Quantum Minimax Theorem in Statistical Decision Theory

November 10, 2014
Public Version

田中冬彦(Fuyuhiko TANAKA)

Graduate School of Engineering Science, Osaka University

If you are interested in details, please see quant-ph 1410.3639 ; citations are also welcome!
Prologue:
What is Bayesian Statistics?
(Non-technical Short Version)

*It is for non-statisticians esp. researchers in quantum information.*
Mission of Statistics

To give a desirable and plausible method according to statistical models, purpose of inference (point estimation, prediction, hypothesis testing, model selection, etc.), and loss functions.

*Precise evaluation of estimation error is secondary.

Difference from Pure Math

Statistics is required to meet various needs in the world

→ Significant problems in statistics have changed according to the times

(This trivial fact does not seem to be known so much except for statisticians.)
Bayesian Statistics

Statistics admitting the usage of **prior distributions**

prior distribution = prob. dist. on the unknown parameter

(We do not care about the interpretation of this prob.)

*Interpretation heavily depends on individual researcher.*
What’s Bayesian Statistics? (2/2)

Relation to frequentist’s statistics (traditional statistics)

1. Bayesian statistics has **enlarged the framework** and satisfies various needs. (Not an alternative of frequentists’ statistics!)

2. Sticking to frequentists’ methods like MLEs is the preference to a specific prior distribution (narrow-minded!)
   (Indeed, MLEs often fail; their independence of loss is unsuitable)

3. Frequentists’ results give good **starting points to Bayesian analysis**

For details, see, e.g., C. Robert, Bayesian Choice Chap.11
Misunderstanding

Since priors are not objective, Bayesian methods should not be used for scientific result.

Objective priors (Noninformative priors) are proposed for such usage.

Bayesian statistics do not deny frequentists’ way and statistical inference based on an objective prior often agrees with frequentists’ one. (ex: MLE = MAP estimate under the uniform prior)

It seems unsolved which noninformative prior should be used.

The essence of this problem is how to choose a good statistical method with small sample. It is NOT inherent to Bayesian statistics, but has been a central issue in traditional statistics.

Frequentists and mathematicians avoid this and only consider mathematically tractable models and/or asymptotic theory (it’s large-sample approximation!)

In Bayesian statistics, this problem is transformed into the choice of noninformative prior
Further Reference

Preface in RIMS Kokyuroku 1834 (Theory and applications of statistical inference in quantum theory)

POINT

- Point out some misunderstanding of Statistics for non-experts (esp. Physicists)
- Define the word “quantum statistical inference” as an academic term

Sorry but it is written in Japanese.
CONTENTS

0. Notation
1. Introduction
2. Quantum Statistical Models and Measurements
3. Quantum Statistical Decision Problems
4. Risk Functions and Optimality Criteria
5. Main Result
6. Minimax POVMs and Least Favorable Priors
7. Summary
0. Notation
Description of Probability in QM

Quantum Mechanics (QM): described by operators in a Hilbert space

Probability: described by density op. and measurement

\[ \rho(\theta) \]


\[ \mu_\theta(dx) = \text{Tr} \rho(\theta)M(dx) \]

Density operator

Probability distributions

Measurement (POVM)

\[ M(dx) \]
Density Operators

Density operators (density matrices, state)

\[ S(\mathcal{H}) := \{ \text{all d.o.s on a Hilbert space } \mathcal{H} \} \]

\[ = \{ \rho \in L(\mathcal{H}) \mid \text{Tr } \rho = 1, \rho \geq 0 \} \]

When \( \dim \mathcal{H} = n < \infty \)

\[ \rho = U \begin{pmatrix} p_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & p_n \end{pmatrix} U^* \]

\[ \begin{array}{c}
    \text{p}_i \geq 0 \\
    \text{Tr } \rho = \sum_{j=1}^{n} p_j = 1
\end{array} \]
Measurements in QM (1/2)

Mathematical model of measurements

= **Positive Operator Valued Measure (POVM)**

**POVM (def. for finite sets)**

Sample space $\Omega = \text{Possible values for a measurement} = \{x_1, x_2, \ldots, x_k\}$

POVM $\{M_j\}_{j=1,\ldots,k} = \text{a family of linear ops. satisfying}$

$$
\begin{aligned}
M_j &:= M(\{x_j\}) \geq O \\
\sum_j M_j &= I \\
\end{aligned}
$$

Positive operator (positive semidefinite matrix)
Measurements in QM (2/2)

