Lüders rule and conditional probability for commuting events

John Harding and Hung Nguyen

Abstract In the context of quantum probability models for causal reasoning in economics, we discuss the mathematics surrounding classical conditional probability and its quantum counterpart for commuting events as given by Lüders rule. While there are many issues in terms of meaning and interpretation in both the classical and quantum setting, there is a transparent connection between the mathematics involved in the two settings.

Key words: Conditional probability, Luders rule, quantum conditional probability, quantum mechanics, quantum probability space.

1 Introduction

The standard mathematical treatment of probability is due to Kolmogorov [7]. A probability space is a triple $(\Omega, \mathscr{A}, p)$ that consists of a set Ω called the event space, a σ-algebra $\mathscr{A}$ of subsets of Ω, and a countably additive probability measure $p : \mathscr{A} \to [0,1]$. A subset A of Ω that belongs to $\mathscr{A}$ is called an event, and $p(A)$ is called the probability of the event.

If B is an event, then the collection $\mathscr{A}_B$ of subsets of B that belong to $\mathscr{A}$ is a σ-algebra on B, and if $p(B) \neq 0$ there is a probability measure p_B on $\mathscr{A}_B$ given by $p_B(C) = p(C)/p(B)$. In the case that $p(B) \neq 0$, for any event A of $\mathscr{A}$ we define the conditional probability $p(A|B)$ of A given B to be $p_B(A \cap B)$. Thus

John Harding
New Mexico State University, USA
e-mail: jharding@nmsu.edu
Hung Nguyen
New Mexico State University, USA and Chiang Mai University, Thailand
e-mail: hunguyen@nmsu.edu

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

The standard von Neumann [15] treatment of quantum mechanics associates to a system a Hilbert space $\mathscr{H}$. Events of the system are given by the projection operators of $\mathscr{H}$. Each maximal set of pairwise commuting projection operators forms a complete Boolean subalgebra of the projection lattice $\mathrm{Proj}(\mathscr{H})$ called a block. A state ρ of the system is a map $\rho : \mathrm{Proj}(\mathscr{H}) \to [0,1]$ that restricts to a countably additive probability measure on each block. For an event Q and state ρ, we call the $\rho(Q)$ the probability of the event Q when in state ρ.

A commonly used quantum counterpart for conditional probability is based on Lüders Rule. Given a system in state ρ, upon a successful test to see that an event R is obtained, Lüders Rule is a description of a state update of the system to a new state ρ_R. For an event Q one can then defines Lüders conditional probability $\rho(Q|R)$ of Q given R when the system in originally in state ρ to be $\rho_R(Q)$.

The primary purpose of this note is expository, to illustrate in as transparent means as possible the mathematical correspondence between classical conditional probability and Lüders conditional probability for commuting events. We make no claims for or against the use of Lüders probability as a tool in aspects of quantum information. Our intended audience is those with modest familiarity with the issues. The philosophy and interpretation of even "ordinary" classical probability is an enormous issue, let alone quantum counterparts. In this direction we will make some comments, but our intent is to address the mathematics.

This note is organized as follows. In the second section, we provide background on finite-dimensional quantum mechanics and Lüders rule in this setting. In the third section we discuss the connection between classical conditional probability as applied to finite sample spaces and Lüders probability in the setting of finite-dimensional quantum mechanics. In the fourth section, we extend this discussion to infinite-dimensional quantum mechanics. The fifth section makes some concluding remarks.

For general background on quantum mechanics, in a spirit presented here, the reader should consult [2, 12, 14, 15]. For discussions of Lüders rule and conditional probability, consult [1, 16]. For general considerations of probability spaces consult any standard text on these matters, such as [7, 8]. For matters related to Hilbert spaces and their projection lattices, consult [5, 11].

2 Finite-dimensional quantum mechanics

Up to isomorphism, a finite-dimensional complex Hilbert space $\mathscr{H}$ is the complex vector space $\mathbb{C}^n$ for some natural number n, with inner product $\langle \cdot | \cdot \rangle$, and with norm $\|v\| = \sqrt{\langle v|v\rangle}$. Throughout this section, our Hilbert spaces will all be finite dimensional. The infinite-dimensional setting is considered later.

A projection of $\mathscr{H}$ is a matrix, or operator, Q that satisfies $Q^2 = Q^\dagger = Q$. The range of a projection Q is a subspace S, and the nullspace is the subspace $S^\perp$ consisting of the vectors orthogonal to each vector in S. For each subspace S there is exactly one projection P whose range is S. We treat projections, or equivalently subspaces, as events of a quantum system.

The collection $\text{Proj}(\mathscr{H})$ of projections plays the role of the set $\mathscr{A}$ of subsets of the sample space Ω in the case of a classical probability space $(\Omega, \mathscr{A}, p)$ on a finite set. Classical events $\mathscr{A}$ form a Boolean algebra where the meet is the conjunction $E \wedge F$ of events, the join $E \vee F$ is their disjunction, and E^c is the negation. Quantum events form a structure called an orthomodular lattice that is a non-distributive generalization of a Boolean algebra. For projections Q and R associated to the subspaces S and T, their meet $Q \wedge R$ is the projection associated with the subspace $S \cap T$, their join $Q \vee R$ is the projection associated with the subspace $S + T$, and Q^c is the projection $1 - Q$ associated with $S^\perp$.

