

**FLACSO / Facultad Latinoamericana de Ciencias
Sociales**

**Maestría en Ciencias Políticas & Sociología
Metodología de la Investigación Social**

Unidad 3 - Clase 7 y 8

**Concepto de covarianza, relaciones bivariadas
Análisis por medio de cuadros de contingencia**

Agustín Salvia / Eduardo Donza

Concepto de cuadro bivariado

Variable x

Variable y

	x_1	x_2	x_3	
y_1				Marginal 1
y_2				Marginal 2
y_3				Marginal 3
	Subtotal 1	Subtotal 2	Subtotal 3	Total

Frecuencias condicionales

Usos de cuadros bivariados

- **Para describir a la población según características de dos variables**
- **Para contrastar hipótesis**

Cuadro bivariado

Sector de inserción de la población según sexo
GBA / EPH 2º trim. de 2010
-Cantidad de personas-

Tabla de contingencia Sector de Inserción * Sexo

Recuento

		Sexo		Total
		Varón	Mujer	
Sector de Inserción	Sector Público	353806	381228	735034
	Sector Formal	1686727	885237	2571964
	Sector Informal	1347934	1157440	2505374
Total		3388467	2423905	5812372

Cuadro bivariado para analizar datos

Sector de inserción de la población según sexo

GBA / EPH 2º trim. de 2010

-Según porcentaje respectivo-

Tabla de contingencia Sector de Inserción * Sexo

			Sexo		Total
			Varón	Mujer	
Sector de Inserción	Sector Público	% dentro de Sector de Inserción	48,1%	51,9%	100,0%
		% dentro de Sexo	10,4%	15,7%	12,6%
		% del total	6,1%	6,6%	12,6%
	Sector Formal	% dentro de Sector de Inserción	65,6%	34,4%	100,0%
		% dentro de Sexo	49,8%	36,5%	44,2%
		% del total	29,0%	15,2%	44,2%
	Sector Informal	% dentro de Sector de Inserción	53,8%	46,2%	100,0%
		% dentro de Sexo	39,8%	47,8%	43,1%
		% del total	23,2%	19,9%	43,1%
Total	% dentro de Sector de Inserción	58,3%	41,7%	100,0%	
	% dentro de Sexo	100,0%	100,0%	100,0%	
	% del total	58,3%	41,7%	100,0%	

Cuadro bivariado para verificar hipótesis / Covarianza

Concepto de covarianza / Contrastación de hipótesis

Relación entre variables

- Fuerza
- Sentido
- Forma
- Grado

Roles: $x \longrightarrow y$

$x \longleftarrow y$

Tipos de hipótesis

- Diagonales
- Rinconales

Posibles resultado al analizar la covarianza

- Intermedia
- Nula
- Total

Cuadro bivariado para verificar hipótesis

	Varón	Mujer
Sector transporte	90%	20%
Otros sectores	10%	80%
	100%	100%

Cuadro bivariado para verificar hipótesis

	Varón	Mujer
Sector transporte	90%	20%
Otros sectores	10%	80%
	100%	100%

$$d\% = 70\%$$

Relación
intermedia entre las
variables

Cuadro bivariado para verificar hipótesis

	Varón	Mujer
Sector transporte	60%	60%
Otros sectores	40%	40%
	100%	100%

$$d\% = 0\%$$

Independencia
estadística entre las
variables

Cuadro bivariado para verificar hipótesis

	Varón	Mujer
Sector transporte	100%	0%
Otros sectores	0%	100%
	100%	100%

$d\% = 100\%$

Relación perfecta
entre las variables

Cuadros bivariados para verificar hipótesis

Reglas para el procedimiento

1. Colocar la variable independiente en el cabezal del cuadro
2. Si son variables ordinales, verificar divergencia o convergencia de las categorías
3. Realizar porcentaje por columnas
4. Comparar por filas

Cuadro bivariado para verificar hipótesis

Condición de actividad por sexo

GBA / EPH 2º trim. de 2010

-En porcentaje-

Tabla de contingencia Condición de actividad * Sexo

% dentro de Sexo

		Sexo		Total
		Varón	Mujer	
Condición de actividad	ocupado	93,1%	90,3%	91,9%
	desocupado	6,9%	9,7%	8,1%
Total		100,0%	100,0%	100,0%

Pasos:

- Var. Independiente en el cabezal
- Orden de categorías
- Porcentajes por columnas
- Comparar por fila

d% = 2,8%

Cuadro bivariado para verificar hipótesis

Sector de inserción de la población según sexo
GBA / EPH 2º trim. de 2010
-En porcentaje-

Tabla de contingencia Sector de Inserción * Sexo

% dentro de Sexo

		Sexo		Total
		Varón	Mujer	
Sector de Inserción	Sector Público	10,4%	15,7%	12,6%
	Sector Formal	49,8%	36,5%	44,2%
	Sector Informal	39,8%	47,8%	43,1%
Total		100,0%	100,0%	100,0%

