

SEMINARIO DE POSGRADO

TÉCNICAS AVANZADAS DE INVESTIGACIÓN SOCIAL

MÓDULO 2.2 A

**Agustín Salvia
Santiago Poy**

**Análisis multivariado
Modelo Lazarsfeld**

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA

EL PAPEL DEL ANÁLISIS MULTIVARIABLE ES PROPORCIONAR LOS SUSTITUTOS LÓGICOS DEL CONTROL EXPERIMENTAL Y PONER A PRUEBA HIPÓTESIS ROBUSTAS ACERCA DE FENÓMENOS SOCIALES COMPLEJOS.

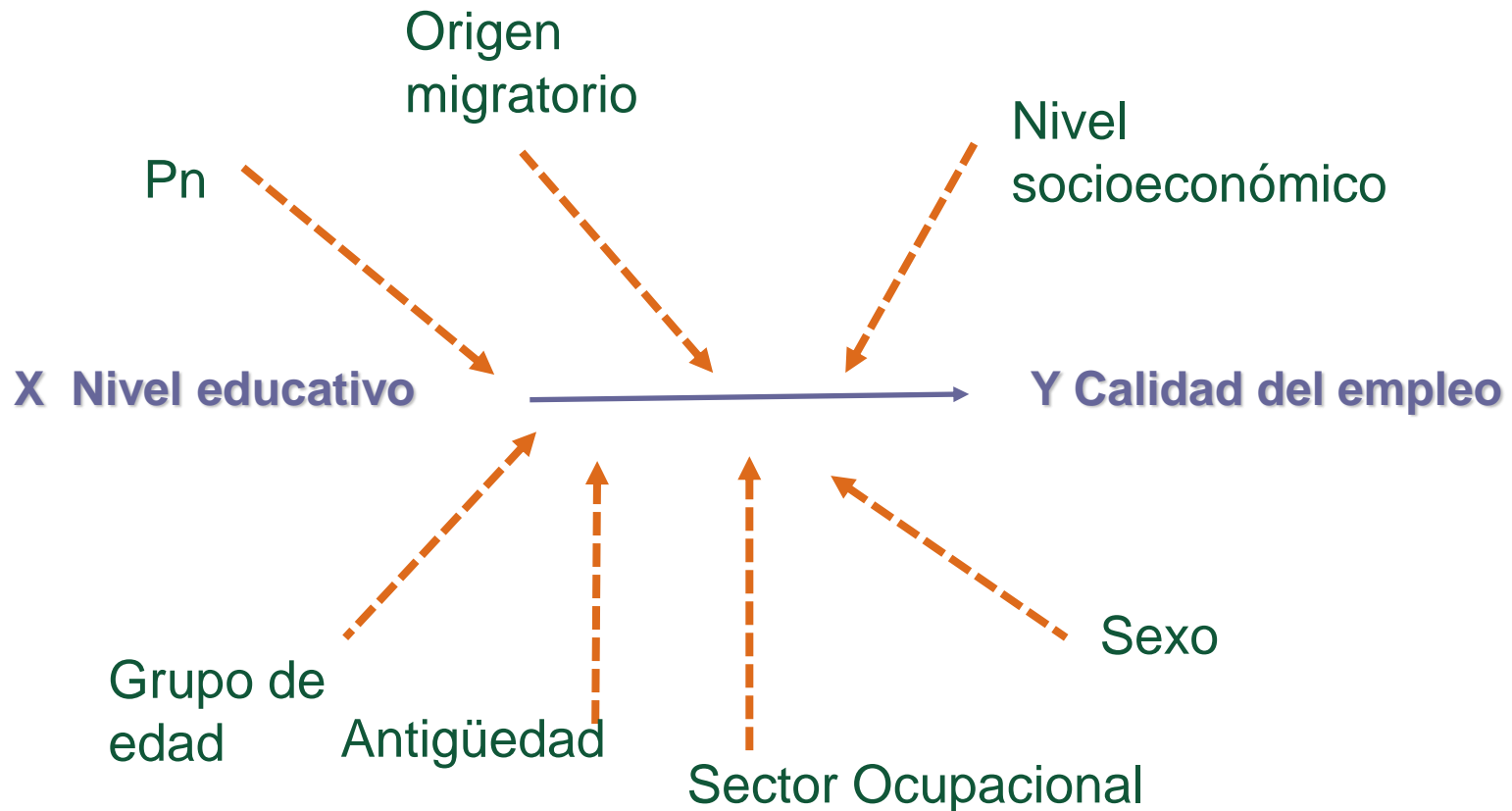
DOS TIPOS DE PROBLEMAS ENFRENTA EL ANÁLISIS MULTIVARIADO

- **Análisis de los datos: ¿cómo manipular la información, resumirla, identificar y evaluar las diferentes relaciones?**
- **Interpretación de los datos: ¿cómo diferenciar los efectos particulares de los de interacción y cómo evaluar de manera racional el sentido de las regularidades empíricas?**

