A note on probability and Hilbert’s VI problem.

A. Gandolfi
email: ag189@nyu.edu - gandolfi@math.unifi.it

Abstract

We start from various shortcomings, errors and dilemmas which arise in applications of probability. We argue that, together with similar issues in mechanics, they constitute the core of Hilbert’s VI problem. In the paper we indicate a solution to these problems. This leads to a reformulation of the axioms of probability and, in the finite case, to algebrization of probability theory.

The overall outcome is that while so far in most mathematical theories the focus is on unburdening general existence results, a shift is needed towards abstract discriminating existential conditions.

Key words and phrases: discrete probabilities, conditional probabilities, independence, conditional independence, existence theorems, Boole, De Finetti, Kolmogorov, axioms.
AMS subject classification: 60C05, 60K35.

1 Introduction

In axiomatic probability theory the axioms fix the rules about probability spaces and random variables. Consequently, probability spaces and random variables (and/or their distributions) are determined according to the basic conditions; concepts such as independence, or expressions like conditional probabilities, moments, correlations etc. are then defined within the theory.

*NYU Abu Dhabi and Dipartimento di Matematica e Informatica U. Dini, Università di Firenze, Firenze
The development of the theory consists chiefly of determining probabilities of other events or of computing the value of some of the above mentioned expressions. In mathematical terms, this is a direct problem, and in the recent decades it has lead to superb achievements.

On the other hand, this is not the way probability theory is used in applications or in exercises inspired by practical problems: often, it is not the probability space or the random variables or the distributions which are given, but rather some assumptions about independence or about the value of certain probabilities, conditional probabilities, moments etc. These assumptions are taken as starting point, and calculations of other probabilities or expectations are then carried out from there. Regardless of the possible motivations behind the assumptions, such as symmetry, physical or biological models, statistical data from past experiments or personal valuations, we want to focus here on the fact that the procedure of starting from specific assumptions does not fall within the framework of the classical axiomatic theory. One example for all: if independence is defined in the theory as the factorization of some probability, within that theory it can only be verified by checking such factorization (even indirectly); any a priori assumption of independence does not fall within that theory.

The practice of assuming some given features, in fact, has the disadvantage that nothing insures that they might be contradictory, so as to make our deductions meaningless. To give an example, we illustrate in Appendix A a very simple exercise, taken from an excellent and widely used textbook: given the hypothesis, many calculations are possible and apparently interesting, but an extra one shows that the set of assumptions is actually inconsistent.

In concrete situations, the iterated use of well tested assumptions, some safe procedures such as adding additional independent events or random variables to existing models, the ability in capturing real phenomena while making a model, or even some luck, often prevent the use of inconsistent sets of assumptions, but this does not relieve from the need of a mathematical foundation of probability.

Consistency is equivalent to satisfiability, i.e. the existence of at least one probability space in which the assumptions are actually true features or calculated values. This would lead to the inverse problem of determining probability space, events and random variables given the assumptions. But if one was to solve such problem by an actual explicit construction, then the proof of consistency would have been achieved at the price of losing the
calculative power of the assumptions. To continue the first example above, one often wants to assume independence and then use such assumption to compute \( P(A \cap B) \) in terms of \( P(A)P(B) \); but if the independence of \( A \) and \( B \) needs to be checked in advance, then we need a calculation of \( P(A \cap B) \) not in terms of \( P(A)P(B) \); after this calculation is done, the calculative power of independence is lost. Independence, conditional probabilities and concepts alike would become interesting verifications of intuition, but lose most of their calculative power.

This issue has been raised, in various forms and with diverse partial answers, at least by Boole, Hilbert and De Finetti. Some historical remarks on such contributions can be found in Appendix C.

In this paper I propose a solution to this dilemma. What is needed in order to insure consistency while retaining the calculative power of the assumptions is an abstract verification of consistency. The word "abstract" is the key feature: the more abstract the verification is, the better it is, as it is likely to be "simple" and to leave the entire calculative power to the assumptions. A direct construction of a suitable model would be a much less effective overshooting. As simple as it may seem, the observation that existence is sufficient is capable of solving a number of long standing dilemmas. Basically, it paves the way to solving Hilbert VI problem. Purely existential results for elementary problems, including assumptions about independence of events, are illustrated on Section 3, where it is shown that for a large class of finite probability problems existence is decidable.

It should be noticed that we came to the need of an abstract verification of existence stimulated by the apparently opposite need to apply the theory to concrete problems. In this respect, it is not unconceivable that one might require that the calculations based on the model be the most constructive possible; yet, the justification of the model has some advantage in being the most abstract possible. In other words, in order to justify possibly very constructive forward calculations we are lead to the search of very existential backward results. In a sense, this settles the conflict between constructive and non constructive mathematics: the two interpretations apply to different stages of mathematical modelling (see Section 2.7) below.

Clearly, there is plenty of existential results in probability theory; many models, such as Poisson distribution or Brownian motion just to quote two examples, can be or have been defined axiomatically, with their existence
verified afterwards. But what is missing is a systematic approach, applicable even to specific detailed assumptions, and aimed to abstract consistency proofs.

For such an approach it is, in fact, not enough to introduce assumptions and then produce a model in which the assumptions can be verified, as we often do not want to "produce" the model and we need to allow ourselves the possibility verifying consistency without a model at hand. We achieve this by formalizing assumptions, which we then call "requirements"; it is then natural to observe that also the properties of probability can be treated as "requirements", let’s call them foundational; and that all of this can be done on abstract variables which take sets or functions as their values. At this level of abstraction, all requirements together, the foundational and the additional ones, can be requested, and sometimes proven, to be consistent for the variables without mentioning any specific model.

This is analogous to the passage from real numbers to real variables, for which equations can be written and sometimes proven solvable or unsolvable without reference to any specific real number. This analogy takes a very concrete form for problems with a finite number of events, which can be algebrized to systems of equations (see Section 4).

The situation is also analogous to that of varieties, with concrete models taking the place of atlases of charts, and the collection of requirements being like the intrinsic properties of a variety. This analogy has been brought up by Tao ([T]) in relation to the more limited operation of enlarging a probability space (which is like enlarging the variety): clearly, any two concrete models satisfying given requirements have isomorphic portions (regardless of whether one is larger than the other or not), so our definition includes the "dogma" proposed by Tao (see [T] and Section 2.2 below).

2 Probability

2.1 Requirements

A boolean function is a function taking only values 0, which can be interpreted as false, and 1, interpreted as true.

**Definition 2.1.** A requirement on a class is a boolean function acting on each element of the class.

An element \( \mathfrak{c} \in \mathfrak{C} \) in the class satisfies the requirement \( r \) if \( r(\mathfrak{c}) \) holds.
Definition 2.2. For a given class \( C \), a \( C \)-variable \( c \) is a variable which takes value in the elements of the class. Each element \( \tau \) of the class is a realization of the variable.

We are interested in set variables, set function variables and function variables, which take sets, set function and functions as values, respectively. Real variables are the usual variables over \( \mathbb{R} \).

Definition 2.3. Given classes \( C_i, i \in I \), and a relation \( R \) on a subset of \( \prod_{i \in I} C_i \), the statement that \( C_i \)-variables \( c_i, i \in I \), satisfy the relation \( R \) is defined as the requirement that for each simultaneous realizations \( \tau_i, i \in I \), of the variables \( c_i \) belonging to the subset, \( R(\prod_{i \in I} \tau_i) = 1 \).

By means of suitable requirements all the elementary operations on sets, families of sets, set functions and functions can be lifted to the corresponding class variables. Here are some examples of how to interpret operations for class variables using requirements; this convention is adopted below for every relation or function on class variables.

Example 1. Given two set variables \( A \) and \( \Omega \), the assertion that \( A \subseteq \Omega \) is the requirement that in each realization \( \overline{A} \) and \( \overline{\Omega} \) of \( A \) and \( \Omega \), respectively, \( \overline{A} \subseteq \overline{\Omega} \).

The assertion that two set variables are equal is the requirement that for each realization the set realizing the first set variable is also a realization of the second, and vice versa.

The assertion that a set variable \( A \cap B \) (or \( A \cup B \), respectively) is the intersection (or the union) of two set variables \( A \) and \( B \) is the requirement that in any realization \( \overline{A}, \overline{B} \) and \( \overline{A \cap B} \) of \( A, B \) and \( A \cap B \) (or \( \overline{A \cup B} \) of \( A \cup B \), respectively) one has \( \overline{A \cap B} = \overline{A} \cap \overline{B} \) (or \( \overline{A \cup B} = \overline{A} \cup \overline{B} \), respectively). Notice that this implies that the sets realizing the set variables must all be subset of a common set.

Given a family \( \mathcal{A} \) of set variables, the assertion that a set function variable \( P \) is defined on \( \mathcal{A} \) is the requirement that for each realization \( \overline{P}, \overline{\mathcal{A}} \) of \( P \) and \( \mathcal{A} \), respectively, and every \( \overline{A} \in \overline{\mathcal{A}} \), \( \overline{P}(\overline{A}) \) is a real number.

For brevity, we say that the requirements lifting relations or operations to the class variables are \textbf{operational requirements}. Adopting operational requirements we no longer need to refer to realizations of class variables.
Example 2. (a) The assertion that given a set variable $\Omega$, a collection $\mathcal{A}$ of set variables, such that $A \subseteq \Omega$ for all $A \in \mathcal{A}$, is a $\sigma$-algebra is the requirement that $\Omega \in \mathcal{A}$, for each $A \in \mathcal{A}$ also $A^c \in \mathcal{A}$ and, finally, for each countable family $A_i$ of set variables in $\mathcal{A}$ one has $\bigcup_i A_i \in \mathcal{A}$.

(b) Given a set variable $\Omega$ and a $\sigma$-algebra of set variables $\mathcal{A}$ in $\Omega$, the assertion that a set function variable $P$ is nonnegative is the requirement that $P(A) \geq 0$ for all $A \in \mathcal{A}$; the assertion that $P$ is normalized is the requirement that $P(\Omega) = 1$; the assertion that $P$ is countably additive is the requirement that for every countable collection of disjoint set variables $A_i \in \mathcal{A}, i = 1, \ldots$, $P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

(c) Given a set variable $A$ and a set function variable $P$, the assertion that the value of $P$ on $A$ is less than $1/2$ is the requirement that $P(A) < 1/2$.

Requirements in the last example are used to lift standard statements about probabilities to class variables, and we refer to these and all others acting in the same way as relational requirements.

2.2 Probability environments

The definition of probability is going to be given at the level of class variables with satisfiable requirements. In order to specify satisfiability we need to select a definition of probability spaces, and we rely here on Kolmogorov definition as it captures most of the interesting development of probability so far; however, alternatives are possible. Kolmogorov definition can be lifted to class variables using relational requirements, exactly those in parts (a) and (b) of Example 2. These are called foundational requirements for a probability space from now on. Starting from these, the entire theory of probability which stems from Kolmogorov definition can be developed on class variables; one only needs to use relational requirements to give meaning to the various definitions and operations.

