

Introduction to Probability and Statistics
Prof. G. Srinivasan
Department of Management Studies
Indian Institute of Technology, Madras

Lecture – 20
Normal Distribution


In this lecture we study the normal probability model, where the random variable follows a Normal Distribution.

(Refer Slide Time: 00:29)

Normal random variables have bell shaped histograms. The probability distribution of a normal variable is the bell curve.

The probability distribution of any random variable that is the sum of *enough* independent random variables is bell shaped.

If the random variables to be summed have a normal distribution, the sum has a normal distribution. Sum of just about any random variable are eventually normally distributed.




So, we first try to understand the normal distribution. So, normal random variables have bell shaped histograms, all of us have seen the bell shaped curve we will also be showing the bell shaped curve in this lecture. The probability distribution of a normal variable is the bell curve. The probability of distribution of any random variable that is the sum of enough independent random variables is also bell shaped, we will see that. If random variables to be summed have a normal distribution, then the sum has a normal distribution. Sum of just about any random variable are eventually normally distributed.

So, if we take a random variable and keep summing them we at some point would come to the normal distribution. So, you seen 2 or 3 sentences and we will tried explain these sentences suitably.

(Refer Slide Time: 01:22)

Consider tossing a coin 10 times and compute the probability of 0 heads to 10 heads . $p = 0.5$ and $n = 10$.

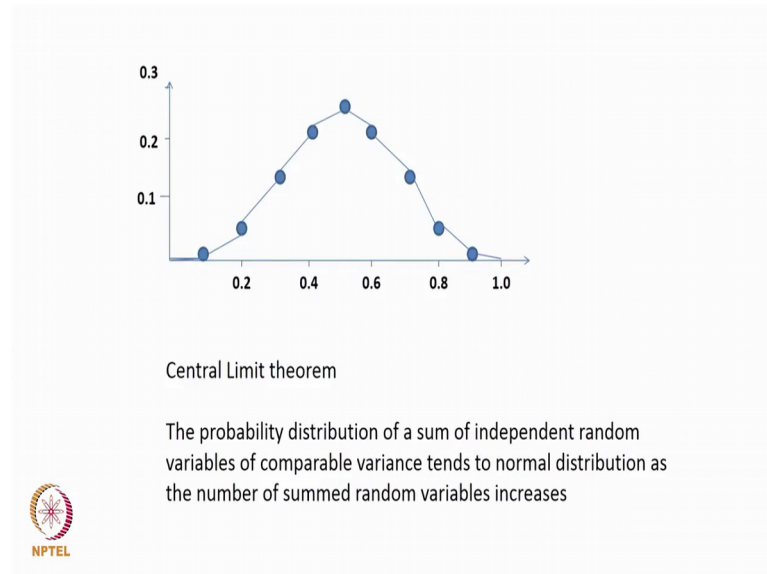
$P(0) = 0.0009766 = P(10)$
 $P(1) = 0.009766 = P(9)$
 $P(2) = 0.044 = P(8)$
 $P(3) = 0.117 = P(7)$
 $P(4) = 0.205 = P(6)$
 $P(5) = 0.246$



For example, you consider tossing a coin 10 times and compute the probability of 0 heads to 10 heads. So, p is equal to 0.5 and n is equal to 10. We realized that p is 0.5 probability of success is also equal to probability of failure and therefore, P of 0 which means probability of getting 0 heads assuming head is a success can be calculated by $n C X P^X Q^{n-x}$. Q is $1 - P$ we have seen that in the previous lectures on binomial. So, P is also equal to Q therefore, P of 0 you see $n C 0$ which is one P to the power 0 0 heads and Q to the power 10 and since P and Q are equal. You would have P to the power 10 in all these cases, except the $n C X$ will change.