Análisis de asociación

hipótesis

Entre



Análisis de asociación

Variables explicativas

X Nivel ed \longrightarrow Y empleo pleno

Variables aleatorizadas

Variables controladas

Variables perturbadoras

Pn

P1

P2

Sexo

Origen migratorio

Nivel socio económico

Nivel educativo



ASOCIACIÓN ESTADÍSTICA MULTIVARIADA

NECESIDAD DE UN MÉTODO QUE PERMITA

- **Explicar una relación descubriendo las conexiones causales existente entre las variables.**
- **Identificar condiciones bajo las cuales una relación tiene lugar.**
- **Identificar factores o condiciones independientes que operan sobre una misma variable.**
- **Evaluar la existencia de relaciones espurias entre variables.**

Modelo Lazarsfeld

Hipótesis

Relación Original

X nivel ed. \longrightarrow Y empleo pleno

	Hasta SI	SC y más
Resto de activos	XXXXXX	
Empleo pleno de derechos		XXXXXX

Relaciones Parciales

VARONES	Hasta SI	SC y más
Resto de activos		
Empleo pleno de derechos		XXXX

(XY) Z1

MUJERES	Hasta SI	SC y más
Resto de activos	XXXX	
Empleo pleno de derechos		

(XY) Z2

Análisis Multivariados

Relaciones marginales

	VARONES	MUJERES
Resto de activos		XXX
Empleo pleno de derechos	XXX	

XZ

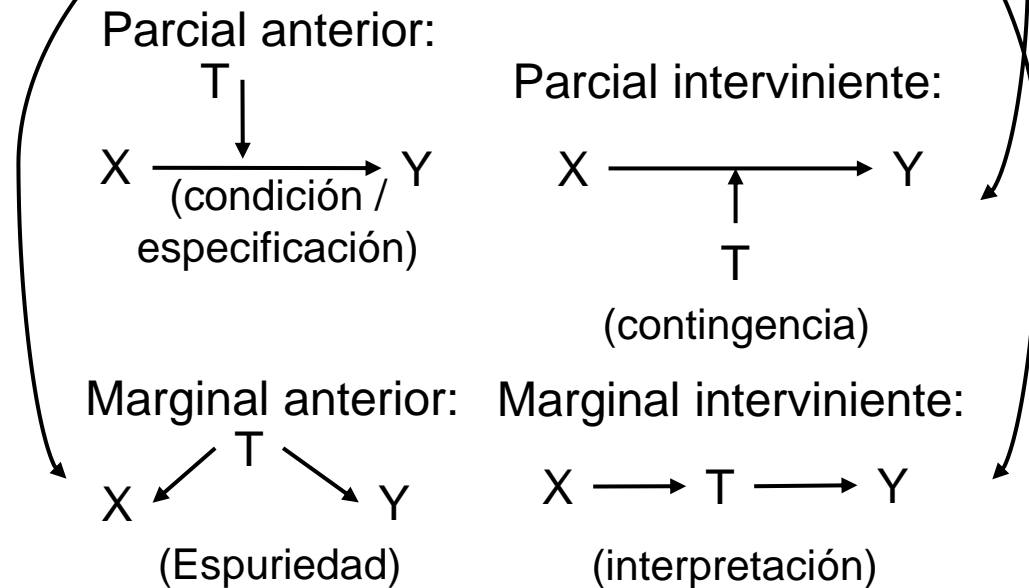
	VARONES	MUJERES
Hasta SI		XXX
SC y más	XXX	

ZY

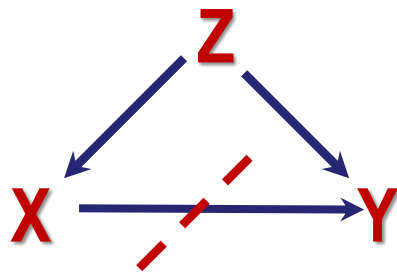
ASOCIACIÓN ESTADÍSTICA MULTIVARIADA EL MODELO LAZARSFELD

- Forma estadística
 - Por parciales
 - Por marginales
- Temporalidad
 - Antecedente
 - Interviniente

	Antecedente	Interviniente
Parcial	PA	PI
Marginal	MA	MI



Relación Marginal anterior



Para interpretar correlaciones equívocas.

