Quantum Bayes' rule affirming consistency in measurement inferences in quantum mechanics

Arnav Metrani

IISER Mohali

17th November 2024

Objectives of the paper

- To establish a Bayes rule analog for quantum mechanical processes
- To display the inadequacies of classical Bayes rule when applied to certain system evolutions in QM.
- To provide resolutions to Frauchiger-Renner's and Hardy's paradoxes.

Why can't we simply apply classical probability theory (CPT) directly onto QM?

- CPT does not rely on causal relations of events. In QM, experiments involving timelike and spacelike seperation regions are modelled differently:
  - Spacelike: $\mathcal{H}_1 \otimes \mathcal{H}_2$
  - Timelike: $\Lambda : \mathcal{L}(\mathcal{H}_1) \to \mathcal{L}(\mathcal{H}_2)$
- Joint probability distributions for a physical system cannot be defined for non-commuting operators.
- Entangled systems cannot be described via classical distributions (controversial?)

Setup

PHYSICAL REVIEW A **88**, 052130 (2013)

Towards a formulation of quantum theory as a causally neutral theory of Bayesian inference

M. S. Leifer[1,2,*] and Robert W. Spekkens[2,†]
[1]*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*
[2]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario, Canada N2L 2Y5*
(Received 19 June 2012; revised manuscript received 30 October 2013; published 27 November 2013)

Quantum theory can be viewed as a generalization of classical probability theory, but the analogy as it has been developed so far is not complete. Whereas the manner in which inferences are made in classical probability theory is independent of the causal relation that holds between the conditioned variable and the conditioning variable, in the conventional quantum formalism, there is a significant difference between how one treats experiments involving two systems at a single time and those involving a single system at two times. In this article, we develop the formalism of ***quantum conditional states***, which provides a unified description of these two sorts of experiment. In addition, concepts that are distinct in the conventional formalism become unified: Channels, sets of states, and positive operator valued measures are all seen to be instances of conditional states; the action of a channel on a state, ensemble averaging, the Born rule, the composition of channels, and nonselective state-update rules are all seen to be instances of belief propagation. Using a quantum generalization of Bayes' theorem and the associated notion of Bayesian conditioning, we also show that the remote steering of quantum states can be described within our formalism as a mere updating of beliefs about one system given new information about another, and retrodictive inferences can be expressed using the same belief propagation rule as is used for predictive inferences. Finally, we show that previous arguments for interpreting the projection postulate as a quantum generalization of Bayesian conditioning are based on a misleading analogy and that it is best understood as a combination of belief propagation (corresponding to the nonselective state-update map) and conditioning on the measurement outcome.

DOI: 10.1103/PhysRevA.88.052130 PACS number(s): 03.65.Ca, 03.65.Ta, 03.67.−a

Definitions

- **J,K,L**: Classical random variables
- **R,S,T**: "Snapshot" of the system (classical region)
- **A,B,C**: Quantum regions

What are J,K,L

"J" corresponds to the probability distribution of the observable wrt. the system at that time. In the QM formalism this simply corresponds to the probability distribution corresponding to the measurement outcomes of an observable. $P(Spin) = \{p(\uparrow), p(\downarrow)\}$

What are R,S,T

"R" corresponds to the set of all the observable quantities for the system. For particles in a box, R would contain the positions set and the momentum set.

"P(R)" corresponds to the set of probability distributions of all the observable quantities for the system at that time.
$P(R) = \{P(\hat{x}), P(\hat{p}), ...\}$

What are A,B,C?

"A" denotes a particular physical system, at a particular position in space, at a particular time.
Why do this?

Now $\mathcal{H}_{\mathcal{AB}} = \mathcal{H}_{\mathcal{A}} \otimes \mathcal{H}_{\mathcal{B}}$ for two disjoint regions, irrespective of what their nature of "disjointedness" is. (Different systems, same system in different time)

What are A,B,C?

What about states corresponding to timelike seperated regions? Why write them as a tensor product of an unrelated region?

This would simply now be $\rho_{AB} = \rho_A \otimes \mathbb{I}$
Thus evolution corresponding to a single physical system in two different times becomes $\rho_A \otimes \mathbb{I} \rightarrow \mathbb{I} \otimes \rho_B$. Furthermore, in this case $\mathcal{H}_A = \mathcal{H}_B$.

