# MUESTREO ESTADÍSTICO

## Muestreo Aleatorio Estratificado (M.A.E)

## SEMANA-6

Raúl Alberto Pérez

Universidad Nacional de Colombia, Escuela de Estadística, 2021-I

## Resúmen del Tamaño de muestra para estimar a: $\mu$ $\tau$ ó $p$

1. Si se desea estimar la media ($\mu$): $D = \frac{B^2}{Z^2_{1-\frac{\alpha}{2}}}$.

2. En general, haciendo un procedimiento análogo, para estimar el total, $\tau$, se llega a la misma expresión dada anteriormente, con:

$$D = \frac{B^2}{Z^2_{1-\frac{\alpha}{2}} N^2}.$$

3. Si se desea estimar la proporción, $p$, entonces

$$\sigma_h^2 = p_h(1-p_h) \quad y \quad D = \frac{B^2}{Z^2_{1-\frac{\alpha}{2}}}.$$

## Algunas observaciones importantes acerca del tamaño de muestra

1. Las varianzas $\sigma_h^2$ se aproximan por varianzas muestrales ($s_h^2$) obtenidas de estudios previos o a partir de un estudio piloto.

2. También si se tiene información acerca de la amplitud de las observaciones dentro de cada estrato se puede usar la siguiente aproximación para las desviaciones estándar en cada estrato:

$$\sigma_h = \frac{R_i}{6}, \quad R_h = Y_{h\,\text{máx}} - Y_{h\,\text{mín}}$$

3. Existen diferentes formas de determinar los $w_h$, los cuales se conocen como afijaciones o asignaciones de la muestra. Varias consideraciones al respecto se ven a continuación

## Afijación de la Muestra $n$ entre los estratos

Recordar que el objetivo de una encuesta por muestreo consiste en proporcionar estimadores con pequeñas varianzas al menor costo posible .

Después de elegir el tamaño de muestra $n$, existen diferentes maneras de repartir $n$ entre los diferentes estratos individualmente, ie. de hallar: $n_1, n_2, \ldots, n_H$

Cada división puede originar una varianza diferente para la media muestral, por lo tanto el objetivo consiste en utilizar una asignación que presente una cantidad especificada de información a un costo mínimo.

**En Resumen:**

El mejor esquema de asignación de $n$ a los estratos está influenciado por los siguientes tres factores:

1. El número total de elementos en cada estrato ($N_h$)

2. La variabilidad de las observaciones dentro de cada estrato ($\sigma_h^2$)

3. El costo de obtener una observación en cada estrato ($C_h$)

## Criterios para determinar las afijaciones $w_h$

Además se observan los siguientes criterios:

**a.** El número de elementos en cada estrato afecta la cantidad de información en la muestra.

**b.** Se deben fijar tamaños de muestra grandes a estratos grandes.

**c.** Tener en cuenta la variabilidad de los estratos, a mayor variabilidad el tamaño de muestra debe ser mayor.

**d.** Si el costo de obtener una observación varía de un estrato a otro se tomarán muestras pequeñas en estratos con costo alto con el fin de minimizar el costo.

**1. Asignación aproximada que minimiza el costo $C_h$ para un valor fijo de $Var[\overline{y}_{ets}] = D$ o que minimiza la varianza $Var[\overline{y}_{ets}] = D$ a un costo fijo $C_h$**

$$w_h = \left( \frac{\frac{N_h \sigma_h}{\sqrt{C_h}}}{\sum_{h=1}^{H} \left( \frac{N_h \sigma_h}{\sqrt{C_h}} \right)} \right)$$

donde, $N_h$-tamaño del $h$-ésimo estrato, $\sigma_h^2$-varianza poblacional del estrato $h$ y $C_h$-costo de obtener una observación en el estrato $h$.

Este valor de los $w_h$ se reemplaza en la expresión anterior para hallar el tamaño de muestra $n$.

En este caso el tamaño de muestra asociado a cada estrato $n_h$-es directamente proporcional al producto de $N_h$ y $\sigma_h$ e inversamente proporcional a $\sqrt{C_h}$

A esta asignación también se le conoce como **Asignación Óptima del Tamaño de Muestra**.

## 2. Asignación Óptima de Neyman (Costos iguales)

Si los costos por observación en cada estrato son desconocidos, se pueden asumir todos iguales, ie. $C_1 = C_2 = \cdots = C_H = C$, entonces:

$$w_h = \left( \frac{N_h \sigma_h}{\sum_{h=1}^{H} N_h \sigma_h} \right)$$

En este caso el tamaño de muestra asociado a cada estrato $n_h$, es directamente proporcional al producto de $N_h$ y $\sigma_h$.

