

Ejemplos Resueltos Tema 2

2012

Ejemplos de tablas de contingencia

La tabla de datos siguiente muestra la población en Granada (de 16 años o más) clasificada por Nacionalidad y Relación con la Actividad Laboral en el años 2005 y en el 2011.

DATOS

Los datos se han obtenido de la Encuesta de Población Activa para los periodos 2005 y 2011 y se muestran en la tabla siguiente formada por 4 columnas:

DATOS: Población en Granada de 16 o más años según Nacionalidad y Relación con la Actividad (años 2005 y 2011)

Nacionalidad	RelActi	Año	Pobmas16
Español	ocupado	2005	291441
Español	parado	2005	42343
Español	inactivo	2005	331378
Extranjero	ocupado	2005	22445
Extranjero	parado	2005	4154
Extranjero	inactivo	2005	14317
Español	ocupado	2011	271094
Español	parado	2011	107695
Español	inactivo	2011	306920
Extranjero	ocupado	2011	28616
Extranjero	parado	2011	18537
Extranjero	inactivo	2011	17994

Construiremos la distribución bidimensional de Nacionalidad y Relación con Actividad y las correspondientes distribuciones de frecuencias undimensionales:

Marginales:

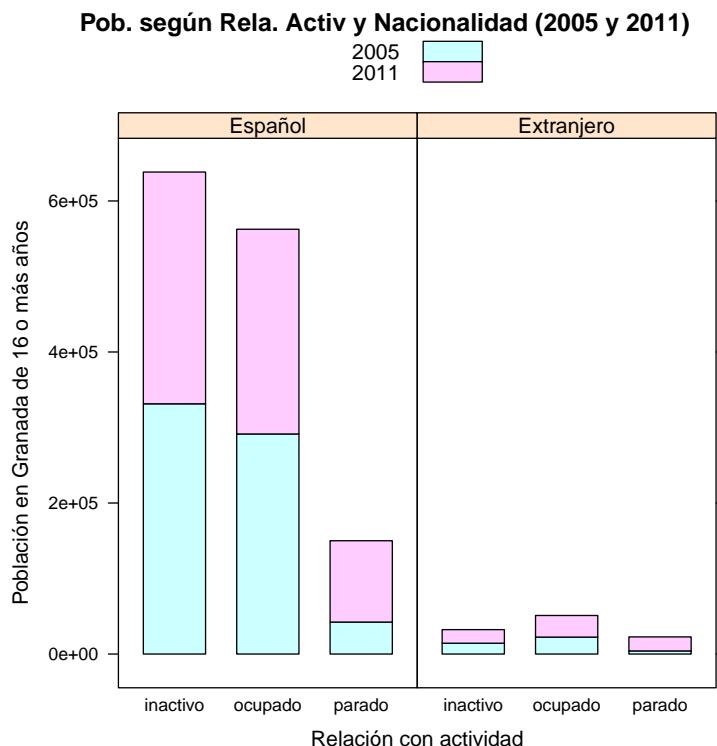
- 1. de Nacionalidad y
- 2. de Relación con la Actividad

Condicionadas:

- 1. de Nacionalidad a los valores de Relación con la Actividad y
- 2. de Relación con la Actividad a los valores de Nacionalidad.

Expresaremos en frecuencias relativas la distribución bidimesional, condicionadas y marginales.

0.1. Representación gráfica de los datos



Distribución bidimensional en frecuencias absolutas y en frecuencias relativas

Distribución bidimensional de Nacionalidad y Relación Actividad (año 2005):

Nacionalidad		
RelActi	Español	Extranjero
inactivo	331378	14317
ocupado	291441	22445
parado	42343	4154

Distribución bidimensional de Nacionalidad y Relación Actividad (año 2011):

Nacionalidad		
RelActi	Español	Extranjero
inactivo	306920	17994
ocupado	271094	28616
parado	107695	18537

Distribución bidimensional de Nacionalidad y Relación Actividad (año 2011) en frecuencias relativas.

Nacionalidad			
RelActi	Español	Extranjero	Suma
inactivo	0.409	0.024	0.433
ocupado	0.361	0.038	0.399
parado	0.143	0.025	0.168
Suma	0.913	0.087	1.000

Distribuciones marginales en frecuencias relativas y absolutas(año 2011)

Marginal de la variable Relación con la Actividad (RelActi)

RelActi	FrecAbs
inactivo	324914.00
ocupado	299710.00
parado	126232.00

RelActi	FrecRel
inactivo	0.43
ocupado	0.40
parado	0.17

Marginal de la variable Nacionalidad

Nacionalidad	FrecAbs
Español	685709.00
Extranjero	65147.00

Nacionalidad	FrecRel
Español	0.91
Extranjero	0.09

Distribuciones marginales en frecuencias relativas y absolutas(año 2005)

De modo similar se obtienen las distribuciones marginales correspondientes a la bidimensional del año 2005. Lo dejamos propuesto.

