DS501 STATISTICAL AND MATHEMATICAL METHODS FOR DATA SCIENCE

Dr. Muhammad Wasim

PhD, MS, M.Phil, M.Sc, MCS

Certified Data Analyst [KARACHI.AI]

Lecture Week 04

- Probability

PROBABILITY

Mutually Exclusive/Disjoint Events

- In probability theory, two events are said to be mutually exclusive if they cannot occur at the same time or simultaneously. In other words, mutually exclusive events are called disjoint events.
- If two events are considered as disjoint events, then the probability of both events occurring at the same time will be zero.
- If A and B are the two events, then the probability of probability is written by

Probability of Disjoint (or) Mutually Exclusive Event = **P (A and B) = 0**

PROBABILITY

Mutually Exclusive/Disjoint Events

- In probability, the specific addition rule is valid when two events are mutually exclusive.
- It states that the probability of either event occurring is the sum of probabilities of each event occurring.
- If A and B are said to be mutually exclusive events then the probability of an event A occurring or the probability of event B occurring is given as P(A) + P(B)

$$\mathbf{P\ (A\ or\ B) = P(A) + P(B)}$$

PROBABILITY

Mutually Exclusive/Disjoint Events

- Some of the examples of the mutually exclusive events are:
- When tossing a coin, the event of getting head and tail are mutually exclusive. Because the probability of getting head and tail simultaneously is 0.
- In a six-sided die, the events “2” and “5” are mutually exclusive. We cannot get both the events 2 and 5 at the same time when we threw one die.
- In a deck of 52 cards, drawing a red card and drawing a club are mutually exclusive events because all the clubs are black.
- If the events A and B are not mutually exclusive, the probability of getting A or B is given as

P (A or B) = P(A) + P(B) – P (A and B)

PROBABILITY

Dependent and Independent Events:

- Two events are said to be dependent if the occurrence of one event changes the probability of another event.
- Two events are said to be independent events if the probability of one event that does not affect the probability of another event.
- If two events are mutually exclusive, they are not independent and also independent events cannot be mutually exclusive.

PROBABILITY

Mutually Exclusive Events Probability Rules:

From the definition of mutually exclusive events, certain rules for the probability are concluded.

1. Addition Rule: $P (A + B) = 1$
2. Subtraction Rule: $P (A \cup B)' = 0$
3. Multiplication Rule: $P (A \cap B) = 0$

PROBABILITY

Conditional Probability for Mutually Exclusive Events:

Conditional probability is stated as the probability of an event A, given that another event B has occurred. Conditional Probability for two independent events B given A is denoted by the expression $P(B|A)$ and it is defined using the equation

$P(B|A)=P(A∩B)P(A)$

Redefine the above equation using multiplication rule: $P(A∩B)=0$

$P(B|A)=0\times P(A)$

So the conditional probability formula for mutually exclusive events is

$P(B|A)=0$

PROBABILITY

Solved Problem: (mutually exclusive events)

What is the probability of a dice showing a number 3 or number 5?

Solution: Let,

P(3) is the probability of getting a number 3

P(5) is the probability of getting a number 5

$P(3) = 1/6$ and $P(5) = 1/6$

So, $P(3 \text{ or } 5) = P(3) + P(5)$

$P(3 \text{ or } 5) = (1/6) + (1/6) = 2/6$

$P(3 \text{ or } 5) = 1/3$

Therefore, the probability of a die showing 3 or 5 is 1/3

PROBABILITY

What is Probability without Replacement or Dependent Probability?

- In some experiments, the sample space may change for the different events.
  For example, a marble may be taken from a bag with 20 marbles and then a second marble is taken **without replacing** the first marble.
- The sample space for the second event is then 19 marbles instead of 20 marbles.
- This is called probability without replacement or dependent probability.
- We can use a tree diagram to help us find the probability without replacement.

PROBABILITY

How to find the Probability without Replacement or Dependent Probability?

Step 1: Draw the Probability Tree Diagram and write the probability of each branch. (Remember that the objects are not replaced)

Step 2: Look for all the available paths (or branches) of a particular outcome.