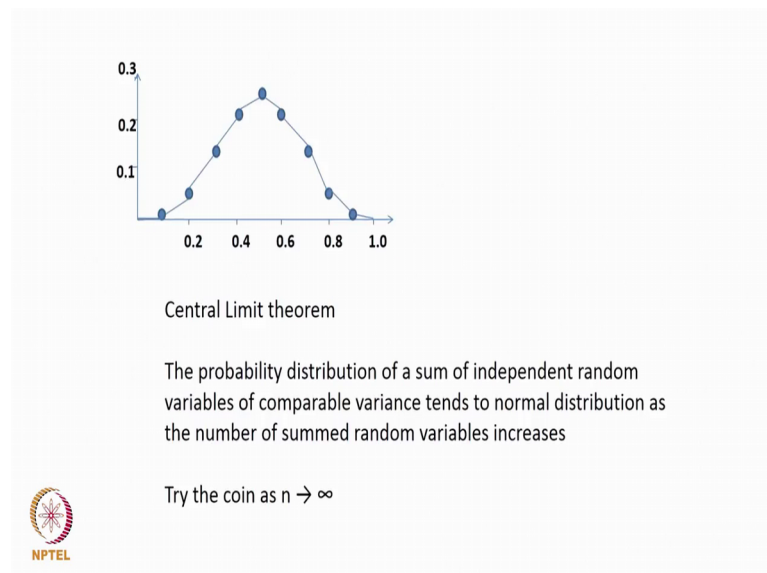
So, probability of 0 heads will be equal to probability of getting 10 heads, which is equal to 0.0009766. Similarly, probability of getting one head will be equal to the probability of getting 9 heads. Please note in this case, because P is equal to Q this happens and both are equal to 0.5. So, 0.009766, P of 2 is equal to P of 8 0.044, P of getting 3 heads is equal to P of getting 7 heads which is 0.117, P of 4 is equal to P of 6, 0.205 and P 5 is 0.246. So, if we add from P of 0 to P of 10, we will get 1. So, this comes from the binomial distribution doing.

(Refer Slide Time: 03:06)



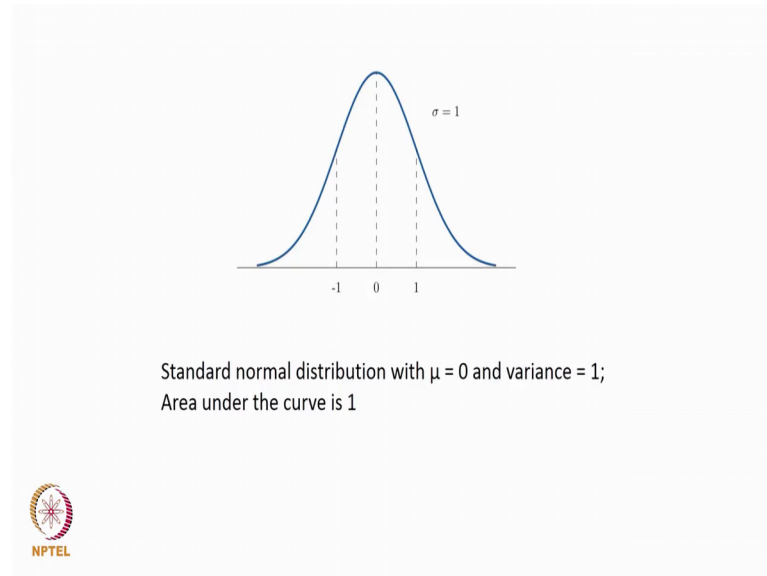
And if we try to plot this, we try to get a picture which is like this, which is very similar to the normal curve. So, the central limit theorem which is a very important theorem; says, the probability distribution of a sum of independent random variables of comparable variance tends to normal distribution as the number of summed variables increases.

(Refer Slide Time: 03:28)



So, try tossing the coin as n tends to infinity we will get this.

(Refer Slide Time: 03:34)



Now, what we will look at more in this lecture is, the standard normal distribution, we will study this further. There are some more equations that describe the normal distribution which we would possibly not do in this introductory course on probability and statistics. Perhaps, the first level course we would look at all of them. This is the normal distribution curve or the bell shaped curve. This is also the standard normal we will also see the difference. A typically standard normal distribution has mu equal to 0 and variance equal to 1. Area under the curve will be equal to 1. So, in otherwise you would have a mu here, now we have 0 here. Now also realize that this does not touch the X axis from either side. It can assume to vertically converges it just goes on and on. Therefore, in principle the random variable can take any value.

Now, remember that both the normal curve as well as the standard normal curve look similar they shape is the same, except that we have mu equal to 0 in the standard normal and the corresponding mu in the normal distribution. We will see examples to understand all of this.

So, which of the following can be treated as normal. So, whenever the normal comes one has to understand the symmetry, one has to understand the peak in the middle and so on. So, when we plot so, what kind of a curve can we get? And from that curve can we say that something is normal. So, simple characteristics are the random variable can take any


value, there is a peak at μ and then is a bell shaped curve and then there is a symmetry. So, we will look at all these partners then and try to answer these questions.

(Refer Slide Time: 05:21)

Exercise

Which of the following are normal?

