Is there a probability theory of non-classical events?

Vincenzo Marra
Dipartimento di Informatica e Comunicazione
Università degli Studi di Milano
marra@dico.unimi.it

November 2\textsuperscript{nd}, 2009
International Workshop on Probability, Uncertainty, and Rationality
Certosa di Pontignano
Is there a probability theory of non-classical events?

Outline

1. Prologue
2. From logic to probability: a sketch
3. A many-valued logic
4. Many-valued probability?
Is there a probability theory of non-classical events?

Outline

1. Prologue

2. From logic to probability: a sketch

3. A many-valued logic

4. Many-valued probability?
A conversation I heard some time ago at a mathematics conference:
A conversation I heard some time ago at a mathematics conference:

First mathematician – A good talk should always contain at least one theorem and one joke.
A conversation I heard some time ago at a mathematics conference:

First mathematician – *A good talk should always contain at least one theorem and one joke.*

Second mathematician – *And you better make sure it is clear which is which.*
A conversation I heard some time ago at a mathematics conference:

First mathematician – *A good talk should always contain at least one theorem and one joke.*

Second mathematician – *And you better make sure it is clear which is which.*

There will be no theorem on many-valued probability in this talk, lest you might mistake it for a joke.
A conversation I heard some time ago at a mathematics conference:

First mathematician – A good talk should always contain at least one theorem and one joke.

Second mathematician – And you better make sure it is clear which is which.

There will be no theorem on many-valued probability in this talk, lest you might mistake it for a joke. In fact, I am really doubtful whether there even is such a thing as many-valued probability at all.
A conversation I heard some time ago at a mathematics conference:

First mathematician – *A good talk should always contain at least one theorem and one joke.*

Second mathematician – *And you better make sure it is clear which is which.*

There will be no theorem on many-valued probability in this talk, lest you might mistake it for a joke. In fact, I am really doubtful whether there even *is* such a thing as many-valued probability at all. So I guess we better start on the safe side.
Is there a probability theory of non-classical events?

Prologue

Historiographic cliché: Probability theory begins in 1654, with the correspondence between Fermat and Pascal on the *problem of points*, proposed to them by the Chevalier de Méré (born Antoine Gombaud).
Is there a probability theory of non-classical events?

Prologue

Fermat to Pascal:

“Monsieur,
If I undertake to make a point with a single die in eight throws, and if we agree after the money is put at stake, that I shall not cast the first throw, it is necessary by my theory that I take $\frac{1}{6}$ of the total sum to be impartial because of the aforesaid first throw. And if we agree after that that I shall not play the second throw, I should, for my share, take the sixth of the remainder that is $\frac{5}{36}$ of the total. If, after that, we agree that I shall not play the third throw, . . .”
Part of Pascal’s reply:

2. Votre méthode est très-sûre et est celle qui m’est la première venue à la pensée dans cette recherche; mais, parce que la peine des combinaisons est excessive, j’en ai trouvé un abrégé et proprement une autre méthode bien plus courte et plus nette, que je voudrois vous pouvoir dire ici en peu de mots; car je voudrois désormais vous ouvrir mon coeur, s’il se pouvoit, tant j’ai de joie de voir notre rencontre. Je vois bien que la vérité est la même à Toulouse et à Paris.
Part of Pascal’s reply:

2. Votre méthode est très-sûre et est celle qui m’est la première venue à la pensée dans cette recherche; mais, parce que la peine des combinaisons est excessive, j’en ai trouvé un abrégé et proprement une autre méthode bien plus courte et plus nette, que je voudrois vous pouvoir dire ici en peu de mots : car je voudrois désormais vous ouvrir mon cœur, s’il se pouvoit, tant j’ai de joie de voir notre rencontre. Je vois bien que la vérité est la même à Toulouse et à Paris.

“I see the truth is the same in Toulouse as in Paris.”
Is there a probability theory of non-classical events?

Part of Pascal’s reply:

2. Votre méthode est très-sûre et est celle qui m’est la première venue à la pensée dans cette recherche; mais, parce que la peine des combinaisons est excessive, j’en ai trouvé un abrégé et proprement une autre méthode bien plus courte et plus nette, que je voudrois vous pouvoir dire ici en peu de mots: car je voudrois désormais vous ouvrir mon cœur, s’il se pouvoit, tant j’ai de joie de voir notre rencontre. Je vois bien que la vérité est la même à Toulouse et à Paris.

“I see the truth is the same in Toulouse as in Paris.”

Apart from Pascal’s reference to truth, what does logic have to do with probability?
Outline

1 Prologue

2 From logic to probability: a sketch
   ■ The rôle of logic in probability theory
   ■ Recap on classical propositional logic
   ■ Probability assignments

3 A many-valued logic

4 Many-valued probability?
One key primitive notion in probability theory is that of event.
One key primitive notion in probability theory is that of event. An important tradition in the subject regards probability theory as an attempt to model the possible outcomes of idealised experiments.
One key primitive notion in probability theory is that of event. An important tradition in the subject regards probability theory as an attempt to model the possible outcomes of idealised experiments. Thus, in the letter from Fermat to Pascal quoted above, the experiment is a sequence of eight throws of a die (with faces numbered from 1 to 6).
One key primitive notion in probability theory is that of \textit{event}. An important tradition in the subject regards probability theory as an attempt to model the \textit{possible outcomes of idealised experiments}.

Thus, in the letter from Fermat to Pascal quoted above, the \textit{experiment} is a sequence of eight throws of a die (with faces numbered from 1 to 6).

The \textit{possible outcomes} of the experiment are all possible sequences of points in the eight throws; one such, e.g., is

\begin{equation*}
1, 1, 3, 6, 2, 4, 6, 4.
\end{equation*}
One key primitive notion in probability theory is that of event. An important tradition in the subject regards probability theory as an attempt to model the possible outcomes of idealised experiments. Thus, in the letter from Fermat to Pascal quoted above, the experiment is a sequence of eight throws of a die (with faces numbered from 1 to 6). The possible outcomes of the experiment are all possible sequences of points in the eight throws; one such, e.g., is

\[1, 1, 3, 6, 2, 4, 6, 4.\]

The set \(S\) of all possible outcomes is called the sample space.
One key primitive notion in probability theory is that of event. An important tradition in the subject regards probability theory as an attempt to model the possible outcomes of idealised experiments.

