



# TAMAÑO DE MUESTRA

**M<sup>a</sup> Purificación Galindo Villardón**

[pgalindo@usal.es](mailto:pgalindo@usal.es)

# DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

Para determinar el tamaño de la muestra es necesario que precisemos:

- El error de estimación

$$|\hat{p} - p| \leq E$$

- El nivel de confianza

$$[\text{prob}(\hat{p} - p) < E] = 1 - \alpha$$

# DETERMINACIÓN DEL TAMAÑO DE LA MUESTRA

**Por Error de estimación**, se entiende la máxima diferencia admitida entre el verdadero valor del parámetro y el valor estimado a partir de la muestra.

$$|\hat{p} - p| \leq E$$

*El nivel de error suele estar entre un 3 y un 5%.*

Por **nivel de confianza** se entiende la proporción de veces que, en un muestreo repetido, se requerirá que el error de estimación sea menor que la cantidad prefijada. Generalmente esta información se expresa en términos probabilísticos,

*Los valores más comúnmente aceptados por la comunidad científica son el 95 o el 99%.*

$$1 - \alpha$$

# TAMAÑO DE MUESTRA UN MUESTREO ESTRATIFICADO / CUOTAS

## Determinación del tamaño de la muestra

- ❖ 1. *El primer problema a resolver será la estimación del tamaño adecuado de la muestra para que ésta sea representativa.*
- ❖ 2. *El segundo problema es estimar el tamaño de cada estrato*

# POBLACIONES GRANDES (INFINITAS)

# CALCULO DEL TAMAÑO MUESTRAL

## POBLACIONES GRANDES

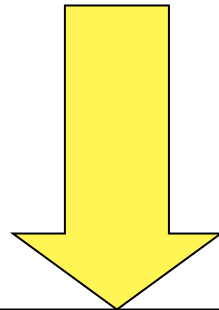
$$I_p^{1-\alpha} = \left[ \hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n}} \right]$$

$$E = z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

# CALCULO DEL TAMAÑO MUESTRAL

## POBLACIONES GRANDES (infinitas)

$$E = z_{\alpha/2} \sqrt{\frac{pq}{n}}$$

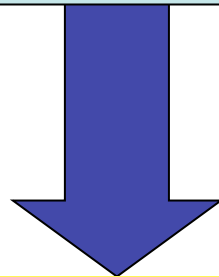


$$n = \frac{z_{\alpha/2}^2 p q}{E^2}$$

# CALCULO DEL TAMAÑO MUESTRAL

## POBLACIONES GRANDES (infinitas)

$$n = \frac{z_{\alpha/2}^2 p q}{E^2}$$



Tomando:  $p=q=0.5$   
Nivel de confianza=  $95.5\%$

$$n = \frac{1}{E^2}$$



# POBLACIONES FINITAS

# CALCULO DEL TAMAÑO MUESTRAL

## POBLACIONES FINITAS

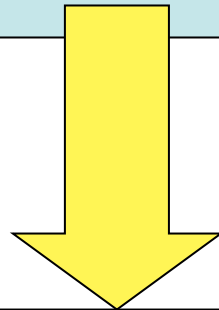
$$I_p^{1-\alpha} = \left[ \hat{p} \pm z_{\alpha/2} \sqrt{\frac{pq}{n} \frac{N-n}{N-1}} \right]$$

$$E = z_{\alpha/2} \sqrt{\frac{pq}{n} \frac{N-n}{N-1}}$$

# CALCULO DEL TAMAÑO MUESTRAL

## POBLACIONES FINITAS

$$E = z_{\alpha/2} \sqrt{\frac{pq}{n} \frac{N-n}{N-1}}$$

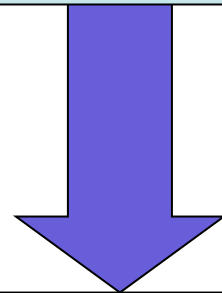


$$n = \frac{z_{\alpha/2}^2 N p q}{E^2 (N-1) + z_{\alpha/2}^2 p q}$$

# CALCULO DEL TAMAÑO MUESTRAL

## POBLACIONES FINITAS

$$n = \frac{z_{\alpha/2}^2 N p q}{E^2 (N - 1) + z_{\alpha/2}^2 p q}$$



Tomando:  $p=q=0.5$   
Nivel de confianza= **95.5%**

$$z_{\alpha/2}^2 p q = 1$$

$$n = \frac{N}{E^2 (N - 1) + 1}$$

# ERROR MUESTRAL Y TAMAÑO DE MUESTRA

La relación entre el error elegido y el tamaño de muestra necesario para conseguirlo no es lineal, sino **hiperbólica**; por tanto, *pequeñas disminuciones en el error producen aumentos enormes en el tamaño de muestra con el consiguiente incremento en el presupuesto.*

