UNIVERSIDAD NACIONAL DE COLOMBIA

Retardo Cosmológico Temporal en Teorías de Gravedad Modificada $f(R)$

Julián Orlando Jiménez Cárdenas

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Física
Bogotá D. C. , Colombia
2019

Retardo Cosmológico Temporal en Teorías de Gravedad Modificada $f(R)$

Julián Orlando Jiménez Cárdenas

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Físico

Director(a):
Ph.D. Leonardo Castañeda Colorado

Línea de Investigación:
Astrofísica, Gravitación y Cosmologia
Grupo de Investigación:
Grupo de Galaxias, Gravitación y Cosmologia

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de Física
Bogotá D. C. ,Colombia
2019

Resumen

En la primera parte de este texto se presenta una introducción a la geometría diferencial como herramienta matemática para la Relatividad General. Se estudia la gravedad linealizada y el papel que esta desempeña en la radiación gravitacional, profundizando así en los conceptos de gauge, energía y contribución cuadrupolar. Seguidamente se presentan las ecuaciones de Einstein relajadas como una generalización para el estudio de la radiación gravitacional y se obtienen expresiones generales para la energía, el momentum lineal y angular. Posteriormente se muestra la relación entre las expresiones de flujo de energía, momentum lineal y angular con el tensor de Weyl.

Palabras clave: Radiación Gravitacional, Gravedad Linealizada, Ecuaciones de Einstein relajadas, Tensor de Weyl

Abstract

In the first part of this text an introduction to differential geometry is presented as a tool for General Relativity. The linearized gravity is studied and the role that this one plays in the gravitational radiation, deepening in the gauge, energy and quadrupolar contribution concepts. After this the relaxed Einstein equations are presented as a generalization for the study of gravitational radiation and general expression of energy, linear and angular momentum are obtained. Later it is shown the relation between the flux of energy, lineal and angular momentum with the Weyl tensor.

Keywords: Gravitational Radiation, Linearized Gravity, Relaxed Einstein field equations, Weyl tensor

Contenido

1 Relatividad General

1.1. Introducción

Este capítulo es una breve introducción a la teoría de la relatividad general, partiendo desde el concepto de variedad, el espacio tangente y cotangente, el concepto de curvatura y el papel que juega la gravedad en todas estas ideas matemáticas. Las referencias clave de este capítulo son [2],[3].

1.2. Variedades

1.2.1. Mapas

Definición 1.2.1 (Mapa)**.** *Dados dos conjuntos A y B, un mapa $\phi : M \to N$ es una relación que asigna cada elemento $x \in M$ a un único elemento $y \in N$. En este caso, se denota como $\phi(x) = y$.*

Definición 1.2.2 (Composición de Mapas)**.** *Con dos mapas $\phi : A \to B$ y $\Psi : B \to C$, se define la composición de ambos mapas, $\Psi \circ \phi : A \to C$, por su acción sobre los elementos de A:*

$$(\Psi \circ \phi)(a) = \Psi(\phi(a)).$$

$$\begin{tikzcd}
A \arrow[rr, "\Psi\circ\phi"] \arrow[dr, "\phi"'] & & C \\
& B \arrow[ur, "\Psi"'] &
\end{tikzcd}$$

Un mapa $\phi : A \to B$ se dice inyectivo (uno a uno) si cada elemento de B tiene a lo sumo un elemento de A que es mapeado a él. Este mapa se dice sobreyectivo si cada elemento de B tiene al menos un elemento de A mapeado a él. A se conoce como el dominio del mapa ϕ, y su imagen es

$$Im\, \phi := \{y \in B : \exists x \in A \text{ tal que } \phi(x) = y\}.$$

La preimagen de un conjunto $U \subseteq B$ bajo la función ϕ se define como

$$\phi^{-1}(U) := \{x \in A : \exists y \in U \text{ tal que } \phi(x) = y\}.$$

Un mapa $\phi : A \to B$ que es inyectivo y sobreyectivo a la vez se conoce como invertible (biyectivo). En este caso, se define el mapa inverso $\phi^{-1} : B \to A$ de modo que se satisfaga que, para todo $y \in B$ $(\phi \circ \phi^{-1})(y) = y$.

$$\begin{tikzcd} A \arrow[r, bend right, "\phi"'] & B \arrow[l, bend right, "\phi^{-1}"'] \end{tikzcd}$$

Un mapa f de $\mathbb{R}^m$ a $\mathbb{R}^n$ toma una $m-$tupla $(x^1, x^2, \ldots, x^m)$ y la envía a una $n-$tupla $(y^1, y^2, \ldots, y^n)$, de modo que se puede pensar como una colección de n funciones ϕ^i de m variables:

$$y^i = \phi^i(x^1, \ldots, x^m) \text{ con } i = 1, \ldots, n,$$

de modo que

$$f(x^1, \ldots, x^m) = (\phi^1(x^1, \ldots, x^m), \ldots, \phi^n(x^1, \ldots, x^m)).$$

Se referirá a cada una de las funciones ϕ^i como C^p si son continuas y $p-$veces diferenciables, y al mapa entero $f : \mathbb{R}^m \to \mathbb{R}^n$ como C^p si cada uno de los campos escalares $\phi^i, i = 1, \ldots, n$ es al menos C^p.
Un mapa C^0 es continuo pero no necesariamente diferenciable, mientras que un mapa C^∞ es continuo y puede ser diferenciado cuantas veces se desee. Los mapas C^∞ se llaman suaves.

Definición 1.2.3 (Difeomorfismo)**.** *El mapa $\phi : A \to B$ se conoce como difeomorfismo si es biyectivo, y tanto él como su inversa son C^∞. Se dice entonces que los conjuntos A y B son difeomorfos.*

1.2.2. Regla de la cadena

Si tiene dos mapas $f : \mathbb{R}^m \to \mathbb{R}^n$ y $g : \mathbb{R}^n \to \mathbb{R}^l$, que se componen en $(g \circ f) : \mathbb{R}^m \to \mathbb{R}^l$, represente cada espacio en términos de coordenadas: x^a en $\mathbb{R}^m$, y^b en $\mathbb{R}^n$ y z^c en $\mathbb{R}^l$, donde los índices a, b, c varían sobre los valores apropiados.

$$\begin{tikzcd} \mathbb{R}^m \arrow[rr, "g \circ f"] \arrow[dr, "f"] & & \mathbb{R}^l \\ & \mathbb{R}^n \arrow[ur, "g"] & \end{tikzcd}$$

La regla de la cadena relaciona las derivadas parciales de la composición $(g \circ f)$ con las derivadas parciales de los mapas f y g de la siguiente manera

$$\frac{\partial}{\partial x^a}(g \circ f)^c = \sum_{b=1}^{n} \frac{\partial f^b}{\partial x^a} \frac{\partial g^c}{\partial y^b}.$$

1.2.3. Variedades

En el capítulo de estabilidad se trató el concepto de variedad como un espacio métrico homeomorfo localmente a la bola abierta. En este capítulo se tomará una variedad más general: la variedad topológica, para lo cual se introducirá la idea de topología, y demás conceptos necesarios en términos de la topología de la variedad.

