EYP1027 Modelos Probabilísticos

Clase 9 y 10

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Contenido I

Distribuciones continuas

A continuación se presentan algunas de las distribuciones continuas más comunes, pero que de ninguna manera constituyen todas las distribuciones utilizadas en estadística. De hecho, como ya fue mencionado, cualquier función integrable no negativa se puede transformar en una pdf.

Distribución uniforme

Definición 1.1

La distribución uniforme continua se define extendiendo la masa uniformemente en un intervalo (a, b). Su fdp está dada por,

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in (a, b) \\ 0 & \text{eoc} \end{cases}$$

Notación: $X \sim U(a, b)$ (o $X \sim U[a, b]$).

Teorema 1.1

Si $X \sim U(a,b)$, entonces,

i) $\mathrm{E}(X) = \dfrac{a+b}{2}$

ii) $\mathrm{Var}(X) = \dfrac{(b-a)^2}{12}$

iii) $M(t) = \dfrac{e^{bt}-e^{at}}{t(b-a)}, \quad t \neq 0.$

Demostración 1.1

Se deja como ejercicio para el lector.

Nota: Para $a=0$ y $b=1$, se tiene que $X \sim U(0,1)$, la cual juega un rol clave en la generación de números aleatorios.

Distribuciones continuas

Ejemplo 1.1

Sea $X \sim U(-3, 2)$. Calcule $P(X \geq 0)$ y $P(-5 \leq X \leq \frac{1}{2})$. En este caso la fdp de X está dada por,

$$f(x) = \begin{cases} \dfrac{1}{5}, & \text{si } -3 < x < 2 \\ 0, & \text{eoc} \end{cases}$$

Por lo tanto,

$$P(X \geq 0) = \int_0^2 \frac{1}{5} dx = \frac{2}{5},$$

y

$$P(-5 \leq X \leq \frac{1}{2}) = \int_{-3}^{\frac{1}{2}} \frac{1}{5} dx = \frac{1}{5}\left(\frac{1}{2} + 3\right) = \frac{7}{10}.$$

Distribución normal
La distribución normal es quizás una de las más importantes, debido a que tiene un rol central tanto en la teoría de probabilidad como en la teoría estadística. También se le llama distribución gaussiana en honor a Gauss, a quien se le considera el "padre" de esta distribución. La importancia de la distribución normal se debe al famoso teorema del límite central.

Definición 1.2

Se dice que una variable aleatoria X tiene una distribución normal con parámetros μ (un número real) y σ (un real positivo), si su fdp está dada por,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right\}, \ x \in \mathbb{R}.$$

Notación: $X \sim N(\mu, \sigma^2)$.

Distribuciones continuas

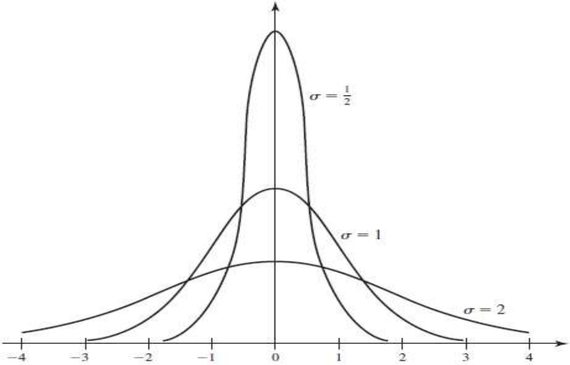

Figura 1: Gráfico de la fdp normal para $\mu = 0$ y $\sigma = 1/2, 1, 2$.

Definición 1.3

Distribución normal estándar Si $Z \sim N(0,1)$ se dice que Z tiene una distribución normal estándar. Su fdp está dada por,

$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\}, \ z \in \mathbb{R}.$$

Su fda, denota por $\Phi(\cdot)$, esta dada por,

$$\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}x^2\right\} dx, \ z \in \mathbb{R},$$

la cual está tabulada para varios valores de $z \in \mathbb{R}$; usualmente para $-4 < z < 4$.

Note que $\Phi(0) = 1/2$ y que $\Phi(z) + \Phi(-z) = 1, \forall \ z \in \mathbb{R}$.

