Pioneer Journal of Theoretical and Applied Statistics
Volume 11, Numbers 1-2, 2016, Pages 25-35
This paper is available online at http://www.pspchv.com/content_PJTAS.html

MEASUREMENT THEORETICAL APPROACH TO REGRESSION ANALYSIS

KOHSHI KIKUCHI

Faculty of Liberal Studies
National Institute of Technology
Kumamoto College
2659-2 Suya, Koshi
Kumamoto 861-1102
Japan

Abstract

To make a unified understanding of statistic theory, we have proposed a measurement theory which is stated with an introduction of two axioms based on the principle in quantum mechanics, Born's probabilistic interpretation and Heisenberg's picture representation of simultaneous measurements. The objective of this paper is to examine regression analyses through the measurement theory.

0. Introduction

Adopting, as axioms, Born's probabilistic interpretation and Heisenberg's simultaneous measurements in quantum mechanics, Ishikawa [2] introduced a measurement theory with a view to setting a framework to make a unified

Received August 9, 2016; Revised September 13, 2016

2010 Mathematics Subject Classification: 62A01, 62F10, 81P15.

Keywords and phrases: measurement theory, observables, Fisher's maximum likelihood method, simultaneous measurements, Born's probabilistic interpretation, Heisenberg picture.

understanding of various kinds of research fields in so-called system theories. The objective of the present paper is to apply the method to give a good understanding of inferred issues in statistics, where we rely on Gauss-Fisher's principle via Bayes theory.

Throughout in this paper, the symbol Ω denotes a compact Hausdorff space with the Borel field $\mathcal{B}_\Omega$. The space $C(\Omega)$ denotes the Banach-algebra:

$$C(\Omega) := \{f : \Omega \to \mathbb{R} \mid f \text{ is continuous on } \Omega\}.$$

endowed with canonical structures. We use the notation: for functions f, g, the function f_g denotes the product of the functions and the inequality $f \leq g$ is defined by

$$f(\omega) \leq g(\omega) \text{ for } \omega \in \Omega.$$

The function 0 and 1 denotes the zero function and the constant 1 function, respectively.

Let $C(\Omega)^*$ be the dual Banach space of $C(\Omega)$ with the operator norm and with the product $\langle \hat{\rho}, f \rangle$ for $\hat{\rho} \in C(\Omega)^*$ and $f \in C(\Omega)$. We denote by $\mathcal{M}(\Omega)$ the set of all measures on $\mathcal{B}_\Omega$. The Riesz theorem says that $C(\Omega)^*$ and $\mathcal{M}(\Omega)$ can be identified in the following sense: there exists an isometrically continuous, linear and bijective map $\Psi : \mathcal{M}(\Omega) \to C(\Omega)^*$ with the identity

$$\Psi(\rho)(f) = \int_\Omega f(\omega)\rho(d\omega) \text{ for } f \in C(\Omega) \text{ and } \rho \in \mathcal{M}(\Omega).$$

We shall introduce two subclasses of $\mathcal{M}(\Omega)$: the mixed state class $\mathcal{M}^m(\Omega)$ and the pure state class $\mathcal{M}^p(\Omega)$ defined by

$$\mathcal{M}^m(\Omega) := \{\rho \in \mathcal{M}(\Omega) \mid \rho(A) \geq 0 \text{ for } A \in \mathcal{B}_\Omega \text{ and } \rho(\Omega) = 1\},$$

$$\mathcal{M}^p(\Omega) := \{\delta_\omega \in \mathcal{M}(\Omega) \mid \omega \in \Omega\},$$

where δ_ω is a point measure at $\omega \in \Omega$, i.e.,

$$\delta_\omega(A) = 1 \text{ for } \omega \in A \text{ and } 0 \text{ for } \omega \notin A.$$

The space $\mathcal{M}^p(\Omega)$ with the weak* topology can be identified with the Ω, $\mathcal{M}^p(\Omega)$ is called a state space under the identification.

