BITS Pilani

DIGITAL

Excellence made yours

Probability and Statistics

Week 4 Live Session

Dr. Pushpendra Gupta

Recap

- 📊 Chebyshev Inequality
- 📏 Distribution Analysis
- 📈 Data Visualization
- 🎯 Choosing the Right "Average"

Today’s Focus

- 🧠 Recap
- 🎯 Probability
- ⚖️ Conditional Probability
- 💻 Hands-on Visualization using Google Colab

What is Probability?

Probability measures how likely an event is to occur.

Expressed as a number between 0 and 1:

- 0: Impossible event
- 1: Certain event
- 0.5: Equal chance (e.g., coin flip)

Core Concept

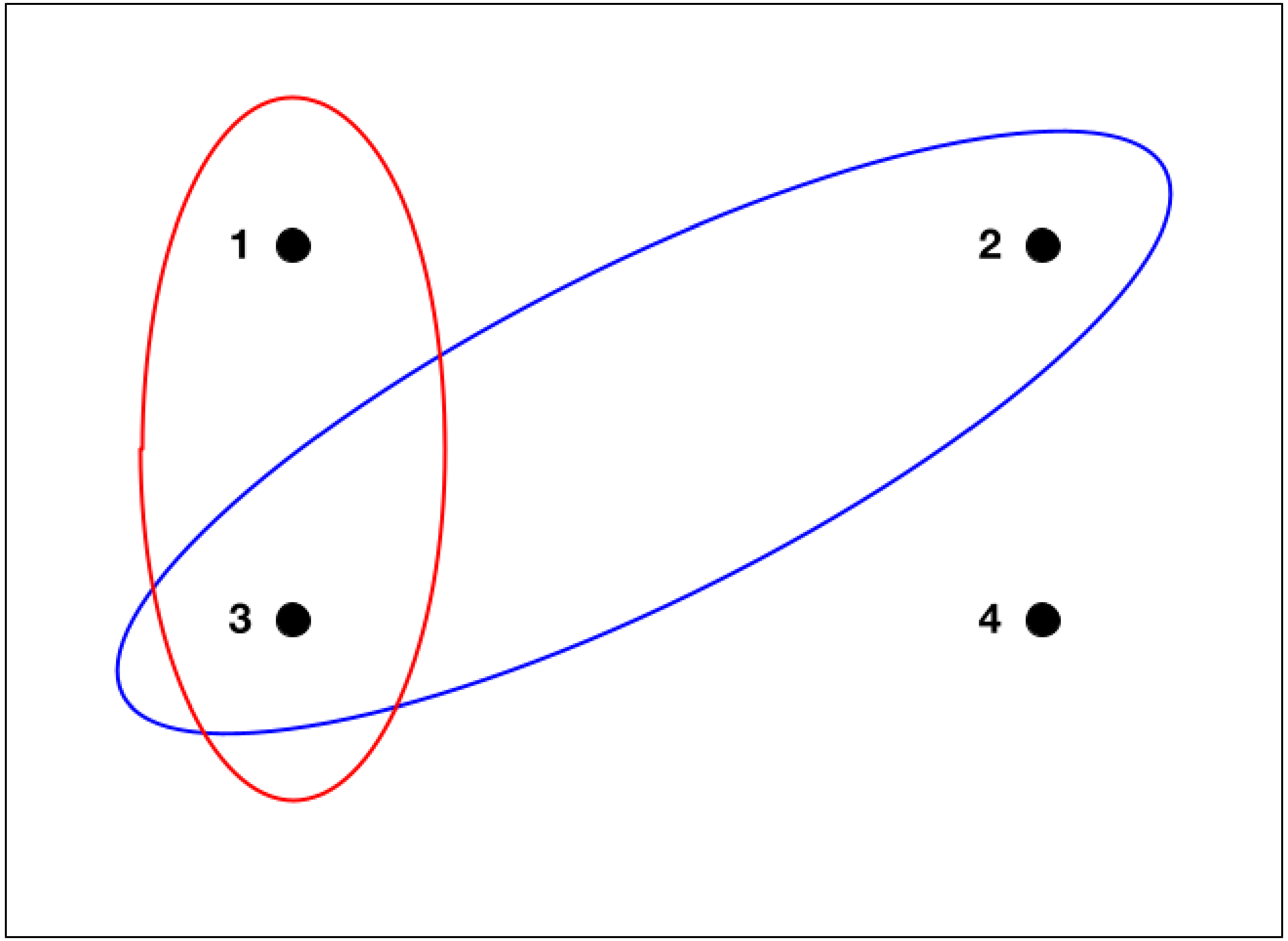
$$P(A) = \frac{Number\ of\ Favorable}{Number\ of\ Possible\ Outcomes}$$

Experiment: Action with observable results (e.g., rolling a die).

