Dealing with Materials Data

Professor M P Gururanjan
Professor Hina A Gokhale
Department of Metallurgical Engineering and Materials Science
Indian Institute of Technology, Bombay

Lecture 51

Parameter Estimator 2

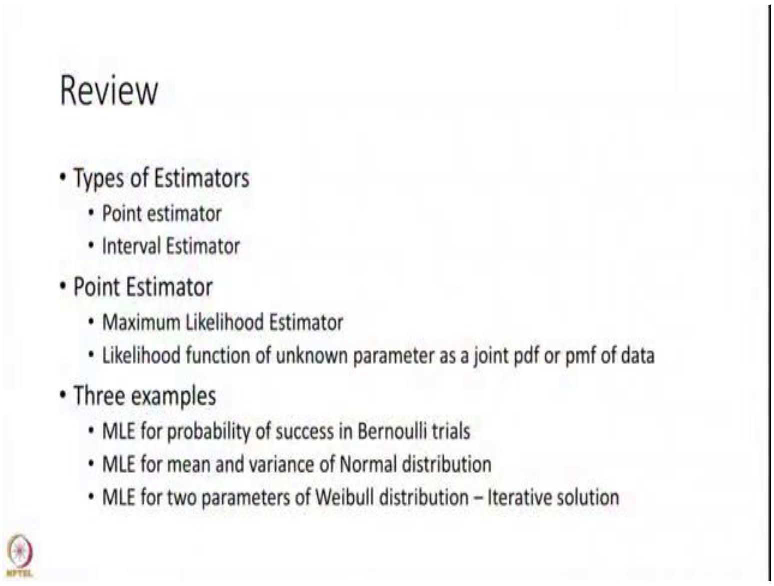
Hello and welcome to the course on Dealing with Materials Data. As I mentioned in the previous session we are now going through a process of explaining parametric estimation. Briefly the idea is that there is a population for which we know very little and we would like to know more about the population in order to make more inferences about the population, through observing a small sample.

It is a very common practice and I would like to briefly explain here, why do we do that because in the area of production of some metallic components. We would like to have, we would like to guarantee a certain quality or a certain parametric value for that product that we give.

For example, if we are giving a certain flat product and we may say that the flat product has a certain strength property. Now, we cannot test each and every product to have what strength property it has so all the product produced in a industry is the population. But, what we do is we draw a randomly a small sample, test their strength properties and then we declare that this is the kind of strength property our population, our production will have.

How to go about analysing this data? The small sample realisation that we get to come up with the values for the strength properties, this is the question we are trying to answer through parametric estimation.

(Refer Slide Time: 02:28)



So, we talked in the previous session that there are two types of estimator. One is called a point estimator which would for example give you the exact value of say yield strength of the product.

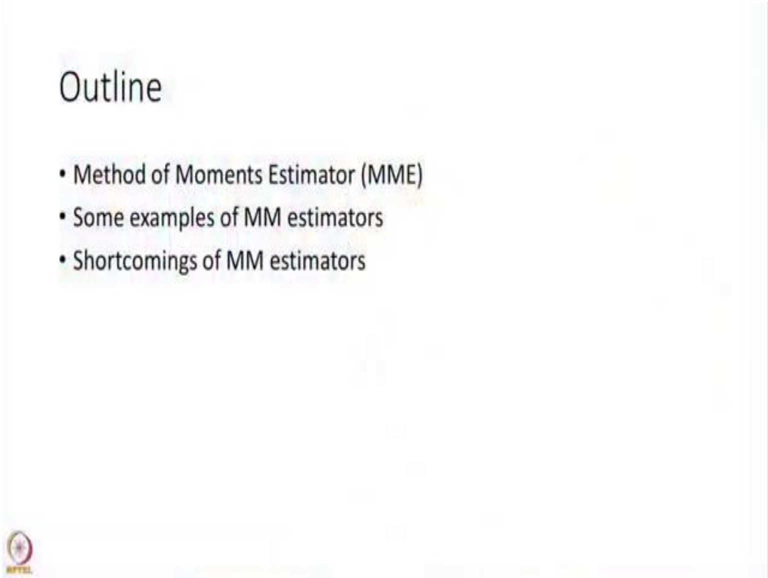
There could be another estimator in which we give an interval estimator. So, we say that instead of saying that the yield strength will be exactly 1300 MPa we instead say that it would lie in a certain interval before 95 percent of the time. It means that 95 percent of the product which will come out of this industry will have real strength falling in this interval.

