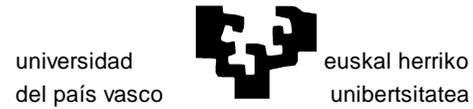


eman ta zabal zazu



**Departamento de Economía financiera II**  
(ECONOMÍA FINANCIERA Y CONTABILIDAD, COMERCIALIZACIÓN E INVESTIGACIÓN DE MERCADOS)

## INVESTIGACIÓN COMERCIAL: INTRODUCCIÓN

### TEMA 6:

### *INTRODUCCIÓN AL MUESTREO*

Jon Charterina Abando

## 6.1.- Conveniencia y limitaciones del muestreo

- Objeto de la IC: Obtener información a cerca de características propias de la población que se desee estudiar
- **Población:** conjunto de elementos que comparten una serie de características y que representan el universo de cuyo estudio es objeto el trabajo de investigación
- La información sobre dicha población se podría obtener de dos maneras:
  - Realizando un censo, estudiando todos los elementos que la componen
  - Trabajando con muestras

- **Censo:** Enumeración completa de todos los elementos de una población.

Los parámetros de toda la población podrían calcularse directamente después de haber estudiado su valor en todos los individuos

- **Muestra:** Subgrupo de la población seleccionado y estudiado para obtener los parámetros de ésta a través de la inferencia estadística.

Las conclusiones serán válidas para toda la población si se ha escogido una muestra representativa

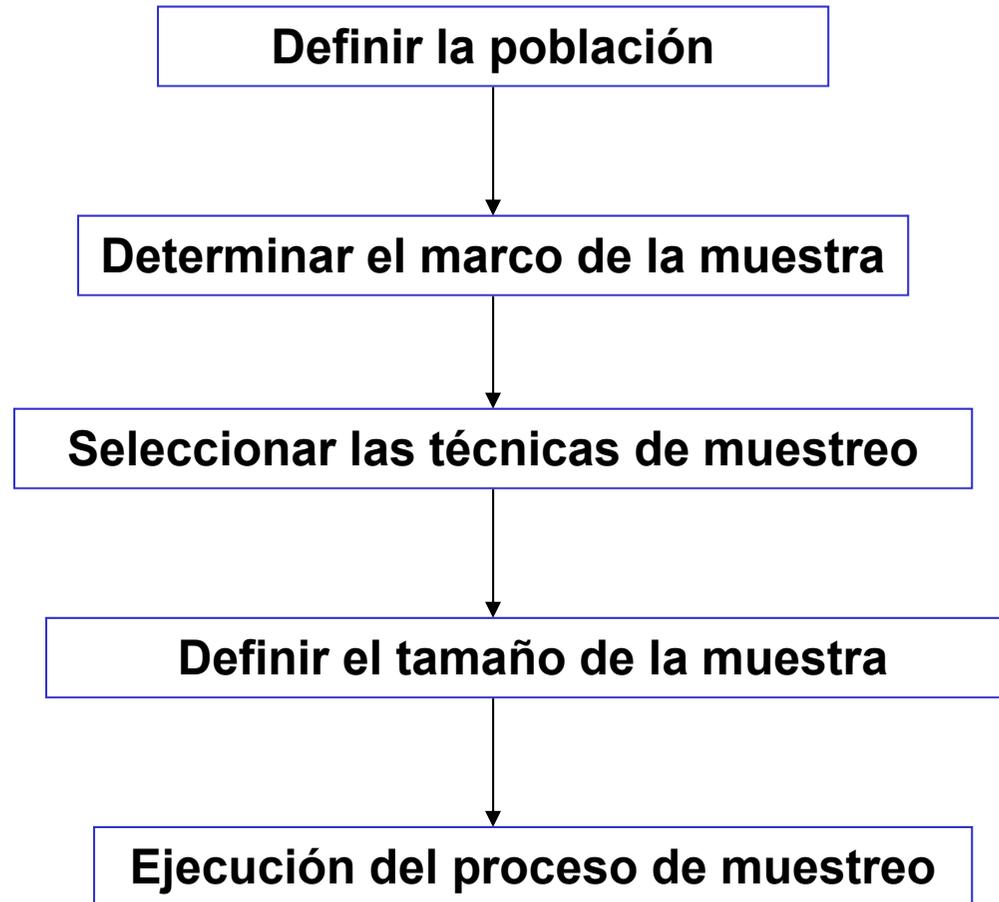
## Censo o muestra ¿de qué depende?

- **Tamaño de la población:** Si es muy grande el censo es inabordable. Si es muy pequeña no merece la pena hacer muestreo
- **Variabilidad de los miembros de la población:** Si las variables a estudiar tienen grandes variaciones mejor realizar censo
- **Razones temporales:** Si la población es muy grande o dispersa un censo se tarda mucho en hacer. Se producirían errores sistemáticos por obsolescencia de las primeras observaciones
- **Razones de definición:** Si no se sabe bien quien integra la población no se puede trabajar con censos

- **Tipos de errores que se pueden cometer:**
  - **Errores muestrales:**
    - No pueden ser eliminados, se deben al hecho de trabajar con muestras
    - Se pueden acotar
  - **Errores ajenos al muestreo (sistemáticos):**
    - Debidos a mala definición del problema de investigación, fallos en la redacción de las preguntas del cuestionario, mala selección de los miembros de la muestra, errores de análisis, etc
    - Si la población es muy grande los errores sistemáticos son muy importantes. Por eso se acepta asumir errores aleatorios y trabajar con muestras

	<b>Censo</b>	<b>Muestra</b>
<b>Tamaño de la población</b>	<b>Si es pequeña</b>	<b>Si es grande</b>
<b>Variabilidad de la población</b>	<b>Si es grande</b>	<b>Si es pequeña</b>
<b>Tiempo disponible</b>	<b>Si es grande</b>	<b>Si es escaso</b>
<b>Definición población</b>	<b>Si es buena</b>	<b>Si es mala o inexistente</b>
<b>Error aleatorio</b>	<b>Inexistente</b>	<b>En principio tolerable y acotado</b>
<b>Errores sistemáticos</b>	<b>Mayor riesgo</b>	<b>Menor riesgo</b>

## **6.2.- El proceso de muestreo**



# **1º Definir la población objeto de estudio:**

- **Debe quedar claro quiénes deben ser incluidos y quiénes no**
- **Definir la población en términos de 4 parámetros:**
  - **Los elementos**
  - **Las unidades muestrales**
  - **La extensión o zona**
  - **El tiempo**

- **Elemento**: Es el objeto acerca del cual o a partir del cual se desea obtener una información.

En IC, el elemento de la población suele ser el entrevistado

- **Unidad muestral**: Es un elemento, o una unidad que contiene un elemento, y que está disponible para la selección del proceso de muestreo en algún momento en particular.

Sirve como paso previo para identificar y acceder al elemento de la población.

