Summary of inductive logic

Jason Swanson
University of Central Florida
Orlando, FL
`jason.swanson@ucf.edu`

April 22, 2026

Abstract

In the beginning, probability was part of logic. That was the view of Leibniz, Bernoulli, de Moivre, Laplace, and even Boole. It was a logic that was concerned with degrees of evidentiary support. These degrees were called probabilities. The idea that probabilities are empirical properties of sequences is a modern one. It emerged from the logical empiricism of the early 20th century. (Logical empiricism itself died out in the 1960s.) This new idea came to be known as frequentism. Not everyone accepted the new way of thinking. In the 1950s, the frequentists coined a derogatory nickname for their opponents. They called them Bayesians.

Although shunned by the mainstream of the 20th century, Bayesian ideas now flourish. Modern science has become more probabilistic today than ever before. Out of necessity, people are returning to the original way of thinking. Probability is a form of reasoning.

In the 1930s, Kolmogorov established measure theory as the foundation for probability. Kolmogorov was a frequentist, and this motivated his work. But the mathematical apparatus he constructed was not bound to frequentism. It is perfectly compatible with the original idea of probability as logic. In fact, modern probabilists use measure theory every day. But they do not behave as if they are frequentists. They behave as if they are performing a kind of logic. They work exclusively by deriving probabilistic statements from probabilistic premises. But the logic that underlies their practice is unformalized. Or rather, it was unformalized, until now.

In *The Principle of Probability* [1], a formal system of probabilistic reasoning is presented. This system fully captures the modern practice of probability. It is the first logical system to accomplish this. Embedded in this system is Kolmogorov's measure-theoretic treatment. In this way, we see that measure theory is but the tip of the probability iceberg. Below it lies a foundation of logic. In this short note, we give a very brief overview and summary of this logical system.

1 Syntax

Let L be an extralogical signature. That is, L is a set of constant symbols, relation symbols, and function symbols. Let Var be an uncountable set of individual variables. More precisely, $\mathsf{Var} = \{\boldsymbol{x}_\alpha \mid \alpha < \boldsymbol{\omega}_1\}$, where $\boldsymbol{\omega}_1$ is the first uncountable ordinal. We use the logical symbols, $\neg$, $\bigwedge$, $\forall$, and $=$. The symbols in L are called extralogical symbols.

Terms are constructed as in first-order logic. Each $c \in L$ and each $x \in \mathsf{Var}$ is a term, and if $f \in L$ is an n-ary function symbol, then $ft_1 \cdots t_n$ is a term whenever $t_1, \ldots, t_n$ are terms. The set of terms is denoted by $\mathcal{T}$.

Atomic formulas are also constructed as in first-order logic. If s and t are terms, then $s = t$ is an atomic formula, and if $r \in L$ is an n-ary relation symbol, then $rt_1 \cdots t_n$ is an atomic formula whenever $t_1, \ldots, t_n$ are terms. More generally, formulas are constructed from atomic formulas using the logical operators. If φ is a formula and $x \in \mathsf{Var}$, then $\neg\varphi$ and $\forall x\varphi$ are also formulas. Additionally, if Φ is a countable set of formulas, then $\bigwedge \Phi$ is also a formula. The only difference between this and first-order logic is that we allow countable conjunctions. Common logical operations not specified above are defined using shorthand. For example, $\bigvee \Phi$ is shorthand for $\neg \bigwedge \neg\Phi$, where $\neg\Phi = \{\neg\varphi \mid \varphi \in \Phi\}$. Other examples include

$$\begin{aligned} (\varphi \wedge \psi) &= \textstyle\bigwedge\{\varphi, \psi\}, & (\varphi \vee \psi) &= \textstyle\bigvee\{\varphi, \psi\}, \\ (\varphi \to \psi) &= (\neg\varphi \vee \psi), & (\varphi \leftrightarrow \psi) &= (\varphi \to \psi) \wedge (\psi \to \varphi), \\ (s \neq t) &= \neg(s = t), & \exists x\varphi &= \neg\forall\neg\varphi. \end{aligned}$$

The set of formulas is denoted by $\mathcal{L}_{\boldsymbol{\omega}_1,\boldsymbol{\omega}}$ in the literature. For simplicity, we denote it simply by $\mathcal{L}$.

A variable x in a formula φ is bound if it lies within the scope of a quantifier $\forall x$. Otherwise, it is free. The set of free variables in a formula φ is denoted by $\operatorname{free}\varphi$. More generally, for $X \subseteq \mathcal{L}$, we write $\operatorname{free} X = \bigcup_{\varphi \in X} \operatorname{free}\varphi$. A sentence is a formula without any free variables. The set of sentences is denoted by $\mathcal{L}^0$. Given a formula φ and a term t, we write $\varphi(t/x)$ to denote the formula obtained by replacing every free occurrence of x in φ with the term t. We say that t is free for x in φ if none of the variables in t become bound after performing the substitution $\varphi(t/x)$.

