PROBABILITY AND STANDARD PROBABILITY DISTRIBUTIONS

BASIC PROBABILITY CONCEPTS

- *Marginal Probability*
- *Joint Probability*
- *Conditional Probability*
- *Probability Trees and Bayes' Theorem*

PROBABILITY

- Probability quantifies the uncertainty of the outcomes of a random variable.
- Or, it quantifies likelihood or possibilities of an event.
- Specifically, it quantifies how likely a specific outcome is for a random variable, such as the flip of a coin, the roll of a dice, or drawing a playing card from a deck.
- For a random variable x, $P(x)$ is a function that assigns a probability to all values of x.
- Probability Density of x = P(x)

PROBABILITY

- The probability of a specific event *A* for a random variable x is denoted as *P(x=A)*, or simply as *P(A).*
- Probability of Event A = P(A)
- Probability is calculated as the number of desired outcomes divided by the total possible outcomes, in the case where all outcomes are equally likely.
- Probability = (number of desired outcomes) / (total number of possible outcomes)

PROBABILITY

The probability of an event not occurring, is called the complement.

This can be calculated by one minus the probability of the event, or *1 – P(A)*.

For example, the probability of not rolling a 5 would be

1 – P(5) or 1 – 0.166 or about 0.833 or about 83.333%.

Probability of Not Event A = 1 – P(A)

Probability can range in from 0 to 1, where 0 means the event to be an impossible one and 1 indicates a certain event.

The probability of all the events in a sample space adds up to 1.

PROBLEMS AND SOLUTIONS ON PROBABILITY

Question 1: Find the probability of 'getting 3 on rolling a die'.

Solution:

- Sample Space = S = {1, 2, 3, 4, 5, 6}
- Total number of outcomes = n(S) = 6
- Let A be the event of getting 3.
- Number of favourable outcomes = n(A) = 1
- i.e. A = {3}
- Probability, P(A) = n(A)/n(S) = 1/6
- Hence, P(getting 3 on rolling a die) = 1/6

QUESTION 2: DRAW A RANDOM CARD FROM A PACK OF CARDS. WHAT IS THE PROBABILITY THAT THE CARD DRAWN IS A FACE CARD? SOLUTION:

A standard deck has 52 cards.

Total number of outcomes = n(S) = 52

Let E be the event of drawing a face card.

Number of favourable events = n(E) = 4 x 3 = 12 (considered Jack, Queen and King only)

Probability, P = Number of Favourable Outcomes/Total Number of Outcomes

P(E) = n(E)/n(S)

= 12/52

= 3/13

P(the card drawn is a face card) = 3/13

QUESTION 3: A VESSEL CONTAINS 4 BLUE BALLS, 5 RED BALLS AND 11 WHITE BALLS. IF THREE BALLS ARE DRAWN FROM THE VESSEL AT RANDOM, WHAT IS THE PROBABILITY THAT THE FIRST BALL IS RED, THE SECOND BALL IS BLUE, AND THE THIRD BALL IS WHITE?

Solution:

The probability to get the first ball is red or the first event is 5/20.

Since we have drawn a ball for the first event to occur, then the number of possibilities left for the second event to occur is 20 – 1 = 19.

Hence, the probability of getting the second ball as blue or the second event is 4/19.

Again with the first and second event occurring, the number of possibilities left for the third event to occur is 19 – 1 = 18.

And the probability of the third ball is white or the third event is 11/18.

Therefore, the probability is 5/20 x 4/19 x 11/18 = 44/1368 = 0.032.

Or we can express it as: P = 3.2%.

Question 4: Two dice are rolled, find the probability that the sum is:

1. **equal to 1 2. less than 13**

Solution:

To find the probability that the sum is equal to 1 we have to first determine the sample space S of two dice as shown below.

S = { (1,1),(1,2),(1,3),(1,4),(1,5),(1,6)

(2,1),(2,2),(2,3),(2,4),(2,5),(2,6)

(3,1),(3,2),(3,3),(3,4),(3,5),(3,6)

(4,1),(4,2),(4,3),(4,4),(4,5),(4,6)

(5,1),(5,2),(5,3),(5,4),(5,5),(5,6)

(6,1),(6,2),(6,3),(6,4),(6,5),(6,6) }

So, n(S) = 36

1) Let E be the event "sum equal to 1". Since, there are no outcomes which where a sum is equal to 1, hence, P(E) = n(E) / n(S) = 0 / 36 = 0

2) Let B be the event of getting the sum of numbers on dice is less than 13.

From the sample space, we can see all possible outcomes for the event B, which gives a sum less than B. Like: (1,1) or (1,6) or (2,6) or (6,6). So you can see the limit of an event to occur is when both dies have number 6, i.e. (6,6). Thus, n(B) = 36

Hence, P(B) = n(B) / n(S) = 36 / 36 = 1

EQUALLY LIKELY EVENTS

When the events have the same theoretical probability of happening, then they are called equally likely events.

The results of a sample space are called equally likely if all of them have the same probability of occurring.

- Getting 3 and 5 on throwing a die
- Getting an even number and an odd number on a die
- Getting 1, 2 or 3 on rolling a die

COMPLEMENTARY EVENTS

The possibility that there will be only two outcomes which states that an event will occur or not.

Basically, the complement of an event occurring in the exact opposite that the probability of it is not occurring. Some more examples are:

- It will rain or not rain today
- The student will pass the exam or not pass.
- You win the lottery or you don’t.

INDEPENDENT EVENTS

If the probability of occurrence of an event A is not affected by the occurrence of another event B, then A and B are said to be independent events.

Consider an example of rolling a die.

If A is the event 'the number appearing is odd' and B be the event 'the number appearing is a multiple of 3', then

$P(A) = 3/6 = 1/2$ and $P(B) = 2/6 = 1/3$

Also A and B is the event 'the number appearing is odd and a multiple of 3' so that

$P(A \cap B) = 1/6$

$P(A | B) = P(A \cap B)/ P(B) = (1/6) / (1/3) = 1/2$

$P(A) = P(A | B) = 1/2$, which implies that the occurrence of event B has not affected the probability of occurrence of the event A .

If A and B are independent events, then $P(A | B) = P(A)$

Using Multiplication rule of probability, $P(A \cap B) = P(B) .P(A | B)$

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

MUTUALLY EXCLUSIVE EVENTS

Two events are said to be mutually exclusive if they cannot occur at the same time or simultaneously. In

They are also called disjoint events.

If two events are considered disjoint events, then the probability of both events occurring at the same time will be zero.

If the events A and B are not mutually exclusive, the probability of getting A or B that is $P(A \cup B)$ formula is given as follows:

$$P(A \cup B) = P(A) + P(B) - P(A \text{ and } B)$$

When tossing a coin, the event of getting head and tail are mutually exclusive.

In a six-sided die, the events “2” and “5” are mutually exclusive.

MARGINAL PROBABILITY

The probability of an event occurring (p(A)), it may be thought of as an unconditional probability.

It is not conditioned on another event.

Example: the probability that a card drawn is red

(p(red) = 0.5).

Another example: the probability that a card drawn is 4

(p(four)=1/13).

JOINT PROBABILITY

It is the probability of two different event A and event B occurring at the same time.

It is the probability of the intersection of two or more events.

The probability of the intersection of A and B may be written $p(A \cap B)$ or p(A and B).

Example: the probability that a card is a four and red =

p(four and red) = 2/52=1/26.

(There are two red fours in a deck of 52, the 4 of hearts and the 4 of diamonds).

CONDITIONAL PROBABILITY

It is the probability of event A occurring, given that event B occurs.

Mathematically, it is represented as P(X | Y).

This is read as “probability of X given/conditioned on Y”.

Example: given that you drew a red card, what’s the probability that it’s a four (p(four|red))=2/26=1/13.

So out of the 26 red cards (given a red card), there are two fours so 2/26=1/13.

BAYES' THEOREM

It describes the probability of an event, based on prior knowledge of conditions that might be related to that event.

It can also be considered for conditional probability examples.

It is used where the probability of occurrence of a particular event is calculated based on other conditions which are also called conditional probability.

For example: There are 3 bags, each containing some white marbles and some black marbles in each bag. If a white marble is drawn at random. With probability to find that this white marble is from the first bag. In cases like such, we use the Bayes' Theorem.

