CS365
Foundations of Data Science

Lectures 2, 3
(1/23, 25)

Charalampos E. Tsourakakis
ctsourak@bu.edu

What is a fair coin?

$\frac{1}{2}$

Are all coin flips random?

- Can a coin flip be “rigged”?
  - Yes!

- [Dynamical Bias in the Coin Toss by P Diaconis, S Holmes, R Montgomery]

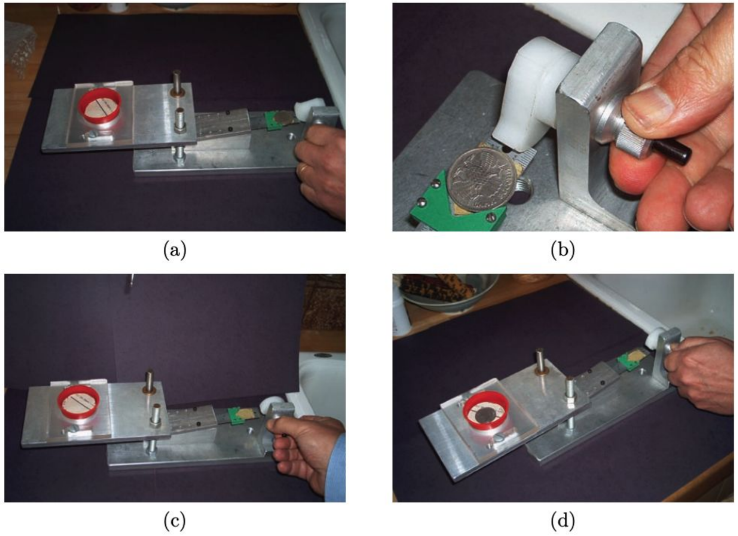
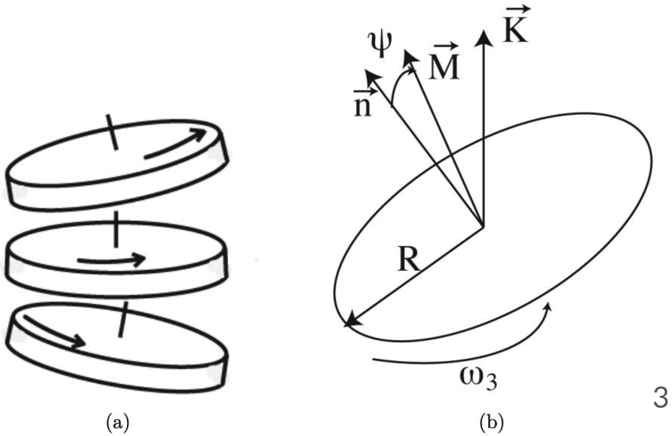
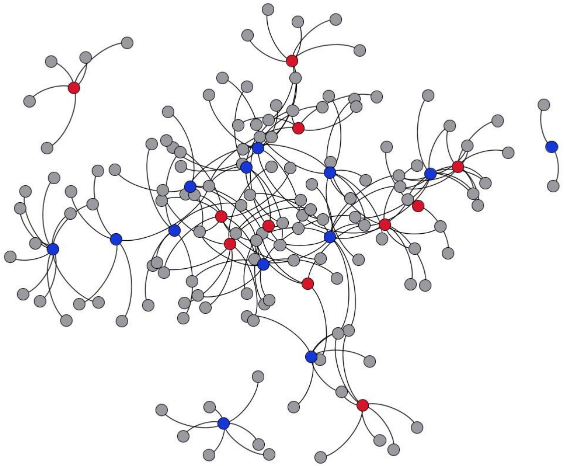


