Mathematica Slovaca

Ewa Czkwianianc

On the factorization criterion for quantum statistics

Mathematica Slovaca, Vol. 55 (2005), No. 3, 373--380

Persistent URL: http://dml.cz/dmlcz/136919

Terms of use:

© Mathematical Institute of the Slovak Academy of Sciences, 2005

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.

This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* http://project.dml.cz

Math. Slovaca, **55** (2005), No. 3, 373–380

Mathematica Slovaca
©2005
Mathematical Institute
Slovak Academy of Sciences

ON THE FACTORIZATION CRITERION FOR QUANTUM STATISTICS

EWA CZKWIANIANC

(Communicated by Anatolij Dvurečenskij)

ABSTRACT. We discuss some fundamental statistical notions for a quantum system in the simplest case when only a finite number of different states is possible and all states are pure. Thus our non-commutative statistical space is described by a Hilbert space H with a finite system of unit vectors $(\varphi_\theta,\ \theta = 1,\dots,n)$. Each statistic is described by a self-adjoint operator S in H. We define the sufficiency of S in some natural way motivated by the classical factorization criterion. We show that a generalization of Blackwell-Rao theorem cannot be obtained. Moreover, some linear space containing all non-biased estimators with minimal variance is described in the paper. In fact, this space is huge, compared to the commutative case.

1. Introduction

Our concept of a non-commutative sufficient statistic is motivated by the following observation. Let a system of distributions in a classical statistical space be given by densities $(f_\theta,\ \theta\in\Theta)$ with respect to a given measure μ on a sample space $(X,\mathcal{F})$. Let S be any statistic on $(X,\mathcal{F})$. The sufficiency of S can be described using geometry of the space $L^2 = L^2(X,\mathcal{F},\mu)$. The densities f_θ can be described by unit vectors $\varphi_\theta = f_\theta^{\frac{1}{2}}$ in L^2. A real statistic S can be identified with a self-adjoint operator in L^2. The classical factorization criterion for sufficiency of S can be formulated as follows (see [7]).

THEOREM 1.1. *The statistic S is sufficient if and only if there exist a function $\chi\in L^2$ and Borel functions $\Phi_\theta\colon\mathbb{R}\to\mathbb{R}$ satisfying*

$$\varphi_\theta = \sqrt{f_\theta} = \Phi_\theta(S)\chi\,, \qquad \theta\in\Theta\,.$$

In the simplest quantum case, probability distributions are described by unit vectors in a Hilbert space H. Random variables are replaced by self-adjoint

2000 Mathematics Subject Classification: Primary 81P18, 81Q10, 62B05.
Keywords: quantum measurement, pure states, non biased estimator, sufficient statistic.

operators, called *observables*, and a distribution of an observable T in a state $\varphi \in H$, $\|\varphi\| = 1$, is given by the formula $p_{T_\varphi}(Z) = \langle e_T(Z)\varphi, \varphi\rangle$, $Z \in \operatorname{Borel}\mathbb{R}$, with $e_T(\cdot)$ being the spectral measure of the operator T. Thus the following definition seems natural.

DEFINITION 1.2. Let $(\varphi_\theta,\ \theta \in \Theta)$ be a set of unit vectors in a separable Hilbert space H. A self-adjoint operator $T = \int_{-\infty}^{\infty} \lambda\, e_T(\mathrm{d}\lambda)$ is a *sufficient statistic* if there exists a vector $\chi \in H$, $\|\chi\| = 1$, and Borel functions Φ_θ, $\theta \in \Theta$, satisfying

$$\varphi_\theta = (\Phi_\theta T)\chi\,, \qquad \theta \in \Theta\,.$$

Formally, Definition 1.2 gives an exact generalization of classical sufficient statistic, via factorization criterion. We add some elementary discussion of the new notion in the language of quantum measurement and relative information. Assume that a quantum system is in a state $\Phi_\theta(S)\chi$. We measure an observable S, obtaining a value s, and *after this* we obtain a value t measuring an observable T. Let

$$S = \int_{\mathbb{R}} \lambda\, e_S(\mathrm{d}\lambda)\,, \qquad T = \int_{\mathbb{R}} \lambda\, e_T(\mathrm{d}\lambda)$$

be the spectral representations of these self-adjoint operators. A non-commuting pair $ST \neq TS$ is allowed. According to the classical von Neumann theory of measurement, the pair (s,t) belongs to $A \times B$, $A, B \in \operatorname{Borel}\mathbb{R}$, with probability

$$p_{\theta,(S,T)}(A \times B) = \int_A \langle e_T(B)\Phi_\theta(s)e_S(\mathrm{d}s)\chi,\ \Phi_\theta(s)e_S(\mathrm{d}s)\chi\rangle\,.$$

