Vector aleatorio

3. Se debe seleccionar un comité de tres personas elegidas al azar de un grupo constituido por cuatro docentes y cinco estudiantes. Sea X_1: número de docentes en el comité y X_2: número de estudiantes en el comité.

a) Determine la distribución de probabilidad conjunta de X_1 y X_2.
b) Determine las distribuciones marginales de X_1 y X_2.
c) ¿Son X_1 y X_2 independientes?
d) Calcule $P(X_1 = 1/X_2 \geq 1)$.

X_1:"número de docentes en el comité"
X_2:"número de estudiantes en el comité"
Primero calcularemos las distribuciones marginales y luego la de probabilidad conjunta.

b) Las distribuciones marginales de X_1 y X_2 son hipergeométricas (¿por qué?), es decir
$X_1 \sim H(9,4,3)$

$$p_{X_1}(x_1) = \frac{\binom{4}{x_1}\binom{5}{3-x_1}}{\binom{9}{3}} \qquad x_1 = \overline{0,3}$$

$X_2 \sim H(9,5,3)$

$$p_{X_2}(x_2) = \frac{\binom{5}{x_2}\binom{4}{3-x_2}}{\binom{9}{3}} \qquad x_2 = \overline{0,3}$$

Luego

$$p_{X_1(0)} = \frac{\binom{4}{0}\binom{5}{3}}{\binom{9}{3}} = \frac{1 \cdot 10}{84} = \frac{5}{42}$$
$$p_{X_1(1)} = \frac{\binom{4}{1}\binom{5}{2}}{84} = \frac{4 \cdot 10}{84} = \frac{10}{21}$$
$$p_{X_1(2)} = \frac{\binom{4}{2}\binom{5}{1}}{84} = \frac{6 \cdot 5}{84} = \frac{5}{14}$$
$$p_{X_1(3)} = \frac{\binom{4}{3}\binom{5}{0}}{84} = \frac{4 \cdot 1}{84} = \frac{1}{21}$$

$$p_{X_2(0)} = \frac{\binom{5}{0}\binom{4}{3}}{84} = \frac{1 \cdot 4}{84} = \frac{1}{21}$$
$$p_{X_2(1)} = \frac{\binom{5}{1}\binom{4}{2}}{84} = \frac{5 \cdot 6}{84} = \frac{5}{14}$$
$$p_{X_2(2)} = \frac{\binom{5}{2}\binom{4}{1}}{84} = \frac{10 \cdot 4}{84} = \frac{10}{21}$$
$$p_{X_2(3)} = \frac{\binom{5}{3}\binom{4}{0}}{84} = \frac{1 \cdot 10}{84} = \frac{5}{42}$$

a)
Sabemos que $X_1+X_2=3$, por lo tanto la distribución de probabilidad conjunta
$p(x_1, x_2) = 0 \ \forall\, x_1, x_2 /\ x_1 + x_2 \neq 3$

Armamos una tabla completando con 0 donde $x_1+x_2 \neq 3$ y colocando las distribuciones marginales en la última columna y fila. Teniendo en cuenta la definición de probabilidad marginal, podemos completar las celdas faltantes de la tabla.

$x_2 \backslash x_1$	0	1	2	3	$p_{X_2}(x_2)$
0	0	0	0	$\frac{1}{21}$	$\frac{1}{21}$
1	0	0	$\frac{5}{14}$	0	$\frac{5}{14}$
2	0	$\frac{10}{21}$	0	0	$\frac{10}{21}$
3	$\frac{5}{42}$	0	0	0	$\frac{5}{42}$
$p_{X_1}(x_1)$	$\frac{5}{42}$	$\frac{10}{21}$	$\frac{5}{14}$	$\frac{1}{21}$	1

Finalmente, podemos expresar la distribución de probabilidad conjunta p(x_1,x_2) como
$p(x_1, x_2) = 0 \;\forall\; x_1, x_2 /\; x_1 + x_2 \neq 3$

$p(3{,}0) = \frac{1}{21} \qquad p(2{,}1) = \frac{5}{14} \qquad p(1{,}2) = \frac{10}{21} \qquad p(0{,}3) = \frac{5}{42}$

c)Para que X_1 y X_2 sean independientes, debe cumplirse alguna de las tres condiciones vistas en la teoría para todos los valores posibles de X_1 y X_2.

$$\left.\begin{aligned} p(3{,}0) &= \tfrac{1}{21} \\ p_{X_1}(3)\, p_{X_2}(0) &= \tfrac{1}{21}\cdot\tfrac{1}{21} = \tfrac{1}{441} \end{aligned}\right\} \Rightarrow p(3{,}0) \neq p_{X_1}(3)\, p_{X_2}(0) \Rightarrow X_1 \text{ y } X_2 \text{ no son independientes}$$

Podría haberse probado la no independencia evaluando esta condición con cualquier otro valor de X_1 y X_2.

d)

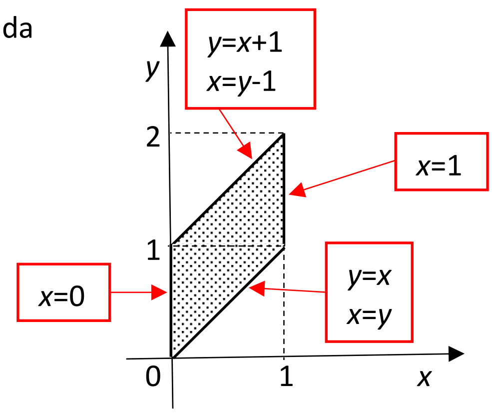
$$P(X_1 = 1 /\; X_2 \geq 1) = \frac{P(X_1 = 1 \;\cap\; X_2 \geq 1)}{P(X_2 \geq 1)} = \frac{p(1{,}1) + p(1{,}2) + p(1{,}3)}{p_{X_2}(1) + p_{X_2}(2) + p_{X_2}(3)} = \frac{1}{2}$$

5. Sea f la función densidad de probabilidad conjunta del vector aleatorio (X, Y), dada por

$$f(x, y) = \begin{cases} k & \text{si } 0 < x < 1,\; x < y < x + 1 \\ 0 & \text{en otro caso} \end{cases}$$

a) Determine el valor de k.
b) Obtenga las distribuciones marginales de X e Y respectivamente.
c) Calcule el coeficiente de correlación entre X e Y.

En el gráfico de la derecha, el área sombreada representa los valores posibles de x e y. También se muestran las ecuaciones de las rectas que limitan el área sombreada. Este gráfico de coordenadas es de mucha utilidad en el momento de resolver las integrales y determinar los extremos de integración.



a) Determine el valor de *k*.

