

## Muestreo. Problemas.

### Diseños Bi-etápicos- limitados “conglomerados”

#### Problema 4

Cierto fabricante de lavadoras quiere conocer el equipamiento de sus clientes en otros aparatos, en particular en lavaplatos. Tiene 100 000 clientes en su fichero de clientes. Dicho fichero viene organizado en 2000 "grupos" de 50 clientes que viven en zonas geográficas suficientemente contiguas. El fabricante pide que se haga un sondeo a 1000 clientes. Es interesante, por razones económicas, efectuar dicho sondeo por conglomerados, aprovechando la organización del fichero. Por lo tanto, se decide extraer 20 grupos de clientes, mediante una extracción aleatoria simple sin reposición entre los 2000 grupos. A continuación, se entrevistan todos los clientes de los grupos seleccionados.

Se denota  $T_i$  el número de poseedores de un lavaplatos en el grupo  $i$ . Se obtienen en la muestra los siguientes resultados:  $\Sigma T_i = 124$   $\Sigma T_i^2 = 1818$

1. Estimar por punto y por intervalo el número total de poseedores de lavaplatos entre los clientes de la marca de lavadoras.
2. Estimar por punto y por intervalo la proporción  $p$  de poseedores de lavaplatos entre los clientes. En los dos casos, emplear un nivel de confianza del 95%.

#### Problema 5

Un cierto periódico se distribuye en 5000 quioscos. El servicio se hace con motoristas. Cada motorista atiende a 50 quioscos, relativamente cercanos. Se desea conocer el número de quiosqueros partidarios de abrir los domingos por la tarde. La dirección del periódico encarga un estudio destinado a estimar dicho número.

Se decide observar 1000 quioscos, en 20 conglomerados, aprovechando el reagrupamiento geográfico establecido.

En los 20 conglomerados seleccionados se observa los efectivos de partidarios de abrir los domingos por la tarde

nr	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	16	18	19	20	suma
efec	10	9	12	8	34	2	7	22	10	12	8	11	10	10	32	4	5	24	8	14	252

1. Estimar el número total de quiosqueros partidarios de abrir los domingos por la tarde
2. Sea  $P$  la proporción de quiosqueros partidarios de abrir los domingos por la tarde. Estimar  $P$ .
3. Opinen sobre la buena/mala calidad de la estimación y de las razones

4. ¿Cuál habría sido la precisión de la estimación con una muestreo aleatorio simple? Calculen  $\hat{V}(\hat{p})$  y el efecto del diseño DEFF.

### Problema 6

Un banco tiene 39800 clientes-empresas en sus ficheros informáticos repartidos en 3980 agentes. Cada agente gestiona exactamente 10 clientes. Se desea estimar la proporción de clientes a los cuales el banco ha hecho un préstamo.

Al azar (MAS), se seleccionan 40 agentes. Para cada agente seleccionado, se observa el número de clientes que han pedido un prestamos. Se obtiene:

nr	1		40	
efec	...		...	185
Efec <sup>2</sup>	...		...	1263

Estimar la proporción de clientes a los cuales el banco ha hecho un préstamo.

### Problema 7

Un club de excursionistas cuenta con 40000 socios. Se desea conocer el número de socios que son abonados a una determinada revista.

Existe una ficha para cada socio del club, y el fichero viene clasificado por orden geográfico de manera que las zonas geográficas se siguen las unas a las otras.

Se decide hacer una encuesta a 800 socios, tomados en conglomerados de 10. Así, se ahorran gastos de transporte, dado que un encuestador puede realizar 10 entrevistas al día, por término medio, si los entrevistados viven en una área geográfica cercana.

Se pueden considerar las  $N=40000$  fichas como una base compuesta por  $M=4000$  conglomerados de 10 unidades cada uno.

Se selecciona, con un procedimiento ASSR, 80 conglomerados. En cada conglomerado, se observa  $T_i$ , el número de socios abonados a la revista ( $0 \leq T_i \leq 10$ ). Se obtiene:

$$\Sigma T_i = 370$$

$$\Sigma T_i^2 = 2536$$

- Calcular un intervalo de confianza al 95% para el número total de abonados a la revista entre los socios del club
- Sea  $P$ , la proporción de socios abonados a la revista. Estimar  $P$ , por punto y por intervalo.
- ¿Cuál sería la precisión obtenida en la estimación de la proporción con un diseño ASSR?

- d) ¿Cómo explica la diferencia entre las amplitudes de los intervalos de confianza calculados en las anteriores preguntas?
- e) ¿Qué sugerencia se podría formular para mejorar la precisión de la selección bietápica empleada en la selección de la muestra?