



INSTITUTO NACIONAL DE ESTADISTICA



# Encuestas Demográficas.

Diseño de la muestra

*INE*

*24 de Junio de 2.016*

Prácticas Encuestas Demográficas. EPA



## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



# 1.INTRODUCCIÓN

## EPA

✓La Encuesta de Población Activa es una encuesta de **tipo continuo**, con **periodicidad trimestral** que se viene realizando ininterrumpidamente desde 1964.

✓Su objetivo es el conocimiento de la actividad económica del país en los relativo al componente humano.

✓La desagregación de los resultados

- Detallados a escala nacional
- Principales características a nivel de CCAA y Prov

## EES2014

✓La Encuesta Europea de Salud es una encuesta de **tipo “esporádico”**, con **periodicidad quinquenal** que se viene realizando ininterrumpidamente desde 1964.

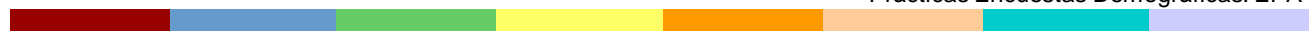
✓Su objetivo es proporcionar información sobre salud de la población española, de manera armonizada y comparable a nivel europeo.

✓La desagregación de los resultados

- Detallados a escala nacional
- Principales características a nivel de CCAA

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



## 2. Ámbito de la encuesta

### EPA

### EES2014

#### Poblacional

- Población residente en viviendas familiares principales( **16 o más años**)

- Población residente en viviendas familiares principales( **15 o más años**)



**Se excluyen hogares colectivos**

#### Geográfico

- Territorio Nacional

- Territorio Nacional

#### Temporal

- Referencia Resultados:  
**Trimestre**
- Referencia Información:  
Semana previa a la recogida  
**(Semana de Referencia)**

- Referencia Resultados:  
**anual (año 2014)**
- Referencia Información:  
**Últimos 12 meses/4 semanas... antes de la entrevista**

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



### 3.Marco de la Encuesta

Distinguimos dos marcos:

#### Marco áreas geográficas

- Secciones censales, áreas geográficas en que se divide el territorio nacional utilizadas con fines estadísticos y electorales.

**EPA**

*Censo 2011.  
Actualizado  
periódicamente*

#### Marco Viviendas

- Relación de viviendas en las secciones seleccionadas para la muestra.

**EES2014**

*Padrón*

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES

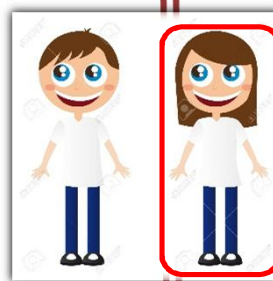
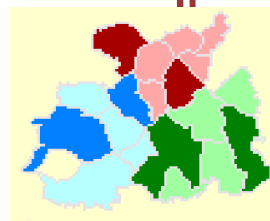




## 4. Tipo de muestreo

### EPA

- Muestreo *bietápico* con estratificación de unidades de primera etapa
- Unidades de primera etapa : Secciones censales
- Unidades de Segunda etapa: Viviendas Principales



### EES2014

- Muestreo *trietápico* con estratificación de unidades de primera etapa
- Unidades de primera etapa : Secciones censales
- Unidades de Segunda etapa: Viviendas Principales
- Unidades de tercera etapa: Un adulto de 15 o más años

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA.AFIJACIÓN
6. CRITERIO DE ESTRATIFICACION
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
- 11.ESTIMADORES
- 12.ERRORES



## 5. Tamaño de la muestra. Afijación

### EPA

- **TAMAÑO:**

Secciones Censales =3822

Viviendas ≈ 20 titulares

*(\*) EPA usa viviendas reservas, para sustitución en primera entrevista*

- **AFIJACIÓN:** Objetivos: Estimaciones provinciales y nacionales fiables.  
En **provincias** : De compromiso entre uniforme y proporcional  
En **estratos** : Estrictamente proporcional

### EES2014

- **TAMAÑO:**

Secciones Censales =2500

Viviendas =15

*(\*) Sin viviendas reservas desde 2014. Se aumentó el tamaño para conseguir el mismo tamaño de muestra efectiva que con reservas (12 tit y 12 reservas)*

- **AFIJACIÓN:** Objetivos: Estimaciones provinciales y nacionales fiables.  
En **CCAA**: De compromiso entre uniforme y proporcional  
En **estratos** : Estrictamente proporcional

## 5. Tamaño de la muestra. Afijación

17:13 Tuesday, May 19, 2015

### DISTRIBUCION SECCIONES EPA POR CPRO/ESTRATO

	ESTRATO									Todo
	1	2	3	4	5	6	7	8	9	
	N	N	N	N	N	N	N	N	N	
PROVIN										
01-Álava	30	.	.	.	.	3	6	.	.	39
02-Albacete	18	.	.	.	8	.	4	5	4	39
03-Alicante/Alacant	17	10	.	16	21	6	5	3	.	78
04-Almería	16	.	.	9	.	7	3	4	.	39
05-Ávila	15	.	.	.	.	.	6	5	13	39
06-Badajoz	20	.	.	6	10	6	12	14	10	78
07-Balears (Illes)	42	.	.	9	28	12	9	4	.	104
08-Barcelona	50	.	33	24	22	12	9	3	3	156
09-Burgos	20	.	.	.	7	.	3	.	9	39
10-Cáceres	19	.	.	.	7	8	9	10	25	78
11-Cádiz	12	13	6	26	9	6	6	.	.	78
12-Castellón/Castelló	26	.	.	.	29	5	7	4	7	78
13-Ciudad Real	13	9	.	.	14	16	11	7	8	78
14-Córdoba	34	.	.	.	17	7	10	10	.	78
15-Coruña (A)	42	14	12	.	30	16	28	14	.	156
16-Cuenca	12	.	.	.	.	.	8	5	14	39
17-Girona	15	.	.	.	24	13	10	8	8	78
18-Granada	26	.	.	5	10	14	9	8	6	78
19-Guadalajara	16	.	.	.	8	.	.	8	7	39
20-Guipuzcoa	26	.	.	6	12	18	10	6	.	78
21-Huelva	14	.	.	.	8	8	.	9	.	39
22-Huesca	12	.	.	.	.	15	.	.	12	39
23-Ibiza	17	8	.	.	15	12	10	16	.	78