In practical problems, however, it is at the start that one introduces additional requirements of more elaborated probabilistic nature, and the issue of consistency with the basic ones arises. The framework that we have elaborated is now suitable for such introduction; this is done by using additional relational requirements imposed before hand. Several of these requirements are discussed starting from next section, relying upon the survey in Appendix B; they can be used to determine probability spaces for individual problems. The definition of probability, however, should not rely upon spe-
cific choices of the additional requirements, as it should incorporate future potential unforeseeable needs, with the only constrained that all requirements be consistent together.

**Definition 2.4.** A collection of requirements for a collection of class variables is admissible if there exists a collection of realizations, one for each of the class variables, which satisfy all the requirements.

An empty family of requirements is always admissible. Here is the main definition.

**Definition 2.5.** Let a collection of requirements for a set variable, a collection of set variables and a set function variable be given. A **probability environment** is a triple \((\Omega, \mathcal{A}, P)\), where \(\Omega\) is a set variable, \(\mathcal{A}\) is a collection of set variables, \(P\) is a set function variable with the properties that:

1. there is a collection of set variables \(\mathcal{A}\) such that \(\Omega, \mathcal{A}\) and \(P\) satisfy the foundational requirements for a probability space;
2. \(\mathcal{A} \subseteq \mathcal{A}\);
3. \(\Omega, \mathcal{A}\) and \(P\) satisfies all the requirements in the given collection;
4. all the requirements together, the foundational, \(\mathcal{A} \subseteq \mathcal{A}\) and the given ones, form an admissible family for the class variables \(\Omega, \mathcal{A}, \mathcal{A}\) and \(P\).

In such case, \(\Omega\) is called universe, the elements \(A \in \mathcal{A}\) are called events, and \(P\) is called probability, and we say that the probabilistic requirements determine the probability environment.

A realization of \((\Omega, \mathcal{A}, P)\) is any quadruple \((\overline{\Omega}, \overline{\mathcal{A}}, \overline{\mathcal{A}}, \overline{P})\), where \(\overline{\Omega}, \overline{\mathcal{A}}\) and \(\overline{P}\) are realizations of \(\Omega, \mathcal{A}\) and \(P\), respectively, satisfying the combination of all foundational, \(\mathcal{A} \subseteq \mathcal{A}\), and given requirements.

The assertion “given a probability environment” means ” given a collection of requirements which, together with the foundational ones, are admissible”.

Requirements for a set variable, a collection of set variables and a set function variable, and other which are discussed later will be indicated as **specific probabilistic requirements**

Any probability space \((\overline{\Omega}, \overline{\mathcal{A}}, \overline{P})\) satisfying Kolmogorov axioms is a probability environment; in fact, it satisfies the basic requirements and the specific
probabilistic requirements that $\Omega = \overline{\Omega}$, $\mathcal{A} = \overline{\mathcal{A}}$ and $P = \overline{P}$; the probability space itself is a proof that the combination of all requirements is admissible. We call this a \textbf{concrete probability space}.

The existence part of the classical extension theorem is a proof of admissibility for the following collection of additional requirements: no condition is asked on the universe, $\mathcal{A}$ is a field (or a semi-ring) of events and $P$ is nonnegative, normalized, countably additive on $\mathcal{A}$. Likewise, the existence part of Kolmogorov’s existence theorem and of similar results are proofs of admissibility for specific requirements.

Between any two realizations $(\Omega_1, \mathcal{A}_1, \mathcal{A}_1, P_1)$ and $(\Omega_2, \mathcal{A}_2, \mathcal{A}_2, P_2)$ there is an obvious relation, a one to one map between $\mathcal{A}_1$ and $\mathcal{A}_2$ preserving the properties determined by the requirements. In some cases such map can be extended further. One particular case of this, when there are no probabilistic requirements and one realization is a probability subspace of the other, the extension of the map is the one considered in the dogma in [T].

### 2.3 Specific probabilistic requirements

The specific probabilistic requirements which can be set preliminarily to the definition of a probability environment are those giving shape to the environment. They are meant to incorporate in the theory the informal requirements surveyed in Appendix B. Indeed, they can be imposed on the set variables or on the set function variables, separately or jointly. Based on Appendix B, we now review the most typical ones and illustrate how informal assertions can they be imposed through requirements.

1. Any assertion on the set variable $\Omega$ is the requirement that each realization $\overline{\Omega}$ of $\Omega$ satisfies that assertion. This covers all possible assertions like being finite, or countable, or have any particular feature.

2. Likewise, any assertion on the collection of set variables $\mathcal{A}$ can be stated as requirement. This is the case for assertions concerning the cardinality or the composition of $\mathcal{A}$.

3. Requirements of more probabilistic nature are generally imposed on $P$. For instance, one often requires that $P(A) = a$ or $P(A)$ be in some subset of the unit interval. We mention two more requirements as they naturally resolve the conflict between definitions and assumptions.

The first one is independence:
Definition 2.6. The assertion that two events $A_1$ and $A_2$ are independent under a probability $P$ is the requirement that $P(A_1 \cap A_2) = P(A_1)P(A_2)$.

The assertion that the events of a family $A_i, i \in I$, are collectively independent is the requirement that for all finite set of indices $i_1, \ldots, i_k \in I$

$$P(\cap_{j=1}^{k} A_{i_j}) = \prod_{j=1}^{k} P(A_{i_j}).$$

Notice that, although the definition is unchanged from the one offered in the standard axiomatic theory, it means something different: the first part of the definition, for instance, really consists of the requirement that in every realization $\overline{A_1}, \overline{A_2}$ and $\overline{P}$ of $A_1, A_2$ and $P$, respectively, $\overline{P}(\overline{A_1} \cap \overline{A_2}) = \overline{P}(\overline{A_1})\overline{P}(\overline{A_2})$, and that this requirement is consistent with all the others which are being imposed.

Definition 2.7 (Conditional probability). For two events $A_1$ and $A_2$, a probability $P$ and a real number $a \in \mathbb{R}$, the assertion that $P(A_1|A_2) = a$ is the requirement $P(A_1 \cap A_2) = aP(A_2)$ and $P(A_2) \neq 0$.

Definition 2.8 (Conditional independence). For three events $A_1, A_2, A_3$ and a probability $P$, the assertion that $A_1$ and $A_2$ are independent given $A_3$ is the requirement that

$$\frac{P(A_1 \cap A_2 \cap A_3)}{P(A_3)} = \frac{P(A_1 \cap A_3)}{P(A_3)} \frac{P(A_2 \cap A_3)}{P(A_3)}$$

and $P(A_3) \neq 0$.

2.4 Random maps and variables

Similarly to what done for probabilities, it is possible to lift random maps and variables to class variables. To this purpose we consider map variables, which are variables taking values in maps from one set to another. Usual map properties and operations are lifted to map variables by means of operational requirements.

Example 3. The statement that a map variable $f$ is real valued is the requirement that for each realization $\overline{f}$ of $f$ takes values in $\mathbb{R}$. 

9
The statement that a real valued map variable \(-f\) is the additive inverse of a real valued map variable \(f\) is the requirement that for each realizations \(-\bar{f}\) and \(\bar{f}\) of \(-f\) and \(f\), respectively, \(-\bar{f} = -\bar{f}\).

The statement that a real valued map variable \(f + g\) is the sum of two real valued map variables \(f\) and \(g\) is the requirement that in every realizations \(\bar{f}, \bar{g}\) and \(\bar{f} + \bar{g}\) of \(f, g\) and \(f + g\), respectively, \(\bar{f} + \bar{g} = \bar{f} + \bar{g}\).

Relational requirements allow to introduce measurability.

**Example 4.** Given two set variables \(\Omega, \Omega'\) and \(\sigma\)-algebras \(A, A'\) of set variables, all contained in \(\Omega\) and \(\Omega'\), respectively, the statement that a map variable \(f\) defined on \(\Omega\) and taking values in \(\Omega'\) is \(A - A'\)-measurable is the requirement that for all \(A' \in A'\) we have \(f^{-1}(A) \in A\).

**Definition 2.9.** Given a probability environment \((\Omega, A, P)\) with requirements \(R\), and a further collection \(R'\) of requirements for a map variable, a random map \(X\) is a map variable on \(\Omega\), measurable with respect to a \(\sigma\)-algebra \(A\) of set variables, all subsets of \(\Omega\), with \(A\) containing \(A\), with the properties that the combination of all requirements \(R\) and \(R'\) together form an admissible family.

If one of the requirements is that the random map takes value in \(\mathbb{R}\) then \(X\) is called random variable.

### 2.5 Probabilistic requirements for random maps and variables

One can produce an analogous list of requirements commonly imposed in random maps and variables.

A collection of random maps determines a joint distribution, which is a nonnegative, countably additive normalized set function variable \(P\), and requirements on this can be phrased in the context of requirements for probabilities.

Further conditions are imposed on moments or conditional expectations, which can be defined as follows.

**Example 5.** Given a set variable \(\Omega\) and a \(\sigma\)-algebra \(A\) of set variables, all contained in \(\Omega\); a nonnegative, countably additive set function variable \(\mu\) on \(A\); and a real valued, \(A - B\)-measurable map variable \(f\) defined on \(\Omega\), with \(B\) the real Borel \(\sigma\)-algebra, the statement that a real variable \(\int f d\mu\) is the
integral of $f$ with respect to $\mu$ is the requirement that for every realizations $\int fd\mu$, $\mu$ and $f$, respectively, $\int f \, d\mu = \int f \, d\mu$.

**Example 6.** Given a set variable $\Omega$, two $\sigma$-algebras $\mathcal{A}$ and $\mathcal{A}'$ of set variables, all contained in $\Omega$, with $\mathcal{A}' \subseteq \mathcal{A}$, nonnegative, countably additive normalized set function variable $P$ on $\mathcal{A}$, and a real valued, $\mathcal{A} - \mathcal{B}$-measurable map variable $f$ defined on $\Omega$, the statement that a real valued, map variable $E_P(f|\mathcal{A}')$ is the conditional expectation of $f$ given $\mathcal{A}'$ is the requirement that $E_P(f|\mathcal{A}')$ is $\mathcal{A}' - \mathcal{B}$-measurable and for all $A \in \mathcal{A}'$, $\int_A f \, dP = \int_A E_P(f|\mathcal{A}') \, dP$.

# 3 Comments and examples

## 3.1 Consequencies and constructivism

There are two types of consequencies which can be drawn from the definition of probability environments and random maps:

**Definition 3.1.** We say that a statement is a possible consequence of a probabilistic environment $(\Omega, \mathcal{A}, P)$ or a random map $X$ if the statements holds for at least one of its realizations $(\Omega, \mathcal{A}, P)$ or $X$; a statement is a necessary consequence if it holds for all of the realizations.

It is a good time to discuss constructivism in applied mathematics. As natural, applied mathematics is more conveniently developed using restricted sets of axioms forcing the explicite construction of the mathematical entities appearing. This is the origin of phylosophical considerations asserting that constructive mathematics has a higher status than the nonconstructive one, and axioms like the law of excluded middle and the axiom of choice come under criticism (see, for instance, [Br] or [TvD]).

