

# PROBABILITY

We shall be concerned with **probability** in much of the course: Bayesian nets, Markov chains, entropy, etc.

The reason is that probability is the main tool in dealing with **uncertainty**, and this is a central feature of how real cognitive beings reason and act.

This week will contain most of the background, with more later as needed.

In addition to the standard material I'll have much more *conceptual* discussion.

If this goes too fast, I recommend getting a textbook.

One possibility is the M118 book Finite Mathematics.

A more advanced treatment is in Mathematical Methods for AI, the textbook I used the last time.

## SOME PROBLEMS

If we flip a fair coin, the probability of Heads is  $1/2$ .

Suppose we flipped a coin 100 times. If the number of Heads is between 40 and 60, I'll give you \$10. If it's not, you give me \$1. Would it be smart to take this bet?

## SOME PROBLEMS

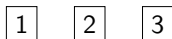
Suppose we flip a fair coin repeatedly, starting with flip 1.  
We stop when we get a Heads.

Either we stop on an even numbered flip or an odd one  
(or we go on forever).

What is the probability that we stop on an **even numbered flip**?

# MONTY HALL PROBLEM

Monty Hall stands with you in front of three doors.



One of these has a car, two are empty.

Monty knows which door has the car.

You pick one door.

Monty opens one door with nothing behind it.

He then offers you the chance to switch to the third door.

Question: Is it to your advantage to switch, or does it make no difference to you?

# MONTY HALL PROBLEM, CONTINUED

*One line of reasoning:*

The car is behind each door with probability  $1/3$ .

No matter where it is, and no matter what you pick, it's always possible to show an open door. Now that we have two closed doors with no additional knowledge, the probability that the car is behind each is  $1/2$ .

Therefore it cannot change the probability that the car is behind any given door.

So it cannot possibly help you to change doors.

# MONTY HALL PROBLEM, CONTINUED

*Another line of reasoning:*

Let's assume that *your first guess is random* as is *Monty's choice*.

Let's write the pairs as (reality, your guess).

(1, 1)	(1, 2)	(1, 3)
(2, 1)	(2, 2)	(2, 3)
(3, 1)	(3, 2)	(3, 3)

Each pair has probability  $1/9$ , and  $3 \times (1/9) = 1/3$ .

Your initial guess has  $1/3$  of a chance of being right.

The *no-switching strategy* succeeds  $1/3$  of the time.

If you switch, then a correct first guess becomes incorrect.

More importantly, an incorrect first guess becomes correct.

So the *switching strategy* succeeds  $2/3$  of the time.

# DISCRETE PROBABILITY SPACES

In these notes,  $\mathbb{R}$  is the set of real numbers, and  $\mathbb{R}^{\geq 0}$  is the set of non-negative reals.

A *probability space* is a set  $S$  and a function  $\Pr : S \rightarrow \mathbb{R}^{\geq 0}$

$$\sum_{s \in S} \Pr(s) = 1.$$

An *event* is a subset  $A \subseteq S$ .

For example,  $\emptyset$  and  $S$  are events.

We define

$$\Pr(A) = \sum_{s \in A} \Pr(s).$$

So  $\Pr(\emptyset) = 0$  and  $\Pr(S) = 1$ .

## SOME BACKGROUND ON SETS

We also recall some basic definitions about sets:

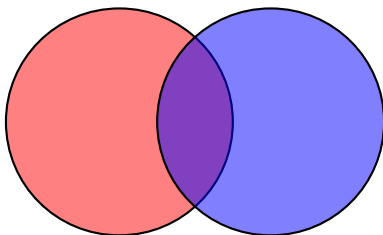
$A \cup B$	$A$ union $B$	things in $A$ or $B$ or both
$A \cap B$	$A$ intersect $B$	things in both $A$ and $B$
$-A$	complement of $A$	things in $S$ but out of $A$

Other notation for  $\cup$  is  $\vee$ .

Other notation for  $\cap$  is  $\wedge$ , and in older books &.

