EYP1027 Modelos Probabilísticos

Clase 2

Profesor: Reinaldo B. Arellano-Valle

Departamento de Estadística
Pontificia Universidad Católica de Chile

Segundo Semestre 2020

Contenido I

1. Modelo de Probabilidad
   - Espacio Muestral
   - Eventos
   - Medidad de Probabilidad

Modelo Probabilidad: Un modelo o espacio de probabilidad es una **representación matemática** de una situación con resultado **incierto** (fenómeno aleatorio).

Tales situaciones son representadas o estudiadas mediantes **experimentos aleatorios**.

Definición 1.1

Experimento Aleatorio: Un experimento se dice aleatorio si sus resultados no pueden determinarse de antemano.

Se supone **conocido** al conjunto de todos los posibles resultados de un experimento aleatorio, pero no se tiene certeza acerca de cual de los posibles resultados ocurrirá al efectuar el experimento.

Definición 1.2

Espacio Muestral: Al conjunto Ω de todos los posibles resultados de un experimento aleatorio se le denomina espacio muestral.

Cada elemento $\omega \in \Omega$ representa un posible resultado del experimento, y se denomina **punto muestral** o **suceso elemental**.

El espacio muestral Ω es el **primer componente** de un modelo de probabilidad, y su especificación puede ser tan abstracta como se requiera; la única condición es que sea un conjunto no vacío.

Ejemplo 1.1

Experimento: lanzar una moneda al aire. Los posibles resultados en este caso son cara c y sello s, es decir,

$$\Omega = \{c, s\}$$

Ejemplo 1.2

Experimento: tirar un dado ordinario tres veces consecutivas. En este caso los posibles resultados son tripletas de la forma (a, b, c) con $a, b, c \in \{1, 2, 3, 4, 5, 6\}$, es decir,

$$\Omega = \{(a, b, c) : a, b, c \in \{1, 2, 3, 4, 5, 6\}\}$$

Ejemplo 1.3

Experimento: Los artículos que salen de una línea de producción son marcados como defectuosos (D) o aceptables (A). Se observan los artículos y se anota su estado. Esto continúa hasta que se produzcan dos defectos consecutivos o se hayan verificado cuatro artículos, lo que ocurra primero; en este caso,

$$\Omega = \{DD, ADD, ADAD, AADD, AADA, AAAA \\ AAAD, ADAA, DAAA, DADA, DAAD, DADD\}$$

Ejemplo 1.4

Experimento: Observar el número de llamadas en curso en una central telefónica; en este caso,

$$\Omega = \{0, 1, 2, \cdots\}$$

Ejemplo 1.5

Experimento: Escoger un punto *al azar* dentro del círculo unitario; en este caso,

$$\Omega = \{(x, y) : x^2 + y^2 \leq 1\}.$$

A pesar de la simplicidad de los ejemplos anteriores, se desprende a partir de ellos que un espacio muestral puede ser de diferente naturaleza.

Definición 1.3

Espacio Muestral Discreto: Un espacio muestral Ω se dice **discreto** si es finito o infinito contable (o numerable).

Espacio Muestral Continuo: Un espacio muestral Ω se dice **continuo** si es infinito no contable.

Esta distinción es esencial para establecer la medida de probabilidad que asociaremos a nuestro modelo de probabilidad.

En los Ejemplos 1.1-1.4, el espacio muestral Ω es discreto; mientras que en el Ejemplo 1.5, el espacio muestral Ω es continuo.

Nota: Situaciones más complejas donde Ω comparta ambas características también pueden presentarse (al menos en teoría).

Informalmente, cualquier subconjunto A de Ω al cual podemos atribuirle una probabilidad ("medir") se llama **evento** (o evento aleatorio).

Al realizar el experimento, decimos que un evento A **occurre** ssi el resultado del experimento es algún $\omega \in A$.

Así, el objetivo se reduce cuantificar la chance de ocurrencia del evento A.

Sería deseable poder atribuir una probabilidad a cualquier subconjunto A de un espacio muestral Ω.

Esto puede hacerse cuando Ω es discreto, pero en general (al menos en teoría) no es posible "medir-a cualquier subconjuto de Ω.

Modelo de Probabilidad

Eventos

La colección de subconjuntos de Ω sobre la cual se define la chance de ocurrencia es una $\sigma-$algebra (o $\sigma-$álgebra), y constituye el segundo componente formal de un modelo de probabilidad.

