

Mètodes de Mostratge

MUESTREO EN POBLACIONES FINITAS

(TEORÍA DEL MUESTREO/
SAMPLING THEORY)

Profesora: Cristina Núñez del Toro
Despacho UPC Campus Nord C5-224

consultas: viernes 8-10h

con cita previa

(cristina.nunez@upc.edu)

Qué esperáis ver en esta asignatura?

Qué es la teoría del muestreo?

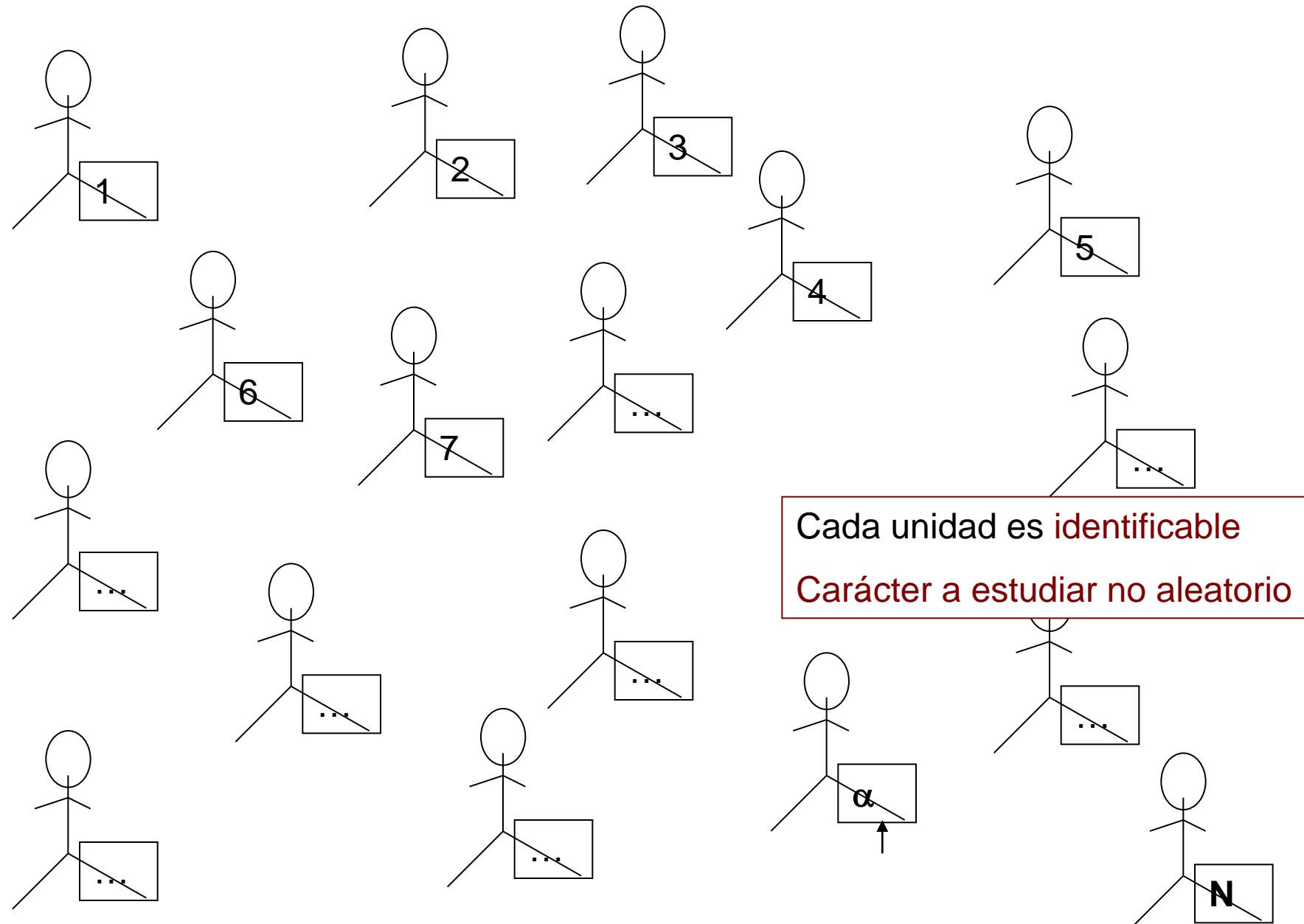
- Técnica para la selección de una muestra a partir de una población.
- Se espera conseguir que sus propiedades sean extrapolables a la población.
- Ahorrar recursos, y obtener resultados parecidos a los que se alcanzarían si se realizase un estudio de toda la población.
- No solo hacer estimaciones de la población sino estimar también los márgenes de error correspondientes a dichas estimaciones.
- Nunca podremos estar enteramente seguros de que el resultado sea una muestra representativa, pero sí podemos actuar de manera que esta condición se alcance con una probabilidad alta.
- Al conjunto de muestras que se pueden obtener de la población se denomina espacio muestral. La variable que asocia a cada muestra su probabilidad de extracción, sigue la llamada distribución muestral.