Step 3: Multiply along the branches and add vertically to find the probability of the outcome.

PROBABILITY

Example:

A jar consists of 21 sweets. 12 are green and 9 are blue. William picked two sweets at random.

a) Draw a tree diagram to represent the experiment.

b) Find the probability that

   i) both sweets are blue.

   ii) one sweet is blue and one sweet is green.

c) William randomly took a third sweet. Find the probability that:

   i) all three sweets are green?

   ii) at least one of the sweet is blue?

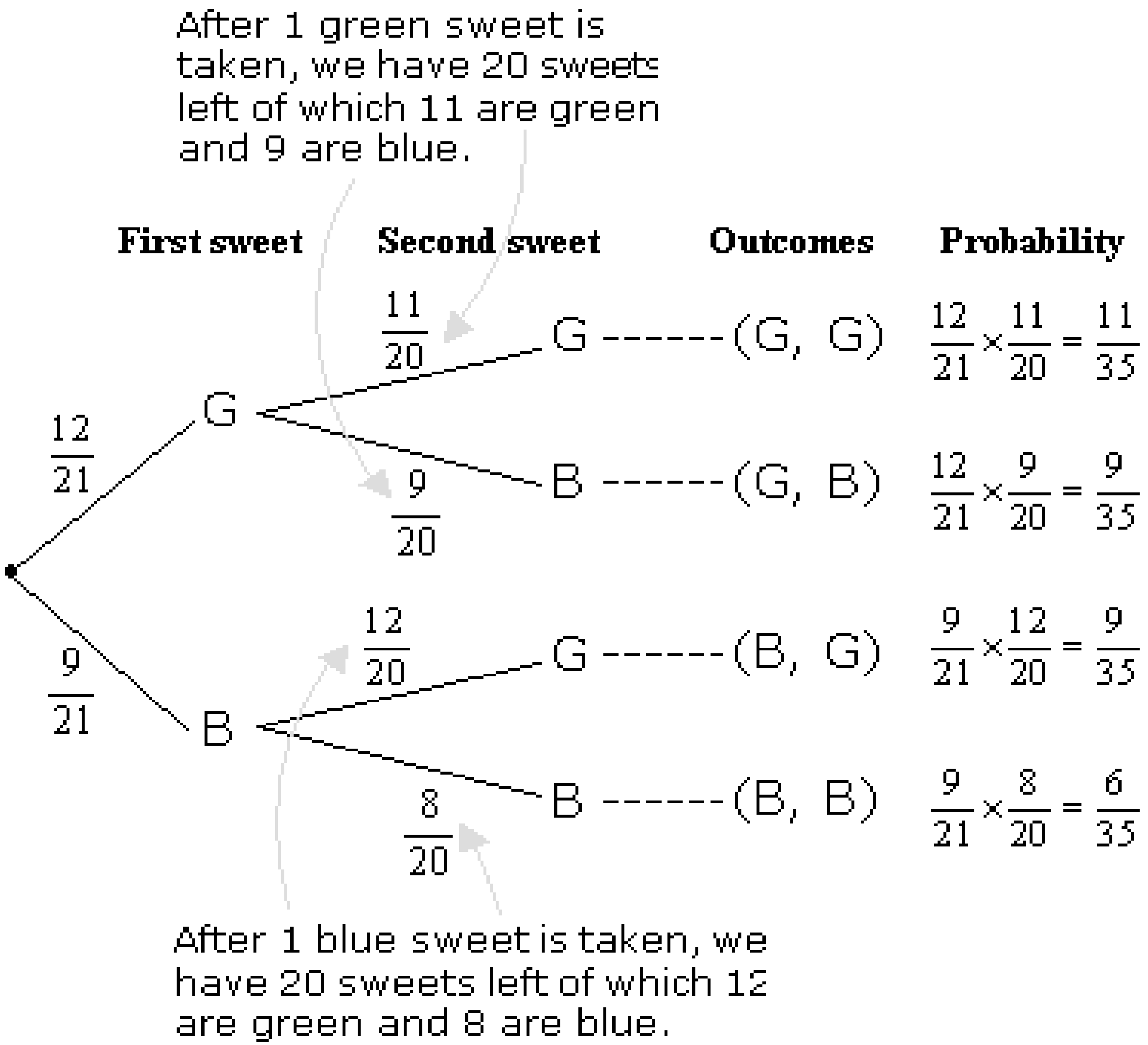
PROBABILITY

Solution:

a) Although both sweets were Taken together it is similar to picking one sweet and then the second sweet without replacing the first sweet.

