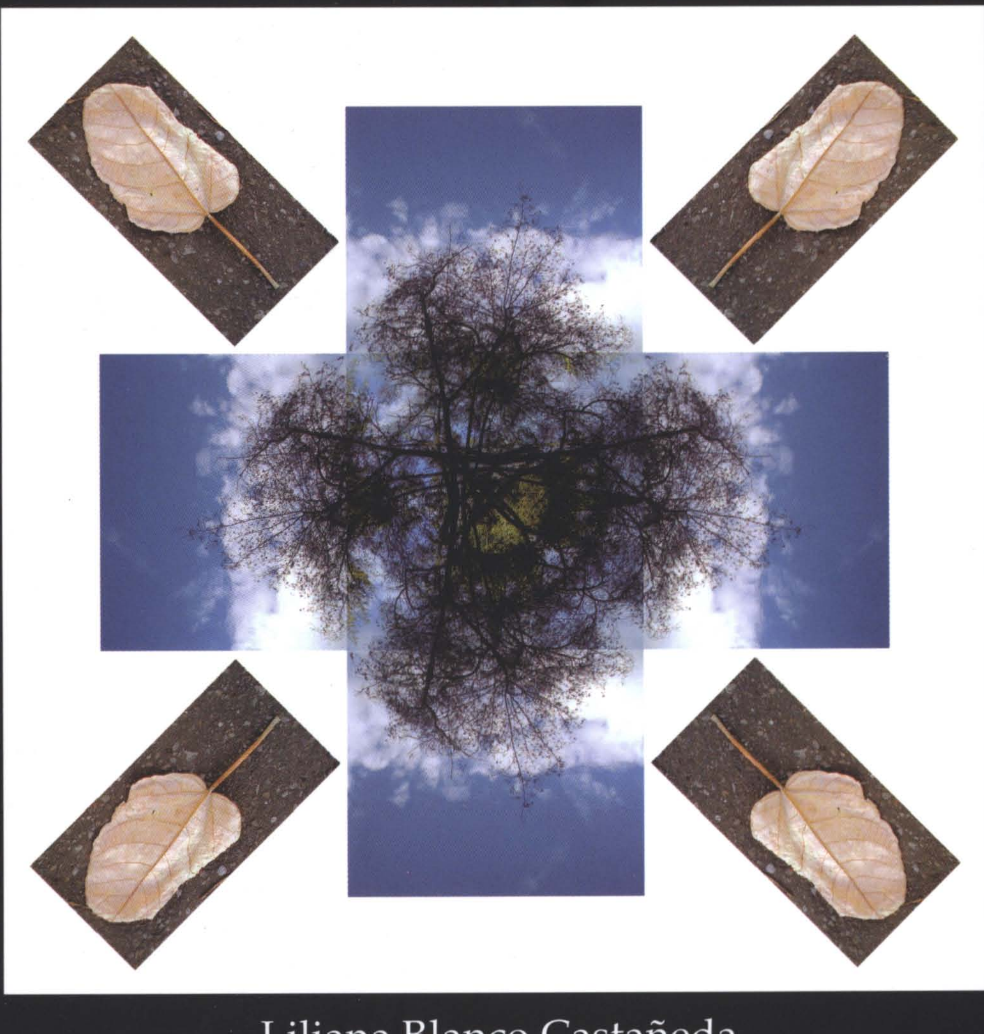
Colección Texto

PROBABILIDAD



Liliana Blanco Castañeda

UNIVERSIDAD NACIONAL DE COLOMBIA

Facultad de Ciencias

Probabilidad

Probabilidad

Liliana Blanco Castañeda
Dr. rer.nat
Profesora Asociada

Universidad Nacional de Colombia

Facultad de Ciencias
Sede Bogotá

Blanco, Liliana,
Probabilidad / Liliana Blanco Castañeda
Universidad Nacional de Colombia, Sede Bogotá
Facultad de Ciencias, 2004

vi, 380 p. : 32 il.

ISBN: 958-701-449-9

1. Teoría de Probabilidad
2. Procesos Estocásticos

MATHEMATICS SUBJECT CLASSIFICATIO: 2000:60-01
Universidad Nacional de Colombia

PROBABILIDAD

Decano: Moisés Wasserman
Vicedecana: Natalia Ruiz
Director de Publicaciones: Gustavo Rubiano

Primera Edición, 2004

ISBN: 958-701-449-9

Impreso por:
Universidad Nacional de Colombia
UNIBIBLOS
Bogotá, Colombia
2004

A mis hijos Paula y Sebastián

y a mi sobrina Catalina

Índice general

Prefacio

Este texto está diseñado para un primer curso de teoría de probabilidad en las carreras de matemáticas y estadística, así como para los estudiantes de la especialización en estadística que no cuenten con conocimientos previos en el área. La idea de escribir este texto surgió hace ya varios años, como respuesta a la necesidad de contar con un texto en español, que abarcara los principales temas que deben ser vistos por los estudiantes de las carreras antes mencionadas. Fue así como en el año 1998 publiqué una primera versión del presente texto, en la colección "Notas de clase" de la Facultad de Ciencias de la Universidad Nacional de Colombia. Luego, a partir de la experiencia otorgada por la enseñanza del curso de probabilidad durante varios semestres, en los cuales se detectaron, gracias al aporte crítico de colegas y estudiantes, varias deficiencias y carencias en esa primera versión, ésta fue corregida y aumentada hasta dar origen al presente trabajo.

El texto está dividido en ocho capítulos y cuatro apéndices. En el primer capítulo se presentan los conceptos básicos de la teoría tales como: espacios de probabilidad, eventos independientes, probabilidad condicional y regla de Bayes. En el segundo se discuten los conceptos de variable aleatoria, distribución y función de distribución de una variable aleatoria, valor esperado, varianza, funciones generadora de momentos y característica. En el tercer y cuarto capítulo se presentan, respectivamente, las distribuciones de tipo discreto y continuo de uso más frecuente en las aplicaciones. El quinto capítulo está dedicado al estudio de los vectores aleatorios y sus distribuciones; se estudian entre otros temas, la distribución conjunta de variables aleatorias, los conceptos de independencia de variables aleatorias y de covarianza y coeficiente de correlación, se trabaja también la distribución de una función de un vector aleatorio, obteniendo como caso particular, las distribuciones de la suma, diferencia, producto y cociente de variables aleatorias. Finalmente se dedican las últimas secciones al estudio de los conceptos de valor esperado de un vector aleatorio, matriz de varianzas y covarianzas, funciones generadoras de momentos y características conjuntas y se termina con una introducción al estudio de la distribución normal multivariada. En el sexto capítulo se trabajan los conceptos de probabilidad y esperanza condicional. El estudio de los teoremas límites es el objetivo del séptimo capítulo, en él se estudian tres tipos de convergencia de sucesiones de variables aleatorias: convergencia estocástica, casi siempre y en distribución, se establecen relaciones entre ellas y se presentan las leyes débil y fuerte de los grandes números, así como el Teorema Central del Límite. En el octavo y último capítulo se hace una introducción a la modelación aleatoria. Se explica cómo obtener algoritmos que permiten calcular probabilidades, valores esperados, así como simular valores de variables aleatorias con un número contable de posibles resultados y de variables aleatorias continuas con función de distribución conocida. En el apéndice uno se hace una breve introducción a la teoría de conjuntos, haciendo especial énfasis en los conceptos de imagen directa e inversa de una función. El el segundo apéndice se presentan los conceptos básicos del análisis combinatorio, desarrollando ejemplos ilustrativos del principio fundamental del conteo. En el tercer apéndice se presenta un resumen de los conceptos del álgebra lineal usados a lo largo del texto. Por último, el cuarto apéndice contiene las tablas de las funciones de distribución más usadas en las aplicaciones. Estas tablas fueron generadas con el programa MATLAB.

Al final de cada capítulo hay una serie de ejercicios, con ellos el lector podrá verificar su comprensión de los temas desarrollados y encontrará, en algunos casos, material adicional de estudio.

Para la comprensión del texto, el lector debe tener conocimientos sólidos del cálculo diferencial e integral en una y varias variables.

Deseo agradecer de manera especial al profesor Ignacio Mantilla, quien además de brindarme su apoyo moral, colaboró no sólo en el desarrollo del octavo capítulo, escribiendo e implementando los algoritmos que en él aparecen, sino también en la generación de las tablas que aparecen en el cuarto apéndice.

Por último deseo agradecer a la Universidad Nacional de Colombia, y en especial a su Departamento de Estadística, por brindarme el tiempo y las facilidades necesarias que hicieron posible la elaboración de este texto; a mis colegas Fabio Nieto y Humberto Mayorga de la Universidad Nacional y Fernando Ruiz del IDEAM, quienes colaboraron en la corrección del manuscrito, así como al Profesor Gustavo Rubiano, Director de Publicaciones de la Facultad de Ciencias, quien me motivó permanentemente.

Capítulo 1

Conceptos básicos

La teoría de la probabilidad ha sido relacionada desde sus comienzos con los juegos de azar, de hecho ya en los tiempos del primer emperador romano Augusto (63 A.C-14 D.C), eran comunes los juegos de azar y se hacían tablas de mortandad. Éste fue el origen de la probabilidad y la estadística. Posteriormente estas dos disciplinas se fueron separando debido a sus distintos objetivos, pero sin dejar de estar relacionadas. En el siglo XVI se sostuvieron discusiones filosóficas sobre la probabilidad, se destaca en esta época el filósofo italiano Gerolamo Cardano quien fue uno de los primeros en hacer un tratamiento matemático del azar. Entre los siglos XVII y XVIII se hicieron importantes avances en la teoría de la probabilidad debido en parte al desarrollo del cálculo infinitesimal; de este período sobresalen, entre otros resultados, los siguientes: la ley de los grandes números, debida a James Bernoulli, esta ley es un teorema de límite básico en la teoría moderna de probabilidad, una interpretación sencilla de ella establece que si se realiza un experimento aleatorio en el que hay sólo dos posibles resultados: éxito y fracaso, entonces la proporción de éxitos obtenidos tiende a estabilizarse alrededor de un número entre 0 y 1 (que resulta ser la probabilidad de éxito), a medida que aumenta el número de repeticiones; y el teorema de DeMoivre-Laplace de 1733, 1785 y 1812, el cual establece que cuando n es suficientemente grande, una variable aleatoria binomial con parámetros n y p tiene aproximadamente la misma distribución que una variable aleatoria normal con media np y varianza $np(1-p)$. Este resultado fue probado inicialmente por DeMoivre para el caso $p = \frac{1}{2}$ en el año 1733 y luego extendido, al caso $0 < p < 1$ arbitrario por Laplace en el año 1812. A pesar de haber obtenido resultados teóricos de la envergadura de los mencionados anteriormente, es necesario anotar que en esta época, no había claridad en los conceptos básicos. Para entonces, ya era conocida la famosa definición de Laplace del concepto de probabilidad como el cociente de casos favorables sobre casos posibles, suponiendo que todos los resultados del experimento aleatorio subyacente son igualmente probables. Pero ¿qué significaba: "igualmente probables"? En el año 1892, el matemático alemán Karl Stumpf interpreta esta expresión afirmando que los eventos son igualmente probables, cuando no se tiene ningún conocimiento de cuál de los distintos resultados del experimento en cuestión va a ocurrir. En contraposición con este punto de vista, se tiene el del matemático alemán Johann von Kriess, quien consideraba que para determinar los resultados igualmente probables de un experimento, era necesario tener un conocimiento objetivo de éste. Así por ejemplo, si sólo sabe que una urna contiene bolas negras y blancas, entonces, según Strumpf, para la primera extracción es igualmente probable sacar una bola blanca o una bola negra, en tanto que von Kriess admitiría este supuesto sólo si se conoce que el número de bolas negras en la urna es igual al número de bolas blancas. Se dice que el mismo Markov tuvo dificultades al respecto: según Krengel [Kre] en el libro de texto de Markov (1912), se encuentra el siguiente ejemplo: "supóngase que en una urna hay bolas de cuatro colores $1, 2, 3$ y 4 con frecuencias desconocidas a, b, c y d, entonces la probabilidad de extraer una bola de color 1 es igual a $\frac{1}{4}$, pues los cuatro colores son igualmente probables". Como se observa, no había claridad, aún en el caso de experimentos aleatorios con un número finito de posibles resultados, acerca de su modelación matemática.

A comienzos del siglo XX y a pesar de haberse ocupado del tema algunos de los más famosos matemáticos, como por ejemplo Cardano, Fermat, Bernoulli, Laplace, Poisson y Gauss, de cuestiones probabilísticas, la teoría de la probabilidad no era reconocida dentro del medio académico como una disciplina matemática y se discutía si más bien se trataba de una disciplina empírica. En el famoso Segundo Congreso Internacional de Matemáticas, realizado en agosto del año 1900, David Hilbert plantea, en su transcendental conferencia del 8 de agosto, como sexto problema la tarea de axiomatizar el cálculo de probabilidades. En el año 1901 G. Bohlmann formuló una primera aproximación a la axiomatización de la probabilidad [Kre]: él define la probabilidad de un evento E como un número no negativo $p(E)$ para el cual se satisface:

i) Si E es el evento seguro entonces $p(E) = 1$.

ii) Si E_1 y E_2 son dos eventos, tales que ellos ocurren simultáneamente con probabilidad cero, entonces la probabilidad de que E_1 o E_2 ocurran es igual a $p(E_1) + p(E_2)$.

En el año 1907 el italiano Ugo Broggi escribió, bajo la dirección de

Hilbert, su trabajo de doctorado titulado "Die Axiome der Wahrscheinlichkeitsrechnung" (*Los axiomas del cálculo de probabilidades*). La definición de evento se presenta de manera difusa y se afirma que la aditividad y la $\sigma-$aditividad son equivalentes (la demostración de este resultado falso contiene tal cantidad de errores, que es de suponer que Hilbert no leyó en detalle el trabajo). Sin embargo, puede considerarse este trabajo como el predecesor de los trabajos de Kolmogorov.

En el congreso internacional de matemáticas llevado a cabo en Roma en el año 1908, Bohlman define la independencia de eventos en la forma conocida actualmente y muestra la diferencia de este concepto con el de la independencia dos a dos. Cabe anotar que aún no se había dado una definición precisa de evento.

De acuerdo a lo relatado por Krengel [Kre], en el año 1901 el matemático sueco Wiman utiliza el concepto de medida en su definición de probabilidad geométrica. A este respecto, dice Borel en el año 1905 lo siguiente: "Cuando se usa la convención: la probabilidad de un conjunto es proporcional a su longitud, área o volumen, entonces se debe ser explícito y aclarar que esto es sólo una convención más y no una definición de probabilidad".

Gracias a los trabajos de Fréchet y Caratheodory, quienes "liberaron" la teoría de la medida de su interpretación geométrica, se abrió el camino para la axiomatización de la teoría de la probabilidad tal y como se la conoce hoy en día. En el famoso libro "Grundbegriffe der Wahrscheinlichkeitsrechnung" (Fundamentos de la teoría de probabilidad) publicado por primera vez en el año 1933, el matemático ruso Andrei Nikolaevich Kolmogorov (1903-1987) axiomatizó la teoría de la probabilidad haciendo uso de la teoría de la medida, quedando rigurosamente definidos los conceptos de espacio de probabilidad, evento, variable aleatoria, independencia de eventos y de variables aleatorias, probabilidad condicional, entre otros. Pero en el trabajo realizado por Kolmogorov no sólo se establecieron de manera explícita los axiomas y definiciones básicas de la teoría del cálculo de probabilidades, sino que además se formularon las bases de la teoría de los procesos estocásticos, en particular se realizaron importantes contribuciones en el desarrollo de los procesos de Markov y de los procesos de ramificación. Uno de los resultados más importantes presentados por Kolmogorov es el teorema de consistencia, el cual es fundamental cuando se desea garantizar la existencia de procesos estocásticos como elementos aleatorios de espacios de dimensión infinita.

La teoría de la probabilidad no sólo es atractiva por ser una teoría matemática compleja sino por sus múltiples aplicaciones a otros campos del conocimiento científico. Su amplia gama de aplicaciones abarca tópicos en física, química, genética, ecología, comunicaciones, demografía y finanzas, entre otros muchos.

A principios del siglo XX, uno de los problemas científicos más importantes era comprender el llamado movimiento browniano, nombrado así en honor al botánico inglés Robert Brown, quien observó que las partículas de polen suspendidas en un líquido se mueven de manera incesante e irregular. Brown pensó, en un comienzo, que el movimiento se debía a la naturaleza orgánica del polen, sin embargo el mismo refuta esta teoría al verificar, con un simple experimento, que el movimiento se presentaba también en sustancias inorgánicas.

Desde los trabajos realizados por Brown y hasta finales del siglo XIX no se sabe de otras investigaciones acerca del movimiento browniano. En 1905 Einstein publica en su artículo titulado "Über die von der molekularkinetischen Theorie der Warme gefordete Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen" (Acerca de la teoría de la cinética molecular del movimiento, inducido por el calor, de partículas suspendidas en un líquido) [Kah] los aspectos principales del movimiento browniano, él prueba que el movimiento de la partícula en el instante t se puede modelar por medio de la distribución normal y concluye además que el movimiento se debe a los continuos choques de la partícula con las moléculas del líquido en el cual se halla suspendida. Cabe anotar sin embargo, como el mismo Einstein lo reconoce, que él desconocía los trabajos de Brown [Nel]. La primera investigación matemática acerca del movimiento browniano es debida a Louis Bachelier quien en el año 1900 propusó en su tesis doc-toral "Theorie de la spéculation" (Teoría de la especulación) al movimiento browniano como modelo asociado a los precios especulativos. Una de las imperfecciones del modelo propuesto por Bachelier fue que en éste los precios podían ser negativos y por esto el modelo cayó en el olvido durante largo tiempo. En el año 1960 el economista Samuelson (premio Nobel de economía del año 1970) propuso la exponencial del movimiento browniano para modelar el comportamiento de los precios que están sujetos a incertidumbre.

La estructura matemática del movimiento browniano, tal y como se la conoce hoy en día, es debida al famoso matemático norteamericano Norbert Wiener (1894-1964). Por esta razón el movimiento browniano también es llamado proceso de Wiener. Los primeros artículos de Wiener sobre movimiento browniano son muy difíciles de leer y sólo el renombrado matemático francés Paul Lévy (1886-1971) pudo reconocer su importancia. Paul Lévy contribuyó, de manera notable, al desarrollo de la teoría de la probabilidad y entre sus contribuciones más importantes se destacan: el concepto de martingala, los procesos de Lévy, que incluyen, entre otros, al movimiento browniano y al de Poisson, y el teorema de continuidad de funciones características. Además, Lévy dedujo varias de las propiedades más importantes del movimiento browniano. Se dice (ver [Gor]) que ha ocurrido muchas veces, que cosas que se creían nuevos descubrimientos importantes en teoría de probabilidad ya estaban contenidos, de alguna manera, en los trabajos de Lévy.

En los años setenta del siglo pasado, se describió la fórmula de Black-Scholes y Merton, que permite la valoración de opciones de compra y venta europeas. Por este trabajo se les otorgó a Scholes y Merton el premio Nobel en economía en el año 1997 (para esta fecha Black ya había fallecido). El desarrollo dado por Black-Scholes y Merton, sin embargo, no hubiese sido posible de no haber sido por los trabajos realizados por el matemático japónes Kyosi Itô (1915-) publicados en los años 1940 y 1946, en los cuales presenta una serie de artículos donde introduce dos de los conceptos más importantes de la teoría moderna de probabilidad: la integral estocástica y las ecuaciones diferenciales estocásticas. Estos conceptos han sido además una herramienta esencial en muchos campos de la matemática, como por ejemplo en la teoría de las ecuaciones diferenciales parciales, así como en las aplicaciones, no sólo a las matemáticas financieras, sino a la física teórica, la biología y la ingeniería (ver [Kor]).

Como se ve, la teoría de la probabilidad es no sólo una teoría matemática hermosa sino que ofrece una amplia gama de aplicaciones a problemas concretos de la vida real, ya que muchos fenómenos que se presentan en la naturaleza son de origen aleatorio y sólo con técnicas probabilistas pueden ser modelados.

Espero que este libro de texto motive a los estudiantes de matemáticas y estadística al estudio de la teoría de la probabilidad, no sólo por su riqueza teórica y su gran cantidad de aplicaciones, sino porque conjuga las distintas teorías matemáticas: álgebra, ecuaciones diferenciales, análisis funcional con un sólo fin: tratar de entender el mundo en el que vivimos.

1.1. Espacios de probabilidad

En esta sección se va a desarrollar el concepto de medida de probabilidad y a presentar sus propiedades básicas.

Cuando se lanza un dado corriente una vez, no se puede predecir cuál será el resultado que se va a obtener, sin embargo, se sabe que el conjunto de posibles resultados es $\{1,2,3,4,5,6\}$. Un experimento de éste tipo es lo que se llama *experimento aleatorio.* Esto es:

Definición 1.1 (experimento aleatorio) *Un experimento se dice aleatorio si su resultado no puede ser determinado de antemano.*

Se supone que el conjunto de todos los posibles resultados de un experimento aleatorio es conocido. A este conjunto se le llama *espacio muestral.* Más precisamente:

Definición 1.2 (espacio muestral) *El conjunto Ω de todos los posibles resultados de un experimento aleatorio se llama espacio muestral. Los elementos $\omega \in \Omega$ son llamados puntos muestrales.*

Ejemplo 1.3 *Experimento: lanzamiento de una moneda corriente. Los posibles resultados, en este caso, son "cara" $= c$ y "sello" $= s$. Esto es,* $\Omega = \{c, s\}$. ▲

Ejemplo 1.4 *Experimento: lanzamiento de un dado corriente 3 veces consecutivas. En este caso, los posibles resultados son triplas (a, b, c) con $a, b, c \in \{1, 2, 3, 4, 5, 6\}$. Esto es,*

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

Ejemplo 1.5 *Experimento: se observa el número de veces que es necesario lanzar una moneda corriente hasta obtener por primera vez cara. En este caso:*

$$\Omega = \{1, 2, ..., \infty\} = \mathbb{N} \cup \{\infty\},$$

obsérvese que el punto muestral "1" indica que se obtuvo cara en el primer lanzamiento, el punto muestral "4" indica que en los tres primeros lanzamientos se obtuvo "sello" y en el cuarto "cara", el punto muestral "∞" indica que nunca se obtiene "cara". ▲

Ejemplo 1.6 *Experimento: se observa la posición de una partícula que se mueve aleatoriamente sobre el eje real. En este caso $\Omega = \mathbb{R}$.* ▲

Se observa que los puntos muestrales pueden ser números, vectores, símbolos, etc. los cuales están determinados por el experimento considerado.

Definición 1.7 *El espacio muestral Ω se llama discreto si es finito o numerable. Un experimento aleatorio se llama finito (discreto) si su espacio muestral es finito (discreto).*

Volviendo al ejemplo 1.4, una pregunta que surge de manera natural es: ¿cuál es el "chance" que tiene un "evento", tal como "la suma de los resultados obtenidos es mayor o igual que 2", de ocurrir?. Esto es, ¿cuál es el "chance" de ocurrir de

$$A := \{(a,b,c) \in \Omega : a+b+c \geq 2\}?$$

Pero, ¿qué es un evento? Si se sigue la idea expuesta anteriormente, se puede pensar que un evento es simplemente un subconjunto del espacio muestral, pero entonces ¿son todos los subconjuntos de un espacio muestral eventos? La respuesta a esta pregunta es no. La clase de subconjuntos del espacio muestral para los que estará definido el "chance" que tienen de ocurrir debe tener estructura de $\sigma-$álgebra, concepto que se precisará a continuación.

Definición 1.8 (σ-álgebra) *Sea $\Omega \neq \Phi$. Una colección $\Im$ de subconjuntos de Ω es una $\sigma-$álgebra sobre Ω, si:*

i) $\Omega \in \Im$.

ii) Si $A \in \Im$ entonces $A^c \in \Im$.

iii) Si $A_1, A_2, ... \in \Im$ entonces $\bigcup_{i=1}^{\infty} A_i \in \Im$.

Los elementos de $\Im$ se llaman eventos.

Ejemplo 1.9 *Sea $\Omega \neq \Phi$. Entonces $\Im_0 = \{\Phi, \Omega\}$ y $\wp(\Omega) := \{A : A \subseteq \Omega\}$ son $\sigma-$álgebras sobre Ω. $\Im_0$ recibe el nombre de $\sigma-$álgebra trivial y $\wp(\Omega)$ recibe el nombre de $\sigma-$álgebra total.* ▲

Ejemplo 1.10 *Sea $\Omega = \{1,2,3\}$. Entonces $\Im = \{\Phi, \{1\}, \{2,3\}, \Omega\}$ es una $\sigma-$álgebra sobre Ω, en tanto que $\Re = \{\Phi, \{1\}, \{2\}, \{3\}, \Omega\}$ no lo es.* ▲

Ejemplo 1.11 *Sean Ω y $\widetilde{\Omega}$ conjuntos no vacios y $\widetilde{\Im}$ una $\sigma-$álgebra sobre $\widetilde{\Omega}$. Si $T : \Omega \to \widetilde{\Omega}$ es una función, entonces la colección $T^{-1}(\widetilde{\Im}) = \{T^{-1}(A) : A \in \widetilde{\Im}\}$ es una $\sigma-$álgebra sobre Ω.* ▲

Ejemplo 1.12 *Sean $\Omega \neq \Phi$ y $\overline{\Omega}$ un subconjunto no vacio de Ω. Si $\Im$ es una $\sigma-$álgebra sobre Ω, entonces $\overline{\Im} = \{A \cap \overline{\Omega} : A \in \Im\}$ es una σ-álgebra sobre $\overline{\Omega}$, llamada la huella de $\Im$ sobre $\overline{\Omega}$.* ▲

Ejemplo 1.13 *Sean $\Omega \neq \Phi$ finito o numerable y $\Im$ una $\sigma-$álgebra sobre Ω que contiene todos los conjuntos de la forma $\{\omega\}$ con $\omega \in \Omega$. Entonces $\Im = \wp(\Omega)$.* ▲

Ejemplo 1.14 *Si* $\Omega \neq \Phi$ *y* $\Im_1, \Im_2, \ldots$ *son* $\sigma-$*álgebras sobre* Ω *entonces* $\bigcap\limits_{i=1}^{\infty} \Im_i$ *es una* $\sigma-$*álgebra sobre* Ω. ▲

Ejemplo 1.15 ($\sigma-$álgebra generada) *Sean* $\Omega \neq \Phi$ *y* L *una colección de subconjuntos de* Ω. *Sea* $\mathcal{M} = \{\Im : \Im$ *es una* $\sigma-$*álgebra sobre* Ω *que contiene a* $L\}$. *Entonces, por el ejemplo anterior, se tiene que* $\sigma(L) := \bigcap\limits_{\Im \in \mathcal{M}} \Im$ *es la menor* $\sigma-$*álgebra sobre* Ω *que contiene a* L. *Esta* $\sigma-$*álgebra se llama* $\sigma-$*álgebra generada por* L. ▲

Ejemplo 1.16 ($\sigma-$álgebra de Borel) *La menor* $\sigma-$*álgebra sobre* $\mathbb{R}$ *que contiene todos los intervalos de la forma* $(-\infty, a]$ *con* $a \in \mathbb{R}$ *se llama* $\sigma-$*álgebra de Borel y se denota por* $\mathcal{B}$. *Puesto que* $\mathcal{B}$ *es una* $\sigma-$*álgebra entonces los siguientes conjuntos también pertenece a* $\mathcal{B}$:

$$\begin{aligned}
(a, \infty) &= \mathbb{R} - (-\infty, a], \\
(a, b] &= (-\infty, b] \cap (a, \infty), \\
(-\infty, a) &= \bigcup_{n=1}^{\infty} (-\infty, a - \frac{1}{n}], \\
[a, \infty) &= \mathbb{R} - (-\infty, a), \\
(a, b) &= (-\infty, b) \cap (a, \infty), \\
[a, b] &= \mathbb{R} - ((-\infty, a) \cup (b, \infty)), \\
\{a\} &= [a, a], \\
\mathbb{N} &= \bigcup_{n=0}^{\infty} \{n\}, \\
\mathbb{Q} &= \bigcup_{r \in \mathbb{Q}} \{r\}, \\
\mathbb{Q}^c &= \mathbb{R} - \mathbb{Q}.
\end{aligned}$$

donde $a, b \in \mathbb{R}$ *y* $a < b$. *¿Son entonces todos los subconjuntos de* $\mathbb{R}$ *conjuntos boreleanos? La respuesta a esta pregunta es no. En el libro de Royden [Roy] hay un ejemplo de un subconjunto de* $\mathbb{R}$ *que no es boreleano.* ▲

Ejemplo 1.17 ($\sigma-$álgebra de Borel en $\mathbb{R}^n$) *Sean* $a = (a_1, \cdots, a_n)$ *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* $\sigma-$*álgebra generada por todos los intervalos de la forma:*

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

se llama $\sigma-$*álgebra de Borel en* $\mathbb{R}^n$. ▲

Definición 1.18 (espacio medible) *Sean $\Omega \neq \Phi$ y $\Im$ una $\sigma-$álgebra sobre Ω. La pareja $(\Omega, \Im)$ se llama espacio medible.*

Es claro, a partir de la definición, que Ω y Φ pertenecen a cualquier $\sigma-$álgebra definida sobre Ω. Ω se llama evento seguro y Φ se llama evento imposible. Un evento de la forma $\{\omega\}$ con $\omega \in \Omega$ se llama evento elemental.

Decir que un evento A ocurre significa que el resultado obtenido, al realizar el experimento aleatorio cuyo espacio muestral es Ω, es un elemento de A. Por lo tanto si A y B son eventos entonces:

i) $A \cup B$ es un evento que ocurre, si y sólo si, A o B o ambos ocurren.
ii) $A \cap B$ es un evento que ocurre, si y sólo, si A y B ocurren.
iii) A^c es un evento que ocurre, si y sólo si, A no ocurre.
iv) $A - B$ es un evento que ocurre, si y sólo si, A ocurre pero B no.

Ejemplo 1.19 *Si en el ejemplo 1.4 se consideran los eventos: $A =$ "el resultado del primer lanzamiento es un número primo" y $B =$ " la suma de los resultados obtenidos es menor o igual a 4", entonces:*

$$A \cup B = \{(a, b, c) \in \Omega : a \in \{2, 3, 5\} \text{ o } (a + b + c) \leq 4\},$$

así por ejemplo, $(2, 1, 1), (5, 3, 4), (1, 1, 1)$ son elementos de $A \cup B$. Por otra parte

$$A \cap B = \{(a, b, c) : a \in \{2, 3, 5\} \text{ y } (a + b + c) \leq 4\} = \{(2, 1, 1)\}.$$

¿A qué son iguales $A - B$ y A^c? (Ejercicio!). ▲

Definición 1.20 (eventos mutuamente excluyentes) *Dos eventos A y B se dicen mutuamente excluyentes si $A \cap B = \Phi$.*

Ejemplo 1.21 *Se lanza una moneda corriente una vez. Sean $A =$ "el resultado obtenido es cara" y $B =$ "el resultado obtenido es sello". Es claro que A y B son mutuamente excluyentes.* ▲

Ejemplo 1.22 *Se lanza una moneda corriente tantas veces como sea necesario hasta obtener cara por primera vez y se cuenta el número de lanzamientos necesarios para ello. Si*

$A :=$ *"no se obtiene cara antes del tercer lanzamiento"* $= \{3, 4, 5, ..., \infty\}$ *y*

$B :=$ *"no se obtiene cara antes del segundo lanzamiento"* $= \{2, 4, 5, ..., \infty\}$

entonces, A y B no son mutuamente excluyentes. ▲

1.1.1. Concepto de probabilidad

El objetivo a continuación es asignar a cada evento A un número real no negativo que indique el "chance" que tiene A de ocurrir. Supóngase que se realiza un experimento aleatorio n veces y que las condiciones en que éste se ejecuta se mantienen más o menos constantes.

Definición 1.23 (frecuencia relativa) *Para cada evento A, el número $fr(A) := \frac{n(A)}{n}$ se llama frecuencia relativa de A, donde $n(A)$ indica el número de veces que ocurre el evento A.*

Ejemplo 1.24 *Supóngase que se lanza una moneda corriente 100 veces y que en 60 de esos lanzamientos el resultado es "cara", entonces las frecuencias relativas de los eventos $A :=$ "el resultado es cara" y $B :=$ "el resultado es sello" son respectivamente $\frac{3}{5}$ y $\frac{2}{5}$.* ▲

Ejemplo 1.25 *Se lanza un dado corriente 500 veces y se obtiene 83 veces la cara con el número 3. En este caso la frecuencia relativa del evento*

$$A := \text{“el resultado obtenido es 3”}$$

es igual a $\frac{83}{500}$. ▲

Desafortunadamente se tiene que para cada A fijo, $fr(A)$ no es constante pues su valor depende de n; sin embargo se ha observado que cuando un experimento aleatorio se realiza un número suficientemente grande de veces, bajo condiciones similares, la frecuencia relativa $fr(A)$ se estabiliza alrededor de un valor específico entre 0 y 1.

Ejemplo 1.26 *Supóngase que se lanza un dado corriente un número n de veces y sea*

$$A := \text{“el resultado obtenido es 3”}.$$

La tabla siguiente 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

. ▲

La estabilización de la frecuencia relativa es lo que se conoce como "regularidad estadística" y es lo que le permite al hombre hacer predicciones que eliminan, aunque sea parcialmente, la incertidumbre presente en los fenómenos impredecibles.

El valor $P(A)$ alrededor del cual se estabiliza la frecuencia relativa de un evento indica el "chance" que tiene éste de ocurrir. Interesa saber cuáles propiedades tiene este número $P(A)$. En primer lugar se observa que puesto que $n(A) \geq 0$ entonces $P(A)$ debe ser mayor o igual a cero, como $n(\Omega) = n$ entonces $fr(\Omega) = 1$ y en consecuencia $P(\Omega) = 1$. Por otra parte si A y B son eventos mutuamente excluyentes entonces se satisface que $n(A \cup B) = n(A) + n(B)$ y por lo tanto $fr(A \cup B) = fr(A) + fr(B)$. Esto implica que $P(A \cup B) = P(A) + P(B)$ si $A \cap B = \Phi$. Las anteriores observaciones inducen a dar la siguiente definición:

Definición 1.27 (espacio de probabilidad) *Sea* $(\Omega, \Im)$ *un espacio medible. Una función* P *definida sobre* $\Im$ *y de valor real que satisface las siguientes condiciones:*

i) $P(A) \geq 0$ *para todo* $A \in \Im$.

ii) $P(\Omega) = 1$

iii) *Si* $A_1, A_2, \ldots$ *son elementos de* $\Im$ *mutuamente excluyentes, esto es*

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

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 tripla $(\Omega, \Im, P)$ *se llama espacio de probabilidad.*

Ejemplo 1.28 *Sean* $\Omega = \{1, 2, 3\}$, $\Im = \{\Phi, \Omega, \{1\}, \{2, 3\}\}$ *y* P *la aplicación definida sobre* $\Im$ *como sigue:*

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

Es fácil ver que P *es una medida de probabilidad sobre* $(\Omega, \Im)$. ▲

Ejemplo 1.29 *Sean* $\Omega = \{1,2\}$, $\Im = \wp(\Omega)$ *y* P *la aplicación definida sobre* $\Im$ *por:*

$$P(A) = \begin{cases} 0 & si\ A = \Phi \\ \frac{1}{3} & si\ A = \{1\} \\ \frac{2}{3} & si\ A = \{2\} \\ 1 & si\ A = \{1,2\} \end{cases}$$

P es una medida de probabilidad. ▲

Ejemplo 1.30 (medida de Lebesgue) *Sea* $\Omega = [0,1]$, $\overline{\mathcal{B}}$ *la huella de* $\mathcal{B}$ *en* $[0,1]$ *y* λ *la medida de probabilidad definida sobre* $\overline{\mathcal{B}}$ *que a cada intervalo finito contenido en* $[0,1]$ *le asigna su longitud (una construcción precisa de* λ *va más allá de los objetivos de este texto. Se supondrá simplemente que tal medida existe). La tripla* $([0,1], \overline{\mathcal{B}}, \lambda)$ *es un espacio de probabilidad. La medida de probabilidad* λ *se llama medida de Lebesgue restringida al intervalo* $[0,1]$.

Definición 1.31 (evento nulo) *Sea* $(\Omega, \Im, P)$ *un espacio de probabilidad. Cualquier evento* A *con probabilidad* 0 *se llama evento nulo.*

Ejemplo 1.32 *Sea* $\Omega = \{1,2,3\}$, $\Im = \{\Phi, \{2\}, \{1,3\}, \Omega\}$ *y* P *dada por:*

$$P(A) = \begin{cases} 1 & si\ \ 2 \in A \\ 0 & si\ \ 2 \notin A \end{cases}.$$

Los eventos Φ *y* $\{1,3\}$ *son eventos nulos.* ▲

Definición 1.33 (espacio completo) *Un espacio de probabilidad* $(\Omega, \Im, P)$ *se dice completo si todos los subconjuntos de eventos nulos son eventos nulos.*

Ejemplo 1.34 *El espacio de probabilidad del ejemplo anterior no es completo, pues* $\{1\} \subset \{1,3\}$ *pero* $\{1\}$ *no es un evento y por consiguiente* $P(\{1\})$ *no está definida.* ▲

Nota 1.35 *Todos los espacios que se consideren de ahora en adelante (y mientras no se diga lo contrario) se suponen completos.*

A continuación se establecen las propiedades más importantes de una medida de probabilidad P.

Teorema 1.36 *Sea $(\Omega, \Im, P)$ un espacio de probabilidad. Entonces:*

1. $P(\Phi) = 0$

2. *Si $A, B \in \Im$ y $A \cap B = \Phi$ entonces $P(A \cup B) = P(A) + P(B)$.*

3. $P(A^c) = 1 - P(A)$.

4. *Si $A \subseteq B$ entonces $P(A) \leq P(B)$ y $P(B - A) = P(B) - P(A)$. En particular se tiene que $P(A) \leq 1$ para todo $A \in \Im$.*

5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

6. *Sea $(A_n)_n$ una sucesión creciente de elementos de $\Im$, esto es $A_n \in \Im$ para todo n y además $A_n \subseteq A_{n+1}$ para todo $n = 1, 2, \cdots$, entonces:*

$$P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$$

   donde $\lim\limits_{n\to\infty} A_n := \bigcup\limits_{n=1}^{\infty} A_n$.

7. *Sea $(A_n)_n$ una sucesión decreciente de elementos de $\Im$, esto es $A_n \in \Im$ para todo n y además $A_n \supseteq A_{n+1}$ para todo $n = 1, 2, ...$, entonces:*

$$P(\lim_{n\to\infty} A_n) = \lim_{n\to\infty} P(A_n)$$

   donde $\lim\limits_{n\to\infty} A_n := \bigcap\limits_{n=1}^{\infty} A_n$.

Demostración.

1. $1 = P(\Omega) = P(\Omega \cup \Phi \cup \Phi \cup \cdots) = P(\Omega) + P(\Phi) + P(\Phi) + \cdots$. Entonces $0 \geq P(\Phi) \geq 0$ de donde se concluye $P(\Phi) = 0$.

2. $A \cup B = A \cup B \cup \Phi \cup \Phi \cup \cdots$. Por la propiedad iii) de medida de probabilidad y el resultado anterior, se obtiene lo pedido.

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

4. $B = A \cup (B - A)$. Por 2. se sigue que $P(B) = P(A) + P(B - A) \geq P(A)$.

5. Como ejercicio.

6. Sean $C_1 = A_1,\ C_2 = A_2 - A_1, \cdots, C_n = A_n - A_{n-1}$. Es claro que:

$$\bigcup_{n=1}^{\infty} C_n = \bigcup_{n=1}^{\infty} A_n$$

Además como $C_i \cap C_j = \Phi$ para todo $i \neq j$, se sigue de la propiedad iii) de medida de probabilidad que:

$$\begin{aligned} P\left(\bigcup_{n=1}^{\infty} A_n\right) &= P(\bigcup_{n=1}^{\infty} C_n) \\ &= \sum_{n=1}^{\infty} P(C_n) \\ &= \lim_{n\to\infty} \sum_{k=1}^{n} P(C_k) \\ &= \lim_{n\to\infty} P(\bigcup_{k=1}^{n} C_k) \\ &= \lim_{n\to\infty} P(A_n) \end{aligned}$$

7. Como ejercicio.

■

Nota 1.37 *Sean A, B y C eventos. Entonces, aplicando el teorema anterior se obtiene:*

$$\begin{aligned} P(A \cup B \cup C) &= P(A) + P(B \cup C) - P(A \cap (B \cup C)) \\ &= P(A) + P(B) + P(C) - P(B \cap C) - P(A \cap B) \\ &\qquad - P(A \cap C) + P(A \cap B \cap C) \end{aligned}$$

Haciendo un razonamiento de tipo inductivo, se deduce que si $A_1, A_2, ..., A_n$ son eventos, entonces:

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

donde la suma

$$\sum_{i_1<i_2<...<i_r} P(A_{i_1} \cap A_{i_2} \cap \cdots \cap A_{i_r})$$

se toma sobre todos los posibles subconjuntos de tamaño r del conjunto $\{1, 2, ..., n\}$.

Nota 1.38 *Sean $(\Omega, \Im, P)$ un espacio de probabilidad con Ω finito o numerable y $\Im = \wp(\Omega)$. Sea $\Phi \neq A \in \Im$. Es claro que*

$$A = \underset{\omega \in A}{\cup} \{\omega\}$$

y por consiguiente

$$P(A) = \sum_{\omega \in A} P(\omega)$$

donde $P(\omega) := P(\{\omega\})$. Esto es P está completamente determinada por $p_j := P(\omega_j)$ donde los ω_j, $j = 1, 2, ...$,denotan los elementos de Ω.

Es claro que el vector $p := (p_1, p_2, ...)$ de dimensión $|\Omega|$, donde $|\Omega|$ denota el número de elementos de Ω, satisface las condiciones:

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

Un vector p que satisface las condiciones anteriores se llama vector de probabilidades.

Nota 1.39 *Sean $\Phi \neq \Omega = \{\omega_1, \omega_2, ...\}$ un conjunto finito o numerable, $\wp(\Omega)$ la $\sigma-$álgebra total sobre Ω y p un vector de probabilidades $|\Omega|-$dimensional. Es fácil ver que la aplicación P definida sobre $\wp(\Omega)$ por:*

$$\begin{aligned} P(\Phi) &= 0 \\ P(\omega_j) &= p_j \ , \ j = 1, 2, ... \\ P(A) &= \sum_{\{j \, : \, \omega_j \in A\}} p_j \quad \text{para } \Phi \neq A \subseteq \Omega \end{aligned}$$

es una medida de probabilidad. El espacio de probabilidad $(\Omega, \wp(\Omega), P)$ se llama espacio de probabilidad discreto.

Ejemplo 1.40 *Sea $(\Omega, \Im, P)$ un espacio de probabilidad con:*

$$\begin{aligned} \Omega &= \{1, 2, 3, 4\} \\ \Im &= \{\Phi, \Omega, \{1\}, \{2, 3\}, \{4\}, \{1, 2, 3\}, \{2, 3, 4\}, \{1, 4\}\} \end{aligned}$$

$$P(\{1\}) = \frac{1}{4}, \quad P(\{2, 3\}) = \frac{1}{2} \quad y \quad P(\{4\}) = \frac{1}{4}$$

Entonces:

$$\begin{aligned} P(\{1,2,3\}) &= \frac{3}{4} \\ P(\{2,3,4\}) &= \frac{3}{4} \\ P(\{1,4\}) &= \frac{1}{2}. \quad \blacktriangle \end{aligned}$$

Ejemplo 1.41 *Sea* $(\Omega, \wp(\Omega), P)$ *un espacio de probabilidad discreto con* $\Omega = \{a, b, c\}$ *y* P *determinada por el vector de probabilidades* $p = (\frac{1}{7}, \frac{4}{7}, \frac{2}{7})$. *Entonces:*

$$P(\{a,b\}) = \frac{5}{7}, \quad P(\{b,c\}) = \frac{6}{7} \quad y \quad P(\{a,c\}) = \frac{3}{7}. \quad \blacktriangle$$

Ejemplo 1.42 *Sea* $(\Omega, \Im, P)$ *un espacio de probabilidad. Si* A *y* B *son eventos con* $P(A) = p$, $P(B) = q$ *y* $P(A \cup B) = r$, *entonces:*

$$\begin{aligned} P(A \cap B) &= p + q - r \\ P(A - B) &= r - q \\ P(A^c \cap B^c) &= 1 - r \\ P(A \cup B^c) &= p - r + 1 \quad \blacktriangle \end{aligned}$$

Ejemplo 1.43 *Sean* $(\Omega, \Im, P)$ *un espacio de probabilidad y* A *y* B *elementos de* $\Im$ *con* $P(A) = \frac{1}{2}$ *y* $P(B) = \frac{2}{3}$. *Entonces:*

$$\frac{1}{6} \leq P(A \cap B) \leq \frac{1}{2}$$

pues $P(A \cap B) \leq P(A) = \frac{1}{2}$ *y* $P(A \cup B) \leq 1$ $\blacktriangle$

Ejemplo 1.44 *Se lanza un dado cargado una vez. Supóngase que:*

j	1	2	3	4	5	6
p_j	$\frac{1}{32}$	$\frac{5}{32}$	$\frac{7}{32}$	$\frac{3}{32}$	$\frac{8}{32}$	$\frac{8}{32}$

Entonces la probabilidad de obtener un número que no sea divisible por 3 *y cuyo cuadrado sea menor que* 20 *es igual a* $\frac{9}{32}$; *en tanto que la probabilidad de obtener un número* i *tal que* $|i - 5| \leq 3$ *es igual a* $\frac{31}{32}$. $\blacktriangle$

1.1.2. Espacios de probabilidad laplacianos

Dentro de los experimentos aleatorios los más fáciles de analizar son los llamados experimentos laplacianos. Estos son experimentos que tienen un número finito de posibles resultados, cada uno con la misma probabilidad de ser obtenido.

El lanzamiento de una moneda corriente o el de un dado corriente, un número finito de veces, son ejemplos de experimentos laplacianos.

Definición 1.45 (espacio laplaciano)
Un espacio de probabilidad $(\Omega, \Im, P)$ *se llama laplaciano, si* Ω *es finito,* $\Im = \wp(\Omega)$ *y* $P(\omega) = \frac{1}{|\Omega|}$ *para todo* $\omega \in \Omega$. *La medida de probabilidad* P *se llama distribución laplaciana (o uniforme o clásica) en* Ω.

Nota 1.46 *Si* $(\Omega, \Im, P)$ *es un espacio de probabilidad laplaciano y* $A \subseteq \Omega$ *entonces:*

$$P(A) = P\left(\bigcup_{\omega \in A} \{\omega\}\right) = \sum_{\omega \in A} \frac{1}{|\Omega|} = \frac{|A|}{|\Omega|}.$$

En otras palabras,

$$P(A) = \frac{\text{“número de casos favorables a } A\text{”}}{\text{“número de casos posibles”}}.$$

Esta última expresión no es de ningún modo una definición de probabilidad sino simplemente una consecuencia de suponer que todos los resultados son igualmente probables.

En un espacio de probabilidad laplaciano se tiene por lo tanto, que el cálculo de probabilidades se reduce a contar el número de elementos de un conjunto finito, es decir, se llega a un problema de análisis combinatorio. Para los lectores no familiarizados con este tema se incluye un apéndice con los conceptos y resultados básicos de esta teoría.

Ejemplo 1.47 *En una lotería se escogen seis números de* 49. *La probabilidad de que los números* 1, 2, 3, 4, 5 *y* 6 *sean escogidos es igual a*

$$\frac{1}{\binom{49}{6}} = 7.1511 \times 10^{-8}.$$

Observe que esta probabilidad es igual a la probabilidad de que los números 4, 23, 24, 35, 40 *y* 45 *sean escogidos.*

La probabilidad p *de que* 44 *sea uno de los números escogidos es igual a:*

$$p = \frac{\binom{48}{5}}{\binom{49}{6}} = 0.12245. \quad \blacktriangle$$

Ejemplo 1.48 *Diez personas se encuentran sentadas aleatoriamente en una mesa redonda. La probabilidad p de que dos miembros de una pareja en particular estén sentados juntos es igual a:*

$$p = \frac{2!8!}{9!} = \frac{2}{9}. \quad \blacktriangle$$

Ejemplo 1.49 *En un taller de reparación de electrodomésticos se encuentran 10 televisores para reparación, de los cuales 3 son de marca A, 3 de marca B y 4 de marca C. El orden en el cual los televisores son reparados es aleatorio.*

La probabilidad p_1 de que el primer televisor en ser reparado sea de marca A es igual a:

$$p_1 = \frac{3.9!}{10!} = 0.3.$$

La probabilidad p_2 de que los tres primeros televisores en ser reparados sean de marca A es igual a:

$$p_2 = \frac{3.3.7!}{10!} = \frac{1}{120}.$$

La probabilidad p_3 de que el orden de reparación de los televisores sea $CABCABCABC$ es igual a:

$$p_3 = \frac{4.3^3.2^3}{10!} = \frac{1}{4200}. \quad \blacktriangle$$

Ejemplo 1.50 *En el juego de bridge, se reparte la baraja completa de 52 cartas entre 4 jugadores. Se desea calcular la probabilidad de que uno de los jugadores reciba todas las 13 picas.*

En este caso se tiene que, el número total de formas posibles de repartir la baraja entre los cuatro jugadores es igual a:

$$\binom{52}{13, 13, 13, 13} = \binom{52}{13}\binom{39}{13}\binom{26}{13}\binom{13}{13}$$

y el número total de formas de repartir la baraja para que uno de los jugadores tenga todas las picas es igual a:

$$\binom{4}{1}\binom{39}{13}\binom{26}{13}\binom{13}{13}.$$

Por lo tanto, la probabilidad pedida p es igual a:

$$p = \frac{4}{\binom{52}{13}} = 6.2991 \times 10^{-12}. \quad \blacktriangle$$

Ejemplo 1.51 *Supóngase que los cumpleaños de las personas pueden ocurrir con igual probabilidad en cualquiera de los 365 días del año (se ignoran años bisiestos y el hecho de que las tasas de natalidad no son uniformes durante el año). La probabilidad p de que no haya dos personas, en un grupo de n personas, con el mismo día de cumpleaños es igual a:*

$$\begin{aligned} p &= \frac{365 \times 364 \times ... \times (365 - n + 1)}{365^n} \\ &= \left(1 - \frac{1}{365}\right)\left(1 - \frac{2}{365}\right)...\left(1 - \frac{n-1}{365}\right). \quad \blacktriangle \end{aligned}$$

Ejemplo 1.52 (modelos de urna) *En una urna se encuentran N bolas del mismo tipo, R de las bolas son rojas y $(N - R)$ son blancas. Se extraen al azar n bolas.Se desea calcular la probabilidad de que exactamente $k \leq n$ de las bolas extraídas sean rojas.*

Para facilitar el razonamiento, se va a suponer que las bolas están numeradas del 1 al N y que la numeración de las bolas rojas va del 1 a R. Para resolver el problema, es necesario distinguir dos casos importantes: extracción sin sustitución y extracción con sustitución. En el primer caso se deben considerar a su vez dos alternativas: las bolas son extraídas una a una y las bolas son extraídas todas al tiempo.

1. *Extracción sin sustitución (alternativa una a una): Las n bolas son extraídas una a una de la urna y dejadas fuera de ésta. En este caso el espacio muestral está dado por:*

$$\begin{aligned} \Omega = \{(a_1, a_2, ..., a_n) : a_j \in \{1, 2, ..., N\}, a_i \neq a_j \\ \textit{para todo } i \neq j,\ j = 1, 2, ..., n\}. \end{aligned}$$

Sea

$$A_k := \text{“}\textit{exactamente } k \leq n \textit{ de las bolas extraídas son rojas}\text{”}.$$

Es claro que A_k consta de todas las n-plas de Ω con exactamente k componentes menores o iguales a R. Por lo tanto:

$$|\Omega| = N \times (N - 1) \times \cdots \times (N - (n - 1)) =: (N)_n$$

y

$$\begin{aligned} |A_k| &= \binom{n}{k} R \times (R - 1) \times \cdots \times (R - k + 1) \times (N - R) \times \\ &\qquad \cdots \times (N - R - n + k + 1) \\ &= \binom{n}{k} (R)_k \times (N - R)_{(n-k)}. \end{aligned}$$

Luego:

$$P(A_k) = \frac{|A_k|}{|\Omega|} = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}.$$

2. *Extracción sin sustitución (alternativa todas al tiempo): El espacio muestral es en este caso igual a:*

$$\Omega = \{T : T \subseteq \{1, 2, ..., N\} \text{ con } |T| = n\}.$$

Aqui A_k consta de todos los elementos de Ω que contienen exactamente k elementos menores o iguales a R.

Por lo tanto:

$$|\Omega| = \binom{N}{n} \quad y \quad |A_k| = \binom{R}{k}\binom{N-R}{n-k},$$

luego:

$$P(A_k) = \frac{|A_k|}{|\Omega|} = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}.$$

Como se ve, en el caso de la extracción sin sustitución, es irrelevante para el cálculo de la probabilidad del evento A_k, si las bolas fueron extraídas una a una o todas al tiempo.

3. *Extracción con sustitución: En este caso cada bola extraída es devuelta inmediatamente a la urna, después de mezclar nuevamente las bolas, se extrae aleatoriamente la siguiente bola y así sucesivamente.*

El espacio muestral, en esta situación, está dado por:

$$\Omega = \{(a_1, a_2, ..., a_n) : a_j \in \{1, 2, ..., N\},\ j = 1, 2, ..., N\}.$$

El evento A_k está constituido por todas las n-plas de Ω que tienen k componentes menores o iguales a R.

Entonces:

$$|\Omega| = N^n \ y \ |A_k| = \binom{n}{k} R^k (N-R)^{n-k}$$

y en consecuencia:

$$P(A_k) = \frac{|A_k|}{|\Omega|} = \binom{n}{k} p^k q^{n-k} \qquad \text{donde } p = \frac{R}{N} \quad y \quad q = 1 - p. \quad \blacktriangle$$

Ejemplo 1.53 *Supóngase que n bolas están distribuidas en n urnas, de tal manera que los n^n arreglos posibles son igualmente probables. Hallar la probabilidad de que solamente la primera urna esté vacia.*

Solución: Sea A el evento que indica que sólo la primera urna está vacía. Este evento ocurre sólo si las n bolas están distribuidas en las $(n-1)$ urnas restantes de tal manera que ninguna de ellas se encuentre vacía. Así exactamente una de las $(n-1)$ urnas debe tener exactamente dos bolas y las restantes $(n-2)$ deben tener una bola cada una. Sea B_j el evento que indica que la urna j, con $j = 2, \cdots, n$, tiene dos bolas, la primera urna está vacía y las restantes $(n-2)$ urnas contienen exactamente una bola cada una. Es claro que los eventos B_j son dos a dos disyuntos y que su unión es A. Para calcular $P(B_j)$ se observa que las dos bolas colocadas en la urna j pueden ser escogidas de $\binom{n}{2}$ formas y que las restantes $(n-2)$ bolas pueden ser distribuidas en las restantes urnas de $(n-2)!$ maneras. Por lo tanto,

$$P(B_j) = \frac{\binom{n}{2}(n-2)!}{n^n}$$

y así:

$$P(A) = \sum_{j=2}^{n} P(B_j) = \frac{(n-1)\binom{n}{2}(n-2)!}{n^n}. \quad \blacktriangle$$

1.2. Probabilidad condicional e independencia de eventos

Muchas veces se obtiene una información parcial acerca de un experimento aleatorio antes de que sea conocido el resultado final. Con base en esta información se cambia, por lo general, la estructura probabilística de los posibles resultados. Por ejemplo, un jugador de póker puede en un momento dado ver las cartas de un compañero de juego. Supóngase que el tiempo fue tan corto, que el jugador sólo pudo observar que todas las cartas de su compañero son de color rojo, esto es, corazones o diamantes. El jugador sabe entonces que su compañero no puede tener los cuatro reyes, evento que en un principio tenía probabilidad positiva. Por otra parte, tiene la sospecha que el evento "tener todas las cartas de la misma pinta" es más probable de lo que era antes de obtener la información adicional.

A continuación se va a analizar la situación desde el punto de vista de las frecuencias relativas. Sea B un evento cuyo chance de ocurrir debe ser medido bajo la suposición de que un evento A ha sido observado. Si el experimento se repite, bajo las mismas condiciones, n veces entonces la frecuencia relativa de B bajo la condición A se define como sigue:

$$fr(B \mid A) := \frac{n(A \cap B)}{n(A)}; \text{ si } n(A) > 0$$

siendo $n(A \cap B)$ el número de casos favorables a $A \cap B$.

Es claro que $fr(B \mid A)$ depende de n. Sin embargo, cuando el experimento se realiza un número suficientemente grande de veces, las frecuencias relativas tienden a estabilizarse alrededor de un valor específico entre 0 y 1, conocido como probabilidad condicional del evento B bajo la condición A.

Se observa que:

$$fr(B \mid A) = \frac{\frac{n(A\cap B)}{n}}{\frac{n(A)}{n}} = \frac{fr(A \cap B)}{fr(A)}; \text{ si } n(A) > 0.$$

Cuando n es suficientemente grande, el numerador en la anterior expresión tiende a $P(A \cap B)$ y el denominador a $P(A)$. Esta observación motiva la siguiente definición:

Definición 1.54 (probabilidad condicional) *Sea* $(\Omega, \Im, P)$ *un espacio de probabilidad. Si* $A, B \in \Im$ *con* $P(A) > 0$, *entonces se define la probabilidad del evento* B *bajo la condición* A *como sigue:*

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

Ejemplo 1.55 *Se lanzan dos dados corrientes una vez. La probabilidad de que al menos uno de los resultados sea* 6, *dado que los resultados obtenidos son diferentes, es igual a* $\frac{1}{3}$, *como lo muestra el siguiente razonamiento: Sea* B *el evento "los resultados son diferentes" y* A *el evento "por lo menos uno de los resultados es* 6*". Es claro que*

$$\begin{aligned} A &= \{(a,6) : a \in \{1,2,...,6\}\} \cup \{(6,b) : b \in \{1,2,...,6\}\} \quad y \\ B &= \{(a,b) : a \in \{1,2,...,6\},\ a \neq b\}, \end{aligned}$$

entonces:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{10}{36}}{\frac{30}{36}} = \frac{1}{3}. \quad \blacktriangle$$

Ejemplo 1.56 *Una urna contiene* 12 *bolas de las cuales* 8 *son blancas. Se extrae una muestra de tamaño* 4 *sin reemplazo. Entonces la probabilidad de que la primera y la tercera bola extraídas sean blancas dado que la muestra contiene exactamente tres bolas blancas es igual a* $\frac{1}{2}$. *En efecto, si se supone que las bolas están numeradas del* 1 *al* 12, *entonces*

$$\Omega = \{(a_1, a_2, a_3, a_4) : a_i \in \{1, 2, ..., 12\}, \quad a_i \neq a_j \text{ para todo } i \neq j\}.$$

Sean

$$A := \text{“exactamente tres de las bolas extraídas son blancas”}$$
$$B := \text{“la primera y la tercera bola extraídas son blancas”}$$

Es claro que:

$$P(B \mid A) = \frac{P(A \cap B)}{P(A)} = \frac{\frac{n(A\cap B)}{n(\Omega)}}{\frac{n(A)}{n(\Omega)}} = \frac{n(A \cap B)}{n(A)} = \frac{\binom{2}{1}8.6.7.4}{\binom{4}{3}4.8.7.6} = \frac{1}{2}. \quad \blacktriangle$$

Ejemplo 1.57 *Sea* $(\Omega, \Im, P)$ *un espacio de probabilidad con*

$$\Omega = \{a, b, c, d, e, f\}, \; \Im = \wp(\Omega)$$

y

ω	a	b	c	d	e	f
$P(\omega)$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{3}{16}$	$\frac{1}{4}$	$\frac{5}{16}$

.

Sean $A = \{a, c, e\}$, $B = \{c, d, e, f\}$ *y* $C = \{b, c, f\}$. *Entonces*

$$\begin{aligned} P(A \mid B^c \cup C) &= \frac{P(A \cap (B^c \cup C))}{P(B^c \cup C)} \\ &= \frac{P(\{a, c\})}{P(\{a, b, c, f\})} \\ &= \frac{1}{3}. \quad \blacktriangle \end{aligned}$$

El siguiente teorema da las principales propiedades de la probabilidad condicional:

Teorema 1.58 (medida de probabilidad condicional) *Sea* $(\Omega, \Im, P)$ *un espacio de probabilidad y* $A \in \Im$ *con* $P(A) > 0$. *Entonces:*

1. $P(\cdot \mid A)$ *es una medida de probabilidad sobre* Ω, *que está concentrada en* A, *esto es* $P(A \mid A) = 1$.

2. *Si* $A \cap B = \Phi$ *entonces* $P(B \mid A) = 0$.

3. $P(B \cap C \mid A) = P(B \mid A \cap C)\, P(C \mid A)$, *si* $P(A \cap C) > 0$.

4. *Si* $A_1, A_2, ..., A_n \in \Im$ *con* $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$ *entonces*

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

Demostración.

1. Se deben verificar las tres condiciones que caracterizan a una medida de probabilidad:

   i) Es claro que $P(B \mid A) \geq 0$ para todo $B \in \Im$.

   ii) $P(\Omega \mid A) = \frac{P(A \cap \Omega)}{P(A)} = \frac{P(A)}{P(A)} = 1$. Análogamente se verifica que $P(A \mid A) = 1$.

   iii) Sean $A_1, A_2, ...$ elementos de $\Im$ disyuntos dos a dos. Entonces:

$$\begin{aligned} P\left(\bigcup_{i=1}^{\infty} A_i \mid A\right) &= \frac{P\left(A \cap \left(\bigcup_{i=1}^{\infty} A_i\right)\right)}{P(A)} \\ &= \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap A)\right)}{P(A)} \\ &= \frac{\sum_{i=1}^{\infty} P(A_i \cap A)}{P(A)} \\ &= \sum_{i=1}^{\infty} P(A_i \mid A) \end{aligned}$$

2. Como ejercicio.

3.

$$\begin{aligned} P(B \cap C \mid A) &= \frac{P(B \cap C \cap A)}{P(A)} \\ &= \frac{P(B \cap C \cap A)}{P(C \cap A)} \times \frac{P(C \cap A)}{P(A)} \\ &= P(B \mid A \cap C)P(C \mid A). \end{aligned}$$

4. Ejercicio.

■

Ejemplo 1.59 *Una urna contiene* 12 *bolas de las cuales* 4 *son negras y* 8 *son blancas. Se juega el siguiente juego: Se extrae una bola al azar, se anota su color y se devuelve a la urna junto con dos bolas adicionales del mismo color. Calcular la probabilidad de que en las tres primeras repeticiones del juego sean extraídas bolas negras.*

Solución: Para $i = 1, 2, 3$ se define:

$A_i :=$ "en la i-ésima repetición del juego fue extraída una bola negra".

Es claro que:

$$\begin{aligned} P(A_1 \cap A_2 \cap A_3) &= P(A_3 \mid A_2 \cap A_1)P(A_2 \mid A_1)P(A_1) \\ &= \frac{8}{16} \times \frac{6}{14} \times \frac{4}{12} \\ &= \frac{1}{14}. \quad \blacktriangle \end{aligned}$$

Ejemplo 1.60 *Tres adolescentes desean entrar a una película para mayores de 18 años. El encargado de la taquilla para verificar la edad, les pide su documento de identidad. Los jóvenes lo entregan y el dependiente, luego de revisarlos y negarles la entrada, se los devuelve al azar. Calcular la probabilidad de que ninguno de los jóvenes reciba su propio documento.*

Solución: Sean

$$\begin{aligned} A &:= \text{"ningún jóven recibe su propio documento"} \\ B_i &:= \text{"el i-ésimo jóven recibe su propio documento"} \end{aligned}$$

Es claro que la probabilidad buscada es :

$$\begin{aligned} P(A) &= P(B_1^c \cap B_2^c \cap B_3^c) \\ &= 1 - P(B_1 \cup B_2 \cup B_3) \\ &= 1 - P(B_1) - P(B_2) - P(B_3) + P(B_1 \cap B_2) + \\ &\quad P(B_2 \cap B_3) + P(B_1 \cap B_3) - P(B_1 \cap B_2 \cap B_3). \end{aligned}$$

Como hay tres casos posibles y uno favorable, se tiene que:

$$P(B_i) = \frac{1}{3} \text{ para } i = 1, 2, 3.$$

Por otra parte

$$P(B_i \cap B_j) = P(B_i)P(B_j \mid B_i) = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$$

ya que al haber entregado al $i-$ésimo joven correctamente su documento de identidad, para el $j-$ésimo queda sólo una opción favorable entre dos posibles. Análogamente se verifica que

$$\begin{aligned} P(B_1 \cap B_2 \cap B_3) &= P(B_1)P(B_2 \mid B_1)P(B_3 \mid B_1 \cap B_2) \\ &= \frac{1}{3} \times \frac{1}{2} \times 1 = \frac{1}{6}. \end{aligned}$$

Por lo tanto

$$P(A) \simeq 0.333. \quad \blacktriangle$$

Los siguientes dos resultados son de suma importancia en las aplicaciones:

Teorema 1.61 (teorema de probabilidad total) *Sea $A_1, A_2, ...$ una partición finita o numerable de Ω, es decir, $A_i \cap A_j = \Phi$ para todo $i \neq j$ y $\bigcup_{i=1}^{\infty} A_i = \Omega$, tal que $P(A_i) > 0$ para todo i. Entonces para cualquier $B \in \Im$ se satisface:*

$$P(B) = \sum_n P(B \mid A_n)P(A_n).$$

Demostración.

$$\begin{aligned} P(B) &= P(B \cap \Omega) \\ &= P(\bigcup_n (B \cap A_n)) \\ &= \sum_n P(B \cap A_n) \\ &= \sum_n P(B \mid A_n)P(A_n) \end{aligned}$$

con lo cual queda probado el teorema. ■

Como corolario del teorema anterior se obtiene un resultado conocido como regla de Bayes, el cual constituye la base de una teoría estadística muy importante llamada Bayesiana.

Corolario 1.62 (regla de Bayes) *Sea $A_1, A_2, \cdots$ una partición finita o numerable de Ω con $P(A_i) > 0$ para todo i, entonces se satisface para todo $B \in \Im$ con $P(B) > 0$ lo siguiente:*

$$P(A_i \mid B) = \frac{P(A_i)P(B \mid A_i)}{\sum\limits_j P(B \mid A_j)P(A_j)}; \text{ para todo } i.$$

Demostración.

$$\begin{aligned} P(A_i \mid B) &= \frac{P(A_i \cap B)}{P(B)} \\ &= \frac{P(A_i)P(B \mid A_i)}{P(B)} \\ &= \frac{P(A_i)P(B \mid A_i)}{\sum\limits_j P(B \mid A_j)P(A_j)} \end{aligned}$$

■

Para dar una interpretación a la regla de Bayes, supóngase que los eventos $A_1, A_2, ...$ son todas las posibles causas, mutuamente excluyentes, de un evento B. Bajo el supuesto que el evento B ha sido observado, la fórmula de Bayes permite conocer cuál de estas causas es la más probable de haber producido el evento B.

Ejemplo 1.63 *El señor Rodríguez sabe que hay un 40 % de posibilidad de que la empresa en la cual labora abra una sucursal en Montevideo (Uruguay). Si lo hace, la probabilidad de que él sea nombrado gerente de dicha sucursal es de un 80 %. Si no lo hace, la probabilidad de que el Sr. Rodríguez sea nombrado gerente en otra sucursal es de tan sólo 10 %. Se desea calcular la probabilidad de que el Sr. Rodríguez sea nombrado gerente de una sucursal de su empresa.*

Sean:

$$\begin{aligned} G &:= \text{“el Sr. Rodríguez es nombrado gerente”} \\ A &:= \text{“la empresa abre una sucursal en Montevideo”} \end{aligned}$$

entonces,

$$\begin{aligned} P(G) &= P(G \mid A)P(A) + P(G \mid A^c)P(A^c) \\ &= 0.8 \times 0.4 + 0.10 \times 0.60 \\ &= 0.38. \quad \blacktriangle \end{aligned}$$

Ejemplo 1.64 *Para el ejemplo anterior, si se sabe que el Sr. Rodríguez fue nombrado gerente de una sucursal de su empresa ¿cuál es la probabilidad de que la empresa haya abierto una sucursal en Montevideo?*

Es claro de la regla de Bayes que:

$$\begin{aligned} P(A \mid G) &= \frac{P(A)P(G \mid A)}{P(G)} \\ &= \frac{0.4 \times 0.8}{0.38} \\ &= 0.84211. \quad \blacktriangle \end{aligned}$$

Ejemplo 1.65 *La cuarta parte de una población está vacunada contra una enfermedad contagiosa. En el transcurso de una epidemia debida a tal enfermedad, se observa que de cada cinco enfermos sólo uno está vacunado. Se sabe además que de cada doce vacunados sólo uno está enfermo. Se quiere calcular la probabilidad de que un no vacunado esté enfermo.*

Sean:

$$\begin{aligned} V &:= \text{“La persona está vacunada”} \\ E &:= \text{“La persona está enferma”} \end{aligned}$$

De acuerdo a la información suministrada se tiene que:

$$\begin{aligned} P(V \mid E) &= \frac{1}{5} \\ P(E \mid V) &= \frac{1}{12} \\ P(V) &= \frac{1}{4}. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} x &:= P(E \mid V^c) \\ &= \frac{P(V^c \mid E)P(E)}{P(V^c)} \\ &= \frac{\frac{4}{5}(\frac{1}{12}\cdot\frac{1}{4} + \frac{3}{4}x)}{\frac{3}{4}}, \end{aligned}$$

de donde se obtiene

$$x = \frac{1}{9}. \quad \blacktriangle$$

Definición 1.66 (distribución a priori y a posteriori)
Sea $A_1, A_2, \ldots$ *una partición finita o numerable de* Ω *con* $P(A_i) > 0$ *para todo* i. *Si* B *es un elemento de* $\Im$ *con* $P(B) > 0$, *entonces* $(P(A_n))_n$ *se llama distribución "a priori", esto es, antes de que suceda* B, *y* $(P(A_n \mid B))_n$ *se llama distribución "a posteriori", es decir, después de que sucede* B.

Ejemplo 1.67 *En una ciudad se llevan a cabo pruebas para detectar cierta enfermedad. Supóngase que el* $1\,\%$ *de las personas sanas son registradas como enfermas, que el* $0.1\,\%$ *de la población está realmente enferma y que el* $90\,\%$ *de los enfermos son reportados como tales. Se desea calcular la probabilidad de que una persona, seleccionada al azar y reportada como enferma, esté realmente enferma.*

Si se definen los eventos:

$$\begin{aligned} E &:= \text{"la persona está realmente enferma"} \\ T &:= \text{"la persona es reportada como enferma"} \end{aligned}$$

enonces de la información suministrada se sabe que:

$$\begin{aligned} P(E) &= 0.001 \\ P(T \mid E^c) &= 0.01 \\ P(T \mid E) &= 0.9. \end{aligned}$$

Por lo tanto,

$$\begin{aligned} P(E \mid T) &= \frac{P(T \mid E)P(E)}{P(T \mid E^c)P(E^c) + P(T \mid E)P(E)} \\ &= \frac{0.9 \times 0.001}{0.01 \times 0.999 + 0.9 \times 0.001} \\ &= 8.2645 \times 10^{-2} \approx 0.083. \end{aligned}$$

En este caso:

$$\begin{aligned} \textit{Distribución a priori} &= (0.001, 0.999) \\ \textit{Distribución a posteriori} &= (0.083, 0.917). \quad \blacktriangle \end{aligned}$$

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

$$P(A \mid B) = P(A) \tag{1.1}$$

En este caso se dice que el evento A es "independiente" del evento B. La "definición" dada en (1.1) tiene como inconveniente que sólo es válida si $P(B) > 0$. Para evitar esta dependencia de la probabilidad del evento B, se asume como definición de independencia la siguiente:

Definición 1.68 (eventos independientes) *Dos eventos A y B son independientes, si y sólo si,*

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

En caso contrario se dice que los eventos son dependientes.

Ejemplo 1.69 *Supóngase que se lanza un dado corriente dos veces consecutivas. Sean:*

$A :=$ *"la suma de los resultados obtenidos es un número par"*

$B :=$ *"el resultado del segundo lanzamiento es par"*

En este caso:

$$P(A) = P(B) = \frac{1}{2}$$

y además $P(A \cap B) = \frac{1}{4}$. Por lo tanto, los eventos son independientes. ▲

Ejemplo 1.70 *Se carga un dado de manera que la probabilidad de obtener un número par es igual a $\frac{2}{5}$. Sean A y B como en el ejemplo anterior. En este caso se tiene que:*

$$P(A) = \frac{13}{25}$$

$$P(B) = \frac{2}{5}$$

$$P(A \cap B) = \frac{4}{25}$$

por lo tanto, A y B no son independientes. ▲

Nota 1.71 *Un error, que a menudo se comente, es asegurar que los eventos son independientes si son mutuamente excluyentes. Observe que esto no es cierto. Por ejemplo, si se lanza una moneda corriente una vez y se consideran los eventos:*

$A :=$ *"el resultado obtenido es cara"*

$B :=$ *"el resultado obtenido es sello"* *entonces, es claro que, A y B son mutuamente excluyentes. Sin embargo no son independientes, pues*

$$0 = P(A \cap B) \neq \frac{1}{4} = P(A)P(B).$$

Teorema 1.72 *Sean A y B eventos independientes. Entonces:*

1. *A y B^c son independientes (y por simetría A^c y B son independientes).*
2. *A^c y B^c son independientes.*

Demostración.

1.

$$\begin{aligned} P(A) &= P(A \cap B^c) + P(A \cap B) \\ &= P(A \cap B^c) + P(A)P(B), \end{aligned}$$

por lo tanto,

$$P(A \cap B^c) = P(A)[1 - P(B)] = P(A)P(B^c).$$

2.

$$\begin{aligned} P(A^c \cap B^c) &= 1 - P(A \cup B) \\ &= 1 - P(A) - P(B) + P(A \cap B) \\ &= 1 - P(A) - P(B) - P(A)P(B) \\ &= [1 - P(A)][1 - P(B)] \\ &= P(A^c)P(B^c). \end{aligned}$$

■

En muchos casos se requiere analizar la independencia de dos o más eventos, por ello es necesario generalizar el concepto de independencia dado.

Definición 1.73 (familia independiente) *Una familia $\{A_i : i \in I\}$ de eventos se dice independiente si:*

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

para todo subconjunto finito J de I.

Definición 1.74 (familia independiente dos a dos) *Una familia* $\{A_i : i \in I\}$ *de eventos se dice independiente dos a dos si:*

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

La independencia dos a dos de una familia no implica la independencia de ésta, como lo muestra el siguiente ejemplo:

Ejemplo 1.75 *Se lanza un dado corriente dos veces consecutivas. Sean A, B y C los eventos definidos como sigue:*

$$\begin{aligned} A &:= \text{“en el primer lanzamiento se obtuvo } 2\text{ ”} \\ B &:= \text{“en el segundo lanzamiento se obtuvo } 5\text{ ”} \\ C &:= \text{“ la suma de los resultados es } 7\text{ ”} \end{aligned}$$

Es claro que:

$$\begin{aligned} P(A) = P(B) = P(C) &= \frac{1}{6} \\ P(A \cap B) = P(A \cap C) = P(B \cap C) &= \frac{1}{36} \\ P(A \cap B \cap C) &= \frac{1}{36} \end{aligned}$$

Por lo tanto, los eventos son independientes dos a dos pero no son independientes, pues:

$$P(A \cap B \cap C) \neq P(A)P(B)P(C). \quad \blacktriangle$$

1.3. Probabilidad Geométrica

Sea $(\Omega, \Im, P)$ un espacio de probabilidad y supóngase que sobre $(\Omega, \Im)$ hay definida una medida geométrica m tal como longitud, área o volumen. Se define la probabilidad del evento A como sigue:

$$P(A) := \frac{m(A)}{m(\Omega)}.$$

A continuación se darán algunos ejemplos de cálculo de probabilidades geométricas.

Ejemplo 1.76 *María Victoria y Carlos acordaron encontrarse en el centro de la ciudad entre las 12m y la 1pm. Los dos llegan en cualquier momento* *dentro de ese intervalo de tiempo. Suponiendo que los tiempos de llegada, de cada uno de ellos, son independientes, calcular:*

i) La probabilidad de que Carlos y María Victoria se encuentren si cada uno espera a lo más diez minutos.

ii) La probabilidad de que Carlos y María Victoria se encuentren, si María Victoria sólo espera cinco minutos y Carlos sólo veinte.

Solución

i. Sean X y Y las variables aleatorias definidas como sigue:

$$X := \text{“Hora de llegada de María Victoria”}$$
$$Y := \text{“Hora de llegada de Carlos”}$$

El espacio muestral, en este caso, está dado por:

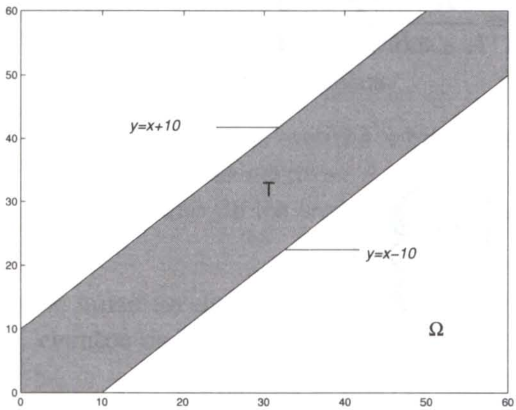
$$\Omega := \{(x,y) \in \mathbb{R}^2 : 0 \leq x \leq 60,\ 0 \leq y \leq 60\}$$

El evento al cual le queremos calcular la probabilidad es:

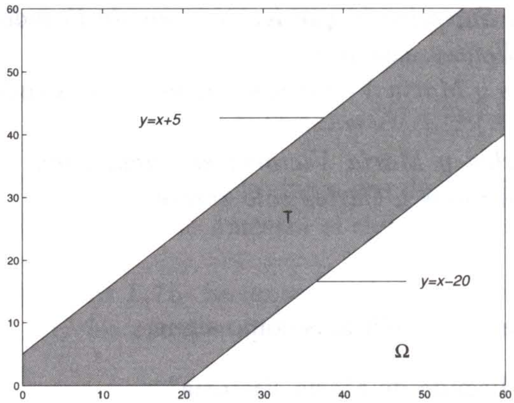
$$T = \{(x,y) \in \Omega : |x-y| \leq 10\}$$

Por lo tanto (ver figura):

$$P(T) = \frac{\text{área de } T}{\text{área de } \Omega} = \frac{11}{36}$$



ii. El conjunto de puntos T de Ω que representan los tiempos de llegada de Carlos y María Victoria que permiten un encuentro de los dos, está representado en la figura:



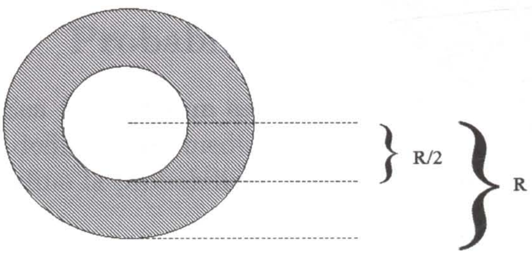
Por lo tanto la probabilidad pedida es igual a:

$$P(T) = \frac{103}{288}. \quad \blacktriangle$$

Ejemplo 1.77 *Se escoge al azar un punto de un círculo de radio R. Calcular la probabilidad de que el punto esté más cerca de la circunferencia que del centro del círculo.*

Solución: Sea $\Omega :=$ conjunto de todos los puntos en el interior del círculo de radio R. Es claro que el evento al cual se le quiere hallar su probabilidad es: $A :=$ conjunto de puntos en Ω, que pertenecen al anillo que se muestra en la figura. Por lo tanto la probabilidad pedida es igual a :

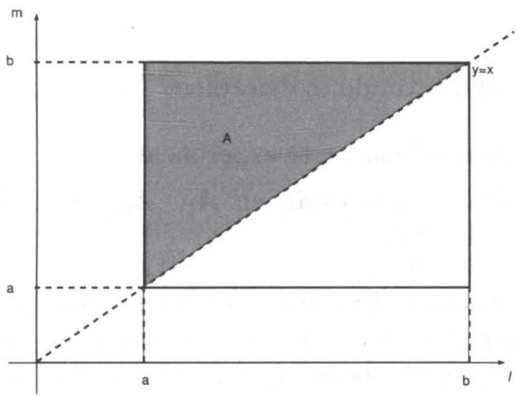
$$p = \frac{\pi R^2 - \frac{\pi R^2}{4}}{\pi R^2} = \frac{3}{4}$$



Ejemplo 1.78 *Sobre un segmento de recta $\overline{ab}$ se marcan aleatoriamente dos puntos l y m. Calcular la probabilidad de que l esté más cerca de a que m.*

Solución: En este caso el espacio muestral es $\Omega = [a, b] \times [a, b]$ y el evento al cual se le quiere calcular su probabilidad es $A = \{(l, m) \in \Omega : l < m\}$. De la figura se obtiene que la probabilidad pedida p es:

$$p = \frac{\text{área de } A}{\text{área de } \Omega} = \frac{1}{2}.$$



1.4. Ejercicios

1. Se lanza una moneda corriente tres veces consecutivas. Sean:

$$A := \text{“En el primer lanzamiento se obtiene cara”}$$
$$B := \text{“En el tercer lanzamiento se obtiene sello”}$$

Describir en palabras los eventos $A \cap B$, $A \cup B$, A^c, $A^c \cap B^c$, $A \cap B^c$ y determinar sus elementos.

2. Sean A, B y C tres eventos arbitrarios. Describir en términos de A, B y C los siguientes eventos: A y B pero no C; todos los tres; sólo A; por lo menos uno de los tres; a lo más uno de los tres; a lo más dos de los tres.

3. Se lanza un dado corriente dos veces consecutivas. Sean A, B y C los eventos dados por:

$$A := \text{“el primer resultado obtenido es un número par”}$$
$$B := \text{“la suma de los resultados obtenidos es menor que 7”}$$
$$C := \text{“el segundo resultado obtenido es un número primo”}$$

Determinar los elementos que pertenecen a los siguiente eventos:

a) $A \cap B \cap C$

b) $B \cup (A \cap C^c)$

c) $(A \cap C) \cap [(A \cup B)^c]$

4. Un experimento aleatorio consiste en extraer tres bombillos al azar y clasificarlos como defectuoso "D" o no defectuoso "N". Considere los eventos

$$A_i := \text{"el } i\text{-ésimo bombillo extraído es defectuoso"}, \ i = 1, 2, 3$$

a) Describir el espacio muestral para este experimento.

b) Listar todos los resultados que están en A_1, A_2, $A_1 \cup A_3$, $A_1^c \cap A_2^c \cap A_3$ y $(A_1 \cup A_2^c) \cap A_3$.

5. Un trabajador fabrica n artículos. El evento "el i-ésimo artículo es defectuoso" se denotará por A_i con $i = 1, 2, \cdots, n$. Describir los siguientes eventos usando los A_i y las operaciones usuales entre eventos:

a) $B :=$ "Por lo menos un artículo es defectuoso".

b) $C :=$ "Ninguno de los n artículos es defectuoso".

c) $D :=$ "Exactamente un artículo es defectuoso".

d) $E :=$ "A lo más un artículo es defectuoso".

6. Sean A, B y C eventos arbitrarios. Describir en términos de A, B y C los siguientes eventos:

a) $E_1 :=$ "Por lo menos uno de los eventos A, B, C ocurre".

b) $E_2 :=$ "Exactamente dos de los eventos A, B, C ocurren".

c) $E_3 :=$ "Por lo menos dos de los eventos A, B, C ocurren".

d) $E_4 :=$ "Al menos uno de los eventos A, B, C ocurre".

7. Un total de 35 % de los estudiantes de la Universidad Nacional de Colombia están inscritos en un curso de inglés, 7 % están inscritos en un curso de alemán y 2 % están inscritos en cursos de inglés y alemán. ¿Qué porcentaje de los estudiantes están inscritos en cursos de inglés pero no de alemán? ¿Qué porcentaje de los estudiantes no está inscrito en inglés ni en alemán?

8. Sea $\Omega \neq \Phi$ y $\Im$ una σ-álgebra en Ω. Demostrar:

a) Si $A_1, A_2, \ldots \in \Im$ entonces $\bigcap_{i=1}^{\infty} A_i \in \Im$.

b) Si $A, B \in \Im$ entonces $A \cup B$, $A \cap B$, $A - B$ y $A \bigtriangleup B$ pertenecen a $\Im$.

9.

a) Sean Ω y $\widetilde{\Omega}$ conjuntos no vacios, $\widetilde{\Im}$ una σ-álgebra en $\widetilde{\Omega}$ y $T : \Omega \to \widetilde{\Omega}$ una función. Demostrar que:

$$T^{-1}(\widetilde{\Im}) := \{T^{-1}(\widetilde{A}) : \widetilde{A} \in \widetilde{\Im}\}$$

es una σ-álgebra en Ω.

b) Determinar la $\sigma-$álgebra sobre $\Omega = \{1, 2, 3\}$ generada por $\{\{2\}, \{3\}\}$.

c) Sean $\mathcal{C}$ y $\mathcal{D}$ dos familias de subconjuntos de un conjunto no vacio Ω con $\mathcal{C} \subseteq \mathcal{D}$. ¿Es $\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{D})$? Explicar.

d) Sea $\Omega = \{1, 2\}$, $\Im = \{\Phi, \Omega, \{1\}, \{2\}\}$ y μ definida sobre $\Im$ por:

$$\begin{aligned}\mu(\Phi) &= 0\\ \mu(\Omega) &= 1\\ \mu(\{1\}) &= \frac{1}{3}\\ \mu(\{2\}) &= \frac{2}{3}\end{aligned}$$

¿Es $(\Omega, \Im, \mu)$ un espacio de probabilidad? Explicar.

10. Sean $\Omega = \{a, b, c, d\}$, $\Im = \{\Phi, \Omega, \{a\}, \{b, c\}, \{d\}, \{a, b, c\}, \{b, c, d\}, \{a, d\}\}$ y P una aplicación de $\Im$ en $[0, 1]$ con $P(\{a\}) = \frac{2}{7}$, $P(\{b, c\}) = \frac{4}{9}$ y $P(\{d\}) = \alpha$.

    a) Determinar el valor de α de tal manera que P sea una medida de probabilidad sobre $(\Omega, \Im)$.

    b) Calcular $P(\{a, b, c\})$, $P(\{b, c, d\})$ y $P(\{a, d\})$.

11. Sean $(P_n)_{n \in \mathbb{N}}$ una sucesión de medidas de probabilidad definidas sobre el espacio medible $(\Omega, \Im)$, $(\alpha_n)_{n \in \mathbb{N}}$ una sucesión de números reales no negativos tales que $\sum_{n=1}^{\infty} \alpha_n = 1$ y $P : \Im \to \mathbb{R}$ definida por:

$$P(A) := \sum_{n \in \mathbb{N}} \alpha_n P_n(A) \qquad \text{para todo } A \in \Im$$

Demostrar que P es una medida de probabilidad definida sobre $(\Omega, \Im)$.

12. Determinar si las siguientes proposiciones son falsas o verdaderas. Justificar brevemente la respuesta:

   a) Si $P(A) = 0$ entonces $A = \Phi$

   b) Si $P(A) = P(B) = 0$ entonces $P(A \cup B) = 0$.

   c) Si $P(A) = \frac{1}{2}$ y $P(B) = \frac{1}{3}$ entonces $\frac{1}{2} \leq P(A \cup B) \leq \frac{5}{6}$

   d) Si $P(A) = P(B) = p$ entonces $P(A \cap B) \leq p^2$

   e) $P(A \bigtriangleup B) = P(A) + P(B) - 2P(A \cap B)$

   f) Si $P(A) = 0.5$, $P(B) = 0.4$ y $P(A \cup B) = 0.8$ entonces

$$P(A^c \cap B) = 0.1$$

13. Sean A y B dos eventos con $P(A) = \frac{1}{2}$ y $P(B^c) = \frac{1}{4}$ ¿Pueden ser A y B mutuamente excluyentes? Explicar.

14. Se carga un dado de manera que los números pares tienen el doble de probabilidad de salir que los impares. ¿A qué es igual la probabilidad de obtener un número par?, ¿un número primo?, ¿un número primo impar?

15. De los 100 estudiantes de la carrera de Filología y Lenguas clásicas del Departamento de Idiomas de una universidad, se tiene que 28 estudiantes asisten a clases de latín, 26 a clases de griego y 16 a clases de hebreo. Hay 12 estudiantes que asisten tanto a las clases de latín como a las clases de griego, 4 están asistiendo a las de latín y hebreo y 6 están en las de griego y hebreo. Además se tiene que 2 estudiantes asisten a los tres cursos mencionados.

   a) Si se escoge un estudiante de Filología y Lenguas clásicas, aleatoriamente, ¿cuál es la probabilidad de él o ella asista sólo a clase de hebreo?

   b) Si se escoge un estudiante de Filología y Lenguas clásicas aleatoriamente, ¿cuál es la probabilidad de que él o ella asista a clases de griego y de hebreo pero no de latín?

   c) Si dos estudiantes de Filología y Lenguas clásicas son escogidos al azar, ¿cuál es la probabilidad de que al menos uno de ellos asista a uno de los cursos?

16. Se contrató una empresa para realizar una encuesta entre los 1000 suscriptores de una revista. Los datos suministrados en el informe presentado fueron los siguientes: 550 suscriptores son profesionales, 630 son casados, 650 son mayores de 35 años, 127 son profesionales y mayores de 35 años, 218 son casados y mayores de 35 años, 152 son profesionales y casados y 100 son casados, profesionales y ma-yores de 35 años. ¿Son los datos presentados en este reporte correctos? Explicar la respuesta.

17. Se lanza una moneda corriente n veces. Sea

$$A_k := \text{“se obtiene cara por primera vez en el } k - \text{ésimo lanzamiento”}$$

donde $k = 1, 2, \cdots, n$. ¿A qué es igual $P(A_k)$?

18. Se distribuyen 10 bolas distinguibles en 7 urnas distinguibles. ¿A qué es igual la probabilidad de que todas las urnas tengan por lo menos una bola?, ¿Cuál es la probabilidad de que exactamente dos de las urnas queden vacías?

19. Un grupo de 40 estudiantes está conformado por 20 hombres y 20 mujeres. Si se divide al grupo en dos grupos iguales. ¿Cuál es la probabilidad de que cada grupo tenga el mismo número de hombres que de mujeres?

20. Los coeficientes a, b y c de la ecuación cuadrática $ax^2 + bx + c = 0$ son determinados lanzando un dado corriente tres veces consecutivas. ¿A qué es igual la probabilidad de que las dos raíces de la ecuación sean reales?. ¿A qué es igual la probabilidad de que las dos raíces sean complejas?

21. Se tienen dos urnas A y B. La urna A contiene 3 bolas rojas y 4 negras y la urna B contiene 5 bolas rojas y 7 negras. Se selecciona al azar una bola de cada urna. ¿A qué es igual la probabilidad de que las dos bolas seleccionadas sean del mismo color?

22. Una urna contiene 3 bolas rojas y 7 negras. Los jugadores A y B extraen consecutivamente una bola de la urna, hasta que una bola roja es seleccionada.¿A qué es igual la probabilidad de que el jugador A extraiga la bola roja? Suponga que la extracción se hace sin sustitución y que el jugador A inicia el juego.

23. En un lago hay 200 peces ornamentales. Se capturan 50 de ellos, se marcan y se regresan al lago. Días más tarde, se capturan 40 peces, ¿cuál es la probabilidad de que 20 de esos 40 peces estén marcados?

24. Se ordenan 5 hombres y 5 mujeres de acuerdo con sus calificaciones en un examen. Suponga que no hay dos calificaciones iguales y que los 10! órdenes posibles son igualmente probables. ¿A qué es igual la probabilidad de que la posición más alta alcanzada por un hombre sea la cuarta?

25. Un estudiante está tan entusiasmado con su curso de probabilidad que decide planear sus actividades de fin de semana de acuerdo con el resultado que obtenga al lanzar un dado corriente una vez. Si el resultado del lanzamiento es menor o igual a 4 el entonces sale de rumba con sus amigos, si el resultado es 5 se queda en casa estudiando probabilidad y si el resultado es 6 invita a su pareja a cine.

    Para hacerse a una idea de cómo será su actividad de fin de semana, el estudiante divide el año en trece periodos de cuatro semanas cada uno y se interesa por las probabilidades de los siguientes eventos:

    a) estudiar probabilidad por lo menos una vez.
    b) ir dos veces a cine.
    c) salir cuatro veces de rumba con sus amigos.
    d) realizar cada actividad por lo menos una vez.

    ¿Puede usted ayudarlo a calcular estas probabilidades?

26. Para iluminar una escalera han sido colocadas 7 lámparas y se han literado con las letras $A, B, \cdots, G$. Para garantizar la iluminación de la escalera deben funcionar las lámparas A o B y las lámparas F o G, o alguna de las lámparas C, D o E. La probabilidad de estar fuera de servicio es igual a $\frac{1}{5}$ para todas las lámparas.

    a) ¿A qué es igual la probabilidad de que la iluminación de la escalera esté garantizada?
    b) ¿Cómo cambia la probabilidad pedida en **a**., si la lámpara D no se usa?

27. ¿A qué es igual la probabilidad de que entre 25 personas por lo menos dos tengan cumpleaños el mismo día? Supóngase que cada año tiene

365 días y que todos los días tienen la misma probabilidad de ser un día de cumpleaños.

28. A una fiesta de navidad asisten n personas cada una de las cuales lleva un regalo. Los regalos se introducen en una bolsa y se mezclan homogéneamente, luego cada persona extrae al azar un regalo de la bolsa. ¿Cuál es la probabilidad de que ninguna persona extraiga su propio regalo?

29. En el juego del bridge se reparte la baraja completa de 52 cartas entre 4 jugadores.

    a) ¿A qué es igual la probabilidad de que uno de los jugadores reciba todas las trece picas?

    b) ¿A qué es igual la probabilidad de que cada jugador reciba un as?

30. Una urna contiene 15 bolas de las cuales 9 son rojas y 6 son blancas. Se juega el siguiente juego: se extrae una bola al azar de la urna, se anota su color y se devuelve a la urna junto con dos bolas adicionales del otro color. ¿A qué es igual la probabilidad de que en las primeras tres repeticiones del juego se hayan extraído bolas blancas?

