Chapter 5

Fisher's statistics I (under Axiom 1)

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

$$\underset{[\text{Axiom 1 (2.37)}]}{\text{PMT} = \text{measurement}} + \underset{[\text{Axiom 2 (3.26)}]}{\text{the relation among systems}} \quad \text{in } C^*\text{-algebra} . \tag{5.1}$$
(=(1.4))

In this chapter we intend to understand Fisher's statistics in Axiom 1. The reader will see that Fisher's maximum likelihood method is a direct consequence of Axiom 1.[1] And further, we discuss "inference interval" and "testing statistical hypothesis" in Axiom 1. By the results obtained in this chapter (and in the next chapter), we conclude that Fisher's statistics is theoretically true. (Cf. "Declaration (1.11)" in §1.4.)[2]

5.1 Introduction

The first attempt of the measurement theoretical approach to statistics was proposed in [44]. Although the argument in [44] is not deep, at least it convinces us of the good possibility of the axiomatic formulation (i.e., the measurement theoretical formulation) of statistics.

Most statisticians consider that statistics is closely related to "measurements", or, statistics is the study to analyze "measured data" for some purpose. Therefore, PMT should be immediately examined in comparison with statistics. The purpose of this chapter is to execute it, in other words, to propose a measurement theoretical formulation of statistics. We think that statistics is mainly related to the following aspect of measurement theory:

[1]Readers are not required to have much knowledge of statistics.

[2]We believe that the philosophy of statistics should be more discussed in statistics, (Cf. [61]). That is because it is indispensable for the understanding of "statistics (= mathematics + something)". It should be noted that "to formulate statistics in the framework of MT" implies "to introduce the philosophy of MT into statistics".

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

Let $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]}\big)$ be a measurement formulated in a C^*-algebra $\mathcal{A}$. Recall the (III) in §2.5 [Remarks], that is, the measurement $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]}\big)$ always determines the sample space $\Big(X, \overline{\mathcal{F}}, {}_{\mathcal{A}^*}\Big\langle \rho^p, F(\,\cdot\,)\Big\rangle_{\mathcal{A}}\Big)$. Here note that the mathematical structure of the sample space $\Big\{{}_{\mathcal{A}^*}\Big\langle \rho^p, F(\Xi)\Big\rangle_{\mathcal{A}}\Big\}_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*), \Xi \in \overline{\mathcal{F}}}$ is the same as that of the conventional formulation of statistics $\Big($i.e, $\Big\{P(\Xi, \theta)\Big\}_{\theta \in \Theta, \Xi \in \overline{\mathcal{F}}}$, where, for each θ in a parameter space Θ, $P(\cdot, \theta)$ is a probability measure on a measurable space $(X, \overline{\mathcal{F}})$, *cf.* [86]$\Big)$. Therefore, there is good hope that statistics can be described in terms of measurements. Also, this is precisely our motivation in this chapter. Following the common knowledge of quantum mechanics, we believe that any scientific statement including the term "probability" is not meaningful without the concept of "measurement". (*cf.* §2.5. Remarks). As mentioned in the above, the term "state" in measurement theory corresponds to the term "parameter" in statistics. The reason that we use the term "state" is due to that we want to stress that PMT is constructed modeled on mechanics.[3]

5.2 Fisher's maximum likelihood method

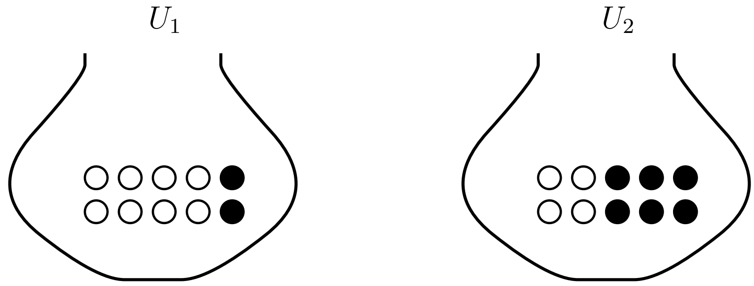
The purpose of this section is to study and understand "Fisher's maximum likelihood method" completely under Axiom 1 (of measurement theory). The following Problem 5.1 is the most typical in all examples of "Fisher's maximum likelihood method".

5.2.1 Fisher's maximum likelihood method

Problem 5.1. [The urn problem by Fisher's maximum likelihood method]. There are two urns U_1 and U_2. The urn U_1 [resp. U_2] contains 8 white and 2 black balls [resp. 4 white and 6 black balls].

[3]This means that we study statistics by an analogy of "mechanics". Note the following correspondence:

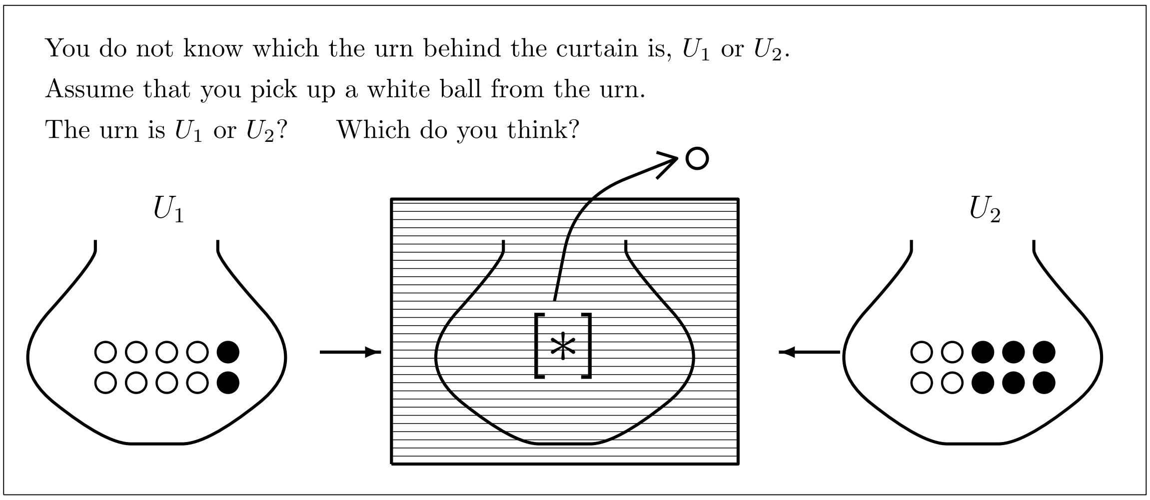
system $S_{[\rho^p]}$ (in PMT)	$\Longleftrightarrow$	population (in the conventional statistics)
[represented by pure state]		[represented by parameter]



Here consider the following procedures (P_1) and (P_2).

(P_1) One of the two (i.e., U_1 or U_2) is chosen and is settled behind a curtain. Note, for completeness, that you do not know whether it is U_1 or U_2.

(P_2) Pick up a ball out of the urn chosen by the procedure (P_1). And you find that the ball is white.



Now we have the following question:

(Q) Which is the chosen urn (behind the curtain), U_1 or U_2?

This is quite easy. That is, everyone will immediately infer "the urn behind the curtain = U_1". However, it is just "Fisher's maximum likelihood method". Cf. Example 5.8.

■

We begin with the following definition.

Notation 5.2. $[\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})]$. *Consider a measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho^p]})$ *formulated in a* C^**-algebra* $\mathcal{A}$. *In most measurements, it is usual to think 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, when we want to stress 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 state our present problem as follows:

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

In order to answer this problem, in [44] we introduced Fisher's method (precisely, Fisher's max mum likelihood method) as follows: $\Big($Strictly speaking, Theorem 5.3 should not be called "theorem" but "assertion", since it is not a purely mathematical result but a consequence of Axiom 1.$\Big)$

Theorem 5.3. [Fisher's maximum likelihood method in classical and quantum measurements, *cf.* [44]]. *Consider a measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[*]})$ *formulated in a* C^**-algebra* $\mathcal{A}$. *When we know that the measured value obtained by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$ *belongs to* Ξ $(\in \mathcal{F})$, *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:*

$$ {}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \rangle_{\mathcal{A}} = \max_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)} {}_{\mathcal{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathcal{A}}. \tag{5.2} $$

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 5.6 later.*$\Big)$ *Although the* ρ_0^p *in (5.2) is not generally determined uniquely, in this book we usually assume the uniqueness.*

Proof (*or, Explanation*). Let ρ_1^p and ρ_2^p be elements in $\mathfrak{S}^p(\mathcal{A}^*)$. Assume that "$[\,*\,] = \rho_1^p$ " or "$[\,*\,] = \rho_2^p$ ". And 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. Thus, there is a reason to regard the unknown state $[*]$ as the state ρ_2^p and not as the state ρ_1^p. Also, 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. □

Remark 5.4. [Radon-Nikodým derivative]. Assume that there exists a measure ν on

$(X, \overline{\mathcal{F}})$ (*cf.* (III) in §2.5) and $f(\cdot, \rho^p) \in L^1(\Omega, \nu)$ $(\forall \rho^p \in \mathfrak{S}^p(\mathcal{A}^*))$ such that:

$$\rho^p(F(\Xi)) = \int_\Xi f(x, \rho^p)\nu(dx) \quad (\forall \Xi \in \overline{\mathcal{F}}, \forall \rho^p \in \mathfrak{S}^p(\mathcal{A}^*)). \tag{5.3}$$

Then, even if $\Xi = \{x_0\}$ and $\rho^p(F(\{x_0\}) = 0$ $(\forall \rho^p \in \mathfrak{S}^p(\mathcal{A}^*)$ $)$ in Theorem 5.3, we may calculate as follows:

$$\frac{\rho_1^p(F(\{x_0\}))}{\rho_2^p(F(\{x_0\}))} = \lim_{\Xi \to \{x_0\}} \frac{\rho_1^p(F(\Xi))}{\rho_2^p(F(\Xi))} = \frac{f(x_0, \rho_1^p)}{f(x_0, \rho_2^p)}. \tag{5.4}$$

In this sense (or, in the sense of "Radon-Nikodým derivative"), we can compare $\rho_1^p(F(\{x_0\}))$ with $\rho_2^p(F(\{x_0\}))$, even when $\rho_1^p(F(\{x_0\})) = \rho_2^p(F(\{x_0\})) = 0$. When we know that the measured value x_0 $(\in X)$ is obtained by the measurement $\mathbf{M}_{\mathcal{A}}(\mathbf{O}, S_{[*]})$, by the same reason in Theorem 5.3, we can infer that the state $[*]$ of the system S is equal to ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^*))$ such that:

$$f(x_0, \rho_0^p) = \max_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)} f(x_0, \rho^p).$$

Here, the map $E : X \to \mathfrak{S}^p(\mathcal{A}^*)$, (i.e., $X \ni x_0 \mapsto \rho_0^p \in \mathfrak{S}^p(\mathcal{A}^*)$), is called "Fisher's estimator".

■

We begin with the following corollary, which is used in the proof of Corollary 5.6 and our main assertion $\Big($i.e., Regression Analysis II (in Chapter 6)$\Big)$.

Corollary 5.5. [The conditional probability representation of Fisher's method, *cf.* [55]]. *Let* $\mathbf{O} \equiv (X, \mathcal{F}, F)$ *and* $\mathbf{O}' \equiv (Y, \mathcal{G}, G)$ *be observables in* $\mathcal{A}$. *Let* $\widehat{\mathbf{O}}$ *be a quasi-product observable of* $\mathbf{O}$ *and* $\mathbf{O}'$, *that is,* $\widehat{\mathbf{O}} \equiv \mathbf{O} \overset{\text{qp}}{\times} \mathbf{O}' = (X \times Y, \mathcal{F} \bigotimes \mathcal{G}, F \overset{\text{qp}}{\times} G)$. *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* $\Xi \times Y$ $(\in \mathcal{F} \bigotimes \mathcal{G})$. *Then, there is a reason to infer that the unknown measured value* y $(\in Y)$ *is distributed under the conditional probability* $P_\Xi(\cdot)$, *where*

$$P_\Xi(\Gamma) = \frac{{}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \overset{\text{qp}}{\times} G(\Gamma) \rangle_{\mathcal{A}}}{{}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \rangle_{\mathcal{A}}} \left(= \frac{\rho_0^p(F(\Xi) \overset{\text{qp}}{\times} G(\Gamma))}{\rho_0^p(F(\Xi))} \right) \quad (\forall \Gamma \in \mathcal{G}), \tag{5.5}$$

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

$${}_{\mathcal{A}^*}\langle \rho_0^p, F(\Xi) \rangle_{\mathcal{A}} = \max_{\rho^p \in \mathfrak{S}^p(\mathcal{A}^*)} {}_{\mathcal{A}^*}\langle \rho^p, F(\Xi) \rangle_{\mathcal{A}}. \tag{5.6}$$

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 $\Xi \times Y$ $(\in \mathcal{F} \bigotimes \mathcal{G})$, we can infer, by Theorem 5.3

(Fisher's method) and the equality $F(\Xi) = F(\Xi) \overset{\mathrm{qp}}{\times} G(Y)$, that the $[\ast]$ (in $\mathbf{M}_{\mathcal{A}}(\widehat{\mathbf{O}}, S_{[\ast]})$) is equal to ρ_0^p $(\in \mathfrak{S}^p(\mathcal{A}^\ast))$. Thus, the conditional probability that $P_\Xi(\cdot)$ under the condition that we know that $(x, y) \in \Xi \times Y$ is given by

$$P_\Xi(\Gamma) = \frac{\rho_0^p(F(\Xi) \overset{\mathrm{qp}}{\times} G(\Gamma))}{\rho_0^p(F(\Xi) \overset{\mathrm{qp}}{\times} G(Y))} = \frac{\rho_0^p(F(\Xi) \overset{\mathrm{qp}}{\times} G(\Gamma))}{\rho_0^p(F(\Xi))}. \tag{5.7}$$

This completes the proof. □

The following corollary is the most essential in classical measurements. That is because what we want to infer is usually the state after the measurement (*cf.* Theorem 5.3).

