Capitulo IV:

IV.1 Variables Aleatorias Discretas

Problema IV.1.1

Sea $A = \{1, 2, 3\}$*. Se realiza el siguiente experimento, se extrae al azar un subconjunto de* A*.*

(a) Encuentre $(\Omega, \mathcal{F}, P)$

(b) Se define la variable aleatoria $X : \Omega \to I\!R$ *como:*

$$\omega \to X(\omega) = \begin{cases} \sum_{i \in \omega} i & \omega \neq \phi \\ 0 & \omega = \phi \end{cases}$$

Pruebe que X *es variable aleatoria discreta, obtenga su soporte y su ley inducida.*

Solución:

(a) $\Omega = \mathcal{P}(A) \qquad \mathcal{F} = \mathcal{P}(\mathcal{P}(A)) \qquad |\Omega| = 2^3 = 8.$

De hecho $\Omega = \{\{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}, \phi\}$

Dado $\omega \in \Omega \qquad I\!P(\omega) = \frac{1}{|\Omega|} = \frac{1}{8}.$

(b) Como $\mathcal{F} = \mathcal{P}(\mathcal{P}(A))$ toda función $X : \Omega \to I\!R$ será medible pues esa σ-algebra contiene a todos los subconjuntos.

Hagamos el cálculo de $X(\omega)$

ω	ϕ	$\{1\}$	$\{2\}$	$\{3\}$	$\{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$X(\omega)$	0	1	2	3	3	4	5	6

$X(\omega) \in \{0, 1, 2, 3, 4, 5, 6\}$ por lo tanto es una variable aleatoria discreta.

La ley inducida de la variable aleatoria discreta X está determinada por los números $I\!P[X^{-1}(k)]$ con $k \in \{0, 1, 2, ..., 6\}$

k	0	1	2	3	4	5	6
$X^{-1}(k)$	ϕ	$\{1\}$	$\{2\}$	$\{3\} \cup \{1,2\}$	$\{1,3\}$	$\{2,3\}$	$\{1,2,3\}$
$I\!P[X^{-1}(k)]$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Problema IV.1.2

Se hacen n lanzamientos independientes con un dado ordinario de 6 lados. Calcule la probabilidad que:

(a) El mayor de los números obtenidos sea k con $k \in \{1, ..., 6\}$.
(b) El menor de los números obtenidos sea k con $k \in \{1, ..., 6\}$.

Solución:

Es claro que la variable aleatoria a examinar tiene un soporte discreto.

El espacio de probabilidad será $\Omega = \{1, ..., 6\}^n \quad \mathcal{F} = \mathcal{P}(\Omega)$
Dado $w \in \Omega \quad I\!P(\{\omega\}) = \frac{1}{6^n}$, es decir estamos considerando un dado equilibrado con lanzamientos independientes.

(a) Llamemos X_i el valor del i-esimo lanzamiento $\forall i = 1, ..., n$.

Es sencillo de entender que:

$$I\!P\left[\max_{i=1,\dots,n} X_i \leq k\right] = I\!P(X_i \leq k \ \forall i = 1, ..., n)$$

Por independencia entre lanzamientos lo anterior será igual a:

$$\prod_{i=1}^{n} I\!P(X_i \leq k)$$

Como todas las variables X_i tienen igual distribución basta calcular:

$$I\!P(X_i \leq k) = I\!P\left[\bigcup_{i=1}^{k}\{X_i = k\}\right] = \sum_{i=1}^{k} I\!P(X_i = k) = \sum_{i=1}^{k} \frac{1}{6} = \frac{k}{6}$$

Por lo tanto $I\!P\left[\max_{i=1,\dots,n} X_i \leq k\right] = \left(\frac{k}{6}\right)^n \quad \forall k = 0, ..., 6$

Además:

$$\begin{aligned} I\!P\left[\max_{i=1,\dots,n} X_i \leq k\right] &= I\!P\left[\{\max_{i=1,\dots,n} X_i \leq k-1\} \cup \{\max_{i=1,\dots,n} X_i = k\}\right] \\ &= \left(\frac{k-1}{6}\right)^n + I\!P(\max_{i=1,\dots,n} X_i = k) \quad \forall k = 1, ..., 6 \end{aligned}$$

Por lo tanto, al despejar:

$$\mathbb{P}\left[\max_{i=1,\ldots,n} X_i = k\right] = \left(\frac{k}{6}\right)^n - \left(\frac{k-1}{6}\right)^n$$

(b) Es extremadamente cómodo notar que:

$$\min_{i=1,\ldots,n} X_i = \max_{i=1,\ldots,n} (7 - X_i)$$

En consecuencia:

$$\mathbb{P}\left[\min_{i=1,\ldots,n} X_i = k\right] = \mathbb{P}(7 - \max_{i=1,\ldots,n} X_i = k) = \mathbb{P}\left[\max_{i=1,\ldots,n} X_i = 7 - k\right]$$

Por lo tanto:

$$\mathbb{P}\left[\min_{i=1,\ldots,n} X_i = k\right] = \left(\frac{7-k}{6}\right)^n - \left(\frac{6-k}{6}\right)^n = \left(1 - \frac{(k-1)}{6}\right)^n - \left(1 - \frac{k}{6}\right)^n$$

Si uno no se da cuenta de la relación anterior para resolver (b) se puede calcular $\mathbb{P}\left[\min_{i=1,\ldots,n} X_i \geq k\right] = \mathbb{P}(X_i \geq k \ \forall i = 1, ..., n)$ y razonar como en (a).

Observación:

El método utilizado para calcular la distribución del máximo se puede utilizar para cualquier familia de variable aleatorias iid (independientes idénticamente distribuidas) y lo mismo para el mínimo.

Problema IV.1.3

Tres puntos son escogidos al azar e independientemente en el intervalo $[0, 1]$. *Sea* X *el número de puntos que pertenecen al inervalo* $[0, c]$ *donde* $c \in [0, 1]$ *es un número real fijo. Se pide que determine para cada* $i \in \{0, 1, 2, 3\}$: $P[X = i]$.

Solución:

El experimento consistente en extraer tres puntos al azar y de manera independiente del intervalo $[0, 1]$ se modela con el espacio muestral producto $\Omega := [0, 1] \times [0, 1] \times [0, 1]$

dotado de la σ-álgebra producto $\mathcal{A} := \beta([0,1]) \otimes \beta([0,1]) \otimes \beta([0,1])$ y de la medida de probabilidad P tal que

$$(\forall A_1, A_2, A_3 \in \beta([0,1])) : \quad P(A_1 \times A_2 \times A_3) = \text{largo}(A_1) \cdot \text{largo}(A_2) \cdot \text{largo}(A_3)$$

Consideremos la función $X : \Omega \to \{0,1,2,3\}$ definida por

$$X(\omega) := \text{ No. de coordenadas de } \omega \text{ pertenecientes al intervalo } [0,c].$$

Como

$$
\begin{aligned}
&X^{-1}(\{0\}) =]c,1] \times]c,1] \times]c,1] \in \mathcal{A} \\
&X^{-1}(\{1\}) = [0,c] \times]c,1] \times]c,1] \cup]c,1] \times [0,c] \times]c,1] \cup]c,1] \times]c,1] \times [0,c] \in \mathcal{A} \\
&X^{-1}(\{2\}) = [0,c] \times [0,c] \times]c,1] \ \cup \ [0,c] \times]c,1] \times [0,c] \cup]c,1] \times [0,c] \times [0,c] \in \mathcal{A} \\
&X^{-1}(\{3\}) = [0,c] \times [0,c] \times]0,c] \in \mathcal{A}
\end{aligned}
$$

si en el conjunto $\{0,1,2,3\}$ definimos la σ-álgebra $\mathcal{A}' := \mathcal{P}(\{0,1,2,3\})$ se tendrá que $X : (\Omega, \mathcal{A}) \to (\{0,1,2,3\}, \mathcal{A}'))$ resulta medible y

$$
\begin{aligned}
&P[X=0] = (1-c) \cdot (1-c) \cdot (1-c) = (1-c)^3 \\
&P[X=1] = c \cdot (1-c) \cdot (1-c) + (1-c) \cdot c \cdot (1-c) + (1-c) \cdot (1-c) \cdot c = 3c(1-c)^2 \\
&P[X=2] = c \cdot c \cdot (1-c) + c \cdot (1-c) \cdot c + (1-c) \cdot c \cdot c = c^2(1-c) \\
&P[X=3] = c \cdot c \cdot c = c^3
\end{aligned}
$$

Problema IV.1.4

Una urna contiene diez cartas enumeradas de 1 hasta 6; se extraen sucesivamente cartas de la urna y sin reposición. Considere el experimento (E) consistente en determinar en que extracción aparece la primera carta con número par. Se pide que:

(a) Modele el experimento (E) como la evaluación de cierta variable aleatoria X definida en cierto espacio probabilístico.

(b) Determine para cada $n \in \{1,...,6\}$: $P[X=n]$.

Solución:

(a) Consideremos el conjunto $\Omega := \{\omega = (\omega_1, ..., \omega_6) \in \{1,...,6\}^6 : i \neq j \Rightarrow \omega_i \neq \omega_j\}$ y la función $X : \Omega \to \{1,...,6\}$ definida por

$$(\forall \omega \in \Omega) : \quad X(\omega) := \min\{i \in \{1,...,6\} : \omega_i \text{ es par }\}$$

El experimento (E) puede ser modelado como la evaluación de la función X en un punto ω, extraído al azar y de manera equiprobable en el conjunto Ω. Lo anterior justifica que en el e.m. Ω (de cardinal finito) definamos la σ-álgebra $\mathcal{A} := \mathcal{P}(\Omega)$ y la medida de probabilidad

$$(\forall A \in \mathcal{A}): \quad P(A) := \frac{|A|}{|\Omega|} = \frac{|A|}{6!}$$

Al poner en el espacio de llegada de X la σ-álgebra $\mathcal{P}(\{1, ..., 6\})$ vemos que X es una función medible e.d. X es una variable aleatoria .

(b) Claramente se tiene que si

$$6 \geq n > 4 \Rightarrow [X = n] := \{\omega \in \Omega : X(\omega) = n\} = \phi$$

luego

$$P[X = 5] = 0$$

$$P[X = 6] = 0$$

Por otro lado fácilmente se demuestra que:

$$|[X = 1]| = |\{\omega \in \Omega : \omega_1 \text{ es par }\}| = 3 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 5! \cdot 3$$
$$|[X = 2]| = |\{\omega \in \Omega : \omega_2 \text{ es impar } \wedge \omega_2 \text{ es par }\}| = 3 \cdot 3 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 4! \cdot 9$$
$$|[X = 3]| = |\{\omega \in \Omega : \omega_1, \omega_2 \text{ son impares } \wedge \omega_3 \text{ es par }\}| = 3 \cdot 2 \cdot 3 \cdot 3 \cdot 2 \cdot 1 = 3! \cdot 18$$
$$|[X = 4]| = |\{\omega \in \Omega : \omega_1, \omega_2, \omega_3 \text{ son impares } \wedge \omega_4 \text{ es par }\}| = 3 \cdot 2 \cdot 1 \cdot 3 \cdot 2 \cdot 1 = 3! \cdot 6$$

y por lo tanto

$$P[X = 1] = \frac{360}{720}$$
$$P[X = 2] = \frac{216}{720}$$
$$P[X = 3] = \frac{108}{720}$$
$$P[X = 4] = \frac{36}{720}$$

Problema IV.1.5

Considere el experimento (E') de tirar dos veces e independientemente un dado y a continuación sumar los resultados obtenidos en cada tirada. Se le pide que:

(a) Modele el experimento (E') como la evaluación de cierta v.a. X definida en cierto espacio probabilístico $(\Omega, \mathcal{A}, P)$ y a valores en cierto espacio de probabilidad $(\Omega', \mathcal{A}')$.