## 3. Asignación proporcional al tamaño del estrato

**(Costos y Varianzas iguales)**

Además de encontrar costos iguales , ie. $C_1 = C_2 = \cdots = C_H = C$,

en alguna ocasiones resultan las mismas varianzas, ie. $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_H^2 = \sigma^2$

En este canso se tiene que:

$$w_h = \left( \frac{N_h}{\sum_{h=1}^{H} N_h} \right)$$

A esta asignación se le conoce como **Asignación Proporcional**

## 4. Asignación Igual por estrato
## Costos, Varianzas y estratos iguales

Si además de costos iguales, todos los estratos tienen aproximadamente el mismo tamaño, ie. $N_1 = N_2 = \cdots = N_h$

y no hay información disponible acerca de la variabilidad existente dentro de los estratos, ie. se asume que: $\sigma_1^2 = \sigma_1^2 = \cdots = \sigma_h^2 = \sigma^2$

entonces, lo más sencillo y aconsejable es asignar a cada estrato el mismo tamaño de muestra, lo cual se conoce como **Asignación Igual**, la cual está dada por:

$$w_h = \left(\frac{1}{H}\right), \;\; ie. \;\; n_h = \frac{n}{H}$$

donde, $n$-es el tamaño de muestra global y $H$-es el número de estratos.

Esta es la asignación menos adecuada debido a que no tienen en cuenta ni tamaños, ni variabilidad, ni costos en los estratos.

**EJEMPLO:** Una empresa publicitaria está interesada en determinar cuánto debe enfatizar la publicidad televisiva en una determinada región, y decide realizar una encuesta por muestreo para estimar el número promedio de horas por semana que se ve la televisión en los hogares de la región. La región comprende dos pueblos, pueblo A y pueblo B, y un área rural. El pueblo A está ubicado en torno a una fábrica, y la mayoría de los hogares son de trabajadores industriales con niños de edad escolar. El pueblo B es un suburbio exclusivo de una ciudad vecina y consta de habitantes más viejos con pocos niños en casa. Existen 155 hogares en el pueblo A, 62 hogares en el pueblo B y 93 hogares en el área rural.

Una encuesta anterior a los hogares de dicha región, sugiere que las varianzas de cada uno de los estratos ( Pueblo-A, Pueblo-B y Área Rural ) son aproximadamente: $\sigma_1^2 \approx 25, \sigma_2^2 \approx 225$ y $\sigma_3^2 \approx 100$.

Ahora, la empresa publicitaria encontró que cuesta más obtener una observación del área rural que una del pueblo A o del pueblo B. El incremento se debe a los costos de traslados de un hogar a otro.

El costo por observación en cada pueblo se ha asignado en 9-dólares, ie. $C_1 = C_2 = 9$, mientras que los costos por observación en el área rural se han estimado en 16-dólares, ie. $C_3 = 16$.

Hallar el tamaño de muestra total $n$ y las respectivas asignaciones a cada estrato, ie. $n_1, n_2$ y $n_3$ que permitan a la empresa estimar, con el mínimo costo, el tiempo promedio que se ve TV en los hogares de la región, con un límite para el error de estimación de 2-horas, es decir, con una varianza para el estimador de:

$$B = 2\sqrt{Var[\overline{y}_{est}]} = 2 \ , \ \text{ o bien } \ Var[\overline{y}_{est}] = 1 = D$$

En este caso el tamaño de muestra $n$ se obtiene, reemplazando a $w_h$ por:

$$w_h = \frac{\frac{N_h\sigma_h}{\sqrt{C_h}}}{\sum_{h=1}^{H}\left(\frac{N_h\sigma_h}{\sqrt{C_h}}\right)}$$ en:

$$n = \frac{\sum_{h=1}^{H}\left(N_h^2\sigma_h^2/w_h\right)}{N^2D + \sum_{h=1}^{H} N_h\sigma_h^2}$$

