



# TABLAS DE CONTINGENCIA

M<sup>a</sup> Purificación VICENTE GALINDO

[purivic@yahoo.com](mailto:purivic@yahoo.com)

Departamento de Estadística.

Universidad de Salamanca. España

# Análisis de la relación entre dos variables

• Cuantitativas

• Cualitativas

Ho: Las 2 variables son independientes  
Ha: Las variables están relacionadas

Coef. Correlación Pearson

Test Chi-cuadrado



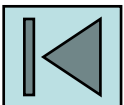
**p-valor < 0.05 => Rechazo Ho**

- **HIPOTESIS DE PARTIDA:  $H_0$**   
**Las dos variables en estudio  
son independientes**
- **HIPOTESIS ALTERNATIVA:  $H_a$**   
**Las dos variables en estudio  
están relacionadas**

# Un ejemplo:

## Tabla de frecuencias observadas

	<i>Peor</i>	<i>Igual</i>	<i>Mejor</i>	
<i>Trat1</i>	7	28	115	
<i>Trat2</i>	15	20	85	
<i>Trat3</i>	10	30	90	
<i>Trat4</i>	5	40	115	
				560



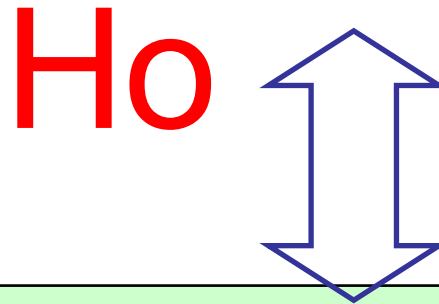
## Frecuencias marginales

	<i>Peor</i>	<i>Igual</i>	<i>Mejor</i>	TOTAL
<i>Trat1</i>	<b>7</b>	<b>28</b>	<b>115</b>	150
<i>Trat2</i>	<b>15</b>	<b>20</b>	<b>85</b>	120
<i>Trat3</i>	<b>10</b>	<b>30</b>	<b>90</b>	130
<i>Trat4</i>	<b>5</b>	<b>40</b>	<b>115</b>	160
TOTAL	37	118	405	<b>560</b>



# ¿Cómo se contrasta?

- Partimos de una tabla de frecuencias observadas
- Se calculan las frecuencias que cabría esperar si las dos variables fueran independientes



$$fe_{ij} = (\text{Total fila } i\text{-ésima}) (\text{Total columna } j\text{-ésima}) / \text{Total global}$$

## Cálculo de la frecuencia esperada $fe_{11}$

	<i>Peor</i>	<i>Igual</i>	<i>Mejor</i>	TOTAL
<i>Trat1</i>	<b>7</b>			<b>150</b>
<i>Trat2</i>				
<i>Trat3</i>				
<i>Trat4</i>				
TOTAL	<b>37</b>			<b>560</b>

$$f_{o_{11}} = 7 \quad fe_{11} = (150 \times 37) / 560 = 9.91$$



# TABLA DE FRECUENCIAS ESPERADAS

	Peor	Igual	Mejor	TOTAL
Trat 1	9,91	31,61	108,48	150
Trat 2	7,93	25,28	86,79	120
Trat 3	8,59	27,39	94,02	130
Trat 4	10,57	33,72	115,71	160
TOTAL	37	118	405	560





# CÓMO MEDIR LAS DISCREPANCIAS

Se calcula la diferencia entre ambas magnitudes ( $\mathbf{fo}_{ij} - \mathbf{fe}_{ij}$ ), para todas y cada una de las casillas de la tabla

## Estadígrafo de contraste

$$\chi^2 = \sum_i \frac{(\mathbf{fo}_{ij} - \mathbf{fe}_{ij})^2}{\mathbf{fe}_{ij}}$$

$\mathbf{fo}_{ij}$  = frecuencia observada para la ij-ésima casilla.

$\mathbf{fe}_{ij}$  = frecuencia esperada para la ij-ésima casilla.