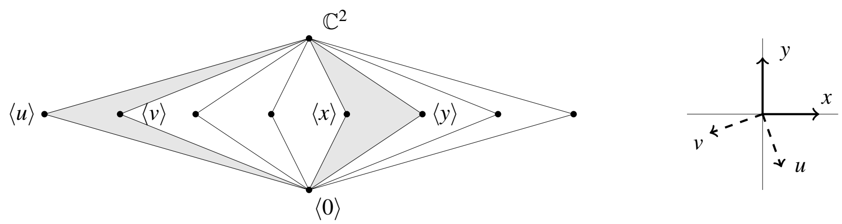


Fig. 1 A portion of the lattice $\text{Proj}(\mathbb{C}^2)$ showing two maximal 4-element Boolean subalgebras

While the lattice $\text{Proj}(\mathscr{H})$ itself is not Boolean, it is built from a collection of Boolean subalgebras "glued" together. Figure 1 shows the case of the projection lattice of $\mathbb{C}^2$. Here the labels are of the subspaces associated to projections, and we use $\langle x \rangle$ for the 1-dimensional subspace spanned by the vector x. The general situation is described in the following.

Definition 1. A maximal Boolean subalgebra of $\text{Proj}(\mathscr{H})$ is called a block.

Every finite Boolean algebra is isomorphic to the power set Boolean algebra of a finite set. In a finite Boolean algebra, an element that covers the 0 with nothing in between is called an atom. The atoms of the power set of the set $\{x_1, \ldots, x_n\}$ are the singleton sets $\{x_1\}, \ldots, \{x_n\}$.

Theorem 1. *If $e_1, \ldots, e_n$ is an orthonormal basis of $\mathscr{H}$, then the projections onto the subspaces $\langle e_1 \rangle, \ldots, \langle e_n \rangle$ are the atoms of a block of* $\text{Proj}(\mathscr{H})$*, this block is isomorphic to the power set of $\{e_1, \ldots, e_n\}$, and every block arises this way.*

Figure 1 shows the block for the orthonormal basis x, y of $\mathbb{C}^2$, and also the block for the orthonormal basis u, v. Each of these blocks is isomorphic to the power set of a 2-element set. For $\text{Proj}(\mathbb{C}^3)$, each block has 3 atoms and is isomorphic to the power set of a 3-element set. Blocks in this case can overlap in more interesting ways. For u,v in the x,y-plane as in Figure 1, the blocks for the orthonormal basis x,y,z and for the orthonomal basis u,v,z will share an atom and its complement.

Blocks play a vital role, and we require the following well-known result to work effectively with them.

Definition 2. Two events Q,R are compatible if they lie in a common block of $\mathscr{H}$.

Proposition 1. *Two events Q,R compatible iff they commute, that is, if $PQ = QP$; and in this case the product $PQ = P \wedge Q$ is their meet.*

We have two thirds of the quantum analog of a classical finite probability space $(\Omega,\mathscr{A},p)$. The sample space Ω is replaced by a finite-dimensional Hilbert space $\mathscr{H}$, and the power set Boolean algebra $\mathscr{A}$ of all subsets of Ω is replaced by the lattice $\mathrm{Proj}(\mathscr{H})$ of projections, or equivalently all subspaces of $\mathscr{H}$. It remains to treat the quantum analog of the classical probability measure $p : \mathscr{A} \to [0,1]$.

Definition 3. A map $\rho : \mathrm{Proj}(\mathscr{H}) \to [0,1]$ is a quantum probability measure if it restricts to a classical probability measure ρ_B on each block B of $\mathrm{Proj}(\mathscr{H})$.

Somewhat remarkably, there is a clear description of all quantum probability measures on $\mathrm{Proj}(\mathscr{H})$, at least when $\dim \mathscr{H} \geq 3$. We begin with the following.

Definition 4. For a vector $v \neq 0$ in $\mathscr{H}$, define the pure state $\rho_v : \mathrm{Proj}(\mathscr{H}) \to [0,1]$ by setting

$$\rho_v(Q) \;=\; \frac{\|Qv\|^2}{\|v\|^2} \tag{1}$$

A convex combination $\rho = \lambda_1\rho_{v_1} + \cdots + \lambda_k\rho_{v_k}$ of pure states is called a mixed state, or simply a state. Of course, every pure state is a state, but not conversely.

Remark 1. Any two vectors generating the same 1-dimensional subspace give the same pure state. When v is a unit vector, the state is given by $\rho_v(Q) = \|Qv\|^2$.

Remark 2. States can be represented by matrices. Let v be a unit vector, $|v\rangle$ be v considered as a column matrix, and $\langle v|$ be the adjoint of $|v\rangle$. Then $|v\rangle\langle v|$ is the projection onto the subspace spanned by v. A short calculation shows that the value of $\rho_v(Q)$ can be computed through the trace by

$$\rho_v(Q) = tr(|v\rangle\langle v|Q)$$

For a mixed state ρ given by a convex combination of pure states, one takes the convex combination D of the matrices for the pure states and has $\rho(Q) = tr(DQ)$. Matrices D that occur in this way are exactly density matrices.

Theorem 2 (Gleason [3]). *If* $\dim \mathscr{H} \geq 3$*, the states of $\mathscr{H}$ are exactly the quantum probability measures on* $\mathrm{Proj}(\mathscr{H})$.

