

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 11: CONTRASTES NO PARAMÉTRICOS

- 11.1.- Contraste de rachas. por intervalos
- 11.2.- Contraste de normalidad.
- 11.3.- Contraste de independencia.



TEMA 11. COMPETENCIAS

Una vez superado el tema los alumnos serán capaces de:

- Resolver contrastes relativos a la aleatoriedad de la muestra.
- Aplicar contrastes de normalidad.
- Contrastar la independencia entre dos características



MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 11: CONTRASTES NO PARAMÉTRICOS

11.1.- Contraste de rachas.



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Contraste de rachas

H_0 : Los datos constituyen una muestra aleatoria

Resolución del contraste

- Se clasifican los elementos de la muestra en dos categorías, a las que se asocia una variable dicotómica
- Si la variable X estudiada es cuantitativa, la variable dicotómica se define como: 0, si $x < Me$ y 1, si $x > Me$
- Se calcula el número de rachas R
Una racha es una secuencia de observaciones consecutivas de la misma categoría
- Para n elevado se tiene la discrepancia:

$$d_R = \frac{R - 2np(1-p)}{2\sqrt{np(1-p)}} \rightarrow N(0,1) \quad \leftarrow \begin{matrix} p = \text{Proporción observada de '1'} \end{matrix}$$

La regla de decisión se basa en el nº de rachas observado (r^*), y conduce a rechazar la hipótesis de aleatoriedad si r^* es muy grande o muy pequeño (contraste bilateral)



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Ilustración

H_0 : Los datos constituyen una muestra aleatoria

Ejemplo

Muestra observada:

X	2	1	2,5	3	1,5	3	(Me = 2,25)
V. Dicotómica	0	0	1	1	0	1	

nº rachas: $r^* = 4$

$$d_R^* = \frac{r^* - 2np(1-p)}{2\sqrt{np(1-p)}} = \frac{4 - 2(6)(0,5)(0,5)}{2\sqrt{6(0,5)(0,5)}} = 0,89$$

$$p = P(N(0,1) > 0,89) = 0,37$$

No se rechaza H_0



MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 11: CONTRASTES NO PARAMÉTRICOS

11.2.- Contraste de normalidad.



Test de Jarque y Bera

Hipótesis: Normalidad

Se compara la forma (asimetría y apuntamiento) de la distribución muestral con la de una normal

	Asimetría	Apuntamiento
	$g_1 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{S^3}$	$g_2 = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{S^4} - 3$
$g=0$	Distribución simétrica	Distribución mesocúrtica
$g>0$	Distribución con asimetría positiva	Distribución leptocúrtica
$g<0$	Distribución con asimetría negativa	Distribución platicúrtica

Discrepancia

$$d_{JB} = \frac{n}{6} \left(g_1^2 + \frac{1}{4} g_2^2 \right) \quad \text{Bajo } H_0, d_{JB} \approx \chi^2_2$$

Nivel crítico

$$p = P(d_{JB} > d_{JB}^* / H_0)$$

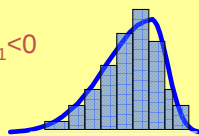


Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Test de Jarque y Bera

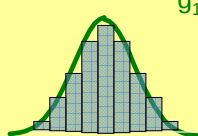
Asimetría negativa

$$g_1 < 0$$



Simétrica

$$g_1 = 0$$



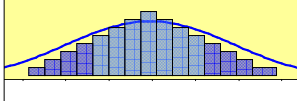
Asimetría positiva

$$g_1 > 0$$



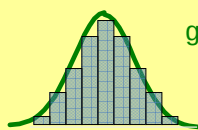
Platicúrtica

$$g_2 < 0$$



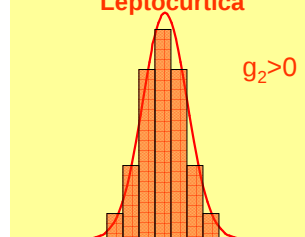
Mesocúrtica

$$g_2 = 0$$



Leptocúrtica

$$g_2 > 0$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

MÉTODOS ESTADÍSTICOS PARA LA EMPRESA

TEMA 10: CONTRASTES NO PARAMÉTRICOS

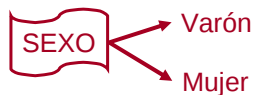
11.3.- Contraste de independencia.



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Independencia entre dos poblaciones

→ X variable cualitativa con r categorías $A_1, A_2, \dots, A_r$



→ Y variable cualitativa con s categorías $B_1, B_2, \dots, B_s$



¿Existe relación entre X (SEXO) e Y (HÁBITO)?

Frecuencias conjuntas observadas

X	A_1	A_2	...	A_r
Y				
B_1	n_{11}	n_{21}	...	n_{r1}
B_2	n_{12}	n_{22}	...	n_{r2}
...	...	...	...	...
B_s	n_{1s}	n_{2s}	...	n_{rs}

$$n_{i.} = \sum_{j=1}^s n_{ij}$$

$$n_{.j} = \sum_{i=1}^r n_{ij}$$

$$\sum_{i=1}^r \sum_{j=1}^s n_{ij} = n$$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Contraste de independencia χ^2

H_0 : X e Y son v.a. independientes

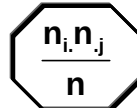
→ X e Y tienen r y s categorías, respectivamente

→ Se comparan las frecuencias conjuntas observadas y teóricas

Frecuencias observadas



Frecuencias teóricas



$$d_{IND} = \sum_{i=1}^r \sum_{j=1}^s \frac{\left(n_{ij} - \frac{n_i \cdot n_j}{n} \right)^2}{\frac{n_i \cdot n_j}{n}} \rightarrow \chi^2_{(r-1)(s-1)}$$

Aproximación $\frac{n_i \cdot n_j}{n} \geq 5$

Nivel crítico → $p = P(d_{IND} > d_{IND}^* / H_0)$



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa

Test de independencia. Ilustración

H_0 : El sexo y el hábito son independientes

Frecuencias observadas

SEXO	Varón	Mujer	n_j
HÁBITO			
Fumador	20	19	39
No fumador	10	16	26
n_i	30	35	65

Discrepancia

$$d_{IND}^* = \sum_{i=1}^2 \sum_{j=1}^2 \frac{\left(n_{ij} - \frac{n_i \cdot n_j}{n} \right)^2}{\frac{n_i \cdot n_j}{n}} = 1,03$$

Nivel crítico

$$p = P(\chi_1^2 > 1,03) = 0,31$$

Conclusión

No se rechaza H_0

Frecuencias marginales

Frecuencias teóricas

SEXO	Varón	Mujer
HÁBITO		
Fumador	18	21
No fumador	12	14



Universidad de Oviedo. Facultad de Economía y Empresa. Grado en ADE. Métodos Estadísticos para la Empresa