UNIT 2 SIGNIFICANCE OF MEAN DIFFERENCES, STANDARD ERROR OF THE MEAN

Structure

2.0 INTRODUCTION

The main function of statistical analysis in behavioural sciences is to draw inferences or made generalisation regarding the population on the basis of results obtained. Therefore the infrential statistics is that branch of statistics which primarily deals with inferences from a sample to a large population from which the sample has been taken. This depends on the fact that how good is the sample estimate. If the sample estimate is not good i.e. having the considerable error or not reliable, we could not be able to draw correct or good inference about the parent population. Thus before to draw inference about the whole population or to made generalisation, it is essential first to determine the reliability or trustworthiness of computed sample mean or other descriptive statistical measures obtained on the basis of a sample taken from a large population.

As an implication of the trustworthiness of the sample measures, we are concerned also with the comparison of two sample estimates with a view to find out if they come from the same population. In other words, the two sample estimates of a given trail of the population do not differ significantly from each other.

Here significant difference means a difference larger than expected by chance or due to sampling fluctuations.

Thus the present unit, highlights the concept of standard error of a sample mean and to compare the two sample means drawn randomly from large population. So that we may be able to test our null hypothesis scientifically, which is made in relation to our experiment or study and to draw the inferences about the population authentically.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Define and explain the meaning of inference;
- Describe the concept of statistics and parameters;
- Distinguish between statistics and parameters;
- Explain the meaning of significance, significance level;
- Elucidate their role and importance to draw inference and to make generalisation about the population;
- Explain and differentiate between Sampling Error, Measurement Error and Standard;
- Error of Mean value obtained on the basis of a sample from a population;
- Analyse the ‘t’ distribution and its role in inferential statistics;
- Describe the standard error of large and small size sample means;
- Analyse the mean of the population on the basis of the mean of a sample taken from the population with certain level of confidence;
- Determine the appropriate sample size for a experimental study or for a research work Compare the means of two sample means obtained from the same population;
- Differentiate between independent sample means and correlated sample means;
- Test the null hypothesis (Ho) made in relation to an experimental study; and
- Analyse the errors made in relation to testing the null hypothesis.

2.2 THE CONCEPT OF PARAMETERS AND STATISTICS AND THEIR SYMBOLIC REPRESENTATION

Suppose you have administered a verbal test of intelligence on a group of 50 students studying in class VIII of a school of your city. Further, suppose you find the mean I.Q. of this specified group is “105”. Can you from this data or information obtained on the relatively small group, say any thing about the I.Q. of all the VIII class students studying in your city. The answer is “Yes” but under certain conditions. The specified condition is “the degree to which sample mean (M) which is also known as “Estimate” represents its parent population mean which is known as “True Mean” or “Parameter”. Therefore the two terms Estimates and Parameters are defined as given below.

2.2.1 Estimate

The statistical measurements e.g. measures of central tendency, measures of variations, and measures of relationships obtained on the basis of a sample are known as “Estimates” or Statistics. Symbolically, these are generally represented by using the English alphabets e.g.

Mean = M, Standard Deviation = S.D. or σ, Correlation = r etc.

2.2.2 Parameter

The statistical measurements obtained on the basis of entire population are known as “True Measures” or “Parameters”.

Symbolically, these are represented by putting over the bar (-) over corresponding English alphabets or represented by Greek letters e.g.

True Mean or Population Mean = $\overline{M}$ or μ (Mu)

True S.D. or Population S.D. = $\overline{S.D.}$ or $\bar{\sigma}$

True or Population correlation = $\bar{r}$ or η

It is rarely if ever possible to measure all the units or members of a given population. Therefore, practically or for case we draw a small segment of the population with convenient specified number of units or members, which is known as the sample of the population.

Therefore, we do not know the parameters of a given population. But we can under specified condition, forecast the parameters from our sample statistics or estimates with known degree of accuracy.

2.3 SIGNIFICANCE AND LEVEL OF SIGNIFICANCE OF THE STATISTICS

Ordinarily, we draw only a single sample from its parent population. However, our problem becomes one of determining how we can infer or estimate the mean of the population (Mpop) on the basis of the sample mean (M). Thus the degree to which a sample mean (M) represents its parameter is an index of the “Significance” or Trustworthiness of the computed sample mean.

When we draw a sample from the population, the observed statistics or estimate that is the mean of the sample obtained, may be some time large or small to the mean of the population (Mpop). The difference may have arisen "by chance" due to the differences in the composition of our sample, or due to its selection method or the procedure followed in the sample selection. The gap between the two measures sample mean (M) and population mean (Mpop), if is low and negligible the sample mean is considered to be trustworthy and we can forecast or estimate the population mean (Mpop) successfully.

Therefore, a sample mean (M) is statistically trustworthy or significant to forecast the mean of the population (Mpop), depending upon the probability that the difference between the two measures i.e. Mpop and M could have been arisen "by chance".

And the confidence level to which this forecast has been made is known as level of confidence or level of significance.

In simpler terms, the level of significance or level of confidence is a degree to which we accept or reject or predict a happening or incidence with confidence.

There are a number of levels of confidence or levels of significance e.g. 100%, 99%, 95%, 90% .... 50% etc. In psychology and other behavioural sciences, generally, we consider only two levels of significance viz. the 99% level of significance or level of confidence and 95% level of significance a level of confidence.

The amount 99% and 95% confidence is also termed as 0.01 and 0.05 level of confidence. The 0.01 level means, if we repeatedly draw a sample or conduct an experiment 100 times, only on one occasion, the obtained sample mean or results will fallout side the limits Mpop ± 2.58 S.E.

Here the term S.E. means the standard error exists in the estimate or sample statistics.

Similarly 0.05 level means, if repeatedly draw a sample or conduct an experiment 100 times, only on five occasions the obtained sample mean will fall out side the limits Mpop ± 1.96 S.E.

The value 1.96 and 2.58 have been taken from the 't' distribution or 't' table (P...).

Keeping large size sample in view.

The 0.01 level is more rigorous and higher in terms of standard, as compared to the 0.05 level and would require a high level of accuracy and precision. Hence, if an obtained value (on the basis of a sample or an experiment) is significant at 0.01 level, it is automatically significant at 0.05 level but the reverse is not always true.

2.4 SAMPLING ERROR AND STANDARD ERROR

The score of an individual of a group obtained on a certain test consists of two types of errors (i) measurement error and (ii) statistics error.

In other words,

True Score X_T = observed score or obtained score (X_0) ± error (E)

and error E is = Measurement Error + Statistics Error

The measurement error is caused by a measuring instrument used to measure a trait or variable and personal observation made by the individual on the instrument.

The measurement error is due to the reliability of a test, as no test is perfectly reliable specially in behavioural sciences. In other words no test or a measuring instrument gives us 100% accurate measurement. The personal error is dependent upon the accurate perception and attention of the individual to take observations or measurement on the measuring instrument.

The statistics error refers to the errors of sample statistical measurements or estimates obtained on the basis of a sample drawn from a population.

As it is not possible to have perfectly true representative sample of a population in behaviourable sciences. The statistics error is of two types: (i) Sampling Error (ii) Standard Error of statistics i.e. statistical measurements. Now let us see what are these two errors in detail.

2.4.1 Sampling Errors

Sampling error refers to the difference between the mean of the entire population and the mean obtained of the sample taken from the population.

Thus sampling Error = Mpop ~ M or $\overline{M}$ - M

As the difference is low the mean obtained on the basis of sample is near to the population mean and sample mean is considered to be representing the population mean (or Mpop)

2.4.2 Standard Error

The standard error is nothing but the intra differences in the sample measurements of number of samples taken from a single population.

As the intra differences in the number of sampling observation i.e. the statistics of the same parent population is less and tending to zero, we may say the obtained sample statistics is quite reliable and can be considered as representative of the Mpop or $\overline{M}$.

For more clarification, suppose you wish to determine the I.Q. level of the high school going students studying in the various schools of your district. Is it possible for you to administer the intelligence test on all the high school going students of your district and get the Average I.Q.? Your answer may be certainly not.

The easiest method is to select a sample of 10 schools each from urban and rural areas of your district by using random method and administer the intelligence test to the students studying in high school class of these selected 20 schools in total. In such condition, you may have approximately 20 samples of high school going students and have 20 means of intelligence scores obtained on the intelligence test which you have administered. It is possible that all the mean values you have obtained are not equal. Some may be small and some may be large in their values. Theoretically, there should be no difference in the mean value obtained and all the mean values should be equal as all the samples are taken from the same parent population by using random method of sample selection. Thus the inter variation lies within the values of 20 mean values indicating that the error lies within the various observations taken.

Further, to have the mean I.Q. of all the high school going students you may calculate combined mean or average mean of all 20 means obtained on the basis of samples taken. This obtained combined or average mean value is the I.Q. of the parent population i.e. the high school going students of your district. If you compare all the

20 sample means to the obtained combined mean value i.e. the population mean, you will find that some of sample means are lesser than this Mpop value, and some are higher.

