Journal of Mathematical Psychology 85 (2018) 17–24

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

journal homepage: www.elsevier.com/locate/jmp

On universality of classical probability with contextually labeled random variables

Ehtibar N. Dzhafarov *, Maria Kon

Purdue University, United States

HIGHLIGHTS

- Classical probability theory is adequate for describing all empirical phenomena.
- Random variables however should be identified (labeled) contextually.
- Examples used are Bell inequalities, order effects, and additivity of probabilities.
- Question-order effect is described by a noncontextual cyclic system of rank 2.
- Double-slit experiment is described by a noncontextual cyclic system of rank 4.

ARTICLE INFO

Article history:
Received 22 October 2017
Received in revised form 23 April 2018

Keywords:
Classical probability
Contextuality
Contextual labeling
Double-slit experiment
Question-order effects
Random variables

ABSTRACT

One can often encounter claims that classical (Kolmogorovian) probability theory cannot handle, or even is contradicted by, certain empirical findings or substantive theories. This note joins several previous attempts to explain that these claims are unjustified, illustrating this on the issues of (non)existence of joint distributions, probabilities of ordered events, and additivity of probabilities. The specific focus of this note is on showing that the mistakes underlying these claims can be precluded by labeling all random variables involved contextually. Moreover, contextual labeling also enables a valuable additional way of analyzing probabilistic aspects of empirical situations: determining whether the random variables involved form a contextual system, in the sense generalized from quantum mechanics. Thus, to the extent the Wang–Busemeyer QQ equality for the question order effect holds, the system describing them is noncontextual. The double-slit experiment and its behavioral analogues also turn out to form a noncontextual system, having the same probabilistic format (cyclic system of rank 4) as the one describing spins of two entangled electrons.

In the literature on foundations of quantum physics (Accardi, 1982; Feynman, 1951; Feynman, Leighton, & Sands, 1975; Khrennikov, 2009b) and, more recently, psychology (Aerts, 2009, 2014; Broekaert, Basieva, Blasiak, & Pothos, 2017; Busemeyer & Bruza, 2012; Moreira & Wichert, 2016; Pothos & Busemeyer, 2013), one can encounter statements that classical (Kolmogorovian) probability theory does not have adequate conceptual means to handle (sometimes, even, is contradicted by) this or that empirical fact.

Three of the most widespread assertions of this kind are as follows:

Statement 1: Classical probability requires that certain (e.g., Bell-type) inequalities hold for certain systems of random variables, but we know from quantum mechanics and from behavioral experiments that they may be violated.

Statement 2: In classical probability, the joint occurrence of two events is commutative, but we know from quantum mechanics and from behavioral experiments that the order of two events generally matters for their joint probability.

Statement 3: Classical probability is additive (equivalently, obeys the law of total probability), but we know from quantum mechanics and from behavioral experiments that this additivity (the law of total probability) can be violated.

This note has three objectives: (1) to show that the three statements above are based on misidentification of the random variables involved, due to ignoring their inherently contextual labeling; (2) to show that contextual labeling is a principled way to "automatically" ensure correct applicability of classical probability theory to an empirical situation; and (3) to demonstrate how the use of contextual labeling enables so-called contextuality analysis of systems of random variables, a relatively new form of probabilistic analysis of considerable interest in empirical applications. Contextual labeling of random variables is the departing

* Correspondence to: Department of Psychological Sciences, Purdue University, 703 Third Street West Lafayette, IN 47907, USA.
E-mail address: ehtibar@purdue.edu (E.N. Dzhafarov).

https://doi.org/10.1016/j.jmp.2018.06.001

principle of Khrennikov's Växjö Model (Khrennikov, 2009a) and of the Contextuality-by-Default theory (Dzhafarov, 2017; Dzhafarov, Cervantes, & Kujala, 2017; Dzhafarov & Kujala, 2014a, 2015, 2016a, 2017a, b; Dzhafarov, Kujala, & Cervantes, 2016).

Let us preamble this discussion by stating our view of classical probability theory (CPT), one that we are not prepared to defend in complete generality, confining ourselves instead to merely illustrating it on the three statements above. This view is that CPT, on a par with classical logic and set theory, is a universal abstract mathematical theory. As an abstract mathematical theory, it does not make empirically testable predictions, because of which it cannot be contradicted by any empirical observation. As a universal theory, for any empirical situation, it has conceptual means to adequately describe anything that can be qualified as this situation's probabilistic features (in the frequentist sense). Moreover, as a conceptual tool, in the same way as classical logic and set theory, it is indispensable and irreplaceable in dealing with probabilistic problems: at the end, the results of any non-classical probabilistic analysis have to be formulated in terms of classical (frequentist) probabilities, distributions, and random variables. However, when applied to an empirical situation, CPT can (even must) be complemented by special-purpose computations identifying some of the random variables, distributions, and probabilities in this particular situation. To give a very simple example, CPT provides methods for deriving probabilities of events defined on the outcomes of rolling a die from a distribution of these outcomes, but it cannot predict this distribution. A special theory is needed to know, e.g., that if a die is manufactured in a particular way, then the distribution of its outcomes is uniform. We view quantum probability as such a special-purpose theory complementing classical probability. This mathematical formalism is indispensable in quantum mechanics and has significant achievements to its credit in psychology (e.g., Wang & Busemeyer, 2013). It can be formalized and presented as an abstract calculus alternative to or even generalizing the calculus of CPT, in the same way one can formalize a paraconsistent logic as a generalization of classical logic. However, just as one cannot replace classical logic with paraconsistent logic in analyzing anything, including the very paraconsistent logic itself, one cannot dispense with classical probability when discussing and analyzing quantum probability computations and relating them to data.

