So as I was mentioning we had I was mentioning about estimating parameters you can have if you know the form of the density function but if you do not know the values of the parameters then how do you estimate them well there are very many distributions in the world and here I will give you example of only one distribution that is normal distribution normal distribution are Gaussian distribution I will tell you ways of estimating the parameters then the next question is what is the meaning of estimation.

What is the meaning of estimation there are several estimates several types of estimates possible there is something called unbiased estimate there is something called maximum likelihood estimation there is something called consistent estimate like this depending on the uses are depending on the type of the data I mean one can have several types of estimates since ,I am only going to talk about normal distribution not about any other distribution and if you are got finitely many points from a normal distribution be it uni-variate or multivariate normal distribution.

Then the mean of these finitely many points it is unbiased estimate consistent estimate and also a maximum likelihood estimate of the population mean I will repeat it and I will write on the board if you have finitely many poor from a normal distribution the normal distribution can be uni-variate or uni-variate then the mean of these finitely many points is unbiased estimate consistent estimate and maximum likelihood estimate of the population mean.
(Refer Slide Time: 02:45)
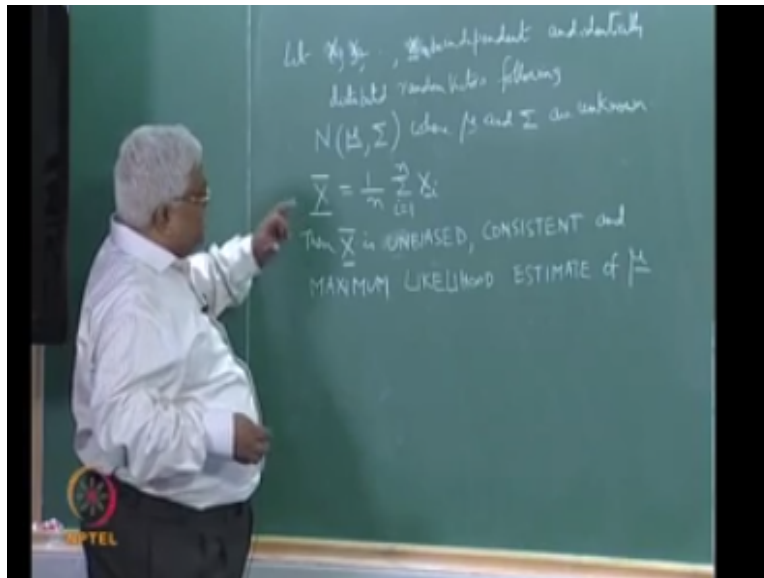
These are endpoints coming from Gaussian distribution independent and identically distributed this is one terminology you are going to find at many places independent and identically distributed do you understand the meaning of this let me tell you the meaning the meaning of one value being our one variable being independent of the another variable you take a coin you toss it once you will get head the same coin you toss it second time you will get some other outcome do you think the first outcome has any role to play.

On the second outcome the answer is no so the result of the first trial does not have any role to play on the result of the second trial then these two trials are independent whatever may be the outcomes of these two trials they are independent now the meaning of identically distributed you are tossing the same coin the probability of head of the coin has not changed the probability of tail of the coin has not changed so the distribution has not changed so in the first trial and second trial you have the same distribution.

But they are independent so then we say that these are independent and identically distributed random variables and why the word variable is used because you are going to get the value as H or T if you are going to get a vector having some number of components then we would call it independent and identically distributed random vectors independent and identically distributed random vectors so let now let me write the statistical way of writing it Let X 1 X 2 X n be independent and identically distributed random vectors following normal distribution with menu and variance covariance matrix Σ where μ and Σ are unknown.

You might be having a question the question is I wrote here random vectors usually we get the observations usually we get for one data set you let us just say you have got 10 vectors usually we get let us just say each one of them is 4dimensional vectors so you will get 10such 4 dimensional vectors ok but then here it is written random vectors you see those vectors that you have got they are taken to be realizations of these 10random vectors if you have got 10vectors in your data set then the small l value is 10 and those vectors are taken to be realizations of these 10vectors.

I will tell you once again with an example say I have tossed a coin my first toss is resulting in head say second toss is also resulting in head this is one realization of that means my b X 1 value is H here X 2 value is also H but note that my X 1 value could have been P my X 2 value could have been T so this X 1 X 2 pair I might have got h, h HT t HT t any one of the four pairs there possible but then in that instant I have got X 1 as a X 2's also H so they are realizations of the vectors but then all possible values of X 1 they can be many all possible values of X 2 can be many.