An inductive statement is a triple (X, φ, p), where $X \subseteq \mathcal{L}^0$, $\varphi \in \mathcal{L}^0$, and $p \in [0,1]$. We interpret (X, φ, p) as asserting that X partially entails φ, and that p is the degree of this partial entailment. In an inductive statement, X is called the antecedent, φ is called the consequent, and p is called the probability. The set of inductive statements is $\mathcal{L}^{\mathrm{IS}} = \mathfrak{P}\mathcal{L}^0 \times \mathcal{L}^0 \times [0,1]$, where $\mathfrak{P}$ denotes the power set. If $P \subseteq \mathcal{L}^{\mathrm{IS}}$, we define

$$\operatorname{ante} P = \{X \subseteq \mathcal{L}^0 : (X, \varphi, p) \in P \text{ for some } \varphi \in \mathcal{L}^0 \text{ and } p \in [0,1]\}.$$

That is, $\operatorname{ante} P$ is the set of antecedents that occur in P.

2 Inferential calculus

The turnstile symbol, $\vdash$, is used for both the deductive and inductive derivability relations. We write $X \vdash \varphi$ to mean that the formula φ can be derived from the set of formulas $X \subseteq \mathcal{L}$.

But we also write $Q \vdash (X, \varphi, p)$ to mean the inductive statement (X, φ, p) can derived from the set $Q \subseteq \mathcal{L}^{\mathrm{IS}}$.

2.1 The deductive calculus

For deductive derivability, we use the classical system of natural deduction, with the obvious adjustment for countable conjunctions. That is, the relation $X \vdash \varphi$ is defined by the following rules of deductive inference:

(a) $\varphi \vdash \varphi$,

(b) if $X \vdash \varphi$ and $X \subseteq X'$, then $X' \vdash \varphi$,

(c) if $X \vdash \bigwedge \Phi$, then $X \vdash \theta$ for all $\theta \in \Phi$,

(d) if $X \vdash \theta$ for all $\theta \in \Phi$, then $X \vdash \bigwedge \Phi$,

(e) if $X \vdash \varphi$ and $X \vdash \neg\varphi$, then $X \vdash \psi$,

(f) if $X, \varphi \vdash \psi$ and $X, \neg\varphi \vdash \psi$, then $X \vdash \psi$.

(g) if $X \vdash \forall x\varphi$, then $X \vdash \varphi(t/x)$ when t is free for x in φ,

(h) if $x \notin \operatorname{free} X$ and $X \vdash \varphi$, then $X \vdash \forall x\varphi$,

(i) $\vdash t = t$ for all terms t, and

(j) if $X \vdash s = t, \varphi(s/x)$, then $X \vdash \varphi(t/x)$ when s and t are free for x in φ.

If $Y \subseteq \mathcal{L}$, we write $X \vdash Y$ to mean $X \vdash \psi$ for all $\psi \in Y$. Two formulas φ and ψ are (logically) equivalent, written $\varphi \equiv \psi$, if $\varphi \vdash \psi$ and $\psi \vdash \varphi$. More generally, $X \equiv Y$ means $X \vdash Y$ and $Y \vdash X$. Two formulas φ and ψ are equivalent given X, written $\varphi \equiv_X \psi$, if $X, \varphi \vdash \psi$ and $X, \psi \vdash \varphi$.

A (deductive) theory is a set $T \subseteq \mathcal{L}^0$ that is closed under deductive inference. That is, $T \vdash \varphi$ implies $\varphi \in T$ for all $\varphi \in \mathcal{L}^0$. If $X \subseteq \mathcal{L}^0$, the (deductive) theory generated by X is $T(X) = \{\varphi \in \mathcal{L}^0 \mid X \vdash \varphi\}$. We also denote this by T_X, and it is the smallest theory having X as a subset.

2.2 The inductive calculus

Inductive derivability is defined using nine rules of inductive inference. Each rule is expressed by saying what it means for a set $P \subseteq \mathcal{L}^{\mathrm{IS}}$ to be closed under that rule. The first rule is

(R1) (*the rule of logical equivalence*) If $(X, \varphi, p) \in P$, $X' \equiv X$, and $\varphi' \equiv_X \varphi$, then $(X', \varphi', p) \in P$ and there is no other value p' such that $(X', \varphi', p') \in P$.

A set P is called admissible if it satisfies (R1). If P is admissible, then for a given X and φ, there can be at most one p such that $(X, \varphi, p) \in P$. We therefore write $P(\varphi \mid X) = p$ as shorthand for $(X, \varphi, p) \in P$. Commas to the right of the bar are understood as unions, so that, for example $P(\varphi \mid X, \psi)$ means $P(\varphi \mid X \cup \{\psi\})$. When $X = \emptyset$, we often omit it, leaving only $P(\varphi)$ or $P(\varphi \mid \psi)$. For admissible P, if $(X, \varphi, p) \in P$, then we say $P(\varphi \mid X)$ exists or is defined.

