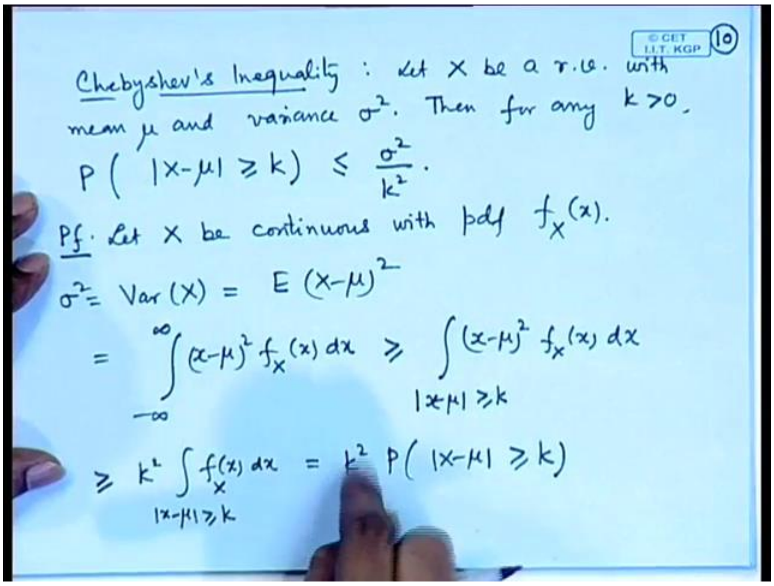
Probability and Statistics
Prof. Somesh Kumar
Department of Mathematics
Indian Institute of Technology, Kharagpur

Lecture – 18
Characteristics of Distributions – II

We may not have substantial knowledge about probabilities of various intervals or random variable taking value less than something or greater than something. So in such cases, we have certain probability in equalities which are useful if we know only a certain moments say mean or variance or one particular moment.

(Refer Slide Time: 00:50)

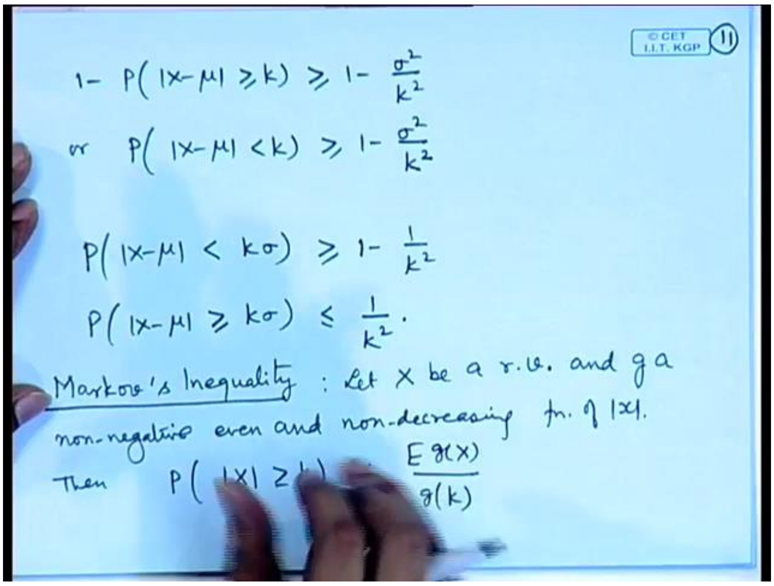


So, these are known as one of the first one in this direction is called Chebyshev's inequality. Let x be a random variable with mean mu and variance sigma square, then for any k positive probability of modulus x minus mu greater than or equal to k is less than or equal to sigma square by k square. So, you can see this denotes the probability of x lying in certain interval or lying outside a certain interval. We do not have full information about the random variable except its mean and variance; nevertheless we are able to give certain bound for this probability.