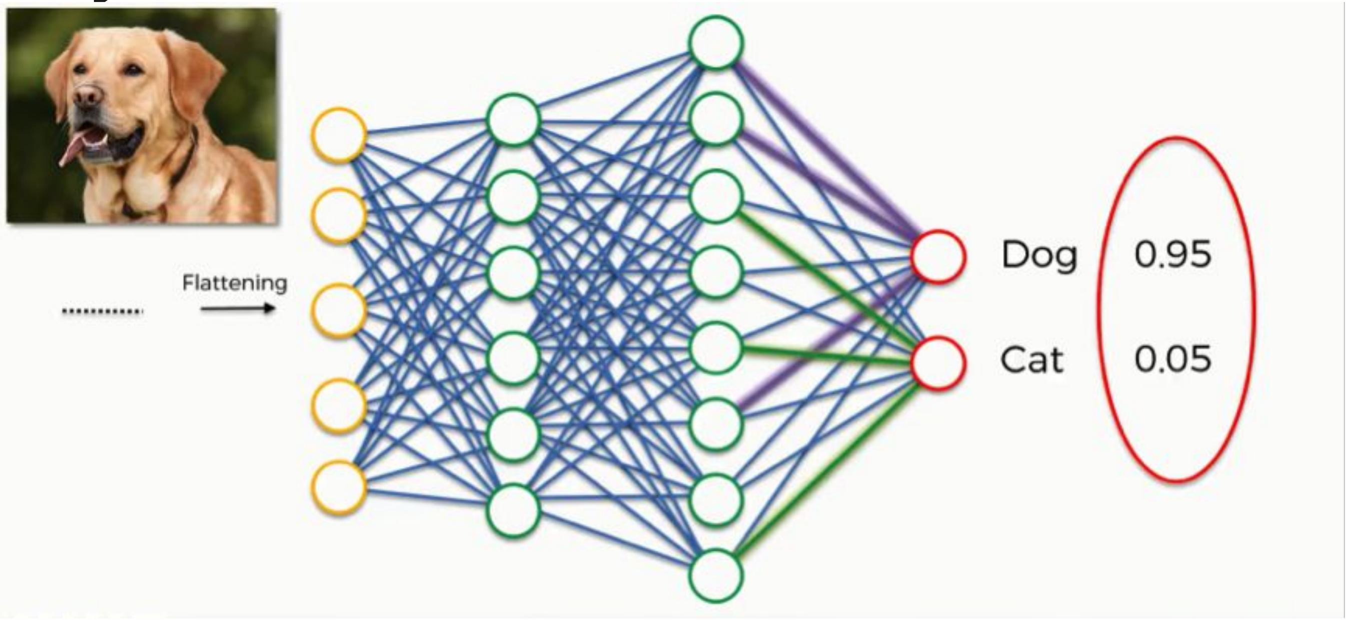
Fig. I



Modeling uncertainty



Disease spreading



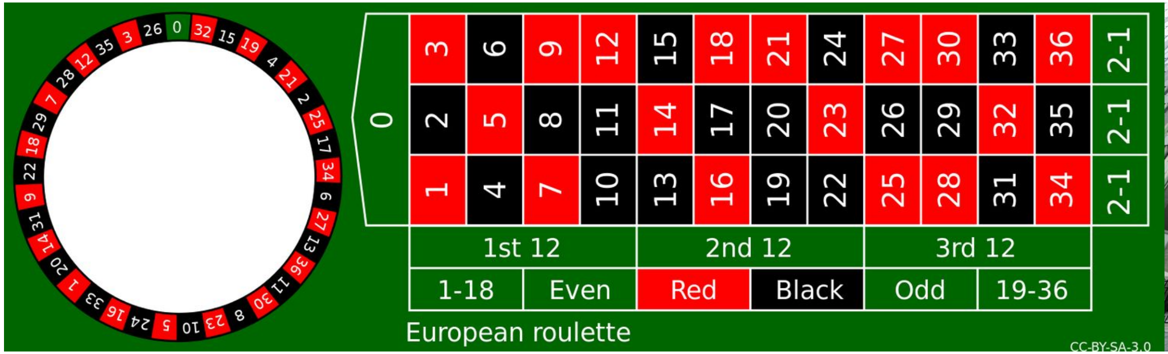
Source: Blog

covariance?

x, Y

$\vec{x} = (x, x)$

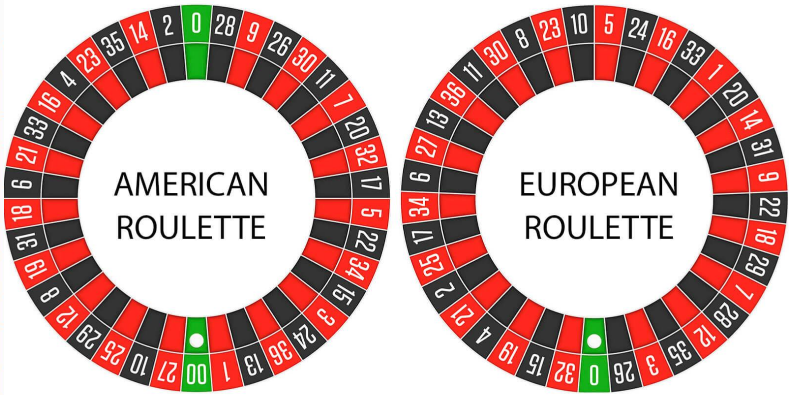
Modeling uncertainty



- Information theory, modeling the reliability of numerous complex systems, insurance companies, investments etc.
- **Today's agenda**: reminders of prereq probability material through problem solving.

