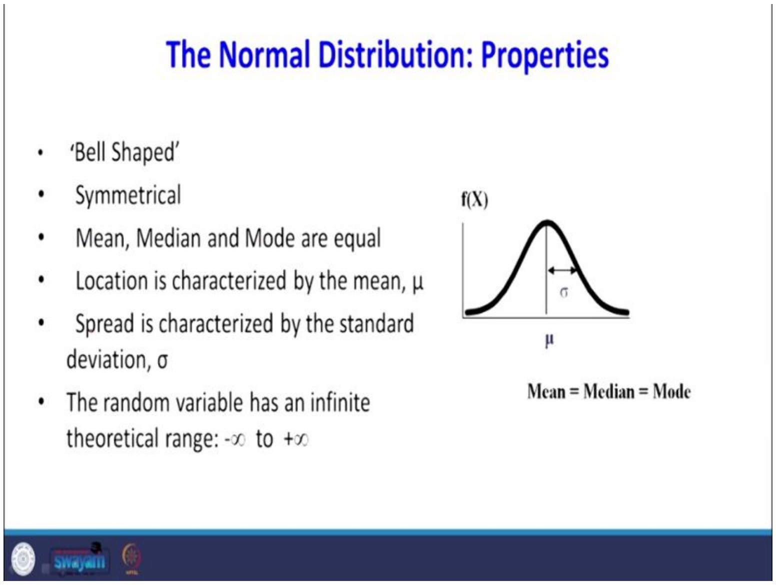
Data Analytics with Python
Prof. Ramesh Anbanandam
Department of Management Studies
Indian Institute of Technology Roorkee

Lecture – 10
Probability Distributions - III

Welcome back students now we are going to discuss another important continuous distribution and that is normal distribution, normal distribution can be called as mother of all distribution because, if you any phenomena if you are not aware about the nature of the distributions, you can assume that it follow normal distribution, most of the statistical test or whatever analytical tools which are going to use in this course.

Good to have some assumptions that did follow normal distributions knowing the properties and behavior and assumptions about the normal distribution is very important for this course. Some of the properties normal distribution is Bell shaped curved.

(Refer Slide Time: 01:16)

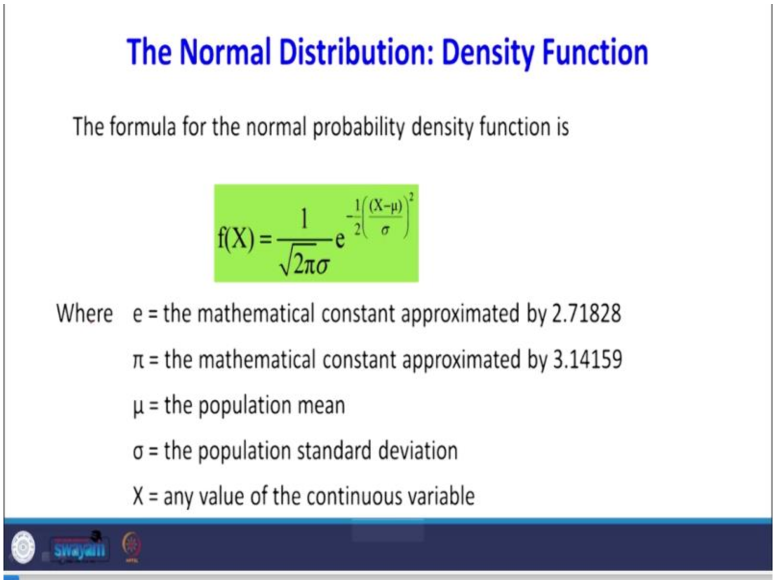


The Normal Distribution: Properties

- 'Bell Shaped'
- Symmetrical
- Mean, Median and Mode are equal
- Location is characterized by the mean, μ
- Spread is characterized by the standard deviation, σ
- The random variable has an infinite theoretical range: -∞ to +∞

Right we curve form a bell shaped curve it is symmetrical, you can fold it so after folding both the side are same another important property mean, median and modes are equal the location is characterized by its mean mu the spread is characterized by standard deviation. The random variable has an infinite theoretical rates that is a minus infinity to plus infinity.

(Refer Slide Time: 01:48)



The Normal Distribution: Density Function

The formula for the normal probability density function is

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

Where e = the mathematical constant approximated by 2.71828

π = the mathematical constant approximated by 3.14159

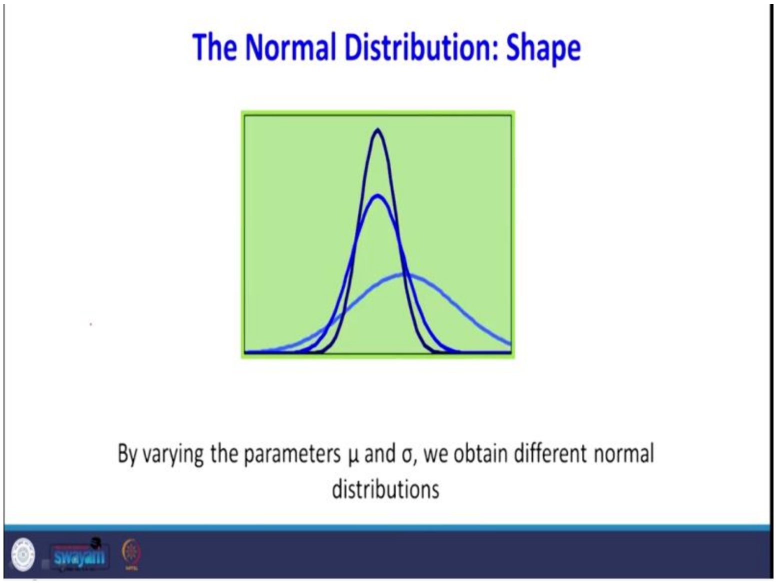
μ = the population mean

σ = the population standard deviation

X = any value of the continuous variable

The formula for normal probability density function is f of X = 1 / 2 pi into sigma e to the more one minus 1 / 2 X - mu / sigma whole square, where e is the mathematical constant, the value is 2.71828 pi is the mathematical constant the value is 3.14 mu is the population mean sigma is the population standard deviation, X is any value of the continuous variable.

(Refer Slide Time: 02:18)

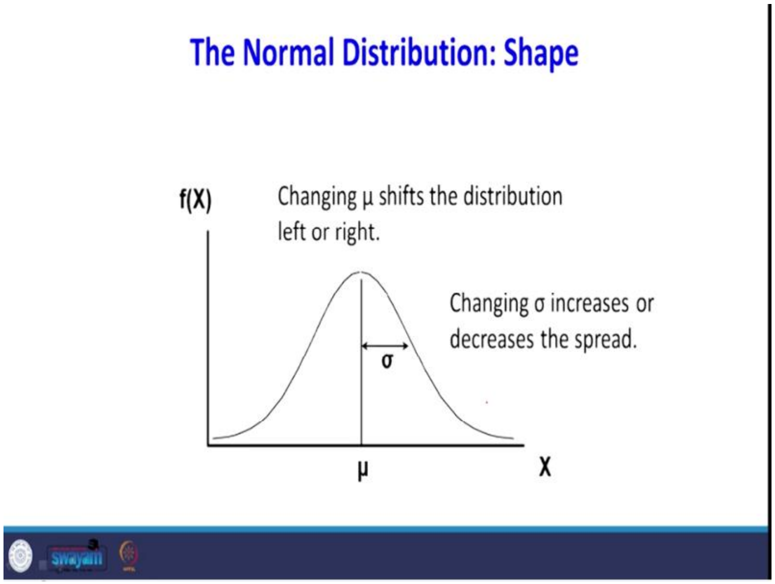


The Normal Distribution: Shape

By varying the parameters μ and σ, we obtain different normal distributions

The shape of the normal distribution will change based on its spread by varying parameters mu and sigma big obtain different normal distributions. For example this one where the sigma is very low this case sigma is literal normal, this is sigma is very big.

