Clase 7: Variables aleatorias continuas

Matías Carrasco

6 de agosto de 2019

Resumen Definimos las variables continuas basándonos en la noción de densidad de probabilidad. Luego generalizamos a más dimensiones, definiendo así la densidad conjunta de dos variables, y el concepto de independencia entre variables continuas.

Índice

Densidades de probabilidad

De las variables que no son discretas, las más importantes para nosotros serán aquellas cuya distribución se puede describir mediante *densidades* de probabilidad. Estas variables se llaman absolutamente continuas, aunque nosotros las llamaremos simplemente continuas.[1] Entre ellas destaca la distribución normal, o campana de Gauss, que estudiaremos más adelante.

[1] Existen variables que no son ni discretas, ni absolutamente continuas.

El concepto de densidad de probabilidad es similar al concepto de densidad de masa, o de densidad de carga (en caso de cargas positivas), que conocemos de los cursos de física y química. Por ejemplo, la densidad de masa es la cantidad de masa por unidad de volumen de una sustancia. Ver la Figura 1.

$$\text{densidad} = \frac{\text{masa}}{\text{volumen}}$$

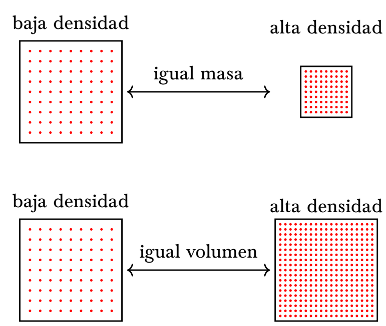


Figura 1: La densidad mide el amontonamiento de partículas.

Recordar que la densidad media de un volumen V es por definión

$$p_{\text{masa}} = \frac{m}{V} = \frac{\text{Masa del cuerpo}}{\text{Volumen que ocupa}}.$$

En general, un cuerpo está formado por una enormidad de partículas (del orden de 10^{24}), y cada una de ellas tiene una masa muy pequeña. Si no deseamos estudiar las propiedades moleculares de la materia, muchas veces usamos modelos en los cuales suponemos que las partículas "se aglomeran" para formar un continuo. De este modo, cada punto x del sólido tiene masa cero, pero la masa de un volumen finito, por más chico que sea, es positiva.

En general nos interesan cuerpos en los cuales la densidad de masa no es uniforme. Esto es, en el cual la densidad alrededor de un punto puede diferir de la densidad media; en algunos lugares la densidad es mayor que en otros. Ver Figuara 2.

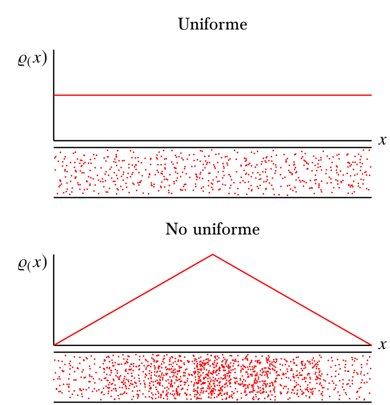


Figura 2: Una densidad no uniforme.

Si $\Delta_k V$ son pequeños volúmenes decrecientes centrados alrededor de un punto x, la densidad en el punto es

$$p_{\text{masa}}(x) = \lim_{k\to\infty} \frac{\Delta_k m}{\Delta_k V} = \frac{dm}{dV}(x).$$

La masa total del cuerpo viene entonces dada por la integral

$$\text{Masa}(\text{Cuerpo}) = \int_V p(x)dV(x),$$

en donde V es la región de $\mathbb{R}^3$ que ocupa el cuerpo.

Una densidad de probabilidad juega el mismo papel que la densidad de masa, cambiado masa por probabilidad. Como nos centraremos primero en variables aleatorias a valores reales, las densidades de probabilidad por el momento serán densidades en un espacio de dimensión uno. Esto equivale a densidades lineales en física, para las cuales la masa de un pequeño intervalo Δx centrado en el punto x es $\Delta m(x) \approx p(x)\Delta x$, ver la Figura 3. La única diferencia entre las densidades de masa y las densidades de probabilidad, es que estas últimas deben integrar siempre 1 pues la probabilidad total del espacio es siempre 1.

