VARIABLE ALEATORIA

FELIPE OSORIO

1. Conceptos Preliminares

Considere el espacio de probabilidad $(\Omega, \mathcal{F}, \mathsf{P})$. Recuerde que Ω representa el conjunto de todos los resultados posibles de un experimento aleatorio, el que se denomina espacio muestral. El objetivo es convertir cada resultado de un experimento aleatorio en un número, lo que motiva la siguiente definición.

Definición 1.1. Una variable aleatoria X es una función $X : \Omega \to \mathbb{R}$ que transforma un resultado $\omega \in \Omega$ a un número $X(\omega)$ en la recta real.

Ejemplo 1.2. Suponga que se lanza una moneda dos veces. En este caso el espacio muestral Ω es
$$\Omega = \{(C, C), (C, S), (S, C), (S, S)\}.$$
Suponga que X es una variable aleatoria representando la suma de las caras, es decir $X(\cdot) :=$ número de caras. Tenemos que el espacio muestral tiene 4 elementos, a saber
$$\omega_1 = (C, C), \quad \omega_2 = (C, S), \quad \omega_3 = (S, C), \quad \omega_4 = (S, S),$$
luego,
$$X(\omega_1) = 2, \quad X(\omega_2) = 1, \quad X(\omega_3) = 1, \quad X(\omega_4) = 0.$$
En este caso, podemos notar que $X(\omega)$ no es una función uno-a-uno.

Para relacionar una variable aleatoria con el concepto de probabilidad, definiremos el evento,
$$A = \{\omega \in \Omega : X(\omega) = a\},$$
este es el conjunto de todos los resultados posibles en Ω, es decir, todos los ω's tal que $X(\omega) = a$. Deseamos conocer el "tamaño"de A. En efecto, diremos que determinar la probabilidad de $X = a$, es equivalente a calcular la probabilidad de $A = \{\omega \in \Omega : X(\omega) = a\}$. En efecto, debemos hallar la preimagen $X^{-1}(a)$ definida como
$$X^{-1}(a) := \{\omega \in \Omega : X(\omega) = a\}.$$
Es decir,
$$\mathsf{P}(X = a) = \mathsf{P}(X^{-1}(a)) = \mathsf{P}(\{\omega \in \Omega : X(\omega) = a\}).$$
En nuestro ejemplo del lanzamiento de la moneda, tenemos:
$$X^{-1}(0) = \{\omega_4\}, \qquad X^{-1}(1) = \{\omega_2, \omega_3\}, \qquad X^{-1}(2) = \{\omega_1\}.$$
De este modo,
$$\mathsf{P}(X = 1) = \mathsf{P}(X^{-1}(1)) = \mathsf{P}(\{\omega_2, \omega_3\}).$$

Ejemplo 1.3. Suponga que se lanzamos un dado. En este caso el espacio muestral es dado por
$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$
De este modo, podemos considerar
$$\begin{aligned} X(1) &= 1, & X(2) &= 2, & X(3) &= 3, \\ X(4) &= 4, & X(5) &= 5, & X(6) &= 6. \end{aligned}$$
Así, por ejemplo
$$\begin{aligned} \mathsf{P}(X \leq 3) &= \mathsf{P}(X = 1) + \mathsf{P}(X = 2) + \mathsf{P}(X = 3) \\ &= \mathsf{P}(X^{-1}(1)) + \mathsf{P}(X^{-1}(2)) + \mathsf{P}(X^{-1}(3)) \\ &= \mathsf{P}(\{1\}) + \mathsf{P}(\{2\}) + \mathsf{P}(\{3\}) \\ &= \frac{1}{3} + \frac{1}{3} + \frac{1}{3} = \frac{1}{2}. \end{aligned}$$
Evidentemente podemos definir probabilidades de forma más general, como
$$\mathsf{P}(X \in B) = \mathsf{P}(\{\omega \in \Omega : X(\omega) \in B\}),$$
y en particular
$$\mathsf{P}(X \leq x) = \mathsf{P}(\{\omega \in \Omega : X(\omega) \leq x\}).$$

Definición 1.4. La *función de distribución acumulada* (CDF) de una variable aleatoria X, denotada por $F_X(x)$, es definida por
$$F_X(x) = \mathsf{P}(X \leq x).$$
El conjunto de todos los resultados posibles de X es denotado como $\mathcal{X} := X(\Omega)$, y será llamado *espacio muestral*.

Propiedad 1.5. La función $F(x)$ es una CDF si y sólo si, se satisfacen las condiciones

(i) $\lim_{x\to-\infty} F(x) = 0$, y $\lim_{x\to+\infty} F(x) = 1$.

