1. Introduction

Welcome to the series of e-learning modules on Standard Error and Estimation of Standard Errors. In this module we are going cover the basic concept of standard error and the results related to the estimation of standard errors of mean, total and proportions.

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

- Standard Errors
- Variance of sample mean and total under Simple Random Sampling With Replacement and Standard Errors
- Variance of sample mean and total under Simple Random Sampling Without Replacement and Standard Errors
- Standard Errors for proportions

In most statistical problems, there is no estimator that is guaranteed to give the right answer because the value of the estimator typically depends on the sample.The standard error measures the long-run average spread of estimated values in the same hypothetical scenario.

Both bias and standard errors contribute to the average size of the error of an estimator. If the bias is large, on an average the estimator overshoots or undershoots the truth by a large amount. If the standard error is large, the estimator typically is far from truth, even if its average is close to truth.

The formulae for the standard errors of the estimated population mean and total are used primarily for three purposes:
One, to compare the precision obtained by Simple Random Sampling With that given by other methods of sampling.
Two, to estimate the size of the sample needed in a survey that is being planned, and
Three, to estimate the precision actually attained in a survey that has been completed.

Standard error can be defined as the standard deviation of a given distribution.
For example: Suppose ‘t’ is an estimator, then, standard error of ‘t’ equals to square root of variance of t.

2. Theorems on Simple Random Sampling With Replacement

We shall now look at some theorems on Simple Random Sampling With Replacement. Here, theorem 1 state that, under Simple Random Sampling With Replacement, variance of sample mean is given by sigma square by n.

Proof:

Let y one, y two, etc yn be a Simple Random Sampling With Replacement sample of size ‘n’ drawn from a population of size N. We know that population mean Y bar equals to summation Yi by N

And sample mean y bar is equal to summation yi by n

Also we have population variance sigma square is equal to summation Yi minus Y bar whole square divided by N,

which is equal to summation Yi square by N minus Y bar square and the expected value of y bar is equal to population mean Y bar.

Now, consider:

Variance of sample mean y bar is equal to variance of summation yi by n which is equal to variance of summation yi by n square

Is equal to one by n square into summation of variance of yi plus summation i not equal to j covariance of yi yj.

Here, covariance of yi, yj is equal to zero, since the y1 units selected at i^{th} draw is independent of yj units selected at the j^{th} draw.

Therefore, variance of y bar is equal to one by n square into summation variance yi. Call this equation as one.

Variance of yi is equal to expected value of yi square minus Expected value of yi whole square.

But, the expected value of yi square is equal to summation Yi square by N which is equal to summation Yi square by N.

Expected value of yi equals to summation Yi by N equal to the population mean Y bar.

Therefore, variance of yi is equal to summation Yi square by N minus Y bar square which is equal to sigma square.

By substituting the values in equation 1, we get, Variance of y bar is equal to one by n square summation variance of yi,

Which is equal to one by n square summation sigma square

Equal to n sigma square by n square, this is equal to sigma square by n.

Theorem 2: Under Simple Random Sampling With Replacement, sample mean square is an unbiased estimator of sigma square, that is, to show that expected value of s square is equal to sigma square.

Proof: We know that:

Sigma square is equal to summation Yi minus Y bar whole square divided by N which is equal to summation Yi square by N minus Y bar square.
Which implies summation Yi square by N is equal to sigma square plus Y bar square. Call this as star.
S square is equal to summation yi minus y bar whole square by n minus 1
Is equal to one by n minus 1 into summation yi square minus n into y bar square.
Consider expected value of s square. Equal to expected value of summation yi minus y bar whole square by n-1
Which is equal to expected value of summation yi square minus n ybar square divided by n minus one
Which is equal to summation expected value of yi square minus n expected value of ybar square divided by n minus one . Call this as 1.
Expected value of yi square is equal to summation Y i square by N which is equal to sigma square plus Y bar square. Call this as 2.
Expected value of y bar square is equal to variance of y bar plus expected value of y bar whole square which is equal to sigma square by n plus Y bar square. Call this as 3.

Substituting equations <u>2</u> and <u>3</u> in equation <u>1</u>, we get,
Expected value of s square is equal to summation of sigma square plus Y bar square minus n into sigma square by n plus Y bar square, divided by n minus one
Which is equal to n into sigma square plus Y bar square minus n into sigma square by n plus Y bar square divided by n minus one
which is equal to n minus one into sigma square by n minus one, which is equal to sigma square.