- Si la entrevista es personal a pie de calle el elemento es igual a la unidad muestral.
- Si p.e. se escoge aleatoriamente el domicilio, este es la unidad muestral y el elemento es cada entrevistado del domicilio

# Conceptos básicos de muestreo

**Variables:** Las variables pueden ser de tres tipos: métricas o cuantitativas, ordinales y nominales.

$U$	Piso	Puerta	En propiedad	Personas
$u_1$	1	Izq.	No	1
$u_2$	1	Dcha.	Sí	4
$u_3$	2	Izq.	Sí	2
$u_4$	2	Dcha.	No	3
...	...	...	...	...
$u_{16}$	8	Dcha.	Sí	1

# Conceptos básicos de muestreo

**Parámetro:** Un parámetro es todo valor que describe de manera resumida una población.

Total de personas que viven en cada portal:  $T = \sum_1^N X_i = 1 + 4 + 2 + 3 + \dots + 1 = 39$

Media de personas que viven en cada vivienda:  $\mu = T / N = 39 / 16 = 2,44$

Total de clase C de viviendas en propiedad:  $T = \sum_1^N Y_i = 0 + 1 + 1 + 0 + \dots + 1 = 11$

Proporción de viviendas en propiedad:  $P = C / N = 11 / 16 = 0,6875$

# Conceptos básicos de muestreo

**Estadístico o estimador:** Es una función de los valores muestrales, una descripción resumida de la muestra. A diferencia de los parámetros, los estadísticos son aleatorios. Por ejemplo, si estimáramos el total de clase (recuento de propietarios de vivienda) y la proporción:

**Estimación del total de clase C:** 
$$\hat{C} = N \cdot p = N \frac{1}{n} \sum_{i=1}^n y_i$$

Donde:

N: Número de elementos de la población

n: número de elementos de la muestra

p: proporción de los elementos de la muestra que pertenecen a la clase C

$y_i$ : variable nominal que indica la pertenencia a la clase C, tal que:

$$y_i = 1 \quad \text{si } u_i \in C$$

$$y_i = 0 \quad \text{si } u_i \notin C$$

Si la muestra es una planta del portal, para el piso 1,  $p = (0 + 1)/2 = 0,5$

Si se mantiene el tamaño muestral  $n$  en 2,  $p$  podría oscilar entre  $\{0; 0.5 ; 1\}$

# Conceptos básicos de muestreo

La *aleatoriedad* de los estadísticos dependerá del proceso de muestreo que se siga. En el ejemplo de los vecinos del portal, suponiendo que éste fuese toda la población, **¿cuántas muestras de 8 viviendas sería posible extraer?**

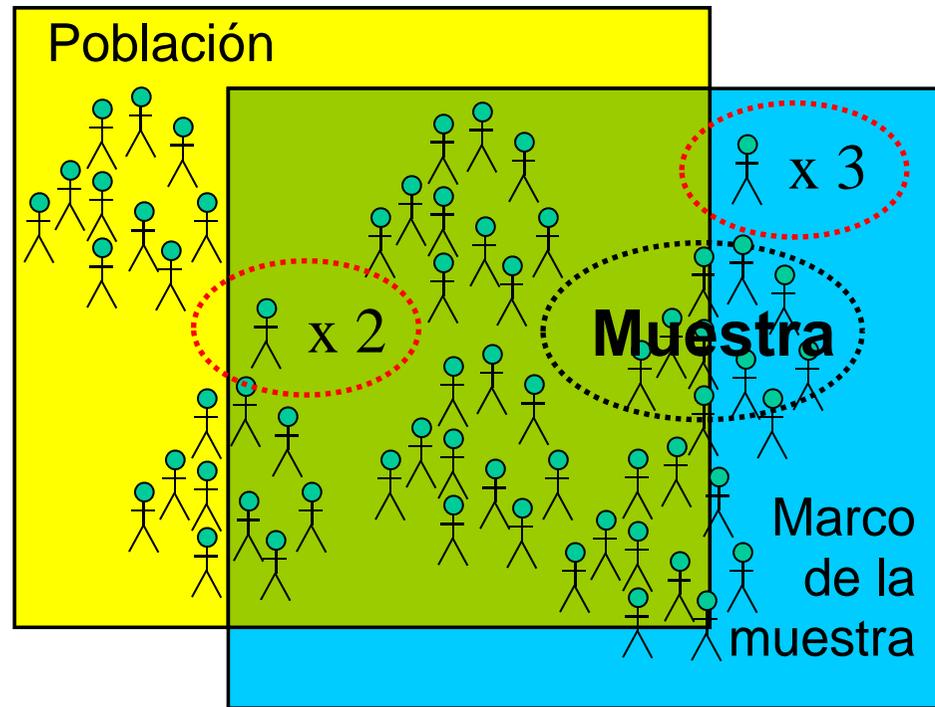
$n$	$\binom{16}{n} = \frac{16!}{(16-n)! \cdot n!}$
1	16
2	120
3	560
4	1820
5	4368
6	8008
7	11440
8	12870
9	11440
10	8008
11	4368
12	1820
13	560
14	120
15	16
16	1

**El número de combinaciones de 16 tomados de  $n$  en  $n$**

## **2º Determinar el marco de muestreo o marco de la muestra**

- **El marco de muestreo es la representación de los elementos de la población objetivo**
- **Es una lista de los elementos de la población, o bien un conjunto de instrucciones para identificar a ésta**
- **Ejemplo: para la entrevistas telefónicas en una provincia, un marco de muestreo es el listín**
- **Ejemplo: si la población es el conjunto de vecinos de una ciudad, un marco de muestreo es el padrón**
- **En el marco del muestreo, puede que no estén todos los que son, ni sean todos los que están**

## Problemas del muestreo en la práctica



**Desfases entre la población  
y el marco de la muestra**

## El error aleatorio de la muestra

**Ejemplo: Precios estimados por 40 personas para un producto particular**

P1 = 1.200	P9 = 1.500	P17 = 4.100	P25 = 2.500	P33 = 1.430
P2 = 1.340	P10 = 2.400	P18 = 3.110	P26 = 3.000	P34 = 1.900
P3 = 2.150	P11 = 3.200	P19 = 4.980	P27 = 3.540	P35 = 1.680
P4 = 3.200	P12 = 3.400	P20 = 4.270	P28 = 2.600	P36 = 2.150
P5 = 2.130	P13 = 4.900	P21 = 2.180	P29 = 2.750	P37 = 1.430
P6 = 3.210	P14 = 1.750	P22 = 2.890	P30 = 3.190	P38 = 2.160
P7 = 2.240	P15 = 1.800	P23 = 2.200	P31 = 3.600	P39 = 2.190
P8 = 3.150	P16 = 2.450	P24 = 2.000	P32 = 5.000	P40 = 2.240