Other notation for  $-$  is  $\neg$ ,  $\sim$ , or  $\bar{A}$ .

## PICTURES



The union is the part shaded with one at least one color.

This allows for both colors.

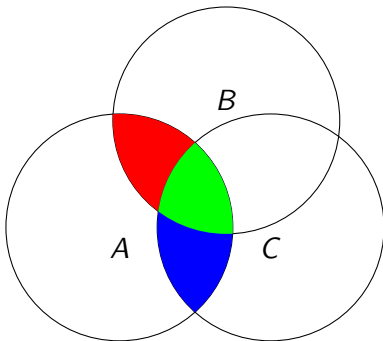
The intersection is the part in purple.

The complement of the red set would be everything else in whatever “big set” we’re talking about.

This “big set” is also called “the universe” and it is usually *implicit*.

## PRACTICE WITH PICTURES

The green part is  $A \cap B \cap C$ . What are the red and blue parts?



What expression gives us the union of the three colored parts?

## LAWS OF BOOLEAN ALGEBRA

$$A \cap B = B \cap A$$

$$A \cup B = B \cup A$$

$$A \cap \emptyset = \emptyset$$

$$A \cup \emptyset = A$$

$$A \cap S = A$$

$$A \cup S = S$$

$$A \cap -A = \emptyset$$

$$A \cup -A = S$$

$$-(A \cap B) = -A \cup -B \quad -(A \cup B) = -A \cap -B$$

$$(A \cap B) \cap C = A \cap (B \cap C)$$

$$(A \cup B) \cup C = A \cup (B \cup C)$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

Also:  $A \subseteq B$  if and only if  $A \cap B = A$  if and only if  $A \cup B = B$ .

# FACTS ABOUT PROBABILITY

For all events  $A \subseteq S$ ,  $0 \leq \Pr(A) \leq 1$ .

If  $A \subseteq B$ , then  $\Pr(A) \leq \Pr(B)$ .

$\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$ .

Hence: if  $A \cap B = \emptyset$ , then  $\Pr(A \cup B) = \Pr(A) + \Pr(B)$ .

Hence:  $\Pr(A \cap B) + \Pr(A \cap -B) = \Pr(A)$ .

## EXAMPLE SPACE

Roll two fair dice.

$S$  is the set of all 36 possible pairs  $(i, j)$ .

By fairness,  $\Pr(s) = 1/36$  for all  $s \in S$ .

So  $\sum_{s \in S} \Pr(s) = 36 \cdot 1/36 = 1$ .

And now we know that we have a probability space.

For each  $k$ , let  $A_k$  be the set of rolls which sum to  $k$ .

event	list	prob
$A_1$	$\emptyset$	0
$A_2$	$\{(1, 1)\}$	$1/36$
$A_3$	$\{(1, 2), (2, 1)\}$	$2/36 = 1/18$
$A_4$	$\{(1, 3), (2, 2), (3, 1)\}$	$3/36 = 1/12$
$A_5$	$\{(1, 4), (2, 3), (3, 2), (4, 1)\}$	$4/36 = 1/9$
$A_6$	$\{(1, 5), (2, 4), (3, 3), (4, 2), (5, 1)\}$	$5/36$

# DISCRETE RANDOM VARIABLES

We have a probability space  $(S, \Pr)$ .

Also functions  $X : S \rightarrow V$  for some  $V$ . (Usually  $V = \mathbb{R}$ .)

Let the image of  $X$  be  $x_1, \dots, x_n$ .

$X$  is a *random variable*. Usually  $S$  is implicit.

$X$  induces a **partition** of  $S$ : for  $x_i$  we have

$$X = x_i = \{s \in S : X(s) = x_i\}$$

and we also have  $\Pr[X = x_i]$ .

(These are standard kinds of notation. The important point is that  $X = x_i$  is an *event*.)

Let  $p_i = \Pr[X = x_i]$ ,

If  $V$  is the set of reals, we get an **expectation**

$$E(X) = \sum_{i=1}^n p_i x_i = \sum_{s \in S} X(s) \Pr(s).$$

# PARTITIONS

We just mentioned partitions.

