

Teoría de las Probabilidades

Mercedes Arriojas

Índice general

1. Modelos Probabilísticos. Teoría Combinatoria.	5
1.1. Introducción.	5
1.2. Breve reseña histórica.	6
1.3. Situaciones y Ejemplos que Conducen a Problemas Probabilísticos. . .	6
1.4. Modelo Probabilístico	8
1.5. Teoría de Conjuntos.	17
1.5.1. Definiciones Básicas y Notaciones.	17
1.5.2. Operaciones con Conjuntos.	18
1.6. Espacio de Probabilidad.	24
1.6.1. Espacios de Probabilidad.	25
1.7. Propiedades de la Probabilidad.	25
1.8. Introducción a la Teoría Combinatoria.	31
1.8.1. Muestreo.	33
2. Probabilidad Condicional, Fórmula de Bayes e Independencia.	53
2.1. Probabilidad Condicional.	53
2.1.1. Propiedades básicas.	55
2.1.2. Resultados Básicos sobre Probabilidad Condicional.	58
2.2. La Fórmula de Bayes.	60
2.3. Independencia.	62
3. Variables Aleatorias	69
3.1. Introducción.	69
3.2. Función de Distribución.	74
3.2.1. Propiedades de la Función de Distribución.	74
3.3. Algunas Distribuciones Importantes.	82
3.4. Densidades. Variables Aleatorias Continuas.	92
3.4.1. Distribuciones de Algunas Variables Aleatorias Continuas. . .	95
3.5. Esperanza y Varianza de Variables Aleatorias.	104
3.5.1. Caso Discreto.	104
3.5.2. Caso Continuo.	105
3.5.3. Aplicaciones.	106

3.5.4. Esperanza de una Función de una Variable Aleatoria	107
4. Variables Aleatorias Conjuntas	113
4.1. Distribución Conjunta de Variables Aleatorias.	113
4.2. Esperanza de Funciones de Variables Aleatorias.	118
4.3. Distribución Condicional de Variables Aleatorias.	123
5. Variables Aleatorias Independientes.	129
5.1. Indenpendencia	129
5.1.1. Distribución Conjunta de Variables Aleatorias Independientes.	130
5.1.2. Suma de Variables Aleatorias Independientes.	131
5.2. Funciones de Varias Variables Aleatorias.	135
5.3. Función Generadora de Momentos.	141
5.3.1. Cálculo de Momentos.	145
6. Teoremas Limite	151
6.1. Leyes de Grandes Números.	151
6.1.1. Introducción	151
6.1.2. Desigualdad de Markov	152
6.2. Teorema Central del Límite	158
6.2.1. Convergencia de Variables Aleatorias.	158
6.2.2. Teorema Central del Límite	159
6.2.3. Teorema de De Moivre-Laplace	160
6.3. Ley Fuerte de los Grandes Números.	168
6.3.1. Preliminares.	168
6.3.2. Ley fuerte de los grandes números para variables aleatorias de Bernoulli.	172
Bibliografía	176

Capítulo 1

Modelos Probabilísticos. Teoría Combinatoria.

1.1. Introducción.

Este curso es una introducción a la teoría de las probabilidades. Se tiene como objetivo fundamental propiciar la comprensión de la naturaleza de las situaciones que comunmente pueden ser descritas mediante modelos probabilísticos. Se estudiará la forma en que tales situaciones pueden expresarse en términos matemáticos; es decir, la construcción de un modelo probabilístico. Consideraremos además algunas de las muchísimas aplicaciones de las probabilidades en la descripción y el estudio del mundo que nos rodea.

Uno de los aspectos característicos de la teoría de las probabilidades es el amplio campo de aplicaciones que posee, el cual se extiende a todas las ramas de las ciencias naturales, la tecnología y las ciencias sociales. Esto se debe a que la teoría de las probabilidades permite estudiar y *medir la incertidumbre* que forma parte de casi todo lo que ocurre a nuestro alrededor. De hecho, ya sea por falta de instrumentos de medición adecuados o precisos, por desconocimiento de las leyes que rigen los fenómenos o por no poder manejar todas las variables que intervienen en los mismos, normalmente nuestro conocimiento de la realidad circundante es sólo parcial; siendo la parte que ignoramos muy relevante, cuando no de vital importancia.

Muchos problemas interesantes en probabilidades pueden formularse con herramientas matemáticas muy básicas. Pese a esto y aún cuando en ocasiones dichos problemas resulten pintorescos y hasta graciosos su solución puede no ser obvia y en algunos casos puede resultar bastante difícil de obtener.

En efecto, desde su origen, la teoría de las probabilidades se asoció a problemas de la vida cotidiana, no necesariamente vinculados con aspectos académicos de las matemáticas. Resulta por lo tanto ilustrativo considerar los inicios de esta importante área de conocimiento.

1.2. Breve reseña histórica.

El inicio de la teoría de las probabilidades se produce oficialmente a mediados del siglo XVII y está estrechamente relacionado con los juegos de azar; los cuales, en la sociedad francesa de 1650, eran una actividad generalizada entre las personas adineradas. Surgían constantemente nuevos y cada vez más complicados juegos con dados, cartas, ruletas, etc., en los que se apostaban grandes sumas de dinero. Entre los jugadores, surgían al mismo tiempo preguntas relacionadas con métodos que permitieran calcular el nivel de riesgo que asumían en una determinada apuesta y las posibilidades de ganar. Uno de estos jugadores, De Méré, consultó en París con el conocido matemático y filósofo Blaise Pascal sobre algunas de estas preguntas. Esto originó una famosa correspondencia entre Pascal y algunos de sus amigos matemáticos de la época, entre quienes se encontraba Pierre Fermat. De allí surgieron comentarios y soluciones a muchos de los problemas planteados iniciándose de esa forma lo que hoy se conoce como la teoría de las probabilidades.

Aún cuando las aplicaciones de la teoría de las probabilidades se extendieron con rapidez a diferentes áreas de la ciencia, durante mucho tiempo se consideraban las probabilidades como una disciplina menor, que a falta de mayor información permitía predecir o aproximar lo que no se podía calcular. Esta concepción cambió en el siglo XX cuando Werner Heisenberg (premio Nobel de física, 1958) postuló el "*principio de incertidumbre*" según el cual para partículas muy pequeñas es imposible conocer la velocidad y la posición simultáneamente: mientras más nos acercamos a la posición exacta más difusa se hace la idea de velocidad. Lo mejor que puede hacerse es dar una respuesta probabilística en esta situación. Este principio que en la actualidad es generalmente aceptado, cambió la forma en que se pensaba acerca de las probabilidades transformándola en una disciplina independiente y fundamental en el desarrollo de la ciencia.

1.3. Situaciones y Ejemplos que Conducen a Problemas Probabilísticos.

Antes de formular los conceptos básicos relacionados con la construcción de un modelo probabilístico, consideraremos varias situaciones concretas que nos permitirán apreciar la relevancia de dichos conceptos e ilustrar la variedad de problemas que pueden formularse en el contexto de la teoría de las probabilidades.

Ejemplos 1.1 .

Apuestas y Juegos.

1. En un casino se lanzan dos dados distintos. Un jugador apuesta a un número y gana si ese número es la suma de las caras mostradas por los dados. El horóscopo de un jugador dice que sus números de suerte son 3 y 6, pero sólo puede apostar a un número. ¿Cuál es su mejor opción?
2. *El carro y las cabras (The New York Times, 1991)*. Suponga que usted está en la fase final de un concurso que se transmite por televisión. En este momento crucial el animador del concurso lo coloca ante tres puertas, una de las cuales esconde un lujoso automóvil nuevo. Detrás de cada una de las otras dos puertas se encuentra una linda cabra. Usted tiene que escoger una puerta y se ganara lo que este tras ella. Usted se decide y anuncia su escogencia. Entonces el animador abre una de las puertas que oculta a una cabra y le pregunta si le gustaría cambiar su elección inicial. Asumiendo que usted quiere llevarse el carro, ¿seria un cambio de puerta ventajoso para usted? **Nota:** *Observe que la solución a los problemas anteriores depende de que podamos medir el chance de ganar en cada caso. En el ejemplo 1 se puede ver fácilmente que se tiene un mayor chance de ganar si se apuesta al número 6 (¿por qué?)*
3. *El problema de los puntos*. Suponga que dos personas estan jugando un juego en el que el ganador recibirá una cantidad a de dinero. En cada partida el ganador recibe un punto y gana el que primero acumule n puntos. Si por alguna razón el juego debe ser suspendido antes de que haya un ganador, ¿cómo debería repartirse el premio final entre los concursantes?

Colocación de r Bolas en n Celdas.

1. *Colocación de bolas en celdas*. Se tienen 3 bolas ($r = 3$) etiquetadas **A**, **B** y **C** y tres celdas ($n = 3$) enumeradas **1**, **2** y **3**. Las bolas se colocan al azar en las celdas. ¿Cuál es la probabilidad de que exactamente una celda quede vacia?
2. *Ocurrencia de accidentes*. Ocurren r accidentes (bolas) en 7 dias (celdas) de la semana. ¿Hay alguna posibilidad de que no ocurran accidentes el lunes y el miércoles?
3. *Estudio del efecto genético de la radiación*. Las partículas α son las bolas y los cromosomas representan las celdas.
4. *Fotografía*. Una placa fotográfica está cubierta por granos sensibles a la luz. Interesa saber el número de cuantos de luz (bolas) que golpea un grano (celda).

Otros.

1. *Se hacen dos marcas en una varilla de madera, de manera aleatoria e independiente una de la otra. después de esto la varilla se divide en n pedazos. ¿Que tan factible es que ambas marcas queden en el mismo pedazo?*
2. *Cumpleaños. ¿Cuántas personas deberían estar presentes en un salón para que la probabilidad de que al menos dos celebren su cumpleaños el mismo día sea mayor que $1/2$?*
3. *El Dilema de los Prisioneros.* Tres prisioneros, que llamaremos **A**, **B** y **C**, saben que dos de ellos serán liberados y el tercero permanecerá preso por un largo tiempo. Puesto que no se tiene ningún criterio para decidir quien es liberado y quien no, ninguno de ellos sabe quienes serán los elegidos. El prisionero **A** le pide al guardia el nombre de uno de los prisioneros, distinto de él mismo, a ser liberado. El guardia se niega con el siguiente argumento: *En este momento tú probabilidad de ser liberado es de $2/3$. Sin embargo, si te digo el nombre de uno de los prisioneros que será liberado pasarás a ser uno de dos prisioneros a ser liberados con lo que tu chance de salir libre se reducirá a $1/2$. Como no quiero perjudicarte no te diré nada.* Ante esta extraña situación, cabe preguntarse si el razonamiento del guardia es correcto. ¿De qué manera la información suministrada por el guardia afecta el chance que el preso tiene de ser liberado? ¿Cómo podemos analizar este problema?

1.4. Modelo Probabilístico

Un modelo matemático es una descripción de un fenómeno natural o de un sistema que utiliza conceptos y lenguaje matemáticos. Tal modelo matemático permite simular la evolución del fenómeno o sistema que describe con la finalidad de hacer predicciones acerca de su comportamiento futuro.

Los modelos matemáticos son ampliamente usados en muchas áreas incluyendo biología, química, física, ingeniería y economía.

Se pueden distinguir dos tipos fundamentales de modelos matemáticos *determinísticos* y *aleatorios*.

Un modelo determinístico describe una situación en la que, bajo un conjunto de condiciones dadas, se produce siempre el mismo resultado. Por ejemplo si, bajo ciertas condiciones de presión, se calienta un líquido hasta alcanzar su punto de ebullición, este pasa a estado gaseoso.

Por otra parte, los modelos aleatorios describen situaciones en las que, bajo un conjunto de *condiciones* dadas, se produce un resultado perteneciente a una familia de resultados posibles. En este caso no es posible predecir con exactitud qué resultado se obtendrá. Por ejemplo, si se lanza una moneda con dos caras distintas, aún

cuando conocemos todos los posibles resultados, no es posible predecir cual de ellos se obtendrá en un próximo lanzamiento.

La teoría de las probabilidades tiene como objetivos la formulación, descripción y estudio de modelos de situaciones aleatorias; es decir, modelos aleatorios.

Las situaciones que se estudian en probabilidades se caracterizan por la incertidumbre, esto es la imposibilidad de predecir lo que ocurrirá en el futuro inmediato, aún cuando conozcamos cuales son los resultados posibles. *La probabilidad brinda la posibilidad de "medir", en un sentido que precisaremos más adelante, la incertidumbre inherente a un fenómeno dado.*

Observación 1.1 *Las situaciones planteadas en los ejemplos de la sección anterior, son casos particulares de la siguiente situación más general: **Se considera una acción que puede ser repetida cuantas veces se quiera en las mismas condiciones y de la cual, en cada repetición, se obtiene uno y sólo uno de m resultados posibles. Suponemos además que no se tiene control sobre el resultado a producirse en cada realización de la acción, por lo que no puede determinarse con anticipación qué resultado se obtendrá.***

A continuación presentamos algunos ejemplos adicionales de esta situación general y del conjunto de resultados en cada caso.

Ejemplos 1.2

1. **Acción:** Lanzar un dado con seis caras distintas.

Conjunto de resultados: $\Omega = \{1, 2, 3, 4, 5, 6\}$.

2. **Acción:** Lanzar una moneda con dos caras distintas.

Conjunto de resultados: $\Omega = \{C, S\}$.

3. **Acción:** Seleccionar un número de lotería.

Conjunto de resultados: $\Omega = \{1, 2, \dots, 10000\}$.

4. **Acción:** Seleccionar seis números en el juego de Lotto.

Conjunto de resultados:

$$\Omega = \{\{\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6\} : \omega_i \in \{1, \dots, 50\}, \omega_i \neq \omega_j \text{ } i \neq j\}.$$

En cada uno de los casos anteriores, *aún cuando sabemos cuales son los resultados posibles, no podemos predecir con exactitud lo que ocurrirá una vez que se repita la acción*. Además, si el dado es perfectamente simétrico, la moneda es perfecta y la selección de los números de lotería se hace sin trampas estaremos de acuerdo en que todos los resultados son igualmente plausibles cada vez que se realiza el experimento:

- Al lanzar el dado la probabilidad de obtener un i es igual a $1/6$, $i = 1, \dots, 6$.
- Al lanzar la moneda la probabilidad de obtener una cara es de $1/2$.
- Al seleccionar un número la probabilidad de obtener j es $1/10000$, $j = 1, \dots, 10000$.

NOTA: Más adelante aprenderemos a determinar la probabilidad de una combinación de números en el Lotto.

Observación 1.2 *Las acciones anteriores pueden ser repetidas bajo las mismas condiciones. Son por lo tanto ejemplos de lo que se conoce como **experimento aleatorio** (Ver Observación 1.1.). El conjunto de resultados posibles de un experimento aleatorio se denomina **espacio muestral**. Más adelante exploraremos estos conceptos con mayor detalle.*

Consideremos ahora la solución de algunos de los problemas planteados en ejemplos anteriores.

El dilema de los prisioneros.

Comentario. Uno de los aspectos más básicos de este problema es que, al ser liberados dos prisioneros, se tienen tres resultados posibles: A y B son liberados, A y C son liberados ó B y C son liberados. Denotaremos estos resultados como $\{A, B\}$, $\{A, C\}$ y $\{B, C\}$ respectivamente.

Otro aspecto importante, es que si asumimos que los presos a ser liberados se escogen al azar, es decir, sin que prive ningún criterio de selección, la intuición sugiere que cada uno de los posibles resultados debe tener la misma posibilidad de ocurrir. Por lo tanto, si hemos de asignar un número a la *posibilidad de ocurrencia* de cada resultado y si asumimos que se le asigna el número 1 a un resultado que tiene un chance de ocurrencia de 100 %, este número debería ser el mismo para todos los resultados, en este caso $1/3$ (ó un chance de 33 %).

Consideremos el experimento que consiste en seleccionar al azar un par de presos para ser liberados. Esto equivale a seleccionar al azar uno de los tres pares (no ordenados) $\{A, B\}$, $\{A, C\}$ o $\{B, C\}$. Como cada par tiene probabilidad $1/3$ de ser seleccionado, la probabilidad de que A sea liberado es la probabilidad de que ocurra $\{A, B\}$ o $\{A, C\}$, es decir $1/3 + 1/3 = 2/3$ que es lo que el guardia afirma. Sin

embargo en este contexto no nos es posible determinar si el hecho de que el guardia dé a conocer al prisionero A la identidad de uno de los que será liberado afecta sus posibilidades de salir libre. Para responder a esto debemos considerar un experimento con un espacio muestral (resultados) diferente. Sean

$O_1 = \{\text{quedan libres } A \text{ y } B, \text{ el guardia informa que } B \text{ será liberado}\}$

$O_2 = \{\text{quedan libres } A \text{ y } C, \text{ el guardia informa que } C \text{ será liberado}\}$

$O_3 = \{\text{quedan libres } B \text{ y } C, \text{ el guardia informa que } B \text{ será liberado}\}$

$O_4 = \{\text{quedan libres } B \text{ y } C, \text{ el guardia informa que } C \text{ será liberado}\}$

En esta situación A es liberado si ocurre O_1 u O_2 . Observe que O_1 es equivalente a seleccionar $\{A, B\}$ por lo que O_1 tiene probabilidad $1/3$ y análogamente O_2 es equivalente a seleccionar $\{A, C\}$ que también tiene probabilidad $1/3$. Entonces la probabilidad de liberar a A sigue siendo $1/3 + 1/3 = 2/3$; y concluimos que la información suministrada por el guardia no altera las posibilidades de A de salir libre. Más adelante analizaremos este problema en mayor profundidad.

El carro y las cabras. Asumiremos que el carro se encuentra detrás de la puerta **1**, con esto no se pierde generalidad ya que los otros casos pueden ser tratados de manera similar. Consideremos el experimento que consiste en los siguientes pasos:

Paso 1. El participante elige una puerta.

Paso 2. El conductor del programa descubre una de las cabras y le da al participante la opción de cambiar de puerta.

Paso 3. El participante hace su elección definitiva.

Un resultado de este experimento puede representarse como una terna

(elección 1, cabra, elección 2)

y dependiendo de la elección final del participante se tienen dos situaciones posibles: o bien el concursante cambia su elección inicial o la ratifica. Lo que queremos determinar es **si estas dos situaciones diferentes le dan al participante diferentes posibilidades de ganar el carro**. Dado que el participante no tiene ninguna idea de la ubicación del carro, lo único que podemos afirmar es que las tres puertas tienen la misma posibilidad de ser seleccionadas en la primera elección, esto es $1/3$.

Situación 1: se cambia de puerta. Observe que este paso es parte del experimento, en cuanto a que determina el espacio muestral.

En este caso el espacio muestral asociado al experimento está dado por:

$$\Omega = \{(1, 2, 3), (1, 3, 2), (2, 3, 1), (3, 2, 1)\}.$$

El participante gana el carro en el evento $\{(2, 3, 1), (3, 2, 1)\}$ y como $\{(2, 3, 1)\}$ es el evento de elegir la puerta 2 y $\{(3, 2, 1)\}$ es el evento de elegir la puerta 3, se tiene la posibilidad de ganar el carro es igual a $1/3 + 1/3 = 2/3$.

Situación 2: se ratifica la elección inicial. En este caso el espacio muestral asociado al experimento esta dado por:

$$\Omega = \{(1, 2, 1), (1, 3, 1), (2, 3, 2), (3, 2, 3)\}.$$

El participante gana el carro si ocurre $\{(1, 2, 1), (1, 3, 1)\}$ que es el evento de elegir la puerta 1. Por lo tanto en este caso la posibilidad de ganar es de $1/3$.

Formulación del modelo probabilístico

Dado un fenómeno natural se formula un modelo matemático que lo aproxime, del estudio de este modelo se sacan conclusiones que son aplicadas a la situación real.

1. FENOMENO NATURAL	2. MODELO MATEMÁTICO	3. CONCLUSIONES
------------------------	-------------------------	-----------------

A continuación describiremos los elementos básicos en la formulación de un modelo probabilístico.

1. Un ***experimento aleatorio*** es una acción que puede ser repetida bajo las mismas condiciones tantas veces como se quiera; del cual se conocen todos los resultados posibles sin que se pueda predecir con exactitud el resultado que se obtendrá en la siguiente repetición.
2. El ***espacio muestral asociado a un experimento aleatorio*** es el conjunto de los posibles resultados del experimento. Un espacio muestral se dice ***discreto*** si posee una cantidad *finita o numerablemente infinita* de elementos. En caso contrario se dice que el espacio muestral es *continuo*.
3. Un ***evento asociado a un experimento*** es un *subconjunto del espacio muestral*, esto es un conjunto de resultados.
4. **Medida de Probabilidad. Definición preliminar.** Dado un espacio muestral Ω correspondiente a un experimento aleatorio, una ***probabilidad*** (también denominada ***distribución de probabilidad, medida de probabilidad o función de probabilidad***) es una función que asigna a cada evento A en Ω un número $P(A) \in [0, 1]$ y que satisface ciertas condiciones que serán detalladas más adelante. La probabilidad de un evento se interpreta como una medida de la plausibilidad de que dicho evento ocurra.

En las secciones de ejemplos siguientes se ilustran los conceptos de espacio muestral y evento.

Ejemplos 1.3

1. Se lanzan dos dados, uno rojo y uno verde, simultáneamente y se anota el resultado. El espacio muestral asociado a este experimento es el conjunto:

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

Este espacio puede también ser representado como

$R \setminus V$	1	2	3	4	5	6
1	(1,1)	(1,2)	(1,3)	(1,4)	(1,5)	(1,6)
2	(2,1)	(2,2)	(2,3)	(2,4)	(2,5)	(2,6)
3	(3,1)	(3,2)	(3,3)	(3,4)	(3,5)	(3,6)
4	(4,1)	(4,2)	(4,3)	(4,4)	(4,5)	(4,6)
5	(5,1)	(5,2)	(5,3)	(5,4)	(5,5)	(5,6)
6	(6,1)	(6,2)	(6,3)	(6,4)	(6,5)	(6,6)

2. Se lanza una moneda dos veces y se anota el resultado de cada lanzamiento. El espacio muestral es el conjunto

$$\Omega = \{(c, c), (c, s), (s, c), (s, s)\}.$$

3. Se lanza una moneda hasta obtener una cara. El espacio muestral es el conjunto, infinito numerable

$$\Omega = \{(c), (s, c), (s, s, c), (s, s, s, c), \dots\}.$$

4. En una fábrica se toma uno de los artículos producidos y se prueba para determinar si es defectuoso. En este caso el espacio muestral se puede escribir como $\Omega = \{B, D\}$, donde B significa bueno y D defectuoso. Si en cambio se extraen n artículos y se prueban, el espacio muestral puede ser representado por el conjunto

$$\Omega = \{(\epsilon_1, \dots, \epsilon_n) : \epsilon_i \in \{B, D\}, i = 1, \dots, n\},$$

donde $\epsilon_i = 0$ si el i -ésimo artículo está bueno y $\epsilon_i = 1$ si el i -ésimo artículo está defectuoso.

5. Se escoge un punto al azar lanzando un dardo a un disco de un metro. El espacio muestral es en este caso el conjunto de puntos del plano que se encuentran dentro de la circunferencia de radio 1, es decir

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

Comentario. *El espacio muestral, como abstracción de una situación real, no es único ya que depende de lo que la persona que realiza el experimento considere importante. Por ejemplo, al lanzar dos dados (ejemplo 1.3.1) se podría considerar el caso en que uno de los dados cae fuera de nuestro rango visual y decidir que en este caso su valor no puede ser registrado. Entonces habría que incluir en el espacio muestral elementos de la forma: $(b, 1), \dots$ y $(1, b), \dots$; con lo que se obtiene un espacio muestral diferente asociado al mismo experimento.*

6. Se lanzan dos dados distintos y se apuesta a que la suma es tres; esto es, se apuesta al evento $A = \{(1, 2), (2, 1)\}$. El evento la suma es 5 está dado por $B = \{(1, 4), (2, 3), (3, 2), (4, 1)\}$. El evento el dado rojo cae en 1 y el verde entre 3 y 6 es el conjunto $C = \{(1, 3), (1, 4), (1, 5), (1, 6)\}$.
7. Se lanza una moneda dos veces. El evento **\emph{se obtiene al menos una cara}** es el conjunto $D = \{(c, s), (s, c), (c, c)\}$.

Comentario.

- Al realizar un experimento decimos que un evento A ocurre si el resultado obtenido pertenece a A .
- Los resultados individuales de un experimento aleatorio son llamados también **eventos simples**.
- Cuando se tenga una cantidad finita de resultados posibles; es decir, un espacio muestral finito, cualquier subconjunto del espacio muestral se considerará un evento. Si el espacio muestral es infinito esto no será necesariamente cierto, por razones que serán expuestas más adelante.

Antes de formular la definición precisa de probabilidad, discutiremos algunas nociones intuitivas que han sido utilizadas para tratar de definir esta medida.

Proporción Relativa.

Suponga que se tiene una caja de mangos y se sabe que hay algunos mangos dañados en ella. A efectos de fijar un valor por la caja, resulta más razonable tomar en cuenta la proporción relativa de frutas dañadas que su número absoluto. Esto es, si sabemos que hay 35 mangos dañados, la importancia de esta cantidad ciertamente

dependerá del número total de mangos en la caja: si hay sólo 50 mangos una gran cantidad del total no sirve mientras que si hay 500 mangos 35 malos es una cantidad menos significativa. Esto es de importancia si consideramos el experimento aleatorio consistente en sacar un mango al azar de la caja y el evento

$$A = \text{el mango está dañado.}$$

Dado que el espacio muestral asociado a este experimento es el conjunto de todos los mangos y que cualquier mango puede ser seleccionado por igual, una medida de la plausibilidad de este evento podría estar dada por la proporción de mangos dañados:

$$P(A) = \frac{|A|}{|\Omega|},$$

donde las barras indican el número de elementos del conjunto.

Esta noción puede extenderse a espacios muestrales infinitos. Si consideramos el espacio muestral del ejemplo 1.3.5, y para un subconjunto A de Ω definimos $|A| = \text{área de } A$ podemos entonces definir la probabilidad de A como la razón $P(A) = |A|/|\Omega|$.

Note que, aún cuando los conjuntos considerados en el ejemplo anterior son infinitos, su medida (área) es finita.

Frecuencia Relativa.

Suponga que un experimento con espacio muestral Ω se realiza repetidamente en idénticas condiciones. Para cada evento A se define $n(A)$ como el número de veces que el evento A ocurre en las n primeras repeticiones. Entonces se puede definir

$$P(A) = \lim_{n \rightarrow \infty} \frac{n(A)}{n}.$$

Aquí se define la probabilidad de un evento como el **límite de los promedios de las veces que el evento ocurre**, es decir, **el límite de las frecuencias relativas del evento**. Esta definición tiene varios inconvenientes. En primer lugar no podemos garantizar la existencia del límite. En segundo lugar es posible que sucesiones de frecuencias correspondientes a distintas sucesiones de realizaciones del experimento tengan límites diferentes, con lo cual la definición no tendría sentido.

En particular, si el espacio muestral es finito y tiene m elementos, y todos los resultados son igualmente probables, se tiene que para n grande cada uno de los m resultados ocurrirá aproximadamente n/m veces. Si A es un evento que posee a elementos entonces en n realizaciones del experimento A ocurrirá $n(A) = \frac{an}{m}$ veces aproximadamente. Entonces la frecuencia relativa de A satisface

$$\frac{n(A)}{n} = \frac{a}{m} = \frac{|A|}{|\Omega|} = P(A)$$

si consideramos la proporción relativa como medida del evento A . Es decir, en los casos en que todos los resultados tienen la misma probabilidad la frecuencia relativa coincide con la proporción relativa.

Ejercicios 1.1

1. *Un dado es lanzado y, luego, se lanza una moneda dos veces.*
 - (a) *Describa un espacio muestral S que contenga todos los resultados de este experimento.*
 - (b) *Asuma que todos los resultados en (a) tienen la misma probabilidad. Encuentre la probabilidad de los siguientes eventos:*
 - i) *se obtiene un seis y al menos una cara;*
 - ii) *se obtiene un número par y una cara en el segundo lanzamiento de la moneda;*
 - iii) *se obtiene al menos una cara y un número menor que 5.*
2. *Variación del problema del carro y la cabra. En esta oportunidad hay cuatro puertas: tres cabras y un carro. Usted escoge una puerta al azar y entonces el animador abre una de las puertas que oculta una cabra. Suponga que usted se cambia a una de las otras puertas (escogiendola al azar). ¿Cuál es la probabilidad ahora de ganar el carro? ¿Cuál es la probabilidad de ganar el carro si usted no cambia de puerta?*
3. *Suponga que la alarma de su reloj suena en algún momento entre las 6 y las 7 a.m. Describa un espacio muestral con cada uno de los posibles instantes de tiempo en que la alarma puede sonar. ¿Es el espacio continuo ó discreto?*
4. *En el ejercicio anterior, suponga que la alarma sólo puede sonar en intervalos de 5 minutos: 6 am, 6:05 am, 6:10 am, etc. Describa el espacio muestral con cada posible instante en que la alarma suena. Explique porqué es este espacio discreto.*
5. *Asuma que en el espacio discreto del ejercicio 4 todos los resultados son igualmente probables (distribución uniforme) y que usted siempre sueña lo mismo, si está dormido, entre las 6:18 y las 6:43 am. Suponga que está dormido cuando la alarma lo despierta un día. Encuentre la probabilidad de que la alarma interrumpa su sueño.*
6. *Describa el espacio muestral asociado a cada uno de los experimentos siguientes:*
 - (a) *Se mide el tiempo de vida de un transistor cada hora.*
 - (b) *Se sacan aleatoriamente 3 bolas de una bolsa que contiene 3 bolas rojas, 2*

blancas y 4 negras.

(c) Una caja contiene 3 metras, 1 roja, 1 verde y una azul. Se toma una metra de la caja, se reemplaza la metra en la caja y luego se toma otra metra.

(d) Considere el caso en que el experimento anterior se realiza sacando la segunda metra sin reemplazar la primera.

(e) Se lanza un dado continuamente hasta que aparezca un 6.

7. Una caja tiene 2 bolas blancas, 1 negra y 1 roja. Se saca una bola, se anota el color y se introducen en la caja dos bolas, cada una de uno de los colores **distinto** al extraído. Luego se saca una segunda bola.

(a) Describa el espacio muestral asociado a este experimento.

(b) Describa los eventos A = al menos una bola blanca y B = a lo sumo una bola blanca.

8. Se tienen dos monedas perfectamente simétricas. Se asigna a cara el valor 1 y a sello el valor 2. Se lanzan las monedas simultáneamente y se registra la suma de las caras mostradas.

(a) Describa el espacio muestral asociado a este experimento.

(b) Describa el evento ninguna cara y un número par.

9. Considere el experimento que consiste en colocar 3 bolas etiquetadas como A , B y C en tres celdas enumeradas como 1, 2 y 3.

(a) Halle el espacio muestral asociado a este experimento.

(b) Halle los eventos: V_i = la celda i está vacía $i \in \{1, 2, 3\}$; U_i = la celda i contiene una bola.

1.5. Teoría de Conjuntos.

1.5.1. Definiciones Básicas y Notaciones.

Hemos observado que la probabilidad puede considerarse como una función de conjuntos, haremos por lo tanto un repaso de las nociones básicas de teoría de conjuntos que serán esenciales en el resto del curso.

Un **conjunto** es una colección de objetos determinada por una regla que especifica exactamente que objetos pertenecen a la colección.

Los objetos pertenecientes a un conjunto se denominan **elementos** o **puntos** del conjunto.

Notación.

- Se utilizarán letras mayúsculas, latinas o griegas, para denotar a los conjuntos y minúsculas para denotar a los elementos. En particular denotamos por $\mathbf{R}$ al conjunto de los números reales y por $\mathbf{N}$ al conjunto de los números naturales.
- Dado un conjunto S , $a \in S$ se lee **a pertenece a S** y $a \notin S$ se lee **a no pertenece a S** .

a S .

- El símbolo $\emptyset$ denota al conjunto vacío (conjunto que no tiene elementos).

Si E y F son conjuntos decimos que E **está contenido en** F o que E **es un subconjunto de** F , denotado $E \subseteq F$, si todo elemento de E es también elemento de F . Si $E \subseteq F$ y E y F **no** son iguales, se dice que E **está estrictamente contenido en** F , lo que se denota $E \subset F$. Dos conjuntos E y F son iguales si poseen exactamente los mismos elementos; esto ocurre si y solamente si $E \subseteq F$ y $F \subseteq E$.

Los conjuntos pueden representarse de varias maneras:

- Describiendo verbalmente la regla que define a sus elementos: *el conjunto de los estudiantes de matemática en la UCV*.
- Dando una lista de todos los elementos que pertenecen al conjunto: $A = \{1, 3, 5\} =$ *el conjunto de los impares entre 0 y 6*.
- Escribiendo de manera simbólica la propiedad que caracteriza a sus elementos: Si $\wp(a)$ significa que el elemento a satisface la propiedad $A = \{a \in \mathbf{N} : a \text{ impar}, 1 \leq a \leq 5\}$. En general $A = \{a \in \mathbf{N} : \wp(a)\}$.

Nota. Un conjunto E se dice **numerablemente infinito** si se puede establecer una biyección entre E y el conjunto de los números naturales $\mathbf{N}$. Un conjunto es **numerable** o **contable** si es finito o numerablemente infinito.

1.5.2. Operaciones con Conjuntos.

Sea $\mathbb{E}$ un conjunto (no vacío) de elementos. En lo que sigue el conjunto $\mathbb{E}$ será denominado *espacio* y sus elementos *puntos*. Dado un subconjunto E de $\mathbb{E}$, **complemento de E** en $\mathbb{E}$ se define como

$$E^c = \{x \in \mathbb{E} : x \notin E\}.$$

Dados los conjuntos E y F , se definen:

La **unión** de E y F es el conjunto

$$E \cup F = \{x \in \mathbb{E} : x \in E \text{ ó } x \in F\}.$$

La **intersección** de E y F es el conjunto

$$E \cap F = \{x : x \in E \text{ y } x \in F\}.$$

La **diferencia** de E y F es el conjunto

$$E \setminus F = \{x \in E : x \notin F\} = E \cap F^c.$$

Cuando $F \subseteq E$ se usa la notación $E \setminus F = E - F$. Note que $A^c = \mathbb{E} - A$. Finalmente, la **diferencia simétrica** de E y F se define como

$$E \triangle F = (E \setminus F) \cup (F \setminus E) = (E \cup F) \setminus (E \cap F).$$

Dos conjuntos E y F son **disjuntos** si no tienen elementos en común, es decir su intersección es vacía: $A \cap B = \emptyset$. Por ejemplo si Ω es el espacio muestral correspondiente al lanzamiento de dos dados, son conjuntos disjuntos los eventos: $A = \text{la suma es } 8$ y $B = \text{al menos un uno}$.

Propiedades.

- $A \cup A = A = A \cap A$.

- $(A^c)^c = A$.

- *Leyes Distributivas:*

- i) $(A \cup B) \cap C = (A \cap C) \cup (B \cap C)$.

Demostración: Si $x \in (A \cup B) \cap C$ entonces $x \in A \cup B$ y $x \in C$; por lo tanto $(x \in A \text{ ó } x \in B)$ y $x \in C$, de donde deducimos que $(x \in A \cap C)$ o $x \in B \cap C$, lo que implica que $(A \cup B) \cap C \subseteq (A \cap C) \cup (B \cap C)$. Recíprocamente si $x \in (A \cap C) \cup (B \cap C)$, entonces x ciertamente pertenece a C y además vemos que $x \in A$ o $x \in B$. Lo que demuestra la otra contención.

- ii) $(A \cap B) \cup C = (A \cup C) \cap (B \cup C)$. La demostración es similar y se deja como ejercicio.

- *Leyes de Morgan:*

- i) $(A \cup B)^c = A^c \cap B^c$.

Demostración: Si $x \in (A \cup B)^c$ entonces $x \notin A \cup B$ lo que implica que $x \notin A$ y $x \notin B$ es decir $x \in A^c \cap B^c$. Recíprocamente si $x \in A^c \cap B^c$ entonces como $x \notin A$ y $x \notin B$ se tiene que $x \in (A \cup B)^c$.

- ii) $(A \cap B)^c = A^c \cup B^c$. Ejercicio: Demuestre esta igualdad a partir de la anterior.