A conditional probability of the event $(t \in B)$ given a value s can be obtain as a Radon Nikodym derivative

$$\frac{\mathrm{d}p_{\theta,(S,T)}(\cdot \times B)}{\mathrm{d}p_{\theta,(S,T)}(\cdot \times \mathbb{R})} = \frac{\langle e_T(B)e_S(\mathrm{d}s)\chi,\ e_S(\mathrm{d}s)\chi\rangle}{\langle e_S(\mathrm{d}s)\chi,\ e_S(\mathrm{d}s)\chi\rangle}$$

and *does not depend on* θ. This leads to the following remark.

Remark 1.3. Let S be a sufficient statistic according to Definition 1.2. If any observable T is measured after the measurement of S, then the value t of T does not contain any new information on the parameter θ if only s is known.

The opposite implication is also rather obvious and can be obtained by contradiction.

Remark 1.4. Assume that T is measured after S. Suppose that any conditional distribution of the value t of T given the value s of S does not depend on θ. Then S is a sufficient statistic according to Definition 1.2.

The problem of the existence of a sufficient statistic for a class of states $(\varphi_1, \dots, \varphi_n)$ can be also explained.

THEOREM 1.5. *For a given system of states $(\varphi_1, \dots, \varphi_n)$ there exists a sufficient statistic if and only if*

$$\langle \varphi_i, \varphi_j \rangle \in \mathbb{R} \qquad \text{for any} \quad i, j \leq n . \tag{1}$$

Proof. The necessity of (1) is obvious. Assume that the condition (1) is satisfied and let $(e_1, \dots, e_n)$ be an orthonormal system obtained by a Gramm-Schmidt orthogonalization of $\varphi_1, \dots, \varphi_n$. It is enough to use $e_1, \dots, e_n$ as eigenvectors of S. □

Obviously a sufficient statistic is not unique.

Remark 1.6. If for a system of linearly independent states $\varphi_1, \dots, \varphi_n$, $n \geq 2$, there exists a sufficient statistic S, then one can find another sufficient statistic T and moreover $ST \neq TS$. Indeed, let $(e_1, \dots, e_n)$, $(f_1, \dots, f_n)$ be any different orthonormal systems such that coefficients of φ_i in (e_j) as well as in (f_j) are real for $1 \leq i \leq n$. Then always (e_j) and (f_j) can be taken as eigenvectors of S and T, respectively.

2. Sufficient statistics and non-biased estimators

As usually, see [4] or [2], we say that S is a non biased estimator of a parametric function $\gamma \colon \theta \mapsto \mathbb{R}$ if

$$\langle S\varphi_\theta, \varphi_\theta \rangle = \gamma(\theta) , \qquad \theta \in \Theta .$$

The variance of a non biased estimator is given by the formula

$$D_\theta^2 S = \big\langle \big(S - \gamma(\theta)\big)^2 \varphi_\theta, \varphi_\theta \big\rangle = \langle S^2 \varphi_\theta, \varphi_\theta \rangle - \gamma(\theta)^2 , \qquad \theta \in \Theta .$$

Obviously, S should be non biased and of minimal variance to be used as a good estimator of $\gamma(\cdot)$.

Let S be a sufficient statistic. Roughly speaking we show that a linear space B spanned by all good estimators has to be drastically larger then the space of all Borel functions of S, see examples below. Thus a generalization of the classical Blackwell-Rao theorem cannot be obtained ([7]). A complete description of such linear space B is given in this section.

The literature on quantum statistics is now extremally diverse from abstract theory of operator algebras to specialized applications. The well-known concept of sufficiency is connected with the theory of conditional expectation given a subalgebra of an operator algebra M. Namely, one assumes that this conditional expectation does not depend on a state (see [3], [8]). Our concept seems to be different. It is more elementary and connected with classical descriptions of quantum measurement ([6], [1], [5]). The estimation theory is widely developed also for continuous parameters (see monographs [4], [2]). But a general notion of sufficiency is not discussed there.

We denote by $\hat{e}$ the one dimensional projection $x \mapsto \langle x, e\rangle e$ given by a unit vector $e \in H$. The indicator of any Borel set Z is denoted by 1_Z.