1. Marks obtained out of 100 by 200 students in a subject
2. Money value of each purchase in a supermarket in a day
3. Career score in ascending order of a cricketer
4. Number of visitors in a day to the department



Marks obtained out of 100 by 200 students in a subject. Generally, we could look at this kind of has a normal distribution in the sense that, there are some interesting a reasons at just why we need not. Because we just saw that the mark, if we assumed to be normal it can take very large value, it can take a very small value as well, but then when we are talking about an exam. In a subject we have clearly defined boundaries. It is a 0 to 100 and therefore, we do not have a value of X equal to 1 or 1 and so on. But in spite of that we could expect a reasonable amount of symmetry. And we could think of this as close to a normal. Money value of each purchase in a supermarket in a day or may not be close to normal. What will happen is the average it will not peak at the average. We could have few a very large purchases, we could have a large number of small purchases and so on

So, it could be a skewed distribution. So, skewed distribution would not be symmetric, we have seen skewed distributions are skewed distribution earlier in this course. So, it will taper to the right the peak will shift will be to the left if it is right skewed and the other way. If it is left skewed, career score in ascending order is of a cricketer, it talks about individual cores scores. So, we would not be able to do that. But if we sort this career scores in some order and try to build a histogram and so on.

So, one might try to get a picture that reasonably close. But again in this case that will be a small number of very large scores, and a large number of small scores. So, we could expect some amount of skewedness in the data and therefore, we need not treated as close to normal. Number of visitors in a day to a department. Again may not be very close to normal we could have some days. You could have simply a bunch of visitors, and we could have 40 or 50 visitors on some days, and on some other days we would have a small. So, we will have a large number of days with small number of visitors and small number of days with a large number of visitors and therefore, would not be close to normal.

Now, what is the relationship between the normal distribution and the standard normal distribution? So, we will be working with a standard normal distribution most of the times. So, a given normal distribution will have a given mu and a given sigma. The standard normal changes that to mu equal to 0 and variance equal to 1. So, what is the difference? So, what is the difference is here?

(Refer Slide Time: 08:12)


Standardizing the normal distribution – z score

z score measures the number of standard deviations that separates the value from the mean

$$z = \frac{x - \mu}{\sigma}$$

The average mark in a class of 200 students is assumed to be normal (60, 20). Find the probability that a randomly chosen student has a mark > 70?

$Z = (70 - 60)/20 = 0.5$
 $P(x > 70) = P(z > 0.5)$
From standard normal tables area for $z = 0.5$ is



Z score measures the number of standard deviations that separates a given value from the mean. So, if mu is the mean and sigma is a standard deviation and x is a given value we calculate what is called x minus mu by sigma. So, x minus mu is the difference divided by sigma is the difference, divided by the standard deviation which tells us the number of standard deviations that separates the value from the mean.

For example, if z equal to 2 then x minus μ is equal to 2 σ . Therefore, the number of standard deviations that separates the value of from the mean is 2. So, z is x minus μ by σ so, quickly to do a computation. Average mark in a class of 200 students is assumed to be normal with 60, 20. So, μ is 60 and standard deviation is 20. Find the probability that a randomly chosen student has mark greater than 70.

So, we first find out z . So, in this case we normally used small z lowercase z . So, have shown it is upper case Z here, but we use z is equal to x minus μ by σ , μ is 60 σ is 20 x is 70. So, z corresponding to x equal to 70 is x minus μ by σ 70 minus 60 which is 10 divided by 20 which is 0.5. So, probability of x greater than 70 is the same as probability of z greater than 0.5. Because we have now reduced or converted the given μ and σ into a z score, and we will start working using the z score and using the standard normal table and the z score the area would correspond to the probable.

So, what we have to understand this given μ and σ . x is related to z and z is equal to x minus μ by σ . So, for a given x we can calculate z , and then for the z value get some figures from the standard normal and then use it to solve for the given x . That is something which we will do. From standard normal tables, the area we will compute and show. So, there is the standard normal table and the area under the standard normal table. We will use that area to compute.

(Refer Slide Time: 10:46)

STANDARD NORMAL DISTRIBUTION: Table Values Represent AREA to the LEFT of the Z score.

Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.9	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.8	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0006	.0006	.0005
-3.7	.0011	.0011	.0010	.0010	.0009	.0009	.0009	.0008	.0008	.0008
-3.6	.0016	.0015	.0015	.0014	.0014	.0013	.0013	.0012	.0012	.0011
-3.5	.0023	.0022	.0022	.0021	.0020	.0019	.0019	.0018	.0017	.0017
-3.4	.0034	.0032	.0031	.0030	.0029	.0028	.0027	.0026	.0025	.0024
-3.3	.0048	.0047	.0045	.0044	.0043	.0042	.0041	.0040	.0039	.0038
-3.2	.0069	.0066	.0064	.0062	.0060	.0058	.0056	.0054	.0052	.0050
-3.1	.0097	.0094	.0090	.0087	.0084	.0082	.0079	.0076	.0074	.0071
-3.0	.0135	.0131	.0126	.0122	.0118	.0114	.0111	.0107	.0104	.0101
-2.9	.0187	.0181	.0175	.0169	.0164	.0159	.0154	.0149	.0144	.0139
-2.8	.0256	.0248	.0240	.0233	.0226	.0219	.0212	.0205	.0199	.0193
-2.7	.0347	.0336	.0326	.0317	.0307	.0298	.0289	.0280	.0272	.0264
-2.6	.0466	.0453	.0440	.0427	.0415	.0402	.0391	.0379	.0368	.0357
-2.5	.0621	.0604	.0587	.0570	.0554	.0539	.0523	.0508	.0494	.0480
-2.4	.0820	.0798	.0776	.0755	.0734	.0714	.0695	.0676	.0657	.0639
-2.3	.1072	.1044	.1017	.0990	.0964	.0939	.0914	.0889	.0866	.0842
-2.2	.1390	.1355	.1321	.1287	.1255	.1222	.1191	.1160	.1130	.1101
-2.1	.1786	.1743	.1700	.1659	.1618	.1578	.1539	.1500	.1463	.1426
-2.0	.2275	.2222	.2169	.2118	.2068	.2018	.1970	.1923	.1876	.1831
-1.9	.2872	.2807	.2743	.2680	.2619	.2559	.2500	.2442	.2385	.2330
-1.8	.3598	.3515	.3428	.3342	.3258	.3176	.3093	.3014	.2934	.2856
-1.7	.4457	.4363	.4272	.4182	.4093	.4006	.3920	.3836	.3754	.3673
-1.6	.5480	.5370	.5262	.5155	.5050	.4947	.4846	.4746	.4648	.4551
-1.5	.6681	.6552	.6426	.6301	.6178	.6057	.5938	.5821	.5705	.5592
-1.4	.8076	.7927	.7780	.7636	.7493	.7353	.7215	.7078	.6944	.6811
-1.3	.9596	.9510	.9428	.9349	.9271	.9193	.9116	.9041	.8967	.8894
-1.2	1.1507	1.1314	1.1123	1.0935	1.0749	1.0565	1.0383	1.0204	1.0027	.9853
-1.1	1.3567	1.3350	1.3136	1.2924	1.2714	1.2507	1.2302	1.2100	1.1900	1.1702
-1.0	1.5866	1.5625	1.5386	1.5151	1.4917	1.4686	1.4457	1.4231	1.4007	1.3786
-0.9	1.8406	1.8141	1.7879	1.7619	1.7361	1.7106	1.6853	1.6602	1.6354	1.6109
-0.8	2.1186	2.0897	2.0611	2.0327	2.0045	1.9766	1.9489	1.9215	1.8943	1.8673
-0.7	2.4196	2.3885	2.3576	2.3270	2.2965	2.2663	2.2363	2.2065	2.1770	2.1476
-0.6	2.7425	2.7093	2.6763	2.6435	2.6109	2.5785	2.5463	2.5143	2.4825	2.4510
-0.5	3.0854	3.0503	3.0153	2.9806	2.9460	2.9116	2.8774	2.8434	2.8096	2.7760
-0.4	3.4458	3.4090	3.3724	3.3360	3.2997	3.2636	3.2276	3.1918	3.1561	3.1207
-0.3	3.8209	3.7828	3.7448	3.7070	3.6693	3.6317	3.5943	3.5569	3.5197	3.4827
-0.2	4.2074	4.1683	4.1294	4.0905	4.0517	4.0129	3.9743	3.9358	3.8974	3.8591
-0.1	4.6017	4.5620	4.5224	4.4828	4.4433	4.4038	4.3644	4.3251	4.2858	4.2465
0.0	5.0000	4.9601	4.9202	4.8803	4.8405	4.8006	4.7608	4.7210	4.6812	4.6414



Now, how do we do that? I have just shown these 2 tables cannot see the internet. So, I acknowledge that, and these tables are available in open source these tables are available in most statistics books, and it is not difficult to get these tables. You will see a clutter of numbers, and sometimes you will see a small picture which also tells you what this number represents.

