Capítulo 1

Independencia

1.1. Preliminares

En esta sección recordamos algunos resultados que serán utilizados en las demostraciones que siguen a continuación.

Definición 1.1 Sea X una variable aleatoria definida sobre el espacio $(\Omega, \mathcal{F}, P)$. La σ-álgebra generada por X, $\sigma(X)$ es la menor σ-álgebra contenida en $\mathcal{F}$ que hace medible a X.

Como X es una v.a., X es medible respecto a $\mathcal{F}$ y en consecuencia la colección de σálgebras que hacen medible a X no es vacía. Si tomamos la intersección de todas ellas el resultado es $\sigma(X)$: Si X es medible respecto a la σ-álgebra $\mathcal{H}$ entonces $\sigma(X) \subset \mathcal{H}$.

Ejemplo 1.1
Sea $\Omega = \{1, 2, 3, 4, 5, 6\}$ y $X : \Omega \to \mathbb{R}$ definida por

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \text{es par,} \\ 0 & \text{si } \omega \text{es impar,.} \end{cases}$$

En este caso es sencillo verificar que

$$\sigma(X) = \{\Omega, \emptyset, \{1, 3, 5\}, \{2, 4, 6\}\}$$

Definición 1.2 Una colección $\mathcal{C}$ de subconjuntos de Ω es un sistema π si es cerrada bajo intersecciones:

$$A, B \in \mathcal{C} \Rightarrow A \cap B \in \mathcal{C}.$$

Una colección $\mathcal{L}$ de subconjuntos de Ω es un sistema λ o una clase de Dynkin si satisface las siguientes condiciones:

(a) $\Omega \in \mathcal{L}$.

(b) $A, B \in \mathcal{L}, A \subset B \Rightarrow B - A \in \mathcal{L}$.

(c) $\mathcal{L}$ es cerrada bajo límites crecientes: $A_n \in \mathcal{L}, n \geq 1, A_n \subset A_{n+1}$ para todo $n \geq 1 \Rightarrow \cup_{n\geq 1} A_n \in \mathcal{L}$.

Es inmediato que toda σ-álgebra es un sistema λ.

Ejercicio Demostrar que (b) y (c) son equivalentes a

(b') $A \in \mathcal{L} \Rightarrow A^c \in \mathcal{L}$.

(c') $\mathcal{L}$ es cerrada abajo uniones disjuntas: Si $A_n \in \mathcal{L}, n \geq 1$ y $A_i \cap A_j = \emptyset$ siempre que $i \neq j$ entonces $\cup_{n\geq 1} A_n \in \mathcal{L}$.

Proposición 1.1 *Sea $\mathcal{F} \subset \mathcal{P}(\Omega)$. $\mathcal{F}$ es una σ-álgebra si y solo si $\mathcal{F}$ es un álgebra y satisface*

$$A_n \in \mathcal{F}, A_n \subset A_{n+1}, n \geq 1 \Rightarrow \cup_{n\geq 1} A_n \in \mathcal{F}.$$

Teorema 1.1 (Dynkin) *(a) Si $\mathcal{C}$ es un sistema π y $\mathcal{L}$ es un sistema λ tal que $\mathcal{C} \subset \mathcal{L}$ entonces $\sigma(\mathcal{C}) \subset \mathcal{L}$.*
(b) Si $\mathcal{C}$ es un sistema π

$$\sigma(\mathcal{C}) = \mathcal{L}(\mathcal{C}),$$

es decir, la σ-álgebra generada por $\mathcal{C}$ coincide con el sistema λ generado por $\mathcal{C}$.

Demostración. (a) Sea $\mathcal{L}(\mathcal{C})$ el sistema λ generado por $\mathcal{C}$, es decir, la intersección de todos los sistemas λ que contienen a $\mathcal{C}$. Si $\mathcal{L}(\mathcal{C})$ fuese un sistema π entonces sería una σ-álgebra, y en consecuencia $\sigma(\mathcal{C}) \subset \mathcal{L}(\mathcal{C}) \subset \mathcal{L}$. Por lo tanto basta mostrar que $\mathcal{L}(\mathcal{C})$ es un sistema π.

Para cada conjunto A definimos $\mathcal{G}_A$ como la clase de los conjuntos B tales que $A \cap B \in \mathcal{L}(\mathcal{C})$. Si $A \in \mathcal{C}$ entonces $\mathcal{G}_A$ es un sistema λ:

(i) $\Omega \in \mathcal{G}_A$ porque $\Omega \cap A = A \in \mathcal{C} \subset \mathcal{L}(\mathcal{C})$.

(ii) Supongamos que $B \subset C,\ B, C \in \mathcal{G}_A$, es decir, $B \cap A \in \mathcal{L}(\mathcal{C}), C \cap A \in \mathcal{L}(\mathcal{C})$. Entonces $(C - B) \cap A = (C \cap A) - (B \cap A)$ y como esto es una diferencia propia, está en $\mathcal{L}(\mathcal{C})$. Por lo tanto $C - B \in \mathcal{G}_A$.

(iii) Sea $B_n \uparrow B,\ B_n \in \mathcal{G}_A$, entonces $B_n \cap A \in \mathcal{L}(\mathcal{C}),\ B_n \cap A \uparrow B \cap A$ y como $\mathcal{L}(\mathcal{C})$ es cerrada bajo límites crecientes, $B \cap A \in \mathcal{L}(\mathcal{C})$ y $B \in \mathcal{G}_A$.

Veamos ahora que $\mathcal{L}(\mathcal{C})$ es un sistema π. Si $A \in \mathcal{C},\ B \in \mathcal{C}$ entonces $A \cap B \in \mathcal{C} \subset \mathcal{L}(\mathcal{C})$. Por lo tanto si $A \in \mathcal{C}$, tenemos que $\mathcal{C} \subset \mathcal{G}_A$ y como $\mathcal{G}_A$ es un sistema λ, $\mathcal{L}(\mathcal{C}) \subset \mathcal{G}_A$. Por lo tanto $A \in \mathcal{C} \Rightarrow \mathcal{L}(\mathcal{C}) \subset \mathcal{G}_A$, es decir,

$$A \in \mathcal{C},\ B \in \mathcal{L}(\mathcal{C}) \Rightarrow A \cap B \in \mathcal{L}(C).$$

Pero esto quiere decir que si $B \in \mathcal{L}(\mathcal{C})$ entonces $\mathcal{C} \subset \mathcal{G}_B$, y como $\mathcal{G}_B$ es un sistema λ, tenemos que $\mathcal{L}(\mathcal{C}) \subset \mathcal{G}_B$. En consecuencia

$$B \in \mathcal{L}(\mathcal{C}),\ C \in \mathcal{L}(\mathcal{C}) \Rightarrow B \cap C \in \mathcal{L}(C)$$

y $\mathcal{L}(\mathcal{C})$ es un sistema π.

(b) Es consecuencia de (a). Como $\mathcal{C} \subset \mathcal{L}(\mathcal{C})$ tenemos que $\sigma(\mathcal{C}) \subset \mathcal{L}(\mathcal{C})$. Por otro lado $\sigma(\mathcal{C})$ es una σ-álgebra y por lo tanto un sistema λ que contiene a $\mathcal{C}$. En consecuencia $\mathcal{L}(\mathcal{C}) \subset \sigma(\mathcal{C})$. ■

1.2. Eventos Independientes

La noción de independencia surge en el ámbito de la Teoría de Probabilidades, de modo que en este capítulo todas las medidas son probabilidades.

Definición 1.3 Sea $(\Omega, \mathcal{F}, P)$ un espacio de probabilidad. Los eventos $A, B \in \mathcal{F}$ son independientes si

$$P(A \cap B) = P(A)P(B).$$

Definición 1.4 Los eventos $A_1, \dots, A_n$ son independientes si

$$P\big(\cap_{i\in I} A_i\big) = \prod_{i\in I} P(A_i) \quad \forall I \subset \{1, \dots, n\}. \tag{1.1}$$

Esta ecuación equivale a

$$P\big(B_1 \cap B_2 \cap \cdots \cap B_n\big) = \prod_{i=1}^{n} P(B_i) \tag{1.2}$$

donde para cada $i = 1, \ldots, n$, B_i es A_i o Ω.

Si para todo par de índices i, j se tiene que

$$P(A_i \cap A_j) = P(A_i)P(A_j)$$

decimos que los eventos $A_1, \ldots, A_n$ son independientes a pares.

Independencia a pares no implica independencia, como lo muestra el siguiente ejemplo.

Ejemplo 1.2
Sea P la distribución uniforme en $\Omega = \{1, 2, 3, 4\}$. Es fácil verificar que los eventos $A_i = \{i, 4\}$ para $i = 1, 2, 3$ son independientes a pares pero no son independientes.

Definición 1.5 Sea $\mathcal{C}_i \subset \mathcal{F}$, $i = 1, \ldots, n$. Las clases $\mathcal{C}_i$ son independientes si para cada selección $A_1, \ldots, A_n$ con $A_i \in \mathcal{C}_i$, $i = 1, \ldots, n$, tenemos que los eventos A_i, $i = 1, \ldots, n$ son independientes.

Teorema 1.2 *Si para cada $i = 1, \ldots, n$, $\mathcal{C}_i$ es una clase no vacía de eventos que satisface*

1. *$\mathcal{C}_i$ es un sistema π para $i = 1, \ldots, n$,*
2. *$\mathcal{C}_i$ son independientes para $i = 1, \ldots, n$,*

entonces $\sigma(\mathcal{C}_1), \ldots, \sigma(\mathcal{C}_n)$ son independientes.

Demostración. Haremos la demostración para el caso $n = 2$. El resultado general se puede obtener por inducción. Fijamos $A_2 \in \mathcal{C}_2$ y ponemos

$$\mathcal{L} = \{A \in \mathcal{F} : P(A \cap A_2) = P(A)P(A_2)\}.$$

Veamos que $\mathcal{L}$ es un sistema λ:
(a) $\Omega \in \mathcal{L}$ porque $P(\Omega \cap A_2) = P(A_2) = P(\Omega)P(A_2)$.
(b) $A, B \in \mathcal{L}$, $A \subset B$ entonces $B - A \in \mathcal{L}$:

$$\begin{aligned} P((B - A) \cap A_2) &= P((B \cap A_2) - (A \cap A_2)) = P(B \cap A_2) - P(A \cap A_2) \\ &= P(B)P(A_2) - P(A)P(A_2) = (P(B) - P(A))P(A_2) \\ &= P(B - A)P(A_2). \end{aligned}$$

(c) Si $B_n \in \mathcal{L}$, $B_n \uparrow B$ entonces $B \in \mathcal{L}$.

$$\begin{aligned} P(B \cap A_2) &= P(\lim_n B_n \cap A_2) = P(\lim_n (B_n \cap A_2)) \\ &= \lim_n P(B_n \cap A_2) = \lim_n P(B_n)P(A_2) \\ &= P(B)P(A_2) \end{aligned}$$

Además $\mathcal{C}_1 \subset \mathcal{L}$ y por el Teorema de Dynkin $\sigma(\mathcal{C}_1) \subset \mathcal{L}$. En consecuencia $\sigma(\mathcal{C}_1)$ y $\mathcal{C}_2$ son independientes. De manera similar se muestra ahora que $\sigma(\mathcal{C}_1)$ y $\sigma(\mathcal{C}_2)$ son independientes. ■

Definición 1.6 Sea T un conjunto arbitrario de índices. Las clases $\mathcal{C}_t,\ t \in T$ son independientes si para cada $I \subset T$ finito, las clases $\mathcal{C}_t,\ t \in I$ son independientes.

Corolario 1.1 *Si $\{\mathcal{C}_t,\ t \in T\}$ son sistemas π no vacíos e independientes entonces $\{\sigma(\mathcal{C}_t),\ t \in T\}$ son independientes.*

Notación: Usaremos la notación $\perp$ para indicar independencia.

1.3. Variables Aleatorias Independientes

Definición 1.7 $\{X_t,\ t\in T\}$ es una familia de variables aleatorias independientes (v.a.i.) si $\{\sigma(X_t),\ t\in T\}$ son σ-álgebras independientes.

Ejemplo 1.3
Como $\sigma(\mathbf{1}_A)=\{\emptyset,\Omega,A,A^c\}$ tenemos que $\mathbf{1}_{A_1},\dots,\mathbf{1}_{A_n}$ son independientes si y sólo si $A_1,\dots,A_n$ son independientes.

Sea $\{X_t,\ t\in T\}$ una colección de v.a. y sea J un subconjunto finito de T. Usaremos la notación F_J para indicar la función de distribución conjunta de las variables X_j para $j\in J$. Estas funciones F_J se conocen como las funciones de distribución finito-dimensionales de la colección $\{X_t,\ t\in T\}$. Tenemos

$$F_J(x_t,\ t\in J)=P(X_t\le x_t,\ t\in J) \tag{1.3}$$

Teorema 1.3 (Criterio de Factorización) *Una familia de v.a. $\{X_t,\ t\in T\}$ es independiente si y sólo si para todo $J\subset T$ finito se tiene que*

$$F_J(x_t,\ t\in J)=\prod_{t\in J}P(X_t\le x_t) \tag{1.4}$$

para todo $x_t\in\mathbb{R},\ t\in J$.

