INDIAN INSTITUTE OF TECHNOLOGY KANPUR

Probability Models

Prof. Abhinava Tripathi

Introduction

Introduction

- Probability models empower statisticians to draw inferences from sample data
- The average salary of a data scientist in India is INR 25 Lakhs plus-minus INR 10 Lakh , with 90% of confidence
- How do we get to these conclusions?
- Do we extensively examine the entire population of data scientists

Random Experiments and Sample Space

Random Experiments

A random experiment is a process that generates well-defined experimental outcomes. On any single repetition or trial, the outcome that occurs is determined completely by chance

1. The experimental outcomes are well defined, and in many cases can even be listed prior to conducting the experiment
2. On any single repetition or trial of the experiment, one and only one of the possible experimental outcomes will occur
3. The experimental outcome that occurs on any trial is determined solely by chance

Random Experiments and Sample Space

- Whether it’s cricket, football or soccer or basketball, tossing a coin is a very common feature
- In cricket, a coin might be tossed to decide which team gets to bat first or ball first

Random Experiments and Sample Space

- Let’s take an example of a fair coin
- A fair coin is a coin that’ll have the structure or the make identical for both the heads and tails side of the coin

- When you toss it, the chance of heads is the same as the chance of tails
- The captain of the Indian team calls for heads.
- What do you think will be the chance that the captain wins?

Random Experiments and Sample Space

- You can have only 2 outcomes: heads or tails
- The chance that he wins is computed like this:

- Number of favorable outcome / Total outcomes = 1(heads)/2(heads and tails) = 0.5
- The chances of something happening is known as the probability

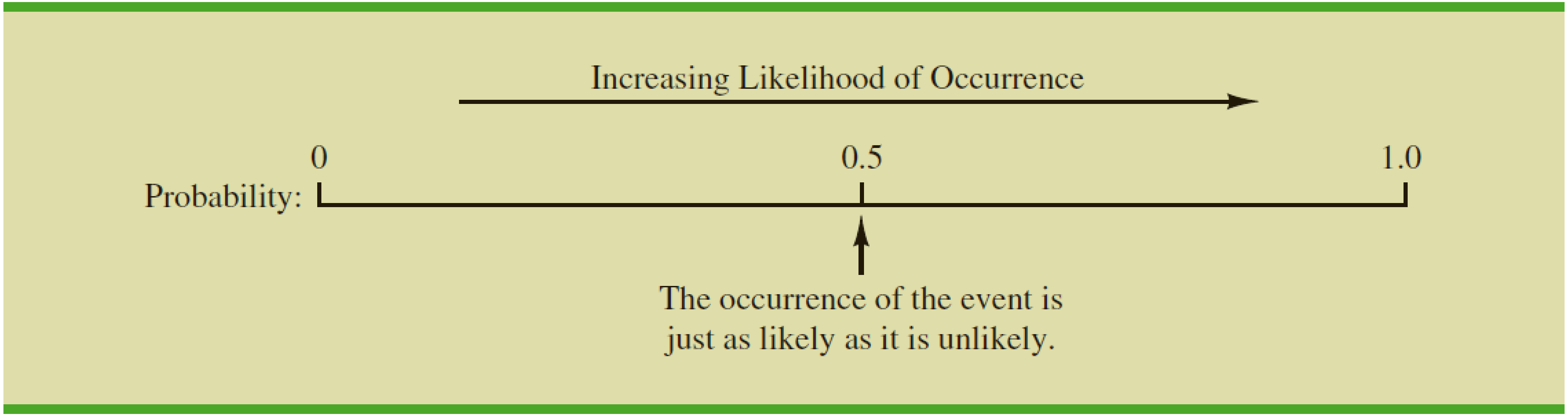
Random Experiments and Sample Space

- Probability is really the language of uncertainty
- The coin toss game here is known as an experiment
- The sample space for a random experiment is the set of all experimental outcomes
- The set of possible outcomes {Head, Tails} is known as the sample space
- Sample space can be mathematically represented as Ω (omega) or S = {Heads, Tails}
- In our example, the probability of the captain winning is known as the event

Probability Events

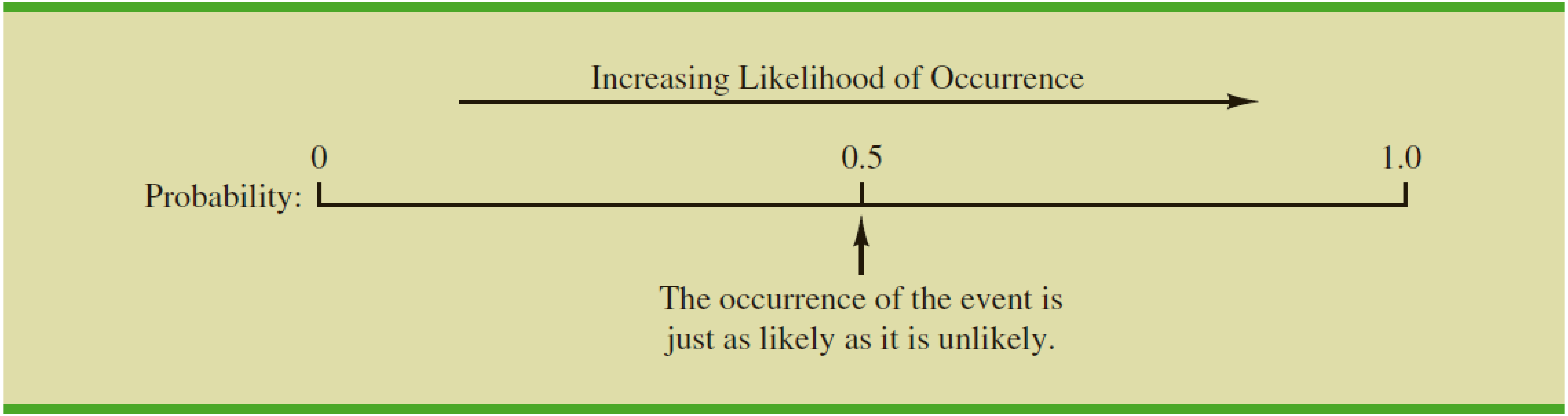
Probability Events

- Probability is a numerical measure of the likelihood that an event will occur
- Thus, probabilities can be used as measures of the degree of uncertainty associated with an event



Probability Events

- Probability values are always assigned on a scale from 0 to 1
- A probability near zero indicates an event is unlikely to occur; a probability near 1 indicates an event is almost



Probability Events

- Let us consider the example of rolling dice with 6 faces
- What is the random experiment here?
- Sample space: S = {1, 2, 3, 4, 5, 6}
- Number of favorable outcome / Total outcomes = 1(getting a 1)/ 6(1, 2, 3, 4, 5, 6) = ⅙
- What about the probability of getting an even number, that is, 2, 4, 6: 3/6 = 0.5.

Probability Events

- In this example, Rolling a dice is a random experiment
- ‘Rolling a one’ or the ‘Rolling an even number’ are the events for which we find the probability
- A random experiment is a process that leads to one of several possible outcomes
- The probability formula we discussed may not apply at all the places

Probability Events

- Let’s consider the marks scored out of 100 by a student in a math test
- What can be the outcomes of the test: any number between 0 and 100
- If evaluation is in whole numbers, the sample space becomes the set of whole numbers between 0 and 100, which is {0, 1, 2, ….,100}
- What is the probability of the event of “Student scoring 80 in a math test”
- According to our formula: Number of favorable outcomes / Total outcomes = 1(getting 80)/ 101(0 to 100) = 1/101
- Does it look right?

Probability Events

- Not all experiments are designed to make all outcomes equally likely
- You can have a test that is designed such that people get at least a passing grade
- How would we calculate probability in that scenario: look past data
- For example, if there were 50 students in the class who took the math test
  - Out of which 10 students scored 80
  - 10 students scored 90 and the rest scored 60
- Then what is the probability of a student scoring 80 marks?

