# Unidad 3. Variables aleatorias continuas

El siguiente material es sólo orientativo, está sujeto a correcciones y se irá completando con el intercambio en el aula.

## 1. Conceptos generales

**Definición 1.** *Una variable aleatoria $X$ se llama continua sii existe una función*

$$f_X : \mathbb{R} \to \mathbb{R}_{\geq 0}$$

*tal que la función de distribución $F_X$ satisface*

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

*A la función $f_X$ se le denomina función de densidad.*

Salvo que sea necesario de aquí en más, para simplificar, omitiremos el subíndice $X$.

**Ejemplo 1.** *Sea $X$ una variable aleatoria con función de distribución $F$ dada por $F(x) = \int_{-\infty}^{x} f(t)dt$ donde la densidad $f$ viene dada por*

$$f(x) = \begin{cases} 1/2 & \text{si } 1 \leq x \leq 3 \\ 0 & \text{en caso contrario} \end{cases}$$

*Para hallar la forma explícita de $F$ notar que:*

- *si $x < 1$, $F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{x} 0dt = 0$*
- *si $1 \leq x \leq 3$, $F(x) = \int_{-\infty}^{x} f(t)dt = \int_{-\infty}^{1} 0dt + \int_{1}^{x} \frac{1}{2}dt = \frac{1}{2}y\Big|_1^x = \frac{x-1}{2}$*
- *si $x \geq 3$, $F(x) = 1$.*

*Así*

$$F(x) = \begin{cases} 0 & \text{si } x \leq 1 \\ \frac{x-1}{2} & \text{si } 1 \leq x \leq 3 \\ 1 & \text{si } x \geq 3 \end{cases} \tag{1}$$

*En la Figura 1 se muestran los gráficos de $f$ y $F$*

Notar entonces que, por el Teorema Fundamental del Cálculo y por ser por ser $F$ una función de distribución podemos establecer las siguientes afirmaciones.

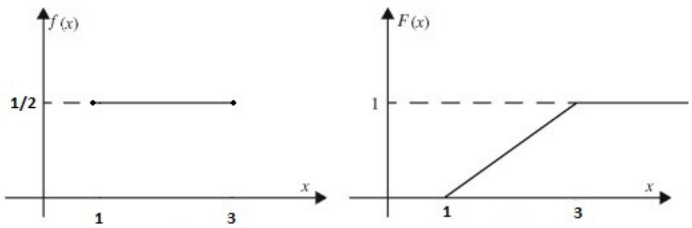


**Figura 1:** Gráficos de la densidad $f$ y la función de distribución $F$ del Ejemplo 1.

**Proposición 1.** *Si $X$ es una variable aleatoria continua con función de distribución $F$ entonces:*

*i) $F$ resulta una función continua en toda la recta real y,*

*ii) Si $x_0$ es un punto de continuidad de la densidad $f$ entonces*

$$F'(x_0) = f(x_0),$$

*iii) Además,*

$$\int_{-\infty}^{\infty} f(t)dt = 1.$$

Cuando $X$ es una variable aleatoria discreta, la función de probabilidad de masa $f$ en $x$ es un valor de probabilidad $P(X = x) > 0$ si $x$ pertenece al rango. Si ahora $f$ es la densidad asociada a una variable aleatoria continua ¿es posible que $\forall x : f(x) = P(X = x)$? La respuesta es no. Analicemos el porqué: en la Unidad 2 establecimos que cualquiera sea una función de distribución $F$ se satisface que

$$\forall x \in \mathbb{R} : P(X = x) = F(x) - \lim_{y \to x^-} F(y),$$

implicando que, si $F$ es continua, $\forall x \in \mathbb{R} : P(X = x) = 0$. Si fuera $\forall x : f(x) = P(X = x)$ entonces sería $f \equiv 0$ por lo tanto también, por la Definición 1, $F \equiv 0$ lo cual es absurdo ya que $F$ es una función de distribución y $\lim_{x\to\infty} F(x) = 1$.

Resumimos esta discusión en la siguiente proposición y su corolario.

**Proposición 2.** *Si $X$ es una variable aleatoria continua entonces*

$$\forall x \in \mathbb{R} : P(X = x) = 0.$$

**Corolario 1.** *Si $X$ es una variable aleatoria discreta entonces $X$ no es una variable aleatoria continua.*

En la Unidad 2 enunciamos y probamos que para cualquier función de distribución $F$ se cumple que

$$\forall a, b \in \mathbb{R} : \text{ si } a < b \text{ entonces } P(a < X \leq b) = F(b) - F(a).$$

De esta afirmación y de la Proposición 2 se establece la validez de la proposición enunciada a continuación y cuya demostración dejamos como ejercicio.

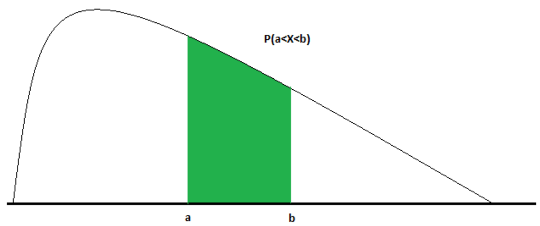


**Figura 2:** Representación de la probabilidad $P(a < X < b)$ como una área, en verde, bajo la curva de la densidad $f$ en el intervalo $[a, b]$.

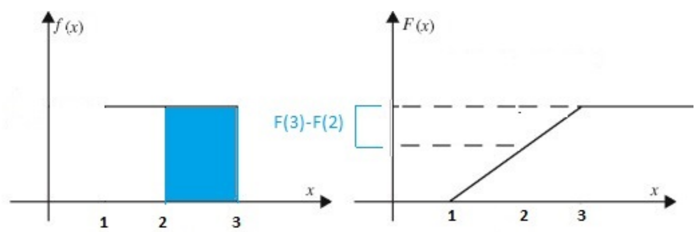


**Figura 3:** $P(2 \leq X < 3) = \frac{1}{2}$ marcada en color azul:en el panel de la izquierda representa un área y en el de la derecha una diferencia de valores de $F$.