Definición 1.4

$\sigma-$**Algebra:** Sea $\Omega \neq \emptyset$. Una colección $\mathcal{A}$ de subconjuntos de Ω se llama $\sigma-$algebra si:

A1) $\Omega \in \mathcal{A}$,

A2) Si $A \in \mathcal{A}$, entonces $A^c \in \mathcal{A}$,

A3) Si $A_1, A_2, \dots \in \mathcal{A}$, entonces $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Eventos: Los elementos de $\mathcal{A}$ se llaman **eventos** (subconjuntos "medibles" de Ω).

Ejemplo 1.6

Sea $\Omega \neq \emptyset$. Entonces:

1) $\mathcal{a}_0 = \{\emptyset, \Omega\}$ es conocida como la $\sigma-$algebra trivial (la menor $\sigma-$algebra posible sobre Ω)
3) $\mathcal{a}_1 = \{\emptyset, A, A^c, \Omega\}$ es una $\sigma-$algebra para todo $A \subset \Omega$
2) $\mathcal{P}(\Omega) := \{A : A \subseteq \Omega\}$ (conjunto potencia) es conocida como $\sigma-$algebra total sobre Ω (la mayor $\sigma-$algebra posible sobre Ω).

Ejemplo 1.7

Experimento: Lanzar una moneda (corriente). Aquí $\Omega = \{c, s\}$. Luego, $\mathcal{a}_0 = \{\emptyset, \Omega\}$ es la $\sigma-$algebra trivial sobre Ω, y $\mathcal{P}(\Omega) = \{\emptyset, \{c\}, \{s\}, \Omega\}$ es la $\sigma-$algebra total sobre Ω; mientras que $\mathcal{A} = \{\emptyset, \{C\}\}$ no es una $\sigma-$algebra de subconjutos de Ω.

Ejemplo 1.8

Experimento: Lanzar dos monedas (corrientes). En este caso, $\Omega = \{cc, cs, sc, ss\}$. Entonces, $\mathcal{a} = \{\emptyset, \{cc, cs\}, \{sc, ss\}, \Omega\}$ es una $\sigma-$algebra sobre Ω.

Tarea: Describa la $\sigma-$algebra total $(\mathcal{P}(\Omega))$ en este ejemplo.

Ejemplo 1.9

Experimento: Lanzar un dado ordinario tres veces consecutivas. Entonces, $\Omega = \{(a, b, c) : a, b, c \in \{1, 2, 3, 4, 5, 6\}\}$, de modo que la colección

$$\mathcal{F} = \{\emptyset, \{(1, 2, 3)\}, \Omega \backslash \{(1, 2, 3)\}, \Omega\}$$

es una $\sigma-$algebra sobre Ω; mientras que la coloección

$$\mathcal{G} = \{(1, 2, 3), (1, 1, 1)\}$$

no es una $\sigma-$algebra sobre Ω.

Teorema 1.1

Sea $\Omega \neq \emptyset$. Si $\mathcal{A}_1$, $\mathcal{A}_2$, $\cdots$ son $\sigma-$algebras sobre Ω, entonces $\bigcap_{i=1}^{\infty} \mathcal{A}_i$ también es una $\sigma-$algebra sobre Ω.

Demostración 1.1

A1) $\Omega \in \bigcap_j \mathcal{A}_j$ ya que $\Omega \in \mathcal{A}_j\ \forall\, j$. A2) Sea $A \in \bigcap_j \mathcal{A}_j$. Entonces, $A \in \mathcal{A}_j$ para cada j; luego $A^c \in \mathcal{A}_j\ \forall\, j$. Por lo tanto $A^c \in \bigcap_j \mathcal{A}_j$. A3) Sean $A_1, A_2, \ldots \in \bigcap_j \mathcal{A}_j$. Entonces, $A_1, A_2, \ldots \in \mathcal{A}_j\ \forall\, j$, de modo que $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}_j\ \forall\, j$. Así, concluimos que $\bigcup_{i=1}^{\infty} A_i \in \bigcap_j \mathcal{A}_j$.