POBLACIÓN FINITA

versus

POBLACIÓN INFINITA

Población finita:



Dado que cada unidad es identificable, se puede tener un **marco muestral** en el cual cada fila-unidad está enriquecida por la **información auxiliar** conocida

Variables auxiliares

Identificador	Var1 Sexo	Var2 Edad
Unidad 1				
....				
Unidad α				
...				
Unidad N				

Se desea estimar la media poblacional de la variable: horas semanales dedicadas al trabajo doméstico

En contraste, volvamos a lo que es una población infinita: un ejemplo

Una máquina fabrica barras de acero.

En tanto que a un estadístico, le piden estimar la media poblacional de las piezas producidas por la máquina.

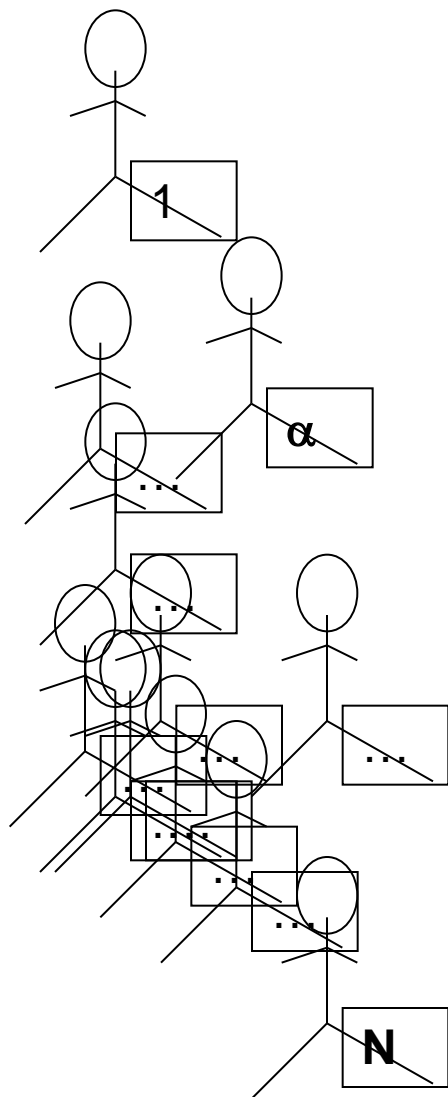


Las unidades no tienen “identificador”,
son de alguna manera intercambiables
Carácter a estudiar aleatorio

Ahora

POBLACIÓN FINITA

Población finita



Objetivo

¿cuál es la media \bar{Y} (y/o el total T) de un carácter (o variable de interés) Y en esta población

