

L U M B R E



Chi-cuadrada y pruebas sin campana

Materia: Ciencia de Datos e IA

11/10/2026

Antes de empezar

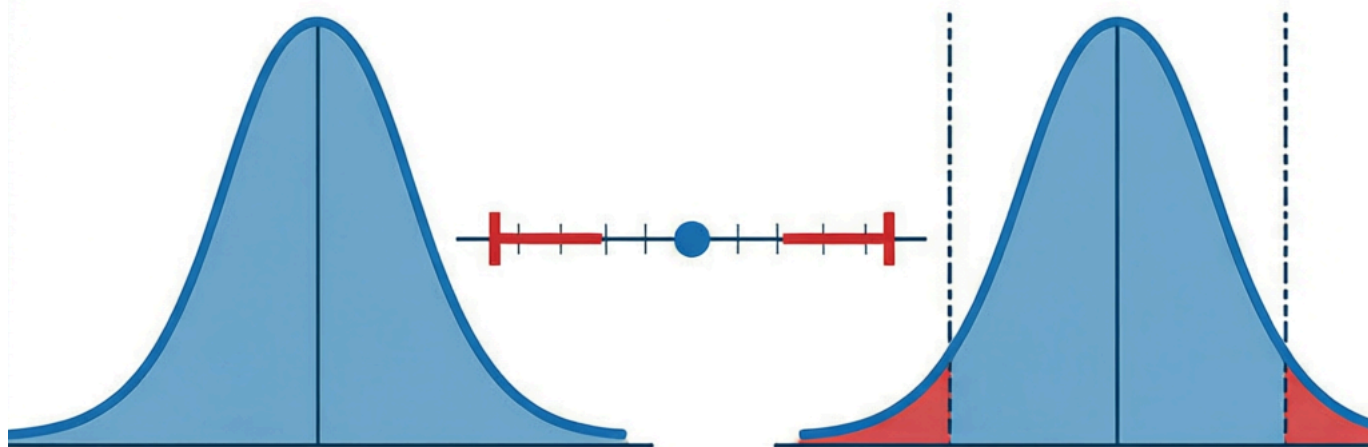
Dos contrastes chi-cuadrada resueltos paso a mano (dado cargado y tabla 2x2 medica), condiciones y Yates, mapa de la familia no parametrica con ejemplo de signo resuelto y autoexamen.

QUÉ APRENDERÁS

- Ejecutar contrastes ji-cuadrada de bondad de ajuste e independencia con aritmetica completa, verificar condiciones y elegir el test no parametrico adecuado.

Chi-cuadrada y pruebas sin campana: cuando los datos no son numericos

Los test z y t que ya dominas piden medidas y normalidad. Media poblacion habla en CATEGORIAS: compr/no compro, turno dia/noche, variante A/B/C con tasas de conversion. Y a veces la medida existe pero la distribucion es una barbaridad asimetrica. Ahi entran la ji-cuadrada —la escala entre lo esperado y lo observado— y la familia no parametrica, que trabaja con rangos y signos. Este objeto resuelve de punta a cola los dos test chi-cuadrada tipicos y mapa cuando usar cada alternativa.



La ji-cuadrada mide sorpresas: cuanto se desvia lo observado de lo esperado, en unidades de la variabilidad típica.

La idea en una línea

$\chi^2 = \sum (\text{observado} - \text{esperado})^2 / \text{esperado}$, sumado sobre todas las casillas. Si el azar explicara todo, χ^2 sería pequeño; valores grandes contra la tabla con sus grados de libertad dicen: el azar no alcanza.

Bondad de ajuste: ¿truca la ruleta?

Problema 1. Un dado se lanza 60 veces: caras [8, 12, 13, 7, 12, 8]. ¿Esta cargado? $\alpha = 0.05$.

ACTIVIDAD INTERACTIVA DISPONIBLE EN
LUMBRE

1. Contraste de bondad de ajuste



Test de independencia: ¿el tratamiento influye?

Problema 2. Ensayo observacional: 100 pacientes, tabla 2x2 entre tratamiento (T, 50) y control (C, 50) contra recuperación: T+rec 40, T+no 10, C+rec 25, C+no 25. Contrasta independencia, $\alpha = 0.05$.