Corollary 5.6. [Fisher's maximum likelihood method in classical measurements, *cf.* [55]]. *Let* $\mathbf{O} \equiv (X, \mathcal{F}, F)$ *be an observable in a commutative* $C^\ast$*-algebra* $C(\Omega)$*. Assume that we know that the measured value obtained by a measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\ast]})$ *belongs to* Ξ $(\in \mathcal{F})$*. Then, we can assert the following (i) and (ii):*

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

$$[F(\Xi)](\omega_0) = \max_{\omega \in \Omega} [F(\Xi)](\omega), \tag{5.8}$$

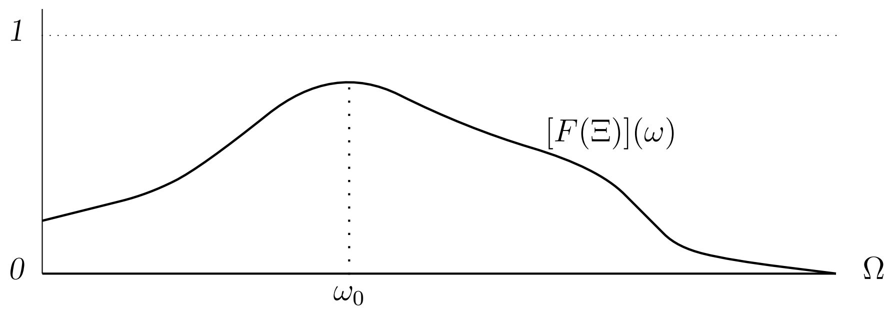
and,

(ii) *there is a reason to infer that the state after the measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\ast]})$ *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*

$$[\ast] = \text{"the state after the measurement } \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\ast]})\text{"} = \delta_{\omega_0}. \tag{5.9}$$



Proof. The (i) is the special case of Fisher's maximum likelihood method (*cf.* Theorem 5.3), i.e., $\mathcal{A} = C(\Omega)$. Thus it suffices to prove (ii) as follows: $\Big($This (ii) will be, under the definition of "S-state" (*cf.* Definition 6.7), proved in Remark 6.12 as a special case of Lemma 6.11 later. In this sense, the proof mentioned here is temporary.$\Big)$ Let $\mathbf{O}' \equiv (Y, \mathcal{G}, G)$ be any observable in $C(\Omega)$. Let $\widehat{\mathbf{O}}$ be the product observable of $\mathbf{O}$ and $\mathbf{O}'$, that is, $\widehat{\mathbf{O}} \equiv \mathbf{O} \times \mathbf{O}' = (X \times Y, \mathcal{F} \bigotimes \mathcal{G}, F \times G)$. Consider a measurement $\mathbf{M}_{C(\Omega)}(\widehat{\mathbf{O}} \equiv (X \times Y, \mathcal{F} \bigotimes \mathcal{G}, F \times G), S_{[*]})$. And assume

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

Corollary 5.5 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

$$P_\Xi(\Gamma)[F(\Xi)](\omega_0) = [G(\Gamma)](\omega_0) \quad (\forall \Gamma \in \mathcal{G}), \tag{5.10}$$

where ω_0 $(\in \Omega)$ is defined in (5.8). Also note that the above (A) can be represented by the following two steps (A_1) and (A_2) $\Big($i.e., $(A) = (A_1) + (A_2)\Big)$:

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

and

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

$\Big($The above is somewhat metaphorical since "two measurements" seem to appear (*cf.* §2.5[Remarks (II)]).$\Big)$ Comparing (A) and "$(A_1) + (A_2)$", we see, by (5.10), that

$$\text{"the probability that } y \text{ belongs to } \Gamma\ (\in \mathcal{G}) \text{ in } (A_2)\text{"} = [G(\Gamma)](\omega_0) \quad (\forall \Gamma \in \mathcal{G}) \tag{5.11}$$

That is, we get the sample space $(Y, \mathcal{G}, [G(\cdot)](\omega_0))$. Therefore, we say, from the arbitrariness of $\mathbf{O}' \equiv (Y, \mathcal{G}, G)$, that

(A_3) the state after the (A_1) (i.e., the state after the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$) is equal to δ_{ω_0}.

This completes the proof. $\Big($This corollary does not hold in quantum measurements, since the product observable $\widehat{\mathbf{O}} \equiv \mathbf{O} \times \mathbf{O}' = (X \times Y, \mathcal{F} \bigotimes \mathcal{G}, F \times G)$ 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 5.7)" is hidden in the above proof. This will be more clarified in Remark 6.12 later.

Remark 5.7. [Bayes operator]. Let $\mathbf{O} \equiv (X, \mathcal{F}, F)$ be an observable in $C(\Omega)$. For each Ξ ($\in \mathcal{F}$), define the continuous linear operator $B_\Xi^{(0,0)}$ (or, $B_\Xi^{\mathbf{O}}$, $B_\Xi^{\mathbf{O},(0,0)}$) $: C(\Omega) \to C(\Omega)$ such that:

$$B_\Xi^{(0,0)}(g)\Big(\equiv B_\Xi^{\mathbf{O}}(g) \equiv B_\Xi^{\mathbf{O},(0,0)}(g)\Big) = F(\Xi) \ \cdot \ g \qquad (\forall g \in C(\Omega)), \tag{5.12}$$

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

(i) for any observable $\mathbf{O}_1 \equiv (Y, \mathcal{G}, G)$, there exists an observable $\widehat{\mathbf{O}} \equiv (X \times Y, \mathcal{F} \bigotimes \mathcal{G}, \widehat{F})$ in $C(\Omega)$ such that:

$$\widehat{F}(\Xi \times \Gamma) = B_\Xi^{(0,0)}(G(\Gamma)) \qquad (\Xi \in \mathcal{F}, \Gamma \in \mathcal{G}).$$

That is because it suffices to define $\widehat{\mathbf{O}}$ by the product observable $\mathbf{O} \times \mathbf{O}_1$. Define the map $R_\Xi^{(0,0)} : \mathcal{M}_{+1}^m(\Omega) \to \mathcal{M}_{+1}^m(\Omega)$ (called *"normalized Bayes dual operator"*) such that:

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

where $[B_\Xi^{(0,0)}]^* : \mathcal{M}(\Omega) \to \mathcal{M}(\Omega)$ is the dual operator of $[B_\Xi^{(0,0)}]$, that is,

$$[R_\Xi^{(0,0)}(\nu)](D_0) = \frac{\int_{D_0} [F(\Xi)](\omega)\nu(d\omega)}{\int_\Omega [F(\Xi)](\omega)\nu(d\omega)} \qquad (\forall D_0 \in \mathcal{B}_\Omega). \tag{5.13}$$

Thus, we can describe the well known Bayes theorem (*cf.* [86]) such as

$$\mathcal{M}_{+1}^m(\Omega) \ni \nu \ (= \textit{pretest state}) \mapsto \ (\textit{posttest state} =) R_\Xi^{(0,0)}(\nu) \in \mathcal{M}_{+1}^m(\Omega)^4 \tag{5.14}$$

Note that this says that (i)⇒(ii) in Corollary 5.6. That is because a simple calculation shows that $R_\Xi^{(0,0)}(\delta_{\omega_0}) = \delta_{\omega_0}$ in the case of Corollary 5.6. In §6.2, the reader will again

[4] The pretest state [resp. posttest state] may be usually called "priori state" [resp. "posterior state"].

study the Bayes operator in more general situations.

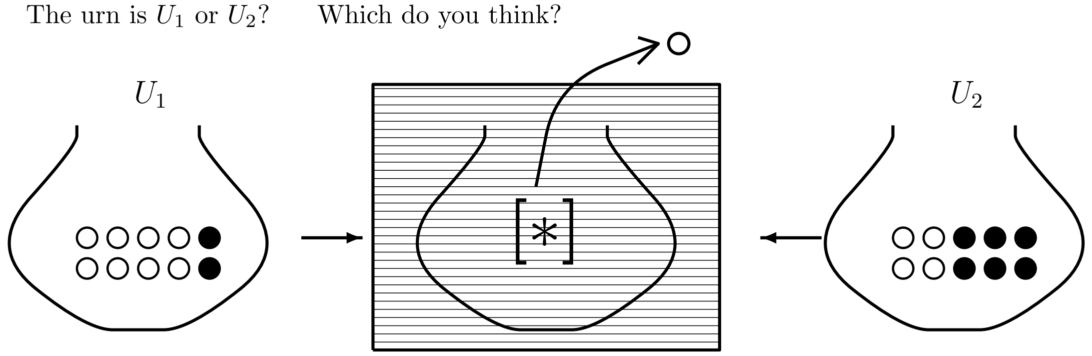
■

Example 5.8. [Continued from Problem 5.1 (Urn problem)[5]]. Recall Example 5.1. That is, consider the following procedures (P_1) and (P_2).

(P_1) One of the two (i.e., U_1 or U_2) is chosen and is settled behind a curtain. Note, for completeness, that you do not know whether it is U_1 or U_2.[6]

(P_2) Pick up a ball out of the urn chosen by the procedure (P_1). And you find that the ball is white.

You do not know which the urn behind the curtain is, U_1 or U_2.
Assume that you pick up a white ball from the urn.
The urn is U_1 or U_2? Which do you think?



Now we have the following question:

(Q) Which is the chosen urn (behind the curtain), U_1 or U_2?

[*Answer*]. Put $\Omega = \{\omega_1, \omega_2\}$. Here,

$$\begin{cases} \omega_1 & \cdots\cdots \text{the state that the urn } U_1 \text{ is behind the curtain} \\ \omega_2 & \cdots\cdots \text{the state that the urn } U_2 \text{ is behind the curtain.} \end{cases} \tag{5.15}$$

In this sense, we frequently use the following identification:

$$U_1 \approx \omega_1, \qquad U_2 \approx \omega_2. \tag{5.16}$$

[5]As mentioned in Example 2.16, we believe that "urn problem" is the most fundamental in all examples of statistics.

[6]Here we are not concerned with SMT_{PEP} (i.e., the principle of equal probability, *cf.* §11.4)

And define the observable $\mathbf{O}\big(\equiv (X \equiv \{w,b\}, 2^{\{w,b\}}, F)\big)$ in $C(\Omega)$ where

$$[F(\{w\})](\omega_1) = 0.8, \qquad [F(\{b\})](\omega_1) = 0.2,$$
$$[F(\{w\})](\omega_2) = 0.4, \qquad [F(\{b\})](\omega_2) = 0.6.$$

Since we do not know whether the state is ω_1 or ω_2, we have the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$. Thus, out situation is

- a measured value "w" is obtained by the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$.

Then, we conclude, by Fisher's maximum likelihood method, that

- the urn behind the curtain is U_1.

That is because

$$[F(\{w\})](\omega_1) = 0.8 = \max\{[F(\{w\})](\omega_1), [F(\{w\})](\omega_2)\}.$$

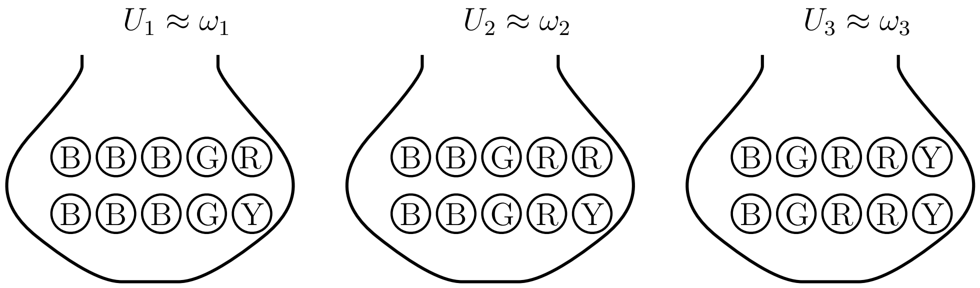
■

Example 5.9. [Urn problem]. Let U_j, $j = 1,2,3$, be urns that contain sufficiently many colored balls as follows:

	blue balls	green balls	red balls	yellow balls
urn U_1	60%	20%	10%	10%
urn U_2	40%	20%	30%	10%
urn U_3	20%	20%	40%	20%

(5.17)

Put $\mathbf{U} = \{U_1, U_2, U_3\}$. We consider the state space $\Omega \ \big(\equiv \{\omega_1, \omega_2, \omega_3\} \ \big)$ with the discrete topology, which is identified with $\mathbf{U}$, that is, $\mathbf{U} \ni U_j \leftrightarrow \omega_j \in \Omega \approx \mathcal{M}^p_{+1}(\Omega)$.[7]



[7]Strictly speaking, we must consider the identification as (5.15).

Define the observable $\mathbf{O} \equiv (X = \{b, g, r, y\}, \mathcal{P}(X), F_{(\cdot)})$ in $C(\Omega)$ by the usual way. That is,

$$\begin{array}{llll} F_{\{b\}}(\omega_1) = 6/10 & F_{\{g\}}(\omega_1) = 2/10 & F_{\{r\}}(\omega_1) = 1/10 & F_{\{y\}}(\omega_1) = 1/10 \\ F_{\{b\}}(\omega_2) = 4/10 & F_{\{g\}}(\omega_2) = 2/10 & F_{\{r\}}(\omega_2) = 3/10 & F_{\{y\}}(\omega_2) = 1/10 \\ F_{\{b\}}(\omega_3) = 2/10 & F_{\{g\}}(\omega_3) = 2/10 & F_{\{r\}}(\omega_3) = 4/10 & F_{\{y\}}(\omega_3) = 2/10. \end{array} \tag{5.18}$$

Then we have the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$.