Further one step more, you calculate the difference of these sample means from the population mean obtained, i.e. find (Mpop – M_1), (Mpop – M_2) .... (Mpop – M_{20}). You will find that some of these difference values are negative and some are positive. The mean of all these differences should be zero and the standard deviation of all these differences should be 1.

As the errors are normally distributed in the universe, in simple terms we can say that we have a normal distribution of specific statistics or sample statistical measurement, which is also known as sampling distribution.

Therefore, theoretically the standard error of the statistics (sample statistical measurements) is the standard deviation of the sampling distribution of the statistics and is represented by the symbol $S.E._M$.

Standard Error of sampling measurements on statistics is calculated by using the formula given below :

$$\text{S.E.M or } \sigma_M = \frac{\bar{\sigma}}{\sqrt{N}}$$

Where, S.E.M = Standard Error of sampling measurement

$\bar{\sigma}$ = Standard deviation of the scores obtained from the population.

N = Size of the sample or total number of units in a sample

Look carefully and study the formula given above, you will find, the standard error of any statistics depends mathematically upon two characteristics.

i) the variability or spread of scores around the mean of the population and

ii) the number of units or cases in the sample taken from the population.

As there is low variability in the scores of population, i.e. the population is homogeneous on the trait being measured, and also the number of cases in the sample are too large, the standard error of the statistics is tending to zero.

In the formula, standard error of statistics is directly proportionate to the standard deviation (σ) of the scores of population and inversely proportionate to the size of sample or number of cases in the sample (N).

Thus in brief, it can be said that if the population is homogeneous to the variable or trait being measured and a large size of sample (say more than the 500 units), taken from the population; in such condition the sample drawn will be representative to its parent population and is highly reliable.

Self Assessment Questions

1) Explain the following terms:

   a) Estimate

   ....

   ....

b) Parameter

....

....

c) Statistics

....

....

d) Sampling Error

....

....

e) Measurement Error

....

....

f) Standard Error

....

....

2) What is the general formula to know the standard errors of the various statistical measures?

....

....

....

3) What do you mean by significance and levels of significance?

....

....

....

4) In behavioural sciences, which levels of confidence are considered

....

....

....

5) What is the difference between significance of statistics and confidence interval for true statistics?

....

....

....

2.5 't' RATIO AND 't' RATIO DISTRIBUTION TABLE

2.5.1 't' Ratio

So far we have studied two types of distributions viz.,

1) Distribution of scores – Normal distribution (unit I)
2) Distribution of 'statistics' or sample statistical measures (p -)

For more clarification, again going back. suppose we have drawn 100 samples of equal size (say n = 500) from a parent population and we calculate the mean value of the scores of a trait of the population obtained from each sample. Thus we have a distribution of 100 means.

Of course all these sample means will not be alike. Some may have comparatively large values and some may have small. If we draw a frequency polygon of the mean values or the "statistics" obtained, the curve will be "Bell – Shaped" i.e. the normal curve and having the same characteristics or properties as the normal probability curve has.

The distribution of statistics values or sample statistical measurements is known as the "sampling – distribution" of the statistics.

The corresponding standard score formula i.e. $z = \frac{X-M}{\sigma}$, may now become as –

$$t = \frac{M-M_{pop}}{S.E._M} \quad \text{or} \quad \frac{M-\overline{M}}{S.E.\overline{M}}$$

where t = standard score of the sample measures or statistics and termed as "t.Ratio"

M = Mean of specific statistics or sample measure.

$\overline{M}$ or Mpop = Mean of the parameter value of the specific statistics or mean of the specific statistics of the population

$S.E._M$ = Standard Error of the statistics i.e. the standard deviation of the sampling distribution of the statistics.

Actually, t is defined as we have defined the z. It is the ratio of deviation from the mean or other parameter, in a distribution of sample statistics, to the standard error of that distribution.

To distinguish z score of the sampling distribution of sample statistics, we use "t" which is also known as "student's t".

The "t" ratio was discovered by an English statistician, W.S. Gossett in 1908 under the pen name "student". Therefore, the "t" ratio is also know as "student's t" and its distribution is known as "student's t distribution".

As the "t" ratio is the standard score (like z score) with mean = 0 and standard deviation ±1, therefore the t ratio is a deviation of sample mean (M) from population mean ($\overline{M}$ or Mpop).

If this deviation is large the sample statistics mean is not reliable or trustworthy and if the deviation is small, the sample statistics mean is reliable and representative to the mean of its parent population ($\overline{M}$).

2.5.2 The Sampling Distribution of "t" Distribution

Just now we have studied about the sampling distribution of sample statistics and the "t" ratio. Imagine that we have taken number of independent samples with equal size from a population. Let us say we have computed the "t" ratio for every sample statistics with N constant. Thus a frequency distribution of these ratios would be a sampling distribution of "t" and is known as "t" distribution. The mean of all "t" ratios is zero and standard deviation of all "t" ratios i.e. σ is always $\pm\sigma_t$.

It has been observed that if the sample size varies the sampling distribution of "t" also varies, though it is normal distribution. The sampling distribution of "t" may vary in kurtosis. Student's t distribution becomes increasingly leptokurtic as the size of sample decreases.

As the size of sample is tending to be large, the distribution of "t" approaches to the normal distribution. Thus we have a family of "t" distributions, rather to one and the σt values varies on the x axis.

Fisher has prepared a table of "t" distribution having N, i.e. the size of sample different for different levels of significance. The details of the same are given below:

Table 2.5.1 : Table of "t" for use in determining the significance of statistics

Degrees of Freedom	Probability (P) 0.10	0.05	0.02	0.01
1	$t = 6.34$	$t = 12.71$	$t = 31.82$	$t = 63.66$
2	2.92	4.30	6.96	9.92
3	2.35	3.18	4.54	5.84
4	2.13	2.78	3.75	4.60
5	2.02	2.57	3.36	4.03
6	1.94	2.45	3.14	3.71
7	1.90	2.36	3.00	3.50
8	1.86	2.31	2.90	3.36
9	1.83	2.26	2.82	3.25
10	1.81	2.23	2.76	3.17
11	1.80	2.20	2.72	3.11
12	1.78	2.18	2.68	3.06
13	1.77	2.16	2.65	3.01
14	1.76	2.14	2.62	2.98
15	1.75	2.13	2.60	2.95
16	1.75	2.12	2.58	2.92
17	1.74	2.11	2.57	2.90
18	1.73	2.10	2.55	2.88
19	1.73	2.09	2.54	2.86
20	1.72	2.09	2.53	2.84
21	1.72	2.08	2.52	2.83
22	1.72	2.07	2.51	2.82
23	1.71	2.07	2.50	2.81
24	1.71	2.06	2.49	2.80
25	1.71	2.06	2.48	2.79
26	1.71	2.06	2.48	2.78
27	1.70	2.05	2.47	2.77
28	1.70	2.05	2.47	2.76
29	1.70	2.04	2.46	2.76
30	1.70	2.04	2.46	2.75
35	1.69	2.03	2.44	2.72
40	1.68	2.02	2.42	2.71
45	1.68	2.02	2.41	2.69
50	1.68	2.01	2.40	2.68
60	1.67	2.00	2.39	2.66
70	1.67	2.00	2.38	2.65
80	1.66	1.99	2.38	2.64
90	1.66	1.99	2.37	2.63
100	1.66	1.98	2.36	2.63
125	1.66	1.98	2.36	2.62
150	1.66	1.98	2.35	2.61
200	1.65	1.97	2.35	2.60
300	1.65	1.97	2.34	2.59
400	1.65	1.97	2.34	2.59
500	1.65	1.96	2.33	2.59
1000	1.65	1.96	2.33	2.58
∞	1.65	1.96	2.33	2.58

Let us now take an example. Let us say there are 26 subjects. N = 26.

Example: When N = 26, the corresponding degree of freedom (df) is N-1 i.e. 25.

In column 3 at 0.05 level of significance the t value is 2.06.

It means that five times in 100 trials a divergence of sample mean or statistics obtained may be expected at a 2.05 σ to M, that is to its mean population either to its left or right side.

The "t" distribution table has great significance in inferential statistics testing the null hypothesis framed in relation to various experiments made in psychology and education.

2.6 STANDARD ERROR OF SAMPLE STATISTICS – THE SAMPLE MEAN

The standard error of sample statistics or the statistical measurements of a sample has great importance in inferential statistics. With the help of the standard error statistics we can determine the reliability or trustworthiness of the descriptive statistics e.g. proportion percentage, measures of central tendency (mean, median and mode) measures of variability (standard deviation & quartile deviation), Measures of correlation (r, p and R) etc.

For convenience here we discuss only the significance of means which are detained as under :

2.6.1 Meaning of Standard Error of Mean

The Standard error of mean measures the degree to which the mean is affected by the errors of measurement as well as by the errors of Sampling or Sampling fluctuations from one random sample to the other. In other words how dependable is the mean obtained from a sample to its parameter i.e. population Mean (M pop).