Probability Events

- Incorrect to use the formula number of favorable outcomes divided by Total outcomes = 1(getting 80)/ 3(60, 80,90) = ⅓
- The formula changes here as follows: Frequency of favourable outcome / Total frequency of all outcomes
- Frequency is essentially the number of times an experiment is repeated with the outcome
- In our case, the experiment is repeated for 50 students and the Frequency of favorable outcomes is 10: probability is 10/50= 0.20

Probability Events

- The formula: “Number of favorable outcome divided by Total outcomes” worked for the coin toss and rolling a die experiments because different outcomes of experiments do not depend on what happened in the past
- The different outcomes the case of coin toss and in case of die roll are equally likely
- However, this is usually not the case with most experiments
- That is not the case with the Math exam, which is designed in a way that one outcome, say, scoring 0, is not as likely as scoring a 100, or scoring 60 or 80

Probability Definition

Probability Definition

- We calculate probability when we want to understand the chances of an event happening in a random experiment
- Probability is associated with events that have unsure outcomes
- It might or might not rain on a given day; hence there are two possible events
- Probability = Frequency of favorable outcomes / Total frequency of all outcomes
- Probabilities are always between 0 and 1

Probability Definition

- Probability of getting 7 in a die roll can be zero, that is minimum
- Conversely, if the event is defined as getting 1, 2, 3, 4, 5, or 6, then the probability can be 1, that is maximum
- Hence the probability of an event in a random experiment can only be between 0 and 1
- A probability of 0 means a certainty that an event will not occur
- A probability of 1 means that an event is certain to occur

Probability Definition

- Probability is also represented as percentages: 0 corresponds to 0% and 1 corresponds to 100%
- Often, the news on weather channel tells you that there is a 20% or 30% chance of rainfall on a given day
- An investment in financial markets may have a low or high probability of providing high returns
- Similar probabilities of success and failure exists in medicine, business, and life itself

Combining two or more experiments

Combining two or more experiments

- Let's now look at some more complex experiments
- Let us flip two coins together: what would be the sample space here
- The experiment of tossing two coins can be thought of as a two-step experiment in which step 1 is the tossing of the first coin and step 2 is the tossing of the second coin
- If we use H to denote head and T to denote tail, then then our sample space becomes: S= HH, HT, TH, and TT

Combining two or more experiments

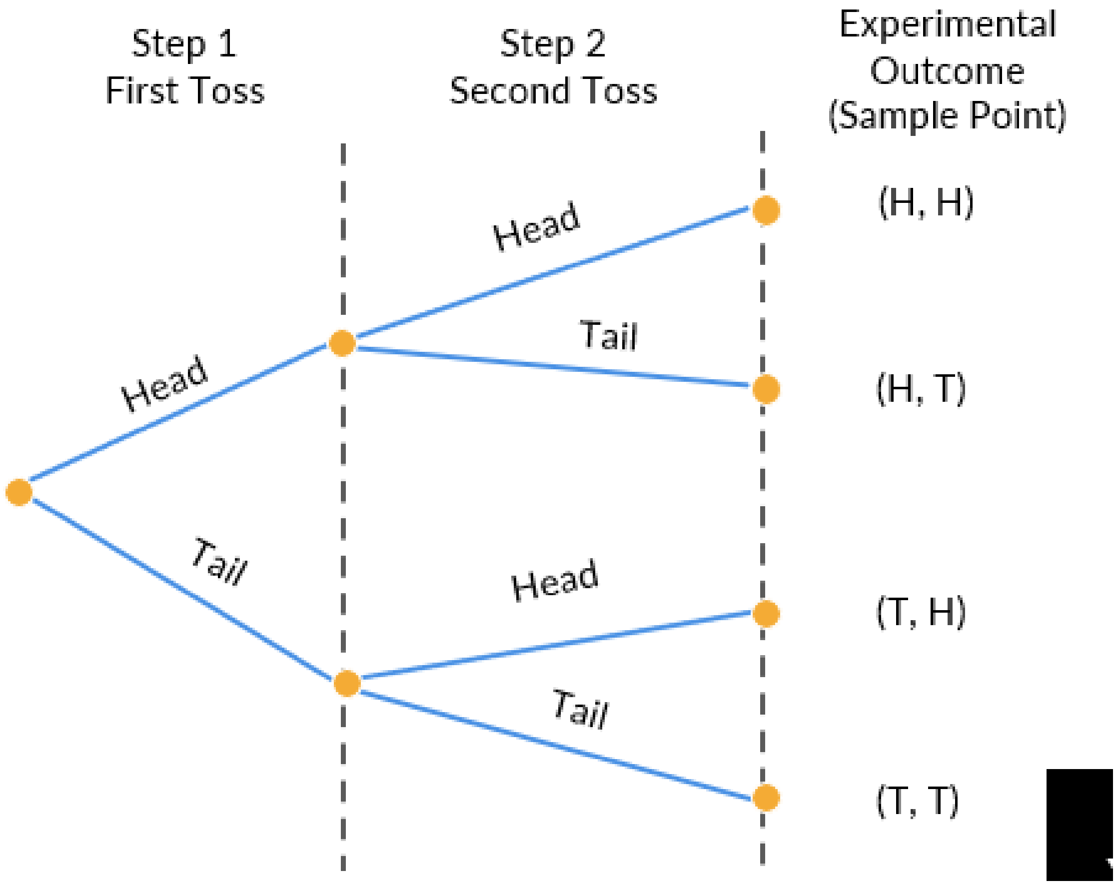
- In this experiment, the total number of possible outcomes are 2x2=4
- Here 2 (H, T) is the sample space of first and second coins
- Let us generalize this with a random event in k-steps, where the possible outcomes are n1, n2,…nk
- In such multi-step experiments, the total number of experimental outcomes are: S = n1 x n2 x …xnk
- Such events can be shown in the form of tree diagram

Combining two or more experiments

- Tree diagram for a simple coin toss game is provided here

Tree-Diagram example

- Consider tossing a balanced coin twice



Combining two or more experiments

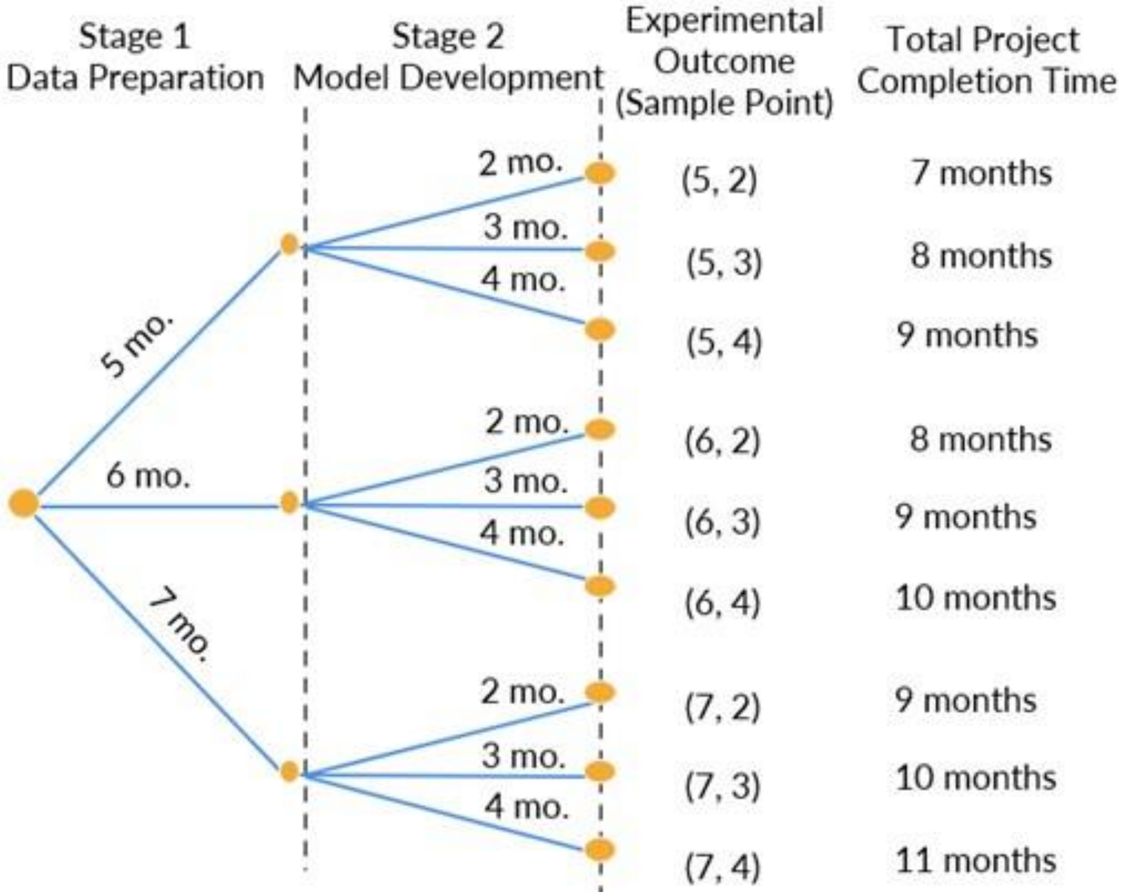
- The experiments need not be only a single step, they can be multistep
- The sample outcome can be a complex representation of the multistep experiment
- Consider yourself as data scientist on a two stage project: (1) Data Preparation; (2) Model Development
- How to estimate the possible completion time for that project?