EPA

## 5. Tamaño de la muestra. Afijación

Comunidades Autónomas	Secciones
01 Andalucía	296
02 Aragón	108
03 Asturias (Princ. de)	96
04 Baleares (Islas)	100
05 Canarias	124
06 Cantabria	88
07 Castilla y León	148
08 Castilla-La Mancha	128
09 Cataluña	276
10 Comunidad Valenciana	204
11 Extremadura	100
12 Galicia	144
13 Madrid (Comunidad de)	240
14 Murcia (Región de)	108
15 Navarra (Com. Foral)	88
16 País Vasco	128
17 La Rioja	76
18 Ceuta	24
19 Melilla	24
<b>Total</b>	<b>2500</b>



**EES14**

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



## 6. Criterio de estratificación

### EPA

### EES2014

**1. Geográfico:** Según la importancia demográfica del municipio al que pertenecen las unidades primarias

**1. Geográfico:** Según la importancia demográfica del municipio al que pertenecen las unidades primarias

INSTITUTO NACIONAL DE ESTADÍSTICA

Municipios Autorrepresentados	1	• Capitales de provincia	0	• Municipios de más de 500.000 habitantes
	2	• Municipios importantes en relación con	1	• Capitales de provincia(excepto los anteriores)
	3	• Municipios Mayores de 100.000 habita	2	• Municipios Mayores de 100.000 habitantes
Municipios Correpresentados	4	• 50.000-100.000	3	• 50.000-100.000
	5	• 20.000-50.000	4	• 20.000-50.000
	6	• 10.000-20.000	5	• 10.000-20.000
	7	• 5.000-10.000	6	• Menos de 10.000 habitantes
	8	• 2.000-5.000		
	9	• <2.000		

**2. Socioeconómico:** Dentro de cada estrato las secciones se clasifican según la categoría socioeconómica de la población activa de la sección

## 6.1 Criterios de subestratificación (I)

### Socioeconómico(subestratos)

EPA

La metodología utilizada para definir los nuevos subestratos se ha modificado en relación con la empleada en el anterior diseño(hasta final de 2004).

Las secciones censales se han agrupado dentro de cada estrato en subestratos, de la forma más homogénea posible. Para ello se han tenido en cuenta aquellas características que se consideran más correlacionadas con las variables de interés de la encuesta.

La información sobre las variables de subestratificación al nivel de sección censal procede del Censo 2011 y de la Agencia Tributaria



## 6.1 Criterios de subestratificación (II)

### Socioeconómico(subestratos)

**EPA**

En el proceso de subestratificación se han considerado dos grupos de secciones:

- Las de los estratos 7, 8 y 9 a las que se les asigna como subestrato la comarca (NUTS4) del municipio al que pertenecen
- Las del resto de estratos a las que se procede a agrupar, dentro de sus estratos, aplicando técnicas de análisis de conglomerados (cluster).

## 6.1 Criterios de subestratificación (III)

En la subestratificación de las secciones de los estratos 1 a 6, se han utilizado las siguientes **variables**:

- Porcentaje de parados en la sección.
- Porcentaje de inactivos.
- Porcentaje de ocupados.
- Porcentaje de extranjeros.
- Porcentaje de personas entre 0 y 14 años.
- Porcentaje de personas entre 15 y 24 años.
- Porcentaje de personas de 65 o más años.
- Porcentaje de personas con nivel de estudios realizado de analfabetos, sin estudios o nivel de estudios de primer grado.
- Porcentaje de personas con nivel de estudios realizado de ESO, EGB, Bachillerato, FP.
- Porcentaje de personas con nivel de estudios realizado de diplomatura, licenciatura o doctorado.



