# Random Variables – without Basic Space

Götz Kersting *

June 11, 2007

**Abstract**

The common definition of a random variable as a measurable function works well 'in practice', but has conceptual shortcomings, as was pointed out by several authors. Here we treat random variables not as derived quantities but as mathematical objects, whose basic properties are given by intuitive axioms. This requires that their target spaces fulfil a minimal regularity condition saying that the diagonal in the product space is measurable. From the axioms we deduce the basic properties of random variables and events.

*MSC 2000 subject classifications.* Primary 60A05
*Key words and phrases.* Random variable, event, a.s. equality, convergence, measurable space with denumerable separation

## 1 Introduction

In this paper we define the concept of a *stochastic ensemble*. It is our intention thereby to give an intuitive axiomatic approach to the concept of a random variable. The primary ingredient is a sufficiently rich collection of random variables (with 'good' target spaces). The set of observable events will be derived from it.

Among the notions of probability it is the *random variable*, which in our view constitutes the fundamental object of modern probability theory. Albeit in the history of mathematical probability *events* came first, random variables are closer to the roots of understanding nondeterministic phenomena. Nowadays events typically refer to random variables and are no longer studied for their own sake, and for *distributions* the situation ist not much different. Moreover, random variables turn out to be flexible mathematical objects. They can be handled in other ways than events or distributions (think of couplings), and these ways often conform to intuition. 'Probabilistic', 'pathwise' methods gain importance and combinatorial constructions with random variables can substitute (or nicely prepare) analytic methods. It was a common believe that first of

*Fachbereich Informatik und Mathematik, Universität Frankfurt, Fach 187, D-60054 Frankfurt am Main, Germany, kersting@math.uni-frankfurt.de

all the distributions of random variables matter in probability, but this believe is outdated.

Today it is customary to adapt random variables to a context from measure theory. Yet the feeling has persisted that random variables are objects in their own right. This was manifest, when measure theory took over in probability: According to J. Doob (interviewed by Snell [9]) "it was a shock for probabilists to realize that a function is glorified into a random variable as soon as its domain is assigned a probability distribution with respect to which the function is measurable." Later the experts insisted that it is the idea of random variables, which conforms to intuition. Legendary is L. Breiman's [2] statement: "Probability theory has a right and a left hand. On the right is the rigorous foundational work using the tools of measure theory. The left hand 'thinks probabilistically,' reduces problems to gambling situations, coin-tossing, and motions of a physical particle." In applications of probability the concept of a random variable never lost its appeal. We may quote D. Mumford [8]: "There are two approaches to developing the basic theory of probability. One is to use wherever possible the reduction to measure theory, eliminating the probabilistical language ... The other is to put the concept of 'random variable' on center stage and work with manipulations of random variables wherever possible". And, "for my part, I find the second way ... infinitely clearer".

**Example.** To illustrate this assertion let us consider different proofs of the central limit theorem saying that $(X_1 + \ldots + X_n)/\sqrt{n}$ is asymptotically normal for iid random variables $X_1, X_2, \ldots$ with mean 0 and variance 1. There is the established analytic approach via characteristic functions. In contrast let us recall a coupling method taken from [2], which shortly speaking consists in replacing $X_1, \ldots, X_n$ one after the other by independent standard normal random variables $Y_1, \ldots, Y_n$. In more detail this looks as follows: Let $f : \mathbb{R} \to \mathbb{R}$ be thrice differentiable, bounded and with bounded derivatives. Then it is sufficient to show that

$$\mathbf{E}\Big[f\Big(\frac{X_1 + \cdots + X_n}{\sqrt{n}}\Big) - f\Big(\frac{Y_1 + \cdots + Y_n}{\sqrt{n}}\Big)\Big]$$

converges to zero. The integrand may be expanded into

$$\sum_{i=1}^{n} \Big[f\Big(\frac{Z_i}{\sqrt{n}}\Big) - f\Big(\frac{Z_{i-1}}{\sqrt{n}}\Big)\Big]$$

with $Z_i := X_1 + \cdots + X_i + Y_{i+1} + \cdots + Y_n$. By means of two Taylor expansions around $U_i := X_1 + \cdots + X_{i-1} + Y_{i+1} + \cdots + Y_n$ the summands turn into

$$\frac{X_i - Y_i}{\sqrt{n}} f'(U_i) + \frac{X_i^2 - Y_i^2}{2n} f''(U_i) + O_P(n^{-3/2})$$

Taking expectations the first two terms vanish because of independence, and a closer look at the remainder gives the assertion (see [2], page 168). $\square$

From an architectual point of view these considerations and statements suggest to try and start from random variables in the presentation of probability theory and therewith to bring intuition and methods closer together – rather than to gain random variables as derived quantities in the accustomed measure-theoretic manner. We like to show that this can be accomplished without much technical effort. For this purpose we may leave aside distributions in this paper.

Let us comment on the difference of our approach to the customary one choosing a certain $\sigma$-field $\mathcal{E}$ on some basic set $\Omega$ as a starting point and then to identify events and random variables with measurable sets and measurable functions. In our view this is a set-theoretic *model* of the probabilistic notions.

To explain this first by analogy let us recall, how natural numbers are treated in mathematics. There are two ways: Either one starts from the well-known Peano axioms. Then the *set* of natural numbers is the object of study, and a natural number is nothing more than an element of this structured set. Or natural numbers are introduced by a set-theoretic construction, e.g. $0 := \emptyset, 1 := \{0\}, \dots, n+1 := n \cup \{n\}, \dots$ (see [7]). This setting exhibits aspects, which are completely irrelevant for natural numbers (such as $n \in n+1$ and $n \subset n+1$) and which stress that we are dealing with a *model* of the natural numbers. Thus, if one considers different models, they will not be similar (isomorphic) in any respect.

Analoguous observations can be made in our context, if events and random variables are represented by a measure space $(\Omega, \mathcal{E})$ and associated measurable functions. Note the following: There are the subsets of $\Omega$ not belonging to $\mathcal{E}$, which are totally irrelevant. To some extent this is also true for the elements $\omega$ of $\Omega$ (as also Mumford [8] pointed out; already Caratheodory considered integration on spaces without points in his theory of soma [3]). The 'small omegas' do not show up in any relevant result of probability theory, and one could do without them, if they were not needed to *define* measurable functions. Next the notion of a random variable is ambiguous: There are random variables and a.s. defined random variables, represented by measurable functions and equivalence classes of measurable functions. This distinction, though unavoidable in the traditional setting, is somewhat annoying. Finally note that probabilists leave aside the question of isomorphy of measurable spaces.

All these observations indicate that measurable spaces and mappings indeed make up a model of events and random variables. This is not to say that such models should be avoided, but one should not overlook that they might mislead. Aspects like the construction of non-Borel-sets are of no relevance in probability and may detract beginners. Also one should be cautious in giving the elements of $\Omega$ some undue relevance ('state of the world'), which may create misconceptions.