Para que f(*x*,*y*) sea una función de densidad de probabilidad conjunta, debe cumplir la condición de cierre.

$$\int_0^1 \int_x^{x+1} \mathrm{f}(x,y)\ dy\ dx = \int_0^1 \int_x^{x+1} k\ dy\ dx = k \int_0^1 \left(y|_x^{x+1}\right)\ dx = k \int_0^1 dx = k \implies k = 1$$

Notar que la integral externa se evalúa sobre la variable *x*, es decir, en "sentido horizontal" si nos referimos al gráfico de coordenadas. *x* varía entre 0 y 1.
La integral interna se evalúa sobre la variable *y*, es decir, en "sentido vertical" si nos referimos al gráfico de coordenadas. *y* varía entre las rectas *y*=*x* e *y*=*x*+1.

b) Obtenga las distribuciones marginales de X e Y respectivamente.
Para calcular la distribución marginal $\mathrm{f_X}(x)$ se debe integrar sobre todos los valores de *y*, es decir, en "sentido vertical" si nos referimos al gráfico de coordenadas.

$$\mathrm{f_X}(x) = \int_{-\infty}^{+\infty} \mathrm{f}(x,y)dy = \int_x^{x+1} dy = 1$$

Finalmente

$$\mathrm{f_X}(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{en otro caso} \end{cases}$$

Para calcular la distribución marginal $\mathrm{f_Y}(y)$ se debe integrar sobre todos los valores de *x*, es decir, en "sentido horizontal" si nos referimos al gráfico de coordenadas. Por lo tanto, debemos separar el cálculo en dos partes.
Si $0 < y < 1 \Rightarrow 0 < x < y$

$$\mathrm{f_Y}(y) = \int_0^y dx = y$$

Si $1 < y < 2 \Rightarrow y - 1 < x < 1$

$$\mathrm{f_Y}(y) = \int_{y-1}^1 dx = 2 - y$$

Finalmente

$$\mathrm{f_Y}(y) = \begin{cases} y & 0 < y < 1 \\ 2 - y & 1 < y < 2 \\ 0 & \text{en otro caso} \end{cases}$$

c) Calcule el coeficiente de correlación ρ_{XY}.

$$\rho_{\mathrm{XY}} = \frac{\mathrm{cov(X,Y)}}{(\mathrm{V(X)\ V(Y)})^{1/2}} \qquad \mathrm{cov(X,Y) = E(X\ Y) - E(X)\ E(Y)}$$

$$\mathrm{E(X)} = \int_{-\infty}^{+\infty} x\ \mathrm{f_X}(x)\ dx = \int_{0}^{1} x\ dx = \left.\frac{x^2}{2}\right|_0^1 = \frac{1}{2}$$

$$\mathrm{E(Y)} = \int_{0}^{1} y\ y\ dy + \int_{1}^{2} y\ (2-y)\ dy = \left.\frac{y^3}{3}\right|_0^1 + \left.\left(y^2 - \frac{y^3}{3}\right)\right|_1^2 = \frac{1}{3} + 4 - \frac{8}{3} - 1 + \frac{1}{3} = 1$$

$$\mathrm{E(X\ Y)} = \iint_{\mathrm{R_{X,Y}}} x\ y\ \mathrm{f}(x,y)\ dx\ dy = \int_{0}^{1}\int_{x}^{x+1} x\ y\ dy\ dx = \int_{0}^{1} x \int_{x}^{x+1} y\ dy\ dx$$

Calculamos la integral interna

$$\int_{x}^{x+1} y\ dy = \left[\frac{y^2}{2}\right]_x^{x+1} = \frac{((x+1)^2 - x^2)}{2} = \frac{1}{2}\left(x^2 + 2x + 1 - x^2\right) = \frac{2x+1}{2}$$

Luego calculamos la integral externa

$$\int_{0}^{1} \frac{x}{2}\ (2x+1)\ dx = \int_0^1 \frac{1}{2}\ (2x^2+x)\ dx = \frac{1}{2}\left[2\frac{x^3}{3}\right]_0^1 + \left[\frac{x^2}{2}\right]_0^1 = \frac{1}{2}\left(\frac{2}{3} + \frac{1}{2}\right) = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$$

Luego: E(X Y)=7/12 cov(X,Y)=7/12-1/2=1/12

Para calcular ρ_{XY} nos resta determinar V(X) y V(Y)

$$\mathrm{E(X^2)} = \int_0^1 x^2\ dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$

$$\mathrm{E(Y^2)} = \int_0^1 y^2\ y\ dy + \int_1^2 y^2\ (2-y)\ dy = \left[\frac{y^4}{4}\right]_0^1 + \left[2\frac{y^3}{3}\right]_1^2 - \left[\frac{y^4}{4}\right]_1^2 = \frac{7}{6}$$

$$\mathrm{V(X)} = \mathrm{E(X^2)} - \left(\mathrm{E(X)}\right)^2 = \frac{1}{3} - \frac{1}{4} = \frac{1}{12} \qquad \mathrm{V(Y)} = \mathrm{E(Y^2)} - \left(\mathrm{E(Y)}\right)^2 = \frac{7}{6} - 1 = \frac{1}{6}$$

Finalmente

$$\rho_{\mathrm{XY}} = \frac{\mathrm{cov(X,Y)}}{(\mathrm{V(X)\ V(Y)})^{1/2}} = \frac{1}{\sqrt{2}}$$

6. Un sistema electrónico opera con dos tipos de componentes. Sean T_1 y T_2 las variables aleatorias tiempo de duración (en horas) de las componentes de tipo 1 y 2 respectivamente. La función densidad de probabilidad conjunta del vector (T_1, T_2) es

$$f(t_1, t_2) = \begin{cases} \frac{t_1}{8} e^{-\left(\frac{t_1+t_2}{2}\right)} & \text{si } 0 < t_1,\ 0 < t_2 \\ 0 & \text{en otro caso} \end{cases}$$

a) Determine $P(T_1 \geq 1, T_2 \geq 2)$. Nota: la notación anterior es equivalente a $P(T_1 \geq 1 \cap T_2 \geq 2)$.
b) Calcule la probabilidad de que una componente de tipo 2 tenga una duración mayor que 2 horas.

