VANDERBILT
School *of* Engineering

CE 3300-01 – RISK, RELIABILITY, AND RESILIENCE ENGINEERING

Bayesian Learning

Book Reference: Chapters 2

TA: Mohamad Kazma

January 28th-29th, 2025

Instructor: Dr. Hiba Baroud

Probability Rules 2 – Recap on Objectives

- Distinguish probability views
  - Events and Probability
  - Probability Problem Characteristics

- Describe events, outcomes, and corresponding Probabilities
  - Estimating Probabilities
  - Mathematical Operations of Sets

- Calculate marginal, conditional, and joint probabilities
  - The Addition Rule
  - Conditional Probability
  - The Multiplication Rule
  - The Theorem of Total Probability
  - **The Bayes' Theorem**

Bayesian Learning – Learning Objectives

- Discuss Bayesian reasoning

- Explain Bayes Theorem
  - **The Bayes' Theorem**

- Calculating and updating conditional probabilities

Recall: Epistemic Vs. Aleatory Uncertainty

Aleatory Uncertainty – The product of the inherent variability in natural processes.

- An example is the variability of the loads that the structure has to withstand.

Epistemic Uncertainty – The result of the lack of enough knowledge or information about the analyzed system

- This lack of information can be produced by deficiency of data or because the structure's behavior is not correctly represented.

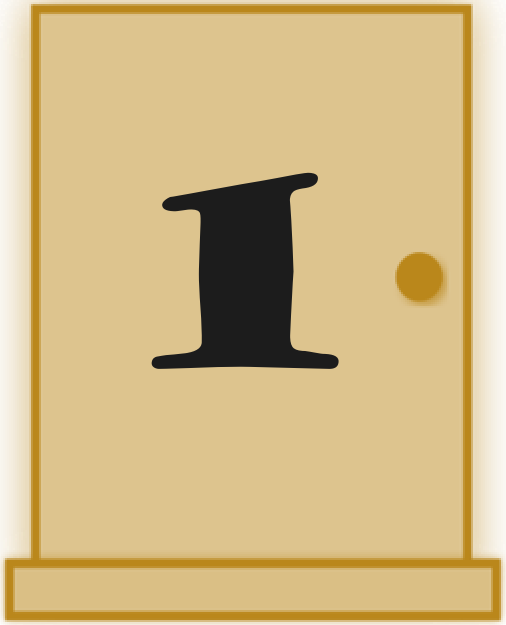
- The more knowledge that is available about a structure or system, the more this type of uncertainty can be reduced.

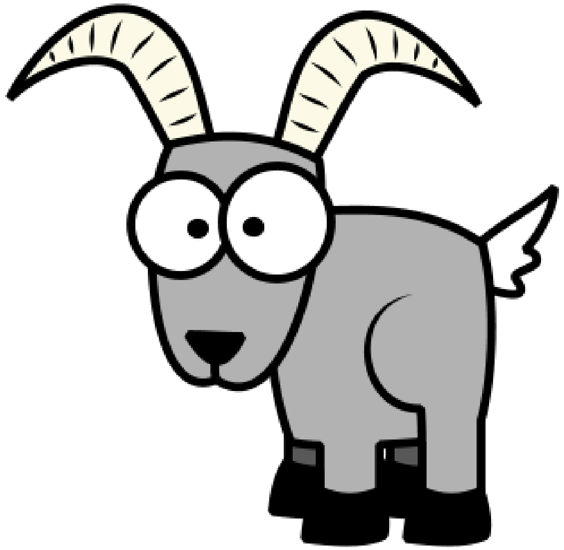
The Monty Hall Problem

The problem was named after Monty Hall, the host of the American television game show **Let's Make a Deal.**

The Monty Hall Problem

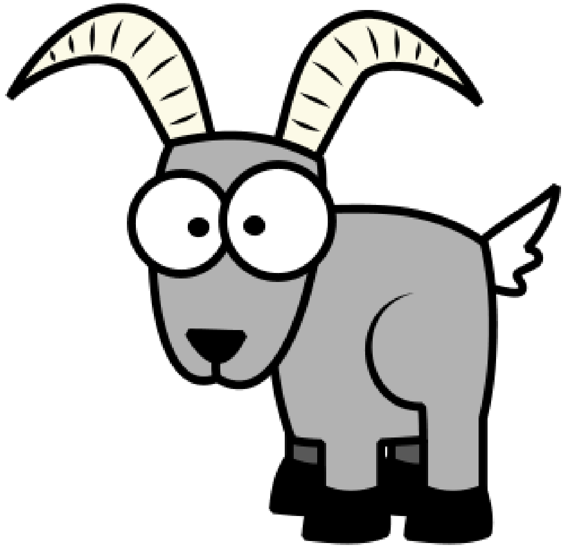
One door is revealed.
Two choices – either stay or swap