- La diferencia de conjuntos **no** es ni asociativa ni conmutativa; es decir, en general $(A \setminus B) \setminus C \neq A \setminus (B \setminus C)$ y $A \setminus B \neq B \setminus A$. Dé un ejemplo.

Operaciones con Familias Arbitrarias de Conjuntos.

Sea I un conjunto arbitrario de elementos llamados *índices*, usualmente se asume $I \subseteq \mathbf{R}$ que puede ser finito, infinito numerable o infinito no numerable, y para cada $\alpha \in I$ sea E_α un conjunto.

Notación. La colección de los E_α se denota $\{E_\alpha\}_{\alpha \in I}$.

Definimos la **unión de una colección de conjuntos**, denotada por

$$\bigcup_{\alpha \in I} E_{\alpha},$$

como el conjunto consistente de los elementos que pertenecen a E_{α} para al menos un $\alpha \in I$.

De manera análoga, la **intersección de una colección de conjuntos** se denota

$$\bigcap_{\alpha \in I} E_{\alpha},$$

y se define como el conjunto conformado por los elementos que pertenecen a E_{α} para todo $\alpha \in I$.

Cuando el conjunto de índices I es finito, podemos decir que $I = \{1, 2, \dots, n\}$ para algún $n \in \mathbf{N}$. En este caso la unión y la intersección se escriben como:

$$\bigcup_{i=1}^n E_i = \{x : x \in E_i \text{ para algún } i \in \{1, \dots, n\}\}$$

y

$$\bigcap_{i=1}^n E_i = \{x : x \in E_i \text{ para todo } i \in \{1, \dots, n\}\}$$

respectivamente. Cuando I es numerable se tiene

$$\bigcup_{i=1}^{\infty} E_i = \{x : x \in E_i \text{ para algún } i \in \mathbf{N}\}$$

y

$$\bigcap_{i=1}^{\infty} E_i = \{x : x \in E_i \text{ para todo } i \in \mathbf{N}\}.$$

Las *leyes distributivas* se generalizan de manera natural para familias arbitrarias de conjuntos:

$$F \cap \left(\bigcup_{\alpha \in I} E_{\alpha} \right) = \bigcup_{\alpha \in I} (F \cap E_{\alpha}),$$

$$F \cup \left(\bigcap_{\alpha \in I} E_{\alpha} \right) = \bigcap_{\alpha \in I} (F \cup E_{\alpha}).$$

Se tienen también las *leyes de Morgan*:

$$\left(\bigcup_{\alpha \in I} E_{\alpha} \right)^c = \bigcap_{\alpha \in I} E_{\alpha}^c,$$

$$\left(\bigcap_{\alpha \in I} E_{\alpha}\right)^c = \bigcup_{\alpha \in I} E_{\alpha}^c.$$

Función Indicatríz.

Dado un conjunto $A \subseteq \mathbb{E}$ la **función indicatríz de A** (también llamada **función indicadora o característica de A**) se define de la siguiente manera:

$$I_A : \mathbb{E} \mapsto \{0, 1\},$$

$$I_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \notin A \end{cases} \quad (1.1)$$

Notación. Dados dos números reales a y b :

- el *mínimo* entre a y b se denota por $a \wedge b$.
- el *máximo* entre a y b se denota por $a \vee b$.

La función indicatríz satisface las siguientes propiedades:

- $I_{A \cap B}(\omega) = I_A(\omega) \wedge I_B(\omega) = I_A(\omega)I_B(\omega)$.
- $I_{A \cup B}(\omega) = I_A(\omega) \vee I_B(\omega)$.
- $I_{\Omega} = I_A(\omega) + I_{A^c}(\omega)$.

Ejercicios 1.2

Nota. En algunos casos se usará la notación $A \cap B = AB$.

1. Demuestre que
 - (a) $E \cap F \subseteq E \subseteq E \cup F$.
 - (b) Si $E \subseteq F$, entonces $F^c \subseteq E^c$.
 - (c) $F = (F \cap E) \cup (F \cap E^c)$.
 - (d) $E \cup F = E \cup (F \setminus E)$.
 - (e) $F = (E \cup F) \cap (E^c \cup F) \cap (E \cup F^c)$.
2. Demuestre las leyes distributivas y las leyes de Morgan para familias infinitas de conjuntos.
3. Demuestre que si dos conjuntos tienen complementos idénticos ellos son iguales.
4. Considere los conjuntos: A =artículos de vestir, B =artículos de bajo costo y C =artículos de buena calidad. Describa, verbalmente, lo que simbolizan los conjuntos: a) $A \cup (B \cap C)$; b) $(A - B) - C$; c) $A - (B - C)$.
5. Demuestre que $(A \cup B) \cap C \neq A \cup (B \cap C)$.

6. Demuestre que $A \subseteq B$ si y sólo si $A \cap B = A$ ó $A \cup B = B$.
7. Demuestre que A y B son disjuntos si y sólo si $A - B = A$ ó $A \cup B = A \Delta B$.
8. Demuestre que $(A \cup B) - (C \cup D) \subseteq (A - C) \cup (B - D)$.
9. Considere la operación de conjuntos denotada como $/$ y definida de la siguiente forma: $A/B = A^c \cup B$. Demuestre que: a) $(A/B) \cap (B/C) \subseteq A/C$.
b) $(A/B) \cap (A/C) = A/(B \cap C)$.
c) $(A/B) \cap (B/A) = (A \Delta B)^c$.
10. . Expresa las operaciones de unión, intersección y complementación de conjuntos en términos de diferencia de conjuntos.
11. Demuestre que:
a) $A \subseteq B$ si y sólo si $I_A \leq I_B$.
b) $A \cap B = \emptyset$ si y sólo si $I_A I_B = 0$.
12. Jugadores A , B y C lanzan una moneda por turnos, primero lanza A , después B y por último C . Gana el primero en obtener una cara. Dé una representación del espacio muestral y defina los siguientes eventos: i) $A = A$ gana; ii) $B = B$ gana; iii) $(A \cup B)^c$.
13. Si E , F y G son conjuntos dados, demuestre que $E - (F \cup G) = (E - F) \cap (E - G)$.
14. Demuestre que $E \cup F = (E \Delta F) \Delta (E \cap F)$ y que $E - F = E \Delta (E \cap F)$.
15. Se lanzan dos dados. Sea E el evento de que la suma sea impar; F el evento de que al menos uno de los dados cae en 1; y G el evento de que la suma es 5. Describa los eventos $E \cap F$, $E \cup F$, $F \cap G$, $E \cap F^c$ y $E \cap F \cap G$.
16. . Se lanzan dos dados distintos. Halle los eventos $E =$ la suma es no menor que 9, $F =$ la suma es no menor a 4, $G =$ al menos un dado cae en 3, $H =$ la suma es par, $E \cap F$, $F \cap H$, $F \setminus G$, $G \setminus H$ y $E \cap G \cap H$.
17. En un juego de cartas sea $k \in \{1, 2, 3, 4\}$. Sea N_k el evento el jugador sentado hacia el norte tiene al menos k ases. Sean S_k , E_k y O_k eventos análogos para sur, este y oeste. Diga que número de ases tiene el jugador sentado hacia el oeste en cada uno de los siguientes eventos:
a) O_1^c ; b) $N_2 S_2$; c) $N_1^c S_1^c E_1^c$; d) $O_2 - O_3$; e) $N_1 S_1 E_1 O_1$; f) $N_3 O_1$; g) $(N_2 \cup S_2) E_2$.
18. En el ejemplo anterior verifique que: a) $S_3 \subseteq S_2$; b) $S_3 O_2 = \emptyset$; c) $N_2 S_1 E_1 O_1 = \emptyset$; d) $N_2 S_2 \subseteq O_1^c$; e) $(N_2 \cup S_2) O_3 = \emptyset$; f) $O_4 = N_1^c S_1^c E_1^c$.

19. Considere el experimento que consiste en colocar 3 bolas etiquetadas como **A**, **B** y **C** en tres celdas enumeradas como **1**, **2** y **3**. Considere los eventos: V_i = la celda i está vacía; U_i = la celda i contiene una bola; D_i = la celda i contiene dos bolas y T_i = la celda i contiene tres bolas para $i \in \{1, 2, 3\}$.
 (a) Demuestre que: (i) $U_1U_2 \subseteq U_3$. (ii) $V_1V_2 = T_3$. (iii) $D_1D_2 = \emptyset$.
 (iv) $T_1 \subseteq T_2$. (v) $U_2D_1 \subseteq V_3 \setminus D_2U_1$.
 (b) ¿Cuántos elementos contiene U_1U_2 ?
20. Sean E , F y G tres eventos. Encuentre expresiones para los eventos siguientes en términos de E , F y G :
 (a) sólo E ocurre; (b) E y G ambos ocurren pero no F ;
 (c) al menos uno de los eventos ocurre; (d) al menos dos de los eventos ocurren;
 (e) los tres eventos ocurren; (f) ninguno de los eventos ocurre;
 (g) a lo sumo uno de ellos ocurre; (h) a lo sumo dos de los eventos ocurre
 (i) exactamente dos de ellos ocurren; (j) a lo sumo tres de ellos ocurren.
21. Dos grupos comparten ciertos miembros. Suponga que el grupo A tiene 123 miembros, que el grupo B tiene 78 y que el número total de individuos en ambos grupos es 184. ¿Cuántos miembros pertenecen a ambos grupos?
22. Los grupos A , B y C tienen respectivamente 57, 49 y 43 miembros. A y B comparten 13, A y C comparten 7 y B y C comparten 4 miembros. Hay sólo un individuo perteneciente a los tres grupos. Halle el número total de personas pertenecientes a la unión de los tres grupos.
23. Una ciudad con una población de 100.000 habitantes tiene tres diarios: I , II y III . La proporción de gente en la ciudad que lee estos diarios es: $I = 10\%$, $II = 30\%$, $III = 5\%$, I y $II = 8\%$, I y $III = 2\%$, II y $III = 4\%$ y I , II y $III = 1\%$.
 a) Encuentre el número de personas que lee sólo un diario.
 b) ¿Cuánta gente lee al menos dos diarios?
 c) Si I y III son diarios matutinos y II es un diario vespertino, ¿Cuántas personas lee al menos un diario matutino y un diario vespertino?
 d) ¿Cuántas personas leen sólo un diario matutino y un diario vespertino?
24. Los siguientes datos fueron suministrados por un estudio de un grupo de 1000 suscriptores a una revista. En lo relativo a sexo, estado civil y educación habían: 312 hombres, 470 personas casadas, 525 graduados universitarios, 42 hombres graduados universitarios, 147 graduados universitarios casados, 87 hombres casados y 25 hombres casados graduados universitarios. Demuestre que los números reportados en el estudio son incorrectos.

1.6. Espacio de Probabilidad.

En la práctica, cuando se realiza un experimento nos interesará saber si ciertos subconjuntos del espacio muestral Ω , conjunto de resultados, ocurren o no. A estos subconjuntos *de interés* los llamamos **eventos o sucesos**.

Por ejemplo, si consideramos la situación descrita en el ejemplo 1.3.4, podemos estar interesados en el subconjunto:

$$d \text{ objetos defectuosos} = \{(a_1, \dots, a_n) : a_i \in \{0, 1\}; i = 1, \dots, n\}$$

para algún d natural.

Recordemos que un *evento* $A \subseteq \Omega$ *ocurre*, al realizar un experimento, si el resultado ω obtenido es tal que $\omega \in A$. Denotaremos a la familia de subconjuntos de Ω formada por los eventos, como $\mathcal{A}$. Como se mencionó antes, en general **no** todos los subconjuntos de un espacio muestral son eventos. Aún cuando la familia $\mathcal{A}$ de *subconjuntos de interés* depende de la situación estudiada, esta siempre satisface las siguientes propiedades:

e1. $\Omega \in \mathcal{A}$, es decir que la realización del experimento produce un resultado. A Ω lo llamaremos *evento cierto o seguro*, ya que este evento siempre ocurre.

e2. $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$, es decir, si A es un evento entonces A *no ocurre* también es un evento.

e3. Si $A_n \in \mathcal{A}$, para $n = 1, 2, \dots$ entonces $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$, es decir que la unión numerable de eventos es un evento. Note que al realizar un experimento, la unión de eventos ocurre si el resultado pertenece a alguno de los eventos que integran la unión.

Definición 1.1 Una familia $\mathcal{A}$ de subconjuntos de Ω que satisface las propiedades e1, e2 y e3 se denomina **σ -álgebra** de subconjuntos de Ω .

Algunas de las propiedades que se deducen inmediatamente de la definición anterior son

ei. El conjunto vacío, $\emptyset$, es un evento ya que $\emptyset = \Omega^c$.

eii. La unión finita de eventos es un evento: dados $A_1, \dots, A_n$ en $\mathcal{A}$, usando ei basta considerar $A_{n+1} = A_{n+2} = \dots = \emptyset$, luego se puede aplicar e3.

eiii. Si $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$, entonces $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$; es decir, la intersección numerable de eventos es un evento. Esto se deduce fácilmente usando las leyes de Morgan, ya que

$$\bigcap_{n=1}^{\infty} A_n = (\bigcup_{n=1}^{\infty} A_n^c)^c.$$

El resultado se obtiene al aplicar e2 y e3.

1.6.1. Espacios de Probabilidad.

Definición 1.2 Sea Ω un espacio muestral y $\mathcal{A}$ una familia de eventos de Ω . Una **medida de probabilidad o probabilidad** sobre Ω , es una función P definida sobre $\mathcal{A}$ que satisface los siguientes axiomas:

p1. P es no negativa, es decir para todo evento $A \in \mathcal{A}$ se cumple que $P(A) \geq 0$.

p2. P es **contablemente aditiva o σ -aditiva**; es decir, si $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$ son disjuntos dos a dos, lo que significa que $A_i \cap A_j = \emptyset$ para cada $i \neq j$, entonces $P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$.

p3. El evento cierto tiene probabilidad 1: $P(\Omega) = 1$.

La terna $(\Omega, \mathcal{A}, P)$ se denomina un **espacio de probabilidad**, el valor $P(A)$ se denomina **probabilidad de A** .

Así como el espacio muestral Ω se define de acuerdo con la situación que se desea modelar, la asignación de probabilidades a los eventos considerados sobre Ω , esto es, la definición de una medida de probabilidad para un problema dado, también depende de la situación particular estudiada.

Por ejemplo, si consideramos como experimento el lanzar una moneda, la probabilidad de cara o sello dependerá de si se tiene una moneda simétrica o no.

1.7. Propiedades de la Probabilidad.

Las siguientes propiedades son consecuencia directa de la definición de probabilidad.

pi. $P(\emptyset) = 0$.

Demostración: Para demostrar esto consideremos la sucesión $\{A_n\}_{n \geq 1} \subseteq \mathcal{A}$ donde $A_1 = \Omega$ y $A_i = \emptyset, i = 2, 3, \dots$. Estos conjuntos son claramente disjuntos dos a dos: para $i \neq j$ se cumple $A_i \cap A_j = \emptyset$. Por lo tanto, aplicando *p2.* se tiene que:

$$P(\Omega) = P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) = P(\Omega) + \sum_{i=2}^{\infty} P(A_i),$$

lo que implica que $\sum_{i=2}^{\infty} P(A_i) = 0$. Como para todo i $P(A_i) \geq 0$, se tiene entonces que $P(A_i) = 0$ para todo $i \geq 2$; es decir, $P(\emptyset) = 0$.

pii. P es **finitamente aditiva**; esto es, si $\{A_i\}_{i \in \{1, \dots, k\}} \subseteq \mathcal{A}$ son tales que $A_i \cap A_j = \emptyset$,

para $i \neq j$, entonces

$$P(\cup_{i=1}^k A_i) = \sum_{i=1}^k P(A_i).$$

Demostración: Para demostrar esto se considera la sucesión $\{A_n\}_{n \geq 1}$ donde $A_n = \emptyset$ para $n \geq k+1$ y se aplican las propiedades $p2$ y pi .

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

Demostración: Dado que $\Omega = A \cup A^c$ y que $A \cap A^c = \emptyset$, se tiene que

$$P(\Omega) = P(A) + P(A^c) = 1 \implies P(A^c) = 1 - P(A).$$

piv. Si $A \subseteq B$ entonces $P(A) \leq P(B)$.

Demostración: Para demostrar esto podemos escribir B como unión de conjuntos disjuntos: $B = A \cup (B - A) = A \cup (B \cap A^c)$ y así tenemos que

$$P(B) = P(A) + P(B - A) \geq P(A),$$

ya que $P(B - A) \geq 0$.

Nota. Observe que si $A \subseteq B$ entonces $P(B - A) = P(B) - P(A)$.

pvi. Para todo $A \in \mathcal{A}$ se cumple que $P(A) \leq 1$.

Demostración: Como $A \subseteq \Omega$ esta propiedad es una consecuencia inmediata de la anterior.

Observe que **toda medida de probabilidad toma valores entre 0 y 1**.

pvi. Dados eventos A y B , no necesariamente disjuntos, se cumple que

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

Demostración: Observemos que

$$A \cup B = A \cup (B \setminus A) \text{ y } B = (B \cap A) \cup (B \cap A^c),$$

aplicando la propiedad de aditividad finita se obtiene

$$P(A \cup B) = P(A) + P(B \setminus A) \text{ y } P(B) = P(B \cap A) + P(B \setminus A),$$

de donde

$$P(A \cup B) - P(A) = P(B \setminus A) = P(B) - P(B \cap A)$$

lo que implica que

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

pvi. **Continuidad desde Abajo.** Si $\{A_n\}_{n \geq 1}$ es una sucesión creciente de eventos, es decir que satisface

$$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots,$$

entonces,

$$P(\cup_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n).$$

Para demostrar esta propiedad definamos la sucesión $\{B_n\}_{n \geq 1}$ como:

$$B_1 = A_1, B_2 = A_2 - A_1, \dots, B_n = A_n - A_{n-1}, \dots$$

Los conjuntos de esta nueva sucesión son dos a dos disjuntos: sean $i \neq j$, podemos suponer sin pérdida de generalidad que $i > j$, entonces, observando que

$$A_{j-1} \subseteq A_j \subseteq A_{i-1},$$

se tiene que

$$A_j - A_{j-1} \subseteq A_{i-1}.$$

Por lo tanto

$$A_{i-1}^c \cap (A_j - A_{j-1}) = \emptyset.$$

y

$$B_i \cap B_j = (A_i - A_{i-1}) \cap (A_j - A_{j-1}) = (A_i \cap A_{i-1}^c) \cap (A_j \cap A_{j-1}^c) = A_i \cap \emptyset = \emptyset;$$

Por otra parte es claro que cada B_i es un evento y que

$$\cup_{n=1}^{\infty} A_n = \cup_{n=1}^{\infty} B_n.$$

Verifiquelo.

Utilizando la ecuación obtenida en *piv*, se tiene que, para todo $i \geq 2$,

$$P(B_i) = P(A_i - A_{i-1}) = P(A_i) - P(A_{i-1}).$$

Entonces

$$\begin{aligned} P(\cup_{n=1}^{\infty} A_n) &= P(\cup_{n=1}^{\infty} B_n) = \sum_{i=1}^{\infty} P(B_i) = \sum_{i=1}^{\infty} (P(A_i) - P(A_{i-1})) \\ &= \lim_{n \rightarrow \infty} \{P(A_1) + \sum_{i=2}^n (P(A_i) - P(A_{i-1}))\} = \lim_{n \rightarrow \infty} P(A_n). \end{aligned}$$

pviii. Continuidad desde Arriba. Si $\{A_n\}_{n \geq 1}$ es una sucesión decreciente de eventos, es decir que satisface

$$A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots,$$

entonces,

$$P(\cap_{n=1}^{\infty} A_n) = \lim_{n \rightarrow \infty} P(A_n).$$

En este caso la sucesión $\{A_n^c\}$ es creciente; aplicando *pvi* se obtiene

$$\begin{aligned} P(\cap_{n=1}^{\infty} A_n) &= P((\cup_{n=1}^{\infty} A_n^c)^c) = 1 - P(\cup_{n=1}^{\infty} A_n^c) \\ &= 1 - \lim_{n \rightarrow \infty} P(A_n^c) = 1 - \lim_{n \rightarrow \infty} (1 - P(A_n)) = \lim_{n \rightarrow \infty} P(A_n). \end{aligned}$$

pix. Dados los eventos A , B y C se cumple que

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

Se aplica *pvi* reiteradamente.

a los eventos $A \cup B$ y C , obtenemos

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

Aplicando nuevamente *pvi* a las uniones entre paréntesis, se tiene

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

px. Sea $\{A_i\}_{i \in \{1, \dots, k\}} \subseteq \mathcal{A}$ una familia de eventos no necesariamente disjunta. Entonces

$$P(\cup_{i=1}^k A_i) = \sum_{i=1}^k P(A_i) - \sum_{i < j=2}^k P(A_i \cap A_j) + \dots + (-1)^{k+1} P(A_1 \cap \dots \cap A_k).$$

Esta propiedad se puede demostrarse por inducción sobre k , haciendo uso de *pvi*.

Ejemplos 1.4

1. Lanzamiento de dos dados distintos. *El espacio muestral correspondiente a este experimento fué descrito en el ejemplo 1.3.6. Suponiendo que los dados son perfectamente simétricos, tenemos que todos los resultados son igualmente probables y podemos considerar $\mathcal{A} = \wp(\Omega)$. Así que definiendo para un evento $A \subseteq \Omega$:*

$$P(A) = \frac{|A|}{|\Omega|}$$

se obtiene una medida de probabilidad. Verifíquelo.

2. Si se tiene un espacio muestral finito $\Omega = \{\omega_1, \dots, \omega_n\}$, en que todos los resultados son igualmente probables la función del ejemplo anterior define una medida de probabilidad en este espacio.

3. Se tiene un experimento en el que ocurren dos posibles resultados: 0 ó 1. Entonces $\Omega = \{0, 1\}$. Definimos $P(0) = \frac{2}{3}$ y $P(1) = \frac{1}{3}$. En general tomando $0 < p < 1$, se obtiene una medida de probabilidad definiendo $P(0) = p$ y $P(1) = 1 - p$.

Nota: La familia de eventos en este caso es $\mathcal{A} = \{\Omega, \emptyset, \{0\}, \{1\}\}$.

4. Sean $\Omega = \{\omega_1, \dots, \omega_n\}$, $\mathcal{A} = \wp(\Omega)$ y $p_i \geq 0$, $i = 1, \dots, n$ tales que $\sum_{i=1}^n p_i = 1$. Definimos $P(\omega_i) = p_i$, $i = 1, \dots, n$. P es una probabilidad sobre $\mathcal{A}$.

5. Sea $\Omega = \{\omega_0, \dots\}$ un conjunto infinito numerable, sea $\lambda > 0$ constante

$$P(\omega_i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, \dots$$

define una medida de probabilidad. Verifíquelo. Esta función se denomina función de probabilidad de Poisson.

6. $\Omega = \{\omega_1, \dots\}$ un conjunto infinito numerable y sea $0 < p < 1$. Defina

$$P(\omega_i) = (1 - p)^{i-1} p, \quad i = 1, 2, \dots$$

Se tiene que

$$P(\Omega) = \sum_{i=1}^{\infty} P(\omega_i) = p \sum_{i=1}^{\infty} (1 - p)^{i-1} = 1,$$

por lo que P es una probabilidad.

Ejercicios 1.3

Nota. En algunos casos se usará la notación $A \cap B = AB$.

1. Se tiene un cesto de manzanas. Supongamos que cada manzana cuesta una moneda, mientras que las manzanas podridas son gratis. Denotemos por Ω el conjunto de todas las manzanas, por R el conjunto de las manzanas podridas y sea S un conjunto cualquiera sacado del cesto. Definamos

$$Q(S) = \frac{|S \setminus R|}{|\Omega - R|}, \quad |A| = \# \text{ elementos en } A.$$

Entonces Q es el valor relativo de S con respecto al valor del cesto. Demuestre que Q es una medida de probabilidad en Ω .

2. Supongamos que las tierras de una hacienda H de forma cuadrada se encuentran divididas en tres franjas de igual superficie A , B y C ; y que los valores de las respectivas unidades de terreno están en la relación 1:3:2. El valor relativo de una parcela medible S de esa hacienda con respecto al valor total del terreno es:

$$V(S) = \frac{P(SA) + 3P(SB) + 2P(SC)}{2}, \quad \text{donde } P(SA) = \frac{\text{area}(S \cap A)}{\text{area}(H)}.$$

Demuestre que V es una medida de probabilidad.

3. Demuestre que si P y Q son dos medidas de probabilidad definidas sobre un mismo espacio muestral numerable, entonces $aP + bQ$ es también una medida probabilística siempre que a y b sean no negativos y se verifique $a + b = 1$. Dé un ejemplo de esto.
4. Demuestre que para eventos A , B y C

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

5. Sea P una probabilidad. Demuestre que si $P(A) = 0,9$ y $P(B) = 0,8$ entonces $P(A \cap B) \geq 0,7$. En general demuestre que para dos eventos A y B cualesquiera se tiene que $P(A \cap B) \geq P(A) + P(B) - 1$.
6. Demuestre que:
- (a) La probabilidad de que exactamente uno de los eventos E o F ocurra es igual a: $P(E) + P(F) - 2P(E \cap F)$.
 - (b) La probabilidad de que ocurra E pero no F es igual a: $P(E \setminus F) = P(E) - P(E \cap F)$.
 - (c) La probabilidad de que no ocurran ni E ni F es igual a: $P(E^c \setminus F) = 1 - P(E) - P(F) + P(E \cap F)$.
7. ¿Cuál es la probabilidad de que la suma de dos dados distintos y perfectos sea 6 ó menos de 6?
8. Se lanzan dos dados perfectos. Halle la probabilidad de cada uno de los eventos señalados en el ejercicio 1.2.14.
9. Sea $\Omega = \{\omega_1, \omega_2, \dots\}$ un espacio muestral discreto. En cada uno de los casos siguientes, diga si $(\Omega, \mathcal{P}(\Omega), P)$ es un espacio de probabilidad. En caso negativo, diga si se puede modificar P para que la respuesta sea afirmativa. Diga como se define $P(A)$ para $A \in \mathcal{P}(\Omega)$. Justifique sus respuestas.
- (a) $P(\{\omega_n\}) = \frac{1}{n(n+1)}$.
 - (b) $P(\{\omega_n\}) = \frac{2}{3^{n-1}}$.
 - (c) $P(\{\omega_n\}) = \frac{2}{(2n-1)(2n+1)}$.

10. Sea $\Omega = \{\omega_1, \omega_2, \dots, \omega_m\}$ un espacio muestral. Sea $p \in (0, 1)$ y $q = 1 - p$. Para una constante c y $1 \leq n \leq m$ defina $P(\{\omega_n\}) = cq^{n-1}p$. Determine el valor de c para que P sea una medida de probabilidad.

1.8. Introducción a la Teoría Combinatoria.

Como se ha visto para los espacios muestrales finitos, en los que es razonable asumir que todos los resultados son igualmente probables, el cálculo de las probabilidades requiere, en muchos casos, el determinar el número de elementos del espacio muestral y del número de elementos en los eventos considerados. Es por esto importante desarrollar técnicas para contar los elementos de un conjunto dado en los casos en que, debido al número de posibles resultados y/o a la naturaleza del experimento, no sea posible, ó práctico, elaborar una lista con todos los elementos del espacio muestral. La teoría matemática que proporciona las herramientas que permiten contar los elementos de estos conjuntos, se conoce como Analisis Combinatorio.

Ejemplos 1.5

1. Se lanzan seis dados distintos. a) ¿De cuántas maneras distintas pueden salir sus caras? b) ¿De cuántas maneras puede suceder que todos muestren caras distintas?
2. En una pequeña comunidad hay 10 hombres cada uno de los cuales tiene 3 hijos. Se elegirá un hombre y a uno de sus hijos como el padre y el hijo del año, ¿cuántas parejas de padre-hijo son posibles?
3. Se forma una comisión de estudiantes de ciencias que se distribuyen por carreras de la siguiente manera 3 de biología, 4 de computación, 5 de física y 2 de matemática. Se quiere formar una subcomisión de cuatro estudiantes, consistente de 1 estudiante de cada carrera. ¿Cuántas subcomisiones diferentes son posibles?
4. ¿Cuántas funciones $f : \{x_1, x_2, \dots, x_n\} \mapsto \{0, 1\}$ pueden ser definidas ?
5. ¿Cuántas placas diferentes de 7 caracteres se pueden formar si los primeros 3 caracteres son letras y los últimos 4 son números?
6. ¿Cuántas iniciales diferentes pueden formarse con dos letras del alfabeto si las letras **no** pueden ser repetidas?
7. En un juego del Loto se gana con seis aciertos escogidos de un total de 54 números. ¿Cuántas posibles maneras hay de seleccionar una lista de números ganadores?

8. Un jurado internacional consta de 2 chilenos, 1 peruanos, 3 venezolanos y 2 colombianos. Se elige una comisión conformada por un representante de cada país. ¿Cuántas posibles comisiones hay?

A continuación enunciamos una de las reglas fundamentales de conteo.

Principio Básico de Conteo.

Supongamos que se realiza una selección múltiple de r componentes tales que: la primera componente se selecciona de un conjunto de m_1 elementos; la segunda se selecciona de un conjunto de m_2 elementos y así sucesivamente. Entonces el número total de posibles selecciones (ó resultados) es igual a

$$m_1 \times m_2 \times \dots \times m_r.$$

Nota. Observe que la situación anterior puede también presentarse como la realización de r experimentos con $m_1, m_2, \dots, m_r$ resultados posibles, respectivamente.

Ejemplos 1.6

1. Seis dados. Al lanzar seis dados distintos se tienen

a) $|\Omega| = 6 \times 6 \times 6 \times 6 \times 6 \times 6 = 6^6$ resultados posibles.

b) El número de resultados en que todos los dados muestran caras distintas es igual a $6 \times 5 \times 4 \times 3 \times 2 = 6!$.

NOTA. Si A es el evento todos los dados presentan caras distintas, entonces, suponiendo que todos los dados son perfectamente simétricos y que por lo tanto todos los resultados son igualmente probables, se tiene que

$$P(A) = \frac{|A|}{|\Omega|} = \frac{6!}{6^6}$$

2. Padre-hijo del año. Dado que el padre puede escogerse de 10 posibles maneras y para cada elección del padre hay 3 posibles elecciones de hijo se tiene un total de $10 \times 3 = 30$ parejas posibles.

3. Subcomisiones de Estudiantes. Considerando nuevamente la selección de la subcomisión como un resultado de la realización de 4 experimentos (ó selecciones) se tiene que hay $3 \times 4 \times 5 \times 2 = 120$ posibles subcomisiones.

Nota. Observe que, en los dos ejemplos anteriores, los individuos en cada grupo tienen oposiciones o roles diferentes.

4. Funciones. *Puesto que $f(x_i)$ tiene que ser 0 ó 1 para cada $i = 1, \dots, n$ se tiene que hay 2^n funciones posibles.*

5. Placas de auto. *Asumiendo que las letras y los números pueden repetirse en una misma placa, tendremos resultados de la forma $(l_1, l_2, l_3, n_1, n_2, n_3, n_4)$ y*

$$|\Omega| = 27 \times 27 \times 10 \times 10 \times 10 \times 10 = 7,290,000.$$

6. Iniciales. *¿Cuántas iniciales diferentes pueden formarse con dos letras del alfabeto si las letras no pueden ser repetidas?*

Hay $27 \times 26 = 702$ iniciales posibles.

7. Comisión internacional. *Hay $2 \times 1 \times 2 \times 2 = 8$ comisiones posibles.*

1.8.1. Muestreo.

Una **muestra ó muestra aleatoria** es una porción de determinada población que se selecciona al azar y se analiza para sacar conclusiones acerca de toda la población. Esto es particularmente útil cuando, debido a la gran cantidad de individuos en la población, resulta poco práctico o imposible hacer el análisis considerando a todos los miembros de la misma, como ocurre por ejemplo al realizar una encuesta de opinión a la población o sondear un lote de la producción de una fábrica para controlar la calidad.

El proceso de selección de una muestra se denomina **muestreo**. Existen varias maneras de seleccionar una muestra aleatoria; es decir, existen varios tipos de muestreo, que permiten considerar situaciones diferentes. Presentaremos los dos enfoques principales que nos permitirán describir los diferentes tipos de muestreo que consideraremos más adelante: **extracción de bolas de una urna o caja y colocación de bolas en celdas**. Además, introduciremos algunos de los elementos y notaciones del análisis combinatorio que, conjuntamente con el principio básico de conteo constituirán las herramientas fundamentales para la determinación del número de elementos de un evento.

Extracción de bolas de una urna

En este enfoque se considera la siguiente situación: Una urna contiene m bolas enumeradas, de 1 a m , todas diferentes mientras no se especifique lo contrario. De la urna se extraen n bolas. Esto puede hacerse de diferentes maneras, dependiendo de si cada bola extraída se repone o no en la urna (muestreo con o sin reposición) y de si la extracción se realiza o no en cierto orden (muestreo ordenado o sin orden).

Colocación de bolas en celdas

Se tienen m cajas o celdas numeradas y n bolas (numeradas o no) que se colocan aleatoriamente en las cajas.

EXTRACCION	COLOCACIÓN
Se tienen m bolas numeradas en una urna de la cual se extraen n	Se tienen m cajas numeradas y n bolas (numeradas ó no) que se colocan aleatoriamente en las cajas.

Tipos de Muestreo.

1) **Muestreo Ordenado con Reposición ó Variaciones con Repetición.**
 Se extraen *sucesivamente* n bolas, devolviendo a la urna cada bola antes de la siguiente extracción. Se anotan *en orden* los números de las bolas extraídas: $(a_1, a_2, \dots, a_n)$ donde $a_j \in \{1, \dots, m\}$ para $j = 1, \dots, n$. Utilizando el principio básico de conteo se obtienen $m \times m \times \dots \times m = m^n$ posibles resultados, ya que cada bola extraída podrá seleccionarse de m maneras posibles.

NOTA: Observe que las componentes del vector de resultados pueden repetirse.

Ejemplos 1.7

1. Se lanzan seis dados distintos. *Los resultados se pueden expresar como sextuplas $(c_1, c_2, c_3, c_4, c_5, c_6)$ donde c_i representa el número obtenido en el i -ésimo dado. Como se observó antes $|\Omega| = 6^6$.*
2. Placas de auto. *Asumiendo que las letras y los números pueden repetirse en una misma placa, tendremos resultados de la forma $(l_1, l_2, n_1, n_2, n_3, l_3, l_4)$ y*

$$|\Omega| = 27 \times 27 \times 10 \times 10 \times 10 \times 27 \times 27 = 531,441,000.$$

3. Iniciales. *¿Cuántas iniciales diferentes pueden formarse con dos ó tres letras del alfabeto si las letras pueden ser repetidas?*
 - $(27 \times 27) + (27 \times 27 \times 27) = 20,412$.

I) Muestreo Ordenado con Reposición

EXTRACCION	COLOCACIÓN
▪ (a) <i>Muestreo con Reposición</i> ▪ (b) <i>Ordenado</i>	▪ (a') <i>Cada caja puede contener más de una bola</i> ▪ (b') <i>Las bolas estan numeradas.</i>
Resultado $(a_1, \dots, a_n)$. ▪ En la i-ésima extracción se obtiene el número $k = a_i$ ▪ El número k puede repetirse.	▪ En la k-ésima caja se coloca la bola número i. ▪ La k-ésima caja puede recibir más de una bola.

II) *Muestreo Ordenado sin Reposición ó Variaciones.*

a) $n < m$. Se extraen de la urna n bolas en forma sucesiva, sin reponer en la urna las bolas extraídas. Se anotan los números de las bolas extraídas, en el *orden* en que se extraen, por lo que un resultado puede representarse como: $(a_1, a_2, \dots, a_n)$ se pueden elegir de

$$m \times (m - 1) \times (m - 2) \times \dots \times (m - (n - 1)) = \frac{m!}{(m - n)!} = V_n^m$$

maneras. La notación V_n^m se lee *variaciones de m elementos en grupos de n*.

NOTA: Observe que las componentes del vector de resultados son todas diferentes.

b) $n = m$: **Permutaciones.** Cuando $m = n$ se están considerando las posibles maneras de ordenar (permutar, jerarquizar) m elementos: $\# \text{ permutaciones de } m \text{ elementos} = m! = V_m^m$.