Remark 3. For dimension 2, there are quantum probability measures that are not states. It is because the dimension is too small to allow the blocks to overlap. These probability measures are too wild to be of physical interest, and one deals solely with states rather than quantum probability measures. Of course, except in the case of $\mathbb{C}^2$, these concepts agree.

Definition 5. A quantum probability space $(\mathscr{H}, \mathrm{Proj}(\mathscr{H}), \rho)$ consists of a Hilbert space $\mathscr{H}$, its projection lattice, and a state ρ on $\mathscr{H}$.

Given a block B of $\mathrm{Proj}(\mathscr{H})$, there is an orthonormal basis $e_1, \ldots, e_n$ with B isomorphic to the power set of $\Omega_B = \{\langle e_1\rangle, \ldots, \langle e_n\rangle\}$. Then the restriction ρ_B of the state ρ to B is a classical probability measure. If we allow the slight abuse of identifying B and the power set of Ω_B, the following summarizes the situation.

Theorem 3. *A quantum probability space $(\mathscr{H}, \mathrm{Proj}(\mathscr{H}), \rho)$ is the union over all its blocks B of the overlapping classical probability spaces (Ω_B, B, ρ_B).*

3 Lüders conditional probability

In this section, we remain in the finite-dimensional setting. We begin with Lüders rule for "updating" a state in quantum mechanics. For a system in state ρ, after it is observed that event R occurs, the system is updated to a new state ρ_R. The description of this new state is a refinement, really a correction, of von Neumann's "projection postulate". We begin with matters for a pure state.

Definition 6. For a pure state $\rho = \rho_v$ and event R with $\rho(R) \neq 0$, set $\rho_R = \rho_{Rv}$.

When the system is in a pure state ρ_v, the idea is simple. After R is found to occur, the system is updated to the pure state given by the projection Rv of v onto R. A mixed state can be viewed, among other ways, as a statistical mixture of pure states. The formula for the update of a mixed state is exactly as one would expect.

Definition 7. For a state $\rho = \lambda_1\rho_{v_1} + \cdots + \lambda_k\rho_{v_k}$ and event R with $\rho(R) \neq 0$, let

$$\rho_R = \mu_1\rho_{Rv_1} + \cdots + \mu_k\rho_{Rv_k} \quad \text{where} \quad \mu_i = \frac{\lambda_i\rho_{v_i}(R)}{\rho(R)} \tag{2}$$

Remark 4. Note that $|Rv\rangle\langle Rv| = R|v\rangle\langle v|R$ and the trace $tr(R|v\rangle\langle v|R) = tr(|v\rangle\langle v|R)$ is the probability $\rho_v(R)$. So the matrix ρ_{Rv} is $(R|v\rangle\langle v|R)/\rho_v(R)$. Suppose the state ρ in Definition 7 has matrix D, then the matrix for $\mu_1\rho_{Rv_1} + \cdots + \mu_k\rho_{Rv_k}$ is

$$\frac{\lambda_1}{tr(DR)} R|v_1\rangle\langle v_1|R + \cdots + \frac{\lambda_k}{tr(DR)} R|v_k\rangle\langle v_k|R$$

This is equal to $RDR/tr(DR)$, showing that the definition of ρ_R is independent of the choice of representation as a convex combination.

Definition 8. For ρ a state and events Q,R with $\rho(R) \neq 0$, let Lüders conditional probability of Q given R be given by

$$\rho^L(Q|R) \;=\; \rho_R(Q) \tag{3}$$

Our aim, broadly, is to consider the relationship between Lüders conditional probability and classical conditional probability. As a first step, we show that for a quantum probability space $(\mathscr{H}, \mathrm{Proj}(\mathscr{H}), \rho)$, Lüders conditional probability and classical conditional probability agree on each of the classical probability spaces (Ω_B, B, ρ_B), for B a block of $\mathscr{H}$, that comprise the quantum probability space.

Theorem 4. *If ρ is a state on $\mathscr{H}$ and Q,R belong to a block B of $\mathscr{H}$, then Lüders conditional probability $\rho^L(Q|R)$ is equal to the classical probability $\rho_B(Q|R)$.*

Proof. Note first that the condition $\rho(R) \neq 0$ for Lüders conditional probability $\rho^L(Q|R)$ to be defined is the same condition required for the classical conditional probability $\rho_B(Q|R)$ to be defined. For a pure state $\rho = \rho_v$ with v a unit vector, having Q,R belong to a block means that they commute and $QR = Q \wedge R$. Using this we obtain

$$\rho^L(Q|R) = \rho_{Rv}(Q) = \frac{\|QRv\|^2}{\|Rv\|^2} = \frac{\|(Q\wedge R)v\|^2}{\|Rv\|^2} = \frac{\rho_v(Q\wedge R)}{\rho_v(R)} = \rho_B(Q|R) \tag{4}$$

The result for a mixed state follows from equations (2) and (3) and the result for the pure states. The calculations and reasoning behind them are similar to those one would would find for the classical conditional probability of a statistical mixture of probabilities.

Remark 5. The discussion that Lüders conditional probability agrees with classical conditional probability is usually given in terms of an algebraic calculation using the density operator representation of a state. We give the argument above to illustrate it is a consequence of the very transparent result using the projection postulate for pure states.