Demostración. Por la definición 1.6 basta ver que para todo subconjunto finito de índices J, $\{X_t,\ t\in J\}$ son independientes si y sólo si (1.4) vale. Definimos $\mathcal{C}_t=\{\{X_t\le x\},\ x\in\mathbb{R}\}$. Entonces

(i) $\mathcal{C}_t$ es un sistema π porque
$$\{X_t\le x\}\cap\{X_t\le y\}=\{X_t\le x\wedge y\}$$

(ii) $\sigma(\mathcal{C}_t)=\sigma(X_t)$.

Pero (1.4) dice que $\{\mathcal{C}_t,\ t\in J\}$ es una familia independiente y por el teorema 1.2 tenemos que $\{\sigma(\mathcal{C}_t)=\sigma(X_t):t\in J\}$ son independientes. ■

Corolario 1.2 *La colección finita $X_1,\dots,X_k$ de v.a. es independiente si y sólo si*

$$F_{1,\dots,k}(x_1,\dots,x_k)=P(X_1\le x_1,\dots,X_k\le x_k)=\prod_{i=1}^{k}P(X_i\le x_i)=\prod_{i=1}^{k}F_i(x_i) \tag{1.5}$$

para todo $x_i\in\mathbb{R},\ i=1,\dots,k$.

Corolario 1.3 *Las v.a. $X_1,\dots,X_k$ discretas con rango numerable $\mathcal{R}$ son independientes si y sólo si*

$$P(X_1=x_1,\dots,X_k=x_k)=\prod_{i=1}^{k}P(X_i=x_i) \tag{1.6}$$

para todo $x_i\in\mathcal{R},\ i=1,\dots,k$.

Demostración. Si $X_1,\dots,X_k$ son independientes entonces $\sigma(X_i)$, $i=1,\dots,n$ son independientes y como $\{X_i=x_i\}\in\sigma(X_i)$ tenemos que $\{X_i=x_i\},\ i=1,\dots,k$ son eventos independientes, de modo que (1.6) es válida.

Recíprocamente, supongamos que (1.6) vale. Sea $\mathbf{z}, \mathbf{x}$ vectores en $\mathcal{R}^k$. Usaremos la notación $\mathbf{z} \leq \mathbf{x}$ para indicar que $z_i \leq x_i$ para todo $i = 1, \ldots, k$, y $\mathbf{x}_{-i} = (x_1, \ldots, x_{i-1}, x_{i+1}, \ldots x_k)$, el vector que se obtiene eliminando la i-ésima componente de $\mathbf{x}$. Entonces

$$\begin{aligned}
P(X_i \leq x_i,\ i = 1, \ldots, k) &= \sum_{\mathbf{z} \leq \mathbf{x}} P(X_i = z_i,\ i = 1, \ldots k) \\
&= \sum_{\mathbf{z} \leq \mathbf{x}} \prod_{i=1}^{k} P(X_i = z_i) \\
&= \sum_{\mathbf{z}_{-1} \leq \mathbf{x}_{-1}} \sum_{z_1 \leq x_1,\ z_1 \in \mathcal{R}} P(X_1 = z_1) \prod_{i=2}^{k} P(X_i = z_i) \\
&= P(X_1 \leq x_1) \sum_{\mathbf{z}_{-1} \leq \mathbf{x}_{-1}} \prod_{i=2}^{k} P(X_i = z_i) \\
&= \prod_{i=1}^{k} P(X_i \leq x_i).
\end{aligned}$$

■

Corolario 1.4 *Si las variables $X_1, \ldots, X_k$ tienen distribución conjunta que es absolutamente continua con respecto a la medida de Lebesgue en $\mathbb{R}^k$, entonces ellas son independientes si y solo si*

$$f_{1,\ldots,k}(x_1, \ldots, x_k) = \prod_{i=1}^{k} f_i(x_i) \tag{1.7}$$

c.s. respecto a la medida de Lebesgue en $\mathbb{R}^k$, donde $f_{1,\ldots,k}$ es la densidad de la distribución conjunta del vector $(X_1, \ldots, X_k)$ mientras que f_i es la densidad de la variable X_i.

Demostración. Ejercicio.

Proposición 1.2 *Sea $\{X_1, \ldots, X_k\}$ una colección de variables aleatorias definidas en el espacio $(\Omega, \mathcal{F}, P)$. Las variables son independientes si y solo si*

$$\mathrm{E} \prod_{i=1}^{k} h_i(X_i) = \prod_{i=1}^{k} \mathrm{E}\, h_i(X_i) \tag{1.8}$$

para todas las funciones $h_i : \mathbb{R} \to \mathbb{R}$, $i = 1, \ldots, k$ medibles y acotadas.

Demostración. *Si (1.8) es válida, tomamos $h_i = \mathbf{1}_{B_i}$ con $B_i \in \mathcal{B}$, $i = 1, 2, \ldots, k$ y se obtiene la independencia de las variables.*

Recíprocamente, si $\{X_1, \ldots, X_k\}$ son independientes, entonces (1.8) es válida si tomamos $h_i = \mathbf{1}_{B_i}$ con $B_i \in \mathcal{B}$, $i = 1, 2, \ldots, k$. Por linealidad también es válida para funciones simples y siguiendo el procedimiento usual para esta demostraciones la ecuación (1.8) es válida para cualesquiera funciones medibles $h_i, i = 1, \ldots, k$. ■

Ejemplo 1.4 (Records, Rangos y el Teorema de Rényi)
Sea $(X_n,\ n \geq 1)$ una sucesión de variables aleatorias independientes identicamente distribuidas (v.a.i.i.d.) con f.d. común continua $F(x)$. Por continuidad

$$P(X_i = x_i) = 0 \tag{1.9}$$

de modo que si definimos $Empates = \cup_{i \neq j}\{X_i = X_j\}$, entonces $P(Empates) = 0$

Decimos que X_n es un *record* de la sucesión si

$$X_n > \max_{1 \le i < n} X_i$$

y definimos $A_n = \{X_n \text{ es un record}\}$.

El rango relativo de X_n en $X_1, \dots, X_n$ se define por

$$R_n = \sum_{j=1}^{n} \mathbf{1}_{\{X_j \ge X_n\}}$$

de modo que

$$\begin{aligned} R_n &= 1 \text{ sii } X_n \text{ es un record,} \\ &= 2 \text{ sii } X_n \text{ es el segundo valor en } X_1, \dots, X_n, \end{aligned}$$

y así sucesivamente.

Teorema 1.4 (Renyi) *Supongamos que $(X_n,\ n \ge 1)$ son i.i.d. con f.d. común $F(x)$ continua.*

(a) La sucesión de v.a. $(R_n,\ n \ge 1)$ es independiente y $P(R_n = k) = 1/n$, $k = 1, \dots, n$.

(b) La sucesión de eventos $(A_n,\ n \ge 1)$ es independiente y $P(A_n) = 1/n$.

Demostración. (b) es consecuencia de (a) porque $A_n = \{R_n = 1\}$.

(a) Hay $n!$ ordenamientos de $X_1, \dots, X_n$. Por simetría, como $X_1, \dots, X_n$ son i.i.d., los ordenamientos posibles son equiprobables con probabilidad igual a $1/n!$. Cada realización de $R_1, \dots, R_n$ determina, de manera única, un orden: Por ejemplo, si $n = 3$,

$$\begin{aligned} R_1(\omega) = 1,\ R_2(\omega) = 1,\ R_3(\omega) = 1 \quad &\Rightarrow \quad X_1(\omega) < X_2(\omega) < X_3(\omega) \\ R_1(\omega) = 1,\ R_2(\omega) = 2,\ R_3(\omega) = 3 \quad &\Rightarrow \quad X_3(\omega) < X_2(\omega) < X_1(\omega) \end{aligned}$$

En consecuencia, cada realización de $R_1, \dots, R_n$ tiene igual probabilidad que un ordenamiento particular de $X_1, \dots, X_n$. Por lo tanto

$$P(R_1 = r_1, \dots, R_n = r_n) = 1/n!$$

para $r_i \in \{1, \dots, i\}$, $i = 1, \dots, n$.

Observemos que

$$P(R_n = r_n) = \sum_{r_1, \dots, r_{n-1}} P(R_1 = r_1, \dots, R_{n-1} = r_{n-1}, R_n = r_n) = \sum_{r_1, \dots, r_{n-1}} 1/n!$$

Como r_i tiene i valores posibles, el número de términos en la suma es $1 \cdot 2 \cdot 3 \cdots (n-1) = (n-1)!$ y

$$P(R_n = r_n) = \frac{(n-1)!}{n!} = \frac{1}{n}, \quad n = 1, 2, \dots$$

Así,

$$P(R_1 = r_1, \dots, R_n = r_n) = \frac{1}{n!} = P(R_1 = r_1) \cdots P(R_n = r_n).$$

■

Ejemplo 1.5 (Desarrollo diádico)
Sea $(\Omega, \mathcal{F}, P) = ((0,1], \mathcal{B}, \lambda)$, el espacio de Lebesgue. Escribimos $\omega \in (0,1]$ usando su desarrollo binario

$$\omega = \sum_{n=1}^{\infty} \frac{d_n(\omega)}{2^n} = 0.d_1(\omega)d_2(\omega)d_3(\omega)\dots$$

donde cada $d_n(\omega)$ es 0 ó 1. Escribimos 1 como

$$0.111\cdots = \sum_{n=1}^{\infty} 2^{-n} = \frac{1/2}{1-1/2} = 1,$$

y para números como 1/2 que tienen dos desarrollos posibles:

$$\frac{1}{2} = \sum_{n=2}^{\infty} \frac{1}{2^n} = 0.0111\cdots \qquad\qquad \frac{1}{2} = \frac{1}{2} + 0 + 0\cdots = 0.1000\cdots$$

convenimos en usar el desarrollo que no termina, es decir, usamos el primer desarrollo.

Cada d_n es una variable aleatoria: Como d_n es discreta con valores $0,\ 1$, basta verificar que

$$\{d_n = 0\} \in \mathcal{B}, \qquad\qquad \{d_n = 1\} \in \mathcal{B}.$$

Comenzamos por considerar el caso $n = 1$,

$$\{d_1 = 0\} = (0.0000\cdots, 0.0111\cdots] = (0, 1/2] \in \mathcal{B},$$
$$\{d_1 = 1\} = (0.1000\cdots, 0.111\cdots] = (1/2, 1] \in \mathcal{B}.$$

El extremo izquierdo es abierto por la convención que hemos adoptado. Observamos que $P(d_1 = 1) = P(d_1 = 0) = 1/2$.

Ahora procedemos al caso general $n \geq 2$,

$$\{d_n = 1\} = \bigcup_{(u_1,\dots,u_{n-1})\in\{0,1\}^{n-1}} (0.u_1u_2\cdots u_{n-1}1000\cdots, 0.u_1u_2\cdots u_{n-1}111\cdots] \tag{1.10}$$

que es una unión disjunta de 2^{n-1} intervalos en $\mathcal{B}$. Por ejemplo $\{d_2 = 1\} = (\frac{1}{4}, \frac{2}{4}] \cup (\frac{3}{4}, 1]$.

Podemos usar (1.10) para calcular la distribución de d_n. Tenemos

$$\begin{aligned}
P(d_n = 1) &= \sum_{(u_1,\dots,u_{n-1})\in\{0,1\}^{n-1}} \lambda(0.u_1u_2\cdots u_{n-1}1000\cdots, 0.u_1u_2\cdots u_{n-1}111\cdots] \\
&= \sum_{(u_1,\dots,u_{n-1})\in\{0,1\}^{n-1}} [\sum_{i=1}^{n-1} \frac{u_i}{2^i} + \sum_{i=n}^{\infty} 2^{-i} - \sum_{i=1}^{n-1} \frac{u_i}{2^i} + 2^{-n}] \\
&= 2^{n-1} \sum_{i=n+1}^{\infty} 2^{-i} = \frac{1}{2}.
\end{aligned}$$

Concluimos que

$$P(d_n = 0) = P(d_n = 1) = \frac{1}{2}. \tag{1.11}$$

La sucesión $(d_n,\ n \geq 1)$ es i.i.d.: Ya vimos en (1.11) que son idénticamente distribuidas y falta ver que son independientes. Para esto basta tomar $n \geq 1$ y probar que $\{d_1, \dots, d_n\}$ son independientes. Para $(u_1, \dots, u_n) \in \{0,1\}^n$ tenemos

$$\cap_{i=1}^{n} \{d_i = u_i\} = (0.u_1u_2\cdots u_n000\cdots, 0.u_1u_2\cdots u_n111\cdots].$$

De nuevo el extremo izquierdo es abierto por la convención que adoptamos. Como la probabilidad de un intervalo es su longitud

$$\begin{aligned} P(\cap_{i=1}^n \{d_i = u_i\}) &= \sum_{i=1}^n \frac{u_i}{2^i} + \sum_{i=n+1}^\infty \frac{1}{2^i} - \sum_{i=1}^n \frac{u_i}{2^i} \\ &= \frac{2^{-(n+1)}}{1-\frac{1}{2}} = \frac{1}{2^n} = \prod_{i=1}^n P(d_i = u_i) \end{aligned}$$

y las variables son independientes.