Combining two or more experiments

- Let us make use of past data from similar projects
- Past data suggests completion time of **Data Preparation** stage as 5, 6, or 7 months, equally likely
- For model **Development Stage**, the same is 2, 3, or 4 months, equally likely
- What is the random experiment here?
- What are the possible different events here?
- Let us try to find the probabilities of some of these different events

Combining two or more experiments

- Let us understand this with the tree diagram
- The number of sample outcomes = (number of outcomes in step 1) into (number of outcomes in step 2) = 3 * 3 = 9
- What is the probability of completing data preparation in 5-Months: 3/9=1/3

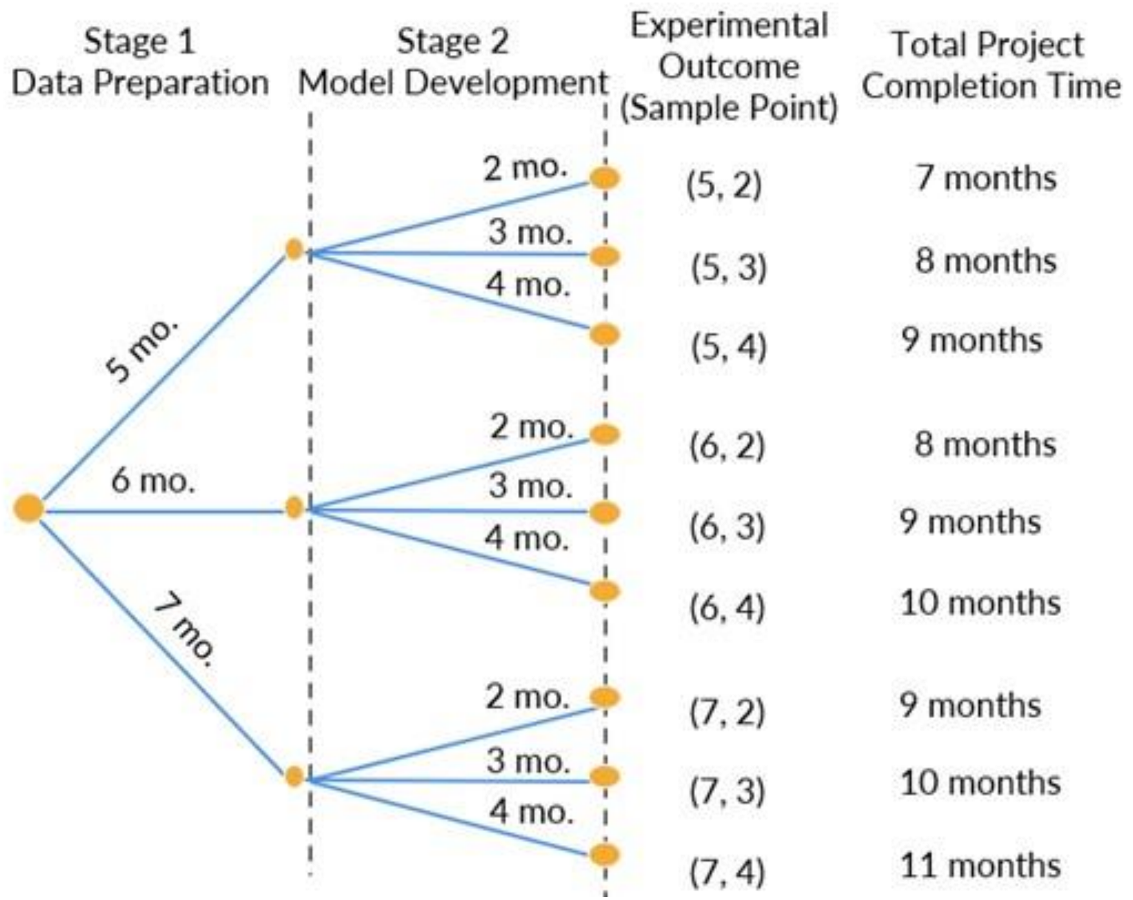


Combining two or more experiments

- For this question, you do not need the information about the next stage

- What is the probability of completing model development in exactly 3 months: 1/3

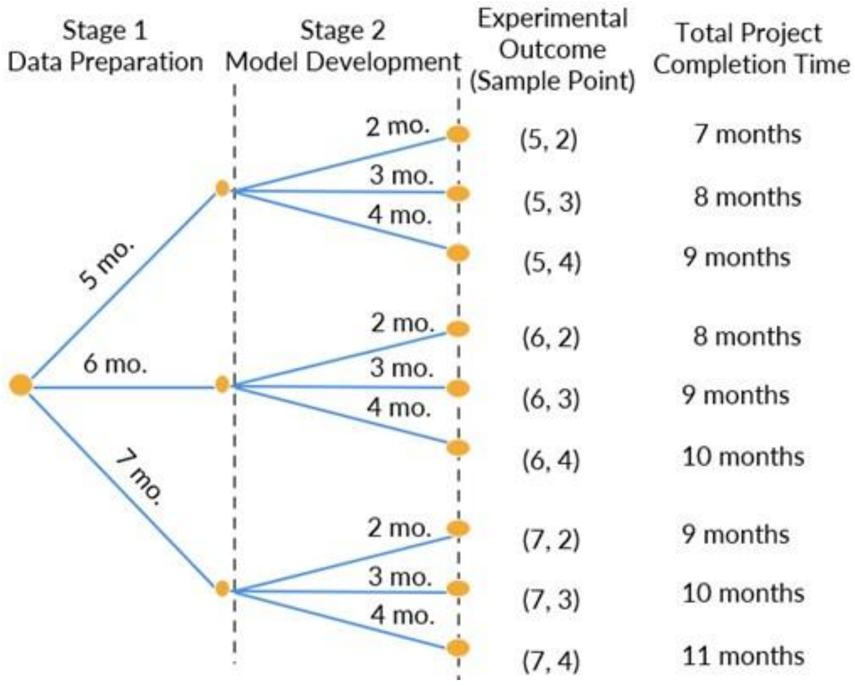
- What is the probability that project will be completed in 8-months: 2/9



Types of Events

Types of Events

- If we are estimating the probability that project will be finished in 9-months or less
- There are 6 out of the possible 9 outcomes where it happens: hence the probability = 6/9= 66.67%
- What is the probability that the work will be completed in more than 9-months; this question is also some how related to previous question



Types of Events

- The probability of completing the project in more than 9-months is 3/9=1/3
- This is so because out of the 9 possible outcomes, 3 satisfy this condition
- In the remaining six outcomes, the work will be completed in 9-months or less; thus, the probability 1-1/3= 2/3
- If we call these events A and B respectively, then A and B are complementary events
- That B can be referred to as 'Ac' or complementary to A

