that this is a special case of Fig. (3.20).) Assume that:

(1) The principal of this high school knows the both functions h and w. That is, he knows the exact data of the height and weight concerning all students.

Also, assume that:

(2) Some day, a certain student helped a drowned girl. But, he left without reporting the name. Thus, all information that the principal knows is as follows:

 (i) he is a student of his high school.

 (ii) his height [resp. weight] is about 170 cm [resp. about 80 kg].

Now we have the following question:

- Under the above assumption (1) and (2), how does the principal infer who is he?

This is just what regression analysis says. For the solution, see Regression Analysis I (6.7) later.

■

In order to explain our main assertion, let us begin with the following Example 6.2 (the conventional argument of regression analysis in Fisher's maximum likelihood method), which is easy and well-known.

Example 6.2. [The conventional argument of regression analysis in Fisher's method]. We have a rectangular water tank filled with water. Assume that the height of water at time t is given by the following function $h(t)$:

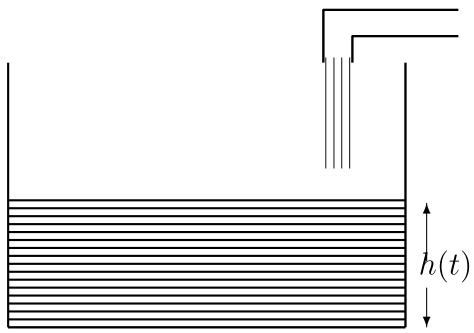
$$h(t) = \alpha_0 + \beta_0 t, \tag{6.2}$$

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 unit time. The measured height $h_m(t)$ of water at time t is assumed to be represented by

$$h_m(t) = \alpha_0 + \beta_0 t + e(t), \tag{6.3}$$

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:

$$h_m(1) = 1.9, \quad h_m(2) = 3.0, \quad h_m(3) = 4.7. \tag{6.4}$$



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 (6.4).

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

$$(\alpha_0, \beta_0) = (0.4, 1.4). \tag{6.5}$$

$\Big($For the derivation of (6.5) from (6.4), see Example 6.4 (6.16) later.$\Big)$ And next, we can infer that

$$h(2) = 3.2, \tag{6.6}$$

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 6.2 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 6.2 as follows:

(ii) What kinds of axioms are hidden behind the argument in Example 6.2? And moreover, justify the argument in Example 6.2 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 6.2 is true (or not). That is because there is no justification without an axiomatic formulation. In this sense, we believe that the above question (ii) is the most important problem in theoretical statistics. Also, if some know 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.* [71], [84]), it is a matter of course that they want 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 (*cf.* Corollary 5.6). Also, the derivation of the (6.6) from the (6.5) 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 chapter. However, in order to describe the above (iii) precisely, we must make vast preparations.

Our standing point of this book is extremely theoretical (and not practical). However we expect that many statisticians will be interested in our proposal. That is because we believe that every statistician may want to know the justification of both the (6.5) and the (6.6) in Example 6.2.

6.1.2 Regression analysis I in measurements

By the results in the previous chapters (i.e., Theorem 3.7 and Corollary 5.6), we can easily propose:

REGRESSION ANALYSIS I [The conventional regression analysis in PMT]. (6.7)
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 a* C^**-algebra* $C(\Omega_t)$ *be given for each* $t \in T$. *Let* $\widetilde{\mathbf{O}}_0$ *be the Heisenberg picture representation of the sequential observable* $[\{\mathbf{O}_t\}_{t\in T}, \{\Phi_{t,\pi(t)} : C(\Omega_t) \to C(\Omega_{\pi(t)})\}_{t\in T\setminus\{0\}}\ ]$ *in* $C(\Omega_0)$. *Then, we have a measurement*

$$\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0), S_{[*]}). \qquad (\textit{cf. Theorem 3.7}).$$

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} $(\in \mathcal{M}_{+1}^p(\Omega))$ $\Big($*defined by (6.9)*$\Big)$ *are equal. That is, Corollary 5.6 says that there is a reason to infer that*

$$[\,*\,] = \text{“the state after the measurement } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = \delta_{\omega_0}. \qquad (6.8)$$

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

$$[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega_0) = \max_{\omega\in\Omega_0} [\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega). \qquad (6.9)$$

■

Remark 6.3. [Regression analysis I]. The above regression analysis is quite applicable. For example, note that the “$\Phi_{\pi(t),t} : C(\Omega_t) \to C(\Omega_{\pi(t)})$” is generally assumed to be Markov operators (and not homomorphisms). In this sense, Regression analysis I may not be “conventional”.

■

Now we shall review Example 6.2 in the light of Regression Analysis I.