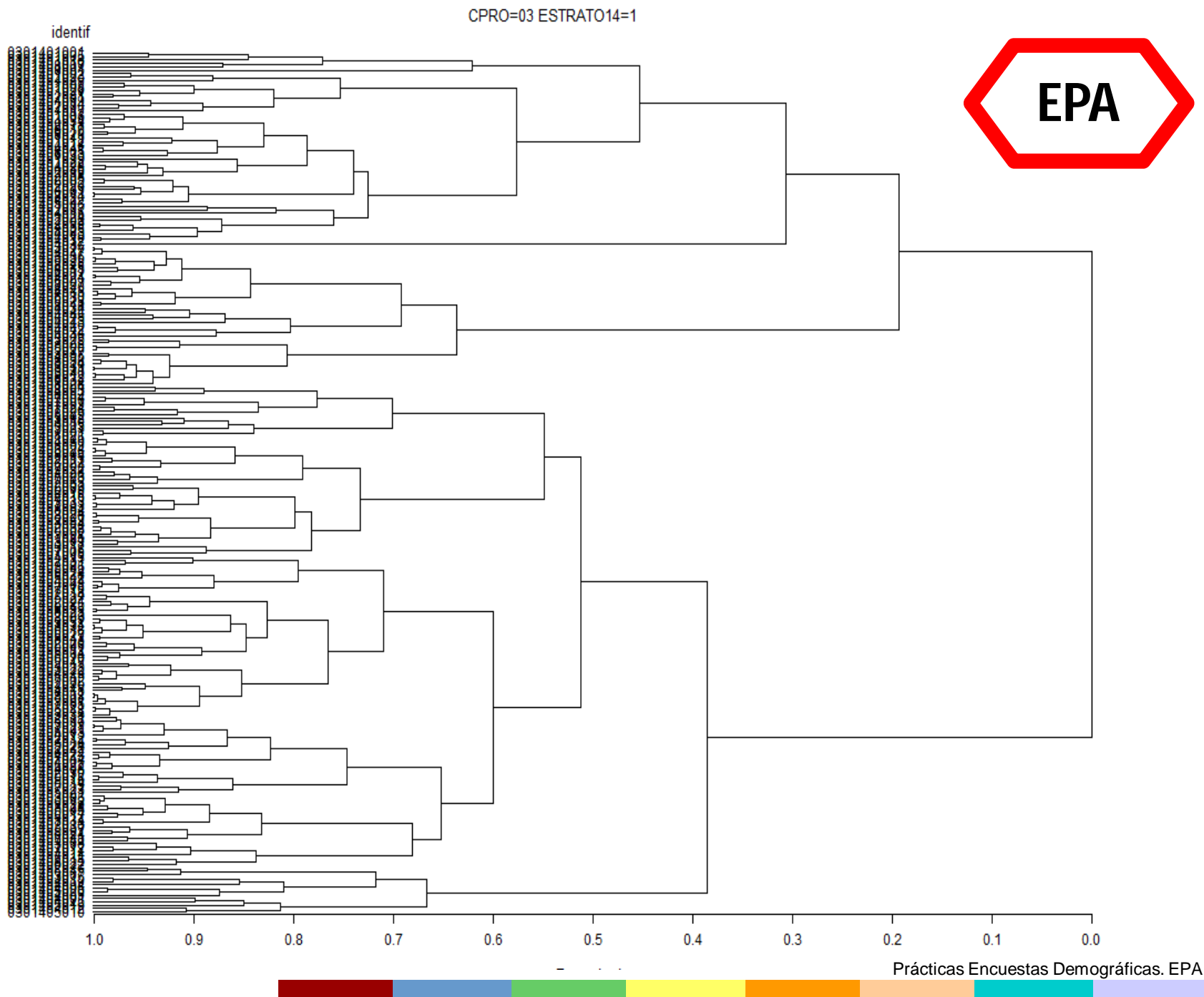
**EPA**

## 6.1 Criterios de subestratificación (IV)

**EPA**

Por último las variables fiscales que se han utilizado son:

- Renta total por vivienda con perceptores.
- Renta Capital mobiliario e inmobiliario sobre renta total.
- Renta agraria sobre renta total.
- El **algoritmo** usado para obtener los subestratos(conglomerados) ha sido el de Ward (JASA 1963). Este es un algoritmo multivariante de análisis de conglomerados jerarquizado basado en la minimización de las distancias dentro de los conglomerados. Este método está disponible en el procedimiento CLUSTER, del módulo SAS/STAT de SAS.



CPRO=03 ESTRATO14=1

CLUSTER	_TYPE_	_FREQ_	P_ANTES_ PARADOS_ Mean	P_ANTES_ INACTIVOS_ Mean	P_ANTES_ OCUPADOS_ Mean	P_ANTES_ EXTRANJEROS_ Mean	P_ANTES_ P0_14_ Mean	P_ANTES_ P15_24_ Mean
.	0	253	21.3048	36.8467	41.8484	13.7915	13.9044	10.2709
1	1	41	16.7165	24.2530	59.0306	7.6305	21.3306	9.7917
2	1	155	25.6296	37.2519	37.1185	16.8766	13.4830	10.7066
3	1	57	12.8447	44.8038	42.3515	9.8339	9.7087	9.4305
P_ANTES_ P65_Mean	P_ANTES_ ESREAL1_ 2_3_Mean	P_ANTES_ ESREAL4_5_ 6_7_Mean	P_ANTES_ ESREAL8_9_ 10_Mean	ANTES_RTO_ PERV_Mean	ANTES_ RDE_RTO_ Mean	ANTES_ RCMRCI_ RTO_Mean	ANTES_ RAA_RTO_ Mean	
18.6249	21.7727	55.5778	22.6496	31415.00	4.83001	6.4587	0.036575	
7.5581	10.1944	55.7793	34.0263	40019.96	3.08899	4.6002	0.018274	
19.1538	28.0470	58.4836	13.4694	24156.95	6.35474	4.6685	0.032680	
25.1472	13.0391	47.5309	39.4301	44962.30	1.93615	12.6636	0.060332	

CPRO=03 ESTRATO14=2

CLUSTER	_TYPE_	_FREQ_	P_ANTES_ PARADOS_ Mean	P_ANTES_ INACTIVOS_ Mean	P_ANTES_ OCUPADOS_ Mean	P_ANTES_ EXTRANJEROS_ Mean	P_ANTES_ P0_14_ Mean	P_ANTES_ P15_24_ Mean
.	0	173	27.1331	34.0857	38.7812	10.5090	15.5527	11.2109
1	1	69	26.5070	27.8128	45.6802	6.9895	19.8462	9.3967
P_ANTES_ P65_Mean	P_ANTES_ ESREAL1_ 2_3_Mean	P_ANTES_ ESREAL4_5_ 6_7_Mean	P_ANTES_ ESREAL8_9_ 10_Mean	ANTES_RTO_ PERV_Mean	ANTES_ RDE_RTO_ Mean	ANTES_ RCMRCI_ RTO_Mean	ANTES_ RAA_RTO_ Mean	
15.3562	32.8721	53.8614	13.2665	21117.24	8.1380	5.19912	0.21806	
10.9810	25.6313	58.7723	15.5965	23674.80	6.7217	4.96181	0.22790	

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



## 7. Selección

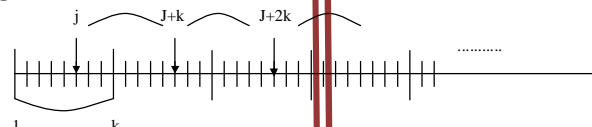
### EPA

#### Muestreo Bietápico

- 1ª Etapa: Probabilidad proporcional al tamaño, medido por el número de viviendas

$$p_j = \frac{V_j}{\sum_{j \in h} V_j}$$

- 2ª Etapa: Selección sistemática con arranque aleatorio y probabilidades iguales

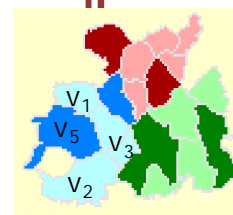


**Muestras autoponderadas a nivel de estrato**

### EES2014

#### Muestreo Trietápico

- 1ª Etapa: Probabilidad proporcional al tamaño, medido por el número de viviendas



- 2ª Etapa: Selección sistemática con arranque aleatorio y probabilidades iguales

$$P(V_{ijh}) = P(S_{jh}) \cdot P(V_{ijh} / S_{jh}) = K_h \cdot \frac{V_{jh}}{V_h} \cdot \frac{m}{V_{jh}} = \frac{K_h \cdot m}{V_h}$$

- 3ª Etapa: Un adulto de 15 o más años (Tabla Kish)

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES





## 8. Distribución de la muestra

### EPA

Cada período de la encuesta es de un **trimestre** siendo cada una de las secciones de la muestra visitada en una de las 13 semanas del mismo.

