Introducción a la Estadística y Probabilidades CM-274

Variables Aleatorias

Enlazar espacios muestrales y eventos a los datos, es proporcionado por el concepto de **Variables Aleatorias**.

Una **variable aleatoria** es una función $X : \Omega \to \mathbb{R}$ que asigna a un número real para cada resultado ω.

Ejemplo Lanzamos una moneda 10 veces. Sea $X(w)$ el número de caras en la secuencia ω. Por ejemplo si $\omega = HHTHHHTHHTT$, entonces $X(\omega) = 6$.

Ejemplo Lancemos dos monedas y sea X el número de caras. Entonces, $\mathbb{P}(X = 0) = (\{TT\}) = 1/4$, $\mathbb{P}(X = 1) = \mathbb{P}(\{HT, TH\}) = 1/2$ y $\mathbb{P}(X = 2) = \mathbb{P}(\{HH\}) = 1/4$. Las variables aleatorias y su distribución puede ser resumida como sigue:

ω	$\mathbb{P}(\{\omega\})$	$X(\omega)$
TT	1/4	0
TH	1/4	1
HT	1/4	1
HH	1/4	2

x	$\mathbb{P}(X = x)$
0	1/4
1	1/2
2	1/4

Dada una variable aleatoria X y un subconjunto A de la línea real, definamos $X^{-}(A) = \{\omega \in \Omega : X(\omega) \in A\}$ y sea

$$\mathbb{P}(X \in A) = \mathbb{P}(X^{-1}(A)) = \mathbb{P}(\{\omega \in \Omega; X(\omega) \in A\})$$
$$\mathbb{P}(X = x) = \mathbb{P}(X^{-1}(x)) = \mathbb{P}(\{\omega \in \Omega; X(\omega) = x\}).$$

Función Distribución y Función de Probabilidad

Definición La *Función Distribución Acumulativa* o `CDF`, es la función $F_X : \mathbb{R} \to [0, 1]$ es definido como

$$F_X(x) = \mathbb{P}(X \leq x). \tag{1}$$

Los siguientes resultados muestran que la `CDF` completamente determina la distribución de una variable aleatoria.

1. Sea X que tiene un `CDF` F y sea Y que tiene `CDF` G. Si $F(x) = G(x)$ para todo x, entonces $\mathbb{P}(X \in A) = \mathbb{P}(Y \in A)$ para todo A.
2. Una función $F : \mathbb{R} \to [0, 1]$ es un `CDF` para alguna probabilidad $\mathbb{P}$ si y sólo si se cumplen las siguientes condiciones

(a) F es no decreciente: $x_1 < x_2$ implica $F(x_1) \leq F(x_2)$.

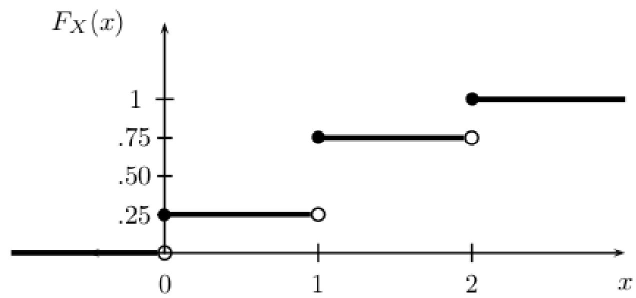
(b) F es normalizado:
$$\lim_{x\to-\infty} F(x) = 0$$
y
$$\lim_{x\to\infty} F(x) = 1$$

(c) F es continua por la derecha: $F(x) = F(x^+)$ para todo x, donde
$$F(x^+) = \lim_{y\to x} F(y) \qquad \text{si} \qquad y > x.$$

Ejemplo Lanzamos una moneda dos veces y sea X el número de caras. Entonces $\mathbb{P}(X = 0) = \mathbb{P}(X = 2) = 1/4$ y $\mathbb{P}(X = 1) = 1/2$. La función distribución es

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1/4 & 0 \leq x < 1 \\ 3/4 & 1 \leq x < 2 \\ 1 & x \geq 2. \end{cases}$$

y el `CDF` es mostrado en la siguiente figura. Se debe notar que la función es continua por la derecha,no decreciente y es definida para todo x, incluso si la variable aleatoria toma los valores $0, 1$ y 2.



Definición Si X es discreta si toma valores contables $\{x_1, x_2, \ldots\}$. Definimos la **función probabilidad** o **pmf** para X por $f_X(x) = \mathbb{P}(X = x)$.

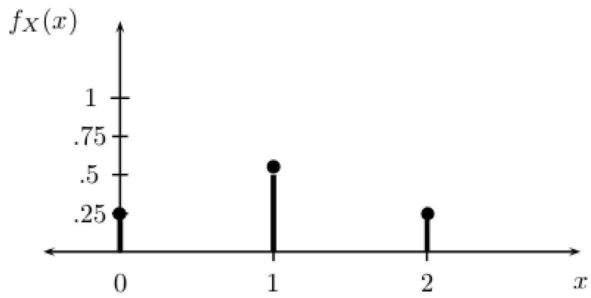
Así, $f_X(x) \geq 0$ para todo $x \in \mathbb{R}$ y $\sum_i f_X(x_i) = 1$. El `CDF` es relacionado a f_X por

$$F_X(x) = \mathbb{P}(X \leq x) = \sum_{x_i \leq x} f_X(x_i).$$

La función probabilidad para el ejemplo anterior es

$$f_X(x) = \begin{cases} 1/ & x = 0 \\ 1/2 & x = 1 \\ 1/4 & x = 2 \\ 0 & \text{en otros casos.} \end{cases}$$

Como indica la siguiente figura



Definición Una variable aleatoria es **continua** si existe una función f_X tal que $f_X(x) \geq 0$ para todo x, $\int_{-\infty}^{\infty} f_X(x)dx = 1$ y para todo $a \leq b$,

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x)dx. \tag{2}$$

La función f_X es llamada **función densidad de probabilidad (PDF)**. Tenemos que

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

y $f_X(x) = F_X'(x)$ para todos los puntos x donde F_X es diferenciable.