Thus, in the letter from Fermat to Pascal quoted above, the experiment is a sequence of eight throws of a die (with faces numbered from 1 to 6). The possible outcomes of the experiment are all possible sequences of points in the eight throws; one such, e.g., is $1, 1, 3, 6, 2, 4, 6, 4$.

The set $S$ of all possible outcomes is called the sample space. Certain subsets of $S$ (not necessarily all) are then selected as having special interest for the problem at hand; they form the collection $\mathcal{E}$ of events.
One key primitive notion in probability theory is that of event. An important tradition in the subject regards probability theory as an attempt to model the possible outcomes of idealised experiments.
Thus, in the letter from Fermat to Pascal quoted above, the experiment is a sequence of eight throws of a die (with faces numbered from 1 to 6). The possible outcomes of the experiment are all possible sequences of points in the eight throws; one such, e.g., is 1, 1, 3, 6, 2, 4, 6, 4.

The set $S$ of all possible outcomes is called the sample space. Certain subsets of $S$ (not necessarily all) are then selected as having special interest for the problem at hand; they form the collection $\mathcal{E}$ of events. (It is customary to ask that $\mathcal{E}$ satisfy some additional assumptions, and we will return to this point later.)
There is, however, a second approach to the notion of event that also has a substantial tradition.
There is, however, a second approach to the notion of event that also has a substantial tradition. It consists in taking propositions as the primitive notion, and in defining events as a derived notion.
There is, however, a second approach to the notion of event that also has a substantial tradition. It consists in taking propositions as the primitive notion, and in defining events as a derived notion.

An event is then whatever may be described by a proposition.
There is, however, a second approach to the notion of event that also has a substantial tradition. It consists in taking propositions as the primitive notion, and in defining events as a derived notion.

An event is then whatever may be described by a proposition.

Thus, returning to Fermat’s example, the set consisting of the two sequence of points

\[1, 1, 1, 1, 1, 1, 1, 1\quad\text{and}\quad6, 6, 6, 6, 6, 6, 6, 6\]

corresponds to an event, because it may be described by a proposition.
There is, however, a second approach to the notion of event that also has a substantial tradition. It consists in taking propositions as the primitive notion, and in defining events as a derived notion.

An event is then whatever may be described by a proposition.

Thus, returning to Fermat’s example, the set consisting of the two sequence of points

\[1, 1, 1, 1, 1, 1, 1, 1 \quad \text{and} \quad 6, 6, 6, 6, 6, 6, 6, 6\]

corresponds to an event, because it may be described by a proposition. Say,

“Either one observes, as the outcome of the experiment, the smallest possible point at each throw, or else one observes the largest possible point at each throw.”
Boole on events vs. propositions:

6. Before we proceed to estimate to what extent known methods may be applied to the solution of problems such as the above, it will be advantageous to notice, that there is another form under which all questions in the theory of probabilities may be viewed; and this form consists in substituting for \textit{events} the propositions which assert that those events have occurred, or will occur; and viewing the element of numerical probability as having reference to the \textit{truth} of those \textit{propositions}, not to the oc-

George Boole, \textit{An Investigation of The Laws of Thought}, pp. 247–248, Dublin 1854
Is there a probability theory of non-classical events?

From logic to probability: a sketch

The rôle of logic in probability theory

Keynes on events vs. propositions:

CH. I

FUNDAMENTAL IDEAS

4. With the term “event,” which has taken hitherto so important a place in the phraseology of the subject, I shall dispense altogether.\(^1\) Writers on Probability have generally dealt with what they term the “happening” of “events.” In the problems which they first studied this did not involve much departure from common usage. But these expressions are now used in a way which is vague and ambiguous; and it will be more than a verbal improvement to discuss the truth and the probability of propositions instead of the occurrence and the probability of events.\(^2\)

Boole remarks that when using propositions in place of (some mathematical model of) events, probabilities will refer to the truth of propositions, rather than to the occurrence of the event.
Boole remarks that when using propositions in place of (some mathematical model of) events, probabilities will refer to the truth of propositions, rather than to the occurrence of the event.

Keynes openly states that this is “more than a verbal improvement”. How is it so?
Boole remarks that when using propositions in place of (some mathematical model of) events, probabilities will refer to the truth of propositions, rather than to the occurrence of the event.

Keynes openly states that this is "more than a verbal improvement". How is it so?

Well, one advantage of the approach based on propositions is the following. Suppose we were concerned with the questions: What are events? And how does one observe the occurrence of an event?
Boole remarks that when using propositions in place of (some mathematical model of) events, probabilities will refer to the truth of propositions, rather than to the occurrence of the event.

Keynes openly states that this is “more than a verbal improvement”. How is it so?

Well, one advantage of the approach based on propositions is the following. Suppose we were concerned with the questions: What are events? And how does one observe the occurrence of an event? Now we would then be concerned with the new questions: what are propositions? And how does one verify the truth of a proposition?
This shift of perspective is more than a verbal improvement precisely in that our new questions have nothing to do with the theory of probability *per se*. 
This shift of perspective is more than a verbal improvement precisely in that our new questions have nothing to do with the theory of probability per se.

We are now in the realm of logic, and this might give us a safer vantage point – after all, logic is about 4 times older than probability theory.
This shift of perspective is more than a verbal improvement precisely in that our new questions have nothing to do with the theory of probability per se.

We are now in the realm of logic, and this might give us a safer vantage point – after all, logic is about 4 times older than probability theory.

There indeed are widely accepted answers provided by logicians to the questions above, and to these we now turn.
This shift of perspective is more than a verbal improvement precisely in that our new questions have nothing to do with the theory of probability *per se*.

We are now in the realm of logic, and this might give us a safer vantage point – after all, logic is about 4 times older than probability theory.

There indeed are widely accepted answers provided by logicians to the questions above, and to these we now turn.

To summarise so far:

*The rôle of logic in the theory of probability is to provide a formal model for the notion of event.*
This shift of perspective is more than a verbal improvement precisely in that our new questions have nothing to do with the theory of probability per se.

We are now in the realm of logic, and this might give us a safer vantage point – after all, logic is about 4 times older than probability theory.

There indeed are widely accepted answers provided by logicians to the questions above, and to these we now turn.