(ii) Para $x \leq y \Rightarrow F(x) \leq F(y)$, es decir $F(x)$ es función no decreciente.

(iii) $F(x)$ es continua a la derecha, esto es, para todo x_0, tenemos
$$\lim_{x\to x_0^+} F(x) = F(x_0).$$

Demostración. Considere $x \leq y \Rightarrow \{X \leq x\} \subseteq \{Y \leq y\}$, lo que lleva a $\mathsf{P}(X \leq x) \leq \mathsf{P}(Y \leq y)$, es decir $F(x) \leq F(y)$, lo que muestra (ii).

Para notar (iii), considere una secuencia decreciente $x_n \downarrow x$, entonces $\{X \leq x_n\}$ es una secuencia decreciente de eventos, luego
$$\bigcap_{n=1}^{\infty} \{X \leq x_n\} = \{X \leq x\}.$$
Es decir, $\{X \leq x_n\} \downarrow \{X \leq x\}$ y por la continuidad de la función de probabilidad, sigue que
$$F(x_n) = \mathsf{P}(X \leq x_n) \downarrow \mathsf{P}(X \leq x) = F(x).$$
Finalmente, para mostrar (i) note que para $x_n \downarrow -\infty$, entonces
$$\{X \leq x_n\} \downarrow \varnothing \quad \Rightarrow \quad F(x_n) = \mathsf{P}(X \leq x_n) \downarrow 0.$$

Mientras que, si $x_n \uparrow \infty$, entonces

$$\{X \leq x_n\} \uparrow \Omega \quad \Rightarrow \quad F(x_n) = \mathsf{P}(X \leq x_n) \uparrow 1.$$

$\square$

Observación 1.6. Cualquier función $F : \mathbb{R} \to [0,1]$ satisfaciendo las propiedades anteriores corresponde a una CDF.

Ejemplo 1.7. Suponga que se lanza una moneda hasta que obtenemos una *cara*. Defina la variable aleatoria

$$X : \text{número de lanzamientos hasta obtener una cara.}$$

Sea p la probabilidad de obtener una cara en *un único lanzamiento* de la moneda ($0 < p < 1$). Entonces, para $x = 1, 2, \dots,$

$$\mathsf{P}(X = x) = (1-p)^{x-1}p,$$

pues debemos obtener $x - 1$ *sellos* antes que aparezca la primera *cara*. De este modo,

$$\mathsf{P}(X \leq x) = \sum_{k=1}^{x} P(X = k) = \sum_{k=1}^{x} (1-p)^{k-1}p,$$

es decir,

$$\begin{aligned} \mathsf{P}(X \leq x) &= p\sum_{k=1}^{x}(1-p)^{k-1} = p\{1 + (1-p) + (1-p)^2 + \cdots + (1-p)^{x-1}\} \\ &= p\Big(\frac{1-(1-p)^x}{1-(1-p)}\Big) = 1 - (1-p)^x, \end{aligned}$$

para $x \in \{1, 2, \dots\}$. Es fácil notar que $F(x)$ para $p \in (0,1)$ es una CDF. En efecto,

$$\lim_{x \to -\infty} F(x) = 0,$$

pues $F(x) = 0$ para $x < 0$, y

$$\lim_{x \to +\infty} F(x) = \lim_{x \to +\infty} 1 - (1-p)^x = 1.$$

Para verificar que $F(x)$ es no decreciente, basta verificar que la suma en $F(x)$ contiene más términos positivos conforme x aumenta. (Además, para todo x, $F(x + \epsilon) = F(x)$ si $\epsilon > 0$ es suficientemente pequeño sigue que

$$\lim_{\epsilon \to 0^+} F(x + \epsilon) = F(x),$$

es continua por la derecha).

Observación 1.8. Evidentemente, podemos escribir:

$$F_X(x) := \mathsf{P}_X((-\infty, x]) = \mathsf{P}(X \leq x).$$

Definición 1.9 (σ-álgebra generada)**.** Sea $\mathcal{C}$ una colección de conjuntos en Ω. Entonces $\sigma(\mathcal{C})$, el σ-álgebra generada por $\mathcal{C}$, es la σ-álgebra más pequeña $\mathcal{F}$ definida en Ω tal que $\mathcal{C} \subset \mathcal{F}$.