II) Muestreo Ordenado sin Reposición

EXTRACCION	COLOCACIÓN
▪ <i>(a)</i> Muestreo sin Reposición ▪ <i>(b)</i> Ordenado	$n \leq m$ ▪ <i>(a')</i> Ninguna caja puede contener más de una bola ▪ <i>(b')</i> Las bolas estan numeradas.
(I) $n < m$ Resultado $(a_1, \dots, a_n)$. ▪ En la i-ésima extracción se obtiene el número $k = a_i$ ▪ El número $k \in \{1, \dots, m\}$ no puede repetirse.	(I) $n < m$. ▪ En la k-ésima caja se coloca la bola número i ▪ En la k-ésima caja se coloca a lo sumo una bola.
(II) $n = m$ Permutaciones Resultado $(a_1, \dots, a_n)$ ▪ Cada número $k \in \{1, \dots, m\}$ aparece en el resultado exactamente una vez.	(II) $n = m$ ▪ Cada caja recibe exactamente una ficha numerada

Ejemplos 1.8

1. Placas de auto. Asumiendo que las letras y los números no pueden repetirse en una misma placa, tendremos resultados de la forma $(l_1, l_2, n_1, n_2, n_3, l_3, l_4)$ y

$$|\Omega| = 27 \times 26 \times 10 \times 9 \times 8 \times 25 \times 24 = 303,264,000.$$

2. Iniciales. ¿Cuántas iniciales diferentes pueden formarse con dos ó tres letras del alfabeto si las letras **no** pueden ser repetidas?

- $(27 \times 26) + (27 \times 26 \times 25) = 17,550$.

3. Una junta directiva formada por un presidente, un tesorero y un secretario se elige de un grupo de 10 personas. ¿Cuántas juntas diferentes son posibles?

- Hay $V_3^{10} = \frac{10!}{7!}$ juntas posibles.

4. José, Jaime, Jacobo y Jesús forman una banda musical con 4 instrumentos. Si cada uno de ellos puede tocar los cuatro instrumentos:

a) ¿De cuántas maneras puede organizarse la banda?

b) ¿De cuántas si José y Jaime tocan los cuatro instrumentos y Jacobo y Jesús sólo dos de ellos?

- a) $4 \times 3 \times 2 \times 1 = 4!$

- b) $(2 \times 1) \times (2 \times 1) = 4$.

5. ¿Cuántos ordenes de bateo son posibles en un equipo de beisbol que consiste de 9 jugadores?

- Hay $9! = 362,880$ posibles ordenes de bateo.

6. Un estudiante tiene 10 libros de los cuales 4 son de matemáticas, 3 de química, 2 de historia y 1 de literatura. Si el estudiante quiere arreglar los libros en un estante de forma tal que los de una misma materia queden juntos ¿cuántos arreglos posibles hay?

- Se tienen $4! \times 3! \times 2! \times 1!$ maneras de arreglar los libros de forma tal que los de matemática se encuentren primero, luego los de química, después los de historia y por último el de literatura. También las asignaturas se pueden ordenar de $4!$ maneras y para cada ordenamiento de las asignaturas se pueden tener $4! \times 3! \times 2! \times 1!$ ordenamientos de los libros. Por lo tanto hay $4! \times 4! \times 3! \times 2! \times 1!$ maneras de colocar los libros en el estante.

III) Muestreo sin Orden y sin Reposición.

a) **Combinaciones.** De una urna que contiene m bolas se extraen n bolas ($n \leq m$) sin reponerlas en la urna y sin tomar en cuenta el orden de extracción. Cada uno de los resultados de este muestreo se denomina **combinación**.

Dado un conjunto de m elementos veremos de cuántas maneras pueden combinarse estos elementos en grupos de n . Por una parte, se tiene que cada muestra (combinación) de tamaño n produce $n!$ muestras ordenadas, es decir $n!$ variaciones de n elementos en grupos de n ; además toda muestra ordenada (de n elementos) es una de las $n!$ ordenaciones obtenidas permutando un grupo de n elementos. Es decir, las

variaciones se obtienen al considerar todas las permutaciones posibles de las combinaciones. Entonces, se tiene que si x es el número de combinaciones de m elementos en grupos de n ,

$$V_n^m = \frac{m!}{(m-n)!} = n!x;$$

por lo que

$$x = \frac{m!}{n!(m-n)!} = C_n^m = \binom{m}{n},$$

donde la cantidad a la derecha de la ecuación denota

$C_n^m = \#$ de subconjuntos de n elementos extraídos de un conjunto de m elementos.

Ejemplos 1.9

1. De un grupo de 20 personas se forma una comisión de 3. ¿Cuántas comisiones diferentes pueden formarse? Hay $C_3^{20} = \binom{20}{3} = \frac{20!}{3!(20-3)!} = \frac{20 \times 19 \times 18}{3 \times 2 \times 1} = 1140$ comisiones posibles.
2. Juego del Loto. Como a efectos del juego el orden en que se obtienen los números no afecta el resultado, hay

$$C_6^{54} = \binom{54}{6} = \frac{54 \times 53 \times 52 \times 51 \times 50 \times 49 \times 48!}{6! \times 48!} = 25,827,165$$

posibles juegos o resultados.

3. De un grupo de 5 hombres y 7 mujeres, a) ¿cuántos posibles comités consistentes de 2 hombres y 3 mujeres pueden formarse? b) ¿Cuántos si dos de las mujeres están peleadas y se rehusan a participar juntas en el mismo comité?
 - i. Las mujeres pueden elegirse de $\binom{7}{3}$ maneras y los hombres de $\binom{5}{2}$ maneras, por lo tanto, usando el principio básico de conteo hay

$$\binom{7}{3} \times \binom{5}{2} = \frac{(7 \times 6 \times 5)}{(3 \times 2 \times 1)} \frac{(5 \times 4)}{(2 \times 1)} = 350$$

- ii. Si dos de las mujeres no quieren trabajar juntas hay $\binom{2}{0} \times \binom{5}{3}$ grupos de 3 mujeres en que ninguna de las mujeres peleadas participa y $\binom{2}{1} \times \binom{5}{2}$ grupos de 3 mujeres en que sólo una de ellas participa. Como hay $\binom{5}{2}$ maneras de elegir a los hombres, se tienen

$$\left[\binom{2}{0} \times \binom{5}{3} + \binom{2}{1} \times \binom{5}{2} \right] \times \binom{5}{2} = 300$$

posibles comites.

Observe que otra manera de encontrar el número de comités en este caso es restando del resultado en la parte i el número de comités en que participan las dos mujeres, que es igual a $\left[\binom{2}{2} \times \binom{5}{1} \right] \times \binom{5}{2} = 50$.

b) Permutaciones de m Bolas Distinguibles por Grupos. Supongamos que se tiene m_1 bolas de color 1, m_2 bolas de color 2, $\dots, m_r$ bolas de color r ; donde $m_1 + m_2 + \dots + m_r = m$. Supongamos además que los colores son diferentes pero no hay manera de distinguir entre bolas de un mismo color. ¿Cuántas disposiciones diferentes de las bolas son posibles?

Como hemos visto las m bolas pueden ordenarse de $m!$ maneras. Sin embargo, en cada una de estas disposiciones, los ordenamientos de bolas del mismo color no son distinguibles; y como las bolas de color 1 se pueden ordenar de $m_1!$ maneras, las bolas de color 2 de $m_2!$ maneras, $\dots$ las de color r de $m_r!$ maneras, se tiene un total de $m_1! \times m_2! \times \dots \times m_r!$ ordenamientos indistinguibles. Por lo tanto hay

$$\frac{m!}{m_1! \times m_2! \times \dots \times m_r!} = \binom{m}{m_1, m_2, \dots, m_r}$$

disposiciones distinguibles.

Ejemplos 1.10

1. En un torneo de ajedrez hay 10 participantes de los cuales 4 son rusos, 3 estadounidenses, 2 británicos y 1 brasileño. Si el resultado del torneo sólo muestra la nacionalidad del jugador en el orden en que se colocó, ¿cuántos posibles resultados hay?

• Como no se distingue entre los participantes de una misma nacionalidad, se tienen

$$\binom{10}{4,3,2} = \frac{10!}{4! \times 3! \times 2!} = 12,600$$

posibles ordenamientos.

2. ¿Cuántas señales diferentes pueden formarse con 9 banderas de las cuales 5 son blancas, 2 rojas y 2 azules si banderas de un mismo color son idénticas?

• Como no se distingue entre banderas de un mismo color, se tienen $\binom{9}{5,2,2} = \frac{9!}{5!2!2!} = 756$ señales diferentes.

III) Muestreo sin Reposición y sin Orden

EXTRACCION	COLOCACIÓN
▪ (a) Muestreo sin Reposición ▪ (b) No Ordenado	$n \leq m$ ▪ (a') Ninguna caja puede contener más de una bola ▪ (b') Las bolas no están numeradas.
$n < m$ Resultado $a_1, \dots, a_n$ donde $a_i \neq a_j \ i \neq j$.	▪ Se coloca una bola al azar en cada una de las n cajas.

3. ¿Cuántos posibles arreglos pueden formarse usando las letras de la palabra PEPPER?

- Análogamente se tienen $\binom{6}{3,2} = \frac{6!}{3!2!} = 60$ arreglos posibles.

4. Juegos de Cartas. P = pinta y D = denominación o valor.

$P \setminus D$	A	2	3	4	5	6	7	8	9	10	J	Q	K
♣	A	2	3	4	5	6	7	8	9	10	J	Q	K
♦	A	2	3	4	5	6	7	8	9	10	J	Q	K
♥	A	2	3	4	5	6	7	8	9	10	J	Q	K
♠	A	2	3	4	5	6	7	8	9	10	J	Q	K

Determine el número de manos de cartas con las características especificadas y calcule la probabilidad de obtener los juegos correspondientes asumiendo que el mazo de cartas se baraja concienzudamente y que no se hace trampa al distribuir las:

a) Ω = **Posibles manos de poker**. Como una mano consta de 5 cartas y el orden en que las cartas aparecen no es importante, se tienen $\binom{52}{5}$ posibles

manos.

b) $A = \mathbf{5 \text{ cartas de una pinta.}}$ Hay 4 pintas con 13 denominaciones (valores) diferentes. Entonces para cada pinta se tienen $\binom{13}{5}$ maneras de seleccionar 5 cartas de esa pinta. Por consiguiente, el número total de manos es en este caso

$$|A| = 4 \times \binom{13}{5} \quad y \quad P(A) = \frac{|A|}{|\Omega|} = \frac{4 \times \binom{13}{5}}{\binom{52}{5}}$$

c) $B = \mathbf{Un \text{ par: } \{a, a, b, c, d\}; \text{ donde } a, b, c \text{ y } d \text{ son valores distintos.}}$

- El valor a lo elegimos entre 13 valores posibles.
- Para cada valor de a hay 4 pintas posibles de las que podemos elegir 2 pintas de $\binom{4}{2}$ maneras.
- Una vez hecha esta elección, podemos escoger los valores b, c y d entre los 12 restantes de $\binom{12}{3}$ maneras.
- Finalmente observemos que cada una de las cartas con valores b, c y d puede presentarse en 4 pintas diferentes, por lo que

$$|B| = 13 \times \binom{4}{2} \times \binom{12}{3} \times 4^3 \quad y \quad P(B) = \frac{13 \times \binom{4}{2} \times \binom{12}{3} \times 4^3}{\binom{52}{5}}$$

posibles manos con un par.

d) $C = \mathbf{Dos \text{ pares: } \{a, a, b, b, c\} \text{ donde } a, b \text{ y } c \text{ son valores distintos.}}$

- Hay $\binom{13}{2}$ denominaciones (valores) posibles para los pares.
- El primer par se elige de entre 4 pintas de $\binom{4}{2}$ maneras. De forma análoga se elige el segundo par.
- Para la última carta hay 11 valores posibles cada uno en 4 pintas diferentes. Por lo tanto el número de manos con dos pares es igual a

$$|C| = \binom{13}{2} \times \binom{4}{2} \times \binom{4}{2} \times 11 \times 4 \quad y \quad P(C) = \frac{\binom{13}{2} \times \binom{4}{2} \times \binom{4}{2} \times 11 \times 4}{\binom{52}{5}}.$$

e) $D = \mathbf{Un \text{ trio: } \{a, a, a, b, c\} \text{ donde } a, b \text{ y } c \text{ son valores distintos.}}$

- La denominación (valor) de a puede seleccionarse de 13 maneras.
- Para cada una de estas denominaciones las pintas se pueden escoger de $\binom{4}{3}$ maneras.
- Análogamente las denominaciones b y c se pueden elegir de $\binom{12}{2}$ maneras

diferentes y para cada una de estas denominaciones se tienen 4 pintas posibles.
Por lo tanto

$$|D| = 13 \times \binom{4}{3} \times \binom{12}{2} \times 4^2 \quad \text{trios y} \quad P(D) = \frac{13 \times \binom{4}{3} \times \binom{12}{2} \times 4^2}{\binom{52}{5}}$$

f) $E =$ **Un cuarteto:** $\{a, a, a, a, b\}$ donde a y b son valores distintos.

- Hay 13 maneras de elegir el valor de a que aparecerá en todas las pintas, esto es $\binom{4}{4} = 1$ formas.

- Para la carta restante hay 12 maneras de escoger el valor b y 4 pintas posibles.
Por lo tanto

$$|E| = 13 \times \binom{4}{4} \times 12 \times 4 \quad \text{y} \quad P(E) = \frac{13 \times \binom{4}{4} \times 12 \times 4}{\binom{52}{5}}$$

g) $F =$ **Un par y un trio (casa llena):** $\{a, a, a, b, b\}$ donde a y b son valores distintos.

- Hay 13 maneras de elegir el valor a y 12 de elegir el valor b .

- Para cada elección de a hay $\binom{4}{3}$ maneras de elegir las pintas del trio y por cada valor de b $\binom{4}{2}$ por lo tanto

$$|F| = 13 \times \binom{4}{3} \times 12 \times \binom{4}{2} \quad \text{y} \quad P(D) = \frac{13 \times \binom{4}{3} \times 12 \times \binom{4}{2}}{\binom{52}{5}}.$$

NOTA: Observe como se implementa el principio básico de conteo en los ejemplos anteriores.

IV) Muestreo sin Orden y con Reposición.

Consideramos una caja que contiene m bolas numeradas. Se extrae una bola, se anota el número y se devuelve la bola a la caja. Este proceso se repite r veces y los números extraídos se anotan sin tomar en cuenta el orden en que aparecen.

En primer lugar determinaremos el número de elementos del espacio muestral asociado a éste experimento. Una manera natural de enfocar esta situación es considerando el problema equivalente de *colocar r bolas indistinguibles en m cajas numeradas*, permitiendo más de una bola en una caja:

j_k bolas en la caja $\# k$ equivale a sacar la bola con el número k j_k veces.

Veamos que en efecto ambos problemas son equivalentes. Un resultado del espacio muestral asociado al muestreo sin orden y con reposición es de la forma $(a_{j_1}, \dots, a_{j_r})$ donde $1 \leq j_k \leq m$, $1 \leq k \leq r$. La notación vectorial no significa orden en la extracción, sin embargo resulta conveniente a efectos de calcular el número de elementos. Dado que el muestreo es con reposición, los j'_k s pueden repetirse:

a_{j_1} significa que se coloca una bola en la caja j_1 ,

a_{j_2} significa que se coloca una bola en la caja j_2 ,

a_{j_3} significa que se coloca una bola en la caja j_3 ,

...

a_{j_r} significa que se coloca una bola en la caja j_r .

NOTA: Cada vez que el índice j_i aparezca se coloca una bola en la caja j_i .

Podemos imaginar que sacamos aleatoriamente y con reposición r fichas, numeradas del 1 al m , contenidas en una caja. El número obtenido en una extracción es el número de la caja donde se colocará la bola siguiente.

En un resultado en el que el objeto a_i se repite r_i veces, $i = 1, \dots, m$, se tiene que

$$\sum_{i=1}^m r_i = r$$

y la caja i contiene r_i bolas o equivalentemente el número i se sacó r_i veces.

Si $r < m$ entonces al menos $m - r$ de los r_i son iguales a cero; esto equivale a que al menos $m - r$ cajas queden vacías. Si $r \geq m$ todavía puede haber cajas vacías (r_i 's iguales a cero). Calculemos el número de elementos del espacio muestral.

Podemos representar las m cajas por $m - 1$ barras verticales (|) y las bolas por estrellas (*). Un posible elemento en el caso en que $m = 8$ y $r = 8$ sería

$$\{ * || * * * | * | * * || * | \}.$$

En la representación anterior este elemento es $(a_1, a_3, a_3, a_3, a_4, a_5, a_5, a_7)$ por lo que

$$r_1 = 1, \quad r_2 = 0, \quad r_3 = 3, \quad r_4 = 1, \quad r_5 = 2, \quad r_6 = 0, \quad r_7 = 1, \quad y, \quad r_8 = 0.$$

Toda configuración consta de $m - 1$ barras y r estrellas, por lo tanto la situación es análoga a tener banderas de dos colores combinadas al azar para formar señales: **hay un total de $m - 1 + r$ banderas de las cuales $m - 1$ son del color 1 y r son del color 2.** El número de elementos del espacio muestral es el número de posibles

señales que se pueden formar con estos dos simbolos; es decir,

$$|\Omega| = \frac{(m+r-1)!}{(m-1)!r!} = \binom{m+r-1}{r} = \binom{m+r-1}{m-1}.$$

IV) Muestreo con Reposición y sin Orden

EXTRACCION	COLOCACIÓN
▪ (a) Muestreo con Reposición ▪ (b) No Ordenado	$n \leq m$ ▪ (a') Cada caja puede contener más de una bola ▪ (b') Las bolas no estan numeradas.
$n < m$ Resultado $a_1, \dots, a_n$. ▪ Un elemento en el conjunto de resultados puede repetirse	▪ Se colocan n bolas en m cajas seleccionadas al azar permitiendo la colocación de más de una bola en una caja ▪ Las bolas no estan enumeradas

Ejemplos 1.11

1. Supongamos que se tienen 6 bolas idénticas para ser distribuidas en 4 cajas numeradas del 1 al 4. Es decir se tienen las etiquetas con los números de las cajas en una bolsa y se sacan aleatoriamente y con reposición 6 veces. Si representamos $a_{j_i} = j_i$, el resultado $(3, 4, 4, 3, 2, 1)$ significa que se colocaron una bola en la caja 1, una en la caja 2, dos en la caja 3 y dos en la caja 4. Note que, aún cuando el orden en que se colocan las bolas está registrado en el resultado, el mismo resulta irrelevante ya que lo que importa es el número de bolas que al final tiene cada caja. Por ejemplo $(3, 4, 4, 3, 2, 1)$ y $(4, 4, 3, 3, 1, 2)$ representan el mismo resultado. Esta notación resultará conveniente más adelante cuando se consideren bolas distinguibles.

2. Se tienen r dados idénticos. ¿Cuántos posibles resultados hay en un lanzamiento?

Podemos representar los resultados mediante un conjunto $\{a_{j_1}, a_{j_2}, \dots, a_{j_r}\}$, donde a_{j_i} significa un dado cae en el número j_i , $j_i \in \{1, \dots, 6\}$. Esto a su vez equivale a colocar una bola en la caja j_i .

Como hay 6 cajas y r bolas, el número de resultados posibles es igual a

$$\binom{6+r-1}{r} = \binom{6+r-1}{6-1} = \binom{5+r}{5}.$$

3. En una piñata hay 10 juguetes idénticos y 8 niños. Si cualquiera de ellos puede recoger más de un juguete, ¿de cuántas maneras pueden quedar repartidos los juguetes?

Identificamos a cada niño con una caja y a cada juguete con una bola. Entonces $n = 8$ y $r = 10$ por lo que hay

$$\binom{8+10-1}{10} = \binom{17}{10}$$

maneras de repartir los juguetes.

Algunos casos especiales.

a) **Número de colocaciones en las que no queda ninguna caja vacía.** Se tienen m cajas y se extraen r bolas, donde $r \geq n$. Si representamos un resultado como antes $\{*| \dots *|*\}$ *ninguna caja vacía* significa que no quedan dos barras contiguas. Para contar el número de elementos en este evento podemos pensar en colocar las r bolas (sin las barras): $\{***** \dots *\}$ con lo que quedan $r - 1$ espacios para ser ocupados por las $m - 1$ barras. Una disposición de bolas en cajas se obtiene cuando se seleccionan $m - 1$ espacios, de entre los $r - 1$ disponibles, para colocar las barras. Esto puede hacerse de $\binom{r-1}{m-1}$ formas, por lo que éste es el número de elementos en el evento considerado.

Ejemplos 1.12

1. En el ejemplo anterior (3) ¿cuál es la probabilidad de que no quede ningún niño sin juguete?

$$\frac{\binom{9}{7}}{\binom{17}{10}}.$$

b) **Número de colocaciones en que ninguna caja contiene más de una bola.** Esto equivale a extraer r bolas, $r \leq m$ sin reemplazo. Un resultado típico de este evento se caracteriza por no presentar dos estrellas contiguas.

Si se disponen las $m - 1$ barras

$$\{\star | \star | \star | \dots | \star\}$$

que generan m espacios donde se colocaran r bolas; las disposiciones (resultados) posibles, se obtienen escogiendo r espacios (cajas) entre los m disponibles. Dado que el orden en que se seleccionan los espacios (cajas) no es relevante, cada uno de estos resultados es una combinación de m espacios (cajas) en grupos de r , por lo que hay $\binom{m}{r}$ disposiciones posibles.

c) **Número de colocaciones de r bolas distinguibles en m cajas.** Se colocan al azar r bolas numeradas en m cajas numeradas. Esto puede hacerse mediante la extracción aleatoria de r etiquetas, **con orden y con reposición**, de una urna que contiene m etiquetas enumeradas de 1 hasta m . Un resultado puede representarse como una r -tupla ordenada

$$(b_1, b_2, \dots, b_r)$$

donde la i -ésima coordenada corresponde a la caja donde se colocó la bola i . Por la forma en que se realiza el muestreo, es claro que una caja puede contener más de una bola, ya que al hacerse el muestreo con reposición el número de una caja puede extraerse varias veces.

Usando el principio básico de conteo, se concluye que el espacio muestral correspondiente a este experimento aleatorio contiene m^r elementos.

Ejemplos 1.13

Consideremos el evento r_i bolas en la caja j_i donde $r_1 + r_2 + \dots + r_k = r$, $r_i \geq 0$ y $1 \leq j_i \leq m$ para $i = 1, \dots, k$. El número de elementos en este evento es

$$\frac{m!}{r_1! r_2! \dots r_k!}.$$

Verifiquelo.

Fórmulas Combinatorias.

En esta sección exploraremos algunas propiedades de las combinaciones de m elementos en grupos de n , que pueden resultar de utilidad para la realización de cálculos y la verificación de otras propiedades.

Se considera un conjunto $C = \{c_1, \dots, c_m\}$ del cual se extrae un subconjunto de n elementos. Entonces,

- c1. $\binom{m}{m} = 1$.
 c2. $\binom{m}{1} = m$.
 c3. $\binom{m}{n} = \binom{m}{m-n}$, para $0 \leq n \leq m$.

La igualdad en c3 se puede verificar observando que todo subconjunto de n elementos produce un, y sólo un, subconjunto de $m - n$ elementos, su complemento. Por otra parte, todo subconjunto de $m - n$ elementos es complemento de un, y sólo un, conjunto de n elementos. Esto implica que elegir un subconjunto de n elementos equivale a elegir uno de $m - n$ elementos y reciprocamente.

- c4. *Fórmula de Steaffectel.* $\binom{m}{n} = \binom{m-1}{n-1} + \binom{m-1}{n}$ para $0 \leq n \leq m - 1$.

Denotemos como $\mathcal{D}$ la familia de todos los subconjuntos de C que contienen n elementos. Entonces para un elemento c_1 de C podemos descomponer $\mathcal{D}$ como la unión disjunta $\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2$ donde $\mathcal{D}_1$ son los subconjuntos de C que contienen a c_1 y $\mathcal{D}_2$ son los subconjuntos de C que no contienen a c_1 . Por lo tanto

$$\binom{m}{n} = |\mathcal{D}| = |\mathcal{D}_1| + |\mathcal{D}_2| = \binom{m-1}{n-1} + \binom{m-1}{n}$$

- c5. $\binom{n+m}{r} = \binom{n}{0}\binom{m}{r} + \binom{n}{1}\binom{m}{r-1} + \binom{n}{2}\binom{m}{r-2} + \dots + \binom{n}{r}\binom{m}{0}$.

- c6. *Fórmula del binomio de Newton.* $(a+b)^m = \sum_{k=0}^m \binom{m}{k} a^k b^{m-k}$. Esto se demuestra usando inducción sobre m .

- c7. *Corolario de la fórmula anterior:* $2^m = \sum_{k=0}^m \binom{m}{k} = (1+1)^m$.

Ejercicios 1.4

1. Un grupo de seis personas va al teatro. ¿De cuántas maneras diferentes pueden sentarse en seis asientos adyacentes?
2. Si el grupo en el ejercicio anterior está conformado por tres hombres y tres mujeres, calcule la probabilidad de que dos personas del mismo sexo nunca queden en asientos contiguos.
3. Suponga que siete espacios adyacentes van a ser ocupados por las letras c y s de forma tal que c aparece cuatro veces y s tres veces. ¿Cuántos arreglos diferentes pueden formarse de esta manera? Nota: recuerde que no se distingue entre dos letras

c ó dos letras s.

4. Considere un juego de lotto en el que se gana el premio máximo acertando seis números tomados de un panel de números del 1 al 40. Calcule la probabilidad de ganar con un juego.

5. Si 4 venezolanos, 3 franceses y 3 italianos se sientan en una fila, ¿cuántas disposiciones son posibles si personas de la misma nacionalidad deben sentarse juntas?

6. a) ¿De cuántas maneras pueden 3 niños y 3 niñas sentarse en una fila?

b) ¿De cuántas maneras pueden 3 niños y 3 niñas sentarse en una fila si los niños deben sentarse juntos y las niñas deben sentarse juntas?

c) ¿De cuántas maneras si sólo los niños deben sentarse juntos?

d) ¿De cuántas maneras si no se permite que dos personas del mismo sexo se sienten juntas?

7. ¿De cuántas maneras pueden sentarse 8 personas en fila en las condiciones siguientes: a) no hay restricciones;

b) las personas A y B deben sentarse juntas;

c) hay 4 hombres y 4 mujeres y no se sientan juntas dos personas del mismo sexo;

d) hay 5 hombres y ellos deben sentarse juntos;

e) hay 4 parejas casadas y cada pareja debe sentarse junta?

8. Un niño tiene 12 bloques de madera de los cuales 6 son negros, 4 son rojos, 1 es blanco y 1 es azul. Si el niño pone los bloques en línea, ¿cuántos arreglos son posibles?

9. ¿Cuántos arreglos pueden hacerse con las letras de :

a) FACIL; b) MANZANO; c) MISSISSIPPI; d) COMARCA.

10. Un departamento de policía en un pueblo pequeño tiene 10 oficiales. Si el departamento tiene 5 oficiales patrullando las calles, 2 trabajando en la estación y 3 en reserva en la estación, ¿cuántas divisiones diferentes de los 10 oficiales en los 3 grupos son posibles?

11. Una junta de 7 se forma con 2 estudiantes, 2 empleados y 3 profesores que se escogen de un grupo de 5 estudiantes, 6 empleados y 4 profesores. ¿Cuántas juntas son posibles?

12. Un presidente un tesorero y un secretario van a ser elegidos de un grupo de 10 personas. Dé el número de elecciones posibles en cada uno de los siguientes casos:

a) No hay restricciones.

b) A y B no trabajan juntos.

- c) C y D trabajan juntos o no trabajarán.
- d) E debe tener uno de los cargos.
- e) F aceptará sólo el cargo de presidente?

13. Se dividen 12 personas en 3 comisiones de tamaño 3, 4 y 5. ¿Cuántas comisiones diferentes pueden formarse?

14. Una prueba consta de 12 preguntas con respuestas V (verdadero) y F (falso). Un estudiante que no estudió pide ayuda al profesor y éste le dice que 6 de las preguntas son verdaderas y 6 falsas. ¿Cuál es la probabilidad de que el estudiante responda correctamente a todas las preguntas?

15. Una persona llega a una reunión en la que no conoce a nadie. Si hay 6 mujeres y 4 hombres y se sabe que hay 4 matrimonios, ¿de cuántas maneras puede la persona figurarse como están formados los matrimonios? ¿De cuántas maneras si hay 3 matrimonios?

16. Un comité de 5 se selecciona aleatoriamente de un grupo de 6 hombres y 9 mujeres. ¿Cuál es la probabilidad de elegir 3 hombres y 2 mujeres?

17. Si dos bolas se extraen al azar y al mismo tiempo de una caja que contiene 6 bolas blancas y 5 bolas negras, ¿cuál es la probabilidad de que una de las bolas sea blanca y la otra negra?

18. Se escogen al azar 4 zapatos de un conjunto de 5 pares. ¿Cuál es la probabilidad de que se forme por lo menos un par?

19. Suponga que Marcos está asistiendo a una convención para nacidos en el mes de Enero. Marcos y otras siete personas se quedan encerrados en un ascensor y para pasar el tiempo Marcos trata de calcular la probabilidad de que todas las ocho personas atrapadas tengan cumpleaños diferentes. ¿Qué respuesta debería obtener Marcos? ¿Cuál es la probabilidad de que al menos dos personas tengan el mismo cumpleaños? Nota: Asuma que las ocho personas son nacidas en enero.

20. Si r personas se encuentran en un edificio público, ¿cuál es la probabilidad de que al menos dos de ellas tengan la misma fecha (día/mes) de cumpleaños?

21. Si se lanzan dos dados, ¿cuál es la probabilidad de que la suma de las caras sea igual a $i \in \{2, 3, \dots, 12\}$?

22. ¿Cuál es la probabilidad de sacar 6 ó menos de 6 con tres dados distintos? ¿Cuál es la probabilidad si los dados son idénticos? Suponga que los dados son perfectamente

simétricos.

23. Una urna contiene 3 bolas rojas y 7 negras. Jugadores A y B extraen bolas de la urna consecutivamente y comenzando por A, hasta que una bola roja sea seleccionada. Encuentre la probabilidad de que A saque la bola roja. Sugerencia. Considere los eventos: $E_A = A$ saca la bola roja y $E_i =$ la bola roja se saca en la i -ésima extracción.

24. Se lanza una moneda hasta que aparezca por primera vez el mismo resultado dos veces. Asumiendo que cada cara de la moneda tiene igual probabilidad, halle la probabilidad de los siguientes eventos: $A =$ El experimento termina antes del sexto lanzamiento. $B =$ Se requiere un número par de lanzamientos.

25. Una urna contiene 5 bolas rojas, 6 bolas azules y 8 bolas verdes. Si un conjunto de 3 bolas es extraído aleatoriamente, ¿cuál es la probabilidad de que todas las bolas sean del mismo color? ¿Cuál es la probabilidad de que todas sean de colores diferentes? Nota: considere los casos en que las extracciones se realizan con reemplazamiento y sin reemplazamiento.

26. Un grupo de 6 hombres y 6 mujeres se dividen aleatoriamente en dos grupos de 6 personas cada uno. Encuentre la probabilidad de que cada grupo tenga el mismo número de hombres.

27. Demuestre la fórmula de Steaffel.

28. Dé un argumento combinatorio para justificar la fórmula c5. Demuestre analíticamente la fórmula.

29. Demuestre la fórmula del binomio de Newton.

30. Se reparten al azar 5 bolas distinguibles en 3 casillas. Determine la probabilidad de que por lo menos haya una en cada casilla.

31. ¿De cuántas maneras pueden colocarse n bolas idénticas en n cajas numeradas de tal modo que exactamente dos cajas queden vacías? ¿De cuántas maneras de tal modo que una caja quede vacía?

32. Se colocan n bolas distinguibles en n cajas numeradas. Encuentre la probabilidad de que exactamente una caja quede vacía.

33. Si se colocan n bolas numeradas en n cajas numeradas, ¿cuál es la probabilidad de que cada caja contenga una bola?

34. Los números $1, 2, 3, \dots, n$ están colocados en un orden aleatorio, encuentre la probabilidad de que

a) los dígitos 1 y 2 estén en su posición correcta.

b) los dígitos 1, 2 y 3 estén en su posición correcta.

35. Se tienen $2k$ cajas numeradas (desde 1 hasta $2k$) y k bolas numeradas (desde 1 hasta k). Se distribuyen las k bolas en las $2k$ cajas de modo que cada caja contenga a lo sumo una bola.

a) Halle el número de configuraciones tales que la caja #1 resulte ocupada.

b) Halle el número de configuraciones tales que la caja #1 resulte vacía.

c) Halle el número de configuraciones tales que la caja #1 resulte vacía y la caja #2 contenga la bola #2.

36. Un ascensor con 7 pasajeros se detiene en 10 pisos. ¿Cuál es la probabilidad de que 2 pasajeros no se bajen en el mismo piso?

37. Se disponen en fila 3 bolas blancas y 6 bolas negras de manera que no haya dos bolas blancas consecutivas. ¿Cuántas maneras hay de hacerlo?

38. Se lanzan 6 dados distinguibles.

a) ¿De cuántas maneras pueden salir sus caras superiores?

b) ¿De cuántas maneras puede ocurrir que sus caras superiores sean todas distintas?

c) ¿Cuál es la probabilidad de que todas sus caras superiores sean iguales?

d) ¿Cuál es la probabilidad de que al menos dos dados muestren sus caras superiores iguales?

39. ¿Cuántos enteros hay entre 10^6 (un millón) y 10^7 (diez millones) tales que

a) no tengan dos dígitos consecutivos iguales?

b) tengan todos sus dígitos diferentes?

40. Consideremos el conjunto de todas las poligonales (n, S_n) con $n = 0, 1, 2, \dots$, que parten del origen, es decir, $S_0 = 0$ y que en cada paso saltan una unidad hacia arriba o una unidad hacia abajo; es decir,

$$S_n = X_1 + X_2 + \dots + X_n,$$

donde $X_k = 1$ ó -1 ($k = 1, 2, \dots, n$). $\{S_n\}_{n \geq 1}$ así definido se denomina paseo al azar o caminata aleatoria.

a) ¿Cuántas poligonales se pueden construir en el intervalo de tiempo $[0, n]$?

b) ¿Cuántas poligonales son tales que $S_n = 0$?

41. Dos cajas contienen $2n$ bolas cada una, numeradas desde 1 hasta $2n$. Seleccionamos n bolas de cada caja. Halle la probabilidad de que haya a lo sumo una bola con el mismo número en ambos conjuntos?

42. Se tiene un lote con N productos entre los cuales hay n defectuosos. Halle la probabilidad de que en una muestra aleatoria de tamaño k se encuentren:

a) j piezas defectuosas, $j \leq n$ y $j \leq k$.

b) Al menos j piezas defectuosas, $j \leq n$ y $j \leq k$.

c) A lo sumo j piezas defectuosas, $j \leq n$ y $j \leq k$.

43. Se tienen 6 palos de igual longitud y cada uno de ellos se parte en dos: uno largo y uno corto. Con los 12 palos así obtenidos se vuelven a formar 6 palos reagrupándolos al azar en parejas.

a) Halle la probabilidad de juntar los trozos en la forma original.

b) Halle la probabilidad de que cada trozo largo sea unido a uno corto.

Capítulo 2

Probabilidad Condicional, Fórmula de Bayes e Independencia.

2.1. Probabilidad Condicional.

Consideramos situaciones en las que para calcular la probabilidad de un evento se tiene información adicional como, por ejemplo la ocurrencia de algún otro evento. Nos preguntamos ¿en qué medida afecta esta información la probabilidad buscada? Ilustraremos esta situación con algunos ejemplos.

Ejemplos 2.1

1. Considere el espacio muestral correspondiente al lanzamiento de dos dados (uno rojo y uno verde) y supongamos que todos los resultados, denotados como (r, v) , son igualmente probables.

Sean A = suma igual a 8; B = dado rojo muestra 2 y C = suma mayor que 10. Supongamos que sabemos que el dado rojo muestra 2; es decir, ocurre B . Entonces, dada esta información, ¿cuál es la probabilidad de que la suma sea igual a 8? ¿cuál es la probabilidad de que la suma sea mayor que 10?

