Frequently Asked Questions

1. What do you mean by Standard Error?

Answer:

The **STANDARD ERROR** measures the long-run average spread of the estimated values in the same hypothetical scenario.

Both bias and standard error 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 the truth, even if its average is close to the truth. Standard error can be defined as the standard deviation of the distribution.

For example, suppose t is an estimator then Standard Error of t equals to square root of variance of t.

2. What are the purposes of obtaining standard errors?

Answer:

The formulae for the standard errors of the estimated population mean and total are used primarily for three purposes:

1) To compare the precision obtained by Simple Random Sampling with that given by other methods of sampling
2) To estimate the size of the sample needed in a survey that is being planned
3) To estimate the precision actually attained in a survey that has been completed.

3. Under SRSWR variance of sample mean is given by sigma square by n.

i.e., $V(\bar{y}) = \sigma^2 / n$

Answer:

Let $y_1, y_2, \ldots, y_n$ be SRSWR sample of size n drawn from a population of size N. We know $\bar{Y} = \dfrac{\sum_{i=1}^{N} Y_i}{N}$ and $\bar{y} = \dfrac{\sum_{i=1}^{n} y_i}{n}$.

Also we have $\sigma^2 = \dfrac{\sum (Y_i - \bar{Y})^2}{N} = \dfrac{\sum Y_i^2}{N} - \bar{Y}^2$ and $E(\bar{y}) = \bar{Y}$

Consider $V(\bar{y}) = V(\sum y_i / n)$

$$= \frac{V(\sum y_i)}{n^2}$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} V(y_i) + \sum_{i \neq j} \operatorname{cov}(y_i, y_j)\right]$$

Here cov $(y_i, y_j) = 0$ since the y_i units selected at ith draw is independent of y_j units selected at the jth draw.

Therefore $V(\bar{y}) = \dfrac{1}{n^2}\left[\sum_{i=1}^{n} V(y_i)\right] \longrightarrow$ (1)

$$V(y_i) = E(y_i^2) - [E(y_i)]^2$$

But $E(y_i^2) = \sum_{i=1}^{N} Y_i^2 \bullet \dfrac{1}{N} = \dfrac{\sum Y_i^2}{N}$

$$E(y_i) = \sum_{i=1}^{N} Y_i \bullet \frac{1}{N} = \bar{Y}$$

Therefore $V(y_i) = \dfrac{\sum Y_i^2}{N} - (\bar{Y})^2 = \sigma^2$

By substituting the above in equation 1, we get

$$V(\bar{y}) = \frac{1}{n^2}\sum V(y_i)$$

$$= \frac{1}{n^2}\sum_{i=1}^{n}\sigma^2 = \frac{n\sigma^2}{n^2} = \sigma^2/n$$

4. Under SRSWR, sample mean square is an unbiased estimator of σ^2 i.e. to show that $E(s^2) = \sigma^2$

Answer:

Let $y_1, y_2, \ldots, y_n$ be SRSWR sample of size n drawn from a population of size N. We know $\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}$ and $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$.

We have $E(\bar{y}) = \bar{Y}$ and $V(\bar{y}) = \sigma^2 / n$

We kmow that $\sigma^2 = \frac{\sum (Y_i - \bar{Y})^2}{N} = \frac{\sum Y_i^2}{N} - \bar{Y}^2$

$$\Rightarrow \frac{\sum Y_i^2}{N} = \sigma^2 + \bar{Y}^2 \qquad \text{———} \quad *$$

$$s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1} = \frac{1}{n-1}\left[\sum y_i^2 - n\bar{y}^2\right]$$

Consider $E(s^2) = E(\left[\frac{1}{n-1} \times \sum (y_i - \bar{y}^2)\right]$

$$= \frac{1}{n-1} \times E\left[\sum y_i^2 - n\bar{y}^2\right]$$

$$= \frac{1}{n-1} \times \left[\sum_{i=1}^{n} E(y_i^2) - nE(\bar{y}^2)\right] \qquad \text{———} \quad 1$$

$$E(y_i^2) = \sum_{i=1}^{N} Y_i^2 \times \frac{1}{N} = \frac{\sum Y_i^2}{N} = \sigma^2 + \bar{Y}^2 \qquad \text{———} \qquad 2$$

$$E(\bar{y}^2) = V(\bar{y}) + [E(\bar{y})]^2 = \sigma^2/n + \bar{Y}^2 \qquad \text{———} \qquad 3$$