These are ways to divide up a set into parts with two features:

- 1 All the parts must be **pairwise disjoint**: no overlap.
- 2 Every point in the big universe must belong to one of the parts.

## EXAMPLE AGAIN

In our space for rolling two dice, let  $X(i, j) = i + j$ .

Then the event we called  $A_k$  is now called " $X = k$ ".

This is the name of a set.

$$p_1 = 0, p_2 = 1/36, p_3 = 2/36, p_4 = 3/36, p_5 = 4/36, p_6 = 5/36, \\ p_7 = 6/36, p_8 = 5/36, p_9 = 4/36, p_{10} = 3/36, p_{11} = 2/36, p_{12} = 1/36.$$

The expectation of  $X$  can be calculated in two ways.

First,

$$E(X) = 1(0/36) + 2(1/36) + \cdots + 12(1/36) = 7.$$

Second, we could use  $E(X) = \sum_{s \in S} X(s) \Pr(s)$ .

This would be a sum of 36 numbers.

# CONDITIONAL PROBABILITY

The conditional probability of  $A$  given  $B$  is defined by:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$$

We **only** use this notation when  $\Pr(B) \neq 0$ .

Let  $S$  be from the dice example, let  $B =$  first is even.

We calculate  $\Pr(X = 8|B)$ .

The intersection is  $\{(2, 6), (4, 4), (6, 2)\}$ .

$\Pr(B) = 1/2$ ; you can check this.

So  $\Pr(X = 8|B) = (3/36)/(1/2) = 3/18$ .

# INDEPENDENCE

Events  $A$  and  $B$  are called *independent* if

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B).$$

This is equivalent to saying:  $\Pr(B) = 0$ , or  $\Pr(A|B) = \Pr(A)$ .

The idea: Knowing  $B$  occurred gives no information about the occurrence of  $A$ .

With 2 dice: “Sum = 8” and “first even” are not independent.

But “Sum = 5” and “first even” are independent.

So clearly we have to be careful with this concept!

## WHY IS THIS?

Let  $X$  be the random variable for the sum of the two dice.

Let  $B$  be “first die is even”.  $\Pr(B) = 1/2$ .

$\Pr(X = 8) = 5/36$ , and  $\Pr(X = 5) = 4/36$ .

Then  $(X = 8) \cap B = \{(2, 6), (4, 4), (6, 2)\}$ . So

$\Pr((X = 8) \cap B) = 3/36$ .

Try to do the rest yourself.

## WHY IS THIS?

Let  $X$  be the random variable for the sum of the two dice.

Let  $B$  be “first die is even”.  $\Pr(B) = 1/2$ .

$\Pr(X = 8) = 5/36$ , and  $\Pr(X = 5) = 4/36$ .

Then  $(X = 8) \cap B = \{(2, 6), (4, 4), (6, 2)\}$ . So

$\Pr((X = 8) \cap B) = 3/36$ .

Also,  $(X = 5) \cap B = \{(2, 3), (4, 1)\}$ ,

so  $\Pr((X = 5) \cap B) = 2/36$ .

And we see that

$$\begin{array}{lll} \Pr((X = 8) \cap B) & = & \Pr(X = 8) \cdot \Pr(B) \\ 3/36 & & 5/36 \cdot 1/2 & \text{false} \\ \Pr((X = 5) \cap B) & = & \Pr(X = 5) \cdot \Pr(B) \\ 2/36 & & 4/36 \cdot 1/2 & \text{true} \end{array}$$

# BACK TO THE MONTY HALL PROBLEM

We assume that given a choice of two doors to open, Monty opens each with probability  $1/2$ .

Let's write the triples of situations after the host opens as

(reality, your guess, Monty opens).

