

Stochastic Structural Dynamics
Prof. Dr. C. S. Manohar
Department of Civil Engineering
Indian Institute of Science, Bangalore

Lecture No. # 3

Scalar random variables-2

(Refer Slide Time: 00:23)

Recall

Random variable is a function from sample space into real line such that

(1) for every $x \in R$, $\{\omega : X(\omega) \leq x\}$ is an event,

(2) $P(\omega : X(\omega) = \pm\infty) = 0$

Discrete
Continuous
Mixed

Descriptions

- Probability distribution function
- Probability density function
- Probability mass function

PDF of a Mixed RV

The graph shows a curve labeled $F_X(x)$ on the vertical axis and x on the horizontal axis. The curve starts at the origin and increases. It has a jump discontinuity at a point x_0 , where the function value increases abruptly. The jump is represented by a vertical dashed line with a solid dot at the top and an open circle at the bottom. The curve is smooth and continuous before and after the jump.

NPTEL

Will begin today's lecture by, quickly recalling what we were doing in the previous lecture. We defined, what is the random variable? Random variable is a function from sample space to the real line and we clarified the meaning of the set of all ω , where x of ω is less than or equal to x , as a sub set of the sample space on which we assigned probabilities. We describe random variable through functions; three functions - probability distribution function, probability density function, and probability mass function. A typical probability distribution function is shown here, this is a probability distribution function for a mixed random variable; a random variable can be discrete. Where the probability distribution proceeds through only jumps and in continuous random variable, it proceeds without any jumps and in mixed random variable, there can be structures, where it proceeds without any jumps and as well we can have jumps as shown here.

(Refer Slide Time: 01:43)

Recall (continued)

Bernoulli's random variable

$$P[X=0] = 1-p$$
$$P[X=1] = p$$

Binomial random variable

$$P[N=k] = {}^N C_k p^k (1-p)^{N-k} \quad k=0,1,2,\dots,N$$

Geometric random variable

$$P[X=k] = (1-p)^{k-1} p \quad k=1,2,\dots,\infty$$

NPTEL

We defined probability density function, as the derivative of this distribution function. Probability mass function is defined for discrete random variables, we considered several random variables, Bernoulli random variable is the basic building block. Here, we basically, deal with random variables, which take only 2 states, this is 0 is called success and 1 is called failure, and we assign probability P and $1 - p$. So, a sample space here is discrete in a binomial random variable, we deal with n repeated Bernoulli trials and ask, what probability of K successes in n trials is and we have shown, this to be a function of this kind, where K takes values from $0, 1, 2, \dots, n$.

(Refer Slide Time: 03:20)

Models for rare events: Poisson random variable

(a) We are looking for occurrence of an isolated phenomenon in a time/space continuum.

(b) We cannot put an upper bound on the number of occurrences.

(c) Actual number of occurrences is relatively small.

Examples: goals in football match (time continuum), defect in a yarn (1-d space continuum), typos in a manuscript (2-d continuum), defect in a solid (3-d continuum). **Stress at a point exceeding elastic limit during the life time of a structure.**

$$P(X=k) = \exp(-a) \frac{a^k}{k!}, k=0,1,2,\dots,\infty$$

Check

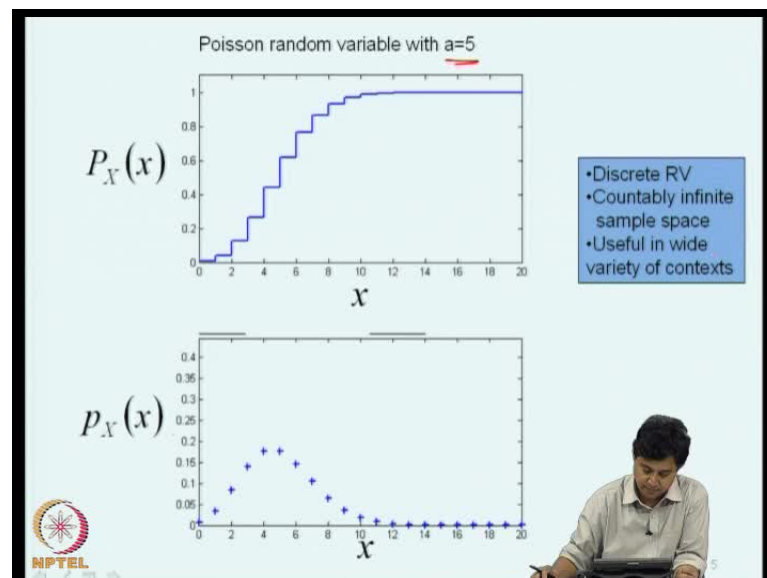
$$P(X \leq \infty) = \sum_{k=0}^{\infty} \exp(-a) \frac{a^k}{k!} = \exp(-a) \sum_{k=0}^{\infty} \frac{a^k}{k!} = \exp(-a) e^a = 1$$

NPTEL

So, here again, the sample space is finite, discrete and number of successes to first failure. We called the geometric random variable number of trials to the first success that is defined, suppose K_x is the number of trails for first success, if we get first success on the K trial, we have got failures on the first K minus 1 trial and success on the K trial here K runs from 1, 2 and it can go all the way up to infinity. So, here the sample space is countably infinite.

We also talked about models, for rare event that is where we were at the end of the last lecture here, we are looking for occurrence of isolated phenomena in time or space continuum, we cannot put an upper bound on the number of this isolated phenomena, but actual number of such occurrences is relatively small. Some of the examples are for example, number of goals scored in a football match, where we are looking at isolated phenomena in time continuum or we look for defect in a yarn in 1-dimensional space. This is 1-d space continuum or a typo graphical errors in a manuscript it is 2-d continuum or defects in a solid it is 3-d continuum or stress at a point exceeding, elastic limit during the life time of the structure.

(Refer Slide Time: 04:45)

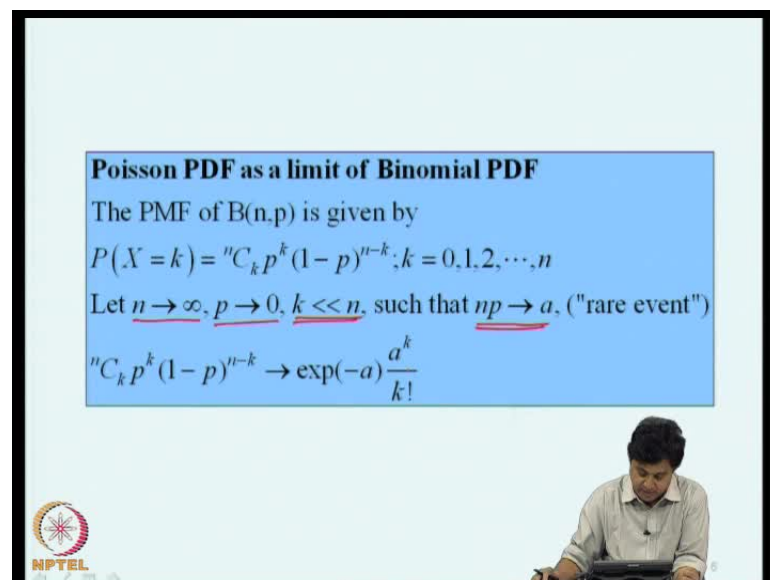


So, here we are looking for isolated phenomena in time continuum. In this class of problems the underlying random variables namely, the number of isolated phenomena is given by this particular function, here a is a parameter and K takes values of from 0, 1, 2 up to infinity.

So, the sample space here is countably infinite and λ is the parameter of the probability distribution function. We can verify that the probability of X less than or equal to infinity, we should be the probability of sample space is indeed equal to 1. This is how the probability distribution function for a Poisson random variable with parameter λ equal to 5 looks like and this is the associated probability density function expressed in terms of set of Dirac's delta function.

So, this is the discrete random variable with countably infinite sample space and it is going to be very useful in a variety of contexts and will come across this random variable quite often in later discussions.

(Refer Slide Time: 05:19)



Poisson PDF as a limit of Binomial PDF

The PMF of $B(n, p)$ is given by

$$P(X = k) = {}^n C_k p^k (1 - p)^{n-k}, k = 0, 1, 2, \dots, n$$

Let $n \rightarrow \infty$, $p \rightarrow 0$, $k \ll n$, such that $np \rightarrow a$, ("rare event")

$${}^n C_k p^k (1 - p)^{n-k} \rightarrow \exp(-a) \frac{a^k}{k!}$$

NPTEL

(Refer Slide Time: 06:33)

Proof

$${}^nC_k p^k (1-p)^{n-k} = \frac{n!}{(n-k)!k!} \left(\frac{a}{n}\right)^k \left(1-\frac{a}{n}\right)^{n-k}$$
$${}^nC_k \approx \frac{n!}{(n-k)!k!} \approx \frac{n^k}{k!}$$
$$p^k = \left(\frac{a}{n}\right)^k$$
$$1-p \approx \exp(-p)$$
$$(1-p)^{n-k} \approx \exp[-(n-k)p] = \exp(-a) \exp(kp) \approx \exp(-a)$$
$$\Rightarrow {}^nC_k p^k (1-p)^{n-k} \rightarrow \frac{n^k}{k!} \left(\frac{a}{n}\right)^k \exp(-a) = \frac{a^k}{k!} \exp(-a)$$

Now, we can obtain the Poisson probability distribution function, as a limit of the binomial probability distribution function, here for a binomial probability distribution function that is the for a binomial random variable with parameters n and p probability of k successes in n trials is given by this. Now, what we will do is we will make the number of trials to become very large and probability of success in any given trial going to 0, such that the product $n p$ goes to a constant and we will look at realizations of x for small values of k .