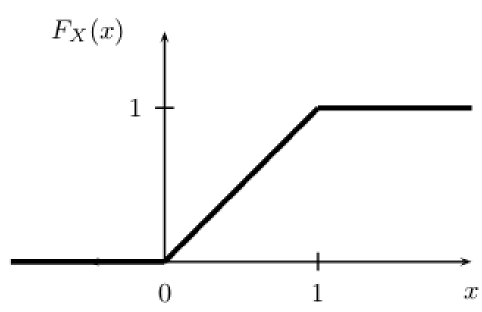
Ejemplo Supongase que X tiene `PDF`

$$f_X(x) = \begin{cases} 1 & \text{para } \ 0 \leq x < 1 \\ 0 & \text{en otros casos} \end{cases}$$

Claramente $f_X \geq 0$ y $\int f_X(x)dx = 1$. Una variable aleatoria con esta densidad se dice que tiene una distribución uniforme (Uniform(0,1)), lo que captura la idea, de escoger un punto entre 0 y 1. El `CDF` está dado por

$$F_X(x) = \begin{cases} 0 & x < 0 \\ x & 0 \leq x \leq 1 \\ 1 & x > 1. \end{cases}$$

y cuyo gráfico es



Ejemplo Supongamos que X tiene PDF

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{1}{(1+x)^2} & \text{en otros casos.} \end{cases}$$

Desde que $\int f(x)dx = 1$ esta es una bien definida PDF.

Observaciones

1. Si X es continua, entonces $\mathbb{P}(X = x) = 0$ para cada x.
2. No se debe pensar $f(x)$ como $\mathbb{P}(X = x)$. Esto se cumple, para variables aleatorias discretas.

Ejemplo Sea

$$f(x) = \begin{cases} 0 & x < 0 \\ \frac{1}{(1+x)} & \text{en otros casos.} \end{cases}$$

Esto no es un PDF, desde que $\int f(x)dx = \log \infty = \infty$.

Teorema Sea F el CDF una variable aleatoria X. Entonces

1. $\mathbb{P}(X = x) = F(x) - F(x^-)$, donde $F(x^-) = \lim_{y \to x} F(y)$ si $y < x$ $(\lim_{y \uparrow x} F(y))$.
2. $\mathbb{P}(x < X \leq y = F(y) - F(x)$
3. $\mathbb{P}(X > x) = 1 - F(x)$
4. Si X es continua, entonces

$$\begin{aligned} F(b) - F(a) &= \mathbb{P}(a < X < b) = \mathbb{P}(a \leq X < b) \\ &= \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X \leq b) \end{aligned}$$

Definamos la inversa CDF, o **función cuantil**

Definición Sea X una variable aleatoria con CDF F. La *inversa CDF* o *función cuantil* es definida como

$$F^{-1}(q) = \inf\{x : F(x) > q\}$$

para $q \in [0, 1]$. Si F es estrictamente creciente y continua, entonces $F^{-1}(q)$ es el único número real x tal que $F(x) = q$.

Llamamos $F^{-1}(1/4)$ el *el primer cuartil*, $F^{-}(1/2)$ la *mediana* o *segundo cuartil* y $F^{-1}(3/4)$, el tercer cuartil.

Dos variables aleatorias X y Y son **iguales en distribución** si $F_X(x) = F_Y(y)$ para todo x. Esto significa que las aseveraciones de probabilidad acerca de X y Y deben ser las mismas.

Algunas importantes Variables Aleatorias Discretas

1.- Distribución Uniforme Discreta Sea $k > 1$ un número entero. Supongamos que X tiene `pmf` dado por

$$f(x) = \begin{cases} 1/k & \text{para } 1, \ldots, k \\ 0 & \text{en otros casos.} \end{cases}$$

Decimos que X tiene una distribución uniforme sobre $\{1, \ldots, k\}$.

2.-La Distribución de Bernoulli Sea X que representa una lanzamiento de una moneda. Entonces $\mathbb{P}(X = 1) = p$ y $\mathbb{P}(X = 0) = 1 - p$ para algún $p \in [0, 1]$, entonces decimos que X tiene una distribución de Bernoulli, escrita como $X \sim$ Bernoulli. La función probabilidad es $f(x) = p^x(1 - p)^{1-x}$ para $x \in \{0, 1\}$.

3.- Distribución Binomial Supongamos que tenemos una moneda, que cae cara con una probabilidad p para algún $0 \leq p \leq 1$. Lanzamos la moneda n veces y sea X el número de caras. Asumimos que esos lanzamientos son independientes. Sea $f(x) = \mathbb{P}(X = x)$, el `pmf` . Se puede mostrar que

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^x & \text{para } x = 0, \ldots, n \\ 0 & \text{en otros casos.} \end{cases}$$

Una variable aleatoria con este `pmf` es llamada variable aleatoria Binomial y se puede escribir como $X \sim \text{Binomial}(n_1, p)$ y $X_2 \sim \text{Binomial}(n_2, p)$ entonces
$X_1 + X_2 \sim \text{Binomial}(n_1 + n_2, p)$.

