1. Introduction

Welcome to the series of E-learning modules on moments and quantiles. Here we discuss about raw and central moments, their relationships, uses etc. There are many quantiles in general but here in particular we discuss about median, quartiles, deciles and percentiles.

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

- Know about raw and central moments
- Understand the conditions for existence of moments
- Understand the relationship between raw and central moments
- Explain the uses of the moments
- Describe Quantiles, that is, median, quartiles, deciles and percentiles.

Introduction

- The concept of moments was borrowed from physics
- The characteristics of a frequency distribution are described by its moments
- The r^{th} moment of a set of values about any constant is the mean of the r^{th} powers of the deviations of the values from the constant
- Moments about any constant can be found
- The moments about the arithmetic mean are called central moments
- The moments about any other constant are called raw moments

If X is a continuous random variable, then r^{th} raw moment is given by,
μ r dash is equal to Expectation of (X power r)
If the continuous random variable X has probability density function f of (x), then mu r dash is equal to integral over x, x power r into f of x dx
(08)
r^{th} moment about any point A is given by,
mu r dash is equal to expectation of X minus A whole power r
Is equal to integral over x, x minus A power r into f of x dx
Suppose in above expression if we take A is equal to Expectation of (X) then we get moment about mean or central moment which is denoted by μ_r and is given by,
Mu r is equal to expectation of X minus E of X the whole square
This is equal to integral over x, x minus E of X the whole square into f of x d x.

2. Existence of Moments

Expectation of modulus of X always exists in the extended real numbers. R is equal to R union infinity union minus infinity and Expectation of modulus of X belongs to the closed interval zero comma infinity; that is, either Expectation of modulus of X is a non-negative real number or Expectation of mod X is equal to infinity.

- Expectation of X exists and is finite that is, Expectation of mod X is less than infinity∞.
- Expectation of mod X is less than infinity, implies mod expectation of X is less than or equal to expectation of mod X which is finite.
- If one less than or equal to r less than or equal to s, then
  Expectation of mod X power s is less than infinity implies expectation of mod X power r is less than infinity.

Note that the first two conditions have already been proved in the previous module. Now let us prove the third one.

Given expectation of mod X power 's' is less than infinity then we need to prove than expectation of X power r also exists, that is it is less than infinity.
Consider integral from minus infinity to infinity mod x power r into f of x d x.
This integral is divided as integral from minus one to one and minus infinity to minus one and one to infinity, which are combined as mod x greater than one.
That is, integral from minus one to one mod x power r into f of x d x plus integral mod x greater than one mod x power r into f of x d x
If r is less than s then mod x power r is less than mod x power z for mod x greater than one

Therefore
integral from minus infinity to infinity mod x power r into f of x d x is less than or equal to integral over minus one to one mod x power r into f of x d x plus integral mod x greater than one, mod x to the power s into f of x d x
is less than or equal to integral from minus one to one f of x d x plus integral mod x greater than one, mod x power to the power s into f of x d x
Since for minus one less than one, mod x power r is less than one
Integral over minus infinity to infinity mod x power r into f of x d x is less than or equal to one plus integral over mod x greater than one, mod x power s into f of x d x is less than infinity.
Implies, expectation of mod X power r exists for all one less than or equal to r less than or equal to s.

Note that This result states that if the moments of a specified order exist, then all the lower order moments automatically exist. However, the converse is not true. That is we may have a distributions for which all moments of a specified order exist but no higher order moment exist. We prove this with an example.

Let X be a random variable with probability density function
f of x is equal to two divided by x cube, for x greater than or equal to one and equal to zero for x less than one
We have Expectation of X is equal to integral over one to infinity x into f of x d x.
By substituting f of x and simplifying we get,
two into integral form one to infinity, x power minus two into d x

Is equal to minus two divided by x ranges from one to infinity is equal to two.

Now consider expectation of x square is equal to integral over one to infinity, x square f of x d x
This is equal to two into integral from one to infinity x power minus one d x.
Is equal to log x ranges from one to infinity which is equal to infinity.
Thus for the above distribution, 1^{st} order moment exists but 2^{nd} order moment does not exist.

