

Statistical Learning for Reliability Analysis
Dr. Monalisa Sarma
Subir Chowdhury of Quality and Reliability
Indian Institute of Technology – Kharagpur

Lecture – 14
Tutorial on Continuous Probability Distributions Functions (Part 1)

Hello guys. So, today we will be discussing some problems on the topics which we have covered in the last 2 lectures basically, that is lecture 12 and lecture 13 that is, we have discussed the different continuous probability distribution based on that we will be covering some of the problems so it is a tutorial class.

(Refer Slide Time: 00:44)

The slide is titled "Concepts Covered" and lists the following items:

- Solving objective type questions
 - To test the level of understanding from Lecture 12-13
- Problems to ponder
 - To build problem solving aptitude

The slide also features a video feed of the instructor, Dr. Monalisa Sarma, in the bottom right corner. At the bottom of the slide, there are logos for NPTEL and IIT Kharagpur, along with the text "Monalisa Sarma IIT KHARAGPUR".

So, first we will be solving some of the objective type question which will be essentially test the level of understanding and then we will solve few problems.

(Refer Slide Time: 00:56)

The slide is titled "Question-4.1" and contains the following text:

T 4.1: Classify the following random variables as discrete or continuous:

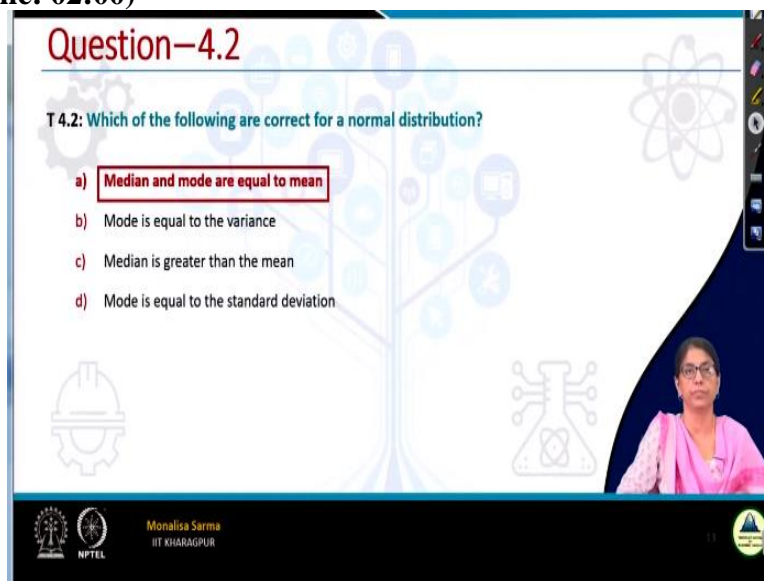
- a) X : the number of automobile accidents per year in a city. [Discrete]
- b) Y : the length of time to play 18 holes of golf. [Continuous]
- c) M : the amount of milk produced yearly by a particular cow. [Continuous]
- d) N : the number of eggs laid each month by a hen. [Discrete]

The slide also features a video feed of the instructor, Dr. Monalisa Sarma, in the bottom right corner. At the bottom of the slide, there are logos for NPTEL and IIT Kharagpur, along with the text "Monalisa Sarma IIT KHARAGPUR".

So, first the objective type equations the very easy questions, I am sure all of you will be able to answer it, but again, my request is that like in all the tutorial classes be do not see the answer immediately first try to answer yourself if you cannot, then go to the answer. So, classify the following random variable as discrete or continuous X the number of automobile accidents per year in a day.

So, it is a discrete variable or continuous variable number of automobile accidents. So, number of automobile accident, it cannot take any possible values within an interval. So, it has to be a discrete variable. Then next is length of the time to play 18 holes of golf. So, when we talk of a length of time so, time it can take any value in an interval, so it has to be a continuous variable. Similarly, the amount of milk produced merely by a particular cow that also it is a continuous variable. Then the number of eggs laid each month by a hen that number of $e m x$ that has to be an integer value. So, it has to be a discrete variable.

(Refer Slide Time: 02:00)



The screenshot shows a presentation slide titled "Question-4.2" with the question: "T 4.2: Which of the following are correct for a normal distribution?". There are four options listed: a) Median and mode are equal to mean, b) Mode is equal to the variance, c) Median is greater than the mean, and d) Mode is equal to the standard deviation. Option (a) is highlighted with a red box. The slide also features a small video inset of a woman in the bottom right corner and logos for NPTEL and IIT Kharagpur at the bottom.

So, which of the following are correct for a normal distribution, for a normal distribution we have seen the mean is equals to the median is equals to the mode. Mode is always the highest value and in normal distribution, we have seen it is symmetric on towards the mean and 50% of the value is on the left side of the mean, 50% of the value is on the right side of the mean. So, that is why that is also the median. So, we have already seen at median we get the high speed. So, median and mode are equal to mean that is the correct answer.

(Refer Slide Time: 02:34)

Question-4.3

T 4.3: Let X be the distance in miles from their present homes to University campus for a group of student. Then X is:

- A categorical (nominal) variable
- A continuous variable**
- A discrete variable
- A parameter

Then let X be a distance in miles from their present homes to university campus for a group of students. Then X is what is X ? X basically is the distance from their space wherever they stay the hostel or whatever it is from the hostel to the university campus. So, if X is a distance. Distance means it can take any value within a range. So, it has to be a continuous variable.