Definición 1.2.4 (Espacio topologico)**.** *Tome A como un conjunto arbitrario. τ es una topología para el conjunto A si satisface las siguientes condiciones:*

1. $\emptyset, A \in \tau$,

2. *Si $\{U_\alpha\}_{\alpha\in I} \subset \tau$ es una familia arbitraria de elementos de τ, entonces la unión de toda esta familia pertenece a τ, es decir, $\bigcup_{\alpha\in I} U_\alpha \in \tau$, y*

3. *Si $\{U_n\}_{n=1}^m \subset \tau$ es una familia finita de elementos de τ, entonces la intersección de todos sus elementos también es un elemento de τ, es decir, $\bigcap_{n=1}^m U_n \in \tau$.*

En este caso se dice que la pareja (A, τ) es un espacio topológico. Los elementos de τ se llaman abiertos y sus complementos se llaman cerrados.

Definición 1.2.5 (Carta o sistema coordenado)**.** *Considere un espacio topológico (M, τ). Una carta o sistema coordenado (U, ϕ) consiste de un conjunto abierto $U \subset M$, junto con un mapa inyectivo $\phi : U \to \mathbb{R}^n$, tal que $\phi(U)$ es abierto en $(\mathbb{R}^n, \tau_u)$*[1].

Definición 1.2.6 (Atlas C^r)**.** *Un atlas C^r es una colección indexada de cartas $\{(U_\alpha.\phi_\alpha)\}_{\alpha\in I}$, con ϕ_α siendo al menos C^r, para todo $\alpha \in I$, que satisface las siguientes condiciones*

1. *$\bigcup_{\alpha\in I} U_\alpha = M$, es decir, $\{U_\alpha\}_{\alpha\in I}$ es un cubrimiento abierto para M y*

2. *si para algunos $\alpha, \beta \in I$ $(\alpha \neq \beta)$, $U_\alpha \cap U_\beta \neq \emptyset$, entonces el mapa $(\phi_\alpha \circ \phi_\beta^{-1}) : \phi_\beta(U_\alpha \cap U_\beta) \to \phi_\alpha(U_\alpha \cap U_\beta)$ toma puntos en $\phi_\beta(U_\alpha \cap U_\beta) \subseteq \mathbb{R}^n$ y los envía a puntos en $\phi_\alpha(U_\alpha \cap U_\beta)$, y viceversa. Ambas composiciones deben ser C^r. Si se satisface esta condición se dice que los mapas ϕ_α y ϕ_β son compatibles*

Un atlas se dice maximal si contiene todas las posibles cartas compatibles.

$$\begin{tikzcd}
U_\alpha \cap U_\beta \subset M \arrow[r, "\phi_\alpha"] \arrow[d, "\phi_\beta"] & \phi_\alpha(U_\alpha \cap U_\beta) \subset \mathbb{R}^n \arrow[dl, bend left, "\phi_\beta\circ\phi_\alpha^{-1}"] \\
\phi_\beta(U_\alpha \cap U_\beta) \subset \mathbb{R}^n \arrow[ur, "\phi_\alpha\circ\phi_\beta^{-1}"] &
\end{tikzcd}$$

Definición 1.2.7 (C^r Variedad $n-$dimensional)**.** *Una C^r variedad $n-$dimensional es un espacio topológico (M, τ) junto con un atlas maximal C^r.*

[1] τ_u denota la topología usual sobre $\mathbb{R}^n$.

El hecho de que una variedad sea localmente como $\mathbb{R}^n$ (a través de las cartas) introduce la posibilidad de usar herramientas del cálculo real sobre ella. Tome por ejemplo dos C^∞ variedades (M, τ_M) y (N, τ_N) de dimensión m y n, respectivamente. Por simplicidad, pero sin pérdida de generalidad, tome $\phi : M \to \mathbb{R}^m$ y $\Psi : N \to \mathbb{R}^n$ como las cartas coordenadas de M y N, respectivamente. Si $f : M \to N$ es una función entre ambas variedades,

$$\begin{tikzcd} M \arrow[r, "f"] & N \arrow[d, "\Psi"] \\ \mathbb{R}^m \arrow[u, "\phi^{-1}"] \arrow[r, "\Psi\circ f\circ\phi^{-1}"] & \mathbb{R}^n \end{tikzcd}$$

se puede introducir el concepto de diferenciación sobre el mapa f, construyendo el mapa

$$(\Psi \circ f \circ \phi^{-1}) : \mathbb{R}^m \to \mathbb{R}^n,$$

de modo que el operador $\frac{\partial f}{\partial x^\mu}$ quede definido como

$$\frac{\partial f}{\partial x^\mu} := \frac{\partial}{\partial x^\mu}(\Psi \circ f \circ \phi^{-1}),$$

donde $\mu = 1, \ldots, m$.

1.2.4. Espacio tangente y cotangente

Tome $\mathcal{F}$ como el espacio de todas las funciones suaves $f : M \to \mathbb{R}$ ($\phi^{-1} \circ f$ es de clase C^∞, siendo ϕ la carta coordenada de M). Cada curva $\gamma : \mathbb{R} \to M$ que pasa por algún punto $p \in M$ define un operador sobre el espacio, la derivada direccional, que mapea f a

$$\left.\frac{\mathrm{d}f}{d\lambda}\right|_{\lambda:\gamma(\lambda)=p} := \frac{\mathrm{d}}{d\lambda}(f \circ \gamma)(\lambda)$$

(evaluada en p).

$$\begin{tikzcd} \mathbb{R} \arrow[r, "\gamma"] \arrow[rr, bend right, "f\circ\gamma"'] & M \arrow[r, "f"] & \mathbb{R} \end{tikzcd}$$

Definición 1.2.8 (Espacio tangente). *El espacio tangente T_pM a un punto $p \in M$ es el espacio de los operadores derivadas direccionales dados por todas las curvas que pasan por el punto p. Este espacio resulta ser un espacio vectorial.*

El espacio tangente T_pM posee una base natural, $\{\partial_\mu\}$. Cada uno de estos operadores está definido en términos de la curva generada por la carta coordenada del punto p. Es decir, si (U, ϕ) es una carta coordenada tal que $p \in U$, se toma $(\phi^{-1})^\mu : \mathbb{R} \to M$ como la restricción de la función ϕ^{-1} a una única variable, x^μ, $\mu = 1, \ldots, m$, con el objetivo de que esta nueva función sea una curva sobre M que pase por p, para que defina la derivada direccional ∂_μ. Para ver que efectivamente es una base del espacio tangente T_pM, considere una variedad m-dimensional suave M, una carta coordenada (U, ϕ), una curva $\gamma : \mathbb{R} \to M$ y una función $f : M \to \mathbb{R}$.

