# 1 Variables aleatorias multivariadas

Las relaciones que pueden existir entre dos o más variables aleatorias, suele ser un problema de modelado más común que uno que involucra una única variable aleatoria. Por ejemplo, podemos estar interesados en relacionar la variable aleatoria $X$ que mide la intensidad de una señal de teléfono celular en algún lugar y la variable aleatoria $Y$ que mide la distancia de ese lugar hacia el local de cobertura más cercano. O podemos estar interesados en relacionar una variable aleatoria $X$ que define el proceso de llegada de los pacientes al consultorio de un médico y la variable aleatoria $Y$ que mide el tiempo que pasan los clientes en la sala de espera. Como ejemplo final, podemos estar interesados, en examinar la situación en la que los trabajos que llegan a un fábrica en el que las máquinas están sujetas a fallas. Se puede usar una primera variable aleatoria para indicar el número de trabajos presentes y se puede usar un segunda variable para representar el número de máquinas de trabajo.

Las sumas, las diferencias y los productos de las variables aleatorias suelen ser importantes, al igual que las relaciones como el máximo y el mínimo. Considere la posibilidad de ejecutar un programa de computadora que consta de diferentes secciones cuyo tiempo para completar depende de la cantidad y el tipo de datos proporcionados.

De ello se deduce que el tiempo necesario para ejecutar todo el programa depende también de los datos. El tiempo de ejecución de la i-ésima sección puede ser representado por una variable aleatoria $X_i$. Sea la variable aleatoria $X$ la que represente el tiempo necesario para ejecutar todo el programa. Si las secciones se ejecutan una tras otra, entonces $X$ está dado por $X = \sum_i X_i$. Si las secciones se ejecutan en paralelo, entonces $X = \max\{X_i, i = 1, 2, \dots\}$.

## 1.1 Funciones de masa de probabilidad conjunta

La función de masa de probabilidad para una variable aleatoria $X$, es denotada por $p_X(x)$ y es la probabilidad que $X$ toma el valor de $x$, esto es, $p_X(x)$ es la probabilidad de los eventos que consisten de todas las salidas que son mapeadas en el valor $x$. Si $X$ y $Y$ son variables aleatorias, una salida es un punto en plano $(x, y)$ y un evento es un subconjunto de esos puntos del plano: $[X = x, Y = y]$ es el evento que consiste de todas las salidas para el cual $X$ es mapeada a $x$ y $Y$ es mapeada a $y$. La `función de masa de probabilidad conjunta` denotada como $p_{X,Y}$ es igual a la probabilidad del evento $[X = x, Y = y]$ y es la función $p_{X,Y} : \mathbb{R}^2 \to [0, 1]$ definida por

$$p_{X,Y}(x, y) = \mathbb{P}(\omega \in \Omega : X(\omega) = x, Y(\omega) = y)$$

o de forma abreviada :

$$p_{X,Y}(x, y) = \mathbb{P}(X = x, Y = y)$$

La función de masa de probabilidad conjunta de dos variables aleatorias discretas puede representarse convenientemente en forma tabular. Por ejemplo, podemos colocar valores realizables de $X$ en la parte superior de una tabla y los de $Y$ en el lado izquierdo. El elemento de la tabla correspondiente a la columna $X = x$ y la fila $Y = y$ denota la probabilidad del evento $[X = x, Y = y]$,

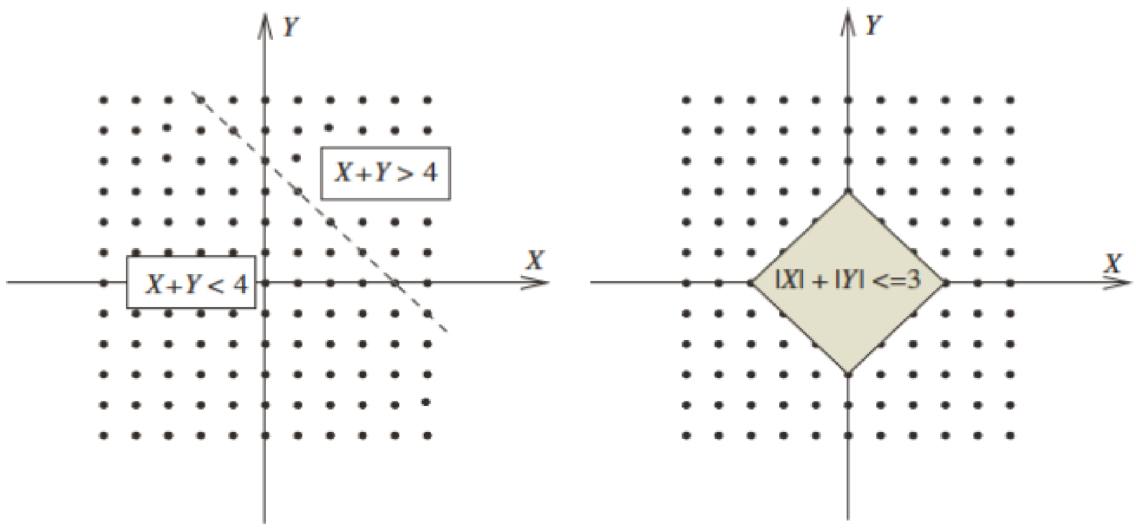
**Ejemplo 1.1** Sea $X$ una variable aleatoria que tiene los valores en $\{1, 2, 3, 4, 5\}$ y $Y$ una variable aleatoria con valores en $\{-1, 0, 1, 2\}$ y sea la función de masa de probabilidad conjunta dada por $p_{X,Y}(x, y) = \alpha$, para todos los posibles valores de $x$ e $y$. Para calcular los valores de $\alpha$, observamos que sólo 20 punto en el plano $(x, y)$ tiene probabilidad positiva y además la probabilidad de cada uno es la misma. Esto nos dá $\alpha = 1/20 = 0.05$. Las siguientes tablas muestran la función de masa de probabilidad conjunta de $X$ y $Y$.

