3. Discrete Probability

CSE 312
Spring 2015
W.L. Ruzzo
Probability theory:

“an aberration of the intellect”

and

“ignorance coined into science”

– John Stuart Mill
**Sample space:**  $S$ is a set of all potential outcomes of an experiment (often $\Omega$ in textbooks—Greek uppercase omega)

- **Coin flip:**  $S = \{\text{Heads}, \text{Tails}\}$
- **Flipping two coins:**  $S = \{(\text{H,H}), (\text{H,T}), (\text{T,H}), (\text{T,T})\}$
- **Roll of one 6-sided die:**  $S = \{1, 2, 3, 4, 5, 6\}$
- **# emails in a day:**  $S = \{ x : x \in \mathbb{Z}, \ x \geq 0 \}$
- **YouTube hrs. in a day:**  $S = \{ x : x \in \mathbb{R}, 0 \leq x \leq 24 \}$

Some fine print: “sample space” for an experiment isn’t uniquely defined, & “potential” outcomes may include literally are impossible ones, e.g., $S=\{1,2,3,4,5,6,7\}$ for a 6-sided die; it’s all OK if you’re sensible and consistent, e.g., if you make probability(7)=0. Rare to see things quite this wacky, but bottom line: a sample space is just a set, any set.
**Events:** $E \subseteq S$ is an arbitrary subset of the sample space

- Coin flip is heads: $E = \{\text{Head}\}$
- At least one head in 2 flips: $E = \{(H,H), (H,T), (T,H)\}$
- Roll of die is odd: $E = \{1, 3, 5\}$
- # emails in a day < 20: $E = \{x : x \in \mathbb{Z}, 0 \leq x < 20\}$
- # emails in a day is prime: $E = \{2, 3, 5, 7, 11, 13, \ldots\}$
- Wasted day (>5 YT hrs): $E = \{x : x \in \mathbb{R}, x > 5\}$

Note: an event is not an outcome, it is a set of outcomes. E.g., the outcome of rolling a die is always a single number in 1..6; “roll is odd” aggregates 3 potential outcomes as one event; “roll is >5” aggregates 1 potential outcome as the event $E = \{6\}$ (a singleton set).
set operations on events

E and F are events in the sample space S

\[ S \]

\[ E \quad \cap \quad F \]
E and F are events in the sample space $S$

Event “E OR F”, written $E \cup F$

$S = \{1,2,3,4,5,6\}$
outcome of one die roll

$E = \{1,2\}$, $F = \{2,3\}$
$E \cup F = \{1,2,3\}$
set operations on events

E and F are events in the sample space S

Event “E AND F”, written $E \cap F$ or $EF$

$S = \{1,2,3,4,5,6\}$
outcome of one die roll

$E = \{1,2\}, \ \ F = \{2,3\}$
$E \cap F = \{2\}$
set operations on events

E and F are events in the sample space S

\[ EF = \emptyset \iff E, F \text{ are “mutually exclusive”} \]

\[ S = \{1,2,3,4,5,6\} \]

outcome of one die roll

\[ E = \{1,2\}, \quad F = \{2,3\}, \quad G = \{5,6\} \]

\[ EF = \{2\}, \text{ not mutually exclusive, but } E, G \text{ and } F, G \text{ are} \]
set operations on events

E and F are events in the sample space S

Event “not E,” written $\overline{E}$ or $\neg E$

$S = \{1, 2, 3, 4, 5, 6\}$
outcome of one die roll

$E = \{1, 2\}$
$\neg E = \{3, 4, 5, 6\}$
set operations on events

DeMorgan’s Laws

\[
\overline{E \cup F} = \overline{E} \cap \overline{F}
\]

\[
\overline{E \cap F} = \overline{E} \cup \overline{F}
\]
Intuition: Probability as the relative frequency of an event

\[ \Pr(E) = \lim_{n \to \infty} \frac{\# \text{ of occurrences of } E \text{ in } n \text{ trials}}{n} \]

Mathematically, this proves messy to deal with.