Sample Space (S): All possible outcomes (e.g., {1,2,3,4,5,6} for a die).

Event (A): Subset of sample space (e.g., rolling an even number).

Favorable Outcome: Results that satisfy the event.

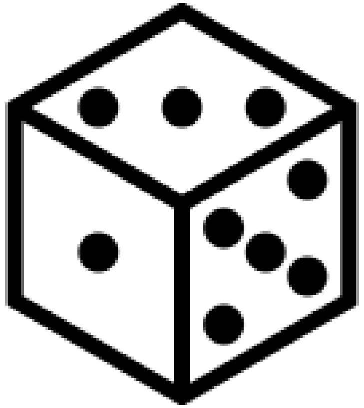


https://en.wikipedia.org/wiki/Sample_space

Simple Events: Rolling a Die

$$P(A) = \frac{Number\ of\ Favorable}{Number\ of\ Possible\ Outcomes}$$

Experiment: Roll a fair six-sided die.



Sample Space: {1, 2, 3, 4, 5, 6} (Total outcomes: 6).

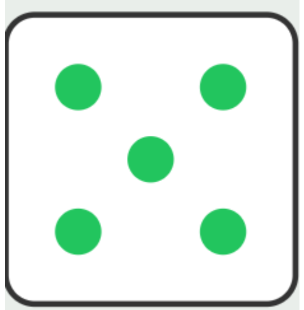
We'll calculate probabilities for specific events.

Rolling Greater Than 4

BITS Pilani DIGITAL Excellence made yours

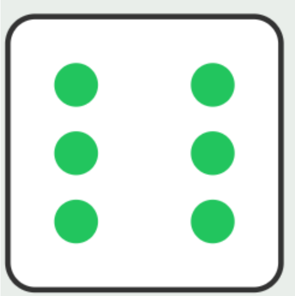
Event: Roll a number > 4.

Favorable Outcomes: {5, 6} (2 outcomes).



$$P(>4) = \frac{2}{6} = \frac{1}{3} \approx 0.3333\ (33.33\%)$$

```python
favorable_outcomes = 2
total_outcomes = 6
probability = favorable_outcomes / total_outcomes
print(f"Probability: {probability}")
```



Rolling an Even Number or Rolling <=3

Event: Roll an even number.

Favorable Outcomes: {2, 4, 6} (3 outcomes).

$$P(even) = \frac{3}{6} = \frac{1}{2} \approx 0.5\ (50\%)$$

Event: Roll a number ≤ 3

Favorable Outcomes: {1, 2, 3} (3 outcomes).

$$P(\leq 3) = \frac{3}{6} = \frac{1}{2} \approx 0.5\ (50\%)$$

Probability of 1 to 6

Event: Roll a number between 1 and 6 (inclusive).

Favorable Outcomes: {1, 2, 3, 4, 5, 6} (6 outcomes).

$$P(1\ to\ 6) = \frac{6}{6} = 1\ (100\%)$$

Part B – Probability Using Combinations

Use combinations for selecting items (when order does not matter)

$$C(n,r) = C_r^n = \binom{n}{r} = \frac{n!}{r!\,(n-r)!}$$

C(5,2) = 10 meaning there are 10 ways to choose 2 from 5.

