

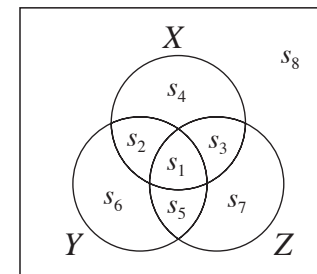
Overview of Finite Propositional Boolean Algebras I

- Consider a logical language \mathcal{L} containing n atomic sentences. These may be sentence letters ($X, Y, Z, \text{etc.}$), or they may be atomic sentences of monadic or relational predicate calculus ($Fa, Gb, Rab, Hcd, \text{etc.}$).
- The Boolean Algebra $\mathcal{B}_{\mathcal{L}}$ set-up by such a language will be such that:
 - $\mathcal{B}_{\mathcal{L}}$ will have 2^n states (corresponding to the state descriptions of \mathcal{L})
 - $\mathcal{B}_{\mathcal{L}}$ will contain 2^{2^n} propositions, in total.
 - This is because each proposition p in $\mathcal{B}_{\mathcal{L}}$ is equivalent to a disjunction of state descriptions. Thus, each subset of the set of state descriptions of \mathcal{L} corresponds to a proposition of $\mathcal{B}_{\mathcal{L}}$.
 - Note: there are 2^{2^n} subsets of a set of size 2^n .
 - The empty set \emptyset of state descriptions corresponds to “the empty disjunction”, which corresponds to the logical falsehood: \perp .
 - Singelton sets of state descriptions correspond to “disjunctions with one member”. [All other subsets are “normal” disjunctions.]

Overview of Finite Propositional Boolean Algebras II

- Example. Let \mathcal{L} have three atomic sentences: $X, Y,$ and Z . Then, $\mathcal{B}_{\mathcal{L}}$ is:

X	Y	Z	States
T	T	T	s_1
T	T	F	s_2
T	F	T	s_3
T	F	F	s_4
F	T	T	s_5
F	T	F	s_6
F	F	T	s_7
F	F	F	s_8



- Examples of reduction to disjunctions of state descriptions of \mathcal{L} :
 - $\neg X \ \& \ \sim X$ is equivalent to the empty disjunction: \perp .
 - $\neg X \ \& \ (\sim Y \ \& \ Z)$ is equivalent to the singleton disjunction: s_3 .
 - $\neg X \equiv (Y \vee Z)$ is equivalent to: $s_1 \vee s_2 \vee s_3 \vee s_8$.
- In general: $p \models \bigvee \{s_i \mid s_i \models p\}$. And, if $\{s_i \mid s_i \models p\} = \emptyset$, then $p \models \perp$.

The Probability Calculus: An Algebraic Approach I

- Once we grasp the concept of a finite Boolean algebra of propositions, understanding the probability calculus algebraically is very easy.
- The central concept is a finite probability model. A finite probability model \mathcal{M} is a finite Boolean algebra of propositions \mathcal{B} , together with a function $\text{Pr}(\cdot)$ which maps elements of \mathcal{B} to the unit interval $[0, 1] \in \mathbb{R}$.
- This function $\text{Pr}(\cdot)$ must be a probability function. It turns out that a probability function $\text{Pr}(\cdot)$ on \mathcal{B} is just a function that assigns a real number on $[0, 1]$ to each state s_i of \mathcal{B} , such that $\sum_i \text{Pr}(s_i) = 1$.
- Once we have $\text{Pr}(\cdot)$'s basic assignments to the states of \mathcal{B} (s.d.'s of \mathcal{L}), we define $\text{Pr}(p)$ for any statement \mathcal{L} of the language of \mathcal{B} , as follows:

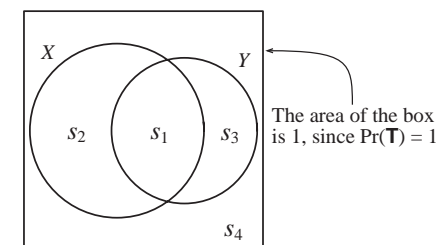
$$\text{Pr}(p) = \sum_{s_i \models p} \text{Pr}(s_i) \quad [\text{note: if } p \models \perp, \text{ then } \text{Pr}(p) = 0]$$

- In other words, $\text{Pr}(p)$ is the sum of the probabilities of the state descriptions in p 's (equivalent) disjunction of state descriptions.