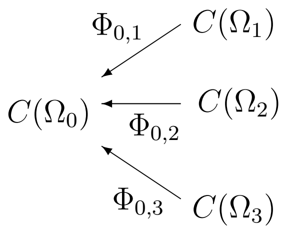
Example 6.4. [Continued from Example 6.2, the conventional argument of regression analysis in Fisher's method]. Put $\Omega_0 = [0.0,\ 1.0] \times [0.0,\ 2.0]$, and put $\Omega_1 = \Omega_2 = \Omega_3 = [0.0,\ 10.0]$. For each t $(\in \{1, 2, 3\})$, define a continuous map $\phi_{0,t} : \Omega_0 \to \Omega_t$ such that:

$$\Omega_0(\ \equiv [0.0,\ 1.0] \times [0.0,\ 2.0]) \ni \omega \equiv (\alpha, \beta) \underset{\phi_{0,t}}{\mapsto} \alpha + \beta t \in \Omega_t(\ \equiv [0.0,\ 10.0]). \qquad (6.10)$$

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

$$[\Phi_{0,t} f_t](\omega) = f_t(\phi_{0,t}(\omega)) \qquad (\forall\omega \in \Omega_0, \forall f_t \in C(\Omega_t)). \qquad (6.11)$$

It is usual to assume that regression analysis is applied to the system with a parallel structure such as in the figure (6.12). $\Big($From the peculiarity of this problem, we can also assume that this system has a series structure. However, we are not concerned with it.$\Big)$



(6.12)

For each $t \in \{1,2,3\}$, consider the discrete Gaussian observable $\mathbf{O}_{\sigma^2,N} \equiv (X_N, 2^{X_N}, F_{\sigma,N})$ in $C(\Omega_t)$, (*cf.*(2.60) in Example 2.18). That is,

$$\Omega_t = [0.0,\ 10.0], \qquad X_N = \{\frac{k}{N} \mid k = 0, \pm 1, \pm 2, ..., \pm N^2\},$$

and

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

(*cf.* (2.aa60) in Example 2.18)

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:

$$
\begin{aligned}
&[\widetilde{F}_0(\Xi_1 \times \Xi_2 \times \Xi_3)](\omega) = [\Phi_{0,1} F_{\sigma^2,N}](\omega) \cdot [\Phi_{0,2} F_{\sigma^2,N}](\omega) \cdot [\Phi_{0,3} F_{\sigma^2,N}](\omega) \\
=&[F_{\sigma^2,N}(\Xi_1)](\phi_{0,1}(\omega)) \cdot [F_{\sigma^2,N}(\Xi_2)](\phi_{0,2}(\omega)) \cdot [F_{\sigma^2,N}(\Xi_3)](\phi_{0,3}(\omega)) \\
&(\forall \Xi_1, \Xi_2, \Xi_3 \in 2^{X_N},\ \forall \omega = (\alpha, \beta) \in \Omega_0 = [0.0,\ 1.0] \times [0.0,\ 2.0]). \qquad (6.13)
\end{aligned}
$$

Then, we have the measurement $\mathbf{M}_{C(\Omega_0)}(\ \widetilde{\mathbf{O}}_0, S_{[*]})$. The (6.4) says that the measured value obtained by the measurement $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})$ is equal to

$$(1.9,\ 3.0,\ 4.7)\ (\in X_N^3). \qquad (6.14)$$

Here, Fisher's method (Corollary 5.6) says that it suffices to solve the problem

$$\text{“Find } (\alpha_0, \beta_0) \text{ such as } \max_{(\alpha,\beta)\in\Omega_0} [\widetilde{F}_0(\{1.9\} \times \{3.0\} \times \{4.7\}(\alpha, \beta)\text{”}. \qquad (6.15)$$

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

$$\begin{aligned}
&(6.15) \Rightarrow \max_{(\alpha,\beta)\in\Omega_0} \frac{1}{\sqrt{2\pi\sigma^2}^3} \iiint_{\Xi_1\times\Xi_2\times\Xi_3} e^{[-\frac{(x_1-(\alpha+\beta))^2+(x_2-(\alpha+2\beta))^2+(x_3-(\alpha+3\beta))^2}{2\sigma^2}]} dx_1 dx_2 dx_3 \\
&\Rightarrow \max_{(\alpha,\beta)\in\Omega_0} \exp\Big(- [(1.9-(\alpha+\beta))^2 + (3.0-(\alpha+2\beta))^2 + (4.7-(\alpha+3\beta))^2]/(2\sigma^2)\Big) \\
&\Rightarrow \min_{(\alpha,\beta)\in\Omega_0} [(1.9-(\alpha+\beta))^2 + (3.0-(\alpha+2\beta))^2 + (4.7-(\alpha+3\beta))^2] \\
&\qquad\qquad\qquad \text{(by the least squares method)} \\
&\Rightarrow \begin{cases} (1.9-(\alpha+\beta)) + (3.0-(\alpha+2\beta)) + (4.7-(\alpha+3\beta)) = 0 \\ (1.9-(\alpha+\beta)) + 2(3.0-(\alpha+2\beta)) + 3(4.7-(\alpha+3\beta)) = 0 \end{cases} \\
&\Rightarrow \quad (\alpha_0, \beta_0) = (0.4, 1.4). \qquad (6.16)
\end{aligned}$$

This is the conclusion of Regression Analysis I (6.7). 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 (6.16) $\Big(= (6.5)$ in Example 6.2$\Big)$.

This fact will be important in §6.3.

