Probabilidad y Estadística (93.24) Mezcla de variables aleatorias, funciones de variable aleatoria y distribuciones conjuntas

Índice

1. Repaso de algunos conceptos

Función de variable aleatoria.

Sea X una variable aleatoria con función de distribución (acumulada) $\mathrm{F}_X : \mathbb{R} \to [0,1]$ y $g : \mathbb{R} \to \mathbb{R}$ una función dada. Si se define una nueva variable aleatoria $Y = g(X)$, la función de distribución de Y está dada por

$$\mathrm{F}_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(g(X) \leq y). \tag{1}$$

Generación de variables aleatorias

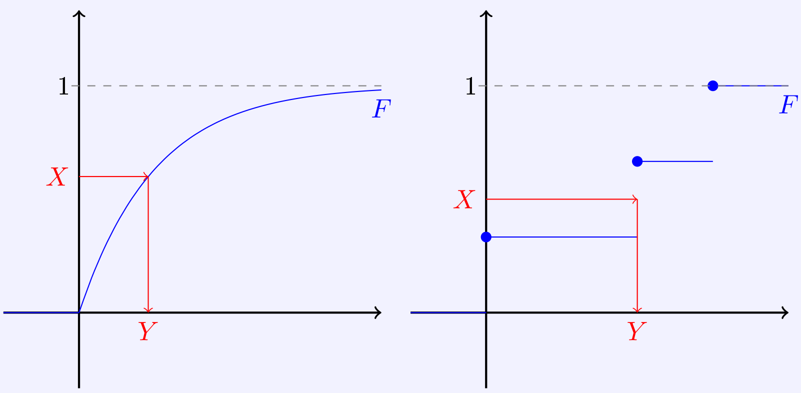
Un caso particular de interés es el siguiente. Sea $X \sim U(0,1)$ y F una función de distribución de probabilidad (acumulada) cualquiera. Definamos g de la siguiente manera:

$$g(x) = \min\{y \in \mathbb{R} : \ F(y) \geq x\}. \tag{2}$$

Recordemos que F es no-decreciente y continua por derecha, por lo que esta definición de g es correcta. En particular, si existe un intervalo $[a,b] \subset \mathbb{R}$ en el cual F sea estrictamente creciente, entonces $g(x) = F^{-1}(x)$ para todo $x \in F([a,b])$. Luego, si definimos una nueva variable aleatoria $Y = g(X)$, es fácil demostrar que (ver los gráficos más abajo)

$$\mathrm{F}_Y(y) = \mathrm{P}(Y \leq y) = \mathrm{P}(g(X) \leq y) = F(y). \tag{3}$$

Es decir, la distribución de probabilidad de Y viene dada por F. **Estas ecuaciones proveen una forma de simular una variable aleatoria con una distribución cualquiera a partir de una variable aleatoria con distribución uniforme.**



Transformaciones afines

Sea X una variable aleatoria y $a, b \in \mathbb{R}$. Si definimos $Y = aX + b$, es fácil demostrar que

$$\mathrm{E}[Y] = a\mathrm{E}[X] + b, \tag{4}$$
$$\mathrm{Var}[Y] = a^2\mathrm{Var}[X], \tag{5}$$

siempre y cuando los momentos en los lados derechos de las ecuaciones anteriores estén definidos.
Si X es una variable aleatoria normal y $a \neq 0$, entonces Y también es una variable aleatoria normal. En particular,

$$Z \sim \mathcal{N}(0,1) \Rightarrow \qquad X = \sigma_X Z + \mu_X \sim \mathcal{N}(\mu_X, \sigma_X), \tag{6}$$
$$X \sim \mathcal{N}(\mu_X, \sigma_X) \Rightarrow Z = \frac{X - \mu_X}{\sigma_X} \sim \mathcal{N}(0,1) \qquad \text{(estandarización)}. \tag{7}$$

Variables bidimensionales discretas

Sean X e Y dos variables aleatorias discretas. Su comportamiento puede ser descripto por la **función de probabilidad conjunta**

$$\mathrm{p}_{X,Y}(x,y) = \mathrm{P}(X = x, Y = y) \qquad x \in \mathcal{R}_X,\, y \in \mathcal{R}_Y. \tag{8}$$

La funciones de probabilidad **marginales** pueden ser obtenidas como

$$\mathrm{p}_X(x) = \mathrm{P}(X = x) = \sum_{y \in \mathcal{R}_Y} \mathrm{p}_{X,Y}(x,y) \qquad x \in \mathcal{R}_X, \tag{9}$$
$$\mathrm{p}_Y(y) = \mathrm{P}(Y = y) = \sum_{x \in \mathcal{R}_X} \mathrm{p}_{X,Y}(x,y) \qquad y \in \mathcal{R}_Y. \tag{10}$$

Las variables aleatorias son **independientes** sii

$$\mathrm{p}_{X,Y}(x,y) = \mathrm{p}_X(x)\mathrm{p}_Y(y) \qquad x \in \mathcal{R}_X,\, y \in \mathcal{R}_Y. \tag{11}$$

Es fácil ver que, en este caso,

$$\mathrm{P}((X,Y) \in [a,b] \times \in [c,d]) = \mathrm{P}(X \in [a,b], Y \in [c,d]) = \mathrm{P}(X \in [a,b]) \cdot \mathrm{P}(Y \in [c,d]), \tag{12}$$

para todo par de intervalos $[a,b], [c,d] \subset \mathbb{R}$.
El valor esperado de una función h de ambas variables aleatorias ($h : \mathbb{R}^2 \to \mathbb{R}$) viene dado por

$$\mathrm{E}[h(X,Y)] = \sum_{(x,y) \in \mathcal{R}_X \times \mathcal{R}_Y} h(x,y)\mathrm{p}_{X,Y}(x,y) = \sum_{x \in \mathcal{R}_X} \sum_{y \in \mathcal{R}_Y} h(x,y)\mathrm{p}_{X,Y}(x,y). \tag{13}$$

Sean dos funciones $k, m : \mathbb{R} \to \mathbb{R}$. Luego, si X e Y son variables aleatorias independientes,

$$\mathrm{E}[k(X)m(Y)] = \sum_{x \in \mathcal{R}_X} \sum_{y \in \mathcal{R}_Y} k(X)m(Y)\mathrm{p}_X(x)\mathrm{p}_Y(y) = \mathrm{E}[k(X)]\,\mathrm{E}[m(Y)]. \tag{14}$$

Variables bidimensionales continuas

El comportamiento de dos variables aleatorias continuas X e Y suele se descripto por una **densidad de probabilidad conjunta** $\mathrm{f}_{X,Y} : \mathbb{R}^2 \to \mathbb{R}^{\geq 0}$. A partir de ésta, pueden calcularse todas las probabilidades relacionadas a ambas variables como

$$\mathrm{P}\left((X,Y) \in B\right) = \iint\limits_{(x,y)\in B} \mathrm{f}_{X,Y}(x,y)\, dxdy \qquad B \subseteq \mathbb{R}^2. \tag{15}$$

Allí donde las derivadas tienen sentido, la densidad conjunta puede escribirse como

$$\mathrm{f}_{X,Y}(x,y) = \frac{\partial^2}{\partial x \partial y} \mathrm{P}\left(X \leq x, Y \leq y\right). \tag{16}$$

La densidades de probabilidad **marginales** pueden ser obtenidas como

$$\mathrm{f}_X(x) = \int\limits_{\mathbb{R}} \mathrm{f}_{X,Y}(x,y)\, dy, \qquad \mathrm{f}_Y(y) = \int\limits_{\mathbb{R}} \mathrm{f}_{X,Y}(x,y)\, dx. \tag{17}$$

Las variables aleatorias son **independientes** sii

$$\mathrm{f}_{X,Y}(x,y) = \mathrm{f}_X(x)\mathrm{f}_Y(y) \qquad (x,y) \in \mathbb{R}^2. \tag{18}$$

Es fácil ver que, en este caso,

$$\mathrm{P}\left((X,Y) \in [a,b] \times \in [c,d]\right) = \mathrm{P}\left(X \in [a,b], Y \in [c,d]\right) = \mathrm{P}\left(X \in [a,b]\right) \cdot \mathrm{P}\left(Y \in [c,d]\right), \tag{19}$$

para todo par de intervalos $[a,b], [c,d] \subset \mathbb{R}$.
El valor esperado de una función h de ambas variables aleatorias ($h : \mathbb{R}^2 \to \mathbb{R}$) viene dado por

$$\mathrm{E}\left[h(X,Y)\right] = \iint\limits_{\mathbb{R}^2} h(x,y)\mathrm{f}_{X,Y}(x,y)\, dxdy. \tag{20}$$

Sean dos funciones $k, m : \mathbb{R} \to \mathbb{R}$. Luego, si X e Y son variables aleatorias independientes,

$$\mathrm{E}\left[k(X)m(Y)\right] = \iint\limits_{\mathbb{R}^2} k(x)m(y)\mathrm{f}_X(x)\mathrm{f}_Y(y)\, dxdy = \mathrm{E}\left[k(X)\right]\mathrm{E}\left[m(Y)\right]. \tag{21}$$

Covarianza y correlación

La **covarianza** de dos variables aleatorias está definida por

$$\operatorname{Cov}[X,Y] = \operatorname{E}[(X-\operatorname{E}[X])(Y-\operatorname{E}[Y])] = \operatorname{E}[XY] - \operatorname{E}[X]\operatorname{E}[Y]. \tag{22}$$

Es fácil ver que, si X e Y son variables aleatorias independientes, entonces $\operatorname{Cov}[X,Y]=0$. Sin embargo, el hecho que $\operatorname{Cov}[X,Y]=0$ no implica que X e Y sean independientes.
El **coeficiente de correlación** de dos variables aleatorias está definido como

$$\rho(X,Y) = \frac{\operatorname{Cov}[X,Y]}{\sqrt{\operatorname{Var}[X]\operatorname{Var}[Y]}}. \tag{23}$$

Se puede demostrar que $\rho(X,Y)=\pm 1$ sii $\exists a,b\in\mathbb{R}$ tales que

$$\operatorname{P}(Y = aX+b) = 1. \tag{24}$$

Es decir, $\rho(X,Y)=\pm 1$ sii con probabilidad 1 existe una relación lineal entre ambas variables aleatorias. Más aún, $\operatorname{sign}(a)=\operatorname{sign}(\rho(X,Y))$.

Suma de dos variables aleatorias

Sean X e Y dos variables aleatorias. Luego,

$$\begin{aligned} \operatorname{E}[X+Y] &= \operatorname{E}[X]+\operatorname{E}[Y] \\ \operatorname{E}\left[(X+Y)^2\right] &= \operatorname{E}\left[X^2\right] + 2\operatorname{E}[XY] + \operatorname{E}\left[Y^2\right], \end{aligned}$$

siempre que los momentos en los lados derechos estén definidos. A partir de estas ecuaciones, es fácil ver que

$$\operatorname{Var}[X+Y] = \operatorname{Var}[X] + 2\operatorname{Cov}[X,Y] + \operatorname{Var}[Y]. \tag{25}$$

En el caso particular de variables independientes, tenemos

$$\operatorname{Var}[X+Y] = \operatorname{Var}[X] + \operatorname{Var}[Y]. \tag{26}$$

Si X e Y son variables aleatorias discretas, entonces

$$\operatorname{P}(X+Y=z) = \sum_{\{(x,y)\in\mathcal{R}_X\times\mathcal{R}_Y : x+y=z\}} \operatorname{p}_{X,Y}(x,y). \tag{27}$$

En el caso particular en que X e Y independientes,

$$\operatorname{P}(X+Y=z) = \sum_{x\in\mathcal{R}_X} \operatorname{p}_X(x)\operatorname{p}_Y(z-x) = \sum_{y\in\mathcal{R}_Y} \operatorname{p}_Y(y)\operatorname{p}_X(z-y). \tag{28}$$

De manera análoga, si X e Y son variables aleatorias continuas e independientes, tenemos

$$\operatorname{f}_{X+Y}(z) = \int_{\mathbb{R}} \operatorname{f}_X(x)\operatorname{f}_Y(z-x)\,dx = \int_{\mathbb{R}} \operatorname{f}_Y(y)\operatorname{f}_X(z-y)\,dy. \tag{29}$$

Variables bidimensionales discretas - distribuciones condicionales

Sean X e Y dos variables aleatorias discretas. La **función de masa de probabilidad condicional** de Y dado X está dada por

$$\mathrm{p}_{X|Y}(x|y) = \mathrm{P}\big(X = x|\, Y = y\big) = \frac{\mathrm{p}_{X,Y}(x,y)}{\mathrm{p}_Y(y)} \qquad \forall x \in \mathcal{R}_X,\ \forall y \in \mathcal{R}_Y \text{ tal que } \mathrm{p}_Y(y) > 0. \tag{30}$$

No importa la definición para y tales que $\mathrm{p}_Y(y) = 0$, pero, por conveniencia, podemos definir $\mathrm{p}_{X|Y} = 0$ en esos casos.
Se sigue inmediatamente que las variables aleatorias son **independientes** sii

$$\mathrm{p}_{X|Y}(x|y) = \mathrm{p}_X(x) \qquad \forall y \in \mathcal{R}_Y \text{ tal que } \mathrm{p}_Y(y) > 0. \tag{31}$$

Es simple demostrar que $\mathrm{p}_{X|Y}$ es verdaderamente una función de masa de probabilidad. De esta forma, podemos considerar $X|Y$ como una variable aleatoria. En particular, podemos calcular **valores esperados condicionales**

$$\mathrm{E}\left[h(X)|Y = y\right] = \sum_{x\in\mathcal{R}_X} h(x)\mathrm{p}_{X|Y}(x|y). \tag{32}$$

Este valor esperado depende del valor de Y. Es decir que es una función de Y y, por tanto, ¡es una variable aleatoria! De hecho, si no fijamos el valor $Y = y$, tenemos:

$$\mathrm{E}\left[h(X)|Y\right] = \sum_{x\in\mathcal{R}_X} h(x)\mathrm{p}_{X|Y}(x|Y). \tag{33}$$

El valor esperado de esta variable aleatoria es:

$$\begin{aligned}\mathrm{E}\left[\mathrm{E}\left[h(X)|Y\right]\right] &= \sum_{y\in\mathcal{R}_Y} \mathrm{E}\left[h(X)|Y = y\right]\mathrm{p}_Y(y) = \sum_{y\in\mathcal{R}_Y}\sum_{x\in\mathcal{R}_X} h(x)\mathrm{p}_{X|Y}(x|y)\mathrm{p}_Y(y)\\ &= \sum_{y\in\mathcal{R}_Y}\sum_{x\in\mathcal{R}_X} h(x)\mathrm{p}_{X,Y}(x,y)\\ &= \mathrm{E}\left[h(X)\right].\end{aligned} \tag{34}$$

En este sentido, podemos definir

$$\mu_{X|Y} = \mathrm{E}\left[X|Y\right], \qquad \mathrm{Var}\left[X|Y\right] = \sigma^2_{X|Y} = \mathrm{E}\left[(X-\mu_{X|Y})^2|Y\right] = \mathrm{E}\left[X^2|Y\right] - \mu^2_{X|Y}. \tag{35}$$

Más aún,

$$\mu_X = \mathrm{E}\left[\mu_{X|Y}\right], \qquad \sigma^2_X = \mathrm{E}\left[\sigma^2_{X|Y}\right] + \mathrm{Var}\left[\mu_{X|Y}\right]. \tag{36}$$

Variables bidimensionales continuas - distribuciones condicionales

Sean X e Y dos variables aleatorias continuas. La **densidad de probabilidad condicional** de Y dado X está dada por

$$\mathrm{f}_{X|Y}(x|y) = \frac{\mathrm{f}_{X,Y}(x,y)}{\mathrm{f}_Y(y)} \qquad \forall x \in \mathbb{R}, \forall y \in \mathbb{R} \text{ tal que } \mathrm{f}_Y(y) > 0. \tag{37}$$

