Lecture 2: Conditional Probability

Wen Dingzhu

School of Information Science and Technology (SIST)
ShanghaiTech University

wendzh@shanghaitech.edu.cn

February 13, 2023

上海科技大学
ShanghaiTech University

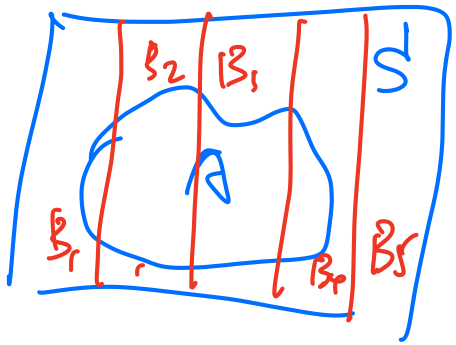
Overview

Outline

Thinking Conditionally

- New data & information may affect our uncertainties.
- Conditional probability: how to update our belief?
- All probabilities are conditional! (explicit/implicit background info or assumption)

Thinking Conditionally

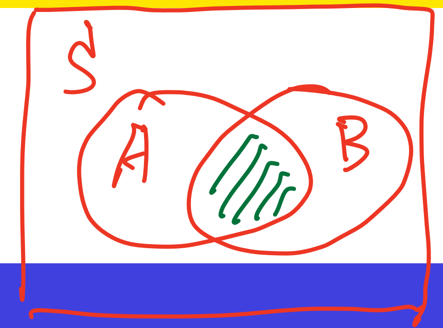


- Conditioning: a powerful problem-solving strategy.
- Reducing a complicated probability problem to a bunch of simpler conditional probability problems.
- First-step analysis: obtain recursive solution to multi-stage problems.
- **Conditioning is the soul of statistics!**

Definition of Conditional Probability

$P(A)$

B



Definition

If A and B are events with $P(B) > 0$, then the conditional probability of A given B, denoted by $P(A|B)$, is defined as

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

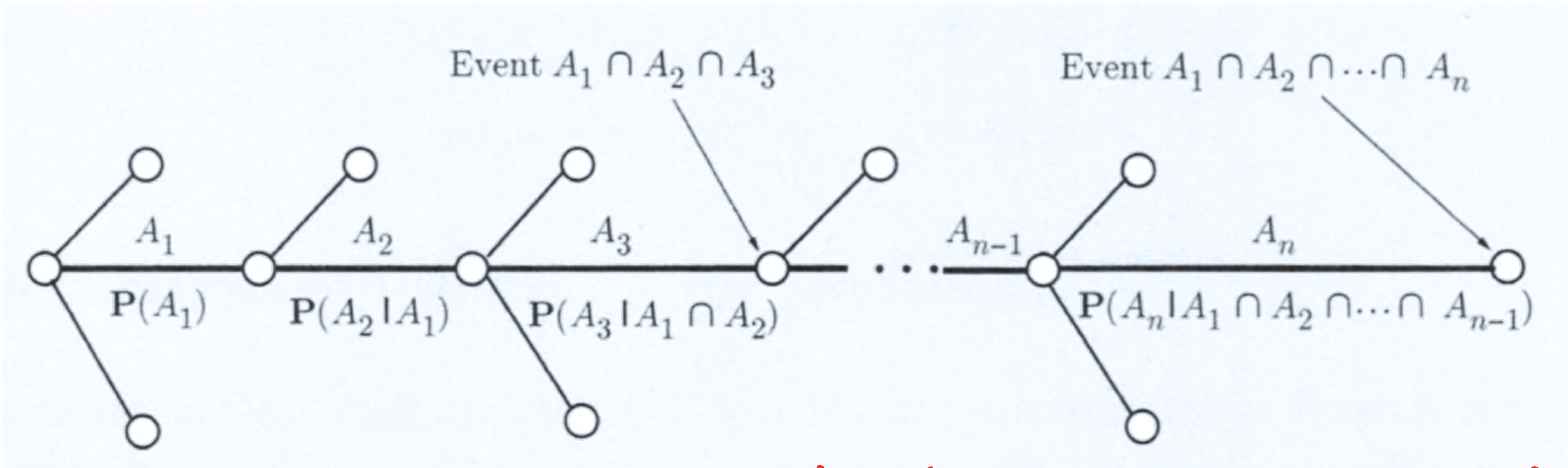
- $P(A)$: prior probability of A.
- $P(A|B)$: posterior probability of A.

Outline

Chain Rule

$P(A \cap B) = P(A, B)$

① $n=2$, $P(A_1, A_2) = P(A_1) \cdot P(A_2 | A_1)$



② $n=3$, $P(A_1, A_2, A_3) = P(A_1, A_2) P(A_3 | A_1, A_2) = P(A_1) P(A_2 | A_1) P(A_3 | A_1, A_2)$.

Theorem

For any events $A_1, \ldots, A_n$ with positive probabilities,

$$P(A_1, \ldots, A_n) = P(A_1) P(A_2 | A_1) P(A_3 | A_1, A_2) \cdots P(A_n | A_1, \ldots, A_{n-1}).$$

$\log P(A_1, \cdots, A_n) = \log P(A_1) + \log P(A_2 | A_1) + \cdots + \log P(A_n | A_1, \cdots, A_{n-1})$

Bayes' Rule

$$P(A|B) = \frac{P(A\cap B)}{P(B)} \Rightarrow P(A\cap B) = \underline{P(A|B)P(B)}$$

$$\|$$

$$P(B|A) = \frac{P(B\cap A)}{P(A)} \Rightarrow P(B\cap A) = \underline{P(B|A)P(A)}$$

Theorem

For any events A and B with positive probabilities,

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

A = "guilty"

B = "evidence"