Relaciones espurias.
“Explicación”

- Relaciones parciales tendientes a cero. El interés se centra en las relaciones marginales.
- Ej: X = Educación
Y = Calidad del Empleo
Z = Regiones Económicas

•

Relación Marginal Intermedia

- Hipótesis: El **estado civil** de las mujeres condiciona la **tasa de ausentismo** laboral, de modo tal que las mujeres casadas presentarán elevadas tasas de ausentismo. Se propone como variable de control la importancia de las **responsabilidades familiares**, ya que se supone que son éstas las que determinan el nivel de ausentismo y, a su vez, se encuentran estrechamente vinculadas al estado civil.

estado civil → **responsabilidades familiares** → **tasa de ausentismo**

Otro ejemplo:

Ocupado en una gran empresa → **premios por productividad** → **salario alto**

Relación Marginal Intermedia

“Interpretación”



- El ausentismo no se encuentra vinculado al estado civil,
- Relaciones parciales tendientes a cero. El interés se centra en las relaciones marginales.

Ej: X = Estado civil

Y = Ausentismo

Z = Responsabilidades familiares

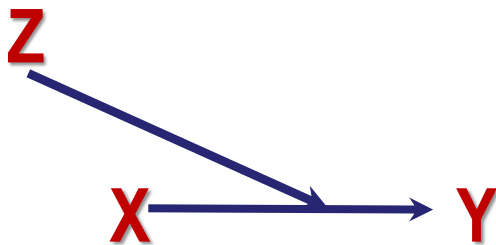
Relación Parcial anterior

- Hipótesis: El **nivel educativo** se vincula con las **oportunidades de empleo**. De manera tal que los jóvenes que tienen niveles de instrucción más altos tienen también mayores oportunidades de empleo.

Sin embargo, si se considera el origen social de los jóvenes la relación en cuestión se verifica para los de **nivel socioeconómico** más alto y se diluye para los de estratos más bajos.



Relación Parcial Anterior



Vble test o
control

como **“condición”**

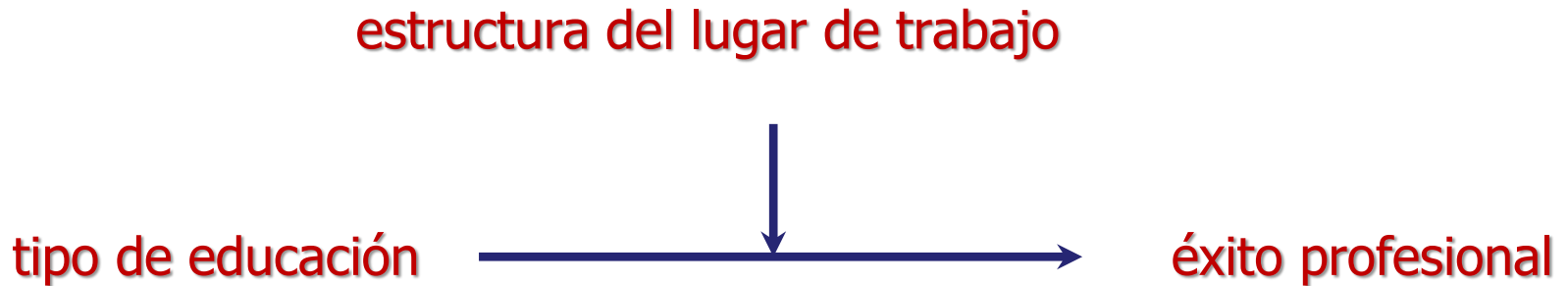
- Ej: X = nivel educativo
Y = oportunidades de empleo
Z = origen social

- Función de especificación.
- Una de las relaciones condicionales es mayor a la relación original.