Substituting eqns 2 and 3 in eqn 1, we get

$$E(s^2) = \frac{1}{n-1} \times \left[\sum_{i=1}^{n} (\sigma^2 + \bar{Y}^2) - n(\frac{\sigma^2}{n} + \bar{Y}^2) \right]$$

$$= \frac{1}{n-1} \times \left[n(\sigma^2 + \bar{Y}^2) - n(\sigma^2/n + \bar{Y}^2 \right]$$

$$= \frac{1}{n-1} \times \left[n\sigma^2 + n\bar{Y}^2 - n\sigma^2/n - n\bar{Y}^2 \right]$$

$$= \frac{1}{n-1} \times (n-1)\sigma^2$$

$$= \sigma^2$$

Therefore sample mean square is an unbiased estimator of population variance.

5. The variance of population total under SRSWR is given by $N^2\sigma^2/n$ i.e. to prove that $V(\hat{Y}) = N^2\sigma^2/n$

Answer:

Let $y_1, y_2, \ldots, y_n$ be SRSWR sample of size n drawn from a population of size N.

$$V(\hat{Y}) = V(N\bar{y}) = N^2 V(\bar{y}) = N^2\sigma^2/n$$

6. An unbiased estimator of variance of estimated population total is given by $N^2 s^2 / n$ i.e. to prove that $V(\hat{\hat{Y}}) = N^2 s^2 / n$

Answer:

Consider $\hat{V}(\hat{Y}) = (N^2 \hat{\sigma}^2 / n) = \frac{N^2}{n} \times \hat{\sigma}^2 = \frac{N^2 s^2}{n}$

$[\because E(s^2) = \sigma^2 \Rightarrow \hat{\sigma}^2 = s^2]$.

7. Under SRSWOR variance of the unbiased estimator of population mean is given by $V(\bar{y}) = (\frac{N-n}{N-1}) \times (\frac{\sigma^2}{n})$ or

$$V(\bar{y}) = (\frac{N-n}{N}) \times (\frac{S^2}{n})$$

Answer:

We know that

Also we have $\sigma^2 = \frac{\sum (Y_i - \bar{Y})^2}{N} = \frac{\sum Y_i^2}{N} - \bar{Y}^2$ and

$$S^2 = \frac{\sum (Y_i - \bar{Y})^2}{N-1} \Rightarrow (N-1)S^2 = N\sigma^2 \Rightarrow \sigma^2 = \frac{(N-1)S^2}{N}$$

Consider $V(\bar{y}) = V(\sum y_i / n)$

$$= \frac{V(\sum y_i)}{n^2}$$

$$= \frac{1}{n^2}\left[\sum_{i=1}^{n} V(y_i) + \sum_{i \neq j} \text{cov}(y_i, y_j)\right] \quad \text{———— } 1$$

$$V(y_i) = E(y_i^2) - [E(y_i)]^2$$

But $E(y_i^2) = \sum_{i=1}^{N} Y_i^2 \bullet \frac{1}{N} = \frac{\sum Y_i^2}{N}$

$$E(y_i) = \sum_{i=1}^{N} Y_i \bullet \frac{1}{N} = \overline{Y}$$

Therefore $V(y_i) = \frac{\sum Y_i^2}{N} - (\overline{Y})^2 = \sigma^2$ ———— 2

$$\text{cov}(y_i, y_j) = E(y_i, y_j) - E(y_i) \times E(y_j)$$

$$E(y_i) = \sum_{i=1}^{N} Y_i \bullet \frac{1}{N} = \overline{Y}$$

$$E(y_j) = \sum_{j=1}^{N} Y_j \bullet \frac{1}{N} = \overline{Y}$$

Since yi and yj takes values Yi and Yj with probability 1/N each ,therefore

$\text{cov}(y_i, y_j) = E(y_i, y_j) - \overline{Y}^2$ —— 3

$$E(y_i, y_j) = \sum_{i \neq j}^{N} Y_i Y_j \times p(y_i, y_j)$$

$= \sum_{i \neq j}^{N} Y_i Y_j \times \frac{1}{N(N-1)}$ —— 4

We know that $\sum (Y_i - \overline{Y}) = 0$

$$\Rightarrow \left[\sum (Y_i - \overline{Y})\right]^2 = 0$$

$$\Rightarrow \left[\sum Y_i - N\overline{Y}\right]^2 = 0$$

$$\Rightarrow (\sum Y_i)^2 + N^2 \overline{Y}^2 - 2N\overline{Y}(\sum Y_i) = 0$$

$$\Rightarrow (\sum Y_i)^2 + N^2 \overline{Y}^2 - 2N\overline{Y}N\overline{Y} = 0$$