(Refer Slide Time: 02:39)



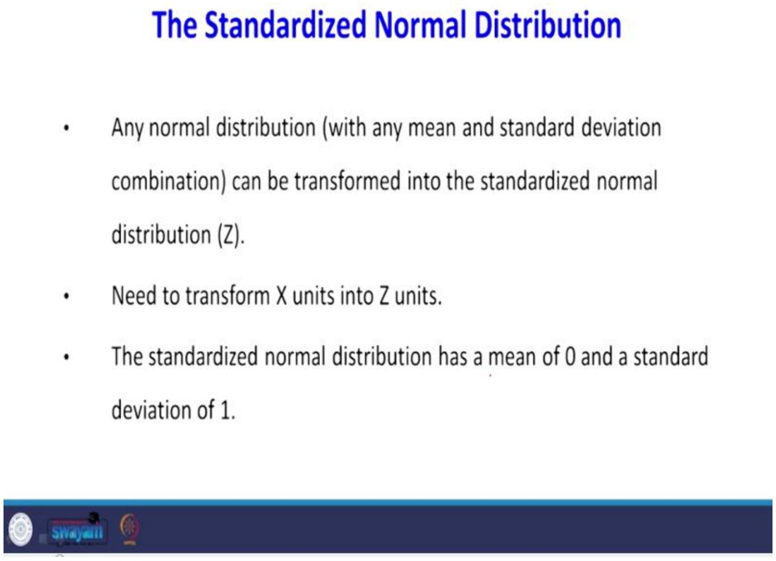
The Normal Distribution: Shape

Changing μ shifts the distribution left or right.

Changing σ increases or decreases the spread.

Changing mu shift to the distribution left or right if you increase the value of mu, it can go right side or left side. Changing sigma standard deviation increases or decreases the spread generally, when you decrease the sigma the spread will decrease when you increase the sigma the spread will increase.

(Refer Slide Time: 03:01)

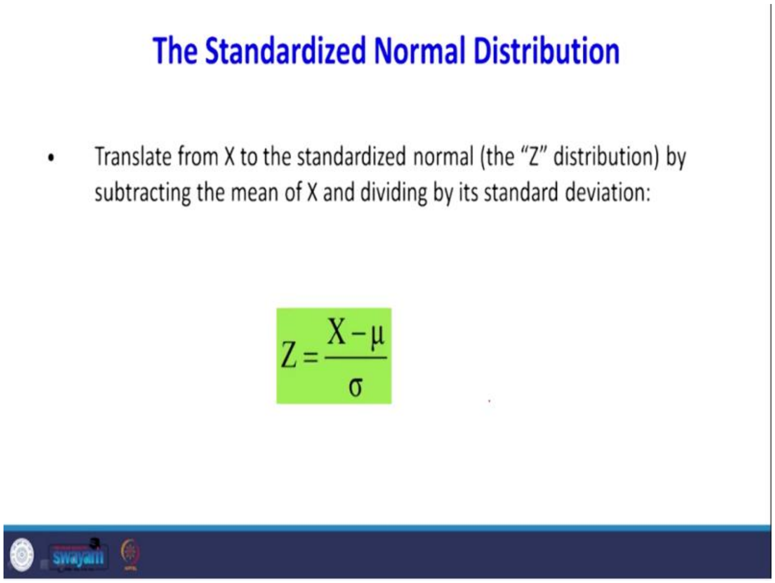


The Standardized Normal Distribution

- Any normal distribution (with any mean and standard deviation combination) can be transformed into the standardized normal distribution (Z).
- Need to transform X units into Z units.
- The standardized normal distribution has a mean of 0 and a standard deviation of 1.

There is another normal distribution standardized normal distribution, any normal distribution with the mean and standard deviation combination can be transformed into standardized normal distribution. One thing what you had to do, we need to transform X unit into Z units, Z is nothing but the conversion method is X - mu / sigma the standardized normal distribution as means 0 and the variance or standard deviation is 1.

(Refer Slide Time: 03:36)



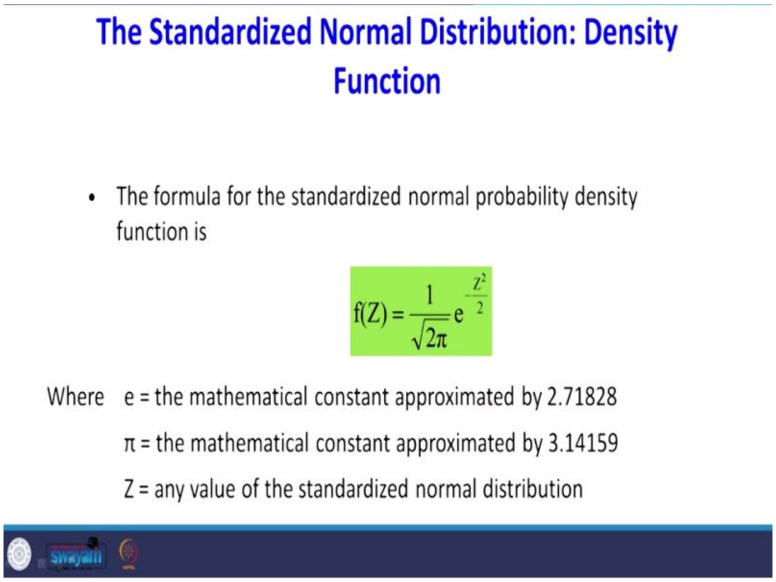
The Standardized Normal Distribution

- Translate from X to the standardized normal (the “Z” distribution) by subtracting the mean of X and dividing by its standard deviation:

$$Z = \frac{X - \mu}{\sigma}$$

The translate from X to the standardized normal that is the Z distribution by subtracting the mean of the X and dividing by standard deviation. So, that conversion from it is normal distribution to standardized normal distribution is done with the help of this Z transformation Where Z = X – mu / sigma, X is a random variable mu is the mean of the population. sigma is the standard deviation of the population.

(Refer Slide Time: 04:05)



The Standardized Normal Distribution: Density Function

- The formula for the standardized normal probability density function is

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

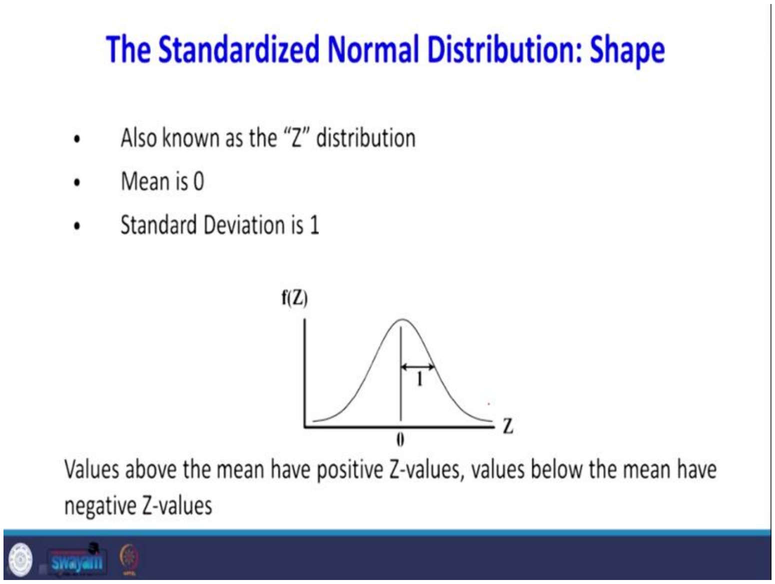
Where e = the mathematical constant approximated by 2.71828

π = the mathematical constant approximated by 3.14159

Z = any value of the standardized normal distribution

The formula for the standardized normal probability density function, if you substitute Z equal to X – mu / sigma in our previous equation and it become if f of Z = 1 / root of 2 pi e power - z square / 2, where pi is the mathematical constant z is any value of this standardized normal distribution.

(Refer Slide Time: 04:26)



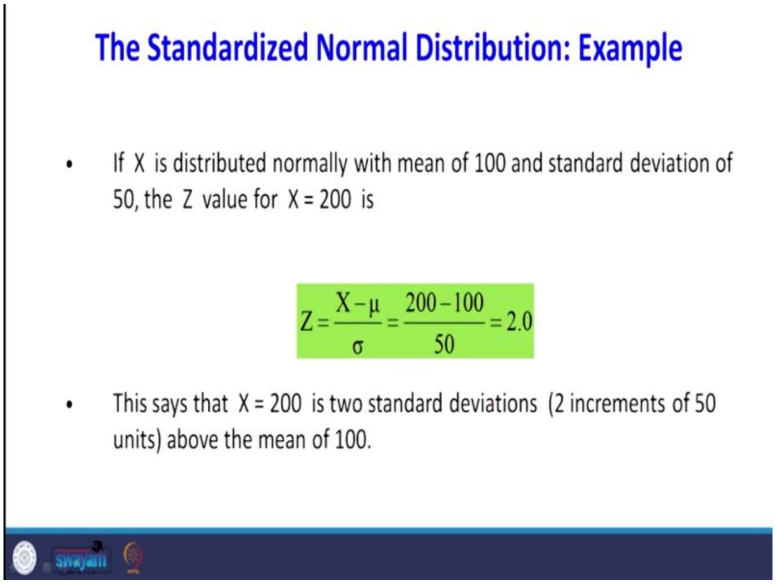
The Standardized Normal Distribution: Shape

- Also known as the “Z” distribution
- Mean is 0
- Standard Deviation is 1

Values above the mean have positive Z-values, values below the mean have negative Z-values

Standardized normal distribution the shape how they look like also known as Z distribution mean is 0 standard deviation is 1, the value above the mean how positives Z value, values below the mean will have negatives Z value.

