EYP1027 Modelos Probabilísticos

Clase 7

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Segundo Semestre 2020

Contenido I

Momentos y Funciones Generadoras

Recordemos que si X es una variable aleatoria en $(\Omega, \mathcal{A}, P)$ y g es una función con dominio y recorrido en los reales, entonces $g(X)$ también es una variable aleatoria en $(\Omega, \mathcal{A}, P)$, y su esperanza puede ser calculada como,

$$\mathrm{E}\{g(X)\} = \begin{cases} \sum_{x\in\mathcal{X}} g(x) f_X(x) & \text{si } X \text{ es discreta },\\ \int_{-\infty}^{\infty} g(x) f_X(x) dx & \text{si } X \text{ es continua },\end{cases}$$

siempre que exista la $\sum$ o la $\int$. Como fue dicho antes, si $\mathrm{E}(|g(X)|) = \infty$, decimos que $\mathrm{E}\{g(X)\}$ no existe.

En particular, la media $\mu_X := \mathrm{E}(X)$ y la varianza $\sigma_X^2 := \mathrm{E}\{(X-\mu_X)^2\}$ se obtienen para $g(x) = x$ y $g(x) = (x-\mu_X)^2$, respectivamente. Tales características constituyen casos especiales de los denominados momentos no centrados y centrados de la variable aleatoria X (o su distribución).

Nota: Si la función g es muy complicada, la evaluación de E$\{g(X)\}$ o Var$\{g(X)\}$ puede conducir a integraciones (o sumas) complicadas. De aqui que sean muy útiles las aproximaciones siguientes.

Teorema 1.1

Sea una variable aleatoria X con $\mathrm{E}(X) = \mu$, y $\mathrm{Var}(X) = \sigma^2$. Si g es una función a lo menos diferenciable dos veces en el punto $x = \mu$, entonces,

$$\mathrm{E}(Y) \simeq g(\mu) + \frac{g''(\mu)}{2}\sigma^2$$
$$\mathrm{Var}(Y) \simeq \left[g'(\mu)\right]^2 \sigma^2.$$

Los momentos de una distribución o variable aleatoria son una clase especial de esperanzas.

Definición 1.1

Sea X una variable aleatoria en $(\Omega, \mathcal{A}, P)$, y asuma que existen las sumatorias o integrales requeridas para cada entero positivo k:

i) El k-ésimo momento no centrado de X, se define como,

$$\mathrm{E}(X^k) = \begin{cases} \sum_{x \in \mathcal{X}} x^k f_X(x) & \text{si } X \text{es discreta,} \\ \int_{-\infty}^{\infty} x^k f_X(x) dx & \text{si } X \text{es continua.} \end{cases}$$

$k = 1 \Longrightarrow \mathrm{E}(X) = \mu_X$ es la media de X.

ii) El k-ésimo momento centrado de X, se define como,

$$\mathrm{E}\{(X - \mu_X)^k\} = \begin{cases} \sum_{x \in \mathcal{X}} (x - \mu_X)^k f_X(x) & \text{si } X \text{es discreta,} \\ \int_{-\infty}^{\infty} (x - \mu_X)^k f_X(x) dx & \text{si } X \text{es continua.} \end{cases}$$

$k = 2 \Longrightarrow \mathrm{E}\{(X - \mu_X)^2\} = \sigma_X^2$ es la varianza de X.

Momentos y Funciones generadoras

Definición 1.2

A parte de la media $\mu_X = \mathrm{E}(X)$ y la varianza $\sigma_X^2 = \mathrm{E}\{(X-\mu_X)^2\}$, las que son medidas de localización y dispersión de la distribución X, respectivamente, otras dos carasterísticas de gran interés son los índices de asimetría (γ_X) y curtosis (κ_X) de la distribución de X:

a) $\gamma_X = \mathrm{E}\left\{\left(\dfrac{X-\mu_X}{\sigma_X}\right)^3\right\} \leftarrow$ asimetría de la distribución de X,

a) $\kappa_X = \mathrm{E}\left\{\left(\dfrac{X-\mu_X}{\sigma_X}\right)^4\right\} \leftarrow$ curtosis de la distribución de X.

Note que tales índices coresponden al los momentos tercero y cuarto, respectivamente, de la variable aleatoria estandarizada $Z = \dfrac{X-\mu_X}{\sigma_X}$, la cual es tal que $\mathsf{E}(Z) = 0$ y $\mathsf{Var}(Z) = 1$.

De forma similar se definen $\mathsf{E}(|X|^k)$ y $\mathsf{E}(|X-\mu_X|^k)$, llamados momentos absolutos no centrado y centrado de orden k, respectivamente.