No importa la definición para y tales que $\mathrm{f}_Y(y) = 0$, pero, por conveniencia, podemos definir $\mathrm{f}_{X|Y} = 0$ en esos casos.
Se sigue inmediatamente que las variables aleatorias son **independientes** sii

$$\mathrm{f}_{X|Y}(x|y) = \mathrm{f}_X(x) \qquad \forall y \in \mathbb{R} \text{ tal que } \mathrm{f}_Y(y) > 0. \tag{38}$$

Es simple demostrar que $\mathrm{f}_{X|Y}$ es verdaderamente una densidad de probabilidad. De esta forma, podemos considerar $X|Y$ como una variable aleatoria. En particular, podemos calcular **valores esperados condicionales**

$$\mathrm{E}\left[h(X)|Y=y\right] = \int_{-\infty}^{+\infty} h(x)\mathrm{f}_{X|Y}(x|y)dx. \tag{39}$$

Este valor esperado depende del valor de Y. Es decir que es una función de Y y, por tanto, ¡es una variable aleatoria! De hecho, si no fijamos el valor $Y = y$, tenemos:

$$\mathrm{E}\left[h(X)|Y\right] = \int_{-\infty}^{+\infty} h(x)\mathrm{f}_{X|Y}(x|Y)dx. \tag{40}$$

El valor esperado de esta variable aleatoria es:

$$\begin{aligned}\mathrm{E}\left[\mathrm{E}\left[h(X)|Y\right]\right] &= \int_{-\infty}^{+\infty} \mathrm{E}\left[h(X)|Y=y\right]\mathrm{f}_Y(y)dy = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x)\mathrm{f}_{X|Y}(x|y)\mathrm{f}_Y(y)dxdy \\ &= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} h(x)\mathrm{f}_{X,Y}(x,y)dxdy \\ &= \mathrm{E}\left[h(X)\right]. \end{aligned} \tag{41}$$

En este sentido, podemos definir

$$\mu_{X|Y} = \mathrm{E}\left[X|Y\right], \qquad \mathrm{Var}\left[X|Y\right] = \sigma^2_{X|Y} = \mathrm{E}\left[(X-\mu_{X|Y})^2|Y\right] = \mathrm{E}\left[X^2|Y\right] - \mu^2_{X|Y}. \tag{42}$$

Más aún,

$$\mu_X = \mathrm{E}\left[\mu_{X|Y}\right], \qquad \sigma^2_X = \mathrm{E}\left[\sigma^2_{X|Y}\right] + \mathrm{Var}\left[\mu_{X|Y}\right]. \tag{43}$$

Variables bidimensionales mixtas: continua condicionada a discreta

Sea Y una variable aleatoria discreta y X una variable aleatoria continua descripta por la función de distribución acumulada condicional

$$\mathrm{F}_{X|Y}(x|y) = \mathrm{P}\big(X \leq x|\, Y = y\big) \qquad \forall x \in \mathbb{R},\, y \in \mathcal{R}_Y. \tag{44}$$

La densidad de probabilidad condicional $\mathrm{f}_{X|Y}$ y viene dada por

$$\mathrm{f}_{X|Y}(x|y) = \frac{d}{dx}\mathrm{F}_{X|Y}(x|y). \tag{45}$$

La función de distribución acumulada de X es, simplemente,

$$\mathrm{F}_X(x) = \mathrm{P}\left(X \leq x\right) = \sum_{y\in\mathcal{R}_Y} \mathrm{P}\big(X \leq x|\, Y = y\big)\mathrm{P}\left(Y = y\right) = \sum_{y\in\mathcal{R}_Y} \mathrm{F}_{X|Y}(x|y)\mathrm{p}_Y(y). \tag{46}$$

Análogamente,

$$\mathrm{f}_X(x) = \sum_{y\in\mathcal{R}_Y} \mathrm{f}_{X|Y}(x|y)\mathrm{p}_Y(y). \tag{47}$$

De esta forma, tenemos que los valores esperados son

$$\begin{aligned}
\mathrm{E}\left[h(X)\right] &= \int_{\mathbb{R}} h(x)\mathrm{f}_X(x)dx = \sum_{y\in\mathcal{R}_Y}\left[\int_{\mathbb{R}} h(x)\mathrm{f}_{X|Y}(x|y)dx\right]\mathrm{p}_Y(y) && (48)\\
&= \sum_{y\in\mathcal{R}_Y} \mathrm{E}\left[h(X)|Y = y\right]\mathrm{p}_Y(y) && (49)\\
&= \mathrm{E}\left[\mathrm{E}\left[h(X)|Y\right]\right]. && (50)
\end{aligned}$$

A veces, este caso se denomina de *mezcla de variables aleatorias.*

Variables bidimensionales mixtas: discreta condicionada a continua

Sea Y una variable aleatoria continua y X una variable aleatoria discreta descripta por la función de masa de probabilidad condicional

$$\mathrm{p}_{X|Y}(x|y) = \mathrm{P}\big(X = x|\, Y = y\big) \qquad \forall x \in \mathcal{R}_Y,\, y \in \mathbb{R}. \tag{51}$$

La función de masa de probabilidad marginal de X es, simplemente,

$$\mathrm{p}_X(x) = \mathrm{P}\left(X = x\right) = \int_{\mathbb{R}} \mathrm{p}_{X|Y}(x|y)\mathrm{f}_Y(y)dy. \tag{52}$$

De esta forma, tenemos que los valores esperados son

$$\begin{aligned}
\mathrm{E}\left[h(X)\right] &= \sum_{x\in\mathcal{R}_X} h(x)\mathrm{p}_X(x) = \int_{\mathbb{R}}\left[\sum_{x\in\mathcal{R}_X} h(x)\mathrm{p}_{X|Y}(x|y)\right]\mathrm{f}_Y(y)dy && (53)\\
&= \int_{\mathbb{R}} \mathrm{E}\left[h(X)|Y = y\right]\mathrm{f}_Y(y)dy && (54)\\
&= \mathrm{E}\left[\mathrm{E}\left[h(X)|Y\right]\right]. && (55)
\end{aligned}$$

Ideas sobre integrales dobles

1. Concepto físico 1: El techo de un edificio está descripto por $w : \mathbb{R}^2 \to \mathbb{R}^+$ (campo escalar no-negativo). La base del edificio es el conjunto $A \subseteq \mathbb{R}^2$. Luego

$$\text{Volumen debajo del techo } w(x, y) = \iint_A w(x, y)dxdy. \tag{56}$$

Ver Fig. 1.

2. Concepto físico 2: Una placa cubre la superficie descripta por $A \subseteq \mathbb{R}^2$. Si $w : \mathbb{R}^2 \to \mathbb{R}^+$ (campo escalar no-negativo) representa la densidad superficial de la placa, luego

$$\text{Masa de la placa} = m = \iint_A w(x, y)dxdy, \tag{57}$$

$$\text{Abscisa del baricentro} = x_g = \frac{1}{m}\iint_A xw(x, y)dxdy, \tag{58}$$

$$\text{Ordenada del baricentro} = y_g = \frac{1}{m}\iint_A yw(x, y)dxdy. \tag{59}$$

3. <u>Suma de Riemann</u>: Una forma de entender la integral doble sobre un rectángulo $A = [a, b] \times [c, d]$ es la siguiente. Para $n \in \mathbb{N}$, definimos $\Delta x = (b - a)/n$ y $\Delta y = (d - c)/n$. Sea $w(x, y)$ una función acotada en A. Luego,

$$\iint_{[a,b]\times[c,d]} w(x, y)dxdy = \lim_{n\to\infty} \sum_{i,k=0}^{n-1} w(\tilde{x}_{i,k}, \tilde{y}_{i,k})\Delta x\Delta y, \tag{60}$$

sii el límite del lado derecho existe para cualquier elección de $(\tilde{x}_{i,k}, \tilde{y}_{i,k}) \in [a+i\Delta x, a+(i+1)\Delta x] \times [c + k\Delta y, c + (k + 1)\Delta y]$.

4. <u>Teorema de Fubini</u>: Sea $A = [a, b] \times [c, d]$ y Sea $w(x, y)$ una función continua en A. Luego,

$$\iint_{[a,b]\times[c,d]} w(x, y)dxdy = \int_{[a,b]} \left[\int_{[c,d]} w(x, y)dy\right] dx = \int_{[c,d]} \left[\int_{[a,b]} w(x, y)dx\right] dy. \tag{61}$$

5. Sean $A = \{(x, y) \in \mathbb{R}^2 : \phi_l(x) \leq y \leq \phi_u(x),\ x \in [a, b]\}$ (ver Fig. 2) y $w(x, y)$ una función continua en A. Luego,

$$\iint_A w(x, y)dxdy = \int_{[a,b]} \left[\int_{\phi_l(x)}^{\phi_u(x)} w(x, y)dy\right] dx. \tag{62}$$

6. Sean $A = \{(x, y) \in \mathbb{R}^2 : \psi_l(x) \leq x \leq \psi_u(x),\ y \in [a, b]\}$ y $w(x, y)$ una función continua en A. Luego,

$$\iint_A w(x, y)dxdy = \int_{[c,d]} \left[\int_{\psi_l(y)}^{\psi_u(y)} w(x, y)dx\right] dy. \tag{63}$$

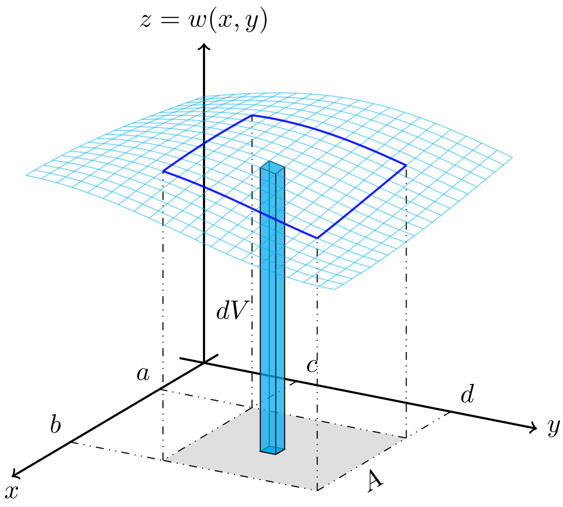


Figura 1: Adaptado de `https://tikz.net/differential-of-volume-rectangular-coordinates/`, por Efraín Soto Apolinar.

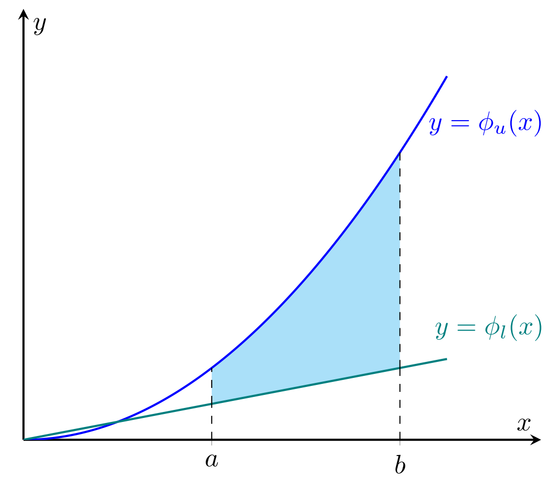


Figura 2: Dominio x-simple.

2. Guía de ejercicios

1. La variable aleatoria discreta X tiene un recorrido y función de probabilidad como indica la siguiente tabla:

x	-2	-1	1	2
$\mathrm{p}_X(x)$	0.1	0.3	0.2	0.4

Obtener la función de probabilidad, el valor esperado y varianza de Y si:
a) $Y = 3X$ b) $Y = X^2$ c) $Y = 2\,X^3$.

2. La variable aleatoria Z tiene distribución normal standard.
   a) Obtener la distribución de probabilidades de la variable $X = m + sZ$ donde m y s son constantes reales y $s > 0$.
   b) Obtener la función densidad de probabilidad de $X = Z^2$.
   Sugerencia: $\mathrm{P}\,(X < x) = \mathrm{P}\,(-\sqrt{x} < Z < \sqrt{x})$.

3. La variable aleatoria U tiene distribución uniforme en (0,1).
   a) Obtenga la distribución de probabilidades de $W = U^2$.
   b) Demostrar que $X = \frac{-1}{\lambda}\ln(1-U)$, con $\lambda > 0$, es una variable aleatoria con distribución exponencial de parámetro λ. Realice una simulación generando 10 valores de X si $\lambda = 5$ usando el generador de números aleatorios (en realidad pseudo-aleatorios) de su calculadora de mano.

4. La variable aleatoria X tiene distribución uniforme en (1,3).
   a) Obtener las funciones densidad de probabilidad de $Y = 3X + 4$ y $Z = e^X$.
   b) Calcular los valores esperados de Y y Z.

5. El radio R de una esfera se considera una variable aleatoria continua. Supongamos que R tiene una función densidad de probabilidad $\mathrm{f}_R(r) = 6\,r\,(1-r) \quad 0 < r < 1$. Obtener la función densidad de probabilidad del volumen V y del área A de la esfera.

   Resolución aquí. **Video con el ejercicio explicado.**

6. El beneficio total de una empresa está dado por $B = 10\,Q - 5\,Q^2$ (en miles de pesos) si Q es la cantidad vendida, que se supone es una variable aleatoria continua que toma valores entre 0 y 2 con función densidad de probabilidad: $\mathrm{f}_Q(q) = q$ si $0 < q < 1$ y $\mathrm{f}_Q(q) = -q + 2$ si $1 < q < 2$.
   a) Calcular la probabilidad de obtener un beneficio superior a los 3000 pesos.
   b) Calcular el valor esperado de B.
   c) Obtener la función de distribución de B.

   Resolución aquí.

7. El costo C para producir una cantidad X de un producto viene dado por $C(X) = (X - 4)(X - 8) + 10$ (en miles de pesos). Suponga que X es una variable aleatoria continua con distribución uniforme en (4, 8).
   a) Calcular la probabilidad de que el costo sea mayor que 9.

b) Calcular el valor esperado de C.

c) Obtener la función de distribución de C.

8. Una corriente eléctrica variable I se puede considerar como una variable aleatoria distribuida uniformemente en el intervalo (9, 11). Si esta corriente pasa por una resistencia de 2 Ω entonces la potencia disipada W viene dada por $W = 2\,I^2$.

   a) Obtener la función densidad de probabilidad de W.

   b) Calcular $\mathrm{E}[W]$ y $\mathrm{Var}[W]$.

   c) Calcular $\mathrm{P}(W > \mathrm{E}[W])$.

9. La variable aleatoria continua X tiene una función densidad de probabilidad tal que $\mathrm{f}_X(x) = x + 1$ para $x \in (-1, 0)$, $\mathrm{f}_X(x) = -x + 1$ en $(0, 1)$ y cero fuera del intervalo $(-1, 1)$.

   a) Obtener las funciones de distribución y densidad de probabilidad de $Y = X^2$.

   b) Determinar $\mathrm{E}[Y]$ y $\mathrm{Var}[Y]$.

10. Sean $X \sim \text{Uniforme}(0, 1)$ e $Y = w(X)$, donde $w : [0, 1] \to \mathbb{R}$ viene dada por

$$w(x) = \begin{cases} -1 & x \leq 0.3 \\ +0 & 0.3 < x \leq 0.7 \\ +1 & x > 0.7. \end{cases} \tag{64}$$

   a) ¿Y es una variable aleatoria discreta o continua?

   b) Determine la función de masa de probabilidad o la función de densidad de probabilidad de Y, según corresponda.

   c) Calcule $\mathrm{E}[Y]$ y $\mathrm{Var}[Y]$.

11. Sea $X \sim \text{Uniforme}(0, 1)$. Encuentre una función $w : [0, 1] \to \mathbb{R}$ tal que $Y = w(X) \sim \text{Binomial}(3, 0.5)$.