Rules (R2)–(R4) describe the relationship between inductive and deductive inference.

(R2) (*the rule of logical implication*) If $X \in \operatorname{ante} P$ and $X \vdash \varphi$, then $P(\varphi \mid X) = 1$.

(R3) (*the rule of material implication*) If $X \in \operatorname{ante} P$ and $P(\psi \mid X, \varphi) = 1$, then $P(\varphi \to \psi \mid X) = 1$.

(R4) (*the rule of deductive transitivity*) If $P(\varphi \mid X) = 1$ and $\varphi \vdash \psi$, then $P(\psi \mid X) = 1$. Also, for any $X' \in \operatorname{ante} P$, if $X' \vdash X$ and $P(\varphi \mid X) = 1$, then $P(\varphi \mid X') = 1$.

Rules (R5)–(R7) are the familiar, basic rules of probability.

(R5) (*the addition rule*) Let $X \vdash \neg(\varphi \wedge \psi)$. Consider the equation,

$$P(\varphi \vee \psi \mid X) = P(\varphi \mid X) + P(\psi \mid X). \tag{1}$$

If two of the above probabilities exist, then so does the third and (1) holds.

(R6) (*the multiplication rule*) Consider the equation,

$$P(\varphi \wedge \psi \mid X) = P(\varphi \mid X) P(\psi \mid X, \varphi). \tag{2}$$

If two of the above probabilities exist and are positive, then the third exists and (2) holds.

(R7) (*the continuity rule*) If $P(\varphi_n \mid X)$ exists and $X, \varphi_n \vdash \varphi_{n+1}$ for all $n \in \mathbb{N}$, then

$$P(\textstyle\bigvee_n \varphi_n \mid X) = \lim_n P(\varphi_n \mid X).$$

A set P is called entire if it satisfies (R1)–(R7). It can be shown from these rules that in (R6), the existing probabilities need not be positive. It is only required that when solving for the third probability, we are not dividing by zero.

Rule (R8) says that we can use uniqueness to make inferences. That is, if there is a unique way to assign a probability without violating (R1)–(R7), then we may infer that probability. To state (R8), we first need a definition. If $P \subseteq \overline{P} \subseteq \mathcal{L}^{\mathrm{IS}}$, then $\overline{P}$ is called a completion of P if $\overline{P}$ is entire and also satisfies the following conditions:

(i) If $\overline{P}(\varphi \mid X)$ and $\overline{P}(\psi \mid X)$ exist, then $\overline{P}(\varphi \wedge \psi \mid X)$ exists.

(ii) If $X \in \operatorname{ante} \overline{P}$ and $X \cup \{\varphi\} \in \operatorname{ante} \overline{P}$, then $\overline{P}(\varphi \mid X)$ exists.

With this definition, we have

(R8) (*the rule of inductive extension*) If $\overline{P}(\varphi \mid X) = p$ for every completion $\overline{P}$ of P, then $P(\varphi \mid X) = p$.

As set P is called semi-closed if it satisfies (R1)–(R8).

The final rule is

(R9) (*the rule of deductive extension*) If $S \subseteq \mathcal{L}^0$ is nonempty and $P(\theta \mid X) = 1$ for all $\theta \in S$, then $X \cup S \in \operatorname{ante} P$ and $P(\,\cdot \mid X, S) = P(\,\cdot \mid X)$.

Rule (R9) says that we may freely add formulas of probability one to our antecedents. A set P is called closed if it satisfies (R1)–(R9).

A set $Q \subseteq \mathcal{L}^{\mathrm{IS}}$ is called connected if there exists $\widehat{Q} \subseteq Q$ such that

(i) every $\widehat{X} \in \operatorname{ante} \widehat{Q}$ is countable axiomatizable over a common $X_0 \in \operatorname{ante} \widehat{Q}$, and

(ii) every $X \in \operatorname{ante} Q$ is logically equivalent to $\widehat{X} \cup S$ for some $\widehat{X} \in \operatorname{ante} \widehat{Q}$ and some $S \subseteq \{\theta \in \mathcal{L}^0 \mid (\widehat{X}, \theta, 1) \in \widehat{Q}\}$.

When we require that a set Q is connected, we ensure that the inductive statements in Q can be potentially related to one another by countably many applications of the rules of inductive inference. Given a connected set Q, the set X_0 in (i) is unique up to logical equivalence. We use T_0 to denote the deductive theory generated by X_0 and call T_0 the root of Q.