31. Calcular la probabilidad de que en un grupo de trece cartas, de una baraja normal de 52 cartas, haya exactamente dos reyes y un as. ¿A qué es igual la probabilidad de que en dicho grupo haya exactamente un as dado que que el grupo contiene exactamente dos reyes?

32. Sean A y B eventos tales que $P(A) = 0.5$, $P(B) = 0.3$ y $P(A \cap B) = 0.1$. Calcular $P(A \mid B)$, $P(A \mid B^c)$, $P(A \mid A \cap B)$, $P(A^c \mid A \cup B)$ y $P(A \cap B \mid A \cup B)$.

33. Un estudiante de matemáticas tiene que presentar el mismo día un examen de probabilidad y uno de álgebra. Sean:

    $A :=$ "el estudiante reprueba el examen de probabilidad "

    $B :=$ "el estudiante reprueba el examen de álgebra"

    Si $P(A) = 0.4$, $P(B) = 0.3$ y $P(A \cap B) = 0.2$. ¿A qué es igual la probabilidad de que el estudiante apruebe el examen de álgebra dado que aprobó el de probabilidad?, ¿a qué es igual que el estudiante apruebe el examen de probabilidad dado que reprobó el de álgebra?

34. A través de una encuesta se pudo establecer lo siguiente:

 90 % de las familias bogotanas poseen radio y televisor.

 8 % de las familias bogotanas poseen radio pero no poseen televisor.

 2 % de las familias bogotanas poseen televisor pero no poseen radio.

 95 % de las familias bogotanas que poseen radio y televisor saben quién es el alcalde de la ciudad.

 80 % de las familias bogotanas que poseen radio pero no televisor saben quién es el alcalde de la ciudad.

 1 % de las familias bogotanas que poseen televisor pero no radio, no saben quién es el alcalde de la ciudad.

 Se escoge una familia bogotana al azar. Sean:

$$\begin{aligned} T &:= \text{“La familia tiene televisor”} \\ R &:= \text{“La familia tiene radio”} \\ B &:= \text{“La familia sabe quién es el alcalde de la ciudad”} \end{aligned}$$

 Calcular las siguientes probabilidades:

 a) $P(T \cup R)$

 b) $P(B \cap T)$

 c) $P(T \mid B)$

35. Un número aleatorio N de dados corrientes es lanzado. Sea A_i el evento de que $N = i$ y suponga que $P(A_i) = 2^{-i}$ con $i \geq 1$. Sea S la suma de los resultados. Determinar la probabilidad de que:

 a) $N = 2$ dado que $S = 4$.

 b) $S = 4$ dado que N es par.

 c) $N = 2$ dado que $S = 4$ y el resultado del primer dado es 1.

36. Supóngase que se tiene una población que se desarrolla de la siguiente manera: una partícula inicial, que constituye la 0-ésima generación, tiene 0, 1 o 2 hijas con probabilidades $\frac{1}{6}$, $\frac{2}{3}$ y $\frac{1}{6}$ respectivamente. Luego de reproducirse la partícula muere. Las hijas se reproducen independientemente unas de otras e independientemente de la historia

familiar, de la misma manera que la partícula original. La primera generación está compuesta por las hijas de la partícula inicial, la segunda por las nietas y así sucesivamente. Dado que en la segunda generación hay una partícula, ¿a qué es igual la probabilidad de que en la primera haya habido dos partículas?. ¿Cuál es la probabilidad de que en la segunda generación haya por lo menos una partícula?

37. El 5 % de las personas de una población sufren de tensión arterial alta. De las personas con tensión arterial alta se tiene que el 75 % son consumidores asiduos de bebidas alcohólicas, mientras que sólo el 50 % de las personas con tensión arterial alta consumen frecuentemente bebidas alcohólicas. ¿Cuál es el porcentaje de personas con tensión arterial alta que consumen asiduamente bebidas alcohólicas?

38. Se tienen dos urnas A y B. La urna A contiene 7 bolas rojas y 5 blancas y la urna B contiene 2 bolas rojas y 4 blancas. Se lanza un dado corriente. Si se obtiene un 3 o un 6 se toma una bola de B y se coloca en A y luego se toma una bola de A; si aparece otro número se toma una bola de A y se coloca en B y luego de extrae una bola de B. ¿A qué es igual la probabilidad de que las dos bolas extraídas sean rojas?

39. Supóngase que usted le pide a un compañero de curso que lo inscriba en la asignatura "Matemáticas sin esfuerzo" que se ofrecerá el próximo semestre en su universidad. Si su compañero olvida hacer la inscripción en los plazos estipulados por el Departamento de Matemáticas, la probabilidad de que usted consiga cupo en dicha asignatura es de sólo el 2 %, en tanto que si su compañero hace la inscripción a tiempo, la probabilidad de que usted consiga cupo es del 80 %. Usted está seguro, en un 95 %, de que su compañero hará la inscripción a tiempo. Si usted no obtuvo cupo, ¿a qué es igual la probabilidad de que su compañero haya olvidado inscribirlo a tiempo?

40. La probabilidad de que en un parto gemelar ambos bebés sean de género masculino es de 0.32, en tanto que la probabilidad de que sean ambos de género femenino es de 0.28. ¿A qué es igual la probabilidad de que en un parto gemelar, el segundo niño en nacer sea de género masculino dado que el primero en nacer es de género masculino? Supóngase que es tan probable que el primer niño en nacer sea de género femenino como de género masculino.

41. Un inversionista está pensando en comprar un número muy grande de acciones de una compañía. La cotización de acciones en la bolsa du-

rante los seis meses anteriores es de gran interés para el inversionista. Con base en esta información observa que la cotización se relaciona con el producto nacional bruto (PNB). Si el PNB aumenta, la probabilidad de que el precio de las acciones aumente es de 0.7. Si el PNB es el mismo, la probabilidad de que las acciones aumenten su valor es de 0.2, en tanto que si el PNB disminuye entonces la probabilidad de que las acciones aumenten su valor es de sólo 0.1. Si las probabilidades de que el PNB aumente, siga siendo el mismo o disminuya son respectivamente 0.5, 0.3 y 0.2, ¿a qué es igual la probabilidad de que las acciones aumenten su valor? Si las acciones aumentaron su valor, ¿cuál es la probabilidad de que el PNB haya aumentado?

42. En una urna hay ocho monedas. Dos de ellas tienen dos sellos, tres monedas son corrientes y tres están "cargadas" de tal manera que la probabilidad de obtener sello es igual a $\frac{3}{5}$. Se escoge una moneda al azar de la urna y se lanza. Si el resultado del lanzamiento es "cara", ¿cuál es la probabilidad de que haya sido lanzada una moneda corriente?

43. Sea $\Omega = \{1, 2, 3, 4\}$, $\Im = \wp(\Omega)$ y $P(\omega) = \frac{1}{4}$ para todo $\omega \in \Omega$. Sea $A = \{2, 3\}$. Determinar todos los elementos $B \in \Im$ tales que A y B sean independientes.

44. Sea $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\Im = \wp(\Omega)$ y $P(\omega) = \frac{1}{6}$ para todo $\omega \in \Omega$. Demostrar que si A y B son elementos de $\Im$ independientes y A tiene 3 elementos entonces B debe tener un número par de elementos.

45. Demostrar que si $A = B$ y A y B son independientes (esto es A es independiente de si mismo) entonces $P(A) = 0$ o $P(A) = 1$.

46. Sea A un evento. Demostrar que las siguientes condiciones son equivalentes:

    a) A y B son independientes para cualquier evento B.

    b) $P(A) = 0$ o $P(A) = 1$.

47. Sean A, B y C eventos independientes. Demostrar que A y $B \cup C$, A y $B \cap C$, A y $(B - C)$ son independientes.

48. Demostrar que los eventos $A_1, A_2, \cdots, A_n$ son independientes si y sólo si

$$P(B_1 \cap B_2 \cap \cdots \cap B_n) = P(B_1)P(B_2) \cdots P(B_n)$$

para toda posible elección de $B_1, B_2, \cdots, B_n$ con $B_i = A_i$ o $B_i = A_i^c$ para $i = 1, 2, \cdots, n$.

49. Una moneda corriente se lanza tres veces consecutivas. Considerar los siguientes eventos:

$$A := \text{“los resultados de los lanzamientos 1 y 2 son diferentes”}$$
$$B := \text{“los resultados de los lanzamientos 2 y 3 son diferentes”}$$
$$C := \text{“los resultados de los lanzamientos 1 y 3 son diferentes”}$$

   a) Verificar que $P(A) = P(A \mid B) = P(A \mid C)$ y que $P(A) \neq P(A \mid B \cap C)$.

   b) ¿Son A, B y C dos a dos independientes?, ¿son A, B y C independientes? Explicar las respuestas.

50. Sean A, B y C eventos independientes con $P(A) = P(B) = P(C) = \frac{1}{3}$. Calcular la probabilidad de que:

   a) Por lo menos uno de los eventos ocurra.

   b) Por lo menos dos de los eventos ocurran.

   c) Exactamente dos de los eventos ocurran.

51. Determinar la probabilidad de que entre siete personas:

   a) No hay dos que hayan nacido el mismo día de la semana (domingo, lunes, martes, etc.)

   b) Por lo menos dos nacieron el mismo día.

   c) Hay dos que nacieron el domingo y dos el martes.

52. Supóngase que se tiene una urna que contiene r bolas numeradas del 1 al N. Una muestra aleatoria, sin reemplazo, de tamaño n, es sacada de la urna y se anotan los números de las bolas extraídas. Luego de devolver las bolas extraídas a la urna, se extrae una segunda muestra, sin reemplazo, de tamaño m. Calcular la probabilidad de que las dos muestras tengan k bolas en común.

53. Una urna contiene bolas numeradas del 1 al n. Se selecciona una bola al azar.

   a) ¿A qué es igual la probabilidad de que el número que aparece en la bola sea divisible por 3 o 4 ?

b) ¿Qué ocurre con la probabilidad calculada en el literal anterior cuando $n \to \infty$?

Capítulo 2

Variables aleatorias y sus distribuciones

En un experimento aleatorio hay frecuentemente más interés por ciertos valores numéricos que se pueden deducir de los resultados de un experimento aleatorio que por los resultados en sí. Supóngase por ejemplo, que se lanza una moneda corriente seis veces consecutivas y que interesa conocer el número de caras obtenidas. En este caso se tiene que el espacio muestral es igual a:

$$\Omega = \{(a_1, a_2, a_3, a_4, a_5, a_6) : a_i \in \{C, S\}, i = 1, \cdots, 6\}.$$

Si se define $X :=$ "número de caras obtenidas", entonces se tiene que X es una aplicación de Ω en $\widetilde{\Omega}$ con $\widetilde{\Omega} = \{1, 2, \cdots, 6\}$. Así por ejemplo, se tiene que $X((C, C, S, C, C, S)) = 4$. La aplicación X es un ejemplo de variable aleatoria, concepto que se precisa a continuación.

Definición 2.1 (variable aleatoria) *Sean $(\Omega, \Im, P)$ un espacio de probabilidad y $(\widetilde{\Omega}, \widetilde{\Im})$ un espacio medible. Una $\Im - \widetilde{\Im}-$variable aleatoria es una aplicación $X : \Omega \longrightarrow \widetilde{\Omega}$ tal que, para todo $A \in \widetilde{\Im}$ se tiene que $X^{-1}(A) \in \Im$.*

Si $(\widetilde{\Omega}, \widetilde{\Im}) = (\mathbb{R}, \mathcal{B})$, entonces, se dice que X es una variable aleatoria real.

Ejemplo 2.2 *Sean $\Omega = \{1, 2, 3\}$, $\Im = \{\Phi, \{1\}, \{2, 3\}, \Omega\}$, P una medida de probabilidad arbitraria definida sobre $\Im$. Supóngase que $(\widetilde{\Omega}, \widetilde{\Im})$ es un espacio medible con $\widetilde{\Omega} = \{a, b\}$ y $\widetilde{\Im} = \wp(\widetilde{\Omega})$. La aplicación $X : \Omega \longrightarrow \widetilde{\Omega}$ dada por:*

$$X(\omega) = \begin{cases} a & si \quad \omega = 1 \\ b & si \quad \omega = 2 \text{ o } \omega = 3 \end{cases}$$

es una $\Im - \widetilde{\Im}$*−variable aleatoria, pues:*

$$X^{-1}(\Phi) = \Phi,\ X^{-1}(\{a\}) = \{1\},\ X^{-1}(\{b\}) = \{2,3\},\ X^{-1}(\widetilde{\Omega}) = \Omega,$$

en tanto que, la aplicación $Y : \Omega \longrightarrow \widetilde{\Omega}$ *dada por:*

$$Y(\omega) = \begin{cases} a & si \quad \omega = 2 \\ b & si \quad \omega = 1 \ o\ \omega = 3 \end{cases}$$

no es una $\Im - \widetilde{\Im}$*−variable aleatoria pues:*

$$Y^{-1}(\{a\}) = \{2\} \notin \Im.$$

Si se toma como σ*-álgebra en* Ω*, a* $\Im' = \{\Phi, \{2\}, \{1,3\}, \Omega\}$*, se tiene que* Y *es una* $\Im' - \widetilde{\Im}$*−variable aleatoria.* ▲

Ejemplo 2.3 *Sean* $(\Omega, \Im, P)$ *un espacio de probabilidad con* $\Im = \wp(\Omega)$ *y* $(\widetilde{\Omega}, \widetilde{\Im})$ *un espacio medible. Es claro que cualquier aplicación* $X : \Omega \longrightarrow \widetilde{\Omega}$ *es una* $\Im - \widetilde{\Im}$*−variable aleatoria.* ▲

Nota 2.4 *Sea* $(\Omega, \Im, P)$ *un espacio de probabilidad arbitrario. Puesto que la* σ*-álgebra de Borel* $\mathcal{B}$ *en* $\mathbb{R}$ *es generada por el conjunto de todos los intervalos de la forma* $(-\infty, x]$ *con* $x \in \mathbb{R}$*, se puede demostrar que una aplicación* $X : \Omega \longrightarrow \mathbb{R}$ *es una* $\Im - \mathcal{B}$ *-variable aleatoria, si y sólo si,* $X^{-1}((-\infty, x]) \in \Im$ *para todo* $x \in \mathbb{R}$.

Ejemplo 2.5 *Sean* $(\Omega, \Im, P)$ *un espacio de probabilidad y* $A \in \Im$ *fijo. La aplicación* $\mathcal{X}_A : \Omega \longrightarrow \mathbb{R}$ *dada por:*

$$\mathcal{X}_A(\omega) = \begin{cases} 1 & si \quad \omega \in A \\ 0 & si \quad \omega \notin A \end{cases}$$

es una variable aleatoria real. En efecto, si $\alpha \in \mathbb{R}$ *entonces:*

$$\{\omega \in \Omega : \mathcal{X}_A(\omega) \leq \alpha\} = \begin{cases} \Phi & si \quad \alpha < 0 \\ A^c & si \quad 0 \leq \alpha < 1 \\ \Omega & si \quad \alpha \geq 1 \end{cases}$$

La aplicación $\mathcal{X}_A$ *se llama la función indicadora de* A*. Otras notaciones de uso frecuente, para esta aplicación, son:* I_A *y* 1_A. ▲

Haciendo uso de las propiedades de la $\sigma-$álgebra y de la definición de variable aleatoria real, se puede demostrar que: si X y Y son variables aleatorias reales definidas sobre el mismo espacio de probabilidad, y α es una constante real, entonces, $X+Y$, αX, XY, $\frac{X}{Y}$ (si está definida), $\max\{X, Y\}$ y $\min\{X, Y\}$ son también variables aleatorias.

Notación 2.6 *Sea X una variable aleatoria definida sobre un espacio de probabilidad $(\Omega, \Im, P)$ y con valores en un espacio medible $(\widetilde{\Omega}, \widetilde{\Im})$. Se define:*

$$\{X \in B\} := \{\omega \in \Omega : X(\omega) \in B\} \quad con\ B \in \widetilde{\Im}$$

Teorema 2.7 *Supóngase que X es una variable aleatoria definida sobre el espacio de probabilidad $(\Omega, \Im, P)$ y con valores en el espacio medible $(\widetilde{\Omega}, \widetilde{\Im})$. La función P_X definida sobre la σ-álgebra $\widetilde{\Im}$ por medio de:*

$$P_X(B) := P(\{X \in B\}); \text{ para todo } B \in \widetilde{\Im}$$

es una medida de probabilidad sobre $(\widetilde{\Omega}, \widetilde{\Im})$ llamada distribución de la variable aleatoria X.

Demostración. *Se debe verificar que P_X satisface las tres condiciones que definen una medida de probabilidad.*

1. *Es claro que para todo $B \in \widetilde{\Im}$, $P_X(B) \geq 0$.*
2. *$P_X(\widetilde{\Omega}) = P(\{\omega \in \Omega : X(\omega) \in \widetilde{\Omega}\} = P(\Omega) = 1$.*
3. *Sean $A_1, A_2, \cdots$ elementos de $\widetilde{\Im}$ con $A_i \cap A_j = \Phi$ para todo $i \neq j$. Entonces,*

$$\begin{aligned}
P_X\left(\bigcup_{i=1}^{\infty} A_i\right) &= P\left(\left\{X \in \bigcup_{i=1}^{\infty} A_i\right\}\right) \\
&= P\left(\bigcup_{i=1}^{\infty} \{X \in A_i\}\right) \\
&= \sum_{i=1}^{\infty} P(\{X \in A_i\}) \\
&= \sum_{i=1}^{\infty} P_X(A_i).
\end{aligned}$$

■

Ejemplo 2.8 *Sean $\Omega = \{1, 2, 3\}, \Im = \{\Phi, \{1\}, \{2, 3\}, \Omega\}$ y P dada por:*

$$\begin{aligned}
P(\Phi) &= 0 \\
P(\{1\}) &= \frac{1}{5} \\
P(\{2, 3\}) &= \frac{4}{5} \\
P(\Omega) &= 1
\end{aligned}$$

Si $\widetilde{\Omega} = \{a, b\}$, $\widetilde{\Im} = \wp(\widetilde{\Omega})$ *y* $X : \Omega \longrightarrow \widetilde{\Omega}$ *está dada por:*

$$X(\omega) = \begin{cases} a & si \quad \omega = 1 \\ b & si \quad \omega = 2 \ \ o \ \ \omega = 3 \end{cases}$$

entonces la distribución P_X *de* X *es igual a:*

$$\begin{aligned} P_X(\Phi) &= \quad 0 \\ P_X(\{a\}) &= \ P(\{1\}) \quad = \tfrac{1}{5} \\ P_X(\{b\}) &= P(\{2,3\}) = \tfrac{4}{5} \\ P_X(\widetilde{\Omega}) &= \quad 1. \qquad \blacktriangle \end{aligned}$$

De ahora en adelante y mientras no se establezca lo contrario se trabajará con variables aleatorias reales exclusivamente.

Definición 2.9 (función de distribución) *Sea* X *una variable aleatoria real. La función* F_X *definida sobre* $\mathbb{R}$ *por medio de:*

$$\begin{aligned} F_X(x) &:= P_X((-\infty, x]) \\ &= P(X \leq x) \end{aligned}$$

se llama función de distribución (o de distribución acumulativa) de la variable aleatoria X.

Ejemplo 2.10 *Sean* $(\Omega, \Im, P)$ *un espacio de probabilidad arbitrario y* $A \in \Im$ *fijo. La función de distribución* F *de la variable aleatoria* $\mathcal{X}_A$ *está dada por:*

$$F(x) = \begin{cases} 0 & si \quad x < 0 \\ P(A^c) & si \quad 0 \leq x < 1 \\ 1 & si \quad x \geq 1 \end{cases} \qquad \blacktriangle$$

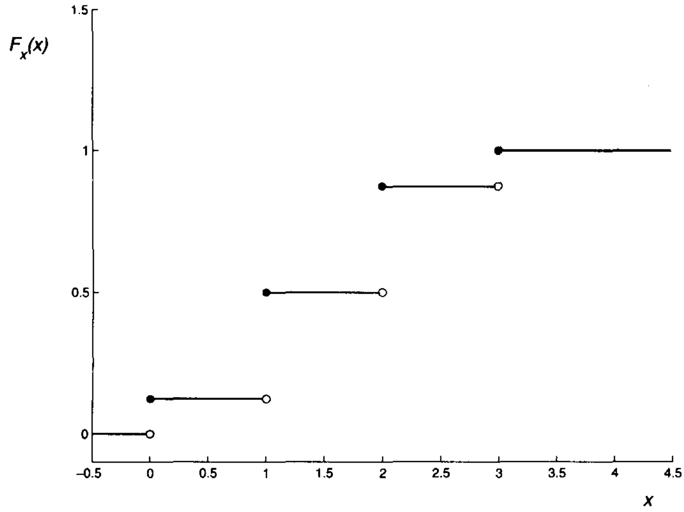
Ejemplo 2.11 *Supóngase que se lanza una moneda corriente tres veces consecutivas y sea* X *la variable aleatoria definida por:*

$$X := \text{“número de caras obtenidas”}$$

En este caso se tiene que la función de distribución de X *es igual a:*

$$F_X(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{1}{8} & si \quad 0 \leq x < 1 \\ \frac{1}{2} & si \quad 1 \leq x < 2 \\ \frac{7}{8} & si \quad 2 \leq x < 3 \\ 1 & si \quad x \geq 3 \end{cases}$$

cuya gráfica indica la figura.



Ejemplo 2.12 *Supóngase que se lanza un dado corriente dos veces consecutivas. Sea Y la variable aleatoria dada por:*

$$Y := \text{"diferencia absoluta de los resultados obtenidos"}.$$

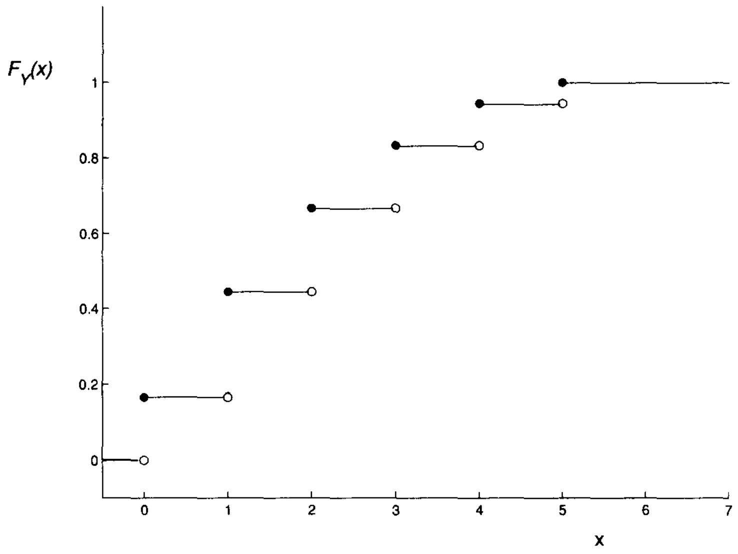
Esto es:

$$Y((x,y)) = |x - y| \qquad \text{con } x, y \in \{1,2,3,4,5,6\};$$

en este caso la función de distribución de la variable aleatoria Y está dada por:

$$F_Y(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{6}{36} & si \quad 0 \le x < 1 \\ \frac{16}{36} & si \quad 1 \le x < 2 \\ \frac{24}{36} & si \quad 2 \le x < 3 \\ \frac{30}{36} & si \quad 3 \le x < 4 \\ \frac{34}{36} & si \quad 4 \le x < 5 \\ 1 & si \quad x \ge 5. \end{cases}$$

cuya gráfica se presenta en la figura siguiente:



Un resultado importante de la teoría de la probabilidad, cuya demostración va más allá de los objetivos de estas notas, establece que la distribución P_X de una variable aleatoria real X queda completamente determinada por su función de distribución F_X [Mun]. Es por esto que, en el caso de variables aleatorias reales se acostumbra identificar la distribución de la variable, la cual como se vio anteriormente es una medida de probabilidad sobre $(\mathbb{R}, \mathcal{B})$, con su función de distribución.

Se observa en los ejemplos anteriores, que la función de distribución de las variables aleatorias en consideración tienen ciertas características comunes como, por ejemplo: todas ellas son no decrecientes y continuas a derecha, el límite cuando x tiende a ∞, en todos los casos es 1; en tanto que el límite cuando x tiende a $-\infty$, en todos los casos es igual a 0. Estas propiedades son características de toda función de distribución como lo demuestra el teorema siguiente.

Teorema 2.13 *Sea X una variable aleatoria real definida sobre $(\Omega, \Im, P)$. La función de distribución F_X satisface las siguientes condiciones:*

1. *Si $x \leq y$ entonces $F_X(x) \leq F_X(y)$*

2. $F_X(x^+) := \lim_{h \to 0^+} F_X(x+h) = F_X(x)$ *para todo $x \in \mathbb{R}$.*

3. $\lim_{x \to \infty} F_X(x) = 1$

4. $\lim_{x\to-\infty} F_X(x) = 0$

Demostración.

1. Si $x \leq y$ entonces

$$\{\omega \in \Omega : X(\omega) \leq x\} \subseteq \{\omega \in \Omega : X(\omega) \leq y\}$$

por lo tanto:

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

2. Sea $x \in \mathbb{R}$ fijo. Supóngase que $(x_n)_n$ es una sucesión decreciente de números reales cuyo límite es x. Esto es:

$$x_1 \geq x_2 \geq \cdots > x \quad \text{y} \quad \lim_{n\to\infty} x_n = x$$

se observa que:

$$\{X \leq x_1\} \supseteq \{X \leq x_2\} \supseteq \cdots$$

y que

$$\bigcap_{n=1}^{\infty}\{X \leq x_n\} = \{X \leq x\}.$$

Por lo tanto, de las propiedades de la medida de probabilidad P, se sigue que:

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

3. Es evidente, que para todo $n \in \mathbb{N}$, se satisface:

$$\{X \leq n\} \subseteq \{X \leq (n+1)\}$$

además

$$\Omega = \bigcup_{n=1}^{\infty}\{X \leq n\}.$$

Por lo tanto,

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

4. Es claro que, para todo $n \in \mathbb{N}$, se satisface:

$$\{X \leq -n\} \supseteq \{X \leq -(n+1)\}$$

y que además,

$$\Phi = \bigcap_{n=1}^{\infty} \{X \leq -n\}.$$

Por lo tanto:

$$\lim_{x \to -\infty} F_X(x) = \lim_{n \to \infty} F_X(-n) = \lim_{n \to \infty} P(X \leq -n) = P(\Phi) = 0.$$

■

Corolario 2.14 *Sean X una variable aleatoria real definida sobre $(\Omega, \Im, P)$, F_X su función de distribución y $a, b \in \mathbb{R}$ con $a < b$, entonces:*

1. $F_X(x^-) := \lim_{h \to 0^+} F_X(x - h) = P(X < x)$
2. $P(a \leq X \leq b) = F_X(b) - F_X(a^-)$
3. $P(a < X \leq b) = F_X(b) - F_X(a)$
4. $P(a \leq X < b) = F_X(b^-) - F_X(a^-)$
5. $P(a < X < b) = F_X(b^-) - F_X(a)$
6. $P(X = a) = F_X(a) - F_X(a^-)$
7. Si $P(a < X < b) = 0$ entonces F_X es constante en el intervalo (a, b).

Demostración. Se hacen sólo las demostraciones de los numerales 1. y 2., las demás quedan como ejercicio.

1. Sea $x \in \mathbb{R}$ fijo. Supóngase que $(x_n)_n$ es una sucesión creciente de números reales cuyo límite es x. Esto es:

$$x_1 \leq x_2 \leq \cdots < x \quad \text{y} \quad \lim_{n \to \infty} x_n = x$$

es claro que:

$$\{X \leq x_1\} \subseteq \{X \leq x_2\} \subseteq \cdots$$

y que

$$\bigcup_{n=1}^{\infty} \{X \leq x_n\} = \{X < x\}.$$

Por lo tanto:

$$F_X(x^-) = \lim_{n\to\infty} F_X(x_n) = \lim_{n\to\infty} P(X \leq x_n) = P(X < x)$$

2. Puesto que:

$$\Omega = \{X < a\} \cup \{a \leq X \leq b\} \cup \{X > b\}$$

se obtiene:

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

Esto es:

$$P(a \leq X \leq b\} = P(X \leq b) - P(X < a) = F_X(b) - F_X(a^-)$$

■

Las variables aleatorias se clasifican de acuerdo con su función de distribución. Si la función de distribución F_X, de la variable aleatoria X, es una función escalonada, entonces se dice que X es una variable aleatoria *discreta,* si F_X es continua, entonces, se dice que X es una variable aleatoria *continua,* y si F_X se puede expresar como una combinación lineal de una función escalonada y una función continua, entonces, se dice que X es una variable aleatoria *mixta.*

Ejemplo 2.15 *Sea X una variable aleatoria real cuya función de distribución está dada por:*

$$F_X(x) = \begin{cases} 0 & si \quad x < -\sqrt{2} \\ \frac{3}{5} & si \quad -\sqrt{2} \leq x < \pi \\ 1 & si \quad x \geq \pi \end{cases}$$

entonces, X es una variable aleatoria discreta. Obsérvese que X toma sólo los valores $-\sqrt{2}$ y π con probabilidades $\frac{3}{5}$ y $\frac{2}{5}$ respectivamente. ▲

Ejemplo 2.16 *Sea X una variable aleatoria cuya función de distribución está dada por:*

$$F_X(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{x^2}{4} & si \quad 0 \leq x < 2 \\ 1 & si \quad x \geq 2 \end{cases}$$

puesto que F_X es una función continua, se tiene que X es una variable aleatoria continua. ▲

Ejemplo 2.17 *Sea X la variable aleatoria cuya función de distribución está dada por:*

$$F_X(x) = \frac{1}{2}\mathcal{X}_{[0,\infty)}(x) + \frac{1}{2}(1 - e^{-3x})\mathcal{X}_{[0,\infty)}(x),$$

esto es:

$$F_X(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{1}{2} + \frac{1}{2}(1 - e^{-3x}) & si \quad x \geq 0 \end{cases}$$

Es claro que X es una variable aleatoria mixta. ▲

Nota 2.18 *Si X es una variable aleatoria real continua definida sobre un espacio de probabilidad $(\Omega, \Im, P)$, entonces, $P(X = a) = 0$; para todo $a \in \mathbb{R}$.*

En estas notas se trabajarán básicamente variables aleatorias discretas y variables aleatorias continuas.

2.1. Variables aleatorias discretas.

Las funciones de distribución escalonadas tienen discontinuidades que son llamadas *saltos*. A continuación se establece de manera precisa, dicho concepto.

Definición 2.19 *Sean X una variable aleatoria real y F_X su función de distribución. Se dice que F_X presenta un salto en el punto $a \in \mathbb{R}$ si*

$$F_X(a) - F_X(a^-) \neq 0$$

La diferencia $F_X(a)-F_X(a^-)$ se llama magnitud del salto y por las propiedades desarrolladas anteriormente se tiene que es igual a $P(X = a)$.

Ejemplo 2.20 *En el ejemplo 2.12 se observa que la variable aleatoria Y presenta saltos en los puntos $x = i$ con $i = 0, \cdots, 5$. Las magnitudes de dichos saltos son, $\frac{1}{6}, \frac{5}{18}, \frac{2}{9}, \frac{1}{6}, \frac{1}{9}$ y $\frac{1}{18}$ respectivamente.* ▲

Nota 2.21 *Si X es una variable aleatoria real continua entonces el conjunto de saltos de F_X es el conjunto vacio.*

El resultado siguiente es muy importante pues, garantiza que el número de saltos de una variable aleatoria real discreta es a lo más numerable.

Teorema 2.22 *Sean X una variable real discreta definida sobre el espacio de probabilidad $(\Omega, \Im, P)$ y F_X su función de distribución entonces el número de saltos de F_X es a lo más numerable.*

Demostración. Puesto que la magnitud de cada salto es un elemento del intervalo $(0, 1]$ y la colección de intervalos I_n de la forma

$$I_n = (\frac{1}{2^{n+1}}, \frac{1}{2^n}] \quad \text{con } n = 0, 1, ...$$

forma una partición de $(0, 1]$, se tiene que la magnitud de cada salto debe pertenecer a alguno de los intervalos I_n. Como las magnitudes de los saltos son probabilidades, es claro que a lo más hay un salto cuya magnitud está en el intervalo I_0, hay a lo más tres saltos cuya magnitud está en el intervalo I_1, hay a lo más 7 saltos con magnitud en el intervalo I_2 y en general que hay a lo más $(2^{n+1} - 1)$ saltos cuya magnitud está en el intervalo I_n. Por lo tanto, como existe un número numerable de intervalos I_n y a lo más $(2^{n+1} - 1)$ saltos en el intervalo I_n, se concluye que el número de saltos es a lo más numerable. ■

Del resultado anterior se concluye que el rango de una variable aleatoria real discreta es a lo más un conjunto numerable.

Sea X una variable aleatoria real discreta y supóngase que X toma los valores $x_1, x_2, \cdots$ (todos diferentes). Sea x un número real arbitrario, entonces:

$$\begin{aligned} F_X(x) &= P(X \leq x) \\ &= P\left(\bigcup_{x_i \leq x} (X = x_i)\right) \\ &= \sum_{x_i \leq x} P(X = x_i). \end{aligned}$$

Esto es, la función de distribución de X queda completamente determinada por los valores p_i con $i = 1, 2 \cdots$ donde $p_i := P(X = x_i)$. Esta observación motiva la definición siguiente:

Definición 2.23 (función de densidad discreta) *Sea X una variable aleatoria real discreta con valores $x_1, x_2, \cdots$ (todosdiferentes). La función f_X, definida sobre $\mathbb{R}$, mediante:*

$$f_X(x) = \begin{cases} P(X = x_i) & \text{si} \quad x = x_1, x_2, \cdots \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

se llama función de densidad de la variable aleatoria X.

Ejemplo 2.24 *Supóngase que se lanza un dado corriente una vez y sea X la variable aleatoria que indica el resultado obtenido. En este caso se tiene que los posibles valores de X son $1, \cdots, 6$. La función de densidad de X está dada por:*

$$f_X(x) = \begin{cases} \frac{1}{6} & si \quad x = 1, \cdots, 6 \\ 0 & en \quad otro\ caso. \quad \blacktriangle \end{cases}$$

Se ha visto que la función de densidad, de la variable aleatoria X, determina completamente su función de distribución. Recíprocamente se tiene que para cualquier $x \in \mathbb{R}$ se satisface:

$$P(X = x) = F_X(x) - F_X(x^-),$$

esto es, la función de distribución de la variable aleatoria X, determina completamente su función de densidad.

Ejemplo 2.25 *Sea X una variable aleatoria cuya función de distribución está dada por:*

$$F_X(x) = \begin{cases} 0 & si \quad x < -2 \\ \frac{1}{7} & si \quad -2 \leq x < \frac{1}{2} \\ \frac{4}{7} & si \quad \frac{1}{2} \leq x < \sqrt{2} \\ 1 & si \quad x \geq \sqrt{2} \end{cases}$$

En este caso se tiene que la función de densidad, de la variable aleatoria X, está dada por:

$$f_X(x) = \begin{cases} \frac{1}{7} & si\ x = -2 \\ \frac{3}{7} & si\ x = \frac{1}{2} \\ \frac{3}{7} & si\ x = \sqrt{2} \\ 0 & en\ otro\ caso. \quad \blacktriangle \end{cases}$$

Ejemplo 2.26 *Sea X una variable aleatoria discreta con función de densidad dada por:*

$$f_X(x) = \begin{cases} \frac{5}{9} & si \quad x = -1 \\ \frac{4}{9} & si \quad x = \frac{3}{2} \\ 0 & en \quad otro\ caso \end{cases}$$

En este caso se tiene que la función de distribución de la variable aleatoria X está dada por:

$$F_X(x) = \begin{cases} 0 & si \quad x < -1 \\ \frac{5}{9} & si \quad -1 \leq x < \frac{3}{2} \\ 1 & si \quad x \geq \frac{3}{2}. \quad \blacktriangle \end{cases}$$

2.2. Variables aleatorias continuas

En el estudio de las variables aleatorias continuas, se hace especial énfasis en las absolutamente continuas, definidas como sigue:

Definición 2.27 (variable aleatoria absolutamente continua) *Sea X una variable aleatoria real definida sobre el espacio de probabilidad $(\Omega, \Im, P)$. Se dice que X es absolutamente continua, si y sólo si, existe una función real no negativa e integrable f_X tal que, para todo $x \in \mathbb{R}$, se satisface:*

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt. \tag{2.1}$$

La función f_X recibe el nombre de función de densidad (fdd) de la variable aleatoria X.

Nota 2.28 *La integral dada en (2.1) es una integral de Riemann.*

Nota 2.29 *En la definición de función de densidad, para variables aleatorias absolutamente continuas, se habla de una función de densidad, ya que ésta no es única. De hecho, si $f_X(\cdot)$ es una fdd de X, entonces, cualquier función que se obtenga al modificar $f_X(\cdot)$ en un número finito de puntos es también una fdd de X.*

Ejemplo 2.30 *Sea X una variable aleatoria con función de distribución dada por:*

$$F_X(x) = \begin{cases} 0 & si \quad x < 0 \\ x & si \quad 0 \leq x < 1 \\ 1 & si \quad x \geq 1 \end{cases}$$

La función f_X definida por:

$$f_X(x) = \begin{cases} 1 & si \quad 0 \leq x < 1 \\ 0 & en \quad otro\ caso \end{cases}$$

es una función de densidad de la variable aleatoria X. ▲

Ejemplo 2.31 *Sea X una variable aleatoria con función de distribución dada por:*

$$F_X(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{x^2}{4} & si \quad 0 \leq x \leq 2 \\ 1 & si \quad x > 2 \end{cases}$$

Es fácil verificar que la función f_X dada por:

$$f_X(x) = \begin{cases} \frac{1}{2}x & si \quad 0 \leq x \leq 2 \\ 0 & en \quad otro\ caso \end{cases}$$

es una función de densidad de la variable aleatoria X. ▲

Supóngase que X es una variable aleatoria absolutamente continua con función de densidad f_X. A continuación se van a deducir algunas de las propiedades que satisface dicha función.

Se tiene:

$$\begin{aligned} \int_{-\infty}^{\infty} f_X(t)dt &= \lim_{x \to \infty} \int_{-\infty}^{x} f_X(t)dt \\ &= \lim_{x \to \infty} F_X(x) = 1 \end{aligned}$$

Si $a, b \in \mathbb{R}$ entonces:

$$\begin{aligned} P(a < X \leq b) &= F_X(b) - F_X(a) \\ &= \int_{-\infty}^{b} f_X(x)dx - \int_{-\infty}^{a} f_X(x)dx \\ &= \int_{a}^{b} f_X(x)dx. \end{aligned}$$

Obsérvese que por ser F_X una función continua, la integral anterior es igual también a $P(a \leq X \leq b)$, $P(a < X < b)$, etc. Más aún, se tiene que:

$$P(X \in B) = \int_B f_X(x)dx \tag{2.2}$$

para todo conjunto Borel B. La demostración de este resultado va más allá de los objetivos de este curso. Es necesario aclarar que la integral dada en (2.2) debe ser interpretada como una integral de Lebesgue, pues la integral de Riemann no está definida para todos los conjuntos boreleanos.

Sin embargo, si B es un intervalo o una unión de intervalos la integral de Riemann tiene sentido, lo cual, para efectos prácticos, es suficiente.

Finalmente, para todo $x \in \mathbb{R}$, donde F_X sea derivable, se satisface:

$$\frac{d}{dx}F_X(x) = f_X(x).$$

Esta última propiedad implica que, para $\Delta x \approx 0$, se tiene:

$$\begin{aligned} P\left(x - \Delta x < X \leq x + \Delta x\right) &= F_X(x + \Delta x) - F_X(x - \Delta x) \\ &\approx 2\Delta x f_X(x) \end{aligned}$$

esto es, la probabilidad de que X esté en un intervalo de longitud pequeña alrededor de x es igual a la función de densidad de X evaluada en x por la longitud del intervalo.

Cuando sea claro del contexto a qué variable aleatoria se hace referencia, se suprimirá el subíndice X tanto en la función de distribución como en la de densidad.

Ejemplo 2.32 *Sea X una variable aleatoria con función de densidad dada por:*

$$f(x) = \begin{cases} \alpha x(2-x) & si \quad 0 \leq x \leq 2 \\ 0 & en \quad otro\ caso \end{cases}$$

Determinar:

1. *El valor de α.*

2. *La función de distribución de la variable aleatoria X.*

3. $P(-1 \leq X \leq \frac{1}{2})$.

Solución:

1. Puesto que

$$1 = \int_{-\infty}^{\infty} f(x)dx = \alpha \int_0^2 (2x - x^2)dx = \alpha\frac{4}{3}$$

se obtiene que $\alpha = \frac{3}{4}$.

2.

$$
\begin{aligned}
F(x) &= \int_{-\infty}^{x} f(t)dt \\
&= \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{3}{4}\int_0^x (2t - t^2)dt & \text{si} \quad 0 \le x \le 2 \\ 1 & \text{si} \quad x > 2 \end{cases} \\
&= \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{4}x^2(3 - x) & \text{si} \quad 0 \le x \le 2 \\ 1 & \text{si} \quad x > 2 \end{cases}
\end{aligned}
$$

3.

$$
\begin{aligned}
P\left(-1 \le X \le \frac{1}{2}\right) &= F(\frac{1}{2}) - F(-1) \\
&= \frac{5}{64}. \quad \blacktriangle
\end{aligned}
$$

Ejemplo 2.33 *Sea X una variable aleatoria cuya función de distribución está dada por:*

$$
F(x) = \begin{cases} 1 - (1 + x)e^{-x} & si \quad x \ge 0 \\ 0 & en \quad otro\ caso \end{cases}
$$

1. Determinar una función de densidad X.

2. Calcular $P(X \le \frac{1}{3})$.

Solución:

1.

$$
f(x) = \begin{cases} xe^{-x} & \text{si} \quad x > 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}
$$

2.

$$
\begin{aligned}
P\left(X \le \frac{1}{3}\right) &= F(\frac{1}{3}) \\
&= 1 - \frac{4}{3}e^{-\frac{1}{3}} \\
&= 4.4625 \times 10^{-2}. \quad \blacktriangle
\end{aligned}
$$

2.3. Distribución de una función de una variable aleatoria

Supóngase que X es una variable aleatoria real definida sobre el espacio de probabilidad $(\Omega, \Im, P)$ y sea g una función tal que $Y = g(X)$ es una variable aleatoria definida sobre $(\Omega, \Im, P)$. Interesa determinar (si es posible) la función de distribución de la variable aleatoria Y en términos de la función de distribución de la variable aleatoria X.

Para indicar el procedimiento se analizarán algunos casos particulares.

Ejemplo 2.34 *Sea X una variable aleatoria discreta con distribución dada por:*

x	-1	0	1	2	3
$P(X = x)$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

Sea $Y = X^2$. Se observa que los posibles valores de Y son $0, 1, 4$ y 9. Además:

y	0	1	4	9
$P(Y = y)$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{2}{7}$	$\frac{1}{7}$

Las correspondientes funciones de distribución de X y de Y son:

$$F_X(x) = \begin{cases} 0 & si \quad x < -1 \\ \frac{1}{7} & si \quad -1 \leq x < 0 \\ \frac{3}{7} & si \quad 0 \leq x < 1 \\ \frac{4}{7} & si \quad 1 \leq x < 2 \\ \frac{6}{7} & si \quad 2 \leq x < 3 \\ 1 & si \quad x \geq 3 \end{cases}$$

y

$$F_Y(y) = \begin{cases} 0 & si \quad x < 0 \\ \frac{2}{7} & si \quad 0 \leq x < 1 \\ \frac{4}{7} & si \quad 1 \leq x < 4 \\ \frac{6}{7} & si \quad 4 \leq x < 9 \\ 1 & si \quad x \geq 9. \quad \blacktriangle \end{cases}$$

Ejemplo 2.35 *Sea X una variable aleatoria real discreta o continua. Sea $Y := aX + b$ donde $a, b \in \mathbb{R}$ y $a \neq 0$. Es claro que:*

$$
\begin{aligned}
F_Y(x) &= P(Y \leq x) \\
&= P(aX + b \leq x) \\
&= \begin{cases} P(X \leq \frac{x-b}{a}) & si \quad a > 0 \\ P(X \geq \frac{x-b}{a}) & si \quad a < 0 \end{cases} \\
&= \begin{cases} F_X(\frac{x-b}{a}) & si \quad a > 0 \\ 1 - F_X\left(\left(\frac{x-b}{a}\right)^-\right) & si \quad a < 0. \quad \blacktriangle \end{cases}
\end{aligned}
$$

Ejemplo 2.36 *Sea X una variable aleatoria real con función de densidad dada por:*

$$
f_X(x) = \begin{cases} \frac{1}{2} & si \quad -1 < x < 1 \\ 0 & en \quad otro\ caso \end{cases}
$$

Sea $Y = |X|$. Entonces:

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= \begin{cases} P(|X| \leq y) & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases} \\
&= \begin{cases} P(-y \leq X \leq y) & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases} \\
&= \begin{cases} F_X(y) - F_X(-y) & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases}
\end{aligned}
$$

Por lo tanto, una función de densidad de la variable aleatoria Y está dada por:

$$
\begin{aligned}
f_Y(y) &= \begin{cases} f_X(y) + f_X(-y) & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases} \\
&= \begin{cases} 1 & si \quad 0 < y < 1 \\ 0 & en \quad otro\ caso. \quad \blacktriangle \end{cases}
\end{aligned}
$$

Ejemplo 2.37 *Sea X una variable aleatoria real con función de densidad dada por:*

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

Sea $Y = e^X$. En este caso se tiene que:

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= \begin{cases} P(X \leq \ln y) & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases} \\
&= \begin{cases} F_X(\ln y) & si \quad y > 0 \\ 0 & si \quad y \leq 0 \end{cases}
\end{aligned}
$$

Por lo tanto, una función de densidad de la variable aleatoria Y está dada por:

$$
\begin{aligned}
f_Y(y) &= \begin{cases} \frac{1}{y} f_X(\ln y) & si \quad y > 0 \\ 0 & en \quad otro\ caso \end{cases} \\
&= \begin{cases} \frac{1}{y} & si \quad 0 < \ln y < 1 \\ 0 & en \quad otro\ caso. \end{cases} \quad \blacktriangle
\end{aligned}
$$

El ejemplo anterior es un caso particular del siguiente teorema.

Teorema 2.38 *Sea X una variable aleatoria real absolutamente continua, con función de densidad f_X. Si h es una función estrictamente monótona y diferenciable, entonces, una función de densidad de la variable aleatoria $Y = h(X)$ está dada por:*

$$
f_Y(y) = \begin{cases} f_X(h^{-1}(y)) \left| \frac{d}{dy} h^{-1}(y) \right| & si \quad y = h(x) \quad para\ algún\ x \\ 0 & si \quad y \neq h(x) \quad para\ todo\ x \end{cases}
$$

donde $h^{-1}(\cdot)$ es la inversa de $h(\cdot)$.

Demostración. Supóngase que h es una función estrictamente creciente y sea $y \in \mathbb{R}$ tal que $y = h(x)$ para algún x.

Entonces:

$$
\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(h(X) \leq y) \\
&= P(X \leq h^{-1}(y)) \\
&= F_X(h^{-1}(y))
\end{aligned}
$$

diferenciando se obtiene:

$$\begin{aligned}
f_Y(y) &= f_X(h^{-1}(y)).\frac{d}{dy}h^{-1}(y) \\
&= f_X(h^{-1}(y)).\left|\frac{d}{dy}h^{-1}(y)\right|
\end{aligned}$$

pues la derivada de h es positiva.

Sean h una función estrictamente decreciente y $y \in \mathbb{R}$ tal que $y = h(x)$ para algún x. Entonces:

$$\begin{aligned}
F_Y(y) &= P(Y \leq y) \\
&= P(h(X) \leq y) \\
&= P(X \geq h^{-1}(y)) \\
&= 1 - F_X(h^{-1}(y)).
\end{aligned}$$

diferenciando se obtiene:

$$\begin{aligned}
f_Y(y) &= -f_X(h^{-1}(y)).\frac{d}{dy}h^{-1}(y) \\
&= f_X(h^{-1}(y))\left|\frac{d}{dy}h^{-1}(y)\right|
\end{aligned}$$

pues en este caso la derivada de h es negativa.

Si $y \in \mathbb{R}$ es tal que $y \neq h(x)$ para todo x, entonces, $F_Y(y) = 0$ o $F_Y(y) = 1$, esto es, $f_Y(y) = 0$. ■

2.4. Valor esperado y varianza de una variable aleatoria

Sea X una variable aleatoria real. Se sabe que sus propiedades de tipo pro-babilístico están determinadas por su función de distribución $F_X(\cdot)$. Sin embargo, interesa conocer un par de valores que "resuman" en cierta forma dicha información. Por ejemplo, se quiere definir un "promedio" de los valores asumidos por X y una cantidad que "mida" que tanto varían los valores de X con respecto a ese valor. Esta información estará dada por dos valores reales llamados esperanza (media, valor esperado) y varianza de la variable aleatoria, respectivamente. A continuación se precisan estos conceptos.

Definición 2.39 (valor esperado) *Sea X una variable aleatoria real definida sobre el espacio de probabilidad $(\Omega, \Im, P)$.*

1. *Si X es una variable aleatoria discreta, con valores $x_1, x_2, \cdots$, se dice que X posee un valor esperado si:*

$$\sum_{k=1}^{\infty} |x_k| \, P(X = x_k) < \infty.$$

En tal caso, se define el valor esperado EX (esperanza matemática, media) de X como:

$$EX = \sum_{k=1}^{\infty} x_k P(X = x_k).$$

2. *Si X es una variable aleatoria continua con función de densidad f_X , se dice que X posee un valor esperado si*

$$\int_{-\infty}^{\infty} |x| \, f_X(x) dx < \infty.$$

En tal caso se define el valor esperado EX (esperanza matemática, media) de X como:

$$EX = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Nota 2.40 *Si X es una variable aleatoria real que toma sólo un número finito de valores, entonces EX existe siempre.*

Ejemplo 2.41 *Supóngase que se lanza un dado corriente una vez y sea X la variable aleatoria que denota el resultado obtenido. Es claro que:*

$$EX = \sum_{k=1}^{6} \frac{k}{6} = \frac{21}{6}. \quad \blacktriangle$$

Ejemplo 2.42 *Sean $(\Omega, \Im, P)$ un espacio de probabilidad arbitrario y $A \in \Im$ fijo. Se sabe que $X := \mathcal{X}_A$ es una variable aleatoria discreta que toma sólo los valores 0 y 1 con probabilidades $P(A^c)$ y $P(A)$ respectivamente. Por lo tanto EX existe y es igual a $P(A)$.* $\blacktriangle$

Ejemplo 2.43 *Sea X una variable discreta con función de densidad dada por:*

$$f(x) = \begin{cases} e^{-3}\frac{3^x}{x!} & si \quad x = 0, 1, 2, \cdots \\ 0 & en \quad otro\ caso \end{cases}$$

entonces:

$$\begin{aligned} EX &= \sum_{k=1}^{\infty} k e^{-3} \frac{3^k}{k!} \\ &= e^{-3} \sum_{k=1}^{\infty} \frac{3^k}{(k-1)!} \\ &= e^{-3} \sum_{j=0}^{\infty} \frac{3^{j+1}}{j!} \\ &= 3e^{-3} \sum_{j=0}^{\infty} \frac{3^j}{j!} \\ &= 3e^{-3} e^3 = 3. \quad \blacktriangle \end{aligned}$$

Ejemplo 2.44 *Sea X una variable aleatoria con fdd dada por:*

$$f(x) = \begin{cases} 2e^{-2x} & si \quad x > 0 \\ 0 & en \quad otro\ caso \end{cases}.$$

puesto que:

$$\int_{-\infty}^{\infty} |x| \, f(x) dx = \int_{0}^{\infty} 2xe^{-2x} = \frac{1}{2}.$$

se concluye que, el valor esperado de X existe y es igual a $\frac{1}{2}$. ▲

Ejemplo 2.45 *Sea X una variable aleatoria con valores en $\mathbb{Z}$. Supóngase que:*

$$P(X = j) = \begin{cases} \frac{c}{j^2} & si \quad j \neq 0 \\ 0 & si \quad j = 0 \end{cases}$$

donde $c > 0$ es una constante tal que

$$\sum_j \frac{c}{j^2} = 1.$$

puesto que:

$$\sum_j |j| \, P(X = j) = \infty,$$

EX no existe. ▲

Ejemplo 2.46 *Sea X una variable aleatoria absolutamente continua con función de densidad dada por:*

$$f(x) = \frac{\alpha}{\pi(\alpha^2 + x^2)}; \quad si\ x \in \mathbb{R}$$

donde $\alpha > 0$ es una constante.

Entonces:

$$\int_{-\infty}^{\infty} |x|\, f(x)dx = \frac{2\alpha}{\pi} \int_0^{\infty} \frac{x}{\alpha^2 + x^2} dx = \infty,$$

es decir EX no existe. ▲

El resultado siguiente permite hallar el valor esperado de una variable aleatoria absolutamente continua a partir de su función de distribución.

Lema 2.47 *Sea Y una variable aleatoria absolutamente continua con función de densidad f. Si EY existe, entonces:*

$$EY = \int_0^{\infty} [1 - F_Y(y)]dy - \int_0^{\infty} F_Y(-y)dy.$$

Demostración.

$$\begin{aligned}
EY &= \int_{-\infty}^{\infty} xf(x)dx \\
&= \int_0^{\infty} xf(x)dx + \int_{-\infty}^{0} xf(x)dx \\
&= \int_0^{\infty} \left(\int_0^{x} dy \right) f(x)dx - \int_{-\infty}^{0} \left(\int_0^{-x} dy \right) f(x)dx \\
&= \int_0^{\infty} \int_y^{\infty} f(x)dxdy - \int_0^{\infty} \int_{-\infty}^{-y} f(x)dxdy \\
&= \int_0^{\infty} P(Y > y)dy - \int_0^{\infty} P(Y \leq -y)dy \\
&= \int_0^{\infty} [1 - F_Y(y)]dy - \int_0^{\infty} F_Y(-y)dy.
\end{aligned}$$

■

Supóngase que X es una variable aleatoria y que $g : \mathbb{R} \longrightarrow \mathbb{R}$ es una función tal que $Y = g(X)$ es una variable aleatoria cuyo valor esperado existe. Es claro que, para calcular EY, se necesita hallar la función de densidad f_Y de Y. Este proceso puede ser complicado, afortunadamente existe un método, que permite hacer los cálculos de una manera más sencilla, conocido como *"ley del estadístico inconciente"*. Dicho método se presenta en el lema siguiente.

Lema 2.48 *Sea X una variable aleatoria real con función de densidad f_X y $g : \mathbb{R} \longrightarrow \mathbb{R}$ una función tal que $Y = g(X)$ es una variable aleatoria. Entonces:*

$$E(g(X)) = \begin{cases} \displaystyle\sum_x g(x)f_X(x) & \text{si} \quad X \text{ es una variable aleatoria discreta} \\ \int_{-\infty}^{\infty} g(x)f_X(x)dx & \text{si} \quad X \text{ es una variable aleatoria absolutamente continua} \end{cases}$$

siempre y cuando la suma, en el caso discreto, o la integral, en el caso continuo, converjan absolutamente.

Demostración.

1. Supóngase que X es una variable aleatoria discreta que toma los valores $x_1, x_2, \cdots$. En este caso la función de densidad de X está dada por:

$$f_X(x) = \begin{cases} P(X = x) & \text{si} \quad x = x_1, x_2, \cdots \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

La variable aleatoria $Y = g(X)$ toma los valores $g(x_1), g(x_2), \cdots$. Es claro que algunos de esos valores pueden ser iguales. Se va a suponer que y_j con $j \geq 1$, representan los diferentes valores de $g(x_i)$. Entonces agrupando todos los $g(x_i)$ que tengan el mismo valor se obtiene:

$$\begin{aligned} \sum_i g(x_i)f_X(x_i) &= \sum_j \sum_{i:g(x_i)=y_j} g(x_i)f_X(x_i) \\ &= \sum_j y_j \sum_{i:g(x_i)=y_j} f_X(x_i) \\ &= \sum_j y_j P(g(X) = y_j) \\ &= \sum_j y_j P(Y = y_j) \\ &= EY = E(g(X)). \end{aligned}$$

2. Sea X una variable aleatoria absolutamente continua con función de densidad f_X. Supóngase que g es una función no negativa, entonces:

$$\begin{aligned} E(g(X)) &= \int_0^\infty P(g(X) > y)dy - \int_0^\infty P(g(X) \leq -y)dy \\ &= \int_0^\infty P(g(X) > y)dy \\ &= \int_0^\infty \left(\int_B f_X(x)dx \right) dy \end{aligned}$$

donde $B := \{x : g(x) > y\}$. Por lo tanto:

$$\begin{aligned} E(g(X)) &= \int_0^\infty \int_0^{g(x)} f_X(x)dydx \\ &= \int_0^\infty g(x)f_X(x)dx. \end{aligned}$$

La demostración del caso general requiere de resultados que van más allá de los objetivos de estas notas. El lector interesado puede consultar el texto de R. Ash [Ash].

■

Ejemplo 2.49 *Sea X una variable con función de densidad dada por:*

$$f(x) = \begin{cases} 2x & si \quad 0 < x < 1 \\ 0 & en \quad otro\ caso \end{cases}$$

Entonces:

$$\begin{aligned} EX^3 &= \int_{-\infty}^\infty x^3 f(x)dx \\ &= \int_0^1 2x^4 dx \\ &= \frac{2}{5}. \quad \blacktriangle \end{aligned}$$

A continuación se presentan algunas de las propiedades más importantes del valor esperado de una variable aleatoria.

Teorema 2.50 *Sea X una variable aleatoria real.*

1. *Si $P(X \geq 0) = 1$ y EX existe, entonces, $EX \geq 0$.*

2. $E\alpha = \alpha$; *para cada constante* α.

3. *Si* X *es acotada, esto es, si existe una constante real* $M > 0$ *tal que* $P(|X| \leq M) = 1$, *entonces,* EX *existe.*

4. *Si* α *y* β *son constantes y si* g *y* h *son funciones, tales que* $g(X)$ *y* $h(X)$ *son variables aleatorias cuyos valores esperados existen, entonces el valor esperado de* $(\alpha g(X) + \beta h(X))$ *existe y*

$$E(\alpha g(X) + \beta h(X)) = \alpha E(g(X)) + \beta E(h(X)).$$

5. *Si* g *y* h *son funciones tales que* $g(X)$ *y* $h(X)$ *son variables aleatorias cuyos valores esperados existen y si* $g(x) \leq h(x)$ *para todo* x, *entonces*

$$E(g(X)) \leq E(h(X)).$$

En particular se tiene que:

$$|EX| \leq E|X|$$

Demostración.

1. Supóngase que X es una variable aleatoria discreta que toma los valores $x_1, x_2, \cdots$. Como $P(X < 0) = 0$, se tiene que $P(X = x_j) = 0$ para todo $x_j < 0$. Por lo tanto:

$$\begin{aligned} EX &= \sum_i x_i P(X = x_i) \\ &= \sum_{i:x_i \geq 0} x_i P(X = x_i) \geq 0 \end{aligned}$$

Si X es una variable aleatoria absolutamente continua con función de densidad f, entonces,

$$\begin{aligned} EX &= \int_0^\infty [1 - F_X(x)]dx - \int_0^\infty F_X(-x)dx \\ &= \int_0^\infty [1 - F_X(x)]dx \\ &= \int_0^\infty P(X > x)dx \geq 0 \end{aligned}$$

2.

$$
\begin{aligned}
E\alpha &= \alpha P(X = \alpha) \\
&= \alpha
\end{aligned}
$$

3. Si X es una variable aleatoria discreta que toma valores $x_1, x_2, \cdots$, entonces, como $P(|X| > M) = 0$ se puede suponer que $\{x_1, x_2, \cdots\} \subseteq [-M, M]$. Por lo tanto,

$$\sum_i |x_i| \, P(X = x_i) \leq M \sum_i P(X = x_i) = M < \infty$$

Si X es una variable aleatoria absolutamente continua con función de densidad f entonces, como $P(|X| > M) = 0$, se puede suponer que $f(x) = 0$ para todo $x \notin [-M, M]$. Por lo tanto:

$$
\begin{aligned}
\int_{-\infty}^{\infty} |x| \, f(x) dx &= \int_{-M}^{M} |x| \, f(x) dx \\
&\leq M \int_{-M}^{M} f(x) dx = M < \infty
\end{aligned}
$$

4. Se hace la demostración para el caso continuo. El caso discreto queda como ejercicio. Si X es una variable aleatoria absolutamente continua con función de densidad f, entonces,

$$
\begin{aligned}
\int_{-\infty}^{\infty} |\alpha g(x) + \beta h(x)| \, f(x) dx &\leq \int_{-\infty}^{\infty} |a g(x)| \, f(x) dx \\
&\quad + \int_{-\infty}^{\infty} |\beta h(x)| \, f(x) dx \\
&= |\alpha| \int_{-\infty}^{\infty} |g(x)| \, f(x) dx \\
&\quad + |\beta| \int_{-\infty}^{\infty} |h(x)| \, f(x) dx < \infty.
\end{aligned}
$$

Luego el valor esperado de $(\alpha g(X) + \beta h(X))$ existe y es claro que:

$$
\begin{aligned}
\int_{-\infty}^{\infty} [\alpha g(x) + \beta h(x)] f(x) dx &= \int_{-\infty}^{\infty} [a g(x)] f(x) dx + \int_{-\infty}^{\infty} [\beta h(x)] f(x) dx \\
&= \alpha E(g(X)) + \beta E(h(X)).
\end{aligned}
$$

5. Se hace la demostración para el caso continuo. El caso discreto queda como ejercicio. Supóngase que X es una variable aleatoria absolutamente continua con función de densidad f. Entonces,

$$E(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)dx \leq \int_{-\infty}^{\infty} h(x)f(x)dx = E(h(X)).$$

Ahora bien, puesto que

$$-X \leq |X| \leq X$$

se obtiene:

$$\begin{aligned} E(-X) &\leq E\,|X| \leq EX \\ -EX &\leq E\,|X| \leq EX \end{aligned}$$

esto es:

$$E\,|X| \leq |EX|$$

■

El valor esperado de una variable aleatoria es el "primer momento central" de la variable alrededor de cero. En general, se tiene que los momentos centrales alrededor de cero de una variable aleatoria, son los valores esperados de las potencias de la variable. Más precisamente:

Definición 2.51 (momento central alrededor de *cero***)**
Sea X una variable aleatoria real. El $r-$ésimo momento central de X alrededor de cero, denotado por $\mu_r^{'}$, se define como sigue:

$$\mu_r^{'} := EX^r$$

siempre y cuando el valor esperado exista.

El resultado siguiente permite afirmar que si $s < r$ y $\mu_r^{'}$ existe, entonces $\mu_s^{'}$ existe.

Lema 2.52 *Sea X una variable aleatoria tal que $\mu_r^{'}$ existe, entonces, $\mu_s^{'}$ existe para todo $s < r$.*
Demostración. *Se hace la demostración para el caso continuo. El caso discreto queda como ejercicio para el lector.*

Puesto que

$$|x^s| < 1 + |x^r| \qquad \textit{para } x \in \mathbb{R}$$

se tiene:

$$\begin{aligned}\int_{-\infty}^{\infty} |x^s|\, f_X(x)dx &\leq \int_{-\infty}^{\infty} [1+|x^r|] f_X(x)dx \\ &= 1 + \int_{-\infty}^{\infty} |x^r|\, f_X(x)dx < \infty.\end{aligned}$$

■

Definición 2.53 (momento central alrededor de la media) *Sea X una variable aleatoria real cuyo valor esperado existe. El $r-$ésimo momento central de X alrededor de EX se define como:*

$$\mu_r := E([X - EX]^r)$$

siempre y cuando el valor esperado exista.

El resultado siguiente permite relacionar los momentos de orden r, alrededor de la media, con los momentos de orden r, alrededor del origen.

Lema 2.54 *Sea X una variable aleatoria cuyo valor esperado existe. Si μ_r existe, entonces,*

$$\mu_r = \sum_{k=0}^{r} \binom{r}{k} \mu_k^{'} (-\mu_1^{'})^{r-k}.$$

Demostración.

$$\begin{aligned}\mu_r &= E(X - EX)^r \\ &= E(X - \mu_1^{'})^r \\ &= E(\sum_{k=0}^{r} \binom{r}{k} X^k (-\mu_1^{'})^{r-k}) \\ &= \sum_{k=0}^{r} \binom{r}{k} EX^k (-\mu_1^{'})^{r-k} \\ &= \sum_{k=0}^{r} \binom{r}{k} \mu_k^{'} (-\mu_1^{'})^{r-k}.\end{aligned}$$

■

Se observa que para cualquier variable aleatoria, cuyo valor esperado existe, se satisface que $\mu_0 = 1$ y $\mu_1 = 0$. El segundo momento central de X, alrededor de su media, recibe el nombre de *varianza* de la variable aleatoria X y se denota generalmente por σ_X^2; la raíz cuadrada de la varianza se llama *desviación estándar* de X y se denota usualmente por σ_X. Otras notaciones de uso frecuente para la varianza de X son: $Var(X)$ y $V(X)$. En este texto se usará indistintamente cualesquiera de ellas.

La varianza mide la dispersión de los valores de la variable alrededor de su media. El término $(X - EX)^2$ es el cuadrado de la distancia de X a EX y por lo tanto $E(X - EX)^2$ representa el promedio de los cuadrados de las distancias de cada valor de X a EX. Luego si una variable aleatoria tiene una varianza pequeña los valores posibles de X están próximos a la media, en tanto que si X tiene una varianza grande los valores de X tienden a estar lejos de la media.

En las aplicaciones, ver [Osp] se acostumbra a utilizar una medida de dispersión relativa, llamada coeficiente de variación , definida como:

$$CV(X) := \frac{\sigma_X}{EX}; \text{ si } EX \neq 0.$$

Cuando $|EX|$ no está próximo a cero, se usa $CV(X)$ como un indicador de qué tan grande es la varianza. Empíricamente, se ha observado, que cuando $|CV(X)| < 0.1$ la varianza, por lo general, no es grande.

A continuación se presentan algunas de las propiedades más importantes de la varianza de una variable aleatoria.

Teorema 2.55 *Sean X una variable aleatoria, cuyo valor esperado existe, y $\alpha, \beta \in \mathbb{R}$ constantes. Entonces:*

1. $Var(X) \geq 0$
2. $Var(\alpha) = 0$
3. $Var(\alpha X) = \alpha^2 Var(X)$
4. $Var(X + \beta) = Var(X)$
5. $Var(X) = 0$, *si y sólo si,* $P(X = EX) = 1$.

Demostración.

1. Es claro a partir de la definición de varianza y de las propiedades del valor esperado.

2. $$Var(\alpha) = E(\alpha - E(\alpha))^2 = E(0) = 0$$

3.

$$\begin{aligned} Var(\alpha X) &= E[\alpha X - E(\alpha X)]^2 \\ &= E(\alpha X - \alpha EX)^2 \\ &= \alpha^2 E(X - EX)^2 \\ &= \alpha^2 VarX \end{aligned}$$

4.

$$\begin{aligned} Var(X + \beta) &= E[(X + \beta) - E(X + \beta)]^2 \\ &= E[X + \beta - EX - \beta]^2 \\ &= Var(X) \end{aligned}$$

5.

a) Si $X = EX$ con probabilidad 1, es claro que $Var(X) = 0$.

b) Supóngase que $Var(X) = 0$, y sea $a := EX$.
Si $P(X = a) < 1$ entonces, existe $c > 0$ tal que

$$P((X - a)^2 > c) > 0.$$

puesto que

$$(x - a)^2 \geq c\mathcal{X}_{\{(x-a)^2>c\}}$$

entonces

$$\begin{aligned} E(X - a)^2 &\geq E(c\mathcal{X}_{\{(X-a)^2>c\}}) \\ Var(X) &\geq cE(\mathcal{X}_{\{(X-a)^2>c\}}) \\ Var(X) &\geq cP((X - a)^2 > c) > 0 \end{aligned}$$

lo cual es una contradicción.
Por lo tanto $P(X = EX) = 1$.

■

Para el cálculo de la varianza de una variable aleatoria resulta muy útil el resultado siguiente.

Lema 2.56 *Sea X una variable aleatoria cuyo valor esperado existe. Entonces:*

$$Var(X) = EX^2 - (EX)^2$$

Demostración.

$$\begin{aligned} Var(X) &= E(X-EX)^2 \\ &= E(X^2 - 2XEX + (EX)^2) \\ &= EX^2 - 2EXEX + E(EX)^2 \\ &= EX^2 - 2(EX)^2 + (EX)^2 \\ &= EX^2 - (EX)^2 \end{aligned}$$

■

Ejemplo 2.57 *Supóngase que se lanza un dado corriente una vez y sea X la variable aleatoria que denota el resultado obtenido. Se sabe que $EX = \frac{21}{6}$ y como*

$$EX^2 = \sum_{k=1}^{6} \frac{1}{6} k^2 = \frac{91}{6},$$

entonces,

$$Var(X) = \frac{35}{12} \approx 2.92. \quad \blacktriangle$$

Ejemplo 2.58 *Sea X como en el ejemplo 2.42, entonces:*

$$EX^2 = P(X=1) = P(A).$$

Por lo tanto:

$$\begin{aligned} Var(X) &= P(A) - [P(A)]^2. \\ &= P(A)P(A^c). \quad \blacktriangle \end{aligned}$$

Ejemplo 2.59 *Sea X como en el ejemplo 2.43, en este caso:*

$$\begin{aligned} EX^2 &= \sum_{k=0}^{\infty} k^2 e^{-3} \frac{3^k}{k!} \\ &= e^{-3} \sum_{k=1}^{\infty} k \frac{3^k}{(k-1)!} \\ &= e^{-3} \sum_{j=0}^{\infty} (j+1) \frac{3^{j+1}}{j!} \\ &= e^{-3}[9e^3 + 3e^3] \\ &= 12. \end{aligned}$$

Por lo tanto,

$$Var(X) = 12 - 9 = 3. \quad \blacktriangle$$

Ejemplo 2.60 *Sea X como en 2.44. Entonces:*

$$\begin{aligned} EX^2 &= \int_{-\infty}^{\infty} x^2 f(x)dx \\ &= \int_{0}^{\infty} 2x^2 e^{-2x} dx \\ &= \frac{1}{2} \end{aligned}$$

y por lo tanto:

$$Var(X) = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}. \quad \blacktriangle$$

Ejemplo 2.61 *La calificación promedio en una prueba estadística fue de* 62.5 *con una desviación estándar de* 10. *El profesor sospecha que el examen fue difícil y por lo tanto desea ajustar las calificaciones de manera que el promedio sea* 70 *y la desviación estándar* 8. *Si X representa la calificación obtenida por los estudiantes en la prueba. ¿Qué ajuste del tipo $\alpha X + \beta$ debe utilizar?*

Solución: En este caso se tiene que $EX = 62.5$ y que $Var(X) = 100$. Sea $Y = \alpha X + \beta$. Entonces:

$$\begin{aligned} 70 &= E(Y) = \alpha EX + \beta = 62.5\alpha + \beta \\ 64 &= Var(Y) = \alpha^2 Var(X) = 100\alpha^2. \end{aligned}$$

Por lo tanto,

$$|\alpha| = \frac{4}{5}$$

de donde se deduce que hay dos posibles ajustes del tipo pedido:

$$Y = \frac{4}{5}X + 20$$

o

$$Y = -\frac{4}{5}X + 120. \quad \blacktriangle$$

Los momentos de una variable aleatoria X desempeñan un papel muy importante en la estadística tanto teórica como aplicada, por ello resulta muy conveniente tener un mecanismo que permita calcular fácilmente, en lo posible, los momentos de una variable aleatoria. Este mecanismo es proporcionado por la llamada función generadora de momentos, la cual se define a continuación.

Definición 2.62 (función generadora de momentos) *Sea X una variable aleatoria tal que $E(e^{tX})$ es finito para todo $t \in (-\alpha, \alpha)$, con α real positivo. Se define la función generadora de momentos de X, denotada por $m_X(\cdot)$ como:*

$$m_X(t) = E(e^{tX}); \quad \text{con } t \in (-\alpha, \alpha).$$

Esto es:

$$m_X(t) = \begin{cases} \sum\limits_k e^{tx_k} P(X = x_k) & \text{si} \quad \begin{array}{l} X \text{ es una v.a discreta} \\ \text{con valores} \quad x_1, x_2, x_3, \cdots \end{array} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx & \text{si} \quad \begin{array}{l} X \text{ es una v.a continua con función} \\ \text{de densidad } f \end{array} \end{cases}$$

Es importante observar que no todas las distribuciones de probabilidad tienen asociada una función generadora de momentos (f.g.m). Más adelante se darán algunos ejemplos que ratifican esta afirmación.

Antes de dar las propiedades más importantes de la f.g.m, de una variable aleatoria X, se presentan algunos ejemplos.

Ejemplo 2.63 *Sea X una variable aleatoria discreta con función de densidad dada por:*

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{si} \quad x = 0, 1, \cdots, n \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

donde n es un entero positivo y $p \in (0, 1)$.
En este caso:

$$\begin{aligned} m_X(t) &= \sum_{k=0}^{n} e^{tk} \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^{n} \binom{n}{k} (pe^t)^k (1-p)^{n-k} \\ &= (pe^t + q)^n \quad \text{donde } q := 1 - p. \quad \blacktriangle \end{aligned}$$

Nota 2.64 *La distribución dada en el ejemplo 2.63 recibe el nombre de distribución binomial de parámetros n y p. En el próximo capítulo se estudiará esta distribución en detalle.*

Ejemplo 2.65 *Sea X la variable aleatoria dada en 2.43. En este caso:*

$$\begin{aligned} m_X(t) &= \sum_{k=0}^{\infty} e^{tk} e^{-3} \frac{3^k}{k!} \\ &= e^{-3} \sum_{k=0}^{\infty} \frac{(3e^t)^k}{k!} \\ &= e^{-3} \exp(3e^t) \\ &= \exp(3(e^t - 1)). \quad \blacktriangle \end{aligned}$$

Ejemplo 2.66 *Sea X una variable aleatoria con función de densidad dada por:*

$$f(x) = \begin{cases} 2e^{-2x} & si \quad x > 0 \\ 0 & en \quad otro\ caso \end{cases}$$

Es claro que:

$$\begin{aligned} m_X(t) &= \int_0^{\infty} e^{tx} 2e^{-2x} dx \\ &= (1 - \frac{t}{2})^{-1}; \qquad con \quad t < 2. \quad \blacktriangle \end{aligned}$$

Ejemplo 2.67 *Sea X una variable aleatoria con función de densidad dada por:*

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi x}} \exp\{-\frac{1}{2}[\frac{\ln(x)-\mu}{\sigma}]^2\} & si \quad x > 0 \\ 0 & en \quad otro\ caso \end{cases}$$

donde $\mu \in \mathbb{R}$ y $\sigma > 0$ son constantes.

Entonces:

$$E(e^{tX}) = \frac{1}{\sigma\sqrt{2\pi}} \int_0^{\infty} \frac{1}{x} \exp\{tx - \frac{1}{2}[\frac{\ln(x) - \mu}{\sigma}]\} dx$$

haciendo el cambio de variable

$$u = \frac{\ln(x) - \mu}{\sigma}$$

se obtiene que

$$\begin{aligned} E(e^{tX}) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{te^{\mu+\sigma u} - \frac{u^2}{2}\right\} du \\ &\geq \frac{1}{\sqrt{2\pi}} \int_0^{\infty} \exp\left\{te^{\mu+\sigma u} - \frac{u^2}{2}\right\} du \end{aligned}$$

como $\exp(\sigma u) > \frac{(\sigma u)^3}{3!}$ *se deduce que para* $t > 0$

$$te^{\mu+\sigma u} - \frac{u^2}{2} > \frac{u^2}{2}\left[\frac{te^{\mu}\sigma^3}{3}u - 1\right].$$

Por lo tanto, si se toma $u > \frac{6}{te^{\mu}\sigma^3}$ *se obtiene que:*

$$\left[\frac{te^{\mu}\sigma^3}{3}u - 1\right] > 1$$

y en consecuencia tomando

$$a = \frac{6}{te^{\mu}\sigma^3}$$

se obtiene:

$$E(e^{tX}) \geq \frac{1}{\sqrt{2\pi}}\int_a^{\infty} \exp(\frac{u^2}{2})du = \infty,$$

es decir, no existe una vecindad del origen en la que $E(e^{tX})$ *sea finito y en consecuencia la función generadora de momentos de* X *no existe.* ▲

La existencia de la función generadora de momentos de una v.a X garantiza la existencia de todos los momentos de orden r alrededor de cero de X. Es más, si la función generadora de momentos existe, entonces, es diferenciable en una vecindad del origen y se satisface que:

$$\left.\frac{d^r}{dt^r}m_X(t)\right|_{t=0} = EX^r.$$

Más precisamente se tienen los dos teoremas siguientes, cuyas demostraciones pueden ser consultadas en [Her].

Teorema 2.68 *Si* X *es una variable aleatoria cuya función generadora de momentos existe, entonces,* EX^r *existe para toda* $r \in \mathbb{N}_0$

Es importante aclarar que el recíproco del teorema anterior no es válido: el hecho de que existan todos los momentos de orden r de una v.a X no garantiza la existencia de la f.g.m de la variable. Por ejemplo, para la v.a dada en el ejemplo 2.67, se tiene que, para toda $r \in \mathbb{N}_0$:

$$EX^r = \exp\left\{r\mu + \frac{(r\sigma)^2}{2}\right\}.$$

Teorema 2.69 *Si X es una variable aleatoria cuya función generadora de momentos $m_X(.)$ existe, entonces, existe $h \in (0, \infty)$ tal que:*

$$m_X(t) = \sum_{k=0}^{\infty} E(X^k)\frac{t^k}{k!}; \quad \textit{para todo } t \in (-h, h)$$

y por lo tanto,

$$EX^r = \left.\frac{d^r}{dt^r} m_X(t)\right|_{t=0}$$

Una propiedad muy importante de la f.g.m de una variable aleatoria es que, cuando existe, ella caracteriza la distribución de la variable. Más precisamente se tiene el teorema siguiente cuya demostración va más allá de los objetivos de estas notas. El lector interesado puede consultar [Ash]

Teorema 2.70 *Sean X y Y variables aleatorias cuyas funciones generadoras de momentos existen. Si*

$$m_X(t) = m_Y(t); \ \textit{para todo } t,$$

entonces, X y Y tienen la misma distribución.

Como se ve, la función generadora de momentos de una variable aleatoria, cuando existe, es una herramienta muy útil para calcular los momentos de orden r alrededor de cero, de la variable. Desafortunadamente esta función no siempre existe, por esto, se hace necesario introducir una nueva clase de funciones que son igualmente útiles y que siempre existen.

Definición 2.71 *Sea X una variable aleatoria. La función característica de X es la función $\phi_X : \mathbb{R} \longrightarrow \mathbb{C}$ definida por:*

$$\phi_X(t) := E(e^{itX}) = E(\cos tX) + iE(\sin tX); \quad \textit{donde } i = \sqrt{-1}.$$

Ejemplo 2.72 *Sea X una variable aleatoria con*

$$P(X = 1) = P(X = -1) = \frac{1}{2}.$$

Entonces:

$$\begin{aligned} \phi_X(t) &= \frac{1}{2}(\cos t + \cos(-t)) + \frac{i}{2}(\sin t + \sin(-t)) \\ &= \cos t. \quad \blacktriangle \end{aligned}$$

Ejemplo 2.73 *Sea X una variable aleatoria discreta con función de densidad dada en el ejemplo 2.43. Entonces:*

$$\begin{aligned}\phi_X(t) &= E(e^{itX}) \\ &= \sum_{k=0}^{\infty} e^{itk} e^{-3} \frac{3^k}{k!} \\ &= e^{-3} \sum_{k=0}^{\infty} \frac{(3e^{it})^k}{k!} \\ &= e^{-3} \exp(3e^{it}) \\ &= \exp[3(e^{it} - 1)]. \quad \blacktriangle\end{aligned}$$

Ejemplo 2.74 *Sea X una variable aleatoria continua con función de densidad f dada por:*

$$f(x) = \begin{cases} 1 & si \quad 0 < x < 1 \\ 0 & en \quad otro\ caso \end{cases}$$

Entonces, para $t \neq 0$:

$$\begin{aligned}\phi_X(t) &= \int_0^1 \cos(tx)dx + i \int_0^1 \sin(tx)dx \\ &= \frac{1}{t}(\sin t - i\cos t + i) \\ &= \frac{1}{it}(e^{it} - 1)\end{aligned}$$

y para $t = 0$

$$\phi_X(0) = E(e^0) = 1. \quad \blacktriangle$$

Se presentan a continuación, sin demostración, las principales propiedades de la función característica. El lector interesado puede consultar las demostraciones en [Gri] y [Her].

Teorema 2.75 *Si X es una variable aleatoria discreta o absolutamente continua entonces $E(e^{itX})$ existe para todo $t \in \mathbb{R}$.*

Teorema 2.76 *Sea X una variable aleatoria. La función característica $\phi_X(.)$ de X satisface:*

1. $\phi_X(0) = 1.$
2. $|\phi_X(t)| \leq 1$ *para todo t.*

3. Si EX^k existe entonces:

$$\frac{d^k}{dt^k}\phi_X(t)\mid_{t=0}= i^k EX^k.$$

Por último, es importante observar que la función característica de una variable aleatoria, al igual que la f.g.m. (cuando ésta existe), determina la distribución de la variable. Esto es, se satisface:

Teorema 2.77 *Si X y Y son variables aleatorias y*

$$\phi_X(t) = \phi_Y(t);\ \textit{para todo } t,$$

entonces, X y Y tienen la misma distribución.

2.5. Ejercicios

1. Sean $\Omega = \{1,2,3\}$, $\Im = \{\Phi, \{1\}, \{2,3\}, \Omega\}$. Para $A \in \Im$ se define

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

¿Es $X : \Omega \longrightarrow \mathbb{R}$ definida por $X(\omega) = \omega^2$ una variable aleatoria real? Justificar.

2. Sea $\Omega = \{1,2,3,4\}$. Determinar la menor σ-álgebra sobre Ω con respecto a la cual $X(\omega) := \omega + 1$ es una variable aleatoria real.

3. Se carga una moneda de tal manera que $P(C) = \frac{3}{7}$ y $P(S) = \frac{4}{7}$. Supóngase que se lanza la moneda tres veces consecutivas y sea X la variable aleatoria que indica el número de caras obtenidas. Hallar la función de distribución de la variable aleatoria X y calcular EX.

4. El espacio muestral de un experimento aleatorio es $\Omega = \{a, b, c, d, e, f\}$ y cada resultado es igualmente probable. Se define una variable aleatoria X como sigue:

ω	a	b	c	d	e	f
$X(\omega)$	0	0	1.5	1.5	2	3

Calcular las siguientes probabilidades:

a) $P(X = 1.5)$.

b) $P(|X-1| \leq 1.5)$.

c) $P(X \geq 0 \ \vee \ X < 2)$.

5. Demostrar que si todos los valores de una variable aleatoria real X están en el intervalo $[a, b]$ con $a < b$, entonces $F_X(x) = 0$ para todo $x < a$ y $F_X(x) = 1$ para todo $x \geq b$.

6. Se lanza una moneda corriente cuatro veces consecutivas. Sea X la variable aleatoria que denota el número de caras obtenidas. Hallar y graficar la función de distribución de la variable aleatoria $Y := X - 2$.

7. Sea $(\Omega, \Im, P)$ un espacio de probabilidad definido como sigue:

$$\Omega := \{(1,2,3), (2,3,1), (3,1,2)\}, \ \Im := \wp(\Omega) \text{ y } P(\omega) := \frac{1}{3}$$

para todo $\omega \in \Omega$. Considérense las variables aleatorias X_n definidas sobre Ω, por:

$$X_n(\omega) = \omega_n \text{ para } n = 1, 2, 3 \text{ y } \omega = (\omega_1, \omega_2, \omega_3) \in \Omega$$

a) Determinar el conjunto de valores S que toman las variables aleatorias X_n.

b) Verificar que

$$P(X_3 = 3 \mid X_2 \in \{1,2\}, \ X_1 = 3) \neq P(X_3 = 3 \mid X_2 \in \{1,2\}).$$

8. Una urna contiene 5 bolas blancas y 10 negras. Un dado corriente es lanzado. Se extrae de la urna un número de bolas igual al resultado obtenido en el lanzamiento del dado. ¿A qué es igual la probabilidad de que todas las bolas extraídas sean blancas?, ¿a qué es igual la probabilidad de que el resultado obtenido al lanzar el dado sea 3 si todas las bolas extraídas son blancas?

9. Sea X una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} 0.2 & \text{si} \quad -1 < x \leq 0 \\ 0.2 + cx & \text{si} \quad 0 < x \leq 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

a) Determinar el valor de c.

b) Obtener la función de distribución de la variable aleatoria X.

c) Calcular $P(0 \leq X < 0.5)$

d) Determinar $P(X > 0.5 \mid X > 0.1)$

e) Calcular la función de distribución y la función de densidad de la variable aleatoria $Y := 2X + 3$

10. Sea X una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} c(2-x) & \text{si} \quad 0 \leq x \leq 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

a) Calcular el valor de c.

b) Determinar la función de distribución de la variable aleatoria X.

c) Calcular $P(|X| \geq 0.2)$

11. La función de distribución acumulativa de una variable aleatoria X está dada por:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 2x - x^2 & \text{si} \quad 0 \leq x \leq 1 \\ 1 & \text{si} \quad x > 1 \end{cases}$$

a) Calcular $P\left(X \geq \frac{3}{2}\right)$ y $P\left(-2 \leq X \leq \frac{3}{4}\right)$.

b) Determinar una función de densidad $f_X(\cdot)$.

12. Sean X, Y y Z variables aleatorias cuyas funciones de distribución son respectivamente:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad -\infty < x < -1 \\ 0.2 & \text{si} \quad -1 \leq x < 0 \\ 0.7 & \text{si} \quad 0 \leq x < 1 \\ 0.8 & \text{si} \quad 1 \leq x < 2 \\ 1 & \text{si} \quad x \geq 2 \end{cases}$$

$$F_Y(x) = \begin{cases} 0 & \text{si} \quad -\infty < x < 0 \\ \frac{1}{2} + \frac{1}{2}x & \text{si} \quad 0 \leq x < 1 \\ 1 & \text{si} \quad x \geq 1 \end{cases}$$

$$F_Z(x) = \begin{cases} 0 & \text{si} \quad -\infty < x < -1 \\ \frac{1}{2} + \frac{1}{2}x & \text{si} \quad -1 \leq x < 1 \\ 1 & \text{si} \quad x \geq 1 \end{cases}$$

a) ¿Cuáles variables son de tipo discreto?, ¿cuáles son de tipo continuo? Explicar.

b) Calcular $P(X = 0)$, $P(\frac{1}{2} < X \leq 2)$ y $P(X > 1.5)$

c) Calcular $P(Y = 0)$, $P(\frac{1}{2} < Y \leq 2)$ y $P(Y > 0)$

d) Calcular $P(Z = 0)$, $P(-\frac{1}{2} < Z \leq \frac{1}{2})$ y $P(Z \geq 2)$

13. Un tablero circular de radio 1 se zonifica en n discos concéntricos de radios $\frac{1}{n}, \frac{2}{n}, \cdots, 1$. Un dardo se lanza, al azar, dentro del círculo. Si cae en el anillo entre los círculos con radios $\frac{i}{n}$ y $\frac{(i+1)}{n}$ para, $i = 0, \cdots, n-1$, se ganan $(n-i)$ unidades monetarias. Sea X la variable aleatoria que denota la cantidad de dinero ganada. Hallar la función de densidad de la variable aleatoria X.

14. En cada uno de los ejercicios siguientes determinar los valores de la constante C, para los cuales las funciones dadas, son funciones de densidad discretas sobre los enteros positivos:

a) $f(x) = C2^{-x}$

b) $f(x) = \frac{C2^{-x}}{x}$

c) $f(x) = Cx^{-2}$

d) $f(x) = \frac{C2^{x}}{x!}$

15. En cada uno de los ejercicios siguientes determinar los valores de la constante C, para los cuales las funciones dadas, son funciones de densidad:

a) $f(x) = C\{x(1-x)\}^{-\frac{1}{2}}$

b) $f(x) = C\exp(-x - e^{-x}),\ x \in \mathbb{R}$

16. Una variable aleatoria absolutamente continua X toma valores en el intervalo $[0, 4]$ y su función de densidad está dada por:

$$f(x) = \frac{1}{2} - cx$$

a) Determinar el valor de c.

b) Calcular $P(\frac{1}{2} \leq X < 3)$.

17. Sea X una variable aleatoria absolutamente continua con función de densidad f. Demostrar que las variables aleatorias X y $(-X)$ tienen la misma función de distribución, si y sólo si, $f(x) = f(-x)$ para todo $x \in \mathbb{R}$.

18. En los casos siguientes, determinar la función de distribución de la variable aleatoria real discreta X cuya función de densidad está dada por:

    a) $P(X = k) = pq^{k-1}, \ k = 1, 2, \cdots$ con $p \in (0, 1)$ fijo y $q := 1 - p$.

    b)
    $$P(X = k) = \frac{6k^2}{n(n+1)(2n+1)}, \ k = 1, \cdots, n$$

19. Una persona pide prestado un llavero con siete llaves y no sabe cuál de ellas abre un candado. Por lo tanto, intenta con cada una hasta que consigue abrirlo. Sea X la variable aleatoria que indica el número de intentos necesarios para conseguir abrir el candado.

    a) Determinar la función de densidad de la variable aleatoria X.

    b) Calcular $P(X \leq 2)$ y $P(X = 5)$.

20. Cuatro bolas se extraen aleatoriamente y sin reemplazo de una urna que contiene 25 bolas numeradas del 1 al 25. Si usted apuesta a que por lo menos una de las cuatro bolas tiene una numeración menor o igual a 5, ¿cuál es la probabilidad de que usted gane la apuesta?

21. Un jugador extrae simultánea y aleatoriamente dos bolas de una urna que contiene 8 bolas blancas, 5 bolas negras y 3 bolas azules. Suponga que el jugador gana 5000 pesos por cada bola negra seleccionada y pierde 3000 pesos por cada bola blanca seleccionada. Sea X la variable aleatoria que denota la fortuna del jugador. Hallar la función de densidad de la variable aleatoria X.

22. Un comerciante tiene dos puntos de venta de computadores. La probabilidad de que venda, en un día, un computador en el primer punto de venta es de 0.4 e independientemente, la probabilidad de que venda, en un día, un computador en el segundo punto de venta es de 0.7. Supóngase además, que es igualmente probable que él venda un computador de marca o un clon. Un computador de marca tiene un

costo de 1800 dólares, en tanto que un clon, con las mismas especificaciones, tiene un costo de 1000 dólares. Sea X la cantidad, en dólares, que el comerciante vende en un día. Hallar la distribución de la variable aleatoria X.

23. Demostrar que

$$\int_{-\infty}^{\infty} \exp(-x^2)dx = \sqrt{\pi}$$

y usar el resultado para demostrar que

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}; \quad x \in \mathbb{R}$$

es una función de densidad si $\sigma > 0$.

24. Supóngase que f y g son funciones de densidad y que $0 < \lambda < 1$ es una constante. ¿Es $\lambda f + (1-\lambda)g$ una función de densidad?, ¿es fg una función de densidad? Explicar.

25. Sea X una variable aleatoria con función de distribución acumulativa dada por:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 1 \\ \frac{4}{3}x - 1 & \text{si} \quad 1 \leq x < \frac{5}{4} \\ 1 & \text{si} \quad x \geq \frac{5}{4} \end{cases}$$

Determinar una función de distribución acumulativa discreta $F_d(\cdot)$ y una continua $F_c(\cdot)$, así como constantes α y β tales que:

$$F_X(x) = \alpha F_d(x) + \beta F_c(x)$$

26. Sea X una variable aleatoria con función de distribución acumulativa dada por:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ x^2 + \frac{1}{4} & \text{si} \quad 0 \leq x < \frac{1}{2} \\ \frac{3}{4} & \text{si} \quad \frac{1}{2} \leq x < 2 \\ 1 & \text{si} \quad x \geq 2 \end{cases}$$

Determinar una función de distribución acumulativa discreta $F_d(\cdot)$ y una continua $F_c(\cdot)$, así como constantes α y β tales que:

$$F_X(x) = \alpha F_d(x) + \beta F_c(x)$$

27. Se dice que una variable aleatoria real discreta X tiene distribución logarítmica de Fisher, de parámetro θ, si su función de densidad está dada por:

$$f(x) = \begin{cases} \frac{1}{|\ln(1-\theta)|}\frac{\theta^x}{x} & \text{si} \quad x = 1, 2, \cdots \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

con $\theta \in (0, 1)$. Verificar que f es una función de densidad.

28. Sea X una variable aleatoria con función de densidad dada por:

$$f(x) = C.\frac{1}{1+x^2} \text{ si } x \in \mathbb{R}$$

 a) Determinar el valor de C.
 b) Calcular $P(X \geq 0)$.
 c) Hallar (si existen) EX y $Var(X)$.
 d) Hallar la función de distribución de X.

29. Se lanza un dado corriente dos veces consecutivas. Sean X y Y las variables aleatorias definidas por:

$$X := \text{“resultado del primer lanzamiento”}$$
$$Y := \text{“resultado del segundo lanzamiento”}$$

Calcular $E(\text{máx}\{X, Y\})$ y $E(\text{mín}\{X, Y\})$

30. Sea X una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} x^3 & \text{si} \quad 0 < x \leq 1 \\ (2-x)^3 & \text{si} \quad 1 < x \leq 2 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

 a) Calcular $\mu := EX$ y $\sigma^2 := Var(X)$.
 b) Hallar $P(\mu - 2\sigma < X < \mu + 2\sigma)$.

31. Sea X una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} 12x^3 - 21x^2 + 10x & \text{si} \quad 0 \leq x \leq 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

 a) Calcular $\mu := EX$ y $\sigma^2 := Var(X)$.

b) Determinar el valor de c tal que $P(X > c) = \frac{7}{16}$.

32. Sea X una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} C\sin\left(\frac{1}{5}\pi x\right) & \text{si} \quad 0 \leq x \leq 5 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

a) Determinar el valor de C.

b) Hallar (si existen) la media y la varianza de X.

33. Sea X una variable aleatoria con función de distribución acumulativa dada por:

$$F_X(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{2}x & \text{si} \quad 0 \leq x \leq 2 \\ 1 & \text{si} \quad x > 2 \end{cases}$$

y sea $Y := X^2$.

a) Determinar la función de distribución acumulativa de Y.

b) Calcular $P(Y \leq X)$, $P(X + Y \leq \frac{3}{4})$ y $P(X \leq 2Y)$.

