

Función generadora de momentos:

Definición: Si X es una variable aleatoria, el momento de orden k de X se define como

$$E(X^k)$$

siempre que la esperanza exista.

Notemos que

$E(X) = \mu$	1er momento: posición
$E(X^2) = \sigma^2 + \mu^2$	2do momento: dispersión
$E(X^3)$	3er momento: relacionado con una medida de asimetría
$E(X^4)$	4to momento: relacionado con la kurtosis

Definición: La función generadora de momentos de una v.a. X es una función a valores reales $M_X(t)$, definida como

$$M_X(t) = E(e^{tX}) = \begin{cases} \sum_{x \in R_X} e^{tx} p_X(x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} e^{tx} f_X(x) dx & \text{si } X \text{ es continua} \end{cases}$$

siempre que el valor esperado exista para todo $t \in (-h, h)$, $h > 0$. Esta última es una condición técnica necesaria para que $M_X(t)$ sea diferenciable en 0.

Se denomina función generadora de momentos porque los momentos de X ($E(X^n)$) pueden ser obtenidos derivando esta función y evaluando la derivada en $t = 0$, tal como lo establece el siguiente teorema.

Teorema: Sea X una v.a. para la cual existe la función generadora de momentos $M_X(t)$, entonces

$$E(X^n) = \left. \frac{\partial^n}{\partial t^n} M_X(t) \right|_{t=0}$$

La demostración se basa en el siguiente lema de Cálculo avanzado (ver por ejemplo, Advanced Calculus, D. Widder (1961)):

Lema: Si la función $g(t)$ definida por

$$g(t) = \sum_x e^{tx} p(x) \quad \text{ó} \quad g(t) = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

converge para todo $t \in (-h, h)$ para algún $h > 0$, entonces existen las derivadas de orden n de $g(t)$ para todo $t \in (-h, h)$ y para todo n entero positivo y se obtienen como

$$\frac{\partial^n g(t)}{\partial t^n} = \sum_x \frac{\partial^n e^{tx}}{\partial t^n} p(x) \quad \text{ó} \quad \frac{\partial^n g(t)}{\partial t^n} = \int_{-\infty}^{\infty} \frac{\partial^n e^{tx}}{\partial t^n} f(x) dx$$

Demostración del Teorema: Si la función generadora de momentos existe para todo $t \in (-h, h)$ para algún $h > 0$, aplicando el lema,

$$\frac{\partial^n M_X(t)}{\partial t^n} = \sum_x \frac{\partial^n e^{tx}}{\partial t^n} p(x) \quad \text{ó} \quad \frac{\partial^n M_X(t)}{\partial t^n} = \int_{-\infty}^{\infty} \frac{\partial^n e^{tx}}{\partial t^n} f(x) dx$$

$$\frac{\partial^n M_X(t)}{\partial t^n} = \sum_x x^n e^{tx} p(x) \quad \text{ó} \quad \frac{\partial^n M_X(t)}{\partial t^n} = \int_{-\infty}^{\infty} x^n e^{tx} f(x) dx$$

Evaluando estas derivadas en 0 ,

$$\left. \frac{\partial^n M_X(t)}{\partial t^n} \right|_{t=0} = \sum_x x^n p(x) = E(X^n) \quad \text{ó} \quad \left. \frac{\partial^n M_X(t)}{\partial t^n} \right|_{t=0} = \int_{-\infty}^{\infty} x^n f(x) dx = E(X^n)$$

Ejemplos: 1) Sea X una v.a. con distribución exponencial de parámetro λ , o sea con densidad

$$f_X(x) = \lambda e^{-\lambda x} I_{(0, \infty)}(x)$$

$$M_X(t) = E(e^{tX}) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t} \int_0^{\infty} (\lambda-t) e^{-(\lambda-t)x} dx = \frac{\lambda}{\lambda-t}$$

siempre que $t < \lambda$.