Consider another example.
A random variable X has probability density function
f of x is equal to r plus one into ‘a’ to the power ‘r’ plus one divided by x plus ‘a’ power ‘r’ plus two where x is greater than or equal to zero and a is positive.
Now consider r^{th} raw moment,
Mu r dash is equal to expectation of X power ‘r’ is equal to ‘r’ plus one into ‘a’ power r plus one integral over zero to infinity x power r divided by x plus a power r plus two d x
substituting x is equal to a into y in the above integral and using beta integral function
Integral over zero to infinity, x power m plus one divided by one plus x power m plus n d x is equal to beta of m, n.

On simplification,
Mu r dash is equal to r plus one into a power r into beta of r plus one and one is equal to a power r, where beta m, n is equal to gamma m into gamma n divided by gamma m plus n.
However
Mu r plus one dash is equal to expectation of X power r plus one is equal to r plus one into a power r plus one into integral over zero to infinity x power r plus one divided by x plus a power r plus two d x tends to infinity.
As the integral is not convergent. Hence in this case only moments up to r th order exist and higher order moments do not exists.

3. Relationship between Raw and Central Moments

Now let us discuss the relation between raw moments and central moments.
We have already discussed the topic moments in the first semester. Let us revise the same.
The relation between raw and central moments is given by the expression,
Mu r is equal to Mu r dash minus r C one Mu r minus dash into μ one dash plus r C two Mu r minus two into Mu one dash square minus etc plus minus one power four into Mu one dash to the power r.
In particular on substituting r is equal to two, three and four in above and simplifying we get
Mu two is equal to Mu two dash minus Mu one dash square
Mu three is equal to Mu three dash minus three Mu two dash into Mu one dash plus two Mu one dash cube and
Mu four is equal to Mu four dash minus four into Mu three dash into Mu one dash plus six Mu two dash into Mu one dash square minus three Mu one dash power four.

Uses of first four moments

- The first moment about zero is the arithmetic mean
- The second central moment is the variance of the distribution
- The third central moment is a measure of skewness
- The fourth central moment is a measure of kurtosis

The measure of skewness is given by, beta one is equal to Mu three square divided by Mu two cube and the measure of kurtosis is given by beta two is equal to Mu four divided by Mu two square.

Based on the measures of skewness and kurtosis we can write the nature of the distribution. As beta one is always positive we use the sign of Mu three to decide whether the distribution is positively skewed or negatively skewed.

- If Mu three is positive, the distribution is positively skewed
- If Mu three is negative, the distribution is negatively skewed and
- If Mu three is equal to zero the distribution is symmetric

Similarly beta two gives the “peakedness” or flatness of the distribution. That is

- If beta two is less than three, the distribution is platykurtic,
- If beta two is greater than three the distribution is leptokurtic and
- If beta two is equal to three then the distribution has normal curve

4. Quantiles and Median

Now let us discuss about the quantiles.
Quantiles are points taken at regular intervals from the cumulative distribution function of a random variable.
Dividing ordered data into ‘q’ essentially equal-sized data subsets is the motivation for q-quantiles.
Quantiles are the data values marking the boundaries between consecutive subsets.

Put another way, the k^{th} q-quantile for a random variable is the value x such that the probability that the random variable will be less than x is at most k by q and the probability that the random variable will be more than x is at most (q minus k) divided by q is equal to one minus (k by q). There are q minus one,
of the q -quantiles, one for each integer k satisfying zero less than k less than q.

Some q-quantiles have special names

- The second quantile is called the median
- The three-quantiles are called tertiles or terciles represented byT
- The four-quantiles are called quartiles denoted by Q
- The five-quantiles are called quintiles denoted by QU
- The six-quantiles are called sextiles represented by S
- The 10-quantiles are called deciles denoted by D
- The 12-quantiles are called duo-deciles denoted by Dd
- The 20-quantiles are called vigintiles denoted by V
- The 100-quantiles are called percentiles denoted by P
- The 1000-quantiles are called permiles denoted by Pr

The median, quartiles deciles and percentiles most commonly used quantiles.

Median
Now let us discuss about the median.
We know that median divides the whole distribution into two equal parts.
Hence if a continuous random variable X has probability mass function f of x and if median of the distribution is taken as M, then
Integral from minus infinity to M f of x d x is equal to integral from M to infinity is equal to half.