■

The above may be the standard argument of the conventional regression analysis in measurement theory. However, our problem (i) in Example 6.2 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 6.2, we usually consider that it suffices to calculate $h(2)$ $\Big(\equiv \phi_{0,2}(0.4, 1.4)\Big)$ in the following:

$$h(2) = 0.4 + 1.4 \times 2 = 3.2. \qquad (6.17)$$

However, this is doubtful. $\Big($In fact, this (6.17) is not always true in general situations. (*cf.* Regression analysis II (6.51) later).$\Big)$ We should not rely on "a common sense" but Axioms 1 and 2. That is, we must solve the problem:

- How can the above (6.17) $\Big(= (6.6)$ in Example 6.2$\Big)$ be deduced from Axioms 1 and 2?

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

6.2 Bayes operator, Schrödinger picture, and S-states

In order to improve Regression Analysis I (introduced in the previous section), in this section we make some preparations (i.e., Bayes operator, Schrödinger picture, S-state, etc.). Our main assertion (Regression Analysis II) will be mentioned in §6.3. We begin with the following definition, which is a general form of "Bayes operator" in Remark 5.7.

Definition 6.5. [Bayes operator (or precisely, Bayes-Kalman 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* $\mathbf{O}_t \equiv (X_t, \mathcal{F}_t, F_t)$ *in* $C(\Omega_t)$ *be given for each* $t \in T$*. Let* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, \bigotimes_{t\in T} \mathcal{F}_t, \widetilde{F}_0)$ *be as in Theorem 3.7 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* $B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} : C(\Omega_\tau) \to C(\Omega_0)$ *satisfies the following condition* (BO), *we call* $\{B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in \mathcal{F}_t \ (\forall t \in T)\}$ [*resp.* $B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}$] *a family of Bayes operators* [*resp. a Bayes operator*]*:*

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

- (i) $\widehat{\mathbf{O}}_0$ *is the Heisenberg picture representation 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) $\widehat{F}_0((\prod_{t\in T} \Xi_t) \times \Gamma_\tau) = B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau)) \quad (\forall \Xi_t \in \mathcal{F}_t \ (\forall t \in T), \forall \Gamma_\tau \in \mathcal{G}_\tau)$*,*
- (iii) $\widehat{F}_0((\prod_{t\in T} \Xi_t) \times Y_\tau) = \widetilde{F}_0(\prod_{t\in T} \Xi_t)\Big(\equiv B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)\Big)$*,* $(\forall \Xi_t \in \mathcal{F}_t \ (\forall t \in T))$*, 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:*

$$R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\nu) = \frac{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\nu)}{\|(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\nu)\|_{\mathcal{M}(\Omega_\tau)}} \qquad (\forall \nu \in \mathcal{M}^m_{+1}(\Omega_0)), \tag{6.18}$$

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* $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}$ *is called a "normalized dual Bayes operator". Bayes operator is also called "Bayes-Kalman operator".*

■

We see

$$B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(g_\tau) \leq \Phi_{0,\tau} g_\tau \quad (\forall g_\tau \in C(\Omega_\tau) \text{ such that } g_\tau \geq 0), \tag{6.19}$$

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

$$\begin{aligned} B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau)) &= \widehat{F}_0((\prod_{t\in T}\Xi_t)\times\Gamma_\tau) \leq \widehat{F}_0((\prod_{t\in T}X_t)\times\Gamma_\tau) \\ &= \Phi_{0,\tau}G_\tau(\Gamma_\tau)\Big(=B^{(0,\tau)}_{\Pi_{t\in T}X_t}(G_\tau(\Gamma_\tau))\Big) \qquad (\forall \Gamma_\tau \in \mathcal{F}_\tau). \end{aligned} \tag{6.20}$$

The following theorem is essential to Regression Analysis II later.

Theorem 6.6. [The existence theorem of the Bayes operator (*cf.* [46, 55])]. *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 3.7 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* $\widetilde{\mathbf{O}}_s \equiv (\prod_{t\in T_s}X_t, 2^{\prod_{t\in T_s}X_t}, \widetilde{F}_s)$ *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 3.7). 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 [46]. The proof in [46] is essentially true, but it is not complete. That is because the definition of "Bayes operator" (i.e., Definition 6.5) was not mentioned in [46]. Thus, we add the complete proof in what follows. 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:

$$\widehat{B}^{(\tau,\tau)}_{\Pi_{t\in T_\tau}\Xi_t}(g_\tau) = \widetilde{F}_\tau(\Pi_{t\in T_\tau}\Xi_t) \times g_\tau \quad (\forall g_\tau \in C(\Omega_\tau)), \tag{6.21}$$

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

$$\widehat{F}_\tau(\Pi_{t\in T_\tau}\Xi_t \times \Gamma_\tau) = \widehat{B}^{(\tau,\tau)}_{\Pi_{t\in T_\tau}\Xi_t}(G_\tau(\Gamma_\tau)) \qquad (\forall \Gamma_\tau \in 2^{Y_\tau}), \tag{6.22}$$

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 $\widetilde{\mathbf{O}}_\tau$ $\Big(= (\Pi_{t\in T_\tau}X_t, 2^{\Pi_{t\in T_\tau}X_t}, \widetilde{F}_\tau)\Big)$, which is uniquely determined.