1. Born Interpretations

Following Davies [1], we shall introduce a concept of observables. In all descriptions of this paper, the symbol X denotes a set composed of finite elements and $\mathcal{F}_X = 2^X$ the power set of X. We call a triplet $(X, \mathcal{F}_X, F)_{C(\Omega)}$ with an index $C(\Omega)$ an observable if $F : \mathcal{F}_X \to C(\Omega)$ satisfies

(i) $0 \leq F(E) \leq 1$ for $E \in \mathcal{F}_X$, $F(\emptyset) = 0$ and $F(X) = 1$,

(ii) $F(E_1 \cup E_2) = F(E_1) + F(E_2)$ for disjoint sets E_1, $E_2 \in \mathcal{F}_X$.

By abstracting some concepts in mathematical engineerings, we shall introduce concepts of systems, general systems and so on, by use of which we can state or understand various kinds of phenomena through our measurement theory endowed with two axioms.

A *system* is denoted by S, and a system with a state $\omega \in \Omega$ is denoted by $S_{(\omega)}$.

Under this setting, we introduce concepts of measurements of observables for systems.

A *measurement of observable* $O := (X, \mathcal{F}_X, F)_{C(\Omega)}$ for system $S_{(\omega)}$ is denoted by $M(O, S_{(\omega)})$, by which a *measured-value* is obtained as an element of X. The notation $M(O, S_{(*)})$ is used for $M(O, S_{(\omega)})$, when we regard $\omega \in \Omega_0$ as a seemingly unknown state.

Under this understanding of these terminologies, observables and measurements, we adopt Born's probabilistic interpretation of quantum mechanics as an axiom in a foundation of measurement theory.

Axiom 1. The probability that a value measured by $M((X, \mathcal{F}_X, F)_{C(\Omega)}, S_{(\omega)})$ belongs to a set $E \in \mathcal{F}_X$ is given by $F(E)(\omega)$.

Axiom 1 together with the identification $\Omega \equiv \mathcal{M}^p(\Omega)$ reads

A: If $\delta_\omega \in \mathcal{M}^p(\Omega)$, the probability that a value measured by $M((X, \mathcal{F}_X, F)_{C(\Omega)}, S_{(\omega)})$ belongs to $E \in \mathcal{F}_X$ is given by $\langle \delta_\omega, F(E) \rangle$.

2. Heisenberg Pictures

In this section, we study a relation among systems, which is a fundamental concept in our measurement theory.

For a given observable $O_k := (X_k, \mathcal{F}_{X_k}, F_k)_{C(\Omega)} (k = 1, 2, ..., n)$, the triplet

$$O := \left(\prod_{k=1}^{n} X_k, \mathcal{F}_{\prod_{k=1}^{n} X_k}, F \right)_{C(\Omega)}$$

is called a *product observable* of $\{O_k\}_{k=1}^n$, if

$$F\left(\prod_{k=1}^{n} E_k \right) = \prod_{k=1}^{n} F_k(E_k) \text{ for } \prod_{k=1}^{n} E_k \in \prod_{k=1}^{n} \mathcal{F}_{X_k}; \tag{1}$$

the mapping F is denoted by $\prod_{k=1}^{n} F_k$.

Let $T = \{1, 2, ..., n\}$ and 1_t be the constant 1 in $C(\Omega_t)$.

A continuous linear operator $\Phi_t : C(\Omega_t) \to C(\Omega_0)$ for $t \in T$ is called a *Markov operator*, if

(i) $\Phi_t(f) \geq 0$ for any positive function f in $C(\Omega_t)$,

(ii) $\Phi_t(1_t) = 1_0$.

A Markov operator $\Phi_t : C(\Omega_t) \to C(\Omega_0)$ is called a *homomorphic Markov operator*, if it is homomorphic, that is,

$$\Phi_t(f_g) = \Phi_t(f)\Phi_t(g) \text{ for } f,\ g \in C(\Omega_t).$$

Let $O_t = (X_t, \mathcal{F}_{X_t}, F_t)_{C(\Omega_t)}$ $(t \in T)$ be an observable. The triplet

$\Phi_t O_t := (X_t, \mathcal{F}_{X_t}, \Phi_{t} \circ F_t)_{C(\Omega_0)}$ is verified to be an observable, and called a *pull-back observable* of O_t by a Markov operator Φ_t.