So, we are essentially, simulating the prescriptions for rare event namely, the probability of success is small, there is actually no bound on number of successes, but in practice that what we are looking for is number of success to be quite small related to n , in such a case the binomial distribution can be shown to become Poisson distribution. A simple proof for this can be constructed by manipulating the binomial distribution and imposing the limit that we are putting that is n goes to infinite p goes to 0, such that $n p$ goes to a constant, if we do that indeed 1 can prove that the binomial distribution goes to a Poisson distribution.

(Refer Slide Time: 06:59)

Example

- Consider $B(1000, 10^{-3})$
- $N = 1000; p = 10^{-3}$
- $P(X = 0) = {}^{1000}C_0 (10^{-3})^0 (1 - 10^{-3})^{1000-0} = 0.36769 //$

Poisson limit

$np = 1000 \times 10^{-3} = 1$

$k = 0 \ll 1000$

$P[X = 0] = \exp(-a) \frac{a^0}{0!} = \exp(-1) = 0.36787$ ←

NPTEL

We will consider some numerical example, so that we understand what exactly is meant here. Let us consider, a binomial distribution, where we are conducting 1000 trials and probability of success in any single trial is 10 to the power of minus 3. So, if you faithfully follow the binomial model a N is 1000 p is 10 to the power of minus 3 and, if I am seeking the probability of no successes in 1000 trials, this is a rare event by the way, this will be $1000 C_0 10$ to the power of minus 0 and this $1 - p$ to the power of 1000 the number is 0.36769. Now, this particular choice of parameters in N and p are suitable for application of Poisson limit simply, because n is sufficiently large p is sufficiently small and n p, which is parameter a is 1 here. And k that I am looking for a 0 that is indeed quite small compared with 1000. So, this is situation, where I could think of applying Poisson model for the binomial distribution and according to Poisson model, probability of x equal to 0, is exponential minus a. a to the power of 0 by 0 factorial and this turns out to be 0.36787, which compares reasonably with the number that we get from the binomial distribution.

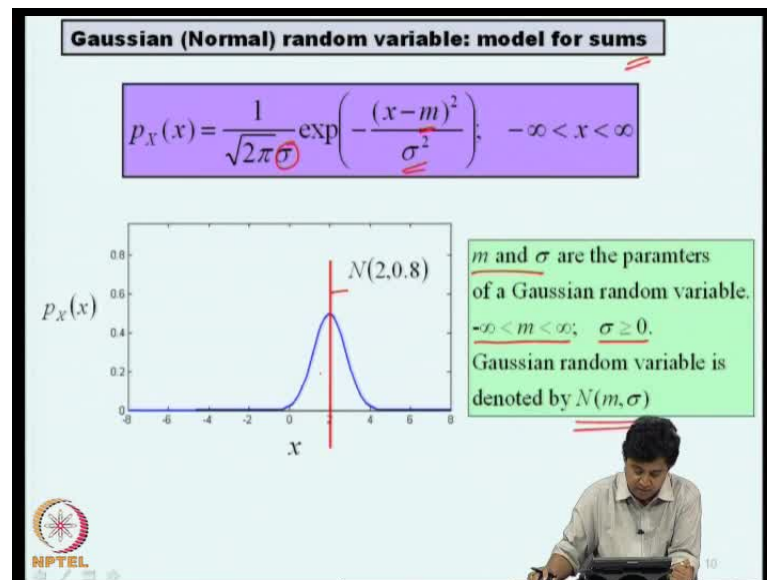
(Refer Slide Time: 08:30)

Example
Consider $B(3000, 10^{-3})$
 $P(X > 5) = ?$
 $P(X > 5) = 1 - P(X \leq 5) = \sum_{k=0}^5 {}^{3000}C_k (10^{-3})^k (1 - 10^{-3})^{3000-k}$
Try evaluating this!!

Poisson's limit
 $N = 3000; p = 10^{-3}; a = np = 3$
 $k \leq 5 \ll 3000$
 $P(X > 5) = 1 - \sum_{k=0}^5 \exp(-3) \frac{3^k}{k!} = 0.084$

Will consider another example, will consider a binomial random variable, with N equal to 3000 and probability of success is 10 to the power of minus 3. Now, the question I am asking is what is the probability of number of successes or greater than 5, here probability of X greater than 5 is nothing but 1 minus probability of X less than or equal to 5; so, you have to sum the probabilities of this is 1 minus k equal to 0 to 5, this I am summing from 0 to 5. So, this can be in principle can be evaluated, but you will see that you will have few numerical difficult is in doing this, but we can now explore if Poisson's limit is appropriate for this situation. So, number of trials is 3000 and probability of success is 10 to the power of minus 3. So, n appears sufficiently large P is sufficiently small. So, the limit n p the product n p is 3 and the range of k that I am interested in is k less than or equal to 5, which is quite small compared with 3000. So, the application of Poisson's limit ((C)) c appears reasonable.

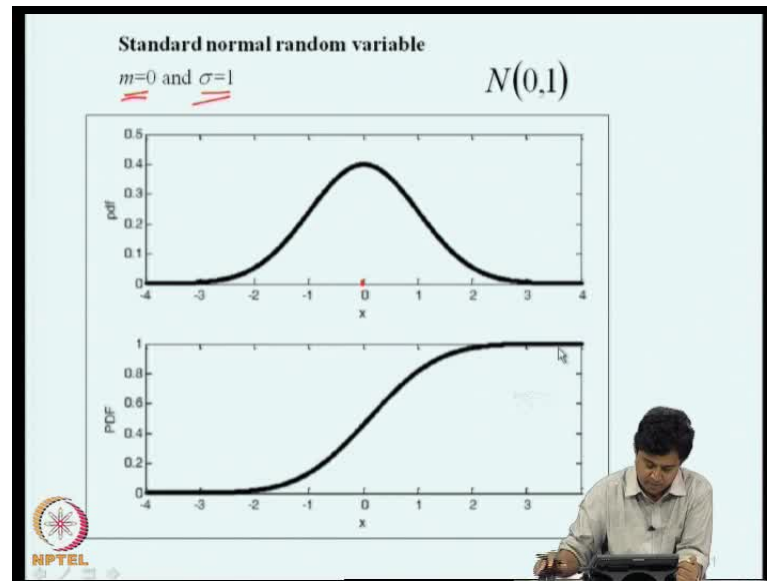
(Refer Slide Time: 10:08)



So, if we evaluate this probability, we get a number 0.084 and you can compare this number by evaluating, actually this expression. Now, I will consider another random variable, the Gaussian random variable also known as the normal random variable. We will see shortly that it helps us to model situations where, uncertainty emerge out of a summation mechanism where, different sources of uncertainties contribute to the final uncertainty in a parameter of interest and they all add up, I will make this statements more precise, as we go along but right now, we can look at the probability density function of a Gaussian random variable. So, here there is an exponential function and a square of the state variable is here. There are 2 parameters 1 is appearing, here that is m and other is σ , which also appears here, x takes value from minus infinity to plus infinity.

So, here, the sample space is the real line and there are 2 parameters m and σ . So, m and σ are the parameters and m itself can take values from minus infinity to plus infinity, whereas this parameter σ has to be non-negative.

(Refer Slide Time: 12:10)



A Gaussian random variable is denoted by $n\ m, \sigma$; n is for this word normal and this is m, σ . A typical plot of a normal probability density function for m equal to 2 and σ equal to 0.8 is shown here. This is a familiar bell like curve it which is symmetric about this line. This line is nothing but m equal to 2, you can see that this function is symmetric about m , that is, what it is shows and as x tends to infinity plus minus infinity the function decays to 0 and area under this curve is going to be 1. Will see, how to prove that shortly More on this, as we go along write now, we will try to see get familiarize with this probability density function is specific form of a normal random variable is when the parameter m is 0. And this parameter σ is 1. This is, This is the $n\ 0, 1$ normal random variable and the probability density function is here and it is symmetric about this line, it is symmetric about this line and the probability distribution function is a continuous curve you starts from 0 and goes to 1.

(Refer Slide Time: 12:52)

Gaussian random variable as a limit of Binomial random variable

Under the conditions
 $n \rightarrow \infty, npq \gg 1, np - \sqrt{npq} < k < np + \sqrt{npq}$ $q = 1-p$

it can be shown that

$$\binom{n}{k} p^k (1-p)^{n-k} \rightarrow \frac{1}{\sqrt{2\pi npq}} \exp\left[-\frac{1}{2} \left(\frac{k-np}{\sqrt{npq}}\right)^2\right]$$

The Gaussian random variable can also be obtained, as a limit of binomial random variable. Here, the limit that we consider is again, we take n to be large that is number of trials in the binomial random variable to be large and q is $1 - p$ and we demand that the product $n p q$ should be greater than or equal to far greater than 1 and k itself takes values in the neighborhood of $n p$; $n p$ plus square root $n p q$ and $n p$ minus square root $n p q$. And if these conditions are satisfied, we can show that the binomial distribution can be approximated by this Gaussian curve.

So, you recall now is started with a Bernoulli trial and constructed a binomial random variable, and I am using binomial random variable to construct Poisson and Gaussian random variables in that sense, a Bernoulli random variable is the building block for constructing various probability distribution functions.