An inductive theory is a set $P \subseteq \mathcal{L}^{\mathrm{IS}}$ that is both closed and connected. If T_0 is the root of P, then we define

$$T_P = \{\theta \in \mathcal{L}^0 \mid P(\theta \mid T_0) = 1\}. \tag{3}$$

The set T_P is a deductive theory which we call the deductive theory determined by P. It has the property that $\theta \in T_P$ if and only if $P(\theta \mid X) = 1$ for all $X \in \operatorname{ante} P$. In other words, T_P is the set of sentences that are true in P, regardless of the antecedent. Moreover, for a fixed $X \in \operatorname{ante} P$, we have $P(\varphi \mid X) = 1$ if and only if $X, T_P \vdash \varphi$.

A set $Q \subseteq \mathcal{L}^{\mathrm{IS}}$ is consistent if it is connected and can be extended to an inductive theory. In that case, Q has a unique smallest extension to an inductive theory, which we denote by $\mathbf{P}(Q)$, or $\mathbf{P}_Q$, and call the inductive theory generated by Q. We write $Q \vdash (X, \varphi, p)$ to mean that Q is consistent and $(X, \varphi, p) \in \mathbf{P}(Q)$.

An inductive condition is a set $\mathcal{C}$ of inductive theories with a common root. If that common root is T_0, then we say T_0 is the root of $\mathcal{C}$. An inductive condition is consistent if it is nonempty. If $\mathcal{C}$ is a consistent inductive condition with root T_0, then the inductive theory generated by $\mathcal{C}$, which we denote by $\mathbf{P}(\mathcal{C})$ or $\mathbf{P}_{\mathcal{C}}$, is the largest inductive theory with root T_0 that is contained in $\bigcap \mathcal{C}$.

We write $\mathcal{C} \vdash (X, \varphi, p)$ to mean that $(X, \varphi, p) \in \mathbf{P}(\mathcal{C})$. These two uses of inductive derivability are related in the following way. If Q is connected with root T_0, then let $\mathcal{C}_Q$ be the set of inductive theories P with root T_0 such that $Q \subseteq P$. Then $\mathcal{C}_Q$ is consistent if and only if Q is consistent, and in that case, $\mathbf{P}(\mathcal{C}_Q) = \mathbf{P}(Q)$. Hence, $\mathcal{C}_Q \vdash (X, \varphi, p)$ if and only if $Q \vdash (X, \varphi, p)$.

2.3 Independence

Fix a set $X \subseteq \mathcal{L}^0$, an inductive theory $P \subseteq \mathcal{L}^{\mathrm{IS}}$ with root T_0, and two sentences $\varphi, \psi \in \mathcal{L}^0$. Assume that $P(\varphi \mid X)$ and $P(\psi \mid X)$ both exist.

W say that φ is dependent on ψ (given X, under P) if $P(\varphi \mid X, \psi)$ exists and is not equal to $P(\varphi \mid X)$. We say that φ is independent of ψ if either $P(\psi \mid X) = 0$ or $P(\varphi \mid X, \psi) = P(\varphi \mid X)$. In any inductive theory P, we have $X \cup \{\psi\} \in \operatorname{ante} P$ if and only if $P(\psi \mid X) > 0$. Hence, it cannot happen that φ is both dependent on and independent of ψ. Conversely, if $P(\varphi \mid X, \psi)$ exists, then one of the two conditions holds.

It can be shown that φ is dependent on ψ if and only if ψ is dependent on φ. The same is also true for independence. In fact, φ is independent of ψ if and only if

$$P(\varphi \wedge \psi \mid X) = P(\varphi \mid X)P(\psi \mid X).$$

Consequently, we can change our language and say that φ and ψ are dependent, or that φ and ψ are independent.

To define independence for more than two sentences, we must introduce dialog sets. A dialog set is a nonempty set of sentences that is closed under negation, countable conjunction, and logical equivalence. If $Y \subseteq \mathcal{L}^0$, then $\delta(Y)$ denotes the smallest dialog set containing Y, and is called the dialog set generated by Y. The dialog set $\delta(Y)$ consists of all sentences that can be formed from Y using negation, conjunction, and equivalence.

Now, let $\langle \varphi_i \mid i \in I \rangle$ be an indexed collection of two or more formulas. Assume that $P(\varphi_i \mid X)$ exists for each i. Such a collection is independent (given X, under P) if φ and ψ are independent whenever $\varphi \in \delta(\{\varphi_i \mid i \in I_1\})$ and $\psi \in \delta(\{\varphi_i \mid i \in I_2\})$, where I_1 and I_2 are nonempty disjoint subsets of I. In other words, whenever φ and ψ are formed from nonoverlapping subsets of $\langle \varphi_i \mid i \in I \rangle$, they are independent. This definition generalizes the previous one in the sense that φ and ψ are independent if and only if $\langle \varphi, \psi \rangle$ is independent. Moreover, it can be shown that $\langle \varphi_i \mid i \in I \rangle$ is independent if and only if

$$P(\textstyle\bigwedge_{j \in J} \varphi_j \mid X) = \prod_{j \in J} P(\varphi_j \mid X)$$

for all finite $J \subseteq I$.

