DSE 210: Probability and Statistics

Overview

Outline

- Introduction
  - Myself
  - TAs
  - Course
  - Probability and Statistics Overview

Myself

- Volkan Vural (vvural@ucsd.edu)
- Ph.D. in Machine Learning
- Research in Classification/ Computer Aided Diagnosis
- 6+ years of experience in Investment Technologies
- Boston College / MBA classes
  - Forecasting in Business and Economics
  - Machine Learning for Business Intelligence

TAs

- **Ismail Ocak (iocak@eng.ucsd.edu)**
  - UCSD ECE Grad Student

- **Rithesh R N (rramapur@eng.ucsd.edu)**
  - UCSD CSE Grad Student

Course

- Probability and Statistics
  - Important concepts & methods
  - Both theory and applications
- Practical and application oriented
- Python will be used for implementations

- Important links
  - Website: https://mas-dse.github.io/DSE210/
  - GitHub: https://mas-dse.github.io/DSE210/
  - Piazza: https://piazza.com/ucsd/winter2020/dse210
  - Gradescope: https://gradescope.com (Entry code : MZGGBE)
  - Canvas: https://canvas.ucsd.edu

Course Structure & Grading

- Course Structure
  - Lectures : 1/3, 1/17, 1/31, 2/14, 2/28, 3/13
  - 4 HWs (2 weeks/ HW)
  - One final
- Grading
  - HW1 (release: 1/3, deadline: 1/16) - 15%
  - HW2 (release: 1/17, deadline: 1/30) - 15%
  - HW3 (release: 1/31, deadline: 2/13) - 15%
  - HW4 (release: 2/14, deadline: 2/27) - 15%
  - Final - 40%

The kinds of questions we'll study

- Design a spam filter.
- What fraction of San Diegans occasionally smoke pot?
- Categorize New York Times articles by their underlying topics.
- Two new malaria vaccines are under consideration. How can we determine which is better?
- We've obtained user ratings of many movies. Visualize them.
- A dating service asks each user to answer 200 multiple choice questions. Summarize each user's responses by a few numbers.

Intermediate-level questions

- Regression. How do you fit a line to a set of points?
- Clustering. Given a bunch of data points, partition them into groups that are distinct from each other.
- Laws of large numbers. A drunk starts off from a bar and at each time step, takes either a step to the right or a step to the left. Where will he be, approximately, after n time steps?
- Hypothesis testing. You are given two alternatives and wish to test which is better. Design an experiment to do this.
- Dimensionality reduction. Find the primary axes of variation in a data set.

Low-level questions

- If you toss a coin 10 times, what is the chance of getting heads every time?
- Throw 20 balls into 20 bins at random. What is the probability that at least one of the bins remains empty?
- If each cereal box contains one of k action figures, how many boxes do you need to buy, on average, before getting all the figures?
- What fraction of a bell curve lies at least one standard deviation away from the mean?
- Find a concise description of a data matrix.

Outline : Lecture 1

- Sets and Counting
  - Sets, Tuples
  - Set Operations - Union, Intersection
  - Permutations
  - Combinations
  - Hands on
  - Self - practice
- Probability Space
  - Sample space and probability of Outcomes
  - Examples
  - Hands on
  - Self - practice
- Multiple events, conditioning, and independence
  - Conditional Probability
  - Summation Rule
  - Bayes’ rule
  - Independence
  - Hands on
  - Self - practice

Outline : Lecture 2

- Random Variables, Expectation and Variance
  - Random Variable
  - Expected Value
  - Variance
  - Sampling
  - Hands on
  - Self - practice
- Modeling data with Probability distributions
  - Binomial distribution
  - MLE
  - Normal Distribution
  - Multinomial Distribution
  - Poisson Distribution
  - Hands on
  - Self - practice
- Classification with Generative Models 1
  - Classification and Regression
  - Generative Models
  - Univariate Gaussian
  - Hands on
  - Self - practice

Outline : Lecture 3

- Classification with Generative Models 2
  - Bivariate Distributions
  - Covariance and Correlation
  - Bivariate Gaussian
  - Hands on
  - Self - practice
- Linear Algebra Primer
  - Vectors and Matrices
  - Quadratic Functions
  - Hands on
  - Self - practice
- Classification with Generative Models 3
  - Multivariate Gaussian
  - Multi nominal Naive Bayes
  - Hands on
  - Self - practice

Outline : Lecture 4

- Clustering
  - K- means
  - Mixtures of Gaussians
  - The EM algorithm
  - Linkage methods
  - Hands on
  - Self - practice
- Informative Projections
  - Dimensionality reduction
  - Projections
  - Hands on
  - Self - practice
- Singular Value Decomposition (SVD)
  - Eigen Values and Eigen Vectors
  - SVD
  - Hands on
  - Self - practice