$P(A|B)$

$P(B|A) \to 1$

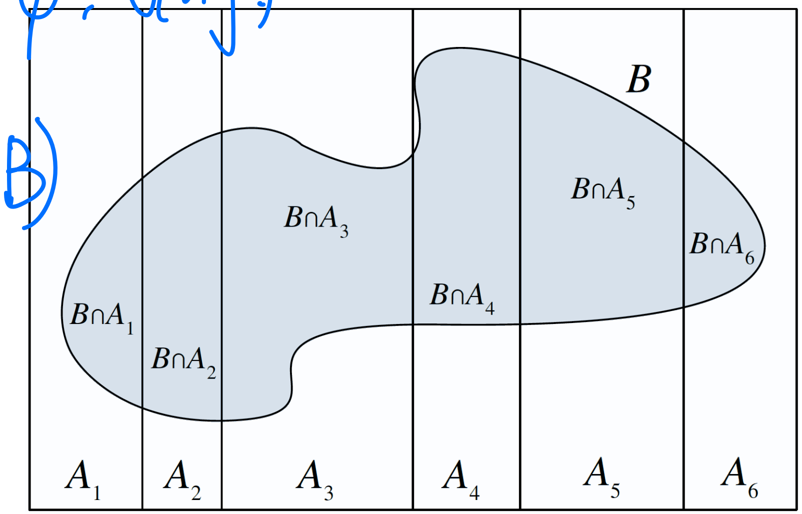
$P(B|A^c) \to 1$

The Law of Total Probability (LOTP)

$(A_i, B) \cap (A_j, B) = \phi, \ \forall (i,j)$

$P[(A_i \cap B) \cup (A_j \cap B)]$

$= P(A_i, B) + P(A_j, B)$



$B = \bigcup_{i=1}^{n} (A_i \cap B).$

$P(B) = \sum_{i=1}^{n} P(A_i \cap B).$

$= \sum_{i=1}^{n} P(A_i, B)$

Theorem

Let $A_1, \ldots, A_n$ be a partition of the sample space S (i.e., the A_i are disjoint events and their union is S), with $P(A_i) > 0$ for all i. Then

$$P(B) = \sum_{i=1}^{n} \overbrace{P(B|A_i)P(A_i)}^{P(A_i, B)}.$$

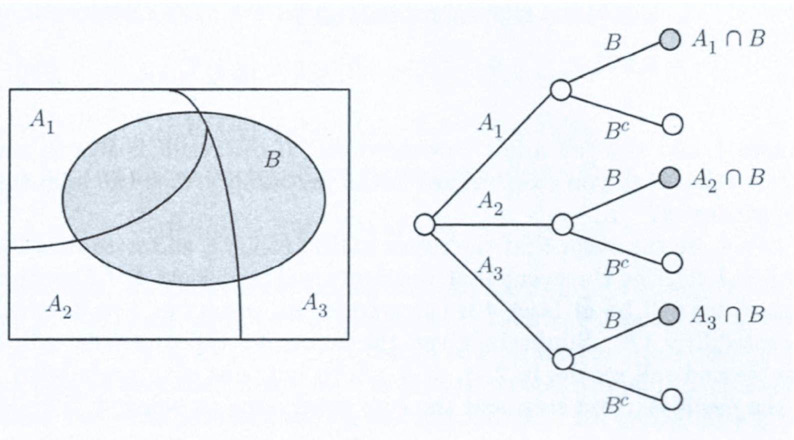
Inference & Bayes' Rule

A_1 : "coin is fair"

A_2 : "coin is unfair".

B : data/info

$P(A_1|B)$ $P(A_2|B)$



Theorem

Let $A_1, \ldots, A_n$ be a partition of the sample space S (i.e., the A_i are disjoint events and their union is S), with $P(A_i) > 0$ for all i. Then for any event B such that $P(B) > 0$, we have

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + \cdots + P(A_n)P(B|A_n)}.$$

Example: Bayes Spam Filter

$A =$ "an email is spam". $P(A) = 0.8$

$B =$ "an email mentions 'free money'" $P(B|A) = 0.1$

A spam filter is designed by looking at commonly occurring phrases in spam. Suppose that 80% of email is spam. In 10% of the spam emails, the phrase "free money" is used, whereas this phrase is only used in 1% of non-spam emails. A new email has just arrived, which does mention "free money". What is the probability that it is spam?

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

$$= \frac{0.1 \times 0.8}{0.1 \times 0.8 + 0.01 \times 0.2} = \frac{80}{82} \approx 0.9756$$

$P(A|B) > P(A^c|B)$.

Example: Bayes Spam Filter

B_1 = "an email mentions 'free money'"

B_2 = " . . . . 'fun' "

$$\log\left[\frac{P(A|B_1,B_2)}{P(A^c|B_1,B_2)}\right] > 0 \text{ spam}, < 0$$

Bayes' rule is used several times in the context of spam:

- Compute the probability that the message is spam, knowing that a given word appears in this message.
- Compute the probability that the message is spam, taking into consideration all of its words (or a relevant subset of them).
- Sometimes a third time, to deal with rare words.
- Reference: `https://en.wikipedia.org/wiki/Naive_Bayes_spam_filtering`

$$P(A|B_1,B_2) = \frac{P(B_1,B_2|A)P(A)}{P(B_1,B_2)} \qquad \frac{P(A|B_1,B_2)}{P(A^c|B_1,B_2)} = \frac{P(B_1,B_2|A)P(A)}{P(B_1,B_2|A^c)P(A^c)}$$

Example: Random Coin

A: "the choosen coin lands Heads 3 times"

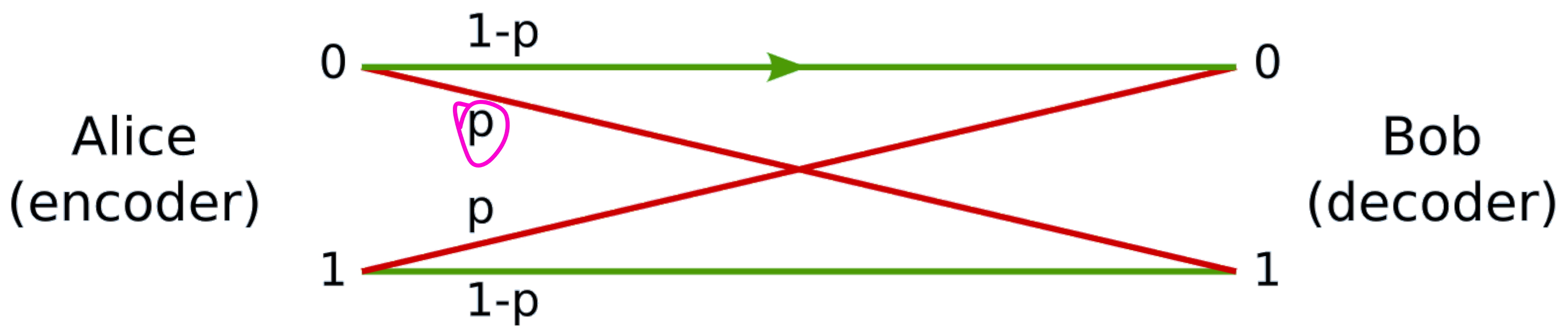
B: "the picked coin is fair".

