$\mathbb{X}$ v.a. que nos ayudará a modelar un fenómeno aleatorio

$f(x)$ función de distribución de probabilidad. Nos muestra el patrón agregado que $\mathbb{X}$ es capaz de describir

$F(x)$ función de distribución acumulada. Nos ayuda a calcular $P_{\mathbb{X}}(\mathbb{X} \in A)$.

$\mathbb{E}(\mathbb{X})$ indica el valor esperado o promedio que toma la v.a

$$\mathbb{E}(\mathbb{X}) = \begin{cases} \sum_{j \geq 1} x_j f(x_j), & \mathbb{X} \text{ v.a. discreta} \\ \int_{-\infty}^{\infty} x f(x) dx, & \mathbb{X} \text{ v.a. continua.} \end{cases}$$

Además, se tiene que

$$\mathbb{E}(g(\mathbb{X})) = \begin{cases} \sum_{j \geq 1} g(x_j) f(x_j), & \mathbb{X} \text{ v.a. discreta} \\ \int_{-\infty}^{\infty} g(x) f(x) dx, & \mathbb{X} \text{ v.a. continua} \end{cases}$$

se le conoce como el teorema del estadístico inconsciente. (TEI)

En particular se tiene que para funciones lineales

$$g(X) = aX + b$$

$$\mathbb{E}(aX + b) = a\,\mathbb{E}(X) + b$$

Dem.

$$\mathbb{E}(aX + b) \overset{\text{TEI}}{=} \int (ax + b) f(x)\, dx$$

$$= a\int x f(x)\, dx + b\cancelto{1}{\int f(x)\, dx}$$

$$= a\,\mathbb{E}(X) + b$$

Momentos de una distribución

Para cualquier entero k, el k-ésimo momento de la v.a. X se define como

$$m_k = \mathbb{E}(X^k)$$

En donde por el TEI sabemos que

$$\mathbb{E}(X^k) = \int x^k f(x)\, dx$$

Siempre que $\mathbb{E}(|X|^K) < \infty$ tendremos que el K-ésimo momento de la v.a. X existe.

El primer y segundo momento serán relevantes

$$m_1 = \mathbb{E}(X) = \mu \qquad y \qquad m_2 = \mathbb{E}(X^2)$$

Momentos alrededor de la media $= \mathbb{E}(X)$

$$\mu_K = \mathbb{E}((X - \mu)^K)$$

también son conocidos como momentos centrales de la v.a. X

Segundo momento central : varianza

$$\mu_2 = \mathbb{E}((X-\mu)^2) = Var(X) = \sigma^2$$

nos da información de la desviación promedio entre los valores que toma la v.a. X y su valor promedio [unidades de la v.a. al cuadrado]

Es fácil demostrar que

$$\sigma^2 = Var(X) = \mathbb{E}((X-\mu)^2)$$
$$= \mathbb{E}(X^2) - \mu^2$$
$$= m_2 - m_1^2$$

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

$$= \int (x^2 - 2x\mu + \mu^2) f(x)\, dx$$

$$= \int x^2 f(x)\, dx - 2\mu \int x f(x)\, dx + \mu^2 \int f(x)\, dx$$

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

$$= \mathbb{E}(X^2) - \mu^2$$

En el caso de la varianza de X tenemos que

Para una constante b, $Var(b) = 0$

$$Var(b) = \mathbb{E}(b^2) - \mathbb{E}(b)^2$$

$$= b^2 - b^2 = 0$$

Una constante tiene varianza cero, no se desvía con respecto a su valor promedio.

Para constantes a y b $Var(aX+b) = a^2 Var(X)$

$$Var(aX+b) = \mathbb{E}((aX+b)^2) - \mathbb{E}(aX+b)^2$$

$$= \mathbb{E}\left(a^2 X^2 + 2abX + b^2\right) - (a\mu + b)^2$$

$$= a^2\,\mathbb{E}(X^2) + \cancel{2ab\mu} + \cancel{b^2} - \left(a^2\mu^2 + \cancel{2a\mu b} + \cancel{b^2}\right)$$

$$= a^2\left(\mathbb{E}(X^2) - \mu^2\right) = a^2\,Var(X) \perp$$

$$\sigma = \sqrt{\sigma^2} = \sqrt{Var(X)}$$ es conocida

como la desviación estandar de X y de la misma forma que σ^2 nos da información acerca de las desviaciones entre los valores que toma X y su valor promedio. La diferencia es que $\underline{\sigma}$ esta en las mismas unidades que X por lo que en general se interpreta de forma más clara y fácil.