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "\gamma"] \arrow[rr, bend left, "f\circ\gamma"] \arrow[dr, "\phi\circ\gamma"'] & U \subseteq M \arrow[r, "f"] \arrow[d, bend right, "\phi"'] & \mathbb{R} \\
& \mathbb{R}^m \arrow[u, bend right, "\phi^{-1}"'] \arrow[ur, "f\circ\phi^{-1}"'] &
\end{tikzcd}$$

Si λ es el parámetro de la curva γ, se expande el operador $\frac{\mathrm{d}}{\mathrm{d}\lambda}$ en términos de los operadores ∂_μ aplicando la regla de la cadena:

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = \frac{\mathrm{d}}{\mathrm{d}\lambda}(f\circ\gamma) = \frac{\mathrm{d}}{\mathrm{d}\lambda}((f\circ\phi^{-1})\circ(\phi\circ\gamma)) = \frac{\mathrm{d}(\phi\circ\gamma)^\mu}{\mathrm{d}\lambda}\frac{\partial(f\circ\phi^{-1})}{\partial x^\mu} = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\partial_\mu f.$$

Como la función f es arbitraria,

$$\frac{\mathrm{d}}{\mathrm{d}\lambda} = \frac{\mathrm{d}x^\mu}{\mathrm{d}\lambda}\partial_\mu,$$

con lo que los operadores derivada direccional $\{\partial_\mu\}$ son una base para T_pM, conocida como base coordenada. Además, esto implica que el espacio tangente T_pM tiene la misma dimensión de la variedad.

Una de las ventajas de este punto de vista de los vectores como operadores diferenciales es que la ley de transformación es inmediata. Como los vectores de la base son $\hat{e}_{(\mu)} = \partial_\mu$, los vectores de la base en un nuevo sistema coordenado $x^{\mu'}$ están dadas por la regla de la cadena [2]

$$\partial_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu$$

La ley de transformación de vectores se introduce de tal forma que un vector del espacio tangente $V = V^\mu\partial_\mu$ permanezca invariante bajo un cambio de base, es decir,

$$V^\mu\partial_\mu = V^{\mu'}\partial_{\mu'} = V^{\mu'}\frac{\partial x^\mu}{\partial x^{\mu'}}\partial_\mu,$$

y como la matriz $\frac{\partial x^{\mu'}}{\partial x^\mu}$ es la inversa de $\frac{\partial x^\mu}{\partial x^{\mu'}}$, la ley de transformación es

$$V^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}V^\mu. \tag{1-1}$$

Definición 1.2.9 (Espacio cotangente)**.** *El espacio cotangente T_p^*M de una variedad M en un punto $p \in M$ es el conjunto de los mapas lineales $\omega : T_pM \to \mathbb{R}$. Los elementos de este espacio se conocen como 1-formas.*

El ejemplo canónico de 1-forma es el gradiente de una función $f : M \to \mathbb{R}$, denotado por $\mathrm{d}f$. Su acción sobre un vector $\frac{\mathrm{d}}{\mathrm{d}\lambda}$ del espacio tangente es exactamente la derivada direccional sobre la función f:

$$\mathrm{d}f\left(\frac{\mathrm{d}}{\mathrm{d}\lambda}\right) = \left.\frac{\mathrm{d}f}{\mathrm{d}\lambda}\right|_p.$$

Justo como las derivadas parciales a lo largo de los ejes coordenados proveen una base natural para el espacio tangente, los gradientes de las funciones coordenadas x^μ proveen una base natural para el espacio cotangente $\{\mathrm{d}x^\mu\}$, conocida como base dual. Observe que, al aplicar $\mathrm{d}x^\mu$ a ∂_η se obtiene que

$$\mathrm{d}x^\mu(\partial_\nu) = \frac{\partial x^\mu}{\partial x^\nu} = \frac{\partial}{\partial x^\nu}(x^\mu \circ (x^\nu)^{-1}) = \frac{\partial}{\partial x^\nu}((x^\mu \circ \phi^{-1}) \circ (\phi \circ (x^\nu)^{-1})).$$

$$\begin{tikzcd}
\mathbb{R} \arrow[r, "(x^\nu)^{-1}"] \arrow[dr, "\phi\circ(x^\nu)^{-1}"'] & M \arrow[d, "\phi"] & \\
 & \mathbb{R}^m \arrow[r, "\phi^{-1}"] \arrow[dr, "x^\mu\circ\phi^{-1}"'] & M \arrow[d, "x^\mu"] \\
 & & \mathbb{R}
\end{tikzcd}$$

Aplicando la regla de la cadena,

$$\frac{\partial}{\partial x^\nu}((x^\mu \circ \phi^{-1}) \circ (\phi \circ (x^\nu)^{-1})) = \frac{\partial(\phi \circ (x^\nu)^{-1})^\eta}{\partial x^\nu}\frac{\partial(x^\mu \circ \phi^{-1})}{\partial x^\eta}.$$

Intuitivamente, $x^\mu \circ \phi^{-1} = x^\mu$, donde x^μ del lado izquierdo de la igualdad es la μ-ésima coordenada de la carta, y al lado derecho es la μ-ésima componente de $\mathbb{R}^m$. Por otro lado, $(\phi \circ (x^\nu)^{-1})^\eta = x^\eta$, con lo que

$$\frac{\partial(\phi \circ (x^\nu)^{-1})^\eta}{\partial x^\nu}\frac{\partial(x^\mu \circ \phi^{-1})}{\partial x^\eta} = \frac{\partial x^\eta}{\partial x^\nu}\frac{\partial x^\mu}{\partial x^\eta} = \delta^\mu_\eta \frac{\partial x^\eta}{\partial x^\nu} = \frac{\partial x^\mu}{\partial x^\nu} = \delta^\mu_\nu.$$

En resumen,

$$\boxed{\mathrm{d}x^\mu(\partial_\nu) = \delta^\mu_\nu.} \tag{1-2}$$

Esta condición determina que $\{\mathrm{d}x^\mu\}$ es una base para el espacio cotangente T_p^*M [2]. De este modo, cualquier 1-forma ω se puede expandir en sus componentes: $\omega = \omega_\mu \mathrm{d}x^\mu$. Las propiedades de transformación de los vectores de la base dual y las componentes de una 1-forma se siguen de la misma forma que en el caso del espacio tangente:

$$\mathrm{d}x^{\mu'} = \frac{\partial x^{\mu'}}{\partial x^\mu}\mathrm{d}x^\mu;\ \omega_{\mu'} = \frac{\partial x^\mu}{\partial x^{\mu'}}\omega_\mu.$$

Definición 1.2.10 (Espacio producto cartesiano)**.** *Se define el espacio producto cartesiano Π^k_l respecto a un punto $p \in M$ de la variedad como:*

$$\Pi^k_l := \underbrace{T_p^*M \times \cdots \times T_p^*M}_{l-veces} \times \underbrace{T_pM \times \cdots \times T_pM}_{k-veces},\ \textit{es decir,}$$

$$\Pi^k_l = \{(\omega^1, \omega^2, \ldots, \omega^l, \mathbb{X}_1, \mathbb{X}_2, \ldots, \mathbb{X}_k) : \omega^i \in T_p^*M;\ \mathbb{X}_i \in T_pM\}.$$

Este espacio es un espacio vectorial con la suma y el producto usuales.