Definición 1.10 (σ-álgebra de Borel)**.** Considere el intervalo semi-cerrado,
$$(-\infty, b] := \{x : -\infty < x \leq b\}.$$
El σ-álgebra de Borel, denotada por $\mathcal{B}$ es una σ-álgebra generada a partir de los intervalos semi-cerrados $(-\infty, b]$. Es decir,
$$\mathcal{B} = \sigma((-\infty, b]).$$

Observación 1.11. La σ-álgebra de Borel contiene elementos de la forma:
$$(a, b), [a, b], (a, b], [a, b), [a, \infty), (a, \infty), (-\infty, b], \{b\}.$$
Así que escribimos $\mathcal{B} = \sigma(\mathbb{R})$.

Resultado 1.12. *Suponga que X es una variable aleatoria definida sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathsf{P})$. La función $F_X(\cdot)$ definida sobre la σ-álgebra de Borel, como:*
$$\mathsf{P}_X(B) = \mathsf{P}(\{X \in B\}), \qquad \forall\, B \in \mathcal{B},$$
es una probabilidad sobre $(\mathbb{R}, \mathcal{B})$ y es llamada distribución *de la variable X.*

Ejemplo 1.13. Considere la función,
$$F(x) = \frac{1}{1 + e^{-x}},$$
que satisface las condiciones anteriores, pues
$$\lim_{x \to -\infty} e^{-x} = \infty \quad \Rightarrow \quad \lim_{x \to -\infty} F(x) = 0$$
$$\lim_{x \to +\infty} e^{-x} = 0 \quad \Rightarrow \quad \lim_{x \to +\infty} F(x) = 1.$$

Note que, diferenciando $F(x)$ obtenemos
$$\frac{\mathsf{d}\, F(x)}{\mathsf{d}\, x} = \frac{e^{-x}}{(1 + e^{-x})^2} > 0,$$
es decir, $F(x)$ es creciente. Adicionalmente, notamos que $F(x)$ es función continua.

Observación 1.14. Una variable aleatoria X se dice *continua* si $F_X(x)$ es continua. X es *discreta* si $F_X(x)$ es una función escalón de x.

Definición 1.15. Las variables aleatorias X e Y son *idénticamente distribuídas* si para todo $A \in \mathcal{B}$,
$$\mathsf{P}(X \in A) = \mathsf{P}(Y \in A),$$
en cuyo caso anotamos $X \stackrel{\mathsf{d}}{=} Y$.

Resultado 1.16. *Las siguientes dos condiciones son equivalentes:*
(i) *Las variables aleatorias X e Y son idénticamente distribuídas.*
(ii) *$F_X(x) = F_Y(x)$ para todo x.*

Demostración. Probaremos solamente que (i) $\Rightarrow$ (ii). Como $X \stackrel{\mathsf{d}}{=} Y$ tenemos que para $A \in \mathcal{B}$,
$$\mathsf{P}(X \in A) = \mathsf{P}(Y \in A),$$
y en particular para $(-\infty, x] \in \mathcal{B}$, es decir
$$F_X(x) = \mathsf{P}(X \in (-\infty, x]) = \mathsf{P}(Y \in (-\infty, x]) = F_Y(x).$$
□

Ejemplo 1.17. Sea $(\Omega, \mathcal{F}, \mathsf{P})$ un espacio de probabilidad y $A \in \mathcal{F}$. La función $I_A : \Omega \to \mathbb{R}$, dada por

$$I_A(\omega) = \begin{cases} 1, & \text{si } \omega \in A, \\ 0, & \text{si } \omega \notin A, \end{cases}$$

es una variable aleatoria (real). En efecto, para $\alpha \in \mathbb{R}$,

$$\{\omega \in \Omega : I_A(\omega) \leq \alpha\} = \begin{cases} \varnothing, & \alpha < 0, \\ A^c, & 0 \leq \alpha < 1, \\ \Omega, & \alpha \geq 1. \end{cases}$$

Definición 1.18. Una variable aleatoria es discreta si $\mathcal{X} = X(\Omega)$ es un conjunto finito o numerable. Es decir, $X(\omega) \in \mathcal{X} = \{x_1, x_2, \dots\}$ para todo $\omega \in \Omega$. La función,

$$p_X(x) = \mathsf{P}(X = x), \qquad \forall x \in \mathcal{X},$$

es llamada *función de (masa de) probabilidad* (pmf) de X.