Nota: En general, la unión de $\sigma-$algebras sobre Ω no es una $\sigma-$algebra sobre Ω. Por ejemplo, si $\Omega = \{1, 2, 3\}$, entonces $\mathcal{A}_1$=$\{\emptyset, \{1\}, \{2, 3\}, \Omega\}$ y $\mathcal{A}_2$=$\{\emptyset, \{1, 2\}, \{3\}, \Omega\}$ son claramente $\sigma-$algebras sobre Ω, pero $\mathcal{A}_1 \cup \mathcal{A}_2 = \{\emptyset, \{1\}, \{1, 2\}, \{3\}, \{2, 3\}, \Omega\}$ no es una $\sigma-$algebra sobre Ω (Tarea!).

Definición 1.5

$\sigma-$**Algebra Generada:** Sea $\Omega \neq \emptyset$. Sea $\mathcal{A}$ una colección de subconjuntos de Ω, y sea

$\mathcal{M} = \{a\colon a$ es una σ-algebra sobre Ω que contiene a $\mathcal{A}\}$.

Entonces,

$$\sigma(\mathcal{A}) = \bigcap_{a \in \mathcal{M}} a$$

es la $\sigma-$algebra más pequeña sobre Ω que contiene a $\mathcal{A}$. A $\sigma(\mathcal{A})$ se le llama $\sigma-$algebra generada por $\mathcal{A}$.

Definición 1.6

$\sigma-$**Algebra de Borel:** La $\sigma-$algebra más pequeña sobre $\mathbb{R}$ generada por la clase de intervalos de la forma $(-\infty, a]$, $a \in \mathbb{R}$, se llama $\sigma-$**algebra de Borel**, y se denota por $\mathcal{B}$ (o $\mathcal{B}(\mathbb{R})$). Si $A \in \mathcal{B}$, se dice que A es un **boreliano** (o subconjunto de Borel de $\mathbb{R}$). Ya que $\mathcal{B}$ es una $\sigma-$algebra, si tomamos $a, b \in \mathbb{R}$ con $a < b$, entonces los siguientes subconjuntos de $\mathbb{R}$ son borelianos:

$$(a, \infty) = \mathbb{R} \backslash (-\infty, a]; \quad (a, b] = (-\infty, b] \cap (a, \infty);$$

$$(-\infty, a) = \bigcup_{n=1}^{\infty} \left(-\infty, a - \frac{1}{n}\right];$$

$$[a, \infty) = \mathbb{R} \backslash (-\infty, a); \quad (a, b) = (-\infty, b) \cap (a, \infty);$$

$$[a, b] = \mathbb{R} \backslash ((-\infty, a) \cup (b, \infty)); \quad \{a\} = [a, a]$$

Definición 1.7

$\sigma-$**Algebra de Borel sobre** $\mathbb{R}^n$**:** Sea $a = (a_1, \cdots, a_n)$ y $b = (b_1, \cdots, b_n)$ elementos de $\mathbb{R}^n$ con $a \leq b$, es decir, $a_i \leq b_i$ para todo $i = 1, \cdots, n$. La $\sigma-$algebra, denotada por $\mathcal{B}_n$, generada por todos los intervalos (rectangulos) de la forma

$$(a, b] := \{x = (x_1, \cdots, x_n) \in \mathbb{R}^n : a_i < x_i \leq b_i, i = 1, \cdots, n\}$$

es llamada $\sigma-$algebra de Borel sobre $\mathbb{R}^n$.

Definición 1.8

Espacio Medible: Sea $\Omega \neq \emptyset$ y $\mathcal{A}$ una $\sigma-$algebra sobre Ω. Al par $(\Omega, \mathcal{A})$ se le denomina **espacio medible o espacio de sucesos**.

De la definición de $\sigma-$algebra, queda claro que ambos Ω y $\emptyset$ son eventos ya que pertenecen a cualquier $\sigma-$algebra definida sobre Ω.

$\emptyset$ se llama **evento imposible**, Ω se llama **evento seguro**, y un evento de la forma $\{\omega\}$ con $\omega \in \Omega$ se llama **evento simple** o **elemental**.

Decimos que el evento A de Ω ocurre si al de realizar el experimento aleatorio, el resultado obtenido pertenece al conjunto A.

En este sentido, si A y B son dos eventos de Ω, entonces:

i) El evento $A \cup B$ ocurre si y solo si A o B o ambos ocurren,
ii) El evento $A \cap B$ ocurre si y solo si A y B ocurren,
iii) El evento A^c ocurre si y solo si A no ocurre,
iv) El evento $A \backslash B$ ocurre si y solo si A ocurre pero B no.