Roulette

Odds & Payouts at European & American Roulette			
Roulette Bet	**Payout**	**European Roulette Odds**	**American Roulette Odds**
Single Number	35 to 1	2.70%	2.60%
2 Number Combination	17 to 1	5.4%	5.3%
3 Number Combination	11 to 1	8.1%	7.9%
4 Number Combination	8 to 1	10.8%	10.5%
5 Number Combination	6 to 1	13.5%	13.2%
6 Number Combination	5 to 1	16.2%	15.8%
Column	2 to 1	32.40%	31.6%
Dozen	2 to 1	32.40%	31.6%
Even/Odd	1 to 1	48.60%	47.4%
Red/Black	1 to 1	48.60%	47.4%
Low/High	1 to 1	48.60%	47.4%



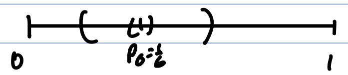
- Even American: {2,4,6,8,...,34,36}, hence Pr(even)=18/38=0.47368
- Even European: same favorable outcomes {2,4,6,8,...,34,36}, but Pr(even)=18/37=0.4864

전체개수가 달라서

확률이 다르다.

$$\hat{p}_6 = \frac{\#\text{ times I observe } 6}{n}$$

$$\hat{p}_6 \xrightarrow{n \to \infty} \frac{1}{6}$$



$$\Pr(|\hat{p}_6 - p_6| \geq 0.01) \leq 0.05$$

when $n \to \infty$ is 0 since $\hat{p}_6 = p_6$

Probability space

Sample space → Event space. Probability

$$(\Omega, \mathcal{F}, \mathbb{P})$$

Probability assign every single event to numbers

Sample space

event space

probability measure

$$\mathbb{P} : \mathcal{F} \to [0, 1]$$

Questions: what is a random variable? What is the difference between continuous and discrete random variables?

Random variable : Toss coin twice

Ω { $\overset{w_1}{HH}$, $\overset{w_2}{TT}$, $\overset{w_3}{HT}$, $\overset{w_4}{TH}$ }

Discrete $X : \Omega \to D$ (Domain D) D $X(w)$: # heads in w

$X(w_1) = 2$ $X(w_2) = 1$ $X(w_3) = 1$ $X(w_4) = 0$

$P(\Omega) = 1$

countable

Central limit theorem

Bayes

A B C D

$\frac{1}{3}$ $\frac{1}{3}$ $\frac{1}{3}$

Monty-hall problem

Conditional

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. A, and the host, who knows what's behind the doors, opens another door, say No. C, which has a goat.
- He then says to you, "Do you want to pick door No. B?" Is it to your advantage to switch your choice?

$$\frac{P(A\cap D)}{P(D)} = \frac{P(D\cap A)}{PA} \times \frac{PA}{P(D)}$$

$\frac{2}{6}$ $\frac{1}{6}$ $+\frac{1}{3} =$

$\frac{1}{2}\times\frac{1}{3} + 1\times\frac{1}{3} + 0$

open C

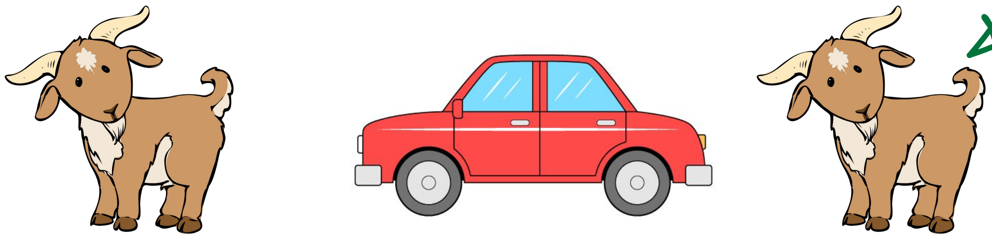
A : car $\quad P(D|A) = \frac{P(D\cap A)}{P(A)} = \frac{1}{2}$

B : car $\quad P(D|B) = 1$ C를 열수밖에

C : car $\quad P(D|C) = 0$

$$P(A|D) = \frac{P(D|A)P(A)}{P(D|A)P(A) + P(D|B)P(B) + P(D|C)P(C)}$$

$$= \frac{P(D|A)P(A)}{P(D)}$$



P(A|D)

$$= \frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}\times\frac{1}{3} + 1\times\frac{1}{3} + 0\times\frac{1}{3}} = \frac{\frac{1}{6}}{\frac{3}{6}} = \frac{1}{3}$$

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

$$= \frac{1\times\frac{1}{3}}{\frac{3}{6}} = \frac{1}{3}\times\frac{6}{3} = \frac{2}{3}$$

Assumptions

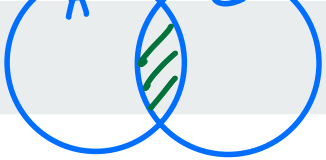
Let's make the problem concrete by specifying certain assumptions.