Definimos las variables aleatorias
T_1:"duración en horas de componente tipo 1"
T_2:"duración en horas de componente tipo 2"

$$f(t_1, t_2) = \begin{cases} \frac{t_1}{8} e^{-\left(\frac{t_1+t_2}{2}\right)} & \text{si } 0 < t_1,\ 0 < t_2 \\ 0 & \text{en otro caso} \end{cases}$$

a)

$$P(T_1 \geq 1, T_2 \geq 2) = \int_1^\infty \int_2^\infty \frac{t_1}{8}\ e^{-\frac{t1+t2}{2}}\ dt_1\ dt_2 = \frac{1}{8} \int_1^\infty t_1\ e^{-\frac{t_1}{2}} dt_1\ \int_2^\infty e^{-\frac{t_2}{2}}\ dt_2$$

Para calcular la integral de t_1, usaremos la técnica de integración por partes

$$\int u\ dv = u\, v - \int v\ du$$

Identificando a *u* y *dv*

$$u = t_1 \Rightarrow du = dt_1 \qquad dv = e^{-\frac{t_1}{2}}\ dt_1 \Rightarrow v = -2\ e^{-\frac{t_1}{2}}$$

Luego

$$\int t_1\ e^{-\frac{t_1}{2}}\ dt_1 = -2t_1 e^{-\frac{t1}{2}} + 2 \int e^{-\frac{t1}{2}}\ dt_1 = -2e^{-\frac{t_1}{2}}\ (t_1 + 2)$$

Finalmente

$$P(T_1 \geq 1\ \cap\ T_2 \geq 2) = \frac{1}{8} \int_1^\infty t_1\ e^{-\frac{t_1}{2}} dt_1\ \int_2^\infty e^{-\frac{t_2}{2}} dt_2 = \frac{3}{4}\ e^{-\frac{1}{2}}\ 2\ e^{-1} = \frac{3}{2} e^{-\frac{3}{2}} \cong 0{,}3347$$

b) Calcule la probabilidad de que una componente de tipo 2 tenga una duración mayor que 2 horas.

Debemos calcular la densidad de probabilidad marginal de T_2

$$\mathrm{f}_{\mathrm{T}_2}(t_2) = \int_0^{\infty} \mathrm{f}(t_1,t_2)\ dt_1 = \int_0^{\infty} \frac{t_1}{8}\ e^{-\frac{t_1+t_2}{2}} dt_1 = \frac{e^{-\frac{t_2}{2}}}{8} \int_0^{\infty} t_1 e^{-\frac{t_1}{2}} dt_1 =$$

$$= \frac{e^{-t_2}}{8}\left[-2t_1 e^{-\frac{t_1}{2}} - 4e^{-\frac{t_1}{2}}\right]_0^{\infty} = \frac{e^{-\frac{t_2}{2}}}{2}$$

Finalmente

$$\mathrm{P}(\mathrm{T}_2 > 2) = \int_2^{\infty} \mathrm{f}_{\mathrm{T}_2}(t_2)\cdot dt_2 = \int_2^{\infty} \frac{e^{-\frac{t_2}{2}}}{2} dt_2 = \frac{1}{2}(-2)\left[e^{-\frac{t_2}{2}}\right]_2^{\infty} = e^{-1}$$

7. Los tiempos en horas T_a y T_b que dos estudiantes A y B demoran en resolver un problema son variables aleatorias independientes con distribución exponencial y esperanza 0,5 hs. Calcule la probabilidad de que el estudiante A demore a lo sumo una hora y el estudiante B demore como máximo 45 minutos.

Definimos las variables aleatorias
T_a:"tiempo en que el estudiante A demora en hacer un ejercicio"
T_b:"tiempo en que el estudiante B demora en hacer un ejercicio"

Se pide calcular $P(T_a<1, T_b<0{,}75)$ que es lo mismo que $P(T_a<1 \cap T_b<0{,}75)$.

Del enunciado sabemos que
$E(T_a)=0{,}5$ $\qquad$ $E(T_b)=0{,}5$

$T_a \sim Exp(2)$

$$\mathrm{f}_{\mathrm{T_a}}(t_\mathrm{a}) = \begin{cases} 2e^{-2t_\mathrm{a}} & \text{si } t_\mathrm{a} > 0 \\ 0 & \text{en otro caso} \end{cases}$$

$T_b \sim Exp(2)$

$$\mathrm{f}_{\mathrm{T_b}}(t_\mathrm{b}) = \begin{cases} 2e^{-2t_\mathrm{b}} & \text{si } t_\mathrm{b} > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Como sabemos que T_a y T_b son independientes, la distribución de probabilidad conjunta es
$\mathrm{f}(t_\mathrm{a}, t_\mathrm{b}) = \mathrm{f}_{\mathrm{T_a}}(t_\mathrm{a}) \cdot \mathrm{f}_{\mathrm{T_b}}(t_\mathrm{b})$

Por lo tanto

$$\mathrm{f}_{(t_a,t_b)} = \begin{cases} 4e^{-2(t_a+t_b)} & t_a > 0, t_b > 0 \\ 0 & \text{en otro caso} \end{cases}$$

Finalmente

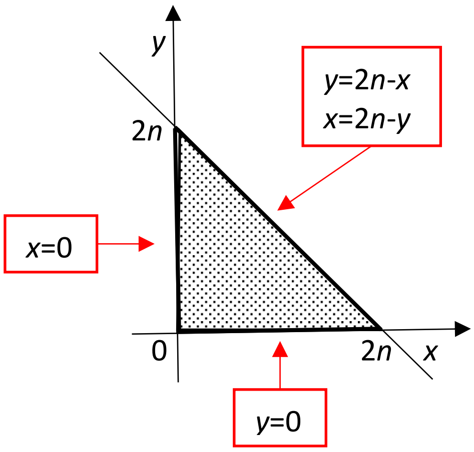
$$P(T_a < 1, T_b < \frac{3}{4}) = 4\int_0^1 \int_0^{\frac{3}{4}} e^{-2(t_a+t_b)}\ dt_a\ dt_b = 4\int_0^1 e^{-2t_a}\ dt_a \int_0^{\frac{3}{4}} e^{-2t_b}\ dt_b =$$

$$= 4\left[\frac{e^{-2t_a}}{-2}\right]_0^1 \left[\frac{e^{-2t_b}}{-2}\right]_0^{\frac{3}{4}} = (1-e^{-2})(1-e^{-\frac{3}{2}}) \approx 0,6717$$

Otra forma de resolver el problema:
Como sabemos que T_a y T_b son independientes, es posible calcular $P(T_a<1)$ y $P(T_b<0,75)$ mediante sus distribuciones marginales y luego multiplicar ambos resultados para calcular $P(T_a<1, T_b<0,75)$.