En este caso se tiene que:

$$\sum_{h=1}^{3}\frac{N_h\sigma_h}{\sqrt{C_h}} = \frac{N_1\sigma_1}{\sqrt{C_1}} + \frac{N_2\sigma_2}{\sqrt{C_2}} + \frac{N_3\sigma_3}{\sqrt{C_3}}$$

$$= \frac{(155)5}{\sqrt{9}} + \frac{(62)15}{\sqrt{9}} + \frac{(93)10}{\sqrt{16}}$$

$$= 800.83$$

luego,

$$w_1 = \left( \frac{\frac{N_1 \sigma_1}{\sqrt{C_1}}}{\sum_{h=1}^{3} \left( \frac{N_h \sigma_h}{\sqrt{C_h}} \right)} \right) = \left[ \frac{155(5)/3}{800.83} \right] = 0.322795618 = 0.32$$

$$w_2 = \left( \frac{\frac{N_2 \sigma_2}{\sqrt{C_2}}}{\sum_{h=1}^{3} \left( \frac{N_h \sigma_h}{\sqrt{C_h}} \right)} \right) = \left[ \frac{62(15)/3}{800.83} \right] = 0.387354742 = 0.39$$

$$w_3 = \left( \frac{\frac{N_3 \sigma_3}{\sqrt{C_3}}}{\sum_{h=1}^{3} \left( \frac{N_h \sigma_h}{\sqrt{C_h}} \right)} \right) = \left[ \frac{93(10)/4}{800.83} \right] = 0.290516056 = 0.29$$

De donde para aplicar:

$$n = \frac{\sum_{h=1}^{H} \left(N_h^2 \sigma_h^2 / w_h\right)}{N^2 D + \sum_{h=1}^{H} N_h \sigma_h^2}$$

primero hallamos:

$$\sum_{h=1}^{3} N_h^2 \sigma_h^2 / w_h = \frac{N_1^2 \sigma_1^2}{w_1} + \frac{N_2^2 \sigma_2^2}{w_2} + \frac{N_3^2 \sigma_3^2}{w_3}$$

$$= \frac{(155)^2 (5)^2}{0.32} + \frac{(62)^2 (15)^2}{0.39} + \frac{(93)^2 (10)^2}{0.29} = 7077059.226$$

$$\sum_{h=1}^{H} N_h \sigma_h^2 = N_1 \sigma_1^2 + N_2 \sigma_2^2 + N_3 \sigma_3^2$$

$$= 155(5)^2 + 62(25)^2 + 93(10)^2 = 27125, \quad \text{y}$$

$$N^2 D^2 = (310)^2 (1) = 96100$$

$$n = \frac{\sum_{h=1}^{H} \left( N_h^2 \sigma_h^2 / w_h \right)}{N^2 D + \sum_{h=1}^{H} N_h \sigma_h^2} = \frac{7077059{,}226}{96100 + 27125} = 57.43200832 = 58$$

Ahora, la asignación para cada estrato sería:

$$n_1 = n w_1 = (58)(0.32) = 18.5 \approx 18$$

$$n_2 = n w_2 = (58)(0.39) = 22.6 \approx 23$$

$$n_3 = n w_3 = (58)(0.29) = 16.8 \approx 17$$

Claramente,

$$n = n_1 + n_2 + n_3 = 18 + 23 + 17 = 58$$

En este caso el costo del muestreo es:

$$c_1(n_1) + c2(n_2) + c_3(n_3) = 9(18) + 9(23) + 16(17) = 641$$

**Ejemplo:** Supongamos que la empresa publicitaria sólo cuenta con los recursos de 500 dólares para gastar en el muestreo, Cuál deberá ser el tamaño de muestra en este caso?

**Solución:** La asignación de la muestra continua siendo la misma, ie.

$w_1 = 0.32$, $w_2 = 0.39$ y $w_3 = 0.29$.

Como el costo total debe ser igual a 500, entonces de debe cumplir:

$$c_1 n_1 + c_2 n_2 + c_3 n_3 = 500$$

es decir:

$$9n_1 + 9n_2 + 16n_3 = 500$$

Además, como: $n_i = na_i$, entonces se tiene:

$$9na_1 + 9na_2 + 16na_3 = 500 \iff 9n(0.32) + 9n(0.39) + 16n(0.29) = 500$$

de donde,

$$11.03n = 500 \iff n = \frac{500}{11.03} = 45.33 \approx 45$$

NOTA: se deben tomar 45 (y no 46), para asegurar que el costo sea menor a 500.

La asignación correspondiente está dada por:

$$\begin{aligned} n_1 &= na_1 = 45(0.32) = 14 \\ n_2 &= na_2 = 45(0.39) = 18 \\ n_3 &= na_3 = 45(0.29) = 13 \end{aligned}$$

Con esta muestra los recursos totales gastados son de:

$$9n_1 + 9n_2 + 16n_3 \iff 9(14) + 9(18) + 16(13) = 496 \leq 500$$

## CONTINUACIÓN EJEMPLO ANTERIOR:

La firma publicitaria decide utilizar entrevistas por teléfono en lugar de entrevistas personales, porque todos los hogares de la región tienen teléfono y este método reduce los costos del estudio.