Analogs

Classical probability	Quantum theory
$P(R)$	ρ_A
$P(R,S)$	ρ_{AB}
$P(S) = \sum_R P(R,S)$	$\rho_B = \mathsf{Tr}_A(\rho_{AB})$
$\sum_S P(S\|R) = 1$	$\mathsf{Tr}_B(\rho_{B\|A}) = I_A$

(The summation statement is slightly misleading. Given the formalisms above,we would rather say:
$P(X) = \sum_Y P(X,Y) : \ \forall X \in R, \forall Y \in S$. The way written above is just for representing the generality.)

Analog for conditional probability

$$\sum_S P(S|R) = 1$$

An acausal conditional state for B given A is $\rho_{B|A}$ defined on $\mathcal{H}_{AB}$, such that $Tr_B(\rho_{B|A}) = \mathbb{I}_A$. This is satisfied by
$\rho_{AB} = (\sqrt{\rho_A} \otimes \mathbb{I}_B)\rho_{B|A}(\sqrt{\rho_A} \otimes \mathbb{I}_B)$
(Can be shown quite easily via partial traces on ρ_{AB}). By defining a

non-commutative and non-associative operator $M \star N = \sqrt{N} M \sqrt{N}$,
$\rho_{AB} = \rho_{B|A} \star \rho_A$

Quantum Bayes rule

Classical Bayes theorem:

$$P(R|S) = \frac{P(S|R)P(R)}{P(S)}$$

Quantum Bayes theorem:

$$\rho_{A|B} = \rho_{B|A} \star \left(\rho_A \rho_B^{-1}\right)$$

All analogs

Table: Translation of concepts and equations from conventional notation to the conditional states formalism.

	Conventional notation	Conditional states formalism
Probability distribution of X	$P(X)$	ρ_X
Probability that $X = x$	$P(X = x)$	$\rho_{X=x}$
Set of states on A	$\{\rho_x^A\}$	$\varrho_{A\|X}$
Individual state on A	ρ_x^A	$\varrho_{A\|X=x}$
POVM on A	$\{E_y^A\}$	$\varrho_{Y\|A}$
Individual effect on A	E_y^A	$\varrho_{Y=y\|A}$
Channel from A to B	$\mathcal{E}_{B\|A}$	$\varrho_{B\|A}$
Instrument	$\{\mathcal{E}^y_{B\|A}\}$	$\varrho_{Y,B\|A}$
Individual operation	$\mathcal{E}^y_{B\|A}$	$\varrho_{Y=y,B\|A}$
The Born rule	$\forall y : P(Y = y) = \mathrm{Tr}_A(E_y^A \rho_A)$	$\rho_Y = \mathrm{Tr}_A(\varrho_{Y\|A}\rho_A)$
Ensemble averaging	$\rho_A = \sum_x P(X = x)\rho_x^A$	$\rho_A = \mathrm{Tr}_X(\varrho_{A\|X}\rho_X)$
Action of a channel (Schrödinger)	$\rho_B = \mathcal{E}_{B\|A}(\rho_A)$	$\rho_B = \mathrm{Tr}_A(\varrho_{B\|A}\rho_A)$
Composition of channels	$\mathcal{E}_{C\|A} = \mathcal{E}_{C\|B} \circ \mathcal{E}_{B\|A}$	$\varrho_{C\|A} = \mathrm{Tr}_B(\varrho_{C\|B}\varrho_{B\|A})$
Action of a channel (Heisenberg)	$E_y^A = (\mathcal{E}_{B\|A})^*(E_y^B)$	$\varrho_{Y\|A} = \mathrm{Tr}_B(\varrho_{Y\|B}\varrho_{B\|A})$
Nonselective state-update rule	$\forall y : P(Y = y)\rho_y^B = \mathcal{E}^y_{B\|A}(\rho_A)$	$\rho_{Y,B} = \mathrm{Tr}_A(\varrho_{Y,B\|A}\rho_A)$

Example of discrepancy b/w classical and quantum probability

In this example we are talking about the evolution of two systems: One is the physical system under investigation, and the second is the apparatus which we utilise to obtain measurements and make inferences of the evolution of the physical system.