To summarise so far:

*The rôle of logic in the theory of probability is to provide a formal model for the notion of event.*

*Added in Proof.* There would be much to say about a comparison between the approach based on propositions and the one based on events. If I have time I will return to this by way of an Appendix to the main talk.
Is there a probability theory of non-classical events?

From logic to probability: a sketch

Recap on classical propositional logic

Taking stock of available knowledge about the logical analysis of propositions, we use classical propositional logic without further ado as a formal means of modeling events.
Taking stock of available knowledge about the logical analysis of propositions, we use classical propositional logic without further ado as a formal means of modeling events. We start with a (finite or infinite) set of propositional variables, or atomic formulæ, that are to stand for propositions. Say, if we content ourselves with countably many:

\[ X_1, X_2, \ldots, X_n, \ldots \]
Taking stock of available knowledge about the logical analysis of propositions, we use classical propositional logic without further ado as a formal means of modeling events. We start with a (finite or infinite) set of propositional variables, or atomic formulæ, that are to stand for propositions. Say, if we content ourselves with countably many:

\[ X_1, X_2, \ldots, X_n, \ldots \].

To these we adjoin two symbols \( \top \) and \( \bot \), say, that are to stand for a proposition that is always true (the verum), and one that is always false (the falsum), respectively.
Taking stock of available knowledge about the logical analysis of propositions, we use **classical propositional logic** without further ado as a formal means of modeling events. We start with a (finite or infinite) set of **propositional variables**, or **atomic formulæ**, that are to stand for propositions. Say, if we content ourselves with countably many:

\[ X_1, X_2, \ldots, X_n, \ldots. \]

To these we adjoin two symbols \( \top \) and \( \bot \), say, that are to stand for a proposition that is always true (the **verum**), and one that is always false (the **falsum**), respectively. To construct compound formulæ we use the **logical connectives**:

- \( \lor \), for **disjunction**;
- \( \land \), for **conjunction**;
- \( \rightarrow \), for **implication**; and
- \( \neg \), for **negation**.
The usual recursive definition of general formulæ now reads as follows.
The usual recursive definition of general formulæ now reads as follows.

- \( \top \) and \( \bot \) are formulæ.
- All propositional variables are formulæ.
- If \( \alpha \) and \( \beta \) are formulæ, so are \( (\alpha \lor \beta) \), \( (\alpha \land \beta) \), \( (\alpha \rightarrow \beta) \), and \( \neg \alpha \).
The usual recursive definition of general formulæ now reads as follows.

- \( \top \) and \( \bot \) are formulæ.
- All propositional variables are formulæ.
- If \( \alpha \) and \( \beta \) are formulæ, so are \( (\alpha \lor \beta) \), \( (\alpha \land \beta) \), \( (\alpha \rightarrow \beta) \), and \( \neg \alpha \).

Let us write \( \text{FORM} \) for the set of all formulæ constructed over the countable language \( X_1, \ldots, X_n, \ldots \), and \( \text{FORM}_n \) for the set of formulæ constructed over the first \( n \) propositional variables \( X_1, \ldots, X_n \), at the most.
As usual, we construct a formal semantics in order to interpret formulæ.
As usual, we construct a formal semantics in order to interpret formulæ. Namely, we consider assignments of truth-values, or evaluations, or interpretations.
As usual, we construct a formal semantics in order to interpret formulæ. Namely, we consider assignments of truth-values, or evaluations, or interpretations. These are functions

\[ w : \text{FORM} \rightarrow \{0, 1\} \]

subject to the usual conditions:
As usual, we construct a formal semantics in order to interpret formulæ. Namely, we consider assignments of truth-values, or evaluations, or interpretations. These are functions

$$w : \text{FORM} \rightarrow \{0, 1\}$$

subject to the usual conditions:

- $w(\top) = 1, w(\bot) = 0$;
As usual, we construct a formal semantics in order to interpret formulæ. Namely, we consider assignments of truth-values, or evaluations, or interpretations. These are functions

\[ w: \text{FORM} \rightarrow \{0, 1\} \]

subject to the usual conditions:

- \[ w(\top) = 1, w(\bot) = 0; \]
- \[ w(\alpha \land \beta) = 1 \text{ if and only if both } w(\alpha) = 1 \text{ and } w(\beta) = 1; \]
As usual, we construct a formal semantics in order to interpret formulæ. Namely, we consider assignments of truth-values, or evaluations, or interpretations. These are functions

\[ w : \text{FORM} \to \{0, 1\} \]

subject to the usual conditions:

- \( w(\top) = 1, w(\bot) = 0 \);
- \( w(\alpha \land \beta) = 1 \) if and only if both \( w(\alpha) = 1 \) and \( w(\beta) = 1 \);
- \( w(\alpha \lor \beta) = 1 \) if and only if either \( w(\alpha) = 1 \) or \( w(\beta) = 1 \) (or both);
As usual, we construct a formal semantics in order to interpret formulæ. Namely, we consider assignments of truth-values, or evaluations, or interpretations. These are functions

$$w: \text{FORM} \to \{0, 1\}$$

subject to the usual conditions:

- $w(\top) = 1, w(\bot) = 0$;
- $w(\alpha \land \beta) = 1$ if and only if both $w(\alpha) = 1$ and $w(\beta) = 1$;
- $w(\alpha \lor \beta) = 1$ if and only if either $w(\alpha) = 1$ or $w(\beta) = 1$ (or both);
- $w(\alpha \to \beta) = 0$ if and only if $w(\alpha) = 0$ and $w(\beta) = 1$;
As usual, we construct a formal semantics in order to interpret formulæ.
Namely, we consider assignments of truth-values, or evaluations, or interpretations.
These are functions

\[ w : \text{FORM} \rightarrow \{0, 1\} \]

subject to the usual conditions:

- \( w(\top) = 1, w(\bot) = 0; \)
- \( w(\alpha \land \beta) = 1 \) if and only if both \( w(\alpha) = 1 \) and \( w(\beta) = 1; \)
- \( w(\alpha \lor \beta) = 1 \) if and only if either \( w(\alpha) = 1 \) or \( w(\beta) = 1 \) (or both);
- \( w(\alpha \rightarrow \beta) = 0 \) if and only if \( w(\alpha) = 0 \) and \( w(\beta) = 1; \)
- \( w(\neg \alpha) = 1 \) if and only if \( w(\alpha) = 0. \)
A first easy observation about these definitions is that a valuation \( w \) is subject to no restrictions concerning the values it assigns to propositional variables.
A first easy observation about these definitions is that a valuation $w$ is subject to no restrictions concerning the values it assigns to propositional variables. Moreover, it is also apparent that any valuation $w$ is uniquely determined by such values $w(X_1), \ldots, w(X_n), \ldots$. 
A first easy observation about these definitions is that a valuation \( w \) is subject to no restrictions concerning the values it assigns to propositional variables. Moreover, it is also apparent that any valuation \( w \) is uniquely determined by such values \( w(X_1), \ldots, w(X_n), \ldots \). This is the principle of truth-functionality (after G. Frege) for classical propositional logic.
A first easy observation about these definitions is that a valuation *w is subject to no restrictions concerning the values it assigns to propositional variables*. Moreover, it is also apparent that *any valuation w is uniquely determined by such values w(X₁),...,w(Xₙ),....* This is the principle of truth-functionality (after G. Frege) for classical propositional logic. Following a well-established tradition going back at least to G. W. Leibniz, it is useful to think of an evaluation *w: FORM → {0, 1}* as a possible world in which each proposition α either holds – i.e., *w(α) = 1* – or fails to hold – i.e., *w(α) = 0*. 
A first easy observation about these definitions is that a valuation \( w \) is subject to no restrictions concerning the values it assigns to propositional variables. Moreover, it is also apparent that any valuation \( w \) is uniquely determined by such values \( w(X_1), \ldots, w(X_n), \ldots \).

This is the principle of truth-functionality (after G. Frege) for classical propositional logic. Following a well-established tradition going back at least to G. W. Leibniz, it is useful to think of an evaluation \( w : \text{FORM} \to \{0, 1\} \) as a possible world in which each proposition \( \alpha \) either holds – i.e., \( w(\alpha) = 1 \) – or fails to hold – i.e., \( w(\alpha) = 0 \).

Because of truth-functionality, then, a classical logician’s world is something very ethereal indeed: it is uniquely determined by the collection of true atomic propositions that can be uttered.
Analytic truths, or **tautologies** after L. Wittgenstein, are now defined as those formulæ \( \alpha \in \text{FORM} \) that are *true in every possible world*, i.e. such that \( w(\alpha) = 1 \) for any assignment \( w \).
Analytic truths, or tautologies after L. Wittgenstein, are now defined as those formulæ $\alpha \in \text{FORM}$ that are true in every possible world, i.e. such that $w(\alpha) = 1$ for any assignment $w$. Tautologies are of central interest in logic proper, but not in probability theory – they are just linguistic descriptions of the sure event, an event that always occurs. Nonetheless, let us recall why they are important to logic, omitting details.
Is there a probability theory of non-classical events?

Recap on classical propositional logic

From logic to probability: a sketch

Analytic truths, or tautologies after L. Wittgenstein, are now defined as those formulæ $\alpha \in \text{FORM}$ that are true in every possible world, i.e. such that $w(\alpha) = 1$ for any assignment $w$.

Tautologies are of central interest in logic proper, but not in probability theory – they are just linguistic descriptions of the sure event, an event that always occurs. Nonetheless, let us recall why they are important to logic, omitting details.

Selecting an appropriate collection of formulæ as axioms, and using modus ponens as deduction rule, one defines the notion of provable formula. We do not need to recall details here.
Analytic truths, or **tautologies** after L. Wittgenstein, are now defined as those formulæ $\alpha \in \text{FORM}$ that are *true in every possible world*, i.e. such that $w(\alpha) = 1$ for any assignment $w$.

Tautologies are of central interest in logic proper, but not in probability theory – they are just linguistic descriptions of the **sure event**, an event that always occurs. Nonetheless, let us recall why they are important to logic, omitting details.

Selecting an appropriate collection of formulæ as **axioms**, and using **modus ponens** as deduction rule, one defines the notion of **provable formula**. We do not need to recall details here. Writing $\vdash \alpha$ to mean that a formula $\alpha$ is provable, and writing $\models \alpha$ to mean that $\alpha$ is a tautology, we can now state the all-important (soundness and) **completeness theorem**: for any $\alpha \in \text{FORM}$,

$$\vdash \alpha \quad \text{if and only if} \quad \models \alpha .$$
Finally, let me recall that in classical logic there also is a stronger version of completeness than the above one.
Finally, let me recall that in classical logic there also is a stronger version of completeness than the above one. By a *theory* $\Theta$ one means a deductively closed set of formulæ.
Finally, let me recall that in classical logic there also is a stronger version of completeness than the above one. By a theory $\Theta$ one means a deductively closed set of formulæ. One generalises syntactic and semantic consequence to theories.
Finally, let me recall that in classical logic there also is a stronger version of completeness than the above one. By a theory $\Theta$ one means a deductively closed set of formulæ. One generalises syntactic and semantic consequence to theories. Thus, for $\alpha \in \text{FORM}$,

$$\Theta \vdash \alpha$$

means that $\alpha$ is deducible from $\Theta$ via *modus ponens*, whereas

$$\Theta \models \alpha$$

means that $\alpha$ is true in each possible world wherein each formula of $\Theta$ is true.
Finally, let me recall that in classical logic there also is a stronger version of completeness than the above one. By a **theory** \( \Theta \) one means a deductively closed set of formulæ. One generalises syntactic and semantic consequence to theories. Thus, for \( \alpha \in \text{FORM} \),

\[ \Theta \vdash \alpha \]

means that \( \alpha \) is deducible from \( \Theta \) via *modus ponens*, whereas

\[ \Theta \models \alpha \]

means that \( \alpha \) is true in each possible world wherein each formula of \( \Theta \) is true.

The **completeness theorem** for theories now reads: for any \( \alpha \in \text{FORM} \) and any theory \( \Theta \),

\[ \Theta \vdash \alpha \quad \text{if and only if} \quad \Theta \models \alpha . \]
Having summarised our notation for classical logic, let us move on to probabilities.
Having summarised our notation for classical logic, let us move on to probabilities.

Fix a theory $\Theta$. A **probability assignment** (relative to $\Theta$) is a function

$$P : \text{FORM} \rightarrow \mathbb{R}$$

that satisfies Kolmogorov’s axioms:
Having summarised our notation for classical logic, let us move on to probabilities.