Instead, we define “Probability” via a function from subsets of \( S \) (“events”) to real numbers

\[ \Pr: 2^S \to \mathbb{R} \]

satisfying the properties (axioms) below.
Intuition: Probability as the relative frequency of an event

\[ \Pr(E) = \lim_{n \to \infty} \left( \frac{\text{# of occurrences of } E \text{ in } n \text{ trials}}{n} \right) \]

Axiom 1 (Non-negativity): \( 0 \leq \Pr(E) \)

Axiom 2 (Normalization): \( \Pr(S) = 1 \)

Axiom 3 (Additivity):

If \( E \) and \( F \) are mutually exclusive (\( EF = \emptyset \)), then

\[ \Pr(E \cup F) = \Pr(E) + \Pr(F) \]

For any sequence \( E_1, E_2, \ldots, E_n \) of mutually exclusive events,

\[ \Pr \left( \bigcup_{i=1}^{n} E_i \right) = \Pr(E_1) + \cdots + \Pr(E_n) \]
implications of axioms

\[ \Pr(\overline{E}) = 1 - \Pr(E) \]

1 = \Pr(S) = \Pr(E \cup \overline{E}) = \Pr(E) + \Pr(\overline{E})

If \( E \subseteq F \), then \( \Pr(E) \leq \Pr(F) \)

\[ \Pr(F) = \Pr(E) + \Pr(F - E) \geq \Pr(E) \]

\[ \Pr(E \cup F) = \Pr(E) + \Pr(F) - \Pr(EF) \]

inclusion-exclusion

\[ \Pr(E) \leq 1 \]

exercise

And many others
**Sample space:** $S =$ set of all potential outcomes of experiment

E.g., flip two coins: $S = \{(H,H), (H,T), (T,H), (T,T)\}$

**Events:** $E \subseteq S$ is an arbitrary subset of the sample space

$\geq 1$ head in 2 flips: $E = \{(H,H), (H,T), (T,H)\}$

**Probability:**

A function from subsets of $S$ to real numbers – $\Pr: 2^S \rightarrow \mathbb{R}$

**Probability Axioms:**

Axiom 1 (Non-negativity): $0 \leq \Pr(E)$

Axiom 2 (Normalization): $\Pr(S) = 1$

Axiom 3 (Additivity): $EF = \emptyset \Rightarrow \Pr(E \cup F) = \Pr(E) + \Pr(F)$
Simplest case: sample spaces with equally likely outcomes.

- Coin flips: \( S = \{\text{Heads}, \text{Tails}\} \)
- Flipping two coins: \( S = \{(\text{H,H}), (\text{H,T}), (\text{T,H}), (\text{T,T})\} \)
- Roll of 6-sided die: \( S = \{1, 2, 3, 4, 5, 6\} \)

\[
\Pr(\text{each outcome}) = \frac{1}{|S|}
\]

In that case,

\[
\Pr(E) = \frac{\text{number of outcomes in } E}{\text{number of outcomes in } S} = \frac{|E|}{|S|}
\]

Why? Axiom 3 plus fact that \( E \) = union of singletons in \( E \)
Roll two 6-sided dice. What is $\text{Pr}(\text{sum of dice} = 7)$?

$S = \{(1,1), (1,2), (1,3), (1,4), (1,5), (1,6), (2,1), (2,2), (2,3), (2,4), (2,5), (2,6), (3,1), (3,2), (3,3), (3,4), (3,5), (3,6), (4,1), (4,2), (4,3), (4,4), (4,5), (4,6), (5,1), (5,2), (5,3), (5,4), (5,5), (5,6), (6,1), (6,2), (6,3), (6,4), (6,5), (6,6)\}$