(Refer Slide Time: 03:01)

Question-4.4

T 4.4: Specify the distribution for the following:

- Random variable specifying the time to first occurrence of an event. [Exponential Distribution]
- Random variable specifying the time to k th occurrence of an event. **[Gamma Distribution]**

Specify the distribution for the following, random variable specifying the time to first occurrence of an event time to first occurrence of an event we have seen what distribution it is this? This is a exponential distribution than random variables specifying the time to k th occurrence of an event, k th occurrence of an event mentioned in a time we know that we are interested in finding out a specific number of events that will happen, when a specific number of events that will happen at done it has to be a gamma distribution.

Similar kind of thing we have seen in negative binomial distribution is not it? Very negative binomial distribution, it was the number of trials were fixed and we are interested in a fixed number of events, fixed number of success or failure whatever it is in a fixed number of trials are sorry number of trials are not fixed and negative binomial number of trials are not fixed, but the number of events are fixed. So, here also random variable is the time so here the time is not fixed, but we are interested till we see kth event k events so that is the gamma distribution.

(Refer Slide Time: 04:12)

Question-4.5

T 4.5: The time between the arrival of one customer and the arrival of the next customer has an exponential distribution and is independent of previous arrivals. The number of customers that will arrive during the next hour has:

- a) A Poisson distribution
- b) A Gamma distribution
- c) An exponential distribution
- d) A binomial distribution

Monalisa Sarma
IIT KHARAGPUR
NPTEL

Interesting question the time between the arrival of 1 customer and arrival of the next customer has an exponential distribution and is independent of the previous arrivals. That is given the time between the arrival of 1 customer and arrival of the next customer it has an exponential distribution, that is an exponential distribution arrival between 2 events. We already know arrival between 2 events or time for arrival of the very first event that is basically we consider rate of the event it is all these are the model by exponential distribution.

We already know this and what it is asking the number of customers that will arrival during the next hour question is something else it has nothing to do with exponential distribution. It has specified data time between 2 events as it follows exponential distribution the number of customer that will arrive during the next hour. So, the number of customers that will arrive during the next hour that we have. So, what this is definitely Poisson distribution?

We know Poisson distribution we have seen given the average arrival rate we are interested in the number of events that will arrive in a particular length of time or in a space whatever it

is so that is Poisson distribution here also, it is the same question the number of customers that will arrive during the next hour, so it is a Poisson of distribution.

(Refer Slide Time: 05:32)

The screenshot shows a presentation slide titled "Question-4.6". The text on the slide is as follows:

Question-4.6

T 4.6: State whether the given statements are True or False:

- a) The probability that a continuous random variable lies in the interval 4 to 7, inclusively, is the sum of $P(4) + P(5) + P(6) + P(7)$.
- b) Gamma distribution is a special case of the exponential distribution.

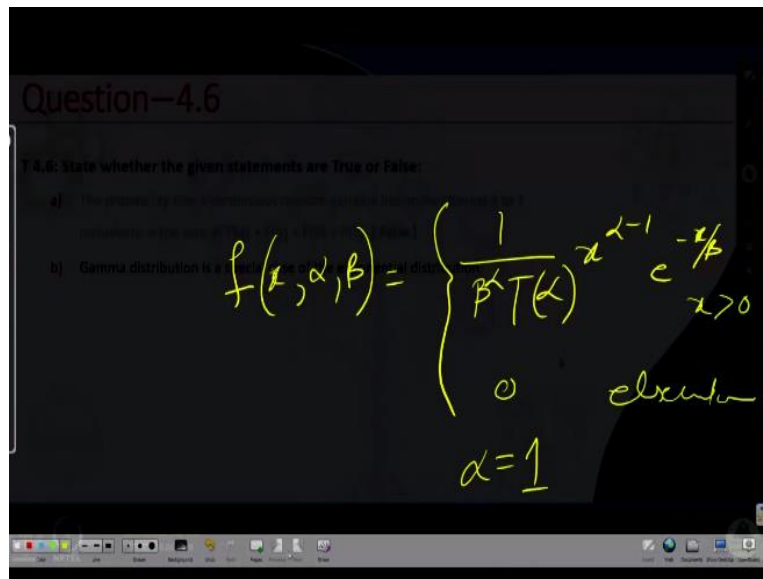
The slide also features a video inset of a woman in a pink shirt, and logos for NPTEL and IIT Kharagpur at the bottom.

Here we to say true or false. The probability that a continuous random variable lies in the interval 4 to 7 inclusively is the sum of $P 4 + P 5 + P 6 + P 7$ it true or this false true? No, it is false why when we are talking about continuous random variable the probability cannot be in a continuous random variable we cannot specify the probability at a specific point I get the probability at P 4 will be equal to 0, probability at P 5 will be equal to 0, we cannot specify probability at a specific point whenever we have to specify probability it has to be within an interval.

So, when you are talking about probability that lies in the interval 4 to 7, it has to be integration of the probability density function whatever we have integration of probability density function is F_x . So, integration of F_x from 4 to 7 that will be the probability but not this whatever it is given here. So, this is false. Now, next is gamma distribution is a special case of exponential distribution.

Is gamma distribution is a special case of exponential distribution? No, gamma is not a special case of exponential distribution. In fact, exponential distribution is a special case of gamma distribution.

(Refer Slide Time: 06:45)



Let us see why if you can remember we have already done this, but still if you let us see whether you can remember it or not. So, like remember the formula for gamma distribution, so, we are interested in say time t and gamma distribution at 2 parameter α and β . So, what was the expression it is $1 / \beta^\alpha$ there is a gamma function of $\alpha x^{\alpha-1} e^{-x/\beta}$. This is if you have considered here x then this will also be an x and let us make this x .