[I] Now we consider the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$. And assume that we get the measured value 'b' by the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$. Then Fisher's maximum likelihood method (i.e., Corollary 5.6) says that there is a reason to infer that

$$[*] = \omega_1$$

since

$$F_{\{b\}}(\omega_1) = 0.6 = \max_{\omega \in \Omega} F_{\{b\}}(\omega) = \max\{0.6,\ 0.4,\ 0.2\}.$$

That is, the unknown urn $[*]$ is U_1.

[II] Also, consider the (iterated) measurement $\mathbf{M}_{C(\Omega)}(\times_{k=1}^2 \mathbf{O} \equiv (X^2, \mathcal{P}(X^2), \times_{k=1}^2 F)$, $S_{[*]})$ where $(\times_{k=1}^2 F)_{\Xi_1 \times \Xi_2}(\omega) = F_{\Xi_1}(\omega) \cdot F_{\Xi_2}(\omega)$. Also, assume that

- the measured value (b, r) is obtained by the iterated measurement $\mathbf{M}_{C(\Omega)}(\times_{k=1}^2 \mathbf{O}, S_{[*]})$.

Applying Fisher's method (= Corollary 5.6), we get the conclusion as follows: Put

$$E(\omega) = F_{\{b\}}(\omega) F_{\{r\}}(\omega).$$

Clearly it holds that $E(\omega_1) = 6 \cdot 1/10^2 = 0.06$, $E(\omega_2) = 4 \cdot 3/10^2 = 0.12$ and $E(\omega_3) = 2 \cdot 4/10^2 = 0.08$. Therefore, there is a very reason to think that $[\,*\,] = \delta_{\omega_2}$, that is, the unknown urn $[*]$ is U_2.

[III; Remark (moment method)]. Here, let us consider the above [II] by the moment method (*cf.* Definition 2.27). Define the distance Δ on $\mathcal{M}_{+1}^m(X)$ such that:

$$\begin{aligned} \Delta(\nu_1, \nu_2) =& \sum_{x \in X \equiv \{b,g,r,y\}} |\nu_1(\{x\}) - \nu_2(\{x\})| \\ =& |\nu_1(\{b\}) - \nu_2(\{b\})| + |\nu_1(\{g\}) - \nu_2(\{g\})| + |\nu_1(\{r\}) - \nu_2(\{r\})| + |\nu_1(\{y\}) - \nu_2(\{y\})|. \end{aligned}$$

Note that ${}_{\mathcal{M}(\Omega)}\big\langle \delta_{\omega_1}, F_{\{b\}} \big\rangle_{C(\Omega)} = \delta_{\omega_1}(F_{\{b\}}) = F_{\{b\}}(\omega_1) = 6/10$, and similarly (*cf.* (5.18)),

$$
\begin{array}{llll}
\delta_{\omega_1}(F_{\{b\}}) = 6/10 & \delta_{\omega_1}(F_{\{g\}}) = 2/10 & \delta_{\omega_1}(F_{\{r\}}) = 1/10 & \delta_{\omega_1}(F_{\{y\}}) = 1/10 \\
\delta_{\omega_2}(F_{\{b\}}) = 4/10 & \delta_{\omega_2}(F_{\{g\}}) = 2/10 & \delta_{\omega_2}(F_{\{r\}}) = 3/10 & \delta_{\omega_2}(F_{\{y\}}) = 1/10 \\
\delta_{\omega_3}(F_{\{b\}}) = 2/10 & \delta_{\omega_3}(F_{\{g\}}) = 2/10 & \delta_{\omega_3}(F_{\{r\}}) = 4/10 & \delta_{\omega_3}(F_{\{y\}}) = 2/10.
\end{array}
$$

Since the measured value (b, r) is obtained, we have the sample space $(X, 2^X, \nu)$ such that

$$
\nu(\{b\}) = 1/2, \qquad \nu(\{g\}) = 0, \qquad \nu(\{r\}) = 1/2, \qquad \nu(\{y\}) = 0.
$$

Then, we see that

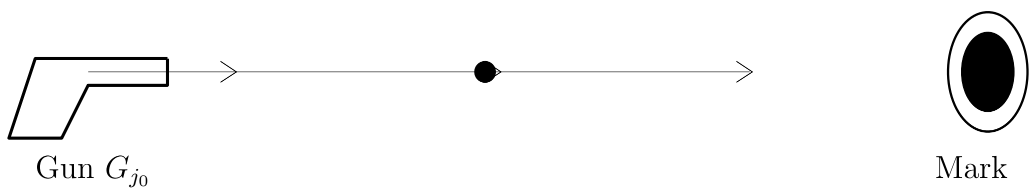
$$
\begin{aligned}
\Delta(\delta_{\omega_1}(F_{\{\cdot\}}), \nu) &= |6/10 - 1/2| + |2/10 - 0| + |1/10 - 1/2| + |1/10 - 0| = 8/10 \\
\Delta(\delta_{\omega_2}(F_{\{\cdot\}}), \nu) &= |4/10 - 1/2| + |2/10 - 0| + |3/10 - 1/2| + |1/10 - 0| = 6/10 \\
\Delta(\delta_{\omega_3}(F_{\{\cdot\}}), \nu) &= |2/10 - 1/2| + |2/10 - 0| + |4/10 - 1/2| + |2/10 - 0| = 8/10.
\end{aligned}
$$

Thus, the moment method says that the unknown urn [$*$] is U_2.

■

Example 5.10. [At a gun shop, [44]]. Let $\mathbf{G} \equiv \{G_1, ..., G_{50}\}$ be a set of guns in a gun shop. Assume that

$$
\text{the percentage of "hits of a gun } G_j\text{"} = \begin{cases} 80\% & \text{if} \quad 1 \le j \le 30, \\ 70\% & \text{if} \quad 31 \le j \le 40, \\ 10\% & \text{if} \quad 41 \le j \le 50. \end{cases} \tag{5.19}
$$



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

(i) Some one picks up a certain gun G_{j_0} from $\mathbf{G}$. He does not know the information concerning the j_0.

(ii) He shoots the gun G_{j_0} three times. First and second he hits the mark, and third he misses the mark.

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

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

$$F_{\{1\}}(\omega_j) = \begin{cases} 0.8 & \text{if} \quad 1 \le j \le 30, \\ 0.7 & \text{if} \quad 31 \le j \le 40, \\ 0.1 & \text{if} \quad 41 \le j \le 50 \end{cases} \tag{5.20}$$

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

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

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

$$(\mathop{\times}_{k=1}^{3} F)_{\Xi_1 \times \Xi_2 \times \Xi_3}(\omega) = F_{\Xi_1}(\omega) F_{\Xi_2}(\omega) F_{\Xi_3}(\omega) \qquad (\forall \Xi_1 \times \Xi_2 \times \Xi_3 \in \mathcal{P}(X^3), \forall \omega \in \Omega). \tag{5.21}$$

Clearly, the above statement (ii) implies that the measured value $(1,1,0)$ is obtained by $\mathbf{M}_{C(\Omega)}\big(\times_{k=1}^3 \mathbf{O}, S_{[*]}\big)$. $\big($ The observer does not know that $[*] = \delta_{\omega_{j_0}}$. $\big)$ By a simple calculation, we see

$$F_{\{1\}}(\omega_j) F_{\{1\}}(\omega_j) F_{\{0\}}(\omega_j) = \begin{cases} 0.128 & \text{if} \quad 1 \le j \le 30, \\ 0.147 & \text{if} \quad 31 \le j \le 40, \\ 0.009 & \text{if} \quad 41 \le j \le 50. \end{cases} \tag{5.22}$$

Therefore, by Fisher's method (= Corollary 5.6), there is a very reason to consider that $31 \le j_0 \le 40$.

■

Example 5.11. [(i): Gaussian observable]. Consider a commutative C^*-algebra $C_0(\mathbf{R})$. And define the Gaussian observable $\mathbf{O}_{\sigma^2} \equiv (\mathbf{R}, \mathcal{B}_{\mathbf{R}}^{\mathrm{bd}}, F_{(\cdot)}^{\sigma^2})$ in $C_0(\mathbf{R})$ such that:

$$F_{\Xi}^{\sigma^2}(\mu) = \frac{1}{\sqrt{2\pi}\sigma} \int_{\Xi} \exp[-\frac{1}{2\sigma^2}(x-\mu)^2] dx \qquad (\forall \Xi \in \mathcal{B}_{\mathbf{R}}^{\mathrm{bd}}, \quad \forall \mu \in \mathbf{R}). \tag{5.23}$$

Further, consider the product observable $\times_{k=1}^3 \mathbf{O}$ (or in short, $\mathbf{O}_{\sigma^2}^3$) $\equiv (\mathbf{R}^3, \mathcal{B}_{\mathbf{R}^3}^{\mathrm{bd}}, F_{(\cdot)}^{\sigma^2,3})$ in $C_0(\mathbf{R})$ such that:

$$
\begin{aligned}
&F^{\sigma^2,3}_{\Xi_1\times\Xi_2\times\Xi_3}(\mu) = F^{\sigma^2}_{\Xi_1}(\mu)\cdot F^{\sigma^2}_{\Xi_2}(\mu)\cdot F^{\sigma^2}_{\Xi_3}(\mu)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^3}\int_{\Xi_1\times\Xi_2\times\Xi_3}\exp[-\frac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2}{2\sigma^2}]dx_1dx_2dx_3\\
&\qquad(\forall\Xi = k\in\mathcal{B}^{\rm bd}_{\mathbf{R}}, k=1,2,3, \quad \forall\mu\in\mathbf{R}).
\end{aligned}
\tag{5.24}
$$

Here consider the measurement $\mathbf{M}_{C_0(\mathbf{R})}(\mathbf{O}^3_{\sigma^2}, S_{[*]})$. And assume that

- the measured value (x^0_1, x^0_2, x^0_3) ($\in\mathbf{R}^3$) is obtained by the $\mathbf{M}_{C_0(\mathbf{R})}(\mathbf{O}^3_{\sigma^2}, S_{[*]})$.

Then, Fisher's method (=Corollary 5.6) and Remark 5.4 say that there is a reason to think that the unknown state $[\,*\,]=\mu_0$, where

$$
\begin{aligned}
&\frac{1}{(\sqrt{2\pi}\sigma)^3}\exp[-\frac{(x^0_1-\mu_0)^2+(x^0_2-\mu_0)^2+(x^0_3-\mu_0)^2}{2\sigma^2}]\\
=&\max_{\mu\in\mathbf{R}}\Big[\frac{1}{(\sqrt{2\pi}\sigma)^3}\exp[-\frac{(x^0_1-\mu)^2+(x^0_2-\mu)^2+(x^0_3-\mu)^2}{2\sigma^2}]\Big],
\end{aligned}
\tag{5.25}
$$

which is equivalent to

$$
\begin{aligned}
&(x^0_1-\mu_0)^2+(x^0_2-\mu_0)^2+(x^0_3-\mu_0)^2\\
=&\min_{\mu\in\mathbf{R}}[(x^0_1-\mu)^2+(x^0_2-\mu)^2+(x^0_3-\mu)^2]
\end{aligned}
\tag{5.26}
$$

and moreover, equivalently,

$$
\mu_0 = (x^0_1+x^0_2+x^0_3)/3. \tag{5.27}
$$

[(ii): Gaussian observable]. Consider a commutative C^*-algebra $C([0,100])$, where $[0,100]$ $\equiv\{\mu\in\mathbf{R}\mid 0\le\mu\le 100\}$. And define the Gaussian observable $\mathbf{O}_{\sigma^2}\equiv(\mathbf{R},\mathcal{B}^{\rm bd}_{\mathbf{R}},F^{\sigma^2}_{(\cdot)})$ in $C([0,100])$ such that:

$$
F^{\sigma^2}_{\Xi}(\mu)=\frac{1}{\sqrt{2\pi}\sigma}\int_{\Xi}\exp[-\frac{1}{2\sigma^2}(x-\mu)^2]dx \qquad (\forall\Xi\in\mathcal{B}^{\rm bd}_{\mathbf{R}}, \quad \forall\mu\in[0,100]). \tag{5.28}
$$

Further, consider the product observable $\mathbf{O}^3_{\sigma^2}\equiv(\mathbf{R}^3,\mathcal{B}^{\rm bd}_{\mathbf{R}^3},F^{\sigma^2,3}_{(\cdot)})$ in $C_0([0,100])$ such that:

$$
\begin{aligned}
&F^{\sigma^2,3}_{\Xi_1\times\Xi_2\times\Xi_3}(\mu)\\
=&\frac{1}{(\sqrt{2\pi}\sigma)^3}\int_{\Xi_1\times\Xi_2\times\Xi_3}\exp[-\frac{(x_1-\mu)^2+(x_2-\mu)^2+(x_3-\mu)^2}{2\sigma^2}]dx_1dx_2dx_3\\
&\qquad(\forall\Xi = k\in\mathcal{B}^{\rm bd}_{\mathbf{R}}, k=1,2,3, \quad \forall\mu\in[0,100]).
\end{aligned}
\tag{5.29}
$$

Here consider the measurement $\mathbf{M}_{C([0,100])}(\mathbf{O}^3_{\sigma^2}, S_{[*]})$. And assume that

- the measured value (x_1^0, x_2^0, x_3^0) $(\in \mathbf{R}^3)$ is obtained by the $\mathbf{M}_{C([0,100])}(\mathbf{O}^3_{\sigma^2}, S_{[*]})$