Teorema 1.2

Sea $X \sim N(\mu, \sigma^2)$. Entonces,

i) $\mathrm{E}(X) = \mu$

ii) $\mathrm{Var}(X) = \sigma^2$

iii) $M(t) = \exp\left\{\mu t + \frac{1}{2}\,\sigma^2 t^2\right\}.$

Demostración 1.2

Calculemos la fgm de X. En efecto,

$$
\begin{aligned}
M(t) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{tx - \frac{1}{2\sigma^2}\,(x-\mu)^2\right\} dx \\
&= \exp\left\{\mu t + \frac{1}{2}\,\sigma^2 t^2\right\} \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty} \exp\left\{-\frac{1}{2\sigma^2}\,\left(x-(\mu+\sigma^2 t)\right)^2\right\} dx \\
&= \exp\left\{\mu t + \frac{1}{2}\,\sigma^2 t^2\right\},\ \ t \in \mathbb{R}.
\end{aligned}
$$

Teorema 1.3

Sea $X \sim N(\mu, \sigma^2)$. Entonces,

i) $Y = a + bX \sim N(a + b\mu, b^2\sigma^2)$

ii) $Z = (X - \mu)/\sigma \sim N(0, 1)$.

Demostración 1.3

Basta notar que,

$$\begin{aligned} M_Y(t) &= \exp\{at\} M_X(bt) = \exp\{at\} \exp\left\{\mu bt + \frac{}{2}\sigma^2 b^2 t^2\right\} \\ &= \exp\left\{(a + b\mu)t + \frac{1}{2}(\sigma b)^2 t^2\right\}, \end{aligned}$$

que corresponde a la fgm de una distribución normal con media $a + b\mu$ y varianza $\sigma^2 b^2$. Luego $Y \sim N(a + b\mu, b^2\sigma^2)$.

Ejemplo 1.2

Sea $X \sim N(1,4)$. Calcule, $P(0 \leq X < 1)$ y $P(X^2 > 4)$. Entonces,

$$\begin{aligned} P(0 \leq X < 1) &= P(-\frac{1}{2} \leq \frac{X-1}{2} < 0) \\ &= P(-\frac{1}{2} \leq Z < 0) \\ &= \Phi(0) - \Phi(-\frac{1}{2}) \\ &= 0.5 - 0.30854 \\ &= 0.19146. \end{aligned}$$

Similarmente,

$$\begin{aligned}
P\left(X^2 > 4\right) &= 1 - P(|X| \leq 2)\\
&= 1 - P(-2 \leq X \leq 2)\\
&= 1 - P(-\frac{3}{2} \leq \frac{X-1}{2} \leq \frac{1}{2})\\
&= 1 - P(-\frac{3}{2} \leq Z \leq \frac{1}{2})\\
&= 1 - \left\{\Phi(\frac{1}{2}) - \Phi(-\frac{3}{2})\right\}\\
&= 1 - \{0.69146 - 0.06681\}\\
&= 0.37535.
\end{aligned}$$

Ejemplo 1.3

Sean X_1 y X_2 las duraciones de dos dispositivos electrónicos. Asuma que $X_1 \sim N(40, 36)$ y $X_2 \sim N(45, 9)$. Si el dispositivo se va a utilizar durante 45 horas, qué dispositivo sería el preferido ? Si se va a utilizar durante 42 horas, cuál debería preferirse ?
Hay que encontrar qué dispositivo tiene una mayor probabilidad de vida útil de más de 45 horas:

$$\begin{aligned} P(X_1 > 45) &= P(Z > \frac{45-40}{6}) = 1 - \Phi(5/6) \\ &= 1 - 0.7995 = 0.2005 \end{aligned}$$

y

$$\begin{aligned} P(X_2 > 45) &= P(Z > \frac{45-45}{3}) = 1 - \Phi(0) \\ &= 1 - 0.5 = 0.5. \end{aligned}$$

Por lo tanto, se debe preferir el dispositivo X_2. Ahora, hay que encontrar qué dispositivo tiene una mayor probabilidad de vida útil de más de 42 horas. Cálculos similares producen,

$$\begin{aligned} P(X_1 > 42) &= 1 - \Phi(1/3) \\ &= 0.3707, \\ P(X_2 > 42) &= 1 - \Phi(-1) \\ &= 0.8413 \end{aligned}$$

En este caso también se debe preferir el dispositivo X_2.