**Proposición 3.** *Si $X$ es una variable aleatoria continua con densidad $f$ y función de distribución $F$, para cualesquiera $a$ y $b$ con $a < b$:*

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t)dt \tag{2}$$

*y además $P(a < X \leq b) = P(a \leq X < b) = P(a \leq X \leq b) = P(a < X < b)$.*

Notar que (2) establece la relación entre una probabilidad y el área bajo la curva de la densidad y, en la Figura 2 se muestra una representación de esta relación.

**Problema 1.** *Continuación del Ejemplo 1*

*Calcular $P(2 \leq X < 3)$ y asociar este valor a los gráficos de $f$ y $F$.*

**Solución.**Utilizando la Proposición 3 obtenemos

$$P(2 \leq X < 3) = F(3) - F(2) = \frac{3-1}{2} - \frac{2-1}{2} = 1/2,$$

valor que se corresponde con un área marcada en azul en el panel izquierdo de la Figura 3 y para una diferencia de valores de $F$ indicada en el panel de la derecha de esa misma figura, también en azul.

**Ejemplo 2.** *Tiempo de vida*

El tiempo de vida $X$, en años, de un componente electrónico es una variable aleatoria continua con densidad

$$f(x) = \begin{cases} \frac{k}{x^3} & \text{si} \quad x \geq 1 \\ 0 & \text{si} \quad x < 1 \end{cases}$$

Nos preguntamos

- ¿Cuál es valor de $k$ tal que $f$ sea una función de densidad? y,
- ¿Cuál es la probabilidad de que el tiempo de vida sobrepase los 5 años?.

En primer lugar notar que de

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{1}^{+\infty} \frac{k}{x^3}dx = -\left.\frac{k}{2x^2}\right|_{1}^{+\infty} = \frac{k}{2} = 1$$

se obtiene $k = 2$. Luego

$$F(x) = 1 - \frac{1}{x^2}$$

para $x > 1$ y $F(x) = 0$ si $x \leq 1$. De aquí concluimos que

$$P(X > 5) = 1 - F(5) = \frac{1}{25}.$$

El Corolario 1 nos está sugiriendo que los conceptos de discreto y continuo se refieren a objetos bien diferentes y van a permitir modelar situaciones distintas de la realidad.

En la Unidad 2 utilizamos la variable $X$ con rango $R_X = \{0, 1\}$ definida a partir del experimento aleatorio de lanzar una moneda, donde 1 representa "cara" y, además asociamos un modelo físico. ¿Cuál es un modelo físico asociado a una variable aleatoria continua?

En primer lugar tengamos en cuenta la siguiente proposición.

**Proposición 4.** *Sea $X$ una variable aleatoria continua y sea $f$ su función de densidad. Si $f$ es continua en $x_0$, entonces*

$$\lim_{h \to 0} \frac{P\left(x_0 - h \leq X \leq x_0 + h\right)}{2h} = \lim_{h \to 0} \frac{1}{2h} \int_{x_0 - h}^{x_0 + h} f(t)dt = f\left(x_0\right)$$

**Demostración.** Ver apéndice.

Es importante hacer notar que ii) de la Proposición 1 es una consecuencia de esta proposición.

Más informalmente podemos escribir que para $h > 0$ pequeño

$$P\left(x_0 - h \leq X \leq x_0 + h\right) \approx 2hf(x_0)$$

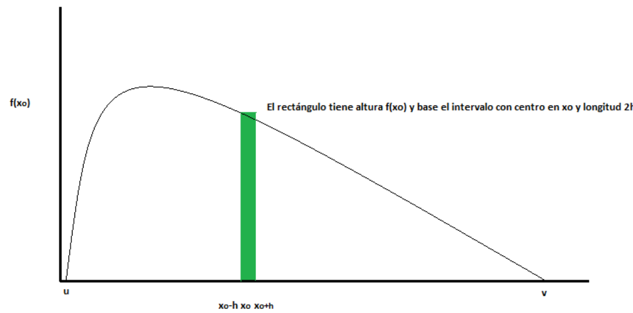


**Figura 4:** Representación de la probabilidad $P\,(x_0 - h \leq X \leq x_0 + h)$ como “cantidad de masa por unidad de longitud”.

La Figura 4 da una representación de esta relación: el área del rectángulo, en verde, es aproximadamente el área bajo la curva de la densidad $f$ en el intervalo $[x_0 - h, x_0 + h]$.

Para obtener un modelo físico podemos representar con el segmento $[u, v]$ una varilla de metal y la densidad $f$ en $x_0$ puede interpretarse como “cantidad de masa por unidad de longitud” en un punto cualquiera $x_0$ de la varilla. En un contexto probabilístico, podría decirse que $f$ indica la probabilidad por unidad de longitud en las cercanías del punto $x_0$ (Yohai, 2006, p. 51).

Queremos ahora generalizar la noción de variable aleatoria continua a la de un vector aleatorio continuo. Dadas dos variables aleatorias, discretas o continuas, un vector aleatorio de dimensión 2 es el par $(X, Y)$ y su función de distribución conjunta es $F(x, y) = P(X \leq x, Y \leq y)$ para cualesquiera $x, y \in \mathbb{R}$. La función $F$ resulta monótona creciente en cada componente y

$$\begin{aligned} \lim_{x\to\infty, y\to\infty} F(x,y) &= 1 \\ \forall x, y : \lim_{u\to-\infty} F(x,u) &= \lim_{v\to-\infty} F(v,y) = 0. \end{aligned}$$

Las funciones de distribución de $X$ y de $Y$, $F_X$ y $F_Y$, se denominan marginales.

La noción de variable aleatoria continua se generaliza del siguiente modo.

**Definición 2.** *Un vector aleatorio $(X, Y)$ se le llama vector aleatorio continuo si existe $f : \mathbb{R}^2 \to \mathbb{R}_{\geq 0}$ tal que su función de distribución $F$ satisface*

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(u, v)dudv.$$

*A $f$ se le llama función de densidad.*

Cuando $f$ es continua en un punto $(x_0, y_0)$ podemos establecer la siguiente relación

$$f(x_0, y_0) = \frac{\partial^2}{\partial y \partial x} F(x, y) \mid_{(x_0, y_0)} . \tag{3}$$

Si $(X, Y)$ es un vector aleatorio continuo entonces (ver Yohai, 2006, Teorema 5.9, pág. 101 ) $X$ e $Y$ resultan variables aleatorias continuas y sus funciones de densidad (marginales) satisfacen

$$\forall x, y \in \mathbb{R}: \; f_X(x) = \int_{-\infty}^{\infty} f(x, v)dv \tag{4}$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(u, y)du \tag{5}$$

Las definiciones y resultados anteriores se extienden a vectores $(X_1, \ldots, X_p)$ con $p > 2$.