12. Sean $X \sim \text{Uniforme}(0, 1)$ e $Y = w(X)$, donde $w : [0, 1] \to \mathbb{R}$ viene dada por

$$w(x) = \begin{cases} 0 & x \leq 0.5 \\ x - 0.5 & x > 0.5. \end{cases} \tag{65}$$

   a) ¿Y es una variable aleatoria discreta o continua?

   b) Calcule $\mathrm{E}[Y]$ y $\mathrm{Var}[Y]$.

13. La tabla siguiente representa la distribución de probabilidades conjunta $\mathrm{p}_{X,Y}(x, y)$. Calcule todas las distribuciones marginales y condicionales.

$y \downarrow$ \ $x \rightarrow$	1	2	3	4	5
1	0.10	0.10	0.05	0.15	0.10
2	0.05	0.05	0.05	0.15	0.00
3	0.05	0.05	0.00	0.10	0.00

14. Considere la siguiente tabla de la función de masa de probabilidad conjunta $p_{X,Y}(x,y)$ de las variables aleatorias discretas X e Y:

$y \downarrow$ \ $x \rightarrow$	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

Obtener:

a) las distribuciones de probabilidad marginales;

b) la distribución de probabilidades de $X+Y$ y XY;

c) el valor esperado y la varianza de $X+Y$ y de XY.

d) la covarianza $\operatorname{Cov}[X,Y]$.

Resolución aquí.

15. Dada la variable aleatoria bidimensional

$$(X,Y) = \{(-2,4), (-1,1), (1,1), (2,4)\}$$

tal que todos los pares tienen probabilidad $1/4$,

a) obtener las distribuciones marginales de X y de Y;

b) calcular $\mathrm{E}[X]$, $\mathrm{E}[Y]$, $\mathrm{E}[XY]$;

c) verificar que X e Y no son independientes (existe entre ellas una relación funcional sencilla).

16. Una urna contiene 10 bolillas blancas y 5 negras. Se tira una moneda al aire. Si sale cara se agregan 10 bolillas negras a la urna; si sale ceca se agregan 10 bolillas blancas. Se saca una bolilla al azar y se observa el color.
Se definen dos variables aleatorias discretas: X toma el valor 0 si sale ceca al arrojar la moneda y 1 si sale cara, e Y que toma el valor 0 si la bolilla extraída es negra y 1 si es blanca. Obtener:

a) la distribución de probabilidad condicional $p_{Y|X}(y|x)$;

b) la función de masa de probabilidad conjunta $p_{X,Y}(x,y)$;

c) las funciones de masa de probabilidad marginales $p_X(x)$ y $p_Y(y)$;

d) la covarianza $\operatorname{Cov}[X,Y]$.

17. Se tira un dado y se anota el número X que sale. Se tira una moneda X veces y se anota el número Y de caras que salen.

a) Obtener $p_{Y|X}$ para X tomando los valores 2, 3 y 4.

b) Obtener la distribución conjunta de probabilidades para X tomando los valores 1, 2, 3 y 4.

18. Si la distribución de probabilidad conjunta de la variable aleatoria bidimensional (X,Y) viene dada por $p_{X,Y}(x,y) = (x+y)/30$ para $x \in \{0, 1, 2, 3\}$, $y \in \{0, 1, 2\}$. Obtener:

a) $P(X \leq 2, Y = 1)$;
b) $P(X > 2, Y \leq 1)$;
c) $P(X > Y)$;
d) $P(X + Y = 4)$;
e) $P(Y = 1|X = 2)$;
f) la función de masa de probabilidad condicional $p_{Y|X}(y|x)$;
g) las funciones de masa de probabilidad de X e Y;
h) $E[X]$ y $E[Y]$;
i) $E[X + Y]$ y $\mathrm{Var}[X + Y]$.

19. Cierto supermercado tiene una caja de atención común y otra caja rápida. Supongamos que X es el número de clientes que están en espera en la caja común en un momento particular del día, y que Y es el número de clientes que están en espera en la caja rápida al mismo tiempo. La distribución de probabilidades conjunta de (X, Y) se resume en la siguiente tabla:

$x \downarrow$ \ $y \rightarrow$	0	1	2	3
0	0.08	0.07	0.04	0.00
1	0.06	0.15	0.05	0.04
2	0.05	0.04	0.1	0.06
3	0.00	0.03	0.04	0.07
4	0.00	0.01	0.05	0.06

a) ¿Cuál es la probabilidad de que haya exactamente un cliente en cada línea de espera?
b) ¿Cuál es la probabilidad de que la cantidad de clientes en cada cola sea la misma?
c) Calcular la probabilidad de que haya por lo menos dos clientes más en una cola de espera que en la otra.
d) ¿Cuál es la probabilidad de que la cantidad de clientes en ambas colas sea exactamente 4?
e) Calcular el valor esperado del número de clientes en cada caja.
f) ¿Son X e Y variables aleatorias independientes?

Resolución aquí. Video con el ejercicio explicado.

20. Construya las tablas de la función de probabilidad de dos variables aleatorias discretas X e Y siendo el recorrido de ambas finito, la primera variable X de recorrido $\mathcal{R}_X = \{0\ 1\ 2\ 3\ 4\}$ y la segunda, Y, de recorrido $\mathcal{R}_Y = \{2\ 4\ 6\ 8\}$. Obtenga la distribución de probabilidades de la variable aleatoria $Z = X + Y$ si X e Y son independientes. Extraiga conclusiones respecto de la función de probabilidad de Z.
Sugerencia: Observe como la función de probabilidad de Z se obtiene como los coeficientes de un polinomio que es producto de otros dos polinomios.

21. Dos máquinas A y B, que se accionan independientemente pueden tener un cierto número de fallas cada día. La distribución de probabilidades del número de fallas X e Y, respectivamente de las máquinas A y B, es:

Nro. de fallas k	0	1	2	3	4	5	6
$P(X = k)$	0.10	0.20	0.30	0.20	0.09	0.07	0.04
$P(Y = k)$	0.30	0.10	0.10	0.10	0.10	0.15	0.15

Expresar los sucesos que siguen en términos de X e Y y calcular sus probabilidades:

a) A y B tienen el mismo número de fallas;

b) el número de fallas entre ambas máquinas es menor que 4;

c) A tiene más fallas que B;

d) B tiene el doble de fallas que A;

e) B tiene 4 fallas, cuando se sabe que B tiene por lo menos 2 fallas;

f) el número mínimo de fallas de las dos máquinas es 3.

22. Un examen consta de dos partes. Para un estudiante seleccionado al azar se definen las variables aleatorias: X = número de puntos en la primera parte, Y = número de puntos en la segunda parte. Supongamos que la siguiente tabla muestra la distribución conjunta $\mathrm{p}_{X,Y}(x,y)$ de X e Y:

$x \downarrow$ \ $y \rightarrow$	0	5	10	15
0	0.02	0.06	0.02	0.10
5	0.04	0.15	0.20	0.10
10	0.01	0.15	0.14	0.01

La clasificación registrada es la suma $X+Y$ de los puntajes obtenidos en cada parte.

a) Hallar la distribución de probabilidades de $X+Y$ y la esperanza de la clasificación registrada.

b) ¿Son independientes las calificaciones de las dos partes? Justificar la respuesta.

23. Las variables aleatorias discretas X e Y son independientes y cada una tiene distribución binomial de parámetros $n=3$ y $p=0.5$.

a) Calcular $\mathrm{P}(2\,X+3\,Y<10)$ y $\mathrm{P}(X=Y)$.

b) Obtener la distribución de probabilidades de $Z=X+Y$. Verificar que Z tiene distribución binomial e identificar sus parámetros. Calcular $\mathrm{E}[Z]$ y $\mathrm{Var}(Z)$.

Resolución aquí. Video con el ejercicio explicado.

24. Considere las tablas de distribución conjunta de probabilidades:

Tabla 1

$y \downarrow$ \ $x \rightarrow$	-1	0	1
-1	0.0	0.1	0.0
0	0.1	0.6	0.1
1	0.0	0.1	0.0

Tabla 2

$y \downarrow$ \ $x \rightarrow$	-1	0	1
-1	0.2	0.0	0.0
0	0.0	0.6	0.0
1	0.0	0.0	0.2

Tabla 3

$y \downarrow$ \ $x \rightarrow$	-1	0	1
-1	0.0	0.0	0.2
0	0.0	0.6	0.0
1	0.2	0.0	0.0

En cada caso, analice la independencia de las variables X e Y y calcule:

a) $\mathrm{E}[X]$, $\mathrm{E}[Y]$, $\mathrm{E}[XY]$;

b) $\mathrm{Var}[X]$ y $\mathrm{Var}[Y]$;

c) $\mathrm{Cov}[X,Y]$ y el coeficiente de correlación $\rho(X,Y)$.

25. La función de densidad de probabilidad conjunta de (X,Y) es

$$\mathrm{f}_{X,Y}(x,y) = \begin{cases} a(x+y) & 0<x<3,\, 0<y<3 \\ 0 & x \notin (0,3)\times(0,3) \end{cases}$$

a) Calcular $\mathrm{P}(1<X<2, 1<Y<2)$

b) Calcular $\mathrm{E}[X]$, $\mathrm{E}[Y]$, σ_X y σ_Y.

Resolución aquí.

26. La cantidad de combustible, en miles de litros, que contiene un tanque al principio del día es una variable aleatoria X, de la cual una cantidad aleatoria Y se vende durante el día. Se supone que el tanque no se rellena durante el día, de forma tal que la función de densidad de probabilidad conjunta es:

$$\mathrm{f}_{X,Y}(x,y) = \begin{cases} 2 & 0<x<1,\, 0<y<x \\ 0 & \text{en otro caso.} \end{cases}$$

a) Determine si X e Y son independientes.

b) Calcule la probabilidad de que se hayan vendido en un día más de 500 litros, si la cantidad de combustible que había en el tanque al principio del día era mayor a 700 litros.

27. Dos componentes electrónicos tienen la siguiente distribución de probabilidades conjunta para sus tiempos de duración X e Y:

$$\mathrm{f}_{X,Y}(x,y) = \begin{cases} x\exp[-x(1+y)] & 0<x,\, 0<y \\ 0 & x<0 \text{ o } y<0. \end{cases}$$

a) ¿Cuál es la probabilidad de que la duración X del primer componente sea mayor que 3?

b) Obtener las distribuciones marginales de probabilidad de cada variable y analizar si son independientes.

c) Calcular la covarianza de X e Y.

28. Se cortan chapas rectangulares de lados X (largo) e Y (ancho) que se suponen variables aleatorias con densidad de probabilidad conjunta

$$\mathrm{f}_{X,Y}(x,y) = \begin{cases} k\,x^2\,y & 1<x<2,\, 1<y<x \\ 0 & \text{en otro caso.} \end{cases}$$

a) Obtenga el valor de k.

b) Calcule la probabilidad de ocurrencia de cada uno de estos eventos: *el largo es mayor que 1.8*, *el largo es un 20 % mayor que el ancho* y *el área de la chapa es mayor a 2.*

c) Calcule el ancho y el largo medio.

d) ¿ Son el ancho y largo variables aleatorias independientes?

29. El precio de mercado de cierto bien se supone que es una variable aleatoria P con distribución uniforme en el intervalo (10, 12). La demanda D de ese producto depende del precio P con una densidad condicional de probabilidad dada por $\mathrm{f}_{D|P}(d|p) = e^{-(d-p)}\ d > p$.

 a) Obtenga la función densidad de probabilidad de la demanda.

 b) Calcule la probabilidad de que la demanda supere 13.

30. Considere dos dispositivos que forman parte de un sistema y sean X e Y sus tiempos de vida o duración (en miles de horas) hasta que fallen. Si X e Y son variables aleatorias independientes con distribución exponencial de parámetro λ:

 a) Obtenga la distribución de probabilidades conjunta de (X, Y).

 b) ¿Cuál es la probabilidad de que cada dispositivo dure a lo sumo $1/\lambda$?
 (*Sugerencia:* considere $\mathrm{P}(X < 1/\lambda,\ Y < 1/\lambda)$).

 c) ¿Cuál es la probabilidad de que la duración total $(X + Y)$ sea a lo sumo t miles de horas?
 (Sugerencia: considere la región $A = \{(x, y) : x > 0,\ y > 0,\ x + y < t\}$ e integre).

31. Dos personas han quedado citadas en un lugar. El instante de llegada de cada una a ese lugar puede suponerse una variable aleatoria con distribución uniforme en (0, 30) y que esos instantes de llegada son independientes.

 a) Obtener la función de distribución y de densidad de probabilidad del tiempo T que una persona espera a la otra.

 b) Calcular el valor esperado y el desvío estándar de T.

 Resolución aquí. Video con el ejercicio explicado.

32. Se eligen dos números al azar en (0,1). Se supone que la elección de uno es independiente de la del otro. Calcular la probabilidad de que:

 a) la suma sea menor que 1;

 b) la suma de sus cuadrados sea menor que 1;

 c) su producto no exceda $\frac{2}{9}$.

33. Suponga que X e Y son variables aleatorias. Demostrar que

 a) si X e Y son independientes entonces $\mathrm{Cov}\,[X, Y] = \rho(X, Y) = 0$;

 b) $\mathrm{Cov}\,[aX + b, cY + d] = a\,c\ \mathrm{Cov}\,[X, Y]$ para a, b, c y d constantes;

 c) $\rho(aX + b,\ cY + d) = \rho(X, Y)$ para a, b, c y d constantes y $\mathrm{sign}(a) = \mathrm{sign}(c)$;

 d) Si $Y = aX + b$ con a y b constantes y $(a \neq 0)$ entonces $\rho(X, Y) = \mathrm{sign}(a)$.

34. Sea γ (> 0) un punto sobre el eje de las ordenadas y $\Theta \sim$ Uniforme$(-\pi/2, +\pi/2)$. Definimos la variable X como el lugar donde la semi recta con origen en γ y ángulo Θ respecto del eje de las ordenadas intersecta el eje de las abscisas. Es decir,

$$X = \gamma \tan(\Theta). \tag{66}$$

a) Demuestre que $X \sim$ Cauchy$(0, \gamma)$, es decir,

$$\mathrm{f}_X(x) = \frac{1}{\pi} \frac{\gamma}{\gamma^2 + x^2}. \tag{67}$$

b) Muestre que $\mathrm{E}\left[X^k\right]$ no existe para ningún $k > 0$.

35. Sean X e Y dos variables aleatorias independientes e idénticamente distribuidas (i.i.d.) y sea $W = X + Y$. Determine la función de masa de probabilidad (si son discretas) o la densidad de probabilidad (si son continuas) de W en los siguientes casos:

a) $X \sim$ Bernoulli(p).

b) $X \sim$ Binomial(n, p).

c) $X \sim$ Poisson(λ).

d) $X \sim$ Geométrica(p).

e) $X \sim$ BinoNeg(n, p).

f) $X \sim$ Hipergeo$(10, 5, 2)$.

g) $X \sim N(0, 1)$. *Ayuda: complete cuadrados.*

h) $X \sim$ Cauchy(0,1). *Ayuda: use la identidad de Lagrange, i.e.,*

$$(a^2 + b^2) \times (c^2 + d^2) = (ac + bd)^2 + (ad - bc)^2.$$

i) $X \sim$ Uniforme$(0, 1)$.

j) $X \sim$ Exp(1).

k) $X = Z^2$, con $Z \sim N(0, 1)$. *Ayuda: para* $w > 0$

$$\begin{aligned} \mathrm{F}_W(w) = \mathrm{P}\left(X^2 + Y^2 \leq w\right) &= \iint\limits_{x^2+y^2 \leq w} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dxdy \\ &= \frac{1}{2\pi} \int_0^{2\pi} \int_0^{\sqrt{w}} e^{-\frac{r^2}{2}} r dr d\theta, \end{aligned}$$

donde se ha pasado a coordenadas polares.