Si la Hipótesis nula es cierta,  $\chi^2$  sigue una distribución Chi cuadrado con  $(I-1)(J-1)$  grados de libertad

• **Rechazaremos  $H_0$  cuando  $\chi^2$  experimental  $>$   $\chi^2$  crítico**

## CÁLCULO DEL VALOR EXPERIMENTAL

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(fo_{ij} - fe_{ij})^2}{fe_{ij}}$$

➤ *Aplicado a nuestro ejemplo el resultado sería:*

$$\chi^2_{\text{exp}} = \frac{(7 - 9.91)^2}{9.91} + \dots + \frac{(115 - 115.71)^2}{115.71} = 13.87$$

## TABLA DE LA JI-CUADRADO

	0.9950	0.9750	0.950	0.900	0.200	0.10	0.050	0.025	0.010	0.001
1	0.0000393	0.000982	0.00393	0.0158	1.642	2.706	3.841	5.024	6.635	10.828
2	0.010	0.0506	0.103	0.211	3.219	4.605	5.991	7.378	9.510	13.816
3	0.0717	0.216	0.352	0.584	4.642	6.251	7.879	9.348	11.345	16.266
4	0.207	0.484	0.711	1.064	5.989	7.779	9.488	11.143	13.277	18.467
5	0.412	0.831	1.145	1.610	7.289	9.236	11.070	12.833	15.086	20.515
6	0.676	1.237	1.635	2.204	8.558	10.645	12.592	14.449	16.812	22.458
7	0.989	1.690	2.167	2.833	9.803	12.017	14.067	16.013	18.475	24.322
8	1.344	2.180	2.733	3.490	11.030	13.362	15.507	17.535	20.090	26.124
9	1.735	2.700	3.325	4.168	12.242	14.684	16.919	19.023	21.666	27.877
10	2.156	3.247	3.940	4.865	13.442	15.987	18.307	20.483	23.209	29.588
11	2.603	3.816	4.575	5.578	14.631	17.275	19.675	21.920	24.725	31.264
12	3.074	4.404	5.226	6.304	15.812	18.549	21.026	23.337	26.217	32.909
13	3.565	5.009	5.892	7.042	16.985	19.812	22.362	24.736	27.688	34.528
14	4.075	5.629	6.571	7.790	18.151	21.064	23.685	26.119	29.141	36.123
15	4.601	6.262	7.261	8.547	19.311	22.307	24.996	27.488	30.578	37.697
16	5.142	6.908	7.962	9.312	20.465	23.452	26.296	28.645	32.000	39.252
17	5.697	7.564	8.672	10.085	21.615	24.769	27.587	30.191	33.409	40.790
18	6.265	8.231	9.390	10.865	22.760	25.989	28.869	31.526	34.805	42.312
19	6.844	8.907	10.117	11.651	23.900	27.204	30.144	32.852	36.191	43.820
20	7.434	9.591	10.851	12.443	25.038	28.412	31.410	34.170	37.566	45.315
21	8.034	10.283	11.591	13.240	26.171	29.615	32.671	35.479	38.932	46.979
22	8.643	10.982	12.338	14.041	27.301	30.813	33.924	36.781	40.289	48.268
23	9.260	11.689	13.091	14.848	28.429	32.007	35.172	38.076	41.638	49.728
24	9.886	12.401	13.848	15.659	29.553	33.196	36.415	39.364	42.980	51.179
25	10.520	13.120	14.611	16.473	30.675	34.382	37.652	40.646	44.314	52.620
26	11.160	13.844	15.379	17.292	31.795	35.563	38.885	41.923	45.642	54.052
27	11.808	14.573	16.151	18.114	32.912	36.741	40.113	43.195	46.963	55.476
28	12.461	15.308	16.928	18.939	34.026	37.916	41.337	44.461	48.278	56.892
29	13.121	16.047	17.708	19.769	35.139	39.087	42.557	45.722	49.588	58.301
30	13.787	16.791	18.493	20.599	36.250	40.256	43.773	46.979	50.892	59.703

## CÁLCULO DEL VALOR EXPERIMENTAL

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(fo_{ij} - fe_{ij})^2}{fe_{ij}}$$

