Short Note on

Probability

FOR DATA SCIENCE

Agenda

1. Terminologies
2. Types of Events
3. Probability
4. Random Variable
5. Types of Probability
6. Bayes' Theorem

1. Terminologies

RANDOM EXPERIMENT	A random experiment is an action or process that can have multiple possible outcomes, but the result cannot be predicted in advance.
TRIAL	A trial is a single performance of a random experiment, which results in one outcome.
OUTCOME	An outcome is the result obtained from a single trial of a random experiment.
SAMPLE SPACE	The sample space is the complete set of all possible outcomes of a random experiment.
EVENT	An event is a specific group of outcomes from a random experiment, and it can include one or more outcomes from the sample space.

Terminologies Example

RANDOM EXPERIMENT	Tossing a Coin
TRIAL	Each time you toss the coin once, it's a trial
OUTCOME	The result of a single coin toss (a trial) is an outcome, e.g., Heads.
SAMPLE SPACE	The sample space is the set of all possible outcomes of the experiment: S = {Heads, Tails}
EVENT	An event is a subset of the sample space. For example, Event E = getting a Head So, E = {Heads} Another example: Event F = getting either Heads or Tails → F = {Heads, Tails}

Types of events

Simple Event

A simple event, also called an elementary event, is an event that has only one possible outcome.

For Example:
Rolling a fair six-sided die, getting a 3 is a simple event

Compound Event

A compound event is made up of two or more simple events.

For Example:
When rolling a die, the event "rolling an odd number" is a compound event because it includes getting a 1, 3, or 5.

Independent Event

Two events are called independent if the result of one event does not change the chance of the other event happening.

For Example:
Flipping a coin and rolling a die are independent events because the coin flip doesn't affect the die roll.

Dependent Event

Events are dependent if one event affects the chance of the other happening.

For Example:
Drawing two cards without replacement is dependent because the first draw changes the cards left for the second draw.

Types of events

Mutually Exclusive Event

Two events are mutually exclusive if they can't happen at the same time.

For Example:
R Rolling a die and getting a 2 or a 4 only one can happen in a single roll.

Exhaustive Events

A set of events is exhaustive if one of them must happen when the experiment is done.

For Example:
Rolling a die and getting either an even or an odd number—one of them will always happen.

Impossible Event & Certain Event

Impossible Event: An event that can never happen.

Example: Rolling a die and getting a 7.

Certain Event: An event that always happens.

Example: Rolling a die and getting a number between 1 and 6.

PROBABILITY

Probability is a way to measure how likely an event is to happen. It's a basic idea in statistics that helps us make predictions and smart decisions in areas like science, engineering, medicine, economics, and social sciences.

Probability is shown as a number between 0 and 1.
A probability of 0 means the event will not happen, while 1 means it will definitely happen.
A probability of 0.5 means the event is equally likely to happen or not happen.

Empirical Vs Theoretical Probability

Empirical Probability

Experimental Probability (also called observed probability) is based on actual results from experiments or trials. It's calculated by dividing the number of times an event happens by the total number of trials.

Theoretical Probability

Theoretical Probability is used when all outcomes are equally likely. It's found by dividing the number of favorable outcomes by the total number of possible outcomes.
Formula: Probability of event A = (Outcomes in A) / (Total outcomes in sample space)

Random Variable

In probability theory, a random variable is a rule or function that turns the outcomes of a random process into numbers. It helps us represent random outcomes in a way we can calculate and analyze.

- Input: The input to a random variable is an outcome from the sample space (all the possible results of an experiment or random process).
- Output: The output is a real number that we assign to each of those outcomes.

Imagine you roll a six-sided die. The sample space is {1, 2, 3, 4, 5, 6}.
We can define a random variable X such that:

- If the die shows 1, then X = 1
- If the die shows 2, then X = 2
- ... and so on.

In this case, the random variable X simply assigns the number that appears on the die.

Types of Random Variable:
1) Discrete Random Variable
2) Continuous Variable

1. Discrete Random Variable:

- Takes a countable number of values.
- These values are often whole numbers like 0, 1, 2, 3, etc.
- Example: Number of heads in 5 coin tosses, number of cars passing a street in an hour.
- The outcomes are distinct and separate.