Definición 1.2.11 (Tensores)**.** *Un tensor* (k,l) $\mathbb{T} : \Pi_l^k \to \mathbb{R}$ *es un mapa multilineal (lineal en cada uno de sus argumentos). Este tensor se puede expandir en términos de las bases del espacio tangente y cotangente de la siguiente forma:*

$$\mathbb{T} = T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}\partial_{\mu_1} \otimes \cdots \otimes \partial_{\mu_k} \otimes \mathrm{d}x^{\nu_1} \otimes \cdots \otimes \mathrm{d}x^{\nu_l},$$

donde $T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}$ *son los coeficientes del tensor.*

De modo similar al caso de los vectores, los tensores transforman coordenadas en cada uno de sus índices de la siguiente forma:

$$T^{\mu'_1\cdots\mu'_k}{}_{\nu'_1\cdots\nu'_l} = \frac{\partial x^{\mu'_1}}{\partial x^{\mu_1}}\cdots\frac{\partial x^{\mu'_k}}{\partial x^{\mu_k}}\frac{\partial x^{\nu_1}}{\partial x^{\nu'_1}}\cdots\frac{\partial x^{\nu_l}}{\partial x^{\nu'_l}}T^{\mu_1\cdots\mu_k}{}_{\nu_1\cdots\nu_l}.$$

Infortunadamente, la derivada parcial de un tensor no es, en general, un tensor (no cumple esta regla de transformación de coordenadas). Esto motivará posteriormente la derivada covariante, que preservará el carácter tensorial tras aplicarse sobre un tensor.

Definición 1.2.12 (Tensor métrico)**.** *El tensor métrico* $g_{\mu\nu}$ *es un tensor simétrico* $(0,2)$*, cuya representación matricial tiene determinante no nulo (*$g = |g_{\mu\nu}| \neq 0$*), y satisface la relación*

$$g^{\mu\nu}g_{\nu\sigma} = \delta^\mu{}_\sigma. \tag{1-3}$$

La simetría de $g_{\mu\nu}$ *implica la simetría de* $g^{\mu\nu}$*, y la relación* (1-3) *permite que el tensor métrico se pueda usar para subir o bajar índices.*

Definición 1.2.13 (Elemento de línea)**.** *El elemento de línea se define de la siguiente forma*

$$\mathrm{d}s^2 = g_{\mu\nu}\mathrm{d}x^\mu\mathrm{d}x^\nu.$$

2 Estadística bayesiana

En este capítulo se hace una breve introducción a la teoría bayesiana, con el fin de obtener las herramientas necesarias para ajustar los parámetros de un modelo dado a través de un conjunto de datos experimentales. Algunas referencias clave de este capítulo son [1], [4].

2.1. Preliminares

Antes de enunciar la regla de Bayes, es necesario definir lo que es un espacio de probabilidad, y la probabilidad condicional. Esta sección busca fundamentar las bases de la teoría de la probabilidad, para comprender a cabalidad temas posteriores.

2.1.1. Espacio de probabilidad

La probabilidad es una teoría matemática que busca medir de alguna forma la posibilidad de que ocurra un evento contenido en un conjunto de posibles eventos, resultados todos de un experimento. Por supuesto, no se conoce de forma determinista cuál será el resultado tras la ejecución del experimento, de modo que sólo se puede hablar de posibilidades de que ocurra algún evento. Este tipo de experimentos se conocerán como experimentos aleatorios.

Definición 2.1.1 (Experimento Aleatorio)**.** *Un experimento se dice aleatorio si su resultado no se puede determinar de antemano.*

Definición 2.1.2 (Espacio de Muestra)**.** *El conjunto Ω de todos los posibles resultados de un experimento aleatorio se llama espacio de muestra. Un elemento $\omega \in \Omega$ se llama resultado o muestra. Ω se dice discreto si es finito o contable.*

Ahora se requiere definir lo que se entenderá por evento, para lo cual se definirá una estructura sobre el espacio de muestra, conocida como $\sigma-$álgebra, que dará cuenta de los eventos de interés tras la ejecución de un experimento aleatorio.

Definición 2.1.3 (σ-álgebra)**.** *Tome $\Omega \neq \emptyset$. Una colección $\Im$ de subconjuntos de Ω se llama $\sigma-$álgebra sobre Ω si:*

1. $\Omega \in \Im$,

2. *Si $A \in \Im$, entonces $A^c \in \Im$ y,*

3. *Si* $A_1, A_2, \ldots \in \Im$, *entonces* $\bigcup_{i=1}^{\infty} A_i \in \Im$.

Los elementos de $\Im$ *se llaman eventos.*

El siguiente teorema será de utilidad para construir una σ-álgebra a partir de un conjunto finito o contable de σ-álgebras.

Teorema 2.1.1. *Si* $\Omega \neq \emptyset$ *y* $\Im_1, \Im_2, \ldots$ *son* $\sigma-$*álgebras sobre* Ω, *entonces* $\bigcap_{i=1}^{\infty} \Im_i$ *es una* σ-*álgebra sobre* Ω.

Demostración. Como $\Omega \in \Im_j$, para $j = 1, 2, \ldots$, $\Omega \in \bigcap_{j=1}^{\infty} \Im_j$. Si $A \in \bigcap_{j=1}^{\infty} \Im_j$, $A \in \Im_j$, para $j = 1, 2, \ldots$, de modo que $A^c \in \Im_j$, y $A^c \in \bigcap_{j=1}^{\infty} \Im_j$. Por último, si

$$A_1, A_2, \ldots \in \bigcap_{j=1}^{\infty} \Im_j,$$

para todo $j = 1, 2, \ldots$, $A_1, A_2, \ldots \in \Im_j$, de modo que

$$\bigcup_{i=1}^{\infty} A_i \in \Im_j \text{ y } \bigcup_{i=1}^{\infty} A_i \in \bigcap_{j=1}^{\infty} \Im_j.$$

□

Con este teorema, se puede definir la σ-álgebra más pequeña[1] que contiene un subconjunto de Ω.

Definición 2.1.4 (σ-álgebra generada)**.** *Tome* $\Omega \neq \emptyset$ *y* $\mathcal{A}$ *como una colección de subconjuntos de* Ω. *Si* $\mathcal{M} := \{\Im : \Im$ *es una* $\sigma -$ *álgebra sobre* Ω *que contiene a* $\mathcal{A}\}$,

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

es la $\sigma-$*álgebra más pequeña sobre* Ω *que contiene a* $\mathcal{A}$. *Esta* $\sigma-$*álgebra se conoce como* σ-*álgebra generada por* $\mathcal{A}$.