Teorema 1.2

a) $|\mathrm{E}(X^k)| \leq \mathrm{E}(|X|^k)$.

b) Si $\mathrm{E}(X^k)$, $k>0$, es finito, entonces $\mathrm{E}\{(X-\mu_X)^k\}$ es finito.

c) Si $\mathrm{E}(X^k)$, $k>0$, es finito, entonces $\mathrm{E}(X^k)$ es finito $\forall\, 0<m<k$.

d) Desigualdad de Markov: $P(|X|\geq\lambda)\leq\frac{\mathrm{E}(|X|^k)}{\lambda^k}$ $\forall\,\lambda>0$ y $\forall\, k>0$. En particular, si la varianza de X es finita, entonces,

$$P(|X-\mathrm{E}(X)|\geq\varepsilon)\leq\frac{\mathrm{Var}(X)}{\varepsilon^2}\quad\forall\,\varepsilon>0.$$

e) Si X es una variable aleatoria arbitraria, con k-ésimo momento finito, entonces, para $k=1,2,\ldots$, se tiene que,

$$\mathrm{E}(X^k)=k\left\{\int_0^\infty x^{k-1}\left(1-F_X(x)\right)dx-\int_{-\infty}^0 x^{k-1}F_X(x)dx\right\}.$$

Demostración 1.1

Demostraremos sólo la propiedad c) asumiendo que X es continua; luego,

$$\begin{aligned}
\mathrm{E}(|X|^k) &= \int_{-\infty}^{\infty} |x|^k f_X(x)dx \\
&= \int_{|x|<\lambda}^{\infty} |x|^k f_X(x)dx + \int_{|x|\geq\lambda}^{\infty} |x|^k f_X(x)dx \quad \forall\, \lambda > 0 \\
&\geq \int_{|x|\geq\lambda}^{\infty} |x|^k f_X(x)dx \quad \forall\, \lambda > 0 \\
&\geq \int_{|x|\geq\lambda}^{\infty} \lambda^k f_X(x)dx = \lambda^k \int_{|x|\geq\lambda}^{\infty} f_X(x)dx \quad \forall\, \lambda > 0.
\end{aligned}$$

Dividiendo en ambos lados por λ^k, se obtiene el resultado. La demostración cuando X discreta es similar.

Ejemplo 1.1

Bajo ciertas condiciones la tensión superficial de un líquido (dina/cm) está dada por $S = 2(1 - 0.005T)^{1.2}$ donde T es la temperatura del líquido (grados centigrados). Suponga que T es una variable aleatoria continua con la siguiente fdp:

$$f_T(t) = \begin{cases} 3000t^{-4}, & t \geq 10 \\ 0, & \text{eoc.} \end{cases}$$

Luego,

$$\begin{aligned} \mathrm{E}(T) &= \int_{10}^{\infty} 3000t^{-3}dt = 15 \text{ (grados centigrados), y} \\ \mathrm{Var}(T) &= \mathrm{E}\left(T^2\right) - (15)^2 \\ &= \int_{10}^{\infty} 3000t^{-2}dt - 225 = 75 \text{ (grados centigrados)}^2. \end{aligned}$$

Para calcular $\mathrm{E}(S)$ y $\mathrm{Var}(S)$ hay que calcular las integrales siguientes:

$$\int_{10}^{\infty}(1-0.005t)^{1.2}t^{-4}dt$$
$$\int_{10}^{\infty}(1-0.005t)^{2.4}t^{-4}dt$$

En vez de evaluar esas expresiones, se obtendrán aproximaciones para $\mathrm{E}(S)$ y $\mathrm{Var}(S)$ usando el Teorema 1.1. Para poder usar esas fórmulas hay que calcular $g'(15)$ y $g''(15)$, en donde $g(t) = 2(1-0.005t)^{1.2}$. Para la primera derivada, se tiene que,

$$g'(t) = 2.4(1-0.005t)^{0.2}(-0.005) = -0.012(1-0.005t)^{0.2}.$$

Luego

$$g(15) = 1.82 \text{ y } g'(15) = 0.01.$$

De un modo semejante,

$$g''(t) = -0.0024(1 - 0.005t)^{-0.8}(-0.005) = 0.000012(1 - 0.005t)^{-0.8}.$$

Por lo tanto,

$$g''(15) = \frac{0.000012}{(0.925)^{0.8}} = 0^{+}.$$

Así, se tiene que,

$$\begin{aligned} \mathrm{E}(S) &\simeq g(15) + \tfrac{1}{2}75g''(15) = 1.82 \text{ (dinas /cm)} \\ \mathrm{Var}(S) &\simeq 75\left[g'(15)\right]^2 = 0.87 \text{ (dinas/cm) }^2. \end{aligned}$$