How to construct a probability space (Axiom of QM)

<table>
<thead>
<tr>
<th>d.o.</th>
<th>$\rho \in S(\mathcal{H})$</th>
</tr>
</thead>
<tbody>
<tr>
<td>sample space</td>
<td>$\Omega = {x_1, x_2, \ldots, x_k}$</td>
</tr>
<tr>
<td>POVM (Measurement)</td>
<td>${M_j}_{j=1,\ldots,k}$</td>
</tr>
</tbody>
</table>

Then, the probability of obtaining a measurement outcome $x_j \in \Omega$

$$p_j := \text{Tr} \rho M_j$$

* The above probability satisfies conditions (positive and summing to one) due to both definitions of d.o.s and POVMs.

$$\begin{cases} \rho \geq 0 \\ \text{Tr} \rho = 1 \end{cases} \quad \text{and} \quad \begin{cases} M_j \geq O \\ \sum_j M_j = I \end{cases}$$
1. Introduction
Quantum Bayesian Hypothesis Testing (1/2)

1. Alice chooses one alphabet (denoted by an integer) and sends one corresponding quantum state (described by a density operator)

\[ \{1, 2, \ldots, k\} \rightarrow \{\rho_1, \rho_2, \ldots, \rho_k\} \subseteq S(H) \]

2. Bob receives the quantum state and perform a measurement and guesses the alphabet Alice really sent to him. The whole process is described by a POVM \( \{M_j\}_{j=1,2,\ldots,k} \)

3. The proportion of alphabets is given by a distribution (called a prior distribution) \( \pi_1 + \cdots + \pi_k = 1, \pi_1, \ldots, \pi_k \geq 0 \)
Quantum Bayesian Hypothesis Testing (2/2)

* The probability that Bob guesses $j$ when Alice sends $i$.

$$p(B = j \mid A = i) = \text{Tr} \ \rho_i M_j$$

* Bob’s loss (estimation error) when he guesses $j$ while Alice sends $i$.

$$w(i, j)$$

* Bob’s risk for the $i$-th alphabet

$$R_M(i) = \sum_{j=1}^{k} w(i, j) p(B = j \mid A = i)$$

* Bob’s average risk (his task is to minimize by using a good POVM)

$$\sum_{i=1}^{k} \pi_i R_M(i)$$
Summary of Q-BHT

Choose one alphabet and sends one quantum state

\[ \{1, 2, \ldots, k\} \rightarrow \{\rho_1, \rho_2, \ldots, \rho_k\} \subseteq S(H) \]

Proportion of each alphabet
(prior distribution)

\[ \pi_1 + \cdots + \pi_k = 1, \pi_k \geq 0 \]

Guess the alphabet by choosing a measurement (POVM)

\[ \{M_j\}_{j=1, 2, \ldots, k} \]

Average risk

\[ r_{\pi, M} := \sum_{i=1}^{k} \pi_i R_M(i) \]

\[ R_M(i) = \sum_{j=1}^{k} w(i, j) \text{Tr} \rho_i M_j \]

Bayes POVM w.r.t. \( \pi \) = POVM minimizing the average risk
Minimax POVM

- Chooses one alphabet and sends one quantum state

\[ \{1, 2, \ldots, k\} \rightarrow \{\rho_1, \rho_2, \ldots, \rho_k\} \subseteq S(H) \]

- Guess the alphabet by choosing a measurement (POVM)

\[ \{M_j\}_{j=1, 2, \ldots, k} \]

Bob has no prior information \(\rightarrow\) Minimize the worst-case risk

The worst-case risk

\[ r^*_M := \sup_i R_M(i) \]

Minimax POVM = POVM minimizing the worst-case risk
Quantum Minimax Theorem (Simple Ver.)

Choose one alphabet and sends one quantum state
\[ \{1, 2, \ldots, k\} \rightarrow \{\rho_1, \rho_2, \ldots, \rho_k\} \subseteq S(H) \]

Guess the alphabet by choosing a measurement (POVM)
\[ \{M_j\}_{j=1,2,\ldots,k} \]

Theorem (Hirota and Ikehara, 1982)
\[
\inf_M \sup_i R_M(i) = \sup \inf_{\pi} \sum_{i=1}^{k} \pi_i R_M(i)
\]

Minimax POVM agrees with the Bayes POVM w.r.t. the worst case prior.