## 2. Esperanza y Varianza

Al igual que en las variables aleatorias discretas hay características numéricas asociadas a ellas, como la esperanza, la varianza y el desvío.

**Definición 3.** *Si $X$ es una variable aleatoria continua entonces la esperanza se define como*

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x)dx$$

*si la integral $\int_{-\infty}^{\infty} |x| f(x)dx$ es finita.*

**Problema 2.** *Hallar $E(X)$ para la variable aleatoria indroducida en el Ejemplo 1*

La densidad de $X$ es

$$f(x) = \begin{cases} 1/2 & \text{si } 1 \leq x \leq 3 \\ 0 & \text{en caso contrario} \end{cases}$$

y, en consecuencia,

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x)dx \\ &= \int_1^3 x \frac{1}{2} dx \\ &= \frac{1}{2} \frac{x^2}{2} \Big|_1^3 \\ &= 2 \end{aligned}$$

De este modo, la media o esperanza de $X$ es 2 $\square$

**Problema 3.** *En el contexto del Ejemplo 2, hallar el tiempo de vida medio de un componente electrónico.*

Por definición

$$\mathrm{E}\left(X\right)=\int_{1}^{\infty}x2x^{-3}dx=-2x^{-1}\Big|_{1}^{\infty}=2$$

Así, la esperanza de vida es de dos años. □

El concepto de esperanza también tiene en este contexto la misma interpretación física, la de ser el *centro de equilibrio o de gravedad*, tal como fue explicitado en la Unidad 2. En el Ejemplo 1 el centro de gravedad es el punto medio del intervalo $[1,3]$ porque la densidad es uniforme, pero en el Ejemplo 2 el centro de equilibro se encuentra en $x=2$ aún la densidad sea positiva en la semirecta $[1,\infty]$.

A continuación enunciamos la ley del inconsciente estadístico para variables aleatorias continuas. La enunciación formal completa y su demostración están fuera del alcance de este curso.

**Proposición 5.** ***Ley del inconsciente estadístico***

*Dada una variable aleatoria $X$ continua con función de densidad $f$ y sea $g:\mathbb{R}\to\mathbb{R}$ una función, entonces*

$$\mathrm{E}\left[g(X)\right]=\int_{-\infty}^{\infty}g(x)f(x)dx.$$

Esta proposición se puede generalizar al caso multivariado. En efecto, si $(X_1,\ldots,X_p)$ es un vector aleatorio continuo con con función de densidad $f$ y $g:\mathbb{R}^p\to\mathbb{R}$ entonces

$$\mathrm{E}\left[g(X_1,\ldots,X_p)\right]=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}g(x_1,\ldots,x_p)dx_1\ldots dx_p. \tag{6}$$

Al igual que antes se definen los conceptos de varianza y desvío

**Definición 4.** *Si $X$ es una variable aleatoria continua entonces la varianza de $X$ se define como*

$$\textsc{var}(X)=\mathrm{E}\left[\left(X-\mathrm{E}\left(X\right)\right)^2\right].$$

*y a $\textsc{desv}(X)=\sqrt{\textsc{var}(X)}$ se le denomina desvío de la variable aleatoria .*

Y por la Proposición 5 se tiene que

$$\textsc{var}\left(X\right)=\int_{-\infty}^{\infty}(x-\mathrm{E}\left(X\right))^2f(x)dx, \tag{7}$$

si $f$ es la densidad de la variable aleatoria continua $X$. Y, de modo análogo a como establecimos en la Unidad 2, es válida la siguiente igualdad

$$\mathrm{VAR}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 \,. \tag{8}$$

Las siguientes propiedades ya enunciadas para variables aleatorias discretas también valen para continuas.

**Proposición 6.** *Son válidas las siguientes afirmaciones.*

*a)* *Si* $X \geq 0$[1] *entonces* $\mathrm{E}(X) \geq 0$.

*b)* *Para cualesquiera constantes* $a$ *y* $b$

$$\begin{aligned} \mathrm{E}(aX + bY) &= a\mathrm{E}(X) + b\mathrm{E}(Y) & (9) \\ \mathrm{VAR}(aX + bY) &= a^2\mathrm{VAR}(X) + b^2\mathrm{VAR}(Y) + 2ab\mathrm{E}\left[(X - \mathrm{E}(X))\,(Y - \mathrm{E}(Y))\right] & (10) \end{aligned}$$

**Demostración.** Ejercicio.

# 3. Familias de variables aleatorias continuas

Existe una enorme variedad de familias de variables aleatorias continuas, nosotros estudiaremos aquí algunas de ellas. Cada familia está indexada por uno o más parámetros.

## 3.1. Distribución uniforme

Dado un intervalo $[a, b] \subset \mathbb{R}$, una variable aleatoria $X$ se dice que tiene *distribución uniforme en* $[a, b]$ si su función de densidad es

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{en caso contrario} \end{cases} \tag{11}$$

Notar que $f$ se puede escribir también como $f = (1/(b-a))I_{[a,b]}$ donde $I$ denota la función indicadora.[2]

Los límites $a$ y $b$ del intervalo son los *parámetros* del modelo. Utilizaremos la notación $X \sim$ uniforme$[a, b]$.[3] Cuando $a = 0$ y $b = 1$ la distribución se denomina *uniforme estándar*.

La función de distribución $F$ está dada por la siguiente expresión:

[1] O sea, si la función de densidad se anula en los reales negativos.
[2] Dado un subconjunto $A$, definimos la función indicadora $I_A(x)$ como la función que vale 1 si $x \in A$ y 0 en caso contrario.
[3] Hay que tener en cuenta que la distribución puede definirse también sobre un intervalo abierto o semiabierto.

$$F(x) = \begin{cases} 0 & \text{si } x \leq a \\ \frac{x-a}{b-a} & \text{si } a \leq x \leq b \\ 1 & \text{si } x \geq b \end{cases} \tag{12}$$

El Ejemplo 1 introdujo un miembro de esta familia de distribuciones uniformes.