We started discussing in the previous session about the point estimator and we talked about maximum likelihood estimator in which, we said that we know the population distribution. But, we know the form of the distribution but we do not know the parameter of the distribution. So, that parameter is say it is if it is theta then we said that take a joint density function of all the sample realization that we have got which is independent and identically distributed.

So, it is simply a product of their individual densities and then this function contains all possible information that you can get from the sample for the population with respect to theta. This is what we called a likelihood function of theta, unknown parameter and then we maximize this likelihood and the point that gave you the maximum value we called it a point estimator, maximum likelihood point estimator of the parameter theta.

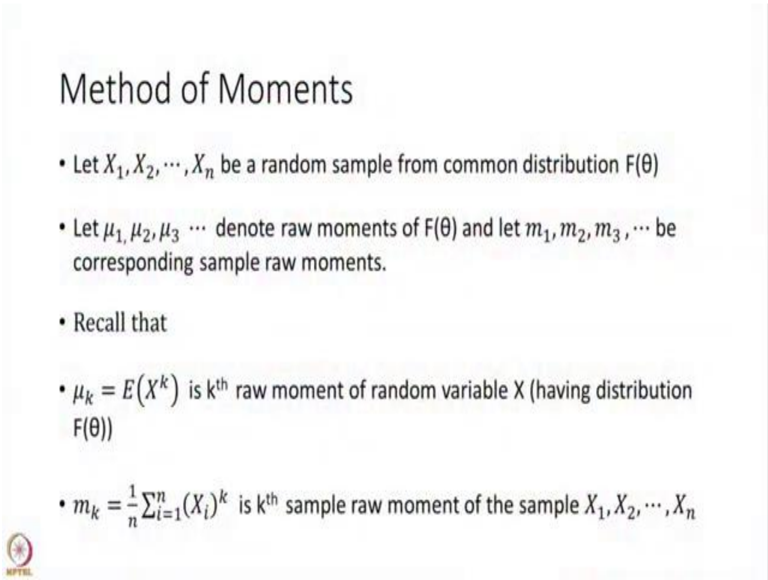
We talked three examples, two examples Bernoulli trials and normal very simple and straight forward. Where you take the derivative of the maximum likelihood function, equate it to zero and then you find, what is the value of the one or two unknown parameters. But, then we also took the case of Weibull distribution in which we found that you have to find a solution using numerical methods or iterative methods.

(Refer Slide Time: 05:03)



So, now this time we want to talk about method of moments estimator, we will again give the examples of Bernoulli and normal. But, we will also like to point out a short coming of method of moments estimator by giving one example.

(Refer Slide Time: 05:23)



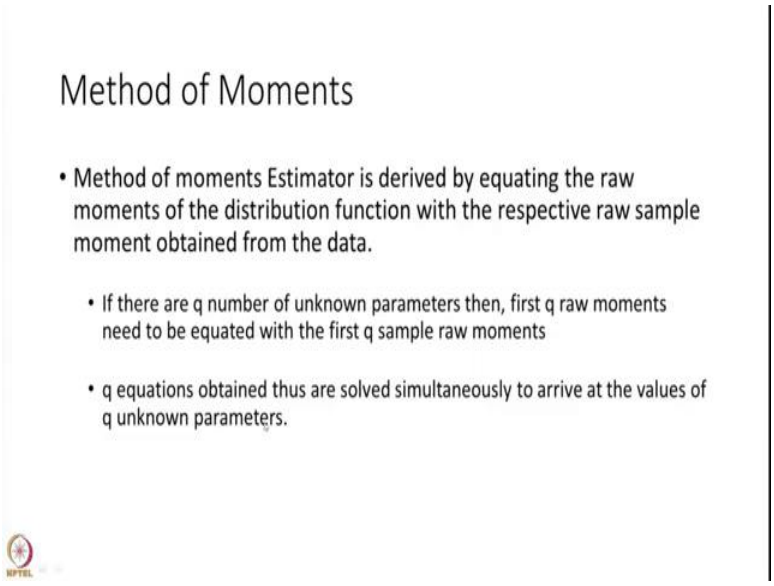
Suppose we have a random sample Let $X_1, X_2, \cdots, X_n$ be a random sample from common distribution F(θ)

