STAT 350 Lecture 2: Randomness and Probability

Probability Theory

(Chapter 2 of WMMY)

Outline

Frequency vs Probability

Flip a fair coin four times and we care about the number of Heads. Regard the population as the outcomes of infinite independent trials of flipping the coin 4 times. We are interested in the distribution of the variable: number of Heads by flipping a fair coin 4 times.

1. Collect a data set by sampling from the population.
2. The variable is discrete.
3. Draw line graph to see the distribution of the variable based on the sample.

Frequency vs Probability

Think about the following:

1. The distribution is actually for the variable defined on the whole population.
2. The line graph drawn based on the sample is just to get an approximation of the distribution of the variable. And the larger the sample size is, the better the approximation is.
3. Actually the distribution of the variable can be obtained as the limit of the distribution calculated from the sample as the sample size goes to infinity.
4. Could you use your math to come up with the distribution exactly in a clever way without any limitation process?

Think Hard

Flip a fair coin 4 times, what's the probability to have 2 heads?

1. We are doing an experiment: flip a fair coin 4 times. (Different times get different results.)
2. We can enumerate all of the possible results for such experiment.
3. Each result is equally likely to happen.
4. Count how many such results give out 2 heads.
5. Count how many results in total.
6. The ratio of the two counts is the probability!

Outline

Experiments

- An *experiment* is a procedure by which an observation or measurement is made.
- A single execution of an experiment is called a *trial*.
- The collection of all possible *outcomes* of an experiment is called the *sample space*, and is denoted by $\mathcal{S}$.
- What is $\mathcal{S}$ in the following examples?
  - Experiment $\rightarrow$ Flip a coin until Heads appear.
  - Experiment $\rightarrow$ Toss two dice.
  - Experiment $\rightarrow$ Sum of the numbers showing up after tossing two dice.

Events

- A collection of possible outcomes of an experiment is called an *event*, and usually denoted by upper-case letters A, B, C, etc.
- Can be defined by writing out the particular outcomes in a list, or by a phrase describing those outcomes.
- Example.
  - Experiment $\rightarrow$ Roll a six-sided die.
  - The sample space is

$$\mathcal{S} = \{1, 2, 3, 4, 5, 6\}.$$

  - $A = \{\text{“even numbers”}\} = \{2, 4, 6\}$
  - $B = \{\text{“not a five”}\} = \{1, 2, 3, 4, 6\}$
  - $C = \{\text{“greater than 3”}\} = \{4, 5, 6\}$

Set operations

- Events are nothing but subsets of $\mathcal{S}$.
- *Set theory* tells us how to deal with these objects.
- Basic set operations:
  - **Union:** $A \cup B = \{\text{stuff in either } A \text{ or } B\}$
  - **Intersection:** $A \cap B = \{\text{stuff in both } A \text{ and } B\}$
  - **Complement:** $A^c = \{\text{stuff not in } A\}$
- Note: Ross writes AB instead of $A \cap B$!
- Just make the following connections:

$$\begin{aligned}
\text{Union} &\iff \cup \iff \text{“or”} \\
\text{Intersection} &\iff \cap \iff \text{“and”} \\
\text{Complement} &\iff (\cdot)^c \iff \text{“not”}
\end{aligned}$$

Example – die roll

- Experiment $\rightarrow$ Roll a six-sided die.
- Recall: the sample space is $\mathcal{S} = \{1, 2, 3, 4, 5, 6\}$.
- Define three events:

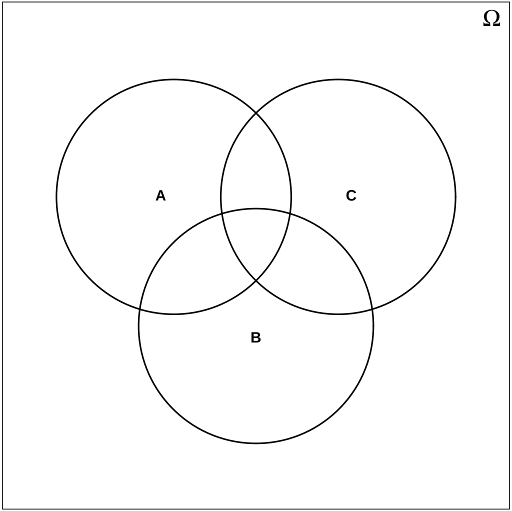
$$\begin{aligned} A &= \{\text{“even numbers”}\} = \{2, 4, 6\} \\ B &= \{\text{“not a five”}\} = \{1, 2, 3, 4, 6\} \\ C &= \{\text{“greater than 3”}\} = \{4, 5, 6\} \end{aligned}$$

- Then we can do the following calculations:
  - $A \cap C = \{4, 6\}$
  - $A^c = \{1, 3, 5\}$
  - $A \cup C = \{2, 4, 5, 6\}$
  - $A \cup B = B$

More terminology

- Sometimes all the outcomes of one event E are also outcomes in another event F (i.e., $x \in E$ implies $x \in F$).
- We then write $E \subset F$ ("E is a subset of F").
- Sometimes two events E and F have no elements in common.
- Then nothing is in both events and we write $E \cap F = \emptyset$, the "empty set".
- Two events E and F with $E \cap F = \emptyset$ are called *disjoint* or *mutually exclusive*.