(b) Determine en (E') la probabilidad de que ocurra el suceso "la suma de los resultados obtenidos en cada tirada es par".

(c) *Determine si en (E') el suceso "los resultados obtenidos en cada tirada del dado son números consecutivos" puede o no ser descrito a partir de los valores que toma la v.a. X.*

Solución:

(a) El experimento (E') es natural modelarlo con el e.m. $\Omega' := I\!N$ y la σ-álgebra de sucesos de interés $\mathcal{A}' := \mathcal{P}(\Omega')$. Por otro lado, el experimento (E) consistente en tirar dos veces e independientemente una moneda se modela naturalmente con el espacio probabilístico $(\Omega, \mathcal{A}, P)$ donde $\Omega := \{(i,j) : i,j \in \{1,...,6\}\}, \mathcal{A} := \mathcal{P}(\Omega)$ y $P(A) := \frac{|A|}{36}$ para todo $A \in \mathcal{A}$.

Es claro que cualquier pregunta concerniente al resultado de (E') se puede formular en términos de los valores que toma la función $X : \Omega \to \Omega'$ definida por:

$$X(\omega) = i + j \quad \text{si } \omega = (i,j) \in \Omega$$

Este hecho es lo que intenta rescatar en la definición de medibilidad de una función. En nuestro caso en particular es claro que

$$(\forall A \in \mathcal{A}') : \quad X^{-1}(A) \in \mathcal{A}$$

y por lo tanto la función $X : (\Omega, \mathcal{A}) \to (\Omega, \mathcal{A}')$ es medible, es decir, X es una variable aleatoria.

(b) Denotando por A al suceso en (E') definido por "la suma de los resultados obtenidos en cada tirada es par" se tiene que A tiene asociado en Ω' el conjunto

$$A := \{2n : n \in I\!N\} \in \mathcal{A}'$$

Es claro que la frecuencia de ocurrencia del suceso A en el experimento (E') coincide con la frecuencia de ocurrencia del suceso $X^{-1}(A)$ en el experimento (E). Es por este motivo que la probabilidad de ocurrencia del suceso A en (E') viene dada por el número $P[X \in A] := P(X^{-1}(A))$,que está bien definido pues $X^{-1}(A) \in \mathcal{A}$. De este modo como

$$X^{-1}(A) = \{(1,1),(1,3),(1,5),(2,2),(2,4),(2,4),(2,6),(3,1),(3,3),(3,5),$$

$$(4,2),(4,4),(4,6),(5,1),(5,3),(5,5),(6,2),(6,4),(6,6)\}$$

entonces

$$P[X \in A] = \frac{|X^{-1}(A)|}{36} = \frac{18}{36} = \frac{1}{2}$$

(c) El suceso B en (E) definido por "los resultados obtenidos en cada tirada del dado son números consecutivos" tiene asociado en Ω el conjunto

$$B := \{(1,2),(2,1),(2,3),(3,2),(3,4),(4,3),(4,5),(5,4),(5,6),(6,5)\}$$

¿A partir del valor que toma $X(\omega)$ se puede discriminar si $\omega \in B$ o $\omega \notin B$?

Claramente no pues si $X(\omega) = 5$ entonces ω puede ser $(2,3) \in B$ o $(1,4) \notin B$. Formalmente lo anterior se expresa como

$$(\forall B' \in \mathcal{A}') : X^{-1}(B') \neq B$$

Lo anterior muestra que la v.a. X resume lo suficiente la información del resultado del experimento (E) como para eventualmente no poder saber a partir del valor que toma X si el resultado del experimento (E) pertenece o no a B.

Si decimos que en (E) un suceso es X-observable, si a partir del valor que toma la v.a. X podemos decidir si el resultado de (E) pertenece o no al suceso en cuestión, se tendrá de lo discutido que los únicos sucesos X-observables en (E) son aquellos que pertenecen al conjunto definido por

$$\{A \in \mathcal{A} : A = X^{-1}(A') \text{ para algún } A' \in \mathcal{A}'\} =: \sigma(X)$$

Problema IV.1.6

Sea $(\Omega, \mathcal{A}, P)$ *un espacio probabilístico y* $X, Y : \Omega \to \{1,2,3\}$ *dos variables aleatorias tales que*

$$(\forall i,j \in \{1,2,3\}) : \quad P[X = i, Y = j] = a_{ij}$$

donde $A = (a_{ij})_{1 \le i,j \le 3}$ *es la matriz dada por*

$$A = \begin{bmatrix} 0 & 1/5 & 0 \\ 1/5 & 1/5 & 1/5 \\ 0 & 1/5 & 0 \end{bmatrix}$$

Determine para cada $j \in \{1,2,3\} : P[X = j]$ *y* $P[Y = j]$.

Solución:

Es implícito del enunciado que X e Y son funciones medibles de $\mathcal{A}$ en $\mathcal{P}(\{1,2,3\})$ y por lo tanto

$$(\forall j \in \{1,2,3\}) : [X = j] := X^{-1}(\{j\}) \in \mathcal{A} \quad , \quad [Y = j] := Y^{-1}(\{j\}) \in \mathcal{A}$$

Por otro lado es evidente que:

$$\Omega = \{\omega \in \Omega : X(\omega) \in \{1,2,3\}\} = \{\omega \in \Omega : Y(\omega) \in \{1,2,3\}\}$$

y por lo tanto

$$\Omega \;=\; [X=1] \cup [X=2] \cup [X=3] \;=\; [Y=1] \cup [Y=2] \cup [Y=3] \tag{1}$$

Consideremos $j \in \{1,2,3\}$ fijo, de (1) vemos que:

$$[X=j] = \bigcup_{i=1}^{3}([X=j] \cap [Y=i]) = \bigcup_{i=1}^{3}[X=j, Y=i]$$

de donde se concluye que

$$(\forall j \in \{1,2,3\}): \quad P[X=j] = \sum_{i=1}^{3} P[X=j, Y=i] = \sum_{i=1}^{3} p_{ji} \tag{2}$$

De manera análoga se demuestra que

$$(\forall j \in \{1,2,3\}): \quad P[Y=j] = \sum_{i=1}^{3} P[X=i, Y=j] = \sum_{i=1}^{3} p_{ij} \tag{3}$$

Finalmente en virtud de (2) y (3) obtenemos que:

$$\begin{array}{ccccc} P[X=1]=1/5 & , & P[X=2]=3/5 & , & P[X=3]=1/5 \\ P[Y=1]=1/5 & , & P[Y=2]=3/5 & , & P[Y=3]=1/5 \end{array}$$

y por lo tanto $(\forall j \in \{1,2,3\}) : P[X=j] = P[Y=j]$. Ésto no significa necesariamente que $(\forall \omega \in \Omega) : X(\omega) = Y(\omega)$, sólo quiere decir que si (E_x) (respec. (E_y)) es el experimento consistente en evaluar la función X (respec. Y) en un punto extraído al azar en Ω siguiendo la ley de probailidad P, entonces los experimentos (E_x) y (E_y) están regidos por la misma ley probabilística i.e. un observador que observe la realización consecutiva de uno solo de éstos experimentos sin saber si se trata de (E_x) o (E_y), no sería capaz de distinguirlos y los vería como parte de un mismo fenónemo aleatorio.

Problema IV.1.7

Dadas $X_1, ..., X_n$ variables aleatorias independientes de Poisson con parámetros $\lambda_1, ..., \lambda_n$ respectivamente.

Muestre que $Y = \sum_{i=1}^{n} X_i$ es una variable aleatoria de Poisson de parámetro $\sum_{i=1}^{n} \lambda_i$

Solucion: Se probará por inducción sobre n.

(Caso $n = 2$)

Como las variables X_1, X_2 son independientes, la probabilidad condicional de X_2 con respecto a X_1 es igual a la probabilidad de X_2. Además ambas variables tienen su soporte en $I\!N$. Dado $k \in I\!N$ Al aplicar probabilidades totales se obtiene:

$$I\!P(X_1 + X_2 = k) = I\!P(X_2 = k - X_1) = \sum_{m=0}^{\infty} I\!P\left[X_2 = k - X_1/X_1 = m\right] I\!P(X_1 = m)$$

Es decir:

$$I\!P(X_1 + X_2 = k) = \sum_{m=0}^{\infty} I\!P(X_2 = k - m) I\!P(X_1 = m)$$

Como $I\!P(X_i = k) = e^{\lambda_i} \frac{\lambda_i^k}{k!}$ cuando $k \in I\!N$ y 0 en cualquier otro caso, al reemplazar tendremos que:

$$I\!P(X_1 + X_2 = k) = \sum_{m=0}^{k} e^{-(\lambda_1+\lambda_2)} \frac{\lambda_1^m \lambda_2^{k-m}}{(k-m)!}$$

Al multiplicar por $\frac{k!}{k!}$ se obtiene:

$$I\!P(X_1 + X_2 = k) = \frac{e^{-(\lambda_1+\lambda_2)}}{k!} (\lambda_1 + \lambda_2)^k$$

Esto debido a que por el teorema del Binomio de Newton se sabe que:

$$\sum_{m=0}^{k} \binom{k}{m} \lambda_1^m \lambda_2^{k-m} = (\lambda_1 + \lambda_2)^k$$

Esto muestra que:

$$I\!P(X_1 + X_2 = k) = \frac{e^{-(\lambda_1+\lambda_2)}}{k!} (\lambda_1 + \lambda_2)^k$$

Es decir $X_1 + X_2$ tiene distribución de Poisson de parámetro $\lambda_1 + \lambda_2$.

($n \Rightarrow n + 1$)

Si $X_1, ..., X_{n+1}$ son independientes $\Rightarrow$ $X_1, ..., X_n$ son independientes. Al aplicar la hipótesis de inducción tengo que $Y = \sum_{i=1}^{n} X_i$ es una variable aleatoria de Poisson de parámetro $\sum_{i=1}^{n} \lambda_i$.

Siendo Y función solo de $X_1, ..., X_n$ entonces será independiente de X_{n+1}, luego aplicando nuevamente la hipótesis de inducción para el caso $n = 2$ obtenemos que $X_1 + ... + X_{n+1}$ es una variable aleatoria de Poisson de parámetro $\sum_{i=1}^{n} \lambda_i + \lambda_{n+1}$.

