

ESTADÍSTICA



GRADO TURISMO

TEMA 4: ANÁLISIS CUALITATIVO DEL TURISMO



TEMA 4: ANÁLISIS CUALITATIVO DEL TURISMO

1. Asociación entre variables nominales
2. Correlación entre variables ordinales

1.- ASOCIACIÓN ENTRE VARIABLES NOMINALES

74127 turistas extranjeros llegados en agosto de 2009 a la Comunitat Valenciana han sido clasificados según la **Zona Turística de destino** y su **Procedencia** dando lugar a la siguiente tabla de contingencia:

Zona Turística	Procedencia	Francia	Reino Unido
Costa Blanca (Alicante)		12136	52111
Costa Azahar (Castellón)		7081	1464
Costa Valencia (Valencia)		1085	250

Fuente: Encuesta de ocupación hotelera 2009, INE.

Objetivo: Calcular coeficientes que nos midan el grado de asociación existente entre las variables nominales o atributos: **Zona Turística de destino** y **Procedencia**. Para ello definimos los siguientes tipos de frecuencias:

- Frecuencias Observadas conjuntas: O_{ij}
- Frecuencias marginales: n_i y n_j → Si las calculamos, la tabla quedaría:

Zona Turística	Procedencia	Francia	Reino Unido	Totales n_i
Costa Blanca (Alicante)		12136	52111	64247
Costa Azahar (Castellón)		7081	1464	8545
Costa Valencia (Valencia)		1085	250	1335
Totales	n_j	20302	53825	74127=N

- Frecuencias Esperadas o Teóricas: E_{ij} serían las frecuencias conjuntas que obtendríamos en el caso hipotético de que no hubiera ninguna relación entre las variables, es decir que las variables fueran **INDEPENDIENTES**. Estas frecuencias esperadas se calculan como:
$$\frac{n_i \times n_j}{N}$$
 La tabla quedaría:

Zona Turística	Procedencia	Francia	Reino Unido
Costa Blanca (Alicante)		17596,1	46650,9
Costa Azahar (Castellón)		2340,3	6204,7
Costa Valencia (Valencia)		365,6	969,4

Medidas de Asociación

1º calculamos el **estadístico** χ^2 (Chi-Cuadrado) que compara para cada celda de la tabla de contingencia las frecuencias observadas con las esperadas. Se define:

$$\chi^2 = \sum_{i=1}^I \sum_{j=1}^J \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

Interpretación del **estadístico** χ^2 :

- Si las variables fueran independientes, entonces $O_{ij} = E_{ij}$ y el **estadístico** $\chi^2 = 0$
- Si el **estadístico** $\chi^2 \neq 0$ existe algún tipo de asociación entre las variables.

$$\begin{aligned} \chi^2 &= \frac{(12136 - 17596,1)^2}{17596,1} + \frac{(52111 - 46650,9)^2}{46650,9} + \frac{(7081 - 2340,3)^2}{2340,3} \\ &= \frac{(1464 - 6204,7)^2}{2340,3} + \frac{(1085 - 365,6)^2}{365,6} + \frac{(250 - 969,4)^2}{969,4} = 17507,6 \end{aligned}$$

Inconveniente de χ^2 : su límite inferior es cero, pero no tiene límite superior.

El estadístico χ^2 se puede calcular con la Caest α :



CaEst 1.5

DESCRIPTIVA ▼ PROBABILIDAD ▼ INFERENCIA ▼ MUESTRAS ▼ ANOVA ▼ MISCELÁNEO

- INTERVALOS DE CONFIANZA ▶
- CONTRASTES PARAMÉTRICOS ▶
- CONTRASTES NO PARAMÉTRICOS ▶

- Contraste de homogeneidad , tabla de 2x2
- **Contraste de homogeneidad , tabla de 5x5 máximo**
- Contraste de bondad del ajuste , test de la CHI2
- Prueba de Mann-Whiyney
- Prueba de Kruskal-Wallis, tres muestras
- Prueba de Kruskal-Wallis, cuatro muestras
- Test de signos

procede **CEA**
CES

CONTRASTE DE HOMOGENEIDAD (válido par independencia) DE LA CHI-2:

las correcciones de Yates son automáticas

Introduce el número de filas y columnas y después sus valores

filas

columnas

activar

	y_1	y_2	y_3	y_4	y_5	Total
x_1	<input type="text" value="12136"/>	<input type="text" value="52111"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="64247"/>
x_2	<input type="text" value="7081"/>	<input type="text" value="1464"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="8545"/>
x_3	<input type="text" value="1085"/>	<input type="text" value="250"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="1335"/>
x_4	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>
x_5	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>
Total	<input type="text" value="20302"/>	<input type="text" value="53825"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="----"/>	<input type="text" value="74127"/>

calcular estadístico

Estadístico

Coeficiente de Contingencia C de Pearson

$$C = \sqrt{\frac{\chi^2}{N + \chi^2}} = \sqrt{\frac{17507,6}{74127 + 17507,6}} = 0,437 \quad \boxed{0 \leq C < 1}$$

Interpretación del coeficiente de contingencia C:

- Si las variables fueran independientes, entonces $C=0$.
- Cuanto más cerca de 0, menos grado de asociación entre las variables.
- Cuanto mayor sea el valor de C, más grado de asociación habrá entre las variables.

