The rule of conditional probability is valid in quantum theory

P.G.L. Porta Mana
Kavli Institute, Trondheim <pgl@portamana.org>

***; updated 13 July 2020

***

1 ***

In a recent manuscript, Gelman & Yao[1] claim that "the usual rules of conditional probability fail in the quantum realm" and purport to support such statement with an example. Such a statement is false. I would like to recall some literature that shows why it is false, and to sum up the fallacy behind their example.

Similar claims with similar examples have appeared before in the quantum literature. Bernard O. Koopman, for example, discussed the falsity of such claims in 1957. The introduction to his work[2] is very explicit:

> Ever since the advent of modern quantum mechanics in the late 1920's, the idea has been prevalent that the classical laws of probability cease, in some sense, to be valid in the new theory. More or less explicit statements to this effect have been made in large number and by many of the most eminent workers in the new physics [. . .]. Some authors have even gone farther and stated that the formal structure of logic must be altered to conform to the terms of reference of quantum physics [. . .].
>
> Such a thesis is surprising, to say the least, to anyone holding more or less conventional views regarding the positions of logic, probability, and experimental science: many of us have been apt – perhaps too naively – to assume that experiments can lead to conclusions only when worked up

[1] Gelman & Yao 2020. [2] Koopman 1957.

> by means of logic and probability, whose laws seem to be on a different level from those of physical science.
>
> The primary object of this presentation is to show that the thesis in question is entirely without validity and is the product of a confused view of the laws of probability.

A claim similar to Gelman & Yao's, with a similar supporting example, was made by Brukner & Zeilinger[3], although their focus was on the validity of some properties of the Shannon entropy in quantum theory. That work has some fame in the quantum community.

The fallacy in their reasoning, which was based on incorrect calculations of conditional probabilities, was shown by Porta Mana[4] through a step-by-step analysis and calculation. This work also showed, through simple examples[5], that the same incorrect conclusions could be obtained with completely classical systems, such as drawing from an urn, if the same incorrect way of computing conditional probabilities was applied.

I refer to the cited work for the full analysis and counterexamples. Here I summarize the basic fallacy and give a simpler counterexample.

The basic fallacy is the confusion of the probability conditional with a temporal conditional. Take the conditional probability $\mathrm{P}(A \mid B)$, where the statement or event A refers to a time t_2, and B to a time t_1 preceding t_2. This probability is related to the reverse conditional $\mathrm{P}(B \mid A)$ by

$$\mathrm{P}(A \mid B)\,\mathrm{P}(B) = \mathrm{P}(B \mid A)\,\mathrm{P}(A) = \mathrm{P}(A \land B)\ . \tag{1}$$

It goes without saying that *the statements or events A and B must be the same on both sides of these equations*. In particular, in $\mathrm{P}(B \mid A)$ the statement B still refers to the time t_1, and A to the time t_2, with t_1 preceding t_2. We can represent these times explicitly and rewrite (2) as

$$\mathrm{P}(A_2 \mid B_1)\,\mathrm{P}(B_1) = \mathrm{P}(B_1 \mid A_2)\,\mathrm{P}(A_2) = \mathrm{P}(A_2 \land B_1)\ . \tag{2}$$

The fallacy is to think that, in calculating $\mathrm{P}(B \mid A)$, we should now ensure that B refers to time t_2, and A to time t_1, swapping the times. But these would be *different* events, not the original events. In symbols, we would be calculating $\mathrm{P}(B_2 \mid A_1)$, which is different from $\mathrm{P}(B_1 \mid A_2)$, to which the conditional-probability rule (2) refers.

[3] Brukner & Zeilinger 2001. [4] Porta Mana 2004. [5] Porta Mana 2004 § IV.

Probability theory has nothing to say, a priori, about the relation between $\mathrm{P}(B_1 \mid A_2)$ and $\mathrm{P}(B_2 \mid A_1)$. A simple example can illustrate this point.

Consider an urn with Red and Blue balls. The drawing scheme is with replacement for blue, and without replacement for red. That is, if a blue ball is drawn, it is put back in the urn (and the urn shaken) before the next draw. If a red ball is drawn, it is not put back before the next draw.