Drawing Colored Balls

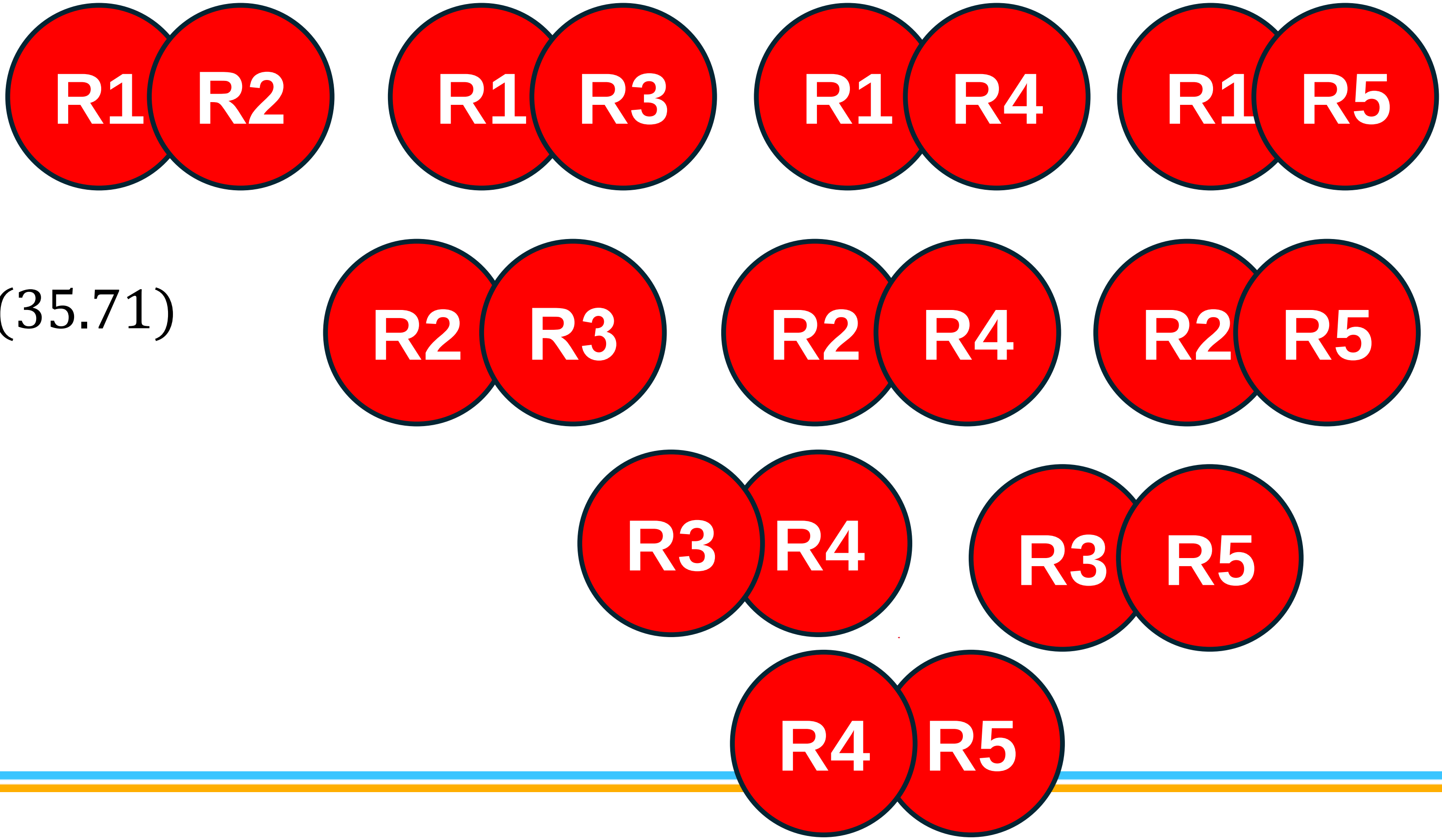
Bag: 5 red balls + 3 blue balls (Total: 8)

Event: Draw 2 balls at random. Probability both red?

Total ways: C(8,2) = 28.

Favorable C(5,2) = 10.

$$P(both\ red) = \frac{10}{28} = \frac{5}{14} \approx 0.3571\ (35.71)$$



Conditional Probability

Conditional probability is the probability of an event occurring given that another event has already occurred. We write this as

$$P(A|B),$$

which reads as "the probability of A given B."