Check that the probabilities in the last column add up to 1.

$$\frac{11}{35}+\frac{9}{35}+\frac{9}{35}+\frac{6}{35}=1$$

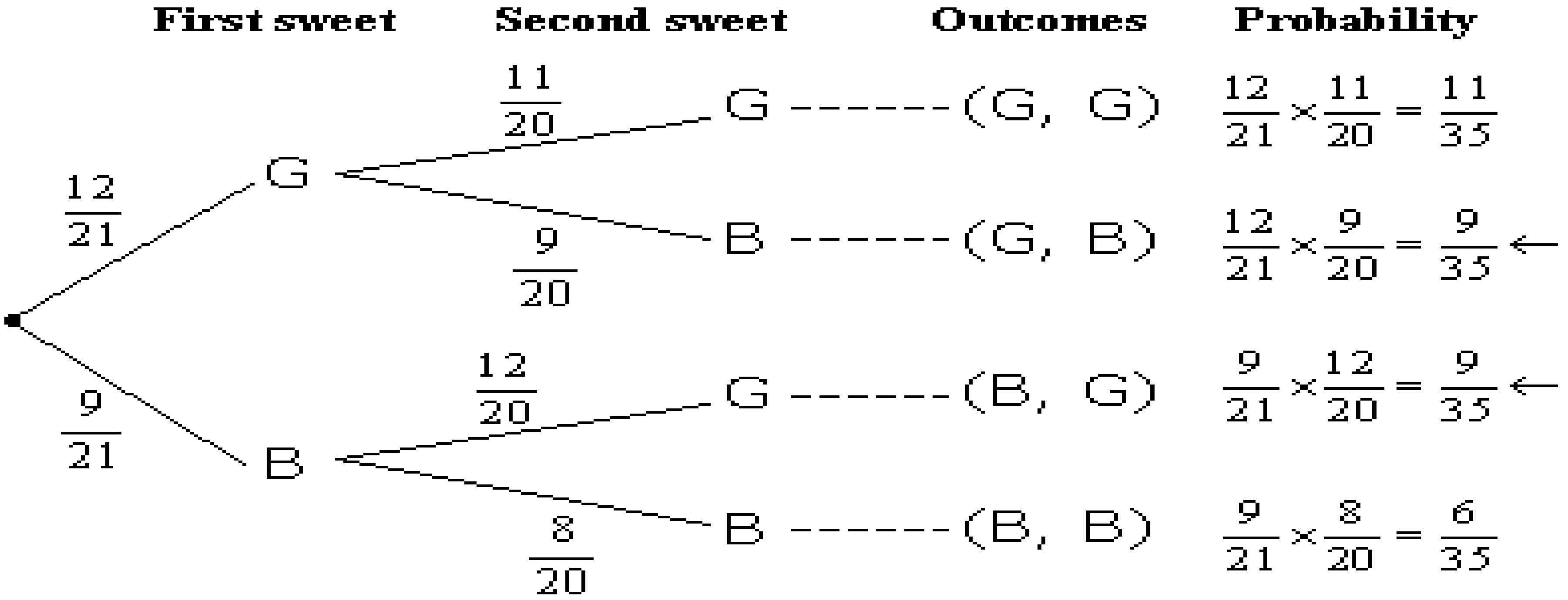


PROBABILITY

b) i)

P(both sweets are blue) = P(B, B)

$$= \frac{9}{21} \times \frac{8}{20} = \frac{6}{35}$$



PROBABILITY

ii)