8. Se escoge al azar un punto de coordenadas reales en el triángulo limitado por $y=0$, $x=0$, $y=2n-x$. Suponiendo que todos los puntos tienen la misma probabilidad de ser elegidos, determine
a) la distribución conjunta de las coordenadas (X, Y)
b) las distribuciones marginales de X e Y
c) la distribución condicional de X dado Y=*y*

Hacemos un gráfico cartesiano similar al del ejercicio 5 para visualizar el dominio de integración.



a) Del enunciado, sabemos que la función de distribución de probabilidad conjunta de X e Y es uniforme. Luego, planteamos la ecuación de f(*x*,*y*)

$$f(x,y) = \begin{cases} k & 0 < x < 2n, \quad 0 < y < 2n - x \\ 0 & \text{en otro caso} \end{cases}$$

Nota: se podría expresar a f(*x*,*y*) considerando que $0<y<2n$ y, consecuentemente, $0<x<2n-y$. (Ver ejercicio resuelto 9)

Como f(*x*,*y*) es una f.d.p. conjunta, debe cumplir con la condición de cierre.
Entonces

$$\int_0^{2n}\int_0^{2n-x} k\ dy\ dx = k\int_0^{2n}\int_0^{2n-x}\ dy\ dx = k\int_0^{2n} [y]_0^{2n-x}\ dx =$$

$$k\int_0^{2n}(2n-x)\ dx = k\left[2nx - \frac{x^2}{2}\right]_0^{2n} = k(4n^2 - \frac{4n^2}{2}) = \frac{k}{2}4n^2 = 2kn^2$$

Para cumplir con la condición de cierre debe ser

$$2kn^2 = 1 \implies k = \frac{1}{2n^2}$$

Finalmente

$$\mathrm{f}(x,y) = \begin{cases} \frac{1}{2n^2} & 0 < x < 2n, \quad 0 < y < 2n - x \\ 0 & \text{en otro caso} \end{cases}$$

b) las distribuciones marginales de X e Y

La distribución marginal de X es $\mathrm{f_X}(x)$ y calculamos la integral

$$\mathrm{f_X}(x) = \int_0^{2n-x} \mathrm{f}(x,y)\ dy = \int_0^{2n-x} \frac{1}{2n^2}\ dy = \frac{1}{2n^2}(2n - x)$$

Luego

$$\mathrm{f_X}(x) = \begin{cases} \frac{2n-x}{2n^2} & 0 < x < 2n \\ 0 & \text{en otro caso} \end{cases}$$

Para el caso de la distribución marginal $\mathrm{f_Y}(y)$ calculamos

$$\mathrm{f_Y}(y) = \int_0^{2n-y} \mathrm{f}(x,y)\ dx = \int_0^{2n-y} \frac{1}{2n^2}\ dx = \frac{2n-y}{2n^2}$$

Luego

$$\mathrm{f_Y}(y) = \begin{cases} \frac{2n-y}{2n^2} & 0 < y < 2n \\ 0 & \text{en otro caso} \end{cases}$$

A su vez, las distribuciones de probabilidad marginales recién calculadas deben ambas cumplir la condición de cierre, es decir

$$\int_{-\infty}^{+\infty} \mathrm{f_X}(x)\, dx = \int_0^{2n} \mathrm{f_X}(x)\, dx = 1 \qquad \int_{-\infty}^{+\infty} \mathrm{f_Y}(y)\, dy = \int_0^{2n} \mathrm{f_Y}(y)\, dy = 1$$

c) la distribución condicional de X dado Y=*y*

Para calcular la distribución condicional usamos la definición

$$\mathrm{f}_{\mathrm{X/Y}=y}(x) = \frac{\mathrm{f}(x,y)}{\mathrm{f_Y}(y)}$$

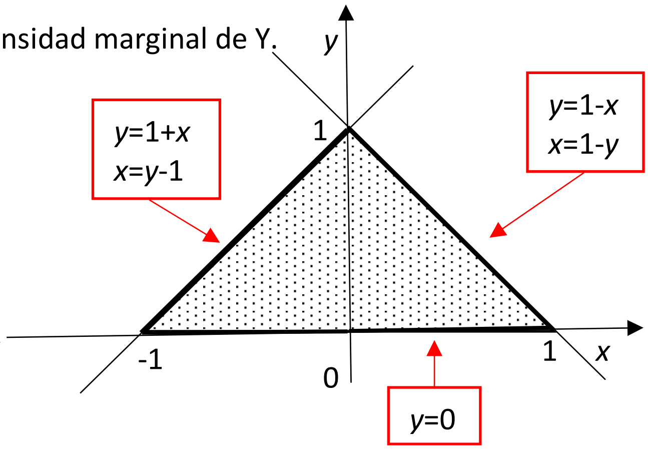
$$\mathrm{f}_{\mathrm{X/Y}=y}(x) = \begin{cases} \dfrac{1}{2n-y} & 0 < y < 2n, 0 < x < 2n - y \\ 0 & \text{en otro caso} \end{cases}$$

El enunciado no lo pide, pero podemos calcular de manera similar $\mathrm{f}_{\mathrm{Y/X}=x}(y)$

$$\mathrm{f}_{\mathrm{Y/X}=x}(x) = \frac{\mathrm{f}(x,y)}{\mathrm{f_X}(x)} \longrightarrow \mathrm{f}_{\mathrm{Y/X}=x}(y) = \begin{cases} \frac{1}{2n-x} & 0 < x < 2n,\ 0 < y < 2n - x \\ 0 & \text{en otro caso} \end{cases}$$

9. Sean X e Y variables aleatorias cuya distribución conjunta es uniforme en el triángulo de vértices (-1,0), (0,1) y (1,0).
a) Encuentre la densidad marginal de X y la densidad marginal de Y.
b) Halle Cov(X,Y).

Primero hacemos un gráfico cartesiano para analizar el dominio de f(*x*,*y*).



Del enunciado, sabemos que la función de distribución de probabilidad conjunta de X e Y es uniforme.