**Media:  $\mu = 2.678 \text{ €}$**

**Desv.típica:  $\sigma = 994$**

## El error aleatorio de la muestra

Eligiendo muestras de 5 entrevistados al azar, tenemos las siguientes medias (muestrales):

$$m1 = \{P1, P7, P10, P32, P21\} = 2.604 \text{ €}$$

$$m2 = \{P11, P31, P12, P3, P17\} = 3.290 \text{ €}$$

$$m3 = \{P21, P17, P12, P35, P40\} = 2.720 \text{ €}$$

$$m4 = \{P13, P7, P30, P3, P38\} = 2.928 \text{ €}$$

$$m5 = \{P16, P5, P1, P29, P24\} = 2.106 \text{ €}$$

## El error aleatorio de la muestra

Dichas medias muestrales  $m_i$  tendrán una desviación positiva o negativa respecto del precio medio para los 40 entrevistados, esto es, la población( $\mu$ ):

$$e1 = \mu - m1 = 2.678 - 2.604 = 74$$

$$e2 = \mu - m2 = 2.678 - 3.290 = - 612$$

$$e3 = \mu - m3 = 2.678 - 2.720 = - 42$$

$$e4 = \mu - m4 = 2.678 - 2.928 = - 250$$

$$e5 = \mu - m5 = 2.678 - 2.106 = 572$$

A estas desviaciones se les llama sesgo o error.

## El error aleatorio de la muestra

El sesgo o error entre el precio muestral y el precio de la población estará dentro de un intervalo (llamado *intervalo de confianza*) definido entre  $e + e$  y  $e - e$  con una probabilidad igual a  $1 - \alpha$ , denominada como *nivel de confianza*.

$$P \left[ -k \leq \frac{m_i - \mu}{\sigma_{m_i}} \leq k \right] = 1 - \alpha$$

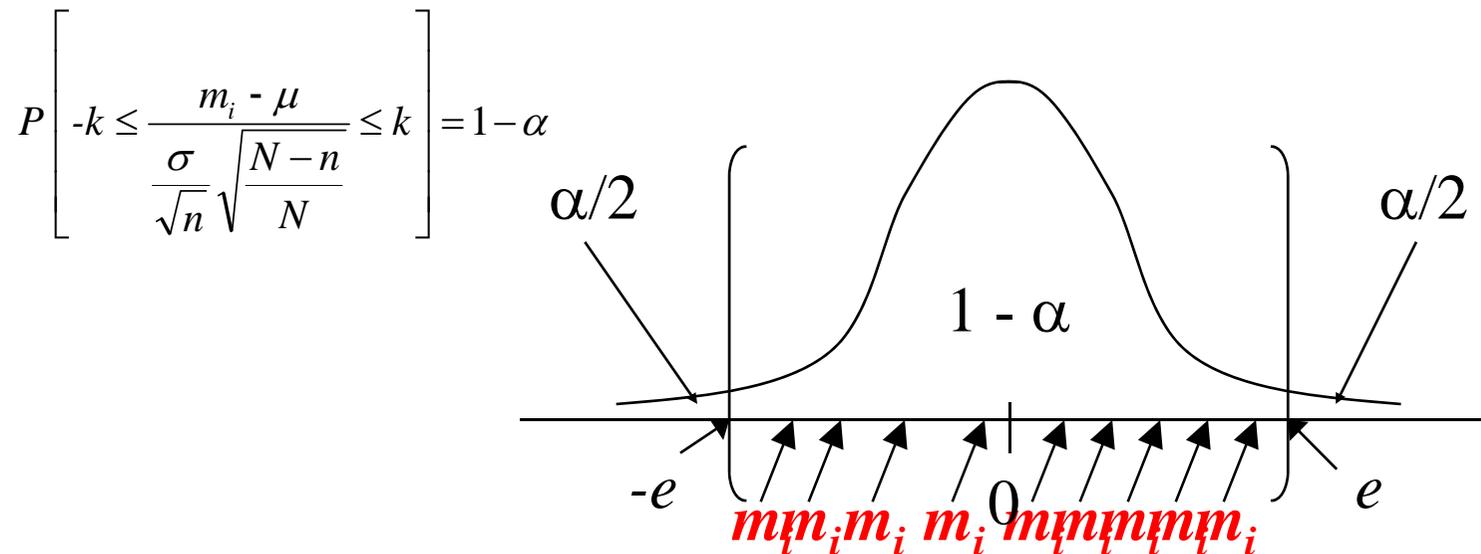
## El error aleatorio de la muestra

El sesgo o error entre el precio muestral y el precio de la población estará dentro de un intervalo (llamado *intervalo de confianza*) definido entre  $e$  y  $-e$  con una probabilidad igual a  $1 - \alpha$ , denominada como *nivel de confianza*.

$$P \left[ -k \leq \frac{m_i - \mu}{\frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N}}} \leq k \right] = 1 - \alpha$$

# El error aleatorio de la muestra

El sesgo o error entre el precio muestral y el precio de la población estará dentro de un intervalo (llamado *intervalo de confianza*) definido entre  $e$  y  $-e$  con una probabilidad igual a  $1 - \alpha$ , denominada como *nivel de confianza*.



## El error aleatorio de la muestra

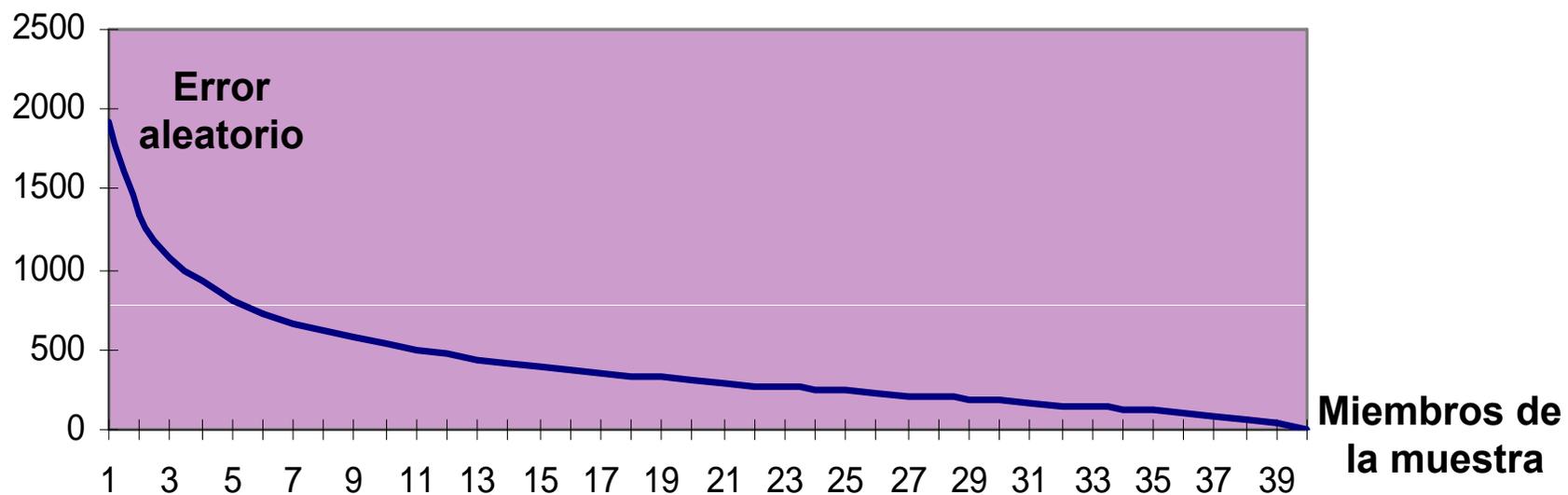
Por tanto, despejando  $\mu$  se obtiene el intervalo de confianza:

$$P \left[ m_i - \underbrace{k \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N}}}_{\text{Error aleatorio de la muestra}} \leq \mu \leq m_i + \underbrace{k \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N-n}{N}}}_{\text{Error aleatorio de la muestra}} \right] = 1 - \alpha$$

Debido al sesgo, no va a poder hallarse un valor concreto, sino dos valores que limitan un intervalo donde se encuentra el verdadero valor  $\mu$ .

# El error aleatorio de la muestra

## Error del muestreo según el número de personas escogidas.



## El error aleatorio de la muestra

En principio, cuantas más referencias de las 5 iniciales se elijan para calcular la media de todas, se irán obteniendo promedios más cercanos a  $\mu$  (2.678€).

En consecuencia, se trata de **saber cuántas observaciones hay que tomar** para no pasar de un nivel de error prefijado.

## **Tamaño óptimo (mínimo) cuando la variable a estimar es una proporción:**

$$n_p = \frac{pqNk^2}{e^2(N-1) + pqk^2}$$

**Cuando la población sea muy grande, se emplea la siguiente relación:**

$$n_0 = \frac{pqk^2}{e^2}$$

- **Si las discrepancias entre la población y el marco muestral son pequeñas, podrían ignorarse**
- **Sino, se deberá reconocer y tratar dicho error**
- **Varias maneras:**
  - **Redefinir la población en términos del marco muestral**
  - **Examinar aquellos elementos escogidos para ver se cumplen con las características exigibles a los elementos de la población.**

### **3º Seleccionar una técnica de muestreo:**

#### **a).- Muestreo no probabilístico:**

- a) La elección de la muestra se hace sin seguir una norma prefijada, de cualquier manera y de forma cómoda**
- b) Es el responsable de la investigación quien dice la forma de escoger los elementos de la muestra y su composición**

## **a.1).- Muestreo de conveniencia:**

- **Técnica sencilla, económica y que consume menos tiempo**
- **La persona que realiza las entrevistas selecciona a su criterio los elementos que han de conformar la muestra**
- **No recomendable en estudios que pretendan hacer inferencia hacia toda la población**
- **Válida para trabajos exploratorios y sondeos pilotos**
- **Ejemplo: estudio de consumo de tabaco entre los jóvenes. Seleccionar por conveniencia ciertos puntos de la ciudad donde se concentran jóvenes y entrevistar a un número determinado elegido sin utilizar ningún criterio**

## **a.2.)- Muestreo de bola de nieve:**

- **Técnica adecuada para poblaciones pequeñas y difíciles de determinar**
- **Consiste en entrevistar a un número de personas que componen la población, y través de éstas indagar otras que se encuentren en las mismas circunstancias, es decir, que formen parte de dicha población**
- **Coste relativamente bajo y facultad de dar con poblaciones raras**
- **Ejemplo: Estudio de coleccionistas de juguetes antiguos. Los primeros captados pueden identificar a otros**

## **a.3.)- Muestreo por cuotas**

### **Dos fases diferenciadas:**

- **Selección de los criterios y asignación de cuotas:**
  - **Desarrollar proporciones o cuotas de subgrupos de la población.**
  - **Estas cuotas dependen de características típicamente demográficas que el investigador selecciona en función de su juicio y experiencia: variables como el sexo, la edad o el lugar de residencia se emplean como criterio**
  - **Las cuotas son asignadas respetando la proporción que guardan a escala poblacional. Para evitar sesgos en la composición**

- **Selección de personas:**
  - El entrevistador escoge a las personas respetando las cuotas que le han sido asignadas

### **Gran ventaja:**

- Dar con una muestra directamente y sin requerir un marco muestral. Procedimiento económico en tiempo y dinero

### **Inconveniente:**

- El entrevistador aplica su criterio propio en la selección de la muestra, lo que puede provocar que lo haga en función de su comodidad

### **Ejemplo:**

- Medir la eficacia de una campaña de la que se considera que sus efectos pueden variar mucho con el sexo. Poblacion: 55% M, 45% H. Utilizar mismas cuotas

## **b) Muestreo probabilístico o aleatorio:**

- **Se escogen a las unidades muestrales mediante un proceso aleatorio o de azar, con lo que podemos calcular la probabilidad de cada una de las unidades, la precisión y los errores cometidos**

## **b.1.)- Muestreo probabilístico con reemplazamiento:**

- **También se llama Muestreo Aleatorio Simple (MAS)**
- **Las unidades son devueltas a la población una vez analizadas, por tanto pueden volver a formar parte de la muestra**
- **Implica que la probabilidad de extracción de un elemento no depende de la extracción anterior de otros elementos**
- **Cuando la población es muy grande la extracción de algunos elementos no afecta a la probabilidad de salida de los que quedan. Se puede asimilar como MAS**

## **b.2.) Muestreo aleatorio sin reemplazamiento:**

- **Cada individuo muestreado no es reintegrado al colectivo**
- **Por tanto, ya no tendrá probabilidad de volver a ser elegido para la muestra**
- **La probabilidad de extracción de un elemento sí depende aquí de la extracción de elementos anteriores**
- **Las fórmulas son más complejas matemáticamente, por esta razón, para poblaciones lo suficientemente grandes, el muestreo sin reemplazamiento se asume igual al muestreo con reemplazamiento**

## **4º Determinar el tamaño de la muestra:**

- **El tamaño de la muestra es el número de elementos de la población que componen la muestra de manera que sea estadísticamente representativa**
- **Determinarlo implica consideraciones de tipo cuantitativo y cualitativo**

## **5º Ejecución del proceso de muestreo:**

- **La ejecución del procedimiento del muestreo viene condicionada por la población, el marco muestral, la unidad muestral, la técnica escogida y el tamaño definido**
- **Habrà que establecer procedimientos para determinar la unidad muestral**

## 6.3.- El muestreo aleatorio simple (M.A.S.)

### 6.3.1- Conceptos previos

- **Población** de **N elementos** identificados
- Se escoge aleatoriamente una **muestra** de **n elementos**
- **Variables**: Nominales, ordinales, y métricas (de intervalo y de razón)
- **Parámetros**: Todo valor que describa de manera resumida la población. Son los valores desconocidos que se quieren hallar a través de la inferencia estadística, mediante los estimadores.  
Ejemplos:
  - La media: la edad media e los estudiantes de una clase.
  - El total: Unidades totales de producto consumidas
  - La proporción: porcentaje de fumadores

- **Estadístico o estimador:** Es una función de los valores muestrales, una descripción resumida de la muestra. Ejemplos:
  - la media de la renta de las personas de la muestra
  - proporción de la muestra que consumen un producto

A diferencia de los parámetros, los estadísticos son aleatorios: no todas las muestras proporcionan el mismo valor para un estadístico.