The sample space

(1, 1, 2)	(1, 1, 3)	(1, 2, 3)	(1, 3, 2)
(2, 1, 3)	(2, 2, 1)	(2, 2, 3)	(2, 3, 1)
(3, 1, 2)	(3, 2, 1)	(3, 3, 1)	(3, 3, 2)

The pairs with the first two entries the same have probability,  $1/18$ .

The other pairs all have probability  $1/9$ .

$\Pr(\text{your first guess is correct}) = 1/3$ .

$\Pr(\text{you would be correct after switching}) = 2/3$ .

# IT'S A GIRL!

I have two children.

Suppose I bring my daughter to class one day.

What's the probability that both children are girls?

This one is tricky because it's not clear what the sample space should be!

# IT'S A GIRL!

I have two children. Suppose I bring my daughter to class one day. What's the probability that both are girls?

The best way is to think of the children *in some order*, say with the one shown in class first.

$$S = \{(m, m), (m, f), (f, m), (f, f)\}.$$

Let  $A$  be the event where the first is  $f$ , and let  $B$  be the event where both are  $f$ . We need  $\Pr(B|A)$ .

$$A = \{(f, m), (f, f)\}. \quad B = \{(f, f)\}.$$

$$\text{So } \Pr(B|A) = \Pr(B \cap A) / \Pr(A) = 1/2.$$

## VARIATION ON “IT’S A GIRL!”

Now suppose you sit next to a woman at a meeting.  
Speaker asks: who has exactly 2 children? She raises her hand.  
Then asks: who has a daughter playing soccer?  
She raises her hand again. What's the prob of *both girls*?

## VARIATION ON “IT’S A GIRL!”

We use the same sample space,

$$S = \{(m, m), (m, f), (f, m), (f, f)\}.$$

But we want  $\Pr(B|C)$ , where  $C$  is the event *at least one f*.

$$C = \{(m, f), (f, m), (f, f)\}.$$

And now we see that

$$\Pr(B|C) = \Pr(B \cap C) / \Pr(C) = 1/3.$$

# THE CHAIN RULE

Recall that  $\Pr(B|A) = \Pr(A \cap B) / \Pr(A)$ .

It follows that  $\Pr(A \cap B) = \Pr(A) \times \Pr(B|A)$ .

Try in two minutes to prove that

$\Pr(A \cap B|C) = \Pr(A|C) \times \Pr(B|A \cap C)$ .

Now we get the generalized underlined fact or three events:

$$\begin{aligned}\Pr(X \cap Y \cap Z) &= \Pr(X \cap (Y \cap Z)) \\ &= \Pr(X) \times \Pr((Y \cap Z)|X) \\ &= \Pr(X) \times \Pr(Y|X) \times \Pr(Z|X \cap Y)\end{aligned}$$

# INDEPENDENCE

We next come to one of the main definitions in probability:  
**independence**.

There is an **intuitive idea**, and a **formal definition**.

This happens again and again, and it's important not to be confused or fooled.

The intuitive idea is that **independent events** means that information about whether one happened gives no information about whether the other did.

# INDEPENDENCE

Here is the **formal definition**:

$A$  and  $B$  are *independent* if  $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$ .

Equivalently,  $\Pr(A|B) = \Pr(A)$ , or else  $\Pr(B) = 0$ .

Why is this equivalence true?

# INDEPENDENCE

$A$  and  $B$  are *independent* if  $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$ .

Equivalently,  $\Pr(A|B) = \Pr(A)$ , or else  $\Pr(B) = 0$ .

First, assume  $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$ .

If  $\Pr(B) \neq 0$ , then

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(A) \cdot \Pr(B)}{\Pr(B)} = \Pr(A)$$

Second, assume  $\Pr(A|B) = \Pr(A)$  or  $\Pr(B) = 0$ .

If  $\Pr(B) = 0$ , then both  $\Pr(A \cap B)$  and  $\Pr(A) \Pr(B)$  are 0.

If  $\Pr(A|B) = \Pr(A)$ , then

$\Pr(A \cap B) = \Pr(A|B) \Pr(B) = \Pr(A) \Pr(B)$ .