➤ *Aplicado a nuestro ejemplo el resultado sería:*

$$\chi^2_{\text{exp}} = \frac{(7 - 9.91)^2}{9.91} + \dots + \frac{(115 - 115.71)^2}{115.71} = 13.87$$

$$\chi^2_{0.05,6} = 12.59 \quad 13.87 > 12.59$$

**=> La respuesta depende del tratamiento**

# Y ahora con el SPSS...



# ¿Cómo meter los datos?

**Los datos pueden estar dispuestos:**

- **En su formato bruto (tantas columnas como variables categóricas deseemos analizar)**

**Como tabla de contingencia ya construida.**

# ¿Cómo meter los datos?

## Datos

- Como tabla de contingencia ya construida.

	Peor	Igual	Mejor	
Trat1	7	28	115	
Trat2	15	20	85	
Trat3	10	30	90	
Trat4	5	40	115	
				560

TABLA.sav - SPSS Editor de datos

Archivo Edición Ver Datos Transformar Analizar Gráficos Utilidades Ventana

15 : DIAGNÓSTICO

	DOSIS	DIAGNÓSTICO	FRECUENCIAS
1	TRATAM_1	1_PEOR	7
2	TRATAM_1	2_IGUAL	28
3	TRATAM_1	3_MEJOR	115
4	TRATAM_2	1_PEOR	15
5	TRATAM_2	2_IGUAL	20
6	TRATAM_2	3_MEJOR	85
7	TRATAM_3	1_PEOR	10
8	TRATAM_3	2_IGUAL	30
9	TRATAM_3	3_MEJOR	90
10	TRATAM_4	1_PEOR	5
11	TRATAM_4	2_IGUAL	40

Tabla de contingencia DOSIS \* DIAGNÓSTICO

Recuento

		DIAGNÓSTICO			Total
		1 PEOR	2 IGUAL	3 MEJOR	
DOSIS	TRATAM_1	7	28	115	150
	TRATAM_2	15	20	85	120
	TRATAM_3	10	30	90	130
	TRATAM_4	5	40	115	160
Total		37	118	405	560

# ¿Cómo realizar el análisis?

The image displays the SPSS software interface with the 'Data' menu open. A red arrow points from the 'Weight Cases...' option in the 'Data' menu to the 'Ponderar casos' dialog box. The 'Analizar' menu is also open, showing the path: 'Estadísticos descriptivos' > 'Tablas de contingencia...'. The 'Tablas de contingencia' dialog box is shown with 'Frecuencias' selected in the 'Filas' list. The 'Ponderar casos' dialog box is also shown, with 'Ponderar casos mediante' selected and 'frecuencias' in the 'Variable de frecuencia' field. The 'Tratamiento' and 'Diagnostico' buttons are highlighted in the 'Tablas de contingencia' dialog box.

**Data** Transform Analyze Graphs

- Define Variable Properties...
- Copy Data Properties...
- New Custom Attribute...
- Define Dates...
- Define Multiple Response Sets...
- Validation
  - Identify Duplicate Cases...
  - Identify Unusual Cases...
- Sort Cases...
- Sort Variables...
- Transpose...
- Restructure...
- Merge Files
- Aggregate...
- Orthogonal Design
- Copy Dataset
- Split File...
- Select Cases...
- Weight Cases...

**Analizar**

- Informes
- Estadísticos descriptivos
  - Frecuencias...
  - Descriptivos...
  - Explorar...
  - Tablas de contingencia...
  - Razón...
- Tablas
- Comparar medias
- Modelo lineal general
- Modelos mixtos
- Correlaciones
- Regresión
- Loglineal
- Clasificar
- Reducción de datos
- Escalas
- Pruebas no paramétricas
- Serie temporales
- Supervivencia
- Respuesta múltiple

**Tablas de contingencia**

Filas: FRECUENCIAS

Columnas: Tratamiento, Diagnostico

Capa 1 de 1

Anterior Siguiente

☐ Mostrar los gráficos de barras agrupadas

☐ Suprimir tablas

Exactas... Estadísticos... Casillas... Formato...