So, this is for x greater than 0 and it is 0 elsewhere. So, if x is the random variable and this is the gamma probability density function, so what is α here for gamma distribution we have 2 parameters α and β what is α here? α is the specific number of events that occur in time t α shape talks about a specific number of events that occurs in time t a fixed number of events as I told you we can draw an analogy from the negative binomial distribution so that is α .

And what is β ? β is the rate of events rate of events we can say means time that was the very first event has occurred or the time between 2 events that is basically this is gamma distribution. Now, in this gamma distribution, if $\alpha = 1$ if we specify $\alpha = 1$ then the resulting distribution is nothing but an exponential distribution. So, we say exponential distribution is a special case of gamma distribution. So, this is also false

(Refer Slide Time: 08:35)

Problem-4.7

T 4.7: $f(x) = \begin{cases} \frac{1}{b-a}, & \text{for } a \leq x \leq b \\ 0, & \text{elsewhere} \end{cases}$

Find $F(x)$.

Monalisa Sarma
IIT KHARAGPUR

So, now we will be solving some bigger problems. So, here this is equation here $f(x)$ is given that is called the probability density function is given this is a probability density function for uniform distribution we can see it is a uniform distribution and it is a range of our random variable that is here random variable is x random variable it ranges from a to b . So, in this range of a to b random variable can take values within the interval a to b and what is the probability density function? The probability the density function is $1 / b - a$.

So, now, what we are asked? We are asked to find the $F(x)$ that is the cumulative distribution function; we have already learned what is the cumulative distribution function? Cumulative distribution function means the value of the function t from $-\infty$ to that level to $2x$ when you are considering x $F(x)$ means what is the probability of getting a value from $-\infty$ to x probability to getting a value from x taking a value from $-\infty$ to x so that is cumulative distribution function.

So, here there are see 3 different reasons one is before a that is the problem the density function $F(x) = 0$ before it is 0 and after this what happens? After this is also 0 . Now we have to find out when you are considering effects so accordingly we will consider it in 3 different states. First we will find out till a what it will be?

(Refer Slide Time: 10:09)

Problem-4.7 : Solution

T.4.7: $f(x) = \begin{cases} \frac{1}{b-a}, & \text{for } a \leq x \leq b \\ 0, & \text{elsewhere} \end{cases}$
Find $F(x)$.

$f(x) = \frac{d-c}{b-a}$

Then,

$$F(x) = \begin{cases} 0, & \text{for } x < a \\ \frac{x-a}{b-a}, & \text{for } a \leq x < b \\ 1, & \text{for } x > b \end{cases}$$

Monalika Sarma
IIT KHARAGPUR

So, for that to find it out first thing is that it is better if we find out take an interval set from c to d , suppose if you are interested in finding out the probability density function probability of in the area of c to d or it is called the density function of the c to d how we will find out this is $d - c / b - a$ is not it? So, now, bringing it in 3 different steps first service from till a to 0 definitely, because still all the values are 0 . So, even if it were to get cumulative, this even if you make it cumulative 0 this whole reason is 0 .

So, if we get 0 till x is less than a and then x values goes from till a to b what it will be $x - a$ whatever x takes value whichever suppose x is value here than what it is $x - a / b - a$ if it takes here whatever it is $x - a / b - a$, so, till b this will be the value of $x - a / b - a$ this will be the F_x that is cumulative value and we know area of the curve is always one, area under the curve is always one that means, when we call this $F_x dx$.

So, if we take integration of $f_x dx$ and here it is ranging from a to b this will be equals to 1 other portion it is everything is 0 so this will be definitely equal to 1 . So, that means, once it reaches b it has become 1 . So, after b definitely it will be 1 only after b all the values are 0 , but till b it is 1 so, when we add it up with b it becomes 1 so, $x = 1$ for x greater than b . So, here this is the figure for the cumulative distribution, this is a fixed a we got 0 than it raises this way this value is $x - a / b - a$ and then from here it reaches 1 so and it continues.

(Refer Slide Time: 12:24)

Problem-4.8

T 4.8: A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

$P(t < 2.3)$

So, simple question on normal distribution already we have seen normal distribution 1, 2 problems in our class only. So, here again we will be seeing 1 2 problems. So, a certain type of storage but battery lasts on average 3 years and standard deviation is 0.5 that means my μ is 3, my μ is 3 years and my standard deviation that is this is 0.5 assuming that the battery life is normally distributed find the probability that a given battery will last less than 2.3 years.

So, I want to find probability that might if I take a random variable whatever random variable you take the x t y any can any variable can do if my take a random variable as say t so, my t is would be less than 2.3. This is the probability I want to find a probability of t less than 2.3. So, how do I find out I will have to take it now to finding this out what I have to take the normal distribution probability density function.

Remember probability density function normal distribution function was a complicated function then we have to integrate it. So, less than 2.3 years means we have to integrate it from 0 to 2.3 years. So, integration of that will be complex. So, what remedy we had? We had a remedy of converting the normal random variable 2 standard normal random variable that is just converting it to Z value.

And then we have the standard normal tables available in the literature books and net anywhere it is freely available downloadable, so we can cancel those lookup table and can get the value. So for that, we will have to find out the Z value corresponding to this 2.3.