4.- Distribución Geométrica X tiene una distribución geométrica con parámetro $p \in (0, 1)$, escrito como $X \sim \text{Geom}(p)$, si

$$\mathbb{P}(X = k) = p(1 - p)^{k-1}, \ \ k \geq 1.$$

Tenemos a partir de esto que

$$\sum_{k=1}^{\infty} \mathbb{P}(X = k) = p \sum_{k=1}^{\infty} (1 - p)^k = \frac{p}{1 - (1 - p)} = 1.$$

Piensa en X como el número de lanzamientos necesarios hasta la primera cara, cuando una moneda es lanzada.

5.-La Distribución de Poisson X tiene una distribución de Poisson con parámetro λ escrita como $X \sim \text{Poisson}(\lambda)$ si

$$f(x) = e^{-\lambda} \frac{\lambda^x}{x!} \ \ x \geq 0.$$

Nota que

$$\sum_{k=0}^{\infty} f(x) = e^{-\lambda} \sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = e^{-\lambda} e^{\lambda} = 1.$$

La distribución de Poisson es usado a menudo como un modelo para eventos, como decaimiento radiactivo y accidentes de tráfico. Si $X_1 \sim \text{Poisson}(\lambda_1)$ y $X_2 \sim \text{Poisson}(\lambda_2)$ entonces $X_1 + X_2 \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

Definimos las variables aleatorias como una aplicación de una espacio muestral Ω a $\mathbb{R}$ pero no mencionamos el espacio muestral en cualquiera de las distribuciones anteriores. El espacio muestral a menudo "desaparece", pero es realmente así Vamos a construir un espacio de muestra explícito de una variable aleatoria de Bernoulli. Sea $\Omega = [0,1]$ y definamos $\mathbb{P}$ saisfaciendo $\mathbf{P}([a,b]) = b - a$ para $0 \leq a \leq b \leq 1$. Fijemos $p \in [0,1]$ y definamos

$$X(\omega) = \begin{cases} 1 & \omega \leq p \\ 0 & \omega > p \end{cases}$$

Entonces $\mathbb{P}(X = 1) = \mathbb{P}(\omega \leq p) = \mathbf{P}([0,p]) = p$ y $\mathbb{P}(X = 0) = 1 - p$. Así $X \sim \text{Bernoulli}(p)$.

Algunas importantes Variables Aleatorias Continuas.

1.- La Distribución Uniforme X tiene una distribución Uniforme en (a,b), escrita como $X \sim \text{Uniform}(a,b)$ si

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{para } x \in [a,b] \\ 0 & \text{en otro caso.} \end{cases}$$

donde $a < b$. La función distribución es es

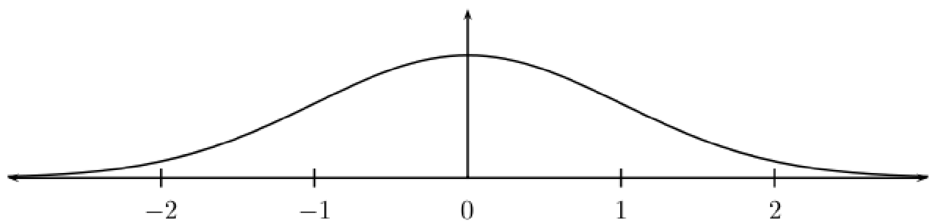
$$F(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & x \in [a,b] \\ 1 & x > b \end{cases}$$

2.-Normal (Gaussiana) X tiene una distribución **Normal (Gaussiana)** con paramétros μ y σ, denotado por $X \sim N(\mu, \sigma^2)$, si

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}, \quad x \in \mathbb{R} \tag{3}$$

donde $\mu \in \mathbb{R}$ y $\sigma > 0$. El parámetro μ es el "centro" (o media) de la distribución y σ es la desviación estándar) de la distribución. La distribución normal juega un papel importante en probabilidad y estadística pues muchos fenómenos en la naturaleza tienen distribuciones aproximadamente normales. Uno de los resultados más importantes de la Estadística el **Teorema del Límite Central** dice que la distribución de una suma de variables aleatorias puede aproximarse por una distribución normal, lo que hace a la distribución Gaussiana muy importante.

Decimos que X tiene una **Distribución Normal Estándar** si $\mu = 0$ y $\sigma = 1$, que se denota como Z. La PDF y CDF de una distribución normal es denotada por $\phi(z)$ y $\Phi(z)$. El PDF de esta distribución es dibujado como



Aquí algunas propiedades

1. Si $X \sim N(\mu, \sigma^2)$, entonces $Z = (X - \mu)/\sigma \sim N(0,1)$.
2. Si $Z \sim N(0,1)$, entonces $X = \mu + \sigma Z \sim N(\mu, \sigma^2)$.
3. Si $X_i \sim N(\mu_i, \sigma_i^2), i = 1, \dots, n$ son independientes, entonces

$$\sum_{i=1}^{n} X_i \sim N\Big(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\Big).$$

Se sigue desde 1. que si $X \sim N(\mu, \sigma^2)$ entonces

$$\begin{aligned}\mathbb{P}(a < X < b) &= \mathbb{P}\Big(\frac{a-\mu}{\sigma} < Z < \frac{b-\mu}{\sigma}\Big) \\ &= \Phi\Big(\frac{b-\mu}{\sigma}\Big) - \Phi\Big(\frac{a-\mu}{\sigma}\Big)\end{aligned}$$

Así podemos calcular las probabilidades, siempre que podamos obtener el CDF de una distribución normal estándar.