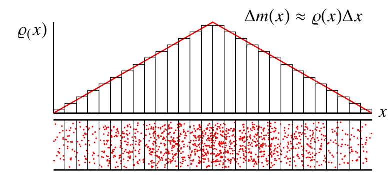


Figura 3: La masa de un pequeño intervalo Δx centrado en un punto x de un objeto con dimensión lineal está dada por $\Delta m(x) \approx p(x)\Delta x$.

Antes de comenzar con los ejemplos probabilísticos, es importante notar que las unidades de la densidad de masa no son las unidades de masa. De hecho, la densidad lineal tiene unidades de masa por unidad de longitud (Kg/m). Del mismo modo, las densidades de probabilidad no tendrán unidades de probabilidad (las unidades de probabilidad se pueden pensar como porcentajes), sino que tendrán unidades de, por ejemplo, probabilidad por unidad de longitud ($\%/m$). Por tanto, las densidades de probabilidad no son probabilidades, pero están relacionadas a estas de la misma forma que la densidad de masa está relacionada a la masa.

La distribución de Cauchy

Imaginemos un gas en un contenedor que tiene un pequeño agujero como se muestra en la Figura 4. Las moléculas del gas chocan contra las paredes del contenedor, pero algunas logran escapar por el agujero. La zona inmediata a las paredes del contenedor se mantiene al vacío, de modo que las moléculas que salen por el agujero viajan en linea recta hasta que chocan contra una pantalla que funciona como detector. La pantalla está a distancia L de las paredes del contenedor.

Nos gustaría conocer la probabilidad de que una molécula, que sale del contenedor por el agujero, choque a la pantalla a una distancia x del centro de la figura. Por supuesto, la probabilidad de que la molécula alcance a la pantalla a una distancia exactamente igual a x es cero, así que debemos discretizar el modelo y dividir el rango de valores posibles en intervalos.

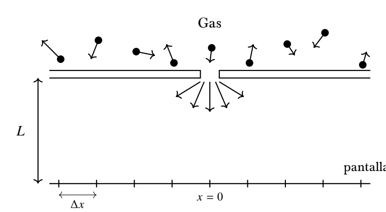


Figura 4: Gas en un contenedor que tiene un pequeño agujero.

Llamemos X a la variable aleatoria que indica la distancia al centro (con signo negativo del lado izquierdo) a la cual choca una molécula del gas con la pantalla. No existen límites, ni superior ni inferior, para los valores de X, de modo que el recorrido de la variable coincide con la recta real $\mathbb{R}$.

Dividimos la recta en intervalos, todos de longitud Δx, con extremos en los puntos $0, \pm\Delta x, \pm 2\Delta x, \ldots$. Llamemos a estos intervalos

$$I_k = [k\Delta x, (k+1)\Delta x],$$

con k entero. Denotamos además por $P_k(\Delta x) = P(X \in I_k)$ la probabilidad de que una molécula choque a la pantalla en el intervalo I_k.

Para calcular $P_k(\Delta x)$, vamos a suponer que las direcciones con las cuales salen las moléculas por el agujero son todas igualmente probables. Las moléculas entonces no salen con una dirección preferida, sino que la distribución de los ángulos de salida de las moléculas es uniforme en un intervalo de longitud π.

Debemos calcular entonces el ángulo que sustenta el intervalo I_k visto desde el centro del agujero. Llamemos a este ángulo θ_k (ver la Figura 5).

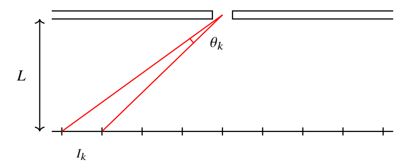


Figura 5: Ángulo que sustenta I_k.

Un poco de trigonometría permite concluir que

$$\theta_k \approx \frac{L\Delta x}{L^2 + (k\Delta x)^2},$$

siendo la aproximación cada vez mejor a medida que Δx se hace más pequeño. Entonces

$$P_k(\Delta x) = \frac{\theta_k}{\pi} \approx \frac{L\Delta x}{\pi\left(L^2 + (k\Delta x)^2\right)}.$$

Notar que al achicar Δx, la probabilidad de que X caiga en I_k se achica también. De hecho, tenemos que

$$P_k(\Delta x) \leq \frac{\Delta x}{\pi L} \to 0,$$

cuando Δx tiende a cero. Esto no es sorprendente pues sabemos que la probabilidad de que X valga un determinado valor puntual es cero.