EXAMPLE 2.1. There exists a parametric function $\gamma(\theta)$ with a non biased estimator for which each estimator being a function of a given sufficient statistic is biased.

Let (e_1, e_2) be an orthonormal basis in a Hilbert space $H = \mathbb{C}^2$. For the system of states $\varphi_1 = e_1$, $\varphi_2 = e_2$, $\varphi_3 = (e_1 + e_2)/\sqrt{2}$, and for the statistic $S = \hat{e}_1$, we have $\varphi_i = \Phi_i(S)\chi$, $i = 1, 2, 3$, with $\Phi_1 = \sqrt{2} \cdot 1_{\{1\}}$, $\Phi_2 = \sqrt{2} \cdot 1_{\{0\}}$, $\Phi_3 = 1$, $\chi = \varphi_3$. Each function $\gamma\colon \{1,2,3\} \to \mathbb{R}$ has a non biased estimator T given by the matrix

$$T = \begin{bmatrix} \gamma(1) & \gamma(3) - \frac{1}{2}(\gamma(1) + \gamma(2)) \\ \gamma(3) - \frac{1}{2}(\gamma(1) + \gamma(2)) & \gamma(2) \end{bmatrix}.$$

On the other hand, γ has a non biased estimator being a function of $S = \hat{e}_1$ if and only if $\gamma(3) = \frac{1}{2}\big(\gamma(1) + \gamma(2)\big)$.

EXAMPLE 2.2. There exists a sufficient statistic S with the following property. For any statistic T, a function $\Phi\colon \mathbb{R} \to \mathbb{R}$ can be found so that T and $\Phi(S)$ are non biased estimators of the same parametric function. But, for some T, the variance of $\Phi(S)$ has to be drastically larger than the variance of T.

Let $H = \mathbb{C}^3$, $S = \sum\limits_{i \le 3} i\hat{e}_i$, with the system of states $\varphi_1 = e_1$, $\varphi_2 = e_2$, $\varphi_3 = \sqrt{\frac{1}{2} - \frac{\varepsilon}{2}}\,(e_1 + e_2) + \sqrt{\varepsilon} e_3$, $0 < \varepsilon < 1$. The function of S being a non biased estimator of the index θ, for all states $\{\varphi_\theta : \theta = 1,2,3\}$, should be of the form

$$\Phi(S) = \hat{e}_1 + 2\hat{e}_2 + \Big(\frac{3}{2\varepsilon} + \frac{3}{2}\Big)\hat{e}_3 .$$

In the state φ_3, the variance of $\Phi(S)$ is given by

$$D^2_{\varphi_3}\Phi(S) = \frac{9}{4\varepsilon} - 2 - \frac{1}{4}\varepsilon ,$$

and tends to infinity for $\varepsilon \to 0$. Obviously,

$$T = \begin{bmatrix} 1 & \frac{3}{2}\frac{1+\varepsilon}{1-\varepsilon} & 0 \\ \frac{3}{2}\frac{1+\varepsilon}{1-\varepsilon} & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is also a non biased estimator of the parameter, and $D^2_{\varphi_3}T = -2 + 2\varepsilon + \frac{9}{4}\frac{(1+\varepsilon)^2}{1-\varepsilon} \to \frac{1}{4}$ as $\varepsilon \to 0$.

3. On minimal spaces of non biased estimators

The phenomenon pointed out in the former examples is not accidental, it proves to be a general property of quantum systems.

Proposition 3.1. *Let $TS \neq ST$ for some self-adjoint operators S, T. There exists states φ_1, φ_2, φ_3 such that S is a sufficient statistic, but each function $\Phi(S)$ is a biased estimator of the function $\gamma(i) = \langle T\varphi_i, \varphi_i\rangle$, $i = 1, 2, 3$.*

Proof. Let $S = \int_{-\infty}^{\infty} \lambda\, e(\mathrm{d}\lambda)$ be a spectral decomposition of S. It is sufficient to choose unit vectors φ_1, φ_2 satisfying $\varphi_1 \in e(Z_1)H$, $\varphi_2 \in e(Z_2)H$ for some sets Z_1, Z_2, $Z_1 \cap Z_2 = \emptyset$, and such that $\Re\langle T\varphi_1, \varphi_2\rangle \neq 0$. Taking additionally $\varphi_3 = \chi = (\varphi_1 + \varphi_2)/\sqrt{2}$ we have

$$\varphi_i = \sqrt{2}\cdot 1_{Z_i}(S)\chi, \qquad i = 1, 2,$$
$$\varphi_3 = 1\cdot\chi.$$