| | x = 1 | x = 2 | x=3 | x =4 | x =5 |
|---|---|---|---|---|---|
| y = -1 | 1/20 | 1/20 | 1/20 | 1/20 | 1/20 |
| y = 0 | 1/20 | 1/20 | 1/20 | 1/20 | 1/20 |
| y = 1 | 1/20 | 1/20 | 1/20 | 1/20 | 1/20 |
| y = 2 | 1/20 | 1/20 | 1/20 | 1/20 | 1/20 |

Cada evento $A$ es un subconjunto de puntos en el plano $(x, y)$, por lo que podemos escribir como:

$$\mathbb{P}(A) = \sum_{(x,y)\in A} p_{X,Y}(x, y).$$

Donde la suma es sobre todos los puntos $(x, y) \in A$. La siguiente figura, muestra dos eventos $A = [X + Y < 4]$ y $B = [|X| + |Y| \leq 3]$, donde todos los posibles valores de $X$ e $Y$ están marcados por puntos.




**Ejemplo 1.2** Sean $X$ e $Y$ variables aleatorias continuas cuya función de masa de probabilidad conjunta es dada por el anterior ejemplo. Sea $A$ el evento que $X$ no tiene valores mayores que $Y$, escrito como $[X \leq Y]$. Los puntos en el plano, teniendo una probabilidad positiva, para el cual, el valor de $X$ es menor o igual al valor de $Y$ son $(1, 1), (1, 2)$ y $(2, 2)$. Por tanto la probabilidad de este evento es dado por $\mathbb{P}(A) = \mathbb{P}(X \leq Y) = 3/20$.

Sea ahora el evento $[XY = 0]$. Este evento contiene, cinco salidas: $(1, 0), (2, 0), (3, 0), (4, 0)$ y $(5, 0)$ y así $\mathbb{P}(B) = \mathbb{P}(XY = 0) = 5/20$.

La función de masa de probabilidad conjunta, especifica la probabilidad de eventos conjuntos, denotados por $[X = x, Y = y]$. Es posible considerar la probabilidad de eventos que se refiere a una de las dos variables aleatorias, tales como $[X = x]$. En efecto considerando el ejemplo (1.1), el evento $[X = 2]$ consiste de cuatro salidas $(2, -1), (2, 0), (2, 1)$ y $(2, 2)$ y tienen una probabilidad igual a $4/20$. En este caso, usamos la función de masa de probabilidad conjunta de las dos variables aleatorias para construir la función de masa de probabilidad para cada una de las variables aleatorias. Esas funciones son las `funciones de masa probabilidad marginal`. La función de masa de probabilidad marginal de $X$ es obtenida como:

$$p_X(x) = \mathbb{P}(X = x) = \sum_k \mathbb{P}(X = x, Y = y_k) = \sum_k p_{X,Y}(x, y_k). \quad (1)$$

Y es una consecuencia directa, del hecho que para un evento $A$, $\mathbb{P}(A) = \sum_{(x,y)\in A} p_{X,Y}(x,y)$.

La distribución marginal de $Y$ es obtenida de manera similar.

**Definición 1.1** Las variables aleatorias $X$ e $Y$ se dicen que son independientes si se cumple :

$$\mathbb{P}(X = x_i, Y = y_i) = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_i)$$

Para todo $x_i$ y $y_i$. En otras palabras, $X$ e $Y$ son independientes si su función de masa conjunta, se factoriza en el producto de sus funciones de masa de probabilidad individualmente.

**Ejemplo 1.3** Sea una distribución para dos variables aleatorias $X$ y $Y$ que toman los valores 0 - 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 |

Entonces $p_{X,Y}(1,1) = \mathbb{P}(X = 1, y = 1) = 4/9$.

**Ejemplo 1.4** Sean $X$ e $Y$ dos variables aleatorias discretas con una función de masa de probabilidad conjunta :

| | X = -1 | X = 0 | X=1 |
|---|---|---|---|
| Y = -1 | 1/12 | 3/12 | 1/12 |
| Y = 0 | 1/12 | 0/12 | 1/12 |
| Y = 1 | 1/12 | 3/12 | 1/12 |

Así, ambas variables aleatorias $X$ y $Y$ toman los valores, desde el conjunto $\{-1, 0, 1\}$. La tabla indica que $\mathbb{P}(X = 0, Y = 0) = 0$, que $\mathbb{P}(X = 0, Y = 1) = 1/4$ y así sucesivamente. La función de masa de probabilidad marginal de $X$, $p_X(x)$ es obtenida sumando las entradas en cada columna de la tabla, la de $Y$, $p_Y(y)$ se encuentra sumando las entradas en cada fila de la tabla. Por lo que se tiene la siguiente tabla:

| | X = -1 | X = 0 | X=1 | $p_Y(y)$ |
|---|---|---|---|---|
| Y = -1 | 1/12 | 3/12 | 1/12 | 5/12 |
| Y = 0 | 1/12 | 0/12 | 1/12 | 1/6 |
| Y = 1 | 1/12 | 3/12 | 1/12 | 5/12 |
| $p_X(x)$ | 1/4 | 1/2 | 1/4 | 1 |

En este ejemplo, $X$ e $Y$ no son independientes, desde que :

$$1/12 = \mathbb{P}(X = 1, Y = 1) \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 1) = 3/12 \times 5/12.$$

De manera similar, se aplica a $X = (X_1, X_2, \ldots, X_n)$ un vector de variables aleatorias discretas. Para este caso la función de masa de probabilidad conjunta de $X$ es definido como la función $p_X$ definida por:

$$p_X(x) = \mathbb{P}(X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

para $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

## 1.2 Función de distribución acumulativa conjunta

La función de distribución acumulativa conjunta de dos variables $X$ e $Y$ es dada por:

$$F_X, Y(x, y) = \mathbb{P}(X \leq x, Y \leq y); \ \ -\infty < x < \infty, \ -\infty < y < \infty.$$