Let $\Phi_t^* : \mathcal{M}^m(\Omega_0) \to \mathcal{M}^m(\Omega_t)$ be the dual Markov operator of

$$\Phi_t : C(\Omega_t) \to C(\Omega_0)$$

defined by

$$\Phi_t^*(\rho)(f) = \rho(\Phi_t(f)) \text{ for } \rho \in \mathcal{M}^m(\Omega_0) \text{ and } f \in C(\Omega_t). \tag{2}$$

Then, the following inclusion is known to hold (cf. [3], [6]):

(i) $\Phi_t^*(\mathcal{M}^m(\Omega_0)) \subset \mathcal{M}^m(\Omega_t)$,

(ii) $\Phi_t^*(\mathcal{M}^p(\Omega_0)) \subset \mathcal{M}^p(\Omega_t)$, provided Φ_t is homomorphic.

Under the identifications $\mathcal{M}^p(\Omega) \equiv \Omega$, the above property (i) reads that the dual operator Φ_t^* induces a transition probability rule $M(\omega, B)$ defined by

$$M(\omega, B) := \Phi_t^*(\delta_\omega)(B) \text{ for } \omega \in \Omega_0 \text{ and } B \in \mathcal{B}_{\Omega_t},$$

and (ii) does that Φ_t^* induces a continuous linear map $\phi_t : \Omega_0 \to \Omega_t$ defined by

$$\delta_{\phi_t(\omega)} := \Phi_t^*(\delta_\omega) \quad \text{for } \omega \in \Omega_0, \tag{3}$$

where we adopt the identification $\omega \equiv \delta_{(\omega)}$.

Let an observable $O_t := (X_t, \mathcal{F}_{X_t}, F_t)_{C(\Omega_t)}$ be given for each $t \in T$, and a tree (4)

$$\begin{array}{ccc} & \overset{\Phi_1}{\swarrow} & C(\Omega_1) \\ & \overset{\Phi_2}{\leftarrow} & C(\Omega_2) \\ C(\Omega_0) & & \vdots \\ & \overset{\Phi_n}{\nwarrow} & C(\Omega_n) \end{array} \tag{4}$$

with Markov operators Φ_t $(t = 1, 2, \ldots, n)$.

Definition 1 (General systems)**.** A tree in figure (4) together with a state $\omega_0 \in \Omega_0$ is called a *general system with an initial state* ω_0 and denoted by $(S_{(\omega_0)}, \{\Phi_t\}_{t\in T})$.

As an axiom for setting a measurement theory, Ishikawa [2] adopts the Heisenberg principle of taking simultaneous measurements. We cite the formulation stated in [5].

Axiom 2. For a general system $(S_{(\omega_0)}, \{\Phi_t\}_{t\in T})$ with an initial state $\omega_0 \in \Omega_0$ and a family $\{(X_t, \mathcal{F}_{X_t}, F_t)_{C(\Omega_t)}\}_{t\in T}$ of observables, there is taken the pull-back and product observable

$$\left(\prod_{t\in T} X_t, \mathcal{F}_{\prod_{t\in T} X_t}, \prod_{t\in T} \Phi_t \circ F_t\right)_{C(\Omega_0)}.$$

Remark 2 (Simultaneous measurements)**.** The product observable is used to take only one measurement for more than one observables. For example, given two observables O_1 and O_2 and $\omega \in \Omega$, we take a *simultaneous measurement* $M(O_1 \times O_2, S_{(\omega)})$, not $M(O_1, S_{(\omega)}) \times M(O_2, S_{(\omega)})$, the fundamental principle in quantum mechanics:

• Only one measurement is permitted to take even in the classical measurement theory.