Si el dado rojo muestra 2, entonces los resultados posibles son:

$$S' = \{(2, 1), (2, 2), (2, 3), (2, 4), (2, 5), (2, 6)\};$$

siendo S' el espacio muestral actualizado, en el que todos los resultados son igualmente probables. Se tiene que la probabilidad de A dado B es igual a $\frac{1}{6}$. Note que esta probabilidad es diferente de $P(A) = \frac{5}{36}$.

Observando el espacio muestral actualizado S' , se verifica que la probabilidad de que ocurra C dado que ha ocurrido B es igual a 0, que también es diferente de $P(C)$.

2. ¿Tiene el rey una hermana? El rey viene de una familia de dos niños. ¿Cuál es la probabilidad de que tenga una hermana?

Podemos describir el espacio muestral como $\{(h, h), (h, v), (v, h), (v, v)\}$. Asumimos que todos los resultados son igualmente probables. Como sabemos que uno de los niños es el rey, el espacio muestral actualizado es $S' = \{(h, v), (v, h), (v, v)\}$, por lo que la probabilidad de que el rey tenga una hermana es $\frac{2}{3}$.

3. Considere una población de 20 estudiantes entre los cuales hay 14 que estudian Ciencias y 6 que estudian Ingeniería. Se escogen al azar, en forma sucesiva y sin reposición dos estudiantes, de forma que el espacio muestral puede representarse como un conjunto de pares (e_1, e_2) . Supongamos que el primer estudiante es de Ciencias, ¿cuál es la probabilidad de que el segundo también lo sea?

Observe que, dado que el muestreo es sin reposición, el número de eventos elementales (ó resultados) es 20×19 .

Sean E_1 =el primer estudiante es de Ciencias y

E_2 =el segundo estudiante es de Ciencias.

El espacio muestral actualizado es: E_1 , este espacio tiene 14×19 posibles resultados de los cuales 14×13 son favorables a E_2 , por lo tanto la probabilidad de que ocurra E_2 dado que ocurre E_1 es

$$\frac{14 \times 13}{14 \times 19} = \frac{\frac{(14 \times 13)}{(20 \times 19)}}{\frac{(14 \times 19)}{(20 \times 19)}} = \frac{P(E_1 \cap E_2)}{P(E_1)}.$$

Esta probabilidad se denomina probabilidad condicional de E_2 dado E_1 y se denota $P(E_2|E_1)$. Una vez más observe que $P(E_2) = \frac{(14 \times 19)}{(20 \times 19)} = \frac{7}{10} \neq \frac{13}{19}$.

Si las selecciones se realizan con reposición el resultado de la primera extracción no nos da información alguna en cuanto al resultado de la segunda, por lo que

$$P(E_2|E_1) = P(E_2) = \frac{14 \times 20}{20 \times 20} = \frac{7}{10}.$$

Definición 2.1 Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $B \in \mathcal{A}$ tal que $P(B) > 0$. Definimos la función de conjuntos probabilidad condicional dado B, denotado $P(\bullet|B)$ mediante

$$P(\bullet|B) : \mathcal{A} \longrightarrow [0, 1],$$

donde

$$P(A|B) = \frac{P(A \cap B)}{P(B)}, \text{ para todo } A \in \mathcal{A}.$$

NOTA: en los ejemplos anteriores se calcularon: $P(A|B)$, $P(h|v)$ y $P(E_2|E_1)$.

2.1.1. Propiedades básicas.

Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $B \in \mathcal{A}$ tal que $P(B) > 0$. Entonces

1. La función $P(\bullet|B)$ es una probabilidad. En efecto

i) $P(A|B) \geq 0$ para cada $A \in \mathcal{A}$.

ii) $P(\Omega|B) = \frac{P(\Omega \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

iii) Sea $(A_n)_{n \geq 1}$ una sucesión de eventos disjuntos dos a dos. Entonces

$$\begin{aligned} P(\cup_{n \geq 1} A_n | B) &= \frac{P(\cup_{n \geq 1} A_n \cap B)}{P(B)} = \frac{P(\cup_{n \geq 1} (A_n \cap B))}{P(B)} \\ &= \sum_{n \geq 1} \frac{P(A_n \cap B)}{P(B)} = \sum_{n \geq 1} P(A_n | B). \end{aligned}$$

2. Si A y B son disjuntos, entonces $P(A|B) = \frac{P(A \cap B)}{P(B)} = 0$.

3. Si $B \subseteq A$ entonces $P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B)}{P(B)} = 1$.

Ejemplos 2.2

1. A y B juegan el siguiente juego de dados: Se lanzan dos dados distintos, si la suma de los dados es igual a 7 A gana y si la suma es igual a 8 B gana. El juego termina al obtener un ganador. ¿Cuál es la probabilidad de que A gane?

Solución 1. Sea p_n la probabilidad de que A gane en n lanzamientos. Esto significa que en los lanzamientos $1, 2, \dots, n-1$ no se obtiene un 7 o un 8. Denotamos el evento la suma es 8 por $E_8 = \{(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)\}$ y el evento la suma es 7 por $E_7 = \{(1, 6), (6, 1), (2, 5), (5, 2), (3, 4), (4, 3)\}$. Tenemos entonces que $p_n = \frac{1}{6}(\frac{25}{36})^{n-1}$. Con lo cual

$$P(A) = \sum_{n \geq 1} p_n = \frac{1}{6} \sum_{n \geq 1} (\frac{25}{36})^{n-1} = \frac{1}{6} \sum_{n \geq 0} (\frac{25}{36})^n = \frac{6}{11}.$$

Solución 2. Lo anterior es equivalente a decir que $P(A)$ es la probabilidad condicional de que la suma sea 7 dado que el juego termina en algún lanzamiento, es decir

$$P(A) = P(\{7\} | \{7, 8\}) = \frac{P(\{7\})}{P(\{7, 8\})} = \frac{6}{11}.$$

2. Se lanza un dado dos veces. a) Si la suma de los resultados es 8, ¿cuál es la probabilidad de que el primer resultado sea $k \in \{1, 2, \dots, 6\}$?
 b) Si el primer lanzamiento resultó en un 3, ¿cuál es la probabilidad de que el segundo sea $k \in \{1, 2, \dots, 6\}$?
 c) Si el primer lanzamiento resultó en un 3, ¿cuál es la probabilidad de que la suma sea 7?

Solución. Denotemos por X el resultado del primer lanzamiento y por Y el resultado del segundo lanzamiento. Tenemos que $P(X = k) = \frac{1}{6} = P(Y = k)$.

a) Se quiere $P(X = k | X + Y = 8) = \frac{P((X=k) \cap (X+Y=8))}{P(X+Y=8)}$. Dado que

$$(X + Y = 8) = \{(2, 6), (6, 2), (3, 5), (5, 3), (4, 4)\}$$

se tiene que $P(X + Y = 8) = \frac{5}{36}$. Además

$$P((X = k) \cap (X + Y = 8)) = \begin{cases} \frac{1}{36} & \text{si } 2 \leq k \leq 6 \\ 0 & \text{si } k = 1. \end{cases} \quad (2.1)$$

Por lo tanto

$$P(X = k | X + Y = 8) = \begin{cases} \frac{1}{5} & \text{si } 2 \leq k \leq 6 \\ 0 & \text{si } k = 1 \end{cases} \quad (2.2)$$

$$b) P(Y = k | X = 3) = \frac{P((Y=k) \cap (X=3))}{P(X=3)} = \frac{1/36}{1/6} = \frac{1}{6} = P(Y = k).$$

Observe que el resultado del primer lanzamiento no afecta la probabilidad del resultado del segundo lanzamiento.

$$c) P(X + Y = 7 | X = 3) = \frac{P((X+Y=7) \cap (X=3))}{P(X=3)} = \frac{1/36}{1/6} = \frac{1}{6}.$$

3. Para controlar una cierta enfermedad en una población donde la proporción de enfermos es p se usa un determinado examen médico para detectar enfermos. Se sabe que la probabilidad de que al aplicar el examen a un enfermo lo muestre como tal es de 0,9, y que la probabilidad de que al aplicarlo a una persona sana la muestre como enferma es de 0.01. Calcule la probabilidad de que una persona este realmente enferma si el examen la muestra como tal.

Solución. Sean E : la persona está enferma y R : el examen resulta positivo (detecta la enfermedad). Se desea calcular

$$P(E|R) = \frac{P(E \cap R)}{P(R)}.$$

Sabemos que:

$$P(E) = p,$$

$$P(R|E) = \frac{P(R \cap E)}{P(E)} = 0,9 \Rightarrow P(E \cap R) = 0,9p$$

$$P(R|E^c) = \frac{P(R \cap E^c)}{P(E^c)} = 0,01 \Rightarrow P(E^c \cap R) = 0,01(1 - p).$$

Por lo tanto $P(R) = P(R \cap E) + P(R \cap E^c) = 0,9p + 0,01(1 - p)$ con lo que

$$P(E|R) = \frac{0,9p}{0,9p + 0,01(1 - p)} = \frac{90p}{89p + 1}.$$

En particular si $p = \frac{1}{30}$ se tiene que $P(E|R) \simeq 0,76$. Note que si p es pequeño la efectividad del examen para detectar la enfermedad es dudosa.

4. Tiempo de vida de un aparato que no envejece. Consideremos un aparato eléctrico cuyo funcionamiento no es afectado por el paso del tiempo, sino que se destruye por alguna perturbación aleatoria. Esto puede expresarse de la siguiente manera: Dado que el aparato está funcionando en el instante s su tiempo de vida a partir de ese instante evoluciona como si el aparato hubiera comenzado a funcionar en s . Si denotamos por T al tiempo de vida del aparato lo anterior significa que

$$P(T > s + h | T > s) = P(T > h) \quad h > 0,$$

es decir

$$\frac{P(\{T > s + h\} \cap \{T > s\})}{P(T > s)} = \frac{P(T > s + h)}{P(T > s)} = P(T > h).$$

Si denotamos $\phi(t) = P(T > t)$ la igualdad anterior significa que

$$\phi(s + h) = \phi(s) \times \phi(h).$$

Se puede comprobar que

$$\phi(t) = e^{-\lambda t}; \quad \text{donde } t \geq 0, \lambda \geq 0.$$

2.1.2. Resultados Básicos sobre Probabilidad Condicional.

Proposición 2.1 (*Probabilidad de la intersección*). Sean $A_1, \dots, A_n$ una colección finita de eventos tales que $P(A_1 \cap A_2 \cap \dots \cap A_n) > 0$, entonces se verifica que

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

Demostración. Observemos que

$$P(A_2|A_1) = \frac{P(A_1 \cap A_2)}{P(A_1)}, P(A_3|A_1 \cap A_2) = \frac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)},$$

y de manera similar

$$P(A_k|A_1 \cap A_2 \cap \dots \cap A_{k-1}) = \frac{P(A_1 \cap A_2 \cap \dots \cap A_k)}{P(A_1 \cap A_2 \cap \dots \cap A_{k-1})}.$$

Como

$$0 < P(A_1 \cap A_2 \cap \dots \cap A_n) \leq P(A_1 \cap A_2 \dots A_{n-1}) \leq \dots \leq P(A_1),$$

entonces todas las probabilidades condicionales están bien definidas. El resultado se obtiene por inducción sobre n .

Ejemplos 2.3

Calcule la probabilidad de que al lanzar seis dados distintos y perfectos se obtengan seis caras distintas.

Solución. Consideremos los siguientes eventos:

A_1 = el resultado del primer dado es $a_1 \in \{1, 2, \dots, 6\}$

A_2 = el resultado del segundo dado es $a_2 \neq a_1$

A_3 = el resultado del tercer dado es $a_3 \neq a_1, a_2$

A_4 = el resultado del cuarto dado es $a_4 \neq a_1, a_2, a_3$

A_5 = el resultado del quinto dado es $a_5 \neq a_1, a_2, a_3, a_4$

A_6 = el resultado del sexto dado es $a_6 \neq a_1, a_2, a_3, a_4, a_5$.

Entonces queremos calcular $P(A_1 \cap A_2 \dots \cap A_6)$ donde $P(A_1) = 1$, $P(A_2|A_1) = \frac{5}{6}$, $P(A_3|A_1 \cap A_2) = \frac{4}{6}$, $P(A_4|A_1 \cap A_2 \cap A_3) = \frac{3}{6}$, $P(A_5|A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{2}{6}$ y $P(A_6|A_1 \cap A_2 \cap A_3 \cap A_4 \cap A_5) = \frac{1}{6}$. Esto implica que

$$P(A_1 \cap A_2 \dots \cap A_6) = 1 \times \frac{5}{6} \times \frac{4}{6} \times \frac{3}{6} \times \frac{2}{6} \times \frac{1}{6} = \frac{5!}{6^5} = \frac{6!}{6^6}$$

Nota. Compare con la situación en que se sientan seis personas en seis sillas: es la misma situación si se asume que las sillas están enumeradas y las seis personas representan los seis dados.

Proposición 2.2 (*Fórmula de la Probabilidad Total.*) Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad y sea $(A_n)_{n \geq 1} \subseteq \mathcal{A}$ una partición del espacio Ω , esto es

$$\Omega = \bigcup A_n$$

y $A_i \cap A_j = \emptyset$, para $i \neq j$. Entonces para todo evento B se cumple que

$$P(B) = \sum_{i \geq 1} P(A_i)P(B|A_i).$$

Demostración. Escribimos

$$B = B \cap \Omega = B \cap \left(\bigcup_{i \geq 1} A_i \right) = \bigcup_{i \geq 1} (B \cap A_i).$$

Entonces, como la familia $(B \cap A_i)_{i \geq 1}$ es disjunta, tenemos que

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

Comentarios.

1. Si para algún n , $P(A_n) = 0$, entonces $P(B \cap A_n) = 0$, por lo que no importa que $P(B|A_n)$ no este definido, ya que el término correspondiente no aparece en la fórmula.
2. La fórmula anterior permite descomponer el evento B en eventos cuyas probabilidades pueden, en algunos casos, ser calculadas más fácilmente.

Ejemplos 2.4

1. Una urna contiene 6 bolas rojas y 4 negras. Se extraen al azar y sin reposición 2 bolas. Considere los siguientes eventos:

A_1 = la primera bola es roja; A_2 = la segunda bola es roja. Calcule $P(A_2)$.

Solución. Tenemos que $P(A_1) = \frac{6}{10} = \frac{3}{5}$ y $P(A_1 \cap A_2) = P(A_1)P(A_2|A_1) = \frac{3}{5} \cdot \frac{5}{9} = \frac{1}{3}$. Por lo tanto

$$P(A_2) = P(A_1)P(A_2|A_1) + P(A_1^c)P(A_2|A_1^c) = \frac{1}{3} + \frac{2}{5} \cdot \frac{6}{9} = \frac{3}{5}.$$

2. Una compañía de seguros clasifica a las personas en dos grupos: aquellos que son propensos a accidentes y aquellos que no lo son. Las estadísticas muestran que una persona propensa a accidentes tendrá un accidente en algún momento en un período de un año con probabilidad 0.4, mientras que la probabilidad del mismo evento para una persona no propensa a accidentes es 0.2. Se asume que el 30 % de la población es propensa a accidentes. a) Calcule la probabilidad de que un nuevo asegurado tenga un accidente en el año siguiente a la compra de la póliza, b) Suponga que un nuevo

asegurado tiene un accidente en el año siguiente a la compra de la póliza. Calcule la probabilidad de que el asegurado sea propenso a accidentes.

Solución. a) Sea A = el asegurado tiene un accidente en el año siguiente a la compra de la póliza y B = el asegurado es propenso a accidentes. Observe que $P(A|B) = 0,4$; $P(A|B^c) = 0,2$ y $P(B) = 0,3$.

Entonces

$$\begin{aligned} P(A) &= P(A \cap B) + P(A \cap B^c) = P(B)P(A|B) + P(B^c)P(A|B^c) \\ &= 0,3 \times 0,4 + 0,7 \times 0,2 = 0,12 + 0,14 = 0,26. \end{aligned}$$

b) Se desea calcular

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{P(B)P(A|B)}{P(A)} = \frac{0,3 \times 0,4}{0,26} = \frac{6}{13}.$$

2.2. La Fórmula de Bayes.

En la fórmula de la probabilidad total, supongamos que el evento B ha ocurrido. Si los eventos $(A_n)_{n \geq 1}$ se interpretan como **hipótesis** acerca de una determinada situación, puede ser de interés el saber de qué manera la ocurrencia de B afecta la probabilidad de las hipótesis. Es decir, consideramos una situación inversa a la que se presenta en la proposición anterior.

Teorema 2.1 (Teorema de Bayes) Sea $(A_n)_{n \geq 1}$ una sucesión disjunta de eventos en un espacio de probabilidad $(\Omega, \mathcal{A}, P)$ tales que $\Omega = \bigcup A_n$. Sea $B \in \mathcal{A}$ tal que $P(B) > 0$. Entonces para cada n se cumple que

$$P(A_n|B) = \frac{P(A_n)P(B|A_n)}{\sum_{i \geq 1} P(A_i)P(B|A_i)}.$$

Demostración. Es inmediata utilizando la fórmula de la probabilidad total y la definición de probabilidad condicional, ya que

$$P(A_n|B) = \frac{P(A_n \cap B)}{P(B)},$$

donde

$$P(A_n \cap B) = P(A_n)P(B|A_n);$$

y

$$P(B) = \sum_{i \geq 1} P(A_i \cap B) = \sum_{i \geq 1} P(A_i)P(B|A_i).$$

Comentario. Otra importante interpretación se obtiene si se considera $(A_n)_{n \geq 1}$ como posibles causas del evento B . La fórmula expresa la contribución de cada una de las causas en la ocurrencia de B , *dado que B ocurre*. Por ejemplo, se consideran como posibles causas de una inundación $A_1 = \text{llueve en exceso}$, $A_2 = \text{se echan desperdicios en los desagüaderos}$, $A_3 = \text{no se hace mantenimiento a las quebradas}$, $A_4 = \text{se modifica el cauce de los ríos por siembras o construcciones}$, etc.

Ejemplos 2.5

1. Un paciente sufre de una enfermedad que es fatal en el 50 % de los casos. Un posible tratamiento es quirúrgico. Se ha demostrado que de los sobrevivientes el 40 % ha tenido cirugía y de los no sobrevivientes 10 % ha tenido cirugía. Encuentre la probabilidad de que un paciente sobreviva si ha tenido cirugía.

Solución. Sean S = el paciente sobrevive y C = el paciente tuvo cirugía. Entonces $P(S) = 0,5 = P(S^c)$. Además sabemos que $P(C|S) = 0,4$ y $P(C|S^c) = 0,1$. Entonces la probabilidad buscada es

$$\begin{aligned} P(S|C) &= \frac{P(S)P(C|S)}{P(S)P(C|S) + P(S^c)P(C|S^c)} = \frac{0,4 \times 0,5}{(0,4 \times 0,5) + (0,1 \times 0,5)} \\ &= \frac{0,2}{0,2 + 0,05} = 0,8 \end{aligned}$$

2. Un examen para detectar la presencia de un virus, detecta al virus en un 95 % de los casos, cuando este está presente; pero puede presentar un falso resultado positivo en un 5 % de los casos. Si en determinada población la enfermedad aparece en alrededor de 0.3 % de las personas, calcule la probabilidad de que una persona tenga la enfermedad dado que el resultado fué positivo.

Solución. Sean R = el examen resultó positivo; E = la enfermedad está presente. Sabemos que $P(R|E) = 0,95$, $P(R|E^c) = 0,05$ y $P(E) = 0,003$. Además,

$$P(R^c|E) = 1 - P(R|E) = 1 - 0,95 = 0,05.$$

Entonces,

$$P(E|R) = \frac{P(E)P(R|E)}{P(E)P(R|E) + P(E^c)P(R|E^c)} = \frac{0,95 \times 0,003}{0,95 \times 0,003 + 0,05 \times 0,997} = 0,05.$$

3. El Dilema del Prisionero.

(a) Espacio muestral considerado por el guardia: $\{\{A, B\}; \{A, C\}; \{B, C\}\}$. Si se asume una probabilidad uniforme (todos los resultados son igualmente probables), entonces el guardia tiene razón al asignar a A la probabilidad de $\frac{2}{3}$ de ser liberado.

(b) En este espacio no es posible sin embargo considerar la probabilidad condicional de que A sea liberado dado que el guardia señala a B (o a C) como uno de los presos a ser liberados, ya que en el espacio muestral considerado no se tiene el evento el guardia dice que B (o C) será liberado.

(c) Considerar espacio muestral más complejo que incluya los eventos sugeridos por el guardia. Calcular las probabilidades en este caso bajo dos conjuntos de supisiciones: (i) el guardia señala indistintamente a B ó C cuando estos son los liberados; (ii) el guardia siempre identifica a B .

(d) Conclusiones: En el caso (i) la probabilidad de que A sea liberado no cambia una vez que el guardia identifica a uno de los prisioneros.

2.3. Independencia.

Se estudiará la situación en que la ocurrencia de un evento B **no** afecta en modo alguno la ocurrencia de otro evento A ; tal y como se intuye en el ejemplo del dilema del prisionero. En este caso ocurre que la probabilidad condicional de A dado B es la misma probabilidad de A . Esto lo observamos en el ejemplo 2.2.2 (se lanza un dado dos veces y se tiene que

$$P(Y = k|X = 3) = \frac{P((Y = k) \cap (X = 3))}{P(X = 3)} = \frac{1/36}{1/6} = \frac{1}{6} = P(Y = k),$$

lo que significa que el resultado del primer lanzamiento no influencia el resultado en el segundo lanzamiento. Veamos otro ejemplo.

Ejemplos 2.6

Un lote de 10 objetos contiene cuatro defectuosos y seis en buen estado. Se extraen dos objetos sucesivamente y con reposición. Considere los eventos:

D_1 = el primer objeto es defectuoso y D_2 = el segundo objeto es defectuoso .

En este caso $P(D_1|D_2) = P(D_1) = \frac{4}{10} = P(D_2) = P(D_2|D_1)$, ya que el espacio muestral no cambia cuando ocurre alguno de estos eventos.

Definición 2.2 Dos eventos A y B son independientes si $P(A|B) = P(A)$. Esto es equivalente a tener que $P(A \cap B) = P(A)P(B)$.

Observe que Ω y $\emptyset$ son independientes de cualquier evento.

Proposición 2.3 Si A y B son independientes, también lo son: a) A y B^c , b) A^c y B^c .

Demostración. a) Dado que $P(A) = P(A \cap B) + P(A \cap B^c)$ tenemos que

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

b) En este caso tenemos que

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

Ejemplos 2.7

Suponga que en el ejemplo 2.6 la extracción se realiza sin reposición. En este caso calculamos

$$\begin{aligned} P(D_2) &= P(D_2 \cap D_1) + P(D_2 \cap D_1^c) = P(D_2|D_1)P(D_1) + P(D_2|D_1^c)P(D_1^c) \\ &= \frac{3}{9} \frac{4}{10} + \frac{4}{9} \frac{6}{10} = \frac{36}{90} = \frac{2}{5}. \end{aligned}$$

Como, por otra parte $P(D_2|D_1) = \frac{3}{9} \neq \frac{2}{5}$ tenemos que D_1 y D_2 no son independientes.

Definición 2.3 Una familia arbitraria de eventos $\mathcal{C} = \{A_i : i \in I\}$ se dice independiente si para cualquier colección finita de eventos $A_{i_1}, A_{i_2}, \dots, A_{i_n}$ de $\mathcal{C}$ se cumple:

$$P\left(\bigcap_{j=1}^n A_{i_j}\right) = \prod_{j=1}^n P(A_{i_j}).$$

Ejemplos 2.8

Calcule la probabilidad de obtener tres 6 al lanzar ocho dados.

El problema equivale a calcular la probabilidad de obtener tres 6 al lanzar ocho veces el mismo dado. Sea E_i = se obtiene un 6 en el i -ésimo lanzamiento. Por independencia tenemos que

$$P(E_1 \cap E_2 \cap E_3 \cap E_4^c \cap E_5^c \cap E_6^c \cap E_7^c \cap E_8^c) = \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^5.$$

Se puede verificar fácilmente que esta es la probabilidad de cualquier conjunto de ocho lanzamientos de un dado en que se produzcan exactamente tres 6. Finalmente como existen $\binom{8}{3}$ conjuntos de este tipo, la probabilidad buscada es $\binom{8}{3} \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^5$.

Ejercicios 2.1

1. Una caja contiene metras de las cuales 10 son blancas, 5 amarillas y 10 negras. Se escoge una metra al azar de la caja y se observa que no es negra. ¿Cuál es la probabilidad de que sea amarilla?
2. Se lanzan dos dados perfectos. ¿Cuál es la probabilidad condicional de que al menos uno de ellos caiga en un 6 dado que caen en números diferentes?
3. Si dos dados perfectos son lanzados, ¿Cuál es la probabilidad condicional de que el primero caiga en un 6 dado que la suma de los dados es i ? Calcule la probabilidad para todos los valores i en tre 2 y 12.
4. Una urna contiene 6 bolas blancas y 9 negras. Si 4 bolas se extraen aleatoriamente y sin reposición, ¿cuál es la probabilidad de que las dos primeras sean blancas y las dos últimas negras?
5. Una urna contiene 12 bolas de las cuales 8 son blancas. Se extrae una muestra de 4 bolas con reemplazamiento (sin reemplazamiento). ¿Cuál es la probabilidad condicional (en ambos casos) de que la primera y la tercera bolas seleccionadas sean blancas, dado que la muestra contiene exactamente 3 bolas blancas?
6. El dilema de un médico. Un médico refiere a un colega la siguiente situación: "Si tengo la certeza de al menos un 80 % de que mi paciente sufre de esta determinada enfermedad, entonces siempre recomiendo cirugía, mientras que si no estoy tan seguro recomiendo evaluaciones adicionales que resultan costosas y, en algunos casos, dolorosas. Con el señor Gonzalez inicialmente tuve sólo un 60 % de certidumbre de que sufría la enfermedad, por lo cual ordené la serie de evaluaciones A que siempre dan un resultado positivo cuando el paciente tiene la enfermedad y casi nunca (da resultado positivo) si el paciente no tiene la enfermedad. El resultado fue positivo, pero cuando estaba listo para recomendar cirugía el paciente me informa por primera vez que es diabético. Esto complica las cosas porque, aunque no cambia mi 60 % de certidumbre inicial, afecta la interpretación de los resultados de la evaluación A, el cual no produce resultados positivos cuando el paciente es saludable, pero desafortunadamente produce resultados positivos en 30 % de los pacientes diabéticos que no sufren la enfermedad. Ahora ¿qué debo hacer? ¿Más evaluaciones o cirugía de inmediato?
7. Un examen de sangre de un laboratorio tiene efectividad de 95 % al detectar una cierta enfermedad cuando esta está en efecto presente. Sin embargo el examen puede también dar un "falso positivo" de 1 % para personas sanas. Si 0,5 % de la población tiene la enfermedad, ¿cuál es la probabilidad de que una persona tenga la enfermedad dado que el examen resulta positivo?

8. *Un detector de mentiras muestra una señal positiva (indicando una mentira) 10 % de las veces que un sujeto dice la verdad y 94 % de las veces que miente. Si dos personas son sospechosas de un crimen que se sabe ha cometido uno sólo de ellos, y ambos dicen ser inocentes, ¿cuál es la probabilidad de que una señal positiva del detector corresponda al culpable?*
9. *Se obtiene una muestra de cuatro bolas a partir de una bolsa que contiene doce, de las cuales ocho son blancas. Si el muestreo es sin reposición, halle la probabilidad de que la tercera bola sea blanca, si sabemos que la muestra tiene tres bolas blancas. ¿Que sucede si el muestreo es con reposición?*
10. *Se lanzan dos dados simétricos. Calcule la probabilidad de que la suma sea 7 dado que: a) La suma es impar. b) La suma es mayor que 6. c) El resultado del primer dado fue impar. d) El resultado del segundo dado fue par. e) El resultado de al menos un dado fue impar. f) Los dos dados tuvieron el mismo resultado. g) Los dos dados tuvieron resultados distintos. h) La suma de los dos dados fue 13.*
11. *Una bolsa contiene cuatro bolas blancas y dos negras y otra contiene tres de cada color. Se escoge una bolsa al azar y luego se selecciona una bola, también al azar. ¿Cuál es la probabilidad de que sea blanca?*
12. *Una caja contiene 10 bombillos, cuatro de los cuales están malos. Los bombillos se prueban extrayéndolos al azar, uno a uno y sin reposición hasta sacar los cuatro en mal estado. ¿Cuál es la probabilidad de que el último en mal estado salga en la quinta prueba? ¿Y en la décima?*
13. *Cierta vacuna brinda protección contra una enfermedad, de modo que una persona vacunada tiene probabilidad 0,4 de contraer la enfermedad, mientras que una sin vacunar tiene probabilidad de 0,8 de enfermarse. 75 % de la población está vacunada. ¿Cuál es la probabilidad de que una persona que tiene la enfermedad esté vacunada?*
14. *En un hospital se aplica una prueba para detectar cáncer. Se ha determinado que en 97 % de los pacientes cancerosos el resultado de la prueba es positivo y sólo en un 5 % de los pacientes sanos la prueba arroja un resultado positivo. Si 2 % de los pacientes del hospital sufre de cáncer, ¿cuál es la probabilidad de que un paciente elegido al azar que reacciona positivamente a la prueba tenga realmente cáncer?*
15. *Suponga que 5 % de los hombres y 25 % de cada 10.000 mujeres son daltónicos. Se escoge un daltónico al azar, ¿cuál es la probabilidad de que sea hombre?*
16. *Una ferretería tiene tres cajas con igual cantidad de tornillos. Dos de las cajas tienen 5 % de tornillos malos y la otra contiene 10 %. Se escogen dos cajas al azar,*

se mezclan en otra caja y de ella se sacan cinco tornillos, uno de los cuales es defectuoso. ¿Cuál es la probabilidad de que una de las cajas de la mezcla sea la que contenía 10 % de tornillos defectuosos?

17. Tres sucursales de una tienda tienen 8, 12 y 14 empleados de los cuales 4, 7 y 10 son mujeres, respectivamente. a) Se escoge una sucursal al azar y de ella se escoge un empleado. Si éste es una mujer, ¿cuál es la probabilidad de que ella trabaje en la sucursal con 12 empleados? b) Si se escoge un segundo empleado de la misma sucursal, ¿cuál es la probabilidad de que se escoja una mujer?
18. Las señales telegráficas de **punto** y **raya** se envían en proporción de 3:4. Debido a ciertas condiciones que causan una transmisión errática, un punto se transforma en raya con probabilidad $1/4$ y una raya en punto con probabilidad $1/3$. Si se recibe un punto, ¿cuál es la probabilidad de que haya sido enviado un punto?
19. Si $A_1, \dots, A_n$ son eventos independientes, demuestre que

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = 1 - (1 - P(A_1)) \times (1 - P(A_2)) \dots \times (1 - P(A_n)).$$

20. Un dado A tiene 4 caras rojas y 2 blancas, mientras que un dado B tiene 2 caras rojas y 4 blancas. Se lanza una moneda legal. Si cae en cara se lanza el dado A y si cae en sello se lanza el B .
 - (a) Demuestre que la probabilidad de rojo en cualquier lanzamiento es $1/2$.
 - (b) Si los primeros dos lanzamientos resultan en rojo, ¿cuál es la probabilidad de que en el tercer lanzamiento se obtenga un rojo?
 - (c) Si los primeros dos lanzamientos resultan en rojo, ¿cuál es la probabilidad de que se este usando el dado A ?
21. Sean A y B disjuntos y de probabilidad positiva. ¿Pueden A y B ser eventos independientes?
22. Un dado perfecto es lanzado sucesivamente. Halle la probabilidad de obtener al menos un 6 en cuatro lanzamientos.
23. Una muestra de tamaño 4 se extrae, con reposición, de una bolsa que contiene 6 bolas de las cuales 4 son blancas. Sea A = exactamente una de las dos primeras extraídas es blanca y B = la cuarta bola es blanca. (a) ¿Son A y B independientes? ¿Qué sucede si el muestreo se realiza sin reposición? (b) Definamos C = exactamente dos de las bolas extraídas son blancas. ¿Son A , B y C independientes? ¿Son B y C independientes?. Considere ambos muestreos: con y sin reposición.
24. Un fabricante de motos inscribe tres corredores en una carrera. Sea A_i = el i -ésimo corredor finaliza la carrera en alguno de los tres primeros lugares, $i =$

- 1, 2, 3. Si los eventos A_1 , A_2 y A_3 son independientes y $P(A_1) = P(A_2) = P(A_3) = 0,1$; calcule las probabilidades de los siguientes eventos:
- (a) Ninguno de los corredores termina entre los tres primeros.
 - (b) Al menos uno de los corredores termina entre los tres primeros.
 - (c) Al menos dos terminan entre los tres primeros.
 - (d) Todos terminan entre los tres primeros.
25. Sean A y B eventos independientes tales que con probabilidad $1/6$ ocurren simultáneamente y con probabilidad $1/3$ ninguno de ellos ocurre. Halle $P(A)$ y $P(B)$. ¿Están estas probabilidades determinadas de manera única?
26. Sean A y B eventos independientes tales que con probabilidad $1/6$ ocurren simultáneamente y con probabilidad $1/3$ A ocurre y B no ocurre. Halle $P(A)$ y $P(B)$. ¿Están estas probabilidades determinadas de manera única?
27. ¿Cuál es el menor valor de n para el cual la probabilidad de obtener al menos un 6 en una serie de n lanzamientos de un dado perfecto es mayor que $3/4$?
28. Los eventos $A_1, A_2, \dots$ son independientes y $P(A_j) = p$, $j = 1, 2, \dots$. Halle el menor valor de n para el cual $P(\cup_{i=1}^n A_k) \geq p_0$ para un número fijo p_0 .
29. Una máquina consta de 4 componentes conectados en paralelo, de manera que la máquina falla sólo si los cuatro componentes fallan. Supongamos que los componentes son independientes entre sí y sus probabilidades de fallar cuando la máquina es encendida son $0,1$; $0,2$; $0,3$ y $0,4$. ¿Cuál es la probabilidad de que funcione al encenderla?

Capítulo 3

Variables Aleatorias

3.1. Introducción.

En algunas situaciones nos interesa asociar a cada resultado de un espacio muestral un valor numérico. Como veremos en los ejemplos siguientes esto equivale a definir una función sobre los elementos del espacio muestral.

Ejemplos 3.1

En cada ejemplo se considera el espacio de probabilidad $(\Omega, \mathcal{A}, P)$ correspondiente.

1. Una moneda, con probabilidad p de cara, se lanza n veces y se quiere saber la probabilidad de obtener r caras, $r = 0, 1, \dots, n$. Recordemos que en este caso el espacio muestral está dado por

$$\Omega = \{(a_1, a_2, \dots, a_n) : a_i = c \text{ ó } a_i = s; i = 1, \dots, n\};$$

dado que este es un espacio muestral finito, definimos

$$\mathcal{A} = \wp P(\Omega).$$

Si asociamos a los resultados los valores $c = 1$ y $s = 0$ y definimos

$$Y = Y(\omega) = \# \text{ de caras en } n \text{ lanzamientos,}$$

el evento r caras en n lanzamientos se puede expresar como $\{\omega \in \Omega : Y(\omega) = r\}$, donde $Y \in \{0, 1, \dots, n\}$.

Cuando $n = 3$ se tiene $(Y = 3) = \{ccc\}$; $(Y = 2) = \{ccs, csc, scc\}$; $(Y = 1) = \{ssc, scs, css\}$; $(Y = 0) = \{sss\}$.

Observe que $\cup_{k=0}^3 (Y = k) = \Omega$.

2. Se extrae una muestra de n objetos de una línea de producción. Se escribe un posible resultado como $\{e_1, \dots, e_n\}$ donde para cada $i \in \{1, \dots, n\}$

$$e_i = \begin{cases} 0 & \text{si el objeto está en buen estado} \\ 1 & \text{si el objeto es defectuoso} \end{cases} \quad (3.1)$$

Sea

$$X = \# \text{ de objetos defectuosos} = \sum_{i=1}^n e_i.$$

El evento r objetos defectuosos se denota como $(X = r)$ consiste en todos los resultados $\{e_1, \dots, e_n\}$ en los que el número 1 parece r veces.

3. Se lanza un dado dos veces. Se definen: X = resultado del primer lanzamiento, Y = resultado del segundo lanzamiento, W = máximo resultado en los dos lanzamientos y Z = mínimo de los resultados en los dos lanzamientos.
4. En una central telefónica, se plantea la pregunta: ¿Cuántas llamadas pueden esperarse en un período de 5 minutos? Sea X = el número de llamadas en $[0, 5]$. Note que $X \in \{0, 1, 2, \dots\}$.