$E = \{(6,1), (5,2), (4,3), (3,4), (2,5), (1,6)\}$

$\text{Pr}(\text{sum} = 7) = |E|/|S| = 6/36 = 1/6.$
Roll two 6-sided dice. What is \( \text{Pr}(\text{sum of dice} = 7) \) ?

\[
S = \{ (1,1), (1,2), (1,3), (1,4), (1,5), (1,6), \\
(2,1), (2,2), (2,3), (2,4), (2,5), (2,6), \\
(3,1), \\
(4,1), \\
(5,1), \\
(6,1), \}
\]

\[
E = \{ (1,6), (2,5), (3,4), (4,3), (5,2), (6,1) \}
\]

\[
\text{Pr}(\text{sum} = 7) = \frac{|E|}{|S|} = \frac{6}{36} = \frac{1}{6}.
\]

**SIDEBAR**

It’s perhaps tempting to try \( S=\{2,3,\ldots,12\} \) and \( E=\{7\} \) for this problem. This isn’t wrong, but note that it doesn’t fit the “equally likely outcomes” scenario. E.g., \( \text{Pr}\{2\}=1/36 \neq 1/6=\text{Pr}\{7\} \). Plus, it’s usually best to make “\( S \)” a simple representation of the “experiment at hand,” e.g., an ordered pair reflecting the 2 dice rolls, rather than a more complex derivative of it, like their sum. The later makes it easy to express this event (“sum is 7”), but makes it difficult or impossible to express other events of potential interest (“product is odd,” for example).
twinkies and ding dongs
twinkies and ding dongs

4 Twinkies and 3 DingDongs in a bag. 3 drawn. What is \( \text{Pr}(\text{one Twinkie and two DingDongs drawn}) \)?

Ordered: (\(S\) ordered triple with 3 of 7 distinguishable objects)

• Pick 3, one after another: \( |S| = 7 \cdot 6 \cdot 5 = 210 \)
• Pick Twinkie as either 1\(^{\text{st}}, 2^{\text{nd}}, \) or 3\(^{\text{rd}}\) item:
  \( |E| = (4 \cdot 3 \cdot 2) + (3 \cdot 4 \cdot 2) + (3 \cdot 2 \cdot 4) = 72 \)
• \( \text{Pr}(\text{1 Twinkie and 2 DingDongs}) = 72/210 = 12/35 \).

Unordered: (\(S\) unordered triple with 3 of 7 distinguishable objects)

• Grab 3 at once: \( |S| = \binom{7}{3} = 35 \)
• \( |E| = \binom{4}{1} \binom{3}{2} = 12 \)
• \( \text{Pr}(\text{1 Twinkie and 2 DingDongs}) = 12/35 \).

Exercise: a 3\(^{\text{rd}}\) way – \(S\) is ordered list of 7, \(E\) is “1\(^{\text{st}}\) 3 OK”; same answer?
birthdays
What is the probability that, of \( n \) people, none share the same birthday?

What are \( S \), \( E \)??

\[ |S| = (365)^n \]
\[ |E| = (365)(364)(363)\cdots(365-n+1) \]

\[ \Pr(\text{no matching birthdays}) = \frac{|E|}{|S|} = \frac{(365)(364)\cdots(365-n+1)}{(365)^n} \]

Some values of \( n \)...

\( n = 23 \): \( \Pr(\text{no matching birthdays}) < 0.5 \)
\( n = 77 \): \( \Pr(\text{no matching birthdays}) < \frac{1}{5000} \)
\( n = 90 \): \( \Pr(\text{no matching birthdays}) < \frac{1}{162,000} \)
\( n = 100 \): \( \Pr(\text{no matching birthdays}) < \frac{1}{3,000,000} \)
\( n = 150 \): \( \Pr(...) < \frac{1}{3,000,000,000,000,000} \)
n = 366?

Pr = 0

Above formula gives this, since

\[
\frac{(365)(364)\ldots(365-n+1)}{(365)^n} = 0
\]

when \( n = 366 \) (or greater).