*The worst-case prior to Bob is called a least favorable prior.*
Quantum states

\[
\begin{align*}
\rho_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \rho_2 &= \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \rho_3 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{align*}
\]

0-1 loss \( w(i, j) = 1 - \delta_{ij} \)

Minimax POVM

\[
\begin{align*}
M_1 &= \frac{1}{2} \begin{pmatrix} 1+1/\sqrt{2} & -1/\sqrt{2} & 0 \\ -1/\sqrt{2} & 1-1/\sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, & M_2 &= \frac{1}{2} \begin{pmatrix} 1-1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & 1+1/\sqrt{2} & 0 \\ 0 & 0 & 0 \end{pmatrix}, & M_3 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.
\end{align*}
\]

*This is a counterexample of Theorem 2 in Hirota and Ikehara (1982), where their proof seems not mathematically rigorous.*
Quantum states

\[
\rho_1 = \begin{pmatrix}
1 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 0
\end{pmatrix}, \quad \rho_2 = \begin{pmatrix}
1/2 & 1/2 & 0 \\
1/2 & 1/2 & 0 \\
0 & 0 & 0
\end{pmatrix}, \quad \rho_3 = \begin{pmatrix}
0 & 0 & 0 \\
0 & 0 & 0 \\
0 & 0 & 1
\end{pmatrix}.
\]

0-1 loss \( w(i, j) = 1 - \delta_{ij} \)

LFP

\[
\pi_{LF}(1) = \pi_{LF}(2) = 1/2, \quad \pi_{LF}(3) = 0
\]

Minimax POVM is constructed as a Bayes POVM w.r.t. LFP.

Important Implication

When completely unknown, it is not necessarily wise to find an optimal POVM with the uniform prior.

(Although Statisticians already know this fact long decades ago…)
Rewritten in Technical Terms

Nature

Parameter space \( \Theta = \{1, 2, \ldots, k\} \)
\( \{\rho_\theta\} \subseteq S(H) \)

Experimenter and Statistician

Decision space \( U = \{1, 2, \ldots, k\} \)
\( \{M_u\}_{u \in U} \)

Loss function
Risk function
\( w(\theta, u) \)
\( R_M(\theta) = \sum_{u \in U} w(\theta, u) \text{Tr} \rho_\theta M_u \)

Theorem

\[
\inf_{M} \sup_{\theta} R_M(\theta) = \sup_{\pi} \inf_{M} \sum_{\theta \in \Theta} \pi(\theta) R_M(\theta)
\]
Main Result (Brief Summary)

Quantum Minimax Theorem

\[
\inf_{M \in P_0} \sup_{\theta \in \Theta} R_M(\theta) = \sup_{\pi \in P(\Theta)} \inf_{M \in P_0} \int_{\Theta} R_M(\theta) \pi(d\theta)
\]

where \( R_M(\theta) := \int_U w(\theta, u) \text{Tr} \rho(\theta) M(du) \)

\[
\text{cf)} \quad \inf_{\delta \in D} \sup_{\theta \in \Theta} R_\delta(\theta) = \sup_{\pi \in P(\Theta)} \inf_{\delta \in D} \int_{\Theta} R_\delta(\theta) \pi(d\theta)
\]

Conditions, assumptions, and essence of proof are explained.
# Previous Works

<table>
<thead>
<tr>
<th>Author</th>
<th>Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wald (1950)</td>
<td>Statistical Decision Theory</td>
</tr>
<tr>
<td>Holevo (1973)</td>
<td>Quantum Statistical Decision Theory</td>
</tr>
</tbody>
</table>

Recent results and applications (many!)

- Kumagai and Hayashi (2013)
- Guta and Jencova (2007)
- Hayashi (1998)

---

Le Cam (1964)

(Classical) Minimax Theorem in statistical decision theory  
First ver. is given by Wald (1950)

---

We show

**Quantum Minimax Theorem** in quantum statistical decision theory
2. Quantum Statistical Models and Measurements
Formal Definition of Statistical Models

(naïve ) Quantum statistical model = A parametric family of density operators

Ex. \[ \rho(\theta) = \frac{1}{2} \begin{pmatrix} 1 + \theta_3 & \theta_1 - i\theta_2 \\ \theta_1 + i\theta_2 & 1 - \theta_3 \end{pmatrix} \]

\[ 0 \leq (\theta_1)^2 + (\theta_2)^2 + (\theta_3)^2 \leq 1 \]

\[ \theta_1, \theta_2, \theta_3 \in \mathbb{R} \]

\[ \rho(r, s; \theta) = \langle r | \rho(\theta) | s \rangle = \psi(r - \theta)\overline{\psi}(s - \theta) \quad \theta \in \mathbb{R} \]