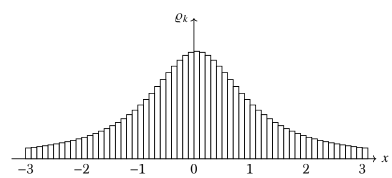


Figura 6: Aproximación de $P_k(\Delta x)$ con $\Delta x = 0.1$ para el valor de $L = 1$.

Que estas probabilidades sean pequeñas dificulta su visualización gráfica. Sin embargo, el siguiente truco nos permite evadir la dificultad. En lugar de graficar directamente los valores de $P_k(\Delta x)$, podemos hacer un gráfico de barras en el cual estas probabilidades vengan dadas por las áreas de las barras. Esto equivale a graficar barras rectangulares, cuyas bases sean los intervalos I_k, que son todos de longitud Δx, y sus alturas sean

$$p_k = \frac{P_k(\Delta x)}{\Delta x} \approx \frac{L}{\pi\left(L^2 + (k\Delta x)^2\right)}.$$

Por tanto, las probabilidades se leerán del gráfico midiendo áreas en lugar de alturas.

En la Figura 6 se muestra la aproximación de $P_k(\Delta x)$ con $\Delta x = 0.1$ para el valor de $L = 1$. Más adelante veremos como el parámetro L influye sobre la forma de la distribución.

En este gráfico las áreas de las barras representan las probabilidades $P_k(\Delta x)$. Notar que el máximo se da en $k = 0$ y vale $1/\pi$. Las alturas de las barras representan p_k la densidad de probabilidad media de cada intervalo I_k.

Veamos a qué converge este gráfico cuando Δx tiende a cero. Fijemos x un real cualquiera y tomemos k un entero tal que x pertenezca al intervalo I_k. No es difícil ver que para este valor de k, se tiene que $k\Delta x \approx x$. Esta aproximación es cada vez mejor a medida que decrece Δx. Entonces

$$p_k = \frac{P_k(\Delta x)}{\Delta x} \to \frac{L}{\pi\,(L^2 + x^2)},$$

cuando Δx tiende a cero. Llamando a este límite $p(x)$, concluimos que el gráfico de barras converge a medida que Δx tiende a cero al gráfico de la función $p(x)$. Ver la Figura 7.

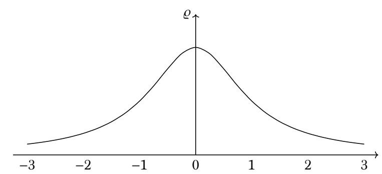


Figura 7: Gráfico de la densidad límite.

Para resumir, si en lugar de graficar directamente las probabilidades $P_k(\Delta x)$, graficamos las densidades medias de probabilidad p_k, entonces el gráfico converge al gráfico de la densidad de probabilidad p que está definida por

$$p(x) = \lim_{\Delta x \to 0} \frac{P_k(\Delta x)}{\Delta x} = \frac{L}{\pi\,(L^2 + x^2)}.$$

Esta densidad se conoce con el nombre de *densidad de Cauchy*.

Una manera informal de interpretar la densidad $p(x)$ es la siguiente: al tomar límite cuando Δx tiende a cero, podemos reemplazar Δx por el infinitésimo dx, de modo que $p(x)dx$ representa la probabilidad de que la variable X caiga en un intervalo infinitesimal centrado en el punto x de longitud dx.

Consideremos un intervalo $I = [a, b]$ en $\mathbb{R}$, y calculemos la probabilidad de que X caiga en I. Para un cierto valor fijo de Δx, tomemos $k_1\Delta x$ el extremo más cercano a a y $k_2\Delta x$ el más cercano a b. Entonces (ver la Figura 8)

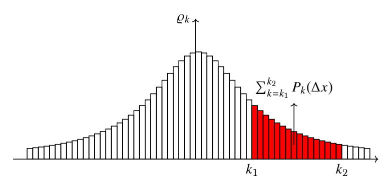


Figura 8: Aproximación discreta de la probabilidad de que X caiga en el intervalo $[a, b]$.

$$P\,(X \in I) \approx \sum_{k=k_1}^{k_2} P\,(X \in I_k) = \sum_{k=k_1}^{k_2} P_k(\Delta x) = \sum_{k=k_1}^{k_2} \frac{L}{\pi\,(L^2 + (k\Delta x)^2)}\Delta x.$$

Estas son sumas de Riemann de la función $p(x)$, por lo que al hacer Δx tender a cero obtenemos (ver Figura 9)

$$P\,(X \in I) = \int_a^b p(x)dx.$$

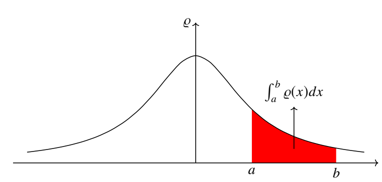


Figura 9: La probabilidad de que X caiga en el intervalo $[a, b]$ está dada por el área bajo el gráfico de la densidad $p(x)$.