36. Sean X_1, X_2 dos variables aleatorias continuas con densidad de probabilidad conjunta

$$\mathrm{f}_{X_1,X_2}(x_1, x_2) = \frac{1}{2\pi\sqrt{\det(\Sigma)}} \exp\left\{-\frac{1}{2}\left(\vec{x} - \vec{\mu}\right)^T \Sigma^{-1} \left(\vec{x} - \vec{\mu}\right)\right\}, \tag{68}$$

donde

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \; \vec{\mu} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \; \Sigma = \begin{pmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{pmatrix}, \tag{69}$$

donde $\sigma_1^2, \sigma_2^2 > 0$, $\det(\Sigma) > 0$.

Demuestre que:

a) $X_i \sim N(\mu_i, \sigma_i)$ para $i = 1, 2$.

b) $\text{Cov}[X_1, X_2] = \sigma_{1,2}$.

c) X_1, X_2 son independientes sii $\sigma_{1,2} = 0$.

37. La tensión de salida V del receptor de un sistema de comunicación digital cuando se recibe un **0**, es Gaussiana (sólo ruido) con media 0 y dispersión 0.3. Cuando se recibe un **1** binario la tensión también tiene distribución Gaussiana (ahora señal más ruido) de media 0.9 y y dispersión 0.25. La lógica del receptor decide que si $V > 0.50$ entonces se recibió un **1** y un **0** en caso contrario. Si los ceros y unos llegan al sistema en igual proporción:

    a) Calcular la probabilidad de que se cometa error en la detección del bit que se decodifica,

    b) Obtener la función de distribución de V y su valor esperado.
    Sugerencia: Construya un árbol y use el teorema de la probabilidad total para obtener la función de distribución de V.

38. Dos máquinas A y B producen piezas cuyos pesos (en gramos) se pueden suponer variables aleatorias continuas con las siguientes funciones densidad de probabilidad: $\text{f}_A(x) = (16 - 2\,x)/49$ si $x \in (1, 8)$ y 0 fuera de ese intervalo, y $\text{f}_B(x) = (2\,x - 2)/49$ si $x \in (1, 8)$ y 0 fuera de ese intervalo. La máquina A produce el 40 % de las piezas y el resto lo produce la máquina B.

    a) Obtener las funciones de distribución y densidad de probabilidad del peso X de las piezas de la producción total.

    b) Determinar el valor esperado $\text{E}[X]$ y el desvío estándar σ_X.

39. La cantidad de órdenes de compra diarias en un negocio se puede considerar una variable aleatoria N con distribución Poisson de parámetro Λ. Dado que muchos factores influyen en la cantidad de pedidos, el parámetro de la distribución Poisson es también considerado una variable aleatoria $\Lambda \sim \text{Uniforme}(0, 10)$.

    a) Calcule $\text{P}(N = 5)$. Muestre que es diferente a $\text{P}\big(N = 5 |\, \Lambda = \text{E}[\Lambda]\big)$.

    b) Determine $\text{E}[N|\Lambda]$ y $\text{Var}[N|\Lambda]$.

    c) Encuentre $\text{E}[N]$ y $\text{Var}[N]$.

    Ayuda:

$$\int x^5 e^{-x} dx = -e^{-x}\left(x^5 + 5x^4 + 20x^3 + 60x^2 + 120x + 120\right) + \text{constante}.$$

40. La cantidad de órdenes de compra diarias en otro negocio también se puede considerar una variable aleatoria N con distribución Poisson de parámetro Λ. Sin embargo, en este caso, el parámetro de la distribución Poisson es considerado una variable aleatoria $\Lambda \sim \text{Exp}(0.2)$.

    a) Calcule $\text{P}(N = k)$ para $k \in \mathbb{N}^0$.

    b) Determine $\text{E}[N|\Lambda]$ y $\text{Var}[N|\Lambda]$.

    c) Encuentre $\text{E}[N]$ y $\text{Var}[N]$.

    Ayuda:

$$\int_0^{+\infty} x^k e^{-ax} dx = \frac{k!}{a^{k+1}} \qquad \forall k \in \mathbb{N}^0, \forall a \in \mathbb{R}^+.$$

41. Sean $U_i \sim \text{Uniforme}(0,1)$, $i = 1, 2$, variables aleatorias independientes. Definamos

$$Z_1 = \sqrt{-2\ln(U_1)}\,\cos(2\pi U_2), \tag{70}$$

$$Z_2 = \sqrt{-2\ln(U_1)}\,\sin(2\pi U_2). \tag{71}$$

Demuestre que $Z_i \sim \mathcal{N}(0,1)$, $i = 1, 2$, i.i.d.

3. Respuestas

1. a) $\mathrm{E}[Y] = 1.5$, $\mathrm{Var}[Y] = 20.25$; b) $\mathrm{E}[Y] = 2.5$, $\mathrm{Var}[Y] = 2.25$;
c) $\mathrm{E}[Y] = 4.6$, $\mathrm{Var}[Y] = 108.84$.

2. a) La variable aleatoria X tiene distribución normal de media m y dispersión s; b) $\mathrm{f}_X(x) = \frac{1}{\sqrt{2\,\pi\,x}}\; e^{-x/2}\, I_{(0,\infty)}(x)$.

3. a) $\mathrm{f}_W(w) = \frac{1}{2\sqrt{w}}\, I_{(0,1)}(x)$.

4. $g(y) = \frac{1}{6}\, I_{[7,13)}$, $\mathrm{E}[Y] = 10$; $h(z) = \frac{1}{2\,z} \quad , z \in (e, e^3)$, $\mathrm{E}[Z] = (e^3 - e)/2$.

5. $h(v) = \frac{3}{2\,\pi}\, \left((\frac{3\,v}{4\,\pi})^{\frac{-1}{3}} - 1 \right)\; 0 < v < \frac{4\,\pi}{3}$.

6. a) 0.865; b) $\frac{25}{6}$; c) $\mathrm{P}(B < b) = (1 - (1 - 0.2\,b)^{0.5})^2$.

7. a) 0.13397; b) 7.33; c) $\mathrm{F}_C(c) = 0.5\,\sqrt{c-6}\; I_{[6,10)}(c) + 1\, I_{[10,\infty)}(c)$.

8. a) $f(w) = \frac{1}{4\,\sqrt{2\,w}} \quad 162 < w < 242$; b) $\mathrm{E}[W] = \frac{602}{3}$ y $\mathrm{Var}[W] = \frac{24016}{45}$; c) $\mathrm{P}(W > \frac{602}{3}) \approx 0.492$.

9. a) La función densidad de probabilidad de Y es

$$\mathrm{f}_Y(y) = \left(\frac{1}{\sqrt{y}} - 1 \right)\; I_{(0,1)}(y);$$

b) $\mathrm{E}[Y] = \frac{1}{6}$ y $\mathrm{Var}[Y] = \frac{7}{180}$.

10. c) $\mathrm{E}[X] = 0$, $\mathrm{Var}[X] = 0.6$.
12. b) $\mathrm{E}[X] = 1/8$, $\mathrm{Var}[X] = \frac{5}{192}$.
13. La distribución marginal de X:

x	1	2	3	4	5
$\mathrm{p}_X(x)$	0.2	0.2	0.1	0.4	0.1

La distribución marginal de Y:

y	1	2	3
$\mathrm{p}_Y(y)$	0.5	0.3	0.2

La distribución condicional $\mathrm{p}_{X|Y}(x|y)$:

$y \rightarrow$ / $x \downarrow$	1	2	3
1	0.20	1/6	0.25
2	0.20	1/6	0.25
3	0.10	1/6	0.00
4	0.30	0.50	0.50
5	0.20	0.00	0.00

La distribución condicional $\mathrm{p}_{Y|X}(y|x)$:

$y \rightarrow$ / $x \downarrow$	1	2	3
1	0.500	0.250	0.250
2	0.500	0.250	0.250
3	0.500	0.500	0.000
4	0.375	0.375	0.250
5	1.000	0.000	0.000

14. a) La distribución marginal de X:

x	0	1	2	3	4	5
$\mathrm{p}_X(x)$	0.03	0.08	0.16	0.21	0.24	0.28

La distribución marginal de Y:

y	0	1	2	3
$\mathrm{p}_Y(y)$	0.25	0.26	0.25	0.24

b) La distribución $\mathrm{p}_Z(z)$ de $Z = X + Y$:

z	1	2	3	4	5	6	7	8
$\mathrm{p}_Z(z)$	0.02	0.06	0.13	0.19	0.24	0.19	0.12	0.05

La distribución $\mathrm{p}_W(w)$ de $W = X\,Y$:

w	0	1	2	3	4	5	6	8	9	10	12	15
$\mathrm{p}_W(w)$	0.28	0.02	0.07	0.07	0.11	0.08	0.09	0.05	0.06	0.06	0.06	0.05

c) $\mathrm{E}[X+Y] = 4.87$, $\mathrm{Var}[X+Y] = 2.6731$, $\mathrm{E}[X\,Y] = 4.76$, $\mathrm{Var}[X\,Y] = 19.2224$;
d) -0.2572.

15. a) La distribución marginal $\mathrm{p}_X(x)$:

x	-2	-1	$+1$	$+2$
$\mathrm{p}_X(x)$	0.25	0.25	0.25	0.25

La distribución marginal $\mathrm{p}_Y(y)$:

y	1	4
$\mathrm{p}_Y(y)$	0.50	0.50

b) $\mathrm{E}[X] = 0$, $\mathrm{E}[Y] = 2.5$ y $\mathrm{E}[X\,Y] = 0$;
c) Se cumple $Y = X^2$;

16. b) La distribución conjunta $\mathrm{p}_{X,Y}(x, y)$:

$x \downarrow$ \ $y \rightarrow$	0	1
0	0.1	0.4
1	0.3	0.2

c) La distribución marginal $\mathrm{p}_X(x)$:

x	0	1
$\mathrm{p}_X(x)$	0.5	0.5

La distribución marginal $\mathrm{p}_Y(y)$:

y	0	1
$\mathrm{p}_Y(y)$	0.4	0.6

d) -0.1.

17. a) La función de probabilidad condicional $\mathrm{p}_{Y|X}$ viene dada por:

$$\mathrm{p}_{Y|X}(r|k) = \binom{k}{r} \frac{1}{2^k} \text{ para } 0 \leq r \leq k,\ k \in \{1, 2, 3, 4, 5, 6\};$$

b) La función de probabilidad conjunta es

$$p_{X,Y}(k,r) = \binom{k}{r} \frac{1}{6\,2^k} \text{ para } 0 \le r \le k,\ k \in \{1,2,3,4,5,6\}.$$

18. a) 0.2; b) $\frac{7}{30}$; c) 0.6, d) $\frac{4}{15}$; e) $\frac{1}{3}$;
f) La probabilidad condicional $p_{Y|X}(y|x)$):

$x \downarrow$ \ $y \rightarrow$	0	1	2
0	0.00	1/3	2/3
1	1/6	1/3	1/2
2	2/9	1/3	4/9
3	1/4	1/3	5/12

g) La distribución marginal $p_X(x)$:

x	0	1	2	3
$p_X(x)$	0.1	0.2	0.3	0.4

La distribución marginal $p_Y(y)$:

y	0	1	2
$p_Y(y)$	1/5	1/3	7/15

h) $E[X] = 2$, $E[Y] = \frac{19}{15}$; i) $E[X+Y] = \frac{49}{15}$, $\text{Var}[X+Y] = \frac{299}{225} \approx 1.3289$.

19. a) 0.15; b) 0.4; c) 0.22; d) 0.17; e) $E[X] = 1.7$, $E[Y] = 1.55$; f) X e Y no son independientes.

20. a) 0.1255; b) 0.34; c) 0.417; d) 0.08 (no tenga en cuenta el par (0, 0)); e) $\frac{1}{6}$; f) 0.12.

21. a) Sea $Z = X + Y$. La función de probabilidad p_Z es:

z	0	5	10	15	20	25
$p_Z(z)$	0.02	0.10	0.18	0.45	0.24	0.01

(0, 0.02), (5, 0.1), (10, 0.18), (15, 0.45), (20, 0.24) y (25, 0.01).
El valor esperado de Z es 14.1S.
b) Las variables X e Y no son independientes.

22. a) $\frac{45}{64}$ y $\frac{5}{16}$; b) La variable Z tiene distribución binomial de parámetros $n = 6$ y $p = 0.5$.

23.

	Tabla 1	Tabla 2	Tabla 3
$E[X]$	0	0	0
$E[Y]$	0	0	0
$E(XY)$	0	0.4	-0.4
$\text{Var}[X]$	0.2	0.4	0.4
$\text{Var}[Y]$	0.2	0.4	0.4
$\text{Cov}[X,Y]$	0	0.4	-0.4
$\rho(X,Y)$	0	1	-1
Independencia	NO	NO	NO

24. a) $\frac{1}{9}$ b) $E[X] = E[Y] = \frac{7}{4}$ $\text{Var}[X] = \text{Var}[Y] = \frac{11}{16}$.

25. a) No son independientes; b) 0.412.

26. a) e^{-3}; b) La variable X tiene distribución exponencial de parámetro 1, y la variable Y tiene a $\frac{1}{(1+y)^2}$ $\forall\, y \in \mathbb{R}$ como función densidad de probabilidad. Las variables aleatorias X e Y no son independientes; c) El valor esperado de Y no es finito.

27. a) $\frac{15}{29}$; b) Las probabilidades solicitadas son 0.4909, 0.5194 y 0.7566; c) $\frac{405}{232}$ y $\frac{245}{174}$; d) No son independientes.

28. La función densidad de probabilidad de D viene dada por

$$\mathrm{f}_D(d) = ((1 - e^{(10-d)})/2)\, I_{(10,12)}(x) + ((e^{(12-d)} - e^{(10-d)})/2)\, I_{(12,\infty)}(x);$$

b) $\mathrm{P}(D > 13) = (e^{-1} - e^{-3})/2 \approx 0.1590$.

29. a) $\mathrm{f}(x, y) = \lambda^2 \exp(-\lambda\,(x + y))$ si $x > 0$ e $y > 0$; b) $(1 - \exp(-1))^2$; c) $1 - \exp(-\lambda\, t) - \lambda\, t \exp(-\lambda\, t)$.

30. a) $\mathrm{F}_T(t) = \frac{1}{900}\,(60\, t - t^2)\, I_{[0,30]}(t) + I_{(30,\infty)}(t)$ y la función densidad de probabilidad de T es $\mathrm{f}_T(t) = \frac{1}{900}\,(60 - 2\, t)\, I_{(0,30)}(t)$; b) El valor esperado de T es 10 y la varianza es 50.

31. a) 0.5; b) $\frac{\pi}{4}$; b) $\frac{2}{9}\,(1 + \ln \frac{9}{2})$.
34.

a) $W \sim \text{Binomial}(2, p)$.

b) $W \sim \text{Binomial}(2n, p)$.

c) $W \sim \text{Poisson}(2\lambda)$.

d) $W \sim \text{BinoNeg}(2, p)$.

e) $W \sim \text{BinoNeg}(2n, p)$.

f) $\mathcal{R}_W = \{0, 1, 2, 3, 4\}$, $p_W = \{0.0494, 0.2469, 0.4074, 0.2469, 0.0494\}$ (en el mismo orden).

g) $W \sim N(0, \sqrt{2})$.

h) $W \sim \text{Cauchy}(0, 2)$.

i)

$$\mathrm{f}_W(w) = \begin{cases} w & w \in (0,1), \\ 2 - w & w \in (1,2), \\ 0 & w \notin (0,2). \end{cases}$$

j) $W \sim \text{Exp}(2)$. También se dice que tiene distribución *ji*-cuadrado con 2 grados de libertad. Es decir, $W \sim \chi^2_2$.

36. a) 0.0513; b) $\mathrm{F}_V(v) = \frac{1}{2}\,(\Phi(v/0.3) + \Phi((v - 0.9)/0.25))$, donde Φ es la función de distribución de la variable aleatoria con distribución normal estándar. El valor esperado de V es $\mathrm{E}\,[V] = 0.45$.