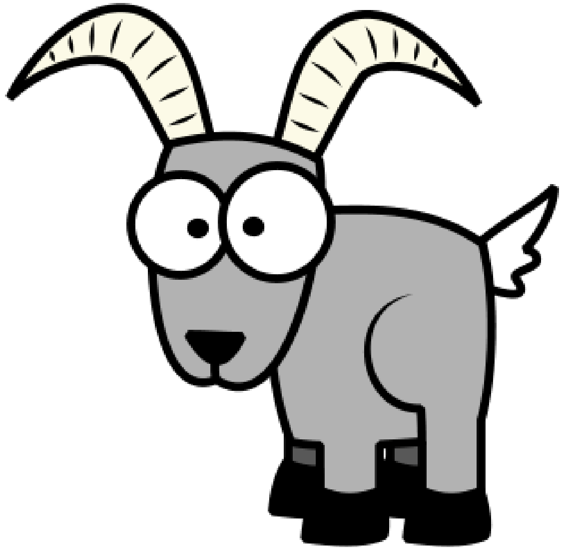




The Monty Hall Problem

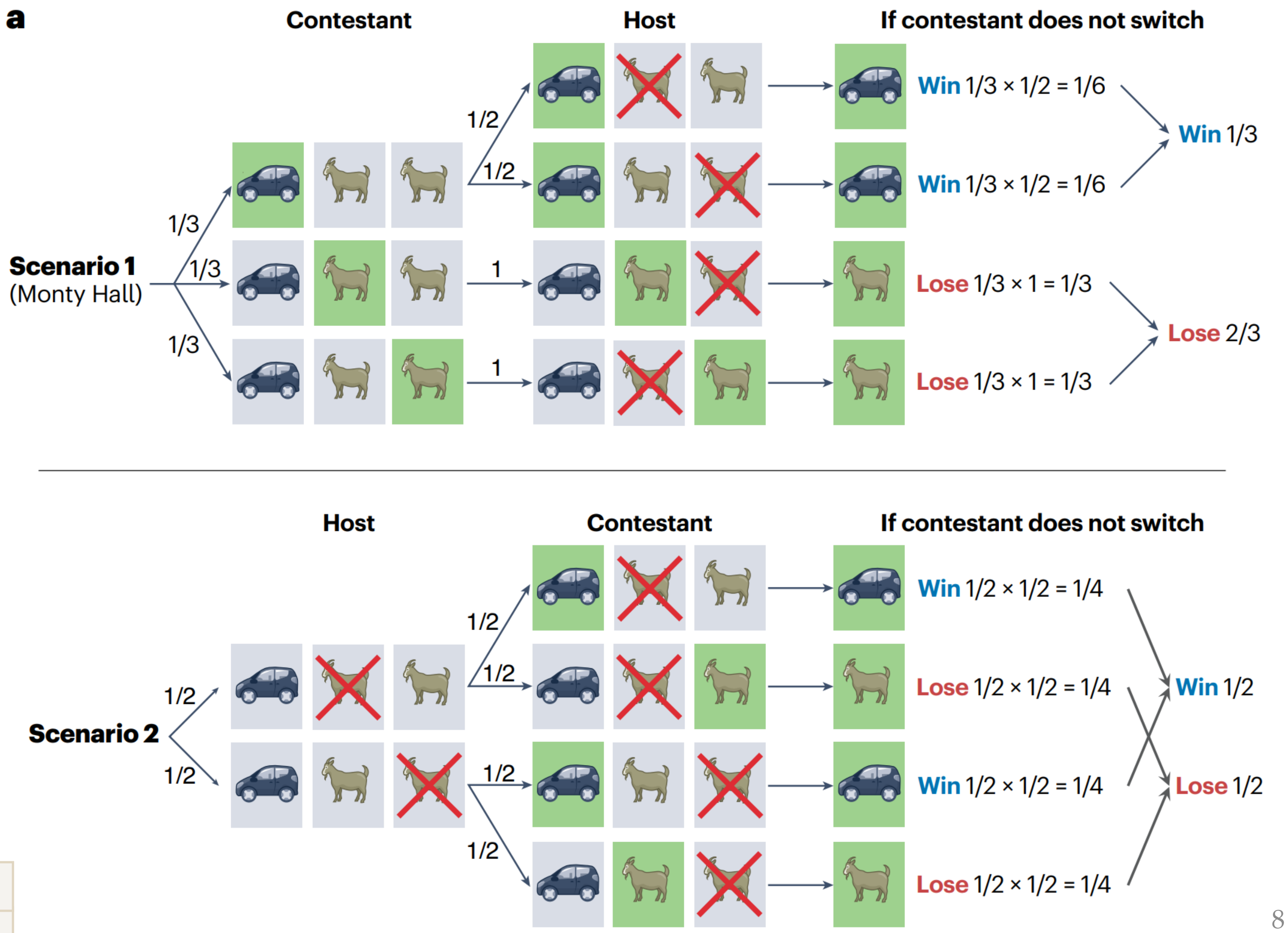
- Would switching increase the contestant's chance of winning?
- What are the probabilities of winning given the **swap or stay** actions?





Let’s Analyze: Two Scenarios

- Order of events matters
- The action after observing a prior event matters



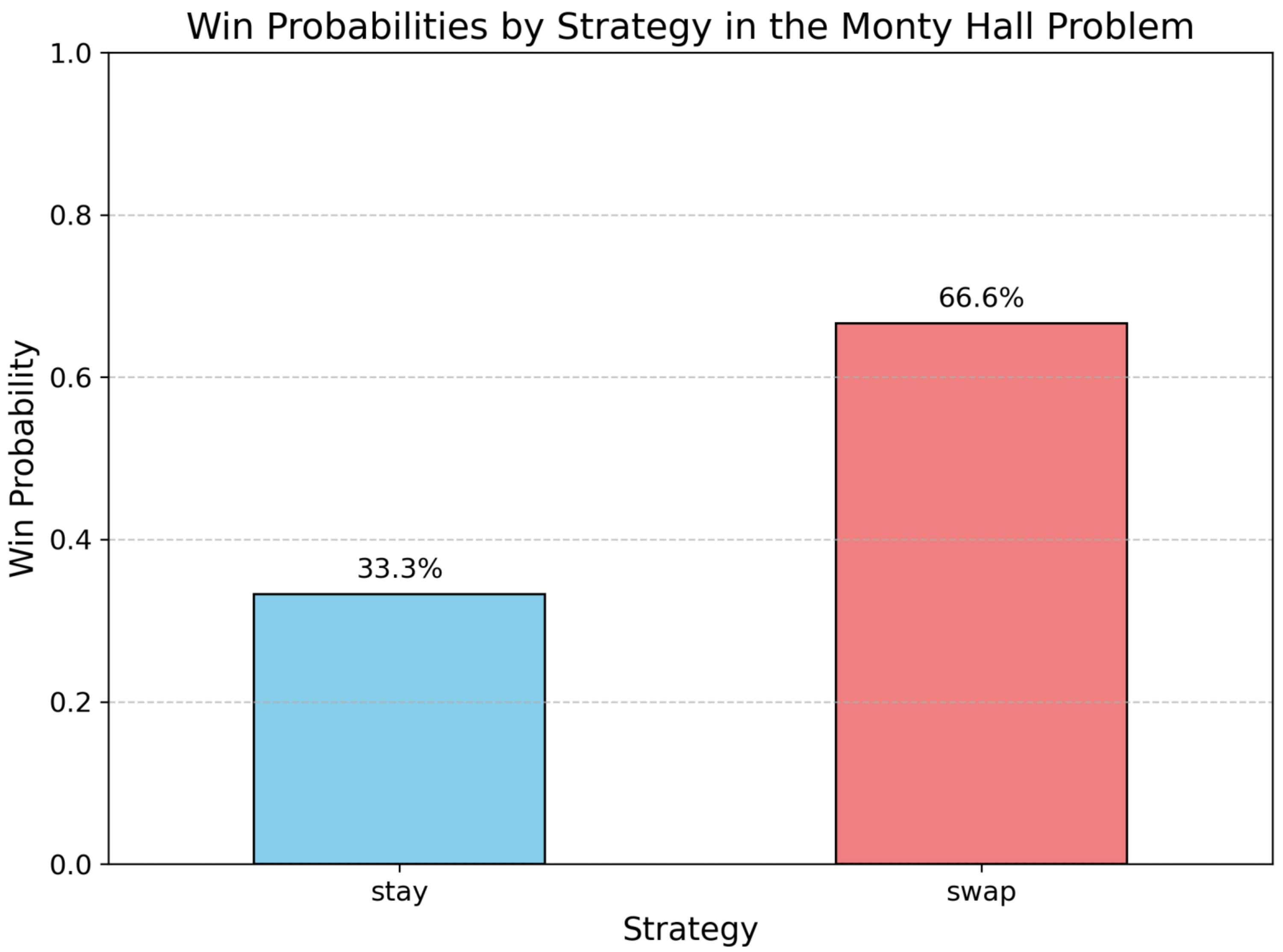
Reading Material

https://doi.org/10.1038/s41587-023-01794-9

The Winning Probabilities

prize_location	initial_choice	reveal_door	stay_or_swap	final_choice	win	win_if_swap	win_if_stay
2	2	1	swap	3	FALSE	FALSE	FALSE
3	3	2	stay	3	TRUE	FALSE	TRUE
3	3	1	stay	3	TRUE	FALSE	TRUE
1	2	3	swap	1	TRUE	TRUE	FALSE
1	3	2	stay	3	FALSE	FALSE	FALSE
1	2	3	swap	1	TRUE	TRUE	FALSE
1	2	3	swap	1	TRUE	TRUE	FALSE
1	3	2	stay	3	FALSE	FALSE	FALSE
1	1	2	swap	3	FALSE	FALSE	FALSE

The winning Probabilities



Intuition: Frequentist Vs. Bayesian Approach

The difference is in the interpretation of uncertainty.

- Observed phenomena are generated by an unknown but fixed process.
- Population parameters are unknown constants, given that complete and exact knowledge about the sample space is not available.