The Probability Calculus: An Algebraic Approach II

- Here's an example of a finite probability model \mathcal{M} , whose algebra \mathcal{B} is characterized by a language \mathcal{L} with two atomic letters “ X ” and “ Y ”:

X	Y	States	$\text{Pr}(s_i)$
T	T	s_1	$\frac{1}{6}$
T	F	s_2	$\frac{1}{4}$
F	T	s_3	$\frac{1}{8}$
F	F	s_4	$\frac{11}{24}$



- On the left, a stochastic truth-table (STT) representation of \mathcal{M} ; on the right, a stochastic Venn Diagram (SVD) representation, in which area is proportional to probability. This is a regular model: $\text{Pr}(s_i) > 0$, for all i .
- \mathcal{M} determines a numerical probability for each p in \mathcal{L} . Examples?
- We can also use STTs to furnish an algebraic method for proving general facts about all probability models — the algebraic method.

The Probability Calculus: An Algebraic Approach III

- Let $a_i = \Pr(s_i)$ be the probability [under the probability assignment $\Pr(\cdot)$] of state s_i in \mathcal{B} — i.e., the area of region s_i in our SVD.
- Once we have real variables (a_i) for each of the basic probabilities, we can not only calculate probabilities relative to *specific* numerical models — we can say **general things**, using only simple high-school algebra.
- That is, we can *translate* any expression ' $\Pr(p)$ ' into a *sum* of some of the a_i , and thus we can *reduce probabilistic* claims about the p 's in \mathcal{B}/\mathcal{L} into simple, high-school-algebraic claims about the real variables a_i .
- This allows us to be able to prove general claims about *probability functions*, by proving their corresponding *algebraic theorems*.
- Method: translate the probability claim into a claim involving sums of the a_i , and determine whether the corresponding claim is a theorem of algebra (assuming only that the a_i are on $[0, 1]$ and that they sum to 1).

The Probability Calculus: An Algebraic Approach IV

- Here are two simple/obvious examples involving two atomic sentences:

Theorem. $\Pr(X \vee Y) = \Pr(X) + \Pr(Y) - \Pr(X \& Y)$.

Proof. $\Pr(X \vee Y) = a_1 + a_2 + a_3 = (a_1 + a_2) + (a_1 + a_3) - a_1$.

Theorem. $\Pr(X) = \Pr(X \& Y) + \Pr(X \& \sim Y)$.

Proof. $a_1 + a_2 = a_1 + a_2$.

- Here are two general facts that are also obvious from the set-up:

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

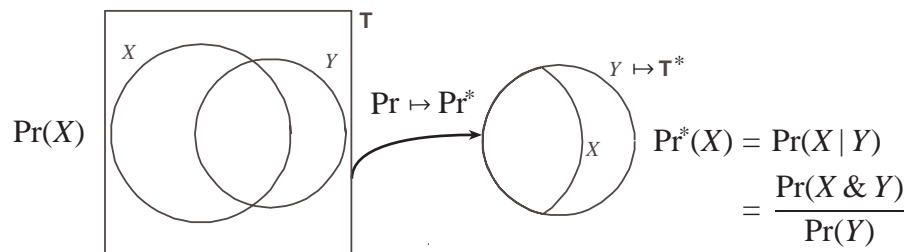
Proof. Obvious, since the same regions always have the same areas, and the algebraic translation is *the same* for logically equivalent p/q .

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

Proof. Since $p \models q$, the set of state descriptions entailing p is a subset of the set of state descriptions entailing q . Thus, the set of a_i in the summation for $\Pr(p)$ will be a subset of the a_i in the summation for $\Pr(q)$. Thus, since all the $a_i \geq 0$, $\Pr(p) \leq \Pr(q)$.

