

Philosophy 148 — Announcements & Such

- Administrative Stuff
 - Branden’s office hours today will be 3–4.
 - We have a permanent location for the Tues. section: 206 Wheeler.
 - We have a permanent location for the Wed. section: 2304 Tolman.
- Last Time: Algebraic Probability (Part II)
- Today’s Agenda
 - An Algebraic Approach to Probability Calculus, Continued
 - * PrSAT: A user-friendly decision Procedure for probability calculus
 - * Systematic vs Extra-Systematic Logical Relations in Algebraic PC
 - Then: The (Orthodox) Axiomatic Approach to Probability Calculus
 - * The axioms, and axiomatic proofs of some basic theorems
 - * Alternative axiomatizations (conditional probability as primitive)
 - * Probabilistic independence and correlation
- Next: Some “Interpretations” or “Kinds” of Probability.

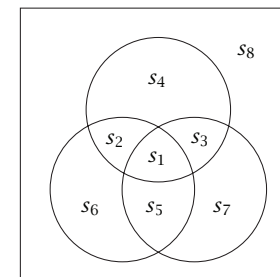
The Probability Calculus: An Algebraic Approach XII

- There are *decision procedures* for Boolean propositional logic, based on truth-tables. These methods are *exponential* in the number of atomic sentences (n), because truth-tables grow exponentially in n (2^n).
- It would be nice if there were a decision procedure for probability calculus, too. In algebraic terms, this would require a decision procedure for the salient fragment of high-school (real) algebra.
- As it turns out, high-school (real) algebra (HSA) is a decidable theory. This was shown by Tarski in the 1920’s. But, it’s only been very recently that computationally feasible procedures have been developed.
- In my “A Decision Procedure for Probability Calculus with Applications”, I describe a user-friendly decision procedure (called PrSAT) for probability calculus, based on recent HSA procedures.
- My implementation is written in *Mathematica* (a general-purpose mathematics computer programming framework). It is freely downloadable from my website, at: <http://fitelson.org/PrSAT/>.

The Probability Calculus: An Algebraic Approach XI

- The algebraic approach for *refuting* general claims involves two steps:
 1. Translate the claim from probability notation into algebraic terms.
 2. Find a (numerical) probability model on which the translation is *false*.
- Show that $\Pr(X | Y \& Z) = \Pr(X | Y \vee Z)$ can be *false*. Here’s a model \mathcal{M} :

X	Y	Z	States	$\Pr(s_i)$
T	T	T	s_1	$a_1 = 1/6$
T	T	F	s_2	$a_2 = 1/6$
T	F	T	s_3	$a_3 = 1/4$
T	F	F	s_4	$a_4 = 1/16$
F	T	T	s_5	$a_5 = 1/6$
F	T	F	s_6	$a_6 = 1/12$
F	F	T	s_7	$a_7 = 1/24$
F	F	F	s_8	$a_8 = 1/16$



(1) Algebraic Translation: $\frac{a_1}{a_1 + a_5} = \frac{a_1 + a_2 + a_3}{a_1 + a_2 + a_3 + a_5 + a_6 + a_7}$.

(2) This claim is *false* on \mathcal{M} , since $1/2 \neq 2/3$. I used PrSAT to find \mathcal{M} .

The Probability Calculus: An Algebraic Approach XIII

- I encourage the use of PrSAT as a tool for finding counter-models and for establishing theorems of probability calculus. It is not a requirement of the course, but it is a useful tool that is worth learning.
- PrSAT doesn’t give readable proofs of theorems. But, it will find concrete numerical counter-models for claims that are not theorems.
- PrSAT will also allow you to calculate probabilities that are determined by a *given* probability assignment. And, it will allow you to do algebraic and numerical “scratch work” without making errors.
- I have posted a *Mathematica* notebook which contains the examples from algebraic probability calculus that we have seen in this lecture. I will be posting further notebooks as the course goes along.
- Let’s have a look at this first notebook (`examples_1.nb`). I will now go through the examples in this notebook, and demonstrate some of the features of PrSAT. I encourage you to play around with it.

Systematic vs Extra-Systematic Logical Relations I

- The entailment relation \models that we've been talking about is just the Boolean entailment relation that is in force *within* the algebra over which $\text{Pr}(\cdot)$ is defined. I will call this relation **systematic entailment**.