**Ponderar casos**

artículo años

☐ No ponderar los casos

☒ Ponderar casos mediante

Variable de frecuencia: frecuencias

Estado actual: No ponderar casos

Aceptar Pegar Restablecer Cancelar Ayuda



# ¿Cómo realizar el análisis descriptivo?



Tabla de contingencia DOSIS \* DIAGNÓSTICO

**Casillas**

Recuento		DIAGNÓSTICO			Total
		1 PEOR	2 IGUAL	3 MEJOR	
DOSIS	TRATAM_1	7	28	115	150
	TRATAM_2	15	20	85	120
	TRATAM_3	10	30	90	130
	TRATAM_4	5	40	115	160
Total		37	118	405	560

## Frecuencias o Recuentos

Tablas de contingencia: Mostrar en las casillas

Frecuencias

☒ Observadas

☐ Esperadas

Continuar

Cancelar

Ayuda

Tabla de contingencia DOSIS \* DIAGNÓSTICO

Frecuencia esperada

		Diagnostico			Total
		1 PEOR	2 IGUAL	3 MEJOR	
Tratamiento/Dosis	TRATAM_1	9,9	31,6	108,5	150,0
	TRATAM_2	7,9	25,3	86,8	120,0
	TRATAM_3	8,6	27,4	94,0	130,0
	TRATAM_4	10,6	33,7	115,7	160,0
Total		37,0	118,0	405,0	560,0

# Casillas

## Porcentajes

### Tablas de contingencia: Mostrar en las casillas

Frecuencias

☒ Observadas

☐ Esperadas

Porcentajes

☒ Fila

☒ Columna

☒ Total

Residuos

☐ No tipificados

☐ Tipificados

☐ Tipificados con

Ponderaciones no enteras

☒ Redondear frecuencias de casillas

☐ Redondear

☐ Truncar frecuencias de casillas

☐ Truncar

☐ No efectuar correcciones

Tabla de contingencia DOSIS \* DIAGNÓSTICO

% Tratamiento/Dosis

		DIAGNÓSTICO			Total
		1 PEOR	2 IGUAL	3 MEJOR	
DOSIS	TRATAM_1	4,7%	18,7%	76,7%	100,0%
	TRATAM_2	12,5%	16,7%	70,8%	100,0%
	TRATAM_3	7,7%	23,1%	69,2%	100,0%
	TRATAM_4	3,1%	25,0%	71,9%	100,0%
Total		6,6%	21,1%	72,3%	100,0%

