

ESTADÍSTICA



GRAU TURISME

TEMA 4: ANÀLISI QUALITATIVA DEL TURISME

Prof. Rosario Martínez Verdú



TEMA 4: ANÀLISI QUALITATIVA DEL TURISME

1. Associació entre variables nominals
2. Correlació entre variables ordinals

1. ASSOCIACIÓ ENTRE VARIABLES NOMINALS

S'han classificat per **zona turística de destinació** i per **procedència** 74.127 turistes estrangers que van arribar l'agost de 2009 a la Comunitat Valenciana. I s'ha elaborat la taula de contingència següent:

Zona turística	Procedència	França	Regne Unit
Costa Blanca (Alacant)		12136	52111
Costa Tarongers (Castelló)		7081	1464
Costa València (València)		1085	250

Font: Enquesta d'ocupació hotelera 2009, INE.

Objectiu: Calcular coeficients que mesuren el grau d'associació existent entre les variables nominals o atributs: **zona turística de destinació** i **procedència**. Per a això definim els següents tipus de freqüències:

- Freqüències observades conjuntes: O_{ij}
- Freqüències marginals: n_i y n_j → Si les calculem, la taula quedaria així:

O_{ij}	Zona turística	Procedència	França	Regne Unit	Totals $n_{i.}$
	Costa Blanca (Alacant)		12136	52111	64247
	Costa Tarongers (Castelló)		7081	1.464	8545
	Costa València (València)		1085	250	1335
	Totals	$n_{.j}$	20302	53825	74127=N

- Freqüències esperades o teòriques: E_{ij} serien les freqüències conjuntes que obtindríem en el cas hipotètic que no hi haguera cap relació entre les variables, és a dir, que les variables foren **INDEPENDENTS**. Aquestes freqüències esperades es calculen com: $\frac{n_{i.} \times n_{.j}}{N}$ La taula quedaria així:

E_{ij}	Zona turística	Procedència	França	Regne Unit
	Costa Blanca (Alacant)		17596,1	46650,9
	Costa Tarongers Castelló		2340,3	6204,7
	Costa València (València)		365,6	969,4

Mesures d'associació

1r. Calculem l'estadístic χ^2 (khi quadrat) que compara per a cada cel·la de la taula de contingència les freqüències observades amb les esperades. Es defineix:

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

Interpretació de l'estadístic χ^2 :

- Si les variables foren independents, aleshores $O_{ij} = E_{ij}$ i l'estadístic $\chi^2 = 0$
- Si l'estadístic $\chi^2 \neq 0$ existeix algun tipus d'associació entre les variables.

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

Inconvenient de χ^2 : el seu límit inferior és zero, però no en té de superior.

L'estadístic χ^2 es pot calcular amb la Caest  :



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-Whitney
 - Prueba de Kruskal-Wallis, tres muestras
 - Prueba de Kruskal-Wallis, cuatro muestras
 - Test de signos

proyecto **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						
	<input type="text" value="3"/>						
	columnas						
	<input type="text" value="2"/>						
	<input type="button" value="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"/>	
<input type="button" value="calcular estadístico"/>		Estadístico <input type="text" value="17507.61"/>					

Coeficient de contingència C de Pearson

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

Interpretació del coeficient de contingència C:

- Si les variables foren independents, aleshores $C=0$.
- Com més prop de 0, menys grau d'associació entre les variables.
- Com major siga el valor de C, més grau d'associació hi haurà entre les variables.

Inconvenient de C: no assoleix mai el valor 1 en el cas d'associació perfecta.

Coeficient 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 és el mínim entre el nombre de files (=3) i el nombre de columnes (=2)

- El coeficient V s'interpreta igual que el coeficient de contingència C, amb l'avantatge que pot assolir el valor 1 en el cas d'associació perfecta entre les variables (major grau d'associació possible).
- Si les variables foren independents, aleshores $V=0$ (menor grau d'associació possible).

2. CORRELACIÓ ENTRE VARIABLES ORDINALS

Es disposa d'informació de 2009 de les N=17 comunitats autònomes sobre:

- Turistes internacionals segons comunitat autònoma de destinació principal, en milers de persones.
- Despesa total dels turistes segons comunitat autònoma de destinació principal, en milions d'euros.

CA	Turistes	Despesa total
Andalusia	7359,47	7337,69
Aragó	352,16	365,91
Astúries	167,96	212,64
Balears	9037,39	8790,67
Canàries	8203,56	9082,28
Cantàbria	315,03	323,24
Castella-La Manxa	187,19	216,89
Castella i Lleó	982,36	680,69
Catalunya	12769,13	9643,02
Comunitat Valenciana	5108,89	4101,83
Extremadura	186,76	118,48
Galícia	762,34	682,79
Madrid	4895,78	5226,86
Múrcia	845,63	775,10
Navarra	197,80	142,98
País Basc	803,63	498,04
Rioja (La)	56,03	43,33

Font: IET. *Moviments turístics en fronteres* (Frontur) i *Enquesta de despesa turística* (Egatur) 2009.

Si ordenem les CA en ordre descendent:
•1r. respecte als turistes internacionals.



S'obtenen les ordenacions de CA
següents, que en diem **rangs**:

CA	Turistes	Despesa total	Rangs turistes
Catalunya	12769,13	9643,02	1
Balears	9037,39	8790,67	2
Canàries	8203,56	9082,28	3
Andalusia	7359,47	7337,69	4
Comunitat Valenciana	5108,89	4101,83	5
Madrid	4895,78	5226,86	6
Castella i Lleó	982,36	680,69	7
Múrcia	845,63	775,10	8
País Basc	803,63	498,04	9
Galícia	762,34	682,79	10
Aragó	352,16	365,91	11
Cantàbria	315,03	323,24	12
Navarra	197,80	142,98	13
Castella-La Manxa	187,19	216,89	14
Extremadura	186,76	118,48	15
Astúries	167,96	212,64	16
Rioja (La)	56,03	43,33	17

Font: IET. Frontur i Egatur 2009.