Luego, podemos plantear la ecuación de f(*x*,*y*) considerando que *x* varía en el intervalo [-1,1] e *y* varía entre 0 y la recta *y*=1+*x* o *y*=1-*x* según corresponda.

$$\mathrm{f}(x,y)=\begin{cases} k & -1<x<0, \quad 0<y<1+x \\ k & 0<x<-1, \quad 0<y<1-x \\ 0 & \text{en otro caso} \end{cases} \qquad (9.1)$$

De manera similar, podemos plantear la ecuación de f(*x*,*y*) considerando que *y* varía en el intervalo [0,1] y *x* varía entre la recta *x*=*y*-1 y la recta *x*=1-*y*.

$$\mathrm{f}(x,y)=\begin{cases} k & 0<y<1, \quad y-1<x<1-y \\ 0 & \text{en otro caso} \end{cases} \qquad (9.2)$$

Las expresiones (9.1) y (9.2) son equivalentes, y podemos usar la expresión más conveniente para el análisis.
Como f(*x*,*y*) es una f.d.p. conjunta, debe cumplir con la condición de cierre.
Entonces, usando (9.2) por mayor simplicidad

$$\int_0^1 \int_{y-1}^{1-y} k\ dx\ dy = k\int_0^1 \left[x|_{y-1}^{1-y}\right]\, dy = k\int_0^1 [(1-y)-(y-1)]\, dy = 2k\int_0^1 (1-y)\, dy = k = 1$$

Luego, *k*=1. Usando (9.1) se llega al mismo resultado.
Finalmente

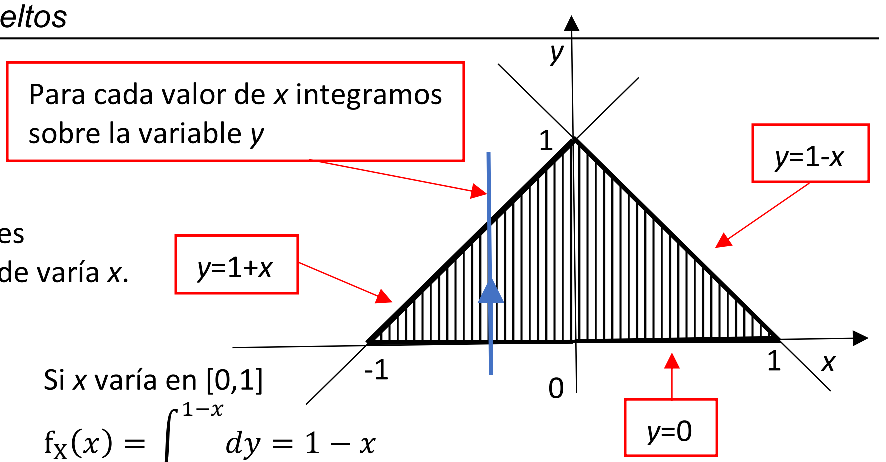
$$\mathrm{f}(x,y)=\begin{cases} 1 & -1<x<0, \quad 0<y<1+x \\ 1 & 0<x<-1, \quad 0<y<1-x \\ 0 & \text{en otro caso} \end{cases} \qquad (9.3)$$

o equivalentemente

$$\mathrm{f}(x,y)=\begin{cases} 1 & 0<y<1, \quad y-1<x<1-y \\ 0 & \text{en otro caso} \end{cases} \qquad (9.4)$$

a) Encuentre la densidad marginal de X y la densidad marginal de Y.
Para calcular $f_X(x)$ debemos integrar a lo largo del eje *y* para cada valor de *x*.

En este caso, conviene usar (9.3) para calcular $f_X(x)$.



Debemos tener en cuenta tres casos según el intervalo donde varía *x*.

Si *x* varía en [-1,0]

$$f_X(x) = \int_0^{1+x} dy = 1 + x$$

Si *x* varía en [0,1]

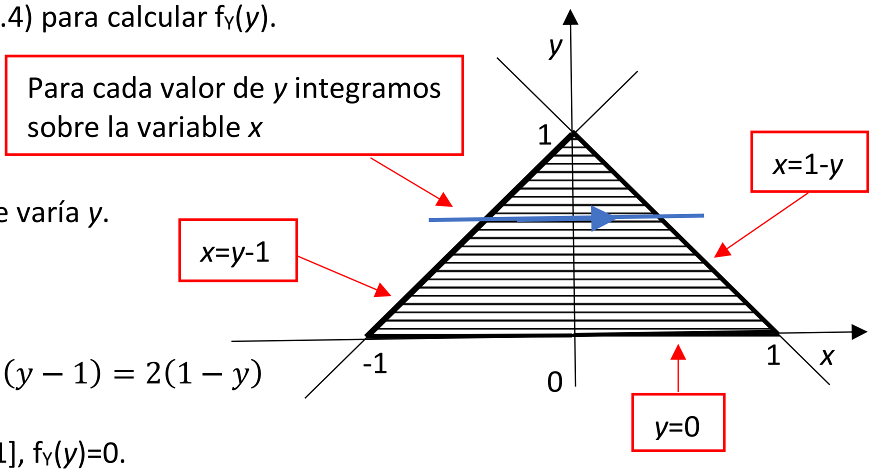
$$f_X(x) = \int_0^{1-x} dy = 1 - x$$

Si *x* no está en el intervalo [-1,1], $f_X(x)=0$.

Por lo tanto

$$f_X(x) = \begin{cases} 1+x & -1 < x < 0 \\ 1-x & 0 < x < 1 \\ 0 & \text{en otro caso} \end{cases} \qquad (9.5)$$

Para calcular $f_Y(y)$ debemos integrar a lo largo del eje *x* para cada valor de *y*.
En este caso, conviene usar (9.4) para calcular $f_Y(y)$.



Debemos tener en cuenta dos casos según el intervalo donde varía *y*.

Si *y* varía en el intervalo [0, 1]

$$f_Y(y) = \int_{y-1}^{1-y} dx = (1-y) - (y-1) = 2(1-y)$$

Si *y* no está en el intervalo [0,1], $f_Y(y)=0$.

Por lo tanto

$$f_Y(y) = \begin{cases} 2(1-y) & 0 < y < 1 \\ 0 & \text{en otro caso} \end{cases} \qquad (9.6)$$

A su vez, las distribuciones de probabilidad marginales recién calculadas deben ambas cumplir la condición de cierre, es decir

$$\int_{-\infty}^{+\infty} f_X(x)\, dx = \int_{-1}^{0} (1+x)\, dx + \int_{0}^{1} (1-x)\, dx = 1 \qquad \int_{-\infty}^{+\infty} f_Y(y)\, dy = \int_{0}^{1} 2(1-y)\, dy = 1$$

b) Halle Cov(X,Y).