- Let's say the car is placed uniformly at random (**uar**) behind a door.
- Our initial guess is also **uar** uniformly at random $\frac{1}{3}$ $\frac{1}{3}$
- The host opens a door with a goat. When there exist two such doors, i.e., our guess is the car, he chooses **uar**.

$$P(A|C) = \frac{P(C|A)\,P(A)}{P(C)} = \frac{P(C|A)\,P(A)}{P(C|A)\,P(A) + P(C|A^c)\,P(A^c)}$$

$P(C) =$

$P(A \cap C) + P(C \cap A^c)$

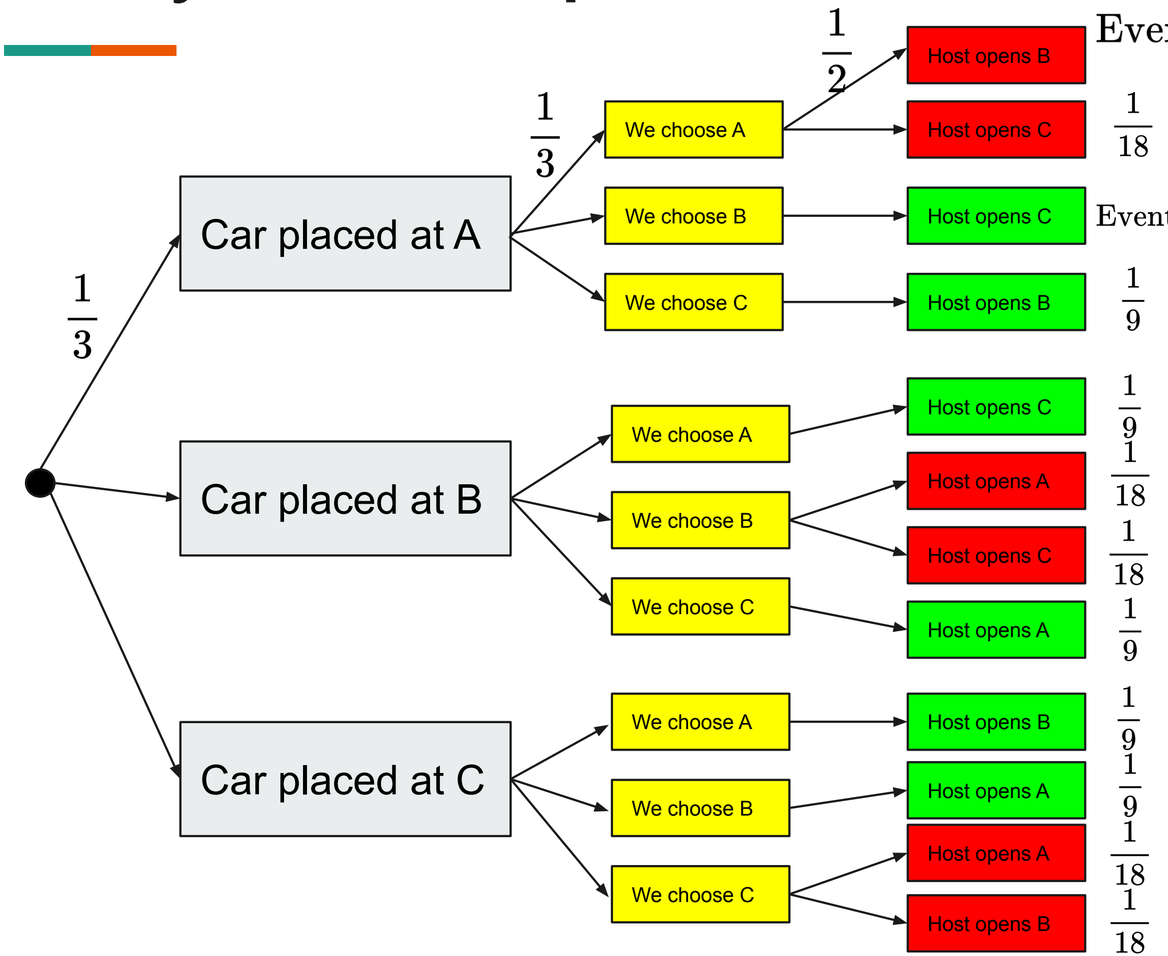


$\frac{P(C \cap A)}{P(A)} \times P(A) + \frac{P(C \cap A^c)}{P(A^c)} \times P(A^c)$

