Intro to Probability

Robert Y. Lewis

CS 0220 2024

April 4, 2024

Overview

Definitions of probability spaces

Definition: A countable *sample space* $\mathcal{S}$ is a nonempty countable(*) set.

Definition: An element $\omega \in \mathcal{S}$ is called an *outcome*.

Definition: A subset of $\mathcal{S}$ is called an *event*.

Definition: A *probability function* on a sample space $\mathcal{S}$ is a function $\Pr : \mathcal{S} \to \mathbb{R}$ such that

- $\Pr[\omega] \geq 0$ for all $\omega \in \mathcal{S}$, and
- $\sum_{\omega \in \mathcal{S}} \Pr[\omega] = 1$.

Definition: A sample space together with a probability function is called a *probability space*. For any event $E \subseteq \mathcal{S}$, the probability of E is defined to be the sum of the probabilities of outcomes in E:

$$\Pr[E] ::= \sum_{\omega \in E} \Pr[\omega].$$

Countability

Side note: a set S is *countable* if it is finite or there is a bijective function $\mathbb{N} \to S$.

Intuition: we can "list" the elements of S. $\{s_0, s_1, s_2, \ldots\}$. Maybe the list ends, maybe it doesn't...

We'll mostly deal with finite probability spaces.

Sum rule

Rule: If $\{E_0, E_1, \ldots,\}$ is collection of disjoint events, then

$$\Pr\left[\bigcup_{n \in \mathbb{N}} E_n\right] = \sum_{n \in \mathbb{N}} \Pr[E_n].$$

Like the sum rule in counting.

Example: I counted plants at the CS22 nursery. 60% of the plants had red flowers, 30% had yellow flowers, and 10% had no flowers. If I pick a plant at random, the probability that it has flowers is 90%.

What is sample space? What are events?

Complement rule

A new day, a new collection of plants:

set	type	Pr
A	Red flowers	0.60
B	Yellow flowers	0.25
$A \cap B$	both	0.15
$\overline{A} \cap \overline{B}$	neither	0.30

Since we know that the sets A and $\overline{A}$ are disjoint and cover all possibilities, the sum rule tells us that $\Pr[A] + \Pr[\overline{A}] = 1$.

Rule: $\Pr[\overline{A}] = 1 - \Pr[A]$.

Example: The chance that a plant does not have red flowers is (in symbols) $\Pr[\overline{A}] = 1 - \Pr[A] = 0.40$.

Difference Rule

set	type	Pr
A	Red flowers	0.60
B	Yellow flowers	0.25
$A \cap B$	both	0.15
$\overline{A} \cap \overline{B}$	neither	0.30

Rule: $\Pr[B \setminus A] = \Pr[B] - \Pr[A \cap B]$

Example: The chance that a plant has red flowers but not yellow flowers is (in symbols) $\Pr[A \setminus B] = \Pr[A] - \Pr[A \cap B] = 0.45$.

Proof: Follows from the Sum Rule because B is the union of the disjoint sets $B - A$ and $A \cap B$.

Inclusion-Exclusion

set	type	Pr
A	Red flowers	0.60
B	Yellow flowers	0.25
$A \cap B$	both	0.15
$\overline{A} \cap \overline{B}$	neither	0.30

Rule: $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B]$

Example: The chance that a plant has flowers is (in symbols) $\Pr[A \cup B] = \Pr[A] + \Pr[B] - \Pr[A \cap B] = 0.70$.

Proof: Follows from the Sum and Difference Rules, because $A \cup B$ is the union of the disjoint sets A and $B \setminus A$.

Boole's Inequality

set	type	Pr
A	Red flowers	0.60
B	Yellow flowers	0.25
$A \cap B$	both	0.15
$\overline{A} \cap \overline{B}$	neither	0.30

Rule: $\Pr[A \cup B] \leq \Pr[A] + \Pr[B]$

Example: The chance that a plant has flowers is (in symbols) $\Pr[A \cup B] \leq \Pr[A] + \Pr[B] = 0.85$.

Proof by inclusion-exclusion and the fact that probabilities are non-negative.

Monotonicity Rule

set	type	Pr
A	Red flowers	0.60
B	Yellow flowers	0.25
$A \cap B$	both	0.15
$\overline{A} \cap \overline{B}$	neither	0.30

Rule: If $A \subseteq B$, then $\Pr[A] \leq \Pr[B]$

Example: The chance that a plant has yellow flowers must be at least as big as the chance that it has both red and yellow flowers.