1.3.1. Agrupamientos

Lema 1.1 *Sea $\{\mathcal{F}_t, t \in T\}$ una familia de σ-álgebras independientes. Sea S un conjunto de índices y supongamos que para cada $s \in S$ tenemos un conjunto $T_s \subset T$ y $\{T_s, s \in S\}$ son disjuntos dos a dos. Definimos*

$$\mathcal{F}_{T_s} = \sigma\big(\cup_{t\in T_s} \mathcal{F}_t\big).$$

Entonces $\{\mathcal{F}_{T_s}, s \in S\}$ es una familia de σ-álgebras independientes.

Demostración. Podemos suponer, sin pérdida de generalidad, que S es finito. Definimos

$$\mathcal{C}_{T_s} = \{\cap_{\alpha\in K} B_\alpha : B_\alpha \in \mathcal{F}_\alpha,\ K \subset T_s,\ K \text{ finito } \}.$$

Entonces $\mathcal{C}_{T_s}$ es un sistema π para cada s y $\{\mathcal{C}_{T_s},\ s \in S\}$ son clases independientes. Por lo tanto sólo hay que ver que $\sigma(\mathcal{C}_{T_s}) = \mathcal{F}_{T_s}$. Está claro que $\mathcal{C}_{T_s} \subset \mathcal{F}_{T_s}$ y por lo tanto $\sigma(\mathcal{C}_{T_s}) \subset \mathcal{F}_{T_s}$. Por otro lado, $\mathcal{F}_\alpha \subset \mathcal{C}_{T_s}$, para todo $\alpha \in T_s$, y por lo tanto $\mathcal{F}_\alpha \subset \sigma(\mathcal{C}_{T_s})$, para todo $\alpha \in T_s$. En consecuencia

$$\cup_{\alpha\in T_s} \mathcal{F}_\alpha \subset \sigma(\mathcal{C}_{T_s})$$

y por lo tanto

$$\mathcal{F}_{T_s} = \sigma\big(\cup_{\alpha\in T_s} \mathcal{F}_\alpha\big) \subset \sigma(\mathcal{C}_{T_s}).$$

■

Ejemplos 1.6
Sea $\{X_n,\ n \geq 1\}$ v.a.i. Entonces

$$\begin{aligned} \sigma(X_j,\ j \leq n) \quad &\perp \quad \sigma(X_j,\ j > n), \\ \sum_{i=1}^n X_i \quad &\perp \quad \sum_{i=n+1}^{n+k} X_i, \\ \max_{1\leq i\leq n} X_i \quad &\perp \quad \max_{n+1\leq i\leq n+k} X_i. \end{aligned}$$

Si $A_n,\ n \geq 1$ son eventos independientes entonces $\cup_{j=1}^n A_j$ y $\cup_{j=n+1}^\infty A_j$ son independientes.

1.4. El Lema de Borel-Cantelli

Definición 1.8 Sea $\{A_n,\ n \geq 1\}$ una sucesión de subconjuntos de Ω. Definimos los siguientes conjuntos:

$$\inf_{k\geq n} A_k = \bigcap_{k=n}^\infty A_k, \qquad \sup_{k\geq n} A_k = \bigcup_{k=n}^\infty A_k$$

$$\liminf_{n\to\infty} A_n = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k,$$
$$\limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

Los últimos dos conjuntos se conocen como el límite inferior y límite superior de la sucesión de conjuntos.

A partir de las definiciones tenemos las siguientes caracterizaciones de estos conjuntos. Un punto x pertenece al conjunto $\limsup A_n$ si y sólo si pertenece a infinitos conjuntos de la sucesión. En cambio x pertenece a $\liminf A_n$ si y sólo si existe un entero $p = p(x)$ tal que $x \in A_k$ para todo $k \geq p$.

Ahora es fácil ver que

$$\liminf_{n\to\infty} A_n \subset \limsup_{n\to\infty} A_n \tag{1.12}$$

ya que

$$\{x : x \in A_k,\ \text{para todo } k \geq p(x)\} \subset \{x : x \in A_k \text{ para infinitos } k\}.$$

Si la contención en sentido contrario a (1.12) es válida, es decir, si

$$\liminf_{n\to\infty} A_n = \limsup_{n\to\infty} A_n = A$$

decimos que la sucesión A_n es convergente y que su límite es A. En este caso escribimos $\lim_{n\to\infty} A_n = A$ o $A_n \to A$.

Lema 1.2 *Sea* $\{A_n,\ n \geq 1\}$ *una sucesión de subconjuntos de* Ω.

(a)

$$\limsup_{n\to\infty} A_n = \Big\{x : \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x) = \infty\Big\}$$
$$= \{x : x \in A_{k_n}, n \geq 1\}$$

para alguna subsucesión infinita k_n*, que depende de* x.

(b)

$$\liminf_{n\to\infty} A_n = \Big\{x : \sum_{n=1}^{\infty} \mathbf{1}_{A_n^c}(x) < \infty\Big\}$$
$$= \{x : x \in A_k, \forall k \geq n(x)\}$$

Demostración. (a) Si

$$x \in \limsup_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k$$

entonces para todo n, $x \in \cup_{k\geq n} A_k$, por lo tanto para todo n existe $k_n \geq n$ tal que $x \in A_{k_n}$, y en consecuencia

$$\sum_{j=1}^{\infty} \mathbf{1}_{A_j}(x) \geq \sum_{n} \mathbf{1}_{A_{k_n}}(x) = \infty,$$

lo cual implica que

$$x \in \Big\{x : \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x) = \infty\Big\}.$$

En consecuencia,

$$\limsup_{n\to\infty} A_n \subset \{x : \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x) = \infty\}.$$

Recíprocamente, si

$$x \in \Big\{x : \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x) = \infty\Big\}.$$

entonces existe $k_n \to \infty$ tal que $x \in A_{k_n}$, y por lo tanto para todo n, $x \in \cup_{j \geq n} A_j$, de modo que $x \in \limsup_{n\to\infty} A_n$. En consecuencia

$$\{x : \sum_{n=1}^{\infty} \mathbf{1}_{A_n}(x) = \infty\} \subset \limsup_{n\to\infty} A_n.$$

La demostración de (b) es similar. ■

Usaremos la notación $\{A_n, i.v.\}$ para $\limsup_n A_n$, donde $i.v.$ significa infinitas veces.

Teorema 1.5 (Borel-Cantelli) *Sea* $(A_n,\ n \geq 1)$ *una sucesión de eventos. Si* $\sum_n P(A_n) < \infty$ *entonces*

$$P(A_n\ i.v.) = P(\limsup_n A_n) = 0.$$

Demostración.

$$\begin{aligned} P(\limsup_n A_n) &= P(\lim_m \cup_{k=m}^{\infty} A_k) = \lim_m P(\cup_{k=m}^{\infty} A_k) \\ &\leq \lim_m \sum_{k=m}^{\infty} P(A_k) = 0. \end{aligned}$$

■

Ejemplo 1.7
Sea $(X_n,\ n \geq 1)$ v.a. de Bernoulli con

$$P(X_n = 1) = p_n = 1 - P(X_n = 0).$$

No suponemos independencia. Tenemos que

$$\sum_n p_n < \infty \quad \Rightarrow \quad P(\lim_n X_n = 0) = 1.$$

Como $p_n = P(X_n = 1)$, el Lema de Borel Cantelli implica que $P(X_n = 1\ i.v.) = 0$. Por lo tanto,

$$1 = P(\liminf_n \{X_n = 1\}^c) = P(\liminf_n \{X_n = 0\}).$$

En consecuencia, con probabilidad 1 las variables X_n valen 0 a partir de cierto índice y convergen a 0 trivialmente.

Teorema 1.6 (Borel-Cantelli) *Si* $(A_n,\ n \geq 1)$ *es una sucesión de eventos independientes y* $\sum P(A_n) = \infty$ *entonces* $P(\limsup_n A_n) = 1$.

Demostración. Tenemos

$$P(\limsup_n A_n) = 1 - P((\limsup_n A_n)^c) = 1 - P(\liminf_n A_n^c).$$

Por lo tanto basta ver que

$$P(\liminf_n A_n^c) = P(\cup_{n\geq 1} \cap_{k\geq n} A_k^c) = 0$$

y para esto basta ver que $P(\cap_{k\geq n} A_k^c) = 0$ para todo n. Como $1 - x \leq e^{-x}$, tenemos

$$P(\cap_{k=n}^{n+j} A_k^c) = \prod_{k=n}^{n+j} (1 - P(A_k)) \leq \prod_{k=n}^{n+j} e^{-P(A_k)} = \exp\{-\sum_{k=n}^{n+j} P(A_k)\}.$$

Como $\sum_k P(A_k) = \infty$, esta última expresión tiende a 0 cuando $j \to \infty$ y por lo tanto

$$P(\cap_{k\geq n} A_k^c) = \lim_j P(\cap_{k=n}^{n+j} A_k^c) = 0.$$

■

Corolario 1.5 (Ley 0-1) *Si $(A_n,\ n \geq 1)$ es una sucesión de eventos independientes entonces*

$$P(A_n\ i.v.) = \begin{cases} 1 & si\ \sum P(A_n) = \infty, \\ 0 & si\ \sum P(A_n) < \infty, \end{cases}$$

Ejemplo 1.8
Supongamos ahora que $(X_n,\ n \geq 1)$ son independientes Bernoulli con

$$P(X_n = 1) = p_n = 1 - P(X_n = 0),$$

entonces

$$P(X_n \to 0) = 1 \qquad \text{si y sólo si} \qquad \sum_k p_k < \infty.$$

Para esto basta observar que

$$P(X_n = 1\ i.v.) = 0 \quad \Leftrightarrow \quad \sum_k p_k < \infty.$$

Ejemplo 1.9 (Variables Exponenciales)
Supongamos que $(X_n,\ n \geq 1)$ son v.a.i. exponenciales de parámetro 1, es decir,

$$P(X_n > x) = e^{-x}, \quad x > 0.$$

Entonces

$$P\Big(\limsup_n \frac{X_n}{\log n} = 1\Big) = 1. \tag{1.13}$$

Para demostrar (1.13) usaremos el siguiente resultado, que es un ejercicio: Si $P(B_k) = 1,\ k \geq 1$, entonces $P(\cap_k B_k) = 1$. Veamos (1.13): para $\omega \in \Omega$,

$$\limsup_n \frac{X_n(\omega)}{\log n} = 1$$

quiere decir dos cosas:

1. Para todo $\varepsilon > 0$, $\frac{X_n(\omega)}{\log n} \leq 1 + \varepsilon$ para todo n suficientemente grande.

2. Para todo $\varepsilon > 0$, $\frac{X_n(\omega)}{\log n} > 1 - \varepsilon$ para infinitos valores de n

Veamos (1). Para cualquier $\varepsilon > 0$,

$$\sum_n P(X_n \geq (1+\varepsilon)\log n) = \sum_n \exp\{-(1+\varepsilon)\log n\} = \sum_n n^{-(1+\varepsilon)} < \infty$$

Por lo tanto, los eventos $\{X_n > (1+\varepsilon)\log n\}$ ocurren infinitas veces con probabilidad 0, es decir, para casi todo ω existe $N(\omega)$ tal que si $n \geq N(\omega)$,

$$X_n(\omega) < (1+\varepsilon)\log n$$

y esto es (1). Veamos (2): Para cualquier $\varepsilon > 0$,

$$\sum_n P(X_n > (1-\varepsilon)\log n) = \sum_n \exp\{-(1-\varepsilon)\log n\} = \sum_n n^{-(1-\varepsilon)} = \infty.$$

Como las X_n son independientes, los eventos $\{E_n > (1-\varepsilon)\log n\}$ ocurren infinitas veces c. p. 1.

1.5. La Ley 0-1 de Kolmogorov

Definición 1.9 Sea $(X_n,\ n \geq 1)$ una sucesión de v.a. Definimos

$$\mathcal{F}'_n = \sigma(X_{n+1}, X_{n+2}, \dots)$$

para $n = 1, 2, \dots$. Definimos la σ-álgebra cola como

$$\mathcal{T} = \cap_n \mathcal{F}'_n = \lim_{n\to\infty} \sigma(X_n, X_{n+1}, \dots).$$

Los eventos en esta σ-álgebra sólo dependen de la cola de la sucesión (X_n). Si $A \in \mathcal{T}$ decimos que A es un evento cola y si una variable es medible respecto a $\mathcal{T}$ decimos que es una variable cola.

Ejemplos 1.10

1. $\{\omega : \sum_{n\geq 1} X_n(\omega) \text{ converge }\} \in \mathcal{T}$.