Let $\mu_1, \mu_2, \mu_3 \cdots$ denote raw moments of F(θ) and let $m_1, m_2, m_3, \cdots$ be corresponding sample raw moments.

$\mu_k = E(X^k)$ is k^th^ raw moment of random variable X (having distribution F(θ))

$m_k = \frac{1}{n}\sum_{i=1}^{n}(X_i)^k$ is k^th^ sample raw moment of the sample $X_1, X_2, \cdots, X_n$

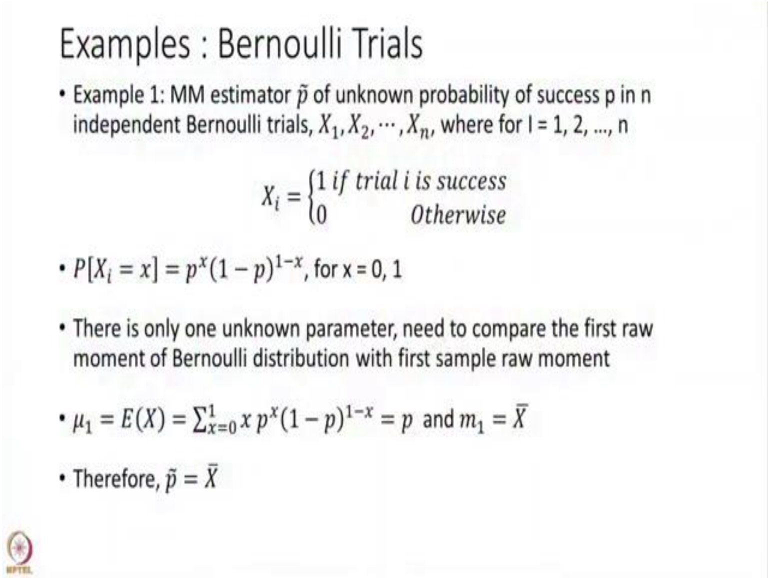
(Refer Slide Time: 06:23)



Now, what we do in method of moments is that, we equate the population moments with the sample raw moments. We take the population first raw moment and equate it to the sample first raw moment. Population second moment, raw moment and we equate it to the sample second raw moment like this. Now, if there are q number of parameters then we take the first q raw moments from the population and first q raw moments from the sample and we equate them.

And we solve these equations because on one side, when you talk about the population moments there is an unknown parameter theta in it. While the sample raw moments are not unknown because there is a realisation of data and from data we can calculate it. So, we have an equation with q unknowns, q equations with q unknowns and we need to solve them simultaneously.

(Refer Slide Time: 07:39)



Let us take an example, of Bernoulli trial which is the simplest of the kind. X1, X2, X3, Xn come from a Bernoulli population where the probability of success is p and Xi is 1 if trial is success and Xi is 0 if trial is not successful.

$$X_i = \begin{cases} 1 \; if \; trial \; i \; is \; success \\ 0 \quad\quad\quad Otherwise \end{cases}$$