- Bayesian statistics assumes that population parameters, though unknown, are quantifiable random variables.
- Uncertainty about them can be described by probability distributions

Bayes' Theorem

x: "hidden variable", the estimate
y: observations, measurements, data

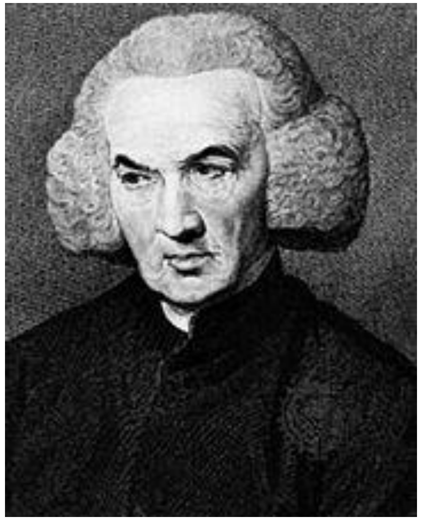
$$p(\mathbf{x} \mid \mathbf{y}) = \frac{p(\mathbf{y}|\mathbf{x})p(\mathbf{x})}{p(\mathbf{y})}$$

- $p(\mathbf{x} \mid \mathbf{y})$ is the **posterior probability**, function of $\mathbf{y}$
- $p(\mathbf{x})$ is the **prior or marginal probability** of $\mathbf{x}$; prior in the sense that it does not take into account the data.
- $p(\mathbf{y})$ is a **normalizing factor**, the data.
- $p(\mathbf{y} \mid \mathbf{x})$ is the **likelihood of the observation y** given $\mathbf{x}$

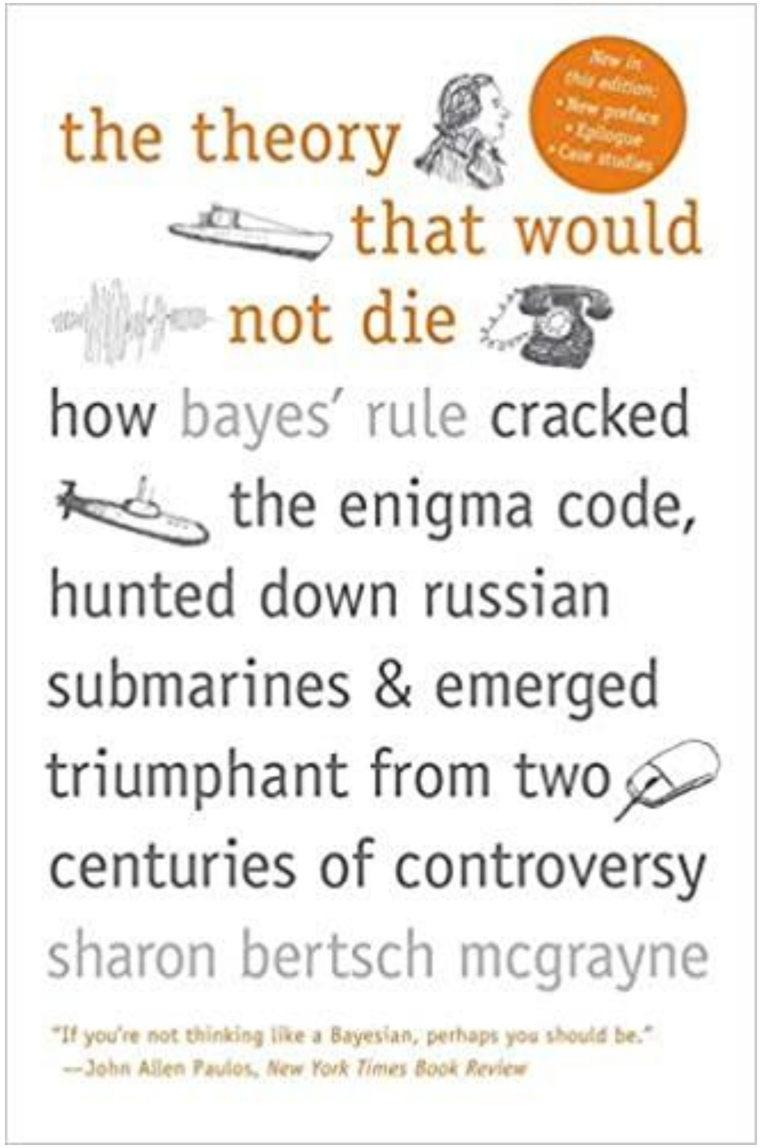
A Bit of History

"An essay towards solving a problem
in the Doctrine of Chances"

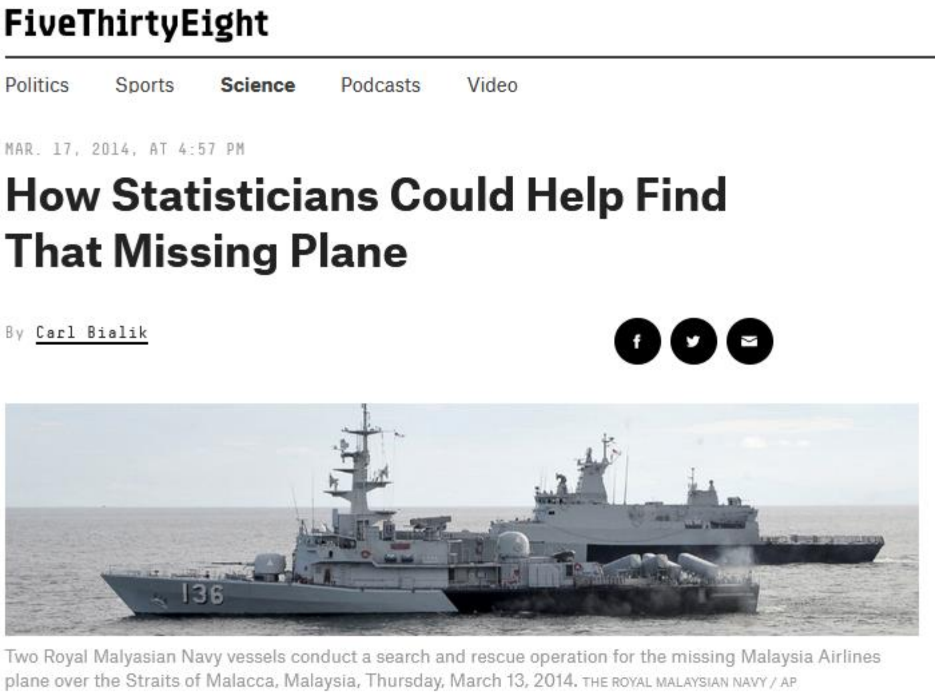
(1763)



A Bit of History



- First publication of Bayesian methods for applied scientists: 1963
- Bayesian statistics is a way of thinking



Example 1 – Monty Hall: A Bayesian Approach

Scenario:

1. You select one of the doors, say Door A.

2. The host, who knows what is behind each door, then opens one of the two remaining doors (e.g., Door B) to reveal a goat.

3. You are then offered a choice: either stick with your original selection (Door A) or switch to the other unopened door (Door C).