Types of Events

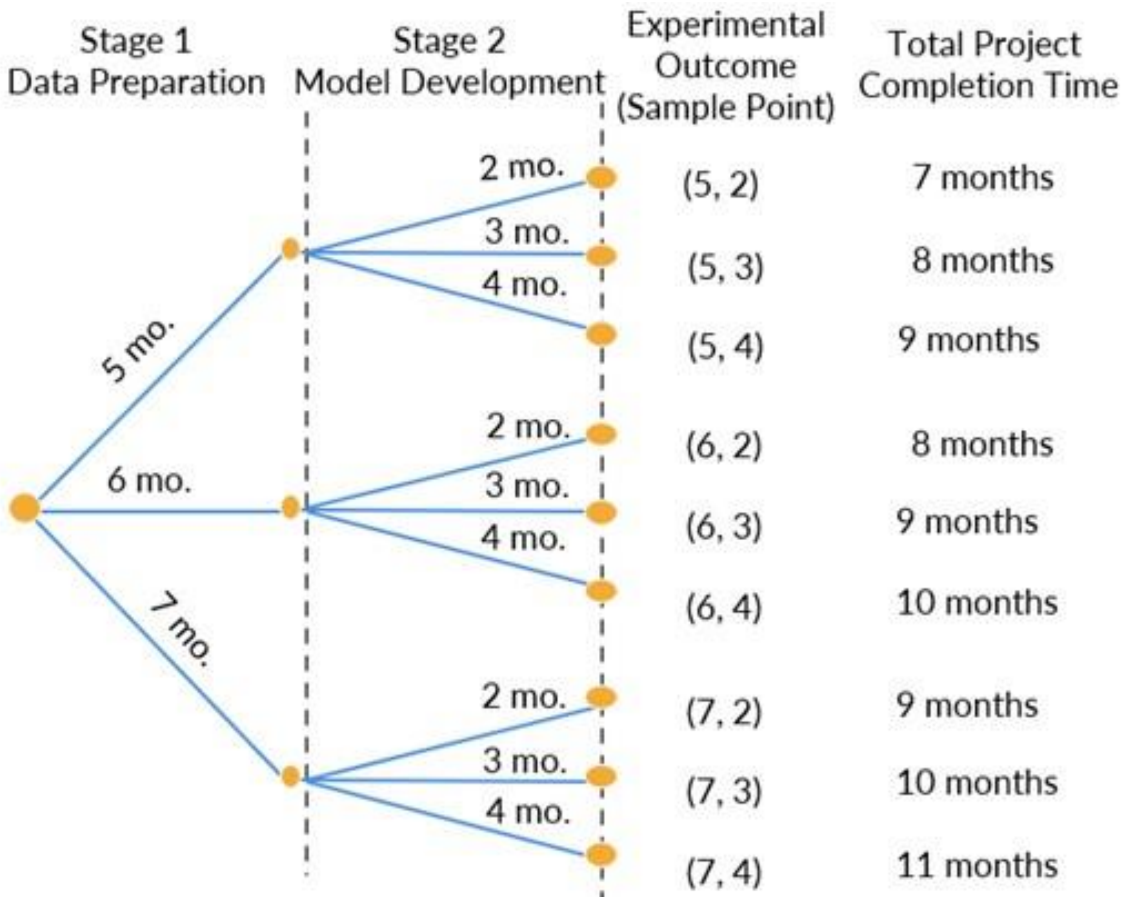
- This can very well be represented as a Venn diagram
- The complement of event A, that is, event Ac, which in our case is Event B

- Event B, is represented by the area that is not in A, which is the white portion within the rectangle
- Since the probability of complete sample space is 1: P(A)+P(Ac)=1
- This takes us to the basic rule of probability, that is, the sum of the probabilities of all the events always add up to 1

Intersection Events

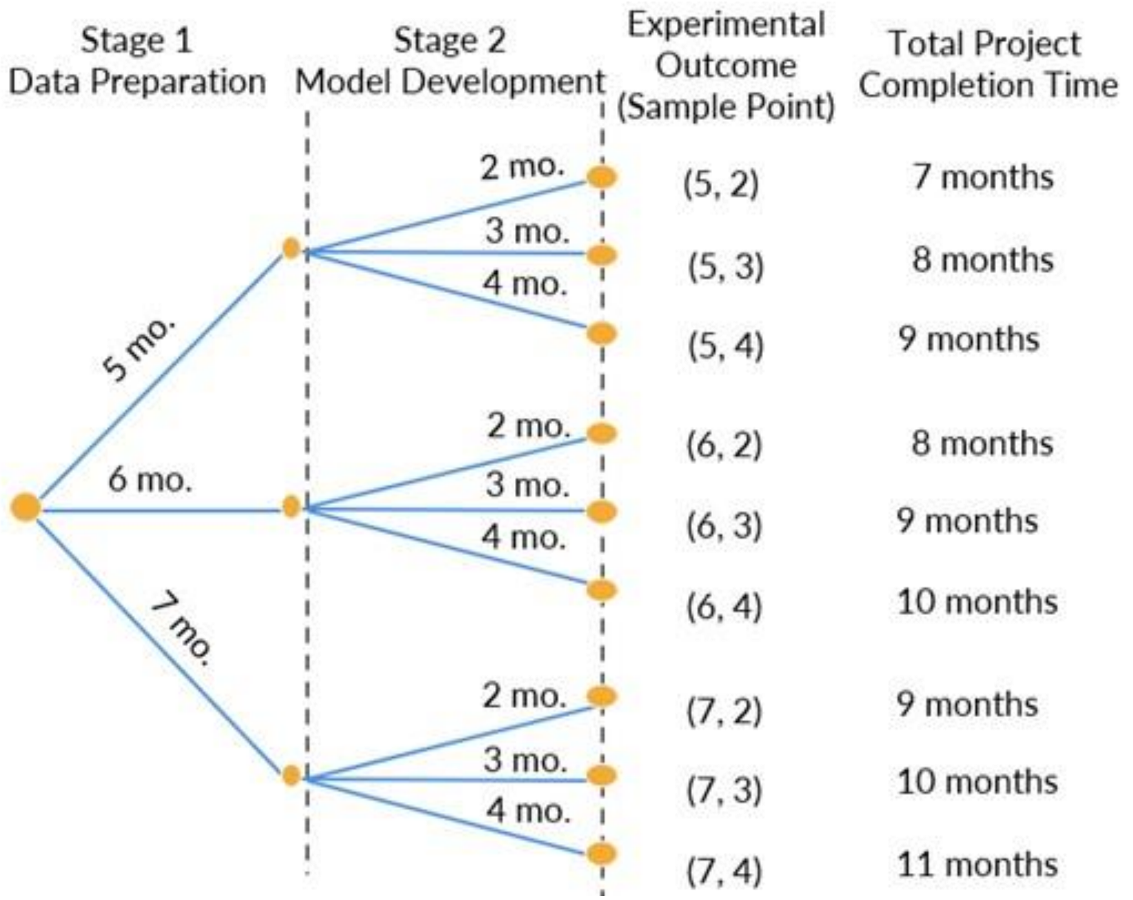
Intersection Events

- An event is a subset of all the possible outcomes of a random experiment
- Within an event, a simple event is with one possible outcome and a compound event is a combination of two or more simple events



Intersection Events

- Our machine learning projects, which is a multi-step experiment, a compound event like data preparation can be broken into 3 possible simple events
- This is how we compute the probabilities for compound events

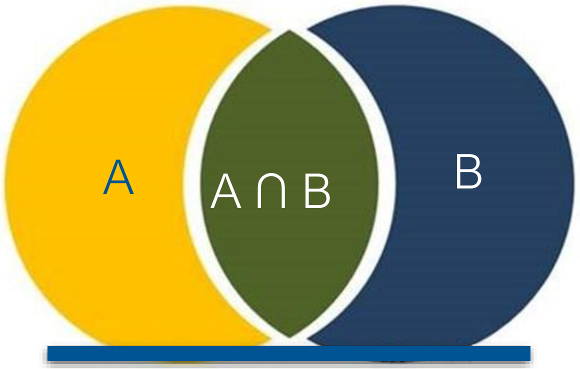


Intersection Events

- In the example of rolling a dice, the event, ‘Rolling a one’ is an example of a simple event
- The event, ‘Rolling an odd number’ will be an example of a compound event
- It can be broken down into three simple events, ‘Rolling a one’, ‘Rolling a three’ and ‘Rolling a five’
- Consider the following event: Completing the Data preparation Stage in 5 months and the Model Development stage in 2

Intersection Events

- This is a composite event with intersection of two events



- Event A: completing the Data preparation Stage in 5 months
- Event B: completing the Model Development stage in 2 months
- For Event A, the sample outcomes are (5, 2), (5, 3) and (5, 4)
- For Event B, the sample outcomes are (5, 2), (6, 2) and (7, 2)
- The intersection or the common of these 2 events is (5, 2): $A \cap B$

Intersection Events

- Hence, the probability of Completing the Data preparation Stage in 5 months and the Model Development stage in 2= 1/9
- Consider another example, where two dice are rolled simultaneously
- How many possible outcomes in the sample space: 6x6
- Consider event A where first die has number 1 and event B, where second dice has the number 5

