

# *Introducción*

# Tablas de contingencia

- Cuando tienes variables categóricas o variables ordinales con pocos valores, los análisis suelen pasar por hacer recuentos de frecuencias.
  - Para dos variables esto suele terminar en una tabla de contingencia. Por ejemplo, en los datos de supervivencia del Titanic nos podríamos plantear si los hombres sobrevivieron más que las mujeres (las mujeres y los niños primero).

Gender\_ \* Survive\_ Crosstabulation

Count

		Survive_		Total
		Died	Lived	
Gender_	Fema	126	344	470
	Male	1364	367	1731
Total		1490	711	2201

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- La interpretación de una tabla de este tipo no es muy simple ya que hay que tener en cuenta los totales. Lo más sencillo es calcular porcentajes por filas.

Gender\_ \* Survive\_ Crosstabulation  
% within Gender\_

		Survive_		Total
		Died	Lived	
Gender_ Fema		26,8%	73,2%	100,0%
Male		78,8%	21,2%	100,0%
Total		67,7%	32,3%	100,0%

- Aquí la comparación interesante es con la fila inferior de Total que podemos interpretar como un promedio. En el Titanic vemos que un 67,7% murió pero si lo dividimos por género, la mortalidad entre las mujeres fue de un 26.8% y para los hombres fue de un 78.8%.

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Chi cuadrado

- En este momento puede ser interesante realizar una prueba de hipótesis para probar si efectivamente las diferencias que observamos son estadísticamente significativas. Esto suele hacerse mediante la prueba de Chi cuadrado (SPSS nos regala un montón de cosas de paso)

Chi-Square Tests					
	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	456,874 <sup>a</sup>	1	,000		
Continuity Correction <sup>b</sup>	454,500	1	,000		
Likelihood Ratio	434,469	1	,000		
Fisher's Exact Test				,000	,000
Linear-by-Linear Association	456,667	1	,000		
N of Valid Cases	2201				

a. 0 cells (.0%) have expected count less than 5. The minimum expected count is 151,83.

b. Computed only for a 2x2 table

- Chi cuadrado es una medida de desajuste
- El nivel de significación prueba la hipótesis de que el desajuste es cero
- Rechazar la hipótesis nula significa (en este caso) hombres y mujeres tuvieron unos niveles diferentes de supervivencia

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Residuales estandarizados

- Chi cuadrado está basado en la diferencia entre las puntuaciones esperadas y las observadas para cada una de las celdas de la tabla. La suma de las diferencias produce un valor alto de Chi-cuadrado. Si ese valor es alto rechazamos la hipótesis nula y por tanto pensamos que hay interacción entre las variables

**Gender\_ \* Survive\_ Crosstabulation**

			Survive_		Total
			Died	Lived	
Gender_	Fema	Count	126	344	470
		Expected Count	318,2	151,8	470,0
		Residual	-192,2	192,2	
	Male	Count	1364	367	1731
		Expected Count	1171,8	559,2	1731,0
		Residual	192,2	-192,2	
Total		Count	1490	711	2201
		Expected Count	1490,0	711,0	2201,0

- Las puntuaciones esperadas no son más que el total por fila multiplicado por el total por columna y dividido por el total. Por ejemplo, Mujer | Murió

$$(470 \times 1490) / 2201 = 318.2$$

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Este caso es sencillo de interpretar pero a veces es necesario examinar los residuales para ver de donde provienen las diferencias y en qué medida. Por ejemplo, si analizamos la clase en la que viajaban (¿viajar en primera te da más posibilidades de sobrevivir en caso de accidente?) tenemos la siguiente tabla

Class\_ \* Survive\_ Crosstabulation

			Survive_		Total
			Died	Lived	
Class_	1st	Count	122	203	325
		Expected Count	220,0	105,0	325,0
		Residual	-98,0	98,0	
	2nd	Count	167	118	285
		Expected Count	192,9	92,1	285,0
		Residual	-25,9	25,9	
	3rd	Count	528	178	706
		Expected Count	477,9	228,1	706,0
		Residual	50,1	-50,1	
	Cre	Count	673	212	885
		Expected Count	599,1	285,9	885,0
		Residual	73,9	-73,9	
Total		Count	1490	711	2201
		Expected Count	1490,0	711,0	2201,0

- Parece que en primera y en segunda se sobrevivió más que en tercera pero estos residuales no tienen en cuenta los distintos tamaños. Para valorar mejor los residuales necesitaríamos estandarizarlos

**Introducción**

Tablas de contingencia  
 Chi cuadrado  
 Residuales estandarizados  
 Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
 Modelos para dos variables  
 Modelos para tres variables  
 Notación  
 Modelos para Titanic  
 Modelos para cuatro variables  
 Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
 Ejemplo Berkeley  
 Planteamiento del modelo  
 Output  
 Interpretación de los resultados  
 Examinando los residuales  
 Interpretando los coeficientes  
 Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
 Ejemplo: Fumar y dolor de cabeza  
 Métodos  
 Procedimiento automático  
 Procedimiento manual  
 Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
 Casillas vacías  
 Covariantes a nivel de celda  
 Modelos ordinales  
 Modelos especiales

- Los residuales corregidos ofrecen esto, estos valores pueden interpretarse como puntuaciones z con los márgenes habituales

Class\_ \* Survive\_ Crosstabulation  
Adjusted Residual

		Survive_	
		Died	Lived
Class_	1st	-12,6	12,6
	2nd	-3,5	3,5
	3rd	4,9	-4,9
	Cre	6,9	-6,9

- En este caso todos los valores son muy marcados, destacando enormemente el valor de supervivencia en la primera clase frente a sobre todo la tripulación que sufrió los peores efectos (el capitán es el último en abandonar el barco pero también los marineros al parecer)

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Resumen

- Lo descrito anteriormente es apropiado para situaciones con dos variables
- Cuando tenemos más de dos variables, la situación se hace bastante más compleja rápidamente

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

[Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic

- Modelos para cuatro variables

- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal

- Ejemplo Berkeley

  - Planteamiento del modelo

  - Output

- Interpretación de los resultados

  - Examinando los residuales

  - Interpretando los coeficientes

- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción

- Ejemplo: Fumar y dolor de cabeza

- Métodos

  - Procedimiento automático

  - Procedimiento manual

- Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

- Modelos logit

- Casillas vacías

- Covariantes a nivel de celda

- Modelos ordinales

- Modelos especiales

## Actividades

1. Calcular las tablas de residuales ajustados para Género y Supervivencia en el archivo Titanic3.sav

*Esto se puede hacer con Analizar/Descriptivos/Tablas de Contingencia.*

**Gender\_ \* Survive\_ Crosstabulation**

			Survive_		Total
			Died	Lived	
Gender_	Fema	Count	126	344	470
		Adjusted Residual	-21,4	21,4	
	Male	Count	1364	367	1731
		Adjusted Residual	21,4	-21,4	
Total		Count	1490	711	2201

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

2. Usando el comando Titanic3.sav y el comando Tablas personalizadas en el menu Tablas calcula una tabla para la Supervivencia en el Titanic en función de la Clase y el Género.

*Una regla que suelo recomendar es poner las variables independientes en las filas y las dependientes en columnas y calcular porcentajes por filas. Esa presentación suele ser más fácil de entender (al menos para mí). La tabla que teneis que conseguir es la de abajo.*

				Survive_	
				Died	Lived
				Row N %	Row N %
Class_ 1st	Gender_ Fema		2,8%	97,2%	
		Male	65,6%	34,4%	
2nd	Gender_ Fema		12,3%	87,7%	
		Male	86,0%	14,0%	
3rd	Gender_ Fema		54,1%	45,9%	
		Male	82,7%	17,3%	
Cre	Gender_ Fema		13,0%	87,0%	
		Male	77,7%	22,3%	

## Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

## Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerarquicos o anidados

## Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

## Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

## Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

***Modelos  
loglineales***

# Modelos para dos y tres variables

- La prueba Chi cuadrado puede entenderse como un modelo loglineal para dos variables.
  - No obstante, cuando hay más de dos variables, esta prueba ya no es muy útil
- Los modelos loglineales se hacen mucho más complejos según el número de variables
  - Dos variables es un modelo casi trivial (es material de primer curso)
  - Tres o más la complejidad aumenta bastante

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes

Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Modelos para dos variables

- Empezaremos con la situación anterior para introducir el problema.
  - Para calcular las puntuaciones esperadas hicimos:

$$m_{ij} = \frac{(m_{fila} m_{columna})}{m_t}$$

- Por ejemplo, para calcular el número esperado de mujeres que sobrevivieron

Gender\_ \* Survive\_ Crosstabulation

Count

		Survive_		Total
		Died	Lived	
Gender_	Fema	126	344	<b>470</b>
	Male	1364	367	1731
Total		1490	<b>711</b>	<b>2201</b>

$$m_{ij} = \frac{470 \times 711}{2201} = 151.8$$

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El cálculo anterior no es más que la aplicación de una regla de probabilidad básica (la probabilidad de un suceso compuesto es igual a la probabilidad de los sucesos individuales cuando hay independencia)

$$p_{ij} = \frac{470}{2201} \times \frac{711}{2201} = 0.0689; 0.0689 \times 2201 = 151.8$$

- Para valorar si hay independencia entonces tenemos que comparar el resultado esperado con el observado, si son iguales, entonces los sucesos individuales son independientes, sino son dependientes.
- En nuestro caso sobrevivieron 344 mujeres así que el valor esperado se queda corto

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Lo anterior puede expresarse como logaritmos lo cual presenta muchas ventajas de cálculo

$$\log(m_{ij}) = \log(m_{fila}) + \log(m_{columna}) - \log(m_{Total})$$

- En nuestro caso,

$$5.02 = 6.15 + 6.56 - 7.69$$

- Que es lo mismo que hemos obtenido antes al calcular el antilogaritmo

$$\exp(5.02) = 151.8$$

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Hacer los cálculos en logaritmos simplifica mucho el analizar modelos más complejos, así una forma de ver el análisis loglineal es como un ANOVA con el logaritmo de las frecuencias como variable dependiente:

- Lo anterior se suele expresar así

$$\log(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B$$

- Un modelo de no-independencia de las variables tendría que tener en cuenta la interacción entre las variables

$$\log(m_{ij}) = \mu + \lambda_i^A + \lambda_j^B + \lambda_i^A \cdot \lambda_j^B$$

- Este último modelo se denomina saturado o completo porque las puntuaciones esperadas coinciden exactamente con las observadas
- Este planteamiento ayuda a entender la idea de la prueba Chi cuadrado. Esta prueba compara el modelo completo con el modelo de independencia. Si no hay diferencia, entonces las variables son independientes.