**Example.** This example of possible misconception is taken from the textbook [1] (Example 4.6 and 33.11). Let $\Omega = [0, 1]$, endowed with the Borel-$\sigma$-field and Lebesgue-measure $\lambda$. Let $\mathcal{F}$ be the sub-$\sigma$-field of sets $B$ with $\lambda(B) = 0$ or 1. Then $\mathcal{F}$ presents an observer, who lacks information. It is mistaken to argue that $\mathcal{F}$ presents full information, because it contains all one point sets such that the observer can recognize, which event $\{\omega\}$ takes place and which 'state' $\omega$ is valid. Therefore for any Borel-set $E \subset \Omega$ the conditional probability $\lambda(E|\mathcal{F})$ ist $\lambda(E)$ a.s., and in general not $\mathbf{1}_E$ a.s. $\square$

The eminent geometer H. Coxeter pinpoints such delusion due to models in stating: "When using models, it is desirable to have two rather than one, so as to avoid the temptation to give either of them undue prominence. Our ... reasoning should all depend on the axioms. The models, having served their purpose of establishing relative consistency, are no more essential than diagrams" (see section 16.2 in [4]). Coxeter has the circle and halfplane models of hyperbolic geometry in mind, but certainly his remark applies more generally.

An axiomatic concept of random variables should avoid the asserted flaws. The reader may judge our approach from this viewpoint. This paper owes a lot to discussions with Hermann Dinges, who put forward related ideas already in [5] (jointly with H. Rost). For further discussion we refer to H. Dinges [6] and D. Mumford [8].

The paper is organized as follows. In section 2 we have a look on those properties of events and random variables irrespective of a measure-theoretic representation (this section may be skipped). In section 3 we discuss the class of measurable spaces, which are suitable to serve as target spaces of random variables. Section 4 contains the axioms for general systems of random variables, which we call stochastic ensembles. In section 5 we derive events and deduce their properties from these axioms. In section 6 we discuss equality and a.s. equality of random variables. In section 7 we address convergence of random variables in order to exemplify how to work within our framework of axioms.

# 2 Events and random variables – an outline

Random variables and events rely on each other. Random variables can be examined from the perspective of events, and vice versa. In this section we describe this interplay in an non-systematic manner and detached from the measure-theoretic model, in order to introduce into our subject.

The field $\mathcal{E}$ of events is a $\sigma$-complete Boolean lattice. In particular:

- Each event $E$ possesses a complementary event $E^c$.
- For any finite or infinite sequence $E_1, E_2, \ldots$ of events there exists its union $\bigcup_n E_n$ and its intersection $\bigcap_n E_n$.
- There are the sure and the impossible events $E_{\text{sure}}$ and $E_{\text{imp}}$.

Also $E_1 \subset E_2$, iff $E_1 \cap E_2 = E_1$. Since events are no longer considered as subsets of some space, unions and intersections have to be interpreted here in the lattice-theoretic manner.

A random variable $X$ first of all has a *target space* $S$ equipped with a $\sigma$-field $\mathcal{B}$. Intuitively $S$ is the set, where $X$ may take its values. Collections of random variables obey the following simple rules:

- To each random variable $X$ with target space $S$ and to each measurable $\varphi : S \to S'$ a random variable with target space $S'$ is uniquely associated, denoted by $\varphi(X)$.
- To each sequence $X_1, X_2, \ldots$ of random variables with target spaces $S_1, S_2, \ldots$ a random variable with target space $S_1 \times S_2 \times \cdots$ equipped with the product-$\sigma$-field is uniquely associated, denoted by $(X_1, X_2, \ldots)$.

The corresponding calculation rules are obvious, we come back to them. We point out that not every measurable space is suitable as a target space – a minimal condition will be given in the next section. Uncountable products $\otimes_{i \in I}(S_i, \mathcal{B}_i)$ of measurable spaces are in general no admissible target spaces. This conforms to the fact that in probability an uncountable family of random variables $(X_i)_{i \in I}$ is at most provisionally considered as a single random variable with values in the product space, before proceeding to a better suited target space.

The connection between random variables and events is established by the remark that to any random variable $X$ and to any measurable subset $B$ of its target space $S$ an event

$$\{X \in B\}$$

is uniquely associated. The events $\{X \in B\}$ uniquely determine $X$, where $B$ runs through the measurable subsets of $S$. The calculation rules are

$$\Big\{X \in \bigcup_n B_n\Big\} \;=\; \bigcup_n \{X \in B_n\} \;,\; \Big\{X \in \bigcap_n B_n\Big\} \;=\; \bigcap_n \{X \in B_n\} \;,$$
$$\{X \in B^c\} \;=\; \{X \in B\}^c \;,\; \{X \in S\} \;=\; E_{\text{sure}} \;,\; \{X \in \emptyset\} \;=\; E_{\text{imp}} \;,$$

where $B, B_1, B_2, \ldots$ are measurable subsets of the target space of $X$. If these properties hold, the mapping $B \mapsto \{X \in B\}$ is called a *$\sigma$-homomorphism*. Moreover

$$\{\varphi(X) \in B'\} \;=\; \{X \in B\} \;, \quad \text{where } B = \varphi^{-1}(B')$$
$$\{(X_1, X_2, \ldots) \in B_1 \times B_2 \times \cdots\} \;=\; \bigcap_n \{X_n \in B_n\} \;.$$

From the perspective of events the connection to random variables is as follows: For any event $E$ there is a random variable $I_E$ with values in $\{0, 1\}$, the indicator variable of $E$, fulfilling

$$\{I_E = 1\} \;=\; E \;,\; \{I_E = 0\} \;=\; E^c \;.$$

For any infinite sequence $E_1, E_2, \ldots$ of disjoint events there is a random variable $N = \min\{n : E_n \text{ occurs}\}$ with values in $\{1, 2, \ldots, \infty\}$ such that

$$\{N = n\} \ = \ E_n \ , \ \{N = \infty\} \ = \ \bigcap_n E_n^c \ .$$

For any infinite sequence $E_1, E_2, \ldots$ of events (disjoint or not) there is a random variable $X$ and measurable subsets $B_1, B_2, \ldots$ of its target space such that

$$\{X \in B_n\} \ = \ E_n$$

for all $n$ (see section 5).

This is about all, what mathematically can be stated about events and random variables. A systematic treatment requires an axiomatic approach. There are two possibilities, namely to start from events or from random variables.

Either the starting point is the field of events, which is assumed to be a $\sigma$-complete Boolean lattice $\mathcal{E}$. Then a random variable $X$ with target space $(S, \mathcal{B})$ is nothing else but a $\sigma$-homomorphism from $\mathcal{B}$ to $\mathcal{E}$. It is convenient to denote it as $B \mapsto \{X \in B\}$ again. In this approach some technical efforts are required to show that any sequence $X_1, X_2, \ldots$ of random variables may be combined to a single random variable $(X_1, X_2, \ldots)$.

Starting from random variables instead is closer to intuition to our taste. Also it circumvents the technical efforts just mentioned. This approach will be put forward in the following sections.

# 3 Spaces with denumerable separation

Not every measurable space qualifies as a possible target space. We require that there exists a denumerable system of measurable sets separating points.