THEOREM OF TOTAL PROBABILITY

Let E_1, E_2,…………..E_n is mutually exclusive and exhaustive events associated with a random experiment and lets E be an event that occurs with some E_i.

Then,

$P(E) = {}^{n}\sum_{i=1} P(E/E_i) \,.\, P(E_i)$

THERE ARE THREE URNS CONTAINING 3 WHITE AND 2 BLACK BALLS; 2 WHITE AND 3 BLACK BALLS; 1 BLACK AND 4 WHITE BALLS RESPECTIVELY. THERE IS AN EQUAL PROBABILITY OF EACH URN BEING CHOSEN. ONE BALL IS EQUAL PROBABILITY CHOSEN AT RANDOM. WHAT IS THE PROBABILITY THAT A WHITE BALL IS DRAWN?

Let E1, E2, and E3 be the events of choosing the first, second, and third urn respectively. Then,

$P(E1) = P(E2) = P(E3) = 1/3$

Let E be the event that a white ball is drawn. Then,

$P(E/E1) = 3/5$, $P(E/E2) = 2/5$, $P(E/E3) = 4/5$

By theorem of total probability, we have

$P(E) = P(E/E1) . P(E1) + P(E/E2) . P(E2) + P(E/E3) . P(E3)$

$= (3/5 * 1/3) + (2/5 * 1/3) + (4/5 * 1/3)$

$= 9/15 = 3/5$

BAYES' THEOREM

The Bayes' theorem is expressed in the following formula:

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

Where:

- P(A|B) – the probability of event A occurring, given event B has occurred
- P(B|A) – the probability of event B occurring, given event A has occurred
- P(A) – the probability of event A
- P(B) – the probability of event B

Note that events A and B are independent events.

60% of the companies that increased their share price by more than 5% in the last three years replaced their CEOs during the period.

At the same time, only 35% of the companies that did not increase their share price by more than 5% in the same period replaced their CEOs. Knowing that the probability that the stock prices grow by more than 5% is 4%, find the probability that the shares of a company that fires its CEO will increase by more than 5%.

Before finding the probabilities, you must first define the notation of the probabilities.

- P(A) – the probability that the stock price increases by 5%
- P(B) – the probability that the CEO is replaced
- P(A|B) – the probability of the stock price increases by 5% given that the CEO has been replaced
- P(B|A) – the probability of the CEO replacement given the stock price has increased by 5%.

USING THE BAYES' THEOREM, WE CAN FIND THE REQUIRED PROBABILITY:

$$P(A|B) = \frac{0.60 \times 0.04}{0.60 \times 0.04 + 0.35 \times (1 - 0.04)} = 0.067 \text{ or } 6.67\%$$

RANDOM VARIABLE

A variable is defined as any symbol that can take any particular set of values.

If the value of a variable depends upon the outcome of a random experiment, it is a random variable and can take up any real value.

Such an experiment, where we know the set of all possible results but find its impossible to predict one at any particular execution, is a random experiment.

Mathematically, a random variable is a real-valued function whose domain is a sample space S of a random experiment.

Random variable is always denoted by capital letter like X,Y,M.

Lowercase letters like x, y, z, m etc. represent its value.

X denotes the Probability Distribution of random variable X.
P(X) denotes the Probability of X.
p(X=x) denotes the Probability that random variable X is equivalent to any particular value, represented by x.

Experiment is tossing a coin 2 times.

Sample space(S) is {HH, TH, HT, TT}.

X(Random Variable) is the number of both heads when we toss a coin 2 times.

P(X=HT)=0.25, P(X=TT)=0.25, P(X=HH)=0.25, P(X=HT)=0.25

For outcome {HT},

Then, X(HH) = 0, X(TH) = 0, X(HT) = 1, X(TT) = 0.

RANDOM VARIABLE

Since there are two forms of data, discrete and continuous, there are two types of random variables.

It can be categorized into two types:

Discrete Random Variable

Continuous Random variable

DISCRETE RANDOM VARIABLE

A discrete random variable is a random variable with a finite or countably infinite range or it takes values that are countable.

Examples: Birth year, total phone calls, number cars etc.

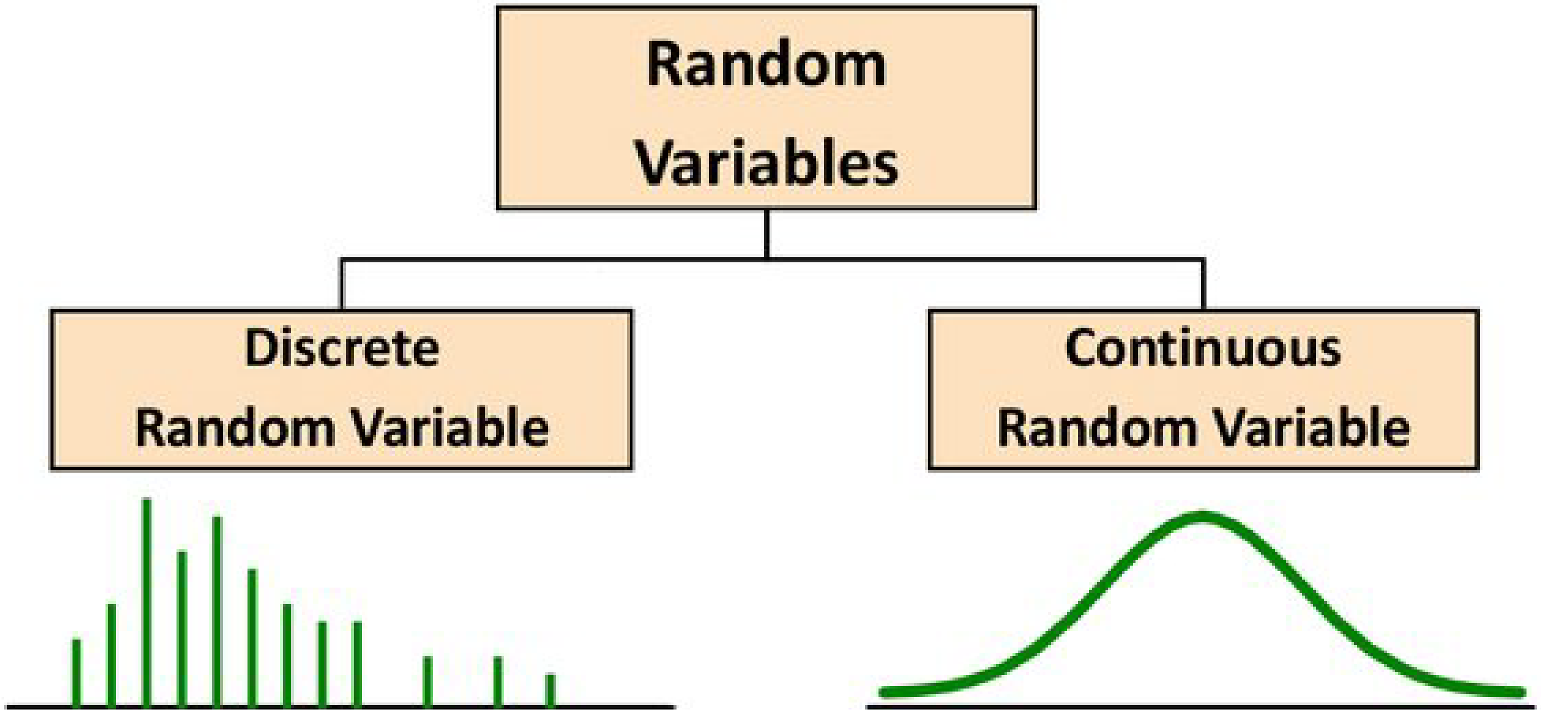
A probability distribution is used to determine what values a random variable can take and how frequently it does so.

CONTINUOUS RANDOM VARIABLE

When the range of a random variable is on a continuous scale, it is called a continuous random variable or it takes an infinite number of possible values. A variable like this is defined over a range of values rather than a single value.

Examples: Height, Weight, Amount of rainfall, etc.

A probability density function is used to describe a continuous random variable because the probability that it will take on an exact value is zero.


Random Variables
Discrete Random Variable
Continuous Random Variable

PROBABILITY DISTRIBUTIONS

A probability distribution is a function that calculates the likelihood of all possible values for a random variable.

For any event of a random experiment, we can find its corresponding probability.

For different values of the random variable, we can find its respective probability.