Ejemplo Supongamos que $X \sim N(3,5)$. Encuentra $\mathbb{P}(X > 1)$. La solución es

$$\mathbb{P}(X > 1) = 1 - \mathbb{P}(X < 1) = 1 - \mathbb{P}\Big(Z < \frac{1-3}{\sqrt{5}}\Big) = 1 - \Phi(-0.8944) = 0.81.$$

Ahora encontremos $q = \Phi^{-1}(0.2)$. Esto significa que tenemos que encontrar q tal que $\mathbb{P}(X < q) = 0.2$. Podemos resolver esto, escribiendo

$$0.2 = \mathbb{P}(X < q) = \mathbb{P}\Big(Z < \frac{q-\mu}{\sigma}\Big) = \Phi\Big(\frac{q-\mu}{\sigma}\Big).$$

Desde la Tabla Normal, $\Phi(-0.8416) = 0.2$. Por tanto

$$-0.8416 = \frac{q-\mu}{\sigma} = \frac{q-3}{\sqrt{5}}$$

y así $q = 3 - 0.8416\sqrt{5} = 1.1181$.

3.-Distribución Exponencial X tiene una distribución Exponencial con perímetro β y denotado por $X \sim \text{Exp}(\beta)$, si

$$f(x) = \frac{1}{\beta} e^{-x/\beta} \quad, x > 0$$

donde $\beta > 0$. La distribución exponencial es usado para modelar el tiempo de vida de componentes electrónicos y el tiempo de espera entre ciertos eventos.

4.- Distribución Gamma Para $\alpha > 0$, la **Función Gamma** es definida como $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$. X tiene una distribución Gamma con parámetros α y β denotado por $X \sim \text{Gamma}(\alpha, \beta)$, si

$$f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad x > 0$$

donde $\alpha, \beta > 0$. La distribución exponencial es sólo la distribución $\Gamma(1, \beta)$. Si $X_i \sim \text{Gamma}(\alpha_i, \beta)$ son independientes, entonces $\sum_{i=1}^n X_i \sim \text{Gamma}(\sum_{i=1}^n \alpha_i, \beta)$.

5.-Distribución Beta X tiene una distribución Beta con parámetros $\alpha > 0$ y $\beta > 0$, denotado por $X \sim \text{Beta}(\alpha, \beta)$, si

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad 0 < x < 1.$$

6.-Distribución t y de Cauchy X tiene una distribución t con ν grados de libertad, escrito como $X \sim t_\nu$ si

$$f(x) = \frac{\Gamma(\frac{\nu+1}{2})}{\Gamma(\frac{\nu}{2})} = \frac{1}{(1 + \frac{x^2}{\nu})^{(\nu+1)/2}}.$$

La distribución Cauchy es un caso especial de la distribución t, correspondiendo a $\nu = 1$. La densidad es

$$f(x) = \frac{1}{\pi(1+x^2)}$$

$$\begin{aligned} \int_{-\infty}^{\infty} f(x) dx = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)} &= \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{d \tan^{-1}(x)}{dx} \\ &= \frac{1}{\pi} [\tan^{-1}(\infty) - \tan^{-1}(-\infty)] = \frac{1}{\pi} [\frac{\pi}{2} - (-\frac{\pi}{2})] = 1. \end{aligned}$$

7.-La Distribución χ X tiene una distribución χ^2 con p grados de libertad, que se escribe como $X \sim \chi^2_p$ si

$$f(x) = \frac{1}{\Gamma(p/2)2^{p/2}} x^{(p/2)-1} e^{-x/2}, \quad x > 0.$$

Si $Z_1, \ldots, Z_p$ son variables aleatorias normales estándar independientes entonces $\sum_{i=1}^{p} Z_i^2 \sim \chi_2^p$.

Distribuciones Bivariadas

Dado un par de variables aleatorias discretas X y Y, se define el `jdf` *(joint mass function)* por $f(x,y) = \mathbb{P}(X = x, Y = y)$.

Ejemplo Aquí una distribución bivariada para dos variables aleatorias X y Y tomando los valores 0 o 1:

	Y = 0	Y = 1	
X = 0	1/9	2/9	1/3
X = 1	2/9	4/9	2/3
	1/3	2/3	1

Así, $f(1,1) = \mathbb{P}(X = 1, y = 1) = 4/9$.

Definición En el caso continuo, llamamos a una función $f(x,y)$ un `PDF` para una variable aleatoria (X,Y) si

1. $f(x,y) \geq 0$ para todo (x,y).
2. $\int_{\infty}^{\infty} \int_{\infty}^{\infty} f(x,y)dxdy = 1$ y
3. Para algún conjunto $A \subset \mathbb{R} \times \mathbb{R}, \mathbb{P}((X,Y) \in A) = \int\int_A f(x,y)dxdy$.