Outline : Lecture 5

- Sampling
  - Central Limit Theorem
  - Sampling design
  - Hands on
  - Self - practice
- Experimental Design and Hypothesis testing
  - Experimental Design
  - Hypothesis testing : z-statistics and X^2 statistics
  - Hands on
  - Self - practice

Sets and counting

DSE 210

Outline : Lecture 1

- **Sets and Counting**
  - **Sets, Tuples**
  - **Set Operations - Union, Intersection**
  - **Permutations**
  - **Combinations**
  - **Hands on**
  - **Self - practice**
- Probability Space
  - Sample space and probability of Outcomes
  - Examples
  - Hands on
  - Self - practice
- Multiple events, conditioning, and independence
  - Conditional Probability
  - Summation Rule
  - Bayes' rule
  - Independence
  - Hands on
  - Self - practice

Sets

$A = \{a, b, c, \ldots, z\} \quad |A| = 26$

$B = \{0, 1\} \quad |B| = 2$

$E = \{\text{all even integers}\} \quad |E| = \infty$

$S = \{x \in E : x \text{ is a multiple of } 3\}$

$I = [0, 1] = \{x : 0 \leq x \leq 1\}$

In a set, the *order* of elements doesn't matter:

$$\{0, 1, 2\} = \{2, 0, 1\}$$

and there are no duplicates.

Tuples

Let $C = \{H, T\}$.

All pairs of elements from C:

$$\{(H, H), (H, T), (T, H), (T, T)\} = C \times C = C^2$$

All triples of elements of C:

$$\{(H, H, H), (H, H, T), (H, T, H), \ldots\} = C \times C \times C = C^3$$

All sequences of k elements from C: denoted $C^k = C \times C \times \cdots \times C$.

How many sequences of length k are there? $|C^k| = |C|^k = 2^k$.

In a sequence, the order of elements matters:

$$(H, T) \neq (T, H).$$

Let $A = \{a, b, c, \ldots, z\}$.

How many sequences of length 2? 26^2

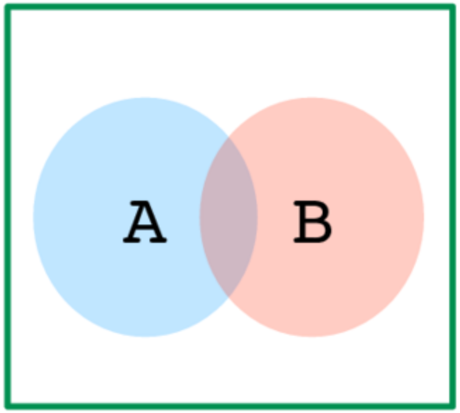
How many sequences of length 10? 26^{10}

How many sequences of length n? 26^n

An alien language has an alphabet of size 10. Every sequence of ≤ 5 of these characters is a valid word. How many words are there in this language?

$$10^1 + 10^2 + 10^3 + 10^4 + 10^5 = 10 + 100 + 1000 + 10000 + 100000 = 111110.$$

Union and intersection



$A \cup B = \{\text{any element in } A \text{ or in } B \text{ or in both}\}$

$A \cap B = \{\text{any element in } A \text{ and in } B\}$

$M = \{2, 3, 5, 7, 11\}$ and $N = \{1, 3, 5, 7, 9\}$

$M \cup N = \{1, 2, 3, 5, 7, 9, 11\}$

$M \cap N = \{3, 5, 7\}$

$S = \{\text{all even integers}\}$ and $T = \{\text{all odd integers}\}$

$S \cup T = \{\text{all integers}\}$

$S \cap T = \emptyset$

Permutations

How many ways to order the three letters A, B, C?

$ABC, ACB, BAC, BCA, CAB, CBA$
3 choices for the first, 2 choices for the second, 1 choice for the third
$3 \times 2 \times 1 = 6$. Call this 3!

How many ways to order A, B, C, D, E?

$5 \times 4 \times 3 \times 2 \times 1 = 5! = 120$

How many ways to place 6 men in a line-up?

$6 \times 5 \times 4 \times 3 \times 2 \times 1 = 6! = 720$

How many possible outcomes of shuffling a deck of cards?

52!

General rule: The number of ways to order n distinct items is:

$$n! = n(n-1)(n-2)\cdots 1.$$

Combinations

An ice-cream parlor has flavors $\{\text{chocolate}, \text{vanilla}, \text{strawberry}, \text{pecan}\}$. You are allowed to pick two of them. How many options do you have?

$$CV, CS, CP, VS, VP, SP$$

In general, the number of ways to pick k items out of n is:

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

For instance, $\binom{4}{2} = \frac{4\cdot 3}{2!} = 6$.