Using Bayes' theorem,

1. Determine if you should stick with your initial choice or switch doors to maximize your chances of winning.

2. Calculate the probability of winning in each case.

Example 1 – Monty Hall: A Bayesian Approach

$$P(H \mid E) = \frac{P(E|H)P(H)}{P(E)}$$

Bayes' Theorem

- H_1: The car is behind Door A
- H_2: The car is behind Door B
- H_3: The car is behind Door C

Before any doors are opened, the probabilities of the car being behind each door are equal

Priors

$$P(H_1) = P(H_2) = P(H_3) = \frac{1}{3}$$

Example 1 – Monty Hall: A Bayesian Approach

The **evidence** is that the host opens Door B, revealing a goat. The **likelihood** depends on where the car is:

Likelihood $P(E \mid H)$

- If H_1 (car is behind Door A): $P(E \mid H_1) =$
- If H_2 (car is behind Door B): $P(E \mid H_2) =$
- If H_3 (car is behind Door C): $P(E \mid H_3) =$

Example 1 – Monty Hall: A Bayesian Approach

The **evidence** is that the host opens Door B, revealing a goat. The **likelihood** depends on where the car is:

Likelihood $P(E \mid H)$

- If H_1 (car is behind Door A): $P(E \mid H_1) = \frac{1}{2}$
- If H_2 (car is behind Door B): $P(E \mid H_2) = 0$
- If H_3 (car is behind Door C): $P(E \mid H_3) = 1$

The **total probability** of the evidence is computed by summing over all H:

$$P(E) = P(E \mid H_1)\,P(H_1) + P(E \mid H_2)\,P(H_2) + P(E \mid H_3)\,P(H_3)$$
$$P(E) = \frac{1}{2}$$

Example 1 – Monty Hall: A Bayesian Approach

Posterior Probabilities

Bayes' Theorem

$$P\left(H_1 \mid E\right) = \frac{P(E|H_1)P(H_1)}{P(E)} =$$

$$P\left(H_3 \mid E\right) = \frac{P(E|H_3)P(H_3)}{P(E)} =$$

Example 1 – Monty Hall: A Bayesian Approach

Posterior Probabilities

Bayes' Theorem

$$P\left(H_1 \mid E\right) = \frac{P(E|H_1)P(H_1)}{P(E)} = \frac{\frac{1}{2}\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

$$P\left(H_3 \mid E\right) = \frac{P(E|H_3)P(H_3)}{P(E)} = \frac{1\cdot\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

Switch to maximize your chances of winning

Conditional and Total Probabilities

The conditional probability of A given B, denoted as $P(A \mid B)$, is defined as:

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)} \quad \text{if } P(B) > 0$$

Similarly, the conditional probability of B given A, denoted as $P(B \mid A)$, is:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} \quad \text{if } P(A) > 0$$

We can relate these conditional probabilities by first writing the joint probability of selecting

Since $P(A \cap B)$ is symmetric (the event " A and B' is the same as " B and A'),

Conditional and Total Probabilities

To express $P(A \mid B)$ in terms of $P(B \mid A)$, rearrange the equation:

$$P(A \mid B) = \frac{P(B \mid A)P(A)}{P(B)}$$

Using **the law of total probability** if A is part of a partition of the sample space.

$$P\left(A_{\mathrm{i}} \mid B\right) = \frac{P(B|A_i)P(A_i)}{\sum_{j=1}^{\mathrm{n}} P(B|A_j)P(A_j)}$$

Example 2 - Cousin's Gender

Your distant cousin has two kids with genders equally likely to occur.

a. What is the probability that both are girls?
b. If you remember that one is a girl, what the probability that both are girls?

(a) $\{XX, XX\}$ $\{XX, XY\}$ $\{XY, XY\}$ $\{XY, XX\}$

$$P(a = XX, b = XX) = \frac{1}{4}$$

(b)

$$P(b = XX \mid a = XX) = \frac{1}{3}$$

Mammogram Problem

- Let's say 1% of women have breast cancer
- Suggesting that 99% of women do not

$$P(\text{ cancer }) = 0.01$$
$$P(\text{ no cancer }) = 0.99$$

- And let's say 80% of mammograms detect breast cancer when it is there (true positive)
- Suggesting that 20% miss the cancer (false negative)

Conditional Probabilities

$$P(\text{test positive} \mid \text{cancer}) = 0.80$$
$$P(\text{test negative} \mid \text{cancer}) = 0.20$$

- And let's say 9.6% of mammograms detect breast cancer when it is not there (false positive)
- Suggesting that 90.4% correctly return a negative result (true negative)

$$P(\text{test positive} \mid \text{no cancer}) = 0.096$$
$$P(\text{test negative} \mid \text{no cancer}) = 0.904$$

The question

Question
If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?

$0\% - 30\% \qquad 30\% - 60\% \qquad 60\% - 100\%$

Answer
If a woman at age 40 is tested as positive, what is the probability that she indeed has breast cancer?

$\boxed{0\% - 30\%} \qquad 30\% - 60\% \qquad 60\% - 100\%$

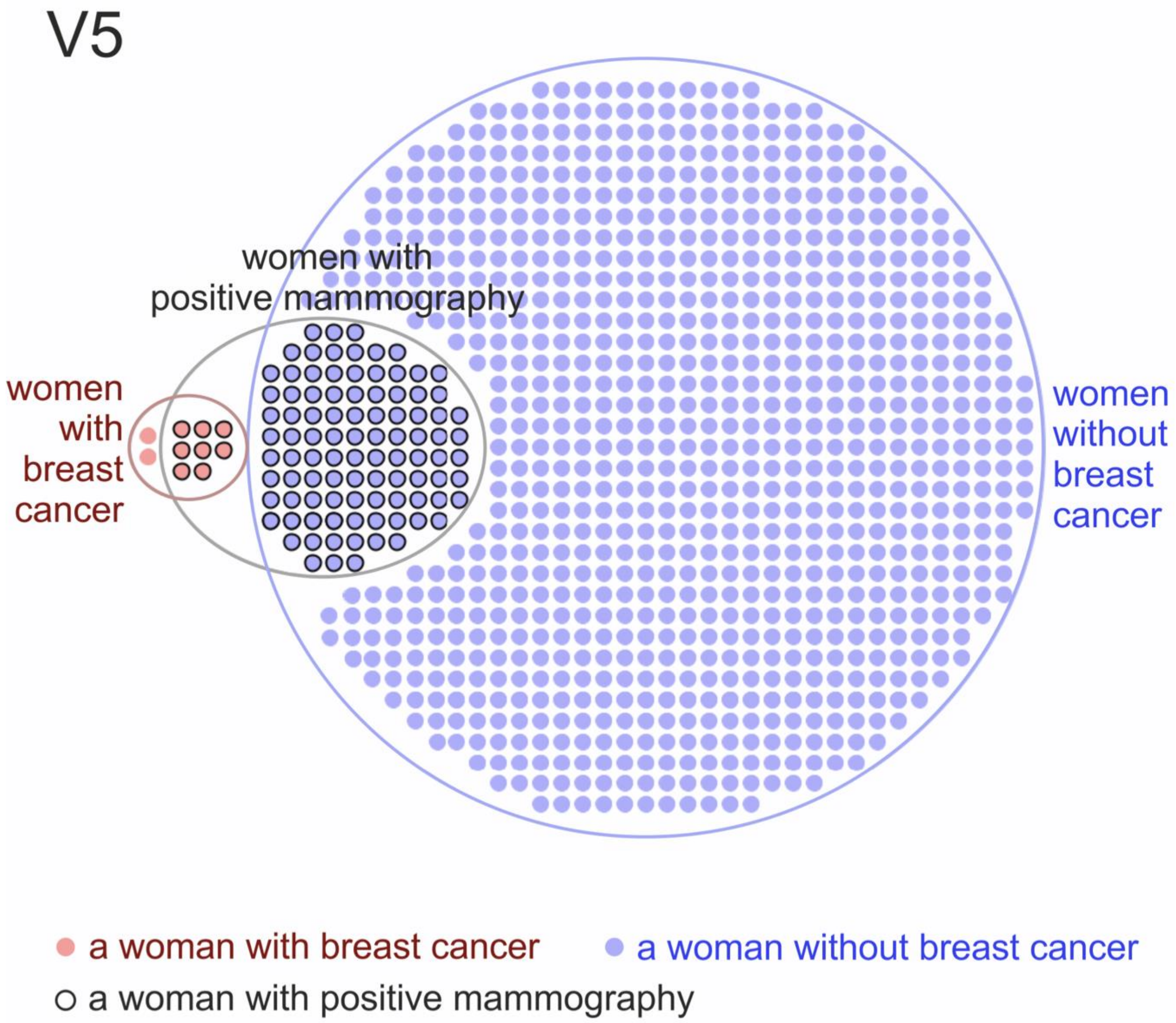
The Doctors' Answers

Question
Why is the actual probability lower?