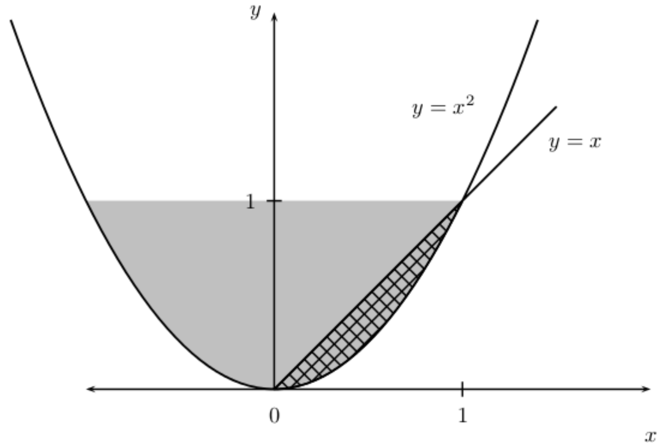
Ejemplo Sea (X,Y) que tiene densidad

$$f(x,y) = \begin{cases} cx^2y & \text{si } x^2 \leq y \leq 1 \\ 0 & \text{en otro caso.} \end{cases}$$

Notemos primero que $-1 \leq x \leq 1$. Encontremos el valor de c. Fijemos un valor x y dejemos que y varie en su rango, el cuál es $x^2 \leq y \leq 1$.

Así

$$1 = \int\int f(x,y)dxdy = c\int_{-1}^{1}\int_{x^2}^{1} x^2 ydydx$$
$$= c\int_{-1}^{1} x^2 \Big[\int_{x^2}^{1} ydy\Big] dx = c\int_{-1}^{1} x^2 \frac{1-x^4}{2} dx = \frac{4c}{21}.$$



Ahora sea $c = 21/4$. Calculemos $\mathbb{P}(X \geq Y)$. Esto corresponde al conjunto $A = \{(x,y); 0 \leq x \leq 1, x^2 \leq y \leq x\}$. (Se puede ver esto, en el diagrama de arriba). Así

$$\begin{aligned}\mathbb{P}(X \geq Y) = \frac{21}{4}\int_0^1 \int_{x^2}^x x^2 y dy dx &= \frac{21}{4}\int_0^1 x^2 \Big[\int_{x^2}^x y dy\Big] dx \\ &= \frac{21}{4}\int_0^1 x^2 \Big(\frac{x^2 - x^4}{2}\Big) dx = \frac{3}{20}.\end{aligned}$$

Definición Si (X,Y) tiene una distribución bivariada con *jdf* $f_{X,Y}$, entonces el `mmf` *marginal mass function para* X es definido por

$$f_X(x) = \mathbb{P}(X = x) = \sum_y \mathbb{P}(X = x, Y = y) = \sum_y f(x,y) \tag{4}$$

y la *marginal mass function para* Y es definido por

$$f_Y(y) = \mathbb{P}(Y = y) = \sum_x \mathbb{P}(X = x, Y = y) = \sum_x f(x,y) \tag{5}$$

Ejemplo Supongamos que $f_{X,Y}$ es dado por la tabla siguiente. La distribución marginal para X corresponde a las filas y la distribución marginal para Y corresponde a las columnas.

	Y = 0	Y = 1	
X = 0	1/10	2/10	3/10
X = 1	3/10	4/10	7/10
	4/10	6/10	1

Por ejemplo, $f_X(0) = 3/10$ y $f_X(1) = 7/10$.

Definición Para variables aleatorias continuas, las densidades marginales son

$$f_X(x) = \int f(x,y)dy \quad f_Y(y) = \int f(x,y)dx \tag{6}$$

Las correspondientes funciones de distribución marginal, son denotadas por F_X y F_Y

Ejemplo Sea (X,Y) tiene densidad

$$f(x,y) = \begin{cases} \frac{21}{4}x^2y & \text{si } x^2 \leq y \leq 1 \\ 0 & \text{en otros casos.} \end{cases}$$

Así

$$f_X(x) = \int f(x,y)dy = \frac{21}{4}x^2 \int_{x^2}^{1} ydy = \frac{21}{8}x^2(1-x^4)$$

para $-1 \leq x \leq 1$ y $f_X(x) = 0$ en otro caso.

Definición Dos variables aleatorias X y Y son independientes, si para cada A y B

$$\mathbb{P}(X \in A, Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B) \tag{7}$$

Teorema Sean X y Y tienen un `PDF` $f_{X,Y}$. Entonces X y Y son independientes si y sólo si $f_{X,Y}(x,y) = f_X(x)f_Y(y)$ para todos los valores de x y y.

Ejemplo Sean X y Y tienen la siguiente distribución

	Y = 0	Y = 1	
X = 0	1/4	1/4	1/2
X = 1	1/4	1/4	1/2
	1/2	1/2	1

Entonces, $f_X(0) = f_X(1) = 1/2$ y $f_Y(0) = f_Y(1) = 1/2$. X y Y son independientes, ya que $f_X(0)f_Y(0) = f(0,0), f_X(0)f_Y(1) = f(0,1), f_X(1)f_Y(0) = f(1,0), f_X(1)f_Y(1) = f(1,1)$. Pero si X y Y tienen la misma distribución

	Y = 0	Y = 1	
X = 0	1/2	0	1/2
X = 1	0	1/2	1/2
	1/2	1/2	1

ellas no son independientes, ya que $f_X(0)f_Y(1) = (1/2)(1/2) = 1/4$, pero $f(0,1) = 0$.