(Refer Slide Time: 14:15)

Example

Let $n = 1000$; $p = 0.5$; $k = 500$.


$$npq = \frac{1000}{4} = 250; \sqrt{npq} = 15.81;$$

$$np - \sqrt{npq} = 500 - 15.81 \approx 485$$

$$np + \sqrt{npq} = 500 + 15.81 \approx 515.$$

It is verified that $485 < k = 500 < 515$.

$$P(X = 500) = {}^{1000}C_{500} (0.5)^{500} (1-0.5)^{1000-500}$$

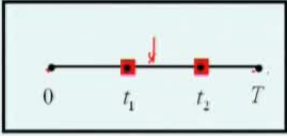


$$\approx \frac{1}{\sqrt{2\pi}15.81} \exp\left[-\frac{1}{2}\left\{\frac{(500-500)^2}{15.81}\right\}\right] = 0.1030$$


13

(Refer Slide Time: 15:40)

Random points:
Poisson models for counting & exponential models for waiting times

Consider a time interval $0, T$.
 Let us place n points randomly in $0, T$.
 Let $0 < t_1 < t_2 < T$.
 Define {success} = {point lies in the sub-interval t_1, t_2 }
 $P\{\text{success}\} = \frac{t_2 - t_1}{T} = p$; $P\{\text{failure}\} = 1 - p$
 Define: $X =$ number of points in t_1, t_2 .
 $P(X = k) = {}^n C_k p^k (1-p)^{n-k}$
 Now, consider $n \rightarrow \infty$, $\frac{t_2 - t_1}{T} \rightarrow 0$, $\frac{n(t_2 - t_1)}{T} \rightarrow a$
 $\Rightarrow {}^n C_k p^k (1-p)^{n-k} \rightarrow \exp(-a) \frac{a^k}{k!}$
 Let $\frac{n}{T} = \lambda$; $(t_2 - t_1) = t_a \Rightarrow a = \lambda t_a$
 $P(X = k) = \exp(-\lambda t_a) \frac{(\lambda t_a)^k}{k!}; k = 0, 1, 2, \dots, \infty$

Let's consider an example, where a n is 1000 p is point 5 and k is 1000 500 $n p q$ is q is also 0.5, because it is 1; 1 minus p $n p q$ tends have to be 250 and square root of that is around 15. Now, $n p$ minus square root $n p q$ is 485 and $n p$ plus square root $n p q$ is around 515. So, according to the Gaussian limits for binomial distribution, if k lies mind you, our, we are interested in k of 500, we are interested knowing what is x equal to 500. So, k is between 485 to 515 so, wherein a situation where the binomial distribution can be approximated by a Gaussian curve and we are getting this number 0.1030 this can be compared by actually, evaluating this.

Now, from the Poisson model, we derive what is known as an exponential random variable, how we will construct that? We will show that an exponential model helps us to model the waiting times between numbers of occurrence of isolated phenomena in Poisson model. So, let us consider a time interval 0 to t , time interval 0 to t and consider a subinterval t_1 to t_2 within this interval, suppose if I place randomly n points in 0 to t .

Now, we define the success, if that point lands in this interval, we say that it is success that is point lies in sub interval t_1 to t_2 . So, probability of success is t_2 minus capital t , which is p , because where placing the point randomly 0 to t . So, it can land anywhere with the same, probability of failure, where success is the point lies in t_1 to t_2 . The probability of failure is point lies outside t_1 and t_2 is 1 minus p . Now, if we place n number of points and we now ask the question what is the probability distribution function of number of points in t_1 to t_2 after we have finish placing n points in 0 to t . So, this follows binomial distribution and probability of x is equal to k is given by $n C_k p^k (1-p)^{n-k}$.

Now, we will consider the Poisson limit of this binomial distribution, we will allow the number of points to become large and we will allow the time interval t_2 to t_1 to become small. In such a way, that their product goes to a constant λ $(t_2 - t_1) / t$ by capital t goes to ∞ that mean, the $n p$ goes to λ in such a case, we know that the binomial distribution goes to the Poisson model. Now, we denote by n by t , we denote as λ and $t_2 - t_1$ as a , with that notation, λ becomes $\lambda t a$. Therefore, probability of x equal to k is given by the Poisson model $\frac{e^{-\lambda t a} (\lambda t a)^k}{k!}$, where k run from 0 to ∞ . So, this the Poisson model for random points occurring in time continuum 0 to capital T .

(Refer Slide Time: 18:45)

Exponential random variable

Define T^* = time for the first arrival

$P(T^* > t) = P(\text{no points in } 0, t) = \exp(-\lambda t)$

$\Rightarrow P_{T^*}(t) = 1 - \exp(-\lambda t)$

$\Rightarrow p_{T^*}(t) = \lambda \exp(-\lambda t); \lambda > 0$

Remarks

(a) T^* can be used to model life time of a structure.

(b) $P_{T^*}(t)$ is the generalization of the geometric random variable to the continuous case.

(c) Consider $P(T^* \leq t | T^* > t_0) = \frac{P(T^* \leq t \cap T^* > t_0)}{P(T^* > t_0)}$

$= \frac{P(t_0 < T^* \leq t)}{P(T^* > t_0)} = \frac{\{1 - \exp(-\lambda t)\} - \{1 - \exp(-\lambda t_0)\}}{\exp(-\lambda t_0)}; t \geq t_0$

$\Rightarrow P_{T^*}(t | T^* > t_0) = 1 - \exp\{-\lambda(t - t_0)\}; t_0 < t < \infty$

$\Rightarrow p_{T^*}(t | T^* > t_0) = \lambda \exp\{-\lambda(t - t_0)\}; t_0 < t < \infty$

Key Words: failure to observe an event up to time t_0 does not alter one's prediction of the length of time (from t_0) before the event will occur. The future is not influenced by the past.

Now, we define another random variable, T^* which is a time for the first arrival, you start from 0, what is the probability that T^* is greater than some time t . If the first point is going to arrive after T^* there will be no points in 0 to t . Therefore, probability of no points in 0 to t is given by exponential minus λt , where k is 0 in the Poisson model. Now, therefore the probability distribution function of T^* is given by 1 minus this, which is 1 minus exponential minus λt . And the associated density function is an exponential function shown here.

So, here I have shown in this plot, the probability density function for some value of λ , it is the exponentially decaying function. The area under this curve is 1 and the associated probability distribution function is this monotone non-decreasing function which starts with 0 and goes to 1, as this state variable goes to infinity.

This T^* can be used to model life time of a structure. So, in that definition of Poisson event, if we define the event to be, for example, extra set a point in the structure exceeding elastic limit and that if we all agree that it is the definition of failure, then the time that we need to wait till that happens in any way is a life time of the structure.

So, life time of the structure can be modeled as exponential random variables, if the occurrence of the failure events follow Poisson distribution, we came across a similar random variable, when we talked about discrete random variable, namely the geometric random variable. Geometric random variable was the number of trials for the first

success; So, the exponential random variable can be thought of as a generalization of the geometric random variable for the continuous case. This exponential distribution has an interesting property that can be clarified here, by considering the probability of T^* less than or equal to t condition on the fact that up to t_0 there was no event. So, that is by definition of conditional probability it is T^* less than or equal to t intersection T^* greater than or equal to t_0 and divided by probability of t greater than t_0 .

Now, if you use the exponential model for this and follow through the arguments here, you can show that the probability distribution of T^* given that there was no event, till t_0 is given by this exponential distribution. Here t runs from t_0 to infinity that means, it is continuous to be an exponential distribution except that it has now shifted to t_0 .

(Refer Slide Time: 22:20)

The slide is titled "Transformation of random variables". It features a diagram of a cantilever beam of length L , fixed at the left end and free at the right end. A downward point load P is applied at the free end. The beam's material properties are given as E, I . To the right of the beam is a rectangular cross-section with depth d and width b . Below the diagram, a text box asks: "If P, L, b, d, E are RVs, what is the pdf of $\delta = \frac{PL^3}{3EI}$?" To the right of this text box is a small diagram of a square with a dot in the center, with arrows pointing to it labeled "Uncertain" and an arrow pointing away labeled "Uncertain". In the bottom right corner, there is a small inset photo of a man sitting at a desk, looking at a laptop.

In words, if you want to express, what this actually is telling us, what it means is failure to observe an event up to time t_0 , does not alter once prediction of the length of the time from t_0 before an event will occur, the future is not influenced by the past that is the message, so this is also known as memory less property.

Now, we come to a new topic, namely transformation of random variables to clarify, what exactly is the problem here, we can consider a simple example of a cantilever beam carrying a tip load P , the cross section of the beam is rectangular, where d is the depth and

b is the breadth, l is the length, e is the young's modulus, I is the moment of inertia, area moment of inertia. Now, we knew that the tip deflection here is given by PL cube by 3EI. Now, if P L b d e are all random variables, then it automatically follows that delta is also a random variable. So, we will typically be interested in knowing, if P and E are random variables and if we know their probability distribution functions, what can we say about probability distribution of delta. This is the problem that we have to answer in structural engineering.