CS131 Reminder: Four step-method

1. Find the sample space
2. Define events of interest
3. Determine outcome probabilities
4. Compute event probabilities

Monty hall - 4 steps in one slide

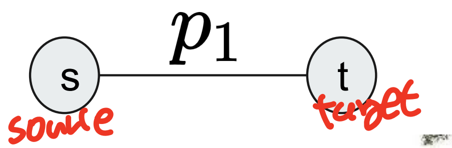


$$\text{Event (AAB)} : \frac{1}{3} \times \frac{1}{3} \times \frac{1}{2} = \frac{1}{18}$$

$$\text{Event (ABC)}: \frac{1}{3} \times \frac{1}{3} \times 1 = \frac{1}{9}$$

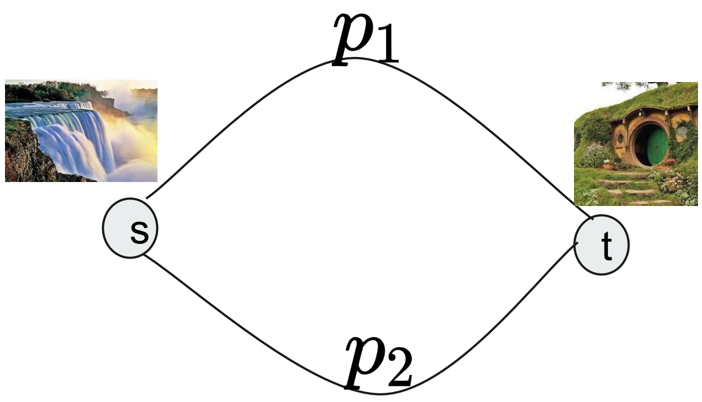
Therefore, switching yields
Probability of winning 6/9=⅔
Vs not switching 6/18=⅓

Transfer water



- Consider a water source **s** and a destination village **t**.
- Each pipe *i* has probability of failure p_i. Pipes fail ***independently***.
- **Question**: What is the probability we cannot get water from s to t? In other words:
    - when is the village **t** not reachable from the water source **s**?

Exercise 1



- Clearly, there is no path if both pipes fail

- Since they are independent, the probability of this event is the product of the probabilities of the individual events

Thus, failure probability is p_1p_2

Reminders: Independent events, conditional probability

Intuitive two events A,B are dependent if A's occurrence or non-occurrence provides us with some information about event B.

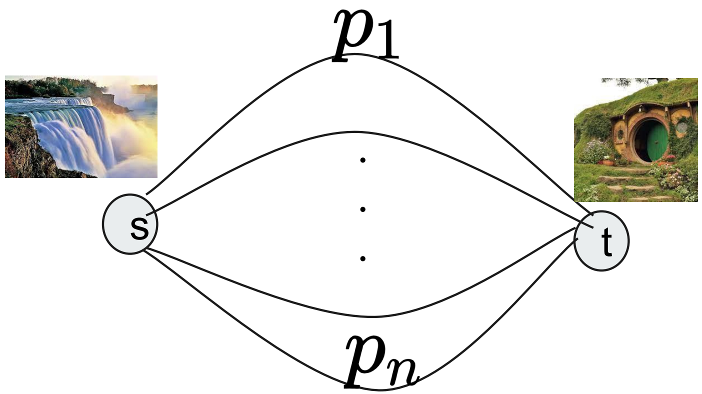
Formally, A,B are independent if and only iff $\boxed{\Pr(A \cap B) = \Pr(A)\Pr(B)}$

By rearranging we get $\boxed{\Pr(A) = \frac{\Pr(A \cap B)}{\Pr(B)}}$

Recall that by the law of conditional probability $\boxed{\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}}$

Therefore, when A,B are independent Pr(A)=Pr(A|B) and of course Pr(B)=Pr(B|A).