**Definition.** *A measurable space* $(S, \mathcal{B})$ *is called* measurable space with denumerable separation (mSdS)*, if there is a denumerable* $\mathcal{C} \subset \mathcal{B}$ *such that for any pair* $x \neq y$ *of elements in* $S$ *there is a* $C \in \mathcal{C}$ *such that* $x \in C$ *and* $y \notin C$.

**Examples.**

1. Any separable metric space together with its Borel-$\sigma$-algebra is a mSdS. This includes the case of denumerable $S$ and in fact any relevant target space of random variables considered in probability.

2. If $(S_1, \mathcal{B}_1), (S_2, \mathcal{B}_2), \ldots$ is a sequence of mSdS, then also the product space $\otimes_n (S_n, \mathcal{B}_n)$ is a mSdS. Indeed, if $\mathcal{C}_1, \mathcal{C}_2, \ldots$ are the separating systems, then

$$\mathcal{C} \ := \ \bigcup_n \{S_1 \times \cdots \times S_{n-1} \times C_n \times S_{n+1} \times \cdots \ : \ C_n \in \mathcal{C}_n\}$$

is denumerable and separating in the product space.

3. An uncountable product of measurable spaces is no mSdS (up to trivial cases). The reason is that these product-$\sigma$-field does not contain the one point sets (see below). $\square$

A mSdS $(S, \mathcal{B})$ has two important properties. Firstly one point subsets $\{x\}$ are measurable, since

$$\{x\} = \bigcap_{C \in \mathcal{C}\,,\, x \in C} C$$

for all $x \in S$. Secondly the 'diagonal'

$$D := \{(x, y) \in S^2 \ : \ x = y\}$$

is measurable in the product space $(S^2, \mathcal{B}^2)$, since

$$D = \bigcap_{C \in \mathcal{C}} C \times C \cup C^c \times C^c \ . \tag{1}$$

These properties are crucial for target spaces of random variables. Remarkably the second one is characteristic for mSdS.

**Proposition 1.** *A measurable space $(S, \mathcal{B})$ is mSdS, if and only if $D \in \mathcal{B}^2$.*

*Proof.* It remains to prove that $D \in \mathcal{B}^2$ implies the existence of a denumerable separating system $\mathcal{C}$. Let

$$\mathcal{F} := \bigcup_{\mathcal{C}} \sigma(\mathcal{C}) \otimes \sigma(\mathcal{C}) \ ,$$

where $\sigma(\mathcal{C})$ is the $\sigma$-field generated by $\mathcal{C}$ and the union is taken over all denumerable $\mathcal{C} \subset \mathcal{B}$. $\mathcal{F}$ is a sub-$\sigma$-field of $\mathcal{B} \otimes \mathcal{B}$ containing all $B_1 \times B_2$ with $B_1, B_2 \in \mathcal{B}$, thus

$$\mathcal{B} \otimes \mathcal{B} = \bigcup_{\mathcal{C}} \sigma(\mathcal{C}) \otimes \sigma(\mathcal{C}) \ .$$

By assumption it follows that $D \in \sigma(\mathcal{C}) \otimes \sigma(\mathcal{C})$ for some denumerable $\mathcal{C} \subset \mathcal{B}$. We show that $\mathcal{C} \cup \{C^c : C \in \mathcal{C}\}$ is a separating system. Let $x, y \in S, x \neq y$. Then $D$ does not belong to the $\sigma$-field

$$\mathcal{G} := \{B \in \sigma(\mathcal{C}) \otimes \sigma(\mathcal{C}) \ : \ \{(x,x),(y,x)\} \subset B \text{ or } \{(x,x),(y,x)\} \subset B^c\} \ .$$

It follows $\mathcal{G} \neq \sigma(\mathcal{C}) \otimes \sigma(\mathcal{C})$, thus there are $B_1, B_2 \in \sigma(\mathcal{C})$ such that $B_1 \times B_2 \notin \mathcal{G}$. Thus $B_1$ contains $x$ or $y$, but not both, and consequently is not an element of the $\sigma$-field

$$\mathcal{H} := \{B \in \sigma(\mathcal{C}) : \{x, y\} \subset B \text{ or } \{x, y\} \subset B^c\} \ .$$

Thus $\mathcal{H} \neq \sigma(\mathcal{C})$, therefore there is a $C \in \mathcal{C}$ such that $x$ or $y$ are elements of $C$, but not both. This finishes the proof. $\square$

The property of denumerable separation proves useful also in the study of $\sigma$-homomorphisms between measurable spaces.

**Proposition 2.** *Let $(S, \mathcal{B})$ be a mSdS, let $(\Omega, \mathcal{E})$ be a measurable space and let $h : \mathcal{B} \to \mathcal{E}$ be a $\sigma$-homomorphism. Then there is a unique measurable function $\eta : \Omega \to S$ such that $\eta^{-1}(B) = h(B)$ for all $B \in \mathcal{B}$.*

*Proof.* First we prove that $h$ is not only a $\sigma$-homomorphism but a $\tau$-homomorphism, that is

$$h(B) \;=\; \bigcup_{x \in B} h(\{x\}) \tag{2}$$

for all $B \in \mathcal{B}$. For the proof let $\{C_1, C_2, \ldots\}$ be a separating system of $\mathcal{B}$. Because $h$ is a $\sigma$-homomorphism,

$$h(B) \;=\; \bigcap_n h(B \cap C_n) \cup h(B \cap C_n^c) \;.$$

Since we consider sets here, this expression may be further transformed by general distributivity: Denoting $C_n^+ := C_n$ and $C_n^- := C_n^c$

$$h(B) \;=\; \bigcup_\chi \bigcap_n h(B \cap C_n^{\chi(n)}) \;=\; \bigcup_\chi h\Big(B \cap \bigcap_n C_n^{\chi(n)}\Big) \;,$$

where the union is taken over all mappings $\chi : \mathbb{N} \to \{+, -\}$. Since $\{C_1, C_2, \ldots\}$ is a separating system, $\bigcap_n C_n^{\chi(n)}$ contains at most one element, and for each $x \in S$ there is exactly one $\chi$ such that $\{x\} = \bigcap_n C_n^{\chi(n)}$. Therefore (2) follows.

In particular $\Omega = \bigcup_{x \in S} h(\{x\})$. This enables us to define $\eta$ by means of

$$\eta(\omega) = x \quad :\Leftrightarrow \quad \omega \in h(\{x\}) \;,$$

that is $\eta^{-1}(\{x\}) = h(\{x\})$. From (2)

$$h(B) \;=\; \bigcup_{x \in B} \eta^{-1}(\{x\}) \;=\; \eta^{-1}(B) \;.$$

In particular $\eta$ is measurable. □

# 4 The axioms for random variables

In this section we introduce the concept of a *stochastic ensemble* $\mathcal{R}_S$, $S \in \mathcal{T}$. We require the following properties:

$\mathcal{T}$ is a collection of elements, which we call *target spaces*. They are assumed to be measurable spaces $(S, \mathcal{B})$ with denumerable separation. Since in concrete cases it is always clear, which $\sigma$-field $\mathcal{B}$ is used within $S$, we often take the liberty to call $S$ the target space and to suppress $\mathcal{B}$.