**Ejemplo:** Clase de 20 personas. Se toma una muestra de 5 personas para estudiar la estatura media. Número posible de combinaciones y por tanto de medias distintas:

$$\binom{N}{n} = \frac{N!}{n!(N-n)!} = \frac{20 \cdot 19 \cdot 18 \cdot 17 \cdot 16}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5} = 15.504$$

## **6.3.2. Determinación de los estimadores en el M.A.S.**

- **Los parámetros de la población que típicamente se busca estimar a través de la muestra son:**
  - **Medias**
  - **Totales**
  - **Proporciones**

## **6.3.3. Tamaño de la muestra en el M.A.S.**

## Estimadores de los tamaños óptimos de la muestra

Parámetro	Población finita	Población infinita
<b>Media</b> ( $\mu$ )	<b>Tamaño.Media muestral</b> $n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$	<b>Tamaño.Media muestral</b> $n_0 = \frac{S^2 k^2}{S^2}$
<b>Total</b> (T)	<b>Tamaño.total muestral</b> $n_t = \frac{S^2 N^2 k^2}{e^2 + N k^2 S^2}$	
<b>Proporción</b> (P)	<b>Tamaño.prop.muestral</b> $n_p = \frac{pq N k^2}{e^2 (N - 1) + pq k^2}$	<b>Tamaño.prop.muestral</b> $n_0 = \frac{pq k^2}{e^2}$

## Estimadores de los tamaños óptimos de la muestra

Parámetro	Población finita	Población infinita
Media ( $\mu$ )	Tamaño.Media muestral $n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$	Tamaño.Media muestral $n_0 = \frac{S^2 k^2}{S^2}$
Total (T)	$n_t = \frac{e^2 N}{e^2 + N k^2 S^2}$	
Proporción (P)	Tamaño.prop.muestral $n_p = \frac{pq N k^2}{e^2 (N - 1) + pq k^2}$	Tamaño.prop.muestral $n_0 = \frac{pq k^2}{e^2}$

Los más empleados en la práctica

## Conceptos básicos:

¿Qué factores influyen en el tamaño de una muestra?:

✓ El tamaño de la población (N)

$$n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$$

## Conceptos básicos:

¿Qué factores influyen en el tamaño de una muestra?:

✓ El error aleatorio (e)

$$n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$$

## Conceptos básicos:

¿Qué factores influyen en el tamaño de una muestra?:

- ✓ La dispersión esperada en las respuestas (S ó p)

$$n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$$

## Conceptos básicos:

¿Qué factores influyen en el tamaño de una muestra?:

- ✓ El nivel de confianza deseado ( $1-\alpha$ ), traducido en su valor de tablas ( $k$ )

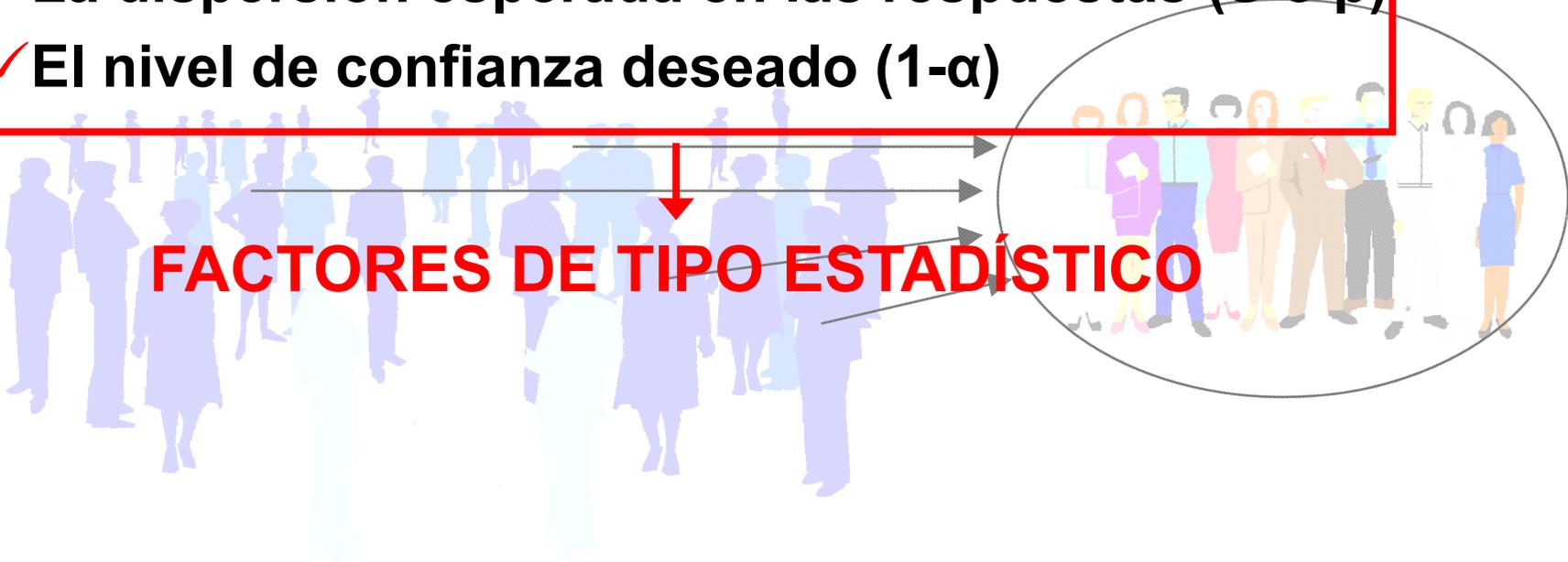
$$n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$$