   Para ver esto observamos que para todo m, $\sum_{n=1}^{\infty} X_n(\omega)$ converge si y sólo si $\sum_{n=m}^{\infty} X_n(\omega)$ converge. Por lo tanto, para todo m,

   $$\Big\{\sum_n X_n \text{ converge}\Big\} = \Big\{\sum_{n\geq m+1} X_n \text{ converge}\Big\} \in \mathcal{F}'_m$$

   y en consecuencia este evento está en la intersección de estas σ-álgebras.

2. $\limsup_{n\to\infty} X_n \in \mathcal{T}, \quad \liminf_{n\to\infty} X_n \in \mathcal{T}, \quad \{\omega : \lim_{n\to\infty} X_n(\omega) \text{ existe }\} \in \mathcal{T}$.

3. Si $S_n = X_1 + \cdots + X_n$ entonces

   $$\Big\{\omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = 0\Big\} \in \mathcal{T}$$

   ya que para cualquier m,

   $$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i(\omega) = \lim_{n\to\infty} \frac{1}{n}\sum_{i=m+1}^{n} X_i(\omega) \in \mathcal{F}'_m$$

Teorema 1.7 (Ley 0-1 de Kolmogorov) *Sea $(X_n,\ n \geq 1)$ una sucesión de v.a.i. con σ-álgebra cola $\mathcal{T}$. Entonces si $\Lambda \in \mathcal{T}$ se tiene que $P(\Lambda) = 0$ ó 1.*

Demostración. Sea $\Lambda \in \mathcal{T}$, demostraremos que Λ es independiente de si mismo, de modo que

$$P(\Lambda) = P(\Lambda \cap \Lambda) = P(\Lambda)P(\Lambda).$$

En consecuencia $P(\Lambda) = P(\Lambda)^2$ y $P(\Lambda) = 0$ ó $P(\Lambda) = 1$.

Para ver que Λ es independiente de si mismo definimos

$$\mathcal{F}_n = \sigma(X_1, \dots, X_n)$$

de modo que $\mathcal{F}_n$ es una sucesión creciente y

$$\mathcal{F}_\infty = \sigma(X_1, X_2, \dots) = \sigma(\cup_{n=1}^\infty \mathcal{F}_n) = \sigma(\cup_{n=1}^\infty \sigma(X_n)).$$

Observamos que

$$\Lambda \in \mathcal{T} \subset \mathcal{F}'_n = \sigma(X_{n+1}, X_{n+2}, \dots) \subset \sigma(X_1, X_2, \dots) = \mathcal{F}_\infty.$$

Por otro lado, para todo n tenemos

$$\Lambda \in \mathcal{F}'_n$$

y como $\mathcal{F}_n \perp \mathcal{F}'_n$, tenemos $\Lambda \perp \mathcal{F}_n$ para todo n y en consecuencia

$$\Lambda \perp \cup_n \mathcal{F}_n.$$

Sea $\mathcal{C}_1 = \{\Lambda\}$ y $\mathcal{C}_2 = \cup_n \mathcal{F}_n$, entonces $\mathcal{C}_i$ es un sistema π para $i = 1, 2$, $\mathcal{C}_1 \perp \mathcal{C}_2$ y por el Criterio Básico

$$\sigma(\mathcal{C}_1) = \{\emptyset, \Omega, \Lambda, \Lambda^c\}, \quad \text{y} \quad \sigma(\mathcal{C}_2) = \sigma(\cup_n \mathcal{F}_n) = \mathcal{F}_\infty$$

son independientes. Pero $\Lambda \in \mathcal{C}_1 \subset \sigma(\mathcal{C}_1)$ y $\Lambda \in \sigma(\mathcal{C}_2) = \mathcal{F}_\infty$. Por lo tanto Λ es independiente de Λ. ■

Definición 1.10 Una σ-álgebra es *casi trivial* si todos sus eventos tienen probabilidad 0 ó 1.

Lema 1.3 *Sea $\mathcal{G}$ una σ-álgebra casi trivial y $X \in \mathcal{G}$. Entonces existe c tal que $P(X = c) = 1$.*

Demostración. $F(x) = P(X \le x)$ es no-decreciente y $\{X \le x\} \in \sigma(X) \subset \mathcal{G}$, de modo que $F(x) = 0$ ó 1, para cada $x \in \mathbb{R}$. Sea

$$c = \sup\{x : F(x) = 0\}.$$

La f. d. debe tener un salto de tamaño 1 en c y por lo tanto $P(X = c) = 1$. ■

Corolario 1.6 *Sea $(X_n,\ n \ge 1)$ una sucesión de v.a. Las siguientes proposiciones son ciertas*

(a) $\{\sum_n X_n$ *converge* $\}$ *tiene probabilidad* 0 *ó* 1.

(b) Las variables $\limsup_n X_n$ *y* $\liminf_n X_n$ *son constantes c. p. 1.*

(c) $\{\omega : \frac{1}{n} S_n(\omega) \to 0\}$ *tiene probabilidad* 0 *ó* 1.

Demostración. Ejercicio.

1.6. Espacios Producto

Sean $\Omega_1,\ \Omega_2$ dos espacios, definimos el espacio producto como

$$\Omega_1 \times \Omega_2 = \{(\omega_1, \omega_2) : \omega_i \in \Omega_i, i = 1, 2\}$$

y definimos los operadores de proyección por

$$\pi_i(\omega_1, \omega_2) = \omega_i, \quad i = 1, 2$$

de modo que $\pi_i : \Omega_1 \times \Omega_2 \to \Omega_i$. Si $A \subset \Omega_1 \times \Omega_2$ definimos las *secciones* de A por

$$A_{\omega_1} = \{\omega_2 : (\omega_1, \omega_2) \in A\} \subset \Omega_2; \qquad A_{\omega_2} = \{\omega_1 : (\omega_1, \omega_2) \in A\} \subset \Omega_1$$

Propiedades

(i) Si $A \subset \Omega_1 \times \Omega_2$ entonces $(A^c)_{\omega_1} = (A_{\omega_1})^c$.
(ii) Si para un conjunto de índices T tenemos $A_t \subset \Omega_1 \times \Omega_2$ para todo $t \in T$, entonces

$$\big(\cup_t A_t\big)_{\omega_1} = \cup_t (A_t)_{\omega_1}, \qquad \big(\cap_t A_t\big)_{\omega_1} = \cap_t (A_t)_{\omega_1}.$$

Consideremos ahora una función X con dominio $\Omega_1 \times \Omega_2$ y valores en algún espacio S, que suponemos es un espacio métrico. Definimos la *sección* de X como

$$X_{\omega_1}(\omega_2) = X(\omega_1, \omega_2)$$

de modo que $X_{\omega_1} : \Omega_2 \to S$. En general ω_1 está fijo y la sección es una función de ω_2. Decimos que X_{ω_1} es la *sección* de X en ω_1. De manera similar definimos la sección $X_{\omega_2} : \Omega_1 \to S$.

Algunas propiedades básicas de las secciones son las siguientes:
(i) $(\mathbf{1}_A)_{\omega_1} = \mathbf{1}_{A_{\omega_1}}$
(ii) Si $S = \mathbb{R}^k$ para $k \geq 1$ y si para $i = 1, 2$ tenemos $X_i : \Omega_1 \times \Omega_2 \to S$, entonces

$$(X_1 + X_2)_{\omega_1} = (X_1)_{\omega_1} + (X_2)_{\omega_2}.$$

(iii) Sea S un espacio métrico, $X_n : \Omega_1 \times \Omega_2 \to S$ y $\lim_{n\to\infty} X_n$ existe. Entonces

$$\lim_{n\to\infty} (X_n)_{\omega_1} = \lim_{n\to\infty} (X_n)_{\omega_1}.$$

Un *rectángulo* en $\Omega_1 \times \Omega_2$ es un subconjunto de $\Omega_1 \times \Omega_2$ de la forma $A_1 \times A_2$ con $A_i \subset \Omega_i$ para $i = 1, 2$. Llamamos a A_1 y A_2 los *lados* del rectángulo. El rectángulo es vacío si al menos uno de los lados es vacío.

Si $(\Omega_i, \mathcal{F}_i)$ son espacios medibles para $i = 1, 2$ decimos que un rectángulo es *medible* si es de la forma $A_1 \times A_2$ con $A_i \in \mathcal{F}_i$, $i = 1, 2$. La clase de los rectángulos medibles es una semiálgebra que denotaremos $\mathcal{R}$ o por $\mathcal{F}_1 \times \mathcal{F}_2$ si deseamos describir de manera explícita las σ-álgebras. Verifiquemos que $\mathcal{R}$ es una semiálgebra.
(i) $\emptyset, \Omega \in \mathcal{R}$.
(ii) $\mathcal{R}$ es un sistema π: Si $A_1 \times A_2,\ B_1 \times B_2 \in \mathcal{R}$ entonces $(A_1 \times A_2) \cap (B_1 \times B_2) = A_1 \cap B_1 \times A_2 \cap B_2 \in \mathcal{R}$.
(iii) $\mathcal{R}$ es cerrada bajo complementos. Supongamos que $A_1 \times A_2 \in \mathcal{R}$. Entonces

$$\Omega_1 \times \Omega_2 - A_1 \times A_2 = (\Omega_1 - A_1) \times A_2 + A_1 \times (\Omega_2 - A_2) + A_1^c \times A_2^c.$$

Ahora definimos la σ-álgebra $\mathcal{F}_1 \otimes \mathcal{F}_2$ como la menor σ-álgebra en $\Omega_1 \times \Omega_2$ que contiene a $\mathcal{R}$ y la llamamos la σ-álgebra producto:

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{R}) = \sigma(\mathcal{F}_1 \times \mathcal{F}_2).$$

Observamos que si $\Omega_1 = \Omega_2 = \mathbb{R}$ entonces

$$\mathcal{B}_1 \otimes \mathcal{B}_2 = \sigma(A_1 \times A_2 : A_i \in \mathcal{B}(\mathbb{R}), i = 1, 2).$$

Hay otras maneras de generar la σ-álgebra producto en $\mathbb{R}^2$. Si $\mathcal{C}$ es la clase de los intervalos de la forma $(a, b]$, usando un argumento de inducción es posible probar que

$$\mathcal{B}_1 \otimes \mathcal{B}_2 = \sigma(\{I_1 \times I_2 : I_j \in \mathcal{C}, j = 1, 2\}).$$

Lema 1.4 *Las secciones de conjuntos medibles son medibles. Si $A \in \mathcal{F}_1 \otimes \mathcal{F}_2$ entonces para todo $\omega_1 \in \Omega_1$,*

$$A_{\omega_1} \in \mathcal{F}_2.$$

Demostración. Definimos

$$\mathcal{C}_{\omega_1} = \{A \subset \Omega_1 \times \Omega_2,\ A_{\omega_1} \in \mathcal{F}_2\}.$$

Si $A \in \mathcal{R}$ y $A = A_1 \times A_2$ donde $A_i \in \mathcal{F}_i$, entonces

$$\begin{aligned} A_{\omega_1} &= \{\omega_2 : (\omega_1, \omega_2) \in A_1 \times A_2\} \\ &= \begin{cases} A_2 \in \mathcal{F}_2, & \text{si } \omega_1 \in A_1 \\ \emptyset \in \mathcal{F}_2, & \text{ai } \omega_1 \notin A_1. \end{cases} \end{aligned}$$

Por lo tanto $A_{\omega_1} \in \mathcal{C}_{\omega_1}$ lo que implica que $\mathcal{R} \subset \mathcal{C}_{\omega_1}$. Además $\mathcal{C}_{\omega_1}$ es un sistema λ:

(a) $\Omega_1 \times \Omega_2 \in \mathcal{C}_{\omega_1}$ porque $\Omega_1 \times \Omega_2 \in \mathcal{R}$.

(b) Si $A \in \mathcal{C}_{\omega_1}$ entonces $A^c \in \mathcal{C}_{\omega_1}$ ya que $(A^c)_{\omega_1} = (A_{\omega_1})^c$ y $A \in \mathcal{C}_{\omega_1}$ implica $A_{\omega_1} \in \mathcal{F}_2$ y por lo tanto $(A_{\omega_1})^c \in \mathcal{F}_2$ Esto implica que $(A^c)_{\omega_1} \in \mathcal{C}_{\omega_1}$.

(c) Si $A_n \in \mathcal{C}_{\omega_1}$ para $n \geq 1$ con (A_n) disjuntos, entonces $(A_n)_{\omega_1} \in \mathcal{F}_2$ implica $\cup_n (A_n)_{\omega_1} \in \mathcal{F}_2$. Pero

$$\cup_{n=1}^{\infty} (A_n)_{\omega_1} = \Big(\cup_{n=1}^{\infty} A_n\Big)_{\omega_1} \in \mathcal{F}_2$$

y por lo tanto $\cup_n (A_n)_{\omega_1} \in \mathcal{C}_{\omega_1}$.