La siguiente afirmación muestra que cualquier miembro de la familia puede ser transformado en la distribución uniforme estándar a través de una aplicación lineal.

**Proposición 7.** *Sean $a$ y $b$ dos números reales con $a < b$. Es válida la siguiente equivalencia:*

$$X \sim \textit{uniforme}[a, b] \textit{ si y sólo si } Y = \frac{X-a}{b-a} \sim \textit{uniforme}[0, 1].$$

**Demostración**. Ejercicio.

### 3.1.1. Esperanza y varianza.

Notar que

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \int_a^b x \frac{1}{b-a} dx \\ &= \frac{1}{b-a} \frac{x^2}{2} \Big|_a^b \\ &= \frac{a+b}{2} \end{aligned}$$

Utilizando (8) se prueba fácilmente que $\text{VAR}(X) = \dfrac{(b-a)^2}{12}$. Resumiendo,

$$\mathrm{E}(X) = \frac{a+b}{2} \text{ y } \text{VAR}(X) = \frac{(b-a)^2}{12}.$$

## 3.2. Distribución gamma

La función "gamma" definida para los reales positivos, denotada por $\Gamma(t)$, se define como

$$\Gamma(t) = \int_0^{\infty} x^{t-1} \mathrm{e}^{-x} dx, \ \forall t > 0.$$

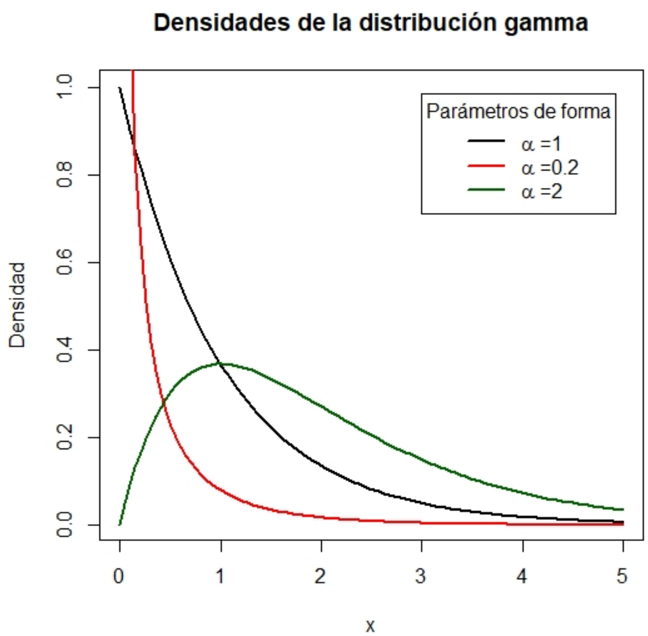


**Figura 5:** Densidades de $X \sim \text{gamma}(\alpha, \lambda)$ con $\lambda = 1$ y $\alpha = 0.2, 1, 2$.

Esta función tiene la propiedad que cualquiera sea $t > 0 : \Gamma(t+1) = t\Gamma(t)$ y, en consecuencia

$$\forall n \in I\!N : \Gamma(n+1) = n!,$$

lo que constituye una generalización del factorial.

Diremos que una variable aleatoria $X$ tiene *distribución gamma de parámetros* $\alpha, \lambda > 0$, $X \sim \text{gamma}(\alpha, \lambda)$ si su función de densidad definida para los reales positivos es

$$f(x) = \frac{\lambda^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \quad x > 0. \tag{13}$$

A $\alpha$ se le llama *parámetro de forma* y $\lambda$ es un *parámetro de escala.*

**Problema 4.** *Comprobar que $f$ definida en* (13) *es una función de densidad.*

**Solución:** ejercicio.

En la Figura 5 se grafican, para un parámetro de escala fijo, las curvas de diferentes densidades variando el parámetro de forma.

La función de distribución de $X \sim \text{gamma}(\alpha, \lambda)$ no tiene una fórmula explícita y para hallar un valor $F(x)$ se utiliza una aproximación numérica.

**Ejemplo 3.** *Los usuarios de internet visitan cierto sitio en un promedio de* 12 *visitas por minuto.*

*Cada seis visitantes uno recibe una promoción por medio de un banner intermitente. Entonces el tiempo consecutivo entre promociones puede modelarse por una distribución gamma con parámetros* $\alpha = 6$ *y* $\lambda = 12$.

### 3.2.1. Esperanza y varianza

Teniendo en cuenta que $\forall \alpha > 0$ y $\lambda > 0$:

$$\int_0^\infty x^{\alpha-1} e^{-\lambda x} dx = \frac{\Gamma(\alpha)}{\lambda^\alpha}$$

entonces

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} x f(x) dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^\alpha e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+1)}{\lambda^{\alpha+1}} = \frac{\alpha}{\lambda}. \end{aligned}$$

Análogamente

$$\begin{aligned} \mathrm{E}\left(X^2\right) &= \int_{-\infty}^{\infty} x^2 f(x) dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^\infty x^{\alpha+1} e^{-\lambda x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha+2)}{\lambda^{\alpha+2}} = \frac{(\alpha+1)\alpha}{\lambda^2} \end{aligned}$$

de donde,

$$\textsc{var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = \frac{(\alpha+1)\alpha - \alpha^2}{\lambda^2} = \frac{\alpha}{\lambda^2}.$$

Resumiendo si $X \sim \text{gamma}(\alpha, \lambda)$,

$$\mathrm{E}(X) = \frac{\alpha}{\lambda} \text{ y } \textsc{var}(X) = \frac{\alpha}{\lambda^2}$$

son dos de sus características numéricas.

## 3.3. Distribución exponencial

Diremos que una variable aleatoria $X$ tiene *distribución exponencial* con *parámetro de escala* $\lambda > 0$, $X \sim \text{exponencial}(\lambda)$, si $X \sim \text{gamma}(1, \lambda)$. De este modo, su densidad es

$$f(x) = \lambda e^{-\lambda x}, \quad x > 0. \tag{14}$$

En la Figura 5 se grafica la curva de una densidad de esta familia.