Respuesta

Censo: observar las N unidades...costoso

Teoría de muestreo=

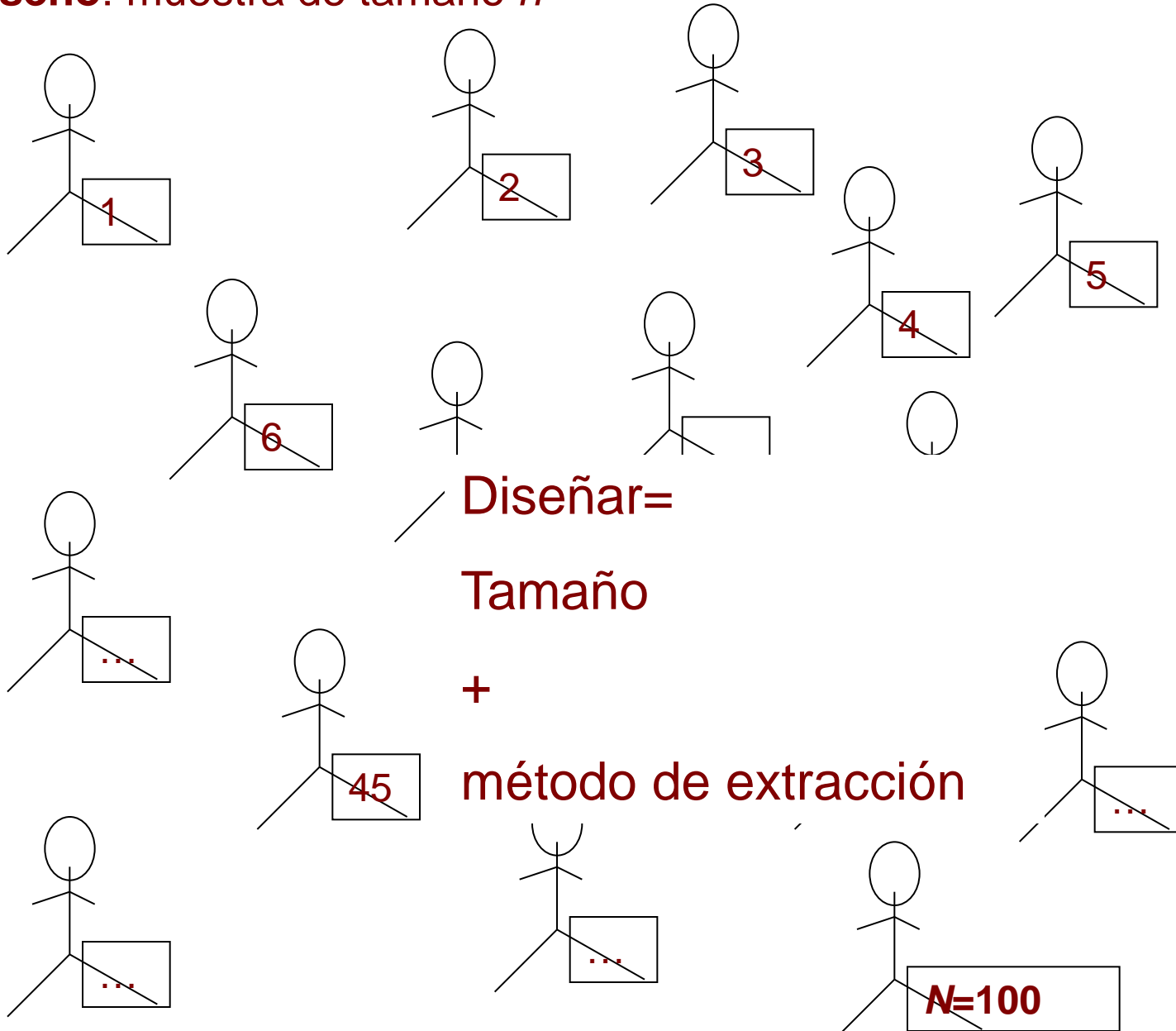
Diseño de la muestra

+

Estimación a partir de la muestra (=escoger el estadístico + hace falta conocer su distribución)

$n=4$

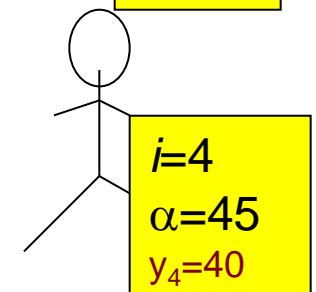
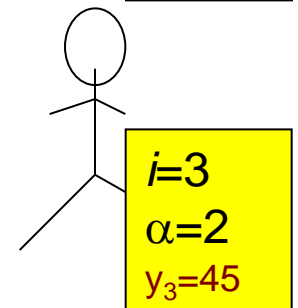
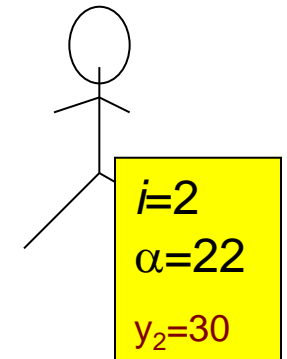
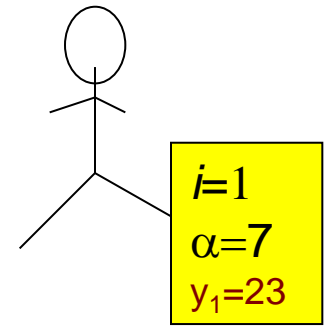
Diseño: muestra de tamaño n



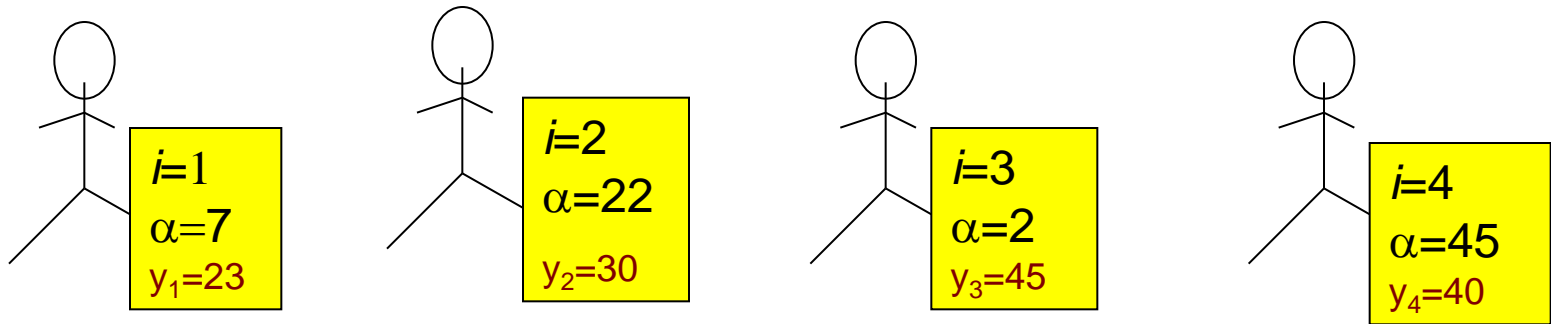
Trabajo de campo:

observar el valor de la variable Y sobre los individuos de la muestra.