Tenemos, en consecuencia, que $\mathcal{C}_{\omega_1}$ es un sistema λ y $\mathcal{R} \subset \mathcal{C}_{\omega_1}$, y por el teorema de Dynkin

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{R}) \subset \mathcal{C}_{\omega_1}.$$

■

Corolario 1.7 *Las secciones de funciones medibles son medibles, es decir, si $X : (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2) \to (S, \mathcal{S})$ entonces $X_{\omega_1} \in \mathcal{F}_2$.*

Demostración. Como X es $\mathcal{F}_1 \otimes \mathcal{F}_2/\mathcal{S}$ medible, tenemos para $\Lambda \in \mathcal{S}$ que

$$\{(\omega_1, \omega_2) : X(\omega_1, \omega_2) \in \Lambda\} = X^{-1}(\Lambda) \in \mathcal{F}_1 \otimes \mathcal{F}_2,$$

y por los resultados anteriores

$$(X^{-1}(\Lambda))_{\omega_1} \in \mathcal{F}_2.$$

Sin embargo

$$\begin{aligned} (X^{-1}(\Lambda))_{\omega_1} &= \{\omega_2 : X(\omega_1, \omega_2) \in \Lambda\} \\ &= \{\omega_2 : X_{\omega_1}(\omega_2) \in \Lambda\} = (X_{\omega_1}^{-1}(\Lambda)), \end{aligned}$$

lo que nos dice que X_{ω_1} es $\mathcal{F}_2/\mathcal{S}$ medible. ■

1.7. Medidas en Espacios Producto

Definición 1.11 Una función $K(\omega_1, A_2) : \Omega_1 \times \mathcal{F}_2 \to [0,1]$ es una función (o núcleo o kernel) de transición si

(i) $\forall \omega_1$, $K(\omega_1, \cdot)$ es una medida de probabilidad en $\mathcal{F}_2$.
(ii) $\forall A_2 \in \mathcal{F}_2$, $K(\cdot, A_2)$ es una función $(\mathcal{F}_1, \mathcal{B}[0,1])$-medible.

Las funciones de transición se usan para definir los procesos de Markov a tiempo discreto. En este caso $K(\omega_1, A_2)$ representa la probabilidad condicional de ir al conjunto de estados A_2 si partimos del estado ω_1.

Teorema 1.8 *Sea P_1 una probabilidad sobre $\mathcal{F}_1$ y $K : \Omega_1 \times \mathcal{F}_2 \to [0,1]$ una función de transición. Entonces K y P_1 determinan una única probabilidad sobre $\mathcal{F}_1 \otimes \mathcal{F}_2$ por la fórmula*

$$P(A_1 \times A_2) = \int_{A_1} K(\omega_1, A_2)\, dP_1(\omega_1) \tag{1.14}$$

para todo $A_1 \times A_2 \in \mathcal{F}_1 \otimes \mathcal{F}_2$.

Demostración. La medida P está definida por (1.14) sobre $\mathcal{F}_1 \times \mathcal{F}_2$ y debemos verificar que se satisfacen las condiciones del Teorema de Extensión. Veamos que P es σ-aditiva en $\mathcal{F}_1 \times \mathcal{F}_2$. Sea $\{A^{(n)} \times B^{(n)}\}$ una sucesión de conjuntos disjuntos en $\mathcal{F}_1 \times \mathcal{F}_2$, cuya unión también es un conjunto en $\mathcal{F}_1 \times \mathcal{F}_2$:

$$\bigcup_n (A^{(n)} \times B^{(n)}) = A \times B.$$

Tenemos

$$\mathbf{1}_A(\omega_1)\mathbf{1}_B(\omega_2) = \mathbf{1}_{A\times B}(\omega_1, \omega_2) = \sum_n \mathbf{1}_{A^{(n)}\times B^{(n)}}(\omega_1, \omega_2) = \sum_n \mathbf{1}_{A^{(n)}}(\omega_1)\mathbf{1}_{B^{(n)}}(\omega_2)$$

Usando el Teorema de Convergencia Monótona para series tenemos

$$\begin{aligned}
P(A \times B) &= \int_{\Omega_1} \mathbf{1}_A(\omega_1) K(\omega_1, B)\, dP_1(\omega_1) = \int_{\Omega_1} \big(\int_{\Omega_2} \mathbf{1}_A(\omega_1)\mathbf{1}_B(\omega_2) K(\omega_1, d\omega_2)\big)\, dP_1(\omega_1) \\
&= \int_{\Omega_1} \big(\int_{\Omega_2} \sum_n \mathbf{1}_{A^{(n)}}(\omega_1)\mathbf{1}_{B^{(n)}}(\omega_2) K(\omega_1, d\omega_2)\big)\, dP_1(\omega_1) \\
&= \int_{\Omega_1} \sum_n \int_{\Omega_2} \mathbf{1}_{A^{(n)}}(\omega_1)\mathbf{1}_{B^{(n)}}(\omega_2) K(\omega_1, d\omega_2)\, dP_1(\omega_1) \\
&= \sum_n \int_{\Omega_1} \mathbf{1}_{A^{(n)}}(\omega_1)\big(\int_{\Omega_2} \mathbf{1}_{B^{(n)}}(\omega_2) K(\omega_1, d\omega_2)\big)\, dP_1(\omega_1) \\
&= \sum_n \int_{\Omega_1} \mathbf{1}_{A^{(n)}}(\omega_1) K(\omega_1, B^{(n)})\, dP_1(\omega_1) \\
&= \sum_n \int_{A^{(n)}} K(\omega_1, B^{(n)})\, dP_1(\omega_1) = \sum_n P(A^{(n)} \times B^{(n)}).
\end{aligned}$$

■

Caso Especial

Supongamos que para alguna medida de probabilidad P_2 en $\mathcal{F}_2$ tenemos $K(\omega_1, B) = P_2(B)$ para todo $\omega_1 \in \Omega_1$, entonces la medida P definida por (1.14) sobre $\mathcal{F}_1 \otimes \mathcal{F}_2$ es

$$P(A \times B) = P_1(A)P_2(B)$$

Denotamos esta probabilidad por $P_1 \times P_2$ y la llamamos la medida producto. Definimos las siguientes σ-álgebras en $\Omega_1 \times \Omega_2$:

$$\mathcal{F}_1^\sharp = \{A \times \Omega_2,\ A \in \mathcal{F}_1\}, \qquad \mathcal{F}_2^\sharp = \{\Omega_1 \times B,\ B \in \mathcal{F}_2\}.$$

Con respecto a la medida producto $P = P_1 \times P_2$ tenemos $\mathcal{F}_1^\sharp \perp \mathcal{F}_2^\sharp$ ya que

$$P(A \times \Omega_2 \cap \Omega_1 \times B) = P(A \times B) = P_1(A)P_2(B) = P(A \times \Omega_2)P(\Omega_1 \times B).$$

Supongamos $X_i : (\Omega_i, \mathcal{F}_i) \to (\mathbb{R}, \mathcal{B})$ es una v.a. sobre Ω_i para $i = 1, 2$. Definimos las siguientes funciones en $\Omega_1 \times \Omega_2$:

$$X_1^\sharp(\omega_1, \omega_2) = X_1(\omega_1) = X_1 \circ \pi_1(\omega_1, \omega_2), \qquad X_2^\sharp(\omega_1, \omega_2) = X_2(\omega_2) = X_2 \circ \pi_2(\omega_1, \omega_2).$$

Con respecto a $P = P_1 \times P_2$, las variables $X_1^\sharp$ y $X_2^\sharp$ son independientes ya que

$$\begin{aligned}
P(X_1^\sharp \leq x,\ X_2^\sharp \leq y) &= P_1 \times P_2\big(\{(\omega_1, \omega_2) : X_1(\omega_1) \leq x,\ X_2(\omega_2) \leq y\}\big) \\
&= P_1 \times P_2\big(\{\omega_1 : X_1(\omega_1) \leq x\} \times \{\omega_1 : X_2(\omega_2) \leq y\}\big) \\
&= P_1\big(\{\omega_1 : X_1(\omega_1) \leq x\}\big) \times P_2\big(\{\omega_2 : X_2(\omega_2) \leq y\}\big) \\
&= P\big(\{(\omega_1, \omega_2) : X_1^\sharp(\omega_1, \omega_2) \leq x\}\big) \times P\big(\{(\omega_1, \omega_2) : X_2^\sharp(\omega_1, \omega_2) \leq y\}\big) \\
&= P\big(X_1^\sharp \leq x\big) \times P\big(X_2^\sharp \leq y\}\big).
\end{aligned}$$

Vemos que la independencia es parte del modelo cuando se usa la medida producto. Podemos afirmar que las cantidades que dependen de componentes disjuntas son independientes.

Es posible extender estas construcciones de 2 a $d \geq 2$ factores y definir la medida producto $P_1 \times \cdots \times P_d$. Más aún, es posible definir la medida producto sobre espacios producto arbitrarios $\prod_{i \in I} \Omega_i$: Si $(\Omega_i, \mathcal{F}_i, P_i)$ son espacios de probabilidad existe una única medida P definida sobre la σ-álgebra producto $\mathcal{F}$ de subconjuntos de $\Omega = \prod_{i \in I} \Omega_i$ que es generada por los cilindros de la forma

$$E_1 \times E_2 \times \cdots \times E_n \times \prod_{i=n+1}^{\infty} \Omega_i, \quad E_i \in \mathcal{F}_i,\ i \geq 1,$$

tal que

$$P(E_1 \times E_2 \times \cdots \times E_n \times \prod_{i=n+1}^{\infty} \Omega_i) = P_1(E_1) \cdots P_n(E_n).$$

1.8. El Teorema de Fubini

Teorema 1.9 *Sea P_1 una probabilidad sobre $(\Omega_1, \mathcal{F}_1)$ y sea $K : \Omega_1 \times \mathcal{F}_2 \to [0, 1]$ un kernel de transición. Definimos P sobre el espacio producto $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2)$ por*

$$P(A_1 \times A_2) = \int_{A_1} K(\omega_1, A_2)\, dP_1(\omega_1). \tag{1.15}$$

Supongamos que $X : (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2) \to (\mathbb{R}, \mathcal{B})$ *es positiva o integrable, entonces*

$$Y(\omega_1) = \int_{\Omega_2} X_{\omega_1}(\omega_2) K(\omega_1, d\omega_2)$$

tiene las siguientes propiedades:
(a) Y *está bien definida.*
(b) $Y \in \mathcal{F}_1$.
(c) $Y \geq 0$, *o* $Y \in L^1(P_1)$.
y además

$$\int_{\Omega_1 \times \Omega_2} X\, dP = \int_{\Omega_1} Y(\omega_1)\, dP_1(\omega_1) = \int_{\Omega_1} \big(\int_{\Omega_2} X_{\omega_1}(\omega_2) K(\omega_1, d\omega_2) \big)\, dP_1(\omega_1). \tag{1.16}$$

Demostración. Para ω_1 fijo tenemos que $X_{\omega_1}(\omega_2)$ es $\mathcal{F}_2$-medible, de modo que Y está bien definida. No es difícil ver que Y es $\mathcal{F}_1$-medible. Veamos (1.16) bajo la hipótesis $X \geq 0$. Definimos

$$A = \int_{\Omega_1 \times \Omega_2} X\, dP, \qquad \text{y} \qquad B = \int_{\Omega_1} Y(\omega_1)\, dP_1(\omega_1).$$

Supongamos inicialmente que $X = \mathbf{1}_{A_1 \times A_2}$ con $A_1 \times A_2 \in \mathcal{F}_1 \times \mathcal{F}_2$. Entonces

$$A = \int_{A_1 \times A_2} dP = P(A_1 \times A_2)$$

y

$$\begin{aligned} B &= \int_{\Omega_1} \big(\int_{\Omega_2} \mathbf{1}_{A_1}(\omega_1) \mathbf{1}_{A_2}(\omega_2) K(\omega_1, d\omega_2) \big) dP_1(\omega_1) \\ &= \int_{A_1} K(\omega_1, A_2) \big) dP_1(\omega_1) = P(A_1 \times A_2) \end{aligned}$$

de modo que (1.16) vale para indicadores de rectángulos medibles. Sea

$$\mathcal{C} = \{C \in \mathcal{F} \otimes \mathcal{F}_2 :\ (1.16) \text{ vale para } X = \mathbf{1}_C\}$$

y sabemos que $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{C}$. Veamos que $\mathcal{C}$ es un sistema λ:
(i) $\Omega_1 \times \Omega_2 \in \mathcal{C}$ porque $\Omega_1 \times \Omega_2 \in \mathcal{F}_1 \times \mathcal{F}_2$.
(ii) Si $C \in \mathcal{C}$ entonces para $X = \mathbf{1}_{C^c}$ tenemos $A = P(C^c) = 1 - P(C)$. Por lo tanto

$$\begin{aligned} A &= 1 - \iint \mathbf{1}_{C_{\omega_1}}(\omega_2) K(\omega_1, d\omega_2)\, dP_1(\omega_1) \\ &= \iint (1 - \mathbf{1}_{C_{\omega_1}}(\omega_2)) K(\omega_1, d\omega_2)\, dP_1(\omega_1) \\ &= \iint (\mathbf{1}_{(C^c)_{\omega_1}}(\omega_2)) K(\omega_1, d\omega_2)\, dP_1(\omega_1) = B, \end{aligned}$$

de modo que $C^c \in \mathcal{C}$.
(iii) Si $C_n \in \mathcal{B}$, $n \geq 1$ son eventos disjuntos

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} \mathbf{1}_{\cup_{n=1}^\infty C_n} dP &= P(\cup_{n=1}^\infty C_n) = \sum_{n=1}^\infty P(C_n) \\ &= \sum_{n=1}^\infty \iint \mathbf{1}_{(C_n)_{\omega_1}}(\omega_2) K(\omega_1, d\omega_2)\, dP_1(\omega_1) \end{aligned}$$

porque $C_n \in \mathcal{C}$. Usando el TCM tenemos que esta expresión es

$$
\begin{aligned}
&= \iint \sum_{n=1}^{\infty} \mathbf{1}_{(C_n)_{\omega_1}}(\omega_2) K(\omega_1, d\omega_2)\, dP_1(\omega_1) \\
&= \iint \mathbf{1}_{\cup_n (C_n)_{\omega_1}}(\omega_2) K(\omega_1, d\omega_2)\, dP_1(\omega_1) = B
\end{aligned}
$$

y por lo tanto $\cup_n C_n \in \mathcal{C}$.