Observación: Este es uno de los muchos ejemplos de familias de variables aleatorias independientes que son cerradas bajo la suma. Es decir la suma de variables aleatorias en la familia es otro elemento de la misma familia. Un ejercicio relacionado sería encontrar otra familia con la misma propiedad, por ejemplo variables aleatorias normales con media y varianza desconocidas.

Problema IV.1.8

Sea $X_1, ..., X_n$ una familia de variables aleatorias exponenciales independientes con esperanzas $\alpha_1, ..., \alpha_n$ respectivamente.

(a) Pruebe que $Y = \min_{i=1,\ldots,n} X_i$ es variable aleatoria exponencial. ¿Cuál es su parámetro?

(b) Muestre que $I\!P\left[X_k = \min_{i=1,\ldots,n} X_i\right] = \frac{\alpha_k}{\alpha_1 + \ldots + \alpha_n}$

Solución:

(a) Calculemos directamente la distribución $F_Y(y) = I\!P(Y \leq y) = 1 - I\!P(Y > y)$. Siempre que $y \geq 0$.

Es fácil convencerse que $\min_{i=1,\ldots,n} X_i > y \Leftrightarrow X_i > y \quad \forall i = 1, ..., n$.

Esto prueba que $I\!P(Y > y) = I\!P(X_1 > y, ..., X_2 > y)$, por independencia de los X_i tendremos que:

$$F_Y(y) = 1 - \prod_{i=1}^{n} I\!P(X_i > y)$$

Además como $I\!P(X_i > y) = \int_y^\infty \alpha_i \cdot e^{-\alpha_i x} dx = \frac{e^{-\alpha_i x}}{\alpha_i} \alpha_i \Big|_\infty^y = \frac{e^{-\alpha_i y}}{\alpha_i} \cdot \alpha_i$

Al reemplazar tendremos que $F_Y(y) = 1 - e^{-\sum_{i=1}^{n} \alpha_i \cdot y} = \int_0^y (\sum_{i=1}^{n} \alpha_i) e^{-\sum_{i=1}^{n} \alpha_i x} dx$.

Es decir la variable Y es exponencial de parámetro $\sum_{i=1}^{n} \alpha_i$.

(b) Es claro que $X_k = \min_{i=1,\ldots,n} X_i \Leftrightarrow X_k \leq \min_{i \neq k} X_i$.

Además X_k y $\min_{i \neq k} X_i$ son independientes exponenciales de parámetro α_k y $\sum_{i \neq k} \alpha_i$ respectivamente.

De manera general si X e Y son variables aleatorias exponenciales independientes de parámetros α y β. Entonces $I\!P(X \leq Y) = \int\limits_0^\infty \int\limits_0^y \alpha\beta e^{-\alpha x} e^{-\beta y} dx dy = \int\limits_0^\infty \beta e^{-\beta y}(1 - e^{-\alpha y}) dy$

Por lo tanto $I\!P(X \leq Y) = 1 - \frac{\beta}{(\alpha+\beta)} = \frac{\alpha}{\alpha+\beta}$.

En nuestro caso particular tendremos que $I\!P\left[X_k \leq \min\limits_{i \neq k} X_i\right] = \frac{\alpha_k}{\alpha_k + \sum\limits_{i \neq k} \alpha_i} = \frac{\alpha_k}{\alpha_1 + \ldots + \alpha_n}$

IV.2 Funciones de Distribución y densidades de Variables Aleatorias

Problema IV.2.1

Sea $(\Omega, \mathcal{F}, I\!P)$ espacio de probabilidad y $T : \Omega \to \Omega$ función medible sobreyectiva tal que:

$$\forall A \in \mathcal{F} \quad I\!P\big[T^{-1}(A)\big] = I\!P(A), \text{ es decir } I\!P\text{-invariante}$$

Sean $J_0 = \{A \in \mathcal{F} / T^{-1}(A) = A\}$ y $J = \{A \in \mathcal{F} / I\!P\big[T^{-1}(A) \triangle A\big] = 0\}$.
Muestre que J_0 y J son σ-álgebras.

Solución:

Primero veamos que J_0 es σ-álgebra.

Como $T^{-1}(\phi) = \phi$ y $T^{-1}(\Omega) = \Omega \Rightarrow \phi, \Omega \in J_0$

Dado $A \in J_0 \quad T^{-1}(A^c) = T^{-1}(A)^c = A^c \Rightarrow A^c \in J_0$

Dados $A, B \in J_0 \Rightarrow T^{-1}(A \cap B) = T^{-1}(A) \cap T^{-1}(B) = A \cap B \Leftrightarrow A \cap B \in J_0$

Dado $\{A_\lambda\}_{\lambda \in \Lambda} \subseteq J_0 \Rightarrow T^{-1}\left(\bigcup\limits_{\lambda \in \Lambda} A_\lambda\right) = \bigcup\limits_{\lambda \in \Lambda} T^{-1}(A_\lambda) = \bigcup\limits_{\lambda \in \Lambda} A_\lambda \Rightarrow \bigcup\limits_{\lambda \in \Lambda} A_\lambda \in J_0$

Lo anterior muestra que J_0 es σ-álgebra.

Para probarlo para J recordemos que $A \triangle B = (A \setminus B) \cup (B \setminus A)$ [Diferencia simétrica].

Entonces $\phi \triangle \phi = \phi$ y $\Omega \triangle \Omega = \phi \Rightarrow \phi, \Omega \in J$.

Una representación equivalente es, $A \triangle B = (A \cup B) \setminus (A \cap B)$.

Dados $A, B \in J$:

$$T^{-1}(A\cap B)\triangle(A\cap B) = \left[T^{-1}(A\cap B)\setminus(A\cap B)\right]\cup\left[(A\cap B)\setminus T^{-1}(A\cap B)\right]$$

$$\begin{aligned}
T^{-1}(A\cap B)\setminus(A\cap B) &= T^{-1}(A)\cap T^{-1}(B)\cap(A^c\cup B^c)\\
&= \left(T^{-1}(B)\cap\left[T^{-1}(A)\setminus A\right]\right)\cup\left(T^{-1}(A)\cap\left[T^{-1}(B)\setminus B\right]\right)\\
(A\cap B)\setminus T^{-1}(A\cap B) &= A\cap B\cap\left[T^{-1}(A)^c\cup T^{-1}(B)^c\right]\\
&= \left[B\cap\left(A\setminus T^{-1}(A)\right)\right]\cup\left[A\cap\left(B\setminus T^{-1}(B)\right)\right]
\end{aligned}$$

Como $I\!P(A\cup B)\leq I\!P(A)+I\!P(B) \quad \forall A,B\in\mathcal{F}$ las igualdades anteriores muestran que:

$$I\!P\left[A\setminus T^{-1}(A)\right] = I\!P\left[T^{-1}(A)\setminus A\right] = I\!P\left[T^{-1}(B)\setminus B\right] = I\!P\left[B\setminus T^{-1}(B)\right] = 0$$

Aseguran que:

$$I\!P\left[T^{-1}(A\cap B)\triangle(A\cap B)\right] = 0$$

Solo resta ver que se tienen esas condiciones.

$A,B\in J\Rightarrow I\!P\left[T^{-1}(A)\triangle A\right] = I\!P\left[T^{-1}(B)\triangle B\right] = 0$

Por definición:

$$I\!P\left[T^{-1}(A)\triangle A\right] = I\!P\left[T^{-1}(A)\setminus A\right] + I\!P\left[A\setminus T^{-1}(A)\right]$$

$$\Rightarrow I\!P\left[T^{-1}(A)\setminus A\right] = I\!P\left[A\setminus T^{-1}(A)\right] = 0$$

De manera análoga se prueba para B, luego $A\cap B\in J$

Dado $A\in J$

$$\begin{aligned}
T^{-1}(A)^c\triangle A^c &= [A^c\cup T^{-1}(A^c)]\setminus[A^c\cap T^{-1}(A)^c] = [A^c\cup T^{-1}(A)^c]\cap[A\cup T^{-1}(A)]\\
&= [T^{-1}(A)\cap A^c]\cup[A\cap T^{-1}(A)^c]\\
&= [T^{-1}(A)\setminus A]\cup[A\setminus T^{-1}(A)]\\
&= [T^{-1}(A)\triangle A]
\end{aligned}$$

Entonces $I\!P\left[T^{-1}(A^c)\triangle A^c\right] = 0$ ssi $I\!P\left[T^{-1}(A)\triangle A\right] = 0 \Rightarrow A^c\in J$

Dado $\{A_\lambda\}_{\lambda\in\Lambda}\subseteq J$

$$T^{-1}\left(\bigcup_{\lambda\in\Lambda}A_\lambda\right)\triangle\left(\bigcup_{\lambda\in\Lambda}A_\lambda\right) = \left[T^{-1}\left(\bigcup_{\lambda\in\Lambda}A_\lambda\right)\cap\bigcap_{\lambda\in\Lambda}A_\lambda^c\right]\cup\left[\bigcup_{\lambda\in\Lambda}A_\lambda\cap\bigcap_{\lambda\in\Lambda}T^{-1}(A_\lambda^c)\right]$$

Sabemos que $\forall\lambda\in\Lambda \quad I\!P\left[T^{-1}(A_\lambda)\triangle A_\lambda\right] = 0$

Es decir $I\!P\left[T^{-1}(A_\lambda) \setminus A_\lambda\right] = I\!P\left[A_\lambda \setminus T^{-1}(A_\lambda)\right] = 0$

Como $\bigcap\limits_{\lambda\in\Lambda} A_\lambda^c \subseteq A_{\lambda_0}^c \forall\lambda_0 \in \Lambda \Rightarrow T^{-1}\left(\bigcup\limits_{\lambda\in\Lambda} A_\lambda\right) \cap \bigcap\limits_{\lambda\in\Lambda} A_\lambda^c \subseteq \bigcup\limits_{\lambda\in\Lambda}\left[T^{-1}(A_\lambda) \cap A_\lambda^c\right]$

Por lo tanto $I\!P\left[T^{-1}\left(\bigcup\limits_{\lambda\in\Lambda} A_\lambda\right) \cap \bigcup\limits_{\lambda\in\Lambda} A_\lambda^c\right] = 0$

De manera análoga $\bigcup\limits_{\lambda\in\Lambda} A_\lambda \cap \bigcup\limits_{\lambda\in\Lambda} T^{-1}(A_\lambda)^c \subseteq \bigcup\limits_{\lambda\in\Lambda}[A_\lambda \cap T^{-1}(A_\lambda)^c]$

Luego $I\!P\left[\bigcup\limits_{\lambda\in\Lambda} A_\lambda \cap \bigcup\limits_{\lambda\in\Lambda} T^{-1}(A_\lambda)^c\right] = 0$ con lo que:

$$I\!P\left[T^{-1}\left(\bigcup_{\lambda\in\Lambda} A_\lambda\right) \triangle \left(\bigcup_{\lambda\in\Lambda} A_\lambda\right)\right] = 0 \text{ es decir } \bigcup_{\lambda\in\Lambda} A_\lambda \in J$$

Por lo tanto J es σ-álgebra.

Observación: Además como $[T^{-1}(A) = A \Rightarrow I\!P(T^{-1}(A)\triangle A) = 0] \Rightarrow J_0 \subseteq J$.

