1. Introduction

Welcome to the series of E-learning modules on Method of Moments. In this module we are going to cover the basic principle of method of moments- Advantages, disadvantages and properties of moment estimators, estimation of certain population parameters by the method of moments.

By the end of this session, you will be able to understand and explain:

- The basic principle of method of moments
- Properties of moment estimators
- Advantages and disadvantages of the method
- The procedure to estimate certain parameters of the population by the method of moments

Introduction :
Any statistical investigation aims at making generalizations from sample to population. Moreover selecting a random sample is essential for drawing valid conclusions about the population Methods of estimation develops theoretical basis of connection between sample information and population model. This in turn permits inference about the population

In statistics, the method of moments is a method of estimation of population parameters such as mean, variance, median, etc. (which need not be moments), by equating sample moments with unobservable population moments and then solving those equations for the quantities to be estimated.

The method of moments is the oldest method of deriving point estimators. It almost always produces some asymptotically unbiased estimators, although they may not be the best estimators.

The method of moments consists in equating the first few moments of the population with the corresponding moments of the sample.

The method of moments in mathematical statistics is one of the general methods for finding statistical estimators of unknown parameters of a probability distribution from results of observations. The method of moments was first used to this end by Karl Pearson in eighteen ninety four to solve the problem of the approximation of an empirical distribution by a system of Pearson distributions

Since parameters enter into the population moments these relations when solved for the parameters give estimates by the method of moments. Of course the method is applicable only when the population moments exist. The method is generally applied for fitting theoretical distributions to the observed data.

2. Procedure of the Method of Moments

The procedure in the method of moments is this: The moments of the empirical distribution are determined (the sample moments), equal in number to the number of parameters to be estimated; they are then equated to the corresponding moments of the probability distribution, which are functions of the unknown parameters; the system of equations thus obtained is solved for the parameters and the solutions are the required estimates.

In practice the method of moments often leads to very simple calculations. Under fairly general conditions the method of moments allows one to find estimators that are asymptotically normal, have mathematical expectation that differs from the true value of the parameter only by a quantity of order 1/n and standard deviation that deviates by a quantity of order 1/√n

However, the estimators found by the method of moments need not be best possible from the point of view of efficiency: their variance need not be minimal. For a normal distribution the method of moments leads to estimators that coincide with the estimators of the maximum-likelihood method, that is, with asymptotically-unbiased asymptotically-efficient estimators.

Scientifically the method of moments is carried out as follows:
Let f of (x, theta one, theta two, etc theta k) is the density function of the population under consideration. To estimate the unknown parameters theta one , theta two, etc theta k if mu r dash denotes the r^{th} moment (about zero) then by definition
Mu r dash is equal to integral of x to the power r into f of (x, theta one , theta two, etc theta k) dx where r is equal to one, two,.. etc

That is, mu 1 dash is equal to integral of x into f of (x, theta one , theta two, etc theta k) dx mu 2 dash is equal to integral of x to the power 2 into f of (x, theta one , theta two, etc theta k) dx etc. Mu k dash is equal to integral of x to the power k into f of (x, theta one, theta two, etc theta k) dx. Mu 1 dash, mu 2 dash, etc mu k dash is in general functions of the parameters theta 1, theta 2, till theta k. Thus the above is a set of k equations involving k unknown parameters theta 1, theta 2, etc till theta k.

Now solving the equations theta 1, theta 2, etc theta k, can be written as the functions of Mu 1 dash, mu 2 dash, etc till mu k dash. But in general Mu 1 dash, mu 2 dash, etc mu k dash are unknown and hence their estimators by sample moments m 1 dash, m 2 dash, etc m k dash respectively where m r dash is equal to summation xi to the power r by n, x one , x two, till xn being sample observations are determined.

In case of frequency distribution of the sample observations $r^{th\ sample}$ moment m r dash is given by m r dash is equal to summation f i into x i to the power r by N. Thus the method of moments consists in equating r^{th} raw moments about the origin in the population to the r^{th} raw moments about the origin in the sample by giving values r is equal to one, two, etc. And obtaining various equations containing parameters and solving these equations to obtain the estimate of the parameters.