Hemos verificado que $\mathcal{C}$ es un sistema λ y $\mathcal{F}_1 \times \mathcal{F}_2 \subset \mathcal{C}$, lo que implica por el teorema de Dynkin que

$$\mathcal{F}_1 \otimes \mathcal{F}_2 = \sigma(\mathcal{F}_1 \times \mathcal{F}_2) \subset \mathcal{C}.$$

Concluimos que si $C \in \mathcal{F}_1 \otimes \mathcal{F}_2$ y $X = \mathbf{1}_C$ entonces (1.16) vale para X.

Ahora escribimos (1.16) como $A(X) = B(X)$. Tanto $A(X)$ como $B(X)$ son lineales en X, de modo que (1.16) vale para funciones simples de la forma

$$X = \sum_{i=1}^{k} c_i \mathbf{1}_{C_i}, \quad C_i \in \mathcal{F}_1 \otimes \mathcal{F}_2.$$

Para $X \geq 0$ existe una sucesión de v.a. simples X_n con $0 \leq X_n \uparrow X$. Tenemos que $A(X_n) = B(X_n)$ y por convergencia monótona

$$A(X_n) \uparrow A(X).$$

Para B, usando el TCM dos veces obtenemos

$$
\begin{aligned}
\lim_{n\to\infty} B(X_n) &= \lim_{n\to\infty} \int_{\Omega_1} \big(\int_{\Omega_2} (X_n)_{\omega_1}(\omega_2) K(\omega_1, d\omega_2) \big)\, dP_1(\omega_1) \\
&= \int_{\Omega_1} \big(\lim_{n\to\infty} \int_{\Omega_2} (X_n)_{\omega_1}(\omega_2) K(\omega_1, d\omega_2) \big)\, dP_1(\omega_1) \\
&= \int_{\Omega_1} \big(\int_{\Omega_2} \lim_{n\to\infty} (X_n)_{\omega_1}(\omega_2) K(\omega_1, d\omega_2) \big)\, dP_1(\omega_1) \\
&= \int_{\Omega_1} \big(\int_{\Omega_2} (X)_{\omega_1}(\omega_2) K(\omega_1, d\omega_2) \big)\, dP_1(\omega_1) = B(X).
\end{aligned}
$$

■

Teorema 1.10 (Fubini-Tonelli) *Sea $P = P_1 \times P_2$ la medida producto. Si X es $\mathcal{F}_1 \otimes \mathcal{F}_2$ medible y es no-negativa o integrable respecto a P entonces*

$$
\begin{aligned}
\int_{\Omega_1 \times \Omega_2} X\, dP &= \int_{\Omega_1} \big(\int_{\Omega_2} X_{\omega_1}(\omega_2)\, dP_2(\omega_2) \big)\, dP_1(\omega_1) = \int_{\Omega_1} \big(\int_{\Omega_2} X(\omega_1, \omega_2)\, dP_2(\omega_2) \big)\, dP_1(\omega_1) \\
&= \int_{\Omega_2} \big(\int_{\Omega_1} X_{\omega_2}(\omega_1)\, dP_1(\omega_1) \big)\, dP_2(\omega_2) = \int_{\Omega_2} \big(\int_{\Omega_1} X(\omega_2, \omega_1)\, dP_1(\omega_1) \big)\, dP_2(\omega_2).
\end{aligned}
$$

Demostración. Sea $K(\omega_1, A_2) = P_2(\omega_2)$, entonces P_1 y K determinan a $P = P_1 \times P_2$ en $\mathcal{F}_1 \otimes \mathcal{F}_2$ y

$$\int_{\Omega_1 \times \Omega_2} X\, dP = \int_{\Omega_1} \big(\int_{\Omega_2} X_{\omega_1}(\omega_2)\, dP_2(\omega_2) \big)\, dP_1(\omega_1) = \int_{\Omega_2} \big(\int_{\Omega_1} X_{\omega_2}(\omega_1)\, dP_1(\omega_1) \big)\, dP_2(\omega_2).$$

Por otro lado si $\widetilde{K}(\omega_2, A_1) = P_1(A_1)$ es una función de transición $\widetilde{K} : \Omega_2 \times \mathcal{F}_1 \to [0, 1]$ entonces $\widetilde{K}$ y P_1 también determinan a $P = P_1 \times P_2$ y

$$\int_{\Omega_1 \times \Omega_2} X\, dP = \int_{\Omega_2} \big(\int_{\Omega_1} X_{\omega_2}(\omega_1)\, dP_1(\omega_1) \big)\, dP_2(\omega_2) = \int_{\Omega_1} \big(\int_{\Omega_2} X_{\omega_1}(\omega_2)\, dP_2(\omega_2) \big)\, dP_1(\omega_1)$$

■

Teorema 1.11 *Si $(\Omega_1, \mathcal{F}_1, \mu_1)$ y $(\Omega_2, \mathcal{F}_2, \mu_2)$ son espacios de medida σ-finitos. Entonces*

$$\pi(E) = \int_{\Omega_1} \nu\{\omega_2 : (\omega_1, \omega_2) \in E\}\, d\mu(\omega_1) = \int_{\Omega_2} \mu\{\omega_1 : (\omega_1, \omega_2) \in E\}\, d\nu(\omega_2)$$

define una medida σ-finita sobre $\mathcal{F}_1 \otimes \mathcal{F}_2$. Es la única medida para la cual $\pi(A \times B) = \mu(A)\nu(B)$ para rectángulos medibles.

Teorema 1.12 *Bajo las hipótesis del teorema anterior, si $f : \Omega_1 \times \Omega_2 \to \mathbb{R}$ satisface $f \geq 0$, las funciones*

$$\int_{\Omega_2} f(\omega_1, \omega_2)\, d\nu(\omega_2), \qquad \int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu(\omega_1) \tag{1.17}$$

son medibles respecto a $\mathcal{F}_1$ y $\mathcal{F}_2$, respectivamente y

$$\begin{aligned} \int_{\Omega_1 \times \Omega_2} f(\omega_1, \omega_2)\, d\pi(\omega_1, \omega_2) &= \int_{\Omega_2} \big(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\mu(\omega_1) \big)\, d\nu(\omega_2) \\ &= \int_{\Omega_2} \big(\int_{\Omega_1} f(\omega_1, \omega_2)\, d\nu(\omega_2) \big)\, d\mu(\omega_1). \end{aligned} \tag{1.18}$$

Si f es integrable respecto a π, las dos funciones (1.17) son medibles y finitas en A_0 y B_0 respectivamente, donde $\mu(\Omega_1 - A_0) = \nu(\Omega_2 - B_0) = 0$ y de nuevo (1.18) vale.

Como aplicación del teorema de Fubini veamos la demostración de la fórmula de integración por partes

Teorema 1.13 *Sea $a < b$ en $\mathbb{R}$ y supongamos que F, G son funciones de Stieltjes que no tienen puntos de discontinuidad en común en $(a, b]$. Entonces*

$$\int_a^b G(x)\, dF(x) = G(b)F(b) - G(a)F(a) - \int_a^b F(x)\, dG(x).$$

Si además G es absolutamente continua con densidad g entonces

$$\int_a^b G(x)\, dF(x) = G(b)F(b) - G(a)F(a) - \int_a^b F(x)g(x)\, dx.$$

Demostración. Observamos que si la fórmula vale para F y G entonces, por linealidad también vale para transformaciones lineales $\alpha F + \beta$ y $\gamma G + \delta$, ya que

$$\begin{aligned} \int_a^b \gamma G(x)\, d(\alpha F(x)) &= \gamma\alpha \int_a^b G(x)\, dF(x) \\ &= \gamma\alpha\Big(G(b)F(b) - G(a)F(a) - \int_a^b F(x)\, dG(x) \Big) \\ &= (\gamma G(b))(\alpha F(b)) - (\gamma G(a))(\alpha F(a)) - \int_a^b (\alpha F(x))\, d(\gamma G(x)), \end{aligned}$$

y

$$\begin{aligned} \int_a^b (G(x) + \delta)\, d(F(x) + \beta) &= \int_a^b G(x)\, dF(x) + \delta(F(b) - F(a)) \\ &= G(b)F(b) - G(a)F(a) - \int_a^b F(x)\, dG(x) + \delta(F(b) - F(a)) \\ &= (G(b) + \delta)F(b) - (G(a) + \delta)F(a) - \int_a^b F(x)\, d(G(x) + \delta). \end{aligned}$$

Por lo tanto no hay pérdida de generalidad al suponer que F y G son funciones de distribución, que asociamos con las variables aleatorias X e Y, respectivamente. Usando ahora el teorema de Fubini obtenemos, por un lado, que

$$\begin{aligned} P(a < X \le b, a < Y \le b) &= \int_a^b \int_a^b d(F \times G)(x,y) = \int_a^b \int_a^b dF(x)\, dG(y) \\ &= \int_a^b dF(x) \int_a^b dG(y) = \big(F(b) - F(a)\big)\big(G(b) - G(a)\big) \end{aligned}$$

Por otro lado, la probabilidad de que el punto (X, Y) caiga dentro del cuadrado $(a, b] \times (a, b]$ la podemos dividir en tres y obtenemos

$$\begin{aligned} P(a < X \le b, a < Y \le b) &= P(a < X < Y \le b) + P(a < Y < X \le b) + P(a < Y = X \le b) \\ &= \int_a^b \int_a^b d(F \times G)(x,y) + \int_a^b \int_a^b d(G \times F)(x,y) + 0 \\ &= \int_a^b \Big(\int_a^b dF(y)\Big) dG(x) + \int_a^b \Big(\int_a^b dG(y)\Big) dF(x) \\ &= \int_a^b (F(x) - F(a))\, dG(x) + \int_a^b (G(x) - G(a))\, dF(x) \\ &= \int_a^b F(x)\, dG(x) + \int_a^b G(x)\, dF(x) - F(a)(G(b) - G(a)) - G(a)(F(b) - F(a)). \end{aligned}$$

La fórmula para integración parcial sigue igualando las dos expresiones. La conclusión para el caso especial cuando G es absolutamente continua sigue del hecho que

$$\int_a^b F(x)\, dG(x) = \int_a^b F(x) g(x)\, dx.$$

■

Ejemplo 1.11 (Tiempos de Ocupación)
Sea $\{X(t, \omega),\ t \in [0, 1]\}$ un proceso sobre $(\Omega, \mathcal{F}, P)$ que satisface

(a) El proceso toma valores en $\mathbb{R}$.

(b) El proceso es medible simultáneamente en ambas variables: Para $\Lambda \in \mathcal{B},\ X^{-1}(\Lambda) = \{(t, \omega) : X(t, \omega) \in \Lambda\} \in \mathcal{B} \otimes \mathcal{F}$.

Fijamos $\Lambda \in \mathcal{B}$ y nos interesa el 'tiempo de ocupación' de X en Λ durante lel período $t \in A$ para $A \in \mathcal{B}[0, 1]$. Como $\Lambda \in \mathcal{B}(\mathbb{R})$ la función

$$\mathbf{1}_\Lambda : (\mathbb{R}, \mathcal{B}(\mathbb{R})) \to (\{0, 1\}, \big\{\emptyset, \{0, 1\}, \{0\}, \{1\}\big\})$$

es medible y por lo tanto

$$\mathbf{1}_\Lambda\big(X(s, \omega)\big) : ([0, 1] \times \Omega, \mathcal{B}[0, 1] \times \mathcal{F}) \to (\{0, 1\}, \mathcal{F}(\{0, 1\})).$$

Definimos la medida aleatoria

$$\chi(A, \omega) = \int_A \mathbf{1}_\Lambda\big(X(s, \omega)\big)\, ds$$

que llamaremos el tiempo de ocupación en Λ para $t \in A$. Tenemos

$$\mathrm{E}[\chi(A, \omega)] = \int_\Omega \Big(\int_A \mathbf{1}_\Lambda\big(X(s, \omega)\big)\, ds\Big)\, dP$$

y usando Fubini esto es

$$= \int_A \Big(\int_\Omega A \mathbf{1}_\Lambda \big(X(s,\omega)\big)\, dP\Big)\, ds = \int_A P(X(s) \in \Lambda)\, ds.$$

■

1.9. Espacios Producto e Independencia

Teorema 1.14 *Sean X, Y v.a. sobre $(\Omega, \mathcal{A}, P)$ con valores en $(E, \mathcal{E})$ y $(F, \mathcal{F})$ respectivamente. Para que X e Y sean independientes es necesario y suficiente que se satisfaga alguna de las siguientes proposiciones*

(a) $f(X)$ y $g(Y)$ son v.a.i. para cualquier par de funciones medibles $f,\ g$.