34. Sea X una variable aleatoria con función de distribución acumulativa continua y estrictamente creciente.

a) Determinar una función de densidad de la variable aleatoria $Y := |X|$.

b) Hallar la función de distribución acumulativa de la variable aleatoria $Y := X^3$.

35. Se tienen tres urnas A, B y C. La urna A contiene 5 bolas rojas y 4 blancas, la urna B contiene 8 bolas rojas y 5 blancas y la urna C contiene 2 bolas rojas y 6 blancas. Se extrae al azar una bola de cada urna. Sea X el número de bolas blancas extraídas. Calcular la función de densidad de X.

36. Sea X una variable aleatoria con $P(X = 0) = 0$ y tal que EX existe y es diferente de cero.

a) ¿Es válido en general que

$$E\left(\frac{1}{X}\right) = \frac{1}{EX}\ ? \tag{2.3}$$

b) ¿Existe alguna variable aleatoria X para la cual (2.3) se satisface? Explicar.

37. Sea X una variable aleatoria discreta con valores en los enteros no negativos y tal que EX y EX^2 existen.

    a) Demostrar que:

    $$EX = \sum_{k=0}^{\infty} P(X \geq k).$$

    b) Verificar que:

    $$\sum_{k=0}^{\infty} kP(X > k) = \frac{1}{2}\left(EX^2 - EX\right).$$

38. Determinar si las proposiciones siguientes son falsas o verdaderas. Justificar la respuesta.

    a) Si $P(X > Y) = 1$, entonces, $EX > EY$.

    b) Si $EX > EY$, entonces, $P(X > Y) = 1$

    c) Si $Y = X + 1$, entonces, $F_X(x) = F_Y(x+1)$; para todo x.

39. Sea X una variable aleatoria con valores en $\mathbb{Z}^+$con $P(X = k) = \frac{C}{3^k}$.

    a) Determinar el valor de C.

    b) Hallar (si existe) EX.

40. Sea X una variable aleatoria con valores $\frac{3^k}{2^k}$; $k = 0, 1, \cdots$ y tal que $P\left(X = \frac{3^k}{2^k}\right) = \frac{1}{2^k}$. ¿Existe EX?, ¿existe $Var(X)$? Explicar.

41. (Desigualdad de Markov) Sea X una variable aleatoria real con $X \geq 0$ y tal que EX existe. Demostrar que para todo $\alpha > 0$ se satisface:

    $$P(X \geq \alpha) \leq \frac{EX}{\alpha}.$$

42. (Desigualdad de Chebyschev) Sea X una variable aleatoria con media μ y varianza σ^2.

a) Demostrar que para todo $\epsilon > 0$ se satisface:

$$P(|X-\mu| \geq \epsilon) \leq \frac{\sigma^2}{\epsilon^2}.$$

b) Si $\mu = \sigma^2 = 20$, ¿qué se puede decir acerca de $P(0 \leq X \leq 40)$?

43. Sea X una variable aleatoria con media 11 y varianza 9. Utilizar la desigualdad de Chebyschef para encontrar (si es posible):

a) Un límite inferior para $P(6 < X < 16)$

b) El valor de k para el cual $P(|X-11| \geq k) \leq 0.09$

44. Sea X una variable aleatoria con media μ y varianza σ^2. Supóngase que H es una función dos veces diferenciable en $x = \mu$ y que $Y := H(X)$ es una variable aleatoria tal que EY y EY^2 existen.

a) Demostrar las siguientes aproximaciones para EY y $Var(Y)$:

$$E(Y) \approx H(\mu) + \frac{H''(\mu)}{2}\sigma^2$$
$$Var(Y) \approx (H'(\mu))^2 \sigma^2$$

b) Haciendo uso de la parte a., calcular (de manera aproximada) el valor esperado y la varianza de la variable aleatoria

$$Y := 2(1-0.005X)^{1.2}$$

donde X es una variable aleatoria cuya función de densidad está dada por:

$$f_X(x) = 3000x^{-4}\mathcal{X}_{[10,\infty)}(x)$$

45. Hallar la función característica de una variable aleatoria X tal que $P(X=1) = \frac{2}{5}$ y $P(X=0) = \frac{3}{5}$.

46. Determinar la función característica de una variable aleatoria X con función de densidad dada por:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si} \quad a \leq x \leq b \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

donde $a, b \in \mathbb{R}$ y $a < b$.

47. Determinar la función característica de una variable aleatoria X con función de densidad dada por:

$$f(x) = \begin{cases} 1 - |x| & \text{si} \quad -1 \leq x \leq 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

48. (Cuantil de orden q) Para cualquier variable aleatoria X, un cuantil de orden q, con $0 < q < 1$, es cualquier número denotado por x_q, que satisface simultáneamente las condiciones:

$$\begin{aligned} &\text{i) } P(X \leq x_q) \geq q \\ &\text{ii) } P(X \geq x_q) \geq 1 - q \end{aligned}$$

Los cuantiles de uso más frecuente son $x_{0.5}$, $x_{0.25}$ y $x_{0.75}$, llamados, respectivamente *mediana, cuartil inferior* y *cuartil superior.*

Un cuantil x_q no es necesariamente único. Cuando un cuantil no es único, existe un intervalo en el que cada punto satisface las condiciones i) y ii). En este caso, algunos autores sugieren considerar el menor valor del intervalo y otros sugieren el punto medio del intervalo. (ver [Her]). Con base en esta información, resolver los ejercicios planteados a continuación:

a) Sea X una variable aleatoria discreta con función de densidad dada por:

$$f_X(x) = \begin{cases} \binom{5}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{5-x} & \text{si} \quad x = 0, \cdots, 5. \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

Determinar un cuantil inferior, un cuantil superior y la mediana de X.

b) Sea X una variable aleatoria con función de densidad dada por:

$$f_X(x) = \begin{cases} \frac{1}{800} \exp\left(-\frac{x}{800}\right) & \text{si} \quad x > 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

Determinar la media y la mediana de X.

c) Sea X una variable aleatoria con función de densidad dada por:

$$f_X(x) = \begin{cases} \frac{10}{x^2} & \text{si} \quad x > 10 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

Determinar (si existen) la media y la mediana de X.

49. (Moda) Sea X una variable aleatoria con función de densidad $f(.)$.Una moda de X (si existe) es un número real ζ tal que

$$f(\zeta) \geq f(x) \text{ para todo } x \in \mathbb{R}$$

a) Sea X una variable aleatoria discreta con función de densidad dada por:

x	1	2	3	4	5	6
$P(X=x)$	$\frac{9}{40}$	$\frac{1}{5}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{3}{20}$	$\frac{7}{40}$

Hallar (si existe) la moda de X

b) Supóngase que la edad a la que un hombre contrae matrimonio, por primera vez, es una variable aleatoria X con función de densidad dada por:

$$f(x) = \begin{cases} \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} \exp(-\lambda x) & \text{si} \quad x \geq 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

donde $r > 0$ y $\lambda > 0$ son constantes y Γ es la función definida por:

$$\Gamma(r) := \int_0^\infty t^{r-1} \exp(-t)\, dt$$

Si la edad promedio es de 28 años y lo más común es que el hombre se case a los 24 años, determinar los valores de r y λ.

c) Verificar que si X es una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si} \quad 0 < x < b \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

con $a, b \in \mathbb{R},\ a < b$, entonces cualquier $\zeta \in (a,b)$ es una moda de X.

d) Verificar que si X es una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} \frac{1}{2} x^{-\frac{3}{2}} \exp\left(-x^{\frac{1}{2}}\right) & \text{si} \quad x > 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

entonces la moda de X no existe.

50. Supóngase que el tiempo (en minutos) que dura una llamada teléfonica es una variable aleatoria con función de densidad dada por:

$$f(t) = \begin{cases} \frac{1}{5}\exp\left(-\frac{t}{5}\right) & \text{si} \quad t > 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

Sea $C(t)$ el monto (en pesos) que debe pagar un usuario por una llamada de t minutos de duración. Si

$$C(t) = \begin{cases} 500 & \text{si} \quad 0 < t \leq 5 \\ 750 & \text{si} \quad 5 < t \leq 10 \\ 100t & \text{si} \quad t > 10 \end{cases}$$

Calcular el costo medio de una llamada.

Capítulo 3

Algunas distribuciones discretas

En este capítulo se presentan algunas de las distribuciones discretas de uso más frecuente.

3.1. Distribuciones discreta uniforme, binomial y de Bernoulli

Definición 3.1 (distribución discreta uniforme) *Se dice que una variable aleatoria X tiene distribución discreta uniforme de parámetro N, donde N es un entero positivo, si su función de densidad está dada por:*

$$f(x) = \begin{cases} \frac{1}{N} & si \quad x = 1, 2, \cdots, N \\ 0 & en \quad otro\ caso \end{cases}$$

La función de densidad tiene la forma que presenta la figura 3.1.

Teorema 3.2 (propiedades de la distribución discreta) *Si X es una variable aleatoria con distribución discreta uniforme de parámetro N entonces:*

1. $EX = \frac{N+1}{2}$.
2. $Var(X) = \frac{(N^2-1)}{12}$.
3. $m_X(t) = \sum_{k=1}^{N} \frac{1}{N} e^{tk}$.

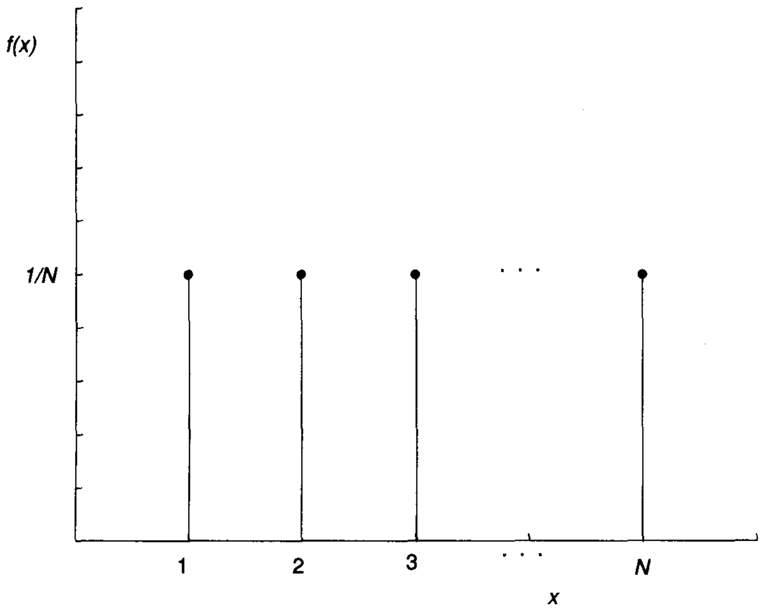


Figura 3.1: Función de densidad de una distribución uniforme discreta

Demostración.

1.
$$EX = \sum_{k=1}^{N} \frac{k}{N} = \frac{1}{N} \times \frac{N(N+1)}{2} = \frac{N+1}{2}.$$

2. Como ejercicio.

3. Inmediata a partir de la definición de f.g.m.

■

Definición 3.3 (distribuciones binomial y Bernoulli) *Se dice que una variable aleatoria X tiene distribución binomial de parámetros n y p, si su función de densidad está dada por:*

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & si \quad x = 0, 1, \cdots, n \\ 0 & en \quad otro\ caso \end{cases}$$

donde n es un entero positivo y $0 < p < 1$.

Si $n = 1$, la distribución binomial recibe el nombre de distribución Bernoulli de parámetro p.

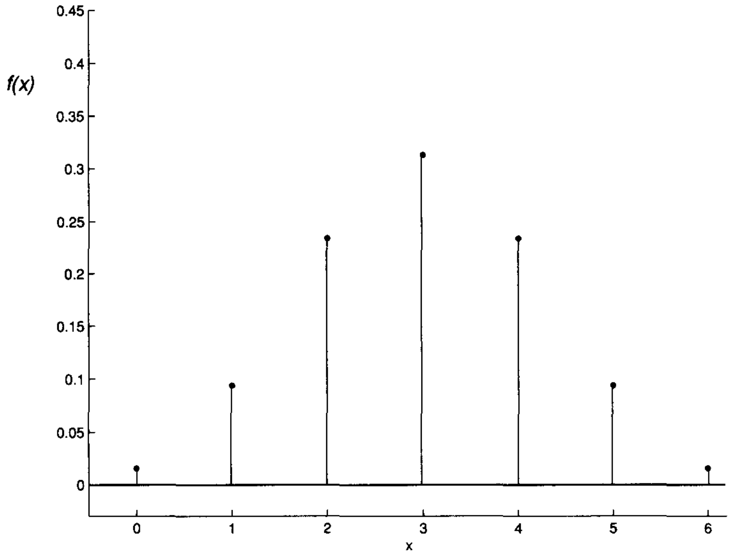


Figura 3.2: Función de densidad de una distribución binomial con parámetros $n = 6, \quad p = 0.5$.

Notación 3.4 *Se escribe* $X \stackrel{d}{=} \mathcal{B}(n, p)$ *para indicar que la variable aleatoria* X *tiene distribución binomial de parámetros* n *y* p.

La distribución binomial se usa frecuentemente en la descripción de aquellos experimentos en los que el resultado es la ocurrencia o no ocurrencia de un suceso. Si la variable aleatoria X denota el número de éxitos en n ensayos independientes del experimento, entonces X tiene distribución binomial de parámetros n y p; siendo p la probabilidad de éxito, es decir, la probabilidad de ocurrencia del suceso. Se acostumbra denotar la probabilidad de fracaso $(1 - p)$ con la letra q.

Ejemplo 3.5 *Se lanza un dado corriente cinco veces consecutivas. Sea* X *la variable aleatoria que denota el número de veces que se obtiene el número* 5 *como resultado. Hallar la función de densidad de* X.

Solución: La variable aleatoria X tiene distribución binomial de parámetros 5 y $\frac{1}{6}$. Por lo tanto,

$$P(X = 0) = \binom{5}{0} \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^5 = 0.40188$$

$$\begin{aligned}
P(X=1) &= \binom{5}{1} \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^4 = 0.40188 \\
P(X=2) &= \binom{5}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^3 = 0.16075 \\
P(X=3) &= \binom{5}{3} \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^2 = 0.03215 \\
P(X=4) &= \binom{5}{4} \times \left(\frac{1}{6}\right)^4 \times \left(\frac{5}{6}\right) = 3.215 \times 10^{-3} \\
P(X=5) &= \binom{5}{5} \times \left(\frac{1}{6}\right)^5 \times \left(\frac{5}{6}\right)^0 = 1.286 \times 10^{-4}. \quad \blacktriangle
\end{aligned}$$

Ejemplo 3.6 *Un vendedor de planes de turismo al Caribe, sabe, por experiencia, que la oportunidad de vender un plan es mayor, mientras más contactos realice con los clientes potenciales. Empíricamente él estableció que la probabilidad de que una persona compre un plan, luego de su visita, es constante e igual a* 0.01. *Si el conjunto de visitas que realiza el vendedor constituye un conjunto independiente de ensayos, ¿cuántos compradores potenciales debe visitar para que la probabilidad de vender por lo menos un plan sea igual a* 0.85*?*

Solución: Sea

$$X := \text{“Número de personas que compran un plan luego de la visita del vendedor”}$$

Se tiene que $X \stackrel{d}{=} \mathcal{B}(n, 0.1)$. Se busca determinar n de tal manera que

$$P(X \geq 1) = 0.85$$

o equivalentemente

$$\begin{aligned}
0.85 &= 1 - P(X=0) \\
&= 1 - \binom{n}{0}(0.01)^0(0.99)^n.
\end{aligned}$$

Entonces:

$$\begin{aligned}
(0.99)^n &= 0.15 \\
n \ln(0.99) &= \ln(0.15) \\
n &= \frac{\ln(0.15)}{\ln(0.99)} \approx 188.76.
\end{aligned}$$

Por lo tanto, el vendedor debe visitar por lo menos a 189 personas para lograr su objetivo. ▲

Ejemplo 3.7 *La mejor amiga de Paula la invitó a una fiesta. Como Paula es aún muy pequeña, sus padres condicionaron el permiso a que su hermano la acompañe. El hermano de Paula le propone el siguiente trato: "tú escoges un número, el que quieras, entre* 1 *y* 6*; luego lanzamos cuatro veces un dado corriente, si el número que tú escogiste aparece por lo menos* 2 *veces, entonces, te acompaño a la fiesta. En caso contrario no te acompaño". ¿Cuál es la probabilidad de que Paula pueda ir a la fiesta?*

Solución: Sea

$$X := \text{"número de veces que aparece el número escogido por Paula"}$$

Es claro que $X \stackrel{d}{=} \mathcal{B}(4, \frac{1}{6})$. Entonces, la probabilidad pedida p es igual a:

$$\begin{aligned} p &= P(X \geq 2) \\ &= 1 - P(X=0) - P(X=1) \\ &= 1 - \binom{4}{0} \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^4 - \binom{4}{1} \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^3 \\ &= 0.13194. \quad \blacktriangle \end{aligned}$$

Ejemplo 3.8 *Un jugador apuesta a uno de los números del* 1 *al* 6*. Una vez apuesta, se lanzan tres dados corrientes. Si el número apostado por el jugador aparece i veces; con $i = 1, 2, 3$, entonces el jugador gana $2i$ unidades monetarias. Si el número apostado por el jugador no aparece en ninguno de los dados, el jugador pierde* 3 *unidades monetarias. ¿Éste es un juego justo para el jugador?. Explique su respuesta.*

Solución: Sea X la variable aleatoria que denota la fortuna del jugador. Los valores que puede tomar X son $-3, 2, 4$ y 6. Es claro que:

$$\begin{aligned} P(X = -3) &= \binom{3}{0} \times \left(\frac{1}{6}\right)^0 \times \left(\frac{5}{6}\right)^3 = \frac{125}{216} \\ P(X = 2) &= \binom{3}{1} \times \left(\frac{1}{6}\right)^1 \times \left(\frac{5}{6}\right)^2 = \frac{75}{216} \\ P(X = 4) &= \binom{3}{2} \times \left(\frac{1}{6}\right)^2 \times \left(\frac{5}{6}\right)^1 = \frac{15}{216} \\ P(X = 6) &= \binom{3}{3} \times \left(\frac{1}{6}\right)^3 \times \left(\frac{5}{6}\right)^0 = \frac{1}{216}. \end{aligned}$$

Por lo tanto:

$$EX = \frac{(-3) \times 125 + 2 \times 75 + 4 \times 15 + 6}{216} = \frac{-159}{216}.$$

Esto indica que, a la postre, el jugador pierde 159 unidades monetarias por cada 216 juegos que realice. Por lo tanto, el juego no le es favorable.▲

Ejemplo 3.9 *Supóngase que n bolas se distribuyen al azar en r urnas. Hallar la probabilidad de que haya exactamente k bolas en las r_1 primeras urnas.*

Solución: Sea $X :=$ "número de bolas en las primeras r_1 urnas". Puesto que $X \stackrel{d}{=} \mathcal{B}(n,p)$ con $p = \frac{r_1}{r}$ entonces:

$$P(X = k) = \binom{n}{k} \left(\frac{r_1}{r}\right)^k \left(1 - \frac{r_1}{r}\right)^{n-k}. \quad \blacktriangle$$

A continuación se presentan las propiedades más importantes de la distribución binomial.

Teorema 3.10 (propiedades de la distribución binomial) *Sea X una variable aleatoria con distribución binomial de parámetros n y p. Entonces:*

1. $EX = np$.
2. $Var(X) = npq$; *donde* $q := 1 - p$.
3. $m_X(t) = \left(pe^t + q\right)^n$.

Demostración. En el ejemplo 2.63 se verificó que la f.g.m de una variable aleatoria con distribución binomial de parámetros n y p, está dada por $m_X(t) = \left(pe^t + q\right)^n$.

Por lo tanto:

$$EX = \frac{d}{dt} m_X(t) \mid_{t=0} = npe^t(pe^t + q)^{n-1} \mid_{t=0} = np$$

Además se tiene que:

$$\begin{aligned} EX^2 &= \left.\frac{d^2}{dt^2} m_X(t)\right|_{t=0} \\ &= \left.\left[n(n-1)(pe^t)^2(pe^t + q)^{n-2} + npe^t(pe^t + q)^{n-1}\right]\right|_{t=0} \\ &= n(n-1)p^2 + np \end{aligned}$$

de donde se deduce:

$$\begin{aligned}Var(X) &= n^2p^2 - np^2 + np - n^2p^2 \\ &= np(1-p) \\ &= npq.\end{aligned}$$

■

Nota 3.11 *Supóngase que X es una variable aleatoria con distribución binomial de parámetros n y p. Sea*

$$B(k) := \binom{n}{k} p^k q^{n-k}$$

Puesto que:

$$\binom{n}{k} = \frac{n-k+1}{k}\binom{n}{k-1}$$

se obtiene para $k = 1, \cdots, n$:

$$\begin{aligned}B(k) &= \frac{n-k+1}{k}\binom{n}{k-1} p \times p^{k-1} \times q^{n-k+1} \times \frac{1}{q} \\ &= \frac{n-k+1}{k} \times \frac{p}{q} \times B(k-1).\end{aligned}$$

Luego comenzando con $B(0) = q^n$ se obtienen en forma recurrente los valores de $B(k)$ para $k = 1, \cdots, n$.

Algoritmo 3.12

Entrada: *p, n, siendo n el número de términos.*

Salida: *$B(k)$ para $k = 0(1)n$.*

Inicialización:

$$\begin{aligned}&q := (1-p), \\ &B(0) := q^n.\end{aligned}$$

Iteración:

Para $k = 0(1)n$ haga:

$$B(k+1) = \tfrac{n-k}{k+1} \times \tfrac{p}{q} \times B(k)$$ ▲

Así por ejemplo, con el algoritmo anterior, para $n = 5$ y $p = 0.3$ se obtiene:

k	$B(k)$
0	0.16807
1	0.36015
2	0.30870
3	0.13230
4	0.02835
5	0.00243

Nota 3.13 *De la observación anterior se obtiene que:*

$$B(k) > B(k-1), \text{ si y sólo si, } \frac{n-k+1}{k} \times \frac{p}{q} > 1$$

esto es:

$$B(k) > B(k-1) \text{ si y sólo si } (n+1)p > k$$

Por lo tanto, si $X \stackrel{d}{=} \mathcal{B}(n,p)$, *entonces,* $P(X = k)$ *crece inicialmente de manera monótona y luego decrece monótonamente, alcanzando su valor mayor cuando* $k = [(n+1)p]$ *donde* $[a]$ *denota la parte entera de* a.

3.2. Distribuciones hipergeométrica y Poisson

En el capítulo 1. se vió que si se tiene una urna con N bolas en total de las cuales R son rojas y $(N-R)$ blancas y si se extrae una muestra de tamaño r, sin reemplazo, entonces la probabilidad $P(A_k)$ de que exactamente k de las bolas extraídas son rojas es igual a:

$$P(k) := P(A_k) = \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}}.$$

Si únicamente interesa el número k de bolas rojas entre las n bolas extraídas, se tiene que $P(k)$ define una medida de probabilidad sobre el conjunto $\{0, 1, \cdots, n\}$ llamada distribución hipergeométrica de parámetros n, R y N. Más precisamente se tiene la definición siguiente:

Definición 3.14 (distribución hipergeométrica) *Se dice que una variable aleatoria* X *tiene distribución hipergeométrica de parámetros* n, R *y* N, *si su función de densidad está dada por:*

$$f_X(x) = \begin{cases} \frac{\binom{R}{x}\binom{N-R}{n-x}}{\binom{N}{n}} & \text{si} \quad x = 0, \cdots, n \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

donde N es un entero positivo, R es un entero no negativo menor o igual a N y n es un entero positivo menor o igual a N.

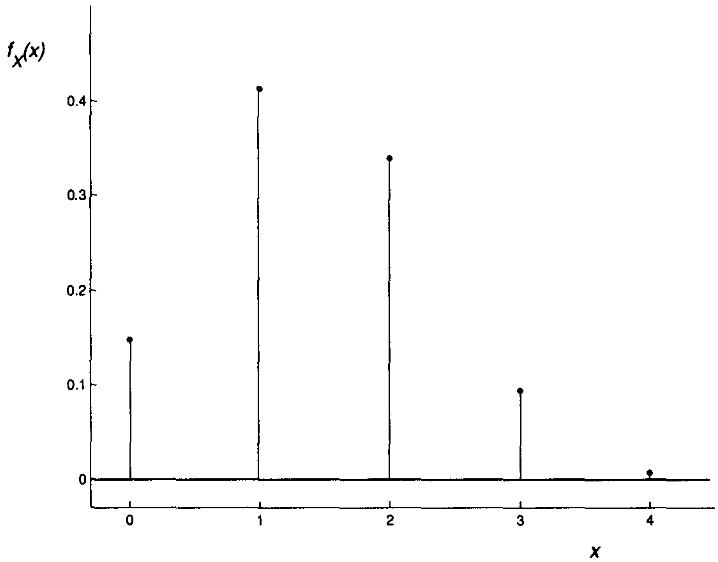


Figura 3.3: Función de densidad de una distribución hipergeométrica con parámetros: $n = 4,\ R = 7,\ N = 20$.

Notación 3.15 *La expresión $X \stackrel{d}{=} Hg(n, R, N)$ significa que la variable aleatoria X tiene una distribución hipergeométrica de parámetros n, R y N.*

Ejemplo 3.16 *La división de vigilancia de una institución universitaria ha adquirido* 50 *equipos de comunicación con el fin de optimizar el servicio en sus predios. Se seleccionan aleatoriamente* 8 *equipos y se someten a prueba para encontrar posibles defectos. Si* 3 *de los* 50 *equipos están defectuosos, ¿cuál es la probabilidad de que la muestra contenga a lo más dos equipos defectuosos?*

Solución: Sea X :=“número de equipos defectuosos en la muestra”. Es claro que $X \stackrel{d}{=} Hg(8, 3, 50)$. Por lo tanto,

$$P\,(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$

$$= \frac{\binom{3}{0}\binom{47}{8}}{\binom{50}{8}} + \frac{\binom{3}{1}\binom{47}{7}}{\binom{50}{8}} + \frac{\binom{3}{2}\binom{47}{6}}{\binom{50}{8}}$$
$$= 0.99714. \quad \blacktriangle$$

Ejemplo 3.17 *Un equipo de trabajo, establecido por el ministerio del medio ambiente, programó visitas a* 25 *fábricas para investigar posibles violaciones a los reglamentos para el control de la contaminación ambiental. Sin embargo, los recortes presupuestales han reducido drásticamente el tamaño del equipo de trabajo, por lo que, solamente se podrán investigar* 5 *de las* 25 *fábricas. Si se sabe que* 10 *de las fábricas están operando sin cumplir los reglamentos, calcule la probabilidad de que al menos una de las fábricas muestreadas esté operando en contravención a los reglamentos.*

Solución: Sea X := "número de fábricas en la muestra que operan sin cumplir los reglamentos". Se tiene que $X \stackrel{d}{=} Hg(5, 10, 25)$. Entonces:

$$\begin{aligned} P(X \geq 1) &= 1 - P(X = 0) \\ &= 1 - \frac{\binom{10}{0} \times \binom{15}{5}}{\binom{25}{5}} \\ &= 0.94348. \quad \blacktriangle \end{aligned}$$

Nota 3.18 *Supóngase que el tamaño N de una población es desconocido. Se desea determinar N sin tener que contar los individuos uno a uno. Un método para hacerlo es el llamado método de captura-recaptura, el cual consiste en capturar R individuos de la población, marcarlos y luego retornarlos a la población. Una vez los individuos, marcados y no-marcados, se mezclan homogéneamente, se toma una muestra de tamaño n. Sea X la variable aleatoria que denota al número de individuos marcados en la muestra. Es claro que: $X \stackrel{d}{=} Hg(n, R, N)$. Supóngase que se observa que X es igual a k. Entonces, $P_k(N) := P(X = k)$ representa la probabilidad de que en la muestra haya exactamente k individuos marcados cuando el tamaño de la población es N. Por lo tanto un estimador $\widehat{N}$ de N es aquel que máximiza la probabilidad de que X sea igual a k. Tal estimador se conoce como estimador de máxima verosimilitud de N. Para determinarlo, se observa que:*

$$P_k(N) \geq P_k(N-1), \text{ si y sólo si, } \frac{(N-R) \times (N-n)}{N \times (N-R-n+k)} \geq 1$$

esto es,

$$P_k(N) \geq P_k(N-1), \text{ si y sólo si, } \frac{R \times n}{k} \geq N$$

por lo tanto,

$$\widehat{N} = \left[\frac{R \times n}{k}\right].$$

Ejemplo 3.19 *Para establecer cuántos peces hay en un lago se procede como sigue: se capturan y marcan* 1000 *peces y luego se devuelven al lago. Días después se capturan de* 150 *peces y se observa que* 10 *de ellos están marcados. Entonces, de acuerdo a la observación anterior, se tiene que el estimador de máxima verosimilitud del tamaño N de la población es:*

$$\widehat{N} = \left[\frac{1000 \times 150}{10}\right] = 15000 \qquad \blacktriangle$$

A continuación se presentan algunas de las propiedades de la distribución hipergeométrica.

Teorema 3.20 (propiedades de la distribución hipergeométrica) *Sea $X \stackrel{d}{=} Hg(n, R, N)$, entonces:*

1. $EX = \frac{nR}{N}$.
2. $Var(X) = n \times \frac{R}{N} \times \frac{N-R}{N} \times \frac{N-n}{N-1}$.

Demostración.

1.

$$\begin{aligned} EX &= \sum_{x=0}^{n} x \frac{\binom{R}{x}\binom{N-R}{n-x}}{\binom{N}{n}} \\ &= \sum_{x=1}^{n} n \frac{R}{N} \frac{\binom{R-1}{x-1}\binom{N-R}{n-x}}{\binom{N-1}{n-1}} \\ &= n \frac{R}{N} \sum_{k=0}^{n-1} \frac{\binom{R-1}{k}\binom{N-R}{n-k-1}}{\binom{N-1}{n-1}} \\ &= n \frac{R}{N}, \end{aligned}$$

puesto que:

$$\sum_{k=0}^{n-1} \binom{R-1}{k} \binom{N-R}{n-k-1} = \binom{N-1}{n-1}.$$

2. Como

$$Var(X) = E(X(X-1)) + EX - (EX)^2$$

se va a calcular $E(X(X-1))$.

Se tiene:

$$\begin{aligned}
E(X(X-1)) &= \sum_{x=0}^{n} x(x-1)P(X=x) \\
&= \sum_{x=2}^{n} \left[x(x-1)\frac{R(R-1)(R-2)!}{(R-x)!x(x-1)(x-2)!} \times \right. \\
&\qquad \left. \frac{(N-n)!n(n-1)(n-2)!}{N(N-1)(N-2)!} \times \binom{N-R}{n-x} \right] \\
&= n(n-1)\frac{R(R-1)}{N(N-1)} \sum_{x=2}^{n} \left[\frac{\binom{R-2}{x-2}\binom{N-R}{n-x}}{\binom{N-2}{n-2}} \right] \\
&= n(n-1)\frac{R(R-1)}{N(N-1)} \sum_{k=0}^{n-2} \left[\frac{\binom{R-2}{k}\binom{N-R}{n-k-2}}{\binom{N-2}{n-2}} \right] \\
&= n(n-1)\frac{R(R-1)}{N(N-1)}
\end{aligned}$$

pues

$$\sum_{k=0}^{n-2} \binom{R-2}{k}\binom{N-R}{n-k-2} = \binom{N-2}{n-2}.$$

Por lo tanto,

$$\begin{aligned}
Var(X) &= n(n-1)\frac{R(R-1)}{N(N-1)} + \frac{nR}{N} - \frac{n^2R^2}{N^2} \\
&= n\frac{R}{N}\left[\frac{(N-R)(N-n)}{N(N-1)}\right].
\end{aligned}$$

■

A continuación se verá que si el tamaño de la población N es grande, en comparación con el tamaño de la muestra n, entonces, la distribución hipergeométrica puede aproximarse por una distribución binomial. Más precisamente se tiene el siguiente resultado:

Teorema 3.21 *Sea $0 < p < 1$. Si $N, R \longrightarrow \infty$ de tal forma que $\frac{R}{N} \longrightarrow p$, entonces:*

$$Hg(n,R,N)(k) := \frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}} \longrightarrow \mathcal{B}(n,p)(k) := \binom{n}{k} p^k (1-p)^{n-k}.$$

Demostración. Puesto que $\frac{N-R}{N} \longrightarrow (1-p) = q > 0$ cuando $N, R \longrightarrow \infty$, entonces, $(N-R) \longrightarrow \infty$ cuando $N \longrightarrow \infty$. Por lo tanto:

$$\begin{aligned}
&\frac{\binom{R}{k}\binom{N-R}{n-k}}{\binom{N}{n}} \\
&= \binom{n}{k} \frac{R(R-1)\cdots(R-k+1)(N-R)\cdots(N-R-(n-k)+1)}{N(N-1)\cdots(N-n+1)} \\
&= \binom{n}{k}\left(\frac{R}{N}\right)^k \left(\frac{N-R}{N}\right)^{n-k} \frac{R(R-1)\cdots(R-k+1)}{R^k} \times \\
&\quad \frac{(N-R)\cdots(N-R-(n-k)+1)}{(N-R)^{n-k}} \times \frac{N^n}{N(N-1)\cdots(N-n+1)} \\
&\underset{N,R\longrightarrow\infty}{\longrightarrow} \binom{n}{k} p^k (1-p)^{n-k}.
\end{aligned}$$

■

Ejemplo 3.22 *En una ciudad de 2 millones de habitantes, se tiene que el 60 % pertenecen al partido político A. Se eligen al azar 100 personas. La distribución del número de personas, entre los 100 elegidos, que pertenecen al partido A es una distribución hipergeométrica de parámetros 100, 1′200.000 y 2′000.000. Aplicando el resultado anterior se puede aproximar esta distribución por una binomial de parámetros $n = 1000$ y $p = 0.6$. Así, por ejemplo, la probabilidad de que entre los 100 elegidos exactamente 40 pertenezcan al partido A es igual a:*

$$\binom{100}{40}(0.6)^{40}(0.4)^{60} = 2.4425 \times 10^{-5}. \quad \blacktriangle$$

La tabla siguiente compara las distribuciones binomial e hipergeométrica.

k	$Hg(4, 60, 100)$	$Hg(4, 600, 1000)$	$Hg(4, 6000, 10000)$	$\mathcal{B}\left(4, \frac{3}{5}\right)$
0	0.02331	0.02537	0.02558	0.0256
1	0.15118	0.15337	0.15358	0.1536
2	0.35208	0.34624	0.34566	0.3456
3	0.34907	0.34595	0.34563	0.3456
4	0.12436	0.12908	0.12955	0.1296

Definición 3.23 (distribución Poisson) *Se dice que una variable aleatoria X tiene distribución Poisson de parámetro $\lambda > 0$, si su función de densidad está dada por:*

$$f(x) = \begin{cases} e^{-\lambda}\frac{\lambda^x}{x!} & si \quad x = 0, 1, \cdots \\ 0 & en \quad otro\ caso \end{cases}$$

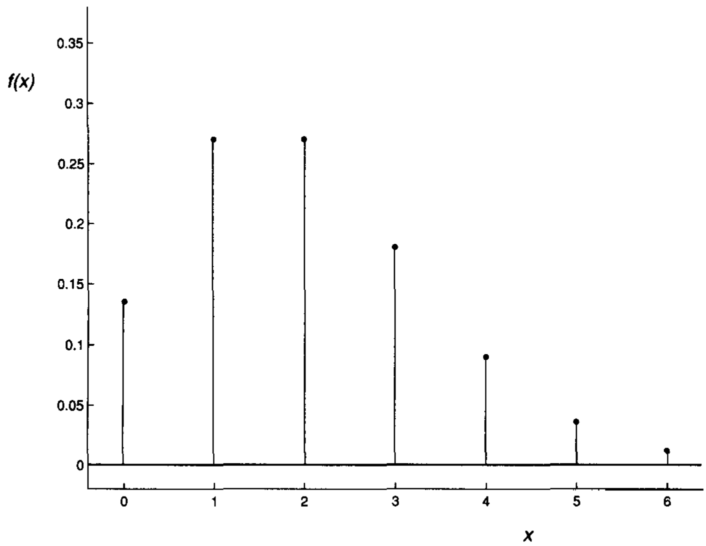


Figura 3.4: Función de densidad de una distribución Poisson con parámetro $\lambda = 2$.

Notación 3.24 *Sea X una variable aleatoria. Se escribe $X \stackrel{d}{=} \mathcal{P}(\lambda)$ para indicar que X tiene una distribución Poisson de parámetro λ.*

Nota 3.25 *Sea* $X \overset{d}{=} \mathcal{P}(\lambda)$ *y sea* $\mathcal{P}(\lambda)(k) := P(X = k)$. *Es fácil verificar que para* $k = 1, 2, \cdots$

$$\mathcal{P}(\lambda)(k) = \frac{\lambda}{k}\mathcal{P}(\lambda)(k-1),$$

luego, comenzando con $\mathcal{P}(\lambda)(0) = e^{-\lambda}$ *se obtienen, recurrentemente, los valores de* $\mathcal{P}(\lambda)(k)$ *para* $k = 1, 2, \cdots$

Un algoritmo sencillo para calcular los valores de la distribución Poisson, que utiliza la fórmula anterior, está dado por:

Algoritmo 3.26 (Cálculo de $\mathcal{P}(\lambda)(k)$)

Entrada: λ, n, *siendo* n *el númerode términos.*

Salida: $\mathcal{P}(\lambda)(k)$ *para* $k = 0(1)n$.

Inicialización: $\mathcal{P}(\lambda)(0) = e^{-\lambda}$.

Iteración:

Para $k = 0(1)n$ *haga:*

$\mathcal{P}(\lambda)(k+1) = \frac{\lambda}{k+1}\mathcal{P}(\lambda)(k)$. ▲

Al analizar la gráfica de la función de densidad de una variable aleatoria con distribución Poisson, se observa que las probabilidades individuales son cada vez más pequeñas a medida que la variable toma valores cada vez más grandes. Ésta es, precisamente, una característica general de la distribución Poisson. Otras características de la distribución Poisson están dadas en el teorema siguiente.

Teorema 3.27 *Sea* X *una variable aleatoria con distribución Poisson de parámetro* λ. *Entonces:*

1. $EX = \lambda$.
2. $Var(X) = \lambda$.
3. $m_X(t) = \exp(\lambda(e^t - 1))$.

Demostración. Se hace, inicialmente, la demostración de 3. Luego, aplicando las propiedades de la f.g.m, se deducen 1. y 2. Se tiene que:

$$\begin{aligned}
m_X(t) &= \sum_{k=0}^{\infty} e^{tk} e^{-\lambda} \frac{\lambda^k}{k!} \\
&= e^{-\lambda} \sum_{k=0}^{\infty} \frac{\left(\lambda e^t\right)^k}{k!} \\
&= \exp(\lambda(e^t - 1)).
\end{aligned}$$

por lo tanto,

$$\begin{aligned}
EX &= \frac{d}{dt} m_X(t) \mid_{t=0} \\
&= \lambda e^t \exp\left(\lambda(e^t - 1)\right) \mid_{t=0} \\
&= \lambda
\end{aligned}$$

y

$$\begin{aligned}
EX^2 &= \frac{d^2}{dt^2} m_X(t) \mid_{t=0} \\
&= \left[\lambda e^t \exp\left(\lambda(e^t - 1)\right) + \lambda^2 e^{2t} \exp\left(\lambda(e^t - 1)\right)\right] \mid_{t=0} \\
&= \lambda(\lambda + 1),
\end{aligned}$$

Entonces,

$$Var(X) = \lambda.$$

■

Ejemplo 3.28 *El número de personas que ingresan diariamente a la unidad de urgencias de un hospital, tiene una distribución Poisson de media* 10. *¿A qué es igual la probabilidad de que, en un día en particular, el número de pacientes que ingresen a la unidad de urgencias, en dicho hospital, sea menor o igual a* 3*?*

Solución: Sea $X :=$ "número de pacientes que ingresan diariamente a la unidad de urgencias". Por los datos del problema, se sabe que $X \stackrel{d}{=} \mathcal{P}(10)$. Por lo tanto,

$$\begin{aligned}
P(X \leq 3) &= P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3) \\
&= e^{-10} + 10e^{-10} + 50e^{-10} + \frac{1000}{6} e^{-10} \\
&= 1.0336 \times 10^{-2}. \quad \blacktriangle
\end{aligned}$$

Se verá a continuación que la distribución de Poisson es una forma límite de la distribución binomial, cuando n es suficientemente grande y p suficiente y adecuadamente pequeño.

Teorema 3.29 *Si $p(n)$ es una sucesión con $0 < p(n) < 1$ y $n\,(p(n)) \longrightarrow \lambda$ cuando $n \longrightarrow \infty$, entonces*

$$\mathcal{B}_{n,p(n)}(k) := \binom{n}{k} (p(n))^k (1-p(n))^{n-k} \underset{n\to\infty}{\longrightarrow} e^{-\lambda}\frac{\lambda^k}{k!} =: \mathcal{P}(\lambda)(k)$$

Demostración. Sea $\lambda_n = n\,(p(n))$ entonces

$$\mathcal{B}_{n,p(n)}(k) = \frac{1}{k!} \times \frac{n}{n} \times \cdots \times \frac{n-k+1}{n} \times \lambda_n^k \times \left(1-\frac{\lambda_n}{n}\right)^n \times \left(1-\frac{\lambda_n}{n}\right)^{-k}$$

Tomando el límite cuando $n \to \infty$ se tiene que los cocientes $\frac{n}{n}, \cdots, \frac{n-k+1}{n}$ y el factor $\left(1-\frac{\lambda_n}{n}\right)^{-k}$ tienden a 1, mientras que el factor $\left(1-\frac{\lambda_n}{n}\right)^n$ tiende a $e^{-\lambda}$.

Por lo tanto,

$$\mathcal{B}_{n,p(n)}(k) \underset{n\to\infty}{\longrightarrow} \mathcal{P}(\lambda)(k)$$

■

En la tabla siguiente se observa la "bondad" de la aproximación cuando $\lambda = np = 1$:

k	$\mathcal{P}(\lambda)(k)$	$\mathcal{B}_{100,\frac{1}{100}}(k)$	$\mathcal{B}_{10,\frac{1}{10}}(k)$
0	0.3678	0.3660	0.3486
1	0.3678	0.3697	0.3874
2	0.1839	0.1848	0.1937
3	0.0613	0.0609	0.0574
4	0.0153	0.0149	0.0112

El teorema anterior implica que la distribución de Poisson ofrece un modelo probabilístico adecuado para todos aquellos experimentos aleatorios en los que las repeticiones son independientes unas de otras y en los que sólo hay dos posibles resultados: éxito o fracaso, con probabilidad de éxito pequeña, y en los que el interés se centra en conocer el número de éxitos obtenidos al realizar el experimento un número suficientemente grande de veces. Empíricamente se ha establecido, que la aproximación se puede aplicar con seguridad si $n \geq 100$, $p \leq 0.01$ y $np \leq 20$.

Ejemplo 3.30 *En un auditorio se encuentran* 135 *estudiantes. La probabilidad de que uno de los estudiantes se encuentre hoy de cumpleaños es igual a* $\frac{1}{365}$. *¿Cuál es la probabilidad de que en el auditorio dos o más estudiantes estén hoy de cumpleaños?*

Solución: Sea $X :=$ "número de estudiantes que están de cumpleaños hoy". Se sabe que $X \stackrel{d}{=} \mathcal{B}(135, \frac{1}{365})$. Ésta última distribución se puede aproximar por una Poisson con parámetro $\lambda = \frac{27}{73}$. Por lo tanto,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - e^{-\frac{27}{73}} - \frac{27}{73}e^{-\frac{27}{73}} \\ &= 5.3659 \times 10^{-2}. \quad \blacktriangle \end{aligned}$$

Una aplicación, muy importante, de la distribución Poisson se presenta en relación con la ocurrencia de un cierto tipo de eventos en un intervalo de tiempo determinado. Así por ejemplo, la distribución Poisson ha sido utilizada para describir la distribución del número de partículas α que llegan a un determinado punto del espacio, durante un período de tiempo t, y que son emitidas por una sustancia radioactiva. También se ha usado para hallar la distribución del número de individuos que llegan a una línea de espera, en un período de tiempo t. En estos casos se supone que se inicia el conteo en el tiempo $t = 0$ y que se hacen las suposiciones siguientes en relación con la llegada de los individuos (partículas):

1. Existe un parámetro $\lambda > 0$ tal que, para cualquier intervalo de longitud pequeña Δt, se tiene que la probabilidad de que llegue exactamente un individuo (partícula) es igual a $\lambda\Delta t + o(\Delta t)$; donde $o(\Delta t)$ es una cantidad tal que $\frac{o(\Delta t)}{\Delta t} \to 0$ cuando $\Delta t \to 0$.

2. La probabilidad de que no lleguen individuos (partículas) en un intervalo de longitud pequeña Δt es igual a $1 - \lambda\Delta t + o(\Delta t)$.

3. La probabilidad de que llegue más de un individuo (partícula) en un intervalo de longitud pequeña Δt es del orden $o(\Delta t)$.

4. El número de individuos (partículas) que llegan durante un intervalo de tiempo es independiente del número de individuos (partículas) que llegan antes de éste..

Si X_t es la variable aleatoria que denota el número de individuos (partículas) que llegan en el intervalo de tiempo $(0, t]$, entonces, se tiene que (ver[Ros]) X_t tiene una distribución Poisson con parámetro λt.

Ejemplo 3.31 *Supóngase que el número de llamadas que entran a una central teléfonica es de 30 por hora. ¿Cuál es la probabilidad de que no se reciban llamadas en un período de 3 minutos?. ¿Cuál es la probabilidad de que se reciban más de cinco llamadas en un intervalo de 5 minutos?*

Solución: Sea $X_t :=$ "número de llamadas que se reciben en el intervalo de tiempo $(0, t]$", donde el tiempo está medido en minutos.

Por los datos del problema, $X_t \stackrel{d}{=} \mathcal{P}(\frac{1}{2}t)$. Luego:

$$P(X_3 = 0) = e^{-0.5\times 3} = 0.22313$$

y

$$P(X_5 \geq 6) = \sum_{k=6}^{\infty} e^{-2.5}\frac{(2.5)^k}{k!} = 0.042. \quad \blacktriangle$$

3.3. Distribuciones geométrica y binomial negativa.

Al trabajar la distribución binomial, se procedió como sigue: se repitió el experimento aleatorio n veces y se calculó la probabilidad de obtener de manera exacta k éxitos. En este caso, el número de repeticiones permanece constante mientras que el número de éxitos es aleatorio. Supóngase que ahora la pregunta es la siguiente: ¿cuál es la probabilidad de tener que repetir el experimento n veces para obtener de manera exacta k éxitos?. Es decir, ahora el número de éxitos permanece constante en tanto que el número de repeticiones X es una variable aleatoria.

¿A qué es igual la probabilidad de que X tome el valor j con $j = k, k+1, \cdots$? Si $X = j$ entonces el j-ésimo resultado es necesariamente un éxito, por lo tanto los restantes $(k-1)$ éxitos se obtienen en las restantes $(j-1)$ repeticiones del experimento. Esto es;

$$P(X = j) = \binom{j-1}{k-1} p^k (1-p)^{j-k}; \ j = k, k+1, \cdots$$

donde $0 < p < 1$ es la probabilidad de éxito.

Definición 3.32 (distribución binomial negativa y geométrica) *Se dice que una variable aleatoria X tiene distribución binomial negativa de parámetros k y p, si su función de densidad está dada por:*

$$f(x) = \begin{cases} \binom{x-1}{k-1} p^k (1-p)^{x-k} & si \quad x = k, k+1, \cdots \\ 0 & en \quad otro\ caso \end{cases}$$

En el caso especial $k = 1$, se dice que la variable aleatoria tiene distribución geométrica de parámetro p.

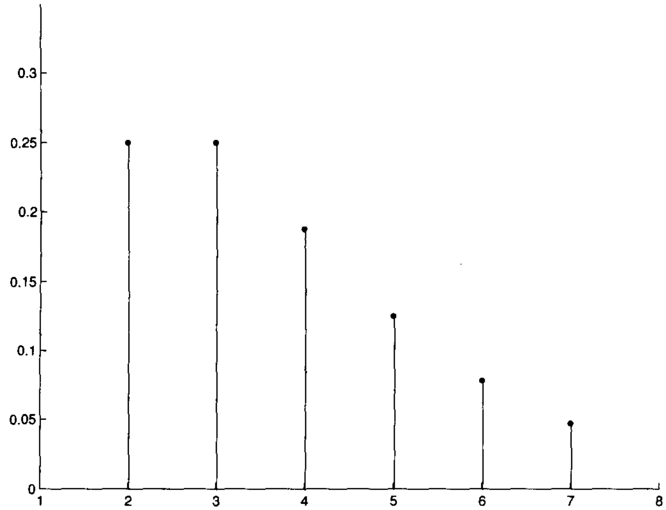


Figura 3.5: Función de densidad de una distribución binomial negativa con parámetros $k = 2$ y $p = \frac{1}{2}$.

Notación 3.33 *Las expresiones $X \overset{d}{=} \mathcal{B}_N(k,p)$ y $X \overset{d}{=} \mathcal{G}(p)$ indican que X tiene distribución binomial negativa de parámetros k y p y X tiene distribución geométrica de parámetro p, respectivamente.*

Nota 3.34 *Supóngase que no interesa conocer el número de ensayos necesarios hasta obtener de manera exacta k éxitos si no el número de fracasos Y ocurridos antes de obtener de manera exacta k éxitos. En este caso, se tiene que $X = k + Y$ y por lo tanto:*

$$P(Y = j) = \binom{k+j-1}{k-1} p^k (1-p)^j \quad con \;\; j = 0, 1, 2, \cdots$$

Algunos autores llaman a esta última la distribución binomial negativa y a la dada en la definición anterior la llaman distribución Pascal.

Ejemplo 3.35 *En un departamento de control de calidad se inspeccionan las unidades terminadas que provienen de una línea de ensamble. Si la proporción de unidades defectuosas es de 0.03, ¿cuál es la probabilidad de que la vigésima unidad inspeccionada sea la tercera que se encuentra defectuosa?*

Solución: *Sea* X := *"número de unidades que es necesario inspeccionar hasta obtener de exactamente tres unidades defectuosas".*

Es claro que $X \stackrel{d}{=} \mathcal{B}_N(3, 0.03)$. *Por lo tanto,*

$$\begin{aligned} P(X = 20) &= \binom{19}{2} (0.03)^3 (1 - 0.03)^{17} \\ &= 2.7509 \times 10^{-3}. \quad \blacktriangle \end{aligned}$$

A continuación se va a calcular el valor esperado y la varianza de una variable aleatoria con distribución binomial negativa.

Teorema 3.36 *Sea* X *una variable aleatoria con distribución binomial negativa de parámetros* k *y* p. *Entonces:*

1. $EX = \frac{k}{p}$.

2. $Var(X) = \frac{k(1-p)}{p^2}$.

3. $m_X(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]^k$.

Demostración. El $r-$ésimo momento de X, alrededor de 0, está dado por:

$$\begin{aligned} EX^r &= \sum_{j=k}^{\infty} j^r \binom{j-1}{k-1} p^k (1-p)^{j-k} \\ &= \frac{k}{p} \sum_{j=k}^{\infty} j^{r-1} \binom{j}{k} p^{k+1} (1-p)^{j-k} \\ &= \frac{k}{p} \sum_{n=k+1}^{\infty} (n-1)^{r-1} \binom{n-1}{k} p^{k+1} (1-p)^{n-1-k} \\ &= \frac{k}{p} E\left((Y-1)^{r-1}\right) \end{aligned}$$

donde $Y \stackrel{d}{=} \mathcal{B}_N(k+1, p)$. Por lo tanto,

$$EX = \frac{k}{p}$$

y

$$\begin{aligned}
EX^2 &= \frac{k}{p}E(Y-1) \\
&= \frac{k}{p} \times \left[\frac{k+1}{p} - 1\right] \\
&= \frac{k}{p} \times \frac{k+1-p}{p}.
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
Var(X) &= \frac{k^2+k-kp}{p^2} - \frac{k^2}{p^2} \\
&= \frac{k(1-p)}{p^2}.
\end{aligned}$$

Por otra parte,

$$\begin{aligned}
m_X(t) &= \sum_{j=k}^{\infty} e^{tj}\binom{j-1}{k-1}p^k(1-p)^{j-k} \\
&= e^{tk}\sum_{l=0}^{\infty} e^{tl}\binom{l+k-1}{k-1}p^k(1-p)^l \\
&= e^{tk}\sum_{l=0}^{\infty}(-1)^l\binom{-k}{l}p^k\left((1-p)\,e^t\right)^l
\end{aligned}$$

pues

$$\begin{aligned}
\binom{-k}{l} &:= (-1)^l\frac{k(k+1)\cdots(k+l-1)}{l!} \\
&= (-1)^l\binom{l+k-1}{k-1}.
\end{aligned}$$

Haciendo uso del desarrollo en serie de Taylor de la función $g(x) := (1-x)^{-k}$, alrededor de 0, se obtiene:

$$(1-x)^{-k} = \sum_{l=0}^{\infty}\binom{-k}{j}(-x)^l.$$

Por consiguiente:

$$
\begin{aligned}
m_X(t) &= \left(pe^t\right)^k \left(1-(1-p)\,e^t\right)^{-k} \\
&= \left[\frac{pe^t}{1-(1-p)e^t}\right]^k .
\end{aligned}
$$

■

Corolario 3.37 *Si X es una variable aleatoria con distribución geométrica de parámetro p entonces:*

1. $EX = \frac{1}{p}$.
2. $Var(X) = \frac{1-p}{p^2}$.
3. $m_X(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]$.

Nota 3.38 *Para la variable aleatoria $Y = X - k$ se obtiene que:*

$$
\begin{aligned}
EY &= \frac{k}{p} - k = \frac{k(1-p)}{p} \\
Var(Y) &= Var(X) = \frac{k(1-p)}{p^2} \\
m_Y(t) &= \left[\frac{p}{1-(1-p)e^t}\right]^k
\end{aligned}
$$

Nota 3.39 *Supóngase que el tamaño N de una población es desconocido. Se desea determinar N sin tener que contar los individuos uno a uno. Un método para hacerlo es el llamado método inverso de captura-recaptura, el cual consiste en capturar R individuos de la población, marcarlos y luego retornarlos a la población. Una vez los individuos marcados y no-marcados se mezclan homogéneamente, se extraen individuos de la población hasta que un número predeterminado de individuos marcados, digamos k, se obtiene. Sea X la variable aleatoria que denota el número de extracciones necesarias hasta obtener k individuos marcados. Es claro que: $X \stackrel{d}{=} \mathcal{B}_N(k,p)$; donde $p = \frac{R}{N}$. Supóngase que se observa que X es igual a j. Entonces, $P_k(N) := P(X = j)$ representa la probabilidad de que haya sido necesario hacer j extracciones para obtener exactamente k individuos marcados, cuando el tamaño de la población es N. Por lo tanto, un estimador $\widehat{N}$ de N es aquel* *que maximiza la probabilidad de que X sea igual a j. Tal estimador, llamado estimador de máxima verosimilitud de N, es igual a*

$$\widehat{N} = \left[\frac{R \times j}{k}\right].$$

Ejemplo 3.40 *Para establecer cuántos peces hay en un lago se procede como sigue: se capturan y marcan* 1000 *peces y se luego se devuelven al lago. Días después se capturan peces hasta obtener* 15 *peces marcados. Si se necesitaron* 120 *extracciones para obtener los* 15 *peces marcados, entonces, de acuerdo a la observación anterior, se tiene que el estimador de máxima verosimilitud del tamaño N de la población es:*

$$\widehat{N} = \left[\frac{1000 \times 120}{15}\right] = 8000. \quad \blacktriangle$$

Nota 3.41 *Sean X una variable aleatoria con distribución binomial negativa con parámetros k y p y Y una variable aleatoria con distribución binomial de parámetros n y p. Entonces:*

$$\begin{aligned} P(X > n) &= \sum_{j=n+1}^{\infty} P(X = j) \\ &= \sum_{j=n+1}^{\infty} \binom{j-1}{k-1} p^k (1-p)^{j-k} \\ &= \sum_{j=0}^{k-1} \binom{n}{j} p^j (1-p)^{n-j} \\ &= P(Y < k). \end{aligned}$$

3.4. Ejercicios

1. Se lanza una moneda corriente un número par de veces. ¿A qué es igual la probabilidad de que la mitad de las veces se obtenga como resultado "cara"?

2. En un club nacional de automovilistas comienza una campaña telefónica con el propósito de aumentar su número de miembros. Con base en la experiencia previa, se sabe que una de cada 10 personas que reciben la llamada se une al club. Si en un día 20 personas reciben la llamada teléfonica, cuál es la probabilidad de que por lo menos dos de ellas se inscriban al club?, ¿cuál es el número esperado?

3. Se lanza un dado corriente 10 veces consecutivas. Calcular la probabilidad de obtener por lo menos 3 veces el número 5.

4. En una sala de cómputo, de un centro educativo, hay 20 computadores. La probabilidad de que en horas pico un computador esté ocupado es de 0.9. ¿Cuál es la probabilidad de encontrar al menos un computador desocupado en horas pico?, ¿cuál es la probabilidad de que todos los computadores estén ocupados en horas pico?

5. Supóngase que el número de clientes X que llegan a un banco en una hora es una variable aleatoria Poisson, y que $P(X = 0) = 0.02$. Calcular la media y la varianza de X.

6. Sea X una variable aleatoria con distribución Poisson de parámetro λ. Si $P(X = 0) = 0.4$, hallar $P(X \leq 3)$.

7. Sea X una variable aleatoria con distribución Poisson de parámetro λ. Si:
$$P(X = 1) = \frac{1}{5}P(X = 2)$$
Calcular $P(X = 0)$ y $P(X \geq 4)$.

8. En un departamento de control de calidad, se inspeccionan las unidades terminadas que provienen de una linea de ensamble. Se piensa que la proporción de unidades defectuosas es de 0.01. ¿Cuál es la probabilidad de que la quinta unidad inspeccionada sea la segunda que se encuentre defectuosa?

9. Las ruletas que se usan en los casinos, tienen 38 lugares de los cuales 18 son negros, 18 son rojos y 2 son verdes. Sea X la variable aleatoria que denota el número de veces que es necesario hacer girar la ruleta, para obtener por primera vez un número rojo. Hallar la función de densidad de X.

10. Un examen de opción múltiple contiene 30 preguntas, cada una con cinco respuestas posibles. Supóngase que un estudiante sólo adivina las respuestas.

    a) ¿Cuál es la probabilidad de que el estudiante conteste de manera correcta más de 20 preguntas?

    b) ¿Cuál es la probabilidad de que el estudiante conteste de manera correcta menos de 3 preguntas?

c) ¿Cuál es la probabilidad de que el estudiante conteste todas las preguntas de manera incorrecta?

11. Mediante estudios recientes se determinó que la probabilidad de morir a causa de cierta vacuna contra la gripe es de 0.00002. Si se administra la vacuna a 100000 personas y se supone que éstas constituyen un conjunto independiente de ensayos, ¿cuál es la probabilidad de que mueran a lo más dos personas a causa de la vacuna?

12. Un profesor desea hacer el examen final de su materia en forma de test. El examen debe contener 15 preguntas y, por reglamento interno, un estudiante lo aprueba si el número de respuestas correctas es de por lo menos 10. Para minimizar, en lo posible, el "riesgo" de que un estudiante apruebe el examen adivinando las respuestas, el profesor desea colocar k posibles respuestas a cada pregunta siendo sólo una de ellas correcta. ¿qué valor debe tener k para que un estudiante, que responda al azar todas las preguntas, tenga una probabilidad igual a 0.001 de aprobar el examen?

13. Para un experimento médico se requieren 15 personas que sean daltónicas, si se sabe que sólo el 0.1 % de la población tiene ésta característica, ¿cuál es el número esperado de personas que deben ser entrevistadas para encontrar las 15 requeridas?

14. En un lago se encuentran N peces de los cuales R ($\leq N$) están marcados. Supongamos que se pescan n ($\leq R$) peces uno a uno, sin sustitución y sea $M_i :=$ "el $i - \acute{e}simo$ pez capturado está marcado", $i = 1, 2, ..., n$. ¿A qué es igual $p(M_i)$?, ¿son M_1 y M_3 independientes?

15. En una central teléfonica se reciben llamadas, según las leyes de un proceso de Poisson, con un promedio de diez llamadas por hora, ¿cuál es la probabilidad de que ninguna llamada sea recibida entre las 8 am y las 12m?

16. En cierta entidad estatal, la probabilidad de que una llamada sea atendida en menos de 30 segundos es de 0.25. Supóngase que las llamadas son independientes.

    a) Si una persona llama 10 veces, ¿cuál es la probabilidad de que 9 de las llamadas sean contestadas en menos de 30 segundos?

    b) Si una persona llama 20 veces, ¿cuál es el número promedio de llamadas que serán contestadas en menos de 30 segundos?

c) ¿Cuál es el número promedio de llamadas que hay que hacer, hasta tener una respuesta en menos de 30 segundos?

d) ¿Cuál es la probabilidad de tener que llamar seis veces para que dos de las llamadas sean contestadas en menos de 30 segundos?

17. Supóngase que el 5 % de los artículos producidos en una fábrica son defectuosos. Se escogen 15 artículos al azar y se inspeccionan, ¿cuál es la probabilidad de que se encuentren a lo sumo tres artículos defectuosos?

18. Supóngase que X es una variable aleatoria con distribución geométrica de parámetro p. Hallar la función de densidad de X^2 y de $X+3$.

19. En una biblioteca hay tres ejemplares del libro "Hágase rico en menos de 24 horas" los cuales son prestados a los usuarios por un día. El número de solicitudes diarias por un ejemplar, es una variable aleatoria con distribución Poisson de media 5. Calcular:

    a) La proporción de días en los que la demanda por un ejemplar del libro es cero.

    b) La proporción de días en los que la demanda por un ejemplar del libro supera la oferta.

20. Cierto sistema de un vehículo espacial debe funcionar correctamente para que la nave pueda reingresar en la atmósfera terrestre. Un componente del sistema opera sin problemas sólo el 85 % de las veces. Al fin de aumentar la confiabilidad del sistema, cuatro de estos componentes se instalan de modo tal que el sistema opere sin problemas, si por lo menos uno de los componentes está funcionando sin problemas.

    i) ¿Qué probabilidad hay de que falle el sistema? Suponga que los componentes operan de forma independiente.

    ii) Si el sistema falla, ¿qué se infiere acerca de la tasa de éxito de 85 % que se dice tiene un solo componente?

21. Una compañía de seguros descubrió que sólo alrededor del 0.1 % de la población tiene cierto tipo de accidente cada año. Si los 10000 asegurados fueran seleccionados aleatoriamente, ¿cuál es la probabilidad de que no más de 5 de estos clientes tenga un accidente de este tipo el próximo año?

22. En un bosque hay 20 osos de anteojos de los cuales 5 son capturados, marcados y dejados nuevamente en libertad. Unas semanas más tarde, 4 de los 20 osos son capturados. Calcular la probabilidad de que a lo más dos de los osos capturados estén marcados.

23. Una compañía analiza los embarques de sus proveedores con la finalidad de detectar productos que no cumplen con las especificaciones. Se sabe que el 3 % de los productos no satisfacen las específicaciones de la compañía. ¿Cuál debe ser el tamaño de la muestra para que la probabilidad de seleccionar al menos un artículo que no cumple con las especificaciones, sea al menos 0.90? Supóngase que en este caso resulta adecuado el uso de la aproximación de la distribución hipergeométrica por la binomial.

24. Un libro de 300 páginas tiene 253 errores tipográficos. Supóngase que cada página tiene 5000 caracteres. ¿Cuál es la probabilidad de que en la primera página no haya errores tipográficos?, ¿cuál es la probabilidad de que en la primera página haya por lo menos un error tipográfico?

25. En una fábrica se producen cartuchos de tinta para estilógrafo. Uno de cada 30 cartuchos producidos por la fábrica resulta ser defectuoso. Los cartuchos son empacados en cajitas de seis cartuchos cada una. Calcular el número esperado de cajitas que contienen, respectivamente, ningún cartucho defectuoso, un cartucho defectuoso, dos o más cartuchos defectuosos, en un envío de 1000 paquetes.

26. La policía sospecha que en un camión cargado con 40 bultos de arroz se han camuflado paquetes de cocaína. Para confirmar su sospecha, la policía escoge al azar 5 bultos para inspeccionarlos. Si en efecto, de los 40 bultos de arroz, que contiene el camión, 10 tienen camuflada cocaína, ¿cuál es la probabilidad de que por lo menos uno de los bultos de la muestra contenga cocaína?

27. En un concurso se propone el siguiente juego: en una urna se encuentran 5 boletas premiadas y 9 boletas no premiadas. El concursante debe extraer simultáneamente y al azar 3 boletas; si las tres boletas están premiadas, el concursante tiene dos opciones: puede escoger uno de los premios y retirarse del concurso o bien puede repetir el ensayo tres veces más y si en las tres repeticiones ocurre que las 3 boletas extraídas están premiadas entonces el concursante recibe un automóvil último modelo, además de los premios indicados en las boletas. Si un

concursante, opta por esta última opción, ¿cuál es la probabilidad de que gane?

28. El promedio de homicidios mensuales en un país es de 1 por cada 100000 habitantes.

    a) Determinar la probabilidad de que en una ciudad de dicho país, de 400000 habitantes, haya 8 o más homicidios en un mes dado.

    b) Calcular la probabilidad de que haya, por lo menos, dos meses durante el año en los que, en dicha ciudad, ocurran 8 o más homicidios.

    c) Contando el presente mes como el mes número 1, ¿cuál es la probabilidad de que el primer mes en tener 5 o más homicidios sea el cuarto?

29. El número de veces que una persona contrae un resfriado en un año constituye una variable aleatoria con distribución Poisson de parámetro $\lambda = 3$. Supóngase que acaba de salir al mercado un nuevo medicamento (basado en grandes cantidades de vitamina C) que reduce el parámetro de Poisson, en el 85 % de la población, a $\lambda = 2$, y en el 15 % restante no tiene efectos apreciables sobre resfriados. Si una persona toma el medicamento durante un año y en este lapso tiene 2 resfriados, ¿qué tan posible es que el medicamento no haya surtido efecto en esa persona?

30. Determinar el número esperado y la varianza del número de veces que es necesario lanzar un dado corriente hasta que el resultado "1" ocurra 4 veces.

31. Un jugador tiene la siguiente estrategia de juego a la ruleta: él apuesta dos unidades monetarias al color rojo, si en el primer giro de la ruleta aparece un número rojo, él toma el dinero ganado y se retira del juego; si en el primer giro de la ruleta aparece un número negro o verde, él hace girar la ruleta dos veces más y apuesta cada vez tres unidades monetarias al color rojo y luego se retira del juego. Sea X la variable aleatoria que denota la fortuna del jugador. Calcular EX (ver problema 9).

32. Para financiarse sus estudios en la universidad, un joven ha decidido vender emparedados de jamón y queso a sus compañeros. El costo de elaboración de cada emparedado es de 500 pesos y el joven los vende

a 1500 pesos. Sin embargo, los emparedados que no logra vender en un día, no los puede vender al día siguiente. Si la demanda diaria de emparedados es una variable aleatoria con distribución binomial de parámetros $n = 20$ y $p = \frac{2}{3}$, ¿cuántos emparedados debe elaborar diariamente para maximizar su ganancia diaria esperada?

33. Se lanza una moneda corriente tantas veces como sea necesario hasta obtener por primera vez "cara". Sea X el número de lanzamientos requeridos.

    a) Calcular $E(X)$.

    b) Hallar (si existe) $E(2^X)$.

34. Se carga una moneda de tal manera que la probabilidad de obtener "cara" es igual a 0.4. Usar la desigualdad de Chebyschev para determinar cuántas veces debe lanzarse la moneda, de tal manera que, con una probabilidad de por lo menos 0.9, el cociente entre el número de veces que se obtiene cara y el número total de lanzamientos, esté entre 0.3 y 0.5.

35. Sea X una variable aleatoria con distribución Poisson de parámetro λ. Encontrar el valor de λ para el cual $P(X = k)$ sea máxima.

36. Sea X una variable aleatoria con distribución Poisson de parámetro λ. Demostrar que:

$$P(X \leq n) = \frac{1}{n!}\int_{\lambda}^{\infty} e^{-x}x^n dx \qquad n = 0, 1, \cdots$$

37. Sea X una variable aleatoria con distribución binomial de parámetros n y p. Demostrar que:

$$E\left(\frac{1}{X+1}\right) = \frac{1-(1-p)^{n+1}}{(n+1)p}.$$

38. Sea X una variable aleatoria con distribución binomial de parámetros n y p. Encontrar el valor de p que maximiza $P(X = k)$ para $k = 0, \cdots, n$.

39. Sea X una variable aleatoria con distribución hipergeométrica de parámetros n, R y N.

a) Demostrar que:

$$P(X = j+1) = \frac{(R-j)(n-j)}{(j+1)((N-R)-n+j+1)} P(X = j).$$

b) Verificar que:

$$P(X = j+1) > P(X = j),$$

si y sólo si,

$$j < \frac{(n+1)(R+1)}{N+2} - 1.$$

40. Sea X una variable aleatoria con distribución Poisson de parámetro λ. Demostrar que:

$$E(X^n) = \lambda E\left((X+1)^{n-1}\right).$$

41. Sea X una variable aleatoria con distribución geométrica de parámetro p. Demostrar que:

$$P(X = n+k \mid X > n) = P(X = k).$$

42. Sea X una variable aleatoria con distribución geométrica de parámetro p. Determinar $E\left(\frac{1}{X}\right)$.

43. Sea X una variable aleatoria con distribución Poisson de parámetro λ. Demostrar:

    a) $EX^2 = \lambda E(X+1)$.

    b) Si $\lambda = 1$ entonces $E(|X-1|) = \frac{2}{e}$.

44. Determinar la función característica de una variable aleatoria con distribución binomial de parámetros n y p.

45. Determinar la función característica de una variable aleatoria con distribución Poisson de parámetro λ.

46. ¿Cuántos hijos debe tener una pareja para que, con probabilidad 0.95, ella tenga por lo menos un hijo y una hija?

Capítulo 4

Algunas distribuciones continuas

En este capítulo se estudiarán algunas de las distribuciones de tipo absolutamente continuo de uso más frecuente.

4.1. Distribución uniforme

Supóngase que un bus escolar llega siempre a cierto paradero entre las 6 a.m y las $6:10$ a.m y que la probabilidad de que el bus llegue en cualquier subintervalo de tiempo, del intervalo $[0,10]$, es sólo proporcional a la longitud del subintervalo. Es decir, es igualmente probable que el bus llegue entre las $6:00$ a.m y las $6:02$ a.m, a que llegue entre las $6:07$ a.m y las $6:09$ a.m. Sea X el tiempo, medido en minutos, que un escolar debe esperar en el paradero del bus, si llega exactamente a las $6:00$ a.m. Si se miden cuidadosamente, durante muchas mañanas, la hora en la que llega el bus, se puede construir, con los datos obtenidos, un histograma de frecuencias relativas. De la descripción anterior, se tiene que las frecuencias relativas con que se observa a X entre $6:00$ y $6:02$, y entre $6:07$ y $6:09$, son practicamente las mismas. La variable X es un ejemplo de una variable aleatoria con distribución uniforme. Más precisamente se tiene la definición siguiente:

Definición 4.1 (distribución uniforme) *Se dice que una variable aleatoria X está distribuida uniformemente sobre el intervalo $[a,b]$, con $a<b$*

números reales, si su función de densidad está dada por:

$$f(x) = \begin{cases} \frac{1}{b-a} & si \quad a \leq x \leq b \\ 0 & en \quad otro\ caso \end{cases}$$

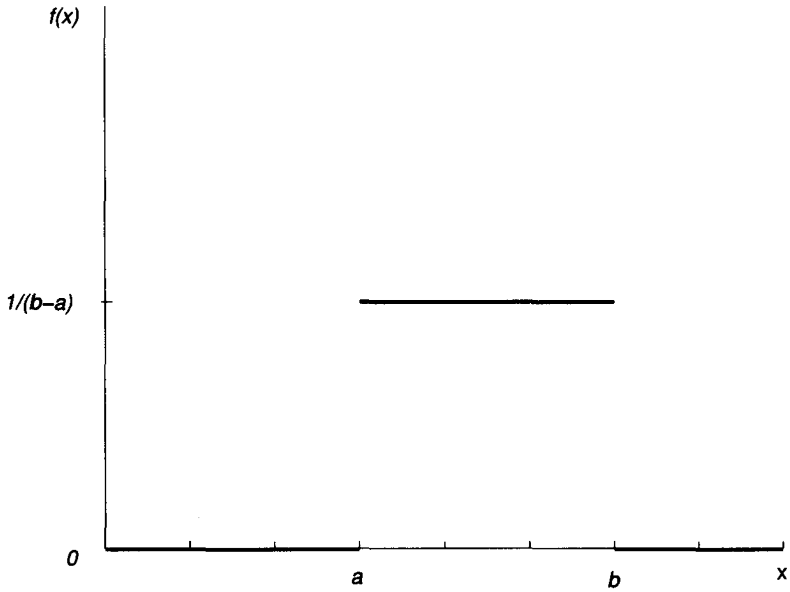


Figura 4.1: Función de densidad de una distribución uniforme

Notación 4.2 *La expresión* $X \stackrel{d}{=} \mathcal{U}[a,b]$ *significa que la variable aleatoria* X *tiene distribución uniforme sobre el intervalo* $[a,b]$.

Es fácil verificar que si $X \stackrel{d}{=} \mathcal{U}[a,b]$, entonces, la función de distribución de X está dada por:

$$F(x) = \begin{cases} 0 & \text{si} \quad x < a \\ \frac{x-a}{b-a} & \text{si} \quad a \leq x \leq b \\ 1 & \text{si} \quad x > b. \end{cases}$$

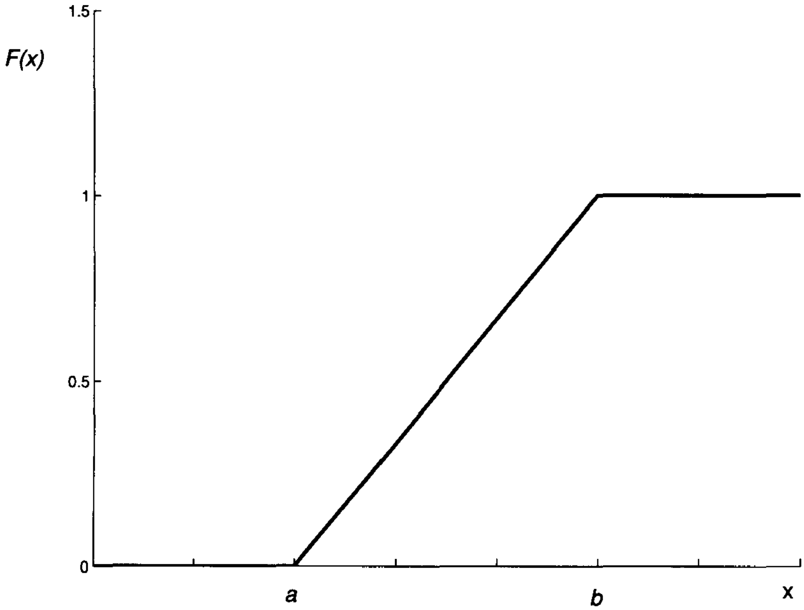


Figura 4.2: Función de distribución de una variable aleatoria con distribución uniforme

Ejemplo 4.3 *Sea* $X \overset{d}{=} \mathcal{U}[-3,2]$. *Calcular:*

1. $P(X \geq 0)$
2. $P(-5 \leq X \leq \frac{1}{2})$

Solución: En este caso la función de densidad de la variable aleatoria X está dada por:

$$f(x) = \begin{cases} \frac{1}{5} & \text{si} \quad -3 \leq x \leq 2 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

por lo tanto,

$$P(X \geq 0) = \int_0^2 \frac{1}{5} dx = \frac{2}{5}$$

y

$$P\left(-5 \leq X \leq \frac{1}{2}\right) = \int_{-3}^{\frac{1}{2}} \frac{1}{5} dx = \frac{1}{5}\left(\frac{1}{2} + 3\right) = \frac{7}{10}. \quad \blacktriangle$$

Ejemplo 4.4 *Sean $a, b \in \mathbb{R}$ fijos, con $a < b$. Se escoge un número X, al azar, en el intervalo $[a, b]$. Esto significa que cualquier subintervalo de $[a, b]$, de longitud τ tiene la misma probabilidad de contener a X. Por lo tanto, para cualquier $a \leq x \leq y \leq b$, se tiene que $P(x \leq X \leq y)$ depende sólo de $(y - x)$. Si f es la función de densidad de la variable aleatoria X, entonces,*

$$kdx = P(x < X \leq x + dx) = f(x)dx$$

es decir, $f(x) = k$ siendo k una constante apropiada. Puesto que

$$1 = \int_{-\infty}^{\infty} f(x)dx = \int_a^b kdx$$

se deduce que $k = \frac{1}{b-a}$. Esto es, $X \stackrel{d}{=} \mathcal{U}[a, b]$. ▲

Ejemplo 4.5 *Se escoge un número al azar en el intervalo $[1, 3]$. ¿Cuál es la probabilidad de que el primer dígito al lado derecho del punto decimal sea el 5? ¿Cuál es la probabilidad de que el segundo dígito a la derecha del punto decimal sea el 2?*

Solución: Sea $X :=$ "número escogido al azar en el intervalo $[1, 3]$". La función de densidad de la variable aleatoria X de acuerdo al ejemplo anterior, es:

$$f(x) = \begin{cases} \frac{1}{2} & \text{si} \quad 1 \leq x \leq 3 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

Por lo tanto,

$$\begin{aligned} &P(\text{"primer dígito al lado derecho del punto decimal de } X \text{ es 5"}) \\ &= P(1.5 \leq X < 1.6) + P(2.5 \leq X < 2.6) \\ &= \int_{1.5}^{1.6} \frac{1}{2} dx + \int_{2.5}^{2.6} \frac{1}{2} dx = 0.1 \end{aligned}$$

y

$$\begin{aligned} &P(\text{"segundo dígito al lado derecho del punto decimal de } X \text{ es 2"}) \\ &= P\left(X \in \bigcup_{k=0}^{9} \{[1.k2, 1.k3) \cup [2.k2, 2.k3)\} \right) \\ &= 20 \times \frac{1}{2} \times 0.01 = 0.1. \quad \blacktriangle \end{aligned}$$

Teorema 4.6 *Si X es una variable aleatoria con distribución uniforme sobre el intervalo $[a, b]$ entonces:*

1. $EX = \frac{a+b}{2}$.

2. $Var(X) = \frac{(b-a)^2}{12}$.

Demostración.

1.

$$\begin{aligned} EX &= \int_a^b x \frac{1}{b-a} dx \\ &= \frac{a+b}{2}. \end{aligned}$$

2.

$$\begin{aligned} EX^2 &= \int_a^b x^2 \frac{1}{b-a} dx \\ &= \frac{b^2 + ab + a^2}{3} \end{aligned}$$

por lo tanto,

$$\begin{aligned} Var(X) &= \frac{b^2 - 2ab + a^2}{12} \\ &= \frac{(b-a)^2}{12}. \end{aligned}$$

■

Ejemplo 4.7 *Supóngase que $X \stackrel{d}{=} \mathcal{U}[a, b]$ y que $EX = 2$ y $VarX = \frac{3}{4}$. Calcular $P(X \leq 1)$.*

Solución: Se tiene que $\frac{a+b}{2} = 2$ y $\frac{(b-a)^2}{12} = \frac{3}{4}$. Por lo tanto $a = \frac{1}{2}$ y $b = \frac{7}{2}$. Luego:

$$P(X \leq 1) = \int_{\frac{1}{2}}^{1} \frac{1}{3} dx = \frac{1}{6}. \quad \blacktriangle$$

Nota 4.8 *Supóngase que X es una variable aleatoria con función de distribución continua y creciente F. Sea Y una variable aleatoria con distribución uniforme en el intervalo $(0,1)$ y sea Z la variable aleatoria definida como $Z := F^{-1}(Y)$. La función de distribución de la variable aleatoria Z está dada por:*

$$\begin{aligned} F_Z(z) &= P(Z \leq z) \\ &= P\left(F^{-1}(Y) \leq z\right) \\ &= P(Y \leq F(z)) \\ &= F(z). \end{aligned}$$

Esto es, las variables aleatorias X y Z tienen la misma distribución de probabilidades. Por lo tanto, se pueden simular variables aleatorias con distribución continua y creciente, conocida. En el capítulo 8 se explicará este procedimiento.

4.2. Distribución normal.

La distribución normal es una de las más importantes y de mayor uso tanto en la teoría de la probabilidad, como en la teoría estadística.

Algunos autores la llaman distribución gaussiana, en honor a Gauss, a quien se considera el "padre" de ésta distribución.

La importancia de la distribución normal, radica en el famoso Teorema Central del Límite, el cual se discutirá en el capítulo 7.

Definición 4.9 *Se dice que la variable aleatoria X tiene distribución normal de parámetros μ y σ, donde μ es un número real y σ es un número real positivo, si su función de densidad está dada por:*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]; \; x \in \mathbb{R}.$$

Se deja como ejercicio, al lector, verificar que, efectivamente, f es una función de densidad. Esto es que f es no negativa y que

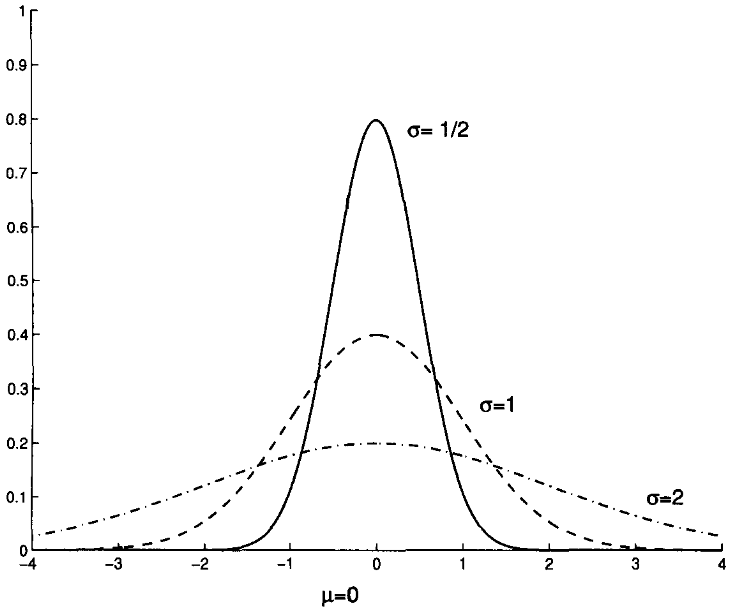
$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

El parámetro μ es un parámetro de localización y σ es un parámetro de escala. Conceptos que se precisan a continuación.

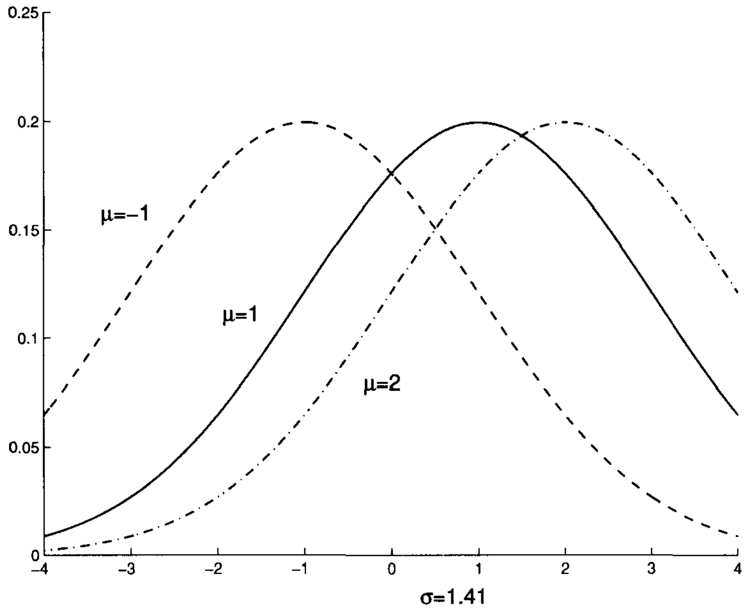
Definición 4.10 (parámetros de localización y de escala) *Sea Y una variable aleatoria. Se dice que θ_1 es un parámetro de localización, si para todo $c \in \mathbb{R}$, se tiene que la variable aleatoria $Z := Y + c$ tiene parámetro $\theta_1 + c$. Esto es, si $f_Y(\cdot; \theta_1, \theta_2)$ es la función de densidad de Y entonces la función de densidad de Z es $f_Z(\cdot; \theta_1 + c, \cdot)$. Se dice que θ_2 es un parámetro de escala, si $\theta_2 > 0$ y para todo $c \in \mathbb{R}$, la variable aleatoria $W := cY$ tiene parámetro $|c|\,\theta_2$. Esto es, si $f_Y(\cdot; \theta_1, \theta_2)$ es la función de densidad de Y, entonces la función de densidad de W es $f_W(\cdot; \cdot, |c|\,\theta_2)$.*

Notación 4.11 *Se escribe $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ para indicar que X es una variable aleatoria con distribución normal de parámetros μ y σ.*

En la gráfica siguiente se muestra la función de densidad de una variable aleatoria X con distribución normal con $\mu = 0$ y valores diferentes de σ.



La siguiente gráfica muestra la función de densidad de una variable aleatoria $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ para $\sigma = 1.41$ y valores diferentes de μ.



La función de distribución de una variable aleatoria $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ está dada por:

$$F(x) = \frac{1}{\sqrt{2\pi}\sigma} \int_{-\infty}^{x} \exp\left[-\frac{1}{2}\left(\frac{u-\mu}{\sigma}\right)^2\right] du.$$

La gráfica de F, con $\mu = 0$ y $\sigma = 1$, está dada en la figura 4.3.

Definición 4.12 (distribución normal estándar) *Si* $X \stackrel{d}{=} \mathcal{N}(0,1)$, *entonces se dice que* X *tiene distribución normal estándar. La función de densidad y la función de esta variable aleatoria se denotan por* $\phi(\cdot)$ *y* $\Phi(\cdot)$ *respectivamente.*

Nota 4.13 *La función de densidad de una variable aleatoria normal estándar es simétrica con respecto al eje* y. *Por lo tanto, para todo* $z < 0$ *se satisface que:*

$$\Phi(z) = 1 - \Phi(-z).$$

Nota 4.14 *Sea* $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ *y sea* $Y := aX + b$ *donde* a *y* b *son constantes reales con* $a \neq 0$. *Por lo visto, en el capítulo 2, se sabe que una función de*

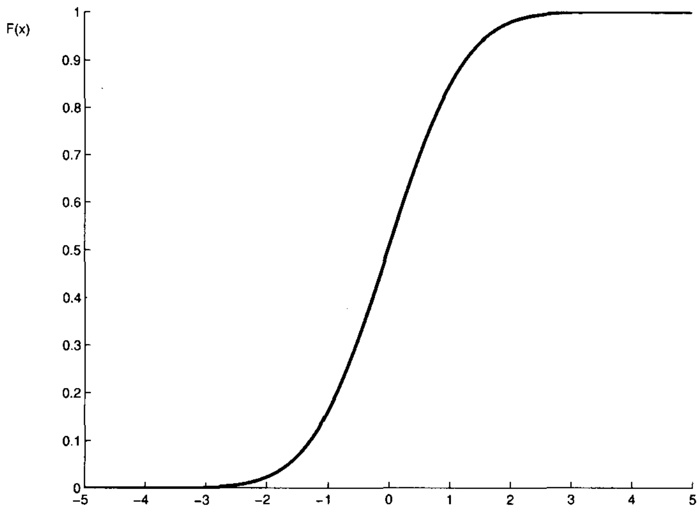


Figura 4.3: Función de distribución de la normal estándar

densidad de la variable aleatoria Y está dada por:

$$f_Y(x) = \frac{1}{|a|} f_X(\frac{x-b}{a})$$

$$= \frac{1}{|a|} \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\left(\frac{x-(a\mu+\sigma)}{a\sigma}\right)^2\right].$$

Esto es, Y tiene distribución normal de parámetro de localización $a\mu + b$ y parámetro de escala $|a|\,\sigma$. En particular, si $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$ entonces $Y = \frac{X-\mu}{\sigma}$ tiene distribución normal estándar. Esto es, para conocer los valores de la función de distribución de una variable aleatoria con distribución normal arbitraria, basta conocer los de una variable aleatoria con distribución normal estándar. En el apéndice D.3 se encuentra una tabla con los valores de la función de distribución normal estándar.

Ejemplo 4.15 *Sea $X \stackrel{d}{=} \mathcal{N}(1, 4)$. Calcular:*

1. $P(0 \leq X < 1)$.
2. $P(X^2 > 4)$.

Solución: Se tiene que:

$$\begin{aligned}
P(0 \leq X < 1) &= P\left(-\frac{1}{2} \leq \frac{X-1}{2} < 0\right) \\
&= \Phi(0) - \Phi\left(-\frac{1}{2}\right) \\
&= 0.5 - 0.30854 \\
&= 0.19146
\end{aligned}$$

y

$$\begin{aligned}
P\left(X^2 > 4\right) &= 1 - P(|X| \leq 2) \\
&= 1 - P(-2 \leq X \leq 2) \\
&= 1 - P\left(-\frac{3}{2} \leq \frac{X-1}{2} \leq \frac{1}{2}\right) \\
&= 1 - \left[\Phi(\frac{1}{2}) - \Phi(-\frac{3}{2})\right] \\
&= 1 - [0.69146 - 0.06681] \\
&= 0.37535. \quad \blacktriangle
\end{aligned}$$

Ejemplo 4.16 *Supóngase que un instructor asume que las calificaciones finales de los estudiantes son los valores de una variable aleatoria X con distribución normal de parámetros μ y σ. El instructor decide asignar la calificación A a aquellos estudiantes cuyo puntaje exceda a $\mu + \sigma$, la calificación B a aquellos estudiantes cuyos puntajes estén entre μ y $\mu + \sigma$, la calificación C a aquellos estudiantes cuyos puntajes estén entre $\mu - \sigma$ y μ, la calificación D a aquellos estudiantes cuyos puntajes estén entre $\mu - 2\sigma$ y $\mu - \sigma$, y la calificación F a aquellos estudiantes cuyos puntajes sean inferiores a $\mu - 2\sigma$. Encontrar el porcentaje de estudiantes que obtienen como calificación A, B, C, D o F.*

Solución: Se tiene que:

$$\begin{aligned}
P(X \geq \mu + \sigma) &= P\left(\frac{X-\mu}{\sigma} \geq 1\right) \\
&= 1 - \Phi(1) = 0.1587. \\
P(\mu \leq X < \mu + \sigma) &= P\left(0 \leq \frac{X-\mu}{\sigma} < 1\right) \\
&= \Phi(1) - \Phi(0) = 0.34134.
\end{aligned}$$

$$\begin{aligned} P(\mu - \sigma \leq X < \mu) &= P\left(-1 \leq \frac{X-\mu}{\sigma} < 0\right) \\ &= \Phi(0) - \Phi(-1) = 0.34134. \\ P(\mu - 2\sigma \leq X < \mu - \sigma) &= P\left(-2 \leq \frac{X-\mu}{\sigma} < -1\right) \\ &= \Phi(-1) - \Phi(-2) = 0.13591. \\ P(X < \mu - 2\sigma) &= P\left(\frac{X-\mu}{\sigma} < -2\right) \\ &= \Phi(-2) = 0.02275. \end{aligned}$$

Por lo tanto el 15.87 % de los estudiantes obtienen como calificación A; el 34.13 % B; el 34.13 % C; el 13.59 % D y el restante 2.28 % F. ▲

Ejemplo 4.17 *Sea* $X \stackrel{d}{=} \mathcal{N}(12, 4)$. *Encontrar el valor de c para el cual* $P(X > c) = 0.10$.

Solución:

$$\begin{aligned} P(X > c) &= 1 - P(X \leq c) \\ &= 1 - P(\frac{X-12}{2} \leq \frac{c-12}{2}) \\ &= 1 - \Phi\left(\frac{c-12}{2}\right). \end{aligned}$$

esto es,

$$\Phi\left(\frac{c-12}{2}\right) = 0.9,$$

luego de los valores dados en la tabla se tiene que:

$$\frac{c-12}{2} = 1.285,$$

es decir, $c = 14.57$. ▲

Nota 4.18 *Supóngase que* $X \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$. *Entonces:*

$$\begin{aligned} P(\mu - 3\sigma \leq X \leq \mu + 3\sigma) &= P\left(\frac{\mu - 3\sigma - \mu}{\sigma} \leq \frac{X-\mu}{\sigma} \leq \frac{\mu + 3\sigma - \mu}{\sigma}\right) \\ &= \Phi(3) - \Phi(-3) \\ &= 0.99865 - 0.00135 \\ &= 0.9973 \end{aligned}$$

o equivalentemente, $P(|X-\mu|>3\sigma)=0.0027$. *Es decir, los valores alejados de* μ *tienen probabilidad pequeña de ocurrir. Por esta razón, se dice que la distribución normal tiene la característica de tener colas ligeras.*

A continuación se va a hallar el valor esperado, la varianza y la función generadora de momentos de una variable aleatoria con distribución normal.

Teorema 4.19 *Sea* $X \stackrel{d}{=} \mathcal{N}(\mu,\sigma^2)$. *Entonces:*

1. $EX=\mu$.
2. $Var(X)=\sigma^2$.
3. $m_X(t)=\exp\left[\mu t+\frac{\sigma^2t^2}{2}\right]$.