Calculemos ahora $E(X)$ y $V(X)$.

$$E(X) = \left. \frac{\partial M_X(t)}{\partial t} \right|_{t=0} = \left. \frac{\partial}{\partial t} \left(\frac{\lambda}{\lambda - t} \right) \right|_{t=0} = \left. \frac{\lambda}{(\lambda - t)^2} \right|_{t=0} = \frac{1}{\lambda}.$$

Como $V(X) = E(X^2) - (E(X))^2$, calculemos $E(X^2)$.

$$E(X^2) = \left. \frac{\partial^2 M_X(t)}{\partial t^2} \right|_{t=0} = \left. \frac{\partial}{\partial t} \left(\frac{\lambda}{(\lambda - t)^2} \right) \right|_{t=0} = \left. \frac{2\lambda(\lambda - t)}{(\lambda - t)^4} \right|_{t=0} = \frac{2}{\lambda^2}$$

entonces, $V(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$.

2) Sea X una v.a. con distribución Binomial de parámetros, n y p , o sea $X \sim \text{Bi}(n, p)$. Su función de probabilidad puntual es

$$p_X(k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{si } 0 \leq k \leq n$$

$$M_X(t) = E(e^{tX}) = \sum_{k=0}^n e^{tk} \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \binom{n}{k} (e^t p)^k (1-p)^{n-k} = (e^t p + 1 - p)^n.$$

Calculemos ahora $E(X)$ y $V(X)$.

$$E(X) = \left. \frac{\partial M_X(t)}{\partial t} \right|_{t=0} = \left. \frac{\partial (e^t p + 1 - p)^n}{\partial t} \right|_{t=0} = n(e^t p + 1 - p)^{n-1} p e^t \Big|_{t=0} = np.$$

$$E(X^2) = \left. \frac{\partial^2 M_X(t)}{\partial t^2} \right|_{t=0} = \left. \frac{\partial}{\partial t} (n(e^t p + 1 - p)^{n-1} p e^t) \right|_{t=0} =$$

$$= \left. (n(n-1)(e^t p + 1 - p)^{n-2} (p e^t)^2 + n(e^t p + 1 - p)^{n-1} p e^t) \right|_{t=0} = n(n-1)p^2 + np.$$

Entonces, $V(X) = E(X^2) - (E(X))^2 = n(n-1)p^2 + np - (np)^2 = -np^2 + np = np(1-p)$.

Propiedad: Sea X una v.a. con función generadora de momentos $M_X(t)$, entonces si

$$Y = aX + b, \text{ entonces } M_Y(t) = e^{bt} M_X(at).$$

Dem: Ejercicio.

Unicidad de $M_X(t)$: Además de permitir calcular momentos de una v.a., la función generadora de momentos permite identificar la función de densidad o de probabilidad de una v.a. debido a la propiedad de **unicidad**, la cual establece que hay una correspondencia uno a uno entre funciones de densidad o probabilidad y funciones generadoras de momentos.

Teorema de Unicidad: Si existe la función generadora de momentos de una variable aleatoria, es única. Además la función generadora de momentos determina a la función de densidad o probabilidad de la v.a. salvo a lo sumo en un conjunto de probabilidad 0.

A continuación, presentamos una tabla con la función generadora de momentos de algunas de las familias de distribución que hemos estudiado.

Distribución	$M_X(t)$
Bi(n,p)	$(e^t p + 1 - p)^n$
P(λ)	$e^{\lambda(e^t - 1)}$
N(μ, σ^2)	$e^{\frac{\sigma^2 t^2}{2} + \mu t}$
E(λ)	$\frac{\lambda}{\lambda - t}$
G(α, λ)	$\left(\frac{\lambda}{\lambda - t} \right)^\alpha$
U(a,b)	$\frac{e^{tb} - e^{ta}}{t(b - a)}$
G(p)	$\frac{p e^t}{1 - (1 - p) e^t}$
BN(r,p)	$\left(\frac{p e^t}{1 - (1 - p) e^t} \right)^r$

Ejercicio: ¿Para qué valores de t existe cada una de las funciones generadoras de momentos de la tabla anterior?