The values of random variables along with the corresponding probabilities are the probability distribution of the random variable.

A probability distribution and probability mass functions can both be used to define a discrete probability distribution.

A continuous probability distribution is described using a probability distribution function and a probability density function.

FOR DISCRETE RANDOM VARIABLES

The probability distribution of a discrete random variable is a list of probabilities associated with each of its possible values.

It is also sometimes called the probability function or the probability mass function.

More formally, the probability distribution of a discrete random variable X is a function which gives the probability p(xi) that the random variable equals xi, for each value xi:

p(xi) = P(X=xi)

It satisfies the following conditions:

- 0 <= p(xi) <= 1
- sum of all p(xi) is 1

FOR DISCRETE RANDOM VARIABLES

Consider a random variable X= number of heads after tossing a coin thrice.

x $\in$ *{0,1,2,3}*.

All the possible outcomes after a coin is flipped thrice are, {HHH,HHT,HTT,TTT,TTH,THH,THT,HTH}.

What will be the probability that 0 heads occur?

We denote it as P(X=0)=1/8=0.125

probability of getting exactly 1 head=P(X=1)=3/8=0.375

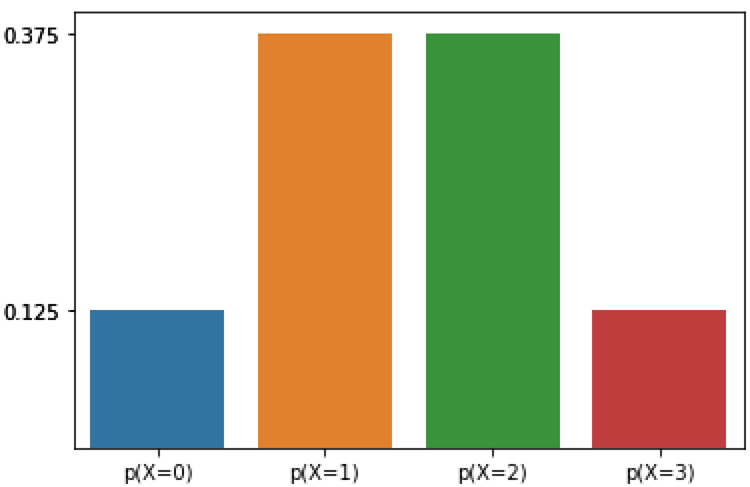
P(X=2)=3/8=0.375

P(x=3)=1/8=0.125

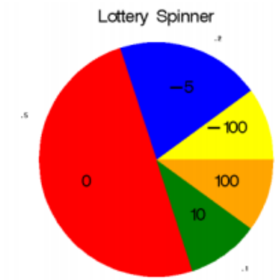
If we sum up the probabilities of all outcomes, it will be equal to one. This gives us the Probability Distribution of that random variable.

In the case of Discrete Random Variables, the function that denotes the probability of the random variable for each x in the range of X is known as the **Probability Distribution Function(PDF).**

It can be shown using tables or graph or mathematical equation.



Color	Y	$P(Y)$
Yellow	-100	.10
Blue	-5	.20
Red	0	.50
Green	10	.10
Tan	100	.10



> note that the values of x take on all possible cases. and the sum of the probabilities add to 1. mathematically, this can be written as f(x) = p(x = x). the set of ordered pairs (x, f(x)) is called the probability function, probability mass function or probability distribution function of the discrete random variable x. f(x) is considered a probability mass function if it satisfies the following conditions:

In case of rolling of a die, the probability of each value X can take is the same. So the probability distribution in this case will be:
P(X=1) = 1/6, P(X=2) = 1/6 and so on.

x	1	2	3	4	5	6
P(X=x)	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Note that the values of x take on all possible cases. And the sum of the probabilities add to 1.

Mathematically, this can be written as $f(x) = P(X = x)$.

The set of ordered pairs $(x, f(x))$ is called the **probability function**, **probability mass function** or **probability distribution function** of the discrete random variable X.

$f(x)$ is considered a probability mass function if it satisfies the following conditions:

1. $f(x) \geq 0$,
2. $\sum_x f(x) = 1$,
3. $P(X = x) = f(x)$.

However, many times we may wish to compute the probability that the random variable X be lesser than or equal to some real number *x.*

Writing F(*x*) = P(X ≤ *x*) for every real number *x*, we define F(*x*) to be the **cumulative distribution function** of the random variable X.

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty.$$

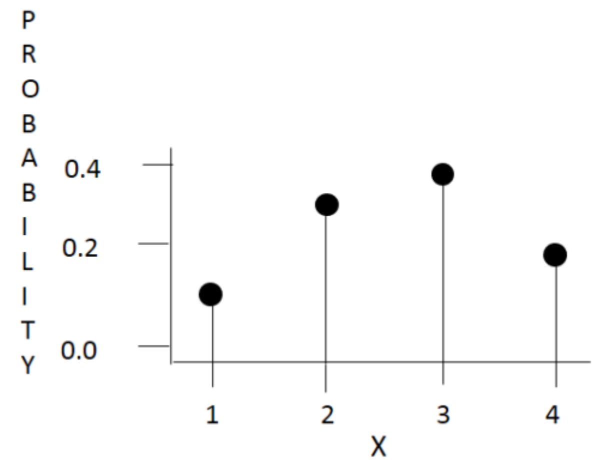
Suppose a random variable X has the following probability distribution p(xi):

xi	0	1	2	3	4	5
p(xi)	1/32	5/32	10/32	10/32	5/32	1/32

The cumulative distribution function F(x) is then:

xi	0	1	2	3	4	5
F(xi)	1/32	6/32	16/32	26/32	31/32	32/32

BASED ON BELOW PICTURE WHAT IS THE PROBABILITY THAT X TAKES A VALUE OF EITHER 2 OR 3.



P(X=2 or X=3) = P(X=2) + P(x=3) = 0.3+0.4 =0.7

Here union of probabilities = sum of probabilities.

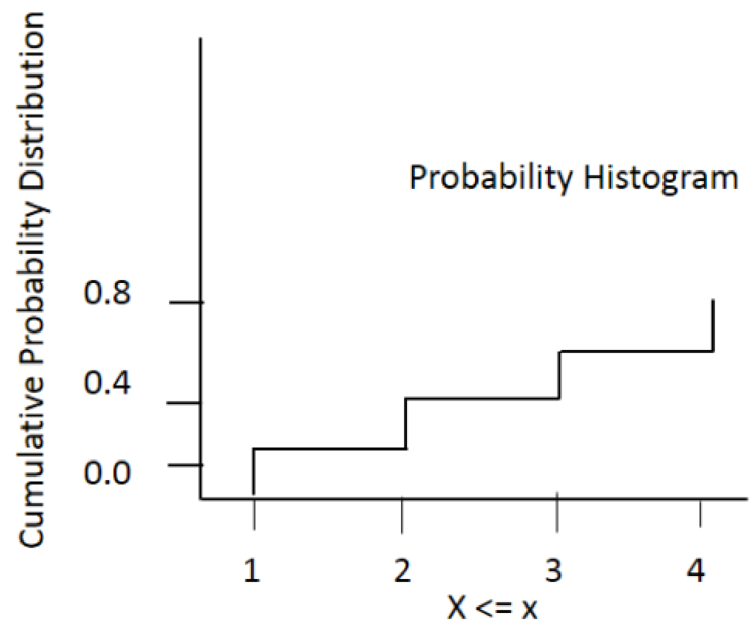
What is the probability that X is greater than 1 ?

P(X=2 or X=3 or X=4) = 0.3+0.4+ 0.2 =0.9

Or, P(X=2 or X=3 or X=4)= 1- P(x=1) = 1- 0.1 = 0.9 both the answer are same. Here we are finding from complementary rule.

THE PROBABILITY OF X IS LESS THAN OR 1 IS 0.1. SIMILARLY, PROBABILITY OF X IS LESS THAN OR EQUAL TO 2 IS (0.1+0.3) =0.4 AND SO ON.

Outcome	1	2	3	4
Probability	0.1	0.3	0.4	0.2
Cumulative Probability	0.1	0.4	0.8	1
		0.1 + 0.3	0.1 + 0.3 +0.4	0.1 + 0.3 + 0.4 + 0.2



FOR CONTINUOUS RANDOM VARIABLES

The probability of a continuous random variable assuming exactly any of its values is 0.