Keeping in mind the Sample Size; there are two situations:

i) Large Sample

ii) Small Sample

2.6.2 The Standard Error of Mean of Large Sample

When we say large sample, the number of items in the sample will be more than 30. That is N>30. In such condition the Standard error of the Mean is determined by using the formula given below:

$S.E._M = \sigma / \sqrt{N}$ when N= >30

Where

$S.E._M$= Standard Error of the Mean of the scores of a large sample

σ = Standard deviation of the scores of a population

N = Size of the sample or number of cases in the sample

This formula is used when the population parameter of standard deviation (σ) is known. But in practice it is not possible to have the value of Ïó. Having the situation that the sample is selected from the population by using random method of sample selection, σ can be replaced by the ó. The value of the standard deviation of the scores of the sample taken. Therefore, in the above formula the σ can be replaced by ó.

$S.E._M = ó / \sqrt{N}$ when N >30

where

$S.E._M$ = Standard Error of the Mean of the scores of a Sample

ó = Standard Deviation of the scores of a sample.

N = Number of units or cases in the Sample.

Example 1: A reasoning test was administered on a sample of 225 boys of age group 15 + years. The mean of the scores obtained on the test is 40 and the standard deviation is 12. Determine how dependable the mean of sample is.

Given : N=225. M=40 and ó = 12

To find : The trustworthyness of the sample mean we know that standard error of the mean, when N>30 is determined by using the formula-

$S.E._M = ó / \sqrt{N}$

$S.E._M = 12 / \sqrt{225}$

$= 12 / 15 = 0.80$

Or $S.E_M = 0.80$

i.e. = 0.80

Interpretation of the Result

Keeping in mind the logic of sampling distribution, that is if we draw 100 samples, each sample has 225 units from a large population of boys of age group 15+ years, we will have 100 sample means falling into a normal distribution around the Mpop and $ó_M$ (the standard deviation of sampling distribution of Means i.e. the standard error of Mean)

As per properties of Normal Distribution, in 95% cases the sample means will lie within the range of ±1.96 in to the Mpop (see Z table in unit I). Conversely out of 100, the 99 sample means having equal size, will be within the range of ±2.57 (2.57 ×0.80) of the Mpop.

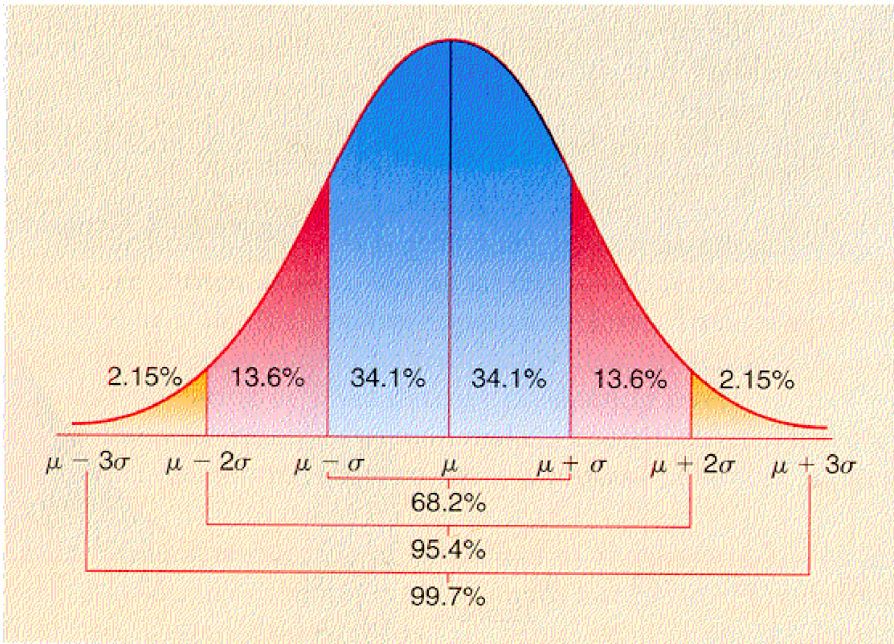

Fig. 2.6.1: Image graph

(**Source:** kmblog.rmutp.ac.th/.../28/normal-distribution/)

From the figure it is quite evident that the S.E. of the M=40. Sample of 225 having

bó = 12 lie within the acceptable region of the N.P.C.(Normal Probability curve). Thus the sample mean obtained is quite trustworthy with the confidence of 95% probability. There are only 5% chances that the sample mean obtained will lie in the area of the rejection of M.P.C.

In simplest term we can say that, there is 95% probability the maximum possibility of the standard error of the sample mean (40) is ± 1.57 (1.96×0.80) which is less than the value of T=1.96 at .05 level of confidence for df=224 (N-1) Thus the obtained sample mean (40) is quite dependable to its Mpop with the confidence level of 95%.

Example 2: In the example 1, suppose in place of N=225, we have a sample of 625 units and the remaining observations are the same. Determine how good an estimate is it of the population mean?

Solution

Given : N=625, M=40 and ó=12

To find : Dependency of sample Mean or reliability of sample mean

We know that

$$ó_M \,/\, S.E._M = \frac{\sigma}{\sqrt{N}}$$

$$= \frac{12}{\sqrt{625}}$$

$$= \frac{12}{25} = 0.48$$

Or $S.E._M$=0.48.

Interpretation of Result

The maximum standard error of sample M=40 and $\sum$=12 having 625 units is ±0.94 (1.96×0.48) at 95% level of confidence which is much less than the value of t.05 = ± 1.96. Therefore, the obtained sample mean is reliable and to the considered as representative to its Mpop at 95% level of confidence.

Self Assessment Questions

1) Compare the two results obtained from Example no 1 and 2 respectively. What you have observed and what is your conclusion.

....

....

....

....

....

2) The mean achievement score of a random sample of 400 psychology students is 57 and D.D. is 15? Determine how dependable is the sample mean?

....

....

....

....

3) A sample of 80 subjects has the mean = 21.40 and standard deviation 4.90. Determine how far the sample mean is trustworthy to its Mpop.

....

....

....

....

2.6.3 Degree of Freedom

Before to proceed for the standard error of small sample mean, it is imperative here to understand the concept of Degrees of Freedom.

The expression degrees of freedom is abbreviated from the full expression "Degrees of Freedom to Vary". A point in the space has unlimited freedom to move in any direction, but a point on a straight line has only one freedom to move on the line i.e. it is free to move in one dimension only. A point on the plane has two degrees of freedom. A point on a three dimension space having three degree of freedom to move.

This shows that the degree of freedom is associated to the number of restrictions imposed upon the observations.

The degree of freedom is a mathematical concept and is a of key importance in inferential statistics. Almost all test of significance require the calculation of degree of freedom.

When a sample statistics is used to estimate a parameter, the number of degrees of freedom depends upon the number of restrictions placed upon the scores, each restriction reducing one degree of freedom (df).

For example, we have four numbers 4, 5, 8 and 3, the sum of these numbers is 20 which is fixed. e.g. here we have restricted freedom to get sum 20, as we have only to change the values of first three figures and not the last one, as it depends upon the fixed sum 20,

i.e.

6 + 6 + 9 + ———-20

Or 2 + 7 + 10 + ———20

Or 6 + 7 + 5 + ———20

In the above sum expressions we have last one freedom to determine the value of last 4th numeral figure as to get fixed sum 20. Therefore in the above expressions, we are bound to take forcefully the numeral figures 4,1 and 2 respectively.

Like wise, in Statistics, when we calculate Mean or S.D. of a given distribution of scores we lose one degree of freedom to get fixed sums. Therefore we have N-1 df to compute statistical measures , specifically the standard deviation of the scores given and N-2 in case to compute the co-efficient of correlation. It means the degree of freedom is not always (N-1) however, but will vary with the problem and the restrictions imposed.

In the case of very large sample used in behavioural sciences or social sciences no appreciable difference takes place in the value of ó$_M$ by N-1 instead of N. The use of N or N-1 thus remains a matter of arbitrary decision. But in the case of small samples having number of units or cases below 30 ó$_M$ no correction (N-1) has been applied in computation of S.D and thus a considerable variation occurs on the value of ó$_M$ … Therefore, it is imperative to use N-1 in place of N in computation of ó$_M$ of the small sample.

Further you have to study the t" distribution table (table No-2.5.1) very carefully. In the process, you will find that as the size of sample or degree of freedom approaches to 500 or above, the "t" value approaches to the value of 95% and 99% are 1.96 and 2.58 respectively and remain constant. It means the "t" distribution becomes normal distribution or Z-distribution. When the size of sample decreases especially below 30 you will find the "t" values are gradually increasing at 95% level and 99% level considerably. In such condition the ó$_M$ values gives us wrong information and we may interpret the results inappropriately.

2.6.4 The Standard Error of Means of Small Sample

The small sample means, when size of sample (N) is about 30 or less, is treated as small sample. The formula for the standard error of small sample mean score is as follows–

S.E.$_M$ on S$_M$= ó / $\sqrt{N-1}$

As here S.E$_M$= Standard error of Mean of Small sample

Standard deviation of the population

N = Size of the sample i.e. 30 or below

Note : For practice we replace -ó by ó i.e. standard deviation of the sample. Because of the reason ó is not possible to obtain for whole population.

Example 3: A randomly selected group of 17 students were given a word cancellation test. The mean and S.D obtained for cancelling the words per minute is 58 and 8 respectively. Determine how far sample mean is acceptable to represent the Mean of the population?