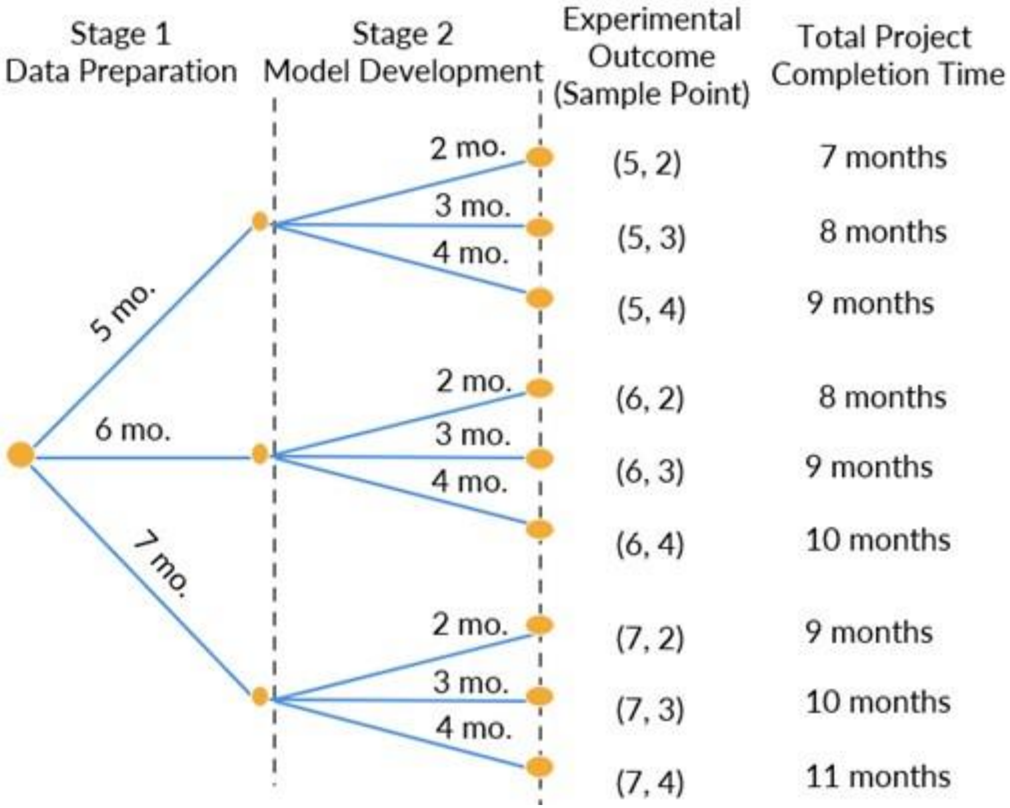
Intersection Events

- Each outcome has two values as they represent the individual outcomes of each of the two dice
- There are six possible outcomes for event A: {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)}
- Again the possible outcomes for event B are: {(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5)}
- The intersection of events A and B (A ∩ B) of these events is (1,5)

Union of Events

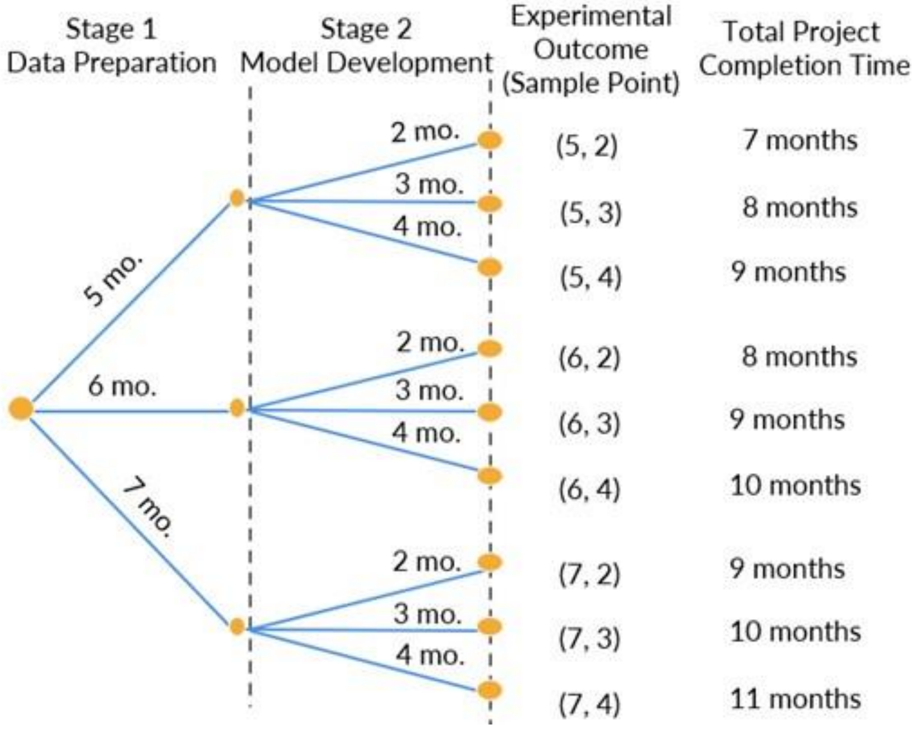
Union of Events

- For our Machine Learning project, what is the probability of completing the data preparation Stage in 5 months (Event A) or completing the Model Development stage in 2 months (Event B) ?
- For Event A, the sample outcomes are (5, 2), (5, 3) and (5, 4)
- For Event B, the sample outcomes are (5, 2), (6, 2) and (7, 2)



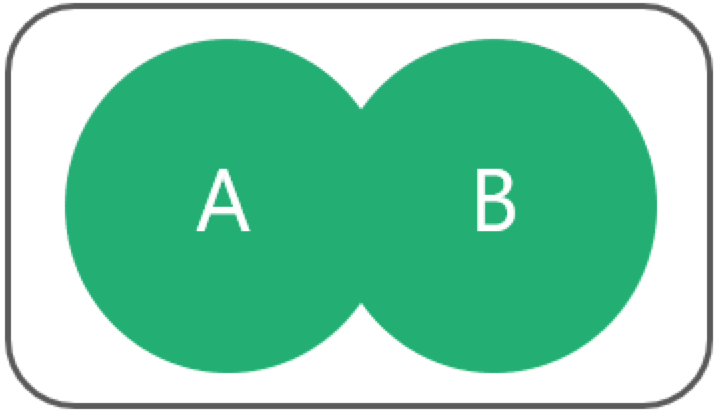
Union of Events

- The ‘or’ criteria represents the union of these events and we determine this by writing all the outcomes that belong to these 2 events which are {(5, 2), (5, 3), (5, 4), (6, 2) and (7, 2)}
- Mathematically, the union of events A and B can be represented as A ∪ (union) B or ‘A or B’



Union of Events

- The shaded green region between the events A and B represents union of A and B



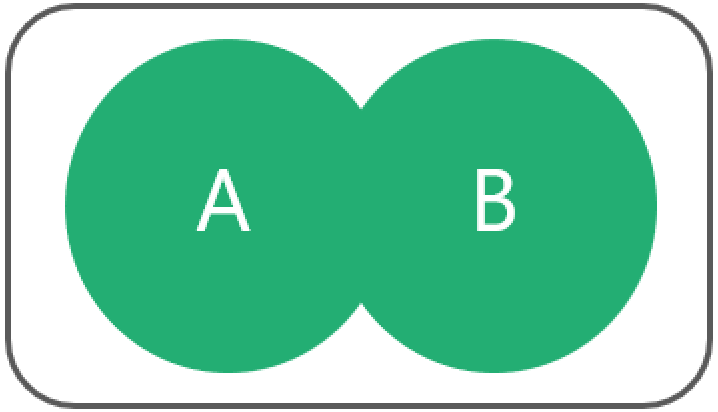
- Let us again consider the example of tossing two dies simultaneously

- Event A, where the first dice rolls 1 (6 possible outcomes): {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)}

- Event B, where the second dice rolls 5 (6 possible outcomes):{(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5)}

Union of Events

- Event A, where the first dice rolls 1 (6 possible outcomes): {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)}
- Event B, where the second dice rolls 5 (6 possible outcomes):{(1, 5), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5)}

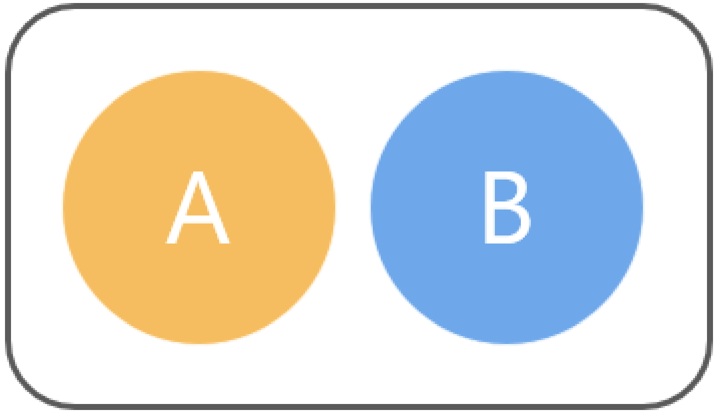


- All the possible AUB outcomes are listed here: {(1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6), (2, 5), (3, 5), (4, 5), (5, 5), (6, 5) }