37. La función densidad de probabilidad de X es $\mathrm{f}_X(x) = \frac{1}{245}\,(2\, x + 26)\, I_{(1,8)}(x)$, donde $I_{(a,b)}(x)$ toma valor 1 si $x \in (a, b)$ y 0 en caso contrario. El valor esperado de X es $\frac{71}{15} \approx 4.73$, la varianza $\frac{1813}{450} \approx 4.029$ y el desvío estándar es $\sigma(X) \approx 2.007$.

39. a) 0.0933, 0.1755; b) $\mathrm{E}\,[N|\Lambda] = \mathrm{Var}\,[N|\Lambda] = \Lambda$; c) $\mathrm{E}\,[N] = 5$, $\mathrm{Var}\,[N] = 40/3$.

4. Ejercicios resueltos

Ejercicio 7

El radio R de una esfera se considera una variable aleatoria continua. Supongamos que R tiene una función densidad de probabilidad $\mathrm{f}_R(r) = 6r(1-r),\ 0 < r < 1$. Obtener la función densidad de probabilidad del volumen V y del área A de la esfera.

Resolución

Variable aleatoria

Las variables aleatorias nos vienen dadas y son:

- R = radio de una esfera seleccionada al azar.
- V = volumen de una esfera seleccionada al azar.
- A = área externa de una esfera seleccionada al azar.

Eventos y Datos

Tenemos como datos la densidad de R y además, que V y A son funciones de R:

- $V(R) = \dfrac{4\pi}{3} \cdot R^3$
- $A(R) = 4\pi \cdot R^2$

Además, vale rescatar que todas estas variables aleatorias son continuas.

Nos piden la densidad de V y A. Para eso, podemos empezar calculando sus funciones de distribución y luego derivarlas. Además, si bien no tenemos la distribución de R, sí tenemos su densidad y podemos utilizar esa información ya que V y A son funciones de R.

Distribución de V

Por definición, la distribución de V viene dada por:

$$\mathrm{F}_V(v) = \mathrm{P}(V \leq v), \quad v \in \mathbb{R} \tag{72}$$

Recordemos ahora que $V = V(R) = \dfrac{4\pi}{3} \cdot R^3$. Por lo tanto:

$$\mathrm{F}_V(v) = \mathrm{P}(V \leq v) = \mathrm{P}(\frac{4\pi}{3} \cdot R^3 \leq v) = P\left(R^3 \leq \frac{3v}{4\pi}\right) = P\left(R \leq \sqrt[3]{\frac{3v}{4\pi}}\right) = \mathrm{F}_R\left(\sqrt[3]{\frac{3v}{4\pi}}\right) \tag{73}$$

Es decir, hemos logrado escribir la distribución de V en función de la distribución de R. No tenemos la distribución de R pero veamos que no es necesario. Como nos piden la densidad de V, podemos derivar la expresión obtenida:

$$\begin{aligned}
\mathrm{f}_V(v) &= \frac{d}{dv}[\mathrm{F}_V(v)] = \frac{d}{dv}\left[\mathrm{F}_R\left(\sqrt[3]{\frac{3v}{4\pi}}\right)\right] \\
&= \mathrm{f}_R\left(\sqrt[3]{\frac{3v}{4\pi}}\right) \cdot \underbrace{\frac{d}{dv}\left[\sqrt[3]{\frac{3v}{4\pi}}\right]}_{\substack{\text{Regla de} \\ \text{la cadena}}} = \mathrm{f}_R\left(\sqrt[3]{\frac{3v}{4\pi}}\right) \cdot \frac{1}{3}\left(\frac{3v}{4\pi}\right)^{-\frac{2}{3}} \cdot \underbrace{\frac{3}{4\pi}}_{\substack{\text{Regla de} \\ \text{la cadena}}}
\end{aligned} \tag{74}$$

Notar que aún no hemos determinado valores para v, por lo que la igualdad es para todo $v \in \mathbb{R}$, aunque muchos valores tendrán densidad nula. Intuitivamente, si $0 < r < 1$, entonces $0 < v < \frac{4\pi}{3}$. Sin embargo, esta determinación no siempre es tan intuitiva, por lo que debemos aprender a hacerlo con cuidado para esos casos:

$$\mathrm{f}_R(r) = \begin{cases} 6r(1-r) & \text{si } 0 < r < 1 \\ 0 & \text{en caso contrario} \end{cases} \Rightarrow \tag{75}$$

$$\begin{aligned} \Rightarrow \mathrm{f}_V(v) &= \mathrm{f}_R\left(\sqrt[3]{\frac{3v}{4\pi}}\right) \cdot \frac{1}{3}\left(\frac{3v}{4\pi}\right)^{-\frac{2}{3}} \cdot \frac{3}{4\pi} \\ &= \begin{cases} 6\sqrt[3]{\frac{3v}{4\pi}}\left(1-\sqrt[3]{\frac{3v}{4\pi}}\right)\frac{1}{3}\left(\frac{3v}{4\pi}\right)^{-\frac{2}{3}} \cdot \frac{3}{4\pi} & \text{si } 0 < \sqrt[3]{\frac{3v}{4\pi}} < 1 \\ 0 & \text{en caso contrario} \end{cases} \end{aligned} \tag{76}$$

Aglomerando términos (están exceptuados de la distancia social), tenemos la siguiente expresión:

$$\begin{aligned} \mathrm{f}_V(v) &= \begin{cases} \frac{3}{2\pi}\left(\left(\frac{3v}{4\pi}\right)^{-\frac{1}{3}} - 1\right) & \text{si } 0^3 \cdot \frac{4\pi}{3} < v < 1^3 \cdot \frac{4\pi}{3} \\ 0 & \text{en caso contrario} \end{cases} \Rightarrow \\ \Rightarrow \mathrm{f}_V(v) &= \begin{cases} \frac{3}{2\pi}\left(\sqrt[3]{\frac{4\pi}{3v}} - 1\right) & \text{si } 0 < v < \frac{4\pi}{3} \\ 0 & \text{en caso contrario} \end{cases} \end{aligned} \tag{77}$$

Observación: Notar que cuando $v \to 0$, $\mathrm{f}_V(v) \to +\infty$, sin embargo, se puede verificar que la integral impropia de esta densidad da 1 como resultado.

Distribución de A

Nuevamente, comenzamos con la distribución de A:

$$\mathrm{F}_A(a) = \mathrm{P}(A \leq a) = \mathrm{P}(4\pi R^2 \leq a) = P\left(R^2 \leq \frac{a}{4\pi}\right) \tag{78}$$

Para poder tomar raíz a ambos lados, el miembro derecho debe ser mayor o igual que cero. Esto no significa que para $a < 0$ no se pueda calcular:

$$\text{Si } a < 0 \Rightarrow \mathrm{F}_A(a) = P\left(\underbrace{R^2}_{\geq 0} \leq \underbrace{\frac{a}{4\pi}}_{<0}\right) = 0 \Rightarrow \text{Si } a < 0 : \mathrm{f}_A(a) = 0 \tag{79}$$

Ahora, tomando $a \geq 0$, podemos tomar la raíz cuadrada a ambos lados:

$$\begin{aligned} P\left(R^2 \leq \frac{a}{4\pi}\right) = P\left(\sqrt{R^2} \leq \sqrt{\frac{a}{4\pi}}\right) &= P\left(|R| \leq \frac{\sqrt{a}}{2\sqrt{\pi}}\right) = P\left(-\frac{\sqrt{a}}{2\sqrt{\pi}} \leq R \leq \frac{\sqrt{a}}{2\sqrt{\pi}}\right) \\ &= \mathrm{F}_R\left(\frac{\sqrt{a}}{2\sqrt{\pi}}\right) - \mathrm{F}_R\left(-\frac{\sqrt{a}}{2\sqrt{\pi}}\right) \end{aligned} \tag{80}$$

Por otro lado, podemos derivar esta expresión respecto de a:

$$\begin{aligned} \mathrm{f}_A(a) &= \frac{d}{da}\left[\mathrm{F}_R\left(\frac{\sqrt{a}}{2\sqrt{\pi}}\right) - \mathrm{F}_R\left(-\frac{\sqrt{a}}{2\sqrt{\pi}}\right)\right] \\ &= \mathrm{f}_R\left(\frac{\sqrt{a}}{2\sqrt{\pi}}\right) \cdot \left(\frac{1}{4\sqrt{\pi a}}\right) - \mathrm{f}_R\left(-\frac{\sqrt{a}}{2\sqrt{\pi}}\right)\left(-\frac{1}{4\sqrt{\pi a}}\right) \end{aligned} \tag{81}$$

Tengamos en cuenta que $-\dfrac{\sqrt{a}}{2\sqrt{\pi}} < 0$, $\mathrm{f}_R\left(-\dfrac{\sqrt{a}}{2\sqrt{\pi}}\right) = 0$. Entonces,

$$\mathrm{f}_A(a) = \frac{1}{4\sqrt{\pi a}}\mathrm{f}_R\left(\frac{\sqrt{a}}{2\sqrt{\pi}}\right) = \begin{cases} \dfrac{1}{4\sqrt{\pi a}}6\dfrac{\sqrt{a}}{2\sqrt{\pi}}\left(1-\dfrac{\sqrt{a}}{2\sqrt{\pi}}\right) & \text{si } 0 < \frac{\sqrt{a}}{2\sqrt{\pi}} < 1 \\ 0 & \text{en caso contrario} \end{cases} \tag{82}$$

Luego:

$$\mathrm{f}_A(a) = \begin{cases} \dfrac{3}{4\pi}\left(1-\dfrac{\sqrt{a}}{2\sqrt{\pi}}\right) & \text{si } 0 < a < 4\pi \\ 0 & \text{en caso contrario} \end{cases} \tag{83}$$

Ejercicio 8.

El beneficio total de una empresa está dado por $B = 10\,Q - 5\,Q^2$ (en miles de pesos) si Q es la cantidad vendida, que se supone es una variable aleatoria continua que toma valores entre 0 y 2 con función densidad de probabilidad: $\mathrm{f}_Q(q) = q$ si $0 < q < 1$ y $\mathrm{f}_Q(q) = -q + 2$ si $1 < q < 2$.

1. Calcular la probabilidad de obtener un beneficio superior a los 3000 pesos.
2. Calcular el valor esperado de B.
3. Obtener la función de distribución de B.

Resolución

En todo el desarrollo, asumiremos que trabajamos con las cantidades en miles de pesos. El primer paso es determinar la función de distribución de la variable aleatoria Q. Para ello, es conveniente graficar la densidad de probabilidad, como se hace en la Fig. 3. Dado que el área de un triángulo es 0.5×`base`×`altura`, es fácil ver que

$$\mathrm{F}_Q(q) = \begin{cases} 0 & q \leq 0 \\ \frac{q^2}{2} & q \in [0,1] \\ 1 - \frac{(2-q)^2}{2} & q \in [1,2] \\ 1 & q \geq 2, \end{cases} \tag{84}$$

donde hemos usado el hecho $\mathrm{P}\left(Q < 0\right) = \mathrm{P}\left(Q > 2\right) = 0$.

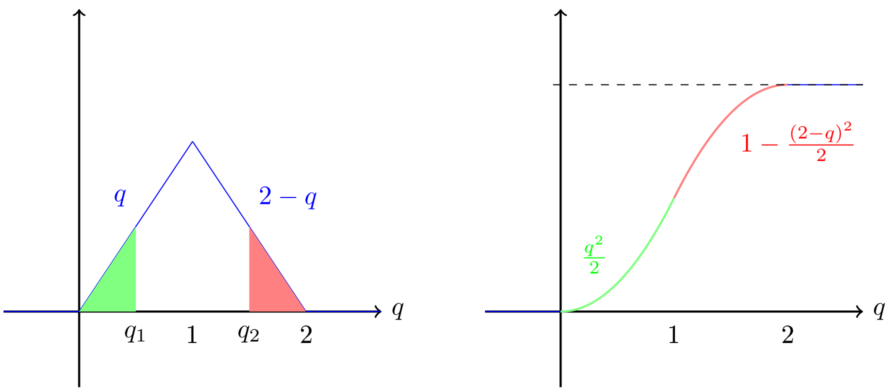


Figura 3: Densidad de probabilidad (izquierda) y función de distribución (derecha) de Q, Ejercicio 8 de la guía 5

Para calcular la función de distribución del beneficio B, también ayuda realizar un gráfico, como el que aparece en la Fig. 4. A partir de esta figura, para $b_\alpha \in [0,5]$

$$\mathrm{F}_B\left(b_\alpha\right) = \mathrm{P}\left(B \leq b_\alpha\right) = \mathrm{P}\left(Q \leq q_{\alpha,1}\right) + \mathrm{P}\left(Q \geq q_{\alpha,2}\right) = \frac{q_{\alpha,1}^2}{2} + 1 - \left[1 - \frac{\left(2-q_{\alpha,2}\right)^2}{2}\right] \tag{85}$$

$$= \frac{q_{\alpha,1}^2}{2} + \frac{\left(2-q_{\alpha,2}\right)^2}{2}, \tag{86}$$

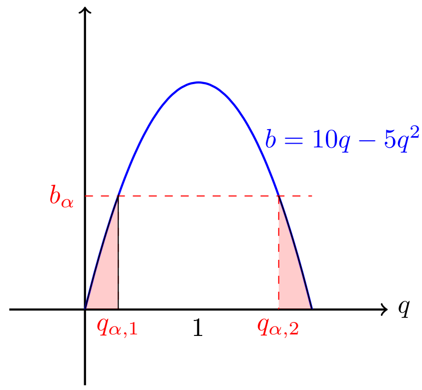


Figura 4: Beneficio en función del costo, Ejercicio 8 de la guía 5

donde $q_{\alpha,1}$ y $q_{\alpha,2}$ pueden ser obtenidos como solución de la ecuación

$$b_\alpha = 10q_\alpha - 5q_\alpha^2 \Rightarrow q_{\alpha,1}, q_{\alpha,2} = 1 \pm \sqrt{1 - \frac{b_\alpha}{5}}. \tag{87}$$

Finalmente, tenemos

$$\mathrm{F}_B(b) = \begin{cases} 0 & b \leq 0 \\ \left(1 - \sqrt{1 - \frac{b}{5}}\right)^2 & b \in [0, 5] \\ 1 & b \geq 5 \end{cases} \tag{88}$$

Obsérvese que es fácil determinar la densidad de probabilidad a partir de este resultado:

$$\mathrm{f}_B(b) = \begin{cases} \frac{1}{5}\left(\frac{1}{\sqrt{1-\frac{b}{5}}} - 1\right) & b \in (0, 5) \\ 0 & b \notin (0, 5) \end{cases} \tag{89}$$

A partir de la Ec. (88), es fácil calcular la probabilidad de obtener un beneficio superior a 3000 pesos:

$$\mathrm{P}(B > 3) = 1 - \mathrm{F}_B(3) = 1 - \left(1 - \sqrt{1 - \frac{3}{5}}\right)^2 \approx 0.8649. \tag{90}$$

El valor esperado del beneficio puede obtenerse de varias formas, por ejemplo, mediante las siguientes integrales:

$$\mathrm{E}[B] = \int_{\mathbb{R}} b\mathrm{f}_B(b)db = \int_{\mathbb{R}} \left(10q - 5q^2\right) \mathrm{f}_Q(q)dq. \tag{91}$$

En este problema, parece más sencillo utilizar la segunda de las integrales. Por lo tanto, tenemos

$$\begin{aligned}
\mathrm{E}[B] &= \int_0^1 \left(10q - 5q^2\right) q dq + \int_1^2 \left(10q - 5q^2\right)(2-q)\, dq \\
&= \int_0^1 \left(10q^2 - 5q^3\right) q dq + \int_1^2 \left(20q - 20q^2 + 5q^3\right) dq \\
&= \left(\frac{10}{3}q^3 - \frac{5}{4}q^4\right)\Bigg|_0^1 + \left(10q^2 - \frac{20}{3}q^3 + \frac{5}{4}q^4\right)\Bigg|_1^2 \\
&= \left[\left(\frac{10}{3} - \frac{5}{4}\right) - (0)\right] + \left[\left(40 - \frac{160}{3} + \frac{80}{4}\right) - \left(10 - \frac{20}{3} + \frac{5}{4}\right)\right] \\
&= \frac{25}{6}.
\end{aligned}$$

Ejercicio 13

Considere la siguiente tabla de la función de probabilidad conjunta $\mathrm{p}_{X,Y}(x,y)$ de las variables aleatorias discretas X e Y:

$y \downarrow$ \ $x \rightarrow$	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

Obtener:

1. las distribuciones de probabilidad marginales;
2. la distribución de probabilidades de $X+Y$ y XY;
3. el valor esperado y la varianza de $X+Y$ y de XY;
4. la covarianza $\mathrm{Cov}[X,Y]$.