Ejemplo 1.2

Suponga que X es una variable aleatoria continua con fdp,

$$f_X(x) = \begin{cases} 1+x, & \text{si } \; -1 \leq x \leq 0, \\ 1-x, & \text{si } \; 0 \leq x \leq 1, \\ 0, & \text{si no.} \end{cases}$$

Como $|X| \leq 1$ (con probabilidad 1), entonces, para todo $k > 0$, se tiene que $|X|^k \leq 1$ (con probabilidad 1). Luego, todos los momentos de X (centrados o no) son finitos. Además, debido a la simetría de $f_X(\cdot)$ en torno al cero, es decir, $f_X(-x) = f_X(x)$ para todo x, se tiene que $\mu_X = \mathrm{E}(X) = 0$, de modo que los momentos centrados y no centrados coinciden. La simetría de X también implica que los momentos de orden impar son todos nulos, es decir, $\mathrm{E}(X^{2k-1}) = 0$ para todo $k = 1, 2, \ldots$. Consecuentemente, $\gamma_X = 0$.

Para los momentos de orden par, se tiene que,

$$
\begin{aligned}
\mathrm{E}(X^{2k}) &= \int_{-1}^{0} x^{2k}(1+x)dx + \int_{0}^{1} x^{2k}(1-x)dx \\
&= \int_{-1}^{0} x^{2k}dx + \int_{-1}^{0} x^{2k+1}dx + \int_{0}^{1} x^{2k}dx - \int_{0}^{1} x^{2k+1}dx \\
&= 2\int_{0}^{1} y^{2k}dy - 2\int_{0}^{1} y^{2k+1}dy \quad (y=-x) \\
&= \frac{1}{(k+1)(2k+1)}, \quad k = 1, 2, \ldots.
\end{aligned}
$$

Así, $k=1 \Longrightarrow \mathrm{Var}(X) = \mathrm{E}(X^2) = 1/6$ y $k=2 \Longrightarrow \kappa_X = 36/15 = 2.4$.

Ejemplo 1.3

Sea X = suma de valores en las caras superiores de dos dados arrojados al azar. Aquí, $\mathcal{X} = \{2, 3, ..., 12\}$, con fmp dada por,

$$f_X(x) = P(X = x) = \begin{cases} \frac{1}{36} \times \text{mín}(x-1, 13-x), & \text{si } x \in \mathcal{X}, \\ 0, & \text{si no.} \end{cases}$$

Luego,

i) $\mu_X = \frac{1}{36} \sum_{x=2}^{12} x \times \text{mín}(x-1, 13-x) = 7.$

ii) $\sigma_X^2 = \frac{1}{36} \sum_{x=2}^{12} (x-7)^2 \times \text{mín}(x-1, 13-x) = 5.833$

iii) $\text{E}\{(X-7)^3\} = \frac{1}{36} \sum_{x=2}^{12} (x-7)^3 \times \text{mín}(x-1, 13-x) = 0$. Luego $\gamma_X = 0$, y la distribución es simétrica.

iv) $\text{E}\{(X-7)^4\} = \frac{1}{36} \sum_{x=2}^{12} (x-7)^4 \times \text{mín}(x-1, 13-x) = 80.5$ Luego $\kappa_X = 80.5/2.415^4 = 2.366$.

La distribución de probabilidades de una variable aleatoria X también puede ser caracterizada mediante la denominada función generadora de momento (fgm).

La existencia de la fgm determina la existencia de los momentos de la variable aleatoria X, y como su nombre lo indica, ella genera los momentos de X.

Aunque en muchos casos puede ser más fácil calcular los momentos directamente de su definición, la utilidad de la fgm no es sólo generar momentos, sino que también ayudar a caracterizar una distribución.

En otras palabras, cada distribucíon tiene asociada una única fgm (provisto que exista) y viceversa.

Definición 1.3

Función generadora de momentos: Sea X una variable aleatoria con fda F_X. La función generadora de momento (fgm) de X (o F_X), denotada por $M_X(t)$, se define como,

$$M_X(t) = \mathrm{E}(e^{tX}),$$

siempre que exista la esperanza para t en alguna vecindad de 0. Es decir, existe $h > 0$ tal que, para todo t en $(-h,\ h)$, $\mathrm{E}(e^{tX})$ exista.

Si la esperanza no existe en una vecindad de 0, decimos que la fgm de X no existe.

