

5. Muestreo por Conglomerados

- 5.1 Necesidad y ventajas del muestreo por conglomerados.
- 5.2 Formación de los conglomerados. Conglomerados y estratos.
- 5.3 Estimación de la media, proporción y total poblacionales.
- 5.4 Determinación del tamaño muestral.

1

1

5. Muestreo por Conglomerados

- 5.1 Necesidad y ventajas del muestreo por conglomerados.
- 5.2 Formación de los conglomerados. Conglomerados y estratos. Notación

N = conglomerados en la población.

n = conglomerados en la muestra.

m_i = elementos en el conglomerado i

y_i = suma de las observaciones en el conglomerado i

$M = \sum_{i=1}^N m_i$ = elementos en la población

$m = \sum_{i=1}^n m_i$ = elementos en la muestra

$\bar{M} = \frac{1}{N} \sum_{i=1}^N m_i$ = tamaño medio de los conglomerados de la población.

$\bar{m} = \frac{1}{n} \sum_{i=1}^n m_i$ = tamaño medio de los conglomerados de la muestra.

2

2

5. Muestreo por Conglomerados

5.3 Estimación de la media, proporción y total poblacionales.

$$\hat{\mu} = \bar{y} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n m_i} \quad \hat{V}(\bar{y}) = \frac{1}{M^2} \frac{N-n}{N} \frac{S_c^2}{n} \quad S_c^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}m_i)^2$$

$$\hat{\tau} = M\bar{y} \quad \hat{V}(\hat{\tau}) = M^2 \hat{V}(\bar{y})$$

$$\hat{\tau}_t = N\bar{y}_t \quad \hat{V}(\hat{\tau}_t) = N^2 \hat{V}(\bar{y}_t) = N(N-n) \frac{S_t^2}{n}$$

$$\bar{y}_t = \frac{1}{n} \sum_{i=1}^n y_i \quad \hat{V}(\bar{y}_t) = \frac{N-n}{N} \frac{S_t^2}{n} \quad S_t^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_t)^2$$

Si $m_1 = m_2 = \dots = m_N \Rightarrow M\bar{y} = N\bar{y}_t$ 2

3

5. Muestreo por Conglomerados

5.4 Determinación del tamaño muestral.

$$\hat{\tau} = M\bar{y} \quad \hat{\mu} = \bar{y} = \frac{\sum_{i=1}^n y_i}{\sum_{i=1}^n m_i}$$

$$n = \frac{N\sigma_c^2}{ND + \sigma_c^2} \quad \hat{\sigma}_c^2 = S_c^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}m_i)^2 \quad D = \frac{B^2 \bar{M}^2}{4} \quad (\text{media})$$

$$D = \frac{B^2}{4N^2} \quad (\text{total})$$

$$\hat{\tau}_t = N\bar{y}_t \quad \bar{y}_t = \frac{1}{n} \sum_{i=1}^n y_i \quad \text{m.a.s. sobre los totales de los conglomerados}$$

$$n = \frac{N\sigma_t^2}{ND + \sigma_t^2} \quad \hat{\sigma}_t^2 = S_t^2 = \frac{1}{n-1} \sum_{i=1}^n (y_i - \bar{y}_t)^2 \quad D = \frac{B^2}{4N^2} \quad 3$$

4

5. Muestreo por Conglomerados

Ejemplo de población con 36 elementos:

101	201	301	401	501	601
102	202	302	402	502	602
103	203	303	403	503	603
104	204	304	404	504	604
105	205	305	405	505	605
106	206	306	406	506	606

Media: 353,5
Varianza: 29169,58
N: 36

Si estimamos la media de la anterior población usando muestreo aleatorio simple con un tamaño muestral de $n=24$, la varianza del estimador media muestral es 416,71.