Definición 2.1.5 (Espacio de medida)**.** *Tome* $\Omega \neq \emptyset$ *y sea* $\Im$ *una* σ-*álgebra sobre* Ω. *La pareja* $(\Omega, \Im)$ *se llama espacio de medida.*

Al evento $\emptyset$ se le conoce como evento imposible; Ω es el evento seguro y $\{\omega\}$, con $\omega \in \Omega$ es un evento simple. Se dice que el evento A ocurre después de llevar a cabo el experimento aleatorio si se obtiene un resultado en A, esto es, A ocurre si el resultado es algún $\omega \in A$.

1. El evento $A \cup B$ ocurre si y sólo si A ocurre, B pasa, o ambos ocurren.

[1] Es la más pequeña en el sentido de que es la que requiere menos elementos para satisfacer las condiciones necesarias para ser una σ-álgebra.

2. El evento $A \cap B$ ocurre si y sólo si A y B ocurren a la vez.

3. El evento A^c ocurre si y sólo si A no ocurre.

4. El evento $A - B$ ocurre si y sólo si A ocurre pero B no ocurre.

Si dos eventos no tienen eventos simples en común, se dirá que son eventos mutuamente excluyentes:

Definición 2.1.6 (Eventos mutuamente excluyentes). *Dos eventos A y B se dicen mutuamente excluyentes si $A \cap B = \emptyset$.*

Antes de introducir la función de probabilidad, que medirá la posibilidad de que ocurra un evento de la $\sigma-$álgebra, es necesario definir por completez la frecuencia relativa, pues ella determina la posibilidad de que ocurra un evento al cabo de n repeticiones del experimento aleatorio.

Definición 2.1.7 (Frecuencia relativa). *Para cada evento A, el número $f_r(A) := \frac{n(A)}{n}$ se llama la frecuencia relativa de A, donde $n(A)$ indica el número de veces que ocurre A en n repeticiones del experimento aleatorio.*

Cuando $n \to \infty$, se puede hablar intuitivamente de la probabilidad de que ocurra el evento A, normalizada de 0 a 1. Por supuesto, es imposible realizar infinitas veces un experimento aleatorio para determinar la probabilidad de ocurrencia de todos los eventos de la $\sigma-$álgebra, por lo que se introduce de antemano la función de probabilidad, suponiendo que ella da cuenta del comportamiento de la frecuencia relativa cuando $n \to \infty$.

Definición 2.1.8 (Espacio de probabilidad). *Tome $(\Omega, \Im)$ como un espacio de medida. Una función real P sobre $\Im$ que satisface las siguientes condiciones:*

1. *$P(A) \geq 0$ para todo $A \in \Im$ (no negativa),*

2. *$P(\Omega) = 1$ (normalizada) y,*

3. *si $A_1, A_2, \ldots$ son eventos mutuamente excluyentes en $\Im$, esto es, si*

$$A_i \cap A_j = \emptyset \text{ para todo } i \neq j, \text{ entonces}$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i),$$

se llama medida de probabilidad sobre $(\Omega, \Im)$. La tripleta $(\Omega, \Im, P)$ se llama espacio de probabilidad.

El siguiente teorema caracteriza las propiedades más importantes de un espacio de probabilidad.

Teorema 2.1.2. *Si $(\Omega, \Im, P)$ es un espacio de probabilidad, entonces*

1. $P(\emptyset) = 0$.
2. *Si $A, B \in \Im$ y $A \cap B = \emptyset$, entonces $P(A \cup B) = P(A) + P(B)$.*
3. *Para todo $A \in \Im$, $P(A^c) = 1 - P(A)$.*
4. *Si $A \subseteq B$, entonces $P(A) \leq P(B)$ y $P(B - A) = P(B) - P(A)$. En particular, $P(A) \leq 1$ para todo $A \in \Im$.*
5. *Para todo $A, B \in \Im$, $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.*
6. *Tome $\{A_n\}_n \subseteq \Im$ como una sucesión creciente, esto es, $A_n \subseteq A_{n+1}, \forall n \in \mathbb{N}$; entonces*
$$P\left(\lim_{m\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n), \text{ donde } \lim_{n\to\infty} A_n := \bigcup_{i=1}^{\infty} A_i.$$
7. *Tome $\{A_n\}_n \subseteq \Im$ como una sucesión decreciente, esto es, $A_n \supseteq A_{n+1}, \forall n \in \mathbb{N}$; entonces*
$$P\left(\lim_{m\to\infty} A_n\right) = \lim_{n\to\infty} P(A_n), \text{ donde } \lim_{n\to\infty} A_n := \bigcap_{i=1}^{\infty} A_i.$$

Demostración.

1. $1 = P(\Omega \cup \emptyset \cup \emptyset \cup \cdots) = P(\Omega) + P(\emptyset) + P(\emptyset) + \cdots = 1 + P(\emptyset) + \cdots \implies P(\emptyset) = 0$.
2. $P(A \cup B) = P(A \cup B \cup \emptyset \cup \emptyset \cup \cdots) = P(A) + P(B)$.
3. $P(A) + P(A^c) = P(A \cup A^c) = P(\Omega) = 1 \implies P(A^c) = 1 - P(A)$.
4. Si $A \subseteq B$, $B = A \cup (B - A)$, de modo que $P(B) = P(A) + P(B - A)$. Como $P \geq 0$, $P(B) \geq P(A)$ y $P(B - A) = P(B) - P(A)$. Si $B = \Omega$, $P(A) \leq 1$.
5. Use el hecho de que $A \cup B = [A - (A \cap B)] \cup [B - (A \cap B)] \cup [A \cap B]$.
6. Tome la sucesión $C_1 = A_1, C_2 = A_2 - A_1, \ldots, C_r = A_r - A_{r-1}, \ldots$. Es claro que
$$\bigcup_{i=1}^{\infty} C_i = \bigcup_{i=1}^{\infty} A_i.$$
Más aún, como $C_i \cap C_j = \emptyset \ \forall i \neq j$, se sigue que
$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{n=1}^{\infty} C_n\right) = \sum_{n=1}^{\infty} P(C_n) = \lim_{n\to\infty} \sum_{k=1}^{n} P(C_k)$$
$$= \lim_{n\to\infty} P\left(\bigcup_{k=1}^{n} C_k\right) = \lim_{n\to\infty} P(A_n).$$

7. Tome la sucesión $\{B_n = A_n^c\}_n$ y aplique el resultado anterior.

$\square$

Aplicando el teorema anterior de forma inductiva, para algunos eventos $A_1, A_2, \dots, A_n \in \Im$:

$$\begin{aligned} P(A_1 \cup A_2 \cup \dots \cup A_n) = &\sum_{i=1}^{n} P(A_i) - \sum_{i_1 < i_2} P(A_{i_1} \cap A_{i_2}) + \cdots \\ &+ (-1)^{r+1} \sum_{i_1 < i_2 < \dots < i_r} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_r}) \\ &+ \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n). \end{aligned}$$

Por otro lado, tome $(\Omega, \Im, P)$ como un espacio de probabilidad con Ω finito o contable y $\Im = \mathbb{P}(\Omega)$. Tome $\emptyset \neq A \in \Im$. Es claro que

$$A = \bigcup_{\omega \in A} \{\omega\}, \text{ de modo que}$$

$$P(A) = \sum_{\omega \in A} P(\omega), \text{ donde } P(\omega) := P(\{\omega\}).$$

Así, P queda completamente definido por $p_j := P(\omega_j)$, donde $\omega_j \in \Omega$. El vector $|\Omega|-$dimensional $p := (p_1, p_2, \dots)$ satisface las siguientes condiciones:

- $p_j \geq 0$ y
- $\sum_{j=1}^{\infty} p_j = 1$.

