

Clase 10

Tests estadísticos para el análisis bivariante



Le Sphinx

CAMPUS

El pack Sphinx dedicado a la enseñanza de métodos y técnicas para encuestas

Objetivos

Objetivos metodológicos

- Ver cómo las pruebas estadísticas para el análisis bivalente ayudarán a concluir los resultados de una encuesta de manera más precisa y fiable.
- Comprender el principio general de un test estadístico.
- Conocer las principales pruebas estadísticas útiles para el análisis bivalente de los resultados de la encuesta: chi-cuadrado, Fisher, coeficiente de correlación.
- Ser capaz de interpretar los resultados de las principales pruebas para el análisis bivariado y traducirlos en forma de comentarios explícitos, útiles para las conclusiones del estudio.

Objetivos Sphinx

- Cómo activar los diversos tests estadísticos en Sphinx Campus
- Ser capaz de interpretar correctamente los parámetros estadísticos

Tabla de contenidos

1. Las etapas sucesivas de un test
2. Los diferentes tipos de análisis cruzado y los tests correspondientes
3. Test chi cuadrado
4. Análisis de la varianza y test F de Fisher
5. Regresión y correlación

1. Las etapas sucesivas de un test

- En el caso de los análisis bivariantes, se pueden utilizar pruebas estadísticas para validar la existencia o no de una **relación** o correspondencia **válida** entre dos variables o dos categorías.
- Un análisis cruzado (o bivariado) requiere un test de tres pasos:
 1. **¿Hay** una relación significativa entre las dos variables?
→ Esto nos lleva a establecer que los resultados nos permiten establecer probabilidades que son fiables.
 2. En caso afirmativo, ¿esta relación es moderada, **fuerte** o muy fuerte?
→ ¿La probabilidad de nuestra conclusión es **moderada** o fuerte?
 3. En caso afirmativo, ¿cómo se **traduce** esto en hechos?
→ Este es el paso más importante porque se trata de indicar qué características se asocian: edad/satisfacción; hotel/residencia, etc. Es la información más **útil** que produce este tipo de análisis.

2. Los diferentes tipos de análisis cruzado y los tests correspondientes

- Esta serie de tratamientos depende de la naturaleza de las preguntas -las variables- que se cruzan.
- Se aplican 3 métodos de procesamiento cruzado según el tipo de respuestas registradas:
 - cerradas (variables nominales)
 - abiertas numéricas

Caso	Ejemplo	Tratamiento acorde	Test estadístico que valida la relación
Cerrada x cerrada	<i>Sexo x Deporte practicado</i>	Cruce	Chi cuadrado
Cerrada x numérica	<i>Sexo x Objetivo salarial</i>	Análisis de varianza	Test de Fisher
Numérica x numérica	<i>Edad x Objetivo salarial</i>	Regresión lineal	Coeficiente de correlación

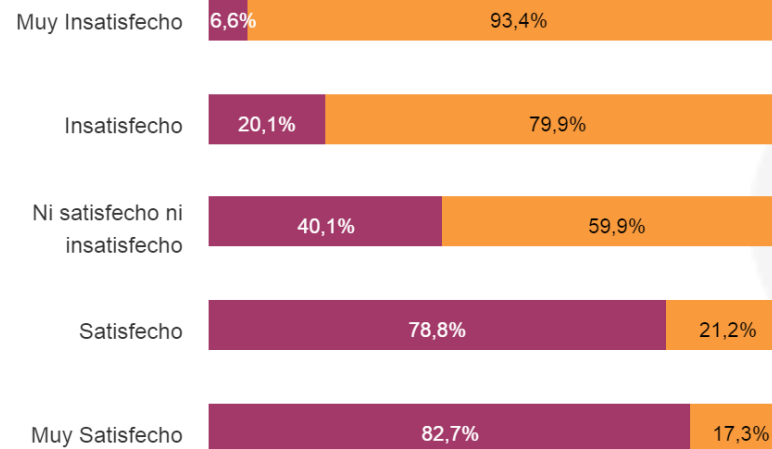
3. Test chi cuadrado

3.1 Principios generales

- Se trata de una prueba de independencia, para ver si la distribución de los términos de la variable explicativa es una función de la distribución de los términos de la variable explicada.
- El punto de partida es, por lo tanto, una tabla cruzada, o tabla de contingencia.