Demostración. Se va a calcular la función generadora de momentos. A partir de ella se hallan, fácilmente, el valor esperado y la varianza.

$$\begin{aligned}
m_X(t) &= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp\left[tx-\frac{(x-\mu)^2}{2\sigma^2}\right]dx\\
&= \frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}\exp\left[-\frac{(x-\mu-\sigma^2t)^2}{2\sigma^2}+\mu t+\frac{\sigma^2t^2}{2}\right]dx\\
&= \frac{1}{\sqrt{2\pi}\sigma}\exp\left[\mu t+\frac{\sigma^2t^2}{2}\right]\int_{-\infty}^{\infty}\exp\left[-\frac{\left(x-\{\mu+\sigma^2t\}\right)^2}{2\sigma^2}\right]dx\\
&= \exp\left[\mu t+\frac{\sigma^2t^2}{2}\right]
\end{aligned}$$

por consiguiente:

$$EX=\left[\mu+\sigma^2t\right]\exp\left[\mu t+\frac{\sigma^2t^2}{2}\right]\Bigg|_{t=0}=\mu$$

y

$$EX^2=\left(\sigma^2+\left[\mu+\sigma^2t\right]^2\right)\exp\left[\mu t+\frac{\sigma^2t^2}{2}\right]\Bigg|_{t=0}=\sigma^2+\mu^2$$

luego,

$$Var(X)=\sigma^2.$$

■

Nota 4.20 *Se puede verificar que la función característica de una variable aleatoria X, con distribución normal de parámetros μ y σ, está dada por:*

$$\Phi_X(t) = \exp\left[i\mu t - \frac{\sigma^2 t^2}{2}\right].$$

Nota 4.21 *La distribución normal es otra forma límite de la distribución binomial, siempre y cuando se satisfagan las siguientes condiciones sobre los parámetros n y p de la distribución binomial: $n \longrightarrow \infty$ y, ni p ni $q = 1-p$ son muy pequeños. En efecto,*

Supóngase que $X \stackrel{d}{=} \mathcal{B}(n,p)$. Entonces:

$$P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, \cdots, n$$

Cuando $n \longrightarrow \infty$, se tiene que $x \longrightarrow \infty$ y además

$$n! \approx \sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}} \qquad \text{(Fórmula de Stirling)}$$

por lo tanto,

$$\begin{aligned}
\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} &= \lim_{n\to\infty} \frac{\sqrt{2\pi} e^{-n} n^{n+\frac{1}{2}} p^x (1-p)^{n-x}}{\sqrt{2\pi} e^{-x} x^{x+\frac{1}{2}} \sqrt{2\pi} e^{-(n-x)} (n-x)^{(n-x)+\frac{1}{2}}} \\
&= \lim_{n\to\infty} \frac{n^{n+\frac{1}{2}} p^x (1-p)^{n-x} \sqrt{np(1-p)}}{\sqrt{2\pi} x^{x+\frac{1}{2}} (n-x)^{(n-x)+\frac{1}{2}} \sqrt{np(1-p)}} \\
&= \lim_{n\to\infty} \frac{(np)^{x+\frac{1}{2}} \left(n(1-p)\right)^{n-x+\frac{1}{2}}}{\sqrt{2\pi} x^{x+\frac{1}{2}} (n-x)^{(n-x)+\frac{1}{2}} \sqrt{np(1-p)}} \\
&= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}\sqrt{np(1-p)}} \left(\frac{np}{x}\right)^{x+\frac{1}{2}} \left(\frac{n(1-p)}{n-x}\right)^{(n-x)+\frac{1}{2}}.
\end{aligned}$$

Sea $\frac{1}{N} := \left(\frac{np}{x}\right)^{x+\frac{1}{2}} \left(\frac{n(1-p)}{n-x}\right)^{(n-x)+\frac{1}{2}}$. Es claro que:

$$\ln N = \left(x + \frac{1}{2}\right) \ln\left(\frac{x}{np}\right) + \left(n - x + \frac{1}{2}\right) \ln\left(\frac{n-x}{n(1-p)}\right) \tag{4.1}$$

Si se toma $Z := \frac{X-np}{\sqrt{np(1-p)}}$, se tiene que Z toma los valores $z = \frac{x-np}{\sqrt{np(1-p)}}$. Al tomar el límite cuando $n \longrightarrow \infty$, se tiene que Z toma todos los valores *de* $-\infty$ *a* ∞. *Despejando* x *en la ecuación anterior se obtiene que* $x = z\sqrt{np(1-p)} + np$. *Reemplazando en (4.1) se llega a:*

$$\begin{aligned}
\ln N &= \left(z\sqrt{np(1-p)} + np + \frac{1}{2}\right) \ln\left(\frac{z\sqrt{np(1-p)} + np}{np}\right) \\
&+ \left(n - \left(z\sqrt{np(1-p)} + np\right) + \frac{1}{2}\right) \ln\left(\frac{n - \left(z\sqrt{np(1-p)} + np\right)}{n(1-p)}\right) \\
&= \left(np + z\sqrt{np(1-p)} + \frac{1}{2}\right) \ln\left(1 + z\sqrt{\frac{1-p}{np}}\right) \\
&+ \left(n(1-p) - z\sqrt{np(1-p)} + \frac{1}{2}\right) \ln\left(1 - z\sqrt{\frac{p}{n(1-p)}}\right).
\end{aligned}$$

Desarrollando en serie la función $h(x) = \ln(1+x)$ *se obtiene:*

$$\begin{aligned}
\ln N &= \left(z\sqrt{np(1-p)} + np + \frac{1}{2}\right) \left[z\sqrt{\frac{1-p}{np}} - \frac{1}{2}z^2\left(\frac{1-p}{np}\right) + \cdots\right] \\
&+ \left(n(1-p) - z\sqrt{np(1-p)} + \frac{1}{2}\right) \\
&\qquad \cdot \left[-z\sqrt{\frac{p}{n(1-p)}} - \frac{1}{2}z^2\left(\frac{p}{n(1-p)}\right) - \cdots\right] \\
&= \left[z^2(1-p) - \frac{1}{2}z^3\sqrt{\frac{(1-p)^3}{np}} + z\sqrt{np(1-p)} - \frac{1}{2}z^2(1-p)\right. \\
&\qquad \left. + \frac{1}{2}z\sqrt{\frac{1-p}{np}} - \frac{1}{4}z^2\left(\frac{1-p}{np}\right) + \cdots\right] \\
&+ \left[-z\sqrt{np(1-p)} - \frac{1}{2}z^2p + z^2p + \frac{1}{2}z^3\sqrt{\frac{p^3}{n(1-p)}} - \frac{1}{2}z\sqrt{\frac{p}{n(1-p)}}\right. \\
&\qquad \left. - \frac{1}{4}z^2\left(\frac{p}{n(1-p)}\right) - \cdots\right].
\end{aligned}$$

esto es,

$$\ln N = -\frac{1}{2}z^2 + z^2 + \frac{z}{2\sqrt{n}}\left(\sqrt{\frac{1-p}{p}} + \sqrt{\frac{p}{1-p}}\right) + o(n^{-\frac{1}{2}}).$$

por lo tanto,

$$\lim_{n\to\infty} \log N = \frac{1}{2}z^2$$

y en consecuencia:

$$\lim_{n\to\infty} N = e^{\frac{1}{2}z^2}.$$

Puesto que,

$$\begin{aligned}
P(X = x) &= P(x < X \leq x + dx) \\
&= P\left(\frac{x - np}{\sqrt{np(1-p)}} < \frac{X - np}{\sqrt{np(1-p)}} \leq \frac{x + dx - np}{\sqrt{np(1-p)}}\right) \\
&= P(z < Z \leq z + dz) \\
&\approx g(z)dz
\end{aligned}$$

donde $g(\cdot)$ es la función de densidad de la variable aleatoria Z, entonces,

$$\begin{aligned}
g(z) &= \lim_{n\to\infty} \frac{1}{\sqrt{2\pi}} \times \frac{1}{N} \\
&= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}z^2\right).
\end{aligned}$$

Esto es, $Z \stackrel{d}{=} \mathcal{N}(0,1)$. En otras palabras, si n es suficientemente grande $\mathcal{B}(n,p) \approx \mathcal{N}(np, np(1-p))$. En la práctica, la aproximación es, por lo general, aceptable cuando $p \in \left(0, \frac{1}{2}\right)$ y $np\,(1-p) > 9$ o $np > 5$, o si $p \in \left(\frac{1}{2}, 1\right)$ y $n\,(1-p) > 5$. En el caso $p = \frac{1}{2}$ se obtiene una aproximación bastante buena aún en el caso en que n sea "pequeña" (ver [Her]).

El resultado que se acaba de deducir se conoce como teorema de Moivre-Laplace y como se verá más adelante es un caso particular del Teorema Central del Límite.

Ejemplo 4.22 *Se lanza un dado corriente* 1000 *veces consecutivas. Calcular la probabilidad de que el número* 6 *aparezca entre* 150 *y* 200 *veces. ¿A qué es igual la probabilidad de que el número* 6 *aparezca exactamente* 150 *veces?*

Solución: Sea $X :=$"Número de veces que se obtiene 6 como resultado".
Es claro que $X \stackrel{d}{=} \mathcal{B}(1000, \frac{1}{6})$. De acuerdo con el resultado anterior, se puede suponer que X tiene una distribución normal de parámetros

$\mu = \frac{500}{3}$ y $\sigma^2 = \frac{1250}{9}$. Por lo tanto,

$$P(150 \le X \le 200) = P\left(\frac{150 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{X - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{200 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}}\right)$$

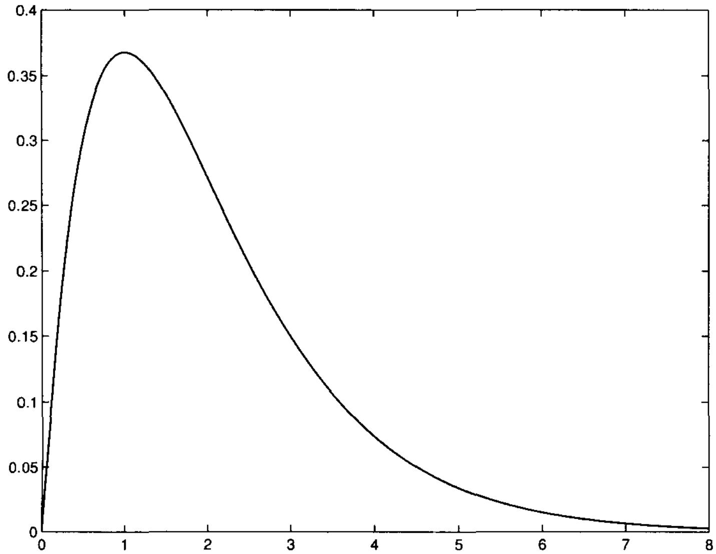
$$\begin{aligned}
&= P(-1.14142 \le Z \le 2.8284) \quad \text{donde } Z \stackrel{d}{=} \mathcal{N}(0,1)\\
&= \Phi(2.8284) - \Phi(-1.14142)\\
&= 0.9976 - 0.07927\\
&= 0.91833.
\end{aligned}$$

Para responder a la segunda pregunta, se observa que, como la distribución binomial es discreta y la distribución normal es continua, una aproximación adecuada se obtiene al considerar lo siguiente:

$$\begin{aligned}
P(X = 150) &= P(149.5 \le X \le 150.5)\\
&= P\left(\frac{149.5 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{X - \frac{500}{3}}{\sqrt{\frac{1250}{9}}} \le \frac{150.5 - \frac{500}{3}}{\sqrt{\frac{1250}{9}}}\right)\\
&= P(-1.4566 \le Z \le -1.3718)\\
&= \Phi(-1.3718) - \Phi(-1.4566)\\
&= 0.08534 - 0.07215\\
&= 0.01319. \quad \blacktriangle
\end{aligned}$$

4.3. Distribución gamma.

Algunas variables aleatorias son no negativas siempre y tienen distribuciones que son sesgadas a la derecha, es decir, la mayor parte del área bajo la gráfica, de la función de densidad, se encuentra cerca del origen y los valores de la función de densidad disminuyen gradualmente cuando x aumenta. Un ejemplo de tales distribuciones es la distribución gamma, cuya función de densidad tiene la siguiente forma:



La distribución gamma se emplea, de manera extensa, en una gran diversidad de áreas, como por ejemplo, para describir los intervalos de tiempo entre dos fallas consecutivas del motor de un avión, o los intervalos de tiempo entre las llegadas de clientes a la cola del punto de pago en un supermercado.

La distribución gamma es la generalización de tres casos particulares que, históricamente, surgieron primero: la distribución exponencial, la distribución Erlang y la distribución Ji- cuadrada.

Definición 4.23 (distribución gamma) *Se dice que una variable aleatoria X tiene distribución gamma de parámetros $r > 0$ y $\lambda > 0$, si su función de densidad está dada por:*

$$f(x) = \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} \exp(-\lambda x) \mathcal{X}_{(0,\infty)}(x)$$

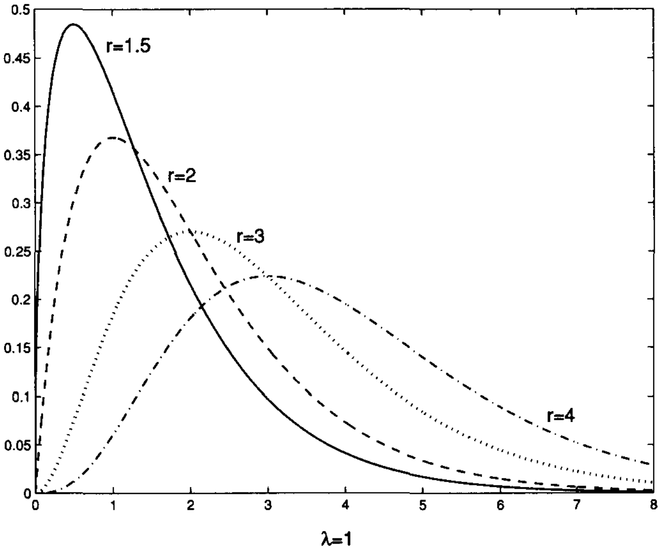
donde $\Gamma(\cdot)$ es la función gamma, esto es,

$$\Gamma(r) := \int_0^\infty t^{r-1} \exp(-t) dt.$$

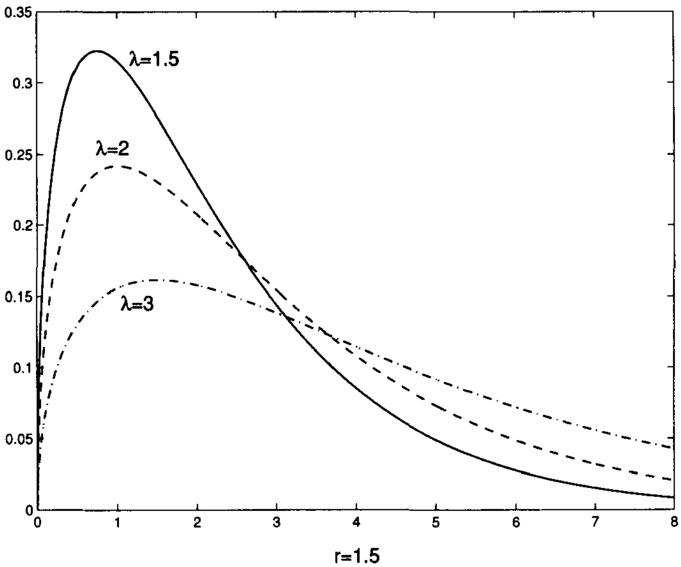
El orden de los parámetros es importante ya que r es el parámetro de forma, en tanto que λ es el parámetro de escala.

La verificación de que f es una función de densidad, se le deja como ejercicio al lector.

La gráfica siguiente muestra la forma de la función de densidad gamma para $\lambda = 1$ y diferentes valores de r.



La gráfica siguiente muestra la forma de la función de densidad gamma para $r = \frac{3}{2}$ y diferentes valores de λ



Notación 4.24 *La expresión* $X \stackrel{d}{=} \Gamma(r, \lambda)$ *significa que* X *tiene distribución gamma de parámetros* r *y* λ.

La función de distribución de una variable aleatoria X con distribución gamma de parámetros r y λ está dada por:

$$\begin{aligned} F(x) &= \int_0^x \frac{\lambda}{\Gamma(r)} (\lambda t)^{r-1} \exp(-\lambda t)\, dt \\ &= \frac{1}{\Gamma(r)} \int_0^{\lambda x} u^{r-1} \exp(-u) du. \end{aligned}$$

Cuando r es un entero positivo, se tiene que $\Gamma(r) = (r-1)!$ y en tal caso se puede verificar que (ejercicio!):

$$F(x) = 1 - \exp(-\lambda x) \sum_{k=0}^{r-1} \frac{(\lambda x)^k}{k!}.$$

Se observa que el lado de la derecha de la ecuación anterior corresponde a $P(Y \geq r)$ donde $Y \stackrel{d}{=} \mathcal{P}(\lambda x)$. ¿Será que existe algún tipo de relación entre la distribución de Poisson y la distribución gamma?. La respuesta a esta pregunta, es sí. En el capítulo 2 se vio que, bajo ciertas condiciones sobre la forma de llegada de los individuos a cierta línea de espera, se tiene que la variable aleatoria X_t, que denota el número de individuos (partículas), que llegan en el intervalo de tiempo $(0, t]$, tiene una distribución Poisson de parámetro λt. Se puede demostrar que si Z es la variable aleatoria que denota el tiempo transcurrido desde el momento en que se inicia la observación, hasta el momento en que llega el $n-$ésimo individuo, entonces $Z \stackrel{d}{=} \Gamma(n, \lambda)$.

En el teorema siguiente se determinan el valor esperado, la varianza y la función generadora de momentos de una variable aleatoria con distribución gamma.

Teorema 4.25 *Si* $X \stackrel{d}{=} \Gamma(r, \lambda)$, *entonces:*

1. $EX = \frac{r}{\lambda}$.
2. $Var(X) = \frac{r}{\lambda^2}$.
3. $m_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r$; *si* $t < \lambda$.

Demostración. *Se va a calcular la función generadora de momentos de* X *y luego, a partir de ella, se hallarán* EX *y* $Var(X)$. *Se tiene que:*

$$\begin{aligned} m_X(t) &= \int_0^\infty \exp(tx) \frac{\lambda}{\Gamma(r)} (\lambda x)^{r-1} \exp(-\lambda x)\, dx \\ &= \left(\frac{\lambda}{\lambda - t}\right)^r \int_0^\infty \frac{(\lambda - t)^r}{\Gamma(r)} x^{r-1} \exp(-(\lambda - t)x) dx. \end{aligned}$$

Si $(\lambda - t) > 0$ *entonces*

$$g(x) := \frac{(\lambda - t)^r}{\Gamma(r)} x^{r-1} \exp(-(\lambda - t)x)\mathcal{X}_{(0,\infty)}(x)$$

es una función de densidad tipo gamma, y por lo tanto

$$\int_0^\infty \frac{(\lambda - t)^r}{\Gamma(r)} x^{r-1} \exp(-(\lambda - t)x)dx = 1$$

luego,

$$m_X(t) = \left(\frac{\lambda}{\lambda - t}\right)^r; \quad si\ t < \lambda.$$

Por otra parte,

$$EX = \left.\frac{d}{dt} m_X(t)\right|_{t=0} = \frac{r}{\lambda}$$

y

$$EX^2 = \left.\frac{d^2}{dt^2} m_X(t)\right|_{t=0} = \frac{r^2 + r}{\lambda^2}.$$

■

En los casos particulares $r = 1$ y $\lambda > 0$ arbitrario; $\lambda = \frac{1}{2}$ y $r = \frac{k}{2}$ con k entero positivo, y $r > 1$ y $\lambda > 0$ arbitrarios, se obtienen, respectivamente, los tipos especiales de distribuciones que se habían mencionado anteriormente, es decir, distribución exponencial, la distribución ji-cuadrada con k grados de libertad y la distribución Erlang.

Notación 4.26 *La expresión* $X \stackrel{d}{=} Exp(\lambda)$ *indica que* X *tiene distribución exponencial de parámetro* λ.

La expresión $X \stackrel{d}{=} \mathcal{X}^2_{(k)}$ *indica que* X *tiene distribución ji-cuadrada con* k *grados de libertad.*

La expresión $X \stackrel{d}{=} Erlang(r, \lambda)$ *indica que* X *tiene distribución Erlang de parámetros* r *y* λ.

Nota 4.27 *Si* $X \stackrel{d}{=} Exp(\lambda)$, *entonces,* $EX = \frac{1}{\lambda}$; $Var(X) = \frac{1}{\lambda^2}$ *y* $m_X(t) = \frac{\lambda}{\lambda - t}$; *para* $t < \lambda$.

Si $X \stackrel{d}{=} \mathcal{X}^2_{(k)}$, *entonces,* $EX = k$; $Var(X) = 2k$ *y* $m_X(t) = (1 - 2t)^{-\frac{k}{2}}$; *para* $t < \frac{1}{2}$.

La distribución exponencial se usa con frecuencia como modelo para la descripción de la distribución del tiempo transcurrido entre las ocurrencias sucesivas de un determinado suceso, como es el caso de los clientes que llegan a una entidad bancaria, llamadas que entran a una central teléfonica, etc. También se usa, la distribución exponencial, para modelar la distribución de la duración de componentes que ni se deterioran ni mejoran con la edad, esto es, aquellos para los cuales la distribución de la duración restante del componente es independiente de la edad actual. Por lo tanto, este modelo se ajusta a la realidad sólo si la distribución del tiempo de vida que le queda al elemento en cuestión no depende de su edad. Más precisamente se tiene el resultado siguiente:

Teorema 4.28 *Sea X una variable aleatoria tal que $P(X > 0) > 0$. Entonces:*

$$X \stackrel{d}{=} Exp(\lambda), \text{ si y sólo si, } P(X > x + t \mid X > t) = P(X > x)$$

para todo $x, t \in [0, \infty)$.

Demostración. *$\Longrightarrow$) Supóngase que $X \stackrel{d}{=} Exp(\lambda)$, entonces:*

$$\begin{aligned} P(X > x + t \mid X > t) &= \frac{P(X > x + t)}{P(X > t)} \\ &= \frac{\exp(-\lambda(x + t))}{\exp(-\lambda t)} \\ &= \exp(-\lambda x) \\ &= P(X > x). \end{aligned}$$

$\Longleftarrow$) Sea $G(x) = P(X > x)$. Entonces, por hipótesis:

$$G(x + t) = G(x)G(t)$$

lo cual implica que $G(x) = \exp(-\lambda x)$; con λ constante mayor que 0. En efecto,

$$\underbrace{G\left(\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}\right)}_{n-veces} = \underbrace{G\left(\frac{1}{n}\right) G\left(\frac{1}{n}\right) \cdots G\left(\frac{1}{n}\right)}_{n-veces}.$$

esto es,

$$G(1) = \left[G\left(\frac{1}{n}\right)\right]^n$$

o equivalentemente,

$$G\left(\frac{1}{n}\right) = [G(1)]^{\frac{1}{n}}.$$

Análogamente se obtiene para $m, n \in \mathbb{N}$ lo siguiente:

$$\begin{aligned} G\left(\frac{m}{n}\right) &= \left[G\left(\frac{1}{n}\right)\right]^m \\ &= [G(1)]^{\frac{m}{n}}. \end{aligned}$$

Como G es una función continua a derecha, se concluye que

$$G(x) = [G(1)]^x$$

Por otra parte, se tiene que $0 < G(1) < 1$. En efecto, si $G(1) = 1$, entonces, $G(x) = 1$, lo cual contradice que $G(\infty) = 0$. Si $G(1) = 0$ entonces $G\left(\frac{1}{m}\right) = 0$ y por continuidad a derecha se concluye que $G(0) = 0$, lo cual contradice la hipótesis. Por lo tanto, se puede tomar $\lambda := -\ln[G(1)]$ con lo cual se obtiene el resultado. ■

Ejemplo 4.29 *La duración X, en horas, de cierto componente tiene una distribución exponencial de media 100 horas. Calcular la probabilidad de que el componente dure por lo menos 200 horas.*

Solución: La función de densidad de la variable aleatoria X está dada por:

$$f(x) = \frac{1}{100}\exp(-\frac{1}{100}x)\mathcal{X}_{(0,\infty)}(x),$$

por lo tanto,

$$\begin{aligned} P(X \geq 200) &= 1 - P(X < 200) \\ &= 1 - \int_0^{200} \frac{1}{100}\exp(-\frac{1}{100}x)dx \\ &= 1 - [-\exp(-2) + 1] \\ &= \exp(-2) = 0.13534. \quad \blacktriangle \end{aligned}$$

4.4. Distribución beta

La distribución que se presenta a continuación se utiliza frecuentemente como un modelo matemático para representar variables físicas cuyos valores se encuentran restringidos en un intervalo de longitud finita, o como modelo para fracciones, tales como la proporción de impurezas en un producto químico o la fracción de tiempo que dura la reparación de una máquina.

Definición 4.30 (distribución beta) *Se dice que la variable aleatoria X tiene distribución beta de parámetros $a > 0$ y $b > 0$ si su función de densidad está dada por:*

$$f(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}\mathcal{X}_{(0,1)}(x)$$

donde $B(a,b)$ es la función beta .Esto es,

$$B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}dx.$$

Notación 4.31 *La expresión $X \stackrel{d}{=} \beta(a,b)$ significa que X tiene distribución beta de parámetros a y b.*

Las funciones beta y gamma se relacionan mediante la expresión siguiente :

$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)},$$

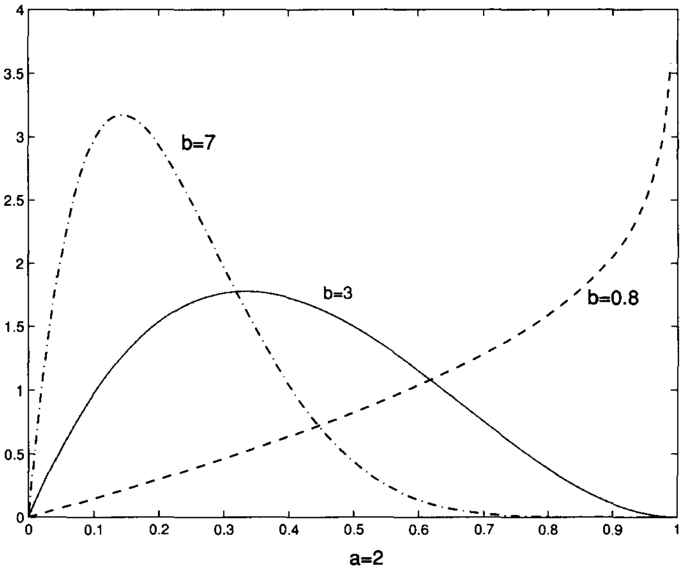
por lo tanto, la función de densidad se expresa de la forma:

$$f(x) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1}\mathcal{X}_{(0,1)}(x).$$

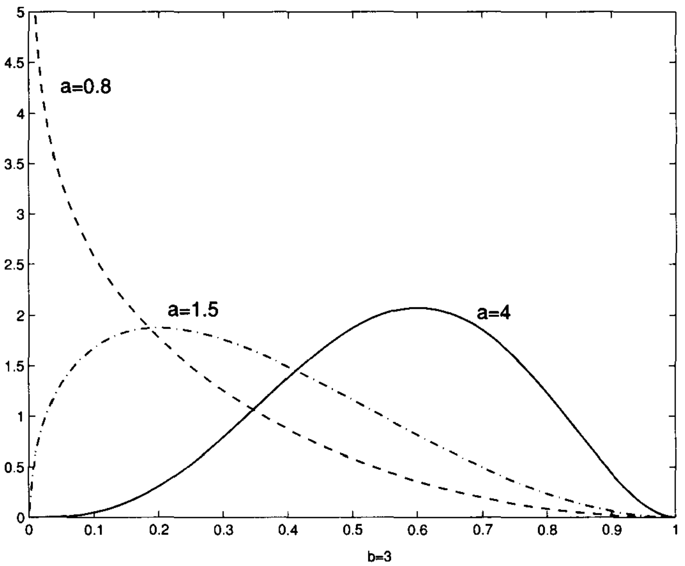
Si a y b son enteros positivos, entonces:

$$f(x) = \frac{(a+b-1)!}{(a-1)!(b-1)!} x^{a-1}(1-x)^{b-1}\mathcal{X}_{(0,1)}(x).$$

La figura siguiente muestra las gráficas de la función de densidad beta para $a = 2$ y diferentes valores de b.



La figura siguiente muestra las gráficas de la función de densidad beta para $b = 3$ y diferentes valores de a.



Es claro que si $a = b = 1$, entonces, la distribución beta coincide con la distribución uniforme sobre el intervalo $(0, 1)$. Además se tiene que:

1. Si $a > 1$ y $b > 1$, la función f tiene un máximo global.

2. Si $a > 1$ y $b < 1$, la función f es creciente.

3. Si $a < 1$ y $b > 1$, la función f es decreciente.
4. Si $a < 1$ y $b < 1$, la gráfica de f tiene forma de U.

La función de distribución de una variable aleatoria con distribución beta está dada por:

$$F(x) = \left[\int_0^x \frac{1}{B(a,b)} t^{a-1}(1-t)^{b-1} dt \right] \mathcal{X}_{(0,1)}(x) + \mathcal{X}_{[1,\infty)}(x).$$

La función generadora de momentos de una variable aleatoria con distribución beta no tiene una forma simple. Por esto, resulta conveniente encontrar sus momentos a partir de la definición.

Teorema 4.32 *Sea* $X \overset{d}{=} \beta(a,b)$, *entonces:*

1. $EX = \frac{a}{a+b}$.

2. $Var(X) = \frac{ab}{(a+b+1)(a+b)^2}$.

Demostración.

$$\begin{aligned} EX^k &= \frac{1}{B(a,b)} \int_0^1 x^{k+a-1}(1-x)^{b-1} dx \\ &= \frac{B(a+k,b)}{B(a,b)} \\ &= \frac{\Gamma(a+k)\Gamma(a+b)}{\Gamma(a+k+b)\Gamma(a)}, \end{aligned}$$

por lo tanto,

$$\begin{aligned} EX &= \frac{\Gamma(a+1)\Gamma(a+b)}{\Gamma(a+1+b)\Gamma(a)} \\ &= \frac{a\Gamma(a)\Gamma(a+b)}{(a+b+1)\Gamma(a+b)\Gamma(a)} \\ &= \frac{a}{a+b+1} \end{aligned}$$

y

$$\begin{aligned} EX^2 &= \frac{\Gamma(a+2)\Gamma(a+b)}{\Gamma(a+2+b)\Gamma(a)} \\ &= \frac{a(a+1)}{(a+b)\,(a+b+1)}. \end{aligned}$$

■

Ejemplo 4.33 *Un distribuidor mayorista de gasolina dispone de tanques de almacenamiento que contienen una cantidad fija de gasolina y que se llenan cada lunes. La proporción de esta reserva, que se vende durante la semana, es de sumo interés para el distribuidor. Mediante observaciones realizadas durante muchas semanas, se encontró que el modelo adecuado para representar esta proporción es una distribución beta con parámetros $a = 4$ y $b = 2$. Hallar la probabilidad de que el mayorista venda al menos el $90\,\%$ de su reserva durante una semana dada.*

Solución: Sea $X :=$"proporción de la reserva que se vende durante la semana". Puesto que $X \overset{d}{=} \beta(4, 2)$ se tiene que:

$$\begin{aligned} P(X \geq 0.9) &= 1 - P(X < 0.9) \\ &= 1 - \int_0^{0.9} 20x^3(1-x)dx \\ &= 0.08146. \quad \blacktriangle \end{aligned}$$

4.5. Distribución Weibull

La distribución Weibull es ampliamente usada en la ingeniería como modelo para la descripción del tiempo de duración de un componente. Esta distribución fue introducida por el científico sueco del mismo nombre, quien demostró que el esfuerzo al que se someten los materiales puede modelarse mediante el empleo de esta distribución.

Sea T la variable aleatoria que denota el tiempo de duración de un componente dado y sea f su función de densidad. Es claro que T es no negativa. Supóngase que se desea conocer la probabilidad de que el componente falle durante las próximas Δt unidades de tiempo, dado que está funcionando correctamente hasta el tiempo t. Si F es la función de distribución de la variable aleatoria T y si $F(t) < 1$,entonces,

$$\begin{aligned} P\left(t \leq T \leq t + \Delta t \mid T > t\right) &= \frac{P(t \leq T \leq t + \Delta t)}{1 - P(T \leq t)} \\ &\approx \frac{f(t)\Delta t}{1 - F(t)} =: \lambda(t). \end{aligned}$$

La función $\lambda(t)$ se conoce como *función de riesgo o tasa de falla* asociada a la variable aleatoria T. La función $R(t) := 1 - F(t)$ se conoce como *función de confiabilidad.* La expresión anterior indica que si se conoce la función de densidad del tiempo de duración del componente, entonces, se conoce su tasa de falla. A continuación se verá que el recíproco también es válido. Puesto que,

$$\lambda(t) = \frac{\frac{d}{dt}F(t)}{1 - F(t)},$$

integrando en ambos lados de la ecuación anterior, se obtiene:

$$\ln(1 - F(t)) = -\int_0^t \lambda(s)ds + C,$$

esto es:

$$F(t) = 1 - \exp(C)\exp\left(-\int_0^t \lambda(s)ds\right).$$

Es razonable suponer que $F(0) = 0$, esto es, que la probabilidad de falla instantánea del componente es cero. En tal caso se tiene que $C = 0$ y por lo tanto:

$$F(t) = \; 1 - \exp\left(-\int_0^t \lambda(s)ds\right); \quad \text{si} \quad t \geq 0$$

lo cual permite conocer la función de distribución de la variable aleatoria T, a partir de su función de riesgo. Si se supone que la tasa de falla es constante e igual a $\lambda > 0$, se ve que para $t \geq 0$:

$$F(t) = 1 - \exp(-\lambda t).$$

Esto indica que la variable aleatoria T tiene una distribución exponencial de parámetro λ.

Supóngase ahora que la función de riesgo de la variable aleatoria T está dada por:

$$\lambda(t) = \alpha\beta t^{\beta-1}$$

donde α y β son constantes positivas. En tal caso se tiene que:

$$\begin{aligned} F(t) &= \begin{cases} 1 - \exp\left(-\int_0^t \alpha\beta s^{\beta-1}ds\right) & \text{si} \quad t \geq 0 \\ 0 & \text{si} \quad t < 0 \end{cases} \\ &= \begin{cases} 1 - \exp(-\alpha t^{\beta}) & \text{si} \quad t \geq 0 \\ 0 & \text{si} \quad t < 0. \end{cases} \end{aligned}$$

La función de densidad de T está dada por:

$$f(t) = \begin{cases} \alpha\beta t^{\beta-1}\exp(-\alpha t^{\beta}) & \text{si} \quad t \geq 0 \\ 0 & \text{si} \quad t < 0. \end{cases}$$

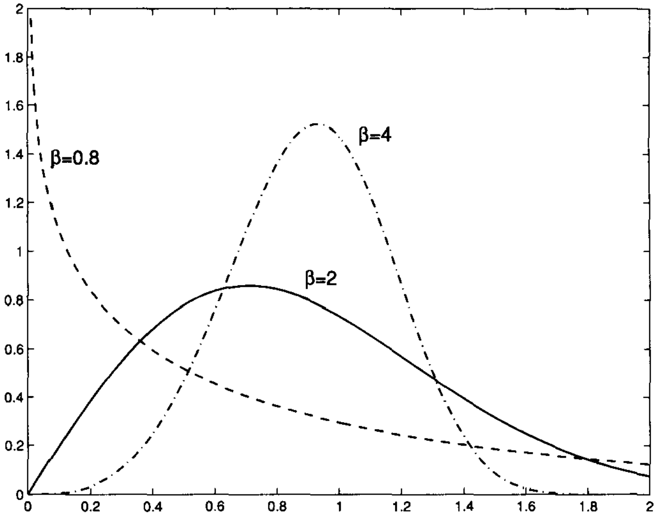
Una variable aleatoria con la función de densidad anterior recibe un nombre especial:

Definición 4.34 (distribución Weibull) *Se dice que una variable aleatoria X tiene distribución Weibull de parámetros α y β, si su función de densidad está dada por:*

$$f(x) = \begin{cases} \alpha\beta x^{\beta-1}\exp(-\alpha x^{\beta}) & si \quad x \geq 0 \\ 0 & si \quad x < 0. \end{cases}$$

Notación 4.35 *La expresión $X \stackrel{d}{=} W(\alpha, \beta)$ indica que la variable aleatoria X tiene una distribución Weibull de parámetros α y β.*

La figura siguiente muestra la gráfica de la función de densidad Weibull para $\alpha = 1$ y diferentes valores de β



Teorema 4.36 *Sea $X \stackrel{d}{=} W(\alpha, \beta)$, entonces:*

1. $EX = \left(\frac{1}{\alpha}\right)^{\frac{1}{\beta}} \Gamma(1 + \frac{1}{\beta})$

2. $Var(X) = \left(\frac{1}{\alpha}\right)^{\frac{2}{\beta}} \left[\Gamma(1 + \frac{2}{\beta}) - \Gamma^2(1 + \frac{1}{\beta})\right]$

Demostración. *Como ejercicio.* ■

Nota 4.37 *Algunos autores, como por ejemplo S. Ross (ver [Ros2]) y F. M. Hernández (ver [Her]) definen la función de densidad de una distribución Weibull considerando tres parámetros: un parámetro de localización c, un* *párametro de escala a y un parámetro de forma b y dicen que una variable aleatoria X tiene distribución Weibull de parámetros a, b y c, si su función de densidad está dada por:*

$$f(x) = \begin{cases} \frac{b}{a}\left[\frac{x}{a}\right]^{b-1} \exp\left[-\left(\frac{x}{a}\right)^{b}\right] & si \quad x > c \\ 0 & si \quad x \leq c \end{cases}.$$

Sin embargo, en la mayoría de las aplicaciones se acostumbra a hacer $c = 0$, *obteniéndose la función de densidad considerada inicialmente, al tomar* $\alpha = \left[\frac{1}{a}\right]^{b}$ *y* $\beta = b$.

4.6. Otras distribuciones continuas

Definición 4.38 (distribución Cauchy) *Se dice que una variable aleatoria X tiene distribución Cauchy de parámetros* θ *y* β, $\theta \in \mathbb{R}$ *y* $\beta \in \mathbb{R}^{+}$, *si su función de densidad está dada por:*

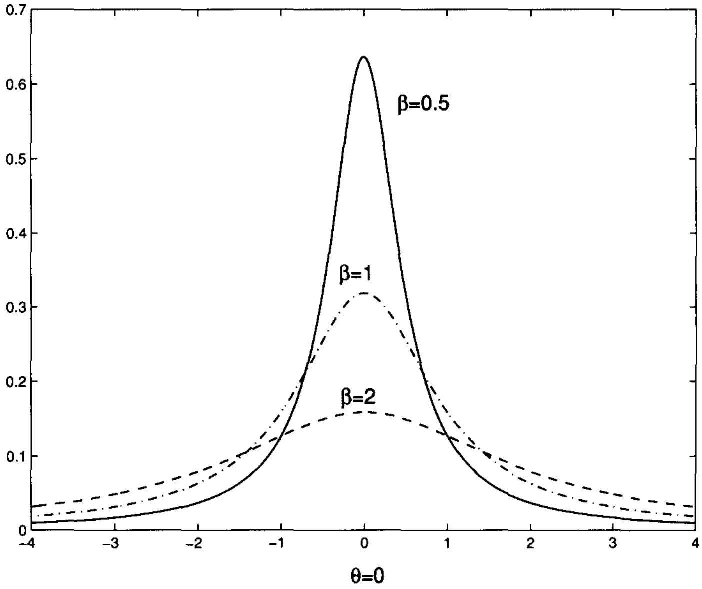
$$f(x) = \frac{1}{\pi\beta}\frac{1}{1+\left(\frac{x-\theta}{\beta}\right)^{2}}, \quad x \in \mathbb{R}.$$

Cuando $\theta = 0$ *y* $\beta = 1$ *se obtiene:*

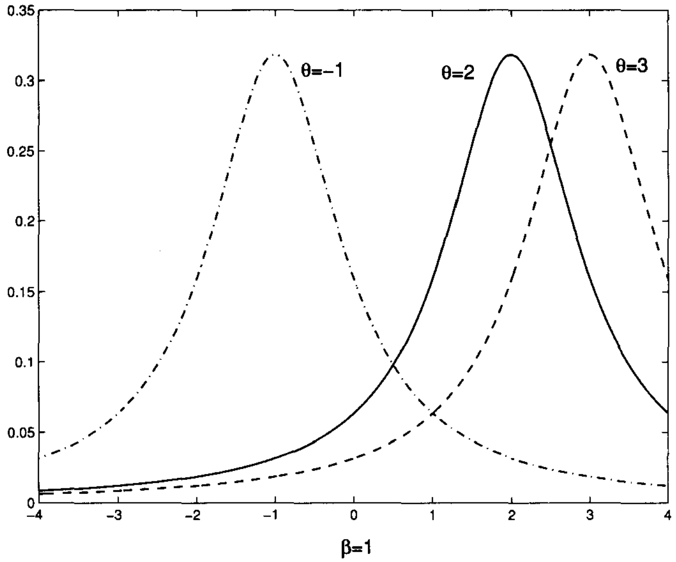
$$f(x) = \frac{1}{\pi(1+x^{2})}$$

la cual se conoce como función de densidad Cauchy estandarizada.

La figura siguiente muestra las gráficas de f para $\theta = 0$ y algunos valores de β.



La figura siguiente muestra las gráficas de f para $\beta = 1$ y algunos valores de θ.



La función de distribución de una variable aleatoria con distribución Cauchy está dada por:

$$F(x) = \frac{1}{2} + \frac{1}{\pi} \arctan\left(\frac{x - \theta}{\beta}\right).$$

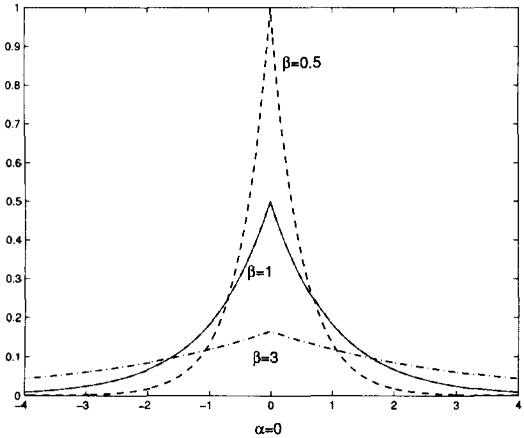
La distribución Cauchy tiene la característica de tener colas pesadas, esto significa que valores lejanos a θ tienen probabilidades grandes de ocurrir. Es por esto, que esta distribución presenta un comportamiento atípico en varios sentidos y es un excelente contraejemplo para varias afirmaciones que en principio resultarían razonables. Recuerde por ejemplo que, en el capítulo 2., se probó que el valor esperado de una variable aleatoria con distribución Cauchy no existe.

Definición 4.39 (distribución de Laplace) *Se dice que una variable aleatoria X tiene una distribución de Laplace o exponencial doble de parámetros α, β, si su función de densidad está dada por:*

$$f(x) = \frac{1}{2\beta} \exp\left(-\frac{|x-\alpha|}{\beta}\right),$$

donde $\alpha \in \mathbb{R}$ y $\beta \in \mathbb{R}^+$.

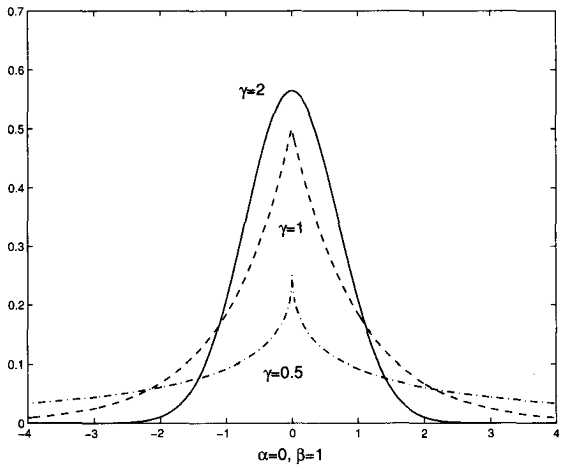
La figura siguiente muestra la gráfica de una función de densidad de Laplace para $\alpha = 0$ y valores diferentes de β.



Definición 4.40 (potencia exponencial) *Se dice que una variable aleatoria X se distribuye como una potencia exponencial de parámetros α, β y γ; con $\alpha \in \mathbb{R}$ y $\beta, \gamma \in \mathbb{R}^+$, si su función de densidad está dada por:*

$$f(x) = \frac{1}{2\beta\Gamma\left(1+\frac{1}{\gamma}\right)} \exp\left[-\left|\frac{x-\alpha}{\beta}\right|^{\gamma}\right].$$

La figura siguiente muestra las gráficas de una función de densidad de una variable aleatoria con distribución potencia exponencial para $\alpha = 0$, $\beta = 1$ y valores diferentes de γ.



Si $\gamma = 2$ entonces

$$f(x) = \frac{1}{\sqrt{\pi}\beta} \exp\left[-\left(\frac{x-\alpha}{\beta}\right)^2\right],$$

esto es, se obtiene la función de densidad de una variable aleatoria con distribución normal de parámetros $\mu = \alpha$ y $\sigma = \frac{\beta}{\sqrt{2}}$.

Si $\gamma = 1$ se obtiene

$$f(x) = \frac{1}{2\beta} \exp\left(-\frac{|x-\alpha|}{\beta}\right)$$

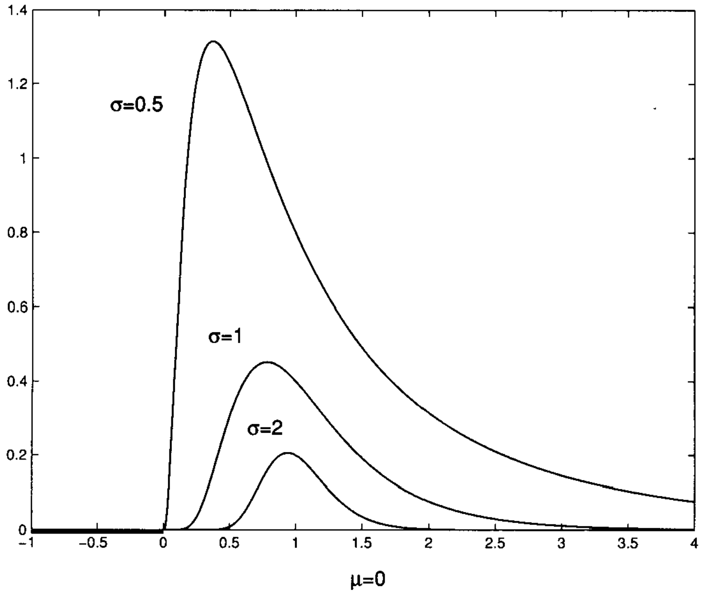
la cual es la función de densidad de Laplace de parámetros α y β.

Definición 4.41 (distribución lognormal) *Sean X una variable no negativa y $Y := \ln X$. Si la variable aleatoria Y tiene distribución normal de parámetros μ y σ,entonces, se dice que Y tiene distribución lognormal de parámetros μ y σ.*

Es claro que si Y tiene distribución lognormal, su función de densidad está dada por:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma x} \exp\left[-\left(\frac{\ln x - \mu}{\sqrt{2}\sigma}\right)^2\right] \mathcal{X}_{(0,\infty)}(x).$$

La figura siguiente muestra la gráfica de una función de densidad de lognormal para $\mu = 0$ y valores diferentes de σ.



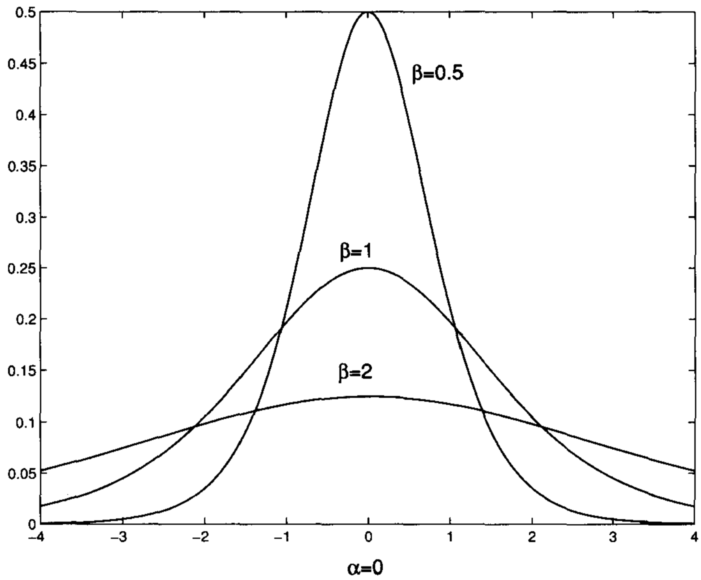
Definición 4.42 (distribución logística) *Se dice que una variable aleatoria X tiene distribución logística de parámetros α y β, con $\alpha \in \mathbb{R}$ y $\beta \in \mathbb{R}^+$ si su función de densidad está dada por:*

$$f(x) = \frac{1}{\beta} \frac{\exp\left[-\left(\frac{x-\alpha}{\beta}\right)\right]}{\left[1 + \exp\left[-\left(\frac{x-\alpha}{\beta}\right)\right]\right]^2} \; ; \quad x \in \mathbb{R}$$

La función de distribución de una variable aleatoria con distribución logística de parámetros α y β está dada por:

$$F(x) = \frac{1}{1 + \exp\left[-\left(\frac{x-\alpha}{\beta}\right)\right]}.$$

La figura siguiente muestra la gráfica de una función de densidad logística para $\alpha = 0$ y valores diferentes de β.



4.7. Ejercicios

1. Sea X una variable aleatoria con distribución uniforme continua en el intervalo $\left[-\frac{1}{2}, \frac{3}{2}\right]$.

   a) Calcular: la media, la varianza y la desviación estándar de X.

   b) Determinar el valor de x para el cual $P\left(|X| < x\right) = 0.9$.

2. Se escoge un número al azar en el intervalo $(0,1)$. Calcular:

   a) La probabilidad de que el primer dígito a la derecha del punto decimal es 6.

   b) La probabilidad de que el segundo dígito a la derecha del punto decimal es 1.

   c) La probabilidad de que el segundo dígito a la derecha del punto decimal es 8 dado que el primer dígito es 3.

3. Sea $X \stackrel{d}{=} \mathcal{U}(a,b)$. Si $EX = 2$ y $Var(X) = \frac{3}{4}$, ¿cuáles son los valores de los parámetros a y b?

4. Un escolar llega al paradero de su bus a las 6 : 00 am en punto, sabiendo que el bus llega en algún momento, distribuido uniformemente

entre las $6:00$ am y las $6:20$ am, ¿cuál es la probabilidad de que el escolar tenga que esperar más de cinco minutos? Si a las $6:10$ am no ha pasado el bus todavía, ¿cuál es la probabilidad de que el escolar tenga que esperar por lo menos 5 minutos más?

5. Supóngase que X es una variable aleatoria con distribución uniforme sobre el intervalo $(0,5)$. Calcular la probabilidad de que las raíces de la ecuación $4x^2+4xX+X+2=0$ sean ambas reales.

6. Sea X una variable aleatoria con distribución uniforme sobre $(-1,1)$. Hallar:

   a) $P\left(|X|>\frac{1}{2}\right)$
   b) Una función de densidad de la variable aleatoria $Y=|X|$.

7. Sea X una variable aleatoria con distribución uniforme sobre $(0,1)$. Hallar funciones de densidad para las siguientes variables aleatorias:

   a) $Y:=\ln X$
   b) $Z:=X^3+2$
   c) $W:=\frac{1}{X}$

8. Un jugador lanza un dardo a una diana. Supóngase que el jugador recibe 10 puntos si su lanzamiento cae a 2 cm. del blanco, 5 puntos si cae entre 2 cm. y 6 cm. del blanco y 3 puntos si cae entre 6 cm. y 10 cm. del blanco. Hallar el número esperado de puntos que obtiene el jugador, sabiendo que la distancia entre el sitio en que cae su lanzamiento y el blanco es una variable aleatoria con distribución uniforme en $[0,10]$.

9. Sea $X\overset{d}{=}\mathcal{N}(0,1)$. Calcular:

   a) $P(0\leq X\leq 1.42)$
   b) $P(-0.73\leq X\leq 0)$
   c) $P(-1.37\leq X\leq 2.01)$
   d) $P(X\geq 1.13)$
   e) $P(|X|\leq 0.5)$.

10. Sea $X\overset{d}{=}\mathcal{N}(5,16)$. En cada uno de los ejercicios siguientes, obtener el valor de x que resuelve la ecuación:

a) $P(X > x) = 0.5$

b) $P(x < X < 9) = 0.2$

c) $P(X \geq x) = 0.01$

11. En los ejercicios que siguen a continuación encontrar los valores de c que satisfacen las igualdades.

a) $P(W \leq c) = 0.95$ si $W \stackrel{d}{=} \chi_8^2$

b) $P(|X - 2| > 5) = c$ si $X \stackrel{d}{=} \mathcal{N}(1, 4)$

c) $P(W > c) = 0.25$ si $W \stackrel{d}{=} \chi_{10}^2$

d) $P(X < c) = 0.05$ si $X \stackrel{d}{=} \mathcal{N}(10, 5)$

e) $P\left(\chi_k^2 \leq c\right) = 0.025$; para $k = 6$, $k = 15$, $k = 25$

12. El volumen que una máquina de llenado automático deposita en latas de una bebida gaseosa tiene una distribución normal de media 12.4 onzas de líquido y desviación estándar de 0.1 onzas de líquido.

a) ¿Cuál es la probabilidad de que el volumen depositado sea menor que 12 onzas de líquido?

b) Si se desechan todas las latas que tienen menos de 12.1 o más de 12.6 onzas de líquido,¿cuál es la proporción de latas desechadas?

13. Supóngase que los puntajes de un examen están distribuidos normalmente con media 76 y desviación estándar 15. El 15 % de los estudiantes, los mejores, obtienen A como nota y el 10 %, los peores, pierden el curso y obtienen P.

a) Hallar el puntaje mínimo para obtener A como calificación.

b) Hallar el puntaje mínimo para aprobar.

14. El tiempo de vida útil de una impresora de burbuja es una variable aleatoria normal de media 5.2 años y desviación estándar 1.4 años. ¿Qué porcentaje de estas impresoras tienen una vida útil de menos de 7 años, menos de 3 años y entre 3 y 7 años?.

15. Sea X una variable aleatoria con distribución normal de media μ y varianza σ^2. Hallar la distribución de la variable aleatoria $Y := 5X - 1$.

16. Sea X una variable aleatoria con distribución binomial de parámetros $n = 30$ y $p = 0.3$. ¿Es razonable aproximar esta distribución a una normal de parámetros $\mu = 9$ y varianza $\sigma^2 = 6.3$? Explicar.

17. Se lanza un dado corriente 1000 veces consecutivas. Calcular la probabilidad de que el número 3 se obtenga menos de 150 veces dado que el número 1 se obtuvo 200 veces exactamente.

18. Sea $X \stackrel{d}{=} \mathcal{N}(3,4)$. Encontrar un número α tal que

$$P(X > \alpha) = 2P(X \leq \alpha)$$

19. Determinar los déciles de la distribución normal estándar, esto es los valores $x_{0.1}, x_{0.2}, \cdots, x_{0.9}$ tales que $\Phi(x_{0.i}) = 0.i$ para $i = 1, \cdots, 9$.

20. Un estudio de mercadeo determinó que la demanda diaria por un reconocido diario capitalino es una variable aleatoria con distribución normal de media $\mu = 50000$ y desviación estándar $\sigma = 12500$. Cada periódico que se vende deja una ganancia de 500 pesos, en tanto que cada periódico no vendido deja una pérdida de 300 pesos. ¿Cuánto debe ser el tiraje diario del periódico para que la ganancia neta esperada sea máxima?

21. Sea X una variable aleatoria con :

    a) distribución uniforme sobre $[-1,1]$

    b) distribución exponencial de parámetro λ

    c) distribución normal de parámetros μ y σ

    Calcular la función de distribución F_Y y una función de densidad f_Y de la variable aleatoria $Y = aX + b$, donde a y b son números reales y $a \neq 0$.

22. En una oficina de reclamos de una empresa de servicio público, se tiene que el tiempo (en minutos) que dura el empleado en atender el reclamo de un usuario, es una variable aleatoria con distribución exponencial de media 15 minutos. Si usted llega a las 12 en punto a la oficina de reclamos y en ese momento no hay cola de espera, pero el empleado está atendiendo a un usuario ¿cuál es la probabilidad de que tenga que esperar menos de 5 minutos en ser atendido?

23. Supóngase que el número de kilómetros que puede recorrer un automóvil antes de que se le acabe la batería está distribuido exponencialmente con un valor promedio de 10000 Km. Si una persona quiere realizar un viaje de 5000 Km, ¿cuál es la probabilidad de que llegue al final de su viaje sin tener que cambiar la batería?

24. El tiempo que transcurre entre las llamadas a una fábrica de artículos deportivos tiene una distribución exponencial con un tiempo promedio entre llamadas de 15 minutos.

    a) ¿Cuál es la probabilidad de que no haya llamadas en un lapso de 30 minutos?

    b) ¿Cuál es la probabilidad de recibir al menos una llamada en un intervalo de 10 minutos?

    c) ¿Cuál es la probabilidad de recibir la primera llamada entre 5 y 10 minutos después de abrir la fábrica?

25. La duración T de un componente electrónico es una variable aleatoria con distribución exponencial de parámetro λ.

    a) Determinar la probabilidad de que el componente electrónico funcione por lo menos hasta $t = 3\lambda^{-1}$

    b) ¿Cuál es la probabilidad de que el componente electrónico funcione por lo menos hasta $t = k\lambda^{-1}$ si hasta el tiempo $t = (k-1)\lambda^{-1}$ funciona?

26. Sea X una variable aleatoria con distribución exponencial de parámetro λ. Hallar funciones de densidad para las siguientes variables aleatorias:

    a) $Y := \ln X$.

    b) $Z := X^2 + 1$.

    c) $W := \frac{1}{X}$.

27. a. El actuario Benjamín Gompertz propuso en 1825 la función de riesgo $\lambda(t)$ dada por:

$$\lambda(t) = \alpha\beta^t \mathcal{X}_{(0,\infty)}(t), \text{ con } \alpha > 0 \text{ y } \beta > 1,$$

para modelar el tiempo de vida de los seres humanos. Determinar la función de distribución F correspondiente a la función de falla $\lambda(t)$.

b. En el año 1860 Makeham modificó la función de riesgo propuesta por Gompertz y sugirió la siguiente:

$$\lambda(t) = \left(\gamma + \alpha\beta^t\right) \mathcal{X}_{(0,\infty)}(t) \quad \text{con } \alpha, \beta \in \mathbb{R}^+ \text{ y } \beta > 1.$$

Determinar la función de distribución F correspondiente a la función de falla $\lambda(t)$ (esta distribución recibe el nombre de distribución de Makeham)

28. Calcular la función de falla de una distribución exponencial de parámetro μ.

29. Determinar la distribución F cuya función de falla está dada por:

$$\lambda(t) = \alpha + \beta t; \; t \geq 0$$

donde α y β son constantes.

30. Supóngase que T denota el tiempo de duración de cierto componente y que la función de riesgo asociada a T está dada por:

$$\lambda(t) = \frac{1}{(t+1)}.$$

Hallar la distribución de T.

31. Supóngase que T denota el tiempo de duración de cierto componente y que la función de riesgo asociada a T está dada por:

$$\lambda(t) = \alpha t^{\beta}$$

donde $\alpha > 0$ y $\beta > 0$ son constantes. Hallar la distribución de T.

32. El tiempo de duración de cierto componente electrónico tiene una distribución Weibull con $\beta = 0.5$ y una vida media de 600 horas. Calcular la probabilidad de que el componente dure al menos 500 horas.

33. El tiempo transcurrido (en meses después del mantenimiento) antes de fallar el equipo de vigilancia por circuito cerrado de TV de un banco, tiene una distribución Weibull con $\alpha = 2$ y $\beta = \frac{1}{60}$. Si el banco quiere que la probabilidad de un daño, antes del siguiente mantenimiento programado, sea de 0.05, ¿con qué frecuencia debe recibir mantenimiento periódico el equipo de vigilancia?

34. Una investigación sobre los gastos para el control de la contaminación incurridos por empresas industriales reveló, que el porcentaje anual del cierre de la capacidad de planta atribuible a la reglamentación ambiental y de seguridad, tiene una distribución beta aproximada con $a = 1$ y $b = 25$.

    a) Calcular la media y la varianza del porcentaje anual de cierre de la capacidad de planta atribuible a la reglamentación ambiental y de seguridad.

    b) Calcular la probabilidad de que más del $1\,\%$ del cierre de la capacidad de planta sea atribuible a la reglamentación ambiental y de seguridad.

35. Sea X una variable aleatoria con distribución normal estándar. Hallar $E(|X|)$.

36. Sea X una variable aleatoria continua con función de distribución estrictamente creciente F. Sea $Y = F(X)$. Demostrar que Y tiene distribución uniforme sobre $(0,1)$.

37. Sea X una variable aleatoria continua con función de distribución estrictamente creciente F. ¿Qué tipo de distribución tiene la variable aleatoria $Y := -\ln(F(X))$?

38. Sea W una variable aleatoria con distribución uniforme en el intervalo $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Determinar una función de densidad de la variable aleatoria $Z := \alpha \tan W$; donde $\alpha \in (0,\infty)$ es una constante fija.

39. Sea X una variable aleatoria con distribución uniforme sobre $(0,1)$. Hallar una función $g : \mathbb{R} \longrightarrow \mathbb{R}$, tal que $Y = g(X)$ tenga distribución normal estándar.

40. Demostrar que si X es una variable aleatoria con distribución gamma de parámetros r y λ, con r entero positivo, entonces:

$$F_X(x) = 1 - \sum_{j=0}^{r-1} \frac{(\lambda x)^j \exp(-\lambda x)}{j!}.$$

41. Demostrar que $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$.

42. Demostrar que:
$$B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

43. Sea X una variable aleatoria con distribución Cauchy estándar. Demostrar que EX no existe.

44. Sea X una variable aleatoria con distribución Cauchy estándar, ¿qué tipo de distribución tiene la variable aleatoria $Y := \frac{1}{X}$?

45. Sea X una variable aleatoria con distribución normal estándar. Demostrar que:
$$\left(x^{-1} - x^{-3}\right)\exp\left(-\frac{1}{2}x^2\right) < \sqrt{2\pi}\left[P(X > x)\right] < x^{-1}\exp\left(-\frac{1}{2}x^2\right),$$
para $x > 0$.

Capítulo 5

Vectores Aleatorios

En este capítulo se estudiará el comportamiento conjunto de dos o más variables aleatorias.

5.1. Distribución conjunta de variables aleatorias

En muchos casos es necesario considerar el comportamiento conjunto de dos o más variables. Supóngase, por ejemplo, que se lanza una moneda corriente tres veces consecutivas y que se desea analizar el comportamiento conjunto de las variables aleatorias X e Y definidas como:

$X :=$ "número de caras obtenidas en los primeros dos lanzamientos"

$Y :=$ "número de caras obtenidas en los dos últimos lanzamientos"

Es claro que:

$$
\begin{aligned}
P(X = 0, Y = 0) &= P((S, S, S)) = \frac{1}{8} \\
P(X = 0, Y = 1) &= P((S, S, C)) = \frac{1}{8} \\
P(X = 1, Y = 0) &= P((C, S, S)) = \frac{1}{8} \\
P(X = 1, Y = 1) &= P(\{(C, S, C), (S, C, S)\}) = \frac{1}{4} \\
P(X = 1, Y = 2) &= P((S, C, C)) = \frac{1}{8} \\
P(X = 2, Y = 1) &= P((C, C, S)) = \frac{1}{8} \\
P(X = 2, Y = 2) &= P((C, C, C)) = \frac{1}{8}
\end{aligned}
$$

Esta información se puede resumir en la tabla siguiente:

$X\backslash Y$	0	1	2
0	$\frac{1}{8}$	$\frac{1}{8}$	0
1	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$
2	0	$\frac{1}{8}$	$\frac{1}{8}$

Definición 5.1 (vector aleatorio $n-$dimensional) *Sean $X_1, X_2, \cdots, X_n$, n variables aleatorias reales definidas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$. La $\Im - \mathcal{B}_n-$variable aleatoria $\mathbf{X} : \Omega \longrightarrow \mathbb{R}^n$ definida por:*

$$\mathbf{X}(\omega) := (X_1(\omega), \cdots, X_n(\omega))$$

se llama vector aleatorio $n-$dimensional.

Definición 5.2 (distribución de un vector aleatorio) *Sea $\mathbf{X}$ un vector aleatorio n-dimensional. La medida de probabilidad definida sobre $\mathcal{B}_n$ por:*

$$P_{\mathbf{X}}(B) := P(\mathbf{X} \in B) \ ; \ B \in \mathcal{B}_n$$

se llama distribución del vector aleatorio $\mathbf{X}$.

Definición 5.3 (función de densidad conjunta)
Sea $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ un vector aleatorio $n-$dimensional. Si las variables aleatorias X_i; con $i = 1, \cdots, n$, son todas discretas, se dice que el vector aleatorio $\mathbf{X}$ es discreto. En tal caso, se define la función de densidad de $\mathbf{X}$, o función de densidad de probabilidad conjunta de las variables aleatorias $X_1, X_2, \cdots, X_n$, como:

$$f_{\mathbf{X}}(\mathbf{x}) := \begin{cases} P(\mathbf{X} = \mathbf{x}) & si \quad \mathbf{x} \text{ pertenece al rango de } \mathbf{X} \\ 0 & en \quad \text{otro caso.} \end{cases}$$

Nota 5.4 *Sean X_1 y X_2 variables aleatorias discretas. Entonces:*

$$\begin{aligned} P(X_1 = x) &= P\left((X_1 = x) \cap \bigcup_y (X_2 = y)\right) \\ &= P\left(\bigcup_y (X_1 = x, X_2 = y)\right) \\ &= \sum_y P(X_1 = x, X_2 = y). \end{aligned}$$

En general, se tiene que:

Teorema 5.5 *Sea* $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ *un vector aleatorio* $n-$*dimensional discreto. Entonces, para todo* $j = 1, \cdots, n$ *se satisface:*

$$P(X_j = x) = \sum_{x_1} \cdots \sum_{x_{j-1}} \sum_{x_{j+1}} \cdots \sum_{x_n} P(X_1 = x_1, \cdots, X_{j-1} = x_{j-1}, X_j = x, X_{j+1} = x_{j+1}, \cdots, X_n = x_n).$$

La función

$$f_{X_j}(x) := \begin{cases} P(X_j = x) & si \quad x \text{ pertenece al rango de } X_j \\ 0 & en \quad \text{otro caso.} \end{cases}$$

se llama distribución marginal de la variable aleatoria X_j.

Ejemplo 5.6 *Sean* X *e* Y *variables aleatorias discretas con distribución conjunta dada por:*

$X\backslash Y$	0	1	2	3
-1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{1}{16}$	$\frac{5}{16}$
1	$\frac{1}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	0

.

Las distribuciones marginales de X *y de* Y *están dadas, respectivamente, por:*

x	-1	1
$P(X = x)$	$\frac{10}{16}$	$\frac{6}{16}$

y

y	0	1	2	3
$P(Y = y)$	$\frac{2}{16}$	$\frac{6}{16}$	$\frac{3}{16}$	$\frac{5}{16}$

▲

Ejemplo 5.7 *Supóngase que se lanza una moneda corriente tres veces consecutivas y sean* X *e* Y *las variables aleatorias definidas como sigue:*

$X :=$ *"número de caras obtenidas"*

$Y :=$ *"número de lanzamiento en el que se obtiene por primera vez cara",* *(*$Y = 0$ *si no hay).*

1. *Hallar la distribución conjunta de* X *e* Y.

2. *Hallar las distribuciones marginales de* X *e* Y.

3. *Calcular* $P(X \leq 2, Y = 1)$, $P(X \leq 2, Y \leq 1)$ *y* $P(X \leq 2 \text{ o } Y \leq 1)$.

Solución:

1. La distribución conjunta de X e Y está dada por:

$X\backslash Y$	0	1	2	3
0	$\frac{1}{8}$	0	0	0
1	0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
2	0	$\frac{2}{8}$	$\frac{1}{8}$	0
3	0	$\frac{1}{8}$	0	0

.

2. Las distribuciones marginales de X e Y están dadas, respectivamente, por:

x	0	1	2	3
$P(X = x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

y

y	0	1	2	3
$P(Y = y)$	$\frac{1}{8}$	$\frac{1}{2}$	$\frac{1}{4}$	$\frac{1}{8}$

.

3. De los numerales anteriores se deduce:

$$
\begin{aligned}
P(X \leq 2, Y = 1) &= P(X = 0, Y = 1) + P(X = 1, Y = 1) \\
&\quad + P(X = 2, Y = 1) \\
&= \frac{3}{8} \\
P(X \leq 2, Y \leq 1) &= P(X \leq 2, Y = 1) + P(X \leq 2, Y = 0) \\
&= \frac{3}{8} + P(X = 0, Y = 0) + P(X = 1, Y = 0) \\
&\quad + P(X = 2, Y = 0) \\
&= \frac{1}{2} \\
P(X \leq 2 \text{ o } Y \leq 1) &= P(X \leq 2) + P(Y \leq 1) - P(X \leq 2, Y \leq 1) \\
&= P(X = 0) + P(X = 1) + P(X = 2) \\
&\quad + P(Y = 0) + P(Y = 1) - \frac{1}{2} \\
&= \frac{7}{8} + \frac{5}{8} - \frac{1}{2} = 1. \quad \blacktriangle
\end{aligned}
$$

Ejemplo 5.8 *Una caja contiene tres puntillas, cuatro tachuelas y dos tornillos. Se extraen, al azar, tres objetos(sin reemplazo). Sea X el número de tachuelas y Y el número de puntillas extraídas. Hallar la distribución conjunta de X e Y.*

Solución: Puesto que:

$$P(X = x, Y = y) = \frac{\binom{4}{x}\binom{3}{y}\binom{2}{3-(x+y)}}{\binom{9}{3}}; \quad \text{para } x, y = 0, 1, 2, 3,$$

se tiene que la distribución conjunta de las variables X e Y está dada por:

$X\backslash Y$	0	1	2	3
0	0	$\frac{3}{84}$	$\frac{6}{84}$	$\frac{1}{84}$
1	$\frac{4}{84}$	$\frac{24}{84}$	$\frac{12}{84}$	0
2	$\frac{12}{84}$	$\frac{18}{84}$	0	0
3	$\frac{4}{84}$	0	0	0

Las distribuciones marginales de X e Y son, respectivamente,

x	0	1	2	3
$P(X = x)$	$\frac{10}{84}$	$\frac{40}{84}$	$\frac{30}{84}$	$\frac{4}{84}$

y

y	0	1	2	3
$P(Y = y)$	$\frac{20}{84}$	$\frac{45}{84}$	$\frac{18}{84}$	$\frac{1}{84}$

▲

Ejemplo 5.9 *Se lanza un dado corriente dos veces consecutivas. Sean:*

X_1 := *"máximo valor obtenido"*

X_2 := *"suma de los resultados obtenidos"*

La función de densidad del vector aleatorio $\mathbf{X} = (X_1, X_2)$ *está dada por:*

$X_1\backslash X_2$	2	3	4	5	6	7	8	9	10	11	12
1	$\frac{1}{36}$	0	0	0	0	0	0	0	0	0	0
2	0	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0	0	0
3	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0	0	0
4	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0	0	0
5	0	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	0	0
6	0	0	0	0	0	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Las distribuciones marginales de X_1 *y* X_2 *son, respectivamente,*

x	1	2	3	4	5	6
$P(X_1 = x)$	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{5}{36}$	$\frac{7}{36}$	$\frac{9}{36}$	$\frac{11}{36}$

y

y	2	3	4	5	6	7	8	9	10	11	12
$P(X_2 = y)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

.

Asi mismo, se tiene que:

$$
\begin{aligned}
P\left(X_1 \leq \frac{3}{2}, X_2 \leq \frac{11}{3}\right) &= P\left(X_1 = 1, X_2 = 2\right) + P\left(X_1 = 1, X_2 = 3\right) \\
&= \frac{1}{36} \\
P\left(X_1 \leq \pi, X_2 \leq 2\right) &= P\left(X_1 = 1, X_2 = 2\right) + P\left(X_1 = 2, X_2 = 2\right) \\
&\quad + P\left(X_1 = 3, X_2 = 2\right) \\
&= \frac{3}{36}.
\end{aligned}
$$

y en general:

$X_2 \backslash X_1$	$x_1 < 1$	$1 \leq x_1 < 2$	$2 \leq x_1 < 3$	$3 \leq x_1 < 4$	$4 \leq x_1 < 5$	$5 \leq x_1 < 6$	$x_1 \geq 6$
$x_2 < 2$	0	0	0	0	0	0	0
$2 \leq x_2 < 3$	0	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
$3 \leq x_2 < 4$	0	$\frac{1}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$	$\frac{3}{36}$
$4 \leq x_2 < 5$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$	$\frac{6}{36}$
$5 \leq x_2 < 6$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{8}{36}$	$\frac{10}{36}$	$\frac{10}{36}$	$\frac{10}{36}$
$6 \leq x_2 < 7$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{13}{36}$	$\frac{15}{36}$	$\frac{15}{36}$
$7 \leq x_2 < 8$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{15}{36}$	$\frac{19}{36}$	$\frac{21}{36}$
$8 \leq x_2 < 9$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{22}{36}$	$\frac{26}{36}$
$9 \leq x_2 < 10$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{24}{36}$	$\frac{30}{36}$
$10 \leq x_2 < 11$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{25}{36}$	$\frac{33}{36}$
$11 \leq x_2 < 12$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{25}{36}$	$\frac{35}{36}$
$x_2 \geq 12$	0	$\frac{1}{36}$	$\frac{4}{36}$	$\frac{9}{36}$	$\frac{16}{36}$	$\frac{25}{36}$	1

donde cada componente, de la tabla, indica la probabilidad de que

$$P(X_1 \leq x_1, X_2 \leq x_2)$$

para los valores de x_1 *y* x_2*, indicados en la primera fila y primera columna, respectivamente.* ▲

Lo anterior, motiva la definición siguiente:

Definición 5.10 (función de distribución acumulativa conjunta) *Sea* $\mathbf{X} = (X_1, X_2, ..., X_n)$ *un vector aleatorio* $n-$*dimensional. La función definida por:*

$$F(x_1, x_2, ..., x_n) := P(X_1 \leq x_1, X_2 \leq x_2, ..., X_n \leq x_n),$$

para todo $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$*; se llama función de distribución acumulativa conjunta de las variables aleatorias* $X_1, X_2, ..., X_n$ *o función de distribución del vector aleatorio* $n-$*dimensional* $\mathbf{X}$.

Nota 5.11 *Al igual que en el caso unidimensional, se tiene que, la distribución del vector aleatorio* $\mathbf{X}$ *queda completamente determinada por su función de distribución.*

Nota 5.12 *Sean* X_1 *y* X_2 *variables aleatorias con función de distribución acumulativa conjunta* F*. Entonces:*

$$\begin{aligned}
F_{X_1}(x) &= P\left(X_1 \leq x\right) \\
&= P\left((X_1 \leq x) \cap \bigcup_k (X_2 \leq k)\right) \\
&= P\left(\bigcup_k (X_1 \leq x, X_2 \leq k)\right) \\
&= \lim_{k \to \infty} P\left(X_1 \leq x, X_2 \leq k\right) \\
&= \lim_{k \to \infty} F(x, k).
\end{aligned}$$

Análogamente, se deduce que $F_{X_2}(y) = \lim_{x \to \infty} F(x, y)$*. En general, se tiene el teorema siguiente:*

Teorema 5.13 *Sea* $\mathbf{X} = (X_1, X_2, ..., X_n)$ *un vector aleatorio* $n-$*dimensional, con función de distribución acumulativa conjunta* F*. Para cada* $j = 1, \cdots, n$*; se tiene que la función de distribución acumulativa de la variable aleatoria* X_j *está dada por:*

$$F_{X_j}(x) = \lim_{x_1 \to \infty} \cdots \lim_{x_{j-1} \to \infty} \lim_{x_{j+1} \to \infty} \cdots \lim_{x_n \to \infty} F(x_1, \cdots, x_n).$$

La función de distribución F_{X_j} *se llama función de distribución acumulativa marginal de la variable aleatoria* X_j.

El teorema anterior indica entonces que, si se conoce la función de distribución acumulativa conjunta de las variables aleatorias $X_1, \cdots, X_n$, se conocen las marginales. El recíproco, en general, no es válido.

A continuación se presentan algunas de las propiedades de la función de distribución conjunta.

Teorema 5.14 *Sea $\mathbf{X} = (X_1, X_2, ..., X_n)$ un vector aleatorio n-dimensional. La función de distribución acumulativa conjunta F de las variables aleatorias $X_1, X_2, ..., X_n$ posee las siguientes propiedades:*

1. $\Delta_a^b F \geq 0$; *para todo* $a = (a_1, \cdots, a_n)$, $b = (b_1, \cdots, b_n) \in \mathbb{R}^n$, *con* $a \leq b$,*donde:*

$$\Delta_a^b F := \sum_{\{\epsilon_1, \cdots, \epsilon_n\} \in \{0,1\}^n} (-1)^{\left(\sum_{i=1}^{n} \epsilon_i\right)} F\left(\epsilon_1 a_1 + (1-\epsilon_1) b_1, \cdots, \epsilon_n a_n + (1-\epsilon_n) b_n\right).$$

2. *F es continua a derecha en cada componente.*

3. *Para todo $a_1, \cdots, a_{i-1}, a_{i+1}, \cdots a_n \in \mathbb{R}$; $i = 1, \cdots, n$, se satisface:*

$$\lim_{x \searrow -\infty} F\left(a_1, \cdots, a_{i-1}, \underset{\substack{\downarrow \\ \textit{lugar } i}}{x}, a_{i+1}, \cdots a_n\right) = 0.$$

4. $$\lim_{(x_1, \cdots, x_n) \to (\infty, \cdots, \infty)} F(x_1, \cdots, x_n) = 1.$$

Demostración. Se hace la demostración para el caso $n = 2$. El caso general se trabaja de manera análoga.

1. Sean $a = (a_1, a_2)$, $b = (b_1, b_2)$; con $a_1 < b_1$ y $a_2 < b_2$, y:

$$\begin{aligned}
A &:= \{(x, y) \in \mathbb{R}^2 : x \leq b_1, y \leq b_2\} \\
B &:= \{(x, y) \in \mathbb{R}^2 : x \leq a_1, y \leq a_2\} \\
C &:= \{(x, y) \in \mathbb{R}^2 : x \leq a_1, y \leq b_2\} \\
D &:= \{(x, y) \in \mathbb{R}^2 : x \leq b_1, y \leq a_2).
\end{aligned}$$

Si $I := (A - C) - (D - B)$, entonces, es claro que:

$$\begin{aligned} 0 &\leq P_{\mathbf{X}}(I) = (P_{\mathbf{X}}(A) - P_{\mathbf{X}}(C)) - (P_{\mathbf{X}}(D) - P_{\mathbf{X}}(B)) \\ &= P_{\mathbf{X}}(A) + P_{\mathbf{X}}(B) - P_{\mathbf{X}}(C) - P_{\mathbf{X}}(D) \\ &= F(b_1, b_2) + F(a_1, a_2) - F(a_1, b_2) - F(b_1, a_2) \\ &= \Delta_a^b. \end{aligned}$$

2. Se debe verificar que:

$$\lim_{x \searrow x_0} F(x, y) = F(x_0, y) \tag{5.1}$$

y que

$$\lim_{y \searrow y_0} F(x, y) = F(x, y_0).$$

Se hace la demostración de 5.1. La otra queda como ejercicio para el lector.

$$\begin{aligned} \lim_{x \searrow x_0} F(x, y) &= \lim_{x \searrow x_0} P(X \leq x, Y \leq y) \\ &= P\left(\lim_{x \searrow x_0} (X \leq x, Y \leq y)\right) \\ &= P(X \leq x_0, Y \leq y) \\ &= F(x_0, y) \end{aligned}$$

3. Puesto que:

$$\begin{aligned} \lim_{x \searrow -\infty} [X \leq x, Y \leq y] &= \lim_{x \searrow -\infty} ([X \leq x] \cap [Y \leq y]) \\ &= \Phi \cap [Y \leq y] = \Phi \end{aligned}$$

se tiene:

$$\begin{aligned} \lim_{x \searrow -\infty} P[X \leq x, Y \leq y] &= P\left(\lim_{x \searrow -\infty} [X \leq x, Y \leq y]\right) \\ &= P(\Phi) = 0. \end{aligned}$$

Análogamente, se verifica que:

$$\lim_{y \searrow -\infty} P[X \leq x, Y \leq y] = 0$$

4. Es claro que:

$$\begin{aligned} \lim_{(x,y)\to(\infty,\infty)} F(x,y) &= \lim_{x\to\infty}\left(\lim_{y\to\infty} P(X\le x, Y\le y)\right) \\ &= \lim_{x\to\infty} P(X\le x) = 1. \end{aligned}$$

■

Definición 5.15 (Variables aleatorias conjuntamente continuas) *Sean $X_1, X_2, ..., X_n$, n variables aleatorias reales definidas sobre el mismo espacio de probabilidad. Se dice que las variables son conjuntamente continuas, si existe una función f definida para todo $(x_1, x_2, ..., x_n) \in \mathbb{R}^n$, no negativa e integrable, tal que:*

$$P((X_1, X_2, ..., X_n) \in C) = \int_C \cdots \int f(x_1, x_2, ..., x_n) dx_1 ... dx_n,$$

para todo conjunto Borel C de $\mathbb{R}^n$. La función f se llama función de densidad de probabilidad conjunta de las variables aleatorias $X_1, X_2, \cdots, X_n$.

Nota 5.16 *De la definición anterior se tiene, en particular, que:*

1. $\int_{\mathbb{R}^n} \cdots \int f(x_1, x_2, ..., x_n) dx_1 ... dx_n = 1.$

2.
$$P(X_1 \le x_1, \cdots, X_n \le x_n) = \int_{-\infty}^{x_n} \cdots \int_{-\infty}^{x_1} f(t_1, ..., t_n) dt_1 .. dt_n. \quad (5.2)$$

La observación anterior indica que, si se conoce la función de densidad de probabilidad conjunta f, de las variables aleatorias X_1, $X_2, ..., X_n$, entonces, se conoce la función de distribución conjunta F. ¿Es válido el recíproco? Esto es, ¿es posible determinar, a partir de la función de distribución conjunta F, la función de densidad conjunta f? La respuesta a esta pregunta la ofrece el teorema siguiente:

Teorema 5.17 *Sean X e Y variables aleatorias con función de distribución conjunta F y función de densidad conjunta f. Entonces:*

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2 F(x,y)}{\partial y \partial x}$$

en los puntos (x, y) donde $f(x, y)$ es continua.
Demostración. *Al aplicar el teorema fundamental del cálculo a 5.2 se obtiene:*

$$\frac{\partial F(x,y)}{\partial x} = \int_{-\infty}^{y} f(x,v)dv$$

luego:

$$\frac{\partial}{\partial y}\left(\frac{\partial F(x,y)}{\partial x}\right) = f(x,y)$$

Puesto que, $\frac{\partial^2 F(x,y)}{\partial x \partial y}$ *y* $\frac{\partial^2 F(x,y)}{\partial y \partial x}$ *existen y son continuas, entonces,*

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2 F(x,y)}{\partial y \partial x}$$

con lo cual queda demostrado el teorema. ■

Es más, si $X_1, \cdots, X_n$ son n variables aleatorias con función de distribución conjunta F, se tiene que la función $g(\cdot, \cdots, \cdot)$, definida sobre $\mathbb{R}^n$, por:

$$g(u_1, \cdots, u_n) := \begin{cases} \left.\frac{\partial^n F(x_1,x_2,\ldots,x_n)}{\partial x_1 \partial x_2 \ldots \partial x_n}\right|_{(x_1,x_2,\ldots,x_n)=(u_1,\cdots,u_n)}, & \text{donde las derivadas existen} \\ 0, & \text{donde las derivadas no existen.} \end{cases}$$

es una función de densidad de probabilidad conjunta de las variables aleatorias $X_1, \cdots, X_n$.

Supóngase ahora que X e Y son variables aleatorias con función de densidad de probabilidad conjunta f y sea g la función,de variable real, definida por:

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

Es claro que:

$$\begin{aligned} \int_{-\infty}^{x} g(u)du &= \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(u,y)dydu \\ &= \lim_{t\to\infty} \int_{-\infty}^{x}\int_{-\infty}^{t} f(u,y)dydu \\ &= \lim_{t\to\infty} F(x,t) = F_X(x). \end{aligned}$$

Además, como $\int_{-\infty}^{\infty} g(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dydx = 1$, entonces, g es una función de densidad de probabilidad de la variable aleatoria X, llamada *función de densidad marginal de X* y denotada por $f_X(x)$.

Análogamente,

$$f_Y(y) := \int_{-\infty}^{\infty} f(x,y)dx$$

es una función de densidad de la variable aleatoria Y.

En general, se tiene el resultado siguiente:

Teorema 5.18 *Si X_1, X_2, ..., X_n son n variables aleatorias reales, con fdp conjunta f, entonces,*

$$f_{X_j}(x) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1, ..., x_{j-1}, x, x_{j+1}, ..., x_n)dx_1...dx_{j-1}dx_{j+1}...dx_n$$

para $j = 1, 2, ..., n$; es una función de densidad de la variable aleatoria X_j.

Ejemplo 5.19 *Sean X e Y variables aleatorias con fdp conjunta dada por:*

$$f(x,y) = \begin{cases} 6xy^2 & si \quad 0 < x < 1, \quad 0 < y < 1 \\ 0 & en \quad otro\ caso. \end{cases}$$

1. *Calcular $P\left(\frac{1}{2} < X \leq \frac{3}{4},\ 0 < Y \leq \frac{1}{3}\right)$.*
2. *Determinar $P(\frac{1}{2} < X \leq \frac{3}{4})$.*

Solución:

1.

$$\begin{aligned} P\left(\frac{1}{2} < X \leq \frac{3}{4},\ 0 < Y \leq \frac{1}{3}\right) &= \int_0^{\frac{1}{3}} \int_{\frac{1}{2}}^{\frac{3}{4}} 6xy^2 dx dy \\ &= 1.1574 \times 10^{-2} \end{aligned}$$

2.

$$\begin{aligned} P\left(\frac{1}{2} < X \leq \frac{3}{4}\right) &= \int_{\frac{1}{2}}^{\frac{3}{4}} \int_0^1 6xy^2 dy dx \\ &= 0.3125. \quad \blacktriangle \end{aligned}$$

Ejemplo 5.20 *Sean X e Y variables aleatorias con fdp conjunta dada por:*

$$f(x,y) = \begin{cases} \frac{4x}{y^3} & si \quad 0 < x < 1;\ y > 1 \\ 0 & en \quad otro\ caso \end{cases}$$

Determinar:

1. $P\left(X < \frac{1}{2} \mid Y > 6\right)$.
2. *Las funciones de densidad marginales de* X *e* Y.

Solución:

1.

$$P\left(X < \frac{1}{2} \mid Y > 6\right) = \frac{P(X < \frac{1}{2}, Y > 6)}{P(Y > 6)}$$
$$= \frac{\int_0^{\frac{1}{2}} \int_6^{\infty} \frac{4x}{y^3} dy dx}{\int_6^{\infty} \int_0^1 \frac{4x}{y^3} dx dy} = \frac{6.9444 \times 10^{-3}}{2.7778 \times 10^{-2}} = 0.25$$

2.

$$f_X(x) = \int_{-\infty}^{\infty} f(x,y) dy$$
$$= \begin{cases} \int_1^{\infty} \frac{4x}{y^3} dy & \text{si} \quad 0 < x < 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$
$$= \begin{cases} 2x & \text{si} \quad 0 < x < 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

y

$$f_Y(y) = \int_{-\infty}^{\infty} f(x,y) dx$$
$$= \begin{cases} \int_0^1 \frac{4x}{y^3} dx & \text{si} \quad y > 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$
$$= \begin{cases} \frac{2}{y^3} & \text{si} \quad y > 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}. \quad \blacktriangle$$

Ejemplo 5.21 *Sean* X *e* Y *variables aleatorias con función de densidad de probabilidad conjunta dada por:*

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

Hallar:

1. *La función de distribución acumulativa conjunta de* X *e* Y.

2. Las funciones de densidad marginales de X e Y.

Solución:

1.

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,t)dudt$$
$$= \begin{cases} \frac{x^2y+xy^2}{2} & \text{si} \quad 0<x<1;\ 0<y<1 \\ \frac{x^2+x}{2} & \text{si} \quad 0<x<1;\ y\geq 1 \\ \frac{y+y^2}{2} & \text{si} \quad x\geq 1;\ 0<y<1 \\ 1 & \text{si} \quad x\geq 1;\ y\geq 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

2.

$$f_X(x) = \begin{cases} x+\frac{1}{2} & \text{si} \quad 0<x<1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

y

$$f_Y(y) = \begin{cases} y+\frac{1}{2} & \text{si} \quad 0<y<1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}. \qquad \blacktriangle$$

Ejemplo 5.22 *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:*

$$f(x,y) = \begin{cases} k & si \quad 0\leq x\leq 2;\ 0\leq y\leq 1 \ y \ 2y\leq x \\ 0 & en \quad otro\ caso \end{cases}$$

Hallar:

1. *El valor de k.*
2. *La función de distribución acumulativa conjunta de X e Y.*
3. *Las funciones de densidad marginales de X y de Y.*
4. *$P(X\leq 3Y)$.*

Solución:

1.

$$1 = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dxdy = \int_0^2\int_0^{\frac{x}{2}} kdydx = k$$

2.

$$F_{XY}(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f(u,t)dudt$$
$$= \begin{cases} 0 & \text{si} \quad x<0,\ y<0 \\ xy-y^2 & \text{si} \quad 0\le x\le 2,\ 0\le y\le \frac{x}{2},\ 0\le y\le 1 \\ \frac{x^2}{4} & \text{si} \quad 0\le x\le 2,\ \frac{x}{2}\le y\le 1 \\ 2y-y^2 & \text{si} \quad x>2,\ 0\le y\le 1 \\ \frac{x^2}{4} & \text{si} \quad 0\le x\le 2,\ y\ge 1 \\ 1 & \text{si} \quad x\ge 2,\ y\ge 1 \end{cases}$$

3.

$$f_X(x) = \begin{cases} \int_0^{\frac{x}{2}} dy & \text{si} \quad 0\le x\le 2 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$
$$= \begin{cases} \frac{x}{2} & \text{si} \quad 0\le x\le 2 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

y

$$f_Y(y) = \begin{cases} \int_{2y}^{2} dx & \text{si} \quad 0\le y\le 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$
$$= \begin{cases} 2-2y & \text{si} \quad 0\le y\le 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

4.

$$P(X\le 3Y) = \int_0^2\int_{\frac{x}{3}}^{\frac{x}{2}} dydx = \frac{1}{3}. \quad \blacktriangle$$

5.2. Variables aleatorias independientes

Definición 5.23 (variables aleatorias independientes) *Sean X e Y dos variables aleatorias reales definidas sobre el mismo espacio de probabilidad, se dice que ellas son independientes, si se satisface que:*

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

para todo A y B conjuntos Borel de $\mathbb{R}$.

Nota 5.24 (vectores aleatorios independientes) *La definición de variables aleatorias independientes puede generalizarse a vectores aleatorios, de la siguiente manera: dos vectores aleatorios n -dimensionales $\mathbf{X}$ e $\mathbf{Y}$ definidos sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$, se dicen independientes, si se satisface que:*

$$P(\mathbf{X} \in A,\ \mathbf{Y} \in B) = P(\mathbf{X} \in A)\, P(\mathbf{Y} \in B),$$

para todo A y B conjuntos Borel de $\mathbb{R}^n$.

Supóngase que X e Y son variables aleatorias independientes, entonces, de la definición anterior, se sigue que:

$$F(x,y) = P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y); \text{ para todo } x, y \in \mathbb{R},$$

esto es,

$$F(x,y) = F_X(x)F_Y(y); \text{ para todo } x, y \in \mathbb{R}. \tag{5.3}$$

Recíprocamente, si la condición 5.3 se satisface, entonces, las variables son independientes.

Supóngase ahora que X e Y son variables aleatorias discretas independientes, entonces:

$$P(X = x, Y = y) = P(X = x)P(Y = y); \tag{5.4}$$

para todo x en el rango de X y todo y en el rango de Y.

Recíprocamente, si la condición 5.4 se satisface, entonces, las variables aleatorias son independientes.

Si X e Y son variables aleatorias independientes, con función de densidad conjunta $f(x,y)$, entonces:

$$P(x < X \leq x+dx,\ y < Y \leq y+dy) = P(x < X \leq x+dx)P(y < Y \leq y+dy).$$

esto es:

$$f(x,y) = f_X(x)f_Y(y); \text{ para todo } x, y \in \mathbb{R} \tag{5.5}$$

Recíprocamente, si la condición 5.5 se satisface, entonces, las variables son independientes. Más en general, se tiene que las variables aleatorias X e Y son independientes, si y sólo si, su función de densidad de probabilidad conjunta $f(x,y)$ puede expresarse como el producto de dos funciones: una en términos solamente de la primera variable y otra en términos sólo de la segunda variable. En otras palabras, X e Y son independientes si existen $h(x)$ y $g(y)$, funciones reales, tales que:

$$f(x,y) = h(x)g(y); \quad -\infty < x < \infty,\ -\infty < y < \infty.$$

Supóngase ahora que X e Y son variables aleatorias independientes cuyas funciones generadoras de momentos existen. Sea $Z := X+Y$, entonces, se tiene que:

$$\begin{aligned} m_Z(t) &= E(\exp tZ) \\ &= E(\exp(tX + tY)) \\ &= E(\exp(tX)\exp(tY)) \\ &= E(\exp(tX))E(\exp(tY)) \\ &= m_X(t)m_Y(t), \end{aligned}$$

esto es, la función generadora de momentos de Z existe y es igual al producto de las funciones generadoras de momentos de X e Y.

A continuación se presentan ejemplos en los que se aplican los resultados anteriores.

Ejemplo 5.25 *En una urna hay tres bolas rojas y seis azules. Se extrae una muestra aleatoria de tamaño dos (con reemplazo). Sean X e Y las variables aleatorias definidas por:*

$$X := \begin{cases} 1 & \textit{si la primera bola extraída es roja} \\ 0 & \textit{si la primera bola extraída es azul} \end{cases}$$

$$Y := \begin{cases} 1 & \textit{si la segunda bola extraída es roja} \\ 0 & \textit{si la segunda bola extraída es azul} \end{cases}$$

¿Son X e Y independientes? Explicar.

Solución: La distribución conjunta de X e Y está dada por:

$X\backslash Y$	0	1
0	$\frac{4}{9}$	$\frac{2}{9}$
1	$\frac{2}{9}$	$\frac{1}{9}$

.