En cada una de las situaciones anteriores hemos asociado un valor numérico único a cada uno de los elementos del espacio muestral; en algunos casos esto se hace de manera muy natural. Es de importancia el que los eventos, asociados al espacio muestral correspondiente, se pueden expresar en términos de estos valores numéricos. Comenzamos formalizando esta idea en el caso de espacios de probabilidad discretos, lo que permitirá entender más claramente el caso general.

Definición 3.1 Sea $(\Omega, \mathcal{P}(\Omega), P)$ un espacio de probabilidad discreto. Una variable aleatoria o variable aleatoria real es una función

$$X : \Omega \rightarrow \mathbf{R}$$

que a cada $\omega \in \Omega$ le asocia un único valor real $X(\omega)$.

Ejemplos 3.2

En cada ejemplo se considera un espacio muestral finito.

1. En el ejemplo 3.1.1 tenemos $Y((ccc)) = 3$, $Y((csc)) = 2$, $Y((sss)) = 0$, etc. Si asumimos que la moneda es perfectamente simétrica; es decir, $p = \frac{1}{2}$, entonces $P(Y = 0) = 1/8$, $P(Y = 1) = 3/8$, $P(Y = 2) = 3/8$ y $P(Y = 3) = 1/8$. (Verifíquelo)

2. Una urna contiene bolas numeradas del 1 al 20. Se seleccionan 3 al azar sin reemplazamiento. Si apostamos que al menos una de las bolas tiene un número mayor ó igual a 17, ¿cuál es la probabilidad de que ganemos la apuesta? Definamos la variable aleatoria X = el mayor número seleccionado. Entonces X es una v.a. que toma los valores $3, 4, \dots, 20$. Si suponemos que todas las bolas son igualmente probables, entonces se tiene que

$$P(X = i) = \frac{\binom{i-1}{2}}{\binom{20}{3}}, \quad i = 3, 4, \dots, 20.$$

Por lo tanto

$$P(\text{ganar}) = P(X = 20) + P(X = 19) + P(X = 18) + P(X = 17) = 0,508.$$

Definición 3.2 Sea $(\Omega, \mathcal{A}, P)$ un espacio de probabilidad. Una variable aleatoria o variable aleatoria real es una función

$$X : \Omega \rightarrow \mathbf{R}$$

que satisface la propiedad de que para todo intervalo $I \subseteq \mathbf{R}$ el conjunto

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} \in \mathcal{A},$$

es decir, es un evento.

El conjunto $X^{-1}(I)$ se denomina preimágen de I por la función X ; por lo que se dice que la función X es una variable aleatoria si la preimágen de cualquier intervalo es un evento.

Definición 3.3 Una variable aleatoria se denomina **discreta**, si su rango es un conjunto discreto (finito ó numerable). En este caso existe un conjunto $\{x_n\}_{n \geq 1}$ (conjunto de valores de X) tal que $\sum_{n \geq 1} P(X = x_n) = 1$.

En particular, las variables aleatorias definidas sobre espacios de probabilidad discretos son discretas. Sin embargo una variable aleatoria definida sobre un espacio de probabilidad no discreto puede ser discreta.

Observe que las variables consideradas en los ejemplos 3.2 son discretas.

Ejemplos 3.3

1. Se escoge al azar un número en $[0, 1]$. En este caso $\Omega = [0, 1]$ y si se define $X(\omega) = \omega \in [0, 1]$, se tiene que X **no** es una variable aleatoria discreta.

2. Sea $c \in \mathbf{R}$ un número fijo. Defina $X(\omega) = c$ para todo $\omega \in \Omega$, es decir X es constante. Entonces para todo intervalo $I \subseteq \mathbf{R}$ se tiene que

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} = \begin{cases} \Omega & \text{si } c \in I \\ \emptyset & \text{si } c \in I^c \end{cases} \quad (3.2)$$

Por lo tanto

$$P(X^{-1}(I)) = P(\{\omega \in \Omega : X(\omega) \in I\}) = \begin{cases} 1 & \text{si } c \in I \\ 0 & \text{si } c \in I^c \end{cases} \quad (3.3)$$

3. Sea $A \in \mathcal{A}$ un evento dado y

$$X(\omega) = \chi_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \in A^c \end{cases} \quad (3.4)$$

Entonces para todo intervalo $I \subseteq \mathbf{R}$ se tiene que

$$X^{-1}(I) = \{\omega \in \Omega : X(\omega) \in I\} = \begin{cases} \Omega & \text{si } \{0, 1\} \subseteq I \\ \emptyset & \text{si } \{0, 1\} \subseteq I^c \\ A & \text{si } 0 \notin I \text{ y } 1 \in I \\ A^c & \text{si } 0 \in I \text{ y } 1 \notin I \end{cases}$$

Por lo tanto

$$P(X^{-1}(I)) = P(\{\omega \in \Omega : X(\omega) \in I\}) = \begin{cases} 1 & \text{si } \{0, 1\} \subseteq I \\ 0 & \text{si } \{0, 1\} \subseteq I^c \\ P(A) & \text{si } 0 \notin I \text{ y } 1 \in I \\ P(A^c) & \text{si } 0 \in I \text{ y } 1 \notin I \end{cases}$$

4. Si en los dos ejemplos anteriores se considera un espacio muestral $\Omega = [a, b] \subseteq \mathbf{R}$, donde $a < b$, se obtienen variables aleatorias discretas definidas sobre espacios no discretos.

Definición 3.4 Dada una variable aleatoria discreta X cuyo conjunto de valores es un conjunto $\{x_n\}_{n \geq 1}$, la función

$$\begin{aligned} p : \{x_n\}_{n \geq 1} &\longmapsto [0, 1] \\ x_n &\mapsto p(n) = p_n = P(X = x_n) \end{aligned} \quad (3.5)$$

se denomina **función de probabilidad de la variable aleatoria X** ó **función de masa de probabilidad de la variable aleatoria X** .

Nota. Esta función satisface la condición $\sum_{n \geq 1} p_n = \sum_{n \geq 1} P(X = x_n) = 1$.

Proposición 3.1 Si X y Y son variables aleatorias entonces también lo son $X+Y$, $X-Y$, XY y X/Y cuando $Y \neq 0$.

Definición 3.5 Un intervalo o rectángulo en $\mathbf{R}^m$ es un conjunto de la forma

$$\{(x_1, \dots, x_m) : x_1 \in J_1, \dots, x_m \in J_m\} = J_1 \times J_2 \times \dots \times J_m,$$

donde, para $i = 1, \dots, m$, $J_i \subseteq \mathbf{R}$ es un intervalo.

Definición 3.6 Una variable aleatoria vectorial es una función

$$Y : \Omega \rightarrow \mathbf{R}^m$$

tal que la preimágen de cualquier intervalo J de $\mathbf{R}^m$ es un evento.

Proposición 3.2 Sean $X : \Omega \rightarrow \mathbf{R}^m$ y $g : \mathbf{R}^m \rightarrow \mathbf{R}$. Consideremos sobre $\mathbf{R}^m$ la σ -álgebra de Borel $\mathcal{B}$ (la menor σ -álgebra que contiene a todos los intervalos). Si X y g son variables aleatorias, entonces la función definida por $Y = g \circ X$ es también una variable aleatoria.

Demostración. Es suficiente demostrar que la preimágen de cualquier intervalo I es un evento.

$$Y^{-1}(I) = \{\omega : Y \in I\} = \{\omega : g(X(\omega)) \in I\} = \{\omega : X(\omega) \in g^{-1}(I)\} = X^{-1}(g^{-1}(I)) \in \mathcal{A}$$

3.2. Función de Distribución.

Definición 3.7 Si X es una variable aleatoria, se define la función de distribución de X como la función $F_X : \mathbf{R} \rightarrow [0, 1]$ tal que para cada $y \in \mathbf{R}$,

$$F_X(y) = P(X \in (-\infty, y]) = P(\{\omega \in \Omega : X(\omega) \leq y\}) = P(X \leq y).$$

Comentarios.

1. Si no hay ambigüedad, se utiliza la notación F en lugar de F_X .
2. Si X es una variable aleatoria discreta tal que $X(\Omega) = \{x_1, \dots, x_n, \dots\}$, entonces se define

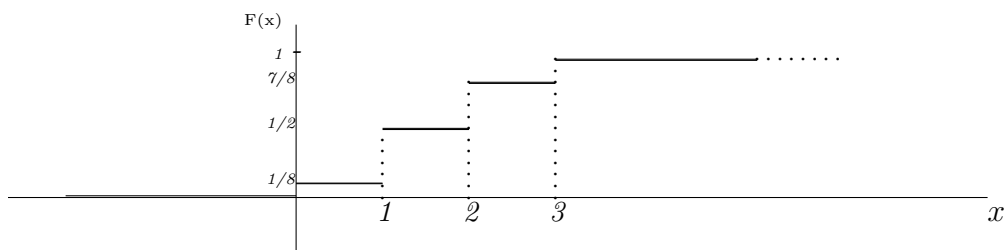
$$F_X(x) = \sum_{x_n \leq x} p_n$$

donde $p_n = P(X = x_n)$.

Ejemplos 3.4

1. Considere la variable aleatoria definida en el ejemplo 3.1.1: $X = \#$ de caras en 3 lanzamientos. En este caso la función de distribución está dada de la siguiente forma:

$$F(0) = \frac{1}{8}; F(1) = \frac{1}{8} + \frac{3}{8} = \frac{1}{2}; F(1,5) = F(1); F(2) = \frac{1}{2} + \frac{3}{8} = \frac{7}{8}; F(3) = 1; F(4) = 1.$$



3.2.1. Propiedades de la Función de Distribución.

Teorema 3.1 Si F es la función de distribución de una variable aleatoria X , entonces:

- a) $0 \leq F(t) \leq 1$ para todo $t \in \mathbf{R}$.
- b) $P(a < X \leq b) = F(b) - F(a)$ para $a < b$.
- c) $F(a) \leq F(b)$ si $a < b$ (no decreciente).

d) $\lim_{t \rightarrow a^+} F(t) = F(a)$ (continua a la derecha)

e) $\lim_{t \rightarrow a^-} F(t) = F(a) - P(X = a)$.

f) $\lim_{t \rightarrow \infty} F(t) = 1$ y $\lim_{t \rightarrow -\infty} F(t) = 0$.

Demostración.

a) Es inmediato ya que por definición $F_X(y) = P(X \leq y) \in [0, 1]$.

b) Escribimos $\{\omega \in \Omega : X(\omega) \leq b\}$ como la unión disjunta

$$\{\omega \in \Omega : X(\omega) \leq b\} = \{\omega \in \Omega : X(\omega) \leq a\} \cup \{\omega \in \Omega : a < X(\omega) \leq b\}.$$

Aplicando la aditividad de la probabilidad obtenemos

$$P(\{\omega \in \Omega : X(\omega) \leq b\}) = P(\{\omega \in \Omega : X(\omega) \leq a\}) + P(\{\omega \in \Omega : a < X(\omega) \leq b\});$$

es decir,

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

c) Se deduce directamente de b) ya que $0 < P(a < X \leq b) = F(b) - F(a)$ para $a < b$.

d) Como F es no decreciente y acotada, el límite existe. Por b) sabemos que

$$F(t) - F(a) = P(a < X \leq t) \text{ para } t > a.$$

Entonces

$$\lim_{t \rightarrow a^+} F(t) = F(a) + \lim_{t \rightarrow a^+} P(a < X \leq t).$$

Veremos que $\lim_{t \rightarrow a^+} P(a < X \leq t) = 0$. Para $n \geq 1$ sea

$$q_n = P(\{\omega \in \Omega : a < X(\omega) \leq a + 1/n\}),$$

entonces

$$\lim_{t \rightarrow a^+} P(a < X \leq t) = \lim_{n \rightarrow \infty} P(a < X \leq a + 1/n) = \lim_{n \rightarrow \infty} q_n.$$

Consideremos ahora la sucesión de eventos

$$S_n = \{\omega \in \Omega : a + \frac{1}{n+1} < X(\omega) \leq a + \frac{1}{n}\},$$

para $n \geq 1$. La sucesión $\{S_n\}_{n \geq 1}$ consta de eventos dos a dos disjuntos tales que:

di. $\{\omega \in \Omega : a < X(\omega) \leq a + 1\} = \cup_{n \geq 1} \{\omega \in \Omega : a + \frac{1}{n+1} < X(\omega) \leq a + \frac{1}{n}\} = \cup_{n \geq 1} S_n$.

dii. $P(S_n) = P(a < X \leq a + \frac{1}{n}) - P(a < X \leq a + \frac{1}{n+1}) = q_n - q_{n+1}$.

diii. $q_1 = P(a < X \leq a+1) = \sum_{n \geq 1} P(S_n) = \sum_{n \geq 1} (q_n - q_{n+1}) = q_1 - \lim_{n \rightarrow \infty} q_n$.
Por lo tanto

$$0 = \lim_{n \rightarrow \infty} q_n = \lim_{t \rightarrow a^+} P(a < X \leq t).$$

e) Por b) tenemos que $F(a) - F(t) = P(t < X \leq a)$; por lo tanto basta demostrar que

$$\lim_{t \rightarrow a^-} P(t < X \leq a) = P(X = a).$$

Puesto que $P(t < X \leq a) = P(t < X < a) + P(X = a)$, será suficiente demostrar que

$$\lim_{t \rightarrow a^-} P(t < X < a) = 0.$$

Para demostrar esto, veremos que para cualquier sucesión creciente $\{y_n\}_{n \geq 1}$ que satisface: $y_n < a$, para todo $n \geq 1$, y $y_n \rightarrow a$ cuando $n \rightarrow \infty$; se cumple que $\lim_{n \rightarrow \infty} P(y_n < X < a) = 0$.

Sean $A_n = \{\omega \in \Omega : y_n < X < a\}$ y $q_n = P(A_n)$. Como la sucesión $\{y_n\}_{n \geq 1}$ es creciente se tiene que, para todo $n \geq 1$, $A_{n+1} \subseteq A_n$ (verifíquelo). Por lo tanto

$$T_n = \{y_n < X \leq y_{n+1}\} = \{y_n < X < a\} - \{y_{n+1} < X < a\} = A_n - A_{n+1},$$

es tal que $P(T_n) = P(A_n) - P(A_{n+1}) = q_n - q_{n+1}$. Además, tenemos que $\{T_n\}_{n \geq 1}$ es una sucesión disjunta de eventos, por lo que

$$\begin{aligned} q_1 &= P(y_1 < X < a) = P(\cup_{n \geq 1} T_n) = \sum_{n \geq 1} P(T_n) = \sum_{n \geq 1} (q_n - q_{n+1}) \\ &= \lim_{n \rightarrow \infty} \sum_{k=1}^n (q_k - q_{k+1}) = q_1 - \lim_{n \rightarrow \infty} q_n \implies \lim_{n \rightarrow \infty} q_n = 0. \end{aligned}$$

f) i) Para demostrar que $\lim_{x \rightarrow -\infty} F(x) = 0$ consideremos cualquier sucesión decreciente $\{x_n\}_{n \geq 1}$ tal que $\lim_{n \rightarrow \infty} x_n = -\infty$. Sea $B_n = \{\omega \in \Omega : X \leq x_n\}$. Por ser $\{x_n\}$ una sucesión decreciente, la sucesión $\{B_n\}$ también es decreciente y es tal que $\cap_{n \geq 1} B_n = \emptyset$ (verifíquelo). Por lo tanto se tiene que

$$\lim_{x \rightarrow -\infty} F(x) = \lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} P(X \leq x_n) = \lim_{n \rightarrow \infty} P(B_n) = P(\cap_{n \geq 1} B_n) = 0.$$

ii) Análogamente, para demostrar que $\lim_{x \rightarrow \infty} F(x) = 1$, consideramos cualquier sucesión creciente $\{x_n\}_{n \geq 1}$ tal que $\lim_{n \rightarrow \infty} x_n = \infty$. Entonces definimos la sucesión de eventos $B_n = \{\omega \in \Omega : X \leq x_n\}$, la cual es creciente y tal que $\cup_{n \geq 1} B_n = \Omega$ (verifíquelo). Se tiene entonces que

$$\lim_{x \rightarrow \infty} F(x) = \lim_{n \rightarrow \infty} F(x_n) = \lim_{n \rightarrow \infty} P(X \leq x_n) = \lim_{n \rightarrow \infty} P(B_n) = P(\cup_{n \geq 1} B_n) = 1.$$

Comentarios.

1. La propiedad e) indica que para todo $a \in \mathbf{R}$, $P(X = a)$ es igual al salto de

la función de distribución F en el punto a . Por lo tanto, cuando a es un punto de continuidad de F , se cumple que $P(X = a) = 0$.

2. Las propiedades c), d) y f) caracterizan a las funciones de distribución. Es decir, si una función G satisface estas propiedades, entonces existe una variable aleatoria Y que tiene a G como función de distribución.

La función de distribución de una variable aleatoria permite calcular las probabilidades de los eventos de interés determinados por la variable aleatoria.

Ejemplos 3.5

1. Supongamos que la función de distribución de una variable aleatoria X está dada por la función

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ \frac{x}{2} & \text{si } 0 \leq x < 1 \\ \frac{2}{3} & \text{si } 1 \leq x < 2 \\ \frac{11}{12} & \text{si } 2 \leq x < 3 \\ 1 & \text{si } 3 \leq x \end{cases}$$

Tenemos entonces que:

a. $P(X < 3) = P(X \leq 3) - P(X = 3) = F(3) - P(X = 3) = 1 - (1 - \frac{11}{12}) = \frac{11}{12}.$

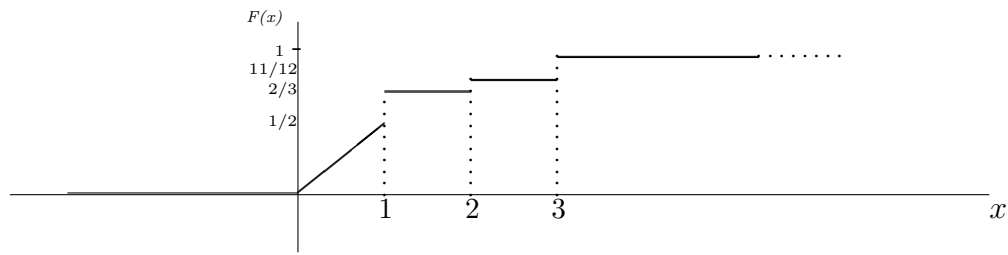
b. $P(X = 1) = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$

c. $P(X > \frac{1}{2}) = 1 - P(x \leq \frac{1}{2}) = 1 - \frac{1}{4} = \frac{3}{4}.$

d. $P(2 < X \leq 4) = F(4) - F(2) = 1 - \frac{11}{12} = \frac{1}{12}.$

Ejercicios 3.1

1. Dos dados diferentes se lanzan dos veces. Encuentre los posibles valores que puede tomar cada una de las variables aleatorias siguientes:

Figura 3.1: Gráfico de $F(x)$.

- a) El máximo valor en los dos lanzamientos.
 - b) El mínimo valor en los dos lanzamientos.
 - c) Suma de los dos lanzamientos.
 - d) Valor del primer lanzamiento menos valor del segundo lanzamiento.
2. Dos bolas se escogen de una urna que contiene 8 bolas blancas, 4 negras y 2 anaranjadas. La escogencia se hace al azar y se ganan dos puntos por cada bola negra y se pierde un punto por cada bola blanca seleccionada. Defina la variable aleatoria X como los puntos obtenidos. Encuentre la función de probabilidad de X , esto es, halle los posibles valores de X y las probabilidades de ocurrencia de estos valores. Sugerencia: Determine el espacio muestral asociado al experimento y halle los posibles valores de X .
 3. Se lanzan 3 dados perfectamente simétricos. Encuentre la función de probabilidad de la variable aleatoria de $X =$ suma de los 3 dados.
 4. Cinco hombres y cinco mujeres son clasificados de acuerdo a los resultados de un examen en el cual no hay dos puntajes iguales y los $10!$ posibles resultados son igualmente probables. Sea X el más alto rango recibido por una mujer ($X = 1$ si una mujer queda en primer lugar, $X = 2$ si un hombre queda en primer lugar y una mujer en el segundo). Encuentre $P(X = i)$, $i = 1, \dots, 10$.
 5. La función de distribución de una variable aleatoria es

$$F(t) = \begin{cases} 0 & \text{si } t < -1 \\ \frac{1}{4} & \text{si } -1 \leq t < -\frac{1}{2} \\ \frac{t}{2} + \frac{1}{2} & \text{si } -\frac{1}{2} \leq t < 0 \\ \frac{t}{4} + \frac{3}{4} & \text{si } 0 \leq t < \frac{1}{2} \\ 1 & \text{si } \frac{1}{2} \leq t \end{cases} \quad (3.6)$$

Halle:

a) La gráfica de F .

b) $P(X = -2)$, $P(X = -1)$, $P(X = \frac{-1}{2})$, $P(X = 0)$.

c) $P(\frac{-1}{2} < X < \frac{1}{4})$, $P(X > \frac{1}{2})$, $P(X \leq \frac{-1}{2})$, $P(X < 0)$.

6. Suponga que la función de distribución de una variable aleatoria X está dada por

$$F(b) = \begin{cases} 0 & \text{si } b < 0 \\ \frac{b}{4} & \text{si } 0 \leq b < 1 \\ \frac{1}{2} + \frac{b-1}{4} & \text{si } 1 \leq b < 2 \\ \frac{11}{12} & \text{si } 2 \leq b < 3 \\ 1 & \text{si } 3 \leq b \end{cases} \quad (3.7)$$

a) Grafique la función F .

b) Encuentre $P(X = i)$, $i = 1, 2, 3$.

c) Encuentre $P(\frac{1}{2} < X < \frac{3}{2})$.

7. Sea X una variable aleatoria con función de distribución

$$F(b) = \begin{cases} 0 & \text{si } b < 0 \\ \frac{1}{2} & \text{si } 0 \leq b < 1 \\ \frac{3}{5} & \text{si } 1 \leq b < 2 \\ \frac{4}{5} & \text{si } 2 \leq b < 3 \\ \frac{9}{10} & \text{si } 3 \leq b < 3,5 \\ 1 & \text{si } 3,5 \leq b \end{cases} \quad (3.8)$$

Halle los posibles valores de X y la función de masa de probabilidad (o función de probabilidad) de X . Grafique la función F .

8. La función de distribución F de una variable aleatoria X viene definida por

$$F(t) = \begin{cases} 0 & \text{si } t < -2 \\ \frac{1}{2} & \text{si } -2 \leq t < 0 \\ \frac{3}{4} & \text{si } 0 \leq t < 2 \\ 1 & \text{si } 2 \leq t \end{cases} \quad (3.9)$$

a) Grafique la función F .

b) Describa la función de masa de probabilidad p (función de probabilidad de X) y represéntela gráficamente.

c) Calcule las siguientes probabilidades: $P(X = 1)$, $P(X \leq 1)$, $P(X < 1)$, $P(X = 2)$, $P(X \leq 2)$, $P(0 < X < 2)$, $P(0 < X \leq 2)$, $P(1 \leq X \leq 2)$.

9. La función de masa de probabilidad p de una variable aleatoria X es nula salvo en los puntos $t = 0, 1, 2$, donde toma los valores $p(0) = 3c^3$, $p(1) = 4c - 10c^2$ y $p(2) = 5c - 1$ para un $c > 0$.

a) Determine el valor de c .

b) Halle y grafique la función de distribución F .

c) Calcule las probabilidades: $P(X < 1)$, $P(X < 2)$, $P(1 < X \leq 2)$ y $P(0 < X < 3)$.

d) Halle el mayor valor de t tal que $F(t) < \frac{1}{2}$.

e) Halle el menor valor de t tal que $F(t) > \frac{1}{3}$.

10. Una variable aleatoria tiene función de distribución continua F , dada por

$$F(t) = \begin{cases} 0 & \text{si } t \leq 0 \\ ct & \text{si } 0 \leq t \leq 1 \\ 1 & \text{si } 1 < t \end{cases} \quad (3.10)$$

a) Determine el valor de c .

b) Calcule las probabilidades: $P(X = \frac{1}{3})$, $P(X < \frac{1}{3})$, $P(|X| < \frac{1}{3})$.

11. Sea X una variable aleatoria con función de distribución

$$F(t) = \begin{cases} 0 & \text{si } t < -3 \\ \frac{1}{8} & \text{si } -3 \leq t < -2 \\ \frac{3}{16} & \text{si } -2 \leq t < -1,5 \\ \frac{5}{16} & \text{si } -1,5 \leq t < 0 \\ \frac{9}{16} & \text{si } 0 \leq t < 1 \\ \frac{10}{16} & \text{si } 1 \leq t < 2,5 \\ \frac{14}{16} & \text{si } 2,5 \leq t < 3 \\ 1 & \text{si } 3 \leq t \end{cases} \quad (3.11)$$

i) Grafique F .

ii) Calcule $P(X = i)$ para $i \in \{-3, -2, -\frac{3}{2}, -1, 0, 1, \frac{3}{2}, 2, \frac{5}{2}, 3, \frac{7}{2}, 4\}$; $P(X \leq 2)$; $P(-1 < X \leq 3)$; $P(X > 1,5)$; $P(X < 0)$.

iii) Halle la función de probabilidad de X .

12. Sea X una variable aleatoria con función de distribución

$$F(t) = \begin{cases} 0 & \text{si } t < -2 \\ \frac{1}{5} & \text{si } -2 \leq t < 0 \\ \frac{3t^2}{10} + \frac{1}{5} & \text{si } 0 \leq t < 1 \\ \frac{2}{3} & \text{si } 1 \leq t < 2 \\ \frac{t}{6} + \frac{1}{3} & \text{si } 2 \leq t < 3 \\ 1 & \text{si } 3 \leq t \end{cases} \quad (3.12)$$

Grafique F y calcule $P(X = i)$ para $i \in \{-3, -2, -1, 0, \frac{1}{2}, 1, 2, 3, 4\}$; $P(X = 1)$; $P(0 \leq X \leq 1)$; $P(X \leq 2,5)$; $P(X < -1)$ y $P(X > 4)$.

13. Sea X una variable aleatoria con función de distribución

$$F(t) = \begin{cases} 0 & \text{si } t < -1 \\ \frac{1}{6} & \text{si } -1 \leq t < 0,5 \\ \frac{1}{2} & \text{si } 0,5 \leq t < 2 \\ \frac{2}{3} & \text{si } 2 \leq t < 3 \\ \frac{5}{6} & \text{si } 3 \leq t < 4 \\ 1 & \text{si } 4 \leq t \end{cases} \quad (3.13)$$

i) Grafique F .

ii) Calcule $P(X = i)$ para $i \in \{-3, -2, -\frac{3}{2}, -1, 0, 1, \frac{1}{2}, 2, \frac{5}{2}, 3, \frac{7}{2}, 4\}$; $P(X \geq 2,5)$; $P(-1 < X \leq 3)$; $P(X > 0)$; $P(X < 3)$.

iii) Halle la función de probabilidad de X .

Comentario. Como el dominio de una variable aleatoria es un espacio muestral, el valor que toma la variable no se puede predecir con exactitud. Por esto estaremos interesados en conocer, más que los valores de la variable, la probabilidad de que el valor que toma la variable aleatoria pertenezca a un subconjunto de $\mathbb{R}$ dado. Esto se conocerá si se conoce la función de distribución de la variable.

3.3. Algunas Distribuciones Importantes.

1. Variable Aleatoria de Bernoulli.

Un experimento que tiene dos posibles resultados se denomina **ensayo de Bernoulli**. Identifiquemos los resultados de un ensayo de Bernoulli como *éxito* y *fracaso*; y sea $p \in (0, 1)$ la probabilidad de obtener éxito en una realización del experimento. Si definimos la v.a. $X = 1$ cuando el resultado es éxito y $X = 0$ cuando el resultado es fracaso tenemos que X tiene función de probabilidad dada por

$$\begin{aligned} P(X = 1) &= p \\ P(X = 0) &= 1 - p. \end{aligned} \quad (3.14)$$

Una variable aleatoria discreta con función de probabilidad anterior se denomina **variable aleatoria de Bernoulli**.

2. Variable Aleatoria Binomial.

Supongamos que se realizan n ensayos de Bernoulli independientes y con probabilidad de éxito p . La variable aleatoria definida como $Y = \#$ de éxitos en los n

ensayos, se denomina **variable aleatoria binomial con parámetros n y p** . Esto también se expresa mediante la notación $Y \sim b(n, p)$. El conjunto de valores de Y es $\{0, 1, \dots, n\}$ y su distribución está dada por

$$p_i = b(i; n, p) = P(Y = i) = \binom{n}{i} p^i (1-p)^{n-i}, \quad i = 0, 1, \dots, n.$$

La función dada por $\{p_i\}_{i=1}^n$ es una función de probabilidad ya que

$$\sum_{i=0}^n p_i = \sum_{i=1}^n \binom{n}{i} p^i (1-p)^{n-i} = (p + 1 - p)^n = 1.$$

Ejemplos 3.6

1. Se lanzan 5 monedas perfectas. Sea $X = \#$ de caras, entonces X es una variable aleatoria binomial con parámetros $n = 5$ y $p = 0,5$. Es decir $X \sim b(5, 0,5)$ con función de probabilidad dada por $P(X = i) = \binom{5}{i} (0,5)^i (0,5)^{n-i}$. Los valores de esta función se dan en la siguiente tabla:

<i>i</i>	<i>0</i>	<i>1</i>	<i>2</i>	<i>3</i>	<i>4</i>	<i>5</i>
$P(X = i)$	$1/32$	$5/32$	$10/32$	$10/32$	$5/32$	$1/32$

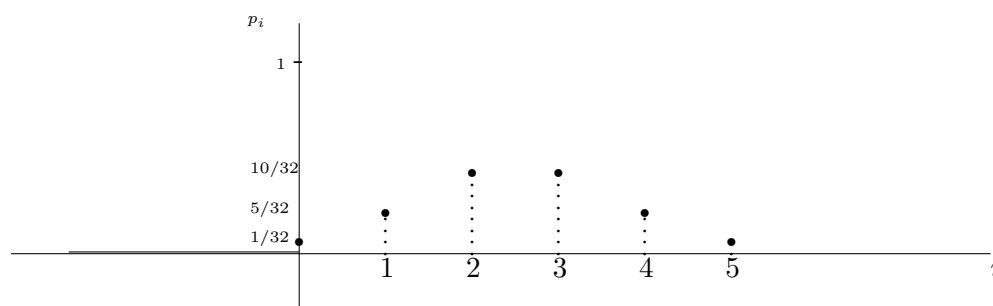


Figura 3.2: Función de probabilidad de X .

2. Se sabe que los tornillos que produce una fábrica son defectuosos con probabilidad 0.01, independientemente entre si. Se venden los tornillos en paquetes de 10 y se ofrece garantía de que a lo sumo 1 de los tornillos es defectuoso. ¿Qué proporción de paquetes serán devueltos?

Sea $X = \#$ defectuosos en un paquete. Entonces $X \sim b(10, 0,01)$ y la probabilidad de que un paquete sea devuelto es

$$\begin{aligned} P(X > 1) &= 1 - P(X \leq 1) = 1 - P(X = 0) - P(X = 1) \\ &= 1 - \binom{10}{0}(0,01)^0(0,99)^{10} - \binom{10}{1}(0,01)^1(0,99)^9 = 0,004. \end{aligned}$$

Por lo tanto la proporción de cajas devueltas es 0.4 %.

3. Variable Aleatoria Geométrica.

Suponga que ensayos de Bernoulli, independientes y con probabilidad $p \in (0, 1)$ de éxito, se realizan hasta que se obtiene un éxito. La variable aleatoria $X = \#$ de ensayos requeridos para obtener un éxito se denomina **variable aleatoria geométrica**. El conjunto de valores de X es $\{1, 2, \dots\}$ y su función de probabilidad está dada por

$$p_n = P(X = n) = (1 - p)^{n-1}p; \quad n = 1, 2, \dots,$$

que ciertamente satisface

$$\sum_{k=1}^{\infty} p_k = \sum_{k=1}^{\infty} (1 - p)^{k-1}p = \sum_{k=0}^{\infty} (1 - p)^k p = \frac{p}{1 - (1 - p)} = 1.$$

Ejemplos 3.7

1. Una urna contiene N bolas blancas y M bolas negras. Se seleccionan bolas aleatoriamente, una a la vez y con reposición, hasta que una bola negra es obtenida. ¿Cuál es la probabilidad de que se necesiten exactamente n extracciones? ¿Cuál es la probabilidad de que se necesiten al menos k extracciones?

Sea $X = \#$ extracciones necesarias para obtener una bola negra.

La v.a. X es geométrica con probabilidad de éxito $p = \frac{M}{M+N}$. Observe que la independencia entre los ensayos se obtiene al hacer la extracción con reposición. Se desea calcular

$$i) P(X = n) = \frac{M}{M + N} \left(\frac{N}{M + N} \right)^{n-1} = \frac{MN^{n-1}}{(M + N)^n}.$$

$$ii) P(X \geq k) = \frac{M}{M + N} \sum_{n=k}^{\infty} \left(\frac{N}{M + N} \right)^{n-1} = \left(\frac{N}{M + N} \right)^{k-1}. \text{ Demuéstrelo.}$$

4. Variable Aleatoria de Poisson.

Sea X una variable aleatoria que tiene el conjunto de valores $\{0, 1, 2, \dots\}$. Se dice que X es una **variable aleatoria de Poisson con parámetro** λ , $\lambda \in (0, \infty)$, si

la función de probabilidad de X está dada por

$$p_i = P_o(i) = P(X = i) = e^{-\lambda} \frac{\lambda^i}{i!}, \quad i = 0, 1, 2, \dots$$

La función P_o define una medida de probabilidad ya que

$$\sum_{k=0}^{\infty} P_o(k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \times e^{\lambda} = 1.$$

En este caso se denota $X \sim \mathcal{P}(\lambda)$.

Aplicaciones de la V.A. de Poisson.

La distribución de Poisson se utiliza, entre otras cosas para modelar situaciones en las que eventos ocurren un cierto número de veces en intervalos (usualmente de tiempos) dados. El parámetro λ se interpreta usualmente como *el número de ocurrencias del evento por unidad de tiempo* o como *el valor promedio de la variable aleatoria*, en un sentido que será aclarado posteriormente.

Ejemplos 3.8

1. Suponga que el número de errores tipográficos en una página de un libro es una v.a. de Poisson con parámetro $\lambda = 1/2$. Si $X = \#$ errores por página, entonces la probabilidad de que haya al menos un error en una página es

$$P(X \geq 1) = 1 - P(X < 1) = 1 - P(X = 0) = 1 - e^{-\lambda} \frac{\lambda^0}{0!} = 1 - e^{-1/2} = 0,393.$$

Otras situaciones en que se usa la v.a. de Poisson.

P1. $X = \#$ de llamadas telefónicas que llega a una central en un período de tiempo.

$\lambda = \#$ llamadas por unidad de tiempo.

P2. $X = \#$ clientes que entra en un banco en un día. $\lambda = \#$ promedio de clientes que entra al banco en un día

P3. $X = \#$ de personas en una comunidad que vive más de 100 años.

P4. La distribución de Poisson puede utilizarse en algunos casos para dar una aproximación de una distribución binomial. Sea $X \sim b(n, p)$ tal que n es un valor grande y p es pequeño. Definamos $\lambda > 0$ como $\lambda = p \times n$, entonces

$$P(X = k) = \binom{n}{k} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n!}{k!(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$\begin{aligned}
&= \frac{(n-k+1)(n-k+2)\dots(n-1)n}{k!} \left(\frac{\lambda^k}{n^k}\right) \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k} \\
&= \frac{(n-k+1)(n-k+2)\dots(n-1)n}{n^k} \left(\frac{\lambda^k}{k!}\right) \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k} \\
&= \frac{(n-k+1)}{n} \frac{(n-k+2)}{n} \dots \frac{(n-1)}{n} \frac{n}{n} \left(\frac{\lambda^k}{k!}\right) \frac{(1-\frac{\lambda}{n})^n}{(1-\frac{\lambda}{n})^k}
\end{aligned}$$

Tomando $\lim_{n \rightarrow \infty}$ tenemos que

$$P(X = k) \rightarrow 1 \times 1 \times \dots \times 1 \times \left(\frac{\lambda^k}{k!}\right) \times \frac{\lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^n}{\lim_{n \rightarrow \infty} (1 - \frac{\lambda}{n})^k} = \left(\frac{\lambda^k}{k!}\right) e^{-\lambda}.$$

Es decir que para n **suficientemente grande** la distribución binomial $b(n, p)$ puede ser aproximada por una distribución de Poisson con parámetro $\lambda = np$.