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes

Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

## *Modelos para tres variables*

- Para dos variables (A y B), veíamos que sólo hay dos posibles modelos.
  - En el que hay interacción entre las dos variables (completo o saturado)
  - En el que no hay interacción entre las dos variables (de independencia o efectos principales)
- No obstante, con tres variables (A, B y C) hay más posibilidades:
  - Ninguna interacción
  - Interacción entre A y B
  - Interacción entre A y C
  - Interacción entre B y C
  - Interacción entre A y B, B y C
  - Interacción entre A y B, A y C
  - etc.

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- En el ejemplo del Titanic algunos posibles modelos son:
  - El Género influyó en la Supervivencia y la Clase Social también
  - El Género influyo en la Supervivencia pero la clase Social no
  - El Género y la Clase Social influyeron en la Supervivencia pero además hubo interacción, así que ciertos géneros tuvieron ventaja si viajaban en ciertas clases pero no en todas
  - No hubo relación entre Género, Supervivencia y Clase social
  - etc.

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Notación

- Para simplificar, se puede utilizar la siguiente notación
  - $[ABC]$  significa todas las interacciones posibles entre las tres variables. Se trata del modelo completo.
  - $[AB]$  Significa todas las interacciones posibles entre A y B
  - $[AB][BC]$  Significa todas las interacciones posibles entre A y B, y entre B y C, pero no las de A y C
  - $[AB][BC][AC]$  Significa todas las interacciones posibles entre dos variables pero excluyendo las que incluyen tres variables
  - $[A][B][C]$  No hay interacciones, se trata del modelo de efectos principales
  - También es posible plantear modelos sin alguno de los efectos principales  $[A][B]$  por ejemplo, pero no son habituales

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Con tres variables, si comparamos el modelo completo [ABC] con el modelo de independencia [A][B][C] y el resultado es de rechazo, hay tantas fuentes de variación que dejamos fuera que prácticamente no significa nada. De hecho el modelo [A][B][C] es casi imposible que ajuste y la mayoría de las veces hay que considerar interacciones

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

## Modelos para Titanic

- En el ejemplo del Titanic (G=Género, C=Clase, S=Supervivencia)  $[G][C][S]$  no ajusta por lo que hay que considerar las interacciones. Por ejemplo,
  - $[GS][CS]$  significaría que tanto la clase como el género influyen en la supervivencia, pero no hay interacción entre clase y género (no ajusta, en realidad, en primera clase viajaban muchas mujeres)
  - $[GS][CS][GC]$  en este caso tenemos lo anterior y además la interacción entre clase y género.

Significa que según la clase en la que se viajaba hombres y mujeres tuvieron diferentes niveles de supervivencia (no ajusta, las mujeres tuvieron ventaja en primera y en segunda clase pero no en tercera)

- $[GCS]$  incluye la interacción de tercer orden.

Como veremos será el único modelo que ajusta en nuestro caso

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- En el caso del Titanic, el único modelo que ajusta es el saturado o completo.
  - No obstante, en otros casos puede ajustar un modelo más simple, con menos elementos
  - El objetivo de nuestro análisis es conseguir un modelo que sea lo más simple posible pero que ajuste. Una vez obtenido es posible realizar diagnósticos y realizar interpretaciones del resultado

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Modelos para cuatro variables

- El número de interacciones crece mucho
- Veremos algún ejemplo pero dado el tiempo disponible no los trabajaremos en detalle

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes

Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual

Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Modelos jerarquicos o anidados

- En un análisis loglineal, lo habitual es que los modelos sean jerarquicos o anidados.
  - Eso significa que si se incluye en el modelo el término de interacción de un nivel superior, automáticamente se incluyen todos aquellos de nivel inferior en los que estén incluidas las variables
  - Por ejemplo, si incluimos la interacción entre Género, Departamento y Admisión, no podemos excluir la interacción entre Género y Admisión
- Esa limitación puede evitarse utilizando modelos no jerarquicos pero estos modelos son mucho más complicados de interpretar y se suelen evitar (al menos en cursos introductorios)
  - Al final del material hay unas notas sobre estos otros modelos

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# *Admisión en Berkeley*

# Pasos para un análisis loglineal

- Los pasos son muy parecidos a los que vimos para regresión
  - Ajuste del modelo
  - Diagnóstico
  - Interpretación
- No obstante, esto habitualmente no es un proceso lineal sino que a menudo implica una cierta cantidad de prueba y ensayo

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Ejemplo Berkeley

- En los años 70, en la universidad de Berkeley se preocuparon mucho al encontrarse con estos datos

Berkeley Admissions Dataset

		<i>Gender</i>			
		Male		Female	
		<i>Admission</i>			
		Yes	No	Yes	No
Department	A	512	313	89	19
	B	353	207	17	8
	C	120	205	202	391
	D	138	279	131	244
	E	53	138	94	299
	F	22	351	24	317

- Género
- Admisión (número de estudiantes admitidos o rechazados)
- Departamento (que equivale a Facultad en España)

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Lo más preocupante es cuando se hace el análisis de Género y Admisión

**Gender \* Admission Crosstabulation**  
% within Gender

		Admission		Total
		N	Y	
Gender	F	69,6%	30,4%	100,0%
	M	55,5%	44,5%	100,0%
Total		61,2%	38,8%	100,0%

- Casi un 70% de las mujeres son rechazadas, frente a un 55.5% los hombres
- Esto parece evidencia de discriminación sexual lo cual supongo que causó mucho revuelo

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- No obstante, si tenemos en cuenta la Facultad (Department) varía bastante:

				Admission	
				N	Y
				Row N %	Row N %
Gender	F	Department	A	17,6%	82,4%
			B	32,0%	68,0%
			C	65,9%	34,1%
			D	65,1%	34,9%
			E	76,1%	23,9%
			F	93,0%	7,0%
	M	Department	A	37,9%	62,1%
			B	37,0%	63,0%
			C	63,1%	36,9%
			D	66,9%	33,1%
			E	72,3%	27,7%
			F	94,1%	5,9%

- Aquí podemos ver que los porcentajes de admisión para las mujeres son similares o incluso mejores que los de los hombres en algunos departamentos.
- Claramente, hay algo que se nos escapa y que podemos explorar con análisis loglineal

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

## Planteamiento del modelo

- Como hemos visto antes, un primero paso es encontrar un modelo que ajuste
  - Este modelo debería ser el más simple posible, que no incluyera términos que sean innecesarios
  - Generalmente, se empieza con un modelo saturado y se van quitando elementos hasta que el modelo deja de ajustar
- Dos medidas de ajuste son la Chi cuadrado y la Desvianza
  - Ambas son muy parecidas, pero la Desvianza tiene algunas propiedades que la hacen más deseable así que es la que mencionaré más habitualmente.
  - Recordar, valores bajos de desvianza indican ajuste y niveles de significación altos también indican ajuste (la hipótesis nula es que el modelo ajusta por lo que no queremos rechazarla)

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

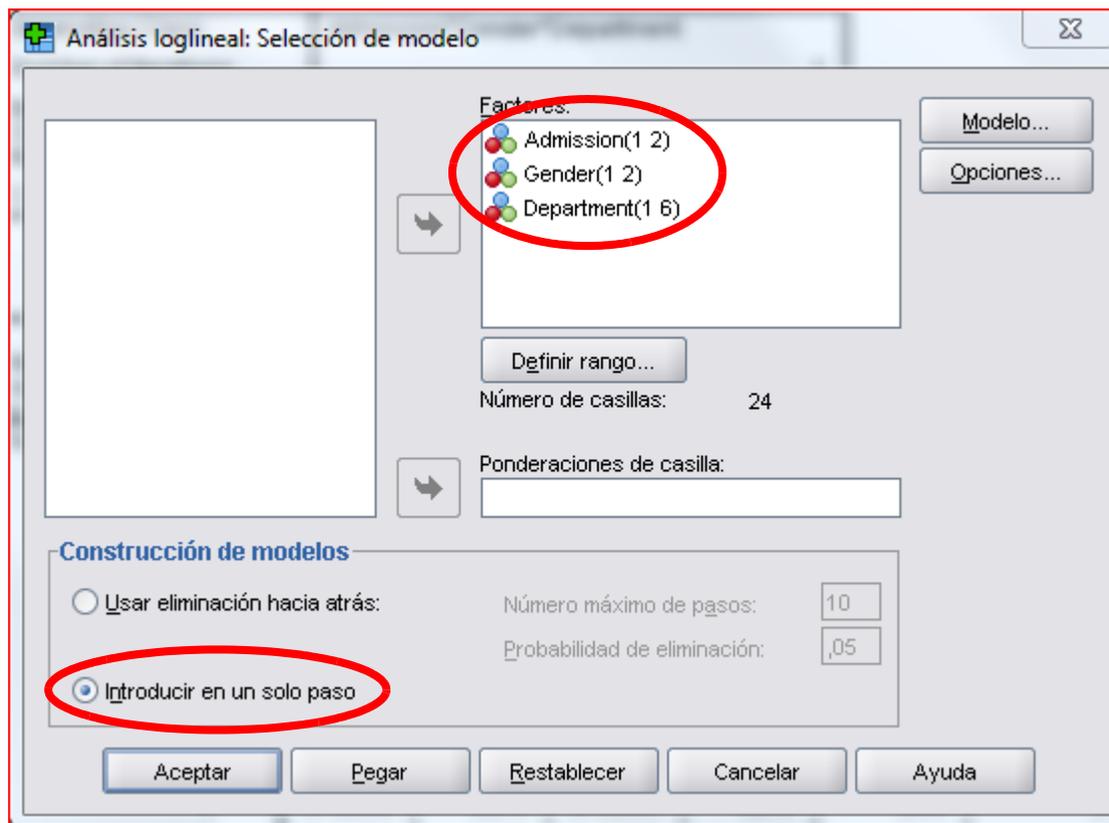
### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- De los módulos disponibles, el más apropiado es el de Selección de Modelo