Solution

Given : N=17, M=58 and ó = 8

To find : dependency of the sample mean

In the problem the size of sample is less than 30. Therefore to find the standard error of sample mean is

$$S.E_M = \frac{\sigma}{\sqrt{N-1}}$$

$$S.E_M = \frac{8}{\sqrt{17-1}}$$

$$= \frac{8}{\sqrt{16}}$$

$$= 8 / 4 = 2$$

In the "t" table (table no. . 2.5.1) at .01 level, the value of "t" for 16 df is 2.92 and the obtained value of t = 2.00. which is less in comparison to the "t" value given in the able. Therefore, the obtained sample mean (58) is quite trustworthy and representing its Mean population by 99% confidence. There is only one chance out of 100, that sample mean is low or high.

Self Assessment Questions

1) What is the concept of Degree of Freedom (df)?

....

....

....

....

2) Why we consider the (df) which determining the reliability or trustworthiness of the statistics.

....

....

....

....

3) What is the difference to calculate the standard error of Mean of Large Size and Small Size samples.

....

....

....

....

2.7 APPLICATION OF THE STANDARD ERROR OF MEAN

2.7.1 Estimation of the Population Statistics – The Mpop

The wider use or application of $S.E._M$ is to estimate the population Statistical measurements i.e. the M pop. Here we are concerned only with the mean. Therefore we will discuss to estimate the mean of population on the basis of standard error of the mean obtained on either large size sample or small size sample.

Here, it is important to note that the estimation is always in range rather to point estimation because exact single value of any measurement is not possible. For example a student can not forecast with confidence that he will secure 85 to 95 marks out of 100 in statistics in the final examination.

Therefore, estimation of Mpop is always in range rather to point. Thus the limits obtained of Mpop (The lower and upper limits) are also known as Fiduciary limits.

The term Fiduciary limits was used by R.A. Fisher for the confidence interval of parameter and the confidence placed in the interval defined as Fiduciary Probability.

The simplest formula to estimate the Mpop is as under:

Mpop or M = M± 2.58 ó$_M$ ——— (at .01 level of significance)

Mpop or M= M±1.96 ó$_M$ ————(at .05 level of significance)

For more classification study the following examples carefully:-

Example 4: One language test was given to 400 boys of VIII class, the mean of their performance is 56 and the standard deviation is 14. What will be the Mean of the population of 99% level of confidence?

Solution

Given: N=400, M=56 and ó =14

To find out: Estimation of population mean at 99% level of confidence.

We know that Mpop at .01 level or at 99% confidence level is

Mpop .01 or M.01= M±2.58 ó$_M$

Where

M .01: Mean of the population at 99% confidence level of confidence

M : Mean of the sample

ó$_M$: Standard Error of the sample mean.

In the problem the values of Mean and N are known and the value of ó$_M$ is unknown. The value of ó$_M$ can be determined by using the formula.

$$ó_M = \frac{\sigma}{\sqrt{N}}$$

$$\therefore ó_M = \frac{14}{\sqrt{400}} = \frac{14}{\sqrt{20}}$$

Or ó$_M$ = 0.70

Thus $\overline{M}$.01 = 56± 2.58×0.70

= 56±1.806

= 54.194-57.806

Or $\overline{M}$.01 = 54-58

The Mean of the population at 99% level of confidence will be within the limits 54 to 58. In other words there are 99% chances that the Mean of the population lie within the range 54-58 scores. There is only 1% chance that mean of the population lie beyond this limit.

Example 5: A randomly selected group of 26 VI grade students having a weight of 35 kg and S.D = 10 kg. How well does this value estimate the average weight of all VI grade students at .99 and .95 level of confidence?

Solution

Given : N =26, M = 35kg and ó= 10 kg.

To find out : The fiduciary limits of the population mean at .05 and .01 levels.

In the problem the given sample size is below 30, Therefore to have the standard error of sample mean we will use the formula:

S_M = 2.0

And = N – 1 = 26 – 1

Or df = 25

i) Fiduciary limits of M at .01 level of confidence

By consulting the t table, level of confidence, the value of "t" for 25 df is 2.79

Thus, the Fiduciary limit of M at .01 or 99% level is

$= \bar{M} \pm 2.79\sigma_M$

$= 3.5\pm 2.79\times2.00$

$= 3.5\pm5.50$

$\therefore \bar{M}$ .01= 29.42-38.58 = (9.16)

ii) The fiduciary limits of $\bar{M}$ at .05 level of confidence

$M_{0.05} = M\pm 2.06\ \sigma_M$

$= 35\pm2.06\times2.0$

$= 35\pm4.12$

Or $M_{0.05}$ = 30.88–39.12

i) Thus The Fiduciary Limits of M .01 = 29.42-38.58

ii) The Fiduciary Limits of M.05 = 30.88-39.12

2.7.2 Determination of the Size of Sample

Standard error of the statistics is also used to estimate the sample size for test results. In order to learn how to use the standard error of the statistics you must go through the following examples.

Example 6

If the standard deviation of a certain population (ó) is 20. How many cases would require in a sample in order that standard error of the mean should not miss by 2.

Solution

Given: ó= 20, and S. $E._M$=20.

To Find Out : No of cases in the sample to be selected i.e. to determine the Size of Sample (N)

We know that

$$\text{S.EM} = \frac{\sigma}{\sqrt{N}}$$

$$\therefore\ 2 = \frac{20}{\sqrt{N}}$$

$$\text{Or } \sqrt{N} = \frac{20}{2} = 10$$

Or N = $(10)^2$

Or N = 100

If the standard error of the sample mean should not be more than 2 in such condition the maximum sample Size Should be 100 i.e. N=100 .

Example 7: The standard deviation of the intelligence scores of an adolescent population is 16. If the maximum acceptable standard error of the mean of the sample should not miss by 1.90, what should be the best sample size at 99% level of confidence?

Solution

Given :ó = 16, SE_M= 1.90

To find out : Sample size which represent its parent population upto the level of 99%.

We know that the Z value of 99% cases is 2.58 (From Z Table)

It means due to chance factors the sample mean would deviate from Mpop by 2.58 $ó_{M.}$ Further in keeping view the measurement and other uncontrolled factors, the measured error in the sample mean we would like to accept is 1.90.

Therefore the maximum error in the sample which we would like to select from the parent population is

$$S.E_M = ó \times \frac{2.58}{\sqrt{N}}$$

$$\text{Or } \sqrt{N} = \frac{\sigma \times 2.58}{SE_M}$$

$$\text{Or N} = \left(\frac{\sigma \times 2.58}{SE_M}\right)^2$$

$$\therefore N = \left(\frac{16 \times 2.58}{1.90}\right)^2$$

Or N = 472

To have a representative sample up to the level of 99% to the parent population, it is good to have a sample size more than 472 cases.

Self Assessment Questions

1) Given M =26.40, ó = 5.20 and N=100 compute

   The fiduciary limits of True Mean at 99% confidence interval

   The fiduciary limits of Population Mean at .95confidence interval.

   ....

   ....

   ....

   ....

2) The mean of 16 independent observations of a certain magnitude is 100 and S.D is 24. At .05 confidence level what are the fiduciary limits of the True Mean.

   ....

   ....

   ....

   ....

3) Suppose it is known that S.D of the scores in a certain population is 20. How many cases would we in a sample in order that the S.E of the sample mean be σ 2.

   ....

   ....

   ....

   ....

2.8 IMPORTANCE AND APPLICATION OF STANDARD ERROR OF MEAN

The Standard error of statistics has wide use in inferential statistics. It helps the experimenter or researcher in drawing concrete conclusions rather than abstract ones.

The various uses of standard error of the statistics are as under:

Various devices are used for determining the reliability of a sample taken from the large population. The reliability of the sample depends upon the reliability of the statistics, which is very easy to calculate.

The main focus of the standard error of statistics is to estimate the population parameters. No sampling device can ensure that the sample selected from a population may be representative. Thus the formula of the standard error of statistics provides us the limits of the parameters, which may remain in an interval of the prefixed confidence interval.

The method of estimating the population parameters the research work feasible, where the population is unknown as impossible to measure. It makes the research work economical from the point of crew of time. Energy and money.

Another application of the standard error of the statistics is to determine the size of the sample for experimental study or a survey study.

The last application of the standard error of statistics to determine the significance of difference of two groups to ascertained by eliminating the sampling or change by estimating the sampling or change errors.

2.9 THE SIGNIFICANCE OF THE DIFFERENCE BETWEEN TWO MEANS

Suppose we wish to study the linguistic ability of the two groups say boys and girls of age group 15 years.

First we have to select two large and representative samples from the two different populations. The population of the Boys of age group 15 years old and second, the population of girls of same age as of the Boys has has to be selected.

To administer the linguistic ability test Battery to both the groups of Boys and Girls selected as sample.

To compute the mean values of like scores obtained by the two groups on the linguistic ability test battery and find the difference between them.

After the above procedural steps, suppose that there is a difference in the means of the two groups and the difference is in favour of the girls of age group 15+ years old. Is this evidence sufficient to draw the conclusion that girls are superior in linguistic ability in comparison to the boys having same age level?