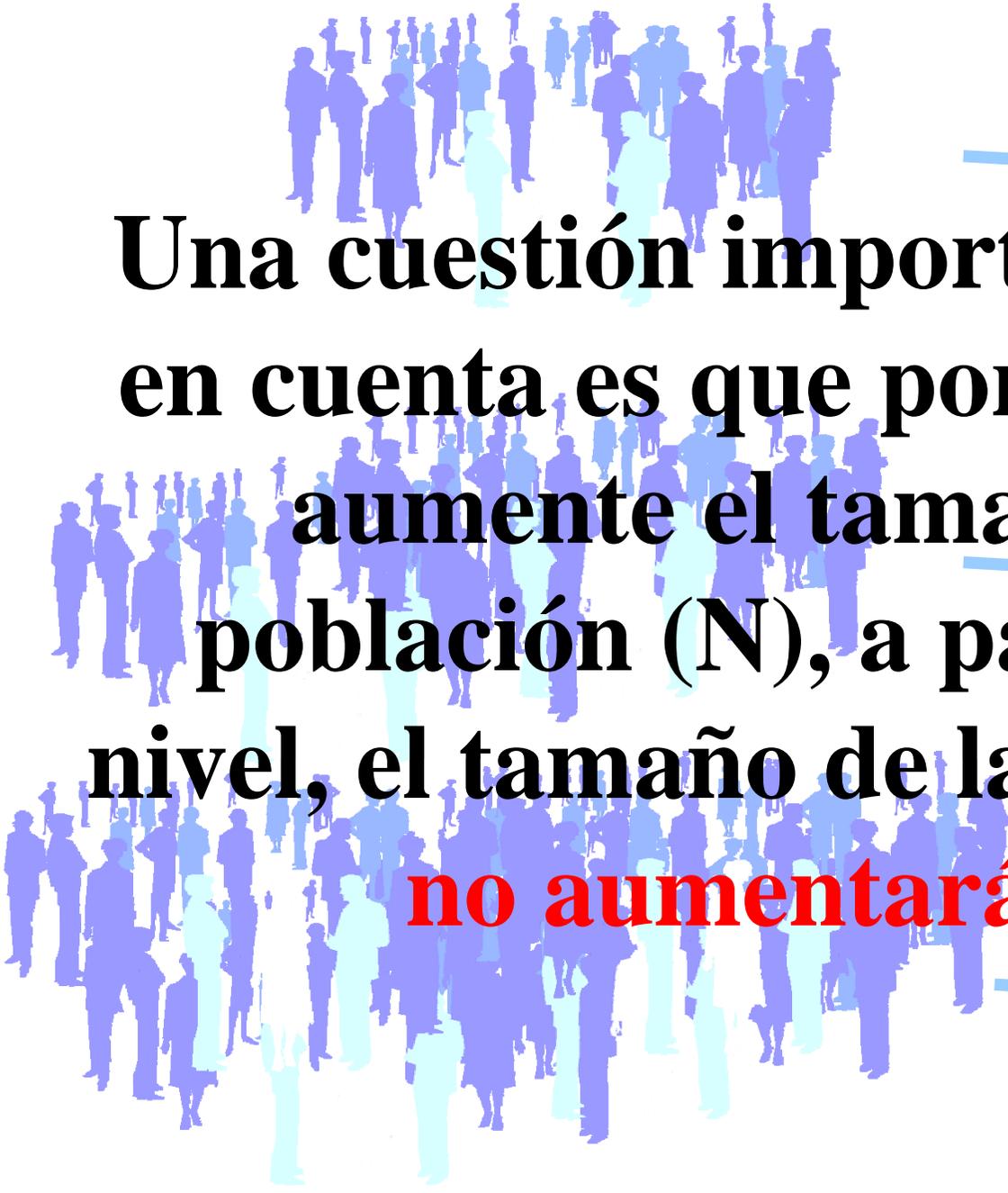
## Conceptos básicos:

¿Qué factores influyen en el tamaño de una muestra?:

- ✓ El tamaño de la población (N)
- ✓ El error aleatorio (e)
- ✓ La dispersión esperada en las respuestas (S ó p)
- ✓ El nivel de confianza deseado ( $1-\alpha$ )

**FACTORES DE TIPO ESTADÍSTICO**





**Una cuestión importante a tener en cuenta es que por mucho que aumente el tamaño de la población ( $N$ ), a partir de un nivel, el tamaño de la muestra ( $n$ ) no aumentará más.**



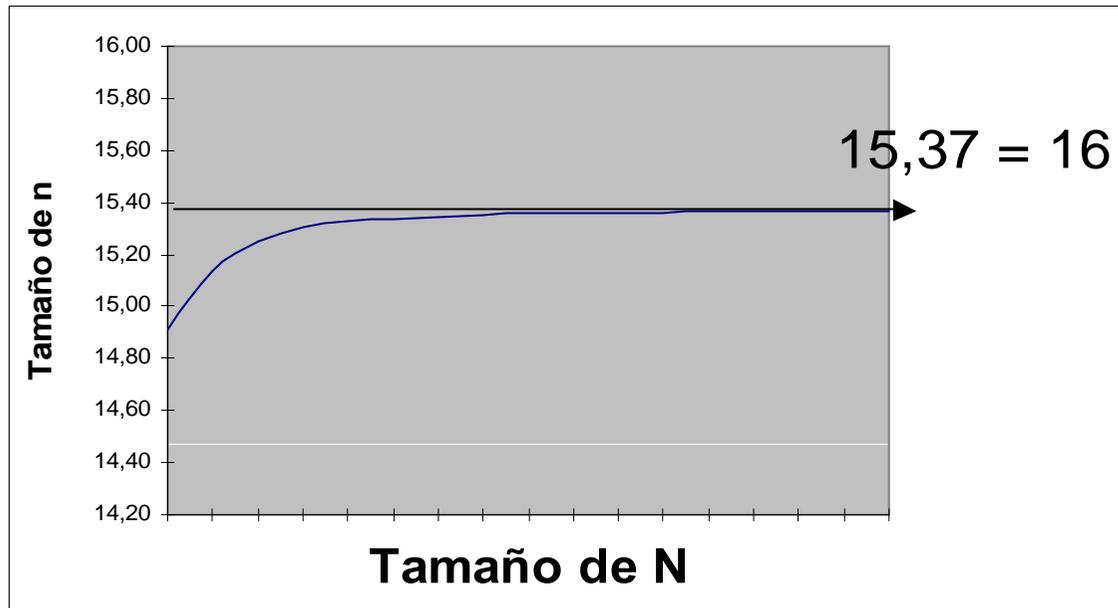
Cuas.Mtral.(S<sup>2</sup>): 1  
 Error<sup>2</sup> (e<sup>2</sup>): 0,25  
 Valor k: 1,96

Población (N)	Muestra(n)
500	14,91
1000	15,13
2000	15,25
4000	15,31
6000	15,33
8000	15,34
10.000	15,34
20.000	15,35
30.000	15,36
40.000	15,36
50.000	15,36
60.000	15,36
70.000	15,36
80.000	15,36
90.000	15,36
100.000	15,36
110.000	15,36