[Step 2] Let s be any element in $T \setminus \{0\}$ such that $s \le \tau$. 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 $\widetilde{\mathbf{O}}_s \Big(= (\Pi_{t\in T_s} X_t, 2^{\Pi_{t\in T_s} X_t}, \widetilde{F}_s)\Big)$. That is, there exists an observable $\widehat{\mathbf{O}}_s \equiv ((\prod_{t\in T_s} X_t) \times Y_\tau, 2^{(\Pi_{t\in T_s} X_t)\times Y_\tau}, \widehat{F}_s)$ in $C(\Omega_s)$ such that

(i) $\widehat{\mathbf{O}}_s$ is the Heisenberg picture representation (*cf.* Theorem 3.7) 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 $t \ne \tau$), $= \mathbf{O}_\tau \times \mathbf{O}'_\tau$ (if $t = \tau$),

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

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

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

$$[\widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\{x_t\}}(g_\tau)](\omega_{\pi(s)}) = \frac{[\widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\})](\omega_{\pi(s)}) \times [\Phi_{\pi(s),s}\widehat{B}^{(s,\tau)}_{\Pi_{t\in T_s}\{x_t\}}(g_\tau)](\omega_{\pi(s)})}{[\widetilde{F}_{\pi(s)}((\Pi_{t\in T_{\pi(s)}\setminus T_s} X_t) \times \prod_{t\in T_s}\{x_t\})](\omega_{\pi(s)})}$$
$$(\forall g_\tau \in C(\Omega_\tau), \quad \forall \omega_{\pi(s)}\ (\in \Omega_{\pi(s)})). \tag{6.23}$$

Here, the above is assumed to be equal to 0 if the denominator of (6.23) 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

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

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:

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

which is clearly the Heisenberg picture representation of $[\{\overline{\mathbf{O}}_t\}_{t\in T_{\pi(s)}}, \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T_{\pi(s)}\setminus\{\pi(s)\}}]$, where $\overline{\mathbf{O}}_t = \mathbf{O}_t$ (if $t \ne \tau$), $= \mathbf{O}_\tau \times \mathbf{O}'_\tau$ (if $t = \tau$). Also, it holds that

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

That is because we see

$$
\begin{aligned}
&\widehat{F}_{\pi(s)}((\Pi_{t\in T_{\pi(s)}}\Xi_t)\times Y_\tau) = \widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\Xi_t}(1_\tau) = \sum_{(x_t)_{t\in T_{\pi(s)}}\in \Pi_{t\in T_{\pi(s)}}\Xi_t} \widehat{B}^{(\pi(s),\tau)}_{\Pi_{t\in T_{\pi(s)}}\{x_t\}}(1_\tau) \\
&= \sum_{(x_t)_{t\in T_{\pi(s)}}\in \Pi_{t\in T_{\pi(s)}}\Xi_t} \frac{\widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\})\times \Phi_{\pi(s),s}\widehat{B}^{(s,\tau)}_{\Pi_{t\in T_s}\{x_t\}}(1_\tau)}{\widetilde{F}_{\pi(s)}((\Pi_{t\in T_{\pi(s)}\setminus T_s}X_t)\times \prod_{t\in T_s}\{x_t\})} \\
&= \sum_{(x_t)_{t\in T_{\pi(s)}}\in \Pi_{t\in T_{\pi(s)}}\Xi_t} \frac{\widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\})\times \widetilde{F}_{\pi(s)}((\Pi_{t\in T_{\pi(s)}\setminus T_s}X_t)\times \prod_{t\in T_s}\{x_t\})}{\widetilde{F}_{\pi(s)}((\Pi_{t\in T_{\pi(s)}\setminus T_s}X_t)\times \prod_{t\in T_s}\{x_t\})} \\
&= \sum_{(x_t)_{t\in T_{\pi(s)}}\in \Pi_{t\in T_{\pi(s)}}\Xi_t} \widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\{x_t\}) = \widetilde{F}_{\pi(s)}(\prod_{t\in T_{\pi(s)}}\Xi_t). \qquad (6.24)
\end{aligned}
$$

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

Let $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T}X_t, 2^{\Pi_{t\in T}X_t}, \widetilde{F}_0)$, $\mathbf{O}'_\tau \equiv (Y_\tau, 2^{Y_\tau}, G_\tau)$, $\{B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t\in 2^{X_t}\ (\forall t\in T)\}$, $\widehat{\mathbf{O}}_0 \equiv ((\prod_{t\in T}X_t)\times Y_\tau, 2^{(\Pi_{t\in T}X_t)\times Y_\tau}, \widehat{F}_0)$ 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 6.5. 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

$$
\frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widehat{F}_0((\prod_{t\in T}\Xi_t)\times\Gamma_\tau)\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widehat{F}_0((\prod_{t\in T}\Xi_t)\times Y_\tau)\rangle_{C(\Omega_0)}} \Big(\equiv \frac{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau))\rangle_{C(\Omega_0)}}{{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)\rangle_{C(\Omega_0)}}\Big). \qquad (6.25)
$$

A simple calculation shows:

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

Therefore, we say that

(C_4) the probability distribution of unknown y $\Big($under (C_2) (=(C_1))$\Big)$ is represented by

$$ {}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}), G_\tau(\Gamma_\tau)\rangle_{C(\Omega_\tau)}. \tag{6.26}$$

Let this (C_4) be, as an abbreviation, denoted (or, called) 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 $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0})$.