Now, in this table the picture is replaced by a sentence which as table values represent area to the left of the Z score. So, if you read this table very carefully. Since I have to show the entire thing in one slide, I have to reduce the font size. So, you will not be able to read it. So, I am reading it for you. Table values represent area to the left of the Z score. So, what is it? So, there is a Z score here, it starts from minus 3.9 and goes to 0 in this picture or table. And you also see 0 0, 0 1, 0 2 0 3 and so on. So, if your Z score is minus 3.43, I am just place in the mouse in that place minus 3.43. So, area to the left of z equal to 3.43, 3.43 is here and that is 00.0003 is the area to the left of minus 3.43.

Now, we go to your the next table which is also a similar table. Again area to the left, and then we realize that here Z varies from 0 to 3.9 on this, and then we have Z on the other side. So, if we look at plus Z is equal to plus 3.25, let us say. So, 3.2 is here 3.25 is here. So, 0.99942 is the area to the left of 3.25. So, if we use these 2 tables, what we understand is given Z value, these tables give us in area to the left of the given Z . Since the total area under the normal standard normal curve is 1, area to the right of the given Z will be 1 minus the area to the left of the given Z . So, we will use this to solve some of these problems.

So now what is we do here? So, somewhere here we said, now we have to find out what is the area for z equal to 0.5. Before we do that let us also understand something from the 2 tables. If you see carefully from the 2 tables, you realize that Z equal to minus 3.9 it is coming to Z equal to 0. And then you realize that z is equal to now if you have to see this little carefully. So, we see z equal to 0.00 the area is 0.5. So, Z is 0.00 is this point, this is Z is equal to 0. So, z is equal to 0.0 the area to the left of this which mean starting from here, right up to this curve right to up to this point, bring it down this area is 0.5 which we know.

Now, let us try to understand because these values are coming and increasing this way. So, what is this? So, this is minus 0.09 you see this is 0.08 and so on. So, you realize that

0-point minus 0.09 will be very, very close to 0 here, and 0.08 would even be slightly to the left of this and so on. And then we realize that the values are actually approaching 0.5. So, at Z is equal to 0 the area is 0.5.

If you look at the next table, once again Z is equal to 0 the area is 0.5. And as we keep increasing Z , the area to the left of it keeps increasing and you realize that at 3.99 the area is 0.99997. So, that is where this 3 sigma becomes important. So, if x minus remember z is equal to x minus μ by σ . So, when Z is 3 then x minus μ is equal to 3 sigma. So, the area under the left of 3 sigma is 0.99997. So, roughly are 3 sigma will be somewhere here. So, this will be z equal to 3. So, almost the entire thing is covered.

So, as now x increases beyond this and goes even to a very large value; the area to the right of it will be negligible and will be close to 0. So, generally we would be looking at z equal to plus 3 to minus 3; and plus 3 effectively covers this entire bell shape of the standard normal, and has area almost equal to 1.99997 and so on.

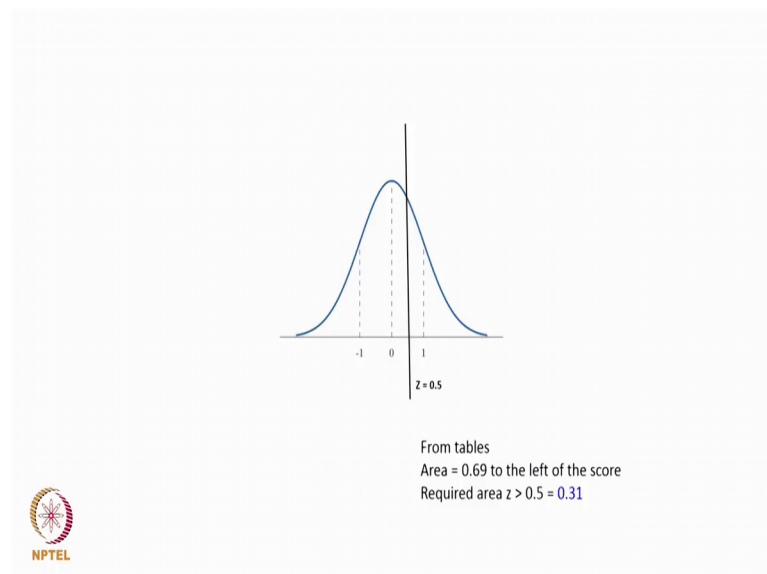
So now let us try to find out what is the value for z equal to 0.5. So, for z equal to the 0.5 we look at this table. So, z equal to the 0.5 is here. So, 0.5 is 0.500 therefore, 0.691 is the area. So, 0.691 is the area; so, from standard normal tables area for z equal to 0.5 is 0.691. So, we are now interested in z greater than 0.5; which means to the right of z is equal to 0.5. So, let us take it us 0.69 for the purpose of discussion. So, area to the left of z equal to 0.5 is 0.69 and area to that right of z equal to 0.5 is 0.31 and therefore, the probability that randomly chosen student has a mark greater than 70 is 0.31.