$\mathcal{R}_S$ is a set for each $S \in \mathcal{T}$. Its elements are called *random variables*, more precisely *random variables with target space $S$*. For $X \in \mathcal{R}_S$ we also say "$X$ takes values in $S$" and write

$$X \curvearrowright S \;.$$

$X, Y \curvearrowright S$ means $X, Y \in \mathcal{R}_S$.

Four axioms are needed to make this concept work. The first two assure that stochastic ensembles are sufficiently rich. (Of course products of target spaces are always endowed with the product-$\sigma$-field.)

**Axiom 1.** $\{0,1\}^n \in \mathcal{T}$ for $n = 1, 2, \ldots, \infty$. Moreover, if $S_1, \ldots, S_n \in \mathcal{T}$, then also $S_1 \times \cdots \times S_n \in \mathcal{T}$.

This axiom contains the minimal assumptions needed for our purposes. In view of Axiom 3 below one might prefer to require that also countably infinite products of target spaces always belong to $\mathcal{T}$. This is a matter of taste – then $\mathcal{T}$ will be enlarged dramatically.

**Axiom 2.** *There are $S \in \mathcal{T}$ and $X, Y \curvearrowright S$ such that $X \neq Y$.*

The other two axioms describe, how to build new random variables from given ones. The next axiom states that random variables transform like ordinary variables.

**Axiom 3.** *To each random variable $X \curvearrowright S$ and to each measurable mapping $\varphi : S \to S'$ with $S' \in \mathcal{T}$ a random variable $X' \curvearrowright S'$ is uniquely associated denoted $X' = \varphi(X)$. These random variables fulfil*

$$\mathrm{id}(X) \;=\; X$$

*and*

$$(\psi \circ \varphi)(X) \;=\; \psi(\varphi(X)) \;,$$

*whenever such expressions may be formed.*

In the next axiom $\pi_i : S_1 \times S_2 \times \cdots \to S_i$ denotes the projection mapping to the $i$-th coordinate,

$$\pi_i(x_1, x_2, \ldots) \;:=\; x_i \;.$$

$\pi_i$ is measurable.

**Axiom 4.** *Let $S_1, S_2, \ldots$ be a finite or infinite sequence of target spaces such that also $S_1 \times S_2 \times \ldots$ belongs to $\mathcal{T}$. Then to any $X_1 \curvearrowright S_1, X_2 \curvearrowright S_2, \ldots$ a random variable $X \curvearrowright S_1 \times S_2 \times \cdots$ is uniquely associated characterized by the property*

$$\pi_i(X) \;=\; X_i$$

*for all $i$. It is denoted $X = (X_1, X_2, \ldots)$ and called the* product variable *of $X_1, X_2, \ldots$.*

Axiom 3 and 4 can be summarized as follows: If $X_1 \curvearrowright S_1, X_2 \curvearrowright S_2, \ldots$ and if $\varphi : S_1 \times S_2 \times \to S$ is measurable, then we may form the random variable

$$\varphi(X_1, X_2, \ldots) \ := \ \phi(X)$$

with $X = (X_1, X_2, \ldots)$, provided the product space belongs to $\mathcal{T}$. Also

$$\psi\big(\varphi_1(X_1, X_2, \ldots), \varphi_2(X_1, X_2, \ldots), \ldots\big) \ = \ \big(\psi \circ (\varphi_1, \varphi_2, \ldots)\big)(X_1, X_2, \ldots) \quad (3)$$

for suitable measurable mappings $\psi, \varphi_1, \varphi_2, \ldots$ Indeed: From Axiom 3 we obtain $\pi_i\big((\varphi_1, \varphi_2, \ldots)(X)\big) = \varphi_i(X)$, thus $(\varphi_1, \varphi_2, \ldots)(X) = (\varphi_1(X), \varphi_2(X), \ldots)$ from Axiom 4 and (3) follows from Axiom 3.

**Examples.**

1. **Real-valued random variables.** As to a concrete example let us look at real-valued random variables. Then $\mathcal{T}$ has to contain $\mathbb{R}^d$ for $d = 1, 2, \ldots$. The ordinary operations within $\mathbb{R}$ transfer to random variables without difficulties. For example, for $X_1, X_2 \curvearrowright \mathbb{R}$ we may define $X_1 + X_2 := \varphi(X_1, X_2)$ using the measurable mapping $\varphi(x_1, x_2) := x_1 + x_2$. The calculation rules transfer from the real numbers to random variables by means of (3), f.e.

$$\begin{aligned} (X_1 + X_2) + X_3 \ &= \ \big(\varphi \circ (\varphi \circ (\pi_1, \pi_2), \pi_3)\big)(X_1, X_2, X_3) \\ &= \ \big(\varphi \circ (\pi_1, \varphi \circ (\pi_2, \pi_3))\big)(X_1, X_2, X_3) \ = \ X_1 + (X_2 + X_3) \end{aligned}$$

   Other operations as $|X|$ and $\max(X_1, X_2)$ are introduced in much the same way.

2. **Random variables with constant value.** Each element $c$ of some target space $S$ may be considered as random variable: Let $\varphi' : S' \to S$ be any mapping taking only the value $c$ and choose any random variable $X' \curvearrowright S'$. $\varphi'$ is measurable. It is easy to show that $\varphi'(X')$ does only depent on $c$: If $\varphi''(X'')$ is another choice, then by Axiom 3 and 4 $\varphi'(X') = (\varphi' \circ \pi')(X', X'') = (\varphi'' \circ \pi'')(X', X'') = \varphi''(X'')$. It is consistent to denote this random variable by $c$ again. Indeed a measurable mapping $\varphi$ may now be applied to $c$ in two different meanings, but the result is the same because of $\varphi(\varphi'(X')) = (\varphi \circ \varphi')(X')$ and the observation that $\varphi \circ \varphi'$ takes the constant value $\phi(c)$.

3. **Measure theoretic models.** For any collection $\mathcal{T}$ of target spaces we get examples of stochastic ensembles by choosing some basic measurable space $(\Omega, \mathcal{E})$ and then letting $\mathcal{R}_S := \{X : \Omega \to S \,|\, X \text{ is measurable}\}$, $S \in \mathcal{T}$, $\varphi(X) := \varphi \circ X$ and $(X_1, X_2, \ldots)$ the product mapping. These are the canonical models. If a probability measure is given on the basic space, then also the sets $\mathcal{R}'_S := \{[X] \,|\, X \in \mathcal{R}_S\}$ make up a stochastic ensemble, where $[X]$ denotes the equivalence class of measurable functions, which are a.e. equal to $X$. Here $\varphi\big([X_1], [X_2], \ldots\big) := \big[\varphi(X_1, X_2, \ldots)\big]$,

$\big([X_1],[X_2],\dots\big) := \big[(X_1,X_2,\ldots)\big]$. The axioms are trivially fulfilled. Note that such stochastic ensembles do not consist of measurable functions in general. $\square$

(3) says in short that ordinary variables may be replaced by random variables in functional relations. The following proposition substantiates this statement.