Even easier to see via pigeon hole principle.
What is the probability that, of n people, none share the same birthday as you?

\[ |S| = (365)^n \]
\[ |E| = (364)^n \]
\[ \Pr(\text{no birthdays = yours}) = \frac{|E|}{|S|} = \frac{(364)^n}{(365)^n} \]

Some values of n...

n = 23:  \quad \Pr(\text{no matching birthdays}) \approx 0.9388
n = 90:  \quad \Pr(\text{no matching birthdays}) \approx 0.7812
n = 253: \quad \Pr(\text{no matching birthdays}) \approx 0.4995

Exercise: \( p^n \) is not linear, but red line looks straight. Why?
chip defect detection
n chips manufactured, one of which is defective
k chips randomly selected from n for testing

What is $\Pr(\text{defective chip is in k selected chips})$?

$|S| = \binom{n}{k}$

$|E| = \binom{1}{1} \binom{n-1}{k-1}$

$\Pr(\text{defective chip is among k selected chips})$

$$= \frac{(1) \binom{n-1}{k-1}}{\binom{n}{k}} = \frac{(n-1)!}{(k-1)!(n-k)!} \cdot \frac{k!}{n!} \cdot \frac{n!}{k!(n-k)!} = \frac{k}{n}$$
n chips manufactured, one of which is defective
k chips randomly selected from n for testing

What is \( \Pr(\text{defective chip is in k selected chips}) \)?

Different analysis:

• Select \( k \) chips at random by permuting all \( n \) chips and then choosing the first \( k \).
• Let \( E_i \) = event that \( i^{th} \) selected chip is defective.
• Events \( E_1, E_2, \ldots, E_k \) are mutually exclusive
• \( \Pr(E_i) = \frac{1}{n} \) for \( i=1,2,\ldots,k \)
• Thus \( \Pr(\text{defective chip is selected}) \)
  \[ = \Pr(E_1) + \cdots + \Pr(E_k) = \frac{k}{n}. \]
n chips manufactured, two of which are defective
k chips randomly selected from n for testing

What is $Pr(\text{a defective chip is in } k \text{ selected chips})$?

$|S| = \binom{n}{k}$  $|E| = (1 \text{ chip defective}) + (2 \text{ chips defective})$

$= \binom{2}{1} \binom{n-2}{k-1} + \binom{2}{2} \binom{n-2}{k-2}$

Pr(\text{a defective chip is in } k \text{ selected chips})

$= \frac{\binom{2}{1} \binom{n-2}{k-1} + \binom{2}{2} \binom{n-2}{k-2}}{\binom{n}{k}}$
n chips manufactured, two of which are defective
k chips randomly selected from n for testing

What is $\Pr(\text{a defective chip is in k selected chips})$?

Another approach:

$\Pr(\text{a defective chip is in k selected chips}) = 1 - \Pr(\text{none})$

$\Pr(\text{none})$:

$$|S| = \binom{n}{k}, |E| = \binom{n - 2}{k}, \Pr(\text{none}) = \frac{n - 2}{\binom{n}{k}}$$

$\Pr(\text{a defective chip is in k selected chips}) = 1 - \frac{n - 2}{\binom{n}{k}}$

(Same as above? Check it!)
poker hands
5 card poker hands (ordinary 52 card deck, no jokers etc.)
flush, 1 pair, 3 of a kind, 2 pairs, full house, …

Sample Space?

Imagine sorted tableau of cards, pick 5:

\[
|S| = \binom{52}{5}
\]
Consider 5 card poker hands.
A “straight” is 5 consecutive rank cards ignoring suit (Ace low or high, but not both. E.g., A,2,3,4,5 or 10,J,Q,K,A)

What is $\Pr(\text{straight})$?

$S$ as on previous slide, $|S| = \binom{52}{5}$

$E = \text{Pick a col A, 2, … 10, then 1 of 4 in next 5 cols (wrapping K→A)}$

$|E| = 10 \cdot \binom{4}{1}^5$

$\Pr(\text{straight}) = \frac{10 \binom{4}{1}^5}{\binom{52}{5}} \approx 0.00394$
card flipping
52 card deck. Cards flipped one at a time.