So, let us actually look at this again. So, you could have z equal to 0.5 somewhere here, and we would have this somewhere here. So, to the left of this is 0.69 to the right of this is 0.31. And therefore, the answer is 0.31. It is also important to note that it is we do not try to answer a question like what is the probability that a randomly chosen student has some mark equal to 70 using this kind of analysis. Which becomes very difficult and much later we will know that such a probability tends to 0 and so on. But as far as the course that we are doing concerned we restrict ourselves to finding out, what is the probability that somebody has mark greater than 70, what is the probability that somebody has mark less than 50, what is the probability that somebody has the mark between 30 and 40. So, all these questions can be answered, but we do not try to answer

a question, what is the probability that the student has a mark equal to 70 or equal to 80 through this kind of a analysis.

So, this is how we relate the given mu and sigma to a z score, and then use the standard normal table to try and answer questions like, what is the probability that mu or x is greater than 70 or z is greater than something. So, for a given x we find out the z score and then area under the standard normal curve helps us to find out the probability.

(Refer Slide Time: 18:57)




So, let us for example, for the same problem which we were discussing. So, you realize that this is z equal to 0.5. So, from the table area to the left of it that is starting from here, going right up to this coming a touching this point come down and here. That areas 0.69 and therefore, area to the right of z equal to 0.5 is 0.31 which is the answer 2 the given question. So, we will continue and take some more examples to understand the standard normal and have a discussion on it.

(Refer Slide Time: 19:33)

Match the following

Number	Column A	Column B	
1	Mean of X	σ^2	2
2	Variance of X	5/6	5
3	Probability of X above 1SD > mean	1/6	3
4	Probability that X < mean	μ	1
5	Probability of z score less than 1	0.5	4



Now, we look at this match the following there are 5 items in column A and 5 items in column B. So, mean of X variance of X so, mean of X is μ , variance of X is σ^2 probability of X above 1 standard deviation greater than the mean. Probability X less than equal to mean and probability of z score less than 1. Probability of X less than equal to mean so, X equal to mean is z equal to 0, $(X - \mu) / \sigma$. So, in X is μ z is 0, now z is 0, it comes in a middle of the bell shaped curve so, 0.5 is the answer. So, probability X less than mean is 0.5 for the 4th one. Probability of X above 1 standard deviation greater than the mean, actually if you go to the normal tables we find that it is area to the left of z equal to 1, 1 standard deviation greater than the mean is z equal to 1.

So, area to the left of z equal to 1 is about 0.84 which we approximate to 5 by 6. Therefore, area above 1 standard deviation is 1 minus 5 by 6 which is 1 by 6 which is given here. And for question number 5 z square less than 1 represents area to the left of z equal to 1; which is about 0.84 which is approximated to 5 Y by 6 in this. And therefore, the answers are μ for the mean σ^2 for the variance above 1 standard deviation 1 by 6, less than equal to μ mean 0.5 below 1 standard deviation of z less than 1 is 5 by 6.

(Refer Slide Time: 21:23)

True or false

The age of 1000 employees in a factory has mean = 40 and SD = 10.

1. More employees are older than 45 than between 35 and 45
2. Most employees are older than 30
3. If the ages are represented in months, they would still be normally distributed
4. If employees retire at 60, how many do you expect to retire soon?
5. If employees below 30 are sent for training, how many do you expect to go for training?
6. If employees above 50 have to go for health test, how many do you expect to go?