Mutually Exclusive Events

Mutually Exclusive Events

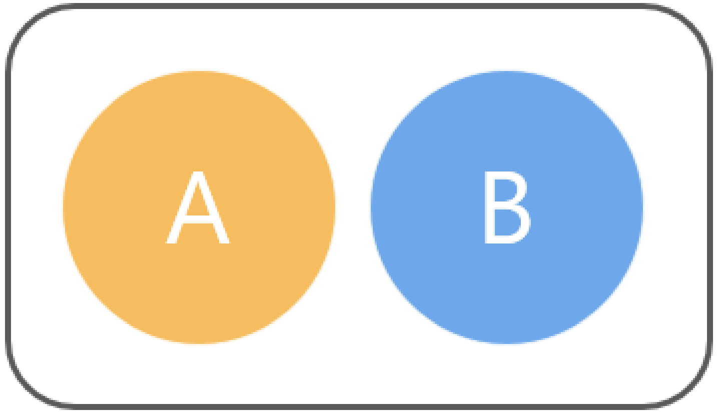
- We can say that two events are mutually exclusive when they do not occur at the same time; that is why they are also called as disjoint events



- For example, if a student has been awarded a grade of C in a subject in a given exam (Event A), she can not be awarded B in the same subject (Event B) in the same exam
- In the Venn diagram shown on the screen, you can see that events A and B do not have an overlapping region

Mutually Exclusive Events

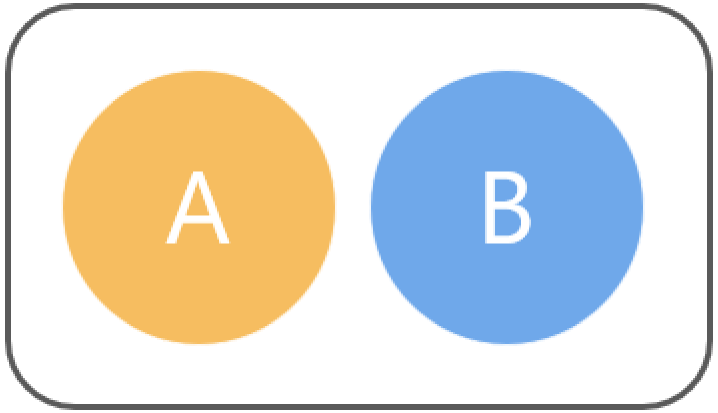
- This is a clear indication that the two events are mutually exclusive or disjoint
- Complement events are always mutually exclusive but not the other way round



- Let us go back to our rolling dice example again
- Event A is the roll where the first die has the number 1. Sample space: (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)
- Event B is Rolls where the first die has the number 3 and the second die has number 5. Sample space: {(3, 5)}.

Mutually Exclusive Events

- Event A is Rolls where the first die has the number 1. Sample space: (1, 1), (1, 2), (1, 3), (1, 4), (1, 5), (1, 6)
- Event B is Rolls where the first die has the number 3 and the second die has number 5. Sample space: {(3, 5)}.



- You can see that these 2 events have nothing in common
- There is no intersection between the two events; hence, these two events are mutually exclusive

Probability for complex events

Probability for complex events

- Let us work through some classic problems related to complex events
- Consider a bin containing 4 balls of different colors, yellow, green, red and blue

- In how many ways you can draw 2 balls from the bin with replacement?
- In the first draw you have 4 balls, so 4 ways
- Similarly, you can draw 4 ways in the second draw. Hence, you can have 4*4 = 16 ways of drawing the 2 balls with replacement

Probability for complex events

- How many ways can you draw 2 balls with replacement such that both are yellow?
- In each draw you have only 1 yellow ball. So there is only 1 way in which both the draws are yellow balls
    - Number of ways to draw 2 balls with replacement is equal to 16
    - Number of ways to draw 2 balls with replacement such that both draws are yellow is equal to 1
- Number of ways to draw 2 balls with replacement such that both draws are yellow is equal to 1/16

Probability for complex events

- There is another interesting way to solve this problem
- What is the probability of drawing a yellow ball? P= ¼
- Drawing 2 yellow balls with replacement is a compound event and is the intersection of the 2 simple events
    - Drawing yellow ball in 1st draw, Event A
    - Drawing yellow ball in 2nd draw, Event B
- And the probability of both events are ¼: thus, the overall event probability is ¼* ¼ =1/6
- In other words P(A∩ (intersection) B) = P(A) * P(B)

Probability for complex events

- P(A∩ (intersection) B) = P(A) * P(B) or multiplication rule
- Multiplication rule works only if the two events are independent
- That is, the two occurrences do not affect each other
- For example, if there are four independent evens A, B, C, and D
- The probability that all these four events will happen together is = P(A)* P(B)*P(C)*P(D)

Probability for complex events

- The multiplication rule can also be applied for dependent events and the formula changes slightly
- This will be dealt along with conditional probabilities
- Independent events and mutually exclusive events are different
- Mutually exclusive events are events that cannot happen together
- Independent events are such that the occurrence of one event is not dependent on the other event

Probability for complex events

- What would happen to the probability of A and B happening together if the events are mutually exclusive?
- The probability of A and B, i.e. P(A and B) or P(A∩B) becomes zero
- Let us consider a more complex example
- The bin now has 20 balls and that out of all the 20 balls, there are 4 yellow balls, 5 green balls, 6 red balls and 5 blue balls
- What is the probability of drawing a yellow ball?
- You can draw 4 ways out of 20, so it's ⅕ = 0.2.

Probability for complex events

- What is the probability of drawing 2 yellow balls with replacement?
- This event is an intersection of 2 events
  - Event A: Drawing yellow ball in 1st draw, Event A
  - Event B: Drawing yellow ball in 2nd draw

- With multiplication rule: $P(A \cap B) = P(A) * P(B) = 4/20 * 4/20 = 1/25 = 0.04$
- What if the events are not independent
- What is the probability of drawing two yellow balls if we don't replace that ball after we have drawn it?

Probability for complex events

- What is the probability of drawing two yellow balls if we don't replace that ball after we have drawn it?

- We still have a one-fifth chance of drawing the first yellow ball
- The same is not the case with second yellow ball
- We have a bin that has 19 balls in it and 3 of them are yellow
- Event B probability of drawing the second yellow ball is 3/19
- The overall probability is =1/5 *3/19=0.032

Conditional Probabilities

Conditional Probabilities

- Let's try to find the probability of drawing a blue ball in the 2nd draw given that you drew a green ball in the first draw without replacement
- Since there is no replacement, the 2nd draw becomes dependent on the 1st draw
- The following events are defined
  - Event A as the Probability of drawing a green ball in draw 1
  - Event B as the Probability of drawing a blue ball in draw 2
- Event B is conditional on event A due to the act of not drawing with replacement
- Mathematically: P(B | A) or P(B given A)

Conditional Probabilities

- The formula for $P(B|A) = P(A \cap B) / P(A)$
- Consider the following example, $P(A)$ is $= 5/20$
- The total number of ways we can draw a green ball in draw 1 and a blue ball in draw 2 is $5 \times 5 = 25$
- Total number of ways of drawing 2 balls $= 20 \times 19$, 19 because there is no replacement
- $P(A \cap B) = 25/380$, and $P(B|A) = P(A \cap B) / P(A) = (25/380)/(5/20) = 5/19$

Conditional Probabilities

- Let us look at another problem: What is the probability of drawing 2 yellow balls without replacement?
- We can define the events as follows: (a) Event A as the Probability of drawing a yellow ball in draw 1; and (b) Event B as the Probability of drawing a yellow ball in draw 2
- Drawing 2 yellow balls is $P(A \cap B)$
- Since we are doing this without replacement, we can use the conditional probability formula: $P(B \text{ given } A) = P(A \cap B) / P(A)$

Conditional Probabilities

- We get P(A ∩ B) = P(B given A) * P(A); where P(A) = 4/20
- Here P(B/A): the probability of drawing a yellow ball in the 2nd draw given we have drawn a yellow ball in the first draw and not replaced it
- So, we have 19 balls left. Since, we have already drawn a yellow ball, 3 yellow balls are left: P(B/A) = 3/19
- So we get P(A ∩ (intersection) B) = 4/20 * 3/19 = 12/380 = 0.032

Conditional Probabilities

- Let’s end this disucssion with 2 interesting notes
- If you look at the formula of $P(A \cap B)$ for dependent and independent events, the only difference is that $P(B)$ gets replaced with P(B given A)
- Second, the condition for the conditional probability formula P(B given A) = $P(A \cap B) / P(A)$ to work is that the denominator is non zero, that is $P(A)$ is not 0.