Then, Fisher's method and Remark 5.4 say that there is a reason to think that the unknown state $[\,*\,] = \mu_0$, where $[\,*\,] = \mu_0$, where

$$
\begin{aligned}
&\frac{1}{(\sqrt{2\pi}\sigma)^3} \exp[-\frac{(x_1^0-\mu_0)^2+(x_2^0-\mu_0)^2+(x_3^0-\mu_0)^2}{2\sigma^2}] \\
=& \max_{\mu\in[0,100]} \Big[\frac{1}{(\sqrt{2\pi}\sigma)^3} \exp[-\frac{(x_1^0-\mu)^2+(x_2^0-\mu)^2+(x_3^0-\mu)^2}{2\sigma^2}]\Big]
\end{aligned} \tag{5.30}
$$

which is equivalent to

$$
\begin{aligned}
&(x_1^0-\mu_0)^2+(x_2^0-\mu_0)^2+(x_3^0-\mu_0)^2 \\
=& \min_{\mu\in[0,100]} [(x_1^0-\mu)^2+(x_2^0-\mu)^2+(x_3^0-\mu)^2]
\end{aligned} \tag{5.31}
$$

and moreover, equivalently,

$$
\mu_0 = \begin{cases} 0 & \text{if } x_1^0+x_2^0+x_3^0 < 0 \\ (x_1^0+x_2^0+x_3^0)/3 & \text{if } 0 \le x_1^0+x_2^0+x_3^0 \le 100 \\ 100 & \text{if } x_1^0+x_2^0+x_3^0 > 100. \end{cases} \tag{5.32}
$$

■

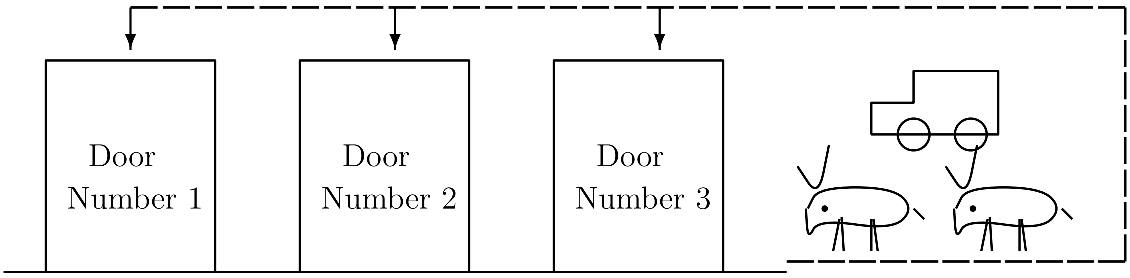
5.2.2 Monty Hall problem in PMT

Problem 5.12. [Monty Hall problem, *cf.*[33]].

The Monty Hall problem is as follows:

(P) Suppose you are on a game show, and you are given the choice of three doors (i.e., "number 1", "number 2", "number 3"). Behind one door is a car, behind the others, goats.

You pick a door, say number 1, and the host, who knows what's behind the doors, opens another door, say "number 3", which has a goat. He says to you, "Do you want to pick door number 2?" Is it to your advantage to switch your choice of doors?



[*Answer*]. Put $\Omega = \{\omega_1, \omega_2, \omega_3\}$, where

$\omega_1 \cdots\cdots$ the state that the car is behind the door number 1

$\omega_2 \cdots\cdots$ the state that the car is behind the door number 2

$\omega_3 \cdots\cdots$ the state that the car is behind the door number 3.

Define the observable $\mathbf{O} \equiv (\{1,2,3\}, 2^{\{1,2,3\}}, F)$ in $C(\Omega)$ such that

$$
\begin{array}{lll}
[F(\{1\})](\omega_1) = 0.0, & [F(\{2\})](\omega_1) = 0.5, & [F(\{3\})](\omega_1) = 0.5, \\
[F(\{1\})](\omega_2) = 0.0, & [F(\{2\})](\omega_2) = 0.0, & [F(\{3\})](\omega_2) = 1.0, \\
[F(\{1\})](\omega_3) = 0.0, & [F(\{2\})](\omega_3) = 1.0, & [F(\{3\})](\omega_3) = 0.0.
\end{array}
$$ [8]

Thus we have a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$. Here, note that

(1) :"measured value 1 is obtained" $\Longleftrightarrow$ The host says "Door (number 1) has a goat",

(2) :"measured value 2 is obtained" $\Longleftrightarrow$ The host says "Door (number 2) has a goat",

(3) :"measured value 3 is obtained" $\Longleftrightarrow$ The host says "Door (number 3) has a goat".

The host said "Door (number 3) has a goat". This implies that you get the measured value "3" by the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$. Therefore, Fisher's maximum likelihood method says that *you should pick door number 2.* That is because we see that

$$
\begin{aligned}
[F(\{3\})](\omega_2) = 1.0 &= \max\{0.5, \ \ 1.0, \ \ 0.0\} \\
&= \max\{[F(\{3\})](\omega_1), [F(\{3\})](\omega_2), [F(\{3\})](\omega_3)\},
\end{aligned}
$$

and thus, $[*] = \delta_{\omega_2}$. However, this is not all of the Monty Hall problem. See Remark 5.13, Problem 8.8 and Problem 11.13 later.

■

[8]Strictly speaking, $F(\{1\})(\omega_1) = 0.5$ and $F(\{2\})(\omega_1) = 0.5$ should be assumed in the problem (P).

Remark 5.13. [Monty Hall problem by the moment method (*cf.* Definition 2.27)]. Here, consider Problem 5.12 by the moment method. Since you get measured value 3, you get the sample space $(\{1,2,3\}, 2^{\{1,2,3\}}, \nu_s)$ such that $\nu_s(\{1\}) = 0$, $\nu_s(\{2\}) = 0$ and $\nu_s(\{3\}) = 1$. For example define the distance Δ such that: for any $\nu_1, \nu_2 \in \mathcal{M}_{+1}^m(\{1,2,3\})$,

$$\Delta(\nu_1,\nu_2) = |\nu_1(\{1\}) - \nu_2(\{1\})| + |\nu_1(\{2\}) - \nu_2(\{2\})| + |\nu_1(\{3\}) - \nu_2(\{3\})|.$$

Then, we see

$$\Delta(\nu_s, [F(\,\cdot\,)](\omega_1)) = |0-0| + |0-0.5| + |1-0.5| = 1,$$
$$\Delta(\nu_s, [F(\,\cdot\,)](\omega_2)) = |0-0| + |0-0| + |1-1| = 0$$

and

$$\Delta(\nu_s, [F(\,\cdot\,)](\omega_3)) = |0-0| + |0-1| + |1-0| = 2.$$

Thus, we can, by the moment method, infer that ω_2 is most possible, that is, the car is behind the door number 2.

■

5.3 Inference interval

Let $\mathbf{O}(\,\equiv (X, \mathcal{F}, F))$ be an observable formulated in a C^*-algebra $\mathcal{A}$. Assume that X has a metric d_X. And assume that the state space $\mathfrak{S}^p(\mathcal{A}^*)$ has the metric $d_{\mathfrak{S}}$, which induces the weak* topology $\sigma(\mathcal{A}^*, \mathcal{A})$. Let $E : X \to \mathfrak{S}^p(\mathcal{A}^*)$ be a continuous map, which is called *"estimator".* Let γ be a real number such that $0 \ll \gamma < 1$, for example, $\gamma = 0.95$. For any $\rho^p(\,\in \mathfrak{S}^p(\mathcal{A}^*))$, define the positive number $\eta^\gamma_{\rho^p}$ $(\,> 0)$ such that:

$$\eta^\gamma_{\rho^p} = \inf\{\eta > 0 : {}_{\mathcal{A}^*}\Big\langle \rho^p, F(E^{-1}(B(\rho^p;\eta)))\Big\rangle_{\mathcal{A}} \geq \gamma\} \tag{5.33}$$

where $B(\rho^p;\eta) = \{\rho^p_1(\,\in \mathfrak{S}^p(\mathcal{A}^*)) : d_{\mathfrak{S}}(\rho^p_1, \rho^p) \leq \eta\}$. For any x $(\,\in X)$, put

$$D^\gamma_x = \{\rho^p(\in \mathfrak{S}^p(\mathcal{A}^*)) : d_{\mathfrak{S}}(E(x), \rho^p) \leq \eta^\gamma_{\rho^p}\}. \tag{5.34}$$

The D^γ_x is called *the (γ)-inference interval of the measured value x.*

Note that,

(A) *for any* $\rho_0^p (\in \mathfrak{S}^p(\mathcal{A}^*))$*, the probability, that the measured value* x *obtained by the measurement* $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho_0^p]}\big)$ *satisfies the following condition* ($\flat$)*, is larger than* γ *(e.g.,* $\gamma = 0.95$*).*

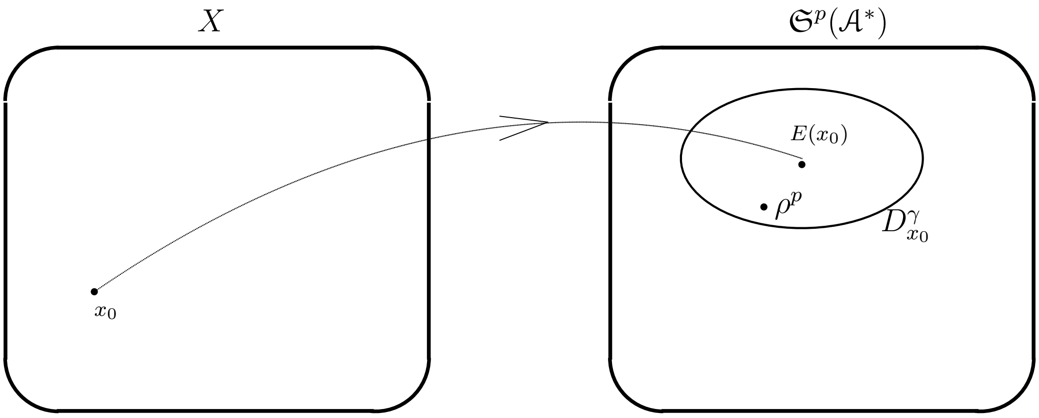
$$(\flat) \qquad E(x) \in B(\rho_0^p; \eta^\gamma_{\rho_0^p}) \qquad \text{or equivalently} \qquad d(E(x), \rho_0^p) \leq \eta^\gamma_{\rho_0^p}.$$

Assume that

(B) we get a measured value x_0 by the measurement $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho_0^p]}\big)$.

Then, we see the following equivalences:

$$(\flat) \quad \Longleftrightarrow \quad d_{\mathfrak{S}}(E(x_0), \rho_0^p) \leq \eta^\gamma_{\rho_0^p} \quad \Longleftrightarrow \quad D^\gamma_{x_0} \ni \rho_0^p.$$



Summing the above argument, we have the following theorem.

Theorem 5.14. [Inference interval]. *Let* $\mathbf{O} \equiv (X, \mathcal{F}, F)$ *be an observable in* $\mathcal{A}$*. Let* ρ_0^p *be any fixed state, i.e.,* $\rho_0^p \in \mathfrak{S}^p(\mathcal{A}^*)$*, Consider a measurement* $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho_0^p]}\big)$*. Let* $E : X \to \mathfrak{S}^p(\mathcal{A}^*)$ *be an estimator. Let* γ *be such as* $0 \ll \gamma < 1$ *(e.g.,* $\gamma = 0.95$*). For any* $x(\in X)$*, define* D^γ_x *as in (5.34). Then, we see,*

($\sharp$) *the probability that the measured value* $x_0(\in X)$ *obtained by the measurement* $\mathbf{M}_{\mathcal{A}}\big(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\rho_0^p]}\big)$ *satisfies the condition that*

$$D^\gamma_{x_0} \ni \rho_0^p \ , \tag{5.35}$$

is larger than γ*.*

■

Example 5.15. [The urn problem]. Put $\Omega = [0,1]$, i.e., the closed interval in $\mathbf{R}$. We assume that each ω ($\in \Omega \equiv [0,1]$) represents an urn that contains a lot of red balls and white balls such that:

$$\frac{\text{the number of white balls in the urn } \omega}{\text{the total number of red and white balls in the urn } \omega} \approx \omega \quad (\forall \omega \in [0,1] \equiv \Omega). \tag{5.36}$$

Define the observable $\mathbf{O} = (X \equiv \{r, w\}, 2^{\{r,w\}}, F)$ in $C(\Omega)$ such that where

$$F(\emptyset)(\omega) = 0, \quad F(\{r\})(\omega) = \omega, \quad F(\{w\})(\omega) = 1 - \omega, \quad F(\{r,w\})(\omega) = 1$$
$$(\forall \omega \in [0,1] \equiv \Omega). \tag{5.37}$$

Here, consider the following measurement M_ω:

$$M_\omega := \text{"Pick out one ball from the urn } \omega\text{, and recognize the color of the ball"} \tag{5.38}$$

That is, we consider

$$M_\omega = \mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[\delta_\omega]}). \tag{5.39}$$

Moreover we define the product observable $\mathbf{O}^N \equiv (X^N, \mathcal{P}(X^N), F^N)$, such that:

$$\begin{aligned}
&[F^N(\Xi_1 \times \Xi_2 \times \cdots \times \Xi_{N-1} \times \Xi_N)](\omega) \\
=&[F(\Xi_1)](\omega) \cdot [F(\Xi_2)](\omega) \cdots [F(\Xi_{N-1})](\omega) \cdot [F(\Xi_N)](\omega) \\
&(\forall \omega \in \Omega \equiv [0,1], \quad \forall \Xi_1, \Xi_2, \cdots, \Xi_N \subseteq X \equiv \{r, w\}).
\end{aligned} \tag{5.40}$$

As mentioned in Definition 2.27, we think that

$$\text{"take a measurement } M_\omega \text{ N times"} \Leftrightarrow \text{"take a measurement } \mathbf{M}_{C(\Omega)}(\mathbf{O}^N, S_{[\delta_\omega]})\text{"} \tag{5.41}$$