La distribución de la muestra es **uniforme** en el tiempo. Para ello se han considerado las variables provincia, estrato, turno de rotación y semana.

La totalidad de la muestra está dividida en **tres submuestras independientes** representativas, cada una de ellas, de toda la población.

Las submuestras correspondientes a cada turno de rotación son representativas, aunque su reducido tamaño impide las estimaciones en dominios medios o pequeños.

### EES2014

La muestra de secciones se **distribuye uniformemente** entre los cuatro trimestres que componen el ámbito temporal de la encuesta.

Dentro de **cada trimestre** se procura asimismo que la distribución de la muestra por semana de referencia sea lo más **homogénea posible**, con el objetivo de que **cualquier periodo del año esté igualmente representado**.

## 8.Distribución de la muestra

Provincia 28 (Madrid)

*Distribución secciones por ESTRATO/TR*

	TR_FIN						Todo
	1	2	3	4	5	6	
	N	N	N	N	N	N	
ESTRATO14							
1	14	15	15	14	15	14	87
3	5	5	6	7	5	6	34
4	3	3	3	2	2	3	16
5	2	1	1	1	2	2	9
6	1	1	.	1	1	.	4
7	1	1	1	1	1	1	6
Todo	26	26	26	26	26	26	156

## 8.Distribución de la muestra

Provincia 28 (Madrid)

10:24 Tuesday, July 15, 2014

*Distribución secciones por ESTRATO/SEMANA*

	SEMANA_FIN													Todo
	01	02	03	04	05	06	07	08	09	10	11	12	13	
	N	N	N	N	N	N	N	N	N	N	N	N	N	
<b>ESTRATO14</b>														
1	7	6	6	7	7	6	7	7	7	7	7	6	7	87
3	3	2	3	3	2	3	2	3	2	3	2	3	3	34
4	1	1	2	1	1	1	2	1	1	1	2	1	1	16
5	.	1	1	1	1	1	.	1	1	1	.	1	.	9
6	.	1	.	.	.	.	1	.	1	.	1	.	.	4
7	1	1	.	.	1	1	.	.	.	.	.	1	1	6
<b>Todo</b>	12	12	12	12	12	12	12	12	12	12	12	12	12	156

## 8.Distribución de la muestra

### Provincia 28 (Madrid)

*Distribución secciones por TR/SEMANA*

10:24 Tuesday, July 15, 2014

	SEMANA_FIN													Todo
	01	02	03	04	05	06	07	08	09	10	11	12	13	
	N	N	N	N	N	N	N	N	N	N	N	N	N	
TR_FIN														
1	2	4	1	2	1	2	2	2	3	1	2	1	3	26
2	3	1	1	3	2	4	1	2	2	3	1	2	1	26
3	2	2	3		2	1	2	3	2	2	3	2	2	26
4	2	2	2	3	3	3	2		2	2	2	2	1	26
5	1	1	2	2	2	1	4	2	1	2	2	2	4	26
6	2	2	3	2	2	1	1	3	2	2	2	3	1	26
Todo	12	12	12	12	12	12	12	12	12	12	12	12	12	156

## 8.Distribución de la muestra

Provincia 28 (Madrid)

*CPRO=28 Distribución secciones por BLOQUE/SEMANA*

	SEMANA_FIN													Todo
	01	02	03	04	05	06	07	08	09	10	11	12	13	
	N	N	N	N	N	N	N	N	N	N	N	N	N	
BLOQUE_FIN														
01	1	1	1	1	1	1	1	1	1	1	1	1	1	13
02	1	1	1	1	1	1	1	1	1	1	1	1	1	13
03	1	1	1	1	1	1	1	1	1	1	1	1	1	13
04	1	1	1	1	1	1	1	1	1	1	1	1	1	13
05	1	1	1	1	1	1	1	1	1	1	1	1	1	13
06	1	1	1	1	1	1	1	1	1	1	1	1	1	13
07	1	1	1	1	1	1	1	1	1	1	1	1	1	13
08	1	1	1	1	1	1	1	1	1	1	1	1	1	13
09	1	1	1	1	1	1	1	1	1	1	1	1	1	13
10	1	1	1	1	1	1	1	1	1	1	1	1	1	13
11	1	1	1	1	1	1	1	1	1	1	1	1	1	13
12	1	1	1	1	1	1	1	1	1	1	1	1	1	13
Todo	12	12	12	12	12	12	12	12	12	12	12	12	12	156

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



## 9. Renovación parcial de la muestra

EPA

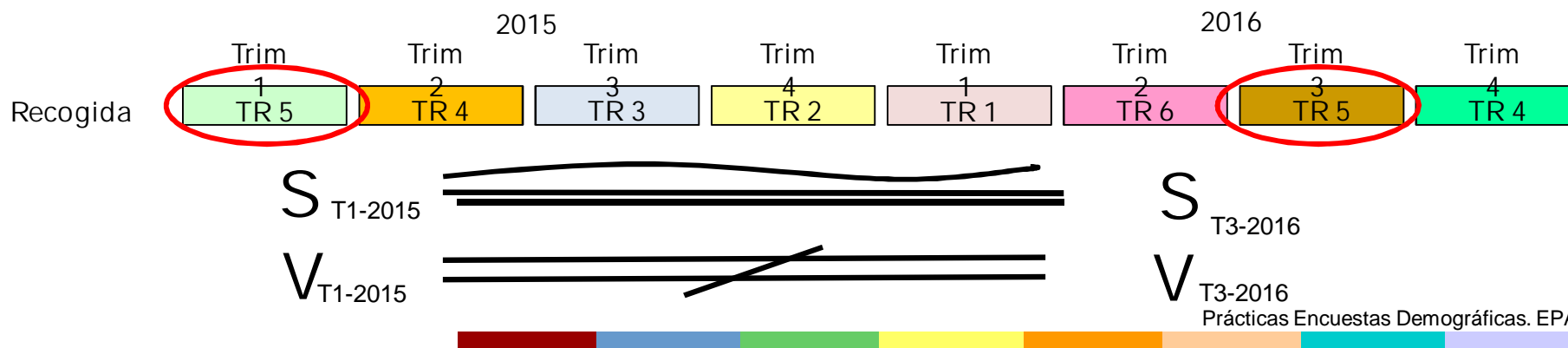
→ Unidades de **primera** etapa: **Secciones censales**. Estas permanecen fijas indefinidamente en la muestra salvo en las excepciones que se señalan :