Ejemplos 3.9

1. Se realiza un experimento que consiste en contar el número de partículas α irradiadas, en un intervalo de 1 segundo, por cada gramo de un material radioactivo. Si la experiencia indica que, en promedio, 3.2 partículas α son irradiadas, ¿cuál es la probabilidad de que no aparezcan más de 2 partículas(en 1 segundo)?

Sea $X = \#$ de partículas de material radioactivo descargadas en un segundo.

Se puede pensar en un gramo de material radioactivo como un gran conjunto de n átomos, cada uno de los cuales tiene probabilidad p de desintegrarse durante el período considerado, emitiendo una partícula α . Entonces X tiene una distribución binomial con parámetros (n, p) donde n es un número muy grande y $p = 3,2/n$, ya que todos los átomos tienen igual probabilidad de descargar una partícula. Por consiguiente el cálculo de las probabilidades p_i pueden ser engorroso, lo que hace preciso considerar $b(n, p) \simeq \mathcal{P}(np) = \mathcal{P}(3,2)$. De esta forma tenemos

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = e^{-3,2} \left(1 + 3,2 + \frac{(3,2)^2}{2}\right) = 0,382.$$

P5. Número de eventos que ocurre en un intervalo de longitud t . Supongamos que un evento dado ocurre en ciertos puntos aleatorios de tiempo de modo que, para algún $\lambda > 0$ se cumplen:

Hi. La probabilidad de que ocurra exactamente un evento en un intervalo de longitud

h es $\lambda h + o(h)$; donde $o(h)$ es un valor pequeño en relación con h . De manera más precisa, $o(h)$ es una función tal que

$$\lim_{h \rightarrow 0} \frac{o(h)}{h} = 0.$$

Hii. La probabilidad de que 2 ó más eventos ocurran en un intervalo de longitud h es $o(h)$.

Hiii. Dados enteros $n, j_1, j_2, \dots, j_n$ y cualquier conjunto de intervalos disjuntos $I_1, \dots, I_n$; si se definen los eventos $E_i = \text{ocurren } j_i \text{ eventos en el intervalo } I_i$ $i = 1, \dots, n$, entonces estos eventos son independientes.

Proposición 3.3 *Supongamos que se cumplen las hipótesis Hi, Hii y Hiii. Sea $N(t) = \#$ de eventos en un intervalo de longitud t , entonces $N(t)$ es una variable aleatoria de Poisson con parámetro λt ; es decir,*

$$P(N(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}, \quad k = 0, 1, 2, \dots$$

Ejemplos 3.10

1. En una región de California, USA, ocurren terremotos de acuerdo a las hipótesis Hi, Hii y Hiii, a una tasa de 2 por semana. Tomando una semana como unidad de tiempo:

i. Encuentre la probabilidad de que al menos 3 terremotos ocurran durante las dos semanas siguientes.

ii. Encuentre la distribución de probabilidad de el tiempo, comenzando desde ahora, hasta el próximo terremoto.

Puesto que λ se interpreta como el número de ocurrencias por unidad de tiempo, tenemos que $\lambda = 2$. Sea $X(t) = \#$ de terremotos en t semanas. Entonces queremos calcular i.

$$\begin{aligned} P(X(2) \geq 3) &= 1 - P(X(2) = 0) - P(X(2) = 1) - P(X(2) = 2) \\ &= 1 - e^{-4}(1 + 4 + 8) = 1 - 13e^{-4}. \end{aligned}$$

ii. Sea $T =$ tiempo hasta el próximo terremoto (medido en semanas). Entonces

$$\begin{aligned} P(T > t) &= P(X(t) = 0) = e^{-\lambda t} \implies \\ F(t) = P(T \leq t) &= 1 - P(T > t) = 1 - e^{-\lambda t} = 1 - e^{-2t}. \end{aligned}$$

5. Variable Aleatoria Hipergeométrica.

Sea $p \in (0, 1)$ dado y suponga que se va a seleccionar, aleatoriamente y sin reposición, una muestra de tamaño n de una urna que contiene N bolas, de las cuales Np

son blancas y $N - Np$ son negras. Si $X = \#$ de bolas blancas seleccionadas, entonces X tiene función de probabilidad dada por

$$p_k = P(X = k) = \frac{\binom{Np}{k} \binom{N-Np}{n-k}}{\binom{N}{n}}, \quad k = 0, 1, \dots, \min(n, Np).$$

En este caso se dice que X es una **variable aleatoria hipergeométrica**.

Ejemplos 3.11

Un comprador de componentes electrónicos los compra en lotes de 10. Su política es inspeccionar 3 componentes y aceptar el lote sólo si los 3 son no defectuosos. Si 30 % de los lotes tienen 4 componentes defectuosas y 70 % sólo 1, ¿qué proporción de lotes rechaza el comprador?

Sean $X = \#$ de componentes defectuosas seleccionadas;

A = el comprador acepta el lote;

B = el lote tiene 4 defectuosas.

Entonces, $A = \{X = 0\}$. Escribiendo $A = A \cap (B \cup B^c)$ se tiene que

$$\begin{aligned} P(X = 0) = P(A) &= P(A \cap B) + P(A \cap B^c) = P(B)P(A|B) + P(B^c)P(A|B^c) \\ &= \frac{3}{10} \frac{\binom{4}{0} \binom{6}{3}}{\binom{10}{3}} + \frac{7}{10} \frac{\binom{1}{0} \binom{9}{3}}{\binom{10}{3}} = \frac{54}{100}. \end{aligned}$$

Por lo que se rechaza el 46 % de los lotes.

Ejercicios 3.2

1. Sea $b(k; n, \frac{\lambda}{n}) = \binom{n}{k} (\frac{\lambda}{n})^k (1 - \frac{\lambda}{n})^{n-k}$ la función de probabilidad de una variable aleatoria binomial con parámetros $(n, \frac{\lambda}{n})$.

i) Demuestre que $\lim_{n \rightarrow \infty} \frac{b(k; n, \frac{\lambda}{n})}{b(k-1; n, \frac{\lambda}{n})} = \frac{\lambda}{k}$, para $k = 0, 1, \dots$

ii) Demuestre que $\lim_{n \rightarrow \infty} b(k; n, \frac{\lambda}{n}) = e^{-\lambda} \frac{\lambda^k}{k!}$.

Sugerencia: Use inducción sobre k y la parte i).

2. Si X es una variable aleatoria binomial con parámetros (n, p) , $0 < p < 1$, demuestre que $f(k) = P(X = k)$ crece al principio monotonamente y luego decrece monotonamente alcanzando su máximo cuando $k = [(n + 1)p]$ donde $[x]$ = parte entera de x . Sugerencia: considere los cocientes:

$$\frac{P(X = k)}{P(X = k - 1)} = \frac{b(k; n, p)}{b(k - 1; n, p)}.$$

3. Considere el siguiente juego: Un apostador apuesta a un número entre 1 y 6. Tres dados son lanzados; si el número al que apostó el jugador aparece i veces, $i = 1, 2, 3$, el jugador gana i unidades, si el número no aparece el jugador pierde una unidad. Halle la distribución de probabilidad de la ganancia del jugador. Si se asume que el juego continua indefinidamente, ¿qué ganancia acumularía el jugador?
4. Suponga que una característica particular de una persona, como el color de ojos o el ser zurdo, se clasifica basandose en un par de genes. Suponga que d representa un gen dominante y r un gen recesivo. Un híbrido es una persona con genes rd . Un niño recibe un gen de cada progenitor. Si, con respecto a una característica particular, dos padres híbridos tienen 4 niños, ¿cuál es la probabilidad de que 3 de los 4 niños tengan la apariencia del gen dominante? Nota: la apariencia de dd es igual a la apariencia de un híbrido
5. Se extrae una bola de una urna que contiene 3 bolas negras y 3 bolas blancas. Se reemplaza la bola extraída y se extrae otra bola. Se prosigue así indefinidamente. ¿Cuál es la probabilidad de que de las 4 primeras bolas extraídas, exactamente 2 sean blancas?
6. En un examen de selección múltiple, con 3 posibles respuestas para cada una de 5 preguntas, ¿cuál es la probabilidad de que un estudiante responda al menos 4 preguntas correctamente si selecciona sus respuestas al azar?
7. Un hombre afirma tener percepción extrasensorial. Para verificarlo, se lanza una moneda insesgada (perfecta) 10 veces y se le pide al hombre predecir el resultado de cada lanzamiento. Si él obtiene 7 aciertos, ¿cuál es la probabilidad de que una persona sin poder extrasensorial obtenga el mismo número de aciertos?
8. Un canal de comunicación transmite dígitos 0 y 1. Debido a estática, el dígito se transmite incorrectamente con probabilidad 0.2. Se quiere transmitir un mensaje importante consistente en uno de los dígitos 0 ó 1. Para reducir la posibilidad de error, se transmite 00000 en lugar de 0 y 11111 en lugar de 1. El que recibe identifica el mensaje como el dígito que aparece un mayor número de veces. ¿Cuál es la probabilidad de que se reciba el mensaje equivocado? Explique qué suposiciones relativas a independencia está haciendo.

9. *Un sistema de satélite consiste de n componentes y funciona en un día determinado si al menos k de las n componentes funcionan. En un día lluvioso, cada una de las n componentes funciona, independientemente, con probabilidad p_1 , mientras que en un día seco cada una de ellas funciona, independientemente, con probabilidad p_2 . Si la probabilidad de lluvia mañana es α , ¿cuál es la probabilidad de que el sistema funcione mañana?*
10. *Suponga que el número de accidentes que ocurre en una autopista cada día es una variable aleatoria de Poisson con parámetro $\lambda = 3$.*
 - i) *Encuentre la probabilidad de que 3 ó más accidentes se produzcan hoy.*
 - ii) *Encuentre la probabilidad de que al menos un accidente se produzca hoy.*
11. *Compare la aproximación de Poisson con la probabilidad binomial correcta en cada uno de los casos siguientes:*
 - i) *$P(X = 2)$ cuando $n = 8$ y $p = 0,1$.*
 - ii) *$P(X = 9)$ cuando $n = 10$ y $p = 0,95$.*
 - iii) *$P(X = 0)$ cuando $n = 10$ y $p = 0,1$.*
 - iv) *$P(X = 4)$ cuando $n = 9$ y $p = 0,2$.*
12. *Si usted compra un ticket de lotería en cada una de 50 loterías diferentes, siendo en cada una la probabilidad de ganar igual a $\frac{1}{100}$, halle la probabilidad (aproximada) de que usted ganará un premio: i) al menos una vez; ii) exactamente una vez; c) al menos dos veces.*
13. *La gente entra a un casino a una rata de 1 cada 2 minutos. Halle la probabilidad de que: i) nadie entre entre 12:00 y 12:05; ii) al menos 4 personas entren en el casino durante ese tiempo.*
14. *El número de llamadas recibidas en una oficina entre 9:00 a.m. y 10:00 a.m. es, en promedio, alrededor de 30. Use la distribución de Poisson para encontrar la probabilidad de que en un día típico no haya ninguna llamada entre 9:45 a.m. y 10:00 a.m. Sugerencia: Observe que $\lambda = 30$ y tome 1 hora como unidad de tiempo.*
15. *Los clientes llegan a un banco de acuerdo con una distribución de Poisson con parámetro 40 (basado en una hora como unidad de tiempo). Dado que 30 clientes llegaron en la primera hora, encuentre la probabilidad de que 60 clientes lleguen durante la primera hora y media.*
16. *Sea X una variable aleatoria de Poisson con parámetro λ . Demuestre que la función de probabilidad de X crece monótonamente y luego decrece monótonamente, alcanzando su máximo valor en $[\lambda]$.*
17. *Halle la probabilidad de que en una compañía en la que trabajan 500 personas exactamente k de ellas cumplan año el mismo día.*

18. Suponga que en una comunidad de 20000 personas se ha determinado que la probabilidad de vivir 100 años es igual a 0.002. Halle la probabilidad (aproximada) de que haya 10 centenarios.
19. Una población de N animales habita en cierta región. El siguiente experimento es realizado: primero se captura un grupo de r animales, se les marca y son devueltos a su medio. Una vez que se han mezclado de nuevo se captura un grupo de n animales. Si todos los animales tienen igual probabilidad de ser capturados, halle la probabilidad de que el número de animales marcados en la segunda captura sea igual a k . ¿Cuáles son los posibles valores de k ?
20. En el proceso de control de calidad, una industria somete a inspección un lote de 10000 piezas de la siguiente forma: se selecciona una muestra de 1500 piezas y se examinan en busca de piezas defectuosas. Si existen en el lote un total de 850 piezas defectuosas, ¿cuál es la probabilidad de que se encuentren 300 piezas defectuosas en la muestra seleccionada?
21. Se lanzan 3 monedas idénticas y se sabe que estas tienen probabilidad de cara igual a $1/3$. Halle la función de probabilidad del número de caras.
22. El número de personas que utiliza un determinado cajero automático en un día, es una variable aleatoria de Poisson con parámetro $\lambda = 10$. a) Halle la probabilidad de que el cajero no sea utilizado en el día de hoy. b) Halle la probabilidad de que al menos dos personas utilicen el cajero en el día de mañana.
23. Una extraña enfermedad aparece en el 5 % de los miembros de una comunidad que consta de 15000 individuos. Calcule, aproximadamente, la probabilidad de que 10 personas sufran la enfermedad.
24. El número de veces que un individuo se resfria en un año dado es una variable aleatoria de Poisson con parámetro $\lambda = 5$. Suponga que un nuevo medicamento reduce el parámetro de Poisson a $\lambda = 3$ en un 75 % de la población. El otro 25 % no se beneficia de la droga. Si una persona usa la droga por un año y tiene dos resfriados en ese tiempo, ¿cuál es la probabilidad de que esta persona sea uno de los que se benefician de la droga ?
25. Un examen de selección múltiple tiene 10 preguntas, cada una de las cuales tiene 5 respuestas posibles. Un estudiante selecciona las respuestas al azar.
 - i) ¿Cuál es la distribución del número de respuestas correctas dadas por el estudiante?.
 - ii) ¿Cuál es la probabilidad de que el estudiante responda al menos 9 preguntas correctamente?
26. El número de inundaciones que se producen anualmente en una localidad es una variable aleatoria de Poisson. Si, en promedio, ocurren 3 inundaciones por año,

halle la probabilidad de que la próxima inundación se produzca en los próximos 6 meses.

27. Un niño practica para un partido de basketball. Si su probabilidad de encestar es de 0,3, ¿cuál es la probabilidad de que tenga que lanzar la pelota 5 veces antes de anotar un punto?
28. En un juego de dados, se gana una moneda al acertar el valor obtenido al lanzar el dado, de lo contrario se pierde una moneda. Si un apostador sólo puede apostar uno de los números ¿cuál es la probabilidad de que tenga que apostar 7 veces para ganar?
29. Una persona tiene cuatro llaves indistinguibles en una bolsa. Si sólo una de ellas abre la puerta que tiene al frente y la persona devuelve a la bolsa cada llave que prueba ¿cuál es la probabilidad de que tenga que probar seis veces para abrir la puerta?

3.4. Densidades. Variables Aleatorias Continuas.

Definición 3.8 Una función $f : \mathbf{R} \rightarrow \mathbf{R}$ es una **función de densidad** sobre $\mathbf{R}$ si y sólo si f satisface las condiciones siguientes:

i) Para todo $x \in \mathbf{R}$, $f(x) \geq 0$.

ii) $\int_{-\infty}^{\infty} f(u)du = 1$.

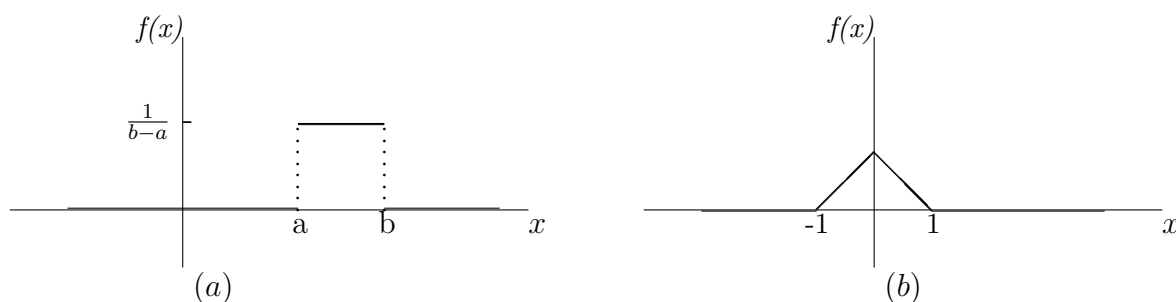


Figura 3.3: Gráficos de funciones de densidad.

Definición 3.9 Una variable aleatoria X es **continua ó absolutamente continua** si existe una función de densidad f tal que para todo $a \in \mathbf{R}$

$$F(a) = P(X \leq a) = \int_{-\infty}^a f(u)du.$$

En este caso f se denomina **función de densidad de probabilidad de X** .

Comentarios.

1. La función de densidad permite responder a cualquier pregunta de carácter probabilístico en relación con la variable aleatoria.
2. De la definición anterior se deduce que la función de distribución F , de una variable aleatoria continua, es continua; por lo que, para todo $x \in \mathbf{R}$, $P(X = x) = 0$.
3. Para cualquier par de números reales $a < b$ se tiene que

$$P(a \leq X \leq b) = P(a < X \leq b) = P(a < X < b) = P(a \leq X < b) = \int_a^b f(u)du.$$

4. Para todo conjunto de Borel B se cumple que

$$P(X \in B) = \int_B f(u)du.$$

Por lo tanto $P(X \in B)$ representa geométricamente el área bajo la curva de f que se genera sobre el conjunto B .

5. Por el teorema fundamental del cálculo la función de distribución F es derivable en cada punto y de continuidad de la función de densidad f . En cada uno de estos puntos se cumple que $F'(y) = f(y)$, es decir que F es la primitiva de f .
6. La función de probabilidad es el concepto análogo a la función de densidad en el caso discreto.

<i>Variable Aleatoria Discreta</i>	<i>Variable Aleatoria Continua</i>
$P(X \leq t) = \sum_{x_n \leq t} p_n$	$P(X \leq t) = \int_{-\infty}^t f(u)du$
$\sum_{n=0}^{\infty} p_n = 1$	$\int_{-\infty}^{\infty} f(u)du = 1$

Ejemplos 3.12

1. Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} c(4x - 2x^2) & \text{si } 0 < x < 2 \\ 0 & \text{si } x \notin (0, 2) \end{cases} \quad (3.15)$$

- i) Halle el valor de c .
- ii) Halle $P(X > 1)$.

Solución: i) El valor c debe satisfacer $c \int_0^2 (4x - 2x^2)dx = 1$. Entonces

$$c \left(4 \frac{x^2}{2} - 2 \frac{x^3}{3} \right) \Big|_{x=0}^{x=2} = c \left(2 \times 4 - \frac{2}{3} \times 8 \right) = c \times \frac{8}{3} \implies c = \frac{3}{8}.$$

$$ii) P(X > 1) = \int_1^{\infty} f(u)du = \frac{3}{8} \int_1^2 (4u - 2u^2)du = \frac{1}{2}.$$

2. La cantidad de tiempo, en horas, que una computadora funciona antes de dañarse es una variable aleatoria continua con función de densidad de probabilidad

$$f(x) = \begin{cases} \lambda e^{-\frac{x}{100}} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (3.16)$$

i) ¿Cuál es la probabilidad de que una computadora funcione entre 50 y 150 horas antes de descomponerse?

ii) ¿Cuál es la probabilidad de que funcione menos de 100 horas?

Debemos en primer lugar determinar el valor de λ .

$$1 = \lambda \int_0^{\infty} e^{-\frac{x}{100}} dx = \lambda (-100e^{-\frac{x}{100}}) \Big|_0^{\infty} = \lambda \times 100 \implies \lambda = \frac{1}{100}.$$

$$i) P(50 \leq X \leq 150) = \frac{1}{100} \int_{50}^{150} e^{-\frac{x}{100}} dx = \frac{1}{100} \times -100e^{-\frac{x}{100}} \Big|_{50}^{150} = e^{-\frac{1}{2}} - e^{-\frac{3}{2}} = 0,384.$$

$$ii) P(X < 100) = \frac{1}{100} \int_0^{100} e^{-\frac{x}{100}} dx = \frac{1}{100} \times -100e^{-\frac{x}{100}} \Big|_0^{100} = -e^{-1} + 1 = 0,633.$$

3. Suponga que la función de densidad de una variable aleatoria X esta dada por

$$f(x) = \begin{cases} \alpha & \text{si } x \in [0, 100] \\ 0 & \text{en otros casos.} \end{cases}$$

Halle $P(-2 \leq X \leq 3)$.

Solución: En primer lugar observamos que

$$\int_0^{100} \alpha = 1 \implies \alpha = \frac{1}{100},$$

por lo que

$$P(-2 \leq X \leq 3) = \int_{-2}^3 f(x)dx = \int_0^3 \alpha dx = \int_0^3 \frac{1}{100} dx = \frac{3}{100}.$$

4. Sea X una variable aleatoria continua con función de densidad

$$f(x) = \begin{cases} \alpha(x - 2x^2) & \text{si } 0 < x < \frac{1}{2} \\ 0 & \text{en otros casos,} \end{cases}$$

Determinamos el valor de α :

$$\alpha \int_0^{\frac{1}{2}} (x - 2x^2) dx = \alpha \times \frac{1}{24} \Rightarrow \alpha = 24.$$

Entonces tenemos que

$$P(X > \frac{1}{4}) = 24 \int_{\frac{1}{4}}^{\frac{1}{2}} (x - 2x^2) dx = 24 \times \frac{1}{48} = \frac{1}{2}.$$

3.4.1. Distribuciones de Algunas Variables Aleatorias Continuas.

1. Variable Aleatoria Uniforme.

Una variable aleatoria X se dice que tiene distribución **uniforme sobre** $[a, b]$ o que está **uniformemente distribuida sobre** $[a, b]$ si su función de densidad de probabilidad está dada por

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{si } x \notin [a, b], \end{cases}$$

(figura 2.3 (a)). En este caso la función de distribución está dada por

$$F(t) = \begin{cases} 0 & \text{si } t < a \\ \frac{t-a}{b-a} & \text{si } t \in [a, b] \\ 1 & \text{si } t \geq b \end{cases}$$

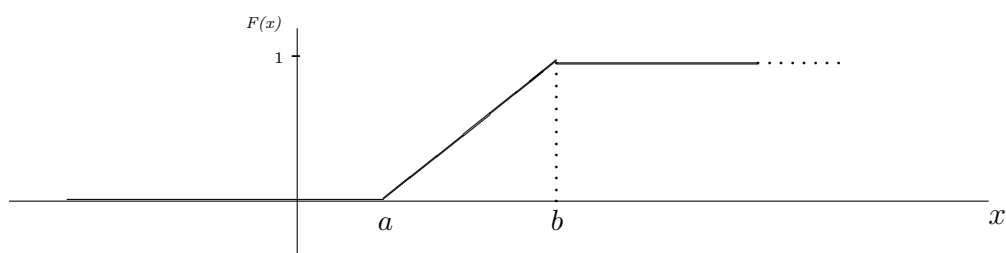


Figura 3.4: Gráfico de $F(x)$.

Ejemplos 3.13

1. Sea X una variable aleatoria continua uniformemente distribuida sobre $[0, 10]$. Entonces

$$P(3 < X < 8) = \int_3^8 \frac{1}{10} dx = \frac{5}{10} = \frac{1}{2};$$

$$P(X > 6) = \int_6^{10} \frac{1}{10} dx = \frac{4}{10} = \frac{2}{5};$$

$$P(X > 11) = 0 \quad y \quad P(X < 11) = 1.$$

2. Variable Aleatoria Normal.

Una variable aleatoria X se dice que tiene distribución **normal** o que está **normalmente distribuida**, con parámetros μ y σ^2 ($\sigma > 0$), si su función de densidad de probabilidad está dada por

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}; \quad \text{para todo } x \in \mathbf{R}.$$

La función de distribución de una variable aleatoria normal es

$$F(t) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx; \quad \text{para todo } t \in \mathbf{R},$$

y utilizaremos la notación $X \sim N(\mu, \sigma^2)$ para decir que X está normalmente distribuida con parámetros μ y σ^2 .

Comentarios.

1. La distribución normal es también denominada distribución de *Laplace-Gauss*.
2. En el caso en que $\mu = 0$ y $\sigma^2 = 1$ se dice que X tiene distribución normal estándar o unitaria y se utiliza la notación

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}; \quad y \quad \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx;$$

para la función de densidad y la función de distribución de X respectivamente.

3. Dado que f es claramente no negativa, para demostrar que es una función de densidad es suficiente demostrar que su integral sobre $\mathbf{R}$ es igual a 1. Es decir,

$$\int_{-\infty}^{\infty} f(x) dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = 1.$$

Mediante el cambio de variables $y = \frac{x-\mu}{\sigma}$ obtenemos: $dy = \frac{1}{\sigma} dx$ y $y \in (-\infty, \infty)$; con esto

$$\int_{-\infty}^{\infty} f(x) dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy.$$

Demostrar que la integral del miembro de la derecha en esta última igualdad es igual a 1, es equivalente a demostrar que

$$\int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy = \sqrt{2\pi}.$$

Para esto consideremos

$$I = \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \implies I^2 = \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx \times \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy,$$

esto es

$$I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} e^{-\frac{y^2}{2}} dx dy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-\frac{(x^2+y^2)}{2}} dx dy.$$

Pasando a coordenadas polares obtenemos $x = r \cos(\theta)$, $y = r \sin(\theta)$ y $dx dy = r dr d\theta$ donde $r \in [0, \infty)$ y $\theta \in [0, 2\pi]$. Entonces

$$I^2 = \int_0^{\infty} \int_0^{2\pi} e^{-\frac{r^2}{2}} r d\theta dr = 2\pi \int_0^{\infty} e^{-\frac{r^2}{2}} r dr = 2\pi.$$

Proposición 3.4 Si X es una variable aleatoria normalmente distribuida con parámetros μ y σ^2 , entonces $Y = \alpha X + \beta$ es una variable aleatoria normalmente distribuida con parámetros $\alpha\mu + \beta$ y $\alpha^2\sigma^2$.

Demostración. Denotemos como F_Y a la función de distribución de Y . Consideramos el caso en que $\alpha > 0$. El caso en que $\alpha < 0$ es análogo y se deja al lector.

$$F_Y(t) = P(Y \leq t) = P(\alpha X + \beta \leq t) = P(X \leq \frac{t - \beta}{\alpha}) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{t - \beta}{\alpha}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} dx.$$

Mediante el cambio de variables $z = \alpha x + \beta$, obtenemos $x = \frac{z - \beta}{\alpha}$ y $dz = \alpha dx$; donde además se tiene que $-\infty < z < t$. Entonces

$$\frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\frac{t - \beta}{\alpha}} e^{-\frac{(x - \mu)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{(z - \beta - \alpha\mu)^2}{2\sigma^2}} \frac{1}{\alpha} dz = \frac{1}{\alpha\sigma\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{(z - \beta - \alpha\mu)^2}{2\alpha^2\sigma^2}} dz,$$

por lo tanto la función de densidad de Y es

$$f_Y(z) = \frac{1}{\alpha\sigma\sqrt{2\pi}} e^{-\frac{(z - \beta - \alpha\mu)^2}{2\alpha^2\sigma^2}}, \quad \text{de donde } Y \sim N(\alpha\mu + \beta, \alpha^2\sigma^2).$$

Un sencillo pero importante corolario de la proposición anterior es que si $X \sim N(\mu, \sigma^2)$, entonces la variable aleatoria

$$Z = \frac{X - \mu}{\sigma} = \frac{1}{\sigma}X + \left(-\frac{\mu}{\sigma}\right) \sim N\left(\frac{1}{\sigma}\mu + \left(-\frac{\mu}{\sigma}\right), \frac{1}{\sigma^2}\sigma^2\right) = N(0, 1).$$

Este hecho resulta muy útil, pues hace posible que la distribución de cualquier variable aleatoria $X \sim N(\mu, \sigma^2)$ pueda expresarse en términos de una normal estándar, ya que

$$F_X(t) = P(X \leq t) = P\left(\frac{X - \mu}{\sigma} \leq \frac{t - \mu}{\sigma}\right) = \Phi\left(\frac{t - \mu}{\sigma}\right).$$

Proposición 3.5 Para la distribución normal estándar se cumple que

$$\Phi(-t) = 1 - \Phi(t) \quad \text{es decir} \quad \Phi(t) + \Phi(-t) = 1.$$

Demostración.

$$\Phi(-t) + \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-t} e^{-\frac{x^2}{2}} dx + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx;$$

haciendo el cambio de variables $y = -x$ se tiene que $dy = -dx$ y $y \in [t, \infty)$. Entonces

$$\Phi(-t) + \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_t^{\infty} e^{-\frac{y^2}{2}} dy + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-\frac{x^2}{2}} dx = 1.$$

Comentario. Las propiedades anteriores permiten calcular probabilidades de variables aleatorias normales usando la distribución normal estándar para valores de $x \geq 0$.

Ejemplos 3.14

1. Sea $X \sim N(3, 9)$, entonces $Z = \frac{X-3}{3} \sim N(0, 1)$. Por lo tanto

$$\begin{aligned} P(2 < X < 5) &= P\left(\frac{2-3}{3} < \frac{X-3}{3} < \frac{5-3}{3}\right) = P\left(\frac{-1}{3} < Z < \frac{2}{3}\right) = \Phi\left(\frac{2}{3}\right) - \Phi\left(\frac{-1}{3}\right) \\ &= \Phi\left(\frac{2}{3}\right) - 1 + \Phi\left(\frac{1}{3}\right) = 0,7454 - 1 + 0,6293 = 0,3747. \end{aligned}$$

$$P(X > 0) = P\left(\frac{X-3}{3} > -1\right) = 1 - \Phi(-1) = \Phi(1) = 0,8413.$$

$$\begin{aligned} P(|X - 3| > 6) &= P(X - 3 > 6) + P(X - 3 < -6) \\ &= P\left(\frac{X-3}{3} > 2\right) + P\left(\frac{X-3}{3} < -2\right) \\ &= P(Z > 2) + P(Z < -2) \\ &= 1 - \Phi(2) + \Phi(-2) = 1 - \Phi(2) + 1 - \Phi(2) \\ &= 2(1 - \Phi(2)) = 2(1 - 0,9772) = 0,0456. \end{aligned}$$

3. Variable Aleatoria Exponencial.

Una variable aleatoria continua X se dice que tiene distribución **exponencial** ó que

está **exponencialmente distribuida**, con parámetro $\lambda > 0$, si su función de densidad es

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (3.17)$$

En este caso la función de distribución está dada por

$$F(t) = P(X \leq t) = \begin{cases} \int_0^t \lambda e^{-\lambda x} dx & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases} = \begin{cases} 1 - e^{-\lambda t} & \text{si } t \geq 0 \\ 0 & \text{si } t < 0 \end{cases} \quad (3.18)$$

Ejemplos 3.15

1. En el ejemplo 2.10.1 se ha definido la variable $X(t) = \# \text{terremotos en } t \text{ semanas}$ y puesto que $X(t) \sim \mathcal{P}(2t)$, se tiene que la variable aleatoria $T = \text{tiempo hasta el próximo terremoto}$ tiene función de distribución dada por $F_T(t) = 1 - e^{-2t}$; es decir, T es una variable aleatoria exponencial con parámetro $\lambda = 2$.

2. La duración de una llamada, en minutos, es una variable aleatoria exponencial con parámetro $\lambda = \frac{1}{10}$. Si una persona llega inmediatamente antes que usted a un teléfono público, encuentre la probabilidad de que usted tenga que esperar: i) más de 10 minutos; ii) entre 10 y 20 minutos.

Solución: Sea $X = \text{la duración de una llamada}$, entonces se quiere calcular:

$$i) P(X > 10) = 1 - P(X \leq 10) = 1 - (1 - e^{-\frac{10}{10}}) = e^{-1} = 0,368.$$

$$ii) P(10 < X < 20) = \int_{10}^{20} \frac{1}{10} e^{-\frac{u}{10}} du = -e^{-\frac{u}{10}} \Big|_{10}^{20} = -e^{-2} + e^{-1} = 0,23255.$$

4. Variable Aleatoria Gamma.

Una variable aleatoria X se dice que tiene distribución **gamma** con parámetros (t, λ) , donde $t > 0$ y $\lambda > 0$, si su función de densidad es

$$f(x) = \begin{cases} \frac{\lambda e^{-\lambda x} (\lambda x)^{t-1}}{\Gamma(t)} & \text{si } x \geq 0 \\ 0 & \text{si } x < 0, \end{cases} \quad (3.19)$$

$$\text{donde } \Gamma(t) = \int_0^\infty e^{-y} y^{t-1} dy.$$

Comentarios.

1. Una variable aleatoria gamma con parámetros $(1, \lambda)$ es una variable aleatoria exponencial con parámetro λ .

2. Si se integra por partes la función Γ se obtiene que

$$\begin{aligned}\Gamma(t) &= \int_0^\infty e^{-y} y^{t-1} dy = -e^{-y} y^{t-1} \Big|_0^\infty + \int_0^\infty e^{-y} (t-1) y^{t-2} dy \\ &= (t-1) \int_0^\infty e^{-y} y^{t-2} dy = (t-1) \Gamma(t-1).\end{aligned}$$

En particular, para $t = n \in \mathbf{N}$, se puede demostrar fácilmente por inducción que $\Gamma(n) = (n-1)!$, ya que

$$\begin{aligned}\Gamma(n) &= (n-1) \Gamma(n-1) = (n-1)(n-2) \Gamma(n-2) = \dots \\ &= (n-1)(n-2) \dots (3)(2) \Gamma(1) = (n-1)!\end{aligned}$$

3. La distribución gamma con parámetros (n, λ) , donde $n \in \mathbf{N}$, aparece en las aplicaciones con frecuencia como la distribución de la variable aleatoria $T_n = \text{tiempo hasta que ocurran } n \text{ eventos}$; la cual es una generalización de la variable T presentada en el ejemplo 3.15.1. En efecto, supongamos que la ocurrencia de un cierto fenómeno satisface las hipótesis H_i , H_{ii} y H_{iii} , con parámetro λ . Como antes definimos $X(t) = \# \text{ de eventos en } [0, t]$ y tenemos que $X(t) \sim \mathcal{P}(\lambda t)$. Si denotamos por F_n la función de distribución de T_n , entonces para cada $t \geq 0$

$$F_n(t) = P(T_n \leq t) = P(X(t) \geq n) = \sum_{k=n}^{\infty} P(X(t) = k) = \sum_{k=n}^{\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} = \sum_{k=n}^{\infty} g_k(t).$$

Como para cada $k \geq 1$ la función g_k es derivable y $\sum_{k=n}^{\infty} g'_k(t)$ converge uniformemente, la función de densidad de T_n , digamos f_n , puede obtenerse derivando término a término la serie que representa a la función de distribución F_n , es decir

$$f_n(t) = F'_n(t) = \sum_{k=n}^{\infty} g'_k(t) = \lambda e^{-\lambda t} \frac{(\lambda t)^{n-1}}{(n-1)!},$$

lo que significa que T_n tiene distribución gamma con parámetros (n, λ) .

4. Cuando n es un entero positivo, la distribución gamma con parámetros $(\frac{n}{2}, \frac{1}{2})$ se denomina *chi cuadrado con n grados de libertad*, lo que se denota como χ_n^2 . Usualmente se usa esta distribución para describir el error cometido al tratar de dar en un blanco en $\mathbf{R}^n$, cuando cada coordenada del error está normalmente distribuida. Estudiaremos más adelante la relación entre esta distribución y la normal.