Venn diagrams



More set operations

Here are some rules about the order of operations

- Order does not matter if all the operations are the same:
  - $(A \cap B) \cap C = A \cap (B \cap C)$
  - $(A \cup B) \cup C = A \cup (B \cup C)$
- But not so easy when operations are mixed:
  - $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$
  - $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$
  - DeMorgan's laws:
    - $(A \cup B)^c = A^c \cap B^c$
    - $(A \cap B)^c = A^c \cup B^c$

Outline

Set functions

- Now that we have a rigorous system for considering experiments, outcomes and events, we need a way to assign probabilities to these events.
- Formally, this is done by defining a function $\Pr(\cdot)$ whose input is an event and its output is a number.
- Functions like this are often referred to as *set functions*.
- If you think of a rectangle in the plane as the set of points it contains, then Area is a set function.

Probability rules/axioms

- To have a theory of probability that is consistent with our intuition, we must impose some rules or axioms on our probability $\Pr(\cdot)$.
- These are Kolmogorov's three axioms:
  1. For any event A, $0 \leq \Pr(A) \leq 1$.
  2. $\Pr(\mathcal{S}) = 1$.
  3. If $A_1, A_2, \ldots$ is a sequence of mutually exclusive (disjoint) events, then

$$\Pr\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} \Pr(A_i), n = 1, 2, \cdots, \infty.$$

- Everything we would ever need to know about probability can be derived from these three simple axioms!

Classical probability

- The first (and simplest) system of probability was based on the assumption that all outcomes are equally likely.
- More specifically, for an event A in a sample space $\mathcal{S}$,

$$\Pr(A) = \frac{\#\text{ of outcomes in } A}{\#\text{ of outcomes in } \mathcal{S}}.$$

- *Claim:* This probability satisfies Kolmogorov's three axioms!
- Can you prove it?

Exercises

Suppose a red die and a green die are rolled. Under the classical probability model, compute the following:

1. $\Pr(\text{Red} = 2, \text{Green} = 4)$
2. $\Pr(\text{both dice are the same})$
3. $\Pr(\text{sum is } 7)$
4. $\Pr(\text{Red} > \text{Green})$

Counting rules

- Under the classical probability model, shortcuts for counting the number of elements in very large sets will be useful.
- There are two special concepts we will use, namely,
  - **Permuations**
  - **Combinations**
- Both are based on the following fundamental principle:

Theorem (Basic Counting Principle – Step by Step)

If experiment i has n_i possible outcomes ($i = 1, 2, \ldots, r$), then

$$\begin{pmatrix} \textit{total number of possible} \\ \textit{outcomes of the } r \textit{ experiments} \end{pmatrix} = \prod_{i=1}^{r} n_i.$$

- E.g. toss a coin 3 times, how many total outcomes? consider possible outcomes of each toss

Example – number of subsets

- Suppose we have a set $E = \{e_1, e_2, \ldots, e_n\}$
- **Q:** How many possible subsets are there?
- **A:** Go step by step:
  - Is e_1 in the subset or not? (Yes/No)
  - Is e_2 in the subset or not? (Yes/No)
  - ...
  - Is e_n in the subset or not? (Yes/No)
- Each sequence of n Yes/No answers defines a subset.
- Therefore, n actions, each with two possible choices, so

$$\text{BCP} \Longrightarrow \# \text{ of subsets of } E = 2^n.$$

Example – Assigning balls to boxes

- Suppose we have n different balls and n labeled boxes.
- **Q:** How many ways can we assign one ball to each box?
- **A:** Go step by step:
  - For the first ball, there are n boxes to choose from.
  - For the second ball, there are $n-1$ boxes to choose from.
  - ...
  - For the $(n-1)^{th}$ ball, there are 2 boxes to choose from.
  - For the n^{th} ball, there is 1 box to choose from.
- $r = n$ actions, $n_i = n - i + 1$, so

$$\text{BCP} \Longrightarrow \#\text{ of assignments} = n(n-1)(n-2)\cdots 2 \cdot 1 = n!.$$