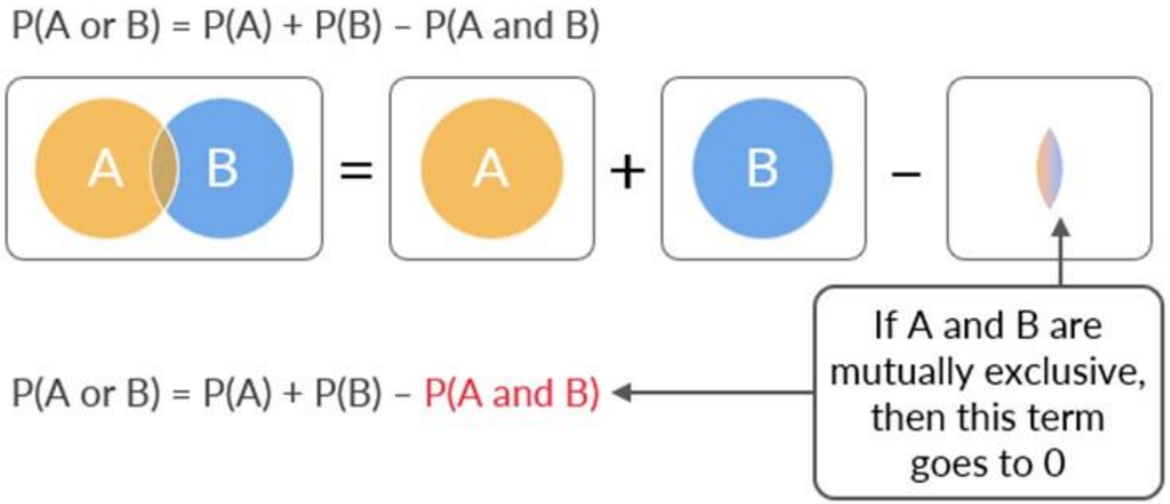
Addition Rule of probability

Addition Rules of probability

- Consider a simple Venn Diagram problem: Your city, which has 100 households, there are two newspapers. Times and Daily news. The circulation departments report that 25 of the city's households have a subscription to the Times and 35 subscribe to the Daily News. A survey reveals that 6 of all households subscribe to both newspapers. How many of the city's households subscribe to either newspaper?

Addition Rules of probability

- A is the number of households reading the Times and B is the number of households reading the Daily News



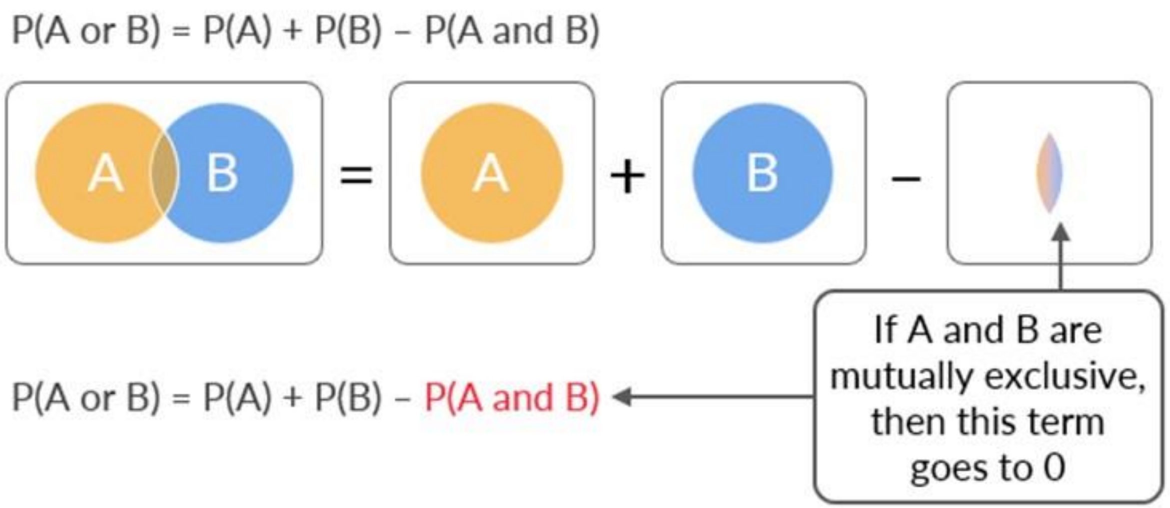
- The shaded region that overlaps A with B, that is A ∩ B indicates the households which have subscription to both

- The shaded region that overlaps A with B, that is A ∩ B indicates the households which have subscription to both

- A = 25, B = 35 and A ∩ B = 6. Our objective is to find A U B

Addition Rules of probability

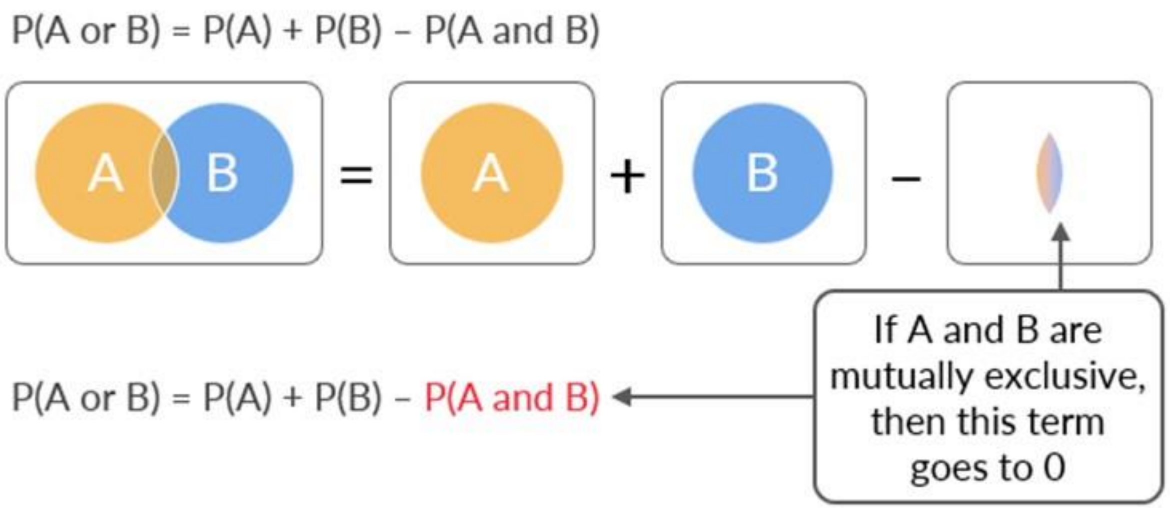
- Divide the three regions as small 'a' as those who read only the Times, small 'b' as those who read only The Daily News and 'c' as those who read both



- A = a + c = 25, B = b + c = 35; A U B = a + b + c and A U B = A + B – c
- Union of A and B (A U B) = A + B - intersection of A and B (A ∩ B)== 25 +35 - 6 = 54

Addition Rules of probability

- Divide by 100 to get the probabilities of these events (why?)
- P(A union B) = 0.54; P(A) = 0.25, P(B) = 0.35; P(A intersection B)= 0.06



- In terms of probabilities: P(A union B) = P(A) + P(B) - P(A intersection B)
- This is known as the addition rule of probability
- If A and B are mutually exclusive, then $P(A \cap B) = 0$: $P(A \cup B) = P(A) + P(B)$.