How many ways to pick three ice-cream flavors?

$$\binom{4}{3} = 4$$

Pick any 4 of your favorite 100 songs. How many ways to do this?

$$\binom{100}{4} = \frac{100 \cdot 99 \cdot 98 \cdot 97}{4 \cdot 3 \cdot 2 \cdot 1}$$

Probability spaces

DSE 210

Outline : Lecture 1

- Sets and Counting
  - Sets, Tuples
  - Set Operations - Union, Intersection
  - Permutations
  - Combinations
  - Hands on
  - Self - practice
- **Probability Space**
  - **Sample space and probability of Outcomes**
  - **Examples**
  - **Hands on**
  - **Self - practice**
- Multiple events, conditioning, and independence
  - Conditional Probability
  - Summation Rule
  - Bayes’ rule
  - Independence
  - Hands on
  - Self - practice

Probability spaces

How to interpret a statement like:

The chance of getting a flush in a five-card poker hand is about 0.20%. (Flush = five of the same suit.)

The underlying **probability space** has two components:

1. The **sample space** (the space of outcomes).

   In the example, $\Omega = \{\text{all possible five-card hands}\}$.

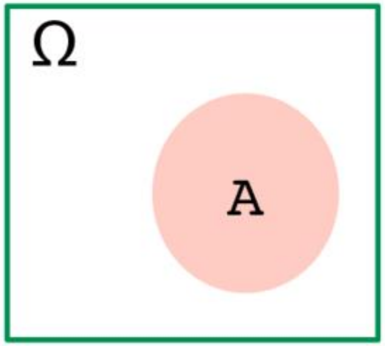
2. The **probabilities of outcomes**.

   In the example, all hands are equally likely: probability $1/|\Omega|$.

Note: $\sum_{\omega \in \Omega} \Pr(\omega) = 1$.

Event of interest: the set of outcomes $A = \{\omega : \omega \text{ is a flush}\} \subset \Omega$.

$$\Pr(A) = \sum_{\omega \in A} \Pr(\omega) = \frac{|A|}{|\Omega|}$$



Examples

Roll a die. What is the chance of getting a number > 3?

Sample space $\Omega = \{1, 2, 3, 4, 5, 6\}$.

Probabilities of outcomes: $\Pr(\omega) = \frac{1}{6}$.

Event of interest: $A = \{4, 5, 6\}$

$$\Pr(A) = \Pr(4) + \Pr(5) + \Pr(6) = \frac{1}{2}.$$

Roll three dice. What is the chance that their sum is 3?

Sample space

$$\begin{aligned}\Omega &= \{(1,1,1),(1,1,2),(1,1,3),\ldots,(6,6,6)\} \\ &= \Omega_o \times \Omega_o \times \Omega_o\end{aligned}$$

where $\Omega_o = \{1,2,3,4,5,6\}$.

Probabilities of outcomes:

$$\Pr(\omega) = \frac{1}{|\Omega|} = \frac{1}{216}$$

Event of interest: $A = \{(1,1,1)\}$. $\Pr(A) = \frac{1}{216}$.

Roll n dice.

Then $\Omega = \Omega_o \times \cdots \times \Omega_o = \Omega_o^n$, where $\Omega_o = \{1,2,3,4,5,6\}$.

What is $|\Omega|$? 6^n.

Probability of an outcome: $\Pr(\omega) = \frac{1}{6^n}$.

Socks in a drawer. A drawer has three blue socks and three red socks. You put your hand in and pull out two socks at random. What is the probability that they match?

Think of grabbing one sock first, then another.

$\Omega = \{(B,B),(B,R),(R,B),(R,R)\} = \{B,R\}^2$.

Probabilities:

$$\begin{aligned}
\Pr((B,B)) &= \frac{1}{2}\cdot\frac{2}{5} = \frac{1}{5} \\
\Pr((B,R)) &= \frac{1}{2}\cdot\frac{3}{5} = \frac{3}{10} \\
\Pr((R,B)) &= \frac{1}{2}\cdot\frac{3}{5} = \frac{3}{10} \\
\Pr((R,R)) &= \frac{1}{2}\cdot\frac{2}{5} = \frac{1}{5}
\end{aligned}$$

Event of interest: $A = \{(B,B),(R,R)\}$. $\Pr(A) = \frac{2}{5}$.

Socks in a drawer, cont'd. This time the drawer has three blue socks and two red socks. You put your hand in and pull out two socks at random. What is the probability that they match?

Sample sample space, $\Omega = \{(B,B),(B,R),(R,B),(R,R)\} = \{B,R\}^2$.

Different probabilities:

$$\begin{aligned}
\Pr((B,B)) &= \frac{3}{5}\cdot\frac{2}{4} = \frac{3}{10}\\
\Pr((B,R)) &= \frac{3}{5}\cdot\frac{2}{4} = \frac{3}{10}\\
\Pr((R,B)) &= \frac{2}{5}\cdot\frac{3}{4} = \frac{3}{10}\\
\Pr((R,R)) &= \frac{2}{5}\cdot\frac{1}{4} = \frac{1}{10}
\end{aligned}$$

Event of interest: $A = \{(B,B),(R,R)\}$. $\Pr(A) = \frac{2}{5}$.