2.4 A coin flipping example

In this example, we build an inductive theory that describes two flips of a coin that is twice as likely to land on heads than tails. Let $L = \{c_1, c_2, h, t\}$, where all symbols in L are constant symbols. The constants h and t denote the heads and tails sides of the coin, and the constants c_1 and c_2 denote the results of the two flips.

Define the sentences

$$\begin{aligned} &\varphi_0 : h \neq t \\ &\varphi_i : c_i = h \vee c_i = t \end{aligned}$$

and let $T_0 = T(\{\varphi_0, \varphi_1, \varphi_2\})$. Then the theory T_0 represents the assumptions that the heads and tails sides of the coin are distinct, and that each flip must result in either heads or tails.

Now let

$$Q = \{(T_0, c_i = h, 2/3) \mid i = 1, 2\}.$$

The inductive statements $Q \subseteq \mathcal{L}^{\mathrm{IS}}$ represent the assumptions that each flip has probability 2/3 of landing on heads.

The set Q is clearly connected with root T_0. It can be shown that Q is consistent. This is done most easily using a model (see Section 3). Once this is done, we may define $\mathbf{P}_Q$, the inductive theory generated by Q.

To illustrate the use of the inductive calculus, let us show that $Q \vdash (T_0, c_1 = t, 1/3)$. In other words, let us demonstrate how to derive, from our assumptions, that the first coin lands on tails with probability 1/3. To do this, we must show that $\mathbf{P}_Q(c_1 = t \mid T_0) = 1/3$. We know that $\mathbf{P}_Q$ satisfies (R1)–(R9). We first use the deductive rules (a)–(j) to show that $T_0 \vdash \neg\psi$, where $\psi = (c_1 = h \wedge c_1 = t)$. Let $x \in \mathit{Var}$ and define $\varphi = (x = t)$. Then c_1 and h are free for x in t. Note that $\varphi(c_1/x) = (c_1 = t)$. By (b) and (c), we have $T_0, \psi \vdash c_1 = h, \varphi(c_1/x)$. Hence, (j) gives $T_0, \psi \vdash h = t$, since $\varphi(h/x) = (h = t)$. Now, $h \neq t$ is shorthand for $\neg h = t$. Thus, by (a), (b), and the fact that $\varphi_0 \in T_0$, we have $T_0, \psi \vdash \neg h = t$. Finally, then, (e) gives $T_0, \psi \vdash \neg\psi$. But (a) and (b) give $T_0, \neg\psi \vdash \neg\psi$, so we may use (f) to obtain $T_0 \vdash \neg\psi$.

With this established, we may now use the addition rule, (R5). Consider the equation

$$\mathbf{P}_Q(c_1 = h \vee c_1 = t \mid T_0) = \mathbf{P}_Q(c_1 = h \mid T_0) + \mathbf{P}_Q(c_1 = t \mid T_0).$$

By (a) and (b), we have $T_0 \vdash c_1 = h \vee c_1 = t$. Hence, (R2) gives $\mathbf{P}_Q(c_1 = h \vee c_1 = t \mid T_0) = 1$. Also, by the definition of Q, we have $\mathbf{P}_Q(c_1 = h \mid T_0) = 2/3$. Therefore, according to (R5), the probability $\mathbf{P}_Q(c_1 = t \mid T_0)$ exists and the above equation holds, which gives us $\mathbf{P}_Q(c_1 = t \mid T_0) = 1/3$.

In the inductive theory $\mathbf{P}_Q$, we did not include the assumption that the two coin flips are independent. We can include this using an inductive condition. Let $\mathcal{C}$ be the set of inductive theories P with root T_0 such that $Q \subseteq P$ and

$$P(c_2 = h \mid T_0, c_1 = h) = P(c_2 = h \mid T_0). \tag{4}$$

Then $\mathcal{C}$ represents all of the previous assumptions, together with the assumption that the sentences $c_1 = h$ and $c_2 = h$ are independent.

As above, using models, it can be shown that $\mathcal{C}$ is consistent. Once this is done, we may define $\mathbf{P}_{\mathcal{C}}$, the inductive theory generated by $\mathcal{C}$. We can now show that

$$\mathcal{C} \vdash (T_0, c_1 = h \wedge c_2 = h, 4/9).$$

In other words, $\mathbf{P}_{\mathcal{C}}(c_1 = h \wedge c_2 = h \mid T_0) = 4/9$.

By definition, $\mathbf{P}_C$ is the largest inductive theory contained in $\bigcap \mathcal{C}$. But it may happen that $\mathbf{P}_{\mathcal{C}}$ is not equal to $\bigcap \mathcal{C}$. This is because the intersection of a collection of inductive theories need not be connected. If we stay close to the root, however, we avoid this problem. More specifically, for any $\varphi, \psi \in \mathcal{L}^0$ and any $p \in [0, 1]$, if $P(\varphi \mid T_0, \psi) = p$ for all $P \in \mathcal{C}$, then $\mathbf{P}_{\mathcal{C}}(\varphi \mid T_0, \psi) = p$. (See [1, Section 3.5.4] for details.) We can now use this to prove our claim.