(Refer Slide Time: 04:44)



The Standardized Normal Distribution: Example

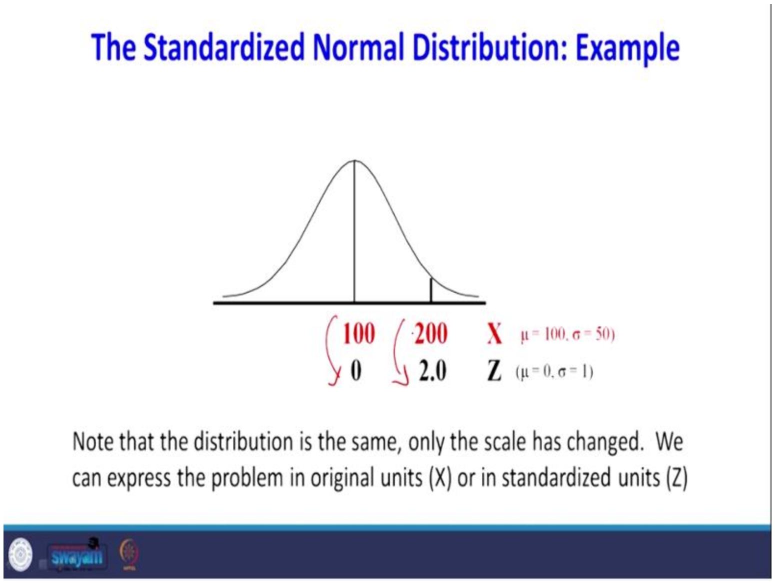
- If X is distributed normally with mean of 100 and standard deviation of 50, the Z value for X = 200 is

$$Z = \frac{X-\mu}{\sigma} = \frac{200-100}{50} = 2.0$$

- This says that X = 200 is two standard deviations (2 increments of 50 units) above the mean of 100.

Let us see how to do that conversion from normal distribution to standardized normal distribution. If X is distributed normally with the mean of 100 and standard deviation of 50 the Z value of X is 200 then corresponding Z value is X - mu / sigma X is 200 - mu 100 divided by sigma 50 equal to 2.0. This says that X 200 is 2 standard deviation above the mean of 100 that is 2 increments of 50 units, the Z value nothing but how many times of it is standard deviation that is nothing but your Z here 2 increments of 50 that is why the Z value is 2.

(Refer Slide Time: 05:41)



The Standardized Normal Distribution: Example

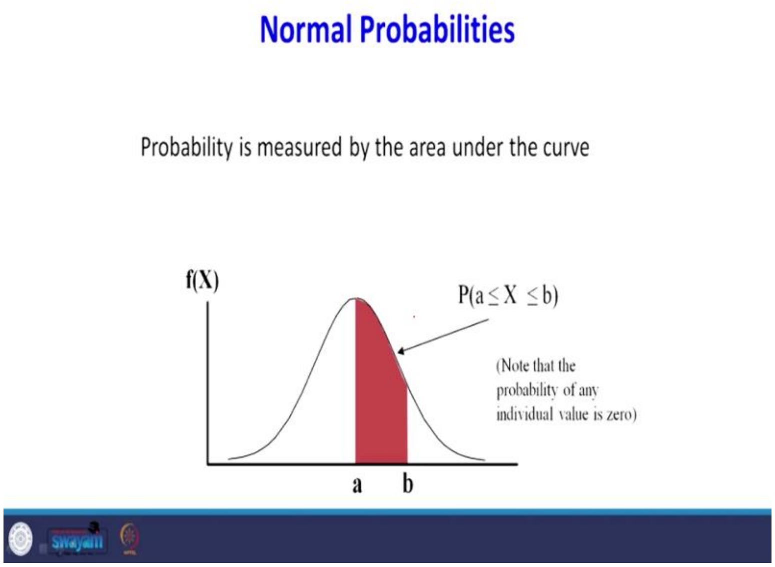
Note that the distribution is the same, only the scale has changed. We can express the problem in original units (X) or in standardized units (Z)

Look at the conversion now this will be so convenient for you the red one where the mean 0 the X 200, we have asked to find out when X = 200 what is the corresponding Z value? The red will shows in the simple normal distribution, the black one shows the standardized normal distribution, you see that the mean of the distribution is under mu standardized scale it becomes 0 when X = 200 in a normal distribution in a standardized normal distribution.

The X and corresponding Z value is 2 where the mean mu equal to 0 sigma equal 1. Note that the distribution is the same only the scale is has changed. We can express the problem in original units are standardized units but there is an advantage why we have to convert into standardized normal distribution sometime you may be required to find out the area of a distributions. Because if you are not standardizing you cannot use that your Z table, Z statistical table every time to know the area you have to integrate.

That is a very compression process that is why every normal distribution is converted to standardized normal distribution for the convenient of looking at the Z value directly from the table that will simplify our task.

(Refer Slide Time: 07:13)



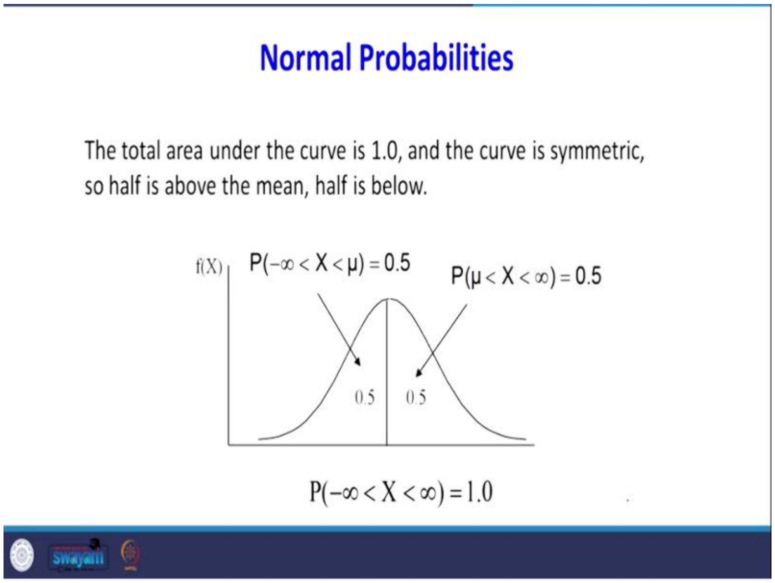
Normal Probabilities

Probability is measured by the area under the curve

(Note that the probability of any individual value is zero)

The probability is measured by area under the curve in a continuous distribution, the probability you know that it is measured area under the curve suppose, always it has to be expressed between A and B. If you want to know the probability exactly at A are exactly B that will not form the area. So, the probability is 0. So in the context of continuous distribution, the meaning of probabilities area under the curve, but if it is a discrete probability distributions, the probability can be red directly by looking at the X and corresponding P of X.

(Refer Slide Time: 07:53)



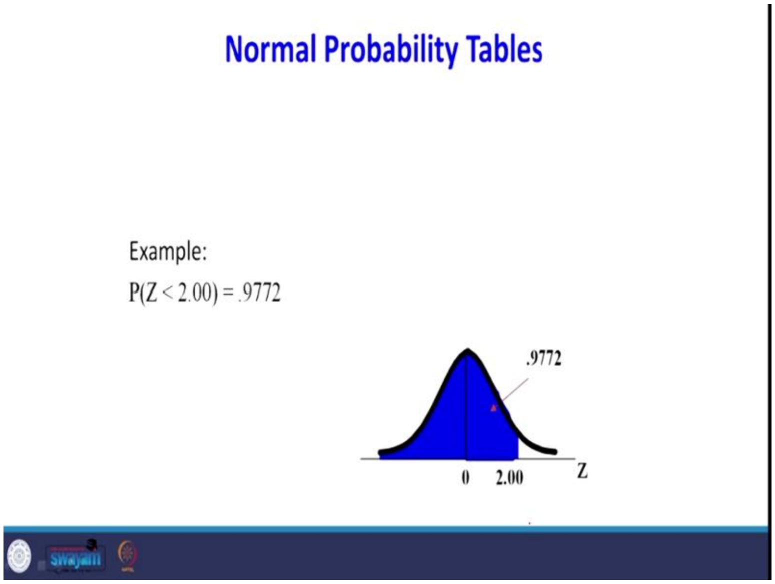
Normal Probabilities

The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below.

$$P(-\infty < X < \infty) = 1.0$$

Total area under the curve is 1 and the curve is symmetric, so half is the above the mean half is below. So P of minus infinity to X less than equal to mu is 0.5 similarly, mu less than equal to X less than equal to plus infinity is 0.5, so the total area is 1.