Shuffle a pack of cards.

Sample space $\Omega = \{\text{all possible orderings of 52 cards}\}$.

What is $|\Omega|$?

$$52! = 52 \cdot 51 \cdot 50 \cdot 49 \cdots 3 \cdot 2 \cdot 1$$

Toss a fair coin 10 times. What is the chance none are heads?

Sample space $\Omega = \{H, T\}^{10}$. It includes, for instance, $(H, T, H, T, H, T, H, T, H, T)$.

What is $|\Omega|$? $2^{10} = 1024$.

For any sequence of coin tosses $\omega \in \Omega$, we have $\Pr(\omega) = \frac{1}{1024}$.

Event of interest: $A = \{(T, T, T, T, T, T, T, T, T, T)\}$. $\Pr(A) = \frac{1}{1024}$.

What is the probability of exactly one head?

Event of interest: $A = \{\omega \in \Omega : \omega \text{ has exactly one } H\}$.

What is $|A|$? 10.
Each sequence in A can be specified by the location of the one H, and there are 10 choices for this.

What is $\Pr(A)$? $\frac{10}{1024}$.

Toss a fair coin 10 times. What is the chance of exactly two heads?

Again, sample space $\Omega = \{H, T\}^{10}$, with $|\Omega| = 2^{10} = 1024$.

For any sequence of coin tosses $\omega \in \Omega$, we have $\Pr(\omega) = \frac{1}{1024}$.

Event of interest: $A = \{\omega \in \Omega : \omega \text{ has exactly two } H\text{'s}\}$.

What is $|A|$? $\binom{10}{2} = \frac{10 \cdot 9}{2} = 45$.
Each sequence in A can be specified by the locations of the two H's and there are $\binom{10}{2}$ choices for these locations.

What is $\Pr(A)$? $\frac{45}{1024}$.

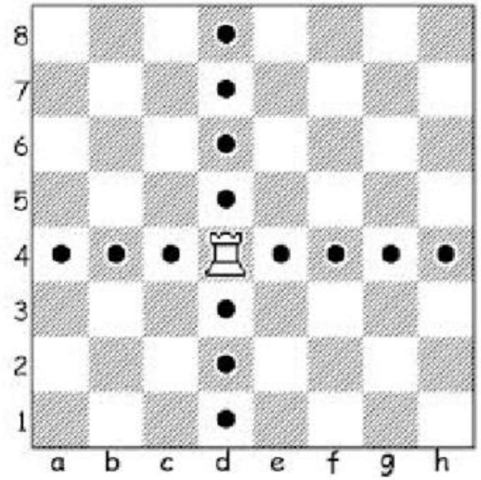
What is the probability of exactly k heads?

Event of interest: $A = \{\omega \in \Omega : \omega \text{ has exactly } k \ H\text{'s}\}$.

What is $|A|$? $\binom{10}{k}$.

What is $\Pr(A)$? $\binom{10}{k}/1024$.

Rooks on a chessboard.



What is the maximum number of rooks you can place so that no rook is attacking any other? There must be a rook in every row and every column: 8.

How many ways are there to place 8 rooks on the board, attacking or not? $\binom{64}{8}$

How many non-attacking placements of 8 rooks are there?
$8 \cdot 7 \cdot 6 \cdot 5 \cdots = 8!$

Randomly place 8 rooks on the board. What is the probability that it is a non-attacking placement?

$$\frac{8!}{\binom{64}{8}}.$$

Five-card poker. You are dealt 5 cards from a deck of 52.

Sample space $\Omega = \{\text{all possible hands}\}$.

Probabilities: each hand is equally likely, $\Pr(\omega) = 1/|\Omega|$.

What is $|\Omega|$? $\binom{52}{5}$.

What is the probability of getting a flush (five of the same suit)?
Event of interest: $F = \{\text{flush hands}\}$. Then $|F| = 4 \times \binom{13}{5}$
Therefore $\Pr(F) = |F|/|\Omega| = 4 \times \binom{13}{5}/\binom{52}{5}$.

What is the chance of a straight flush (flush, and in sequence)?
Let $S = \{\text{straight flush hands}\}$. Then $|S| = 4 \cdot 9 = 36$.
And $\Pr(S) = |S|/|\Omega|$.

Dartboard. A dartboard has radius 1 and its central bullseye has radius 0.1. You throw a dart and it lands at a random location on the board.

Sample space $\Omega = \{(x, y) : x^2 + y^2 \leq 1\}$.
All points are equally likely.

What is the probability of hitting the bullseye?

Event of interest: $B = \{(x, y) : x^2 + y^2 \leq (0.1)^2\}$.