La variable Y es la edad



Estimación



definir el estimador pertinente en función del método de extracción,
estimación puntual y por intervalo con una confianza determinada

$$\hat{\bar{Y}} = \text{función}(y_1, y_2, y_3, y_4)$$

$$\hat{\bar{Y}} \text{ estimador de } \bar{Y}$$

En un trabajo real, se dispone de un **Marco Muestral**

	Identificador
1	Sr Lavallo Granvia, 7
2	Sra. de la Peña, c/Principi, 22
α	...
N	Sr. Pio, c/ Pont, 32

Lista exhaustiva de todas las **unidades** del **universo**

=

Lista exhaustiva de todas los **individuos** de la **población de tal forma que se pueden localizar y entrevistar**

Ej.

Población de los activos de Barcelona

Carácter o variable de interés Y: tiene trabajo (si/no)

El **Marco Muestral** puede contener **variables auxiliares**

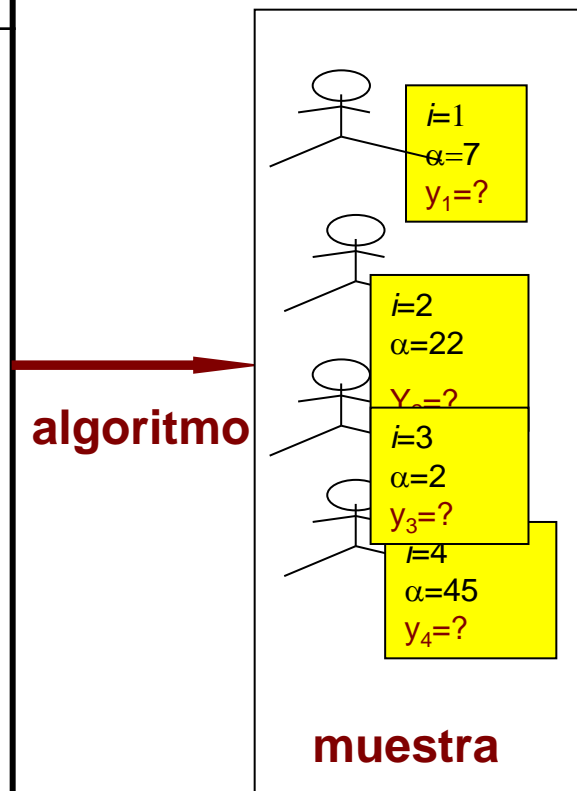
Variables auxiliares

	Identificador	sexo	edad
1	Sr Lavalles Granvia, 7	H	22		
2	Sra. de la Penya, c/Principi, 22	M	39		
...	...				
α					
N	Sr. Pio, c/ Pont, 32	H	51		

Ej.

Población de los activos de Barcelona

Carácter o variable de interés Y : tiene trabajo (si/no)



MUESTREO EN POBLACIONES FINITAS

En resumen:

- Estimar los parámetros de una variable: media, total, ... de una determinada variable de interés en una población finita mediante la observación de una muestra
- Dos fases íntimamente ligadas:
 - Diseño:** método de extracción de la muestra, diseño de la muestra
 - Estimación:** definir el estimador pertinente en función del método de extracción, estimación puntual y por intervalo con una confianza determinada

EJEMPLO

MARFIL ES UN PAÍS DE 6 MUNICIPIOS:

La pregunta

Población de $N=6$ unidades-municipios (=individuos)

Variable de interés Y : “número de habitantes”

Interesa conocer el total y la media de esta variable de interés (T_Y e \bar{Y})

La recogida de datos:

Diseño y extracción de una muestra de tamaño n de municipios

Trabajo de campo: observar el número de habitantes de los municipios seleccionados