Es fácil verificar que:

$$P(X = x, Y = y) = P(X = x)P(Y = y); \text{ para todo } x, y \in \{0, 1\},$$

esto es, X e Y son variables aleatorias independientes, lo cual era de esperarse, pues la composición de la urna, tanto en la primera como en la segunda extracción, es la misma. ▲

Ejemplo 5.26 *Se va a resolver el mismo ejemplo anterior, pero suponiendo que la extracción se hace sin sustitución. En este caso se tiene que la distribución conjunta de X e Y está dada por:*

$X\backslash Y$	0	1
0	$\frac{5}{12}$	$\frac{3}{12}$
1	$\frac{3}{12}$	$\frac{1}{12}$

Como:

$$P(X=0,Y=0)=\frac{5}{12}\neq\frac{2}{3}\times\frac{2}{3}=P(X=0)P(Y=0),$$

entonces, X e Y no son independientes. ▲

Ejemplo 5.27 *Sean X e Y variables aleatorias independientes con funciones de densidad dadas por:*

$$f(x)=\frac{3x^2}{8}\mathcal{X}_{(0,2)}(x)$$
$$g(y)=y^{-2}\mathcal{X}_{(1,\infty)}(y)$$

Calcular $P(XY>1)$.

Solución: Como X e Y son independientes, su función de densidad de probabilidad conjunta está dada por:

$$h(x,y)=\begin{cases}\frac{3x^2}{8y^2} & \text{si} \quad 0<x<2;\ y>1 \\ 0 & \text{en} \quad \text{otro caso}\end{cases}.$$

Por lo tanto,

$$P(XY>1)=\int_1^{\infty}\int_{\frac{1}{y}}^{2}\frac{3x^2}{8y^2}dxdy=\frac{31}{32}. \quad ▲$$

Nota 5.28 *Sean X e Y variables aleatorias discretas independientes. Es claro que:*

$$\begin{aligned}P(X+Y=z)&=\sum_{x=0}^{z}P(X=x,Y=z-x)\\&=\sum_{x=0}^{z}P(X=x)P(Y=z-x).\end{aligned}$$

Ejemplo 5.29 *Supóngase que X e Y son variables aleatorias independientes con $X \stackrel{d}{=} \mathcal{P}(\lambda)$ y $Y \stackrel{d}{=} \mathcal{P}(\mu)$. Hallar la distribución de $Z = X + Y$.*

Solución: Las variables aleatorias X e Y toman los valores $0, 1, \cdots$. Por lo tanto, la variable aleatoria Z toma también los valores $0, 1, \cdots$. Sea $z \in \{0, 1, \cdots\}$, entonces:

$$\begin{aligned}
P\,(Z = z) &= \sum_{x=0}^{z} P(X = x)P(Y = z - x) \\
&= \sum_{x=0}^{z} e^{-\lambda}\frac{\lambda^x}{x!}e^{-\mu}\frac{\mu^{z-x}}{(z-x)!} \\
&= \frac{1}{z!}e^{-(\lambda+\mu)}\sum_{x=0}^{z}\binom{z}{x}\lambda^x\mu^{z-x} \\
&= \frac{1}{z!}e^{-(\lambda+\mu)}(\lambda+\mu)^z.
\end{aligned}$$

Esto es, $Z \stackrel{d}{=} \mathcal{P}(\lambda + \mu)$. ▲

Nota 5.30 *Supóngase que X e Y son variables aleatorias independientes con función de densidad de probabilidad conjunta $f(x, y)$ y funciones de densidad marginales f_X y f_Y, respectivamente. La función de distribución de la variable aleatoria $Z = X + Y$, puede obtenerse como sigue:*

$$\begin{aligned}
F_Z(z) &= P(Z \leq z) \\
&= \iint_{\{x+y\leq\, z\}} f(x, y)dxdy \\
&= \iint_{\{x+y\leq\, z\}} f_X(x)f_Y(y)dxdy
\end{aligned}$$

$$\begin{aligned}
&= \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_X(x)f_Y(y)dxdy \qquad (5.6)\\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{z-y} f_X(x)dx f_Y(y)dy \\
&= \int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy
\end{aligned}$$

Al derivar, la ecuación 5.6, se obtiene que la función de densidad de la *variable aleatoria Z está dada por:*

$$\begin{aligned} f_Z(z) &= \frac{d}{dz}\int_{-\infty}^{\infty} F_X(z-y)f_Y(y)dy \\ &= \int_{-\infty}^{\infty} \frac{d}{dz}F_X(z-y)f_Y(y)dy \\ &= \int_{-\infty}^{\infty} f_X(z-y)f_Y(y)dy. \end{aligned} \tag{5.7}$$

La función de densidad de la variable aleatoria Z se llama la convolución de las funciones de densidad f_X y f_Y y se denota por $f_X * f_Y$.

Ejemplo 5.31 *Sean X y Y variables aleatorias independientes e igualmente distribuidas con distribución exponencial de parámetro $\lambda > 0$. La función de densidad de $Z = X + Y$ está dada por:*

$$\begin{aligned} f_Z(z) &= (f_X * f_Y)(z) \\ &= \int_{-\infty}^{\infty} \lambda e^{-\lambda(z-u)}\mathcal{X}_{(0,\infty)}(z-u)\lambda e^{-\lambda u}\mathcal{X}_{(0,\infty)}(u)du \\ &= \int_{-\infty}^{\infty} \lambda^2 e^{-\lambda z}\mathcal{X}_{(0,\infty)}(z-u)\mathcal{X}_{(0,\infty)}(u)du. \end{aligned}$$

Puesto que:

$$\mathcal{X}_{(0,\infty)}(z-u)\mathcal{X}_{(0,\infty)}(u) = \begin{cases} 1 & si \quad z > u > 0 \\ 0 & en \quad otro\ caso \end{cases}$$

entonces

$$\begin{aligned} (f_X * f_Y)(z) &= \lambda^2 z e^{-\lambda z}\mathcal{X}_{(0,\infty)}(z) \\ &= \frac{\lambda}{\Gamma(2)}(\lambda z)e^{-\lambda z}\mathcal{X}_{(0,\infty)}(z) \end{aligned}$$

esto es, $Z \overset{d}{=} \Gamma(2,\lambda)$. ▲

Ejemplo 5.32 *Sean X y Y variables aleatorias independientes tales que $X \overset{d}{=} \mathcal{U}\,[0,2]$ y $Y \overset{d}{=} Exp(1)$. Calcular $P(X+Y \geq 2.5)$.*

Solución: La función de densidad de probabilidad conjunta de X e Y está dada por:

$$f(x,y) = \begin{cases} \frac{1}{2}e^{-y} & \text{si} \quad 0 \leq x \leq 2;\ y > 0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Por lo tanto,

$$\begin{aligned}
P(X+Y \geq 2.5) &= \iint\limits_{\{(x,y):x+y\geq 2.5\}} f(x,y)dxdy \\
&= \int_0^2 \int_{\frac{5}{2}-x}^{\infty} \frac{1}{2} e^{-y} dydx \\
&= \int_0^2 \frac{1}{2} e^{-\left(\frac{5}{2}-x\right)} dx \\
&= 0.26222. \quad \blacktriangle
\end{aligned}$$

El concepto de independencia puede generalizarse a n variables aleatorias como sigue:

Definición 5.33 (independencia de n variables aleatorias) *n variables aleatorias reales X_1, $X_2, \cdots, X_n$ definidas sobre el mismo espacio de probabilidad se llaman independientes, si y sólo si, se satisface que:*

$$P(X_1 \in A_1, ..., X_n \in A_n) = P(X_1 \in A_1) \cdots P(X_n \in A_n),$$

para todos los conjuntos Borel $A_1, A_2, ..., A_n$ de $\mathbb{R}$.

Nota 5.34 (independencia de n vectores aleatorios) *La definición anterior puede generalizarse a vectores aleatorios de la manera siguiente: n vectores aleatorios k-dimensionales $\mathbf{X}_1$, $\mathbf{X}_2, \cdots, \mathbf{X}_n$, definidos sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$, se dicen independientes, si y sólo si, se satisface que:*

$$P(\mathbf{X}_1 \in A_1, ..., \mathbf{X}_n \in A_n) = P(\mathbf{X}_1 \in A_1)...P(\mathbf{X}_n \in A_n),$$

para todos los conjuntos Borel $A_1, A_2, ..., A_n$ de $\mathbb{R}^n$.

Nota 5.35 *Si $X_1, \cdots, X_n$ son n variables aleatorias independientes cuyas funciones generadoras de momentos existen, entonces, la función generadora de momentos de $Z := X_1 + \cdots + X_n$ también existe y es igual al producto de las funciones generadoras de momentos de las X_i. Un resultado similar se tiene para las funciones características.*

Ejemplo 5.36 *Sean $X_1, X_2, \cdots, X_n$, n variables aleatorias independientes e igualmente distribuidas con distribución exponencial de parámetro $\lambda > 0$, entonces, $Z := X_1 + \cdots + X_n$ tiene distribución gamma de parámetros n y*

λ. En efecto, la función generadora de momentos de $Z := X_1 + \cdots + X_n$ está dada por:

$$
\begin{aligned}
m_Z(t) &= E(\exp tZ) \\
&= \prod_{k=1}^{n} E(\exp tX_k) \\
&= \prod_{k=1}^{n} \left(\frac{\lambda}{\lambda - t}\right); \quad si\ t < \lambda \\
&= \left(\frac{\lambda}{\lambda - t}\right)^n; \quad si\ t < \lambda
\end{aligned}
$$

la cual corresponde a la función generadora de momentos de una variable aleatoria con distribución gamma de parámetros n y λ. ▲

Ejemplo 5.37 *Sean X_1, $X_2, \cdots, X_n$, n variables aleatorias independientes e igualmente distribuidas con distribución normal estándar. Sea*

$$Z := X_1^2 + \cdots + X_n^2.$$

Se sabe que si una variable aleatoria tiene distribución normal estándar, entonces, su cuadrado tiene distribución ji-cuadrado con 1 grado de libertad. Por lo tanto, la función generadora de momentos de Z está dada por:

$$
\begin{aligned}
m_Z(t) &= \prod_{k=1}^{n} m_{X_k}(t) \\
&= \prod_{k=1}^{n} \left(\frac{1}{1 - 2t}\right)^{\frac{1}{2}}; \quad si\ t < \frac{1}{2} \\
&= \left(\frac{1}{1 - 2t}\right)^{\frac{n}{2}}; \quad si\ t < \frac{1}{2}
\end{aligned}
$$

esto es, $Z \stackrel{d}{=} \mathcal{X}^2_{(n)}$. ▲

Ejemplo 5.38 *Sean X_1, $X_2, ..., X_n$, n variables aleatorias independientes y supóngase que $X_k \stackrel{d}{=} \mathcal{N}(\mu_k, \sigma_k^2)$; para $k = 1, \cdots, n$. Sean $\alpha_1, \cdots, \alpha_n$, n constantes reales. Entonces, la función generadora de momentos de la va-*

riable aleatoria $Z := \alpha_1 X_1 + \cdots + \alpha_n X_n$ *está dada por:*

$$\begin{aligned}
m_Z(t) &= E\left(\exp t\left(\sum_{k=1}^{n}\alpha_k X_k\right)\right) \\
&= E\left(\prod_{k=1}^{n}\exp(t\alpha_k X_k)\right) \\
&= \prod_{k=1}^{n}E(\exp(t\alpha_k X_k)) \\
&= \prod_{k=1}^{n}m_{X_k}(t\alpha_k) \\
&= \prod_{k=1}^{n}\exp\left(t\alpha_k\mu_k + \frac{(t\alpha_k)^2}{2}\sigma_k^2\right) \\
&= \exp\left(t\sum_{k=1}^{n}(\alpha_k\mu_k) + \frac{t^2}{2}\sum_{k=1}^{n}(\alpha_k\sigma_k)^2\right)
\end{aligned}$$

esto es, $Z \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$; *donde* $\mu := \sum_{k=1}^{n}(\alpha_k\mu_k)$ *y* $\sigma^2 := \sum_{k=1}^{n}(\alpha_k\sigma_k)^2$. ▲

Ejemplo 5.39 *Sean* $X_1, X_2, \cdots, X_n$, *n variables aleatorias independientes e igualmente distribuidas con distribución normal de parámetros* μ *y* σ^2. *La variable aleatoria* $\overline{X}$ *definida por:*

$$\overline{X} := \frac{1}{n}\sum_{k=1}^{n}X_i$$

tiene distribución normal de parámetros μ *y* $\frac{\sigma^2}{n}$.

En consecuencia:

$$\frac{\sum_{k=1}^{n}X_i - n\mu}{\sigma\sqrt{n}} \stackrel{d}{=} \mathcal{N}(0,1). \quad ▲$$

Ejemplo 5.40 (distribución del máximo y el mínimo)
Sean $X_1, \cdots, X_n$, *n variables aleatorias reales definidas sobre* $(\Omega, \Im, P)$ *y considérense las variables aleatorias* Y *y* Z *definidas como sigue:*

$$\begin{aligned}
Y := \quad \text{máx}(X_1, \cdots, X_n) : \quad \Omega &\longrightarrow \mathbb{R} \\
\omega &\longmapsto \text{máx}(X_1(\omega), \cdots, X_n(\omega))
\end{aligned}$$

y

$$\begin{array}{rlcl} Z := \min(X_1, \cdots, X_n): & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \longmapsto & \min(X_1(\omega), \cdots, X_n(\omega)) \end{array}.$$

Es claro que:

$$F_Y(y) := P(Y \leq y) = P(X_1 \leq y, \cdots, X_n \leq y)$$

y que:

$$F_Z(z) := P(Z \leq z) = 1 - P(X_1 > z, \cdots X_n > z).$$

Por lo tanto, si las variables aleatorias $X_1, \cdots, X_n$ *son independientes, entonces:*

$$F_Y(y) = \prod_{k=1}^{n} P(X_k \leq y) = \prod_{k=1}^{n} F_{X_k}(y)$$

y

$$\begin{aligned} F_Z(z) &= 1 - \prod_{k=1}^{n} [1 - P(X_k \leq z)] \\ &= 1 - \prod_{k=1}^{n} [1 - F_{X_k}(z)] \end{aligned}$$

siendo $F_{X_k}(\cdot)$ *la función de distribución de la variable aleatoria* X_k*, con* $k = 1, \cdots, n$. ▲

Ejemplo 5.41 *Sean X y Y variables aleatorias independientes e igualmente distribuidas con distribución exponencial de parámetro a. Determinar la función de densidad de la variable aleatoria* $Z := \max\{X, Y^3\}$

Solución: La función de densidad de la variable aleatoria $U = Y^3$ está dada por:

$$f_U(u) = \begin{cases} \frac{a}{3\sqrt[3]{u^2}} e^{-a\sqrt[3]{u}} & \text{si } u > 0 \\ 0 & \text{en otro caso.} \end{cases}$$

Como las variables aleatorias X y Y^3 son independientes, entonces, la variable aleatoria $Z = \max\{X, Y^3\}$ tiene como función de distribución acumulativa a:

$$F_Z(z) = F_X(z) F_{Y^3}(z).$$

Por lo tanto, la función de densidad de Z está dada por:

$$f_Z(z) = f_X(z) F_{Y^3}(z) + F_X(z) f_{Y^3}(z),$$

esto es:

$$f_Z(z) = \begin{cases} ae^{-az} + \frac{a}{3\sqrt[3]{z^2}}e^{-a\sqrt[3]{z}} - a(1 + \frac{1}{3\sqrt[3]{z^2}})e^{-a\sqrt[3]{z}-az} & \text{si} \quad z > 0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Nota 5.42 *Si $X_1, X_2, ..., X_n$ son n variables aleatorias independientes, entonces, X_1, $X_2, ..., X_k$;con $k \leq n$, también son independientes. En efecto, si $A_1, A_2, ..., A_k$ son conjuntos Borel de $\mathbb{R}$, entonces:*

$$\begin{aligned} P(X_1 \in A_1, ..., X_k \in A_k) &= P(X_1 \in A_1, \cdots, X_k \in A_k, X_{k+1} \in \mathbb{R}, \\ &\qquad \cdots, X_n \in \mathbb{R}) \\ &= P(X_1 \in A_1) \cdots P(X_k \in A_k) P(X_{k+1} \in \mathbb{R}) \\ &\qquad \cdots P(X_n \in \mathbb{R}) \\ &= P(X_1 \in A_1) \cdots P(X_k \in A_k). \quad \blacktriangle \end{aligned}$$

Supóngase que X_1, X_2 y X_3 son variables aleatorias discretas independientes y sean $Y_1 := X_1 + X_2$ y $Y_2 := X_3^2$. Entonces:

$$\begin{aligned} P(Y_1 = z, Y_2 = w) &= \sum_x P(X_1 = x, X_2 = z - x, X_3^2 = w) \\ &= \sum_x P(X_1 = x)P(X_2 = z - x)P(X_3 = \pm\sqrt{w}) \\ &= P(X_3 = \pm\sqrt{w})\sum_x P(X_1 = x)P(X_2 = z - x) \\ &= P(X_1 + X_2 = z)P(X_3^2 = w). \end{aligned}$$

Esto es, $Y_1 := X_1 + X_2$ y $Y_2 := X_3^2$ son variables aleatorias independientes. ¿Es este resultado válido en general? La respuesta a esta pregunta la da el teorema siguiente, cuya demostración se omite, pues ella requiere de resultados de teoría de la medida.

Teorema 5.43 *Sean $X_1, \cdots, X_n$, n variables aleatorias independientes. Sean Y una variable aleatoria definida en términos de $X_1, \cdots, X_k$ y Z una variable aleatoria definida en términos de $X_{k+1}, \cdots, X_n$; donde $1 \leq k < n$. Entonces, Y y Z son independientes.*

Ejemplo 5.44 *Sean $X_1, \cdots, X_5$ variables aleatorias independientes. Entonces, $Y = X_1X_2 + X_3$ y $Z = e^{X_5} \sin X_4$ son variables aleatorias independientes.* $\blacktriangle$

5.3. Covarianza y coeficiente de correlación

A continuación se define el valor esperado de una función de una variable aleatoria n-dimensional.

Definición 5.45 ***(valor esperado de una función de un vector aleatorio)***
Sean $(X_1, X_2, \cdots, X_n)$ un vector aleatorio n-dimensional y $g(\cdot, \cdots, \cdot)$ una función definida sobre $\mathbb{R}^n$ y de valor real. El valor esperado de la función $g(X_1, \cdots, X_n)$, denotado por $E(g(X_1, \cdots, X_n))$, se define como:

$$E(g(X_1, \cdots, X_n)) :=$$
$$\begin{cases} \sum_{x_1} \cdots \sum_{x_n} g\left(x_1, \cdots, x_n\right) P\left(X_1 = x_1, \cdots, X_n = x_n\right), \\ \hfill \text{en el caso discreto.} \\ \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} g(x_1, \cdots, x_n) f(x_1, \cdots, x_n) dx_1 \cdots dx_n, \\ \hfill \text{en el caso continuo.} \end{cases}$$

siempre y cuando, la suma múltilple, en el caso discreto, o la integral múltiple, en el caso continuo, converjan absolutamente.

Ejemplo 5.46 *Supóngase que se lanza un dado corriente dos veces consecutivas. Sean $X :=$ "máximo valor obtenido" e $Y :=$ "suma de los valores obtenidos". En este caso, $E\left(XY\right) = \frac{1232}{36}$* ▲

Ejemplo 5.47 *Sea (X, Y, Z) un vector aleatorio tridimensional, con función de densidad de probabilidad conjunta dada por:*

$$f(x, y, z) = \begin{cases} 8xyz & si \quad 0 < x < 1;\ 0 < y < 1 \ y\, 0 < z < 1 \\ 0 & \quad en\ otro\ caso \end{cases}.$$

Se tiene que $E(5X - 2Y + Z) = \frac{8}{3}$ y $E\left(XY\right) = \frac{4}{9}$. ▲

Teorema 5.48 *Si X e Y son variables aleatorias cuyos valores esperados existen, entonces, el valor esperado de $X + Y$ también existe y es igual a la suma de los valores esperados.*

Demostración. *Se hace la demostración para el caso continuo. El caso discreto se trabaja de manera análoga.*

Supóngase que f es la función de densidad de probabilidad conjunta de X e Y. Entonces:

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x+y|\, f(x,y)dxdy &\leq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (|x|+|y|)\, f(x,y)dxdy \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x|\, f(x,y)dxdy + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |y|\, f(x,y)dxdy \\
&= \int_{-\infty}^{\infty} |x| \left(\int_{-\infty}^{\infty} f(x,y)dy\right) dx + \int_{-\infty}^{\infty} |y| \left(\int_{-\infty}^{\infty} f(x,y)dx\right) dy \\
&= \int_{-\infty}^{\infty} |x|\, f_X(x)dx + \int_{-\infty}^{\infty} |y|\, f_Y(y)dy < \infty.
\end{aligned}$$

por lo tanto, $E(X+Y)$ existe.

Ahora bien,

$$\begin{aligned}
E(X+Y) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)f(x,y)dxdy \\
&= \int_{-\infty}^{\infty} x \left(\int_{-\infty}^{\infty} f(x,y)dy\right) dx + \int_{-\infty}^{\infty} y \left(\int_{-\infty}^{\infty} f(x,y)dx\right) dy \\
&= \int_{-\infty}^{\infty} x f_X(x)dx + \int_{-\infty}^{\infty} y f_Y(y)dy \\
&= EX + EY.
\end{aligned}$$

■

Nota 5.49 *Si X es una variable aleatoria de tipo discreto y Y es una variable aleatoria de tipo continuo, el resultado sigue siendo válido.*

En general se tiene que:

Teorema 5.50 *Si $X_1, X_2, \cdots, X_n$ son n variables aleatorias cuyos valores esperados existen, entonces, el valor esperado de la suma de las variables aleatorias también existe y es igual a la suma de los valores esperados.*
Demostración. *Como ejercicio.* ■

Ejemplo 5.51 *Se tiene una urna con N bolas de las cuales R son de color rojo y las restantes $(N-R)$ son de color blanco. Se extrae una muestra de tamaño n, sin sustitución. Sea X el número de bolas rojas extraídas en la muestra. De la teoría, vista en el capítulo 3, se sabe que la variable aleatoria X tiene una distribución hipergeométrica de parámetros n, R y N; y que por* *lo tanto* $EX = \frac{nR}{N}$. *A continuación, se va a deducir este mismo resultado, pero expresando a* X *como una suma de variables aleatorias y aplicando el teorema anterior.*

Sean X_i; *con* $i = 1, \cdots n$, *las variables aleatorias definidas como sigue:*

$$X_i := \begin{cases} 1 & si \quad \text{la i-ésima bola extraída es roja} \\ 0 & si \quad \text{la i-ésima bola extraída es blanca} \end{cases}$$

es claro que:

$$X = \sum_{i=1}^{n} X_i$$

por lo tanto,

$$EX = \sum_{i=1}^{n} EX_i = \sum_{i=1}^{n} \frac{R}{N} = \frac{nR}{N}. \quad \blacktriangle$$

Teorema 5.52 *Si* X *e* Y *son variables aleatorias independientes cuyos valores esperados existen, entonces, el valor esperado de* XY *también existe y es igual al producto de los valores esperados.*

Demostración. *Se hace la demostración para el caso continuo. El caso discreto se trabaja de manera análoga.*

Supóngase que f *es la función de densidad de probabilidad conjunta de* X *e* Y.*Entonces:*

$$\begin{aligned} \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |xy|\, f(x,y)dxdy &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} |x|\,|y|\, f_X(x)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty} |x| \left(\int_{-\infty}^{\infty} |y|\, f_Y(y)dy\right) f_X(x)dx \\ &= \left(\int_{-\infty}^{\infty} |x|\, f_X(x)dx\right)\left(\int_{-\infty}^{\infty} |y|\, f_Y(y)dy\right) < \infty \end{aligned}$$

esto es $E(XY)$ *existe.*

Suprimiendo las barras de valor absoluto, en la prueba anterior, se obtiene que:

$$EXY = (EX)(EY).$$

$\blacksquare$

Dos variables aleatorias pueden ser independientes o estar muy relacionadas entre sí. Es posible, por ejemplo, que la variable aleatoria Y tienda a aumentar cuando lo hace X o que la variable aleatoria Y tienda a aumentar cuando X decrece. Los números que se definen a continuación, conocidos como covarianza y coeficiente de correlación, permiten determinar si existe una relación de tipo lineal entre las variables consideradas.

Definición 5.53 (covarianza) *Sean X e Y variables aleatorias definidas sobre el mismo espacio de probabilidad y tales que $EX^2 < \infty$ y $EY^2 < \infty$. La covarianza entre X e Y se define como:*

$$Cov(X,Y) := E((X - EX)(Y - EY)).$$

Nota 5.54 *Sean X e Y variables aleatorias definidas sobre el mismo espacio de probabilidad y tales que $EX^2 < \infty$ y $EY^2 < \infty$. Como $|X| \leq 1 + X^2$, entonces, se obtiene la existencia de EX. Por otra parte, como $|XY| \leq X^2 + Y^2$, entonces, se sigue la existencia del valor esperado de $(X - EY)(Y - EY)$.*

Teorema 5.55 *Sean X e Y variables aleatorias definidas sobre el mismo espacio de probabilidad y tales que $EX^2 < \infty$ y $EY^2 < \infty$. Entonces:*

1. $Cov(X,Y) = EXY - EXEY$.
2. $Cov(X,Y) = Cov(Y,X)$.
3. $VarX = Cov(X,X)$.
4. $Cov(aX + b, Y) = aCov(X,Y)$; *para todo* $a, b \in \mathbb{R}$.

Demostración. *Sólo se hace la demostración de 4. Las demás quedan como ejercicio.*

$$\begin{aligned} Cov(aX + b, Y) &= E(((aX + b) - E(aX + b))(Y - EY)) \\ &= E((aX + b - aEX - b)(Y - EY)) \\ &= aCov(X,Y). \end{aligned}$$

■

Nota 5.56 *De la primera propiedad de la covarianza se deduce que si X e Y son independientes, entonces, $Cov(X,Y) = 0$. El recíproco no es válido en general, como lo demuestra el ejemplo siguiente.*

Ejemplo 5.57 *Sea $\Omega = \{1,2,3,4\}$, $\Im = \wp(\Omega)$ y P dada por: $P(1) = P(2) := \frac{2}{5}$, $P(3) = P(4) := \frac{1}{10}$. Las variables aleatorias X e Y definidas por: $X(1) = Y(2) := 1$, $X(2) = Y(1) := -1$, $X(3) = Y(3) := 2$ y $X(4) = Y(4) := -2$ no son independientes, pero $Cov(X,Y) = 0$.* ▲

Teorema 5.58 (desigualdad de Cauchy-Schwarz) *Sean X e Y dos variables con $EX^2 < \infty$ y $EY^2 < \infty$, entonces, $|EXY|^2 \leq EX^2EY^2$. Se tiene la igualdad, si y sólo, si existen constantes reales a y b no simultáneamente cero tales que $P(aX + bY = 0) = 1$.*

Demostración. *Sea $\alpha = EY^2$ y $\beta = -EXY$. Es claro que $\alpha \geq 0$. Como el resultado se tiene trivialmente para $\alpha = 0$, se considera únicamente el caso $\alpha > 0$. Se tiene que:*

$$\begin{aligned} 0 \leq E((\alpha X + \beta Y)^2) &= E(\alpha^2 X^2 + 2\alpha\beta XY + \beta^2 Y^2) \\ &= \alpha(EX^2EY^2 - E(XY)E(XY)) \end{aligned}$$

como $\alpha > 0$, entonces, se obtiene el resultado.

Si $(EXY)^2 = EX^2EY^2$, entonces, $E(\alpha X + \beta Y)^2 = 0$. Por lo tanto, con probabilidad 1, se tiene que $(\alpha X + \beta Y) = 0$. Si $\alpha > 0$, se puede tomar $a = \alpha$ y $b = \beta$. Si $\alpha = 0$, entonces, se puede tomar $a = 0$ y $b = 1$.

Recíprocamente, si existen números reales a y b no simultáneamente cero tales que $aX + bY = 0$, con probabilidad 1, entonces, $aX = -bY$ con probabilidad 1 y en tal caso es fácil verificar que $|EXY|^2 = EX^2EY^2$. ■

Nota 5.59 *Tomando $|X|$ y $|Y|$ en lugar de X y de Y, en el teorema anterior, se llega a:*

$$E|XY| \leq \sqrt{EX^2}\sqrt{EY^2}$$

Si se aplica el resultado anterior a las variables $X - EX$ y $Y - EY$ se obtiene $|Cov(X,Y)| \leq \sqrt{VarX}\sqrt{VarY}$.

Teorema 5.60 *Si X e Y son variables aleatorias reales con varianzas finitas, entonces, $Var(X + Y) < \infty$ y se tiene que:*

$$Var(X + Y) = VarX + VarY + 2Cov(X,Y). \tag{5.8}$$

Demostración. *Para ver que $Var(X + Y) < \infty$ basta verificar que*

$$E(X + Y)^2 < \infty.$$

Ahora bien:

$$\begin{aligned} E(X + Y)^2 &= EX^2 + 2EXY + EY^2 \\ &\leq EX^2 + 2E|XY| + EY^2 \\ &\leq 2(EX^2 + EY^2) < \infty \end{aligned}$$

aplicando las propiedades del valor esperado se obtiene 5.8. ■

En general se tiene que:

Teorema 5.61 *Si $X_1, X_2, \cdots, X_n$ son n variables aleatorias con varianzas finitas, entonces, $Var\left(\sum_{i=1}^{n} X_i\right) < \infty$ y se satisface:*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + \sum_{i \neq j}\sum Cov(X_i, X_j).$$

Demostración. *Se deja como ejercicio.* ■

Nota 5.62 *Puesto que cada par de índices i, j; con $i \neq j$ aparece dos veces en la suma anterior, se tiene que ésta es equivalente a:*

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum_{i<j}\sum Cov(X_i, X_j).$$

Si $X_1, X_2, \cdots, X_n$ son n variables aleatorias independientes con varianzas finitas, entonces:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

Ejemplo 5.63 *Supóngase que se muestran n fotos de diferentes bebés, cada una correspondiente a una persona de la farándula, actualmente conocida por el público en genera, y se le pregunta a un observador cuál foto corresponde a cada persona. Sea X la variable aleatoria que representa el número de respuestas correctas. Calcular EX y $Var(X)$.*

Solución: Sea

$$X_i := \begin{cases} 1 & \text{si} \quad \text{el observador identifica correctamente la } i\text{-ésima foto.} \\ 0 & \text{en} \quad \text{caso contrario.} \end{cases}$$

Se tiene que:

$$X = \sum_{i=1}^{n} X_i.$$

Por lo tanto:

$$EX = \sum_{i=1}^{n} E(X_i) = \sum_{i=1}^{n} P(X_i = 1) = n\frac{1}{n} = 1.$$

Por otra parte, se tiene que:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) + 2\sum\sum_{i<j} Cov(X_i, X_j),$$

como

$$Var(X_i) = EX_i^2 - (EX_i)^2 = \frac{1}{n} - \left(\frac{1}{n}\right)^2 = \frac{1}{n}\left(\frac{n-1}{n}\right)$$

y

$$\begin{aligned} EX_iX_j &= P(X_i = 1, X_j = 1) \\ &= P(X_j = 1 \mid X_i = 1)P(X_i = 1) \\ &= \frac{1}{n-1} \times \frac{1}{n}, \end{aligned}$$

entonces:

$$\begin{aligned} Cov(X_i, X_j) &= EX_iX_j - EX_iEX_j \\ &= \frac{1}{n-1} \times \frac{1}{n} - \frac{1}{n^2} \\ &= \frac{1}{n}\left[\frac{1}{n(n-1)}\right]. \end{aligned}$$

De esta forma se obtiene:

$$Var(X) = \frac{n(n-1)}{n^2} + 2\binom{n}{2}\frac{1}{n^2(n-1)} = \frac{n-1}{n} + \frac{1}{n} = 1. \quad \blacktriangle$$

La covarianza es una medida de la asociación lineal entre dos variables aleatorias. Una covarianza "grande" indica que con probabilidad 1, hay una relación de tipo lineal entre las dos variables. Pero, ¿qué significa que la covarianza sea "grande"? ¿cómo se puede calificar la magnitud de una covarianza? de hecho, la propiedad 4. de la covarianza establece que el valor de ésta depende de la escala de medida que se utilice. Por esta razón, es difícil, en casos concretos, determinar a simple vista, si una covarianza es "grande" o no. Para eliminar este problema, el matemático inglés Karl Pearson, quien desarrolló la gran mayoría de las técnicas estadísticas modernas, introdujó el concepto siguiente:

Definición 5.64 (coeficiente de correlación) *Sean X e Y variables aleatorias reales, con $0 < VarX < \infty$ y $0 < VarY < \infty$. El coeficiente de correlación entre X e Y se define como:*

$$\rho(X,Y) := \frac{Cov(X,Y)}{\sqrt{VarX}\sqrt{VarY}}.$$

Teorema 5.65 *Sean X e Y variables aleatorias reales con $0 < VarX < \infty$ y $0 < VarY < \infty$.*

1. $\rho(X,Y) = \rho(Y,X)$.
2. $|\rho(X,Y)| \leq 1$.
3. $\rho(X,X) = 1$ *y* $\rho(X,-X) = -1$.
4. $\rho(aX+b,Y) = \rho(X,Y)$; *para* $a,b \in \mathbb{R}$, *con* $a > 0$.
5. $|\rho(X;Y)| = 1$, *si y sólo si, existen constantes* $a,b \in \mathbb{R}$ *no simultáneamente cero tales que* $P(aX+bY=0) = 1$.

Demostración. *Se hace solamente la demostración de los numerales 2. y 5. Los demás quedan como ejercicio para el lector.*

2. *Sean*

$$X^* := \frac{X-EX}{\sqrt{VarX}}$$

y

$$Y^* := \frac{Y-EY}{\sqrt{VarY}}$$

Es claro que $EX^* = EY^* = 0$ *y* $Var(X^*) = Var(Y^*) = 1$. *Por lo tanto:*

$$\begin{aligned}\rho(X^*,Y^*) &= EX^*Y^* \\ &= \frac{EXY - EXEY}{\sqrt{VarX}\sqrt{VarY}} \\ &= \rho(X,Y)\end{aligned}$$

Luego:

$$0 \leq Var(X^* \pm Y^*) = VarX^* \pm 2Cov(X^*,Y^*) + VarY^* = 2(1 \pm \rho(X,Y))$$

de donde se deduce:

$$|\rho(X,Y)| \leq 1$$

5. *Sean X^* e Y^*, como en el numeral 2. Es claro que:*

$$\begin{aligned}
\rho(X,Y)=1 &\iff \rho(X^*,Y^*)=1\\
&\iff (EX^*Y^*)^2=E(X^*)^2E(Y^*)^2\\
&\iff (\exists\alpha,\beta)(\alpha\neq 0\vee\beta\neq 0)(P(\alpha X^*+\beta Y^*=0)=1)\\
&\iff P\left(\alpha\left(\frac{X-EX}{\sqrt{VarX}}\right)+\beta\left(\frac{Y-EY}{\sqrt{VarY}}\right)=0\right)=1\\
&\iff P(\widetilde{\alpha}X+\widetilde{\beta}Y=0)=1,
\end{aligned}$$

donde

$$\widetilde{\alpha}=\frac{\alpha}{\sqrt{VarX}},\ \widetilde{\beta}=\frac{\beta}{\sqrt{VarY}}\ y\ C=\frac{\alpha EX}{\sqrt{VarX}}+\frac{\beta EY}{\sqrt{VarY}},$$

esto es,

$$\rho(X,Y)=1\iff P(Y=aX+b)=1,$$

siendo a,b constantes apropiadas.

Análogamente, se trabaja el caso $\rho(X,Y)=-1$.

■

El item 5., del teorema anterior, indica que si $|\rho(X,Y)|\approx 1$, entonces, $Y(\omega)\approx aX(\omega)+b$; para todo $\omega\in\Omega$. En la práctica, un coeficiente de correlación "grande", en valor absoluto, indica que se puede predecir Y a partir de X y viceversa.

5.4. Distribución de una función de un vector aleatorio.

Sea $\mathbf{X}=(X_1,X_2,\cdots,X_n)$ un vector aleatorio y sean $g_1(\cdot,\cdots,\cdot),\cdots,$ $g_k(\cdot,\cdots,\cdot)$ funciones definidas sobre $A\subseteq\mathbb{R}^n$ y de valor real. Supóngase que $Y_1:=g_1(X_1,\cdots,X_n),\cdots,Y_k:=g_k(X_1,\cdots,X_n)$ son variables aleatorias reales. Se desea determinar la distribución conjunta de las variables aleatorias $Y_1,\cdots,Y_k$, en términos de la distribución conjunta de las variables aleatorias $X_1,\cdots,X_n$.

Supóngase, inicialmente, que las variables aleatorias $X_1, \cdots, X_n$ son discretas y que se conoce su distribución conjunta. Es claro que:

$$\begin{aligned} P(Y_1 = y_1, \cdots, Y_k = y_k) &= P(g_1(X_1, \cdots, X_n), \cdots, g_k(X_1, \cdots, X_n)) \\ &= \sum_{\substack{(x_1,\cdots,x_n)\in\mathbb{R}^n: \\ g_1(x_1,\cdots,x_n)=y_1 \\ \vdots \\ g_k(x_1,\cdots,x_n)=y_k}} P(X_1 = x_1, \cdots, X_n = x_n). \end{aligned}$$

Ejemplo 5.66 *Sean X_1 y X_2 variables aleatorias con distribución conjunta dada por:*

$X_1 \backslash X_2$	0	1
-1	$\frac{1}{7}$	$\frac{1}{7}$
0	$\frac{2}{7}$	$\frac{1}{7}$
1	$\frac{1}{7}$	$\frac{1}{7}$

.

Sean $g_1(x_1, x_2) := x_1 + x_2$ y $g_2(x_1, x_2) = x_1 x_2$. Es claro que, las variables aleatorias

$$Y_1 := g_1(X_1, X_2) = X_1 + X_2 \quad y \quad Y_2 := g_2(X_1, X_2) = X_1 X_2$$

toman, respectivamente, los valores $-1, 0, 1, 2$ y $-1, 0, 1$. La distribución conjunta de Y_1 y Y_2 está dada por:

$Y_1 \backslash Y_2$	-1	0	1
-1	0	$\frac{1}{7}$	0
0	$\frac{1}{7}$	$\frac{2}{7}$	0
1	0	$\frac{2}{7}$	0
2	0	0	$\frac{1}{7}$

▲.

Ejemplo 5.67 *Sean X_1, X_2 y X_3 variables aleatorias con distribución conjunta dada por:*

$\mathbf{x} = (x_1, x_2, x_3)$	$(0,0,0)$	$(0,0,1)$	$(0,1,1)$	$(1,0,1)$	$(1,1,0)$	$(1,1,1)$
$P((X_1, X_2, X_3) = \mathbf{x})$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

Sean $g_1(x_1,x_2,x_3) := x_1 + x_2 + x_3$, $g_2(x_1,x_2,x_3) = |x_3 - x_2|$. *La distribución conjunta de* $Y_1 := g_1(X_1,X_2,X_3)$ *y* $Y_2 := g_2(X_1,X_2,X_3)$ *está dada por:*

$Y_2 \backslash Y_1$	0	1	2	3
0	$\frac{1}{8}$	0	$\frac{1}{8}$	$\frac{1}{8}$
1	0	$\frac{3}{8}$	$\frac{2}{8}$	0

. ▲

Para el caso de variables aleatorias absolutamente continuas se tiene resultado siguiente, el cual se presenta sin demostración. El lector interesado puede consultar la demostración en [Jac].

Teorema 5.68 (teorema de transformación)
Sea $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ *un vector aleatorio con densidad conjunta* $f_{\mathbf{X}}$. *Sea* $g : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *una aplicación inyectiva. Supóngase que tanto* g *como su inversa* $h : A \subseteq \mathbb{R}^n \longrightarrow \mathbb{R}^n$ *son continuas. Si las derivadas parciales de* h *existen y son continuas y si su jacobiano,* J*, es diferente de cero, entonces, el vector aleatorio* $\mathbf{Y} := g(\mathbf{X})$ *tiene función de densidad conjunta* $f_{\mathbf{Y}}$ *dada por:*

$$f_{\mathbf{Y}}(\mathbf{y}) = \begin{cases} |J(\mathbf{y})| \, f_{\mathbf{X}}(h(\mathbf{y})) & \text{si } \mathbf{y} \text{ está en el rango de } g \\ 0 & \text{en otro caso} \end{cases}$$

Ejemplo 5.69 *Sea* $\mathbf{X} = (X_1, X_2)$ *un vector aleatorio con función de densidad de probabilidad conjunta dada por:*

$$f_{\mathbf{X}}(x_1,x_2) = \begin{cases} 1 & \text{si} \quad 0 < x_1 < 1, \ 0 < x_2 < 1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Hallar la función de densidad de probabilidad conjunta de $\mathbf{Y} = (Y_1, Y_2)$, *donde* $Y_1 = X_1 + X_2$ *y* $Y_2 = X_1 - X_2$.

Solución: En este caso se tiene que

$$g(x_1,x_2) = (g_1(x_1,x_2), g_2(x_1,x_2)) = (x_1 + x_2, x_1 - x_2).$$

La transformación inversa está dada por:

$$h(x_1,x_2) = \left(\frac{x_1 + x_2}{2}, \frac{x_1 - x_2}{2} \right).$$

y el jacobiano J de la transformación inversa por:

$$J = \begin{vmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{-1}{2} \end{vmatrix} = -\frac{1}{2}$$

Por lo tanto, la función de densidad de probabilidad conjunta de $\mathbf{Y}$ es:

$$f_{\mathbf{Y}}(y_1, y_2) = \begin{cases} \frac{1}{2} & \text{si } \; 0 < y_1 + y_2 < 2 \;,\; 0 < y_1 - y_2 < 2 \\ 0 & \text{en otro caso.} \end{cases} \quad \blacktriangle$$

Ejemplo 5.70 ***(distribución de la suma y de la diferencia de variables aleatorias)*** *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta f. Sean $Z := X + Y$ y $W := X - Y$. Hallar la función de densidad de probabilidad conjunta de Z y W. Deducir luego, una función de densidad de Z y una función de densidad de W.*

Solución: Al igual que en el ejemplo anterior se tiene que:

$$g(x, y) := (x + y, x - y).$$

La transformación inversa h, está dada por:

$$h(x, y) = \left(\frac{x+y}{2}, \frac{x-y}{2} \right).$$

Por lo tanto, la función de densidad de probabilidad conjunta de Z y W está dada por:

$$f_{(Z,W)}(x, y) = \frac{1}{2} f \left(\frac{x+y}{2}, \frac{x-y}{2} \right).$$

luego,

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{2} f \left(\frac{z+y}{2}, \frac{z-y}{2} \right) dy \\ &= \int_{-\infty}^{\infty} f(u, z-u) du \\ &= \int_{-\infty}^{\infty} f(z-u, u) du. \end{aligned}$$

Análogamente se obtiene:

$$\begin{aligned} f_W(w) &= \int_{-\infty}^{\infty} \frac{1}{2} f \left(\frac{x+w}{2}, \frac{x-w}{2} \right) dx \\ &= \int_{-\infty}^{\infty} f(u, u-w) du \\ &= \int_{-\infty}^{\infty} f(u+w, u) du. \quad \blacktriangle \end{aligned}$$

Nota 5.71 *En el caso, especial, en el que las variables aleatorias X e Y sean independientes, se tiene que la función de densidad para la variable aleatoria $Z := X + Y$ está dada por:*

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(u)f_Y(z-u)du = \int_{-\infty}^{\infty} f_X(z-u)f_Y(u)du \tag{5.9}$$

donde $f_X(\cdot)$ y $f_Y(\cdot)$ denotan las funciones de densidad de X e Y respectivamente. La expresión dada en (5.9) es la convolución de f_X y f_Y, denotada por $f_X * f_Y$ (comparar con (5.7)).

Ejemplo 5.72 *Sean X e Y variables aleatorias independientes con funciones de densidad dadas por:*

$$f_X(x) = \begin{cases} 1 & si \quad 1 < x < 2 \\ 0 & en \quad otro\ caso \end{cases}$$

y

$$f_Y(y) = \begin{cases} \frac{1}{2} & si \quad 3 < y < 5 \\ 0 & en \quad otro\ caso \end{cases}$$

Determinar la función de densidad de $Z = X + Y$.

Solución: Se sabe que:

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(u)f_Y(z-u)du.$$

Como

$$f_X(u)f_Y(z-u) = \begin{cases} \frac{1}{2} & \text{si} \quad 1 < u < 2 \text{ y } u+3 < z < u+5 \\ 0 & \text{en} \quad \text{en otro caso} \end{cases}$$

se obtiene

$$f_Z(z) = \begin{cases} \frac{1}{2}(z-4) & \text{si} \quad 4 < z \le 5 \\ \frac{1}{2} & \text{si} \quad 5 < z < 6 \\ \frac{1}{2}(7-z) & \text{si} \quad 6 \le z < 7 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases} \quad \blacktriangle$$

Ejemplo 5.73 *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:*

$$f(x,y) = \begin{cases} 3x & si \quad 0 \le y \le x \le 1 \\ 0 & en \quad otro\ caso \end{cases}$$

Hallar la función de densidad de $W := X - Y$.

Solución: De acuerdo con el ejemplo anterior:

$$f_W(w) = \int_{-\infty}^{\infty} f(u, u-w)du.$$

Puesto que:

$$f(u, u-w) = \begin{cases} 3u & \text{si} \quad 0 \leq u - w \leq u \leq 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

entonces,

$$f_W(w) = \int_w^1 3u\mathcal{X}_{[0,1]}(w)du = \frac{3}{2}(1-w^2)\mathcal{X}_{[0,1]}(w). \quad \blacktriangle$$

Ejemplo 5.74 ***(distribución del producto de variables aleatorias)*** *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta f. Sean $Z := XY$ y $W := Y$. Hallar la función de densidad de probabilidad conjunta de Z y W. Deducir luego, una función de densidad de Z.*

Solución: Sea g la función dada por:

$$g(x, y) := (g_1(x, y), g_2(x, y)) = (xy, y).$$

La transformación inversa es igual a:

$$h(x, y) = \left(\frac{x}{y}, y\right).$$

El jacobiano, de la transformación inversa, es igual a:

$$J(x, y) = \begin{vmatrix} \frac{1}{y} & \frac{-x}{y^2} \\ 0 & 1 \end{vmatrix} = \frac{1}{y},$$

por lo tanto, la función de densidad de probabilidad conjunta de Z y W está dada por:

$$f_{ZW}(z, w) = \left|\frac{1}{w}\right| f\left(\frac{z}{w}, w\right),$$

de donde se deduce que la función de densidad de $Z = XY$ es:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{|w|} f\left(\frac{z}{w}, w\right) dw \\ &= \int_{-\infty}^{\infty} \frac{1}{|u|} f\left(u, \frac{z}{u}\right) du. \quad \blacktriangle \end{aligned}$$

Ejemplo 5.75 *Sean X e Y variables aleatorias independientes e igualmente distribuidas, con distribución uniforme sobre el intervalo $(0,1)$.*

De acuerdo con el ejemplo anterior, se tiene que una función de densidad de $Z := XY$ está dada por:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} \frac{1}{|u|} f\left(u, \frac{z}{u}\right) du \\ &= \int_{-\infty}^{\infty} \frac{1}{|u|} f_X(u) f_Y\left(\frac{z}{u}\right) du \end{aligned}$$

donde f_X y f_Y denotan las funciones de densidad de X e Y respectivamente.

Puesto que:

$$f_X(u) f_Y\left(\frac{z}{u}\right) = \begin{cases} 1 & si \quad 0 < u < 1 \text{ y } 0 < z < u \\ 0 & en \quad otro\ caso \end{cases}$$

entonces:

$$\begin{aligned} f_Z(z) &= \begin{cases} \int_z^1 \frac{1}{u} du & si \quad 0 < z < 1 \\ 0 & en \quad otro\ caso \end{cases} \\ &= \begin{cases} -\ln z & si \quad 0 < z < 1 \\ 0 & en \quad otro\ caso. \end{cases} \quad \blacktriangle \end{aligned}$$

Ejemplo 5.76 ***(distribución del cociente de variables aleatorias)*** *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta f. Sean $Z := \frac{X}{Y}$ (está definida si $P(Y = 0) = 0$) y $W := Y$. Hallar la función de densidad de probabilidad conjunta de Z y W. Deducir luego, una función de densidad de Z.*

Solución: Considérese la función:

$$g(x, y) := (g_1(x, y), g_2(x, y)) = \left(\frac{x}{y}, y\right).$$

La transformación inversa está dada por:

$$h(x, y) = (xy, y).$$

El jacobiano de la transformación inversa es igual a:

$$J(x, y) = \begin{vmatrix} y & x \\ 0 & 1 \end{vmatrix} = y,$$

entonces, la función de densidad de probabilidad conjunta de Z y W está dada por:

$$f_{ZW}(z,w) = |w| \, f(zw, w).$$

Por lo tanto, una función de densidad de Z es:

$$f_Z(z) = \int_{-\infty}^{\infty} |w| \, f(zw, w) \, dw. \quad \blacktriangle$$

Ejemplo 5.77 *Sean X e Y variables aleatorias independientes e igualmente distribuidas con distribución uniforme sobre el intervalo $(0,1)$.*

De acuerdo con el ejemplo anterior, se tiene que una función de densidad de $Z := \frac{X}{Y}$ está dada por:

$$\begin{aligned} f_Z(z) &= \int_{-\infty}^{\infty} |w| \, f(zw, w) \, dw \\ &= \int_{-\infty}^{\infty} |w| \, f_X(zw) f_Y(w) \, dw \end{aligned}$$

donde f_X y f_Y denotan las funciones de densidad de X e Y respectivamente.

Puesto que:

$$f_X(zw) f_Y(w) = \begin{cases} 1 & si \quad 0 < zw < 1 \ y \ 0 < w < 1 \\ 0 & en \quad otro\ caso \end{cases}$$

entonces:

$$\begin{aligned} f_Z(z) &= \begin{cases} \int_0^1 w dw & si \quad 0 < z < 1 \\ \int_0^{\frac{1}{z}} w dw & si \quad z \geq 1 \\ 0 & en \quad otro\ caso \end{cases} \\ &= \frac{1}{2} \mathcal{X}_{(0,1)}(z) + \frac{1}{2z^2} \mathcal{X}_{[1,\infty)}(z). \quad \blacktriangle \end{aligned}$$

Ejemplo 5.78 *Supóngase que la duración X de un dispositivo electrónico es una variable aleatoria continua con función de densidad de probabilidad dada por:*

$$f(x) = \begin{cases} \frac{1000}{x^2} & si \ \ x > 1000 \\ 0 & en\ otro\ caso \end{cases}$$

Sean X_1 y X_2 dos determinaciones independientes de la anterior variable. Hallar la función de densidad de probabilidad de la variable aleatoria $Z = \frac{X_1}{X_2}$.

Solución: Se sabe que

$$f_Z(z) = \int_{-\infty}^{\infty} |v|\, f(vz) f(v) dv.$$

Puesto que

$$f(vz) = \begin{cases} \frac{1000}{v^2 z^2} & \text{si} \quad vz > 1000 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

entonces:

i) $z > 1$, $v > 1000$ implica $vz > 1000$ y así:

$$f_z(z) = \int_{1000}^{\infty} \frac{1000}{v^2 z^2} \cdot \frac{1000}{v^2} v dv = \frac{(1000)^2}{z^2} \int_{1000}^{\infty} \frac{dv}{v^3} = \frac{1}{2z^2}$$

ii) Si $0 < z < 1$, entonces $vz > 1000$, si y sólo si, $v > \frac{1000}{z}$. En tal caso:

$$f_z(z) = \int_{\frac{1000}{z}}^{\infty} \frac{1000}{v^2 z^2} \cdot \frac{1000}{v^2} v dv = \frac{1}{2}.$$

Esto es:

$$f_z(z) = \begin{cases} \frac{1}{2z^2} & \text{si} \quad z > 1 \\ \frac{1}{2} & \text{si} \quad 0 < z < 1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases} \quad \blacktriangle$$

Ejemplo 5.79 (distribución t-Student) *Sean X e Y variables aleatorias independientes tales que $X \stackrel{d}{=} \mathcal{N}(0,1)$ y $Y \stackrel{d}{=} \mathcal{X}^2_{(k)}$. Considérese la transformación:*

$$g(x,y) := \left(\frac{x}{\sqrt{\frac{y}{k}}}, y \right).$$

La transformación inversa está dada por:

$$h(x,y) = \left(x\sqrt{\frac{y}{k}}, y \right),$$

El jacobiano de la trasformación inversa es:

$$J(x,y) = \sqrt{\frac{y}{k}}.$$

Por lo tanto, la función de densidad de probabilidad conjunta de $Z := \frac{X}{\sqrt{\frac{Y}{k}}}$ *y* $W := Y$ *está dada por:*

$$f_{ZW}(z,w) = \sqrt{\frac{w}{k}} f(z\sqrt{\frac{w}{k}}, w); \quad para\ w > 0$$

donde:

$$f(x,y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right) \left[\Gamma\left(\frac{k}{2}\right)\right]^{-1} \left(\frac{1}{2}\right)^{\frac{k}{2}} y^{\frac{k}{2}-1} \exp\left(-\frac{1}{2}y\right);$$

para $y > 0$.

Integrando $f_{ZW}(z,w)$ *con respecto a* w*, se obtiene que una función de densidad de* Z *está dada por:*

$$f_Z(z) = \frac{1}{\sqrt{k\pi}} \Gamma\left(\frac{k+1}{2}\right) \left[\Gamma\left(\frac{k}{2}\right)\right]^{-1} \frac{1}{\left[1+\frac{z^2}{k}\right]^{\frac{k+1}{2}}}; \ z \in \mathbb{R} \tag{5.10}$$

Se dice que una variable aleatoria real Z *tiene distribución* $t-$*Student con* k *grados de libertad, y se escribe* $Z \stackrel{d}{=} t_{(k)}$*, si su función de densidad está dada por (5.10).* ▲

Ejemplo 5.80 (distribución F) *Sean* X *e* Y *variables aleatorias independientes tales que* $X \stackrel{d}{=} \mathcal{X}^2_{(m)}$ *y* $Y \stackrel{d}{=} \mathcal{X}^2_{(n)}$*. Considérese la aplicación*

$$g(x,y) := \left(\frac{\frac{x}{m}}{\frac{y}{n}}, y\right).$$

Se tiene que su inversa está dada por:

$$h(x,y) = \left(\frac{m}{n}xy, y\right).$$

El jacobiano de la transformación inversa es igual a: $J(x,y) = \frac{m}{n}y$*. Por lo tanto, la función de densidad de probabilidad conjunta de* $Z := \frac{nX}{mY}$ *y* $W := Y$ *está dada por:*

$$f_{ZW}(z,w) = \frac{1}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \left(\frac{1}{2}\right)^{\frac{m}{2}} \left(\frac{1}{2}\right)^{\frac{n}{2}} \left(\frac{m}{n}zw\right)^{\frac{m}{2}-1} w^{\frac{n}{2}-1} \exp\left[-\frac{1}{2}w\left(\frac{m}{n}z+1\right)\right].$$

Al integrar con respecto a w*, se obtiene que una función de densidad de* Z *está dada por:*

$$f_Z(z) = \Gamma\left(\frac{m+n}{2}\right) \left[\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)\right]^{-1} \left(\frac{m}{n}\right)^{\frac{m}{2}} \frac{z^{\frac{m}{2}-1}}{\left(1+\frac{m}{n}z\right)^{\frac{m+n}{2}}}; \ z > 0 \tag{5.11}$$

Se dice que una variable aleatoria Z tiene distribución F, con m grados de libertad en el numerador y n grados de libertad en el denominador, y se escribe $Z \stackrel{d}{=} F_n^m$, si su función de densidad está dada por (5.11). ▲

El teorema de transformación se puede generalizar de la siguiente manera:

Teorema 5.81 *Sea $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ un vector aleatorio con función de densidad de probabilidad conjunta $f_{\mathbf{X}}$. Sea g una aplicación de $\mathbb{R}^n$ en si mismo. Supóngase que $\mathbb{R}^n$ se puede particionar en k conjuntos disyuntos $A_1, \cdots, A_k$ de tal manera que la aplicación g restringida a A_i, para $i = 1, \cdots, k$, es una aplicación uno a uno con inversa h_i. Si las primeras derivadas parciales de h_i existen y son continuas y si los jacobianos J_i son diferentes de cero en el rango de la transformación, para $i = 1, \cdots, k$. Entonces el vector aleatorio $\mathbf{Y} := g(\mathbf{X})$ tiene función de densidad de probabilidad conjunta dada por:*

$$f_{\mathbf{Y}}(\mathbf{y}) = \sum_{i=1}^{m} |J_i(\mathbf{y})|\, f_{\mathbf{X}}\left(h_i(\mathbf{y})\right).$$

Demostración. *Ver [Jac]* ■

Ejemplo 5.82 *Sea X una variable aleatoria con función de densidad f_X. Sea $g : \mathbb{R} \longrightarrow \mathbb{R}$ dada por $g(x) = x^4$. Es claro que $\mathbb{R} = (-\infty, 0) \cup [0, \infty)$ y que las aplicaciones*

$$\begin{array}{rccc} g_1 : & [0,\infty) & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^4 \end{array}$$

y

$$\begin{array}{rccc} g_2 : & (-\infty,0) & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & x^4 \end{array}$$

son invertibles y sus inversas son respectivamente:

$$\begin{array}{rccc} h_1 : & [0,\infty) & \longrightarrow & [0,\infty) \\ & x & \longmapsto & \sqrt[4]{x} \end{array}$$

y

$$\begin{array}{rccc} h_2 : & (0,\infty) & \longrightarrow & (-\infty,0) \\ & x & \longmapsto & -\sqrt[4]{x} \end{array}.$$

Los jacobianos de las transformaciones inversas están dados, respectivamente, por:

$$J_1(x) = \frac{1}{4}x^{-\frac{3}{4}} \qquad y \qquad J_2(x) = -\frac{1}{4}x^{-\frac{3}{4}},$$

por lo tanto, una función de densidad de la variable aleatoria $Y = X^4$ *está dada por:*

$$
\begin{aligned}
f_Y(y) &= \sum_{i=1}^{2} |J_i(y)|\, f_X(h_i(y)) \\
&= \frac{1}{4} y^{-\frac{3}{4}} f_X(\sqrt[4]{y}) + \frac{1}{4} y^{-\frac{3}{4}} f_X(-\sqrt[4]{y}). \quad \blacktriangle
\end{aligned}
$$

5.5. Valor esperado de un vector aleatorio y matriz de varianzas y covarianzas.

A continuación se presenta la generalización de los conceptos de valor esperado y varianza de una variable aleatoria estudiados previamente.

Definición 5.83 (valor esperado de un vector aleatorio)
Sea $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ *un vector aleatorio. La esperanza (o valor esperado) de* $\mathbf{X}$ *, denotada por* $E(\mathbf{X})$*, se define como sigue:*

$$E(\mathbf{X}) := (EX_1, EX_2, \cdots, EX_n)$$

siempre y cuando EX_j *exista; para todo* $j = 1, \cdots, n$.

La anterior definición puede extenderse de la manera siguiente:

Definición 5.84 (valor esperado de una función)
Sea $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ *un vector aleatorio y sea* $h : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *una función dada por:*

$$h(x_1, \cdots, x_n) = (h_1(x_1, \cdots, x_n), \cdots, h_m(x_1, \cdots, x_n))$$

donde las funciones h_i*, para* $i = 1, \cdots, m$*, son funciones de valor real definidas sobre* $\mathbb{R}^n$*. El valor esperado de* $h(\mathbf{X})$ *está dado por:*

$$E(h(\mathbf{X})) := (E(h_1(\mathbf{X})), \cdots, E(h_m(\mathbf{X})))$$

siempre y cuando los valores esperados $E(h_j(\mathbf{X}))$ *para* $j = 1, \cdots, m$ *estén definidos.*

Definición 5.85 (Valor esperado de una matriz aleatoria)
Si X_{ij}*; con* $i = 1, \cdots, m$ *y* $j = 1, \cdots, n$ *son variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad, entonces, la matriz*

$\mathbf{A} = (X_{ij})_{m\times n}$ *se llama matriz aleatoria y se define su valor esperado como la matriz cuyas componentes son los valores esperados de las variables aleatorias* X_{ij}*, esto es:*

$$E\mathbf{A} := (E(X_{ij}))_{m\times n}$$

siempre y cuando $E(X_{ij})$ *exista; para todo* $i = 1, \cdots, m$ *y todo* $j = 1, \cdots, n$*.*

Definición 5.86 (matriz de varianzas y covarianzas)
Sea $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ *un vector aleatorio con* $EX_j^2 < \infty$*; para todo* $j = 1, \cdots, n$*. La matriz de varianzas y covarianzas denotada por* $\sum$ *se define como sigue:*

$$\sum := \begin{pmatrix} Var(X_1) & Cov(X_1, X_2) & \cdots & Cov(X_1, X_n) \\ Cov(X_2, X_1) & Var(X_2) & \cdots & Cov(X_2, X_n) \\ \vdots & \vdots & \vdots & \vdots \\ Cov(X_n, X_1) & Cov(X_n, X_2) & \cdots & Var(X_n) \end{pmatrix}.$$

Nota 5.87 *Obsérvese que* $\sum = E([\mathbf{X}-E\mathbf{X}]^T\,[\mathbf{X}-E\mathbf{X}])$.

Ejemplo 5.88 *Sea* X_1 *y* X_2 *variables aleatorias discretas con distribución conjunta dada por:*

$X_1 \backslash X_2$	-1	0	1
1	$\frac{1}{6}$	$\frac{2}{6}$	0
2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

.

Es claro que el valor esperado de $\mathbf{X} = (X_1, X_2)$ *es igual a:*

$$E(\mathbf{X}) = \left(\frac{3}{2}, -\frac{1}{6}\right)$$

y la matriz de varianzas y covarianzas está dada por:

$$\sum = \begin{pmatrix} \frac{1}{4} & \frac{1}{12} \\ \frac{1}{12} & \frac{17}{36} \end{pmatrix}.$$

Para $a = (a_1, a_2)$*, arbitrario, se obtiene que:*

$$\begin{aligned} a\sum a^T &= (a_1, a_2)\begin{pmatrix} \frac{1}{4} & \frac{1}{12} \\ \frac{1}{12} & \frac{17}{36} \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\ &= \frac{1}{4}a_1^2 + \frac{1}{6}a_1 a_2 + \frac{17}{36}a_2^2 \\ &= \left(\frac{1}{2}a_1 + \frac{1}{6}a_2\right)^2 + \frac{4}{9}a_2^2 \geq 0, \end{aligned}$$

esto es, la matriz $\sum$ es positiva semidefinida. ▲

Ejemplo 5.89 *Sea $\mathbf{X} = (X, Y)$ un vector aleatorio con función de densidad de probabilidad conjunta dada por:*

$$f(x,y) = \begin{cases} \frac{1}{x} & si \quad 0 < x \leq 1, \ 0 < y \leq x \\ 0 & en \qquad otro\ caso \end{cases}.$$

En este caso se tiene que:

$$\begin{aligned}
EXY &= \int_0^1 \int_0^x xy\frac{1}{x}dydx = \frac{1}{6} \\
EX &= \int_0^1 \int_y^1 x\frac{1}{x}dxdy = \frac{1}{2} \\
EY &= \int_0^1 \int_0^x y\frac{1}{x}dydx = \frac{1}{4} \\
EX^2 &= \int_0^1 \int_0^x x^2\frac{1}{x}dydx = \frac{1}{3} \\
EY^2 &= \int_0^1 \int_0^x y^2\frac{1}{x}dydx = \frac{1}{9}.
\end{aligned}$$

Por lo tanto, el valor esperado de $\mathbf{X}$ está dado por:

$$E\mathbf{X} = \left(\frac{1}{2}, \frac{1}{4}\right)$$

y su matriz de varianzas y covarianzas es igual a:

$$\sum = \begin{pmatrix} \frac{1}{12} & \frac{1}{24} \\ \frac{1}{24} & \frac{7}{144} \end{pmatrix}.$$

Por otra parte, para cualquier $a = (a_1, a_2) \in \mathbb{R}^2$ se satisface:

$$\begin{aligned}
a\sum a^T &= (a_1, a_2) \begin{pmatrix} \frac{1}{12} & \frac{1}{24} \\ \frac{1}{24} & \frac{7}{144} \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \\
&= \frac{1}{12}\left(a_1^2 + 2\frac{1}{2}a_1a_2 + \frac{1}{4}a_2^2 - \frac{1}{4}a_2^2 + \frac{7}{12}a_2^2\right) \\
&= \frac{1}{12}\left(a_1 + \frac{1}{2}a_2\right)^2 + \frac{1}{36}a_2^2 \geq 0,
\end{aligned}$$

esto es, $\sum$ es positiva semidefinida. ▲

En general, se tiene el resultado siguiente:

Teorema 5.90 *Sea* $\mathbf{X} = (X_1, \cdots, X_n)$ *un vector aleatorio* $n-$*dimensional. Si* $E(X_j^2) < \infty$; *para todo* $j = 1, \cdots, n$, *entonces, la matriz* $\sum$, *de varianzas y covarianzas de* $\mathbf{X}$, *es positiva semidefinida.*
Demostración. *Sea* $a = (a_1, \cdots, a_n) \in \mathbb{R}^n$ *arbitrario. Considérese la variable aleatoria* Y *definida como sigue:*

$$\begin{aligned} Y &:= (X_1, \cdots, X_n)(a_1, \cdots, a_n)^T \\ &= \sum_{i=1}^{n} a_i X_i. \end{aligned}$$

Se va a verificar que $Var(Y) = a \sum a^T$ *con lo cual queda probado el resultado pues* $Var(Y) \geq 0$.

Se tiene que:

$$Var(Y) = E([Y - EY]^2).$$

Puesto que, $EY = \sum_{i=1}^{n} a_i E(X_i) = \mu a^T$, *siendo* $\mu := E\mathbf{X}$, *y como el transpuesto de un número real es el mismo número, se obtiene:*

$$\begin{aligned} VarY &= E\left(\left[\mathbf{X}a^T - \mu a^T\right]\left[\mathbf{X}a^T - \mu a^T\right]\right) \\ &= E(\left[\mathbf{X}a^T - \mu a^T\right]^T\left[\mathbf{X}a^T - \mu a^T\right]) \\ \\ &= E(\left[a\mathbf{X}^T - a\mu^T\right]\left[\mathbf{X}a^T - \mu a^T\right]) \\ &= E(a[\mathbf{X} - \mu]^T [X - \mu] a^T) \\ &= aE([\mathbf{X} - \mu]^T [X - \mu]) a^T \\ &= a \sum a^T. \end{aligned}$$

■

Nota 5.91 *Es claro que, a partir de la definición de la matriz de varianzas y covarianzas* $\sum$, *si las variables aleatorias* $X_1, X_2, \cdots, X_n$ *son independientes, entonces,* $\sum$ *es una matriz diagonal, donde los elementos de la diagonal son precisamente las varianzas de las variables aleatorias* X_j, $j = 1, \cdots, n$.

Para finalizar esta sección, se introduce el concepto de matriz de correlaciones, el cual es muy importante en el desarrollo de la teoría estadística multivariada.

Definición 5.92 (matriz de correlaciones) *Sea* $\mathbf{X} = (X_1, X_2, \cdots, X_n)$ *un vector aleatorio con* $0 < Var(X_j) < \infty$; *para todo* $j = 1, \cdots, n$. *La matriz de correlaciones de* $\mathbf{X}$, *denotada por* $\mathbf{R}$, *está definida como sigue:*

$$\mathbf{R} := \begin{pmatrix} 1 & \rho(X_1, X_2) & \cdots & \rho(X_1, X_n) \\ \rho(X_2, X_1) & 1 & \cdots & \rho(X_2, X_n) \\ \vdots & \vdots & \vdots & \vdots \\ \rho(X_n, X_1) & \rho(X_n, X_2) & \cdots & 1 \end{pmatrix}.$$

Ejemplo 5.93 *Sea* $\mathbf{X} = (X, Y)$ *un vector aleatorio bidimensional con función de densidad de probabilidad conjunta dada por:*

$$f(x, y) = \begin{cases} 2 & si \quad 0 < x < y, \ 0 < y < 1 \\ 0 & en \quad otro\ caso. \end{cases}$$

Se tiene que:

$$EX = \int_0^1 \int_0^y 2x dx dy = \frac{1}{3}$$

$$EY = \int_0^1 \int_x^1 2y dy dx = \frac{2}{3}$$

$$Var(X) = \int_0^1 \int_0^y 2(x - \frac{1}{3})^2 dx dy = \frac{1}{18}$$

$$Var(Y) = \int_0^1 \int_x^1 2(y - \frac{2}{3})^2 dy dx = \frac{1}{18}$$

$$Cov(X, Y) = \int_0^1 \int_0^y 2(x - \frac{1}{3})(y - \frac{2}{3}) dx dy = \frac{1}{36}$$

Por lo tanto, la matriz de correlaciones está dada por:

$$\mathbf{R} = \begin{pmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{pmatrix}. \quad \blacktriangle$$

Nota 5.94 *La matriz de correlaciones,* $\mathbf{R}$, *hereda las propiedades de la matriz de varianzas y covarianzas,* $\sum$, *ya que,*

$$\mathbf{R} = diag(\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_n}).\sum.diag(\frac{1}{\sigma_1}, \cdots, \frac{1}{\sigma_n})$$

donde $\sigma_j := \sqrt{VarX_j}$, *para* $j = 1, \cdots, n$.

Por lo tanto, $\mathbf{R}$ *es simétrica y positiva semidefinida . Más aún, si* $\sigma_j > 0$; *para todo* j, *entonces,* $\mathbf{R}$ *es no singular, si y sólo si,* $\sum$ *es no singular.*

5.6. Funciones generadoras de momentos y característica conjuntas.

En esta sección se generalizan los conceptos de función generadora de momentos y función característica, presentados en el capítulo 2. para el caso unidimensional, a vectores aleatorios n dimensionales.

Definición 5.95 (función generadora de momentos conjunta)
Sea $\mathbf{X} = (X_1, \cdots, X_n)$ un vector aleatorio n dimensional. Si existe $M > 0$ tal que $E(\exp(\mathbf{X}\mathbf{t}^T))$ es finito para todo $\mathbf{t} = (t_1, \cdots, t_n) \in \mathbb{R}^n$ con

$$\|\mathbf{t}\| := \sqrt{t_1^2 + \cdots + t_n^2} < M,$$

entonces, se define la función generadora de momentos conjunta de $\mathbf{X}$, denotada por $m_{\mathbf{X}}(\mathbf{t})$, como sigue:

$$m_{\mathbf{X}}(\mathbf{t}) := E(\exp(\mathbf{X}\mathbf{t}^T)) \quad \text{para } \|\mathbf{t}\| < M.$$

Ejemplo 5.96 *Sea X_1 y X_2 variables aleatorias discretas con distribución conjunta dada por:*

$X_1 \backslash X_2$	-1	0	1
1	$\frac{1}{6}$	$\frac{2}{6}$	0
2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

.

La función generadora de momentos conjunta de $\mathbf{X} = (X_1, X_2)$ está dada por:

$$\begin{aligned}
m_{\mathbf{X}}(\mathbf{t}) &= E(\exp((X_1, X_2)(t_1, t_2)^T)) \\
&= E(\exp(X_1 t_1 + X_2 t_2)) \\
&= \sum_{x_1}\sum_{x_2} \exp(t_1 x_1 + t_2 x_2) P(X_1 = x_1, X_2 = x_2) \\
&= \frac{1}{6}\exp(t_1 - t_2) + \frac{1}{6}\exp(2t_1 - t_2) + \frac{1}{3}\exp(t_1) + \frac{1}{6}\exp(2t_1) \\
&\qquad + \frac{1}{6}\exp(2t_1 + t_2).
\end{aligned}$$

Se observa que las funciones generadoras de momentos de X_1 y X_2 también *existen y están dadas, respectivamente, por:*

$$\begin{aligned}
m_{X_1}(t_1) &= E(\exp(t_1 X_1)) \\
&= \sum_{x_1} \exp(t_1 x_1) P(X_1 = x_1) \\
&= \frac{1}{2}\exp(t_1) + \frac{1}{2}\exp(2t_1)
\end{aligned}$$

y

$$\begin{aligned}
m_{X_2}(t_2) &= E(\exp(t_2 X_2)) \\
&= \sum_{x_2} \exp(t_2 x_2) P(X_2 = x_2) \\
&= \frac{1}{3}\exp(-t_2) + \frac{1}{6}\exp(t_2) + \frac{1}{2}
\end{aligned}$$

en general se tiene el resultado siguiente:

Teorema 5.97 *Sea* $\mathbf{X} = (X_1, \cdots, X_n)$ *un vector aleatorio n dimensional. La función generadora de momentos conjunta de* $\mathbf{X}$ *existe, si y sólo si, las funciones generadoras de momentos marginales, de las variables aleatorias* X_i *con* $i = 1, \cdots, n$, *existen.*

Demostración. $\Longrightarrow$) *Supóngase, inicialmente, que la función generadora de momentos conjunta de* $\mathbf{X}$ *existe. En tal caso, existe una constante* $M > 0$ *tal que* $m_{\mathbf{X}}(\mathbf{t}) = E(\exp(\mathbf{X}\mathbf{t}^T)) < \infty$ *para todo* $\mathbf{t}$ *con* $\|\mathbf{t}\| < M$. *Luego para todo* $i = 1, \cdots, n$ *se satisface:*

$$\begin{aligned}
m_{X_i}(t_i) &= E(\exp(t_i X_i)) \\
&= m_X(\exp(\mathbf{X}.(0, \cdots, 0, \underset{\substack{\uparrow \\ \textit{lugar } i}}{t_i}, 0, \cdots, 0)^T) \\
&< \infty \quad \text{si } |t_i| < M.
\end{aligned}$$

Esto es, la función generadora de momentos de X_i *existe, para* $i = 1, \cdots, n$.

$\Longleftarrow$) *Supóngase ahora, que las funciones generadoras de momentos marginales existen. En tal caso, para cada* $i = 1, \cdots, n$, *se tiene que existe* $M_i > 0$ *tal que* $m_{X_i}(t_i) = E(\exp(t_i X_i)) < \infty$; *si* $|t_i| < M_i$.

Sea $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}$ *definida por:*

$$\mathbf{h}(\mathbf{t}) := \exp(\mathbf{t}\mathbf{a}^T); \quad \text{con } \mathbf{a} \in \mathbb{R}^n$$

La función $\mathbf{h}$ *es convexa y en consecuencia si* $\mathbf{x}_i \in \mathbb{R}^n$ *con* $i = 1, \cdots, m$ *se satisface que:*

$$\mathbf{h}(\sum_{i=1}^{m} \alpha_i \mathbf{x}_i) \leq \sum_{i=1}^{m} \alpha_i \mathbf{h}(\mathbf{x}_i)$$

donde $\alpha_i \in (0,1)$ *para todo* $i = 1, \cdots, m$ *y* $\sum_{i=1}^{m} \alpha_i = 1$. *Por lo tanto,*

$$\exp\left\{\left(\sum_{i=1}^{m} \alpha_i \mathbf{x}_i\right) \mathbf{a}^T\right\} \leq \sum_{i=1}^{m} \alpha_i \exp(\mathbf{x}_i \mathbf{a}^T)$$

En particular, para $\mathbf{a} = \mathbf{X} = (X_1, \cdots, X_n)$, $\mathbf{x}_i = (0, \cdots, 0, \underset{\substack{\uparrow \\ \text{lugar } i}}{t_i}, 0, \cdots, 0)$ *y* $n = m$ *se obtiene que:*

$$\exp\left\{\sum_{i=1}^{n} \alpha_i t_i X_i\right\} \leq \sum_{i=1}^{n} \alpha_i \exp(t_i X_i)$$

Tomando valores esperados se llega a:

$$E\left(\exp\left\{\sum_{i=1}^{n} \alpha_i t_i X_i\right\}\right) \leq \sum_{i=1}^{n} \alpha_i m_{X_i}(t_i)$$

y por lo tanto, para todo $\mathbf{u} \in \mathcal{R}$ *con*

$$\mathcal{R} := \{\mathbf{u} = (u_1, \cdots, u_n) \in \mathbb{R}^n : u_i = \alpha_i t_i \textit{ con } |t_i| < M_i \textit{ y } \alpha_i \in (0,1), \\ i = 1, \cdots, n\}$$

se satisface que:

$$E(\exp(\mathbf{X}\mathbf{u}^T)) < \infty$$

Sea $\mathbf{t} \in \mathbb{R}^n$ *con* $\|\mathbf{t}\| < M$ *donde* $M := \text{mín}\{\alpha_1 M_1, \cdots, \alpha_n M_n\}$. *Entonces:*

$$E(\exp(\mathbf{X}\mathbf{t}^T)) < \infty$$

es decir, la función generadora de momentos conjunta de $\mathbf{X}$ *existe.* ■

El teorema anterior establece que la función generadora de momentos conjunta, de las variables aleatorias, $X_1, \cdots, X_n$, existe, si y sólo si, las funciones generadoras de momentos marginales existen también. Sin embargo, no dice que se pueda conocer, a partir de las funciones generadoras de momentos marginales, la función generadora de momentos conjunta. Esto es posible si las variables aleatorias son independientes, como lo establece el teorema siguiente:

Teorema 5.98 *Sea* $\mathbf{X} = (X_1, \cdots, X_n)$ *un vector aleatorio n dimensional. Supóngase, que para cada* $i = 1, \cdots, n$, *existe* $M_i > 0$ *con:*

$$m_{X_i}(t) := E(\exp tX_i) < \infty; \quad si \ |t| < M_i$$

Si las variables aleatorias, $X_1, \cdots, X_n$, *son independientes, entonces,* $m_X(\mathbf{t}) < \infty$ *para todo* $\mathbf{t} = (t_1, \cdots, t_n)$; *con* $\|\mathbf{t}\| < M$, *siendo* $M := \text{mín}\{M_1, \cdots, M_n\}$ *y se satisface además que:*

$$m_X(\mathbf{t}) = \prod_{i=1}^{n} m_{X_i}(t_i)$$

Demostración. *Se tiene que:*

$$\begin{aligned}
m_X(\mathbf{t}) &= E(\exp(\mathbf{X}\mathbf{t}^T)) \\
&= E\left\{\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right\} \\
&= E\left\{\prod_{i=1}^{n} \exp(t_i X_i)\right\} \\
&= \prod_{i=1}^{n} E(\exp(t_i X_i)), \ \textit{por la independencia de las variables,} \\
&= \prod_{i=1}^{n} m_{X_i}(t_i).
\end{aligned}$$

■

Al igual que en el caso unidimensional, la función generadora de momentos conjunta permite, cuando existe, calcular los *momentos conjuntos del vector aleatorio* $\mathbf{X}$, *alrededor del origen.* A continuación se define éste último concepto:

Definición 5.99 (momento conjunto) *Sea* $\mathbf{X} = (X_1, \cdots, X_n)$ *un vector aleatorio n dimensional. Se define el momento conjunto de orden* $k_1, \cdots, k_n$ *con* $k_j \in \mathbb{N}$, *del vector aleatorio* $\mathbf{X}$, *alrededor del punto* $\mathbf{a} = (a_1, \cdots, a_n)$, *como:*

$$\mu_{k_1 \cdots k_n} := E\left(\prod_{j=1}^{n} (X_j - a_j)^{k_j}\right)$$

si el valor esperado existe.

El momento conjunto de orden $k_1, \cdots, k_n$ *de* $\mathbf{X}$, *alrededor del origen, se denota por* $\mu'_{k_1 \cdots k_n}$.

Nota 5.100 *Si X e Y son variables aleatorias, cuyos valores esperados existen, entonces, el momento conjunto de orden* $1, 1$, *del vector* $\mathbf{X} = (X, Y)$, *alrededor de* $\mathbf{a} = (EX, EY)$, *es:*

$$\mu_{11} = E((X - EX)(Y - EY)) = Cov(X, Y)$$

Ejemplo 5.101 *Sean X_1 y X_2 con distribución conjunta dada por:*

$X_1 \backslash X_2$	-1	0	1
1	$\frac{1}{6}$	$\frac{2}{6}$	0
2	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

.

Se tiene que:

$$\begin{aligned}
\mu'_{12} &= E(X_1 X_2^2) \\
&= \sum_{x_1} \sum_{x_2} x_1 x_2^2 P(X_1 = x_1, X_2 = x_2) \\
&= \frac{5}{6}.
\end{aligned}$$

Por otra parte,

$$\begin{aligned}
m_X(\mathbf{t}) = \frac{1}{6} \exp(t_1 - t_2) + \frac{1}{6} \exp(2t_1 - t_2) + \frac{1}{3} \exp(t_1) + \frac{1}{6} \exp(2t_1) \\
+ \frac{1}{6} \exp(2t_1 + t_2)
\end{aligned}$$

Además, es fácil verificar que:

$$\begin{aligned}
&\left. \frac{\partial^3 m_{\mathbf{X}}(t_1, t_2)}{\partial t_1 \partial t_2 \partial t_2} \right|_{(t_1, t_2) = (0,0)} \\
&= \left. \left[\frac{1}{6} \exp(t_1 - t_2) + \frac{1}{3} \exp(2t_1 - t_2) + \frac{1}{3} \exp(2t_1 + t_2) \right] \right|_{(t_1, t_2) = (0,0)} \\
&= \frac{5}{6} = \mu'_{12}. \quad \blacktriangle
\end{aligned}$$

El resultado anterior es válido en general, como lo establece el teorema siguiente, el cual se presenta sin demostración.

Teorema 5.102 *Sea* $\mathbf{X} = (X_1, \cdots, X_n)$ *un vector aleatorio n dimensional. Supóngase que la función generadora de momentos conjunta de* $m_{\mathbf{X}}(\mathbf{t})$ *de*

X *existe. Entonces, los momentos conjuntos, alrededor del origen, de todos los ordenes, son finitos y se satisface además que:*

$$\mu'_{k_1 \cdots k_n} = \left. \frac{\partial^{k_1+\cdots+k_n}}{\partial t_1^{k_1} \cdots \partial t_n^{k_n}} m_{\mathbf{X}}(\mathbf{t}) \right|_{(t_1,\cdots,t_n)=(0,\cdots,0)} .$$

Para finalizar esta sección, se presenta la definición de función característica conjunta de un vector aleatorio.

Definición 5.103 (función característica conjunta)
Sean $\mathbf{X} = (X_1, \cdots, X_n)$ *un vector aleatorio* n *dimensional y* $\mathbf{t} = (t_1, \cdots, t_n) \in \mathbb{R}^n$. *La función característica conjunta del vector aleatorio* $\mathbf{X}$, *denotada por* $\varphi_{\mathbf{X}}(\mathbf{t})$, *está definida por:*

$$\varphi_{\mathbf{X}}(\mathbf{t}) := E\left[\exp\left(i\mathbf{X}\mathbf{t}^T\right)\right];$$

donde $i = \sqrt{-1}$.

Al igual que en el caso univariado, se tiene que la función característica conjunta de un vector aleatorio siempre existe. También se satisface que la función característica conjunta de un vector aleatorio caracteriza la distribución del vector. Esto es, dos vectores aleatorios **X** e **Y** tienen la misma función de distribución conjunta, si y sólo si, tienen la misma función característica conjunta. Es más, es posible demostrar que al diferenciar la función característica sucesivamente y evaluar sus derivadas en el origen, se obtiene la expresión siguiente para los momentos conjuntos alrededor del origen:

$$\mu'_{k_1 \cdots k_n} = \frac{1}{i^{k_1+\cdots+k_n}} \left[\frac{\partial^{k_1+\cdots+k_n}}{\partial t_1^{k_1} \cdots \partial t_n^{k_n}} \varphi_{\mathbf{X}}(\mathbf{t}) \mid_{(t_1,\cdots,t_n)=(0,\cdots,0)} \right]$$

siempre y cuando el momento sea finito.

No se presentan las demostraciones de estos resultados, pues van más allá de los objetivos propuestos para este texto. El lector interesado puede ver [Her].

5.7. Distribución normal multivariada.

Al igual que ocurre en el caso univariado, en el que la distribución normal desempeña un papel destacado, se tiene que la distribución que se define a continuación es fundamental en el desarrollo de la teoría estadística multivariada.

Definición 5.104 (distribución normal multivariada) *Se dice que el vector aleatorio n dimensional* $\mathbf{X} = (X_1, \cdots, X_n)$ *tiene distribución normal multivariada si toda combinación lineal,* $\sum_{j=1}^{n} \alpha_j X_j$, *tiene distribución normal univariada, (posiblemente degenerada, como ocurre, por ejemplo, cuando* $\alpha_j = 0$ *para todo* j*).*

Ejemplo 5.105 *Supóngase que* $X_1, \cdots, X_n$ *son* n *variables aleatorias independientes tales que* $X_j \stackrel{d}{=} \mathcal{N}(\mu_j, \sigma_j^2)$ *para* $j = 1, \cdots, n$. *Entonces, si* $Y = \sum_{j=1}^{n} \alpha_j X_j$, *se tiene:*

$$\begin{aligned}
\varphi_Y(t) &= E(e^{itY}) \\
&= E\left(e^{it[\alpha_1 X_1 + \cdots + \alpha_n X_n]}\right) \\
&= \prod_{j=1}^{n} \Phi_{X_j}(\alpha_j t) \\
&= \prod_{j=1}^{n} \exp\left[i\mu_j \alpha_j t - \frac{\alpha_j^2 t^2 \sigma_j^2}{2}\right] \\
&= \exp\left(\sum_{j=1}^{n}\left[i\mu_j \alpha_j t - \frac{\alpha_j^2 t^2 \sigma_j^2}{2}\right]\right) \\
&= \exp\left(i\mu t - \frac{1}{2}\sigma^2 t^2\right)
\end{aligned}$$

donde

$$\mu := \sum_{j=1}^{n} \mu_j \alpha_j \qquad y \quad \sigma^2 := \sum_{j=1}^{n} \alpha_j^2 \sigma_j^2$$

esto es, $Y \stackrel{d}{=} \mathcal{N}(\mu, \sigma^2)$.

Por lo tanto, el vector $\mathbf{X} := (X_1, \cdots, X_n)$ *tiene distribución normal multivariada.* ▲

Teorema 5.106 *Sea* $\mathbf{X} := (X_1, \cdots, X_n)$ *un vector aleatorio.* $\mathbf{X}$ *tiene distribución normal multivariada, si y sólo si, su función característica es de la forma:*

$$\varphi_{\mathbf{X}}(\mathbf{t}) = \exp\left[i\langle \mathbf{t}, \mu\rangle - \frac{1}{2}\langle \mathbf{t}, \mathbf{t}Q\rangle\right]$$

donde $\mu \in \mathbb{R}^n$, Q *es una matriz cuadrada, simétrica y positiva semidefinida y donde* $\langle \cdot, \cdot \rangle$ *denota el producto interno usual de* $\mathbb{R}^n$.

Demostración. $\Longrightarrow$) *Sea* $Y := \sum_{j=1}^{n} \alpha_j X_j$. *Es claro que:*

$$\begin{aligned} EY &= \sum_{j=1}^{n} \alpha_j \mu_j \\ &= \langle \alpha, \mu \rangle \,; \text{ donde } \mu := E\mathbf{X} \text{ y } \alpha := (\alpha_1, \cdots, \alpha_n) \end{aligned}$$

y

$$\begin{aligned} VarY &= Var\left(\sum_{j=1}^{n} \alpha_j X_j\right) \\ &= \sum_{j=1}^{n} Var(\alpha_j X_j) + 2\sum_{j<i}\sum Cov(\alpha_j X_j, \alpha_i X_i) \\ &= \sum_{j=1}^{n} \alpha_j^2 Var(X_j) + 2\sum_{j<i}\sum \alpha_j \alpha_i Cov(X_j, X_i) \\ &= \langle \alpha, \alpha Q \rangle \end{aligned}$$

donde $\alpha := (\alpha_1, \cdots, \alpha_n)$ *y* Q *es la matriz de varianzas y covarianzas de* $\mathbf{X}$.

Por lo tanto, la función característica de Y *es igual a:*

$$\varphi_Y(t) = \exp\left[it \langle \alpha, \mu \rangle - \frac{t^2}{2} \langle \alpha, \alpha Q \rangle\right].$$

Entonces,

$$\begin{aligned} \varphi_{\mathbf{X}}(\alpha) &= E\left[\exp i\mathbf{X}\alpha^T\right] \\ &= E\left[\exp\left(i\sum_{j=1}^{n} \alpha_j X_j\right)\right] \\ &= \varphi_Y(1) \\ &= \exp\left[i \langle \alpha, \mu \rangle - \frac{1}{2} \langle \alpha, \alpha Q \rangle\right] \end{aligned}$$

$\Longleftarrow$) *Sea* $Y := \sum_{j=1}^{n} \alpha_j X_j$. *La función característica de* Y *está dada por:*

$$\begin{aligned}
\varphi_Y(t) &= E(\exp itY) \\
&= E(\exp it\sum_{j=1}^{n}\alpha_j X_j) \\
&= E(\exp i\mathbf{X}\beta^T)
\end{aligned}$$

donde $\beta := t\alpha$. *Esto es,*

$$\begin{aligned}
\varphi_Y(t) = \varphi_{\mathbf{X}}(\beta) &= \varphi_{\mathbf{X}}(t\alpha) \\
&= \exp\left[i\left\langle t\alpha,\mu\right\rangle - \frac{1}{2}\left\langle t\alpha, t\alpha Q\right\rangle\right] \\
&= \exp\left[it\left\langle \alpha,\mu\right\rangle - \frac{1}{2}t^2\left\langle \alpha, \alpha Q\right\rangle\right]
\end{aligned}$$

Entonces, Y *tiene una distribución normal univariada de parámetros* $\langle\alpha,\mu\rangle$ *y* $\langle\alpha,\alpha Q\rangle$. *Por lo tanto,* $\mathbf{X}$ *es normal multivariada.* ■

Nota 5.107 *Es fácil verificar que* μ *es el vector de medias y que* Q *es la matriz de varianzas y covarianzas de* $\mathbf{X}$.

Notación 5.108 *Se escribe* $\mathbf{X} \stackrel{d}{=} \mathcal{N}(\mu, Q)$ *para indicar que* $\mathbf{X}$ *tiene distribución normal multivariada de vector de medias* μ *y matriz de varianzas y covarianzas* Q.

El teorema siguiente afirma que toda distribución normal multivariada se obtiene como una transformación lineal de vectores de variables aleatorias independientes normales univariadas. Para su demostración se requiere del lema siguiente.

Lema 5.109 *Sea* $\mathbf{X} := (X_1, \cdots, X_n)$ *un vector aleatorio tal que* $\mathbf{X} \stackrel{d}{=} \mathcal{N}(\mu, Q)$. *Las componentes;* X_j , $j = 1, \cdots, n$, *son independientes, si y sólo si, la matriz* Q *es diagonal.*

Demostración. $\Longrightarrow$) *Es el resultado dado en 5.91.*

$\Longleftarrow$) *Supóngase que la matriz* Q *es diagonal. Puesto que*

$\mathbf{X} \overset{d}{=} \mathcal{N}(\mu, Q)$, *entonces,*

$$\begin{aligned}
\varphi_{\mathbf{X}}(\mathbf{t}) &= \exp\left[i\langle \mathbf{t}, \mu\rangle - \frac{1}{2}\langle \mathbf{t}, \mathbf{t}Q\rangle\right] \\
&= \exp\left[i\sum_{j=1}^{n}\mu_j t_j - \frac{1}{2}\sum_{j=1}^{n}\sigma_j t_j^2\right] \\
&= \exp\left[i\sum_{j=1}^{n}\left(\mu_j t_j - \frac{1}{2}\sigma_j t_j^2\right)\right] \\
&= \prod_{j=1}^{n}\exp\left[i\left(\mu_j t_j - \frac{1}{2}\sigma_j t_j^2\right)\right] \\
&= \prod_{j=1}^{n}\varphi_{X_j}(t)
\end{aligned}$$

Por lo tanto, las variables aleatorias X_j; $j = 1, \cdots, n$, *son independientes.*
■

Teorema 5.110 *Sea* $\mathbf{X} := (X_1, \cdots, X_n)$ *un vector aleatorio tal que* $\mathbf{X} \overset{d}{=} \mathcal{N}(\mu, Q)$. *Entonces, existen variables aleatorias independientes* $Y_1, \cdots, Y_n$ *tales que o bien* $Y_j = 0$ *o* $Y_j \overset{d}{=} \mathcal{N}(0, \lambda_j)$; *para* $j = 1, \cdots, n$ *y una matriz ortogonal* A *de tal manera que* $\mathbf{X} = \mu + \mathbf{YA}$.
Demostración. *Puesto que* Q *es una matriz simétrica positiva semidefinida se tiene que existe una matriz diagonal* Λ, *cuyas componentes son todas no negativas, y una matriz ortogonal* A *tales que:*

$$Q = A\Lambda A^T$$

Sea $\mathbf{Y} := (\mathbf{X} - \mu)A^T$. *Puesto que* $\mathbf{X}$ *es normal multivariada, entonces,* $\mathbf{Y}$ *lo es también. Además,* Λ *es la matriz de varianzas y covarianzas de* $\mathbf{Y}$. *Como ésta matriz es diagonal se sigue, del lema anterior, que las componentes de* $\mathbf{Y}$ *son independientes. Por último se tiene que:*

$$\mathbf{X} = \mu + \mathbf{YA}.$$

■

Supóngase que $\mathbf{X} = (X_1, \cdots, X_n)$ es un vector aleatorio n dimensional tal que las n variables aleatorias, $X_1, \cdots, X_n$, son independientes e igualmente distribuidas con distribución normal estándar. La función de densidad de probabilidad conjunta de $X_1, \cdots, X_n$ está dada por:

$$\begin{aligned} f(x_1, \cdots, x_n) &= f_{X_1}(x_1) \cdots f_{X_n}(x_n) \\ &= \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_1^2\right) \cdots \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x_n^2\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\sum_{j=1}^{n} x_j^2\right) \\ &= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\mathbf{x}\mathbf{x}^T\right) ; \text{ donde } \mathbf{x} :=(x_1, \cdots, x_n). \end{aligned}$$

Por otra parte, es claro que el vector $\mathbf{X}$ tiene distribución normal multivariada. Surge pues, de manera natural, la pregunta siguiente: Si $\mathbf{X}$ es un vector normal multivariado, ¿bajo qué condiciones se garantiza la existencia de una función de densidad del vector $\mathbf{X}$? La respuesta a esta pregunta la ofrece el teorema siguiente:

Teorema 5.111 *Sea $\mathbf{X} \stackrel{d}{=} \mathcal{N}(\mu, Q)$. Si Q es positiva definida entonces $\mathbf{X}$ posee una función de densidad dada por:*

$$f(\mathbf{x}) = \frac{1}{(2\pi)^{\frac{n}{2}}} (\det Q)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{x} - \mu)Q^{-1}(\mathbf{x} - \mu)^T\right].$$

Demostración. *Como la matriz Q es positiva definida, todos los valores propios de Q son positivos. Además, existe una matriz ortogonal U tal que:*

$$UQU^T = \Lambda$$

donde $\Lambda = diag(\lambda_i)$ siendo $\lambda_1, \cdots, \lambda_n$ los valores propios de Q, esto es, Λ es la matriz diagonal cuyos elementos diagonales son los valores propios de Q.