$\Pr(B) = \text{area}(B)/\text{area}(\Omega) = \dfrac{\pi(0.1)^2}{\pi(1)^2} = 0.01.$

What is the probability of hitting the exact center? The geometric points do not have any length or area. 0

Birthday paradox. A room contains k people. What is the chance that they all have different birthdays?

Number the people $1, 2, \ldots, k$.
Number the days of the year $1, 2, \ldots, 365$.

Let $\omega = (\omega_1, \ldots, \omega_k)$, where $\omega_i \in \{1, 2, \ldots, 365\}$ is the birthday of person i. Thus $\Omega = \{1, 2, \ldots, 365\}^k$.

What is $|\Omega|$? 365^k.

Event of interest: $A = \{(\omega_1, \ldots, \omega_k) : \text{all } \omega_i \text{ different}\}$.
What is $|A|$? $365 \cdot 364 \cdot 363 \cdots (365 - k + 1)$.

Therefore,

$$\Pr(A) = \frac{365 \cdot 364 \cdots (365 - k + 1)}{365^k} = \frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{365 - k + 1}{365}$$

For $k = 23$, this is less than $1/2$. In other words, **in a group of 23 random people, chances are some pair of them have a common birthday!**

Multiple events, conditioning, and independence

DSE 210

Outline : Lecture 1

- Sets and Counting
  - Sets, Tuples
  - Set Operations - Union, Intersection
  - Permutations
  - Combinations
  - Hands on
  - Self - practice
- Probability Space
  - Sample space and probability of Outcomes
  - Examples
  - Hands on
  - Self - practice
- **Multiple events, conditioning, and independence**
  - **Conditional Probability**
  - **Summation Rule**
  - **Bayes’ rule**
  - **Independence**
  - **Hands on**
  - **Self - practice**

People's probability judgements

Experiment by Kahneman-Tversky. Subjects were told:

> *Linda is 31, single, outspoken, and very bright. She majored in philosophy in college. As a student, she was deeply concerned with racial discrimination and other social issues, and participated in anti-nuclear demonstrations.*

They were then asked to rank three possibilities:

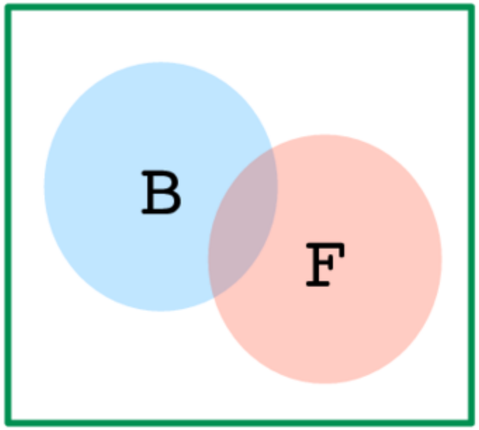
(a) Linda is active in the feminist movement.

(b) Linda is a bank teller.

(c) Linda is a bank teller and is active in the feminist movement.

Over 85% respondents chose (a) $>$ (c) $>$ (b).

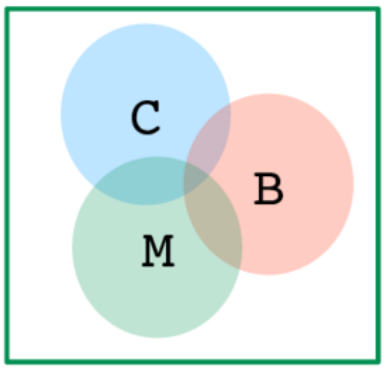
But:
$\Pr(\text{bank teller}, \text{feminist}) \leq \Pr(\text{bank teller})$.



In a city, 60% of people have a car, 20% of people have a bike, and 10% of people have a motorcycle. Anyone without at least one of these walks to work. What is the minimum fraction of people who walk to work?

Let $\Omega = \{\text{people in the town}\}$. Let
$C = \{\text{has car}\}, B = \{\text{has bike}\}, M = \{\text{has motorcyle}\}, W = \{\text{walks}\}$.

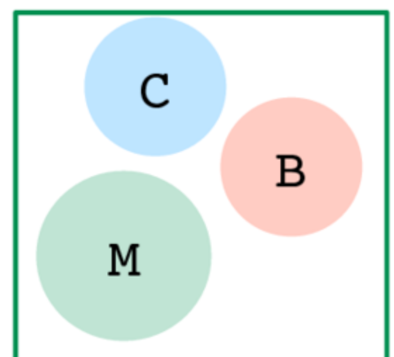
General picture:



$$\Pr(W) \geq 1 - \Pr(C \cup B \cup M)$$

$$\begin{aligned}\Pr(C \cup B \cup M) &\leq \Pr(C) + \Pr(B) + \Pr(M)\\ &= 0.6 + 0.2 + 0.1 = 0.9\end{aligned}$$

and thus $\Pr(W) \geq 0.1$.