Note que, para X variable aleatoria discreta, tenemos

$$\{X \leq x\} = \bigcup_{i:\, x_i \leq x} \{X \leq x_i\},$$

luego

$$\begin{aligned} F_X(x) = \mathsf{P}(X \leq x) &= \mathsf{P}\Big(\bigcup_{i:\, x_i \leq x} \{X \leq x_i\}\Big) \\ &= \sum_{i:\, x_i \leq x} \mathsf{P}(X = x_i) = \sum_{i:\, x_i \leq x} p(x_i). \end{aligned}$$

Definición 1.19. Una variable aleatoria X es dicha continua si existe una función $f(x) \geq 0$, tal que

$$F_X(x) = \int_{-\infty}^{x} f(t) \, \mathsf{d}\, t, \qquad \forall x \in \mathbb{R},$$

en este caso decimos que $f(x)$ es *función de densidad* (pdf) de X.

Resultado 1.20. *Una pmf o pdf deben satisfacer*

$$\sum_{x \in \mathcal{X}} p_X(x) = 1, \qquad \int_{-\infty}^{+\infty} f_X(x) \, \mathsf{d}\, x = 1,$$

respectivamente.

Demostración. Primeramente considere el caso discreto. Tenemos que $\mathcal{X} = X(\Omega)$ corresponde al conjunto de todos los resultados que puede adoptar X. Así,

$$\begin{aligned} \sum_{x \in \mathcal{X}} \mathsf{P}(X = x) = \sum_{x \in \mathcal{X}} \mathsf{P}(\{\omega \in \Omega : X(\omega) = x\}) \\ = \mathsf{P}\Big(\bigcup_{\omega \in \Omega} \{\omega \in \Omega : X(\omega) = x\}\Big) \\ = \mathsf{P}(\Omega) = 1. \end{aligned}$$

Por otro lado, para el caso continuo

$$\int_{\mathcal{X}} f_X(x)\,\mathsf{d}\,x = \int_{-\infty}^{+\infty} f_X(x)\,\mathsf{d}\,x = \lim_{x\to+\infty} F_X(x) - \lim_{x\to-\infty} F_X(x) = 1,$$

por la Propiedad 1.5 (i). $\square$

Ejemplo 1.21. Sea $p_X(k) = c(\frac{1}{2})^k$, $k = 1, 2, \dots$. Sabemos que

$$\sum_{k=1}^{\infty}\Big(\frac{1}{2}\Big)^k = \frac{1}{2} + \Big(\frac{1}{2}\Big)^2 + \Big(\frac{1}{2}\Big)^3 + \cdots = 1.$$

Ahora, como $\sum_{k\in\mathcal{X}} p_X(k) = 1$, tenemos que:

$$\begin{aligned}
\sum_{k=1}^{\infty} c\Big(\frac{1}{2}\Big)^k &= c\sum_{k=1}^{\infty}\Big(\frac{1}{2}\Big)^k = c\Big(\frac{1}{2} + \Big(\frac{1}{2}\Big)^2 + \Big(\frac{1}{2}\Big)^3 + \cdots\Big)\\
&= c\cdot\frac{1}{2}\Big(1 + \frac{1}{2} + \Big(\frac{1}{2}\Big)^2 + \cdots\Big) = \frac{c}{2}\sum_{k=0}^{\infty}\Big(\frac{1}{2}\Big)^k\\
&= \frac{c}{2}\Big(\frac{1}{1-1/2}\Big) = c.
\end{aligned}$$

Es decir, $c = 1$.

Notación 1.22. Usualmente anotamos $X \sim F_X(x)$, o bien $X \sim f_X(x)$ para indicar que X tiene distribución F_X (o bien f_X). Es más, se suele abreviar $X \sim F$ $(X \sim f)$. Además, si X e Y tienen la misma distribución anotamos $X \sim Y$.

Observación 1.23. El Resultado 1.20 también se puede establecer como sigue: las funciones $p_X(x)$ y $f_X(x)$ deben satisfacer:

(i) $p_X(x) \geq 0, \qquad f_X(x) \geq 0.$
(ii) $\displaystyle\sum_{x\in\mathcal{X}} p_X(x) = 1, \qquad \int_{\mathbb{R}} f_X(x)\,\mathsf{d}x = 1.$

Con mayor generalidad podemos hacer, por ejemplo:

$$\mathsf{P}(\{x \in A\}) = \int_A f_X(x)\,\mathsf{d}x.$$

Para el caso continuo, tenemos que:

$$\begin{aligned}
\mathsf{P}(a \leq X \leq b) = \mathsf{P}(X \in [a,b]) &= \int_a^b f_X(x)\,\mathsf{d}x = \int_{-\infty}^b f_X(x)\,\mathsf{d}x - \int_{-\infty}^a f_X(x)\,\mathsf{d}x\\
&= F_X(b) - F_X(a).
\end{aligned}$$

Además, es fácil notar que para este caso,

$$\mathsf{P}(X = a) = \mathsf{P}(a \leq X \leq a) = \int_a^a f_X(x)\,\mathsf{d}x = F_X(a) - F_X(a) = 0.$$