This is certainly an appropriate issue for the consequencies to be derived from the assumptions of a probabilistic environment or a random map. Once the requirements are identified one is naturally lead, in the framework of applications, to make deductions in the safest and most explicit way; this is, in very rough synthesis, the aim of each of the many forms of constructivism. On the other hand, the formalization of the previous sections indicates that there is a preliminary step to be done, namely verification of non contradiction by checking admissibility of the requirements. This step is actually best made in the most abstract and non constructive way: an explicit construction would for the most part be an excessive effort, as we see below;
furthermore, the use of even the most controversial axioms for mathematics is welcomed, as long as it is not itself contradictory, as it guarantees absence of contradiction with a larger set of axioms.

In a sense, the present construction resolves the conflict between constructive and noncostructive mathematics (some similarities can be found in [P]).

3.2 Explicit constructions and uniqueness

As existence of a suitable probability space is the key element above, an explicit construction of a probability space satisfying the requirements is not a relevant issue, and for the most part could be avoided. Of course, there are instances in which an explicit construction is also achievable; if so, this adds additional information: the determination of possible consequencies, for instance, becomes easier.

Similarly, uniqueness of the probability space satisfying the requirements is not a relevant issue. First, it is not required for admissibility or consistency. Second, it requires some effort to specify its precise meaning, as any probability space can be enriched by some independent, or even not independent, irrelevant random structures.

This said, uniqueness results (modulo some class of transformations) about probability spaces satisfying specific requirements are also quite informative: a proof of essential uniqueness can be used to more easily derive necessary consequencies of the requirements.

3.3 Finite vs. countable additivity

For simplicity this exposition has been made in the framework of Kolmogorov axiomatics, including, in particular, countable additivity. One reason for this choice is that the majority of probabilistic works and results is developed under these assumptions.

It is possible, however, to simply assume finite additivity, and some authors and works argue that this is a better choice. From the point of view of probabilistic environment this is simply a modeling choice: the entire construction can be repeated by assuming that concrete probability spaces satisfy finite additivity only; let’s indicate the outcome of this construction
by finitely additive probabilistic environment. Finitely additive probabilistic environment have more models, hence more sets of requirements are admissible, but less consequences can be drawn as compared to using what we merely called probabilistic environment.

It is thus entirely a modeling issue, balancing applicability against consequences. There are no technical difficulties and Sections 2 and 3 can read by replacing finitely additive probability environments to countably additive ones.

The main focus and novelty remains with the interpretation of existence of suitable concrete probability space, whether countably additive or not, and the lack of a main role of uniqueness (see [K]) for instance, in which lack of uniqueness seems to be a big issue).

As a trivial example of necessary consequence, if \( \mathbb{A} \) includes a countable family of disjoint events \( A_i \) and a collection of other requirements is imposed on \( (\Omega, \mathbb{A}, P) \) including requirements on \( P(A_i) \) and not on \( P(\cup_{i=1}^{\infty} A_i) \), then whether we have \( P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \) or not depends on whether we are using countably or just finitely additive probabilistic environments.

### 3.4 Examples

There are many examples in which an existential result has been looked for in order to guarantee that the use of specific assumptions were not contradictory. Here we list just a few of them, to exemplify how the formulation of requirements works in different contexts and how it has been previously treated. Basically, this happens all the times in which probabilistic structure is defined ”axiomatically”. What was missing was a unifying theory assigning a central role to the search for the existential results; for instance, existential results have been often put on the same level as the question of uniqueness (modulo some natural equivalence relation), which has a less relevant role. Another weakness of the current treatment is that the preference has been given to cases in which existence is insured in great generality, as compared to situations in which the values of several parameters might be critical to existence.

1. The first example concerns independent events or random variables: one can prove, as often presented even in elementary texts, that families of any cardinality of collectively independent events with prescribed
probabilities or of independent random variables with prescribed distributions always exist. This guarantees that one can always assume any desidered collection of (collectively) independent objects, and, in case, enrich it with additional independent ones. On the other hand, this falls into the class of problems in which existence is so amply guaranteed as to pose no critical issue about admissibility of the requirements.

2. A situation in which existence is not guaranteed, and hence admissibility conditions must be carefully analyzed, would be that of assuming independence only among some selected subsets of the events, and imposing probabilities for other collections: one result, which could be called "existence theorem for partially independent events", is presented in Section .... below. More complicated situations arise when also conditional probabilities or analogous requirements are imposed.

3. Finite Markov chains are defined by prescribing a finite (for finite time) or, in general, an infinite number of requirements on some conditional probabilities:

\[ P(X_n = x_n \mid X_{n-1} = x_{n-1}, X_{n-2} = x_{n-2}, \ldots) = P(X_n = x_n \mid X_{n-1} = x_{n-1}) = a_{n,x_n,x_{n-1}} \]

for all suitable \( n, x_n, x_{n-1}, x_{n-2}, \ldots \). The existence result is easily proven by selecting suitable matrices and vectors.

4. Likewise, in the theory of Gibbs distributions with finite state space, a family of conditional probabilities or conditional expectations is prescribed, depending on some interaction function, and existence is proven by taking weak limits of appropriate finite volume distributions (see [Ruelle] and [Dobrushin almost everywhere interactions]).

5. The most typical example is the axiomatic definition of Brownian motion, in which only some probabilities, independencies and conditional probabilities are prescribed. Einstein used this assumptions to derive some physical constants, but the issue of non contradiction of his hypothesis, casted in terms of existence of the actual model, has been solved some 20 years later by Wiener (whether his and later proofs were more or less constructive has the relevance discussed in the previous section).
6. In general, any axiomatic definition of a probability space or stochastic process has lead to considerations about its existence; see, for instance, [Varadhan S.R.S., Williams R.J.: Brownian motion in a wedge with oblique reflection. Comm. Pure Appl. Math. 38, 405443 (1985)]

7. The binomial model in finance is also described axiomatically, and then there are solutions (more than one) satisfying the assumptions (see, e.g., Vassiliou...page 142)

8. De Finetti has analyzed the case in which probabilities are assigned to some events obtained by Boolean operations on a finite number of initial events. The question of existence boils down to solving a system of linear equations. The complexity has then led to the formulation of satisfiability problem called SAT.

9. The exponential distribution can be found to be a memoryless distribution with measurable density. Similarly, the Poisson point process can be axiomatically defined.

10. The recent development on scaling limits in percolation and other random structures started from Schramm hypothesis of the existence of scaling limit with certain properties: per se they might have been contradictory. Using these properties one derives the SLE construction [O. Schramm (2000), Scaling limits of loop-erased random walks and uniform spanning trees, Israel J. Math. 118, 221-288]; subsequently, existence has been proven [S. Rohde, O. Schramm (2005), Basic properties of SLE, Ann. Math. 161, 879-920, basically Th. 3.6].

11. The classical moment problem or cumulants. See the generalized moment problem in Jean Bernard Lasserre: Moments, Positive Polynomials and Their Applications - Imperial College Press, World Scientific, Singapore, 2010

There are, however, no broad results which allow for a wide selection of requirements. As a partial answer, we see in the next section that the case of finitely many events is quite often suitable of a definite answer to the admissibility problem.
4 The finite case

4.1 Reduction to systems of real equations and inequations

Many applied problems can be treated by considering only a finite number of events. As discussed in Appendix B, the informal requirements on a finite number of events can be expressed by taking relations between probabilities of boolean combinations of such events and imposing that an equality or inequality be satisfied. We now formalize such requirements, as done in the previous section, to encompass the fact that they must be phrased with no reference to existing events or probabilities; we see that this results in solving a system (with potentially infinitely many terms, see the example below) of equations and inequations involving real valued functions of real variables. In most cases relations are polynomials and this leads to decidability of a large collection of problems.

Consider a probability environment $(\Omega, \mathcal{A}, P)$ as in Definition 2.5 with the following requirements:

1. no requirements are imposed on $\Omega$;
2. the requirement on $\mathcal{A}$ is that there are $n$ events $A_1, \ldots, A_n$ in $\mathcal{A}$;
3. there are requirements on $P$ given in terms of $k$ boolean combinations $B_1, \ldots, B_k$, each expressed in terms of unions, intersections and complements of the $A_i$’s; for each boolean combination $B_j$ consider a variable $x_j$; consider then a family of real valued functions $g_r = g_r(x_1, \ldots, x_k), r \in R$, in the variables $x_j$, $j = 1, \ldots, k$, with $R$ a set of indices of any possible cardinality, and equations or inequations $g_r = g_r(x_1, \ldots, x_k) \leq 0$, $r \in R$; the requirements on $P$ are that if $x_j = P(B_j)$ for all $j = 1, \ldots, k$, then the equations and inequations are all satisfied. More explicitly, the above requirements mean that for each realization $(\Omega, \overline{\mathcal{A}}, \overline{P})$ of $(\Omega, \mathcal{A}, P)$, if $\overline{B}_j$ is the $j$-th boolean combination of the $\overline{A}_i$’s expressed with each $A_i$ replaced by $\overline{A}_i$ and $\overline{x}_j = \overline{P}(\overline{B}_j)$, then $\overline{g}_r := g_r(\overline{x}_1, \ldots, \overline{x}_k) \leq 0$ for all $r \in R$.

This can be called finite probability environment. Existence of the environment, i.e. admissibility of the appropriate collection of requirements, is discussed in the next theorem.
Lemma 4.1. The above family of requirements determines a probability environment if and only if the following happens.

For every \( j = 1, \ldots, k \) let \( B_j \) be expressed in disjunctive normal form \( B_j = \bigcup_{\alpha \in \Sigma_j} A^\alpha \) for the appropriate \( \Sigma_i \), where \( \alpha \in \Sigma = \{0, 1\}^n \), \( A^0 = A^c \), \( A^1 = A \) and \( A^\alpha = \bigcap_{i=1}^n A_i^\alpha \). Consider then the change of variables \( x_j = \sum_{\alpha \in \Sigma_j} y_\alpha \), using the \( 2^n \) variables \( y = \{y_\alpha\}_{\alpha \in \Sigma} \), one for each of the \( A^\alpha \). Then the family of requirements is admissibile if and only if the system of equations and inequations

\[
g(r(x_1(y), \ldots, x_k(y))) = g_r(\sum_{\alpha \in \Sigma_1} y_\alpha, \ldots, \sum_{\alpha \in \Sigma_k} y_\alpha) < 0,
\]

obtained by the change of variables \( x_j = x_j(y_\alpha) \), together with the additional normalizing condition \( \sum_{\alpha \in \{0,1\}^n} y_\alpha = 1 \) on the \( y_\alpha \)'s, admits solution \( y = (y_\alpha)_{\alpha \in \Sigma} \) with \( y_\alpha \geq 0 \).

Proof.