I hope you have understood the subtle difference I hope you have understood the subtle difference between what a variable is and the data points that you have got so independent and identically distributed random vectors following this but μ and Σ are unknown but we have these are following this so we have some realization of this then how do you estimate this mean μ and variance covariance matrix Σ on the basis of this then so the answer is you write X bar this bar is denoting.

The mean this bar is denoting the vector one over n summation I = 1 tonic you take the mean of these n vectors so in one realization you will get one mean in another realization of these n vectors you might get another mean are you understanding that is why this is also a variable in one realization you may get one mean in another realization of these n vectors you may get another mean so this is also a variable then this is the statement that I made unbiased you and consistent and maximum likelihood estimate of μ positions these things.

I am getting but Evan Oliveira and price there are n trial interests small n files so in one set of n trials you will get one by one set of values in another set of n trials you will get another setoff n values so that is why this is a variable this is a variable then you might then your natural question is what is the meaning of unbiased estimate what is the meaning of consistent estimate and what

is the meaning of maximum likelihood estimate I think that I will not go into it I will not go into then I have to go deep into statistics which is not my aim okay.

I just want to state the results so about mean there is actually no confusion in the sense that you just take the sample mean and observations you just take the mean and that is a good estimate of population mean whatever maybe the mean but about variance you have small problem which was what I was in my one of the previous classes I was mentioning okay so this is about mean now about the covariance matrix.

(Refer Slide Time: 12:02)
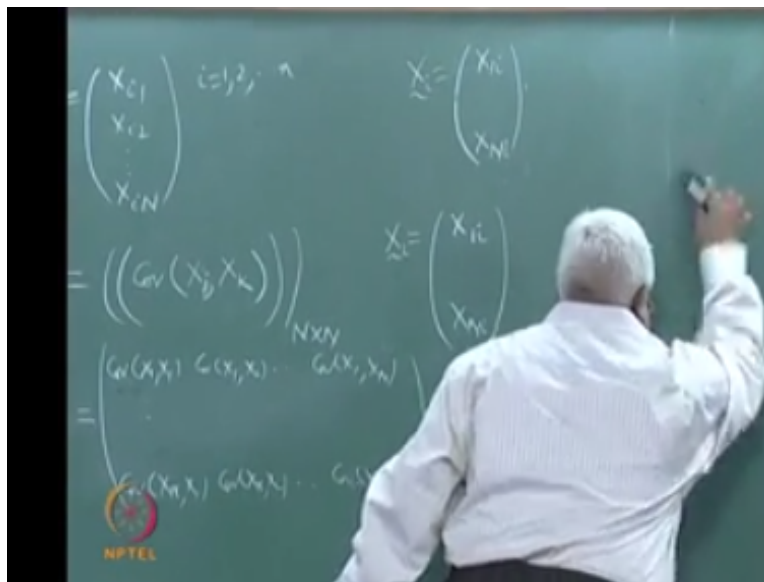
Let me write this vector X I has let me write this vector X I as xi1 xi2x in I is equal to so can we just use right then maybe I can take this thing as GIU Gaussian distribution Gaussian distribution with mean μ and variance covariance matrix Σ in general I a using capital n for the number of features capital n for the number of features so there are capital N number of features that means you are in capital N dimensional space and your number of points is small n number of points is smaller so this is vector I am writing it as xi1 xi2 x in now this covariance matrix Σ will be this is covariance of what this will be the covariance of here you.

Have basically you have one vector X which is occurring which has the same distribution you have one vector X which has same distribution throughout so it is giving rise to the first one X 1 the second one X to the third one X 3 and soon and so forth so you are trying to find out covariance between the eighth component of this vector and the jet component of this vector like

that you are going to find the covariance between the covariance matrix will be a capital N by capital N matrix since you have capital N number of variables your covariance matrix will be a capital N by capital n matrix.

So if your covariance matrix is capital N by capital n matrix then you are basically going to have covariance between the jet component and gate component you might be wondering what this J is and what this case J is 1 2 n K is 1 ton and this J and K are your virginal vector is X you have X 1 to xn and the IH realization X I it is basically here it is X 1 I here it is X n I this is your X I this X I is same as these x this X I is same as this X I X1 if you write I is equal to 1 here you will get X 1, 1 2 xn 1 I is equal to 2 X1 2 ,2 xn 2 and I is equal to small n X 1small n X capital n small n and you will find the covariance between.