Probably the answer to this question may either be "yes" or "No" depending upon further testing of the difference of the means of two groups whether it is statistically significant or not. In other words, it is essential to test further, that how far the difference exists in the mean values of two groups is due to " chance" factor or is it "real and dependable"

This question involves the standard error of the difference that exists between the mean values of the two groups and the same as significant or not. Therefore, in order to test the significance of an obtained difference, we must first have a S.E. of the difference of sample means. Then from the difference between the sample means and the standard error of the difference of sample means ($S.E._{D.M.}$), we can determine whether a difference probably exists between the population means.

A difference is called *Significant* when the probability is high that it cannot be attributed to chance (i.e. by temporary and accidental factors or by sampling fluctuations) and hence represents a true difference between population means. And a difference is *non-significant* or chance, when it appears reasonably certain that it could be easily have arisen from sampling fluctuations and hence implies no real or true difference between the population means.

Thus the above discussion leads us to conclude that the significance of the difference between two sample means obtained from the two populations either independent or correlated depend upon two factors, viz.,

i) Standard Error of Difference between the two means, and

ii) The levels at which $S.E_{DM}$ is significant.

iii) Standard Error of the difference of two means (Σ_{DM}) and critical ratio (C.R)

2.9.1 Standard Error of the Difference of Two Means and Critical Ratio (CR)

Suppose we have two independent large populations, say A and B, and let us say that we have taken several numbers of samples (say two) from each population. Now if we compute the mean values of the scores of a trait of the two populations, we have 100 sample means obtained from the population A and 100 sample means obtained from the population B, and let us say that we find that there is a difference between the two sample means of population A and B. Thus in this way we have 100 differences of sample means. If we plot the frequency polygon of these hundred samples, certainly we will have a normal curve, and the distribution of the sample mean differences will be known as sampling *Distribution of Mean Differences.*

The standard error of the sample mean differences can be obtained by computing standard deviation of the sampling distribution of mean differences. This can be computed by using the formula:

$$S.E_M \text{ or } \Sigma_{DM} = \frac{\Sigma 1^2 + \Sigma 2^2}{N_1 + N_2} \text{ (In case of two independent population)}$$

Where

Σ_1 = Standard Deviation of the scores of a trait of the sample-1

Σ_2 = Standard Deviation of the scores of a trait of the sample-2

N_1= Number of cases in sample-1

N_2= Number of cases in sample-2

After having the standard error of the sample mean differences, the next step is to decide how far the particular sample mean difference is deviating from the two population mean differences ($M_1 \sim M_2$) on the normal probability curve scale. For the purpose we have to calculate Z score of the particular two sample mean differences, using the formula

$$Z = \frac{X–M}{\sigma_{DM}} \quad \text{(see unit-I)}$$

or

$$Z = \frac{(M_1 \sim M_2)–(M_1 \sim M_2)}{\frac{\sqrt{\Sigma 1^2 + \sigma^2}}{N_1+N_2}}$$

To distinguish the Z score of the difference of two sample means, the symbol C.R (Critical Ratio) is used. Therefore

$$C.R = \frac{(M_1 \sim M_2)–(M_1 \sim M_2)}{\Sigma_{DM}}$$

If the two independent populations are alike or same about a trait being measured, then

$M_1 \sim M_2 \quad = 0$

$$\therefore \text{C.R.} = \frac{(M_1 \sim M_2) - 0}{\sigma_{DM}}$$

$$\text{Or} \quad \text{C.R} = \frac{(M_1 \sim M_2)}{\sigma_{DM}}$$

This is the general formula to decide the significance of the difference exists in the two sample means taken from the two independent populations.

The formula of C.R. clearly indicates that it is a simple ratio between difference of the two sample means and the standard error of the sample mean differences. Further it is nothing but a Z score, which indicates how far the two sample mean difference is deviating from the two parent population mean difference, which is Zero.

2.9.2 Levels of Significance

Whether a difference in the two statistic i.e. the statistical measures obtained or the parameters are to be considered as statistically significant?

It depends upon the probability that the given difference could have arisen "by chance.".

It also depends upon the purposes of experiment, usually, a difference is marked "significant", when the gap between the two sample means points to or signifies a real difference between parameters of the population from which the sample are drawn.

The research workers as the experimenters have an option to choose several arbitrary standards called levels of significance of which the .05 and .01 levels are most often used. The confidence with which an experimenter or research worker, rejects or retains (accept) a null hypothesis, depends upon the level of significance adopted.

You carefully look at the table Z distribution presented in unit 1, table no. 1.6.1, you will find that at the point ±1.96 the total 95% cases fall.

If we take ±1.96 at the base line of normal distribution curve as two points, we find that total 95% area of the curve lie between these two points.

Remaining 5% area lies left on to the right side of the curve i.e. 2½ % area lies to the left and 2 ½ % area lies to the right side of the curve.

If the Z Score of the mean difference which is also known as "C. R. value of t ratio" is 1.96 or below, it means the difference of the two means lies within the acceptance area of the normal distribution of the sampling distribution of the area differences.

Hence the null hypotheses (Ho) are retained. "CR" or "t ratio" is higher than 1.96, means the mean difference falls within the area of rejection, hence the null hypotheses (Ho) is rejected.

Further, you see the table 1.6.1 again, you will find that at the point ±2.58ó on the base of the normal distribution curve total 99% area of the curve or cases lie within the range -2.58 ó to + 2.58ó.

Only 1% area of the curve or cases lie beyond these two units. If the CR value or "t ratio" is below the value of 2.58, i.e. within the area of acceptance of 99% level the obtained mean difference is significant at .01 level or 99% level.

If the CR value of t. ratio is obtained above to the 2.58% , the null hypothesis (Ho) said to be rejected at 99% level or .01 level of significance.

2.9.3 The Null Hypothesis

In the above paragraphs a term Null Hypothesis is used in relation to determining of the significance of difference between the two means. Before we proceed further, it is essential to know about the null hypothesis and its rate in determining the significance of the difference of "Zero difference" or "No difference" in the relative specific parameters of the population and symbolically it is denoted as Ho.

Hypothesis is a suggested or pre determined relation of a problem which is tested on the basis of the evidences collected. Null hypothesis is a useful tool in testing the significance of the difference.

The null hypothesis a states that there is no true difference between two population means, and that the difference found between sample means, therefore, are only by chance i.e. accidental and unimportant.

The null hypothesis is based on the simple logic that a man is innocent until he is proved guilty. It constitutes or brings a challenge before the experimenter to call the necessary evidences to reject or retain the null hypothesis which he has framed.

After rejection of the null hypothesis automatically the alternative hypotheses will he accepted, For Example: In a study of Linguistic ability of boys and girls of group 14+years, the researcher has framed the following two hypothesis-

Ho: There is no difference in the means of linguistic ability scores of male and female adolescents of age group 14+ years.

$H_{A:}$ The mean of the linguistic ability scores is in favour of the adolescents girls than the boys of age group 14+ -16 + years.

It is obvious that if the null hypothesis (Ho) is rejected on the basis of statistical treatment made on the related evidences collected, the alternative hypothesis (H_A) will be accepted. If the null hypothesis (Ho) is accepted on the basis of the evidences collected, in such condition the alternative hypothesis (H_A) will be rejected.

2.9.4 Basic Assumption of Testing of Significance difference between the Two Sample Means

The formula which is used to test the significance of the difference between the two sample means (see 2.9.1) is based on certain basic assumptions. The assumptions are as under:

1) The variable or the trait being measured or studied is normally distributed in the universe.

2) There is no difference in the Means of the two or more populations. i.e. My=Mz If there is a violation or deviation in the above assumptions in testing the significant difference in the two means, we can not use "C.R" or "t" test of significance. In such condition, there are other methods, which are used for the purpose.

3) The samples are drawn from the population using random method of sample selection.

4) The size of the sample drawn from the population is relatively large.

2.9.5 Two Tailed and One Tailed Test of Significance

Under the null hypothesis, difference between the obtained sample means from two populations i.e. M_1 and M_2 may be either plus or minus and is often in one direction.

In yet another case, the differences between true parameters may be difference of Zero i.e. M_1-M_2=0, or Mz_{DM}=0, so that in determining probabilities we consider both traits of the sampling distribution. This two tailed test, as it is sometimes called, is generally used when we wish to discover whether two groups have conceivably been drawn from the same population with respect to the trait being measured.

In many research studies or experiments the primary concern is with the direction of the difference rather than with its existence in absolute terms. This situation arises when we are not interested in negative differences or in the losses made as this has no importance practically. However, we are much interested in the positive directions i.e. gains or growth or developments. For example, suppose we want to study the effect of extrinsic motivation on solving the mathematical problems or in sentence construction, it is unlikely that extrinsic motivation leads to loss in either solving the mathematical problems correctly or framing the sentences correctly.

Thus here only we are interested the positive effect of motivation and we study both gain made by the learners in solving the mathematical problems or constructing the sentences correctly. In such condition we use one tail of the normal probability curve that is the positive side and the Z or ó values will be changed after 95% and 99% level of significance far. In case of large samples the Z or ó values for 95% level it becomes 12.33 ó in place of 2.586.