Let x one, x two, etc, till xn be a random sample of size n from a population with probability density function(x, theta). Then xi (i equal to 1, 2 etc., till n) are i. i. d. Hence if Expected value of x_i to the power r exists , then by Weak Law of Large Numbers we get
1 by n summation xi to the power r tends in probability to Expected value of xi to the power r which implies m r dash tends in probability to mu r dash
Hence sample moments are consistent estimators of the corresponding population moments.

Method of Moments (MOM) is a numerical technique used to approximately solve linear operator equations such as differential equations, or integral equations. The unknown function is approximated by a finite series of known expansion functions with unknown expansion coefficients.

The approximate function is substituted in the original operator equation and the resulting approximate equation is tested so that the residual is minimized in some sense. This results into a number of simultaneous algebraic equations for the unknown coefficients. These equations are then solved using matrix calculus.

3. Properties, Advantages and Disadvantages of Moment Estimators

Properties of Moment Estimators:

- Moment estimators are consistent provided the population moments exists
- Moment estimators need not be unbiased
- Under certain conditions moment estimators have an asymptotic Normal distribution
- Moment estimators are less efficient than maximum likelihood estimators

Advantages and disadvantages of this method:

In some respects, when estimating parameters of a known family of probability distributions, this method was superseded by Fisher's method of maximum likelihood, because maximum likelihood estimators have higher probability of being close to the quantities to be estimated.

However, in some cases, as in the example of the gamma distribution, the likelihood equations may be intractable without computers, whereas the method-of-moments estimators can be quickly and easily calculated by hand.

Estimates by the method of moments may be used as the first approximation to the solutions of the likelihood equations, and successive improved approximations may then be found by the Newton–Raphson method. In this way the method of moments and the method of maximum likelihood are symbiotic.

In some cases, infrequent with large samples but not so infrequent with small samples, the estimates given by the method of moments are outside of the parameter space; it does not make sense to rely on them then. That problem never arises in the method of maximum likelihood. Also, estimates by the method of moments are not necessarily sufficient statistics, i.e., they sometimes fail to take into account all relevant information in the sample.

When estimating other structural parameters (e.g., parameters of a utility function, instead of parameters of a known probability distribution), appropriate probability distributions may not be known, and moment-based estimates may be preferred to Maximum Likelihood Estimation. Method of Moments has been used to solve vast number of electromagnetic problems during the last five decades.

The Method of Moments technique, as applied to problems in electromagnetic theory, was introduced by Roger F.Harrington in his nineteen sixty seven seminar paper, “Matrix Methods for Field Problems”. The implementation of the Method of Moments, by Poggio and Burke at Lawrence Livermore National Labs during the nineteen seventy’s, established this solution technique as a mainstay in the design of wire and wire array antennas.

4. Illustrative Examples

Illustrative examples
Estimate the parameter p in sampling from a Binomial population with parameter n known by the method of moments
Solution:
Since only one parameter is unknown we need to find mu one dash only. But for a Binomial variable with parameters n and p Mu one dash is equal to Expected value of X which is equal to n into p. Now, from the method of moments unknown parameter p is estimated such that mu one dash is equal to m one dash.

But m one dash is equal to summation xi by k which is equal to x bar where x one, x two , etc till xk are the k observations drawn from the Binomial population.
Hence mu one dash is equal to n into p which is equal to m one dash which is equal to x bar.
Therefore Moment estimator of p is given by p cap which is equal to m one dash by n which is equal to x bar by n

1) Obtain an moment estimator of theta of a Uniform population with parameters zero and theta
   We know that when xi follows Uniform distribution with range zero and theta
   The probability density function f of x is equal to one divided by theta while xi can range from zero to theta
   The first population moment mu one dash is equal to theta by two
   Now from the method of moments unknown parameter theta is estimated such that mu one dash is equal to m one dash

   But m one dash is equal to summation xi by n, i runs from one to n, which is equal to x bar. Hence mu one dash is equal to theta by two which is equal to m one dash which is equal to x bar. Therefore Moment estimator of theta is given by theta cap is equal to two into x bar

2) Find the moment estimator of theta in an exponential distribution with mean theta we know that when x i follows exponential distribution with mean theta, the probability density function f of x is equal to one by theta into e to the power minus xi by theta, zero less than xi less than infinity.
   The first population moment mu one dash is equal to theta. Now from the method of moments unknown parameter theta is estimated such that mu one dash is equal to m one dash