Gráfico de barras para la distribución marginal de Relación con Actividad (2011)

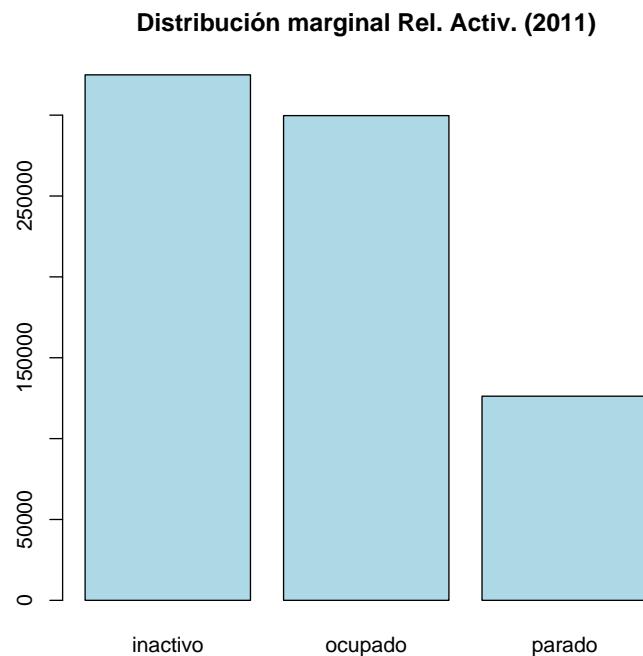
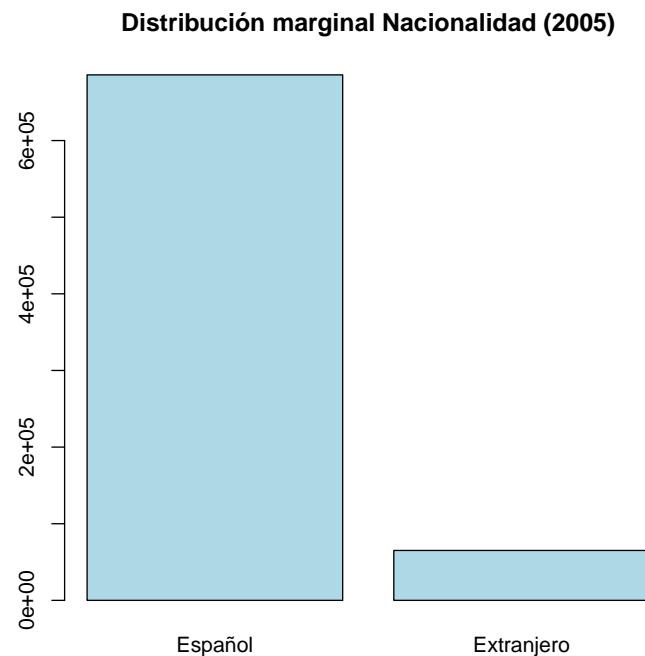


Gráfico de barras para la distribución marginal de Nacionalidad (2011)



0.2. Distribuciones condicionadas

Distribuciones Condicionadas De Relación Actividad a los valores de Nacionalidad (2011)

	Español	Extranjero
inactivo	0.448	0.276
ocupado	0.395	0.439
parado	0.157	0.285
Sum	1.000	1.000

Distribuciones Condicionadas De Nacionalidad a los valores de Relación Actividad (2011)

	Español	Extranjero	Sum
inactivo	0.945	0.055	1.000
ocupado	0.905	0.095	1.000
parado	0.853	0.147	1.000

Estadístico de asociación Chi-cuadrado

El estadístico Chi-cuadrado permite estudiar la asociación entre dos variables en una tabla de contingencia.

$$\chi^2 = \sum \frac{(n_{ij} - t_{ij})^2}{t_{ij}}$$

donde las frecuencias teóricas t_{ij} representan los valores frecuenciales correspondientes a la hipótesis de independencia de las variables. Se obtienen como el producto de las frecuencias marginales dividido por el total, N, de la población

$$t_{ij} = \frac{n_{i*}n_{*j}}{N}$$

Estadístico de asociación Chi-cuadrado en la distribución bidimensional de RelActi y Nacionalidad

Para la tabla de contingencia dada por

		Nacionalidad		Sum
RelActi		Español	Extranjero	
inactivo		306920	17994	324914
ocupado		271094	28616	299710
parado		107695	18537	126232
Sum		685709	65147	750856

se obtienen las frecuencias teóricas o frecuencias estimadas bajo independencia siguientes:

		Nacionalidad		Sum
RelActi		Español	Extranjero	
inactivo		296723.3	28190.72	
ocupado		273706.1	26003.93	
parado		115279.7	10952.35	

Por ejemplo, $t_{11} = \frac{324914 \times 685709}{750856} = 296723,3$, $t_{12} = \frac{324914 \times 65147}{750856} = 28190,72$, etc.

Se obtiene cada sumando del estadístico Chi-cuadrado, $r_{ij} = \frac{(n_{ij} - t_{ij})^2}{t_{ij}}$, la suma de los cuales da lugar al estadístico Chi-cuadrado, 10077,41.

nij	tij	rij
306920	296723,28	350,40
271094	273706,07	24,93
107695	115279,65	499,02
17994	28190,72	3688,20
28616	26003,93	262,38
18537	10952,35	5252,48
		10077,41

$$\chi^2 = \frac{(306920 - 296723,28)^2}{296723,28} + \frac{(271094 - 273706,07)^2}{273706,07} + \dots + \frac{(18537 - 10952,35)^2}{10952,35} =$$

$$= 350,40 + 24,93 + \dots + 5252,48 = 10077,41$$