Dicho en palabras, la probabilidad de que X caiga en un intervalo I es igual al área de la región comprendida entre I y el gráfico de la función p.

En particular, tomando el intervalo I igual a toda la recta, vemos que

$$\int_{-\infty}^{+\infty} p(x)dx = P\,(X \in \mathbb{R}) = 1.$$

Esto simplemente refleja el hecho de que las moléculas tocarán la pantalla en algún lugar de la recta. Esta propiedad es general: el área debajo la curva de una densidad de probabilidad es siempre igual a 1.

Podemos verificar directamente que la integral es uno en este caso. Notar primero que la primitiva de p es

$$\int_a^b \frac{L}{\pi\left(L^2+x^2\right)}dx = \left[\frac{1}{\pi}\arctan\left(\frac{x}{L}\right)\right]_a^b.$$

Luego, al tomar límites cuando $a \to -\infty$ y $b \to +\infty$ obtenemos

$$\int_{-\infty}^{+\infty} \frac{L}{\pi\left(L^2+x^2\right)}dx = 1.$$

Podemos intuir la influencia de L en la forma de la distribución de X, pensando físicamente en casos extremos para los cuales L es pequeño o respectivamente grande. Si la pantalla está muy cerca del agujero, la mayoría de las moléculas chocaran cerca del centro, por lo que la distribución de X se concentra más sobre $x = 0$. Recíprocamente, cuando la pantalla está muy lejos del agujero, una pequeña diferencia de ángulos hace que las moléculas toquen la pantalla en lugares muy distantes. Esto da como resultado una distribución de X mucho más esparcida, en donde el ancho de la campana es mayor. Esto lo podemos ver fácilmente en la Figura 10.

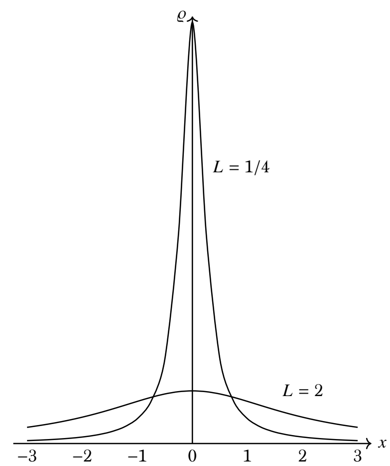


Figura 10: El parámetro L indica el "ancho" de la distribución.

De hecho, L es un indicador de cuán lejos caen las moléculas sobre la pantalla. Por ejemplo

$$P\left(|X| \leq kL\right) = \int_{-kL}^{kL} \frac{L}{\pi\left(L^2+x^2\right)}dx = \frac{\arctan(k) - \arctan(-k)}{\pi}.$$

Para $k = 1$ el resultado es $P\left(|X| \leq L\right) = 1/2$, por lo que esperamos ver que la mitad de las moléculas caen a una distancia menor que L del origen. Para $k = 2$ vemos que $P\left(|X| \leq 2L\right) \approx 0.70$ y para $k = 3$ es aproximadamente 0.80.

Variables continuas

Una densidad de probabilidad es una función $p : \mathbb{R} \to \mathbb{R}$ integrable que cumple las dos condiciones siguientes:

1. Es positiva: $p(x) \geq 0$ para todo $x \in \mathbb{R}$;
2. Integra uno:
$$\int_{-\infty}^{+\infty} p(x)dx = 1.$$

Es importante hacer notar dos cosas:

- La primera es que una densidad de probabilidad no tiene porqué ser menor o igual a uno. Por ejemplo, la densidad de Cauchy alcanza el máximo en $x = 0$ y este vale $1/\pi L$, que es mayor que 1 si $L < 1/\pi$.