(Refer Slide Time: 08:16)



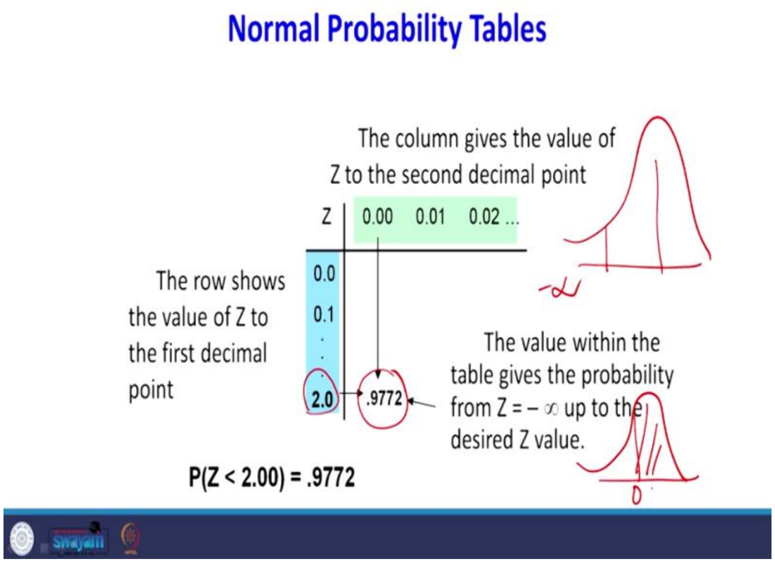
Normal Probability Tables

Example:

P(Z < 2.00) = .9772

Suppose, if you want to know the area Z less than 2.00 see this was when the Z is lesser this area is 0.9772.

(Refer Slide Time: 08:30)



Normal Probability Tables

The column gives the value of Z to the second decimal point

The row shows the value of Z to the first decimal point

The value within the table gives the probability from Z = – ∞ up to the desired Z value.

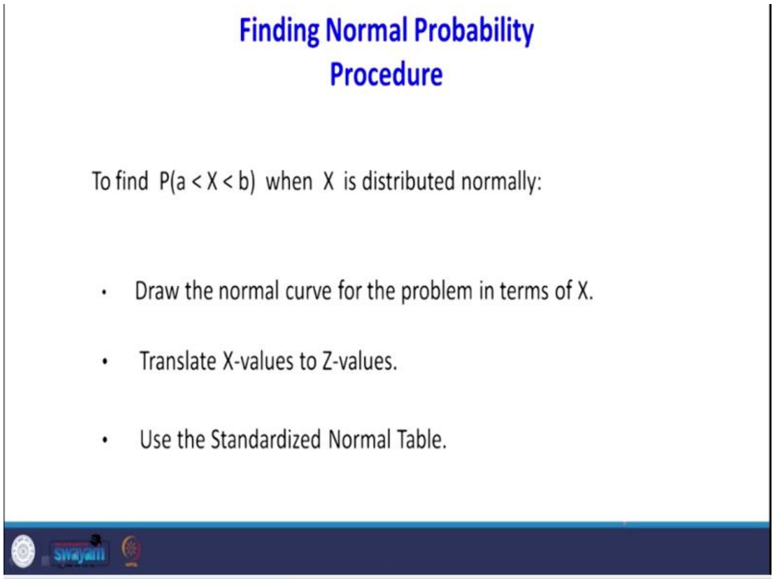
P(Z < 2.00) = .9772

One way you can read it directly from the Z table suppose in the rows, the Z value is given the column the decimal of it is given. Suppose if you want to know Z 2.00 you have to look at in row 2.00, the corresponding area is this one. See, the rows shows the value of Z to the first decimal

.the column gives the value of Z to the second decimal .the value within the table gives the probability from Z minus infinity up to desired Z value.

When we look at the table, statistical table especially Z table it should be very careful whether the area is given minus infinity there are 2 possibilities sometime the area may be given minus infinity to plus X value, sometime area may be given only the positive value, this side value 0 positively values of Z is given. If you want to know if you want to read the negative value of Z because it is symmetric, so you all can read justly, just only the positivity then we can take that value to the negative side.

(Refer Slide Time: 09:53)



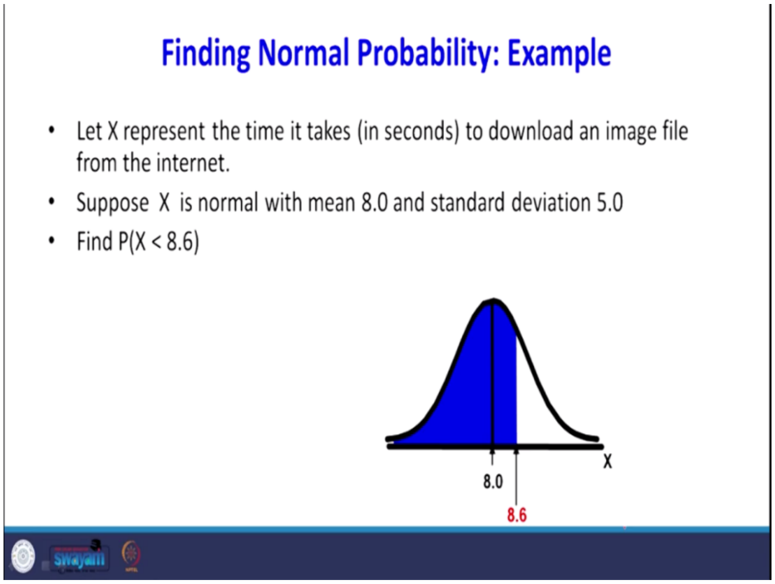
Finding Normal Probability Procedure

To find $P(a < X < b)$ when X is distributed normally:

- Draw the normal curve for the problem in terms of X.
- Translate X-values to Z-values.
- Use the Standardized Normal Table.

So, finding normal probability procedure we will see one problem to find P of A < X < b when X is distributed normally, the first one is draw the normal curve of the problem in terms of X whenever you are going to find out area, it is always good to draw the distribution draw the normal distribution then you can intuitively you can read from the picture, so the next step is translate X value to Z values then use standardized normal table where you can get the area.

(Refer Slide Time: 10:34)

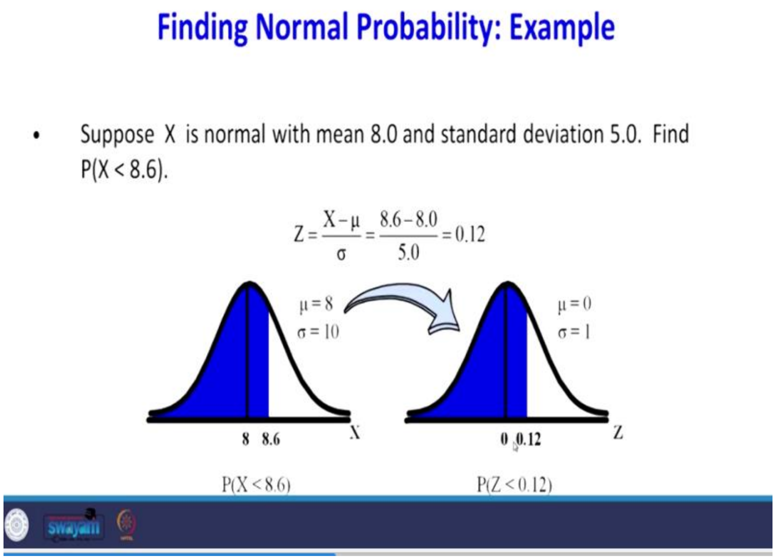


Finding Normal Probability: Example

- Let X represent the time it takes (in seconds) to download an image file from the internet.
- Suppose X is normal with mean 8.0 and standard deviation 5.0
- Find P(X < 8.6)

Let X represents the time it takes to download an image file from the internet. Suppose X is a normal with mean 8 and standard deviation 5. If we want to know what is the probability of X less than 8.6 that means what is the probability of downloading time is below 8.6 So first you have to mark the mean then you are to find out this X values 8.6, so since it is asked less than 8.6, the left side area, so the first steps is 8.6 has to be converted into, you can integrated by using normal distributions, you can substitute to minus infinity to 8.6 mean you can integrated, we will get the area there is no problem, but it is very time consuming process.