(Refer Slide Time: 14:15)

Problem-4.8 : Solution

- The given figure showing the distribution of battery lives and the desired area.
- To find $P(X < 2.3)$, we need to evaluate the area under the normal curve to the left of 2.3
- This is accomplished by finding the area to the left of the corresponding Z value.
- Hence, we find that

$$Z = \frac{2.3 - 3}{0.5} = -1.4,$$

and then, using Table 1, we have

$$P(X < 2.3) = P(Z < -1.4) = 0.0808$$

T 4.8: A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

Monalisa Sarma
IIT KHARAGPUR

So, this is what this is the mean, this is the standard deviation 0.5. So, this 2.3 is essentially we want to find this probability that our value is less than 2.3 means essentially we want to find out this area, probability the area under this curve is nothing but the area under this curve. So, this is we have to calculate the Z value how do we calculate the Z value? Z value is $x - \mu / \sigma$. So, we got the Z value is - 1.4.

So, this 2.3 is nothing but if we draw a standard normal distribution, so standard normal distribution, this is 0 mean is 0 and standard deviation is 1. So, this -1.4 will be somewhere here, -1.4. So, we need to find out this area, so we can cancel it from the table. So, how do we get it from the table?

(Refer Slide Time: 15:17)

Problem-4.8 : Solution

Table 1

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.496	0.492	0.488	0.484	0.4801	0.4761	0.4721	0.4681	0.4641
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.2	0.4207	0.4168	0.4129	0.409	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.3	0.3968	0.3929	0.3890	0.3851	0.3812	0.3773	0.3734	0.3695	0.3656	0.3617
-0.4	0.3577	0.3538	0.3499	0.3460	0.3421	0.3382	0.3343	0.3304	0.3265	0.3226
-0.5	0.3186	0.3147	0.3108	0.3069	0.3030	0.2991	0.2952	0.2913	0.2874	0.2835
-0.6	0.2794	0.2755	0.2716	0.2677	0.2638	0.2599	0.2560	0.2521	0.2482	0.2443
-0.7	0.2403	0.2364	0.2325	0.2286	0.2247	0.2208	0.2169	0.2130	0.2091	0.2052
-0.8	0.2013	0.1974	0.1935	0.1896	0.1857	0.1818	0.1779	0.1740	0.1701	0.1662
-0.9	0.1623	0.1584	0.1545	0.1506	0.1467	0.1428	0.1389	0.1350	0.1311	0.1272
-1.0	0.1232	0.1193	0.1154	0.1115	0.1076	0.1037	0.0998	0.0959	0.0920	0.0881
-1.1	0.0842	0.0803	0.0764	0.0725	0.0686	0.0647	0.0608	0.0569	0.0530	0.0491
-1.2	0.0451	0.0412	0.0373	0.0334	0.0295	0.0256	0.0217	0.0178	0.0139	0.0100
-1.3	0.0060	0.0021	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-1.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-2.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

T 4.8: A certain type of storage battery lasts, on average, 3.0 years with a standard deviation of 0.5 year. Assuming that battery life is normally distributed, find the probability that a given battery will last less than 2.3 years.

Monalisa Sarma
IIT KHARAGPUR

So, in the table, you will see 1 - 1.4 corresponding to - 1.4 this value is 0.0808. Again, I am telling you this table is for cumulative value that means when I am taking is -1.4 that means

from $-\infty$ to -1.4 where probability is 0.0808. And that is what we wanted? We wanted this value this is from $-\infty$ to if this value corresponds to -1.4 $-\infty$ to -1.4 that is the area, this value from the table we are getting is 0.0808 so that is the probability.

(Refer Slide Time: 15:50)

Problem-4.9

T 4.9: In an industrial process, the diameter of a ball bearing is an important measurement. The buyer sets specifications for the diameter to be 3.0 ± 0.01 cm. The implication is that no part falling outside these specifications will be accepted. It is known that in the process the diameter of a ball bearing has a normal distribution with mean $\mu = 3.0$ and standard deviation $\sigma = 0.005$. On average, how many manufactured ball bearings will be scrapped?

Handwritten notes on the slide include a normal distribution curve with a shaded area between 2.99 and 3.01, and a box containing the inequality $2.99 \leq d \leq 3.01$ with z_1 and z_2 written below it.

So, that is how we solve this problem? So, this is the second problem is also similar line I will not go into the details. So, just for what is practice, if you have this problem. So, here the industrial process, the diameter a ball bearing is an important measurement, the buyer set specification for the diameter to be 3 plus - 0.01. That means my diameters should be the buyer sets the specification that the diameter buyer sets that my diameter should be in the range of if I am telling its diameter is d.

So, my diameter will be between 3.01 to 2.99 so, my diameter should be in this range. So, essentially if it is normally distributed, so mean is here it is given mean is 3, so, this is mean is 3. So, I want basically what is the probability it will fall in this area this is 3.01 this is 2.99. So, now corresponding to 2.9 and what is the Z value corresponding to 3.01 what is the Z value if we can find a Z value then from the table we can find out the corresponding probability of this.

If I tell it says Z 1 if it is a Z 2 from the table if I can find out what is the Z value? Z 2 will be set to value will be this and Z 1 value will be dispersion. So, Z 2 - Z 1 will give me this probability.