Exercise 2

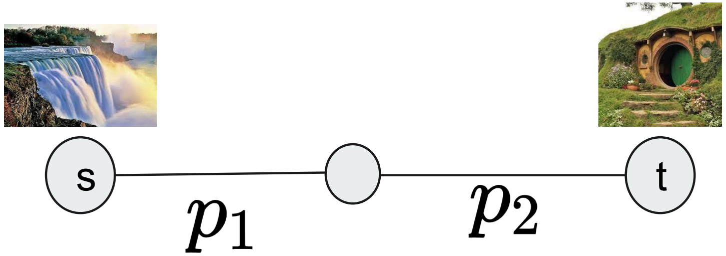


- Clearly, there is no path if **all** pipes fail

- Since they are independent, the probability of this event is the product of the probabilities of the individual events

Thus, failure probability is $p_1 p_2 \ldots p_n$

Exercise 3



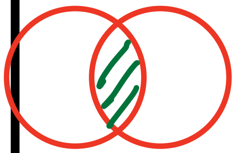
- Clearly, there is no path if **at least one of the pipes fail**

- Let A_i be the event that pipe i fails.

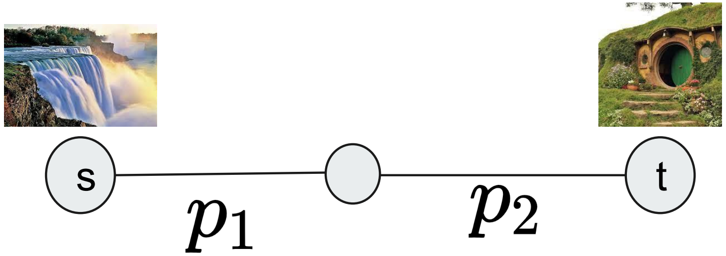
- Then,

$$\begin{aligned}\Pr(A_1 \cup A_2) &= \Pr(A_1) + \Pr(A_2) - \Pr(A_1 \cap A_2) \\ &= p_1 + p_2 - \Pr(A_1)\Pr(A_2) \\ &= p_1 + p_2 - p_1 p_2\end{aligned}$$

independent.



Exercise 3



- **Using conditional probability**

- We condition on whether the one of the two pipes (say the first) is broken or not.

- Let A_i be the event that pipe i fails.

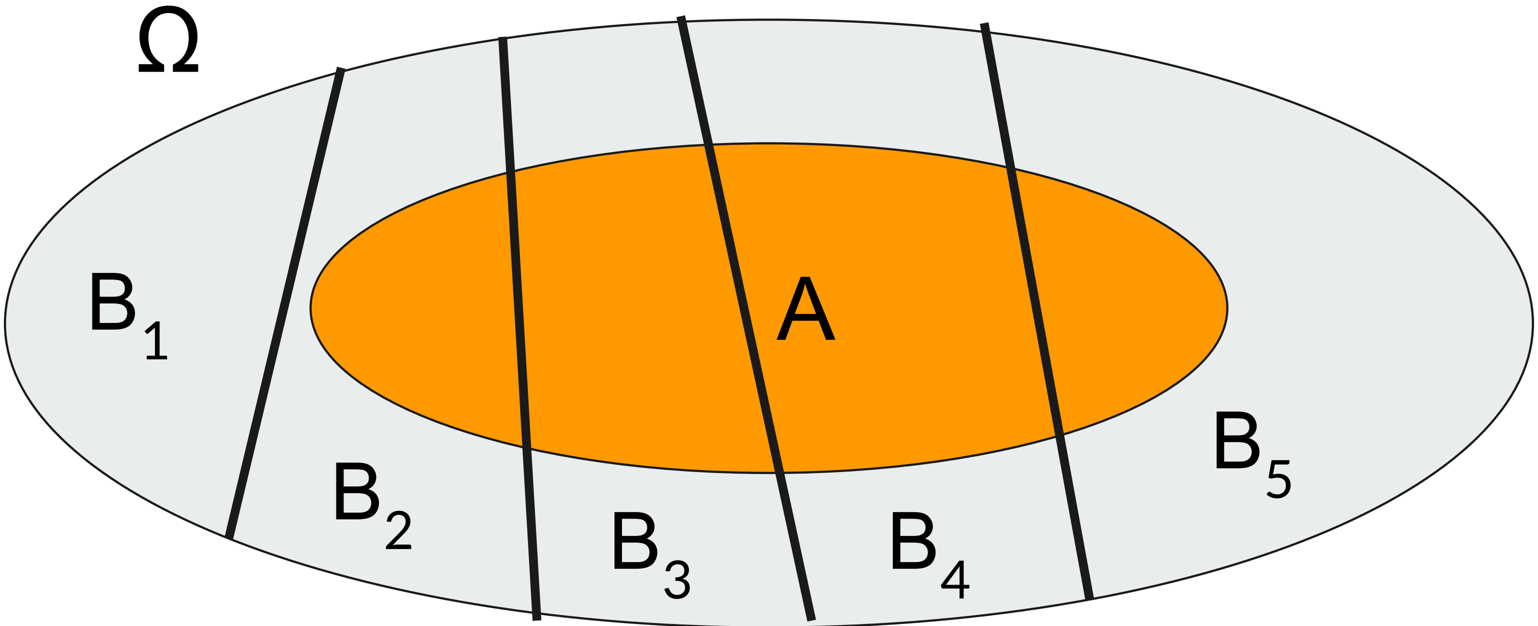
$$\begin{aligned}\Pr(A_1 \cup A_2) &= \Pr(A_1) + (1 - \Pr(A_1)) \Pr(A_2) \\ &= p_1 + (1 - p_1) p_2 \\ &= p_1 + p_2 - p_1 p_2\end{aligned}$$