- En este caso utilizaremos introducir en un solo paso
- Este módulo tiene el inconveniente de que hay que introducir los valores de categorías para las variables

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

**Ajuste de modelos**

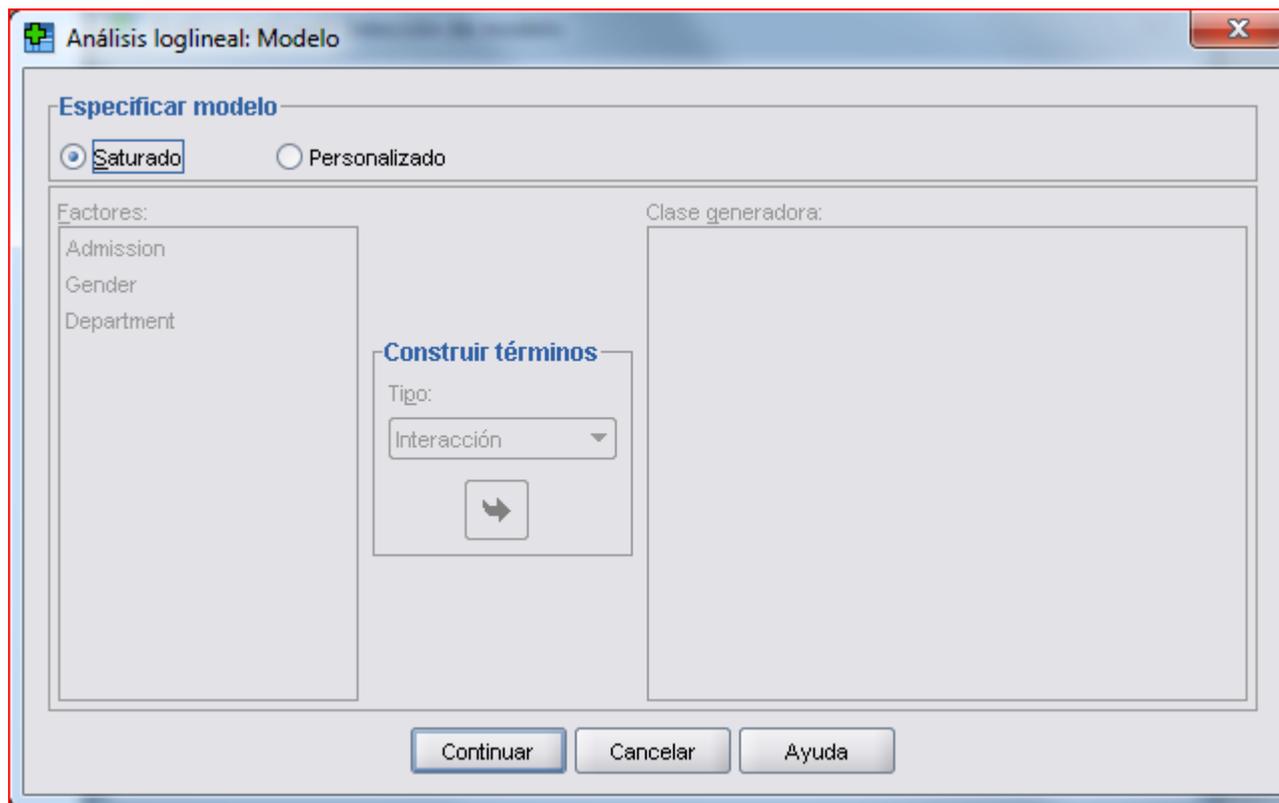
Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El cuadro de diálogo de Modelo lo dejamos así



- Ajustaremos el modelo saturado (el cual siempre ajusta perfectamente)

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

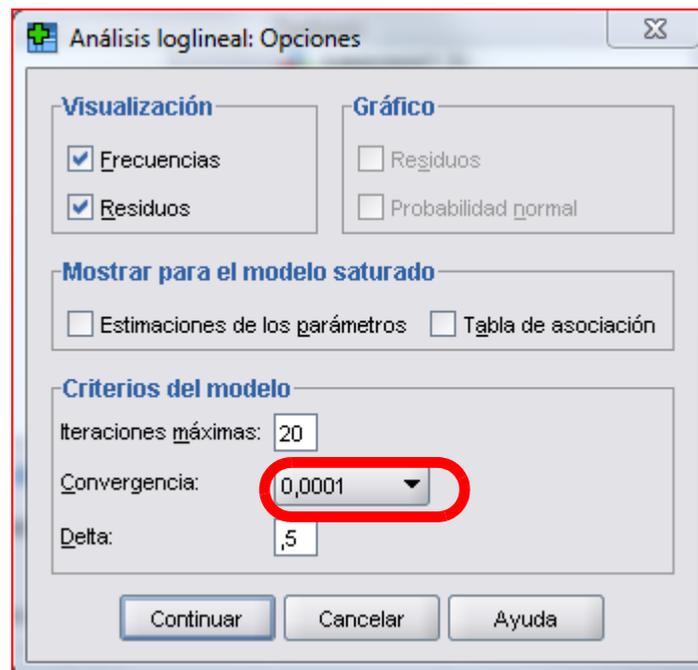
**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El cuadro de diálogo de opciones lo dejamos así:



- Atención al círculo marcado en rojo. Hay un bug en SPSS y si dejamos “valor por defecto” no funciona

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Output

- Información general (para revisar por si hay algo mal)

Data Information		
		N
Cases	Valid	4526
	Out of Range <sup>a</sup>	0
	Missing	0
	Weighted Valid	4526
Categories	Admission	2
	Gender	2
	Department	6

a. Cases rejected because of out of range factor values.

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerarquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
  - Planteamiento del modelo
  - Output
- Interpretación de los resultados
  - Examinando los residuales
  - Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
  - Procedimiento automático
  - Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

- **Residuales y demás (en el caso del modelo saturado no tienen mucho interés)**
  - **No obstante, en modelos no saturados incorporan información importante**

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables

Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal

Ejemplo Berkeley

Planteamiento del modelo

Output

Interpretación de los resultados

Examinando los residuales

Interpretando los coeficientes

Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción

Ejemplo: Fumar y dolor de cabeza

Métodos

Procedimiento automático

Procedimiento manual

Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit

Casillas vacías

Covariantes a nivel de celda

Modelos ordinales

Modelos especiales

- Lo mismo con los resultados de bondad de ajuste

Goodness-of-Fit Tests								
			Chi-Square	df	Sig.			
Likelihood Ratio			nnn	nl				
Cell Counts and Residuals								
Admission	Gender	Department	Observed		Expected		Residuals	Std. Residuals
			Count <sup>a</sup>	%	Count	%		
N	F	A	19,500	,4%	19,500	,4%	,000	,000
		B	8,500	,2%	8,500	,2%	,000	,000
		C	391,500	8,7%	391,500	8,7%	,000	,000
		D	244,500	5,4%	244,500	5,4%	,000	,000
		E	299,500	6,6%	299,500	6,6%	,000	,000
		F	317,500	7,0%	317,500	7,0%	,000	,000
	M	A	313,500	6,9%	313,500	6,9%	,000	,000
		B	207,500	4,6%	207,500	4,6%	,000	,000
		C	205,500	4,5%	205,500	4,5%	,000	,000
		D	279,500	6,2%	279,500	6,2%	,000	,000
		E	138,500	3,1%	138,500	3,1%	,000	,000
		F	351,500	7,8%	351,500	7,8%	,000	,000
Y	F	A	89,500	2,0%	89,500	2,0%	,000	,000
		B	17,500	,4%	17,500	,4%	,000	,000
		C	202,500	4,5%	202,500	4,5%	,000	,000
		D	131,500	2,9%	131,500	2,9%	,000	,000
		E	94,500	2,1%	94,500	2,1%	,000	,000
		F	24,500	,5%	24,500	,5%	,000	,000
	M	A	512,500	11,3%	512,500	11,3%	,000	,000
		B	353,500	7,8%	353,500	7,8%	,000	,000
		C	120,500	2,7%	120,500	2,7%	,000	,000
		D	138,500	3,1%	138,500	3,1%	,000	,000
		E	53,500	1,2%	53,500	1,2%	,000	,000
		F	22,500	,5%	22,500	,5%	,000	,000

a. For saturated models, .500 has been added to all observed cells.

# Contenidos

## Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- Actividades

## Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

## Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley
- Actividades

## Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- Actividades

## Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

- El test para los distintos submodelos sí que es interesante

K-Way and Higher-Order Effects							
	K	df	Likelihood Ratio		Pearson		Number of Iterations
			Chi-Square	Sig.	Chi-Square	Sig.	
K-way and Higher Order Effects <sup>a</sup>	1	23	2650,095	,000	2392,628	,000	0
	2	16	2097,671	,000	2000,328	,000	2
	3	5	20,205	,001	18,815	,002	6
K-way Effects <sup>b</sup>	1	7	552,424	,000	392,300	,000	0
	2	11	2077,466	,000	1981,513	,000	0
	3	5	20,205	,001	18,815	,002	0

a. Tests that k-way and higher order effects are zero.

b. Tests that k-way effects are zero.