Momentos y Funciones generadoras

Más explicitamente, la fgm de X se calcula como,

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

Para ver cómo la fgm genera los momentos de X, primero note que si $|t| < h$, algún $h > 0$, entonces,

$$e^{tx} = \sum_{k=0}^{\infty} \frac{t^k}{k!} x^k \quad \text{(converge para todo } x\text{)}$$

$$\Longrightarrow M_X(t) = \mathsf{E}\left(e^{tX}\right) = \mathsf{E}\left(\sum_{k=0}^{\infty} \frac{t^k}{k!} X^k\right) = \sum_{k=0}^{\infty} \frac{t^k}{k!} \mathsf{E}(X^k)$$

$$= 1 + \frac{t}{1!}\mathsf{E}(X) + \frac{t^2}{2!}\mathsf{E}(X^2) + \cdots + \frac{t^k}{k!}\mathsf{E}(X^k) + \frac{t^{k+1}}{(k+1)!}\mathsf{E}(X^{k+1})$$

Dirivando en ambos lados k veces con respecto a t, y evaluando en el resultado en $t = 0$, se obtiene el siguiente teorema.

Teorema 1.3

Si X tiene fgm $M_X(t)$, entonces,

$$\mathrm{E}(X^k) = M_X^{(k)}(0),$$

donde,

$$M_X^{(k)}(0) = \frac{d^k}{dt^k} M_X(t)\,|_{t=0}$$

Es decir, el k-ésimo momento de X es igual a la k-ésima derivada de $M_X(t)$ evaluada en $t = 0$.

A continuación, se entrega una demostración alternativa de este resultado.

Demostración 1.2

Suponiendo que podemos diferenciar bajo el signo integral, tenemos

$$\begin{aligned}
\frac{d}{dt} M_X(t) &= \frac{d}{dt} \int_{-\infty}^{\infty} e^{tx} f_X(x) dx \\
&= \int_{-\infty}^{\infty} \left(\frac{d}{dt} e^{tx} \right) f_X(x) dx \\
&= \int_{-\infty}^{\infty} \left(x e^{tx} \right) f_X(x) dx \\
&= \mathrm{E}(X e^{tX}).
\end{aligned}$$

Así,

$$\frac{d}{dt} M_X(t)|_{t=0} = \mathrm{E}(X e^{tX}|_{t=0}) = \mathrm{E}(X).$$

Procediendo de manera análoga, podemos probar que,

$$\frac{d^k}{dt^k} M_X(t)|_{t=0} = \mathrm{E}(X^k e^{tX}|_{t=0}) = \mathrm{E}(X^k).$$

Ejemplo 1.4

Sea X una variable aleatoria con fdp dada por,

$$f_X(x) = \begin{cases} 2e^{-2x} & \text{si } x \geq 0 \\ 0 & \text{eoc} \end{cases}$$

Luego,

$$\begin{aligned} M_X(t) &= \int_0^\infty e^{tx} 2e^{-2x} dx \\ &= (1 - t/2)^{-1}, \quad \text{con } t < 2. \end{aligned}$$

Ejemplo 1.5

Se dice que una variable aleatoria continua X tiene distribución normal estándar, denotada por $X \sim N(0,1)$, si la fdp X esta dada por,

$$f_X(x) := \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}, \quad -\infty < x < \infty.$$

Note que $\phi(-x) = \phi(x) \; \forall \, x$, de modo que la distribución $N(0,1)$ es símétrica con respecto al 0; luego, si sus momentos impares existen, todos ellos deben ser nulos. En este caso, para obtener el k-ésimo momento de X, hay que resolver la integral,

$$\mathrm{E}(X^k) = \int_{-\infty}^{\infty} x^k \phi(x) dx$$

para lo cual hay que integrar por parte en forma recursiva.

En cambio, al notar que $e^{tx}\phi(x) = e^{t^2/2}\phi(x-t)$, el cálculo de la fgm de la distribución normal estandar es bastante simple, ya que,

$$\begin{aligned} M_X(t) = \mathrm{E}(e^{tX}) &= \int_{-\infty}^{\infty} e^{tx}\phi(x)dx \\ &= \int_{-\infty}^{\infty} e^{t^2/2}\phi(x-t)\phi(x)dx \\ &= e^{t^2/2}\int_{-\infty}^{\infty} \phi(y)dy \quad (y = x-t) \\ &= e^{t^2/2}, \end{aligned}$$

la cual existe para todo $t \in \mathbb{R}$. Derivar esta función es bastante simple, como se ilustra a continuación.