(Refer Slide Time: 17:40)

Problem-4.9 : Solution

The values corresponding to the specification limits are $X_1 = 2.99$ and $X_2 = 3.01$

The corresponding z values are

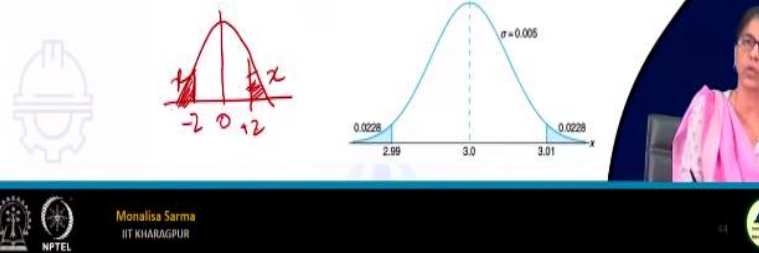
$$z_1 = \frac{2.99 - 3.0}{0.005} = -2.0 \quad \text{and} \quad z_2 = \frac{3.01 - 3.0}{0.005} = +2.0$$

Hence,

$$P(2.99 < X < 3.01) = P(-2.0 < Z < 2.0)$$

From Table 4, $P(Z < -2.0) = ?$

T 4.9: In an industrial process, the diameter of a ball bearing is an important measurement. The buyer sets specifications for the diameter to be 3.0 ± 0.01 cm. The implication is that no part falling outside these specifications will be accepted. It is known that in the process the diameter of a ball bearing has a normal distribution with mean $\mu = 3.0$ and standard deviation $\sigma = 0.005$. On average, how many manufactured ball bearings will be scrapped?



So, this is how it is been shown here, but I am not going into the details whatever I have discussed, it is a we will be able to solve it and show, here one thing I have just wanted to mention. See here we will calculating the Z value what I got Z 1? Z 1 when I use 2.99 it is to between 2.99 to 3.01 and I use 2.91 my Z 1 value is -2.0 and Z 2 value is 2.0. Here I mentioned one point like see normal distribution is very much a symmetric distribution.

So, if this is and say normal distribution if this is a standard normal distribution, so if I get whatever, say this is -2 and this is +2 from the table, I got whatever value is this whatever probability is this suppose this probability is x. If this probability is x, this probability will also be x because this is symmetric, so whatever value you got from by looking at the table form - 2.0, again you do not need to look up in a table to find out the value for 2.0 you will be same value since it is symmetric.

(Refer Slide Time: 19:02)

Problem-4.10

T 4.10: A machine produces rubber balls whose diameters are normally distributed with mean 5.50 cm and standard deviation 0.08 cm.

- What proportion of balls will have diameters less than 5.60 cm,
- between 5.34 and 5.44 cm?
- The balls are packed in cylindrical tubes whose internal diameters are normally distributed with mean 5.70 cm and standard deviation 0.12 cm. If a ball, selected at random, is placed in a tube, selected at random, what is the distribution of the clearance? (The clearance is the internal diameter of the tube minus the diameter of the ball.) What is the probability that the clearance is between 0.05 cm and 0.25 cm?

So that is what now the next question say a machine produces rubber balls whose diameters are normally distributed with mean of 5.50 centimetre and standard deviation of 0.08 centimetre. Some rubber balls are given was the mean and standard division are given, mean and standard deviation of the diameter, some rubber balls and the diameter. So, this is a rubber ball, it is diameters mean of rubber balls, mean of the diameters as well as the standard division of the diameters is become. What proportion of the balls will have a diameter less than 5.60 centimetre that will be able to solve it. We have done many questions.

So, I am not going to this, then what proportion of the balls will have between 5.34 and 5.4 per centimetre. That also you will be able to do into I have given the solution here, but I will not be discussing it will be able to do it. It is the same thing. Why I have kept this question is I want to discuss this portion. So, see the point c the balls are packed in cylindrical tubes whose internal diameters are normally distributed with mean of 5.70 centimetre and standard deviation of 0.12 centimetre.

So, this balls are kept in some central cylindrical tubes, the tubes diameter mean of this diameters are given 5.70 and standard deviation is 0.12 means, there are many tubes. So, when we find out what is the on an average what is the diameter of the tubes that way we find out the mean. So, the mean is 5.70 and standard deviation is 0.12 centimetre. If a ball selected at random is placed in a tube selected a random tube also have selected at random ball also have selected.

A random means balls diameter size is given mean diameter size is given on an average that is the diameter size similarly, there are many 2 tubes on an average the diameter size is given this is its diameter sizes pipe I mean average diameter size is 5.70 that is the mean. So, I have picked randomly I pick on ball and I picked what to say cylindrical tube also randomly and I have put the balls in the tube now mind it.

The cylindrical tube is also normally distributed the diameter is normally distributed and the balls have also normally distributed this we have already seen. Now, it is asking what is the distribution of the clearance? So, what is clearance when it is made, that clearance is the internal diameter of the tube minus the diameter of the ball. So, this is we have a tube so, here I am putting the balls the internal clearance, what is a clearance?

Clearance is this diameter when I put the ball there will be some clearance someplace. So that clearance I have to find out that clearance. What is the probability that a clearance is between 0.25 and 0.05 and 0.25 centimetre my main intention of giving this question is to bring to your notice one important point here like what I want to say is that here if suppose X is a normal random variable, pay attention if X is a normal random variable, if Y is a normal random variable X is a normal random variable Y is a normal random variable.

When I say X is a normal random variable, what do I mean by that X is a normal random variable means all the values that the x will take along with its probabilities, it will follow a normal distribution means when we plot it, it will follow a bell shaped curve, that is why we call it X is a normal distribution X is a normal random variable. Specifying a variable to be a normal random variable meaning is that only. So, it applies to all distributions similarly, so now X is a normal random variable Y is a normal random variable.

If then in such a case $X + Y$ or $X - Y$ both will be a normal random variable $X + Y$. So, let me take it as say another variable say M say this is another variable N . So, M is also a normal random variable N is also a normal random variable. Keep give notice to this if X is a normal random variable Y is another normal random variable, it is arithmetic operation where there may be plus or minus the resulting will be again in normal random variable.

