**Algunos apuntes sobre chi-cuadrado y bondad de ajuste**

La distribución chi-cuadrado se puede emplear para pruebas de bondad de ajuste, en las que se comparan unas frecuencias empíricas (observadas) frente a unas frecuencias teóricas.

La idea es la siguiente. Pensemos un dado con CINCO caras, y queremos saber si está balanceado (es decir, si todas las caras con equiprobables; esta es es la “hipótesis nula”) o no. Para poner a prueba nuestra hipótesis, lanzamos el dado 100 veces y contamos el número de unos, doses, treses, cuatros y cincos. Y que las frecuencias empíricas sean:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 |
| frec.empírica | 30 | 10 | 20 | 15 | 25 |

¿son estos datos fácilmente compatibles con nuestra hipótesis de que el dado está bien construido? Necesitamos tener las frecuencias teóricas y ver las similitudes/diferencias. Las frecuencias teóricas las deducimos de la hipótesis de que el dado esté balanceado: de 100 lanzamientos, como todas las caras se asumen equiprobables, pues es 1/5 en cada cara, o sea, 20 en cada cara (100\*1/5):

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 |
| frec.empírica | 30 | 10 | 20 | 15 | 25 |
| frec.teórica | 20 | 20 | 20 | 20 | 20 |

Y cómo juzgamos su parecido. Pues mediante la prueba “chi cuadrado”, cuya fórmula está a continuación—que se realiza para cada casilla:



|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 |
| frec.empírica | 30 | 10 | 20 | 15 | 25 |
| frec.teórica | 20 | 20 | 20 | 20 | 20 |
|  |  |  |  |  |  |
|  | 5 | 5 | 0 | 1.25 | 1.25 |
|  |  |  |  |  |  |
|  | SUMA | 12.5 |  |  |  |

Ahora bien, esta suma de 12.5 (chi-cuadrado empírica), ¿qué quiere decir? Perfecto el ajuste no es, ¿pero es compatible con la idea de que el dado está equilibrado?

La clave es que, si realmente el dado estuviese equilibrado, la “chi-cuadrado” empírica seguiría una distribución teórica conocida: UNA CHI-CUADRADO CON CUATRO GRADOS DE LIBERTAD. Y de la cual tenemos tablas, podemos calcular percentiles, etc. (Lo de 4 grados de libertad es porque en realidad sabiendo cuatro de las frecuencias de las casillas, la quinta no puede variar, dado que la suma de frecuencias de las casillas en el ejemplo debe ser 100).

NOTA: si el dado estuviera bien hecho y hacemos la prueba 1000000000 de veces, y hacemos un histograma, y lo suavizamos, pues encajaría perfectamente con la distribución teórica chi-cuadrado.



Ahora bien, ¿cómo decidimos si quedarnos con la “hipótesis nula” (dado equilibrado) o no? Obviamente, una chi-cuadrado muy pequeña es un indicio de que el dado está equilibrado, y una chi-cuadrado muy alta es un indicio de que el dado no lo está. ¿Dónde ponemos el límite? Pues se suele poner (en psicología) en el percentil 95:

El criterio es, si la “chi-cuadrado” empírica supera el percentil 95, entonces decimos que el ajuste es demasiado pobre, por lo que rechazamos la “hipótesis nula” y concluimos que el dado no está equilibrado.

En cambio, si la “chi-cuadrado” empírica es menor al percentil 95, entonces decimos que el ajuste es aedcuado, mantenemos la “hipótesis nula” y concluimos que el dado está equilibrado.

¿Y en el ejemplo?



Pues el percentil 95 (en la tabla está la proporción de datos a la derecha) es 9.49.

Dado que la chi-cuadrado empírica es 12.5, es mayor que el percentil 95, por lo que concluimos que el dado no está equilibrado.