2.9.6 Uncorrelated (Independent) and Correlated (Dependant) Sample Means

When we are interested to test whether two groups differ significantly on a trait or characteristics measured, the two situations arises with respect to differences between means

1) Uncorrelated or Independent two sample means

2) Correlated or Dependent two sample means

The two sample means are uncorrelated or independent when computed from different samples selected by using random method of sample selection from one population or from different populations or from uncorrelated tests administered to the same sample.

The two sample means are correlated when a single group of population is tested in two different situations by using the same test. In other words when one test is used on a single group before the experiment and after the experiment or when the units of the Group or the population from which two sample are drawn are not mutually exclusive.

In the latter situation, the modified formula for calculating the standard error of the difference of two sample means is applied.

Thus to test the significance of sample means, there are always following four s situations:

1) Two Large Independent Sample. i.e. when N_1 and N_2 >30

2) Two Small Independent Samples i.e. when N_1 and N_2 <30

3) Two Large correlated samples.

4) Two Small correlated samples.

2.10 SIGNIFICANCE OF THE TWO LARGE INDEPENDENT OR UNCORRELATED SAMPLE MEANS

The formula for testing the significance of two large independent sample means is as follows:

$CR=M_1 \sim M_2$ where $ó_{DM} = \sqrt{\sigma M_1^2 + \sigma M_2^2}$ or

$= \sqrt{\sigma 1^2}$ / N1 + ó2$^{2\,/}$ N2

Self Assessment Questions

1) What do you mean by significance of the difference in two means?

....

....

....

....

2) Define Standard Error of the difference of the two sample means.

....

....

....

....

3) Define Sampling distribution of the differences of Means of two Samples.

....

....

....

....

4) What should be the mean value of sampling distribution of the difference of the means?

....

....

....

....

5) What indicates $S.E._{DM}$ or $\_\_\__{DM}$?

....

....

....

....

6) What do you mean by Ho, Define it.

....

....

....

....

7) What are the assumptions on which testing of the difference of two Mean is based?

....

....

....

....

8) What do you mean by One Tail Test and Two Tail Test? When these two tails are used?

....

....

....

....

9) What is meant by uncorrelated and correlated sample means?

....

....

....

....

Thus

$$CR = M_1 \sim M_2 \ / \ ó \ D$$

$$= \frac{M_1 \sim M_2}{\sqrt{\sigma 1^2 + \sigma 2^2}}$$

$$N_1 \qquad N_2$$

Where

CR : Critical Ratio

M_1 : Mean of the Sample or Group 1

M_2 : Mean of the Sample or Group 2

ó[1] : Standard Deviation of the Scores of sample 1

ó[2] : Standard Deviation of the Scores of sample 2

N_1 : Number of cases in Sample 1

N_2 : Number of cases in sample 2

Example 8: An Intelligence test was administered on the two groups of Boys and Girls. These two groups were drawn from the two populations independently by using random method of sample selection. After administration of the test, the following statistics was obtained

Groups	N	M	Σ
Boys	65	52	13
Girls	60	48	12

Determine the difference between the mean values of Boys and Girls significant?

Solution

In the given problem, the two samples are quite large and independent. Therefore, to test the significance difference in the mean values of Boys and Girls. First we have to determine the null hypothesis which is

$Ho=M_B=M_G$ i.e.

There is no significant difference in the mean value of the Boys and Girls and the two groups are taken from the same population

$$\text{C.R.} = \frac{M_1 \sim M_2 - 0}{\sigma D}$$

$$= \frac{(M_1 \sim M_2)}{\frac{\sqrt{\sigma 1^2}}{N_1} + \frac{\sqrt{\sigma 2^2}}{N_2}}$$

$$\text{C.R.} = \frac{(52-48)}{\frac{\sqrt{13^2}}{65} + \frac{\sqrt{12^2}}{60}} = \frac{4}{\frac{\sqrt{169}}{65} + \frac{\sqrt{144}}{60}} = \frac{4}{\sqrt{5}}$$

Or C.R. = 1.79

df = $(N_1\text{-}1) + (N_2\text{-}1)$

= (65-1) + (60-1)

= 123

To test the null hypothesis, which is framed, we will use two tail test. In the "t" distribution table (sub heading no. 2.5.2) at 123 df the "t" value at .05 level and .01 level is 1.98 and 2.62 respectively (The "t" table has 100 and 125 df, but df 123 is not given, therefore nearest of 123 i.e. 125df is considered). The obtained t value (1.79) is much less than these two values, hence it is not significant and null hypothesis is accepted at any level of significance.

Interpretation of the Results

Since our null hypothesis is retained, we can say that Boys and Girls do not differ significantly in their level of intelligence. Whatever difference is observed in the obtained mean values of two samples is due to chance factors and sampling fluctuations. Thus we can say with 99% level of confidence that no sex difference exists in the intelligence level of the population.

2.11 SIGNIFICANCE OF THE TWO SMALL INDEPENDENT ON UNCORRELATED SAMPLE MEANS

When the N's of two independent samples are small (less than 30), the S.E. $_{DM}$ or -ó$_{DM}$ (standard error of the difference between two means) should depend upon the Standard Deviation (S.D. or ó) values computed by using the correction

The formula used to test the significance of the difference of two means of small independent samples is :

$$t = \frac{M_1 \sim M_2}{S.E_{\ DM}}$$

Where

$$S.E_{\ DM} = S.D \ \frac{\sqrt{N_1 + N_2}}{N_1 \times N_2}$$

$$\text{And S.D} = \frac{\sqrt{\Sigma(x1\text{-}M1)^2 + \Sigma(x2\text{-}M2)^2}}{(N_1 - 1)(N_2 - 1)}$$

For simplification the above formula can also be written as

$$t = \frac{M1 \sim M2}{\frac{\Sigma d2 + \Sigma d2^2}{N1+N2-2} \times \frac{N1+N2}{N1N2}} \quad \text{— (i)}$$

Where

$D_1 = (S_1\text{-}M_1)$, and

$D_2 = (x_2\text{-}M_2)$

Here, X_1 and X_2 are the new scores of two groups, M_1 and M_2 are given in relation to the two samples or groups having the small number units or cases.

When the raw data are not given and we have statistics or the estimates of two small size sample, in such condition, we use the formula-

The corresponding Mean values of the scores of the two groups N_1 and N_2 are the number of the units or cases in the two groups t is also a critical ratio in which more exact estimate of the $ó_{DM}$ is used. Here 't' in place C.R. is used because sampling distribution of "t" is not normal when N is small i.e. <30, "t" is a critical ratio (C.R.), but all C.R's are not "t's.

$$t = \frac{M_1 \sim M_2}{\sqrt{\sigma 1^2 \left(N_1 - 1\right) \sigma 2^2 \left(N_2 - 1\right) \times \left(N_1 + N_2\right) / N_1 N_2}} \qquad \text{.......(ii)}$$

Where

M_1 = Mean of the scores of sample -1

M_2 = Mean of the scores of sample -2

$ó^1$ = Standard Deviation of the scores of sample-1

$ó^2$ = Standard Deviation of the scores of sample-2

N_1 = Number of units or cases on the sample-1

N_2 = Number of units or cases in the sample-2

For more clarification study the following examples very carefully.

Example 9: An attitude test regarding a vocational course was given to 10 urban boys and 5 rural boys. The obtained scores are as under-

Urban Boys (x1) = 6, 7, 8, 10, 15, 16, 9, 10, 0, 9

Rural Boys (x2) = 4, 3, 2, 1, 5

Determine at .05 level of significance that its there a significant difference in the attitude of boys belonging to rural and urban areas in relation to a vocational course?

Solution

Ho=b1=b2 : H1=b1 $\neq$ b2

Level of significance = .05

For acceptance or rejection of null hypothesis at .05 level of significance, the two tail test is used.

Thus

Urban Boys | **Rural Boys**

X1	d1=(x1-m1)	d12
6	-4	16
7	-3	9
8	-2	4
10	0	0
15	+5	25
16	+6	36
9	-1	1
10	0	0
10	0	0
9	-1	1
$\sum x_1=100$		$\sum d1^2=92$

X2	d2=(X2-M2)	d22
4	+1	1
3	0	0
2	-1	1
1	-2	4
5	+2	4
$\sum x_2=15$		$\sum d_2^2=10$

M= 15/5
M= 3

M= $\sum$x/N
= 100/10
M=10

We know that

$$t = \frac{M_1 \sim M_2}{\sqrt{\frac{\Sigma d1^2 + \Sigma d2^2}{N_1+N_2-2} \times \frac{N_1+N_2}{N_1N_2}}}$$

$$= \frac{10-3}{\sqrt{\frac{92+10}{10+5-2} \times \frac{10+5}{10\times5}}}$$

$$= \frac{7}{\sqrt{7.8\times0.30}} = \frac{7}{\sqrt{2.34}} = \frac{7}{1.54}$$

Or t = 4.46

df = $(N_1-1)+(N_2-1)$

= 9+4

=13

In "t" distribution table (table 2.5.1), the t value for 13 df at .05 level is 2.16. The obtained t value 4.46 is much greater than this value. Hence null hypothesis is rejected.

