Un Enfoque en la Ensenãnza de la Distribución Normal

Alberto L. Maltz

Resumen

En este artículo se desarrolla una de las formas posibles de motivar la aparición de la Distribución Normal en los primeros cursos universitartios de Probabilidades.

Introducción

Como es conocido por los lectores se dice que la distribución de una variable aleatoria X es Normal, cuando su función de densidad viene dada por

$$f(x) = \frac{\exp(-(x-\mu)^2/2\sigma^2)}{\sigma\sqrt{2\pi}} \quad ,$$

resultando ser que μ es su media y σ^2 su varianza. Suele abreviarse diciendo que la distribución es $N(\mu, \sigma^2)$.

Se tiene entonces

$$P(a \leq X \leq b) = \int_a^b f(x)dx \quad ,$$

para todo $a \leq b$.

Cuando $\mu = 0$ y $\sigma^2 = 1$ se tiene la Distribución Normal estándar, con la cual puede manejarse los cálculos de todas las normales.

El hallazgo de la distribución Normal fue una de las exquisitas contribuciones de Abraham de Moivre (1667-1754) a la Matemática y a la Ciencia en general.

Como si fuera poco, la vía de este descubrimiento fue la prueba de la primera versión del importante Teorema Central del Límite (TCL). Un tratamiento accesible puede encontrarse en el Capítulo 2 de [1].

A partir de allí el tema acaparó la atención de grandes científicos, en algunos casos para las aplicaciones y en otros para hacer nuevas construcciones o formular desarrollos teóricos.

A través del tiempo fueron apareciendo nuevas visiones, alguna de las cuales son tan potentes que por si solas pueden convencernos de la importancia de la Normal.

Son inevitables los nombres de Pierre Simon de Laplace (1749-1827) y Carl Friedrich Gauss (1777-1855), quien propuso a la Normal como distribución de errores.

En el siglo XX aparecen construcciones de la Normal basadas en la iluminadora propiedad de maximizar la entropía de Shannon, que puede interpretarse como el grado de incertidumbre o de equidad. Para este tema referimos al Capítulo 9 de [2].

Una pregunta interesante, desde el ángulo de la educación matemática, es cómo introducir la distribucin Normal en un primer curso de Probabilidades.

Muchos autores la definen sin motivarla y esperan que sus buenas propiedades convenzan a los alumnos de su importancia.

Otros optan por deducirla en base a algunas de sus buenas propiedades analíticas, como por ejemplo las dos siguientes: La distribución conjunta de dos variables aleatorias independientes con distribución $N(0, \sigma^2)$ es invariante por rotaciones centradas en el origen; la suma de dos variables aleatorias independientes con distribuciones $N(\mu_1, \sigma_1^2), N(\mu_2, \sigma_2^2)$ es $N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$.

La motivaciones basadas en el TCL o el concepto de entropía no son descartables, pero suelen ser sofisticadas para un primer curso.

En esta nota quiero referirme a un esquema que era usual hace varias décadas (basado en el tipo de ideas estadísticas de Gauss) y que presentaré con un pequeño maquillaje.

Es un intento más de que el nuevo objeto matemático pueda incorporarse de manera amigable a los alumnos.

1. Muestras ideales

Supongamos que X es una variable aleatoria contínua con densidad f tal que

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx = 0$$

$$Var(X) = \int_{-\infty}^{\infty} x^2 f(x) dx = 1$$

(si los valores 0 y 1 fueran otros sería viable una idea similar).

Agregamos a f la condición de que es positiva, derivable en $\Re$, decreciente en $[0, \infty)$ y $f(x) = f(-x)$ para todo x. Todas las condiciones anteriores tienen motivaciones sencillas e indiscutidas en cualquier esquema didáctico.

Una muestra $x_1, ..., x_n$ de X podría llamarse "ideal" si las estimaciones habituales de media y varianza coincidieran con los valores exactos. Es decir

$$g_1(x_1, ..., x_n) = \frac{x_1 + ... + x_n}{n} = 0 \qquad , \qquad g_2(x_1, ..., x_n) = \frac{x_1^2 + ... + x_n^2}{n-1} = 1$$

Una propiedad que parece bueno pedirle a X para que sea "normal" (en el sentido cotidiano de la palabra) es que para cada n las muestras ideales sean las de mayor "chance". Esta maximalidad se traduce en la maximalidad de $\prod_{i=1}^{n} f(x_i)$ o equivalentemente la de su logaritmo $h(x_1, ..., x_n) = \sum_{i=1}^{n} ln(f(x_i))$.

Hemos armado un problema de extremos condicionados en n variables. Al atacarlo con el método de los Multiplicadores de Lagrange, a las ecuaciones $g_1(x_1, ..., x_n) = 0$ y $g_2(x_1, ..., x_n) = 0$, con incógnitas $x_1, ..., x_n$, se agregan las n siguientes (que junto con las nuevas incógnitas λ_1 y λ_2 provienen del planteo $\nabla h = \lambda_1 \nabla g_1 + \lambda_2 \nabla g_2$) :

$$\frac{f'(x_i)}{f(x_i)} = \frac{\lambda_1}{n} + 2\frac{\lambda_2}{n-1} x_i \quad i = 1, ..., n.$$

Un modo de que esto funcione bien (el único con la información que tenemos) es que f satisfaga la ecuación diferencial

$$\frac{f'(x)}{f(x)} = a_1 + a_2 x$$

donde $a_1 = \lambda_1/n$, $a_2 = 2\lambda_2/(n-1)$.

La solución general de esta ecuación es

$$f(x) = a_0 \exp(a_1 x + a_2 x^2/2),$$

que agerega la nueva incógnita a_0.

De las condiciones pedidas para f y utilizando la conocida igualdad $\int_{-\infty}^{\infty} \exp(-x^2)dx = \sqrt{\pi}$ se obtiene $a_0 = 1/\sqrt{2\pi}$, $a_1 = 0$, $a_2 = -1$, con lo cual

$$f(x) = \frac{\exp(-x^2/2)}{\sqrt{2\pi}} \ ,$$

densidad de una $N(0,1)$.

Referencias

1. Durrett, R. Probability: Theory and Examples. Duxury Press.

2. Rényi, A. Cálculo de Probabilidades. Editorial Reverté.

Universidad Nacional de La Plata.