La variable aleatoria Gaussiana (o normal)

La v.a. Gaussiana es por mucho la distribución más usada para describir fenómenos aleatorios. Debe su nombre al matemático Aleman Carl Friedrich Gauss quien la utilizo para modelar datos astronómicos. Sin embargo no fue el primero en usarla existen registros de que de Moivre y Laplace la usaron para modelar errores de medición.

Se le llama normal pues en los 1,800 se tenía la creencia de que la distribución Gaussiana representaba las "leyes de la naturaleza" pues muchos fenomenos aleatorios pueden describirse usando esta v.a.

- Alturas, peso, circunferencia de la cabeza, longitud huesos, tamaño de craneo, frecuencia cardiaca, presión arterial, tamaño de los órganos internos, niveles de glucosa en sangre, etc.

- Errores de medición
- Tiempo de reacción
- Tiempo de ejecución de programas informaticos
- Calificaciones o scores.

etc

Además, por el Teorema Central de Límite sabemos que la distribución normal describe la distribución de los promedios (bajo algunos supuestos).

El problema quizá es que se usa demasiado, incluso en casos en los que es claro que no se obtendrá una buena aproximación.

- No describiremos como se obtiene, pero hay distintas maneras

$$X \sim Bin(n, p)$$

$$Z_n = \frac{X - np}{\sqrt{np(1-p)}} \xrightarrow[n \to \infty]{} Z \sim N(0,1)$$

- Construyendo un modelo para describir errores de medición.
- Construyendo un modelo para describir desviaciones de un objetivo.

Decimos que $X \sim N(\mu, \sigma^2)$, si su función de distribución de probabilidad está dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}(x-\mu)^2}, \quad -\infty < x < \infty$$

Como hemos mencionado, esta distribución es una modo para describir observaciones que representan desviaciones con respecto a un valor central. Las desviaciones pueden ser negativas o positivas

¿Es una distribución de probabilidad?

Claramente $f(x) \geq 0$

$$\int_{-\infty}^{\infty} f(x)\,dx = 1$$

Vamos a trabajar el caso en el que $Z \sim N(0,1)$, i.e.

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty$$

la normal estandar

$$A^2 = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z_1^2/2}\, dz_1\right)\left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-z_2^2/2}\, dz_2\right)$$

$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} e^{-(z_1^2+z_2^2)/2}\, dz_1\right) dz_2$$

Coordenadas polares $z_1 = r\cos(\theta)$ $\quad r \in (0,\infty)$

$z_2 = r\,\text{sen}(\theta)$ $\quad \theta \in [0, 2\pi]$

$|J| = r$

$$\Rightarrow A = \frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty} r e^{-r^2/2}\, dr\, d\theta$$

$$= \int_0^{\infty} r e^{-r^2/2}\, dr = -\int_0^{-\infty} e^{u}\, du$$

$$u = -r^2/2 \Rightarrow du = -r\, dr \qquad = \int_{-\infty}^{0} e^{u}\, du = 1$$

$$A^2 = 1 \iff \underline{A = 1} \quad \text{ó} \quad -1$$

$$\text{Si } Z \sim N(0,1) \Rightarrow X = \mu + \sigma Z \sim N(\mu, \sigma^2)$$

Dem

$$F_X(x) = P(X \leq x) \quad \leftarrow \text{transformación vía la función de distribución}$$

$$= P(\mu + \sigma Z \leq x)$$

$$= P\left(Z \leq \frac{x-\mu}{\sigma}\right) = F_Z\left(\frac{x-\mu}{\sigma}\right)$$

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{1}{\sigma} f_Z\left(\frac{x-\mu}{\sigma}\right)$$

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad z \in (-\infty, \infty)$$

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in (-\infty, \infty)$$

Si: $X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P\left(\frac{X-\mu}{\sigma} \leq z\right) \\ &= P(X \leq \mu + \sigma z) \end{aligned}$$

$$\begin{aligned} f(z) &= \frac{d}{dz} P(X \leq \mu + \sigma z) \\ &= \sigma f_X(\mu + \sigma z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < \infty \end{aligned}$$