For any Borel function Φ the equalities

$$\langle\Phi(S)\varphi_i, \varphi_i\rangle = \langle T\varphi_i, \varphi_i\rangle, \qquad i = 1, 2,$$

imply

$$\begin{aligned}\langle\Phi(S)\varphi_3, \varphi_3\rangle &= \langle(1_{Z_1}\Phi)(S)\varphi_3, \varphi_3\rangle + \langle(1_{Z_2}\Phi)(S)\varphi_3, \varphi_3\rangle \\ &= \frac{1}{2}\langle\Phi(S)\varphi_1, \varphi_1\rangle + \frac{1}{2}\langle\Phi(S)\varphi_2, \varphi_2\rangle = \frac{1}{2}\langle T\varphi_1, \varphi_1\rangle + \frac{1}{2}\langle T\varphi_2, \varphi_2\rangle \\ &\neq \frac{1}{2}\langle T\varphi_1, \varphi_1\rangle + \frac{1}{2}\langle T\varphi_2, \varphi_2\rangle + \mathrm{Re}\langle T\varphi_1, \varphi_2\rangle = \langle T\varphi_3, \varphi_3\rangle.\end{aligned}$$

□

Let now a system of states $\{\varphi_\theta : \theta \in \Theta\}$ and a sufficient statistic S be fixed. Two questions are of special statistical importance. How can the minimal space M of observables be characterized to satisfy

$$(\forall T = T^*)(\exists U \in M)(\forall\theta \in \Theta)\big(\langle T\varphi_\theta, \varphi_\theta\rangle = \langle U\varphi_\theta, \varphi_\theta\rangle\big)\,?$$

And also how can one characterize such minimal space N that for any observable T there exists an observable $U \in N$ satisfying

$$(\forall\theta \in \Theta)\big(\langle T\varphi_\theta, \varphi_\theta\rangle = \langle U\varphi_\theta, \varphi_\theta\rangle\big)$$

and

$$D^2_{\varphi_\theta} U \le D^2_{\varphi_\theta} T \qquad \text{for all} \qquad \theta \in \Theta$$

(or

$$D^2_{\varphi_\theta} U < D^2_{\varphi_\theta} T \qquad \text{for some} \quad \theta \in \Theta\,)\,?$$

The questions lead to tedious details and we postpone the discussion to subsequent papers. We finish the section with rather expected observation.

PROPOSITION 3.2. *Let P be the orthogonal projection on the closure of the space* $\operatorname{lin}\{\varphi_\theta : \theta \in \Theta\}$ *spanned by a given system of states. Then for any observable T the statistic $U = PTP$ satisfies*

$$\begin{aligned} \langle U\varphi_\theta, \varphi_\theta\rangle &= \langle T\varphi_\theta, \varphi_\theta\rangle\,, \qquad && \theta \in \Theta\,, \qquad (2 \\ D^2_{\varphi_\theta} U &\le D^2_{\varphi_\theta} T\,, && \theta \in \Theta\,. \end{aligned}$$

P r o o f . The equality (2) is obvious and

$$\langle U^2\varphi_\theta, \varphi_\theta\rangle = \langle PTP\varphi_\theta, PTP\varphi_\theta\rangle = \langle PT\varphi_\theta, PT\varphi_\theta\rangle = \|PT\varphi_\theta\|^2 \le \|T\varphi_\theta\|^2\,.$$

□

4. Comparison with non-commutative sufficient subalgebras

From the point of view of statistical information a statistic $S = S^*$ can be identified with a commutative von Neumann algebra generated by S. Thus Definition 1.2 can be formulated as follows.

A commutative subalgebra $\mathcal{N}$ of $B(H)$ is a sufficient statistic for a system of states $\varphi_1, \dots, \varphi_n$ if there exists a vector $\chi \in H$ and self-adjoint elements $S_1, \dots, S_n$ in $\mathcal{N}$ satisfying $\varphi_i = S_i\chi$, $1 \le i \le n$.

On the other hand, the idea of H i a i et al., based on conditional expectation. leads to the following definition.

DEFINITION 4.1. A commutative subalgebra $\mathcal{N} \subset B(H)$ is a *Hiai-Ohya-Tsukada sufficient* (*HOT-sufficient*) *statistic* for states $\varphi_1, \dots, \varphi_n$ if for any $i \le n$ there exists a conditional expectation $\mathbb{E}_i$ given $\mathcal{N}$ preserving state φ_i, and $\mathbb{E}_i$ does not depend on i.