95 doctors out of 100

said the answer is between 70% to 80%

Mammogram Problem – Ven Diagram



Given that a woman tested positive.

What is the probability she has cancer?

Mammogram Problem – Bayes' Theorem

$$P(\text{ cancer } \mid \text{ test positive }) =$$

$$= \frac{P(\text{ test positive } \mid \text{ cancer })P(\text{ cancer })}{P(\text{ test positive })}$$

$$= \frac{(0.8)(0.01)}{(0.8)(0.01) + (0.096)(0.99)} = 0.0776$$

$$P(B \cap C) = P(B \mid C)P(C)$$

$$P(B \cap \bar{C}) = P(B)$$

8 OUT OF 100 PEOPLE WHO TEST POSITIVE HAVE CANCER

Mammogram Problem – Total Probability

$$P(A \mid B) = 1 - P(\bar{A} \mid B)$$

$$P(\text{ no cancer } \mid \text{ test positive }) = \frac{P(\text{ test positive } \mid \text{ no cancer })P(\text{ no cancer })}{P(\text{ test positive })}$$

$$= \frac{(0.096)(0.99)}{(0.096)(0.99) + (0.80)(0.01)} = 0.9224$$

Shortcut: $P\,(\text{no cancer} \mid \text{test positive}) = 1 - P(\text{ cancer} \mid \text{test positive })$

Example 3 – Highway Safety

- A recent highway safety study found that 77% of accidents the driver was wearing a seat belt.
- Accident reports indicated that 92% of those drivers escaped serious injury but only 63% of the non-belted drivers were so fortunate.

What is the probability that a driver who was seriously injured wasn't wearing a belt?

Let B: wearing a belt, P: serious injury

$$P(B) = 0.77 \quad P(\bar{I} \mid B) = 0.92$$

$$\Rightarrow P(\bar{I} \mid \bar{B}) = 0.63$$

$$P(\bar{B} \mid I) = ?$$

Example 3 – Continued

$$P(I \mid \bar{B}) = 0.37$$

$$\begin{aligned} P(I \cap \bar{B}) &= P(I \mid \bar{B})P(\bar{B}) \\ &= (1 - 0.63)(1 - 0.72) \end{aligned}$$

$$P(\bar{B} \mid I) = \frac{P(I \mid \bar{B})P(\bar{B})}{P(I)}$$

$$P(I) = P(I \cap B) + P(I \cap \bar{B})$$

$$= \frac{(1 - 0.63)(1 - 0.77)}{(1 - 0.63)(1 - 0.77) + (1 - 0.92) \times 0.77}$$

$$P(\bar{B} \mid I) = 0.58$$

Example 4 – Updating Probabilities 1

Biased Coin Toss Problem

We have two coins, one fair (C) and one biased (Cb).

- The probability of tossing heads with the fair coin is 50%.

- The probability of tossing heads with the biased coin, $P(H|Cb) = 0.75$.