The Probability Calculus: An Algebraic Approach V

- **Conditional Probability.** $\Pr(p | q) \stackrel{\text{def}}{=} \frac{\Pr(p \& q)}{\Pr(q)}$, provided that $\Pr(q) > 0$.
- Intuitively, $\Pr(p | q)$ is supposed to be the probability of p **given that q is true**. So, *conditionalizing* on q is like “supposing q to be true”.
- Using Venn diagrams, we can explain: “Supposing Y to be true” is like “treating the Y -circle as if it is the bounding box of the Venn Diagram”.
- This is like “moving to a new $\Pr^*(\cdot)$ such that $\Pr^*(Y) = 1$.” Picture:



The Probability Calculus: An Algebraic Approach VI

- There may be other ways of defining conditional probability, which may also seem to capture the “supposing q to be true” intuition.
- But, any such definition must make $\Pr(\cdot | q)$ itself a *probability function*, for all propositions q . This proves to be quite a strong constraint.
- Algebraically, we can see just how strong this constraint is. Recall:

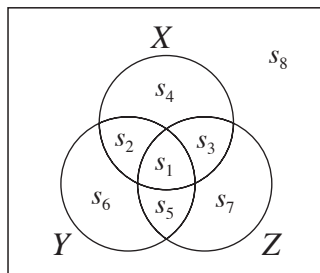
$$\Pr(X \vee Y) = \Pr(X) + \Pr(Y) - \Pr(X \& Y)$$

- Therefore, if $\Pr(\cdot | q)$ is to be a *probability function* for all q , then we must also have the following equality (in general), for all Z :

$$\Pr(X \vee Y | Z) = \Pr(X | Z) + \Pr(Y | Z) - \Pr(X \& Y | Z)$$

- Using our algebraic method, we can *prove* this. We just need to remind ourselves of what the 3-atomic sentence algebra looks like, and how the algebraic translation of this equation would go. Let’s do that ...

X	Y	Z	States	Pr(s_i)
T	T	T	s_1	a_1
T	T	F	s_2	a_2
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 our definition of conditional probability, we have:

$$\Pr(X \vee Y | Z) = \frac{\Pr((X \vee Y) \& Z)}{\Pr(Z)} = \frac{\Pr((X \& Z) \vee (Y \& Z))}{\Pr(Z)} = \frac{a_1 + a_3 + a_5}{a_1 + a_3 + a_5 + a_7}$$

and

$$\begin{aligned} \Pr(X | Z) + \Pr(Y | Z) - \Pr(X \& Y | Z) &= \frac{\Pr(X \& Z)}{\Pr(Z)} + \frac{\Pr(Y \& Z)}{\Pr(Z)} - \frac{\Pr(X \& Y \& Z)}{\Pr(Z)} \\ &= \frac{\Pr(X \& Z) + \Pr(Y \& Z) - \Pr(X \& Y \& Z)}{\Pr(Z)} \\ &= \frac{(a_1 + a_3) + (a_1 + a_5) - a_1}{a_1 + a_3 + a_5 + a_7} = \frac{a_1 + a_3 + a_5}{a_1 + a_3 + a_5 + a_7} \end{aligned}$$

The Probability Calculus: An Algebraic Approach VII

- We can use our algebraic method to demonstrate that our definition of $\Pr(\cdot | q)$ yields a probability function, *for all* q , in the following way.
- Intuitively, think about what an “unconditional” and a “conditional” stochastic truth-table must look like, for any pair of sentences p and q .

p	q	Pr(s_i)	$\cdot q$	p	q	Pr($s_i q$)
T	T	a_1	\rightarrow	T	T	$\Pr(s_1 q) \stackrel{\text{def}}{=} \frac{\Pr(s_1 \& q)}{\Pr(q)} = \frac{a_1}{a_1 + a_3}$
T	F	a_2		T	F	$\Pr(s_2 q) \stackrel{\text{def}}{=} \frac{\Pr(s_2 \& q)}{\Pr(q)} = 0$
F	T	a_3		F	T	$\Pr(s_3 q) \stackrel{\text{def}}{=} \frac{\Pr(s_3 \& q)}{\Pr(q)} = \frac{a_3}{a_1 + a_3}$
F	F	a_4		F	F	$\Pr(s_4 q) \stackrel{\text{def}}{=} \frac{\Pr(s_4 \& q)}{\Pr(q)} = 0$

- Note: the new basic probabilities assigned to the state descriptions, under our “conditionalized” $\Pr(\cdot | q)$ satisfy the requirements for being a *probability* function, since $\frac{a_1}{a_1 + a_3} + \frac{a_3}{a_1 + a_3} = 1$, and $\frac{a_1}{a_1 + a_3}, \frac{a_3}{a_1 + a_3} \in [0, 1]$.

The Probability Calculus: An Algebraic Approach VIII

- Here’s a neat theorem of the probability calculus, proved algebraically.