Distribución Gama

La distribución gama se usa de manera extensa en una variedad de áreas como, por ejemplo, para describir los intervalos de tiempo entre dos fallas consecutivas en el motor de un avión, o los intervalos de tiempo entre las llegadas de clientes a una cola en el cajero de un supermercado.

Definición 1.4

Se dice que la variable aleatoria X tiene una distribución gama con los parámetros $\alpha > 0$ y $\lambda > 0$ si su fdp está dada por

$$f(x) = \begin{cases} \frac{\lambda}{\Gamma(\alpha)}(\lambda x)^{r-1}\exp(-\lambda x), & \text{si } x > 0, \\ 0, & \text{eoc} \end{cases}$$

donde $\Gamma(\cdot)$ es la función gamma, es decir, $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1}\exp(-t)dt$, con $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$, $\Gamma(1/2) = \sqrt{\pi}$ y $\Gamma(n) = (n-1)!$, $n = 1, 2, \ldots$.

Notación: $X \sim Gama(\alpha, \lambda)$.

Distribuciones continuas

El orden de los parámetros es importante ya que α es un parámetro de forma, mientras que λ es un parámetro de escala.

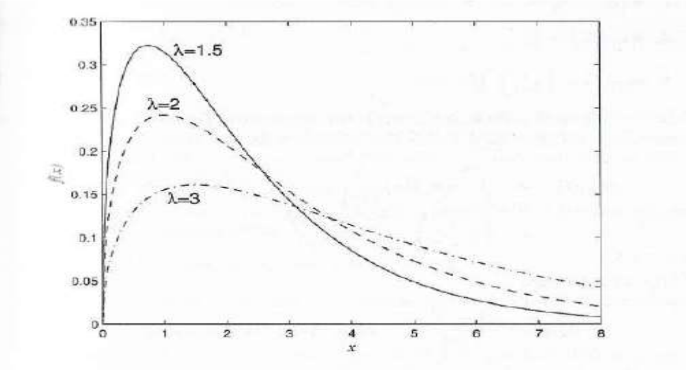

Figura 2: Gráfico de la fdp gama para $\alpha = 1.5$ y $\lambda = 1.5, 2, 3$.

Teorema 1.4

Si $X \sim Gama(\alpha, \lambda)$, entonces,

i) $\mathrm{E}(X) = \alpha/\lambda$

ii) $\mathrm{Var}(X) = \alpha/\lambda^2$

iii) $M(t) = (1 - t/\lambda)^{-\alpha}$, si $t < \lambda$.

Demostración 1.4

Calculemos la fgm de X,

$$
\begin{aligned}
M(t) &= \int_0^{\infty} \exp(tx) \frac{\lambda}{\Gamma(r)} (\lambda x)^{\alpha-1} \exp(-\lambda x) dx \\
&= \left(\frac{\lambda}{\lambda - t}\right)^{\alpha} \int_0^{\infty} \frac{(\lambda - t)^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} \exp\{-(\lambda - t)x\} dx \\
&= \left(1 - \frac{t}{\lambda}\right)^{-\alpha}, \text{ si } \lambda - t > 0.
\end{aligned}
$$

Luego,

$$\mathrm{E}(X) = \frac{d}{dt}M(t)\Big|_{t=0} = \frac{\alpha}{\lambda} \text{ y } \mathrm{E}\left(X^2\right) = \frac{d^2}{dt^2}M(t)\Big|_{t=0} = \frac{\alpha^2+\alpha}{\lambda^2},$$

de donde sigue que $\mathrm{Var}(X) = \alpha/\lambda^2$.

Casos particulares:

i) Si $\alpha = 1 \Longrightarrow X \sim Exp(\lambda)$, distribución exponencial; en este caso

$$E(X) = \frac{1}{\lambda}, \quad \mathrm{Var}(X) = \frac{1}{\lambda^2}, \quad M(t) = (1 - \frac{t}{\lambda})^{-1},\ t < \lambda$$

ii) $\alpha = \frac{\nu}{2}$ y $\lambda = \frac{1}{2} \Longrightarrow X \sim \chi^2_\nu$, distribución chi-cuadrado con ν grados de libertad; para este caso,

$$\mathrm{E}(X) = \nu, \quad \mathrm{Var}(X) = 2\nu, \quad M(t) = (1-2t)^{-\nu/2},\ t < 1/2.$$

iii) Si $\alpha > 1 \Longrightarrow X \sim Erlag(\alpha, \lambda)$, distribución de Erlang.