Resolución

Distribuciones marginales: Es sencillo calcular las distribuciones marginales sumando las filas (marginal de Y) y las columnas (marginal de X) de la tabla:

$y \downarrow$ \ $x \rightarrow$	0	1	2	3	4	5	p_Y
0	0.00	0.01	0.03	0.05	0.07	0.09	0.25
1	0.01	0.02	0.04	0.05	0.06	0.08	0.26
2	0.01	0.03	0.05	0.05	0.05	0.06	0.25
3	0.01	0.02	0.04	0.06	0.06	0.05	0.24
p_X	0.03	0.08	0.16	0.21	0.24	0.28	—

A partir de esta tabla, es fácil ver que las variables no son independientes. En efecto, tenemos, por ejemplo,

$$\mathrm{p}_{X,Y}(0,0) = 0 \neq 0.03 \times 0.25 = \mathrm{p}_X(0)\mathrm{p}_Y(0). \tag{92}$$

Los valores esperados de cada variable aleatoria se pueden calcular como

$$\mathrm{E}[X] = \sum_{x=0}^{5} x\mathrm{p}_X(x) = 3.39, \tag{93}$$

$$\mathrm{E}[X] = \sum_{y=0}^{3} y\mathrm{p}_Y(y) = 1.48, \tag{94}$$

$$\mathrm{E}\left[X^2\right] = \sum_{x=0}^{5} x^2\mathrm{p}_X(x) = 13.45, \tag{95}$$

$$\mathrm{E}\left[Y^2\right] = \sum_{x=0}^{3} y^2\mathrm{p}_Y(y) = 3.42, \tag{96}$$

$$\mathrm{Var}[X] = (13.45) - (3.39)^2 = 1.9579 \Rightarrow \sigma_X \approx 1.8412, \tag{97}$$

$$\mathrm{Var}[Y] = (3.42) - (1.48)^2 = 1.2296 \Rightarrow \sigma_Y \approx 1.3569. \tag{98}$$

Distribución de XY:

Llamemos $W = XY$. Luego, $\mathcal{R}_Z = \{0, 1, 2, 3, 4, 5, 6, 8, 9, 10, 12, 15\}$. Para calcular la función de probabilidad de W, es conveniente superponer a la tabla de probabilidades conjuntas la siguiente tabla con los valores de W:

$y \downarrow$ \ $x \rightarrow$	0	1	2	3	4	5
0	0	0	0	0	0	0
1	0	1	2	3	4	5
2	0	2	4	6	8	10
3	0	3	6	9	12	15

Marcando con distintos colores los distintos casos que nos interesan de la tabla de probabilidades conjuntas:

$y \downarrow$ \ $x \rightarrow$	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

Luego,

w	0	1	2	3	4	5	6	8	9	10	12	15
p_W	0.28	0.02	0.07	0.07	0.11	0.08	0.09	0.05	0.06	0.06	0.06	0.05

Los momentos de W:

$$\begin{aligned}
\mathrm{E}[W] &= \sum_{w \in \mathcal{R}_W} w \mathrm{p}_W(w) = 4.76, \\
\mathrm{E}\left[W^2\right] &= \sum_{w \in \mathcal{R}_W} w^2 \mathrm{p}_W(w) = 41.88, \\
\mathrm{Var}[W] &= (41.88) - (4.76)^2 = 19.2224 \Rightarrow \sigma_W \approx 4.3843.
\end{aligned}$$

A partir de los resultados hasta aquí obtenidos, podemos calcular la covarianza de las variables X e Y:

$$\mathrm{Cov}[X, Y] = \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] = -0.2572. \tag{99}$$

El coeficiente de correlación es:

$$\rho(X, Y) = \frac{\mathrm{Cov}[X, Y]}{\sigma_X \sigma_Y} \approx -0.1658. \tag{100}$$

Distribución de $X + Y$:

Definamos $Z = X + Y$. Luego, $\mathcal{R}_Z = \{0, 1, 2, \cdots, 8\}$. La función de probabilidad de Z se puede encontrar sumando los valores en las anti-diagonales de la tabla de probabilidades conjuntas, escritas con distintos colores:

$y \backslash x$	0	1	2	3	4	5
0	0.00	0.01	0.03	0.05	0.07	0.09
1	0.01	0.02	0.04	0.05	0.06	0.08
2	0.01	0.03	0.05	0.05	0.05	0.06
3	0.01	0.02	0.04	0.06	0.06	0.05

Por tanto,

z	0	1	2	3	4	5	6	7	8
p_Z	0.00	0.02	0.06	0.13	0.19	0.24	0.19	0.12	0.05

Ahora podemos calcular los momentos de Z:

$$\mathrm{E}[Z] = \sum_{z=0}^{8} z\,\mathrm{p}_Z(z) = 4.87,$$
$$\mathrm{E}\left[Z^2\right] = \sum_{z=0}^{8} z^2\,\mathrm{p}_Z(z) = 26.39,$$
$$\mathrm{Var}[Z] = (26.39) - (4.87)^2 = 2.6731 \Rightarrow \sigma_Z \approx 1.6350.$$

Obsérvese que:

$$\mathrm{E}[X] + \mathrm{E}[Y] = 3.39 + 1.48 = 4.87 = \mathrm{E}[Z], \tag{101}$$
$$\mathrm{Var}[X] + 2\mathrm{Cov}[X,Y] + \mathrm{Var}[Y] = 1.9579 + 2(-0.2572) + 1.2296 = 2.5731 = \mathrm{Var}[Z]. \tag{102}$$

Ejercicio 18

Cierto supermercado tiene una caja de atención común y otra caja rápida. Supongamos que X es el número de clientes que están en espera en la caja común en un momento particular del día, y que Y es el número de clientes que están en espera en la caja rápida al mismo tiempo. La distribución de probabilidades conjunta de (X, Y) se resume en la siguente tabla:

$x \downarrow$ \ $y \rightarrow$	0	1	2	3
0	0.08	0.07	0.04	0.00
1	0.06	0.15	0.05	0.04
2	0.05	0.04	0.1	0.06
3	0.00	0.03	0.04	0.07
4	0.00	0.01	0.05	0.06

a) ¿Cuál es la probabilidad de que haya exactamente un cliente en cada línea de espera?

b) ¿Cuál es la probabilidad de que la cantidad de clientes en cada cola sea la misma?

c) Calcular la probabilidad de que haya por lo menos dos clientes más en una cola de espera que en la otra.

d) ¿Cuál es la probabilidad de que la cantidad de clientes en ambas colas sea exactamente 4?

e) Calcular el valor esperado del número de clientes en cada caja.

f) ¿Son X e Y variables aleatorias independientes?

Resolución

Variables Aleatorias

Las variables son las definidas en el enunciado:

- X = cantidad de clientes en la caja común.
- Y = cantidad de clientes en la caja rápida.

Datos

Sabemos la función de probabilidad conjunta $\mathrm{p}_{XY}(x, y)$.

Ítem a

La probabilidad de que haya un cliente en cada línea se corresponde con el evento $X = 1, Y = 1$:

$$\mathrm{P}(X = 1, Y = 1) = \mathrm{p}_{XY}(1, 1) = 0.15 \tag{103}$$

Ítem b

Que ambas filas tengan la misma cantidad de gente se traduce en el evento $X = Y$. Para calcular este evento, podemos usar el teorema de probabilidad total, basado en la partición dada por $R_Y = \{0, 1, 2, 3\}$:

$$\mathrm{P}(X = Y) = \sum_{j \in R_Y} \mathrm{P}(X = Y, Y = j) = \sum_{j=0}^{3} \mathrm{P}(X = j, Y = j) = 0.08 + 0.15 + 0.1 + 0.07 = 0.4 \tag{104}$$

Ítem c

Consideramos el siguiente evento:

$$A = \text{hay al menos dos personas más en una fila que en otra.} \tag{105}$$

Este evento se puede dar en dos circunstancias, en las que $X > Y$ o $X < Y$. Notar que no pueden ocurrir simultáneamente A y $X = Y$. Es decir,

$$\mathrm{P}(A) = \mathrm{P}(A \cap X > Y) + \mathrm{P}(A \cap X < Y) \tag{106}$$

Calcularemos ambos términos por separado:

$$\begin{aligned} \mathrm{P}(A \cap X > Y) =& \underbrace{\mathrm{P}(X = 2, Y = 0)}_{0.05} + \underbrace{\mathrm{P}(X = 3, Y = 0)}_{0} + \underbrace{\mathrm{P}(X = 4, Y = 0)}_{0} + \\ &+ \underbrace{\mathrm{P}(X = 3, Y = 1)}_{0.03} + \underbrace{\mathrm{P}(X = 4, Y = 1)}_{0.01} + \underbrace{\mathrm{P}(X = 4, Y = 2)}_{0.05} = 0.14 \end{aligned} \tag{107}$$

Por otro lado,

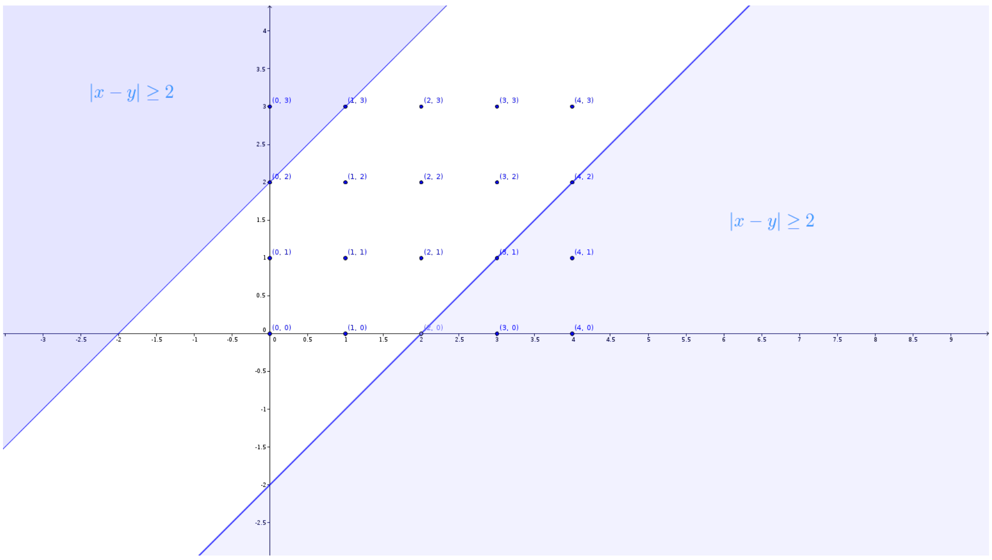
$$\mathrm{P}(A \cap X < Y) = \underbrace{\mathrm{P}(X = 0, Y = 2)}_{0.04} + \underbrace{\mathrm{P}(X = 0, Y = 3)}_{0} + \underbrace{\mathrm{P}(X = 1, Y = 3)}_{0.04} = 0.08 \tag{108}$$

Por lo tanto,

$$\mathrm{P}(A) = 0.14 + 0.08 = 0.22 \tag{109}$$

Notemos que tuvimos que contabilizar una gran cantidad de celdas de la tabla de contingencia y no es tan inmediato cuáles de estas celdas incluir en el cálculo. Otra forma de pensarlo es mediante un gráfico bidimensional que capture las características probabilísticas de ambas variables, nos puede facilitar cómo determinar las probabilidades a considerar.

Antes tenemos que pensar el evento A en términos de X e Y. Notemos que una diferencia entre ambos de dos o más personas, se traduce en el evento $|X - Y| \geq 2$. Por lo tanto, caracterizando los puntos de la tabla como coordenadas en el plano cartesiano, podemos identificar fácilmente los que pertenecen a la región:



Es decir, podemos determinar los puntos a considerar sólo mirando el gráfico:

- $(X,Y) = (0,2)$
- $(X,Y) = (0,3)$
- $(X,Y) = (1,3)$
- $(X,Y) = (2,0)$
- $(X,Y) = (3,0)$
- $(X,Y) = (4,0)$
- $(X,Y) = (3,1)$
- $(X,Y) = (4,1)$
- $(X,Y) = (4,2)$

Más aún, podíamos excluir de antemano los puntos de probabilidad nula del gráfico y obtener únicamente: Es decir, podemos determinar los puntos a considerar sólo mirando el gráfico:

- $(X,Y) = (0,2)$
- $(X,Y) = (1,3)$
- $(X,Y) = (2,0)$
- $(X,Y) = (3,1)$
- $(X,Y) = (4,1)$
- $(X,Y) = (4,2)$

Ítem d

Aquí nos piden la probabilidad de que entre ambas colas haya 4 clientes. Entonces, usando el teorema de probabilidad total:

$$\mathrm{P}(X+Y=4) = \sum_{j\in R_Y} \mathrm{P}(X+Y=4, Y=j) = \sum_{j=0}^{3} \mathrm{P}(X=4-j, Y=j) = \tag{110}$$

$$= \mathrm{P}(X=4,Y=0)+\mathrm{P}(X=3,Y=1)+\mathrm{P}(X=2,Y=2)+\mathrm{P}(X=1,Y=3) = 0+0.03+0.1+0.04 = 0.17 \tag{111}$$

Ítem e

Para calcular el valor medio de cada variable, podríamos pensar dos funciones $g(X,Y) = X$ y $h(X,Y) = Y$. Entonces,

- $\mathrm{E}[X] = \mathrm{E}(g(X,Y)) = \sum_{i\in R_X}\sum_{j\in R_Y} g(i,j)\cdot \mathrm{p}_{XY}(i,j) = \sum_{i\in R_X}\sum_{j\in R_Y} i\cdot \mathrm{p}_{XY}(i,j)$
- $\mathrm{E}[Y] = \mathrm{E}(h(X,Y)) = \sum_{i\in R_X}\sum_{j\in R_Y} h(i,j)\cdot \mathrm{p}_{XY}(i,j) = \sum_{i\in R_X}\sum_{j\in R_Y} j\cdot \mathrm{p}_{XY}(i,j)$

Notar que si bien esta opción es válida, debemos sumar sobre toda la tabla. Sin embargo, podemos utilizar la siguiente propiedad:

- $\mathrm{E}[X] = \sum_{i\in R_X} i\cdot \underbrace{\sum_{j\in R_Y} \mathrm{p}_{XY}(i,j)}_{\mathrm{p}_X(i)} = \sum_{i=0}^{4} i\cdot \mathrm{p}_X(i)$
- $\mathrm{E}[Y] = \sum_{j\in R_Y} j\cdot \underbrace{\sum_{i\in R_X} \mathrm{p}_{XY}(i,j)}_{\mathrm{p}_Y(j)} = \sum_{j=0}^{3} j\cdot \mathrm{p}_Y(j)$

Al ser variables discretas y de pocos valores, quizás resulta más sencillo utilizar las probabilidades marginales, ya que su cálculo se obtiene de sumar respecto de las filas o las columnas de la tabla, ubicando los resultados al margen (dándole el nombre de marginales):

$X \setminus Y$	0	1	2	3	$\mathrm{p}_X(x)$
0	0.08	0.07	0.04	0	0.19
1	0.06	0.15	0.05	0.04	0.3
2	0.05	0.04	0.1	0.06	0.25
3	0	0.03	0.04	0.07	0.14
4	0	0.01	0.05	0.06	0.12
$\mathrm{p}_Y(y)$	0.19	0.3	0.28	0.23	1

(112)

Por lo tanto,

- $\mathrm{E}\left[X\right] = \sum_{i=0}^{4} i \cdot \mathrm{p}_X(i) = 0 \cdot 0.19 + 1 \cdot 0.3 + 2 \cdot 0.25 + 3 \cdot 0.14 + 4 \cdot 0.12 = 1.7$
- $\mathrm{E}\left[Y\right] = \sum_{j=0}^{3} j \cdot \mathrm{p}_Y(j) = 0 \cdot 0.19 + 1 \cdot 0.3 + 2 \cdot 0.28 + 3 \cdot 0.23 = 1.55$

Ítem f

Si X e Y son independientes, $\mathrm{p}_{X,Y}(x, y) = \mathrm{p}_X(x) \cdot \mathrm{p}_Y(y)$. Esto debe cumplirse para todo valor de x e y.

