Función generadora de momentos y suma de variables independientes

Kevin Piterman

9 de Junio, 2020

Resumen.

Definición. La función generadora de momentos de una variable aleatoria X es la función

$$M_X(t) = \mathbb{E}(e^{tX}),\ t \in \mathbb{R} \quad \text{(si existe)}.$$

El n-ésimo momento de X es $\mathbb{E}(X^n)$.

Interpretación de los momentos:

- $\mathbb{E}(X) = \mu$, primer momento: posición (la media).
- $\mathbb{E}(X^2) = \mu^2 + \sigma^2$, segundo momento: dispersión (con respecto a la media).
- $\mathbb{E}(X^3)$, tercer momento: se relaciona con una medida de asimetría (cuánto se desplaza la distribución en una determinada dirección).
- $\mathbb{E}(X^4)$, cuarto momento: se relaciona con la kurtosis (cuán "puntiaguda" es la distribución).

Teorema. (Relación momentos y M_X) $\mathbb{E}(X^n) = \frac{\partial^n M_X(t)}{\partial t^n}(0)$.

Proof. Separar en casos: continua y discreta, y derivar. □

¿Para qué nos sirve la función generadora de momentos? Porque **caracteriza la distribución**.

Teorema. (Unicidad) Si X e Y son dos variables aleatorias tales que admiten funciones generadoras de momentos, entonces $M_X = M_Y$ si y solo si $X \sim Y$. Es decir, la función generadora de momentos solo depende de la distribución.

Uso: calcular distribuciones. Por ejemplo, suma de variables aleatorias independientes.

Proposición. Sean X_1, …, X_n variables aleatorias independientes con funciones generadoras de momentos M_{X_1}, …, M_{X_n}. Entonces $M_{\sum_i X_i}(t) = M_{X_1}(t). \dots .M_{X_n}(t)$.

Proof. La esperanza de variables independientes es el producto de las esperanzas:

$$M_{\sum_i X_i}(t) = \mathbb{E}(e^{t\sum_i X_i}) = \mathbb{E}(e^{tX_1} \dots e^{tX_n}) = \mathbb{E}(e^{tX_1}) \dots \mathbb{E}(e^{tX_n}) = M_{X_1}(t). \dots .M_{X_n}(t).$$

□

Teorema. Sean $X_1, \dots, X_n$ variables aleatorias independientes y sea $X = \sum_{i=1}^n X_i$.

1. Si $X_i \sim Bi(n_i, p)$ para todo i, entonces $X \sim Bi(\sum_i n_i, p)$.
2. Si $X_i \sim Ge(p)$ para todo i, entonces $X \sim BN(n, p)$.
3. Si $X_i \sim \mathcal{P}(\lambda_i)$ para todo i, entonces $X \sim \mathcal{P}(\sum_i \lambda_i)$.
4. Si $X_i \sim \Gamma(\alpha_i, \lambda)$ para todo i, entonces $X \sim \Gamma(\sum_i \alpha_i, \lambda)$.

5. Si $X_i \sim N(\mu_i, \sigma_i^2)$ para todo i, entonces $X \sim N(\sum_i \mu_i, \sum_i \sigma_i^2)$.
6. Si $X_i \sim U(0,1)$ para todo i, entonces X tiene la distribución *Irwin-Hall* (nos fuimos de tema...).

Proof. Ejercicio! Sugerencia: calcular primero las funciones generadoras de momento de cada distribución y utilizar la proposición anterior. $\square$

Ejercicio 1. Probar que suma de Gammas independientes es Gamma como dice el teorema anterior.

Sugerencia: probar primero que $M_Y(t) = \frac{\lambda^\alpha}{(\lambda-t)^\alpha}$ si $Y \sim \Gamma(\alpha, \lambda)$.

Solución.

Calculemos primero $M_Y(t)$, con $Y \sim \Gamma(\alpha, \lambda)$. Por definición,

$$\begin{aligned}
M_Y(t) &= \mathbb{E}(e^{tY}) \\
&= \int e^{ty} f_Y(y) dy \\
&= \int_0^{+\infty} e^{ty} e^{-\lambda y} \lambda^\alpha y^{\alpha-1} \frac{1}{\Gamma(\alpha)} dy \\
&= \int_0^{+\infty} e^{-(\lambda-t)y} y^{\alpha-1} \frac{(\lambda-t)^\alpha}{\Gamma(\alpha)} \frac{\lambda^\alpha}{(\lambda-t)^\alpha} dy \\
&= \int_0^{+\infty} f_{\Gamma(\lambda-t,\alpha)}(y) \frac{\lambda^\alpha}{(\lambda-t)^\alpha} dy \\
&= \frac{\lambda^\alpha}{(\lambda-t)^\alpha}.
\end{aligned}$$

Notar que esto vale para $\lambda \neq t$.