- La segunda es que una densidad de probabilidad no tiene porque ser continua. Consideremos el siguiente ejemplo: supongamos que el experimento consiste en elegir un punto X con distribución uniforme en un intervalo I. Sabemos que para todo intervalo J, contenido o no en I, la probabilidad de que X caiga en J viene dada por

$$P(X \in J) = \frac{\mathrm{Long}(J \cap I)}{\mathrm{Long}(I)}.$$

  Esto lo podemos escribir de la siguiente manera

$$P(X \in J) = \int_J p(x)dx, \quad \text{en donde} \quad p(x) = \begin{cases} \frac{1}{\mathrm{Long}(I)} & \text{si } x \in I; \\ 0 & \text{si } x \notin I. \end{cases}$$

  Claramente p no es continua en los extremos de los intervalos, pues tiene un salto de cero a una constante positiva en dichos puntos. Ver la Figura 11.

  En general, si el recorrido de una variable continua está contenido en un intervalo, la densidad es igual a cero fuera del intervalo.

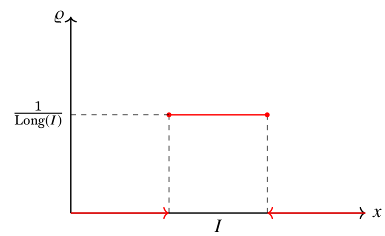


Figura 11: Densidad de la distribución uniforme.

Ahora que disponemos de la noción de densidad de probabilidad, podemos definir las variables aleatorias absolutamente continuas.

Variable aleatoria absolutamente continua

Una variable aleatoria X es absolutamente continua si existe una densidad de probabilidad p tal que

$$P(X \in I) = \int_I p(x)dx,$$

para todo intervalo I de $\mathbb{R}$. Para indicar que p es la densidad de X a veces escribiremos p_X.

Ejemplo 1 Sea X una variable aleatoria con densidad $p(x) = Cx^2$ en $[0,1]$. Hallar el valor de la constante C.

Como la probabilidad total debe ser igual a 1, tenemos

$$\int_0^1 p(x)dx = 1 \Leftrightarrow \int_0^1 Cx^2 dx = 1.$$

Evaluando la integral, la ecuación derecha se convierte en $C/3 = 1$, es decir $C = 3$.

Calcular $P(X \leq 1/2)$. Por definición, debemos integrar

$$P(X \leq 1/2) = \int_0^{1/2} p(x)dx = \int_0^{1/2} 3x^2 dx = 1/8.$$

■

Ejemplo 2 Sean X e Y las coordenadas de un punto elegido al azar en el interior del polígono que se muestra en la Figura 12. Hallar la densidad $Z = |X| + |Y|$.

Observar primero que Z toma valores en $[0, 1]$. Para $z \in [0, 1]$, el evento $\{z \leq Z \leq z + dz\}$ consiste de aquellos puntos que están dentro de la banda que se muestra en la Figura 13.

El lado exterior de la banda es igual a $\sqrt{2}(z + dz)$, y análogamente el lado interior es $\sqrt{2}z$. Como el área del polígono es 2,

$$P(z \leq Z \leq z + dz) = \frac{2(z + dz)^2 - z^2}{2} = 2z\,dz + dz^2.$$

Dividiendo por dz y tomando límite cuando dz tiende a cero, obtenemos

$$p(z) = \lim_{dz \to 0} \frac{2z\,dz + dz^2}{dz} = \lim_{dz \to 0} 2z + dz = 2z.$$

La densidad de Z es entonces $2z$ en $[0, 1]$. ■

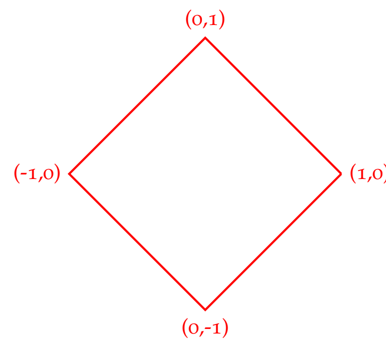


Figura 12: Elegimos un punto al azar dentro del polígono.

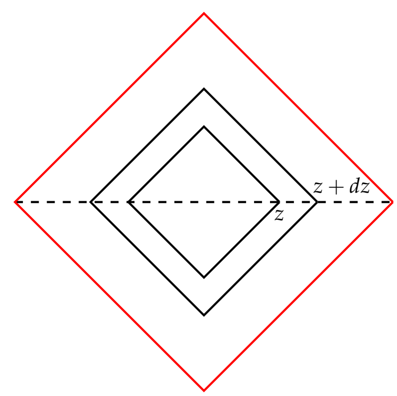


Figura 13: El evento $\{z \leq Z \leq z + dz\}$.