2. Continuous Random Variable:

- Can take any value within a range, including decimals.
- There are infinitely many possible values.
- Example: The height of students, time taken to run a marathon, temperature.
- The values are measured, not counted.

PROBABILITY DISTRIBUTION OF A RANDOM VARIABLE

A probability distribution shows all the possible outcomes of a random variable along with the probability of each outcome. It helps us understand how the values of a random variable are likely to behave.

Example (Discrete Random Variable)
Tossing a fair coin:

- Sample Space = {H, T}
- Define Random Variable X such that:
- X = 1 if Head (H)
- X = 0 if Tail (T)

Probability Distribution Table:

X:	1	0
P(X):	½	½

For random variables with many possible values, especially continuous ones, we often express the relationship between the value (X) and its probability (P) using a function:
y = f(X), this is called a Probability Distribution Function.

For discrete we use probability mass function and for continuous we use probability density function.

MEAN OF A RANDOM VARIABLE

The mean of a random variable, also known as its expected value (E[X]), represents the long-run average outcome of a random process if it were repeated many times.

In simple terms, it answers the question:
"What value should we expect on average?"

More precisely, it is a weighted average of all the possible values that the random variable can take, where each value is weighted by its probability of occurring.

Example: Rolling a fair six-sided die

X:	1	2	3	4	5	6
P(X):	⅙	⅙	⅙	⅙	⅙	⅙

So, we multiply the random variable with its probability
Mean (E[X]) = $\sum (x \times P(x))$
Mean = 1*1/6 + 2*1/6 + 3*1/6 + 4*1/6 + 5*1/6 + 6*1/6

So, the expected value of a die roll is 3.5, which isn't an actual outcome, but represents the average value over many rolls.

VARIANCE OF A RANDOM VARIABLE

Variance is a measure of how much the values of a random variable spread out or deviate from the mean (expected value). It helps us understand the variability or consistency in the outcomes of a random process.

Formula:
For a discrete random variable X with mean $\mu = E[X]$,
the variance is calculated as:
$Var(X) = E[(X - \mu)^2] = \sum (x_i - \mu)^2 \times P(x_i)$
Alternatively, using a shortcut formula:
$Var(X) = E[X^2] - (E[X])^2$

Example: Rolling a Fair Die
Let X be the outcome when rolling a fair 6-sided die.

- Possible values: X = {1, 2, 3, 4, 5, 6}
- Each outcome has probability: P(X) = 1/6
- Mean (E[X]) = (1+2+3+4+5+6)/6 = 3.5

Using the formula, Variance = 2.92

So, the variance is about 2.92, meaning that the outcomes typically deviate from the mean (3.5) by a squared distance of about 2.92 units.
A higher variance means more spread out values, while a lower variance means the values are closer to the mean.

TYPES OF PROBABILITY

1. **JOINT PROBABILITY**
2. MARGINAL PROBABILITY
3. CONDITIONAL PROBABILITY

When we have two random variables, say X and Y, the joint probability denoted as P(X = x, Y = y) is the probability that X takes the value x and Y takes the value y at the same time.

Example (Titanic Dataset):
Let's X be a random variable associated with the Pclass of a passenger.
X: Passenger class (Pclass) => values: 1, 2, 3
Let's Y be a random variable associated with the survival status of a passenger.
Y: Survival status => 0 = did not survive, 1 = survived

So, probability distribution of X and Y will look like:

X / Y	1	2	3
0	80	97	372
1	136	87	119

Total Passenger = 891

joint probability Calculation:
Example:
P(X = 1, Y = 0) = 80/891 = 0.0897
Now, let's calculate all joint probabilities:

X / Y	1	2	3
0	0.0897	0.1088	0.4175
1	0.1526	0.0976	0.1335

Each value in this table represents the joint probability for a specific combination of X and Y. The complete table is called the joint probability distribution.