Es decir, que el costo de obtener una observación es el mismo en lso tres estratos, ie. $C_1 = C_2 = C_3 = C$

Hallar el tamaño de muestra total $n$ y las respectivas asignaciones a cada estrato, ie. $n_1, n_2$ y $n_3$ que permitan a la empresa estimar el tiempo promedio que se ve TV en los hogares de la región, con un límite para el error de estimación de 2-horas, es decir, con una varianza para el estimador de:

$$B = 2\sqrt{Var[\overline{y}_{est}]} = 2\ , \quad \text{o bien} \quad Var[\overline{y}_{est}] = 1 = D$$

En este caso el tamaño de muestra $n$ se obtiene, reemplazando a $w_h$:

$$w_h = \frac{N_h \sigma_h}{\sum_{h=1}^{H} N_h \sigma_h}, \quad \text{en :} \quad n = \frac{\sum_{h=1}^{H} \left( N_h^2 \sigma_h^2 / w_h \right)}{N^2 D + \sum_{h=1}^{H} N_h \sigma_h^2}$$

En este caso se tiene que:

$$\sum_{h=1}^{3} N_h\sigma_h = N_1\sigma_1 + N_2\sigma_2 + N_3\sigma_3$$

$$= (155)5 + (62)15 + (93)10 = 2635$$

luego,

$$w_1 = \frac{N_1\sigma_1}{\sum_{h=1}^{3} N_h\sigma_h} = \left[\frac{155(5)}{2635}\right] = 0.294117647 = 0.30$$

$$w_2 = \frac{N_2\sigma_2}{\sum_{h=1}^{3} N_h\sigma_h} = \left[\frac{62(15)}{2635}\right] = 0.352941176 = 0.35$$

$$w_3 = \frac{N_3\sigma_3}{\sum_{h=1}^{3} N_h\sigma_h} = \left[\frac{93(10)}{2635}\right] = 0.352941176 = 0.35$$

De donde para aplicar:

$$n = \frac{\sum_{h=1}^{H}\left(N_h^2\sigma_h^2/w_h\right)}{N^2D + \sum_{h=1}^{H} N_h\sigma_h^2}$$

primero hallamos:

$$\sum_{h=1}^{3} N_h^2\sigma_h^2/w_h = \frac{N_1^2\sigma_1^2}{w_1} + \frac{N_2^2\sigma_2^2}{w_2} + \frac{N_3^2\sigma_3^2}{w_3}$$

$$= \frac{(155)^2(5)^2}{0.30} + \frac{(62)^2(15)^2}{0.35} + \frac{(93)^2(10)^2}{0.25} = 6944369.048$$

$$\sum_{h=1}^{H} N_h\sigma_h^2 = N_1\sigma_1^2 + N_2\sigma_2^2 + N_3\sigma_3^2$$

$$= 155(5)^2 + 62(25)^2 + 93(10)^2 = 27125, \quad \text{y}$$

$$N^2D^2 = (310)^2(1) = 96100$$

$$n = \frac{\sum_{h=1}^{H}\left(N_h^2\sigma_h^2/w_h\right)}{N^2D + \sum_{h=1}^{H} N_h\sigma_h^2} = \frac{6944369.048}{96100 + 27125} = 56.35519617 = 57$$

Ahora, la asignación para cada estrato sería:

$$n_1 = nw_1 = (57)(0.30) = 17.1 \approx 17$$

$$n_2 = nw_2 = (57)(0.35) = 19.95 \approx 20$$

$$n_3 = nw_3 = (57)(0.35) = 19.95 \approx 20$$

Claramente,

$$n = n_1 + n_2 + n_3 = 17 + 20 + 10 = 57$$

En este caso el costo del muestreo es:

$$c(n_1) + c(n_2) + c(n_3) = 9(17 + 20 + 20) = c(57)$$

## CONTINUACIÓN EJEMPLO ANTERIOR:

Considere que las varianzas aproximadas que se utilizaron anteriormente, ie. $\sigma_1^2 \approx 25, \sigma_2^2 \approx 225$ y $\sigma_3^2 \approx 100$, son erróneas y suponga que dichas varianzas son iguales, ie. $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma^2 = 10$.