Define the estimator $E : X^N(\equiv \{r, w\}^N) \to \Omega(\equiv [0,1])$

$$E(x_1, x_2, \cdots, x_{N-1}, x_N) = \frac{\sharp[\{n \in \{1, 2, \cdots, N\} \mid x_n = r\}]}{N}$$
$$(\forall x = (x_1, x_2, \cdots, x_{N-1}, x_N) \in X^N \equiv \{r, w\}^N). \tag{5.42}$$

For each $\omega(\in [0,1] \equiv \Omega)$, define the positive number η^γ_ω such that:

$$\begin{aligned}
&\eta^\gamma_\omega \\
=&\inf \Big\{ \eta > 0 \;\Big|\; [F^N(\{(x_1, x_2, \cdots, x_N) \mid \omega - \eta \leq E(x_1, x_2, \cdots, x_N) \leq \omega + \eta\})](\omega) > 0.95 \Big\} \\
=&\inf_{[F^N(\{(x_1,x_2,\cdots,x_N) | \omega-\eta \leq E(x_1,x_2,\cdots,x_N) \leq \omega+\eta\})](\omega) > 0.95} \eta.
\end{aligned} \tag{5.43}$$

Put

$$D_x^\gamma = \{\omega (\in \Omega) : \ |E(x) - \omega| \le \eta_\omega^\gamma\}. \tag{5.44}$$

For example, assume that N is sufficiently large and $\gamma = 0.95$. Then we see, by (2.58), that

$$\eta_\omega^{0.95} \approx 1.96\sqrt{\frac{\omega(1-\omega)}{N}}$$

and

$$D_x^{0.95} = [E(x) - \eta_-, E(x) + \eta_+] \tag{5.45}$$

where

$$\eta_- = \eta_{E(x)-\eta_-}^{0.95}, \quad \eta_+ = \eta_{E(x)+\eta_+}^{0.95}. \tag{5.46}$$

Under the assumption that N is sufficiently large, we can consider that

$$\eta_- \approx \eta_+ \approx \eta_{E(x)}^{0.95} \approx 1.96\sqrt{\frac{E(x)(1-E(x))}{N}}.$$

Then we can conclude that

- *for any urn* $\omega (\in \Omega \equiv [0,1])$*, the probability, that the measured value* $x = (x_1, x_2, \cdots, x_N)$ *obtained by the measurement* $\mathbf{M}_{\mathcal{A}}(\mathbf{O}^N, S_{[\delta_\omega]})$ *satisfies the following condition* ($\sharp$)*, is larger than* γ *(e.g.,* $\gamma = 0.95$*).*

  ($\sharp$) $\qquad E(x) - 1.96\sqrt{\frac{E(x)(1-E(x))}{N}} \le \omega \le E(x) + 1.96\sqrt{\frac{E(x)(1-E(x))}{N}}.$

  where E *is defined by (5.42).*

■

5.4 Testing statistical hypothesis

Now we study "testing statistical hypothesis", that is, answer the following question.

Problem 5.16. [Testing statistical hypothesis]. *Consider a measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[*]})$ *formulated in* $C(\Omega)$*. Let* $E : X \to \Omega$ *be Fisher's estimator. Assume the following hypothesis:*

(H) the unknown state $[*]$ *belongs to a closed set* C_H $(\subseteq \Omega)$*.*

And further assume that we see the following fact:

(F) a measured value $x_0(\in X)$ *is obtained by measurement* $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[*]})$*.*

Here, our present purpose is to propose an algorithm that decides whether the above hypothesis (H) can be denied by the fact (F). This algorithm is called "the testing statistical hypothesis."

In the above problem, it is usually expected that the hypothesis (H) is not true. In this sense, the above (H) is called *the null hypothesis.*

Now we provide two answers (i.e., Answer 1 and Answer 2). Answer 1 (likelihood ratio test) is, of course, well-known and authorized. Also, in order to solve the question: "Is there another answer?", we add Answer 2 after Answer 1.

Answer 1. [Likelihood ratio test]. Consider a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[*]})$ formulated in $C(\Omega)$. Let $E : X \to \Omega$ be Fisher's estimator, i.e., it is defined by

$$E(x) = \lim_{\Xi_n \to \{x\}} \omega_n \qquad (\forall x \in X),$$

where ω_n $(\in \Omega)$ is chosen such that it satisfies

$$\frac{[F(\Xi_n)](\omega_n)}{\max_{\omega \in \Omega}[F(\Xi_n)](\omega)} = 1.$$

$\Big($For the exact argument, see Remark 5.4 (Radon-Nikodým derivative).$\Big)$ Assume both (H) and (F) in Problem 5.16. Consider a real number α such that $0 < \alpha \ll 1$ (e.g. $\alpha = 0.05$, which may be called *a significance level.* Let ω be in Ω. Then, by Axiom 1, we have a sample probability measure P_ω on X (of the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\delta_\omega]}))$ such that:

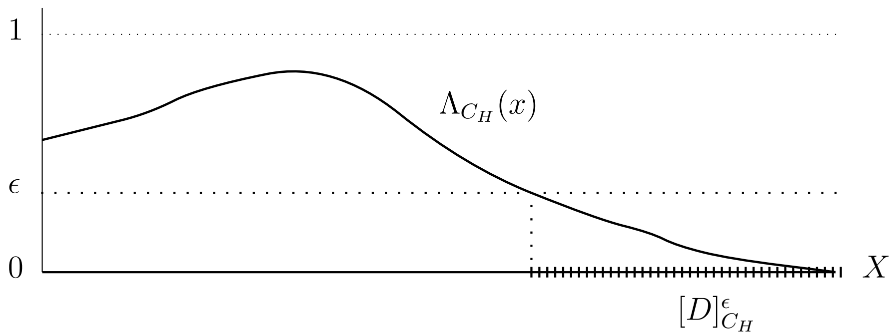
$$P_\omega(\Xi) = [F(\Xi)](\omega) \quad (\forall \Xi \in \mathcal{F}). \tag{5.47}$$

Here define the function $\Lambda_{C_H} : X \to [0, 1]$ such that:

$$\Lambda_{C_H}(x) = \lim_{\Xi \to \{x\}} \frac{\sup_{\omega \in C_H} P_\omega(\Xi)}{\sup_{\omega \in \Omega} P_\omega(\Xi)} \quad (\forall x \in X). \tag{5.48}$$

Also, for any ϵ $(0 < \epsilon \leq 1)$, define $[D]^{\epsilon}_{C_H}$ $(\in \mathcal{F})$ such that:

$$[D]^{\epsilon}_{C_H} = \{x \in X \mid \Lambda_{C_H}(x) < \epsilon\}. \tag{5.49}$$



Thus we can define $\epsilon^{0.05}_{\max}$ such that:

$$\epsilon^{0.05}_{\max} = \sup\{\epsilon \mid \sup_{\omega_0 \in C_H} P_{\omega_0}([D]^{\epsilon}_{C_H}) \leq 0.05\}. \tag{5.50}$$

Now we can conclude that

Answer 1

$$\begin{cases} \text{if } x_0 \in [D]^{\epsilon^{0.05}_{\max}}_{C_H}, \text{ then the hypothesis (H) can be denied} \\ \text{if } x_0 \notin [D]^{\epsilon^{0.05}_{\max}}_{C_H}, \text{ then the hypothesis (H) can not be denied} \end{cases} \tag{5.51}$$

□

Next we shall propose "Answer 2". Before this, we must prepare the following well-known lemma.

Lemma 5.17. [Neyman-Pearson theorem, α-influential domain of ν_1 for ν_2, *cf.* [59]]. *Let $(X, \mathcal{F})$ be a measurable space. Let ν_1 and ν_2 be probability measures on X. Define the Radon-Nikodým derivative $\frac{d\nu_1}{d\nu_2} : X \to [0, \infty)$ such that:*

$$\frac{d\nu_1}{d\nu_2}(x) = \lim_{\Xi \to x} \frac{\nu_1(\Xi)}{\nu_2(\Xi)} \quad (x \in X). \tag{5.52}$$

Put

$$[D](\epsilon, \frac{d\nu_1}{d\nu_2}) = \{x \in X \mid \frac{d\nu_1}{d\nu_2}(x) < \epsilon\}, \qquad (0 \leq \epsilon \leq \infty). \tag{5.53}$$

Thus we can define $\breve{\epsilon}^{0.05}_{\max}$ *such that:*

$$\breve{\epsilon}^{0.05}_{\max}\Big(\equiv \breve{\epsilon}^{\alpha=0.05}_{\max}\Big) = \sup\{\epsilon \mid \nu_1([D](\epsilon, \frac{d\nu_1}{d\nu_2})) \le 0.05\}. \tag{5.54}$$

Now we have the

$$[D](\breve{\epsilon}^{0.05}_{\max}, \frac{d\nu_1}{d\nu_2}), \tag{5.55}$$

which is called *"the* 0.05*-influential domain of* ν_1 *for* ν_2*."*

■

Answer 2. [A test using Neyman-Pearson theorem]. Consider a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[*]})$ formulated in $C(\Omega)$. Let $E : X \to \Omega$ be Fisher's estimator. Assume both (H) and (F) in Problem 5.16. Consider a real number α such that $0 < \alpha \ll 1$ (e.g. $\alpha = 0.05$ which may be also called *a significance level.* Let ω be in C_H. Consider a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O} \equiv (X, \mathcal{F}, F), S_{[\delta_\omega]})$. Let x be in X. Then, we have two sample probability measures P_ω and $P_{E(x)}$ on X such that:

$$\nu_\omega(\Xi) = P_\omega(\Xi) = [F(\Xi)](\omega) \qquad (\forall \Xi \in \mathcal{F})$$

and

$$\nu_{E(x)} = P_{E(x)}(\Xi) = [F(\Xi)](E(x)) \qquad (\forall \Xi \in \mathcal{F}). \tag{5.56}$$

Thus, we have "the 0.05-influential domain of ν_1 for ν_2" such that:

$$[D](\breve{\epsilon}^{0.05}_{\max}, \frac{d\nu_\omega}{d\nu_{E(x)}}). \tag{5.57}$$

Put

$$[D]^{\breve{\epsilon}^{0.05}_{\max}}_{C_H,x} = \cap_{\omega \in C_H} [D](\breve{\epsilon}^{0.05}_{\max}, \frac{d\nu_\omega}{d\nu_{E(x)}}). \tag{5.58}$$

Lastly, we put

$$[D]^{\breve{\epsilon}^{0.05}_{\max}}_{C_H} = \{x \in X \mid x \in [D]^{\breve{\epsilon}^{0.05}_{\max}}_{C_H,x}\}. \tag{5.59}$$

Now we can conclude that

Answer 2

$$\begin{cases} \text{if } x_0 \in [D]^{\breve{\epsilon}^{0.05}_{\max}}_{C_H}, \text{ then the hypothesis (H) can be denied} \\ \\ \text{if } x_0 \notin [D]^{\breve{\epsilon}^{0.05}_{\max}}_{C_H}, \text{ then the hypothesis (H) can not be denied} \end{cases} \tag{5.60}$$

Remark 5.18. [*Answers 1 and 2*]. We believe that the above two answers 1 and 2 are proper though the meanings of "significant level" is different in each answer (*cf.* [II; $C_H = [0,\infty)$] in Examples 5.16 and 5.17). We do not know whether there is another proper answer.

■

Example 5.19. [Likelihood ratio test for the Gaussian observable]. Put $\Omega = \mathbf{R}$, $\mathcal{A} = C_0(\Omega)$, $\mathbf{O}_{\sigma^2} \equiv (\mathbf{R}, \mathcal{B}_{\mathbf{R}}^{\text{bd}}, F_{(\cdot)}^{\sigma^2})$ in $C_0(\Omega)$ such that:

$$F_{\Xi}^{\sigma^2}(\omega) = \frac{1}{\sqrt{2\pi}\sigma}\int_{\Xi} \exp[-\frac{(x-\omega)^2}{2\sigma^2}]du \qquad (\forall \Xi \in \mathcal{B}_{\mathbf{R}}^{\text{bd}}, \quad \forall \omega \in \Omega = \mathbf{R}). \tag{5.61}$$

And thus. consider the product observable $\mathbf{O}_{\sigma}^2 \equiv (\mathbf{R}^2, \mathcal{B}_{\mathbf{R}^2}^{\text{bd}}, F_{(\cdot)}^{\sigma^2}) \times F_{(\cdot)}^{\sigma^2})$ in $C_0(\Omega)$. That is,

$$(F_{\Xi_1}^{\sigma^2} \times F_{\Xi_2}^{\sigma^2})(\omega) = \frac{1}{(\sqrt{2\pi}\sigma)^2}\iint_{\Xi_1 \times \Xi_2} \exp[-\frac{(x_1-\omega)^2 + (x_2-\omega)^2}{2\sigma^2}]dx_1 dx_2$$

$$(\forall \Xi_k \in \mathcal{B}_{\mathbf{R}}^{\text{bd}} (k=1,2), \quad \forall \omega \in \Omega = \mathbf{R}). \tag{5.62}$$

[Case(I): Two sided test, i.e., $C_H = \{\omega_0\}$]. Assume that $C_H = \{\omega_0\}$, $\omega_0 \in \Omega = \mathbf{R}$. Then,

$$\begin{aligned}
\Lambda_{\{\omega_0\}}(x_1,x_2) &= \lim_{\Xi_1 \times \Xi_2 \to \{(x_1,x_2)\}} \frac{\sup_{\omega \in \{\omega_0\}} P_\omega(\Xi_1 \times \Xi_2)}{\sup_{\omega \in \Omega} P_\omega(\Xi_1 \times \Xi_2)} \\
&= \frac{\exp[-\frac{(x_1-\omega_0)^2 + (x_2-\omega_0)^2}{2\sigma^2}]}{\exp[-\frac{(x_1-(x_1+x_2)/2)^2 + (x_2-(x_1+x_2)/2)^2}{2\sigma^2}]} \\
&= \exp[-\frac{[(x_1+x_2)-2\omega_0]^2}{4\sigma^2}] = \exp[-\frac{[(x_1+x_2)/2-\omega_0]^2}{2(\sigma/\sqrt{2})^2}] \\
&\quad (\forall (x_1,x_2) \in \mathbf{R}^2).
\end{aligned} \tag{5.63}$$