Problema IV.2.2

Si Y sigue una distribución uniforme en $[0,5]$, ¿ Cuál es la probabilidad de que ambas raíces de la ecuación

$$4x^2 + 4xY + Y = 0$$

sean reales?

Solución:

Por definición, si Y es una v.a. con distribución uniforme en $[0,5]$ entonces Y es v.a. con densidad

$$(\forall t \in I\!R): \quad f_Y(t) = \frac{1}{5}1_{[0,5]}(t)$$

Al considerar que Y es una función medible que está definida en cierto espacio de probabilidad Ω y a valores en $I\!R$, se tendrá que dado $\omega \in \Omega$ la ecuación en la indeterminada x:

$$4x^2 + 4xY(\omega) + Y(\omega) = 0$$

tiene ambas de sus raíces reales si y sólo si

$$[4Y(\omega)]^2 - 4\cdot 4\cdot Y(\omega) \geq 0$$

i.e.

$$Y(\omega)\cdot[Y(\omega) - 1] \geq 0$$

De lo anterior, concluimos que al definir en Ω el suceso

$$R := \{\omega \in \Omega : \text{ la ecuación } 4x^2 + 4xY(\omega) + Y(\omega) = 0 \text{ tiene ambas raíces reales } \}$$

entonces

$$\begin{aligned} R &= \{\omega \in \Omega : Y(\omega)[Y(\omega) - 1] \geq 0\} \\ &= \{\omega \in \Omega : Y(\omega) \geq 0 \wedge Y(\omega) \geq 1\} \cup \{\omega \in \Omega : Y(\omega) \leq 0 \wedge Y(\omega) \leq 1\} \end{aligned} \tag{1}$$

Pero como

$$P\left(\{\omega \in \Omega : Y(\omega) \leq 0 \wedge Y(\omega) \leq 1\}\right) = P(\{\omega \in \Omega : Y(\omega) \leq 0\}) = \int_{-\infty}^{0} f_Y(t) = 0$$

de (1) concluimos que:

$$\begin{aligned} P(R) &= P(\{\omega \in \Omega : Y(\omega) \geq 0 \wedge Y(\omega) \geq 1\}) = P(\{\omega \in \Omega : Y(\omega) \geq 1\}) \\ &= P[Y \in [1, +\infty]] = \int_{1}^{+\infty} f_Y(t)dt = \int_{1}^{5} \frac{1}{5}dt = \frac{4}{5} \end{aligned}$$

Antes de finalizar notamos que la igualdad (1), asegura que al conjunto R tiene sentido calcularle probabilidad. En efecto, si en Ω se encuentra definida una σ-álgebra $\mathcal{A}$ respecto de la cual la función $Y : \Omega \to \mathbb{R}$ es medible, entonces de (1) es directo que:

$$R = [Y \geq 1] \cup [Y \leq 0] \tag{2}$$

Como Y es función medible se tiene que $[Y \geq 1] \in \mathcal{A}$ e $[Y \leq 0] \in \mathcal{A}$, de donde por (2) resulta evidente que $R \in \mathcal{A}$.

Problema IV.2.3

Dadas X, Y, Z variables aleatorias independientes con distribución uniforme en $[0,1]$. ¿Cuál es la probabilidad que la ecuación cuadrática $Xt^2 + Yt + Z = 0$ tenga raices reales?

Solución:

Por independencia la densidad conjunta del vector (X, Y, Z) será:

$$f_{XYZ}(x, y, z) = f_X(x)f_Y(y)f_Z(z) = \mathbb{1}_{[0,1]}(x) \cdot \mathbb{1}_{[0,1]}(y) \cdot \mathbb{1}_{[0,1]}(z)$$

$Xt^2 + Yt + Z$ tendrá raíces reales si y solo si $Y^2 - 4XZ \geq 0$.

Por lo tanto la solución es calcular el area de la intersección en $\mathbb{R}^3$ entre el cubo $[0,1]^3$ y la superficie $Y^2 - 4XZ \geq 0$ ya que hay una densidad uniforme en el cubo.

Como $X, Y, Z \geq 0$ entonces al despejar de $Y^2 - 4XZ \geq 0$ obtenemos $Y \geq 2\sqrt{XZ}$.

Para poder establecer un parametrización del dominio, de ahora en adelante asumiremos que $Y \geq 2\sqrt{XZ}$ y que X, Y, $Z \in [0,1]$.

Para que $Y \in [0,1] \Rightarrow \quad 2\sqrt{XZ} \leq 1$

Mas precisamente $Y \in [2\sqrt{XZ}, 1]$ cuando $2\sqrt{XZ} \leq 1$

Claramente tendremos que suponer además que $2\sqrt{XZ} \leq 1$

$$2\sqrt{XZ} \leq 1 \Rightarrow XZ \leq \frac{1}{4} \Rightarrow Z(X) \leq \frac{1}{4X}$$

$$\begin{cases} Z \in [0, \frac{1}{4X}] \text{ si } X \in [\frac{1}{4}, 1] \\ Z \in [0,1] \quad \text{ si } X \in [0, \frac{1}{4}] \end{cases}$$

Con el razonamiento anterior estamos en condiciones de dar una parametrización de la superficie. Esta será:

$$([0, \frac{1}{4}], [2\sqrt{XZ}, 1], [0,1]) \cup ([\frac{1}{4}, 1], [2\sqrt{XZ}, 1], [0, \frac{1}{4X}])$$

En consecuencia la probabilidad a calcular será:

$$\begin{aligned} \mathbb{P}[Y^2 - 4XZ \geq 0] &= \int_0^{\frac{1}{4}} \int_0^1 \int_{2\sqrt{XZ}}^1 dydzdx + \int_{\frac{1}{4}}^1 \int_0^{\frac{1}{4X}} \int_{2\sqrt{XZ}}^1 dydzdx \\ &= \int_0^{\frac{1}{4}} \int_0^1 (1 - 2\sqrt{XZ}) dzdx + \int_{\frac{1}{4}}^1 \int_0^{\frac{1}{4X}} (1 - 2\sqrt{XZ}) dzdx \\ &= \int_0^{\frac{1}{4}} (1 - \frac{4}{3}\sqrt{X}) dx + \int_{\frac{1}{4}}^1 (\frac{1}{4X} - \frac{1}{6X}) dx \\ &= \frac{1}{4} - \frac{4}{3} \cdot \frac{2}{3} \cdot (\frac{1}{4})^{\frac{3}{2}} + \frac{1}{12} \ln 4 \approx 0.25441342 \end{aligned}$$

Problema IV.2.4

Sea X una v.a. con función de densidad f_X. Se define la v.a. $Y := \frac{X-a}{b}$ donde $a, b \in I\!R$ están dados y $b > 0$.Demuestre que la v.a. Y también tiene función de densidad.

Solución:

Debemos encontrar una función $f_Y : I\!R \to I\!R_+ \cup \{0\}$ integrable en $I\!R$ tal que

$$(\forall y \in I\!R) : F_Y(y) = \int_{-\infty}^{y} f_Y(u)du$$

donde F_Y es la función de distribución de la v.a. Y.

En efecto, dado $y \in I\!R$ tiene que

$$F_Y(y) := P[\frac{X-a}{b} \leq y] = P[X - a \leq y \cdot b] = P[X \leq y \cdot b + a]$$

y por lo tanto

$$F_Y(y) = \int_{-\infty}^{y \cdot b + a} f_X(v)dv \tag{1}$$

Consideremos el cambio de variable lineal $v = u \cdot b + a$. La expresión en (1) queda:

$$F_Y(y) = \int_{-\infty}^{y} b \cdot f_X(u \cdot b + a)du$$

Al definir $(\forall y \in I\!R) : \quad f_Y(y) := b \cdot f_X(y \cdot b + a)$ de lo anterior obtenemos que:

$$(\forall y \in I\!R) : \quad F_Y(y) = \int_{-\infty}^{y} f_Y(y)dy$$

lo que demuestra que la v.a. Y tiene función de densidad.

Para finalizar destacamos que la medibilidad de la función Y está garantizada por la medibilidad de la v.a. X, y además la integrabilidad en $I\!R$ de la función f_Y es consecuencia de la integrabilidad en $I\!R$ de la función f_X.

Problema IV.2.5

Sea X una v.a. simétrica en torno a cero i.e. $(\forall x > 0) : P[X > x] = P[X < -x]$. Muestre que si X es v.a con función de densidad f_X entonces la v.a definida por $Y := X^2$ tiene función de densidad dada por

$$f_Y(y) := \begin{cases} 0 & , y \leq 0 \\ y^{-\frac{1}{2}} f_X(y^{\frac{1}{2}}) & , y > 0 \end{cases}$$

Solución:

Demostremos primero que:

(1) Si X es v.a. simétrica en torno a cero entonces

$$(\forall x \geq 0) : \quad P[0 \leq X \leq x] = P[-x \leq X \leq 0]$$

(2) Si X es v.a. simétrica en torno a cero con función de densidad entonces

$$(\forall x \geq 0) : \quad P[|X| \leq x] = 2P[0 \leq X \leq x]$$

En efecto, si X es v.a. simétrica en torno a cero entonces $P[X > 0] = P[X < 0]$ y por lo tanto $P[X \leq 0] = P[X \geq 0]$. Dado $x > 0$ se verifica que:

$$\begin{aligned} P[0 \leq X \leq x] &= P[X \geq 0] - P[X > x] \\ &= P[X \leq 0] - P[X < -x] \\ &= P[-x \leq X \leq 0] \end{aligned}$$

luego (1) es válido si $x > 0$, como (1) es evidente si $x = 0$ concluimos .

Por otro lado, si X es v.a. simétrica en torno a cero con función de densidad, dado $x \geq 0$ es claro que:

$$P[|X| \leq x] = P[-x \leq X \leq x] = P[-x \leq X < 0] + P[0 \leq X \leq x] \tag{3}$$

Como X tiene función de densidad en particular se tiene que $P[X = 0] = 0$, luego de (3) obtenemos que:

$$P[|X| \leq x] = \{P[-x \leq X < 0] + P[X = 0]\} + P[0 \leq X \leq x]$$

i.e.

$$P[|X| \leq x] = P[-x \leq X \leq 0] + P[0 \leq X \leq x]$$

de donde por (1) concluimos que

$$(\forall x \geq 0) : \quad P[|X| \leq x] = 2P[0 \leq X \leq x]$$

lo que demuestra (2).