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

The Fruit Basket Problem

Imagine a basket contains 10 pieces of fruit:

2 Red Apples
3 Green Apples
1 Green Banana
4 Yellow Bananas

Total: 10 fruits

Question 1: What is the probability of picking an apple?

P(Apple) = Number of Apples / Total Fruits

$$= (2 + 3) / 10$$
$$= 5/10$$
$$= 0.5 \text{ or } 50\%$$

There's a 50% chance of picking an apple.

The Fruit Basket Problem

Question 1: What is the probability of picking an apple?

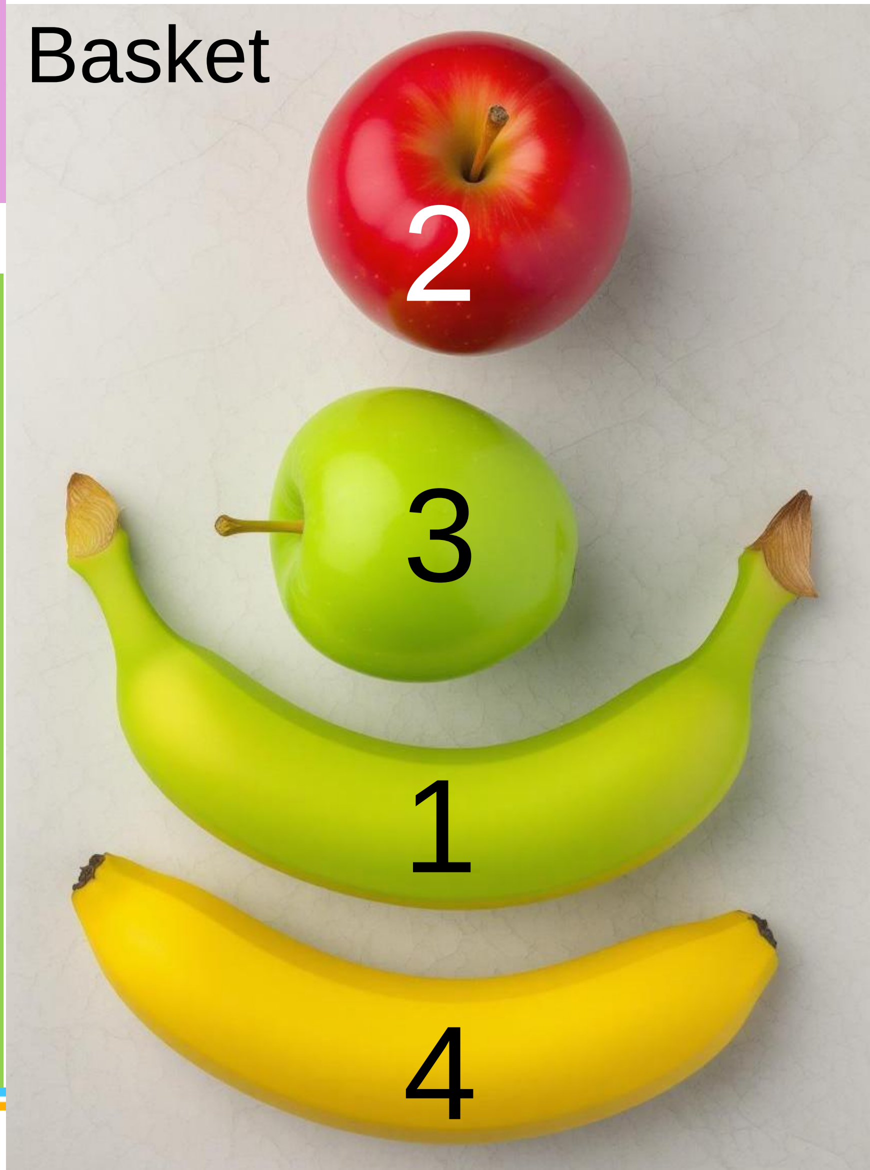
P(Apple) $= \frac{5}{10} = 0.5\ or\ 50\%$

Question 2: Now You Know It's Green! Now, what is the probability that it's an apple?

Method 1: Now we only consider GREEN fruits: 3 Green Apples and 1 Green Banana

Total Green Fruits: 4

P(Apple | Green)
$=$ Number of Green Apples / Total Green Fruits
$= 3 / 4 \qquad = 0.75\ or\ 75\%$



The Fruit Basket Problem

Question 1: What is the probability of picking an apple?