For a quantum probability space $(\mathscr{H}, \mathrm{Proj}(\mathscr{H}), \rho)$, Theorem 4 shows that for each block B, the classical probability space (Ω_B, B, ρ_B) has its classical conditional probability given by Lüders rule. There remains the possibility that the classical probability spaces (Ω_B, B, ρ_B) have some special properties not usually enjoyed by finite classical probability spaces. We next show that this is not the case.

Theorem 5. *Let $(\Omega, \mathscr{A}, p)$ be a classical probability space with $|\Omega| = n$. Then there is a quantum probability space $(\mathscr{H}, \mathrm{Proj}(\mathscr{H}), \rho)$, where $\mathscr{H}$ has dimension n, a unit vector $v \in \mathscr{H}$ with $\rho = \rho_v$, and a block B of $\mathscr{H}$ so that $(\Omega, \mathscr{A}, p) \simeq (\Omega_B, B, \rho_B)$, and so that classical conditional probability in (Ω_B, B, ρ_B) is given by Lüders rule.*

Proof. Suppose $\Omega = \{x_1, \ldots, x_n\}$. Let $\mathscr{H}$ be an n-dimensional Hilbert space with $e_1, \ldots, e_n$ an orthonormal basis. Let B be the block whose atoms $P_1, \ldots, P_n$ are the projections onto the 1-dimensional subspaces $\langle e_1 \rangle, \ldots, \langle e_n \rangle$. Then B is isomorphic to the power set of Ω. Let $\lambda_i = \sqrt{p(x_i)}$ for $i = 1, \dots, n$ and set $v = \lambda_1 e_1 + \cdots + \lambda_n e_n$. Set $\rho = \rho_v$. Then for $1 \leq i \leq n$

$$\rho(P_i) = \|P_i v\|^2 = \lambda_i^2 = p(x_i)$$

So $(\Omega, \mathscr{A}, p) \simeq (\Omega_B, B, \rho_B)$, and by Theorem 4 classical conditional probability in (Ω_B, B, ρ_B) is given by Lüders rule.

4 The infinite-dimensional setting

We continue our investigation of the previous section in the setting of a separable infinite-dimensional Hilbert space $\mathscr{H}$. As before, $\text{Proj}(\mathscr{H})$ is an orthomodular lattice, and we call its maximal Boolean subalgebras blocks.

Proposition 2. *Each block of a Hilbert space is a complete Boolean algebra.*

In the finite-dimensional setting, a block is isomorphic to the power set of its atoms. In the general setting, blocks are complete, but are not usually sigma algebras of subsets of a set. Still, one can conduct much of probability theory in this setting.

Definition 9. A map $p : B \to [0,1]$ on a σ-complete Boolean algebra is a σ-additive measure if $p(0) = 0$, $p(1) = 1$, and for each family $(x_n)_{\mathbb{N}}$ where $x_m \leq x_n'$ for $m \neq n$

$$p(\bigvee_n x_n) = \sum_n p(x_n)$$

Definition 10. A generalized classical probability space is a pair (B, p) with B a σ-complete Boolean algebra and p a σ-additive measure. Members of B are called events, and for events S, T with $p(T) \neq 0$ the generalized classical conditional probability is given by

$$p(S|T) = \frac{p(S \wedge T)}{p(T)}$$

Example 1. The key example is the Boolean algebra B of Lebesgue measurable sets modulo sets of measure zero and the map $p : B \to [0,1]$ given by Lebesgue measure of an equivalence class of Lebesgue measurable sets modulo measure zero.

Remark 6. It is not the case that every σ-complete Boolean algebra is isomorphic to a σ-algebra of subsets of a set. But the Loomis-Sikorski theorem says that each is isomorphic to a σ-algebra of sets modulo a σ-ideal of null sets.

In the infinite-dimensional setting, a quantum probability measure is a map $\rho : \text{Proj}(\mathscr{H}) \to [0,1]$ that restricts to a σ-additive measure on each block. Pure states are defined as before, and mixed states as σ-convex combinations of pure states. Gleason's theorem remains valid with these definitions. The interpretation of states in terms of density matrices must be revised to density operators.

Definition 11. A quantum probability space is a system $(\mathscr{H}, \text{Proj}(\mathscr{H}), \rho)$ where $\mathscr{H}$ is a separable Hilbert space and ρ is a state.

Theorem 6. *A quantum probability space $(\mathscr{H}, \text{Proj}(\mathscr{H}), \rho)$ consists of a family (B, ρ_B) of generalized classical probability spaces where B ranges over the blocks of $\mathscr{H}$. For each block B, generalized classical conditional probability on (B, ρ_B) is given by Lüders rule.*

Proof. Since $\text{Proj}(\mathscr{H})$ is an orthomodular lattice, it is covered by its blocks. Since each block is complete and each state ρ gives a quantum probability measure, each (B, ρ_B) is a generalized classical probability space. It remains only to show that Lüders rule gives generalized conditional probability on each (B, ρ_B). For $\rho = \rho_v$ a pure state for a unit vector v, the argument is exactly as in (4). The argument for a general state is exactly as in the finite-dimensional setting, replacing a convex combination with a σ-convex combination.