Interpretation of the Result

Our null hypothesis is rejected at .05 level of significance for 13 df. Thus we can say that in 95% cases significant difference in the attitude of the urban and rural boys regarding a vocational course. There are only 5% chances out of 100 that the two groups have same attitude towards a vocational course.

Example 10: music interest test was administered on 15 + years did boys and girls sample taken independently from the two populations. The following statistics was obtained:

	Mean	S.D.	N
Girls	40.39	8.69	30
Boys	35.81	8.33	25

Is the mean difference is in favour of girls?

Solution:

$Ho = b_1 = b_2$

$H_1 = b_1 \neq b_2$

In the given problem, the row scores of the two groups are not given. Therefore we will use the following formula for testing of the difference of means of two uncorrelated sample means:

$$t = \frac{M_1 \sim M_2}{\sqrt{\frac{\Sigma 1^2 (N_1 - 1) + \Sigma 2^2 (N_2 - 1)}{N_1 + N_2 - 2} \times \frac{N_1 + N_2}{N_1 \times N_2}}}$$

$$t = \frac{40.39 – 35.81}{\sqrt{\frac{(8.69)^2 (30-1) + (8.33)^2 (25-1)}{30+25–2} \times \frac{30+25}{30 \times 25}}}$$

$$= \frac{4.58}{\sqrt{75.516 \times 29 + 69.389 \times 24 \times 55}}$$

$$= \frac{4.58}{\sqrt{7274 \times .073}} = \frac{4.58}{2.309}$$

Or t $= 1.98$

d f= $(N_1 – 1) + (N_2 – 1) = 53$

In the t distribution table for 53 df the t value at .05 level is 2.01. Our calculated t value 1.98 is less than this value. Therefore, the null hypothesis is retained.

Interpretation of the Results

Since our null hypothesis is accepted at .05 level of significance. Therefore it can be said that in 95 cases out of 100, there is no significant difference in the mean values of boys and girls regarding their interest in music. There are only 5% chances that the two groups do not have equal interest in music. Hence with 95% confidence, we can say that both boys and girls have equal interest in music. Whatever difference is deserved in the mean values of the groups is by chance or due to sampling of fluctuations.

2.12 SIGNIFICANCE OF THE TWO LARGE CORRELATED SAMPLES

In some of the experimental studies a single group is tested in two different conditions and the observations are in pairs. Or two groups are used in the experimental condition, but they are matched by using pairs method. In these conditions, a modified formula of standard error of the difference of means is used. Therefore the formula for testing of the difference of two means of large correlated samples is –

$$t = \frac{M_1 \sim M_2}{\sqrt{\sigma M_1^{\ 2} + \sigma M_2^{\ 2} - 2r_{12}\sigma M_1\ \sigma M_2}}$$

In the formula

M_1 = Mean of the scores of sample -1

M_1 = Mean of the scores of sample -2

óM_1 = Standard Error of the Mean of sample-1

i.e. ó$M_1 = \frac{\Sigma 1}{\sqrt{N_1}}$

ó$M_2 = \frac{\Sigma 2}{\sqrt{N_1}}$

and r1,2 = correlation between two sets of scores.

For more classification go through the following examples

Example 11: An Intelligence test was administered on a group of 400 students twice after an interval of 2 months. The data obtained are as under-

		M	S.D
Testing –I	:	25	8
Testing-II	:	30	5
N	:	400	
r_{12}	:	65	

Test if there is a significant difference in the means of intelligence scores obtained in two testing conditions.

Solution:

Ho => b1=b2 and H1=>b1 $\neq$ b2

$$\therefore\ t = \frac{M_1 \sim M_2}{\sqrt{\sigma M_1^{\ 2} + \sigma M_2^{\ 2} - 2r_{12}\sigma M_1\ \sigma M_2}}$$

According the formula all values are given except S.E of means ($\sum$ M). Therefore first we have to calculate standard errors of the means of the two sets of scores

$$\therefore\ \text{ó}M_1 = \frac{\Sigma 1^2}{\sqrt{N_1}} = \frac{8^2}{\sqrt{400}} = \frac{64}{20}$$

Or $óM_1 = 3.20$

Similarly

$$óM_2 = \frac{\sigma_2^{\ 2}}{\sqrt{N_2}} = \frac{5^2}{\sqrt{400}} = \frac{25}{20}$$

Or $óM_2 = 1.25$

Thus

$$t = \frac{30–25}{\sqrt{(3.20)^2 + (1.25)^2 - 2 \times .65 \times 3.20 \times 1.25}}$$

$$= \frac{5}{\sqrt{10.24 + 1.5625 - 5.20}}$$

$$= \frac{5}{\sqrt{6.6025}} = \frac{5}{2.57}$$

$t = 1.95$

df = N-1 = 400-1

df = N-1= 400-1 (In the example N is same i.e. the single group is tested in two different time intervals)

a df = 399

According to "t" distribution table (Table no-2.5.1) the value of t for 399 df at .01 level is 2.59. Our calculated value of t is 1.95, which is smaller than the value of t given in "t" distribution table. Hence the obtained t value is not significant even at .05 level. Therefore our null hypothesis is retained at .01 level of significance.

Interpretation of the Results

Since the obtained t value is found insignificant level for 399 df; thus the difference in the mean values of the intelligence scores of a group, tested after an interval of two months is not significant in 99 conditions out of 100, there is only 1% hence that the difference in two means is significant at .01 level.

Example 12: In a vocational training course an achievement test was administered on 64 students at the time of admission. After training of one year the same achievement test was administered. The results obtained are as under:

		M	ó
Before Training	:	52.50	7.25
After Training	:	58.70	5.30

Is the gain, after training significant?

Solution:

$H_o = b_1 = b_2$ (The gain after training is insignificant)

$H_1 = b_1 \neq b_2$

(*Note:* Read the problem carefully, here we will use one tail test rather to use two tail test. Because here we are interested in gain due to training, not in the loss. That is we are interested in one side of the B.P.C which is +ve side. 99% confidence and .05 for 95% confidence. See the table no-2.5.1 carefully and read the footnote)

We know that formula of testing the difference between two large correlated means is–

$$\text{Formula } t = \frac{M_1 \sim M_2}{\sqrt{\sigma M_1^2 + \sigma M_2^2 - 2r_{12}\sigma M_1 \; \sigma M_2}}$$

Where

$$\sigma M_1 = \frac{\sigma 1}{\sqrt{N}} = \frac{7.25}{\sqrt{100}} = \frac{7.25}{10}$$

Or σM_1 = .725

$$\text{And } \sigma M_2 \quad \frac{\sigma 2}{\sqrt{N}} = \frac{530}{\sqrt{N}} = \frac{5.30}{\sqrt{100}} = \frac{5.30}{10} = .53$$

Or σM_2 = .53

$$t = \frac{58.70\text{–}52.50}{\sqrt{(.725)^2 + (.53)^2 - 2 \times .50 \times .725 \times .53}}$$

$$= \frac{6.2}{\sqrt{0.4223}} = \frac{6.2}{.65}$$

t = 9.54

df = (100-1)

= 99

In the‘t’ distribution table (table No. 2.5.1) at .02 level the t value for 99 df is 2.36 and out obtained t value is 9.54, which is much greater than the “t” value of the table. Thus the obtained t value is significant at 99% level of significance. Therefore our null hypothesis is rejected.

Interpretation of the Results

Since the obtained “t” value is found significant at .02 level for 99df. Thus we can say that gain on the achievement test made by the students after training is highly significant. Therefore we can say with 99% confidence that given vocational training is quite effective. There is only 1 chance out of hundred, the vocational training is ineffective.

2.13 SIGNIFICANCE OF TWO SMALL CORRELATED MEANS

In case of determining the significance difference between the two correlated small sample mean we have two methods, which are as under

i) **Direct Method:** i.e. we have to calculate Mean Values standard deviation values of the two groups and the coefficient of correlation (r_{12}) between the scores of two groups. In such condition the formula used to test significant difference in the two means is –

$$t = \frac{M_1 \sim M_2}{\frac{\sqrt{\sigma 1^2}}{N-1} + \frac{\sqrt{\sigma 2^2}}{N-1} - 2r_{12}\frac{\sigma 1 \sigma 2}{N-1}}$$

$$t = \frac{M_1 \sim M_2}{\sqrt{Sm_1{}^2 + Sm_2{}^2 - 2r_{12}Sm_1Sm_2}}$$

Where

$$Sm_1 = \frac{\sigma 1}{\sqrt{N-1}}$$ (standard error of the small sample mean)

$$Sm_2 = \frac{\sigma 2}{\sqrt{N-2}}$$

ii) **Difference Method:** In this method we have the raw data of two small groups or sample and we are not calculate coefficient of correlation (r_{12}) between the two sets of scores.

Examples 13: A pre test and past test are given to 12 subjects. The scores obtained are as under–

S. No.-	1	2	3	4	5	6	7	8	9	10	11	12
Pre-Test:	42	50	51	26	35	42	60	41	70	38	62	55
Past-Test:	40	62	61	35	30	52	68	51	84	50	72	63

Determine if the gain on past test score significant?