Función de distribución acumulada

Al igual que para variables discretas, definimos la función de distribución (acumulada) de una variable continua X como

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} p(u)du.$$

Igual que antes, escribimos F_X cuando queremos resaltar la variable X. Notar que la única diferencia es que remplazamos la suma por una integral. Visualmente la f.d.a. es el área debajo del gráfico de la densidad de X a la izquierda de x.

La f.d.a. cumple las misma propiedades, que si recuerdan bien fueron probadas en total generalidad usando los axiomas de Kolmogorov. Recordemos las principales:

- $0 \leq F(x) \leq 1$
- $F(x)$ es no-decreciente, i.e. si $x \leq y$ entonces $F(x) \leq F(y)$.
- $\lim_{x \to -\infty} F(x) = 0$ y $\lim_{x \to +\infty} F(x) = 1$
- $P(a < X \leq b) = F(b) - F(a)$

Además, la f.d.a. de una variable continua verifica

- $F(x)$ es continua, pues $P(X = x) = 0$ para todo $x \in \mathbb{R}$;
- $F'(x) = p(x)$ en todo punto x en donde F es derivable.

Esta segundo propiedad es muy útil para calcular la densidad de una variable.

Ejemplo 3 Sea X la variable definida en el Ejemplo 1. Hallar su f.d.a..

La densidad de X es $p(x) = 3x^2$ en $[0,1]$. Luego, para $x \in [0,1]$, tenemos

$$F(x) = \int_0^x p(u)du = \int_0^x 3u^2 du = x^3.$$

Además, si $x < 0$ entonces $F(x) = 0$, y si $x > 1$ entonces $F(x) = 1$. Así que

$$F(x) = \begin{cases} 0 & \text{si } x < 0; \\ x^3 & \text{si } 0 \leq x \leq 1; \\ 1 & \text{si } x > 1. \end{cases}$$

Observar que $P(X \leq 1/2) = F(1/2) = 1/8$ como habíamos calculado antes.

Ejemplo 4 Hagamos de nuevo el Ejemplo 2, pero usando ahora la última propiedad mencionada anteriormente. Para $z \in [0,1]$, el evento $\{Z \leq z\}$ consiste de aquellos puntos en el interior del polígono de vértice z, ver la Figura 14.

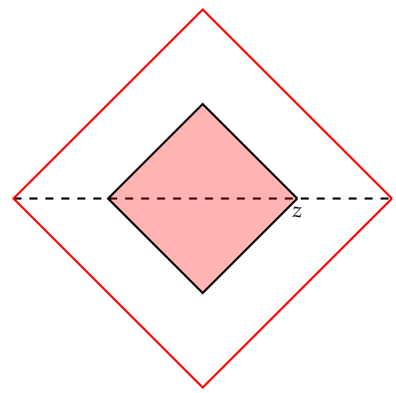


Figura 14: El evento $\{Z \leq z\}$.

Como el área del polígono es $2z^2$, vemos que $F(z) = P(Z \leq z) = 2z^2/2 = z^2$.

Si $z > 1$ entonces $F(z) = 1$, y si $z < 0$ entonces $F(z) = 0$. Resulta entonces

$$F(z) = \begin{cases} 0 & \text{si } z < 0; \\ z^2 & \text{si } 0 \leq z \leq 1; \\ 1 & \text{si } z > 1. \end{cases}$$

Está función es derivable, excepto en 0 y 1. En todos los demás puntos

$$p(z) = F'(z) = \begin{cases} 2z & \text{si } z \in (0,1); \\ 0 & \text{si } z \notin [0,1]. \end{cases}$$

Obtenemos así el mismo resultado que aplicando la definición de densidad.

Densidad conjunta e independencia

La noción de densidad de probabilidad $p(x,y)$ en dos variables es una extensión natural de la idea de densidad de probabilidad en dimension uno. La función $p(x,y)$ da la densidad de probabilidad por unidad de área para valores de (X,Y) cercanos a (x,y). Ver la Figura 15.