TYPES OF PROBABILITY

1. JOINT PROBABILITY
2. MARGINAL PROBABILITY
3. CONDITIONAL PROBABILITY

Marginal probability refers to the probability of an event occurring irrespective of the outcome of some other event. When dealing with random variables, the marginal probability of a random variable is simply the probability of that variable taking a certain value, regardless of the values of other variables.

Example (Titanic Dataset):
Let's X be a random variable associated with the Pclass of a passenger.
X: Passenger class (Pclass) => values: 1, 2, 3
Let's Y be a random variable associated with the survival status of a passenger.
Y: Survival status => 0 = did not survive, 1 = survived

So, probability distribution of X and Y will look like:

X	1	2	3	All
Y				
0	80	97	372	549
1	136	87	119	342
All	216	184	491	891

Now, let's calculate all marginal probabilities:

X	1	2	3	All
Y				
0	0.0897	0.1088	0.4175	0.6161
1	0.1526	0.0976	0.1335	0.3838
All	0.2424	0.2065	0.5510	1.0000

From the above table:

- The marginal probability of being in passenger class 1 (X = 1) is P(X = 1) = 0.2424

The above distribution is called marginal probability distribution, also known as unconditional probability.

TYPES OF PROBABILITY

1. JOINT PROBABILITY
2. MARGINAL PROBABILITY
3. **CONDITIONAL PROBABILITY**

Conditional probability is a measure of the likelihood of an event occurring given that another event has already occurred. If the event of interest is A and event B is known to have occurred, the conditional probability of A given B is denoted as p(A|B).

Formula for conditional probability:
P(A|B) = P(A,B) / P(B)

Intuition behind conditional probability:
The intuition behind conditional probability lies in reducing the sample space. When we say we want to calculate the probability of A given B, we are focusing only on the cases where B has occurred. In other words, B becomes our new sample space.

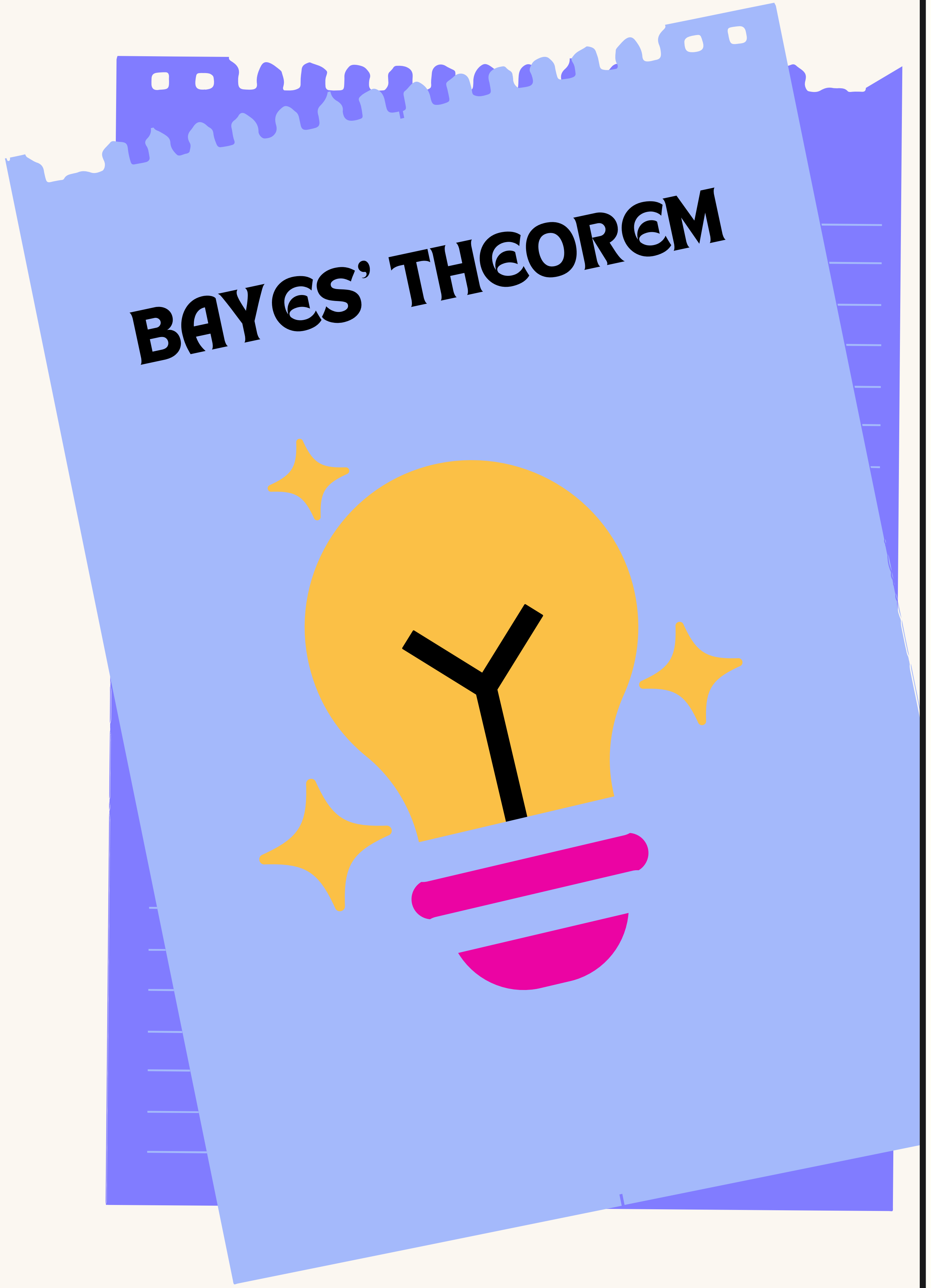
- Denominator (P(B)): This represents the probability of event B occurring. Since B is given, we are only considering outcomes within B.
- Numerator (P(A,B)): This represents the probability that both A and B occur together.