P(one sweet is blue and one sweet is green) = P(G, B) or P(B, G)

$$= \frac{9}{35} + \frac{9}{35} = \frac{18}{35}$$

c) i) P(all three sweets are green) = P(G, G, G)

$$= \frac{12}{21} \times \frac{11}{20} \times \frac{10}{19} = \frac{22}{133}$$

ii) P(at least 1 sweet is blue) = 1 – P(all three sweets are green)

$$= 1 - \frac{22}{133}$$

$$= \frac{111}{133}$$

PROBABILITY

On a library shelf, three geometry and five algebra books. Books are not replaced after someone borrow it. If two books are taken then what is the probability that first is of geometry and other of algebra.

Solution:

The probability of first book to be of geometry = $3/8$

The probability of second book to be of algebra = $5/7$

As in total number of books one is already taken out, the total has become 7.

Probability of both events occurring = $3/8 \times 5/7 = 15/56$

PROBABILITY

Addition Rule

1. Mutually Exclusive/disjoint

   $P(A\cup B)=P(A)+P(B)$

2. Not Mutually Exclusive/joint

   $P(A\cup B)=P(A)+P(B)-P(A\cap B)$

Multiplication Rule

1. Independent Events

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

2. Dependent Events

   $P(A\cap B)=P(A)\times P(B/A)$

PROBABILITY

Examples of Not Mutually Exclusive

Ex-1:

total cards = 52

King = 4

Heart = 13

Common = 1

Find P(King of Heart)?

P(King or Heart)= $\frac{4}{52} + \frac{13}{52} - \frac{1}{52}$

$$= \frac{16}{52}$$

Examples of Mutually Exclusive

Ex-1:

total cards = 52

King = 4

Ace = 4

Common = 0

Find P(Ace of Heart)?

P(Ace or Heart)= $\frac{4}{52} + \frac{4}{52} = \frac{8}{52}$

$$= \frac{2}{13}$$

PROBABILITY

Examples of Not Mutually Exclusive

Ex-2:

total cards = 52

Queen = 4

Black = 26

Common = 2

Find P(Queen or Black)?

$$P(\text{Queen or Black}) = \frac{4}{52} + \frac{26}{52} - \frac{2}{52}$$

$$= \frac{28}{52}$$

Examples of Mutually Exclusive

Ex-2:

total cards = 52

Heart = 13

Spade = 13

Common = 0

Find P(Heart or Spade)?

$$P(\text{Heart or Spade}) = \frac{13}{52} + \frac{13}{52} = \frac{26}{52}$$

$$= \frac{1}{2}$$

PROBABILITY

Additive Rules

$P(A \cup B)$

$P(A+B)$

P(A or B)

P(either A or B)

P(at least 1)

Additive Rules

$P(A \cap B)$

$P(AB)$

P(A as well as B)

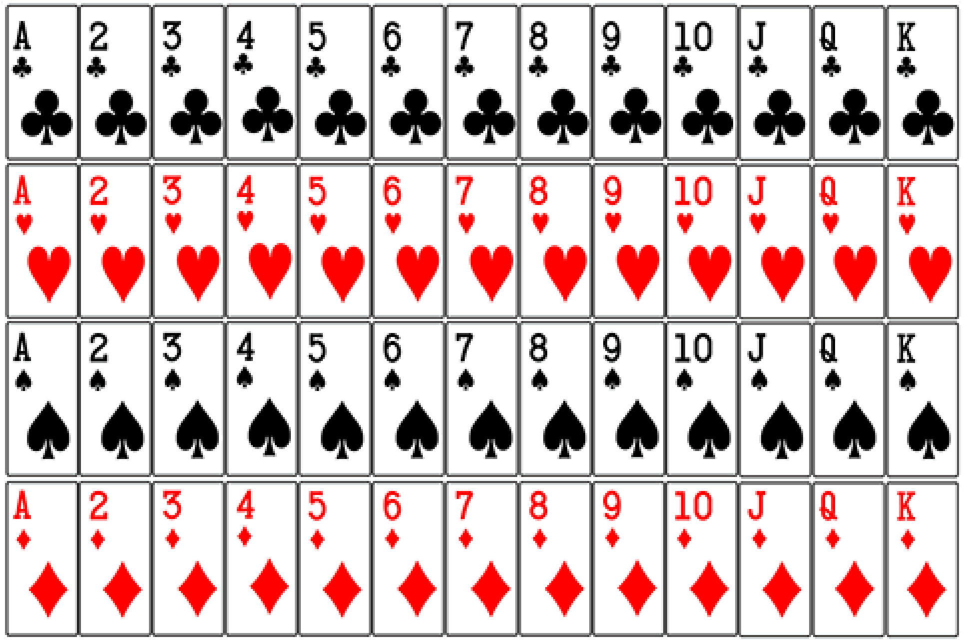
P(Both)