Un vector que satisface las anteriores condiciones se llama **vector de probabilidad**.

2.1.2. Probabilidad condicional

Tome B como un evento cuya opción de ocurrir debe ser medida bajo la suposición de que otro evento A fue observado. Si el experimento se repite n veces bajo las mismas circunstancias, entonces la frecuencia relativa de B bajo la condición A se define como

$$f_r(B|A) := \frac{n(A \cap B)}{n(A)} = \frac{\frac{n(A \cap B)}{n}}{\frac{n(A)}{n}} = \frac{f_r(A \cap B)}{f_r(A)}, \text{ si } n(A) > 0.$$

Esto motiva la definición de probabilidad condicional, como el comportamiento de esta frecuencia relativa cuando $n \to \infty$.

Definición 2.1.9 (Probabilidad condicional)**.** *Tome $(\Omega, \Im, P)$ como un espacio de probabilidad. Si $A, B \in \Im$, con $P(A) > 0$, entonces la probabilidad del evento B bajo la condición A se define como sigue*

$$P(B|A) := \frac{P(A \cap B)}{P(A).}$$

El siguiente teorema provee algunas propiedades de la probabilidad condicional.

Teorema 2.1.3 (Medida de probabilidad condicional)**.** *Tome* $(\Omega, \Im, P)$ *como un espacio de probabilidad y* $A \in \Im$*, con* $P(A) > 0$*. Entonces:*

1. $P(\cdot|A)$ *es una medida de probabilidad sobre* Ω *centrada en* A*, esto es,* $P(A|A) = 1$*.*
2. *Si* $A \cap B = \emptyset$*, entonces* $P(B|A) = 0$*.*
3. $P(B \cap C|A) = P(B|A \cap C)P(C|A)$ *si* $P(A \cap C) > 0$*.*
4. *Si* $A_1, A_2, \ldots, A_n \in \Im$*, con* $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$*, entonces*

   $$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Demostración.

1. Las tres propiedades de una medida de probabilidad deben ser verificadas.
   a) Claramente, $P(B|A) \geq 0$ para todo $B \in \Im$.
   b) $P(\Omega|A) = \frac{P(\Omega \cap A)}{P(A)} = \frac{P(A)}{P(A)} = 1$. También se tiene que $P(A|A) = 1$.
   c) Tome $A_1, A_2, \ldots \in \Im$ una sucesión de conjuntos disyuntos. Entonces

   $$P\left(\bigcup_{i=1}^{\infty} A_i|A\right) = \frac{P\left(A \cap \bigcup_{i=1}^{\infty} A_i\right)}{P(A)} = \frac{P\left(\bigcup_{i=1}^{\infty} A \cap A_i\right)}{P(A)}$$
   $$= \sum_{i=1}^{\infty} \frac{P(A \cap A_i)}{P(A)} = \sum_{i=1}^{\infty} P(A_i|A).$$

2. Si $A \cap B = \emptyset$, $P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(\emptyset)}{P(A)} = 0$.
3. $P(B \cap C|A) = \frac{P(A \cap B \cap C)}{P(A)} = \frac{P(B \cap C \cap A)}{P(A \cap C)} \frac{P(C \cap A)}{P(A)} = P(B|A \cap C)P(C|A)$.
4. $P(A_1 \cap \cdots \cap A_n) = \frac{P(A_1 \cap \cdots \cap A_n)}{P(A_1 \cap \cdots \cap A_{n-1})} \frac{P(A_1 \cap \cdots \cap A_{n-1})}{P(A_1 \cap \cdots \cap A_{n-2})} \cdots \frac{P(A_1 \cap A_2)}{P(A_1)} P(A_1) = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1})$.

□

El siguiente teorema permitirá probar la regla de Bayes.

Teorema 2.1.4 (Teorema de probabilidad total)**.** *Tome* $A_1, A_2, \ldots$ *como una partición finita o contable de* Ω*, esto es,* $A_i \cap A_j = \emptyset, \forall i \neq j$ *y* $\bigcup_{i=1}^{\infty} A_i = \Omega$*, tal que* $P(A_i) > 0$*, para todo* $A_i \in \Im$*. Entonces, para todo* $B \in \Im$*:*

$$P(B) = \sum_i P(B|A_i)P(A_i).$$

Demostración. Observe que

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{\infty} A_i \right) = \bigcup_{i=1}^{\infty} B \cap A_i,$$

de modo que

$$P(B) = P\left(\bigcup_{i=1}^{\infty} B \cap A_i \right) = \sum_{i=1}^{\infty} P(B \cap A_i) = \sum_{i=1}^{\infty} P(B|A_i)P(A_i).$$

□

Matemáticamente, este teorema se puede interpretar como que la probabilidad de que ocurra B se puede medir en términos de una partición de Ω en el sentido de que B, como subconjunto de Ω puede ocurrir cuando ocurran algunos elementos de la partición, los cuales tendrán mayor peso en el término $P(B|A_i)$ de la suma.
Como corolario del teorema anterior, se obtiene la **regla de Bayes**, que constituye la base para la **teoría Bayesiana**.

Corolario 2.1.1 (Regla de Bayes)**.** *Tome $A_1, A_2, \ldots$ como una partición finita o contable de Ω con $P(A_i) > 0$, para todo i; entonces, para todo $B \in \Im$ con $P(B) > 0$:*

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)P(A_j)}, \forall i.$$

Demostración.

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_j P(B|A_j)P(A_j)}.$$

Con la partición $A_1 = A, A_2 = A^c$ se obtiene la forma usual de la regla de Bayes. □

A continuación se definen las distribuciones a priori y a posteriori, que hacen referencia a la probabilidad de que ocurran ciertos eventos de una partición de Ω antes de que ocurra un evento B y al cabo de que este evento B ocurra.

Definición 2.1.10 (Distribuciones a priori y a posteriori)**.** *Tome $A_1, A_2, \ldots$ como una partición finita o contable de Ω, con $P(A_i) > 0$, para todo i. Si $P(B) > 0$, con $B \in \Im$, entonces $\{P(A_n)\}_n$ se llama distribución a priori (antes de que B ocurra), y $\{P(A_n|B)\}_n$ se llama distribución a posteriori (después de que B ocurra).*

Algunas veces, la ocurrencia de un evento B no afecta la probabilidad de un evento A, es decir,

$$P(A|B) = P(A).$$

En este caso, se dice que el evento A es independiente del evento B. Esto motiva la siguiente definición.