The above theorem suggests that finite probability environments can be classified based on the number of equations and inequations and the type of functions involved in them. Some problems involve uncountably many conditions and hence uncountably many \( g \)'s; this is the case when looking for maximum entropy probability distributions in a certain class: here, for instance, one requires that \( \Omega \) be of some finite size \( m \) and that the probabilities \( p_i \)'s satisfy \( \sum_{i=1}^m p_i \log(p_i) \geq \sum_{i=1}^m p'_i \log(p'_i) \) for all \( p'_i \)'s in the appropriate uncountable family.

Most problems, and hence most finite probability environments, involve, however, only finitely many equations and inequations. In some cases such relations are linear, as when one requires given probabilities or conditional probabilities or moments; in most cases they are polynomial, as when one requires independence or cumulants. The problems with a finite number of polynomial equations or inequations can be called polynomial finite probabilities or polynomial finite probability environments. They constitute the vast majority of the probability environments appearing in exercises or applications of finite probabilities.

Notice that instead of performing the \( x-y \) change of variables and then look for a solution to the \( g(x(y)) < 0 \)'s, one could look for a solution \( \bar{x} \) of the \( g(x) < 0 \)'s first, and only later solve \( x(y) = \bar{x} \) with the additional constraints on \( y \). However, this method is not so well suited for proving existence of solutions, as the mere existence of \( \bar{x} \) is in general not suitable for solving the next step.
This method could, on the other hand, be quite convenient to show absence of a solution and hence non admissibility: if a system with even a subset of the $g(x) \triangleleft 0$’s has no solution, then the requirements are not admissible. This looks, however, a lot like hunting for a contradiction by almost randomly drawing consequences from the assumptions, so it really adds no new value to algebrization. By using the contradiction already observed in Appendix 1, at the end of Section 3.4 we show an example of how some of the $g(x) \triangleleft 0$’s can be combined to show nonexistence of a solution.

4.2 Algebrization and decidability of the (large) class of polynomial finite probabilities

The theory for polynomial finite probability environments is suitable for an algebraization which, through Tarski-Seidenberg elimination theorem, leads to a proof of its decidability.

As above, consider a polynomial finite probability environment $(\Omega, A, P)$ in which no requirements are imposed on $\Omega$, and the requirement on $A$ is that there are $n$ events $A_1, \ldots, A_n$ in $A$. The requirements on $P$ are given as above in terms of functions $g_r$ of variables $x_j$, one for each boolean combination $B_j$ (necessarily a finite number), but now assume that the $g_r$ are in finite number, i.e. $r = 1, \ldots, m$, and are polynomials.

**Corollary 4.2.** The problem of determining if the family of requirements of a polynomial finite probability environment is admissible is decidable.

**Proof.** By the Lemma in the previous section, the admissibility problem now requires the nonemptiness of the semi-algebraic variety $g_r = g_r(x_1, \ldots, x_k) \triangleleft 0$, $r = 1, \ldots, m$, in the variables $x_j$’s.

Using Tarski-Seidenberg elimination theorem and Sturm’s theorem, the existence of a solution is decidable. 

It should be noticed that Tarski-Seidenberg theorem, and, even more, Sturm’s theorem are purely existential. In fact, they predict existence or absence of solutions of polynomial equations and inequations in certain intervals even in cases in which the solutions cannot be found by radicals. There are other purely existential method to determine the existence of a solution, for instance via the positivstellensatz (see next section). These other methods
are often more easily computable, as, although the determination remains NP complete, for many problems there is a polynomial solution (see [Pa]).

This is exactly a situation in which the existential result is available, not difficult to apply and with an answer which is, in principle at least, definite, while an explicit determination of the probabilities which would satisfy the requirements is sometimes not possible; even the efforts of finding suitable approximations would be a waste of energy, as they would hardly ensure existence, and most of the probabilities being approximated would be useless anyway.

4.3 Example

The following example, formulated in the language of an applied problem or exercise, has been constructed to provide an extremely simple situation in which one can see the role played by purely existential results.

Example 7. Show that if five equiprobable collectively independent events are such that the probability of each exceeds that of the overall intersection by 0.5, then the probability that at least one events occurs can be controlled (i.e. is bounded) by the sum of the probabilities of any three of them (instead of all five as would follow from subadditivity). Show that if the excess is 0.55 instead of 0.5, then it is controlled even by sum of the probabilities of two of them.

Incorrect Solution 4.3. Let $A_i$, $i = 1, \ldots, 5$ indicate the five events, and let $a = P(A_i) - P(\bigcap_{i=1}^{5}A_i)$ be the indicated excess; thus, we start from $a = 0.5$. By inclusion-exclusion, independence and equiprobability, the probability that at least one occurs satisfies

$$P(\bigcup_{i=1}^{5}A_i) = \sum P(A_i) - \sum_{i_1 \neq i_2} P(A_{i_1} \cap A_{i_2}) + \sum_{i_1 \neq i_2 \neq i_3} P(A_{i_1} \cap A_{i_2} \cap A_{i_3})$$

$$= 5P(A_1) - 10P(A_1)^2 + 10P(A_1)^3 - 5P(A_1)^4 + P(A_1)^5$$

$$= 6P(A_1) - 10P(A_1)^2 + 10P(A_1)^3 - 5P(A_1)^4 - a.$$

If $z = P(A_1)$, then one can see (for instance by observing that all roots are complex and the value in $z = 1$ is positive) that $3z - (6z - 10z^2 + 10z^3 - 5z^4 - 0.5) > 0$ for all $z$.  

19
Similarly, if \( a = 0.55 \) then \( 2z - (6z^2 + 10z^3 - 5z^4 - 0.55) > 0 \) for all \( z \). Hence, the claimed conclusion holds for both \( a = 0.5 \) and \( a = 0.55 \).

**Solution 4.4.** We need first to assess admissibility of the requirements for a probability environment \((\Omega, A, P)\). There are no requirements on \( \Omega \) and \( A \) is required to be of size 5. So let \( A = \{A_i, i = 1, \ldots, 5\} \), let \( a = P(A_i) - P(\cap_{i=1}^5 A_i) \) be the indicated excess, and consider the boolean combinations \( B_\beta = \cap_{i \in \beta} A_i \) for \( \beta \subseteq \{1, 2, \ldots, n\} \). The stated requirements are then expressed by the following system

\[
\begin{align*}
  x_{i,1,2,4,5} + a - x_{i,1} &= 0, & i &= 1, \ldots, 5 \\
  x_\beta - \prod_{i \in \beta} x_i &= 0, & \beta &\subseteq \{1, 2, \ldots, n\} \\
  x_{i,1} - x_{j,1} &= 0, & i, j &= 1, \ldots, 5, i \neq j.
\end{align*}
\]

Perform the change of variables \( x_\beta = \sum_{\alpha: \alpha_i = 1 \text{ for } i \in \beta} y_a \) and add the normalization equation to get the system

\[
\begin{align*}
  y_{1,1,1,1} + a - \sum_{\alpha: \alpha_i = 1 \text{ for } i \in \beta} y_a &= 0, & i &= 1, \ldots, 5 \\
  \sum_{\alpha: \alpha_i = 1 \text{ for } i \in \beta} y_a - \prod_{i \in \beta} \sum_{\alpha: \alpha_i = 1} y_a &= 0, & \beta &\subseteq \{1, 2, \ldots, n\} \\
  \sum_{\alpha: \alpha_i = 1} y_a - \sum_{\alpha: \alpha_j = 1} y_a &= 0, & i, j &= 1, \ldots, 5, i \neq j \\
  \sum_{\alpha \in \{0,1\}^n} y_a - 1 &= 0.
\end{align*}
\]

The nonemptiness of the semi-algebraic variety determined by the last system can be determined by cylindrical reduction according to Tarsky-Seidenberg elimination, but it is shorter to proceed with some substitutions. Indicating by \( z \) the common value of the \( \sum_{\alpha: \alpha_i = 1} y_a \)'s one gets to

\[
\begin{align*}
  y_{1,1,1,1} + a - z &= 0, & i &= 1, \ldots, 5 \\
  z - a - z^5 &= 0 \\
  \sum_{\alpha \in \{0,1\}^n} y_a - 1 &= 0.
\end{align*}
\]

For \( a = 0.5 \) we cannot solve the second equation by radicals, but methods like Sturm’s theorem or VCA or VAS allow to determine that there are indeed two solutions \( \tau \in [0,1] \). Using one such solution, of which we only know the existence, one can define \( \overline{\Omega} = \{0,1\}^5, \overline{A} = P(\overline{\Omega}), \overline{A}_i = \{\alpha \in \overline{\Omega} : \alpha_i = 1\} \subseteq \overline{A} \), and \( P(\overline{A}_i) = \tau \), extended by independence to all other concrete events. \((\overline{\Omega}, \overline{A}, \overline{P})\) is a concrete probability space, so the requirements are admissible.
The calculation (1) shows that the claimed conclusion holds as a necessary consequence of the requirements.

For \( a = 0.55 \), however, the same methods show that there are no roots in \([0, 1]\), hence the requirements are not admissible; the claimed assumptions are contradictory and the calculation in (1) does not make any sense.

**Remark 4.5.** Sturm’s theorem gives also a condition on \( a \) for the existence of a solution in \([0, 1]\) of \( z - a - z^5 = 0 \). The Sturm sequence in \( z = 0 \) is \( a, -1, -a, 1 - \frac{3125a^4}{256} \) and in \( z = 1 \) is \( a, 4, 4/5 - a, 1 - \frac{3125a^4}{256} \). Hence, there are two solutions in \([0, 1]\) for \( a \in \left[0, \frac{4}{35/4}\right) \), one for \( a = \frac{4}{35/4} \) and none for \( a > \frac{4}{35/4} \). Hence, the requirements are admissible if and only if \( a \leq \frac{4}{35/4} \approx 0.535 \).

In Example 7 as it happens in all the problems of the part of the theory we have named polynomial finite probability, the identification of \( \Omega \) and \( A \) posed no problems, it was the probability which ended up needing a non constructive existential result to ensure of its existence. One can set up an example in which even for the universe only existence can be established by imposing requirements on the set variables as well. For instance, by requiring that the universe be the set solution of some polynomial equations, and the probability uniform with each element of probability \( 1/k \) for a fixed number \( k \). Following Sturm theorem, one can determine how many solutions there are (albeit in some cases cannot determine them) and the family of requirements is admissible if and only if \( k \) is such number.

### 4.4 Positivstellensatz and Dutch Books for polynomial finite probabilities

We have seen that requirements admissibility for polynomial finite probabilities is decidable. We see now that the same problem is amenable of a dual interpretation, in which admissibility is equivalent to the absence of a (sometimes weak) Dutch Book, i.e. a rigging strategy against thebeliever of inadmissible requirements. There are two reasons for developing this dual interpretation: one is that feasibility of a semialgebraic variety, which is the algebraic translation of the admissibility problem, is NP-complete, and the dual interpretation allows the development of several techniques (discussed for instance in [Pa], and application, in Bleckerman, Parrilo and Thomas..... or the book on Moments) which in many instances shorten the calculations; the second reason is that the dual interpretation has a meaning in itself, to
the point that it could even be taken as starting point of the construction of probabilities.