Sabemos que Cov(X,Y)=E(X Y)-E(X) E(Y)

$$E(X) = \int_{-\infty}^{+\infty} x\, f_X(x)\; dx = \int_{-1}^{0} x\,(1+x)\; dx + \int_{0}^{1} x\,(1-x)\; dx = (x + x^2)|_{-1}^{0} + (x - x^2)|_{0}^{1} = 0$$

$$\mathrm{E(Y)} = \int_{-\infty}^{+\infty} y\, \mathrm{f_Y}(y)\ dy = \int_{0}^{1} y\, 2(1-y)\ \ dy = 2\int_{0}^{1} (y-y^2)\ \ dy = 2\left(\frac{y^2}{2} - \frac{y^3}{3}\right)\Bigg|_0^1 = \frac{1}{3}$$

$$\mathrm{E(X\,Y)} = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} x\, y\ \ \mathrm{f}(x,y)\ dx\ dy = \int_{0}^{1}\int_{y-1}^{1-y} x\, y\ \ dx\ dy = \int_{0}^{1} y \int_{y-1}^{1-y} x\ dx\ dy = \cdots = 0$$

Finalmente: Cov(X,Y)=0

Suma de variables aleatorias

10. Un aparato de televisión puede tener dos tipos de roturas: debido a falla de transistores o debido a la falla de condensadores. Ambas fuentes de rotura son independientes. El número de roturas debido a falla de transistores durante el primero año de utilización del aparato es una v.a. que sigue una ley de Poisson con promedio 0,5. El número de roturas debido a la falla de condensadores, durante el mismo período, sigue una ley de Poisson con promedio 1. Calcule la probabilidad de que, en el primer año de utilización del aparato, éste tenga exactamente 2 roturas.

En este problema debemos analizar dos v.a.d. con distribución de Poisson.
T:"número de roturas debido a falla de transistores durante el primer año de uso"
C:"número de roturas debido a falla de condensadores durante el primer año de uso"

De acuerdo al enunciado y recordando los conceptos de distribuciones de Poisson, las esperanzas de C y T son
E(T)=1/2= λ_T
E(C)=1= λ_C
donde λ_T y λ_C son los parámetros de las distribuciones de T y C respectivamente.
Entonces : T ~ Po(1/2) y C ~ Po(1)

Definimos una nueva v.a. que representa la totalidad de roturas en un período de tiempo
X:"número de roturas durante el primer año de uso"
X= T + C

Como T y C son independientes, por propiedad reproductiva de distribuciones de Poisson
X~Po(λ_X) donde $\lambda_X = \lambda_T + \lambda_C$ =3/2

El problema pide calcular la probabilidad de que ocurran 2 roturas durante el primer año de uso.
Entonces: X~Po(3/2)

$$\mathrm{P(X=2)} = e^{-\left(\frac{3}{2}\ \cdot\ 2\right)}\ \frac{\left(\frac{3}{2}\right)^2}{2!} \cong 0{,}056$$

11. Ciertos elementos de un circuito eléctrico se protegen contra el exceso de voltaje por medio de dos relevadores R_1 y R_2 que se ajustan para ser descargados en períodos X_1 y X_2 respectivamente, después que comienza el exceso de voltaje. Estos períodos de descarga varían debido a pequeños factores incontrolables, por lo que se puede suponer que X_1 y X_2 son v.a. normales independientes con tiempos medios de descarga $\mu_1 = 1$ s y μ_2, y varianzas $\sigma_1^2=\sigma_2^2=0{,}1$ s^2. Determine μ_2 de manera tal que la probabilidad de que R_2 sea descargada antes que R_1 sea a lo sumo 0,001.

Definimos las variables aleatorias
X_1:"tiempo de descarga de R_1"
X_2:"tempo de descarga de R_2"
Sabemos que X_1 y X_2 son v.a. independientes y
$X_1 \sim N(\mu_1, \sigma_1)$ $\quad$ $\mu_1=1$ s $\quad$ $\sigma_1^2=0{,}1$ s^2
$X_2 \sim N(\mu_2, \sigma_2)$ $\quad$ $\mu_2=¿?$ $\quad$ $\sigma_2^2=0{,}1$ s^2
El problema pide calcular
$P(X_2<X_1)=P(X_2-X_1<0)=P(Y<0)\leq 0{,}001$
Luego, analizamos la variable aleatoria $Y= X_2-X_1$.

¿Por qué se puede aplicar esta propiedad?

Por propiedad reproductiva de la distribución normal, resulta
$Y \sim N(\mu, \sigma)$ $\quad$ con $\mu= \mu_2-\mu_1= \mu_2-1$ y $\sigma^2=\sigma_1^2+\sigma_2^2=0{,}2$

Definiendo la variable normal estandarizada $Z = \frac{Y-\mu}{\sigma}$ podemos decir

$$P(Y<0) = P\left(Z < z = \frac{-\mu}{\sigma}\right) \leq 0{,}001 \Rightarrow z \leq -3{,}10$$

Consideramos un valor de *z* que cumple con la condición de P(Y<0)≤0,001

$$\frac{-\mu}{\sigma} = \frac{-(\mu_2-\mu_1)}{\sqrt{({\sigma_2}^2+{\sigma_1}^2)}} = \frac{-(\mu_2-1)}{\sqrt{0{,}2}} \leq -3{,}10 \Rightarrow \frac{(\mu_2-1)}{\sqrt{0{,}2}} \geq 3{,}10 \Rightarrow \mu_2 \geq 2{,}3864$$

12. Sean $X_1,X_2,\ldots,X_n$, *n* variables aleatorias independientes distribuidas uniformemente en el intervalo (0,1). Encuentre la función de densidad de probabilidad de la variable aleatoria
$Z= X_1+X_2+\ldots+X_{60}$.

Sabemos que
$X_i \sim U[0,1]$ $\quad$ con $i=1,\ldots,60$

$$f_{X_i}(x_i) = \begin{cases} 1 & 0 \leq x_i \leq 1 \\ 0 & \text{en otro caso} \end{cases} \qquad F_{X_i}(x_i) = \begin{cases} \ldots \\ \ldots \\ \ldots \end{cases}$$

completar

$E(X_i)=1/2$ $\qquad$ $V(X_i)=1/12$

Como $X_1,X_2,\ldots,X_n$ son independientes, $E(X_i)$ y $V(X_i)$ son finitas para todo $i=1,\ldots,60$ y considerando que $n=60$ es un número suficientemente grande, podemos aplicar el teorema central del límite y concluir que la variable $Z= X_1+X_2+\ldots+X_{60}$ tiene una distribución aproximadamente normal de media μ y varianza σ^2.

Luego

$$Z = \sum_{i=1}^{60} X_i \sim N(\mu, \sigma^2)$$

$$\mu = E(Z) = \sum_{i=1}^{60} E(X_i) = 60 \,.\frac{1}{2} = 30$$

$$\sigma^2 = V(Z) = \sum_{i=1}^{60} V(X_i) = 60 \,.\frac{1}{12} = 5$$

Finalmente

$$f_Z(z) \cong \frac{1}{\sqrt{2\pi}\sqrt{5}} \; e^{-\frac{(z-30)^2}{5}}$$

$$R_Z = \{z \in \mathbb{R} / 0 \leq z \leq 60\}$$

¿Por qué el recorrido de Z no es todo el conjunto de los números reales?