¡Importante! Si: $Z \sim N(0,1)$

$\Rightarrow X = \mu + \sigma Z \sim N(\mu, \sigma^2)$

o $X \sim N(\mu, \sigma^2)$

$\Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0,1)$

Por eso todos los resultados para la distribución normal $N(\mu, \sigma^2)$ sólo se prueban para la $N(0,1)$

Función generadora de momentos

$$M_Z(t) = \mathbb{E}(e^{Zt})$$

$$= \int_{-\infty}^{\infty} e^{zt} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \, dz$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z^2 - 2zt)}$$

$$z^2 - 2zt + t^2 - t^2 = (z-t)^2 - t^2$$

$$= e^{t^2/2} \underbrace{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(z-t)^2} dz}_{N(t,1)}$$

$$M_Z(t) = e^{t^2/2}$$

$$M_X(t) = \mathbb{E}(e^{Xt}) = \mathbb{E}(e^{(\mu + Z\sigma)t})$$

$$= e^{\mu t}\, \mathbb{E}(e^{Z\sigma t})$$

$$= e^{\mu t} M_Z(\sigma t)$$

$$= e^{\mu t} e^{\sigma^2 t^2/2}$$

$$= e^{\mu t + \sigma^2 t^2/2}$$

Si $X \sim N(\mu, \sigma^2)$

$$\Rightarrow M_X(t) = e^{\mu t + \sigma^2 t^2/2}$$

$$\mathbb{E}(X) = \frac{d}{dt} M_X(t) \Big|_{t=0}$$

$$= (\mu + t\sigma^2)\, e^{\mu t + \sigma^2 t^2/2} \Big|_{t=0}$$

$$= \mu$$

$$\mathbb{E}(X^2) = \frac{d^2}{dt^2} M_X(t) \Big|_{t=0}$$

$$= \mu(\mu + \sigma^2 t)\, e^{\mu t + \sigma^2 t^2/2}$$
$$+ \sigma^2 e^{\mu t + \sigma^2 t^2/2} + t \{$$

$$= \mu^2 + \sigma^2$$

$$\Rightarrow \mathbb{E}(X^2) = \mu^2 + \sigma^2$$

$$\Rightarrow \operatorname{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2$$
$$= \sigma^2$$

Coeficiente de asimetría $\gamma_3 = 0$
Coeficiente de curtosis $\gamma_4 = 3$

Como la curtosis de la distribución $N(\mu, \sigma^2)$

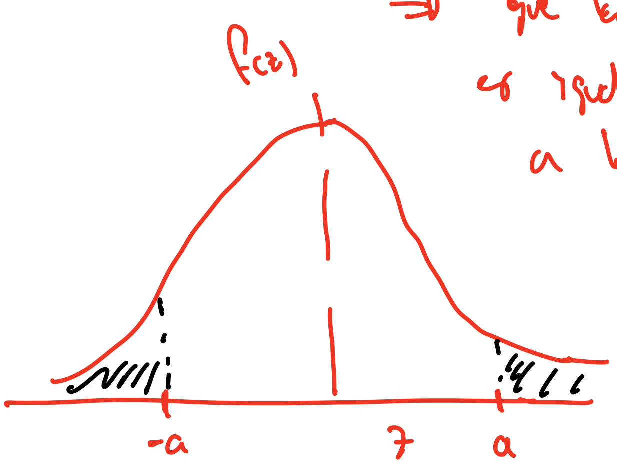
$\Rightarrow$ si obtenemos γ_4 para otra distribución

$\gamma_4 < 3 \Rightarrow$ es una distribución chata

$\gamma_4 > 3 \Rightarrow$ es una distribución muy picuda.

Mas propiedades de la $N(0,1)$ $\gamma_3 = 0$

$\Rightarrow$ que la normal es simétrica
es igualmente probable tener valores
a la izquierda como a la derecha



$F(-a) = 1 - F_Z(a)$, en donde

$$F_Z(a) = \int_{-\infty}^{a} f(z)\, dz$$

Si $X \sim N(\mu, \sigma^2)$ y queremos

$$P(a < X < b) = P\left(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\right)$$

$$= P\left(Z \leq \frac{b-\mu}{\sigma}\right) - P\left(Z \leq \frac{a-\mu}{\sigma}\right)$$

$$= F_Z\left(\frac{b-\mu}{\sigma}\right) - F_Z\left(\frac{a-\mu}{\sigma}\right)$$