Para ser un CDF de buena fe, una función debe poseer las siguientes propiedades, similares a las encontradas en el CDF de una sola variable aleatoria:

- $0 \leq F_{X,Y}(x, y) \leq 1$ para $-\infty < x < \infty, \ -\infty < y < \infty$, desde que $F_{X,Y}(x, y)$ es una probabilidad.
- $F_{X,Y}(x, \infty) = F_{X,Y}(-\infty, y) = 0$
  $F_{X,Y}(-\infty, -\infty) = 0$, $F_{X,Y}(+\infty, +\infty) = 1$.
- $F_{X,Y}$ debe ser una función no decreciente de ambos $x$ e $y$, es decir:

$$F_{X,Y}(x_1, y_1) \leq F_{X,Y}(x_2, y_2) \text{ si } x_1 \leq x_2 \text{ si } \ y_1 \leq y_2.$$

**Ejemplo 1.5** La función distribución acumulativa de una variable aleatoria $X$ dado en forma tabular, es dada por:

| | $-2 < x$ | $-2 \leq x < -1$ | $-1 \leq x < 0$ | $0 \leq x < 1$ | $1 \leq x$ |
|---|---|---|---|---|---|
| $F_x(x)$ | 0 | 1/8 | 1/4 | 1/2 | 1 |

Esto nos dice, por ejemplo que $\mathbb{P}(X < 1.0) = 1/2$. También $\mathbb{P}(X \leq 0.5) = 1/2, \mathbb{P}(X \leq 0.0) = 1/2$ y así. De manera similar $\mathbb{P}(X < -1.0) = 1/8, \mathbb{P}(X \leq -1.5) = 1/8$ . Usando el mismo enfoque, podemos encontrar la función de distribución acumulutativa conjunta de dos variables aleatorias discretas $X$ e $Y$ en forma tabular. Un ejemplo es el siguiente:

| | | | | | |
|---|---|---|---|---|---|
| $y \geq 2$ | 0 | 1/8 | 1/4 | 1/2 | 1 |
| $0 \leq y < 2$ | 0 | 3/32 | 3/16 | 3/8 | 3/4 |
| $-2 \leq y < 0$ | 0 | 1/32 | 1/16 | 1/8 | 1/4 |
| $y < 2$ | 0 | 0 | 0 | 0 | 0 |
| $F_{X,Y}(x, y)$ | $-2 < x$ | $-2 \leq x < -1$ | $-1 \leq x < 0$ | $0 \leq x < 1$ | $1 \leq x$ |

De esta tabla, podemos encontrar por ejemplo que $\mathbb{P}(X < 1, Y < 0) = 1/8$, que $\mathbb{P}(X \leq 0.5, Y < 0) = 1/8 = \mathbb{P}(X \leq 0.0, Y < 0)$, que $\mathbb{P}(X \leq -1.3, Y \leq 1.3) = 3/32$, que $\mathbb{P}(X \geq 1, Y < 1) = 3/4$ y así.

Se observa que las condiciones detalladas anteriormente para que una función sea un CDF de buena fe se mantienen en este ejemplo.

Mientras que una sola variable aleatoria mapea los resultados en puntos de la línea real, la variable aleatoria conjunta mapea los resultados en puntos en un espacio de mayor dimensión. Por ejemplo, dos variables aleatorias asignan resultados en el plano $(x, y)$, tres variables aleatorias mapean los resultados en el espacio tridimensional $(x, y, z)$ y así sucesivamente. Para una sola variable aleatoria, el valor de la función de distribución acumulativa en un punto $x$ es igual a la probabilidad del evento que contiene todos los resultados que son mapeadas en puntos menores o iguales a $x$.

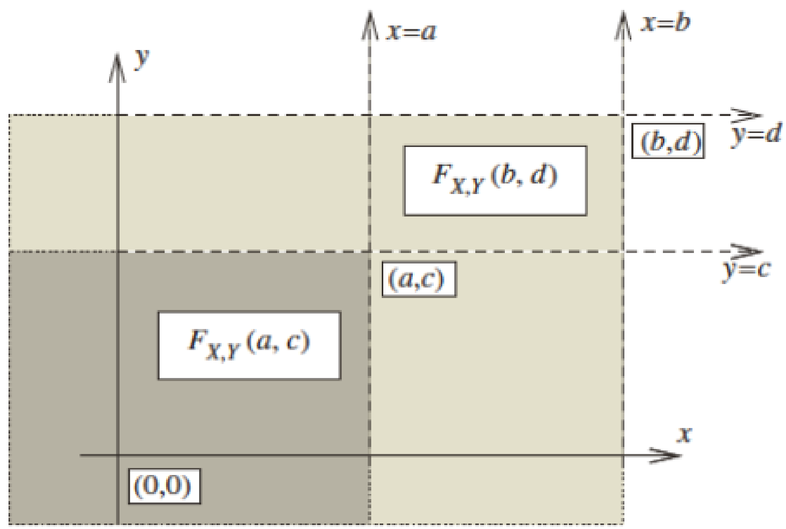
El valor de la distribución acumulativa de dos variables aleatorias $X$ e $Y$ en el punto $(x, y)$ es la probabilidad del evento que contiene todos los resultados que se asignan en el área infinita que se encuentra a la izquierda de una línea vertical a través de $X = x$ y por debajo de la línea horizontal a través de $Y = y$.