14. Un ensamble eléctrico consta de 20 bloques de tipo A y 30 bloques de tipo B conectados en serie, cuyas longitudes varían aleatoriamente con distribución normal. Los ensambles se colocan en recipientes cuya longitud (en cm) varía aleatoriamente con distribución normal, media 65 y desviación estándar 0,5. Se conoce además que la longitud (en cm) de un bloque de tipo A varía aleatoriamente con media 1,95 y desvío estándar 0,01 y la longitud (en cm) de un bloque de tipo B varía aleatoriamente con media 0,83 y desviación estándar 0,02. Calcule la probabilidad de que un ensamble entre en un recipiente, cuando ambos son elegidos al azar.

Definimos las variables aleatorias continuas
A_i:"longitud en cm de un bloque tipo A" $\quad i=1,\ldots,20$
B_j:"longitud en cm de un bloque tipo B" $\quad j=1,\ldots,30$
R:"longitud en cm de un recipiente para el ensamble de bloques"

De acuerdo al enunciado

$A_i \sim N(E(A_i),V(A_i))$	$E(A_i)=1{,}95$	$V(A_i)=0{,}01^2$	$i=1,\ldots,20$
$B_j \sim N(E(B_j),V(B_j))$	$E(B_j)=0{,}83$	$V(B_j)=0{,}02^2$	$j=1,\ldots,30$
$R \sim N(E(R),V(R))$	$E(R)=65$	$V(R)=0{,}5^2$	

Supongamos que R, A_i con $i=1,\ldots,20$ y B_j con $j=1,\ldots,30$, son variables aleatorias independientes entre sí.
Definimos la v.a.c. A:"longitud total en cm de los 20 bloques tipo A usados"

$$A = \sum_{i=1}^{20} A_i$$

Como todas las variables A_i son independientes entre sí y tienen distribución normal, por propiedad reproductiva de la distribución normal la variable A tiene distribución normal:
$A \sim N(E(A),V(A))$

Donde

$$\mathrm{E(A)} = \mu_{\mathrm{A}} = \sum_{i=1}^{20} \mathrm{E}(\mathrm{A}_i) = 20\,.\,1{,}95 = 39 \qquad \mathrm{V(A)} = \sigma_{\mathrm{A}}^2 = \sum_{i=1}^{20} \mathrm{V}(\mathrm{A}_i) = 20\,.\,0{,}01^2 = 0{,}002$$

Similarmente, definimos la v.a.c. B:"longitud total en cm de los 30 bloques tipo B usados"

$$\mathrm{B} = \sum_{j=1}^{30} \mathrm{B}_j$$

Como todas las variables B_j son independientes entre sí y tienen distribución normal, por propiedad reproductiva de las distribuciones normales, la variable B tiene distribución normal:
B ~ N(E(B),V(B))
Donde

$$\mathrm{E(B)} = \mu_{\mathrm{B}} = \sum_{j=1}^{30} \mathrm{E}(\mathrm{B}_j) = 30\,.\;0{,}83 = 24{,}9 \qquad \mathrm{V(B)} = \sigma_{\mathrm{B}}^2 = \sum_{j=1}^{30} \mathrm{V}(\mathrm{B}_j) = 30\,.\,0{,}02^2 = 0{,}012$$

Para que el ensamble completo quepa en un recipiente, debe ser A+B<R. Luego, debemos calcular P(A+B<R). Entonces definimos una nueva v.a.c. Y=A+B-R de manera que P(A+B<R)=P(A+B-R<0)=P(Y<0)

Considerando que A, B y R son v.a. independientes y con distribución normal, por propiedad reproductiva de las distribuciones normales, Y tiene distribución normal: Y ~ N(μ_Y , σ_Y^2)
Donde

$$\mathrm{E(Y)} = \mu_{\mathrm{Y}} = \mu_{\mathrm{A}} + \mu_{\mathrm{B}} - \mu_{\mathrm{R}} = -1{,}1 \qquad \mathrm{V(Y)} = \sigma_{\mathrm{Y}}^2 = \sigma_{\mathrm{A}}^2 + \sigma_{\mathrm{B}}^2 + \sigma_{\mathrm{R}}^2 = 0{,}264$$

Definimos la variable estandarizada Z

$$\mathrm{Z} = \frac{\mathrm{Y} - \mu_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}}$$

Luego

$$\mathrm{P(A + B < R) = P(A + B - R < 0) = P(Y < 0)} = \mathrm{P}\left(\mathrm{Z} < \frac{-\mu_{\mathrm{Y}}}{\sigma_{\mathrm{Y}}}\right) = \mathrm{P(Z < 2{,}14)} = 0{,}9838$$

Finalmente, la probabilidad de que un ensamble quepa en un recipiente, cuando ambos son elegidos al azar es 0,9838.

15. Un sistema está formado por 100 componentes que funcionan independientemente. La probabilidad de que cualquier componente falle durante el período de operación es igual a 0,10. Para que funcione el sistema completo, deben funcionar al menos 85 componentes. Calcule la probabilidad de que el sistema funcione.
En este problema se puede resolver usando una distribución binomial, pero conviene usar el teorema central del límite para hallar la solución aproximada.

Veamos la solución usando la distribución binomial.
Definimos los sucesos
D_i:"falla componente i durante el período de operación" con i=1,…,100
P(D_i)=0,1 de acuerdo al enunciado

Definimos la variable aleatoria discreta
N:"número de componentes que fallan durante el período de operación de un total de 100 componentes"
Como las componentes funcionan independientemente, podemos decir que
$\mathrm{N} \sim \mathrm{Bi}(100, 0{,}1)$

$$\mathrm{p_N}(n) = \binom{100}{n}\ 0{,}1^{n}\ \ 0{,}9^{100-n} \qquad n = \overline{0{,}100}$$

E(N)=10 V(N)=9

Sabemos que para que funcione el sistema, deben funcionar al menos 85 componentes. Es decir que, como máximo, pueden fallar 100-85=15 componentes. Entonces debemos calcular

$$\mathrm{P}(\mathrm{N} \leq 15) = \sum_{n=0}^{15} \binom{100}{y}\ 0{,}1^{n}\ \ 0{,}9^{100-n} = \cdots$$

Otra forma: aproximaremos el cálculo usando el teorema central del límite.
Definimos los sucesos
D_i:"falla componente i durante el período de operación" con i=1,…,100
$P(D_i)$=0,1 de acuerdo al enunciado

Definimos las variables aleatorias discretas asociadas a cada D_i de la siguiente manera

$$\mathrm{X}_i = \begin{cases} 0 & \text{si } \mathrm{D}_i \text{ no falla} \\ 1 & \quad \text{si } \mathrm{D}_i \text{ falla} \end{cases} \qquad \text{con } i = 1, \dots, 100$$

donde cada X_i es independiente del resto de acuerdo al enunciado.