Example – Permutations

- A club has n individuals, and the goal is to choose a committee of k members, where each member has a different title (e.g., a president, vice president, and treasurer for $k = 3$).
- **Q:** How many such committees are there (denoted by $(n)_k$)?
- **A:** This is just like the previous example.
  - For the 1^{st} title, there are n person to choose from
  - For the 2^{nd} title, there are $n - 1$ person to choose from ...
  - For the k^{th} title, there are $n - (k - 1)$ person to choose from
  - Therefore, by the BCP

$$(n)_k = n(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}.$$

- **Summary**: In a pool with n different objects, **choose** k objects from the pool and **arrange** them in a specific order.—number of different results is $(n)_k$—**Permutation of** k **objects from** n.

Example – Combinations

- A club has n individuals, and the goal is to choose a committee of k members without differentiation.
- **Q:** How many such committees are there (denoted by $\binom{n}{k}$)?
- **A:** This can be solved by combining the previous two examples.
  - If with different titles, then $(n)_k$.
  - For a group with k members, how many different labels for the k titles? Assign k balls to k boxes, which leads to $k!$.
  - Then we have the result

$$\binom{n}{k} = \frac{(n)_k}{k!} = \frac{n!}{k!(n-k)!}.$$

- **Summary**: In a pool with n different objects, **choose** k objects from the pool and **without arrangement**.—number of different results is $\binom{n}{k}$—**Combination of k objects from n.**

Exercise – coin tossing

Toss a fair coin five times.

1. What's the probability of getting exactly three H's?
2. What's the probability of getting exactly two H's?
3. Explain why the answers to parts (1) and (2) are the same.

Exercise – poker hands

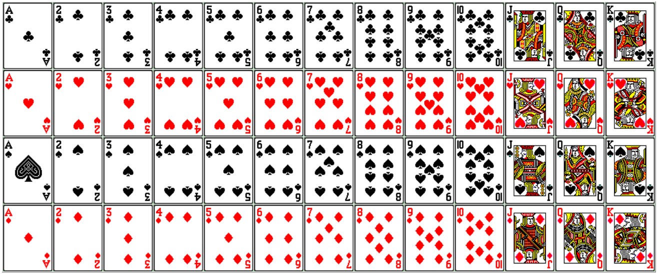


Figure: A 52-card deck has 13 ranks of each of the four suits, clubs (black),heart (red), spades (black), and diamonds (red).

Deal out five cards from a 52-card deck at random.

1. How many distinct hands are possible?
2. $\Pr(\text{flush}) = \binom{4}{1}\binom{13}{5}/\binom{52}{5}$ (all 5 cards from the same suit)
3. $\Pr(\text{two pair}) = \binom{13}{2}\binom{4}{2}\binom{4}{2}\binom{11}{1}\binom{4}{1}/\binom{52}{5}$ (this hand has the pattern AABBC with A, B, C distinct rank)

Important properties (6 Rules) of General Probability

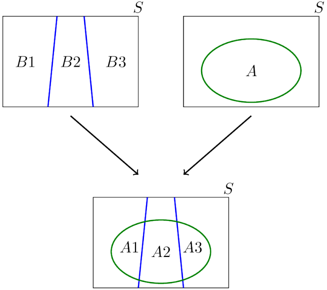
3 Rules of General Probability:

- (Rule 1) Complementation rule: $\Pr(A^c) = 1 - \Pr(A)$.
- (Rule 2) Monotonicity rule: If $A \subset B$, then $\Pr(A) \leq \Pr(B)$.
- (Rule 3) Partition rule: $\Pr(B) = \Pr(B \cap A) + \Pr(B \cap A^c)$.

Note: The general partition rule is stated as follows. A sequence of events $A_1, A_2, \ldots$ is a partition of $\mathcal{S}$ if $A_i \cap A_j = \emptyset$ for each $i \neq j$, and $\bigcup_{n=1}^{\infty} A_n = \mathcal{S}$.
If $A_1, A_2, \ldots$ is a partition of $\mathcal{S}$ and B is any event, then

$$\Pr(B) = \sum_{n=1}^{\infty} \Pr(B \cap A_n).$$

Powerful idea 1: solving things case by case



Partition Rule: Case by Case

If $B_1, B_2, \ldots$ is a partition of $\mathcal{S}$ and A is any event, then

$$\Pr(A) = \sum_{n=1}^{\infty} \Pr(A \cap B_n).$$

Exercises

1. Toss a fair coin five times.
   1. What is the probability of seeing exactly one H?
   2. What is the probability of seeing **at least** two H's?
2. In a community, 23% are male democrat and 25% are female democrats. If a person is chosen at random, what is the probability he/she is a democrat?