La respuesta

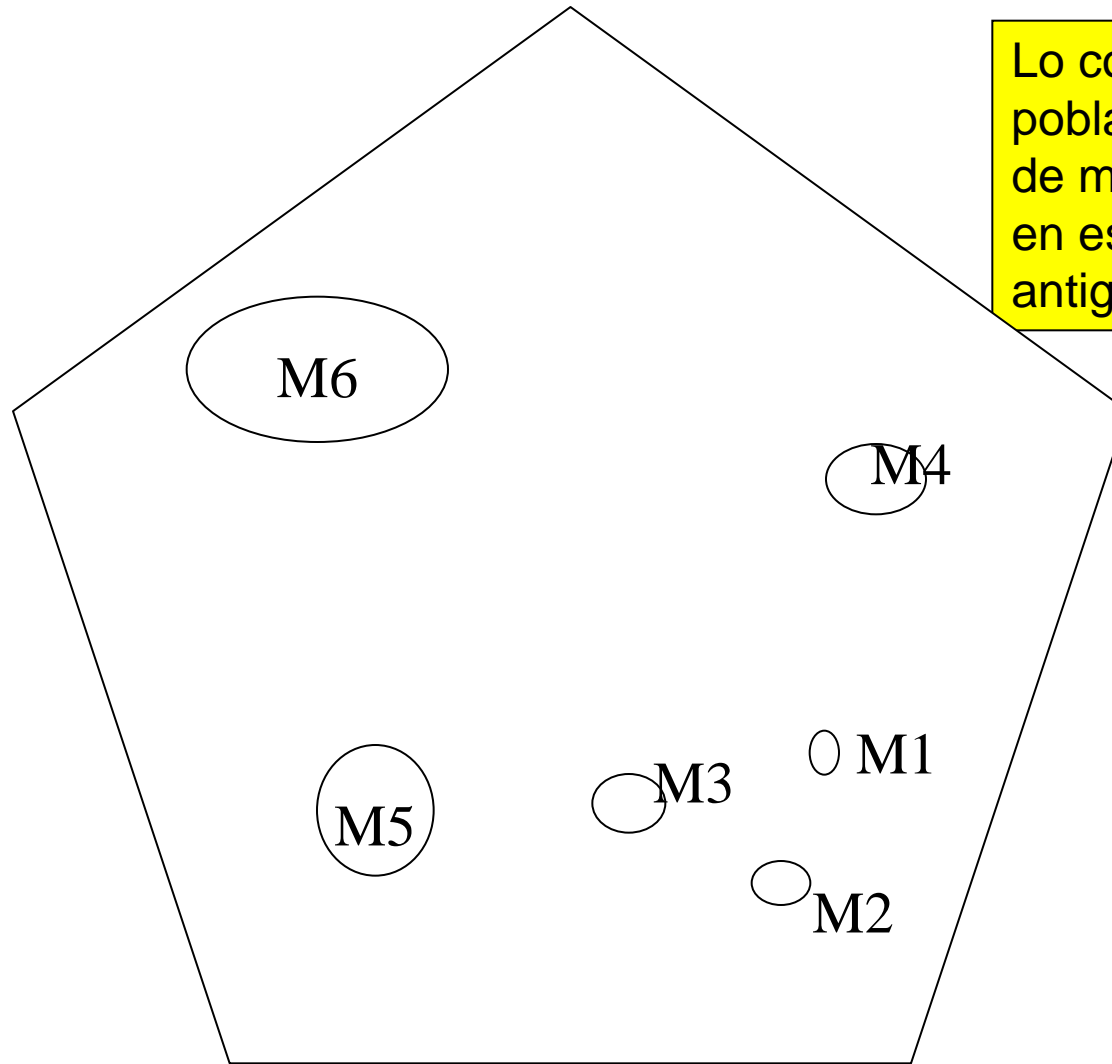
Estimación

- del total de habitantes T_Y : valor que toma el estimador: \hat{T}_Y sobre la muestra
- de la media \bar{Y} de habitantes por municipio valor que toma el estimador: $\hat{\bar{Y}}$ sobre la muestra

MARFIL:

UNIVERSO (POBLACIÓN) compuesto de
6 UNIDADES (o individuos, los municipios)
Se estudia la **VARIABLE DE INTERÉS**: Número
de habitantes Y

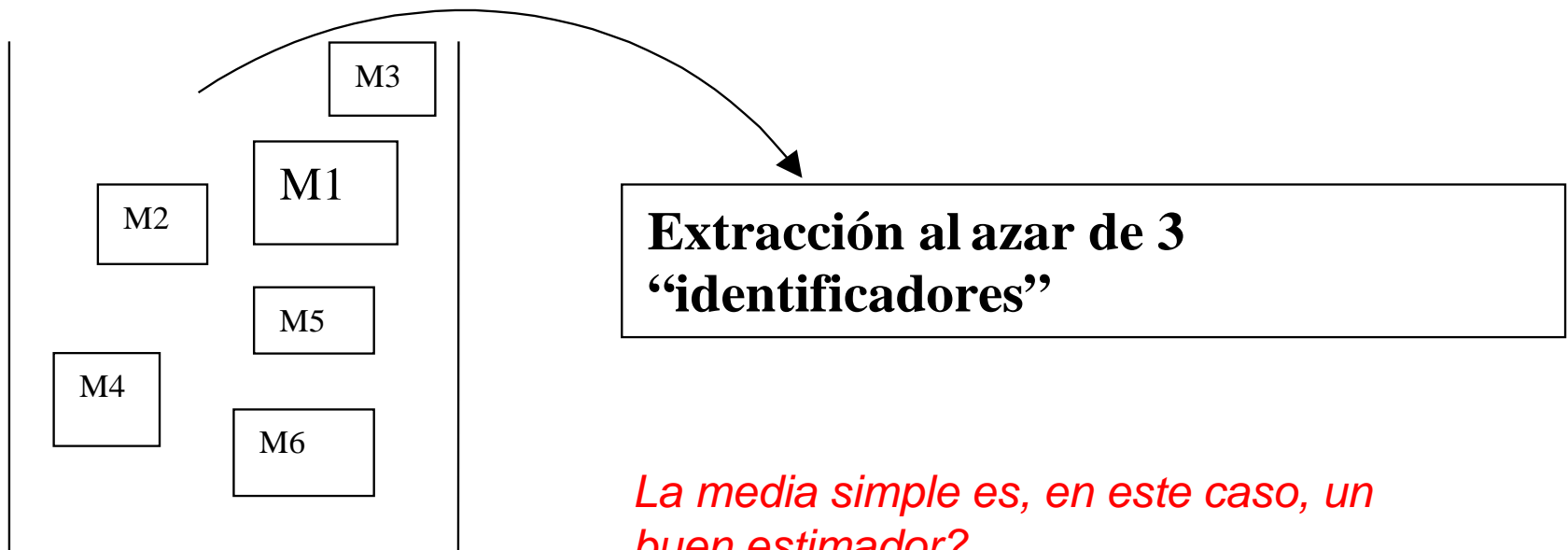
Lo conocido sobre la
población (universo)
de municipios viene
en este mapa
antiguo de 5 años