(Refer Slide Time: 24:15)

Basic problem of transformation of random variables

Let X be a RV; define $Y=g(X)$;
Given pdf of X , what is the pdf of Y ?

$$P(y < Y \leq y + dy) = p_Y(y)dy$$

$$= p_X(x_1) |dx_1| + p_X(x_2) |dx_2| + p_X(x_3) |dx_3|$$

$$\Rightarrow p_Y(y) = \frac{p_X(x_1)}{\left| \frac{dy}{dx} \right|_{x=x_1}} + \frac{p_X(x_2)}{\left| \frac{dy}{dx} \right|_{x=x_2}} + \frac{p_X(x_3)}{\left| \frac{dy}{dx} \right|_{x=x_3}}$$

In general

$$p_Y(y) = \sum_{i=1}^n \frac{p_X(x_i)}{\left| \frac{dy}{dx} \right|_{x=x_i}} \text{ with } x_i = g^{-1}(y)$$

Why modulus?
We wish to add probabilities.

The inputs to a system, if this is answer done, the output will also be answer. So, this input variable pass through our system through the physical loss and governs a system and produce uncertainty in the output. So, given the probability distribution of these uncertainties in the input and the definition of this system, the output is the random variable and what is the probability distribution of that output is the question.

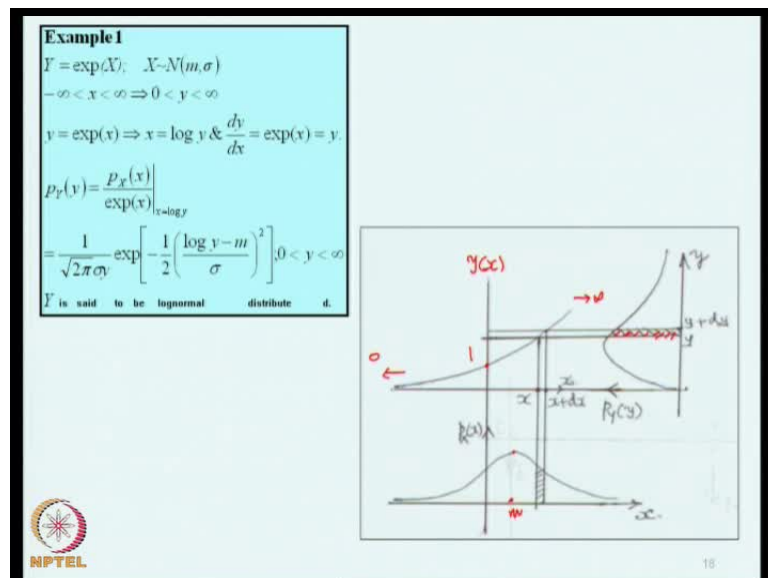
So, this is essentially a problem in transformation of random variables. So, let us make it more specific, let X be a random variable, we will define, we will define a function y is equal to $g X$; here, the question is given the probability density function of X , what is the probability distribution function of y , then of course the function g of X is also specified to make the meaning of this problem clear, what we do is we plot g of X versus x here this is x this is g of X and this function is my g of X and I will plot they the given probability density function of x below this say, let this be the probability density

function and this is the unknown probability density function of y which we need to determine.

Now, I consider the question, what is the probability that y lies between y and $y + dy$ this is the question that I am asking that is given by $p_y(y)$ of y into dy $p_y(y)$ of y is still unknown. So, if y lies between y and $y + dy$ where all x can lie. So, to see that we draw these lines here so it intersects g of X at 3 places this line from y it takes intersects here intersects here intersects here. So, if y lies between y to $y + dy$ x can lie here x_1 minus dx_1 to x_1 x_2 to $x_2 + dx_2$ or x_3 minus dx_3 to x_3

So, whenever x lies in the shaded regions, y will lie in this shaded region and these three events are mutually exclusive. So, this probability $p_y(y)$ of y dy is nothing but the sum of this probability. This the probability density function. Therefore, the probability that x lies in this shaded region is given by probability density function evaluated at x_1 into dx_1 right. So, that is the first step and p_x of x_2 into dx_2 plus p_x of x_3 into dx_3 , therefore, p_y of y is now given by p_x of x_1 divided by dy by dx at x is equal to x_1 and p_x of x_2 divided by dy by dx at x is equal to x_2 and for similarly for a third point

(Refer Slide Time: 27:37)

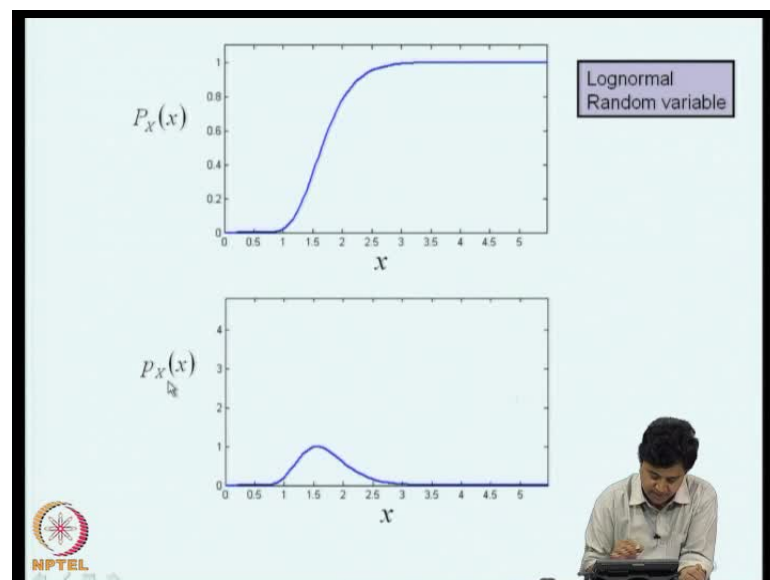


In general, if there are n possible routes like this the probability density function of y is given by the summation over all these routes, where x_i 's are the routes of the equation g of y x equal to y . Now, you have noticed that we are putting modulus, here it is just to ensure that we add the probabilities needs to be positive number. So, we are we wish to

ensure that when adding probabilities, there is no sign associated with probabilities to **make that** to ensure that we are putting this modulus.

So, we can consider in a few couple of examples, and see what we land from there the first example, we consider is we assume that X is a normal random variable with parameter m and σ and the function y is defined as exponential of x . Now, x is normal; therefore, the state variable here takes values from minus infinity to plus infinity, the since X lies between minus infinity to plus infinity based on this rule, we can certain that y lies between 0 to infinity. So, y is to positive y is nonnegative it lies from 0 to infinity.

(Refer Slide Time: 30:13)



(Refer Slide Time: 27:37)

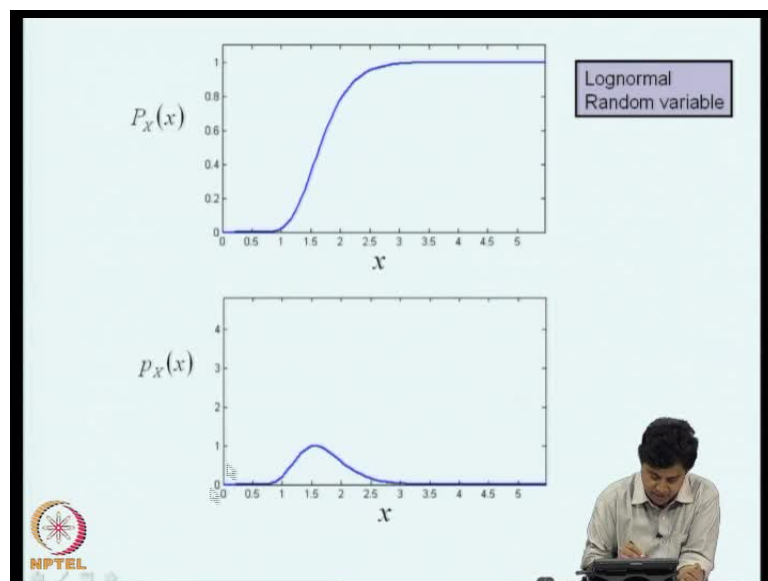
Example 1
 $Y = \exp(X); X \sim N(m, \sigma)$
 $-\infty < x < \infty \Rightarrow 0 < y < \infty$
 $y = \exp(x) \Rightarrow x = \log y \ \& \ \frac{dy}{dx} = \exp(x) = y$

$$p_Y(y) = \frac{p_X(x)}{\exp(x)} \Big|_{x=\log y}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{\log y - m}{\sigma}\right)^2\right], 0 < y < \infty$$

Y is said to be lognormal distribute d.

(Refer Slide Time: 30:13)



Now, we come here, I am plotting y of x versus x here, in this graph. This is my exponential function, so, at x is equal to 0 it is 1 and as x tends to infinity, it goes to infinity and x goes to minus infinity, it goes to 0 now this is my probability density function of normal random variable, this is the mean and parameter is or sigma is also reflected in this diagram, we are interested in knowing the probability that y lies between this shaded region y to y plus dy if y lies between y to y plus dy x will lie in this shaded region.

Therefore, probability density function of y at y is given by p_x of x divided by $\frac{dy}{dx}$ there is exponential x , which is same as y . So, we substitute into the formula here and we get the following this expression for the probability density function of y the state variable y is appearing here, as well as here this random variable y is known as lognormal random variable the name here originates, because if you look at x which is normal is actually $\log y$. so, logarithm of y is normally distributed so, we call it as lognormal random variable.

This is a typical plot of a lognormal random variable, this is the probability density function, although there is at y equal to 0 the y is sitting. Here, in the denominator but $\log y$ is in the exponent and if you apply the **L Hopital's rules**, you can show that as y goes to 0 the probability density function goes to 0. So, we start with 0 and it peaks and this is the probability density function they associated probability distribution function is here it is a 0 at x equal to 0 and as x tends to infinity becomes 1. So, this x here is the lognormal random variable, a typical plot of that.