$P[X_i = x] = p^x(1-p)^{1-x}$, for x = 0, 1

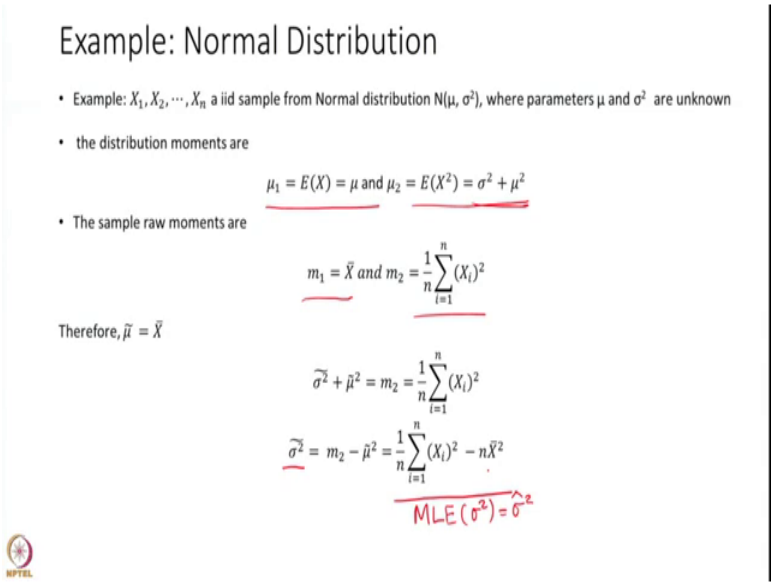
There is only one unknown parameter, need to compare the first raw moment of Bernoulli distribution with first sample raw moment

$$\mu_1 = E(X) = \sum_{x=0}^{1} x\, p^x (1-p)^{1-x} = p \text{ and } m_1 = \bar{X}$$

Therefore, $\tilde{p} = \bar{X}$

Remember, we got the same result in the maximum likelihood estimator, this is not a rule this is an exception. Anyway let us continue another thing which will look like a rule.

(Refer Slide Time: 09:18)



If you take a normal distribution, we take a sample X1, X2, X3, Xn instead of calling it every time random I have chosen now to call it independent identically distributed sample Normal Distribution N (μ, σ^2), where parameters μ and σ^2 are unknown

$\mu_1 = E(X) = \mu$ and $\mu_2 = E(X^2) = \sigma^2 + \mu^2$

The sample raw moments are

$$m_1 = \bar{X}\ and\ m_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i)^2$$

Therefore, $\widetilde{\mu} = \bar{X}$

$$\widetilde{\sigma^2} + \tilde{\mu}^2 = m_2 = \frac{1}{n}\sum_{i=1}^{n}(X_i)^2$$

$$\widetilde{\sigma^2} = m_2 - \tilde{\mu}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i)^2 - n\bar{X}^2$$

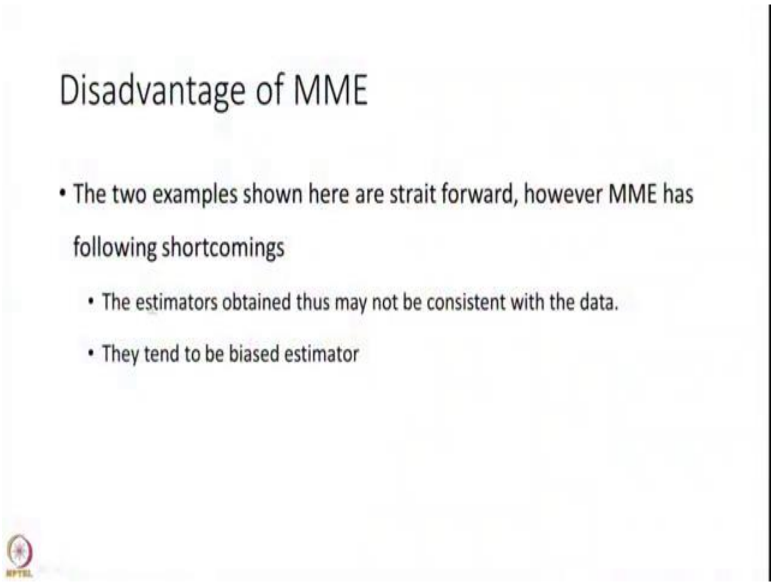
that is the MME estimator or method of moment estimator of mean value of normal population is same as the sample mean or sample average.

Now, you find that you have to equate the second moment that is second moment of population with the second raw moment of the sample. And if you simplify, you will find that the method of moment estimator of variance of normal distribution sigma square curled is equal to, actually this is MLE of sigma square which is we call it sigma square hat, you please confirm this, this also a good case but now let us consider the case where this may not always hold true. And this is why I am going to I have decided to discuss this method here.