Cruce: La calidad de la relación con su asesor / Sexo

RELACION_ASESOR	SEXO		TOTAL
	HOMBRE	MUJER	
Muy Insatisfecho	6	85	91
Insatisfecho	30	119	149
Ni satisfecho ni insatisfecho	77	115	192
Satisfecho	320	86	406
Muy Satisfecho	210	44	254
TOTAL	643	449	1.092



Este ejemplo es una encuesta de satisfacción de un banco. Se han cruzado el sexo con la satisfacción sobre la relación con el asesor. Queremos saber si la satisfacción varía según el sexo.

Con la tabla, los efectivos o los gráficos no podemos concluir «científicamente» sin tener disponible el test estadístico.

3. Test chi cuadrado

- El test de chi cuadrado es una comparación entre:
 - Valor real (los observados en el estudio)
 - Valor teórico (los que deberían haberse obtenido en ausencia de cualquier relación entre las 2 variables).

Valor real

Cruce: ¿Cada cuánto ves la televisión? / Sexo

FRÉQUENCE	SEXE		TOTAL
	MUJER	HOMBRE	
Varias horas al día	105	77	182
Una hora al día	48	71	119
Varias horas a la semana	45	61	106
Más de 5 horas al día	4	4	8
Varias veces al mes	31	30	61
Nunca	3	4	7
TOTAL	236	247	483

Valor teórico

Cruce: ¿Cada cuánto ves la televisión? / Sexo

FRÉQUENCE	SEXE	
	MUJER	HOMBRE
Varias horas al día	88,928	93,072
Una hora al día	58,145	60,855
Varias horas a la semana	51,793	54,207
Más de 5 horas al día	3,909	4,091
Varias veces al mes	29,805	31,195
Nunca	3,420	3,580

Cálculo del valor teórico: 48,86% de mujeres ($n=236$) en la muestra ($n=483$). Si no existiera una relación entre el género y la frecuencia de visión, habría un 48,86% de mujeres en cada categoría de frecuencia.

Por ejemplo, 48,86% de las mujeres ven la televisión varias horas al día, el 48,86% de 182 = 88,9. Lo mismo para los 247 hombres, el 51,14% de la muestra; teóricamente deberían representar el 51,14% de los que ven una hora de televisión al día, es decir, el 51,14% de 119 = 60,9, y así sucesivamente.

3. Test chi cuadrado

- Test chi cuadrado = suma de las diferencias entre el valor teórico y el valor real.
- Si las diferencias son muy pequeñas, estamos cerca del caso en que no hay relación entre las dos variables. Cuanto mayores sean las diferencias, más fuerte será la relación entre las dos variables.
- Fórmula de cálculo:

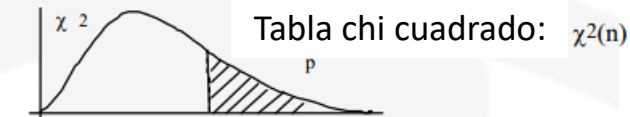
$$X^2_{calculado} = \frac{\sum (y_{efectivo\ teórico} - efectivo\ real)^2}{Efectivo\ teórico}$$

En nuestro ejemplo, esta fórmula de cálculo se aplica 12 veces (2 columnas * 6 filas).

Chi cuadrado calculado: $(88,9 - 105)^2/88,9 + (58,1 - 48)^2/58,1 + \dots + (3,6 - 4)^2/3,6 = 2,92 + 1,76 + \dots + 0,04 = \mathbf{11,1}$

3. 2 Primera etapa de conclusión: validar o no la relación

- Para contrastar si hay diferencias significativas, debemos comparar este chi cuadrado calculado con un valor de referencia de corte (obtenido según las leyes probabilísticas), es decir, el chi cuadrado teórico.
- Esto se obtiene en una tabla según dos criterios (tabla disponible, por ejemplo, en [este enlace](#)):
 - p valor = probabilidad de que un valor estadístico calculado sea posible dada una hipótesis nula cierta. Por defecto es 0,05 que representa solamente el 5 % de riesgo de llegar a una conclusión errónea
 - Grados de libertad: (número de líneas – 1) x (número de columnas – 1). En nuestro caso, los grados de libertad son: (6 – 1) x (2 – 1) = 5.
 - Aquí, el valor de chi cuadrado teórico, p = 0,05 y gdl = 5, es de **11,070**.