The jet and the K component of the original vector X of the original vector X the jet and K component of the original vector X which I am representing it as XJ XK and your covariance matrix will be covariance X 1X 1 covariance X 1 X 2 and covariance X1 X N and this is covariance X n comma X1 covariance X n comma X 2 and this is covariance X n comma X n which is nothing but the variance this is the population variance covariance matrix which is not known to us this is what we would like to estimate.

What we have got are these things what we have got and what we have we have these things now from these we would like to estimate this now how does one estimate it here again you have consistent maximum likelihood and unbiased estimates now there is a small difference here between unbiased and maximum likelihood estimate the unbiased estimate of the covariance matrix covariance between jet and jet and K it component covariance between J and K components.

(Refer Slide Time: 18:06)

$$\widehat{\text{Co}}(X_j, X_k) = \frac{1}{n-1} \sum_{i=1}^{n} (X_{ji} - \bar{X}_j)(X_{ki} - \bar{X}_k) \rightarrow \text{UNBIASED ESTIMATE}$$

$$\text{where } \underline{\bar{X}}' = (\bar{X}_1, \cdots, \bar{X}_N)$$

$$\widehat{\text{Co}}(X_j, X_k) = \frac{1}{n} \sum_{i=1}^{n} (X_{ji} - \bar{X}_j)(X_{ki} - \bar{X}_k) \rightarrow \text{MLE}$$

The estimate is represented by a hat okay this is equal to 1 by n- 1summation I is equal to 1 to small n j is coming here where this is unbiased estimate this is unbiased estimate these are the means of the small n observations for each feature unfortunately this is also a variable for different sets of realizations you might get different means so these are all variables so I had to write represent them in capital letters so for the jet variable I have taken the jet sample mean for the case variable I have taken the case sample mean and I wrote here n - 1 this is unbiased estimate whereas if I write here small n it is going to be maximum likelihood estimate.

Which was what I was telling you in my one of the previous classes if I write there small n it will be maximum likelihood estimate there is that Emily is the standard form maximum likelihood estimate M L e this is how you estimate that again small n mine one by small n minus one is unbiased estimate if I write there one by n it is going to be maximum likelihood estimate MLU now in my previous class one of the previous classes I was mentioning these two things without giving the proper mathematics here this is the proper mathematics I could have written here see were COV XJ XK there is a hat this is1 by n summation.

I is equal to 1 to N x ji - XJ bar into then this is MLE for a very large in whether you write n RN -1 it does not matter and I also mentioned in that class that statisticians would like to take 1 by n minus 1 not 1 by M and I also mentioned that I was telling you about 1 by n because that is what generally people have learnt in their childhood when you learn mean and variance you write their1 by small n not 1 by n- 1 anyway since I have done this thing from now onwards I would

expect you to write 1 by n - 1 not 1 by n do you know the meaning of unbiased estimate the meaning of unbiased estimate is this just one minute please yeah, so I will tell you the definition of unbiased estimate suppose.

(Refer Slide Time: 22:36)
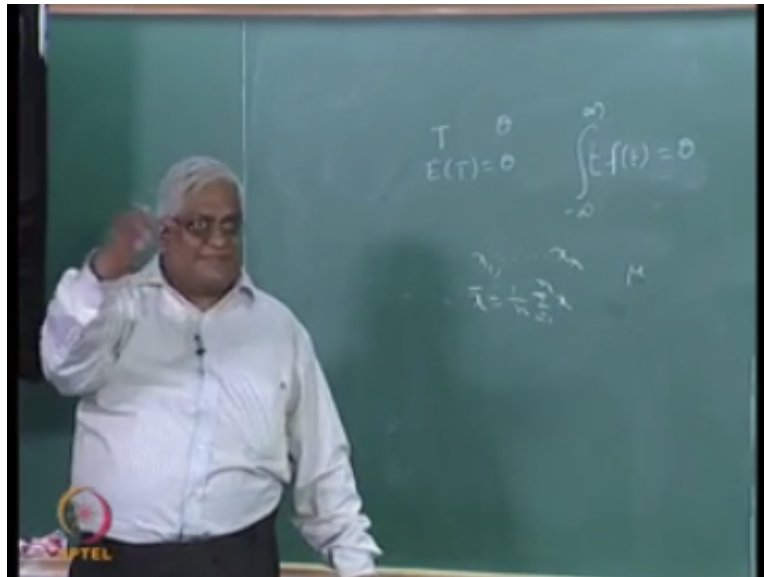
P is a function of the sample observations and you would like to estimate θ then T is said to be unbiased for θ if expected value of T is actually θ expected value of T is actually θ now what is the meaning of expected value of T you are going to find the mean of this T expectation means the mean that means integral T capital T is taking the value of small T with its density function suppose the density are represented by ft and suppose it is ranging over minus ∞+ ∞ a set of observations are a set of a function of the sample observations P is a function of the sample observations θ is the parameter that is to be estimated then P is said to be an unbiased estimator of θ if expected value of p is equal to θ the expected value of t means.