Procedamos a demostrar que Y tiene función de densidad. Notemos que $P[Y \geq 0] = 1$ luego

$$(\forall y < 0): \quad F_Y(y) := P[Y \leq y] = 0 \tag{4}$$

Por otro lado, si $y \geq 0$ entonces

$$F_Y(y) := P[Y \leq y] = P[X^2 \leq y] = P[|X| \leq \sqrt{y}]$$

de donde al denotar por f_X a la función de densidad de la v.a X, por (2) obtenemos que

$$(\forall y \geq 0): \quad F_Y(y) = 2P[0 \leq X \leq \sqrt{y}] = 2\int_0^{\sqrt{y}} f_X(u)du \tag{5}$$

Al hacer en (5) el cambio de variable $v = u^2$ obtenemos que

$$(\forall y \geq 0): \quad F_Y(y) = \int_0^{y} v^{-\frac{1}{2}} f_X(v^{\frac{1}{2}})dv \tag{6}$$

Al definir entonces

$$f_Y(y) := \begin{cases} 0 & , y \leq 0 \\ y^{-\frac{1}{2}} f_x(y^{\frac{1}{2}}), & y > 0 \end{cases}$$

es fácil convencerse a partir de (4) y (6) que

$$(\forall y \in I\!R): F_Y(y) = \int_{-\infty}^{y} f_Y(u)du$$

y por lo tanto la v.a. Y tiene función de densidad.

Antes de finalizar destacamos que la medibilidad de la función Y está garantizada por la medibilidad de la v.a. X. Además la integrabilidad en $I\!R$ de la función f_Y es consecuencia de la integrabilidad en $I\!R$ de f_X y de la igualdad de las integrales que permiten pasar de las expresiones (5) a la (6).

Problema IV.2.6

Sea X una v.a. con función de distribución F_X y sea $c \in I\!R$ una constante dada. Se define la v.a. $Y := \max\{X, c\}$. Se pide que:

(a) Dé una expresión para la función de distribución de la v.a. Y a partir de la función de distribución de la v.a. X .

(b) Demuestre que si X es v.a. con densidad y $P[X \leq c] = 0$ entonces Y también tiene función de densidad ¿Qué puede decir si $P[X \leq c] > 0$? .

Solución:

(a) En lo que sigue denotaremos por F_Y a la función de distribución de la v.a. Y, cuya medibilidad está garantizada por la medibilidad de la v.a X. Claramente si $X : \Omega \to I\!R$ entonces

$$(\forall \omega \in \Omega) : \quad Y(\omega) := \max\{X(\omega), c\} \geq c$$

luego

$$(\forall y < c) : P[Y \leq y] \leq P[Y < c] = 0$$

de donde

$$(\forall y < c) : F_Y(y) = 0 \tag{1}$$

Por otro lado, si $y \geq c$ es claro que:

$$Y(\omega) \leq y \;\Leftrightarrow\; \max\{X(\omega), c\} \leq y \;\Leftrightarrow\; X(\omega) \leq y$$

luego

$$(\forall y \geq c) : F_Y(y) = F_X(y) \tag{2}$$

De (1) y (2) se concluye que:

$$F_Y(y) = \begin{cases} 0 & , \; y < c \\ F_X(y), & y \geq c \end{cases} \tag{3}$$

(b) Denotemos por f_X a la función de densidad de la v.a. X, entonces dado que $P[X \leq c] = 0$ se tendrá de (3) que

$$\begin{aligned} (\forall y \geq c) : \quad F_Y(y) &= F_X(y) := P[X \leq y] \\ &= P[X \leq c] + P[c < X \leq y] \\ &= P[c < X \leq y] \end{aligned} \tag{4}$$

Como X es v.a. con densidad entonces $(\forall x \in I\!R) : P[X = x] = 0$, en particular $P[X = c] = 0$, y por lo tanto de (4) obtenemos que

$$(\forall y \geq c) : F_Y(y) = P[c < x \leq y] + P[X = c] = P[c \leq X \leq y]$$

de donde

$$(\forall y \geq c) : \quad F_Y(y) = P[X \in [c, y]] = \int_c^y f_X(u)du \tag{5}$$

Al definir entonces $f_Y : I\!R \to I\!R_+ \cup \{0\}$ por

$$f_Y(u) = \begin{cases} 0 & , u < c \\ f_X(u), & u \geq c \end{cases}$$

se tendrá que f_Y es una función integrable en $I\!R$ y de (5)

$$(\forall y \geq c): \quad F_Y(y) = \int_{-\infty}^{y} f_Y(u)du$$

Como además

$$(\forall y < c): \quad F_Y(y) = 0 = \int_{-\infty}^{y} f_Y(u)du$$

concluimos que

$$(\forall y \in I\!R): \quad F_Y(y) = \int_{-\infty}^{y} f_Y(u)du$$

lo que demuestra que la v.a. Y tiene función de densidad.

En el caso que $P[X \leq c] > 0$, a partir de (3) obtenemos que

$$F_Y(c) = F_X(c) := P[X \leq c] > 0$$

Como $\quad (\forall y < c) : F_Y(y) = 0 \quad$ es directo que $\quad F_Y(c^-) = 0 \quad$ luego

$$P[Y = c] = F_Y(c) - F_Y(c^-) = F_Y(c) > 0 \tag{6}$$

y por lo tanto la v.a Y no puede tener función de densidad pues de tenerla debiera en particular tenerse que $P[Y = c] = 0$ lo que está en contradicción con (6).

Problema IV.2.7

Sean $X_1, ..., X_n$ variables aleatorias con densidad común de Rayleigh con parámetro $\theta > 0$, dada por:

$$f(x) = \frac{x}{\theta^2} e^{-x^2/2\theta^2} 1\!\!1_{x>0}(x)$$

(a) Determine la densidad conjunta de $Y_1, ..., Y_n$ con $Y_i = X_i^2 \quad \forall i = 1, ..., n$.

(b) ¿Cuál es la distribución de $U = \min\limits_{1 \leq i \leq n} X_i$?

Solución:

(a)

Calculemos directamente la distribución de $Y = X^2$ con X variable aleatoria Rayleigh de parámetro θ.

$$\begin{aligned}
I\!P[Y \leq y] &= I\!P[|x| \leq \sqrt{y}] = I\!P[-\sqrt{y} \leq x \leq \sqrt{y}] \\
&= I\!P[0 \leq x \leq \sqrt{y}] \quad (\text{ Esto pues X es positiva }) \\
F_X(y) &= \int_0^{\sqrt{y}} \frac{x}{\theta^2} e^{-x^2/2\theta^2 e^2} dx = -e^{-\frac{x^2}{2\theta^2}} |_0^{\sqrt{y}} \\
&= (1 - e^{-y/2\theta^2}) \quad (y \geq 0)
\end{aligned}$$

Por lo tanto la ley de la variable aleatoria Y es exponencial.

Si los X_i son independientes $\Rightarrow X_i^2$ son independientes luego:

$$f_{Y_1,...,Y_n}(y_1, ..., y_n) = \prod_{i=1}^{n} \left[\frac{1}{2\theta^2} e^{-y_i/2\theta^2} \right] \text{ con } y_i \geq 0 \quad \forall i = 1, ..., n$$

(b) Ahora escribimos U en términos de los X_i.

$$\begin{aligned}
I\!P[U \leq u] &= 1 - I\!P[U > u] = 1 - I\!P\left[\min_{1 \leq i \leq n} X_i > u\right] \\
&= 1 - I\!P[X_1 > u, X_2 > u, ..., X_n > u] \\
&= 1 - I\!P[X_1 > u] \cdots I\!P[X_n > u] = 1 - I\!P[X_1 > u]^n
\end{aligned}$$

$\uparrow$ Indep. $\qquad\qquad\qquad$ $\uparrow$ igual densidad

Calculemos explícitamente la distribución de una variable aleatoria X con densidad de Raleigh.

$$\begin{aligned}
(u > 0) \quad I\!P[X > u] &= 1 - I\!P[X \leq u] \\
&= 1 - \int_0^u \frac{x}{\theta^2} e^{-x^2/2\theta^2} dx
\end{aligned}$$

$$I\!P[X > u] = 1 + e^{-x^2/2\theta^2} |_0^u = 1 + e^{-u^2/2\theta^2} - 1 = e^{-u^2/2\theta^2}$$

Al reemplazar obtenemos:

$$I\!P[U \leq u] = 1 - e^{-nu^2/2\theta^2}$$

Entonces la distribución de U es también de Rayleigh pero de parámetro $\frac{\theta}{\sqrt{n}}$.

(c) Calculemos la distribución de Z utilizando las probabilidades condicionales.

$$F_Z(z) = I\!P[Z \leq z] = I\!P\left[\frac{X_1}{X_2} \leq z\right] = I\!P[X_1 \leq zX_2]$$
$$= \int_0^\infty I\!P\left[X_1 \leq zX_2/X_2 = x_2\right] f_{X_2}(x_2)dx_2$$

Sabemos que:

$$I\!P\left[X_1 \leq zX_2/X_2 = x_2\right] = \int_0^{zx_2} f_{X_1/X_2}(s, x_2)ds$$

Como X_1 y X_2 son independientes la distribución condicional será:

$$f_{X_1/X_2}(a/b) = \frac{f_{X_1X_2}(a,b)}{f_{X_2}(b)} = f_{X_1}(a)$$
$$I\!P[Z \leq z] = \int_0^\infty \int_0^{zx_2} \frac{s}{\theta^2} e^{-s^2/2\theta^2} ds \frac{x_2}{\theta^2} e^{-x_2^2/2\theta^2} dx_2$$
$$\int_0^\infty (e^{-s^2/2\theta^2}|_{zx_2}^{0}) \frac{x_2}{\theta^2} e^{-x_2^2/2\theta^2} dx_2 = \int_0^\infty \frac{x_2}{\theta^2} e^{-x_2^2/2\theta^2} (1 - e^{-z^2x_2^2/2\theta^2}) dx_2$$
$$= \int_0^\infty \frac{x_2}{\theta^2} e^{-x_2^2/2\theta^2} dx_2 - \int_0^\infty \frac{x_2}{\theta^2} e^{-x_2^2(z^2+1)/2\theta^2} dx_2$$

La función densidad debe integrar 1 para cualquier parámetro por lo que sabemos que:

$$\int_0^\infty \frac{x}{\theta^2} e^{-x^2/2\theta^2} dx = 1 \quad \forall \theta \in I\!R_+$$

Consideremos $\theta^{*2} = \frac{\theta^2}{(z^2+1)} \Rightarrow (z^2+1)\int_0^\infty \frac{x}{\theta^2} e^{-x^2(z^2+1)/2\theta^2} dx = 1$

Reemplazando tendremos que:

$$I\!P[Z \leq z] = 1 - \frac{1}{(z^2+1)} = \frac{z^2}{(z^2+1)}$$

Observación: Sería un buen ejercicio hacer el mismo cálculo utilizando el método del Jacobiano. Otra pregunta interesante sería cambiar el min de la parte (a) por un max. Sigue siendo una variable aleatoria de Rayleigh ?