This last point of view was suggested by De Finetti, who in fact discovered the linear part of what we are going to expose and named it the fundamental theory of probabilities. Although we see here that the dual interpretation carries over to all polynomial finite probabilities, hence to most finite probabilistic problems, and certainly extends further, it turns out that determining the correct dual problem is not so natural at first (see, for instance, Wally et al., where there are inconclusive attempts to extend a linear dual interpretation). Such problem has evolved into PSAT...

Let’s call a believer of inadmissible requirements for polynomial finite probabilities an incorrect evaluator of probabilities. We also say that a weak Dutch Book against an individual is a game in which the individual believes to have an average strictly positive gain, while instead the game is a draw in every single realization. A Dutch Book against an individual is a game in which the individual believes to have an average positive gain, while instead (s)he loses an amount bounded away from 0 in every single realization.

**Theorem 4.6.** A family of requirements for polynomial finite probabilities concerning \( n \) events or random variables is not admissible if and only if, assuming that it is possible to realize a finite but sufficiently large number of independent copies of the collection of events and random variables, it is possible to realize a weak Dutch Book against any incorrect evaluator believing such requirements.

Some care must be used in interpreting the content of this theorem. When talking about an (incorrect) evaluator of finite polynomial probabilities we intend that he/she has determined the set \( \Omega \) to be used as concrete universe and the sets \( A \) to be used as the events on whose probabilities the requirements are being discussed, and that he believes those requirements hold for the pair \((\Omega, A)\) and some probability \( P \) on them. One of the assumptions in the theorem is that it is possible to find or produce a, finite but sufficiently large, number of copies of set \( \Omega_j \), in each of which there is collection of events \( A_j \), in bijection with \( A \), and each equipped with a random mechanism under which the evaluator believes that the events in \( A_j \) have the same probability as the corresponding one in the bijection. Furthermore, the copies are produced in such a way that the evaluator considers the events and random variables in different copies as independent.
More explicitly, there are polynomials $t$, maximal power of the variable $y$, integers $k$, \cite{BochnakCosteRoy:RealAlgebraicGeometry}. A are boolean combinations of the events in the $x$ generated by the $h$, such that all its coefficients are zero. As such, let $\sum y_{\alpha,s}$ the $y_{\alpha,s}$ using the standard normal form; and add the normalizing condition $\sum_{\alpha,s} y_{\alpha,s} = 1$ and equations asking that all $y_{\alpha,s} \geq 0$.

Consider now the system which includes the $g_r(x_i(y_{\alpha,s})) \leq 0$, $r = 1, \ldots, m'$, the normalizing equation and the nonnegativity conditions on $y$; by distinguishing the three possible values of $\triangleleft$, denote it as follows:

$$
\begin{cases}
    f_r(y) = 0, & r = 1, \ldots, m_1 \\
    g_r(y) \geq 0, & r = m_1 + 1, \ldots, m_1 + m_2 \\
    h_r(y) \neq 0, & r = m_1 + m_2 + 1, \ldots, m_1 + m_2 + m_3 = m'.
\end{cases}
$$

It represents a semialgebraic variety, which is feasible, i.e. nonempty, if and only if the requirements are admissible. By the positivstellensatz (see \cite{BochnakCosteRoy:RealAlgebraicGeometry}), the system has no solution if and only if the following happens. There exists a polynomial $F$ in the ideal generated by the $f_r$’s in $\mathbb{R}[y]$, a polynomial $G$ in the cone generated by the $g_r$’s in $\mathbb{R}[y]$ and polynomial $H$ in the multiplicative monoid generated by the $h_r$’s in $\mathbb{R}[y]$ such that

$$
F + G + H = 0.
$$

More explicitly, there are polynomials $t_r \in \mathbb{R}[y], r = 1, \ldots, m_1$; $s_J \in \mathbb{R}[y]$, $J \subseteq \{m_1 + 1, \ldots, m_1 + m_2\}$ which are sum of squares; and even integers $k_r, r = m_1 + m_2 + 1, \ldots, m_1 + m_2 + m_3$, such that

$$
v(y) = \sum_{r=1}^{m_1} t_r f_r + \sum_{J \subseteq \{m_1 + 1, \ldots, m_1 + m_2\}} s_J \prod_{r \in J} g_r + \prod_{r=m_1+m_2+1}^{m_1+m_2+m_3} (h_r)^{k_r} = 0.
$$

\cite{BochnakCosteRoy:RealAlgebraicGeometry}.

We need to investigate the polynomial (4) as a polynomial in the $y_{\alpha,s}$’s before taking into account that all its coefficients are zero. As such, let $\nu_{(\alpha,s)}$ be maximal power of the variable $y_{\alpha,s}$, and consider $\nu = \sum_{(\alpha,s) \in \{0,1\}^{n_1} \times \prod_{s_1}^{n_2} s_{i_2}} \nu_{(\alpha,s)}$;
next, list the \((\alpha, s)\)'s in some fixed order \((\alpha, s)_1, \ldots, (\alpha, s)_{n_1+n_2}\); for each \(t \in \{1, \ldots, n_1 + n_2\}\) let \(\Sigma^{(\alpha, s)_t}\) be the set of all permutations \(\sigma^{(\alpha, s)_t} = (\Sigma^{(\alpha, s)_t}_i), i = 1, \ldots, \nu_{(\alpha, s)}\) of integers \(\{\sum_{t'=1}^{t'-1} \nu_{(\alpha, s)}_{t'}, \ldots, \sum_{t'=1}^{t'} \nu_{(\alpha, s)}_{t'}\}\).

We take \(\nu\) independent copies of the events and random variables identified by the incorrect evaluator. We form now a random variable, basically by replacing each occurrence of the variables \(y_{\alpha, s}\)'s in (4) by indicator function \(\mathbb{I}_{\alpha, s, (j)}\) that the \(j\)-th independent copy of the event \(A^\alpha \cap \{X = s\}\) takes place, and then summing the fully replaced polynomial over all permutations of the indices of the copies. We need to specify how to choose the copy to be used for each replacement. We do this in steps for each selection \(\sigma^{(\alpha, s)}\) of a permutation for each \((\alpha, s)\):

1. consider each of the polynomials \(t_r f_r, s_J \prod_{r \in J} g_r\) and \(\prod_{r'=m_1+m_2+m_3} (h_r)^{k_r}\) separately.

2. In each such polynomial \(u\) consider one of its factors at a time (using the factorization in which they are already expressed, for instance \(t_r\) and \(f_r\) for those in the ideal);

3. expand out completely each such factor into a sum of monomials; consider each monomial separately and if in such monomial the variable \(y_{\alpha, s}\) appears at some power \(m^1_{\alpha, s}\); then replace it by the product \(\prod_{j=1}^{m^1_{\alpha, s}} \mathbb{I}_{\alpha, s, (\sigma^{(\alpha, s)}_j)}\); repeat for all variables in \(y\). Decorate the name of the factor by a tilde to indicate the obtained random variable.

4. Consider the second factor of \(u\); repeat the previous step, with this change: if the variable \(y_{\alpha, s}\) appears at some power \(m^2_{\alpha, s}\); then replace it by the product \(\prod_{j=m^2_{\alpha, s}}^{m^2_{\alpha, s}+1} \mathbb{I}_{\alpha, s, (\sigma^{(\alpha, s)}_j)}\).

5. Repeat, always using \(\mathbb{I}_{\alpha, s, (\sigma^{(\alpha, s)}_j)}\) referred to new \(j\)'s and hence additional copies, till all factors of \(u\) have been changed; notice that the total number of copies of \(A^\alpha \cap \{X = s\}\) used in the procedure is not greater than \(\nu_{\alpha, s}\).

6. Consider the next polynomial from the list in point 1., and repeat steps 2.-5. till all polynomials in 1. have been changed.
The above procedure produces a random variable

\[ \tilde{u} = \sum_{r=1}^{m_1} \tilde{t}_r f_r + \sum_{J \subseteq \{m_1+1, \ldots, m_1+m_2\}} \tilde{s}_J \prod_{r \in J} \tilde{g}_r + \prod_{r=m_1+m_2+1}^{m_1+m_2+m_3} (\tilde{h}_r)^{kr}; \]

which depends on the permutations \( \{\sigma^{(\alpha,s)}\}_{(\alpha,s) \in (0,1)^{n_2} \times \prod_{i_2=1}^{n_2} s_{i_2}} \). Let then \( \tilde{v} = \sum_{\text{all permutations}} \{\sigma^{(\alpha,s)}\}_{(\alpha,s) \in (0,1)^{n_2} \times \prod_{i_2=1}^{n_2} s_{i_2}} \tilde{u} \).

We compute the expected value of \( \tilde{v} \) according to the incorrect evaluator. For each of the polynomials in 1. above, each of its factors contains indicator functions which refer to different copies of the collection of events and random variables: in fact, two indicator functions \( \mathbb{1}_{\alpha,s}(\sigma^{(\alpha,s)}_{j}) \) always refer to different copies if the \( (\alpha,s) \)'s are different; if the \( (\alpha,s) \)'s are the same then the \( j \)'s in \( \sigma^{(\alpha,s)}_{j} \) are different by construction. Therefore, the incorrect evaluator would consider each factor as independent, and factorize the expected value of the product. In each monomial inside each factor, the indicator functions refer to different copies of the collection of events and random variables, and hence they are considered independent by the incorrect evaluator. We have thus that he/she would compute the expected value of the product of the indicator functions replacing the variables of a monomial \( \prod_{\alpha,s} y_{\alpha,s}^{k_{\alpha,s}} \) as

\[ E\left( \prod_{(\alpha,s) \in J_{\alpha,s}} \mathbb{1}_{\alpha,s,\sigma^{(\alpha,s)}_{j}} \right) = \prod_{(\alpha,s) \in J_{\alpha,s}} \left( P(A^\alpha \cap \{X = s\}) \right)^{m_{\alpha,s}} \]

for suitable \( m_{\alpha,s} \geq 0 \) and \( J_{\alpha,s} \) of cardinality \( m_{\alpha,s} \). Hence, the incorrect evaluator would compute \( E(\tilde{u}) \) as the corresponding polynomial in \( \tilde{y} \) with the \( y_{\alpha,s} \)'s replaced by the \( P(A^\alpha \cap \{X = s\}) \)'s. This is the value for which he/she thinks that the relations in (2) hold.