DISEÑO

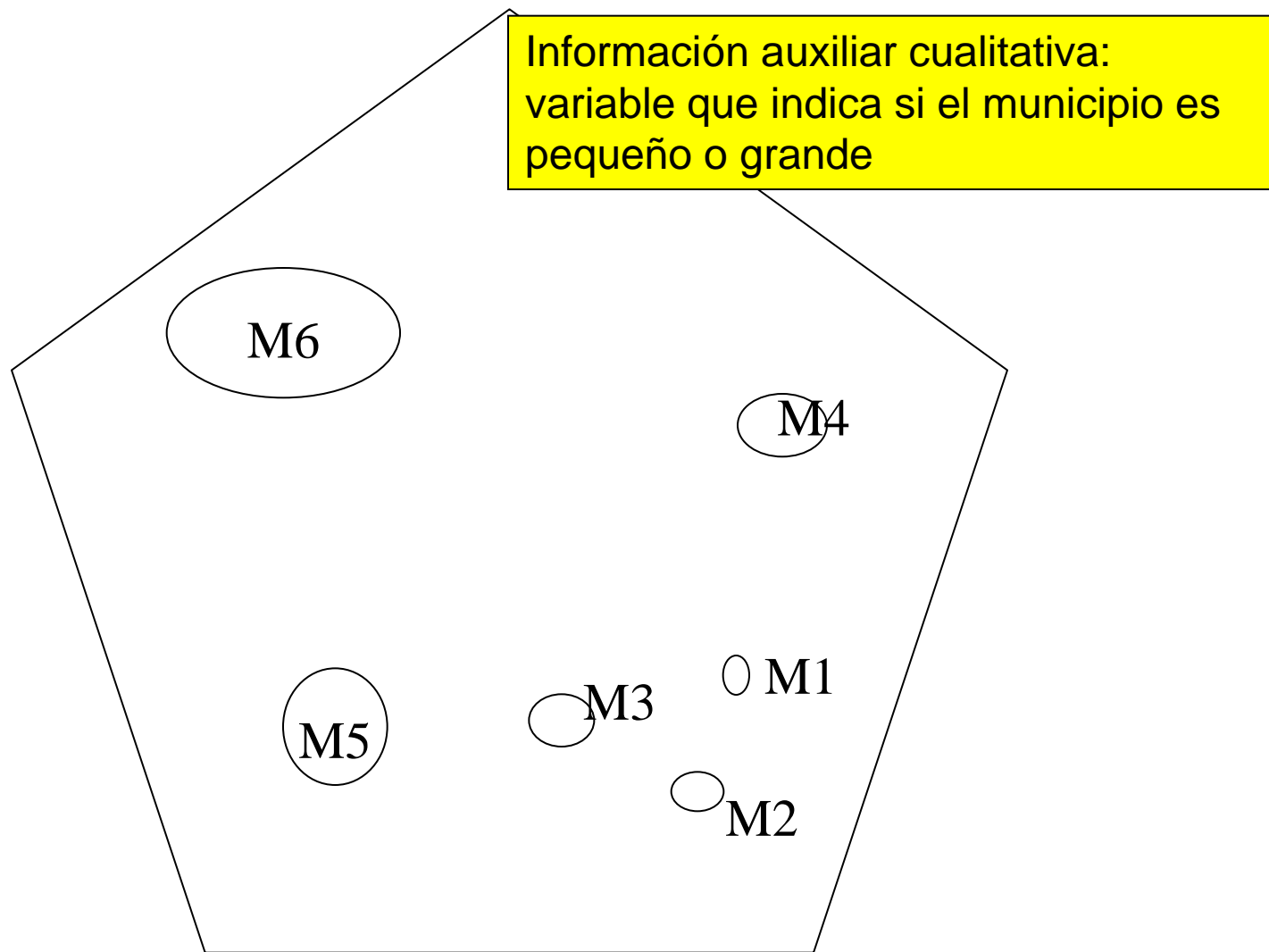
1. DISEÑO ALEATORIO SIMPLE

- con reposición
- sin reposición



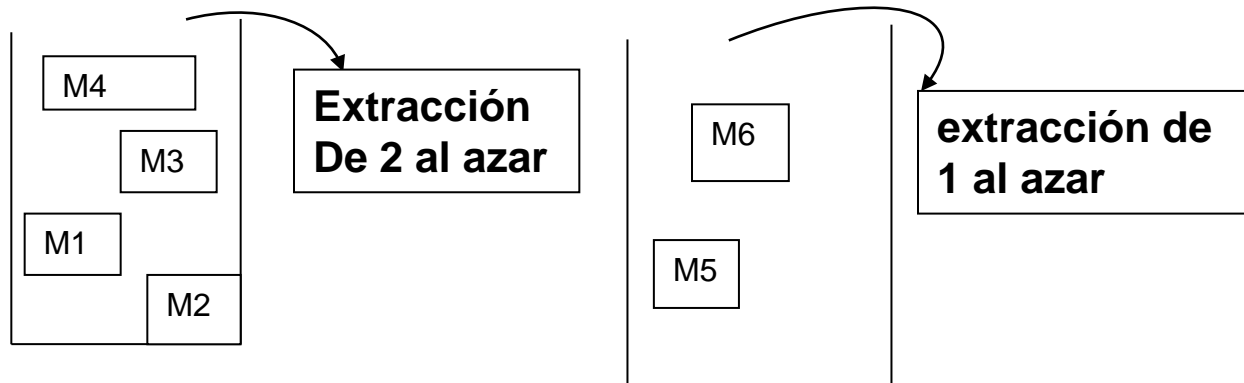
2. DISEÑO ESTRATIFICADO

UTILIZAR INFORMACIÓN AUXILIAR CUALITATIVA:



Por ejemplo:

2 estratos, se reparte la muestra en los 2 estratos de tamaño. El tamaño de la muestra se reparte de la siguiente forma en los estratos: 2 y 1, respectivamente



Se necesita una variable auxiliar cualitativa X **conocida sobre todos los individuos** para poder repartir la población en los estratos