Fix a theory $\Theta$. A **probability assignment** (relative to $\Theta$) is a function

$$P : \text{FORM} \rightarrow \mathbb{R}$$

that satisfies Kolmogorov’s axioms:

$$(K0) \quad P(\alpha) = P(\beta) \text{ whenever } \Theta \vdash (\alpha \rightarrow \beta) \land (\beta \rightarrow \alpha).$$
Having summarised our notation for classical logic, let us move on to probabilities.

Fix a theory $\Theta$. A **probability assignment** (relative to $\Theta$) is a function

$$ P : \text{FORM} \to \mathbb{R} $$

that satisfies Kolmogorov’s axioms:

(K0) $P(\alpha) = P(\beta)$ whenever $\Theta \vdash (\alpha \to \beta) \land (\beta \to \alpha)$.

(K1) $P(\bot) = 0$ and $P(\top) = 1$. 
Having summarised our notation for classical logic, let us move on to probabilities.

Fix a theory $\Theta$. A **probability assignment** (relative to $\Theta$) is a function

$$P : \text{FORM} \to \mathbb{R}$$

that satisfies Kolmogorov’s axioms:

(K0) $P(\alpha) = P(\beta)$ whenever $\Theta \vdash (\alpha \rightarrow \beta) \land (\beta \rightarrow \alpha)$.

(K1) $P(\bot) = 0$ and $P(\top) = 1$.

(K2) $P(\alpha) \leq P(\beta)$ whenever $\alpha \rightarrow \beta$ holds.
Having summarised our notation for classical logic, let us move on to probabilities.

Fix a theory $\Theta$. A **probability assignment** (relative to $\Theta$) is a function

$$P : \text{FORM} \to \mathbb{R}$$

that satisfies Kolmogorov’s axioms:

**K0** $P(\alpha) = P(\beta)$ whenever $\Theta \vdash (\alpha \rightarrow \beta) \land (\beta \rightarrow \alpha)$.

**K1** $P(\bot) = 0$ and $P(\top) = 1$.

**K2** $P(\alpha) \leq P(\beta)$ whenever $\alpha \rightarrow \beta$ holds.

**K3** $P(\alpha \lor \beta) + P(\alpha \land \beta) = P(\alpha) + P(\beta)$.
It should be stressed that these axioms, as is the case for all well-established mathematical notions, did not come out of the blue.
It should be stressed that these axioms, as is the case for all well-established mathematical notions, did not come out of the blue. When Kolmogorov’s celebrated *Grundbegriffe der Wahrscheinlichkeitsrechnung* appeared in 1933, a huge amount of interesting mathematics in measure and probability theory was already available. And it would be ludicrous even to suggest that the importance of Kolmogorov’s work in the field amounts to his axiomatic definition.
It should be stressed that these axioms, as is the case for all well-established mathematical notions, did not come out of the blue. When Kolmogorov’s celebrated *Grundbegriffe der Wahrscheinlichkeitsrechnung* appeared in 1933, a huge amount of interesting mathematics in measure and probability theory was already available. And it would be ludicrous even to suggest that the importance of Kolmogorov’s work in the field amounts to his axiomatic definition. So from certain points of view – the conceptual development of mathematics, for instance – it is entirely appropriate to hold that this definition was already fully justified in 1933.
It should be stressed that these axioms, as is the case for all well-established mathematical notions, did not come out of the blue. When Kolmogorov’s celebrated *Grundbegriffe der Wahrscheinlichkeitsrechnung* appeared in 1933, a huge amount of interesting mathematics in measure and probability theory was already available. And it would be ludicrous even to suggest that the importance of Kolmogorov’s work in the field amounts to his axiomatic definition.

So from certain points of view – the conceptual development of mathematics, for instance – it is entirely appropriate to hold that this definition was already fully justified in 1933. It is equally true, nonetheless, that from different perspectives it is reasonable – even necessary – to ask for a different sort of justification.
Is there a probability theory of non-classical events?

From logic to probability: a sketch

Probability assignments

The literature in this direction is abundant. To single out one example, let me mention the Ramsey-de Finetti Dutch book argument (1926, 1937), along with its later utility-based version by L. Savage (1954), that featured in other talks. Of the various justifications of probability theory I will say no more here.
The literature in this direction is abundant. To single out one example, let me mention the Ramsey-de Finetti Dutch book argument (1926, 1937), along with its later utility-based version by L. Savage (1954), that featured in other talks. Of the various justifications of probability theory I will say no more here.
Outline

1. Prologue
2. From logic to probability: a sketch
3. A many-valued logic
4. Many-valued probability?
Let us now consider non-classical logic. I will look at a single example, a many-valued logic known as Gödel propositional logic.
Let us now consider non-classical logic. I will look at a single example, a many-valued logic known as Gödel propositional logic. This is part of a hierarchy of many-valued logics that includes e.g. Łukasiewicz logic, but I will not discuss the larger picture here.
Let us now consider non-classical logic. I will look at a single example, a many-valued logic known as Gödel propositional logic. This is part of a hierarchy of many-valued logics that includes e.g. Łukasiewicz logic, but I will not discuss the larger picture here. For the syntax, we use exactly the same notation as for classical logic. Concerning semantics, truth-value assignments

\[ w : \text{FORM} \to [0, 1] \]

now take values in the real unit-interval \([0, 1] \subseteq \mathbb{R}\), and are subject to the following conditions for any formulæ \(\alpha\) and \(\beta\).
Is there a probability theory of non-classical events?

A many-valued logic

\[ w(\bot) = 0, \quad w(\top) = 1. \]
Is there a probability theory of non-classical events?

A many-valued logic

- \( w(\bot) = 0, \quad w(\top) = 1. \)
- \( w(\alpha \land \beta) = \min(w(\alpha), w(\beta)). \)
Is there a probability theory of non-classical events?

A many-valued logic

- $w(\perp) = 0$, $w(\top) = 1$.
- $w(\alpha \land \beta) = \min (w(\alpha), w(\beta))$.
- $w(\alpha \lor \beta) = \max (w(\alpha), w(\beta))$.

A tautology is again a formula true in any possible world, i.e. evaluated to 1 by any assignment.
Is there a probability theory of non-classical events?