States of both are represented by two-levels: $(|0\rangle, |1\rangle)$ Causes belonging to physical system induces effects in the apparatus system.

Example of discrepancy b/w classical and quantum probability

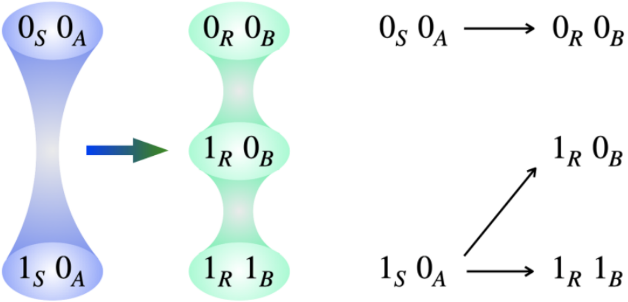

Figure: Quantum and stochastic evolution

Classically, $P(1_R 0_B|_1 S 0_A) = 1/2$ and $P(1_R 1_B|1_S 0_A) = 1/2$
A, B, S, R represent quantum regions for apparatus and physical system respectively.

Example of discrepancy b/w classical and quantum probability

The evolutions of the systems are given by some evolution V st. $|0_S\rangle|0_A\rangle \rightarrow |0_R\rangle|0_B\rangle$, $|1_S\rangle|0_A\rangle \rightarrow |1_R\rangle|+_B\rangle$ We will be observing effects in A, B and inferring their causes in R, S.

- Classical: $P(1_R 1_B | 1_S 0_A) = |\langle 1_R|\langle 1_B|V|1_S\rangle|0_B\rangle|^2 = 1/2$

  Using Born's rule, $P(1_S 0_A | 1_R 1_B) = \frac{P(1_R 1_B | 1_S 0_A) P(1_S 0_A)}{P(1_R 1_B)} = 1$ The effect $|1_R\rangle|1_B\rangle$ can deterministically infer the cause $|1_S\rangle|0_A\rangle$. This, in turn, means that $|1_S\rangle$ in S is the cause for the effect $|1_B\rangle$ observed in B.

Example of discrepancy b/w classical and quantum probability

- Quantum Bayes rule: Via the previous formalism, $\mathrm{Tr}_{RB}\left[\rho^{V^\dagger}_{SA|RB} \star |1_R\rangle\langle 1_R| \otimes |1_B\rangle\langle 1_B|\right] \neq |1_S\rangle\langle 1_S| \otimes |0_S\rangle\langle 0_S|$, or $V^\dagger|1_R\rangle|1_B\rangle = |1_S\rangle|-_A\rangle \neq |1_S\rangle|0_A\rangle$.

No deterministic relation, at most correlation with probability $1/2$.
In direct contradiction with classical probability.

(The evolution of the system is given by
$\rho^V_{SA|RB} = \sum_{m,n} |m_{SA}\rangle\langle m_{SA}| \otimes \Lambda(|m_{S'A'}\rangle\langle n_{S'A'}|)$,
$\mathcal{H}_{S'A'}$ is a copy of H_{SA} $\Lambda : \mathcal{L}(\mathcal{H}_{S'A'}) \mapsto \mathcal{L}(\mathcal{H}_{RB})$

More info

The authors go on to apply this to Hardy's paradox to show that the conclusion drawn by Hardy (that local-realistic interpretation is incorrect since it causes contradictory inferences) is an incorrect conclusion since classical probability cannot be applied to such cases.

Conclusion and discussion

- The authors argue that this formalism helps resolve certain paradoxes in QM. Whether this is true or not relies on finding possible extensions to the $\star$ operator, whose non-associativity and non-commutativity make it impossible give any framework when multiple probability distributions are involved.
- $|1_S\rangle|-_A\rangle$ is obtained even though $|1_S\rangle|1_A\rangle$ hasn't even been defined. This is at the heart of the Bayesian philosophy, that probabilities are not immutable or objective, but rather these are subjective values placed by rational agents who are the observers.