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 ρ^p is always fixed and never moves. In this sense, the ρ^p 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 6.7. [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 (after the measurement $\mathbf{M}_{C(\Omega_0)}$ ($\widehat{\mathbf{O}}_0$, $S_{[\delta_{\omega_0}]}$)) at τ ($\in T$) is equal to $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0})$". 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)⇒(C_5)$\big)$ as the following lemma.

Lemma 6.8. [S-state]. *Let $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0)$, $\{B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t}\ (\forall t\in T)\}$ 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 6.5. Assume that*

- *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

(♯) *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 $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0})$.*

■

The following lemma will be used as Theorem 6.13.

Lemma 6.9. [Inference and S-state]. *Let* $\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0)$, $\{B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t} \mid \Xi_t \in 2^{X_t} \ (\forall t \in T)\}$ *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 6.5. Assume that*

($\bullet$) *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

($\sharp$) *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* $R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0})$.

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

$$[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega_0) = \max_{\omega\in\Omega_0} [\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega). \tag{6.27}$$

Proof. The proof is similar to that of Corollary 5.6. Let $(Y_\tau, 2^{Y_\tau}, G_\tau)$ be any observable in $C(\Omega_\tau)$. Note that the above ($\bullet$) is the same as the following:

($\bullet$)' 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 6.5).

Thus we can infer, by Theorem 5.3 (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 (6.27)$\Big)$. Thus the conditional probability $P_{\Pi_{t\in T}\Xi_t}(\cdot)$ under the condition that we know that $((x_t)_{t\in T}, y) \in (\prod_{t\in T} X_t) \times Y_\tau$ is given by

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

From the equivalence of (C_4) and (C_5), we can conclude the ($\sharp$). □

Now we consider the simplest case 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 2.21). Then, we have the measurement

$$\mathbf{M}_{C(\Omega_0)}\Big(\widetilde{\mathbf{O}}_0 \equiv (\{0,1\}^2, 2^{\{0,1\}^2}, F^{(\mathrm{nl})}_0 \times \Phi_{0,\tau} F^{(\mathrm{nl})}_\tau), S_{[\delta_{\omega_0}]}\Big). \tag{6.28}$$

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

$$(\{0,1\}^2 \times Y_\tau, 2^{\{0,1\}^2 \times Y_\tau}, F_0^{(\mathrm{nl})} \times \Phi_{0,\tau} F_\tau^{(\mathrm{nl})} \times \Phi_{0,\tau} G_\tau) \tag{6.29}$$

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 2.21). Thus, the above (ii) implies that

(iii) the probability distribution of unknown y $\Big($under (ii) $\big(=$ (i)$\big)\Big)$, i.e., the probability that $y \in \Gamma_\tau$, is represented by

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

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

That is because it holds that

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

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 think that (iv) always holds.

From the above argument, we have the following lemma. This will be used in the statement (6.33).

Lemma 6.10. [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,* $\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}_0, S_{[\delta_{\omega_0}]}),$ $\Big)$ *is hidden behind the assertion* ($\sharp$).

■

Also, the following lemma is the formal representation of Corollary 5.6 (ii). $\Big($Cf. Remark 6.12.$\Big)$

Lemma 6.11. [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* $\mathbf{O}^{(\rm nl)}_\tau = (\{0,1\}, 2^{\{0,1\}}, F^{(\rm nl)}_\tau)$ *be the null observable in* $C(\Omega_\tau)$ *(cf. Example 2.21). Consider a measurement* $\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0(\equiv \mathbf{O}_0 \times \Phi_{0,\tau}\mathbf{O}^{(\rm nl)}_\tau), S_{[*]})$*, which is essentially the same as* $\mathbf{M}_{C(\Omega_0)}(\mathbf{O}_0, S_{[*]})$. *Assume that*

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

Then we see that

($\sharp$) *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} ($\in \mathcal{M}^p_{+1}(\Omega_0)$) *is defined by*

$$[F_0(\Xi_0)](\omega_0) = \max_{\omega \in \Omega_0} [F_0(\Xi_0)](\omega). \tag{6.30}$$

Proof. Let $B^{(0,\tau)}_{\Xi_0 \times \{1\}} : C(\Omega_\tau) \to C(\Omega_0)$ and $R^{(0,\tau)}_{\Xi_0} : \mathcal{M}^m_{+1}(\Omega_0) \to \mathcal{M}^m_{+1}(\Omega_\tau)$ be as in Definition 6.5. Here, note that, from the property of null observable, it holds that $F_0(\Xi_0) \times \Phi_{0,\tau}F^{(\rm nl)}_\tau(\{1\}) = F_0(\Xi_0)$. Thus we see that $B^{(0,\tau)}_{\Xi_0 \times \{1\}}(g_\tau) = F_0(\Xi_0) \times \Phi_{0,\tau}g_\tau$ for any g_τ ($\in C(\Omega_\tau)$). By Lemma 6.9, it suffices to prove $R^{(0,\tau)}_{\Xi_0}(\delta_{\omega_0}) = \Phi^*_{0,\tau}(\delta_{\omega_0})$. This is shown as follows:

$$\begin{aligned} &{}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Xi_0 \times \{1\}}(\delta_{\omega_0}), g_\tau \rangle_{C(\Omega_\tau)} = {}_{C(\Omega_\tau)^*}\langle \frac{(B^{(0,\tau)}_{\Xi_0 \times \{1\}})^*(\delta_{\omega_0})}{\|(B^{(0,\tau)}_{\Xi_0 \times \{1\}})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}}, g_\tau \rangle_{C(\Omega_\tau)} \\ =& \frac{1}{\|(B^{(0,\tau)}_{\Xi_0 \times \{1\}})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}} {}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Xi_0 \times \{1\}}(g_\tau) \rangle_{C(\Omega_0)} = \frac{[F_0(\Xi_0)](\omega_0) \times [\Phi_{0,\tau}g_\tau](\omega_0)}{[F_0(\Xi_0)](\omega_0)} \\ =& {}_{C(\Omega_\tau)^*}\langle \Phi^*_{0,\tau}(\delta_{\omega_0}), g_\tau \rangle_{C(\Omega_\tau)} \qquad (\forall g_\tau \in C(\Omega_\tau)). \end{aligned} \tag{6.31}$$

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

The following remark shows that Corollary 5.6 (ii) is a direct consequence of Lemma 6.11.

Remark 6.12. [Continued from Corollary 5.6 (Fisher's maximum likelihood method in classical measurements)]. As mentioned before, the proof of Corollary 5.6 is temporary. Corollary 5.6 should be understood as a corollary of Lemma 6.11 as follows: In Lemma 6.11, 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 5.6 is a consequence of Lemma 6.11. This should be regarded as the formal proof of Corollary 5.6.

■

6.3 Regression analysis II in measurements

Now let us explain the reason why we consider:

(♯) it is worthwhile doubting the derivation of (6.6) $\big(= (6.17)\big)$ from (6.5) $\big(= (6.16)\big)$, i.e., the formula $h(2) = 0.4 + 1.4 \times 2 = 3.2$.

Using the notations in Regression Analysis I (6.7), we recall the statement (R) of Example 6.4 as follows:

(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 (6.5) $\big(=(6.16)\big)$.

And we conclude, by Corollary 5.6 (or Remark 6.12), that

$$
\begin{aligned}
[\,*\,] = &\text{ “the S-state after the measurement } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} \\
= &\ \delta_{\omega_0}.
\end{aligned} \tag{6.32}
$$

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

- the map "$\delta_{\omega_0} \mapsto \Phi^*_{0,\tau}(\delta_{\omega_0})$" $\Big($i.e., the derivation of (6.6) $\big(= (6.17)\big)$ from (6.5) $\big(= (6.16)\big)\Big)$ 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 6.10. (6.33)

Thus, in order to conclude the assertion (6.6) $\big(= (6.17)\big)$, we need the above "two measurements", that is,

$$\text{“}\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\textstyle\prod_{t\in T} X_t, 2^{\prod_{t\in T} X_t}, \widetilde{F}_0), S_{[*]}) \text{ ” and “ } \mathbf{M}_{C(\Omega_0)}(\Phi_{0,\tau}\mathbf{O}'_\tau \equiv (Y_\tau, 2^{Y_\tau}, \Phi_{0,\tau}G_\tau), S_{[\delta_{\omega_0}]})\text{”}. \tag{6.34}$$

However, note that it is forbidden to conduct "two measurements" (*cf.* §2.5(II)). This is the reason that we think that it is worthwhile doubting (6.6) $\big(= (6.17)\big)$. In order to avoid this confusion, it suffices to consider the "simultaneous" measurement:

$$\mathbf{M}_{C(\Omega_0)}(\widehat{\mathbf{O}}_0 \equiv ((\prod_{t\in T} X_t) \times Y_\tau, 2^{(\prod_{t\in T} X_t)\times Y_\tau}, \widehat{F}_0), S_{[*]}), \quad \text{(where } \widehat{\mathbf{O}}_0 \text{ is as in Definition 6.5)}, \tag{6.35}$$

instead of (6.34).

Then, we rewrite Lemma 6.9 as an main theorem as follows:

Theorem 6.13. [= Lemma 6.9, 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 3.7 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 6.5. 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*

$$(\sharp) \quad \text{“the S-state at } \tau\ (\in T) \text{ after } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = R^{(0,\tau)}_{\prod_{t\in T}\Xi_t}(\delta_{\omega_0}). \tag{6.36}$$

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

$$[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega_0) = \max_{\omega\in\Omega_0}[\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega). \tag{6.37}$$

■

Lastly, we prove the following lemma, which justifies the inference (6.6).