So, here the clearance is that means the clearance that value that we will get from the clearance that is also a normal random variable clearance will be minus. Suppose this is the x is the normal random variable their specifying the diameter of the cylinder, why is there normal random variables specifying the diameter of the ball. So, when we talk about the clearance, clearance we talked means that is $X - Y$. So, definitely $X - Y$ is also a normal random variable.

So, this is first question, what is the distribution of the clearance? Distribution of the clearance is normal distribution that is the answer. Now secondly when we; found out the arithmetic combination of 2 normal variables now what about the mean and standard deviation. So, if X is a normal random variable, Y is normal random variable and then $X + Y$, that is M . M is a normal random variable, the mean of M will be mean of $X +$ mean of Y . So, again, I mentioned the mean of M will be mean of $X +$ mean of Y .

Standard Deviation of M will be or variance of M will be variance of X + variance of Y. what will be added for means it will be added for variance also it will be added, but see for N X is a normal random variable Y is a normal random variable X - Y is a normal random variable. Then means of N is μ of X that is mean of X - mean of Y since this is minus mean of X mean of N. Mean of n is mean of X - mean of Y that means, mean of X is taken is a μ X that means, μ X - μ Y.

Whatever the variance of N variance of N will be variance of X + variance of Y it will not be subtracted it will be added though it is minus it will not be subtracted it will be added we have seen the laws of variance when we have discussed Poisson distribution if you have forgotten please refer to those laws variance is always added it is never subtracted even if the random variables are subtracted. So, to bring home this point, this is the intention of giving this question as it is the solving of this is same what we have learned till now.

(Refer Slide Time: 25:50)

Problem-4.10 : Solution

a) Proportion of balls having diameters less than 5.60 cm

$$Z = \frac{(5.60 - 5.50)}{0.08} = 1.25$$

Using Table 5, $P(Z < 1.25) = 0.8944$

b) Between 5.34 and 5.44 cm

$$Z_1 = \frac{(5.34 - 5.50)}{0.08} = -2.0$$

$$Z_2 = \frac{(5.44 - 5.50)}{0.08} = -0.75$$

Hence probability of being between 5.34 cm and 5.44 cm is given by

T 4.10: A machine produces rubber balls whose diameters are normally distributed with mean 5.50 cm and standard deviation 0.08 cm.

b) What proportion of balls will have diameters between 5.34 and 5.44 cm?

Monalisa Sarma
IIT KHARAGPUR

So, this is I am just skipping you will be able to do it second b also am speaking, skipping it in be able to do it.

(Refer Slide Time: 26:01)

Problem—4.10 : Solution

c) The probability that the clearance is between 0.05 cm and 0.25 cm

If X is the diameter of the ball and Y is the diameter of the tube the clearance is $Y - X$. This will be normally distributed with

mean = $5.70 - 5.50 = 0.20$

variance = $0.08^2 + 0.12^2 = 0.0208$

standard deviation = 0.1442

Now

$$z_1 = \frac{(0.05 - 0.20)}{0.1442} = -1.040 \quad \text{and} \quad z_2 = \frac{(0.25 - 0.20)}{0.1442} = 0.347$$


so that probability of the clearance between 0.05 cm and 0.25 cm is given by


$$0.6357 - 0.1492 = 0.4865$$

Note: Interpolation has been used in reading the normal tables (Table 7 and Table 8), but the effect on the final answer is small


T 4.10: A machine produces rubber balls whose diameters are normally distributed with mean 5.50 cm and standard deviation 0.08 cm.

c) The balls are packed in cylindrical tubes whose internal diameters are normally distributed with mean 5.70 cm and standard deviation 0.12 cm. If a ball, selected at random, is placed in a tube, selected at random, what is the distribution of the clearance? (The clearance is the internal diameter of the tube minus the diameter of the ball.) What is the probability that the clearance is between 0.05 cm and 0.25 cm?





Monalisa Sarma
IIT KHARAGPUR



So, now coming to the third point, see here, the probability that a clearance is between we have to find out a probability that a clearance is between this and so now it is a question of simple normal distribution. So, if X is the diameter of the ball, Y is the diameter of the tube so, clearance $Y - X$ this will be normally distributed. Any normal distribution we can to define a normal distribution we will have to need 2 parameters what are those 2 parameters mean and variance or you can say mean and standard deviation whatever it is.

So, to define a normal random variable normal distribution, we need 2 parameters mean and variance. So, what is the mean? Mean is mean of $Y -$ mean of X that is mean this is my mean what is variance? Variance of $Y +$ variance of X so, this is my variance. So, this is variance standard deviation is the square root of variance. So, now I got mean value I got standard deviation value then I need to find out what is the probability that it will the clearance is between this and this value.

Similar type of problems we have done many problems we will be able to do it and very show and even if you cannot this answer is given here solution is given here you can see, but I suggest please do it by yourself. So, in this problem basically we have used an interpolation why interpolation sometimes like here value of Z_2 is 0.347, but in the table, we will not get that 0.347.

(Refer Slide Time: 27:40)

Problem-4.10 : Solution

Table 8

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.6293	0.63307	0.63684	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.6591	0.66276	0.6664	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.7054	0.70884	0.71226	0.71566	0.71904	0.7224
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.7549
0.7	0.75804	0.76115	0.76424	0.7673	0.77035	0.77337	0.77637	0.77935	0.7823	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.8665	0.86864	0.87076	0.87286	0.87493	0.87698	0.879	0.881	0.88298

T 4.10: A machine produces rubber balls whose diameters are normally distributed with mean 5.50 cm and standard deviation 0.08 cm.

c) The balls are packed in cylindrical tubes whose internal diameters are normally distributed with mean 5.70 cm and standard deviation 0.12 cm. If a ball, selected at random, is placed in a tube, selected at random, what is the distribution of the clearance? (The clearance is the internal diameter of the tube minus the diameter of the ball.) What is the probability that the clearance is between 0.05 cm and 0.25 cm?