## TRY THIS

Show that if  $A$  and  $B$  are independent, then  $A$  and  $\overline{B}$  are also independent.

(It's clear **intuitively**, but the point is to check it for the **formal definition**.)

## TRY THIS

Show that if  $A$  and  $B$  are independent, then  $A$  and  $\bar{B}$  are also independent.

(It's clear **intuitively**, but the point is to check it for the **formal definition**.)

Hint: First check that  $(A \cap B)$  and  $(A \cap \bar{B})$  have empty intersection, and their union is  $A$ .

So what does this tell us about  $\Pr(A \cap B)$  and  $\Pr(A \cap \bar{B})$ ?

$$\begin{aligned}\Pr(A \cap \bar{B}) &= \Pr(A) - \Pr(A \cap B) \\ &= \Pr(A) - \Pr(A) \cdot \Pr(B) \\ &= \Pr(A)(1 - \Pr(B)) \\ &= \Pr(A) \cdot \Pr(\bar{B})\end{aligned}$$

# BAYES' THEOREM

$$\Pr(A|B) = \frac{\Pr(B|A) \Pr(A)}{\Pr(B)}$$

(provided  $\Pr(A)$  is not zero).

Reason:

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)} = \frac{\Pr(B \cap A)}{\Pr(B)} = \frac{\Pr(B|A) \Pr(A)}{\Pr(B)}$$

Another useful form: suppose that  $A_1, \dots, A_n$  are a *partition* of the overall space  $S$ . This means that

- ★ For  $i \neq j$ ,  $A_i \cap A_j = \emptyset$ .
- ★  $A_1 \cup \dots \cup A_n = S$ .

Then provided all these events have non-zero probability,

$$\Pr(A_i|B) = \frac{\Pr(B|A_i) \Pr(A_i)}{\sum_j \Pr(B|A_j) \Pr(A_j)}.$$

# A TYPICAL BAYESIAN PROBLEMS

$S$  = symptom

$D$  = a common disease

Some made up numbers:

$$\Pr(S|D) = .6, \Pr(D) = .4, \Pr(S| - D) = .01.$$

Someone comes to a doctor with the symptom. What's the probability of their having the disease?

$$\begin{aligned} \Pr(S) &= \Pr(S \cap D) + \Pr(S \cap -D) \\ &= \Pr(S|D) \Pr(D) + \Pr(S| - D) \Pr(-D) \\ &= (.6)(.4) + (.01)(.6) \\ &= .24 + .006 \\ &= .246 \end{aligned}$$

So  $\Pr(D|S) = \Pr(S \cap D) / \Pr(S) = .24 / .246 = .975 = 97.5\%$ .

## A HINT OF THINGS TO COME

Let  $A$ ,  $B$  and  $C$  be events. The **joint distribution** of these is a table telling all possible probabilities of combinations with  $-$  and  $\cap$ :

event	Prob	event	Prob
$A \cap B \cap C$	.1	$\neg A \cap B \cap C$	.4
$A \cap B \cap \neg C$	.2	$\neg A \cap B \cap \neg C$	.1
$A \cap \neg B \cap C$	0	$\neg A \cap \neg B \cap C$	.2
$A \cap \neg B \cap \neg C$	0	$\neg A \cap \neg B \cap \neg C$	0

The numbers could be anything, but they must add up to 1.

To give distribution takes 7 numbers;  $2^n - 1$  in general.

If we knew that  $A$ ,  $B$ , and  $C$  were independent, then we'd only need 3 numbers:  $\Pr(A)$ ,  $\Pr(B)$ , and  $\Pr(C)$ .