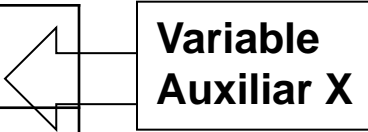
Porqué interesa este tipo de diseño?

La media simple es, en este caso, un buen estimador?

DISEÑO CON PROBABILIDADES DESIGUALES

INFORMACIÓN AUXILIAR CUANTITATIVA:

A partir del último censo, se conoce el número de viviendas de cada municipio:

Munic	Nº viv	
M_1	10	
M_2	10	
M_3	30	
M_4	50	
M_5	100	
M_6	100	

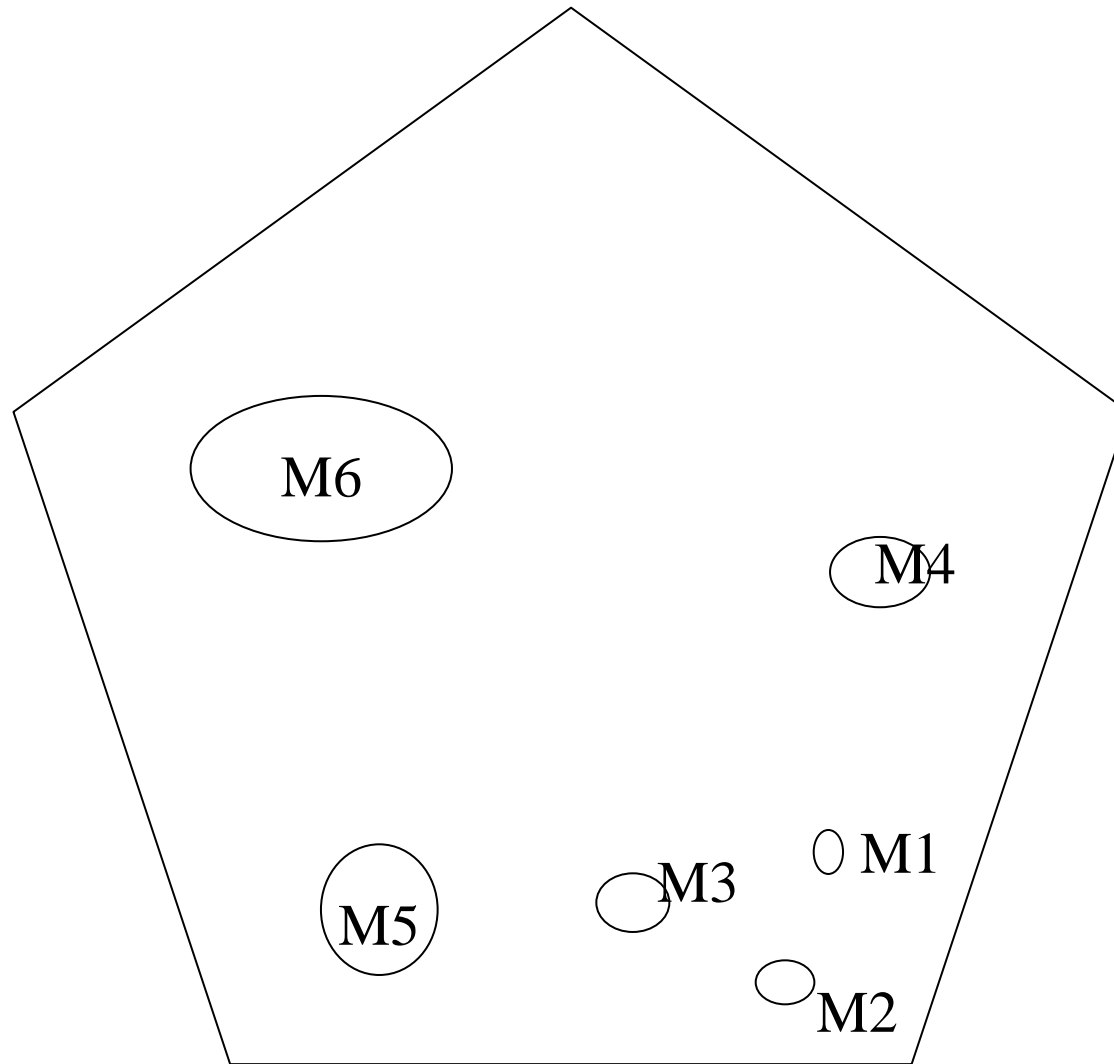
Por ejemplo, se da a cada municipio una probabilidad proporcional al número de viviendas X