Luego, por ser sucesos equivalentes

$\mathrm{P}(\mathrm{X}_i = 0) = \mathrm{P}(\overline{\mathrm{D}_i}) = 0{,}9$ y $\mathrm{P}(\mathrm{X}_i = 1) = \mathrm{P}(\mathrm{D}_i) = 0{,}1$

$E(X_i)$=0,1 $V(X_i)$=0,09

Ahora definimos una nueva variable aleatoria discreta

$$\mathrm{Y} = \sum_{i=1}^{100} \mathrm{X}_i$$

Como cada X_i es independiente del resto, como $E(X_i)$ y $V(X_i)$ son finitas para todo i=1,…,100 y considerando que 100 es un número suficientemente grande, podemos aproximar la distribución de Y a una distribución normal de la siguiente forma
$\mathrm{Y} \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$ donde

$$\mu_{\mathrm{Y}} = \sum_{i=1}^{100} \mathrm{E}(\mathrm{X}_i) = 10 \qquad \sigma_{\mathrm{Y}}^2 = \mathrm{V}(\mathrm{Y}) = \sum_{i=1}^{100} \mathrm{V}(\mathrm{X}_i) = 9 \Rightarrow \sigma_{\mathrm{Y}} = 3$$

Sabemos que para que funcione el sistema, deben funcionar al menos 85 componentes. Es decir que, como máximo, pueden fallar 100-85=15 componentes. Entonces debemos calcular
$P(Y \leq 15)$=… ← completar

16. El consumo de combustible en litros de un ómnibus que realiza el trayecto Rosario-Santa Fe es una variable aleatoria normalmente distribuida con media 20 litros y desviación estándar 2 litros. Calcule la probabilidad de que 640 litros resulten insuficientes para realizar 30 viajes.

En este problema nos interesa analizar las v.a.c.
X_i:"consumo de combustible en el trayecto *i*"
Y:"consumo total de combustible en 30 trayectos"

Sabemos que $Y = \sum_{i=1}^{30} X_i$

y se desea calcular la probabilidad que el gasto total sea mayor que 640 litros, es decir P(Y>640).

Como el consumo en cada trayecto es independiente del consumo en el resto de los trayectos, y que la distribución de probabilidad en cada trayecto es normal, aplicando la propiedad reproductiva de la distribución normal es posible afirmar que Y tiene una distribución normal
$Y \sim N(\mu_Y, \sigma_Y^2)$ donde μ_Y =... y σ_Y^2=...

completar

19. Una empresa dedicada a la venta de repuestos sabe que la demanda diaria D varía aleatoriamente con la siguiente distribución de probabilidad

D	0	1	2
P(D=*d*)	0,25	0,5	0,25

Desde el momento en que se hace el pedido hasta que el mismo ingresa al stock transcurren 90 días. Determine cuántas unidades deben tenerse en existencia en momento de hacer el pedido si se quiere que la probabilidad de que la demanda durante los 90 días supere la existencia sea 0,05.
Definimos las variables
D_i:"demanda en el día *i*"
D_T:"demanda total en un período de 90 días"

La demanda total será $D_T = \sum_{i=1}^{90} D_i$

Calcular μ_i y σ_i

Podemos calcular: $E(D_i)=\mu_i$=... y $V(D_i)=\sigma_i^2$=... . Considerando que la demanda de cada día es independiente de la demanda del resto de los días y que $E(D_i)$ y $V(D_i)$ son finitas para todo *i*=1,...,90, y que las cantidad de días (90) es un número suficientemente grande, podemos aplicar el teorema central del límite y concluir que la variable D_T tiene una distribución aproximadamente normal de media μ y varianza σ^2.
μ=...=90
σ^2=...=45
$D_T \sim N(90, 45)$

completar

Sea S:"unidades en existencia"
de acuerdo al enunciado, la probabilidad que la demanda total supere la existencia
$P(D_T > S) = 0{,}05 \Rightarrow P(D_T < S) = 0{,}95 \Rightarrow \cdots \Rightarrow S \gtrsim 101$

completar

20. La demanda diaria de agua potable por habitante en una población de 10000 habitantes es una v.a. H con E(H) = 0,4 m^3 y σ(H) = 0,09 m^3. La disponibilidad de agua (en m^3) para el consumo almacenado diariamente en una represa es una v.a. Y~N(4500; 450^2).
¿Cuál es la probabilidad de que en un día cualquiera no sea satisfecha la demanda?

Para resolver este problema, buscaremos la demanda total diaria de agua.
Siendo Hi:"demanda diaria de agua del habitante *i*"
$E(H_i)=0,4$ $V(H_i)=0,09^2=0,0081$ con *i*=1, …, 10000

Definimos la variable aleatoria X:"demanda diaria total de agua"

$$X = \sum_{i=1}^{10000} H_i$$

Considerando que la demanda diaria por habitante es independiente de la demanda del resto de los habitantes y que $E(H_i)$ y $V(H_i)$ son finitas para todo *i*=1,…,10000, y que las cantidad de habitantes es un número suficientemente grande, podemos aplicar el teorema central del límite y concluir que la variable X tiene una distribución aproximadamente normal de media μ_X y varianza σ_X^2.
X~N(μ_X , σ_X^2) con E(X)= μ_X =…=4000 y V(X)= σ_X^2=…=81 completar

Por otra parte, sabemos que la disponibilidad diaria de agua potable es una v.a. Y con distribución normal, es decir Y~N(4500; 450^2).

La probabilidad de que en un día cualquiera no sea satisfecha la demanda significa evaluar P(Y<X). Para poder hacer este cálculo, definimos una nueva variable aleatoria R=Y-X.
Luego P(Y<X)=P(R<0).

Como X e Y tienen ambas distribución normal, y consideramos que la demanda diaria X y la disponibilidad diaria de agua Y son independientes, podemos decir que R tiene una distribución normal por la propiedad reproductiva de las distribuciones normales.
En consecuencia
R~N(μ_R , σ_R^2) con E(R)= μ_R = μ_Y - μ_X =500 y V(R)= σ_R^2= σ_Y^2+ σ_X^2 =202581

P(Y<XS)=P(R<0)=…≈0,1335 completar