De este modo, para el caso continuo las siguientes probabilidades son equivalentes

$$\mathsf{P}(a < X \leq b) = \mathsf{P}(a \leq X < b) = \mathsf{P}(a < X < b) = \mathsf{P}(a \leq X \leq b).$$

Además, es fácil notar que:

$$\mathsf{P}(X > a) = 1 - \mathsf{P}(X \leq a) = 1 - F_X(a).$$

Observación 1.24. Suponga que $g(x)$ es función positiva sobre el conjunto A (también conocido como soporte de g), y además

$$\int_{\{x\in A\}} g(x)\,\mathsf{d}x = K < \infty,$$

para $K > 0$ constante. Entonces,

$$f_X(x) = \frac{1}{K}\,g(x),$$

es función de densidad de la variable aleatoria X definida sobre A.

Ejemplo 1.25. Considere la densidad,

$$f_X(x) = 1 \cdot I_{[0,1]}(x),$$

y sea $Y = \text{mín}(X, 1/2)$. Entonces, la variable aleatoria Y es de tipo mixto.

2. Transformación de Variables Aleatorias

Suponga X una variable aleatoria definida sobre el espacio de probabilidad $(\Omega, \mathcal{F}, \mathsf{P})$ y sea g una función tal que $Y = g(X)$. Considere los siguientes ejemplos.

Ejemplo 2.1. Sea X una variable aleatoria con Y definida como $Y = |X|$. Sea F_X la CDF de X. Entonces, la CDF de Y es dada por:

$$\mathsf{P}(Y \leq y) = \mathsf{P}(|X| \leq y) = \mathsf{P}(-y \leq X \leq y).$$

Asumiremos por ahora que X es variable aleatoria continua, luego

$$\mathsf{P}(\leq y) = \mathsf{P}(X \leq y) - \mathsf{P}(X \leq -y) = F_X(y) - F_X(-y).$$

Note también que $y > 0$. Es decir,

$$F_Y(y) = \begin{cases} F_X(y) - F_X(-y), & y > 0, \\ 0, & \text{en otro caso.} \end{cases}$$

Para el caso discreto tenemos que,

$$\begin{aligned} \mathsf{P}(Y \leq y) &= \mathsf{P}(-y \leq X \leq y) \\ &= \mathsf{P}(X \leq y) - \mathsf{P}(X \leq -y) + \mathsf{P}(X = -y). \end{aligned}$$

Ejemplo 2.2. Sea X una variable aleatoria continua y considere la transformación $Y = aX + b$, $a, b \in \mathbb{R}$ con $a \neq 0$. De este modo,

$$\begin{aligned} F_Y(y) &= \mathsf{P}(Y \leq y) = \mathsf{P}(aX + b \leq y) \\ &= \begin{cases} \mathsf{P}(X \leq \frac{y-b}{a}), & a > 0, \\ \mathsf{P}(X \geq \frac{y-b}{a}), & a < 0, \end{cases} \\ &= \begin{cases} F_X(\frac{y-b}{a}), & a > 0, \\ 1 - F_X(\frac{y-b}{a}), & a < 0. \end{cases} \end{aligned}$$

Ejemplo 2.3. Sea X una variable aleatoria con función de densidad

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{en otro caso,} \end{cases}$$

y sea $Y = e^X$. De este modo,

$$\begin{aligned} F_Y(y) &= \mathsf{P}(Y \le y) = \mathsf{P}(e^X \le y) = \mathsf{P}(X \le \log y) \\ &= F_X(\log y), \qquad y > 0. \end{aligned}$$

Además,

$$\begin{aligned} f_Y(y) &= \frac{\mathsf{d}}{\mathsf{d}\, y} F_Y(y) = \frac{\mathsf{d}}{\mathsf{d}\, y} F_X(\log y) = F_X'(\log y) \frac{\mathsf{d}}{\mathsf{d}\, y} \log y \\ &= \frac{1}{y} F_X'(\log y) = \frac{1}{y} f_X(\log y), \end{aligned}$$

para $y > 0$. Note además que $x = \log y$. Así,

$$f_Y(y) = \begin{cases} \frac{1}{y}, & 0 < \log y < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

La idea clave es estudiar el comportamiento de Y en términos de X. En efecto, podemos escribir

$$\mathsf{P}(Y \in A) = \mathsf{P}(g(X) \in A),$$

lo que permite notar que la distribución de Y depende de la función F_X y g.