Addition rule (Rule 4)

- $n = 2$ events:

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

- $n = 3$ events:

$$\begin{aligned} \Pr(A \cup B \cup C) = {} & \Pr(A) + \Pr(B) + \Pr(C) \\ & - \Pr(A \cap B) - \Pr(A \cap C) - \Pr(B \cap C) \\ & + \Pr(A \cap B \cap C) \end{aligned}$$

- The *Inclusion-Exclusion principle* provides a formula for general n, but it's very messy!
- If the events are mutually exclusive, then Axiom 3 says we just add the individual probabilities.

1. Two friends, Adam and Brian, are visiting you from out of town. They are scheduled to arrive at the airport at the same time, but on different flights. Prior experience with these airlines suggests that Adam's flight will be late with probability 0.25, Brian's flight will be late with probability 0.3, and both will be late with probability 0.18. If you are picking up your two friends from the airport, what is the probability you will have to wait?
2. A total of 28% of American males smoke cigarettes, 7% smoke cigars, and 5% smoke both cigarettes and cigars. What % of males smoke neither cigarettes nor cigars?

Outline

Motivating Example 1: Dog and Cat

- What's the probability to get a dog if you randomly pick one from the pool of 10 male dogs, 10 female dogs, 3 male cats and 7 female cats?—**Sample Space?**
- What's the probability to get a dog if you randomly pick one from the pool of 10 male dogs, 10 female dogs, 3 male cats and 7 female cats, given that you are told that the one you picked is female?—**Sample Space?**

Motivating Example 2: Tossing Coin

Toss a fair coin three times, and define

$$A = \{\text{exactly two H's}\} \quad \text{and} \quad B = \{\text{first toss is H}\}.$$

1. What is the sample space?
2. What is $\Pr(A)$?
3. What is $\Pr(B)$?
4. What is $\Pr(A \cap B)$?
5. What is the probability of A given that B already happened?

Definition

- The *conditional probability* of A, given an event B occured, is defined as

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

  (Of course, $\Pr(B)$ must be > 0 since B actually occured!)
- That is, we have shrunk the sample space from $\mathcal{S}$ to B and the conditional probability is just the relative probability of A i.e. the % of B that is also A.

Exercise

Suppose $\Pr(A) = 0.55$, $\Pr(B) = 0.20$, and $\Pr(B \mid A) = 0.30$. Find the following probabilities:

1. $\Pr(A \cap B)$
2. $\Pr(A \cup B)$
3. $\Pr(A^c \cap B^c)$
4. $\Pr(A \mid B)$

Multiplication rule (Rule 5)

- We know that solving things **step by step** is really powerful.
- For the probability of intersection $(A \cap B)$, we can also get it by step and step.
  1. Step 1: Let A happen first.$\rightarrow \Pr(A)$
  2. Step 2: Given A happened, let B happen then.$\rightarrow \Pr(B \mid A)$
- *Multiplication rule*:

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

- Important consequence: by symmetry

$$\Pr(B) \Pr(A \mid B) = \Pr(A) \Pr(B \mid A). \qquad (\star)$$

Powerful idea 2: solving thing step by step

- We can extend the multiplication rule to more than two events following the **step by step** idea.
- Here it is for three events A, B and C:

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

$$A \Rightarrow B \mid A \Rightarrow C \mid A \cap B$$

Birthday problem – statement

Suppose n people are in a room. What is the probability that at least two people share the same birthday (month and day)? Assume that all 365 days (no leap year!) are equally likely.

- What do you think?
- Is this probability high or low?
- It depends on the value of n, right?

Birthday problem – solution

Let E_n be the event that at least two people share a birthday, but consider finding $\Pr(E_n^c)$, the probability that all n have different birthdays. Then

- How many different cases in total: 365^n (each day can be the birthday for any individual);
- How many cases lead to no shared birthday: $365 \times 364 \times 363 \times \cdots \times (365 - n + 1) = (365)_n$ —Permutation of n from 365;
- Result:

$$\Pr(E_n^c) = \frac{(365)_n}{365^n}.$$

Therefore,

$$\Pr(E_n) = 1 - \frac{(365)_n}{365^n}.$$

It's not easy to tell from this formula if the probability is large or small, but we can easily plot this as a function of n in order to visualize what's happening.