Joint and Marginal probability

Joint and Marginal probability

- The three major types of probability include: (a) Joint probability, (b) Marginal probability, and (c) Conditional probability
- Let us consider an experiment to know whether an MBA degree could be a possible factor in the success of a mutual funds manager

	Fund outperforms the market	Fund does not outperform the market
Fund manager graduated from a top-30 MBA program	40	100
Fund manager did not graduate from a top-30 MBA program	16	244

Joint and Marginal probability

This can also be represented in another way

Fund manager graduated from a top-30 MBA program	Fund outperforms the market	40
Fund manager graduated from a top-30 MBA program	Fund does not outperform the market	100
Fund manager did not graduate from a top-30 MBA program	Fund outperforms the market	16
Fund manager did not graduate from a top-30 MBA program	Fund does not outperform the market	244

Joint and Marginal probability

Now before we calculate the different types of probabilities, we need to first convert this into a probability table

Fund manager graduated from a top-30 MBA program	Fund outperforms the market	0.10
Fund manager graduated from a top-30 MBA program	Fund does not outperform the market	0.34
Fund manager did not graduate from a top-30 MBA program	Fund outperforms the market	0.04
Fund manager did not graduate from a top-30 MBA program	Fund does not outperform the market	0.61

Joint and Marginal probability

We have also given notations to each of these events for simplicity

A1	B1	0.1
A1	B2	0.34
A2	B1	0.04
A2	B2	0.61

- The joint probability of event A1 and B1 is equal to P(A1 intersection B1) which is 0.11
- The joint probability of events A1 and B2 is equal to 0.34
- The joint probability of event A2 and B1 is equal to 0.04
- And the joint probability of events A2 and B2 is equal to 0.61

Joint and Marginal probability

Next let us understand the concept of marginal probability

A1	B1	0.1
A1	B2	0.34
A2	B1	0.04
A2	B2	0.61

- What is the probability that a mutual fund outperforms the market?
- The probability that the fund will outperform the market will be nothing but the sum of the probabilities in rows 1 and 3, which is equal to 0.1+0.04 which is 0.14 or 14%
- This is marginal probability

Joint and Marginal probability

Marginal probability describes the probability of an event occurring, irrespective of the knowledge gained or the effect from previous or other external events

A1	B1	0.1
A1	B2	0.34
A2	B1	0.04
A2	B2	0.61

- Marginal probability of B1 is nothing but P(B1) which is equal to 0.14

Joint and Marginal probability

- Probability that a mutual fund will outperform the market given that the fund manager graduated from a top MBA program
- The following formula is used: P(B1 given A1) = P(A1 ∩ B1) / P(A1)
- The probability that a manager graduated from a top MBA program: P(A1) = 0.10 + 0.34 = 0.44
- Here, the joint probability P(A1 ∩ B1) from the table as 0.1
- Hence, P(B1 given A1) = 0.1/0.44 = 0.227 (or 22.7%)

Joint and Marginal probability

- What happens to the conditional probability formula if the events are independent of each other?
- P(A given B) = P(A) since P(A1 ∩ B1) = P(A) * P(B) for independent events
- Or alternatively, P(B given A) = P(B) * P(A)
- If the event, 'Mutual Fund outperforms the market' and the event, 'Fund manager is from a top MBA school' are independent of each other or not?
- In the context of conditional probabilities, Bayes theorem assumes considerable significance

Bayes Theorem-I

Bayes Theorem

- Let us go back to our earlier experiment

Fund manager graduated from a top-30 MBA program	Fund outperforms the market	0.10
Fund manager graduated from a top-30 MBA program	Fund does not outperform the market	0.34
Fund manager did not graduate from a top-30 MBA program	Fund outperforms the market	0.04
Fund manager did not graduate from a top-30 MBA program	Fund does not outperform the market	0.61

Bayes Theorem

- We calculated the probability that a mutual fund will outperform the market given the fund manager graduated from a top MBA program
  - Event A is for Fund manager graduated from a top-30 MBA program
  - Event B is for Fund outperforms the market
  - Also, P(B given A) = P(A ∩ B) / P(A)
- The joint probability of A and B: P(A intersection B) or A∩B = 0.10
- P(A) = 0.44, P(B given A) = 0.1/0.44 = 0.227

Bayes Theorem

- What is the probability that a manager has graduated from a top MBA program given that the mutual fund outperformed the market
- P(A given B) = P(A intersection B) / P(B)
- $P(A \cap B) = 0.1$; $P(B) = 0.10+0.04 = 0.14$
- P(A given B) = P(A intersection B) / P(B) = 0.1 / 0.14 = 0.714
- Similarly, P(B given A) = P(A intersection B) / P(A)= 0.227
- Also, P(B given A) P(A) = P(A intersection B)= P(A given B) P(B)
- Rewriting differently, P(B given A) = P(A given B) P(B) / P(A)

Bayes Theorem-II

Bayes Theorem

- Bayes theorem comes into the picture when a direct calculation of a conditional probability is not possible due to lack of information
- Consider the following example, past research suggests that the probability of a middle-aged female developing breast cancer is 0.01 or 1% (Event B)
- A person can be tested positive for breast cancer but in actuality, one may not have it, and vice-versa
- Here, B is the event that a middle-aged female develops breast cancer; P(B)=0.01

Bayes Theorem

- So effectively, we are also asking how reliable is the testing methodology?
- This is a type of conditional probability because we are measuring the probability of breast cancer given the condition that the test came positive
- The event that a female tests positive as event A
- Thus, the probability that we are interested in measuring is basically, P(B given A)
- In a sample of women who already had breast cancer, it was found that only 90% of these women tested positive. This means that the probability of a woman testing positive given she has breast cancer is 0.9: P(A given B) = 0.9.

Bayes Theorem

- From the Bayes rule, P(B given A) = P(A given B) P(B) / P(A)
- We have P(A given B) and P(B) with us; we need to calculate P(A)
- B', which is the complement of B will be an event that a woman does not have breast cancer
- Hence this probability is nothing but P(A given B')
- Also, $P(A) = P(A \cap B) + P(A \cap B')$

	B	**B'**
A	$P(A \cap B)$	$P(A \cap B')$
A'	$P(A' \cap B)$	$P(A' \cap B')$

Bayes Theorem

- From the Bayes rule, P(B given A) = P(A given B) P(B) / P(A)
- We have P(A given B) and P(B) with us; we need to calculate P(A)
- B', which is the complement of B will be an event that a woman does not have breast cancer
- Hence this probability is nothing but P(A given B')
- Also, $P(A) = P(A \cap B) + P(A \cap B')$

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

Bayes Theorem

- $P(A \cap B) = P(A \text{ given } B) * P(B)$ or $P(A \cap B') = P(A \text{ given } B') * P(B')$
- $P(A) = P(A \cap B) + P(A \cap B') = P(A \text{ given } B) * P(B) + P(A \text{ given } B') * P(B')$
- Also, P(B given A) = P(A given B) P(B) / P(A)
- $P(B|A) = (P(A \text{ given } B)\, P(B)) / (P(A \text{ given } B)\, P(B) + P(A \text{ given } B')\, P(B'))$
- P(A given B) = 0.9 and P(B) = 0.01
- P(A given B') = 0.08
- P(B') = 1-P(B) = 0.99

	B	**B'**
A	$P(A \cap B)$	$P(A \cap B')$
A'	$P(A' \cap B)$	$P(A' \cap B')$

Bayes Theorem

- P(B|A)= (P(A given B) P(B)) / (P(A given B) P(B) + P(A given B') P(B'))
- P(B|A)= (0.9 * 0.01) /((0.9 * 0.01) + (0.08 * 0.99)) = 0.102 or 10.2%
- There is only a 10.2% chance that a woman will develop breast cancer if she tests positive
- Bayes theorem : P(B|A)= (P(A given B) P(B)) / (P(A given B) P(B) + P(A given B') P(B'))
- Let us understand the significance of this equation

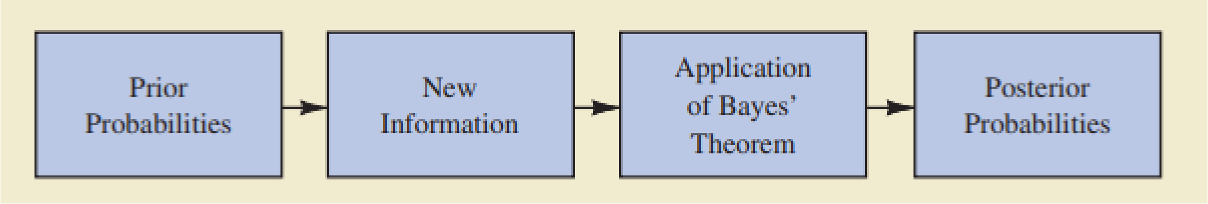
	B	B'
A	P(A∩B)	P(A∩B')
A'	P(A'∩B)	P(A'∩B')

Bayes Theorem

- We start off with a set of initial or prior probabilities
- Prior probability is the probability of an event before we did our experiment and collected our new data
- In our case, the probability of event B, which was the probability of a female developing breast cancer is the prior probability
- We obtained new information about all the events for which we had the prior probabilities
- This information was given to us in the form of the conditional probabilities, P(A given B) and P(A given B')

Bayes Theorem

- Our objective was basically to update these our prior knowledge about females developing breast cancer with this new information and calculate P(B given A)



- This new probability that we calculated is referred to as posterior probabilities and Bayes' theorem is used for making these probability calculations
- In our case, the probability that a woman has or will develop breast cancer if she tests positive was the posterior probability
- $P(B|A) = (P(A \text{ given } B)\, P(B)) \,/\, (P(A \text{ given } B)\, P(B) + P(A \text{ given } B')\, P(B'))$

Thanks!