Ejemplo Supongamos que X y Y son indepedientes y tienen la misma densidad

$$f(x) = \begin{cases} 2x & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otros casos} \end{cases}$$

Encontremos $\mathbb{P}(X+Y\leq 1)$. Usando independencia, tenemos que la `jdf` es

$$f(x,y)=f_X(x)f_Y(y)=\begin{cases}4xy & \text{si } \ 0\leq x\leq 1,\ \ 0\leq y\leq 1\\ 0 & \text{en otros casos.}\end{cases}$$

Ahora,

$$\begin{aligned}\mathbb{P}(X+Y\leq 1) &= \int\int_{x+y\leq 1} f(x,y)dydx\\ &= 4\int_0^1 x\left[\int_0^{1-x} ydy\right]dx\\ &= 4\int_0^1 x\frac{(1-x)^2}{2}dx=\frac{1}{6}.\end{aligned}$$

Teorema Supongamos que el rango de X y Y es un rectángulo (posiblemente infinito). Si $f(x,y)=g(x)h(y)$ para algún g y h (no necesariamente funciones de densidad) entonces X y Y son independientes.

Ejemplo Sean X y Y con densidad

$$f(x,y)=\begin{cases}2e^{-(x+2y)} & \text{si } \ x>0 \ \ y \ \ y>0\\ 0 & \text{en otros casos.}\end{cases}$$

El rango de X y Y es el rectángulo $(0,\infty)\times(0,\infty)$. Podemos escribir $f(x,y)=g(x)h(y)$, donde $g(x)=2e^{-x}$ y $h(y)=e^{-2y}$. Así, X y Y son independientes.

Si X y Y son discretas, entonces podemos calcular la distribución condicional de X dado que hemos observado $Y=y$. Especificamente, $\mathbb{P}(X=x|Y=y)=\mathbb{P}(X=x,Y=y)/\mathbb{P}(Y=y)$. Esto conduce a la definición `cpmf` o *Conditional probability mass function*.

Definición La `cpmf` es

$$f_{X|Y}(x|y)=\mathbb{P}(X=x|Y=y)=\frac{\mathbb{P}(X=x,Y=y)}{\mathbb{P}(Y=y)}=\frac{f_{X,Y}(x,y)}{f_Y(y)}$$

si $f_Y>0$.

Definición Para variables aleatorias continuas, la `cpdf`, *Conditional probability density function* es

$$f_{X|Y}(x|y)=\frac{f_{X|Y}(x,y)}{f_Y(y)}$$

asumiendo que $f_Y(y)>0$. Entonces,

$$\mathbb{P}(X\in A|Y=y)=\int_A f_{X|Y}(x|y)dx.$$

Ejemplo Sean X y Y, que tienen una distribución uniforme sobre un cuadrado unidad. Así $f_{X|Y}(x|y) = 1$, para $0 \leq x \leq 1$ y 0 en otros casos. Dado $Y = y$, X es Uniform(0.1). Podemos escribir esto como $X|Y = y \sim \text{Uniform}(0, 1)$.

Ejemplo Supongamos que $X \sim \text{Uniform}(0, 1)$. Después de obtener un valor de X, generamos $Y|X = x \sim \text{Uniform}(x, 1)$. La distribución marginal de Y se puede calcular a partir de

$$f_X(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{en otros casos.} \end{cases}$$

y

$$f_{Y|X}(y|x) = \begin{cases} \frac{1}{1-x} & \text{si } 0 < x < y < 1 \\ 0 & \text{en otros casos.} \end{cases}$$

Así

$$f_{X,Y}(x, y) = f_{Y|X}(y|x) f_X(x) = \begin{cases} \frac{1}{1-x} & \text{si } 0 < x < y < 1 \\ 0 & \text{en otros casos.} \end{cases}$$

La distribución marginal para Y es

$$f_Y(y) = \int_0^y f_{X,Y}(x, y)dx = \int_0^y \frac{dx}{1-x} = -\int_1^{1-y} \frac{du}{u} = -\log(1-y)$$

para $0 < y < 1$.

Distribuciones Multivariadas y Muestras Idénticamente Distribuidas e Independientes

Sea $X = (X_1, \ldots, X_n)$ donde $X_1, \ldots, X_n$ son variables aleatorias. Llamamos a X **Vector Aleatorio**. $f(x_1, \ldots, x_n)$ denota el PDF. Es posible definir los mismos conceptos, de la misma manera que las distribuciones multivariadas. Decimos que $X_1, \ldots, X_n$ son independientes si, para cada $A_1, \ldots, A_n$

$$\mathbb{P}(X_1 \in A_1, \ldots, X_n \in A_n) = \prod_{i=1}^{n} \mathbb{P}(X_i \in A_i). \tag{8}$$

Es suficiente verificar que $f(x_1, \ldots, x_n) = \prod_{i=1}^{n} f_{X_i}(x_i)$.

Definición Si $X_1, \ldots X_n$ son independientes y cada uno de ellos tiene la misma distribución marginal con CDF F, decimos que $X_1, \ldots, X_n$ son idénticamente distribuidas e independientes y escribimos como

$$X_1, \ldots, X_n \sim F$$

Si F tiene densidad f, escribimos también $X_1, \ldots, X_n \sim f$. Llamamos a $X_1, \ldots, X_n$ una muestra aleatoria de tamaño n desde F.