We next consider the extent to which a classical probability space $(\Omega, \mathscr{A}, p)$ can be realized within (B, ρ_B) for some block B of a quantum probability space. The situation is more delicate than in the finite-dimensional setting, but good progress can be made. Our primary tool is the following standard result that we state for probability spaces, though it is valid for all measure spaces.

Theorem 7. *For a probability space $(\Omega, \mathscr{A}, p)$, the collection $L^2(\Omega, \mathscr{A}, p)$ of square integrable functions $f : \Omega \to \mathbb{C}$ modulo equivalence on sets of probability zero is a Hilbert space under the inner product*

$$\langle f, g \rangle = \int_\Omega \overline{f} g \, dp$$

Definition 12. A probability space $(\Omega, \mathscr{A}, p)$ is separable iff the metric on $\mathscr{A}$ given through symmetric difference by $p(\Delta(S,T))$ is separable.

It is known [2, p. 20] that $L^2(\Omega, \mathscr{A}, p)$ is separable iff $(\Omega, \mathscr{A}, p)$ is separable. Since we wish to stay in the setting of separable Hilbert spaces, we confine attention to separable probability spaces. Additionally, when working with Hilbert spaces $L^2(\Omega, \mathscr{A}, p)$ we use the abbreviation a.e. for working with equivalence on sets of probability zero and $[f]$ for the equivalence class of f modulo equivalence a.e.

Proposition 3. *For a separable probability space $(\Omega, \mathscr{A}, p)$ and $S \in \mathscr{A}$, the set*

$$\hat{S} = \{[f] : f = 0 \text{ a.e. outside of } S\}$$

is a closed subspace of $L^2(\Omega, \mathscr{A}, p)$.

Proof. Clearly $\hat{S}$ is closed under sums and scalar multiplication, so is a subspace. Suppose g is nonzero on a set of positive measure outside of S. Then there is a set T outside of S such that $\int_T |g| \, dp = \varepsilon > 0$. So the distance between $[f]$ and $[g]$ is greater than ε for each $[f] \in \hat{S}$, and therefore $[g]$ does not belong to the closure of $\hat{S}$.

Definition 13. For a separable probability space $(\Omega,\mathscr{A},p)$ and $S\in\mathscr{A}$, let $P_{\hat{S}}$ be the projection onto $\hat{S}$. Define $\Gamma:\mathscr{A}\to\mathrm{Proj}(\mathscr{H})$ by $\Gamma(S)=P_{\hat{S}}$ and let $B_{\mathscr{A}}$ be the image of this map.

Theorem 8. *For $(\Omega,\mathscr{A},p)$ a separable probability space, $B_{\mathscr{A}}$ is a σ-complete Boolean algebra of projections and the map $\Gamma:\mathscr{A}\to B_{\mathscr{A}}$ is a σ-complete Boolean algebra homomorphism whose kernel is the sets of probability zero in $\mathscr{A}$.*

Proof. Clearly Γ preserves bounds. For a countable family S_n in $\mathscr{A}$ with $S=\bigcap_n S_n$ it is easy to see that $\hat{S}=\bigcap_n\hat{S}_n$. So Γ preserves countable meets. For $S\in\mathscr{A}$ with $S^c=\Omega\setminus S$, an element $[f]\in\hat{S}$ and $[g]\in\hat{S^c}$ have the support of f,g disjoint a.e. hence $\langle f,g\rangle=0$. So $\hat{S}$ and $\hat{S^c}$ are orthogonal subspaces. But each function can be expressed as a sum $f+g$ with $[f]\in\hat{S}$ and $[g]\in\hat{S^c}$. Thus $\hat{S}^\perp=\hat{S^c}$. So Γ preserves orthocomplements. It follows that Γ is a σ-complete ortholattice homomorphism and its image is a σ-complete Boolean algebra. It is easy to see that the kernel of Γ is the sets of probability zero in $\mathscr{A}$.

Definition 14. For a separable probability space $(\Omega,\mathscr{A},p)$, let $\mathbb{1}$ be the equivalence class of the function taking constant value 1 on Ω and let $\rho_{\mathbb{1}}$ be the pure state on $L^2(\Omega,\mathscr{A},p)$ given by the unit vector $\mathbb{1}$.

With setup as in Theorem 8 we have the following.

Theorem 9. *The map $\Gamma:(\mathscr{A},p)\to(B_{\mathscr{A}},\rho_{\mathbb{1}})$ preserves probability and conditional probability: for $S,T\in\mathscr{A}$ with $p(T)\neq 0$*

1. $p(S)=\rho_{\mathbb{1}}(\Gamma(S))$,
2. $p(S|T)=\rho_{\mathbb{1}}(\Gamma(S)|\Gamma(T))$.

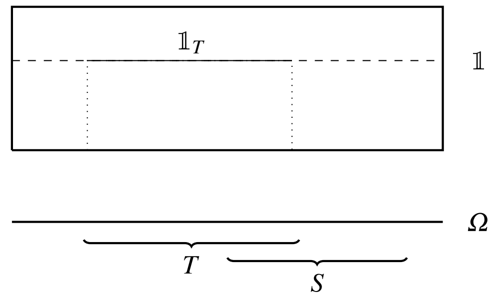
Further, in $(B_{\mathscr{A}},\rho_{\mathbb{1}})$ conditional probability is given by Lüders rule.