To prove this let us consider X to be continuous with certain pdf; f x. So, let us consider the expression for variance; this is equal to expectation of x minus mu square which is equal to integral x minus mu square, f x; d x. Now this particular integral is greater than or equal to modulus x minus mu greater than or equal to k. This is because the integrand is non-negative. So if you reduce the region of integration, the value will become a smaller.

Now, on this region x minus mu whole square is greater than or equal to k square. So, we can replace by that this is nothing, but probability of modulus x minus mu greater than or equal to k. So, as a consequence probability of modulus x minus mu greater than or equal to k is less than or equal to sigma square by k square. You can write down the alternative forms of this inequality by taking the complementary event here.

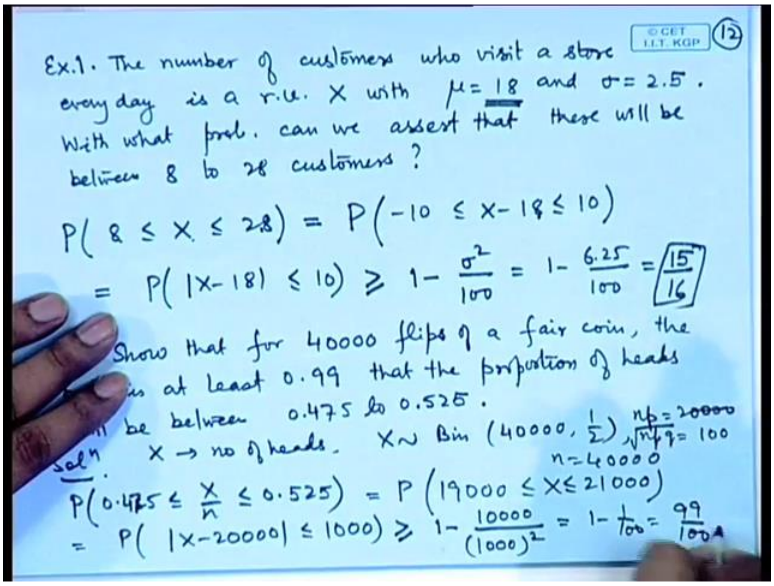
(Refer Slide Time: 03:31)



So, you will have 1 minus probability of modulus x minus mu greater than or equal to k, greater than or equal to 1 minus sigma square by k square or you can write probability of modulus x minus mu less than k is greater than or equal to 1 minus sigma square by k square. Sometimes the form is written in this fashion probability of modulus X minus mu less than k sigma is greater than or equal to 1 minus 1 by k square or probability of modulus x minus mu greater than or equal to k sigma is less than or equal to 1 by k square.

A more general inequality of the same type is called Markov's inequality: let X be a random variable and g a non -negative even and non-decreasing function of modulus X, then probability of modulus x greater than or equal to k is less than or equal to expectation of g X divided by g of k. You can see that if I replace X by X minus mu and take g as the squared function; that is x square then this Markov inequality gives the Chebyshev's inequality.

(Refer Slide Time: 05:46)



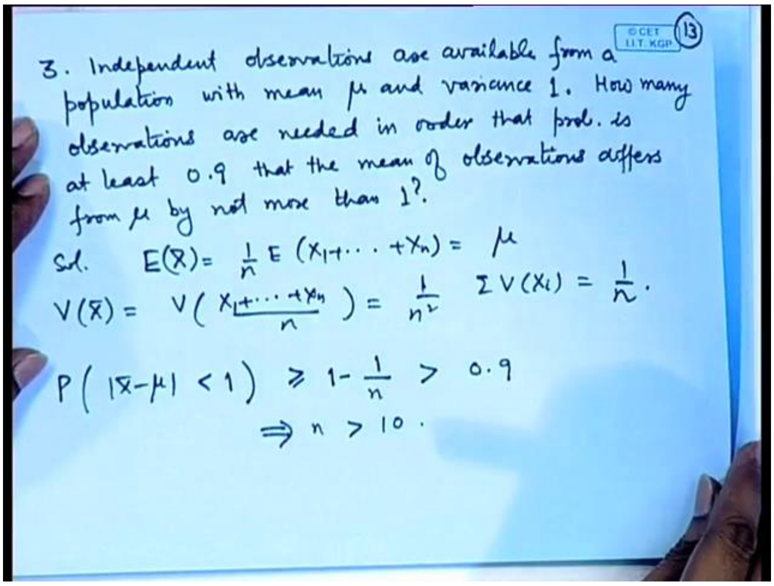
In this one if you replace x by x minus mu and g X is equal to X square then we get exactly this one, so this is a more general inequality of the same type. Let us take some example to explain this; the number of customers who visit a store everyday is a random variable X with mean say 18 and standard deviation is equal to 2.5. With what probability can we assert that there will be between 8 to 28 customers? That means we are interested in an estimate of the probability that the number of customers is between 8 to 28.