Puede observarse de la siguiente figura, que $F_{X,Y}(a, b) = \mathbb{P}(X \leq b, Y \leq d)$ es igual a la probabilidad del evento que contiene todas las salidas dentro de las áreas sombreadas (sombreado ligero y fuerte), mientras que $F_{X,Y}(a,c) = \mathbb{P}(X \leq a, Y \leq c)$ es igual a la probabilidad del evento que contiene todas las salidas dentro del área de sombreado fuerte. Otros resultado siguen, por ejemplo:

$$\mathbb{P}(a < X \leq b, Y \leq c) = F_{X,Y}(b,c) - F_{X,Y}(a,c)$$

y

$$\mathbb{P}(X \leq a, c < Y \leq d) = F_{X,Y}(a,d) - F_{X,Y}(a,c)$$




El uso de la función de distribución acumulativa conjunta de dos variables aleatorias para calcular la probabilidad que $(X, Y)$ esté en un rectángulo en el plano no es excesivamente complicado. Podemos usar el resultado:

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c) + F_{X,Y}(a,c). \tag{2}$$

Lo que requiere que la función de distribución acumulativa conjunta se evalúe en cada uno de los cuatro puntos del rectángulo. Desafortunadamente, cuando la probabilidad requerida se encuentra en una región no rectangular, el cálculo del CDF conjunto se vuelve mucho más difÃcil de usar. Sin embargo, la ecuación (2) es un caso especial útil cuya prueba constituye un ejercicio que ilumina.

Procedemos de la siguiente manera: Sea $A$ que representa un rectángulo con vértices $(a,c), (b,c), (b,d)$ y $(a,d)$ (es decir, el rectángulo dentro del cual buscamos la probabilidad que $X, Y$ pueden ser encontradas). Entonces $\mathbb{P}(A) = \mathbb{P}(a < X \leq b, c < Y \leq d)$. Sea $B$ el rectángulo semi-infinito que se encuentra a la izquierda de $A$ y sea $C$ que representa el rectángulo semi-infinito que está debajo de $A$, como se ilustra en la en la siguiente figura.

Se debe observar que $A, B$ y $C$ son eventos que no tienen puntos en común, es decir, son mutuamente excluyentes, así:

$$\mathbb{P}(A \cup B \cup C) = \mathbb{P}(A) + \mathbb{P}(B) + \mathbb{P}(C),$$

Lo que da como resultado:

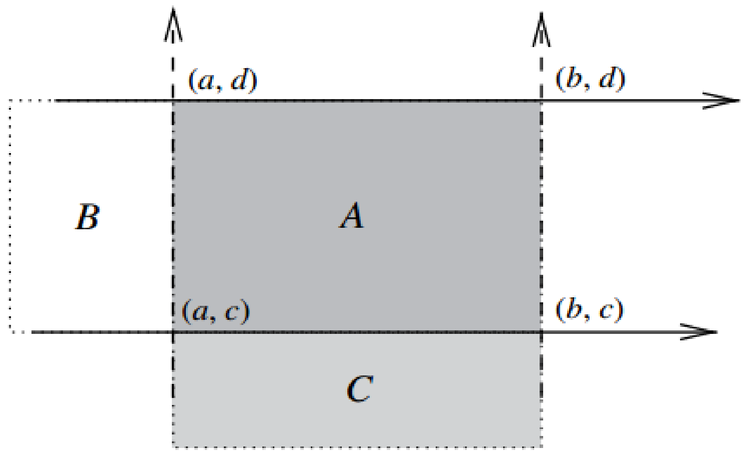
$$\mathbb{P}(A) = \mathbb{P}(A \cup B \cup C) - \mathbb{P}(B) - \mathbb{P}(C)$$

Además de la definicón de CDF conjunta, se tienen los siguientes resultados:

$$\begin{aligned}
\mathbb{P}(B) &= F_{X,Y}(a,d) - F_{X,Y}(a,c), \\
\mathbb{P}(C) &= F_{X,Y}(b,c) - F_{X,Y}(a,c), \\
\mathbb{P}(A \cup B \cup C) &= F_{X,Y}(b,d) - F_{X,Y}(a,c),
\end{aligned}$$

y podemos concluir que :

$$\begin{aligned}
\mathbb{P}(A) &= F_{X,Y}(b,d) - F_{X,Y}(a,c) - [F_{X,Y}(a,d) - F_{X,Y}(a,c) + F_{X,Y}(b,c) - F_{X,Y}(a,c)] \\
&= F_{X,Y}(b,d) - F_{X,Y}(a,d) - F_{X,Y}(b,c) + F_{X,Y}(a,c).
\end{aligned}$$




**Ejemplo 1.6** La función de distribución acumulativa conjunta de dos variables aleatorias continuas es dada por:

$$F_{X,Y}(x,y) = (1 - e^{-\lambda x})(1 - e^{-\mu y}),\ 0 \leq x < \infty, 0 \leq y < \infty,$$

y es cero en otros partes del plano, donde ambos $\lambda$ y $\mu$ son estrictamente positivo. Calculemos la probabilidad $\mathbb{P}(0 \leq x \leq 1, 0 \leq y \leq 1)$. Aademás por (2) tenemos:

$$\begin{aligned}
\mathbb{P}(0 \leq x \leq 1, 0 \leq y \leq 1) &= F_{X,Y}(1,1) - F_{X,Y}(0,1) - F_{X,Y}(1,0) + F_{X,Y}(0,0) \\
&= (1 - e^{-\lambda})(1 - e^{-\mu}) - (1 - e^{-\lambda})(1 - e^{0}) - (1 - e^{0})(1 - e^{-\mu}) + (1 - e^{0})(1 - e^{0}) \\
&= (1 - e^{-\lambda})(1 - e^{-\mu}).
\end{aligned}$$

Ahora volvemos nuestra atención a las funciones de distribución acumulativa marginal derivadas de las CDF conjuntas considerando lo que ocurre cuando a una , pero no a ambas, de las variables aleatorias $X$ e $Y$ se hace tender al infinito. En el límite cuando $y \to \infty$, el evento $[X \leq x, Y \leq y]$ tiende al evento $[X \leq x, Y < \infty] = [X \leq x]$ (se ha utilizado que $Y < \infty$ para indicar que $Y$ puede tomar cualquier valor). Esto implica que la única restricción es $[X \leq x]$, es decir, el evento $[X \leq x, Y < \infty]$ se reduce al evento $[X \leq x]$. Pero ahora, la probabilidad del evento $[X \leq x]$ es simplemente la distribución acumulativa de $X$, así:

$$\lim_{y \to \infty} F_{X,Y}(x,y) = F_X(x).$$

De manera similar:

$$\lim_{x\to\infty} F_{X,Y}(x,y) = F_Y(y).$$

Estas funciones son llamadas `distribuciones marginales` y pueden ser calculadas desde la distribución conjunta.