En casos donde la tabla de probabilidad conjunta tiene valores nulos, podemos sospechar que ambas variables no son independientes, ya que para que lo sean, debería cumplirse que alguno de los valores tenga una probabilidad marginal nula. Esto nos diría que dicho valor no se encuentra en el recorrido.

Más concretamente, como $\mathrm{p}_{X,Y}(4, 0) = 0$, para que sean independientes, debería suceder que $\mathrm{p}_X(4) = 0$ o $\mathrm{p}_Y(0) = 0$. Sin embargo, como $\mathrm{p}_X(4) = 0.12$ o $\mathrm{p}_Y(0) = 0.19$, entonces:

$$\mathrm{p}_{X,Y}(4, 0) = 0 \neq 0.12 \cdot 0.19 = \mathrm{p}_X(4) \cdot \mathrm{p}_Y(0) \tag{113}$$

Por lo tanto, X e Y no son independientes.

Ejercicio 22

Las variables aleatorias discretas X e Y son independientes y cada una tiene distribución binomial de parámetros $n = 3$ y $p = 0.5$

a) Calcular $\mathrm{P}(2X + 3Y < 10)$ y $\mathrm{P}(X = Y)$.

b) Obtener la distribución de probabilidades de $Z = X + Y$. Verificar que Z tiene distribución binomial e identificar sus parámetros. Calcular $\mathrm{E}[Z]$ y $\mathrm{Var}[Z]$.

Resolución

Variables aleatorias y datos

En este caso las variables aleatorias X e Y no representan ningún escenario particular. Lo que sabemos de X e Y es que son independientes y sus distribuciones $(X, Y \sim \mathrm{Bi}(3, p))$.

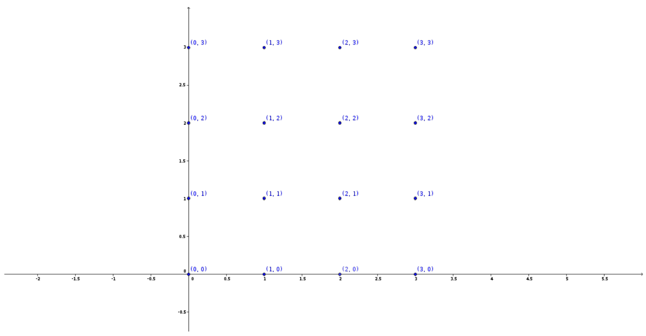
Ítem a

Con estos datos puedo obtener la distribución de probabilidad conjunta de X, Y. Como son independientes, se obtiene:

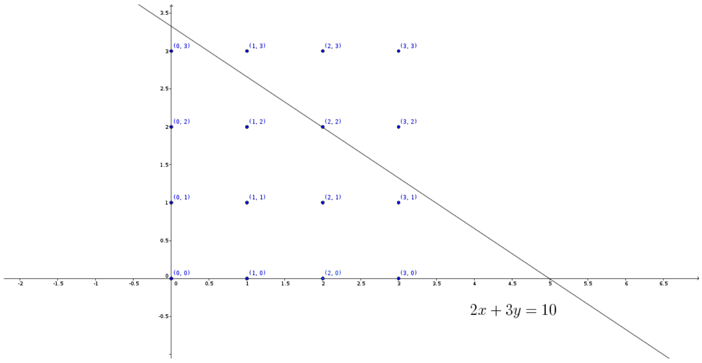
$$\begin{aligned}\mathrm{p}_{XY}(i,j) = \mathrm{P}(X = i, Y = j) &\overset{\text{IND}}{=} \mathrm{P}(X = i) \cdot \mathrm{P}(Y = j) = \binom{3}{i} 0.5^i 0.5^{3-i} \cdot \binom{3}{j} 0.5^j 0.5^{3-j} \\ &= \binom{3}{i}\binom{3}{j} 0.5^6 \end{aligned} \tag{114}$$

Notar que esto tiene sentido sólo cuando $0 \leq i, j \leq 3$, de otro modo $\mathrm{p}_{XY}(i,j) = 0$.

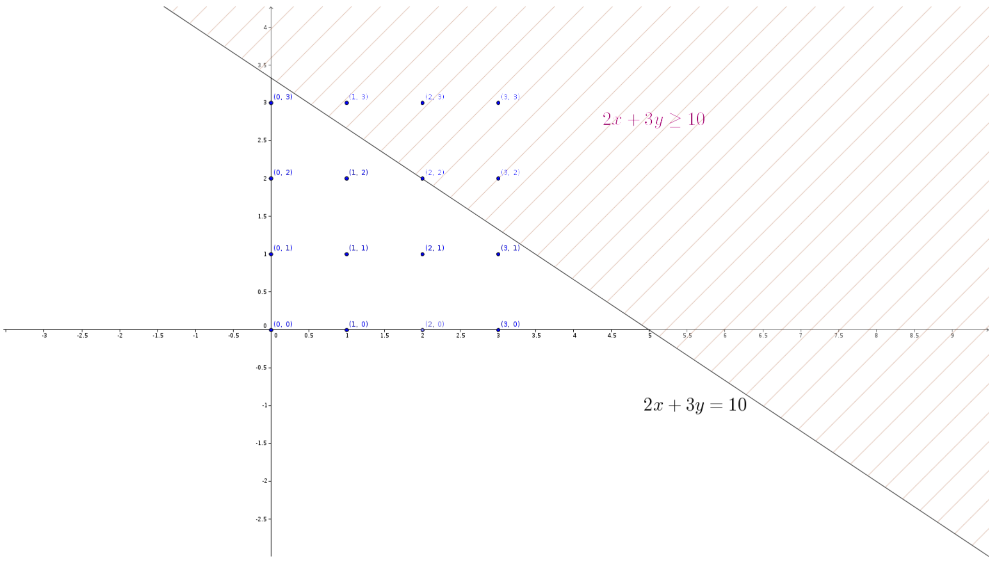
Podemos graficar estos puntos en un plano cartesiano para observar qué puntos debemos tener en cuenta para calcular la probabilidad de cualquier evento representando la región correspondiente:



Como me piden $\mathrm{P}(2X + 3Y < 10)$, puedo graficar la región correspondiente y ver qué puntos debo considerar.



Observo que el semiplano $2x + 3y < 10$ contiene 11 puntos y el semiplano $2x + 3y \geq 10$ contiene sólo 5. Por lo tanto, me conviene calcular $\mathrm{P}(2X + 3Y \geq 10)$ y luego utilizar la regla del complemento.



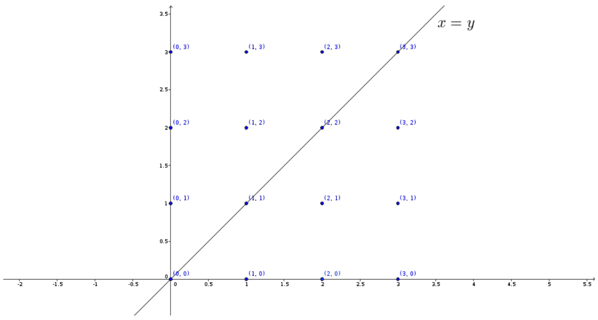
Por lo tanto, se obtiene:

$$\begin{aligned}
\mathrm{P}(2X + 3Y < 10) &= 1 - \mathrm{P}(2X + 3Y \geq 10) \\
&= 1 - \mathrm{p}_{XY}(1,3) - \mathrm{p}_{XY}(2,3) - \mathrm{p}_{XY}(3,3) - \mathrm{p}_{XY}(2,2) - \mathrm{p}_{XY}(2,3) \\
&= 1 - \binom{3}{1}\binom{3}{3}0.5^6 - \binom{3}{2}\binom{3}{3}0.5^6 - \binom{3}{3}\binom{3}{3}0.5^6 - \\
&\qquad - \binom{3}{2}\binom{3}{2}0.5^6 - \binom{3}{3}\binom{3}{2}0.5^6 \\
&= 1 - 0.5^6 \cdot (3 + 3 + 1 + 9 + 3) = 1 - 19 \cdot 0.5^6 = 0.703125 = \frac{45}{64}
\end{aligned} \tag{115}$$

Por otro lado, para calcular $\mathrm{P}(X = Y)$, claramente hay que sumar los siguientes términos:

$$\begin{aligned}
\mathrm{P}(X = Y) &= \mathrm{p}_{XY}(0,0) + \mathrm{p}_{XY}(1,1) + \mathrm{p}_{XY}(2,2) + \mathrm{p}_{XY}(3,3) \\
&= \binom{3}{0}^2 0.5^6 + \binom{3}{1}^2 0.5^6 + \binom{3}{2}^2 0.5^6 + \binom{3}{3}^2 0.5^6 \\
&= 20 \cdot 0.5^6 = \frac{5}{16} = 0.3125
\end{aligned} \tag{116}$$

Los puntos incluidos en el cálculo coinciden con lo que se puede observar gráficamente:



Ítem b

Ahora tenemos la siguiente variable:

$$Z = g(X, Y) = X + Y \tag{117}$$

Como $R_X = \{0, 1, 2, 3\}$ y $R_Y = \{0, 1, 2, 3\}$, considerando todos los resultados de las sumas posibles se tiene:

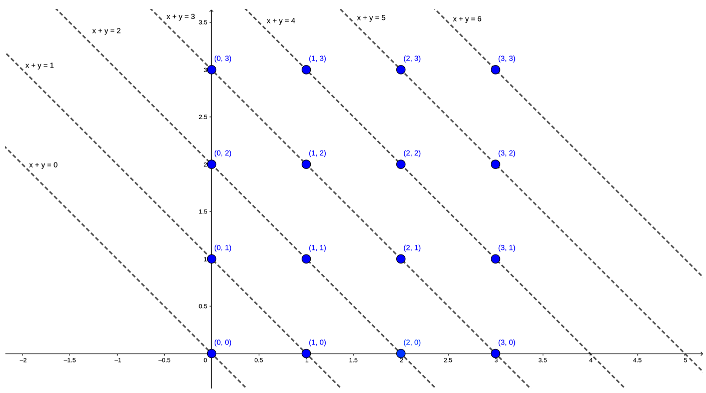
$$R_Z = \{0, 1, 2, 3, 4, 5, 6\} \tag{118}$$

Nos piden verificar que la distribución de Z es binomial. Viendo el recorrido, ya podemos conjeturar al número 6 el primer parámetro. Resta ver la probabilidad de éxito.

Intuitivamente, tanto X como Y cuentan la cantidad de éxitos en 3 intentos independientes con una **misma** probabilidad de éxito $p = 0.5$. Como a su vez cada grupo de tres intentos son independientes, podemos decir que $Z = X + Y$ cuenta la cantidad de éxitos en 6 intentos independientes con una probabilidad de éxito $p = 0.5$.

Sin embargo este razonamiento no demuestra que $Z \sim \text{Bi}(6; 0.5)$ y debemos demostrarlo calculando las probabilidades puntuales y viendo que coinciden con las de dicha binomial.

Podemos calcular la probabilidad de que $Z = k$ graficando la recta $x + y = k$ y sumando sobre los puntos de probabilidad positiva que están sobre la recta. Haciendo variar k en los valores de R_Z se tiene:



Es decir,

- Para $k = 0$:

$$\begin{aligned} \text{p}_Z(0) &= \text{P}(Z = 0) = \text{P}(X + Y = 0) = \text{p}_{XY}(0, 0) \\ &= \binom{3}{0}\binom{3}{0} 0.5^6 = 0.5^6 = \frac{1}{64} \end{aligned} \tag{119}$$

- Para $k = 1$:

$$\begin{aligned} \mathrm{p}_Z(1) &= \mathrm{P}(Z=1) = \mathrm{P}(X+Y=1) = \mathrm{p}_{XY}(1,0) + \mathrm{p}_{XY}(0,1) = \\ &= \binom{3}{1}\binom{3}{0} 0.5^6 + \binom{3}{0}\binom{3}{1} 0.5^6 = 6 \cdot 0.5^6 = \frac{3}{32} \end{aligned} \tag{120}$$

- Para $k = 2$:

$$\begin{aligned} \mathrm{p}_Z(2) &= \mathrm{P}(Z=2) = \mathrm{P}(X+Y=2) = \mathrm{p}_{XY}(2,0) + \mathrm{p}_{XY}(1,1) + \mathrm{p}_{XY}(0,2) \\ &= \binom{3}{2}\binom{3}{0} 0.5^6 + \binom{3}{1}\binom{3}{1} 0.5^6 + \binom{3}{0}\binom{3}{2} 0.5^6 \\ &= 15 \cdot 0.5^6 = \frac{15}{64} \end{aligned} \tag{121}$$

- Para $k = 3$:

$$\begin{aligned} \mathrm{p}_Z(3) &= \mathrm{P}(Z=3) = \mathrm{P}(X+Y=3) = \mathrm{p}_{XY}(3,0) + \mathrm{p}_{XY}(2,1) + \mathrm{p}_{XY}(1,2) + \mathrm{p}_{XY}(0,3) \\ &= \binom{3}{3}\binom{3}{0} 0.5^6 + \binom{3}{2}\binom{3}{1} 0.5^6 + \binom{3}{1}\binom{3}{2} 0.5^6 + \binom{3}{0}\binom{3}{3} 0.5^6 \\ &= 20 \cdot 0.5^6 = \frac{5}{16} \end{aligned} \tag{122}$$

- Para $k = 4$:

$$\begin{aligned} \mathrm{p}_Z(4) &= \mathrm{P}(Z=4) = \mathrm{P}(X+Y=4) = \mathrm{p}_{XY}(3,1) + \mathrm{p}_{XY}(2,2) + \mathrm{p}_{XY}(1,3) \\ &= \binom{3}{3}\binom{3}{1} 0.5^6 + \binom{3}{2}\binom{3}{2} 0.5^6 + \binom{3}{1}\binom{3}{3} 0.5^6 \\ &= 15 \cdot 0.5^6 = \frac{15}{64} \end{aligned} \tag{123}$$

- Para $k = 5$:

$$\begin{aligned} \mathrm{p}_Z(5) &= \mathrm{P}(Z=5) = \mathrm{P}(X+Y=5) = \mathrm{p}_{XY}(2,3) + \mathrm{p}_{XY}(3,2) = \\ &= \binom{3}{2}\binom{3}{3} 0.5^6 + \binom{3}{3}\binom{3}{2} 0.5^6 = 6 \cdot 0.5^6 = \frac{3}{32} \end{aligned} \tag{124}$$

- Para $k = 6$:

$$\begin{aligned} \mathrm{p}_Z(6) &= \mathrm{P}(Z=6) = \mathrm{P}(X+Y=6) = \mathrm{p}_{XY}(3,3) = \\ &= \binom{3}{3}\binom{3}{3} 0.5^6 = 0.5^6 = \frac{1}{64} \end{aligned} \tag{125}$$