By dividing the joint probability P(A , B) by the probability of B, we are effectively asking:
Out of all cases where B happens, how many times does A also happen?

Example (Titanic Dataset, X = Pclass and Y = Survival Status):

X	1	2	3	All
Y				
0	0.0897	0.1088	0.4175	0.6161
1	0.1526	0.0976	0.1335	0.3838
All	0.2424	0.2065	0.5510	1.0000

To find the conditional probability that a passenger did not survive given they were in 3rd class, we calculate:
P(Y=0|X=3) = P(Y=0,X=3)/P(X) = 0.4175/0.5510 = 0.7582
So, about 75.82% of 3rd class passengers did not survive.



Bayes' Theorem is a fundamental principle in probability and statistics that describes how to update the probability of a hypothesis in light of new evidence. It is widely used in areas such as machine learning, statistics, medical diagnostics, natural language processing, and game theory.

Named after Thomas Bayes, who first formulated a version of this rule, the theorem allows us to revise prior beliefs by incorporating observed data. In essence, it provides a formal mechanism to move from prior probability to posterior probability.

$P(A|B) = [P(A) \times P(B|A)] / P(B)$

Where:

- $P(A|B)$ = Posterior probability: the probability of event A given that B has occurred.
- $P(B|A)$ = Likelihood: the probability of event B given that A is true.
- $P(A)$ = Prior probability of event A.
- $P(B)$ = Marginal probability of event B.

<u>Mathematical Proof:</u>

According to conditional probability:

$P(A|B) = P(A,B) / P(B) \quad \ldots (1)$

$P(B|A) = P(A,B) / P(A) \quad \ldots (2)$

From equation (2), we can rearrange to express the joint probability:

$P(A,B) = P(A) \times P(B|A) \quad \ldots (3)$

Now substitute equation (3) into equation (1):

$P(A|B) = [P(A) \times P(B|A)] / P(B)$

This completes the proof of Bayes' Theorem.

Probability isn't just about numbers. it's about how we make sense of uncertainty in everyday life.

From understanding risks to making better decisions in business and tech, the concepts we covered—events, random variables, types of probability, and Bayes' Theorem—are the building blocks of critical thinking in a data-driven world.

Whether you're diving into analytics, exploring machine learning, or simply curious about how things work behind the scenes, learningprobability is a powerful step forward.

Thanks for taking the time to go through this!
Let me know your thoughts.

THANK YOU!