success and fail

We used the law of total probability

Let Ω be a probability space. Let $B_1, \ldots, B_m$ be a partition of Ω. Then,

$$\boxed{\Pr(A) = \sum_{i=1}^{m} \Pr(A \cap B_i) = \sum_{i=1}^{m} \Pr(B_i) \Pr(A|B_i)}$$



Exercise 4

- Instead of thinking the probability that t will not be reachable from s, we think of the probability that it is. Reminder: $\Pr\left(\bar{A}\right) = 1 - \Pr(A)$
  - Let $\bar{A}_i$ be the event that pipe i does not fail.
- The probability of not failing is $\Pr\left(\cap_{i=1}^{n} \bar{A}_i\right) = \prod_{i=1}^{n} \Pr\left(\bar{A}_i\right) = \prod_{i=1}^{n} (1 - p_i)$

Therefore, the right answer is $1 - \prod_{i=1}^{n} (1 - p_i)$.

Reminder: chain rule

Chain rule:

$\frac{\Pr(A_2 \cap A_1)}{\Pr(A_1)}$ "

$\frac{\Pr(A_3 \cap A_2 \cap A_1)}{\Pr(A_2 \cap A_1)}$ "

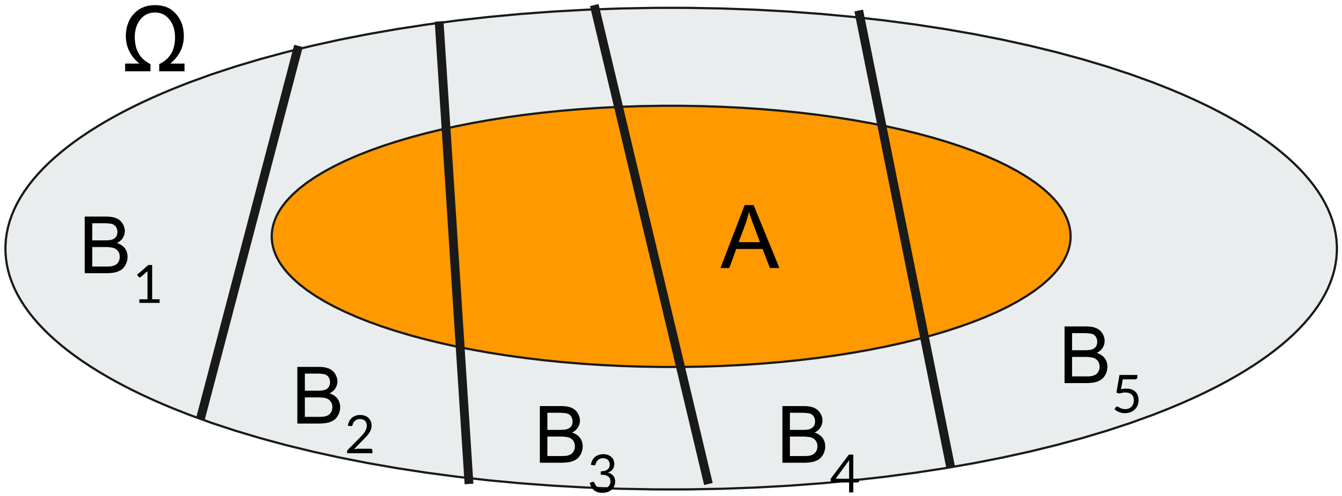
$$\Pr(A_1 \cap \ldots \cap A_n) = \Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_2A_1)\ldots\Pr(A_n|A_{n-1}..A_1)$$

In our case the events are mutually independent, so this simplifies to the product of the individual probabilities of the events A_i.

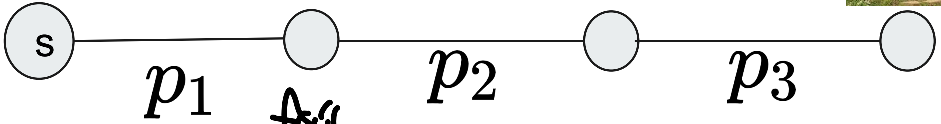
Question: what is the difference between pairwise and mutually independent events?