$P(B|A)$.

You have one fair coin, and one biased coin which lands Heads with probability 3/4. You pick one of the coins at random and flip it three times. It lands Heads all three times. Given this information, what is the probability that the coin you picked is the fair one?

$$P(B|A) = \frac{P(A|B)P(B)}{\underbrace{P(A|B)}_{(\frac{1}{2})^3}\underbrace{P(B)}_{\frac{1}{2}} + \underbrace{P(A|B^c)}_{(\frac{3}{4})^3}\underbrace{P(B^c)}_{\frac{1}{2}}} = \frac{8}{35} < \frac{1}{2}$$

Example: Communication Channel



Example: Communication Channel

$A =$ "Alice sent a '1'" $P(B|A) = 0.95$

$B =$ "Bob received a '1'". $P(A|B)$

Suppose that Alice sends only one bit (a 0 or 1) to Bob, with equal probabilities. If she sends a 0, there is a 5% chance of an error occurring, resulting in Bob receiving a 1; if she sends a 1, there is a 5% chance of an error occurring, resulting in Bob receiving a 0. Given that Bob receives a 1, what is the probability that Alice actually sent a 1?

$P(A) = 0.5$

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)} = \frac{0.95 \times 0.5}{0.95 \times 0.5 + 0.05 \times 0.5} = 0.95$$

Outline

Conditional Probability

$\hat{P}(\cdot) = P(\cdot|E)$.

- Condition on an event E, we update our beliefs to be consistent with this knowledge.
- $P(\cdot|E)$ is also a probability function with sample space S:
  - $0 \leq P(\cdot|E) \leq 1$ with $P(S|E) = 1$ and $P(\emptyset|E) = 0$.
  - If events $A1, A2, \ldots$ are disjoint, then $P(\cup_{j=1}^{\infty} A_j|E) = \sum_{j=1}^{\infty} P(A_j|E)$.
  - $P(A^c|E) = 1 - P(A|E)$.
  - Inclusion-exclusion: $P(A \cup B|E) = A(A|E) + P(B|E) - P(A \cap B|E)$.

Bayes' Rule with Extra Conditioning

① $P(A|B,E) = \frac{P(A,B,E)}{P(B,E)} = \frac{P(B|A,E)\,P(A,E)}{P(B,E)} = \frac{P(B|A,E)\,P(A|E)\,\cancel{P(E)}}{P(B|E)\,\cancel{P(E)}}$

② $\hat{P}(\cdot) = P(\cdot|E)$.

Theorem

Provided that $P(A \cap E) > 0$ and $P(B \cap E) > 0$, we have

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

③ $\hat{P}(A|B) = \frac{\hat{P}(B|A)\,\hat{P}(A)}{\hat{P}(B)}$

$P(A|E)$

$P(B|E)$

~~$P(B|A|E)$~~

LOTP with Extra Conditioning

① LOTP : $P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$. $A_1, \cdots, A_n$ disjoint.

② $\hat{P}(\cdot) = P(\cdot|E)$.

Theorem

Let $A_1, \ldots, A_n$ be a partition of the sample space S (i.e., the A_i are disjoint events and their union is S), with $P(A_i \cap E) > 0$ for all i. Then

$$P(B|E) = \sum_{i=1}^{n} P(B|A_i, E)P(A_i|E).$$

③ $\hat{P}(B) = \sum_{i=1}^{n} \hat{P}(B|A_i)\hat{P}(A_i)$

Example: Random Coin

A: "the chosen coin lands Head 3 times".

F = " the chosen coin is fair " F^c

H = " the chosen coin lands head on the 4th toss".

You have one fair coin, and one biased coin which lands Heads with probability 3/4. You pick one of the coins at random and flip it three times. It lands Heads all three times. If we toss the coin a fourth time, what is the probability that it will land Heads once more?

$$P(H|A) = \underbrace{P(H|F,A)}_{} \cdot \underbrace{P(F|A)}_{} + P(H|F^c,A)\, P(F^c|A).$$

$$\overset{1/2}{P(H|F,A)} \quad \overset{8/35}{P(F|A)} \quad \overset{3/4}{P(H|F^c,A)} \quad \overset{27/35}{P(F^c|A)}$$

$$= \frac{97}{140} > 1/2$$

Example: Random Coin

Approaches for Finding $P(A|B, C)$ $= P(A|C, B)$.

ⓐ $P(A|B,C) = \dfrac{P(A,B,C)}{P(B,C)}$

- Think of B, C as the single event $B \cap C$.
- Use Bayes' rule with extra conditioning on C.
- Use Bayes' rule with extra conditioning on B.

Outline

Independence of Two Events

$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$.

Definition

Events A and B are independent if

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

If $P(A) > 0$ and $P(B) > 0$, then this is equivalent to

$$P(A|B) = P(A),$$
$$P(B|A) = P(B).$$

Independence vs. Disjointness

A, B independent $\quad P(A,B) = P(A)P(B)$.

A, B disjoint $\quad P(A,B) = 0$.

$A \cap B = \phi$.

Independence of Complementary Set

$P(\cdot \mid A)$.

$P(A)>0$, $P(B)>0$

$P(A|B) = P(A)$.

$P(B|A) = P(B)$

$P(B^c|A) = 1 - P(B|A)$.

$= 1 - P(B) = P(B^c)$

Theorem

If A and B are independent, then A and B^c are independent, A^c and B are independent, and A^c and B^c are independent.

Independence of Three Events

Definition

Events A, B and C are independent if all of the following equations hold:

$$\left.\begin{aligned} P(A \cap B) &= P(A)P(B), \\ P(A \cap C) &= P(A)P(C), \\ P(B \cap C) &= P(B)P(C), \end{aligned}\right\} \text{两两独立.}$$

$$\underline{P(A \cap B \cap C) = P(A)P(B)P(C).}$$

Pairwise Independence and Independence

① $P(A) = 1/2$, $P(B) = 1/2$.

② $P(C) = P[\underbrace{(A \cap B) \cup (A^c \cap B^c)}_{\text{disjoint}}] = P(A \cap B) + P(A^c \cap B^c)$.

$= 1/4 + 1/4$

$= 1/2$.

Consider two fair, independent coin tosses.

- A: the event that the first is Heads.
- B: the event that the second is Heads.
- C: the event that both tosses have the same result.