Sea $A := Udiag(\sqrt{\lambda_i})U^T$. Es claro que $A^TA = Q$ y que A es también positiva definida.

Sea $\mathbf{h} : \mathbb{R}^{1\times n} \longrightarrow \mathbb{R}^{1\times n}$, dada por $\mathbf{h}(\mathbf{x}) = \mathbf{x}A + \mu$. La función inversa de $\mathbf{h}$ está dada por $\mathbf{h}^{-1}(\mathbf{x}) = (\mathbf{x} - \mu)A^{-1}$. Por el teorema de transformación se tiene que la función de densidad de $\mathbf{X} := \mathbf{Y}A + \mu$, donde $\mathbf{Y} = (Y_1, \cdots, Y_n)$ es un vector aleatorio n dimensional tal que las variables aleatorias $Y_1, \cdots, Y_n$ son independientes e igualmente distribuidas con *distribución normal estándar, está dada por:*

$$
\begin{aligned}
f_{\mathbf{X}}(\mathbf{x}) &= f_{\mathbf{Y}}(h^{-1}(\mathbf{x})).\left|\det \frac{\partial h^{-1}}{\partial \mathbf{x}}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}[h^{-1}(\mathbf{x})][h^{-1}(\mathbf{x})]^T\right)\left|\det \frac{\partial h^{-1}}{\partial \mathbf{x}}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\left[(\mathbf{x}-\mu)A^{-1}\right]\left[(\mathbf{x}-\mu)A^{-1}\right]^T\right)\left|\det A^{-1}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\left[(\mathbf{x}-\mu)A^{-1}(A^{-1})^T(\mathbf{x}-\mu)^T\right]\right)\left|\det A^{-1}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\left[(\mathbf{x}-\mu)A^{-1}(A^T)^{-1}(\mathbf{x}-\mu)^T\right]\right)\left|\det A^{-1}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\left[(\mathbf{x}-\mu)\left(A^TA\right)^{-1}(\mathbf{x}-\mu)^T\right]\right)\left|\det A^{-1}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{1}{2}\left[(\mathbf{x}-\mu)Q^{-1}(\mathbf{x}-\mu)^T\right]\right)\left|\det A^{-1}\right| \\
&= \frac{1}{(2\pi)^{\frac{n}{2}}} (\det Q)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)Q^{-1}(\mathbf{x}-\mu)^T\right]
\end{aligned}
$$

■

Nota 5.112 (distribución normal bivariada) *Como caso especial, del teorema anterior, supóngase que*

$$\mathbf{X} = (X_1, X_2) \stackrel{d}{=} \mathcal{N}(\mu, Q)\,; \quad donde \;\; \mu = (\mu_1, \mu_2)$$

y

$$Q = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{pmatrix}; \quad donde \;\; \mu_1 := EX_1, \;\; \mu_2 := EX_2.$$

$$\sigma_1^2 := Var\,(X_1)\,, \;\; \sigma_2^2 := Var\,(X_2)\,, \;\; \sigma_{12} = Cov\,(X_1, X_2) = \sigma_{21} = \rho\sigma_1\sigma_2$$

siendo ρ el coeficiente de correlación.

Por lo tanto,

$$f(\mathbf{x}) = \frac{1}{(2\pi)} (\det Q)^{-\frac{1}{2}} \exp\left[-\frac{1}{2}(\mathbf{x}-\mu)Q^{-1}(\mathbf{x}-\mu)^T\right].$$

Puesto que

$$\det Q = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2(1-\rho^2)$$

y

$$Q^{-1} = \frac{1}{\sigma_1^2\sigma_2^2(1-\rho^2)}\begin{pmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{21} & \sigma_1^2 \end{pmatrix}$$

se obtiene:

$$f(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 \right.\right.$$
$$\left.\left. -2\rho\left(\frac{(x_1-\mu_1)(x_2-\mu_2)}{\sigma_1\sigma_2}\right)+\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right\}\right].$$

Se observa además que:

$$\begin{aligned} f_{X_1}(x_1) &= \int_{-\infty}^{\infty} f(x_1,x_2)dx_2 \\ &= \frac{1}{\sqrt{2\pi}\sigma_1}\exp\left[-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2\right] \end{aligned}$$

y

$$\begin{aligned} f_{X_2}(x_2) &= \int_{-\infty}^{\infty} f(x_1,x_2)dx_1 \\ &= \frac{1}{\sqrt{2\pi}\sigma_2}\exp\left[-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]. \end{aligned}$$

Esto es, las distribuciones marginales de $\mathbf{X} = (X_1, X_2)$ *son distribuciones normales univariadas. En general se tiene que:*

Teorema 5.113 *Todas las distribuciones marginales de* $\mathbf{X} \stackrel{d}{=} \mathcal{N}(\mu, Q)$ *son normales multivariadas.*
Demostración. *Supóngase que* $\mathbf{X} = (X_1, \cdots, X_n)$ *y sea*

$$\widetilde{\mathbf{X}} := (X_{k_1}, \cdots, X_{k_l})$$

donde $\{k_1, \cdots, k_l\}$ *es un subconjunto de* $\{1, \cdots, n\}$.

La función característica de $\widetilde{\mathbf{X}}$ *está dada por:*

$$\begin{aligned} \varphi_{\widetilde{\mathbf{X}}}(t_{k_1}, \cdots, t_{k_l}) &= E(i\sum_{r=1}^{l} t_{k_r} X_{t_{k_r}}) \\ &= \varphi_{\mathbf{X}}(t_1, \cdots, t_n); \ \ donde\ t_j = 0\ si\ j \notin \{k_1, \cdots, k_l\}. \end{aligned}$$

Luego $\widetilde{\mathbf{X}}$ *tiene una distribución normal multivariada.* ■

5.8. Ejercicios

1. Supóngase que la distribución conjunta de las variables aleatorias X e Y está dada por:

$X\backslash Y$	1	2	3	4
0	0.1	0	0	0
-1	0.1	0.1	0	0
-2	0.1	0.1	0.1	0
-3	0.1	0.1	0.1	0.1

   Calcular:

   a) $P(X \geq -2, Y \geq 2)$

   b) $P(X \geq -2 \ \vee Y \geq 2)$

   c) Las distribuciones marginales de X y Y.

   d) La distribución de $Z := X + Y$

2. Supóngase que X e Y son variables aleatorias con distribución conjunta dada por:

$X\backslash Y$	0	1	2
1	0.2	α	β
2	γ	0.1	δ
3	η	κ	0.3

   Determinar valores para $\alpha, \beta, \gamma, \delta, \eta$ y κ de tal manera que:

   $P(X = 1) = 0.4,\ P(X = 2) = 0.3,\ P(Y = 0) = 0.2\, y\, P(Y = 2) = 0.6$

3. En una urna se encuentran 4 bolas rojas, 5 negras y 2 azules. Se extraen al azar (sin repetición) dos bolas de la urna. Sea X la variable aleatoria que denota el número de bolas rojas extraídas y Y la variable aleatoria que denota el número de bolas negras extraídas. Hallar:

   a) La distribución conjunta de X e Y

   b) EX y EY

4. Realizar el ejercicio anterior bajo el supuesto de que la extracción se hace con sustitución.

5. Sean X, Y y Z variables aleatorias con distribución conjunta dada por:

$\mathbf{x} = (x, y, z)$	$(1,2,3)$	$(2,1,3)$	$(3,2,1)$	$(1,3,2)$	$(1,2,3)$	$(2,3,1)$
$P(\mathbf{X} = \mathbf{x})$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Calcular:

a) $E(X + Y + Z)$.

b) $E(XYZ)$.

6. (Distribución multinomial) Supóngase que hay $(k+1)$ resultados distintos en un experimento aleatorio con probabilidades de ocurrencia $p_1, \cdots, p_{k+1}$ donde $\sum_{i=1}^{k+1} p_i = 1$. Sea X_i el número de veces que se obtiene el i-ésimo resultado en n repeticiones independientes del experimento. Verificar que la función de densidad conjunta de las variables aleatorias $X_1, \cdots, X_{k+1}$ está dada por:

$$f(x_1, \cdots, x_{k+1}) = \frac{n!}{\prod_{i=1}^{k+1} x_i!} \prod_{i=1}^{k+1} p_i^{x_i}$$

donde $x_i = 0, \cdots, n$ y $\sum_{i=1}^{n} x_i = n$.

7. Las cirugías que se programan, para un mes, en un hospital estatal son clasificadas en 4 categorías de prioridad en su realización: urgente, prioridad normal, baja prioridad y espera. Las directivas del hospital estiman que el 10 % de las cirugías que se programan son clasificadas en la primera categoría, el 50 % en la segunda categoría, el 30 % en la tercera y el restante 10 % en la cuarta. Supóngase que en un mes se programan 30 cirugías. Calcular:

a) La probabilidad de que 5 de las cirugías sean clasificadas en la primera categoría, 15 en la segunda, 7 en la tercera y 3 en la cuarta.

b) El número esperado de cirugías clasificadas en la tercera categoria.

8. Sean X e Y variables aleatorias discretas independientes con distribución conjunta dada por:

$X \backslash Y$	-1	0	1
0	$\frac{1}{6}$	d	$\frac{1}{6}$
1	a	e	k
2	b	f	h

Si $P(X = 1) = P(X = 2) = \frac{1}{3}$, determinar los valores de los datos que faltan en la tabla. ¿ A qué es igual EXY ?

9. En una pasteleria se venden pasteles Gloria a razón promedio de 1.3 por hora, tortas de chocolate a razón promedio de 0.6 por hora y roscones de arequipe a razón promedio de 2.8 por hora. Supóngase que las cantidades vendidas, de cada producto, son independientes y que cada una tiene una distribución Poisson. Calcular:

   a) La distribución del número total de pasteles Gloria, tortas de chocolate y roscones de arequipe vendidos en 2 horas.

   b) La probabilidad de vender por lo menos dos de los productos en un período de 15 minutos.

10. Sean X e Y variables aleatorias con distribución conjunta dada por:

$X \backslash Y$	-1	0	1
0	$\frac{1}{16}$	$\frac{1}{4}$	0
1	$\frac{3}{16}$	$\frac{1}{8}$	$\frac{3}{16}$
2	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$

Hallar la distribución conjunta de $Z := X + Y$ y $W := X - Y$.

11. Sean X e Y variables aleatorias discretas con distribución conjunta dada por:

$X \backslash Y$	-1	0	2
1	$\frac{1}{18}$	$\frac{3}{18}$	$\frac{1}{9}$
2	$\frac{1}{9}$	0	$\frac{1}{6}$
3	0	$\frac{2}{9}$	$\frac{1}{6}$

Calcular $Cov(X,Y)$ y $\rho(X,Y)$.

12. Calcular los valores esperados de las variables aleatorias dadas en los ejercicios 25 y 26 del capítulo 2.

13. Una urna contiene 3 bolas rojas y 2 negras. Se extrae una muestra aleatoria de tamaño 2, sin reemplazo. Sea X el número de bolas rojas seleccionadas y Y el número de bolas negras seleccionadas. Calcular $\rho(X, Y)$.

14. Sean X e Y variables aleatorias independientes con $X \stackrel{d}{=} \mathcal{B}\left(3, \frac{1}{3}\right)$ y $Y \stackrel{d}{=} \mathcal{B}\left(2, \frac{1}{2}\right)$. Calcular $P\left(X = Y\right)$.

15. Sean X e Y variables aleatorias independientes con distribución Bernoulli de parámetro $\frac{1}{2}$. ¿Son $Z := X + Y$ y $W := |X - Y|$ independientes? Explicar.

16. Sean X_1, X_2 y X_3 variables aleatorias independientes con varianzas positivas y finitas σ_1^2, σ_2^2 y σ_3^2 respectivamente.

    Calcular el coeficiente de correlación entre $X_1 - X_2$ y $X_2 + X_3$.

17. Sean X e Y dos variables aleatorias tales que $\rho(X, Y) = \frac{1}{2}$, $VarX = 1$ y $VarY = 2$. Calcular $Var(X - 2Y)$

18. Sean X e Y variables aleatorias independientes. Supóngase que tanto X como Y toman los valores 1 y -1 con probabilidad $\frac{1}{2}$ cada uno. Sea $Z := XY$. ¿Son X, Y y Z dos a dos independientes?, ¿son X, Y y Z independientes? Explicar.

19. Se venden m de $n \geq 2$ billetes de loteria. Supóngase que los billetes están marcados del 1 al n y que cada billete tiene el mismo "chance" de ser vendido. Calcular el valor esperado y la varianza de la variable aleatoria que representa la suma de los números de los billetes vendidos.

20. ¿A qué es igual el número esperado de días del año, para los cuales se satisface que exactamente k de r personas celebran su cumpleaños? Supóngase que cada uno de los 365 días del año tiene la misma probabilidad de ser un día de cumpleaños, además ignorése la existencia de años bisiestos.

21. Bajo los mismos supuestos dados en el problema 20, ¿a qué es igual el número esperado de días del año en los que hay más de un cumpleaños? Verificar, con ayuda de una calculadora, que éste número esperado es, para todo $r \geq 29$, mayor o igual a 1.

22. Sean X e Y variables aleatorias con media 0, varianza 1 y correlación ρ. Demostrar que:

$$E\left(\text{máx}\left\{X^2, Y^2\right\}\right) \leq 1 + \sqrt{1-\rho^2}$$

Sugerencia: $\max\{u,v\} = \frac{1}{2}(u+v) + \frac{1}{2}|u-v|$. Usar la desigualdad de Cauchy-Schwarz.

23. Sean X e Y variables aleatorias con media 0, varianza 1 y covarianza ρ.

    a) Demostrar que las variables aleatorias $Z := X - \rho Y$ y Y son no correlacionadas.

    b) Calcular EZ y $Var\,(Z)$.

24. Sean X , Y y Z variables aleatorias con media 0, varianza 1. Sea $\rho_1 := \rho(X,Y)$, $\rho_2 := \rho(Y,Z)$ y $\rho_3 := \rho(X,Z)$. Demostrar que:

$$\rho_3 \geq \rho_1\rho_2 - \sqrt{1-\rho_1^2}\sqrt{1-\rho_2^2}$$

Sugerencia:

$$XZ = \left[\rho_1 Y + (X - \rho_1 Y)\right]\left[\rho_2 Y + (Z - \rho_2 Y)\right]$$

25. Sean $X \stackrel{d}{=} \mathcal{U}(0,1)$ e $Y = X^2$.

    a) Hallar $\rho(X,Y)$.

    b) ¿Son X e Y independientes? Explicar.

26. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} c & \text{si} \quad 0 < x < 4,\ 0 < y \text{ y } (x-1) < y < (x+1) \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

    a) Determinar el valor de c.

    b) Calcular $P\left(X < \frac{1}{2}, Y < \frac{1}{2}\right)$ y $P\left(X < \frac{1}{2}\right)$.

    c) Determinar EX y EY.

27. Diez clientes entre los que se encuentran Juan y María llegan a un almacén entre las 8 : 00 am y el mediodía. Suponiendo que los clientes llegan independientemente unos de otros y que el tiempo de llegada de cada uno de ellos es una variable aleatoria con distribución uniforme sobre el intervalo $[0,4]$ .Calcular:

    a) La probabilidad de que Juan llegue antes de las 11 : 00 am.

    b) La probabilidad de que Juan y María lleguen ambos antes de las 11 : 00 am.

    c) El número esperado de clientes que llegan antes de las 11 : 00 am

28. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} \exp\left[-(x+y)\right] & \text{si} \quad x>0,\ y>0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

    Calcular:

    a) $P(1 < X+Y < 2)$.

    b) $P(X < Y \mid X > 2Y)$.

    c) $P(X > 1)$.

29. Sean X e Y variables aleatorias con función de distribución acumulativa conjunta dada por:

$$F(x,y) = \begin{cases} (1-\exp(-x))\left(\frac{1}{2} + \frac{1}{\pi}\tan^{-1} y\right) & \text{si} \quad x>0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

    ¿Existe una función de densidad conjunta de X y Y ?. Explicar.

30. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} c\sin(x+y) & \text{si} \quad 0 \le x, y \le \frac{\pi}{2} \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

    Calcular:

    a) El valor de c.

    b) Las funciones de densidad marginales de X y de Y.

31. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} \frac{4x}{y^3} & \text{si} \quad 0 < x < 1,\ 1 < y \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Calcular:

   a) $P\left(\frac{1}{2} < X \leq \frac{3}{4},\ 0 < Y \leq \frac{1}{3}\right)$.
   b) $P(Y > 5)$.

32. Alberto y Sandra han acordado reunirse entre las $7:00$ pm y las $8:00$ pm en un restaurante del centro de la ciudad. Sea X la variable aleatoria que denota el tiempo de llegada de Alberto e Y la variable aleatoria que denota el tiempo de llegada de Sandra. Supóngase que X e Y son independientes e igualmente distribuidas con distribución uniforme sobre el intervalo $[7,8]$.

   a) Determinar la función de densidad conjunta de X e Y.
   b) Calcular la probabilidad de que ambos, Alberto y Sandra, lleguen al restaurante entre las $7:15$ pm y las $8:15$ pm.
   c) Si el primero en llegar espera sólo 10 minutos antes de irse a comer a otra parte, ¿cuál es la probabilidad de que tanto Sandra como Alberto coman en el restaurante inicialmente elegido?

33. La variable aleatoria bidimensional (X,Y) tiene la siguiente función de densidad de probabilidad conjunta:

$$f(t_1,t_2) = \frac{1}{16\pi} e^{-\frac{1}{2}\left(\frac{(t_1-3)^2}{4}+\frac{(t_2+2)^2}{16}\right)}; \quad -\infty < t_1, t_2 < \infty.$$

Calcular:

   a) Las funciones de densidad marginales f_X, f_Y.
   b) EX y EY.
   c) $VarX$ y $VarY$.
   d) $Cov(X,Y)$ y $\rho(X,Y)$.

34. Sea (X, Y) las coordenadas de un punto escogido al azar dentro de un círculo unitario. Esto es, X e Y son variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x, y) = \begin{cases} \frac{1}{\pi} & \text{si} \quad 0 < x^2 + y^2 \leq 1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Calcular $P(X < Y)$.

35. Un punto Q de coordenadas (X, Y) es escogido aleatoriamente en el cuadrado $[0, 1] \times [0, 1]$. Calcular la probabilidad de que Q esté más cerca de $(0, 0)$ que de $\left(\frac{1}{2}, \frac{1}{2}\right)$.

36. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x, y) = \begin{cases} 3x & \text{si} \quad 0 \leq y \leq x \leq 1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Hallar la función de densidad de la variable aleatoria $Z := X - Y$.

37. Sean X e Y variables aleatorias independientes con funciones de densidad dadas por:

$$f_X(x) = \mathcal{X}_{(1,2)}(x) \quad \text{y} \quad f_Y(x) = \frac{1}{2}\mathcal{X}_{(3,5)}(x).$$

Determinar una función de densidad de $Z := X + Y$.

38. Sean X e Y variables aleatorias independientes con $X \stackrel{d}{=} \mathcal{U}(0, 1)$ y $Y \stackrel{d}{=} \mathcal{U}(0, 2)$. Calcular:

   a) $P([X + Y] > 1)$.
   b) $P\left(Y \leq X^2 + 1\right)$.

39. Sean X e Y variables aleatorias independientes. Supóngase que X tiene una distribución exponencial de parámetro λ y que Y tiene densidad $f(x) = 2x\mathcal{X}_{(0,1)}(x)$.

   a) Determinar una densidad de la variable aleatoria $Z := X + Y$.
   b) Calcular la $Cov(X, Z)$.

40. Sean X e Y variables aleatorias independientes e igualmente distribuidas con distribución uniforme sobre el intervalo $(0, 1)$. Calcular:

a) $P\left(|X-Y| \leq \frac{1}{2}\right)$.
b) $P\left(\left|\frac{X}{Y}-1\right| \leq \frac{1}{2}\right)$.
c) $P\left(Y \geq X \mid Y \geq \frac{1}{2}\right)$.

41. Demostrar o refutar: Si X e Y son variables aleatorias tales que la función característica de $(X+Y)$ es igual al producto de las funciones características de X e Y, entonces, X e Y son independientes. Justificar la respuesta.

42. Sean X e Y variables aleatorias independientes e igualmente distribuidas con distribución exponencial de parámetro λ. Determinar funciones de densidad para las siguientes variables aleatorias:

 a) $Z := |X-Y|$
 b) $W := \text{mín}\left\{X, Y^3\right\}$

43. Sean X ,Y y Z variables aleatorias independientes e igualmente distribuidas con distribución uniforme sobre el intervalo $[0,1]$.

 a) Hallar la función de densidad conjunta de $W := XY$ y $V := Z^2$.
 b) Calcular $P(V \leq W)$.

44. Sean X e Y variables aleatorias independientes e igualmente distribuidas con distribución normal estándar. Sean $Y_1 = X+Y$ y $Y_2 = X/Y$. Hallar la función de densidad de probabilidad conjunta de las variables aleatorias Y_1 y Y_2, ¿qué tipo de distribución tiene Y_2?

45. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} 4xy & \text{si} \quad 0<x<1 \text{ y } 0<y<1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Encontrar la función de densidad de probabilidad conjunta de X^2 e Y^2.

46. Sean X e Y dos variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} 9x^2y^2 & \text{si} \quad 0<x<1 \text{ y } 0<y<1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Sean $Y_1 = X^3$ y $Y_2 = Y^3$. Hallar $f_{(Y_1,Y_2)}(y_1,y_2)$.

47. Supóngase que X, Y, Z son variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y,z) = \begin{cases} e^{-(x+y+z)} & \text{si} \quad x>0, y>0, z>0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Hallar f_U, donde $U := \frac{X+Y+Z}{3}$.

48. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} \frac{3x}{2} & \text{si} \quad 0<x<1 \text{ y } -x<y<x \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Encontrar f_Z, donde $Z = X - Y$.

49. Sean $X_1, X_2, ..., X_n$ variables aleatorias *i.i.d* con $\mathcal{N}(\mu, \sigma^2)$. ¿Qué tipo de distribución tienen las variables aleatorias $Z_i := \frac{X_i - \mu}{\sigma}$; $i = 1, 2, ..., n$?, ¿cuál es la distribución de $Z_1^2 + ... + Z_n^2$?

50. Sean $X_1, X_2, ..., X_n$ variables aleatorias *i.i.d* con $\mathcal{N}(\mu, \sigma^2)$. Sea

$$S^2 := \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2.$$

¿Qué tipo de distribución tiene $\frac{(n-1)S^2}{\sigma^2}$? Explicar.

51. La resistencia a la tensión para cierto tipo de alambre se distribuye normalmente con una media desconocida μ y una varianza desconocida σ^2. Se seleccionaron al azar seis segmentos de alambre de un rollo grande y se midió $Y_i :=$ " la resistencia a la tensión para el segmento i"; $i = 1, 2, .., 6$. La media de la población μ y la varianza σ^2 se pueden estimar por $\overline{Y}$ y S^2 respectivamente. Encontrar la probabilidad de que $\overline{Y}$ esté a lo más a $\frac{2S}{\sqrt{n}}$ de la verdadera media poblacional μ.

52. Supóngase que $X \stackrel{d}{=} F_n^m$. Encontrar el valor de x para el cual:

a) $P(X \leq x\ ) = 0.99$ con $m = 7, n = 3$.

b) $P(X \leq x) = 0.005$ con $m = 20$, $n = 30$

c) $P(X \leq x) = 0.95$ con $m = 2$, $n = 9$.

53. Si $X \stackrel{d}{=} t_n$, ¿Qué tipo de distribución tiene X^2?

54. Demostrar que si $X \overset{d}{=} F_n^m$, entonces, $EX = \frac{n}{n-2}$; para $n > 2$ y que $VarX = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$.

    Sugerencia: Supóngase que $X = \frac{U/m}{V/n}$, donde U y V son independientes y $U \overset{d}{=} \mathcal{X}_m^2$ y $V \overset{d}{=} \mathcal{X}_n^2$

55. Sea X una variable aleatoria con distribución $t-$ Student con k grados de libertad. Calcular EX y $VarX$.

56. Sea X una variable aleatoria con distribución normal estándar. ¿Son X y $|X|$ independientes?, ¿son no correlacionadas?. Explicar.

57. Sean $X_1, X_2, ..., X_n$ variables aleatorias $i.i.d$ con $\mathcal{N}(\mu, \sigma^2)$.¿ Qué tipo de distribución tiene la variable aleatoria siguiente

$$\frac{\sqrt{n(n-1)}(\overline{X}-\mu)}{\sqrt{\sum_{i=1}^{n}(X_i-\overline{X})^2}}?$$

    Explicar.

58. Sean X e Y variables aleatorias independientes con funciones generadoras de momentos dadas por:

$$m_X(t) = \exp\left(2e^t - 2\right) \quad \text{y} \quad m_Y(t) = \frac{1}{5}\left(1 + 2e^{-t} + 2e^t\right).$$

    Calcular:

    a) $P([X+Y] = 2)$.
    b) EXY.

59. Sea $\mathbf{X}$ un vector aleatorio n-dimensional con matriz de varianzas y covarianzas $\sum$. Sea A una matriz cuadrada de orden n no singular y sea $\mathbf{Y} := \mathbf{X}A$.

    a) Demostrar que $E(\mathbf{Y}) = E(\mathbf{X})A$
    b) Calcular la matriz de varianzas y covarianzas de $\mathbf{Y}$.

60. Sea $\mathbf{X} = (X, Y)$ un vector aleatorio con distribución normal bivariada con $\mu_X = 3$, $\mu_Y = 7.7$, $\sigma_X = 0.04$, $\sigma_Y = 0.08$ y $\rho = 0$. Calcular:

$$P(2.95 < X < 3.05,\ 7.60 < Y < 7.80)$$

61. Sea $\mathbf{X} = (X_1, \cdots, X_n)$ un vector aleatorio n dimensional con $\mathbf{X} \stackrel{d}{=} \mathcal{N}(\mu, Q)$, siendo Q una matriz no singular. Demostrar que

$$\mathbf{Y} := (\mathbf{X} - \mu) W^{-1}$$

es un vector aleatorio con distribución $\mathcal{N}(\mathbf{0}, I)$, donde I es la matriz identidad de orden n y W es una matriz tal que $W^2 = Q$. En tal caso se dice que el vector $\mathbf{Y}$ tiene distribución normal multivariada estándar.

62. Sea $\mathbf{X} = (X_1, X_2, X_3)$ un vector aleatorio con distribución normal trivariada con $\mu = \mathbf{0}$ y Q dada por:

$$Q = \begin{pmatrix} 1 & \rho_1 & \rho_2 \\ \rho_1 & 1 & \rho_3 \\ \rho_2 & \rho_3 & 1 \end{pmatrix}$$

Calcular $P(X > 0, Y > 0, Z > 0)$.

63. Sea $\mathbf{X} = (X_1, X_2, X_3)$ un vector aleatorio con distribución normal trivariada con $\mu = \mathbf{0}$ y Q dada por:

$$Q = \begin{pmatrix} 5 & 2 & -2 \\ 2 & 6 & 3 \\ -2 & 3 & 8 \end{pmatrix}$$

Hallar la función de densidad $f(x, y, z)$ de $\mathbf{X}$.

Capítulo 6

Probabilidad condicional y esperanza condicional

Uno de los conceptos más importantes y útiles de la teoría de probabilidad es el de esperanza condicional. La razón es doble: en primer lugar, en la práctica interesa frecuentemente calcular probabilidades y valores esperados cuando alguna información parcial está a disposición. Por otra parte, cuando se desea hallar una probabilidad o un valor esperado, muchas veces resulta conveniente condicionar primero con respecto a alguna variable aleatoria apropiada.

6.1. Función de densidad condicional

La relación entre dos variables aleatorias puede verse hallando la distribución condicional de una de ellas, dado el valor de la otra. En el capítulo 1. se definió, para los eventos A y B con probabilidad de B mayor que 0, la probabilidad condicional de A dado B como:

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

Es natural por consiguiente, la definición siguiente:

Definición 6.1 (función de densidad condicional de v.a discretas) *Sean X e Y variables aleatorias discretas. Se define la función de densidad de probabilidad condicional de X dado $Y = y$ como sigue:*

$$f_{X|Y}(x \mid y) := P(X = x \mid Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)};$$

para todos los y *para los cuales* $P(Y = y) > 0$.

Definición 6.2 (función de distribución condicional de v.a discretas) *La función de distribución condicional de* X *dado* $Y = y$ *se define como:*

$$F_{X|Y}(x \mid y) := P(X \leq x \mid Y = y) = \sum_{k \leq x} f_{X|Y}(x \mid y)$$

para todos los y *para los cuales* $P(Y = y) > 0$.

Definición 6.3 (valor esperado condicional de v.a discretas) *El valor esperado condicional de* X *dado* $Y = y$ *se define así:*

$$E(X \mid Y = y) =: \sum_{x} x f_{X|Y}(x \mid y)$$

para todos los y *para los cuales* $P(Y = y) > 0$.

Ejemplo 6.4 *Una urna contiene tres bolas rojas y dos verdes. Se extrae una muestra aleatoria de tamaño 2(con reemplazo). Sean:*

$$X := \begin{cases} 1 & \text{si la primera bola extraída es roja.} \\ 0 & \text{si la primera bola extraída es verde.} \end{cases}$$
$$Y := \begin{cases} 1 & \text{si la segunda bola extraída es roja.} \\ 0 & \text{si la segunda bola extraída es verde.} \end{cases}$$

La distribución de probabilidad conjunta de las variables aleatorias X *e* Y *está dada por:*

$X \backslash Y$	0	1
0	$\frac{4}{25}$	$\frac{6}{25}$
1	$\frac{6}{25}$	$\frac{9}{25}$

.

de esta forma:

$$f_{X|Y}(x \mid 0) = \begin{cases} \frac{2}{5} & si \quad x = 0 \\ \frac{3}{5} & si \quad x = 1 \end{cases}$$

$$f_{X|Y}(x \mid 1) = \begin{cases} \frac{2}{5} & si \quad x = 0 \\ \frac{3}{5} & si \quad x = 1 \end{cases}$$

$$F_{X|Y}(x \mid 0) = \begin{cases} 0 & si \quad x < 0 \\ \frac{2}{5} & si \quad \frac{2}{5} \leq x < 1 \\ 1 & si \quad x \geq 1 \end{cases}$$

y

$$E(X \mid Y = y) = \begin{cases} \frac{3}{5} & si \quad y = 0 \\ \frac{3}{5} & si \quad y = 1. \end{cases} \quad \blacktriangle$$

Ejemplo 6.5 *Sean X e Y variables aleatorias independientes con distribución Poisson de parámetros λ_1 y λ_2 respectivamente. Calcular el valor esperado de X bajo la condición de que $X + Y = n$, con n entero no negativo fijo.*

Solución: Como:

$$f_{X|X+Y}(x \mid n) = P(X = x \mid X + Y = n) = \frac{P(X = x, Y = n - x)}{P(X + Y = n)}$$

$$= \binom{n}{x} \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^x \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{n-x}$$

Esto es, X tiene, bajo la condición $X + Y = n$, una distribución binomial de parámetros n y $p := \frac{\lambda_1}{\lambda_1 + \lambda_2}$. Por lo tanto,

$$E(X \mid X + Y = n) = n\left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right). \quad \blacktriangle$$

Definición 6.6 ***(función de densidad de probabilidad condicional para v.a con fdp conjunta f)*** *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta f. Se define la función de densidad de probabilidad condicional de X dado $Y = y$, como:*

$$f_{X|Y}(x \mid y) := \frac{f(x, y)}{f_Y(y)}$$

para todos los y con $f_Y(y) > 0$.

Definición 6.7 ***(función de distribución condicional para v.a con fdp conjunta f)*** *La función de distribución condicional de X dado $Y = y$, está definida como:*

$$F_{X|Y}(x \mid y) := P(X \leq x \mid Y = y) = \int_{-\infty}^{x} f_{X|Y}(t \mid y)\, dt$$

para todos los y con $f_Y(y) > 0$.

Definición 6.8 ***(valor esperado condicional para v.a con fdp conjunta f)*** *El valor esperado condicional de X dado $Y = y$, se define como sigue:*

$$E(X \mid Y = y) := \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)\, dx$$

para todos los y con $f_Y(y) > 0$.

Ejemplo 6.9 *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:*

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

Para $0 < y < 1$ se tiene que:

$$f_{X|Y}(x \mid y) = \begin{cases} \frac{x+y}{0.5+y} & si \quad 0 < x < 1 \\ 0 & en \quad otro\ caso. \end{cases}$$

$$F_{X|Y}(x \mid y) = \begin{cases} 0 & si \quad x \leq 0 \\ \frac{1}{0.5+y}\left(\frac{x^2}{2} + xy\right) & si \quad 0 \leq x < 1 \\ 1 & si \quad x \geq 1 \end{cases}$$

y

$$E(X \mid Y = y) = \frac{1}{0.5 + y}\left(\frac{1}{3} + \frac{y}{2}\right). \quad \blacktriangle$$

Ejemplo 6.10 *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:*

$$f(x, y) = \begin{cases} e^{-y} & si \quad 0 < x < y \\ 0 & en \quad otro\ caso. \end{cases}$$

Para $y > 0$ se tiene:

$$f_{X|Y}(x \mid y) = \begin{cases} \frac{1}{y} & si \quad 0 < x < y \\ 0 & en \quad otro\ caso. \end{cases}$$

$$F_{X|Y}(x \mid y) = \begin{cases} \frac{x}{y} & si \quad 0 < x < y \\ 1 & si \quad x > 0,\ x > y \\ 0 & en \quad otro\ caso. \end{cases}$$

y

$$E(X \mid Y = y) = \frac{y}{2}. \quad \blacktriangle$$

Ejemplo 6.11 *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:*

$$f(x,y) = \begin{cases} \frac{3}{8}(x + y^2) & si \quad 0 < x < 2\ y\ 0 < y < 1 \\ 0 & en \quad otro\ caso. \end{cases}$$

Calcular $f_{X|Y}(x \mid y)$ y $E(X \mid Y = y)$.

Solución: La función de densidad marginal de Y es igual a:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\ &= \int_0^2 \frac{3}{8}(x + y^2)\mathcal{X}_{(0,1)}(y)dx \\ &= \frac{3}{4}\left(1 + y^2\right)\mathcal{X}_{(0,1)}(y). \end{aligned}$$

luego, para $0 < y < 1$, se tiene que:

$$f_{X|Y}(x \mid y) = \frac{1}{2}\left(\frac{x + y^2}{1 + y^2}\right)\mathcal{X}_{(0,2)}(x)$$

y

$$\begin{aligned} E(X \mid Y = y) &= \int_{-\infty}^{\infty} x f_{X|Y}(x \mid y)dx \\ &= \int_0^2 x\frac{1}{2}\left(\frac{x + y^2}{1 + y^2}\right)dx \\ &= \frac{4 + 3y^2}{3(1 + y^2)}. \quad \blacktriangle \end{aligned}$$

Nota 6.12 *Para todo y; con $f_Y(y) > 0$, y todo conjunto Borel A de $\mathbb{R}$, se tiene:*

$$P(X \in A \mid Y = y) = \int_A f_{X|Y}(x \mid y)dx$$

Ejemplo 6.13 *Si X e Y son variables aleatorias con función de densidad de probabilidad conjunta dada por:*

$$f(x,y) = \frac{e^{-\frac{x}{y}}e^{-y}}{y}\mathcal{X}_{(0,\infty)}(x)\mathcal{X}_{(0,\infty)}(y),$$

entonces, $P(X > 1 \mid Y = y) = e^{-\frac{1}{y}}$; para $y > 0$. ▲

Nota 6.14 *Si X e Y son variables aleatorias independientes, entonces, la densidad condicional de X dado $Y = y$ es igual a la densidad no condicional de X.*

Nota 6.15 (regla de Bayes)

$$f_{X|Y}(x \mid y) = \frac{f(x,y)}{f_Y(y)}, \quad si\ f_Y(y) > 0.$$

Como $f(x,y) = f_X(x)f_{Y|X}(y \mid x)$ y $f_Y(y) = \int_{-\infty}^{\infty} f(x,y)dx$, entonces,

$$f_{X|Y}(x \mid y) = \frac{f_X(x)f_{Y|X}(y \mid x)}{\int_{-\infty}^{\infty} f_X(x)f_{Y|X}(y \mid x)dx}. \quad ▲$$

Se han definido las funciones de densidad condicionales para variables aleatorias discretas y para variables aleatorias con función de densidad de probabilidad conjunta dada f. Supóngase ahora que X es una variable aleatoria absolutamente continua y que N es una variable aleatoria discreta. En este caso:

$$\begin{aligned} f_{X|N}(x \mid n) &= \lim_{\Delta x \to 0} \frac{P(x < X \leq x + \Delta x \mid N = n)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{P(N = n \mid x < X \leq x + \Delta x)}{P(N = n)} \frac{P(x < X \leq x + \Delta x)}{\Delta x} \\ &= \frac{P(N = n \mid X = x)}{P(N = n)} f_X(x). \end{aligned}$$

Ejemplo 6.16 *Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(0,1)$ y N una variable aleatoria con distribución binomial de parámetros $n+m$ y X . Entonces, para $0 < x < 1$; tenemos:*

$$
\begin{aligned}
f_{X|N}(x \mid n) &= \frac{P(N = n \mid X = x)}{P(N = n)} f_X(x) \\
&= \frac{\binom{n+m}{n}}{P(N = n)} x^n (1-x)^m \\
&= C x^n (1-x)^m.
\end{aligned}
$$

Esto es, X tiene bajo la condición $N = n$, una distribución beta de parámetros $(n+1)$ y $(m+1)$. ▲

Ejemplo 6.17 *Sea Y una variable aleatoria con distribución Poisson de parámetro Λ, donde Λ es una variable aleatoria con distribución $\Gamma(\alpha, \beta)$. Calcular $f_{\Lambda|Y}(\lambda \mid y)$.*

Solución: Se tiene que:

$$
f_{Y|\Lambda}(y \mid \lambda) = \begin{cases} \frac{\lambda^y e^{-\lambda}}{y!} & \text{si} \quad y = 0, 1, \cdots \\ 0 & \text{en} \quad \text{otro caso} \end{cases},
$$

entonces:

$$
\begin{aligned}
f(\lambda, y) &= f_{Y|\Lambda}(y \mid \lambda) f_\Lambda(\lambda) \\
&= \begin{cases} f_\Lambda(\lambda) \frac{\lambda^y e^{-\lambda}}{y!} & \text{si} \quad y = 0, 1, \cdots \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}.
\end{aligned}
$$

Puesto que,

$$
f_\Lambda(\lambda) = \frac{\beta^\alpha \lambda^{\alpha-1} \exp(-\lambda\beta)}{\Gamma(\alpha)} \mathcal{X}_{(0,\infty)}(\lambda)
$$

se obtiene:

$$
\begin{aligned}
f_Y(y) &= \int_0^\infty \frac{\beta^\alpha \lambda^{\alpha-1} \exp(-\lambda\beta)}{\Gamma(\alpha)} \frac{\lambda^y e^{-\lambda}}{y!} d\lambda \\
&= \frac{\beta^\alpha}{y! \Gamma(\alpha)} \int_0^\infty \lambda^{\alpha+y-1} \exp(-\lambda(\beta+1)) d\lambda \\
&= \frac{\beta^\alpha \Gamma(\alpha+y)}{y! \, (\beta+1)^{\alpha+y} \, \Gamma(\alpha)}
\end{aligned}
$$

y en consecuencia para $\lambda > 0$ y y entero no negativo,

$$\begin{aligned} f_{\Lambda|Y}(\lambda \mid y) &= \frac{f(\lambda, y)}{f_Y(y)} \\ &= \frac{\beta^\alpha \lambda^{\alpha+y-1} y!(\beta+1)^{\alpha+y}\Gamma(\alpha)\exp(-\lambda(\beta+1))}{y!\beta^\alpha\Gamma(\alpha)\Gamma(\alpha+y)} \\ &= \frac{(\beta+1)^{\alpha+y}\,\lambda^{\alpha+y-1}\exp(-\lambda(\beta+1))}{\Gamma(\alpha+y)}, \end{aligned}$$

esto es, Λ tiene, bajo la condición $Y = y$, una distribución gamma de parámetros $(\alpha + y)$ y $(\beta + 1)$. ▲

Definición 6.18 (valor esperado condicional de una función) *Sean X y Y variables aleatorias reales y h una función real tal que $h(X)$ es una variable aleatoria. Se define:*

$$E(h(X) \mid Y = y) := \begin{cases} \sum\limits_x h(x)P(X = x \mid Y = y) & si \quad \begin{matrix} X \text{ y } Y \text{ son variables} \\ \text{aleatorias discretas} \end{matrix} \\ \int_{-\infty}^{\infty} h(x) f_{X|Y}(x \mid y)dx & si \quad \begin{matrix} X \text{ y } Y \text{ son v.a con} \\ \text{fdp conjunta } f \end{matrix} \end{cases},$$

para todos los valores y de Y para los cuales $P(Y = y) > 0$, en el caso discreto, y $f_Y(y) > 0$ en el caso continuo.

Ejemplo 6.19 *Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:*

$$f(x,y) = \begin{cases} \frac{1}{2}y\exp(-xy) & si \quad 0 < x < \infty;\ 0 < y < 2 \\ 0 & en \quad otro\ caso. \end{cases}$$

Se tiene que:

$$\begin{aligned} f_Y(y) &= \int_{-\infty}^{\infty} f(x,y)dx \\ &= \int_0^{\infty} \frac{1}{2} y \exp(-xy)\mathcal{X}_{(0,2)}(y)dx \\ &= \frac{1}{2}\mathcal{X}_{(0,2)}(y). \end{aligned}$$

Por lo tanto, para $0 < y < 2$; se obtiene:

$$f_{X|Y}(x \mid y) = \frac{\frac{1}{2}y\exp(-xy)\mathcal{X}_{(0,\infty)}(x)}{\frac{1}{2}} = y\exp(-xy)\mathcal{X}_{(0,\infty)}(x),$$

de donde se deduce que:

$$
\begin{aligned}
E\left(\exp\left(\frac{X}{2}\right) \mid Y=1\right) &= \int_{-\infty}^{\infty} \exp\left(\frac{x}{2}\right) f_{X|Y}(x \mid 1)dx \\
&= \int_{0}^{\infty} \exp\left(\frac{x}{2}\right) \exp(-x)dx \\
&= 2. \quad \blacktriangle
\end{aligned}
$$

A partir de la definición anterior, se puede definir una nueva variable aleatoria como sigue:

Definición 6.20 (Valor esperado condicional) *Sean X e Y variables aleatorias reales definidas sobre $(\Omega, \Im, P)$ y h una función real tal que $h(X)$ es una variable aleatoria. La variable aleatoria $E(h(X) \mid Y)$ definida por:*

$$
\begin{array}{rccc}
E(h(X) \mid Y): & \Omega & \longrightarrow & \mathbb{R} \\
 & \omega & \longmapsto & E(h(X) \mid Y = Y(\omega))
\end{array}
$$

se llama valor esperado condicional de $h(X)$ dado Y.

Ejemplo 6.21 *Sea $\Omega = \{a, b, c\}$, $\Im = \wp(\Omega)$ y $P(\omega) = \frac{1}{3}$ para todo $\omega \in \Omega$.*
Considérese las variables aleatorias X e Y definidas como sigue:

$$
X(\omega) = \begin{cases} 1 & si \quad \omega = a, b \\ 0 & si \quad \omega = c \end{cases}
$$

y

$$
Y(\omega) = \begin{cases} \pi & si \quad \omega = a \\ \frac{1}{2} & si \quad \omega = b \\ -1 & si \quad \omega = c \end{cases},
$$

entonces:

$$
E(X \mid Y) = \begin{cases} E(X \mid Y = \pi) & si \quad \omega = a \\ E(X \mid Y = \frac{1}{2}) & si \quad \omega = b \\ E(X \mid Y = -1) & si \quad \omega = c. \end{cases}
$$

Se tiene que:

$$
\begin{aligned}
E(X \mid Y = \pi) &= \sum_{x} xP(X = x \mid Y = \pi) \\
&= P(X = 1 \mid Y = \pi) \\
&= \frac{P(X = 1, Y = \pi)}{P(Y = \pi)} \\
&= \frac{P(a)}{P(a)} = 1,
\end{aligned}
$$

análogamente se verifica que:

$$E\left(X \mid Y = \frac{1}{2}\right) = \frac{P(b)}{P(b)} = 1$$

y

$$E(X \mid Y = -1) = \frac{P(\Phi)}{P(c)} = 0.$$

Entonces:

$$E(X \mid Y) = \begin{cases} 1 & si \quad \omega = a, b \\ 0 & si \quad \omega = c \end{cases}.$$

Se observa, adicionalmente, que:

$$\begin{aligned} E\left(E(X \mid Y)\right) &= \sum_y E(X \mid Y = y) P(Y = y) \\ &= \frac{2}{3} = EX \end{aligned}$$

En el teorema siguiente se demuestra que este resultado se tiene en general. ▲

Teorema 6.22 *Sean X , Y y Z variables aleatorias reales definidas sobre $(\Omega, \Im, P)$ y h una función real tal que $h(X)$ es una variable aleatoria. Si $E(h(X))$ existe, entonces:*

$$E(h(X)) = E(E(h(X)) \mid Y)$$

Demostración. *Supóngase, inicialmente, que X e Y son variables aleatorias discretas. Entonces:*

$$\begin{aligned} E(E(h(X)) \mid Y) &= \sum_y E(h(X) \mid Y = y) P(Y = y) \\ &= \sum_y \sum_x h(x) P(X = x \mid Y = y) P(Y = y) \\ &= \sum_y \sum_x h(x) P(X = x, Y = y) \\ &= \sum_x h(x) \left(\sum_y P(X = x, Y = y) \right) \\ &= \sum_x h(x) P(X = x) \\ &= E(h(X)) \end{aligned}$$

Si X e Y son variables aleatorias con función de densidad de probabilidad conjunta f, entonces,

$$\begin{aligned}
E(E(h(X) \mid Y)) &= \int_{-\infty}^{\infty} E(h(X) \mid Y = y) f_Y(y) dy \\
&= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} h(x) f_{X|Y}(x \mid y) dx \right) f_Y(y) dy \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h(x) f(x,y) dx dy \\
&= \int_{-\infty}^{\infty} h(x) \left(\int_{-\infty}^{\infty} f(x,y) dy \right) dx \\
&= \int_{-\infty}^{\infty} h(x) f_X(x) dx \\
&= E(h(X)).
\end{aligned}$$

■

Ejemplo 6.23 *El número de clientes que llegan a un almacén en un día, es una variable aleatoria con distribución Poisson de media $\lambda = 10$. La cantidad de dinero (en miles de pesos) gastada, por cada cliente, es una variable aleatoria que se distribuye uniformemente sobre el intervalo $(0, 100]$. Determinar el monto de dinero que el almacén espera recaudar en un día.*

Solución: Sean X y M las variables aleatorias definidas por:

$X :=$ "número de clientes que llegan al almacén en un día"

y

$M :=$ "monto de dinero que el almacén recauda en un día".

Es claro que:

$$M = \sum_{i=1}^{X} M_i$$

donde

$M_i :=$ "cantidad de dinero gastada por el i-ésimo cliente"

De acuerdo al teorema anterior, se tiene que:

$$EM = E(E(M \mid X)).$$

Puesto que:

$$\begin{aligned} E\left(M \mid X = k\right) &= E\left(\sum_{i=1}^{k} M_i\right) \\ &= \sum_{i=1}^{k} E(M_i) \\ &= 50k \end{aligned}$$

entonces,

$$EM = E(50X) = 50EX = 500 \text{ miles de pesos.} \quad \blacktriangle$$

Nota 6.24 *En particular se tiene que, si $(\Omega, \Im, P)$ es un espacio de probabilidad arbitrario y si $A \in \Im$ fijo, entonces:*

$$\begin{aligned} P\left(A\right) &= E\left(\mathcal{X}_A\right) \\ &= E\left[E\left(\mathcal{X}_A \mid Y\right)\right] \\ &= \begin{cases} \sum\limits_{y} P\left(A \mid Y = y\right) P\left(Y = y\right) & \text{si} \quad Y \text{ es una v.a discreta} \\ \int_{-\infty}^{\infty} P\left(A \mid Y = y\right) f_Y(y) dy & \text{si} \quad Y \text{ es una v.a con fdp } f_Y. \end{cases} \end{aligned}$$

Ejemplo 6.25 *Sean X e Y variables aleatorias independientes con densidades f_X y f_Y, respectivamente. Calcular $P(X < Y)$.*

Solución: Sea $A := \{X < Y\}$. Entonces:

$$\begin{aligned} P(A) &= E\left(\mathcal{X}_A\right) \\ &= E(E(\mathcal{X}_A \mid Y)) \\ &= \int_{-\infty}^{\infty} E(\mathcal{X}_A \mid Y = y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} P(X < Y \mid Y = y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} P(X < y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} F_X(y) f_Y(y) dy \end{aligned}$$

siendo $F_X(.)$ la función de distribución de X. ▲

Teorema 6.26 *Si X e Y son variables aleatorias reales definidas sobre $(\Omega, \Im, P)$ y si h es una función real tal que $h(X)$ es una variable aleatoria. Entonces, se tiene que la esperanza condicional satisface las siguientes condiciones:*

1. $E(X \mid Y) \geq 0$; *si* $X \geq 0$ *c.s.*
2. $E(1 \mid Y) = 1$.
3. *Si* X *e* Y *son independientes, entonces,* $E(X \mid Y) = EX$.
4. $E(Xh(Y) \mid Y) = h(Y)E(X \mid Y)$.
5. $E(\alpha X + \beta Y \mid Z) = \alpha E(X \mid Z) + \beta E(Y \mid Z)$; *para* $\alpha, \beta \in \mathbb{R}$.

Demostración. Se hacen las demostraciones para el caso discreto. El caso continuo se trabaja de manera similar.

1. Supóngase que X toma los valores $x_1, x_2, \cdots$. Puesto que $P(X < 0) = 0$, entonces, $P(X = x_j) = 0$; para $x_j < 0$. Por lo tanto,

$$\begin{aligned} E(X \mid Y = y) &= \sum_j x_j P(X = x_j \mid Y = y) \\ &= \sum_{j:x_j \geq 0} x_j P(X = x_j \mid Y = y) \geq 0 \end{aligned}$$

y en consecuencia $E(X \mid Y) \geq 0$.

2. Sea $X := 1$ entonces:

$$E(X \mid Y = y) = 1P(X = 1 \mid Y = y) = 1$$

3. Como X y Y son independientes, se tiene que

$$P(X = x \mid Y = y) = P(X = x),$$

para todo y; con $P(Y = y) > 0$. Por lo tanto,

$$\begin{aligned} E\left(X \mid Y = y\right) &= \sum_x x P(X = x \mid Y = y) \\ &= \sum_x x P(X = x) \\ &= EX \end{aligned}$$

4.

$$\begin{aligned} E\left(Xh(Y) \mid Y=y\right) &= \sum_x xh(y)P(X=x \mid Y=y) \\ &= h(y)E(X \mid Y=y) \end{aligned}$$

y en consecuencia:

$$E(Xh(Y) \mid Y) = h(Y)E(X \mid Y).$$

5.

$$\begin{aligned} E\left(\alpha X+\beta Y \mid Z=z\right) &= \sum_{x,y}\left(\alpha x+\beta y\right)P(X=x,Y=y \mid Z=z) \\ &= \alpha\sum_{x,y} xP\left(X=x,Y=y \mid Z=z\right) \\ &\quad +\beta\sum_{x,y} yP\left(X=x,Y=y \mid Z=z\right) \\ &= \alpha\sum_x x\sum_y P(X=x,Y=y \mid Z=z) \\ &\quad +\beta\sum_y y\sum_x P\left(X=x,Y=y \mid Z=z\right) \\ &= \alpha\sum_x xP(X=x \mid Z=z) \\ &\quad +\beta\sum_y yP\left(Y=y \mid Z=z\right) \\ &= \alpha E(X \mid Z=z)+\beta E(Y \mid Z=z). \end{aligned}$$

Por lo tanto,

$$E(\alpha X+\beta Y \mid Z) = \alpha E(X \mid Z)+\beta E(Y \mid Z).$$

■

Ejemplo 6.27 *Considérese una sucesión Bernoulli de longitud $(n+m)$ con probabilidad de éxito p. Calcular el número esperado de éxitos en los primeros n ensayos.*

Solución: Sean Y :="número total de éxitos" y, para cada $i=1,\cdots,n$; sea

$$X_i := \begin{cases} 1 & \text{si} \quad \text{se obtiene éxito en el } i-\text{ésimo ensayo.} \\ 0 & \text{en} \quad \text{caso contrario} \end{cases}.$$

Es claro que

$$X := \sum_{i=1}^{n} X_i = \text{“número de éxitos en los primeros } n \text{ ensayos”},$$

por lo tanto,

$$EX = E(E(X \mid Y)),$$

pero,

$$\begin{aligned} E(X \mid Y = k) &= E\left(\sum_{i=1}^{n} X_i \mid Y = k\right) \\ &= \sum_{i=1}^{n} E(X_i \mid Y = k) \\ &= \sum_{i=1}^{n} P(X_i = 1 \mid Y = k) \\ &= \frac{kn}{n+m}, \end{aligned}$$

luego,

$$EX = E\left(\frac{nY}{n+m}\right) = \frac{n}{n+m} EY = \frac{n}{n+m} p(n+m) = np. \quad \blacktriangle$$

6.2. Esperanza condicional dada una $\sigma-$álgebra.

En esta sección se va a trabajar el concepto de valor esperado condicional de una variable aleatoria con respecto a una $\sigma-$álgebra, el cual generaliza el concepto de esperanza condicional desarrollado en la sección anterior.

Definición 6.28 (esperanza condicional de X dado B) *Sea X una variable aleatoria real definida sobre $(\Omega, \Im, P)$ y sea $B \in \Im$ con $P(B) > 0$. La esperanza condicional de X dado B se define como sigue:*

$$E(X \mid B) := \frac{E(X\mathcal{X}_B)}{P(B)}$$

si el valor esperado de $Y := X\mathcal{X}_B$ existe.

Ejemplo 6.29 *Se lanza un dado corriente dos veces consecutivas. Sea X la variable aleatoria que denota la suma de los resultados obtenidos y B el evento que indica que el resultado del primer lanzamiento es 5. Calcular $E(X \mid B)$.*

Solución El espacio muestral del experimento está dado por:

$$\Omega = \{(a,b) : a,b \in \{1,\cdots,6\}\}$$

y es claro que:

$$(X\mathcal{X}_B)(a,b) = \begin{cases} 5+b & \text{si} \quad a=5,\ b \in \{1,\cdots,6\} \\ 0 & \text{en} \quad \text{otro caso} \end{cases},$$

luego:

$$\begin{aligned} E(X\mathcal{X}_B) &= \sum_{b=1}^{6} (5+b)\, P(X\mathcal{X}_B = 5+b) \\ &= \frac{1}{36}\sum_{b=1}^{6} (5+b) \\ &= \frac{51}{36} \end{aligned}$$

y

$$P(B) = \frac{1}{6}.$$

Por lo tanto,

$$E(X \mid B) = \frac{\frac{51}{36}}{\frac{1}{6}} = \frac{51}{6} = 8.5. \quad \blacktriangle$$

Ejemplo 6.30 *Sea X una variable aleatoria con distribución exponencial de parámetro λ. Calcular $E(X \mid \{X \geq t\})$.*

Solución: Puesto que $X \stackrel{d}{=} Exp(\lambda)$ se tiene que su función de densidad está dada por:

$$f_X(x) = \begin{cases} \lambda \exp(-\lambda x) & \text{si} \quad x \geq 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

por lo tanto,

$$\begin{aligned} P(X \geq t) &= 1 - P(X < t) \\ &= \begin{cases} 1 - \int_0^t \lambda \exp(-\lambda x)dx & \text{si} \quad t \geq 0 \\ 1 & \text{si} \quad t < 0 \end{cases} \\ &= \begin{cases} \exp(-\lambda t) & \text{si} \quad t \geq 0 \\ 1 & \text{si} \quad t < 0 \end{cases}. \end{aligned}$$

Por otra parte,

$$
\begin{aligned}
E(X\mathcal{X}_{\{X\geq t\}}) &= \begin{cases} \int_t^\infty x\lambda\exp(-\lambda x)dx & \text{si} \quad t\geq 0 \\ \int_0^\infty x\lambda\exp(-\lambda x)dx & \text{si} \quad t<0 \end{cases} \\
&= \begin{cases} t\exp(-\lambda t)+\frac{1}{\lambda}\exp(-\lambda t) & \text{si} \quad t\geq 0 \\ \frac{1}{\lambda} & \text{si} \quad t<0 \end{cases}
\end{aligned}
$$

y en consecuencia,

$$
E(X \mid \{X\geq t\}) = \begin{cases} t+\frac{1}{\lambda} & \text{si} \quad t\geq 0 \\ \frac{1}{\lambda} & \text{si} \quad t<0 \end{cases} \qquad \blacktriangle
$$

Ejemplo 6.31 *Sean X, Y y Z variables aleatorias con distribución conjunta dada por:*

$\mathbf{x} := (x,y,z)$	$(0,0,0)$	$(0,0,1)$	$(0,1,1)$	$(1,1,1)$	$(1,1,0)$	$(1,0,1)$
$P(\mathbf{X}=\mathbf{x})$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{9}$	$\frac{1}{9}$	$\frac{2}{9}$

Calcular $E(X \mid Y=0, Z=1)$.

Solución:

$$
\begin{aligned}
E(X \mid Y=0, Z=1) &= \sum_x xP(X=x \mid Y=0, Z=1) \\
&= P(X=1 \mid Y=0, Z=1) \\
&= \frac{P(X=1, Y=0, Z=1)}{P(Y=0, Z=1)} \\
&= \frac{\frac{2}{9}}{\sum_x P(X=x, Y=0, Z=1)} \\
&= \frac{1}{2}. \quad \blacktriangle
\end{aligned}
$$

Definición 6.32 (esperanza condicional de X dado $\mathcal{G}$) *Sea X una variable aleatoria real definida sobre$(\Omega, \Im, P)$ con $EX<\infty$. Sea $\mathcal{G}$ una sub-$\sigma-$álgebra de $\Im$. La esperanza condicional de X dado $\mathcal{G}$, denotada por $E(X \mid \mathcal{G})$, es una variable aleatoria $\mathcal{G}-$medible tal que:*

$$
E([X-E(X \mid \mathcal{G})]\mathcal{X}_G)=0; \text{ para todo } G\in\mathcal{G}. \tag{6.1}
$$

Ejemplo 6.33 *Sea* $\Omega = \{a, b, c\}$, $\Im = \wp(\Omega)$ *y* $P(\omega) = \frac{1}{3}$; *para todo* $\omega \in \Omega$. *Supóngase que* X *es la variable aleatoria real dada por:*

$$X(\omega) = \begin{cases} 0 & si \quad \omega = a, b \\ 2 & si \quad \omega = c \end{cases}$$

y sea $\mathcal{G} := \{\Phi, \{a\}, \{b, c\}, \Omega\}$.

Se tiene que Y *dada por:*

$$Y(\omega) = \begin{cases} 0 & si \quad \omega = a \\ 1 & si \quad \omega = b, c \end{cases}$$

es igual a $E(X \mid \mathcal{G})$. *En efecto,*

1. Y *es* $\mathcal{G}$ *-medible, pues:*

$$Y^{-1}((-\infty, x]) = \begin{cases} \Phi & si \quad x < 0 \\ \{a\} & si \quad 0 \leq x < 1 \\ \Omega & si \quad x \geq 1. \end{cases}$$

2. Y *satisface la condición (6.1), ya que:*

$$(X - Y)(\omega) = \begin{cases} 0 & si \quad \omega = a \\ -1 & si \quad \omega = b \\ 1 & si \quad \omega = c \end{cases}$$

por lo tanto,

$$\begin{aligned} (X - Y)\mathcal{X}_{\{a\}} &= 0 = (X - Y)\mathcal{X}_{\Phi} \\ (X - Y)\mathcal{X}_{\{b,c\}} &= (X - Y) = (X - Y)\mathcal{X}_{\Omega} \end{aligned}$$

y en consecuencia,

$$\begin{aligned} E\left[(X - Y)\mathcal{X}_{\{a\}}\right] &= 0 = E\left[(X - Y)\mathcal{X}_{\Phi}\right] \\ E\left[(X - Y)\mathcal{X}_{\{b,c\}}\right] &= E\left[(X - Y)\mathcal{X}_{\Omega}\right] \\ &= E\left[(X - Y)\right] \\ &= (-1)P(b) + 1P(c) = 0. \quad \blacktriangle \end{aligned}$$

Ejemplo 6.34 *Sea* $\Omega = \{a, b, c\}$, $\Im = \wp(\Omega)$ *y* $P(\omega) = \frac{1}{3}$; *para todo* $\omega \in \Omega$. *Supóngase que* X *es la variable aleatoria real dada por:*

$$X(\omega) = \begin{cases} 0 & si \quad \omega = a \\ -1 & si \quad \omega = b \\ 1 & si \quad \omega = c \end{cases}$$

y sea $\mathcal{H} := \{\Phi, \Omega]$.

Es fácil verificar que $Z := 0$ *es* $\mathcal{H}$*-medible y que satisface la condición (6.1). Por lo tanto* $Z = E(X \mid \mathcal{H})$. ▲

Notación 6.35 *Sean* X, $Y_1, \cdots, Y_n$ *variables aleatorias reales. El valor esperado* $E(X \mid \sigma(Y_1, \cdots, Y_n))$, *siendo* $\sigma(Y_1, \cdots, Y_n)$ *la menor* $\sigma-$*álgebra con respecto a la cual las variables aleatorias* $Y_1, \cdots, Y_n$ *son medibles, se acostumbra denotar mediante* $E(X \mid Y_1, \cdots, Y_n)$.

Nota 6.36 *La esperanza condicional tiene una muy útil aplicación en la teoría bayesiana de la estadística. Un problema clásico, en esta teoría, se obtiene cuando se observan datos* $\mathbf{X} := (X_1, \cdots, X_n)$ *cuya distribución está determinada por el valor de una variable aleatoria* θ, *la cual tiene una distribución específica llamada distribución a priori. Con base en los valores de los datos* $\mathbf{X}$, *el problema que interesa es estimar el valor desconocido de* θ. *Un estimador de* θ *puede ser cualquier función* $d(\mathbf{X})$ *de los datos. En la teoría bayesiana se busca escoger* $d(\mathbf{X})$ *de tal forma que se minimice el valor esperado condicional del cuadrado de la distancia entre el estimador y el parámetro, esto es, se busca que* $E([\theta - d(\mathbf{X})]^2 \mid \mathbf{X})$ *se minimice.*

Puesto que condicionado sobre $\mathbf{X}$ *se tiene que* $d(\mathbf{X})$ *es constante y como para cualquier variable aleatoria* W *se tiene que* $E\left[(W-c)^2\right]$ *se minimiza cuando* $c = EW$ *entonces se concluye que el estimador que minimiza a* $E([\theta - d(\mathbf{X})]^2 \mid \mathbf{X})$ *está dado por* $d(\mathbf{X}) = E(\theta \mid \mathbf{X})$. *Este estimador se llama estimador de Bayes.*

Ejemplo 6.37 *La estatura alcanzada por el hijo de un individuo, con una estatura de* x *cm, es una variable aleatoria con distribución normal de media* $x+3$ *y varianza* 2. *¿Cuál es la mejor predicción de la estatura que se espera que alcance el hijo de un individuo que mide* $170cm$*?*

Solución: Sea X la variable aleatoria que denota la estatura del padre y sea θ la variable aleatoria que denota la estatura del hijo. De acuerdo con la información suministrada $\theta \overset{d}{=} \mathcal{N}(X+3, 2)$. Por la observación anterior, se sabe que el mejor predictor posible de la estatura del hijo es $d(X) = E(\theta \mid X)$. Por lo tanto, si $X = 170$, entonces, $\theta \overset{d}{=} \mathcal{N}(173, 2)$ de donde

$$E(\theta \mid X = 170) = 173 \text{ cm.} \quad \blacktriangle$$

6.3. Ejercicios

1. Supóngase que X e Y son variables aleatorias discretas con distribución conjunta dada por:

$X\backslash Y$	1	2	3	4
1	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
2	0	$\frac{2}{16}$	$\frac{1}{16}$	$\frac{1}{16}$
3	0	0	$\frac{3}{16}$	$\frac{1}{16}$
4	0	0	0	$\frac{4}{16}$

   a) Calcular las distribuciones de X e Y.

   b) Determinar $E(X \mid Y = 1)$ y $E(Y \mid X = 1)$.

2. Supóngase que X e Y son variables aleatorias discretas cuya distribución conjunta de probabilidades está dada por:

$X\backslash Y$	1	2	3
1	$\frac{1}{12}$	$\frac{2}{12}$	0
2	0	$\frac{4}{12}$	$\frac{1}{12}$
3	$\frac{1}{12}$	$\frac{1}{12}$	$\frac{2}{12}$

   Verificar que $E\left(E\left(X \mid Y\right)\right) = EX$ y que $E\left(E\left(Y \mid X\right)\right) = EY$.

3. Se lanza un dado corriente dos veces consecutivas. Sea X la variable aleatoria que denota el número de resultados pares obtenidos y Y la variable aleatoria que denota el número de resultados menores que 4 obtenidos. Hallar $E\left(XE\left(Y \mid X\right)\right)$.

4. Una urna contiene 8 bolas rojas y 5 negras. Se realizan, de manera consecutiva, dos extracciones sin repetición. En la primera extracción se sacan 2 bolas y en la segunda 3. Sean X la variable aleatoria que denota el número de bolas rojas sacadas en la primera extracción y Y la variable aleatoria que denota el número de bolas rojas sacadas en la segunda extracción. Hallar $E\left(Y \mid X = 1\right)$.

5. Un minero se encuentra atrapado en una mina que tiene 3 puertas. La primera puerta conduce a un túnel que lo lleva a la libertad después de dos horas de camino. La segunda puerta conduce a un túnel que lo regresa a la mina después de tres horas de camino. La tercera puerta conduce a un túnel que lo regresa a la mina después de seis horas de

camino. Si el minero escoge la puerta 1 con probabilidad $\frac{2}{7}$, la puerta 2 con probabilidad $\frac{1}{7}$ y la puerta 3 con probabilidad $\frac{4}{7}$, entonces, ¿cuál es la longitud de tiempo esperado hasta que el minero recobre su libertad?

6. Un jugador extrae dos bolas, una tras otra, de una urna que contiene 5 bolas rojas y 4 negras. Por cada bola roja extraída el jugador gana dos unidades monetarias y por cada bola negra extraída el jugador pierde una unidad monetaria. Sea X la variable que denota la fortuna del jugador y Y la variable aleatoria que toma el valor 1 si la primera bola extraída es roja y el valor 0 si la primera bola extraída es de color negro.

   a) Calcular $E(X \mid Y)$.

   b) Usar la parte a. para hallar EX.

7. Supóngase que X es una variable aleatoria discreta con función de densidad dada por $f_X(x) = \frac{x}{3}$; si $x = 1, 2$ y que Y es una variable aleatoria tal que:

$$f_{Y|X}(y \mid x) = \binom{x}{y}\left(\frac{1}{2}\right)^x ; \quad \text{para } y = 0, \cdots, x \text{ y } x = 1, 2.$$

   Hallar:

   a) La distribución conjunta de X y Y.

   b) $E(X \mid Y)$.

8. Si X tiene una distribución Bernoulli de parámetro p y si $E(Y \mid X = 0) = 1$ y $E(Y \mid X = 1) = 2$, ¿a qué es igual EY ?

9. Sean X e Y variables aleatorias uniformemente distribuidas sobre la región del triángulo acotado por :$x = 2$, $y = 0$ y $2y = x$, esto es, la función de densidad conjunta de las variables aleatorias X e Y está dada por:

$$f(x,y) = \begin{cases} \frac{1}{\text{área del triángulo}} & \text{si} \quad (x,y) \text{ está en el triángulo} \\ 0 & \text{en} \quad \text{caso contrario.} \end{cases}$$

   Calcular:

   a) $P\left(Y \leq \frac{X}{3}\right)$.

b) $P(Y \geq 0.5)$.

c) $P(X \leq 1.5 \mid Y = 0.5)$.

10. Supóngase que la función de densidad de probabilidad conjunta de las variables aleatorias X e Y está dada por:

$$f(x, y) = \begin{cases} e^{-y} & \text{para} \quad x > 0,\ y > x \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

a) Calcular $P(X > 2 \mid Y < 4)$.

b) $E(X \mid Y = y)$.

c) $E(Y \mid X = x)$.

d) Verificar que $EX = E(E(X \mid Y))$ y que $EY = E(E(Y \mid X))$.

11. Sean X e Y variables aleatorias independientes. Demostrar que:

$$E(Y \mid X = x) = EY; \qquad \text{para todo } x.$$

12. Demostrar que si $E(Y \mid X = x) = EY$;para todo x, entonces, X e Y son no correlacionadas. Dar un contraejemplo que muestre que el recíproco no es cierto.

Sugerencia: Probar y usar el hecho que $E(XY) = E(XE(Y \mid X))$.

13. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x, y) = \begin{cases} \frac{3}{8}\left(x + y^2\right) & \text{si} \quad 0 < x < 2\,,\ 0 < y < 1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Calcular $E(X \mid Y = \frac{1}{4})$.

14. Sean $Y_1^{(n)}, Y_2^{(n)}, \ldots$ variables aleatorias independientes, igualmente distribuidas, con valores en $\mathbb{N}_0$; $n \in \mathbb{N}$. Supóngase que: $m := E(Y_1^{(n)}) < \infty$ y $0 < \sigma^2 := Var(Y_1^{(n)}) < \infty$. Sean $Z_0 := 1$ y

$$Z_{n+1} := \begin{cases} Y_1^{(n)} + Y_2^{(n)} + \ldots + Y_k^{(n)} & \text{si} \quad Z_n = k, k > 0 \\ 0 & \text{si} \quad Z_n = 0. \end{cases}$$

a) Calcular $E(Z_{n+1} \mid Z_n)$ y EZ_n.

b) Sea $f(s) := E\left(s^{Z_1}\right) = \sum_k P(Z_1 = k)\, s^k$; con $|s| \leq 1$, la función generadora de probabilidades de Z_1. Calcular $f_n(s) := E(s^{Z_n})$ en términos de f.

c) ¿A qué es igual $Var(Z_n)$?

15. La varianza condicional de Y dado $X = x$ está definida por:

$$Var(Y \mid X = x) := E\left(Y^2 \mid X = x\right) - (E(Y \mid X = x))^2 .$$

Demostrar que

$$VarY = E(Var(Y \mid X)) + Var(E(Y \mid X)) .$$

16. Sean X e Y variables aleatorias con distribución conjunta dada por:

$X\backslash Y$	0	1	2
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
1	0	$\frac{1}{8}$	$\frac{1}{8}$
2	0	0	$\frac{3}{8}$

Hallar $Var(Y \mid X)$.

17. Sea Y una variable aleatoria con distribución Poisson de parámetro λ. Supóngase que Z es la variable aleatoria definida por:

$$Z := \sum_{i=1}^{Y} X_i$$

donde las variables aleatorias $X_1, X_2, \cdots$ son independientes entre si e independientes de Y. Supóngase, además, que las variables aleatorias $X_1, X_2, \cdots$ son igualmente distribuidas con distribución Bernoulli de parámetro $p \in (0,1)$. Hallar EZ y $Var(Z)$.

18. Sean X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} \frac{1}{8}\left(x^2 - y^2\right)\exp(-x) & \text{si} \quad x > 0\,,\ |y| < x \\ 0 & \text{en} \quad \text{otro caso.} \end{cases} .$$

Calcular $E(X \mid Y = 1)$.

19. Supóngase que la función de densidad de probabilidad conjunta de las variables aleatorias X y Y está dada por:

$$f(x,y) = \begin{cases} \frac{6}{7}\left(x^2 + \frac{xy}{2}\right) & \text{si} \quad 0 < x < 1,\ 0 < y < 2 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Calcular $E(X \mid Y = y)$.

20. Sea (X, Y) un vector aleatorio con distribución uniforme en el triángulo limitado por: $x \geq 0$, $y \geq 0$ y $(x + y) \leq 2$. Calcular $E(Y \mid X = x)$.

21. Sea X e Y variables aleatorias con función de densidad de probabilidad conjunta dada por:

$$f(x,y) = \begin{cases} 8xy & \text{si} \quad 0 < y \leq x < 1 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}.$$

Calcular:

a) $E(X \mid Y = y)$.
b) $E(X^2 \mid Y = y)$.
c) $Var(X \mid Y = y)$.

22. La función de densidad conjunta de X e Y está dada por:

$$f(x,y) = \begin{cases} xe^{-x(y+1)} & \text{si} \quad x > 0 \text{ y } y > 0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

Hallar la función de densidad condicional de X dado que $Y = y$ y la de Y dado que $X = x$.

23. Sea $\mathbf{X} = (X, Y)$ un vector aleatorio con función de densidad dada por:

$$f(x,y) = \frac{1}{2\pi \times 0.6} \exp\left[-\frac{1}{2 \times 0.36}\left((x-3)^2 - 1.6(x-3)(y-3) + (y-3)^2\right)\right].$$

Calcular:

a) Las funciones de densidad marginales de X e Y.
b) La función de densidad condicional $f_{Y|X}(y \mid X = 2)$.
c) El valor de c para el cual $P(Y > c \mid X = 2) = 0.05$.

24. Se lanzan dos dados corrientes simultáneamente. Sean X la variable aleatoria que denota la suma de los resultados obtenidos y B el evento definido por $B :=$" la suma de los resultados obtenidos es divisible por 3 ". Calcular $E(X \mid B)$.

25. Sean X y Y variables aleatorias independientes e igualmente distribuidas con distribución uniforme sobre el intervalo $(0, 2)$. Calcular:

    a) $P(X \geq 1 \mid (X + Y) \leq 3)$.
    b) $E(X \mid (X + Y) \leq 3)$.

26. Sean X y Y variables aleatorias con función de densidad conjunta dada por:

$$f(x, y) = \begin{cases} \exp(-x - y) & \text{si} \quad x > 0 \text{ y } y > 0 \\ 0 & \text{en} \quad \text{otro caso.} \end{cases}$$

    Calcular $E(X + Y \mid X < Y)$.

27. Sea X una variable aleatoria con distribución uniforme sobre el intervalo $(0, 1)$ y sea Y una variable aleatoria con distribución uniforme sobre $(0, X)$. Determinar:

    a) La función de densidad de probabilidad conjunta de X e Y.
    b) La función de densidad marginal de Y.

28. El número de personas que entran a un ascensor, en el primer piso de un edificio, es una variable aleatoria con distribución Poisson de media 10. Supóngase que hay N pisos, además del primero, y que cada persona abandona, con igual probabilidad, el ascensor en cualquiera de los N pisos, independientemente de donde lo abandonen los demás. Calcular el número esperado de paradas que hace el ascensor hasta quedar desocupado. Se supone que no entran personas al ascensor en los pisos diferentes al primero.

29. Se lanza una moneda corriente hasta obtener cinco "caras" de manera consecutiva. Determinar el número esperado de lanzamientos requeridos.

30. Se lanza un dado corriente de manera sucesiva. Sean X e Y las variables aleatorias que denotan, respectivamente, el número de lanzamientos requeridos para obtener 2 y 4. Calcular:

a) EX.

b) $E(X \mid Y = 1)$.

c) $E(X \mid Y = 5)$.

31. Una urna contiene 6 bolas rojas y 5 blancas. Se extraen, de manera consecutiva, dos muestras, sin reemplazo, de tamaños 3 y 5. Sea X el número de bolas blancas en la primera muestra y Y el número de bolas blancas en la segunda muestra. Calcular $E(X \mid Y = k)$, para $k = 1, 2, 3, 4, 5$.

32. Supóngase que se realiza un experimento de tipo Bernoulli tantas veces sea necesario hasta que se obtiene por primera vez éxito. Si la probabilidad de éxito es una variable aleatoria con distribución uniforme en el intervalo $(0, 1)$, ¿a qué es igual la probabilidad de que sean necesarias n repeticiones del experimento?

33. Sea X una variable aleatoria cuyo valor esperado existe. Demostrar que:

$$E(X) = E(X \mid X < y) P(X < y) + E(X \mid X \geq y) P(X \geq y).$$

34. La covarianza condicional de X y Y dado Z está definida por:

$$Cov(X, Y \mid Z) := E[(X - E(X \mid Z))(Y - E(Y \mid Z)) \mid Z].$$

a) Demostrar que:

$$Cov(X, Y \mid Z) = E(XY \mid Z) - E(X \mid Z) E(Y \mid Z).$$

b) Verificar que:

$$Cov(X, Y) = E[Cov(X, Y \mid Z)] + Cov(E[X \mid Z], E[Y \mid Z]).$$

35. Sea $\mathbf{X} = (X, Y)$ un vector aleatorio con distribución normal bivariada. Demostrar que la distribución condicional de Y, dado que $X = x$, es normal de parámetros μ dado por:

$$\mu = E(Y) + \rho(X, Y) \frac{\sqrt{VarX}}{\sqrt{VarY}} (X - EX)$$

y σ^2 dado por:

$$\sigma^2 = (VarX)(1 - \rho(X, Y)).$$

Capítulo 7

Teoremas Límites

En estadística es muy importante poder sacar conclusiones acerca de una población a partir de una muestra dada y establecer qué tan confiables son dichas conclusiones. Es por ello que resulta de gran importancia poder conocer el comportamiento asintótico de las sucesiones de variables aleatorias.

En este capítulo se estudiarán algunos de los resultados más importantes dentro de los teoremas límites, como son: la ley débil de los grandes números, la ley fuerte de los grandes números y el teorema central del límite, llamado éste último así, por ser el central de los teoremas límites de la teoría de la probabilidad (ver [HerA]).

7.1. Ley débil de los grandes números

Cuando se conoce la distribución de una variable aleatoria X se puede calcular su valor esperado y su varianza. Sin embargo, el conocimiento de estas dos cantidades no permite calcular probabilidades del tipo $P(|X - c| > \epsilon)$. El matemático ruso Chebyschev, demostró una desigualdad, conocida como *desigualdad de Chebyschev* (comparar con el ejercicio 42 del capítulo 2), la cual ofrece una cota superior para tales probabilidades. Aún cuando su uso práctico es limitado, su importancia teórica es indiscutible como se verá más adelante.

La desigualdad de Chebyschev es un caso particular de la *desigualdad de Markov* (comparar con el ejercicio 41 del capítulo 2), la cual se presenta a continuación.

Lema 7.1 (desigualdad de Markov) *Si X es una variable aleatoria no* *negativa cuyo valor esperado existe, entonces, para todo* $a > 0$; *se satisface:*

$$P(X \geq a) \leq \frac{EX}{a}. \tag{7.1}$$

Demostración. *Considérese la variable aleatoria* I *definida como:*

$$I := \begin{cases} 1 & si \quad X \geq a \\ 0 & en \quad otro\ caso. \end{cases}$$

Como $X \geq 0$, *entonces,* $I \leq \frac{X}{a}$. *Tomando valores esperados a ambos lados de la desigualdad se obtiene el resultado.* ■

Ejemplo 7.2 *Por experiencia, un profesor sabe que el puntaje obtenido por un estudiante en el examen final de su materia es una variable aleatoria con media* 75. *Obtener una cota superior para la probabilidad de que el estudiante obtenga un puntaje mayor o igual a* 85.

Solución: Sea X la variable aleatoria definida como:

$$X := \text{“puntaje obtenido por el estudiante en el examen final”}$$

Puesto que X es una variable aleatoria no negativa, entonces, de la desigualdad de Markov se obtiene que:

$$P(X \geq 85) \leq 0.88235. \quad \blacktriangle$$

Ejemplo 7.3 *Supóngase que* X *es una variable aleatoria con distribución binomial de parámetros* 5 *y* $\frac{1}{3}$. *Usar la desigualdad de Markov para encontrar una cota superior para* $P(X \geq 2)$. *Calcular de manera exacta* $P(X \geq 2)$ *y comparar los resultados.*

Solución: Se sabe que $EX = \frac{5}{3}$, por lo tanto,

$$P(X \geq 2) \leq \frac{5}{6}.$$

Por otra parte,

$$\begin{aligned} P(X \geq 2) &= 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{5}{0}\left(\frac{1}{3}\right)^0\left(\frac{2}{3}\right)^5 - \binom{5}{1}\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^4 \\ &= 0.53909 \end{aligned}$$

Esto indica que la información aportada por la desigualdad de Markov no es de mucha utilidad en este caso. ▲

En muchas situaciones no se tiene información específica de la distribución de la variable aleatoria X y es, en tales casos, donde la desigualdad de Chebyschev puede ofrecer información valiosa acerca del comportamiento de la variable.

Teorema 7.4 (desigualdad de Chebyschev) *Sea X una variable aleatoria con $Var(X) < \infty$. Entonces, para todo $\epsilon > 0$; se satisface lo siguiente:*

$$P(|X - EX| \geq \epsilon) \leq \frac{1}{\epsilon^2} Var(X).$$

Demostración. *Sea $Y := |X - EX|^2$ y $a = \epsilon^2$. De la desigualdad de Markov se obtiene:*

$$P(|X - EX| \geq \epsilon) \leq \frac{E\left(|X - EX|^2\right)}{\epsilon^2} = \frac{1}{\epsilon^2} Var(X) \tag{7.2}$$

con lo cual queda demostrado. ∎

Es evidente que (7.2) es equivalente a:

$$P(|X - EX| < \epsilon) \geq 1 - \frac{1}{\epsilon^2} Var(X).$$

Así mismo, si se toma $\epsilon := \sigma k$ con $k > 0$ y $\sigma := \sqrt{Var(X)}$, se obtiene:

$$P(|X - EX| \geq k\epsilon) \leq \frac{1}{k^2}.$$

Si en (7.2) se reemplaza EX por cualquier número real C, se obtiene:

$$P(|X - C| \geq \epsilon) \leq \frac{E\left(|X - C|^2\right)}{\epsilon^2}.$$

Esta última expresión es la que algunos autores llaman desigualdad de Chebyschev. (ver [Mey]).

Nota 7.5 *En (7.1) y (7.2) se pueden reemplazar $P(X \geq a)$ y $P(|X - EX| \geq \epsilon)$ por $P(X > a)$ y $P(|X - EX| > \epsilon)$, respectivamente, y las desigualdades siguen siendo válidas.*

Ejemplo 7.6 *¿Existe alguna variable aleatoria X para la cual se satisfaga:*

$$P(\mu_X - 2\sigma_X \leq X \leq \mu_X + 2\sigma_X) = 0.6 \tag{7.3}$$

siendo μ_X el valor esperado de X y σ_X la desviación estándar de X?

Solución: De la desigualdad de Chebyschev se sigue que:

$$P(\mu_X - 2\sigma_X \leq X \leq \mu_X + 2\sigma_X) = P(|X - \mu_X| \leq 2\sigma_X) \geq \frac{3}{4}.$$

Por lo tanto no puede existir ninguna variable aleatoria X que satisfaga (7.3). ▲

Ejemplo 7.7 *Demostrar que si* $Var(X) = 0$, *entonces,* $P(X = EX) = 1$.

Prueba: Por la desigualdad de Chebyschev se tiene que para cualquier $n \geq 1$; se satisface:

$$P\left(|X - EX| > \frac{1}{n}\right) = 0.$$

Tomando límite cuando $n \to \infty$ se obtiene:

$$\begin{aligned} 0 &= \lim_{n\to\infty} P\left(|X - EX| > \frac{1}{n}\right) = P\left(\lim_{n\to\infty}\left\{|X - EX| > \frac{1}{n}\right\}\right) \\ &= P(X \neq EX). \end{aligned}$$

Esto es, $P(X = EX) = 1$. ▲

Como una aplicación de la desigualdad de Chebyschev se obtiene la *ley débil de los grandes números.* Éste es uno de los resultados teóricos más importantes de la teoría de la probabilidad y fue demostrado inicialmente por Jacobo Bernoulli para un caso particular. La ley débil de los grandes números establece que el valor esperado EX de una variable aleatoria X puede ser considerado como una "idealización", para n "suficientemente grande", del promedio aritmético $\overline{X} := \frac{X_1+\cdots+X_n}{n}$, donde $X_1, \cdots, X_n$ son variables aleatorias independientes e igualmente distribuidas con la misma distribución de X.

Para enunciar esta ley se requiere del siguiente concepto:

Definición 7.8 (variables aleatorias igualmente distribuidas) *Se dice que* $X_1, X_2, \cdots$ *es una sucesión de variables aleatorias independientes e igualmente distribuidas si:*

1. *Para cada* $n \in \mathbb{Z}^+$ *se tiene que* $X_1, \cdots, X_n$ *son variables aleatorias independientes.*

2. *Para todo* $i, j \in \mathbb{Z}^+$; X_i *y* X_j *tienen la misma distribución.*

Nota 7.9 *La definición anterior puede ser generalizada a vectores aleatorios como sigue: se dice que* $\mathbf{X}_1, \mathbf{X}_2, \cdots$ *es una sucesión de vectores aleatorios independientes e igualmente distribuidos si:*

1. *Para cada* $n \in \mathbb{Z}^+$ *se tiene que* $\mathbf{X}_1, \cdots, \mathbf{X}_n$ *son vectores aleatorios independientes.*
2. *Para todo* $i, j \in \mathbb{Z}^+$; $\mathbf{X}_i$ *y* $\mathbf{X}_j$ *tienen la misma distribución.*

Teorema 7.10 (ley débil de los grandes números) *Sea* $X_1, X_2, \cdots$ *una sucesión de variables aleatorias independientes e igualmente distribuidas con media* μ *y varianza finita* σ^2. *Entonces, para todo* $\epsilon > 0$; *se satisface que:*

$$P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\sigma^2}{n\epsilon^2},$$

de donde se deduce que, para todo $\epsilon > 0$; *se tiene que:*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) = 0.$$

Demostración. *Sea* $\overline{X_n} := \frac{X_1+\cdots+X_n}{n}$ *la media aritmética de las primeras* n *variables.*

Es claro que $E\left(\overline{X_n}\right) = \mu$ *y que* $Var\left(\overline{X_n}\right) = \frac{\sigma^2}{n}$. *Por la desigualdad de Chebyschev se tiene que, para todo* $\epsilon > 0$; *se satisface:*

$$P\left(\left|\overline{X_n} - E\overline{X_n}\right| \geq \epsilon\right) \leq \frac{Var\left(\overline{X_n}\right)}{\epsilon^2}$$

que es lo que se quería demostrar. ■

Nota 7.11 *Existe una demostración de la ley débil de los grandes números que no requiere de la suposición de varianza finita. Este resultado, debido al matemático ruso Alexander Khinchin, y cuya demostración puede consultarse en [Her], establece lo siguiente:*

Si $X_1, X_2, \cdots$ *es una sucesión de variables aleatorias independientes e igualmente distribuidas, con media finita* μ, *entonces, para todo* $\epsilon > 0$; *se satisface que:*

$$\lim_{n\to\infty} P\left(\left|\frac{X_1 + \cdots + X_n}{n} - \mu\right| > \epsilon\right) = 0.$$

Como caso especial de la ley débil de los grandes números, se obtiene el siguiente resultado:

Corolario 7.12 (ley de Bernoulli) *Sea $X_1, X_2, \cdots$ una sucesión de variables aleatorias independientes e igualmente distribuidas con distribución Bernoulli con parámetro p. Entonces, para todo $\epsilon > 0$; se tiene que:*

$$P\left(\left|\frac{K_n}{n} - p\right| \geq \epsilon\right) \leq \frac{1}{4n\epsilon^2}; \tag{7.4}$$

donde $K_n := X_1 + \cdots + X_n$.

Demostración. *De la ley débil de los grandes números se obtiene que:*

$$P\left(\left|\frac{K_n}{n} - p\right| \geq \epsilon\right) \leq \frac{p(1-p)}{4n\epsilon^2}.$$

Puesto que $p \in (0,1)$, entonces, $p(1-p) \leq \frac{1}{4}$; y en consecuencia se satisface (7.4). ■

Nota 7.13 *La ley de Bernoulli establece que si se realiza un experimento aleatorio, en el cual sólo hay dos posibles resultados: éxito o fracaso, con probabilidad de éxito igual a p, un número suficientemente grande de veces, entonces, para todo $\epsilon > 0$; el conjunto de los resultados del experimento para los cuales la proporción de éxitos obtenidos dista de la probabilidad de éxito p en más de ϵ, tiende a cero. Obsérvese que la ley de Bernoulli no afirma que la frecuencia relativa de éxito converge a la probabilidad de éxito, cuando se repite el experimento un número suficientemente grande de veces. Aún cuando este resultado es también válido, no se puede deducir directamente a partir de la ley de Bernoulli.*

Nota 7.14 *La demostración de la ley de Bernoulli está basada en la desigualdad de Chebyschev. La demostración original de Bernoulli (ver [Her]), establece que, para $\epsilon > 0$ y $0 < \delta < 1$ arbitrarios; se satisface:*

$$P\left(\left|\frac{K_n}{n} - p\right| > \epsilon\right) > 1 - \delta$$

cuando $n \geq \left[\frac{(1+\epsilon)}{\epsilon^2}\right] \ln\left(\frac{1}{\delta}\right) + \frac{1}{\epsilon}$. Por ejemplo, si $\epsilon = 0.03$ y $\delta = 0.00002$, entonces, $n \geq 12416$ implica:

$$P\left(\left|\frac{K_n}{n} - p\right| > 0.03\right) > 0.99998.$$

Esto es, si se repite el experimento por lo menos 12416 veces, se puede estar seguro de que la proporción de éxitos obtenidos distará de la probabilidad *de éxito p, en menos del $3\,\%$ del total de todos los posibles resultados del experimento.*

Según [Her], Cantelli demostró un resultado más fuerte que el de Bernoulli. Cantelli probó que si $n \geq \left(\frac{2}{\epsilon^2}\right)\ln\left(\frac{4}{\delta\epsilon^2}\right) + 2 := N$, entonces:

$$P\left(\bigcap_{n=N}^{\infty}\left(\left|\frac{K_n}{n} - p\right| > \epsilon\right)\right) > 1 - \delta$$

y así, para $\epsilon = 0.03$ y $\delta = 0.00002$; se tiene que $n \geq 42711$ implica:

$$P\left(\bigcap_{n=N}^{\infty}\left(\left|\frac{K_n}{n} - p\right| > \epsilon\right)\right) > 0.99998.$$

7.2. Convergencia de sucesiones de variables aleatorias.

Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad. En esta sección se definirán tres modos de convergencia de la sucesión $(X_n)_n$ a X. Es importante aclarar, sin embargo, que hay otros tipos de convergencia diferentes a los que se consideran aquí (ver por ejemplo [Bau]).

Definición 7.15 (convergencia en probabilidad)
Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$, se dice que $(X_n)_n$ converge estocásticamente hacia a X (o en probabilidad), y se escribe:

$$X_n \xrightarrow[n\to\infty]{P} X,$$

si, para todo $\epsilon > 0$; se satisface que:

$$\lim_{n\to\infty} P\left(|X_n - X| > \epsilon\right) = 0.$$

Nota 7.16 *Haciendo uso de la definición anterior, se puede expresar la ley débil de los grandes números como sigue: Si $X_1, X_2, \cdots$ es una sucesión de variables aleatorias independientes e igualmente distribuidas con media μ y varianza finita σ^2 entonces:*

$$\overline{X_n} \xrightarrow[n\to\infty]{P} \mu$$

donde $\overline{X_n}$ es la media aritmética de las primeras n variables.

Ejemplo 7.17 *Sea* $(X_n)_{n\in\mathbb{Z}^+}$ *una sucesión de variables aleatorias tales que* $P(X_n = 0) = 1 - \frac{1}{n}$ *y* $P(X_n = n) = \frac{1}{n}$; *con* $n = 1, 2, \cdots$.

Sea $\epsilon > 0$ *arbitrario. Se tiene que:*

$$P(|X_n| > \epsilon) = \begin{cases} \frac{1}{n} & si \quad n > \epsilon \\ 0 & si \quad n \leq \epsilon \end{cases}$$

por lo tanto,

$$\lim_{n\to\infty} P(|X_n| > \epsilon) = 0,$$

esto es,

$$X_n \underset{n\to\infty}{\xrightarrow{P}} 0. \quad \blacktriangle$$

Ejemplo 7.18 *Se lanza una moneda corriente una vez. Para cada* $n = 1, 2, \cdots$; *se define la variable aleatoria* X_n *como sigue:*

$$X_n := \begin{cases} 1 & si \quad \text{el resultado del lanzamiento es sello} \\ 0 & si \quad \text{el resultado del lanzamiento es cara} \end{cases}$$

y sea X *la variable aleatoria dada por:*

$$X := \begin{cases} 1 & si \quad \text{el resultado del lanzamiento es cara} \\ 0 & si \quad \text{el resultado del lanzamiento es sello} \end{cases}$$

Es claro que, para todo $n = 1, 2, \cdots$; *se tiene:*

$$|X_n - X| = 1$$

y por lo tanto,

$$P\left(|X_n - X| > \frac{1}{2}\right) = 1$$

luego,

$$X_n \underset{n\to\infty}{\overset{P}{\not\to}} X. \quad \blacktriangle$$

Un resultado útil, para establecer la convergencia en probabilidad de una sucesión de variables aleatorias, es el teorema siguiente, cuya demostración se omite. El lector interesado, en la demostración, puede consultar [Jac].

Teorema 7.19 *Sean* X, $X_1, X_2, \cdots$ *variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad* $(\Omega, \Im, P)$. *Entonces,* $X_n \underset{n\to\infty}{\xrightarrow{P}} X$, *si y sólo si, se satisface la siguiente condición:*

$$\lim_{n\to\infty} E\left(\frac{|X_n - X|}{1 + |X_n - X|}\right) = 0.$$

Otro de los modos de convergencia que se estudiará en esta sección es la llamada convergencia casi siempre, concepto que se presenta a continuación.