5. Variable Aleatoria Cauchy.

Una variable aleatoria X se dice que tiene distribución de **Cauchy** con parámetro $\alpha \in \mathbf{R}$, si su función de densidad es

$$f(x) = \frac{1}{\pi} \frac{1}{[1 + (x - \alpha)^2]}, \quad \text{donde } -\infty < x < \infty.$$

La función de distribución de Cauchy está dada por

$$F(t) = \frac{1}{\pi} \int_{-\infty}^t \frac{1}{[1 + (x - \alpha)^2]} dx = \frac{1}{\pi} \arctg(x - \alpha)|_{-\infty}^t = \frac{1}{2} + \frac{\arctg(t - \alpha)}{\pi}.$$

Ejemplos 3.16

1. Una aguja que puede girar en torno al punto $(-1, 0)$ del eje x se impulsa para que gire y luego se la deja parar espontáneamente. Sea θ el ángulo que forma la aguja con el eje x y supongamos que θ está uniformemente distribuido sobre $[-\frac{\pi}{2}, \frac{\pi}{2}]$, es decir que

$$F_{\theta}(t) = \begin{cases} 0 & \text{si } t < -\frac{\pi}{2} \\ \frac{t + \frac{\pi}{2}}{\pi} & \text{si } t \in [-\frac{\pi}{2}, \frac{\pi}{2}) \\ 1 & \text{si } t \geq \frac{\pi}{2} \end{cases}$$

Entonces la variable aleatoria $Y(\theta) = \tg(\theta)$ tiene distribución de Cauchy con parámetro $\alpha = 0$. En efecto

$$P(Y(\theta) \leq t) = P(\tg(\theta) \leq t) = P(\theta \leq \arctg(t)) = \frac{1}{2} + \frac{\arctg(t)}{\pi}.$$

6. Variable Aleatoria Beta.

Una variable aleatoria X se dice que tiene distribución de **beta** con parámetros a y b , si su función de densidad está dada por

$$f(x) = \begin{cases} \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} & \text{si } x \in (0, 1) \\ 0 & \text{si } x \notin (0, 1) \end{cases}$$

donde $B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx$ se denomina *función beta*.

Apéndice.

Problema Famoso: La paradoja de Bertrand.

Considere una cuerda aleatoria de un círculo de radio r . ¿Cuál es la probabilidad de que la longitud de la cuerda sea mayor que el lado del triángulo equilátero inscrito en ese círculo?

Para resolver este problema debemos precisar que se entiende por "cuerda aleatoria". Recordemos que, la mediatriz de un segmento es el segmento de recta perpendicular al segmento, en su punto medio, trazada desde un punto dado.

Solución 1: La posición de una cuerda puede ser determinada por su distancia al centro del círculo, que denotaremos por O . Esta distancia puede variar entre 0 y r . Ahora consideremos un triángulo equilátero ABC inscrito en el círculo de radio r y sean a_1 , a_2 y a_3 las distancias del centro a cada uno de los lados. Tenemos que: $OA = OB = OC$, si OA_1 es la mediatriz de AB , OA_2 es la mediatriz de AC y OA_3 es la mediatriz de CB , entonces A_1 es el punto medio de AB , A_2 es el punto medio de AC y A_3 es el punto medio de CB . Sea $m = |OA_1|$; entonces $a_1 = a_2 = a_3 = a = \sqrt{r^2 - m^2}$ y como m satisface $\cos 30^\circ = \frac{m}{r} = \frac{\sqrt{3}}{2}$ entonces $m = r \frac{\sqrt{3}}{2}$ y $a = \sqrt{r^2 - (r \frac{\sqrt{3}}{2})^2} = \frac{r}{2}$.

La longitud de la cuerda es mayor que el lado del triángulo si su distancia al centro es menor que $a = \frac{r}{2}$. Si asumimos entonces que una cuerda aleatoria es aquella cuya distancia D al centro O es una variable aleatoria uniformemente distribuida en $[0, r]$, la probabilidad de que la cuerda sea mayor que el lado del triángulo equilátero inscrito ABC es $P(D < \frac{r}{2}) = \int_0^{\frac{r}{2}} \frac{1}{r} dx = \frac{1}{2}$.

Solución 2. Dada una cuerda arbitraria en el círculo, considere la tangente al círculo que pasa por uno de los extremos de la cuerda. La posición de la cuerda queda determinada por el ángulo θ entre la tangente y la cuerda; el cual varía entre 0 y 180° . En este caso la longitud de la cuerda es mayor que la longitud del lado del triángulo si $60^\circ < \theta < 120^\circ$. Asumiendo θ uniformemente distribuido se tiene que la probabilidad buscada es $P(60^\circ < \theta < 120^\circ) = \frac{120^\circ - 60^\circ}{180^\circ} = \frac{1}{3}$.

Se observa entonces que el resultado varía dependiendo del experimento que se use para caracterizar una cuerda aleatoria.

Ejercicios 3.3

1. Sea X una variable aleatoria con función de densidad de probabilidad

$$f(x) = \begin{cases} c(1 - x^2) & \text{si } -1 < x < 1 \\ 0 & \text{si } x \notin (-1, 1) \end{cases} \quad (3.20)$$

- i) ¿Cuál es el valor de c ?
- ii) ¿Cuál es la función de distribución de X ?

2. Un sistema puede funcionar por una cantidad aleatoria de tiempo X , medida en meses. Si la densidad de X está dada por

$$f(x) = \begin{cases} cxe^{-x/2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0, \end{cases} \quad (3.21)$$

¿cuál es la probabilidad de que el sistema funcione por lo menos 5 meses?

3. A una parada de autobús llegan unidades cada 15 minutos, a partir de las 7:00 a.m. Si un pasajero llega a la parada a una hora que está uniformemente distribuida entre las 7:00 y las 7:30, encuentre la probabilidad de que su espera por el autobús sea de: i) menos de 5 minutos. ii) más de 10 minutos.
4. Los trenes que se dirigen a un punto A llegan a la estación a intervalos de 15 minutos comenzando a las 7:00 a.m., mientras que los trenes que se dirigen al punto B llegan a intervalos de 15 minutos comenzando a las 7:05 a.m. Si un pasajero llega a la estación a una hora uniformemente distribuida entre las 7:00 y las 8:00 a.m. y se monta en el primer tren que llega, ¿qué proporción de veces la persona se dirige al punto A? ¿Qué pasa si el pasajero llega a una hora uniformemente distribuida entre 7:10 y 8:10 a.m.?
5. Si $X \simeq N(10, 36)$ calcule: i) $P(X > 5)$; ii) $P(4 < X < 16)$; iii) $P(X < 8)$; iv) $P(X < 20)$; v) $P(X > 16)$.
6. La cantidad anual de lluvia, en pulgadas, en cierta región está normalmente distribuida con parámetros $\mu = 40$ y $\sigma = 4$. ¿Cuál es la probabilidad de que, comenzando con este año, haya que esperar 10 años antes de tener un caudal de lluvia por encima de 50 pulgadas? ¿Qué suposiciones está haciendo? Sugerencia: Considere la variable aleatoria $T = \#$ años para que lluevan más de 50 pulgadas.
7. Expertos certifican que la duración en días de un embarazo es una variable aleatoria aproximadamente normal con parámetros $\mu = 270$ y $\sigma^2 = 100$. Para la realización de un estudio se seleccionan 10 mujeres embarazadas. Aquellas cuyos embarazos duren menos de 240 días ó más de 290 serán incluidas en un programa de tratamiento especial. ¿Cuál es la probabilidad de que menos de 8 mujeres sean incluidas en el programa?
8. La altura, en pulgadas de un hombre de 25 años es una variable aleatoria normal con parámetros $\mu = 71$ y $\sigma^2 = 6,25$. ¿Qué porcentaje de hombres de 25 años miden más de 6 pies 2 pulgadas? ¿Qué porcentaje de hombres de más de 6 pies miden más de 6 pies 5 pulgadas? Nota: un pie = 12 pulgadas.
9. El ancho, en pulgadas, de una ranura de una pieza de duraluminio forjado está normalmente distribuido con $\mu = 0,9$ y $\sigma = 0,003$. Los límites especificados son $0.9-0.005$ a $0.9+0.005$. ¿Qué porcentaje de piezas serán defectuosas?

¿Cuál es el máximo valor σ_M de σ que permitirá no más de un 1 % de piezas defectuosas cuando los anchos de las ranuras están normalmente distribuidos con parámetros $\mu = 0,9$ y $\sigma = \sigma_M$?

10. Sea X una variable aleatoria con distribución normal estándar. Demuestre que:
 - i) $P(|X| < k) = 2\Phi(k) - 1$ ii) $P(|X| > k) = 2(1 - \Phi(k))$.
11. Si la variable aleatoria $X \simeq N(1, 4)$, calcule: i) $P(-3 \leq X \leq 3)$; ii) $P(-5 \leq X \leq 3)$; iii) $P(-1 \leq X \leq 0)$.
12. Demuestre que si X es una variable aleatoria exponencial con parámetro λ , y $c > 0$; entonces cX es exponencial con parámetro λ/c .
13. El tiempo, en horas, requerido para reparar una máquina es una variable aleatoria exponencialmente distribuida con parámetro $\lambda = 1/2$.
 - i) ¿Cuál es la probabilidad de que el tiempo de reparación sea mayor de 2 horas?
 - ii) ¿Cuál es la probabilidad condicional de que la reparación tome al menos 10 horas, dado que toma más de 9 horas?
14. El número de años que un radio funciona está exponencialmente distribuido con parámetro $\lambda = 1/8$. ¿Cuál es la probabilidad de que un radio funcione por más de 8 años?
15. Sea X una variable aleatoria de Cauchy con función de densidad de dada por:

$$f(x) = \frac{1}{\pi[1 + ((x - \alpha)^2)]}, \quad x \in \mathbf{R}, \quad \alpha > 0.$$

a) Demuestre que f es en efecto una densidad.

b) Calcule: $P(X \leq -1)$; $P(X > 0)$; $P(X < 1)$.

16. Si X es una variable aleatoria normal con parámetros $\mu = 2$, $\sigma^2 = 9$, calcule:
 - a) $P(X > 4)$; b) $P(-3 < X < 5)$; c) $P(8 < X < 12)$.

3.5. Esperanza y Varianza de Variables Aleatorias.

3.5.1. Caso Discreto.

Definición 3.10 Sea X una variable aleatoria discreta tal que su conjunto de valores es $\{x_n\}_{n \geq 1}$ y su función de probabilidad $\{p_n\}_{n \geq 1}$, donde $p_n = P(X = x_n)$. La **esperanza matemática o valor esperado** de X es el número denotado por $E[X]$ y definido como

$$E[X] = \sum_{n=0}^{\infty} x_n P(X = x_n) = \sum_{n=0}^{\infty} x_n p_n;$$

siempre y cuando la serie anterior sea convergente. En este caso se dice que existe la esperanza matemática de X .

Observe que $E[X]$ es un promedio ponderado de los valores de la variable aleatoria X . En algunos casos $E[X]$ se denomina *media* ó *media ponderada* de X .

Ejemplos 3.17

1. Sea X el resultado de lanzar un dado perfecto. Entonces como $p_i = \frac{1}{6}$ para $i = 1, \dots, 6$

$$E[X] = 1p_1 + 2p_2 + 3p_3 + 4p_4 + 5p_5 + 6p_6 = \frac{21}{6} = \frac{7}{2}.$$

2. Sea X una variable aleatoria de Bernoulli con parámetro p = probabilidad de éxito. Es decir, $X = 1$ si ocurre un éxito y $X = 0$ si ocurre un fracaso. Entonces

$$E[X] = 0(1 - p) + 1p = p.$$

3. Sea X una variable aleatoria Binomial con parámetros (n, p) . Entonces

$$\begin{aligned} E[X] &= \sum_{k=0}^n kP(X = k) = \sum_{k=1}^n k \binom{n}{k} p^k (1-p)^{n-k} = \sum_{k=0}^n \frac{kn!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=1}^n \frac{(n-1)!n}{(k-1)!(n-k)!} p p^{k-1} (1-p)^{n-k} \\ &= np \sum_{k=0}^{n-1} \frac{(n-1)!}{k!(n-k-1)!} p^k (1-p)^{n-k-1} \\ &= np \sum_{k=0}^{n-1} \binom{n-1}{k} p^k (1-p)^{n-1-k} = np(p + (1-p))^{n-1} = np. \end{aligned}$$

4. Sea X una variable aleatoria de Poisson con parámetro λ . Entonces $E[X] = \lambda$.

3.5.2. Caso Continuo.

Definición 3.11 Sea X una variable aleatoria continua con función de densidad f . Se define la **esperanza de X** como

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx;$$

siempre que la integral anterior sea finita. En este caso ($E[X] < \infty$) se dice que X tiene esperanza o también que es integrable con respecto a la medida de probabilidad dada.

Ejemplos 3.18

1. Si X es una variable aleatoria uniformemente distribuida sobre $[a, b]$, entonces

$$E[X] = \int_a^b x \frac{1}{b-a} dx = \frac{x^2}{2(b-a)} \Big|_a^b = \frac{a+b}{2}.$$

2. Si X es una variable aleatoria normalmente distribuida con parámetros (μ, σ^2) , entonces

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} x \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu + \mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} (x-\mu) e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \mu e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{(x-\mu)}{\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx + \mu = \mu \end{aligned}$$

3. Sea X una variable aleatoria exponencial con parámetro λ , entonces

$$E[X] = \int_0^{\infty} \lambda x e^{-\lambda x} dx = \lambda \left[\frac{-x}{\lambda} e^{-\lambda x} - \frac{1}{\lambda^2} e^{-\lambda x} \right]_0^{\infty} = \frac{1}{\lambda}.$$

Comentario: La esperanza de una variable aleatoria X también se denomina *media* o *valor medio* de la variable aleatoria X .

3.5.3. Aplicaciones.

El juego de San Petersburgo. Se lanza una moneda perfecta hasta que salga cara. Si esto ocurre en el i -ésimo lanzamiento usted recibirá $\$2^i$. Sea $X = \text{ganancia obtenida}$, entonces X toma los valores $2, 2^2, \dots, 2^k, \dots$. Entonces la ganancia esperada es igual a

$$E[X] = 2 \frac{1}{2} + 2^2 \frac{1}{2^2} + 2^3 \frac{1}{2^3} \dots = \sum_{k=1}^{\infty} 1 = \infty.$$

Es decir la esperanza de X no existe. Sin embargo, si se asume que la *casa* nunca se arruina, esto también podría interpretarse como que la ganancia tiende a crecer sin límites. En este caso, es razonable preguntar ¿cuánto estaría usted dispuesto a pagar

para entrar en el juego? Si usted paga $\$2^{10} = \1024 por entrar, entonces para ganar necesita que haya al menos 9 sellos consecutivos. La probabilidad de este evento es menor o igual a 0,00195; que es bastante pequeño. Se presenta en este juego una situación paradójica, ya que, pese a que se tiene una ganancia esperada infinita, no es razonable arriesgarse pues la probabilidad de ganar es muy pequeña.

Determinación de Primas de Seguros. Una persona de 36 años adquiere un seguro de vida de \$50.000 por 20 años. Sea $p_{36} = P(\text{el individuo sobrevive más de 20 años})$. ¿Cuál deberá ser la prima que cobre la compañía para que la ganancia esperada sea positiva?

Sea L la prima total a pagarse durante los 20 años y sea $Y = \text{ganancia de la compañía}$, entonces

$$E[Y] = Lp_{36} - 50,000(1 - p_{36}) > 0 \iff L > \frac{50,000(1 - p_{36})}{p_{36}}.$$

Juego justo e injusto. Suponga que X es una variable aleatoria que representa la ganancia de un jugador en un juego de azar. Cuando $X < 0$ se considera que hay pérdida o que se paga una cuota para entrar en el juego. Decimos que el juego es *justo* o equitativo, si $E[X] = 0$. Si $E[X] \neq 0$ decimos que el juego es *injusto*, siendo $E[X] > 0$ cuando el juego es favorable al jugador y $E[X] < 0$ cuando el juego es desfavorable al jugador.

Si consideramos el juego de Lotto consistente en acertar 6 números de un total de 54 números, tenemos que la probabilidad de ganar un juego es $p = 1/25,827,165 \simeq 3,8 \times 10^{-8}$. Suponga que la ganancia posible es de \$10.000.000 y que se paga un dólar por 2 juegos. Entonces la oportunidad de perder es $p_1 = \frac{25,827,163}{25,827,165} \simeq 0,999$ y la probabilidad de ganar es $p_2 \simeq 7,6 \times 10^{-8}$. Por lo tanto

$$E[X] = (-1)(0,999) + 10^7 \times 7,6 \times 10^{-8} = -,999 + 0,76 = -0,239,$$

de donde se concluye que, de acuerdo al criterio dado, el juego es injusto y desfavorable al jugador.

3.5.4. Esperanza de una Función de una Variable Aleatoria

Lema 3.1 Si Y es una variable aleatoria continua, entonces

$$E[Y] = \int_0^\infty P\{Y > y\}dy - \int_0^\infty P\{Y < -y\}dy.$$

Demostración. Denotemos por f_Y la función de densidad de Y . Entonces puesto que $P\{Y > y\} = \int_y^\infty f_Y(x)dx$

$$\int_0^\infty P\{Y > y\}dy = \int_0^\infty \int_y^\infty f_Y(x)dx dy.$$

Intercambiando el orden de integración en la igualdad anterior tenemos que

$$\int_0^\infty P\{Y > y\}dy = \int_0^\infty \int_0^x f_Y(x)dydx = \int_0^\infty x f_Y(x)dx. \quad (3.22)$$

De manera análoga se tiene que

$$\int_0^\infty P\{Y < -y\}dy = \int_0^\infty \int_{-\infty}^{-y} f_Y(x)dx dy \quad (3.23)$$

$$(3.24)$$

$$= \int_{-\infty}^0 \int_0^{-x} f_Y(x)dydx \quad (3.25)$$

$$= - \int_{-\infty}^0 x f_Y(x)dx. \quad (3.26)$$

Combinando (3.22) y (3.23) obtenemos

$$\begin{aligned} \int_0^\infty P\{Y > y\}dy - \int_0^\infty P\{Y < -y\}dy &= \int_0^\infty x f_Y(x)dx + \int_{-\infty}^0 x f_Y(x)dx. \\ &= \int_{-\infty}^\infty x f_Y(x)dx = E[Y]. \end{aligned}$$

Proposición 3.6 Si X es una variable aleatoria continua con función de densidad f , entonces para toda función g a valores reales, tal que $g(X)$ es una variable aleatoria, se cumple que

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

Demostración. Por el lema anterior se tiene que

$$\begin{aligned} E[g(X)] &= \int_0^\infty P\{g(X) > y\}dy - \int_0^\infty P\{g(X) < -y\}dy \\ &= \int_0^\infty \int_{\{x:g(x)>y\}} f(x)dx dy - \int_0^\infty \int_{\{x:g(x)<-y\}} f(x)dx dy \\ &= \int_{\{x:g(x)>0\}} \int_0^{g(x)} f(x)dy dx - \int_{\{x:g(x)<0\}} \int_0^{-g(x)} f(x)dy dx \\ &= \int_{\{x:g(x)>0\}} g(x)f(x)dx + \int_{\{x:g(x)<0\}} g(x)f(x)dx \\ &= \int_{-\infty}^\infty g(x)f(x)dx. \end{aligned}$$

Comentario: Existe una versión del resultado anterior en el caso en que X es una variable aleatoria discreta. La demostración es sencilla y se deja como ejercicio.

Proposición 3.7 Si X es una variable aleatoria y a y b son constantes entonces

$$E[aX + b] = aE[X] + b$$

siempre que las esperanzas existan.

Demostración. Caso discreto. Como $p_n = P\{\omega : X = x_n\} = P\{\omega : aX + b = ax_n + b\}$, $n \geq 1$, tenemos que

$$E[aX + b] = \sum_{n \geq 0} (ax_n + b)p_n = a \sum_{n \geq 0} x_n p_n + b \sum_{n \geq 0} p_n = aE[X] + b.$$

Caso continuo. Sea f la función de distribución de X y sea $Y = aX + b$. Por la proposición anterior tenemos que

$$E[aX + b] = \int_{-\infty}^{\infty} (ax + b)f(x)dx = a \int_{-\infty}^{\infty} xf(x)dx + b \int_{-\infty}^{\infty} f(x)dx = aE[X] + b.$$

Definición 3.12 Si X es una variable aleatoria con media $E[X] = \mu$, entonces la varianza de X , denotada $Var[X]$, se define como

$$Var[X] = E[(X - E[X])^2].$$

La varianza de X es una *medida de la dispersión* de los valores de X alrededor de la media.

Proposición 3.8 $Var[X] = E[X^2] - (E[X])^2$.

Demostración. Utilizando la proposición anterior, se tiene

$$\begin{aligned} E[(X - E[X])^2] &= E[X^2 - 2XE[X] + (E[X])^2] \\ &= E[X^2] - 2E[X]E[X] + (E[X])^2 = E[X^2] - (E[X])^2. \end{aligned}$$

Ejemplos 3.19

1. Sea X una variable aleatoria normal, $X \simeq N(\mu, \sigma^2)$. Como $E[X] = \mu$, se tiene que

$$\begin{aligned} Var[X] &= E[(X - E[X])^2] = \frac{\sigma^2}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \frac{(x - \mu)^2}{\sigma^2} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx \\ &= \frac{\sigma^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} y^2 e^{-\frac{y^2}{2}} dy = \frac{\sigma^2}{\sqrt{2\pi}} \left[-ye^{-\frac{y^2}{2}} \Big|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} e^{-\frac{y^2}{2}} dy \right] \\ &= \frac{\sigma^2}{\sqrt{2\pi}} [0 + \sqrt{2\pi}] = \sigma^2. \end{aligned}$$

2. Sea $X \simeq b(n, p)$. Sabemos que $E[X] = np$. Entonces

$$\begin{aligned} E[X^2] &= \sum_{k=0}^n k^2 P(X = k) = \sum_{k=1}^n k^2 \binom{n}{k} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n \frac{k^2 n!}{k!(n-k)!} p^k (1-p)^{n-k} = \sum_{k=0}^n (k^2 - k + k) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \\ &= \sum_{k=0}^n k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} + E[X] \\ &= \sum_{k=2}^n \frac{n!}{(k-2)!(n-k)!} p^k (1-p)^{n-k} + E[X] \\ &= p^2 n(n-1) \sum_{k=2}^n \frac{(n-2)!}{(k-2)!(n-k)!} p^{k-2} (1-p)^{n-k} + E[X] \\ &= p^2 n(n-1) \sum_{k=2}^n \binom{n-2}{k-2} p^{k-2} (1-p)^{n-k} + E[X] \\ &= p^2 n(n-1) (p + (1-p))^{n-2} + E[X] \\ &= p^2 n(n-1) + np = np(1 + (n-1)p). \end{aligned}$$

Por lo tanto

$$Var[X] = E[X^2] - (E[X])^2 = np(1 + (n-1)p) - (np)^2 = np(1-p).$$

Definición 3.13 Dada una variable aleatoria X con varianza σ^2 , la **desviación típica ó estándar de X** es igual a σ .

Este valor da una medida de la desviación de los valores de la variable aleatoria con respecto a su media. Cuanto menor es la desviación estándar más se aglutinan los valores de la variable aleatoria en torno a la media.

El valor $E[|X - E[X|]]$ se denomina **desviación media absoluta**.

Ejercicios 3.4

1. Sea X una variable aleatoria de Cauchy con parámetro $\alpha > 0$; es decir, tal que la función de densidad de X está dada por:

$$f(x) = \frac{1}{\pi[1 + ((x - \alpha)^2)]}, \quad x \in \mathbf{R}.$$

Demuestre que X no tiene esperanza. Sugerencia: considere $\alpha = 0$.

2. Sea X es una variable aleatoria con distribución gamma con parámetros (t, λ) , $t > 0$, $\lambda > 0$; es decir, con función de densidad dada por:

$$f(x) = \begin{cases} 0 & \text{si } x < 0 \\ \lambda e^{-\lambda x} \frac{(\lambda x)^{t-1}}{\Gamma(t)} & \text{si } x \geq 0 \end{cases} \quad (3.27)$$

donde $\Gamma(t) = \int_0^\infty e^{-y} y^{t-1} dy$. a) Demuestre que f es una densidad. b) Halle $E[X]$ y $Var[X]$.

3. Halle la esperanza, la varianza y la desviación estándar de cada una de las siguientes variables aleatorias:
 - i) X geométrica con parámetro $p \in (0, 1)$.
 - ii) $X \simeq b(n, p)$ con $p \in (0, 1)$.
 - iii) $X \simeq \mathcal{P}(\lambda)$, con $\lambda > 0$.
 - iv) X uniforme sobre $[a, b]$.
 - v) X exponencial con parámetro $\lambda > 0$.
 - vi) X gamma con parámetros (t, λ) .
4. Sea $X = \#$ de caras obtenido, cuando se realizan dos lanzamientos independientes de una moneda insesgada (perfecta). Calcule $E[X]$.
5. Binomial Negativa. Si se realizan ensayos independientes de Bernoulli, con probabilidad p de éxito, determine el número esperado de ensayos requeridos para obtener r éxitos. Sugerencia: considere las variables aleatorias siguientes:

$X_1 = \#$ de ensayos para primer éxito;

$X_2 = \#$ de ensayos, después del primer éxito, para el segundo éxito; ...

$X_i = \#$ de ensayos, después del éxito $i - 1$, para obtener el i -ésimo éxito.

6. Si $E[X] = 1$ y $Var[X] = 5$ encuentre: a) $E[(X + 2)^2]$. b) $Var[4 + 3X]$.
Sugerencia: Demuestre que $Var[aX + b] = a^2 Var[X]$.
7. Sea X una variable aleatoria binomial negativa (ver ejercicio previo). Halle $Var[X]$.
8. Si X es uniformemente distribuida sobre $[0, 1]$, calcule $E[e^X]$.
9. Sea X = resultado de lanzar un dado perfecto. Calcule $Var[X]$.
10. Uno de los números del 1 al 10 se selecciona aleatoriamente. Usted trata de adivinar el número haciendo preguntas cuyas respuestas serán **si** ó **no**. Calcule el número esperado de preguntas que necesitará hacer en cada uno de los siguientes casos:
i) Su i -ésima pregunta es: ¿es el número igual a i ?
ii) Con cada pregunta usted trata de eliminar la mitad de los números que quedan.
11. Una muestra de 3 objetos es seleccionada aleatoriamente de una caja que contiene 20 objetos de los cuales 4 son defectuosos. Encuentre el número esperado de objetos defectuosos en la muestra.
12. Calcule $E[X]$ si X tiene función de densidad:

$$i) f(x) = \begin{cases} cxe^{-x/2} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0, \end{cases} \quad (3.28)$$

$$ii) f(x) = \begin{cases} c(1 - x^2) & \text{si } -1 < x < 1 \\ 0 & \text{si } x \notin (-1, 1) \end{cases} \quad (3.29)$$

13. El tiempo de vida, en horas, de tubos electrónicos es una variable aleatoria con función de densidad

$$f(x) = \begin{cases} c^2 xe^{-cx} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0, \end{cases} \quad (3.30)$$

donde $c > 0$. Calcule el tiempo de vida esperado de estos tubos.

14. Sea X una variable aleatoria discreta con función de masa de probabilidad $p(x)$. Demuestre que para cualquier función a valores reales g

$$E[g(x)] = \sum_{\{x:p(x)>0\}} g(x)p(x).$$

Capítulo 4

Variables Aleatorias Conjuntas

4.1. Distribución Conjunta de Variables Aleatorias.

En muchas situaciones es necesario tomar en cuenta el comportamiento de varias variables aleatorias al mismo tiempo. Para describir simultáneamente estas variables es necesario definir la función de distribución conjunta de las mismas. Comenzamos por considerar la función de distribución conjunta de dos variables aleatorias. Este concepto se generaliza naturalmente al caso en que se tiene cualquier cantidad finita de variables aleatorias.

Definición 4.1 *Dadas dos variables aleatorias X e Y , se define la función de distribución conjunta de X e Y como*

$$F(a, b) = P(X \leq a, Y \leq b); \quad \text{donde } -\infty < a, b < \infty.$$

Comentarios.

1. Si definimos $A_n = (X \leq a, Y \leq n)$, podemos verificar que para cada $n \geq 1$ $A_{n-1} \subseteq A_n$, por lo que

$$\lim_{n \rightarrow \infty} F(a, n) = \lim_{n \rightarrow \infty} P(A_n) = P(\lim_{n \rightarrow \infty} A_n) = P(X \leq a, Y < \infty) = P(X \leq a) = F_X(a).$$

Análogamente se tiene que

$$\lim_{n \rightarrow \infty} F(n, b) = F_Y(b);$$

lo que permite expresar las distribuciones de X e Y en términos de la distribución conjunta.

2. El cálculo de probabilidades puede hacerse en términos de la función de distribución conjunta.

$$P(X > a, Y > b) = 1 - P((X > a, Y > b)^c) = 1 - P((X > a)^c \cup (Y > b)^c)$$

$$\begin{aligned}
&= 1 - P((X \leq a) \cup (Y \leq b)) \\
&= 1 - P(X \leq a) - P(Y \leq b) + P(X \leq a, Y \leq b) \\
&= 1 - F_X(a) - F_Y(b) + F(a, b).
\end{aligned}$$

3. Si $F(x, y)$ es una función de distribución conjunta de dos variables X e Y , las distribuciones individuales F_X y F_Y se denominan **distribuciones marginales**.

Propiedades de la función de distribución conjunta.

1. $F(x, y)$ es creciente en cualquiera de las dos variables: Sea $x < x'$, puesto que $(X \leq x) \subseteq (X \leq x')$ se tiene que, para todo y :

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

La verificación para la segunda variable es análoga.

2. $\lim_{x \rightarrow -\infty} F(x, y) = \lim_{y \rightarrow -\infty} F(x, y) = 0$.

3. $\lim_{x \rightarrow \infty} \lim_{y \rightarrow \infty} F(x, y) = 1$.

4. Como F es creciente en ambas variables, entonces $0 \leq F(x, y) \leq 1$ para cada $x \in \mathbf{R}$ y $y \in \mathbf{R}$.

5. F es continua a la derecha en ambas variables.

Nota. Contrariamente a lo que ocurre en el caso de una variable, estas propiedades **no** caracterizan a una función de distribución conjunta; es decir, se puede encontrar una función que satisfaga las propiedades anteriores y que no sea una función de distribución conjunta. Considere por ejemplo la función:

$$F(x, y) = \begin{cases} 1 & \text{si } x + y \geq 0 \\ 0 & \text{si } x + y < 0 \end{cases} \quad (4.1)$$

satisface las condiciones anteriores (verifíquelo), sin embargo si F es una función de distribución, entonces las probabilidades pueden calcularse en términos de F y se cumpliría que:

$$P(-1 < X \leq 3, -1 < Y \leq 3) = F(3, 3) - F(3, -1) - F(-1, 3) + F(-1, -1) = -1$$

lo cual claramente representa una contradicción.

Para que una función F que satisface las condiciones anteriores sea una distribución conjunta debe cumplirse que, para $a < b$ y $c < d$ dados,

$$P(X \leq b, Y \leq d) - P(X \leq b, Y \leq c) - P(X \leq a, Y \leq d) + P(X \leq a, Y \leq c) \geq 0.$$

Definición 4.2 Si X e Y son variables aleatorias discretas, entonces la **función de probabilidad conjunta de (X, Y)** se define como

$$p(x, y) = P(X = x, Y = y).$$

En este caso las funciones de probabilidad marginales de X e Y estan dadas por

$$p_X(t) = P(X = t) = P\left(\cup_{\{y:p(t,y)>0\}} \{X = t, Y = y\}\right) = \sum_{\{y:p(t,y)>0\}} p(t, y);$$

$$p_Y(t) = P(Y = t) = \sum_{\{x:p(x,t)>0\}} p(x, t),$$

respectivamente.

Ejemplos 4.1

1. Suponga que 3 bolas son seleccionadas aleatoriamente de una urna que contiene 3 bolas rojas, 4 blancas y 5 azules. Sean $X = \#$ bolas rojas e $Y = \#$ bolas blancas. Entonces la función de probabilidad conjunta de X e Y asi como también las funciones de probabilidad marginales se pueden resumir en la tabla siguiente

$i \setminus j$	0	1	2	3	$P(X = i)$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
$P(Y = j)$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	

Definición 4.3 Se dice que las variables aleatorias X e Y son **conjuntamente continuas** si existe una función $f : \mathbf{R}^2 \rightarrow [0, \infty)$, tal que para todo $(x, y) \in \mathbf{R}^2$

$$F(x, y) = \int_{-\infty}^y \int_{-\infty}^x f(u, v) du dv.$$

En este caso diremos que F tiene densidad f .

Comentarios.

1. Es claro a partir de esta definición que si F tiene densidad f , entonces para cualquier rectángulo $A = [a, b] \times [c, d]$ se cumple que

$$P((X, Y) \in A) = \int \int_A f(u, v) du dv = \int_a^b \int_c^d f(u, v) dv du.$$

También es posible demostrar que para cualquier conjunto de Borel $C \subseteq \mathbf{R}^2$ se cumple que

$$P((X, Y) \in C) = \int \int_C f(u, v) du dv.$$

En particular dado un rectángulo $C = C_1 \times C_2$ donde C_1 y C_2 son conjuntos de Borel, se tiene que

$$P((X, Y) \in C) = \int_{C_1} \int_{C_2} f(u, v) dv du.$$

2. Si X e Y son conjuntamente continuas ellas son continuas ya que para cada conjunto de borel A

$$P(X \in A) = P(X \in A, Y \in (-\infty, \infty)) = \int_A \int_{-\infty}^{\infty} f(u, v) dv du = \int_A f_X(u) du,$$

donde $f_X(u) = \int_{-\infty}^{\infty} f(u, v) dv$, la cual es efectivamente una densidad, debido a que f lo es. De manera análoga se tiene que $f_Y(v) = \int_{-\infty}^{\infty} f(u, v) du$. Las funciones f_X y f_Y se denominan **densidades marginales** de X e Y respectivamente.

Ejemplos 4.2

1. Supongamos que la función de densidad conjunta de X e Y está dada por

$$f(x, y) = \begin{cases} 2e^{-x-2y} & \text{si } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{en otros casos} \end{cases} \quad (4.2)$$

Entonces

$$\begin{aligned} i. P(X > 1, Y \leq 1) &= \int_0^1 \int_1^{\infty} f(x, y) dx dy = \int_0^1 \int_1^{\infty} 2e^{-x-2y} dx dy \\ &= 2 \int_0^1 e^{-2y} \left(\int_1^{\infty} e^{-x} dx \right) dy = 2 \int_0^1 e^{-2y} \left(-e^{-x} \Big|_1^{\infty} \right) dy \\ &= 2 \int_0^1 e^{-2y} (e^{-1}) dy = -e^{-1} e^{-2y} \Big|_0^1 = e^{-1} (1 - e^{-2}). \end{aligned}$$

$$\begin{aligned}
ii. \quad P(X < Y) &= \int \int_{\{(x,y): x < y\}} f(x,y) dx dy = \int_0^\infty 2 \int_0^y e^{-x-2y} dx dy \\
&= \int_0^\infty 2e^{-2y} \left(-e^{-x} \Big|_0^y \right) dy = \int_0^\infty 2e^{-2y} (1 - e^{-y}) dy \\
&= \int_0^\infty 2e^{-2y} dy - \int_0^\infty 2e^{-3y} dy = -e^{-2y} \Big|_0^\infty + \frac{2}{3} e^{-3y} \Big|_0^\infty = 1 - \frac{2}{3} = \frac{1}{3}.
\end{aligned}$$

$$iii. \quad P(X < a) = \int_0^a \int_0^\infty 2e^{-x-2y} dy dx = \int_0^a e^{-x} dx = -e^{-x} \Big|_0^a = 1 - e^{-a}.$$

2. Supongamos que la función de densidad conjunta de X e Y está dada por

$$f(x,y) = \begin{cases} e^{-(x+y)} & \text{si } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{en otros casos} \end{cases} \quad (4.3)$$

Entonces la función de distribución de la variable $\frac{X}{Y}$, para $t > 0$ es igual a

$$\begin{aligned}
F_{\frac{X}{Y}}(t) &= P\left(\frac{X}{Y} \leq t\right) = \int \int_{\{(x,y): \frac{x}{y} \leq t\}} e^{-(x+y)} dx dy \\
&= \int_0^\infty \int_0^{ty} e^{-(x+y)} dx dy = 1 - \frac{1}{1+t}.
\end{aligned}$$

Diferenciando se obtiene la función de densidad de $\frac{X}{Y}$

$$f_{\frac{X}{Y}}(t) = \frac{1}{(1+t)^2}, \quad t > 0.$$

La noción de distribución conjunta de dos variables aleatorias se generaliza de manera natural al considerar la distribución conjunta de n variables aleatorias $X_1, X_2, \dots, X_n$: para números reales $a_1, a_2, \dots, a_n$ se define

$$F(a_1, a_2, \dots, a_n) = P(X_1 \leq a_1, X_2 \leq a_2, \dots, X_n \leq a_n).$$

En este caso se dice que las variables aleatorias $X_1, X_2, \dots, X_n$ son conjuntamente continuas si existe una función $f: \mathbf{R}^n \rightarrow [0, \infty)$, tal que para todo $(x_1, \dots, x_n) \in \mathbf{R}^n$

$$F(x_1, \dots, x_n) = \int_{-\infty}^{x_1} \dots \int_{-\infty}^{x_n} f(u_1, \dots, u_n) du_1 \dots du_n.$$

La función f se denomina **densidad conjunta** de $(X_1, \dots, X_n)$.