Pasos:

- Var. Independiente en el cabezal
- Orden de categorías
- Porcentajes por columnas
- Comparar por fila

La d% no es
medida resumen
de fuerza de la
relación

Asociación entre variables – Verificación de hipótesis

Procedimientos:

- Coeficientes de asociación
- Pruebas de independencia estadística

Asociación entre variables

Criterios de selección de coeficientes

Cantidad filas y columnas	Nivel de medición del cuadro	Hipótesis diagonales	Hipótesis rinconales
2 x 2	Nominal u ordinal	Phi (-1 a 1)	Gamma (o q de Yule) (-1 a 1)
Más de 2 x 2	Ordinal	Tau-b (-1 a 1)	Gamma (-1 a 1)
	Nominal	V de Cramer (0 a 1)	-----

Asociación entre variables

Prueba de independencia estadística de ji cuadrado:

- Frecuencias observadas
- Frecuencias esperadas bajo supuesto de independencia estadística
- Residuales
- Residuales estandarizados
- Suma de residuales

Asociación entre variables

Prueba de independencia estadística de ji cuadrado

Condición de actividad según edad

GBA / EPH 2° trim. de 2010

(Casos reponderados)

-Frecuencias observadas / esperadas y residuales-

Tabla de contingencia Condición de actividad * Edad

			Edad		Total
			De 18 a 24 años	De 25 a 65 años	
Condición de actividad	ocupado	Recuento	491	2942	3433
		Frecuencia esperada	553,0	2880,0	3433,0
		Residuo	-62,0	62,0	
	desocupado	Recuento	111	193	304
		Frecuencia esperada	49,0	255,0	304,0
		Residuo	62,0	-62,0	
Total		Recuento	602	3135	3737
		Frecuencia esperada	602,0	3135,0	3737,0

Asociación entre variables

Prueba de independencia estadística de ji cuadrado

$$\chi^2 = \sum \frac{(o_i - e_i)^2}{e_i}$$

χ^2 empírico

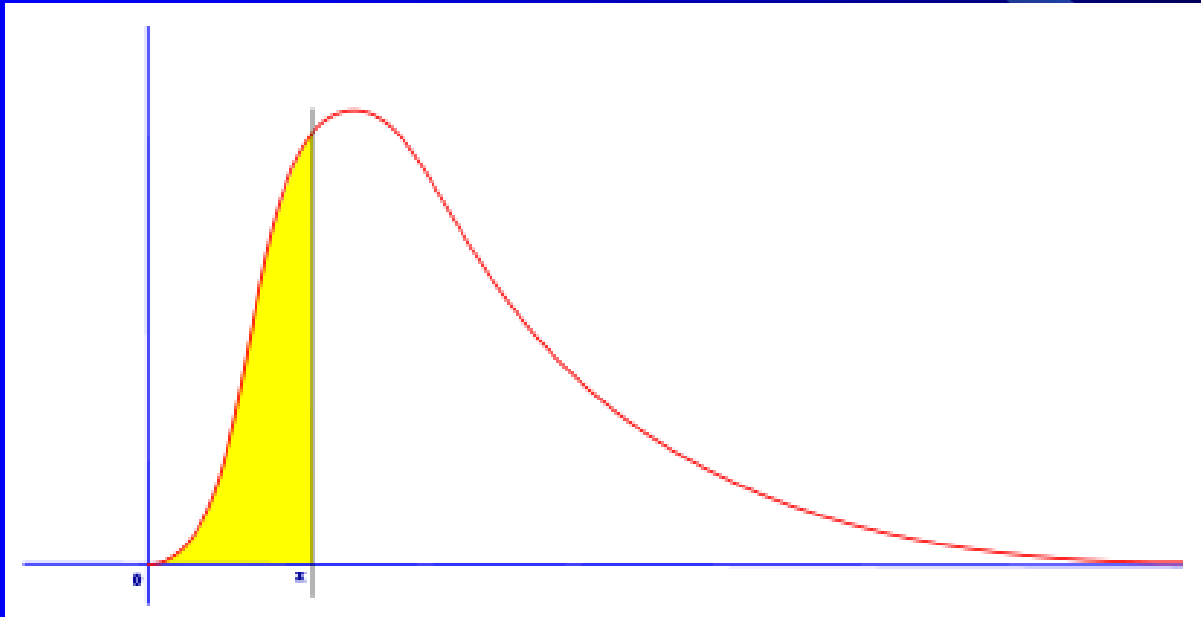
comparación con

χ^2 teórico o de tabla

Asociación entre variables

Prueba de independencia estadística de ji cuadrado

Distribución de χ^2



Efecto de grados de libertad en la forma de la gráfica.