Hence, the probability distribution for a continuous random variable cannot be given in tabular form.

The probability density function of a continuous random variable is a function which can be integrated to obtain the probability that the random variable takes a value in a given interval.

The probability for a continuous random variable is always computed at intervals : $P(a \le X \le b)$.

$$P(a < X < b) = \int_a^b f(x)\, dx.$$

The probability distribution of a continuous random variable can be stated as a formula; and f(*x*) is called the **probability density function**, or simply a **density function,** of X.

In case of a continuous random variable, f(*x*) is considered a probability density function if it satisfies the following conditions:

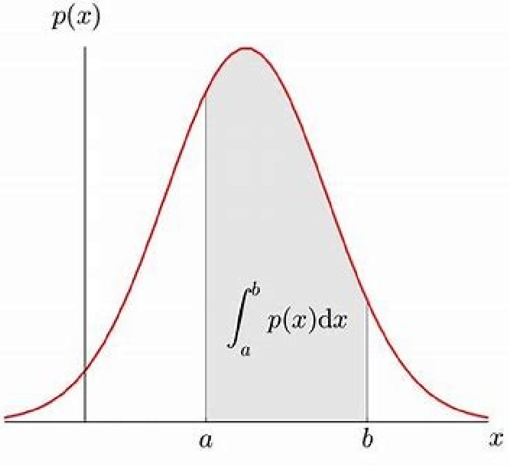
1. $f(x) \geq 0$, for all $x \in R$.
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$.
3. $P(a < X < b) = \int_a^b f(x)\, dx$.

To get **probability** from here you need to consider certain interval under the curve rather than height of the curve at certain location or single value, what we do for probability mass function.

Here probability is given by the surface area under the curve with a interval. To find the probability of a certain interval, say a to b, we find the area under that curve by integrating the PDF in that interval.

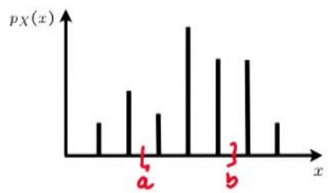
$f_x(x)$ is the Probability Density Function.

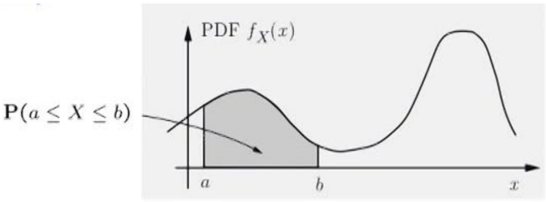
$$P(a < X < b) = \int_a^b f_X(x)\, dx$$



And the **cumulative distribution function** F(*x*) of a continuous random variable is given by:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt, \quad \text{for } -\infty < x < \infty.$$





$$\mathbf{P}(a \le X \le b) = \sum_{x:\, a \le x \le b} p_X(x)$$

$$p_X(x) \ge 0 \qquad \sum_x p_X(x) = 1$$

$$\mathbf{P}(a \le X \le b) = \int_a^b f_X(x)\, dx$$

$$f_X(x) \ge 0 \qquad \int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

Probability Density Function Example

Question:

Let X be a continuous random variable with the PDF given by:

$$f(x) = \begin{cases} x;\ 0 < x < 1 \\ 2 - x;\ 1 < x < 2 \\ 0;\ x > 2 \end{cases}$$

Find P(0.5 < x < 1.5).

Find P (0.5 < x < 1.5).

Solution:

Given PDF is:

$$f(x)=\begin{cases} x;\ 0<x<1 \\ 2-x;\ 1<x<2 \\ 0;\ x>2 \end{cases}$$

$$P(0.5<X<1.5)=\int_{0.5}^{1.5} f(x)dx$$

Let us split the integral by taking the intervals as given below:

$$=\int_{0.5}^{1} f(x)dx+\int_{1}^{1.5} f(x)dx$$

Substituting the corresponding values of f(x) based on the intervals, we get;

$$=\int_{0.5}^{1} xdx+\int_{1}^{1.5}(2-x)dx$$

Integrating the functions, we get;

$$=\left(\frac{x^2}{2}\right)_{0.5}^{1}+\left(2x-\frac{x^2}{2}\right)_{1}^{1.5}$$

$= [(1)^2/2 – (0.5)^2/2] + \{[2(1.5) – (1.5)^2/2] – [2(1) – (1)^2/2]\}$

$= [(\frac{1}{2}) – (\frac{1}{8})] + \{[3 – (9/8)] – [2 – (\frac{1}{2})]\}$

$= (\frac{3}{8}) + [(15/8) – (3/2)]$

$= (3 + 15 – 12)/8$

$= 6/8$

$= 3/4$

PROBABILITY DISTRIBUTION

Depending on type of random variables, its probability distribution can be categorized into :

1) Discrete probability distributions

2) Continuous probability distributions

DISCRETE PROBABILITY DISTRIBUTIONS

These distributions model the probabilities of random variables that can have discrete values as outcomes.

Two discrete probability distribution function are associated with such discrete random variables:

Probability Mass Function (PMF)

Cumulative Distribution Function(CDF)

For example, the possible values for the random variable X that represents the number of heads that can occur when a coin is tossed twice are the set {0, 1, 2} and not any value from 0 to 2 like 0.1 or 1.6.

Examples: Bernoulli, Binomial, Negative Binomial, Hypergeometric, etc.,

DISCRETE PROBABILITY DISTRIBUTIONS

Probability Mass function (PMF) estimate the probability of a particular outcome(discrete) for discrete distributions.

It can be defined as a function that gives the probability of a discrete random variable, X, being exactly equal to some value, x.

$f(x) = P(X = x)$

Cumulative Distribution Function gives the probability that a discrete random variable will be lesser than or equal to a particular value.

$F(x) = P(X \leq x)$

The mean of a discrete probability distribution gives the weighted average of all possible values of the discrete random variable. It is also known as the expected value.

$E[X] = \sum x \, P(X = x)$

The discrete probability distribution variance gives the dispersion of the distribution about the mean.

It can be defined as the average of the squared differences of the distribution from the mean, μ.

$Var[X] = \sum (x - \mu)^2 \, P(X = x)$

HOW TO FIND DISCRETE PROBABILITY DISTRIBUTION?

Step 1: Determine the sample space of the experiment.

Step 2: Define a discrete random variable, X.

Step 3: Identify the possible values that the variable can assume.

Step 4: Calculate the probability associated with each outcome.

Step 5: To get the discrete probability distribution represent the probabilities and the corresponding outcomes in tabular form or in graphical form.

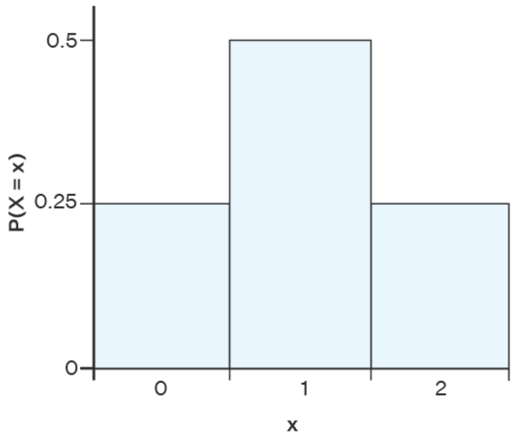
Suppose we flip a coin two times and count the number of heads (successes).

The binomial random variable is the number of heads,

1) sample space is {HH, HT, TH, TT}.
2) X be the number of heads observed.
3) There are 3 possible values of X. These are 0 (no head is observed), 1 (exactly one head is observed), and 2 (the coin lands on heads twice).
4)

x	{TT}	{TH}	{HT}	{HT, HH}
P(X = x)	1 / 4 = 0.25	1 / 4 = 0.25	1 / 4 = 0.25	1 / 4 = 0.25

x	0 {TT}	1 {HT, TH}	2 {HH}
P(X = x)	1 / 4 = 0.25	2 / 4 = 0.5	1 / 4 = 0.25



BERNOULLI DISTRIBUTION

This distribution is generated when we perform an experiment once and it has only two possible outcomes – success and failure.

The trials of this type are called Bernoulli trials, which form the basis for many distributions

Let p be the probability of success and 1 – p is the probability of failure.