- En un análisis loglineal estamos interesados en simplificar el modelo lo más posible (pero que siga ajustando)
- Esta tabla nos da una primera aproximación a esta simplificación

K hace referencia a términos con interacciones de K variables (por ejemplo, K=3 en este caso significa [GDA], mientras que K=2 significa [GD][GA][AD])

La idea es mirar en esta tabla valores no significativos empezando desde arriba hacia abajo. Así, si ocurriera que para K=3 la significación es mayor que 0.05 o el valor que fijemos, nuestro modelo podría excluir esa parte del modelo

En nuestro caso, no podemos excluir ningún término

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Pruebas de asociación parcial

Partial Associations				
Effect	df	Partial Chi-Square	Sig.	Number of Iterations
Admission*Gender	1	1,530	,216	2
Admission*Department	5	763,402	,000	2
Gender*Department	5	1128,696	,000	2
Admission	1	230,028	,000	2
Gender	1	162,874	,000	2
Department	5	159,522	,000	2

- En esta tabla hay tests para cada interacción concreta

La forma de calcularla es ajustar dos modelos, uno con el término y otro sin él y ver si hay diferencia en Chi cuadrado. Cuanto más pequeña la diferencia menos importante es el término y se podría eliminar

- Esto es necesario porque que aunque un grupo de interacciones sea necesario para que el modelo ajuste (en este caso por ejemplo  $K=2$ ), puede que algunas interacciones de ese nivel puedan eliminarse

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Conclusiones sobre el ajuste para los datos de Berkeley
  - Necesitamos un modelo con las interacciones de 3 orden [GDA]
  - La interacción entre Género y Admisión no está clara, pero como los modelos son jerarquicos no hay manera de quitarla fácilmente

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerarquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
  - Planteamiento del modelo
  - Output
- Interpretación de los resultados
  - Examinando los residuales
  - Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
  - Procedimiento automático
  - Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

# Interpretación de los resultados

- Después de encontrar un modelo llega el momento de interpretar. Hay dos maneras para hacer esto:
  - Interpretar los residuales de un modelo que excluye un término en el que estamos interesados
  - Interpretar los coeficientes del modelo que ajusta

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

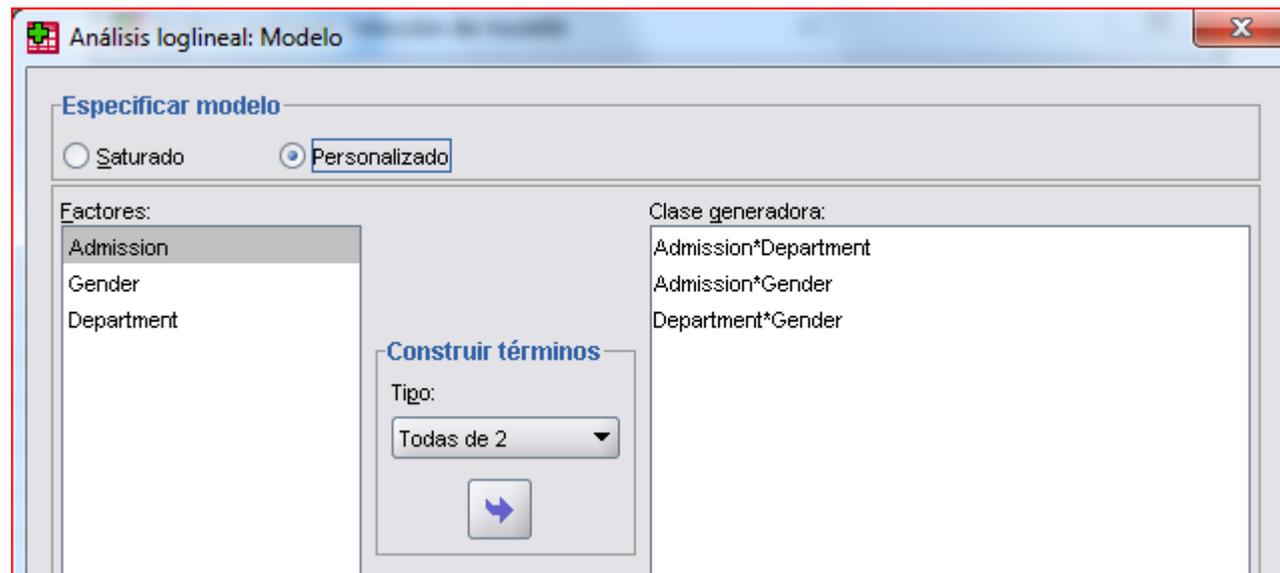
Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

## Examinando los residuales

- La estrategia de mirar los residuales pasa por eliminar un término del modelo que ajusta y ver lo que pasa. De esa manera, los residuales reflejan el efecto de ese término



- Fijaros que en el modelo basta con incluir las interacciones de dos [AD][AG][GD], el programa automáticamente incluye las de nivel inferior

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El modelo no ajusta pero eso es lo que pretendemos

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	20,205	5	,001
Pearson	18,813	5	,002

# Contenidos

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El desajuste viene de los residuales

**Cell Counts and Residuals**

Admission	Gender	Department	Residuals	Std. Residuals
N	F	A	<b>-17,324</b>	<b>-2,874</b>
		B	-,654	-,222
		C	10,784	,553
		D	,744	,048
		E	7,346	,430
		F	-,968	-,054
	M	A	17,170	,998
		B	,529	,037
		C	-10,615	-,723
		D	-,687	-,041
		E	-7,247	-,601
		F	,983	,053
Y	F	A	<b>17,324</b>	<b>2,046</b>
		B	,654	,162
		C	-10,784	-,739
		D	-,744	-,065
		E	-7,346	-,730
		F	,968	,202
	M	A	-17,170	-,746
		B	-,529	-,028
		C	10,615	1,015
		D	,687	,059
		E	7,247	1,071
		F	-,983	-,205

- Los residuales indican que el departamento A aceptó 17 mujeres más que lo esperado. El residual estandarizado para los rechazados mujeres en el grupo A es -2.874 (las mujeres fueron rechazadas menos de lo esperado)

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El resto de los residuales parecen bastante cercanos a las puntuaciones esperadas (hay que fijarse en los valores estandarizados)
  - Esto sugiere que la interacción entre las tres variables es sólo producto de un departamento en el que la discriminación se produjo a favor de las mujeres, no en contra

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes

Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual

Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Interpretando los coeficientes

- En análisis loglineal, como en regresión, la interpretación se puede realizar examinando los coeficientes del modelo. No obstante, en este caso hay una serie de complejidades añadidas.
  - El módulo de selección de modelos sólo calcula los coeficientes para el modelo saturado (en nuestro caso eso no sería problema)
  - La codificación que hace dificulta la interpretación (en loglineal resulta muy interesante utilizar codificación en variables ficticias tal y como hicimos en regresión)
  - Los coeficientes están en logaritmos
- Los antilogaritmos de los coeficientes son cocientes que pueden interpretarse como frecuencia de que algo ocurra respecto a que ocurra otra cosa

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

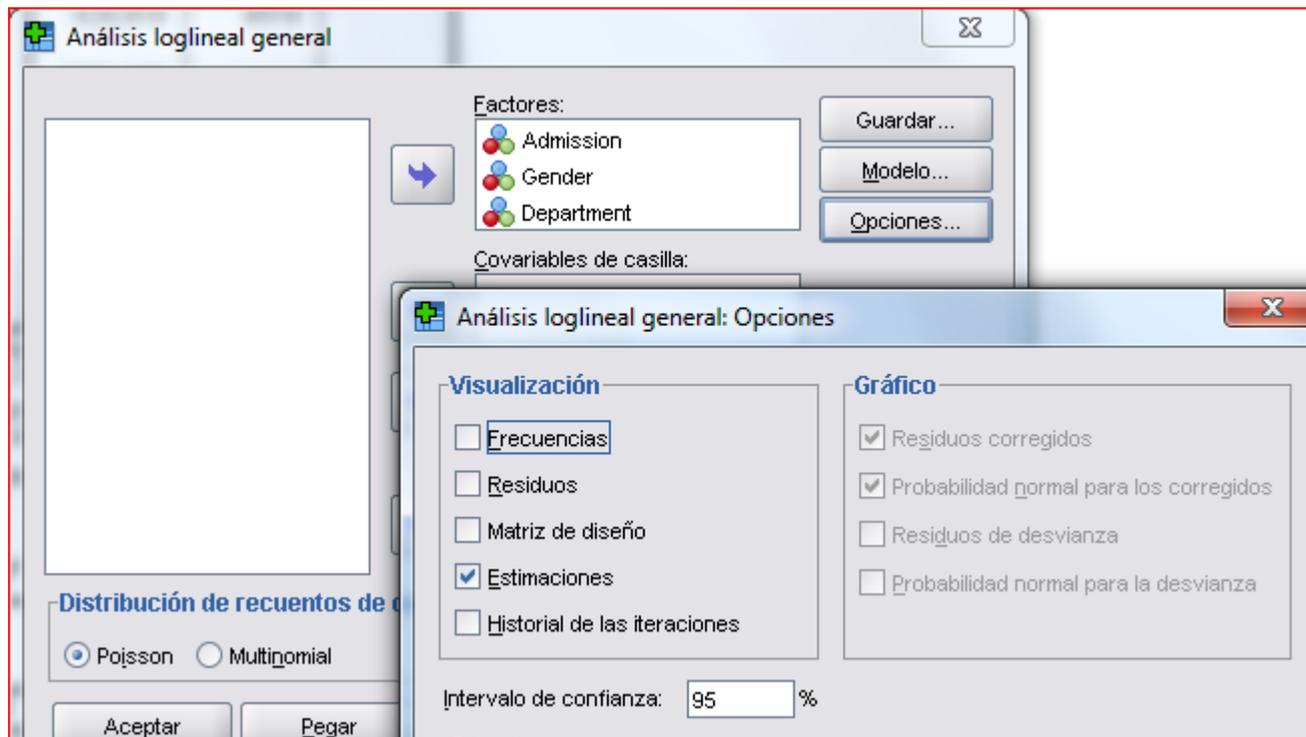
### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Para el módulo de selección de modelos SPSS sólo puede calcular los coeficientes para el modelo saturado (estos coeficientes pueden calcularse con otros módulos)
  - No obstante, la codificación que utiliza hace algo difícil interpretar bien los resultados así que es preferible utilizar el módulo de Loglineal General



## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- La tabla de coeficientes es muy grande y tiene muchos huecos
  - Esto se debe a que para hacer el cálculo es necesario codificar las variables como 1 y 0s pero para cada variable hay que eliminar una de las columnas
  - Así, la variable Género, con dos categorías, se convierte en dos columnas. La primera columna tiene 1 para los hombres y 0 para las mujeres, y la segunda tiene 0 para los hombres y 1 para las mujeres
  - La variable Departamento se convierte en 6 columnas, la primera tiene 1 para el departamento A y 0 para los demás; la segunda tiene 1 para el departamento B y 0 para los demás, etc.
  - Una vez creadas esas columnas, *se elimina la última de las columnas para evitar redundancia (colinealidad). Eso se denomina la matriz de diseño y es lo que permite hacer el cálculo del modelo*
  - Puesto que hay columnas eliminadas, hay coeficientes que no se calculan (aunque eso no significa que no podamos decir nada de ellos)

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Centrándonos en los coeficientes de la interacción para  $K=3$  tenemos lo siguiente:

Parameter	Estimate	Std. Error	Z	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Constant	3,114	,211	14,769	,000	2,700	3,527
[Admission = 1] * [Gender = 1] * [Department = 1]	<b>-,845</b>	<b>,399</b>	<b>-2,121</b>	<b>,034</b>	<b>-1,627</b>	<b>-,064</b>
[Admission = 1] * [Gender = 1] * [Department = 2]	-,002	,523	-,005	,996	-1,028	1,023
[Admission = 1] * [Gender = 1] * [Department = 3]	,312	,335	,934	,350	-,343	,968
[Admission = 1] * [Gender = 1] * [Department = 4]	,105	,337	,311	,756	-,556	,766
[Admission = 1] * [Gender = 1] * [Department = 5]	,389	,362	1,075	,282	-,320	1,099

- Gender=1, Admisión=1, Departamento=1 significa Mujeres, Rechazadas en el departamento A (para ver esto hay que ir a la tabla de datos) tiene el único coeficiente significativo con un valor de -0.845
- Este coeficiente sugiere que las mujeres en el departamento A fueron rechazadas escasamente, pero nos quedamos sin saber en qué medida

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

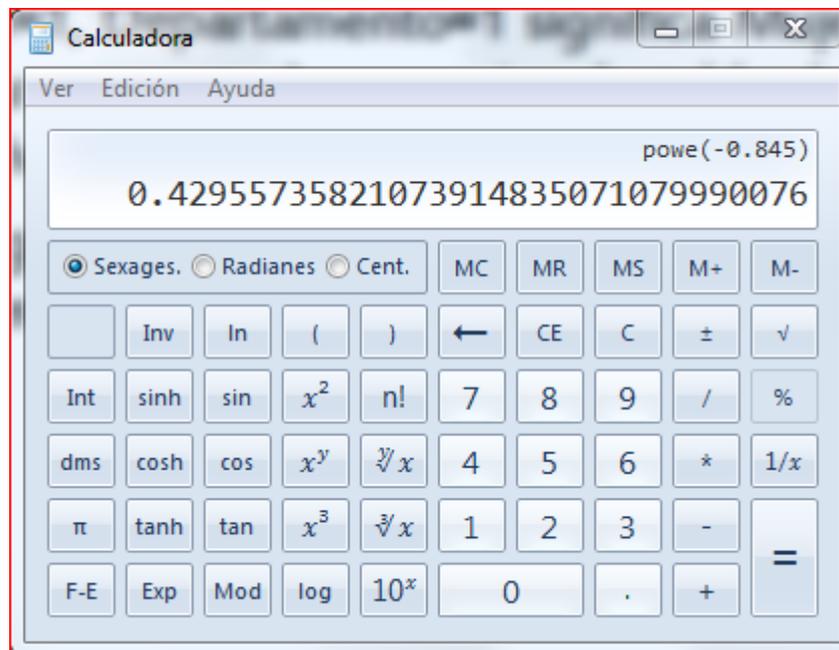
## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Un primer paso para interpretar ese valor es calcular el antilogaritmo del coeficiente



- El valor de 0.43 es lo que se denomina un odds ratio (razón de posibilidades se puede traducir)

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerarquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Unos ejemplos de odds ratio

- El odds ratio de ser rechazado es de  $2771/1755=1.56$  (hay 1.56 veces más rechazados que aceptados)
- El odds ratio ser rechazado para las mujeres versus los hombres es:  
Para las mujeres  $1278/557=2.29$   
Para los hombres  $1493/1198=1.24$
- Mujeres versus hombres  $2.29/1.24=1.83$  (aparentemente hay discriminación)

Gender \* Admission Crosstabulation

Count		Admission		Total
		N	Y	
Gender	F	1278	557	1835
	M	1493	1198	2691
Total		2771	1755	4526

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Odds ratio para mujeres rechazadas versus hombres rechazados para el departamento A versus el departamento F

				Admission	
				N	Y
				Count	Count
Department	A	Gender	F	19	89
			M	313	512
	B	Gender	F	8	17
			M	207	353
	C	Gender	F	391	202
			M	205	120
	D	Gender	F	244	131
			M	279	138
	E	Gender	F	299	94
			M	138	53
	F	Gender	F	317	24
			M	351	22

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Para hacer este cálculo manualmente podemos hacer
  1. Discriminación en el departamento A: Mujeres  $19/89=0.21$ ; Hombres  $313/512=0.61$ . Total 0.35 (la posibilidad de ser rechazada en el departamento A es de 0.35 respecto a los hombres: la discriminación es positiva hacia las mujeres)
  2. Para el departamento F: Mujeres  $317/24=13.2$ ; Hombres  $351/22=16$ . Total  $13.2/16=0.82$  (la discriminación es también positiva hacia las mujeres pero no tan exagerada)
  3. Finalmente,  $0.35/0.82=0.43$  (el valor del coeficiente calculado para el análisis loglineal)
  4. Este valor puede interpretarse como la discriminación hacia las mujeres en el departamento A frente al departamento F

Un valor de 1 hubiera significado igual discriminación. Mayor que 1, más rechazo en F que en A para las mujeres.

El valor de 0.43 significa que en el departamento A el rechazo hacia las mujeres es la mitad que en el F. Es decir, en el A se prefiere a las mujeres más del doble que en el F.

**Introducción**

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

**Modelos loglineales**

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

**Admisión en Berkeley**

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

**Ajuste de modelos**

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

**Casos especiales**

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- **Resumen coeficientes y odds ratio**

- Los coeficientes son los odds ratios de las puntuaciones esperadas para las categorías comparadas con las categorías de referencia
- En SPSS, la categoría de referencia es la última de cada variable
- En nuestro caso, para  $K=3$ , sólo una de los coeficientes es significativo, el del departamento A

Eso significa que el único coeficiente que es diferente del de F es el de A. Los demás tienen coeficientes similares

El valor en la escala logarítmica es negativo (-0.86) lo que significaría de una manera aproximada que se rechaza a las mujeres menos en el A que en el F

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

# Conclusiones para Berkeley

- Se constata que hay un efecto del género y el departamento sobre la admisión pero para un departamento concreto
  - En ese departamento, la admisión favorece a las mujeres
- En los otros departamentos no se aprecia discriminación

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

## Actividades

1. En el ejemplo de 1991 US General Survey Sexo Feliz tenemos dos variables. Este es un ejemplo sencillo, en el que se puede probar el efecto en esa muestra del género sobre la felicidad.

*Una estrategia sencilla es valorar los residuales para ver en qué casillas se producen las diferencias si las hay. Un elemento interesante es calcular los odds ratios para el modelo saturado y ver en qué consiste la diferencia. Para ello se puede ajustar el modelo saturado en Loglineal General. Los dos parámetros más interesantes son  $\text{sexo}=1$ ,  $\text{feliz}=1$  y  $\text{feliz}=2$ . El primero es  $.507$  que su antilogaritmo es  $1.66$  (los hombres dicen que son más felices que no felices  $1.66$  veces más que las mujeres). Ese mismo valor puede calcularse directamente utilizando las frecuencias observadas (206.5, 53.5, 261.5, 112.5)*

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

2. Felicidad (Muy Feliz, Bastante, No demasiado Feliz) en función del Género y la raza (B, N y Otra). Examina el efecto del género y la raza sobre la felicidad indicando qué modelo ajusta y qué términos son más importantes.

*En este caso la interacción entre sexo y raza no convendría quitarla ya que por definición esperamos que sea cero y si no ocurre así es por un error en el muestreo así que conviene incluirla para fijar sus residuales a cero. No obstante el procedimiento automático lo hace. La tabla final muestra lo siguiente. Tanto el Género como la Raza tienen efecto pero no hay interacción.*

Step Summary

Step <sup>a</sup>		Effects	Chi-Square <sup>c</sup>	df	Sig.	Number of Iterations
0	Generating Class <sup>b</sup>	sexo*raza*feliz	,000	0	.	
	Deleted Effect 1	sexo*raza*feliz	3,426	4	,489	3
1	Generating Class <sup>b</sup>	sexo*raza, sexo*feliz, raza*feliz	3,426	4	,489	
	Deleted Effect 1	sexo*raza	3,880	2	,144	2
	2	sexo*feliz	6,702	2	,035	2
	3	raza*feliz	21,212	4	,000	2
2	Generating Class <sup>b</sup>	sexo*feliz, raza*feliz	7,305	6	,294	
	Deleted Effect 1	sexo*feliz	7,936	2	,019	2
	2	raza*feliz	22,446	4	,000	2
3	Generating Class <sup>b</sup>	sexo*feliz, raza*feliz	7,305	6	,294	

a. At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than ,050.

b. Statistics are displayed for the best model at each step after step 0.

c. For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes

Conclusiones para Berkeley

[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos

Procedimiento automático  
Procedimiento manual

Resumen sobre ajuste de modelos

[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# *Ajuste de modelos*

# Introducción

- El proceso de ajustar modelos en análisis loglineal es de mucha importancia
  - Hemos visto, que cuando los modelos tienen muchos términos, su interpretación va haciéndose cada vez más compleja
  - Conseguir modelos que no sean complejos nos va a permitir hacer interpretaciones más simples
  - También, mediante este procedimiento podemos probar hipótesis específicas sobre términos
  - Existen procedimientos automáticos, pero es interesante realizar parte de ese trabajo manualmente
  - Hay interacciones que no es conveniente quitar puesto que nos sirven para controlar variables confundentes

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Ejemplo: Fumar y dolor de cabeza