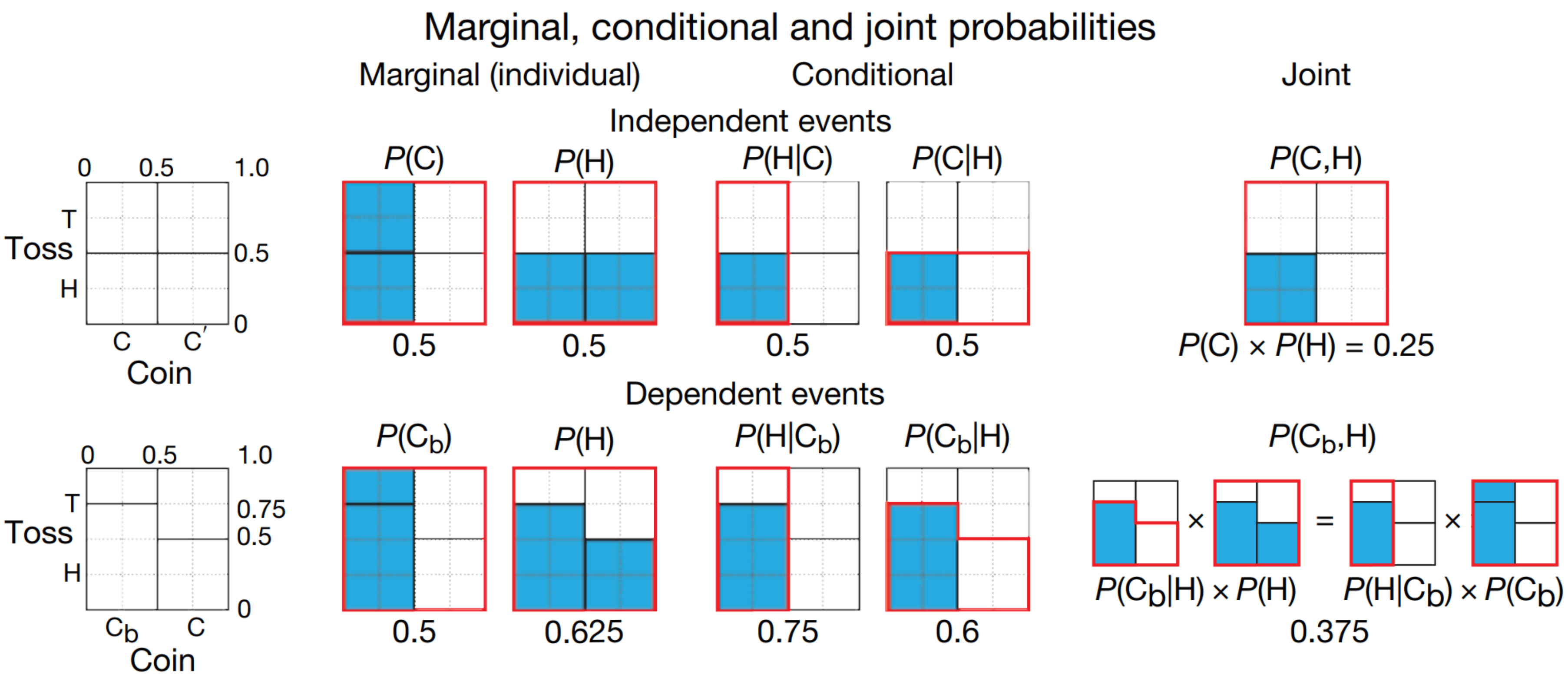
Reading Material

https://doi:10.1038/nmeth.3335

Example 4 – Continued

Fair Coin

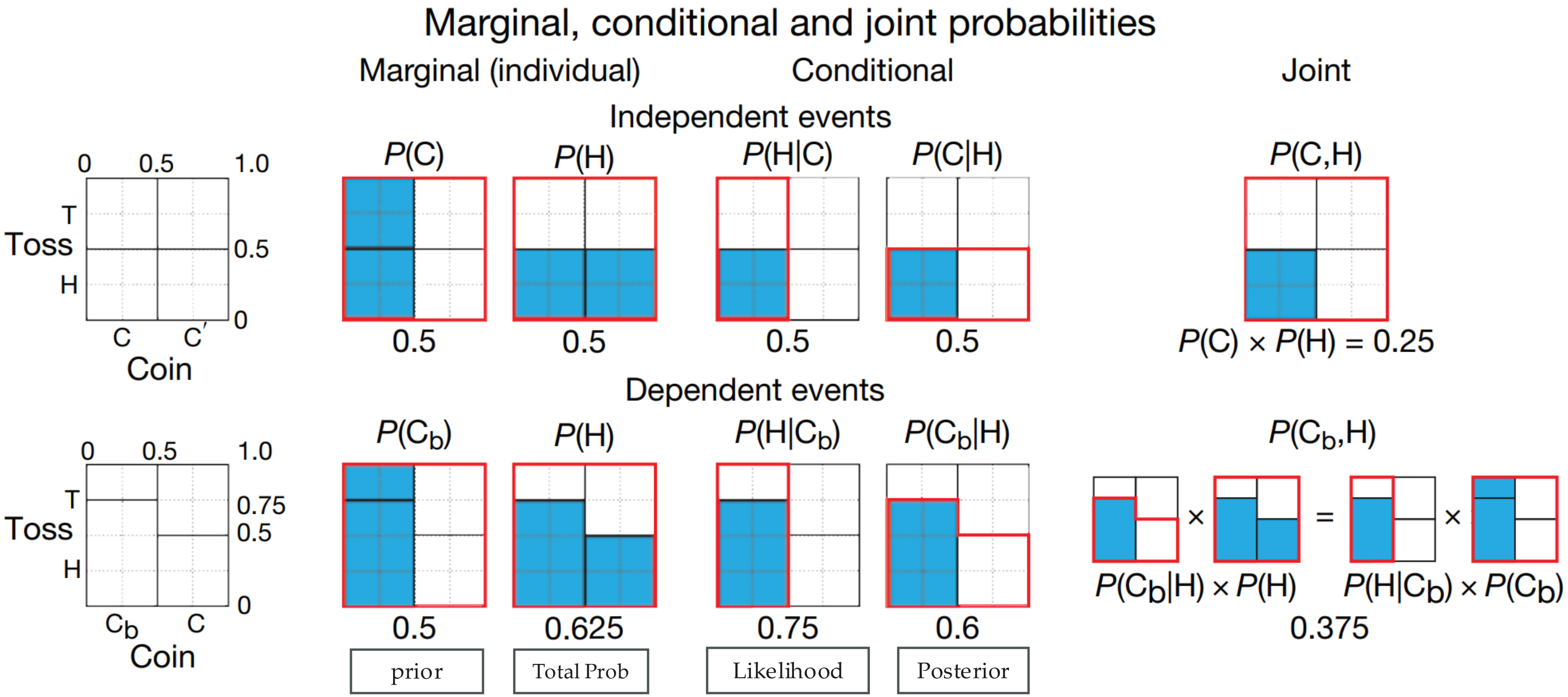
$$P(\mathrm{H}) = 0.5 \qquad P(\mathrm{C},\mathrm{H}) = P(\mathrm{C}) \times P(\mathrm{H})$$



Example 4 – Continued

Biased coin towards heads

$$P(\mathrm{H} \mid \mathrm{C}) = 0.5 \text{ and } P(\mathrm{H} \mid \mathrm{C_b}) = 0.75$$



Leads to Bayes Theorem

Conditional Probabilities

Example 4 – Continued

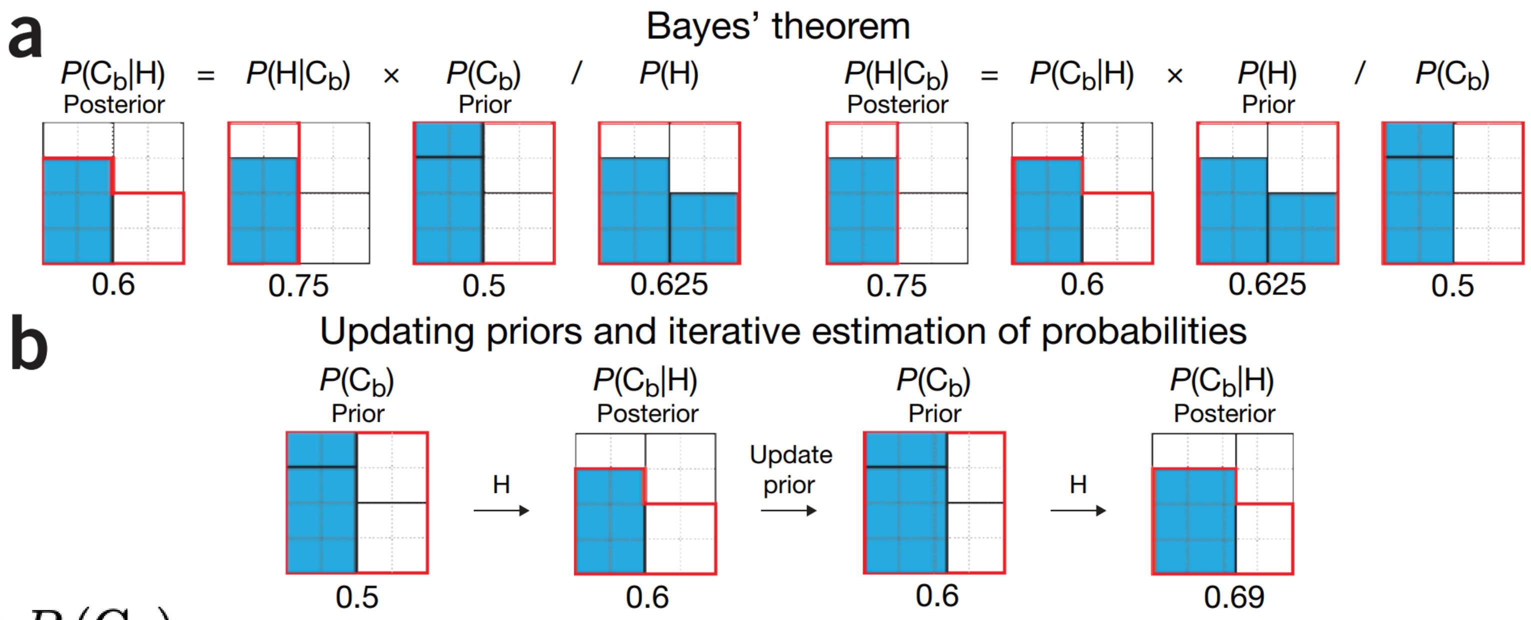
Biased Coin

If we select a coin at random, we can use Bayes' theorem to calculate the probability that the coin we selected was the biased one, given outcome of tosses.

Posterior

$$P\left(\mathrm{C_b} \mid \mathrm{H}\right) = \frac{P(\mathrm{H}|\mathrm{C_b}) \times P(\mathrm{C_b})}{P(\mathrm{H})}$$

$$P\left(\mathrm{C_b} \mid \mathrm{H}\right) = \frac{0.75 \times 0.5}{0.625}$$



Probability of observing heads

$$P(\mathrm{H}) = P(\mathrm{H} \mid \mathrm{C}) \times P(\mathrm{C}) + P\left(\mathrm{H} \mid \mathrm{C_b}\right) \times P\left(\mathrm{C_b}\right)$$
$$= 0.5 \times 0.5 + 0.75 \times 0.5 = 0.625$$

Example 4 – Continued

toss	*toss outcome* 0 = tail, 1 = head	*posterior* P(Cb|H)	*likelihood* P(H|Cb)	*prior* P(Cb)	P(H)
1	1	0.600	0.75	0.500	0.625
2	1	0.692	0.75	0.600	0.650
3	1	0.771	0.75	0.692	0.673
4	1	0.835	0.75	0.771	0.693
5	1	0.884	0.75	0.835	0.709
6	1	0.919	0.75	0.884	0.721
7	1	0.945	0.75	0.919	0.730
8	1	0.962	0.75	0.945	0.736
9	1	0.975	0.75	0.962	0.741
10	1	0.983	0.75	0.975	0.744
11	1	0.989	0.75	0.983	0.746
12	1	0.992	0.75	0.989	0.747