Birthday problem – solution

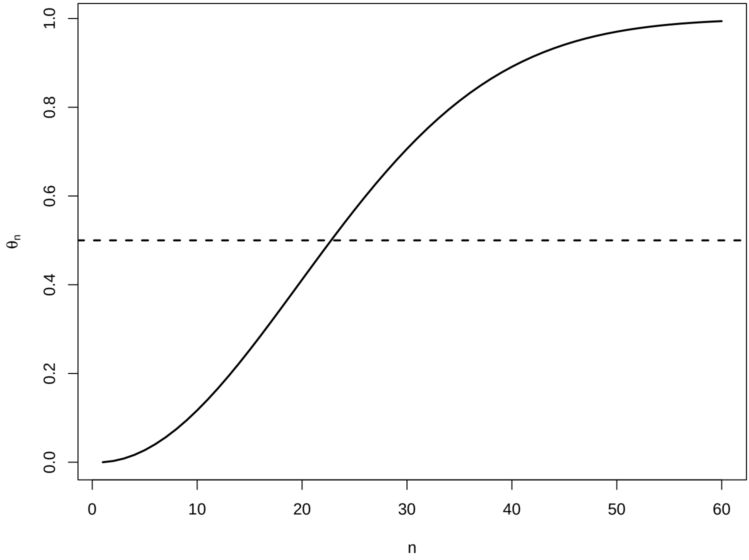
Let E_n be the event that at least two people share a birthday, but consider finding $\Pr(E_n^c)$, the probability that all n have different birthdays. Then solve it using Multiplication Rule:

- A_1: the first person does not share his/her birthday with previously analyzed people—$\Pr(A_1) = 1$
- A_2: the 2nd person does not share his/her birthday with previously analyzed people—$\Pr(A_2|A_1) = 364/365$
- A_3: the 3rd person does not share his/her birthday with previously analyzed people—$\Pr(A_3|A_1 \cap A_2) = 363/365$
- ...
- A_n: the n-th person does not share his/her birthday with previously analyzed people—$\Pr(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}) = (365 - n + 1)/365$

$$\Pr(E_n^c) = \frac{365}{365} \cdot \frac{364}{365} \cdots \frac{365 - n + 1}{365} = \frac{(365)_n}{365^n}$$

Therefore, $\Pr(E_n) = 1 - \frac{(365)_n}{365^n}$.

Birthday problem, cont.



Two powerful ideas together: Law of Total Probability

- Sometimes an experiment can be done case by case.
- And for each case, the experiment can be done step by step.
- Recall: A sequence of events $A_1, A_2, \ldots$ is a partition of $\mathcal{S}$ if $A_i \cap A_j = \emptyset$ for each $i \neq j$, and $\bigcup_{n=1}^{\infty} A_n = \mathcal{S}$.

Theorem (Law of Total Probability: Divide and Conquer)

If $A_1, A_2, \ldots$ is a partition of $\mathcal{S}$ and B is any event, then

$$\Pr(B) = \sum_{n=1}^{\infty} \Pr(B \mid A_n) \Pr(A_n).$$

Exercise: Marble balls

Suppose we have two boxes that contain marble balls. 25% of balls in Box 1 are white, 50% of balls in Box 2 are black. If we randomly choose a box, and then randomly select a ball from the chosen box, what is the probability we get a white marble?

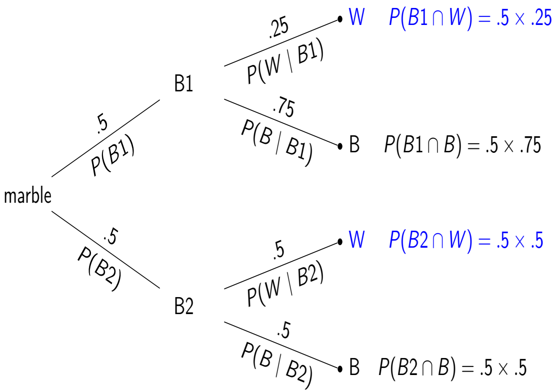
Outline

Motivating Example: Marble balls

Reconsider the box/marble problem before. Suppose we have two boxes that contain marble balls. 25% of balls in Box 1 are white, 50% of balls in Box 2 are black.

1. Given the marble selected was from Box 1, what is the probability we get a white marble?
2. What is the probability we get a white marble?
3. Given the marble selected was white, what is the probability it came from Box 1?

Tree Diagram



1. Given the marble selected was from Box 1, what is the probability we get a white marble?
2. What is the probability we get a white marble?
3. Given the marble selected was white, what is the probability it came from Box 1?