(Refer Slide Time: 30:59)

Example
 $Y = aX^2$
 $\Rightarrow x = \pm \sqrt{\frac{y}{a}}; \frac{dy}{dx} = 2ax$
 $p_Y(y) = \frac{p_X\left(\sqrt{\frac{y}{a}}\right) + p_X\left(-\sqrt{\frac{y}{a}}\right)}{2\sqrt{ay}}, 0 < y < \infty$

The slide also features two graphs. The top graph shows a normal distribution $p_X(x)$ centered at x_2 with a shaded area between $x_1 + dx_1$ and $x_2 + dx_2$. The bottom graph shows the resulting lognormal distribution $p_Y(y)$ for $y > 0$, with a shaded area between y and $y + dy$. The NPTEL logo is visible in the bottom left corner.

We will consider another transformation Y equal to aX square, so, first step is we have to invert this relation and find out the routes of this equation. So, x is plus minus square root y by a and what is $d y$ by $d x$ it is $2 a x$. So, here we are plotting the function y equal to $a x$ square here and the given probability density function of x is shown here and this is the unknown probability density function of y and if x takes values from minus infinity

to plus infinity y will take value from 0 to infinity. So, the probability that y lies between y to y plus dy is nothing but the union probability of the union of these 2 events namely x lies between x 1 to this x 1 minus dx 1 and x 2 to x 2 plus d x 2. So, you have to add these 2 probabilities, if you follow those rules and if we add them up I get the probability density function of y in terms of probability density function of x, which is given and this is the requisite probability density function.

(Refer Slide Time: 32:20)

Example

$$U = \frac{X-m}{\sigma}; X \sim N(m, \sigma)$$

$$p_U(u) = \frac{p_X(\sigma u + m)}{\frac{1}{\sigma}} = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{u^2}{2}\right); -\infty < u < \infty$$

$\Rightarrow U \sim N(0,1)$

Remarks

- (a) $U \sim N(0,1)$
- (b) U is non-dimensional (why?)
- (c) This is an illustration of a general result that Gaussian quantities are closed under linear operations.

Exercise

Let $X \sim N(m, \sigma)$. Define $Y = aX + b$. Show that $Y \sim N(am + b, a\sigma)$.

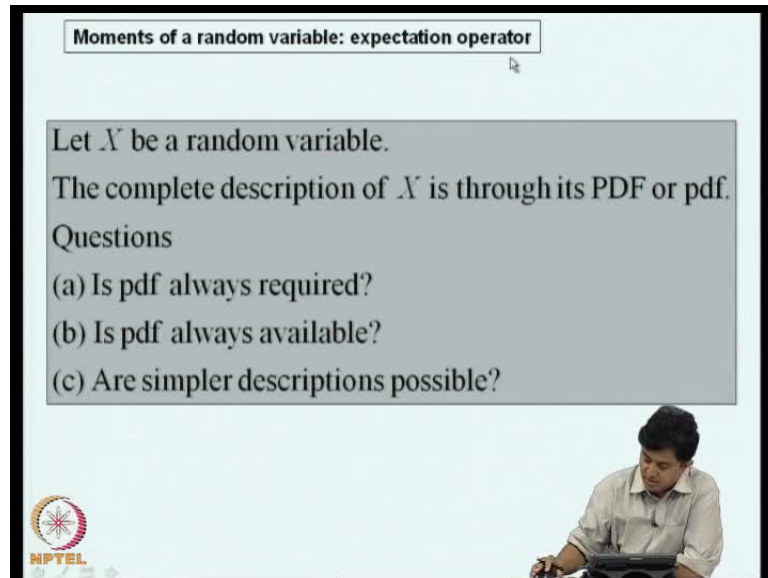
NPTEL

Now, we consider another example, let X be a normal random variable with parameters m and sigma and I define a new random variable u, which is X minus m by sigma. Now, we have to invert this relation x will be sigma u plus m that is the route and what is du by dx it is 1 by sigma. So, the probability density function of u is given by probability density function of x evaluated at the root divided by the modulus of the slope and if we do that we get this function, which is actually another normal random variable, whose parameter is 0 and 1, m is 0 sigma is 1.

So, **we have** what we have done is, we have carried out a linear transformation on a Gaussian random variable and we see that the Gaussian property of the random variable is preserved. U is you can show that U is non-dimensional right now, it may not be clear but as we go along will be able to show that U is non-dimensional and 1 more remark that can be made is that the fact that upon a linear transformation a Gaussian random

variable remains Gaussian, in this particular example, this is an illustration of a general result that Gaussian quantities are closed under Gaussian operations.

(Refer Slide Time: 35:04)



Moments of a random variable: expectation operator

Let X be a random variable.
The complete description of X is through its PDF or pdf.

Questions

- (a) Is pdf always required?
- (b) Is pdf always available?
- (c) Are simpler descriptions possible?

The slide also features the NPTEL logo in the bottom left corner and a small inset image of a person sitting at a desk in the bottom right corner.

Later, in this course, we will see this, this operation, what I mean here can be a differential operation or a matrix operation or another words, if input to a mechanical system or Gaussian and the mechanical system is linear the output will also be Gaussian.

As an exercise, we could consider another problem, where x is again normal with parameter m and σ and we define a linear transformation on x $a x$ plus b - where a and b are deterministic numbers and we can show that y is again normal with mean $a m$ plus b with parameter $a m$ plus b and the second parameter a into σ . This again is an illustration of the fact that a linear transformation of a Gaussian random variable keeps the random variable Gaussian.

Now, we will ask this question, now, let us consider a random variable x , we will ask the question, what constitutes the complete description of this random variable; we have answered this question, for complete description of x we need to specified its probability distribution function or its probability density function.

(Refer Slide Time: 36:13)

Mean of a random variable

$$\eta = \int_{-\infty}^{\infty} x p_X(x) dx \quad (\text{continuous random variable})$$
$$\eta = \sum_{i=1}^n x_i P(X = x_i) \quad (\text{discrete random variable})$$

η is a measure of **central value** of X .

NPTEL

Now, as engineers, we are always interested in simplifying descriptions. So, we can ask the questions is probability density function always required, an associated question is probability density function always available. This leads to the question or simpler descriptions possible for a random variable that takes us to the next topic namely moments of a random variable and that introduces us to what is known as expectation operator.

So, let us start with, what is meant by, mean of a random variable, lets x be a random variable; let it be continuous and $p_X(x)$ be the probability density function by definition the mean of x is given by this integral $\int x p_X(x) dx$, if x is a discrete random variable; where this probability of X equal to x_i is specified for i equal to 1 to n but definition the mean of this random variable x is given by the summation i equal to 1 to n $x_i p_X(x_i)$.

(Refer Slide Time: 37:30)

Is this notion consistent with our intuitive notion of a mean?

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n x_i = \frac{1}{n} \sum_{i=1}^r n_i x_i$$

$$= \sum_{i=1}^r \left(\frac{n_i}{n} \right) x_i$$

$$= \sum_{i=1}^r P(X = x_i) x_i \sim \int_{-\infty}^{\infty} x p_X(x) dx$$

$$= \eta$$

NPTEL

The word mean is we are familiar with, what it means, the question that we can ask is this definition of mean of a random variable consistent, with what we know as mean or the average we claim that eta is a measure of central value of x.

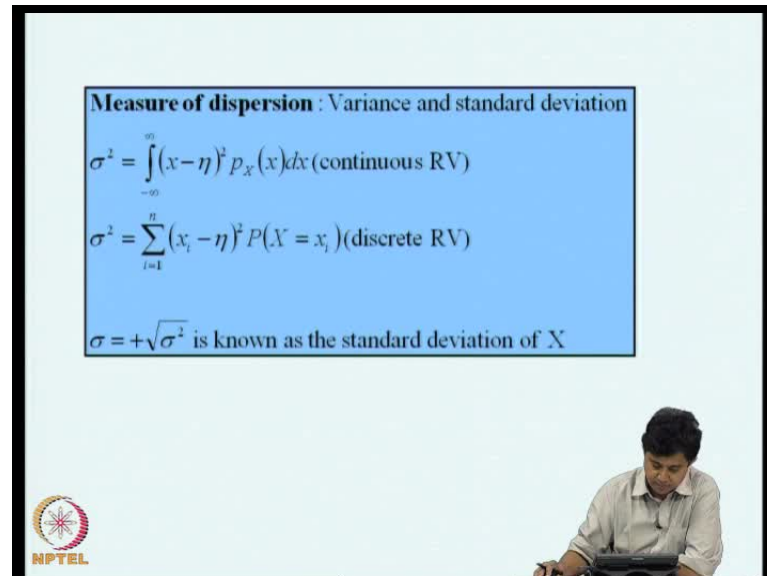
If I am asked to specify a single number, which is most representative of capital x I would be tempted to offer eta is a candidate the question as our saying is this notion consistent with our intuitive notion of a mean by that what i means suppose if you have numbers x i i equal to 1 to x n to x 1 x 2 x 3 x n and I ask the question what is the average of x or mean of x our intuition would say that you have to add all these numbers and divide it by the number of observation that is the mean of x we call it as x bar.

Now, this can be rewritten as i equal to 1 to r n i x i, where n i is number of observations which are identical to x i that means for example, if you are evaluating a class out of 10 marks and you are giving marks 0, 1, 2, 3 up to 10 and if there are 100 students some may be 5 students may get 8 marks, 11 students may get 7 marks, so and so forth. So, this n i is the number of x i 7 minutes n i number of students, we are got 7 marks right so the average is again given by this.