- Tenemos una muestra de sujetos con información sobre fumar, dolor de cabeza y el género

				Headache	
				LESSTHWEEK	MORETHWEEK
				Row N %	Row N %
Smoke	NOT	Gender	MEN	93,9%	6,1%
			WOMEN	82,0%	18,0%
	YES	Gender	MEN	92,9%	7,1%
			WOMEN	72,6%	27,4%

- Mirando la tabla vemos que parece que fumar parece tener efecto sobre el dolor de cabeza, y también las mujeres parecen tener más dolores de cabeza que los hombres

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Métodos

- Probaremos dos métodos
  - El automático con selección hacia atrás
  - El manual para comprobar hipótesis específicas

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables

Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal

Ejemplo Berkeley

Planteamiento del modelo

Output

Interpretación de los resultados

Examinando los residuales

Interpretando los coeficientes

Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción

Ejemplo: Fumar y dolor de cabeza

Métodos

Procedimiento automático

Procedimiento manual

Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit

Casillas vacías

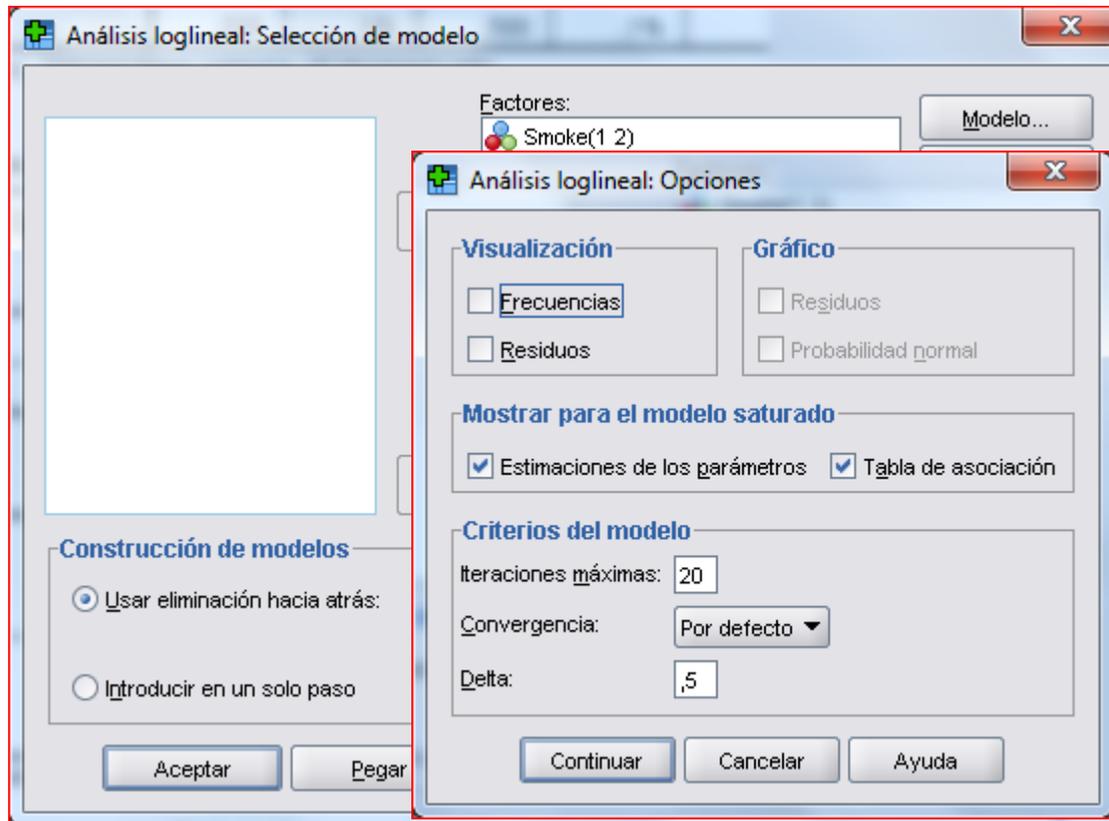
Covariantes a nivel de celda

Modelos ordinales

Modelos especiales

# Procedimiento automático

- En el modelo de selección de modelos hay un procedimiento de eliminación hacia atrás. En opciones elegir estimaciones y tabla de asociación.



## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- El output interesante empieza en esta tabla

	K	df	Likelihood Ratio		Pearson		Number of Iterations
			Chi-Square	Sig.	Chi-Square	Sig.	
K-way and Higher Order	1	7	659,371	,000	612,739	,000	0
Effects <sup>a</sup>	2	4	60,808	,000	61,448	,000	2
	3	1	,825	,364	,837	,360	7
K-way Effects <sup>b</sup>	1	3	598,563	,000	551,291	,000	0
	2	3	59,982	,000	60,611	,000	0
	3	1	,825	,364	,837	,360	0

a. Tests that k-way and higher order effects are zero.

b. Tests that k-way effects are zero.

- En esta tabla vemos que la aportación de las interacciones de 3 variables no son suficientes y las podemos desestimar. Nuestro modelo ajusta sólo con interacciones entre dos variables.
- No obstante, hay varias posibles modelos con interacciones entre dos variables. Para evaluar el adecuado se puede utilizar el output a continuación.

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Tabla de eliminación hacia atrás

Step Summary					
Step <sup>a</sup>	Effects	Chi-Square	df	Sig.	Number of Iterations
0 Generating Class <sup>b</sup>	Smoke*Gender*Headache	,000	0	.	
Deleted Effect 1	<b>Smoke*Gender*Headache</b>	<b>,825</b>	<b>1</b>	<b>,364</b>	7
1 Generating Class <sup>b</sup>	Smoke*Gender, Smoke*Headache, Gender*Headache	,825	1	,364	
Deleted Effect 1	<b>Smoke*Gender</b>	<b>3,226</b>	<b>1</b>	<b>,072</b>	2
Deleted Effect 2	Smoke*Headache	4,807	1	,028	2
Deleted Effect 3	Gender*Headache	54,910	1	,000	2
2 Generating Class <sup>b</sup>	Smoke*Headache, Gender*Headache	4,051	2	,132	
Deleted Effect 1	<b>Smoke*Headache</b>	<b>3,327</b>	<b>1</b>	<b>,068</b>	2
Deleted Effect 2	Gender*Headache	53,430	1	,000	2
3 Generating Class <sup>b</sup>	Gender*Headache, Smoke	7,378	3	,061	
Deleted Effect 1	Gender*Headache	53,430	1	,000	2
Deleted Effect 2	Smoke	6,062	1	,014	2
4 Generating Class <sup>b</sup>	<b>Gender*Headache, Smoke</b>	<b>7,378</b>	<b>3</b>	<b>,061</b>	

a. At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than ,050.

b. Statistics are displayed for the best model at each step after step 0.

c. For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

- Esta tabla es el resumen de una serie de modelos en los que los distintos términos van eliminándose uno por uno por pasos.
- En el primer paso se elimina el término **Smoke\*Gender\*Headache**, en el segundo **Smoke\*Headache** y luego **Smoke\*Gender**
- El modelo final es **[GH][S]**

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Si nos quedamos con ese modelo, no hay relación entre Fumar y Dolor de cabeza, ni tampoco entre Fumar y Género.
- Sí que habría relación entre Género y Dolor de Cabeza.

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

## Procedimiento manual

- El procedimiento anterior nos ha dado como resultado un modelo que excluye dos interacciones de segundo orden
  - No obstante, la interacción de interés central es la de Fumar con Dolor de cabeza.
  - Una vez hemos encontrado un modelo que ajusta podríamos quedarnos con él y no hacer la prueba del término que nos falta, pero eso no sería buena idea
  - Por otro lado, aunque podemos quitar la interacción entre Fumar y Género, puesto que no es de interés directo probar si es distinta o no de cero, es mejor incluirla para de ese modo eliminar la duda de si tiene algún efecto

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Este modelo lo podemos encontrar en la tabla que vimos anteriormente

Step Summary						
Step <sup>a</sup>		Effects	Chi-Square <sup>c</sup>	df	Sig.	Number of Iterations
0	Generating Class <sup>b</sup>	Smoke*Gender*Headache	,000	0	.	
	Deleted Effect	1 Smoke*Gender*Headache	,825	1	,364	7
1	Generating Class <sup>b</sup>	Smoke*Gender, Smoke*Headache, Gender*Headache	,825	1	,364	
	Deleted Effect	1 Smoke*Gender	3,226	1	,072	2
		2 <b>Smoke*Headache</b>	<b>4,807</b>	<b>1</b>	<b>,028</b>	<b>2</b>
		3 Gender*Headache	54,910	1	,000	2
2	Generating Class <sup>b</sup>	Smoke*Headache, Gender*Headache	4,051	2	,132	
	Deleted Effect	1 Smoke*Headache	3,327	1	,068	2
		2 Gender*Headache	53,430	1	,000	2
3	Generating Class <sup>b</sup>	Gender*Headache, Smoke	7,378	3	,061	
	Deleted Effect	1 Gender*Headache	53,430	1	,000	2
		2 Smoke	6,062	1	,014	2
4	Generating Class <sup>b</sup>	Gender*Headache, Smoke	7,378	3	,061	

a. At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than ,050.

b. Statistics are displayed for the best model at each step after step 0.

c. For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

- El término en negrita es un test de la diferencia del modelo [SG][SH][GH] frente a [SG][GH] lo cual nos da [SH], es decir el término que nos interesa
  - Esa diferencia es significativa (Sig .028) lo cual puede interpretarse como que efectivamente existe un efecto de Fumar sobre el dolor de cabeza
  - Ese término pasa a ser no significativo al quitar posteriormente [GH]

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

- Esta técnica, comparar un modelo más completo con uno que difiere solamente en los términos concretos que queremos probar es muy útil
  - Hemos visto que esas comparaciones están hechas automáticamente en la tabla de Step Summary en muchas ocasiones (aunque no siempre)
- Si queremos hacer una comparación más compleja entre dos modelos y esta no está automáticamente en la tabla de Step summary podemos calcularlo manualmente
  - Calculando el Likelihood Ratio para los dos modelos si un modelo está anidado dentro del otro (todos los términos del modelo inferior están también en el modelo superior)
  - Restando la diferencia en el Likelihood Ratio y la diferencia en grados de libertad. Puede examinarse la significación con una calculadora de Chi cuadrado (hay muchas en internet <http://stattrek.com/Tables/ChiSquare.aspx>) o usar mi programa