- Agotamiento de los hogares consultables
- Actualización de probabilidades de selección
- Variaciones del seccionado

→ Unidades de **segunda** etapa: **Viviendas principales**

Las viviendas de la muestra son renovadas parcialmente cada trimestre. Esta renovación afecta a una sexta parte de las secciones (5/6 permanecen de un trimestre a otro) y se realiza mediante el siguiente procedimiento:

- 1.- **Turnos de rotación:** El conjunto de las secciones de la muestra está repartido en 6 grupos llamados turnos de rotación.
- 2.- Cada trimestre, las viviendas correspondientes a las secciones de un determinado turno de rotación son renovadas en su totalidad.

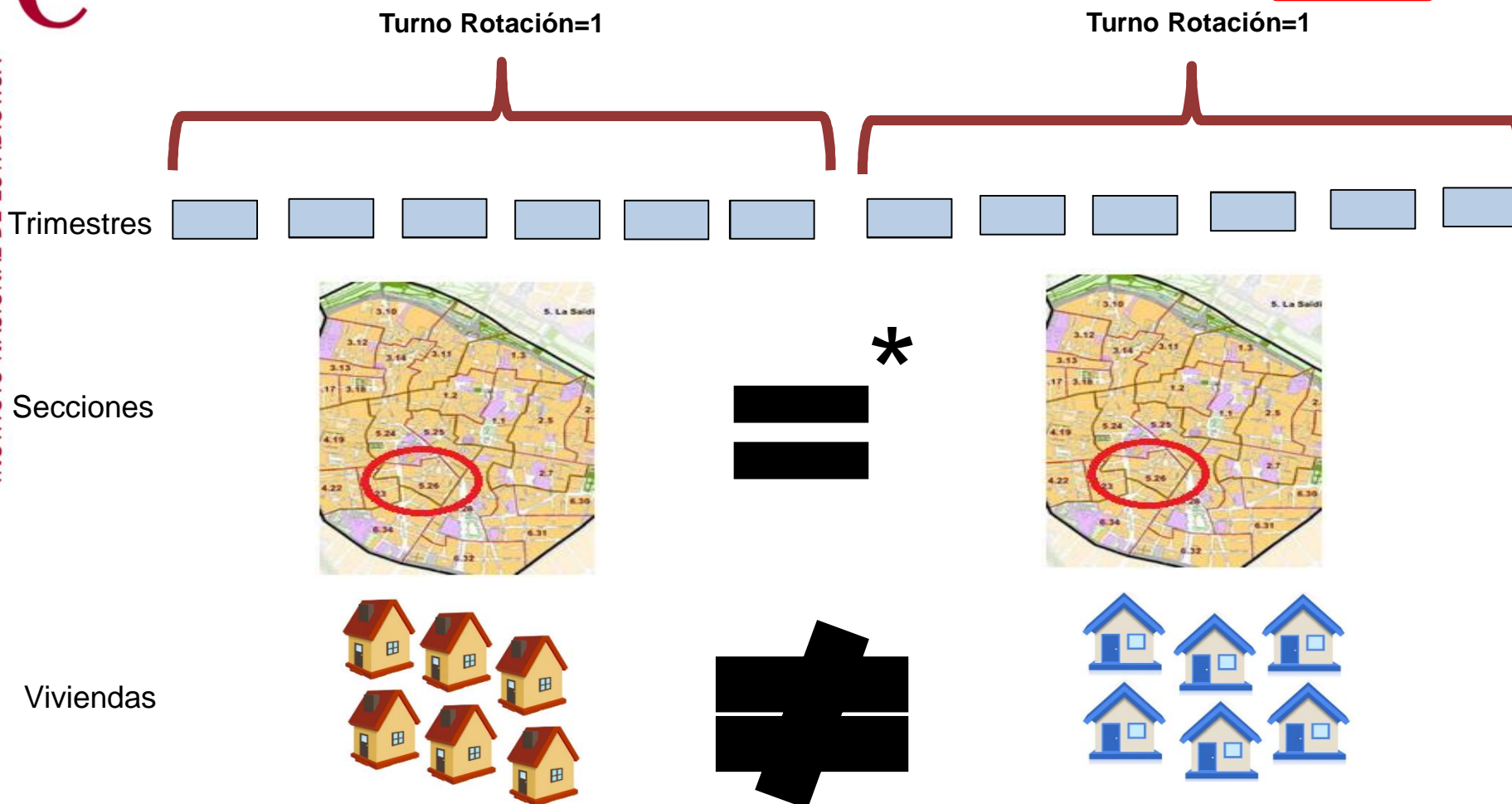




## 9. Renovación parcial de la muestra

EPA

INSTITUTO NACIONAL DE ESTADISTICA



(\*)Salvo por actualización de probabilidades de selección , Variaciones del seccionado ...





## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



## 10.Actualización de las unidades de muestreo

Las continuas variaciones de población ya sea en sus características, o bien en su distribución espacial exigen realizar actualizaciones en los marcos, que repercuten en la estructura muestral.

En la E.P.A. hay que considerar tres tipos de actualizaciones:



EPA

**Actualización de secciones censales** como consecuencia de variaciones(particiones, fusiones,...) en las unidades primarias seleccionadas para la muestra.

