# Test $\chi^2$ y cálculo del p-value

**Samantha Kucher**

Supongamos que tenemos una variable aleatoria con distribución gaussiana: $x \sim N(\mu, \sigma)$. Si a esa variable le restamos su $\mu$ y la dividimos por su $\sigma$, vamos a obtener una variable con distribución $N(0, 1)$. Es decir, $y = \left(\frac{x-\mu}{\sigma}\right) \sim N(0, 1)$. Ahora supongamos que tenemos muchas variables con distribución gaussiana y efectuamos la siguiente operación:

$$z = \sum_{i=1}^{n} \left(\frac{x_i - \mu_i}{\sigma_i}\right)^2 = \sum_{i=1}^{n} y_i^2 \tag{1}$$

La variable $z$ que resulta de sumar $n$ variables con distribución $N(0, 1)$ elevadas al cuadrado, tiene distribución "chi cuadrado con $n$ grados de libertad ": $z \sim \chi_n^2$. En la Fig. 1 pueden observarse algunos ejemplos de distribuciones $\chi^2$ con distintos grados de libertad.

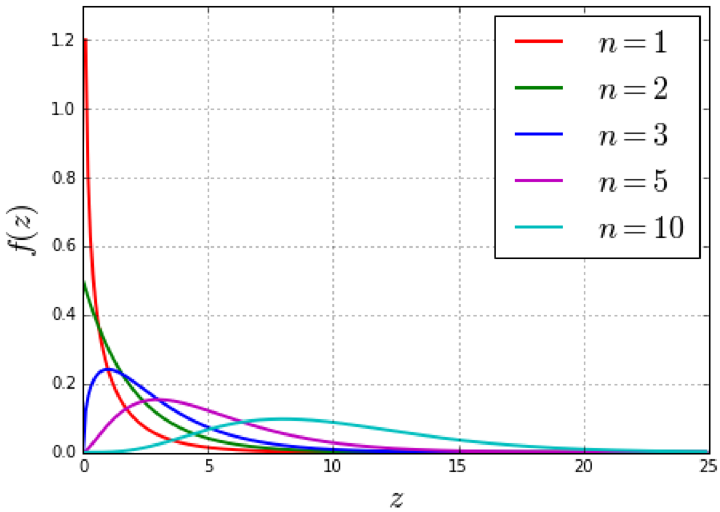


Figura 1: Distribuciones $\chi_n^2$.

Tomemos, por ejemplo, los puntos obtenidos de una medición. Supongamos que los queremos ajustar por una función, para lo cual utilizamos el método de cuadrados mínimos. La hipótesis que queremos testear es: "Los valores medidos fluctúan con distribución gaussiana alrededor de sus valores medios y estos son bien descriptos por la teoría $f(x_i)$". Es decir,

$$H_0 \begin{cases} x_i \sim N(\mu_i, \sigma_i) \\ \hat{\mu}_i = f(x_i) \end{cases}$$

Si la hipótesis $H_0$ es cierta, entonces la variable $w$ definida como

$$w = \sum_{i=1}^{n} \left(\frac{x_i - f(x_i)}{\sigma}\right)^2 \tag{2}$$

tiene distribución $\chi_{n-k}^2$ donde $k$ es la cantidad de parámetros de los que depende la función $f$. En particular, si queremos hacer un ajuste lineal, tenemos que

$$w = \sum_{i=1}^{n} \left(\frac{x_i - (\hat{a}x_i + \hat{b})}{\sigma_i}\right)^2 \sim \chi_{n-2}^2 \tag{3}$$

En la Fig. 2 se puede observar una $\chi_{65}^2$, que fue utilizada para realizar un test $\chi^2$ a un conjunto de 67 datos a los que se les aplicó un ajuste lineal. Cabe aclarar que como es una densidad de probabilidad, su integral está normalizada.

¿Cómo esperamos que sea el valor de $w$? Es la distancia de los puntos a la recta (al cuadrado), medida en $\sigma$, por lo tanto si la hipótesis es cierta (la recta ajusta bien a los puntos, que tienen distribución gaussiana) esperamos que sea "chico". Queremos rechazar, entonces, aquellos valores de $w$ que sean "grandes ". Para eso, definimos una *zona de rechazo*, tal que si nuestro valor de $w$ cae en esa zona, descartamos la hipótesis $H_0$. Para delimitar esa zona, elegimos un $w_{critico}$ de la siguiente manera:

$$\begin{cases} \text{Si } w_{medido} < w_{critico} \longrightarrow \text{Aceptamos } H_0 \\ \text{Si } w_{medido} > w_{critico} \longrightarrow \text{Rechazamos } H_0 \end{cases}$$

En general se habla en términos de probabilidades, no de valores de $w$. La probabilidad de caer en la zona de rechazo es lo que llamamos *significancia* ($\alpha$) del test (ver Fig. 2). En general se suele tomar $\alpha = 5\,\%$.

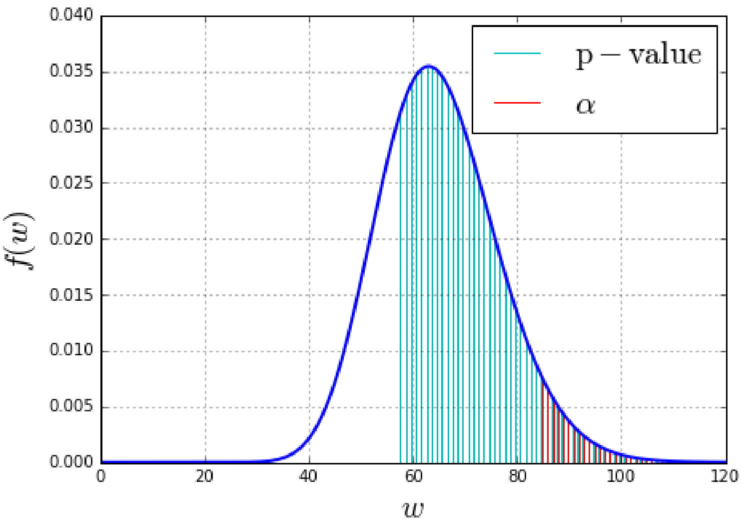


Figura 2: Distribución $\chi^2_{65}$, indicando significancia del test y p-value.

Se define el p-value como

$$p - value = \int_{w_{medido}}^{\infty} f(w)dw = P(w > w_{medido} \mid H_0) \tag{4}$$

Es la probabilidad de obtener un valor como el que obtuvimos, o más extremo, donde "más extremo", en este caso, significa un valor mayor (si $H_0$ es cierta). Es la mayor significancia que puedo pedirle al test sin rechazar $H_0$. Es necesario aclarar que el p-value no es la probabilidad de que $H_0$ sea cierta. El test $\chi^2$ queda definido como

$$\begin{cases} \text{Si p-value} > \alpha \longrightarrow \text{Aceptamos } H_0 \\ \text{Si p-value} < \alpha \longrightarrow \text{Rechazamos } H_0 \end{cases}$$