A many-valued logic

- \( w(\bot) = 0, \ w(\top) = 1 \).
- \( w(\alpha \land \beta) = \min (w(\alpha), w(\beta)) \).
- \( w(\alpha \lor \beta) = \max (w(\alpha), w(\beta)) \).
- \( w(\alpha \rightarrow \beta) = \begin{cases} 1 & \text{if } w(\alpha) \leq w(\beta) \\ w(\beta) & \text{otherwise} \end{cases} \)
Is there a probability theory of non-classical events?

A many-valued logic

- $w(\bot) = 0$, $w(\top) = 1$.
- $w(\alpha \land \beta) = \min (w(\alpha), w(\beta))$.
- $w(\alpha \lor \beta) = \max (w(\alpha), w(\beta))$.
- $w(\alpha \rightarrow \beta) = \begin{cases} 1 & \text{if } w(\alpha) \leq w(\beta) \\ w(\beta) & \text{otherwise} \end{cases}$
- $w(\neg \alpha) = \begin{cases} 1 & \text{if } w(\alpha) = 0 \\ 0 & \text{if } w(\alpha) > 0 \end{cases}$

A tautology is again a formula true in any possible world, i.e. evaluated to $1$ by any assignment.
Is there a probability theory of non-classical events?

---

A many-valued logic

- $w(\bot) = 0$, $w(\top) = 1$.
- $w(\alpha \land \beta) = \min (w(\alpha), w(\beta))$.
- $w(\alpha \lor \beta) = \max (w(\alpha), w(\beta))$.
- $w(\alpha \rightarrow \beta) = \begin{cases} 1 & \text{if } w(\alpha) \leq w(\beta) \\ w(\beta) & \text{otherwise} \end{cases}$
- $w(\neg \alpha) = \begin{cases} 1 & \text{if } w(\alpha) = 0 \\ 0 & \text{if } w(\alpha) > 0 \end{cases}$

A tautology is again a formula true in any possible world, i.e. evaluated to 1 by any assignment.
For an appropriate axiomatisation with *modus ponens* as deduction rule, one obtains completeness with respect to this many-valued semantics.
For an appropriate axiomatisation with *modus ponens* as deduction rule, one obtains completeness with respect to this many-valued semantics. Gödel logic can in fact be axiomatized as the schematic extension of the intuitionistic propositional calculus by the *prelinearity axiom* 

$$(\alpha \rightarrow \beta) \lor (\beta \rightarrow \alpha).$$
For an appropriate axiomatisation with *modus ponens* as deduction rule, one obtains completeness with respect to this many-valued semantics.

Gödel logic can in fact be axiomatized as the schematic extension of the intuitionistic propositional calculus by the *prelinearity axiom*

\[(\alpha \rightarrow \beta) \vee (\beta \rightarrow \alpha) .\]

Actually, Gödel logic is, so to speak, what you get taking the intersection of (schematic extensions of) intuitionistic logic with (schematic extensions of) many-valued logic in the sense of Petr Hájek.
Is there a probability theory of non-classical events?

A many-valued logic

**Figure:** All one-variable formulæ in Gödel logic, up to logical equivalence, ordered by $\alpha \leq \beta$ if and only if $\vdash \alpha \rightarrow \beta$. 
Outline

1 Prologue

2 From logic to probability: a sketch

3 A many-valued logic

4 Many-valued probability?
To help us proceed, forgive the bad form of quoting from the abstract of my own talk:
To help us proceed, forgive the bad form of quoting from the abstract of my own talk:

Since Boolean events may be denoted by formulæ in classical propositional logic, by analogy one is tempted to regard formulæ in a non-classical propositional logic as linguistic descriptions of non-classical events – whatever the latter are. And since Boolean events may be assigned (subjective) probabilities, one is lured into attempts at defining probabilities of non-classical events.
To help us proceed, forgive the bad form of quoting from the abstract of my own talk:

*Since Boolean events may be denoted by formulæ in classical propositional logic, by analogy one is tempted to regard formulæ in a non-classical propositional logic as linguistic descriptions of non-classical events – whatever the latter are. And since Boolean events may be assigned (subjective) probabilities, one is lured into attempts at defining probabilities of non-classical events.*

So each formula in the previous picture should denote one of the six “Gödelian events” that can be described with one propositional variable $X$. 
To help us proceed, forgive the bad form of quoting from the abstract of my own talk:

> Since Boolean events may be denoted by formulæ in classical propositional logic, by analogy one is tempted to regard formulæ in a non-classical propositional logic as linguistic descriptions of non-classical events – whatever the latter are. And since Boolean events may be assigned (subjective) probabilities, one is lured into attempts at defining probabilities of non-classical events.

So each formula in the previous picture should denote one of the six “Gödelian events” that can be described with one propositional variable $X$.

Unfortunately, one can have published several papers on Gödel logic, and still have no clue at all as to what such an event might look like.
To help us proceed, forgive the bad form of quoting from the abstract of my own talk:

Since Boolean events may be denoted by formulæ in classical propositional logic, by analogy one is tempted to regard formulæ in a non-classical propositional logic as linguistic descriptions of non-classical events – whatever the latter are. And since Boolean events may be assigned (subjective) probabilities, one is lured into attempts at defining probabilities of non-classical events.

So each formula in the previous picture should denote one of the six “Gödelian events” that can be described with one propositional variable $X$. Unfortunately, one can have published several papers on Gödel logic, and still have no clue at all as to what such an event might look like. Compare: everybody will start talking about tossing coins as soon as you ask them about classical events, even if they never even heard of classical logic.
Is there a probability theory of non-classical events?

Many-valued probability?

With some amount of hindsight, let’s see what we can do.
With some amount of hindsight, let’s see what we can do. Many-valued logics are supposed to provide a formal model of vague statements.
With some amount of hindsight, let’s see what we can do. Many-valued logics are supposed to provide a formal model of vague statements.

“The coin lands heads” is non-example.
With some amount of hindsight, let’s see what we can do. Many-valued logics are supposed to provide a formal model of vague statements. “The coin lands heads” is non-example. “It rains” might work better. Of course, we can always regard this simply as a classical proposition.
With some amount of hindsight, let’s see what we can do. Many-valued logics are supposed to provide a formal model of vague statements.