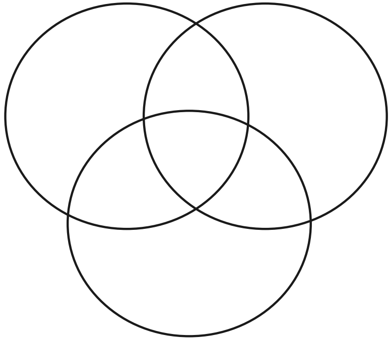
Conditional probability + Law of total probability → Bayes rule

$$\boxed{\Pr(B_i|A) = \frac{\Pr(B_i \cap A)}{\Pr(A)} = \frac{\Pr(B_i)\Pr(A|B_i)}{\sum_{j=1}^{n}\Pr(B_j)\Pr(B_j|A)}}$$



Exercise n=3





fail

$$1-(1-p_1)(1-p_2)(1-p_3) = 1-(1-p_1)(1-p_2-p_3+p_2p_3)$$

success success success

$$= 1-(1-p_2-p_3+p_2p_3-p_1+p_1p_2+p_1p_3-p_1p_2p_3)$$

$$= p_1+p_2+p_3-p_1p_2-p_1p_3-p_2p_3+p_1p_2p_3$$

Does this remind of something from CS131?

Exercise 4

- We condition on whether the one of the two pipes (say the first) is broken or not.

- Let A_i be the event that pipe i fails.

- We are interested in $\Pr(A_1 \cup \ldots \cup A_n)$.

Inclusion-exclusion

각집합의 원소의 수를 이용 합집합의 원소의수 구할때
부호전환 홀수면 더하고
짝수면 뺀다.

$$\Pr(\cup_{i=1}^{n} A_i) = \sum_{\mathcal{J}\subseteq\{1,\ldots,n\};|\mathcal{J}|=k} (-1)^{k+1} P(\bigcap_{i\in\mathcal{J}} A_i)$$

$n=3$

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{I\subseteq U} (-1)^{|I|+1} \left|\bigcap_{i\in I} A_i\right|$$

Proof sketch (inductive proof)

When n=1 the statement is obvious. Use the IH and the fact that

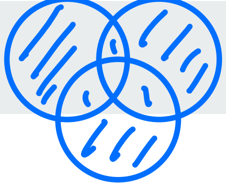
$$P\left(\bigcup_{i=1}^{n+1} A_i\right) = P\left(\bigcup_{i=1}^{n} A_i\right) + P\left(A_{n+1} \setminus \bigcup_{i=1}^{n} A_i\right)$$

$$= P\left(\bigcup_{i=1}^{n} A_i\right) + P(A_{n+1}) - P\left(\bigcup_{i=1}^{n}(A_i \cap A_{n+1})\right).$$

$|A\cup B| = |A| + |B| - |A\cap C|$

$|A\cup B\cup C| = |A| + |B| + |C| - |A\cap B|$

$- |B \cap C| - |C \cap A| + |A \cap B \cap C|$



/// 한번
/// 두번
/// 세번

Inclusion exclusion

Another convenient way to write the IE formula is the following

$$\mathbf{P}\left(\bigcup_{i=1}^{n} A_i\right) = S_1 - S_2 + S_3 - \ldots + (-1)^{n-1} S_n$$

where
$$S_k = \sum_{1 \le i_1 < i_2 < \ldots < i_k \le n} \mathbf{P}(A_{i_1} \cap A_{i_2} \cap \ldots \cap A_{i_k}).$$

In our setting, due to the independence of the events A_i we can write the following expression

$$\Pr(\cup A_i) = \sum_{k=1}^{n} (-1)^{k+1} \sum_{I \subseteq [n], |I|=k} \prod_{i \in I} \Pr(A_i)$$

let's write down some terms

$$\begin{aligned} \Pr(\cup A_i) &= p_1 + \ldots + p_n \\ &\quad -(p_1 p_2 + \ldots + p_{n-1} p_n) \\ &\quad +(p_1 p_2 p_3 + \ldots + p_{n-2} p_{n-1} p_n) \\ &\quad - \ldots \end{aligned}$$

Union bound

Let $A_1,\ldots,A_n$ be events in a probability space. Then, we get the following upper bound on the probability of their union.

$$\boxed{\Pr(A_1 \cup \ldots \cup A_n) \leq \sum_{i=1}^{n} \Pr(A_i)}$$

다 더한 것보다 작습니다