The PMF is given as

$$\text{PMF} = \begin{cases} \text{p}, & \text{Success} \\ 1 - \text{p}, & \text{Failure} \end{cases}$$

Examples:

flipping a coin once. p is the probability of getting ahead and 1 – p is the probability of getting a tail.

PROPERTIES OF BINOMIAL DISTRIBUTION

Properties:

1) Binary: Each trial can have only two outcomes which can be considered as success or failure.
2) Independent : The outcome of each trial must be independent of each other.
3) Number : There must be fixed number of trials.
4) Success :The probability of success must remain the same in the trial.

Outcomes of Binomial distribution

This is generated for random variables with only two possible outcomes.

Let p denote the probability of an event is a success which implies $1 - p$ is the probability of the event being a failure.

Performing the experiment repeatedly and plotting the probability each time gives us the Binomial distribution.

Binomial formula

Note the general pattern emerging → if you have only two possible outcomes (call them 1/0 or yes/no or success/failure) in n independent trials, then the probability of exactly X "successes"=

n = number of trials

$$\boxed{\binom{n}{X} p^{X} (1-p)^{n-X}}$$

$1\text{-}p$ = probability of failure

X = # successes out of n trials

p = probability of success

Example:

Flipping a coin n number of times and calculating the probabilities of getting a particular number of heads.

More real-world examples include the number of successful sales calls for a company or whether a drug works for a disease or not.

- Number of winning lottery tickets when you buy 10 tickets of the same kind
- Number of left-handers in a randomly selected sample of 100 unrelated people

Probability mass function $P\left(X=k\right)=\left(\begin{array}{c} n \\ k \end{array}\right)p^{k}\left(1-p\right)^{n-k}$

Mean: $\mu=E\left(X\right)=np$

Variance: $\sigma^{2}=V\left(X\right)=np\left(1-p\right)$

$n=\#\ of\ \ trials$

$p=probability\ of\ \ success$

$k=\#\ of\ \ successes$

Binomial Distribution Formula

$$P(x)=\binom{n}{x}p^{x}q^{n-x}=\frac{n!}{(n-x)!\,x!}p^{x}q^{n-x}$$

For example, suppose we shuffle a standard deck of cards, and we turn over the top card. We put the card back in the deck and reshuffle. We repeat this process five times. Let X equal the number of Jacks we observe. Is this a binomial distribution?

- **B** – binary – yes, either it's a Jack or it isn't
- **I** – independent – yes, because we replace the card each time, the trials are independent.
- **N** – number of trials fixed in advance – yes, we are told to repeat the process five times.
- **S** – successes (probability of success) are the same – yes, the likelihood of getting a Jack is 4 out of 52 each time you turn over a card.
- Therefore, this is an example of a binomial distribution.

Suppose that Charlie makes a free throw has probability of 0.82 on any one try. Assuming that this probability doesn't change, find the chance that Charlie makes 4 out of the next seven free throws.

$$n = 7 \ \ shots$$

$$p = 0.82$$

$$k = 4 \ \ free\ throws$$

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

$$P(X = 4) = \binom{7}{4} 0.82^4 (1 - 0.82)^{7-4}$$

$$P(X = 4) = \left(\frac{7!}{4!(7-4)!}\right) 0.82^4 (0.18)^3$$

$$P(X = 4) = (35) 0.82^4 (0.18)^3 = 0.0923$$

let's determine the number of free throws Charlie should expect to make and the standard deviation.

Mean:	$\mu = E(X) = np$ $E(X) = (7)(0.82) = 5.74$
Variance:	$\sigma^2 = V(X) = np(1-p)$ $V(X) = (7)(0.82)(1-0.82) = 1.033$
Standard Deviation	$\sigma = \sqrt{\sigma^2} = \sqrt{1.033} = 1.016$

$n = 7 \ \ shots$

$p = 0.82$

$k = 4 \ \ free \ throws$

2. A student takes a ten question multiple choice exam for which there are four choices for each question. The student is unprepared and relies on guessing for each question.

(a) Find the probability that the student guesses half the questions correctly.

$P(x) = {}_nC_x \cdot p^x \cdot q^{n-x}$

$n = 10$
$p = .25$
$q = .75$
$x = 5$

$P(x=5) = {}_{10}C_5 \cdot (.25)^5 \cdot (.75)^5$

$P(x=5) = .058$

(b) Find the probability that the student earns a passing grade of 65% or greater on the exam.

4 5

Binomial distribution: example

- If I toss a coin 20 times, what's the probability of getting of getting 2 or fewer heads?

$$\binom{20}{0}(.5)^0(.5)^{20} = \frac{20!}{20!0!}(.5)^{20} = 9.5x10^{-7} +$$

$$\binom{20}{1}(.5)^1(.5)^{19} = \frac{20!}{19!1!}(.5)^{20} = 20x9.5x10^{-7} = 1.9x10^{-5} +$$

$$\binom{20}{2}(.5)^2(.5)^{18} = \frac{20!}{18!2!}(.5)^{20} = 190x9.5x10^{-7} = 1.8x10^{-4}$$

$$= 1.8x10^{-4}$$

Binomial Distribution Example

- Example: 35% of all voters support Proposition A. If a random sample of 10 voters is polled, what is the probability that exactly three of them support the proposition?

i.e., find P(x = 3) if n = 10 and p = 0.35 :

$$P(x=3) = \frac{n!}{x!(n-x)!} p^x q^{n-x} = \frac{10!}{3!7!}(0.35)^3(0.65)^7 = 0.2522$$

There is a 25.22% chance that exactly 3 out of the 10 voters will support Proposition A

Choosing your numbers:
N = 3 and X = 2,

We get:

(1) Exactly 2 out of 3 adults have college degree =

$$P(2) = \binom{3}{2} 0.7^{3-2} 0.3^2 = 3 \times 0.7 \times 0.09 = 0.189$$

(2) Less than 2 out of 3 adults have college degree:

$$P(i < 2) = \sum_{i=0}^{2-1} \binom{3}{i} 0.7^{3-i} 0.3^i$$

$= 0.7^3 + 3 \text{ X } 0.49 \text{ X } 0.3 = 0.343 + 0.441 = 0.784$

(3) Greater than 2 out of selected 3 adults have college degree:

$$P(i > 2) = \sum_{i=2+1}^{3} \binom{3}{i} 0.7^{3-i} 0.3^i$$

$= 0.3^3 = 0.027$

POISSON DISTRIBUTION

Itdescribes the events that occur in a fixed interval of time or space.

Examples:

Consider the case of the number of calls received by a customer care center per hour. We can estimate the average number of calls per hour but we cannot determine the exact number and the exact time at which there is a call. Each occurrence of an event is independent of the other occurrences.

The PMF is given as,

$$P(X = x) = \frac{e^{-\lambda} \lambda^{x}}{x!}$$

where λ is the average number of times the event has occurred in a certain period of time,

x is the poisson random variable (with desired outcome)

and e is the base of logarithm , Euler's number, and e = 2.71828 (approx).

PROPERTIES OF POISSON DISTRIBUTION

The occurrence of the event are independent in an interval.

An infinite number of occurrences of the of the event are possible in the interval.

The probability of a single event in the interval is propotional to the length of the event.

In an infinitely small portion of the interval, the probability of more than one occurrence of the event is negligible.

The Poisson distribution is limited when the number of trials n is indefinitely large.

If the mean is large, then the Poisson distribution is approximately a normal distribution.

In Poisson distribution, the mean is represented as $\mu = E(X) = \lambda$.

The mean and the variance of Poisson Distribution are equal. It means that $E(X) = V(X)$

Where,

$V(X)$ is the variance.

The standard deviation is always equal to the square root of the mean μ.

APPLICATIONS OF POISSON DISTRIBUTION

- To count the number of defects of a finished product
- To count the number of deaths in a country by any disease or natural calamity
- To count the number of infected plants in the field
- To count the number of bacteria in the organisms or the radioactive decay in atoms
- To calculate the waiting time between the events.

If the random variable X follows a Poisson distribution with mean 3.4, find P(X=6)?