Dependent Events

$$P(A \cap B) = P(A) \times P(B/A)$$

Where $P(B/A)$ mean prob. of event B, when event A has already occurred.

PROBABILITY



Cards (52)

Spade	Club	Diamond	Heart
1 King	1 King	1 King	1 King
1 Queen	1 Queen	1 Queen	1 Queen
1 Jack	1 Jack	1 Jack	1 Jack
1 Ace	1 Ace	1 Ace	1 Ace
2-10 Cards	2-10 Cards	2-10 Cards	2-10 Cards
Total = 13	**Total** = 13	**Total** = 13	**Total** = 13

PROBABILITY

Probability of Independent Events (with replacement)

Two coins are tossed together, what is the probability of head on first and tail on second.

$P(H,T) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

If head on first coin ?????

Similarly,

P(Ace and King) =

$4/52 \times 4/52 = 1/169$

Probability of Dependent Events (without replacement)

There is one packet of card. Draw two cards (without replacement) what is the probability that Ace and King occurs?

$$P(\text{Ace \& King}) = \underbrace{4/52}_{P(A)} \times \underbrace{4/51}_{P(B/A)} = 16/$$

If with replacement ?????

PROBABILITY

The definition of independence:

Events A and B are independent if and only if

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

If A and B are independent:

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

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

PROBABILITY

When A and B have non-zero probabilities of occurring, these 3 statements are equivalent:

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

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

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

If they are true, then A and B are independent.

If they are false, then A and B are not independent.

PROBABILITY

Suppose $P(A) = 0.70$, $P(B) = 0.20$, and A and B are independent.

What is $P(A \cap B)$?

$$P(A \cap B) = 0.70 \cdot 0.20 = 0.14$$

$$P(A|B) = 0.70 \qquad\qquad P(B|A) = 0.20$$

PROBABILITY

Suppose $P(A) = 0.40$, $P(B) = 0.80$, and $P(A \cap B) = 0.36$.

Are A and B independent?

$$P(A)P(B) = 0.40 \cdot 0.80 = 0.32 \neq P(A \cap B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.36}{0.80} = 0.45 \neq P(A)$$

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.36}{0.40} = 0.90 \neq P(B)$$

A and B are not independent

PROBABILITY

Suppose a balanced six-sided die is about to be rolled, and we define the following events:

$$S = \{1,2,3,4,5,6\}$$

$$A = \{1, 2\} \qquad P(A) = \tfrac{1}{3}$$

$$B = \{2, 3, 4, 5, 6\} \qquad P(B) = \tfrac{5}{6} \qquad P(A)P(B) = \tfrac{1}{3} \times \tfrac{5}{6} = \tfrac{5}{18}$$

$$A \cap B = \{2\} \qquad P(A \cap B) = \tfrac{1}{6} \neq P(A)P(B)$$

A and B are not independent

$$P(A|B) = \tfrac{1}{5} \qquad P(B|A) = \tfrac{1}{2}$$

PROBABILITY

Suppose a balanced six-sided die is about to be rolled, and we define the following events:

$A = \{1, 2\}$ $\quad P(A) = \frac{1}{3}$

$C = \{2, 4, 6\}$ $\quad P(C) = \frac{1}{2}$ $\quad \overline{C} = \{1, 3, 5\}$

$$P(A|C) = \frac{1}{3} \qquad P(A|\overline{C}) = \frac{1}{3}$$

$P(A|C) = P(A)$ $\quad A$ and C are independent

$P(A|\overline{C}) = P(A)$ $\quad A$ and $\overline{C}$ are independent

CONDITIONAL PROBABILITY

Recall the conditional probability formula:

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

(provided $P(B) > 0$)

Suppose:

$$P(A) = 0.34,\ P(B) = 0.50,\ P(A \cup B) = 0.70$$

What is $P(A|B)$?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.14}{0.50} = 0.28$$

The addition rule:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$0.70 = 0.34 + 0.50 - P(A \cap B)$$

$$\implies P(A \cap B) = 0.14$$

CONDITIONAL PROBABILITY

Suppose we are about to roll a balanced six-sided die once.

Consider the events:

$A = \{1, 4\}$ $\qquad$ $S = \{1, 2, 3, 4, 5, 6\}$

$B = \{2, 3, 4, 6\}$

$C = \{1, 3, 5\}$ ← Reduced sample space

What is $P(A \cup B | C)$? $\qquad$ $P(A \cup B | C) = \frac{2}{3}$

$A \cup B = \{1, 2, 3, 4, 6\}$

CONDITIONAL PROBABILITY

A study investigated causes of sudden (non-violent) deaths in western regions of Paris, France.

A sample of 523 such deaths revealed the following.

	Cardiovascular	Cerebral	Respiratory	Other	Total
Males	264	38	36	21	359
Females	89	27	29	19	164
Total	353	65	65	40	523

Suppose one of these cases is randomly selected.

What is the probability the person was female?

$$P(F) = \frac{164}{523} \approx 0.31$$

CONDITIONAL PROBABILITY

	Cardiovascular	Cerebral	Respiratory	Other	Total
Males	264	38	36	21	359
Females	89	27	29	19	164
Total	353	65	65	40	523

Suppose one of these cases is randomly selected.

Given the cause was cardiovascular in nature, what is the probability the person was female?

$$P(F|CV) = \frac{89}{353}$$

$$P(F|CV) = \frac{P(F \cap CV)}{P(CV)} = \frac{89/\cancel{523}}{353/\cancel{523}} = \frac{89}{353}$$

CONDITIONAL PROBABILITY

	Cardiovascular	Cerebral	Respiratory	Other	Total
Males	264	38	36	21	359
Females	89	27	29	19	164
Total	353	65	65	40	523

Suppose one of these cases is randomly selected.

Given the person was female, what is the probability the cause was cardiovascular in nature?

$$P(CV|F) = \frac{89}{164}$$

CONDITIONAL PROBABILITY

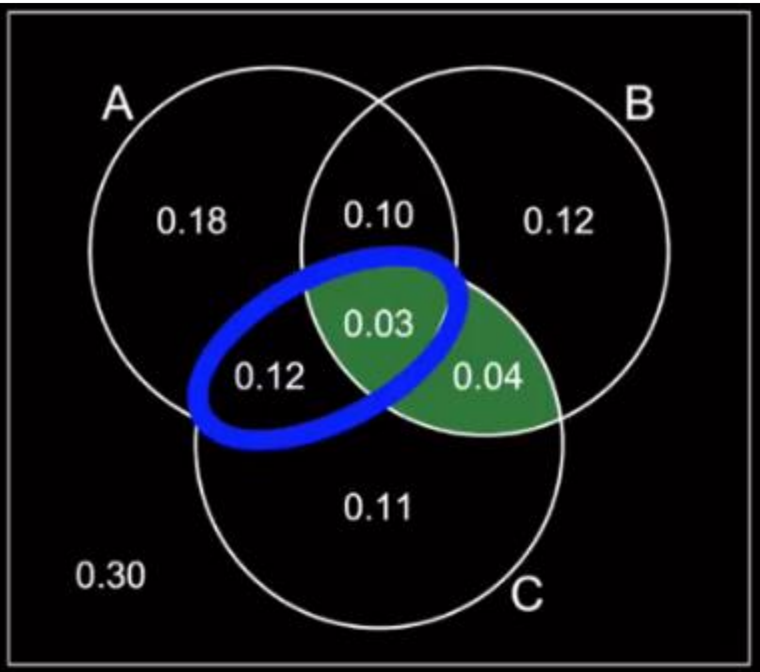
	Cardiovascular	Cerebral	Respiratory	Other	Total
Males	264	38	36	21	359
Females	89	27	29	19	164
Total	353	65	65	40	523

Suppose one of these cases is randomly selected.