n P	0.10	0.05	0.02	0.01
1	2,706	3,841	5,412	6,635
2	4,605	5,991	7,824	9,210
3	6,251	7,815	9,837	11,341
4	7,779	9,488	11,668	13,277
5	9,236	11,070	13,388	15,086
6	10,645	12,592	15,033	16,812
7	12,017	14,067	16,622	18,475
8	13,362	15,507	18,168	20,090
9	14,684	16,919	19,679	21,666
10	15,987	18,307	21,161	23,209

Si el chi cuadrado calculado es mayor que el chi cuadrado teórico, se puede concluir que existe una relación entre las dos variables (en el nivel de error correspondiente). En nuestro ejemplo, tenemos $11.1 > 11.070$. Entonces, con una tasa de error del 5%, podemos decir que la frecuencia de visión varía significativamente según el género.

3. 3 Segunda etapa de conclusión: establecer la fuerza de relación o su fiabilidad

- Con un chi cuadrado teórico del 5%, tenemos una conclusión fiable del 95%. Aumentar el nivel de precisión permite determinar (reduciendo p valor) si nuestra conclusión sigue siendo válida y establecer así su nivel de fiabilidad, es decir, su fuerza.
- La fuerza de la relación está más garantizada cuando el umbral de confianza está cerca del 100%. La siguiente tabla ilustra cómo el chi-cuadrado teórico aumenta con el incremento del nivel de fiabilidad; así podemos ver desde qué nivel la relación sigue siendo válida, siempre a 5 grados de libertad.

p = nivel de confianza	Fiabilidad de la conclusión	Chi cuadrado teórico	Chi cuadrado calculado	Conclusión
0,05	95 %	11,07	11,1	Relación validada
0,025	97,5 %	12,83		Relación NO validada
0,01	99 %	15,08		Relación NO validada
0,005	99,5 %	16,75		Relación NO validada
0,001	99,9 %	16,3		Relación NO validada

La tabla muestra que la relación sigue siendo válida en el 95% (o $p=0,05$); sin embargo, ya no es significativa en el nivel del 97,5% (o $p=0,0025$), en el que el chi-cuadrado calculado es menor que el chi-cuadrado teórico (12,83).

3. 4 Tercera etapa de conclusión: detalles de la relación

- A continuación, es necesario examinar los detalles de la relación para establecer cómo se traduce en la práctica. Para ello, nos referimos a los chi cuadrados por celda (o chi cuadrados por pareja), calculados para cada celda de la tabla, considerando cada combinación.

Ejemplo: Tercera etapa de conclusión: detalles de la relación

Cruce: Sexo / ¿Cada cuánto ves la televisión?

¿CADA CUÁNTO VES LA TELEVISIÓN?	SEXO	
	MUJER	HOMBRE
Varias horas al día	88,93	93,07
Una hora al día	58,14	60,86

La relación es muy significativa. Valor de $P = < 0,01$; $\chi^2 = 9,91$; Grados de libertad = 2.

Cruce: Sexo / ¿Cada cuánto ves la televisión?

¿CADA CUÁNTO VES LA TELEVISIÓN?	SEXO								
	MUJER			HOMBRE			TOTAL		
	N	%	TEST ESTADÍSTICO...	N	%	TEST ESTADÍSTICO...	N	%	TEST ESTADÍSTICO...
Varias horas al día	105	57,69%	+ MS	77	42,31%	- MS	182	100,00%	
Una hora al día	48	40,34%	- S	71	59,66%	+ S	119	100,00%	
Varias horas a la semana	45	42,45%	- PS	61	57,55%	+ PS	106	100,00%	
Más de 5 horas al día	4	50,00%		4	50,00%		8	100,00%	
Varias veces al mes	31	50,82%		30	49,18%		61	100,00%	
Nunca	3	42,86%		4	57,14%		7	100,00%	
TOTAL	200	100,00%		182	100,00%		382	100,00%	

■ Elementos infrarrepresentados ■ Elementos sobrerrepresentados

Por ejemplo, para "mujer" y "varias horas al día", hay un cálculo por pares y una tabla simplificada con 2x2 modalidades para el género y la frecuencia. Se calcula un chi-cuadrado en este único cruce de 2 x 2. En nuestro ejemplo es 9,91 contra un chi-cuadrado teórico de 3,84 (a $p=0,01$). Luego se valida la correspondencia entre las modalidades "mujer" y "varias horas al día". Incluso a $p=0,01$; χ^2 teórico: 6,64.