$P(A \cap B) = 1/4 = P(A) \cdot P(B)$.

$P(A \cap C) = 1/4 = P(A) \cdot P(C)$

$P(B \cap C) = 1/4 = P(B) \cdot P(C)$

$P(A \cap B \cap C) = 1/4$

$\overset{?}{\neq}$.

$P(A) P(B) P(C) = 1/8$

Conditional Independence

$\tilde{P}(\cdot) = P(\cdot \mid E)$

Definition

Events A and B are said to be conditionally independent given E if:

$$P(A \cap B|E) = P(A|E)P(B|E).$$

Example: Conditional Independence $\not\Rightarrow$ Independence

① $P(A_1 \cap A_2 | F) = P(A_1|F)\,P(A_2|F)$. $\quad P(A_1 \cap A_2 | F^c)$

② ? $P(A_1 \cap A_2) \neq P(A_1)P(A_2) = 5/8 \times 5/8 = 25/64$

- We choose either a fair coin or a biased coin (w.p. 3/4 of landing Heads).
- But we do not know which one we have chosen and we flip it twice.
- Event F: "chosen the fair coin".
- Event A_1: "the first coin tosses landing Heads".
- Event A_2: "the second coin tosses landing Heads".

$P(A_1) = P(A_1|F)P(F) + P(A_1|F^c)P(F^c) = 1/2 \times 1/2 + \frac{3}{4} \times 1/2 = 5/8.$

$P(A_1 \cap A_2) = P(A_1 \cap A_2 | F)P(F) + P(A_1 \cap A_2 | F^c)P(F^c)$

$= 1/4 \times 1/2 + \frac{3}{4} \cdot \frac{3}{4} \cdot \frac{1}{2} = 13/32$

Example: Independence $\not\Rightarrow$ Conditional Independence

A, B independent

Given R.

$\hat{P}(\cdot) = P(\cdot \mid R)$.

$P(B \mid A^c, R) = 1$.

$\hat{P}(B \mid A^c) = \hat{P}(B)$.

- Only Alice and Bob call me.
- Each day, they call me independently.
- Event R: "Phone rings".
- Event A: "Alice call me".
- Event B: "Bob call me".

Example: Conditional Independence Given E vs. E^c

$$P(W \cap A \mid E^c) = P(W \mid E^c)\,P(A \mid E^c)$$

$$P(W \cap A \mid E) \neq P(W \mid E)\,P(A \mid E).$$

- There are two classes: good & bad.
- Good: students get grade A with working hard.
- Bad: students get grades randomly regardless of their efforts.
- Event E: "Class is good".
- Event W: "Students working hard".
- Event A: "Students receive grade A".

Outline

$B^c \triangleq \overline{B}$

$A_1, A_2, \cdots, A_n$ independent

$$P\left(\bigcup_{i=1}^{n} A_i\right) = 1 - P\left(\overline{\bigcup_{i=1}^{n} A_i}\right)$$

$$= 1 - P(\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n})$$

$$= 1 - P(A_1^c) \times P(A_2^c) \times \cdots \times P(A_n^c)$$

$$= 1 - [1-P(A_1)] \times [1-P(A_2)] \times \cdots \times [1-P(A_n)]$$

Example: Laplace's Rule of Succession

$P(H|F_n)$

C_i = "the i-th coin is selected initially", $i = 0, 1, 2, \cdots, k$.

F_n = "the first n flips all land head"

H = "the (n+1)-th flip lands head"

There are $k+1$ coins in a box. When flipped, the i^{th} coin will turn up heads with probability $i/k, i = 0, 1, \ldots, k$. A coin is randomly selected from the box and is then repeatedly flipped. If the first n flips all result in heads, what is the conditional probability that the $(n+1)^{st}$ flip will do likewise?

$$P(H|F_n) = \sum_{i=0}^{k} \underbrace{P(H|C_i, F_n)}_{\| \; P(H|C_i) = \frac{i}{k}} P(C_i|F_n) \qquad \frac{(i/k)^n}{\sum_{j=0}^{k}(j/k)^n}$$

$$P(C_i|F_n) = \frac{P(F_n|C_i)P(C_i)}{P(F_n)} = \frac{P(F_n|C_i)P(C_i)}{\sum_{j=0}^{k} P(F_n|C_j)P(C_j)}$$

Solution of Laplace's Rule of Succession

$$P(C_i|F_n) = \frac{P(F_n|C_i)P(C_i)}{\sum_{j=0}^{k} P(F_n|C_j)P(C_j)} = \frac{\left(\frac{i}{k}\right)^n \cdot \cancel{\frac{1}{k+1}}}{\sum_{j=0}^{k}\left(\frac{j}{k}\right)^n \cdot \cancel{\frac{1}{k+1}}} = \frac{(i/k)^n}{\sum_{j=0}^{k}(j/k)^n}$$

$$P(H|F_n) = \frac{\sum_{i=0}^{k}\left(\frac{i}{k}\right)^{n+1}}{\sum_{j=0}^{k}\left(\frac{j}{k}\right)^n}$$