Tabla de contingencia DOSIS \* DIAGNÓSTICO

% Diagnostico

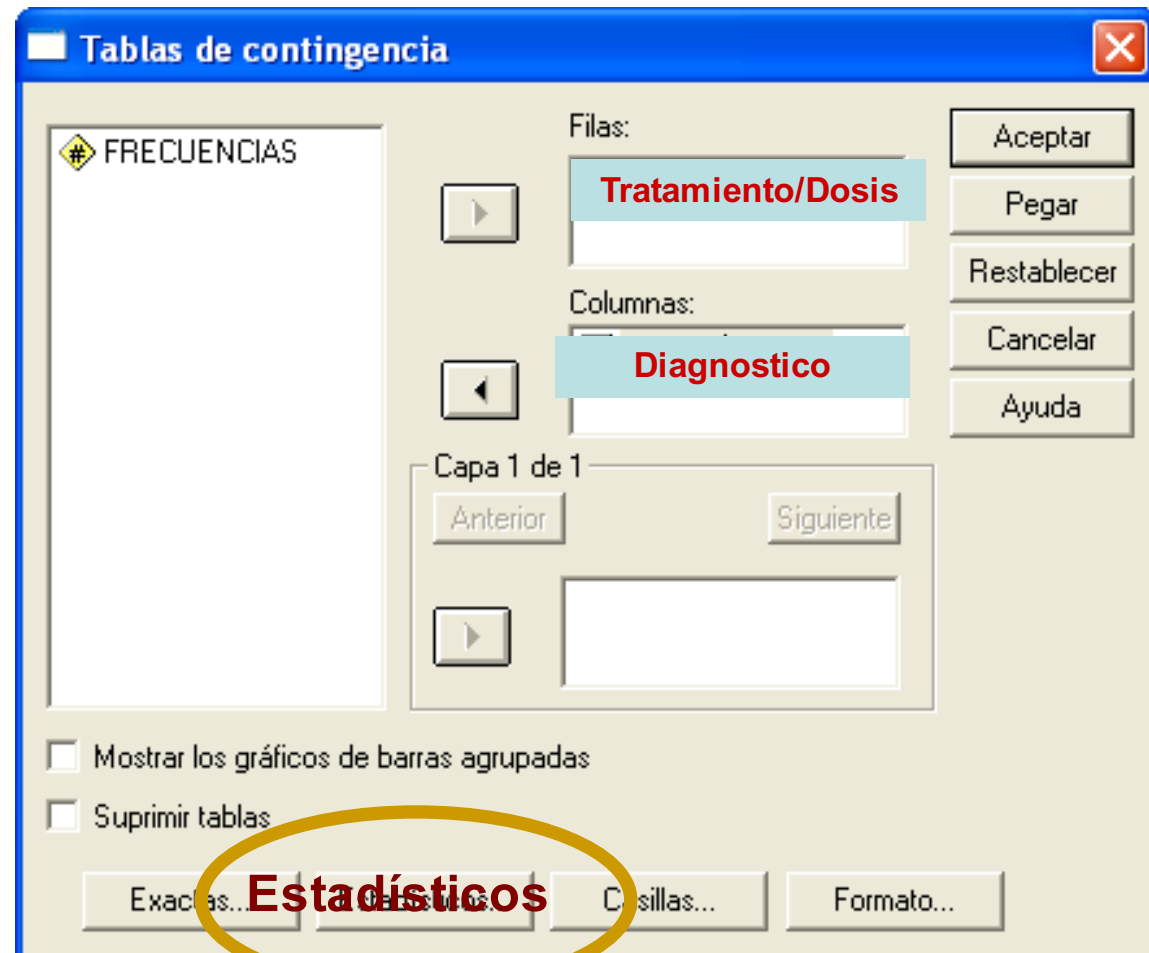
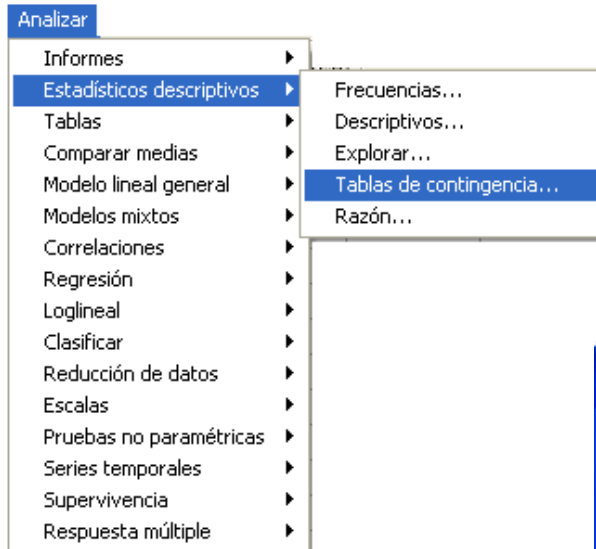
		DIAGNÓSTICO			Total
		1 PEOR	2 IGUAL	3 MEJOR	
DOSIS	TRATAM_1	18,9%	23,7%	28,4%	26,8%
	TRATAM_2	40,5%	16,9%	21,0%	21,4%
	TRATAM_3	27,0%	25,4%	22,2%	23,2%
	TRATAM_4	13,5%	33,9%	28,4%	28,6%
Total		100,0%	100,0%	100,0%	100,0%

Tabla de contingencia DOSIS \* DIAGNÓSTICO

% del total

		DIAGNÓSTICO			Total
		1 PEOR	2 IGUAL	3 MEJOR	
DOSIS	TRATAM_1	1,3%	5,0%	20,5%	26,8%
	TRATAM_2	2,7%	3,6%	15,2%	21,4%
	TRATAM_3	1,8%	5,4%	16,1%	23,2%
	TRATAM_4	,9%	7,1%	20,5%	28,6%
Total		6,6%	21,1%	72,3%	100,0%