5

5

5. Muestreo por Conglomerados

Ejemplo de población con 36 elementos en 6 conglomerados:

Conglomerado						
1	101	201	301	401	501	601
2	102	202	302	402	502	602
3	103	203	303	403	503	603
4	104	204	304	404	504	604
5	105	205	305	405	505	605
6	106	206	306	406	506	606

Media: 353,5
Varianza: 29169,58
M: 36
N: 6

6

6

5. Muestreo por Conglomerados

Muestra	Conglomerados en la muestra, $n=4$				m_i				y_i				ESTIMADOR DE LA MEDIA
	3	4	5	6	6	6	6	6	2118	2124	2130	2136	
1	3	4	5	6	6	6	6	6	2118	2124	2130	2136	354,5
2	2	4	5	6	6	6	6	6	2112	2124	2130	2136	354,25
3	2	3	5	6	6	6	6	6	2112	2118	2130	2136	354
4	2	3	4	6	6	6	6	6	2112	2118	2124	2136	353,75
5	2	3	4	5	6	6	6	6	2112	2118	2124	2130	353,5
6	1	4	5	6	6	6	6	6	2106	2124	2130	2136	354
7	1	3	5	6	6	6	6	6	2106	2118	2130	2136	353,75
8	1	3	4	6	6	6	6	6	2106	2118	2124	2136	353,5
9	1	3	4	5	6	6	6	6	2106	2118	2124	2130	353,25
10	1	2	5	6	6	6	6	6	2106	2112	2130	2136	353,5
11	1	2	4	6	6	6	6	6	2106	2112	2124	2136	353,25
12	1	2	4	5	6	6	6	6	2106	2112	2124	2130	353
13	1	2	3	6	6	6	6	6	2106	2112	2118	2136	353
14	1	2	3	5	6	6	6	6	2106	2112	2118	2130	352,75
15	1	2	3	4	6	6	6	6	2106	2112	2118	2124	352,5

MEDIA DEL ESTIMADOR: 353,5
 VARIANZA DEL ESTIMADOR: 0,29

7

5. Muestreo por Conglomerados

Ejemplo de población con 36 elementos en 6 conglomerados:

Conglomerados					
1	2	3	4	5	6
101	201	301	401	501	601
102	202	302	402	502	602
103	203	303	403	503	603
104	204	304	404	504	604
105	205	305	405	505	605
106	206	306	406	506	606

Media: 353,5
Varianza: 29169,58
M: 36
N: 6

8

8

5. Muestreo por Conglomerados

Muestra	Conglomerados en la muestra, $n=4$				m_i			y_i				ESTIMADOR DE LA MEDIA	
1	3	4	5	6	6	6	6	6	1821	2421	3021	3621	453,5
2	2	4	5	6	6	6	6	6	1221	2421	3021	3621	428,5
3	2	3	5	6	6	6	6	6	1221	1821	3021	3621	403,5
4	2	3	4	6	6	6	6	6	1221	1821	2421	3621	378,5
5	2	3	4	5	6	6	6	6	1221	1821	2421	3021	353,5
6	1	4	5	6	6	6	6	6	621	2421	3021	3621	403,5
7	1	3	5	6	6	6	6	6	621	1821	3021	3621	378,5
8	1	3	4	6	6	6	6	6	621	1821	2421	3621	353,5
9	1	3	4	5	6	6	6	6	621	1821	2421	3021	328,5
10	1	2	5	6	6	6	6	6	621	1221	3021	3621	353,5
11	1	2	4	6	6	6	6	6	621	1221	2421	3621	328,5
12	1	2	4	5	6	6	6	6	621	1221	2421	3021	303,5
13	1	2	3	6	6	6	6	6	621	1221	1821	3621	303,5
14	1	2	3	5	6	6	6	6	621	1221	1821	3021	278,5
15	1	2	3	4	6	6	6	6	621	1221	1821	2421	253,5

MEDIA DEL ESTIMADOR: 353,5
 VARIANZA DEL ESTIMADOR: 2916,67