Given the cause was cerebral or respiratory in nature, what is the probability the person was male?

$$P(M|C \cup R) = \frac{38+36}{65+65} = \frac{74}{130} \approx 0.57$$

CONDITIONAL PROBABILITY

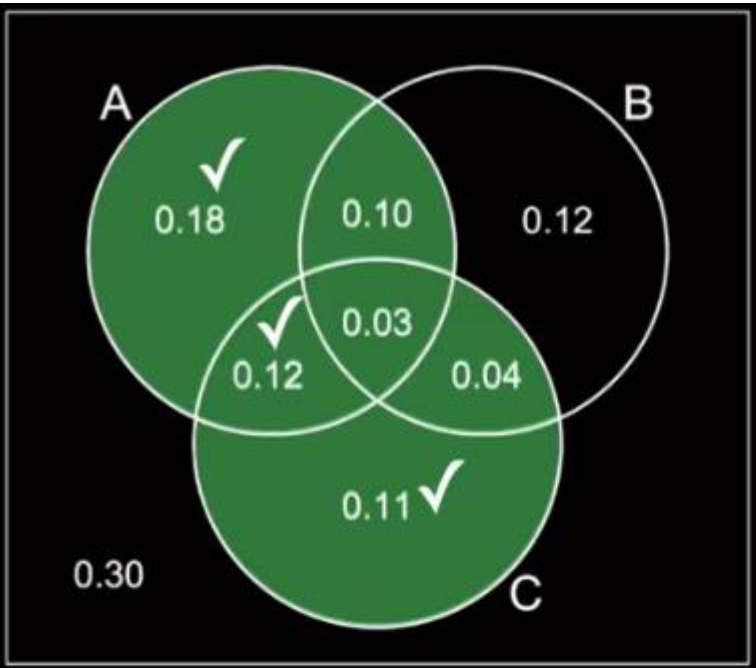


$P(A) = 0.43$
$P(B) = 0.29$
$P(C) = 0.30$
$P(A \cap B) = 0.13$
$P(A \cap C) = 0.15$
$P(B \cap C) = 0.07$
$P(A \cap B \cap C) = 0.03$

What is $P(A \cap C | B \cap C)$?

$$P(A \cap C | B \cap C) = \frac{0.03}{0.03 + 0.04} = \frac{P((A \cap C) \cap (B \cap C))}{P(B \cap C)}$$

CONDITIONAL PROBABILITY



$$P(A) = 0.43$$
$$P(B) = 0.29$$
$$P(C) = 0.30$$
$$P(A \cap B) = 0.13$$
$$P(A \cap C) = 0.15$$
$$P(B \cap C) = 0.07$$
$$P(A \cap B \cap C) = 0.03$$

What is $P(\overline{B}|A \cup C)$?

$$P(\overline{B}|A \cup C) = \frac{0.18+0.12+0.11}{0.18+0.12+0.11+0.10+0.03+0.04} = \frac{P(\overline{B} \cap (A \cup C))}{P(A \cup C)}$$

CONDITIONAL PROBABILITY

A Canadian adult is randomly selected.

A: The person is over 6 feet (183 cm) tall
B: The person is male

Which one of the following is true?

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

$$P(A|B) > P(A) \iff P(B|A) > P(B)$$

$$P(A|B) < P(A)$$

CONDITIONAL PROBABILITY

Suppose A is a subset of B, and $P(A) > 0$.

What can be said of $P(A|B)$ and $P(B|A)$?

$$P(B|A) = 1$$

$$0 < P(A|B) \leq 1$$