### 3.3.1. Esperanza y varianza

Dado que la distribución exponencial es un caso particular de la distribución gamma se tiene que

$$\mathrm{E}(X) = \frac{1}{\lambda} \text{ y } \mathrm{VAR}(X) = \frac{1}{\lambda^2}.$$

### 3.3.2. Tiempos de interarribo

Consideremos una variable aleatoria $X$ que cuenta "el número de eventos raros que ocurren en un intervalo de tiempo $t > 0$" y asumamos que tiene una distribución Poisson con un parámetro proporcional a $t$, digamos $\lambda t$ con $\lambda > 0$.

Consideremos ahora el suceso "que el tiempo, $T$, que transcurre hasta el próximo evento raro sea mayor que $t$", es decir $(T > t)$. Este suceso puede ser descripto también como "cero eventos raros ocurren hasta el instante $t$" o, equivalentemente, $X = 0$. Notar que

$$P(X = 0) = \mathrm{e}^{-\lambda t} \frac{(\lambda t)^0}{0!}$$

y, en consecuencia la función de distribución de $T$ es

$$F_T(t) = 1 - P(T > t) = 1 - P(X = 0) = 1 - \mathrm{e}^{-\lambda t}.$$

Es decir, $T \sim \text{exponencial}(\lambda)$.

**Problema 5.** *Los trabajos son enviados a una impresora a un promedio de 3 por hora. Queremos calcular el tiempo esperado entre trabajos y la probabilidad de que el próximo trabajo sea enviado en el lapso de los próximos cinco minutos.*

**Solución:** ejercicio.

### 3.4. Distribución normal o Gaussiana

Diremos que una variable aleatoria $X$ tiene *distribución normal* con parámetros $\mu \in \mathbb{R}$ y $\sigma^2 > 0$, $X \sim \mathrm{N}(\mu, \sigma^2)$, si su densidad es

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2},\ x \in \mathbb{R}. \tag{15}$$

A $\mu$ se le denomina *parámetro de posición* y a $\sigma$ *parámetro de escala*. Por otro lado hay que tener en cuenta que $x = \mu$ es el eje de simetría de la curva y $\sigma^2$ controla la concentración de probabilidad en torno a ese eje.

**Proposición 8.** *La función* (15) *define una función densidad.*

**Demostración:** Es claro que $\forall x:\ f(x) \geq 0$. En Yohai (2006, p. 60 y sgts.) se prueba que

$$\int \mathrm{e}^{-t^2/2} dt = \sqrt{2\pi}.$$

De esta igualdad y por sustitución se concluye que

$$\int f(x) dx = 1.$$

Hemos probado que $f$ satisface las condiciones para ser una función de densidad. □

En la Figuras 6 y 7 graficamos curvas de densidades de distribuciones Gaussianas. En la primera figura dejamos fijo el parámetro de posición para ilustrar el impacto del parámetro de escala en la gráfica, mientras que en la segunda $\sigma^2$ permanece fijo y variamos $\mu$.

A la distribución $\mathrm{N}(0, 1)$ se le llama *normal estándar* y a su densidad se la denota con $\phi(x)$ y a la función de distribución con $\Phi(x)$.

Al igual que lo que ocurre con otras funciones de distribución no hay una forma explícita para

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2}(\frac{t-\mu}{\sigma})^2} dt \tag{16}$$

y se recurre a una aproximación numérica para calcular el valor de la integral.

**Proposición 9.** *Sean $\mu$ y $\sigma$ dos números reales con $\sigma > 0$. Es válida la siguiente equivalencia:*

$$X \sim N(\mu, \sigma^2) \textit{ si y sólo si } Y = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

**Demostración.** Consideremos $X \sim \mathrm{N}(\mu, \sigma^2)$ y sean $F_X$ y $f_X$ su función de distribución y de densidad respectivamente, dadas por las expresiones (15) y (16).

Consideremos $Y = \frac{X-\mu}{\sigma}$ y demostremos que la función de densidad de $Y$ coincide con la de la

[3] Ver el apéndice para una definición general de parámetro de posición y de escala