1. Older than 45: $z = 0.5$ area = 0.31
Between 35 and 45 is diff between > 45 and > 35 . For mark = 35 $z = -0.5$
Area = 0.31 (left) Required area = $0.69 - 0.31 = 0.38 > 0.31$


2. For age = 30 $z = -1$. Area to left = 0.16; 84% are older than 30

3. Yes

4. For age = 60 $z = 2$ area to left = 0.98; 2% be over 60

5. $Z = -1$. Area to left = 0.16; 16%

6. Age = 50 $z = 1$. Area to left = 0.84; 16%



So, let us look at some more questions. So, let us assumed that age of 1000 employees in a factory, has mean equal to 40 and standard deviation is equal to 10. And to answer these questions we are trying to assume that it is normal. So, when we make this normal assumption one has to keep in mind that, X can take a very large value or X can take a very small value. But then we are going to make this normal assumption and then try to answer these questions.

So, more employees are older than 45, than between 35 and 45. So, mean is 40 standard deviation is 10. So, if we take 45 then z is equal to 45 minus 40 by 10 so, z is equal to 0.5. So, area to the left of 0.5 we just now says 0.69, and area to the right is therefore, 0.31. So, more employees older than 45 is probability is 0.31. And between 35 and 45 is the difference between z equal to 45 and z equal to 35; so, z 35 is minus 0.5. So, when z is minus 0.51, has to understand that the area to the left of it is 0.31. Therefore, the required area is 0.69 minus 0.31, which is 0.38 which happens to be greater than 0.31.

So, let us try to understand this a little bit more. So, let say we have this. So, this is a case where ages 45 so, z is 0.5. So, if we look at a case where ages 45, which is 0.5. So, this area to the left is 0.69 this is 0.31. So, when we look at 35 z is equal to minus 0.5 which will come somewhere here, come somewhere here. And by symmetry this area will also be equal to this area so, area to the left of z equal to minus 0.5 is 0.31. Area to the left of plus 0.5 is 0.69. And therefore, the area between these 2 is the difference which we saw

here as difference between 0.69 and 0.31 is 0.38 which is greater than 0.31. Therefore, more employees between the age group of 35 and 45 than employees with age greater than 45. Most employees are older than 30.

Now, the mean is 40 so, z is minus 1. And because z is minus 1 the area to the left of that z will be less than 0.5, and area to the right of that z will be greater than 0.5. Therefore, we will have more people older than 30. In fact, the average is 40 therefore; we will have more people greater than 30. How many more we can find out? So, z is equal to minus 1; which gives us area to the left is 0.16 from the table. And therefore, to the right is 0.84 and for 84 percent of the people would be older than 30; is the ages are represented in months? Would it still be normally distributed? Yes, so, multiply the random variable by a constant. It will still be randomly distributed is the original one; is the mean will only change. If the employees retired 60, how many do you expect to retire soon as an interesting question.

So, when X is equal to 60, z becomes equal to 2 and area to the left is 0.98. So, 2 percent will be over 60 is the actual answer. But then we have to go back and realize that we have made that normal approximation. So, these 2 people we know now say would either have retired or it is a would say very close to retirement. But 2 percent will be over 60 and therefore, they would have retired by now if 60 is the retirement age. If employee is below 30 years are sent for training, how many would you expect to go for training.

So, we have already seen that when age is 30, z is minus 1 and area to the left is 0.16. Therefore, 16 percent of the people would be going for training. If employees above the age of 50 have to go for a health checkup, how many do you expect to go? So, when X is equal to 50, z is equal to 1 and we know that area to the left is 0.84. So, people with age less than 50, 84 percent, people with age greater than or equal to 50 will be 16 percent. And these 16 percent will go for a health checkup.

So now you see how a normal approximation can help understand certain things in reality. But one can question the normal assumption in this case, but let us assume that we have make that assumption and we have tried to answer these questions.

(Refer Slide Time: 26:43)

Question

If X and Y are both normal with mean μ and SD σ , how would the distributions of $2X$, $2Y$ and $X + Y$ look?


Normal with mean = 2μ . Variance of sum = $2\sigma^2$ and not $4\sigma^2$

The average salary in an office with 2000 people is 20000 with SD = 10000. If the salary is normally distributed, how many have a salary > 50000?

If all get a bonus of 5000 for a festival what happens to mean and SD

If all get 10% rise, what happens to mean and SD

Salary > 50000; $z = 3$ area to left = 0.999 people with salary > 50000 is 0.001% 1 in 1000; mean shifts to 25000; SD is same; both mean and SD increase by 10%.