Method of moments is a very attractive method and as you saw in the two simple examples, they very easily give us in a very much simpler manner, the maximum likelihood estimator for the unknown parameters in the case of Bernoulli trials as well as in the case of normal distribution.

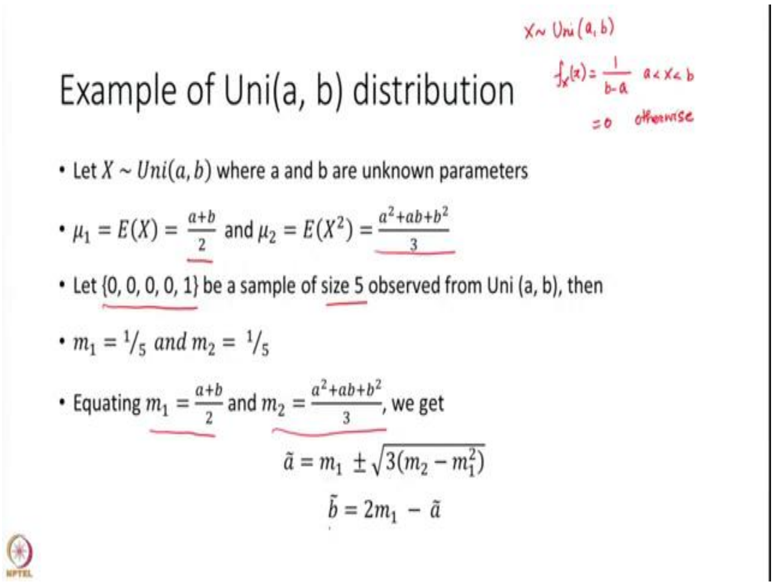
Remember maximum likelihood estimator you have to find a likelihood function then take a log likelihood, then its take derivative then you equate to 0 and then solve the equation. Compare to this in these two examples you must have seen, that finding a method of moments estimator MME is much simpler.

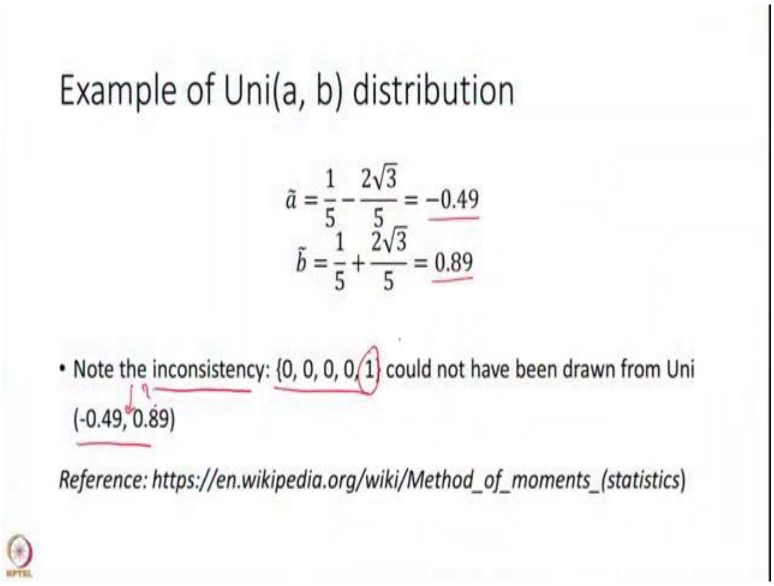
(Refer Slide Time: 12:54)



But, at times there is a disadvantage connected to this method of moments estimator or at times it is called moments matching estimator. Sometimes they are inconsistent with the data and number of times they tend to be biased estimator. This we are going to define later but here I would like to show that, it may not be very consistent with the data that you have got. And one example will sufficed for that.