3. Bayes Formulation

Definition 3 (Bayes operators)**.** Let $\{O_t\}_{t\in T}$ be a family of observables $O_t := (X_t, \mathcal{F}_{X_t}, F_t)_{C(\Omega_t)}$ and $(S_{(\omega_0)}, \{\Phi_t\}_{t\in T})$ a general system with an initial state $\omega_0 \in \Omega_0$. We call a family $\{B^{(\tau)}_{\prod_{t\in T} E_t} : E_t \in \mathcal{F}_{X_t}\}_{t\in T}$ of positive bounded linear operators $B^{(\tau)}_{\prod_{t\in T} E_t} : C(\Omega_t) \to C(\Omega_0)$ a Bayes operator at $\tau \in T$, if

(BO) for any observable $O'_\tau := (Y_\tau, \mathcal{F}_{Y_\tau}, G_\tau)_{C(\Omega_\tau)}$, there exists an observable

$$\hat{O}_0 := \left(Y_\tau \times \prod_{t \in T} X_t,\ \mathcal{F}_{Y_\tau \times \prod_{t \in T} X_t},\ \hat{F} \right)_{C(\Omega_0)}$$

which satisfies, for $E_t \in \mathcal{F}_{X_t}$ $(t \in T)$,

(i) $B^{(\tau)}_{\prod_{t \in T} E_t}(G_\tau(E'_\tau)) = \hat{F}\left(E'_\tau \times \prod_{t \in T} E_t \right)$ for $E'_\tau \in \mathcal{F}_{Y_\tau}$,

(ii) $B^{(\tau)}_{\prod_{t \in T} E_t}(1_\tau) = \prod_{t \in T} \Phi_t F_t\left(\prod_{t \in T} E_t \right)$,

(iii) $B_{E_\tau \times \prod_{t \in T \setminus \{\tau\}} X_t}(G_\tau(E'_\tau)) = \Phi_\tau(F_\tau(E_\tau) G_\tau(E'_\tau))$.

We introduce a *normalized dual Bayes operator* $R^{(\tau)}_{\prod_{t \in T} E_t} : \mathcal{M}^m(\Omega_0) \to \mathcal{M}^m(\Omega_\tau)$ defined by

$$R^{(\tau)}_{\prod_{t \in T} E_t}(\nu) = \frac{\left(B^{(\tau)}_{\prod_{t \in T} E_t} \right)^* (\nu)}{\left\| \left(B^{(\tau)}_{\prod_{t \in T} E_t} \right)^* (\nu) \right\|} \text{ for } \nu \in \mathcal{M}^m(\Omega_0), \tag{5}$$

where $\left(B^{(\tau)}_{\prod_{t \in T} E_t} \right)^*$ is the dual operator of Bayes operator $B^{(\tau)}_{\prod_{t \in T} E_t}$.

Theorem 4 (Existence theorem of Bayes operators)**.** *Let* $\{O_t\}_{t \in T}$ *be a family of observables* $O_t := (X_t, \mathcal{F}_{X_t}, F_t)_{C(\Omega_t)}$ *and* $(S_{(\omega_0)}, \{\Phi_t\}_{t \in T})$ *a general system with an initial state* $\omega_0 \in \Omega_0$. *Then, there exists a Bayes operator* $B^{(\tau)}_{\prod_{t \in T} E_t}$ *at each* $\tau \in T$.

Proof. We construct a pull-back and product observable

$$\hat{O}_0 := \left(Y_\tau \times \prod_{t \in T} X_t,\ \mathcal{F}_{Y_\tau \times \prod_{t \in T} X_t},\ \hat{F}_0 \right)_{C(\Omega_0)},$$

where

$$\hat{F}_0 = \Phi_{\tau^\circ}(F_\tau \times G_\tau) \times \prod_{t \in T \setminus \{\tau\}} \Phi_{t^\circ} F_t. \tag{6}$$

The observable $\hat{O}_0$ introduces a Bayes operator such that

$$B_{\prod_{t \in T} E_t}(G_\tau(E'_\tau)) = \Phi_\tau(F_\tau(E_\tau) G_\tau(E'_\tau)) \prod_{t \in T \setminus \{\tau\}} \Phi_{t^\circ} F_t(E_t). \tag{7}$$