Proof. (1) Since $\rho_{\mathbb{1}}$ is a pure state given by a unit vector, $\rho_{\mathbb{1}}(\Gamma(S))=\|P_{\hat{S}}\mathbb{1}\|^2$. But the projection of $\mathbb{1}$ onto the closed subspace $\hat{S}=\{[f]:f=0\text{ a.e. outside }S\}$ is the equivalence class $\mathbb{1}_S$ of the characteristic function 1_S of S. Using the definition of norm $\|P_{\hat{S}}\mathbb{1}\|^2=\int_\Omega|1_S|^2\,dp=p(S)$. (2) Since $B_{\mathscr{A}}$ is a Boolean subalgebra of the projection lattice, it is contained in a block, by Thoerem 6 conditional probability on $(B_{\mathscr{A}},\rho_{\mathbb{1}})$ is given by Lüders rule. Note that $P_{\hat{T}}(\mathbb{1})$ is the equivalence class $\mathbb{1}_T$ of the characteristic function of T. Thus

$$\rho_{\mathbb{1}}(\Gamma(S)|\Gamma(T))=\rho_{\mathbb{1}}(P_{\hat{S}}|P_{\hat{T}})=\rho_{\mathbb{1}_T}(P_{\hat{S}})=\frac{\|P_{\hat{S}}\mathbb{1}_T\|^2}{\|\mathbb{1}_T\|^2}=\frac{\|\mathbb{1}_{S\cap T}\|^2}{\|\mathbb{1}_T\|^2}=\frac{p(S\cap T)}{p(T)}$$

Remark 7. Notation and fussy details such as equivalence a.e. can get in the way of showing how transparently Lüders rule for conditional probability translates to classical conditional probability in the setting of Theorems 8 and 9. We speak now informally and ignoring modulo a.e..

For a classical probability space $(\Omega,\mathscr{A},p)$, consider $L^2(\Omega,\mathscr{A},p)$. Each $S\in\mathscr{A}$ gives a closed subspace $\hat{S}$ of this Hilbert space consisting of functions vanishing outside of S. The set $B_{\mathscr{A}}$ of projections $P_{\hat{S}}$ onto these subspaces is a Boolean algebra isomorphic with $\mathscr{A}$.



Then the probability $\rho_{\mathbb{1}}(T)$ is the integral of the characteristic function of T with respect to the measure p, so is $p(T)$. Lüders state update after T is observed is to the state $\rho_{\mathbb{1}_T}$ given by the characteristic function $\mathbb{1}_T$. Projecting $\mathbb{1}_T$ onto $\hat{S}$ gives $\mathbb{1}_{S\cap T}$. So $\rho_{\mathbb{1}_T}$ is a measure on $B_{\mathscr{A}}$ with $\rho_{\mathbb{1}_T}(P_{\hat{S}}) = p(S\cap T)/p(T)$.

Remark 8. The results for finite classical probability spaces and finite-dimensional Hilbert spaces are a special case of the situation described above.

Remark 9. In the setup of Theorems 8 and 9, we do not know if $B_{\mathscr{A}}$ is always a block. This does not affect our arguments, but would make things tidier.

Remark 10. It is well known that every separable Hilbert space $\mathscr{H}$ is isomorphic to $L^2(\Omega,\mathscr{A},p)$ for some classical probability space $(\Omega,\mathscr{A},p)$. An affirmative answer to the following question would show that the situation described in Theorems 8 and 9 is the generic one.

Question. Given a separable Hilbert space $\mathscr{H}$, a block B of $\mathrm{Proj}(\mathscr{H})$, and a state ρ of $\mathscr{H}$, is there a separable classical probability space $(\Omega,\mathscr{A},p)$ and a unitary map $U : L^2(\Omega,\mathscr{A},p) \to \mathscr{H}$ such that U induces an isomorphism between $B_{\mathscr{A}}$ and B that takes the restriction of the state $\rho_{\mathbb{1}}$ to $B_{\mathscr{A}}$ to the restriction of the state ρ to B?

We conclude this section with an aside that makes use of our previous results. A type of a "partially frequentist" justification of the rule for classical conditional probability is given in [4, 9] using a countable product of a probability space with itself. To sketch, suppose A,B are events of a classical probability space $(\Omega,\mathscr{A},p)$ with $p(B) \neq 0$. In the countable product $\prod_n(\Omega_n,\mathscr{A}_n,p_n)$ with all factors equal to $(\Omega,\mathscr{A},p)$, consider for each $n \geq 0$ the rectangle

$$(B^c)^n(A\cap B) = \underbrace{B^c \times \cdots \times B^c}_{n \text{ times}} \times (A\cap B) \times \Omega \times \Omega \times \cdots$$

Using the formula for the sum of a geometric series,

$$\sum_{n=0}^{\infty} p((B^c)^n(A\cap B)) = \left(\sum_{n=0}^{\infty} p(B^c)\right) p(A\cap B) = \frac{p(A\cap B)}{p(B)}$$

One can interpret this as repeating many trials of a thing until B is obtained, and having A also obtained on the first time B is attained.