(Refer Slide Time: 13:24)





Please note this example I have picked up from the Wikipedia and I have given the reference at the end. You are also welcome to go through and read through it. It gives a very good description of method of moments. So, let us consider a uniform distribution with two unknown parameters a and b, I hope you recall. When x is distributed uniform with parameter a and b

Let $X \sim Uni(a,\ b)$ where a and b are unknown parameters

$\mu_1 = E(X) = \frac{a+b}{2}$ and $\mu_2 = E(X^2) = \frac{a^2+ab+b^2}{3}$

Let {0, 0, 0, 0, 1} be a sample of size 5 observed from Uni (a, b), then

$$m_1 = {}^1/_5 \ and \ m_2 = {}^1/_5$$

Equating $m_1 = \frac{a+b}{2}$ and $m_2 = \frac{a^2+ab+b^2}{3}$, we get

$\tilde{a} = m_1 \; \pm \sqrt{3(m_2 - m_1^2)}$

$$\tilde{b} = 2m_1 \; - \; \tilde{a}$$

$$\tilde{a} = \frac{1}{5} - \frac{2\sqrt{3}}{5} = -0.49$$

$$\tilde{b} = \frac{1}{5} + \frac{2\sqrt{3}}{5} = 0.89$$

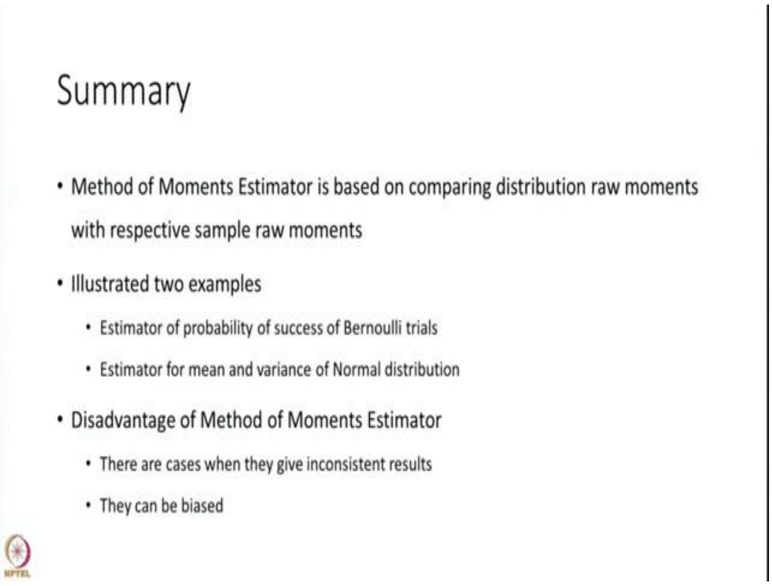
Note the inconsistency: {0, 0, 0, 0, 1} could not have been drawn from Uni (-0.49, 0.89)

I wish to bring it out to you that method of moment estimators are easy to calculate. And very attractive very easily these days available on variety of software including R-programming. But, be careful when you use it. It is much better to use the maximum likelihood estimator compared to matching of moments estimator or method of moments estimator.

Then why do we have it? It is natural question, why do we have this estimator? Well sometimes finding maximum likelihood estimator is difficult, finding any other estimator involves lot of numerical calculations or very complicated equations. At that time we fall back on the method of moments estimator.

But what this example tells us is that, we have to be very careful, we have to solve it. And we cannot say that these are the good estimator unless we put it through certain test and certain observations.

(Refer Slide Time: 19:23)



So, with this we summarized what we discussed just now. We discussed the method of moments estimator and we said that it is based on comparing distribution raw moments with the respective sample raw moments. So, if you have q unknown parameters you take q distribution raw moments and equate them with the corresponding q sample raw moments. You get a q set of equations and you have to solve them simultaneously.

We gave two examples in which, it actually resulted into maximum likelihood estimator only. It was a very simple example of Bernoulli trials where we tried to estimate probability of success. Then we took the normal distribution where we tried to estimate its mean and variance. But, through the uniform distribution giving a one very specific sample that we may observed.

We found that there is a disadvantage connected with the method of moments estimator, that they give inconsistent result. Of course, I have not shown you whether how they become biased but they also tend to be biased, this is an additional information. I have given you the reference to the Wikipedia from where this example is taken, you are welcome to go through it. Thank you.