Theorem. $\Pr(X \supset Y) \geq \Pr(Y | X)$. [Provided that $\Pr(X) > 0$, of course.]

Proof. $\Pr(X \supset Y) = \Pr(\sim X \vee Y) = \Pr(s_1 \vee s_3 \vee s_4) = a_1 + a_3 + a_4$.

$$\Pr(Y | X) = \frac{\Pr(Y \& X)}{\Pr(X)} = \frac{\Pr(s_1)}{\Pr(s_1 \vee s_2)} = \frac{a_1}{a_1 + a_2}$$

So, we need to prove that $a_1 + a_3 + a_4 \geq \frac{a_1}{a_1 + a_2}$.

- First, note that $a_4 = 1 - (a_1 + a_2 + a_3)$, since the a_i ’s must sum to 1.
- Thus, we need to show that $a_1 + a_3 + 1 - a_1 - a_2 - a_3 \geq \frac{a_1}{a_1 + a_2}$.
- By simple algebra, this reduces to showing that $1 - a_2 \geq \frac{a_1}{a_1 + a_2}$.
- If $a_1 + a_2 > 0$ and $a_i \in [0, 1]$, this must hold, since then we must have: $a_2 \geq a_2 \cdot (a_1 + a_2)$, and then the boxed formulas are equivalent. \square

The Probability Calculus: An Algebraic Approach IX

- Here are some further fundamental theorems of probability calculus, involving 2 or 3 atomic sentences and CP. Easy, given defn. of CP.

- **The Law of Total Probability (LTP):**

$$\Pr(X | Y) = \Pr(X | Y \& Z) \cdot \Pr(Z | Y) + \Pr(X | Y \& \sim Z) \cdot \Pr(\sim Z | Y)$$

- Note: $\Pr(X | \top) = \Pr(X)$. Why? So, the LTP has a *special case*:

$$\begin{aligned} \Pr(X | \top) &= \Pr(X) = \Pr(X | \top \& Z) \cdot \Pr(Z | \top) + \Pr(X | \top \& \sim Z) \cdot \Pr(\sim Z | \top) \\ &= \Pr(X | Z) \cdot \Pr(Z) + \Pr(X | \sim Z) \cdot \Pr(\sim Z) \end{aligned}$$

- Two forms of **Bayes’s Theorem**. The second one *follows*, using (LTP):

$$\begin{aligned} \Pr(X | Y) &= \frac{\Pr(Y | X) \cdot \Pr(X)}{\Pr(Y)} \\ &= \frac{\Pr(Y | Z) \cdot \Pr(Z) + \Pr(Y | \sim Z) \cdot \Pr(\sim Z)}{\Pr(Y | Z) \cdot \Pr(Z) + \Pr(Y | \sim Z) \cdot \Pr(\sim Z)} \end{aligned}$$

The Probability Calculus: An Algebraic Approach X

- One more interesting theorem (due to Popper & Miller), algebraically.
- Let $d(X, Y) \stackrel{\text{def}}{=} \Pr(X | Y) - \Pr(X)$. Then, we have the following theorem:

Theorem (PM). $d(X, Y) = d(X \vee Y, Y) + d(X \vee \sim Y, Y)$.

Proof (algebraic, using STT from X/Y language, above).

$$d(X, Y) \stackrel{\text{def}}{=} \Pr(X | Y) - \Pr(X) = \frac{a_1}{a_1 + a_3} - (a_1 + a_2)$$

$$d(X \vee Y, Y) \stackrel{\text{def}}{=} \Pr(X \vee Y | Y) - \Pr(X \vee Y) = 1 - a_1 - a_2 - a_3$$

$$d(X \vee \sim Y, Y) \stackrel{\text{def}}{=} \Pr(X \vee \sim Y | Y) - \Pr(X \vee \sim Y) = \frac{a_1}{a_1 + a_3} - (a_1 + a_2 + a_4)$$

$$\begin{aligned} \therefore d(X \vee Y, Y) + d(X \vee \sim Y, Y) &= 1 - a_1 - a_2 - a_3 + \frac{a_1}{a_1 + a_3} - a_1 - a_2 - a_4 \\ &= \frac{a_1}{a_1 + a_3} + 1 - a_1 - a_2 - a_3 - a_1 - a_2 - (1 - (a_1 + a_2 + a_3)) \\ &= \frac{a_1}{a_1 + a_3} - (a_1 + a_2). \quad \square \end{aligned}$$

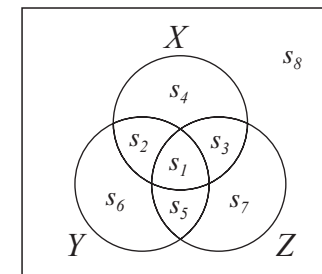
The Probability Calculus: An Algebraic Approach XII

- 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 created a *Mathematica* notebook which contains some examples from algebraic probability calculus that we see in this lecture.
- Let's have a look at this first notebook (`examples_lmu.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.