Complements and unions

The complement of an event.

Let Ω be a sample space and $E \subset \Omega$ an event.
Write E^c for the event that E does not occur, that is, $E^c = \Omega \setminus E$.

$$\Pr(E^c) = 1 - \Pr(E).$$

The union bound.

For any events $E_1, \ldots, E_k$:

$$\Pr(E_1 \cup \cdots \cup E_k) \leq \Pr(E_1) + \cdots + \Pr(E_k).$$

This inequality is exact when the events are disjoint.

Coupon-collector problem

Each cereal box has one of k action figures. How many boxes do you need to buy so that you are likely to get all k figures?

Say we buy n boxes.
Let A_i be the event that the ith action figure is *not* obtained.

$$\begin{aligned} \Pr(A_i) &= \Pr(\text{not in 1st box}) \cdot \Pr(\text{not in 2nd box}) \cdots \Pr(\text{not in } n\text{th box}) \\ &= \left(1 - \frac{1}{k}\right)^n \leq e^{-n/k} \end{aligned}$$

By union bound, the probability of missing some figure is

$$\Pr(A_1 \cup \cdots \cup A_k) \leq \Pr(A_1) + \cdots + \Pr(A_k) \leq ke^{-n/k}.$$

Setting $n \geq k \ln 2k$ makes this $\leq 1/2$.

Therefore: enough to buy $O(k \log k)$ cereal boxes.

Coupon-collector problem - Proof

Each cereal box has one of k action figures. How many boxes do you need to buy so that you are likely to get all k figures?

Let T be the time to collect all n coupons, and let t_i be the time to collect the i-th coupon after $i-1$ coupons have been collected.

The probability of collecting a new coupon is $p_i = \dfrac{n-i+1}{n}$. Therefore, t_i has geometric distribution with expectation $\dfrac{1}{p_i}$ (We will see geometric distribution and its expectation later)

$$\mathrm{E(T)} = \mathrm{E}(t_1)+..+\mathrm{E}(t_n) = \frac{1}{p_1} + ... + \frac{1}{p_n} = \frac{n}{n} + \frac{n}{n-1} + ... + \frac{1}{n}$$
$$= \mathrm{n}\ .\ \left(\frac{1}{1} + ... + \frac{1}{n}\right)$$
$$\approx n\log n + \gamma n + \frac{1}{2}, \text{ where } \gamma \approx 0.5772 \text{ is the Euler–Mascheroni constant.}$$

Conditional probability

You meet a stranger at a bar. What is the chance he votes Republican?

Just use the average for your town: 0.5, say.

Now suppose you find out he plays tennis.

Sample space $\Omega = \{\text{all people in your town}\}$
Two events of interest:
$R = \{\text{votes Republican}\}$
$T = \{\text{plays tennis}\}$
What is $\Pr(R|T)$?

Formula for conditional probability:

$$\Pr(R|T) = \frac{\Pr(R \cap T)}{\Pr(T)} = \frac{\#\text{ people who vote Republican and play tennis}}{\#\text{ people who play tennis}}.$$

Pregnancy test.

The following data is obtained on a pregnancy test:

$\Omega = \{\text{women who use the test}\}$
$P = \{\text{women using the test who are actually pregnant}\}$
$T = \{\text{women for whom the test comes out positive}\}$

Suppose $T \subset P$ and $\Pr(P) = 0.4$ and $\Pr(T) = 0.3$.

Suppose the test comes out positive. What is the chance of pregnancy?
$\Pr(P|T) = \frac{\Pr(P \cap T)}{\Pr(T)} = \frac{0.3}{0.3} = 1$

Suppose the test comes out negative. What is the chance of pregnancy?

$$\Pr(P|T^c) = \frac{\Pr(P \cap T^c)}{\Pr(T^c)} = \frac{\Pr(P) - \Pr(T)}{1 - \Pr(T)} = \frac{0.1}{0.7} = \frac{1}{7}.$$

Rolls of a die.

You roll a die twice. What is the probability that the sum is ≥ 10:

If the first roll is 6?

$$\Pr(\text{sum} \geq 10 | \text{first} = 6) = \Pr(\text{second} \geq 4) = \frac{1}{2}.$$

If the first roll is ≥ 3?

$$\begin{aligned}\Pr(\text{sum} \geq 10 | \text{first} \geq 3) &= \frac{\Pr(\text{sum} \geq 10, \text{first} \geq 3)}{\Pr(\text{first} \geq 3)} \\ &= \frac{\Pr(\{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\})}{2/3} = \frac{1}{4}.\end{aligned}$$