(b) $\mathrm{E}[f(X)g(Y)] = \mathrm{E}[f(X)]\,\mathrm{E}[g(Y)]$ para cualquier par de funciones $f,\ g$ medibles y acotadas o medibles y positivas.

(c) Si E y F son espacios métricos con σ-álgebras de Borel $\mathcal{E}$ y $\mathcal{F}$, entonces $\mathrm{E}[f(X)g(Y)] = \mathrm{E}[f(X)]\,\mathrm{E}[g(Y)]$ para cualquier par de funciones $f,\ g$ continuas y acotadas.

Demostración. Sabemos que X, Y son v.a.i. si

(d) $P(X \in A, Y \in B) = P(X \in A)P(Y \in B)$ para todo $A \in \mathcal{E},\ B \in \mathcal{F}$.

$(a) \Rightarrow (d)$ Basta tomar $f(x) = x,\ g(y) = y$.

$(d) \Rightarrow (a)$ Dadas f, g observamos que

$$\big(f(X)\big)^{-1}(\mathcal{E}) = X^{-1}(f^{-1}(\mathcal{E})) \subset X^{-1}(\mathcal{E})$$

y de manera similar $\big(g(Y)\big)^{-1}(\mathcal{F}) \subset Y^{-1}(\mathcal{F})$ y como $X^{-1}(\mathcal{E})$ y $Y^{-1}(\mathcal{F})$ son independientes, las dos sub-σ-álgebras también lo son.

$(b) \Rightarrow (d)$ Basta tomar $f(x) = \mathbf{1}_A(x),\ g(y) = \mathbf{1}_B(y),\ A \in \mathcal{E},\ B \in \mathcal{F}$.

$(d) \Rightarrow (b)$ Sabemos que (b) vale para funciones indicadoras y por linealidad vale para funciones simples. Si f y g son positivas sean (f_n) y $g_n)$ sucesiones crecientes de funciones simples positivas que convergen a f y g respectivamente. Observamos que $0 \leq f_n(x)g_n(y) \uparrow f(x)g(y)$. Por el TCM

$$\begin{aligned} \mathrm{E}[f(X)g(Y)] &= \mathrm{E}[\lim_n f_n(X)g_n(Y)] = \lim_n \mathrm{E}[f_n(X)g_n(Y)] \\ &= \lim_n \mathrm{E}[f_n(X)]\,\mathrm{E}[g_n(Y)] = \mathrm{E}[f(X)]\,\mathrm{E}[g(Y)] \end{aligned}$$

Esto demuestra el resultado para funciones positivas. Si f y g son acotadas, las descomponemos $f = f^+ - f^-,\ g = g^+ - g^-$ y usamos linealidad.

$(b) \Rightarrow (c)$ Es evidente.

$(c) \Rightarrow (d)$ Basta probar (d) para $A \in \mathcal{C},\ B \in \mathcal{D}$ donde $\mathcal{C}$ y $\mathcal{D}$ son las clases de todos los conjuntos cerrados de E, F, respectivamente, que son sistemas π que generan a $\mathcal{E}$ y $\mathcal{F}$.

Sea $A \subset E$ un conjunto cerrado y sea $d(x, A)$ la distancia entre el punto x y A. Entonces $f_n(x) = \min\{1, nd(x, A)\}$ es continua, satisface $0 \leq f_n(x) \leq 1$ y la sucesión $(1 - f_n)$ decrece a la función indicadora $\mathbf{1}_A$. De manera similar a $B \subset F$ cerrado le asociamos una sucesión de funciones continuas g_n que decrecen a $\mathbf{1}_B$ con $0 \leq g_n \leq 1$. Repetimos ahora el argumento de $(d) \Rightarrow (b)$ reemplazando el TCM por el TCD. ■

Ejemplo 1.12
Sean E y F finitos o numerables. Para (X, Y) sea

$$\begin{aligned} P_{X,Y}(i,j) &= P(X = i, Y = j) = P(\{\omega : X(\omega) = i \text{ y } Y(\omega) = j\}) \\ &= P(\{X = i\} \cap \{Y = j\}). \end{aligned}$$

Entonces X e Y son independientes si y sólo si $P_{X,Y}(i,j) = P_X(i)P_Y(j)$.

Proposición 1.3 *Sea X, Y dos variables aleatorias sobre $(\Omega, \mathcal{A}, P)$ con valores en $(E, \mathcal{E})$ y $(F, \mathcal{F})$ respectivamente. El par $Z = (X, Y)$ es una variable aleatoria con valores en $(E \times F, \mathcal{E} \otimes \mathcal{F})$. Las variables X, Y son independientes si y sólo si la distribución $P_{X,Y}$ del par (X, Y) es igual al producto $P_X \times P_Y$ de las distribuciones de X e Y.*

Demostración. Si $A \times B \in \mathcal{E} \times \mathcal{F}$ tenemos $Z^{-1}(A \times B) = X^{-1}(A) \cap Y^{-1}(B) \in \mathcal{A}$ y como $\mathcal{E} \otimes \mathcal{F} = \sigma(\mathcal{E} \times \mathcal{F})$ esto implica que Z es medible.

X e Y son independientes si y sólo si para cualesquiera $A \in \mathcal{E},\ B \in \mathcal{F}$ tenemos

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

o equivalentemente

$$P_{X,Y}(A \times B) = P_X(A)P_Y(B).$$

Esto equivale a decir que $P_{X,Y}(A \times B) = P_X \times P_Y(A \times B)$ para todos los conjuntos $A \times B \in \mathcal{E} \times \mathcal{F}$ y por unicidad esto equivale a $P_{X,Y} = P_X \times P_Y$ en $\mathcal{E} \otimes \mathcal{F}$. ∎

1.10. Convolución

Consideremos el espacio producto $(\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, P_1 \times P_2)$ y supongamos que (X_1, X_2) es una v.a. bidimensional cuyas f.d. marginales son F_1 y F_2, respectivamente. La fórmula para la convolución nos da la distribución de $X_1 + X_2$.

Teorema 1.15 *Bajo las condiciones anteriores,*

$$F_{X_1+X_2}(u) = \int_{-\infty}^{\infty} F_1(u - y)\, dF_2(y).$$

Si, además, X_2 es absolutamente continua con densidad f_2, entonces

$$F_{X_1+X_2}(u) = \int_{-\infty}^{\infty} F_1(u - y) f_2(y)\, dy.$$

Si X_1 es absolutamente continua con densidad f_1, la densidad de la suma es igual a

$$f_{X_1+X_2}(u) = \int_{-\infty}^{\infty} f_1(u - y)\, dF_2(y).$$

Si ambas son absolutamente continuas, entonces

$$f_{X_1+X_2}(u) = \int_{-\infty}^{\infty} f_1(u - y) f_2(y)\, dy.$$

Demostración.

$$\begin{aligned} F_{X_1+X_2}(u) = P(X_1 + X_2 \le u) &= \iint_{x+y\le u} d(F_1 \times F_2)(x, y) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{u-y} d(F_1 \times F_2)(x, y) = \int_{-\infty}^{\infty} \Big(\int_{-\infty}^{u-y} dF_1(x) \Big) dF_2(y) \\ &= \int_{-\infty}^{\infty} F_1(u - y)\, dF_2(y). \end{aligned}$$

El resto sigue de esta relación. ∎

1.11. Sucesiones de Variables Aleatorias Independientes

Para $n \in \mathbb{N}$ sea X_n una v.a. sobre $(\Omega_n, \mathcal{F}_n, P_n)$. Definimos

$$\Omega = \prod_{n=1}^{\infty} \Omega_n, \qquad \mathcal{F} = \bigotimes_{n=1}^{\infty} \mathcal{F}_n$$

donde $\bigotimes \mathcal{F}_n$ denota la σ-álgebra de Ω generada por los conjuntos de la forma

$$A_1 \times A_2 \times \cdots \times A_k \times \Omega_{k+1} \times \Omega_{k+2} \times \cdots \qquad A_i \in \mathcal{F}_i, 1 \leq i \leq k,\ k \geq 1,$$

es decir, $\mathcal{F}$ es la σ-álgebra generada por productos cartesianos finitos de conjuntos de las σ-álgebras coordenadas, que se conocen como *cilindros*. Es fácil ver que los cilindros forman una semiálgebra.

Teorema 1.16 *Dados los espacios de probabilidad* $(\Omega_n, \mathcal{F}_n, P_n)$, $n \geq 1$ *y* $\Omega = \prod_1^\infty \Omega_n$, $\mathcal{F} = \bigotimes_1^\infty \mathcal{F}_n$, *existe una única probabilidad* P *en* $(\Omega, \mathcal{F})$ *tal que*

$$P(A_1 \times A_2 \times \cdots \times A_k \times \Omega_{k+1} \times \Omega_{k+2} \times \cdots) = \prod_{i=1}^{k} P_i(A_i)$$

para $k \geq 1$ *y* $A_i \in \mathcal{F}_i$.

Demostración.

La función P definida sobre los cilindros es positiva y satisface $P(\Omega) = 1$. Para poder usar el teorema de extensión basta demostrar que esta función es σ-aditiva y por el teorema 2.3 basta ver que P es continua. Supongamos entonces que existe una sucesión de cilindros $(C_n,\ n \geq 1)$ tales que para cada n tenemos $C_{n+1} \subset C_n$ y $P(C_n) \geq \delta > 0$. Probaremos que $\cap_1^\infty C_n \neq \emptyset$.

Observamos que todo cilindro E está determinado por un número finito de coordenadas en el sentido de que existe un entero k finito, que sólo depende de E, tal que si $\omega = (\omega_1, \omega_2, \dots)$ y $\omega' = (\omega'_1, \omega'_2, \dots)$ son dos puntos de Ω con $\omega_j = \omega'_j$ para $1 \leq j \leq k$, entonces o bien ambos puntos están en E o ninguno de los dos está en E. Para simplificar la notación y sin pérdida de generalidad, podemos suponer que para cada n, el conjunto C_n está determinado por las primeras n coordenadas.

Introducimos ahora la siguiente notación: Dado $\omega_1^0 \in \Omega_1$ y un subconjunto cualquiera E de Ω, $(E|\omega_1^0)$ denota al conjunto $\Omega_1 \times E_1$, donde E_1 es el conjunto de puntos $(\omega_2, \omega_3, \dots) \in \prod_{n=2}^\infty \Omega_n$ tales que $(\omega_1^0, \omega_2, \omega_3, \dots) \in E$. Si ω_1^0 no es la primera coordenada de ningún punto de E, entonces $(E|\omega_1^0) = \emptyset$. Observamos que si E es un cilindro entonces $(E|\omega_1^0)$ es un cilindro para todo ω_1^0.