☞ Because **probability zero is not the same thing as systematic logical falsehood**, there is room to emulate *extra-systematic* logical relations using probability models. This is an important "trick" we'll use often.

- Here's an example. Consider a propositional language with three atomic letters: X, Y, Z . This sets-up the standard 3-atomic-sentence Boolean algebra \mathcal{B} that we've seen several times already. Now, we'll add a twist.
- Let's *extra-systematically* interpret ' X ' as $(\forall x)(Rx \rightarrow Bx)$, ' Y ' as Ra , and ' Z ' as Ba . This extra-systematic interpretation of the atomic sentences has no effect on the systematic logical relations in \mathcal{B} .
- But, we can use a suitable $\text{Pr}(\cdot)$ over \mathcal{B} to *emulate* the extra-systematic (MPL) entailment relations (\models) between Ra, Ba , and $(\forall x)(Rx \rightarrow Bx)$.

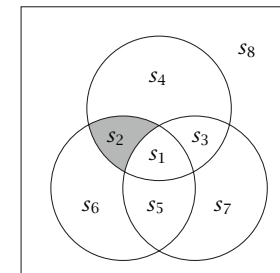
Axiomatic Treatment of Probability Calculus I

- A probability model \mathcal{M} is a Boolean algebra of propositions \mathcal{B} , together with a function $\text{Pr}(\cdot) : \mathcal{B} \rightarrow \mathbb{R}$ satisfying the following three *axioms*.
 - For all $p \in \mathcal{B}$, $\text{Pr}(p) \geq 0$. [non-negativity]
 - $\text{Pr}(\top) = 1$, where \top is the tautological proposition. [normality]
 - For all $p, q \in \mathcal{B}$, if p and q are mutually exclusive (inconsistent), then $\text{Pr}(p \vee q) = \text{Pr}(p) + \text{Pr}(q)$. [additivity axiom]
- Conditional probability is *defined* in terms of unconditional probability in the usual way: $\text{Pr}(p | q) \stackrel{\text{def}}{=} \frac{\text{Pr}(p \& q)}{\text{Pr}(q)}$, provided that $\text{Pr}(q) > 0$.
- We could also state everything in terms of a (propositional) *language* \mathcal{L} with a finite number of atomic *sentences*. Then, we would talk about *sentences* rather than *propositions*, and the axioms would read:
 - For all $p \in \mathcal{L}$, $\text{Pr}(p) \geq 0$.
 - For all $p \in \mathcal{L}$, if $p \models \top$, then $\text{Pr}(p) = 1$.
 - For all $p, q \in \mathcal{L}$, if $p \& q \models \perp$, then $\text{Pr}(p \vee q) = \text{Pr}(p) + \text{Pr}(q)$.

Systematic vs Extra-Systematic Logical Relations II

- Example.** *Extra-systematically*, we have: $(\forall x)(Rx \rightarrow Bx) \& Ra \models Ba$.
- We do *not* have the corresponding *systematic* entailment: $X \& Y \neq Z!$
- But, we can *emulate* this \models relation, by assigning $\text{Pr}(X \& Y \& \sim Z) = 0$.

X	Y	Z	States	$\text{Pr}(s_i)$
T	T	T	s_1	a_1
T	T	F	s_2	$a_2 = 0$
T	F	T	s_3	a_3
T	F	F	s_4	a_4
F	T	T	s_5	a_5
F	T	F	s_6	a_6
F	F	T	s_7	a_7
F	F	F	s_8	a_8



- By enforcing the *extra-systematic constraint* $\text{Pr}(X \& Y \& \sim Z) = 0$, we can investigate features of our extra-systematic (*monadic-predicate-logical*) interpretation of X, Y , and Z , using only *sentential* probability calculus.
- This very useful "trick" will be used throughout the course.

Axiomatic Treatment of Probability Calculus II

- Instead of using the algebraic approach for proving theorems, we can also give *axiomatic* proofs. This is the standard way of proving claims in probability calculus (PrSAT doesn't give proofs, so we need axioms).
- Here are two examples of theorems and their *axiomatic* proofs (see the Eells *Appendix*). Note: these are *trivial* from an *algebraic* point of view!