Dos Importantes Distribuciones Multivariadas

Multinomial La versión multivariada de la Binomial, es llamada `multinomial`. Considere la posibilidad de extraer una bola de una urna que tiene bolas con k colores, etiquetados como `'color 1', 'color 2', ...'color k'`. Sea $p = (p1, \ldots, p_k)$ donde $p_j \geq 0$ y $\sum_{j=1}^{k} p_j = 1$ y supongamos que p_j es la probabilidad de encontrar una bola de color j. Lanzamos n veces (independientes lanzamientos con reemplazamiento) y sea $X = (X_1, \ldots, X_k)$ donde X_j es el número de veces que el color j aparece. Por lo tanto $n = \sum_{j=1}^{k} X_j$, así decimos que X tiene una distribución Multinomial Multinomial(n, p). La función probabilidad es

$$f(x) = \binom{n}{x_1 \ldots x_k} p_1^{x_1} \cdots p_k^{x_k} \tag{9}$$

donde

$$\binom{n}{x_1 \ldots x_k} = \frac{n!}{x_1! \cdots x_k!}.$$

Propiedad Supongamos que $X \sim$ Multinomial(n, p), donde $X = (X_1, \ldots, X_k$ y $p = (p_1, \ldots, p_k)$. La distribución marginal de X_j es Binomial(n, p_j).

Normal Multivariada La distribución normal univariada tiene dos parámetros μ y σ. En la versión multivariada, μ es un vector y σ es reemplazada por una matriz Σ. Para empezar

$$Z = \begin{pmatrix} Z_1 \\ \vdots \\ Z_k \end{pmatrix}$$

donde $Z_1, \ldots, Z_k \sim N(0, 1)$ son independientes. La densidad de Z es

$$\begin{aligned} f(z) = \prod_{i=1}^{k} f(z_i) &= \frac{1}{(2\pi)^{k/2}} \exp\Big\{ -\frac{1}{2} \sum_{j=1}^{k} z_j^2 \Big\} \\ &= \frac{1}{(2\pi)^{k/2}} \exp\Big\{ -\frac{1}{2} z^T z \Big\}. \end{aligned}$$

Decimos que Z tiene una distribución multivariada estándar escrita como $Z \sim N(0, I)$, donde 0 representa el vector de K ceros y I es la matriz identidad de orden $k \times k$. De manera más general, un vector X tiene una distribución Normal Multivariada, denotada por $X \sim N(\mu, \Sigma)$, si tiene una densidad

$$f(x;\mu,\Sigma) = \frac{1}{(2\pi)^{k/2}|(\Sigma)|^{1/2}} \exp\Big\{-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\Big\} \tag{10}$$

donde $|\Sigma|$, denota el determinante de Σ, μ es un vector de longitud k y Σ es una matriz simétrica, definida positiva. Colocando $\mu = 0$ y $\Sigma = I$ tenemos la distribución Normal Estándar.

desde que $|\Sigma|$ es simétrica y definida positiva, existe una matriz $\Sigma^{1/2}$ llamada la raíz cuadrada de Σ con las siguientes propiedades:

1. La matriz $\Sigma^{1/2}$ es simétrica.
2. $\Sigma = \Sigma^{1/2}\Sigma^{1/2}$.
3. $\Sigma^{1/2}\Sigma^{-1/2} = \Sigma^{-1/2}\Sigma^{1/2} = I$ donde $\Sigma^{-1/2} = (\Sigma^{1/2})^{-1}$

Teorema Si $Z \sim N(0,I)$ y $X = \mu + \Sigma^{1/2}Z$ entonces $X \sim N(\mu,\Sigma)$. De otro lado, si $X \sim N(\mu,\Sigma)$ entonces $\Sigma^{-1/2}(X-\mu) \sim N(0,I)$.

Supongamos que particionamos un vector Normal aleatorio X, como $X = (X_a, X_b)$. De manera similar tenemos la partición $\mu = (\mu_a, \mu_b)$ y

$$\Sigma = \begin{pmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{pmatrix}$$

Teorema Sea $X \sim N(\mu,\Sigma)$. Entonces

1. La distribución marginal de X_a es $X_a \sim N(\mu_a, \Sigma_{aa})$.
2. La distribución condicional de X_b dado $X_a = x_a$ es

$$X_b|X_a = x_a \sim N(\mu_b + \Sigma_{ba}\Sigma_{aa}^{-1}(x_a-\mu_a), \Sigma_{bb} - \Sigma_{ba}\Sigma_{aa}^{-1}\Sigma_{ab}).$$

3. Si a es un vector entonces $a^TX \sim N(a^T\mu, a^T\sum a)$.
4. $V = (X-\mu)^T\Sigma^{-1}(X-\mu) \sim \chi^2_k$.

Transformación de Variables Aleatorias

Supongamos que X es una variable aleatoria con `PDF` f_X y `CDF` F_X . Sea $Y = r(X)$ una función de X, llamada una **transformación de X**. Calcular el `PDF` y `CDF` de Y en el caso discreto, por ejemplo tenemos

$$\begin{aligned} f_Y(y) &= \mathbb{P}(Y = y) = \mathbb{P}(r(X) = y) \\ &= \mathbb{P}(\{x; r(x) = y\}) = \mathbb{P}(X \in r^{-1}(y)) \end{aligned}$$