Para probar la distribución de la suma, utilizamos el teorema de unicidad de la función generadora de momentos. O sea, queremos ver que

$$M_{\sum_i X_i}(t) = M_{\Gamma(\sum_i \alpha_i, \lambda)}(t).$$

Por independencia,

$$M_{\sum_i X_i}(t) = \prod_i M_{X_i}(t) = \prod_i \frac{\lambda^{\alpha_i}}{(\lambda-t)^{\alpha_i}} = \left(\frac{\lambda}{\lambda-t}\right)^{\sum_i \alpha_i} = M_{\Gamma(\sum_i \alpha_i, \lambda)}(t).$$

Por lo tanto $\sum_i X_i \sim \Gamma\left(\sum_i \alpha_i, \lambda\right)$.

Ejercicio 2. Sean X, Y variables aleatorias independientes con distribución $U(-1,1)$. Sea $Z = \frac{X+Y}{2}$.

1. Probar que $f_Z(z) = (z+1)I_{(-1,0)}(z) + (1-z)I_{(0,1)}(z)$.
2. Probar que $M_{aW+b} = e^{bt}M_W(at)$ para cualquier variable W.
3. Hallar $M_{U(0,1)}$ y usarlo para calcular $M_{U(-1,1)}$ con el ítem anterior.
4. Hallar M_Z.

Solución. Vamos a resolver el ejercicio usando funciones generadoras de momentos. Sea $g(z) = (z+1)I_{(-1,0)}(z) + (1-z)I_{(0,1)}(z)$. Sea Z_0 una variable aleatoria con densidad g. Para ver que la densidad de Z es la misma que la densidad de Z_0, alcanza con ver que $M_Z = M_{Z_0}$ por el teorema de unicidad de funciones generadoras de momentos.

Calculemos primero M_{Z_0}:

$$
\begin{aligned}
M_{Z_0}(t) &= \mathbb{E}(e^{tZ_0}) \\
&= \int e^{tz} g(z) dz \\
&= \int e^{tz}((z+1)I_{(-1,0)}(z) + (1-z)I_{(0,1)}(z))dz \\
&= \int e^{tz}(z+1)I_{(-1,0)}(z) + e^{tz}(1-z)I_{(0,1)}(z)dz \\
&= \int e^{tz}(z+1)I_{(-1,0)}(z)dz + \int e^{tz}(1-z)I_{(0,1)}(z)dz \\
&= \int_{-1}^{0} e^{tz}(z+1)dz + \int_{0}^{1} e^{tz}(1-z)dz \\
&= \frac{t+e^{-t}-1}{t^2} + \frac{-t+e^{t}-1}{t^2} \\
&= \frac{e^{-t}+e^{t}-2}{t^2}
\end{aligned}
$$

Para calcular M_Z, resolvemos primero los ítems 2 y 3.

$$M_{aW+b}(t) = \mathbb{E}(e^{t(aW+b)}) = \mathbb{E}(e^{taW}e^{tb}) = \mathbb{E}(e^{taW})e^{tb} = M_W(at)e^{tb}.$$

Calculamos ahora la generadora de momentos de $U(0,1)$.

$$M_{U(0,1)}(t) = \int e^{tu} I_{(0,1)}(u) du = \int_0^1 e^{tu} du = \frac{1}{t} e^{tu}\Big|_{u=0}^{u=1} = \frac{1}{t}\left(e^t - 1\right).$$

Si $W \sim (0,1)$ entonces $2W-1 \sim U(-1,1)$, y por la cuenta anterior

$$M_{U(-1,1)}(t) = M_{2W-1}(t) = M_W(2t)e^{t(-1)} = \frac{1}{2t}(e^{2t}-1)e^{-t}.$$

Recordar que $M_X = M_{U(-1,1)} = M_Y$. Luego,

$$
\begin{aligned}
M_Z(t) = M_{\frac{X+Y}{2}}(t) = M_{X+Y}\left(\frac{t}{2}\right) = M_X\left(\frac{t}{2}\right) M_Y\left(\frac{t}{2}\right) = \left(\frac{1}{t}(e^t-1)e^{-\frac{t}{2}}\right)^2 \\
= \frac{e^{-t}+e^{t}-2}{t^2} = M_{Z_0}(t).
\end{aligned}
$$

Como $M_Z = M_{Z_0}$, $f_Z(z) = f_{Z_0}(z) = g(z) = (z+1)I_{(-1,0)}(z) + (1-z)I_{(0,1)}(z)$.

Ejercicio 3. (Tarea?) Si X, Y son variables aleatorias independientes (ambas continuas o ambas discretas), entonces $f_{X+Y} = f_X * f_Y$ (caso continuo) o bien $p_{X+Y} = p_X * p_Y$ (caso discreto), donde $*$ indica *convolución de funciones*. Más precisamente,

$$f_{X+Y}(t) = (f_X * f_Y)(t) = \int_{-\infty}^{+\infty} f_X(s) f_Y(t-s) ds \quad \text{(caso continuo)},$$

$$p_{X+Y}(n) = (p_X * p_Y)(n) = \sum_{m=-\infty}^{+\infty} p_X(m) p_Y(n-m) \quad \text{(caso discreto)}.$$

Tarea para pensar. La operación de convolución es asociativa, es decir, $(f*g)*h = f*(g*h)$, y conmutativa $f*g = g*f$. ¿Tiene elemento neutro (caso continuo, caso discreto)? ¿Tiene inversos?