Now, what is n i by n, which is nothing but probability of x equal to x i by 1 of our definition of probability. So, this summation thus can be given as x i into probability of x equal to x i this is perfectly consistent with the definition of mean of a random variable for discrete random variable, which we just now solved if these number of distinct

observations become very large this summation can be replaced by an integral and this again is consistent with the definition of mean of a random variable, this is eta.

(Refer Slide Time: 39:26)



Measure of dispersion : Variance and standard deviation

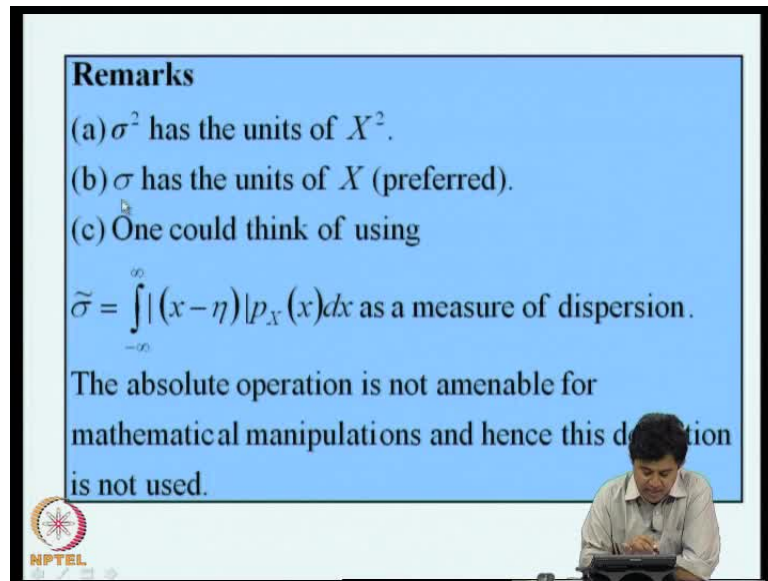
$$\sigma^2 = \int_{-\infty}^{\infty} (x - \eta)^2 p_x(x) dx \text{ (continuous RV)}$$
$$\sigma^2 = \sum_{i=1}^n (x_i - \eta)^2 P(X = x_i) \text{ (discrete RV)}$$

$\sigma = +\sqrt{\sigma^2}$ is known as the standard deviation of X

NPTEL

So, then the definition of mean as a moment of a probability density function is consistent with, what we think as the so called arithmetic mean. After considering as measure of central value, the next natural question to ask is to define a measure of dispersion, this leads us to the notion of variance and standard deviation; again if x is a continuous random variable with probability density function p x of x by definition the variance is defined as integral minus infinity to plus infinity x minus eta, where eta is now, the mean which we already defined square of that p x of x dx for a continuous random variable and an associated definition for a discrete random variable, where this integral is replace by a summation, the positive square root of this variance is known as the standard deviation of capital x, we do not assign any meaning to the negative root this is by convention.

(Refer Slide Time: 40:28)




Remarks

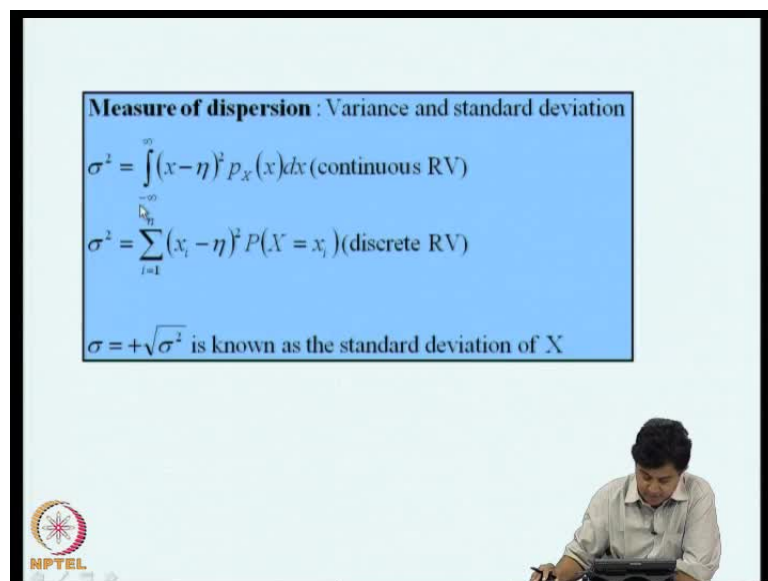
- (a) σ^2 has the units of X^2 .
- (b) σ has the units of X (preferred).
- (c) One could think of using

$$\tilde{\sigma} = \int_{-\infty}^{\infty} |x - \eta| p_X(x) dx \text{ as a measure of dispersion.}$$

The absolute operation is not amenable for mathematical manipulations and hence this definition is not used.




(Refer Slide Time: 40:48)



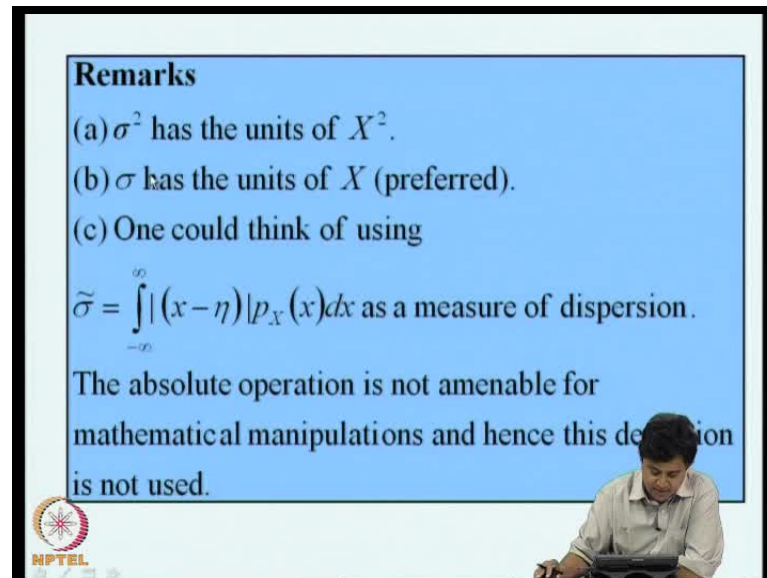
Measure of dispersion : Variance and standard deviation

$$\sigma^2 = \int_{-\infty}^{\infty} (x - \eta)^2 p_X(x) dx \text{ (continuous RV)}$$
$$\sigma^2 = \sum_{i=1}^n (x_i - \eta)^2 P(X = x_i) \text{ (discrete RV)}$$

$\sigma = +\sqrt{\sigma^2}$ is known as the standard deviation of X



(Refer Slide Time: 41:03)



Remarks

(a) σ^2 has the units of X^2 .



(b) σ has the units of X (preferred).

(c) One could think of using

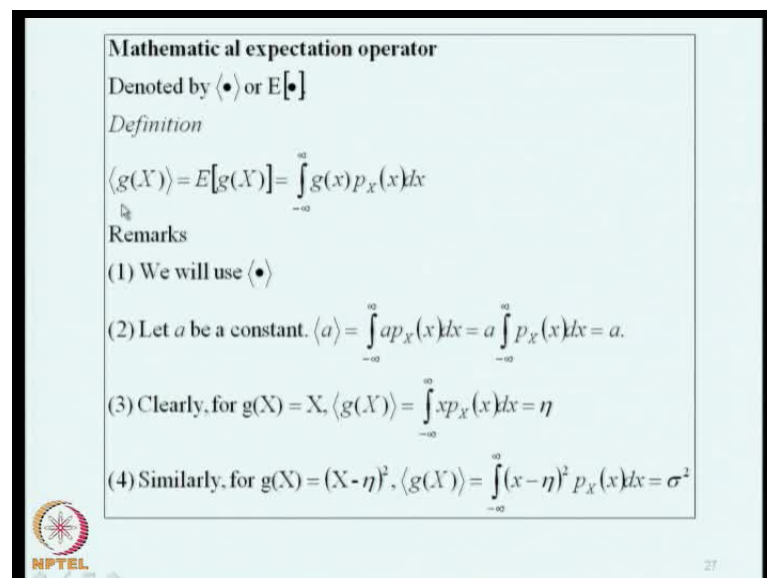
$$\tilde{\sigma} = \int_{-\infty}^{\infty} |(x - \eta)| p_X(x) dx$$

as a measure of dispersion.

The absolute operation is not amenable for mathematical manipulations and hence this definition is not used.



(Refer Slide Time: 41:58)



Mathematical expectation operator

Denoted by $\langle \bullet \rangle$ or $E[\bullet]$

Definition

$$\langle g(X) \rangle = E[g(X)] = \int_{-\infty}^{\infty} g(x) p_X(x) dx$$


Remarks

(1) We will use $\langle \bullet \rangle$

(2) Let a be a constant, $\langle a \rangle = \int_{-\infty}^{\infty} a p_X(x) dx = a \int_{-\infty}^{\infty} p_X(x) dx = a$.

