STAT 522

HW6

Satoshi Ido

Mar 8 2024

1. a. $\bar{y}_u = \frac{t}{N} = \frac{\sum_{h}^{H}\sum_{j}^{N} y_{hj}}{N} = \frac{431}{6} = 71.8$

$$S_h^2 = \sum_{j=1}^{N_h} \frac{(y_{hj} - \bar{y}_{hu})^2}{N_h - 1} = \frac{430.84}{6-1} = 86.168$$

b $\binom{6}{4} = \frac{6 \cdot 5}{2-1} = 15$

c. $\because \bar{y} = \frac{1}{n}\sum^{n} y_j \quad \because Vn(\bar{y}) = \frac{S^2}{n}\left(1 - \frac{n}{N}\right)$

Sample 1 = {66, 59, 70, 83}, $\bar{y}_1 = 69.5$, $V(\bar{y}_1) = 8.47$

Sample 2 = {66, 59, 70, 82}, $\bar{y}_2 = 69.25$, $V(\bar{y}_2) = 7.74$

Sample 3 = {66, 59, 70, 71}, $\bar{y}_3 = 66.50$, $V(\bar{y}_3) = 2.78$

Sample 4 = {66, 59, 83, 82}, $\bar{y}_4 = 72.50$, $V(\bar{y}_4) = 11.08$

Sample 5 = {66, 59, 83, 71} $\bar{y}_5 = 69.75$, $V(\bar{y}_5) = 6.41$

Sample 6 = {66, 59, 82, 71} $\bar{y}_6 = 69.50$, $V(\bar{y}_6) = 8.47$

Sample 7 = {66, 70, 83, 82} $\bar{y}_7 = 75.25$, $V(\bar{y}_7) = 5.41$

Sample 8 = {66, 70, 83, 71} $\bar{y}_8 = 72.50$, $V(\bar{y}_8) = 3.58$

Sample 9 = {66, 70, 82, 71} $\bar{y}_9 = 72.25$, $V(\bar{y}_9) = 2.91$

Sample 10 = {66, 83, 82, 71} $\bar{y}_{10} = 75.50$, $V(\bar{y}_{10}) = 5.35$

Sample 11 = {59, 70, 83, 82} $\bar{y}_{11} = 73.50$, $V(\bar{y}_{11}) = 8.47$

Sample 12 = {59, 70, 83, 71} $\bar{y}_{12} = 70.75$, $V(\bar{y}_{12}) = 5.81$

Sample 13 = {59, 70, 82, 71} $\bar{y}_{13} = 70.50$, $V(\bar{y}_{13}) = 5.14$

Sample 14 = {59, 83, 82, 71} $\bar{y}_{14} = 73.75$, $V(\bar{y}_{14}) = 8.31$

Sample 15 = {70, 83, 82, 71} $\bar{y}_{15} = 76.50$, $V(\bar{y}_{15}) = 2.64$

//

d. $\binom{3}{1} \times \binom{3}{1} = 9$ $\therefore$ 9 startified samples

e.

Possible stratified random samples

= Sample 4, 5, 6, 7, 8, 9, 11, 12, 13 referred to prob C. //

The samples which cannot occur with the stratified design

= sample 1, 2, 3, 10, 14, 15 referred to prob C. //

f. $\because \bar{y}_{str} = \sum_{h=1}^{H} \frac{N_h}{N} \bar{y}_h$ where $N_h = 3$ and $N = 6$

$\therefore V(\bar{y}_{str}) = \sum_{h=1}^{H} \left(1 - \frac{n_h}{N_h}\right)\left(\frac{N_h}{N}\right)^2 \cdot \frac{S_h^2}{n_h}$ where $S_h^2 = \sum_{j \in S_h} \frac{(y_{hj} - \bar{y}_h)^2}{n_h - 1}$

Sample 4,

$$\bar{y}_{str} = \frac{3}{6}(62.5 + 82.5) = 72.5$$

$$V(\bar{y}_{str}) = \left(1 - \frac{2}{3}\right)\left(\frac{3}{6}\right)^2 \cdot \frac{S_{h1}^2 + S_{h2}^2}{2}$$

$$S_{h1}^2 = \frac{1}{2-1}\{(66 - 62.5)^2 + (59 - 62.5)^2\} = 24.5$$

$$S_{h2}^2 = \frac{1}{2-1}\{(83 - 82.5)^2 + (82 - 82.5)^2\} = 0.5$$

$$\therefore V(\bar{y}_{str}) = \frac{1}{24}(24.5 + 0.5) = 1.041\overline{6} \cong 1.042$$

The same process is applied to other samplings.

Sample 5, $\bar{y}_{str} = 69.75$ $V(\bar{y}_{str}) \cong 4.02$

Sample 6, $\bar{y}_{str} = 69.5$ $V(\bar{y}_{str}) \cong 3.54$

Sample 7, $\bar{y}_{str} = 73.25$ $V(\bar{y}_{str}) \cong 0.35$

Sample 8, $\bar{y}_{str} = 72.5$ $V(\bar{y}_{str}) \cong 3.33$

Sample 9, $\bar{y}_{str} = 72.25$ $V(\bar{y}_{str}) \cong 2.65$

Sample 11, $\bar{y}_{str} = 73.5$ $V(\bar{y}_{str}) \cong 2.54$

Sample 12, $\bar{y}_{str} = 70.75$ $V(\bar{y}_{str}) \cong 5.52$

Sample 13, $\bar{y}_{str} = 70.5$ $V(\bar{y}_{str}) \cong 5.04$

Comparing to $V(\bar{y})$ of each sampling respectively, we can see the stratified sampling reduces the variance dramatically. The stratified method has smaller variance.

Due Friday, March 8 or after the break

Complete exercises below. For all questions, show your work.

1. Consider a population of 6 students. Suppose we know the test scores of the students to be:

Student	1	2	3	4	5	6
Score	66	59	70	83	82	71

   a. Find the mean $\bar{y}_U$ and variance S^2 of the population. (4 points)

   b. How many SRS samples of size 4 are possible? (1 point)

   c. List the possible SRS samples. For each, find the sample mean. Find $V(\bar{y})$ using $V(\bar{y}) = \frac{S^2}{n}\left(1 - \frac{n}{N}\right)$ (6 points)

   d. Now let stratum 1 consist of students 1-3 and stratum 2 consist of students 4-6. How many stratified random samples of size 4 are possible in which 2 students are selected from each stratum? (2 points)

   e. List the possible stratified random samples. Which of the samples from c. cannot occur with the stratified design? (4 points)

f. Find $\bar{y}_{str}$ for each possible stratified random sample. Find $V\left(\bar{y}_{str}\right)$ using $V\left(\bar{y}_{str}\right)=\sum_{h=1}^{H}\left(1-\frac{n_h}{N_h}\right)\left(\frac{N_h}{N}\right)^2\frac{S_h^2}{n_h}$, and compare it to $V\left(\bar{y}\right)$. Which sample has smaller variance? (6 points)