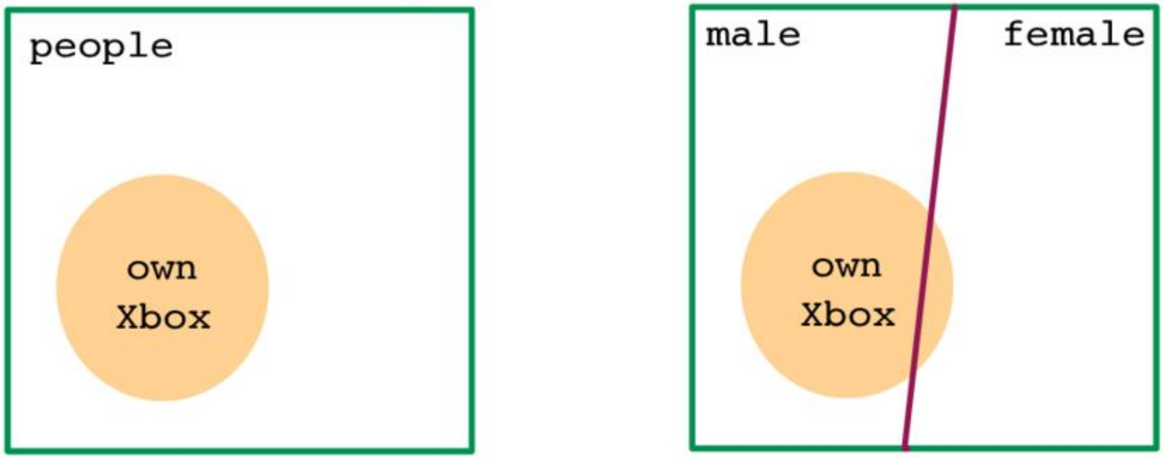
If the first roll is < 6?

$$\begin{aligned}\Pr(\text{sum} \geq 10 | \text{first} < 6) &= \frac{\Pr(\text{sum} \geq 10, \text{first} < 6)}{\Pr(\text{first} < 6)} \\ &= \frac{\Pr(\{(5,5),(5,6),(4,6)\})}{5/6} = \frac{1}{10}.\end{aligned}$$

Summation rule

Breaking down a probability into disjoint pieces.

Example: what fraction of people own an Xbox?



$$
\begin{aligned}
&\Pr(\text{own Xbox}) \\
&= \Pr(\text{own Xbox, male}) + \Pr(\text{own Xbox, female}) \\
&= \Pr(\text{own Xbox}|\text{male})\Pr(\text{male}) + \Pr(\text{own Xbox}|\text{female})\Pr(\text{female})
\end{aligned}
$$

Suppose events $A_1, \ldots, A_k$ are disjoint and $A_1 \cup \cdots \cup A_k = \Omega$: that is, one of these events must occur. Then for any other event E,

$$
\begin{aligned}
\Pr(E) &= \Pr(E, A_1) + \Pr(E, A_2) + \cdots + \Pr(E, A_k) \\
&= \Pr(E|A_1)\Pr(A_1) + \Pr(E|A_2)\Pr(A_2) + \cdots + \Pr(E|A_k)\Pr(A_k)
\end{aligned}
$$

The Monty Hall game

Three doors: one has a treasure chest behind it and the other two have goats. You pick a door and indicate it to Monty. He opens one of the other two doors to reveal a goat. Now, should you stick to your initial choice, or switch to the other unopened door?

You should switch.

First argument:

$\mathrm{Pr}(\text{initial choice has treasure}) = 1/3$
No matter what Monty does, he can't change this fact. So
$\mathrm{Pr}(\text{other unopened door has treasure}) = 2/3$

Second argument:

$$\begin{aligned}
&\mathrm{Pr}(\text{treasure in other door}) \\
&= \mathrm{Pr}(\text{treasure in other door}|\text{initial choice correct})\mathrm{Pr}(\text{initial choice correct}) + \\
&\quad \mathrm{Pr}(\text{treasure in other door}|\text{initial choice wrong})\mathrm{Pr}(\text{initial choice wrong}) \\
&= 0 \cdot \frac{1}{3} + 1 \cdot \frac{2}{3} = \frac{2}{3}.
\end{aligned}$$

Sex bias in graduate admissions

In 1969, there were 12673 applicants for graduate study at Berkeley. 44% of the male applicants were accepted, and 35% of the female applicants.

Define:

- $\Omega = \{\text{all applicants}\}$
- $M = \{\text{male applicants}\}$
- What is M^c? $M^c = \{\text{female applicants}\}$
- $A = \{\text{accepted applicants}\}$

What do the percentages 44% and 35% correspond to?

$\Pr(A|M) = 0.44$ and $\Pr(A|M^c) = 0.35$.

The administration found, however, that in every department, the accept rate for female applicants was at least as high as the accept rate for male applicants. How could this be?