(3) Clearly, for $g(X) = X$, $\langle g(X) \rangle = \int_{-\infty}^{\infty} x p_X(x) dx = \eta$

(4) Similarly, for $g(X) = (X - \eta)^2$, $\langle g(X) \rangle = \int_{-\infty}^{\infty} (x - \eta)^2 p_X(x) dx = \sigma^2$



27

Now, variance of x has the units of square of x that can be verified, whereas standard deviation has units of x ; therefore, it is preferred over variance. What we have done here is to find out how does x deviates from its central value and squared it and summed it over minus infinity to plus infinity. We may be tempted to offer an alternate definition for measure of dispersion, where instead of square squaring and then taking a square root, I can as well consider here the modulus of x minus η and p_X of x d x integral minus infinity to plus infinity as a measure of dispersion this is perfectly valued as a definition, but this is not commonly used because this operation of taking absolute value

is not amenable for algebraic manipulations. Say, a plus b whole square can be expanded as a square plus b square plus 2 a b, whereas the modulus of a plus b state as modulus of a plus b. So, this is 1 reason, why this is not preferred, although this also and looks perfectly acceptable, as a definition. The generalization of the notion of mean and variance can be **of can be** seen by defining what is known as expectation operator.

We consider x to be a random variable with given probability density function p x of x and consider the function g of X, we say this quantity as expected value of g of X, one can use 2 notations to write this 1 is to use this angular brackets or used this brackets and put a e outside, this e is expectation operator. This is to be read as expected value of g of X, this also has be read as expected value of g of X, this by definition is this integral g of X p x of x dx from minus infinity to plus infinity. In our lecture, I would prefer to use this notation, this show this angular brackets. This is the expectation operator that I will follow. Now, let us consider, what is **the...**, a b a constant, what is the expected value of a constant.

What it is this is integral minus infinity to plus infinity a p x of x dx. a is a constant. So, we can take that outside and we are left with a into this minus infinity to plus infinity p x of x dx, what is this minus infinity to plus infinity p x of x d x, it is the probability of the sample space, which is 1. Therefore, this is a; so expected value of a is a.

(Refer Slide Time: 44:20)

Remarks (continued)

(5) Define $m_n = \langle X^n \rangle = \int_{-\infty}^{\infty} x^n p_X(x) dx$; n-th order raw moments.

(6) Define $\mu_n = \langle (X - \eta)^n \rangle = \int_{-\infty}^{\infty} (x - \eta)^n p_X(x) dx$; n-th order central moment.

(7) COV = coefficient of variation = $\frac{\sigma}{\eta}$; $\eta \neq 0$.

$$\eta = \int_{-\infty}^{\infty} (x - \eta) p_X(x) dx$$

$$= \int_{-\infty}^{\infty} x p_X(x) dx - \eta \int_{-\infty}^{\infty} p_X(x) dx$$

$$\eta - \eta = 0$$

(8) $\mu_0 = 1, \mu_1 = 0, \mu_2 = \sigma^2$

(9) skewness = $\frac{\mu_3}{\sigma^3}$; kurtosis = $\frac{\mu_4}{\sigma^4}$; (non-dimensional entities)

(10) $m_2 = \langle X^2 \rangle$ = mean square value, $\sqrt{m_2}$ = root mean square value.


(11) Root mean square value = standard deviation if mean = 0.

(12) $\langle (X - \eta)^2 \rangle = \langle X^2 + \eta^2 - 2X\eta \rangle = \langle X^2 \rangle + \eta^2 - 2\eta \langle X \rangle = \langle X^2 \rangle - \eta^2$

(13) $\sigma^2 + \eta^2 = \langle X^2 \rangle$. Clearly, $\sigma^2 > 0 \Rightarrow \eta^2 < \langle X^2 \rangle$. That is, $m_2 > m_1^2$.

Moments are not arbitrary. They need to satisfy certain inequalities.

(14) Two random variables may have the same mean and variance. This does not imply that they will have the same pdf.



Now, if g of X is x , what is the expected value of g of X ? It is minus infinity to plus infinity $x p x$ of $x dx$, which is nothing but the mean of the random variable x similarly, if I now define g of X as x minus η whole square the expected value of g of X in this case will be the variance of x . So, to define mean and variance, we can use a expectation operator, which is more general, the mean and variance can be obtained as special cases of the general expectation of g of X .

Now, we can give further meanings to g of X , we can say g of X is x to the power of n . We can call them as n th order raw moments, it is given by x to the power of $n p x$ of $x dx$, we can similarly define n th order central moments, where I consider g of X to be x minus η to the power of n . So, this is integral minus infinity to plus infinity x minus η to the power of $n p x$ of $x dx$. This we call it as n th order central moments, we define the ratio of the standard deviation to the mean as the coefficient of variation.

Now, let us consider, what happens in this definition n equal to 0. So, n equal to 0 x minus η to the power of 0 is 1 expected value of 1 is 1. Therefore, μ naught is 1 what is n equal to 1. This is expected value of x minus η , this can be we can consider this specific case of n equal to 1 in some detail. So, this is minus infinity to infinity x minus $\eta p x$ of $x dx$, this is nothing but minus infinity to infinity $x p x$ of $x dx$ minus η minus infinity to infinity $p x$ of $x dx$. This is nothing but η and this is nothing but η . So, this η minus η equal to 0.

So, the first order central moment is 0. The second order central moment is nothing but the variance so on and so forth. Now, we define the ratio of third order moment to the cube of the standard deviation as skewness, similarly ratio of μ_4 by σ to the power of 4 as kurtosis. These are non-dimensional entities coefficient of variance skewness and kurtosis are non-dimensional.

Now, let us come to now this moment m_n m_1 is expected value of x , which is η m_2 is expected value of X square, this and we call this as mean square value in the square root of this is known as root means square value. These are some of the terms that you come across, let us consider a situation, where mean is 0 in which case you can see here m_n will be equal to μ_n because η is 0.

So, in that case m_2 will be nothing but variance and therefore, the root means square in that case is nothing but standard deviation, if means is 0, if mean is not 0, the standard

deviation and root means square values are different the r m s value and standard deviation are synonymous, if and only if mean is 0.

Now, we can consider this second order central moment and if I now, expand x is square plus η square minus $2x\eta$, I can write this as expected value of x is square. This is a constant. Therefore, expected value of constant is η square and expected value of $2x\eta$ is 2η expected value of x and we see that this becomes the variance is nothing but means square value minus square of the mean that means $\sigma^2 + \eta^2$ is mean square value.

Since σ^2 is greater than 0, it implies that the means square value has to be greater than η^2 that is, the second moment m_2 must be greater than or equal to the square of the first moment, or in other words the moments of a random variables or not arbitrary, you cannot arbitrarily assign moments to random variable they have to satisfy certain in equalities.

(Refer Slide Time: 49:13)


Moment generating function

Definition : $\psi_X(s) = \langle \exp(sX) \rangle = \int_{-\infty}^{\infty} \exp(sx) p_X(x) dx$

That is $\psi_X(s)$ is the Laplace transform of the pdf.

$$\exp(sX) = 1 + sX + \frac{(sX)^2}{2!} + \frac{(sX)^3}{3!} + \dots$$

$$\psi_X(s) = 1 + s\langle X \rangle + \frac{s^2}{2!}\langle X^2 \rangle + \frac{s^3}{3!}\langle X^3 \rangle + \dots$$

$$\Rightarrow \left. \frac{d\psi_X}{ds} \right|_{s=0} = \langle X \rangle; \left. \frac{d^2\psi_X}{ds^2} \right|_{s=0} = \langle X^2 \rangle; \dots \left. \frac{d^n\psi_X}{ds^n} \right|_{s=0} = \langle X^n \rangle$$


Now, finally, we can remark that we can have 2 random variables, which may have the same mean and variance, but this does not means that the 2 random variables will have the same probability density function, because there other moments can be different. In discussion of moments, we talk about what is known as a moment generating function, that is defined as expected value of a quantity exponential of $s x$.

So, that is given by minus infinity to plus infinity exponential of $s x$ p of $x dx$, you would see that this is nothing but the Laplace transform of the probability density function. So, we define this function as a moment generating function, how does this word originate to see that we can expand exponential $s x$ in a series $1 + s x + \frac{s^2 x^2}{2!} + \dots$. Now, the moment generating function is an expected value of this, so if you carry out the expectation operation on this, the left hand side expected value of the left hand side is the moment generating function and right hand side is $1 + s$ into mean of x plus square by 2 factorial expected value of x is square so and so forth.

(Refer Slide Time: 51:08)

Characteristic function

Definition : $\varphi_X(\omega) = \langle \exp(i\omega X) \rangle = \int_{-\infty}^{\infty} \exp(i\omega x) p_X(x) dx$

Here ω is real valued. Thus, $\varphi_X(\omega)$ is the Fourier transform of the pdf.

It can be shown that $\left. \frac{1}{i^n} \frac{d^n \varphi_X}{d\omega^n} \right|_{\omega=0} = \langle X^n \rangle$.