P(Apple) = 50%

Question 2: Now You Know It's Green! Now, what is the probability that it's an apple?

Method 1: P(Apple | Green) = ¾ = 75%

Method 2: Using Bayes' P(A|B) = P(A ∩ B) / P(B)

Where:

A = Picking an Apple

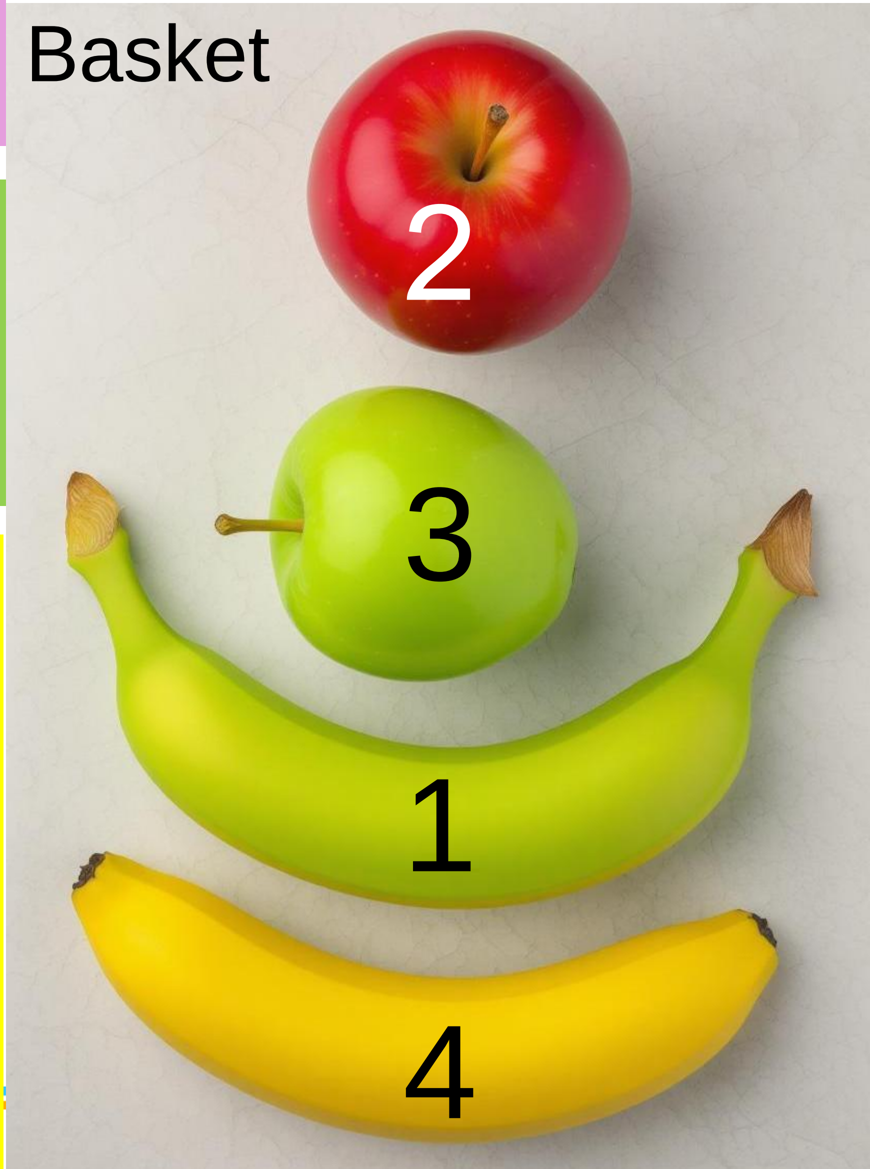
B = Picking a Green fruit

$A \cap B$ = Picking a Green Apple

$P(A \cap B) = P(\text{Green Apple}) = 3/10 = 0.3$

$P(B) = \text{P(Green fruit)} = (3+1)/10 = 4/10 = 0.4$

$P(A|B) = 0.3 / 0.4 = 3/4 = 0.75$ or 75%



The Fruit Basket Problem

Question 1: What is the probability of picking an apple?