□

Lemma 5. *Let* $\tau \in T$. *If* $\Phi_\tau : C(\Omega_\tau) \to C(\Omega_0)$ *is a homomorphic Markov operator, then, it holds, for any observable* $(Y_\tau, \mathcal{F}_{Y_\tau}, G_\tau)_{C(\Omega_\tau)}$, *that*

$$B^{(\tau)}_{\prod_{t \in T} E_t}(G_\tau(E'_\tau)) = \Phi_{\tau^\circ} G_\tau(E'_\tau) \widetilde{F}_0\left(\prod_{t \in T} E_t\right)(E_t \in \mathcal{F}_{X_t}). \tag{8}$$

Proof. From the condition of homomorphic Markov operators, it directly follows that

$$\Phi_\tau(F_\tau(E_\tau) G_\tau(E'_\tau)) = \Phi_{\tau^\circ} F_\tau(E_\tau) \Phi_{\tau^\circ} G_\tau(E'_\tau).$$

Then, (7) implies (8). □

4. Regression Analysis in Measurement Theory

In this section, we try to make an interpretation of regression analysis which infer the state at $\tau \in T$ after taking the measurement.

For observables

$$O_t := (X_t, \mathcal{F}_{X_t}, F_t)_{C(\Omega_t)} (t \in T) \text{ and } O'_\tau := (Y_\tau, \mathcal{F}_{Y_\tau}, G_\tau)_{C(\Omega_\tau)},$$

we let $\widetilde{O}_0$, $B^{(\tau)}_{\prod_{t \in T} E_t}$, $\hat{O}_0$ and $R^{(\tau)}_{\prod_{t \in T} E_t}$ be defined in Definition 3.

In terms of Bayes operators, we shall give an interpretation of the following issue B_1 in inferences.

B_1 : Under the information that a value $\prod_{t \in T} x_t$ measured by $M(\widetilde{O}_0, S_{(*)})$ belongs to $\prod_{t \in T} E_t$, we determine $*$ to be $\omega_0 \in \Omega$ as an inferred value in the system *S*.

The statement B_1 can be regarded as equivalent to the following:

B_2 : Under the information that a value measured by $M(\hat{O}_0, S_{(*)})$ belongs to $Y_\tau \times \prod_{t\in T} E_t$, we determine $*$ to be $\omega_0 \in \Omega$ as an inferred value of the system S.

Under the condition B_2, a value in Y_τ is distributed under the conditional probability, that is, the probability that y belongs to E' is equivalent to $P_{\prod_{t\in T} E_t}(E')$, where

$$P_{\prod_{t\in T} E_t}(E') := \frac{\hat{F}_0\left(E' \times \prod_{t\in T} E_t\right)(\omega_0)}{\tilde{F}_0\left(\prod_{t\in T} E_t\right)(\omega_0)}. \tag{9}$$

From the condition (i) and (ii) of Bayes operator, we see

$$\frac{\hat{F}_0\left(E' \times \prod_{t\in T} E_t\right)(\omega_0)}{\tilde{F}_0\left(\prod_{t\in T} E_t\right)(\omega_0)} = \frac{B^{(\tau)}_{\prod_{t\in T} E_t}(G(E'))(\omega_0)}{B^{(\tau)}_{\prod_{t\in T} E_t}(1_\tau)(\omega_0)}. \tag{10}$$

The identity $B^{(\tau)}_{\prod_{t\in T} E_t}(1_\tau)(\omega_0) = \langle (B^{(\tau)}_{\prod_{t\in T} E_t})^*(\delta_{\omega_0}), 1_\tau \rangle$ implies

$$B^{(\tau)}_{\prod_{t\in T} E_t}(1_\tau)(\omega_0) = \langle \| (B^{(\tau)}_{\prod_{t\in T} E_t})^*(\delta_{\omega_0}) \|, G(E') \rangle. \tag{11}$$

Together with (5), the identifies (10) and (11) implies

$$P_{\prod_{t\in T} E_t}(E') = \langle R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0}), G(E') \rangle. \tag{12}$$

By the identity (12), we see that the following B_3 is equivalent to B_1 and to B_2 :

B_3 : Under the information that a value $\prod_{t\in T} x_t$ measured by $M(\tilde{O}_0, S_{(*)})$ belongs to $\prod_{t\in T} E_t$, the probability that the measured value in Y belongs to $E' \in \mathcal{F}_Y$ is given by $\langle R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0}), G(E') \rangle$.