Theorem ①. $\text{Pr}(\sim p) = 1 - \text{Pr}(p)$.

Proof. Since $p \vee \sim p$ is a tautology, (2) implies $\text{Pr}(p \vee \sim p) = 1$; and since p and $\sim p$ are m.e., (3) implies $\text{Pr}(p \vee \sim p) = \text{Pr}(p) + \text{Pr}(\sim p)$. Therefore, $1 = \text{Pr}(p) + \text{Pr}(\sim p)$, and thus $\text{Pr}(\sim p) = 1 - \text{Pr}(p)$, by simple algebra. \square

Theorem ②. If $p \models q$, then $\text{Pr}(p) = \text{Pr}(q)$. *Proof.* Assume $p \models q$. Then, p and $\sim q$ are mutually exclusive (inconsistent), and $p \vee \sim q \models \top$. So by axioms (2) and (3), and the previous theorem [$\text{Pr}(\sim p) = 1 - \text{Pr}(p)$]:

$$1 = \text{Pr}(p \vee \sim q) = \text{Pr}(p) + \text{Pr}(\sim q) = \text{Pr}(p) + 1 - \text{Pr}(q)$$

So, $1 = \text{Pr}(p) + 1 - \text{Pr}(q)$, and $0 = \text{Pr}(p) - \text{Pr}(q)$. $\therefore \text{Pr}(p) = \text{Pr}(q)$. \square

Axiomatic Treatment of Probability Calculus III

- Here are two more axiomatic proofs:

Theorem ③. If $p \models \perp$, then $\Pr(p) = 0$.

Proof. Assume $p \models \perp$. Then, $\sim p \models \top$, and, by (2), $\Pr(\sim p) = 1$. Then, by the above theorem, $\Pr(\sim p) = 1 - \Pr(p) = 1$, and $\Pr(p) = 0$. \square

Theorem ④. If $p \models q$, then $\Pr(p) \leq \Pr(q)$.

Proof. First, note the following two Boolean equivalences:

$$p \models (p \& q) \vee (p \& \sim q)$$

$$q \models (p \& q) \vee (\sim p \& q)$$

Thus, by our theorem above, we must have the following two identities:

$$\Pr(p) = \Pr[(p \& q) \vee (p \& \sim q)]$$

$$\Pr(q) = \Pr[(p \& q) \vee (\sim p \& q)]$$

By axiom (3), this yields the following two identities:

$$\Pr(p) = \Pr(p \& q) + \Pr(p \& \sim q)$$

$$\Pr(q) = \Pr(p \& q) + \Pr(\sim p \& q)$$

Now, assume $p \models q$. Then, $p \& \sim q \models \perp$. Hence, by our theorem above, $\Pr(p \& \sim q) = 0$. And, under these circumstances, we must have:

$$\Pr(p) = \Pr(p \& q)$$

$$\Pr(q) = \Pr(p \& q) + \Pr(\sim p \& q)$$

That is to say, we must have the following:

$$\Pr(q) = \Pr(p) + \Pr(\sim p \& q)$$

But, by axiom (1), $\Pr(\sim p \& q) \geq 0$. So, by algebra, $\Pr(q) \geq \Pr(p)$. \square

- This gives us an alternative way to prove $p \models q \implies \Pr(p) = \Pr(q)$. We just apply the previous theorem, in both directions (plus algebra).
- You should now be able to prove that $\Pr(p) \in [0, 1]$, for all p .

Axiomatic Treatment of Probability Calculus IV

- Skyrms uses a different-looking set of axioms from ours. One of our first exercises will be to show that Skyrms's axioms are equivalent to ours. Here are the six axioms Skyrms adopts (he calls them "rules").

- If a statement is a tautology, then its probability is equal to 1.
- If a statement is a self-contradiction, then its probability is equal to 0.
- If two statements are logically equivalent, they have the same probability.
- If p and q are mutually exclusive, then $\Pr(p \vee q) = \Pr(p) + \Pr(q)$.
- $\Pr(\sim p) = 1 - \Pr(p)$
- $\Pr(p \vee q) = \Pr(p) + \Pr(q) - \Pr(p \& q)$

- Skyrms uses the same definition of conditional probability as we do, namely, that $\Pr(p | q) = \frac{\Pr(p \& q)}{\Pr(q)}$, if $\Pr(q) > 0$ (otherwise, it's undefined).
- The following problem(s) are good exercises for axiomatic reasoning: Demonstrate that Skyrms's axiomatization is equivalent to ours. This requires giving several proofs, but you've already seen some of them.