Asociación entre variables

Hipótesis

- Hipótesis nula: entre las variables hay independencia estadística

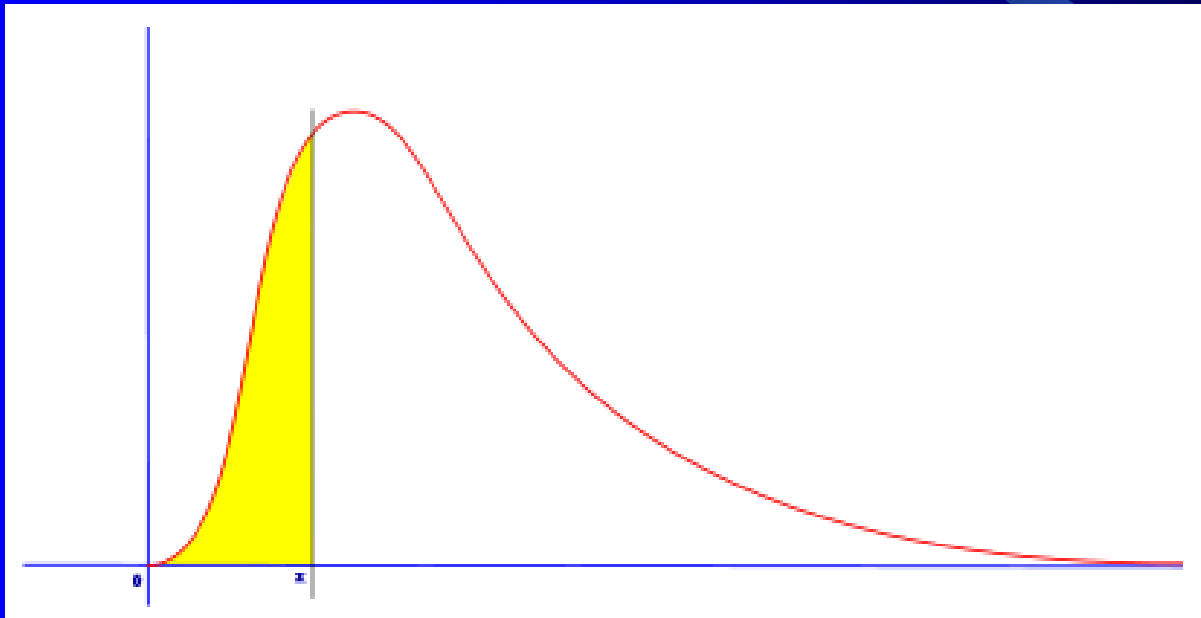
Lógicamente contrarias

- Hipótesis de trabajo: entre las variables hay relación

Asociación entre variables

Prueba de independencia estadística de ji cuadrado
Método en que se utiliza tabla de ji cuadrado

Distribución de χ^2



Si χ^2 empírico es mayor o igual que χ^2 teórico ... Rechazo H_0 ... Acepto H_t

Si χ^2 empírico es menor que χ^2 teórico ... No rechazo H_0 ... Rechazo H_t

Asociación entre variables

Condición de actividad según edad
GBA / EPH 2° trim. de 2010
(Casos reponderados)
-En porcentaje-

Tabla de contingencia Condición de actividad * Edad

% dentro de Edad

		Edad		Total
		De 18 a 24 años	De 25 a 65 años	
Condición de actividad	ocupado	81,6%	93,8%	91,9%
	desocupado	18,4%	6,2%	8,1%
Total		100,0%	100,0%	100,0%

Asociación entre variables

Prueba de independencia estadística de ji cuadrado Salida de SPSS para interpretar significancia

Condición de actividad según edad
GBA / EPH 2° trim. de 2010
(Casos reponderados)

χ^2 empírico

Significancia ó
Error de Tipo 1

Pruebas de chi-cuadrado

	Valor	gl	Sig. asintótica (bilateral)	Sig. exacta (bilateral)	Sig. exacta (unilateral)
Chi-cuadrado de Pearson	101,945 ^a	1	,000		
Corrección por continuidad ^b	100,308	1	,000		
Razón de verosimilitudes	82,646	1	,000		
Estadístico exacto de Fisher				,000	,000
Asociación lineal por lineal	101,917	1	,000		
N de casos válidos	3737				

a. 0 casillas (.0%) tienen una frecuencia esperada inferior a 5. La frecuencia mínima esperada es 48.97.

b. Calculado sólo para una tabla de 2x2.

Asociación entre variables

Coefficientes de asociación

Condición de actividad según edad
GBA / EPH 2º trim. de 2010
(Casos reponderados)

Medidas simétricas

		Valor	Error tip. asint. ^a	T aproximada ^b	Sig. aproximada
Nominal por nominal	Phi	,165			,000
	V de Cramer	,165			,000
Ordinal por ordinal	Tau-b de Kendall	,165	,021	-7,314	,000
	Gamma	,550	,045	-7,314	,000
N de casos válidos		3737			

a. Asumiendo la hipótesis alternativa.

b. Empleando el error típico asintótico basado en la hipótesis nula.

Elegir coeficiente adecuado e interpretar