It follows that, if \( \tilde{y} \) indicates the value of \( y \) with the above substitutions,
the incorrect evaluator would compute, by linearity of the expected value:

\[ E(\tilde{v}) = \sum_{\text{all permutations } \{\sigma^{(a,s)}\}_{(a,s) \in \{0,1\}^{m_1} \times \Pi^{m_2}_{i=1} S_{i_2}}} E(\tilde{u}) \]

\[ = \sum_{\text{all permutations}} \left( \sum_{r=1}^{m_1} E(\tilde{t}_r) E(\tilde{f}_r) + \sum_{J \subseteq \{m_1+1, \ldots, m_1+m_2\}} E(\tilde{s}_J) \prod_{r \in J} E(\tilde{g}_r) \right) \]

\[ = \left( \prod_{(a,s)} \nu_{(a,s)}! \right) \left( \sum_{r=1}^{m_1} t_r(\tilde{y}) f_r(\tilde{y}) + \sum_{J \subseteq \{m_1+1, \ldots, m_1+m_2\}} s_J(\tilde{y}) \prod_{r \in J} g_r(\tilde{y}) \right) \]

\[ + \left( \prod_{r=m_1+m_2+1}^{m_1+m_2+m_3} (h_r(\tilde{y}))^{k_r} \right) > 0 \]

from the equalities and inequalities in (2), and the properties of the polynomials and the powers in (4).

We finally evaluate \( \tilde{v} \) for each possible realization of events and values of the random variables in the collections identified by the incorrect evaluator and in all the copies. First, expand \( \tilde{v} \) completely, and then collect all terms corresponding to random variables which have replaced the same monomial \( \Pi_{a,s} y_{a,s}^{k_{a,s}} \). For each such monomial, there is a certain number \( m \) of terms, with coefficients \( c_1, \ldots, c_m \); we have \( \sum_{i=1}^{m} c_i = 0 \) as the corresponding monomial in the expansion of the l.h.s of (4) has zero coefficient. The sum is not immediately zero with the indicator functions, as they refer to different copies. However, considering the permutation and indicating by \( \sigma^{a,s}_j(i) \) the permutation used in such monomial when the coefficient is \( c_i \), we have that
all the random variables related to the same monomial add up to

\[
\sum_{\text{all permutations}} \sum_{i=1}^{m} c_i \prod_{(\alpha, s) \in J_{\alpha, s}} \prod_{j \in J_{\alpha, s}} \mathbb{I}_{\alpha, s, \sigma_{j}^{\alpha, s}}(i)
\]

\[
= \sum_{i=1}^{m} \sum_{\text{all permutations}} c_i \prod_{(\alpha, s) \in J_{\alpha, s}} \prod_{j \in J_{\alpha, s}} \mathbb{I}_{\alpha, s, \sigma_{j}^{\alpha, s}}(i)
\]

\[
= \sum_{i=1}^{m} c_i \left( \nu - \sum_{(\alpha, s): J_{\alpha, s} \neq \emptyset} (m_{\alpha, s})! \sum_{\text{permutations: } J_{\alpha, s} \neq \emptyset} \prod_{(\alpha, s) \in J_{\alpha, s}} \prod_{j \in J_{\alpha, s}} \mathbb{I}_{\alpha, s, \sigma_{j}^{\alpha, s}}(i) \right) \sum_{i=1}^{m} c_i
\]

\[
= 0
\]

where the penultimate equality derives from the fact that in each product all the terms \( \mathbb{I}_{\alpha, s, \sigma_{j}^{\alpha, s}}(i) \) refer to different copies by construction, and hence the sum over all permutations does not depend on \( i \).

The incorrect evaluator is thus willing to pay an entry fee, whose amount cannot be predicted in advance but only determined when the terms in \( \mathbb{I} \) are computed, to participate in a game which, however, is a constant draw. This is the weak Dutch Book mentioned in the statement of the theorem.

The random variable \( \tilde{v} \) representing the Dutch Book game in the above proof is often not the best possible option, especially because to guarantee that it works a large number of permutations has been introduced. For an actual determination of a game one can often select the copies more carefully so as to set up a game which is more obviously "advantageous" for the incorrect evaluator. Here is a very simple example, worked out completely including an alternative choice for the \( \tilde{v} \).

**Example 8.** An incorrect evaluator of the probabilities of two events \( A_1 \) and \( A_2 \) might assume that they are independent, that \( P(A_1|A_2) = 1/2 \) but that \( P(A_1) \neq 1/2 \). Setting \( x_{\beta_1, \beta_2} = P(A_1^{\beta_1} \cap A_2^{\beta_2}) \) the above requirements give rise
to a system with 4 terms

\[
\begin{cases}
x_{1,1} - x_{1,0}x_{0,1} = 0 \\
x_{1,1} - \frac{1}{2}x_{0,1} = 0 \\
x_{0,1} \neq 0 \\
x_{1,0} - \frac{1}{2} \neq 0.
\end{cases}
\]

After substitution with the \(y_\alpha\)'s one gets a system with 9 terms, \(m_1 = 3\) equations, \(m_2 = 4\) inequations, and \(m_3 = 2\) with \(<\) replaced by \(\neq\). A polynomial certifying that there is no solution is

\[
v = \sum_{r=1}^{2} t_r f_r + \prod_{r=8}^{9} (h_r)^2
\]

\[
= \left( (y_{1,1} + y_{-1,1})(y_{1,1} + y_{1,-1} - \frac{1}{2}) \right) \left( y_{1,1} - (y_{1,1} + y_{1,-1})(y_{1,1} + y_{-1,1}) \right)
\]

\[
+ \left( -(y_{1,1} + y_{-1,1})(y_{1,1} + y_{1,-1} - \frac{1}{2}) \right) \left( y_{1,1} - \frac{1}{2}(y_{1,1} + y_{-1,1}) \right)
\]

\[
+ (y_{1,1} + y_{-1,1})^2(y_{1,1} + y_{1,-1} - \frac{1}{2})^2 \equiv 0.
\]

Now follow steps 1.-6. in the proof of the previous theorem; for a fixed set of permutations, we get

\[
\tilde{t}_2 \tilde{f}_2 = \left( -\frac{1}{2}I_{1,1}(\sigma_{1,1}^{(1,1)})I_{1,1}(\sigma_{1,1}^{(1,1)}) - \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(1,1)})I_{1,1}(\sigma_{1,1}^{(-1,-1)}) + \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(1,1)})^2 \right)
\]

\[
- \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(1,1)})I_{1,1}(\sigma_{1,1}^{(-1,-1)}) - \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(-1,-1)})I_{1,1}(\sigma_{1,1}^{(1,1)}) + \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(-1,-1)})^2
\]

\[
\cdot \left( \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(1,1)}) - \frac{1}{2}I_{1,1}(\sigma_{1,1}^{(-1,-1)}) \right)
\]

and \(\tilde{u}\) and \(\tilde{v}\) accordingly. One can see that there is no easy way of controlling the different copies which are used, except that of taking all permutations.

On the other hand, one can construct a simpler and more direct game by identifying the events as much as possible with the original \(A_i\)'s, rather than with their expansion in standard normal form, and by a clever choice of the copies. Here is a possible form, in which the number of the copy is indicated.
after the event:

\[ \tilde{v} = \left( \mathbb{I}_{A_2,(1)}(\mathbb{I}_{A_1,(3)} - \frac{1}{2}) \right) \left( \mathbb{I}_{A_1 \cap A_2,(2)} - \mathbb{I}_{A_1,(4)} \mathbb{I}_{A_2,(2)} \right) \]
\[ + \left( -\mathbb{I}_{A_2,(1)}(\mathbb{I}_{A_1,(3)} - \frac{1}{2}) \right) \left( \mathbb{I}_{A_1 \cap A_2,(2)} - \frac{1}{2} \mathbb{I}_{A_2,(2)} \right) \]
\[ + \mathbb{I}_{A_2,(1)} \mathbb{I}_{A_2,(2)} \left( \mathbb{I}_{A_1,(3)} - \frac{1}{2} \right) \left( \mathbb{I}_{A_1,(4)} \right) \frac{1}{2} \right) \]

Since in each factor of each multiplication there appear only events belonging to copies which are different from those appearing in the other factors (of the same multiplication), independence can be used to show that \( E(\tilde{v}) > 0 \) for the incorrect evaluator; on the other hand, \( \tilde{v} \equiv 0 \) as checked by simple algebraic expansion.

**Corollary 4.7.** If the inadmissible requirements for polynomial finite probabilities contain no strict inequalities then, assuming the possibility of producing independent copies of the events and the random variables, a Dutch Book can be realized against any incorrect evaluator believing such requirements.

**Proof.** If there are no strict inequalities in the requirements, since normalization and nonnegativity of probabilities correspond also to inequalities which are not strict, there are no strict inequalities in (2). Hence, the only \( h \) can be taken to be \( h = 1 \), which generates the multiplicative monoid. From (1) we have \( F + G = -1 \). Following the same construction as above, except for the terms in the multiplicative monoid, one gets

\[ \tilde{v} = \sum_{r=1}^{m_1} \tilde{t}_r \tilde{f}_r + \sum_{J \subseteq \{m_1+1, \ldots, m_1+m_2\}} \tilde{s}_J \prod_{r \in J} \tilde{g}_r. \]

The incorrect evaluator now estimates \( E(\tilde{v}) = 0 \), while for each realization \( \tilde{v}(\bar{y}) = -1 \).

The incorrect evaluator perceives thus the game as fair, while losing a constant unit amount. This is the Dutch Book mentioned in the statement of the corollary.

Notice that in this case the amount lost is fixed, and as such known in advance of any calculation about the polynomials determining the Dutch Book.
One way of certify inadmissibility of the requirements is by forming a Dutch Book using only some of the $g(x) \triangleleft 0$'s. This is the case for the example in Appendix 1, in which a linear subsystem of the system in the $y$'s cannot be solved. A simple linear programming techniques leads to a linear combination of the equations witnessing inadmissibility of the requirements.

On the other hand, if one knows of a contradiction already, this information can be used to directly construct a Dutch Book.

**Corollary 4.8.** If one can show that there is a chain of equalities or inequalities of the form $a_1 \triangleleft a_2 \cdots \triangleleft a_n$ such that $a_1, a_n \in \mathbb{R}$, $a_1 < a_n$, and $a_k$ can be shown to be greater than or equal to $a_{k+1}$ for all $k = 1, \ldots, n-1$ by means of the $g(x) \triangleleft 0$'s, there one can form a Dutch Book as follows.

Suppose that at step $k$ the relation $g_{i_k} \triangleleft 0$ is used to show that $a_k \geq a_{k+1}$; this means that $a_k - a_{k+1} = c_k g_{i_k}$ for some positive constant $c_k$. \( \frac{1}{a_n - a_1} \sum_{k=1}^{n-1} c_k g_{i_k} \) is a Dutch Book.