Problema IV.2.8

Suponga que $X_1, ..., X_n$ son independientes e identicamente distribuidas con densidad común f. Muestre que la densidad conjunta de $U = \min\limits_{1\le i\le n} X_i$ y $V = \max\limits_{1\le i\le n} X_i$ es:

$$f_{U,V}(u,v) = n(n-1)[F(v) - F(u)]^{n-2} f(u) f(v) \mathbb{1}_{\mathbb{R}_-}(u-v)$$

Solución:

1. Cálculo de $\mathbb{P}(U \ge u, V \le v)$.

$$\mathbb{P}\left[\min_{1\le i\le n} X_i \ge u, \max_{1\le i\le n} X_i \le v\right] = \mathbb{P}[X_i \ge u \quad \forall i = 1, ..., n X_i \le v \quad \forall i = 1, ..., n]$$

$$\Rightarrow \mathbb{P}[U \ge u, V \le v] = \mathbb{P}[u \le X_i \le v \quad \forall i = 1, ..., n]$$

   Por independencia e igual distribución

$$\begin{aligned} \mathbb{P}[u \le X_i \le v \quad \forall i = 1, ..., n] &= \mathbb{P}[u \le X_1 \le v]^n \\ &= [F(v) - F(u)]^n \cdot \mathbb{1}_{\mathbb{R}_-}(u-v) \end{aligned}$$

2. Solo consideraremos $u < v$ ya que en otro caso se prueba rápidamente que vale 0.

$$\mathbb{P}[U \ge \bar{u}, V \le \bar{v}] = \int_{\bar{u}}^{\infty} \int_{-\infty}^{\bar{v}} f_{UV}(u,v)\,dv\,du$$

Sea $g(u, \bar{v}) = \int_{-\infty}^{\bar{v}} f_{UV}(u,v)dv \Rightarrow \mathbb{P}[U \ge \bar{u}, V \le \bar{v}] = \int_{\bar{u}}^{\infty} g(u,\bar{v})du$

Al derivar con respecto a $\bar{u}$ y utilizar el teorema fundamental del cálculo tendremos que:

$$\frac{d}{d\bar{u}}\mathbb{P}[U \ge \bar{u}, V \le \bar{v}] = \frac{d}{d\bar{u}} \int_{\bar{u}}^{\infty} g(u,\bar{v})du = -g(\bar{u},\bar{v}) = \int_{-\infty}^{\bar{v}} f_{UV}(u,v)dv \quad (1)$$

Por otro al derivar (0) tendremos que:

$$\frac{d}{d\bar{u}}\mathbb{P}[U \ge \bar{u}, V \le \bar{v}] = \frac{d}{d\bar{u}}[F(\bar{v}) - F(\bar{u})]^n = n[F(\bar{v}) - F(\bar{u})]^{n-1} f(\bar{u}) \quad (2)$$

Al igualar (1) con (2) se deduce que:

$$\int_{-\infty}^{\bar{v}} f_{UV}(u,v)dv = n[F(\bar{v}) - F(\bar{u})]^{n-1} f(\bar{u})$$

Derivando una vez mas, pero ahora con respecto a $\bar{v}$.

$$\frac{d}{d\bar{v}} \int_{-\infty}^{\bar{v}} f_{UV}(u,v)dv = \frac{d}{d\bar{v}} n[F(\bar{v}) - F(\bar{u})]^{n-1} f(\bar{u})$$

$$f(\bar{u},\bar{v}) = n(n-1)[F(\bar{v}) - F(\bar{u})]^{n-2} f(\bar{u}) f(\bar{v}) \quad \bar{u} < \bar{v}$$

Es claro que existe una fórmula similar a esta para el caso de una variable aleatoria discreta, que es además un buen ejercicio para comprender este problema.

IV.3 Método del Jacobiano

Problema IV.3.1

Sea X variable aleatoria exponencial de parámetro X e Y variable aleatoria con distribución uniforme en $[0,2\pi]$.

Asumiendo que X e Y son independientes, calcule la distribución conjunta de:

$$U = \sqrt{X}\cos Y \quad V = \sqrt{X}\sin Y$$

Muestre que U y V son variables aleatorias independientes y determine sus leyes junto con las de:

$$2\lambda(U^2+V^2) \quad y \quad \frac{U}{V}$$

Solución:

La densidad conjunta de X,Y es:

$$f_{XY}(x,y) = \lambda \cdot e^{-\lambda x} \cdot \mathbb{1}_{[0,\infty)}(x) \cdot \frac{1}{2\pi} \cdot \mathbb{1}_{[0,2\pi]}(y)$$

El cambio de variables que transforma las variables X,Y en las variables U,V es:

$$\begin{pmatrix} x \\ y \end{pmatrix} \to \begin{pmatrix} \sqrt{x}\cos y \\ \sqrt{x}\sin y \end{pmatrix} = \begin{pmatrix} u \\ v \end{pmatrix}$$

Es necesario obtener el cambio de variables inverso, es decir que transforma U,V en X,Y, este cambio de variables que notaremos φ es:

$$\begin{pmatrix} u \\ v \end{pmatrix} \overset{\varphi}{\to} \begin{pmatrix} u^2+v^2 \\ arctg(\frac{v}{u}) \end{pmatrix} = \begin{pmatrix} x \\ y \end{pmatrix}$$

El Jacobiano de φ será:

$$D\varphi(u,v) = \begin{bmatrix} 2u & 2v \\ \frac{1}{1+(\frac{v}{u})^2} \cdot \frac{-v}{u^2} & \frac{1}{1+(\frac{v}{u})^2} \cdot \frac{1}{u} \end{bmatrix}$$

Por lo tanto:

$$|detD\varphi(u,v)| = |\tfrac{2u^2}{u^2+v^2} + \tfrac{2v^2}{u^2+v^2}| = 2$$

Entonces utilizando la regla del Jacobiando que en este caso será:

$$f_{UV}(u,v) = |detD\varphi(u,v)| \cdot f_{XY}(x(u,v), y(u,v))$$

Al reemplazar obtenemos:

$$f_{UV}(u,v) = 2 \cdot e^{-\lambda(u^2+v^2)} \cdot \tfrac{1}{2\pi} \cdot \lambda \cdot 1\!\!1_{[0,\infty)}(u^2+v^2) \cdot 1\!\!1_{[0,2\pi]}(arctg(\tfrac{v}{u}))$$

Para simplificar las funciones indicadoras basta notar que:

$$u^2+v^2 \geq 0 \quad \forall u,v \in I\!R$$

$$arctg(\frac{v}{u}) \in [0,2\pi] \quad \forall u,v \in I\!R$$

$$\Rightarrow f_{UV}(u,v) = \frac{\lambda}{\pi} \cdot e^{-\lambda(u^2+v^2)} \quad \forall u,v \in I\!R$$

Como la densidad conjunta de U,V se escribe como un producto de dos funciones, una que solo depende de U y otra que solo depende de V, entonces las variables U y V son independientes. De hecho esto es un si y solo si cuando consideramos variables aleatorias con densidad. Lo único que hay que determinar es como repartir las constantes para que ambas funciones marginales integren 1. En este caso eso se ve inmediatamente por ser ambas densidades iguales.

Entonces $f_U(u) = \sqrt{\frac{\lambda}{\pi}} e^{-\lambda u^2}$ y $f_V(v) = \sqrt{\frac{\lambda}{\pi}} e^{-\lambda v^2}$

Esto demuestra que:

$$\int_{-\infty}^{\infty} e^{-\lambda u^2} du = \sqrt{\frac{\pi}{\lambda}}$$

Observación:

Normalmente lo anterior se verifica para comprobar los cálculos.

Como U^2 y V^2 son independientes la distribución de la suma tendrá la densidad de la convolución de ambas densidades, es decir:

$$f_{U^2+V^2}(k) = \int_{-\infty}^{\infty} f_U(k-s)f_V(s)ds = \int_{-\infty}^{\infty} \frac{\lambda}{\pi} \cdot e^{-\lambda(k-s)^2} \cdot e^{-\lambda s^2} ds$$

$$\Rightarrow f_{U^2+V^2}(k) = \frac{\lambda}{\pi} \int_{-\infty}^{\infty} e^{-\lambda[(k-s)^2+s^2]ds}$$

Para poder desarrollar la expresión se intenta "armar" la densidad de alguna variable aleatoria conocida en la variable de integración.

Como $(k-s)^2 + s^2 = 2s^2 - 2ks + k^2 = 2(s-\frac{k}{2})^2 + k^2 - \frac{k^2}{2} = 2(s-\frac{k}{2})^2 + \frac{k^2}{2}$.

$$\Rightarrow f_{U^2+V^2}(k) = \frac{\lambda}{\pi} e^{-\lambda\frac{k^2}{2}} \cdot \int_{-\infty}^{\infty} e^{-2\lambda(s-\frac{k}{2})^2} ds = \frac{\lambda}{\pi} e^{-\lambda\frac{k^2}{2}} \int_{-\infty}^{\infty} e^{-2\lambda u^2} du$$

$$= \frac{\lambda}{\pi} \cdot \sqrt{\frac{\pi}{2\lambda}} \cdot e^{-\lambda\frac{k^2}{2}}$$

$$\Rightarrow f_{U^2+V^2}(k) = \sqrt{\frac{\lambda}{2\pi}} \cdot e^{-\frac{\lambda}{2}k^2}$$

La suma de los cuadrados de estas variables aleatorias es una variable aleatoria con la misma densidad pero con parámetro $\frac{\lambda}{2}$.

Para calcular la ley de $2\lambda(U^2+V^2)$ basta ver que:

$$I\!P(2\lambda(U^2+V^2) \leq x) = I\!P(U^2+V^2 \leq \frac{x}{2\lambda}) = \int_{-\infty}^{\frac{x}{2\lambda}} \sqrt{\frac{\lambda}{2\pi}} \cdot e^{-\frac{\lambda}{2}k^2} dk$$

Al hacer el cambio de variables $u = 2\lambda \cdot k$ en la integral se tendrá:

$$\int_{-\infty}^{\frac{x}{2\lambda}} \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\lambda}{2}k^2} dk = \int_{-\infty}^{x} \sqrt{\frac{\lambda}{2\pi}} e^{-\frac{\lambda}{2}\cdot\frac{u^2}{2\lambda^2}} \cdot \frac{dx}{2\lambda}$$

Es decir: $f_{2\lambda(U^2+V^2)}(x) = f_{U^2+V^2}(\frac{x}{2\lambda}) \cdot \frac{1}{2\lambda} = \sqrt{\frac{\lambda}{2\pi\cdot(2\lambda)^2}} \cdot e^{-\frac{u^2}{8\lambda}} = \frac{1}{\sqrt{8\lambda\pi}} \cdot e^{-\frac{u^2}{8\lambda}}$

Para calcular la ley de $\frac{U}{V}$ agregamos la variable V y buscamos el cambio de variables inverso al que transforma U,V en V,$\frac{U}{V}$ es decir:

$$\begin{pmatrix} U \\ V \end{pmatrix} \rightarrow \begin{pmatrix} V \\ \frac{U}{V} \end{pmatrix} \Rightarrow \begin{pmatrix} X \\ Y \end{pmatrix} \rightarrow \begin{pmatrix} X \cdot Y \\ X \end{pmatrix} = \begin{pmatrix} U \\ V \end{pmatrix}$$

El Jacobiano será entonces:

$$\begin{bmatrix} y & x \\ 1 & 0 \end{bmatrix} \Rightarrow |det| = |x|$$

Por lo tanto utilizando nuevamente la regla del Jacobiano:

$$f_{XY}(x,y) = |x| \cdot \sqrt{\frac{\lambda}{\pi}} \cdot e^{-\lambda x^2 y^2} \cdot \sqrt{\frac{\lambda}{\pi}} \cdot e^{-\lambda x^2} = |x|\frac{\lambda}{\pi} e^{-\lambda x^2(1+y^2)}$$

Para obtener la marginal de Y integramos la conjunta con respecto a x.

$$\Rightarrow f_Y(y) = \frac{\lambda}{\pi} \int_{-\infty}^{\infty} |x| e^{-\lambda x^2(1+y^2)} dx = \frac{2\lambda}{\pi} \int_{0}^{\infty} x e^{-\lambda x^2(1+y^2)} dx = \frac{2\lambda}{\pi} \cdot \left. \frac{e^{-\lambda x^2(1+y^2)}}{-2\lambda(1+y^2)} \right|_0^{\infty}$$

$$\Rightarrow f_Y(y) = \frac{1}{\pi(1+y^2)}$$

Por lo tanto $\frac{U}{V}$ tiene distribución de Cauchy.