Proof: $\Pr[B] = \Pr[A \cup (B \setminus A)] = \Pr[A] + \Pr[B - A] \geq \Pr[A]$. First equality because $A \subseteq B$. Then, they add together by the sum rule because the two sets are disjoint. Then, the last inequality holds because probabilities are non-negative.

Puzzle

We surveyed the dinosaurs at Jurassic Park. About 50% of them were carnivores. About 40% of them were poisonous.

If we pick a dinosaur at random, what's the probability that it's neither a carnivore nor poisonous?

Union bound

Rule:

$$\Pr[E_1 \cup E_2 \cup \cdots \cup E_n] \leq \Pr[E_1] + \Pr[E_2] + \cdots + \Pr[E_n].$$

Example: The probability that a student has conflict with an exam is 0.001. What's the probability that *any* of 320 students have a conflict? Can't assume independence because groups of students take classes together, do sports together. Can't get an exact answer with the information provided.

Union bounds says probability of student 1 or 2 or 3 ... 320 less than or equal to $320 \times 0.001 = 0.32$.

Used in machine learning all the time.

Uniform

Definition: A finite probability space $\mathcal{S}$ is said to be *uniform* if $\Pr[\omega]$ is the same for every outcome $\omega \in \mathcal{S}$.

In finite spaces, for any $E \subseteq \mathcal{S}$,

$$\Pr[E] = \frac{|E|}{|\mathcal{S}|}.$$

Examples: Sides of a die, cards in a deck.

Contrast with: Vowels vs. consonants, primes vs. composites.

Counting example

What's the probability that 5 coin flips leads to a palindromic sequence?

What's the space of possibilities $\mathcal{S}$? The results of 5 coin flips: HTTHH. $|\mathcal{S}| = 2^5 = 32$.

What's the event of interest E? Palindromic results: TTHTT. $|E| = 2^3 = 8$. That's because the first 3 flips are "free", then the 4th flip must match the 2nd and the fifth flip must match the first.

The probability, therefore, is $|E|/|\mathcal{S}| = 2^3/2^5 = 1/2^2 = 1/4$.

Three-sided coin

If we want a uniform distribution over two options, we can flip a coin (H/T). If we want a uniform distribution over four options, we can flip two coins (HH/HT/TH/TT). What if we want a uniform distribution over three options?

We could flip two coins and say it's option 1 if both heads, option 2 if both tails, and option 3 if mismatch. Problem? Yes. The probability of mismatch is $\Pr[HT] + \Pr[TH] = 1/2$. Not $1/3$–$1/3$–$1/3$.

HH	1/4
HT	1/4
TH	1/4
TT	1/4

Repeat the trial

We could flip two coins and say it's option 1 if HH, option 2 if TT, option 3 if HT, and *do over* if TH. Problem? Maybe. There is an infinite number of outcomes...

Procedure selects option 1 if:

- HH on the first trial
- or TH on the first trial and HH on the second trial
- or TH on the first two trials and HH on the third trial
- ...

Pr(option 1)
$= 1/4 + 1/4 \times 1/4 + 1/4 \times 1/4 \times 1/4 + \ldots$
$= \sum_{i=1}^{\infty} 1/4^i$
$= 1/4 \times \sum_{i=0}^{\infty} 1/4^i$
$= 1/4 \left(\frac{1}{1-1/4}\right)$
$= 1/4 \times 4/3 = 1/3.$

Aside: Geometric sum

$x =$	$\sum_{i=0}^{\infty} p^i$	the sum we want
$x =$	$p^0 + p^1 + p^2 + p^3 + \dots$	expand
$px =$	$p^1 + p^2 + p^3 + p^4 + \dots$	multiply by p
$p^0 + px =$	$p^0 + p^1 + p^2 + p^3 + p^4 + \dots$	add p^0
$p^0 + px =$	x	defn of x
$p^0 =$	$x - px$	subtract px
$1 =$	$x(1-p)$	factor/simplify
$\frac{1}{1-p} =$	x	divide by $1-p$

Infinite sample space

$\mathcal{S} = \{HH, HT, TT, TH : HH, TH : HT, TH : TT, TH : TH : HH, TH : TH : HT, TH : TH : TT, \ldots\}$
$= \{(TH)^n : HH, (TH)^n : HT, (TH)^n : TT | n \in \mathbb{N}\}$

The probability space is:

$\Pr((TH)^n : HH) = \Pr((TH)^n : HT) = \Pr((TH)^n : TT) = 1/4^{n+1}$.

Note: $\sum_{n=0}^{\infty} 3 \times 1/4^{n+1} = 3/4 \times \frac{1}{1-1/4} = 3/4 \times 4/3 = 1$.

Non-negative and sums to one, valid probability space!