Probabilistic Independence

Definition. p and q are *probabilistically independent*, given r ($p \perp\!\!\!\perp q \mid r$) iff $\Pr(p \& q \mid r) = \Pr(p \mid r) \cdot \Pr(q \mid r)$. [Note: We will use ' $p \perp\!\!\!\perp q$ ' as an abbreviation for ' $p \perp\!\!\!\perp q \mid \top$ '.]

- If $\Pr(p) > 0$ and $\Pr(q) > 0$, then $p \perp\!\!\!\perp q$ is equivalent to all of the following:
 - * $\Pr(p \mid q) = \Pr(p)$ [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.]
- Closely related fact about independence. If $p \perp\!\!\!\perp q$, then we also must have: $p \perp\!\!\!\perp \sim q$, $q \perp\!\!\!\perp \sim p$, and $\sim p \perp\!\!\!\perp \sim q$. See next slide for algebraic set-up.
- 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.

- To wit: is it the case that if $\mathbf{P} = \{p_1, \dots, p_n\}$ is a mutually independent set, then *any* β -functions of any two disjoint subsets of \mathbf{P} are independent?
- So far, we've seen a some *proofs* of *true* general claims about independence, correlation, etc. Now, for some *counterexamples!*
- As always, these are numerical probability models in which some claim *fails*. We have seen two false claims about $\perp\!\!\!\perp$ already. Let's prove them.
- **Theorem.** Pairwise independence of a collection of three propositions $\{X, Y, Z\}$ does not entail mutual independence of the collection. That is to say, there exist probability models in which (1) $\Pr(X \& Y) = \Pr(X) \cdot \Pr(Y)$, (2) $\Pr(X \& Z) = \Pr(X) \cdot \Pr(Z)$, (3) $\Pr(Y \& Z) = \Pr(Y) \cdot \Pr(Z)$, but (4) $\Pr(X \& Y \& Z) \neq \Pr(X) \cdot \Pr(Y) \cdot \Pr(Z)$. *Proof.* Here's a counterexample.
- Suppose a box contains 4 tickets labelled with the following numbers:
 112, 121, 211, 222
 Let us choose one ticket at random (*i.e.*, each ticket has an *equal* probability of being chosen), and consider the following propositions:

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 $\sim q$. Prove this algebraically.
- More generally, if $\{p, q, r\}$ are mutually independent, then p is independent of *any* Boolean function β of q and r , *e.g.*, $p \perp\!\!\!\perp q \vee r$.
- How might one prove this more general theorem? And, is there an even more general theorem to be proved here?

$X = "1"$ occurs at the first place of the chosen ticket.
 $Y = "1"$ occurs at the second place of the chosen ticket.
 $Z = "1"$ occurs at the third place of the chosen ticket.

Since the ticket #'s are 112, 121, 211, 222, we have these probabilities:

$$\Pr(X) = \frac{1}{2}, \Pr(Y) = \frac{1}{2}, \Pr(Z) = \frac{1}{2}$$

Moreover, each of the three conjunctions determines a unique ticket #:

$X \& Y$ = the ticket is labeled #112
 $X \& Z$ = the ticket is labeled #121
 $Y \& Z$ = the ticket is labeled #211

Therefore, since each ticket is equally probable to be chosen, we have:

$$\Pr(X \& Y) = \Pr(X \& Z) = \Pr(Y \& Z) = \frac{1}{4}$$

So, the three events X, Y, Z are pairwise independent (*why?*). But,

$X \& Y \& Z \neq \perp$, since $X, Y,$ and Z are jointly inconsistent.