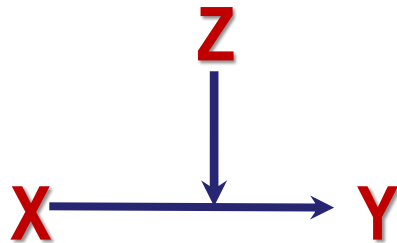
Relación Parcial Intermedia

- Hipótesis: El **tipo de educación** (liberal-autoritaria) se vincula con la **éxito profesional**. De manera tal que los jóvenes que tuvieron una educación liberal tienen altos niveles de éxito profesional.

Se propone como variable de control la **estructura del lugar de trabajo** (estructura liberal- estructura autoritaria) , ya que se supone que los jóvenes educados en una atmósfera liberal tienen menos éxito si el lugar de trabajo es autoritario que si es liberal.



Relación Parcial intermedia



Vble test o control

como
“contingencia”

•

X = Tipo de educación

Y = Éxito profesional

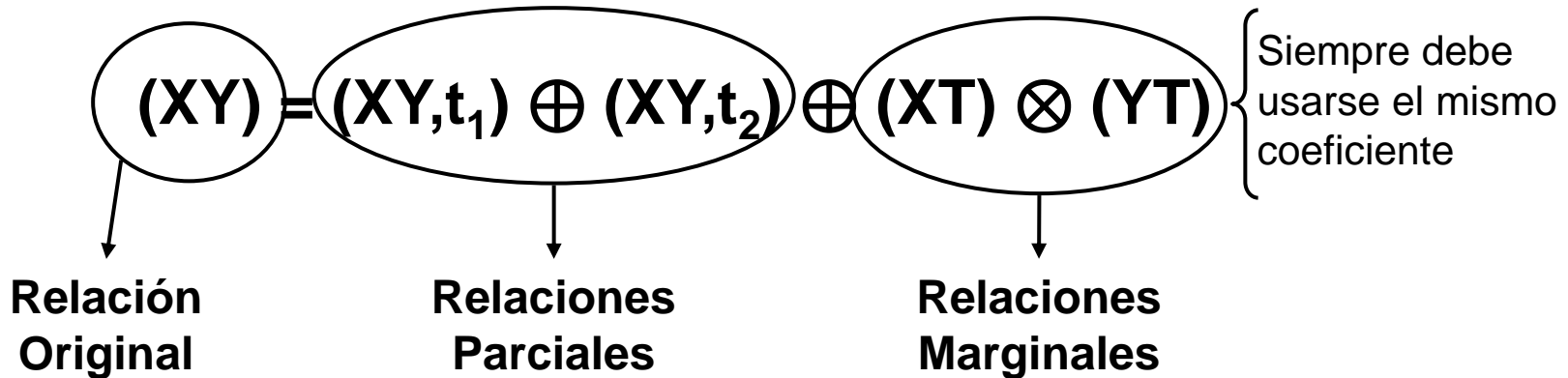
Z = Atmósfera profesional

•

Función de especificación.

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA EL MODELO LAZARSFELD

Ecuación de Covarianzas de Lazarsfeld



“Existe relación causal entre dos variables si, para cualquier factor de prueba antecedente, la relación entre esas variables no desaparece”

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MANERA DE EJEMPLO

Hipótesis Multivariada

- “Entre las personas en edad de alta participación económica (de 25 a 45 años), la tasa de actividad significativamente más elevada entre los varones que entre las mujeres, se explica por la intervención de condiciones familiares”

ESTADÍSTICOS DE ANÁLISIS BIVARIADO

A MODO DE EJEMPLO

PEA * Sexo Crosstabulation

			Sexo		Total
			Varón	Mujer	
PEA	Activo	% within PEA	56,5%	43,5%	100,0%
		% within Sexo	95,2%	65,6%	79,6%
		% of Total	45,0%	34,6%	79,6%
	Inactivo	% within PEA	11,1%	88,9%	100,0%
		% within Sexo	4,8%	34,4%	20,4%
		% of Total	2,3%	18,2%	20,4%
Total	% within PEA	47,2%	52,8%	100,0%	
	% within Sexo	100,0%	100,0%	100,0%	
	% of Total	47,2%	52,8%	100,0%	

PEA * Sexo Crosstabulation

			Sexo		Total	
			Varón	Mujer		
PEA	Activo	Count	6726	5122	11848	
		Expected Count	5588,8	6259,2	11848,0	
		Residual	1137,2	-1137,2		
		Std. Residual	15,2	-14,4		
	Inactivo	Count	401	2860	3261	
		Expected Count	1538,2	1722,8	3261,0	
		Residual	-1137,2	1137,2		
		Std. Residual	-29,0	27,4		
		Total	Count	7127	7982	15109
			Expected Count	7127,0	7982,0	15109,0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2029,509 ^b	1	,000
Likelihood Ratio	2258,729	1	,000
Linear-by-Linear Association	2029,375	1	,000
N of Valid Cases	15109		

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 1538,23.

