IV. Funciones Generadoras de Momentos

M.A. Fernando Jesús Martínez Eissa

02S2019

ITAM
División Académica de Actuaría, Estadística y Matemáticas

Función Generadora de Momentos

Se define a la función generadora de momentos $M_X(t)$ de una variable aleatoria X como:

$$M_X(t) = E[e^{tX}] = \begin{cases} \sum_x e^{tX} p(x) & si\ x\ es\ discreta\ con\ masa\ p(x) \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & si\ x\ es\ continua\ con\ densidad\ f(x) \end{cases}$$

Para t tal que $|t| < b$ para alguna b>0

Se dice que función generadora de momentos no existe si:

- $E[e^{tx}]$ es de la forma ∞ o $\infty - \infty$; o
- No cumple la condición de que $|t| < b$ para alguna b>0

Esta última condición exige que $M_X(t)$ exista en alguna vecindad del cero (pudiendo no estar definida exactamente en t=0) y garantiza que las derivadas de $M_X(t)$ puedan ser evaluadas en t=0

Función Generadora de Momentos

Se llama función generadora de momentos por que todos los momentos de X (no centrales) pueden obtenerse de forma sucesiva diferenciando $M_X(t)$ y evaluando su resultado en t=0.

$$M_X'(t) = \frac{d}{dt} E[e^{tx}] = E\left[\frac{d}{dt} e^{tx}\right] = E[xe^{tx}]$$

Evaluando en t=0 obtenemos:

$$M_X{}'(0) = E[x]$$

De manera a similar, podemos calcular la segunda derivada:

$$M_X{}''(t) = \frac{d}{dt} E[xe^{tx}] = E\left[\frac{d}{dt} xe^{tx}\right] = E[x^2 e^{tx}]$$

Evaluando en t=0 obtenemos:

$$M_X{}''(0) = E[x^2]$$

Función Generadora de Momentos

De forma generalizada

$$M_X{}^{(k)}(t) = E[x^k e^{tx}] \; con \; k \geq 1$$

Lo que implica que:

$$M_X{}^{(k)}(0) = E[x^k] = \mu'_k \; con \; k \geq 1$$

Función Generadora de Momentos

Ejemplo 1

Se define un Experimento Bernoulli cuando su resultado sólo puede ser éxito con probabilidad p, $0<p<1$, o fracaso con probabilidad

$q = 1 - p$, independientemente de otros experimentos. Si la variable X toma valores 0 o 1 considerando el fracaso o éxito de un Experimento Bernoulli, respectivamente, entonces se dice que X tiene una Distribución Bernoulli.

a) Construya la *fmp* de la v.a. X

b) Calcule directamente la esperanza y la varianza de X

c) Calcule la FGM de X

d) Calcule la esperanza y la varianza de X utilizando la FGM

Función Generadora de Momentos

Ejemplo 1

a) Construya la *fmp* de la v.a. X

$$p_X(X) = \begin{cases} p, & x = 1 \\ 1-p, & x = 0 \end{cases}$$

b) Calcule directamente la esperanza y la varianza de X

$$E[X] = 1 * p + 0 * (1-p) = p$$

$$Var[X] = E[X^2] - (E[X])^2 = p - p^2 = p(1-p) = pq$$

c) Calcule la FGM de X

$$M_X(t) = E[e^{tx}] = \sum_x e^{tX} p(x) = e^0(1-p) + e^t p$$

d) Calcule la esperanza y la varianza de X utilizando la FGM

$$M_X'(0) = \frac{d}{dt} E[e^{tx}]\Big|_{t=0} = \frac{d}{dt}\left(e^t * p + e^0 * (1-p)\right)\Big|_{t=0} = e^t p\Big|_{t=0} = p$$

$$M_X^{(2)}(0) = \frac{d^2}{dt^2} E[e^{tx}]\Big|_{t=0} = \frac{d}{dt}(e^t * p)\Big|_{t=0} = e^t p\Big|_{t=0} = p$$