We will apply the multiplication rule, (R6), to an arbitrary $P \in \mathcal{C}$. Consider the equation

$$P(c_1 = h \wedge c_2 = h \mid T_0) = P(c_1 = h \mid T_0) P(c_2 = h \mid T_0, c_1 = h).$$

By the definition of Q, we have $P(c_i = h \mid T_0) = 2/3$. Hence, the first factor on the right of the above equation exists and equals 2/3. Also, using (4), the second factor exists and equals 2/3. Thus, according to (R6), the probability on the left exists and equals 4/9. Since $P \in \mathcal{C}$ was arbitrary, this shows that $\mathbf{P}_{\mathcal{C}}(c_1 = h \wedge c_2 = h \mid T_0) = 4/9$.

3 Models and semantics

Let A be a set, and let us associate each extralogical symbol $\mathsf{s} \in L$ with an actual object or relation in A, that we denote by s^ω. If $L^\omega = \{\mathsf{s}^\omega \mid \mathsf{s} \in L\}$, then the pair $\omega = (A, L^\omega)$ is called a structure, or an L-structure.

An assignment into A is a function $v : \mathit{Var} \to A$. Each assignment has a unique extension $v : \mathcal{T} \to A$ that respects the extralogical signature. More specifically, $v(c) = c^\omega$ for constant symbols $c \in L$ and $v(ft_1 \cdots t_n) = f^\omega(v(t_1), \ldots, v(t_n))$. We call this extension an assignment into ω.

Just as in first order logic, there is a natural way to define what it means for a formula φ to be true in ω when the free variables are assigned values according to v. When this is the case, we say that ω strictly satisfies φ with v, and we write $\omega \models\!\!\equiv \varphi[v]$. The formal definition of strict satisfiability is given recursively by

(i) $\omega \models\!\!\equiv (s = t)[v]$ if and only if $v(s) = v(t)$,

(ii) $\omega \models\!\!\equiv (rt_1 \cdots t_n)[v]$ if and only if $(v(t_1), \ldots, v(t_n)) \in r^\omega$,

(iii) $\omega \models\!\!\equiv (\neg\varphi)[v]$ if and only if $\omega \not\models\!\!\equiv \varphi[v]$,

(iv) $\omega \models\!\!\equiv (\bigwedge \Phi)[v]$ if and only if $\omega \models\!\!\equiv \varphi[v]$ for all $\varphi \in \Phi$, and

(v) $\omega \models\!\!\equiv (\forall x \varphi)[v]$ if and only $\omega \models\!\!\equiv \varphi[v_x^a]$ for all $a \in A$.

In (v), the assignment v_x^a is the one defined by $v_x^a(x) = a$ and $v_x^a(y) = v(y)$ for $y \neq x$.

An L-model, or simply a model, is a probability space $\mathscr{P} = (\Omega, \Sigma, \mathbb{P})$ such that Ω is a set of L-structures. An assignment into $\mathscr{P}$ is an indexed collection $\boldsymbol{v} = \langle v_\omega \mid \omega \in \Omega \rangle$, where v_ω is an assignment into ω for each $\omega \in \Omega$. If $\boldsymbol{v}$ is an assignment into $\mathscr{P}$ and $\varphi \in \boldsymbol{\mathcal{L}}$, then we define

$$\varphi[\boldsymbol{v}]_\Omega = \{\omega \in \Omega \mid \omega \models\!\!\equiv \varphi[v_\omega]\}.$$

We say that $\mathscr{P}$ satisfies φ with $\boldsymbol{v}$, denoted by $\mathscr{P} \vDash \varphi[\boldsymbol{v}]$, if $\varphi[\boldsymbol{v}]_\Omega \in \overline{\Sigma}$ and $\overline{\mathbb{P}}\, \varphi[\boldsymbol{v}]_\Omega = 1$, where $\overline{\mathscr{P}} = (\Omega, \overline{\Sigma}, \overline{\mathbb{P}})$ is the completion of $\mathscr{P}$.

If φ is a sentence, then $\varphi[\boldsymbol{v}]_\Omega$ does not depend on $\boldsymbol{v}$. In this case, we simply write φ_Ω, and we have

$$\varphi_\Omega = \{\omega \in \Omega \mid \omega \models\!\!\equiv \varphi\}.$$

We then write $\mathscr{P} \vDash \varphi$ to mean $\mathscr{P} \vDash \varphi[\boldsymbol{v}]$ for all assignments $\boldsymbol{v}$, and this holds if and only if $\varphi_\Omega \in \overline{\Sigma}$ and $\overline{\mathbb{P}}\, \varphi_\Omega = 1$.