Definición 7.20 (convergencia casi siempre)
Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$, se dice que $(X_n)_n$ converge casi siempre (o con probabilidad uno) hacia a X, y se escribe:

$$X_n \underset{n\to\infty}{\xrightarrow{c.s}} X,$$

si se satisface que:

$$P\left(\lim_{n\to\infty} X_n = X\right) = 1.$$

En otras palabras,

$$X_n \underset{n\to\infty}{\xrightarrow{c.s}} X, \text{ si y sólo si, } P(A) = 1,$$

donde $A := \{\omega \in \Omega : X_n(\omega) \longrightarrow X(\omega)\}$.

Ejemplo 7.21 *Sean $(\Omega, \Im, P)$ un espacio de probabilidad arbitrario y X_n las variables aleatorias definidas sobre $(\Omega, \Im, P)$ como sigue:*

$$\begin{array}{rccc} X_n : & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \longmapsto & \frac{1}{n}+1 \end{array}$$

Es claro que, para todo $\omega \in \Omega$; se satisface que:

$$\lim_{n\to\infty} X_n(\omega) = 1.$$

por lo tanto,

$$P(\{\omega \in \Omega : X_n(\omega) \longrightarrow 1\}) = P(\Omega) = 1$$

esto es,

$$X_n \underset{n\to\infty}{\xrightarrow{c.s}} 1. \quad \blacktriangle$$

7.2.1. Ley fuerte de los grandes números

El resultado siguiente, conocido como *ley fuerte de los grandes números,* es quizás la ley más conocida de la teoría de la probabilidad, ella establece que el promedio de una sucesión de variables aleatorias independientes e igualmente distribuidas converge, con probabilidad 1, a la media de la distribución.

Para demostrar la ley fuerte de los grandes números se requiere del lema siguiente cuya prueba se omite. (ver [Jac]).

Lema 7.22 *Sea $X_1, X_2, \cdots$ una sucesión de variables aleatorias.*

1. *Si todas las variables X_i son positivas, entonces,*

$$E\left(\sum_{n=1}^{\infty} X_n\right) = \sum_{n=1}^{\infty} E(X_n) \tag{7.5}$$

donde ambas expresiones en (7.5) son simultáneamente finitas o infinitas.

2. *Si $\sum_{n=1}^{\infty} E(|X_n|) < \infty$, entonces, $\sum_{n=1}^{\infty} X_n$ converge casi siempre y (7.5) se satisface.*

Teorema 7.23 (ley fuerte de los grandes números) *Sea $X_1, X_2, \cdots$ una sucesión de variables aleatorias independientes e igualmente distribuidas con media finita μ y varianza finita σ^2. Entonces,*

$$\overline{X_n} = \frac{X_1 + X_2 + \cdots X_n}{n} \underset{n \to \infty}{\xrightarrow{c.s}} \mu.$$

Demostración. Sin pérdida de generalidad se puede suponer que $\mu = 0$.

Puesto que $E\left(\overline{X_n}\right) = 0$ y

$$E\left(\left[\,\overline{X_n}\,\right]^2\right) = \frac{1}{n^2} \sum_{1 \leq j,k \leq n} E(X_j X_k) \tag{*}$$

y como para $i \neq j$, $E(X_j X_k) = 0$, entonces, (*) es igual a:

$$\begin{aligned} E\left(\left[\,\overline{X_n}\,\right]^2\right) &= \frac{1}{n^2} \sum_{n=1}^{\infty} E(X_n^2) \\ &= \frac{\sigma^2}{n}, \end{aligned}$$

por lo tanto,

$$\lim_{n \to \infty} E\left(\left[\,\overline{X_n}\,\right]^2\right) = 0.$$

Como $E\left(\left[\,\overline{X_n}\,\right]^2\right) = \frac{\sigma^2}{n}$, entonces,

$$\sum_{n=1}^{\infty} E\left(\left[\,\overline{X_{n^2}}\,\right]^2\right) = \sum_{n=1}^{\infty} \frac{\sigma^2}{n^2} < \infty.$$

Por el lema anterior se tiene que

$$\sum_{n=1}^{\infty} [\,\overline{X_{n^2}}\,]^2 < \infty; \text{ con probabilidad 1.}$$

y en consecuencia, $\lim_{n\to\infty} \overline{X_{n^2}} = 0$; con probabilidad 1.

Sean $n \in \mathbb{N}$ y k_n un entero tales que:

$$[k_n]^2 \leq n < [k_n + 1]^2 ,$$

entonces:

$$\overline{X_n} - \frac{[k_n]^2}{n}\overline{X_{k_n^2}} = \frac{1}{n} \sum_{j=k_n^2+1}^{n} X_j,$$

por lo tanto,

$$\begin{aligned} E\left(\left[\overline{X_n} - \frac{[k_n]^2}{n}\overline{X_{k_n^2}}\right]^2\right) &= \frac{n - k_n^2}{n^2}\sigma^2 \\ &\leq \frac{2k_n + 1}{n^2}\sigma^2 \\ &\leq \frac{2\sqrt{n} + 1}{n^2}\sigma^2 \\ &\leq \frac{3}{\sqrt[2]{n^3}}. \end{aligned}$$

Entonces,

$$\sum_{n=1}^{\infty} E\left(\left[\overline{X_n} - \frac{[k_n]^2}{n}\overline{X_{k_n^2}}\right]^2\right) \leq \sum_{n=1}^{\infty} \frac{3}{\sqrt[2]{n^3}} < \infty.$$

Aplicando el lema anterior se obtiene que:

$$\sum_{n=1}^{\infty} \left[\overline{X_n} - \frac{[k_n]^2}{n}\overline{X_{k_n^2}}\right]^2 < \infty; \text{ con probabilidad 1,}$$

luego,

$$\lim_{n\to\infty} \left[\overline{X_n} - \frac{[k_n]^2}{n}\overline{X_{k_n^2}}\right] = 0; \text{ con probabilidad 1.}$$

Como $\lim_{n\to\infty} \overline{X_{k_n^2}} = 0$ y $\lim_{n\to\infty} \frac{[k_n]^2}{n} = 1$, entonces, se concluye que $\lim_{n\to\infty} \overline{X_n} = 0$ con probabilidad 1, que es lo que se quería demostrar.

Si $\mu \neq 0$ basta considerar, para $i = 1, 2, \cdots$; las variables aleatorias

$$Z_i := X_i - \mu$$

y aplicar para ellas el resultado anterior. ■

La última forma de convergencia que se presenta en este texto, conocida como *convergencia en distribución*, es el modo de convergencia más usado en las aplicaciones.

Definición 7.24 (convergencia en distribución) *Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$ con funciones de distribución $F, F_1, \cdots$, respectivamente. Se dice que $(X_n)_n$ converge en distribución (o converge en ley) a X, y se escribe:*

$$X_n \xrightarrow[n\to\infty]{d} X,$$

si se satisface que:

$$\lim_{n\to\infty} F_n(x) = F(x)\,;\ \textit{para todo } x \textit{ punto de continuidad de } F$$

o, equivalentemente,

$$X_n \xrightarrow[n\to\infty]{d} X,\ \textit{si y sólo si,}\ \lim_{n\to\infty} P(X_n \leq x) = P(X \leq x)$$

para todo x punto de continuidad de F.

Obsérvese que en la convergencia en distribución, al igual que en los otros casos de convergencia considerados anteriormente, se hace referencia a la convergencia de una sucesión de números reales y no a la convergencia de una sucesión de eventos.

Ejemplo 7.25 *Sean $(\Omega, \Im, P)$ un espacio de probabilidad arbitrario y X_n las variables aleatorias definidas sobre $(\Omega, \Im, P)$ como sigue:*

$$\begin{array}{rccc} X_n: & \Omega & \longrightarrow & \mathbb{R} \\ & \omega & \longmapsto & \frac{1}{n} \end{array}$$

y $X = 0$. Puesto que:

$$F_n(x) = \begin{cases} 1 & si \quad \frac{1}{n} \leq x \\ 0 & en \quad otro\ caso \end{cases}$$

y como:

$$F(x) = \begin{cases} 1 & si \quad x \geq 0 \\ 0 & en \quad otro\ caso \end{cases}$$

entonces se satisface, para todo $x \neq 0$; *que:*

$$\lim_{n\to\infty} F_n(x) = F(x)$$

esto es,

$$X_n \xrightarrow[n\to\infty]{d} 0. \quad \blacktriangle$$

Nota 7.26 *Si* X, $X_1, X_2, \cdots$ *son variables aleatorias con valores en* $\mathbb{N}$, *entonces,*

$$X_n \xrightarrow[n\to\infty]{d} X, \text{ si y sólo si, } \lim_{n\to\infty} P(X_n = k) = P(X = k)\,; \text{ para todo } k \in \mathbb{N}.$$

El teorema siguiente relaciona la convergencia en distribución de una sucesión de variables aleatorias con la convergencia de la sucesión de funciones características. La demostración de este resultado va más allá de los objetivos propuestos para este texto. El lector interesado puede consultar [Her] o [Rao].

Teorema 7.27 (teorema de continuidad de Levy & Cramer) *Sean* $X, X_1, X_2, \cdots$ *variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad* $(\Omega, \Im, P)$ *con funciones características* $\phi, \phi_1, \phi_2 \cdots$, *respectivamente.*

1. *Si* $X_n \xrightarrow[n\to\infty]{d} X$ *,entonces,* $\lim_{n\to\infty} \phi_n(t) = \phi(t)$; *para todo* $t \in \mathbb{R}$.
2. *Si* $\lim_{n\to\infty} \phi_n(t) = \phi(t)$; *para todo* $t \in \mathbb{R}$, *y si* $\phi(\cdot)$ *es continua en* $t = 0$, *entonces,* $X_n \xrightarrow[n\to\infty]{d} X$.

A continuación se establecen las principales relaciones entre las distintas formas de convergencia que se han definido. Se verá que:

$$X_n \xrightarrow[n\to\infty]{c.s} X \Longrightarrow X_n \xrightarrow[n\to\infty]{P} X \Longrightarrow X_n \xrightarrow[n\to\infty]{d} X.$$

Teorema 7.28 *Sean* $X, X_1, X_2, \cdots$ *variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad* $(\Omega, \Im, P)$. *Entonces:*

$$X_n \xrightarrow[n\to\infty]{c.s} X \Longrightarrow X_n \xrightarrow[n\to\infty]{P} X.$$

Demostración. *Sea* $\epsilon > 0$. *Los conjuntos*

$$A_k := \{\omega \in \Omega : |X_n(\omega) - X(\omega)| < \epsilon; \text{ para todo } n \geq k\}$$

forman una sucesión creciente de eventos y su unión

$$A_\infty := \bigcup_{k=1}^{\infty} A_k$$

contiene al conjunto

$$A := \left\{\omega \in \Omega : \lim_{n\to\infty} X_n(\omega) = X(\omega)\right\}.$$

Como por hipótesis $P(A) = 1$, *entonces,* $P(A_\infty) = 1$.

Por otra parte, de la continuidad de P, *se sigue que:*

$$\lim_{k\to\infty} P(A_k) = P(A_\infty) = 1.$$

Como

$$0 \leq P(|X_k - X| \geq \epsilon) \leq P(A_k^c)$$

y

$$\lim_{k\to\infty} P(A_k^c) = 0$$

entonces, se obtiene que:

$$X_n \xrightarrow[n\to\infty]{P} X.$$

■

Teorema 7.29 *Sean* $X, X_1, X_2, \cdots$ *variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad* $(\Omega, \Im, P)$ *con funciones de distribución* $F, F_1, \cdots$, *respectivamente, entonces:*

$$X_n \xrightarrow[n\to\infty]{P} X \implies X_n \xrightarrow[n\to\infty]{d} X.$$

Demostración. *Sea* x *un punto de continuidad de* F. *Se quiere demostrar que* $\lim_{n\to\infty} F_n(x) = F(x)$.

Puesto que F *es continua en* x, *entonces, para todo* $\epsilon > 0$; *existe un* $\delta > 0$ *tal que:*

$$F(x+\delta) - F(x) < \frac{\epsilon}{2}$$

y

$$F(x) - F(x-\delta) < \frac{\epsilon}{2}.$$

Por otra parte, como $X_n \xrightarrow[n\to\infty]{P} X$, *entonces, existe* $n(\epsilon) \in \mathbb{N}$ *tal que, para* $n \geq n(\epsilon)$; *se satisface:*

$$P(|X_n - X| > \delta) < \frac{\epsilon}{2}.$$

Se tiene que:

$$\begin{aligned}(X_n \leq x) &= [(X_n \leq x) \cap (|X_n - X| \leq \delta)] \cup [(X_n \leq x) \cap (|X_n - X| > \delta)] \\ &\subseteq (X \leq x + \delta) \cup (|X_n - X| > \delta).\end{aligned}$$

Análogamente:

$$(X \leq x - \delta) \subseteq (X_n \leq x) \cup (|X_n - X| > \delta),$$

por lo tanto,

$$\begin{aligned}F(x) - \epsilon &< F(x - \delta) - \frac{\epsilon}{2} \\ &\leq P(X_n \leq x) + P(|X_n - X| > \delta) - \frac{\epsilon}{2} \\ &\leq P(X_n \leq x) \\ &\leq F(x + \delta) + \frac{\epsilon}{2} \\ &< F(x) + \epsilon.\end{aligned}$$

Es decir, para todo $n \geq n(\epsilon)$; *se satisface que:*

$$|F_n(x) - F(x)| < \epsilon$$

por consiguiente,

$$\lim_{n\to\infty} F_n(x) = F(x).$$

■

Nota 7.30 *El recíproco del teorema anterior no es, en general, válido. Considérese, por ejemplo, las variables aleatorias dadas en el ejemplo 7.18. En ese caso se tiene que* $X_n \xrightarrow[n\to\infty]{d} X$, *pero* $X_n \underset{n\to\infty}{\overset{P}{\nrightarrow}} X$.

Las definiciones de convergencia dadas anteriormente pueden ser generalizadas a vectores aleatorios de la manera siguiente:

Definición 7.31 (convergencia de vectores aleatorios)
Sean $\mathbf{X}, \mathbf{X}_1, \mathbf{X}_2, \cdots$ *vectores aleatorios* $k-$*dimensionales,* $k \in \mathbb{N}$, *con funciones de distribución* $F_{\mathbf{X}}, F_{\mathbf{X}_1}, F_{\mathbf{X}_2}, \cdots$, *respectivamente. Entonces, se dice que:*

1. $(\mathbf{X}_n)_n$ *converge estocásticamente hacia a* $\mathbf{X}$ *(o en probabilidad), y se escribe:* $\mathbf{X}_n \underset{n\to\infty}{\xrightarrow{P}} \mathbf{X}$, *si la* $i-$*ésima componente de* $\mathbf{X}_n$ *converge estocásticamente a la* $i-$*ésima componente de* $\mathbf{X}$; *para* $i = 1, \cdots, k$.

2. $(\mathbf{X}_n)_n$ *converge casi siempre hacia a* $\mathbf{X}$ *(o con probabilidad* 1*), y se escribe:* $\mathbf{X}_n \underset{n\to\infty}{\xrightarrow{c.s}} \mathbf{X}$, *si la* $i-$*ésima componente de* $\mathbf{X}_n$ *converge casi siempre a la* $i-$*ésima componente de* $\mathbf{X}$; *para* $i = 1, \cdots, k$.

3. $(\mathbf{X}_n)_n$ *converge en distribución hacia a* $\mathbf{X}$ *(o en ley), y se escribe:* $\mathbf{X}_n \underset{n\to\infty}{\xrightarrow{d}} \mathbf{X}$, *si* $\lim\limits_{n\to\infty} F_{\mathbf{X}_n}(\mathbf{x}) = F_{\mathbf{X}}(\mathbf{x})$, *para todo* $\mathbf{x} = (x_1, \cdots, x_k) \in \mathbb{R}^k$ *punto de continuidad de* $F_{\mathbf{X}}(\cdot)$.

7.3. Teorema Central del Límite

En esta sección se presenta uno de los resultados más importantes de la teoría de probabilidad. Este resultado, conocido como Teorema Central del Límite (también llamado Teorema del Límite Central), fue demostrado, por primera vez, en el año 1733, por el matemático francés Abraham DeMoivre. En el año 1812 el también matemático francés Pierre Simon Laplace, demostró una versión más general del teorema. La demostración rigurosa del teorema, tal y como se la conoce hoy en día , fue presentada por el matemático ruso Liapounoff en el año 1901.

El Teorema Central del Límite establece que la media aritmética de variables aleatorias independientes e igualmente distribuidas tiene, aproximadamente, una distribución normal cuando el número de variables aleatorias involucradas es "grande" y cuando la varianza es finita y diferente de *cero*.

Teorema 7.32 (teorema central del límite univariado) *Sea* X_1, X_2, $\cdots$ *una sucesión de variables aleatorias independientes e igualmente distribuidas con media* μ *y varianza finita y positiva* σ^2. *Sean* $S_n := \sum\limits_{j=1}^{n} X_j$ *y* $Y_n := \frac{S_n - n\mu}{\sigma\sqrt{n}}$. *Entonces, la sucesión de variables* $Y_1, Y_2, \cdots$ *converge en distribución a una variable aleatoria* Y, *donde* Y *tiene distribución normal estándar.*

Esto es,

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x); \text{ para todo } x \in \mathbb{R}.$$

Demostración. Se puede suponer, sin pérdida de generalidad, que $\mu = 0$. Sea ϕ la función característica de las variables aleatorias $X_1, X_2, \cdots$.

Como las variables aleatorias son independientes se obtiene que:

$$\begin{aligned}
\phi_{Y_n}(t) &= E\left(\exp\left[itY_n\right]\right) \\
&= E\left(\exp\left[it\frac{X_1 + \cdots X_n}{\sigma\sqrt{n}}\right]\right) \\
&= E\left(\prod_{j=1}^{n} \exp\left[it\frac{X_j}{\sigma\sqrt{n}}\right]\right) \\
&= \prod_{j=1}^{n} E\left(\exp\left[it\frac{X_j}{\sigma\sqrt{n}}\right]\right) \\
&= \prod_{j=1}^{n} \phi\left(\frac{t}{\sigma\sqrt{n}}\right) \\
&= \left[\phi\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n
\end{aligned}$$

Al hacer el desarrollo en serie de Taylor de ϕ alrededor de *cero* se obtiene:

$$\phi(t) = 1 + 0 - \frac{\sigma^2 t^2}{2} + t^2 o(t)$$

Por lo tanto,

$$\begin{aligned}
\phi_{Y_n}(t) &= \left[1 - \frac{\sigma^2}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \left(\frac{t}{\sigma\sqrt{n}}\right)^2 o(t)\right]^n \\
&= \exp\left[n \ln\left(1 - \frac{\sigma^2}{2}\left(\frac{t}{\sigma\sqrt{n}}\right)^2 + \left(\frac{t}{\sigma\sqrt{n}}\right)^2 o(t)\right)\right]
\end{aligned}$$

tomando límite, cuando $n \to \infty$, se obtiene:

$$\lim_{n\to\infty} \phi_{Y_n}(t) = \exp\left[-\frac{t^2}{2}\right].$$

Por el teorema de continuidad de Levy & Cramer 7.27 se concluye que

$$Y_n \xrightarrow{d} Y,$$

donde Y es una variable aleatoria con distribución normal estándar. ∎

Nota 7.33 *El teorema de Moivre-Laplace 4.21 es un caso particular del Teorema Central del Límite ya que, si* $X_1, X_2, \cdots$ *son variables aleatorias independientes e igualmente distribuidas con distribución Bernoulli de parámetro* p*, entonces, se satisfacen las condiciones dadas en 7.32.*

Nota 7.34 *Sea* X_1, X_2, $\cdots$ *una sucesión de variables aleatorias independientes e igualmente distribuidas con media* μ *y varianza finita y positiva* σ^2*. El Teorema Central del Límite afirma que, para* n *suficientemente grande, la variable aleatoria* $S_n := \sum_{j=1}^{n} X_j$ *tiene, aproximadamente, una distribución normal de parámetros* $n\mu$ *y* $n\sigma^2$*.*

Nota 7.35 *El Teorema Central del Límite puede ser aplicado a la mayoría de las distribuciones clásicas como por ejemplo: distribución binomial, distribución Poisson, binomial negativa, gamma, Weibull, etc. pues ellas satisfacen las hipótesis del teorema. Sin embargo, no puede ser aplicado a la distribución Cauchy pues ella no satisface las condiciones dadas en éste.*

Nota 7.36 *Existen varias generalizaciones del teorema Central del Límite, dentro de las cuales se destaca el teorema de Lindenberg- Feller que permite que las variables aleatorias involucradas tengan medias y varianzas diferentes (Ver [Bau]).*

Nota 7.37 (teorema central del límite multivariado) *El Teorema del Límite Central puede ser generalizado a sucesiones de vectores aleatorios como sigue:*

Sea $\mathbf{X}_1, \mathbf{X}_2, \cdots$ *una sucesión de vectores aleatorios* $k-$*dimensionales,* $k \in \mathbb{N}$*, independientes e igualmente distribuidos con vector de medias* μ *y matriz de varianzas y covarianzas* Σ*, donde* Σ *es definida positiva. Sea*

$$\overline{\mathbf{X}_n} := \frac{\mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_n}{n}$$

el vector de medias aritméticas. Entonces:

$$\sqrt{n}\left(\overline{\mathbf{X}_n} - \mu\right) \xrightarrow{d} \mathbf{X},$$

donde $\mathbf{X}$ *es un vector aleatorio* $k-$*dimensional con distribución normal multivariada con vector de medias* $\mathbf{0}$ *y matriz de varianzas y covarianzas* Σ*. (ver [Her]).*

Algunas aplicaciones del Teorema Central del Límite son presentadas en los ejemplos siguientes.

Ejemplo 7.38 *Se lanza un dado corriente* 1000 *veces. Calcular la probabilidad de que el número* 4 *aparezca por lo menos* 150 *veces.*

Solución: Sea $X :=$ "Número de veces que se obtiene 4 como resultado". Se sabe que $X \stackrel{d}{=} \mathcal{B}\left(1000,\ \frac{1}{6}\right)$. Aplicando el Teorema de Moivre-Laplace se puede afirmar que X tiene, aproximadamente, una distribución normal de media $\frac{1000}{6}$ y varianza $\frac{5000}{36}$. Por lo tanto,

$$\begin{aligned} P(X \geq 150) &= P\left(\frac{X - \frac{500}{3}}{\frac{25\sqrt{2}}{3}} \geq \frac{150 - \frac{500}{3}}{\frac{25\sqrt{2}}{3}}\right) \\ &= 1 - \Phi(-1.4142) \\ &= 1 - 0.07780 \\ &= 0.9222. \quad \blacktriangle \end{aligned}$$

Ejemplo 7.39 *Un elevador de carga grande puede transportar un máximo de* 10000 *libras. Supóngase que una carga, que contiene* 45 *cajas, se debe transportar mediante el elevador. La experiencia ha demostrado que el peso* X*, de una caja de este tipo de carga, se ajusta a una distribución de probabilidad con una media de* $\mu = 200$ *libras y una desviación estándar de* $\sigma = 55$ *libras. Calcular la probabilidad de que las* 45 *cajas se puedan transportar simultáneamente en el elevador.*

Solución: Sea $X_i :=$ "peso de la i-ésima caja", $i = 1, \cdots, n$. Por los datos del problema se sabe que las variables aleatorias $X_1, X_2, \cdots$ son independientes e igualmente distribuidas con media $\mu = 200$ y varianza $\sigma^2 = 3.025$.Se desea calcular $P\left(\sum\limits_{i=1}^{45} X_i \leq 10000\right)$. Por el Teorema Central del Límite se sabe que $\sum\limits_{i=1}^{45} X_i$ tiene, aproximadamente, una distribución normal de media 9000 y desviación estándar 5500. Por lo tanto,

$$
\begin{aligned}
P\left(\sum_{i=1}^{45} X_i \leq 10000\right) &= P\left(\frac{\sum\limits_{i=1}^{45} X_i - 9000}{55\sqrt{45}} \leq \frac{10000 - 9000}{55\sqrt{45}}\right) \\
&= P\left(\frac{\sum\limits_{i=1}^{45} X_i - 9000}{55\sqrt{45}} \leq 2.71\right) \\
&= \Phi\left(2.71\right) \\
&= 0.9966. \quad \blacktriangle
\end{aligned}
$$

Ejemplo 7.40 *Muchos insumos de producción, como el mineral de hierro, el carbón y el azúcar sin refinar, se muestrean, para determinar su calidad, por un método que implica la toma periódica de muchas pequeñas muestras cuando el material se mueve sobre una banda transportadora. Posteriormente las muestras pequeñas se juntan y mezclan para formar una muestra compuesta. Sea Y_i el volumen de la $i-$ésima muestra pequeña de un lote particular y supóngase que $Y_1, Y_2, ..., Y_n$ es una muestra aleatoria, en donde cada Y_i tiene media μ (en pulgadas cúbicas) y varianza σ^2. El volumen promedio de las muestras, μ, se puede regular ajustando el tamaño del equipo que se utiliza para el muestreo. Supóngase que la varianza de los volúmenes de las muestras, σ^2, es, aproximadamente, 4 para una situación particular. Se requiere que el volumen total de la muestra exceda las 200 pulgadas cúbicas con una probabilidad de 0.95 cuando se seleccionan $n = 50$ muestras pequeñas. Determinar el ajuste de μ que permitirá satisfacer los requerimientos del muestreo.*

Solución: Las variables aleatorias $Y_1, Y_2, \cdots$ son independientes e igualmente distribuidas con media μ y varianza $\sigma^2 = 4$. Por el Teorema Central del Límite, se tiene que $\sum\limits_{i=1}^{50} Y_i$ tiene, aproximadamente, una distribución normal de media 50μ y varianza 200. Se desea determinar

μ de tal manera que $P\left(\sum_{i=1}^{50} Y_i > 200\right) = 0.95$. Se tiene que:

$$\begin{aligned} 0.95 &= P\left(\sum_{i=1}^{50} Y_i > 200\right) \\ &= P\left(\frac{\sum_{i=1}^{50} Y_i - 50\mu}{10\sqrt{2}} > \frac{200-50\mu}{10\sqrt{2}}\right) \\ &= 1 - \Phi\left(\frac{200-50\mu}{10\sqrt{2}}\right). \end{aligned}$$

esto es,

$$\Phi\left(\frac{200-50\mu}{10\sqrt{2}}\right) = 0.05,$$

lo cual implica que,

$$\frac{200-50\mu}{10\sqrt{2}} = -1.64$$

es decir,

$$\mu = 4.46. \quad \blacktriangle$$

7.4. Ejercicios

1. Se lanza una moneda corriente 100 veces consecutivas. Sea X el número de caras obtenidas. Usar la desigualdad de Chebyschev para encontrar una cota inferior de la probabilidad de que $\frac{X}{100}$ difiera de $\frac{1}{2}$ en menos de 0.1.

2. Sean $X_1, X_2, ..., X_{50}$ variables aleatorias independientes e igualmente distribuidas con distribución Poisson de media 1. Calcular
$$P(\sum_{i=1}^{50} X_i > 15).$$

3. **(Cotas de Chernoff)** Supóngase que la función generadora de momentos $m_X(t)$ de una variable aleatoria X existe. Usar la desigualdad de Markov para demostrar que, para cualquier $a \in \mathbb{R}$; se satisface que:

   a) $P(X \geq a) \leq \exp(-ta)\, m_X(t)$ para todo $t > 0$.

b) $P(X \leq a) \leq \exp(-ta)\, m_X(t)$ para todo $t < 0$.

4. **(Desigualdad de Jensen)** Sea f una función convexa, esto es, f es una función real dos veces diferenciable tal que $f''(x) \geq 0$; para todo x. Demostrar que si X es una variable aleatoria real, entonces,

$$E(f(X)) \geq f(E(X)),$$

siempre y cuando los valores esperados existan.

Sugerencia: Hallar el desarrollo en serie de Taylor alrededor de $\mu = EX$ de la función f.

5. Si X es una variable aleatoria no negativa con media 2, ¿qué se puede decir acerca de EX^3 y de $E(\ln X)$?

6. Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$. Demostrar las propiedades siguientes de la convergencia estocástica:

a) $X_n \xrightarrow[n\to\infty]{P} X$, si y sólo si, $(X_n - X) \xrightarrow[n\to\infty]{P} 0$.

b) Si $X_n \xrightarrow[n\to\infty]{P} X$ y $X_n \xrightarrow[n\to\infty]{P} Y$, entonces, $P(X = Y) = 1$.

c) Si $X_n \xrightarrow[n\to\infty]{P} X$, entonces, $(X_n - X_m) \xrightarrow[n,m\to\infty]{P} 0$.

d) Si $X_n \xrightarrow[n\to\infty]{P} X$ y $Y_n \xrightarrow[n\to\infty]{P} Y$, entonces, $(X_n + Y_n) \xrightarrow[n\to\infty]{P} (X + Y)$.

e) Si $X_n \xrightarrow[n\to\infty]{P} X$ y k es una constante, entonces, $kX_n \xrightarrow[n\to\infty]{P} kX$.

f) Si $X_n \xrightarrow[n\to\infty]{P} X$ y $Y_n \xrightarrow[n\to\infty]{P} Y$, entonces, $X_nY_n \xrightarrow[n\to\infty]{P} XY$.

7. Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$. Demostrar que:

a) Si $X_n \xrightarrow[n\to\infty]{c.s} X$ *y* $Y_n \xrightarrow[n\to\infty]{c.s} Y$, entonces, $(X_n + Y_n) \xrightarrow[n\to\infty]{c.s} (X + Y)$.

b) Si $X_n \xrightarrow[n\to\infty]{c.s} X$ *y* $Y_n \xrightarrow[n\to\infty]{c.s} Y$, entonces, $X_nY_n \xrightarrow[n\to\infty]{c.s} XY$.

8. Sean f una función real continua y $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$ .Demostrar que si $X_n \underset{n\to\infty}{\xrightarrow{c.s}} X$, entonces, $f(X_n) \underset{n\to\infty}{\xrightarrow{c.s}} f(X)$.

   Nota: *El resultado sigue siendo válido si en lugar de convergencia casi siempre se considera convergencia estocástica.*

9. Sean $X, X_1, X_2, \cdots$ variables aleatorias reales definidas todas sobre el mismo espacio de probabilidad $(\Omega, \Im, P)$. Demostrar que si $X_n \underset{n\to\infty}{\xrightarrow{d}} k$, donde k es una constante; entonces $X_n \underset{n\to\infty}{\xrightarrow{P}} k$.

10. Supóngase que se tiene una sucesión de Bernoulli de longitud 100 con probabilidad de éxito $p = 0.7$. Sea $X :=$ "número de éxitos obtenidos". Calcular $P(X \in [65, 80])$.

11. Se sabe que una cierta medicina produce reacciones alérgicas en el $1\,\%$ de las personas. Si se aplica esta medicina a un grupo de 500 personas, ¿cuál es la probabilidad de que a lo más 10 de ellas presenten reacciones alérgicas?

12. Al sumar números, un computador aproxima cada número al entero más próximo. Supóngase que todos los errores de aproximación son independientes y distribuidos uniformemente sobre el intervalo $(-0.5, 0.5)$.

    a) Si se suman 1500 números, ¿cuál es la probabilidad de que la magnitud del error total sea mayor que 15?

    b) ¿Cuántos números deben sumarse juntos para que la magnitud del error total sea menor que 10, con probabilidad 0.90?

13. Se lanza un dado corriente tantas veces sea necesario hasta que la suma de los resultados sea mayor que 200. ¿Cuál es la probabilidad de que sean necesarios por lo menos 50 lanzamientos?

14. Una compañia tabacalera afirma que la cantidad de nicotina en sus cigarrillos es una variable aleatoria con media de 2,2 mg y desviación estándar de 3 mg. Sin embargo en 100 cigarrillos, tomados en forma aleatoria, la media muestral de contenido de nicotina fue de 3.1 mg. Si lo que dice la compañía es verdad, ¿cuál es la probabilidad aproximada de que la media muestral sea tan alta como 3.1 o mayor?

15. Los ingenieros civiles creen que el peso W (en miles de libras), que puede soportar un puente vehicular sin que su estructura sufra daño,

tiene una distribución con media 400 y desviación estándar de 40. Supóngase que el peso de un auto (también en miles de libras) es una variable aleatoria con media 3 y desviación estándar de 0.3, ¿cuántos autos tendría que haber en el puente para que la probabilidad de que su estructura sufra daño sea mayor de 0.1?

16. Un profesor sabe por experiencia, que las calificaciones en exámenes de los estudiantes tienen media 77 y desviación estándar 15. El profesor tiene a su cargo este semestre dos grupos: uno con 25 estudiantes y otro con 64.

    a) Aproximar la probabilidad de que en el curso de 25 estudiantes la calificación promedio en exámenes esté entre 72 y 82.
    b) Repetir la parte a. con el curso de 64 estudiantes.
    c) ¿Cuál es la probabilidad de que la calificación promedio en el curso con 25 estudiantes sea mayor que en el curso con 64 estudiantes?
    d) Si se sabe que los promedios de las calificaciones de los cursos son 76 y 83, ¿cuál de los dos cursos, el de 25 estudiantes o el de 64, cree usted que tiene mayor probabilidad de tener el promedio de 83?

17. Sea X una variable aleatoria con distribución gamma de parámetros n y 1.¿Qué tan grande debe ser n para que
$$P\left(\left|\frac{X}{n}-1\right|>0.01\right)<0.01?$$

18. ¿Cuántos lanzamientos de una moneda corriente se requieren para que la probabilidad de que el número promedio de caras obtenidas difiera de 0.5 en a lo más 0.01, sea de por lo menos 0.90?

19. Una máquina que se utiliza para llenar cajas de cereales descarga un promedio de μ onzas por caja. El fabricante desea que la carga real en onzas, X, quede a una onza o menos de μ en por lo menos el 75 % de los casos. Hallar el mayor valor de la desviación estándar σ de X que se puede admitir, si se deben cumplir los requisitos del fabricante.

20. Sea X una variable aleatoria no negativa. Demostrar:
$$E(X)\leq\left[E\left(X^{2}\right)\right]^{\frac{1}{2}}\leq\left[E\left(X^{3}\right)\right]^{\frac{1}{3}}\leq\cdots$$

21. Sea $(X_n)_{n\geq 1}$ una sucesión de variables aleatorias y c una constante tal que, para todo $\epsilon > 0$; se satisface:

$$\lim_{n\to\infty} P\left(|X_n - c| > \epsilon\right) = 0.$$

Demostrar que para cualquier función acotada continua g se satisface:

$$\lim_{n\to\infty} E\left(g\left(X_n\right)\right) = g\left(c\right).$$

Capítulo 8

Simulaciones básicas

En este capítulo se presentan algunos métodos computacionales que permiten simular modelos aleatorios. Se comenzará explicando el método computacional usado en la generación de números aleatorios. Luego se verá cómo utilizar los números aleatorios para simular experimentos tipo Bernoulli, probabilidades de eventos, distribuciones discretas con un número finito de resultados, distribuciones discretas con un número infinito numerable de resultados y variables aleatorias continuas.

8.1. Generación de números aleatorios.

El núcleo de las simulaciones de modelos aleatorios es la capacidad de generar *números aleatorios*, los cuales representan el valor de una variable aleatoria con distribución uniforme.

En un comienzo los números aleatorios se generaban de manera manual utilizando por ejemplo, ruedas giratorias, lanzamientos de dados, barajas, etc. (ver [Ros3]). Actualmente se utilizan métodos computacionales para generarlos.

Un generador de números aleatorios es una fórmula específica que produce números aleatorios en una forma completamente determinística, lo cual es en un principio una contradicción con la idea original de aletoriedad. Es por esta razón que a los números aleatorios generados bien sea, por métodos manuales o computacionales, se les llama frecuentemente *números seudoaleatorios.*

Uno de los métodos más comunes para generar números aleatorios comienza fijando enteros positivos m (multiplicador), a (sumando) y $norma$ (para normalizar). Se inicializa el proceso con un número llamado *"semilla"* el cual satisface que:

$$0 \leq semilla < norma,$$

luego se generan a partir de este valor inicial los demás números aleatorios mediante la fórmula:

$$semilla := (m * semilla + a) \bmod norma. \tag{8.1}$$

Los números *semilla*, m, a y *norma* deben satisfacer ciertas condiciones, las cuales se expondrán a continuación. En primer lugar, la *semilla* debe ser mayor que 1 y no debe exceder el máximo entero que puede ser almacenado por el computador. Puesto que se desea que la secuencia de números sea distinta cada vez que se inicie el generador de números, entonces la *semilla* debe ser un número aleatorio. Una forma de escoger una *semilla*, de tal manera que se garantice la generación de secuencias diferentes, es utilizando el conocido comando "clock", el cual dá como resultado un vector de 6 componentes las cuales representan, respectivamente, el año, mes, día, hora, minuto y segundo, y puesto que cada combinación cambia cada segundo durante todo el año, entonces se generan cada segundo, semillas distintas.

Se observa además que puesto que el siguiente número es generado a partir del anterior, entonces si cualquier número es generado nuevamente, la secuencia completa se repite cíclicamente. Por otra parte, la longitud del ciclo puede ser a lo más igual al valor de *norma*, ya que como la función mod da como resultado el residuo de la división por *norma*, entonces hay *norma* posibles residuos, esto es, *norma* posibles valores del número aleatorio *semilla* dado en 8.1, y por lo tanto una vez uno de los números se repite, la lista completa se inicia nuevamente. Por consiguiente, un buen generador de números aleatorios debe usar números m, a y *norma* que garanticen ciclos de longitud grande, para lograrlo, estos números deben ser escogidos de tal manera que:

$$m * (norma - 1) + a$$

no exceda el máximo entero que puede ser almacenado por el computador (ver [Sol]).

Actualmente los lenguajes de programación y paquetes de cómputo tales como C, Pascal, Fortran, MATLAB, SAS, MAPLE, etc. disponen de funciones que permiten generar números aleatorios.

A continuación se presenta el algoritmo para la generación de números aleatorios descrita arriba. El algoritmo arroja una tabla de números aleatorios que termina cuando se han generado máx números o cuando se repite la semilla inicial.

Algoritmo 8.1 (generación de números aleatorios enteros)

Entrada: a, m, *norma*, *semilla*, máx

Salida: *RANDZ (número aleatorio entero no negativo < norma)*

Inicialización:

$RANDZ := semilla,$
escriba: semilla $= RANDZ.$

Iteración:

para $i = 1(1)\,\text{máx}$ *haga:*
  $RANDZ := (m * RANDZ + a) \bmod norma$
  escriba $RANDZ$
  si $RANDZ = semilla$ *entonces* $PARE$ ▲

Ejemplo 8.2 *Si se eligen los valores siguientes de entrada:*

$$a = 17,\ m = 31,\ norma = 231475,\ semilla = 0,\ \text{máx} = 15$$

el algoritmo produce esta tabla:

semilla	0
1	17
2	544
3	16881
4	60378
5	19935
6	155052
7	177129
8	167091
9	87388
10	162820
11	186462
12	224939
13	28876
14	200748
15	204855

Como se comprobará más adelante, la tarea central es la generación de números aleatorios entre 0 y 1, ya que a partir de ellos es posible la generación de cualquier otro real. Un sencillo algoritmo para tal fin es el siguiente:

Algoritmo 8.3 *(generación de números aleatorios reales entre* 0 *y* 1*)*

Entrada: a, m, $norma$, $semilla$, máx

Salida: RND *(número aleatorio real en* $[0, 1)$*)*

Inicialización:

$S := semilla,$

escriba: semilla $=$ S.

Iteración:

para $i = 1(1)\,\text{máx}$ *haga:*

$S := (m * S + a) \bmod norma$

$RND := \frac{S}{norma}$

escriba RND

si $S = semilla$ *entonces* $PARE$ ▲

Ejemplo 8.4 *Con los mismos datos de entrada del ejemplo anterior,*

$$a = 17,\ m = 31,\ norma = 231475,\ semilla = 0,\ \text{máx} = 15$$

se produce la siguiente tabla:

semilla	0
1	0.0000734421
2	0.0023501458
3	0.0729279620
4	0.2608402635
5	0.0861216114
6	0.6698433956
7	0.7652187061
8	0.7218533319
9	0.3775267307

```
10          0.7034020953
11          0.8055383951
12          0.9717636894
13          0.1247478129
14          0.8672556432
15          0.8849983800
```

Como ilustración sencilla de la aplicación que puede tener el algoritmo 8.3, se presenta a continuación un algoritmo que simula la generación de un experimento del tipo Bernoulli con probabilidad de éxito $p \in (0,1)$.

Algoritmo 8.5 (generación de un experimento Bernoulli)

Entrada: $p \in (0,1)$, máx

Salida: 0 o 1.

Iteración:

para $i = 1(1)$ máx haga:

genere un número aleatorio $RND \in [0,1)$

si $RND < p$ entonces

escriba 1

en caso contrario

escriba 0 ▲

Ejemplo 8.6 *Si se toma* $p = 0.4$, *y se realiza el experimento un número de veces* máx $= 15$, *un típico resultado de la aplicación del algoritmo anterior es el siguiente:*

```
0  1  0  0  0  0  0  1  0  0  1  0  0  1  1
```

8.2. Simulación de probabilidades y de valores esperados.

La interpretación intuitiva de probabilidad es que si un experimento aleatorio laplaciano, es realizado un número suficientemente grande de veces N, entonces la probabilidad de que un evento A ocurra n veces es igual al cociente $\frac{n}{N}$. Para simular $P(A)$ se procede como sigue: se introduce el valor de N y se inicializa con un valor *suma* igual a *cero*. Luego se simula el experimento aleatorio N veces y se suma 1 a *suma* cada vez que la simulación arroje como resultado un elemento de A. Se da como salida a $\frac{suma}{N}$. En los ejemplos siguientes, se ilustra el procedimiento.

Nota 8.7 *En adelante, cuando aparezca RND en un algoritmo, se debe generar un nuevo número aleatorio entre* 0 *y* 1.

Ejemplo 8.8 *En un clóset hay dispuestos dos cajones con zapatos. El primero contiene* 10 *zapatos derechos y el segundo, los* 10 *zapatos izquierdos compañeros. Se sacan uno a uno,* 8 *zapatos del primer cajón y* 8 *zapatos del segundo cajón. ¿Cuál es la probabilidad de que al comparar el $i-$ésimo zapato sacado del primer cajón con el $i-$ésimo zapato extraído del segundo cajón no se obtenga nunca un par completo?*

Solución: El algoritmo siguiente permite hacer la simulación requerida:

Algoritmo 8.9

Entrada: máx (número de ensayos)

Salida: probabilidad de no tener algún par correcto, de los 8 formados.

Inicialización:

$suma := 0$

Sean $D := (d_1, ..., d_8)$, $Z := (z_1, ..., z_8)$ vectores de $\mathbb{R}^8$ (derecho, izquierdo), definidos como $D := 0 =: Z$.

Iteración:

para $i = 1(1)$ máx haga:

Inicialización $i - ésima$:

elección del primer zapato izquierdo

$z_1 := [10 * RND + 1]$ ([$\cdot$] denota la parte entera),

elección del primer zapato derecho

$d_1 := [10 * RND + 1]$

Elección de los demás zapatos izquierdos:

para $j = 2(1)8$ haga:

$z_j := [10 * RND + 1]$

$control := 1;$ (para evitar que se saque de nuevo el mismo zapato izquierdo)

mientras $control = 1$ haga:

$control := 0;$

para $r = 1(1)(j-1)$ haga:

si $z_r = z_j$ entonces

$z_j := [10 * RND + 1]$

$control := 1$

Elección de los demás zapatos derechos:

para $j = 2(1)8$ haga:

$d_j := [10 * RND + 1]$

$control := 1;$ (para evitar que se saque de nuevo el mismo zapato derecho)

mientras $control = 1$ haga:

$control := 0;$

para $r = 1(1)(j-1)$ haga:

si $d_r = d_j$ entonces

$d_j := [10 * RND + 1]$

$control := 1$

comparación de zapatos:

para $j = 1(1)8$ haga

si $z_j = d_j$ entonces

$X := 0$

$PARE.$

en caso contrario

$X := 1$

conteo de aciertos

$suma := suma + X$

$probabilidad := \frac{suma}{\text{máx}}$ ▲

Los siguientes son algunos de los resultados obtenidos para máx = 10000.

0.4483, 0.4461, 0.4413, 0.4481, 0,4383, 0.4464, 0.4424.

Ejemplo 8.10 *En un clóset hay* 10 *pares de zapatos. Se sacan* 8 *zapatos al azar. ¿Cuál es la probabilidad de no sacar ningún par correcto?*

Solución: El algoritmo siguiente da la solución aproximada.

Algoritmo 8.11

Entrada: máx (número de ensayos)

Salida: probabilidad de no poder formar ningún par correcto con los 8 zapatos extraídos.

Inicialización:

$suma := 0$

Sea $Z := (z_1, ..., z_8) \in \mathbb{R}^8$, definido como $Z := 0$.

Iteración:

para $i = 1(1)$ máx haga:

Inicialización $i - ésima$:

elección del primer zapato $z_1 := [20 * RND + 1]$ $([\,\cdot\,]$ denota la parte entera),

Elección de los demás zapatos:

para $j = 2(1)8$ haga:

$z_j := [20 * RND + 1]$

$control := 1$; (para evitar que se saque de nuevo el mismo zapato)

mientras $control = 1$ haga:

$control := 0$;

para $r = 1(1)(j-1)$ haga:

si $z_r = z_j$ entonces

$z_j := [20 * RND + 1]$

$control := 1$

Comparación de los zapatos extraídos: Se supone que los números pares corresponden a zapatos derechos y los números impares a zapatos izquierdos. Los pares de zapatos compañeros se toman como: $(1,2), (3,4), ..., (19,20)$.

$M := 0$ (contador)

$X := 0$ (contador de pares de zapatos encontrados)

para $j = 2(1)8$ haga:

para $r = 1(2)19$ haga:
si $z_j = r$ entonces
para $q = 1(1)8$ haga:
si $z_q = r + 1$ entonces
$X := X + 1$

si $X = 0$ entonces
$M := 1$
$suma := suma + M$

$probabilidad := \frac{suma}{\text{máx}}$ ▲

Con máx $= 10000$, *estos son los resultados de algunos ensayos:*

$$0.0880, 0.0904, 0.0960, 0.0875, 0.0935, 0.0891.$$

Sea Ω el espacio muestral de un experimento aleatorio y X una variable aleatoria real, definida sobre Ω, cuyo valor esperado existe. Se sabe de la ley fuerte de los grandes números que EX es, aproximadamente, igual a la media aritmética de los valores de X cuando el experimento se repite un número suficientemente grande de veces.

Para simular EX se introduce un valor m y se inicializa con un valor *sum* igual a *cero*. Luego, se simula el experimento aleatorio en consideración m veces y se añade el valor de X a *sum*. Finalmente se da como salida $\frac{sum}{m}$.

Ejemplo 8.12 *Supóngase que X es una variable aleatoria con distribución Bernoulli de parámetro p. En 8.5 se dio un algoritmo para simular valores de X. Al sumar el número de unos obtenidos en cada secuencia y dividir por el número de valores simulados, se obtiene una aproximación del valor esperado de X, el cual es igual a p.*

El procedimiento a seguir aparece dado en el algoritmo siguiente:

Algoritmo 8.13

Entrada: $p \in (0, 1)$, máx

Salida: probabilidad de tener: $RND < p$.

Inicialización:

$suma := 0$

Iteración:

para $i = 1(1)\,\text{máx}$ haga:

genere un número aleatorio $RND \in (0,1)$

si $RND < p$ entonces

$suma := suma + 1$

$valor\ esperado = \frac{suma}{\text{máx}}$ ▲

Ejemplo 8.14 *Sean $p = 0.4$ y $\text{máx} = 10.000$. El valor encontrado en este caso es:*

$$valor\ esperado = 0.4028.$$

La repetición del procedimiento arroja estos otros valores:

$$0.4022, \quad 0.3955, \quad 0.4039, \quad 0.3961. \qquad \blacktriangle$$

8.3. Simulación de variables aleatorias discretas con un número finito de resultados.

Supóngase que se tiene un experimento aleatorio que tiene $N+1$ posibles resultados, con probabilidades de ocurrencia $p(i)$, $i = 0, \cdots, N$. Es claro que $p(0) + \cdots + p(N) = 1$.

Se generan números aleatorios en el intervalo $[0,1)$. Puesto que la fracción $p(0)$ de ellos estará en el intervalo $[0, p(0))$, la fracción $p(1)$ de los RND estará en el intervalo $[p(0), p(0) + p(1))$ y así sucesivamente, entonces para simular un resultado i, se usa el siguiente procedimiento: se genera RND y se determina en cuál subintervalo de $[0,1)$ se encuentra, si RND es menor que $p(0)$, se toma $i = 0$, si $RND < p(0) + p(1)$ entonces se toma $i = 1$, y así sucesivamente, esto es el resultado de la simulación es i si i es el menor entero tal que $RND < p(0) + p(1) + \cdots + p(i)$.

Por lo tanto para hacer una simulación con un número finito de posibles resultados se procede como sigue: Se determina N y las probabilidades $p(0), \cdots, p(N)$. Se forman las sumas $sum(i) := p(0) + \cdots + p(i)$ para $i = 0, \cdots, N$. Se genera RND y se encuentra el menor valor de i para el cual $RND < sum(i)$.

Ejemplo 8.15 *Un experimento aleatorio tiene como posibles resultados a $0, \cdots, N$. Dado m el número de simulaciones que se desean realizar, el valor N y las $(N+1)$ probabilidades $p(0) + \cdots + p(N)$, se desea escribir un programa que arroje como resultado todas las m simulaciones del experimento.*

Algoritmo 8.16

Entrada:

vector P con componentes $p(i),\ i = 0, \cdots, N$.

número máx de simulaciones.

Salida: resultados: 0 *o* 1 *o* 2 *o* $\cdots$ *o* N.

Inicialización: *suma* es un vector de $\mathbb{R}^{N+1}$ con componentes $suma(i)$ definidas inicialmente como:

$suma(0) := p(0)$

para $j = 1(1)N$ haga:

$suma(j) := suma(j-1) + p(j)$

Iteración

para $j = 1(1)$ máx haga:

$i := 0;$

mientras $RND \geq suma(i)$ haga:

$i := i + 1$

escriba i. ▲

Ejemplo 8.17 *Generar un experimento que tenga como resultados a* $1, 2, 3$ *y* 4 *con probabilidades* $p(1) = 0.2,\ p(2) = 0.1,\ p(3) = 0.4$ *y* $p(4) = 0.3$.

Implementando el algoritmo anterior se obtuvo, con máx $= 50$, *la secuencia siguiente:*

```
1 2 3 4 2 3 3 4 3 2 3 2 4 3 2 1 4 4 3 1 4
3 4 1 1 3 2 2 2 1 3 1 4 3 4 1 3 3 2 3 2 2
2 1 3 3 4 4 3 4.
```

Ejemplo 8.18 *Simular* 100 *lanzamientos de un dado corriente.*

Solución *En este caso* $P(i) = \frac{1}{6}$, *para* $i = 1, \cdots, 6$ *y* máx $= 100$. *Aplicando el algoritmo se obtiene:*

```
4 4 3 1 2 4 1 4 3 2 1 6 2 3 1 1 4 4 4 4 2
4 4 4 2 3 2 4 2 3 2 2 3 2 2 2 4 4 3 3 2 3
4 3 2 4 2 3 1 1 2 1 3 1 1 5 4 2 3 1 4 2 2
3 1 4 5 3 2 5 3 4 2 2 1 3 1 2 1 3 5 4 3 5
3 4 1 1 3 3 3 3 3 3 2 3 2 3 6 2.
```

Ejemplo 8.19 (Simulación de una variable aleatoria binomial)
Se busca generar el valor de X, una variable aleatoria binomial de parámetros n y p. Esto es, una variable aleatoria X para la cual se satisfaga que:

$$P(X = i) = \binom{n}{i} p^i (1-p)^{n-i} \quad \text{con } i = 0, \cdots, n.$$

Para hacerlo se tiene en cuenta que:

$$P(X = i) = \frac{n-i+1}{i} \frac{p}{(1-p)} P(X = i-1) \quad \text{para } i = 1, 2, \cdots, n$$

El algoritmo propuesto a continuación, verifica primero si $X = 0$, luego si $X = 1$ y asi sucesivamente.

Algoritmo 8.20

Entrada:

parámetros n y p

número máx de simulaciones.

Salida: resultados: 0 *o* 1 *o* 2 *o* 3 *o* $\cdots$ *o* n.

Inicialización:

$q := 1 - p$

P es un vector de $\mathbb{R}^{N+1}$ con componentes $P(i)$ definidas inicialmente como:

$P(0) := q^n$

para $j = 1(1)n$ haga:

$P(j) := \frac{(n-j+1)}{j} * \frac{p}{q} * P(j-1)$

suma es un vector de $\mathbb{R}^{N+1}$ con componentes $suma(i)$ definidas inicialmente como:

$suma(0) := P(0)$
para $j = 1(1)n$ haga:
  $suma(j) := suma(j-1) + P(j)$

Iteración

para $j = 1(1)$ máx haga:
  $i := 0;$
  mientras $RND \geq suma(i)$ haga:
    $i := i + 1$
  escriba i. ▲

Ejemplo 8.21 *Generar valores de una variable aleatoria binomial de parámetros $n = 20$ y $p = 0.3$.*

Solución *Aplicando el algoritmo para* máx $= 100$ *se obtuvo:*

6	4	9	4	4	7	4	5	6	3	3	6	5	7	6	3	2	5	5	6
4	6	5	5	3	4	4	4	5	5	3	3	5	4	6	6	7	5	6	6
6	5	4	8	5	3	7	4	6	3	5	4	6	4	7	7	3	5	5	3
5	8	6	5	2	4	5	4	3	5	5	5	7	2	4	6	5	4	4	6
5	5	4	6	6	6	8	4	4	5	5	5	4	7	3	4	4	5	3	4

8.4. Simulación de variables aleatorias discretas con un número infinito contable de resultados.

Supóngase que X es una variable aleatoria con valores $0, 1, \cdots$ y que $p(i) = P(X = i)$. Para simular un valor de X se requiere dividir el intervalo $[0, 1]$ en un número contable de subintervalos, luego generar RND y por último determinar en cuál subintervalo está localizado RND.

La idea inicial sería generar sumas de la forma $sum(i) := p(0) + \cdots + p(i)$ para $i = 0, \cdots, \infty$. Puesto que esto no es viable, se necesita de un método alternativo. El método propuesto a continuación consiste en computar valores sucesivos de $sum(i)$, parando cuando por primera vez ocurra que $RND < sum(i)$.

Ejemplo 8.22 *(Simulación de una variable aleatoria Poisson). Generar el valor de X, una variable aleatoria Poisson de parámetro λ.*

Solución *Para realizar el algoritmo se tendrá en cuenta que las probabilidades Poisson satisfacen la relación siguiente:*

$$P(X = i) = \frac{\lambda}{i} P\left(X = i - 1\right);\ para\ i = 1, 2, \cdots \tag{8.2}$$

Algoritmo 8.23 (Poisson)

Entrada:

parámetro λ

número máx de simulaciones.

Salida: resultados: $n \in \mathbb{N}$.

para $j = 1(1)$ máx haga:

inicialización:

$i := 0$
$prob := e^{-\lambda}$
$suma := prob$

verificación:

mientras $RND \geq suma$ haga:
$i := i + 1$
$prob := \lambda * prob/i$
$suma := suma + prob$;

escriba i. ▲

Ejemplo 8.24 *Generar valores de una variable aleatoria X con distribución Poisson de parámetro $\lambda = 0.2$.*

Solución *Haciendo uso del algoritmo se obtuvo para* máx $= 100$ *la secuencia de valores siguiente:*

0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	1	0	0	0
0	0	0	0	0	1	0	0	1	0	1	0	0	1	0	0	1	0	0	1
0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1
1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0

Obsérvese que la media muestral $\overline{X}$, en este caso, es exactamente igual a la media poblacional λ. ▲

El algoritmo anterior verifica en forma sucesiva si el valor de la variable Poisson es 0, luego si es 1, luego si es 2 y así sucesivamente y es bastante eficiente cuando λ es pequeño. Desafortunadamente, cuando λ es grande, el computador, debido a errores de redondeo, puede considerar $exp\,(-\lambda)$ igual a *cero*, y en tal caso, la media muestral puede alejarse bastante de la media poblacional λ. Sin embargo, puede hacerse uso del hecho de que una variable aleatoria Poisson con media λ, tiene una gran probabilidad de tomar uno de los valores $[\lambda]$ o $[\lambda]+1$, para desarrollar algoritmos más eficientes que el anterior y que eviten los problemas generados por los errores de redondeo (ver [Ros3]).

8.5. Simulación de distribuciones continuas

En el capítulo 4.(ver 4.8) se observó que si X es una variable aleatoria con función de distribución continua y creciente F y si Y es una variable aleatoria con distribución uniforme en el intervalo $(0,1)$ entonces la variable aleatoria $Z := F^{-1}(Y)$ tiene la misma distribución de X. Este hecho será utilizado a continuación para generar variables aleatorias continuas con función de distribución creciente F.

Sea X una variable aleatoria con función de distribución continua y creciente F. Sea I el intervalo en el cual F es creciente. Entonces, la inversa F^{-1} de F es una función de $(0,1)$ en I. Para simular un valor de X se procede como sigue: primero se genera RND y entonces $F^{-1}(RND)$ es el valor simulado de X.

Ejemplo 8.25 *(Simulación de una variable aleatoria con distribución uniforme sobre el intervalo $[a,b]$). Sea $X \stackrel{d}{=} \mathcal{U}[a,b]$. En este caso la función de distribución F de X está dada por:*

$$F(x) = \begin{cases} 0 & si \quad x < a \\ \frac{x-a}{b-a} & si \quad a \leq x \leq b \\ 1 & si \quad x > b \end{cases}$$

y el intervalo en el cual F es creciente es el intervalo $[a,b]$.

La inversa $F^{-1}(x)$ de F es una función de $(0,1)$ en $[a,b]$ dada por:

$$F^{-1}(x) = (b-a)\,x + a$$

Por lo tanto para generar un valor de X primero se genera RND y luego se toma:

$$X = (b - a) * RND + a$$

Como caso particular, se obtuvo para $a = -1$ y $b = 5$, los siguientes valores de X :

4.094	3.447	1.490	-0.041	0.833
0.294	1.634	3.437	1.899	4.900
3.332	0.381	-0.859	3.523	2.468
-0.260	4.905	4.875	1.751	-0.558
-0.821	0.436	3.093	2.869	4.060
0.184	4.138	1.770	-0.686	2.323
0.255	3.837	2.065	3.543	-0.436
1.739	3.563	4.084	2.407	1.260
-0.735	-0.876	-0.050	0.473	3.580
0.521	4.330	0.255	4.635	3.198

Ejemplo 8.26 *Simular* 50 *valores de una variable aleatoria que tiene como función de densidad a:*

$$f(x) = \begin{cases} 0 & si \quad x < 0 \\ \frac{2}{(x+1)^3} & en \quad otro\ caso \end{cases}$$

Solución *En este caso la función de distribución de la variable aleatoria X está dada por:*

$$F(x) = \begin{cases} 0 & si \quad x < 0 \\ 1 - \frac{1}{(x+1)^2} & si \quad 0 \leq x \end{cases}$$

El intervalo en el cual F es creciente es $(0, \infty)$. Por lo tanto, $F^{-1}(x)$ es una función de $(0, 1)$ en $(0, \infty)$ dada por:

$$F^{-1}(x) = \frac{1}{\sqrt{1 - x}} - 1$$

Por lo tanto para generar un valor de X primero se genera RND y luego se toma:

$$X = \frac{1}{\sqrt{1 - RND}} - 1$$

Para máx $= 50$ *se obtuvieron los siguientes valores de* X :

0.362	0.167	0.774	0.314	1.444
1.300	0.082	0.315	1.130	0.607
0.409	0.468	0.080	0.035	0.221
1.121	0.747	0.095	0.049	1.652
0.607	1.225	0.130	0.472	0.552
0.367	0.151	0.533	0.041	0.136
1.156	3.236	2.814	0.240	2.011
1.304	0.702	0.067	0.832	0.953
0.015	0.574	0.416	0.287	10.762
0.182	0.029	0.086	0.371	0.966

Ejemplo 8.27 *(Generación de una variable aleatoria con distribución exponencial de parámetro λ) Generar* 50 *valores de una variable aleatoria con distribución exponencial de parámetro* $\lambda = 2$.

Solución *Si* $X \overset{d}{=} Exp(\lambda)$ *entonces su función de densidad está dada por:*

$$f(x) = \begin{cases} \lambda \exp(-\lambda x) & si \quad x > 0 \\ 0 & en \quad otro\ caso \end{cases}$$

y su función de distribución es:

$$F(x) = \begin{cases} 1 - \exp(-\lambda x) & si \quad x > 0 \\ 0 & en \quad otro\ caso \end{cases}$$

El intervalo en el cual F *es creciente es* $(0, \infty)$. *Por lo tanto,* $F^{-1}(x)$ *es una función de* $(0,1)$ *en* $(0, \infty)$ *dada por:*

$$F^{-1}(x) = \frac{-\ln(1-x)}{\lambda}$$

Por lo tanto para generar un valor de X *primero se genera* RND *y luego se toma:*

$$X := \frac{-\ln(1-RND)}{\lambda}$$

Al tomar $\lambda = 2$ *se obtuvo:*

0.116	0.462	0.448	0.194	1.166
0.357	0.258	0.169	0.333	0.772

0.160	1.074	0.215	0.085	1.105
0.568	0.239	0.871	0.854	0.692
0.445	0.445	0.223	0.007	0.022
0.088	0.283	0.217	2.233	0.074
0.213	0.012	0.720	0.047	0.037
0.428	0.078	1.221	0.581	0.037
1.950	0.159	0.550	2.120	0.044
0.179	0.100	1.167	1.351	0.339

8.5.1. El método del rechazo

Otro método de uso frecuente para generar valores de variables aleatorias con función de densidad f dada, es el llamado método del rechazo. En este método se supone que se puede generar una variable aleatoria Y con una función de densidad conocida g, luego se genera un número aleatorio RND y se acepta este valor generado con una probabilidad proporcional a $\frac{f(Y)}{g(Y)}$. La técnica es entonces la siguiente:

1. Se genera Y con densidad g.

2. Se genera un número aleatorio RND.

3. Si $RND \leq \frac{f(Y)}{cg(Y)}$, hacer $X = Y$ en caso contrario regresar a 1.

Se puede demostrar que la variable aleatoria generada por el método del rechazo tiene densidad f. (ver [Ros3]).

A continuación se va a ilustrar el método con un ejemplo:

Ejemplo 8.28 *(Simulación de una variable aleatoria con distribución beta) Generar una variable aleatoria con función de densidad dada por:*

$$f(x) = 20x^3(1-x)\mathcal{X}_{(0,1)}(x)$$

Como la función de densidad de esta variable se anula fuera del intervalo $(0,1)$, *se considera el método del rechazo con* $g(x) = \mathcal{X}_{(0,1)}(x)$.

Para determinar la constante c tal que $\frac{f(x)}{g(x)} \leq c$, *se busca el máximo de* $\frac{f(x)}{g(x)} = 20x^3(1-x)$.*Derivando se obtiene que el valor máximo se toma en* $x = \frac{3}{4}$. *Entonces,*

$$20x^3(1-x) \leq \frac{135}{64}$$

por lo tanto,

$$\frac{f(x)}{cg(x)} = \frac{256}{27}x^3(1-x)$$

El procedimiento a seguir sería entonces:

Paso 1. *Generar números aleatorios* $RND1$ *y* $RND2$.

Paso 2. *Si* $RND2 \leq \frac{256}{27}(RND1)^3(1-RND1)$ *detenerse y hacer* $X = RND1$, *en caso contrario regresar al paso 1.*

Al aplicar este procedimiento se obtuvieron los siguientes valores de X :

0.953	0.443	0.945	0.626	0.637
0.649	0.884	0.607	0.835	0.274
0.230	0.764	0.455	0.714	0.771
0.687	0.518	0.525	0.717	0.675
0.829	0.791	0.988	0.744	0.629
0.602	0.862	0.841	0.805	0.310
0.856	0.660	0.567	0.917	0.728
0.757	0.674	0.566	0.827	0.645
0.935	0.573	0.634	0.482	0.678
0.704	0.745	0.609	0.620	0.629

Ejemplo 8.29 *(Generación de una variable aleatoria con distribución normal estándar) Sea* $X \stackrel{d}{=} \mathcal{N}(0,1)$, *esto es,* X *es una variable aleatoria con función de densidad dada por:*

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right);\ x \in \mathbb{R}.$$

Para generar los valores de X *se observa en primer lugar que la variable aleatoria* $W := |X|$ *tiene como función de densidad*

$$h(x) = \frac{2}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right);\ x \in \mathbb{R}.$$

El procedimiento a seguir es el siguiente: se genera primero a la variable aleatoria W, *se obtienen luego los valores de la variable aleatoria* X *al hacer que sea igualmente probable que* W *sea igual a* X *y a* $-X$. *Para generar* W *se usa el método del rechazo con* $g(x) = \exp(-x)$ *para* $x > 0$. *Se observa que* *la función* $\frac{h(x)}{g(x)}$ *toma su valor máximo en* $x = 1$. *Luego* $c = \frac{2}{\sqrt{2\pi}} \exp\left(\frac{1}{2}\right) = \sqrt{\frac{2e}{\pi}}$. *Por lo tanto,*

$$\frac{h(x)}{cg(x)} = \exp\left(-\frac{1}{2}(x-1)^2\right).$$

El algoritmo siguiente permite simular una variable aleatoria con distribución normal estándar.

Algoritmo 8.30 *(Simulación de una variable aleatoria con distribución normal estándar)*

Entrada: número máx de simulaciones.

Salida: valores de Z.

para $i = 1(1)$ máx haga:

inicialización:

genere una variable aleatoria Y con distribución exponencial de parámetro $\lambda = 1$

genere un número aleatorio RND

verificación:

mientras $RND > \exp\left(-\frac{1}{2}(Y-1)^2\right)$ haga:

genere una variable aleatoria Y con distribución exponencial de parámetro $\lambda = 1$

genere un número aleatorio RND

salida: (Para escribir Y o $-Y$ con igual probabilidad)

genere un número aleatorio rnd

si $rnd > 0.5$ entonces

escriba $-Y$

en caso contrario

escriba Y. ▲

Al aplicar el algoritmo anterior para máx $= 60$ se obtuvieron los resultados siguientes:

-0.733	0.748	-0.946	-2.304	0.155
1.743	-0.435	1.143	1.468	-1.818
0.746	0.117	-2.353	-0.042	-0.741
-1.425	-1.684	0.909	0.126	0.889
0.566	0.932	-1.062	-1.233	0.641
0.070	-1.379	-1.026	0.571	-1.014
0.337	-1.431	-0.106	-0.650	1.304
-0.491	0.065	-1.642	-0.462	-0.929
1.334	-0.122	0.893	-1.161	0.295
1.820	1.876	-0.332	1.064	-0.006
-0.136	0.543	-1.585	-1.371	-0.242
0.542	0.462	1.988	-0.039	0.643

En este caso la media muestral $\overline{Z}$ es igual a -0.0818.

8.6. Ejercicios

1. En una ciudad pequeña hay cinco hoteles. Supóngase que llegan tres turistas quienes se alojan aleatoriamente en ellos. Hacer un programa para calcular la probabilidad de que todos los turistas se hospeden en hoteles diferentes. ¿A qué es igual la probabilidad exacta?.
2. En un clóset hay 10 pares de zapatos. Se sacan 8 zapatos al azar. Hacer un programa para calcular la probabilidad de extraer exactamente un par correcto.
3. Escribir un programa para la realización práctica de los algoritmos 8.1 y 8.3.
4. Escribir programas para la realización práctica de los algoritmos 8.9 y 8.11.
5. Escribir un programa para la realización práctica de 8.13.
6. En el ejemplo 8.10, considere la variable aleatoria X :="número de parejas que coinciden". Hacer un programa para calcular EX.
7. Escribir un programa para la realización práctica de 8.16.
8. Escribir un programa para la realización práctica de 8.20.

9. Escribir un programa para la realización práctica de 8.23.

10. Escribir un algoritmo e implementarlo para simular una variable aleatoria con distribución uniforme en el intervalo $[a, b]$.

11. Escribir un algoritmo e implementarlo para simular una variable aleatoria con distribución exponencial con parámetro λ..

12. Escribir un algoritmo e implementarlo para simular una variable aleatoria con función de densidad dada por:

$$f(x) = \begin{cases} 3x^2 & \text{si} \quad 0 < x < 1 \\ 0 & \text{en} \quad \text{otro caso} \end{cases} .$$

13. Escribir un algoritmo e implementarlo para simular una variable aleatoria con distribución beta con parámetros a y b.

14. Escribir un programa para la realización práctica de 8.29.

15. Hacer uso del método del rechazo para hacer un programa que genere los valores de una variable aleatoria X con función de densidad dada por:

$$f(x) = \begin{cases} \frac{2}{\sqrt{\pi}} x^{\frac{1}{2}} \exp(-x) & \text{si} \quad x > 0 \\ 0 & \text{en} \quad \text{otro caso} \end{cases}$$

Usar como función de prueba $g(x) = \frac{2}{3} \exp\left(-\frac{2}{3}x\right)$.

Apéndice A

Nociones de conjuntos

En este apéndice se presentan algunos conceptos y resultados básicos de la teoría de conjuntos, usados a lo largo del texto.

Intuitivamente un conjunto es una agrupación de objetos bien definidos. Los objetos que integran un conjunto son llamados elementos o miembros del conjunto.

Se acostumbra a denotar los conjuntos por las letras mayúsculas del alfabeto: $A, B, C, M, X, \ldots$. Los elementos del conjunto se denotan, usualmente, por letras minúsculas: $a, b, c, x, y, \ldots$

Se escribe $x \in A$, para indicar que el elemento x es un miembro del conjunto A. Si por el contrario, x no es un elemento del conjunto A se escribe $x \notin A$.

Los conjuntos pueden ser descritos enumerando todos sus elementos o enunciando propiedades que deben tener sus elementos. En el primer caso se dice que el conjunto está determinado por extensión y en el segundo por comprensión. Asi por ejemplo, el conjunto

$$A = \{1, 3, 5, 9\}$$

está descrito por extensión, en tanto que el conjunto

$$B = \{x : \ x \text{ es un número racional menor o igual a } 5\}$$

está definido por comprensión.

Un conjunto cuyo número de elementos se puede expresar como un entero no negativo se llama conjunto finito. El conjunto que no tiene elementos se llama conjunto vacio y se denota por Φ. Un conjunto se dice infinito si no es finito.

Se dice que dos conjuntos son iguales, si y sólo si, ellos tienen exactamente los mismos elementos. Si todos los elementos de un conjunto A son elementos de un conjunto B se dice que A está contenido en B (o B contiene a A) y se escribe $A \subseteq B$ $(B \supseteq A)$.

Es claro a partir de la definición que todo conjunto es subconjunto de si mismo y que Φ es subconjunto de todo conjunto.

Si A es subconjunto de B y existe por lo menos un elemento de B que no está en A, se dice que A es subconjunto propio de B y se escribe $A \subset B$.

En las aplicaciones de la teoría de conjuntos usualmente todos los conjuntos que se considerán son subconjuntos de un mismo conjunto. A este conjunto se le llama conjunto universal y se denota por $\mathcal{U}$.

La unión de dos conjuntos A y B, denotada por $A \cup B$, es el conjunto de los elementos que pertenecen a A, o a B o a ambos. Esto es:

$$A \cup B := \{x : x \in A \ \vee \ x \in B\}.$$

Similarmente si $\{A_i\}_{i \in I}$ es una familia de conjuntos y $J \subseteq I$, entonces:

$$\bigcup_{i \in J} A_i := \{x : x \in A_i \text{ para por lo menos un } i \in J\}.$$

La intersección de dos conjuntos A y B, denotada por $A \cap B$, es el conjunto de los elementos que pertenecen a A y a B. Esto es:

$$A \cap B := \{x : x \in A \ \wedge \ x \in B\}.$$

Similarmente si $\{A_i\}_{i \in I}$ es una familia de conjuntos y $J \subseteq I$, entonces:

$$\bigcap_{i \in J} A_i := \{x : x \in A_i \text{ para todo } i \in J\}.$$

Si A y B no tienen elementos en común, es decir si $A \cap B = \Phi$, se dice que ellos son mutuamente excluyentes o disjuntos.

La diferencia de dos conjuntos A y B,denotada por $A - B$, es el conjunto de todos los elementos que pertenecen a A pero que no pertenecen a B, esto es:

$$A - B := \{x : x \in A \ \wedge \ x \notin B\}.$$

El complemento de un conjunto A,denotado por A^c, es la diferencia entre el conjunto universal $\mathcal{U}$ y A. Esto es:

$$A^c := \{x : x \in \mathcal{U} \ \wedge \ x \notin A\}.$$

A continuación se presentan, sin demostración, las propiedades básicas de las operaciones entre conjuntos. El lector interesado puede consultar las demostraciones en [MuJ].

Teorema A.1 *Sean A, B y C subconjuntos de un conjunto universal $\mathcal{U}$. Entonces:*

1. (Leyes conmutativas)

   a) $A \cup B = B \cup A$

   b) $A \cap B = B \cap A$

2. (Leyes asociativas)

   a) $A \cup (B \cup C) = (A \cup B) \cup C$

   b) $A \cap (B \cap C) = (A \cap B) \cap C$

3. (Leyes distributivas)

   a) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

   b) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

4. (Leyes complementarias)

   a) $A \cup A^c = \mathcal{U}$

   b) $A \cap A^c = \Phi$

   c) $A \cup \mathcal{U} = \mathcal{U}$

   d) $A \cap \mathcal{U} = A$

   e) $A \cup \Phi = A$

   f) $A \cap \Phi = \Phi$

5. (Leyes de la diferencia)

   a) $A - B = A \cap B^c$

   b) $A - B = A - (A \cap B) = (A \cup B) - B$

   c) $A - (B - C) = (A - B) \cup (A - C)$

   d) $(A \cup B) - C = (A - C) \cup (B - C)$

   e) $A - (B \cup C) = (A - B) \cap (A - C)$

   f) $(A \cap B) \cup (A - B) = A$

   g) $(A \cap B) \cap (A - B) = \Phi$

6. (Leyes de De-Morgan)

   a) $(A \cap B)^c = (A^c \cup B^c)$

   b) $(A \cup B)^c = (A^c \cap B^c)$

7. (Ley involutiva)

   $(A^c)^c = A$

8. (Leyes de idempotencia)

   a) $A \cup A = A$

   b) $A \cap A = A$

A continuación se presenta uno de los conceptos más importantes de la matemática: Si a cada elemento de un conjunto A se le hace corresponder, de algún modo bien definido, un elemento único de un conjunto B, entonces, se dice que esa correspondencia es una función de A en B. Denotando esa correspondencia por f se escribe:

$$f : A \to B.$$

El conjunto A se llama el dominio de la función y el conjunto B el codominio de f. Si $a \in A$ entonces el elemento de B que le corresponde a a, se llama la imagen de a y se denota por $f(a)$.

Si f es una función de A en B y $b \in B$ entonces se define la imagen recíproca de B como el conjunto de todos los elementos del conjunto A que tienen a b como imagen, esto es:

$$f^{-1}(b) := \{a \in A : f(a) = b\}.$$

Más en general, se tiene que si C es un subconjunto de B, entonces el conjunto de elementos de A cuyas imágenes son elementos de C, se llama imagen recíproca de C por f, y se denota por $f^{-1}(C)$. Esto es:

$$f^{-1}(C) := \{a \in A : f(a) \in C\}.$$

Análogamente, si D es un subconjunto de A, entonces al subconjunto de B cuyos elementos son las imágenes de los elementos de D por f se le llama imagen directa de D por f y se denota por $f(D)$. Esto es:

$$\begin{aligned} f(D) &:= \{b \in B : b = f(a) \text{ para algún } a \in D\} \\ &= \{f(a) : a \in D\}. \end{aligned}$$

Algunas de las propiedades más importantes de las imagenes inversa y directa de conjuntos por una función, están resumidas en el siguiente teorema, cuya demostración puede ser consultada en [MuJ].

Teorema A.2 *Si $f : A \to B$ es una función y N_1, N_2 y N son subconjuntos de B y M es un subconjunto de A entonces:*

1. $f^{-1}(N_1 \cup N_2) = f^{-1}(N_1) \cup f^{-1}(N_2)$.
2. $N_1 \subseteq N_2 \Longrightarrow f^{-1}(N_1) \subseteq f^{-1}(N_2)$.
3. $f^{-1}(N_1 \cap N_2) = f^{-1}(N_1) \cap f^{-1}(N_2)$.
4. $f(f^{-1}(N)) \subseteq N$.

Se dice que un conjunto A es numerable, si y sólo si, existe una función f de A en el conjunto de los números naturales $\mathbb{N}$ tal que:

1. f es 1-1, es decir, si $x, y \in A$ y $x \neq y$, entonces, $f(x) \neq f(y)$.
2. f es sobreyectiva, es decir, $f(A) = \mathbb{N}$.

Apéndice B

Introducción al análisis combinatorio

En el capítulo I se vio que en los espacios laplacianos el cálculo de probabilidades se reduce a contar el número de elementos de un conjunto finito. La teoría matemática del conteo es formalmente conocida como análisis combinatorio.

Las diferentes técnicas de conteo se basan en el siguiente principio fundamental.

Teorema B.1 *(Principio fundamental del conteo) Supóngase que se realizan dos experimentos . Si el primer experimento tiene m posibles resultados y si para cada uno de los resultados del primer experimento hay n posibles resultados del segundo experimento, entonces, el número de posibles resultados de los dos experimentos realizados en el orden indicado es mn.*

Demostración. El principio fundamental del conteo puede demostrarse enumerando todos los posibles resultados de los dos experimentos como sigue:

$$\begin{array}{cccc} (1,1) & (1,2) & \cdots & (1,n) \\ (2,1) & (2,2) & \cdots & (2,n) \\ \vdots & \vdots & \vdots & \vdots \\ (m,1) & (m,2) & \cdots & (m,n) \end{array}$$

Este es un arreglo rectángular con m filas y n columnas, por lo tanto tiene mn componentes. ■

Ejemplo B.2 *En el departamento de Estadística de una universidad hay* 10 *profesores consejeros, cada uno de los cuales tiene* 15 *alumnos a su cargo.*

Si un profesor y uno de sus alumnos van a ser escogidos para representar al departamento en un evento académico, ¿de cuántas maneras puede hacerse la selección?

En este caso hay 10 *formas posibles de escoger al profesor, una vez escogido éste, hay* 15 *formas de seleccionar al estudiante. Por el principio fundamental del conteo hay entonces* 150 *formas de seleccionar la pareja que va a representar al departamento.* ▲

El principio fundamental del conteo puede ser generalizado como sigue:

Teorema B.3 ***(Generalización del principio fundamental del conteo)*** *Si r experimentos son realizados de tal forma que el primero tiene n_1 posibles resultados, y si para cada uno de esos n_1 posibles resultados hay n_2 posibles resultados del segundo experimento, y si para cada uno de los posibles resultados de los dos primeros experimentos hay n_3 posibles resultados del tercer experimento y asi sucesivamente, entonces el número total de resultados de los r experimentos realizados en la forma indicada es $n_1.n_2.\cdots.n_r$.*

Demostración. Como ejercicio. ■

Ejemplo B.4 *El número total de placas para automóviles que pueden hacerse si cada placa debe tener tres letras diferentes y cinco números es igual a* $26 \times 25 \times 24 \times 10^5 = 1.56 \times 10^9$ ▲

Ejemplo B.5 *Si se lanza un dado corriente cuatro veces consecutivas entonces el número total de posibles resultados de este experimento es* $6^4 = 1296$ ▲

Ejercicio B.6 *¿Cuántos números de tres dígitos distintos, menores que* 500*, pueden formarse con los dígitos* $1, 2, 3, 4, 5, 6$ *y* 7 *?*

Ejercicio B.7 *¿Cuántos números de tres dígitos pueden formarse con los dígitos* $1, 4, 8$ *y* 5 *si:*

1. *los tres dígitos son distintos?*
2. *los números deben ser impares?*
3. *los números deben ser divisibles por* 5 *?*

Ejemplo B.8 *Supóngase que hay n bolas distinguibles y r urnas distintas, entonces, el número de maneras en que se pueden distribuir las bolas en las urnas es igual a r^n, ya que la primera bola puede ser colocada en cualquiera de las r urnas, la segunda en cualquiera de las r urnas y así sucesivamente, por lo tanto hay*

$$\underbrace{r.r.\cdots.r}_{n-veces} = r^n$$

formas de colocar las bolas en las urnas. ▲

Ejemplo B.9 ***(Permutaciones)*** *Una permutación es un arreglo en un orden particular de los objetos que forman un conjunto. Por ejemplo, las permutaciones de las letras a, b y c son $abc, acb, bac, bca, cab, cba$. Es decir, hay 6 permutaciones de las tres letras.*

Se puede determinar el número total de permutaciones de los objetos que forman un conjunto sin necesidad de escribir explicítamente la lista de permutaciones, siguiendo el siguiente razonamiento: supóngase que se tiene un conjunto con n elementos, entonces para la primera posición se puede escoger a cualquiera de los n elementos, para la segunda, a cualquiera de los $(n-1)$ elementos restantes, para la tercera a cualquiera de los $(n-2)$ elementos restantes y así sucesivamente. Por lo tanto, el número total $P(n,n)$ de permutaciones de los n elementos es:

$$P(n,n) = n(n-1)(n-2)\cdots 1$$

El producto de un entero positivo n por todos los que le preceden se denota por $n!$ y se llama n factorial. Se define $0! := 1$.

Por lo tanto se tiene que:

$$P(n,n) = n!$$

Una permutación de n objetos tomados $r \leq n$ a la vez, es un arreglo en un orden particular de r de los n objeros. Así por ejemplo, las posibles permutaciones de las letras $a, b.c$ y d tomadas dos a la vez son: $ab, ba, ac, ca, ad, da, bc, cb, bd, db, cd, dc$. Es decir, hay 12 permutaciones de las letras a, b, c y d tomadas dos a la vez. El número total $P(n,r)$ de permutaciones de n objetos tomados r a la vez, puede hallarse siguiendo el siguiente razonamiento: para la primera posición se puede elegir a cualquiera de los n objetos, para la segunda a cualquiera de los $(n-1)$ restantes, y así sucesivamente hasta llegar que para la $r-$ésima posición se puede elegir a cualquiera de los $(n-r+1)$

objetos restantes. Esto es:

$$\begin{aligned} P(n,r) &= n(n-1)\cdots(n-r+1) \\ &= \frac{n!}{(n-r)!}; \quad r \leq n \qquad \blacktriangle \end{aligned}$$

Ejemplo B.10 *Supóngase que se desea sentar a cuatro niñas y tres niños en una fila. Si los niños y las niñas pueden sentarse en cualquier orden entonces habría* $7! = 5040$ *formas de acomodarlos. Si se desea que los niños y las niñas queden alternados entonces habría*

$$4 \times 3 \times 3 \times 2 \times 2 \times 1 \times 1 = 144$$

formas de acomodarlos. Si se desea que tanto los niños como las niñas queden juntos entonces habría $2 \times 4! \times 3! = 288$ *formas de acomodarlos.* ▲

Ejercicio B.11 *¿De cuántas maneras se pueden sentar en una fila cuatro niños y cuatro niñas si los niños y las niñas deben quedar alternados?, ¿de cuántas si los niños se sientan juntos y las niñas también?, ¿de cuántas si sólo las niñas se sientan juntas?*

Ejercicio B.12 *Un inspector revisa seis máquinas diferentes durante el día. A fin de impedir que los operadores sepan cuándo hará la inspección, varía el orden de las visitas, ¿de cuántas maneras puede hacerlo?*

Ejemplo B.13 *Un estudiante desea acomodar* 4 *libros de cálculo,* 2 *de física,* 5 *de probabilidad y* 3 *de álgebra en un estante de tal manera que los libros de la misma materia queden juntos. Se tiene que hay* $4! \times 2! \times 5! \times 3!$ *formas de acomodar los libros de tal manera que los primeros sean los de cálculo, los siguientes los de física, luego los de probabilidad y por último los de álgebra. Como hay* $4!$ *formas de ordenar las materias, entonces hay en total*

$$4! \times 4! \times 2! \times 5! \times 3! = 8.2944 \times 10^5$$

formas de acomodar los libros en los estantes. ▲

Ejemplo B.14 *Se desea calcular el número de maneras de acomodar a* 3 *mexicanos,* 4 *venezolanos,* 3 *argentinos y* 5 *colombianos alrededor de una mesa redonda si las personas de la misma nacionalidad insisten en sentarse juntas. En este caso se tienen cuatro grupos de personas: los mexicanos, los venezolanos, los argentinos y los colombianos. El número de maneras de acomodar los cuatro grupos en la mesa redonda es* $3!$.*Los mexicanos se* *pueden acomodar de* 3! *formas, los venezolanos de* 4! *formas, los argentinos de* 3! *formas y los colombianos de* 5! *formas. Hay, por lo tanto*

$$3! \times 3! \times 4! \times 3! \times 5! = 6.\,220\,8 \times 10^5$$

formas de acomodación de las personas en la mesa. ▲

Ejercicio B.15 *Un grupo de* 5 *alemanes,* 6 *ingleses,* 4 *japoneses y* 6 *colombianos deben ser ubicados en una mesa redonda. ¿De cuántas maneras puede hacerse la ubicación?, ¿de cuántas si las personas de la misma nacionalidad deben quedar juntas?,¿de cuántas si los colombianos deben quedar juntos?.*

Ejemplo B.16 *Sea N el número de permutaciones diferentes de las nueve letras de la palabra "elefantes". Si todas las letras fueran distintas se tendría que el número total de permutaciones es* 9!, *como las tres "e" pueden permutarse entre ellas de* 3! *formas, entonces* $3!N = 9!$. *Esto es*

$$N = \frac{9!}{3!}$$

En general se tiene que: el número total N de formas en que pueden permutarse n objetos de los cuales n_1, $n_2, \cdots, n_r$ son iguales entre si, es:

$$N = \frac{n!}{n_1!n_2! \cdots n_r!} \quad \blacktriangle$$

Ejercicio B.17 *¿Cuántos arreglos diferentes pueden formarse con las letras de la palabra "ganancia"?*

Ejemplo B.18 ***(Combinaciones)*** *Supóngase que se tienen n objetos diferentes. Cada elección de $r \leq n$ de dichos objetos se llama una combinación de orden r. En otras palabras, una combinación de orden r de un conjunto con n elementos es un subconjunto con r elementos del conjunto. Asi por ejemplo, las combinaciones de orden* 2 *de las letras a, b, c y d son: $\{a, b\}, \{a, c\}, \{a, d\}, \{b, c\}, \{b, d\}$ y $\{c, d\}$, esto es, hay* 6 *combinaciones de orden* 2 *de las letras a, b, c y d. Para determinar el número de combinaciones $C(n, r)$ de orden r de n objetos, se observa que si se tuviese en cuenta el orden habría $P(n, r)$ formas de escoger los r objetos, como los r objetos pueden permutarse entre si de $r!$ formas, entonces:*

$$r!C(n, r) = P(n, r)$$
$$C(n, r) = \frac{n!}{(n-r)!r!}.$$

El número $C(n,r)$ se llama " n combinado r " y se denota por $\binom{n}{r}$.

Por convención se define

$$\binom{n}{r} := 0 \quad si\, r < 0 \; o \; r > n \qquad \blacktriangle$$

Ejemplo B.19 *De un grupo de 10 mujeres y 12 hombres se deben escoger cinco parejas, conformadas por hombre y mujer, para un baile. Se desea determinar el número de selecciones posibles.*

Se tiene que hay $\binom{12}{5}$ formas de escoger los hombres que harán parte de la pareja. Una vez escogidos los hombres, se escogen las mujeres, lo cual se puede hacer de $\binom{10}{5}$ formas distintas. Hay por lo tanto

$$\binom{12}{5}\binom{10}{5} = 199584$$

formas de elegir las cinco parejas. $\blacktriangle$

Ejemplo B.20 *De un grupo de 5 hombres y 3 mujeres, se desea escoger un comité conformado por 3 personas, ¿cuántas selecciones son posibles?, ¿cuántas si en el comité debe haber por lo menos una mujer?, ¿cuántas si hay dos hombres que no se llevan bien y no pueden pertenecer ambos al grupo?, ¿cuántas si hay una pareja, hombre-mujer, que sólo aceptan hacer parte del comité si ambos pertenecen a éste?.*

Para el primer caso, se tiene que hay $\binom{8}{3} = 56$ formas de seleccionar el comité. Si en el comité debe haber por lo menos una mujer entonces hay:

$$\binom{3}{1}\binom{5}{2} + \binom{3}{2}\binom{5}{1} + \binom{3}{3}\binom{5}{0} = 46$$

formas de seleccionar el comité. Si hay dos hombres que no se la llevan bien, entonces hay dos opciones: o se incluye a uno de ellos en el comité o se excluye a ambos. En tal caso se tiene:

$$\binom{2}{1}\binom{6}{2} + \binom{6}{3} = 50$$

formas de seleccionar el comité.

Para el último caso se tienen dos opciones o se incluye a las dos personas en el comité o se excluye a ambas. En este caso se tiene:

$$\binom{6}{1} + \binom{6}{3} = 26$$

formas de seleccionar el comité. $\blacktriangle$

Ejercicio B.21 *En un examen de probabilidad un estudiante debe contestar diez de trece preguntas. ¿Cuántas formas de contestar el examen tiene el estudiante?, ¿cuántas si debe contestar por lo menos tres de las primeras cinco preguntas?, ¿cuántas si debe contestar exactamente una de las primeras cinco preguntas?.*

Ejercicio B.22 *Se van a comparar los efectos de dos medicamentos A y B en un estudio farmacéutico en el que participan 50 personas. A 20 personas se les administrará el medicamento A, a 20 el medicamento B y a las restantes 10 se les dará un placebo. ¿De cuántas maneras distintas pueden distribuirse los medicamentos y el placebo?*

Ejemplo B.23 *Un conjunto con n elementos va a ser particionado en m grupos distintos de tamaños $r_1, r_2, \cdots, r_m$ respectivamente siendo*

$$r_1 + r_2 + \cdots + r_m = n.$$

Se desea calcular el número de reparticiones posibles. Se observa que hay $\binom{n}{r_1}$ formas posibles de seleccionar el primer grupo, para cada selección del primer grupo, hay $\binom{n-r_1}{r_2}$ formas de seleccionar el segundo grupo y asi sucesivamente. Por lo tanto hay:

$$\binom{n}{r_1}\binom{n-r_1}{r_2}\cdots\binom{n-r_1-r_2-\cdots r_{m-1}}{r_m} = \frac{n!}{r_{1!} \times r_{2!} \times \cdots r_m!}$$

formas de seleccionar los grupos. Éste número se llama coeficiente multinomial y se denota por:

$$\binom{n}{r_1, r_2, \cdots, r_m} \quad \blacktriangle$$

Ejemplo B.24 *Los treinta alumnos de tercer grado de una escuela deben ser repartidos en 5 grupos con 12, 5, 3, 6 y 4 alumnos respectivamente. ¿Cuántas reparticiones son posibles?*

De acuerdo con el ejemplo anterior se tiene que hay

$$\binom{30}{12, 5, 3, 6, 4} = \frac{30!}{12! \times 5! \times 3! \times 6! \times 4!} = 4.4509 \times 10^{16}$$

reparticiones posibles. ▲

Ejercicio B.25 *El departamento de Estadística de una universidad tiene 33 profesores, los cuales deben ser divididos en cuatro grupos de 15, 8, 7 y 3 profesores. ¿Cuántas particiones son posibles?, ¿cuántas si los cinco miembros del comité asesor del departamento deben pertenecer al primer grupo?.*

Ejercicio B.26 *¿De cuántas maneras se pueden repartir* 7 *regalos entre* 3 *niños si uno de ellos debe recibir* 3 *regalos y los otros niños* 2 *cada uno?*

Ejemplo B.27 *Supóngase que se tiene n bolas indistinguibles entre ellas. ¿De cuántas formas se pueden distribuir las n bolas en r urnas?*

El resultado de este experimento puede ser descrito por un vector $(x_1, \cdots, x_r)$ donde x_i denota el número de bolas colocadas en la urna i. Por lo tanto el problema se reduce a encontrar el número de vectores $(x_1, \cdots, x_r)$, con componentes en los enteros no negativos, tales que $x_1 + \cdots + x_r = n$.

Para resolver este problema, supóngase que se colocan los objetos en una línea horizontal y que se dividen en r grupos de la siguiente forma: hay $(n-1)$ espacios entre los objetos, se escogen $(r-1)$ de esos espacios y allí se trazan líneas divisorias, al hacerlo quedan r grupos ninguno de ellos vacio. Esto es, hay $\binom{n-1}{r-1}$ vectores $(x_1, \cdots, x_r)$, con componentes positivas que satisfacen $x_1 + \cdots + x_r = n$.

Para hallar el número de soluciones no negativas, se observa que el número de soluciones no negativas de $x_1 + \cdots + x_r = n$, es igual al número de soluciones positivas de $y_1 + \cdots + y_r = n + r$, con $y_i = x_i + 1$ para $i = 1, 2, \cdots, r$.

Por lo tanto, el número total de formas en que se pueden distribuir las n bolas indistinguibles en r urnas es:

$$\binom{n+r-1}{r-1} \quad \blacktriangle$$

Ejemplo B.28 *Se reparten* 12 *regalos entre* 7 *niños. ¿Cuántas reparticiones son posibles?, ¿Cuántas si cada niño debe recibir por lo menos un regalo?.*

De acuerdo con el ejemplo anterior, hay $\binom{12+7-1}{7-1} = \binom{18}{6} = 18564$ formas de repartir los regalos entre los niños.

Si cada niño debe recibir por lo menos un regalo, entonces el número total de reparticiones es $\binom{12-1}{7-1} = \binom{11}{6} = 462$ $\blacktriangle$

Ejercicio B.29 *La facultad de ciencias de una universidad ha recibido* 30 *computadores idénticos que deben ser repartidos entre los* 7 *departamentos adscritos a la facultad.¿ Cuántas reparticiones son posibles?, ¿cuántas si cada departamento debe recibir por lo menos dos computadores?*

Ejercicio B.30 *Un inversionista tiene* 300 *millones de pesos para invertir en* 6 *posibles bonos. Cada inversión debe hacerla en millones de pesos. Si el total de los* 300 *millones deben ser invertidos, ¿cuántas estrategias de inversión son posibles?.*

Apéndice C

Tópicos de Álgebra Lineal

En este apéndice se presentan los resultados y conceptos de álgebra lineal, que son utilizados a lo largo del texto. Puesto que se supone que el lector conoce los conceptos básicos de la teoría de matrices, sólo se hará enfásis en algunos términos especiales.