This view is not entirely new. Ballentine (1986) defended a similar position in essentially the same way we are doing here. The difference is in that instead of using random variables, Ballentine confined himself to a more limited language of events, and he used conditionalization in place of the more general contextualization (Dzhafarov & Kujala, 2014b; we discuss conditionalization in Section 3). Khrennikov (2009a), in describing his Växjö contextual model uses Ballentine's conditional-probability notation, but emphasizes that these are not conditional probabilities of CPT. Rather he calls them "contextual probabilities," and explains that "contextual probability [. . .] is not probability that an event, say B, occurs under the condition that another event, say C, has occurred. The contextual probability is probability to get the result $a = \alpha$ under the complex of physical conditions C" (Khrennikov, 2009a, p. 50). This seems to be the same as the contextual labeling used in the Contextuality-by-Default theory. A very clear presentation of a position that is close to ours can be found in the arguments presented in an internet discussion by Maudlin (2013).

The purpose of this paper is to achieve conceptual clarity in understanding CPT, not to criticize specific authors or papers. The latter is an ungrateful task, as most authors' positions are not entirely consistent, are subject to (re)interpretations, and evolve over time. We cite specific papers and occasionally provide quotes only to demonstrate that a reasonable reader may interpret the positions they entail in the spirit of the Statements 1–3 above. Thus, Richard Feynman is often cited as arguing that classical probability is not compatible with quantum mechanics (Accardi, 1982; Costantini, 1993; Khrennikov, 2009b). This interpretation is supported by Feynman's speaking of "the discovery that in nature the laws of combining probabilities were not those of the classical probability theory of Laplace" (Feynman, 1951, p. 533). However, one can also find statements in Feynman's writings that make his point of view less than unequivocal. Thus, we read in the same paper and on the same page that "the concept of probability is not altered in quantum mechanics. When I say the probability of a certain outcome of an experiment is p [. . .] no departure from the concept used in classical statistics is required. What is changed, and changed radically, is the method of calculating probabilities" (*ibid*). This quote is consistent with treating quantum formalisms as special-purpose computations embedded in CPT. We will return to Feynman when discussing the double-slit experiment in Section 3.

1. On Statement 1

> "Classical probability requires that certain (e.g., Bell-type) inequalities hold for certain sets of random variables, but we know from quantum mechanics and from behavioral experiments that they may be violated."

This view is commonly held in both physics and psychology (Aerts, 2009; Aerts & Sozzo, 2011; Bruza, Kitto, Nelson, & McEvoy, 2009; Busemeyer & Bruza, 2012; Filipp & Svozil, 2005; Khrennikov, 2009b; Yearsley & Pothos, 2014). In particular, among those applying quantum probability to behavior and also treating quantum probability theory as an alternative to CPT, there are claims that Bell-type inequalities are violated in experiments involving combinations of concepts (Aerts & Sozzo, 2011; Busemeyer & Bruza, 2012) and memory (Bruza, Kitto, Nelson, & McEvoy, 2009).

We will not recapitulate all the arguments related to this issue, as they have been presented in many previous publications (Dzhafarov, Cervantes, & Kujala, 2017; Dzhafarov & Kujala, 2014a, b, 2016a, 2017a, b; Dzhafarov, Kujala, & Larsson, 2015). We will use just one familiar example. Let R_1, R_2, R_3, R_4 denote a set of binary $(+1/-1)$ random variables with known distributions of $(R_1, R_2), (R_2, R_3), (R_3, R_4)$, and (R_4, R_1). The necessary and sufficient condition for the existence of such a quadruple of random variables is given by the CHSH/Fine inequality (Bell, 1964; Clauser, Horne, Shimony, & Holt, 1969; Fine, 1982):

$$\max_{j=1,\ldots,4}\left|\sum_{i=1}^{4}\langle R_i R_{i\oplus 1}\rangle - 2\left\langle R_j R_{j\oplus 1}\right\rangle\right| \leq 2, \tag{1}$$

where $\oplus 1$ is cyclic shift $1 \to 2 \to 3 \to 4 \to 1$, and $\langle\cdot\rangle$ is expectation. One can easily construct examples of distributions of $(R_i, R_{i\oplus 1})$ for which this inequality is violated, indicating that such R_1, R_2, R_3, R_4 do not exist (essentially by the same logic as in determining that there are no four numbers a, b, c, d with $a = b$, $b = c$, $c = d$, and $d = a + 1$).

The problem arises when we are being told that the existence of such R_1, R_2, R_3, R_4 is predicted by quantum theory and corroborated by experiment. If we believe this, violations of (1) should indeed mean that CPT is inadequate, if not internally contradictory. We should not, however, believe this. R_1, R_2, R_3, R_4 in (1) are random variables in the CPT sense; they are not within the language of quantum theory. To decide what classical random variables should describe outcomes of what quantum measurements, one needs to go outside quantum theory. The general rule is that a random variable is identified by what is being measured and how it is being measured. The latter includes all conditions under which the measurement is made, in particular, all other measurements performed