$$Var[X] = M_X^{(2)}(0) - \grave{M}_X(0) = p - p^2 = p(1-p) = pq$$

Función Generadora de Momentos

Ejemplo 2

Considere la variable aleatoria Y que representa el ingreso de las personas en cierta localidad. Una posible forma de estudiar el comportamiento de Y es suponiendo que tiene una Distribución Pareto con parámetros $\theta > 0$ y $\lambda > 0$, y con función de densidad de probabilidad:

$$f_Y(y) = \frac{\theta\lambda^\theta}{y^{\theta+1}} I_{[\lambda,\infty)}(y)$$

a) Realice un bosquejo de la gráfica de $f_Y(y)$
b) Demostrar que para esta v.a. No existe un valor b tal que $M_Y(t)$ sea finita para $|t| < b;\ \ b > 0$
c) Calcule la esperanza y la varianza de Y.

Binomial

Si $x \sim Bin(n,p)$, entonces

$$\begin{aligned} M(t) &= E[e^{tx}] \\ &= \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \\ &= (pe^t + 1 - p)^n \rightarrow \mathit{resultado\ del\ teorema\ binomial} \end{aligned}$$

Derivando $M'(t) = npe^t(pe^t + 1 - p)^{n-1}$

Evaluando en t=0

$$E[x] = M'^{(0)} = np$$

Derivando por segunda ocasión

$$M''^{(t)} = n(n-1)(pe^t + 1 - p)^{n-2}(pe^t)^2 + npe^t(pe^t + 1 - p)^{n-1}$$

Evaluando en t=0

$$E[x^2] = M''^{(0)} = n(n-1)p^2 + np$$

$$\therefore V[x] = E[x^2] - E^2[x] = np[(n-1)p + 1 - np] = npq$$

Si $x \sim Poisson(\lambda)$, entonces

$$M(t) = E[e^{tx}]$$
$$= \sum_{x=0}^{\infty} e^{tx} \frac{\lambda^x e^{-\lambda}}{x!}$$
$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$
$$= e^{-\lambda} e^{\lambda e^t} \rightarrow e^{\lambda(e^t-1)} resultado\ del\ teorema\ binomial$$

Derivando

$$M'(t) = \lambda e^t\left(e^{\lambda(e^t-1)}\right)$$
$$M''(t) = (\lambda e^t)^2 e^{\lambda(e^t-1)} + \left(\lambda e^t\left(e^{\lambda(e^t-1)}\right)\right)$$

Evaluando en t=0

$$M'(t) = \lambda$$
$$M''(t) = (\lambda)^2 + (\lambda)$$

$$\therefore V[x] = E[x^2] - E^2[x] = \lambda^2 + \lambda - \lambda^2 = \lambda$$

Exponencial

Si $x \sim Exp(\lambda)$, entonces

$$M(t) = E[e^{tx}]$$
$$= \int_0^\infty e^{tx} \lambda e^{-\lambda x} dx$$
$$= \lambda \int_0^\infty e^{-(\lambda - t)x} dx$$
$$= \frac{\lambda}{\lambda - t} \; para \; t < \lambda$$

Derivando

$$M'(t) = \frac{\lambda}{(\lambda - t)^2} \qquad M''(t) = \frac{2\lambda}{(\lambda - t)^3}$$

Evaluando en t=0

$$E[X] = M'(0) = \frac{1}{\lambda} \qquad E[X^2] = M''(0) = \frac{2}{\lambda^2}$$

$$\therefore V[x] = E[x^2] - E^2[x] = \frac{1}{\lambda^2}$$

Si $x \sim N(\mu = 0, \sigma^2 = 1)$, entonces

$$\begin{aligned} M_z(t) &= E[e^{tz}] \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-x^2/2} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-\frac{(x^2-2tx)}{2}} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}+\frac{t^2}{2}} dx \\ &= e^{\frac{t^2}{2}} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-t)^2}{2}} dx = e^{\frac{t^2}{2}} \end{aligned}$$

Dado lo anterior, para obtener la *FGM* de una variable aleatoria arbitraria normal, retomamos que $x=\mu+\sigma Z$ tendrá una distribución normal con parámetros μ y σ^2 siempre que Z sea una $N(0,1)$.