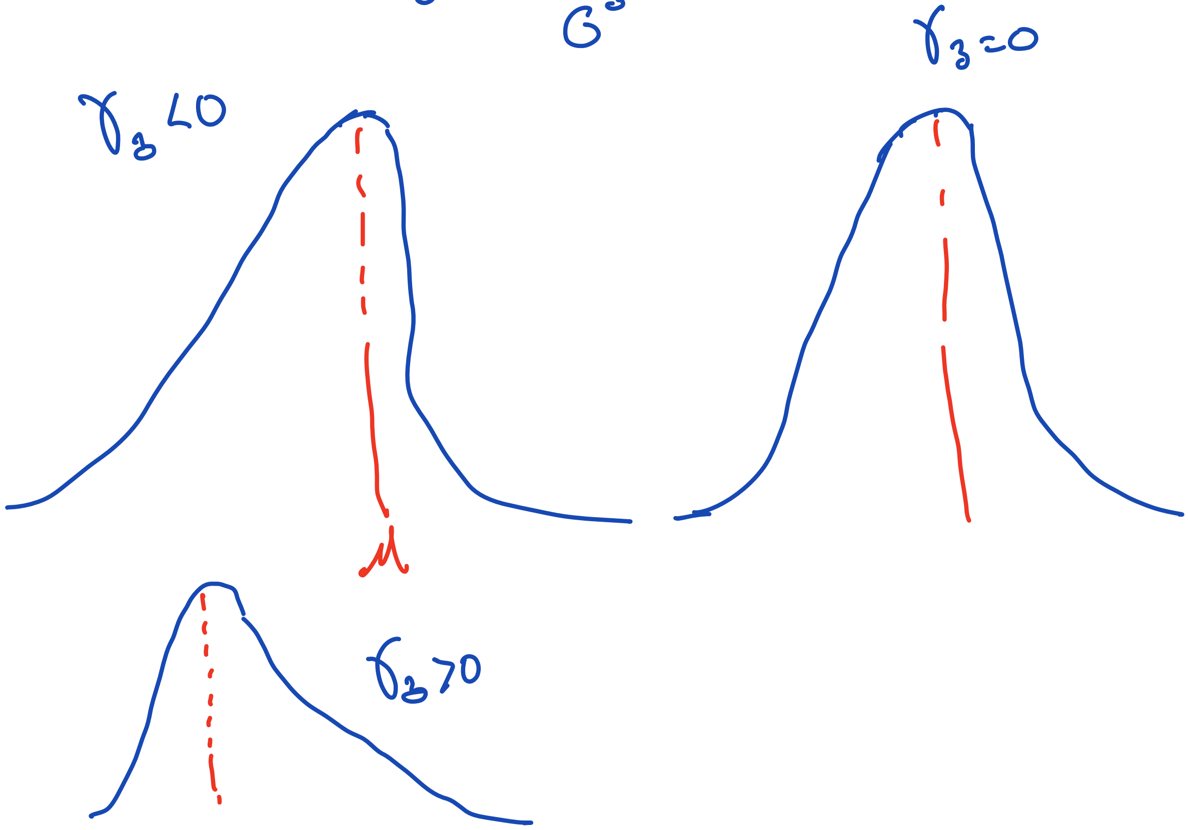
Tercer momento central: Asimetría

Al tercer momento central

$$\mu_3 = \mathbb{E}\left((X - \mu)^3\right)$$

se le llama asimetría de la v.a. y nos da información acerca de si la v.a. toma con mayor probabilidad valores a la izquierda o a la derecha de la media. Para cancelar el efecto de la dimensión suele definirse el coeficiente de asimetría

$$\gamma_3 = \frac{\mu_3}{\sigma^3}$$



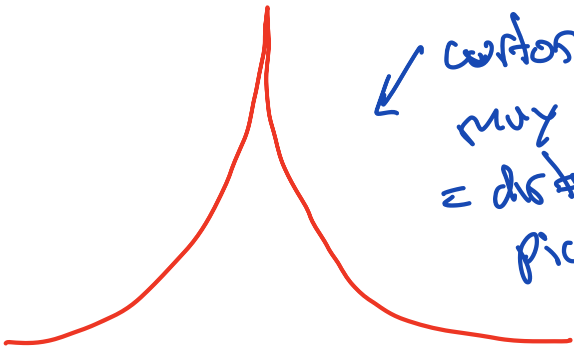
Cuarto momento central: curtosis

$$\mu_4 = \mathbb{E}((X-\mu)^4)$$

para eliminar el efecto de la potencia

$$\gamma_4 = \frac{\mu_4}{\sigma^4} \leftarrow \text{coeficiente de curtosis}$$

Nos da información de que tan picuda o chata es una distribución.



curtosis muy grande = distribución picuda y colas pesadas

curtosis pequeña = distribución chata y colas ligeras

Aplicación: interpretación

μ media — nos da información de la localización de la distribución de X, así como de su valor promedio o esperado. Y es el candidato para el mejor representante de los valores que puede tomar la v.a.

σ^2 y σ varianza — nos da información acerca de que tan juntos o dispersos se encuentran los valores que puede tomar X con respecto a μ.

μ^3 y γ_3 asimetría — nos dice si la v.a. toma con mayor probabilidad valores a la izquierda o a la derecha de su valor esperado.

μ^4 y γ_4
curtosis

Nos indica cuanta varianza se debe a observaciones extremas (la cola).

Mayor curtosis = picos mas agudos y colas mas pesadas $\Rightarrow$ los valores extremos que tome la v.a. contribuyen más a la varianza.

Menor curtosis = picos chatos y colas ligeras $\Rightarrow$ los valores extremos no contribuyen mucho a la varianza y esta esta determinada principalmente por los valores alrededor de μ.

Función generadora de momentos

La función generadora de momentos de una v.a. X, se define como

$$M_X(t) = \mathbb{E}(e^{Xt}) = \begin{cases} \sum_{j\geq 1} e^{x_j t} f(x_j), & X \text{ discreta} \\ \int_{-\infty}^{\infty} e^{xt} f(x)\,dx, & X \text{ continua} \end{cases}$$

Se le llama fgm porque

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

La función generadora de momentos tiene las siguientes propiedades

1. Unicidad. La fgm no existe para todas las v.a. X, pero cuando existe determina de manera única la distribución de la v.a. X

$\Rightarrow$ Si dos v.a. X e Y tienen la misma fgm $(M_X(t) = M_Y(t))$ entonces tienen la misma distribución.

2 = Transformaciones lineales.

Si $Y = aX + b$

$$\Rightarrow M_Y(t) = e^{bt} M_X(at)$$

3 = Sumas de v.a. independientes.

Si $X \perp Y$ y $Z = X + Y$

$$\Rightarrow M_Z(t) = M_X(t) M_Y(t)$$