Distribuciones para Tablas de Contingencia y Prueba de Hipótesis Ji cuadrado

ESTADÍSTICOS DE ANÁLISIS BIVARIADO

A MODO DE EJEMPLO

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b	Approx. Sig.
Nominal by Nominal	Phi	,367			,000
	Cramer's V	,367			,000
	Contingency Coefficient	,344			,000
Ordinal by Ordinal	Gamma	,807	,010	50,109	,000
Measure of Association	Kappa	,292	,006	45,050	,000
N of Valid Cases		15109			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for PEA (Activo / Inactivo)	9,366	8,385	10,462
For cohort Sexo = Varón	4,617	4,207	5,066
For cohort Sexo = Mujer	,493	,481	,505
N of Valid Cases		15109	

Coeficientes de Asociación Tablas

Directional Measures

			Value	Asymp. Std. Error	Approx. T
Nominal by Nominal	Lambda	Symmetric	,154	,010	14,843
		PEA Dependent	,000	,000	.
		Sexo Dependent	,225	,013	14,843
Goodman and Kruskal tau	PEA Dependent	Sexo Dependent	,134	,005	
		Sexo Dependent	,134	,004	
Uncertainty Coefficient	Symmetric	PEA Dependent	,123	,005	26,532
		PEA Dependent	,143	,005	26,532
		Sexo Dependent	,108	,004	26,532

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

INCORPORACIÓN DE UNA VARIABLE TEST O DE CONTROL

Presencia de menores en el hogar

PEA * Sexo * Presencia de menores en el hogar Crosstabulation

Presencia de menores en el hogar				Sexo		Total
				Varón	Mujer	
Al menos un niño de 12 años o menos	PEA Activo	abs.	1671878	1265488	2937366	
		% col	97,3%	59,8%	76,6%	
	Inactivo	abs.	47161	850325	897486	
		% col	2,7%	40,2%	23,4%	
Total		abs.	1719039	2115813	3834852	
		% col	100,0%	100,0%	100,0%	
Sin menores de 12 años	PEA Activo	abs.	984445	778050	1762495	
		% col	91,9%	77,7%	85,0%	
	Inactivo	abs.	86812	223076	309888	
		% col	8,1%	22,3%	15,0%	
Total		abs.	1071257	1001126	2072383	
		% col	100,0%	100,0%	100,0%	

Diferencia porcentual
37,5 p.p.

Diferencia porcentual
14,2 p.p.

Doble Diferencia
23,3 p.p.

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

Explicación por Parciales

Symmetric Measures

Presencia de menores		Value	Asymp. Std. Error ^a	Approx. T ^b	
Al menos un niño de 12 años o menos	Phi	,423			
	Cramer's V	,423			
	Contingency Coefficient	,389			
	Ordinal Measur	Gamma	,898	,008	51,013
	Kappa	,335	,007	42,825	
	N of Valid Cases	10273			
Sin menores de 12 años	Phi	,229			
	Cramer's V	,229			
	Contingency Coefficient	,223			
	Ordinal Measur	Gamma	,555	,028	15,985
	Kappa	,182	,011	15,891	
	N of Valid Cases	4836			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Chi-Square Tests

Presencia de menores en el hogar		Value	df	Asymp. Sig. (2-sided)
Al menos un niño de 12 años o menos	Pearson Chi-Square	741869,1 ^b	1	,000
	N of Valid Cases	3834852		
Sin menores de 12 años	Pearson Chi-Square	81808,083 ^c	1	,000
	N of Valid Cases	2072383		

a. Computed only for a 2x2 table

b. 0 cells (.0%) have expected count less than 5. The minimum expected count is 402313,7.

c. 0 cells (.0%) have expected count less than 5. The minimum expected count is 149700,6.