$k \gg 1$

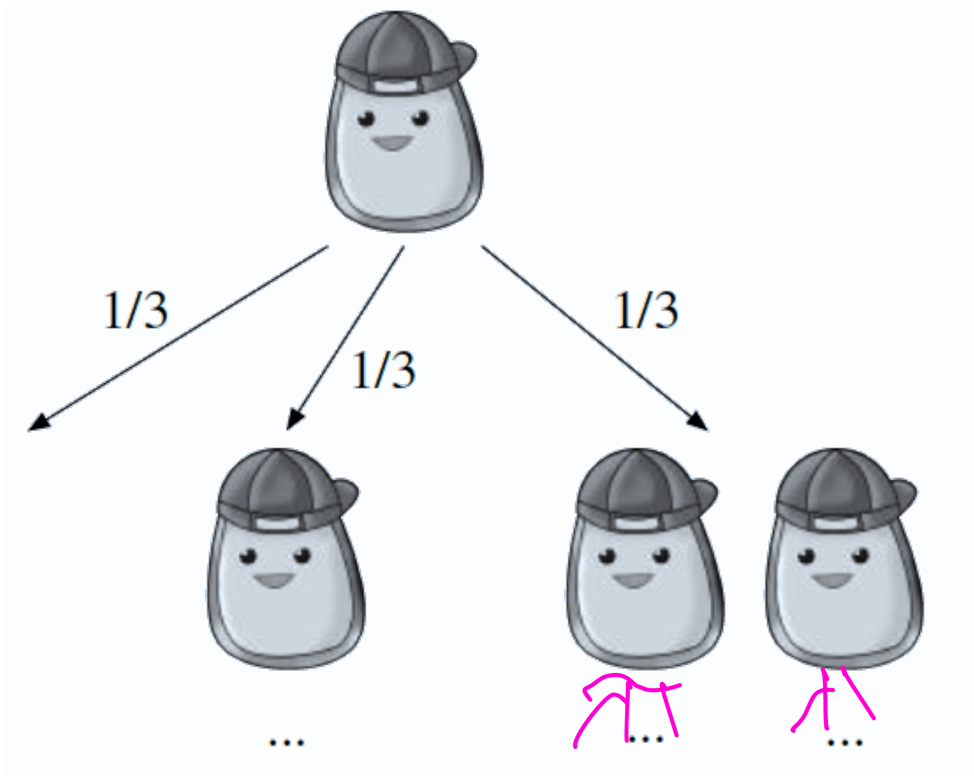
$k \to \infty$,

$$\sum_{i=0}^{k}\frac{1}{k}\left(\frac{i}{k}\right)^{n+1} \simeq \int_0^1 x^{n+1}\,dx = \frac{1}{n+2}; \quad P(H|F_n) = \frac{n+1}{n+2}$$

$$\frac{1}{k}\cdot\sum_{j=0}^{k}\left(\frac{j}{k}\right)^n \simeq \int_0^1 x^n\,dx = \frac{1}{n+1};$$

Example: Branching Process

A single amoeba, Bobo, lives in a pond. After one minute Bobo will either die, split into two amoebas, or stay the same, with equal probability, and in subsequent minutes all living amoebas will behave the same way, independently. What is the probability that the amoeba population will eventually die out?



First-Step Analysis: Branching Process

① D = "the population eventually dies out".

② B_i = "Bobo turns into i amoebas after the first minute",

$i = 0, 1, 2.$

$P(B_i) = 1/3.$

③ $P(D|B_0) = 1.$ $P(D|B_1) = P(D).$

$P(D|B_2) = P^2(D)$

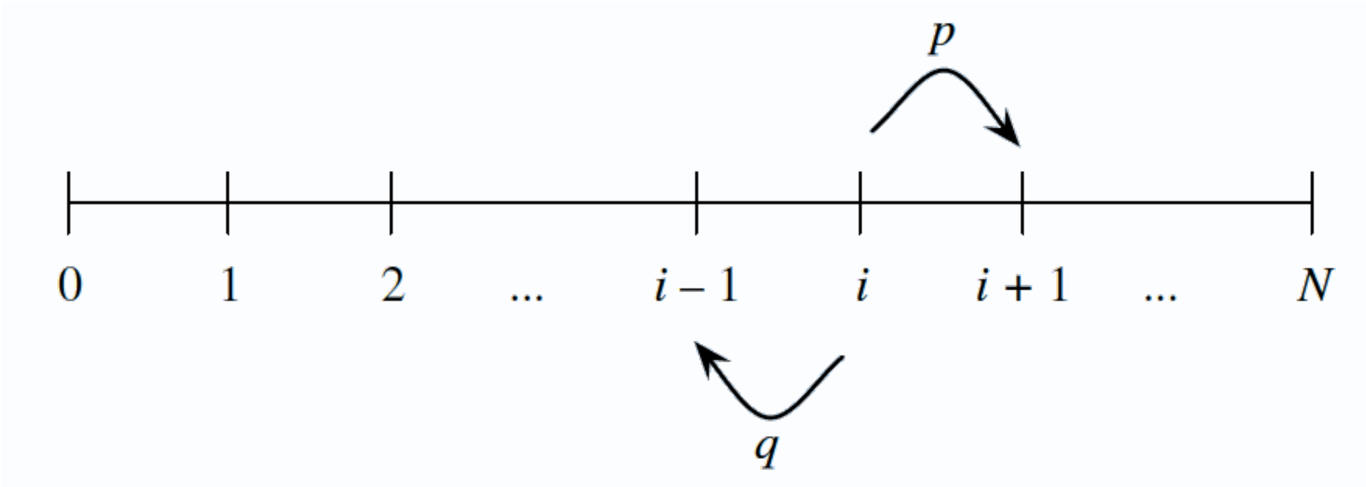
④ By LOTP, $P(D) = P(D|B_0)P(B_0) + P(D|B_1)P(B_1) + P(D|B_2)P(B_2)$

$\underline{P(D) = 1 \times \frac{1}{3} + P(D) \times \frac{1}{3} + P^2(D) \times \frac{1}{3}}$

$\Rightarrow P(D) = 1$

Example: Gambler's Ruin

Two gamblers, A and B, make a sequence of dollar 1 bets. In each bet, gambler A has probability p of winning, and gambler B has probability $q = 1 - p$ of winning. Gambler A starts with i dollars and gambler B starts with $N - i$ dollars; the total wealth between the two remains constant since every time A loses a dollar, the dollar goes to B, and vice versa. The game ends when either A or B is ruined (run out of money). What is the probability that A wins the game (walking away with all the money)?



First-Step Analysis: Gambler's Ruin

A_i: "A starts with i dollars", $i=0,1,\cdots,N$. $p_i = P(A \text{ win} \mid A_i)$.

① $p_0 = P(A \text{ win} \mid A_0) = 0$; $p_N = P(A \text{ win} \mid A_0) = 1$.

② $1 \le i \le N-1$; W_1 = "A wins the first bet". $P(W_1) = p$.