Entonces la FGM de una va normal es:

$$
\begin{aligned}
M_x(t) &= E[e^{tz}] \\
&= E\left[e^{t(\mu+\sigma Z)}\right] \\
&= E[e^{t\mu}e^{t\sigma Z}] \\
&= e^{t\mu}E[e^{t\sigma Z}] \\
&= e^{t\mu}M_z(t\sigma) \\
&= e^{t\mu}e^{\frac{(t\sigma)^2}{2}} \\
&= e^{\frac{\sigma^2 t^2}{2}+\mu t}
\end{aligned}
$$

Derivando $M'_x(t) = (\mu + t\sigma^2)e^{\frac{\sigma^2 t^2}{2}+\mu t}$

$$M''_x(t) = (\mu + t\sigma^2)^2 e^{\frac{\sigma^2 t^2}{2}+\mu t} + o^2 e^{\frac{\sigma^2 t^2}{2}+\mu t}$$

Igualando a cero $M'_x(0) = \mu$

$$M''_x(0) = \mu^2 + o^2$$

$$\therefore V[x] = E[x^2] - E^2[x] = \sigma^2$$

Condiciones de regularidad

Si X es una v.a. Cuya *f.m.p.* (caso discreto) o *f.d.p.* (caso continuo) posee propiedades tales que:

$$\frac{\partial}{\partial t}E[h(X,t)] = E\left[\frac{\partial}{\partial t}h(X,t)\right]$$

Es decir, que los operadores esperanza y derivada se puedan intercambiar, entonces se dice que prevalecen condiciones de regularidad en la distribución de X.

Función característica

Ante la dificultad de la no existencia de la FGM es posible definir a la *función característica*, que siempre existe, pero que se define en el campo de los números complejos ($\mathbb{C}$)

Definición

La función característica de la v.a. X se define por:

$$\Phi_X(t) = E\left[e^{-itX}\right], \text{ donde } i = \sqrt{-1} \in \mathbb{C}$$

La función característica es única y define por completo el comportamiento de la distribución de una variable aleatoria.

Si dos v.a.'s tienen la misma función característica, su distribución de probabilidades es la misma. Esta misma propiedad se puede establecer con la FGM (si existe) y se considerará mas adelante para determinar la distribución de una transformación de una variable aleatoria.

Momentos

Primer momento:

Corresponde a la media o al valor esperado de la variable aleatoria y corresponde al valor sobre el cual tienden a agruparse los datos.

Segundo momento:

Corresponde a la varianza, que es una medida de dispersión.

Tercer momento:

Corresponde al sesgo y se relaciona con la asimetría.

Cuarto momento:

Corresponde a la curtosis que nos indica qué tan “puntiaguda” es la función.

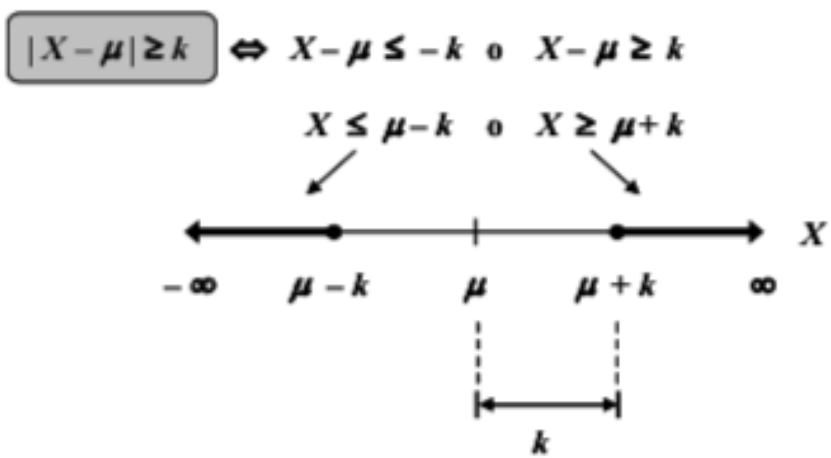
Desigualdad de Tchebysheff

Sea X una v.a. con media μ y varianza σ^2, ambas finitas, entonces para cualquier constante k>0, se tiene:

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Este teorema aplica para cualquier distribución de probabilidades, ya sea que esté acampanada o no.