After first ace (of any suit) appears, consider next card

\[ \Pr(\text{next card} = \text{ace of spades}) < \Pr(\text{next card} = 2 \text{ of clubs})? \]

Maybe, Maybe Not …

\( S = \) all permutations of 52 cards, \(|S| = 52!\)

**Event 1: Next = Ace of Spades.**

Remove A♠, shuffle remaining 51 cards, add A♠ after first Ace

\(|E_1| = 51!\) (only 1 place A♠ can be added)

**Event 2: Next = 2 of Clubs**

Do the same thing with 2♣; \(E_1\) and \(E_2\) have same size

So,

\[ \Pr(E_1) = \Pr(E_2) = \frac{51!}{52!} = \frac{1}{52} \]
Ace of Spades: 2/6

2 of Clubs: 2/6

Theory is the same for a 3-card deck; Pr = 2!/3! = 1/3
hats
n persons at a party throw hats in middle, select at random. What is \( \text{Pr}(\text{no one gets own hat}) \)?

\[
\text{Pr}(\text{no one gets own hat}) = 1 - \text{Pr}(\text{someone gets own hat})
\]

\[
\text{Pr}(\text{someone gets own hat}) = \text{Pr}(\bigcup_{i=1}^{n} E_i), \text{ where } E_i = \text{event that person } i \text{ gets own hat}
\]

\[
\text{Pr}(\bigcup_{i=1}^{n} E_i) = \sum_i P(E_i) - \sum_{i<j} \text{Pr}(E_i E_j) + \sum_{i<j<k} \text{Pr}(E_i E_j E_k) \ldots
\]
Visualizing the sample space $S$:

People: $P_1 \ P_2 \ P_3 \ P_4 \ P_5$
Hats: $H_4 \ H_2 \ H_5 \ H_1 \ H_3$

I.e., a sample point is a *permutation* $\pi$ of $1, \ldots, n$

$\begin{bmatrix} 4 & 2 & 5 & 1 & 3 \end{bmatrix}$

$|S| = n!$
\( E_i = \text{event that person } i \text{ gets own hat: } \pi(i) = i \)

\( |E_i| = (n-1)! \) for all \( i \)

Counting single events:

\( |E_i| = (n-1)! \) for all \( i \)

Counting pairs:

\( E_iE_j : \pi(i) = i \text{ and } \pi(j) = j \)

\( |E_iE_j| = (n-2)! \) for all \( i, j \)
n persons at a party throw hats in middle, select at random. What is $\Pr(\text{no one gets own hat})$?

$E_i = \text{event that person } i \text{ gets own hat}$

$\Pr(\bigcup_{i=1}^{n} E_i) = \sum_i P(E_i) - \sum_{i<j} \Pr(E_i \cap E_j) + \sum_{i<j<k} \Pr(E_i \cap E_j \cap E_k) \ldots$

$\Pr(k \text{ fixed people get own back}) = \frac{(n-k)!}{n!}$

$\binom{n}{k} \text{ times that} = \frac{n!}{k!(n-k)!} \frac{(n-k)!}{n!} = \frac{1}{k!}$

$\Pr(\text{none get own}) = 1 - \Pr(\text{some do}) = 1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \ldots + (-1)^n/n! \approx 1/e \approx .37$
Pr(none get own) = 1 - Pr(some do) =
1 - 1 + 1/2! - 1/3! + 1/4! … + (-1)^n/n! ≈ e^{-1} ≈ .37

Oscillates forever, but quickly converges to 1/e
Sample spaces
Events
Set theory
Axioms
Simple identities
Equally likely outcomes (counting)
Examples
  All good for building your skills
  Birthdays is particularly important for applications
  Hats is important as example of inclusion/exclusion