The number of industrial injuries per working week in a particular factory is known to follow Poisson Distribution with mean 0.5

Find the probability that

1. In a particular week, there will be

   Less than 2 accidents

   More than 2 accidents

2. In a three week period, there will be no accidents

2. In a three week period, there will be no accidents

Solution

Let A be the number of accidents in one week so A- Po (0.5)

(a) (i) P(A<2) = P(A≤1)

= 0.9098 (from tables in Appendix 3(p257), to 4 d.p.)

or, from the formula,

P(A<2) = P(A=0) + P(A=1)

$$= e^{-0.5} + \frac{e^{-0.5}0.5}{1!}$$

$$= \frac{3}{2}e^{-0.5}$$

= 0.9098.

(ii) P(A<2) = 1 - P(A≤2)

= 1 - 0.9856 (from tables)

= 0.0144 (to 4 d.p.)

or

$$1\text{-}P(A=0)+P(A=1)+P(A=2)$$

$$= 1\text{-}e^{-0.5} + e^{-0.5}0.5 + \frac{e^{-0.5}0.5}{2!}$$

(ii) P(A<2) = 1 - P(A≤2)

= 1 - 0.9856 (from tables)

= 0.0144 (to 4 d.p.)

or

1-$P(A=0)+P(A=1)+P(A=2)$

= 1-$e^{-0.5}$ + $e^{-0.5}$0.5 + $\frac{e^{-0.5}0.5}{2!}$

= 1-$e^{-0.5}$(1+0.5+0.125)

= 1-1.625 $e^{-0.5}$

= 0.0144.

(b) P(0 in 3 weeks) = $(e^{-0.5})^3$ = 0.223

In a cafe, the customer arrives at a mean rate of 2 per min. Find the probability of arrival of 5 customers in 1 minute using the Poisson distribution formula.

Solution:

Given: $\lambda = 2$, and $x = 5$.

Using the Poisson distribution formula:

$P(X = x) = (e^{-\lambda} \lambda^{x})/x!$

$P(X = 5) = (e^{-2} 2^{5})/5!$

$P(X = 6) = 0.036$

Answer: The probability of arrival of 5 customers per minute is 3.6%.

Find the mass probability of function at x = 6, if the value of the mean is 3.4.

Solution:

Given: $\lambda = 3.4$, and $x = 6$.

Using the Poisson distribution formula:

$P(X = x) = (e^{-\lambda} \lambda^{x})/x!$

$P(X = 6) = (e^{-3.4} 3.4^{6})/6!$

$P(X = 6) = 0.072$

Answer: The probability of function is 7.2%.

If 3% of electronic units manufactured by a company are defective. Find the probability that in a sample of 200 units, less than 2 bulbs are defective.

Solution:

The probability of defective units p = 3/100 = 0.03

Give n = 200.

We observe that p is small and n is large here. Thus it is a Poisson distribution.

Mean λ= np = 200 × 0.03 = 6

P(X= x) is given by the Poisson Distribution Formula as $(e^{-\lambda} \lambda^{x})/x!$

P(X < 2) = P(X = 0) + P(X= 1)

$=(e^{-6} 6^{0})/0! + (e^{-6}6^{1})/1!$

$= e^{-6} + e^{-6} \times 6$

= 0.00247 + 0.0148

P(X < 2) = 0.01727

Answer: The probability that less than 2 bulbs are defective is 0.01727

CONTINUOUS PROBABILITY DISTRIBUTIONS

These distributions model the probabilities of random variables that can have any possible outcome, also real.

Two continuous probability distribution function are associated with such continuous random variables:

Probability Density Function (PDF)

Cumulative Density/Distribution Function(CDF)

For example, the possible values for the random variable X that represents weights of citizens in a town which can have any value like 34.5, 47.7, etc.,

Examples: Normal, Student’s T, Chi-square, Exponential, etc.,

PROBABILITY DENSITY FUNCTION

Probability Density function (PDF) estimate the probability that it lies within a particular range of values for any given outcome(continuous) for continuous distributions.

CUMULATIVE DISTRIBUTION FUNCTION

The Cumulative Distribution Function of X, evaluated at x is the probability that X will take a value less than or equal to x.

CONTINUOUS PROBABILITY DISTRIBUTIONS

When working with continuous random variables, such as X, we only calculate the probability that X lie within a certain interval;

like $P(X \le k)$ or $P(a \le X \le b)$.

We don't calculate the probability of X being equal to a specific value k.

In fact the result, $P(X=k)=0$, will always be true:

This can be explained by the fact that the total number of possible values of a continuous random variable X is infinite, so the likelihood of any one single outcome tends towards 0.

CONTINUOUS PROBABILITY DISTRIBUTIONS

The idea is to integrate the probability density function $f(x)$ to define a new function $F(x)$, known as the cumulative density function.

To calculate the probability that X be within a certain range,

say $a \le X \le b$, we calculate $F(b) - F(a)$, using the cumulative density function.

Put "simply" we calculate probabilities as:

$$P(a \le X \le b) = \int_a^b f(x)dx$$

where $f(x)$ is the variable's probability density function.

NORMAL DISTRIBUTION

It has two parameters namely mean and standard deviation.

The mean has the highest probability and all other values are distributed equally on either side of the mean in a symmetric fashion.

The standard normal distribution is a special case where the mean is 0 and the standard deviation of 1.

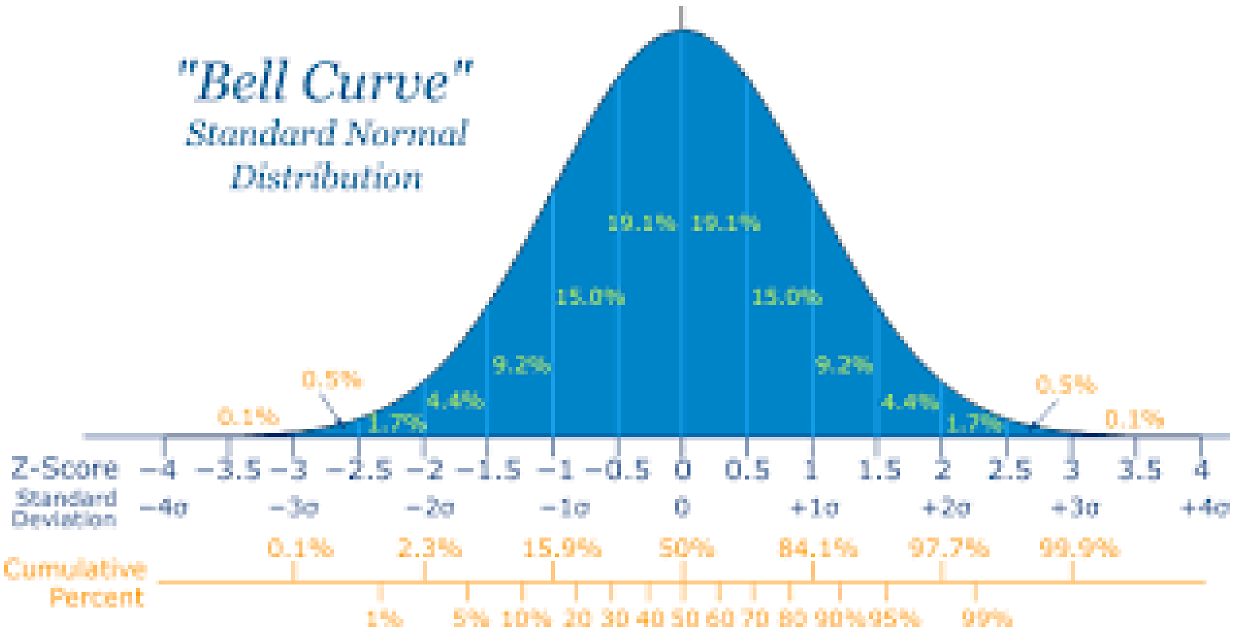
68% of the values are 1 standard deviation away, 95% percent of them are 2 standard deviations away, 99.7% are 3 standard deviations away from mean.

NORMAL DISTRIBUTION

The PDF is given by,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right)$$

where μ is the mean of the random variable X and σ is the standard deviation.



NORMAL DISTRIBUTION

The random variable of a standard normal distribution is known as the standard score or a z-score.