En las tabulaciones cruzadas de le Sphinx se pueden identificar los pares de modalidades para los que la correspondencia es significativa, en azul para las sobrerrepresentaciones y en rojo para las infrarrepresentaciones. MS para muy significativo (en $p=0,01$) o S para significativo (en $p=0,05$). Aquí, las mujeres tienden a ver la televisión durante varias horas al día y los hombres durante una hora.

4. Análisis de la varianza y el test F de Fisher

4.1 Principios generales del análisis

- El análisis de la varianza se ocupa de las variaciones de una variable numérica (explicada) según las categorías de una variable explicativa nominal (o pregunta cerrada). Consiste en descomponer la varianza (todas las variaciones) de la variable numérica en dos partes:
 - **varianza «intergrupo»** que mide las diferencias entre las categorías de la pregunta cerrada; esta variación se debe a la variable explicativa, conocida como el "factor" en el vocabulario del análisis de la varianza.
 - **varianza «intragrupo»** que establece las diferencias dentro de cada categoría de la pregunta cerrada. Esta variación se debe a otras cosas, llamadas el "residual", que se compone de todos los factores distintos de la variable explicativa.

Ejemplo

María tiene un Grado y su nota es 7. El promedio de la clase es de 5,5. El promedio de la clase para los estudiantes de grado es de 6,5. En este caso, descomponemos la varianza de María según el promedio (+3) en dos partes:

- la varianza según el grado: (+2), porque el promedio de grados es de 2 puntos superiores al conjunto;
- la varianza residual, según el resto: (+1), porque María tiene un punto más que la media de los graduados

María tiene un Grado
Nota de María: 7
Nota media de clase: 5,5
Nota media de los graduados en clase: 6,5



4.2 Tabla de análisis de la varianza

- La tabla de análisis de la varianza permite calcular las dos varianzas: la debida al factor y la conocida como «residual»; y de compararlas con la F de Fisher

Varianza	Suma de los cuadrados	Grados de libertad	Cuadrado medio	F de Fisher
Debido al factor (variable explicativa)	SCf	$k - 1$	$CMf = SCf / k - 1$	CMf / CMr
Residual (debido al resto)	SCr	$n - k$	$CMr = SCr / n - k$	
k = número de modalidades del factor (variable explicativa); n = número de observaciones				

Fórmulas de cálculo necesarias:

$$SCf = S ((S x.j)^2 / n_j) - (S x_{ij})^2 / n$$

$S x.j$ = suma de la j-ésima modalidad (columna)

$$SCr = S (x_{ij})^2 - S ((S x.j)^2 / n_j)$$

4.3 Test F de Fisher (1/2)

- Al final de la tabla del análisis de varianza, obtenemos una F de Fisher "calculada", que es una especie de relación entre la varianza debida al factor y la varianza residual.
- Siguiendo el mismo principio que para las demás pruebas estadísticas, el valor obtenido debe compararse con un valor de referencia que se debe superar (obtenido según las leyes probabilísticas), que es la F teórica
- La F teórica se encuentra en la tabla estadística correspondiente (prueba de Fisher ver [este enlace](#)). La F teórica se obtiene en función de la tasa de error (normalmente 0,05); de V1 o $df1 = k-1$ y V2 o $df2 = n-k$

4.3 Test F de Fisher (2/2)

Para el análisis de la varianza y del test F, la conclusión también se extrae en tres pasos:

1. ¿Hay una relación entre las dos variables? Sí, si la F calculada es mayor que la F teórica -para una tasa de error del 5%.
2. Si la respuesta es afirmativa, ¿esta relación es media, fuerte o muy fuerte? Si la F calculada es más alta por defecto cuando $p = 5\%$, ¿es siempre más alta para $p = 1\%$, y entonces para $p = 0,5\%$, y para $p = 0,1\%$? Si la relación se valida siempre cuando p está cerca de 0, es más fiable y, por lo tanto, se puede calificar como "fuerte" o "muy fuerte".
3. Si es así, ¿cómo se traduce a la realidad? ¿Qué categorías de grados tendrían promedios más altos o más bajos que el promedio? Luego miramos los promedios estadísticamente diferentes, de acuerdo con la prueba t de Student (comparación de promedios)

4.4 Los resultados del software

- Se muestra el resultado del test de Fisher, así como la tasa de error.
- Los promedios estadísticamente diferentes se indican con subrayados azules (superior) y rosas (inferior)

Cruce: Selecciona el canal que más sueles ver / Da una nota general al programa desde 0 (muy mala) a 10 (excelente)

SELECCIONA EL CANAL QUE MÁS SUELES VER	DA UNA NOTA GENERAL AL PROGRAMA DESDE 0 (MUY MALA) A 10 (EXCELENTE)		
	MEDIA	DESVIACIÓN TÍPICA	N
▲ Antena3	<u>5.91</u>	1,79	103
Antena 3	5,91	1,79	103
La Sexta	<u>8.25</u>	1,20	51
Telecinco	6,57	1,24	115
▲ Otros	<u>7.34</u>	1,17	132
La 1	6,75	1,23	36
La 2	7,38	0,93	58
Mega	8,33	0,66	21
FDF			0
Neox			0
Divinity	7,33	1,23	15
Gol TV	6,50	3,54	2
Cuatro	6,60	1,22	77
TOTAL	6.82	1.52	478

■ Elementos infrarrepresentados ■ Elementos sobrerrepresentados

La relación es muy significativa. p-value= < 0,01 ; Fisher= 32,26.
Varianza inter= 59,22. Varianza intra= 1,84.

En el ejemplo que se presenta a continuación, los valores de satisfacción varían significativamente dependiendo de la cadena. Más específicamente, La Sexta destaca positivamente del promedio, mientras que Antena3 es mucho menos apreciado.

5. Regresión y correlación

- Con esta técnica, estudiamos los vínculos entre dos variables numéricas. Determinamos si una variable explicada **y** varía en función de otra variable explicativa **x**.

5. 1 Recta de regresión

Buscamos una recta de coordenadas $y = \mathbf{ax} + \mathbf{b}$, que describe cómo evoluciona y en función de x.

Fórmulas necesarias para el cálculo:

$$\mathbf{a} = \text{Cov}(\mathbf{x}, \mathbf{y}) / \text{Var}(\mathbf{x})$$

- $\mathbf{a} = 1/n (\sum x_i y_i) - \bar{x} \bar{y} / 1/n (\sum x_i^2) - \bar{x}^2$

$$\mathbf{b} = \bar{y} - \mathbf{a} \bar{x}$$

- a es la pendiente y b es la constante (la ordenada en el origen)

5. 2 Coeficiente de correlación

- El coeficiente de correlación r indica la fuerza y la dirección de la relación y oscila entre -1 y $+1$. Cerca de 1 (o -1) indica una correlación muy fuerte; cerca de 0 , corresponde a una correlación cero.
- $r = \text{Cov}(x,y) / \text{desviación estándar}(x) \times \text{desviación estándar}(y)$
- Para estimar si la correlación es significativa, comprobamos que la línea de regresión tiene sentido y que podemos usar sus coordenadas para predecir bien y en función de x .
- Comprobamos que el coeficiente tiene $\neq 0$. Conocemos su valor y desviación estándar. Gracias a la desviación estándar de a , podemos estimar p , la probabilidad de equivocarse considerando que es diferente de 0 . Esto equivale a estimar la probabilidad de error de una correlación significativa entre x e y .

5.3 Concluir con regresión y correlación

1. ¿El coeficiente de correlación es significativamente diferente de 0? Si es así, hay una relación entre x e y
2. Si es así, ¿cuál es la fuerza de la correlación? En ciencias sociales y humanidades, $|r| < 0,25$ es débil; $|r|$ entre 0,25 y 0,5, medio; $|r|$ entre 0,5 y 0,75, fuerte; $|r|$ mayor de 0,75, muy fuerte.
3. La línea de regresión muestra cómo se comporta y en función de x . La fiabilidad del pronóstico depende de la fuerza del coeficiente de correlación.