Lemma 6.14. [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 3.7 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 6.5. 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:*

$$B^{(0,\tau)}_{\prod_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau)) = \widetilde{F}_0(\prod_{t\in T} \Xi_t) \times \Phi_{0,\tau}G_\tau(\Gamma_\tau) \quad (\forall \Xi_t \in 2^{X_t}\ (\forall t \in T), \forall \Gamma_\tau \in 2^{Y_\tau}), \tag{6.38}$$

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

$$0 \le g_\tau \le 1,\ g_\tau(\phi_{0,\tau}(\omega_0)) = 0,\ 0 \le h_\tau \le 1, \text{ and } h_\tau(\phi_{0,\tau}(\omega_0)) = 1. \tag{6.39}$$

where $\phi_{0,\tau} : \Omega_0 \to \Omega_\tau$ is defined by (3.14). Then we see, by (6.19), that

$$0 \le [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(g_\tau)](\omega) \le (\Phi_{0,\tau} g_\tau)(\omega) = g_\tau(\phi_{0,\tau}(\omega)) \quad (\forall \omega \in \Omega_0). \tag{6.40}$$

Putting $\omega = \omega_0$ in (6.40), we get, by (6.39), that

$$[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(g_\tau)](\omega_0) = 0. \tag{6.41}$$

Also, from the linearity of Bayes operator and the condition (iii) of Definition 6.5, we get

$$\begin{aligned}
[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau - h_\tau)](\omega) &= [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau)](\omega) - [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(h_\tau)](\omega) \\
&= [\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega) - [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(h_\tau)](\omega) \quad (\forall \omega \in \Omega_0).
\end{aligned} \tag{6.42}$$

Thus, putting $\omega = \omega_0$ in (6.42), we get, by (6.39), that

$$\begin{aligned}
0 &\le [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(1_\tau - h_\tau)](\omega_0) \\
&\le [(\Phi_{0,\tau}(1_\tau - h_\tau))](\omega_0) = 1_\tau(\phi_{0,\tau}(\omega_0)) - h_\tau(\phi_{0,\tau}(\omega_0)) = 1 - 1 = 0.
\end{aligned} \tag{6.43}$$

Then, we obtain

$$[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(h_\tau)](\omega_0) = [\widetilde{F}_0(\prod_{t\in T} \Xi_t)](\omega_0). \tag{6.44}$$

[Step 2]. Let ω_0 be any fixed 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:

$$\begin{aligned}
g_\tau(\omega_\tau) &= \max\{0, f(\omega_\tau) - f(\phi_{0,\tau}(\omega_0))\} \qquad (\forall \omega_\tau \in \Omega_\tau), \\
h_\tau(\omega_\tau) &= \min\{\frac{f(\omega_\tau)}{f(\phi_{0,\tau}(\omega_0))}, 1\} \qquad (\forall \omega_\tau \in \Omega_\tau).
\end{aligned} \tag{6.45}$$

The g_τ and the h_τ clearly satisfy (6.39). And moreover, we see, for any $\omega_\tau \in \Omega_\tau$, that

$$\begin{aligned}
& g_\tau(\omega_\tau) + f(\phi_{0,\tau}(\omega_0)) h_\tau(\omega_\tau) \\
&= \max\{0, f(\omega_\tau) - f(\phi_{0,\tau}(\omega_0))\} + \min\{f(\omega_\tau), f(\phi_{0,\tau}(\omega_0))\} \\
&= \begin{cases} (f(\omega_\tau) - f(\phi_{0,\tau}(\omega_0)) + f(\phi_{0,\tau}(\omega_0)), & \text{if } f(\omega_\tau) \ge f(\phi_{0,\tau}(\omega_0)) \\ 0 + f(\omega_\tau), & \text{if } f(\omega_\tau) \le f(\phi_{0,\tau}(\omega_0)) \end{cases} \\
&= f(\omega_\tau).
\end{aligned} \tag{6.46}$$

[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

$$
\begin{aligned}
&[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(G_\tau(\Gamma_\tau))](\omega) = \Big[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}\Big(\widehat{g}_\tau + [G_\tau(\Gamma_\tau)](\phi_{0,\tau}(\omega_0))\widehat{h}_\tau\Big)\Big](\omega) \\
=&[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{g}_\tau)](\omega) + [G_\tau(\Gamma_\tau)](\phi_{0,\tau}(\omega_0)) \times [B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{h}_\tau)](\omega) \quad (\forall \omega \in \Omega_0). \qquad (6.47)
\end{aligned}
$$

Putting $\omega = \omega_0$, we see, by (6.41) and (6.44), that $[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{g}_\tau)](\omega_0) = 0$ and $[B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\widehat{h}_\tau)](\omega_0)$ $= [\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0)$. And, we see, by (6.47), that

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

Since ω_0 $(\in \Omega_0)$ is arbitrary, we obtain (6.38). This completes the proof. □

Now we can propose our main assertion as follows:

REGRESSION ANALYSIS II [The new proposal of regression analysis, *cf.*[55]]. (6.48)

Let $(T \equiv \{0, 1, ..., N\}, \pi : T\backslash\{0\} \to T)$ *be a tree with root* 0*, and let* $\mathbf{S}_{[*]} \equiv [S_{[*]}; \{C(\Omega_t) \overset{\Phi_{\pi(t),t}}{\to} C(\Omega_{\pi(t)})\}_{t\in T\backslash\{0\}}]$ *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*

$$
\mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0 \equiv (\prod_{t\in T} X_t, 2^{\Pi_{t\in T}X_t}, \widetilde{F}_0), S_{[*]}) \qquad (\textit{cf. Theorem 3.7}). \qquad (6.49)
$$