If we want to utilize the Chebyshev's inequality here the mean is given to be 18, so this becomes probability of x minus 18 lying between minus 10 to 10. That means, probability of modulus x minus 18 less than or equal to 10. So, by Chebyshev's inequality it is greater than or equal to 1 minus sigma square by 100 that is equal to 1 minus 6.25 by 100 or 15 by 16. You can see that this is very high probability for this particular event to be true. So, although here we do not have full information about the probability distribution of the random variable, but we can tell about certain probability. Show that for 40000 flips of a fair coin, the probability is at least 0.99 that the proportion of heads will be between 0.475 to 0.525.

So, here if we consider X to be the number of heads then X follows binomial 40000 and half. We are interested in probability of x by n lying between 0.475 to 0.525. Now n is here 40000, so this is equal to probability that x lies between 19000 to 21000. Now here mean of this distribution here is np that is equal to 20000 and the standard deviation is square root npq which is 100. So, this probability is then modulus of x minus 20000 less than or equal to 1000 which is greater than or equal to 1 minus sigma square by k square which is 1 minus 1 by 100 that is equal to 99 by 100.

So if it is a fair coin, the probability that the proportion of heads is between 0.475 that is 47.5 percent to 52.5 percent heads are there, in 40000 tosses of a fair coin, the probability is at least 0.99. Here we can even get an appropriate exact value of this, but that is too complicated. So, it is a simple solution for a complex looking situation.

(Refer Slide Time: 10:53)

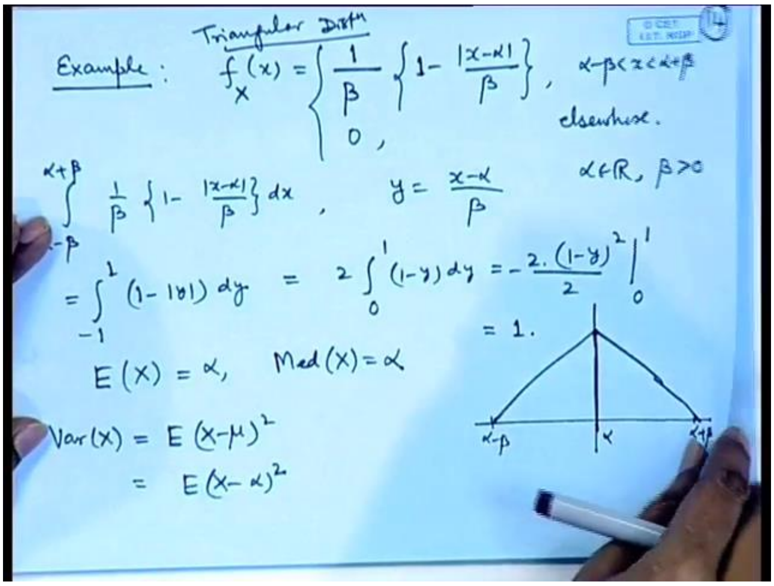


Independent observations are available from a population with mean mu and variance 1. How many observations are needed in order that probability is at least 0.9 that the mean of observations differs from mu by not more than 1?