Inconveniente de C: nunca alcanza el valor 1 en el caso de asociación perfecta.

Coeficiente V de Cramer

$$V = \sqrt{\frac{\chi^2}{N \times (k-1)}} = \sqrt{\frac{17507,6}{74127 \times (2-1)}} = 0,486 \quad \boxed{0 \leq V \leq 1}$$

k es el mínimo entre el n° de filas (=3) y el n° de columnas (=2)

- El coeficiente V se interpreta igual que del coeficiente de contingencia C, con la ventaja de que puede alcanzar el valor 1 en el caso de asociación perfecta entre las variables (mayor grado de asociación posible).
- Si las variables fueran independientes, entonces $V=0$ (menor grado de asociación posible).

2.- CORRELACIÓN ENTRE VARIABLES ORDINALES

Se dispone de información para 2009 para las N=17 Comunidades Autónomas sobre:

- Turistas internacionales según comunidad autónoma de destino principal, en miles de personas.
- Gasto total de los turistas según comunidad autónoma de destino principal, en millones de euros.

CCAA	Turistas	Gasto total
Andalucía	7359,47	7337,69
Aragón	352,16	365,91
Asturias	167,96	212,64
Baleares	9037,39	8790,67
Canarias	8203,56	9082,28
Cantabria	315,03	323,24
Castilla La Mancha	187,19	216,89
Castilla-León	982,36	680,69
Cataluña	12769,13	9643,02
Comunidad Valenciana	5108,89	4101,83
Extremadura	186,76	118,48
Galicia	762,34	682,79
Madrid	4895,78	5226,86
Murcia	845,63	775,10
Navarra	197,80	142,98
País Vasco	803,63	498,04
Rioja (La)	56,03	43,33

Fuente: IET. Movimientos Turísticos en Fronteras (Frontur) y Encuesta de Gasto Turístico (Egatur) 2009.

Si ordenamos las CCAA en orden descendente:
•1º respecto a los Turistas internacionales.



Se obtienen las siguientes ordenaciones o rankings de CCAA que llamamos **Rangos**:

CCAA	Turistas	Gasto total	Rangos Turistas
Cataluña	12769,13	9643,02	1
Baleares	9037,39	8790,67	2
Canarias	8203,56	9082,28	3
Andalucía	7359,47	7337,69	4
Comunidad Valenciana	5108,89	4101,83	5
Madrid	4895,78	5226,86	6
Castilla-León	982,36	680,69	7
Murcia	845,63	775,10	8
País Vasco	803,63	498,04	9
Galicia	762,34	682,79	10
Aragón	352,16	365,91	11
Cantabria	315,03	323,24	12
Navarra	197,80	142,98	13
Castilla La Mancha	187,19	216,89	14
Extremadura	186,76	118,48	15
Asturias	167,96	212,64	16
Rioja (La)	56,03	43,33	17

Fuente: IET. Frontur y Egatur 2009.

Si ordenamos las CCAA en orden descendente:
•2º respecto al Gasto total de los turistas.



Se obtienen las siguientes ordenaciones o rankings de CCAA que llamamos **Rangos**:

CCAA	Turistas	Gasto total	Rangos Turistas	Rangos Gasto total
Cataluña	12769,13	9643,02	1	1
Canarias	8203,56	9082,28	3	2
Baleares	9037,39	8790,67	2	3
Andalucía	7359,47	7337,69	4	4
Madrid	4895,78	5226,86	6	5
Comunidad Valenciana	5108,89	4101,83	5	6
Murcia	845,63	775,10	8	7
Galicia	762,34	682,79	10	8
Castilla-León	982,36	680,69	7	9
País Vasco	803,63	498,04	9	10
Aragón	352,16	365,91	11	11
Cantabria	315,03	323,24	12	12
Castilla La Mancha	187,19	216,89	14	13
Asturias	167,96	212,64	16	14
Navarra	197,80	142,98	13	15
Extremadura	186,76	118,48	15	16
Rioja (La)	56,03	43,33	17	17

Los dos Rangos que hemos obtenido son las variables ordinales u ordenaciones.