Basic Idea (Holevo, 1976)

A quantum statistical model is defined by a map

\[ \rho : \Theta \rightarrow L_1(H) \]

\[ \theta \mapsto \rho(\theta) \]

Next, we see the required conditions for this map
Regular Operator-Valued Function (1/2)

Definition

\( \Theta \) Locally compact metric space
\( K \subseteq \Theta \) compact set

\( K_\delta := \{ (\theta, \eta) \in K \times K : d(\theta, \eta) < \delta \} \quad \forall \delta > 0 \)

\( L_1(H) \) trace-class operators on a Hilbert space

For a map \( T : \Theta \to L_1(H) \), we define

\[ \omega_T(K_\delta) := \inf \left\{ \|X\|_1 : -X \leq T(\theta) - T(\eta) \leq X, \forall (\theta, \eta) \in K_\delta \right\} \]

Definition

An operator-valued function \( T \) is regular

\[ \lim_{\delta \to 0} \omega_T(K_\delta) = 0 \quad \forall K \subseteq \Theta \]
Résumé Opérateur-Valué Function (2/2)

**Remark 1**
\[
\lim_{\delta \to 0} \omega_T (K_\delta) = 0 \quad \forall K \subseteq \Theta
\]

Uniformly continuous w.r.t. trace norm on every compact set
\[
\sup_{(\theta, \eta) \in K_\delta} \left\| T(\theta) - T(\eta) \right\|_1 \to 0 \quad \text{as} \quad \delta \to 0
\]

Converse does not hold if \( \dim H = \infty \)

**Remark 2**

The regularity assures that the following operator-valued integral is well-defined
\[
\int_{\Theta} f(\theta) T(\theta) \pi \left( d \theta \right)
\]

for every \( f \in C_0(\Theta) \) and \( \pi \in P(\Theta) \)

\[
\omega_T (K_\delta) := \inf \left\{ \left\| X \right\|_1 : -X \leq T(\theta) - T(\eta) \leq X, \forall (\theta, \eta) \in K_\delta \right\}
\]
Quantum Statistical Models

Definition

\[ S(H) := \{ X \in L(H) : \text{Tr} X = 1, X \geq 0 \} \]

\( \rho : \Theta \rightarrow S(H) \) is called a quantum statistical model if it satisfies the conditions below.

Conditions

1. Identifiability (one-to-one map) \( \rho(\theta_1) \neq \rho(\theta_2), \theta_1 \neq \theta_2 \)

2. Regularity \( \lim_{\delta \rightarrow 0} \omega_T(K_\delta) = 0 \quad \forall K \subseteq \Theta \)

\[ \uparrow \text{Necessary for our arguments} \]
Measurements (POVMs)

\( U \)  Decision space (locally compact space)
\( A(U) \)  Borel sets (the smallest sigma algebra containing all open sets)

**Definition**

Positive-operator valued function \( M \) is called a POVM if it satisfies

\[
M(B) \geq 0 \quad B \in A(U)
\]

\[
M\left( \bigcup_{j=1}^{\infty} B_j \right) = \sum_{j=1}^{\infty} M(B_j) \quad B_1, B_2, \ldots \in A(U) \quad B_i \cap B_j = \emptyset, i \neq j
\]

\[
M(U) = I
\]

\( Po(U) \)  All POVMs over \( U \)
**Born Rule**

**Axiom of Quantum Mechanics**

For the quantum system described by $\rho$
and a measurement described by a POVM $M(dx)$
measurement outcome is distributed according to

$$x \sim \text{Tr} \rho M(dx) = \mu_{\rho}(dx)$$
3. Quantum Statistical Decision Problems
Basic Setting

**Situation**

For the quantum system specified with the unknown parameter, experimenters extract information for some purposes.

\[ \rho(\theta) \xrightarrow{\text{POVM}} x \xrightarrow{\mu(\theta)(dx)} M(dx) \xrightarrow{\text{POVM}} a(x) \]

**Typical sequence of task**

1. To choose a measurement (POVM)
2. To perform the measurement over the quantum system
3. To take an action \( a(x) \) among choices based on the measurement outcome \( x \)

formally called a *decision function*
Decision Functions

Example

- Estimate the unknown parameter \( a(x) = \frac{x_1 + \cdots + x_n}{n} \)

- Estimate the unknown d.o. rather than parameter

\[
\rho = \begin{pmatrix}
a_1(x) & a_2(x) & 0 \\
a_2^*(x) & a_3(x) & 0 \\
0 & 0 & 1 - a_1(x) - a_3(x)
\end{pmatrix}
\]

- Validate the entanglement/separability \( a(x) = 0,1 \)

- Construct confidence region (credible region) \([a_L(x), a_R(x)]\)
Remarks for Non-statisticians

1. If the quantum state is *completely known*, then the distribution of measurement outcome is also known and the best action is chosen.