Next, Theorem 3 states that, the variance of population total under Simple Random Sampling with Replacement is given by N square sigma square by n. That is, variance of Y cap is equal to N square sigma square by n.

Theorem 4 states that:
An unbiased estimator of variance of estimated population total is given by N square s square by n .That is, the estimate of variance of Y cap is equal to N square into s square by n.

<u>Standard Errors of mean and total under Simple Random Sampling With Replacement</u>
Standard Error can be defined as the Standard Deviation of the distribution i.e. square root of the variance of the distribution

a) The standard error of sample mean is equal to square root of variance of y bar which is equal to square root of sigma square by n
b) An estimate of standard error of sample mean is equal to estimate of the square root of variance of the sample mean which is equal to square root of s square by n

c) A standard error of the estimated population total is equal to square root of variance of estimated population total, which is equal to square root of N square into sigma square by n.
d) An estimate of standard error of estimated population total is equal to estimate of square root of variance of estimated population total, which is equal to square root of N square into s square by n.

3. Theorems on Simple Random Sampling Without Replacement (Part 1)

Theorem 5: Under Simple Random Sampling Without Replacement variance of the unbiased estimator of population mean is given by variance of y bar is equal to N minus n into sigma square divided by n into N minus one or variance of y bar is equal to N minus n into S square divided by n into N.

Proof:

We have the population variance sigma square is equal to summation Yi minus Y bar whole

square divided by N which is equal to summation Yi square by N minus Y bar square and

S square is equal to summation Yi minus Y bar whole square divided by N minus one.

Which implies N minus one into S square is equal to N into sigma square

Which implies sigma square is equal to N minus one into S square divided by N.
Consider variance of ybar is equal to variance of summation yi by n which is equal to variance of summation yi by n square.

This is equal to one by n square into summation of variance of yi plus summation i not equal to j covariance of yi yj. Call this as (1)
Variance of yi is equal to expected value of yi square minus expected value of yi whole square
But, expected value of yi square is equal to summation Yi square by N.

Therefore, the variance of yi is equal to summation Yi square by N minus Y bar square which is equal to sigma square. Call this as 2.
Covariance of yi yj is equal to expected value of yi yj minus expected value of yi into expected value of yj
Expected value of yi is equal to summation Yi by N which is equal to population mean Y bar and similarly expected value of yj is equal to Y bar

Since yi and yj takes values Yi and Yj with probability one by N each
Covariance of yi yj is equal to expected value of yi yj minus Y bar square. Call this as 3.

Expected value of yi yj is equal to summation YiYj into probability of yiyj, which is equal to

summation YiYj by N into N minus one. Call this as four.
We know that summation Yi minus Ybar is equal to zero

Which implies summation Yi minus Ybar whole square is equal to zero
Which implies summation Yi whole square plus N square Ybar square 2 N Y bar into summation Yi is equal to zero.

Which implies summation Yi whole square plus N square Ybar square 2 N Y bar into N into Y bar is equal to zero
Which implies summation Yi whole square plus N square Ybar square two N square Y bar square is equal to zero
Summation Yi square plus Double summation YiYj minus N square Y bar square is equal to zero
Which implies double summation YiYj is equal to N square Y bar square minus summation yi square
Which is equal to N square Y bar square minus N into sigma square plus N into Y bar square, which is equal to N square Y bar square minus N into sigma square plus Y bar square. Call this as five.

By substituting equation 5 in equation 4 we get
Expected value of yi yj is equal to summation YiYj by N into N minus one
Which is equal to N square Ybar square minus N into sigma square plus Ybar square by N into N minus one,
which is equal to N square Ybar square minus sigma square minus Ybar square by N minus one. Call this as six.
By substituting equation 6 in equation 3 we get,
Covariance of yi yj is equal to N square Ybar square minus sigma square minus Ybar square by N minus one minus Y bar square,

which is equal to minus sigma square by N minus one, and let us call this equation 7.

By substituting equations 2 and 7 in equation 1

Variance of y bar is equal to summation sigma square plus summation minus sigma square

by N minus one whole term divided by n square

Which is equal to n sigma square minus n into n minus one sigma square by N minus one

whole term divided by n square

Which is equal to sigma square minus n minus one sigma square by N minus one whole

term divided by n

Which is equal to sigma square into N minus n by N minus one into n. Call this as star.
We know that,
Sigma square is equal to N minus one into S square by N
Hence, variance of y bar is equal to N minus one into S square into N minus n divided by N into n into N minus one
Which is equal to N minus n into s square by N into n.