“The coin lands heads” is non-example. “It rains” might work better. Of course, we can always regard this simply as a classical proposition. However, say we now select one observable feature of the classical event “It rains”; say, the amount of rainfall during a whole day.
With some amount of hindsight, let’s see what we can do. Many-valued logics are supposed to provide a formal model of vague statements. “The coin lands heads” is non-example. “It rains” might work better. Of course, we can always regard this simply as a classical proposition. However, say we now select one observable feature of the classical event “It rains”; say, the amount of rainfall during a whole day. Then we might agree to interpret a sentence such as “It rains a lot” as a many-valued proposition with respect to that observable feature: namely, we agree that this proposition is the more true, the more it rains.
Is there a probability theory of non-classical events?

Many-valued probability?

So we now have two different propositions.
So we now have two different propositions. First, we can utter

\[ X = \text{“It rains a lot”} \]

as a many-valued proposition with respect to the amount of rainfall.
So we now have two different propositions. First, we can utter

$$X = \text{“It rains a lot”}$$

as a many-valued proposition with respect to the amount of rainfall. Second, we can utter

$$\neg \neg X = \text{“It rains”}$$

as a classical proposition: true if it rains, however little or much, and false otherwise.
So we now have two different propositions. First, we can utter

\[ X = \text{“It rains a lot”} \]

as a many-valued proposition with respect to the amount of rainfall.
Second, we can utter

\[ \neg\neg X = \text{“It rains”} \]

as a classical proposition: true if it rains, however little or much, and false otherwise.
Finally, we also consider the proposition

\[ \neg X = \text{“It does not rain”}, \]

and we agree to interpret this as a classical proposition, the complement of \( \neg\neg X \) – we are not interested in degrees of non-rain, in this example.
Is there a probability theory of non-classical events?

Many-valued probability?

Figure: All one-variable formulæ in Gödel logic, up to logical equivalence, ordered by $\alpha \leq \beta$ if and only if $\vdash \alpha \rightarrow \beta$. 
What are the logical relations that we intuitively expect between these three propositions, given the assumptions above?
What are the logical relations that we intuitively expect between these three propositions, given the assumptions above? Nothing to say a about \( \neg
eg X \) and \( \neg X \) – they are the usual Boolean alternative, it rains vs. it does not rain.
What are the logical relations that we intuitively expect between these three propositions, given the assumptions above? Nothing to say a about \( \neg
eg X \) and \( \neg X \) – they are the usual Boolean alternative, it rains vs. it does not rain. What about \( X \) (lots of rain)?
What are the logical relations that we intuitively expect between these three propositions, given the assumptions above? Nothing to say a about \( \neg\neg X \) and \( \neg X \) – they are the usual Boolean alternative, it rains vs. it does not rain. What about \( X \) (lots of rain)? The implication

\[ X \rightarrow \neg\neg X \]

must be true: whenever it rains a lot, then surely it must rain.
What are the logical relations that we **intuitively expect** between these three propositions, given the assumptions above? Nothing to say at all about $\neg\neg X$ and $\neg X$ — they are the usual Boolean alternative, it rains vs. it does not rain. What about $X$ (lots of rain)?

The implication

$$X \rightarrow \neg\neg X$$

must be **true**: whenever it rains a lot, then surely it must rain. But the converse implication is **not necessarily true**: it may well rain, without raining a lot.
What are the logical relations that we intuitively expect between these three propositions, given the assumptions above? Nothing to say a about \( \neg\neg X \) and \( \neg X \) – they are the usual Boolean alternative, it rains vs. it does not rain. What about \( X \) (lots of rain)?

The implication

\[
X \rightarrow \neg\neg X
\]

must be true: whenever it rains a lot, then surely it must rain. But the converse implication is not necessarily true: it may well rain, without raining a lot.

These properties actually are the ones that hold in Gödel logic; so far, so good.
Now that we understand the logic of these three propositions, let us go one step further.
Now that we understand the logic of these three propositions, let us go one step further. How would we attach probabilities to the three (events denoted by the three) propositions?
Now that we understand the logic of these three propositions, let us go one step further.
How would we attach probabilities to the three (events denoted by the three) propositions?
Let us write $p_1$ and $p_2$ for the probabilities of $\neg\neg X$ and $\neg X$, respectively, and $p_1^\top$ for the probability of $X$. Let us assume these numbers lie in $[0, 1]$. 
Now that we understand the logic of these three propositions, let us go one step further. How would we attach probabilities to the three (events denoted by the three) propositions? Let us write $p_1$ and $p_2$ for the probabilities of $\neg\neg X$ and $\neg X$, respectively, and $p_1^1$ for the probability of $X$. Let us assume these numbers lie in $[0, 1]$. Again, nothing special to say about $p_1$ and $p_2$: we must have $p_1 + p_2 = 1$. 
Now that we understand the logic of these three propositions, let us go one step further. How would we attach probabilities to the three (events denoted by the three) propositions? Let us write \( p_1 \) and \( p_2 \) for the probabilities of \( \neg \neg X \) and \( \neg X \), respectively, and \( p_1^1 \) for the probability of \( X \). Let us assume these numbers lie in \([0, 1]\). Again, nothing special to say about \( p_1 \) and \( p_2 \): we must have \( p_1 + p_2 = 1 \). What about \( p_1^1 \)? Well, because of the implication \( X \rightarrow \neg \neg X \), we must have

\[
p_1^1 \leq p_2.
\]
Now that we understand the logic of these three propositions, let us go one step further.
How would we attach probabilities to the three (events denoted by the three) propositions?
Let us write $p_1$ and $p_2$ for the probabilities of $\neg\neg X$ and $\neg X$, respectively, and $p_1^1$ for the probability of $X$. Let us assume these numbers lie in $[0, 1]$.
Again, nothing special to say about $p_1$ and $p_2$: we must have $p_1 + p_2 = 1$.
What about $p_1^1$? Well, because of the implication $X \rightarrow \neg\neg X$, we must have

$$p_1^1 \leq p_2.$$

But now an interesting thing happens: there are two options.
**Option 1.** On the one hand, we can just ask

\[(p_1 - p_1^1) + p_1^1 + p_2 = 1,\]

that is, we treat the three sentences:

"It rains, but not a lot", "It rains a lot", "It does not rain"

as three classical, mutually exclusive cases.
Option 1. On the one hand, we can just ask

\[(p_1 - p_1^1) + p_1^1 + p_2 = 1,\]

that is, we treat the three sentences:

“*It rains, but not a lot*”, “*It rains a lot*”, “*It does not rain*”

as three classical, mutually exclusive cases. It turns out that this corresponds to axiomatising probability assignments to formulæ in Gödel logic by precisely Kolmogorov’s axioms shown above.
**Option 1.** On the one hand, we can just ask

\[(p_1 - p_1^1) + p_1^1 + p_2 = 1,\]

that is, we treat the three sentences:

“It rains, but not a lot”, “It rains a lot”, “It does not rain”

as three classical, mutually exclusive cases. It turns out that this corresponds to axiomatising probability assignments to formulæ in Gödel logic by precisely Kolmogorov’s axioms shown above. But there is a problem here: there is no real connection between Gödel logic and this sort of probability assignments. In particular, you can replicate this probability distribution at once in the classical case – so why on earth use a non-classical logic?
Option 2. We ask that

\[(p_1 - p_1^1) + p_1^1 + p_2 = 1\], but with \(p_1 = 0\) whenever \(p_1^1 = 0\),

that is, we treat the three sentences:

“*It rains*”, “*It rains a lot*”, “*It does not rain*”

as if it is never completely false that whenever it rains, it rains a lot – which is actually the case in Gödel logic.
Option 2. We ask that

$$(p_1 - p_1^1) + p_1^1 + p_2 = 1,$$

but with $p_1 = 0$ whenever $p_1^1 = 0$,

that is, we treat the three sentences:

“It rains”, “It rains a lot”, “It does not rain”

as if it is never completely false that whenever it rains, it rains a lot – which is actually the case in Gödel logic.

It turns out that this corresponds to axiomatising probability assignments to formulæ in Gödel logic by Kolmogorov’s axioms, plus one additional axiom that I will not state explicitly here.
Option 2. We ask that

\[(p_1 - p_1^1) + p_1^1 + p_2 = 1, \text{ but with } p_1 = 0 \text{ whenever } p_1^1 = 0,\]

that is, we treat the three sentences:

“\text{It rains}”, “\text{It rains a lot}”, “\text{It does not rain}”

as if it is never completely false that whenever it rains, it rains a lot – which is actually the case in Gödel logic.

It turns out that this corresponds to axiomatising probability assignments to formulæ in Gödel logic by Kolmogorov’s axioms, plus one additional axiom that I will not state explicitly here. For various reasons, I am convinced that this latter option with the additional axiom is the right one. The axiom is a condition that comes out whether you want it or not from the structure of Gödel logic. One can ignore it, of course, but then one should be doing classical logic, really.
Of course, this leaves us with another problem here: there is no plausible reason why one should estimate zero the probability that it rains, provided one estimates zero the probability that it rains a lot. (Unless, of course, one adds some extra assumption.) Thus, the intended semantics that seemed acceptable to us on logical grounds, it is no longer so when we try to develop probability over this logic.
Of course, this leaves us with another problem here: there is no plausible reason why one should estimate zero the probability that it rains, provided one estimates zero the probability that it rains a lot. (Unless, of course, one adds some extra assumption.) Thus, the intended semantics that seemed acceptable to us on logical grounds, it is no longer so when we try to develop probability over this logic.

The one clear thing is that more thought is needed.
Is there a probability theory of non-classical events?

Outline

5 A mathematical remark
Added in proof: A mathematical remark. I mentioned two traditions in probability theory: one that takes events (=fields of sets) as a primitive notion, and the other that takes propositions (=Boolean algebras) instead. Yesterday, David Makinson also used a propositional language instead of fields of sets, and seemed to suggest in passing that the difference is immaterial. I agree with this completely only insofar as the language is finite. But suppose it is not. Then I would make a case that the two approaches – fields of sets vs. Boolean algebras – are conceptually very different, and in fact lead to different mathematical developments. A related point concerning finite vs. countable additivity was discussed yesterday over lunch with Mariano Giaquinta.
Added in proof: A mathematical remark. I mentioned two traditions in probability theory: one that takes events (=fields of sets) as a primitive notion, and the other that takes propositions (=Boolean algebras) instead. Yesterday, David Makinson also used a propositional language instead of fields of sets, and seemed to suggest in passing that the difference is immaterial. I agree with this completely only insofar as the language is finite. But suppose it is not. Then I would make a case that the two approaches – fields of sets vs. Boolean algebras – are conceptually very different, and in fact lead to different mathematical developments. A related point concerning finite vs. countable additivity was discussed yesterday over lunch with Mariano Giaquinta.

Sure enough, you can always embed any Boolean algebra into the Boolean algebra of all subsets of some set. And you can do so quite canonically after B. Jónsson and A. Tarski, Amer. J. Math. (1951). This embedding, or variations thereof, has been widely used in measure theory to extend finitely additive measures from a given field of sets $B$ to a countably additive measure on the extension of $B$. Since in the countably additive setting one can apply results from analysis, this provides a valuable tool.
But here is another theorem that also leads from finite to countable additivity, in a conceptually very different way. To the given Boolean algebra $B$ one can associate its Stone space $\mathcal{S}(B)$, a compact totally disconnected Hausdorff space. This association is as canonical and natural as possible – it’s part of a contravariant categorical equivalence, or duality. Elements $a \in B$ will then be uniquely represented by an appropriate continuous function $\hat{a}: \mathcal{S}(B) \to \mathbb{R}$.
But here is another theorem that also leads from finite to countable additivity, in a conceptually very different way. To the given Boolean algebra $B$ one can associate its Stone space $\mathcal{S}(B)$, a compact totally disconnected Hausdorff space. This association is as canonical and natural as possible – it’s part of a contravariant categorical equivalence, or duality. Elements $a \in B$ will then be uniquely represented by an appropriate continuous function $\hat{a}: \mathcal{S}(B) \to \mathbb{R}$.

**Theorem**

There is a bijection between finitely additive probability measures on $B$, and regular Borel (hence of course countably additive) probability measures on $\mathcal{S}(B)$. Specifically, to such a measure $\mu$ on $\mathcal{S}(B)$ one associates the finitely additive probability measure $\mathbb{E}$ that, to each element $a \in B$, assigns

$$\mathbb{E}(a) = \int_{\mathcal{S}(B)} \hat{a} d\mu,$$

where $\int \cdot d\mu$ is Lebesgue integration with respect to $\mu$. 
Is there a probability theory of non-classical events?

A mathematical remark

Thank you for your attention.