Considere $y = g(x)$, entonces la función g transforma el espacio muestral $\mathcal{X}$ en un nuevo espacio muestral $\mathcal{Y}$. Es decir, $g : \mathcal{X} \to \mathcal{Y}$. Sea

$$g^{-1}(A) = \{x \in \mathcal{X} : g(x) \in A\},$$ [1]

así

$$\begin{aligned} \mathsf{P}(Y \in A) &= \mathsf{P}(g(X) \in A) = \mathsf{P}(\{x \in \mathcal{X} : g(x) \in A\}) \\ &= \mathsf{P}(X \in g^{-1}(A)). \end{aligned}$$

Suponga que X es una variable aleatoria discreta. De este modo $\mathcal{X}$ es un conjunto numerable y el espacio muestral para $Y = g(X)$ es,

$$\mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\},$$

que también es un conjunto numerable. De este modo, Y tiene función de probabilidades

$$\begin{aligned} p_Y(y) = \mathsf{P}(Y = y) &= \sum_{x \in g^{-1}(y)} \mathsf{P}(X = x) \\ &= \sum_{x \in g^{-1}(y)} p_X(x), \qquad y \in \mathcal{Y}, \end{aligned}$$

y $p_Y(y) = 0$ para $y \notin \mathcal{Y}$.

Es decir, debemos identificar

$$g^{-1}(\{y\}) = \{x \in \mathcal{X} : g(x) = y\} \quad \big(= g^{-1}(y)\big),$$

para cada $y \in \mathcal{Y}$, y sumar las probabilidades apropiadas.

Ejemplo 2.4. Sea X variable aleatoria discreta con función de probabilidades

x	-1	0	1	2	3
$\mathsf{P}(X = x)$	1/7	2/7	1/7	2/7	1/7

[1] que es una transformación de $\mathcal{Y}$ a $\mathcal{X}$.

y sea $Y = X^2$. De este modo,

$$\begin{aligned}\mathcal{Y} &= \{y : y = g(x), x \in \mathcal{X}\} = \{y : y = x^2, x \in \mathcal{X}\}\\ &= \{0, 1, 4, 9\}.\end{aligned}$$

Esto lleva a,

y	0	1	4	9
$\mathsf{P}(Y = y)$	2/7	2/7	2/7	1/7

Ejemplo 2.5. Sea X variable aleatoria con función de probabilidades

$$\mathsf{P}(X = x) = \binom{n}{x}\theta^x(1-\theta)^{n-x},$$

para $x \in \{0, 1, \dots, n\}$, $n \in \mathbb{N}$ y $\theta \in (0, 1)$. Suponga $Y = g(X)$ donde $g(x) = n - x$. Es decir $\mathcal{X} = \{0, 1, \dots, n\}$, $y = n - x$. Así,

$$\mathcal{Y} = \{y : y = g(x), x \in \mathcal{X}\} = \{0, 1, \dots, n\}.$$

De este modo, $g^{-1}(y) = n - y = x$, y

$$\begin{aligned}\mathsf{P}(Y = y) &= \mathsf{P}(n - X = y) = \mathsf{P}(X = n - y) \quad \Big(= \sum_{x \in g^{-1}(y)} p_X(x)\Big)\\ &= \binom{n}{n-y}\theta^{n-y}(1-\theta)^y.\end{aligned}$$

Recordando que

$$\binom{n}{n-y} = \frac{n!}{(n-y)!(n-(n-y))!} = \binom{n}{y},$$

permite obtener,

$$\mathsf{P}(Y = y) = \binom{n}{y}(1-\theta)^y\theta^{n-y},$$

para $y \in \{0, 1, \dots, n\}$.

Para el caso continuo podemos tener fórmulas simples para la CDF y pdf de Y en términos de la CDF y pdf de X. Para $Y = g(X)$, podemos escribir

$$\begin{aligned}F_Y(y) &= \mathsf{P}(Y \le y) = \mathsf{P}(g(X) \le y) = \mathsf{P}(\{x \in \mathcal{X} : g(x) \le y\})\\ &= \int_{\{x \in \mathcal{X} : g(x) \le y\}} f_X(x)\,\mathsf{d}x.\end{aligned}$$

Además, $f_y(y) = F_Y'(y)$.

Observación 2.6. En ocasiones es difícil indentificar el conjunto $\{x \in \mathcal{X} : g(x) \le y\}$ y llevar a cabo la integración sobre esa región.