Este resultado nos permite establecer una cota superior para la probabilidad de que la variable aleatoria X se aleje de su media en k unidades o más como se muestra a continuación:



Desigualdad de Tchebysheff

Dado:

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2}$$

Note que si $\frac{\sigma^2}{k^2} > 1$, la cota resulta poco útil pues $P[X \in A] \leq 1 \; \forall_{A \subset \mathbb{R}}$

La Desigualdad de Tchebysheff es muy útil en la Teoría Asintótica de la Estadística, particularmente para demostrar la ley Débil de los Grandes Números (uno de los principales resultados de la inferencia estadística). La desiguladad de Tcheysheff se puede demostrar fácilmente mediante la Desigualdad de Markov.

Desigualdad de Markov

Sea Z una v.a. no negativa, entonces para a>0:

$$P(Z \geq a) \leq \frac{E[Z]}{a}$$

Demostración

$$E[z] = \int_{-\infty}^{\infty} z f_Z(z) dz = \int_{-\infty}^{a} z f_Z(z) dz + \int_{a}^{\infty} z f_Z(z) dz$$

$$\int_{-\infty}^{a} z f_z(z) dz + \int_{a}^{\infty} z f_z(z) dz \geq \int_{a}^{\infty} z f_z(z) dz \geq \int_{a}^{\infty} a f_Z(z) dz$$

$$\therefore E[z] \geq a \int_{a}^{\infty} f_Z(z) dz = aP(Z \geq a)$$

Versiones de Tchebysheff

Una versión alternativa de la desigualdad de Tchebysheff permite acotar la probabilidad de la v.a. X se aleje r desviaciones estándar de su media (r>0). Si en la desigualdad de Tchebysheff $k=r\sigma$, entonces:

$$P(|X-\mu| \geq r\sigma) \leq \frac{1}{r^2}$$

Considerando la probabilidad del complemento en la desigualdad de Tchbysheff se obtiene una cota inferior para la probabilidad de que la variable aleatoria X se aleje de su media en menos de k unidades, como se muestra a continuación:

$$P(|X-\mu| \geq k) \leq \frac{\sigma^2}{k^2} \Longrightarrow 1 - P(|X-\mu| < k) \leq \frac{\sigma^2}{k^2} \Leftrightarrow P(|X-\mu| < k) \geq 1 - \frac{\sigma^2}{k^2}$$

Distancia entre la v.a. X y su media μ	**Cota superior**	**Cota inferior**
k unidades	$P[\lvert X-\mu\rvert \geq k] \leq \frac{\sigma^2}{k^2}$	$P[\lvert X-\mu\rvert < k] \geq 1-\frac{\sigma^2}{k^2}$
r desviaciones estándar	$P[\lvert X-\mu\rvert \geq r\sigma] \leq \frac{1}{r^2}$	$P[\lvert X-\mu\rvert < r\sigma] \geq 1-\frac{1}{r^2}$

Versiones de Tchebysheff

Ejemplo 3

En un cine se venden refrescos de 300 ml, sin embargo, la máquina que surte el refresco en ocasiones vierte más o menos refresco del debido. Suponga que X es la cantidad (en ml) de refresco adicional (X>0) o faltante (X<0) por vaso, y que su función de densidad de probabilidad es:

$$f_X(x) = \frac{1}{2} e^{-|x-2|} I_{(-\infty,\infty)}(x)$$

Para esta distribución, se puede demostrar que $\mu = E[x] = 2$ y que $\sigma^2 = Var[x] = 2.$

a) Calcular la probabilidad de que el refresco adicional o faltante se aleje del contenido promedio en más de 5 ml.

b) Utilice la desigualdad de Tchebysheff para acotar la probabilidad anterior.

c) ¿Qué probabilidad mínima se puede garantizar a los clientes de que el contenido de refresco por vaso está entre 290 y 314 ml?

d) ¿Cuál es la probabilidad mínima de que el contenido de refresco se aleje menos de 2σ de su μ?