P(Apple) = 50%

Question 2: Now You Know It's Green! Now, what is the probability that it's an apple?

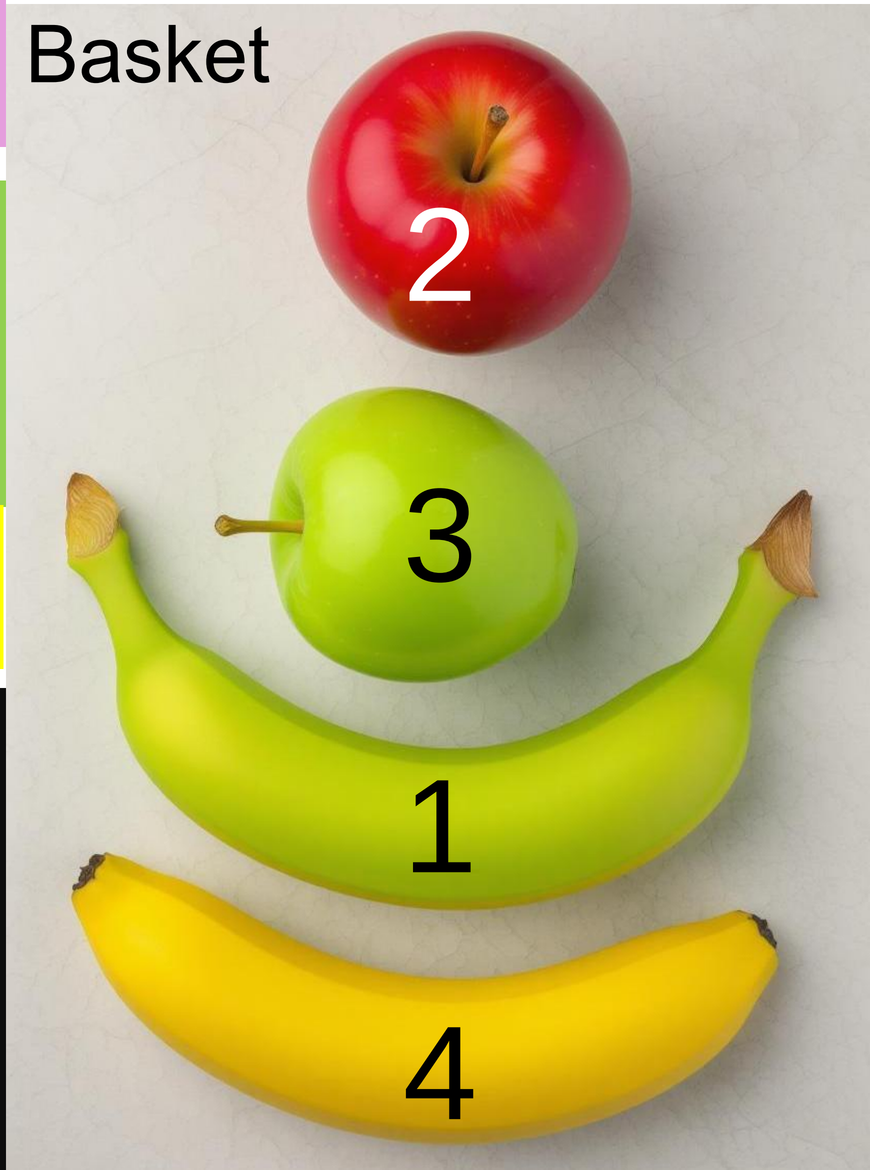
Method 1: P(Apple | Green) = $\frac{3}{4}$ = 75%

Method 2: Using Bayes' P(A|B) = P(A ∩ B) / P(B)

$P(A|B) = 0.3 / 0.4 = 3/4 = 0.75$ or 75%

Before knowing the color: P(Apple) = 50%
After knowing it's green: P(Apple | Green) = 75%

The probability **INCREASED** from 50% to 75%!
Why? Because green apples are more common than green bananas in our basket. The information "it's green" makes it MORE likely to be an apple.



The Fruit Basket Problem

Question 1: What is the probability of picking an apple?