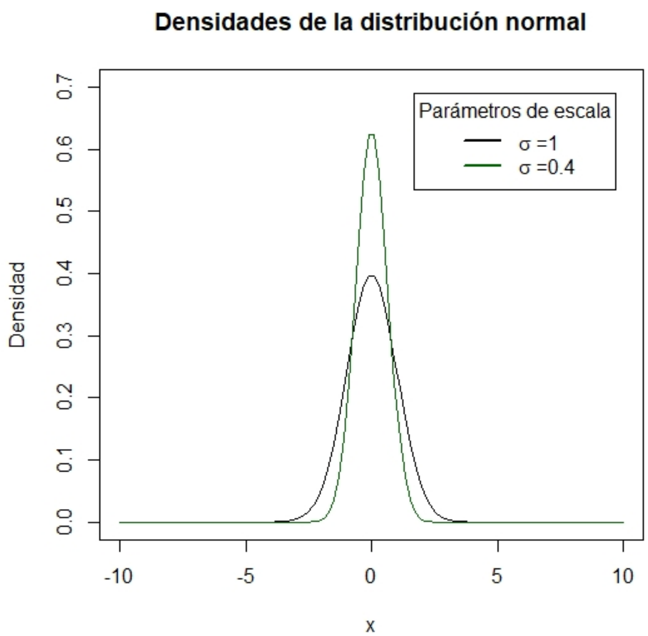


**Figura 6:** Densidades de las distribuciones $\mathrm{N}(0, \sigma^2)$ con $\sigma^2 = 1$ y 0.16

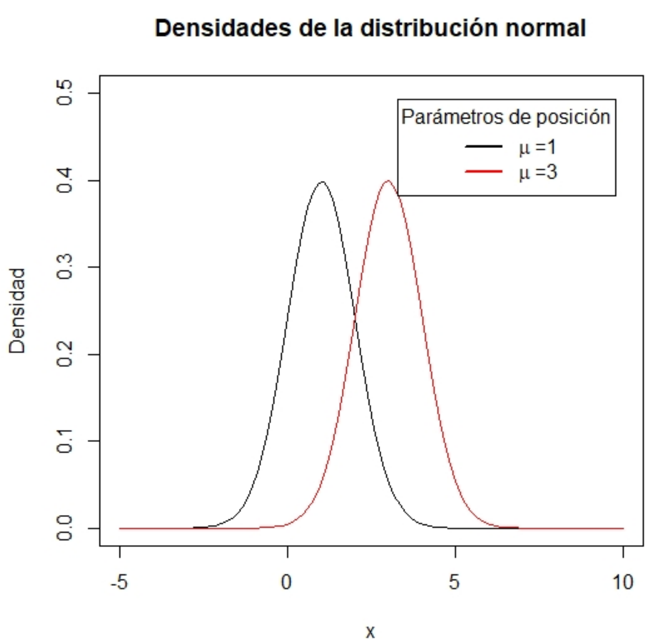


**Figura 7:** Densidades de las distribuciones $\mathrm{N}(\mu, 1)$ con $\mu = 1$ y 3

normal estándar $\phi$. Denotemos con $F_Y$ la función de distribución de $Y$ y sea $y \in \mathbb{R}$,

$$\begin{aligned} F_Y(y) &= P(\frac{X-\mu}{\sigma} \leq y) \\ &= P(X \leq \sigma y + \mu) \\ &= F_X(\sigma y + \mu). \end{aligned}$$

Por la regla de la cadena y simplificando, obtenemos:

$$\begin{aligned} \frac{d}{dy} F_Y(y) &= f_X(\sigma y + \mu)\sigma \\ &= \sigma \frac{1}{\sqrt{2\pi\sigma^2}} \mathrm{e}^{-\frac{1}{2}(\frac{(\sigma y+\mu)-\mu}{\sigma})^2} \\ &= \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{1}{2}y^2} \\ &= \phi(y). \end{aligned}$$

Así, hemos probado que $Y = \frac{X-\mu}{\sigma} \sim \mathrm{N}(0,1)$. La recíproca se deja como ejercicio. □

### 3.4.1. Esperanza y varianza

El cálculo de la esperanza y varianza de una variable aleatoria $X \sim \mathrm{N}(\mu, \sigma^2)$ requiere de contenidos que no han sido abordados hasta el momento en el plan de estudio. Sus valores son

$$\mathrm{E}(X) = \mu \text{ y } \mathrm{VAR}(X) = \sigma^2.$$

# 4. Independencia

## 4.1. Introducción

**Definición 5.** *Sea $(X, Y)$ un vector aleatorio continuo con función de densidad conjunta $f$ y sean $f_X$ y $f_Y$ las funciones de densidad de las variables aleatorias $X$ e $Y$, respectivamente. Diremos que $X$ e $Y$ son independientes si y sólo si*

$$\forall x, y : \ f(x, y) = f_X(x) f_Y(y).$$

Notar que esta definición es análoga cuando las variables aleatorias $X$ e $Y$ son discretas.

**Proposición 10.** *Sea $(X, Y)$ un vector aleatorio continuo con función de distribución conjunta $F$. Las variables aleatorias $X$ e $Y$ son independientes si y sólo si*

$$\forall x, y : \ F(x, y) = F_X(x) F_Y(y).$$

*donde $F_X$ y $F_Y$ son las funciones de distribución de $X$ e $Y$, respectivamente.*

**Demostración.** Ver apéndice.

Es importante tener en cuenta que esta proposición vale también para variables aleatorias discretas, cambiando función de densidad por función de probabilidad de masa.

**Notas.**

a) La Definición 5 y la Proposición 10 se generaliza a cualquier vector $(X_1, \ldots, X_p)$.

b) Dadas funciones $g_1, \ldots, g_p$, si $X_1, \ldots, X_p$ son variables aleatorias independientes entonces $g_1(X_1), \ldots, g_p(X_p)$ resultan variables aleatorias independientes.

## 4.2. Correlación lineal

Análogamente a como lo hicimos para las variables aleatorias discretas, se define la noción de *covarianza* y de *coeficiente de correlación* para variables aleatorias continuas

$$\begin{aligned} \operatorname{cov}(X,Y) &= \mathrm{E}\left[(X-\mathrm{E}(X))\,(Y-\mathrm{E}(Y))\right] \\ \rho &= \frac{\operatorname{cov}(X,Y)}{\operatorname{desv}(X)\operatorname{desv}(Y)}. \end{aligned}$$

Y, vale la misma relación entre las nociones de *no correlación* e *independencia*. Más explícitamente, en primer lugar mencionamos que es válida la siguiente igualdad

$$\operatorname{cov}(X,Y) = \mathrm{E}\,(XY) - \mathrm{E}(X)\mathrm{E}(Y). \tag{17}$$

y por ende concluimos la validez de la siguiente proposición.

**Proposición 11.** *Si $X$ e $Y$ son variables aleatorias independientes entonces*

$$\mathrm{E}(XY) = \mathrm{E}(X)\mathrm{E}(Y)$$

*y, por lo tanto, están no correlacionadas y $\operatorname{cov}(X,Y) = \rho = 0$.*

***Demostración***: ejercicio.

Al igual que el fenómeno que ocurre con las variables aleatorias discretas, independencia y no correlación no son equivalentes.

**Ejemplo 4.** *Contraejemplo: no correlación no implica independencia.*

Consideremos $X \sim$ uniforme$(-1,1)$ con densidad $f(x) = \frac{1}{2}I_{(-1,1)}(x)$ y sea $\theta = \mathrm{E}(X) = 0$. Si definimos $Y = I_{\{|X-\theta|<1\}} = I_{\{|X|<1\}}$, claramente $X$ e $Y$ no son independientes pero sin embargo son no correlacionadas como lo muestra la siguiente igualdad

$$E(XY) = \int_{-\infty}^{\infty} x I_{\{|X-\theta|<1\}} f(x-\theta)dx = \int_{-1}^{1}(t+\theta)f(t)dt = \theta\int_{-1}^{1} f(t)dt = E(X)E(Y) = 0.$$

El contraejemplo es más general, vale también si $f$ se reemplaza por cualquier función simétrica respecto de cero tal que la densidad de $X$ sea $f(x-\theta)$ y por lo tanto $\mathrm{E}(X) = \theta$ para $\theta$ arbitrario. □

A la siguiente proposición, que ya fue enunciada como una observación en la unidad anterior, vale tanto para variables aleatorias continuas como discretas.