Objetivo: Calcular un coeficiente que mida si hay o no relación (correlación o concordancia) entre los dos criterios de ordenación de las 17 CCAA (Rangos Turistas internacionales y Rangos Gasto total).

Coefficiente de correlación ordinal o por rangos de Spearman

Para calcular este coeficiente, primero se calcula para cada CCAA las diferencias entre los rangos obtenidos en cada ordenación. Además, estas diferencias se elevan al cuadrado $\rightarrow d_i^2$ y se suman.

CCAA	Turistas	Gasto total	Rangos Turistas	Rangos Gasto total	d_i^2
Cataluña	12769,13	9643,02	1	1	0
Canarias	8203,56	9082,28	3	2	1
Baleares	9037,39	8790,67	2	3	1
Andalucía	7359,47	7337,69	4	4	0
Madrid	4895,78	5226,86	6	5	1
Comunidad Valenciana	5108,89	4101,83	5	6	1
Murcia	845,63	775,10	8	7	1
Galicia	762,34	682,79	10	8	4
Castilla-León	982,36	680,69	7	9	4
País Vasco	803,63	498,04	9	10	1
Aragón	352,16	365,91	11	11	0
Cantabria	315,03	323,24	12	12	0
Castilla La Mancha	187,19	216,89	14	13	1
Asturias	167,96	212,64	16	14	4
Navarra	197,80	142,98	13	15	4
Extremadura	186,76	118,48	15	16	1
Rioja (La)	56,03	43,33	17	17	0
N=17				suma Σ	24

Cálculo del Coeficiente de correlación ordinal de Spearman:

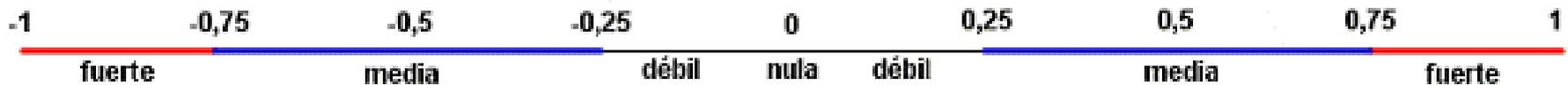
$$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N^3 - N} = 1 - \frac{6 \times 24}{17^3 - 17} = 1 - \frac{144}{4896} = 0,971$$

$$-1 \leq r_s \leq 1$$

Interpretación del coeficiente de correlación ordinal r_s :

- Si $r_s = 1$ → hay concordancia perfecta entre las dos ordenaciones (ambas coinciden para todos los individuos).
- Si $r_s = -1$ → hay discordancia perfecta entre las dos ordenaciones (una ordenación es exactamente la contraria de la otra).
- Si $r_s = 0$ → ausencia de correlación o concordancia entre las dos ordenaciones.
- Si $0 < r_s < 1$ → hay concordancia entre las dos ordenaciones (en el mismo sentido).
- Si $-1 < r_s < 0$ → hay discordancia entre las dos ordenaciones (en sentido contrario).

Interpretación valor r_s



**Discordancia
ordenaciones**

**Concordancia
ordenaciones**

Ejercicio de correlación ordinal:

- A partir de la siguiente información para 2009 para las N=17 Comunidades Autónomas sobre:
- Turistas internacionales según comunidad autónoma de destino principal, en miles de personas.
 - Gasto medio por persona de los turistas según comunidad autónoma de destino principal, en €.

CCAA	Turistas	Gasto medio por persona
Andalucía	7359,47	996,8
Aragón	352,16	1038,7
Asturias	167,96	1264,4
Baleares	9037,39	972,5
Canarias	8203,56	1107,0
Cantabria	315,03	1025,3
Castilla La Mancha	187,19	1170,7
Castilla-León	982,36	700,6
Cataluña	12769,13	760,3
Comunidad Valenciana	5108,89	803,6
Extremadura	186,76	628,3
Galicia	762,34	898,8
Madrid	4895,78	1076,1
Murcia	845,63	915,7
Navarra	197,80	722,7
País Vasco	803,63	624,5
Rioja (La)	56,03	800,9

Fuente: IET. Frontur y Egatur 2009.

- Obtener que el coeficiente de correlación ordinal entre las ordenaciones de las CCAA respecto a los Turistas internacionales y respecto al Gasto medio por persona es igual a -0,012.
- Interpretar el valor de dicho coeficiente.

Más información sobre este tema en:

- PARRA, E; CALERO, F.J.: Estadística para Turismo. Ed. McGraw-Hill, Madrid, 2007. Capítulo 6.
- RONQUILLO, A: Estadística Aplicada al Sector Turístico, Ed Ramón Areces, Madrid, 1997. Capítulo 9.