Ejemplo 2.7. Sea $f_X(x) = e^{-x}$, $x > 0$. De ahí que

$$F_X(x) = \int_0^x e^{-t}\,\mathsf{d}t = -e^{-t}\Big|_0^x = 1 - e^{-x}.$$

Considere $Y = g(X) = \log X$. Note que

$$\{x \in \mathcal{X} : g(x) \le y\} = \{x : \log(x) \le y\} = \{x : x \le e^y\}.$$

Además,

$$\begin{aligned} F_Y(y) &= \mathsf{P}(Y \leq y) = \mathsf{P}(\log X \leq y) = \mathsf{P}(X \leq e^y) \\ &= F_X(e^y) = 1 - \exp(-\exp(y)). \end{aligned}$$

Esto lleva a,

$$\frac{\mathsf{d}}{\mathsf{d}\,y} F_Y(y) = \frac{\mathsf{d}}{\mathsf{d}\,y}\big(1 - \exp(-e^y)\big) = -\frac{\mathsf{d}}{\mathsf{d}\,y}\exp(-e^y) = -\exp(-e^y)\cdot(-e^y),$$

luego,

$$f_Y(y) = e^y \exp(-e^y), \qquad \text{para } y \in \mathbb{R}.$$

La situación es simple cuando y es función monótona. Sea $Y = g(X)$, y

$$\begin{aligned} &\mathcal{X} = \{x : f_X(x) \geq 0\}, \\ &\mathcal{Y} = \{y : y = g(x), \text{para algún } x \in \mathcal{X}\}. \end{aligned}$$

Para g monótona,

(a) creciente: $u > v \Rightarrow g(u) > g(v)$.
(b) decreciente: $u < v \Rightarrow g(u) > g(v)$.

Esto lleva a notar que,

(a) para g creciente, tenemos

$$\begin{aligned} \{x \in \mathcal{X} : g(x) \leq y\} &= \{x \in \mathcal{X} : g^{-1}(g(x)) \leq g^{-1}(y)\} \\ &= \{x \in \mathcal{X} : x \leq g^{-1}(y)\}, \end{aligned}$$

(b) mientras que para g decreciente, sigue que

$$\begin{aligned} \{x \in \mathcal{X} : g(x) \leq y\} &= \{x \in \mathcal{X} : g^{-1}(g(x)) \geq g^{-1}(y)\} \\ &= \{x \in \mathcal{X} : x \geq g^{-1}(y)\}. \end{aligned}$$

De este modo, para g creciente

$$\begin{aligned} F_Y(y) &= \int_{\{x \in \mathcal{X} : x \leq g^{-1}(y)\}} f_X(x)\,\mathsf{d}x = \int_{-\infty}^{g^{-1}(y)} f_X(x)\,\mathsf{d}x \\ &= F_X(g^{-1}(y)). \end{aligned}$$

Por otro lado, para g decreciente

$$F_Y(y) = \int_{g^{-1}(y)}^{\infty} f_X(x)\,\mathsf{d}x = 1 - F_X(g^{-1}(y)).$$

Ejemplo 2.8. Suponga que

$$f_X(x) = \begin{cases} 1, & 0 < x < 1, \\ 0, & \text{en otro caso.} \end{cases}$$

Así,

$$F_X(x) = \int_0^x \mathsf{d}t = t\Big|_0^x = x, \qquad 0 < x < 1.$$

Considere $Y = g(X) = -\log X$, como

$$\frac{\mathsf{d}}{\mathsf{d}\,x} g(x) = \frac{\mathsf{d}}{\mathsf{d}\,x}(-\log x) = -\frac{1}{x} < 0,$$

pues $x \in (0,1)$, es decir g es decreciente. Además, como x fluctua entre 0 y 1, tenemos que $-\log X$ fluctua entre 0 y ∞, es decir $\mathcal{Y} = (0,\infty)$, y

$$x = g^{-1}(y) = e^{-y}, \qquad y > 0.$$

Es decir,

$$F_Y(y) = 1 - F_X(g^{-1}(y)) = 1 - F_X(e^{-y}) = 1 - e^{-y},$$

y

$$f_y(y) = \frac{\mathsf{d}}{\mathsf{d}\,x} F_Y(y) = e^{-y}, \qquad y > 0.$$

Observación 2.9. Note que sólo es necesario verificar que $g(X) = -\log X$ es monótona en el soporte de X (en el ejemplo anterior en el intervalo $(0,1)$).