Bayes' Theorem (Rule 6)

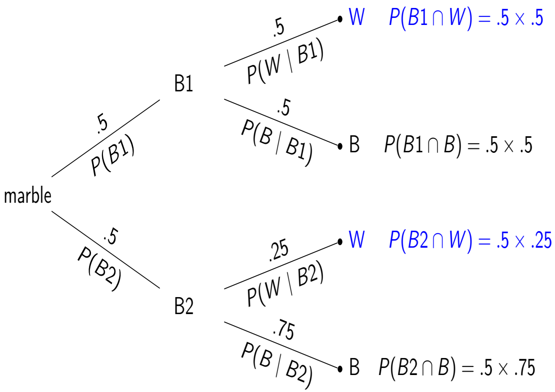
- What if we know $\Pr(B \mid A)$ but want $\Pr(A \mid B)$?
- Bayes' theorem tells us how to do it!
- It's a simple consequence of the symmetry $(\star)$ mentioned above.

Theorem (Bayes)

$$\Pr(A \mid B) = \frac{\Pr(B \mid A)\Pr(A)}{\Pr(B)}.$$

Note: the denominator can be obtained via the law of total probability.

Tree Diagram



1. Given the marble selected was from Box 1, what is the probability we get a white marble?
2. What is the probability we get a white marble?
3. Given the marble selected was white, what is the probability it came from Box 1?

Exercise

Suppose a factory has two machines A and B that produce 60% and 40% of the total output, respectively. Suppose also that 3% of machine A's output and 5% of machine B's output is defective.

1. If an item is chosen at random off the production line, what is the probability it will be defective?
2. If the item chosen is defective, what is the probability it came from machine B?

Outline

Independent events

- If knowing that an event A occurs does not change the probability that B occurs, then A and B are called **independent**.
- More formally, A and B are independent if and only if

$$\Pr(B \mid A) = \Pr(B).$$

- *Special multiplication rule*: Equivalently, A and B are independent if and only if

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

  Can you prove the equivalence?
- Note that independence is symmetric!

Example – coin tossing

Toss a fair coin six times. Clearly, without any calculations, we know that the tosses are independent. Now, if we let A_i be the event that the i^{th} toss is Heads, then

$$\begin{aligned}\Pr(\text{all Heads}) &= \Pr(A_1 \cap A_2 \cap \cdots \cap A_6) \\ &= \prod_{i=1}^{6} \Pr(A_i) = (1/2)^6.\end{aligned}$$

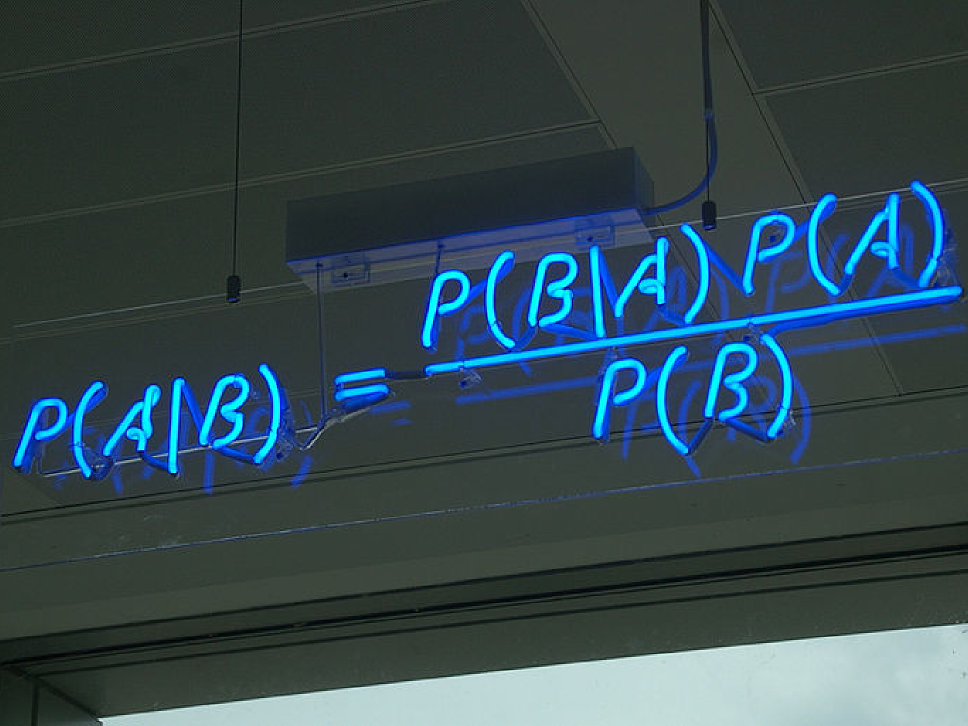
Using the complementation rule, we then obtain

$$\Pr(\text{at least one Tail}) = 1 - \Pr(\text{all Heads}) = 1 - (1/2)^6.$$

Independence and mutual exclusiveness

- Since the use of language in mathematics and everyday life often do not agree, there can be some confusion between independent and mutually exclusive events.
- Remember, A and B are independent if knowing one occurs does not affect the chance that the other occurs.
- However, if A and B are mutually exclusive, knowing one occurs implies the other did not occur.
- **Claim:** If A and B are events with positive probability which are mutually exclusive, then they are *not* independent.