Queremos ver que existe ω_1^0 tal que para todo n se tiene que $P((C_n|\omega_1^0)) \geq \delta/2$. Para esto comenzamos con la ecuación

$$P(C_n) = \int_{\Omega_1} P((C_n|\omega_1))\, dP_1(\omega_1) \tag{1.19}$$

que sigue de la definición de un cilindro. Ponemos ahora

$$B_n = \{\omega_1 : P((C_n|\omega_1)) \geq \delta/2\} \subset \Omega_1.$$

Usando (1.19) tenemos que

$$\delta \leq P(C_n) \leq \int_{B_n} 1\, dP_1(\omega_1) + \int_{B_n^c} \frac{\delta}{2}\, dP_1(\omega_1)$$

y por lo tanto $P(B_n) \geq \delta/2$ para todo $n \geq 1$. Como B_n es decreciente con C_n tenemos $P_1(\cap_1^\infty B_n) \geq \delta/2$. Escogemos ω_1^0 en $\cap_1^\infty B_n$, entonces $P((C_n|\omega_1^0)) \geq \delta/2$. Repitiendo el mismo argumento para el conjunto $(C_n|\omega_1^0)$ vemos que existe un ω_2^0 tal que para todo n, $P((C_n|\omega_1^0, \omega_2^0)) \geq \delta/4$, donde $(C_n|\omega_1^0, \omega_2^0) = ((C_n|\omega_1^0)|\omega_2^0)$ es de la forma $\Omega_1 \times \Omega_2 \times E_3$ y E_3 es el conjunto $(\omega_3, \omega_4, \dots)$ en $\prod_{n=3}^\infty \Omega_n$ tal que $(\omega_1^0, \omega_2^0, \omega_3,$

$\omega_4, \dots) \in C_n$. El argumento continua de esta manera por inducción. Por lo tanto, para cada $k \geq 1$ tenemos ω_k^0 tal que para todo n se tiene

$$P((C_n|\omega_1^0, \dots, \omega_k^0)) \geq \frac{\delta}{2^k}.$$

Consideramos ahora el punto $\omega^0 = (\omega_1^0, \omega_2^0, \dots, \omega_n^0, \dots)$. Como $(C_k|\omega_1^0, \dots, \omega_k^0) \neq \emptyset$, hay un punto en C_k cuyas primeras k coordenadas son las mismas que las de ω^0. Como C_k está determinado por sus primeras k coordenadas, se tiene que $\omega^0 \in C_k$. Esto es cierto para todo k y en consecuencia $\omega^0 \in \cap_1^\infty C_k$, como queríamos demostrar. ■

Si X_n es una v.a. definida en $(\Omega_n, \mathcal{F}_n, P_n)$ llamemos $\widetilde{X}_n$ a su extensión natural a Ω definida como sigue: para $\omega \in \Omega$ sea $\omega = (\omega_1, \omega_2, \dots, \omega_n, \dots)$ con $\omega_i \in \Omega_i$ para cada i. Entonces

$$\widetilde{X}_n(\omega) = X_n(\omega_n)$$

Corolario 1.8 *Sea X_n una v.a. definida en $(\Omega_n, \mathcal{F}_n, P_n)$, $n \geq 1$ y sea $\widetilde{X}_n$ su extensión natural a $(\Omega, \mathcal{F}, P)$. Entonces las variables $(\widetilde{X}_n)_{n\geq 1}$ son independientes y la ley de $\widetilde{X}_n$ en $(\Omega, \mathcal{F}, P)$ es idéntica a la ley de X_n en $(\Omega_n, \mathcal{F}_n, P_n)$.*

Demostración. Tenemos

$$\widetilde{X}^{-1}(B_n) = \Omega_1 \times \cdots \times \Omega_{n-1} \times X_n^{-1}(B_n) \times \Omega_{n+1} \times \cdots$$

y por el teorema 1.16

$$\begin{aligned} P(\cap_{n=1}^k X_n^{-1}(B_n)) &= P(X_1^{-1}(B_1) \times \cdots \times X_k^{-1}(B_k) \times \Omega_{k+1} \times \cdots) \\ &= \prod_{n=1}^k P_n(X_n \in B_n). \end{aligned}$$

■

1.12. Probabilidades en $\mathbb{R}^n$

Denotamos por $\mathcal{B}^n$ a los conjuntos de Borel en $\mathbb{R}^n$. Es sencillo ver que $\mathcal{B}^n$ es generada por los cuadrantes de la forma

$$\prod_{i=1}^n (-\infty, a_i], \qquad a_i \in \mathbb{Q}$$

o por los rectángulos abiertos

$$\prod_{i=1}^n (a_i, b_i), \qquad a_i, b_i \in \mathbb{Q}$$

o por los conjuntos

$$\prod_{i=1}^n (a_i, b_i], \qquad a_i, b_i \in \mathbb{Q}.$$

También tenemos que $\mathcal{B}^n = \mathcal{B} \otimes \mathcal{B} \otimes \cdots \otimes \mathcal{B}$.

La función de distribución n-dimensional de una medida de probabilidad en $(R^n, \mathcal{B}^n)$ se define como

$$F(x_1, \dots, x_n) = P(\prod_{i=1}^n (-\infty, x_i]).$$

Definición 1.12 La medida de Lebesgue λ_n en $(\mathbb{R}^n, \mathcal{B}^n)$ está definida sobre los rectángulos medibles $A_1 \times A_2 \times \cdots \times A_n$ por

$$\lambda_n(\prod_1^n A_i) = \prod_1^n \lambda(A_i), \qquad \forall A_i \in \mathcal{B}.$$

Es posible, a partir de esta fórmula, extender λ_n de manera única a una medida que está caracterizada por la siguiente condición:

$$\lambda_n(\prod_{i=1}^n (a_i, b_i]) = \prod_{i=1}^n (b_i - a_i) \qquad \forall a_i < b_i \in \mathbb{R}.$$

Escribimos

$$\int f(x)\, d\lambda_n(x) = \int f(x)\, dx = \int f(x_1, \dots, x_n)\, dx_1 \cdots dx_n$$

para denotar la integral de f respecto a λ_n y también $\int_A f(x)\, dx$ para la integral del producto $f\mathbf{1}_A$ cuando $A \in \mathcal{B}^n$

Definición 1.13 Una probabilidad P en $(\mathbb{R}^n, \mathcal{B}^n)$ tiene densidad f si $f \geq 0$ es medible en $\mathbb{R}^n$ y

$$P(A) = \int_A f(x)\, dx = \int_A f(x)\mathbf{1}_A(x)\, dx = \int f(x_1, \dots, x_n)\mathbf{1}_A(x_1, \dots, x_n)\, dx_1, \dots dx_n$$

para todo $A \in \mathcal{B}^n$.

Teorema 1.17 *Una función positiva medible según Borel $f : \mathbb{R}^n \to \mathbb{R}$ es la densidad de una medida de probabilidad en $\mathbb{R}^n, \mathcal{B}^n)$ si y sólo si satisface*

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

En este caso f determina totalmente la medida de probabilidad y cualquier otra función medible f' tal que $\lambda_n(f \neq f') = 0$ también es una densidad para la misma medida de probabilidad.

Recíprocamente, una medida de probabilidad en $(\mathbb{R}^n, \mathcal{B}^n)$ absolutamente continua determina su densidad salvo por un conjunto de medida de Lebesgue 0, es decir, si f y f' son dos densidades para esta probabilidad, entonces $\lambda_n(f \neq f') = 0$.

Demostración. Sea f la densidad de la probabilidad P, entonces

$$\int_{-\infty}^x f(y)\, dy = P((-\infty, x]).$$

Haciendo $x \to \infty$ vemos que

$$\int f(y)\, dy = \int_{\mathbb{R}^n} f(y)\, dy = \lim_{x\to\infty} \int_{-\infty}^x f(y)\, dy = 1.$$

Para ver el recíproco sea $f \geq 0$ una función Borel-medible con $\int f(x)\, dx = 1$. Para todo $A \in \mathcal{B}$ ponemos

$$P(A) = \int_A f(y)\, dy = \int \mathbf{1}_A(y) f(y)\, dy, \tag{1.20}$$

esto define una función $P : \mathcal{B}^n \to \mathbb{R}^+$ con $P(\mathbb{R}^n) = 1$. Veamos que es σ-aditiva: Sea $\{A_i, i \geq 1\}$ una sucesión de eventos disjuntos 2 a 2, entonces

$$\begin{aligned} P(\cup_1^\infty A_i) &= \int \mathbf{1}_{\cup_1^\infty A_i} f(x)\, dx = \int \sum_1^\infty \mathbf{1}_{A_i} f(x)\, dx \\ &= \sum_1^\infty \int \mathbf{1}_{A_i} f(x)\, dx = \sum_1^\infty P(A_i) \end{aligned}$$

por el TCM para series.

Veamos ahora que P determina a f salvo por un conjunto de medida de Lebesgue 0. Supongamos que f' es otra densidad para P, entonces f' también satisface (1.20). Sea $\varepsilon > 0$ y $A = \{x : f(x) + \varepsilon \leq f'(x)\}$. Si $\lambda_n(A) > 0$ entonces

$$P(A) + \varepsilon\lambda_n(A) = \int (f(x) + \varepsilon)\mathbf{1}_A(x)\,dx \leq \int f'(x)\mathbf{1}_A(x)\,dx = P(A)$$

lo que es una contradicción, y concluimos que $\lambda_n(\{x : f(x) + \varepsilon \leq f'(x)\}) = 0$. Como

$$\{f + \varepsilon \leq f'\} \uparrow \{f < f'\} \qquad (\varepsilon \downarrow 0)$$

obtenemos que $\lambda_n(\{x : f(x) < f'(x)\}) = 0$. De manera similar se demuestra que $\lambda_n(\{x : f(x) > f'(x)\}) = 0$ y por lo tanto $f' = f$ c.s. ∎

Par simplificar, restringimos nuestra discusión al caso $n = 2$ pero los resultados se extienden fácilmente a $n \geq 3$. Sea X un vector aleatorio en $\mathbb{R}^2$ con componentes $X = (Y, Z)$.

Teorema 1.18 *Supongamos que $X = (Y, Z)$ tiene densidad f en $\mathbb{R}^2$, entonces*

a) Tanto Y como Z tienen densidades en $(\mathbb{R}, \mathcal{B})$ dadas por

$$f_Y(y) = \int_{\mathbb{R}} f(y,z)\,dz, \qquad f_Z(z) = \int_{\mathbb{R}} f(y,z)\,dy.$$

b) Y y Z son independientes si y sólo si $f(y,z) = f_Y(y)f_Z(z)$ c.s.

c) La siguiente fórmula define otra densidad en $\mathbb{R}$ en todo punto $y \in \mathbb{R}$ para el cual $f_Y(y) \neq 0$:

$$f_{Y=y}(z) = \frac{f(y,z)}{f_Y(y)}.$$

Esta función es la densidad condicional de Z dado $Y = y$.

Demostración. a) Para todo $A \in \mathcal{B}$ tenemos

$$\begin{aligned} P(Y \in A) = P(X \in A \times \mathbb{R}) &= \iint_{A\times\mathbb{R}} f(y,z)\,dydz \\ &= \int_A \Big(\int_{\mathbb{R}} f(y,z)\,dz\Big)dy = \int_A f_Y(y)\,dy \end{aligned}$$

y como esto es válido para todo $A \in \mathcal{B}$ y las densidades en $\mathbb{R}$ están caracterizadas por (1.20), $f_Y(y)$ es una densidad para Y.

b) Supongamos que $f(y,z) = f_Y(y)f_Z(z)$, entonces

$$\begin{aligned} P(Y \in A, Z \in B) &= \iint \mathbf{1}_{A\times B}(y,z)f(y,z)\,dydz = \iint \mathbf{1}_A(y)\mathbf{1}_B(z)f(y,z)\,dydz \\ &= \iint \mathbf{1}_A(y)\mathbf{1}_B(z)f_Y(y)f_Z(z)\,dydz = \int \mathbf{1}_A(y)f_Y(y)\,dy \int \mathbf{1}_B(z)f_Z(z)\,dz \\ &= P(Y \in A)P(Z \in B), \end{aligned}$$

y como A y B son borelianos cualesquiera, Y y Z son independientes.

Supongamos ahora que Y y Z son independientes. Sea

$$\mathcal{H} = \Big\{C \in \mathcal{B}^2 : \iint_C f(y,z)\,dydz = \iint_C f_Y(y)f_Z(z)\,dydz\Big\}.$$

Como Y, Z son independientes, si $C = A \times B$ con $A, B \in \mathcal{B}$ entonces

$$P((Y,Z) \in C) = \iint_C f(y,z)\,dydz$$

mientras que

$$\begin{aligned} P((Y,Z) \in C) &= P(Y \in A, Z \in B) = P(Y \in A)P(Z \in B) \\ &= \int_A f_Y(y)\,dy \int_B f_Z(z)\,dz = \iint_{A\times B} f_Y(y) f_Z(z)\,dydz \end{aligned}$$

por el teorema de Fubini. Por lo tanto $\mathcal{H}$ contiene a la clase de los rectángulos medibles $\mathcal{B} \times \mathcal{B}$, que es un sistema π que genera a $\mathcal{B}^2$. Es posible demostrar que $\mathcal{H}$ es un sistema λ y deducimos que $\mathcal{H} = \mathcal{B}^2$. Por lo tanto

$$P(X \in C) = \int_C f(y,z)\,dydz = \int_C f_Y(y) f_Z(z)\,dydz$$

para todo $C \in \mathcal{B}^2$. Por la unicidad de la densidad tenemos $f(y,z) = f_Y(y) f_Z(z)$ c.s.

c) Tenemos

$$\int f_{Y=y}(z)\,dz = \int_{-\infty}^{\infty} \frac{f(y,z)}{f_Y(y)} dz = \frac{1}{f_Y(y)} \int_{-\infty}^{\infty} f/y,z)\,dz = \frac{1}{f_Y(y)} = f_Y(y) = 1.$$

Como $f_{Y=y}(z)$ es positiva, Borel-medible y su integral vale 1, es una densidad. ■

Teorema 1.19 *Sea $X = (X_1, \dots, X_n)$ un vector con densidad f y sea $g : \mathbb{R}^n \to \mathbb{R}^n$ una función inyectiva y continuamente diferenciable cuyo Jacobiano no se anula. Entonces $Y = g(X)$ tiene densidad*

$$f_Y(y) = \begin{cases} f_X(g^{-1}(y))|\det J_{g^{-1}}(y)1| & \textit{si } y \textit{ está en el recorrido de } g \\ 0 & \textit{en otro caso.} \end{cases}$$