Example 5 – Disease and Genomic Markers

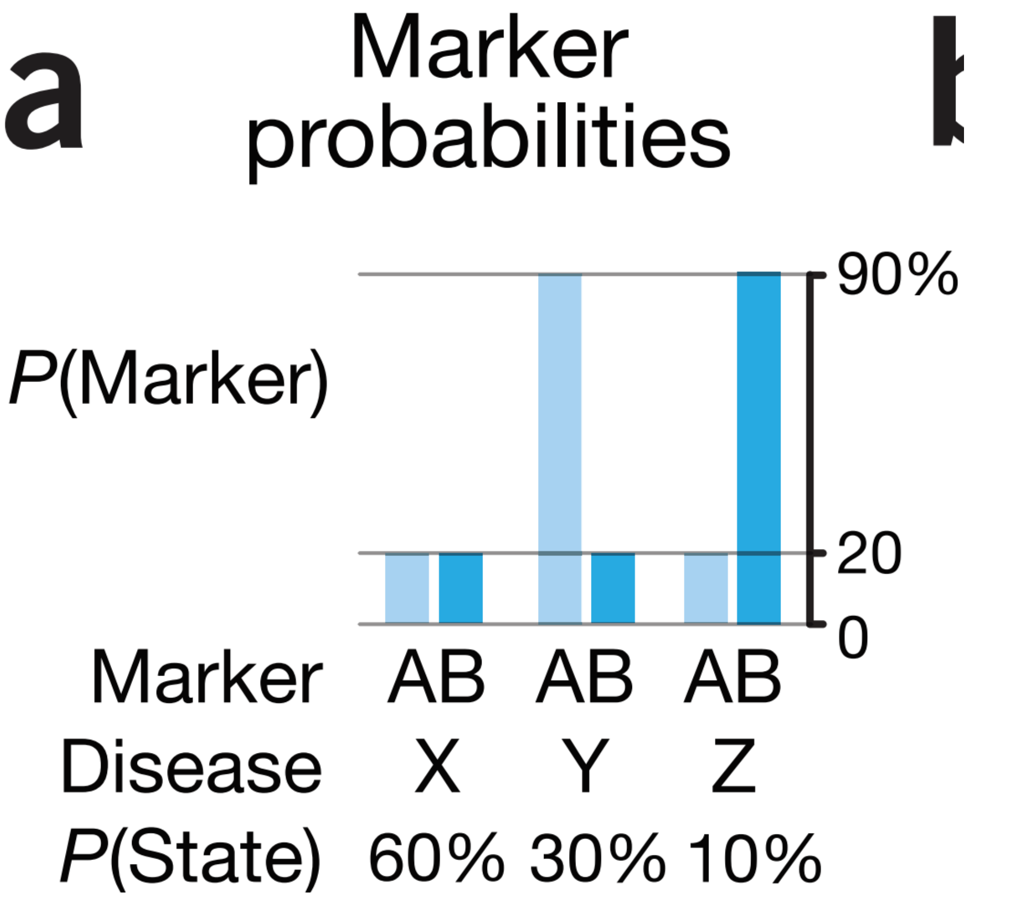
Patient's Disease

We have three diseases D = {X, Y, Z} and two biomarkers M = {A, B}.

- The probability to **find a given marker** in the blood of a patient with a **disease is given** in the first table below.
- Using Bayes' theorem, we can calculate the probability that the **patient has a given disease having observed one of the markers in their blood.**

Example 5 – Continued

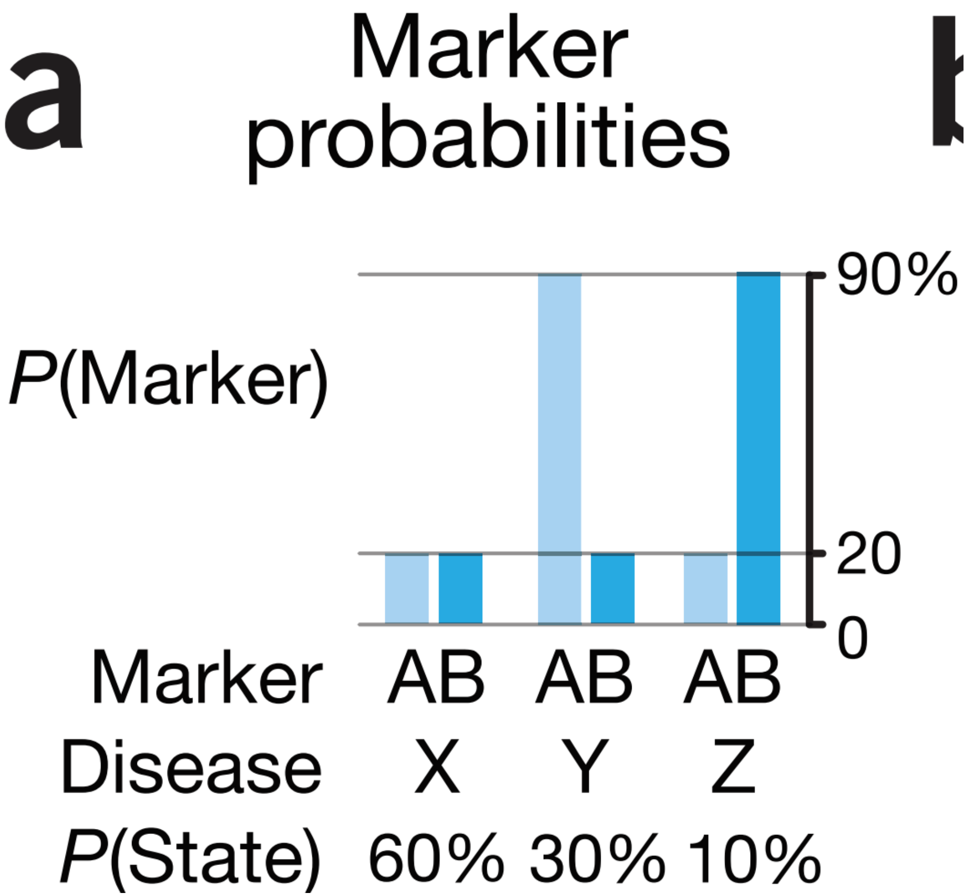
Disease	P(disease)	*Predicted marker*	
		A	**B**
X	0.6	0.2	0.2
Y	0.3	0.9	0.2
Z	0.1	0.2	0.9



Example 5 – Continued

Disease	P(disease)	*Predicted marker*	
		A	**B**
X	0.6	0.2	0.2
Y	0.3	0.9	0.2
Z	0.1	0.2	0.9

$$P(M \mid D)$$



Example 5 – Continued

Probability of disease given that a marker is observed.

$$P(D \mid M) = \frac{P(M|D) \times P(D)}{P(M)}$$

Bayes' Theorem

The probability of disease X given that marker A is observed.

$$P(\mathrm{X} \mid \mathrm{A}) = \frac{P(\mathrm{A}|\mathrm{X}) \times P(\mathrm{X})}{P(\mathrm{A})}$$