You see in the table we will get 0.34 or 0.35 347 we would not get, so you will get the value of 0.34 or 0.35. So, in that case, what we can do? We can just add these 2 values divided by 2 that just interpolation that is all. So, that way we got the value that is all this is the solution of that problem

(Refer Slide Time: 28:03)

Problem-4.11

T 4.12: The average rate of water usage (thousands of gallons per hour) by a certain community is known to involve the lognormal distribution with parameters $\mu = 5$ and $\sigma = 2$. It is important for planning purposes to get a sense of periods of high usage. What is the probability that, for any given hour, 50,000 gallons of water are used?

$Z = \frac{\log(x) - \mu}{\sigma}$



So, you will see if you have any issues you can see it there yourself and not discussing that in details, because many similar problems we have discussed. Now, the next question, the average rate of water usage 1000s of gallons per hour by a certain community is known to involve lognormal distribution with parameter μ and σ , μ is 5 σ is 2 remember lognormal distribution if X is log normally distributed, that means log of X is normally distributed with μ and with whatever μ is specified and whatever σ is specified.

So, when a variable is log normally distributed, the log of that variable will be normally distributed, since the log of that variable will be normally distributed, we have already seen in lecture 13. If it is log normally distributed, then how to find a log normal distribution we have a probability density function again we have to find out the probability of occurrence of a particular value what we have to do we have to either calculate that probability density function means we have to do the integration for particular range.

Or we can look into the standard normal distribution table for lognormal distribution also, we look into the standard normal distribution, how they are our Z value here Z value suppose my normal thing suppose my X is what is a log normally distributed then my log of X will be normally distributed. So, my Z value will be $Z = \frac{\log x - \mu}{\sigma}$. So, from the standard normal distribution table whatever Z values suppose, this is the Z value whatever I got here if I get minus then it will be the left side of the 0 mean.

So, whatever value I get it from the table that will corresponds to the probability of whatever value is given.

(Refer Slide Time: 30:09)

Problem-4.11 : Solution

X follows a lognormal distribution.

$$P(X \geq 50000) = 1 - \Phi\left(\frac{\ln 50000 - 5}{2}\right)$$

$$= 1 - \Phi(2.9099)$$

$$= 1 - 0.9982$$

$$= 0.0018$$

T 4.11: The average rate of water usage (thousands of gallons per hour) by a certain community is known to involve the lognormal distribution with parameters $\mu = 5$ and $\sigma = 2$. It is important for planning purposes to get a sense of periods of high usage. What is the probability that, for any given hour, 50,000 gallons of water are used?

NPTEL
Monalisa Sarma
IIT KHARAGPUR

So, we are interested in finding out what is the probability that any given our 50000 gallons of water uses the probability that is called 50,000 gallons of water is used. So, at least 50,000 gallons of water that means property of X greater equals to 50,000, greater equals to 50,000 means, we have cumulative means we have from $-\infty$ to 50,000. So, in case of log normal it is not minus always it is not lognormal it is for positive scale we have seen that so, in the from the table we can get from one value from 0 to 50,000.

So, when we are interested in finding greater than so, it will be 1 minus of that, so 1 minus of we call it phi I told you yesterday in my last class. So, log of 50,000 - 5 / 2. So, whatever we got.

(Refer Slide Time: 31:04)

Problem-4.11 : Solution

T 4.11: The average rate of water usage (thousands of gallons per hour) by a certain community is known to involve the lognormal distribution with parameters $\mu = 5$ and $\sigma = 2$. It is important for planning purposes to get a sense of periods of high usage. What is the probability that, for any given hour, 50,000 gallons of water are used?

Z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.5279	0.53188	0.53586
0.1	0.53983	0.5438	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61793	0.62179	0.62564	0.62947	0.63329	0.6371	0.64089	0.64467	0.64843	0.65218
0.4	0.65594	0.65969	0.66343	0.66715	0.67086	0.67455	0.67822	0.68188	0.68552	0.68914
0.5	0.69274	0.69632	0.69988	0.70342	0.70693	0.71042	0.71389	0.71734	0.72077	0.72418
0.6	0.72757	0.73094	0.73428	0.7376	0.74089	0.74416	0.74741	0.75064	0.75385	0.75704
0.7	0.76021	0.76337	0.76651	0.76962	0.77271	0.77577	0.77881	0.78183	0.78482	0.78779
0.8	0.79074	0.79369	0.79661	0.79951	0.80238	0.80523	0.80806	0.81087	0.81366	0.81643
0.9	0.81919	0.82196	0.82471	0.82744	0.83015	0.83284	0.83551	0.83816	0.84079	0.8434
1.0	0.84588	0.84854	0.85117	0.85378	0.85637	0.85893	0.86147	0.86398	0.86647	0.86894
1.1	0.87138	0.87384	0.87627	0.87868	0.88107	0.88344	0.88578	0.8881	0.89041	0.8927
1.2	0.89497	0.89721	0.89942	0.90161	0.90377	0.90591	0.90803	0.91013	0.91221	0.91427
1.3	0.91631	0.91831	0.92029	0.92225	0.92419	0.92611	0.92801	0.92989	0.93175	0.93359
1.4	0.93541	0.9372	0.93898	0.94074	0.94248	0.9442	0.94591	0.9476	0.94927	0.95093
1.5	0.95257	0.95421	0.95583	0.95743	0.95901	0.96057	0.96212	0.96365	0.96517	0.96667
1.6	0.96815	0.96961	0.97106	0.97249	0.97391	0.97531	0.9767	0.97807	0.97943	0.98078
1.7	0.98211	0.98344	0.98475	0.98604	0.98731	0.98857	0.98981	0.99103	0.99224	0.99343
1.8	0.9946	0.99577	0.99692	0.99805	0.99916	1.00025	1.00132	1.00237	1.0034	1.00441
1.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
2.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.999
2.1	0.99903	0.99906	0.9991	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929

