ARTICLE IN PRESS

Journal of Mathematical Psychology ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect

Journal of Mathematical Psychology

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

Stochastic unrelatedness, couplings, and contextuality

Ehtibar N. Dzhafarov

Purdue University, United States

HIGHLIGHTS

- Random variables in mutually exclusive contexts possess no joint distribution.
- They can be coupled, i.e., joint distributions can be imposed on them, non-uniquely.
- They can be characterized by what couplings they allow.
- In particular, this is a way to characterize selective influences and contextuality.
- Lack of joint distribution must not be confused with stochastic independence.

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Contextuality
Coupling
Joint distribution
Probability
Random variable
Stochastic relation
Stochastic unrelatedness

ABSTRACT

R. Duncan Luce once mentioned in a conversation that he did not consider Kolmogorov's probability theory well-constructed because it treats stochastic independence as a "numerical accident", while it should be treated as a fundamental relation, more basic than the assignment of numerical probabilities. I argue here that stochastic independence is indeed a "numerical accident", a special form of stochastic dependence between random variables (most broadly defined). The idea that it is fundamental may owe its attractiveness to the confusion of stochastic independence with stochastic unrelatedness, the situation when two or more random variables have no joint distribution, "have nothing to do with each other". Kolmogorov's probability theory cannot be consistently constructed without allowing for stochastic unrelatedness, in fact making it a default situation: any two random variables recorded under mutually incompatible conditions are stochastically unrelated. However, stochastically unrelated random variables can always be probabilistically coupled, i.e., imposed a joint distribution upon, and this generally can be done in an infinity of ways, independent coupling being merely one of them. The notions of stochastic unrelatedness and all possible couplings play a central role in the foundation of probability theory and, especially, in the theory of probabilistic contextuality.

1. Introduction

Almost 15 years ago R. Duncan Luce mentioned in a conversation that the Kolmogorovian probability theory (KPT) was unsatisfactory because it treated stochastic independence as a "numerical accident" rather than a fundamental relation. If I roll a die today in Irvine, California, Duncan said, and on another day you roll a die in Lafayette, Indiana, the fact that the two outcomes are independent cannot be established by checking the multiplication rule. On the contrary, the applicability of the multiplication rule in this case is justified by determining that the two dice are stochastically independent, "have nothing to do with each other".

This simple example (some may think too simple to be of great interest) leads us to the very foundations of probability theory. Let us try to understand it clearly by comparing it to another example. It is about a situation when I repeatedly roll a single die, having defined two random variables:

$$A = \begin{cases} 1 & \text{if the outcome is even} \\ 0 & \text{otherwise,} \end{cases}$$

$$B = \begin{cases} 1 & \text{if the outcome exceeds 3} \\ 0 & \text{otherwise.} \end{cases}$$

These two random variables co-occur in the most obvious empirical meaning: the values of A and B are always observed together, at every roll of the die. Another way of looking at it, the two random variables co-occur because they are functions of one and the same "background" random variable Z, the outcome of rolling the die. As a result, I can estimate from the observations the probabilities $\Pr[A = 1 \text{ and } B = 1]$, $\Pr[A = 1]$, and $\Pr[B = 1]$ (I will use Pr as a symbol for probability throughout this paper): if the

E-mail address: ehtibar@purdue.edu.

http://dx.doi.org/10.1016/j.jmp.2016.01.004

joint probability turns out to be the product of the two marginal ones (statistical issues aside), the two events are determined to be independent. I cannot simply make this determination a priori, as it depends on what die I am rolling: if it is a fair die, A and B are not independent, but if the distribution of the outcomes is

$$\begin{array}{lcccccc} \textit{value}: & 1 & 2 & 3 & 4 & 5 & 6 \\ \textit{pr.mass}: & 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0, \end{array}$$

then A and B are independent.

The difference between this example and that of Duncan Luce's is not in the number of the dice being rolled: my example would not change too much if I roll two dice together, having marked them "Left" and "Right", and define the random variables as

$$A = \begin{cases} 1 & \text{if the Left outcome is even} \\ 0 & \text{otherwise,} \end{cases}$$

$$B = \begin{cases} 1 & \text{if the Right outcome exceeds 3} \\ 0 & \text{otherwise.} \end{cases}$$

The realizations of A and B again come together, this time the empirical meaning of the "togetherness" being "in the same trial", or "simultaneously". Again, one can also say that the two random variables co-occur because they are functions of one and the same "background" random variable Z, only this time it is the pair of values rather than a single one. And again, I can estimate from the observations the probabilities $\Pr[A = 1 \text{ and } B = 1]$, $\Pr[A = 1]$, and $\Pr[B = 1]$ and check their adherence to the multiplication rule. Whether the two random variables are stochastically independent is determined by the outcome of this test: the dice may very well be rigged not to be independent.