# ¿Cómo realizar el análisis?



## Estadísticos

### Tablas de contingencia: Estadísticos

☒ Chi-cuadrado

Nominal

☐ Coeficiente de contingencia

☐ Phi y V de Cramer

☐ Lambda

☐ Coeficiente de incertidumbre

☐ Tau-b de Kendall

☐ Tau-c de Kendall

$$\chi^2_{\text{exp}} = \sum_i \sum_j \frac{(fo_{ij} - fe_{ij})^2}{fe_{ij}}$$

$$G^2 = 2 \sum_{i=1}^I \sum_{j=1}^J fo_{ij} \ln \frac{fo_{ij}}{\hat{fe}_{ij}^{(o)}}$$

☐ Estadísticos de Cochran y de Mantel-Haenszel

Contrastar la razón de ventajas común igual a:

1

Ambos siguen una distribución Chi cuadrado con (I-1) (J-1) grados de libertad

	Peor	Igual	Mejor	
Trat1	7	28	115	
Trat2	15	20	85	
Trat3	10	30	90	
Trat4	5	40	115	
				560

## Resultados

### Tablas de contingencia: Estadísticos

☒ Chi-cuadrado

☐ Correlaciones

Continuar

Cancelar

Ayuda

Nominal

☐ Coeficiente de contingencia

☐ Phi y V de Cramer

☐ Lambda

☐ Coeficiente de incertidumbre

Ordinal

☐ Gamma

☐ d de Somers

☐ Tau-b de Kendall

☐ Tau-c de Kendall

### Pruebas de chi-cuadrado

	Valor	gl	<b>P-valor</b> Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	13,871 <sup>a</sup>	6	,031
Razón de verosimilitud	13,378	6	,037
N de casos válidos	560		

- <sup>a</sup>. 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.  
La frecuencia mínima esperada es 7,93.

## Pruebas de chi-cuadrado

### P-valor

	Valor	gl	Sig. asintótica (bilateral)
Chi-cuadrado de Pearson	13,871 <sup>a</sup>	6	,031
Razón de verosimilitud	13,378	6	,037
N de casos válidos	560		

a. 0 casillas (,0%) tienen una frecuencia esperada inferior a 5.  
La frecuencia mínima esperada es 7,93.

En el ordenador, al lado del valor experimental, aparece el **p-valor**

**El p-valor** nos indica el riesgo que corremos al rechazar la  $H_0$  (independencia) después de haber explorado los datos.

**Si el p-valor es menor de 0.05**, rechazamos  $H_0$  (independencia) y aceptamos la  $H_a$  (relación). Si  $p\text{-valor} > 0.05$ , NO

# GRADO DE ASOCIACIÓN

## Estadísticos

$$CC = \sqrt{\frac{\chi^2}{\chi^2 + n}}$$

$$0 \leq CC = \sqrt{\frac{I-1}{I}}$$

I= Dimensión de la tabla

Tablas de contingencia: Estadísticos

☒ Chi-cuadrado

☐ Correlaciones

☒ Coeficiente de contingencia

☐ Gamma

☐ Phi y V de Cramer

☐ d de Somers

☐ Lambda

☐ Tau-b de Kendall

☐ Coeficiente de incertidumbre

☐ Tau-c de Kendall

☐ Kappa

☐ Riesgo

☐ McNemar

☐ Estadísticos de Cochran y de Mantel-Haenszel

Continuar

Cancelar

Ayuda

igual a: 1

Coeficiente de Contingencia (CC): (Basado en el Chi-cuadrado)

Nominales

Tabla cuadrada

Rol simétrico



	Peor	Igual	Mejor	
Trat1	7	28	115	
Trat2	15	20	85	
Trat3	10	30	90	
Trat4	5	40	115	
				560

### Medidas simétricas

	Valor	Sig. aproximada
Nominal por nominal N de casos válidos	Coeficiente de contingencia ,155 560	,031

- Asumiendo la hipótesis alternativa.
- Empleando el error típico asintótico basado en la hipótesis nula.