(Refer Slide Time: 11:26)



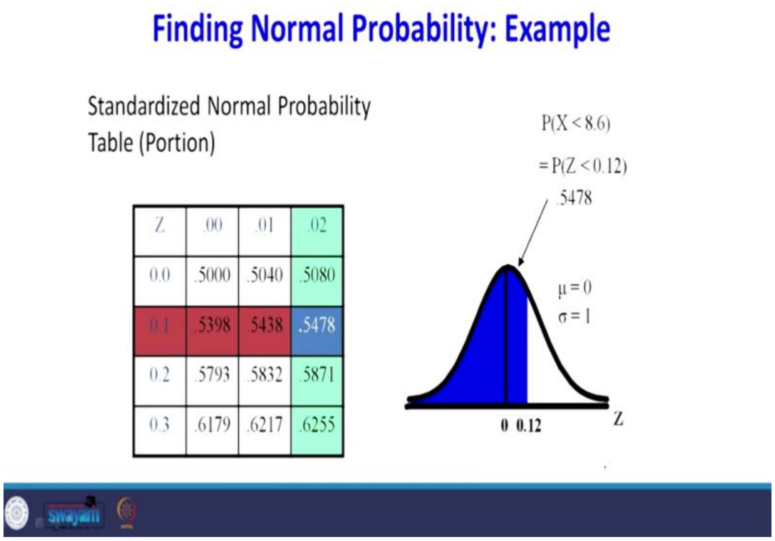
Finding Normal Probability: Example

- Suppose X is normal with mean 8.0 and standard deviation 5.0. Find P(X < 8.6).

$$Z = \frac{X-\mu}{\sigma} = \frac{8.6-8.0}{5.0} = 0.12$$

So, one easy way is you have to convert that normal distribution into standard normal distribution that means the X value has to be converted into Z scale they can read they can use the table to find out the area for the corresponding X value suppose X is the normal with mean 8 or standard deviation 5 X less than 8.6 use the Z equal to X - mu by sigma formula to get Z value when X = 8.6 so we got 0.12, so now when Z value 0.12 you can read this value directly from the normal table to know the probability.

(Refer Slide Time: 12:06)



Finding Normal Probability: Example

Standardized Normal Probability Table (Portion)

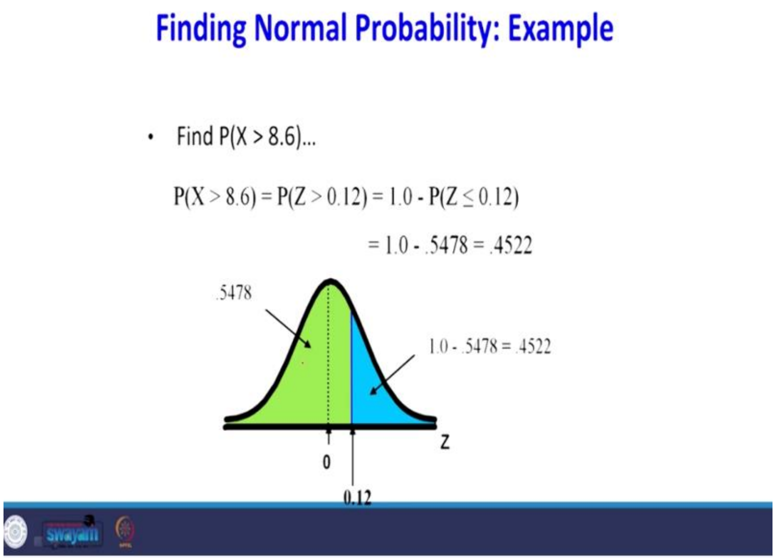
Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

$P(X < 8.6)$

$= P(Z < 0.12)$

Is it that Z value 0.12 so you can Z value 0.12 so this area is 0.5748.

(Refer Slide Time: 12:15)



Finding Normal Probability: Example

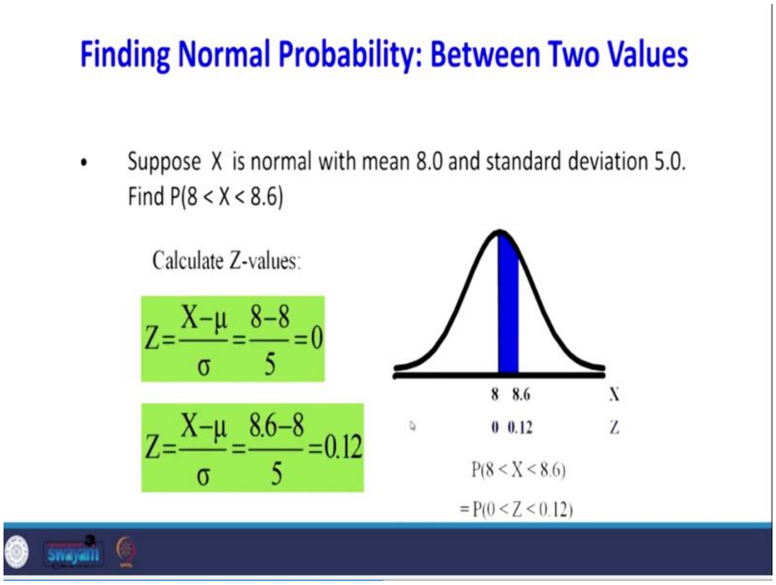
- Find P(X > 8.6)...

$P(X > 8.6) = P(Z > 0.12) = 1.0 - P(Z \leq 0.12)$

$= 1.0 - .5478 = .4522$

Finding normal probability suppose X is greater than 8.6, so now we have to look at the area of the right side so, P of X greater than 8.6 is equal to that we have to convert it to Z scale after getting since it is greater than since the area is 1. 1 - P of Z less than 0.12 will give the, the blue side area. So, one when Z = 0.12 corresponding areas 0.54 so, this side area is after subtracted from one will getting, we are getting 0.4522.

(Refer Slide Time: 12:56)



Finding Normal Probability: Between Two Values

- Suppose X is normal with mean 8.0 and standard deviation 5.0. Find P(8 < X < 8.6)

Calculate Z-values:

$$Z = \frac{X - \mu}{\sigma} = \frac{8 - 8}{5} = 0$$

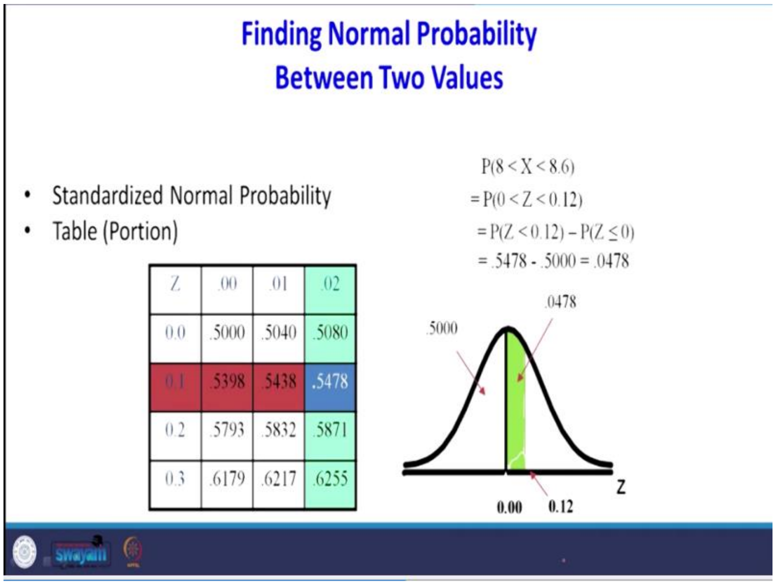
$$Z = \frac{X - \mu}{\sigma} = \frac{8.6 - 8}{5} = 0.12$$

$P(8 < X < 8.6)$

$= P(0 < Z < 0.12)$

Suppose X is a normal with mean 8 standard deviation 5 so fine P of 8 less than X less than 8.6 now, the 2 value of X is given both of values has to convert when X = 8 we are getting Z value 0, when X = 8.6 we are getting Z value 0.12, so now we have to know the area of Z 02 Z 0.12. So that means 02 0.12.