Solution:

S.No. of Subjects	Post Test X_1	Pre Test X_2	Difference (X_2-X_1)	D-MD d	d^2
1	40	42	-2	-10	100
2	62	50	12	4	16
3	61	51	10	2	4
4	35	26	9	1	1
5	30	35	-5	-13	169
6	52	42	10	2	4
7	68	60	8	0	0
8	51	41	10	2	4
9	24	70	14	6	36
10	50	38	12	4	16
11	72	62	10	2	4
12	63	55	8	0	0
			$\sum D = 96$		$\sum d^2 = 354$

$$MD = \frac{\sum D}{N}$$

$$SD = \sqrt{\frac{\sum d^2}{N-1}}$$

$$= \sqrt{\frac{354}{14}}$$

Or SD= 5.67

$$\therefore \quad SE_{DM} = \frac{\sigma D}{\sqrt{N}}$$

Where

SE_{DM} = Standard Error of the Mean of Differences.

ΣD = Standard Deviation of the Differences

And N = Total No. of cases.

$$\text{Thus } SE_{DM} = \frac{5.67}{\sqrt{12}} = \frac{5.67}{3.464}$$

$$= 1.631$$

$$\therefore\ t = \frac{MD}{SE_{DM}} = \frac{8}{1.637} = 4.88,\ df = 11$$

In the "t" distribution table (Table 2.5.1 subheading 2.5.2) for 11 df at .02 level the value is 2.72 and our calculated value of t (4.88) is much greater than the table value. Therefore the null hypothesis is rejected at .01 level of significance.

Interpretation of the Results

Since our null hypothesis is rejected at.01 level of significance, therefore we can say that the gain made by the subject on past test is real in 99 cases out of 100. There are only 1% chance that the gain shown by the subjects in cue to chance factors as by sampling fluctuations.

2.14 POINTS TO BE REMEMBERED WHILE TESTING THE SIGNIFICANCE IN TWO MEANS

When you compare the means of two groups or to compare the means of a single group tested in two different situations or conditions and to know, whether the difference found in the two means is real and significant, or it is due to chance, factors, you should keep in mind the following steps as a process of testing the difference between two means.

i) Set up null hypothesis (Ho) and the alternative hypothesis (H1), according to the requirements of the problem.

ii) Decide about the level of significance for the test, usually in behavioural or social science, .05 and .01 levels are taken into consideration for acceptance or rejection of the null hypothesis.

iii) Decide whether one tailed or two tailed test of significance for independent or the correlated means.

iv) Decide whether the large or samples are involved in the problem or in the experiment.

v) Calculate either C.R value or "t" ratio value as per nature and size of the samples, by using the formulas discussed on the precious pages.

vi) Calculate degree of freedom (df). It should be N1+N2-2 for independent t or uncorrelated samples. While in case of correlated samples it should be N-1.

vii) Consult the "t" distribution table with df keeping in mind the level of significance, which we have decided at step number-11.

viii) Compare the calculated value of "t" with the "t" value given in the table with respect to df and level of significance.

ix) Interpret the Results:

If null hypothesis (Ho) is rejected, there is a significant difference between the two means.

If null hypothesis is accepted, there is no significant difference in the two means. Whatever the difference exists it has arisen due to sampling fluctuations or chance factors only.

Self Assessment Questions

1) How you can define "Critical Ratio" and "t" Ratio?

....

....

....

....

2) What is the difference between "CR" and "t"?

....

....

....

....

3) What is the difference between C.R. and Z Score?

....

....

....

....

4) How can you define the standard Error of the Difference of Means?

....

....

....

....

5) What formula you will use in the following conditions:

a) When two independent large samples are given.

....

....

....

b) When two correlated large samples are given.

....

....

....

c) When two independent small samples are given.

....

....

....

d) When two small independent samples are given.

....

....

....

6) What do you mean by independent samples?

....

....

....

....

7) What do you mean by correlated samples?

....

....

....

....

2.15 ERRORS IN THE INTERPRETATION OF THE RESULTS, WHILE TESTING THE SIGNIFICANT DIFFERENCE BETWEEN TWO MEANS

While interpreting the results obtained from the test of significance of a single mean or the difference between two means, we should take case and not to depend much on the statistical results obtained. Generally while interpreting the results, we may make following two type of Error.

Type-I Error or á Errors:

This Error occurs when the null hypothesis is true, but we reject the same by marking significant difference between the two means.

Type-II Error or â Error:

This error occurs when the null hypothesis is wrong or to be rejected while the same is retained.

This probability of occurrence of type-II error or â error due to finding high level of significance i.e. above to the .01 level of significance which may be .001 or the above.

2.16 LET US SUM UP

The Standard error of the estimates or statistics or sample statistical measures ($S.E._M$) consists –

Error of Sampling, and

Error of Measurement

Fluctuations from sample to sample, the so called sampling error or errors of sampling are not to be thought of as mistakes, features, and the like, but as variations arising from the fact that no two samples are over exactly alike.

Mean values and standard deviations (Σ's) obtained from random samples taken from a population are estimates of their parameters (the true statistical measurements of the population) and the standard error ($S.E_M$) measures the tepidness of these estimates.

If the $S.E_M$ is large, it does not follow necessarily that the obtained statistics is effected by a large sampling error, as much of the variations may be due to error of measurements, when error of measurements are low i.e. the measuring tools or tests are highly reliable, a large $S.E_M$ indicates considerable sampling error.

In the comparative studies or the experiments it is to decide whether the obtained differences of such magnitude is attributed to chance factor or sampling variations or it really exists? For such decisions the standard error of the difference is considered.

The critical or "t" ratio are nothing, but these are the Z scores , which tells how far the sample mean difference derivates to the population mean difference on a normal distribution curve.

"C.R" and "t" are the ratio of the mean difference of the two groups and the standard error of the mean differences.

While deciding the significance of any statistical measure or the difference on the means of two samples or two populations, the degree of freedom and levels of confidence are considered, and in the light of these two we either accept the null hypothesis or to reject the same.

While taking the decision a care is to be taken, so that type-I and type-II error should not be occur.

2.17 UNIT END QUESTIONS

1) Explain the term "Statistical Inference". How is the statistical inference is based upon the estimation of parameters.

2) Indicate the role of standard error for statistical generalisation.

3) Differentiate between significance of statistics and confidence interval of fiduciary limits.

4) Enumerate the various uses of Standard Error of the statistics.

5) What type of errors can occur while interpreting the results based on test of significance? How we can overcome these errors?

6) A Sample of 100 students with mean score 26.40 and SD 5.20 selected randomly from a population. Determine the .95 and .99 for confidence intervals for population true mean.

7) A small sample of 10 cases with mean score 175-50 and $\sum$=5.82 selected randomly. Compute finding limits of parameter mean at .05 and .01 level of confidence.

8) The mean and standard deviation of the intelligence scores obtained on a group of 200 randomly selected students are 102 and 10.20 respectively. How dependable is mean I.Q. of the students?

The following are the data for two independent samples :

	N	M	S.D.
Boys	60	48.50	10.70
Girls	70	53.60	15.40

Is the difference in the mean values of Boys and Girls significant.

A reasoning ability test was given to 8 urban and 6 rural girls of the same Class. The data obtained are differ significantly in there reasoning ability.

Groups	Scores
Urban Girls	16,9,4,23,19,10,5,2
Rural Girls	20,5,1,16,2,4.

The observations given below obtained on 10 subjects in a experiment of Pre and Post test. Is gain trade by the students on post test significant?

Subjects	1	2	3	4	5	6	7	8	9	10
Scores on Pre Test	5	15	9	11	4	9	8	13	6	16
Scores on Post Test	7	9	4	15	6	13	9	5	6	12

9) A group of 10 students was given 5 trials on a test of physical efficiency. Their score on the I and V trials are given below. Test whether there is a significant effect of practice on the improvement made in first to fifth trial.

Subject	A	B	C	D	E	F	G	H	I	J
Trial I	15	16	17	20	25	30	17	18	10	12
Trial V	20	22	22	25	35	30	21	23	17	20

10) A group of 35 students randomly selected was tested before and after experimental treatment. The observations obtained are as under:

	M	ó
Pre Test	15.5	5.2
Post Test	21.6	4.8

Coefficient of Correlation between The scores of Pre and Post Test	0.70

Find out the groups is different significantly on the two testing conditions.

2.18 POINTS FOR DISCUSSION

1) What will happen on the standard error of sample mean if
   a) Sample is homogeneous and large
   b) Sample is heterogeneous and large
   c) Sample is heterogeneous and small
   d) Sample is homogeneous as well as small
2) Differentiate between "t" distribution and Z distribution. What is the basic difference between "t" Test and Z Test.
3) When the "t" distribution and "Z" distribution become conside
4) The necessity of a theoretical distribution model for estimation.

2.19 SUGGESTED READINGS

Aggarwal Y.P. (1990) *Statistical Methods – Concepts Application and Computation*. Sterling Publications Pvt Ltd. New Delhi.

Walker . H.M. and Lev. J. (1965). *Statistical Inference.* Oxford and I B H Publishing Co. Calcutta.