

Assignment-5

S. Rithvik Reddy - cs20btech11049

Download all python codes from

NA

and latex-tikz codes from

<https://github.com/rithvikreddy6300/Assignment-5/blob/main/Assignment-5.tex>

PROBLEM-UGC/MATH 2018(JUNE MATHS SET-A), Q.58

- 1) A simple random variable of size n will be drawn from a class of 125 students, and the mean mathematics score of the sample will be computed, If the standard error of the sample mean for "with replacement sampling" is twice as much as the standard error of the sample mean for "without replacement sampling", the value of n is ?

- a) 32
- b) 63
- c) 79
- d) 94

SOLUTION

Let N be the population size so, $N=120$. The given sample size is n .

Notations :

y : student under consideration.

y_i : Maths marks of i^{th} student in the sample.

Y : student of class.

Y_i : Maths marks of i^{th} student in the class.

$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$: Average of sample class.

$\bar{Y} = \frac{1}{N} \sum_{i=1}^N Y_i$: Average of whole class.

$S^2 = \frac{1}{N-1} \sum_{i=1}^N (Y_i - \bar{Y})^2$: S =Std dev of the class.

$\sigma^2 = \frac{1}{N} \sum_{i=1}^N (Y_i - \bar{Y})^2$: Variance of the class.

Standard error of sample mean $SE_{mean} = \frac{s}{\sqrt{n}}$.

Where

s = standard deviation of sample mean.

n = sample class size.

Variance of the \bar{y}

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2 \quad (0.0.1)$$

$$= E \left[\frac{1}{n} \sum_{i=1}^n (y_i - \bar{Y}) \right]^2 \quad (0.0.2)$$

$$= E \left[\frac{1}{n^2} \sum_{i=1}^n (y_i - \bar{Y})^2 + \frac{1}{n^2} \sum_{1 \leq i \neq j \leq n} (y_i - \bar{Y})(y_j - \bar{Y}) \right] \quad (0.0.3)$$

$$= \frac{1}{n^2} \sum_{i=1}^n E(y_i - \bar{Y})^2 + \frac{1}{n^2} \sum_{1 \leq i \neq j \leq n} E(y_i - \bar{Y})(y_j - \bar{Y}) \quad (0.0.4)$$

$$\text{Let } K = \sum_{1 \leq i \neq j \leq n} E(y_i - \bar{Y})(y_j - \bar{Y}) \quad (0.0.5)$$

$$= \frac{1}{n^2} \sum_{i=1}^n \sigma^2 + \frac{K}{n^2} \quad (0.0.6)$$

$$= \frac{1}{n^2} n \sigma^2 + \frac{K}{n^2} \quad (0.0.7)$$

$$= \frac{N-1}{Nn} S^2 + \frac{K}{n^2} \quad (0.0.8)$$

Finding the value of K in case of Simple random sampling with repetition (SRSWR) and Simple random sampling without repetition (SRSSWOR) allows us to calculate the variance of mean.

K value in case of SRSSWOR

$$K = \sum_{1 \leq i \neq j \leq n} E(y_i - \bar{Y})(y_j - \bar{Y})$$

Consider

$$E(y_i - \bar{Y})(y_j - \bar{Y}) = \frac{1}{N(N-1)} \sum_{1 \leq k \neq l \leq n} E(y_k - \bar{Y})(y_l - \bar{Y})$$

Since

$$\left[\sum_{k=1}^N (y_k - \bar{Y}) \right]^2 = \sum_{i=1}^N (y_i - \bar{Y})^2 + \sum_{1 \leq k \neq l \leq n} E(y_k - \bar{Y})(y_l - \bar{Y})$$

$$\begin{aligned} \Rightarrow 0 &= (N-1)S^2 + \sum_{1 \leq k \neq l \leq n} E(y_k - \bar{Y})(y_l - \bar{Y}) \\ \Rightarrow E(y_i - \bar{Y})(y_j - \bar{Y}) &= \frac{1}{N(N-1)}(N-1)(-S^2) \\ \Rightarrow K &= n(n-1) \frac{(-S^2)}{N} \end{aligned}$$

Putting this value in (0.0.8) gives us

$$V(\bar{y})_{WOR} = \frac{N-1}{Nn} S^2 + \frac{n-1(-S^2)}{Nn} \quad (0.0.9)$$

$$= \frac{N-n}{Nn} S^2 \quad (0.0.10)$$

K value in case of SRSWR

$$K = \sum_{1 \leq i \neq j \leq n} E(y_i - \bar{Y})(y_j - \bar{Y})$$

Since we are selecting the samples with replacements choosing i^{th} and j^{th} sample is independent of each other. So,

$$\begin{aligned} K &= \sum_{1 \leq i \neq j \leq n} E(y_i - \bar{Y})E(y_j - \bar{Y}) \\ &= 0 \end{aligned}$$

(Since deviation about mean is 0)

Putting $K=0$ in (0.0.8) we get

$$V(\bar{y})_{WR} = \frac{N-1}{Nn} S^2 \quad (0.0.11)$$

From equation (0.0.10) standard error of mean of sample class without repetition

$$SE_{WOR} = \frac{s}{\sqrt{n}} \quad (0.0.12)$$

$$= \sqrt{\frac{V(\bar{y})_{WOR}}{n}} \quad (0.0.13)$$

$$= \sqrt{\frac{N-n}{Nn^2}} S \quad (0.0.14)$$

From equation (0.0.11) standard error of mean of sample class with repetition

$$SE_{WR} = \sqrt{\frac{V(\bar{y})_{WR}}{n}} \quad (0.0.15)$$

$$= \sqrt{\frac{N-1}{Nn^2}} S \quad (0.0.16)$$

Given to find the value of n if $2 \times SE_{WOR} = SE_{WR}$.
From (0.0.14) and (0.0.16) we can write

$$2 \sqrt{\frac{N-n}{Nn^2}} S = \sqrt{\frac{N-1}{Nn^2}} S \quad (0.0.17)$$

$$\Rightarrow 4(N-n) = N-1 \quad (0.0.18)$$

$$\Rightarrow 4N + 1 - N = 4n \quad (0.0.19)$$

$$\Rightarrow 4n = 3(125) + 1 \quad (0.0.20)$$

$$\Rightarrow n = 94 \quad (0.0.21)$$

Therefore the sample size for the given condition to be met is $n=94$. **(Option D)**