**Proposition 3.** *Let $S, S', S_1, S_2$ be target spaces and let $\varphi_i : S_1 \times S_2 \to S$, $\psi_i : S_1 \times S_2 \to S'$, $i = 1,2$ be measurable mappings fulfilling*

$$\varphi_1(x,y) = \varphi_2(x,y) \quad \Rightarrow \quad \psi_1(x,y) = \psi_2(x,y)$$

*for all $x \in S_1, y \in S_2$. Let also $X \curvearrowright S_1$, $Y \curvearrowright S_2$ . Then*

$$\varphi_1(X,Y) = \varphi_2(X,Y) \quad \Rightarrow \quad \psi_1(X,Y) = \psi_2(X,Y) \ .$$

The proposition comprises the possibility that $\varphi_i$ and $\psi_i$ do not depend on $x$ or $y$. Note that f.e. $\varphi(x,y) = \varphi'(x)$ for all $x,y$ implies $\varphi(X,Y) = (\varphi' \circ \pi)(X,Y) = \varphi'(X)$ by Axiom 3 and 4.

*Proof.* Let $z = (x,y)$ and $Z = (X,Y)$. Consider $\theta : S \times S \times S' \times S' \to S'$, given by

$$\theta(u,v,u',v') \ := \ \begin{cases} u' \ , & \text{if } u = v \ , \\ v' \ , & \text{if } u \neq v \ . \end{cases}$$

$\theta$ is measurable, due to the fact that the diagonal in $S \times S$ is measurable. Then

$$\theta\big(\varphi_1(z),\varphi_1(z),\psi_1(z),\psi_2(z)\big) \ = \ \psi_1(z) \ ,$$

whereas by assumption

$$\theta\big(\varphi_1(z),\varphi_2(z),\psi_1(z),\psi_2(z)\big) \ = \ \psi_2(z) \ .$$

By means of (3) replace the variable $z$ by the random variable $Z$ in these equations. Then by assumption the lefthand sides coincide, and our claim follows. $\square$

# 5 Events

To each random variable $X \curvearrowright S$ and to each measurable subset $B \subset S$ we associate now an *event*, written as

$$\{X \in B\} \ .$$

In particular, since target spaces contain one point sets, we may form the events

$$\{X = x\} \ := \ \{X \in \{x\}\} \ , \quad x \in S \ .$$

In order to carry out calculations, we have to define equality of events. Here we use that the characteristic function $\mathbf{1}_B(\cdot)$ of measurable $B \subset S$ is a measurable mapping from $S$ to $\{0,1\}$, which allows to apply Axiom 1 and 3.

**Definition.** *Two events* $\{X \in B\}$ *and* $\{X' \in B'\}$ *are said to be equal,*

$$\{X \in B\} \;=\; \{X' \in B'\} \;,$$

if $\mathbf{1}_B(X) = \mathbf{1}_{B'}(X')$.

In other words: In our approach an event is an equivalence class of pairs $(X, B)$. To each event $E$ we may associate its *indicator variable* $I_E$, a random variable with values in $\{0, 1\}$, given by

$$I_E := \mathbf{1}_B(X) \;, \quad \text{if } E = \{X \in B\} \;.$$

Two events with the same indicator variable are equal. The set of events is denoted by $\mathcal{E}$.

**Examples.**

1. The equality
$$\{\varphi(X) \in B\} \;=\; \{X \in \varphi^{-1}(B)\}$$
holds, since $\mathbf{1}_B(\varphi(X)) = \mathbf{1}_B \circ \varphi(X) = \mathbf{1}_{\varphi^{-1}(B)}(X)$ in view of Axiom 3.

2. For any event $E$ the equality
$$\{I_E = 1\} \;=\; E$$
holds, since $\mathbf{1}_{\{1\}} = \mathrm{id}$ on $\{0, 1\}$, thus $\mathbf{1}_{\{1\}}(I_E) = \mathrm{id}(I_E) = I_E$. $\square$

Next we introduce the basic operations with events.

**Proposition 4.** *For any event* $E$ *there exists an unique event, denoted* $E^c$*, such that*

$$\{X \in B\}^c \;=\; \{X \in B^c\}$$

*for any* $X \curvearrowright S$ *and any measurable* $B \subset S$*.*

*Proof.* For $E = \{X \in B\}$ we define $E^c := \{X \in B^c\}$. We only have to show that this definition is unambiguous. Note that $\mathbf{1}_{B^c} = \eta \circ \mathbf{1}_B$ with $\eta(0) = 1, \eta(1) = 0$. Thus $\{X \in B\} = \{X' \in B'\}$ and Axiom 3 imply $\mathbf{1}_{B^c}(X) = \eta(\mathbf{1}_B(X)) = \eta(\mathbf{1}_{B'}(X')) = \mathbf{1}_{(B')^c}(X')$. Therefore $\{X \in B^c\} = \{X' \in (B')^c\}$. $\square$

**Proposition 5.** *For any finite or infinite sequence of events* $E_1, E_2, \ldots$ *there exist two unique events, denoted as* $\bigcup_n E_n$ *and* $\bigcap_n E_n$*, such that*

$$\bigcup_n \{X \in B_n\} \;=\; \Big\{X \in \bigcup_n B_n\Big\} \;,\; \bigcap_n \{X \in B_n\} \;=\; \Big\{X \in \bigcap_n B_n\Big\}$$

*for any* $X \curvearrowright S$ *and any measurable* $B_1, B_2, \ldots \subset S$*.*

*Proof.* We proceed as in the last proof and define $\bigcup_n E_n := \{X \in \bigcup_n B_n\}$, if $E_n = \{X \in B_n\}$. Note that $\mathbf{1}_{\bigcup_n B_n} = \max \circ (\mathbf{1}_{B_1}, \mathbf{1}_{B_2}, \ldots)$ with the measurable function $\max(x_1, x_2, \ldots) := \max_n x_n$ from $\{0,1\}^\ell$ to $\{0,1\}$ ($\ell$ being the length of the sequence). Thus $\{X \in B_n\} = \{X' \in B'_n\}$ implies $\mathbf{1}_{\bigcup_n B_n}(X) = \max(\mathbf{1}_{B_1}(X), \mathbf{1}_{B_2}(X), \ldots) = \max(\mathbf{1}_{B'_1}(X'), \mathbf{1}_{B'_2}(X'), \ldots) = \mathbf{1}_{\bigcup_n B'_n}(X')$ in view of Axiom 3 and 4. Therefore $\{X \in \bigcup_n B_n\} = \{X' \in \bigcup_n B'_n\}$ such that $\bigcup_n E_n$ is well-defined. $\bigcap_n E_n$ is obtained similarly.