Outline


$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}$

Bayesian Reasoning

The essence of Bayesian reasoning is: **updating your beliefs when you get new evidence!**

$$\Pr(H \mid E) = \frac{\Pr(E \mid H)\Pr(H)}{\Pr(E)}.$$

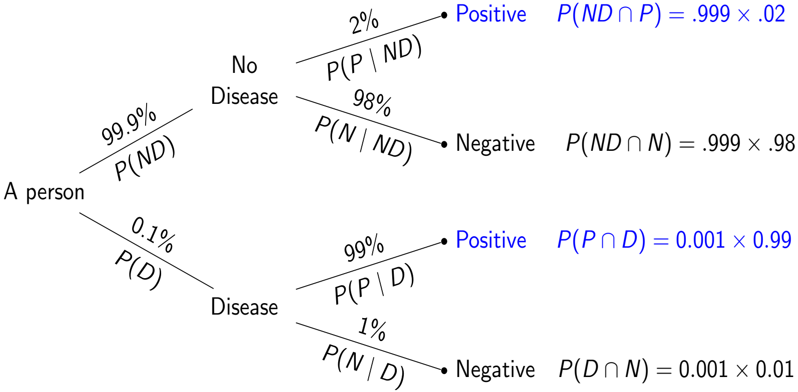
- $\Pr(H)$: **Prior** belief about the hypothesis H.
- $\Pr(E)$: **Marginal** probability for the evidence to happen
- $\Pr(E \mid H)$: **Likelihood** for the evidence to happen given the hypothesis H is true.
- $\Pr(H \mid E)$: **Posterior** belief about the hypothesis H in the light of the observed evidence E.

Example 1: Rare Disease

Suppose that you are worried that you might have a rare disease. You decide to get tested, and suppose that the testing methods for this disease are not perfect: if you have the disease, it shows that you do with 99 percent probability, and if you don't have the disease, it shows that you do not with 98 percent probability.
Suppose this disease is actually quite rare, occurring randomly in the general population in only one of every 1000 people.
If your test results come back positive, what are your chances that you actually have the disease?

$$A. \approx 5\% \qquad B. \approx 50\% \qquad C. 95\%$$

Example 1: Rare Disease



$$\Pr(D \mid P) = \frac{\Pr(P \mid D)\Pr(D)}{\Pr(P)} = \frac{\Pr(P \mid D)\Pr(D)}{\Pr(P \mid D)\Pr(D) + \Pr(P \mid ND)\Pr(ND)}$$

$$= \frac{0.001 \times 0.99}{0.001 \times 0.99 + .999 \times .02} = 4.72\%$$

Example 1: Rare Disease

Example 1: Rare Disease

1. Hypothesis: H_1: the person has no disease; H_2: the person has disease.

Example 1: Rare Disease

1. Hypothesis: H_1: the person has no disease; H_2: the person has disease.
2. Prior: $\Pr(H_1) = 99.9\%; \Pr(H_2) = 0.1\%$.

Example 1: Rare Disease

1. Hypothesis: H_1: the person has no disease; H_2: the person has disease.
2. Prior: $\Pr(H_1) = 99.9\%; \Pr(H_2) = 0.1\%$.
3. Evidence: test is positive.

Example 1: Rare Disease

1. Hypothesis: H_1: the person has no disease; H_2: the person has disease.
2. Prior: $\Pr(H_1) = 99.9\%; \Pr(H_2) = 0.1\%$.
3. Evidence: test is positive.
4. Likelihood: $\Pr(E|H_1) = 2\%, \Pr(E|H_2) = 99\%$.

Example 1: Rare Disease

1. Hypothesis: H_1: the person has no disease; H_2: the person has disease.
2. Prior: $\Pr(H_1) = 99.9\%; \Pr(H_2) = 0.1\%$.
3. Evidence: test is positive.
4. Likelihood: $\Pr(E|H_1) = 2\%, \Pr(E|H_2) = 99\%$.
5. Marginal: $\Pr(E) = \Pr(E|H_1)\Pr(H_1) + \Pr(E|H_2)\Pr(H_2) = 0.001 \times 0.99 + .999 \times .02$.