As usually, a conditional expectation preserving state φ is defined as a norm one idempotent $\mathbb{E}$ with the range $\mathcal{N}$ such that $\mathbb{E}x$ is positive and $\varphi\mathbb{E}x = \varphi x$ for any positive x (cf. [8; §9, §10]).

Remark 4.2. In general a sufficient statistic is not HOT-sufficient.

Hint. Take $S = \lambda_1 \hat{e}_1 + \lambda_2 \hat{e}_2$, $\hat{e}_i = \langle \cdot, e_i \rangle e_i$ for orthonormal $(e_i)_{i=1,2}$, and $\varphi_1 = \frac{1}{\sqrt{2}} e_1 + \frac{1}{\sqrt{2}} e_2$, $n = 1$. One can take $x = \sqrt{1-\varepsilon^2} e_1 + \varepsilon e_2$, $0 < \varepsilon < 1$, to show that a state preserving conditional expectation does not exist.

5. The class of all sufficient statistics

The class of all sufficient statistics for states $\varphi_1, \dots, \varphi_n$ can be described by the use of vector measures and Bochner integrals of vector functions on $\mathbb{R}$.

PROPOSITION 5.1. *For a self-adjoint operator $S = \int_{\mathbb{R}} \lambda \, e(\mathrm{d}\lambda)$, let μ denote a Borel measure*

$$\mu(A) = \sum_{i \le n} \|e(A)\varphi_i\|^2$$

on $\mathbb{R}$ and let Borel measurable vector functions $x_i \colon \mathbb{R} \to H$ be defined by formula

$$e(A)\varphi_i = \int_A x_i(t) \, \mu(\mathrm{d}t) \,, \qquad A \in \operatorname{Borel} \mathbb{R},$$

μ-a.e. Then S is a sufficient statistic if and only if

$$\langle x_i(t), x_j(t) \rangle = \pm \|x_i(t)\| \|x_j(t)\| \,, \qquad i,j \le n \,,$$

μ-a.e..

P r o o f . The sufficiency is equivalent to the existence of some scalar functions $\alpha_i(t)$ and a vector function x independent of i such that $x_i(t) = \alpha_i(t) x(t)$ for $i \le n$. □

In particular, any vector $e \in H$ satisfying $\langle e, \varphi_i \rangle \in \mathbb{R}$, $i \le n$, can be used as an eigenvector for some sufficient statistic S with a non-degenerate eigenvalue. A minimal von Neumann algebra generated by sufficient statistics always coincides with $B(H)$.

We stress that Blackwell-Rao theorem is discussed in our paper in the context of *one* sufficient statistic S. It is connected with a general concept of simultaneous measurements in quantum theory.

A discussion of the interpretation of non-commutative sufficient statistics $S_1, S_2, \dots$ is postponed to the next paper.

Acknowledgement

The author wishes to thank the unknown referee and also professor A. Paszkiewicz for valuable suggestions, which led to essential improvement of the first version of the paper.

REFERENCES

[1] DVUREČENSKIJ, A.: *Gleason's Theorem and its Applications*, Kluwer Academic Publisher, Dordrecht, 1993.

[2] HELSTROM, C. W.: *Quantum Detection and Estimation Theory*, Academic, New York, 1976.

[3] HIAI, F.—OHYA, M.—TSUKADA, M.: *Sufficiency, KMS condition and relative entropy in von Neumann algebras*, Pacific J. Math. **96** (1981), 99–109.

[4] HOLEVO, A. S.: *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland Series in Statistics and Probability, Vol. 1, North-Holland Publ. Comp., Amsterdam-New York-Oxford, 1982.

[5] LUDWIG, G.: *An Axiomatic Basis for Quantum Mechanics, Vol. 1. Derivation of Hilbert space structure*, Springer-Verlag, Berlin, 1985.

[6] von NEUMANN, J.: *Mathematical Foundations of Quantum Mechanics*, Princeton Univ. Press, Princeton, NJ, 1955.

[7] SCHERVISH, M. J.: *Theory of Statistics*, Springer-Verlag, New York, 1995.

[8] STRATILA, S.: *Modular Theory in Operator Algebras*, Editura Academiei/Abacus Press, Bucuresti/Turnbridge Wells, Kent, 1981.

Received April 8, 2003
Revised May 4, 2004

Institute of Mathematics
Technical University of Łódź
Al. Politechniki 11
PL–90-924 Łódź
POLAND

E-mail: malgosiacz@poczta.fm