Ya hemos consideramos un caso particular de densidad de probabilidad en el plano cuando vimos probabilidades geométricas. En ese caso, si Ω es una región acotada del plano, y (X,Y) es un punto elegido al azar en esa región, entonces la densidad de probabilidad conjunta

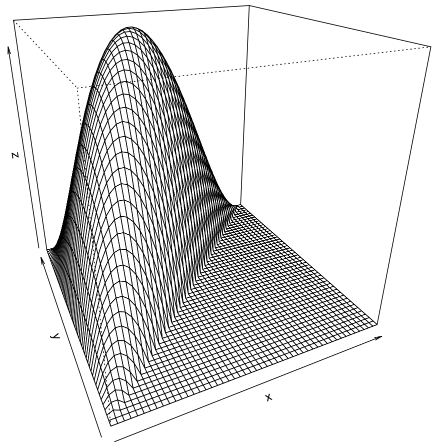


Figura 15: Aquí se muestra una función de densidad conjunta particular cuya fórmula es $p(x,y) = 5!x(y-x)(1-y)$ para $0 < x < y < 1$, y cero en otro caso.

del par (X,Y) es

$$p(x,y) = \begin{cases} 1/\text{Area}(\Omega) & \text{si } (x,y) \in \Omega; \\ 0 & \text{si no.} \end{cases}$$

En general, cualquier evento determinado por dos variables aleatorias X e Y, como el evento $X > 0.25$ e $Y > 0.5$, corresponde a una región del plano. Ahora, en lugar de una distribución uniforme definida por áreas relativas, la probabilidad de la región B se describe por el volumen bajo la superficie de densidad sobre B. Este volumen es una integral

$$P\left((X,Y) \in B\right) = \int\int_B p(x,y)dxdy.$$

Este es el análogo de la familiar interpretación del área bajo la curva para probabilidades obtenidas a partir de densidades en la recta real.

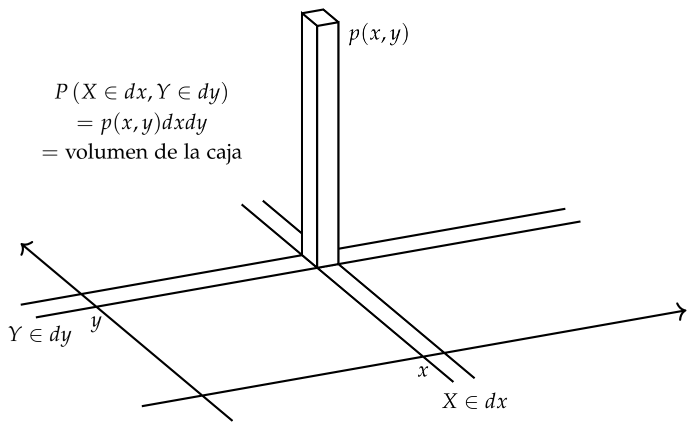
Los ejemplos que vienen muestran cómo tales integrales se pueden calcular mediante integración repetida, cambio de variable o argumentos de simetría.

La distribución uniforme en una región es ahora solo el caso especial en el cual $p(x,y)$ es constante en la región y cero en otro lugar. Como regla general, las fórmulas que involucran densidades conjuntas son análogas a las fórmulas correspondientes para distribuciones conjuntas discretas que vimos anteriormente.

Informalmente, si (X,Y) tienen densidad conjunta $p(x,y)$, entonces la probabilidad infinitesimal es

$$P\left(X \in dx, Y \in dy\right) = p(x,y)dxdy.$$

Esto significa que la probabilidad de que el par (X,Y) caiga en un rectángulo infinitesimal de ancho dx y altura dy cerca del punto (x,y)



es la densidad de probabilidad en (x, y) multiplicada por el área $dxdy$ del rectángulo.

Una densidad de probabilidad bi-variada debe cumplir las siguientes dos condiciones

$$p(x,y) \geq 0, \text{ y } \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} p(x,y)dxdy = 1.$$

Si (X, Y) tiene densidad conjunta $p(x, y)$ en el plano, cada una de las variables aleatorias X e Y tiene densidad en la recta. Estas se llaman *densidades marginales*.