**Proposición 12.** *Dadas dos variables aleatorias $X$ e $Y$, el coeficiente de correlación $\rho$ satisface*

*a)* $-1 \leq \rho \leq 1$

*b)* $|\rho| = 1$ *si y sólo si existen constantes $a$ y $b$ tales que* $P(Y = aX + b) = 1$. *Si* $\rho = 1$ *entonces* $a > 0$ *y si* $\rho = -1$, $a < 0$.

Observar que el inciso $b$ nos dice que el coeficiente de correlación (y también la covarianza) mide el grado de asociación **lineal**: $P(Y = aX + b) = 1$ significa que con probabilidad uno las variables $X$ e $Y$ yacen sobre una recta.

**Ejemplo 5.** *El coeficiente de correlación sólo mide asociaciones lineales.*

Consideremos[4] dos variables aleatorias independientes $X \sim \text{uniforme}(-1, 1)$ y $Z \sim \text{uniforme}(0, 1/10)$ y, definamos la variable aleatoria $Y = X^2 + Z$. Como se argumenta en el Ejemplo 6 del Apéndice $X = x$ e $Y$ no son independientes.

Sin embargo como $X \sim \text{uniforme}(-1, 1)$, $\mathrm{E}(X) = 0 = \mathrm{E}(X^3)$ y dado que $X$ y $Z$ son independientes $\mathrm{E}(XZ) = \mathrm{E}(X)\mathrm{E}(Z) = 0$ y, por lo tanto

$$
\begin{aligned}
\operatorname{cov}(X, Y) &= \mathrm{E}\left(X\left(X^2 + Z\right)\right) - \mathrm{E}(X)\left(\mathrm{E}\left(X^2 + Z\right)\right) \\
&= \mathrm{E}(X^3) + \mathrm{E}(XZ) - 0\mathrm{E}\left(X^2 + Z\right) \\
&= 0 + \mathrm{E}(X)\mathrm{E}(Z) = 0\mathrm{E}(Z) = 0.
\end{aligned}
$$

De este modo $\rho = 0$ y por ende *las variables están no correlacionadas linealmente aunque sean dependientes.* □

Este último ejemplo muestra que cada vez que hablamos de correlación nos referimos siempre a correlación *lineal.*

# 5. Apéndice

## 5.1. Demostraciones pendientes

**Demostración de la Proposición 4.**[5] Considerar que para cualquier $h > 0$ y $x_0 - h \leq x \leq x_0 + h$ se cumple que

[4] Este ejemplo es dado en Casella, G. and Berger, R. (2002). "Statistical Inference. Duxbury (Ejemplo 4.5.9 de pág. 174)

[5] La prueba está basada en el texto de Yohai (2006, p. 52 y sgts.)

$$m_h \leq f(x) \leq M_h,$$

donde $M_h = \max\{f(x) : x \in [x_0 - h, x_0 + h]\}$ y $m_h = \min\{f(x) : x \in [x_0 - h, < x_0 + h]\}$.

Luego

$$\int_{x_0-h}^{x_0+h} m_h \leq \int_{x_0-h}^{x_0+h} f(t)dt \leq \int_{x_0-h}^{x_0+h} M_h$$

y así

$$2hm_h \leq \int_{x_0-h}^{x_0+h} f(t)dt \leq 2hM_h.$$

Si dividimos la última desigualdad por $2h$ obtenemos:

$$m_h \leq \frac{1}{2h}\int_{x_0-h}^{x_0+h} f_X(t)dt \leq M_h,$$

y, por la Proposición 3, la anterior desigualdad se puede reescribir como

$$m_h \leq \frac{P(x_0 - h \leq X \leq x_0 + h)}{2h} \leq M_h.$$

Tomando límite y teniendo en cuenta que por la continuidad de $f$ en $x_0$ se satisface que

$$f(x_0) = \lim_{h\to 0} M_h = \lim_{h\to 0} m_h$$

concluimos que

$$f(x_0) \leq \lim_{h\to 0} \frac{P(x_0 - h \leq X \leq x_0 + h)}{2h} \leq f(x_0)$$

estableciendo la validez de la tesis de la proposición. $\square$

**Demostración de la Proposición 10.**

Si asumimos que las variables aleatorias son independientes entonces $f = f_X f_Y$. Luego

$$\begin{aligned}
F(x,y) &= \int_{-\infty}^{y}\int_{-\infty}^{x} f(u,v)dudv \\
&= \int_{-\infty}^{y}\int_{-\infty}^{x} f_X(u)f_Y(v)dudv \\
&= \int_{-\infty}^{y} f_Y(v)dv \left(\int_{-\infty}^{x} f_X(u)du\right) \\
&= F_X(x)F_Y(y)
\end{aligned}$$

La recíproca es consecuencia de la igualdad (3). $\square$

## 5.2. Familias de posición-escala

**Definición 6.** *Sea $f_0$ una función de densidad. La familia de densidades de posición-escala asociada a $f_0$ es la familia de funciones de la forma*

$$f(x;\mu,\sigma) = \frac{1}{\sigma} f_0\left(\frac{x-\mu}{\sigma}\right)$$

*donde $\sigma > 0$ y $-\infty < \mu < \infty$ y tal que cada una de ellas resulte una función de densidad.*

Un ejemplo de familia de densidades de posición y escala es la familia $\{f(\cdot;\mu,\sigma)\}_{(\mu,\sigma)}$ donde $f(\cdot;\mu,\sigma)$ es la densidad de una variable aleatoria $X \sim \mathrm{N}(\mu,\sigma^2)$. Aquí $f_0$ es la densidad de una $Z \sim \mathrm{N}(0,1)$.