So, if X and Y are both normal with mean μ and standard deviation σ , with the distribution of $2X$, $2Y$ and $X + Y$. How would it look like? So, all of them there will be normal with mean 2μ , but since variance is add item the variance of the sum will be $2\sigma^2$ and not $4\sigma^2$. Therefore, standard deviation will be root to time σ and not to 2σ . The average salary in an office with 2,000 people is 20,000, the standard deviation of 10,000. If the salary is normally distributed how many would have salary greater than 50,000? So, when X is 50,000, $X - \mu$ by σ is 3. So, z is equal to 3. Where you have to the left 0.9999; where you have to the right is 0.001. So, 1 and 1000 would get a salary greater than 50,000.

If all of them get a bonus of 5000 for a festival, what happens to mean standard deviation? So, the mean increases. So, 20,000 will become 25,000 standard deviation will be the same. If all of them get a 10 percent increase, what happens to mean and standard deviation? 10 percent so, multiplied so, both mean and standard deviation will increase by 10 percent in this case.


(Refer Slide Time: 28:06)

Question

Find the probabilities from standard normal tables

1. $P(z < 1)$
2. $P(z > -1)$
3. $P(|z| > 1)$
4. $P(-1 \leq z \leq 1)$

1. $z = 1$ area to left = 0.84 prob = 0.84
2. $Z = -1$ area = 0.84
3. Z to the left of -1 and to the right of +1 area = 0.32
4. $P(-1 \leq z \leq 1) = 0.68$



Find the probability from standard normal tables, z less than 1. So, we already saw that z less than 1 area to the left is 0.84 so, the answer is 0.84. So, question number 2 find the probability of z greater than minus 1? So, we know that z equal to 1 area to the left is 0.84 z equal to 1 area to the right is 0.16. It is symmetric therefore, z equal to minus 1; where you have to the left is 0.16 which is less than z equal to minus 1. Therefore, z greater than equal to minus 1 the answer is 0.84 probability of mod z greater than 1. So, mod z greater than 1 means it is either z greater than 1 or z less than minus 1. So, in z is greater than 1 mod z is greater than 1. When z is less than minus 1 for example, when z is minus 2 the mod z is 2 which is greater than 1.

So, we have already seen z greater than 1 is 0.16. And we also know that z less than minus 1 by symmetry is 0.16 and therefore, what we want here is z greater than 1 and z less than minus 1. So, 0.16 plus 0.16 is equal to 0.32. The forth one is z is between minus 1 and plus 1. So, Z equal to 1 has 0.84 the mean is at 0.5. So, between z equal to 0 and z equal to 1 is 0.34. Between z equal to minus 1 and z equal to 0 is also 0.34. Therefore, between minus 1 and plus 1 a 2 times 0.34; which is 0.68 which is what we show through this pictures this is roughly z equal to 1.

So, to the right is 0.16 to the left is 0.84. So, this is 0.16 this is 0.34, this is 0.34, this is again 0.16. So, you realize that this to put together gives us 0.68. This and this gives us 0.32 and so on.

(Refer Slide Time: 30:24)


Question

There is an investment of Rs 200000. It is expected to grow by 15% with SD = 25%. What would be the change in value if we rule out the worst 2%? What should be the expected return if loss is 20%? If the period of investment is doubled, would the loss double?

$z = -2.05$ for area = 0.02
Change = $\mu + z\sigma = 15 - 2.05 \times 25 = -36.25\%$
Loss of 72500

$\mu - 2.05 \times 25 = -20; \mu = 31.25\%$

No. Risk becomes smaller with longer times.



Now, and there is an investment of 200,000 it is expected to grow by 15 percent with a standard deviation of 25 percent. What would be the change in value if we rule out the worst 2 percent? So, when we have this worst 2 percent. So, for area to the left of 0.02 by the less scenario. So, were area equal to 0.02 which is the worst 2 percent z is minus 2.05. So, they worth can reduce up to the point of z is equal to minus 2.05. So, $\mu + z\sigma$ will be the value of X , because z is equal to X minus μ by σ . So, X will be $\mu + z\sigma$ which can go to minus 36.25 percent.

So, which would give us loss of 72,500, what would be the expected return if the losses 20 percent? So, $\mu - z\sigma = -20$ and then we get a certain value of μ . So, μ is the 2.05 into σ 's minus 20 so, we get the certain value for μ . If the period of investment is doubled would the loss double move, rest become smaller as we move along with longer times. Therefore, if the period of investment is doubled, we also realize the standard deviation does not get doubled. So, the rest become smaller has been increase. And therefore, the loss would not double the loss would be less than double the value. So, this is some examples to study the normal distribution to understand the standard normal table to understand the z values, and to use them to solve some simple problems with the assumption of the normal distribution.