By using inverse Fourier transform, it follows that

$$p_X(x) = \int_{-\infty}^{\infty} \varphi_X(\omega) \exp(-i\omega x) dx.$$

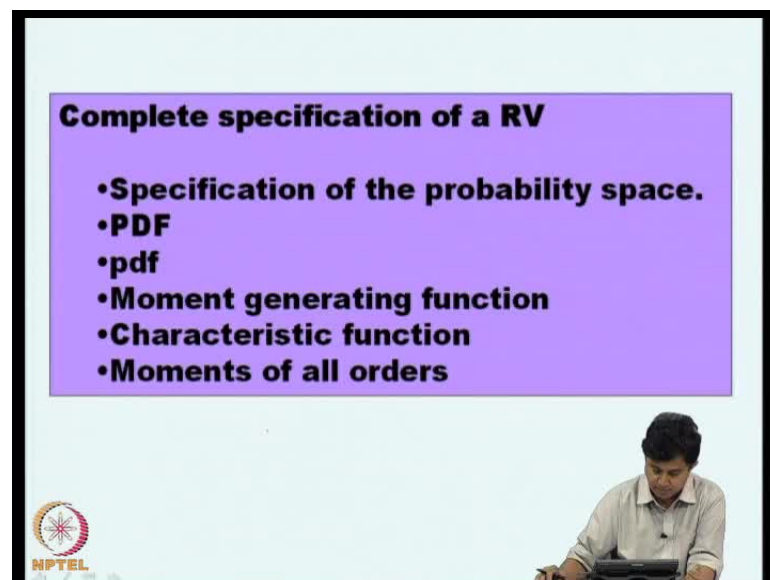
The slide also features the NPTEL logo in the bottom left and a lecturer in the bottom right.

Now, if you differentiate this, now with respect to x and put s is equal to 0, what happens this this is 0 this becomes expected value of x this plus 2 x in to this. This 3 x square into this so and so forth, when I put s equal to 0 and all other higher order terms go to 0. Therefore the mean value can be obtained as $d \psi x$ by $d s$ at s equal to 0, similarly the second derivative at s equal to 0 is the second raw moment, m_2 in general $d^n \psi x$ by $d s^n$ is x to the power of n . So, because of this property, we call this expected value as a moment generating function, here s is a general complex number, if we make s to be a pure imaginary number, we get what is known as a characteristic function. So, we define the expected value of $i \omega x$, where ω is real and i square is minus 1. This is exponential $i \omega s p x$ of $x dx$.

You will recognize that $\phi_X(\omega)$ is the Fourier transform of the probability density function, by expanding $e^{i\omega x}$ in a Taylor series and differentiating that with respect to ω one can show that, the n th order moment is given by the derivative of the characteristic function in this way at $\omega = 0$.

Now, by using inverse Fourier transform, if I know the characteristic function, I can evaluate the probability density function also that would mean a random variable, now can be expressed in terms of its characteristic function also and if you have the characteristic function for a random variable it also constitute the complete description of a random variable.

(Refer Slide Time: 52:25)

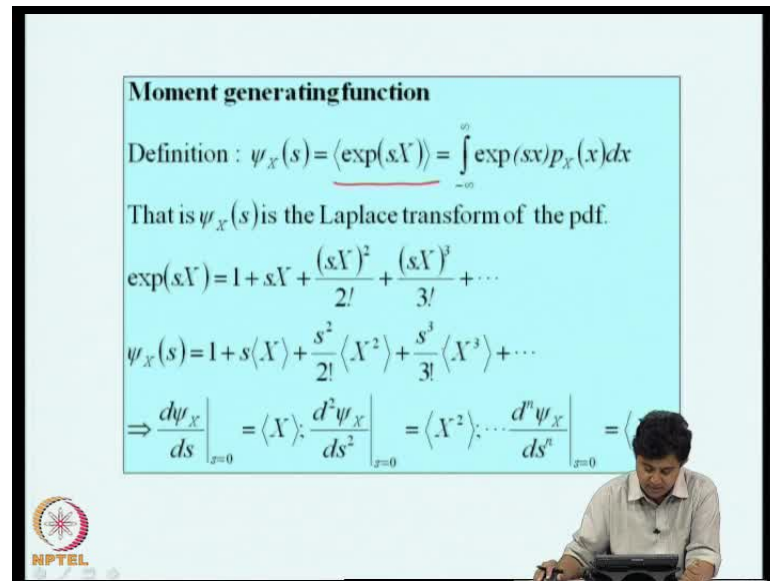


Complete specification of a RV

- **Specification of the probability space.**
- **PDF**
- **pdf**
- **Moment generating function**
- **Characteristic function**
- **Moments of all orders**

NPTEL

(Refer Slide Time: 52:53)




Moment generating function

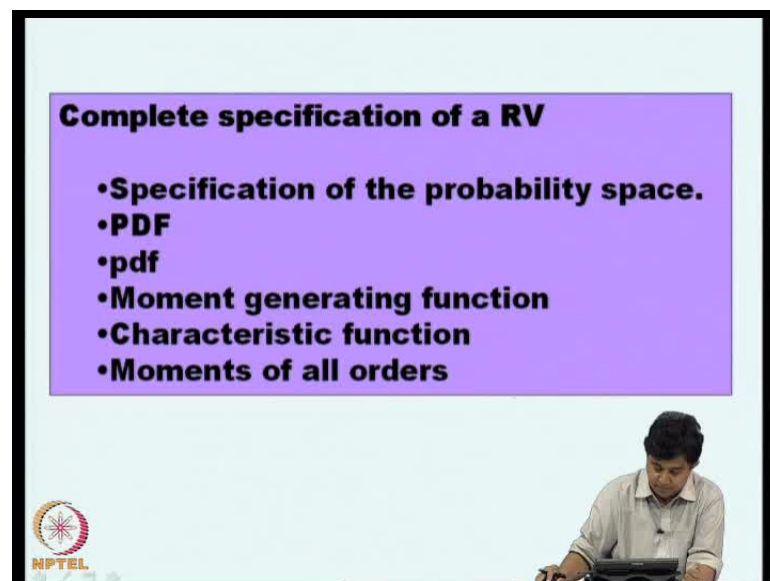
Definition : $\psi_X(s) = \langle \exp(sY) \rangle = \int_{-\infty}^{\infty} \exp(sx) p_X(x) dx$

That is $\psi_X(s)$ is the Laplace transform of the pdf.

$$\exp(sY) = 1 + sY + \frac{(sY)^2}{2!} + \frac{(sY)^3}{3!} + \dots$$
$$\psi_X(s) = 1 + s\langle X \rangle + \frac{s^2}{2!}\langle X^2 \rangle + \frac{s^3}{3!}\langle X^3 \rangle + \dots$$
$$\Rightarrow \left. \frac{d\psi_X}{ds} \right|_{s=0} = \langle X \rangle; \left. \frac{d^2\psi_X}{ds^2} \right|_{s=0} = \langle X^2 \rangle; \dots \left. \frac{d^n\psi_X}{ds^n} \right|_{s=0} = \langle X^n \rangle$$




(Refer Slide Time: 53:09)



Complete specification of a RV

- **Specification of the probability space.**
- **PDF**
- **pdf**
- **Moment generating function**
- **Characteristic function**
- **Moments of all orders**



So, complete specification of random variable is through its probabilities space itself, omega b p or through the probability distribution function or probability density function or the moment generating function or the characteristic function. Alternately, we can specify m n for all orders, if I know m 1, m 2, m 3, m n what I could do is I could go back and construct, the moment generating function using this relation, using this relation and then, I can use universe Laplace or universe Fourier transform and get the probability density function. So, introduction of moment generating function and

characteristic function and moments of a random variables expands our tool kit to model uncertainties.

So, in manipulating mathematical models many times it would be easier to deal with the characteristic function easier to deal with moments not always we need to deal with probability density and distribution functions.

(Refer Slide Time: 53:34)

Bernoulli random variable

$$P(X=0) = p, P(X=1) = 1-p$$

$$\langle X \rangle = \sum_{i=1}^2 x_i P(X=x_i) = 0 \times p + 1 \times (1-p) = 1-p \quad m_1$$

$$\langle X^2 \rangle = \sum_{i=1}^2 x_i^2 P(X=x_i) = 0 \times p + 1 \times (1-p) = 1-p \quad m_2$$

$$\sigma_X^2 = \langle X^2 \rangle - \eta^2 = (1-p) - (1-p)^2 = p(1-p)$$

$$\phi_X(\omega) = \exp(i\omega 0)p + \exp(i\omega)(1-p) = p + (1-p)\exp(i\omega)$$

$$\langle X^n \rangle = 1-p //$$

NPTEL 32

Now, we will consider 1 example, for a Bernoulli random variable let us find out what it is mean and variance. So, for a Bernoulli random variable there are two possible outcomes, x is equal to 0 and x equal to 1 the probability of x is equal to 0 is P and X equal to 1 is 1 minus p.

So, the expected value of x is x i some somewhere i equal to 1 to 2 probability of x equal x i so x 1 is 0 0 into probability of getting 0 is p plus x 2 is 1 and the associated probability 1 minus p. Therefore, mean of x is 1 minus p. similarly the means square value is x i square probability of x equal to x i running from 1 to 2. So, here again 0 square into p plus 1 square into 1 minus p which is 1 minus p ,

Now, what is variance it is mean square value minus square of the mean. So, it is 1 minus p 1 minus p whole square and I get p into 1 minus p. So, square root of this is the standard deviation. The characteristic function for x form a expected value of e exponential of i omega x, so e x p i omega of 0 into p plus e x p into i omega into 1 into

$1 - p$, this is $p + 1 - p$ exponential $i\omega$. And from this, I can show that all m, n is $1 - p$ see you have noticed now, that this is nothing but $m = 1$ which is $1 - p$ this is $1 - p$ and this is $m = 2$, this is $1 - p$, this is $m = 2$ which again $1 - p$ using this characteristic function, you can show that all m, n are $1 - p$ for this random. So, we will close this lecture at this point, we will continue with this in the next lecture.