MUESTREO PROPORCIONAL AL TAMAÑO

Munic	Nº viv	π
M_1	10	0,1
M_2	10	0,1
M_3	30	0,3
M_4	50	0,5
M_5	100	1
M_6	100	1

La media simple es, en este caso, un buen estimador?

DISEÑO MULTIETÁPICO



La media simple es, en este caso, un buen estimador?

ESTIMACIÓN

ESTIMACIÓN

1. El estimador escogido, función de la muestra, depende del método de extracción.
2. Se busca un estimador insesgado y preciso (intervalo de confianza pequeño)

Esto es gran parte del curso....

RECOMPOSICIÓN

Mejorar la estimación obtenida con el estimador “bruto” mediante un estimador “recompuesto” que utiliza información auxiliar

PROGRAMA

0. Repaso: muestreo en población infinita

1. Introducció.

DISSENY CLÀSSICS DE MOSTRATGE

2. Estadístics i estimadors. Repàs estimació en poblacions infinites. Errors en l'estimació.

3. Disseny simple

4. Disseny estratificat

5. Disseny amb probabilitats desiguals

6. Disseny multietàpic.

7. Disseny no probabilista o empíric

RECOMPOSICIÓ

8. Recomposició

9. Tractament de la no-resposta

LO QUE NO SE VE :

10. Mostreig equilibrat. Mostreig indirect. Estudis temporals, Mostreig espacial, Estimació en petits dominis i petites àrees, Precisió de l'estimació en dissenys complexos Mètode bootstrap).

ORGANIZACIÓN

- Clases de teoría Martes 9h00*-10h30 Aula 9B
- Clases de Problemas Miércoles 9h00-10h30 Aula 9B
- Clases de prácticas Viernes 10h-11h, grupo 1 Aula I8
11h-12h grupo 2 Aula I8

DIVISIÓN EN GRUPOS

Por orden alfabético

Un responsable por cada grupo

Ver cambios con ellos, sabiendo que uno se puede cambiar si encuentra con quién intercambiarse.

Responsables?

En clases de prácticas se pasará lista días **aleatorios**

Evaluación:

CONTINUADA

- 3 parciales (1/3 de la nota) con teoría y problemas y **prácticas**. Las preguntas relativas a los tres bloques no se diferencian forzosamente. La práctica se evalúa únicamente mediante preguntas en el parcial.
- Para contestar a determinadas preguntas, es necesario traer los scripts de las prácticas así como los cuadros de resultados.
- Se entregan el bloque de prácticas correspondientes en el parcial.
- Fechas de los parciales en la programación de la asignatura. El último coincide con el examen final (9 de junio).

Evaluación:

ÚNICA

El examen final (evaluación única) comporta una parte de teoría, de problemas y la realización de prácticas en ordenador seleccionadas al azar por cada uno de los candidatos

EXTRAORDINARIA

El examen extraordinario incluye preguntas de teoría, problemas así como la realización de prácticas en ordenador seleccionadas al azar por cada uno de los candidatos.

MATERIAL

- Clairin R., Brion Ph.:
Manual de muestreo, Editorial: La Muralla-Hesperides
INDISPENSABLE
- Transparencias (Atenea)
- Lista de problemas para cada uno de los temas (Atenea)

BIBLIOGRAFÍA ADICIONAL

- La política de los grandes números. Historia de la razón estadística
Alain Desrosières. Edición Melusina

FIN