By LOTP, $p_i = \underline{P(A \text{ wins} \mid A_i)} = P(A \text{ wins} \mid W_1, A_i)\, \underbrace{P(W_1 \mid A_i)}_{p}$

$+ P(A \text{ wins} \mid W_1^c, A_i)\, \underbrace{P(W_1^c \mid A_i)}_{q}$

$P(A \text{ wins} \mid W_1, A_i) = p_{i+1}$ $\quad P(A \text{ wins} \mid W_1^c, A_i) = p_{i-1}$

$$\Rightarrow \begin{cases} p_i = p_{i+1} \cdot p + p_{i-1} \cdot q, & 1 \le i \le N-1 \\ p_0 = 0, \quad p_N = 1, & p+q=1. \end{cases}$$

Difference Equation

First-Step Analysis: Gambler's Ruin

$$\begin{cases} P_i = P_{i+1}\cdot p + P_{i-1}\cdot q. \\ p+q=1 \end{cases}$$

$$P_i\cdot(p+q) = P_{i+1}\cdot p + P_{i-1}\cdot q$$

$$\underbrace{(P_i - P_{i+1})}_{\alpha_i}\cdot p = \underbrace{(P_{i-1} - P_i)}_{\alpha_{i-1}}\, q.$$

$$\alpha_i = \frac{q}{p}\cdot\alpha_{i-1},$$

$$P_i = \begin{cases} \dfrac{1-\left(\frac{q}{p}\right)^i}{1-\left(\frac{q}{p}\right)^N}, & p\neq q \\ \dfrac{i}{N}, & p=q \end{cases}$$

Method of Characteristic Equation

Outline

Simpson's Paradox

- A phenomenon in probability and statistics in which a trend appears in multiple groups of data but disappears or reverses when the groups are combined.
- Mentioned earlier by Karl Pearson in 1899 & Udny Yule in 1903.
- Officially proposed by Edward H. Simpson (1922 – 2019) in 1951.
- Delicate connection between probabilistic reasoning and causal inference.
- Reference: `https://plato.stanford.edu/entries/paradox-simpson/`

Example of Simpson's Paradox

A: "a successful surgery". B = "Dr. Nick is the surgeon".

C = "the surgery is a heart surgery".

	Heart	Band-Aid
Success	70	10
Failure	20	0

80%

Dr. Hibbert

	Heart	Band-Aid
Success	2	81
Failure	8	9

83%

Dr. Nick

$P(A|B, C) = 1/5 < P(A|B^c, C) = \frac{7}{9}$;

$P(A|B, C^c) = \frac{9}{10} < P(A|B^c, C^c) = 1$;

$P(A|B) = 0.83 > P(A|B^c) = 0.8$

Math Behind Simpson's Paradox

By LOTP, $P(A|B) = \underline{P(A|B,C)}\,P(C|B) + \underline{P(A|B,C^c)}\,P(C^c|B)$.

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

If

$$P(A|B,C) < P(A|B^c,C),$$
$$P(A|B,C^c) < P(A|B^c,C^c),$$

$\underset{0.9}{\underline{P(C^c|B)}} \gg \underset{0.1}{\underline{P(C^c|B^c)}}$

then is it possible that

$$\underline{P(A|B) > P(A|B^c)}?$$

Another Example of Simpson's Paradox

Gender discrimination in college admissions: In the 1970s, men were significantly more likely than women to be admitted for graduate study at the University of California, Berkeley, leading to charges of gender discrimination. Yet within most individual departments, women were admitted at a higher rate than men. It was found that women tended to apply to the departments with more competitive admissions, while men tended to apply to less competitive departments.

Monty Hall Problem

On the game show Let's Make a Deal, hosted by Monty Hall, a contestant chooses one of three closed doors, two of which have a goat behind them and one of which has a car. Monty, who knows where the car is, then opens one of the two remaining doors. The door he opens always has a goat behind it (he never reveals the car!). If he has a choice, then he picks a door at random with equal probabilities. Monty then offers the contestant the option of switching to the other unopened door. If the contestant's goal is to get the car, should she switch doors?

Remarks

- Even great mathematician Paul Erdos & Persi Diaconis made mistakes.
- Originally proposed by Steve Selvin, American Statistician 1975.
- Famous when proposed in Marilyn vos Savant's "Ask Marilyn" column, Parade magazine 1990.
- Approximately 10000 readers, including nearly 1000 with PhDs, said no need to switch.

Solution of Monty Hall Problem

(1) Label the three doors 1, 2, 3.

Homogeneous doors, without loss of generality. we assume door 1 is chosen.

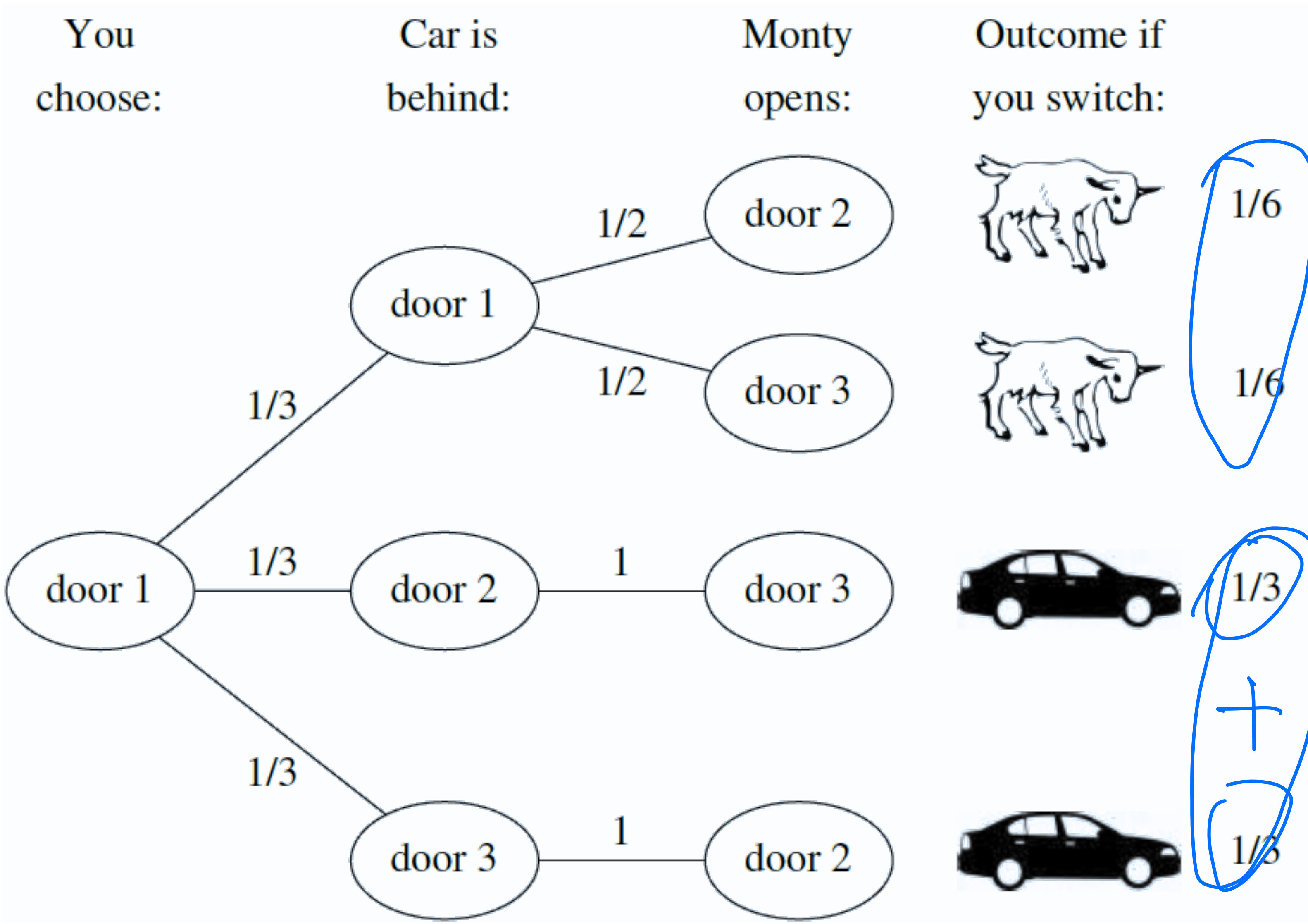
(2) C_i = "the car is behind door i". $i = 1, 2, 3$.