**Ejemplo 1.7** Del ejemplo (1.5), podemos calcular las distribuciones marginales como:

| | $-2 < x$ | $-2 \leq x < -1$ | $-1 \leq x < 0$ | $0 \leq x < 1$ | $1 \leq x$ |
|---|---|---|---|---|---|
| $F_x(x)$ | 0 | 1/8 | 1/4 | 1/2 | 1 |

para $X$, como antes y para $Y$:

| | $-2 < y$ | $-2 \leq y < 0$ | $0 \leq y < 2$ | $y \geq 2$ |
|---|---|---|---|---|
| $F_y(y)$ | 0 | 1/4 | 3/4 | 1 |

De manera similar, las CDF marginales de $X$ y $Y$, derivadas desde la CDF conjunta del ejemplo (1.6), son calculadas fácilmente y son:

$$F_X(x) = (1 - e^{-\lambda x})(1 - 0) = (1 - e^{-\lambda x})$$
$$F_Y(y) = (1 - 0)(1 - e^{-\mu y}) = (1 - e^{-\mu y})$$

**Definición 1.2** Dos variables aleatorias $X$ e $Y$ son independientes si:

$$F_{X,Y}(x,y) = F_X(x)F_Y(y), \quad -\infty < x < \infty, \ -\infty < y < \infty.$$

En otras palabras, si $X$ e $Y$ son variables aleatorias independientes, la función de distribución acumulativa conjunta se factoriza en el producto de sus funciones de distribución acumulativa marginal. Se puede comprobar que las variables aleatorias $X$ e $Y$ de los ejemplos (1.5) y (1.6) son independientes. Si, además, $X$ e $Y$ tienen la misma función de distribución (por ejemplo, cuando $\lambda = \mu$ en el ejemplo (1.6)), entonces el término `independiente e idénticamente distribuido` se le aplica a cada uno de ellos . Normalmente se abreviará a **iid** o **IID**.

La generalización de la definición de la función de distribución acumulativa conjunta a funciones de distribución conjunta de más de dos variables aleatorias, denominados `vectores aleatorios`, es sencilla. Dadas $n$ variables aleatorias $X_1, X_2, \dots X_n$, la función de distribución acumulativa conjunta se define para todo $x_i$ real, $-\infty < x_i < \infty, i = 1, 2, \dots n$ como:

$$F(x_1, x_2, \dots, x_n) = \mathbb{P}(X_1 \leq x_1, X_2 \leq x_2, \dots, X_n \leq x_n).$$

En este caso $F$ debe ser una función no decreciente de todos los $x_i$, $F(-\infty, -\infty, \dots, -\infty) = 0$ y $F(\infty, \infty, \dots, \infty) = 1$. La distribución marginal de $X_i$ es obtenida desde:

$$\mathbb{P}(X_i \leq x_i) = F_{X_i}(x_i) = F(\infty, \dots, \infty, x_i, \infty \dots \infty).$$

Las variables aleatorias $X_1, X_2, \dots X_n$, se dice que son independientes, si para todo los reales $x_1, x_2, \dots x_n$,

$$F(x_1, x_2, \ldots, x_n) = \prod_{i=1}^{n} F_{X_i}(x_i)$$

Si además, todas las $n$ variables aleatorias, $X_i, i = 1, 2, \ldots, n$, tienen la misma función de distribución, se dice que son independientes e idénticamente distribuidas.

**Ejemplo 1.8** Sean $X$ e $Y$ dos variables aleatorias, cada una tomando los valores $\{\ldots, -1, 0, 1, \ldots\}$ con una función de masa de probabilidad conjunta:

$$\mathbb{P}(X = i, Y = j) = p(i, j) \text{ para } i, j = 0, \pm 1, \pm 2, \ldots$$

La función de distribución acumulativa conjunta es dada por:

$$F_{X,Y}(x, y) = \sum_{i \leq x, j \leq y} p(i, j) \text{ para } (x, y) \in \mathbb{R}^2$$

**Ejemplo 1.9** Sean $X$ e $Y$ son variables aleatorias con una función de distribución conjunta:

$$F_{X,Y}(x, y) = \begin{cases} 1 - e^{-x} - e^{-y} + e^{x+y} & \text{si } x, y \geq 0 \\ 0 & \text{en otros casos} \end{cases}$$

La función de distribución marginal de $X$ es:

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x, y) = \begin{cases} 1 - e^{-x} & \text{si } x \geq 0 \\ 0 & \text{en otro casos} \end{cases}.$$

Así $X$ tiene una distribución exponencial con parámetro 1. Un cálculo muestra que $Y$ tiene esta distribución también. Así;

$$F_{X,Y}(x, y) = F_X(x) F_Y(y) \text{ para } x, y \in \mathbb{R}.$$

por tanto $X$ e $Y$ son independientes.