Definición 2.1.11 (Eventos independientes)**.** *Dos eventos A y B se dicen independientes si y sólo si*

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

Si esta condición no se tiene, se dice que los eventos son dependientes.

En algunos casos, es necesario analizar la independencia de dos o más eventos. Para ello, se dan las siguientes definiciones.

Definición 2.1.12 (Familia independiente)**.** *Una familia de eventos $\{A_i : i \in I\}$ se dice independiente si*

$$P\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} P(A_i),$$

para cualquier subconjunto no vacío $J \subseteq I$.

Definición 2.1.13 (Eventos independientes par a par)**.** *Una familia de eventos $\{A_i : i \in I\}$ se dice par a par independiente si*

$$P(A_i \cap A_j) = P(A_i)P(A_j), \text{ para todo } i \neq j.$$

2.1.3. Variables aleatorias

En un experimento aleatorio, generalmente hay mayor interés en determinar ciertos valores numéricos asociados a los resultados del experimento aleatorio, que al resultado mismo del experimento aleatorio. Con esto en mente se define la variable aleatoria.

Definición 2.1.14 (Variable aleatoria)**.** *Tome $(\Omega, \Im, P)$ como un espacio de probabilidad. Una variable aleatoria es un mapa $X : \Omega \to \mathbb{R}$ tal que, para todo $A \in \mathbb{B}$, $X^{-1}(A) \in \Im$, donde $\mathbb{B}$ es la σ-álgebra de Borel sobre $\mathbb{R}$ (σ-álgebra más pequeña que contiene todos los intervalos de la forma $(-\infty, a]$).*
El conjunto de posibles valores de X es $\mathbb{S} := \{x \in \mathbb{R} : \exists \omega \in \Omega \text{ tal que } X(\omega) = x\}$, conocido como ***soporte de la variable aleatoria*** *X.*

Si X es una variable aleatoria definida sobre un espacio de probabilidad $(\Omega, \Im, P)$, se introduce la notación

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\}, \text{ con } B \in \mathbb{B}.$$

Definición 2.1.15 (Variable aleatoria discreta)**.** *Una variable aleatoria X se dice discreta cuando el soporte $\mathbb{S}$ de X es un subconjunto finito o contable de $\mathbb{R}$. Para $x \in \mathbb{S}$, la función $f(x) = P(X = x) := P($ se llama función de densidad de probabilidad (pdf para abreviar).*

Definición 2.1.16 (Variable aleatoria continua)**.** *Una variable aleatoria X se dice continua si el soporte $\mathbb{S}$ de X es la unión de uno o más intervalos y si existe una función no negativa y real $f(x)$ tal que $P(X \leq x) = \int_{-\infty}^{x} f(t)\mathrm{d}t$. La función $f(x)$ se llama función de densidad de probabilidad (pdf).*

Algunas propiedades de la *pdf* discreta son las siguientes:

1. $f(x) \geq 0, \forall x \in \mathbb{S}$ y $f(x) = 0, \forall x \notin \mathbb{S}$.
2. $\sum_{x \in \mathbb{S}} f(x) = 1$.
3. $P(X \in B) = \sum_{x \in B} f(x)$.

Análogamente, para la *pdf* continua:

1. $f(x) \geq 0, \forall x \in \mathbb{S}$ y $f(x) = 0, \forall x \notin \mathbb{S}$.
2. $\int_{\mathbb{S}} f(x)\mathrm{d}x = 1$.
3. $P(X \in B) = \int_B f(x)\mathrm{d}x$.

Definición 2.1.17 (Función de distribución acumulativa)**.** *La función de distribución acumulativa (CDF, para abreviar) de una variable aleatoria se define como la función $F(x) = P(X \leq x)$.*

El siguiente teorema resume algunas propiedades importantes de una *CDF*.

Teorema 2.1.5. *Si X es una variable aleatoria, con CDF $F(x)$, entonces:*

1. $\lim_{x \to -\infty} F(x) = 0$ *y* $\lim_{x \to \infty} F(x) = 1$.
2. *$F(x)$ es no decreciente; esto es, $F(x) \leq F(y)$, siempre que $x \leq y$.*
3. *$F(x)$ es continua por derecha.*
4. *$P(a < X \leq b) = F(b) - F(a)$.*

Demostración. 1. **Cuando $\mathbf{x} \to -\infty$**, para todo $n \in \mathbb{N}$, se satisface que

$$\{X \leq -n\} \supseteq \{X \leq -(n+1)\}, \text{ y en adición,}$$

$$\emptyset = \bigcap_{n=1}^{\infty} \{X \leq -n\}, \text{ con lo que}$$

$$\lim_{x \to -\infty} F(x) = \lim_{n \to \infty} F(-n) = P(\emptyset) = 0.$$

Cuando $\mathbf{x} \to \infty$, para todo $n \in \mathbb{N}$, se satisface que

$$\{X \leq n\} \subseteq \{X \leq n+1\}, \text{ y,}$$

$$\Omega = \bigcup_{n=1}^{\infty} \{X \leq n\}, \text{ de modo que}$$

$$\lim_{x\to\infty} F(x) = \lim_{n\to\infty} = P(\Omega) = 1.$$

2. Si $x \leq y$, entonces

$$\{X \leq x\} \subseteq \{X \leq y\}, \textit{ con lo que}$$

$$F(x) = P(X \leq x) \leq P(X \leq y) = F(y).$$

3. Tome $x \in \mathbb{R}$ fijo. Suponga que $\{x_n\}_{n\in\mathbb{N}}$ es una sucesión decreciente de números reales, cuyo límite va a x. Se puede ver que

$$\{X \leq x_1\} \supseteq \{X \leq x_2\} \supseteq \cdots, \text{ y}$$

$$\bigcap_{n=1}^{\infty} \{X \leq x_n\} = \{X \leq x\}, \text{ de modo que}$$

$$\lim_{y\to x^+} F(y) = \lim_{n\to\infty} F(x_n) = \lim_{n\to\infty} P(X \leq x_n) = P(X \leq x) = F(x).$$

4. Dado que $\Omega = \{X \leq a\} \cup \{a < X \leq b\} \cup \{X > b\}$,

$$1 = P(X \leq a) + P(a < X \leq b) + P(X > b),$$

$$\therefore P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a).$$

$\square$

Los teoremas posteriores indican cómo determinar la *pdf* a partir de la *CDF* de una variable aleatoria.