Axiomatic Treatment of Probability Calculus V

- When it comes to unconditional probability, there is widespread agreement about the axioms we have adopted. Not so, when it comes to *conditional* probability. Some people think, for instance, that $\Pr(p | p)$ should be equal to 1, *even if* $\Pr(p) = 0$. Our theory disagrees.
- For this reason, some people have opted to take *conditional* probability as primitive and then to define *unconditional* probability in terms of it. See Hájek's paper "What Conditional Probability Could Not Be" (on website).
- There are various ways of doing this. These alternative axiomatizations tend to differ on how they treat $\Pr(p | q)$ when q is a *contradiction*. Most theories of conditional probability agree that: $\Pr(p | \perp) = 1$, for all p .
- This is a carry-over from classical deductive logic, which says that $\perp \models p$, for all p . Indeed, classical "logical" interpretations of (conditional) probability go hand-in-hand with this kind of assumption. We'll come back to this in the sections on "logical" probability and inductive logic.

Axiomatic Treatment of Probability Calculus VI

- There are various axiomatizations of conditional probability $\Pr(p \mid q)$. Unfortunately, many of them are *inconsistent*, because they are not careful about the $q = \perp$ case. Salmon & Earman (which we'll read later) use:

(A1) $\Pr(p \mid q) \in [0, 1]$.

(A2) If $q = p$, then $\Pr(p \mid q) = 1$.

(A3) If p and q are mutually exclusive, then $\Pr(p \vee q \mid r) = \Pr(p \mid r) + \Pr(q \mid r)$.

(A4) $\Pr(p \& q \mid r) = \Pr(p \mid r) \cdot \Pr(q \mid r \& p)$.

- That Salmon's (A1)-(A3) are inconsistent can be shown, as follows:
 - By (A2), $\Pr(p \mid \perp) = 1$, and $\Pr(\sim p \mid \perp) = 1$.
 - By (A3), $\Pr(p \vee \sim p \mid \perp) = \Pr(p \mid \perp) + \Pr(\sim p \mid \perp)$.
 - Therefore, $\Pr(p \vee \sim p \mid \perp) = 1 + 1 = 2$, which contradicts (A1). \square
- Consistency can be restored, by amending (A2) and/or (A3) in various ways. Different ways of doing this lead to different axiomatizations.
- See Hájek and Goosens (on website) for discussions of axiomatizing CP.

Axiomatic Treatment of Probability Calculus VII

- Exercise: Prove *axiomatically* that $\Pr(\cdot \mid q)$ is a probability function, assuming our axioms for $\Pr(\cdot)$ and our definition of $\Pr(\cdot \mid \cdot)$, i.e., prove that $\Pr(\cdot \mid q)$ satisfies our axioms for $\Pr(\cdot)$, for all q such that $\Pr(q) > 0$.
 - Axiom (1) requires $\Pr(p \mid q) \geq 0$. By definition, $\Pr(p \mid q) \stackrel{\text{def}}{=} \frac{\Pr(p \& q)}{\Pr(q)}$. Since both $\Pr(p \& q) \geq 0$ and $\Pr(q) > 0$, it follows that $\Pr(p \mid q) \geq 0$. \square
 - Axiom (2) requires $\Pr(\top \mid q) = 1$. By definition, $\Pr(\top \mid q) \stackrel{\text{def}}{=} \frac{\Pr(\top \& q)}{\Pr(q)}$. By theorem ②, $\Pr(\top \& q) = \Pr(q)$. Thus, $\Pr(\top \mid q) = \frac{\Pr(q)}{\Pr(q)} = 1$. \square
 - Axiom (3) requires $\Pr(p \vee r \mid q) = \Pr(p \mid q) + \Pr(r \mid q)$, if p and r are mutually exclusive. Assume p and r are mutually exclusive. Then:

$$\begin{aligned} \Pr(p \vee r \mid q) &= \frac{\Pr[(p \vee r) \& q]}{\Pr(q)} \text{ [Definition of } \Pr(\cdot \mid \cdot)\text{]} \\ &= \frac{\Pr[(p \& q) \vee (r \& q)]}{\Pr(q)} \text{ [Theorem ②]} \\ &= \frac{\Pr(p \& q) + \Pr(r \& q)}{\Pr(q)} \text{ [Axiom (3)]} \\ &= \Pr(p \mid q) + \Pr(r \mid q) \text{ [Definition of } \Pr(\cdot \mid \cdot)\text{]} \quad \square \end{aligned}$$

- So, our definition of $\Pr(\cdot \mid q)$ yields a *probability* function, for all q such that $\Pr(q) > 0$. This is also true for all non-orthodox axiomatizations of $\Pr(\cdot \mid \cdot)$. But, those approaches can *fail* this requirement when $\Pr(q) = 0$.
- E.g., $\Pr(\cdot \mid \perp)$ will *not* be a probability function on many non-orthodox accounts of $\Pr(\cdot \mid \cdot)$, since many of them assume $\Pr(p \mid \perp) = 1$, for *all* p .
- As in Salmon's theory, this leads to $\Pr(p \vee \sim p \mid \perp) = 2$, if we enforce axiom (3) on $\Pr(\cdot \mid \perp)$. Thus, non-orthodox $\Pr(\cdot \mid \perp)$'s are not probabilities.

Independence, Correlation, and Anti-Correlation 1

Definition. p and q are probabilistically independent ($p \perp q$) in a Pr-model \mathcal{M} if $\mathcal{M} = \langle \mathcal{B}, \Pr \rangle$ is such that: $\Pr(p \& q) = \Pr(p) \cdot \Pr(q)$.

- If $\Pr(p) > 0$ and $\Pr(q) > 0$, we can express independence also as follows:
 - * $\Pr(p \mid q) = \Pr(p)$ [Why? Because this is just: $\frac{\Pr(p \& q)}{\Pr(q)} = \Pr(p)$]
 - * $\Pr(q \mid p) = \Pr(q)$ [ditto.]
 - * $\Pr(p \mid q) = \Pr(p \mid \sim q)$ [Not as obvious. See next slide.]
 - * $\Pr(q \mid p) = \Pr(q \mid \sim p)$ [ditto.]
- Exercise: prove this! Closely related fact about independence. If $p \perp q$, then we also must have: $p \perp \sim q$, $q \perp \sim p$, and $\sim p \perp \sim q$. Prove this too!
- A set of propositions $\mathbf{P} = \{p_1, \dots, p_n\}$ is *mutually independent* if all subsets $\{p_i, \dots, p_j\} \subseteq \mathbf{P}$ are s.t. $\Pr(p_i \& \dots \& p_j) = \Pr(p_i) \cdot \dots \cdot \Pr(p_j)$. For sets with 2 propositions, pairwise independence is equivalent to mutual independence. But, not for 3 or more propositions. Example given below.

- Here's an axiomatic proof that $\Pr(p \mid q) = \Pr(p) \cdot \Pr(q) \Leftrightarrow \Pr(p \mid q) = \Pr(p \mid \sim q)$, provided that that $\Pr(q) \in (0, 1)$:

$$\begin{aligned} \Pr(p \mid q) = \Pr(p \mid \sim q) &\Leftrightarrow \frac{\Pr(p \& q)}{\Pr(q)} = \frac{\Pr(p \& \sim q)}{\Pr(\sim q)} \text{ [definition of CP]} \\ \Leftrightarrow \frac{\Pr(p \& q)}{\Pr(q)} - \frac{\Pr(p \& \sim q)}{\Pr(\sim q)} &= 0 \text{ [algebra]} \\ \Leftrightarrow \frac{\Pr(p \& q) \cdot \Pr(\sim q) - \Pr(p \& \sim q) \cdot \Pr(q)}{\Pr(q) \Pr(\sim q)} &= 0 \text{ [algebra]} \\ \Leftrightarrow \Pr(p \& q) \cdot \Pr(\sim q) - \Pr(p \& \sim q) \cdot \Pr(q) &= 0 \text{ [algebra]} \\ \Leftrightarrow \Pr(p \& q) \cdot (1 - \Pr(q)) - \Pr(p \& \sim q) \cdot \Pr(q) &= 0 \text{ [algebra]} \\ \Leftrightarrow \Pr(p \& q) - \Pr(q) \cdot [\Pr(p \& q) + \Pr(p \& \sim q)] &= 0 \text{ [algebra]} \\ \Leftrightarrow \Pr(p \& q) - \Pr(q) \cdot \Pr((p \& q) \vee (p \& \sim q)) &= 0 \text{ [additivity axiom]} \\ \Leftrightarrow \Pr(p \& q) - \Pr(q) \cdot \Pr(p) &= 0 \text{ [Theorem ②]} \\ \Leftrightarrow \Pr(p \& q) = \Pr(p) \cdot \Pr(q) &\text{ [algebra]} \quad \square \end{aligned}$$

- A purely algebraic proof of this theorem can be obtained rather easily:

- More generally, if $\{p, q, r\}$ are mutually independent, then p is independent of any propositional function of q and r , e.g., $p \perp\!\!\!\perp q \vee r$.

Proof. $\Pr(p \& (q \vee r)) = \Pr((p \& q) \vee (p \& r))$ [Theorem ②]

$$\begin{aligned} &= \Pr(p \& q) + \Pr(p \& r) - \Pr(p \& q \& r) \text{ [general additivity]} \\ &= \Pr(p) \cdot \Pr(q) + \Pr(p) \cdot \Pr(r) - \Pr(p) \cdot \Pr(q \& r) \text{ [mutual } \perp\!\!\!\perp] \\ &= \Pr(p) \cdot [\Pr(q) + \Pr(r) - \Pr(q \& r)] \text{ [algebra]} \\ &= \Pr(p) \cdot \Pr(q \vee r) \text{ [general additivity]} \quad \square \end{aligned}$$

- This last proof makes heavy use of general additivity (Skyrms's rule 6): $\Pr(p \vee q) = \Pr(p) + \Pr(q) - \Pr(p \& q)$. This is one of our first exercises involving axiomatic proof (to prove this rule from our axioms).
- How might one prove the more general theorem above: that if $\{p, q, r\}$ are mutually independent, then p is independent of any propositional function of q and r ? And, is there an even more general theorem here?
- To wit: is it the case that if $\mathbf{P} = \{p_1, \dots, p_n\}$ is a mutually independent set, then any p -functions of any two disjoint subsets of \mathbf{P} are independent?

p	q	States	$\Pr(s_i)$
T	T	s_1	a_1
T	F	s_2	a_2
F	T	s_3	a_3
F	F	s_4	$a_4 = 1 - (a_1 + a_2 + a_3)$

$$\begin{aligned} \therefore \Pr(p \mid q) = \Pr(p \mid \sim q) &\Leftrightarrow \frac{a_1}{a_1 + a_3} = \frac{a_2}{a_2 + a_4} = \frac{a_2}{1 - (a_1 + a_3)} \\ &\Leftrightarrow a_1 \cdot (1 - (a_1 + a_3)) = a_2 \cdot (a_1 + a_3) \\ &\Leftrightarrow a_1 = a_2 \cdot (a_1 + a_3) + a_1 \cdot (a_1 + a_3) = (a_2 + a_1) \cdot (a_1 + a_3) \\ &\Leftrightarrow \Pr(p \& q) = \Pr(p) \cdot \Pr(q) \quad \square \end{aligned}$$

- If p and q are independent, then so are p and $\sim q$. Pretty easy proof. Assume that p and q are independent. Then, since $p \models (p \& q) \vee (p \& \sim q)$

$$\begin{aligned} \Pr(p) &= \Pr(p \& q) + \Pr(p \& \sim q) \text{ [Th. ② and additivity axiom]} \\ &= \Pr(p) \cdot \Pr(q) + \Pr(p \& \sim q) \text{ [independence]} \end{aligned}$$

So, $\Pr(p \& \sim q) = \Pr(p) \cdot [1 - \Pr(q)] = \Pr(p) \cdot \Pr(\sim q)$ [algebra & Th. ①] \square