(Refer Slide Time: 13:28)



Finding Normal Probability Between Two Values

- Standardized Normal Probability
- Table (Portion)

Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

$P(8 < X < 8.6)$

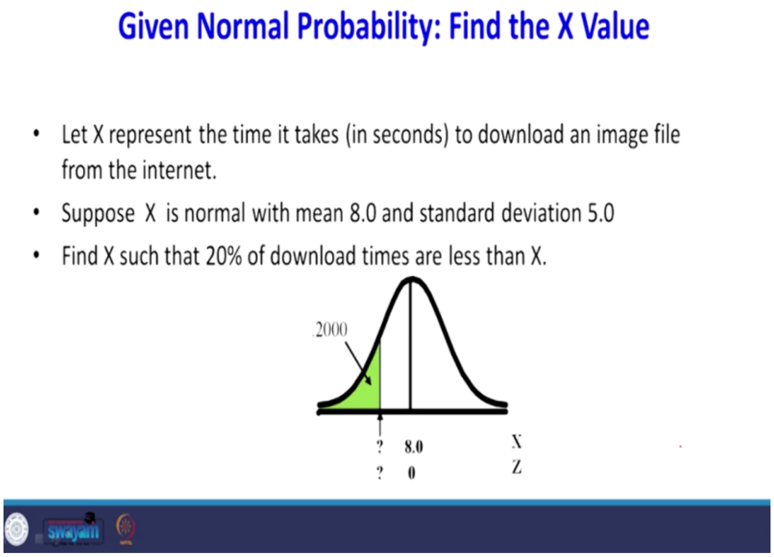
$= P(0 < Z < 0.12)$

$= P(Z < 0.12) - P(Z \leq 0)$

$= .5478 - .5000 = .0478$

One way from the table is, first you find the area up to minus infinity to Z value 0.12. So, we are getting 0.5478 then subtract when Z = 0 left side area we know it is a .5. So, the remaining is 0.0478.

(Refer Slide Time: 13:50)

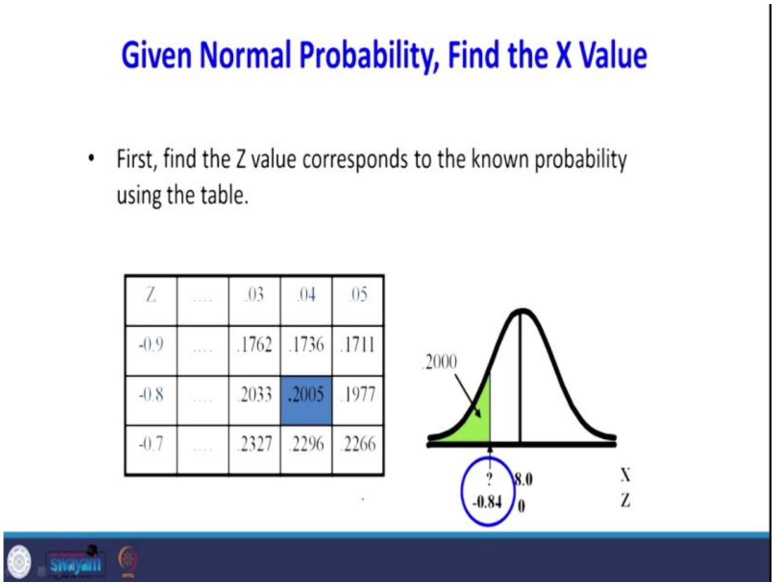


Given Normal Probability: Find the X Value

- Let X represent the time it takes (in seconds) to download an image file from the internet.
- Suppose X is normal with mean 8.0 and standard deviation 5.0
- Find X such that 20% of download times are less than X.

Now, just the reverse of that the probability is given you have to find out the X value, let X represents the time It takes to download an image file from the internet suppose X is normal with mean 8 standard deviation 5 find X such that 20% of the download times are less than X, there are 2 points here one is less than X another one is 20%. So, on the left hand side when area equal 2.2 what is the corresponding X value so, for that first you got to find out Z value from the Z you have to find out the X value.

(Refer Slide Time: 14:35)



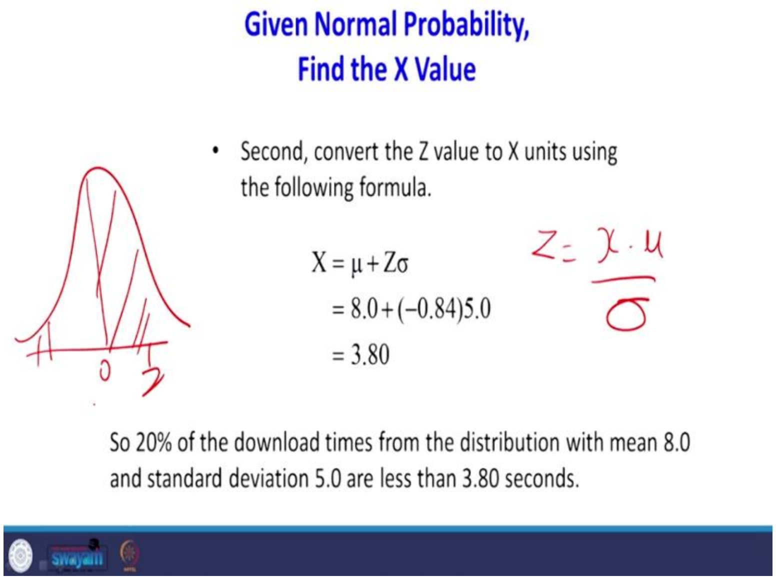
Given Normal Probability, Find the X Value

- First, find the Z value corresponds to the known probability using the table.

Z	....	.03	.04	.05
-0.9	....	.1762	.1736	.1711
-0.8	....	2033	.2005	.1977
-0.7	....	2327	2296	.2266

Now look at the table. So, when area equal to 0.2 corresponding Z values minus 0.84 this is the value of Z.

(Refer Slide Time: 14:45)



Given Normal Probability, Find the X Value

- Second, convert the Z value to X units using the following formula.

$$X = \mu + Z\sigma$$
$$= 8.0 + (-0.84)5.0$$
$$= 3.80$$

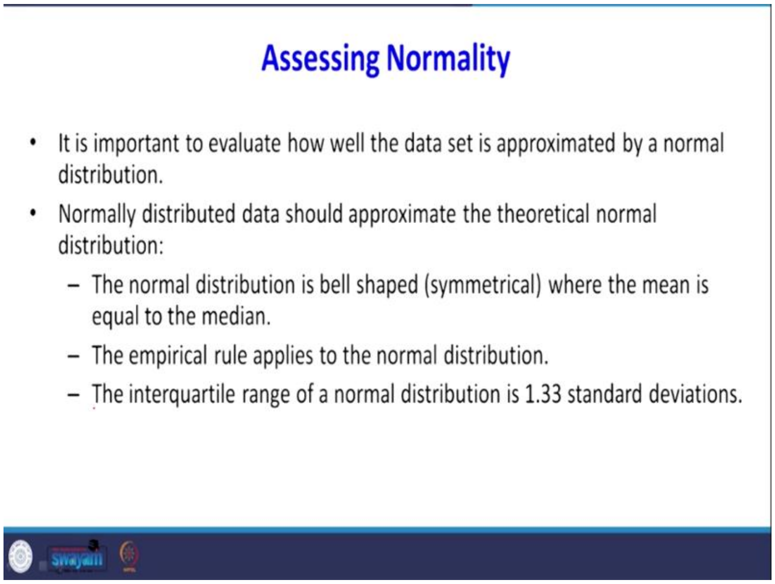
$$Z = \frac{X \cdot \mu}{\sigma}$$

So 20% of the download times from the distribution with mean 8.0 and standard deviation 5.0 are less than 3.80 seconds.

So, we know that Z value is minus 0.84 here this formula has come from this simple formula X - mu / sigma. Now, we know the value of Z from this you have to find out value of X. And one more thing the when you are finding the value of Z, you should be very careful what kind of normal distribution you are using to find out the value of Z if normal distribution is like this that is area is given from 0 to positive Z right. So if you are measuring area on the left hand side, so will get them Z value but have to attach negative side to that. So we should be careful.

So mu = 8.0 plus minus 0.84 multiply by 5, we are getting 3.80. So 20% of the download times from the distribution with the mean 8 and standard deviation 5 are less than 3.8 seconds.