Also, for any $\epsilon (>0)$, define $[D]_{\{\omega_0\}}^{\epsilon}$ $(\in \mathcal{F})$ such that:

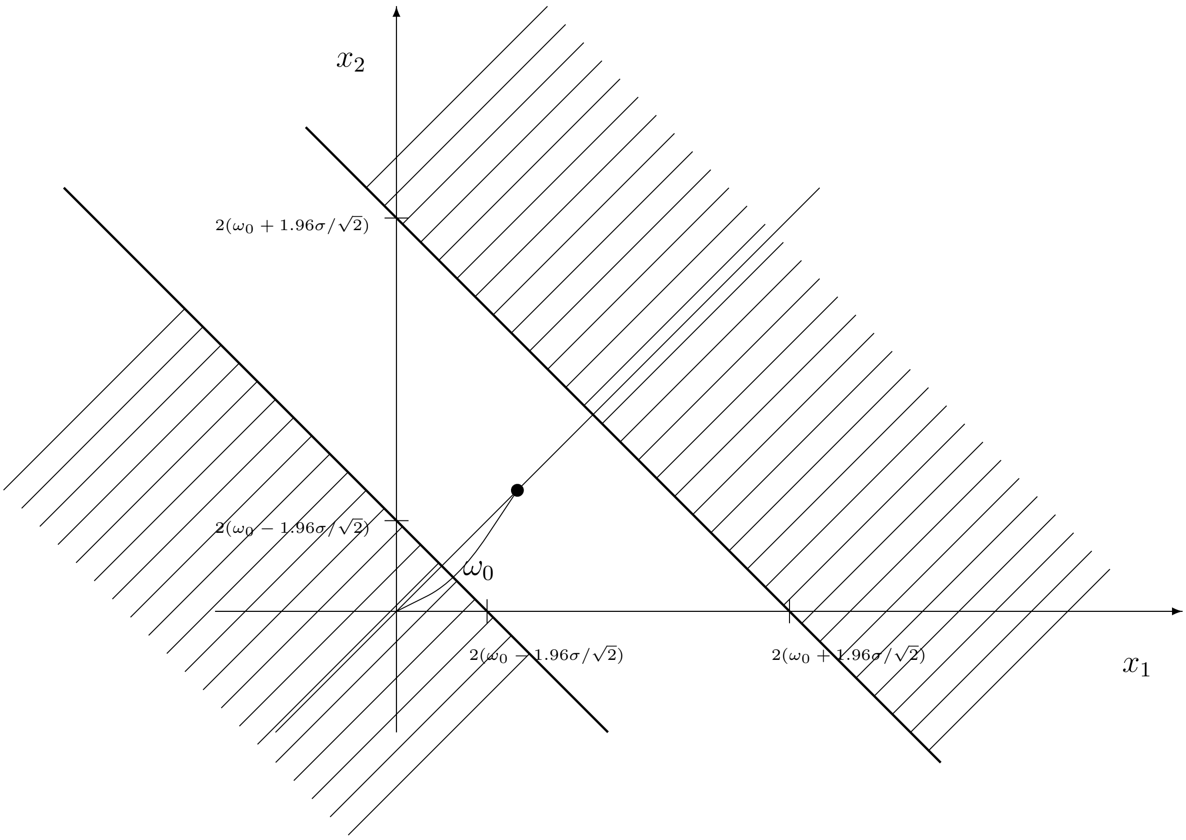
$$[D]_{\{\omega_0\}}^{\epsilon} = \{(x_1,x_2) \in \mathbf{R}^2 \mid \Lambda_{\{\omega_0\}}(x_1,x_2) < \epsilon\}. \tag{5.64}$$

Thus we can define $\epsilon_{\max}^{0.05}$ such that:

$$\epsilon_{\max}^{0.05} = \sup\{\epsilon \mid \sup_{\omega \in \{\omega_0\}} P_\omega([D]_{\{\omega_0\}}^{\epsilon}) \leq 0.05\}. \tag{5.65}$$

Now we can conclude that

$$\begin{aligned}
&[D]_{\{\omega_0\}}^{\epsilon_{\max}^{0.05}} \\
=&\{(x_1,x_2) \in \mathbf{R}^2 \mid (x_1+x_2)/2 \leq \omega_0 - 1.96\sigma/\sqrt{2}\} \\
&\qquad \bigcup \{(x_1,x_2) \in \mathbf{R}^2 \mid (x_1+x_2)/2 \geq \omega_0 + 1.96\sigma/\sqrt{2}\} \\
=&\text{"Slash part in the following figure"}
\end{aligned}$$



[Case(II): One sided test, i.e., $C_H = [\omega_0, \infty)$]. Assume that $C_H = [\omega_0, \infty)$, $\omega_0 \in \Omega = \mathbf{R}$. Then,

$$
\begin{aligned}
\Lambda_{[0,\infty)}(x_1, x_2) &= \lim_{\Xi_1 \times \Xi_2 \to \{(x_1,x_2)\}} \frac{\sup_{\omega \in [\omega_0,\infty)} P_\omega(\Xi_1 \times \Xi_2)}{\sup_{\omega \in \Omega} P_\omega(\Xi_1 \times \Xi_2)} \\
&= \begin{cases} \exp[-\frac{[(x_1+x_2)-2\omega_0]^2}{4\sigma^2}] & (\frac{x_1+x_2}{2} < \omega_0) \\ 1 & (\text{ otherwise }) \end{cases}
\end{aligned} \tag{5.66}
$$

Also, for any $\epsilon(> 0)$, define $[D]^{\epsilon}_{[\omega_0,\infty)}$ $(\in \mathcal{F})$ such that:

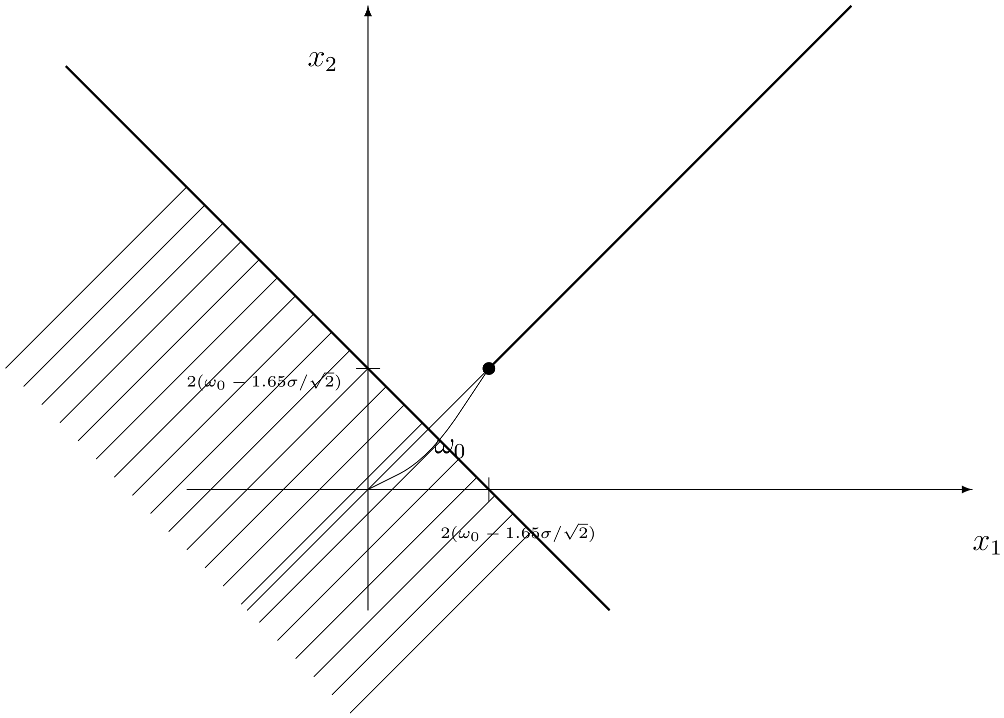
$$
\begin{aligned}
[D]^{\epsilon}_{[0,\infty)} &= \{(x_1, x_2) \in \mathbf{R}^2 \mid \Lambda_{[0,\infty)}(x_1, x_2) \le \epsilon\} \\
&= \{(x_1, x_2) \in \mathbf{R}^2 \mid \frac{x_1 + x_2}{2} - \omega_0 < \sqrt{4\sigma^2 \log \epsilon}\}.
\end{aligned} \tag{5.67}
$$

Thus we can define $\epsilon^{0.05}_{\max}$ such that:

$$
\epsilon^{0.05}_{\max} = \sup\{\epsilon \mid \sup_{\omega_0 \in [0,\infty)} P_{\omega_0}([D]^{\epsilon}_{[0,\infty)}) \le 0.05\}. \tag{5.68}
$$

Therefore, we can conclude that

$$
\begin{aligned}
&[D]^{\epsilon^{0.05}_{\max}}_{[0,\infty)} \\
=&\{(x_1, x_2) \in \mathbf{R}^2 \ \mid \ (x_1 + x_2)/2 \le \omega_0 - 1.65\sigma/\sqrt{2}\}. \qquad (\textit{cf.}\ (2.58)). \\
=&\text{“Slash part in the following figure”}
\end{aligned}
$$



■

Example 5.20. [The test using Neyman-Pearson theorem for the Gaussian observable]. Put $\Omega = \mathbf{R}$, $\mathcal{A} = C_0(\Omega)$, $\mathbf{O}_{\sigma^2} \equiv (\mathbf{R}, \mathcal{B}_{\mathbf{R}}^{\mathrm{bd}}, F_{(\cdot)}^{\sigma^2})$ and $\mathbf{O}_{\sigma^2}^2 \equiv (\mathbf{R}^2, \mathcal{B}_{\mathbf{R}^2}^{\mathrm{bd}}, F_{(\cdot)}^{\sigma^2} \times F_{(\cdot)}^{\sigma^2})$ in $C_0(\Omega)$ are as in the above.

[Case(I): Two sided test, i.e., $C_H = \{\omega_0\}$]. Assume that $C_H = \{\omega_0\}$, $\omega_0 \in \Omega = \mathbf{R}$. Then,

$$\nu_1^{\omega_0}(\Xi_1 \times \Xi_2) = P_{\omega_0}(\Xi_1 \times \Xi_2) = [F(\Xi_1 \times \Xi_2)](\omega_0) \qquad (\forall \Xi_1 \times \Xi_2 \in \mathcal{B}_{\mathbf{R}^2}^{\mathrm{bd}})$$

and

$$\nu_2^{E(x_0)} = P_{E(x_0)}(\Xi_1 \times \Xi_2) = [F(\Xi_1 \times \Xi_2)](E(x_0)) \qquad (\forall \Xi_1 \times \Xi_2 \in \mathcal{B}_{\mathbf{R}^2}^{\mathrm{bd}}). \quad (5.69)$$

Thus, we have "the 0.05-influential domain of ν_1 for ν_2" such that:

$$[D](\breve{\epsilon}_{\max}^{0.05}, \phi_{\nu_1^{\omega_0}/\nu_2^{E((x_1,x_2))}}) = \begin{cases} \{(x_1,x_2) \mid (x_1+x_2)/2 \le \omega_0 - 1.65\sigma/\sqrt{2}\} & (E(x_0) < \omega_0) \\ \{(x_1,x_2) \mid (x_1+x_2)/2 \ge \omega_0 + 1.65\sigma/\sqrt{2}\} & (E(x_0) > \omega_0). \end{cases}$$

Put

$$[D]_{\{\omega_0\},(x_1,x_2)}^{\breve{\epsilon}_{\max}^{0.05}} = \cap_{\omega_0 \in \{\omega_0\}} [D](\breve{\epsilon}_{\max}^{0.05}, \phi_{\nu_1^{\omega_0}/\nu_2^{E((x_1,x_2))}}) = [D](\breve{\epsilon}_{\max}^{0.05}, \phi_{\nu_1^{\omega_0}/\nu_2^{E((x_1,x_2))}}) \quad (\forall (x_1,x_2) \in \mathbf{R}^2). \quad (5.70)$$

Therefore, we can conclude that

$$
\begin{aligned}
[D]^{\check{\epsilon}^{0.05}_{\max}}_{\{\omega_0\}} =& \{(x_1,x_2) \in \mathbf{R}^2 \mid (x_1,x_2) \in [D]^{\check{\epsilon}^{0.05}_{\max}}_{\{\omega_0\},(x_1,x_2)}\} \\
=& \{(x_1,x_2) \ \mid \ (x_1+x_2)/2 \le \omega_0 - 1.65\sigma/\sqrt{2}\} \\
& \qquad \bigcup \{(x_1,x_2) \ \mid \ (x_1+x_2)/2 \ge \omega_0 + 1.65\sigma/\sqrt{2}\}.
\end{aligned}
$$