**Actualización en el marco de viviendas**, restringida a las secciones de la muestra.

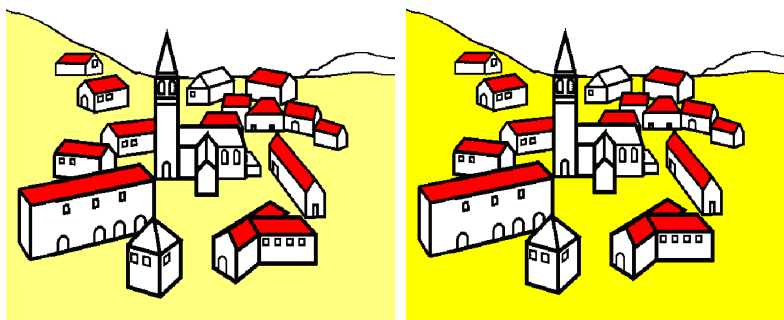
**Actualización de carácter general** relativa a todas las secciones y viviendas, cuando se realizan los Censos de Población.

# 10. Actualización de las unidades de muestreo

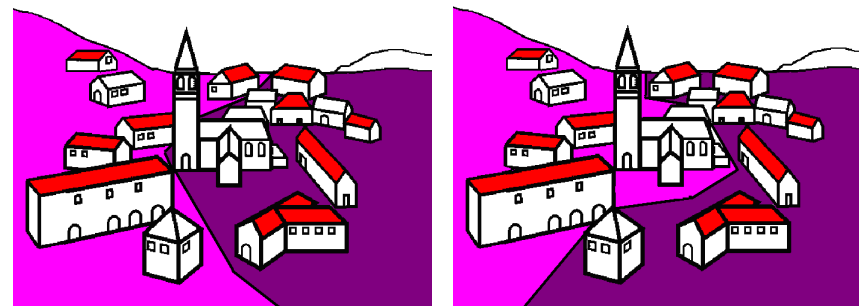
## Actualización de secciones censales

**EPA**

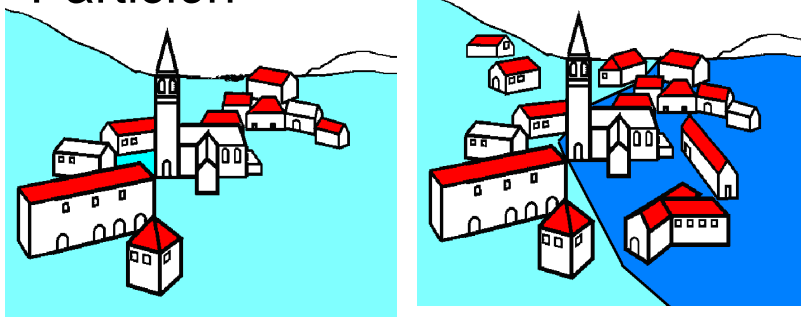
### Cambio de Denominación



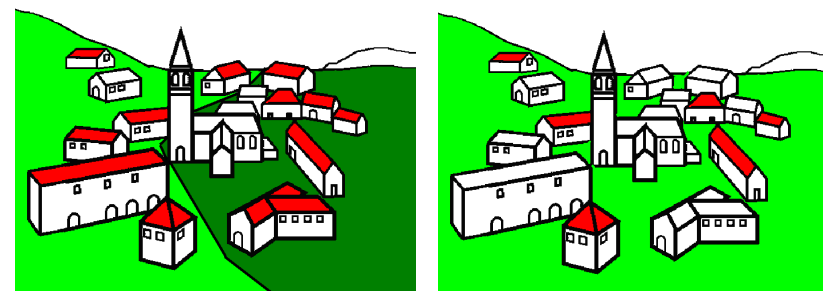
### Delimitación Defectuosa



### Partición

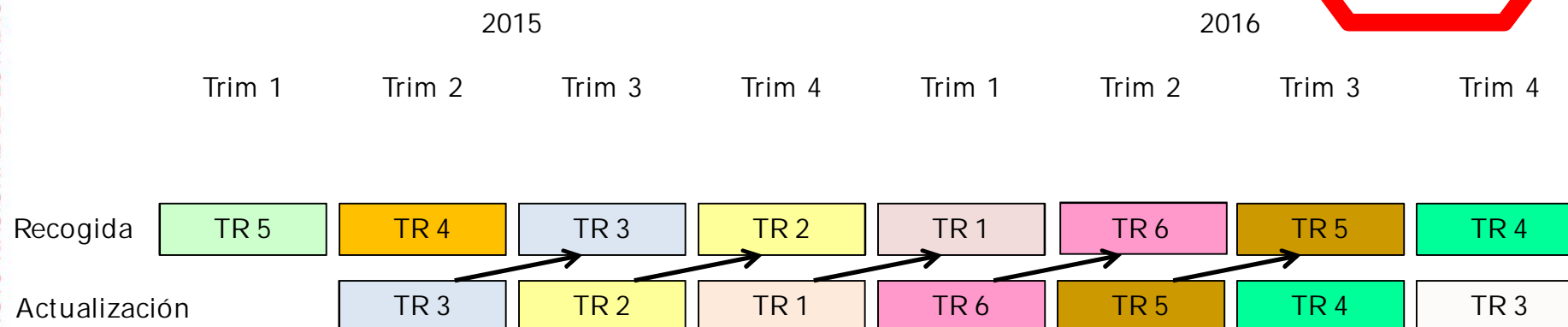


### Fusión



## 10. Actualización de las unidades de muestreo

### Actualización en el marco de viviendas



Se realiza un recorrido exhaustivo de la sección incorporando al directorio las viviendas las nuevas que hayan aparecido

➡ La actualización se concreta básicamente en actuar sobre la constante K

$$K' = K \frac{V'_s}{V_s} \text{ donde}$$

$V'_s = n^\circ \text{ de familias en la sección } S \text{ después de la actualización}$

## 10.Actualización de las unidades de muestreo

Tramas-EPA

Predepuración

Secciones

Actualización ▾

Variaciones

Muestra ▾

Administración ▾

EPA

¡Bienvenido a  
Tramas-EPA!