Funciones de una variable aleatoria

Si X es una variable aleatoria en $(\Omega, \mathcal{A}, P)$, es decir, $X : \Omega \to \mathbb{R}$, entonces, para cualquier función

$$g : \mathbb{R} \to \mathcal{X} \subseteq \mathbb{R},$$

se tiene que $Y = g(X)$ también es una variable aleatoria en $(\Omega, \mathcal{A}, P)$, es decir,

$$Y : \Omega \to \mathcal{Y} \subseteq \mathbb{R},$$

ya que para cada $\omega \in \Omega$, se tiene que $Y(\omega) = g(X(\omega))$ es un número real.

Funciones de una variable aleatoria

Puesto que Y es una función de X, entonces su distribución de probabilidad, P_Y, puede determinarse en términos de la distribución de probabilidad, P_X, de X. Esto es, para cualquier subconjunto B de $\mathcal{Y}$,

$$P_Y(B) = P(Y \in B) = P(g(X) \in B),$$

lo cual muestra que la distribución de Y depende de P_X y la funcion g.

Dependiendo de la elección de g, algunas veces es posible obtener una expresión tratable para el cálculo de estas probabilidades.

Más precisamente, al escribir $y = g(x)$, la función $g(x)$ establece un mapeo del recorrido, $\mathcal{X}$, de la variable aleatoria X, a un subconjunto $\mathcal{Y}$ de $\mathbb{R}$, el cual define el recorrido de la variable aleatoria $Y = g(X)$. Es decir,

$$g : \mathcal{X} \to \mathcal{Y}.$$

Asociado con g, se tiene un mapeo inverso, denotado por g^{-1}, que es un mapeo de subconjuntos de $\mathcal{Y}$ a subconjuntos de $\mathcal{X}$, y está definido por,

$$g^{-1}(B) = \{x \in \mathcal{X} : g(x) \in B\}, \tag{2.1}$$

para cada subconjunto B de $\mathcal{Y}$.

Es decir, $g^{-1}(B)$ es el conjunto de puntos en $\mathcal{X}$ que $g(x)$ transfoma en el conjunto B. Es posible, sin embargo, que B sea un conjunto de un solo punto, digamos $B = \{y\}$. Entonces,

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

En este caso, a menudo se escribe $g^{-1}(y)$ en vez de $g^{-1}(\{y\})$, teniendo en cuenta que $g^{-1}(y)$ puede ser aún un conjunto, si existe más de un x para el cual $g(x) = y$.

Si existe sólo un x para el cual $g(x) = y$, entonces $g^{-1}(y)$ es el conjunto $\{x\}$, y se escribe como $g^{-1}(y) = x$.

Funciones de una variable aleatoria

De esta forma, para cualquier sunconjuto B de $\mathcal{Y}$, la distribución de probabilidad de la variable aleatoria $Y = g(X)$, queda definida como,

$$\begin{aligned} P_Y(B) &:= P(Y \in B) \\ &= P(g(X) \in B) \\ &= P(\{x \in \mathcal{X} : g(x) \in B\}) \\ &= P\left(X \in g^{-1}(B)\right) \\ &:= P_X(g^{-1}(B)). \end{aligned} \tag{2.2}$$

No es difícil probar que $P_Y(B)$, $B \subseteq \mathcal{Y}$, es una medida probabilidad (satisface los Axiomas de Kolmogorov).

Funciones de una variable aleatoria

Caso discreto

Si la variable aleatoria X es discreta, es decir, su recorrido, $\mathcal{X}$, es un subconjunto contable (finito o infinito) de $\mathbb{R}$, entonces, el recorrido la variable aleatoria transformada $Y = g(X)$, es decir,

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

es también un subconjunto contable (finito on infinito) de $\mathbb{R}$; es decir, Y es también una variable aleatoria discreta. Luego, para determinar su distribución de probabilidad, es suficiente encontrar su fmp.