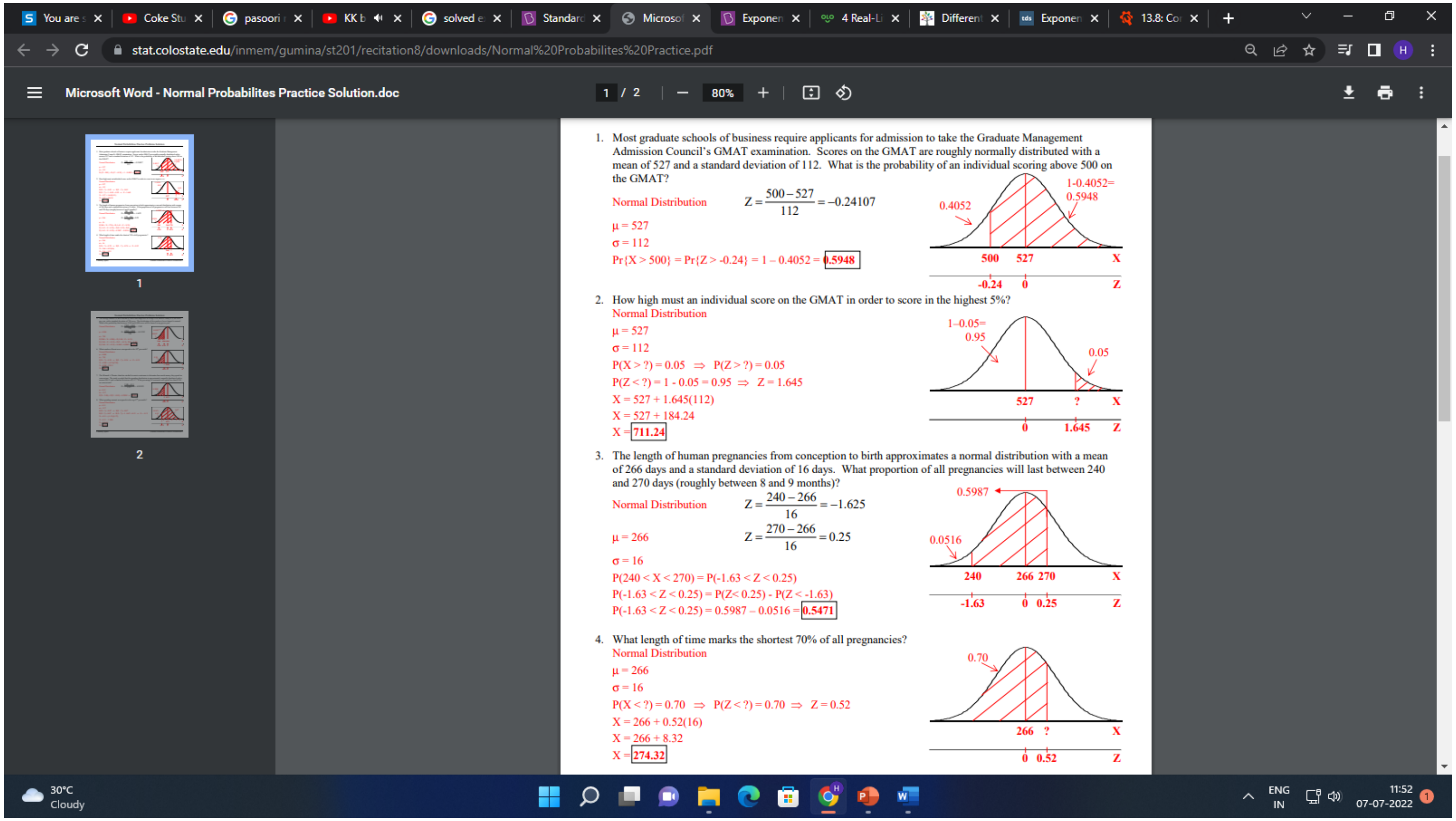
It is possible to transform every normal random variable X into a z score using the following formula:

$z = (X - \mu) / \sigma$

where X is a normal random variable, μ is the mean of X, and σ is the standard deviation of X.

In probability theory, the normal or Gaussian distribution is a very common continuous probability distribution.

NORMAL DISTRIBUTION



1. Most graduate schools of business require applicants for admission to take the Graduate Management Admission Council's GMAT examination. Scores on the GMAT are roughly normally distributed with a mean of 527 and a standard deviation of 112. What is the probability of an individual scoring above 500 on the GMAT?

Normal Distribution

$$Z = \frac{500-527}{112} = -0.24107$$

$\mu = 527$

$\sigma = 112$

$\Pr\{X > 500\} = \Pr\{Z > -0.24\} = 1 - 0.4052 = \boxed{0.5948}$

2. How high must an individual score on the GMAT in order to score in the highest 5%?

Normal Distribution

$\mu = 527$

$\sigma = 112$

$P(X > ?) = 0.05 \Rightarrow P(Z > ?) = 0.05$

$P(Z < ?) = 1 - 0.05 = 0.95 \Rightarrow Z = 1.645$

$X = 527 + 1.645(112)$

$X = 527 + 184.24$

$X = \boxed{711.24}$

3. The length of human pregnancies from conception to birth approximates a normal distribution with a mean of 266 days and a standard deviation of 16 days. What proportion of all pregnancies will last between 240 and 270 days (roughly between 8 and 9 months)?

Normal Distribution

$$Z = \frac{240-266}{16} = -1.625$$

$\mu = 266$

$$Z = \frac{270-266}{16} = 0.25$$

$\sigma = 16$

$P(240 < X < 270) = P(-1.63 < Z < 0.25)$

$P(-1.63 < Z < 0.25) = P(Z < 0.25) - P(Z < -1.63)$

$P(-1.63 < Z < 0.25) = 0.5987 - 0.0516 = \boxed{0.5471}$

4. What length of time marks the shortest 70% of all pregnancies?

Normal Distribution

$\mu = 266$

$\sigma = 16$

$P(X < ?) = 0.70 \Rightarrow P(Z < ?) = 0.70 \Rightarrow Z = 0.52$

$X = 266 + 0.52(16)$

$X = 266 + 8.32$

$X = \boxed{274.32}$

NORMAL DISTRIBUTION

3. The length of human pregnancies from conception to birth approximates a normal distribution with a mean of 266 days and a standard deviation of 16 days. What proportion of all pregnancies will last between 240 and 270 days (roughly between 8 and 9 months)?

Normal Distribution

$$Z = \frac{240 - 266}{16} = -1.625$$

$\mu = 266$

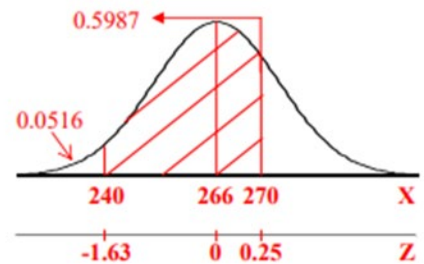
$$Z = \frac{270 - 266}{16} = 0.25$$

$\sigma = 16$

$P(240 < X < 270) = P(-1.63 < Z < 0.25)$

$P(-1.63 < Z < 0.25) = P(Z < 0.25) - P(Z < -1.63)$

$P(-1.63 < Z < 0.25) = 0.5987 - 0.0516 = \boxed{0.5471}$



4. What length of time marks the shortest 70% of all pregnancies?

Normal Distribution

$\mu = 266$

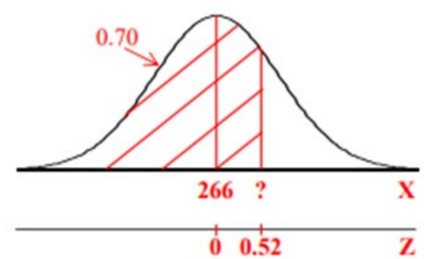
$\sigma = 16$

$P(X < ?) = 0.70 \Rightarrow P(Z < ?) = 0.70 \Rightarrow Z = 0.52$

$X = 266 + 0.52(16)$

$X = 266 + 8.32$

$X = \boxed{274.32}$



NORMAL DISTRIBUTION

5. The average number of acres burned by forest and range fires in a large New Mexico county is 4,300 acres per year, with a standard deviation of 750 acres. The distribution of the number of acres burned is normal. What is the probability that between 2,500 and 4,200 acres will be burned in any given year?