Hence,

$$\Pr(X \& Y \& Z) = \Pr(F) = 1 - \Pr(T) = 0 \neq \Pr(X) \cdot \Pr(Y) \cdot \Pr(Z) = \left(\frac{1}{2}\right)^3 = \frac{1}{8}$$

- This information determines a *unique* probability function. Can you specify it? Algebra (7 equations, 7 unknowns — see STT below).

$$\Pr(X) = a_4 + a_2 + a_3 + a_1 = \frac{1}{2}, \Pr(Y) = a_2 + a_6 + a_1 + a_5 = \frac{1}{2}$$

$$\Pr(Z) = a_3 + a_1 + a_5 + a_7 = \frac{1}{2}, \Pr(X \& Y \& Z) = a_1 = 0$$

$$\Pr(X \& Y) = a_2 + a_1 = \frac{1}{4}, \Pr(X \& Z) = a_3 + a_1 = \frac{1}{4}, \Pr(Y \& Z) = a_1 + a_5 = \frac{1}{4}$$

- Here's the STT. [This (and other models) can be found with PrSAT.]

X	Y	Z	States	Pr(s_i)
T	T	T	s_1	$\Pr(s_1) = a_1 = 0$
T	T	F	s_2	$\Pr(s_2) = a_2 = 1/4$
T	F	T	s_3	$\Pr(s_3) = a_3 = 1/4$
T	F	F	s_4	$\Pr(s_4) = a_4 = 0$
F	T	T	s_5	$\Pr(s_5) = a_5 = 1/4$
F	T	F	s_6	$\Pr(s_6) = a_6 = 0$
F	F	T	s_7	$\Pr(s_7) = a_7 = 0$
F	F	F	s_8	$\Pr(s_8) = a_8 = 1/4$

- Theorem.** \perp is *not* transitive. Example in which $\Pr(X \& Y) = \Pr(X) \cdot \Pr(Y)$, $\Pr(Y \& Z) = \Pr(Y) \cdot \Pr(Z)$, but $\Pr(X \& Z) \neq \Pr(X) \cdot \Pr(Z)$ [$X \neq Y \neq Z$]:

Measures of Confirmation I

- As I mentioned in my previous lectures this week, we can define a notion of “support” or “confirmation” as *probabilistic relevance*.

Definition. E confirms H iff $\Pr(H | E) > \Pr(H)$. E disconfirms H iff $\Pr(H | E) < \Pr(H)$. E is neutral/irrelevant to H iff $\Pr(H | E) = \Pr(H)$.

- Given this qualitative definition of “confirms”, it is natural to think about *quantitative measures of degree of confirmation*.
- This involves adopting some *function* c of $\Pr(H | E)$ and $\Pr(H)$.
- We will use the notation $c(H, E)$ to denote the degree to which E confirms H , according to some function c of $\Pr(H | E)$ and $\Pr(H)$.
- We'll adopt the following convention about the *range* of $c(H, E)$:

$$(R) \quad c(H, E) \in \begin{cases} (0, 1] & \text{if } \Pr(H | E) > \Pr(H), \\ \{0\} & \text{if } \Pr(H | E) = \Pr(H), \\ [-1, 0) & \text{if } \Pr(H | E) < \Pr(H). \end{cases}$$

X	Y	Z	States	Pr(s_i)
T	T	T	s_1	$\Pr(s_1) = a_1 = 3/32$
T	T	F	s_2	$\Pr(s_2) = a_2 = 9/32$
T	F	T	s_3	$\Pr(s_3) = a_3 = 3/32$
T	F	F	s_4	$\Pr(s_4) = a_4 = 9/32$
F	T	T	s_5	$\Pr(s_5) = a_5 = 2/32$
F	T	F	s_6	$\Pr(s_6) = a_6 = 2/32$
F	F	T	s_7	$\Pr(s_7) = a_7 = 2/32$
F	F	F	s_8	$\Pr(s_8) = a_8 = 2/32$

$$\Pr(X \& Y) = a_2 + a_1 = \frac{3}{8} = \frac{3}{4} \cdot \frac{1}{2}$$

$$= (a_4 + a_2 + a_3 + a_1) \cdot (a_2 + a_1 + a_6 + a_5) = \Pr(X) \cdot \Pr(Y)$$

$$\Pr(Y \& Z) = a_1 + a_5 = \frac{5}{32} = \frac{1}{2} \cdot \frac{5}{16}$$

$$= (a_2 + a_1 + a_6 + a_5) \cdot (a_3 + a_1 + a_5 + a_7) = \Pr(Y) \cdot \Pr(Z)$$

$$\Pr(X \& Z) = a_3 + a_1 = \frac{3}{16} \neq \frac{3}{4} \cdot \frac{5}{16}$$

$$= (a_4 + a_2 + a_3 + a_1) \cdot (a_3 + a_1 + a_5 + a_7) = \Pr(X) \cdot \Pr(Z)$$