Portal para la gestión del marco de secciones y huecos



## 10. Actualización de las unidades de muestreo

### Actualización de carácter general

- Cuando se dispone de información procedente de los Censos de Población
- Los cambios se introducen por turnos de rotación
- La forma más directa de actualizar probabilidades de selección es la selección de una nueva muestra a partir del marco disponible más actualizado ➡ GENERA INCONVENIENTES
- Se decide arbitrar un procedimiento que sin distorsionar las probabilidades de selección que realmente corresponden a cada sección, mantenga la muestra de secciones con las mínimas variaciones ➡ Método de Brick, Morganstein y Wolter(1987), basado en el método de Kish y Scott

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES



## 11. Estimadores

El proceso habitual para la obtención de estimadores en encuestas demográficas en general es :

1. **Estimador insesgado de expansión(Horvitz-Thompson):** Compensa las desiguales probabilidades de selección.
2. **Corrección de la falta de respuesta:** Corrige el sesgo producido en las estimaciones por la falta de respuesta total de algunos elementos.
3. **Calibrado con fuentes externas:** Reduce la varianza de las estimaciones mediante la utilización de fuentes auxiliares externas y puede actualizar la estimación en el tiempo.

Como resultado de este proceso se obtiene finalmente un factor de elevación para cada elemento de la muestra efectiva.



## 11. Estimadores


**EPA**

### 1.- Estimador insesgado de expansión(H-T)

Recordamos que la probabilidad de *pertenecer a la muestra* de una vivienda 'i' de la sección 'j' del estrato 'h' viene dada por:

$$P(V_{ijh}) = P(Sec_{jh}) \cdot P(V_{ijh} / Sec_{jh}) = K_h \cdot \frac{V_{jh}}{V_h} \cdot \frac{m}{V_{jh}} = \frac{K_h \cdot m}{V_h}$$

Donde  $K_h$  son las secciones de la muestra en el estrato "h", y "m" es el número de viviendas muestrales por sección.

## 11. Estimadores

Según lo anterior, la probabilidad de pertenecer a la muestra se puede expresar por:

$$P(V_{ijh}) = \frac{v_h^t}{V_h}$$

**EPA**

Siendo  $v_h^t$  el número teórico de viviendas de la muestra en el estrato “h”.

Por tanto el estimador H-T tendrá la expresión:

$$\hat{Y}_{H-T} = \sum_h \frac{V_h}{v_h^t} \cdot \sum_{i \in h} y_i$$

## 11. Estimadores

EPA

### 2.- Corrección de la falta de respuesta

La probabilidad de respuesta por estrato la podemos estimar por:

$$P_{Rh} = \frac{v_h}{v_h^t}$$

Donde  $v_h$  representa la muestra efectiva de viviendas en el estrato  $h$ .

Por tanto el estimador corregido será:

$$\hat{Y}_{H-TCorr} = \sum_h \frac{V_h}{v_h^t} \cdot \frac{v_h^t}{v_h} \sum_{i \in h} y_i = \sum_h \frac{V_h}{v_h} \sum_{i \in h} y_i = \sum_h \hat{Y}_{H-TCorr(h)}$$

## 11. Estimadores

### 3.- Calibrado con fuentes externas(1)

EPA

Planteamiento del problema:

$$\min_{w_k} \sum_{k \in S} \frac{d_k}{q_k} G\left(\frac{w_k}{d_k}\right)$$

Condicionado a

$$\sum_{k \in S} w_k x_k = X$$

$$x'_k = (x_{k1}, x_{k2}, \dots, x_{kJ}) \quad X' = (X_1, X_2, \dots, X_J)$$

donde,  $q_k$  es un término corrector de la posible heterocedasticidad

$G(x)$  **función de distancia G**, positiva y estrictamente convexa:

tal que  $G(1) = G'(1) = 0$ , con objeto de garantizar la solución

$$G''(1) = 1$$

## 11. Estimadores

### 3.- Calibrado con fuentes externas (EPA)

EPA



# IN 11. Estimadores

EES14

En el caso de la EES14, se obtienen dos factores :

- A.\_ Las Estimaciones de Hogares ( y personas residentes)
- B.\_ Las Estimaciones basadas en la persona seleccionada

## A. Estimaciones de hogares

### 1.-Estimador basado en el diseño

$$\hat{Y}_d = \sum_h \sum_{i,j \in h} \frac{1}{K_h \cdot \frac{15}{V_h^{(13)}}} \cdot y_{hij} = \sum_h \sum_{i,j \in h} \frac{V_h^{(13)}}{v_h^t} \cdot y_{hij}$$

### 2.- Corrección de la falta de respuesta/Estimador de razón

$$F_j^{(1)} = \frac{P_h}{p_h^e} \quad \hat{Y}_3 = \sum_h \sum_{i,j \in h} F_j^{(1)} \cdot y_{hij}$$

### 3.-Técnicas de Calibrado

- Grupos de edad y sexo. Hombres y mujeres distribuidos en los siguientes grupos de edad 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65 y más.
- Totales poblacionales por provincia.
- Totales de hogares por provincia
- Hogares por tamaño: 1, 2, 3 y 4 ó más miembros.

## B. Estimaciones de personas

### 1.-Factor inicial $F_j^{(1)}$

### 2.- Factor de la persona seleccionada del hogar j

$$F_{jk}^{(3)} = F_j^{(1)} A_j$$

donde el subíndice jk representa la persona (de 15 y más años) k del hogar j que debe cumplimentar cuestionario individual y donde  $A_j$  es el número de personas de 15 y más años del hogar j.

### 3.-Técnicas de Calibrado

- Población total de 15 y más años por nacionalidad, española o extranjera
- La población por grupos de edad y sexo: Hombres y mujeres de 15-19, 20-24, 25-29, 30-34, 35-39, 40-44, 45-49, 50-54, 55-59, 60-64, 65 y más años.