Problema IV.3.2

Sean X, Y variables aleatorias independientes exponenciales de parámetro λ.
Calcule e identifique la distribución de:

$$Z = \frac{X}{X+Y}$$

Solución:

La densidad conjunta de X, Y será:

$$f_{XY}(x,y) = f_X(x) \cdot f_Y(y) = \lambda^2 \cdot e^{-\lambda(x+y)} \mathbb{1}_{\mathbb{R}^+}(x) \cdot \mathbb{1}_{\mathbb{R}^+}(y)$$

Agregando la variable aleatoria $X+Y$ consideremos el cambio de variables que transforma (X,Y) en $(\frac{X}{X+Y}, X+Y)$.

Nos interesa el cambio de variables inverso, al anotar $z = \frac{x}{x+y}$ y $t = x + y$ se tendrá:

$$(z,t) \stackrel{\varphi}{\to} (x,y) = (zt, t(1-z))$$

El Jacobiano de este cambio de variables será: $|detD\varphi|(z,t) = |\begin{vmatrix} t & z \\ -t & (1-z) \end{vmatrix}| = t$
Por lo tanto al aplicar la regla del Jacobiano:

$$f_{ZT}(z,t) = t \cdot f_{XY}(x(z,t), y(z,t)) = \lambda^2 t \cdot e^{-\lambda(zt+t-zt)} 1\!\!1_{\mathbb{R}^+}(zt) \cdot 1\!\!1_{\mathbb{R}^+}(t(1-z))$$

$$\Rightarrow f_{ZT}(z,t) = \lambda^2 t \cdot e^{-\lambda t} \cdot 1\!\!1_{\mathbb{R}^+}(zt) \cdot 1\!\!1_{\mathbb{R}^+}(t(1-z))$$

Como estamos considerando $z \geq 0$ y $t \geq 0$
$\Rightarrow 1\!\!1_{\mathbb{R}^+}(zt) = 1\!\!1_{\mathbb{R}^+}(t) \quad y \quad 1\!\!1_{\mathbb{R}^+}(t(1-z)) = 1\!\!1_{[0,1]}(z)$
Entonces:

$$f_{ZT}(z,t) = e^{-\lambda t} \cdot t \cdot 1\!\!1_{\mathbb{R}^+}(t) \cdot \lambda^2 \cdot 1\!\!1_{[0,1]}(z)$$

Para obtener la marginal de Z basta integrar la conjunta con respecto a t.

$$f_Z(z) = \left(\int_0^\infty e^{-\lambda t} \cdot t dt \right) \cdot \lambda^2 \cdot 1\!\!1_{[0,1]}(z)$$

Integrando por partes:

$$\int_0^\infty e^{-\lambda t} \cdot t dt = e^{\frac{-\lambda^2 t}{-\lambda}} \Big|_0^\infty - \frac{1}{\lambda^2} e^{-\lambda t} \Big|_0^\infty = \frac{1}{\lambda^2}$$

Entonces $f_Z(z) = 1\!\!1_{[0,1]}(z)$ es decir una variable aleatoria uniforme en $[0,1]$.

Observación: Al obtener $f_{ZT}(z,t) = \left[1\!\!1_{[0,1]}(z)\right] \cdot \left[e^{-\lambda t} \cdot t \cdot 1\!\!1_{[0,1]}(t) \cdot \lambda^2\right]$, un resultado que se escribe como un producto $g_Z(z) \cdot g_T(t)$ se sabe que las variables son independientes y que ambas funciones son múltiplos de $f_Z(z)$ y $f_T(t)$ respectivamente, solo es necesario repartir las constantes para que las marginales integren 1.

IV.4 Normal Multivariada

Problema IV.4.

Un vector aleatorio $\vec{X} = (X_1, ..., X_n) \in \mathbb{R}^n$ se dice que sigue una ley normal multivariada de vector de medias $\mu \in \mathbb{R}^n$ y matriz de varianzas-covarianzas $\Sigma \in \mathbb{R}^{n\times n}$ (lo que se anota $\vec{X} \sim N_n(\mu, \Sigma)$...) si Σ es d.p. y $\vec{X}$ tiene función de densidad

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{det\Sigma}} e^{-\frac{1}{2}(\vec{x}-\mu)^T\Sigma^{-1}(\vec{x}-\mu)}$$

Demuestre que $\vec{X} \sim N_n(\mu, \Sigma)$ si y sólo si $\exists \vec{Z} \sim N_n(0, I)$ y $A \in \mathbb{R}^{n\times n}$ invertible tal que $\vec{X} = A\vec{Z} + \mu$ y $\Sigma = AA^T$.

Solución:

Demostremos primero que si $\vec{Z} \sim N_n(0, I)$ y $A \in \mathbb{R}^{n\times n}$ es invertible entonces el vector aleatorio $\vec{X} := A\vec{Z} + \mu$ tiene ley $N_n(\mu, \Sigma)$ con $\Sigma = AA^T$. En efecto como $\vec{Z} \sim N_n(0, I)$ entonces

$$f_{\vec{Z}}(\vec{z}) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\vec{z}^T\vec{z}}$$

Consideremos la transformación lineal afín $g : \mathbb{R}^n \to \mathbb{R}^n$ dada por $g(\vec{z}) := A\vec{z} + \mu$, es claro que como A es invertible entonces g es biyectiva de hecho $g^{-1}(\vec{x}) = A^{-1}\vec{x} - A^{-1}\mu$. Del método del Jacobiano concluimos que el vector aleatorio $\vec{X} := g(\vec{Z})$ posee función de densidad en $\mathbb{R}^n$ y que ésta viene dada por:

$$f_{\vec{X}}(\vec{x}) = f_{\vec{Z}}(g^{-1}(\vec{x}))|detJ_{g^{-1}}(\vec{x})| = |detA^{-1}|f_{\vec{Z}}(A^{-1}(\vec{x} - \mu))$$

$$= \frac{|detA^{-1}|}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\vec{x}-\mu)^T(A^T)^{-1}A^{-1}(\vec{x}-\mu)}$$

Al definir la matriz $\Sigma := AA^T$ (que es d.p.) se tiene que :

(1) $(A^T)^{-1}A^{-1} = (AA^T)^{-1} = \Sigma^{-1}$
(2) $|detA^{-1}| = |detA|^{-1}$
(3) $\sqrt{det\Sigma} = [(detA)\cdot(detA^T)]^{1/2} = [(detA)^2]^{1/2} = |detA|$

De (1) y (2) concluimos que:

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}|detA|} e^{-\frac{1}{2}(\vec{x}-\mu)^T\Sigma^{-1}(\vec{x}-\mu)}$$

y de (3) obtenemos que:

$$f_{\vec{X}}(\vec{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{det\Sigma}} e^{-\frac{1}{2}(\vec{x}-\mu)^T\Sigma^{-1}(\vec{x}-\mu)}$$

luego $\vec{X} \sim N_n(\mu, \Sigma)$.

Recíprocamente, supongamos que $\vec{X} \sim N_n(\mu, \Sigma)$. Como Σ es matriz d.p. del curso de álgebra lineal sabemos que $\exists A \in I\!R^{n\times n}$ invertible tal que $\Sigma = AA^T$. Consideremos el vector aleatorio $\vec{Z}$ definido por $\vec{Z} := A^{-1}\vec{X} - A^{-1}\mu$ el cual verifica trivialmente la identidad $\vec{X} = A\vec{Z} + \mu$. Para concluir basta demostrar que $\vec{Z} \sim N_n(0, I)$.En efecto, como $\vec{Z} = g^{-1}(\vec{X})$ del método del jacobiano, podemos concluir que $\vec{Z}$ es un vector aleatorio con densidad en $I\!R^n$, dada por

$$\begin{aligned} f_{\vec{Z}}(\vec{z}) &= f_{\vec{X}}(g(\vec{z}))|detJ_g(\vec{z})| = |detA| f_{\vec{X}}(A\vec{z} + \mu) \\ &= \frac{|detA|}{(2\pi)^{n/2}\sqrt{det\Sigma}} e^{-\frac{1}{2}(A\vec{z}+\mu-\mu)^T\Sigma^{-1}(A\vec{z}+\mu-\mu)} \\ &= \frac{|detA|}{(2\pi)^{n/2}\sqrt{det\Sigma}} e^{-\frac{1}{2}\vec{z}^T(A^T\Sigma^{-1}A)\vec{z}} \end{aligned}$$

De (1) es claro que $A^T\Sigma^{-1}A = I$, luego por (3) obtenemos que:

$$f_{\vec{Z}}(\vec{z}) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}\vec{z}^T\vec{z}} = \frac{1}{(2\pi)^{n/2}\sqrt{detI}} e^{-\frac{1}{2}(\vec{z}-0)^T I^{-1}(\vec{z}-0)}$$

i.e. $\vec{Z} \sim N_n(0, I)$.

Problema IV.4.

Demostrar que si $\vec{X} \sim N_n(\mu, \Sigma)$ y $B \in I\!R^{p\times n}$ es de rango completo entonces el vector aleatorio $\vec{Y} \in I\!R^p$ definido por $Y := B\vec{X}$ tiene una ley $N_p(B\mu, B\Sigma B^T)$.

Solución:

Como $\vec{X} \sim N_n(\mu, \Sigma)$ del problema anteriormente resuelto sabemos que $\exists A \in I\!R^{n\times n}$ invertible y $Z \sim N_n(0, I_n)$ tales que $\Sigma = AA^T$ y $\vec{X} = A\vec{Z} + \mu$. Se tiene entonces que:

$$\vec{Y} = (BA)\vec{Z} + B\mu \ \text{ con } \ BA \in I\!R^{p\times n}, B\mu \in I\!R^p$$

Demostremos primero que:

(1) $dim[KerBA] = n - p$
(2) $dim[(KerBA)^\perp] = p$

En efecto como una matriz es de rango completo si y sólo si la aplicación lineal asociada es epiyectiva, es directo del hecho que A sea invertible y B de rango completo que BA es de rango completo. Del teorema del núcleo imagen sale que:

$$dim[I\!R^n] = dim[KerBA] + dim[ImBA] \Leftrightarrow n = dim(KerBA) + p$$

y por lo tanto $dim[KerBA] = n - p$ lo que demuestra (1). Por otro lado como $\mathbb{R}^n = [KerBA] \oplus (KerBA)^{\perp}$ entonces

$$dim[\mathbb{R}^n] = dim[KerBA] + dim[(KerBA)^{\perp}] \Leftrightarrow n = (n-p) + dim[(KerBA)^{\perp}]$$

de donde $dim[(KerBA)^{\perp}] = p$, lo que demuestra (2).