So, if we are looking at expectation of x bar that is equal to 1 by n expectation of X 1 plus X 2 plus X n that is equal to mu, if we are looking at variance of X bar that is variance of X 1 plus X 2 plus X n by n then it is equal to 1 by n square sigma variance of X I; each of this is 1, so it becomes 1 by n. So, probability of modulus x bar minus mu should be less than 1; that is the mean of observations differs from mu by not more than

1 that this probability of modulus x bar minus mu less than 1. So by Chebyshev's inequality, it is greater than or equal to 1 minus 1 by n. Now we want it to be more than 0.9, so this means n should be greater than 10. So, we need at least 10 observations that probability that mean of observations differ from mu by not more than 1.

(Refer Slide Time: 13:39)



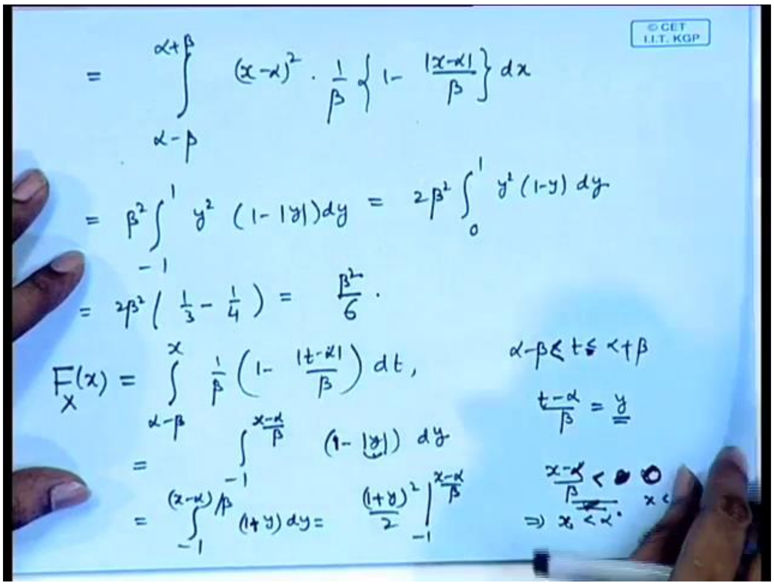
Let us look at some examples of calculation of certain distributions and the moments and other characteristics that we have discussed so far. So, let us consider one example; let x be a continuous random variable with probability density function given by 1 by beta; 1 minus, we will analyze various properties of this distribution. So, if we look at say what are the values of alpha and beta for which this is a valid probability distribution, let us consider. So, first thing we observed that in order that this is a non negative function beta should be positive and 1 minus modulus of x minus alpha by beta. So, here if you look at modulus of x minus alpha by beta is less than 1. Therefore, this quantity is always positive therefore, beta has to be positive in order that this is a density.

Now, we look at the integral of the density over the region, in order to resolve this in a simple way we can consider the transformation y is equal to x minus alpha by beta. Now this is a symmetric function, so it becomes integral 0 to 1; 1 minus y; dy and it is equal to twice 1 minus y whole square by 2 from 0 to 1 and this is simply equal to 1. Therefore, there is no restriction on the range of alpha; alpha can be any real number and beta should be positive real number in order that this is a valid probability density function.

If we want to look at the shape of this distribution in fact, we can see from here for y positive this is 1 minus y and for y negative, it becomes 1 plus y. So, the value at y is equal to plus 1 and minus 1 is 0 and at y is equal to 0, it is 1. So, if we consider this point as alpha, this as alpha minus beta, this as alpha plus beta then the shape of the distribution is triangular, so this is basically a triangular distribution.

Therefore, easily you can see that mean and median of this distribution must be alpha; expectation of x must be alpha, the median of x must be alpha. Therefore, we can consider higher order central movements let us consider say variance, so variance of x is equal to expectation of x minus mu square that is equal to expectation of x minus alpha square.