You are finding the mean of T you are finding the mean of T if that is θ then you say that T is an unbiased estimate of θ now I said that the sample mean if you have say forget about multiple dimensions let us just say we are in single dimension x1 x3 xn this mean x-bar which is 1 over N summation XI this is a this is an unbiased estimate of the population mean μ population mean μ there is no bar what is the meaning of this you take all the realizations of this thing X 1 this is realization of the first random variable capital X 1 X n it is a realization of the N at random variable capital X n which are independent and identically distributed.

Which are independent and identically distributed so you will get many values of this X bar you will get many values of this X bar in one realization you will get one X bar in another lesson you will get another X bar and you will take the average of all these realizations that is actually mean I hope you have understood the basic feeling of unbiased estimate if it is not this if it is something else that means our estimator is not able to estimate this thing properly his statistician would generally like to have would necessarily like to have unbiased estimates because if it is not unbiased then there is a bias you are not able to you are going shifting away from me.

You are shifting away from me there is a bias which a statistician generally does not like that is where statisticians prefer unbiased estimates I hope I have answered your question maximum likelihood estimate is different maximum likelihood estimate the intuition is if you have got this set of observations which particular value of θ will provide with maximum probability the set of observations I am repeating which particular value of the parameter θ will provide you this n points maximum likelihood that is why there were things maximum likelihood so that is maximum likelihood estimate and unbiased estimate.

I have already told you consistent estimate means as the number of points goes to ∞ this function that you have taken is it actually going to the actual value that is consistent estimate as the number of points goes to ∞ this function that you have taken it should if it goes to the actual value then that inconsistent unbiased means on an average you should get this θ maximum likelihood means which particular value of θ will provide will give maximum probability for this sample or maximum likelihood for this sample that is maximum likelihood estimate it is not necessarily true that all these three things should be same and in the case of variance and covariance they are not saying any other question I think stop for this.

End of Module 02-Lecture 3

Online Video Editing /Post Production
M.Karthikeyan
M.V.Ramachandran
P.Baskar

Camera
G.Ramesh
K.Athaullah
K.R.Mahendrababu

K.Vidhya
S.Pradeepa
D.Sabapathi
Soju Francis
Selvam
Sridharan

Studio Assistants
Linuselvan
Krishankumar
A.Saravanan

Additional Post –Production
Kannan Krishnamurthy& Team
Animations
Dvijavanthi

NPTEL Web &Faculty Assistance Team
Allen Jacob Dinesh
Ashok Kumar
Banu.p
Deepa Venkatraman
Dinesh Babu.K.M
Karthick.B
Karthikeyan.A
Lavanya.K
Manikandan.A
Manikandasivam.G
Nandakumar.L
Prasanna Kumar.G
Pradeep Valan.G
Rekha.C
Salomi.J
Santhosh Kumar Singh.P
Saravanakumar.P
Saravanakumar.R
Satishkumar.G
Senthilmurugan.K
Shobana.S
Sivakumar.S
Soundhar Raja Pandian.R
Suman Dominic.J
Udayakumar.C
Vijaya.K.R
Vijayalakshmi
Vinolin Antony Joans

Administrative Assistant

K.S Janakiraman

Principal Project Officer
Usha Nagarajan
Video Producers
K.R.Ravindranath
Kannan Krishnamurty

IIT Madras Production

Funded By
Department of Higher Education
Ministry of Human Resource Development
Government of India

www.nptel.ac.in

Copyrights Reserved