Las densidades marginales se pueden calcular a partir de la densidad conjunta mediante integrales análogas a las fórmulas discretas para probabilidades marginales como sumas de filas y columnas en una tabla de contingencia. La probabilidad de puntos discretos se reemplaza por densidades y sumas por integrales:

$$p_X(x) = \int_{-\infty}^{+\infty} p(x,y)dy;$$
$$p_Y(y) = \int_{-\infty}^{+\infty} p(x,y)dx.$$

En general, dos variables aleatorias X e Y se dicen independientes si

$$P(X \in A, Y \in B) = P(X \in A)\, P(X \in B),$$

para cualquier elección de conjuntos A y B en la recta real.

Variables independientes

Cuando el par (X, Y) tiene densidad conjunta $p(x, y)$, las variables X e Y son independientes si, y solo si, la densidad conjunta es igual al producto de las densidades marginales

$$p(x,y) = p_X(x)p_Y(y) \text{ para todo } x, y \in \mathbb{R}.$$

Intuitivamente, esta ecuación se deduce de la definición tomando A como el intervalo infinitesimal $(x, x + dx)$ y B como el intervalo infinitesimal $(y, y + dy)$, obteniendo así

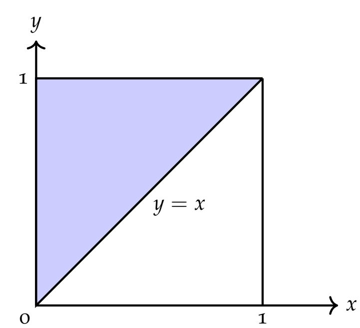
$$P(X \in dx, Y \in dy) = P(X \in dx) P(Y \in dy)$$

por lo que

$$p(x,y)dxdy = p_X(x)dx p_Y(y)dy.$$

Cancelando los infinitésimos se obtiene la formula para las densidades. Recíprocamente, se va de las densidades a las probabilidades integrando.

Ejemplo 5 Supongamos que (X, Y) se distribuye uniformemente en el triángulo $T = \{(x, y) : 0 < x < y < 1\}$. La densidad conjunta de (X, Y) es simplemente



$$p(x,y) = \begin{cases} 2 & \text{si } (x,y) \in T; \\ 0 & \text{si no.} \end{cases}$$

Las densidades marginales son

$$\begin{aligned} p_X(x) &= \int_{-\infty}^{+\infty} p(x,y)dxdy \\ &= \int_{y=x}^{y=1} 2dy \text{ (pues } p(x,y) = 2 \text{ si } 0 < x < y < 1 \text{ y 0 si no)} \\ &= 2(1-x) \text{ (si } 0 < x < 1 \text{ y cero si no)} \end{aligned}$$

$$\begin{aligned} p_Y(y) &= \int_{-\infty}^{+\infty} p(x,y)dxdy \\ &= \int_{x=0}^{x=y} 2dx \text{ (pues } p(x,y) = 2 \text{ si } 0 < x < y < 1 \text{ y 0 si no)} \\ &= 2y \text{ (si } 0 < y < 1 \text{ y cero si no)} \end{aligned}$$

Con esto podemos ver que X e Y no son independientes ya que $p(x,y) \neq p(x)p(y)$. ■

Ejemplo 6 Sean X e Y independientes con densidades dadas por[2]

$$\begin{aligned} p_X(x) &= \lambda e^{-\lambda x} \quad (x > 0), \\ p_Y(y) &= \mu e^{-\mu y} \quad (y > 0). \end{aligned}$$

respectivamente. Calculemos $P(X < Y)$.

La densidad conjunta de (X, Y) es

$$p(x,y) = \left(\lambda e^{-\lambda x}\right)\left(\mu e^{-\mu y}\right) = \lambda\mu e^{-\lambda x - \lambda y}$$

para x e y positivos. Esto se deduce de la independencia de X e Y. Luego, la probabilidad $P(X < Y)$ se calcula integrando $p(x,y)$ sobre

[2] Dejamos como ejercicio verificar que son densidades de probabilidad.

la región $\{(x, y) : x < y\}$:

$$\begin{aligned} P\left(X < Y\right) &= \int\int_{x<y} \lambda\mu e^{-\lambda x - \lambda y} dx dy = \int_{x=0}^{\infty} dx \int_{y=x}^{\infty} \lambda\mu e^{-\lambda x - \lambda y} \\ &= \int_{x=0}^{\infty} \lambda e^{-(\lambda+\mu)x} dx = \frac{\lambda}{\lambda + \mu} \end{aligned}$$

Notar que se obtiene el mismo resultado si se integra en el otro orden, esto siempre es así. ■