Si ordenem les CA en ordre descendent:
•2n respecte a la despesa total dels turistes.



S'obtenen les ordenacions de CA
següents, que en diem **rangs**:

CA	Turistes	Despesa total	Rangs turistes	Rangs despesa total
Catalunya	12769,13	9643,02	1	1
Canàries	8203,56	9082,28	3	2
Balears	9037,39	8790,67	2	3
Andalusia	7359,47	7337,69	4	4
Madrid	4895,78	5226,86	6	5
Comunitat Valenciana	5108,89	4101,83	5	6
Múrcia	845,63	775,10	8	7
Galícia	762,34	682,79	10	8
Castella i Lleó	982,36	680,69	7	9
País Basc	803,63	498,04	9	10
Aragó	352,16	365,91	11	11
Cantàbria	315,03	323,24	12	12
Castella-La Manxa	187,19	216,89	14	13
Astúries	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

Els dos rangs que hem obtingut són les variables ordinals o ordenacions.

Objectiu: Calcular un coeficient que mesure si hi ha o no relació (correlació o concordança) entre els dos criteris d'ordenació de les 17 CA (rangs turistes internacionals i rangs despesa total).

Coeficient de correlació ordinal o per rangs de Spearman

Per calcular aquest coeficient, en primer lloc es calculen per a cada CA les diferències entre els rangs obtinguts en cada ordenació. A més, aquestes diferències s'eleven al quadrat $\rightarrow d_i^2$ i se sumen.

CA	Turistes	Despesa total	Rangs turistes	Rangs despesa total	d_i^2
Catalunya	12769,13	9643,02	1	1	0
Canàries	8203,56	9082,28	3	2	1
Balears	9037,39	8790,67	2	3	1
Andalusia	7359,47	7337,69	4	4	0
Madrid	4895,78	5226,86	6	5	1
Comunitat Valenciana	5108,89	4101,83	5	6	1
Múrcia	845,63	775,10	8	7	1
Galícia	762,34	682,79	10	8	4
Castella i Lleó	982,36	680,69	7	9	4
País Basc	803,63	498,04	9	10	1
Aragó	352,16	365,91	11	11	0
Cantàbria	315,03	323,24	12	12	0
Castella-La Manxa	187,19	216,89	14	13	1
Astúries	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



Càlcul del coeficient de correlació ordinal de Spearman:

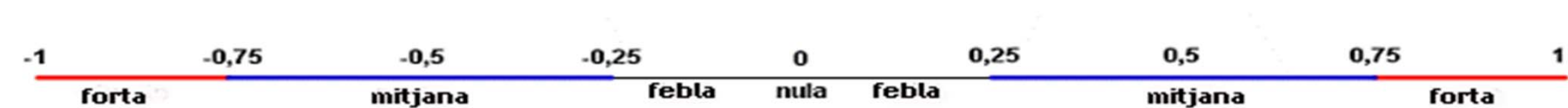
$$r_s = 1 - \frac{6 \sum_{i=1}^N d_i^2}{N^3 - N} = 1 - \frac{6 \times 24}{17^3 - 17} = 1 - \frac{144}{4896} = 0,971$$

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

Interpretació del coeficient de correlació ordinal r_s :

- Si $r_s = 1$ → hi ha concordança perfecta entre les dues ordenacions (ambdues coincideixen per a tots els individus).
- Si $r_s = -1$ → hi ha discordança perfecta entre les dues ordenacions (una ordenació és exactament la contrària de l'altra).
- Si $r_s = 0$ → absència de correlació o concordança entre les dues ordenacions.
- Si $0 < r_s < 1$ → hi ha concordança entre les dues ordenacions (en el mateix sentit).
- Si $-1 < r_s < 0$ → hi ha discordança entre les dues ordenacions (en sentit contrari).

Interpretació valor r_s



**Discordança
ordenacions**

**Concordança
ordenacions**

Exercici de correlació ordinal:

A partir de la informació següent de 2009 de les N=17 comunitats autònomes sobre:

- Turistes internacionals segons comunitat autònoma de destinació principal, en milers de persones.
- Despesa mitjana per persona dels turistes segons comunitat autònoma de destinació principal, en €

CA	Turistes	Despesa mitjana per persona
Andalusia	7359,47	996,8
Aragó	352,16	1038,7
Astúries	167,96	1264,4
Balears	9037,39	972,5
Canàries	8203,56	1107,0
Cantàbria	315,03	1025,3
Castella-La Manxa	187,19	1170,7
Castella i Lleó	982,36	700,6
Catalunya	12769,13	760,3
Comunitat Valenciana	5108,89	803,6
Extremadura	186,76	628,3
Galícia	762,34	898,8
Madrid	4895,78	1076,1
Múrcia	845,63	915,7
Navarra	197,80	722,7
País Basc	803,63	624,5
Rioja (La)	56,03	800,9

Font: IET. Frontur i Egatur 2009.

- Cal obtenir que el coeficient de correlació ordinal entre les ordenacions de les CA respecte dels turistes internacionals i respecte de la despesa mitjana per persona és igual a -0,012.
- Cal interpretar el valor d'aquest coeficient.

Mes informació sobre aquest tema en:

- PARRA, E; CALERO, F. J.: *Estadística para turismo*, Ed. McGraw-Hill, Madrid, 2007. Capítol 6.
- RONQUILLO, A: *Estadística aplicada al sector turístico*, Ed. Ramón Areces, Madrid, 1997. Capítol 9.