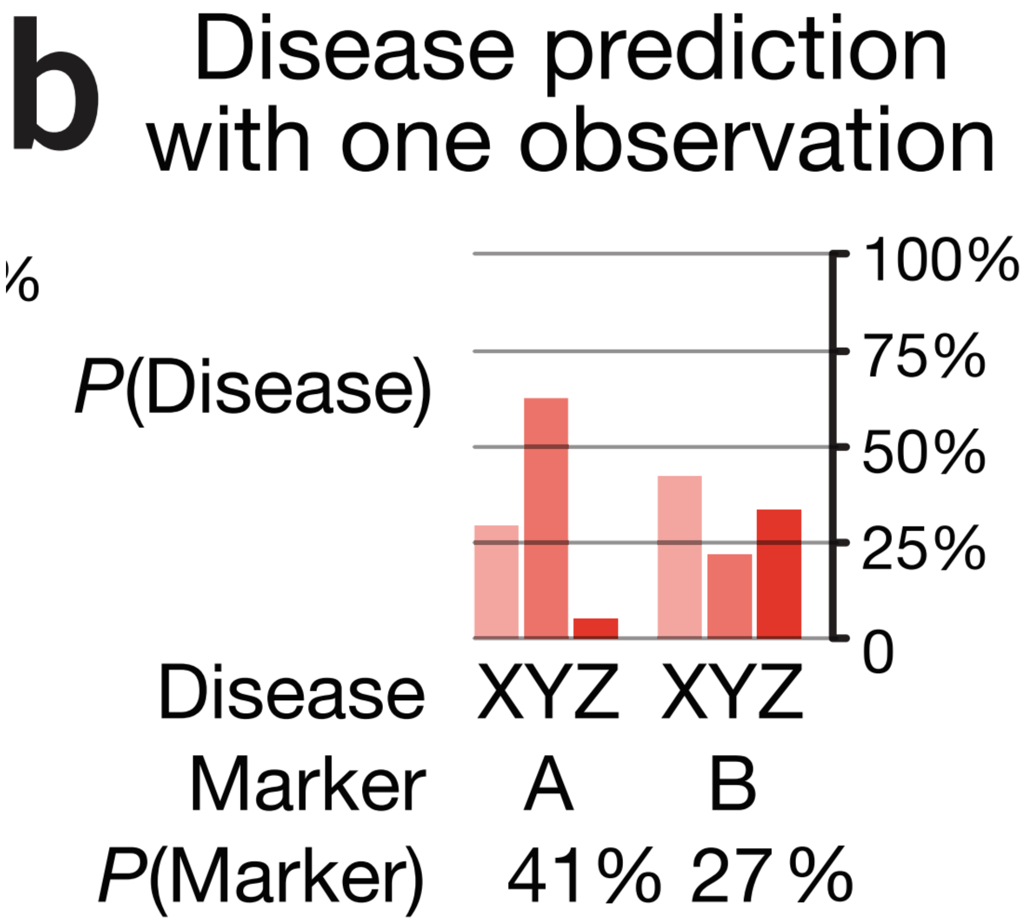
Example 5 – Continued

$$\boxed{P(\mathrm{A} \mid \mathrm{X}) = 0.2}$$

$$\boxed{P(\mathrm{X}) = 0.6}$$

$$P(\mathrm{X} \mid \mathrm{A}) = \frac{P(\mathrm{A}|\mathrm{X}) \times P(\mathrm{X})}{P(\mathrm{A})}$$

$$\boxed{P(A) = 0.6 \times 0.2 + 0.3 \times 0.9 + 0.1 \times 0.2 = 0.41}$$



$$P(\mathrm{X} \mid \mathrm{A}) = 0.2 \times 0.6/0.41 = 0.29$$

Example 5 – Continued: Observation 2

If **marker A is found**, the most likely disease is **Y (65.9%)**.

Disease	Likelihood P(A|d)	Prior P(d)	P(A) = 0.41 P(A|d)xP(d)	Posterior P(d|A)
X	0.2	0.6	0.12	0.29
Y	0.9	0.3	0.27	0.66
Z	0.2	0.1	0.02	0.05

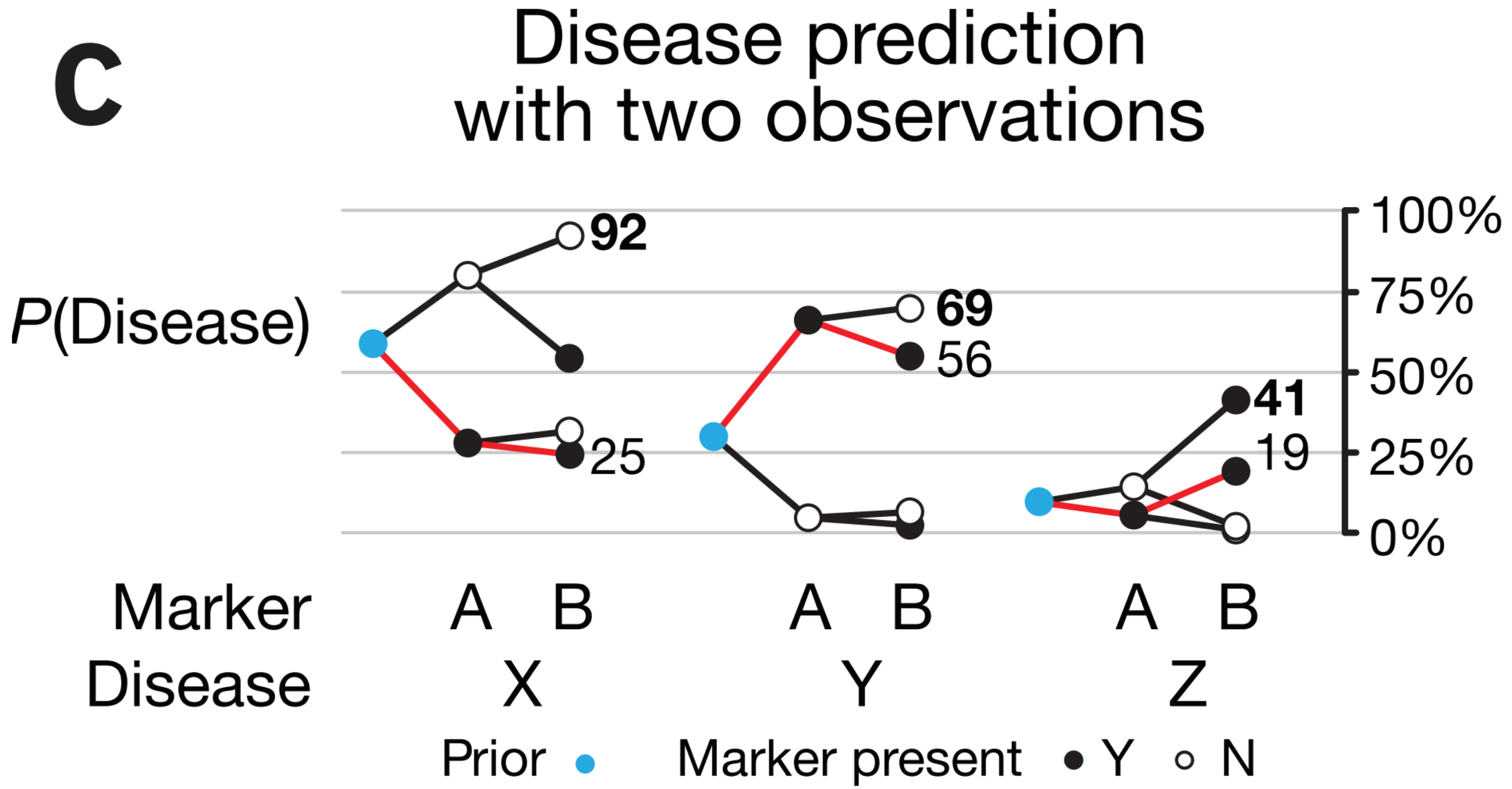
Example 5 – Continued: Observation 2

Having observed A , we could refine our predictions by testing for B.

- We use the posterior probability of the disease after observing A as the new prior.

Disease	Likelihood P(B|d)	Prior P(d)	P(B) = 0.23 P(B|d)xP(d)	Posterior P(d|B)
X	0.2	0.29	0.06	0.25
Y	0.2	0.66	0.04	0.56
Z	0.9	0.05	0.23	0.19

Example 5 – Continued: Observation 2



Next: Discrete Random Variables