Ejemplo Sea $\mathbb{P}(X=-1)=\mathbb{P}(X=1)=1/4$ y $\mathbb{P}(X=0)=1/2$. Sea $Y=X^2$. Entonces, $\mathbb{P}(Y=0)=\mathbb{P}(X=0)=1/2$ y $\mathbb{P}(Y=1)=\mathbb{P}(X=1)+\mathbb{P}(X=-1)=1/2$. Resumiendo

x	$f_X(x)$
-1	1/4
0	1/2
1	1/4

y	$f_Y(y)$
0	1/2
1	1/2

El caso es continuo, es más dificil. Aquí tres pasos para encontrar f_Y:

1. Para cada y, encuentra el conjunto $A_y=\{x : r(x)\leq y\}$.
2. Encontrar el `CDF`

$$\begin{aligned}F_Y(y) &= \mathbb{P}(Y\leq y)=\mathbb{P}(r(X)\leq y)\\ &= \mathbb{P}(\{x : r(x)\leq y\})\\ &= \int_{A_y} f_X(x)dx\end{aligned}$$

3. El `PDF` es $f_Y(y)=F_Y'(y)$.

Ejemplo Sea $f_X(x)=e^{-x}$ para $x>0$. Así, $F_X(x)=\int_0^x f_X(s)ds=1-e^{-x}$. Sea $Y=r(X)=\log(X)$. Entonces $A_y=\{x : x\leq e^y\}$ y

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

Por tanto, $f_Y(y)=e^y e^{-e^y}$ para $y\in\mathbb{R}$.

Si tenemos por ejemplo, X y Y variables aleatorias, veamos como conocer la distribución de $X/Y, X+Y, \max\{X,Y\}$ o $\min\{X,Y\}$. Sea $Z=r(X,Y)$, los pasos para encontrar f_Z son los mismo de antes

1. Para cada z, encuentra el conjunto $A_z=\{(x,y) : r(x,y)\leq z\}$.
2. Encontrar el `CDF`

$$\begin{aligned}F_Z(z) &= \mathbb{P}(Z\leq z)=\mathbb{P}(r(X,Y)\leq z)\\ &= \mathbb{P}(\{(x,y) : r(x,y)\leq z\})\\ &= \int\int_{A_z} f_{X,Y}(x,y)dxdy\end{aligned}$$

3. El `PDF` es $f_Z(z)=F_Z'(z)$.

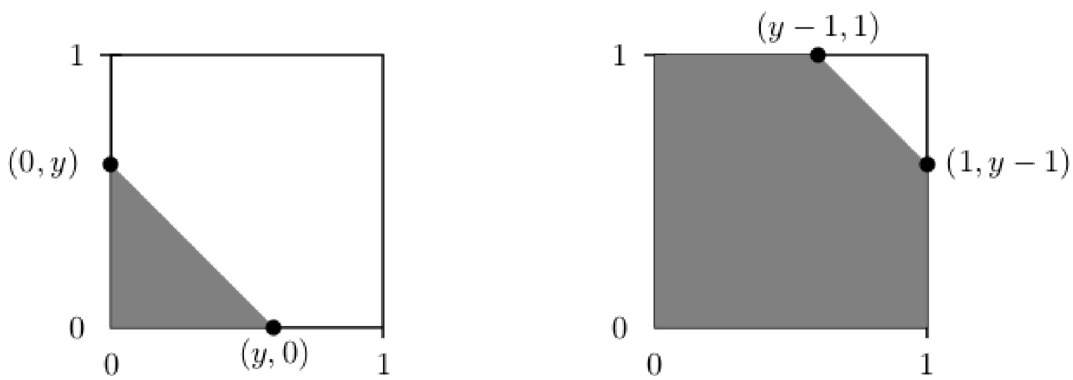
Ejemplo Sea $X_1, X_2 \sim \text{Uniform}(0.1)$ independientes. Encontremos la densidad de $Y = X_1 + X_2$ en efecto la `jdf` de (X_1, X_2) es

$$f(x_1, x_2) = \begin{cases} 1 & 0 < x_1 < 1, 0 < x_2 < 1 \\ 0 & \text{en otros casos} \end{cases}$$

Sea $r(x_1, x_2) = x_1 + x_2$. Ahora

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) = \mathbb{P}(r(X_1, X_2) \leq y) \\ &= \mathbb{P}(\{(x_1, x_2) : r(x_1, x_2) \leq y\}) = \int \int_{A_y} f(x_1, x_2) dx_1 dx_2. \end{aligned}$$

Para encontrar A_y, suponemos primero que $0 < y \leq 1$. Entonces A_y es el triángulo con vértices $(0,0), (y,0)$ y $(0.y)$.. En este caso $\int \int_{A_y} f(x_1, x_2) dx_1 dx_2$ es el área del triángulo, es decir $y^2/2$. Si $1 < y < 2$, entonces A_y es todo el cuadrado unitario, menos el triángulo con vértices $(1, y-1), (1,1), (y-1, 1)$. Este conjunto tiene área $1 - (2-y)^2/2$



Por tanto

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ \frac{y^2}{2} & 0 \leq y < 1 \\ 1 - \frac{(2-y)^2}{2} & 1 \leq y < 2 \\ 1 & y \geq 2 \end{cases}$$

Por diferenciación el `PDF` es

$$f_Y(y) = \begin{cases} y & 0 \leq y \leq 1 \\ 2 - y & 1 \leq y \leq 2 \\ 0 & \text{en otros casos.} \end{cases}$$