$$\Rightarrow (\sum Y_i)^2 + N^2 \overline{Y}^2 - 2N^2 \overline{Y}^2 = 0$$

$$\Rightarrow (\sum Y_i)^2 - N^2 \overline{Y}^2 = 0$$

$$=> \sum Y_i^2 + \sum\sum Y_i Y_j - N^2 \bar{Y}^2 = 0$$

$$=> \sum\sum Y_i Y_j = N^2 \bar{Y}^2 - \sum Y_i^2$$

$$= N^2 \bar{Y}^2 - (N\sigma^2 + N\bar{Y}^2)$$

$$= N^2 \bar{Y}^2 - N(\sigma^2 + \bar{Y}^2) \quad \text{——} \quad 5$$

By substituting eqn 5 in eqn 4 we get

$$E(y_i, y_j) = \sum_{i \neq j}^{N} Y_i Y_j \times \frac{1}{N(N-1)}$$

$$= \left[N^2 \bar{Y}^2 - N(\sigma^2 + \bar{Y}^2)\right] \times \frac{1}{N(N-1)}$$

$$= \frac{N\bar{Y}^2 - \sigma^2 - \bar{Y}^2}{(N-1)} \quad \text{——} \quad 6$$

By substituting eqn 6 in eqn 3 we get

$$Cov(y_i, y_j) = \frac{N\bar{Y}^2 - \sigma^2 - \bar{Y}^2}{(N-1)} - \bar{Y}^2$$

$$= \frac{N\bar{Y}^2 - \sigma^2 - \bar{Y}^2 - N\bar{Y}^2 + \bar{Y}^2}{(N-1)}$$

$$= \frac{-\sigma^2}{(N-1)} \quad \text{——} \quad 7$$

By substituting eqns 2 and 7 in eqn 1, we get

$$V(\bar{y}) = \frac{1}{n^2}\left[\sum_{i=1}^{n} \sigma^2 + \sum_{i \neq j} \frac{-\sigma^2}{N-1}\right]$$

$$= \frac{1}{n^2}\left[n\sigma^2 - n(n-1)\frac{\sigma^2}{N-1}\right]$$

$$= \frac{1}{n}\left[\sigma^2 - (n-1)\frac{\sigma^2}{N-1}\right]$$

$$= \frac{\sigma^2}{n}[1 - \frac{(n-1)}{N-1}]$$

$$= \frac{\sigma^2}{n}[\frac{N-1-n+1}{N-1}]$$

$$= \frac{\sigma^2}{n}[\frac{N-n}{N-1}] \text{ ——————— } \star$$

We know that

$$\sigma^2 = \frac{(N-1)S^2}{N}$$

Hence $V(\bar{y}) = \frac{(N-1)s^2}{Nn}[\frac{N-n}{N-1}]$

$$= \frac{(N-n)s^2}{Nn} \text{ ——————— } \star\star$$

8. A sample mean square s^2 is an unbiased estimator of population mean square S^2 under SRSWOR i.e., $E(s^2) = S^2$

Answer:

Let $y_1, y_2, \ldots, y_n$ be SRSWOR sample of size n drawn from a population of size N. We know $\bar{Y} = \frac{\sum_{i=1}^{N} Y_i}{N}$ and $\bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$.