$P(C_i) = 1/3$.

LOTP: $P_{stay}(win) = \underbrace{P(win \mid C_1)}_{1} \underbrace{P(C_1)}_{\times \frac{1}{3}} + \underbrace{P(win \mid C_2)}_{0} P(C_2) + \underbrace{P(win \mid C_3)}_{0} P(C_3)$

$P_{switch}(win) = 0 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} + 1 \cdot \frac{1}{3} = \frac{2}{3}$.

Tree Diagram of Monty Hall Problem



lose

+

win

Information Conveyed by Monty's Action

door 2 is open by monty

M_j = "Monty opens door j". $j = 1, 2, 3$.

C_i = "Car is behind door i". $i = 1, 2, 3$.

door 1 is chosen by the contestent.

$$\underbrace{P(M_2)}_{1/2} = \underbrace{P(M_2|C_1)}_{\frac{1}{2}}\underbrace{P(C_1)}_{\frac{1}{3}} + \underbrace{P(M_2|C_2)}_{0}\times\underbrace{P(C_2)}_{\frac{1}{3}} + \underbrace{P(M_2|C_3)}_{1}\times\underbrace{P(C_3)}_{\frac{1}{3}}.$$

door 2 is open.

After openning door 2: $P(C_1|M_2) = \frac{P(M_2|C_1)P(C_1)}{P(M_2)} = \frac{\frac{1}{2}\times\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$

$P(C_2|M_2) = 0$, $P(C_3|M_2) = \frac{P(M_2|C_3)P(C_3)}{P(M_2)} = 2/3$.

Outline

Setting of Classical Monty Hall Problem

- Strategies of the contestant: switching & no switching.
- Monty knows the location of the car.
- Homogeneous doors.
- Three doors.

Variant 1: New Strategy

In this scenario, when Monty has choice on picking the door, he picks door 2 with probability p, and door 3 with probability $1 - p$. Now besides the strategies of switching and no switching, you (the contestant) have the third option on your strategy:

- You first pick the door 1.
- If the door 2 is opened, you do not switch.
- If the door 3 is opened, you switch.

What is the winning probability with this strategy?

Variant 2: Unknown Car Location for Monty

As before, Monty shows you three identical doors. One contains a car, the other two contain goats. You choose one of the doors but do not open it. This time, however, Monty does not know the location of the car. He randomly chooses one of the two doors different from your selection and opens it. The door turns out to conceal a goat. He now gives you the options either of sticking with your original door or switching to the other one. What should you do?

Variant 3: Partially Known Car Location for Monty

As before, you are shown three equally likely doors. You choose door one. Monty now points to door two but does not open it. Instead he merely tells you that it conceals a goat. You know that in those cases where the car really is behind door one, Monty chooses randomly between door two and door three. You also know that when the car is behind door two or door three, it is Monty's intention to identify the car's location, but that his assertions regarding the location of the car are only correct with probability p, What should you do now?

Variant 4: Heterogeneous Doors

Suppose the car is not placed randomly behind the three doors. Instead, the car is behind door one with probability p_1, behind door two with probability p_2, and behind door three with probability p_3. Here $p_1 + p_2 + p_3 = 1$ and $p1 \geq p2 \geq p3 > 0$. You are to choose one of the three doors, after which Monty will open a door he knows to conceal a goat. Monty always chooses randomly from among his options in those cases where your initial choice is correct. What strategy should you follow?

Variant 5: Progressive Monty

This time we assume there are n identical doors, where n is an integer satisfying $n \geq 3$. One door conceals a car, the other $n-1$ doors conceal goats. You choose one of the doors at random but do not open it. Monty then opens a door he knows to conceal a goat, always choosing randomly among the available doors. At this point he gives you the option either of sticking with your original door or switching to one of the remaining doors. You make your decision. Monty now eliminates another goat-concealing door (at random) and once more gives you the choice either of sticking or switching. This process continues until only two doors remain in play. What strategy should you follow to maximize your chances of winning?

Variant 6: Two Players

As usual, we are presented with three doors. This time, however, there is a second player in the game. Player one chooses a door, and then player two chooses a different door. If both have chosen goats, then Monty eliminates one of the players at random. If one has chosen the car, then the other player is eliminated. The surviving player knows the other has been eliminated, but does not know the reason for the elimination. After eliminating a player, Monty then opens that player's door and gives the surviving player the options of switching or sticking. What should the player do?

Variant 7: Two Hosts

As before, we are confronted with three identical doors, one concealing a car, the other two concealing goats. We initially choose door one, and Monty then opens door three. This time we know that there are two different hosts who preside over the show, with a coin flip deciding who hosts the show on a given night. The two hosts do not make their decisions in the same way. Coin-Toss Monty chooses his door randomly when your initial choice conceals the car. Three-Obsessed Monty always opens door three when he has the option of doing so. Under these circumstances, is there an advantage to be gained from switching to door two?

Variant 8: Many Cars

As before, this time we still have n doors, but now there are $1 \leq j \leq n-2$ cars and $n-j$ goats. After making your initial choice, Monty opens one of the other doors at random. Should you switch?