## Aspectos generales de la muestra: EPA vs EES14

1. INTRODUCCIÓN
2. ÁMBITO DE LA ENCUESTA
3. MARCO DE LA ENCUESTA
4. TIPO DE MUESTREO
5. TAMAÑO DE LA MUESTRA. AFIJACIÓN
6. CRITERIO DE ESTRATIFICACIÓN
7. SELECCIÓN
8. DISTRIBUCIÓN DE LA MUESTRA
9. RENOVACIÓN PARCIAL DE LA MUESTRA
10. ACTUALIZACIÓN DE LAS UNIDADES DE MUESTREO
11. ESTIMADORES
12. ERRORES

## 12. Errores

EPA

Método de las semimuestras reiteradas

Es un método indirecto de estimación de la varianza, que requiere para su aplicación:

Obtención de 'r' semimuestras de la muestra total

Estimación de la característica 'X' con cada semimuestra

La varianza se estima mediante la expresión:

$$\hat{V}(\hat{X}) = \frac{1}{r} \sum_{i=1}^r (\hat{X}_i - \hat{X})^2$$

donde:

r es el número de semimuestras

$\hat{X}_i$  es la estimación con la i-ésima semimuestra

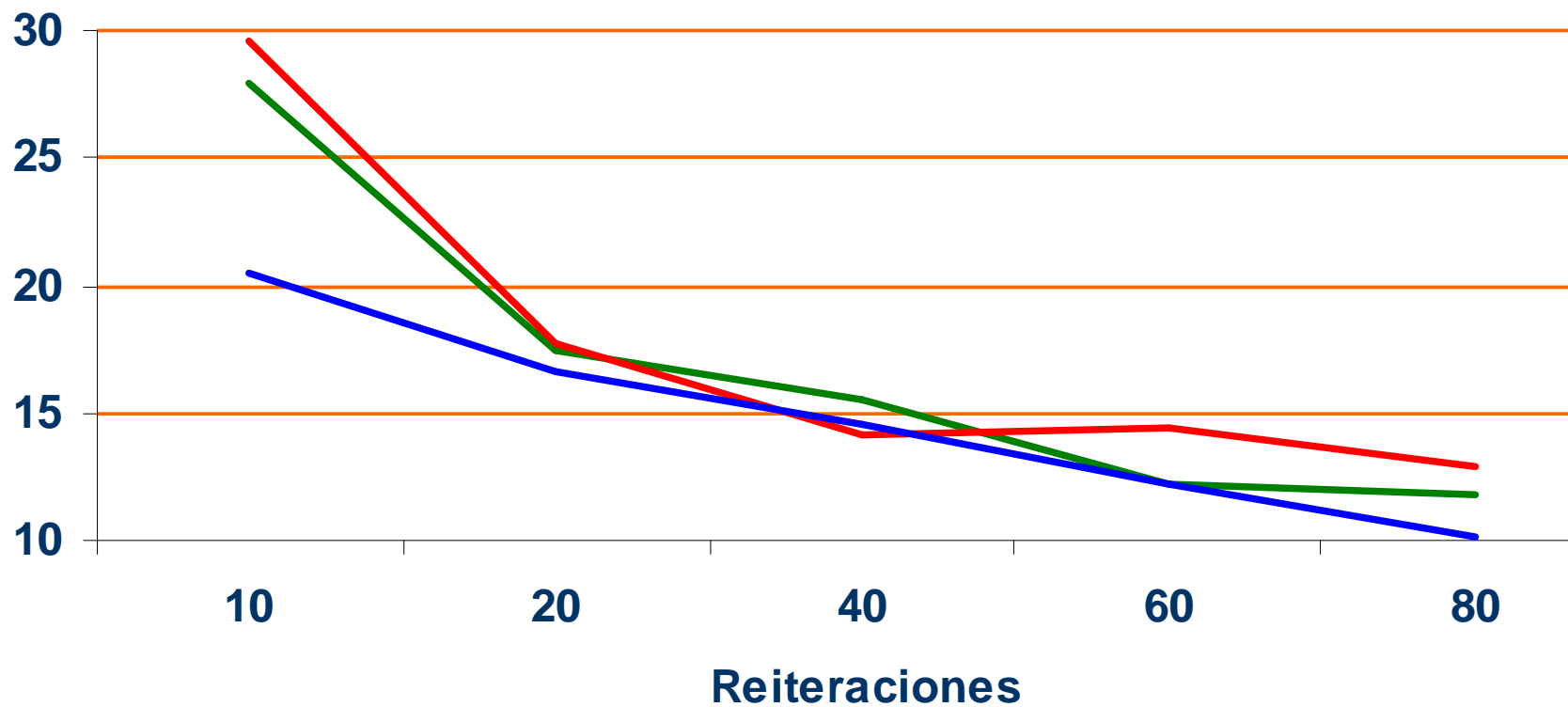
$\hat{X}$  es la estimación obtenida con toda la muestra

En la aplicación de este método se tiene en cuenta que:

- Cada semimuestra queda constituida por un número de secciones equivalente a la mitad de la muestra.
- Cada sección aparece en la mitad de las semimuestras.



## C.V. DE LA VARIABLE CV (PARO)



— Murcia — Asturias — Cantabria

## 12. Errores

EES14

### Método de Jackknife

Este permite obtener la estimación de la varianza del estimador del total de una característica  $Y$  mediante la expresión:

$$\hat{V}(\hat{Y}) = \sum_h \frac{n_h - 1}{n_h} \sum_{i \in h} (\hat{Y}_{(hi)} - \hat{Y})^2$$

siendo:

$\hat{Y}$  la estimación del total de la característica  $Y$  obtenida con la muestra completa

$\hat{Y}_{(hi)}$  la estimación del total de la característica  $Y$  obtenida después de quitar de la muestra las unidades de la sección  $i$  del estrato  $h$ .

$n_h$  el número de secciones de la muestra en el estrato  $h$ .

Para obtener el estimador  $\hat{Y}_{(hi)}$ , y por simplicidad, en vez de recalcular los factores de elevación (con corrección de falta de respuesta, calibrado, etc...), se multiplican los factores del estrato donde se ha quitado la sección, por el factor  $\frac{n_h}{n_h - 1}$

$$\hat{Y} = \sum_{l,j,k \in S} F_{ljk} y_{ljk}$$

$$\hat{Y}_{(hi)} = \sum_{l \neq h} F_{ljk} y_{ljk} + \sum_{\substack{l=h \\ j \neq i}} F_{ljk} \frac{n_h}{n_h - 1} y_{ljk}$$