Ejemplo

Regresión lineal múltiple

Ecuación del modelo $IC = 18.2 + 0.5 \times Edad$

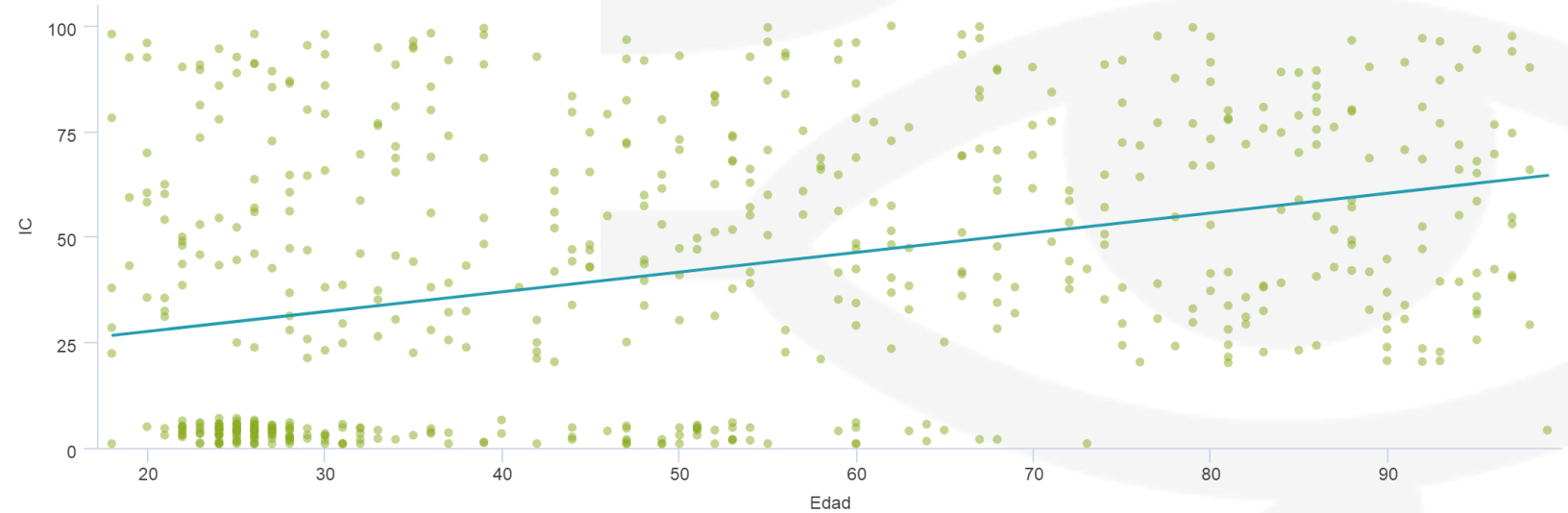
El modelo reproduce 12% de la varianza de la variable a explicar.

Coefficiente de correlación múltiple: $R = 0.3$

Coefficiente de Fisher: $F = 83.2$

p-valor de R : $p(R) = 0$

p-valor de F : $p(F) = 0$



La regresión lineal consiste en producir la ecuación que permite calcular de la manera más exacta el valor de la variable a explicar. Se puede así determinar:

1/ las variables que tienen más influencia (Ninguno al nivel del 10%).

2/ la contribución de las otras variables. Mide la influencia sobre la variable explicativa.

En este ejemplo:

Intención de compra = $18,2 + 0,5 \times Edad$

La recta de regresión indica que la intención de compra aumenta con la edad. En la base (constante), la intención de compra es del 25% y se añade un 5% cada 10 años.

El coeficiente de correlación es de 0,3. El pronóstico indica una pequeña correlación (pero diferente de 0) y por lo tanto no será muy fiable.