With a view of the assertion A of Axiom 1, Bayes formulation of a *posterior state* of measurements is given in the following form:

Bayes formulation. Under the condition B_1, if an initial state $*$ of the system S is inferred to be $\delta_{\omega_0} \in \mathcal{M}^p(\Omega_0)$, then a posterior state at $\tau \in T$ after taking $M(\widetilde{O}_0, S_{(*)})$ is described as $R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0}) \in \mathcal{M}^m(\Omega_\tau)$.

With using Bayes formulation, calculations of posterior states can be represented as the following theorem.

Theorem 6. *For* $\tau \in T$, *we take a family* $\{R^{(\tau)}_{\prod_{t\in T} E_t} \mid E_t \in \mathcal{F}_{X_t}\}_{t\in T}$ *of Bayes operators*. *Then*, *we infer*

(i) *a posterior state at* $\tau \in T$ *after taking* $M(\widetilde{O}_0, S_{(*)})$ *to be* $R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0})$;

(ii) *a posterior state at* $\tau \in T$ *after taking* $M(\widetilde{O}_0, S_{(*)})$ *to be* $\Phi^*_\tau(\delta_{\omega_0})$, *provided* Φ_τ *is homomorphic*.

Proof. The proof of (i) clearly follows from Bayes formulation. We shall give a proof of the statement (ii). From the notation of normalized dual Bayes operator, we see

$$\langle R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0}),\, g_\tau\rangle = \frac{\langle \delta_{\omega_0},\, B^{(\tau)}_{\prod_{t\in T} E_t}(g_\tau)\rangle}{\left\| \left(B^{(\tau)}_{\prod_{t\in T} E_t}\right)^*(\delta_{\omega_0}) \right\|} \ \text{ for } g_\tau \in C(\Omega_\tau). \tag{13}$$

The assertion Lemma 5 gives us

$$\langle R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0}),\, g_\tau\rangle = \frac{\left\langle \delta_{\omega_0},\, \widetilde{F}_0\left(\prod_{t\in T} E_t\right)\Phi_\tau g_\tau \right\rangle}{\widetilde{F}_0\left(\prod_{t\in T} E_t\right)(\omega_0)}$$

$$= \langle \Phi^*_\tau(\delta_{\omega_0}),\, g_\tau\rangle \ \text{ for } g_\tau \in C(\Omega_\tau), \tag{14}$$

which implies the desired identity $R^{(\tau)}_{\prod_{t\in T} E_t}(\delta_{\omega_0}) = \Phi^*_\tau(\delta_{\omega_0})$. □

In inference issue B_1, we rely on Gauss-Fisher's principle to determine an inferred value, so that the scheme to get an inferred value is prescribed in the following theorem.

Theorem 7. *Assume that a value measured by* $M(\tilde{O}_0, S_{(*)})$ *belongs to* $\prod_{t\in T} E_t$. *Then, an inferred value* ω_0 *to the issue* B_1 *is obtained by*

$$\tilde{F}_0\left(\prod_{t\in T} E_t\right)(\omega_0) = \max_{\omega\in\Omega_0} \tilde{F}_0\left(\prod_{t\in T} E_t\right)(\omega). \tag{15}$$

References

[1] E. B. Davies, Quantum Theory of Open Systems, Academic Press, 1976.

[2] S. Ishikawa, Mathematical Foundations of Measurement Theory, Keio University Press, Tokyo, 2006.

[3] R. V. Kadison and J. R. Ringrose, Fundamentals of Theory of Operator Algebras I, Academic Press, 1986.

[4] K. Kikuchi, Regression analysis, Kalman filter and measurement error model in measurement theory, Research Report, KSTS/RR-06/005, Keio University, 2006.

[5] K. Kikuchi, An axiomatic approach to Fisher's maximum likelihood method, Nonlinear Studies 2(18) (2011), 255-262.

[6] S. Ishikawa and K. Kikuchi, Kalman filter in quantum language, arXiv:1404.2664 [math.ST], 2014.

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