El lector interesado en profundizar en estos temas puede consultar el texto de [Sea].

Los vectores de $\mathbb{R}^n$ serán considerados como vectores fila, esto es, matrices de tamaño $1 \times n$.

Si $\mathbf{x} = (x_1, \cdots, x_n)$ y $\mathbf{y} = (y_1, \cdots, y_n)$ son vectores de $\mathbb{R}^n$, entonces, se define su producto interno usual como sigue:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{i=1}^{n} x_i y_i$$

La norma euclidiana de un vector $\mathbf{x} = (x_1, \cdots, x_n)$ de $\mathbb{R}^n$ es el número real $\|\mathbf{x}\| := \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$.

El determinante de una matriz A se denota por $\det(A)$.

La traza de una matriz cuadrada A se denota por $tr(A)$.

La transpuesta de una matriz A se denota por A^T.

Una matriz cuadrada se dice diagonal si todos sus elementos no diagonales son iguales a 0.

Una matriz diagonal cuyos elementos diagonales son todos 1, se llama matriz identidad de orden 1 y se denota por I_n (o por I en caso de que el tamaño sea claro del contexto).

Una matriz cuadrada A se dice simétrica si es igual a su transpuesta.

Una matriz cuadrada de orden n se dice ortogonal si $AA^T = A^T A = I_n$.

Los valores propios de una matriz A de orden n son las soluciones de $\det(A - \lambda I_n) = 0$ donde λ es un número (real o complejo).

Un vector propio de una matriz cuadrada de orden n, asociado al valor propio λ, es un vector $\mathbf{0} \neq \mathbf{v} \in \mathbb{R}^n$ tal que $\mathbf{v}A = \lambda\mathbf{v}$.

Una matriz cuadrada A es singular, si y sólo si, al menos uno de sus valores propios es 0.

Una matriz que no es singular se llama matriz no singular.

Una matriz cuadrada A de orden n se llama positiva definida si para todo vector $\mathbf{x}$ de $\mathbb{R}^n$, con $\mathbf{x} \neq \mathbf{0}$ se satisface que $\mathbf{xAx}^T > 0$.

Una matriz cuadrada A de orden n se llama positiva semidefinida si para todo vector $\mathbf{x}$ de $\mathbb{R}^n$, con $\mathbf{x} \neq \mathbf{0}$ se satisface que $\mathbf{xAx}^T \geq 0$.

En el siguiente teorema se presentan algunas de las propiedades más importantes de las matrices simétricas y positivas definidas:

Teorema C.1 *Sea A una matriz simétrica y positiva definida entonces:*

1. *Existe una matriz W tal que $A = WW^T$.*
2. *Los valores propios de A son positivos (esto implica que la matriz A es no singular).*
3. *La matriz inversa de A es simétrica y positiva definida.*
4. *Existen matrices Λ y V cuadradas tales que $A = V\Lambda V^T$, donde Λ es una matriz diagonal cuyas componentes son los valores propios de A y V es una matriz ortogonal cuyas columnas son los vectores $\frac{\mathbf{v}}{\|\mathbf{v}\|}$, con $\mathbf{v}$ vector propio (este resultado sigue siendo válido cuando A es sólo simétrica).*

Apéndice D

Tablas Estadísticas

D.1. Probabilidades Binomiales

$$P(X \leq t) = \sum_{x=1}^{[t]} \binom{n}{x} p^x \, (1-p)^{n-x}$$

		p						
n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
2	0	0.9025	0.8100	0.7225	0.6400	0.4900	0.3600	0.2500
	1	0.9975	0.9900	0.9775	0.9600	0.9100	0.8400	0.7500
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.8574	0.7290	0.6141	0.5120	0.3430	0.2160	0.1250
	1	0.9927	0.9720	0.9392	0.8960	0.7840	0.6480	0.5000
	2	0.9999	0.9990	0.9966	0.9920	0.9730	0.9360	0.8750
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.8145	0.6561	0.5220	0.4096	0.2401	0.1296	0.0625
	1	0.9860	0.9477	0.8905	0.8192	0.6517	0.4752	0.3125
	2	0.9995	0.9963	0.9880	0.9728	0.9163	0.8208	0.6875
	3	1.0000	0.9999	0.9995	0.9984	0.9919	0.9744	0.9375
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.7738	0.5905	0.4437	0.3277	0.1681	0.0778	0.0313
	1	0.9774	0.9185	0.8352	0.7373	0.5282	0.3370	0.1875
	2	0.9988	0.9914	0.9734	0.9421	0.8369	0.6826	0.5000
	3	1.0000	0.9995	0.9978	0.9933	0.9692	0.9130	0.8125

n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
	4	1.0000	1.0000	0.9999	0.9997	0.9976	0.9898	0.9687
	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.7351	0.5314	0.3771	0.2621	0.1176	0.0467	0.0156
	1	0.9672	0.8857	0.7765	0.6554	0.4202	0.2333	0.1094
	2	0.9978	0.9842	0.9527	0.9011	0.7443	0.5443	0.3438
	3	0.9999	0.9987	0.9941	0.9830	0.9295	0.8208	0.6563
	4	1.0000	0.9999	0.9996	0.9984	0.9891	0.9590	0.8906
	5	1.0000	1.0000	1.0000	0.9999	0.9993	0.9959	0.9844
	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
7	0	0.6983	0.4783	0.3206	0.2097	0.0824	0.0280	0.0078
	1	0.9556	0.8503	0.7166	0.5767	0.3294	0.1586	0.0625
	2	0.9962	0.9743	0.9262	0.8520	0.6471	0.4199	0.2266
	3	0.9998	0.9973	0.9879	0.9667	0.8740	0.7102	0.5000
	4	1.0000	0.9998	0.9988	0.9953	0.9712	0.9037	0.7734
	5	1.0000	1.0000	0.9999	0.9996	0.9962	0.9812	0.9375
	6	1.0000	1.0000	1.0000	1.0000	0.9998	0.9984	0.9922
	7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
8	0	0.6634	0.4305	0.2725	0.1678	0.0576	0.0168	0.0039
	1	0.9428	0.8131	0.6572	0.5033	0.2553	0.1064	0.0352
	2	0.9942	0.9619	0.8948	0.7969	0.5518	0.3154	0.1445
	3	0.9996	0.9950	0.9786	0.9437	0.8059	0.5941	0.3633
	4	1.0000	0.9996	0.9971	0.9896	0.9420	0.8263	0.6367
	5	1.0000	1.0000	0.9998	0.9988	0.9887	0.9502	0.8555
	6	1.0000	1.0000	1.0000	0.9999	0.9987	0.9915	0.9648
	7	1.0000	1.0000	1.0000	1.0000	0.9999	0.9993	0.9961
	8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	0	0.6302	0.3874	0.2316	0.1342	0.0404	0.0101	0.0020
	1	0.9288	0.7748	0.5995	0.4362	0.1960	0.0705	0.0195
	2	0.9916	0.9470	0.8591	0.7382	0.4628	0.2318	0.0898
	3	0.9994	0.9917	0.9661	0.9144	0.7297	0.4826	0.2539
	4	1.0000	0.9991	0.9944	0.9804	0.9012	0.7334	0.5000
	5	1.0000	0.9999	0.9994	0.9969	0.9747	0.9006	0.7461
	6	1.0000	1.0000	1.0000	0.9997	0.9957	0.9750	0.9102
	7	1.0000	1.0000	1.0000	1.0000	0.9996	0.9962	0.9805
	8	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9980
	9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
10	0	0.5987	0.3487	0.1969	0.1074	0.0282	0.0060	0.0010
	1	0.9139	0.7361	0.5443	0.3758	0.1493	0.0464	0.0107
	2	0.9885	0.9298	0.8202	0.6778	0.3828	0.1673	0.0547
	3	0.9990	0.9872	0.9500	0.8791	0.6496	0.3823	0.1719
	4	0.9999	0.9984	0.9901	0.9672	0.8497	0.6331	0.3770
	5	1.0000	0.9999	0.9986	0.9936	0.9527	0.8338	0.6230
	6	1.0000	1.0000	0.9999	0.9991	0.9894	0.9452	0.8281
	7	1.0000	1.0000	1.0000	0.9999	0.9984	0.9877	0.9453
	8	1.0000	1.0000	1.0000	1.0000	0.9999	0.9983	0.9893
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9990
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
11	0	0.5688	0.3138	0.1673	0.0859	0.0198	0.0036	0.0005
	1	0.8981	0.6974	0.4922	0.3221	0.1130	0.0302	0.0059
	2	0.9848	0.9104	0.7788	0.6174	0.3127	0.1189	0.0327
	3	0.9984	0.9815	0.9306	0.8389	0.5696	0.2963	0.1133
	4	0.9999	0.9972	0.9841	0.9496	0.7897	0.5328	0.2744
	5	1.0000	0.9997	0.9973	0.9883	0.9218	0.7535	0.5000
	6	1.0000	1.0000	0.9997	0.9980	0.9784	0.9006	0.7256
	7	1.0000	1.0000	1.0000	0.9998	0.9957	0.9707	0.8867
	8	1.0000	1.0000	1.0000	1.0000	0.9994	0.9941	0.9673
	9	1.0000	1.0000	1.0000	1.0000	1.0000	0.9993	0.9941
	10	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
12	0	0.5404	0.2824	0.1422	0.0687	0.0138	0.0022	0.0002
	1	0.8816	0.6590	0.4435	0.2749	0.0850	0.0196	0.0032
	2	0.9804	0.8891	0.7358	0.5583	0.2528	0.0834	0.0193
	3	0.9978	0.9744	0.9078	0.7946	0.4925	0.2253	0.0730
	4	0.9998	0.9957	0.9761	0.9274	0.7237	0.4382	0.1938
	5	1.0000	0.9995	0.9954	0.9806	0.8822	0.6652	0.3872
	6	1.0000	0.9999	0.9993	0.9961	0.9614	0.8418	0.6128
	7	1.0000	1.0000	0.9999	0.9994	0.9905	0.9427	0.8062
	8	1.0000	1.0000	1.0000	0.9999	0.9983	0.9847	0.9270
	9	1.0000	1.0000	1.0000	1.0000	0.9998	0.9972	0.9807
	10	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9968
	11	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
13	0	0.5133	0.2542	0.1209	0.0550	0.0097	0.0013	0.0001
	1	0.8646	0.6213	0.3983	0.2336	0.0637	0.0126	0.0017
	2	0.9755	0.8661	0.6920	0.5017	0.2025	0.0579	0.0112
	3	0.9969	0.9658	0.8820	0.7473	0.4206	0.1686	0.0461
	4	0.9997	0.9935	0.9658	0.9009	0.6543	0.3530	0.1334
	5	1.0000	0.9991	0.9925	0.9700	0.8346	0.5744	0.2905
	6	1.0000	0.9999	0.9987	0.9930	0.9376	0.7712	0.5000
	7	1.0000	1.0000	0.9998	0.9988	0.9818	0.9023	0.7095
	8	1.0000	1.0000	1.0000	0.9998	0.9960	0.9679	0.8666
	9	1.0000	1.0000	1.0000	1.0000	0.9993	0.9922	0.9539
	10	1.0000	1.0000	1.0000	1.0000	0.9999	0.9987	0.9888
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9983
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
14	0	0.4877	0.2288	0.1028	0.0440	0.0068	0.0008	0.0001
	1	0.8470	0.5846	0.3567	0.1979	0.0475	0.0081	0.0009
	2	0.9699	0.8416	0.6479	0.4481	0.1608	0.0398	0.0065
	3	0.9958	0.9559	0.8535	0.6982	0.3552	0.1243	0.0287
	4	0.9996	0.9908	0.9533	0.8702	0.5842	0.2793	0.0898
	5	1.0000	0.9985	0.9885	0.9561	0.7805	0.4859	0.2120
	6	1.0000	0.9998	0.9978	0.9884	0.9067	0.6925	0.3953
	7	1.0000	1.0000	0.9997	0.9976	0.9685	0.8499	0.6047
	8	1.0000	1.0000	1.0000	0.9996	0.9917	0.9417	0.7880
	9	1.0000	1.0000	1.0000	1.0000	0.9983	0.9825	0.9102
	10	1.0000	1.0000	1.0000	1.0000	0.9998	0.9961	0.9713
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9935
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9991
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
15	0	0.4633	0.2059	0.0874	0.0352	0.0047	0.0005	0.0000
	1	0.8290	0.5490	0.3186	0.1671	0.0353	0.0052	0.0005
	2	0.9638	0.8159	0.6042	0.3980	0.1268	0.0271	0.0037
	3	0.9945	0.9444	0.8227	0.6482	0.2969	0.0905	0.0176
	4	0.9994	0.9873	0.9383	0.8358	0.5155	0.2173	0.0592
	5	0.9999	0.9978	0.9832	0.9389	0.7216	0.4032	0.1509
	6	1.0000	0.9997	0.9964	0.9819	0.8689	0.6098	0.3036

n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
	7	1.0000	1.0000	0.9994	0.9958	0.9500	0.7869	0.5000
	8	1.0000	1.0000	0.9999	0.9992	0.9848	0.9050	0.6964
	9	1.0000	1.0000	1.0000	0.9999	0.9963	0.9662	0.8491
	10	1.0000	1.0000	1.0000	1.0000	0.9993	0.9907	0.9408
	11	1.0000	1.0000	1.0000	1.0000	0.9999	0.9981	0.9824
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9963
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
16	0	0.4401	0.1853	0.0743	0.0281	0.0033	0.0003	0.0000
	1	0.8108	0.5147	0.2839	0.1407	0.0261	0.0033	0.0003
	2	0.9571	0.7892	0.5614	0.3518	0.0994	0.0183	0.0021
	3	0.9930	0.9316	0.7899	0.5981	0.2459	0.0651	0.0106
	4	0.9991	0.9830	0.9209	0.7982	0.4499	0.1666	0.0384
	5	0.9999	0.9967	0.9765	0.9183	0.6598	0.3288	0.1051
	6	1.0000	0.9995	0.9944	0.9733	0.8247	0.5272	0.2272
	7	1.0000	0.9999	0.9989	0.9930	0.9256	0.7161	0.4018
	8	1.0000	1.0000	0.9998	0.9985	0.9743	0.8577	0.5982
	9	1.0000	1.0000	1.0000	0.9998	0.9929	0.9417	0.7728
	10	1.0000	1.0000	1.0000	1.0000	0.9984	0.9809	0.8949
	11	1.0000	1.0000	1.0000	1.0000	0.9997	0.9951	0.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
17	0	0.4181	0.1668	0.0631	0.0225	0.0023	0.0002	0.0000
	1	0.7922	0.4818	0.2525	0.1182	0.0193	0.0021	0.0001
	2	0.9497	0.7618	0.5198	0.3096	0.0774	0.0123	0.0012
	3	0.9912	0.9174	0.7556	0.5489	0.2019	0.0464	0.0064
	4	0.9988	0.9779	0.9013	0.7582	0.3887	0.1260	0.0245
	5	0.9999	0.9953	0.9681	0.8943	0.5968	0.2639	0.0717
	6	1.0000	0.9992	0.9917	0.9623	0.7752	0.4478	0.1662
	7	1.0000	0.9999	0.9983	0.9891	0.8954	0.6405	0.3145
	8	1.0000	1.0000	0.9997	0.9974	0.9597	0.8011	0.5000
	9	1.0000	1.0000	1.0000	0.9995	0.9873	0.9081	0.6855

n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
	10	1.0000	1.0000	1.0000	0.9999	0.9968	0.9652	0.8338
	11	1.0000	1.0000	1.0000	1.0000	0.9993	0.9894	0.9283
	12	1.0000	1.0000	1.0000	1.0000	0.9999	0.9975	0.9755
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9995	0.9936
	14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9988
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
18	0	0.3972	0.1501	0.0536	0.0180	0.0016	0.0001	0.0000
	1	0.7735	0.4503	0.2241	0.0991	0.0142	0.0013	0.0001
	2	0.9419	0.7338	0.4797	0.2713	0.0600	0.0082	0.0007
	3	0.9891	0.9018	0.7202	0.5010	0.1646	0.0328	0.0038
	4	0.9985	0.9718	0.8794	0.7164	0.3327	0.0942	0.0154
	5	0.9998	0.9936	0.9581	0.8671	0.5344	0.2088	0.0481
	6	1.0000	0.9988	0.9882	0.9487	0.7217	0.3743	0.1189
	7	1.0000	0.9998	0.9973	0.9837	0.8593	0.5634	0.2403
	8	1.0000	1.0000	0.9995	0.9957	0.9404	0.7368	0.4073
	9	1.0000	1.0000	0.9999	0.9991	0.9790	0.8653	0.5927
	10	1.0000	1.0000	1.0000	0.9998	0.9939	0.9424	0.7597
	11	1.0000	1.0000	1.0000	1.0000	0.9986	0.9797	0.8811
	12	1.0000	1.0000	1.0000	1.0000	0.9997	0.9942	0.9519
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987	0.9846
	14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998	0.9962
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9993
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
19	0	0.3774	0.1351	0.0456	0.0144	0.0011	0.0001	0.0000
	1	0.7547	0.4203	0.1985	0.0829	0.0104	0.0008	0.0000
	2	0.9335	0.7054	0.4413	0.2369	0.0462	0.0055	0.0004
	3	0.9868	0.8850	0.6841	0.4551	0.1332	0.0230	0.0022
	4	0.9980	0.9648	0.8556	0.6733	0.2822	0.0696	0.0096
	5	0.9998	0.9914	0.9463	0.8369	0.4739	0.1629	0.0318
	6	1.0000	0.9983	0.9837	0.9324	0.6655	0.3081	0.0835
	7	1.0000	0.9997	0.9959	0.9767	0.8180	0.4878	0.1796
	8	1.0000	1.0000	0.9992	0.9933	0.9161	0.6675	0.3238

n	$[t]$	0.05	0.1	0.15	0.20	0.30	0.40	0.50
	9	1.0000	1.0000	0.9999	0.9984	0.9674	0.8139	0.5000
	10	1.0000	1.0000	1.0000	0.9997	0.9895	0.9115	0.6762
	11	1.0000	1.0000	1.0000	1.0000	0.9972	0.9648	0.8204
	12	1.0000	1.0000	1.0000	1.0000	0.9994	0.9884	0.9165
	13	1.0000	1.0000	1.0000	1.0000	0.9999	0.9969	0.9682
	14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9994	0.9904
	15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9978
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9996
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
20	0	0.3585	0.1216	0.0388	0.0115	0.0008	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0076	0.0005	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0355	0.0036	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.1071	0.0160	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.2375	0.0510	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.4164	0.1256	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.6080	0.2500	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.7723	0.4159	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.8867	0.5956	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9520	0.7553	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9829	0.8725	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9949	0.9435	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9987	0.9790	0.8684
	13	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	20	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

D.2. Probabilidades Poisson

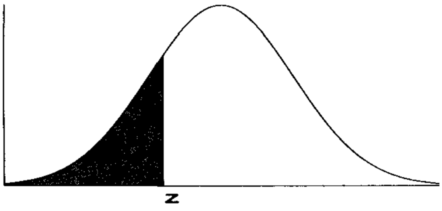
$$P(X \le t) = \sum_{k=0}^{[t]} e^{-\lambda} \frac{\lambda^k}{k!}$$

$\lambda \diagdown [t]$	0	1	2	3	4	5	6
0.1	0.9048	0.9953	0.9998	1.0000	1.0000	1.0000	1.0000
0.2	0.8187	0.9825	0.9989	0.9999	1.0000	1.0000	1.0000
0.3	0.7408	0.9631	0.9964	0.9997	1.0000	1.0000	1.0000
0.4	0.6703	0.9384	0.9921	0.9992	0.9999	1.0000	1.0000
0.5	0.6065	0.9098	0.9856	0.9982	0.9998	1.0000	1.0000
0.6	0.5488	0.8781	0.9769	0.9966	0.9996	1.0000	1.0000
0.7	0.4966	0.8442	0.9659	0.9942	0.9992	0.9999	1.0000
0.8	0.4493	0.8088	0.9526	0.9909	0.9986	0.9998	1.0000
0.9	0.4066	0.7725	0.9371	0.9865	0.9977	0.9997	1.0000
1.0	0.3679	0.7358	0.9197	0.9810	0.9963	0.9994	0.9999
1.2	0.3012	0.6626	0.8795	0.9662	0.9923	0.9985	0.9997
1.4	0.2466	0.5918	0.8335	0.9463	0.9857	0.9968	0.9994
1.6	0.2019	0.5249	0.7834	0.9212	0.9763	0.9940	0.9987
1.8	0.1653	0.4628	0.7306	0.8913	0.9636	0.9896	0.9974
2.0	0.1353	0.4060	0.6767	0.8571	0.9473	0.9834	0.9955
2.5	0.0821	0.2873	0.5438	0.7576	0.8912	0.9580	0.9858
3.0	0.0498	0.1991	0.4232	0.6472	0.8153	0.9161	0.9665
3.5	0.0302	0.1359	0.3208	0.5366	0.7254	0.8576	0.9347
4.0	0.0183	0.0916	0.2381	0.4335	0.6288	0.7851	0.8893
5.0	0.0067	0.0404	0.1247	0.2650	0.4405	0.6160	0.7622
6.0	0.0025	0.0174	0.0620	0.1512	0.2851	0.4457	0.6063
7.0	0.0009	0.0073	0.0296	0.0818	0.1730	0.3007	0.4497
8.0	0.0003	0.0030	0.0138	0.0424	0.0996	0.1912	0.3134
9.0	0.0001	0.0012	0.0062	0.0212	0.0550	0.1157	0.2068
10.0	0.0000	0.0005	0.0028	0.0103	0.0293	0.0671	0.1301
15.0	0.0000	0.0000	0.0000	0.0002	0.0009	0.0028	0.0076
20.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003

$\lambda \diagdown [t]$	7	8	9	10	11	12
0.1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.7	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
0.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
1.4	0.9999	1.0000	1.0000	1.0000	1.0000	1.0000
1.6	0.9997	1.0000	1.0000	1.0000	1.0000	1.0000
1.8	0.9994	0.9999	1.0000	1.0000	1.0000	1.0000
2.0	0.9989	0.9998	1.0000	1.0000	1.0000	1.0000
2.5	0.9958	0.9989	0.9997	0.9999	1.0000	1.0000
3.0	0.9881	0.9962	0.9989	0.9997	0.9999	1.0000
3.5	0.9733	0.9901	0.9967	0.9990	0.9997	0.9999
4.0	0.9489	0.9786	0.9919	0.9972	0.9991	0.9997
5.0	0.8666	0.9319	0.9682	0.9863	0.9945	0.9980
6.0	0.7440	0.8472	0.9161	0.9574	0.9799	0.9912
7.0	0.5987	0.7291	0.8305	0.9015	0.9467	0.9730
8.0	0.4530	0.5925	0.7166	0.8159	0.8881	0.9362
9.0	0.3239	0.4557	0.5874	0.7060	0.8030	0.8758
10.0	0.2202	0.3328	0.4579	0.5830	0.6968	0.7916
15.0	0.0180	0.0374	0.0699	0.1185	0.1848	0.2676
20.0	0.0008	0.0021	0.0050	0.0108	0.0214	0.0390

D.3. Función de Distribución Normal Estándar

$$\Phi(z) = P(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{t^2}{2}}\, dt$$

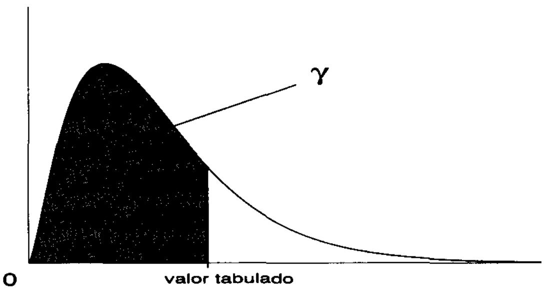


z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990

D.4. Función de Distribución Ji-cuadrado

$$F_{\chi^2}(t) = \int_0^t \frac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)} x^{\frac{n-2}{2}} e^{-\frac{x}{2}} dx = \gamma$$



$n \diagdown \gamma$	0.995	0.990	0.975	0.950	0.900	0.750
1	7.8794	6.6349	5.0239	3.8415	2.7055	1.3233
2	10.5966	9.2103	7.3778	5.9915	4.6052	2.7726
3	12.8382	11.3449	9.3484	7.8147	6.2514	4.1083
4	14.8603	13.2767	11.1433	9.4877	7.7794	5.3853
5	16.7496	15.0863	12.8325	11.0705	9.2364	6.6257
6	18.5476	16.8119	14.4494	12.5916	10.6446	7.8408
7	20.2777	18.4753	16.0128	14.0671	12.0170	9.0371
8	21.9550	20.0902	17.5345	15.5073	13.3616	10.2189
9	23.5894	21.6660	19.0228	16.9190	14.6837	11.3888
10	25.1882	23.2093	20.4832	18.3070	15.9872	12.5489
11	26.7568	24.7250	21.9200	19.6751	17.2750	13.7007

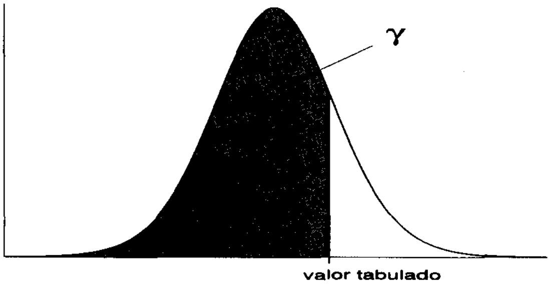
$n \diagdown \gamma$	0.995	0.990	0.975	0.950	0.900	0.750
12	28.2995	26.2170	23.3367	21.0261	18.5493	14.8454
13	29.8195	27.6882	24.7356	22.3620	19.8119	15.9839
14	31.3193	29.1412	26.1189	23.6848	21.0641	17.1169
15	32.8013	30.5779	27.4884	24.9958	22.3071	18.2451
16	34.2672	31.9999	28.8454	26.2962	23.5418	19.3689
17	35.7185	33.4087	30.1910	27.5871	24.7690	20.4887
18	37.1565	34.8053	31.5264	28.8693	25.9894	21.6049
19	38.5823	36.1909	32.8523	30.1435	27.2036	22.7178
20	39.9968	37.5662	34.1696	31.4104	28.4120	23.8277
21	41.4011	38.9322	35.4789	32.6706	29.6151	24.9348
22	42.7957	40.2894	36.7807	33.9244	30.8133	26.0393
23	44.1813	41.6384	38.0756	35.1725	32.0069	27.1413
24	45.5585	42.9798	39.3641	36.4150	33.1962	28.2412
25	46.9279	44.3141	40.6465	37.6525	34.3816	29.3389
26	48.2899	45.6417	41.9232	38.8851	35.5632	30.4346
27	49.6449	46.9629	43.1945	40.1133	36.7412	31.5284
28	50.9934	48.2782	44.4608	41.3371	37.9159	32.6205
29	52.3356	49.5879	45.7223	42.5570	39.0875	33.7109
30	53.6720	50.8922	46.9792	43.7730	40.2560	34.7997
40	66.7660	63.6907	59.3417	55.7585	51.8051	45.6160
50	79.4900	76.1539	71.4202	67.5048	63.1671	56.3336
60	91.9517	88.3794	83.2977	79.0819	74.3970	66.9815
70	104.2149	100.4252	95.0232	90.5312	85.5270	77.5767
80	116.3211	112.3288	106.6286	101.8795	96.5782	88.1303
90	128.2989	124.1163	118.1359	113.1453	107.5650	98.6499
100	140.1695	135.8067	129.5612	124.3421	118.4980	109.1412
200	255.2642	249.4451	241.0579	233.9943	226.0210	213.1022
500	585.2066	576.4928	563.8515	553.1268	540.9303	520.9505
1000	1118.9481	1106.9690	1089.5309	1074.6794	1057.7239	1029.7898

$n \diagdown \gamma$	0.500	0.250	0.100	0.050	0.025	0.010	0.005
1	0.455	0.102	0.016	0.004	0.001	0.000	0.000
2	1.386	0.575	0.211	0.103	0.051	0.020	0.010
3	2.366	1.213	0.584	0.352	0.216	0.115	0.072

$n \setminus \gamma$	0.500	0.250	0.100	0.050	0.025	0.010	0.005
4	3.357	1.923	1.064	0.711	0.484	0.297	0.207
5	4.351	2.675	1.610	1.145	0.831	0.554	0.412
6	5.348	3.455	2.204	1.635	1.237	0.872	0.676
7	6.346	4.255	2.833	2.167	1.690	1.239	0.989
8	7.344	5.071	3.490	2.733	2.180	1.646	1.344
9	8.343	5.899	4.168	3.325	2.700	2.088	1.735
10	9.342	6.737	4.865	3.940	3.247	2.558	2.156
11	10.341	7.584	5.578	4.575	3.816	3.053	2.603
12	11.340	8.438	6.304	5.226	4.404	3.571	3.074
13	12.340	9.299	7.042	5.892	5.009	4.107	3.565
14	13.339	10.165	7.790	6.571	5.629	4.660	4.075
15	14.339	11.037	8.547	7.261	6.262	5.229	4.601
16	15.338	11.912	9.312	7.962	6.908	5.812	5.142
17	16.338	12.792	10.085	8.672	7.564	6.408	5.697
18	17.338	13.675	10.865	9.390	8.231	7.015	6.265
19	18.338	14.562	11.651	10.117	8.907	7.633	6.844
20	19.337	15.452	12.443	10.851	9.591	8.260	7.434
21	20.337	16.344	13.240	11.591	10.283	8.897	8.034
22	21.337	17.240	14.041	12.338	10.982	9.542	8.643
23	22.337	18.137	14.848	13.091	11.689	10.196	9.260
24	23.337	19.037	15.659	13.848	12.401	10.856	9.886
25	24.337	19.939	16.473	14.611	13.120	11.524	10.520
26	25.336	20.843	17.292	15.379	13.844	12.198	11.160
27	26.336	21.749	18.114	16.151	14.573	12.879	11.808
28	27.336	22.657	18.939	16.928	15.308	13.565	12.461
29	28.336	23.567	19.768	17.708	16.047	14.256	13.121
30	29.336	24.478	20.599	18.493	16.791	14.953	13.787
40	39.335	33.660	29.051	26.509	24.433	22.164	20.707
50	49.335	42.942	37.689	34.764	32.357	29.707	27.991
60	59.335	52.294	46.459	43.188	40.482	37.485	35.534
70	69.334	61.698	55.329	51.739	48.758	45.442	43.275
80	79.334	71.145	64.278	60.391	57.153	53.540	51.172
90	89.334	80.625	73.291	69.126	65.647	61.754	59.196
100	99.334	90.133	82.358	77.929	74.222	70.065	67.328
200	199.334	186.172	174.835	168.279	162.728	156.432	152.241
500	499.333	478.323	459.926	449.147	439.936	429.388	422.303
1000	999.333	969.484	943.133	927.594	914.257	898.912	888.564

D.5. Función de Distribución t-Student

$$F_t(x) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\Gamma\left(\frac{n}{2}\right)} \left(1 - u^2\right)^{-\frac{n+1}{2}} du = \gamma$$

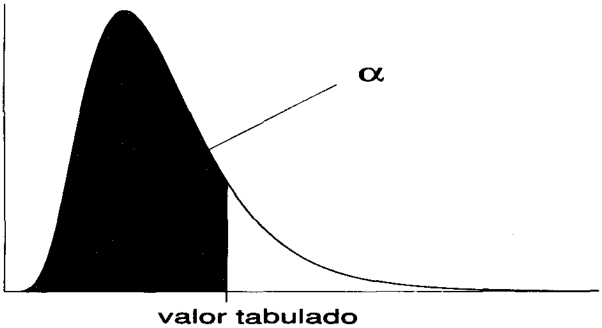


$n \diagdown \gamma$	0.990	0.975	0.950	0.090	0.750
1	31.8205	12.7062	6.3138	3.0777	1.0000
2	6.9646	4.3027	2.9200	1.8856	0.8165
3	4.5407	3.1824	2.3534	1.6377	0.7649
4	3.7469	2.7764	2.1318	1.5332	0.7407
5	3.3649	2.5706	2.0150	1.4759	0.7267
6	3.1427	2.4469	1.9432	1.4398	0.7176
7	2.9980	2.3646	1.8946	1.4149	0.7111
8	2.8965	2.3060	1.8595	1.3968	0.7064
9	2.8214	2.2622	1.8331	1.3830	0.7027
10	2.7638	2.2281	1.8125	1.3722	0.6998
11	2.7181	2.2010	1.7959	1.3634	0.6974
12	2.6810	2.1788	1.7823	1.3562	0.6955
13	2.6503	2.1604	1.7709	1.3502	0.6938
14	2.6245	2.1448	1.7613	1.3450	0.6924
15	2.6025	2.1314	1.7531	1.3406	0.6912
16	2.5835	2.1199	1.7459	1.3368	0.6901
17	2.5669	2.1098	1.7396	1.3334	0.6892
18	2.5524	2.1009	1.7341	1.3304	0.6884
19	2.5395	2.0930	1.7291	1.3277	0.6876
20	2.5280	2.0860	1.7247	1.3253	0.6870
21	2.5176	2.0796	1.7207	1.3232	0.6864
22	2.5083	2.0739	1.7171	1.3212	0.6858

$n \diagdown \gamma$	0.990	0.975	0.950	0.090	0.750
23	2.4999	2.0687	1.7139	1.3195	0.6853
24	2.4922	2.0639	1.7109	1.3178	0.6848
25	2.4851	2.0595	1.7081	1.3163	0.6844
26	2.4786	2.0555	1.7056	1.3150	0.6840
27	2.4727	2.0518	1.7033	1.3137	0.6837
28	2.4671	2.0484	1.7011	1.3125	0.6834
29	2.4620	2.0452	1.6991	1.3114	0.6830
30	2.4573	2.0423	1.6973	1.3104	0.6828
40	2.4233	2.0211	1.6839	1.3031	0.6807
50	2.4033	2.0086	1.6759	1.2987	0.6794
60	2.3901	2.0003	1.6706	1.2958	0.6786
70	2.3808	1.9944	1.6669	1.2938	0.6780
80	2.3739	1.9901	1.6641	1.2922	0.6776
90	2.3685	1.9867	1.6620	1.2910	0.6772
100	2.3642	1.9840	1.6602	1.2901	0.6770
150	2.3515	1.9759	1.6551	1.2872	0.6761
200	2.3451	1.9719	1.6525	1.2858	0.6757
300	2.3388	1.9679	1.6499	1.2844	0.6753
400	2.3357	1.9659	1.6487	1.2837	0.6751
600	2.3326	1.9639	1.6474	1.2830	0.6749
800	2.3310	1.9629	1.6468	1.2826	0.6748
1000	2.3301	1.9623	1.6464	1.2824	0.6747
Inf	2.3263	1.9600	1.6449	1.2816	0.6745

D.6. Función de Distribución F

$$F(x) = \int_{-\infty}^{x} \frac{\Gamma\left(\frac{m+n}{2}\right)}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} m^{\frac{m}{2}} n^{\frac{n}{2}} \frac{u^{\frac{m}{2}-1}}{(n+mu)^{\frac{m+n}{2}}} du$$



D.6.1. Tabla para $\alpha = 0.95$

$n \backslash m$	1	2	3	4	5	6	7	8	9	10
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9	240.5	241.9
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38

$\alpha = 0.95$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
34	4.13	3.28	2.88	2.65	2.49	2.38	2.29	2.23	2.17	2.12
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93
200	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88
500	3.86	3.01	2.62	2.39	2.23	2.12	2.03	1.96	1.90	1.85

$\alpha = 0.95$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
1	243.9	245.4	246.5	247.3	248.0	249.1	250.1	251.1	252.2	253.0
2	19.41	19.42	19.43	19.44	19.45	19.45	19.46	19.47	19.48	19.49
3	8.74	8.71	8.69	8.67	8.66	8.64	8.62	8.59	8.57	8.55
4	5.91	5.87	5.84	5.82	5.80	5.77	5.75	5.72	5.69	5.66
5	4.68	4.64	4.60	4.58	4.56	4.53	4.50	4.46	4.43	4.41
6	4.00	3.96	3.92	3.90	3.87	3.84	3.81	3.77	3.74	3.71
7	3.57	3.53	3.49	3.47	3.44	3.41	3.38	3.34	3.30	3.27
8	3.28	3.24	3.20	3.17	3.15	3.12	3.08	3.04	3.01	2.97
9	3.07	3.03	2.99	2.96	2.94	2.90	2.86	2.83	2.79	2.76
10	2.91	2.86	2.83	2.80	2.77	2.74	2.70	2.66	2.62	2.59

$\alpha = 0.95$

$n \setminus m$	12	14	16	18	20	24	30	40	60	100
11	2.79	2.74	2.70	2.67	2.65	2.61	2.57	2.53	2.49	2.46
12	2.69	2.64	2.60	2.57	2.54	2.51	2.47	2.43	2.38	2.35
13	2.60	2.55	2.51	2.48	2.46	2.42	2.38	2.34	2.30	2.26
14	2.53	2.48	2.44	2.41	2.39	2.35	2.31	2.27	2.22	2.19
15	2.48	2.42	2.38	2.35	2.33	2.29	2.25	2.20	2.16	2.12
16	2.42	2.37	2.33	2.30	2.28	2.24	2.19	2.15	2.11	2.07
17	2.38	2.33	2.29	2.26	2.23	2.19	2.15	2.10	2.06	2.02
18	2.34	2.29	2.25	2.22	2.19	2.15	2.11	2.06	2.02	1.98
19	2.31	2.26	2.21	2.18	2.16	2.11	2.07	2.03	1.98	1.94
20	2.28	2.22	2.18	2.15	2.12	2.08	2.04	1.99	1.95	1.91
21	2.25	2.20	2.16	2.12	2.10	2.05	2.01	1.96	1.92	1.88
22	2.23	2.17	2.13	2.10	2.07	2.03	1.98	1.94	1.89	1.85
23	2.20	2.15	2.11	2.08	2.05	2.01	1.96	1.91	1.86	1.82
24	2.18	2.13	2.09	2.05	2.03	1.98	1.94	1.89	1.84	1.80
25	2.16	2.11	2.07	2.04	2.01	1.96	1.92	1.87	1.82	1.78
26	2.15	2.09	2.05	2.02	1.99	1.95	1.90	1.85	1.80	1.76
27	2.13	2.08	2.04	2.00	1.97	1.93	1.88	1.84	1.79	1.74
28	2.12	2.06	2.02	1.99	1.96	1.91	1.87	1.82	1.77	1.73
29	2.10	2.05	2.01	1.97	1.94	1.90	1.85	1.81	1.75	1.71
30	2.09	2.04	1.99	1.96	1.93	1.89	1.84	1.79	1.74	1.70
34	2.05	1.99	1.95	1.92	1.89	1.84	1.80	1.75	1.69	1.65
40	2.00	1.95	1.90	1.87	1.84	1.79	1.74	1.69	1.64	1.59
50	1.95	1.89	1.85	1.81	1.78	1.74	1.69	1.63	1.58	1.52
70	1.89	1.84	1.79	1.75	1.72	1.67	1.62	1.57	1.50	1.45
100	1.85	1.79	1.75	1.71	1.68	1.63	1.57	1.52	1.45	1.39
200	1.80	1.74	1.69	1.66	1.62	1.57	1.52	1.46	1.39	1.32
500	1.77	1.71	1.66	1.62	1.59	1.54	1.48	1.42	1.35	1.28

D.6.2. Tabla para $\alpha = 0.9$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
1	39.9	49.5	53.6	55.8	57.2	58.2	58.9	59.4	59.9	60.2
2	8.53	9.00	9.16	9.24	9.29	9.33	9.35	9.37	9.38	9.39
3	5.54	5.46	5.39	5.34	5.31	5.28	5.27	5.25	5.24	5.23
4	4.54	4.32	4.19	4.11	4.05	4.01	3.98	3.95	3.94	3.92
5	4.06	3.78	3.62	3.52	3.45	3.40	3.37	3.34	3.32	3.30
6	3.78	3.46	3.29	3.18	3.11	3.05	3.01	2.98	2.96	2.94
7	3.59	3.26	3.07	2.96	2.88	2.83	2.78	2.75	2.72	2.70
8	3.46	3.11	2.92	2.81	2.73	2.67	2.62	2.59	2.56	2.54
9	3.36	3.01	2.81	2.69	2.61	2.55	2.51	2.47	2.44	2.42
10	3.29	2.92	2.73	2.61	2.52	2.46	2.41	2.38	2.35	2.32
11	3.23	2.86	2.66	2.54	2.45	2.39	2.34	2.30	2.27	2.25
12	3.18	2.81	2.61	2.48	2.39	2.33	2.28	2.24	2.21	2.19
13	3.14	2.76	2,56	2.43	2.35	2.28	2.23	2.20	2.16	2.14
14	3.10	2.73	2.52	2.39	2.31	2.24	2.19	2.15	2.12	2.10
15	3.07	2.70	2.49	2.36	2.27	2.21	2.16	2.12	2.09	2.06
16	3.05	2.67	2.46	2.33	2.24	2.18	2.13	2.09	2.06	2.03
17	3.03	2.64	2.44	2.31	2.22	2.15	2.10	2.06	2.03	2.00
18	3.01	2.62	2.42	2.29	2.20	2.13	2.08	2.04	2.00	1.98
19	2.99	2.61	2.40	2.27	2.18	2.11	2.06	2.02	1.98	1.96
20	2.97	2.59	2.38	2.25	2.16	2.09	2.04	2.00	1.96	1.94
21	2.96	2.57	2.36	2.23	2.14	2.08	2.02	1.98	1.95	1.92
22	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90
23	2.94	2.55	2.34	2.21	2.11	2.05	1.99	1.95	1.92	1.89
24	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88
25	2.92	2.53	2.32	2.18	2.09	2.02	1.97	1.93	1.89	1.87
26	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86
27	2.90	2.51	2.30	2.17	2.07	2.00	1.95	1.91	1.87	1.85
28	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84
29	2.89	2.50	2.28	2.15	2.06	1.99	1.93	1.89	1.86	1.83
30	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82
34	2.86	2.47	2.25	2.12	2.02	1.96	1.90	1.86	1.82	1.79
40	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76
50	2.81	2.41	2.20	2.06	1.97	1.90	1.84	1.80	1.76	1.73
70	2.78	2.38	2.16	2.03	1.93	1.86	1.80	1.76	1.72	1.69
100	2.76	2.36	2.14	2.00	1.91	1.83	1.78	1.73	1.69	1.66

$\alpha = 0.9$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
200	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63
500	2.72	2.31	2.09	1.96	1.86	1.79	1.73	1.68	1.64	1.61

$\alpha = 0.9$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
1	60.7	61.1	61.3	61.6	61.7	62.0	62.3	62.5	62.8	63.0
2	9.41	9.42	9.43	9.44	9.44	9.45	9.46	9.47	9.47	9.48
3	5.22	5.20	5.20	5.19	5.18	5.18	5.17	5.16	5.15	5.14
4	3.90	3.88	3.86	3.85	3.84	3.83	3.82	3.80	3.79	3.78
5	3.27	3.25	3.23	3.22	3.21	3.19	3.17	3.16	3.14	3.13
6	2.90	2.88	2.86	2.85	2.84	2.82	2.80	2.78	2.76	2.75
7	2.67	2.64	2.62	2.61	2.59	2.58	2.56	2.54	2.51	2.50
8	2.50	2.48	2.45	2.44	2.42	2.40	2.38	2.36	2.34	2.32
9	2.38	2.35	2.33	2.31	2.30	2.28	2.25	2.23	2.21	2.19
10	2.28	2.26	2.23	2.22	2.20	2.18	2.16	2.13	2.11	2.09
11	2.21	2.18	2.16	2.14	2.12	2.10	2.08	2.05	2.03	2.01
12	2.15	2.12	2.09	2.08	2.06	2.04	2.01	1.99	1.96	1.94
13	2.10	2.07	2.04	2.02	2.01	1.98	1.96	1.93	1.90	1.88
14	2.05	2.02	2.00	1.98	1.96	1.94	1.91	1.89	1.86	1.83
15	2.02	1.99	1.96	1.94	1.92	1.90	1.87	1.85	1.82	1.79
16	1.99	1.95	1.93	1.91	1.89	1.87	1.84	1.81	1.78	1.76
17	1.96	1.93	1.90	1.88	1.86	1.84	1.81	1.78	1.75	1.73
18	1.93	1.90	1.87	1.85	1.84	1.81	1.78	1.75	1.72	1.70
19	1.91	1.88	1.85	1.83	1.81	1.79	1.76	1.73	1.70	1.67
20	1.89	1.86	1.83	1.81	1.79	1.77	1.74	1.71	1.68	1.65
21	1.87	1.84	1.81	1.79	1.78	1.75	1.72	1.69	1.66	1.63
22	1.86	1.83	1.80	1.78	1.76	1.73	1.70	1.67	1.64	1.61
23	1.84	1.81	1.78	1.76	1.74	1.72	1.69	1.66	1.62	1.59
24	1.83	1.80	1.77	1.75	1.73	1.70	1.67	1.64	1.61	1.58
25	1.82	1.79	1.76	1.74	1.72	1.69	1.66	1.63	1.59	1.56
26	1.81	1.77	1.75	1.72	1.71	1.68	1.65	1.61	1.58	1.55

$\alpha = 0.9$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
27	1.80	1.76	1.74	1.71	1.70	1.67	1.64	1.60	1.57	1.54
28	1.79	1.75	1.73	1.70	1.69	1.66	1.63	1.59	1.56	1.53
29	1.78	1.75	1.72	1.69	1.68	1.65	1.62	1.58	1.55	1.52
30	1.77	1.74	1.71	1.69	1.67	1.64	1.61	1.57	1.54	1.51
34	1.75	1.71	1.68	1.66	1.64	1.61	1.58	1.54	1.50	1.47
40	1.71	1.68	1.65	1.62	1.61	1.57	1.54	1.51	1.47	1.43
50	1.68	1.64	1.61	1.59	1.57	1.54	1.50	1.46	1.42	1.39
70	1.64	1.60	1.57	1.55	1.53	1.49	1.46	1.42	1.37	1.34
100	1.61	1.57	1.54	1.52	1.49	1.46	1.42	1.38	1.34	1.29
200	1.58	1.54	1.51	1.48	1.46	1.42	1.38	1.34	1.29	1.24
500	1.56	1.52	1.49	1.46	1.44	1.40	1.36	1.31	1.26	1.21

D.6.3. Tabla para $\alpha = 0.975$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
1	647.8	799.5	864.2	899.6	921.8	937.1	948.2	956.7	963.3	968.6
2	38.51	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40
3	17.44	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42
4	12.22	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84
5	10.01	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62
6	8.81	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46
7	8.07	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76
8	7.57	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30
9	7.21	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96
10	6.94	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72
11	6.72	5.26	4.63	4.28	4.04	3.88	3.76	3.66	3.59	3.53
12	6.55	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37
13	6.41	4.97	4.35	4.00	3.77	3.60	3.48	3.39	3.31	3.25
14	6.30	4.86	4.24	3.89	3.66	3.50	3.38	3.29	3.21	3.15
15	6.20	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06
16	6.12	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99
17	6.04	4.62	4.01	3.66	3.44	3.28	3.16	3.06	2.98	2.92

$\alpha = 0.975$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
18	5.98	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87
19	5.92	4.51	3.90	3.56	3.33	3.17	3.05	2.96	2.88	2.82
20	5.87	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77
21	5.83	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73
22	5.79	4.38	3.78	3.44	3.22	3.05	2.93	2.84	2.76	2.70
23	5.75	4.35	3.75	3.41	3.18	3.02	2.90	2.81	2.73	2.67
24	5.72	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64
25	5.69	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61
26	5.66	4.27	3.67	3.33	3.10	2.94	2.82	2.73	2.65	2.59
27	5.63	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57
28	5.61	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55
29	5.59	4.20	3.61	3.27	3.04	2.88	2.76	2.67	2.59	2.53
30	5.57	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51
34	5.50	4.12	3.53	3.19	2.97	2.81	2.69	2.59	2.52	2.45
40	5.42	4.05	3.46	3.13	2.90	2.74	2.62	2.53	2.45	2.39
50	5.34	3.97	3.39	3.05	2.83	2.67	2.55	2.46	2.38	2.32
70	5.25	3.89	3.31	2.97	2.75	2.59	2.47	2.38	2.30	2.24
100	5.18	3.83	3.25	2.92	2.70	2.54	2.42	2.32	2.24	2.18
200	5.10	3.76	3.18	2.85	2.63	2.47	2.35	2.26	2.18	2.11
500	5.05	3.72	3.14	2.81	2.59	2.43	2.31	2.22	2.14	2.07

$\alpha = 0.975$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
1	976.7	982.5	986.9	990.3	993.1	997	1001	1006	1010	1013
2	39.41	39.43	39.44	39.44	39.45	39.46	39.46	39.47	39.48	39.49
3	14.34	14.28	14.23	14.20	14.17	14.12	14.08	14.04	13.99	13.96
4	8.75	8.68	8.63	8.59	8.56	8.51	8.46	8.41	8.36	8.32
5	6.52	6.46	6.40	6.36	6.33	6.28	6.23	6.18	6.12	6.08
6	5.37	5.30	5.24	5.20	5.17	5.12	5.07	5.01	4.96	4.92
7	4.67	4.60	4.54	4.50	4.47	4.41	4.36	4.31	4.25	4.21
8	4.20	4.13	4.08	4.03	4.00	3.95	3.89	3.84	3.78	3.74

$\alpha = 0.975$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
9	3.87	3.80	3.74	3.70	3.67	3.61	3.56	3.51	3.45	3.40
10	3.62	3.55	3.50	3.45	3.42	3.37	3.31	3.26	3.20	3.15
11	3.43	3.36	3.30	3.26	3.23	3.17	3.12	3.06	3.00	2.96
12	3.28	3.21	3.15	3.11	3.07	3.02	2.96	2.91	2.85	2.80
13	3.15	3.08	3.03	2.98	2.95	2.89	2.84	2.78	2.72	2.67
14	3.05	2.98	2.92	2.88	2.84	2.79	2.73	2.67	2.61	2.56
15	2.96	2.89	2.84	2.79	2.76	2.70	2.64	2.59	2.52	2.47
16	2.89	2.82	2.76	2.72	2.68	2.63	2.57	2.51	2.45	2.40
17	2.82	2.75	2.70	2.65	2.62	2.56	2.50	2.44	2.38	2.33
18	2.77	2.70	2.64	2.60	2.56	2.50	2.44	2.38	2.32	2.27
19	2.72	2.65	2.59	2.55	2.51	2.45	2.39	2.33	2.27	2.22
20	2.68	2.60	2.55	2.50	2.46	2.41	2.35	2.29	2.22	2.17
21	2.64	2.56	2.51	2.46	2.42	2.37	2.31	2.25	2.18	2.13
22	2.60	2.53	2.47	2.43	2.39	2.33	2.27	2.21	2.14	2.09
23	2.57	2.50	2.44	2.39	2.36	2.30	2.24	2.18	2.11	2.06
24	2.54	2.47	2.41	2.36	2.33	2.27	2.21	2.15	2.08	2.02
25	2.51	2.44	2.38	2.34	2.30	2.24	2.18	2.12	2.05	2.00
26	2.49	2.42	2.36	2.31	2.28	2.22	2.16	2.09	2.03	1.97
27	2.47	2.39	2.34	2.29	2.25	2.19	2.13	2.07	2.00	1.94
28	2.45	2.37	2.32	2.27	2.23	2.17	2.11	2.05	1.98	1.92
29	2.43	2.36	2.30	2.25	2.21	2.15	2.09	2.03	1.96	1.90
30	2.41	2.34	2.28	2.23	2.20	2.14	2.07	2.01	1.94	1.88
34	2.35	2.28	2.22	2.17	2.13	2.07	2.01	1.95	1.88	1.82
40	2.29	2.21	2.15	2.11	2.07	2.01	1.94	1.88	1.80	1.74
50	2.22	2.14	2.08	2.03	1.99	1.93	1.87	1.80	1.72	1.66
70	2.14	2.06	2.00	1.95	1.91	1.85	1.78	1.71	1.63	1.56
100	2.08	2.00	1.94	1.89	1.85	1.78	1.71	1.64	1.56	1.48
200	2.01	1.93	1.87	1.82	1.78	1.71	1.64	1.56	1.47	1.39
500	1.97	1.89	1.83	1.78	1.74	1.67	1.60	1.52	1.42	1.34

D.6.4. Tabla para $\alpha = 0.99$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
1	4052	4999	5403	5625	5764	5859	5928	5982	6022	6056
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37	99.39	99.40
3	34.12	30.82	29.46	28.71	28.24	27.91	27.67	27.49	27.35	27.23
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80	14.66	14.55
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29	10.16	10.05
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10	7.98	7.87
7	12.25	9.55	8.45	7.85	7.46	7.19	6.99	6.84	6.72	6.62
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03	5.91	5.81
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47	5.35	5.26
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06	4.94	4.85
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74	4.63	4.54
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50	4.39	4.30
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30	4.19	4.10
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14	4.03	3.94
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00	3.89	3.80
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89	3.78	3.69
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79	3.68	3.59
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71	3.60	3.51
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63	3.52	3.43
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56	3.46	3.37
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51	3.40	3.31
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41	3.30	3.21
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32	3.22	3.13
26	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09
27	7.68	5.49	4.60	4.11	3.78	3.56	3.39	3.26	3.15	3.06
28	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03
29	7.60	5.42	4.54	4.04	3.73	3.50	3.33	3.20	3.09	3.00
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98
34	7.44	5.29	4.42	3.93	3.61	3.39	3.22	3.09	2.98	2.89
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80
50	7.17	5.06	4.20	3.72	3.41	3.19	3.02	2.89	2.78	2.70
70	7.01	4.92	4.07	3.60	3.29	3.07	2.91	2.78	2.67	2.59
100	6.90	4.82	3.98	3.51	3.21	2.99	2.82	2.69	2.59	2.50

$\alpha = 0.99$

$n \diagdown m$	1	2	3	4	5	6	7	8	9	10
200	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41
500	6.69	4.65	3.82	3.36	3.05	2.84	2.68	2.55	2.44	2.36

$\alpha = 0.99$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
1	6106	6143	6170	6191	6209	6235	6261	6287	6313	6334
2	99.42	99.43	99.44	99.44	99.45	99.46	99.47	99.47	99.48	99.49
3	27.05	26.92	26.83	26.75	26.69	26.60	26.50	26.41	26.32	26.24
4	14.37	14.25	14.15	14.08	14.02	13.93	13.84	13.75	13.65	13.58
5	9.89	9.77	9.68	9.61	9.55	9.47	9.38	9.29	9.20	9.13
6	7.72	7.60	7.52	7.45	7.40	7.31	7.23	7.14	7.06	6.99
7	6.47	6.36	6.28	6.21	6.16	6.07	5.99	5.91	5.82	5.75
8	5.67	5.56	5.48	5.41	5.36	5.28	5.20	5.12	5.03	4.96
9	5.11	5.01	4.92	4.86	4.81	4.73	4.65	4.57	4.48	4.41
10	4.71	4.60	4.52	4.46	4.41	4.33	4.25	4.17	4.08	4.01
11	4.40	4.29	4.21	4.15	4.10	4.02	3.94	3.86	3.78	3.71
12	4.16	4.05	3.97	3.91	3.86	3.78	3.70	3.62	3.54	3.47
13	3.96	3.86	3.78	3.72	3.66	3.59	3.51	3.43	3.34	3.27
14	3.80	3.70	3.62	3.56	3.51	3.43	3.35	3.27	3.18	3.11
15	3.67	3.56	3.49	3.42	3.37	3.29	3.21	3.13	3.05	2.98
16	3.55	3.45	3.37	3.31	3.26	3.18	3.10	3.02	2.93	2.86
17	3.46	3.35	3.27	3.21	3.16	3.08	3.00	2.92	2.83	2.76
18	3.37	3.27	3.19	3.13	3.08	3.00	2.92	2.84	2.75	2.68
19	3.30	3.19	3.12	3.05	3.00	2.92	2.84	2.76	2.67	2.60
20	3.23	3.13	3.05	2.99	2.94	2.86	2.78	2.69	2.61	2.54
21	3.17	3.07	2.99	2.93	2.88	2.80	2.72	2.64	2.55	2.48
22	3.12	3.02	2.94	2.88	2.83	2.75	2.67	2.58	2.50	2.42
23	3.07	2.97	2.89	2.83	2.78	2.70	2.62	2.54	2.45	2.37
24	3.03	2.93	2.85	2.79	2.74	2.66	2.58	2.49	2.40	2.33
25	2.99	2.89	2.81	2.75	2.70	2.62	2.54	2.45	2.36	2.29
26	2.96	2.86	2.78	2.72	2.66	2.58	2.50	2.42	2.33	2.25

$\alpha = 0.99$

$n \diagdown m$	12	14	16	18	20	24	30	40	60	100
27	2.93	2.82	2.75	2.68	2.63	2.55	2.47	2.38	2.29	2.22
28	2.90	2.79	2.72	2.65	2.60	2.52	2.44	2.35	2.26	2.19
29	2.87	2.77	2.69	2.63	2.57	2.49	2.41	2.33	2.23	2.16
30	2.84	2.74	2.66	2.60	2.55	2.47	2.39	2.30	2.21	2.13
34	2.76	2.66	2.58	2.51	2.46	2.38	2.30	2.21	2.12	2.04
40	2.66	2.56	2.48	2.42	2.37	2.29	2.20	2.11	2.02	1.94
50	2.56	2.46	2.38	2.32	2.27	2.18	2.10	2.01	1.91	1.82
70	2.45	2.35	2.27	2.20	2.15	2.07	1.98	1.89	1.78	1.70
100	2.37	2.27	2.19	2.12	2.07	1.98	1.89	1.80	1.69	1.60
200	2.27	2.17	2.09	2.03	1.97	1.89	1.79	1.69	1.58	1.48
500	2.22	2.12	2.04	1.97	1.92	1.83	1.74	1.63	1.52	1.41

Respuestas a ejercicios seleccionados

Capítulo 1.

1. $A{\cap}B = \{(c.c, s), (c, s, c)\}$; $A{\cup}B = \{(c, c, c), (c, s, c), (s, c, s), (s, s, s), (c, c, s), (c, s, s)\}$; $A^c = \{(s, c, c), (s, s, c), (s, c, s), (s, s, s)\}$; $A^c \cap B^c = \{(s, c, c),\ (s, s, c)\}$; $A \cap B^c = \{(c, s, c), (c, c, c)\}$.
2. $A \cap B \cap C^c$; $A \cap B \cap C$; $A \cap B^c \cap C^c$; $A \cup B \cup C$; $E = (A \cap B^c \cap C^c){\cup}(A^c \cap B \cap C^c){\cup}(A^c \cap B^c \cap C){\cup}(A^c \cap B^c \cap C^c)$; $D = E \cup (A \cap B \cap C^c) \cup (A^c \cap B \cap C) \cup (A \cap B^c \cap C)$.
5. a) $B = \bigcup_{i=1}^{n} A_i$ **b)** $C = \bigcap_{i=1}^{n} A_i^c$ **c)** $D = \bigcup_{i=1}^{n} \bigcap_{\substack{j=1 \\ i \neq j}}^{n} \left(A_i \cap A_j^c\right)$
d) $E = C \cup D$. **7.** 33 %; 60 %. **10. a)** $\alpha = \frac{17}{63}$ **b)** $\frac{46}{63}$; $\frac{45}{63}$; $\frac{35}{63}$
13. No **14.** $\frac{2}{3}$; $\frac{4}{9}$; $\frac{2}{9}$. **15. a)** $\frac{8}{100}$ **b)** $\frac{4}{100}$ **c)** $\frac{149}{198}$. **16.** No.
17. $P(A_k) = \frac{1}{2^k}$, para $k = 1, \cdots, n$. **18. a)** $\sum_{j=0}^{7} (-1)^j \binom{7}{j} \left(1 - \frac{j}{7}\right)^{10}$
b) 0.37937 **19.** 0.24763 **20. a)** $\frac{43}{216}$ **b)** $\frac{173}{216}$ **21.** $\frac{43}{84}$
22. $0.58\overline{3}$ **23.** 8.3479×10^{-5} **24.** $\frac{5}{84}$ **25. a)** 0.5177 **b)** 0.1157
c) 0.1975 **d)** $0.\overline{2}$ **26. a)** $1 - p^5\left(2 - p^2\right)$ **b)** $1 - p^4\left(2 - p^2\right)$
27. 0.5687 **28.** $1 - 1 + \frac{1}{2!} - \frac{1}{3!} + \cdots + \frac{(-1)^n}{n!} \simeq e^{-1}$
29. a) $6.3{\times}10^{-2}$ **b)** 0.105 **30.** $4.4582{\times}10^{-2}$ **31.** 0.43925 **32.** $\frac{1}{3}$; $\frac{4}{7}$; 1; $\frac{2}{7}$; $\frac{1}{7}$ **33.** $\frac{5}{6}$; $\frac{1}{3}$ **34. a)** 0.1632 **b)** 0.8748 **c)** 0.9318
35. a) 0.1632 **b)** 6.2645×10^{-2} **c)** 0.85207 **36. b)** 0.2824

37. 7.3 % **38.** 0.235 04 **39.** 0.2 **40.** $\frac{8}{13}$ **41.** 0.43; $\frac{35}{43}$ **42.** $\frac{5}{9}$
43. Φ, $\{1,3\}$, $\{1,4\}$, $\{1,2\}$, $\{2,4\}$, $\{1,2,3,4\}$
51. a) 6.12×10^{-3} **b)** 0.9939 **c)** 3.1874×10^{-2} **52.** $\frac{\binom{n}{k}\binom{N-n}{m-k}}{\binom{N}{m}}$
53. a) $p = \frac{1}{n}\left(\left[\frac{n}{3}\right] + \left[\frac{n}{4}\right] - \left[\frac{n}{12}\right]\right)$ **b)** $\lim\limits_{n\to\infty} p = \frac{1}{2}$.

Capítulo 2.

1. No **2.** $\wp(\Omega)$ **3.** $EX = \frac{9}{7}$ **4. a)** $\frac{1}{3}$ **b)** $\frac{5}{6}$ **c)** $\frac{2}{3}$
7. a) $S = \{1,2,3\}$ **7. b)** $P(X_3 = 3 \mid X_2 \in \{1,2\},\ X_1 = 3) = 0$ y $P(X_3 = 3 \mid X_2 \in \{1,2\}) = \frac{1}{2}$ **8. a)** 7.5758×10^{-2}
b) 4.8351×10^{-2} **9. a)** $c = 1.2$ **c)** 0.25 **d)** 0.710
10. a) $c = \frac{2}{3}$ **c)** 0.746 67**11. a)** 0 ; 0.937 5
b) $f_X(x) = (2-2x)\mathcal{X}_{(0,1)}(x)$ } **12. a)** X discreta, Y mixta y Z continua. **b)** 0.5; 0.3; 0.3**12. c)** 0.5; 0.75; 0.5 **d)** 0; 0.5; 0.
13. $P(X = k) = \frac{2(n-k)+1}{n^2}$ para $k = 1, \cdots, n$.
14. a) $c = 1$ **b)** $c = \ln 2$ **c)** $c = \frac{\pi^2}{6}$ **d)** $c = e^2 - 1$
15. a) $c = \pi^{-1}$ **b)** $c = 1$ **16. a)** $c = \frac{1}{8}$ **b)** 0.703 13
18. a) $F_X(x) = \sum\limits_{k=0}^{[x]} pq^{k-1}$ **b)** $F_X(x) = \frac{[x]([x]+1)(2[x]+1)}{n(n+1)(2n+1)}$
19. a) $P(X = i) = \frac{1}{7}$ para $i = 1, \cdots, 7$ **b)** $P(X \le 2) = \frac{2}{7}$; $P(X = 5) = \frac{1}{7}$. **20.** 0.617 **21.** $P(X = -6000) = 0.23333$, $P(X = -3000) = 0.2$, $P(X = 0) = 0.025$, $P(X = 2000) = 0.33333$, $P(X = 5000) = 0.125$, $P(X = 10000) = 0.08334$
22. $P(X = 0) = 0.18$, $P(X = 1000) = 0.27$, $P(X = 1800) = 0.27$, $P(X = 2000) = 0.07$, $P(X = 2800) = 0.14$, $P(X = 3600) = 0.07$ **25.** $F_X(x) = \frac{2}{3}F_d(x) + \frac{1}{3}F_c(x)$ donde

$$F_d(x) = \begin{cases} 0 & \text{si} \quad x < 1 \\ 0.5 & \text{si} \quad 1 \le x < \frac{5}{4} \\ 1 & \text{si} \quad x \ge \frac{5}{4} \end{cases}, \quad F_c(x) = \begin{cases} 0 & \text{si} \quad x < 1 \\ 4x-4 & \text{si} \quad 1 \le x < \frac{5}{4} \\ 1 & \text{si} \quad x \ge \frac{5}{4} \end{cases}$$

26. $F_X(x) = \frac{3}{4}F_d(x) + \frac{1}{4}F_c(x)$ donde

$$F_d(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{3} & \text{si} \quad 0 \le x < \frac{1}{2} \\ \frac{2}{3} & \text{si} \quad \frac{1}{2} \le x < 2 \\ 1 & \text{si} \quad x \ge 2 \end{cases}, \quad F_c(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ 4x^2 & \text{si} \quad 0 \le x < \frac{1}{2} \\ 1 & \text{si} \quad x \ge \frac{1}{2} \end{cases}$$

28. a) $c = 1$ **b)** $\frac{1}{2}$ **c)** $EX = 0$, $Var(X)$ no existe
d) $F_X(x) = \frac{1}{\pi}\left(\tan^{-1} x + \frac{\pi}{2}\right)$ **29.** $\frac{161}{36}$; $\frac{91}{36}$ **30. a)** $\mu = \frac{1}{2}$, $\sigma^2 = \frac{1}{15}$ **b)** 0.266 **31. b)** $c = 0.5$ **32. a)** $c = \frac{1}{10}$

33.a) $F_Y(x) = \begin{cases} 0 & \text{si} \quad x < 0 \\ \frac{1}{2}\sqrt{x} & \text{si} \quad 0 \le x < 4 \\ 1 & \text{si} \quad x \ge 4 \end{cases}$ **b)** $\frac{1}{2}$; $\frac{1}{4}$; $\frac{3}{4}$

35. $P(X = 0) = 0.08546$; $P(X = 1) = 0.37821$; $P(X = 2) = 0.40812$; $P(X = 3) = 0.12821$ **36. a)** No **b)** Sea X con $P(X = -1) = \frac{1}{9}$; $P\left(X = \frac{1}{2}\right) = \frac{4}{9}$ y $P(X = 2) = \frac{4}{9}$
39. a) $c = 2$ **b)** 1.5 **40.** EX existe, $Var(X)$ no existe.
42. b) $P(0 \le X \le 40) \ge \frac{19}{20}$ **43. a)** $\frac{16}{25}$ **b)** $k = 10$
44. b) $EY \simeq 1.82$ **45.** $\phi_X(t) = \frac{2}{5}e^{it} + \frac{3}{5}$
46. $\phi_X(t) = \frac{1}{(b-a)it}\left(e^{itb} - e^{ita}\right)$ si $t \ne 0$, $\phi_X(t) = 1$ si $t = 0$
47. $\phi_X(t) = 1$ si $t = 0$, $\phi_X(t) = \frac{2-2\cos t}{t^2}$ si $t = 0$
48. a) cuantil inferior$= 2$; cuantil superior$= 3$; mediana cualquier valor en $[2, 3]$ **c)** EX no existe, mediana 5. **49. a)** moda 1 **b)** $r = 7$, $\lambda = \frac{1}{4}$.

Capítulo 3

1. $\binom{n}{\frac{n}{2}}\left(\frac{1}{2}\right)^n$ **2.** 0.608; 2. **3.** 0.515 48 **4. a)** 0.878 42 **b)** 0.121 58
5. $EX = Var(X) = 3.9$ **6.** 0.987 01 **7.** e^{-10}; 0.9897
8. $3.881\,2 \times 10^{-4}$ **9.** $X \stackrel{d}{=} \mathcal{G}\left(\frac{9}{19}\right)$ **10. a)** $3.830\,5 \times 10^{-8}$ **b)** $4.417\,9 \times 10^{-2}$ **c)** $1.237\,9 \times 10^{-3}$ **11.** 0.6767 **12.** 5
13. 15000 **16. a)** 2.861×10^{-5} **b)** 5 **c)** 4 **d)** $9.887\,7 \times 10^{-2}$

17. 0.9945 **18.** $P\left(X^2 = j\right) = pq^{\sqrt{j}-1}$ con $j = 1, 4, 9, \cdots$; $P\left(X + 3 = j\right) = pq^{j-4}$ con $j = 4, 5, 6, \cdots$ **19. a)** 0.674 % **b)** 73.5 % **21.** 0.0293 **22.** 0.968 01 **23.** 76 **24.** 0.43; 0.57 **25.** 262; 393; 345 **26.** 0.783 43 **27.** 2.07×10^{-5} **28. a)** 0.1847 **b)** 0.679 24 **c)** 0.100 1 **29.** 6.8064×10^{-2} **30.** 24; 120 **31.** $P(X = 2) = 0.47368$, $P(X = 4) = 0.11809$, $P(X = -2) = 0.26243$, $P(X = -8) = 0.1458$ **32.** 13 **33. a)** 2 **b)** No existe **34.** 240 **35.** $\lambda = k$ **38.** $p = \frac{k}{n}$ **42.** $\frac{-p \ln p}{1-p}$ **44.** $\phi_X(t) = \left(q + pe^{it}\right)^n$ **45.** $\phi_X(t) = \exp\left(\lambda\left(e^{it} - 1\right)\right)$ **46.** 6.

Capítulo 4

1. a) $\frac{1}{2}$; $\frac{1}{3}$; $\frac{\sqrt{3}}{3}$ **b)** $x = 1.3$ **2. a)** 0.1 **b)** 0.1 **c)** 0.1 **4.** $\frac{3}{4}$; $\frac{1}{2}$ **5.** $\frac{3}{5}$ **6. a)** $\frac{1}{2}$ **b)** $f_Y(x) = \mathcal{X}_{(0,1)}(x)$ **7. a)** $f_Y(x) = \exp(x)\,\mathcal{X}_{(-\infty,0)}(x)$ **b)** $f_Y(x) = \frac{1}{3}(x-2)^{-\frac{2}{3}}\,\mathcal{X}_{(2,3)}(x)$ **c)** $f_Y(x) = \frac{1}{x^2}\mathcal{X}_{(1,\infty)}(x)$ **8.** 5.2 **9. a)**0.422 2 **b)**0.267 3 **c)** 0.892 44 **d)** 0.129 24 **e)** 0.382 92 **10. a)** 5 **b)** 6.46 **c)** 14.36 **11. a)** 15.50 **c)** 12.549 **e)** 14.44; 27.488; 40.647 **12. a)** 0 **b)** 2.41 % **13. a)** 97 **b)** 57 **14.** 89.97 %; 5.82 %; 84.2 % **16.** No. Calcular por ejemplo $P(4 \leq X \leq 9)$. **17.** 0.17619 **18.** $\alpha = 2.14$ **20.** 54000 **22.** 0.283 47 **23.** 0.606 53 **25. a)** e^{-3} **b)** e^{-1} **27. a)** $F(t) = 1 - \exp\left(-\eta\left(\beta^t - 1\right)\right)$ con $\eta = \frac{\alpha}{\ln \beta}$ **b)** $F(t) = 1 - \exp\left(-\gamma t - \eta\left(\beta^t - 1\right)\right)$ con $\eta = \frac{\alpha}{\ln \beta}$ **30.** $F(t) = \left(1 - \frac{1}{t+1}\right)$ **31.** $F(t) = 1 - \exp\left(-\frac{\alpha}{\beta+1}t^{\beta+1}\right)$ **32.** 0.275 **34. a)** $\frac{1}{26}$; 1.3697×10^{-3} **b)** 3.555×10^{-4} **35.** $\sqrt{\frac{2}{\pi}}$ **37.** $Exp(1)$ **38.** $f(x) = \frac{a}{\pi(a^2+x^2)}$, $x \in \mathbb{R}$. **39.** g es la inversa de la función de distribución normal estándar **44.** Cauchy estándar.

Capítulo 5

1. a) 0.5 **b)** 0.8 **d)** $P(Z=-2)=0.1,\ P(Z=-1)=0.2,\ P(Z=0)=0.3,\ P(Z=1)=0.4$ **2.** $\alpha=\beta=0.1,\ \gamma=\eta=\kappa=0,\ \delta=0.2$
3. a) $P(X=0,Y=0)=\frac{1}{55},\ P(X=0,Y=1)=\frac{2}{11},$ $P(X=0,Y=2)=\frac{2}{11},\ P(X=1,Y=0)=\frac{8}{55},\ P(X=1,Y=1)=\frac{4}{11},\ P(X=2,Y=0)=\frac{6}{55}$ **b)** $EX=\frac{8}{11},\ EY=\frac{10}{11}$
4. b) $EX=\frac{88}{121},\ EY=\frac{110}{121}$ **5. a)** 6 **b)** 6 **7. a)** 0.373 08 **b)** 9 **8.** $a=b=k=h=\frac{1}{6},\ d=e=f=0$ **9. a)** $\mathcal{P}(9.4)$ **b)** 0.29 **10.** $P(Z=-1,W=1)=\frac{1}{16},\ P(Z=0,W=0)=\frac{1}{4},$ $P(Z=0,W=2)=\frac{3}{16},\ P(Z=1,W=1)=\frac{1}{8},$ $P(Z=1,W=3)=\frac{1}{16},\ P(Z=2,W=0)=\frac{3}{16},$ $P(Z=2,W=2)=\frac{1}{16},\ P(Z=3,W=1)=\frac{1}{16}.$ **11.** $\frac{41}{324}$; 0.125 1
13. -1 **14.** 0.351 85 **15.** No **16.** $\rho=\frac{-\sigma_2^2}{\sqrt{\sigma_2^2+\sigma_3^2}\sqrt{\sigma_1^2+\sigma_2^2}}$
17. $9-2\sqrt{2}$ **19.** $EX=m\left(\frac{n+1}{2}\right)$ **20.** $EX=\binom{r}{k}\left(\frac{1}{365}\right)^{k-1}\left(\frac{364}{365}\right)^{r-k}$
21. $EY=365-\left(\frac{1}{365}\right)^{-1}\left(\frac{364}{365}\right)^{r}-r\left(\frac{364}{365}\right)^{r-1}$
29. Si, $f(x,y)=\frac{e^{-x}}{\pi(1+y^2)},\ x\geq 0,\ y\in\mathbb{R}.$ **30. a)** $\frac{1}{2}$ **b)** $f_X(x)=f_Y(x)=\frac{1}{2}(\cos x+\sin x),\ x\in\left(0,\frac{\pi}{2}\right)$ **31. b)** $\frac{1}{25}$
32. b) $\frac{1}{4}$ **c)**0.305 56 **34.** $\frac{1}{2}$ **35.** $\frac{1}{8}$
37. $f_Z(x)=\frac{1}{2}(x-4)\mathcal{X}_{(4,5]}(x)+\frac{1}{2}\mathcal{X}_{(5,6]}(x)+\frac{1}{2}(7-x)\mathcal{X}_{(6,7]}(x).$
38. a) $\frac{3}{4}$ **b)** $\frac{2}{3}$ **39. a)** $f_Z(x)=2\left\{x-\frac{1}{\lambda}\left[1-\exp(-\lambda x)\right]\right\}\mathcal{X}_{(0,1)}(x)+2\lambda\exp(-\lambda x)\left\{\frac{1}{\lambda}e^{\lambda}-\frac{1}{\lambda^2}\left(e^{\lambda}-1\right)\right\}\mathcal{X}_{[1,\infty)}(x)$ **b)** $\frac{1}{\lambda^2}$ **40. a)** $\frac{3}{4}$ **b)** $\frac{5}{12}$ **c)** $\frac{3}{4}$ **41.** Falso **42. a)** $f_Z(x)=\lambda\exp(-\lambda|x|)\mathcal{X}_{(0,\infty)}(x)$ **b)** $f_W(x)=\lambda\left(1+\frac{1}{3\sqrt[3]{x^2}}\right)\exp\left(-\lambda\sqrt[3]{x}-\lambda x\right)\mathcal{X}_{(0,\infty)}(x)$ **43.** $\frac{4}{9}$
44. Cauchy **46.** $f_{(Y_1,Y_2)}(y_1,y_2)=1\mathcal{X}_{(0,1)}(y_1)\mathcal{X}_{(0,1)}(y_2)$
47. $f_U(x)=\frac{3}{2}(3x)^2\exp(-3x)\mathcal{X}_{(0,\infty)}(x)$ **48.** $f_Z(x)=\frac{3}{4}\left(1-\frac{x^2}{4}\right)\mathcal{X}_{(0,2)}(x)$ **49.** $\mathcal{X}_n^2$ **50.** $\mathcal{X}_{n-1}^2$ **51.** 0.5464 **52. a)** 27.7 **b)** 0.32 **c)** 4.26 **53.** F_n^1 **56.** No son ni independientes ni

correlacionadas. **57.** $t_{(n-1)}$ **62.** $P(X>0,Y>0,Z>0)=$
$\frac{1}{8}+\frac{1}{4\pi}\left\{\sin^{-1}\rho_1+\sin^{-1}\rho_2+\sin^{-1}\rho_3\right\}$
63. $f(x,y,z)=$
$\frac{1}{2\pi\sqrt{230\pi}}\exp\left\{-\frac{1}{230}\left(39x^2+36y^2+26z^2-44xy+36xz-38yz\right)\right\}$.

Capítulo 6.

1. b) $E(X\mid Y=1)=1$; $E(Y\mid X=1)=\frac{5}{2}$ **3.** $\frac{5}{6}$ **4.** $E(X\mid Y=1)=1.4$ **5.** 15 horas **6. a)** $E(X\mid Y=0)=-\frac{1}{8}$; $E(X\mid Y=1)=\frac{5}{2}$ **b)** $EX=\frac{4}{3}$ **7. a)** $P(X=1,Y=0)=\frac{1}{6}$; $P(X=1,Y=1)=\frac{1}{6}$; $P(X=1,Y=2)=0$; $P(X=2,Y=0)=\frac{1}{6}$; $P(X=2,Y=1)=\frac{1}{3}$; $P(X=2,Y=2)=\frac{1}{6}$ **b)** $E(X\mid Y=0)=\frac{3}{2}$; $E(X\mid Y=1)=\frac{5}{3}$; $E(X\mid Y=2)=2$ **8.** $1+p$ **9. a)** $\frac{2}{3}$ **b)** $\frac{1}{4}$ **c)** $\frac{1}{2}$ **10. a)** 8.8492×10^{-2} **b)** $\frac{y}{2}$ para $y>0$ **c)** $(x+1)$ para $x>0$ **13.** $E(X\mid Y=\frac{1}{4})=1.3137$ **16.** $Var(Y\mid X=0)=\frac{2}{3}$; $Var(Y\mid X=1)=\frac{1}{4}$; $Var(Y\mid X=2)=0$ **17.** $EZ=p\lambda=Var(Z)$ **18.** $E(X\mid Y=1)=\frac{1}{4}e$ **20.** $E(Y\mid X=x)=\frac{2-x}{2}$ **21.** $E(X\mid Y=y)=\frac{2(1-y^3)}{3(1-y^2)}$; $E\left(X^2\mid Y=y\right)=\frac{1}{2}\left(1+y^2\right)$; $Var(X\mid Y=0.5)=2.0062\times10^{-2}$ **22.** $f_{Y|X}(y\mid x)=x\exp(-xy)$, $y\geq0$ **23. b)** $f_{Y|X}(y\mid 2)=\frac{1}{\sqrt{2\pi}\times(0.6)}\exp\left\{-\frac{1}{2\times0.36}(y-2.2)^2\right\}$ **c)** $c=3.187$ **24.** 7 **25. a)** $\frac{3}{7}$ **b)** $\frac{19}{21}$ **26.** 2

27. a) $f(x,y)=\begin{cases}\frac{1}{x} & \text{si} \quad 0<y<x<1\\ 0 & \text{en} \quad \text{otro caso}\end{cases}$ **b)** $f_Y(x)=-(\ln x)\mathcal{X}_{(0,1)}(x)$ **28.** $N\left\{1-\exp\left(-\frac{10}{N}\right)\right\}$ **29.** $EX=62$ **30. a)** $EX=6$ **b)** $E(X\mid Y=1)=7$ **c)** $E(X\mid Y=5)=5.8192$ **31.** $E(X\mid Y=1)=\frac{12}{5}$; $E(X\mid Y=2)=\frac{9}{5}$; $E(X\mid Y=3)=\frac{6}{5}$; $E(X\mid Y=4)=\frac{3}{5}$; $E(X\mid Y=5)=0$ **32.** $\frac{1}{n(n+1)}$

Capítulo 7.

1. $\frac{3}{4}$ **2.** 1 **5. a)** $EX^3 \geq 8$ **b)** $E(\ln X) \leq \ln 2$ **10.** 0.8475
11. 0.99324 **13.** 0.9922 **14.** 1.3×10^{-3} **15.** 117 o más
17. 66564 **18.** 4096 **19.** $\sigma = 0.5$

Símbolos

$|\alpha|$ Valor absoluto del número α.

$a \approx b$ a es aproximadamente igual a b

$A - B$ Diferencia de los conjuntos A y B.

$A \cup B$ Unión de los conjuntos A y B.

$A \cap B$ Intersección de los conjuntos A y B.

A^c Complemento del conjunto A.

$|A|$ Número de elementos del conjunto A.

A^T Transpuesta de la matriz A.

$(a_{ij})_{n\times n}$ Matriz de tamaño $m \times n$ cuyo elemento en la fila i y la columna j es a_{ij} .

$\mathcal{B}$ σ-álgebra de Borel en $\mathbb{R}$.

$\mathcal{B}_n$ σ-álgebra de Borel en $\mathbb{R}^n$.

$\binom{n}{r}$ Combinatorio n, r.

$\mathbb{C}$ Conjunto de números complejos.

EX Valor esperado de la variable aleatoria X.

$E(X \mid Y)$ Valor esperado condicional de X dado Y.

$E(X \mid B)$ Esperanza condicional de X dado el evento B.

$E(X \mid \mathcal{G})$ Esperanza condicional de X dada la σ-álgebra $\mathcal{G}$.

$\det(A)$ Determinante de la matriz A.

$\frac{d^r}{dx^r} f(x)$ Derivada de orden r de la función f.

$f_r(A)$ Frecuencia relativa del evento A.

$f(\cdot)$ Función de una variable real.

$f_{X|Y}(\cdot \mid y)$ Función de densidad condicional de X dado $Y = y$.

$F_{X|Y}(\cdot \mid y)$ Función de distribución condicional de X dado $Y = y$.

$f(x^+)$ Límite por la derecha de $f(x)$.

$f(x^-)$ Límite por la izquierda de $f(x)$.

$\phi_X(\cdot)$ Función característica de la variable aleatoria X.

$\Phi(\cdot)$ Función de distribución normal estándar.

$\phi(\cdot)$ Función de densidad de una normal estándar.

$\mathcal{X}_A(\cdot)$ Función característica del conjunto A.

$\ln(x)$ Logaritmo natural de x.

$m_X(\cdot)$ función generadora de momentos de la variable aleatoria X.

μ_r' r-ésimo momento central alrededor de *cero*.

μ_r r-ésimo momento central alrededor de la media.

$\mathbb{N}$ Conjunto de números naturales $\{0, 1, 2, ...\}$.

$O(x_n)$ O grande de x_n.

$o(x_n)$ o pequeño de x_n.

$P(A)$ Probabilidad del evento A.

$P(A \mid B)$ Probabilidad condicional del evento A dado el evento B.

$\langle \mathbf{x}, \mathbf{y} \rangle$ Producto interno de los vectores $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

$\mathbb{Q}$ Conjunto de números racionales.

$\mathbb{R}$ Conjunto de números reales.

σ_X^2 Varianza de la variable aleatoria X.

σ_X Desviación estándar de la variable aleatoria X.

$\sigma(L)$ Menor σ-álgebra que contiene a la colección L.

$[t]$ Parte entera de t.

$tr(A)$ Traza de la matriz A.

$|X|$ Valor absoluto de la variable aleatoria X.

$(x_n)_{n \in \mathbb{N}}$ Sucesión de números reales.

$X \stackrel{d}{=} Y$ significa que las variables aleatorias X y Y tienen la misma distribución.

$X \stackrel{d}{=} \mathcal{B}(n,p)$ indica que la variable aleatoria X tiene distribución binomial de parámetros n y p.

$X \stackrel{d}{=} Hg(n,R,N)$ indica que la variable aleatoria X tiene distribución hipergeométrica de parámetros n, R y N.

$X \stackrel{d}{=} \mathcal{P}(\lambda)$ indica que la variable aleatoria X tiene distribución Poisson de parámetro λ.

$X \stackrel{d}{=} \mathcal{B}_N(k,p)$ indica que la variable aleatoria X tiene distribución binomial negativa de parámetros k y p.

$X \stackrel{d}{=} \mathcal{G}(p)$ indica que la variable aleatoria X tiene distribución geométrica de parámetro p.

$X \stackrel{d}{=} \mathcal{U}[a,b]$ indica que la variable aleatoria X tiene distribución uniforme sobre el intervalo $[a,b]$.

$X \stackrel{d}{=} \mathcal{N}(\mu,\sigma^2)$ indica que la variable aleatoria X tiene distribución normal de media μ y varianza σ^2.

$X \stackrel{d}{=} \Gamma(r,\lambda)$ indica que la variable aleatoria X tiene distribución gamma de parámetros r y λ.

$X \stackrel{d}{=} Exp(\lambda)$ indica que la variable aleatoria X tiene distribución exponencial de parámetro λ.

$X \stackrel{d}{=} \mathcal{X}^2_{(k)}$ indica que la variable aleatoria X tiene distribución ji-cuadrada con k grados de libertad.

$X \overset{d}{=} \beta\,(a,b)$ indica que la variable aleatoria X tiene distribución beta de parámetros a y b.

$X \overset{d}{=} Weibull\,(\alpha,\beta)$ indica que la variable aleatoria X tiene distribución Weibull de parámetros α y β.

$X \overset{d}{=} t_{(k)}$ indica que la variable aleatoria X tiene distribución t-Student con k grados de libertad.

$X \overset{d}{=} F_n^m$ indica que la variable aleatoria X tiene distribución F con m grados de libertad en el numerador y n grados de libertad en el denominador.

$\mathbf{X} \overset{d}{=} \mathcal{N}\,(\mu, Q)$ indica que el vector aleatorio $\mathbf{X}$ tiene distribución normal multivariada de vector de medias μ y matriz de varianzas y covarianzas Q.

$\|\mathbf{X}\|$ Norma euclídea del vector $\mathbf{X}$.

$X_n \underset{n\to\infty}{\xrightarrow{c.s}} X$ Convergencia casi siempre (o con probabilidad 1)de la sucesión de variables aleatorias $(X_n)_n$ a la variable aleatoria X.

$X_n \underset{n\to\infty}{\xrightarrow{d}} X$ Convergencia en distribución (o en ley) de la sucesión de variables aleatorias $(X_n)_n$ a la variable aleatoria X.

$X_n \underset{n\to\infty}{\xrightarrow{P}} X$ Convergencia en probabilidad (o estocástica) de la sucesión de variables aleatorias $(X_n)_n$ a la variable aleatoria X.

$\mathbb{Z}$ Conjunto de números enteros.

$\mathbb{Z}_+$ Conjunto de números enteros positivos $\{1, 2, 3, ...\}$.

$:=$ Remplaza el símbolo "=" cuando se trata de una definición o asignación.

▲ Fin de un ejemplo, algoritmo, nota o definición.

■ Fin de una demostración.

Bibliografía

[Ash] : R. Ash (1972) *Real Analysis and Probability,* Academic Press.

[Bau] : H. Bauer (1991) *Wahrscheinlichkeitstheorie,* Walter de Gruyter.

[Gor] : L.G. Gorostiza (2001) *La probabilidad del siglo XX.* Miscelánea Matemática. Sociedad Matemática Mexicana. Número 33.

[Gri] : G. Grimmett & D. Stirzaker (2001) *Probability and Random Processes,* Oxford University Press.

[HerA] : A. Hernández & O. Hernández (2003) *Elementos de Probabilidad y Estadística,* Textos 21, Nivel Elemental. Sociedad Matemática Mexicana.

[Her] : F. M. Hernández, (2003) *Cálculo de Probabilidades,* Textos 25, Nivel Elemental. Sociedad Matemática Mexicana.

[Heu] : H. Heuser (1990), *Lehrbuch der Analysis 1, 2,* B. G. Teubner Stuttgart.

[Jac] : J. Jacod & P. Protter (2000) *Probability Essentials.* Springer-Verlag.

[Kah] : J.A.Kahane (1997), *A Century of Interplay between Taylor series, Fourier series and Brownian motion.* London Bulletin of the London Mathematical Society. Vol.29. No. 3.

[Kor] : R. Korn & E. Korn. (2000) *Option Pricing and Portafolio Optimization.* American Mathematical Society. Providence, Rhode Island.

[Kre] : U. Krengel (2000), *Einführung in die Wahrscheinlichkeitstheorie und Statistik.* Vieweg Verlag.

[Mey] : P. L. Meyer 1992) *Probabilidad y Aplicaciones Estadísticas.* Adisson Wesley Iberoamericana.

[Mun] : M. Muñoz & L. Blanco (2002), *Introducción a la teoría avanzada de la probabilidad.* Colección Textos. Unibiblos. Universidad Nacional de Colombia.

[Nel] : E. Nelson (1967), *Dinamical Theories of Brownian Motions,* Pricenton University Press.

[MuJ] : J. M. Muñoz (2002), *Introducción a la teoría de conjuntos.* Facultad de Ciencias. Universidad Nacional de Colombia.

[Osp] : D. Ospina (2001), *Introducción al muestreo.* Facultad de Ciencias. Universidad Nacional de Colombia.

[Rao] : C. R. Rao (1973), *Linear Statistical Inference and Its Applications.* John Wiley and Sons, New York.

[Ros] : S. Ross (1996), *Stochastic Processes,* John Wiley & Sons, inc.

[Ros2] : S. Ross (1998), *A first course in Probability,* Prentice Hall.

[Ros3] : S. Ross (1999), *Simulación,* Prentice Hall.

[Roy] : H.L. Royden (1968) *Real Analysis.* Macmillan, New York.

[Sea] :S. R. Searle (1982) *Matrix Algebra Useful for Statistics.* John Wiley and Sons.

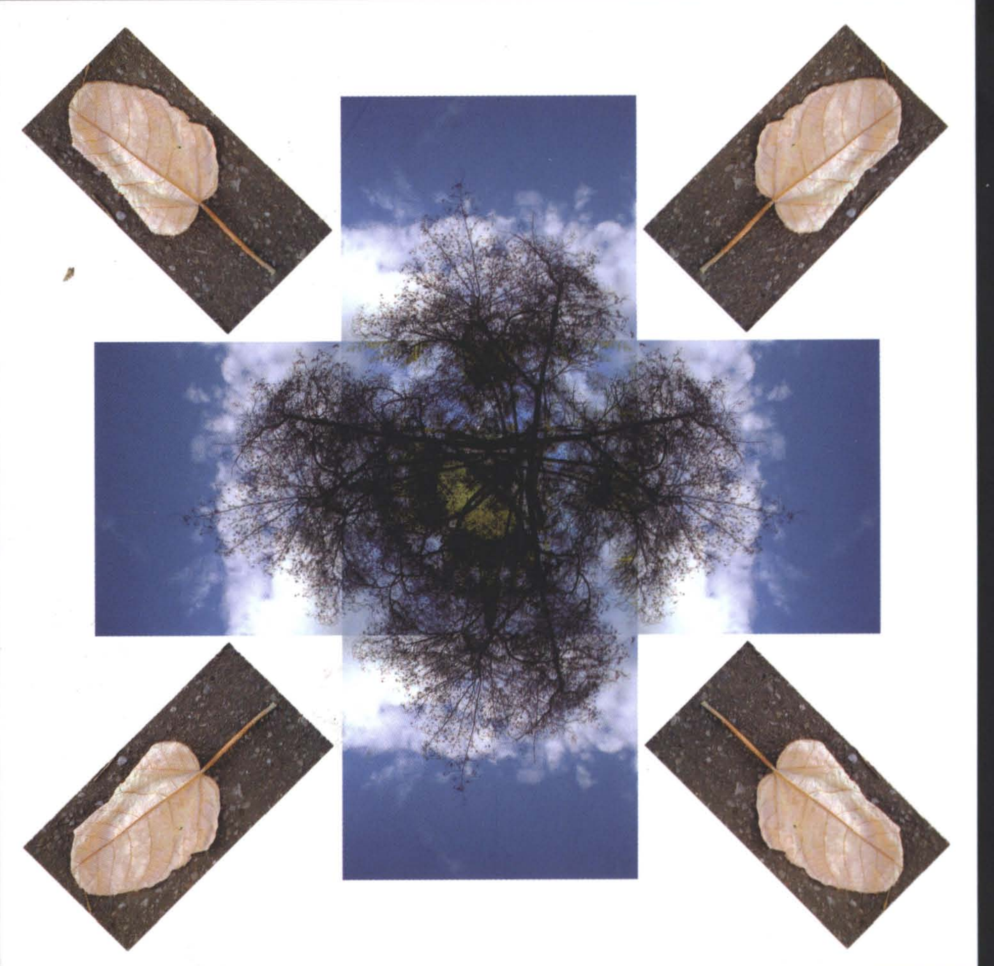
[Sol] : F. Solomon (1987) *Probability and Stochastic Processes.* Prentice-Hall

Índice alfabético

Este libro se terminó de imprimir
en el mes de octubre de 2004
en los talleres gráficos de Unibiblios
en la Universidad Nacional de Colombia
Correo electrónico
dirunibiblo_bog@unal.edu.co
Bogotá, D.C., Colombia

Este texto está diseñado para un primer curso de teoría de probabilidad en las carreras de matemáticas y estadística, así como para los estudiantes de la especialización en estadística que no cuenten con conocimientos previos en el área. La idea de escribir este texto surgió hace ya varios años, como respuesta a la necesidad de contar con un texto en español, que abarcara los principales temas que deben ser vistos por los estudiantes de las carreras antes mencionadas.

Fue así como en el año 1998 publiqué una primera versión del presente texto, en la colección "Notas de clase" de la Facultad de Ciencias de la Universidad Nacional de Colombia. Luego, a partir de la experiencia otorgada por la enseñanza del curso de probabilidad durante varios semestres, en los cuales se detectaron, gracias al aporte crítico de colegas y estudiantes, varias deficiencias en esa primera versión, ésta fue corregida y aumentada hasta dar origen al presente trabajo.



ISBN 958-701-449-9