Example 1: Rare Disease

1. Hypothesis: H_1: the person has no disease; H_2: the person has disease.
2. Prior: $\Pr(H_1) = 99.9\%; \Pr(H_2) = 0.1\%$.
3. Evidence: test is positive.
4. Likelihood: $\Pr(E|H_1) = 2\%, \Pr(E|H_2) = 99\%$.
5. Marginal: $\Pr(E) = \Pr(E|H_1)\Pr(H_1) + \Pr(E|H_2)\Pr(H_2) = 0.001 \times 0.99 + .999 \times .02$.
6. Posterior: $\Pr(H_2|E) = \frac{\Pr(E|H_2)\Pr(H_2)}{\Pr(E)} = \frac{0.001\times0.99}{0.001\times0.99+.999\times.02} = 4.72\%; \Pr(H_1|E) = \frac{\Pr(E|H_1)\Pr(H_1)}{\Pr(E)} = 100\% - 4.72\%$.

Example 1: Rare Disease

1. This result seems counter-intuitive. The diagnostic test appears so accurate that we expect someone with a positive test result to be highly likely to have the disease, whereas the computed posterior probability is only 4.72%. In fact, the rarity of the disease implies that most positive test results arise from errors rather than from diseased individuals. But the chance of having the disease has increased by a multiplicative factor of 47.2. But to claim disease, we have to do further test.
2. For a person whose test shows positive, if we perform the test to him one more time and the result shows positive again, what's the probability that he has the disease? (70.9%)

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

1. Hypothesis: H_1: car in door 1; H_2: car in door 2; H_3: car in door 3.

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

1. Hypothesis: H_1: car in door 1; H_2: car in door 2; H_3: car in door 3.
2. Prior: $\Pr(H_1) = \Pr(H_2) = \Pr(H_3) = 1/3$.

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

1. Hypothesis: H_1: car in door 1; H_2: car in door 2; H_3: car in door 3.
2. Prior: $\Pr(H_1) = \Pr(H_2) = \Pr(H_3) = 1/3$.
3. Evidence: You choose door 1, and the host open door 3, which for sure has goat.

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

1. Hypothesis: H_1: car in door 1; H_2: car in door 2; H_3: car in door 3.
2. Prior: $\Pr(H_1) = \Pr(H_2) = \Pr(H_3) = 1/3$.
3. Evidence: You choose door 1, and the host open door 3, which for sure has goat.
4. Likelihood: $\Pr(E|H_1) = 1/2, \Pr(E|H_2) = 1, \Pr(E|H_3) = 0$.

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

1. Hypothesis: H_1: car in door 1; H_2: car in door 2; H_3: car in door 3.
2. Prior: $\Pr(H_1) = \Pr(H_2) = \Pr(H_3) = 1/3$.
3. Evidence: You choose door 1, and the host open door 3, which for sure has goat.
4. Likelihood: $\Pr(E|H_1) = 1/2, \Pr(E|H_2) = 1, \Pr(E|H_3) = 0$.
5. Marginal: $\Pr(E) =$ $\Pr(E|H_1)\Pr(H_1) + \Pr(E|H_2)\Pr(H_2) + \Pr(E|H_3)\Pr(H_3) = 1/2$.

Example 2: Monty Hall Problem

Monty Hall Problem: Suppose you're on a game show, and you're given the choice of three doors: Behind one door is a car; behind the others, goats. You pick a door, say No. 1, and the host, who knows what's behind the doors, opens another door, say No. 3, which has a goat. He then says to you, "Do you want to pick door No. 2?" Is it to your advantage to switch your choice?

1. Hypothesis: H_1: car in door 1; H_2: car in door 2; H_3: car in door 3.
2. Prior: $\Pr(H_1) = \Pr(H_2) = \Pr(H_3) = 1/3$.
3. Evidence: You choose door 1, and the host open door 3, which for sure has goat.
4. Likelihood: $\Pr(E|H_1) = 1/2, \Pr(E|H_2) = 1, \Pr(E|H_3) = 0$.
5. Marginal: $\Pr(E) = \Pr(E|H_1)\Pr(H_1) + \Pr(E|H_2)\Pr(H_2) + \Pr(E|H_3)\Pr(H_3) = 1/2$.
6. Posterior: $\Pr(H_1|E) = \frac{\Pr(E|H_1)\Pr(H_1)}{\Pr(E)} = 1/3; \Pr(H_2|E) = \frac{\Pr(E|H_2)\Pr(H_2)}{\Pr(E)} = 2/3; \Pr(H_3|E) = \frac{\Pr(E|H_3)\Pr(H_3)}{\Pr(E)} = 0$.