[Case(II): One sided test, i.e., $C_H = [\omega_0, \infty)$]. Assume that $C_H = [\omega_0, \infty)$, $\omega_0 \in \Omega = \mathbf{R}$. Then,

$$
\nu_1^{\omega_0}(\Xi_1 \times \Xi_2) = P_{\omega_0}(\Xi_1 \times \Xi_2) = [F(\Xi_1 \times \Xi_2)](\omega_0) \qquad (\forall \Xi_1 \times \Xi_2 \in \mathcal{B}^{\rm bd}_{\mathbf{R}^2})
$$

and

$$
\nu_2^{E(x_0)}(\Xi_1 \times \Xi_2) = P_{E(x_0)}(\Xi_1 \times \Xi_2) = [F(\Xi_1 \times \Xi_2)](E(x_0)) \qquad (\forall \Xi_1 \times \Xi_2 \in \mathcal{B}^{\rm bd}_{\mathbf{R}^2}). \tag{5.71}
$$

Thus, we have "the 0.05-influential domain of ν_1 for ν_2" such that:

$$
[D](\check{\epsilon}^{0.05}_{\max}, \phi_{\nu_1^{\omega_0}/\nu_2^{E((x_1,x_2))}}) = \left\{ \begin{array}{ll} \{(x_1,x_2) \ \mid \ (x_1+x_2)/2 \le \omega_0 - 1.65\sigma/\sqrt{2}\} & (E(x_0) < \omega_0) \\ \{(x_1,x_2) \ \mid \ (x_1+x_2)/2 \ge \omega_0 + 1.65\sigma/\sqrt{2}\} & (E(x_0) > \omega_0). \end{array} \right.
$$

Put

$$
[D]^{\check{\epsilon}^{0.05}_{\max}}_{[0,\infty),(x_1,x_2)} = \cap_{\omega_0 \in [0,\infty)} [D](\check{\epsilon}^{0.05}_{\max}, \phi_{\nu_1^{\omega_0}/\nu_2^{E(x)}}) \qquad (\forall (x_1,x_2) \in \mathbf{R}^2). \tag{5.72}
$$

Therefore, we can conclude that

$$
\begin{aligned}
[D]^{\check{\epsilon}^{0.05}_{\max}}_{[0,\infty)} =& \{(x_1,x_2) \in \mathbf{R}^2 \mid (x_1,x_2) \in [D]^{\check{\epsilon}^{0.05}_{\max}}_{[0,\infty),(x_1,x_2)}\} \\
=& \{(x_1,x_2) \ \mid \ (x_1+x_2)/2 \le \omega_0 - 1.65\sigma/\sqrt{2}\}.
\end{aligned}
$$

■

5.5 Measurement error model in PMT

Although we have several kinds of measurement error models in statistics (*cf.* Fuller [25], Cheng, etc. [16]), the following may be the simplest one (i.e., with normal distributions (= Gaussian distributions)):

$$
\left\{ \begin{array}{l} \widetilde{y}_n = \theta_0 + \theta_1 x_n + e_n, \\ \widetilde{x}_n = x_n + u_n \\ (e_n, u_n) \sim \mathrm{NI}[\mathrm{average}(0,0), \mathrm{variance}(\sigma_{ee}^2, \sigma_{uu}^2)], {}^{9} \\ \qquad (n = 1, 2, ..., N), \end{array} \right. \tag{5.73}
$$

which, of course, corresponds to the conventional statistics (i.e., the measurement equation in the dynamical system theory (1.2)). 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 value $\{(\widetilde{x}_n, \widetilde{y}_n)\}_{n=1}^N$. Precisely speaking, the purpose of this section is to study this problem in general situations (i.e., without the assumption of normal distributions).

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 an *index state space* (or *parameter space*). Consider a parameterized continuous map ψ^θ : $\Omega_0 \to \Omega_1$, $\theta \in \Theta$, which induces the parameterized homomorphism $\Psi^\theta : C(\Omega_1) \to C(\Omega_0)$ such that (*cf.* (3.14))

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

Consider observables $\mathbf{O}_0 \equiv (X, \mathcal{F}, F)$ in $C(\Omega_0)$ and $\mathbf{O}_1 \equiv (Y, \mathcal{G}, G)$ in $C(\Omega_1)$. And recall that $\Psi^\theta \mathbf{O}_1$ can be identified with the observable in $C(\Omega_0)$ (*cf.* Remark 3.6 (i)). 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 $C(\Omega_0)$. Thus, 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,

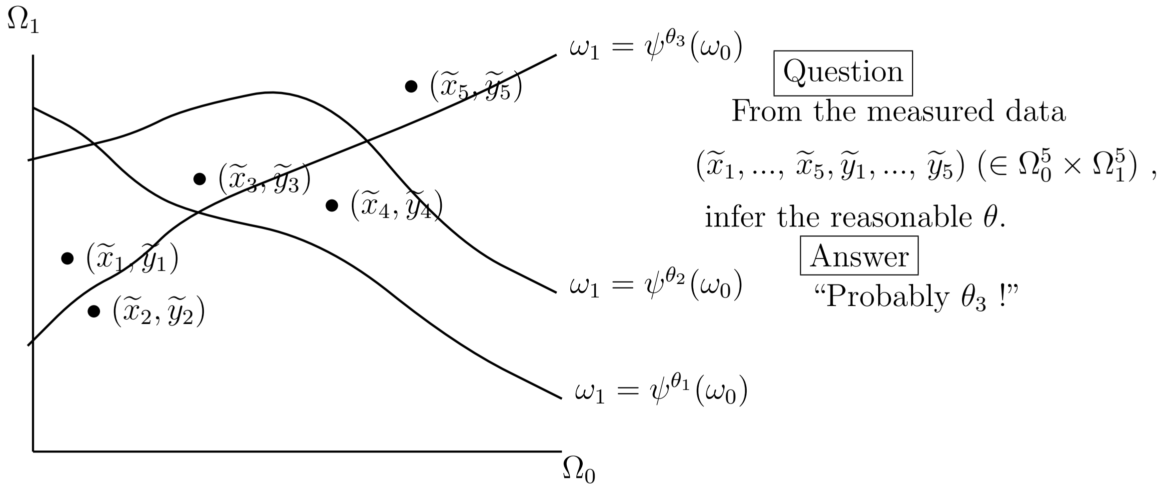
$$\begin{aligned}
&[(\bigotimes_{n=1}^N (F \times \Psi^\theta G))(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, \ldots, \omega_N) \\
=&[F \times \Psi^\theta G(\Xi_1 \times \Gamma_1)](\omega_1) \cdot [F \times \Psi^\theta G(\Xi_2 \times \Gamma_2)](\omega_2) \cdots [F \times \Psi^\theta G(\Xi_N \times \Gamma_N)](\omega_N) \\
&\qquad (\forall \Xi_n \in \mathcal{F}, \forall \Gamma_n \in \mathcal{G}, \forall (\omega_1, \ldots, \omega_N) \in \Omega_0^N). \qquad (5.74)
\end{aligned}$$

Our present problem is as follows:

(♯) 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 $\bar{\omega}_1, \bar{\omega}_2, \ldots, \bar{\omega}_N$ and $\bar{\theta}$ $(\in \Theta)$ are unknown. Assume that we know that the measured value $(\widetilde{x}_1, \ldots, \widetilde{x}_N, \widetilde{y}_1, \ldots, \widetilde{y}_N)$ $(\in X^N \times Y^N)$ 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 $\bar{\omega}_1, \bar{\omega}_2, \ldots, \bar{\omega}_N$ and $\bar{\theta}$ (particularly, $\bar{\theta}$).

[9] Independent random variables with normal distributions

That is, for simplicity under the assumption that $\Omega_0 = X$, $\Omega_1 = Y$, we can illustrate this problem (♯) as follows:



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 $[\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, ..., \omega_N, \theta) = (5.74)$. Note that we have the following identification:

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

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 (*cf.* Corollary 5.6), infer the unknown state $(\otimes_{n=1}^N \delta_{\bar{\omega}_n}) \otimes \delta_{\bar{\theta}}$ such that:

$$\begin{aligned} &[\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\bar{\omega}_1, ..., \bar{\omega}_N, \bar{\theta}) \\ = &\max_{(\omega_1, ..., \omega_N, \theta) \in \Omega_0^N \times \Theta} [\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, ..., \omega_N, \theta). \end{aligned} \tag{5.75}$$

This is the answer to the above problem (♯). It should be noted that the problem (♯) is stated under the very general situations (i.e., Ω_0, Ω_1, X and Y are not necessarily the real lines $\mathbf{R}$).

In the following example, we apply our result (5.75) to the simple measurement error model (5.73) with normal distributions.

Example 5.21. [The simple example of measurement error model (the case that θ_0, θ_1, ω_1 , ..., ω_N are unknown)]. 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:

$$\psi^{(\theta_1, \theta_2)}(\omega) = \theta_1 \omega + \theta_0 \quad (\forall \omega \in \Omega_0, \forall (\theta_0, \theta_1) \in \Theta.$$

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 2.17). Thus, we define the product observable $\widetilde{\mathbf{O}}^{(\theta_0,\theta_1)} = (X \times Y, \mathcal{F} \times \mathcal{G}, H^\theta)$, where $H^\theta \equiv F \times \Psi^\theta G$, in $C(\Omega_0)$ such that:

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

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:

$$[\widehat{H}(\Xi_1 \times \cdots \times \Xi_N \times \Gamma_1 \times \cdots \times \Gamma_N)](\omega_1, ..., \omega_N, \theta_0, \theta_1)$$
$$=(\frac{1}{\sqrt{2\pi\sigma_1\sigma_2}})^{2N} \int\cdots\int_{\Pi_{n=1}^N(\Xi_n\times\Gamma_n)} e^{-\frac{\sum_{n=1}^N(x_n-\omega_n)^2}{2\sigma_1^2} - \frac{\sum_{n=1}^N(y_n-(\theta_1\omega_n+\theta_0))^2}{2\sigma_2^2}} dx_1dy_1\cdots dx_Ndy_N. \tag{5.76}$$

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

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

Then, our main result (5.75) says that

$$\max_{(\omega_1,...,\omega_N,\theta_0,\theta_1)\in\Omega_0^N\times\Theta} [\widehat{H}(\Xi_1^\epsilon \times \cdots \times \Xi_N^\epsilon \times \Gamma_1^\epsilon \times \cdots \times \Gamma_N^\epsilon)](\omega_1, ..., \omega_N, \theta)$$
$$\Longleftrightarrow \min_{(\omega_1,...,\omega_N,\theta_0,\theta_1)\in\Omega_0^N\times\Theta} \Big[\sum_{n=1}^N(\frac{\widetilde{x}_n}{\sigma_1} - \frac{\omega_n}{\sigma_1})^2 + \sum_{n=1}^N(\frac{\widetilde{y}_n}{\sigma_2} - (\frac{\theta_1\sigma_1}{\sigma_2}\frac{\omega_n}{\sigma_1} + \frac{\theta_0}{\sigma_2}))^2\Big] \quad \text{(since } \epsilon \text{ is small)}$$

(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)

$$\Longleftrightarrow \min_{(\theta_0,\theta_1)\in\Theta} \frac{\sum_{n=1}^N(\widetilde{y}_n - \theta_1\widetilde{x}_n - \theta_0)^2}{\sigma_2^2 + \sigma_1^2\theta_1^2} \tag{5.77}$$
$$\Longleftrightarrow \begin{cases} \sum_{n=1}^N(\widetilde{y}_n - \bar{\theta}_1\widetilde{x}_n - \bar{\theta}_0) = 0 & (\leftarrow \frac{\partial}{\partial\theta_0}(5.77) = 0), \\ \sum_{n=1}^N(\bar{\theta}_1\widetilde{y}_n\sigma_1^2 + \widetilde{x}_n\sigma_2^2 - \bar{\theta}_0\bar{\theta}_1\sigma_1^2)(\widetilde{y}_n - \bar{\theta}_1\widetilde{x}_n - \bar{\theta}_0) = 0 & (\leftarrow \frac{\partial}{\partial\theta_1}(5.77) = 0). \end{cases} \tag{5.78}$$

Thus, the unknown parameters $\bar{\theta}_1$ and $\bar{\theta}_2$ are inferred by the solution of this equation (5.78). Note that this is a direct consequence of our main result (5.75).

■

Example 5.22. [The case that θ_0, θ_1, σ_1^2, σ_2^2, ω_1 , ..., ω_N are unknown]. Assume that θ_0, θ_1, σ_1^2, σ_2^2, ω_1 , ..., ω_N are unknown. The log-likelihood is

$$L(\theta_0, \theta_1, \sigma_1^2, \sigma_2^2, \omega_1, ..., \omega_N) = \log[(5.76)]$$
$$= -\frac{N\log\sigma_1^2}{2} - \frac{N\log\sigma_2^2}{2} - \frac{\sum_{n=1}^N(x_n-\omega_n)^2}{2\sigma_1^2} - \frac{\sum_{n=1}^N(y_n-\theta_0-\theta_1\omega_n)^2}{2\sigma_2^2}.$$

Taking partial derivatives with respect to θ_0, θ_1, σ_1^2, σ_2^2 and ω_1, ..., ω_N, and equating the results to zero, gives the likelihood equations,

$$\sum_{n=1}^{N}(y_n - \theta_0 - \theta_1\omega_n) = 0, \qquad \sum_{n=1}^{N}(y_n - \theta_0 - \theta_1\omega_n)\omega_n = 0,$$

$$\frac{\sum_{n=1}^{N}(x_n\omega_n)^2}{N} = \sigma_1^2, \qquad \frac{\sum_{n=1}^{N}(y_n - \theta_0 - \theta_1\omega_n)^2}{N} = \sigma_2^2,$$

$$\frac{(x_n\omega_n)^2}{2\sigma_1^2} - \frac{(y_n - \theta_0 - \theta_1\omega_n)^2}{2\sigma_2^2} = 0, \qquad (n = 1, 2, ..., N).$$

Thus we can easily solve it as follows:

$$\theta_1^2 = \frac{\sigma_2^2}{\sigma_1^2} = \frac{S_{yy}}{S_{xx}}, \quad 2\sigma_1^2 = S_{xx} - \frac{S_{xy}}{\theta_1}, \quad 2\sigma_2^2 = S_{yy} - S_{xy}\theta_1,$$

$$\theta_0 = \bar{y} - \theta_1\bar{x}, \quad 2\omega_n = x_n + \frac{y_n - \theta_0}{\theta_1} = x_n + \bar{x} + \frac{y_n - \bar{y}_n}{\theta_1},$$

where

$$\bar{x} = \frac{x_1 + \cdots + x_N}{N}, \qquad \bar{y} = \frac{y_1 + \cdots + y_N}{N},$$

$$S_{xx} = \frac{(x_1 - \bar{x})^2 + \cdots + (x_N - \bar{x})^2}{N}, \quad S_{yy} = \frac{(y_1 - \bar{y})^2 + \cdots + (y_N - \bar{y})^2}{N},$$

$$S_{xy} = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_N - \bar{x})(y_N - \bar{y})}{N}.$$

(Cf. Cheng, etc. [16]).