4. Theorems on Simple Random Sampling Without Replacement (Part 2)

Theorem 6:
A sample mean square s^2 is an unbiased estimator of population mean square S^2 under Simple Random Sampling Without Replacement.
That is expected value of s square is S square.

Proof:
s square is equal to summation yi minus y bar whole square divided by n minus one and
S square is equal to summation Yi minus Y bar whole square divided by N minus one
Consider Expected value of s square is equal to expected value of summation yi minus ybar whole square divided by n minus one
Which is equal to one by n minus one into expected value of summation yi square minus n into ybar square
Which is equal to one by n minus one into summation expected value of yi square minus n into expected value of ybar square.
Call this equation as one.

But Expected value of yi square is equal to summation Yi square by N.
Which is equal to sigma square plus population mean Y bar square. Call this as two.
Expected value of ybar square is equal to variance of y bar plus expected value of y bar whole square which is equal to

N minus n into sigma square by N minus one into n, plus Y bar square. Call this as three.
Substituting equations <u>2</u> and <u>3</u> in equation <u>1</u>, we get:
Expected value of s square is equal to
Summation sigma square plus Y bar square minus n into N minus n by N minus one into sigma square by n plus Y bar square by n minus one
Which is equal to n into sigma square plus Y bar square minus N minus n by N minus one into sigma square minus n into Y bar square by n minus one.

Which is equal to n into sigma square plus n into Y bar square minus N minus n by N minus one into sigma square minus n into Y bar square by n minus one
Which is equal to sigma square by n minus one whole multiplied by n minus N minus n by N minus one
is equal to sigma square by n minus one whole multiplied by n into N minus one minus N plus n whole divided by N minus one
equals to sigma square by n minus one whole multiplied by N into n minus one whole divided by N minus one
equals to N into sigma square by N minus one
which is equal to S square.

Theorem 7: An estimate of the variance of the sample mean under Simple Random Sampling Without Replacement is given by:
Estimate of variance of y bar is equal to N minus n into s square divided by N into n.

Theorem 8: The variance of estimated population total under Simple Random Sampling Without Replacement is given by N minus n into N square sigma square by n into N minus one
Which is equal to N minus n into N square S square by N into n.

Theorem 9: An unbiased estimator of variance of estimated population total is given by: Cap of variance of Y cap is equal to N square into N minus n into s square by N into n.

Sampling for Attributes.
We know that, S square is equal to N into P into Q divided by N minus one.
Similarly s square is equal to n into p into q divided by n minus one.

5. Theorems on Simple Random Sampling Without Replacement (Part 3)

Theorem 10:
Under Simple Random Sampling Without Replacement variance of sample proportion p is equal to N minus n by N minus one into PQ by n.

<u>Proof:</u>
Already we proved in the last topic that p is equal to sample mean y bar.
Variance of p is equal to variance of y bar which is equal to N minus n into S square divided by N into n
Which is equal to N minus n multiplied by N into P into Q by N into n multiplied by N minus one
Which is equal to N minus n into P into Q by N minus one into n.

Theorem 11:
Estimate of Variance of p is equal to N minus n by n minus one into pq by N.

By making use of already proven result that E of s square equals to S square we get,
Expectation of N minus n into s square by N into n
Is equal to N minus n by N into n multiplied by s square which is equal to variance of p.
Hence, N minus n by p into p into q by n minus 1 gives an unbiased estimate of variance of p as desired.

<u>Standard Errors of the Distribution: Mean , Total and Proportion</u>

a) The Standard Error of sample mean

is equal square root of variance of y bar is equal to square root of N minus n into sigma square divided by n into N minus one
Which is equal to square root of N minus n into S square divided by n into N.

b) An estimate of Standard Error of sample mean

is equal to estimate of square root of variance of y bar is equal to square root of N minus n into s square divided by N minus 1 into n.

c) An estimate of Standard Error of estimated population total

is equal to estimate of square root of variance of Y cap is equal to estimate of square root of N square into variance of y bar is equal to square root of N square into N minus n into s square divided by N into n .

d) The Standard Error of the sample proportion

Standard error of p is equal to square root of variance of p is equal to square root of N minus n by N minus one into PQ by n.

An estimate of Standard Error of sample proportion Estimate of standard error of p is equal to square root of N minus n into pq by N into n minus one.

Here’s a summary of our learning in this session:

- Standard error
- Results related to the estimation of standard errors of mean and total under Simple Random Sampling With Replacement and Simple Random Sampling Without Replacement
- Standard error of proportions