# Estadísticos

$$\phi = \sqrt{\frac{\chi^2}{N}}$$

**Phi :** (Basado en el Chi-cuadrado)

Nominales

Dicotómicas

$$V \text{ Cramer} = \sqrt{\frac{\chi^2}{\min(I-1, J-1)N}}$$

**V de Cramer :** (Basado en el Chi-cuadrado)

Nominales

**Tablas de contingencia: Estadísticos**

☐ Chi-cuadrado

☐ Correlaciones

☐ Nominal

☐ Coeficiente de contingencia

☒ Phi y V de Cramer

☐ Gamma

☐ d de Somers

☐ Tau-b de Kendall

☐ Tau-c de Kendall

☐ Kappa

☐ Riesgo

☐ McNemar

☐ Eta

☐ Estadísticos de Cochran y de Mantel-Haenszel

Contrastar la razón de ventajas común igual a: 1

Continuar

Cancelar

Ayuda

# Casillas

Para ayudarle a descubrir las tramas en los datos que contribuyen a una prueba de chi-cuadrado significativa, cada casilla de la tabla puede contener cualquier combinación de recuentos, porcentajes y residuos seleccionados.

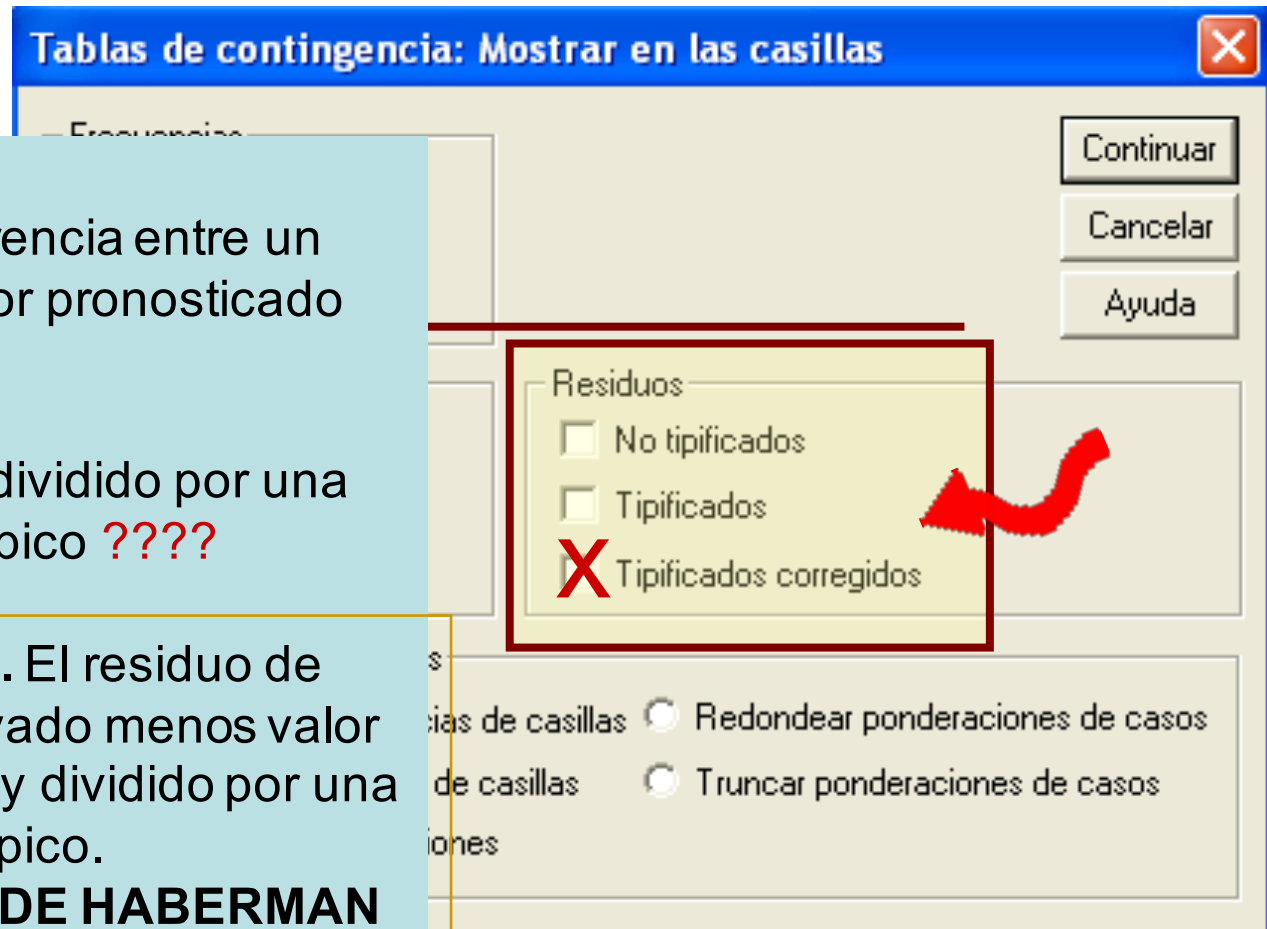
## Residuos.