Funciones de una variable aleatoria

Más precisamente, si X es una variable aleatoria discreta, entonces $Y = g(X)$ es también una variable aleatoria discreta, con fmp dada por,

$$\begin{aligned} f_Y(y) &= P(Y = y) \\ &= P(g(X) = y) \\ &= P(X \in g^{-1}(y)) \\ &= P(\{x \in \mathcal{X} : g(x) = y\}) \\ &= \sum_{\{x \in \mathcal{X} : g(x) = y\}} P(X = x) \\ &= \begin{cases} \sum_{\{x \in \mathcal{X} : g(x) = y\}} f_X(x), & \text{si } y \in \mathcal{Y}, \\ 0, & \text{eoc.} \end{cases} \end{aligned}$$

O sea, para encontrar la fmp de Y, basta con identificar el conjunto $g^{-1}(y) = \{x \in \mathcal{X} : g(x) = y\}$, para cada $y \in \mathcal{Y}$, y luego sumar las probabilidades correspondientes.

Ejemplo 2.1

Sea X una variable aleatoria discreta con fmp dada por,

x	-1	0	1	2	3
$f_X(x)$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

Sea $Y = X^2$. Aquí, $y = g(x) = x^2$, $\mathcal{X} = \{-1, 0, 1, 2, 3\}$, y la variable aleatoria discreta Y toma valores en $\mathcal{Y} = \{0, 1, 4, 9\}$. Luego, la fmp de Y es,

y	0	1	4	9
$f_Y(y)$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

Por ejemplo, $P_Y(Y = 1) = P_X(X = -1) + P_X(X = 1) = \frac{1}{7} + \frac{1}{7} = \frac{2}{7}$.

Ejemplo 2.2

1) Si $X \sim Ber(p)$, entonces $Y = X^2 \sim Ber(p)$. En efecto, como $\mathcal{X} = \{0, 1\}$, con $P(X = 1) = p$, $P(X = 0) = 1 - p$ y $P(X = x) = 0$ para todo $x \notin \mathcal{X}$, se tiene que $\mathcal{Y} = \{0, 1\}$ con

$$\begin{aligned}
P(Y = 1) &= P(X^2 = 1) = P(X = 1) = p, \\
P(Y = 0) &= P(X^2 = 0) = P(X = 0) = 1 - p, \quad \text{y} \\
P(Y = y) &= P(X^2 = y) = 0 \ \text{ para todo } \ y \notin \mathcal{Y}.
\end{aligned}$$

   Es decir, $X^2 \overset{d}{=} X$.

2) Si $X \sim UD(\{-1, 0, 1\})$, entonces $Y = X^2 \sim Ber(2/3)$ (tarea!). Cuál es la distribución de $1 - Y$?

Funciones de una variable aleatoria

Caso continuo

Si la variable aleatoria X es continua, la variable aleatoria $Y = g(X)$ puede ser continua o discreta o mixta de acuerdo con la naturaleza de la trasformación $y = g(x)$. Por ejemplo, para $x \in \mathcal{X} = \mathbb{R}$, las funciones

$$g(x) = ax + b, \quad g(x) = e^x, \quad g(x) = x^2, \quad \text{etc.}$$

definen variables aleatorias continuas; mientras que las funciones

$$g(x) = [x] \;\text{ y }\; g(x) = I_A(x) = \begin{cases} 1, & \text{si } x \in A, \\ 0, & \text{si } x \in A^c, \end{cases}$$

definen variables aleatorias discretas; finalmente, la función

$$g(x) = \max\{0, x\},$$

define una variable aleatoria mixta.

Funciones de una variable aleatoria

Al realizar transformaciones, es importante tener presente los recorridos asociados a las variables aleatorias, para evitar posibles confusiones. Al hacer una transformación de X a $Y = g(X)$, lo más conveniente es usar,

$$\mathcal{X} = \{x : f_X(x) > 0\} \xrightarrow{g} \mathcal{Y} = \{y : y = g(x) \text{ para algún } x \in \mathcal{X}\}, \quad (2.3)$$

que corresponden a los soportes de las distribuciones. La fdp de la variable aleatoria X es positiva solamente sobre el conjunto $\mathcal{X}$, y es cero en otro caso. Similarmente, la fdp (o la fmp) de $Y = g(X)$ es positiva sobre el conjunto $\mathcal{Y}$, y es cero en otro caso.

También es importante analizar la naturaleza de la función $g : \mathcal{X} \to \mathcal{Y}$, en particular, si g es o no una función monótona (biunívoca).