.....  
 .....  
 15,37

**Ejemplo: Estimación del tamaño óptimo de muestra para el n° medio de unidades de consumidas de un producto**

$$n_{\bar{x}} = \frac{S^2 N k^2}{e^2 N + k^2 S^2}$$



### **6.3.4. Elección definitiva de los elementos que compondrán la muestra**

- **Una vez conocido el tamaño óptimo (mínimo) de la muestra:  $n$ , se realiza la elección de los elementos**
- **Partiendo del marco muestral se asigna a cada elemento un número correlativo**
- **A continuación, se eligen los  $n$  elementos a través de una tabla de números aleatorios**

## **6.4. Otros tipos de muestreo probabilístico**

### **6.4.1.- Muestreo estratificado aleatorio**

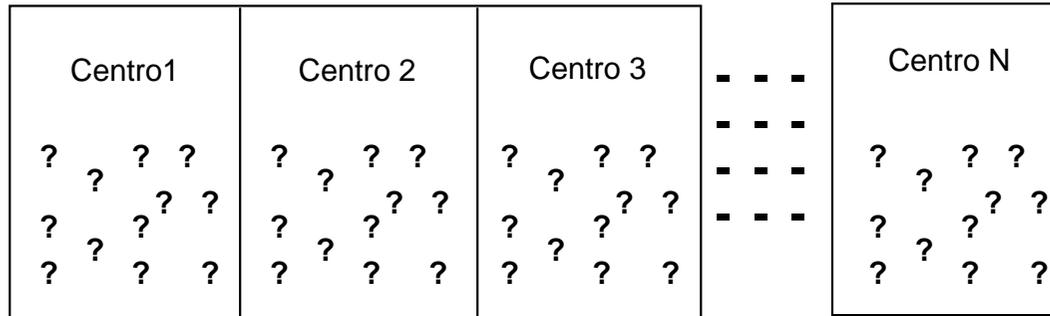
- **Los elementos son clasificados en sub-grupos de unidades con características homogéneas denominados estratos**
- **Luego de cada estrato se obtiene una muestra aleatoria representativa del mismo**
- **Con ello se gana exactitud al particionar el colectivo de acuerdo a algún criterio socioeconómico**

- **Debe cumplir varias condiciones importantes:**
  - **Los estratos deberán ser mutuamente excluyentes y colectivamente exhaustivos**
  - **El criterio para la selección de las variables que nos sirvan para realizar la partición debe ser que los estratos formados sean lo más homogéneos internamente y heterogéneos externamente**
  - **Los criterios empleados tienen que estar bastante vinculados a las características que se estudien**
  - **Las variables escogidas deberán ser fáciles de medir y aplicar, p.e. las demográficas**

- **Normalmente, se comienza calculando el tamaño de la muestra, y luego se procede al reparto entre los distintos estratos: Afijación:**
  - **Afijación simple: Reparto a partes iguales de la muestra entre el número de estratos conocidos**
  - **Afijación proporcional: Reparto de la muestra proporcional al tamaño del estrato**
  - **Afijación óptima: Reparto en función del tamaño del estrato y de su heterogeneidad. Requiere un mayor conocimiento de la población objeto de estudio**

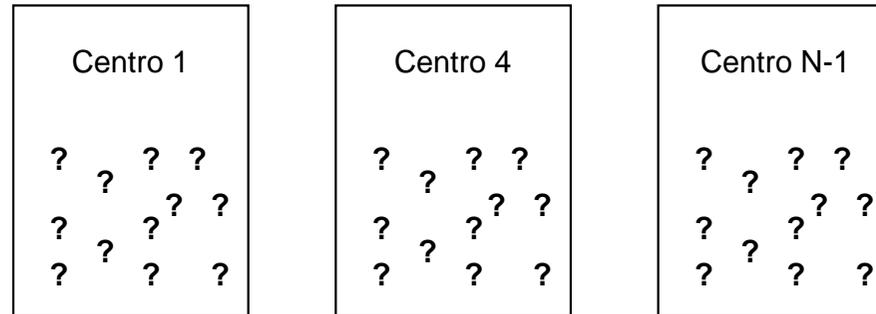
## **6.4.2.- Muestreo por conglomerados o áreas**

- **Útil cuando no se cuenta con listados de los elementos de la población directamente, sino como pertenencia a alguna unidad intermedia para la que sí es posible confeccionar un listado o relación**
- **La unidad de muestreo es un grupo que contiene varios elementos, y se denomina conglomerado**
- **Un conglomerado debe ser tan heterogéneo como la población misma**
- **Ejemplo: estudio de la población universitaria española:**



No sabemos *quiénes* son los estudiantes, aunque sí sabemos cuáles son todos los centros universitarios del Estado (unidad de muestreo).

1º Escogemos un número representativo de centros.



2º Censamos a todos los elementos (estudiantes ) de cada centro escogido.

- |  |
|--|
| <p><b>Centro 4</b><br/> <i>Listado:</i><br/> <i>Alvarez López, Emilio</i><br/> <i>Bello Glez., María</i><br/> <i>Etxeberria Gereño, Enrique</i><br/> <i>Fdez. Fdez., Pedro María</i><br/> <i>Ibarra abando, Miren</i><br/>           .....<br/>           .....<br/>           .....</p> |
|--|

## 6.4.3 Muestreo por etapas

- **Modalidad parecida a la anterior**
- **Presenta dos o más etapas**
- **Se particiona a la población en unidades muestrales que se eligen aleatoriamente**
- **Cada partición de la población conforma una etapa**
- **Para pasar de la unidad muestral a los elementos definitivos no se realiza un censo, sino que se vuelve a realizar una elección aleatoria**
- **Salvo en la última, la unidad muestral no coincide con la unidad que proporciona información**

- **Ejemplo: Siguiendo con el ejemplo anterior, una vez elegidos los centros de estudios, se elegiría aleatoriamente a un número de estudiantes de cada uno. Con los estudiantes escogidos aleatoriamente de todos los centros, se forma la muestra**
- **Cada partición conforma una etapa. El ejemplo es bietápico. En la primera etapa se accede a centros y en la segunda a estudiantes**
- **Tanto el muestreo por conglomerados como el muestreo por etapas tienen como aplicación poder acceder a poblaciones cuyos elementos no se conocen**
- **La diferencia es que la primera tiene mayor precisión al censar a los elementos que componen la unidad muestral, y la segunda es más económica al volver a extraer una muestra de cada unidad intermedia**

## 6.4.4 Muestreo sistemático

- **Consiste en ir eligiendo las unidades muestrales de  $k$  en  $k$  unidades tomando como origen una de ellas elegida al azar entre el elemento que ocupa el primer lugar y el que ocupa el lugar  $k$ . Siendo  $k=N/n$**
- **Se necesita un listado de los elementos de la población o marco muestral. Este debe disponer los elementos según algún criterio que no tenga que ver con el problema de estudio, si no se cometería error sistemático. Ejemplo: estudio del tamaño de las empresas partiendo de listado ordenado por nombre**
- **Tambien se puede emplear sin conocer marco muestral. Ejemplo: entrevistar cada  $i$ -esima persona que sale de una tienda**