Desigualdad de Jensen

Si $g(x)$ es una función convexa, entonces $E[g(X)] \geq g(E[X])$ considerando que los valores esperados existen.

Aplicaciones inmediatas:

- $E\left[\frac{1}{X}\right] \geq \frac{1}{E[X]}$ si $X > 0,\ ya\ que\ g(x) = \frac{1}{x}\ es\ convexa\ \forall_{x>0}$
- $E[X^2] \geq (E[X])^2 \Longrightarrow\ E[X^2] - (E[X])^2 \geq 0 \Longrightarrow Var[X] \geq 0\ \ ya\ que$ $g(x) = x^2\ es\ convexa\ \forall_{x>0}$

Note cómo si $g(x)$ es convexa, entonces $h(x) = -g(x)$ es función cóncava, entonces a partir de la desigualdad de Jensen es posible establecer lo siguiente:

Si $g(x)$ es una función cóncava, entonces $E[g(X)] \leq g(E[X])$, considerando que los valores esperados existen.

Desigualdad de Jensen

Si $g(x)$ es una función lineal, entonces $E[g(X)] = g(E[X])$ considerando que los valores esperados existen.

La proposición anterior, establece que el caso de igualdad de la Desigualdad de Jensen se alcanza cuando $g(x)$ es función lineal.

Demostración

Si $g(x)$ es lineal, entonces g es del estilo $g(x) = a + bx \; con \; a, b \in \mathbb{R}$. Evaluando g en $E[x]$, tenemos:

$$g(E[X]) = a + bE[X]$$

Por otro lado, se tiene:

$$E[g(X)] = E[a + bx] = E[a] + b[E(x)] = a + b[E(x)]$$

Por lo tanto, se concluye que:

$$E[g(x)] = g(E[x])$$

Ejemplo 4

Un inversionista tiene dos alternativas:

1. Invertir su dinero en un instrumento riesgoso que le proporcionará un rendimiento aleatorio X con media μ.
2. Invertir su dinero en instrumentos libres de riesgo que le proporcionan un rendimiento m con probabilidad 1.

Suponga que la decisión del inversionista la realiza maximizando el valor esperado de su función de utilidad $u(\cdot)$¿Qué alternativa debe escoger el inversionista si…

a) … $u(\cdot)$ es convexa?
b) … $u(\cdot)$ es cóncava?

V. Transformaciones de variables aleatorias

M.A. Fernando Jesús Martínez Eissa

02S2019

División Académica de Actuaría, Estadística y Matemáticas

Transformaciones discretas

Teorema

Sea X la v.a. Discreta con soporte x_1, x_2,…,y función de masa de probabilidad $f_X(x)$.

Si $Y = g(X), g: \mathbb{R} \longrightarrow \mathbb{R}$, es una transformación de la v.a. X, entonces:

$$f_Y(y) = \sum_{\{x_i : g(x_i) = y\}} f_X(x_i)$$

Este teorema establece que la $P[Y=y]$, para cada y del soporte de la v.a. discreta Y, se determina sumando las probabilidades $P[X=x_i]$ correspondientes a la o a las masas x_i, que hacen que $g(x_i) = y$

Transformaciones discretas

Ejemplo 5

Un vendedor de autos estima que las ventas de la siguiente semana para un modelo particular (X) tienen las siguientes probabilidades:

X	0	1	2	3
$P[X=x]$	0.2	0.3	0.4	0.1

El vendedor recibe un sueldo semanal de 4 mil pesos más un bono semanal si vende más de un auto. El monto del bono semanal, B(en miles de pesos) se determina por:

$$B = max\{0, bp(X-1)\} = \begin{cases} 0, & x < 1 \\ bp(X-1), & x \geq 1 \end{cases}$$

Donde b=% de bono semanal y p=precio de venta del coche (miles)

Si p es $200,000 y el % de bono semanal es 0.5%

a) Construya la función de masa de probabilidad del ingreso semanal del vendedor.

b) ¿Cuál es el ingreso esperado semanal del vendedor?