■

5.6 Appendix (Iterative likelihood function method)

In this section we study the "Iterative likelihood function method (*cf.* [47])", which will be related to subjective Bayesian statistics (see §8.6 later).

Consider the "measurement" described in the following "step [1]" and "step [2]",

[1] First we take a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}_1 \equiv (X, 2^X, F), S_{[*]})$, and we know that the measured value is equal to x ($\in X$).

[2] And successively, we take a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}_2 \equiv (Y, 2^Y, G), S_{[*]})$, and we know that the measured value is equal to y ($\in Y$).

Note that "[1]+[2]" ie equal to the following [3][10] :

[3] We take a measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O}_1 \times \mathbf{O}_2 \equiv \equiv (X \times Y, \mathcal{F} \times \mathcal{G}, H \equiv F \times G)\ S_{[*]})$, and we know that the measured value obtained by $\mathbf{M}_{C(\Omega)}(\mathbf{O}_1 \times \mathbf{O}_2,\ S_{[*]})$ is equal to $(x, y)\ (\in X \times Y)$.

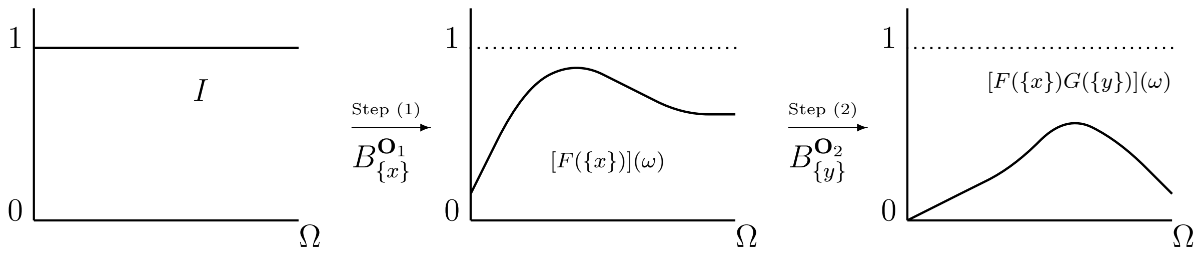
$\Big($A non-negative (real-valued) continuous function $F(\Xi)$ in an observable $(X, \mathcal{F}, F)$ is called *a likelihood function,* or, *a likelihood quantity.*$\Big)$ Then we can say:

[♭] By Step [1], we get the likelihood function $F(\{x\})$. And further by step [2] (i.e., by "[1]+[2]" (=[3])), we get the new likelihood function $F(\{x\})G(\{y\})$ ($\equiv [F \times G](\{x\} \times \{y\})$).

Using the Bayes operator (*cf.* the formula (5.12)), this statement [♭] can be rewritten as follows:

$$I \xrightarrow[B^{\mathbf{O}_1}_{\{x\}}]{\text{Step (1)}} F(\{x\}) \xrightarrow[B^{\mathbf{O}_2}_{\{y\}}]{\text{Step (2)}} F(\{x\})G(\{y\}) \qquad \text{in } C(\Omega), \tag{5.79}$$

where $I(\ \in C(\Omega))$ is the identity element, i.e., the constant function such that $I(\omega) = 1 (\forall \omega \in \Omega)$.



It should be noted that:

(F_1) the constant likelihood function "I" (or "$k \times I$" where $k > 0$) is the likelihood function that represents the fact "we have no information about the system $S_{[*]}$".

Now we introduce the following notation. Cf. [47].

Notation 5.23. $[S_{[*]}(\!(G)\!)_{lq}]$. *The system $S_{[*]}$ (formulated in $C(\Omega)$) such that we know it has the likelihood quantity G ($G \in C(\Omega)$, $0 \le G(\omega)$ $(\forall \omega \in \Omega)$) is denoted by $S_{[*]}(\!(G)\!)_{lq}$.*

[10]Recall §2.5 (Remarks(II)), that is, "Only one measurement is permitted to be conducted". Thus, "[1]+[2]" is a methodological explanation.

Thus, the symbol $\mathbf{M}_{C(\Omega)}\big(\mathbf{O},\, S_{[*]}((kG))_{lq}\big)$ *means "the measurement* $\mathbf{M}_{C(\Omega)}\big(\mathbf{O},\, S_{[*]}\big)$ *under the condition that we know the likelihood quantity of the system* $S_{[*]}$ *is equal to* kG*, where* $G \in C(\Omega)$, $0 \leq G(\omega)$ $(\forall \omega \in \Omega)$*".*

■

Under this notation, the conventional Fisher's maximum likelihood method (i.e., Corollary 5.6) says that:

(F_1') Assume that we first have no information about the system $S_{[*]}$. And we take a measurement $\mathbf{M}_{C(\Omega)}\big(\mathbf{O},\, S_{[*]}\big)$, i.e., $\mathbf{M}_{C(\Omega)}\big(\mathbf{O},\, S_{[*]}((kI))_{lq}\big)$. Then, from the fact that the measured value x $(\in X)$ is obtained by the $\mathbf{M}_{C(\Omega)}\big(\mathbf{O},\, S_{[*]}((kI))_{lq}\big)$, we know that the likelihood quantity of the system $S_{[*]}$ is equal to $k[F(\{x\})](\omega)$. $\Big($Thus, there is a reason to regard the unknown state $[\,*\,]$ as the state $\omega_0(\in \Omega)$ such that $k[F(\{x\})](\omega_0) = \max_{\omega \in \Omega} k[F(\{x\})](\omega).\Big)$

However, it is usual to assume that we have a little bit of information before a measurement. Thus, let us start from the measurement $\mathbf{M}_{C(\Omega)}\big(\mathbf{O} \equiv (X, 2^X, F),\, S_{[*]}((G_0))_{lq}\big)$. Here we have the following problem:

($\mathrm{P_G}$) How to infer the new likelihood quantity of the system $S_{[*]}$ from the fact that the measured value x $(\in X)$ is obtained by the $\mathbf{M}_{C(\Omega)}\big(\mathbf{O},\, S_{[*]}((G_0))_{lq}\big)$.

This is equivalent to the following problem:

($\mathrm{P'_G}$) How to infer the likelihood quantity of the system $S_{[*]}$ from the fact that the measured value (y_0, x) $(\in \{y_0, y_1\} \times X)$ is obtained by the iterated measurement $\mathbf{M}_{C(\Omega)}\big(\mathbf{O}_0 \times \mathbf{O},\, S_{[*]}((kI))_{lq}\big)$, where $\mathbf{O}_0 = (\{y_0, y_1\},\, 2^{\{y_0, y_1\}}, G)$ and $G(\{y_0\}) = G_0$, $G(\{y_1\}) = I - G_0$.

Thus, from (F_1') and "($\mathrm{P_G}$)↔($\mathrm{P'_G}$)", the problem ($\mathrm{P_G}$) is solved as follows:

(F_2) (The answer of the (P_G)): We know that the new likelihood quantity G_{new} of the system $S_{[*]}$ is equal to $B^{\mathbf{O}}_{\{x\}}(G_0)$. Here, Bayes operator $B^{\mathbf{O}}_{\{x\}} : C(\Omega) \to C(\Omega)$ is defined by $B^{\mathbf{O}}_{\{x\}}\,(G) = F(\{x\})G$ $(\forall G \in C(\Omega))$.

Thus we see:

$$S_{[*]}((I))_{lq} \xrightarrow[x \text{ is obtained}]{\mathbf{M}_{C(\Omega)}(\mathbf{O}_1,\, S_{[*]}((I))_{lq})} S_{[*]}((F(\{x\})))_{lq} \xrightarrow[y \text{ is obtained}]{\mathbf{M}_{C(\Omega)}(\mathbf{O}_2,\, S_{[*]}((F(\{x\})))_{lq})} S_{[*]}((F(\{x\})G(\{y\})))_{lq}$$

where $\mathbf{O}_1 = (X, 2^X, F)$ and $\mathbf{O}_2 = (Y, 2^Y, G)$.

Summing up, we can symbolically describe it as follows:

$$\begin{cases} [F_1] \quad \text{No information quantity} \quad \longleftrightarrow \quad kI(\in C(\Omega)) \\ [F_2] \quad S_{[*]}((G))_{lq} \xrightarrow[x \text{ is obtained}]{\mathbf{M}_{C(\Omega)}(\mathbf{O},\, S_{[*]}((G))_{lq})} S_{[*]}((B^{\mathbf{O}}_{\{x\}}G))_{lq}\Big(= S_{[*]}((F(\{x\})G))_{lq}\Big), \end{cases} \tag{5.80}$$

where $\mathbf{O} = (X, 2^X, F)$.

The following example will promote the understanding of "iterative likelihood function method".

Example 5.24. [The urn problem]. There are two urns ω_1 and ω_2. The urn ω_1 [resp. ω_2] contains 8 white and 2 black balls [resp. 4 white and 6 black balls]. Assume that they can not be distinguished in appearance.

- Choose one urn from the two.

Now you sample, randomly, with replacement after each ball.

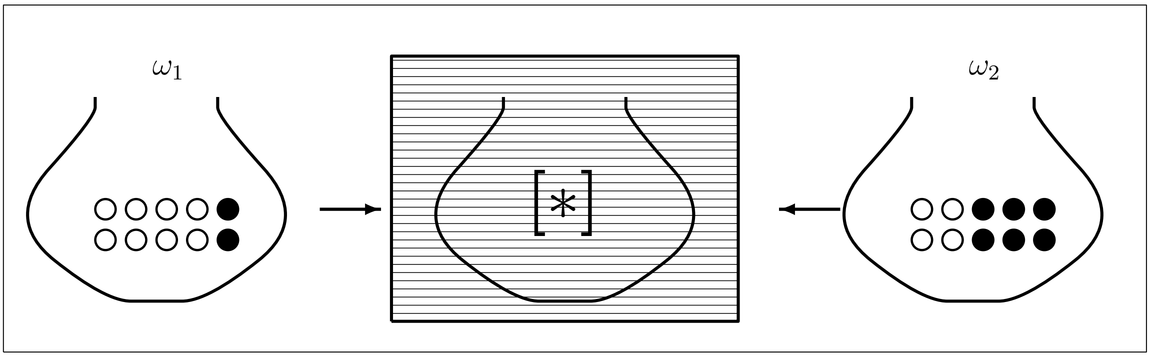
(i). First, you get "white ball".

(Q_1) *Which is the chosen urn,* ω_1 *or* ω_2*?*

(ii). Further, assume that you continuously get "black".

(Q_2) *How about the case? Which is the chosen urn,* ω_1 *or* ω_2*?*

The illustration of $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]})$ (or, $\mathbf{M}_{C(\Omega)}(\mathbf{O}, S_{[*]}((kI))_{lq})$)



[*Answers*]. In what follows this problem is studied in the iterative likelihood function method. Put $\Omega = \{\omega_1, \omega_2\}$. $\mathbf{O} = (\{w, b\}, 2^{\{w,b\}}, F)$ where

$$[F(\{w\})](\omega_1) = 0.8,\ [F(\{b\})](\omega_1) = 0.2,\ [F(\{w\})](\omega_2) = 0.4,\ [F(\{b\})](\omega_2) = 0.6. \tag{5.81}$$

The situation of no information in Fisher's method is represented by kI $(k > 0)$. Thus, it suffices to consider the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O},\ S_{[*]}(\!(kI)\!)_{lq})$. Since the measured value "w" was obtained, the new likelihood quantity G_{new} is given as follows:

$$
\begin{aligned}
G_{\text{new}}(\omega_1)\Big(= kI \cdot [F(\{w\})](\omega_1)\Big) &= 0.8k, \\
G_{\text{new}}(\omega_2)\Big(= kI \cdot [F(\{w\})](\omega_2)\Big) &= 0.4k.
\end{aligned}
\tag{5.82}
$$

Thus, by Fisher's maximum likelihood method, we see that

(A_1) *there is a reason to infer that* $[\,*\,] = \omega_1$.

For the further case, it suffices to consider the measurement $\mathbf{M}_{C(\Omega)}(\mathbf{O},\ S_{[*]}(\!(G_{\text{new}})\!)_{lq})$. Thus we similarly calculate that

$$
\begin{aligned}
G_{\text{new}^2}(\omega_1)\ \Big(= [G_{\text{new}}](\omega_1) \cdot [F(\{b\})](\omega_1)\Big) &= 0.16k, \\
G_{\text{new}^2}(\omega_2)\ \Big(= [G_{\text{new}}](\omega_2) \cdot [F(\{b\})](\omega_2)\Big) &= 0.24k.
\end{aligned}
\tag{5.83}
$$

Thus we, by Fisher's maximum likelihood method, see that

(A_2) *there is a reason to infer that* $[\,*\,] = \omega_2$.

■