(Refer Slide Time: 15:47)



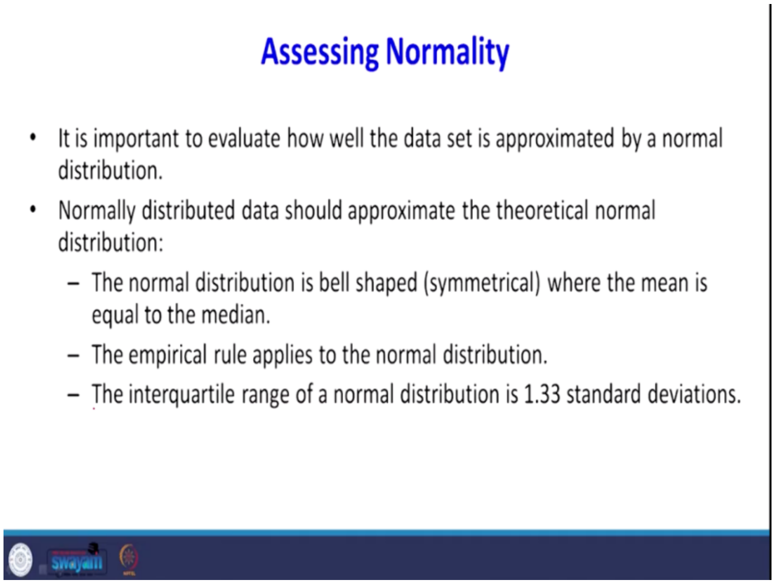
Assessing Normality

- It is important to evaluate how well the data set is approximated by a normal distribution.
- Normally distributed data should approximate the theoretical normal distribution:
  - The normal distribution is bell shaped (symmetrical) where the mean is equal to the median.
  - The empirical rule applies to the normal distribution.
  - The interquartile range of a normal distribution is 1.33 standard deviations.

Another important thing gives us is normality because the normality assumption is very important for other type of inferential statistics. I will tell you why it is important because we will be studying a concept called Central Limit Theorem, where when you do the sampling of the sampling that will follow normal distributions. So, lot of many analytical tools many statistical tools follow the assumption that data should follow normal distributions that is why as soon as you collect the data.

The first step is cleaning the data, when the cleaning in that process is you have to verify whether the data follow normal distribution or not, otherwise, you may not otherwise you will you may end up choosing wrong statistical techniques or analytical techniques.

(Refer Slide Time: 16:35)

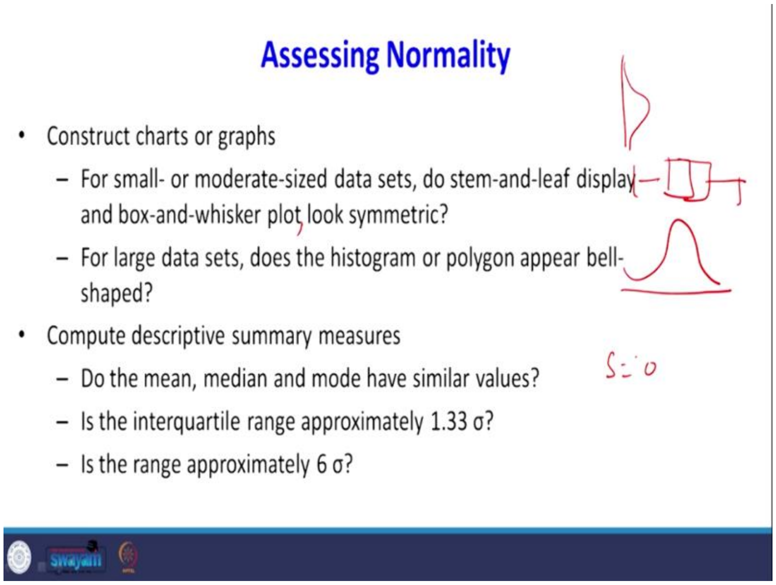


Assessing Normality

- It is important to evaluate how well the data set is approximated by a normal distribution.
- Normally distributed data should approximate the theoretical normal distribution:
  - The normal distribution is bell shaped (symmetrical) where the mean is equal to the median.
  - The empirical rule applies to the normal distribution.
  - The interquartile range of a normal distribution is 1.33 standard deviations.

It is important to evaluate how well the data Z is approximated by a normal distribution. Normally distributed data should approximate theoretical normal distribution, like the normal distribution is bell shaped where the mean is equal to the median. The empirical rule applies to the normal distribution. The interquartile range of a normal distribution is 1.33 standard deviations; these are the way to test the normality.

(Refer Slide Time: 17:04)



Assessing Normality

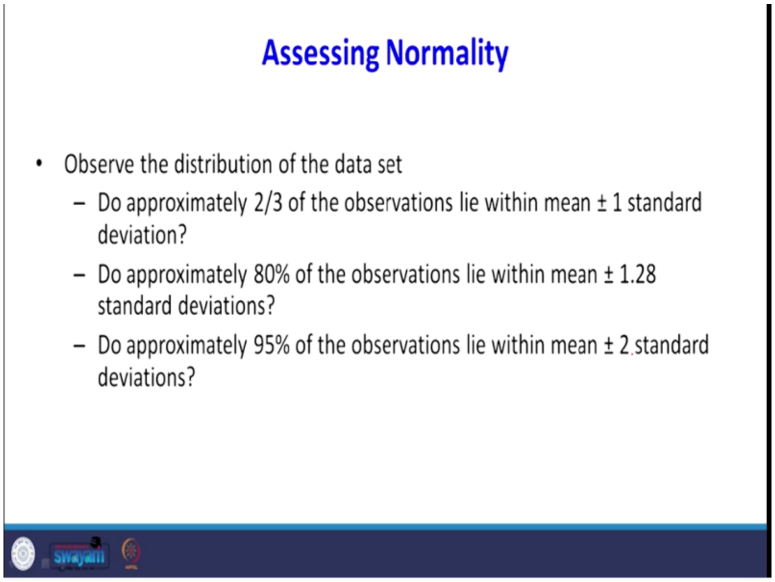
- Construct charts or graphs
  - For small- or moderate-sized data sets, do stem-and-leaf display and box-and-whisker plot look symmetric?
  - For large data sets, does the histogram or polygon appear bell-shaped?
- Compute descriptive summary measures
  - Do the mean, median and mode have similar values?
  - Is the interquartile range approximately 1.33 σ?
  - Is the range approximately 6 σ?

Another way to assess the normality is construct the charts or graph. Now, you can look at the shape of the distribution for small or moderate sized data set, do stem and leaf display and box and whisker plot and check whether it is look symmetrical or not. As I told you in the beginning of the lectures, if you look at the stem and leaf plot, you should follow this kind of shape then we can say it follow normal distributions in the box and whisker plot.

For example, box and whisker plot is like this, right, the middle line that is median line should be middle of the box then only we can say the data Z follow normal distribution for a large data set. That is the histogram or polygon appears bell shaped, you can draw a histogram and also you can verify whether it follows normal distribution other way you can compute descriptive summary measures, whether you can check mean median mode.

How the similar value is the interquartile range approximately 1.33 sigma is the range is approximately 6 sigma, these are the some descriptive measures to check whether the data follow normal distribution or not then you can find the skewness when the skewness is 0 then we can say this data follow normal distribution.

(Refer Slide Time: 18:31)

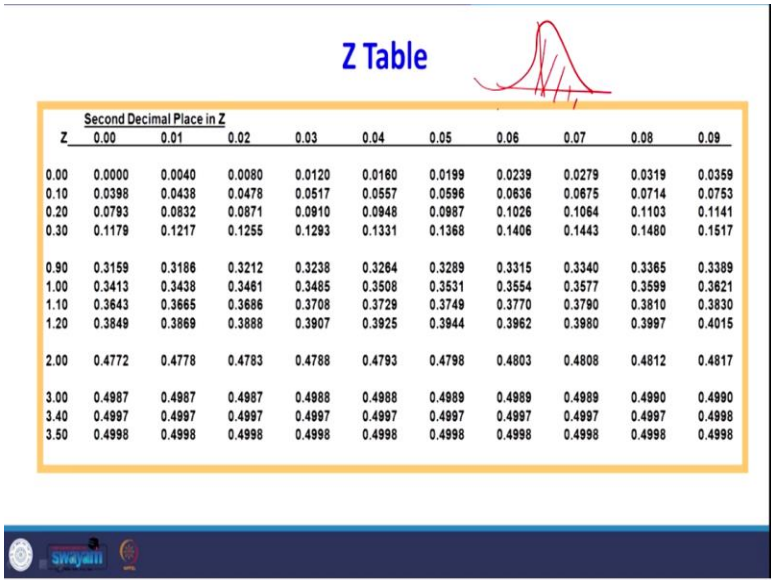


Assessing Normality

- Observe the distribution of the data set
  - Do approximately 2/3 of the observations lie within mean ± 1 standard deviation?
  - Do approximately 80% of the observations lie within mean ± 1.28 standard deviations?
  - Do approximately 95% of the observations lie within mean ± 2 standard deviations?