Suponga, ahora, que tanto X como $Y = g(X)$ son variables aleatorias continuas. En muchos casos, es posible encontrar expresiones simples para la fda de Y en términos de la fda o la fdp de X y la función g. De hecho, la fda de $Y = g(X)$ esta dada por,

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

Aunque en algunos casos resulta difícil identificar la región,

$$g^{-1}((-\infty, y]) = \{x \in \mathcal{X} : g(x) \leq y\},$$

y resolver la integral de $f_X(x)$ bajo esa región, este método es muy útil para encontrar la fdp de la variable aleatoria Y.

Funciones de una variable aleatoria

En particular, si g es monótona en $\mathcal{X}$, entonces,

$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : x \leq g^{-1}(y)\}, \quad \text{si } g \text{ es creciente,}$$
$$\{x \in \mathcal{X} : g(x) \leq y\} = \{x \in \mathcal{X} : x \geq g^{-1}(y)\}, \quad \text{si } g \text{ es decreciente.}$$

Es decir, para cada $y \in \mathcal{Y}$, la fda de $Y = g(X)$ queda definida como,

$$F_Y(y) = \begin{cases} F_X(g^{-1}(y)), & \text{si } \; g \text{ es creciente,} \\ 1 - F_X(g^{-1}(y)), & \text{si } \; g \text{ es decreciente.} \end{cases}$$

Como Y también es una variable aleatoria continua, entonces, derivando esta última expresión mediante la regla de la cadena, su fdp esta dada por,

$$f_Y(y) = \frac{d}{dy}F_Y(y) = \begin{cases} f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y), & \text{si } \ g \text{ es creciente,} \\ -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y), & \text{si } \ g \text{ es decreciente.} \end{cases}$$

Esto prueba el siguiente resultado, al notar que $\frac{d}{dy}g^{-1}(y)$ es positiva si g es creciente y negativa si g es decreciente.

Teorema 2.1

Sean X e $Y = g(X)$ variables aleatorias continuas, tales que X tiene fdp $f_X(x)$ y g una función monótona. Sean $\mathcal{X}$ e $\mathcal{Y}$ definidos como en (2.3). Suponga que $f_X(x)$ es continua sobre $\mathcal{X}$ y que $g^{-1}(y)$ tiene una derivada continua sobre $\mathcal{Y}$. Entonces, la fdp de Y está dada por,

$$f_Y(y) = \begin{cases} f_X\left(g^{-1}(y)\right) \left|\dfrac{d}{dy} g^{-1}(y)\right|, & \text{para } y \in \mathcal{Y}, \\ 0, & \text{en otros casos.} \end{cases} \tag{2.4}$$

Ejemplo 2.3

Sea X una variable aleatoria con distribución exponencial, con fdp

$$f_X(x) = \begin{cases} \exp(-x), & \text{si} \quad x > 0, \\ 0, & \text{eoc.} \end{cases}$$

Se desea obtener la fdp de la variable aleatoria $Y = 2X + 1$. Para esto, primero note que la fda de X es,

$$F_X(x) = \begin{cases} 1 - \exp(-x), & \text{si} \quad x > 0, \\ 0, & \text{eoc.} \end{cases}$$

Funciones de una variable aleatoria

Así, la fda de $Y = 2X + 1$ está dada por,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(2X + 1 \leq y) \\ &= P(X \leq (y-1)/2) \\ &= F_X\left((y-1)/2\right), \quad \big(x = g^{-1}(y) = (y-1)/2\big) \\ &= \begin{cases} 1 - \exp\left(-(y-1)/2\right), & \text{si} \quad (y-1)/2 > 0, \\ 0, & \text{eoc.} \end{cases} \end{aligned}$$

Es decir,

$$F_Y(y) = \begin{cases} 1 - \exp\left(-(y-1)/2\right), & \text{si } \; y \in \mathcal{Y} = (1, \infty), \\ 0, & \text{eoc.} \end{cases}$$

Luego, la fdp de Y está dada por,

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} \frac{1}{2}\exp\{-(y-1)/2\}, & \text{si } \; y \in \mathcal{Y} = (1, \infty), \\ 0, & \text{eoc.} \end{cases}$$