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Resumen sobre ajuste de modelos

- Una parte muy importante de un análisis loglineal es realizar comparaciones entre modelos para ver el efecto de términos específicos
  - Hemos visto que de esa manera hemos podido identificar el efecto (pequeño) de fumar sobre el dolor de cabeza
- Es importante recordar que en el análisis loglineal buscar un modelo que ajuste no es siempre necesario ya que nuestro interés puede estar en un elemento concreto

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

[Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
  - Planteamiento del modelo
  - Output
- Interpretación de los resultados
  - Examinando los residuales
  - Interpretando los coeficientes
- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
  - Procedimiento automático
  - Procedimiento manual
- Resumen sobre ajuste de modelos

[Actividades](#)

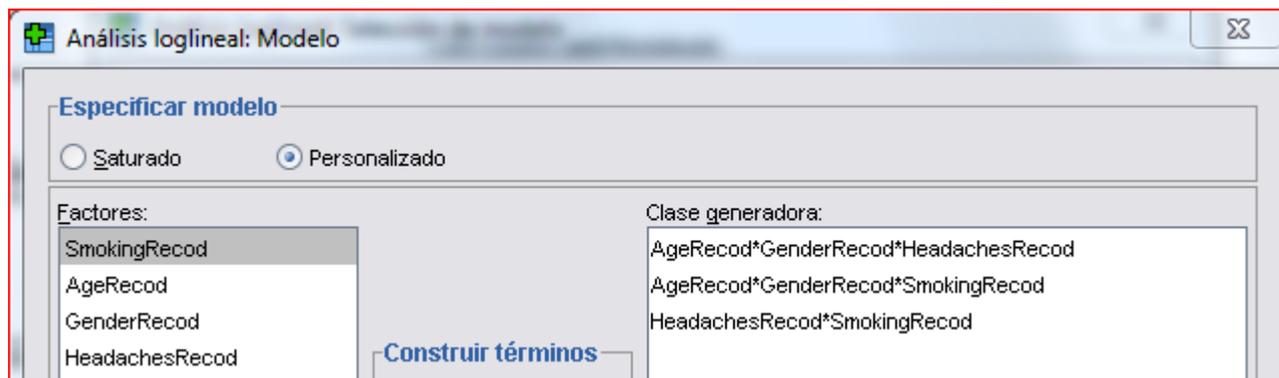
### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

## Actividades

1. Fumar, Dolor de cabeza, Género y Edad. Se encuentra en Smoke4var.sav. En este caso, añadimos una variable al problema

*La interacción más interesante es la de fumar con dolores de cabeza, sin embargo, si hay interacciones a nivel superior quizás tengamos que considerarlas. En este caso el modelo a comprobar sería  $[AGH][SAG][SH]$  con el modelo sin  $[SH]$ . Es posible también obtener un modelo muy simplificado que ajusta con interacciones a nivel 2. Si se ajusta este modelo frente a uno que no tenga la última línea se está haciendo la comparación concreta que estamos interesados.*



*El resultado de los dos modelos es el de las dos tablas de abajo. Haciendo la resta tenemos que la diferencia en devianza es  $5.9471 - 0.971$  y la diferencia de grados de libertad es 1. Se puede ver la significación aquí (el resultado hay que restarlo de 1 y lleva a rechazar la hipótesis nula de igualdad de modelos).*

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	,971	3	,808
Pearson	,980	3	,806

Goodness-of-Fit Tests

	Chi-Square	df	Sig.
Likelihood Ratio	5,741	4	,219
Pearson	5,683	4	,224

# Contenidos

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen  
[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

2. Arthritis3way.sav tiene el efecto de un experimento para mejorar la Artritis. Los resultados están en función del Género. Es interesante también ver si el género también tiene efecto sobre la respuesta (independientemente del tratamiento)

*Calcula también los coeficientes para ver si los tratamientos producen el efecto deseado. Los resultados indican que hay interacción entre tratamiento y respuesta, y que el género no interacciona. No obstante, es cuestionable si la interacción entre Género y tratamiento debería ser eliminada. Si no es eliminada, entonces la interacción entre Género y Respuesta pasa a ser significativa. Este efecto del género sobre la respuesta puede verse en los residuales del modelo que calcular utilizando el método backward (es muy poco pero parece que las mujeres responden más al tratamiento).*

Step Summary

Step <sup>a</sup>		Effects	Chi-Square <sup>c</sup>	df	Sig.	Number of Iterations
0	Generating Class <sup>b</sup>	Genero*Tratamineto*Respuesta	,000	0	.	
	Deleted Effect 1	Genero*Tratamineto*Respuesta	1,705	2	,426	4
1	Generating Class <sup>b</sup>	Genero*Tratamineto, Genero*Respuesta, Tratamineto*Respuesta	1,705	2	,426	
	Deleted Effect 1	Genero*Tratamineto	3,088	1	,079	2
	2	Genero*Respuesta	7,364	2	,025	2
	3	Tratamineto*Respuesta	15,880	2	,000	2

## Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

## Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerarquicos o anidados

## Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes

Conclusiones para Berkeley  
[Actividades](#)

## Ajuste de modelos

Introducción

Ejemplo: Fumar y dolor de cabeza  
Métodos

Procedimiento automático  
Procedimiento manual

Resumen sobre ajuste de modelos  
[Actividades](#)

## Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

2	Generating Class <sup>b</sup>	Genero*Respuesta,	4,793	3	,188	
	Deleted Effect	Tratamineto*Respuesta				
	1	Genero*Respuesta	5,013	2	,082	2
	2	Tratamineto*Respuesta	13,530	2	,001	2
3	Generating Class <sup>b</sup>	Tratamineto*Respuesta, Genero	9,806	5	,081	
	Deleted Effect	Tratamineto*Respuesta	13,530	2	,001	2
	1	Genero	14,165	1	,000	2
	2					
4	Generating Class <sup>b</sup>	Tratamineto*Respuesta, Genero	9,806	5	,081	

- a. At each step, the effect with the largest significance level for the Likelihood Ratio Change is deleted, provided the significance level is larger than ,050.
- b. Statistics are displayed for the best model at each step after step 0.
- c. For 'Deleted Effect', this is the change in the Chi-Square after the effect is deleted from the model.

Cell Counts and Residuals

Genero	Tratamineto	Respuesta	Observed		Expected		Residuals	Std. Residuals
			Count	%	Count	%		
F	PLACEBO	Lots	6,000	7,1%	4,917	5,9%	1,083	,489
		None	19,000	22,6%	20,369	24,2%	-1,369	-,303
		Some	7,000	8,3%	4,917	5,9%	2,083	,940
	TEST	Lots	16,000	19,0%	14,750	17,6%	1,250	,325
		None	6,000	7,1%	9,131	10,9%	-3,131	-1,036
		Some	5,000	6,0%	4,917	5,9%	,083	,038
M	PLACEBO	Lots	1,000	1,2%	2,083	2,5%	-1,083	-,751
		None	10,000	11,9%	8,631	10,3%	1,369	,466
		Some	,000	,0%	2,083	2,5%	-2,083	-1,443
	TEST	Lots	5,000	6,0%	6,250	7,4%	-1,250	-,500
		None	7,000	8,3%	3,869	4,6%	3,131	1,592
		Some	2,000	2,4%	2,083	2,5%	-,083	-,058

# Contenidos

## Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen
- [Actividades](#)

## Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

## Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley
- [Actividades](#)

## Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos
- [Actividades](#)

## Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

3. En el archivo Pain.sav hay tres variables sobre un tratamiento para el dolor. La variable a explicar es Adverso (si o no) y es interesante ver el diagnostico que ha recibido y el tratamiento (4 categorías)

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

[Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic

- Modelos para cuatro variables

- Modelos jerarquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal

- Ejemplo Berkeley

  - Planteamiento del modelo

  - Output

- Interpretación de los resultados

  - Examinando los residuales

  - Interpretando los coeficientes

- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción

- Ejemplo: Fumar y dolor de cabeza

- Métodos

  - Procedimiento automático

  - Procedimiento manual

- Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

- Modelos logit

- Casillas vacías

- Covariantes a nivel de celda

- Modelos ordinales

- Modelos especiales

4. En el archivo `dossintomas.sav` hay información sobre el grupo de edad de unos mine-ros y dos síntomas de enfermedades respiratorias (ahogos y pitidos en los pulmones). Ten en cuenta que los datos están agrupados así que hay que ponderar por la variable `Freq` (comando `Datos>Ponderar Casos`)

*Este es un ejemplo con dos variables dependientes. Aquí es interesante ver si según el grupo de edad los síntomas van a peor. También, si hay interacción entre los síntomas (cuanto más de un síntoma más de otro). La interacción de tercer orden sólo se produce en el grupo de edad de 55-59 años en la que se da una abundancia de casos que tienen una enfermedad o la otra pero no simultáneamente (residuales -0.8, 1.8, 2.2, -1.2). Por lo demás, si analizamos los residuales de un modelo que no incluya la interacción con alguno de los síntomas, se puede ver que a medida que aumenta la edad el síntoma va aumentando claramente. La tabla de abajo muestra el modelo [EA][AP] y por tanto muestra el efecto de la edad sobre los pitidos en los residuales.*

Cell Counts and Residuals

Edad	Ahogo	Pitido	Observed		Expected		Residuals	Std. Residuals
			Count	%	Count	%		
20_24	NOT	NO	1841,000	10,1%	1712,179	9,4%	128,821	3,113
		YES	95,000	,5%	223,821	1,2%	-128,821	-8,611
	YES	NO	7,000	,0%	3,956	,0%	3,044	1,531
		YES	9,000	,0%	12,044	,1%	-3,044	-,877
25_29	NOT	NO	1654,000	9,0%	1555,642	8,5%	98,358	2,494
		YES	105,000	,6%	203,358	1,1%	-98,358	-6,897
	YES	NO	9,000	,0%	7,911	,0%	1,089	,387
		YES	23,000	,1%	24,089	,1%	-1,089	-,222

## Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

- [Actividades](#)

## Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerárquicos o anidados

## Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output
- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes
- Conclusiones para Berkeley

- [Actividades](#)

## Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
- Procedimiento automático
- Procedimiento manual
- Resumen sobre ajuste de modelos

- [Actividades](#)

## Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

30_34	NOT	NO	1863,000	10,2%	1804,155	9,9%	58,845	1,385
		YES	177,000	1,0%	235,845	1,3%	-58,845	-3,832
	YES	NO	19,000	,1%	18,047	,1%	,953	,224
		YES	54,000	,3%	54,953	,3%	-,953	-,129
35_39	NOT	NO	2357,000	12,9%	2311,795	12,6%	45,205	,940
		YES	257,000	1,4%	302,205	1,7%	-45,205	-2,600
	YES	NO	48,000	,3%	41,780	,2%	6,220	,962
		YES	121,000	,7%	127,220	,7%	-6,220	-,551
40_44	NOT	NO	1778,000	9,7%	1813,883	9,9%	-35,883	-,843
		YES	273,000	1,5%	237,117	1,3%	35,883	2,330
	YES	NO	54,000	,3%	55,130	,3%	-1,130	-,152
		YES	169,000	,9%	167,870	,9%	1,130	,087
45_49	NOT	NO	1712,000	9,4%	1800,618	9,8%	-88,618	-2,088
		YES	324,000	1,8%	235,382	1,3%	88,618	5,776
	YES	NO	88,000	,5%	88,257	,5%	-,257	-,027
		YES	269,000	1,5%	268,743	1,5%	,257	,016
50_54	NOT	NO	1324,000	7,2%	1387,608	7,6%	-63,608	-1,708
		YES	245,000	1,3%	181,392	1,0%	63,608	4,723
	YES	NO	117,000	,6%	128,801	,7%	-11,801	-1,040
		YES	404,000	2,2%	392,199	2,1%	11,801	,596
55_59	NOT	NO	967,000	5,3%	1054,193	5,8%	-87,193	-2,685
		YES	225,000	1,2%	137,807	,8%	87,193	7,428
	YES	NO	152,000	,8%	137,948	,8%	14,052	1,196
		YES	406,000	2,2%	420,052	2,3%	-14,052	-,686
60_64	NOT	NO	526,000	2,9%	581,928	3,2%	-55,928	-2,318
		YES	132,000	,7%	76,072	,4%	55,928	6,412
	YES	NO	106,000	,6%	118,171	,6%	-12,171	-1,120
		YES	372,000	2,0%	359,829	2,0%	12,171	,642

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables

Modelos para dos variables

Modelos para tres variables

Notación

Modelos para Titanic

Modelos para cuatro variables

Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal

Ejemplo Berkeley

Planteamiento del modelo

Output

Interpretación de los resultados

Examinando los residuales

Interpretando los coeficientes

Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción

Ejemplo: Fumar y dolor de cabeza

Métodos

Procedimiento automático

Procedimiento manual

Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit

Casillas vacías

Covariantes a nivel de celda

Modelos ordinales

Modelos especiales

## 5. JointIndependence.sav es un ejemplo de libro. Muy fácil de conseguir un modelo que ajuste

*Sólo hay una interacción de segundo orden. Joint Independence es el nombre para uno de los modelos de interacción.*

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

[Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerarquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
  - Planteamiento del modelo
  - Output
- Interpretación de los resultados
  - Examinando los residuales
  - Interpretando los coeficientes
- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
  - Procedimiento automático
  - Procedimiento manual
- Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

6. Melanoma.sav es un archivo sobre tipos de melanoma y lugares en los que ocurre más habitualmente. Lo interesante es detectar en qué sitios se da más cierto tipo.

*Un ejercicio con dos variables. No necesitaríamos loglineal para esto pero también se puede resolver de esa manera*

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

[Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic
- Modelos para cuatro variables
- Modelos jerarquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
  - Planteamiento del modelo
  - Output
- Interpretación de los resultados
  - Examinando los residuales
  - Interpretando los coeficientes
- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción
- Ejemplo: Fumar y dolor de cabeza
- Métodos
  - Procedimiento automático
  - Procedimiento manual
- Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

- Modelos logit
- Casillas vacías
- Covariantes a nivel de celda
- Modelos ordinales
- Modelos especiales

7. Norespuesta.sav se trata de un ejemplo en el que se analiza si la gente no quiso contestar a unas preguntas de encuesta en la calle. Hay tres regiones (copenhage, ciudades pequeñas y en el campo). Es interesante ver qué produjo más no respuesta.

*Generalmente, se piensa que en las ciudades grandes la gente responde menos. ¿Se cumple esto? En este ejemplo es uno de los pocos casos en que se puede eliminar uno de los efectos principales (género), ya que al ser datos de encuesta en realidad hay el mismo número de hombres que de mujeres. Al hacerlo se gana en grados de libertad lo que hace que el modelo [Region Respuesta] ajuste. En este caso, el método automático no es capaz de encontrar este modelo*

## Contenidos

### Introducción

- Tablas de contingencia
- Chi cuadrado
- Residuales estandarizados
- Resumen

[Actividades](#)

### Modelos loglineales

- Modelos para dos y tres variables
- Modelos para dos variables
- Modelos para tres variables
- Notación
- Modelos para Titanic

- Modelos para cuatro variables
- Modelos jerárquicos o anidados

### Admisión en Berkeley

- Pasos para un análisis loglineal
- Ejemplo Berkeley
- Planteamiento del modelo
- Output

- Interpretación de los resultados
- Examinando los residuales
- Interpretando los coeficientes

- Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

- Introducción

- Ejemplo: Fumar y dolor de cabeza

- Métodos

  - Procedimiento automático

  - Procedimiento manual

- Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

- Modelos logit

- Casillas vacías

- Covariantes a nivel de celda

- Modelos ordinales

- Modelos especiales

# *Casos especiales*

# Modelos logit

- Los modelos logit pueden interpretarse como un caso especial de los modelos loglineales cuando establecemos una variable dependiente
- Esto simplifica mucho la interpretación y análisis de los resultados y cuando se tiene más de 3 variables sobre todo resulta muy recomendable
- En un análisis logit, todas las interacciones que incluyen sólo variables independientes se introducen automáticamente en el modelo. Las únicas interacciones disponibles para ser evaluadas son las que incluyen a la variable dependiente (para así ver si explican la variable dependiente)

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Casillas vacías

- Cuando hay celdas con valores iguales a cero se producen problemas de cálculo que puede hacer que:
  - El número de grados de libertad de las pruebas de bondad de ajuste sean incorrectos
  - Algunos parámetros no se pueden calcular, o dan resultados sin sentido (esto sobre todo pasa en el módulo General del análisis loglineal)
  - En SPSS se puede especificar que alguna de las casillas es cero (p.e. niños en la tripulación) pero eso no corrige el problema de los grados de libertad
- Aunque hay algunas propuestas sobre corrección de los grados de libertad (p.e. Clogg and Eliason, 1987) no hay una respuesta muy clara y además los paquetes estadísticos no las suelen contemplar.
- Lo mejor es simplemente colapsar categorías o eliminar categorías de nuestros análisis (si podemos)

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Covariantes a nivel de celda

- SPSS admite variables numéricas que actúan como variables covariantes a nivel de casilla/celda
- Esto permite ajustar modelos que prueban hipótesis que introducen variables numéricas o consideran las variables categóricas como si fueran numéricas
  - Por ejemplo, uno puede utilizar una variable categorizada como niveles de salario como numérica utilizando los valores medios de las categorías
- Aunque este tipo de análisis es posible en SPSS, normalmente los libros sobre análisis loglineal no tratan esta posibilidad
  - Este tipo de problemas pueden tratarse con otro tipo de técnicas

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal

Ejemplo Berkeley

Planteamiento del modelo  
Output

Interpretación de los resultados

Examinando los residuales

Interpretando los coeficientes

Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción

Ejemplo: Fumar y dolor de cabeza

Métodos

Procedimiento automático

Procedimiento manual

Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit

Casillas vacías

Covariantes a nivel de celda

Modelos ordinales

Modelos especiales

# Modelos ordinales

- Eligiendo los valores para la variable que hace de covariante adecuadamente se pueden introducir restricciones que llevan a modelos ordinales
  - Por ejemplo, en Titanic podríamos ver si la clase social tiene un efecto de este tipo
- Este tipo de modelos sí que se suelen ver en conexión con análisis loglineal y no están aquí por que me parecía excesivo para el tiempo dedicado

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic

Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output

Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual

Resumen sobre ajuste de modelos  
[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales

# Modelos especiales

- Hay una serie de modelos especiales que eliminan por ejemplo la diagonal de una tabla de tabulaciones cruzadas, o que comparan la parte de arriba con la parte de abajo de una matriz de frecuencias
  - Estos modelos no los he incluido en este tema pero el SPSS los calcula también utilizando pesos para las celdas
  - Por razones de tiempo y espacio no he incluido los modelos

## Contenidos

### Introducción

Tablas de contingencia  
Chi cuadrado  
Residuales estandarizados  
Resumen

[Actividades](#)

### Modelos loglineales

Modelos para dos y tres variables  
Modelos para dos variables  
Modelos para tres variables  
Notación  
Modelos para Titanic  
Modelos para cuatro variables  
Modelos jerárquicos o anidados

### Admisión en Berkeley

Pasos para un análisis loglineal  
Ejemplo Berkeley  
Planteamiento del modelo  
Output  
Interpretación de los resultados  
Examinando los residuales  
Interpretando los coeficientes  
Conclusiones para Berkeley

[Actividades](#)

### Ajuste de modelos

Introducción  
Ejemplo: Fumar y dolor de cabeza  
Métodos  
Procedimiento automático  
Procedimiento manual  
Resumen sobre ajuste de modelos

[Actividades](#)

### Casos especiales

Modelos logit  
Casillas vacías  
Covariantes a nivel de celda  
Modelos ordinales  
Modelos especiales