**Proof.** Consider the following procedure. Start from $a_1$; if $g_{i_1} \triangleleft 0$ is used to show that $a_1 \geq a_2$, then let $c_1$ be the positive constant such that $a_1 - a_2 = c_1 g_{i_1}$; add and subtract $c_1 g_{i_1}$ to $a_1$ to get $a_1 = a_1 + c_1 g_{i_1} - c_1 g_{i_1} = c_1 g_{i_1} + a_2$; continue adding and subtracting $c_k g_{i_k}$ if $a_k - a_{k+1} = c_k g_{i_k}$; iteratively, one gets $a_1 = c_1 g_{i_1} + c_2 g_{i_2} + a_3 = \cdots = \sum_{k=1}^{n-1} c_k g_{i_k} + a_n$. Thus

$$\sum_{k=1}^{n-1} c_k g_{i_k} = a_1 - a_n < 0$$

and

$$\frac{1}{a_n - a_1} \sum_{k=1}^{n-1} c_k g_{i_k} = -1.$$

On the other hand, the incorrect evaluator believing all the $g(x)\triangleleft 0$'s assumes that $E(\frac{1}{a_n - a_1} \sum_{k=1}^{n-1} c_k g_{i_k}) \geq 0$, as for him/her all terms are nonnegative. \qed

The contradiction found in Appendix 1 for the problem presented there leads to a simple (even linear) certificate that even the system of the $g(x)\triangleleft 0$'s has no solution. A linear subsystem of just the second, third and last equation is certified to have no solution by

$$\bar{v} = \frac{0.6}{0.77} (x_{0,1,0} - 0.7) - \frac{1}{0.77} (0.6 x_{0,1,0} + x_{0,0,1} - 0.6) + \frac{1}{0.77} (x_{0,0,1} - 0.95) \equiv -1.$$
5 Countably many events

The case of (the requirements that) there are at most countably many pre-events is the next step. The following corresponds to the change of variables theorem of the finite case.

**Theorem 5.1.** With the requirement that $\mathbb{A}$ is at most countably infinite, suppose $\mathbb{A} = \{A_n\}_{n \in \mathbb{N}}$ and let the, possibly uncountably many, requirements on the pre-probability be given on the $\sigma$-algebra generated by the the pre-events in $\mathbb{A}$. There the requirements are admissible if and only if the following happens. Let $\Omega' = \{0, 1\}^n$ with the Borel $\sigma$-algebra generated by the cylinders and $A'_n = \{\omega : \omega_n = 1\}$. For any witness that the requirements are admissible, i.e. for any concrete probability space $(\Omega, \mathcal{A}, \mathbb{P})$ containing sets $\{A'_n\}_{n \in \mathbb{N}}$ in bijection with the $\{A_n\}_{n \in \mathbb{N}}$, there is a bijection between the Borel sigma algebra in $\Omega'$ and the sigma-algebra generated by the $A'_n$'s in $\mathcal{A}$. Therefore, each requirement can be expressed in terms of sets in the Borel sigma algebra in $\Omega'$. We must then have that for every finite collection of cylinders there is a Borel concrete probability on the Borel sigma algebra in $\Omega'$ satisfying all the requirements involving those cylinders (and possibly others).

**Proof.**

Particular instances occur when each requirement involves only finitely many events (finitary countable probabilities), in particular when each requirements is polynomial (polynomial finitary countable probabilities). Another instance is when there are only countably many requirements (the countable requirements case).

**Corollary 5.2.** In the countable requirements polynomial finitary countable probabilities case the requirements are admissible if and only if for each finite set of requirements together with the normalization condition there is a solution for the (finitely many) involved cylinders. Each step is decidable, hence it is countably decidable. Moreover, the requirements are admissible if and only if there is no Dutch Book for any finite set of requirements read on the probabilities of cylinders together with normalizations.

A special case of this is when all requirements are linear. An even more special case is the one from probabilistic regress (see Arkinson, Peijnenburg, Herzberg, Meester, Woudenberg) in which conditions are found for a solution to exist (Herzberg) (an easy consequence of the corollary due to the simple
structure of the linear relations). In this case the philosophical interpretation seems to lead to the additional request that the probability of the first event $A_0$ be fully specified by the given conditional probabilities, whatever its value; this requirement is not of the polynomial form. To comment on the philosophical aspects, however, it is quite unclear to me why a determinate value would contradict the epistemic regress problem, while an interval would not, given that in nature we we really deal only with (possibly small) intervals. The requirement that $P(A_0) \neq 0$ would fall within the polynomial framework.

6 Hilbert 6th problem

Very broadly speaking, there is an analogy between what was outlined above for probability and the theory of differential equations. In this analogy, satisfiability in the definition of probability environment plays the role of existence of global solutions of differential equations with constraints, while Kolmogorov spaces would correspond to merely local solutions. Dutch books can be seen as variational formulations.

This analogy can be taken further in relation to mechanics. In his 6-th problem Hilbert asks for the axiomatization of those physical disciplines in which mathematics plays an important role, starting from probability and mechanics. In the additional notes, Hilbert mentions the kinetic theory of gases, which is indeed a link between these two disciplines and on whose description and formalization in mathematical terms many progresses have been made. There is, however, a second link between probability and mechanics, which emerges by considering Hilbert’s comments to the progresses in formalizing probability (see Bohlman).

In fact, the very same problem of confusing definitions and axioms appears in classical mechanics. Newton’s second law states that

$$f = ma,$$  \hspace{1cm} (5)

and it is taken as a definition of force in any attempt of giving a rigorous formalization of mechanics. However, we have the same problem as with independence and conditional probabilities: if (5) is a definition of $f$, then $f$ cannot be known without being computed from the r.h.s. If the force comes from elsewhere, then (5) is an axiom, and it is not clear that it is not in contradiction with other assumptions; to clear the contradiction one
can produce a motion satisfying (5) for the given force, but then when (5) is guaranteed to be usable it becomes useless.

Once again, the solution to this dilemma is a purely existential result, which guarantees the existence of global solutions of (5) for all the masses involved and with all the necessary constraints. No such results can be given with the generality of the local solutions existence theorem, but it is possible to single out classes of situations for which conditions for existence are easier to state. Compared to probability, the situation is simplified by the fact that in mechanics the universe is $\mathbb{R}^d$, and the only question is how to choose $d$ (in itself not an easy task which is deeply studied in lagrangian mechanics); at the same time, it is made more difficult by the fact that solutions of even the simplest differential equations, i.e. linear, are exponentials, so that decidability of existence of solutions for systems of linear ordinary differential equations with polynomial constraints is an open problem (see Tarski). An example of conditions of existence and non-existence is James Serrin, Henghui Zou, Existence and non-existence results for ground states of quasilinear elliptic equations, Archive for Rational Mechanics and Analysis 9. XII. 1992, Issue 2, pp 101-130 or Patrizia Pucci, Raffaella Servadei, Existence, non-existence and regularity of radial ground states for $p$-Laplacian equations with singular weights Annales de l’Institut Henri Poincare (C) Non Linear Analysis Volume 25, Issue 3, MayJune 2008, Pages 505-537.

When fewer constraints are imposed on the differential equations, one ends up in the field of the so called inverse problems. In this entire field existence is systematically assumed without proofs One example of existence proof in inverse problems is Some inverse problems and fractional calculus Yutaka Kamimura, where one seeks the force which generates a motion of prescribed maximum position and half-period. But what we aim to here is not existence for inverse problems per se, but rather a mixed inverse-forward problem, in which existence of the coefficients and global existence of the solution are the bases to move forward in deducing consequences.

7 Appendix A: contradictory set up

We present here a simple exercise taken from a widely used, application oriented, very high quality textbook. In the exercise, which is set up as to mimic a realistic production problem, the authors propose a set of as-
sumptions about probabilities and ask for the calculation of several other probabilities; the required calculations can be carried out without difficulties, although with some efforts suitable for beginners. It is very likely that this exercise has been solved thousands of times. What is problematic, however, is that on carrying out one extra calculation one realizes that the actual set of assumptions is inconsistent. In [AT], Exercise 2.8 page 67 presents the following problem.

**Example 9.** On a given day, casting of concrete structural elements at a construction project depends on the availability of material. The required material may be produced at the job site or delivered from a premixed concrete supplier. However, it is not always certain that these sources of material will be available. Furthermore, whenever it rains at the site, casting cannot be performed. On a given day, define the following elements:

\[ E_1 = \text{there will be no rain} \]
\[ E_2 = \text{production of concrete material at the job site is feasible} \]
\[ E_3 = \text{supply of premixed concrete is available} \]

with the following respective probabilities: \( P(E_1) = 0.8, \ P(E_2) = 0.7, \ P(E_3) = 0.95 \) and \( P(E_3|E_2^c) = 0.6 \) whereas \( E_2 \) and \( E_3 \) are statistically independent of \( E_1 \).

(a) Identify the following events in terms of \( E_1, E_2, \) and \( E_3 \):

(i) \( A = \text{casting of concrete elements can be performed on a given day} \);

(ii) \( B = \text{casting of concrete elements cannot be performed on a given day} \).

(b) determine the probability of the event \( B \).

(c) If production of concrete material at the job site is not feasible, what is the probability that casting of concrete elements can still be performed on a given day?

We just briefly mention the intended solution.

1. \( A = E_1 \cap (E_2 \cup E_3) \) \( B = A^c = E_1^c \cup (E_2^c \cap E_3^c) \);
2. by the independence of (any combination of) $E_2, E_3$ from $E_1$ we have

\[ P(B) = 1 - P(A) = 1 - P(E_1 \cap (E_2 \cup E_3)) \]
\[ = 1 - P(E_1)P(E_2 \cup E_3). \]

Since $E_2 \cup E_3 = E_2 \cup* (E_2^c \cap E_3)$ and $P(E_2^c \cap E_3) = P(E_3|E_2^c)P(E_2^c) = 0,6 \times (1 - 0,7) = 0,18$, we have

\[ P(B) = 1 - P(E_1)P(E_2 \cup E_3) \]
\[ = 1 - (0,8 \times (0,7 + 0,18)) = 0,704. \]

3. we have

\[ P(A|E_2^c) = \frac{P(A \cap E_2^c)}{P(E_2^c)} = \frac{P(E_1 \cap (E_2 \cup E_3) \cap E_2^c)}{P(E_2^c)} = \frac{P(E_1 \cap E_3 \cap E_2^c)}{P(E_2^c)} = \frac{P(E_1)P(E_3 \cap E_2^c)}{P(E_2^c)} = \frac{0,8 \times 0,18}{0,3} = 0,48 \]

However, something is not correct in this set up: we have $P(E_3 \cap E_2^c) = P(E_3|E_2^c)P(E_2^c) = 0.18$ so that

\[ P(E_3 \cap E_2) = P(E_3) - P(E_3 \cap E_2^c) = 0.77 > 0.7 = P(E_2) \]

which is a contradiction.

Notice that the contradiction, although direct consequence of the assumptions, does not appear in all the several intended calculations. This is the main risk: one could compute a number of quantities without realizing that the entire modelization is contradictory.

Notice also that the contradiction shows up with a clever choice of the events to examine. In Appendix B and Section 3 we develop a general method to test for contradictions.
8 Appendix B: Informal requirements

We call informal requirements the assumptions which are made in common applications of probability theory. One of the purposes of this paper is to describe a suitable formalism, so first we need to survey what needs to be formalized. This survey assumes familiarity with facing problems in which randomness appears.

In these situations it turns out to be very useful to have a the universe Ω, a family of events $\mathcal{B}$ in Ω (generally not a $\sigma$-algebra at the start), a probability $P$, and often some maps $X$ defined on Ω. Informal requirements may be imposed on either of these elements.