So, this value we can look at from the table. So, what was the value? Value was 2.9099 we can consider it is a 2.91 so, if you can consider it as 2.91 so, this is 2.91 corresponding this is value 0.99819. So, 1 minus of that will give me the value.

(Refer Slide Time: 31:26)

Problem-4.12

T 4.13: Suppose it has been observed that gun tube failures occur according to the lognormal distribution with $mean = 7$ and $std = 2$. Find the reliability for a 1000 round mission.

$$\frac{F(1000)}{1 - F(x)}$$

So, next question suppose it has been observed that the gun tube failure occur according to the lognormal distribution with mean = 7 standard division is 2 find the reliability for 1000 round mission. We have to find out the reliability for 1000 round mission how do we find out

the reliability for a 1000 round mission? So, it is also that means when we are talking of reliability for 1000 round mission.

First we will find out that cumulative probability that is F of 1000 that is the gun to failure this is cumulative failure till 1000 hours 1000 round mission we have to find out what is the probability of failure till 1000 round mission that is F of 1000. F of 1000 we have seen how we can find it out cumulative we can easily find it out from the lookup table. So, when we got to F of 1000 then reliability is nothing, reliability is what, reliability is just $1 - F(t)$. So, here t is 1000 so $1 - F(1000)$ will give me reliability.

Reliability is failure free if till this if this is the failure question then what is the reliability? Reliability is $1 - F$ of 1000.

(Refer Slide Time: 32:44)

The slide displays the solution to Problem 4.12. It includes a normal distribution curve with a mean of 7 and a standard deviation of 2. The z-value for t=1000 is calculated as $z = \frac{\ln(1000) - 7}{2} = -0.05$. The area under the curve to the left of z = -0.05 is 0.48, which is F(1000). Therefore, the reliability R(1000) is $1 - 0.48 = 0.52$. A woman is visible in the bottom right corner of the slide, and the NPTEL logo is in the bottom left.

Problem-4.12 : Solution

Given, mean = 7 and std (σ) = 2

Now,

At t = 1000 rounds

$$z = \frac{\ln(1000) - 7}{\sigma} = -0.05$$

And,

$$\phi(z) = \phi(-0.05) = 0.48$$

Therefore,

$$F(1000) = 0.48$$
$$R(1000) = 1 - 0.48 = 0.52$$

So that Reliability at 1000 round is **0.52**

T 4.12: Suppose it has been observed that gun tube failures occur according to the lognormal distribution with mean = 7 and std = 2. Find the reliability for a 1000 round mission.

So, at 1000 round we got Z value we have calculated so, this is the Z value so, from the table we will find out what is the value of this.

(Refer Slide Time: 32:56)

Problem-4.12 : Solution

T 4.12: Suppose it has been observed that gun tube failures occur according to the lognormal distribution with $\text{mean} = 7$ and $\text{std} = 2$. Find the reliability for a 1000 round mission.

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5	0.496	0.492	0.488	0.484	0.4801	0.4761	0.4721	0.4681	0.4641
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.2	0.4207	0.4168	0.4129	0.409	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.352	0.3483
-0.4	0.3446	0.3409	0.3372	0.3336	0.33	0.3264	0.3228	0.3192	0.3156	0.3121
-0.5	0.3085	0.305	0.3015	0.2983	0.2946	0.2912	0.2877	0.2843	0.281	0.2776
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.7	0.242	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.8	0.2119	0.209	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.166	0.1635	0.1611
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.123	0.121	0.119	0.117

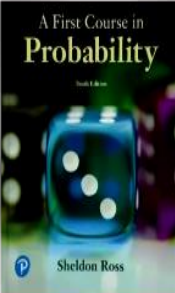
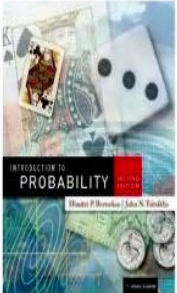
Monalisa Sarma
IIT KHARAGPUR

So, 0.5 this is the value of 0.4801 this table you will be getting in any appendix of any product statistics book or if and you can download it standalone also you can just Google it, you will get it and get a link and you can download it this table are freely available so we got is 0.48. So, F of 1000 is 0.48. So, R of 1000 is $1 - 0.48$, so reliability at 1000 round is 0.52.

(Refer Slide Time: 33:33)

REFERENCES

- ① Sheldon Ross , A First Course in Probability (fourth ed.), Macmillan College Publishing, New York (2013)
- ② D. P. Bertsekas and J. N. Tsitsiklis, Introduction to Probability. Nashua, NH: Athena Scientific, 2008
- ③ W. Feller, An Introduction to Probability Theory and Its Applications, vol. 2. New York: Wiley, 1971

Monalisa Sarma
IIT KHARAGPUR

So that is all in today's lecture there we will be discussing few more problems. And for that I will take one more tutorial class before going to our next topic. So, then these are the references which I have already mentioned before an earlier classes and thank you guys.