## WHAT CAN YOU DO WITH A JOINT?

event	Prob	event	Prob
$A \cap B \cap C$	.1	$\neg A \cap B \cap C$	.4
$A \cap B \cap \neg C$	.2	$\neg A \cap B \cap \neg C$	.1
$A \cap \neg B \cap C$	0	$\neg A \cap \neg B \cap C$	.2
$A \cap \neg B \cap \neg C$	0	$\neg A \cap \neg B \cap \neg C$	0

$$\Pr(A \cap \neg B) = \Pr(A \cap \neg B \cap C) + \Pr(A \cap \neg B \cap \neg C) = 0 + 0 = 0$$

$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap \neg B) = (.1 + .2) + 0 = .3.$$

$$\Pr(B | \neg A) = \Pr(B \cap \neg A) / \Pr(\neg A) = .5 / .7 = .71.$$

# WHAT CAN YOU DO WITH A JOINT?

event	Prob	event	Prob
$A \cap B \cap C$	.1	$\neg A \cap B \cap C$	.4
$A \cap B \cap \neg C$	.2	$\neg A \cap B \cap \neg C$	.1
$A \cap \neg B \cap C$	0	$\neg A \cap \neg B \cap C$	.2
$A \cap \neg B \cap \neg C$	0	$\neg A \cap \neg B \cap \neg C$	0

The point is that with the joint distribution in hand, you can calculate **everything**.

But for large numbers of events/random variables, the joint distribution is too big to get your hands on.

# CONDITIONAL INDEPENDENCE

**Independent** events can become **dependent** in the presence of additional assumptions.

We have seen this already.

Let  $S$  be the two children example:

$$\begin{aligned} S &= \{m, f\} \times \{m, f\} \\ &= \{(m, m), (m, f), (f, m), (f, f)\} \end{aligned}$$

Let  $A$  be the event that the first is a girl.

Let  $B$  be the event that the second is a girl.

These are independent:  $\Pr(A|B) = 1/2 = \Pr(A)$ . But assuming that at least one girl ( $C$ ), they are not:

$$\Pr(A|C) = 2/3, \Pr(B|C) = 2/3,$$

$$\Pr(A \cap B|C) = 1/3.$$

And  $\Pr(A \cap B|C) \neq \Pr(A|C) \cdot \Pr(B|C)$  because

$$1/3 \neq (2/3) \cdot (2/3).$$

## CONDITIONAL INDEPENDENCE DEFINED

We say that  $A$  and  $B$  are **independent given  $C$**  if

$$\Pr(A \cap B | C) = \Pr(A | C) \cdot \Pr(B | C).$$

We also say that  $A$  and  $B$  are **conditionally independent given  $C$** . There are a number of equivalent formulations of this that we'll see in the next set of slides. One is

$$\Pr(A | C) = \Pr(A | B \cap C).$$

We say that  $A_1, \dots, A_n$  are **independent given  $B$**  if

$$\Pr(A_1 \cap \dots \cap A_n | B) = \Pr(A_1 | B) \times \dots \times \Pr(A_n | B).$$

## APPLICATION: DISEASES AND SYMPTOMS AGAIN

Suppose  $D$  has two symptoms:  $S$  and  $T$ .

Suppose that  $S$  and  $T$  are independent given  $D$

and also that  $S$  and  $T$  are independent given  $\bar{D}$

Suppose that

$$\Pr(D) = .3,$$

$$\Pr(S|D) = .4, \Pr(S|\bar{D}) = .1,$$

$$\Pr(T|D) = .6, \Pr(T|\bar{D}) = .2.$$

Suppose a person comes in with  $S$  but not  $T$ . What is the probability that they have  $D$ ?

[Hint: use the Chain Rule, and use conditional independence]

## SOLUTION

$$\begin{aligned}\Pr(D|S \cap \bar{T}) &= \Pr(D \cap S \cap \bar{T}) / \Pr(S \cap \bar{T}) \\ &= \frac{\Pr(D \cap S \cap \bar{T})}{\Pr(D \cap S \cap \bar{T}) + \Pr(\bar{D} \cap S \cap \bar{T})}\end{aligned}$$

$$\begin{aligned}\Pr(D \cap S \cap \bar{T}) &= \Pr(D) \Pr(S \cap \bar{T} | D) \\ &= \Pr(D) \Pr(S | D) \Pr(\bar{T} | D) \quad * \\ &= (.3)(.4)(.4) \\ &= .048\end{aligned}$$

$$\begin{aligned}\Pr(\bar{D} \cap S \cap \bar{T}) &= \Pr(\bar{D}) \Pr(S \cap \bar{T} | \bar{D}) \\ &= \Pr(\bar{D}) \Pr(S | \bar{D}) \Pr(\bar{T} | \bar{D}) \quad * \\ &= (.7)(.1)(.8) \\ &= .056\end{aligned}$$

So we get  $.048 / (.048 + .056) = .462 = 46\%$ .