The urn initially has one blue and one red ball. The probabilities for the first draw are straightforward:

$$\mathrm{P}(B_1) = \tfrac{1}{2} \qquad \mathrm{P}(R_1) = \tfrac{1}{2}\ , \tag{3}$$

as are the conditional probabilities for the second draw, conditional on the first:

$$\mathrm{P}(B_2 \mid B_1) = \tfrac{1}{2} \qquad\qquad \mathrm{P}(R_2 \mid R_1) = 0 \tag{4a}$$

$$\mathrm{P}(R_2 \mid B_1) = \tfrac{1}{2} \qquad\qquad \mathrm{P}(B_2 \mid R_1) = 1\ . \tag{4b}$$

These probabilities can be visualized with the "equiprobable worlds" diagram of fig. 1. In half of the worlds the first draw yields red; in the other half, blue. In all worlds with red first, the second must yield blue. In half of the worlds with blue first, the second yields blue; and in the other half, red.

Let us now calculate the conditional probabilities for the first draw conditional on the second (imagine the first draw was hidden from you, and from seeing the second you're asked to guess what the first was). Using the conditional-probability rule (2) in the form of Bayes's theorem,

$$\mathrm{P}(X_1 \mid Y_2) = \frac{\mathrm{P}(Y_2 \mid X_1)\,\mathrm{P}(X_1)}{\sum_X \mathrm{P}(Y_2 \mid X_1)\,\mathrm{P}(X_1)}\ , \tag{5}$$

$$\begin{array}{lcccc}
 & \multicolumn{4}{c}{\text{equiprobable worlds}} \\
\text{1st draw} & \mathrm{R} & \mathrm{R} & \mathrm{B} & \mathrm{B} \\
 & \downarrow & \downarrow & \downarrow & \downarrow \\
\text{2nd draw} & \mathrm{B} & \mathrm{B} & \mathrm{B} & \mathrm{R}
\end{array}$$

Figure 1

we obtain

$$\mathrm{P}(B_1 \mid B_2) = \tfrac{1}{3} \qquad \mathrm{P}(R_1 \mid R_2) = 0 \tag{6a}$$

$$\mathrm{P}(R_1 \mid B_2) = \tfrac{2}{3} \qquad \mathrm{P}(B_1 \mid R_2) = 1\ . \tag{6b}$$

These conditional probabilities are intuitively correct, as can be checked by simple counting in fig. 1. For example, among all three worlds that have blue at the second draw, two of them have red at the first; hence $\mathrm{P}(R_1 \mid B_2) = 2/3$. Also, if we have red at the second draw, then logically red cannot have been drawn at the first, otherwise there would not have been any red left; hence blue must have been drawn at the first: $\mathrm{P}(B_1 \mid R_2) = 1$.

The following two relations between the conditional probabilities (4b) and (6b) are relevant to our discussion:

$$\mathrm{P}(R_2 \mid B_1) \boldsymbol{\neq} \mathrm{P}(R_1 \mid B_2) \qquad \mathrm{P}(B_2 \mid R_1) \boldsymbol{=} \mathrm{P}(B_1 \mid R_2)\ . \tag{7}$$

As we see, the probability-calculus does not prescribe the equality nor the inequality between events of similar kind but at different times (which therefore are *not* the same event). Any relations of this kind will depend on the specific situation. You can check, for example, that in a scheme of drawing with replacement for both blue and red, we would obtain two equalities in place of (7). An in a scheme of Pólya draws for blue (blue is returned plus an additional blue) and replacement draws for red, we would obtain two *in*equalities in place of (7).

Bibliography

("de *X*" is listed under D, "van *X*" under V, and so on, regardless of national conventions.)

Brukner, Č., Zeilinger, A. (2001): *Conceptual inadequacy of the Shannon information in quantum measurements*. Phys. Rev. A **63**, 022113. See also Porta Mana (2004).

Gelman, A., Yao, Y. (2020): *Holes in Bayesian statistics*. arXiv:2002.06467.

Koopman, B. O. (1957): *Quantum theory and the foundations of probability*. In: MacColl (1957), 97–102.

MacColl, L. A., ed. (1957): *Applied Probability*. (McGraw-Hill, New York).

Porta Mana, P. G. L. (2004): *Consistency of the Shannon entropy in quantum experiments*. Phys. Rev. A **69**^{6}, 062108. Rev. version at arXiv:quant-ph/0302049.