Thus solving
Integral from minus infinity to M f of x d x is equal to half or integral from M to infinity f of x d x is equal to half for M, we get the value of the median.

Now let us consider Quartiles
Since quartiles divides the distribution into four equal parts, we have three quartiles, q one, q two and q three and are given by the equations,

Integral from minus infinity to Q one f of x d x is equal to one divided by four,
Integral from minus infinity to Q two f of x d x is equal to two into one divided by four is equal

to half and

Integral from minus infinity to Q three f of x d x is equal to three into one divided by four is equal to three by four.
Observe that Q two and Median are same.

5. Deciles and Percentiles

Now let us consider Deciles and Percentiles.
As deciles divide the distribution in to 10 equal parts, we have 9 deciles and i^{th} decile is given by the equation,
Integral from minus infinity to D I f of x d x is equal to I divided by 10, where I is equal to one, two etc till nine.

Similarly percentiles divides the distribution into 100 equal parts, we have 99 percentiles and I th percentile is given by the equation,
Integral from minus infinity to P I f of x d x is equal to I divided by 100, where I is equal to one, two etc till 99.

The probability distribution of a random variable X is f of (x) is equal to k into sin (one divided by five) into pi into x, where zero less than or equal to x less than or equal to five. Determine the constant k and obtain the median and quartiles of the distribution.

Solution :
As f of x is a probability density function,
Integral over x f of x d x is equal to one implies k into integral from zero to five sine one divided by five into pi into x d x is equal to one
Let us substitute; one divided by five into pi into x is equal to y.
As x is equal to zero y is equal to zero and as x is equal to five, y is equal to pi and d x is equal to five divided by pi d y.

Hence we get,
K into integral from zero to pi sine y into five divided by pi d y is equal to one
By integrating and simplifying, we get
k is equal to pi divided by 10
Now let us find the median M as follows.
We know that integral from zero to M f of x d x is equal to half implies k into integral from zero to M sine one divided five into pi into x d x is equal to half

Let us substitute (one divided by five)into pi into x is equal to y
As x is equal to zero, y is also equal to zero and as x is equal to M, y is equal to (one divided by five) into pi into M and dx is equal to (five divided by pi) d y
Implies t divided by pi into k into integral from zero to one by five into pi into M sine y d y is equal to half.

On integration we get,
five into k divided by pi into one minus cos y, range is given by, zero to one divided by five into pi into M is equal to half
Implies half into one minus cos one divided by five into pi into M is equal to half
Implies M is equal to five divided by two
Hence median of the distribution is five divided by two.
We know that the second quartile is same as median. Hence we find first and third quartiles only.

The first quartile Q one is given by
Integral from zero to Q one f of x d x is equal to one divided by four
Implies k into integral from zero to Q one, sine one divided by five into pi into x d x is equal to one divided by four.

Considering the same transformation as we have done in the case of median we get,
five divided by pi into k into integral from zero to one divided by five into pi into Q one sine y d y is equal to one divided by four
Implies half into one minus cos one divided by five into pi into Q one is equal to one divided by four.
Therefore cos one divided by five into pi into Q one is equal to half
Is equal to cos pi divided by three
Implies Q one is equal to pi divided by three into five divided by pi is equal to five divided by three.

Similarly the third quartile Q three is given by
Integral from zero to Q three f of x d x is equal to three divided by four
Implies k into integral from zero to Q three, sine one divided by five into pi into x d x is equal to three divided by four.
Considering the same transformation as we have done in the case of median we get,
five divided by pi into k into integral from zero to one divided by five into pi into Q three sine y d y is equal to three divided by four
Implies half into one minus cos one divided by five into pi into Q three is equal to three divided by four.

Therefore cos one divided by five into pi into Q three is equal to minus half is equal to cos pi minus pi divided by three is equal to cos two into pi by three.
Implies Q three is equal to two into pi divided by three into five divided by pi is equal to 10 divided by three.

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

- Understood raw and central moments
- Explained the conditions for existence of moments
- Understood the relationship between raw and central moments
- Described the uses of the moments
- Understood Quantiles especially the median, quartiles, deciles and percentiles