In Duncan Luce's example the situation is very different: the outcomes of rolling the two dice in two different places at two different times have no empirically defined pairing. If I define my random variables as

$$A = \begin{cases} 1 & \text{if on Tuesday in Irvine the outcome is even} \\ 0 & \text{otherwise,} \end{cases}$$

$$B = \begin{cases} 1 & \text{if on Friday in Lafayette the outcome exceeds 3} \\ 0 & \text{otherwise,} \end{cases}$$

then I can estimate empirically the probabilities $\Pr[A = 1]$, and $\Pr[B = 1]$ and find out, e.g., that they are (statistical issues aside) 0.7 and 0.5, respectively. But I cannot estimate empirically $\Pr[A = 1 \text{ and } B = 1]$: the two random variables are not recorded in pairs. The experiment involves no empirical procedure by which one could find which value of B should be paired with which value of A. The two random variables therefore do not have an observable (estimable from frequencies) joint distribution, they cannot be presented as functions of one and the same "background" random variable. What one can do, however, is to *declare* the two random variables stochastically independent, based on one's understanding that they "have nothing to do with each other". If one does so, the validity of $\Pr[A = 1 \text{ and } B = 1]$ being equal to the product of two individual probabilities is true *by construction*, requiring no empirical testing and allowing for no empirical falsification.

This was Duncan Luce's point: while the KPT defines stochastic independence through the multiplication rule, at least in some cases the determination of independence precedes and justifies the applicability of the multiplication rule. In Duncan Luce's opinion, this warranted treating stochastic independence as a fundamental, "qualitative" relation preceding assignment of numerical probabilities. This opinion is in accordance with the general precepts of the representational theory of measurement. Thus, the authors of the first volume of Foundations of Measurement (Krantz, Luce, Suppes, & Tversky, 1971) sympathetically refer to Zoltan Domotor 1969 dissertation in which he axiomatized probability theory treating stochastic independence as a primitive relation. As far as I know, however, it has not translated into a viable alternative to the KPT.

I accept Duncan Luce's example as posing a genuine foundational problem, but I disagree that this problem is about defining independence by means other than the multiplication rule. The position I advocate below in this paper is as follows.

1. Random variables that "have nothing to do with each other" are defined on different domains (sample spaces). Rather than being independent (which is a form of a joint distribution), they are *stochastically unrelated*, i.e., they possess no joint distribution.
2. It is not that we do not know the "true" distribution, or that in "truth" they are independent but we do not know how to justify this. A joint distribution simply is not defined (until imposed by us in one of multiple ways, discussed below).
3. The KPT is consistent with the idea of multiple sample spaces and in fact requires it for internal consistency: the idea of a single sample space for all random variables imaginable is mathematically untenable.
4. Any given set of pairwise stochastically unrelated random variables can always be *coupled*, i.e., imposed a joint distribution on. This is equivalent to inventing a pairing scheme for their realizations, and this can be done in multiple ways, coupling them as independent random variables being just one of them.

2. On random variables, unrelatedness, and independence

2.1. Informal introduction

Stochastic unrelatedness is easy to distinguish from stochastic independence: the latter assumes the existence of a joint distribution, which means that an empirical procedure exists by which each realization of one random variables can be paired (coupled) with that of another. The most familiar forms of empirical coupling are co-occurrence in the same trial and co-relation to the same person. In the table below,

$$\begin{array}{lcccccc} c: & 1 & 2 & 3 & 4 & 5 & \ldots \\ X: & x_1 & x_2 & x_3 & x_4 & x_5 & \ldots \\ Y: & y_1 & y_2 & y_3 & y_4 & y_5 & \ldots, \end{array} \qquad (1)$$

the indexing entity c can be the number of a trial (as in repeatedly rolling two marked dice together) or an ID of a person (as in relating heights and weights, or weights before and after dieting). The random variables X and Y here have a joint distribution: one can, e.g., estimate the probability with which X falls within an event E_X and ("simultaneously") Y falls within an event E_Y; and if

$$\Pr[X \in E_X \& Y \in E_Y] = \Pr[X \in E_X] \Pr[Y \in E_Y], \qquad (2)$$

for any two such events E_X, E_Y, then X and Y are considered independent.

Suppose, however, that the information about c in (1) does not exist, and all one has is some set of values for X and some set of values for Y. Clearly, now the "togetherness" of $X \in E_X$ and $Y \in E_Y$ is undefined. Although $\Pr[X \in E_X]$ and $\Pr[Y \in E_Y]$ have the same meaning as before, $\Pr[X \in E_X \,\&\, Y \in E_Y]$ is undefined, and (2) cannot be tested. This is what stochastic unrelatedness is: lack of a joint distribution. A pair of stochastically unrelated random variables are neither independent nor interdependent, these terms do not apply.

Think, e.g., of a list of weights in some group of people before dieting (X) and a list of weights in some other group of people after dieting (Y): which value of X should be paired with which value