It remains to show that each sequence $E_1, E_2, \ldots$ can be represented as $E_n = \{X \in B_n\}$ with *one* random variable $X$. Letting $X := (I_{E_1}, I_{E_2}, \ldots)$ with target space $S = \{0,1\}^\ell$ and $B_n := \{0,1\}^{n-1} \times \{1\} \times \{0,1\}^{\ell - n}$ we indeed obtain

$$\{X \in B_n\} \;=\; \{X \in \pi_n^{-1}(\{1\})\} \;=\; \{\pi_n(X) = 1\} \;=\; \{I_{E_n} = 1\} \;=\; E_n \;. \quad (4)$$

This is the claim. □

**Proposition 6.** *There are two unique events $E_{\text{sure}} \neq E_{\text{imp}}$ such that*

$$E_{\text{sure}} = \{X \in S\} \;,\; E_{\text{imp}} = \{X \in \emptyset\}$$

*for any $X \curvearrowright S$.*

*Proof.* Again define $E_{\text{sure}} := \{X \in S\}$. Let also $X' \curvearrowright S'$. Then $\mathbf{1}_S \circ \pi = \mathbf{1}_{S'} \circ \pi'$ on $S \times S'$ with projections $\pi, \pi'$. Axiom 3 implies $\mathbf{1}_S(X) = \mathbf{1}_S \circ \pi(X, X') = \mathbf{1}_{S'} \circ \pi'(X, X') = \mathbf{1}_{S'}(X')$, thus $\{X \in S\} = \{X' \in S'\}$. Similarly $\{X \in \emptyset\} = \{X' \in \emptyset\}$, therefore $E_{\text{sure}}$ and $E_{\text{imp}}$ are well-defined.

Now suppose that $E_{\text{sure}} = E_{\text{imp}}$. Then all events are equal, as follows from Proposition 5:

$$\begin{aligned} E &= \{X \in B\} = \{X \in B\} \cap \{X \in S\} \\ &= \{X \in B\} \cap \{X \in \emptyset\} = \{X \in \emptyset\} = E_{\text{imp}} \;. \end{aligned}$$

This implies that any two random variables $X, Y$ with the same target space $S$ are equal. Indeed, let $C_1, C_2, \ldots$ be a separating system in $S$ and let $\varphi := (\mathbf{1}_{C_1}, \mathbf{1}_{C_2}, \ldots)$. Then $\varphi(X) = (I_{\{X \in C_1\}}, I_{\{X \in C_2\}}, \ldots) = (I_{\{Y \in C_1\}}, I_{\{Y \in C_2\}}, \ldots) = \varphi(Y)$. $\varphi$ is injective, thus we may conclude from Proposition 3 that $X = Y$. This contradicts Axiom 2, and our claim follows. □

Thus we have introduced complementary events, unions and intersections of events as well as the sure and the impossible event. In view of the above characterisations and of (4) it is obvious that properties of sets carry over to properties of events. Altogether we end up with a Boolean $\sigma$-lattice, equipped with the order relation

$$E \subset E' \quad :\Leftrightarrow \quad E = E \cap E' \;,$$

with maximal element $E_{\text{sure}}$ and minimal element $E_{\text{imp}}$. It is a standard procedure to obtain the other properties of fields of events.

**Examples.**

1. For $X = (X_1, X_2, \ldots)$ we have from Proposition 5

$$\{X \in B_1 \times B_2 \times \cdots\} \;=\; \Big\{X \in \bigcap_n \pi_n^{-1}(B_n)\Big\} \;=\; \bigcap_n \{X \in \pi_n^{-1}(B_n)\}$$

and from Axiom 4

$$\{(X_1, X_2, \ldots) \in B_1 \times B_2 \times \cdots\} \;=\; \bigcap_n \{X_n \in B_n\} \;.$$

2. Let $\sim$ be any relation on $S$ such that $R := \{(x,y) \in S^2 : x \sim y\}$ is a measurable subset of $S^2$. Then it is natural to define $\{X \sim Y\} := \{(X,Y) \in R\}$ and to write $X \sim Y$, if $\{X \sim Y\} = E_{\text{sure}}$. This is in accordance with Proposition 7 below. □

**Remark: $\sigma$-homomorphisms in stochastic ensembles.** From Proposition 4 to 6 it is immediate that each random variable $X$ induces a $\sigma$-homomorphism from $\mathcal{B}$ to $\mathcal{E}$ given by $B \mapsto h(B) := \{X \in B\}$. From Proposition 7 below it follows that $h$ determines $X$ uniquely. Therefore one may ask, whether in a stochastic ensemble every $\sigma$-homomorphism $h$ from the $\sigma$-field $\mathcal{B}$ of some target space $S$ into $\mathcal{E}$ comes from a random variable. This is true in two cases.

The first case is the classical one that the random variables with target space $S$ are given as above by the system of measurable mappings from some basic measurable space $(\Omega, \mathcal{E})$ into $S$. Then Proposition 2 applies saying that there are no other $\sigma$-homomorphisms.

In the other case we assume that $S$ is a Polish space endowed with its Borel-$\sigma$-field. This case is more profound: Choose a separating sequence $C_1, C_2, \ldots$ Define the measurable function $\varphi := (\mathbf{1}_{C_1}, \mathbf{1}_{C_2}, \ldots)$ from $S$ into $\{0,1\}^\infty$ and the random variable $Y := (I_{h(C_1)}, I_{h(C_2)}, \ldots)$. Then

$$\{Y \in B'\} \;=\; h\big(\varphi^{-1}(B')\big)$$

for each Borel-set $B' \subset \{0,1\}^\infty$. For the proof note that the system of $B'$ fulfilling our claim is a $\sigma$-field containing $\{0,1\}^{m-1} \times \{1\} \times \{0,1\}^\infty$. Now $\varphi$ is an injective measurable mapping from the Polish space $S$ into the Polish space $\{0,1\}^\infty$. A celebrated theorem of Kuratowski says that then the image of each Borel-set is a Borel-set again. An immediate consequence is that there is a mapping $\psi : \{0,1\}^\infty \to S$ such that $\psi \circ \varphi$ is the identity on $S$. Letting $X := \psi(Y)$ one obtains the claim: For each Borel-set $B \subset S$

$$\{X \in B\} \;=\; \{Y \in \psi^{-1}(B)\} \;=\; h(B) \;. \qquad \square$$

# 6 Equality and a.s. equality

Recall from (1) that the diagonal $D \subset S^2$ is measurable in the case of target spaces. Thus we may define for $X, Y \curvearrowright S$

$$\{X = Y\} \;:=\; \{(X,Y) \in D\} \;,\; \{X \neq Y\} \;:=\; \{(X,Y) \in D^c\} \;.$$

**Proposition 7.** *For $X, Y \curvearrowright S$ the following statements are equivalent:*

*i)* $X = Y$,

*ii)* $\{X = Y\} = E_{\text{sure}}$.

*iii)* $\{X \in B\} = \{Y \in B\}$ *for all measurable* $B \subset S$.