**No tipificados.** La diferencia entre un valor observado y el valor pronosticado por el modelo.

**Tipificados.** El residuo dividido por una estimación de su error típico ????

**Tipificados corregidos.** El residuo de una casilla (valor observado menos valor pronosticado), centrado y dividido por una estimación de su error típico.

## RESIDUALES DE HABERMAN



Siguen, asintóticamente, un modelo  $N(0,1)$

# Casillas

## Residuos

Tabla de contingencia DOSIS \* DIAGNÓSTICO

	Peor	Igual	Mejor	
Trat1	7	28	115	
Trat2	15	20	85	
Trat3	10	30	90	
Trat4	5	40	115	
				560

Residuo

		DIAGNÓSTICO		
		1 PEOR	2 IGUAL	3 MEJOR
DOSIS	TRATAM_1	-2,9	-3,6	6,5
	TRATAM_2	7,1	-5,3	-1,8
	TRATAM_3	1,4	2,6	-4,0
	TRATAM_4	-5,6	6,3	-,7

Residuos tipificados

		DIAGNÓSTICO		
		1 PEOR	2 IGUAL	3 MEJOR
DOSIS	TRATAM_1	-,9	-,6	,6
	TRATAM_2	2,5	-1,1	-,2
	TRATAM_3	,5	,5	-,4
	TRATAM_4	-1,7	1,1	-,1

Residuos corregidos

		DIAGNÓSTICO		
		1 PEOR	2 IGUAL	3 MEJOR
DOSIS	TRATAM_1	-1,1	-,8	1,4
	TRATAM_2	2,9	-1,3	-,4
	TRATAM_3	,6	,6	-,9
	TRATAM_4	-2,1	1,4	-,1

Mostrar en las casillas

Continuar

Cancelar

Ayuda

Residuos

☐ No tipificados

☐ Tipificados

☒ Tipificados corregidos

Redondear ponderaciones de casos

Truncar ponderaciones de casos

# GRADO DE ASOCIACIÓN

## Estadísticos

### Tablas de contingencia: Estadísticos

Si las dos variables son **ordinales**, calcula el **coeficiente de Spearman** (correlación entre rangos)

Si las dos variables son **cuantitativas**, calcula el **coeficiente de correlación Pearson**

☒ Correlaciones

Ordinal

☐ Gamma

☐ d de Somers

☐ b de Kendall

☐ c de Kendall

☐ Phi y V de Cramer

☐ Eta

☐ Kappa

☐ Riesgo

☐ McNemar

☐ Estadísticos de Cochran y de Mantel-Haenszel

Contrastar la razón de ventajas común igual a:

1

Continuar

Cancelar

Ayuda



# Análisis de la relación entre dos variables

- **Cuantitativa**

- **Cualitativa**

Ho: Las 2 variables son independientes

Ha: Las variables están relacionadas



**T de STUDENT  
ANOVA**

**p-valor < 0.05 => Rechazo Ho**