Momentos y Funciones generadoras

Por ejemplo, para la primera y segunda derivada de $M_X(t) = e^{t^2/2}$, es decir, de la fgm de $X \sim N(0,1)$, se tiene que,

$$M'_X(t) = tM_X(t) \implies \mathrm{E}(X) = M'_X(0) = 0,$$
$$M''_X(t) = (1+t^2)M_X(t) + tM'_X(t) \implies \mathrm{E}(X^2) = M''_X(0) = 1$$
$$\implies \mathrm{Var}(X) = \mathrm{E}(X^2) - \{\mathrm{E}(X)\}^2 = 1 - 0^2 = 1.$$

Tarea: Verifique que $\gamma_X = \mathrm{E}(X^3) = 0$ (simetría) y $\kappa_X = \mathrm{E}(X^4) = 3$.

Resumiendo,

$$X \sim N(0,1) \implies \begin{cases} \mu_X = 0 & \text{(media)}, \\ \sigma^2_X = 1 & \text{(varianza)}, \\ \gamma_X = 0 & \text{(asimetría)}, \\ \kappa_X = 3 & \text{(curtosis)}, \end{cases}$$

Ejemplo 1.6

Sea X una variable aleatoria discreta con valores en $\mathcal{X} = \{x_1, ..., x_m\}$. Entonces,

$$M_X(t) = \sum_{i=1}^{m} e^{tx_i} P(X = x_i).$$

Si $\mathcal{X} = \{-1, 1\}$, y $P(X = -1) = P(X = 1) = \frac{1}{2}$, entonces,

$$M_X(t) = \frac{1}{2}\{e^{-t} + e^{t}\},$$

para todo $t \in \mathbb{R}$. La media de X está dada por,

$$\mathrm{E}X = \frac{d}{dt} M_X(t)|_{t=0} = \frac{1}{2}\{-e^{-t} + e^{t}\}|_{t=0} = 0.$$

No unicidad de los momentos. Dos variables aletorias, X e Y con distintas fdp, pueden tener los mismos momentos, a menos que tengan soporte acotado. Por otro lado, una importante propiedad de la fgm de una variable aleatoria es que, cuando existe, caracteriza la distribución de la variable. El siguiente teorema resume estos resultados.

Teorema 1.4

Sean $F_X(x)$ y $F_Y(y)$ dos fda cuyos momentos existen.

i) Si X e Y tienen soporte acotado, entonces $F_X(u) = F_Y(u)$ para todo u si y sólo si $\mathrm{E}(X^r) = \mathrm{E}(Y^r)$ para todo $r = 0, 1, 2, \ldots$.

ii) Si la fgm existe y $M_X(t) = M_Y(t)$ para todo t en alguna vecindad de 0, entonces $F_X(u) = F_Y(u)$ para todo u.

Demostración 1.3

Ver Casella & Berger (2002).

Teorema 1.5

Para constantes cualquiera a y b, la fgm de la variable aleatoria $aX + b$ está dada por,

$$M_{aX+b}(t) = e^{bt} M_X(at).$$

Demostración 1.4

Por definición,

$$\begin{aligned} M_{aX+b}(t) &= \mathrm{E}(e^{(aX+b)t}) = \mathrm{E}(e^{(aX)t} e^{bt}) \\ &= e^{bt} \mathrm{E}(e^{(at)X}) \\ &= e^{bt} M_X(at). \end{aligned}$$

Ejemplo 1.7

Sabemos que $X \sim N(0,1) \iff M_X(t) = e^{t^2/2}$, $t \in \mathbb{R}$. Entonces, por el Teorema 1.5, la fgm de $Y = \sigma X + \mu$, donde $\sigma > 0$ (parámetro de escala) y $\mu \in \mathbb{R}$ (parámetro de localización), esta dada por,

$$M_Y(t) = e^{t\mu} e^{t^2\sigma^2/2} = e^{t\mu + t^2\sigma^2/2}, \quad t \in \mathbb{R}.$$

Tal fgm caracteriza la denominada distribución normal con media μ y varianza σ^2, denotada por $Y \sim N(\mu, \sigma^2)$. Recordando que $\mathrm{E}(X) = 0$ y $\mathrm{Var}(X) = 1$, se tiene que

$$\mathrm{E}(Y) = \sigma \mathrm{E}(X) + \mu = \mu \quad \text{y} \quad \mathrm{Var}(Y) = \sigma^2 \mathrm{Var}(X) = \sigma^2.$$

Además, es posible probar que la fdp de $Y = \sigma X + \mu$ está dada por

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}, \quad -\infty < y < \infty.$$

References

Blanco, L., Arunachalam, V. y Dharmaraja, S. (2012). *Introduction to Probability and Stochastic Processes.* John Wiley, New Jersey.

Casella, G. y Berger, R.L. (2002). *Statistical Inference.* Second Edition. Duxbury, California.