Nuevamente se efectúan entrevistas por teléfono, ie. $C_1 = C_2 = C_3 = C$

Hallar el tamaño de muestra total $n$ y las respectivas asignaciones a cada estrato, ie. $n_1, n_2$ y $n_3$ que permitan a la empresa estimar el tiempo promedio que se ve TV en los hogares de la región, con un límite para el error de estimación de 2-horas, es decir, con una varianza para el estimador de:

$$B = 2\sqrt{Var[\overline{y}_{est}]} = 2\ , \quad \text{o bien} \quad Var[\overline{y}_{est}] = 1 = D$$

En este caso el tamaño de muestra $n$ se obtiene, reemplazando a $w_h$:

$$w_h = \frac{N_h}{\sum_{h=1}^{H} N_h}, \quad \text{en :} \quad n = \frac{\sum_{h=1}^{H}\left(N_h^2\sigma_h^2/w_h\right)}{N^2 D + \sum_{h=1}^{H} N_h\sigma_h^2}$$

En este caso se tiene que:

$$\sum_{h=1}^{3} N_h = N_1 + N_2 + N_3 = 155 + 62 + 93 = 310$$

luego,

$$w_1 = \frac{N_1}{\sum_{h=1}^{3} N_h} = \left[\frac{155}{310}\right] = 0.50$$

$$w_2 = \frac{N_2}{\sum_{h=1}^{3} N_h} = \left[\frac{62}{310}\right] = 0.20$$

$$w_3 = \frac{N_3}{\sum_{h=1}^{3} N_h} = \left[\frac{93}{310}\right] = 0.30$$

De donde para aplicar:

$$n = \frac{\sum_{h=1}^{H}\left(N_h^2\sigma_h^2/w_h\right)}{N^2D + \sum_{h=1}^{H} N_h\sigma_h^2}$$

primero hallamos:

$$\sum_{h=1}^{3} N_h^2\sigma_h^2/w_h = \frac{N_1^2\sigma_1^2}{w_1} + \frac{N_2^2\sigma_2^2}{w_2} + \frac{N_3^2\sigma_3^2}{w_3}$$

$$= \frac{(155)^2(5)^2}{0.50} + \frac{(62)^2(15)^2}{0.20} + \frac{(93)^2(10)^2}{0.30} = 8408750$$

$$\sum_{h=1}^{H} N_h\sigma_h^2 = N_1\sigma_1^2 + N_2\sigma_2^2 + N_3\sigma_3^2$$

$$= 155(5)^2 + 62(25)^2 + 93(10)^2 = 27125, \quad \text{y}$$

$$N^2D^2 = (310)^2(1) = 96100$$

$$n = \frac{\sum_{h=1}^{H}\left(N_h^2\sigma_h^2/w_h\right)}{N^2 D + \sum_{h=1}^{H} N_h\sigma_h^2} = \frac{8408750}{96100 + 27125} = 68.23899371 = 69$$

Ahora, la asignación para cada estrato sería:

$$n_1 = nw_1 = (69)(0.50) = 34.5 \approx 34$$

$$n_2 = nw_2 = (69)(0.20) = 13.8 \approx 14$$

$$n_3 = nw_3 = (69)(0.30) = 20.7 \approx 21$$

Claramente,

$$n = n_1 + n_2 + n_3 = 34 + 14 + 21 = 69$$

En este caso el costo del muestreo es:

$$c(n_1) + c(n_2) + c(n_3) = 9(34 + 14 + 21) = c(69)$$

## EJEMPLO PARA PROPORCIONES:

Los siguientes datos se obtuvieron de una muestreo realizado por la empresa publicitaria en años anteriores.

| Estrato | Tamaño n | Número de hogares donde se ve un programa X | $\widehat{p}_h$ |
|---|---|---|---|
| 1 | $n_1 = 20$ | 16 | 0.8 |
| 2 | $n_2 = 8$ | 2 | 0.25 |
| 3 | $n_1 = 12$ | 6 | 0.5 |

La empresa desea efectuar un nuevo estudio en la misma región para estimar la proporción de hogares que ven el programa $X$. Aunque las fracciones $P_h$-son desconocidas, se pueden aproximar mediante las estimaciones del estudio realizado anteriormente, ie. $P_1 = \widehat{p}_1 \approx 0{,}8$, $P_2 = \widehat{p}_2 \approx 0{,}25$ y $P_3 = \widehat{p}_3 \approx 0{,}5$. Los costos de obtener un observación en cada pueblo y en el área rural son nuevamente: $C_1 = C_2 = 9$ y $C_3 = 16$. El número de hogares dentro de cada estrato es: $N_1 = 155$, $N_2 = 62$ y $N_3 = 93$.