En general, si $Y = aX + b$, donde $a \neq 0$ y X tiene fda f_X, entonces

$$f_Y(y) = \begin{cases} \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right), & \text{si } \; y \in \mathcal{Y}, \\ 0, & \text{eoc.} \end{cases}$$

donde de acuerdo con (2.3), $\mathcal{Y} = \{y : y = g(x), \text{ algún } x \in \mathcal{X}\}$. Por ejemplo,

a) Sea $Y = \sigma X + \mu$, donde $\sigma > 0$ y $X \sim N(0, 1)$, entonces, $Y = \sigma X + \mu \sim N(\mu, \sigma^2)$.
b) Sea $X \sim N(\mu, \sigma^2)$, entonces, $Y = \frac{X-\mu}{\sigma} \sim N(0, 1)$.
c) Sea $X \sim U(a, b)$, entonces, $Y = \frac{X-b}{b-a} \sim U(0, 1)$.

Ejemplo 2.4

Sea $Y = \exp(-X)$, donde X es una variable aleatoria continua con fdp,

$$f_X(x) = \begin{cases} 2x, & \text{si} \quad 0 < x < 1 \\ 0, & \text{eoc.} \end{cases}$$

Aquí, $y = g(x) = e^{-x}$, luego $x = g^{-1}(y) = -\ln y \in (0,1) = \mathcal{X}$,de modo que $y \in (e^{-1}, 1) = \mathcal{Y}$ y $dg^{-1}(y)/dy = -1/y$. Entonces, la fdp de $Y = \exp(-X)$ es,

$$f_Y(y) = \begin{cases} -2(\ln y)\ |-1/y|, & \text{si} \quad \exp(-1) < y < 1, \\ 0, & \text{eoc.} \end{cases}$$

o bien,

$$f_Y(y) = \begin{cases} \dfrac{-2\ln y}{y}, & \text{si} \quad \exp(-1) < y < 1, \\ 0, & \text{eoc.} \end{cases}$$

Ejemplo 2.5

Sea $Y = g(X) = X^2$, donde X una variable aleatoria continua con fdp $f_X(x)$. Entonces, para cada $y > 0$, la fdp de Y está dada por,

$$\begin{aligned} F_Y(y) &= P(Y \leq y) \\ &= P(X^2 \leq y) \\ &= P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \quad (X \text{ es continua}) \\ &= \{F_X(x) - F_X(-x)\}|_{x=g^{-1}(y)=\sqrt{y}}. \end{aligned}$$

Como Y también es una variable aleatoria continua, entonces su fdp está dada por,

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \begin{cases} \dfrac{1}{2\sqrt{y}}(f_X(\sqrt{y}) + f_X(-\sqrt{y})), & \text{si } y \in \mathcal{Y} = (0, \infty), \\ 0, & \text{eoc.} \end{cases}$$

Funciones de una variable aleatoria

En el ejemplo anterior la función $g(x) = x^2$ es monótona decreciente para $x < 0$ y monótona creciente para $x > 0$. Luego, si el recorrido de de la variable aleatoria X incluye ambas regiones, la fdp de la variable aleatoria $Y = X^2$ es la suma de dos partes, de modo que no se puede aplicar el Teorema (2.1) directamente. Obviamente, si el recorrido de X incluye un solo lado, por ejemplo, $\mathcal{X} = (0, \infty)$, entonces $f_X(-\sqrt{y}) = 0$ para todo $y > 0$, y el resultado para f_Y puede obtenerse a partir del Teorema 2.1, ya que, en este caso, g es monótona sobre $(0, \infty)$.

Este ejemplo ilustra que, aún sí g no sea ni creciente ni decreciente sobre todo $\mathcal{X}$, a menudo la función g puede ser monótona sobre intervalos disjuntos, cuya unión es igual a $\mathcal{X}$.

Funciones de una variable aleatoria

Teorema 2.2

Sean X e $Y = g(X)$ variables aleatorias continuas, donde X tiene una fdp $f_X(x)$ con recorrido $\mathcal{X}$ definido como en (2.3). Suponga que existe una partición, $A_0, A_1, \ldots, A_k$, de $\mathcal{X}$ tal que $P(X \in A_0) = 0$ y $f_X(x)$ es continua sobre cada A_i. Además, suponga que existen funciones $g_1(x), \ldots, g_k(x)$, definidas sobre $A_1, \ldots, A_k$, respectivamente, tales que:

i) $g(x) = g_i(x)$, para $x \in A_i$;

ii) $g_i(x)$ es monótona sobre A_i;

iii) el conjunto $\mathcal{Y} = \{y : y = g_i(x) \text{ para algún } x \in A_i\}$ es el mismo para cada $i = 1, \ldots, k$; y

iv) $g_i^{-1}(y)$ tiene una derivada continua sobre $\mathcal{Y}$, para cada $i = 1, \ldots, k$.