Risk Estimate

Presencia de menores en el hogar		Value	95% Confidence Interval	
			Lower	Upper
Al menos un niño de 12 años o menos	Odds Ratio for PEA (Activo / Inactivo)	18,687	15,770	22,144
	For cohort Sexo = Varón	8,758	7,504	10,222
	For cohort Sexo = Mujer	,469	,456	,482
	N of Valid Cases	10273		
Sin menores de 12 años	Odds Ratio for PEA (Activo / Inactivo)	3,490	2,974	4,095
	For cohort Sexo = Varón	2,045	1,834	2,280
	For cohort Sexo = Mujer	,586	,554	,619
	N of Valid Cases	4836		

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

Explicación por Marginales Población Activa por Presencia de menores en el hogar

PEA * Presencia de menores en el hogar Crosstabulation

		Presencia de menores en el hogar		Total
		Al menos un niño de 12 años o menos	Sin menores de 12 años	
PEA	Activo	2937366 76,6%	1762495 85,0%	4699861 79,6%
	Inactivo	897486 23,4%	309888 15,0%	1207374 20,4%
Total		3834852 100,0%	2072383 100,0%	5907235 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	51,785 ^b	1	,000
Likelihood Ratio	53,012	1	,000
Linear-by-Linear Association	51,782	1	,000
N of Valid Cases	15109		

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 1043,76.

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

Explicación por Marginales

Población Activa por Presencia de menores en el hogar

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b
Nominal by Nominal	Phi	-,059		
	Cramer's V	,059		
	Contingency Coefficient	,058		
Ordinal by Measure of	Gamma	-,157	,021	-7,445
	Kappa	-,056	,008	-7,196
N of Valid Cases		15109		

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for PEA (Activo / Inactivo)	,729	,668	,795
For cohort Al menos un niño de 12 años o menos	,909	,887	,932
For cohort Sin menores de 12 años	1,248	1,173	1,328
N of Valid Cases	15109		

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

Explicación por Marginales

Presencia de menores en el hogar por Sexo

Presencia de menores en el hogar * Sexo Crosstabulation

		Sexo		Total
		Varón	Mujer	
Presencia de menores en el hogar	Al menos un niño de 12 años o menos	1719039 61,6%	2115813 67,9%	3834852 64,9%
	Sin menores de 12 años	1071257 38,4%	1001126 32,1%	2072383 35,1%
Total		2790296 100,0%	3116939 100,0%	5907235 100,0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	86,242 ^b	1	,000
Likelihood Ratio	86,183	1	,000
Linear-by-Linear Association	86,236	1	,000
N of Valid Cases	15109		

a. Computed only for a 2x2 table

b. 0 cells (,0%) have expected count less than 5. The minimum expected count is 2281,17.

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

Explicación por Marginales Presencia de menores en el hogar por Sexo

Symmetric Measures

		Value	Asymp. Std. Error ^a	Approx. T ^b
Nominal by Nominal	Phi	-,076		
	Cramer's V	,076		
	Contingency Coefficient	,075		
Ordinal by Measure of	Gamma	-,161	,017	-9,282
	Kappa	-,069	,007	-9,287
N of Valid Cases		15109		

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Al menos un niño de 12 años / Sin menores de 12 años	,723	,675	,774
For cohort Sexo = Varón	,846	,818	,876
For cohort Sexo = Mujer	1,171	1,131	1,212
N of Valid Cases	15109		

- Not assuming the null hypothesis.
- Using the asymptotic standard error assuming the null hypothesis.

ASOCIACIÓN ESTADÍSTICA MULTIVARIADA A MODO DE EJEMPLO

ECUACIÓN DE COVARIANZAS DE LAZARSFELD

$$(XY) = (XY_{t_1}) \oplus (XY_{t_2}) \oplus (XT) \otimes (YT)$$

Hipótesis diagonal (PHI):

$$(XY) = (XY_{t_1}) \oplus (XY_{t_2}) \oplus (XT) \otimes (YT)$$
$$0,367 = 0,423 \oplus 0,299 \oplus -0,086 \otimes -0,059$$

Hipótesis rinconal (Gamma):

$$(XY) = (XY_{t_1}) \oplus (XY_{t_2}) \oplus (XT) \otimes (YT)$$
$$0,807 = 0,896 \oplus 0,555 \oplus -0,161 \otimes -0,157$$