We can view this in the quantum setting considering the quantum probability space $(\mathscr{H}, \mathrm{Proj}(\mathscr{H}), \rho_{\mathbb{1}})$ where $\mathscr{H} = L^2(\prod_n (\Omega_n, \mathscr{A}_n, p_n))$. This Hilbert space $\mathscr{H}$ is isomorphic to the countable tensor product $\bigotimes_n L^2(\Omega_n, \mathscr{A}_n, p_n)$ with the ground state of each taken to be $\rho_{\mathbb{1}}$ [10, p. 96, 15.11]. This tensor product is the Fock space $\bigoplus_n \mathscr{H}_n$ where $\mathscr{H}_n$ is the n-fold tensor product of $L^2(\Omega, \mathscr{A}, p)$ with itself [13, 2.5.7].

This leads to the following "partially frequentist" physical interpretation in the quantum setting. Suppose we have an unlimited supply of identically prepared systems that we enumerate as system 1, system 2, and so forth. We conduct an experiment for each system to see if event R is obtained. The event $(R^c)^n(Q \cap R)$ is that system 1, ..., system n did not have R occur, and system $n+1$ had Q and R occur. This last portion can be tested only if Q and R are commuting events.

For non-commuting events Q and R, this "partially frequentist" approach can be used in conjunction with Lüders state update to justify a Lüders rule for conditional probability given Lüders state update. Lüders state update gives that for a system in state ρ, the probability of first obtaining R and then obtaining Q is given by $\rho(R) \cdot \rho_R(Q)$. Using this in place of $p(A \cap B)$ in the classical argument gives

$$\rho(Q|R) = \frac{\rho(R)\rho_R(Q)}{\rho(R)} = \rho_R(Q)$$

This is Lüders rule for quantum conditional probability. This argument applies whether or not Q and R commute. Of course, it is the least problematic part of the situation, it is the state update that truly causes one to ponder.

5 Discussion

Classical probability and conditional probability, as formulated by Kolmogorov [7], is a precise mathematical theory. The connection between this theory and its interpretation in the natural world is a different matter. Our focus in this note is on conditional probability. Consider

$$p(A|B) = \frac{p(A \cap B)}{p(B)}$$

One will see this described in words as "the probability of A given B", or as "the probability that A will occur given that B has occurred", or "the probability that a thing with property B also has property A". An operational view of Kolmogorov's conditional can be given as follows. For a large collection of objects of the same type, view $p(B)$ as the proportion of the objects that have property B, and $p(A|B)$ as the proportion of those objects that have property B that also have property A.

These descriptions differ. Some use tense. Some use the state of knowledge of an observer. Some use the absolute of a thing having one or more properties that are not time-dependent. So is time, or the state of knowledge of an observer, or the existence of an absolute reality, an essential component of classical conditional probability? Likely a common view is that yes, these are different interpretations of the meaning of conditional probability, but they all lead to the same mathematics, so the discussion is at best unnecessary and more suited for philosophers than for people involved in practical matters.

In the quantum world, these are relevant issues. The Kochen-Specker Theorem [6] provides a finite set of events for $\mathbb{C}^3$ that cannot be simultaneously be assigned values of true and false in a consistent manner. So there is no absolute reality for us to rely upon, at least not one that is of a familiar nature to us. The two-slit experiment points to the role of an observer in the system, and the quantum zeno effect points to measurements affecting the subsequent behavior of the system. Unless one knows something special about the state of a system, it is meaningless to ask whether it "has" a given property without testing for the property, a test for the property affects the nature of the system, and in general it is not possible to test for more than one property at a time.

Lüders rule is a mathematical statement. It has a physical interpretation that, to the best of our knowledge, is supported by physical experiments.

> **Lüders rule:** Suppose that a large number of systems are prepared in an identical state ρ. If a non-destructive measurement of a property R is conducted on each system, and then a measurement of a property Q is conducted on those that had property R, then the proportion of those that had property R that subsequently have property Q is given by $\rho_R(Q)$.

This provides a physical interpretation of Lüders rule in terms of sequential measurements. In the preceding sections we have shown a close connection between Lüders rule as applied to compatible events and classical conditional probability. But classical conditional probability does not involve a temporal component in its formulation. What is the reason behind this? The common explanation is that compatible events can be tested "simultaneously". We try to make this more clear.

Proposition 4. *Events Q,R are compatible iff there are four pairwise orthogonal events S,T,U,V whose join is 1 and with $S \vee T = Q$ and $S \vee U = R$.*

Proof. If Q,R are compatible, they belong to a Boolean subalgebra. Then the events $S = Q \wedge R$, $T = Q \wedge R^c$, $U = S^c \wedge R$ and $T = Q^c \wedge R^c$ have the desired properties. The converse follows as pairwise orthogonal events belong to a common Boolean subalgebra.

Remark 11. This provides a means to discuss in precise physical and mathematical terms what it means to test for compatible events Q,R "simultaneously". Once can conduct an experiment with four mutually exclusive and exhaustive outcomes that correspond to events S,T,U,V. Then one views the event Q occurring as either of the events S or T occurring, and the event R occurring as either of the events S or U occurring. This is possible only when Q and R are compatible.

Remark 12. One might wonder why compatibility of events is such an issue. For any two events Q,R their meet $Q \wedge R$ exists in $\mathrm{Proj}(\mathscr{H})$. While this meet is an event, it has little to do logically with the events Q and R. For instance, $Q \wedge R$ and $Q \wedge R^{\perp}$ can both be zero without Q being zero. In fact, this happens for every non-trivial event Q asking about position and every non-trivial event R asking about momentum. In short, $Q \wedge R$ in no sense represents "Q and R", except when Q,R are compatible.