*Proof.* Let $C_1, C_2, \ldots$ separate the elements of $S$. Then $\varphi(x) := (\mathbf{1}_{C_1}, \mathbf{1}_{C_2}, \ldots)$ is a injectiv measurable function from $S$ to $\{0,1\}^\infty$. Thus for $x, y \in S$

$$x = y \quad \Leftrightarrow \quad \mathbf{1}_D(x,y) = \mathbf{1}_{S\times S}(x,y) \quad \Leftrightarrow \quad \varphi(x) = \varphi(y) \ .$$

Proposition 3 implies

$$X = Y \quad \Leftrightarrow \quad \mathbf{1}_D(X,Y) = \mathbf{1}_{S\times S}(X,Y) \quad \Leftrightarrow \quad \varphi(X) = \varphi(Y) \ .$$

Thus $X = Y$ is equivalent to $\{X = Y\} = \{(X,Y) \in S \times S\} = E_{\text{sure}}$ as well as to $\{X \in C_i\} = \{Y \in C_i\}$ for all $i$. Since any measurable $B \subset S$ may be included into the sequence $C_1, C_2, \ldots$, our claim follows. □

Next we discuss the notion of almost sure equality in stochastic ensembles. Shortly speaking this is any equivalence relation compatible with our axioms. As we shall see this conforms to the traditional definition of a.s. equality. On the other hand it will become apparent that in our setting it is no longer necessary to distinguish between random variables and a.s. defined random variables as in the traditional approach. Both give rise to stochastic ensembles.

**Definition.** *An equivalence relation $\sim$ on the collection of random variables of a stochastic ensemble is called an* a.s. equality*, if*

*i)* $X \sim Y$ *implies that* $X, Y$ *have the same target space.*

*ii) There exist* $X, Y \curvearrowright S$ *such that* $X \not\sim Y$.

*iii)* $X \sim Y \quad \Rightarrow \quad \varphi(X) \sim \varphi(Y)$.

*iv)* $X_1 \sim Y_1, X_2 \sim Y_2, \ldots \quad \Rightarrow \quad (X_1, X_2, \ldots) \sim (Y_1, Y_2, \ldots)$.

Let $X^\sim$ denote the equivalence class of the a.s. equality $\sim$ containing $X$. Then we may associate to $X^\sim$ a target space, namely that of $X$. Also we may define

$$\varphi(X^\sim) \ := \ \varphi(X)^\sim \ , \quad \big(X_1^\sim, X_2^\sim, \ldots\big) \ := \ (X_1, X_2, \ldots)^\sim \ .$$

Let

$$\mathcal{R}_S^\sim \ := \ \big\{X^\sim : X \in \mathcal{R}_S\big\}$$

With these conventions the following result is obvious.

**Proposition 8.** *$\mathcal{R}_S^\sim$, $S \in \mathcal{T}$ is a stochastic ensemble.*

In the sequel we show that our definition is intimately connected with the usual definition of a.s. equality. Let us recall the notion of a *null-system* (a *$\sigma$-ideal*). It is a system $\mathcal{N} \subset \mathcal{E}$ of events fulfilling

$$\begin{array}{rcl} E_1, E_2, \ldots \in \mathcal{N} & \Rightarrow & \bigcup_n E_n \in \mathcal{N} \\ E \in \mathcal{N}\ , E' \subset E & \Rightarrow & E' \in \mathcal{N} \\ E_{\text{imp}} \in \mathcal{N} & , & E_{\text{sure}} \notin \mathcal{N} \end{array}$$

An important example of a null-system is the system of events of probability zero in case $\mathcal{E}$ is endowed with a probability measure.

**Proposition 9.** *To each a.s. equality $\sim$ the system of events*

$$\mathcal{N}^\sim := \big\{\{X \neq Y\} : X \sim Y\big\}$$

*is a null-system. It fulfils*

$$X \sim Y \quad \Leftrightarrow \quad \{X \neq Y\} \in \mathcal{N}^\sim$$

*for any $X, Y \curvearrowright S$. The mapping $\sim \mapsto \mathcal{N}^\sim$ establishes a one-to-one correspondence between a.s. equality relations and null-systems.*

*Proof.* From

$$\bigcup_n \{X_n \neq Y_n\} \ = \ \{(X_1, X_2, \ldots) \neq (Y_1, Y_2, \ldots)\}$$

and from condition iv) of the definition it follows that $\mathcal{N}^\sim$ fulfils the first requirement of a null-system. Next let $X, Y \curvearrowright S$ and

$$E \ \subset \ \{X \neq Y\}\ .$$

We have to show $E \in \mathcal{N}^\sim$. Suppose $E = \{Z \in B'\}$ for a suitable $Z \curvearrowright S'$. Define the measurable function $\varphi : S \times S' \to S$ as

$$\varphi(x, z) \ := \ \begin{cases} x\ , & \text{if } z \in B' \\ x_0\ , & \text{if } z \notin B' \end{cases}$$

for some $x_0 \in S$. It follows

$$\{\varphi(X, Z) \neq \varphi(Y, Z)\} \ = \ \{Z \in B'\} \cap \{X \neq Y\} \ = \ E\ .$$

From $X \sim Y$ and condition iii) and iv) of the definition we obtain $\varphi(X, Z) \sim \varphi(Y, Z)$. Thus $E \in \mathcal{N}^\sim$, which means that $\mathcal{N}^\sim$ fulfils the second requirement of a null-system. In particular: If $E_{\text{sure}} \in \mathcal{N}^\sim$, then $\{X \neq Y\} \in \mathcal{N}^\sim$ for all

$X, Y \curvearrowright S$. This contradicts condition ii) of the definition, thus we may conclude $E_{\text{sure}} \notin \mathcal{N}^\sim$. Finally $E_{\text{imp}} = \{X \neq X\} \in \mathcal{N}^\sim$. Therefore $\mathcal{N}^\sim$ is a null-system.

We come to the second claim of the proposition. The implication $\Rightarrow$ is obvious, thus let us assume $\{X \neq Y\} \in \mathcal{N}^\sim$. Then there exist $X', Y' \curvearrowright S$ such that $\{X \neq Y\} = \{X' \neq Y'\}$ and $X' \sim Y'$. We have to show that $X \sim Y$, too. For this purpose we use the measurable mapping $\theta$, defined in the proof of Proposition 3. It fulfils

$$\mathbf{1}_D(x, y) = \mathbf{1}_{D'}(x', y') \quad \Rightarrow \quad \theta(x', y', x, y) = y \ ,$$

where $D$ and $D'$ are the diagonals in $S^2$ and $(S')^2$. By assumption $\mathbf{1}_D(X, Y) = \mathbf{1}_{D'}(X', Y')$, therefore Proposition 3 implies $\theta(X', Y', X, Y) = Y$. Moreover $\theta(x', x', x, y) = x$ and consequently $\theta(X', X', X, Y) = X$. Since $X' \sim Y'$, we obtain $X \sim Y$ in view of iii) and iv) of the definition . Thus also the second statement is proved.