Measures of Confirmation II

- A large number of measures of confirmation have been proposed in the literature (in statistics, cognitive science, philosophy, etc.). Here are the four most popular measures (*up to ordinal equivalence* — see below):

$$- d(H, E) \stackrel{\text{def}}{=} \Pr(H | E) - \Pr(H)$$

$$- r(H, E) \stackrel{\text{def}}{=} \frac{\Pr(H | E) - \Pr(H)}{\Pr(H | E) + \Pr(H)} \doteq \frac{\Pr(H | E)}{\Pr(H)}$$

$$- l(H, E) \stackrel{\text{def}}{=} \frac{\Pr(E | H) - \Pr(E | \sim H)}{\Pr(E | H) + \Pr(E | \sim H)} \doteq \frac{\Pr(H | E) \cdot (1 - \Pr(H))}{(1 - \Pr(H | E)) \cdot \Pr(H)}$$

$$- s(H, E) \stackrel{\text{def}}{=} \Pr(H | E) - \Pr(H | \sim E)$$

- If two measures c_1 and c_2 agree on *all comparisons*, then we say that c_1 and c_2 are *ordinally equivalent* ($c_1 \doteq c_2$). More precisely, we define:

$$c_1 \doteq c_2 \stackrel{\text{def}}{=} c_1(H_1, E_1) \geq c_1(H_2, E_2) \text{ iff and only if } c_2(H_1, E_1) \geq c_2(H_2, E_2)$$
- Exercises:** (i) prove that $\{d, r, l, s\}$ all satisfy (R), and (ii) prove the two “ \doteq ” claims about r and l , above. *Hint.* Use $\frac{x-y}{x+y} = \tanh\left[\frac{1}{2} \log\left(\frac{x}{y}\right)\right]$.

Measures of Confirmation III

- **Fact.** No two of $\{d, r, l, s\}$ are ordinally equivalent. [Use PrSAT!]
- This ordinal disagreement between the most popular measures is what I have called “the plurality of Bayesian measures of confirmation”.
- This was the topic of my dissertation [link on my mathcamp webpage].
- Here are eight important properties of measures of confirmation:
 - (1) If $E \models H_1$ and $E \not\models H_2$, then $c(H_1, E) \geq c(H_2, E)$.
 - (2) If $\Pr(E | H_1) > \Pr(E | H_2)$, then $c(H_1, E) > c(H_2, E)$.
 - (3) If $\Pr(H | E_1) > \Pr(H | E_2)$, then $c(H, E_1) > c(H, E_2)$.
 - (4) $c(H, E) = c(E, H)$.
 - (5) $c(H, E) = -c(H, \sim E)$.
 - (6) $c(H, E) = -c(\sim H, E)$.
 - (7) If $H \models E$, then $c(H, E) > c(H \& X, E)$, for any X .
 - (8) If $\Pr(E | H_1) > \Pr(E | H_2)$ & $\Pr(E | \sim H_1) \leq \Pr(E | \sim H_2)$, then $c(H_1, E) > c(H_2, E)$.

Measures of Confirmation IV

	Does c-Measure have property?							
c-Measures	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$d(H, E)$	NO	NO	YES	NO	NO	YES	YES	YES
$r(H, E)$	NO	YES	YES	YES	NO	NO	NO	YES
$l(H, E)$	YES	NO	YES	NO	NO	YES	YES	YES
$s(H, E)$	NO	NO	NO	NO	YES	YES	YES	YES

- One can settle these (and many other) questions using PrSAT.
- Property (8) is the property that underlies the *robust* theorem about the conjunction fallacy that I discussed at the end of yesterday’s lecture. [It’s one of very few robust properties one finds in the literature.]
- **Exercise.** (iii) Define a relatively simple (\mathcal{R})-measure that *violates* (8).
 - *Hint.* Try $\Pr(H | E)^n - \Pr(H)^n$ for (any) $n > 1$. See:

<http://fitelson.org/crupi.pdf>