Remark 13. With the situation as in Proposition 4, we will compare the probabilities of first obtaining R, then Q, starting in pure state ρ_v with the probability of obtaining $Q \wedge R = S$ when starting in state ρ_v. We already know these will be equal, let us see in another way why. The probability of the first is

$$\rho_v(R)\rho_{Rv}(Q) = \|Rv\|^2 \rho_{Rv}(Q) = \|Rv\|^2 \frac{\|QRv\|^2}{\|Rv\|^2}$$

Now Rv is the projection of v onto the closed subspace for R, and since R is the orthogonal join of S and U we have $Rv = s + u$ for some s,u in closed subspaces for S and U. Then projecting Rv onto the closed subspace for S gives $QRv = s$. But of course, $Sv = s$. Thus

$$\rho_v(R)\rho_{Rv}(Q) = \|s\|^2 = \rho_v(Q \wedge R)$$

This says that starting in state ρ_v, the probability of finding a yes result to a test of R, then to a subsequent test of Q is the same as finding a yes answer directly to a test of $S = R \wedge Q$, and furthermore, the two processes both leave us in the state ρ_s. A symmetrical argument shows that that this is the same result as if Q was conducted first, and then R. In all this, we are assuming that the time between Q and R is infinitesimal, and that no external influences change the state of the system.

To conclude, we remark again that this is largely an expository paper. With the exception the question raised and diversion to product spaces, the material presented is known, although usually presented more from a matrix algebra perspective that we find less illuminating. However, "known" things can become obscured by later developments. Take for example the following passages from the paper [16] that is so germane to the topic of the conference.

> Therefore, the Lüders probability cannot be treated as the generalization of the classical conditional probability that is not necessarily symmetric. ...
>
> But let us remember the quantum-classical correspondence principle, according to which classical theory has to be a particular case of quantum theory. In classical theory, the field of events is commutative. In quantum theory, commuting observables share the same family of eigenvectors. This can be formulated as the property $\langle \alpha_i | n_j \rangle = \delta_{ij}\delta_{\alpha\beta}$. Then, passing to commutative events, for the Lüders probability (2.8) we obtain
>
> $$p^L(A_n|B_\alpha) = \delta_{n\alpha} = p^L(B_\alpha|A_n)$$
>
> This is not merely symmetric, but even trivial. Contrary to this, classical conditional probabilities are neither symmetric nor trivial.

The authors of [16], Yukalov and Sornette, are well aware of the results we have given in previous sections. They clearly have something else in mind in the above passages, although we do not understand precisely what. However, a person less familiar with Lüders conditional probability and its relation to classical probability for commuting events could easily gain an incorrect impression from these passages of [16]. We hope in this note to make the basic situation available in a clear way for those new to the subject.

The larger issue remains, to find an appropriate way, or ways, of dealing with conditional probability in quantum probability models for economics.

References

1. I. G. Bobo, *On quantum conditional probability*, Theoria, 76 (2013), 115-137.
2. A. Dvurecenskij, *Gleason's Theorem and its Applications*, Math. and its Applications (East European Series), 60. Kluwer Academic Publishers Group, Dordrecht; Ister Science Press, Bratislava, 1993.
3. A. Gleason, *Measures on the closed subspaces of a Hilbert space*, J. of Math. and Mech., Indiana Univ. Math. J. 6 (1957), No. 4, 885-893.
4. I. R. Goodman and H. T. Nguyen, *A theory of conditional information for probabilistic inference in intelligent systems PP: Product space approach*, Information Sciences 76 (1994), 13-42
5. R. V. Kadison and J. R. Ringrose, *Fundamentals of the Theory of Operator Algebras. Vol. 1: Elementary Theory*, Academic Press, New York, 1983.
6. S. Kochen and E. Specker, 1967, *The Problem of Hidden Variables in Quantum Mechanics*, Journal of Mathematics and Mechanics, 17 (1967), 59-87.
7. A. Kolmogorov, *Foundations of the Theory of Probability*, Chelsea Publishing Company, New York, 1950.
8. J. Neveu, *Mathematical Foundations of the Calculus of Probability*, Holden-Day, Inc., San Francisco, Calif.-London-Amsterdam, 1965.
9. H. T. Nguyen, *Conditional event algebras: The state-of-the-art*, in Beyond Traditional Probabilistic Data Processing Techniques, (O. Kosheleva et al. Eds.), 545-555, Springer, 2020.
10. K. R. Parthasarathy, *An Introduction to Quantum Stochastic Calculus*, Springer, 1992.
11. P. Pták and S. Pulmannová, *Orthomodular Structures as Quantum Logics*, Fundamental Theories of Physics, 44. Kluwer Academic Publishers Group, Dordrecht, 1991.
12. E. Prugovecki, *Quantum Mechanics in Hilbert Space*, second ed., Academic Press, New York, 1981.
13. N. Weaver, *Mathematical Quantization*, Chapman and Hall, 2001.
14. V. S. Varadarajan, *Geometry of Quantum Theory*, Second Ed., Springer-Verlag, 1985.
15. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955.
16. V. I. Yukalov and D. Sornette, *Quantum probability and quantum decision-making*, Phil. Trans. R. Soc. A 374 (2016), 20150100