We have $E(\bar{y}) = \bar{Y}$ and $V(\bar{y}) = \frac{N-n}{N-1} \times \frac{\sigma^2}{n} = \frac{N-n}{N} \times \frac{S^2}{n}$

Also $\sigma^2 = \frac{\sum (Y_i - \bar{Y})^2}{N} = \frac{\sum Y_i^2}{N} - \bar{Y}^2$

$$s^2 = \frac{\sum (y_i - \bar{y})^2}{n-1} \quad \text{and} \quad S^2 = \frac{\sum (Y_i - \bar{Y})^2}{N-1}$$

Consider $E(s^2) = E(\left[\frac{1}{n-1} \times \sum (y_i - \bar{y}^2)\right]$

$$= \frac{1}{n-1} \times E\left[\sum y_i{}^2 - n\bar{y}^2\right]$$

$$= \frac{1}{n-1} \times \left[\sum_{i=1}^{n} E(y_i{}^2) - nE(\bar{y}^2)\right] \qquad \text{———} \quad 1$$

$$E(y_i{}^2) = \sum_{i=1}^{N} Y_i^2 \times \frac{1}{N} = \frac{\sum Y_i^2}{N} = \sigma^2 + \bar{Y}^2 \qquad \text{———} \quad 2$$

$$E(\bar{y}^2) = V(\bar{y}) + \left[E(\bar{y})\right]^2 = (\frac{N-n}{N-1} \times \sigma^2/n) + \bar{Y}^2 \qquad \text{———} \quad 3$$

Substituting eqns 2 and 3 in eqn 1, we get

$$E(s^2) = \frac{1}{n-1} \times \left[\sum_{i=1}^{n} (\sigma^2 + \bar{Y}^2) - n(\frac{N-n}{N-1} \bullet \frac{\sigma^2}{n} + \bar{Y}^2)\right]$$

$$= \frac{1}{n-1} \times \left[n(\sigma^2 + \bar{Y}^2) - (\frac{(N-n)\sigma^2}{N-1}) - n\bar{Y}^2\right]$$

$$= \frac{1}{n-1} \times \left[n\sigma^2 + n\bar{Y}^2 - (\frac{N-n}{N-1})\sigma^2 - n\bar{Y}^2\right]$$

$$= \frac{\sigma^2}{n-1}\left[n - \frac{N-n}{N-1}\right]$$

$$= \frac{\sigma^2}{n-1}\left[\frac{n(N-1) - N + n}{N-1}\right]$$

$$= \frac{\sigma^2}{n-1}\left[\frac{nN - n - N + n}{N-1}\right]$$

$= \frac{\sigma^2}{n-1}\left[\frac{N(n-1)}{N-1}\right]$

$= \frac{N\sigma^2}{N-1}$

$= S^2$

Therefore sample mean square is an unbiased estimator of population variance.

9. An estimate of the variance of the sample mean under SRSWOR is given by $\hat{V}(\bar{y}) = \frac{N-n}{N} \times \frac{s^2}{n}$

Answer:

Let $y_1, y_2 \ldots y_n$ be SRSWOR sample of size n drawn from a population of size N.

We know that $\hat{V}(\bar{y}) = \frac{N-n}{N} \times \frac{\hat{S}^2}{n}$

$$\Rightarrow \hat{V}(\bar{y}) = \frac{N-n}{N} \times \frac{s^2}{n} \quad [\because E(s^2) = S^2 \Rightarrow \hat{S}^2 = s^2]$$

10. The variance of estimated population total under SRSWOR is given by $\frac{N-n}{N-1} \times N^2\sigma^2/n = \frac{N-n}{N} \times N^2S^2/n$

Answer:

Let $y_1, y_2 \ldots y_n$ be SRSWOR sample of size n drawn from a population of size N.

$$V(\hat{Y}) = V(N\bar{y}) = N^2V(\bar{y}) = N^2 \times \frac{N-n}{N-1} \times \frac{\sigma^2}{n}$$

11. An unbiased estimator of variance of estimated population total is given by $\hat{V}(\hat{Y}) = N^2 \times \frac{N-n}{N-1} \times s^2 \Big/ n$

Answer:

We know that $V(\hat{Y}) = N^2 V(\bar{y})$

$$= N^2 \left(\frac{N-n}{N}\right)\frac{S^2}{n}$$

$$\Rightarrow (\hat{V}(\hat{Y})) = N^2 \left(\frac{N-n}{N}\right)\frac{\hat{S}^2}{n}$$

$$= \Rightarrow N^2 \left(\frac{N-n}{N}\right)\frac{s^2}{n} \qquad [\because E(s^2) = S^2 \Rightarrow \hat{S}^2 = s^2].$$