Bayes' rule

Pearl: You wake up in the middle of the night to the shrill sound of your burglar alarm. What is the chance that a burglary has been attempted?

The facts:

- There is a 95% chance that an attempted burglary will trigger the alarm.

$$\Pr(\text{alarm}|\text{burglary}) = 0.95$$

- There is a 1% chance of a false alarm.

$$\Pr(\text{alarm}|\text{no burglary}) = 0.01$$

- Based on local crime statistics, there is a 1-in-10,000 chance that a given house will be burglarized on a given night.

$$\Pr(\text{burglary}) = 10^{-4}$$

We need to compute

$$\Pr(\text{burglary}|\text{alarm}) = \frac{\Pr(\text{burglary, alarm})}{\Pr(\text{alarm})} = \frac{\Pr(\text{alarm}|\text{burglary})\Pr(\text{burglary})}{\Pr(\text{alarm})}$$

We need to compute

$$\Pr(\text{burglary}|\text{alarm}) = \frac{\Pr(\text{alarm}|\text{burglary})\Pr(\text{burglary})}{\Pr(\text{alarm})}$$

Now,

$$\begin{aligned}\Pr(\text{alarm}) = {} & \Pr(\text{alarm}|\text{burglary})\Pr(\text{burglary}) + \\ & \Pr(\text{alarm}|\text{no burglary})\Pr(\text{no burglary})\end{aligned}$$

Therefore,

$$\Pr(\text{burglary}|\text{alarm}) = \frac{0.95 \times 10^{-4}}{0.95 \times 10^{-4} + 0.01 \times (1 - 10^{-4})} \approx 0.00941$$

The alarm increases one's belief in a burglary hundredfold, from $1/10000$ to roughly $1/100$.

Bayes' rule:

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

Independence

Two events A, B are **independent** if the probability of B occurring is the same whether or not A occurs.

Example: toss two coins.
$A = \{\text{first coin is heads}\}$
$B = \{\text{second coin is heads}\}$

Formally, we say A, B are independent if $\Pr(A \cap B) = \Pr(A)\Pr(B)$.

The independence of A and B implies:

- $\Pr(A|B) = \Pr(A)$
- $\Pr(B|A) = \Pr(B)$
- $\Pr(A|B^c) = \Pr(A)$

Examples: independent or not?

1. You have two children.

   $A = \{\text{first child is a boy}\}$, $B = \{\text{second child is a girl}\}$.

   Independent.

2. You throw two dice.

   $A = \{\text{first is a six}\}$, $B = \{\text{sum} > 10\}$.

   Not independent.

3. You get dealt two cards at random from a deck of 52.

   $A = \{\text{first is a heart}\}$, $B = \{\text{second is a club}\}$.

   Not independent: $\Pr(A) = \Pr(B) = 1/4$, but

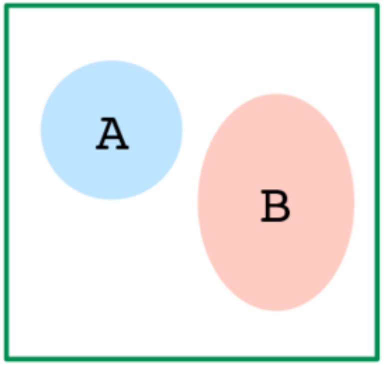
   $$\Pr(A \cap B) = \frac{1}{4} \cdot \frac{13}{51} > \Pr(A)\Pr(B).$$

4. You are dealt two cards.

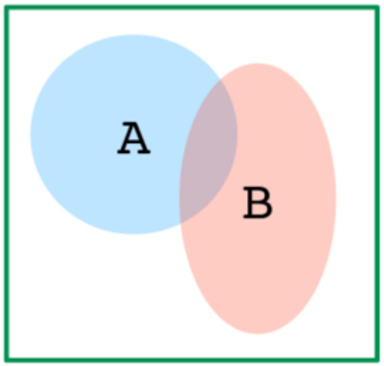
   $A = \{\text{first is a heart}\}$, $B = \{\text{second is a 10}\}$.

   Independent: $\Pr(A) = 1/4, \Pr(B) = 1/13$, and

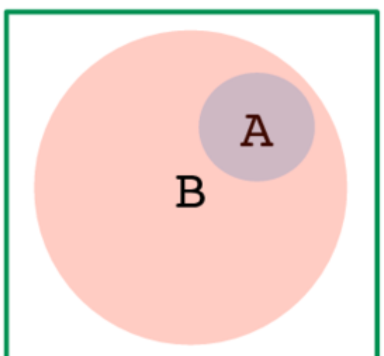
   $$\Pr(A \cap B) = \frac{12}{52} \cdot \frac{4}{51} + \frac{1}{52} \cdot \frac{3}{51} = \frac{1}{52} = \Pr(A)\Pr(B).$$



Not independent



Possibly independent



Not independent