1. On the universe Ω there may be informal requirements about the cardinality (finite, countable etc.) or the characteristics of its elements (integers, real, complex numbers, functions, probability measures etc.). Sometimes, there is no requirement at all on the universe, and any set could be used.

2. The same type of informal requirements about cardinality or composition may be imposed on the events in $\mathcal{B}$ (for instance if one requires that the set of continuous functions be in $\mathcal{B}$). In addition, one can impose conditions on the cardinality of $\mathcal{B}$ (such as asking that it contains only finitely many sets), or impose some conditions on the result of possibly uncountable Boolean operations (the intersection of some sets might be required to be empty, or a property might be required to hold for all $t \in [0, 1]$). Some requirements might include closure of $\mathcal{B}$ with respect to some set operation (as when one wants to use semirings or $\sigma$-algebras of sets). On some occasions, it is just the mere existence of a certain number of sets in $\mathcal{B}$ which is required (as when one is looking for the existence of collectively independent events).

3. Many types of informal requirements are often imposed on $P$. One may require the value of some probabilities of events, often expressed in the form of Boolean operations performed on the events in $\mathcal{B}$, or bounds thereon. Typical informal requirements, which, among other things, single out probability theory from analysis, are about independence, conditional probabilities, conditional independence and the like. Another example would be informal requirements about the entropy.
of a partition of the space. The cardinality of the family of informal requirements ranges from finite to uncountable.

4. Other informal requirements are imposed on random maps. Again, these might be of set theoretical nature, such as selecting the range or asking for measurability, or of probabilistic nature, as when asking for moments, cumulants etc.

These are the informal requirements that a theory of probability should be able to encompass. We underline that the current axiomatic theory is not suitable for if probability is thought of as a map on the subsets of a given set, such concreteness would unduly limit the search of "witnesses" that the assumptions are not contradictory. Chapter 2 develops the abstract theory capable of formally incorporating requirements.

For definiteness, one could explore the mechanism of prior requirements in the simple case in which only a finite number of events $E_1, \ldots, E_n$ are allowed (already a requirement). One might, as in the example in Appendix A, demand that the events have certain probabilities, independencies or conditional probabilities; there the assumptions claim, among other things, that $P(E_1) = 0.8$, $P(E_1 \cap E_2) = P(E_1)P(E_2)$, $P(E_3|E_2^c) = 0.6$. These conditions can actually be interpreted as

$$
\begin{align*}
P(E_1 \cap E_2) - P(E_1)P(E_2) &= 0 \\
P(E_1) - 0.8 &= 0 \\
P(E_3 \cap E_2^c) - 0.6P(E_2^c) &= 0 \\
P(E_2^c) &\neq 0.
\end{align*}
$$

Notice that these are all functional relations among probabilities of finite boolean combinations of the $E_i$'s. This is actually the case in all probability problems involving finitely many events. In Section 3 we show that leads to a functional interpretation, and, in most cases, to algebrization and decidability of the consistency problem.

Next to their appearance in applications, there might be other criteria to evaluate requirements, such as invariance under some family of operations (isomorphisms or embeddings of some sort, for instance): the larger the family the more natural appears to be the requirement. One may wish to restrict
to the more natural requirements, hoping that they are not contradictory. While this is a very relevant issue, we focus here on satisfiability as this is any case a preliminary issue from the mathematical point of view.

9 Appendix C: analysis of the contradictory setup

According to the discussion of Sections 3 (and the more abstract ones in Section 2), the setup described in Exercise 2.8 of [AT] proposes a finite polynomial probability environment, which requires admissibility. By Corollary such issue is decidable as follows.

First, there are three events involved, so that \( A = \{E_1, E_2, E_3\} \). Then select a variable for each of the boolean combinations on which conditions are given; it is convenient to use the following notation (which includes some variables which are not going to be used – the only variables used are those actually appearing in (9) below):

\[
x_{\beta_1, \beta_2, \beta_3} = P(E_1^{\beta_1} \cap E_2^{\beta_2} \cap E_3^{\beta_3})
\]

where \( \beta_m \in \{-1, 0, 1\} \) and \( A^{-1} = A^C, A^0 = \Omega, A^1 = A \). The equations expressing the requirements of the probability environment are the following, where we have assumed that the claimed independence is independence of the algebra generated by \( E_2 \) and \( E_3 \) with \( E_1 \) (as it makes sense that the weather is independent from any combination of human productions):

\[
\begin{align*}
    x_{1,0,0} &= 0.8 \\
    x_{0,1,0} &= 0.7 \\
    x_{0,0,1} &= 0.95 \\
    x_{1,1,1} &= x_{1,1,0} \cdot x_{0,0,1} \\
    x_{1,-1,1} &= x_{1,-1,0} \cdot x_{0,0,1} \\
    x_{-1,1,1} &= x_{-1,1,0} \cdot x_{0,0,1} \\
    x_{-1,-1,1} &= x_{-1,-1,0} \cdot x_{0,0,1} \\
    x_{0,-1,1} &= 0.6 \cdot x_{0,-1,0}
\end{align*}
\] (6)

With the trivial substitution \( x_{0,-1,0} = 1 - x_{0,1,0} \) there are 12 variables in the system.
Now make the change of variables

\[ x_{\beta_1, \beta_2, \beta_3} = \sum_{\alpha_m \in \{\beta_m + |\beta_m| - 1, \beta_m - |\beta_m| + 1\}, m=1,2,3} y_{\alpha_1, \alpha_2, \alpha_3} \]

where \( y_{\alpha_1, \alpha_2, \alpha_3} \) indicates the unknown probability of \( E_1^{\alpha_1} \cap E_2^{\alpha_2} \cap E_3^{\alpha_3} \). After substitution and inclusion of the conditions on the \( y_{\alpha_i} \)'s the system has 9 equations of degree either 1 or 2, and 8 inequations. There is no solution, indicating that the requirements are not admissible; a more explicit calculation is repeated below, and a certificate that the system has no solution is in Section 3.4.

One can actually wonder if it was the value 0.95 required for \( P(E_3) \) which created a problem. In fact, one can leave \( x_{0,0,1} \) as an indeterminate and solve the system in \( y_{\alpha_i} \) for the other 7 variables. The result, before imposing the condition that \( y_{\alpha_i} \geq 0 \) for all \( \alpha_i \)'s, is:

\[
\begin{align*}
    y_{-1,1,1} &= (21 - 50y_{1,1,1})/50 \\
    y_{1,-1,1} &= (12 - 25y_{1,1,1})/25 \\
    y_{1,1,-1} &= 2y_{1,1,1}/3 \\
    y_{-1,-1,1} &= (-3 + 10y_{1,1,1})/10 \\
    y_{-1,1,-1} &= (21 - 50y_{1,1,1})/75 \\
    y_{1,-1,-1} &= (24 - 50y_{1,1,1})/75 \\
    y_{-1,-1,-1} &= (-3 + 10y_{1,1,1})/15.
\end{align*}
\]

With the nonegativity condition one has \( P(E_1 \cap E_2 \cap E_3) = y_{1,1,1} \in [3/10, 21/50] \), and \( P(E_3) = y_{1,1,1} + y_{-1,1,1} + y_{1,-1,1} + y_{-1,-1,1} = 3/5 \). This amounts to the cylindrical decomposition according to Tarsky-Seidenberg reduction theorem, although it is better computed by directly solving the equations first.

In conclusion, only the value \( P(E_3) = 3/5 \) is admissible in this set up.

10 Appendix D: historical remarks

The issues we have discussed here constitute very likely the dilemma to which Hilbert was referring when he stated in his 1905 seminar that "at this stage of the development it is not clear yet which statements are definitions and which statements are axioms." (see [UK], [LC], [DH1], [DH2]) In fact, in the seminar Hilbert had introduced probabilities according to Bohlman axiomatics,
which has conditional probability as an axiom (see [UK, GB]), while in current formulations (and very likely in other attempts in Hilbert’s time) they are definitions. This is the point raised above: with statements like the assignment of a given value to a conditional probability as axioms, we run the risk of contradiction, but if we take them as definitions and verify them within some existing model they lose their calculative power and thus almost any practical purpose.

The same issues were raised by Boole in [Boole1854], [Boole1854a] and [Boole1857], see also [Hailperin] for a thorough analysis. The text is at time of controversial interpretation, but clearly Boole asks for the probability to be assigned to propositions instead of events, which is the same as saying that they are to be assigned to formal expressions rather then subsets of a given set; then he states his fundamental theorem of (finite) probability in which he claims to be able to solve “any” problem: by this Boole clearly means “any set of assignments” in a sense which must be close to the one we have given here. However, Boole states in his axiom VI that events with no relations to others should be taken to be independent (which is an unwarranted assumption) and also does not seem to be able to give a full proof of his fundamental theorem (see Hailperin). What we have given here in Section (4.2) is a proof of Boole fundamental theorem for a very broad class of problems (very likely even larger than he had in mind).

References

[AT] A. H-S. Ang, W. H. Tang, Probability concepts in engineering, Wiley ed., 2nd edition

[UK] U. Krengel, ON THE CONTRIBUTIONS OF GEORG BOHLMANN TO PROBABILITY THEORY, Electronic Journal for History of Probability and Statistics 7 No. 1, 1-13 (2011).

[GB] G. Bohlmann Lebensversicherungsmathematik, Encyklopädie der mathematischen Wissenschaften, Bd I, Teil 2, Artikel I D 4 b (1900), 852917

[Br] Brouwer

[LC] L. Corry, David Hilbert and the Axiomatization of Physics (18981918), Kluwer Academic Publishers, Dordrecht (2004)
[DH1] David Hilbert, Logische Prinzipien des Mathematischen Denkens, Manuskript von Hilberts Vorlesung von 1905, Bibliothek des Math.Inst. Univ. Gottingen, aufgezeichnet von E. Hellinger, 1905.

[DH2] David Hilbert, Logische Prinzipien des Mathematischen Denkens. Max Borns notes of Hilberts lectures of 1905, Handschriftenabteilung der Georg-August-Univ. Gottingen, Signatur: Cod Ms D. Hilbert 558a.

[HM] Hazewinkel, Michiel, ed. (2001), "Disjunctive normal form", Encyclopedia of Mathematics, Springer,

[Kh] Khrennikov, (1999) Interpretations of probability, VSR, Utrecht.

[Pa] Parrilo, Pablo A. "Sum of squares programs and polynomial inequalities." SIAG/OPT Views-and-News: A Forum for the SIAM Activity Group on Optimization. Vol. 15. No. 2. 2004.

[P] H. Poincare (1905), Les mathematiques et la logique. Revue de metaphysique et de morale, 815-835.

[BCR] Bochnak, Jacek; Coste, Michel; Roy, Marie-Francoise. Real Algebraic Geometry. Translated from the 1987 French original. Revised by the authors. Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], 36. Springer-Verlag, Berlin, 1998.

[TvD] ) Anne Sjerp Troelstra, Dirk van Dalen, "Constructivism in Mathematics: An Introduction, Volume 1", 1988.

[T] Tao, T., Topics in random matrix theory, American Mathematical Society, Providence, RI, 2011