12. Under SRSWOR Obtain an expression for the variance of the sample proportion p

Answer:

Already we proved in the last topic that p= $\bar{y}$

We know that

$\sum_{i=1}^{N} Y_i^2$ = A = NP and $\sum_{i=1}^{n} y_i^2$ =a = np

$$S^2 = \frac{\Sigma(Y_i - \bar{Y})^2}{N-1} = \frac{1}{N-1}\left[\sum_i Yi^2 - N\bar{Y}^2\right]$$

$$= \frac{1}{N-1}\left[NP - NP^2\right] = \frac{NPQ}{N-1}$$

Similarly $s^2 = \frac{\Sigma(y_i - \bar{y})^2}{n-1} = \frac{1}{n-1}\left[\sum_i yi^2 - n\bar{y}^2\right] = \frac{npq}{n-1}$

Therefore

$$V(p) = V(\overline{y}) = \frac{N-n}{Nn}S^2 = \frac{N-n}{Nn}\frac{NPQ}{N-1}$$

$$= \frac{(N-n)}{(N-1)}\frac{PQ}{n}$$

13. Derive a expression for the estimate of variance of the sample proportion.

Answer:

We have E (s^2) = S^2 which implies $E\left[\frac{N-n}{Nn}s^2\right] = \frac{N-n}{Nn}S^2$

$$E\left[\frac{(N-n)}{(Nn)}\frac{(npq)}{(n-1)}\right] = V(p)$$

$$E\left[\frac{(N-n)}{(N)}\frac{(pq)}{(n-1)}\right] = V(p)$$

Hence $v(p) = \left[\frac{(N-n)}{N}\frac{(pq)}{(n-1)}\right]$ gives an unbiased estimate of the V(p) as desired

14. How do you estimate the standard errors of the mean and total under SRSWR and get the estimates of the standard errors

Answer:

S.E. can be defined as the S.D. of the distribution i.e. square root of the variance of the distribution

a) The S. E $(\bar{y})$

$$S.E.(\bar{y}) = \sqrt{V(\bar{y})} = \sqrt{\sigma^2/n}$$

b) An estimate of S.E. of sample mean

$$\hat{S.E.}(\bar{y}) = \sqrt{\hat{V}(\bar{y})} = \sqrt{\hat{\sigma}^2/n} = \sqrt{\frac{s^2}{n}}$$

c) A S.E. of the estimated population total

$$S.E.(\hat{Y}) = \sqrt{V(\hat{Y})} = \sqrt{N^2\sigma^2/n}$$

d) An estimate of S.E. of estimated population total

$$\hat{S.E.}(\hat{Y}) = \sqrt{\hat{V}(\hat{Y})} = \sqrt{N^2\hat{\sigma}^2/n} = \sqrt{\frac{N^2\,\hat{\sigma}^2}{n}} = \sqrt{\frac{N^2 s^2}{n}}$$

15. How do you estimate the standard errors of the mean, total and proportion under SRSWOR and get the estimates of the standard errors

Answer:

a) The S.E. of sample mean

$$S.E.(\bar{y}) = \sqrt{V(\bar{y})} = \sqrt{\frac{N-n}{N-1} * \frac{\sigma^2}{n}} = \sqrt{\frac{N-n}{n} * \frac{S^2}{n}}$$

b) An estimate of S.E. of sample mean

$$\hat{S.E.}(\bar{y}) = \sqrt{\hat{V}(\bar{y})} = \sqrt{(\frac{N-n}{N})\frac{s^2}{n}}$$

c) An estimate of S.E. of estimated population total

$$\hat{S.E.}(\hat{Y}) = \sqrt{\hat{V}(\hat{Y})} = \sqrt{N^2\hat{V}(\bar{y})} = \sqrt{N^2 \times \frac{(N-n)}{N} \times \frac{s^2}{n}}$$

d) The S.E of the sample proportion

$$S.E.(p) = \sqrt{V(p)} = \sqrt{\frac{N-n}{N-1} * \frac{PQ}{n}}$$

e)An estimate of S.E. of sample proportion

$$\hat{S.E.}(p) = \sqrt{\hat{V}(p)} = \sqrt{\frac{(N-n)pq}{N(n-1)}}$$