## 1.3 Función densidad de probabilidad conjunta

La función densidad de probabilidad $f_X(x)$ de una variable aleatoria $X$ continua, es definida como la relación:

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

Para un pequeño infinitesimal $dx$, $\mathbb{P}(x < X \leq xx) = f_X(x) dx$, $\displaystyle\int_{-\infty}^{\infty} f_X(x) dx = 1$ y que la probabilidad $\mathbb{P}(a < X \leq b) = \displaystyle\int_{a}^{b} f_X(x) dx$. Todas estas ideas se pueden llevar a dos dimensiones o dimensiones más altas. Si $X$ e $Y$ son variables aleatorias continuas , usualmente podemos encontrar la `función densidad de probabilidad conjunta` $f_{X,Y}(x,y)$ tal que:

$$F_{X,Y}(x,y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u,v)dudv,$$

Y satisface, las condiciones $f_{X,Y}(x,y) \geq 0$ para todo $(x,y)$ y :

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(u,v)dudv = 1.$$

Si las derivadas parciales de $F_{X,Y}(x,y)$ con respecto a $x$ e $y$ existen, entonces:

$$f_{X,Y}(x,y) = \frac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}.$$

Esta relación implica que:

$$\mathbb{P}(a < X \leq b, c < Y \leq d) = \int_{a}^{b} \int_{c}^{d} f_{X,Y}(x,y)dydx.$$

Esto significa que, mientras que la función de densidad de probabilidad de una única variable aleatoria se mide por unidad de longitud, la densidad de probabilidad conjunta de dos variables aleatorias se mide por unidad de área. Además, para infinitesimales $dx$ y $dy$, tenemos:

$$\mathbb{P}(x < X \leq x + dx, y < Y \leq y + dy) \approx f_{X,Y}(x,y)dxdy.$$

Esta propiedad nos lleva a tomar $f_{X,Y}(x,y)$ como el análogo continuo de $\mathbb{P}(X = x, Y = y)$.

Con dos variables aleatorias conjuntas que son continuas, un evento corresponde a un área en el plano $(x,y)$. El evento ocurre si la variable aleatoria $X$ asigna un valor $x_1$ al resultado, $Y$ asigna un valor $y_1$ al resultado y el punto $(x_1, y_1)$ se encuentra dentro del área que define el evento. La probabilidad de la ocurrencia de un evento $A$ viene dada por:

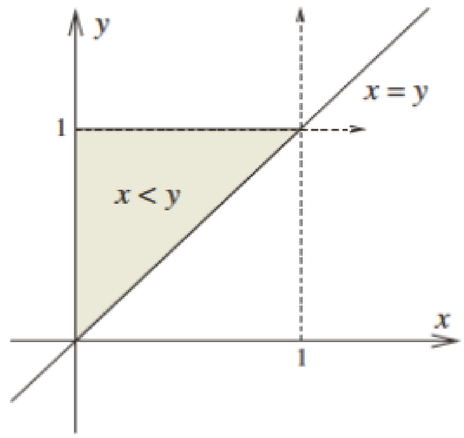
$$\mathbb{P}(A) = \int \int_{\{x,y|(x,y)\in A\}} f_{X,Y}(u,v)dudv = \int \int_{A} f_{X,Y}(u,v)dudv.$$

**Ejemplo 1.10** Sea $X$ e $Y$ dos variables aleatorias con una función densidad de probabilidad conjunta, que tiene la forma:

$$f_{X,Y}(x,y) = \alpha(x+y), \quad 0 \leq x \leq y \leq 1.$$

Y es igual a cero, en otros partes. Encontremos el valor de la constante $\alpha$ y dibujemos la región de probabilidad positiva (el dominio de $f_{X,Y}(x,y)$ en el cual $f_{X,Y}(x,y) > 0$). También, dibujemos la región, donde $X + Y \leq 1$ e integremos sobre esta región para encontrar $\mathbb{P}(X + Y \leq 1)$.

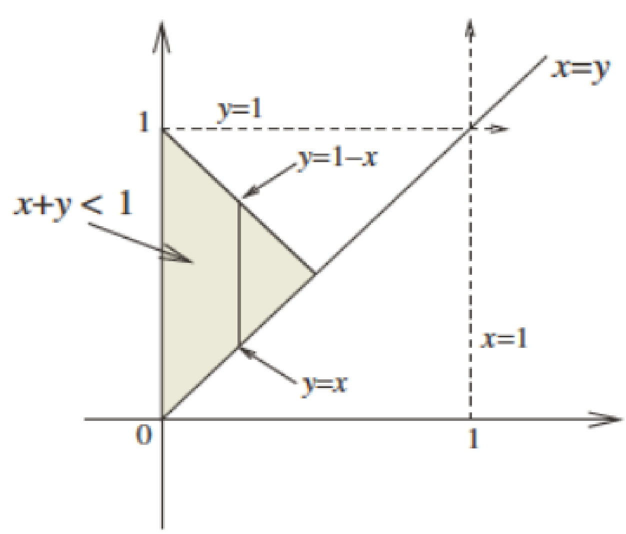
La región del plano, que contiene la probabilidad positiva es mostrada en la siguiente figura:




Para satisfacer, los requerimientos de una función de densidad de probabilidad $f_{X,Y}(x,y)$, debe ser negativa y $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)sdxdy$ es igual a 1. En efecto $f_{X,Y}(x,y) \geq 0$ para todo $x$ e $y$ y además que:

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)dxdy = \int_0^1\int_0^y \alpha(x+y)dxdy = \int_0^1 \left(\alpha x^2/2 + \alpha xy\Big|_0^y\right)dy$$
$$= \int_0^1 \frac{3\alpha y^2}{2}dy = \alpha/2.$$

Se sigue que $\alpha = 2$, para esta función densidad de probabilidad. La región, para el cuál $X + Y \leq 1$ (dentro de la región de probabilidad positiva) es mostradas en la siguiente figura. Esta es toda la región que debemos integrar:



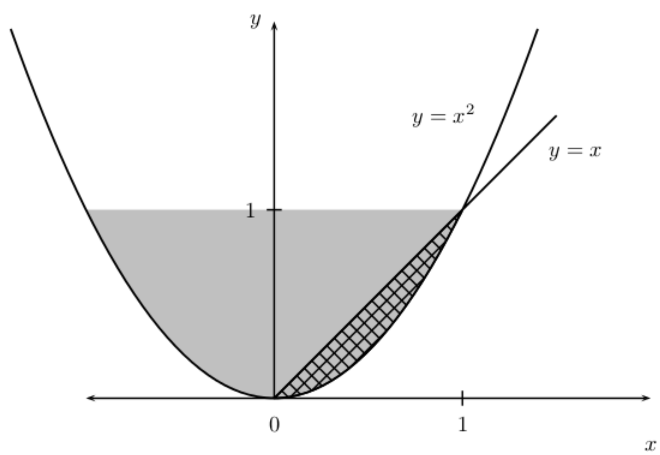

Tenemos:

$$\begin{aligned}\mathbb{P}(X+Y\leq 1) &= \int_{x=0}^{1/2}\int_{y=x}^{1-x}(2x+2y)dydx = \int_{x=0}^{1/2}(2xy+y^2)\Big|_{x}^{1-x}dx\\ &= \int_{x=0}^{1/2}[2x[1-x]+(1-x)^2-2x^2-x^2]dx\\ &= \int_{x=0}^{1/2}(1-4x^2)dx = 1/3\end{aligned}$$

**Ejemplo 1.11** Sean $X$ e $Y$ variables aleatorias continuas, que tienen una función de densidad conjunta:

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

Notemos primero que $-1\leq x\leq 1$. Para encontrar 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:

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

Luego el tenemos que $c=21/4$. 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 gráfico anterior ). Así:

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

## 1.4 Variables aleatorias conjuntas uniformemente distribuidas

Cuando las variables aleatorias X e Y están uniformemente distribuidas sobre una región del plano, es posible utilizar argumentos geométricos para calcular ciertas probabilidades. Por ejemplo, si $X$ e $Y$ son variables aleatorias continuas, uniformemente distribuidas sobre el cuadrado unitario y buscamos la probabilidad $X \leq Y$, observamos que la región $0 \leq x \leq 1, 0 \leq y \leq 1$ y $x \leq y$ constituyen, la mitad de la unidad cuadrada, es decir, el triángulo limitado por el eje $y$, las lńeas $y = 1$ y $x = y$ y así $\mathbb{P}(X \leq Y) = 0.5$. De manera general, si $R$ es una región finita del plano en el que $(X, Y)$ es seguro de pertenecer, la probabilidad de que $(X, Y)$ se encuentre en una subregión $A$ de $R$ es proporcional al tamaño de $A$, con respecto al tamaño de $R$. En otras palabras:

$$\mathbb{P}\Big((X,Y) \in A\Big) = \frac{\text{Area de}A}{\text{Area de}R}.$$

**Ejemplo 1.12** Sean $X$ e $Y$ variables continuas continuas distribuidas uniformemente sobre el cuadrado unitario. Entonces:

$$f_{X,Y}(x,y) = \begin{cases} 1 & \text{si } \ 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0 & \text{en otros casos} \end{cases}$$

Encontremos $\mathbb{P}(X < 1/2, Y < 1/2)$. El evento $A = \{X < 1/2, Y < 1/2\}$ corresponde a un subconjunto del cuadrado unidad. Integrando $f_{X,Y}$ sobre este conjunto, corresponde a calcular el área de conjunto $A$ que es $1/4$. Así $\mathbb{P}(X < 1/2, Y < 1/2) = 1/4$.

**Ejemplo 1.13** Sean $X$ e $Y$ variables aleatorias, uniformemente distribuidas sobre el cuadrado unitario. Encontremos la probabilidad que $\mathbb{P}(X^2 + Y^2 \leq r^2)$ para $r \leq 1$. La función de distribuciíon conjunta de $X$ e $Y$ es dada por:

$$f_{X,Y}(x,y) = \begin{cases} 1 & 0 \leq x \leq 1, \ 0 \leq y \leq 1, \\ 0 & \text{en otros casos.} \end{cases}$$

El área de la región de probabilidad positiva, es decir, el área de $R$, es igual a 1. Se debe notar que $x^2 + y^2 = r^2$ define un círculo centrado en el origen y con radio $r$. El área de este círculo es $\pi r^2$. La porción de la misma que se encuentra en el cuadrante positivo, donde la función de densidad conjunta es positiva, es $\pi r^2/4$. Así, en nuestra terminología, el área de $A$ es $\pi r^2/4$ y por lo tanto:

$$\mathbb{P}(X^2 + Y^2 \leq r^2) = \frac{\pi r^2/4}{1} = \frac{\pi r^2}{4}.$$

**Ejemplo 1.14** La función $f_{X,Y}$ definida por:

$$f_{X,Y}(x,y) = \begin{cases} \dfrac{1}{ab} & \text{si } \ 0 < x < b, 0 < y < b \\ 0 & \text{en otros casos} \end{cases}$$

es una función densidad de probabilidad conjunta. Las variables $X$ e $Y$ son uniformemente distribuidas sobre el rectángulo $B = (0, a) \times (0, b)$. En efecto, para alguna región del plano $A$:

$$\begin{aligned} \mathbb{P}((X,Y) \in A) &= \int\int_A f(x,y)dxdy \\ &= \int\int_{A \cap B} f(x,y)dxdy = \frac{\text{area}(A \cap B)}{\text{area}(B)} \end{aligned}$$

## 1.5 Opcional: Suma de variables aleatorias continuas

Si $X$ e $Y$ son variables aleatorias discretas, con una función de masa de probabilidad conjunta.¿ Cuál es la función de masa probabilidad de $Z = X + Y$?. Para este caso $Z$ toma valores de $z$ si y sólo si $X = x$ y $Y = z - x$ para algún $x$, entonces

$$\begin{aligned}\mathbb{P}(Z = z) &= \mathbb{P}\Big(\bigcup_x (\{X = x\} \cap \{Y = z - x\})\Big) \\ &= \sum_{x \in \texttt{Im X}} \mathbb{P}(X = x, Y = z - x) \;\text{ para }\; z \in \mathbb{R}\end{aligned}$$

**Proposición 1.1** `(Fórmula de Convolución)` Si $X$ e $Y$ son variables aleatorias discretas independientes, entonces $Z = X + Y$ tiene un función masa de probabilidad:

$$\mathbb{P}(Z = z) = \sum_{x \in \texttt{Im X}} \mathbb{P}(X = x)\mathbb{P}(Y = z - x) \;\text{ para }\; z \in \mathbb{R}.$$

La fórmula de convolución dice que la función de masa de $X + Y$ es la `convolución` de la función de $X$ e $Y$.