Conceptos clave

- **Coeficiente de correlación:** indica cómo dos variables numéricas se mueven juntas, una en relación con la otra. Si está cerca de 0, indica casi ninguna correlación; si está cerca de 1, correlación perfecta. Una correlación negativa indica que cuando una variable aumenta, la otra disminuye.
- **Margen de error:** medida de error de una prueba estadística -indica la probabilidad de una conclusión errónea de la prueba. Las pruebas se consideran generalmente concluyentes por debajo de un error del 5% o $p < 0,05$.
- **Significado:** éxito de una prueba estadística, teniendo en cuenta todos sus parámetros, incluido el margen de error establecido (por defecto $p < 0,05$).
- **Prueba de chi cuadrado:** una prueba estadística para estudiar la dependencia entre dos variables categóricas. Se trata de estudiar las discrepancias entre la población encuestada y la población teórica, que se habrían observado en una situación de completa independencia.
- **Test F de Fisher:** prueba estadística para estudiar la dependencia entre una variable categórica y una variable numérica. Se trata de comparar la variación entre categorías (inter) y la variación dentro de las categorías (intra).
- **Test estadístico:** un procedimiento computacional para verificar la veracidad de un fenómeno. A menudo se trata de evaluar la brecha entre lo que se observa y una situación de referencia, en la que nuestra hipótesis sería totalmente rechazada.

Ejercicios

Ejercicio 10a: canales y patrones de visualización

De la encuesta "Televisión", determinar si los motivos de visionado (Q3) varían significativamente dependiendo del canal favorito (Q7). Para este análisis, agrupa los canales citados menos de 20 veces (TSR y LCI).

Ejercicio 10b: frecuencia y opiniones

En la encuesta "Televisión", ¿las opiniones sobre el enriquecimiento (Q22) y el interés cultural de los programas (Q14) varían de acuerdo con la frecuencia de visualización (Q1)? Usa las escalas como variables numéricas. Para este análisis, agrupa las frecuencias extremas (nunca y más de 5 horas al día con categorías cercanas) ya que registran pocas respuestas.

Ejercicio 10c: interés cultural y opinión global

¿Existe un vínculo entre las opiniones de los espectadores sobre el interés cultural de los programas y su satisfacción general con el canal? ¿Y entre la simpatía de los presentadores y la satisfacción general? ¿Qué vínculo es el más fuerte? ¿Qué significa? Usa estas escalas como variables numéricas.

Más recursos

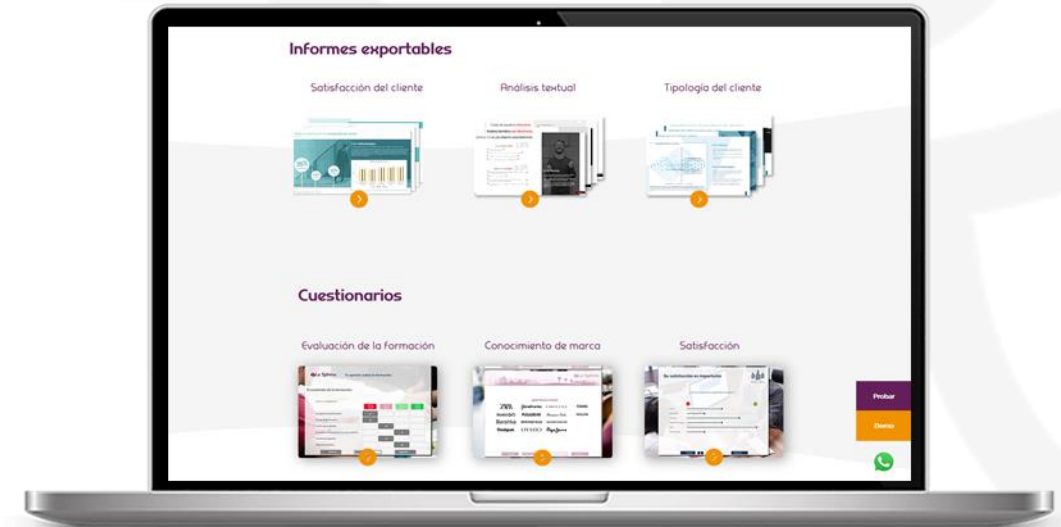
Tutoriales

www.lesphinx.es/videos



Ejemplos

www.lesphinx.es/ejemplos



Descargar una licencia

www.lesphinx.es/probar-gratis