Variant 9: Many Cars

As before, this time we still have n doors, but now there are $1 \leq j \leq n-2$ cars and $n-j$ goats. This time, however, after making your initial choice, Monty tells us that he will reveal a goat with probability p, and will reveal a car with probability $1-p$. The catch is that we must make our decision to stick or switch before knowing which of these possibilities will come to pass. What should we do?

Variant 10: Many Cars, Open Many Doors

As before, this time we still have n doors, but now there are $1 \leq j \leq n-2$ cars and $n-j$ goats. This time, however, after making your initial choice, Monty opens m doors at random, revealing k cars and $m-k$ goats. What should we do?

Three Prisoners Problem

Three prisoners, A, B and C, are in separate cells and sentenced to death. The governor has selected one of them at random to be pardoned. The warden knows which one is pardoned, but is not allowed to tell. Prisoner A begs the warden to let him know the identity of one of the others who is going to be executed. "If B is to be pardoned, give me C's name. If C is to be pardoned, give me B's name. And if I'm to be pardoned, flip a coin to decide whether to name B or C."

Three Prisoners Problem

The warden tells A that B is to be executed. Prisoner A is pleased because he believes that his probability of surviving has gone up from $1/3$ to $1/2$, as it is now between him and C. Prisoner A secretly tells C the news, who is also pleased, because he reasons that A still has a chance of $1/3$ to be the pardoned one, but his chance has gone up to $2/3$. What is the correct answer?

Three Prisoners Problem

- Proposed in Martin Gardner's "Mathematical Games" column, Scientific American, 1959.
- Mathematically equivalent to the Monty Hall problem:
  - car: freedom
  - goat: death execution
  - Monty: warden
- Very few math problems that have been immortalized in verse.

The Prisoner's Paradox Revisited (Richard Bedient)

Awaiting the dawn sat three prisoners wary
   A trio of brigands named Tom, Dick and Mary
Sunrise would signal the death knoll of two
   Just one would survive, the question was who.

Young Mary sat thinking and finally spoke
   To the jailer she said, "You may think this a joke.
But it seems that my odds of surviving 'til tea,
   Are clearly enough just one out of three.

The Prisoner's Paradox Revisited

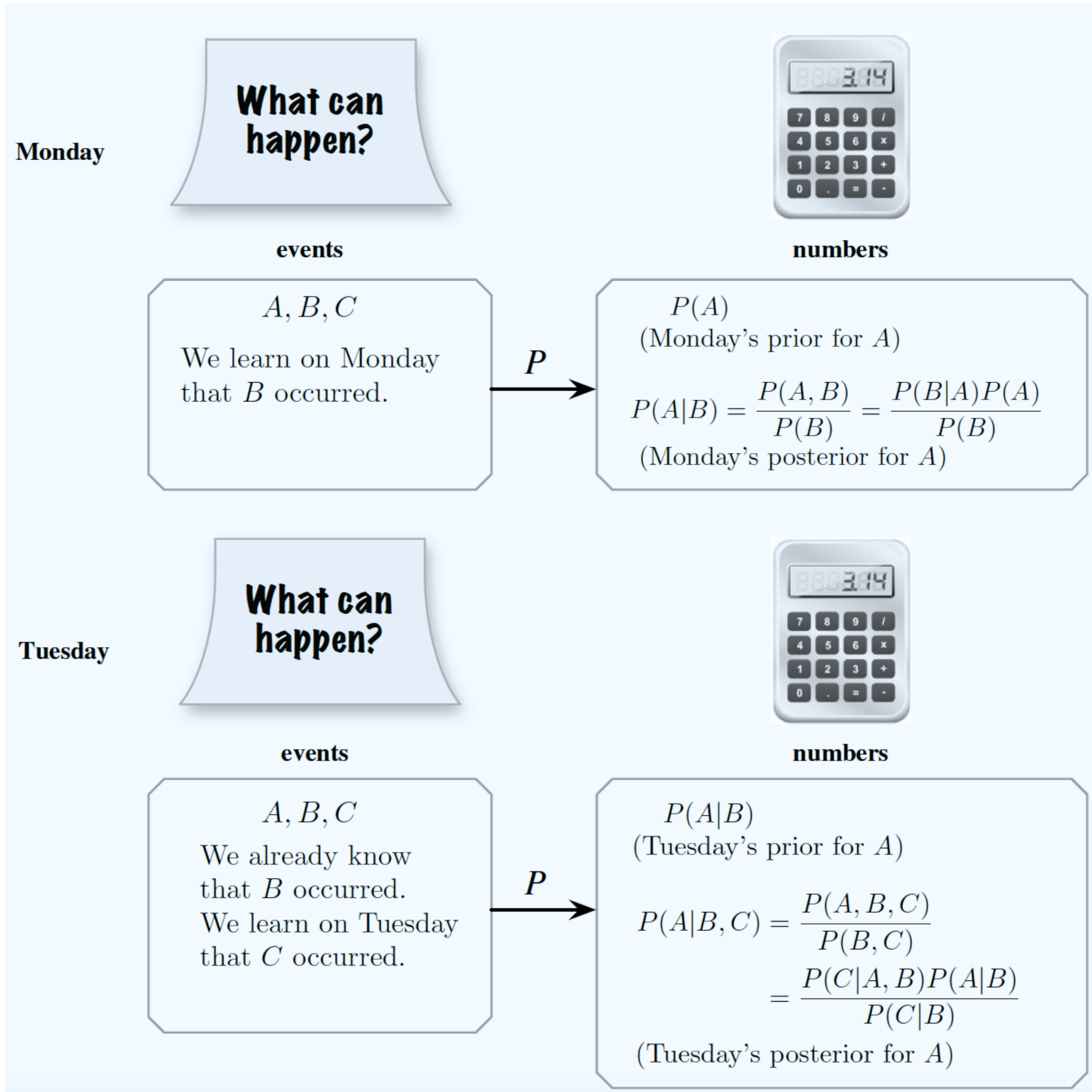
But one of my cohorts must certainly go,
   Without question, that's something I already know.
Telling the name of one who is lost,
   Can't possibly help me. What could it cost?"

That shriveled old jailer himself was no dummy,
   He thought, "But why not?" and pointed to Tommy.
"Now it's just Dick and I," Mary chortled with glee.
   "One in two are my chances, and not one in three!"

Imagine the jailer's chagrin, that old elf.
   She'd tricked him, or had she? Decide for yourself.

Outline

Summary : Conditional Probability & Inference



References

- Chapter 2 of **BH**
- Chapter 1 of **BT**