Entonces,

$$f_Y(y) = \begin{cases} \sum_{i=1}^{k} f_X\left(g_i^{-1}(y)\right)\left|\dfrac{d}{dy} g_i^{-1}(y)\right|, & \text{para } y \in \mathcal{Y} \\ 0, & \text{en otros casos.} \end{cases} \tag{2.5}$$

Note que $\mathcal{X}$ puede ser dividido en conjuntos $A_1, \ldots, A_k$ tales que $g(x)$ es monótona sobre cada A_i. El conjunto A_0 puede ignorarse, ya que $P\left(X \in A_0\right) = 0$. Es importante tener en cuenta que cada $g_i(x)$ es una transformación uno-a-uno desde A_i sobre $\mathcal{Y}$.

Ejemplo 2.6

Sea X una variable aleatoria con distribución normal estándar, con fdp,

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

Considere $Y = X^2$. La función $g(x) = x^2$ es monótona sobre $(-\infty, 0)$ y sobre $(0, \infty)$, y en cada caso $\mathcal{Y} = (0, \infty)$. Para aplicar el Teorema 2.2, basta con tomar $A_0 = \{0\}$,

$$\begin{array}{lll} A_1 = (-\infty, 0), & g_1(x) = x^2, & g_1^{-1}(y) = -\sqrt{y}, \\ A_2 = (0, \infty), & g_2(x) = x^2, & g_2^{-1}(y) = \sqrt{y}, \end{array}$$

tal que $\mathbb{R} = \mathcal{X} = A_0 \cup A_1 \cup A_2$, y A_0, A_1 y A_2 son disjuntos.

Luego, la fdp de Y es,

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2} \left| -\frac{1}{2\sqrt{y}} \right| + \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} \left| \frac{1}{2\sqrt{y}} \right|, & \text{si } \ 0 < y < \infty, \\ 0, & \text{eoc,} \end{cases}$$

es decir,

$$f_Y(y) = \begin{cases} \frac{1}{\sqrt{2\pi}} \frac{1}{\sqrt{y}} e^{-y/2}, & \text{si } \ 0 < y < \infty, \\ 0, & \text{eoc.} \end{cases}$$

Se concluye que si $X \sim N(0,1)$, entonces $Y = X^2 \sim \chi_1^2$, es decir, Y tiene distribución chi-cuadrado con 1 grado de libertad.

Teorema 2.3

Sea X una variable aleatoria con fda $F_X(x)$ continua, y defina la variable aleatoria Y como $Y = F_X(X)$. Entonces, $Y \sim U(0,1)$, es decir, Y tiene una distribución uniforme sobre $(0,\ 1)$, de modo que $P(Y \leq y) = y$ para $0 < y < 1$.

Demostración 2.1

Basta probar que $F_Y(y) = y$ para $0 < y < 1$.

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(F_X(X) \leq y) \\ &= P(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y, \end{aligned}$$

o sea

$$f_Y(y) = \begin{cases} 1, & \text{si } 0 < y < 1, \\ 0, & \text{en otros casos.} \end{cases}$$

Funciones de una variable aleatoria

El resultado anterior es útil en la generación de números aleatorios.

Ejemplo 2.7

Sea $X \sim \exp(1)$, es decir, $F_X(x) = 1 - \mathrm{e}^{-x}$ para $0 < x < \infty$. Entonces, la variable aleatoria $Y = F_X(X) = 1 - \mathrm{e}^{-X}$ tiene distribución $U(0,1)$. Así, para generar un valor x de X, se procede de la siguiente manera:

(a) Genere un número aleatorio $y \in (0,1)$; luego,

$$y = F_X(x) = 1 - \mathrm{e}^{-x} \quad (*)$$

(b) Despeje x de $(*)$; es decir,

$$x = F_X^{-1}(y) = -\log(1-y)$$

Así, al generar n números aleatorios $y_1, \ldots, y_n$ en $(0,1)$, se obtienen n realizaciones $x_1, \ldots, x_n$ de la variable aleatoria X.

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.