De (1) sale que $\exists P \in \mathbb{R}^{n\times(n-p)}$ cuyas columnas son una base ortonormal del s.e.v. $KerBA$ y de (2) también sale que $\exists Q \in \mathbb{R}^{n\times p}$ cuyas columnas son una base ortonormal del s.e.v. $(KerBA)^{\perp}$. Es fácil verificar entonces que:

(3) $P^T P = I_{n-p}$, $Q^T Q = I_p$
(4) $P^T Q = 0 \in \mathbb{R}^{(n-p)\times p}$, $Q^T P = 0 \in \mathbb{R}^{p\times(n-p)}$.

Al definir la matriz por bloques $R := [P|Q] \in \mathbb{R}^{n\times n}$ es evidente que:

$$R^T R = \left[\frac{P^T}{Q^T}\right][P|Q] = \left[\begin{array}{c|c} P^T P & P^T Q \\ \hline Q^T P & Q^T Q \end{array}\right]$$

y por lo tanto de (3) y (4) obtenemos que $R^T R = I_n$, es decir, R es matriz ortonormal (en particular invertible). Definamos el vector aleatorio $\vec{V} := R^T \vec{Z}$, entonces como

$$\vec{V} = R^T \vec{Z} + 0$$

con $R^T \in \mathbb{R}^{n\times n}$ invertible y $\vec{Z} \sim N_n(0, I_n)$, del problema anteriormente resuelto concluimos que $\vec{V} \sim N_n(0, R^T R)$ i.e. $\vec{V} \sim N_n(0, I_n)$. De la identidad trivial $\vec{Z} = R^T \vec{V}$, obtenemos que:

$$\vec{Y} = (BAR)\vec{V} + B\mu$$

Pero $BAR = (BA)[P|Q] = [(BA)P|(BA)Q]$. Como las columnas de Q son elementos del $Ker(BA)$, es fácil deducir que $(BA)Q = 0 \in \mathbb{R}^{p\times p}$, y por lo tanto

$$\vec{Y} = [B\ A\ P|\ 0\ ]\vec{V} + B\mu \tag{5}$$

Si escribimos $\vec{V}$ como $\vec{V} = \left[\frac{\vec{V}_1}{\vec{V}_2}\right]$ donde $\vec{V}_1$ son las primeras p componentes del vector aleatorio $\vec{V}$ y $\vec{V}_2$ las últimas $(n-p)$ componentes del vector aleatorio $\vec{V}$, de (5) obtenemos que

$$\vec{Y} = (BAP)\vec{V}_1 + B\mu \tag{6}$$

Dado que $\vec{V} \sim N_n(0, I_n)$ es fácil deducir que $\vec{V}_1 \sim N_p(0, I_p)$. Además como B, A y P son matrices de rango completo, también lo será la matriz BAP, luego como $BAP \in \mathbb{R}^{p\times p}$, concluimos que BAP es matriz invertible. Del problema anterior se deduce que $\vec{Y} \sim N_p(B\mu, (BAP)(BAP)^T)$, pero

$$(BAP)(BAP)^T = BA(PP^T)A^T B = BAI_n A^T B^T = B(AA^T)B^T = B\Sigma P^T$$

i.e $\vec{Y} \sim N_p(B\mu, B\Sigma B^T)$.

Problema IV.4.

Sea $\vec{X} \sim N_n(0, \sigma^2 I)$ *donde* $\sigma \in I\!R_+$. *Demostrar que si* $A \in I\!R^{n\times n}$ *y* $a \in I\!R^n \setminus \{0\}$ *son tales que* $Aa = 0$ *entonces las variables aleatorias* $\vec{X}^T A\vec{X}$ *y* $a^T\vec{X}$ *son independientes.*

Solución:

En efecto, como $A \in I\!R^{n\times n}$ es simétrica entonces es diagonizable y como $Aa = 0$, entonces a es vector propio de A asociado al valor propio $\lambda = 0$. Se tiene entonces que existe una base ortonormal en $I\!R^n$ de vectores propios de A de la forma:

$$\{v_1, ..., v_{n-1}, \frac{a}{\|a\|}\}$$

donde $(\forall i = 1, ..., n-1) : v_i$ es vector propio de A asociado a algún valor propio λ_i de A. Definamos la matriz

$$P := \left[v_1, ..., v_{n-1}, \frac{a}{\|a\|}\right] \in I\!R^{n\times n}$$

Es fácil convencer que:

(1) $P^T P = PP^T = I_n$

(2) $D := P^T AP = \begin{bmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_{n-1} & \\ & & & 0 \end{bmatrix}$

De (1) sale que P^T es invertible y luego de rango completo, por lo tanto del problema anteriormente resuelto concluimos que el vector aleatorio $\vec{Y} := P^T\vec{X} \sim N_n(P^T 0, P^T(\sigma^2 I)P)$ i.e $\vec{Y} \sim N_n(0, \sigma^2 I)$.
De (2) sale que $A = PDP^T$, luego

$$\vec{X}^T A\vec{X} = \vec{X}^T(PDP^T)\vec{X} = (P^T\vec{X})^T D(P^T\vec{X}) = \vec{Y}^T D\vec{Y} = \sum_{i=1}^{n-1} \lambda_i Y_i^2 \quad (3)$$

$$a^T\vec{X} = a^T(P\vec{Y}) = (a^T P)\vec{Y} = Y_n \quad (4)$$

Si $\vec{Y} \sim N_n(0, \sigma^2 I)$ entonces

$$f_{\vec{Y}}(y_1, ..., y_n) = \frac{1}{(2\pi)^{n/2}\sigma^n} e^{-\frac{1}{2}\sum_{i=1}^{n} y_i^2} = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y_i^2}{2}}$$

Como $(\forall i = 1, ..., n) : \int\limits_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y_i^2}{2}} dy_i = 1$, concluimos que las variables aleatorias $Y_1, ..., Y_n$ son independientes.
Finalmente de (3) y (4), obtenemos que:

$$\begin{aligned} \vec{X}^T A \vec{X} &= \text{ función } (Y_1, ..., Y_{n-1}) \\ a^T \vec{X} &= \text{ función } (Y_n) \end{aligned}$$

Como $Y_1, ..., Y_n$ son v.a. independientes entonces de lo anterior,tambien lo serán las variables aleatorias $\vec{X}^T A \vec{X}$ y $a^T \vec{X}$.

Problema IV.4.

Demostrar que si $X_1, ..., X_n$ son variables aleatorias i.i.d. $N(\mu, \sigma^2)$ entonces las variables aleatorias

$$\bar{X} := \frac{1}{n} \sum_{i=1}^{n} X_i \quad y \quad S_n^2 := \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

son independientes.

Solución:

Definamos para cada $i \in \{1, ..., n\}$ la variable aleatoria $Y_i := X_i - \mu \sim N(0, \sigma^2)$. Es fácil verificar que:

$$\bar{X} = (\frac{1}{n} \sum_{i=1}^{n} Y_i) - \mu$$

$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^{n} (Y_i - \bar{Y})^2$$

luego para demostrar que $\bar{X}$ y S_n^2 son independientes, basta probar que $\frac{1}{n} \sum\limits_{i=1}^{n} Y_i$ y $\frac{1}{n-1} \sum\limits_{i=1}^{n} (Y_i - \bar{Y})$ son independientes. En efecto como $X_1, ..., X_n$ son independientes entonces también lo serán $Y_1, ..., Y_n$ luego al definir el vector aleatorio $\vec{Y} := (Y_1, ..., Y_n)$ se tiene que $\vec{Y}$ poseee función de densidad dada por:

$$f_{\vec{Y}}(\vec{y}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{y_i^2}{2}} = \frac{1}{(2\pi)^{n/2} \sigma^n} e^{-\frac{1}{2} \vec{y}^T \vec{y}}$$

i.e. $\quad \vec{Y} \sim N_n(0, \sigma^2 I)$.

Denotando por $\mathbb{1} := \sum_{i=1}^{n} e_i$ donde $e_1, ..., e_n$ son los vectores de la base canónica de $I\!R^n$, se tiene que:

$$\frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{1}{n}\mathbb{1}^T\vec{Y} = (\frac{\mathbb{1}}{n})^T\vec{Y} \tag{1}$$

y

$$\begin{aligned} \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 &= \frac{1}{n-1}\{\sum_{i=1}^{n} Y_i^2 - 2n\bar{Y}\cdot\frac{1}{n}\sum_{i=1}^{n} Y_i - n(\frac{1}{n}\sum_{i=1}^{n} Y_i)^2\} \\ &= \frac{1}{2}\{\sum_{i=1}^{n} Y_i^2 - n(\frac{1}{n}\sum_{i=1}^{n} Y_i)^2\}. \end{aligned}$$

Además como $\sum_{i=1}^{n} Y_i^2 = \vec{Y}^T\vec{Y}$ y $(\frac{1}{n}\sum_{i=1}^{n} Y_i)^2 = (\frac{1}{n}\vec{Y}^T\mathbb{1})(\frac{1}{n}\mathbb{1}^T\vec{Y}) = \frac{1}{n^2}\vec{y}^T\mathbb{1}\mathbb{1}^T\vec{Y}$ obtenemos que

$$\begin{aligned} \frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2 &= \frac{1}{(n-1)}\{\vec{Y}^T\vec{Y} - \frac{1}{n}\vec{Y}^T\mathbb{1}\mathbb{1}^T\vec{Y}\} \\ &= \vec{Y}^T(\frac{I - n^{-1}\mathbb{1}\mathbb{1}^T}{n-1})\vec{Y} \end{aligned} \tag{2}$$

De las igualdades (1) y (2) como $\vec{Y} \sim N_n(0, \sigma^2, I_n)$ para concluir (según fue demostrado en el problema anterior) que $\frac{1}{n}\sum_{i=1}^{n} Y_i$ e $\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ son independientes, bastaría que:

$$(\frac{I - n^{-1}\mathbb{1}\mathbb{1}^T}{n-1})\frac{\mathbb{1}}{n} = 0$$

En efecto:

$$\begin{aligned} (\frac{I - n^{-1}\mathbb{1}\mathbb{1}^T}{n-1})\frac{\mathbb{1}}{n} &= \frac{1}{n(n-1)}(\mathbb{1} - n^{-1}\mathbb{1}\mathbb{1}^T\mathbb{1}) \\ &= \frac{1}{n(n-1)}(\mathbb{1} - n^{-1}\mathbb{1}n) = \frac{1}{n(n-1)}(\mathbb{1} - \mathbb{1}) = 0 \end{aligned}$$

luego $\frac{1}{n}\sum_{i=1}^{n} Y_i$ e $\frac{1}{n-1}\sum_{i=1}^{n}(Y_i - \bar{Y})^2$ son independientes, y por lo tanto también lo son $\bar{X}$ y S_n^2.