Definición 1.9

Eventos mutuamente excluyentes: Dos eventos A y B se dicen mutuamente excluyentes o incompatibles si $A \cap B = \emptyset$.

Ejemplo 1.10

Se lanza una moneda una vez. Sea $A =$ "el resultado obtenido es cara" y $B =$ "el resultado obtenido es sello". Claramente, los eventos A y B son mutuamente excluyentes.

Definición 1.10

Frecuencia relativa: Para cada evento A, el número $fr(A) = \dfrac{n(A)}{n}$ se llama frecuencia relativa de A, donde $n(A)$ indica la cantidad de veces que ocurrió el evento A en n repeticiones del experimento.

Modelo de Probabilidad

Lamentablemente, para cada evento A fijo, $fr(A)$ no es constante: su valor depende de n; sin embargo, se ha observado que cuando un experimento aleatorio se repite en las mismas condiciones un gran número de veces, la frecuencia relativa $fr(A)$ se estabiliza alrededor de un valor específico entre 0 y 1.

Ejemplo 1.11

Suponga que un dado es lanzado n veces. Sea $A =$ "el resultado obtenido es 3". La siguiente tabla resume los valores obtenidos.

n	Frecuencia	Frecuencia relativa
100	14	0.14
200	29	0.145
300	51	0.17
400	65	0.1625
500	83	0.166

Modelo de Probabilidad

La estabilización de la frecuencia relativa se conoce como "regularidad estadística", lo cual nos permite hacer predicciones que eliminan, aunque parcialmente, la incertidumbre presente en fenómenos imprevisibles.

El valor $P(A)$ alrededor del cual se estabiliza $fr(A)$ es un indicador de la "posibilidad" de ocurrencia del evento A.

Interesa, entonces, describir las propiedades que tal número debería tener:
1. Primero, se observa que desde $n(A) \geq 0$, de modo que $P(A)$ debe ser mayor o igual que cero.
2. Ahora, como $n(\Omega) = n$, entonces $fr(\Omega) = 1$, y por lo tanto $P(\Omega) = 1$.
3. Además, si A y B son eventos mutuamente excluyentes, entonces $n(A \cup B) = n(A) + n(B)$; luego $fr(A \cup B) = fr(A) + fr(B)$, de modo que si $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$.

Estas consideraciones nos llevan a establecer la siguiente definición axiomática (de Kolmogorov) para $P(A)$, $A \in \mathcal{A}$.

Definición 1.11

Medida de Probabilidad Sea $(\Omega, \mathcal{A})$ un espacio medible. Una función real valorada P definida sobre $\mathcal{A}$ que satisface los siguientes axiomas:

A1) $P(A) \geq 0$ para todo $A \in \mathcal{A}$ (no-negativa)

A2) $P(\Omega) = 1$ (normalizada)

A3) Si $A_1, A_2, \cdots \in \mathcal{A}$ son eventos (2 a 2) disjuntos $(A_i \cap A_j = \emptyset \, \forall \, i \neq j)$, entonces $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$ (aditividad contable)

se llama **medida de probabilidad** sobre $(\Omega, \mathcal{A})$.

La terna $(\Omega, \mathcal{A}, P)$ se llama **modelo o espacio de probabilidad**, y el número $P(A)$ se denomina **probabilidad** del evento $A \in \mathcal{A}$.

Ejemplo 1.12

Sea $\Omega = \{1, 2, 3\}$, $\mathcal{A} = \{\emptyset, \{1\}, \{2, 3\}, \Omega\}$. Entonces, la función $P(A)$, $A \in \mathcal{A}$, definida por

$$P(A) = \begin{cases} 1 & \text{si } 3 \in A \\ 0 & \text{si } 3 \notin A \end{cases}$$

es una medida de probabilidad sobre $(\Omega, \mathcal{A})$ (Tarea: verificar esta afirmación!)