A set $X \subseteq \boldsymbol{\mathcal{L}}$ is satisfiable if there exist $\mathscr{P}$ and $\boldsymbol{v}$ such that $\mathscr{P} \vDash \psi[\boldsymbol{v}]$ for all $\psi \in X$. A set X is consistent if and only if it is satisfiable. We write $X \vDash \varphi$ to mean that $\mathscr{P} \vDash \varphi[\boldsymbol{v}]$ whenever $\mathscr{P} \vDash \psi[\boldsymbol{v}]$ for all $\psi \in X$. The relation $\vDash$ is called the consequence relation. When used in this way, we might also call it the deductive consequence relation.

The deductive consequence relation is σ-compact. That is, a set X is satisfiable if and only if every countable subset of X is satisfiable. It is also complete with respect to the deductive derivability relation, meaning that $X \vDash \varphi$ if and only if $X \vdash \varphi$.

Because of deductive completeness, the notions of logical equivalence and connectivity can be formulated semantically. We can therefore use them to formulate the inductive semantics without fear of circularity. Let $(X, \varphi, p) \in \mathcal{L}^{\mathrm{IS}}$ be an inductive statement. Then $\mathscr{P}$ satisfies (X, φ, p), written $\mathscr{P} \vDash (X, \varphi, p)$, if there exists $Y \subseteq \mathcal{L}^0$ and $\psi \in \mathcal{L}^0$ such that

(i) $X \equiv Y \cup \{\psi\}$,

(ii) $\mathscr{P} \vDash Y$, and

(iii) $\dfrac{\overline{\mathbb{P}}\, \varphi_\Omega \cap \psi_\Omega}{\overline{\mathbb{P}}\, \psi_\Omega} = p$.

More generally, $\mathscr{P} \vDash Q \subseteq \mathcal{L}^{\mathrm{IS}}$ if $\mathscr{P} \vDash (X, \varphi, p)$ for all $(X, \varphi, p) \in Q$. We say that Q is satisfiable if $\mathscr{P} \vDash Q$ for some model $\mathscr{P}$. A set Q is consistent if and only if it is connected and satisfiable.

Let Q be connected and satisfiable. If T_0 is the root of Q, then

$$T_Q = \{\theta \in \mathcal{L}^0 \mid \text{for all } \mathscr{P}, \text{ if } \mathscr{P} \vDash Q \text{ and } \mathscr{P} \vDash T_0, \text{ then } \mathscr{P} \vDash \theta\}$$

is a consistent deductive theory with $T_0 \subseteq T_Q$. (If Q is an inductive theory, then this definition is equivalent to (3).) We write $X \hookrightarrow [T_0, T_Q]$ to mean that X is countably axiomatizable over some deductive theory T with $T_0 \subseteq T \subseteq T_Q$. We write $Q \vDash (X, \varphi, p)$ to mean that

(i) Q is connected and satisfiable,

(ii) $X \hookrightarrow [T_0, T_Q]$, where T_0 is the root of Q, and

(iii) $\mathscr{P} \vDash Q$ implies $\mathscr{P} \vDash (X, \varphi, p)$, for all models $\mathscr{P}$.

When we use the consequence relation in this way, we might also call it the inductive consequence relation. The inductive consequence relation is complete with respect to the inductive derivability relation. That is, $Q \vDash (X, \varphi, p)$ if and only if $Q \vdash (X, \varphi, p)$.

We say that $\mathscr{P}$ satisfies an inductive condition $\mathcal{C}$ if $\mathscr{P} \vDash P$ for some $P \in \mathcal{C}$. We define $\mathcal{C} \vDash (X, \varphi, p)$ exactly as above, but with Q replaced by $\mathcal{C}$. (We also omit the connectivity requirement, since connectivity is not defined for inductive conditions.) This extension of the consequence relation is also complete with respect to derivability. That is, $\mathcal{C} \vDash (X, \varphi, p)$ if and only if $\mathcal{C} \vdash (X, \varphi, p)$.

Example 3.1. Returning to the example in Section 2.4, let us now build a model for our coin flips. Let $A = \{0, 1\}$. We first build four structures, ω_0, ω_1, ω_2, and ω_3. Each structure will have domain A, meaning it assigns objects in A to the symbols in L. These assignments are

$$\begin{array}{llll}
h^{\omega_0} = 1 & t^{\omega_0} = 0 & c_1^{\omega_0} = 0 & c_2^{\omega_0} = 0 \\
h^{\omega_1} = 1 & t^{\omega_1} = 0 & c_1^{\omega_1} = 0 & c_2^{\omega_1} = 1 \\
h^{\omega_2} = 1 & t^{\omega_2} = 0 & c_1^{\omega_2} = 1 & c_2^{\omega_2} = 0 \\
h^{\omega_3} = 1 & t^{\omega_3} = 0 & c_1^{\omega_3} = 1 & c_2^{\omega_3} = 1
\end{array}$$

Let $\Omega = \{\omega_0, \omega_1, \omega_2, \omega_3\}$ and $\Sigma = \mathfrak{P}\Omega$, and define $\mathbb{P}$ by

$$\mathbb{P}\{\omega_0\} = 1/9, \qquad \mathbb{P}\{\omega_1\} = \mathbb{P}\{\omega_2\} = 2/9, \qquad \mathbb{P}\{\omega_3\} = 4/9.$$

Then $\mathscr{P} = (\Omega, \Sigma, \mathbb{P})$ is a model.