Resultado 2.10. *Sea X una variable aleatoria continua con función de densidad f_X. Si g es función diferenciable y monótona, entonces la función de densidad de la variable aleatoria $Y = g(X)$ es dada por*

$$f_Y(y) = \Big|\frac{\mathsf{d}}{\mathsf{d}\,y} g^{-1}(y)\Big| f_X(g^{-1}(y)), \qquad y \in \mathcal{Y}.$$

Demostración. Suponga que g es función monótona creciente, luego

$$F_Y(y) = \mathsf{P}(Y \leq y) = \mathsf{P}(g(X) \leq y) = \mathsf{P}(X \leq g^{-1}(y)) = F_X(g^{-1}(y)),$$

diferenciando, obtenemos

$$f_Y(y) = F_X'(g^{-1}(y)) \cdot \frac{\mathsf{d}}{\mathsf{d}\,y} g^{-1}(y) = f_X(g^{-1}(y))\Big|\frac{\mathsf{d}}{\mathsf{d}\,y} g^{-1}(y)\Big|,$$

dado que la derivada de g es positiva.

Ahora, suponga que g es función monótona decreciente, así

$$F_Y(y) = \mathsf{P}(Y \leq y) = \mathsf{P}(g(X) \leq y) = \mathsf{P}(X \geq g^{-1}(y)) = 1 - F_X(g^{-1}(y)),$$

y diferenciando, sigue que

$$f_Y(y) = -f_X(g^{-1}(y)) \cdot \frac{\mathsf{d}}{\mathsf{d}\,y} g^{-1}(y) = f_X(g^{-1}(y))\Big|\frac{\mathsf{d}}{\mathsf{d}\,y} g^{-1}(y)\Big|,$$

debido a que la derivada de g es negativa. □

Corolario 2.11. *Sea g una función monótona por partes. Suponga que existe una partición $A_1, A_2, \dots, A_k$ de $\mathcal{X}$ tal que g es estrictamente monótona y diferenciable en el interior de cada A_i. Sea $Y = g(X)$. Entonces, la densidad de Y es dada por*

$$f_Y(y) = \sum_{i=1}^{k} f_X(g_i^{-1}(y))\Big|\frac{\mathsf{d}}{\mathsf{d}\,y} g_i^{-1}(y)\Big|, \qquad y \in \mathcal{Y},$$

donde g_i^{-1} es la inversa de g en A_i, para $i = 1, \dots, k$.

Ejemplo 2.12. Suponga X variable aleatoria continua, tal que $Y = X^2$. De este modo,

$$\begin{aligned} F_Y(y) &= \mathsf{P}(Y \leq y) = \mathsf{P}(X^2 \leq y) = \mathsf{P}(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= \mathsf{P}(X \leq \sqrt{y}) - \mathsf{P}(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}), \end{aligned}$$

y diferenciando, obtenemos

$$\begin{aligned}
f_Y(y) &= \frac{\mathsf{d}}{\mathsf{d}\,y}\,F_Y(y) = F_X'(\sqrt{y})\cdot\frac{\mathsf{d}}{\mathsf{d}\,y}\,y^{1/2} - F_X'(-\sqrt{y})\cdot\frac{\mathsf{d}}{\mathsf{d}\,y}\big(-y^{1/2}\big)\\
&= f_X(\sqrt{y})\frac{1}{2}\,y^{1/2-1} + f_X(-\sqrt{y})\frac{1}{2}y^{1/2-1} = \frac{1}{2\sqrt{y}}f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}}f_X(-\sqrt{y})\\
&= \frac{1}{2\sqrt{y}}\big[f_X(\sqrt{y}) + f_X(-\sqrt{y})\big].
\end{aligned}$$

Note que esta densidad se escribe como la suma de 2 partes que representan los intervalos donde $g(x) = x^2$ es monótona.

Ejemplo 2.13. Sea X una variable aleatoria con densidad

$$f_X(x) = (2\pi)^{-1/2}\exp(-x^2/2), \qquad x \in \mathbb{R}.$$

Defina $Y = e^X$. En este caso,

$$y = g(x) = e^x \quad \Rightarrow \quad g^{-1}(y) = x = \log y,$$

y

$$\frac{\mathsf{d}}{\mathsf{d}\,y}\,g^{-1}(y) = \frac{\mathsf{d}}{\mathsf{d}\,y}\log y = \frac{1}{y}.$$

De este modo,

$$\begin{aligned}
f_Y(y) &= (2\pi)^{-1/2}\exp\big(-\tfrac{1}{2}(\log y)^2\big)\cdot\frac{1}{y}\\
&= \frac{1}{y\sqrt{2\pi}}\exp\big(-\tfrac{1}{2}(\log y)^2\big), \qquad y > 0.
\end{aligned}$$

Instituto de Estadística, Universidad de Valparaíso, Chile
Orcid ID: 0000-0002-4675-5201
Email address: felipe.osorio@uv.cl