\* \* These lines have some “cheating” that we’ll straighten out soon. Can you see what it is?

# GENERAL PROBABILITY BACKGROUND: SUBSPACES

Let  $S$  be a probability space. So  $\sum_{s \in S} \Pr(s) = 1$ .

Let  $E \subseteq S$  be an event such that  $\Pr(E) > 0$ .

We define a new space  $S|E$ , called  **$S$  relative to  $E$** .

The points are those of  $E$ . We'll write  $e$  for these.

The new probability function is  $\Pr_{new}(e) = \Pr(e) / \Pr(E)$ .

A random variable  $X$  on  $S$  restricts to one on  $S|E$ .

The point is that

$$\Pr_{new}(X = x) = \Pr(X = x|E).$$

Example:  $S$  is this room,  $E$  is the back row.

$\Pr_{new}(\text{Flavor} = \text{chocolate}) = \Pr(\text{Flavor} = \text{chocolate}|E)$ .

## THE VALUE OF SUBSPACES

We already know laws of probability for all events in all spaces. For example:  $\Pr(A \cap B) + \Pr(A \cup B) = \Pr(A) + \Pr(B)$ . An event like  $A$  can be thought of in terms of a random variable  $A$  with  $V(A) = \{t, f\}$ . And for random variables like  $X$  and  $Y$ , we get events  $X = x$  etc. So

$$\begin{aligned} & \Pr(X = x) + \Pr(Y = y) \\ = & \Pr(X = x \cap Y = y) + \Pr(X = x \cup Y = y) \end{aligned}$$

Fix an event  $E$ , and read the law in  $S|E$ . We get

$$\begin{aligned} & \Pr_{new}(X = x) + \Pr_{new}(Y = y) \\ = & \Pr_{new}(X = x \cap Y = y) + \Pr_{new}(X = x \cup Y = y) \end{aligned}$$

And by what we did before,

$$\begin{aligned} & \Pr(X = x|E) + \Pr(Y = y|E) \\ = & \Pr(X = x \cap Y = y|E) + \Pr(X = x \cup Y = y|E) \end{aligned}$$

## THE VALUE OF SUBSPACES, CONTINUED

Now we go back to our original space, and events  $A$ ,  $B$ , and  $E$ .

$A$  and  $B$  are basically the same as random variables with value spaces  $\{t, f\}$ .

More to the point,  $\Pr(A) = \Pr(A = t)$ , and similarly for  $B$ .

So we can apply the last law, taking  $A$  for  $X$ ,  $B$  for  $Y$ , and  $t$  for  $x$  and  $y$ . We get

$$\begin{aligned} & \Pr(A = t \cap B = t|E) + \Pr(A = t \cup Y = t|E) \\ = & \Pr(A = t|E) + \Pr(B = t|E) \end{aligned}$$

And this means that

$$\Pr(A \cap B|E) + \Pr(A \cup B|E) = \Pr(A|E) + \Pr(B|E).$$

Summary: we proved this by reading the original law

$$\Pr(A \cap B) + \Pr(A \cup B) = \Pr(A) + \Pr(B)$$

in the subspace  $E$ .

# HOW NOT TO CHEAT

We saw in a previous class that if  $X$  and  $Y$  are independent, so are  $X$  and  $\bar{Y}$ . By using subspaces, we can also see that if  $X$  and  $Y$  are independent **conditional on  $Z$** , then also  $X$  and  $\bar{Y}$  are independent **conditional on this same event  $Z$** .