2. Action has to be made from *finite data*.

3. Precise estimation of the parameter is only a typical example.
Loss Functions

Performance of decision functions is compared by adopting the loss functions (smaller is better).

Ex: Parameter Estimation

Quantum Statistical Model
\[ \{ \rho(\theta) : \theta \in \Theta \subseteq \mathbb{R}^m \} \]

Action space \( \Theta \)

Loss function (squared error)
\[ w(\theta, a) = |\theta - a|^2 \]

Later we see the formal definition of the loss function.
Measurements Over Decisions

Basic Idea

From the beginning, we only consider POVMs over the action space.
Quantum Statistical Decision Problems

Formal Definitions

\( \Theta \) Parameter space (locally compact space)

\( \rho(\theta) \) Quantum statistical model

\( U \) *Decision space (locally compact space)

\( w : \Theta \times U \to \mathbb{R} \cup \{+\infty\} \) Loss function (lower semi continuous**; bounded below)

The triplet \((\rho(\theta), U, w)\) is called a *quantum statistical decision problem*.

(*we follow Holevo’s terminology instead of saying “action space”.

** we impose a slightly stronger condition on the loss than LeCam, Holevo.*
4. Risk Functions and Optimality Criteria
Comparison of POVMs

The loss (e.g. squared error) depends on both unknown parameter $\theta$ and our decision $u$, which is a random variable.

\[ u \sim \mu_\theta(du) = \text{Tr} \rho(\theta)M(du) \]

In order to compare two POVMs, we focus on the average of the loss w.r.t. the distribution of $u$.

\[ \mathbb{E}[w(\theta, u)] \]

\[ \mathbb{E}'[w(\theta, u')] \]

Compared at the same $\theta$
Risk Functions

**Definition**

The *risk function* of $M \in \mathcal{P}o(U)$

$$R_M(\theta) := \int_U w(\theta, u) \text{Tr} \rho(\theta)M(d\, u) = E_\theta[w(\theta, u)]$$

**Remarks for Non-statisticians**

1. Smaller risk is better.

2. Generally there exists no POVM achieving the uniformly smallest risk among POVMs.

Since $R_M(\theta)$ depends on the unknown parameter, we need additional optimality criteria.
A POVM is said to be *minimax* if it minimizes the supremum of the risk function, i.e.,

\[
\sup_{\theta \in \Theta} R_{M^*}(\theta) = \inf_{M \in Po(U)} \sup_{\theta \in \Theta} R_M(\theta)
\]

**Historical Remark**

Bogomolov showed the existence of a minimax POVM (Bogomolov 1981) in a more general framework.

(Description of quantum states and measurements is different from recent one.)
Optimality Criteria (2/2)

$P(\Theta)$  All probability distributions (defined on Borel sets)

In Bayesian statistics, $\pi \in P(\Theta)$ is called a prior distribution.

**Definition**

The average risk of $M \in Po (U)$ w.r.t. $\pi$

$$r_{\pi,M} := \int_{\Theta} R_M (\theta) \pi (d \theta)$$

A POVM is said to be Bayes if it minimizes the average risk, i.e.,

$$r_{\pi,M} = \inf_{M \in Po (U)} r_{\pi,M}$$

**Historical Remark**

Holevo showed the existence of a Bayes POVM (Holevo 1976; see also Ozawa, 1980) in a more general framework.
Ex. of Loss Functions and Risk Functions

Parameter Estimation

\( \{ \rho(\theta) : \theta \in \Theta \subseteq \mathbb{R}^m \} \)

\( U = \Theta \quad w(\theta, u) = |\theta - u|^2 \)

\( R_M(\theta) = \int_U |\theta - u|^2 \text{Tr} \rho(\theta) M(d\ u) = E_\theta[|\theta - u|^2] \)

Construction of Predictive Density Operator

\( \{ \rho(\theta)^\otimes n : \theta \in \Theta \subseteq \mathbb{R}^p \} \)

\( U = S(H^\otimes m) \quad w(\theta, u) = D(\rho(\theta)^\otimes m \parallel u) \)

\( R_M(\theta) = \int_U D(\rho(\theta)^\otimes m \parallel u) \text{Tr} \rho(\theta)^\otimes m M(d\ u) \)
5. Main Result
Main Result (Quantum Minimax Theorem)

Theorem (*): Let $\Theta$ and $U$ be a compact metric space.