Ejemplo 1.13

Sean $\Omega = \{0, 1, \cdots\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ y P una función definida sobre $\{i\}$ como

$$P(\{i\}) = (1 - q)q^i, \quad i = 0, 1, \cdots, \quad 0 < q < 1.$$

Ya que los tres axiomas de la definición de probabilidad son satisfechos, P es una medida de probabilidad sobre $(\Omega, \mathcal{A})$. Además, la probabilidad de cada evento $A \in \mathcal{A}$ se calcula como

$$P(A) = \sum_{i:i \in A} P(\{i\}) = (1 - q) \sum_{i:i \in A} q^i.$$

Modelo de Probabilidad

Ejemplo 1.14

Sea $(\Omega, \mathcal{A})$, donde $\Omega = \{1, 2, ..., ...\}$ y $\mathcal{A} = \mathcal{P}(\Omega)$. Considere la función definida por

$$P(\{i\}) = \frac{2}{3}\left(\frac{1}{2}\right)^{i-1}, \quad i = 1, 2, \ldots$$

Entonces, P no define una medida de probabilidad sobre $(\Omega, \mathcal{A})$. Ya dado que $\{1\}, \{2\}, \ldots$ es una partición de Ω, entonces $\Omega = \bigcup_{i=1}^{\infty}\{i\}$, de modo que

$$P(\Omega) = P\left(\bigcup_{i=1}^{\infty}\{i\}\right) = \sum_{i=1}^{\infty} P(\{i\}) = \sum_{i=1}^{\infty} \frac{2}{3}\left(\frac{1}{2}\right)^{i-1} = \frac{2}{3}\sum_{i=1}^{\infty}\left(\frac{1}{2}\right)^{i-1} = \frac{2}{3}\,\frac{1}{1-1/2} = \frac{4}{3} > 1,$$

lo cual viola el axioma A3) de una medidad de probabilidad, lo cual prueba la afimación.

Ejemplo 1.15

Sea $\Omega = (0, \infty)$ y $\mathcal{A} = \mathcal{B}$. Defina la medida de probabilidad P como:

$$P(B) = \int_B e^{-x} dx, \quad B \in \mathcal{B}.$$

Entonces:

A1) Claramente $P(B) \geq 0$, $\forall B \in \mathcal{B}$,

A2) $P(\Omega) = \int_\Omega e^{-x} dx = \int_0^\infty e^{-x} dx = 1$,

A3) Sea $B_1, B_2, \cdots$ es una sucesión de intervalos disjuntos en $\mathcal{B}$; luego

$$P\left(\bigcup_{i=1}^{\infty} B_i\right) = \int_{\bigcup_{i=1} B_i} e^{-x} dx = \sum_{i=1}^{\infty} \int_{B_i} e^{-x} dx = \sum_{i=1}^{\infty} P(B_i).$$

Modelo de Probabilidad

Teorema 1.2

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Entonces:

P1) $P(\emptyset) = 0$

P2) Si $A, B \in \mathcal{A}$ y $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$

P3) Para cualquier $A \in \mathcal{A}$, $P(A^c) = 1 - P(A)$

P4) Si $A \subseteq B$, entonces $P(A) \leq P(B)$ y $P(B \backslash A) = P(B) - P(A)$.
En particular $P(A) \leq 1 \overset{(A1)}{\Longrightarrow} 0 \leq P(A) \leq 1$ para todo $A \in \mathcal{A}$

P5) Para cualquier $A, B \in \mathcal{A}$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

P6a) Si $\{A_n\}_{n=1}^{\infty}$ es una secuencia creciente de eventos en $\mathcal{A}$, es decir, $A_n \in \mathcal{A}$ y $A_n \subseteq A_{n+1}$ para todo $n = 1, 2, \ldots$; entonces $P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$ donde $\lim_{n\to\infty} A_n := \bigcup_{n=1}^{\infty} A_n$

P6b) Si $\{A_n\}_{n=1}^{\infty}$ es una secuencia decreciente de eventos en $\mathcal{A}$, es decir, $A_n \in \mathcal{A}$ y $A_n \supseteq A_{n+1}$ para todo $n = 1, 2, \ldots$; entonces $P\left(\lim_{n\to\infty} A_n\right) = \lim_{n\to\infty} P\left(A_n\right)$ donde $\lim_{n\to\infty} A_n := \bigcap_{n=1}^{\infty} A_n$

Demostración 1.2

P1) $1 = P(\Omega) = P(\Omega \cup \emptyset \cup \emptyset \cup \cdots) = P(\Omega) + P(\emptyset) + P(\emptyset) + \cdots$. Entonces $0 \geq P(\emptyset) \geq 0$ y por lo tanto $P(\emptyset) = 0$.