**Definición 1.3** La función indicador de un evento $A$, denotado por $1_A$ y es dado por:

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

La función $1_A$ indica si o no ocurre $A$. Para esta variable aleatoria discreta, la esperanza es dada por:

$$\mathbb{E}(1_A) = \mathbb{P}(A)$$

Algunas propiedades de la función indicador:

- $1_{A \cap B} = 1_A 1_B$
- $1_A + 1_{A^c} = 1$

**Ejemplo 1.15** Sean $A_1, A_2, \ldots, A_n$ eventos y $1_{A_i}$ la función indicador de $A_i$. Por propiedad, la unión $A = A_1 \cup A_2 \ldots, A_n$ tiene como función indicador:

$$1_A = 1 - \prod_{i=1}^{n}(1 - 1_{A_i})$$

El producto puede ser expandido y agrupado para obtener:

$$\begin{aligned}1_A = \sum_i 1_{A_i} - \sum_{i<j} 1_{A_i} 1_{A_j} + \sum_{i<j<k} 1_{A_i} 1_{A_j} 1_{A_k} - \cdots \\ + (-1)^{n+1} 1_{A_1} 1_{A_2} \cdots 1_{A_n}\end{aligned}$$

Tomando esperanzas, obtenemos la `fórmula de inclusión-exclusión`:

$$\mathbb{P}\Big(\bigcup_i A_i\Big) = \sum_i \mathbb{P}(A_i) - \sum_{i<j} \mathbb{P}(A_i \cap A_j) + \sum_{i<j<k} \mathbb{P}(A_i \cap A_j \cap A_k) - \cdots$$
$$\cdots + (-1)^{n+1}\mathbb{P}\Big(\bigcap_i A_i\Big)$$

Supongamos que $X$ e $Y$ tienen una densidad $f_{X,Y}$, entonces la densidad de $Z = X + Y$ se puede calcular como:

$$\begin{aligned}\mathbb{P}(Z \leq z) &= \mathbb{P}(X + Y \leq z)\\ &= \int \int_A f_{X,Y}(x,y)dxdy\end{aligned}$$

Donde $A = \{(x,y) \in \mathbb{R}^2 : x + y \leq z\}$. Expresando esta región en los límites de integración encontramos:

$$\begin{aligned}\mathbb{P}(Z \leq z) &= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{z-x} f_{X,Y}(x,y)dxdy\\ &= \int_{v=-\infty}^{z} \int_{u=-\infty}^{\infty} f_{X,Y}(u, v-u)dudv\end{aligned}$$

con la sustitución $u = x, v = x^2 + y$. Diferenciando esta ecuación con respecto a $z$ , se tiene:

$$f_Z(z) = \int_{-\infty}^{\infty} f_{X,Y}(u, z-u)du.$$

Un caso importante acerca de la suma de variables aleatorias es cuando $X$ e $Y$ son independientes, como indica el siguiente teorema:

**Teorema 1.1** Si las variables aleatorias $X$ e $Y$ son independientes y continuas con funciones densidad $f_X$ y $f_Y$, entonces la función densidad de $f_{X,Y}$ de $Z = X + Y$ es

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx \quad \text{para} \quad z \in \mathbb{R}.$$

**Ejemplo 1.16** Si $X$ e $Y$ son variables aleatorias independientes, con la distribución gamma con parámetros $s$ y $\lambda$ y la distribución gamma parámetros $t$ y $\lambda$. Entonces $Z = X + Y$ tiene una función densidad

$$\begin{aligned}f_Z(z) &= \int_{\infty}^{\infty} f_X(x)f_Y(z-x)\\ &= \begin{cases} \displaystyle\int_0^z f_X(x)f_Y(z-x) & \text{si } \ z > 0\\ 0 & \text{en otros casos}\end{cases}\end{aligned}$$

desde que $f_X(x)f_Y(z-x) = 0$, sin incluir el caso $x > 0$ y $z - x > 0$. Para $z > 0$

$$\begin{aligned} f_Z(z) &= \int_0^z \frac{1}{\Gamma(z)}\lambda(\lambda x)^{s-1}e^{-\lambda x}\frac{1}{\Gamma(t)}\lambda[\lambda(z-x)]^{t-1}e^{-\lambda(z-x)}dx \\ &= Ae^{-\lambda z}\int_0^z x^{s-1}(z-x)^{t-1}dx \end{aligned}$$

donde :

$$A = \frac{1}{\Gamma(s)\Gamma(t)}\lambda^{s+t}$$

Sustituyendo $y = z/x$ en la última integral obtenemos:

$$f_Z(z) = Bz^{s+t-1}e^{-\lambda z} \quad \text{si} \quad z > 0$$

donde $B$ es una constante dado por :

$$B = \frac{1}{\Gamma(s)\Gamma(t)}\lambda^{s+1}\int_0^1 y^{s-1}(1-y)^{t-1}dy$$

La única distribución con función densidad en este cálculo es la distribución gamma con parámetros $s+t$ y $\lambda$ y se sigue que la constante $B$ satisface

$$B = \frac{1}{\Gamma(s+t)}\lambda^{s+t}$$

Luego, la suma de dos variables aleatorias independientes con un distribución gamma con parámetros $s, \lambda$ y $t, \lambda$ respectivamente, tiene una distribución gamma con parámetros $s+t, \lambda$.

Comparando las dos últimas ecuaciones tenemos:

$$\int_0^1 y^{s-1}(1-y)^{t-1}dy = \frac{\Gamma(s)\Gamma(t)}{\Gamma(s+t)} \quad \text{para} \quad s, t > 0$$