Then the following equality holds for every quantum statistical decision problem.

$$\inf_{M \in Po(U)} \sup_{\theta \in \Theta} R_M(\theta) = \sup_{\pi \in P(\Theta)} \inf_{M \in Po(U)} \int_{\Theta} R_M(\theta) \pi(d\theta)$$

where the risk function is given by $R_M(\theta) := \int_{U} w(\theta, u) \text{Tr} \rho(\theta) M(du)$

Corollary: For every closed convex subset $Q \subseteq Po(U)$ the above assertion holds.

*see quant-ph 1410.3639 for proof
Key Lemmas For Theorem

**Compactness of the POVMs**

If the decision space \( U \) is compact, then \( Po(U) \) is also compact.

**Equicontinuity of risk functions**

Loss function \( w : \Theta \times U \rightarrow \mathbb{R} \)

If \( w \) is bounded and continuous, then

\[
F := \{ R_M : M \in Po(U) \} \subseteq C(\Theta)
\]

is (uniformly) equicontinuous.

The equicontinuity implies the compactness of \( F \subseteq C(\Theta) \) under the uniform convergence topology.

We show main theorem by using Le Cam’s result with both lemmas.

However, not a consequence of their previous old results.
6. Minimax POVMs and Least Favorable Priors
Previous Works (LFPs)

Wald (1950) Statistical Decision Theory

Jeffreys (1961) Jeffreys prior

Bernardo (1979, 2005) Reference prior; reference analysis

Komaki (2011) Latent information prior

Tanaka (2012) MDP prior (Reference prior for pure-states model)

Objective priors in Bayesian analysis are indeed least favorable priors.
Main Result (Existence Theorem of LFP)

Definition

If a prior achieves the supremum, i.e., the following holds

\[ \inf_{M \in \text{Po}(U)} \int_{\Theta} R_M(\theta) \pi_{LF} (d\theta) = \sup_{\pi \in P(\Theta)} \inf_{M \in \text{Po}(U)} \int_{\Theta} R_M(\theta) \pi (d\theta) \]

The prior \( \pi_{LF} \) is called a least favorable prior (LFP).

where \( R_M(\theta) := \int_U w(\theta, u) \text{Tr} \rho(\theta) M (du) \)

Theorem

Let \( \Theta \) and \( U \) be a compact metric space.
\( w \): continuous loss function.

Then, for every decision problem, there exists a LFP.
Remark
Even on a compact space, a bounded lower semicontinuous (and not continuous) loss function does not necessarily admit a LFP.

Example \( \Theta = [0,1] \)

\[
R_M(\theta) = L(\theta, u) = \begin{cases} 1 - \theta \\ 0 \end{cases}
\]

\[
1 = \sup_{\theta \in [0,1]} R_M(\theta) = \sup_{\pi} \int_{[0,1]} R_M(\theta) \pi(\text{d} \theta)
\]

But for every prior, \( 1 > \int_{[0,1]} R_M(\theta) \pi(\text{d} \theta) \)
Corollary

Let $\mathcal{H}$ and $U$ be a compact metric space. If a LFP exists for a quantum statistical decision problem, every minimax POVM is a Bayes POVM with respect to the LFP. In particular, if the Bayes POVM w.r.t. the LFP is unique, then it is minimax.

Corollary 2

For every closed convex subset $Q \subseteq Po(U)$, the above assertion holds.

Much of theoretical results are derived from quantum minimax theorem.
7. Summary
Discussion

1. We show quantum minimax theorem, which gives theoretical basis in quantum statistical decision theory, in particular in objective Bayesian analysis.

2. For every closed convex subset of POVMs, all of assertions still hold.

Our result has also practical meanings for experimenters.

Future Works

- Show other decision-theoretic results and previous known results by using quantum minimax theorem

- Propose a practical algorithm of finding a least favorable prior

- Define geometrical structure in the quantum statistical decision problem

If you are interested in details, please see quant-ph 1410.3639; citations are also welcome!