In particular this implies that $\sim \mapsto \mathcal{N}^\sim$ is an injective mapping. It remains to prove surjectivity. Thus let $\mathcal{N}$ be any null-system and define

$$X \sim Y \quad :\Leftrightarrow \quad \{X \neq Y\} \in \mathcal{N} \ ,$$

whenever $X$ and $Y$ have the same target space. We have to show that $\sim$ is an a.s. equality. Since $\{I_{E_{\text{sure}}} \neq I_{E_{\text{imp}}}\} = E_{\text{sure}}$, $I_{E_{\text{sure}}} \not\sim I_{E_{\text{imp}}}$. Thus condition ii) of the definition holds. Condition iii) follows from

$$\{\varphi(X) \neq \varphi(Y)\} \ \subset \ \{X \neq Y\}$$

and condition iv) follows from

$$\{(X_1, X_2, \ldots) \neq (Y_1, Y_2, \ldots)\} \ = \ \bigcup_n \{X_n \neq Y_n\} \ .$$

This finishes the proof. $\Box$

Random variables $X \curvearrowright S$, $X' \curvearrowright S'$ with different target spaces are always unequal in our approach. It might be convenient to call them *indistinguishable*, if $S \cap S'$ is a measurable subset of $S$ and of $S'$ and if

$$\begin{aligned} \{X \in B\} \ &= \ \{X' \in B\} \quad \text{for all measurable } B \subset S \cap S' \ , \\ \{X \in S \cap S'\} \ &= \ \{X' \in S \cap S'\} \ = \ E_{\text{sure}} \ . \end{aligned}$$

In the same manner *a.s. indistinguishability* may be introduced.

# 7 Convergence of random variables

'In practice' the small omegas prove convenient in operating with random variables within the traditional setting. In our approach such manipulations may be reproduced without difficulties within the target spaces. This has been indicated already, here we like to exemplify this briefly in the context of convergence of random variables.

Let $d$ be a metric on the target space $S$. We require

- The mapping $d : S^2 \to \mathbb{R}_+$ is measurable.
- Let $\varphi_1, \varphi_2, \ldots$ be mappings from some target space $S'$ to $S$ converging pointwise to $\varphi : S' \to S$ with respect to $d$. If the $\varphi_n$ are measurable, then also $\varphi$ is measurable.

As is well-known these assumptions are satisfied, if $(S, d)$ is a separable metric space and $\mathcal{B}$ the corresponding Borel-$\sigma$-field on $S$. An example of a non-separable metric space of importance fulfilling both requirements is the space of càdlàg functions endowed with the metric of locally uniform convergence (this metric is used in the theory of stochastic integration).

Now we assume that besides $S$ also $S^\infty$ is a target space. Let us consider the set of convergent sequences with given limit and the set of Cauchy-sequences,

$$\begin{aligned} B_{\text{lim}} &:= \{(x, x_1, x_2, \ldots) \in S \times S^\infty \ : \ x = \lim_n x_n\} \ , \\ B_{\text{Cauchy}} &:= \{(x_1, x_2, \ldots) \in S^\infty \ : \ (x_n) \text{ is Cauchy}\} \ . \end{aligned}$$

Since $B_{\text{lim}} = \bigcap_{k=1}^\infty \bigcup_{m=1}^\infty \bigcap_{n=m}^\infty \{(x, x_1, x_2, \ldots) : d(x_n, x) \le \epsilon_k\}$ and $B_{\text{Cauchy}} = \bigcap_{k=1}^\infty \bigcup_{m=1}^\infty \bigcap_{n=m}^\infty \{(x_1, x_2, \ldots) : d(x_m, x_n) \le \epsilon_k\}$ for any sequence $\epsilon_k \downarrow 0$, these are measurable subsets of $S \times S^\infty$ and $S^\infty$. Thus for any random variables $X, X_1, X_2, \ldots \curvearrowright S$ we may define the events

$$\begin{aligned} \{X_n \to X\} &:= \{(X, X_1, X_2, \ldots) \in B_{\text{lim}}\} \ , \\ \{X_n \text{ is Cauchy}\} &:= \{(X_1, X_2, \ldots) \in B_{\text{Cauchy}}\} \ . \end{aligned}$$

**Proposition 10.** *For any $X, X_1, X_2, \ldots \curvearrowright S$*

$$\{X_n \to X\} \ \subset \ \{X_n \textit{ is Cauchy}\} \ .$$

*Moreover, if $(S, d)$ is a complete metric space, then for any $X_1, X_2, \ldots \curvearrowright S$ there is a $X \curvearrowright S$ such that*

$$\{X_n \to X\} \ = \ \{X_n \text{ is Cauchy}\} \ .$$

*Proof.* Since $B_{\text{lim}} \subset S \times B_{\text{Cauchy}}$

$$\{X_n \to X\} \ \subset \ \{(X, X_1, X_2, \ldots) \in S \times B_{\text{Cauchy}}\} \ = \ \{X_n \text{ is Cauchy}\} \ .$$

Moreover in the case of a complete metric space let $\varphi_n, \varphi : S^\infty \to S$ be given by

$$\begin{aligned} \varphi_n(x_1, x_2, \ldots) &:= \begin{cases} x_n \ , & \text{if } (x_1, x_2, \ldots) \in B_{\text{Cauchy}} \ , \\ z, & \text{otherwise} \ , \end{cases} \\ \varphi(x_1, x_2, \ldots) &:= \begin{cases} \lim_n x_n, & \text{if } (x_1, x_2, \ldots) \in B_{\text{Cauchy}} \ , \\ z, & \text{otherwise} \end{cases} \end{aligned}$$

with some given $z \in S$. $\varphi_n$ ist measurable and $\varphi(x_1, x_2, \ldots) = \lim_n \varphi_n(x_1, x_2, \ldots)$ for all $(x_1, x_2, \ldots)$, thus $\varphi$ is measurable too. Define

$X := \varphi(X_1, X_2, \ldots)$. Because of completeness $\psi^{-1}(B_{\text{lim}}) = B_{\text{Cauchy}}$ for the measurable mapping $\psi : S^\infty \to S \times S^\infty$, given by $\psi(x_1, x_2, \ldots) := (\varphi(x_1, x_2, \ldots), x_1, x_2, \ldots)$. Thus

$$\begin{aligned} \{X_n \to X\} &= \{\psi(X_1, X_2, \ldots) \in B_{\text{lim}}\} \\ &= \{(X_1, X_2, \ldots) \in B_{\text{Cauchy}}\} = \{X_n \text{ is Cauchy}\} , \end{aligned}$$

which is the claim. □

# References


[1] Billingsley, P. (1979). *Probability and Measure*, New York.

[2] Breiman, L. (1968). *Probability*, New York.

[3] Caratheodory, C. (1956). *Mass und Integral und ihre Algebraisierung*, Basel.

[4] Coxeter, H.S.M. (1961). *Introduction to Geometry*, New York.

[5] Dinges, H. and Rost, H. (1982). *Prinzipien der Stochastik*, Stuttgart.

[6] Dinges, H. (2005). Variables, in particular random variables. *Activity and Sign, Grounding Mathematics Education*, M. Hoffmann, J. Lenhard, F. Seeger (Eds.), 305–311.

[7] Halmos, P.R. (1960). *Naive Set Theory*, Princeton.

[8] Mumford, D. (2000). The Dawning of the Age of Stochasticity. *Mathematics: Frontiers and Perspectives*, 197–218.

[9] Snell, L. (1997). A Conversation with Joe Doob. *Stat. Science* **12**, 301–311. http://www.dartmouth.edu/∼chance/Doob/conversation.html