   But m one dash is equal to summation xi by n, i runs from one to n, which is equal to x bar. Hence mu one dash is equal to theta which is equal to m one dash which is equal to x bar. Therefore Moment estimator of theta is given by theta cap is equal to x bar

5. Illustrations Contd

Obtain the moment estimators of mean and variance of a Normal population with parameters mu and sigma square.
Let Xi follow Normal distribution with mean mu and variance sigma square. Since two parameters are unknown we need to find mu one dash and mu two dash. But for a Normal variate with parameters mu and sigma square mu one dash is equal to Expected value of X which is equal to mu and mu two is equal to sigma square

Now from the method of moments unknown parameter mu and sigma square are estimated such that mu one dash is equal to m one dash and mu two dash is equal to m two dash. But m one dash is equal to summation xi by n, i runs from one to n, which is equal to x bar and m two dash is equal to summation xi square by n, i runs from one to n. Hence mu one dash is equal to mu which is equal to m one dash which is equal to x bar. From the method of moments unknown parameter mu is estimated such that mu one dash is equal to m one dash.

Hence an estimate of mu, mu cap is equal to x bar. But since mu two is equal to sigma square which implies mu two is equal to mu two dash minus mu one dash whole square which is equal to sigma square. Hence an estimate of variance Mu two cap is equal to sigma cap square which is equal to m two dash minus m one dash square which is equal to summation x i square by n minus x bar square which is equal to s square. Therefore Moment estimator of parameter mu and sigma square are given as mu cap is equal to x bar and sigma cap square is equal to summation xi square by n minus x bar square which is equal to s square.

Is moment estimate uniquely determined? Explain with an example.
Let x one, x two, etc till xn be a sample taken from a Normal population with mean zero and variance sigma square. Then mu one dash is equal to expected value of x which is equal to zero and mu two is equal to sigma square. But since mu two is equal to sigma square implies mu two is equal to mu two dash minus mu one dash square which is equal to mu two dash, since mean mu one dash is equal to zero. Now from the method of moments unknown parameter mu and sigma square are estimated such that mu one dash is equal to m one dash and mu two dash is equal to m two dash.

By equating mu two dash is equal to m two dash we get sigma square is equal to summation xi square by n. By equating mu two is equal to m two we get sigma cap square is equal to summation xi square by n minus x bar square which is equal to s square. Since m two is equal to m two dash minus m one dash square which is equal to summation xi square by n minus x bar square. If xi follows Normal distribution with mean zero and variance sigma square then the moment estimators of sigma square are
Sigma square is equal to summation xi square by n
Sigma square is equal to s square
Hence Moment estimators are not unique.

Let x one, x two, etc, xn be a sample taken from a Normal population with mean zero and variance theta. Find the moment estimate of theta. Given xi's follows Normal distribution with mean zero and variance theta. The first population moment mu one dash is equal to Expected

value of x which is equal to zero. The second population moment mu two dash is equal to expected value of x square which is equal to variance of x plus expected value of x square which is equal to theta plus zero equals to theta
Hence mu two dash is equal to theta

Now by the method of moments unknown parameter θ is estimated such that mu one dash is equal to m one dash and mu two dash is equal to m two dash
But m one dash is equal to summation xi, i runs from 1 to n by n which is equal to x bar and m two dash is equal to summation xi square, i runs from 1 to n by n. But s square is equal to summation, i runs from 1 to n (xi minus x bar) whole square by n which is equal to summation xi square by n minus x bar square which implies s square plus x bar square is equal to summation xi square by n. Hence mu one dash is equal to mu which is equal to m one dash which is equal to x bar is equal to zero. Call this as (1)

Mu two dash is equal to m two dash which implies theta is equal to s square plus x bar square which is equal to summation xi square by n. From equation one therefore, moment estimator of parameter theta is theta cap which is equal to summation xi square by n

Here's a summary of our learning in this session where we have understood:

- The method of moments
- The basic principles of method of moments
- Properties of the method
- Limitations and advantages of the method of moments
- Illustrative examples to determine the moment estimators