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:*

$$
[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0) = \max_{\omega\in\Omega_0}[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega). \qquad (6.50)
$$

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 6.5.* $\Big($*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 6.6.*$\Big)$ *Then, we see:*

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

$$(\sharp) \quad \text{“The S-state at } \tau\ (\in T) \text{ after } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}). \tag{6.51}$$

Also

(ii). *[The S-state at τ ($\in T$) for 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 $0 < t \le \tau)$ is homomorphic$\Big)$. Then there is a reason to infer that*

$$\text{“the S-state at } \tau\ (\in T) \text{ after } \mathbf{M}_{C(\Omega_0)}(\widetilde{\mathbf{O}}_0, S_{[*]})\text{”} = \Phi^*_{0,\tau}(\delta_{\omega_0}). \tag{6.52}$$

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 (3.14).

Proof. (i). See Theorem 6.13 (= Lemma 6.9).

(ii). We see, by Lemma 6.14, that

$$
\begin{aligned}
&{}_{C(\Omega_\tau)^*}\langle R^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(\delta_{\omega_0}), g_\tau\rangle_{C(\Omega_\tau)} = {}_{C(\Omega_\tau)^*}\langle \frac{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})}{(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})}, g_\tau\rangle_{C(\Omega_\tau)} \\
=&\frac{1}{\|(B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t})^*(\delta_{\omega_0})\|_{\mathcal{M}(\Omega_\tau)}}\,{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, B^{(0,\tau)}_{\Pi_{t\in T}\Xi_t}(g_\tau)\rangle_{C(\Omega_0)} \\
=&\frac{1}{[\widetilde{F}_0(\prod_{t\in T}\Xi_t)](\omega_0)}\,{}_{C(\Omega_0)^*}\langle \delta_{\omega_0}, \widetilde{F}_0(\prod_{t\in T}\Xi_t) \times \Phi_{0,\tau} g_\tau\rangle_{C(\Omega_0)} \quad \text{(by Lemma 6.14)} \\
=&{}_{C(\Omega_\tau)^*}\langle \Phi^*_{0,\tau}(\delta_{\omega_0}), g_\tau\rangle_{C(\Omega_\tau)} \qquad (\forall g_\tau \in C(\Omega_\tau)).
\end{aligned}
$$

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

□

Remark 6.15. [(i) Continued from Example 6.2]. Note that our problem (i) in Example 6.2 was to infer the $h(2)$ and not (α_0, β_0). Regression analysis II (6.52) is applicable to our problem, that is, the above (6.52) says that there is a reason to calculate $h(2)$ in the following:

$$h(2) = \phi_{0,2}(0.4, 1.4) = 0.4 + 1.4 \times 2 = 3.2. \tag{6.53}$$

[(ii) 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 (6.5), 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 (6.5) in Example 6.2 is reasonable, or

(M_2) as the result in the case that $\tau \neq 0$, the conclusion (6.6) in Example 6.2 is reasonable.

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

(M_3) both (6.5) and (6.6) in Example 6.2 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 6.5 (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.* §2.5(II)). This is, of course, interesting, though it is not discussed in this book.

■

6.4 Conclusions

In this chapter we show that regression analysis can be completely understood in PMT as follows (*cf.* [55]):

measurement theory

$$
\Longrightarrow \begin{cases} \text{Axiom 1} \Rightarrow \begin{array}{c}\text{Theorem 5.3} \\ \text{(Fisher's method)}\end{array} \Rightarrow \begin{cases} \text{Corollary 5.5 (conditional probability)} \\ \text{Corollary 5.6 (classical Fisher's method)} \end{cases} \\ \\ \text{Axiom 2} \Rightarrow \begin{cases} \text{Theorem 3.7 (measurability)} \\ \text{Theorem 6.6 (the existence of Bayes operator)} \\ \text{Lemma 6.14 (some property of homomorphic relation).} \end{cases} \end{cases}
$$

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

(i) : "Corollary 5.6" + "Theorem 3.7" ⇒ "Regression Analysis I ",

$$
\text{(ii)}: \quad \text{"Corollary 5.5" + "Theorem 6.6"} \Rightarrow \left.\begin{array}{r} \text{Theorem 3.7} \\ \\ \text{"Theorem 6.13"} \\ \text{(Markov inference)} \\ \\ \text{"Lemma 6.14"} \end{array}\right\} \Rightarrow \text{"Regression Analysis II".}
$$

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 6.15.). It is surprising that both statistics and quantum mechanics can be understood in the same theory, i.e., measurement theory (6.1) (=(1.4)).

We believe that every statistician may want to know the justification of (6.5) and (6.6) in Example 6.2. Thus we expect that many statisticians will be interested in our axiomatic approach. That is because there is no justification without axioms.

We think that the results in Chapters 5 and 6 guarantee that "Fisher's statistics is theoretically true", (*cf.* Declaration (1.11)).