(Refer Slide Time: 17:40)



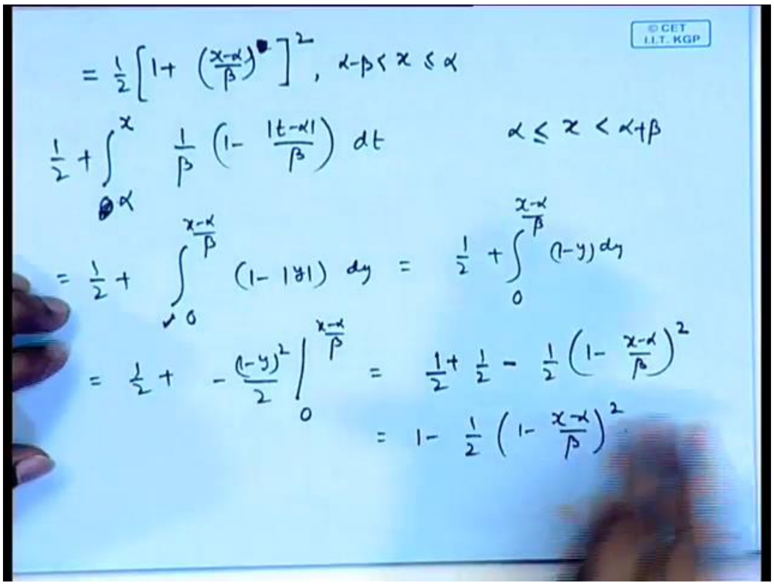
So, this is equal to integral alpha minus beta to alpha plus beta; x minus alpha square 1 by beta, 1 minus modulus x minus alpha by beta d x. So, in order to evaluate this we can consider the same transformation y is equal to x minus alpha by beta. So, after substitution; this turns out to be integral from minus 1 to 1; beta square y square into 1 minus modulus y; d y; as this is a even function, this becomes 2 beta square integral 0 to 1; y square into 1 minus y; d y. So, the integral of this is equal to twice beta square 1 by 3 minus 1 by 4, so that is equal to after simplification beta square by 6, so the variance of this distribution is beta square by 6.

You can see here that if beta is a small value then the variability little high, is little low and if beta is a bigger value then the variability will be more; which is obvious here because this is concentrated from alpha minus beta to alpha plus beta. So if beta is large; this curve will increase further; that means, variability is increasing, if beta is becoming smaller then the variability is becoming less.

We may also look at say quartiles of order 1 by 4 and 3 by 4. So if we consider the point; we can also consider the cumulative distribution function of this. So, if we calculate F X naturally for x less than alpha minus beta it should be 0 and for x greater than alpha plus beta this should be 1. So, we need to concentrate on the integral alpha minus beta to say x 1 by beta 1 minus modulus t minus alpha by beta d t, for t lying between alpha minus beta to alpha plus beta. So, by considering the transformation this becomes minus 1 to x minus alpha by beta, 1 minus modulus y; d y that is by considering the transformation t minus alpha by beta is equal to y.

Now here there are two cases; if x minus alpha by beta is less than 0, in that case is less than alpha; that means, x is less than alpha. So, since we are looking at this 1 less than 0 means x is less than alpha. So, we should consider x minus alpha by beta less than 0 which is basically x is less than alpha; that means, it is before the point of the symmetry. If that is so then this is equal to integral minus 1 to x minus alpha by beta, 1 minus y; dy which is equal to 1 plus y; dy, which is simply equal to 1 plus y square by 2 minus 1 to x minus alpha by beta.