- **Tamaño población: 10.000 individuos.**
- **Tamaño de muestra: 1.000 unidades.**
- **Coeficiente de elevación=  $10.000/1.000=10$**
- **Número elegido al azar de 1 a 10 =3**
  - **1º unidad muestral: número 3 del censo**
  - **2ª unidad muestral:  $3+10=$  número 13 del censo**
  - **3ª unidad muestral:  $13+10$  número 23 del censo**
  - **etc.**

## **6.5. La determinación del tamaño de muestra inicial**

- **El valor  $n$  de la muestra, representa el tamaño neto o final que debe alcanzarse para asegurarnos de que los parámetros son estimados con el nivel de confianza y precisión deseados.**
- **En la práctica, debe considerarse un tamaño mucho mayor de entrevistados potenciales, ya que:**
  - **Ciertas personas del marco no son elegibles**
  - **Habr  personas que no acepten ser entrevistadas, o que la comiencen y no la terminen**

## Conceptos básicos:

### ¿Qué factores influyen en el tamaño de una muestra?:

- ✓ El tamaño de la población (N)
- ✓ El error aleatorio (e)
- ✓ La dispersión esperada en las respuestas (S ó p)
- ✓ El nivel de confianza deseado ( $1-\alpha$ )

- ✓ **El número de negativas por cada aceptación para entrevistarse**
  - ✓ **Cualquier incidencia que acarree una proporción de éxitos respecto a los intentos para contactar con un entrevistado**
- 

**FACTORES DE TIPO PRÁCTICO**

## Conceptos básicos:

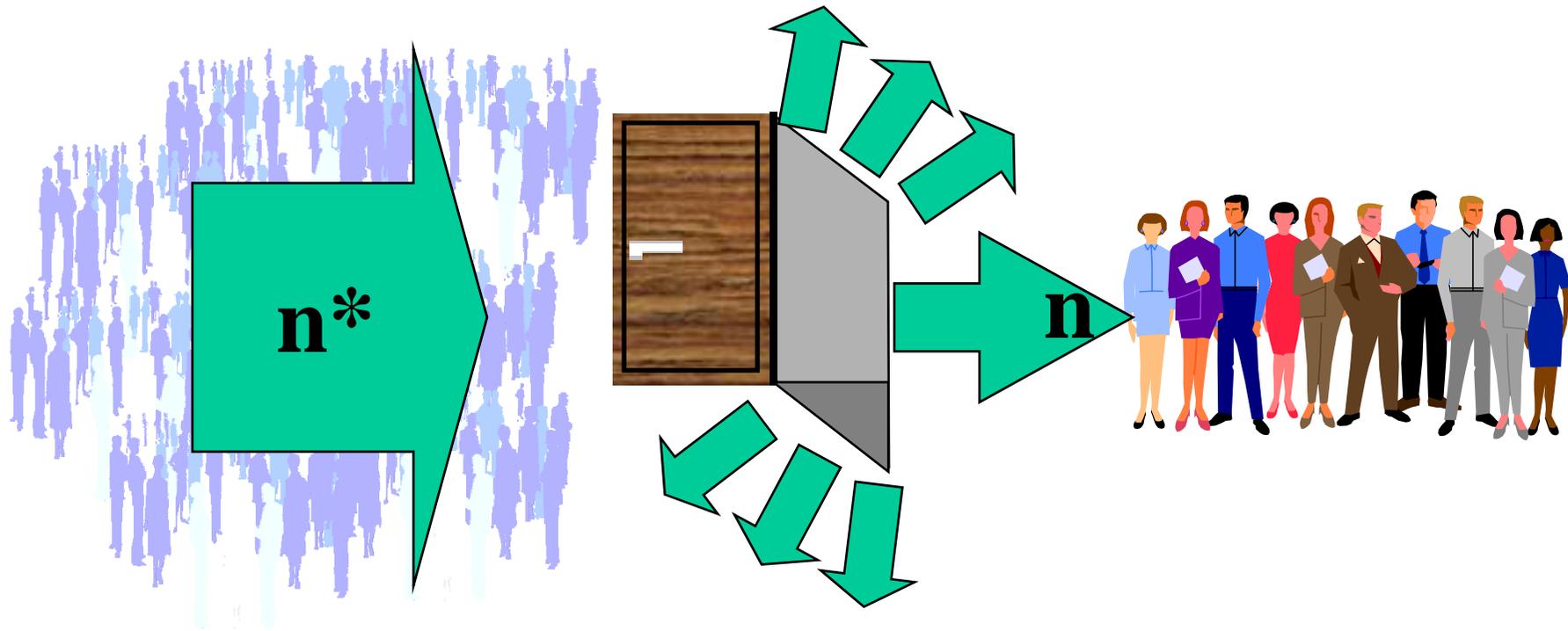
**POR EJEMPLO:** Que el entrevistado...

- ✓ se niegue a hacer la encuesta
- ✓ se niegue a responder a ciertas preguntas
- ✓ se haya ausentado
- ✓ haya cambiado de teléfono, dirección, empresa, etc.
- ✓ se ha ido a vivir a otra parte, ha cambiado de trabajo, etc
- ✓ ha cesado en su actividad
- ✓ ha fallecido

.....



**TODAS ESTAS INCIDENCIAS NOS OBLIGAN A  
AUMENTAR EL TAMAÑO DE LA MUESTRA  
HASTA UN VALOR  $n^*$  PARA QUE AL FINAL NOS  
SALGAN LAS  $n$  ENTREVISTAS DESEADAS**



- **Tendremos que medir dos efectos:**
  - **Ratio de incidencia (Ri): porcentaje de personas elegibles para participar en un estudio**
  - **Ratio de finalización (Rf): porcentaje de personas que siendo aptos para figurar en la muestra desean colaborar para realizarla y acaban la entrevista**

$$\begin{array}{l} \text{Tamaño inicial} \\ \text{de la muestra} \end{array} = \frac{n}{R_i * R_f}$$

## Ejemplo:

- **Opinión de clientes sobre un nuevo detergente para lavavajillas aparecido el último mes**
- **Datos:**
  - **Penetración de los lavavajillas en los hogares: 30%**
  - **Panel de consumidores: 65% de usuarios de lavavajillas compraron detergente el último mes**
  - **Según estudios previos el grado de finalización es del 80%**
- **$R_i = 0,3 * 0,65$ ;  $R_f = 0,8$**
- **Tamaño inicial de muestra =  $n / (0,3 * 0,65 * 0,8) = 6,41 * n$**