Teorema 2.1.6. *Si X es una variable aleatoria discreta con CDF $F(x)$ y soporte $\mathbb{S} = \{x_0, x_1, \ldots\}$, con $x_0 < x_1 < \cdots$, entonces, para $x_k \in \mathbb{S}$,*

$$f(x_k) = F(x_k) - F(x_{k-1}).$$

Demostración. Para $k \geq 1$,

$$f(x_k) = P(X = x_k) = P(x_{k-1} < X \leq x_k) = F(x_k) - F(x_{k-1}).$$

$\square$

Teorema 2.1.7. *Para una variable aleatoria continua, $f(x) = \frac{\mathrm{d}F}{\mathrm{d}x}, \forall x \in \mathbb{R}$.*

Demostración. La prueba se sigue directamente del teorema fundamental del cálculo. $\square$

2.1.4. Vectores aleatorios

En la mayoría de los análisis estadísticos, más de una variable debe ser analizada al cabo de un experimento aleatorio. Cada observación se puede representar como un vector de observaciones, conocido como vector aleatorio.

Definición 2.1.18 (Vector aleatorio)**.** *Un vector aleatorio $\vec{X} = (X_1, X_2, \ldots, X_k)$ es un vector $k-$dimensional, donde $X_1, \ldots, X_k$ son variables aleatorias. Un vector aleatorio se dice discreto cuando cada una de las variables aleatorias que lo conforman son discretas, y continuo cuando son continuas.*

Definición 2.1.19 (Variable aleatoria bivariada)**.** *Un vector aleatorio bidimensional $\vec{X} = (X_1, X_2)$ se llama variable aleatoria bivariada.*

De modo similar al caso de las variables aleatorias, los vectores aleatorios tienen *pdf*, un soporte y una *CDF*. El soporte de un vector aleatorio $k-$dimensional es el conjunto de valores que puede tomar, denotado por $\mathbb{S}_{\vec{X}} \subseteq \mathbb{R}^k$.

Definición 2.1.20 (Función de densidad de probabilidad adjunta discreta)**.** *Tome $\vec{X}$ como un vector aleatorio discreto $k-$dimensional. La pdf adjunta de $\vec{X}$ se define como*

$$f(\vec{x}) := f(x_1, x_2, \ldots, x_k) = P(X_1 = x_1, X_2 = x_2, \ldots, X_k = x_k)$$

para $\vec{x} = (x_1, x_2, \ldots, x_k) \in \mathbb{S}_{\vec{X}}$.

La *pdf* adjunta discreta tiene las siguientes propiedades:

1. $0 \leq f(x_1, x_2, \ldots, x_k) \leq 1, \forall \vec{x} \in \mathbb{S}_{\vec{X}}$.
2. $\sum_{\vec{x} \in \mathbb{S}_{\vec{X}}} f(\vec{x}) = 1$.
3. Para cualquier subconjunto $B \subseteq \mathbb{S}_{\vec{X}},\ P(\vec{X} \in B) = \sum_{\{\vec{x} \in \mathbb{S}_{\vec{X}} : \vec{x} \in B\}} f(\vec{x})$.

Definición 2.1.21 (Función de densidad de probabilidad adjunta continua)**.** *Tome $\vec{X}$ como un vector aleatorio k-dimensional continuo. La pdf continua de $\vec{X}$ se define como cualquier función no negativa $f(\vec{x})$ que satisfaga las siguientes propiedades:*

1. *$f(x_1, \ldots, x_k) > 0, \forall \vec{x} \in \mathbb{S}_{\vec{X}}$.*
2. *$\int_{\mathbb{S}_{\vec{X}}} f(x_1, \ldots, x_k) \mathrm{d}x_1 \cdots \mathrm{d}x_k = 1$.*
3. *Para cualquier subconjunto $B \subset \mathbb{S}_{\vec{X}},\ P(\vec{X} \in B) = \int_B f(x_1, \ldots, x_k) \mathrm{d}x_1 \cdots \mathrm{d}x_k$.*

Un vector aleatorio también tiene *CDF*, conocida en este caso como función de distribución acumulativa adjunta o *CDF* adjunta.

Definición 2.1.22 (Función de distribución acumulativa adjunta)**.** *Tome $\vec{X} = (X_1, X_2, \ldots, X_k)$ como un vector aleatorio $k-$dimensional. La CDF adjunta de $\vec{X}$ se define como*

$$F(x_1, x_2, \ldots, x_k) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_k \leq x_k),\ \forall (x_1, \ldots, x_k) \in \mathbb{R}^k.$$

Las componentes de un vector aleatorio $\vec{X}$ son variables aleatorias, por lo que las *pdf* de cada variable aleatoria X_i de $\vec{X}$ se pueden derivar de la *pdf* adjunta de $\vec{X}$.

Definición 2.1.23 (Función de densidad de probabilidad marginal)**.** *Tome $\vec{X} = (X_1, \ldots, X_k)$ como un vector aleatorio $k-$dimensional. La función de densidad de probabilidad marginal de la variable aleatoria X_i es*

$$f_i(x_i) = \underbrace{\sum_{x_1 \in \mathbb{S}_X} \cdots \sum_{x_k \in \mathbb{S}_{X_k}}}_{\textit{quitando la suma sobre } x_i} f(x_1, \ldots, x_k) \qquad \textit{cuando } \vec{X} \textit{ es discreta y}$$

$$f_i(x_i) = \underbrace{\int_{x_1 \in \mathbb{S}_{X_1}} \cdots \int_{x_k \in \mathbb{S}_{X_k}}}_{\textit{quitando la integral sobre } x_i} f(x_1, \ldots, x_k) \prod_{n \neq i} \mathrm{d}x_n \qquad \textit{cuando } \vec{X} \textit{ es continua.}$$

También se puede definir la distribución condicional dada una variable aleatoria de forma similar al caso de la probabilidad condicional.

Definición 2.1.24 (Función de densidad de probabilidad condicional)**.** *Tome $\vec{X}(X_1, \ldots, X_k)$ como un vector aleatorio $k-$dimensional. Para un valor fijo de x_i, donde $f_i(x_i) > 0$, la función de densidad de probabilidad condicional para $\vec{Y}|X_i$; donde $\vec{Y}$ es un vector aleatorio $(k-1)-$dimensional con todas las variables aleatorias de $\vec{X}$, a excepción de X_i; es*

$$f(\vec{y}|x_i) = \frac{f(x_1, \ldots, x_k)}{f_i(x_i)},$$

donde $\vec{y} \in \mathbb{S}_{\vec{Y}}$.

Bibliografía

[1] L. Blanco, V. Arunachalam, and Dharmaraja S. *Introduction to Probability and Stochastic Processes with Applications*. Wiley, 2012.

[2] S. Carroll. *Lecture Notes on General Relativity*. Institute of Theoretical Physics, University of California, 1997.

[3] J. Munkres. *Topology*. Pearson Education Limited, 2014.

[4] R. Rossi. *Mathematical statistics: an introduction to likelihood based inference*. John Wiley Sons, Inc., 2018.

[5] K. Thorne, J. Wheeler, and C. Misner. *Gravitation*. W.H. Freeman and Company, 1973.