Transformaciones continuas

Para el caso continuo, se presentarán 3 casos para obtener la distribución de la v.a. $Y=g(X)$ a partir de la distribución de la v.a. X:

- Método de la *f.d.a*.
- Método de la transformación monótona
- Método de la *f.g.m*.

Método de la *f.d.a.*

Consiste en calcular la $F_Y(y)$ a partir de $F_X(x)$, donde $Y = g(X)$ y $F_X(x)$ es conocida. Una vez que se tiene $F_Y(y)$ es posible calcular

$$f_Y(y) = \frac{d}{dy} F_Y(y).$$

Transformaciones discretas

Ejemplo 6

Un productor cuenta con un proceso para refinar azúcar que le permite producir hasta 1.5 toneladas por día, pero la cantidad realmente producida X es una v.a. debido a fallas mecánicas y otras contingencias. Suponga que X tiene la función de densidad:

$$f_X(x) = \begin{cases} x & si\ 0 \leq x \leq 1 \\ 1 & si\ 1 < x \leq \dfrac{3}{2} \\ 0 & e.o.c. \end{cases}$$

Se sabe que el productor recibe 300 dólares por tonelada de azúcar refinada pero enfrenta costos fijos de 100 dólares por día.

a) Exprese la utilidad diaria del productor (en cientos de dólares).

b) Obtenga la función de distribución acumulada y la función de densidad de la utilidad diaria del productor.

Ejemplo 7

Si Y tiene la función de densidad de probabilidad:

$$f_Y(y) = (1 - |y|)I_{[-1,1]}(y)$$

a) Obtenga $f_U(u)$ si $U = U(Y) = Y^2$.

b) Indique cómo se relacionan $E[U(Y)]$ y $U(E[Y])$.

Transformaciones continuas

Método de la Transformación Monótona.

Teorema: Distribución de la Transformación Monótona (DTM)

Sea X una v.a. continua con función de densidad de probabilidad $f_X(x)$ y sea $g(x)$ una función estrictamente monótona (creciente o decreciente) y diferenciable (consecuentemente continua) en x. Si $Y = g(X)$ entonces:

i. $f_Y(y) = f_X\left(g^{-1}(y)\right)\left|\frac{d}{dy}g^{-1}(y)\right|$ si $y = g(x)$ para alguna x; y

ii. $f_Y(y) = 0$ si $y \neq g(x)$ para toda x,

Donde $g^{-1}(x)$ es el valor de x tal que $g(x) = y$

Transformaciones continuas

Ejemplo 8

Suponga que X es una v.a. continua con función de densidad $f_X(x)$
Y función de distribución acumulada $F_X(x)$. Obtenga $f_Z(z)$ si

$$Z = g(X) = \frac{1}{X + 1}$$

Transformaciones continuas

Ejemplo 9

Suponga que $X \sim U_{[0,1]}$, es decir, su función de densidad es $f_X(x) = I_{[0,1]}(x)$. Demuestre que $Y = -\beta ln(X), \beta > 0$, tiene una Distribución exponencial con media β.

Transformaciones continuas

Método de la $f.g.m.$

Teorema: Igualdad en distribución via $f.g.m.$

Si X y Y son v.a.'s entonces $M_X(t) = M_Y(t) \Leftrightarrow X \; y \; Y$ tienen la misma distribución de probabilidad.

Ejemplo 10

Se dice que X tiene una distribución normal con parámetros $\mu \in \mathbb{R}$ y $\sigma > 0$, si su función de densidad de probabilidad está dada por:

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

a) Demuestre que $M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$ para $t \in \mathbb{R}$.

b) Demuestre que $E[X] = \mu$ y que $Var[X] = \sigma^2$

c) A la transformación $Z = \frac{x-\mu}{\sigma}$ se le denomina estandarización de la v.a. x. Demuestre que Z tiene una distribución Normal Estándar considerando su fgm.