Normal Distribution $\quad Z = \dfrac{2500 - 4300}{750} = -2.40$

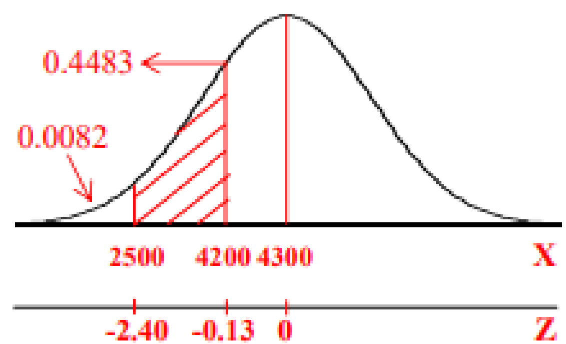
$\mu = 4300 \quad Z = \dfrac{4200 - 4300}{750} = -0.13333$

$\sigma = 750$

$P(2500 < X < 4200) = P(-2.40 < Z < -0.13)$

$P(-2.40 < Z < -0.13) = P(Z < -0.13) - P(Z < -2.40)$

$P(-2.40 < Z < -0.13) = 0.4483 - 0.0082 = \boxed{0.4401}$



6. What number of burnt acres corresponds to the 38th percentile?

Normal Distribution

$\mu = 4300$

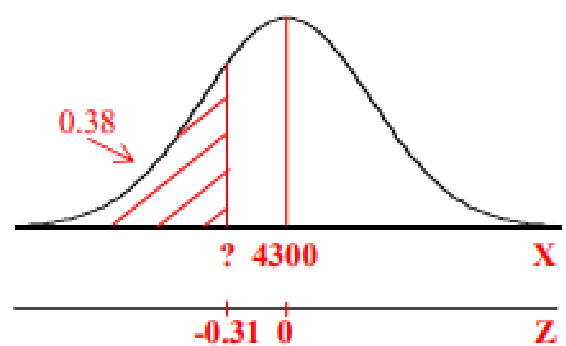
$\sigma = 750$

$P(X < ?) = 0.38 \Rightarrow P(Z < ?) = 0.38 \Rightarrow Z = -0.31$

$X = 4300 + (-0.31)(750)$

$X = 4300 - 232.5$

$X = \boxed{4067.5}$



NORMAL DISTRIBUTION

concessions. The study revealed that the spending distribution is approximately normally distributed with a mean of \$4.11 and a standard deviation of \$1.37. What percentage of customers will spend less than \$3.00 on concessions?

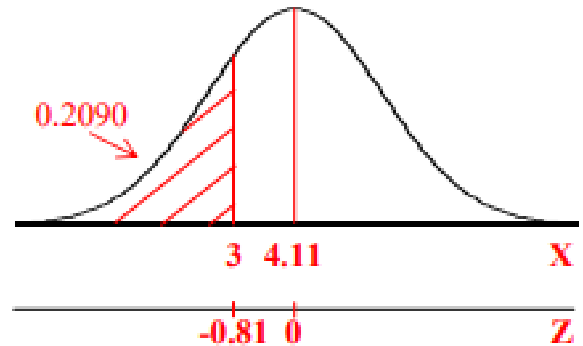
Normal Distribution

$$Z = \frac{3.00 - 4.11}{1.37} = -0.81021$$

$\mu = 4.11$

$\sigma = 1.37$

$P(X < 3.00) = P(Z < -0.81) = 0.2090 \Rightarrow \boxed{20.9\%}$



8. What spending amount corresponds to the top 87th percentile?

Normal Distribution

$\mu = 4.11$

$\sigma = 1.37$

$P(X > ?) = 0.87 \Rightarrow P(Z > ?) = 0.87$

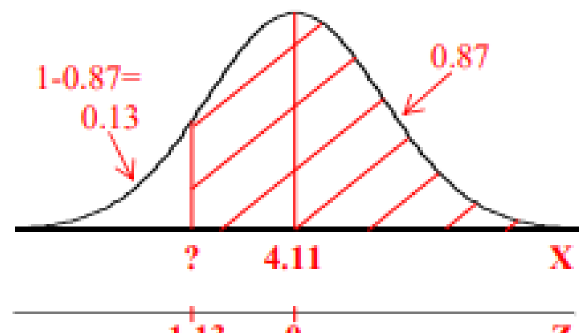
$P(Z > ?) = 0.87 \Rightarrow P(Z < ?) = 1 - 0.87 = 0.13 \Rightarrow Z = -1.13$

$X = 4.11 + (-1.13)(1.37)$

$X = 4.11 - 1.5481$

$X = 2.5619$

$X = \boxed{\$2.56}$



EXPONENTIAL DISTRIBUTION

To predict the amount of waiting time until the next event in a Poisson process (i.e., success, failure, arrival, etc.).

For example, we want to predict the following:

- The amount of time until the customer finishes browsing and actually purchases something in your store (success).
- The amount of time until the hardware on AWS EC2 fails (failure).
- The amount of time you need to wait until the bus arrives (arrival).

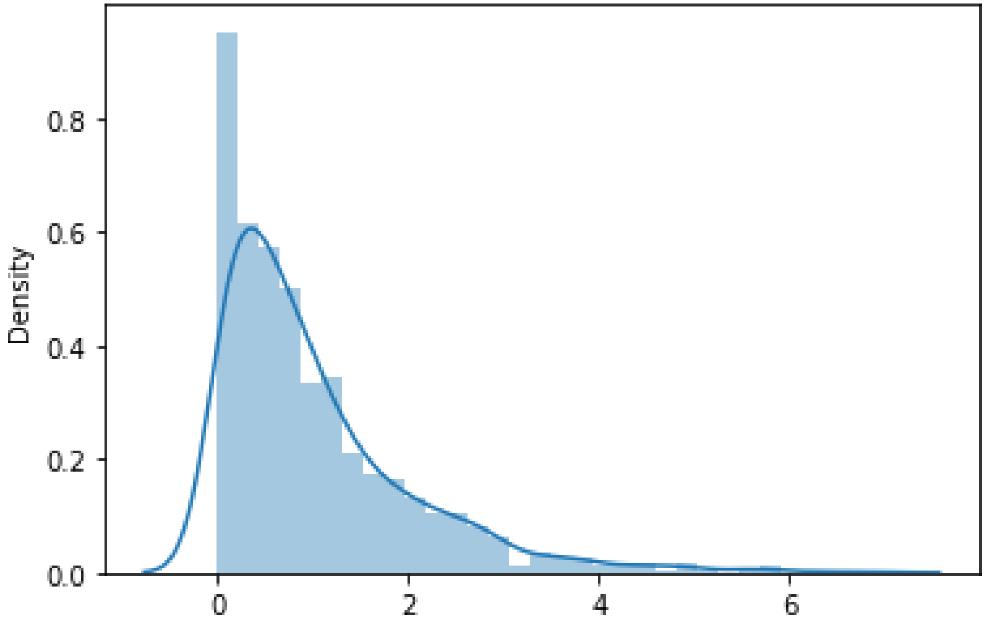
EXPONENTIAL DISTRIBUTION

$$f(x) = \lambda e^{-\lambda x}$$

$$f_X(x|\lambda) = \begin{cases} \lambda e^{-\lambda x} & for\ x > 0 \\ 0 & for\ x \leq 0 \end{cases}$$

The PDF is given by,

where λ is the rate parameter. λ = 1/(average time between events) = 1/μ
and e=2.71828
The mean of the exponential distribution is 1/λ.
And the variance of the exponential distribution is $1/\lambda^2$.



EXPONENTIAL DISTRIBUTION

For example, suppose the mean number of minutes between eruptions for a certain geyser is 40 minutes. If a geyser just erupts, what is the probability that we'll have to wait less than 50 minutes for the next eruption?

To solve this, we need to calculate rate parameter:

- $\lambda = 1/\mu \Rightarrow \lambda = 1/40 \Rightarrow \lambda = .025$
- plug in $\lambda = .025$ and $x = 50$ to the formula for the CDF:
- $P(X \leq x) = 1 - e^{-\lambda x} \Rightarrow P(X \leq 50) = 1 - e^{-.025(50)}$
- $P(X \leq 50) = 0.7135$

EXPONENTIAL DISTRIBUTION

Assume that you usually get 2 phone calls per hour. calculate the probability, that a phone call will come within the next hour.

Solution:

It is given that, 2 phone calls per hour.

So, it would expect that one phone call at every half-an-hour.

So, we can take λ = 0.5

follows:

$$p(0 \leq X \leq 1) = \sum_{x=0}^{1} 0.5e^{-0.5x}$$

$$= 0.393469$$

Therefore, the probability of arriving the phone calls within the next hour is 0.393469