**Proposición 13.** *En el contexto de la definición anterior son válidas las siguientes afirmaciones:*

*a)* $f(x;0,1) = f_0(x)$

*b) Si $X$ es una variable aleatoria con función de densidad $f(\cdot;\mu,\sigma)$, entonces $\frac{X-\mu}{\sigma}$ tiene función de densidad $f(\cdot;0,1) = f_0$.*

*c) Si $Z$ es una variable aleatoria con función de densidad $f(\cdot;0,1)$ entonces $\sigma Z + \mu$ tiene función de densidad $f(\cdot;\mu,\sigma)$.*

**Demostración.** Ejercicio.

En una familia de posición-escala a $(\mu,\sigma)$ se le denomina parámetro de posición-escala.

A continuación introducimos las familias de escala y de posición.

**Definición 7.** *Sea $f_0$ una función de densidad. La familia de densidades de escala asociada a $f_0$ es la familia de funciones de la forma*

$$f(x;\sigma) = \frac{1}{\sigma} f_0\left(\frac{x}{\sigma}\right)$$

*donde $\sigma > 0$ y tal que cada una de ellas resulte una función de densidad.*

Un ejemplo es la familia de densidades correspondientes a la distribución gamma con parámetro de forma fijo.

**Definición 8.** *Sea $f_0$ una función de densidad. La familia de densidades de posición asociada a $f_0$ es la familia de funciones de la forma*

$$f(x;\mu) = f_0\left(x-\mu\right)$$

*donde $\mu \in \mathbb{R}$ y tal que cada una de ellas resulte una función de densidad.*

Observar que la familia de distribuciones $\mathrm{N}(\mu,1)$ dan origen a una familia de densidades de posición y la familia de distribuciones $\mathrm{N}\left(0,\sigma^2\right)$ a una familia de densidades de escala.

## 5.3. Distribuciones condicionales

Implícitamente en la definición que sigue si hablamos de función de densidad o función de probabilidad de masa nos referimos a vectores -y variables- continuos o discretos respectivamente.

**Definición 9.** *Sea $(X,Y)$ un vector aleatorio con $f(x,y)$ y $f_X$ y $f_Y$ la funciones de probabilidad de masa o densidad conjunta y las funciones de probabilidad de masa o densidades marginales. Fijemos $x$, llamaremos función de probabilidad de masa o función de densidad condicional de $Y$ dado $X = x$ a*

$$\forall x, y: \; f(y|x) = \frac{f(x,y)}{f_X(x)} \tag{18}$$

*provisto que $f_X(x) > 0$.*

Supongamos que $(X,Y)$ un vector aleatorio discreto. Notar que fijado $x$ se cumple que:

a) $\forall y : f(y|x) \geq 0$

b) y, además

$$\sum_y f(y|x) = \frac{1}{f_X(x)} \sum_y f(x,y) = \frac{1}{f_X(x)} f_X(x) = 1,$$

con lo cual $f(y|x)$ es una función de probabilidad de masa. Reemplazando sumatorias por integrales concluimos también que, para el caso continuo, $f(y|x)$ es una densidad para cada $x$ fijo.

La función $f(y|x)$ permite definir una función de distribución de manera obvia, la que se denomina *función de distribución condicional* y denotaremos con $F(y|x)$. Llamaremos *distribución condicional de $Y$ dado $X = x$* a la variable aleatoria con función de probabilidad de masa o densidad condicional $f(y|x)$ o función de distribución condicional $F(y|x)$.

**Problema 6.** *Dada la función de probabilidad de masa*

$$f(0,10) = f(0,20) = \tfrac{2}{18}, \; f(1,10) = f(1,30) = \tfrac{3}{18}$$
$$f(1,20) = \tfrac{4}{18}, f(2,30) = \tfrac{4}{18}$$

*del vector $(X,Y)$ hallar la distribución condicional de $Y$ dado $X = x$.*

**Solución.** Por definición la función de probabilidad marginal de $X$ es

$$\begin{aligned} f_X(0) &= f(0,10) + f(0,20) = 4/18 \\ f_X(1) &= f(1,10) + f(1,20) + f(1,30) = 10/18 \\ f_X(2) &= f(2,30) = 4/18. \end{aligned}$$

Para $x = 0$, la función $f(0, y) > 0$ para $y = 10$ o $y = 20$. Luego

$$f(10 \mid 0) = \frac{f(0,10)}{f_X(0)} = (2/18)/(4/18) = 1/2.$$
$$f(20|0) = \frac{f(0,20)}{f_X(0)} = 1/2$$

De este modo, conociendo que $X = 0$, la distribución de $Y$ asigna probabilidad 1/2 a los puntos $y = 20$ e $y = 10$. De modo análogo se pueden hallar $f(y|1)$ y $f(y|2)$. □

La siguiente afirmación es de demostración inmediata pero es importante de ser enfatizada.

**Proposición 14.** *$X$ e $Y$ son variables aleatorias independientes si y sólo si la función de probabilidad de masa o la de densidad condicional de $Y$ dado $X = x$ satisface*

$$\forall y : \; f(y|x) = f_Y(y),$$

*independientemente del valor $x$.*

En otros términos, cuando hay independencia la distribución condicional coincide con la marginal.

**Ejemplo 6.** *Continuación del Ejemplo 4.2*

Mostremos que las variables $X$ e $Y$ introducidas en el Ejemplo 4.2 no son independientes. Se puede demostrar que la densidad conjunta de $(X, Y)$ es

$$f(x, y) = \begin{cases} 5 & \text{si } -1 < x < 1, \; x^2 < y < x^2 + 1/10 \\ 0 & \text{en caso contrario} \end{cases} \tag{19}$$

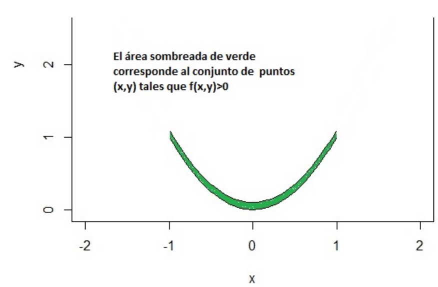


**Figura 8:** El área sombreada en verde corresponde al conjunto del plano en el que la densidad conjunta es positiva

A partir de la expresión de $f(x, y)$, se puede probar que la distribución condicional de $Y = x^2 + Z$ dado $X = x$ es uniforme$(x^2, x^2 + \frac{1}{10})$ con lo cual, por la Proposición 14, las variables no son independientes. □