Some more example, to check the normality observed the distribution of the data set these are the conditions do approximately 2 thirds of the observations live within the plus or minus 1 sigma. Then we can see it follow normal distribution do approximately 80% of the observations live within plus or minus 1.28 standard deviations are do approximately 95% of the observation live within the mean or plus or minus 2 standard deviations, this is the Z table.

(Refer Slide Time: 19:02)



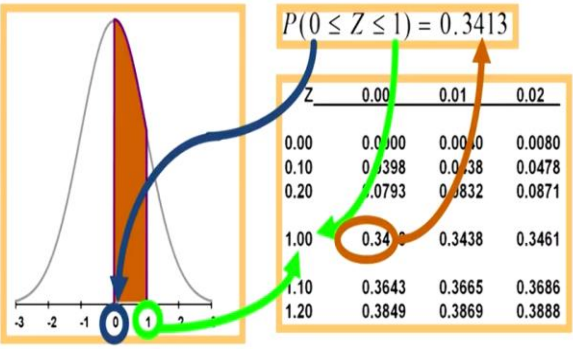
Z Table

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
	Second Decimal Place in Z									
0.00	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.10	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.20	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.30	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.90	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.00	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.10	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.20	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
2.00	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
3.00	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.40	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998
3.50	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998	0.4998

You see the previously the Z table is starting from 0 is not starting from minus infinity. So, this is second decimal, this one suppose Z is 0, the probability 0 healed is given what is given only one side know this area is given only this one. So if you are finding you have to add 0.5 suppose if you want 0.0 you have to add 0.5 to get the, Z table. Another important .which I am planning to willing to share with you.

(Refer Slide Time: 19:39)

Table Lookup of a Standard Normal Probability

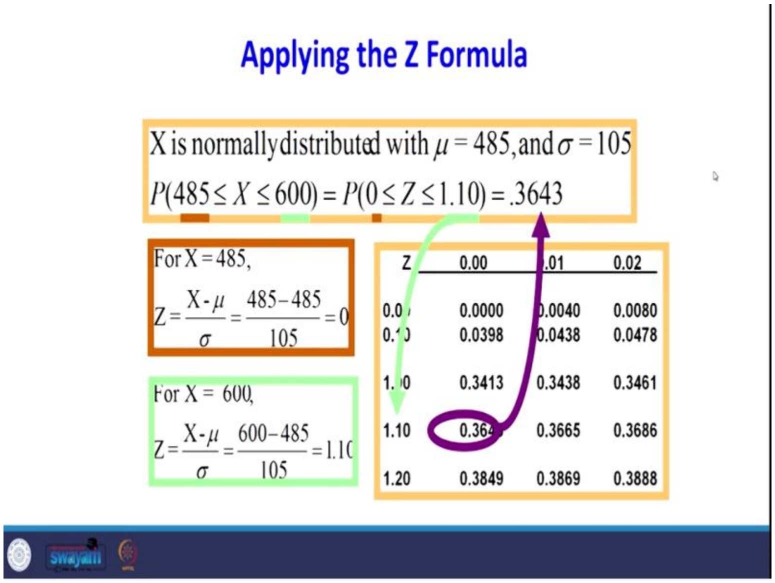


$P(0 \le Z \le 1) = 0.3413$

Z	0.00	0.01	0.02
0.00	0.0000	0.0040	0.0080
0.10	0.0398	0.0438	0.0478
0.20	0.0793	0.0832	0.0871
1.00	0.3413	0.3438	0.3461
1.10	0.3643	0.3665	0.3686
1.20	0.3849	0.3869	0.3888

See this Z = 0 Z = 1 see, this is 0.3413 right between 0 and 0 and one. Suppose if we want to know minus infinity 1 you have to add 0.5 with that plus 0.5. So, we will get the value another one when you see when you look at the normal distribution it will come back to that.

(Refer Slide Time: 20:04)



Applying the Z Formula

X is normally distributed with $\mu = 485$, and $\sigma = 105$

$P(485 \le X \le 600) = P(0 \le Z \le 1.10) = .3643$

For X = 485,

$Z = \frac{X-\mu}{\sigma} = \frac{485-485}{105} = 0$

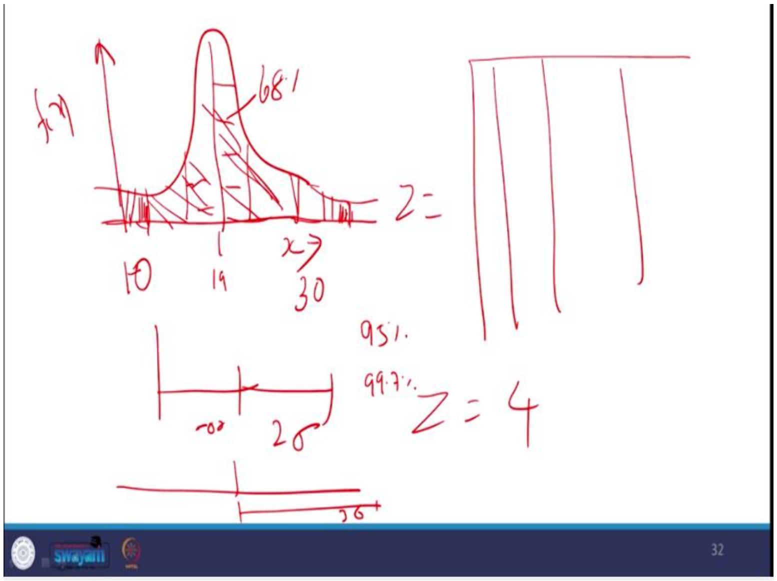
For X = 600,

$Z = \frac{X-\mu}{\sigma} = \frac{600-485}{105} = 1.10$

Z	0.00	0.01	0.02
0.00	0.0000	0.0040	0.0080
0.10	0.0398	0.0438	0.0478
1.00	0.3413	0.3438	0.3461
1.10	0.3643	0.3665	0.3686
1.20	0.3849	0.3869	0.3888

If X is normally distributed with the mu 485 and sigma 105. So, the 485 to 600 when X is 485, you have to convert that scale to 0 when X = 600 corresponding Z values 1.1 so, Z is 0 to 1.4 0.3643 is the area and the curve. Dear students we have seen So, far the properties of normal distribution then we have seen standardized normal distribution a normal distribution how these are interrelated and we have seen how to find out the area with the help of table one property.

(Refer Slide Time: 20:51)



You can look at the normal distribution the normal distribution shape is like this. So, you look at this it would not touch this is x axis y is your probability effects it will not touch the x axis you may how this doubt why it is not touching this distribution normal distribution why it is not touching axis? Because suppose, if I am plotting age of the students in the class follow normal distribution, see the average age is say 19. There is a possibility somebody suppose I am closing this way, some bodies age may be say around 30 somebody age may be around say, around 10.

So, since this normal distribution was drawn with the help of sample, I was not exactly knowing that this kind of rare value of X whether it is X = 30 hours X = 10. So, why I am not closing? Why this normal distribution not touching X axis, because we were given provision for the rare events that means X is maybe very high value X may be very low value, but I am not sure about that. That is why the normal distribution did not touch with the X axis the another doubt you may know when you look at the Z table.

When you look at the Z table, the value of Z most of the time I go back. It will show you see this the value Z is 3.5. The question may come why the value of Z is maximum 4 or 5 in the statistical table you remember the beginning of the class, I will saying from the mean if you travel on either side with one sigma distance you can capture 68% damage if you travel 2 sigma distance from the mean that is this distance 2 sigma distance minus 2 sigma distances.

You can capture 95% of the area of the normal distribution. If you travel 3 sigma distances this extreme distance, I can use some other colour please bear with me. If you travel 3 sigma distance this portion, if you travel 3 sigma distance, you can cover 99.7% of the data. Okay. So, why the value of Z is not beyond 3, the possibility of the Z value is beyond 3 is only 0.3% the same time the probability of x value to become extremely high or extremely low is only 0.3%.

What is the meaning of that only 0.3% change the value of that will be more than 3 that is why all statistical tables given only 3.5 or 4 not beyond that. The another reason and also why we are not closing with the x axis the probability of that extreme events to happen is only 0.3% provided if it is following normal distribution. Now, I will summarize the students that so far via we have seen different type of probability distributions.

The previous class we have seen some of the continuous distribution in this class we have seen an important and normal distribution that is a normal distribution. We have learned properties, normal distribution and a standard normal distribution, how to convert a normal distribution to standard normal distribution, how to refer Z table to find out the area that you have seen. The next class with the help of Python will use how to find out the area under the curve or how to find out the mean standard deviation of different distributions. Thank you very much.