Note that $h^{\omega_i} \neq t^{\omega_i}$, for all i. Hence, $\omega_i \models\!\!\!\equiv \varphi_0$ for all i. This means $(\varphi_0)_\Omega = \Omega$, so that $\overline{\mathbb{P}}(\varphi_0)_\Omega = 1$. Therefore, $\mathscr{P} \vDash \varphi_0$. The same is true for φ_1 and φ_2, so that $\mathscr{P} \vDash T_0$. We also have $(c_1 = h)_\Omega = \{\omega_2, \omega_3\}$. If we define $\top = (\exists \boldsymbol{x}_0\, \boldsymbol{x}_0 = \boldsymbol{x}_0)$, then $T_0 \equiv T_0 \cup \{\top\}$ and $\top_\Omega = \Omega$, so that

$$\frac{\overline{\mathbb{P}}(c_1 = h)_\Omega \cap \top_\Omega}{\overline{\mathbb{P}}\, \top_\Omega} = \mathbb{P}(c_1 = h)_\Omega = \frac{2}{3}.$$

This shows that $\mathscr{P} \vDash (T_0, c_1 = h, 2/3)$. Similarly, $(c_2 = h)_\Omega = \{\omega_1, \omega_3\}$, and we have $\mathscr{P} \vDash (T_0, c_2 = h, 2/3)$. Therefore, $\mathscr{P} \vDash Q$. In other words, Q is satisfiable. Since Q is also connected, this shows that Q is consistent.

To show that $\mathscr{P} \vDash \mathcal{C}$, we must show that $\mathscr{P} \vDash P$ for some $P \in \mathcal{C}$. First note that

$$\frac{\overline{\mathbb{P}}(c_2 = h)_\Omega \cap (c_1 = h)_\Omega}{\overline{\mathbb{P}}(c_1 = h)_\Omega} = \frac{\mathbb{P}\{\omega_3\}}{\mathbb{P}\{\omega_2, \omega_3\}} = \frac{2}{3},$$

which gives

$$\mathscr{P} \vDash (T_0 \cup \{c_1 = h\}, c_2 = h, 2/3). \tag{5}$$

Now, let us define

$$P = \{(X, \varphi, p) \in \mathcal{L}^{\mathrm{IS}} \mid \mathscr{P} \vDash (X, \varphi, p) \text{ and } X \vdash T_0\}.$$

According to [1, Theorem 4.2.6 and Proposition 3.5.10], the set P is an inductive theory with root T_0. From (5), we see that P satisfies (4). Therefore, $P \in \mathcal{C}$, and so $\mathscr{P} \vDash \mathcal{C}$. In other words, $\mathcal{C}$ is satisfiable, which shows that $\mathcal{C}$ is consistent.

4 Embedding of ordinary probability theory

Let (S, Γ, ν) be a probability space and $\langle (R_i, \Gamma_i) \mid i \in I \rangle$ an indexed collection of measurable spaces. For each $i \in I$, let $X_i : S \to R_i$ be a measurable function so that $X = \langle X_i \mid i \in I \rangle$ is an indexed collection of random variables. Assume that Γ is the σ-algebra generated by X. Then (S, Γ, ν, X) is what we call a measure-theoretic probability model. These are the most general types of objects studied in modern mathematical probability theory.

The collection of all measure-theoretic probability models can be naturally and properly embedded into the semantics of inductive logic. There are three embedding theorems in [1] that demonstrate this. They are [1, Theorems 5.4.2, 6.3.4, and 6.4.6]. The last of these is the strongest. Roughly speaking, these theorems say that every measure-theoretic probability model has a natural correspondence to an inductive model $\mathscr{P} = (\Omega, \Sigma, \mathbb{P})$. Outcomes $x \in S$ correspond to structures $\omega \in \Omega$. Events $U \in \Gamma$ correspond to sentences $\varphi \in \mathcal{L}^0$. The set membership relation $x \in U$ corresponds to strict satisfiability $\omega \models\!\!\!\equiv \varphi$. And the random variables correspond to constant symbols in the extralogical signature.

References

[1] J. Swanson. *The Principles of Probability: From Formal Logic to Measure Theory to the Principle of Indifference.* Lecture Notes in Mathematics. Springer Nature Switzerland, 2026.