P(Apple) = 50%

Question 2: Now You Know It's Green! Now, what is the probability that it's an apple?

P(Apple | Green) = ¾ = 75%

The Power of Conditional Probability:

New information can INCREASE probabilities (apples became more likely- from **50% to 75%)**

New information can DECREASE probabilities (bananas became less likely- from **50% to 25%)**

Example

You ask your neighbor to water a sickly plant while you are on vacation. Without water, it will die with probability 0.8; with water, it will die with probability 0.15. You are 90 percent certain that your neighbor will remember to water the plant.

a. What is the probability that the plant will be alive when you return?
b. If it is dead, what is the probability your neighbor forgot to water it?"

Given Information:

P(plant dies | no water) = 0.8
P(plant dies | watered) = 0.15
P(neighbor remembers to water) = 0.9
Therefore, P(neighbor forgets) = 0.1

Part (a)

BITS Pilani DIGITAL
Excellence made yours

a. What is the probability that the plant will be alive when you return?

We need to find P(plant survives).

Step 1: Identify the scenarios

The plant can survive in two ways:

- Neighbor remembers AND plant survives with water
- Neighbor forgets AND plant survives without water

Part (a)

a. What is the probability that the plant will be alive when you return?

Step 2: Calculate probabilities

If watered: P(survives | watered) = 1 - 0.15 = 0.85

If not watered: P(survives | no water) = 1 - 0.8 = 0.2

Step 3: Use law of total probability

$$P(survives)$$
$$= P(survives \mid watered) \times P(watered) + P(survives \mid not\ watered) \times P(not\ watered)$$
$$= (0.85 \times 0.9) + (0.2 \times 0.1)$$
$$= 0.765 + 0.02$$
$$= \mathbf{0.785\ or\ 78.5\%}$$

Part (b)

If the plant is dead, what is the probability your neighbor forgot to water it?

We need to find P(forgot | dead). This is a Bayes' theorem problem.

Step 1: Use Bayes' theorem: P(forgot | dead) = P(dead | forgot) × P(forgot) / P(dead)

Step 2: Find P(dead)

P(dead) = 1 - P(survives) = 1 - 0.785 = 0.215

Or calculated directly:

$$P(dead) = P(dead \mid watered) \times P(watered) + P(dead \mid not\ watered) \times P(not\ watered)$$

$$P(dead) = (0.15 \times 0.9) + (0.8 \times 0.1)$$

$$P(dead) = 0.135 + 0.08 = 0.215$$

Part (b)

BITS Pilani DIGITAL
Excellence made yours

If the plant is dead, what is the probability your neighbor forgot to water it?

Step 3: Calculate P(forgot | dead)

$P(\text{forgot} \mid \text{dead}) = P(\text{dead} \mid \text{forgot}) \times P(\text{forgot}) / P(\text{dead})$

$P(\text{forgot} \mid \text{dead}) = (0.8 \times 0.1) / 0.215$

$P(\text{forgot} \mid \text{dead}) = 0.08 / 0.215$

$P(\text{forgot} \mid \text{dead}) = 0.372$ or 37.2% (approximately 16/43)

Answers

a) Probability plant is alive: 0.785 or 78.5%

b) Probability neighbor forgot (given plant is dead): 0.372 or 37.2%

Even though the plant is dead, there's only a 37.2% chance the neighbor forgot. Why? Because the neighbor remembers 90% of the time, and even when watered, the plant still has a 15% chance of dying. So most of the time when the plant dies, it's actually because it died despite being watered, not because the neighbor forgot.

How to draw Conclusions/Insights

Suppose I am getting the following answer related to example shown during recorded session on Probability of cancer.

P(cancer) = 0.70

P(cancer | elevated PSA) = 0.822 or 82.2%

P(cancer | normal PSA) = 0.664 or 66.4%

What is the insight from the result?

If the test shows elevated PSA, the probability of cancer increases from 70% to 82.2%.If the test shows normal PSA, the probability of cancer decreases from 70% to 66.4%, but it's still quite high! This demonstrates why the PSA test is considered **unreliable**. Even with a normal test result, there's still a 66.4% chance of cancer in this case. The test provides some information but doesn't dramatically change the probability either way.

Thank You