

Ejercicios Muestreo Estratificado

Ejercicio 1. Una muestra estratificada se está diseñando para estimar la prevalencia p de una característica cualquiera. El estrato 1 con N_1 unidades tiene una alta prevalencia de la característica de interés y el estrato 2, con N_2 unidades tiene una prevalencia baja. Asumiendo que el costo de encuestar una unidad es igual en cada uno de los estratos y el tamaño de unidades a ser seleccionadas es 2000 y que dentro de cada estrato se selecciona una muestra bajo un SI

Se pide:

1. Sea p_1 y p_2 las prevalencias en los estratos 1 y 2, respectivamente. Si $p_1 = 0.10$ y $p_2 = 0.03$ y $N_1/N = 0.40$, cuales son n_1 y n_2 bajo asignación óptima?
2. Si $p_1 = 0.10$ y $p_2 = 0.03$ y $N_1/N = 0.40$. Calcular la varianza del estimador de la proporción de la prevalencia para el total de la población bajo asignación proporcional y óptima.
3. Calule los deff del punto 2

Ejercicio 2. Escriba en R una función que toma como argumentos $n_1, n_2, N_1, N_2, S_1, S_2$ y calcula la varianza para el total poblacional bajo un diseño estratificado simple. Elija algunos valores poblacionales para los tamaños de los estratos y las varianzas y grafique dicha función a medida que n_1 y n_2 cambian, de forma de hallar el óptimo. Analizar la sensibilidad de la varianza a medida que los tamaños de muestra por estrato cambian.

Ejercicio 3. Una consultora va a realizar una encuesta de opinión para la que dispone de un presupuesto de 20,000. La consultora sabe que el 90 % de los hogares tiene teléfono fijo. Las encuestas telefónicas tienen un costo de 10 por encuesta, mientras que si todas las encuestas se realizan de forma presencial tienen un costo de 30 por encuesta. Para los hogares que no tienen teléfono el costo de la encuesta presencial asciende a 40. Asumiendo que la varianza de los hogares con o sin teléfono es la misma y el costo fijo C es de 5000. Cuantos hogares por estrato deben ser entrevistados si:

1. Todos los hogares son entrevistados de forma presencial
2. Los hogares con teléfonos son entrevistados de forma telefónica, mientras que los hogares sin teléfono son entrevistados de manera presencial.

Ejercicio 4. Una agencia de publicidad quieren determinar qué tanto deberían intensificar la publicidad de sus cuentas en televisión, en una determinada área. Esta región comprende

dos pequeñas ciudades y una parte rural. La ciudad JL tiene vinculación con una industria textil, la mayoría de sus habitantes trabajan en la fábrica y tienen niños en edad escolar. La ciudad RSC es en realidad un suburbio de nivel socioeconómico alto cercano a una ciudad capital, sus habitantes son mayores y con pocos niños y/o adolescentes en los hogares.

Hay $N_1 = 155$ hogares en la ciudad JL, $N_2 = 62$ en la ciudad RSC y $N_3 = 93$ en la parte rural. En total, son 310 hogares. El objetivo es relevar informacin acerca del tiempo que se ve televisión (horas/ semana). La agencia tiene recursos suficientes para hacer entrevistas a $n = 40$ hogares. Se decide seleccionar muestras bajo un SI de $n_1 = 20$ en JL, $n_2 = 8$ en RSC y $n_3 = 12$ en la parte rural. Los resultados se muestran en la siguiente tabla:

	JL	RSC	Rural
N_h	155	62	93
n_h	20	8	12
\bar{y}_{s_h}	33.9	25.125	19
$S^2_{y_{s_h}}$	35.36	232.41	87.64

1. Estime el promedio de horas por semana que todos los hogares de la región ven televisión.
2. Calcule la estimación de la varianza del punto 1.
3. Estime el total de horas por semana que los hogares de la región están frente al televisor. Calcule el margen de error al 95
4. Calcule la eficiencia ganada por el hecho de haber estratificado los componentes geográficos de la región.

Otra pregunta que hicieron los encuestadores de la agencia fue si los hogares veían el canal CN. Resultó que 16 hogares de JL, 2 de RSC y 6 hogares rurales lo veían.

1. Estime la proporción de hogares que ven el canal CN en toda la región y calcule un IC al 95%
2. Estime el total de hogares de la región que ven el canal ése. Calcule también un IC al 95%.

Ejercicio 5. Se desea realizar una encuesta a hogares de un determinado barrio de Montevideo para conocer el parque vehicular. Se tiene un listado de todos los hogares del barrio (marco muestral) con su ubicación geográfica e ingreso promedio en el último año. Se seleccionó una muestra bajo un diseño STSI en donde se conformaron tres estratos en base al ingreso del hogar (1- bajo, 2 - medio y 3 -alto).

	bajo	medio	alto
N_h	3500	2000	2000
n_h	500	300	200
\bar{y}_{s_h}	0.13	0.45	0.60
$S^2_{y_{s_h}}$	0.123	0.390	0.560

1. Calcular la estimación puntual del total de vehículos y el promedio de vehículos por hogar.
2. Calcular la estimación de la varianza del estimador del total de vehículos y un intervalo de confianza del 95 % para el total de vehículos.
3. Calcular la estimación puntual de la varianza del promedio de vehículos por hogar y un intervalo de confianza del 95 % para el total de vehículos.
4. Considera que el criterio utilizado para la estratificación de los hogares es bueno?
El mismo estudio quiere volver a repetirse dentro de dos años y se planea utilizar el mismo tamaño de muestra total.
 - (a) Calcule el tamaño de muestra por estrato que usted utilizaría si se desea minimizar la varianza del estimador del total de vehículos en el barrio.
 - (b) Calcular la mejora en la eficiencia del diseño bajo esta nueva estrategia de asignación del tamaño de muestra por estrato.