ACTIVIDAD INTERACTIVA DISPONIBLE EN
LUMBRE

2. Chi-cuadrada de contingencia



Condiciones que nadie te recuerda

Esperados por casilla de 5 o mas (si no: agrupa categorias o usa el test exacto de Fisher), observaciones independientes (una fila = un sujeto, no medidas repetidas) y datos categoricos puros. En tablas 2x2 con esperados justos se suele aplicar la correccion de Yates (resta 0.5 a cada desvio): mas conservadora, mismo procedimiento.

La familia no parametrica: rangos y signos

- Signo (antes-despues pareado): ignora magnitudes, cuenta hacia donde virio cada sujeto. Ejemplo: 7 de 8 mejoras tras un taller; bajo H_0 cada signo es una moneda: $P(7 \text{ u } 8 \text{ de } 8) = (8+1)/256 = 0.035$: significativo.
- Wilcoxon signed-rank: el signo con tamaño — rankea los cambios absolutos y suma rangos de un signo. Mas potencia que el signo con datos ordenables.
- Mann-Whitney U: el t de dos muestras sin normalidad: compara RANKS, no medias; robusto a outliers y colas pesadas (ingresos, tiempos de respuesta).
- Kruskal-Wallis: Mann-Whitney para 3 o mas grupos (turnos, regiones). Su post-hoc: Dunn.
- Spearman: Pearson sobre rangos: captura monotonias aunque no linealidad; con empates y outliers se portan bien.

Chi en machine learning

- Selección de características categóricas: χ^2 entre cada feature y la etiqueta (sklearn FeatureSelection chi2) — barato y eficaz pre-modelo.
- Valida tu A/B: la conversión es binaria, el test de proporción z ES el primo hermano de chi-cuadrada (de hecho $z^2 = \chi^2$ para 2x2).
- Goodness-of-fit de modelos: bondad de ajuste de un modelo de frecuencia contra datos agrupados.

Que test sin normalidad uso

- **Categorías y rangos**
 - Una categoría vs teoría
 - Bondad de ajuste χ^2
 - Dos categorías cruzadas
 - Independencia χ^2
 - Fisher si esperados pequeños
 - Pareados no normales
 - Signo: solo dirección
 - Wilcoxon: rango con tamaño
 - Dos grupos o más
 - Mann-Whitney
 - Kruskal-Wallis

El mapa mental interactivo es interactivo en la versión online.



Autoexamen cronometrado (20 minutos)

3. Del problema 1: calcula el estadístico χ^2 del dado (una cifra decimal).

Respuesta:

4. En la tabla 2x2 del problema 2, ¿cuántos grados de libertad tiene el contraste de independencia?

- ☐ 1
- ☐ 2
- ☐ 3
- ☐ 4

5. La chi-cuadrada acepta datos continuos crudos sin agrupar en categorias.☐ Verdadero☐ Falso**6. Del problema 2, calcula el porcentaje de recuperacion del brazo tratamiento (40 de 50).**Respuesta: **7. El test _____ trabaja solo con la DIRECCION del cambio en datos pareados, ignorando magnitudes.****8. Empareja situacion con test.**

8.	Dado sospechoso de carga	-----
8.	Turno vs incidencias	-----
8.	Satisfaccion antes-despues, n pequena y raro	-----
8.	Tres ciudades y tiempos de entrega	-----

Opciones: Bondad de ajuste · Independencia 2x2 o mayor · Wilcoxon firmado · Kruskal-Wallis**Video de repaso**

VIDEO

Inferencia: video complementario

Repaso visual de contrastes, incluidos los categoricos.



Escanea para verlo en Lumbre

<https://www.youtube.com/watch?v=1v1ySMzTQzI>**Para profundizar**

Khan Academy: chi-cuadrada <https://www.khanacademy.org/math/statistics-probability>
 Capitulo de contrastes de Bondad e Independencia con ejercicios.

JMP: test de independencia chi-cuadrada https://www.jmp.com/en_us/statistics-knowledge-portal/chi-square-test/chi-square-test-of-independence.html Explicacion con ejemplo resuelto y formula de esperados (en ingles).

9. El problema 2 encontro asociacion pero NO causa. Redacta el parrafo de conclusion para un stakeholder que quiere "implementar T porque sube la recuperacion 30 puntos": que puede afirmar el numero, que no, y que diseno (randomizado) cerraria la pregunta causal.

Referencia curricular

Materia:	Ciencia de Datos e IA
-----------------	-----------------------

Versión interactiva completa (con actividades, videos y autoevaluación):



<https://univ.lumbre.lat/object/cdia-esti-010-chi-cuadrada-y-test-no-parametricos>

Material educativo generado desde Lumbre · lumbre.lat