Notemos que si se calculan las probabilidades de una binomial de parámetros 6 y 0.5, coinciden con las obtenidas. Por ejemplo,

$$\mathrm{p}_Z(3) = \binom{6}{3} 0.5^6 = \frac{20}{64} = \frac{5}{16} \tag{126}$$

Por otro lado, notemos además que las probabilidades son simétricas respecto de $k = 3$. Es decir,

- $\mathrm{p}_Z(0) = \mathrm{p}_Z(6) = \frac{1}{64}$
- $\mathrm{p}_Z(1) = \mathrm{p}_Z(5) = \frac{3}{32}$
- $\mathrm{p}_Z(2) = \mathrm{p}_Z(4) = \frac{15}{64}$

Esto se debe a que la $p = 1 - p = 0.5$, por lo que las potencias de la binomial resultan siempre idénticas. En cuanto a los combinatorios, recordar que:

- $\binom{3}{0} = \binom{3}{3} = 1$
- $\binom{3}{1} = \binom{3}{2} = 3$

Notar que hay además que siempre que en una probabilidad puntual (por ejemplo, $\mathrm{p}_Z(1)$) aparece uno de estos combinatorios (por ejemplo, $\binom{3}{1}$), en la probabilidad puntual simétrica (en este caso, $\mathrm{p}_Z(5)$) aparece el combinatorio de mismo valor (es decir, $\binom{3}{2}$)

Ejercicio 24

La función de densidad de probabilidad conjunta de (X, Y) es

$$\mathrm{f}_{X,Y}(x,y) = \begin{cases} a(x+y) & 0 < x < 3,\ 0 < y < 3 \\ 0 & x \notin (0,3) \times (0,3) \end{cases}$$

1. Calcular $\mathrm{P}(1 < X < 2, 1 < Y < 2)$
2. Calcular $\mathrm{E}[X]$, $\mathrm{E}[Y]$, σ_X y σ_Y.

Resolución

Primero debemos calcular la constante a. Para ello, notemos que

$$\begin{aligned} 1 &= \iint_{\mathbb{R}^2} \mathrm{f}_{X,Y}(x,y)\,dxdy = \int_0^3 \int_0^3 a(x+y)\,dxdy = \int_0^3 a\left(\frac{x^2}{2} + xy\right)\Big|_{x=0}^{x=3} dy \\ &= \int_0^3 a\left(\frac{9}{2} + 3y\right) dy = a\left(\frac{9}{2}y + \frac{3}{2}y^2\right)\Big|_{y=0}^{y=3} \\ &= 27a \Rightarrow a = \frac{1}{27}. \end{aligned}$$

La probabilidad solicitada se puede calcular como:

$$\begin{aligned} \mathrm{P}(1 < X < 2, 1 < Y < 2) &= \int_1^2 \int_1^2 \frac{1}{27}(x+y)\,dxdy = \int_1^2 \frac{1}{27}\left(\frac{x^2}{2} + xy\right)\Big|_{x=1}^{x=2} dy \\ &= \int_1^2 \frac{1}{27}\left(\frac{3}{2} + y\right) dy = \frac{1}{27}\left(\frac{3}{2}y + \frac{1}{2}y^2\right)\Big|_{y=1}^{y=2} \\ &= \frac{1}{9}. \end{aligned}$$

También podemos calcular

$$\begin{aligned} \mathrm{E}[XY] &= \int_0^3 \int_0^3 xy\frac{1}{27}(x+y)\,dxdy = \int_0^3 \frac{1}{27}\left(\frac{x^3}{3}y + \frac{1}{2}x^2y^2\right)\Big|_{x=0}^{x=3} dy \\ &= \int_0^3 \frac{1}{27}\left(9y + \frac{9}{2}y^2\right) dy = \frac{1}{27}\left(\frac{9}{2}y^2 + \frac{1}{6}y^3\right)\Big|_{y=0}^{y=3} \\ &= \frac{5}{3}. \end{aligned}$$

Densidades marginales:
Las densidad marginal $\mathrm{f}_X(x)$ para $x \in (0,3)$ es:

$$\begin{aligned} \mathrm{f}_X(x) &= \int_{\mathbb{R}} \mathrm{f}_{X,Y}(x,y)\,dy = \int_0^3 \frac{1}{27}(x+y)\,dy = \frac{1}{27}\left(xy + \frac{1}{2}y^2\right)\Big|_{y=0}^{y=3} \\ &= \frac{1}{6} + \frac{x}{9}. \end{aligned}$$

Por tanto, tenemos que

$$\mathrm{f}_X(x) = \begin{cases} \frac{1}{6} + \frac{x}{9} & x \in (0,3) \\ 0 & x \notin (0,3) \end{cases} \tag{127}$$

Es fácil ver que Y tiene la misma densidad que X. Es decir, X e Y están igualmente distribuidas. Por otro lado,

$$\mathrm{f}_{X,Y}(1,1) = \frac{1}{27}(1+1) = \frac{2}{27} \neq \left(\frac{1}{6} + \frac{1}{9}\right)\left(\frac{1}{6} + \frac{1}{9}\right) = \mathrm{f}_X(1)\mathrm{f}_Y(1), \tag{128}$$

por lo que X e Y no son independientes.

Los valores esperados son:

$$\begin{aligned}
\mathrm{E}\left[Y\right] = \mathrm{E}\left[X\right] &= \int_0^3 x\left(\frac{1}{6} + \frac{x}{9}\right)\,dx = \left.\left(\frac{x^2}{12} + \frac{x^3}{27}\right)\right|_{x=0}^{x=3} \\
&= \frac{7}{4}, \\
\mathrm{E}\left[Y^2\right] = \mathrm{E}\left[X^2\right] &= \int_0^3 x^2\left(\frac{1}{6} + \frac{x}{9}\right)\,dx = \left.\left(\frac{x^3}{18} + \frac{x^4}{36}\right)\right|_{x=0}^{x=3} \\
&= \frac{15}{4}, \\
\mathrm{Var}\left[Y\right] = \mathrm{Var}\left[X\right] &= \frac{15}{4} - \left(\frac{7}{4}\right)^2 = \frac{11}{16} \Rightarrow \sigma_Y = \sigma_X \approx 0.8292.
\end{aligned}$$

Con estos resultados, podemos calcular la covarianza de X e Y:

$$\mathrm{Cov}\left[X,Y\right] = \mathrm{E}\left[XY\right] - \mathrm{E}\left[X\right]\mathrm{E}\left[Y\right] = \frac{5}{3} - \frac{7}{4}\cdot\frac{7}{4} = -\frac{67}{48}. \tag{129}$$

Dado que la covarianza es distinta de cero, podemos afirmar que las variables aleatorias no son independientes.

Ejercicio 30

Dos personas han quedado citadas en un lugar. El instante de llegada de cada una a ese lugar puede suponerse una variable aleatoria con distribución uniforme en (0, 30) y que esos instantes de llegada son independientes.

a) Obtener la función de distribución y de densidad de probabilidad del tiempo T que una persona espera a la otra.

b) Calcular el valor esperado y el desvío estándar de T.

Resolución

Variables aleatorias

Consideramos las siguientes variables aleatorias:

- X = tiempo de llegada de la persona A.
- Y = tiempo de llegada de la persona B.
- T = tiempo de espera entre ambos.

Datos

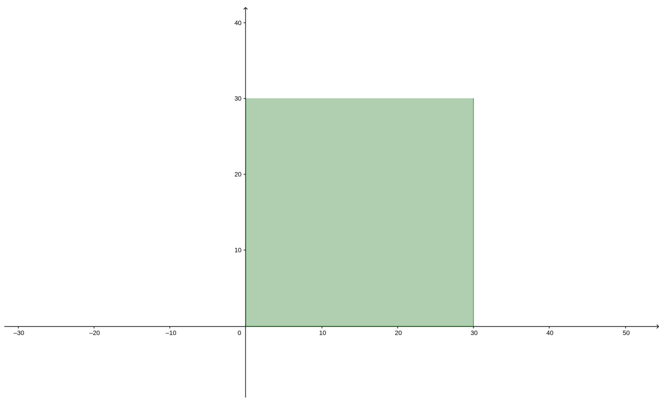
Sabemos que $X, Y \sim \mathcal{U}(0; 30)$, por lo tanto:

$$\mathrm{f}_X(x) = \begin{cases} \frac{1}{30} & \text{si } 0 < x < 30 \\ 0 & \text{en otro caso} \end{cases} \wedge \mathrm{f}_Y(y) = \begin{cases} \frac{1}{30} & \text{si } 0 < y < 30 \\ 0 & \text{en otro caso} \end{cases} \tag{130}$$

Al ser independientes,

$$\mathrm{f}_{X,Y}(x, y) = \mathrm{f}_X(x)\mathrm{f}_Y(y) = \begin{cases} \frac{1}{900} & \text{si } 0 < x, y < 30 \\ 0 & \text{en otro caso} \end{cases} \tag{131}$$

Por lo tanto, todas las consideraciones sobre las probabilidades que dependen de X e Y hay que evaluarlas en este cuadrado:



Ítem a

Debemos calcular $\mathrm{F}_T(t) = \mathrm{P}(T \leq t)$. Claramente, no se pueden esperar menos de 0 minutos y con probabilidad 1 se esperan menos de 30 minutos, por lo que $\mathrm{F}_T(t) = 0$ si $t < 0$ y $\mathrm{F}_T(t) = 1$ si $t > 30$.

Tomaremos el caso más interesante de $t \in (0, 30)$. Notemos que por probabilidad total:

$$\mathrm{F}_T(t) = \mathrm{P}(T \leq t) = \mathrm{P}(T \leq t, X \leq Y) + \mathrm{P}(T \leq t, X > Y) \tag{132}$$

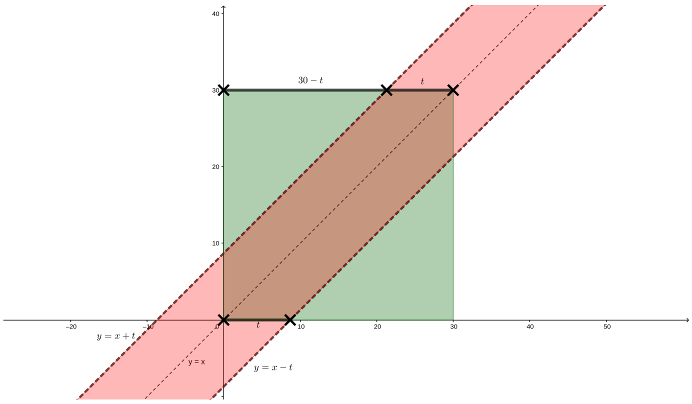
Por otro lado, si $X \leq Y$ significa que la persona A llegó antes que la persona B, por lo que $T = Y - X$. Del mismo modo, si $X > Y \Rightarrow T = X - Y$. Es decir,

$$\mathrm{P}(T \leq t, X \leq Y) + \mathrm{P}(T \leq t, X > Y) = \mathrm{P}(Y - X \leq t, X \leq Y) + \mathrm{P}(X - Y \leq t, X > Y) \tag{133}$$

Con algunos despejes, se tiene:

$$\mathrm{F}_T(t) = \mathrm{P}(X \leq Y \leq X + t) + \mathrm{P}(X - t \leq Y < X) \tag{134}$$

Para calcular estas probabilidades, recurriremos a graficar la región:



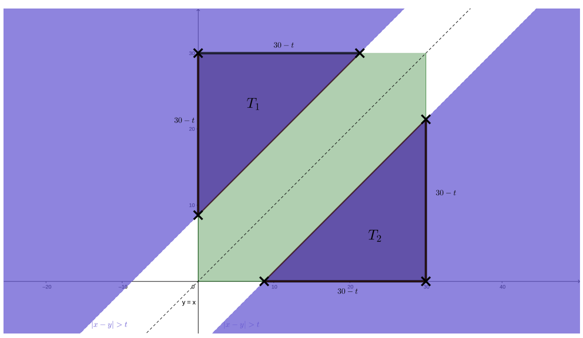
Por lo tanto, como cambian los límites de integración,

$$\mathrm{F}_T(t) = \int_0^t \int_0^{x+t} \mathrm{f}_{X,Y}(x,y)\, dydx + \int_t^{30-t} \int_{x-t}^{x+t} \mathrm{f}_{X,Y}(x,y)\, dydx + \int_{30-t}^{30} \int_{x-t}^{30} \mathrm{f}_{X,Y}(x,y)\, dydx \tag{135}$$

Esto se puede resolver de este modo, sin embargo, hay una forma más sencilla. Por empezar, notemos que dividir entre los casos $X \leq Y$ y $X > Y$ se puede simplificar considerando el tiempo de espera $T = |X - Y|$. Además, podemos calcular $\mathrm{F}_T(t)$ usando el complemento:

$$\mathrm{F}_T(t) = 1 - \mathrm{P}(T > t) = 1 - \mathrm{P}(|X - Y| > t) \tag{136}$$

Esta última expresión da lugar al siguiente gráfico:



Por lo tanto, podemos integrar la densidad sobre las regiones T_1 y T_2:

$$\begin{aligned} \mathrm{P}(T > t) &= \int\int_{T_1} \mathrm{f}_{X,Y}(x,y)\,dydx + \int\int_{T_2} \mathrm{f}_{X,Y}(x,y)\,dydx \\ &= \int\int_{T_1} \frac{1}{900}\,dydx + \int\int_{T_2} \frac{1}{900}\,dydx \end{aligned} \tag{137}$$

Es decir,

$$\begin{aligned} \mathrm{P}(T > t) &= \frac{\int\int_{T_1} 1\,dydx}{900} + \frac{\int\int_{T_2} 1\,dydx}{900} = \frac{\text{Área}(T_1)}{900} + \frac{\text{Área}(T_2)}{900} = \frac{2\text{Área}(T_1)}{900} \\ &= \frac{(30-t)^2}{30^2} = \left(1 - \frac{t}{30}\right)^2 \end{aligned} \tag{138}$$

Por último,

$$\mathrm{F}_T(t) = 1 - \mathrm{P}(T > t) = \begin{cases} 0 & \text{si } t < 0 \\ 1 - \left(1 - \frac{t}{30}\right)^2 & \text{si } 0 \leq t \leq 30 \\ 1 & \text{si } t > 30 \end{cases} \tag{139}$$

Además, nos piden la densidad de probabilidad:

$$\mathrm{f}_T(t) = \frac{d}{dt}[\mathrm{F}_T(t)] = \begin{cases} 2\left(1 - \frac{t}{30}\right)\cdot\frac{1}{30} & \text{si } 0 \leq t \leq 30 \\ 0 & \text{caso contrario} \end{cases} = \begin{cases} \frac{1}{15} - \frac{t}{450} & \text{si } 0 \leq t \leq 30 \\ 0 & \text{caso contrario} \end{cases} \tag{140}$$

Ítem b

Ahora nos piden el valor esperado y el desvío de T. En cuanto al planteo, es inmediato al tener la densidad:

$$\mathrm{E}\left[T\right] = \int_0^{30} t \cdot \mathrm{f}_T(t)\,dt = \int_0^{30} \frac{t}{15} - \frac{t^2}{450}\,dt = \frac{30^2}{30} - \frac{30^3}{1350} = 30 - 20 = 10 \tag{141}$$

Recordando que el desvío cumple $\sigma(T) = \sqrt{\mathrm{E}\left[T^2\right] - \left[\mathrm{E}\left[T\right]\right]^2} = \sqrt{\mathrm{E}(T^2) - 10^2}$, debemos calcular $\mathrm{E}\left[T^2\right]$:

$$\mathrm{E}\left[T^2\right] = \int_0^{30} t^2 \cdot \mathrm{f}_T(t)\,dt = \int_0^{30} \frac{t^2}{15} - \frac{t^3}{450}\,dt = \frac{30^3}{45} - \frac{30^4}{1800} = 600 - 450 = 150 \tag{142}$$

De este modo, el desvío es $\sigma(T) = \sqrt{150 - 100} = \sqrt{50}$