La empresa desea estimar la proporción $P$-de hogares en la región que ven el programa $X$ con un límite para el error de estimación de $B = 0.1$.

Hallar el tamaño de muestra $n$ y la respectiva asignación a cada uno de los estratos, ie. $n_1, n_2$ y $n_3$, de tal forma que el costo de muestreo sea mínimo.

En este caso se reemplazan las $\sigma_h^2$ por: $p_h q_h = p_h(1 - p_h)$.

Para hallar las fracciones $w_h$:

$$w_h = \frac{\frac{N_h \sigma_h}{\sqrt{C_h}}}{\sum_{h=1}^{H} \left(\frac{N_h \sigma_h}{\sqrt{C_h}}\right)} = \frac{N_h \sqrt{p_h q_h / C_h}}{\sum_{h=1}^{H} N_h \sqrt{p_h q_h / C_h}}$$

primero hallamos:

$$\sum_{h=1}^{H} N_h \sqrt{\frac{p_h q_h}{C_h}} = N_1 \sqrt{\frac{p_1 q_1}{C_1}} + N_2 \sqrt{\frac{p_2 q_2}{C_2}} + N_3 \sqrt{\frac{p_3 q_3}{C_3}}$$

$$= 155 \sqrt{\frac{(0.8)(0.2)}{9}} + 62 \sqrt{\frac{(0.25)(0.75)}{9}} + 93 \sqrt{\frac{(0.5)(0.5)}{16}}$$

$$= \frac{62.000}{3} + \frac{26.846}{3} + \frac{46.500}{4}$$

$$= 20.667 + 8.9949 + 11.625$$

$$= 41.241.$$

Luego,

$$w_1 = \frac{N_1\sqrt{p_1 q_1/C_1}}{\sum_{h=1}^{3} N_h\sqrt{P_h Q_h/C_h}} = \left(\frac{20.667}{41.241}\right) = 0.50$$

$$w_2 = \frac{N_2\sqrt{p_2 q_2/C_2}}{\sum_{h=1}^{H} N_h\sqrt{P_h Q_h/C_h}} = \left(\frac{8.949}{41.241}\right) = 0.22$$

$$w_3 = \frac{N_3\sqrt{p_3 q_2/C_2}}{\sum_{h=1}^{H} N_h\sqrt{P_h Q_h/C_h}} = \left(\frac{11.625}{41.241}\right) = 0.28$$

Luego las $w_h$ son:

$$w_1 = 0.5 \quad , \quad w_2 = 0.22 \quad , \quad w_3 = 0.28$$

Ahora, para hallar a $n$ se tiene:

$$n = \frac{\sum_{h=1}^{H} \left( N_h^2 p_h q_h / w_h \right)}{N^2 D + \sum_{h=1}^{H} N_h p_h q_h}$$

Primero hallamos:

$$\sum_{h=1}^{H} \left( N_h^2 p_h q_h / w_h \right) = \frac{N_1^2 p_1 q_1}{w_1} + \frac{N_2^2 p_2 q_2}{w_2} + \frac{N_3^2 p_3 q_3}{w_3}$$

$$= \frac{(155)^2 (0{,}8)(0{,}2)}{0{,}5} + \frac{(62)^2 (0{,}25)(0{,}75)}{0{,}22} + \frac{(93)^2 (0{,}5)(0{,}5)}{0{,}28}$$

$$= 18686.46$$

$$\sum_{h=1}^{H} N_h P_h Q_h = N_1 p_1 q_1 + N_2 p_2 q_2 + N_3 p_3 q_3$$

$$= 155(0.8)(0.2)+62(0.25)(0.75)+93(0.5)(0.5) \; = 59.675$$

Ahora, como $B = 0{,}1$, entonces

$$D = \frac{B^2}{4} = \frac{(0.1)^2}{4} = 0.0025 = Var[\widehat{p}_{est}],$$

luego,

$$n = \frac{\sum_{h=1}^{H}\left(N_h^2 p_h q_h / w_h\right)}{N^2 D + \sum_{h=1}^{H} N_h p_h q_h} = \frac{18686.46}{(310)^2(0.0025) + 59.675} = 62.3 \approx 63$$

y por tanto:

$$n_1 = nw_1 = 63(0.5) = 31\ , \quad n_2 = nw_2 = 63(0.22) = 14$$

$$\text{y} \quad n_1 = nw_3 = 63(0.28) = 18$$