*Esto implica que la disminución en el error sobre los niveles usuales debe estar perfectamente justificada.*

## Población Finita

N	Error	Tamaño muestra
57022	0,01	8508
	0,02	2395
	0,021	2181
	0,022	1994
	0,023	1830
	0,024	1685
	0,025	1556
	0,027	1339
	0,029	1165
	0,03	1090
	0,031	1000
	0,032	960
	0,033	904
	0,034	852
	0,035	805
	0,04	618
	0,05	397
	0,06	276
	0,07	203
	0,08	156

## Población Infinita

Tamaño muestra
10000
2500
2268
2066
1890
1736
1600
1372
1189
1111
1017
977
918
865
816
625
400
278
204
156

# UN EJEMPLO

*Supongamos que queremos realizar un estudio para valorar el nivel socioeconómico de las familias de la ciudad de Salamanca.*

Según el censo de 1981, el nº de familias salmantinas es de 48069.

*Fijaremos la atención en una pregunta cualquiera ya que a los efectos cuantitativos señalados, la problemática sería igual para cualquier otra.*

**¿El cabeza de familia trabaja?**

# Tipo de muestreo

## ***TIPO DE MUESTREO: ESTRATIFICADO ALEATORIO***

De acuerdo con el objetivo planteado, Salamanca no puede considerarse homogénea, por tanto parece razonable dividirla en **estratos** que podrían ser los distintos **barrios** de la ciudad. Proponemos, por tanto, un **muestreo aleatorio estratificado**, (cada uno de los 11 barrios será un estrato) en los cuales se entrevistaría un determinado número de familias salmantinas .

## ***TIPO DE AFIJACION: PROPORCIONAL***

Cada uno de los barrios de la ciudad tiene un censo de familias muy diferente, por tanto el número de encuestas a pasar en cada barrio será proporcional al tamaño del estrato.



# ESTIMACIÓN DEL TAMAÑO DE MUESTRA ADECUADO

El censo de familias de Salamanca según los datos manejados era de 48069 familias.

Error de estimación prefijado = 4%

Nivel de confianza = 95.44 %

$$1 - \alpha = 0,9544 \quad z_{\alpha/2} = 2$$

$$p = 0,5 \quad E = 0,04$$

$$z_{\alpha/2}^2 pq = 1 \quad n = \frac{N}{E^2(N-1) + 1}$$

$$n = \frac{48069}{0,04^2 48068 + 1} = 617$$

# DISTRIBUCION EN LOS ESTRATOS

<i>Barrio</i>	<i>N° de familias</i>	<i>% del total</i>	<i>Entrevistas</i>
<b>Centro-Mercado San Juan</b>	6632	13.79	<b>85</b>
<b>Gran Via-Canalejas</b>	3539	7.34	<b>45</b>
<b>Antiguo</b>	856	1.77	<b>11</b>
<b>Prosperidad-Rollo</b>	5561	11.56	<b>71</b>
<b>Salas Pombo</b>	3318	6.90	<b>43</b>
<b>Carmelitas-Oeste</b>	4330	9.00	<b>56</b>
<b>Pizarrales</b>	3786	7.87	<b>49</b>
<b>Vidal</b>	3104	6.45	<b>40</b>
<b>Garrido</b>	13690	28.47	<b>176</b>
<b>Tejares</b>	976	2.02	<b>12</b>
<b>San Jose-La Vega</b>	2285	4.75	<b>29</b>
<b>TOTAL</b>	48069	100.00	<b>617</b>

# ERROR EN LOS ESTRATOS

*Error global: 4%*

<i>Barrio</i>	<i>N<sub>i</sub></i>	<i>n<sub>i</sub></i>	<i>Error (%)</i>
Centro-Mercado San Juan	6632	85	10,78
Gran Via-Canalejas	3539	45	14,80
Antiguo	856	11	29,01
Prosperidad-Rollo	5561	71	11,80
Salas Pombo	3318	43	15,24
Carmelitas-Oeste	4330	56	13,35
Pizarrales	3786	49	14,27
Vidal	3104	40	15,80
Garrido	13690	176	7,50
Tejares	976	12	29,00
San Jose-La Vega	2285	29	18,56
<b>TOTAL</b>	48069	617	4,00

$$E_i = z_{\alpha/2} \sqrt{\frac{p_i q_i}{n_i} \frac{N_i - n_i}{N_i - 1}}$$