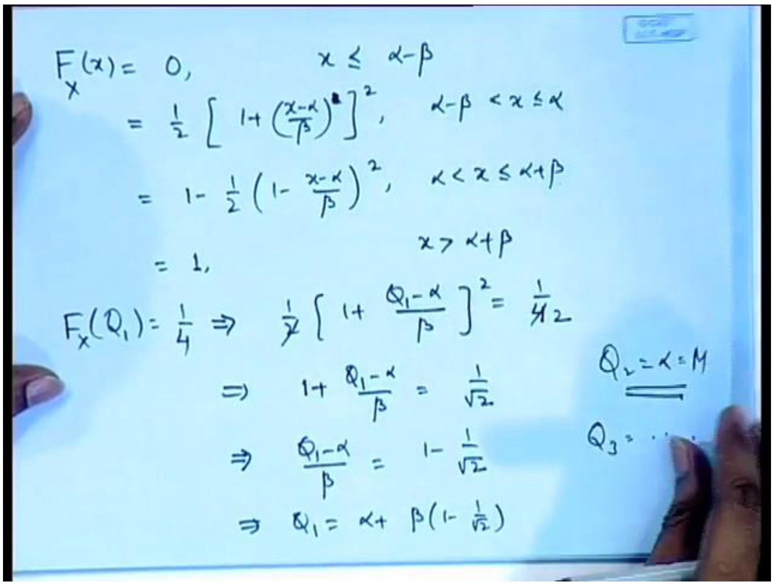
(Refer Slide Time: 21:29)



So, this is evaluated to 1 plus x minus alpha by beta square 1 by 2; 1 plus x minus alpha by beta whole square by 2. So, this is for alpha minus beta less than x less than or equal to alpha. If you look at this value at x is equal to alpha; this is becoming 0, so this value will become exactly equal to half. So, if I am choosing a point x to be greater than or equal to alpha, but less than alpha plus beta, then this will be integral half plus 0 to sorry alpha plus 1 by beta 1 minus t minus alpha by beta dt; that is equal to half plus.

So, this becomes 0 x minus alpha by beta, 1 minus modulus of y dy, that is equal to half plus integral 0 to x minus alpha by beta 1 minus y dy, that is equal to half plus 1 minus y whole square by 2 with a minus sign integral from 0 to x minus alpha by beta; that is equal to 1 minus; so it is half and at 0 this value is becoming half and 1 minus x minus alpha by beta whole square, that is equal to 1 minus 1 by 2, 1 minus x minus alpha by beta square.

(Refer Slide Time: 23:38)



Therefore, we can write the complete description of the CDF as it is 0 for x less than or equal to alpha minus beta, it is equal to half 1 plus x minus alpha by beta whole square alpha minus beta less than x less than or equal to alpha; it is equal to 1 minus half 1 minus x minus alpha by beta square. As a check we can see that the value at the end points of each intervals match because the function is continuous in fact, the function is absolutely continuous. So, if you look at the value at x equal to alpha minus beta here, this is becoming alpha minus beta. So, minus 1 will come here, so 1 minus 1 becomes 0 which is same as the value at x equal to alpha minus beta from the left hand side.

If we look at the value at x equal to alpha here, this value is becoming 0. So, you are getting half here if we put x equal to alpha here this is half so 1 minus half is half. So, the values are matching, if you look at the value at x equal to alpha plus beta of this equation then this value is 0; that means this is equal to 1 and the value here at x equal to alpha plus beta is also 1. So this is satisfying the conditions for the CDF, if we look at the point where the value becomes say 1 by 4 then we need to look at this 1 because the probability at x equal to alpha is equal to half so 1 by 4 will be naturally in this interval alone.

So this means that half; 1 plus Q 1 minus alpha by beta whole square is equal to 1 by 4. So, we can do the simplification here 1 plus Q 1 minus alpha by beta is equal to 1 by root 2; that means, Q 1 minus alpha by beta is equal to 1 minus 1 by root 2. So, Q 1 becomes alpha plus beta into 1 minus 1 by root 2. In a similar way, we can calculate Q 3 also Q 2 is of course alpha that is the median of this distribution.

That is all in today's lecture, we will be considering a special discrete and continuous distributions in the upcoming classes.