P2) $A \cup B = A \cup B \cup \emptyset \cup \emptyset \cup \cdots$. Por lo tanto, la prueba sigue del axioma $A3$) de la definición de medida de probabilidad y del resultado anterior.

P3) $P(A) + P(A^c) = P(A \cup A^c) = P(\Omega) = 1$.

P4) $B = A \cup (B \backslash A)$. Obtenemos $P(B) = P(A) + P(B \backslash A) \geq P(A)$ por $A1$).

P5) Ejercicio para el lector.

P6a) Ejercicio para el lector. Defina $C_1 = A_1, C_2 = A_2 \backslash A_1, \cdots, C_n = A_n \backslash A_{n-1}$. Luego use que $\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n$ y $C_i \cap C_j = \emptyset$ para todo $i \neq j$.

P6b) Ejercicio para el lector.

Teorema 1.3

Sea $\Omega = \{\omega_1, \ldots, \omega_{N(\Omega)}\}$ un espacio muestral finito. Sea $\mathcal{A}$ cualquier $\sigma-$algebra sobre Ω. Sean $p_1, \ldots, p_n$ números no negativos que suman 1. Para cualquier $A \in \mathcal{A}$ defina $P(A)$ por

$$P(A) = \sum_{\{i:\omega_i \in A\}} p_i.$$

Entonces P es una medida de probabilidad sobre $\mathcal{A}$.

Esto sigue siendo cierto si $\Omega = \{\omega_1, \omega_2, \ldots\}$ es un conjunto contable y $p_1, p_2, \ldots$ son números no negativos tal que $\sum_{i=1}^{\infty} p_i = 1$.

En tales casos, (Ω, $\mathcal{A}$ P) se llama **modelo de probabilidad discreto**.

Ejemplo 1.16

En el Teorema 1.3, suponga que $\Omega = \left\{\omega_1, \ldots, \omega_{N(\Omega)}\right\}$ es equiprobable, es decir,

$$p_i = P(\{\omega_i\}) = \frac{1}{N(\Omega)}, \quad \forall\, i = 1, \ldots, N(\Omega).$$

Entonces, para cualquier $A \in \mathcal{A}$ con $N(A)$ elementos (casos favorables al evento A) se tiene que

$$P(A) = \frac{N(A)}{N(\Omega)} = \frac{\text{casos favorables}}{\text{caso posibles}}.$$

En tal caso, $(\Omega\ \mathcal{A},\ P)$ es un **modelo de probabilidad (discreto) equiprobabble.**

Modelo de Probabilidad

Ejercicios: Sea $(\Omega, \mathcal{A}, P)$ un modelo de probabilidad. Pruebe que

1) Si $A_1, \ldots, A_n$ son eventos (2 a 2) disjuntos en $\mathcal{A}$, entonces

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$$

2) Para cualquier secuencia de eventos $A_1, \ldots, A_n$ en $\mathcal{A}$, se tiene que

$$P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i A_j) + \sum_{i<j<k} P(A_i A_j A_k) - \cdots (-1)^{n+1} P(A_1 A_2 \cdots A_n),$$

donde $A_i A_j = A_i \cap A_j, \ldots, A_1 A_2 \cdots A_n = \bigcup_{i=1}^{n} A_i$.

3) Para cualquier secuencia contable de eventos $A_1, A_2 \ldots$ en $\mathcal{A}$, se tiene que

$$P(\bigcup_{i=1}^{\infty} A_i) \leq \sum_{i=1}^{\infty} P(A_i).$$

Nota: Para probar el Ejercicio 3, defina $B_1 = A_1$ y $B_k = A_1^c A_2^c \cdots A_{k-1}^c A_k$ para $k = 2, 3, \ldots$. Note que $B_i B_j = \emptyset \; \forall \, i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = \bigcup_{i=1}^{\infty} B_i$. Luego, aplique el axioma $A3)$ y la propiedad $P2)$ de una medida de probabilidad y que $B_k \subset A_k \; \forall \, k$.

References

Blanco, L., Arunachalam, V. y Dharmaraja, S. (2012). *Introduction to Probability and Stochastic Processes.* John Wiley, New Jersey.

Casella, G. y Berger, R.L. (2002). *Statistical Inference.* Second Edition. Duxbury, California.