4.2. Esperanza de Funciones de Variables Aleatorias.

En esta sección consideramos una familia finita de n variables aleatorias $X_1, X_2, \dots, X_n$ y una función $g : \mathbf{R}^n \rightarrow \mathbf{R}$ lo suficientemente regular para que $Y = g(X_1, X_2, \dots, X_n)$ sea una variable aleatoria.

Preliminares.

1) Esperanza de la Suma de Variables Aleatorias.

Proposición 4.1 *Dadas variables aleatorias X e Y , la suma $X + Y$ es una variable aleatoria cuya esperanza es $E[X + Y] = E[X] + E[Y]$, siempre que todas las esperanzas existan.*

Demostración.

I) **Caso Discreto.** Sean X e Y variables aleatorias discretas con función de probabilidad $p_{ij} = P(X = x_i, Y = y_j)$ para $i, j \geq 0$. Si se define $S = X + Y$, se tiene que S es una variable aleatoria con función de probabilidad

$$q_k = P(S = s_k) = \sum_{\{(i,j): x_i + y_j = s_k\}} p_{ij}.$$

Entonces

$$\begin{aligned} E[S] &= \sum_k s_k P(S = s_k) = \sum_k s_k \sum_{\{(i,j): x_i + y_j = s_k\}} p_{ij} = \sum_k \sum_{\{(i,j): x_i + y_j = s_k\}} s_k p_{ij} \\ &= \sum_k \sum_{\{(i,j): x_i + y_j = s_k\}} (x_i + y_j) p_{ij} = \sum_i \sum_j (x_i + y_j) p_{ij} \\ &= \sum_i x_i \sum_j p_{ij} + \sum_j y_j \sum_i p_{ij} = \sum_i x_i p_X(i) + \sum_j y_j p_Y(j) = E[X] + E[Y]. \end{aligned}$$

II) **Caso Continuo.** Sean X e Y variables aleatorias conjuntamente continuas con función de densidad conjunta f , y sea $S = X + Y$. Entonces la función de distribución de S está dada por

$$\begin{aligned} F_{X+Y}(t) &= P(S \leq t) = P(X + Y \leq t) = \int \int_{\{(x,y): x+y \leq t\}} f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{t-x} f(x, y) dy dx. \end{aligned}$$

Haciendo el cambio de variables $z = x + y$ ($y = z - x$), se tiene que $-\infty < z < t$ por lo que

$$\int_{-\infty}^{t-x} f(x, y) dy = \int_{-\infty}^t f(x, z - x) dz;$$

de aquí que

$$F_{X+Y}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^t f(x, z - x) dz dx = \int_{-\infty}^t \int_{-\infty}^{\infty} f(x, z - x) dx dz = \int_{-\infty}^t f_Z(z) dz;$$

donde $f_Z(z) = \int_{-\infty}^{\infty} f(x, z - x) dx$ ya que $Z = X + Y$. Se tiene finalmente que

$$\begin{aligned} E[X + Y] &= E[Z] = \int_{-\infty}^{\infty} z f_Z(z) dz = \int_{-\infty}^{\infty} z \int_{-\infty}^{\infty} f(x, z - x) dx dz \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} z f(x, z - x) dx dz. \end{aligned}$$

Haciendo el cambio $z = x + y$ se tiene que

$$\begin{aligned} E[X + Y] &= E[Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x + y) f(x, y) dy dx \\ &= \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) dy dx + \int_{-\infty}^{\infty} y \int_{-\infty}^{\infty} f(x, y) dx dy \\ &= \int_{-\infty}^{\infty} x f_X(x) dx + \int_{-\infty}^{\infty} y f_Y(y) dy = E[X] + E[Y]. \end{aligned}$$

2) Monotonía de la Esperanza.

Proposición 4.2 Sean Y y Z variables aleatorias tales que $Y > Z$, entonces $E[Y] \geq E[Z]$.

Demostración. Utilizando la proposición anterior tenemos que

$$E[Y] - E[Z] = E[Y - Z] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (y - z) f(y, z) dy dz \geq 0,$$

ya que $y - z \geq 0$.

Teorema 4.1 Esperanza de una Función de Variables Aleatorias.

Sean $X_1, X_2, \dots, X_n$ variables aleatorias conjuntamente continuas con función de densidad dada por $f : \mathbf{R}^n \rightarrow \mathbf{R}$ y sea $g : \mathbf{R}^n \rightarrow \mathbf{R}$ una función lo suficientemente

regular para que $Y = g(X_1, X_2, \dots, X_n)$ sea una variable aleatoria. Entonces el valor esperado de Y se calcula mediante la fórmula

$$E[Y] = \int \dots \int_{\mathbf{R}^n} g(x_1, x_2, \dots, x_n) f(x_1, x_2, \dots, x_n) dx_1 \dots dx_n,$$

siempre que la integral múltiple sea absolutamente convergente, es decir siempre que

$$\int \dots \int_{\mathbf{R}^n} |g(x_1, x_2, \dots, x_n)| f(x_1, x_2, \dots, x_n) dx_1 \dots dx_n < \infty.$$

Demostración. Paso 1. Observe que, en virtud de la proposición anterior, es suficiente demostrar el teorema en el caso en que $g \geq 0$, ya que si g no es no negativa se puede escribir como una diferencia de funciones no negativas, en efecto $g = g^+ - g^-$, donde

$$g^+(x) = \begin{cases} g(x) & \text{si } g(x) \geq 0 \\ 0 & \text{si } g(x) < 0 \end{cases} \quad (4.4)$$

y

$$g^-(x) = \begin{cases} -g(x) & \text{si } g(x) \leq 0 \\ 0 & \text{si } g(x) > 0 \end{cases} \quad (4.5)$$

son la *parte positiva* y la *parte negativa* de g respectivamente. En este caso se tendrá que, denotando $X = (X_1, \dots, X_n)$ y $x = (x_1, \dots, x_n)$, y usando la proposición 3

$$\begin{aligned} E[g(X)] &= E[g^+(X)] - E[g^-(X)] = \int_{\mathbf{R}^n} g^+(x) f(x) dx - \int_{\mathbf{R}^n} g^-(x) f(x) dx \\ &= \int_{\mathbf{R}^n} (g^+(x) - g^-(x)) f(x) dx = \int_{\mathbf{R}^n} g(x) f(x) dx. \end{aligned}$$

Paso 2. Para $g \geq 0$ definamos para cada entero $m \geq 0$

$$g_m(x) = \frac{k}{2^m} \text{ si } \frac{k}{2^m} \leq g(x) < \frac{k+1}{2^m}, \quad k = 0, 1, \dots$$

Como $g_m(x) = \frac{k}{2^m} \leq g(x) < \frac{k}{2^m} + \frac{1}{2^m}$ se tiene que

$$g(x) - \frac{1}{2^m} \leq \frac{k}{2^m} = g_m(x) \leq g(x), \quad \text{para todo } x \in \mathbf{R}^n. \quad (4.6)$$

Entonces para cada $\omega \in \Omega$

$$g(X(\omega)) - \frac{1}{2^m} \leq g_m(X(\omega)) \leq g(X(\omega)), \quad \text{para todo } \omega, \quad (4.7)$$

donde $X(\omega) = (X_1(\omega), \dots, X_n(\omega))$. Tomando esperanza en esta última expresión y utilizando la monotonía de la esperanza matemática, tenemos que

$$E[g(X)] - \frac{1}{2^m} \leq E[g_m(X)] \leq E[g(X)], \quad (4.8)$$

Multiplicando en (3.12) por f , que es no negativa, obtenemos

$$g(x)f(x) - \frac{1}{2^m}f(x) \leq g_m(x)f(x) \leq g(x)f(x),$$

por lo que

$$\int_{\mathbf{R}^n} g(x)f(x)dx - \int_{\mathbf{R}^n} \frac{1}{2^m}f(x)dx \leq \int_{\mathbf{R}^n} g_m(x)f(x)dx \leq \int_{\mathbf{R}^n} g(x)f(x)dx.$$

Tomando limite en m

$$\lim_{m \rightarrow \infty} \left[\int_{\mathbf{R}^n} g(x)f(x)dx - \frac{1}{2^m} \right] \leq \lim_{m \rightarrow \infty} \int_{\mathbf{R}^n} g_m(x)f(x)dx \leq \lim_{m \rightarrow \infty} \int_{\mathbf{R}^n} g(x)f(x)dx.$$

Esto significa que

$$\lim_{m \rightarrow \infty} \int_{\mathbf{R}^n} g_m(x)f(x)dx = \int_{\mathbf{R}^n} g(x)f(x)dx;$$

por lo que será suficiente demostrar que

$$E[g(X)] = \lim_{m \rightarrow \infty} \int_{\mathbf{R}^n} g_m(x)f(x)dx. \quad (4.9)$$

Por otra parte, tomando limite en m en (3.14) obtenemos que

$$\lim_{m \rightarrow \infty} \left[E[g(X)] - \frac{1}{2^m} \right] \leq \lim_{m \rightarrow \infty} E[g_m(X)] \leq \lim_{m \rightarrow \infty} E[g(X)],$$

es implica que

$$E[g(X)] = \lim_{m \rightarrow \infty} E[g_m(X)].$$

Bastará entonces demostrar que

$$E[g_m(X)] = \int_{\mathbf{R}^n} g_m(x)f(x)dx.$$

Paso 3. Observando que $g_m(X)$ es una variable aleatoria discreta, se tiene

$$\begin{aligned} E[g_m(X)] &= \sum_{k=0}^{\infty} \frac{k}{2^m} P\left(g_m(X) = \frac{k}{2^m}\right) = \sum_{k=0}^{\infty} \frac{k}{2^m} P\left(\frac{k}{2^m} \leq g(X) < \frac{k+1}{2^m}\right) \\ &= \sum_{k=0}^{\infty} \frac{k}{2^m} P\left(X \in g^{-1}\left(\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)\right)\right) = \sum_{k=0}^{\infty} \frac{k}{2^m} \int_{g^{-1}\left(\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)\right)} f(x)dx \\ &= \sum_{k=0}^{\infty} \int_{g^{-1}\left(\left[\frac{k}{2^m}, \frac{k+1}{2^m}\right)\right)} g_m(x)f(x)dx = \int_{\mathbf{R}^n} g_m(x)f(x)dx. \end{aligned}$$

Ejercicios 4.1

1. Se escoge un número X aleatoriamente de el conjunto $A = \{1, 2, 3, 4, 5\}$. Luego se escoge un número Y del subconjunto de A de números no mayores que X . Halle la función de probabilidad conjunta de X e Y .
2. Suponga que 15 % de las familias en una comunidad no tienen hijos, 20 % tienen un hijo, 35 % tienen 2 hijos y 30 % tienen 3 hijos. Suponga además que en cada familia cada niño es con igual probabilidad, e independientemente, hembra o varón. Si se escoge una familia al azar y se definen $H = \#$ de hembras $V = \#$ de varones. Haga una tabla que muestre la función de masa de probabilidad conjunta para (H, V) y las funciones de masa de probabilidad marginales.
3. En una comunidad las familias adquieren sólo dos tipos de mascota: perros y gatos. Cada mascota(perro ó gato) es adquirida independientemente y con la misma probabilidad. Se sabe que un 40 % de las familias no tiene mascota, 30 % tiene 1 mascota, 20 % tiene 2 mascotas y 10 % tiene 3 mascotas. Si $C = \#$ de perros y $G = \#$ de gatos en una familia; halle la función de probabilidad conjunta de (C, G) y calcule el número esperado de perros.
4. Se lanzan dos dados perfectos y distintos. Sea $X = \#$ de dados que caen en un número impar y $Y =$ suma de los dados. (i) Halle la función de probabilidad conjunta de (X, Y) . (ii) Halle la probabilidad condicional de X dado que $Y = 2$.
5. Se lanzan dos dados perfectos, uno rojo y uno verde. Sean $X =$ resultado del dado rojo y $Y =$ menor de los resultados. (i) Halle la función de probabilidad conjunta de (X, Y) . (ii) Calcule $\text{Var}[Y]$.
6. Dos dados simétricos son lanzados. Sean $X =$ menor valor obtenido y $Y =$ mayor valor obtenido. (i) Halle la función de probabilidad conjunta de (X, Y) . (ii) Halle $\text{Var}[Y]$.
7. Considere un grupo de cartas que consiste en J, Q, K y A de las cuatro pintas. Se reparte una mano de 2 cartas. Sean $X = \#$ de diamantes y $Y = \#$ de corazones. Obtebga la función de probabilidad conjunta de (X, Y) y calcule $\text{Var}[X]$.
8. Considere un círculo de radio R y suponga que un punto del círculo se escoge aleatoriamente de manera tal que todas las regiones dentro del círculo tienen igual probabilidad de contener el punto. Es decir, que el punto está uniformemente distribuido dentro del círculo. Suponga que el centro del círculo es el origen de coordenadas y defina X e Y como las coordenadas del punto. Entonces la función de densidad conjunta está dada por

$$f(x, y) = \begin{cases} c & \text{si } x^2 + y^2 \leq R^2 \\ 0 & \text{si } x^2 + y^2 > R^2 \end{cases} \quad (4.10)$$

donde $c > 0$. a) Determine el valor de c . b) Encuentre las funciones de densidad marginales de X y Y . c) Calcule la probabilidad de que la distancia entre el punto y el origen no sea mayor que A .

9. Un grupo de 5 transistores contiene 2 defectuosos. Se prueban los transistores, uno a la vez, hasta encontrar los defectuosos. Sean $N_1 = \#$ de pruebas realizadas hasta que el primer defectuoso es encontrado y $N_2 = \#$ de pruebas adicionales realizadas hasta que el segundo defectuoso es encontrado. Encuentre la función de masa de probabilidad conjunta de (N_1, N_2) . Haga una tabla.

10. La función de densidad de probabilidad conjunta de X y Y está dada por:

$$f(x, y) = c(y^2 - x^2)e^{-y}, \quad -y \leq x \leq y, \quad 0 < y < \infty.$$

- a) Encuentre el valor de c . b) Encuentre las densidades marginales de X y Y .

11. La función de probabilidad conjunta de X e Y está dada por

$$f(x, y) = e^{-(x+y)}, \quad 0 \leq x < \infty, \quad 0 \leq y < \infty.$$

- a) Encuentre $P(X < Y)$; b) $P(X < a)$;

12. La densidad de probabilidad conjunta de X y Y está dada por

$$f(x, y) = \frac{6}{7}(x^2 + \frac{xy}{2}), \quad 0 < x < 1, \quad 0 < y < 2.$$

- i) Halle $P(X > Y)$.
ii) Calcule la función de densidad de X .

13. La función de densidad conjunta de X e Y está dada por

$$f(x, y) = \begin{cases} \frac{1}{y}e^{-(y+\frac{x}{y})} & \text{si } 0 < x < \infty, 0 \leq y < \infty \\ 0 & \text{en otros casos} \end{cases} \quad (4.11)$$

Encuentre $E[X]$ y $E[Y]$.

4.3. Distribución Condicional de Variables Aleatorias.

I) Caso Discreto.

Definición 4.4 Sean X e Y variables aleatorias discretas con funciones de probabilidad p_X y p_Y respectivamente. Se define la **función de probabilidad condicional**

de X dado que $Y = y$ como

$$p_{X|Y}(x|y) = P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)} = \frac{p(x, y)}{p_Y(y)}; \text{ siempre que } p_Y(y) > 0.$$

La **función de probabilidad condicional de Y dado que $X = x$** se define de manera análoga.

Definición 4.5 Sean X e Y variables aleatorias discretas con funciones de probabilidad p_X y p_Y respectivamente. Se define la **función de distribución de probabilidad condicional de X dado que $Y = y$** como

$$F_{X|Y}(x|y) = P(X \leq x|Y = y) = \sum_{s \leq x} p_{X|Y}(s|y) = \sum_{s \leq x} P(X = s|Y = y).$$

Ejemplos 4.3

1. Sean X e Y variables aleatorias con función de probabilidad conjunta $p(x, y)$ definida como

$$p(0, 0) = 0,4; \quad p(0, 1) = 0,2; \quad p(1, 0) = 0,1; \quad p(1, 1) = 0,3.$$

La función de masa de probabilidad condicional de X dado que $Y = 1$ está dada por

$$p_{X|Y}(x|1) = \frac{p(x, 1)}{p_Y(1)}; \quad \text{para } x = 0, 1.$$

En este caso

$$p_Y(1) = p(0, 1) + p(1, 1) = 0,5;$$

por lo que

$$p_{X|Y}(0|1) = \frac{p(0, 1)}{p_Y(1)} = \frac{2}{5} \quad \text{y} \quad p_{X|Y}(1|1) = \frac{p(1, 1)}{p_Y(1)} = \frac{3}{5}.$$

2) Caso Continuo.

Definición 4.6 Sean X e Y variables aleatorias conjuntamente continuas con función de densidad conjunta $f(x, y)$. Entonces la **función de densidad de probabilidad condicional de X dado que $Y = y$** , está dada por

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)}, \quad \text{siempre que } f_Y(y) \neq 0.$$

Definición 4.7 Sean X e Y variables aleatorias conjuntamente continuas con función de densidad conjunta $f(x, y)$. Entonces la **función de distribución condicional de X dado que $Y = y$** , se define como

$$F_{X|Y}(a|y) = P(X \leq a|Y = y) = \int_{-\infty}^a f_{X|Y}(s|y)ds.$$

Para un conjunto de Borel A dado tendremos que

$$P(X \in A|Y = y) = \int_A f_{X|Y}(s|y)ds.$$

Comentario. La expresión $P(X \leq a|Y = y)$ tiene sentido sólo como notación, ya que en el caso continuo $P(Y = y) = 0$. Sin embargo en la definición anterior sólo se requiere que $f_Y(y) \neq 0$.

Ejemplos 4.4

1. La densidad conjunta de X e Y está dada por

$$f(x, y) = \begin{cases} \frac{12}{5}x(2-x-y) & \text{si } 0 < x < 1 \text{ y } 0 < y < 1 \\ 0 & \text{en otro caso} \end{cases}$$

Se tiene entonces que la densidad marginal de Y es igual a

$$f_Y(y) = \int_0^1 \frac{12}{5}x(2-x-y)dx = \frac{12}{5} \frac{(4-3y)}{6}.$$

Por lo tanto

$$f_{X|Y}(x|y) = \frac{\frac{12}{5}x(2-x-y)}{\frac{12}{5} \frac{4-3y}{6}} = \frac{6x(2-x-y)}{4-3y}.$$

siempre que $0 < x < 1$ y $0 < y < 1$

Valores Esperados Condicionales.

En el caso de variables aleatorias conjuntamente discretas X e Y , la función de probabilidad conjunta $p_{X|Y}(x|y)$ es, para cada valor fijo y de Y , una función de probabilidad (verifíquelo!). De manera análoga en el caso continuo la función de densidad condicional $f_{X|Y}(x|y)$ de dos variables aleatorias conjuntamente continuas X e Y , es una función de densidad para y fijo. En ambos casos, esto puede verificarse usando las definiciones correspondientes. Tiene entonces sentido definir la **esperanza condicional de X dado que $Y = y$** como la esperanza de X con respecto a

$p_{X|Y}(x|y)$ ó $f_{X|Y}(x|y)$, según sea el caso. Tenemos entonces que, en el caso discreto, si el conjunto de valores de X es $\{x_n\}_{n \geq 1}$, entonces

$$E[X|Y = y] = \sum_{n \geq 1} x_n p_{X|Y}(x_n|y).$$

Mientras que en el caso continuo, se tiene que

$$E[X|Y = y] = \int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx.$$

Ejemplos 4.5

1. Sean X e Y variables aleatorias conjuntamente continuas con densidad conjunta

$$f(x, y) = \begin{cases} \frac{1}{y} e^{-\frac{x}{y}} e^{-y} & \text{si } 0 < x < \infty, 0 < y < \infty \\ 0 & \text{en otros casos} \end{cases} \quad (4.12)$$

La densidad marginal de Y está dada por:

$$f_Y(y) = \int_0^{\infty} \frac{1}{y} e^{-\frac{x}{y}} e^{-y} dx = e^{-y}.$$

Usando la definición, se tiene que:

$$f_{X|Y}(x|y) = \frac{e^{-\frac{x}{y}}}{y}, \quad \text{si } 0 < x < \infty, 0 < y < \infty.$$

Entonces, para cada y fijo se tiene

$$E[X|Y = y] = \int_0^{\infty} x \frac{e^{-\frac{x}{y}}}{y} dx = y \int_0^{\infty} u e^{-u} du = y.$$

Comentario. Note que la esperanza condicional, con respecto a Y es una función de los valores de Y .

Proposición 4.3 Dadas variables aleatorias X e Y , $E[X] = E[E[X|Y]]$.

Demostración

a) Sean X e Y variables aleatorias discretas, con funciones de probabilidad

$$p_X(n) = P(X = x_n) \quad \text{y} \quad p_Y(n) = P(Y = y_n), \quad n \geq 1$$

respectivamente. Entonces

$$\begin{aligned} E[E[X|Y]] &= \sum_{n \geq 1} E[X|Y = y_n] p_Y(n) = \sum_{n \geq 1} \left(\sum_{m \geq 1} x_m \frac{p(m, n)}{p_Y(n)} \right) p_Y(n) \\ &= \sum_{m \geq 1} x_m \sum_{n \geq 1} p(m, n) = \sum_{m \geq 1} x_m p_X(m) = E[X]. \end{aligned}$$

b) Sean X e Y variables aleatorias conjuntamente continuas con función de densidad conjunta $f(x, y)$. Entonces

$$\begin{aligned} E[E[X|Y]] &= \int_{-\infty}^{\infty} E[X|Y=y] f_Y(y) dy = \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x f_{X|Y}(x|y) dx \right) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} \left(\int_{-\infty}^{\infty} x \frac{f(x, y)}{f_Y(y)} dx \right) f_Y(y) dy = \int_{-\infty}^{\infty} x \int_{-\infty}^{\infty} f(x, y) dy dx \\ &= \int_{-\infty}^{\infty} x f_X(x) dx = E[X]. \end{aligned}$$

Ejercicios 4.2

1. Se escoge un número X aleatoriamente de el conjunto $A = \{1, 2, 3, 4, 5\}$. Luego se escoge un número Y del subconjunto de A de números no mayores que X .
 i) Halle la función de probabilidad condicional de X dado que $Y = i$, $i = 1, \dots, 5$.
 ii) Diga si X e Y son independientes y porqué.
2. Sea U una variable aleatoria uniformemente distribuida sobre $(0, 1)$. Calcule la distribución condicional de U dado que: a) $U > a$; b) $U < a$; donde $0 < a < 1$.
3. La densidad de probabilidad conjunta de X y Y está dada por

$$f(x, y) = \frac{6}{7} \left(x^2 + \frac{xy}{2} \right), \quad 0 < x < 1, \quad 0 < y < 2.$$

Encuentre $P(Y > \frac{1}{2} | X < \frac{1}{2})$.

4. Se lanzan dos dados perfectos y distinguibles. Sea X = el mayor valor obtenido y sea Y = el menor valor obtenido. Halle la función de probabilidad condicional de Y dado que $X = i$, $i = 1, 2, \dots, 6$. Diga si las variables aleatorias son independientes. Justifique su respuesta.
5. La función de probabilidad conjunta de X e Y está dada por:

$$p(1, 1) = \frac{1}{8}; \quad p(1, 2) = \frac{1}{4}; \quad p(2, 1) = \frac{1}{8}; \quad p(2, 2) = \frac{1}{2}.$$

Halle la función de probabilidad de X dado que $Y = i$ para $i = 1, 2$.

6. La función de densidad conjunta de X e Y está dada por

$$f(x, y) = xe^{-x(y+1)}, \quad x > 0, y > 0.$$

- i) Encuentre la densidad condicional de X dado que $Y = y$ y la densidad condicional de Y dado que $X = x$.
 ii) Encuentre la función de densidad de $Z = XY$.

7. La función de densidad conjunta de X e Y está dada por

$$f(x, y) = c(y^2 - x^2)e^{-y}, \quad -y \leq x \leq y, \quad 0 < y < \infty.$$

Encuentre la distribución condicional de X dado que $Y = y$.

8. Un dado perfecto es lanzado sucesivamente. Sean $X = \#$ lanzamientos para obtener un 6 e $Y = \#$ lanzamientos para obtener un 5. Encuentre (a) $E[X]$; (b) $E[X|Y = 1]$; y (c) $E[X|Y = 5]$.
 9. Una urna contiene 4 bolas blancas y 6 negras. Se sacan, sin reemplazamiento, 3 bolas; y luego, también sin reemplazamiento 5 bolas. Sean $X = \#$ bolas blancas en las primeras 3 extracciones e $Y = \#$ bolas blancas en las últimas 5 extracciones. Halle $E[X|Y = i]$, para $i = 1, 2, 3, 4$.

10. La densidad conjunta de X e Y está dada por

$$f(x, y) = \frac{e^{-\frac{x}{y}} e^{-y}}{y}, \quad 0 < x < \infty; \quad 0 < y < \infty.$$

Calcule $E[X^2|Y = y]$.

11. La densidad conjunta de X e Y está dada por

$$f(x, y) = \frac{e^{-y}}{y}, \quad 0 < x < y; \quad 0 < y < \infty.$$

Calcule $E[X^3|Y = y]$.

Capítulo 5

Variables Aleatorias Independientes.

5.1. Indenpendencia

La independencia de variables aleatorias se define en términos de la independencia entre eventos. Recordamos que eventos E_1 y E_2 son independientes si y sólo si

$$P(E_1 \cap E_2) = P(E_1)P(E_2).$$

Definición 5.1 *Dos variables aleatorias X e Y son independientes si y sólo si para conjuntos de Borel cualesquiera $A \subseteq \mathbf{R}$ y $B \subseteq \mathbf{R}$ se cumple que*

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

Observe que de acuerdo a la definición anterior dos variables aleatorias X e Y son independientes si y sólo si lo son los eventos asociados a ellas: $X^{-1}(A)$ y $Y^{-1}(B)$, donde A y B son conjuntos de Borel cualesquiera.

Ejemplos 5.1

1. *Supongamos que se tienen $m + n$ ensayos de Bernoulli independientes. Sea $X = \#$ de éxitos en los n primeros ensayos e $Y = \#$ de éxitos en los últimos m ensayos. Entonces X e Y son independientes. Si definimos $Z = \#$ de éxitos en los $m+n$ ensayos tendremos que X y Z no son independientes, así como tampoco lo son Y y Z .*

2. (Problema del coleccionista) *Un niño está coleccionando 10 muñecos ofrecidos en la promoción de un producto alimenticio de consumo masivo. Si en cada compra del producto se puede obtener con igual probabilidad cualquiera de los muñecos, ¿cuál es el número esperado de compras que necesita hacer el niño para completar la colección?*

Una vez que se obtiene el primer muñeco, se consideran las compras sucesivas como ensayos de Bernoulli en los que éxito significa obtener un muñeco distinto a los anteriores. Por lo tanto definiendo $X_1 = \#$ de compras hasta obtener un muñeco distinto al primero, la probabilidad de éxito en cada compra será igual a $p_1 = \frac{9}{10}$. Una vez que se obtiene el segundo muñeco distinto, las compras sucesivas son ensayos de Bernoulli independientes con probabilidad de éxito igual a $p_2 = \frac{8}{10}$, y si se continua de esta forma hasta obtener todos los muñecos se tendrá que el número de compras necesarias para completar la colección es igual a $X = 1 + X_1 + X_2 + \dots + X_9$ donde $X_i = \#$ compras necesarias para obtener un muñeco distinto a todos los ya obtenidos siendo las X_i 's variables aleatorias independientes geométricamente distribuidas con parámetro (probabilidad de éxito) $p_i = \frac{10-i}{10}$ para $i = 1, \dots, 9$. En vista de esto se tiene que

$$E[X] = 1 + E[X_1] + \dots + E[X_9] = 1 + \frac{10}{9} + \frac{10}{8} + \dots + 10 = 29, 25.$$

Por lo tanto el número esperado de compras necesarias para completar la colección es 29.

5.1.1. Distribución Conjunta de Variables Aleatorias Independientes.

Proposición 5.1 Sean X e Y variables aleatorias independientes con función de distribución conjunta $F(a, b) = P(X \leq a, Y \leq b)$. Entonces

$$F(a, b) = P(X \leq a)P(Y \leq b) = F_X(a)F_Y(b), \text{ para todo par } (a, b).$$

Si X e Y son conjuntamente continuas e independientes con densidad conjunta f entonces se cumple que $f(x, y) = f_X(x)f_Y(y)$.

Ejemplos 5.2

1. Un hombre y una mujer deciden encontrarse en un lugar. Si cada uno de ellos llega a una hora uniformemente distribuida entre las 12 m y la 1 pm, encuentre la probabilidad de que el primero en llegar tenga que esperar más de 10 minutos.

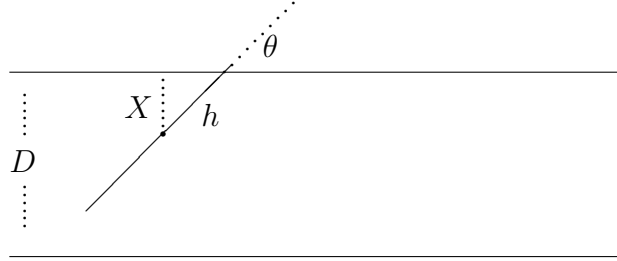
Sean $X =$ tiempo, después de las 12m, que tarda el hombre en llegar y $Y =$ tiempo, después de las 12m, que tarda la mujer en llegar. Entonces X e Y son variables aleatorias independientes, cada una uniformemente distribuida en un intervalo de 60 minutos. Por lo tanto se quiere determinar

$$P(|X-Y| > 10) = P(X-Y > 10) + P(X-Y < -10) = 2 \int_{10}^{60} \int_0^{y-10} \left(\frac{1}{60}\right)^2 dx dy = \frac{25}{36}.$$

2. El Problema de la Aguja de Buffon (1777). Suponga que en una superficie se trazan rectas paralelas de manera tal que la distancia entre dos rectas es igual a una cantidad

fija D . Una aguja de longitud $L \leq D$ se lanza aleatoriamente sobre la superficie. La aguja puede o bien intersectar una de las líneas o caer entre dos de ellas. El problema consiste en encontrar la probabilidad de que la aguja intercepte una de las líneas.

Para resolver el problema observamos que la posición de la aguja con respecto a las líneas está completamente determinada por el ángulo que se forma entre la aguja y las líneas (en dirección positiva $\rightarrow$) y por la distancia entre el punto medio de la aguja y la línea más cercana a ella. Denotaremos estos parámetros por θ y X respectivamente. Tenemos entonces que θ y X son variables aleatorias independientes tales que θ está uniformemente distribuida en $[0, \pi]$ y X está uniformemente distribuida



en $[0, \frac{D}{2}]$.

El segmento perpendicular trazado desde el punto medio de la aguja a la recta más cercana a esta, tiene longitud X . Consideramos el triángulo formado por este segmento, la aguja y la recta 1. Denotando por h la longitud de la hipotenusa de este triángulo, se tiene que la aguja intercepta a la recta si y sólo si $h \leq \frac{L}{2}$. Como $h = \frac{X}{\sin(\theta)}$, la probabilidad de que la aguja intercepte a la recta es

$$\begin{aligned} P\left(\frac{X}{\sin(\theta)} \leq \frac{L}{2}\right) &= \int \int_{\frac{X}{\sin(\theta)} \leq \frac{L}{2}} f(y, x) dy dx = \int \int_{\frac{X}{\sin(\theta)} \leq \frac{L}{2}} f_{\theta}(y) f_X(x) dy dx \\ &= \int_0^{\pi} \int_0^{\frac{L}{2} \sin(y)} \frac{2}{D\pi} dx dy = \frac{2}{D\pi} \int_0^{\pi} \frac{L}{2} \sin(y) dy = \frac{2L}{D\pi}. \end{aligned}$$

5.1.2. Suma de Variables Aleatorias Independientes.

Como vimos antes si X e Y son variables aleatorias conjuntamente continuas con densidad conjunta f y $Z = X + Y$ se tiene que

$$F_{X+Y}(t) = \int_{-\infty}^{\infty} \int_{-\infty}^t f(x, z-x) dz dx = \int_{-\infty}^t \int_{-\infty}^{\infty} f(x, z-x) dx dz = \int_{-\infty}^t f_Z(z) dz;$$

donde $f_Z(z) = \int_{-\infty}^{\infty} f(x, z-x) dx$. Si X e Y son independientes, entonces

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x) dx = f_X * f_Y(z);$$

donde el simbolo $*$ denota el *producto de convolución* de f_X y f_Y .

La siguiente proposición constituye un ejemplo importante de este resultado.

Proposición 5.2 Sean X e Y variables aleatorias independientes con distribución gamma con parámetros (s, λ) y (t, λ) donde s y t son enteros positivos. Entonces $Z = X + Y$ tiene distribución gamma con parámetros $(s + t, \lambda)$.

Demostración. Como X e Y son variables aleatorias independientes, para $a > 0$, se tiene

$$\begin{aligned}
 f_{X+Y}(a) &= \int_{-\infty}^{\infty} f_X(a-y)f_Y(y)dy \\
 &= \frac{1}{\Gamma(s)\Gamma(t)} \int_0^a \lambda e^{-\lambda(a-y)}(\lambda(a-y))^{s-1} \lambda e^{-\lambda y}(\lambda y)^{t-1} dy \\
 &= \frac{\lambda^s \lambda^t}{\Gamma(s)\Gamma(t)} \int_0^a e^{-\lambda a + \lambda y} e^{-\lambda y} (a-y)^{s-1} y^{t-1} dy \\
 &= \frac{\lambda^s \lambda^t e^{-\lambda a}}{\Gamma(s)\Gamma(t)} \int_0^a (a-y)^{s-1} y^{t-1} dy \\
 &= \frac{\lambda^s \lambda^t e^{-\lambda a}}{\Gamma(s)\Gamma(t)} \int_0^a a^{s-1} \left(1 - \frac{y}{a}\right)^{s-1} y^{t-1} dy.
 \end{aligned}$$

Haciendo el cambio de variables $x = \frac{y}{a}$ tenemos que

$$\begin{aligned}
 f_{X+Y}(a) &= \frac{\lambda^s \lambda^t e^{-\lambda a} a^{s+t-1}}{\Gamma(s)\Gamma(t)} \int_0^1 (1-x)^{s-1} x^{t-1} dx \\
 &= e^{-\lambda a} a^{s+t-1} \left(\frac{\lambda^s \lambda^t}{\Gamma(s)\Gamma(t)} \int_0^1 (1-x)^{s-1} x^{t-1} dx \right) = K e^{-\lambda a} a^{s+t-1}.
 \end{aligned}$$

Como f_{X+Y} es una densidad, la constante K debe satisfacer

$$\int_0^{\infty} K e^{-\lambda a} a^{s+t-1} da = 1 \implies K = \frac{1}{\int_0^{\infty} e^{-\lambda a} a^{s+t-1} da};$$

haciendo el cambio de variables $u = \lambda a$ tenemos

$$\begin{aligned}
 \int_0^{\infty} e^{-\lambda a} a^{s+t-1} da &= \int_0^{\infty} e^{-\lambda a} \frac{(\lambda a)^{s+t-1}}{\lambda^{s+t-1}} da = \int_0^{\infty} e^{-u} \frac{u^{s+t-1}}{\lambda^{s+t-1}} \frac{du}{\lambda} \\
 &= \frac{1}{\lambda^{s+t}} \int_0^{\infty} e^{-u} u^{s+t-1} du = \frac{\Gamma(s+t)}{\lambda^{s+t}}.
 \end{aligned}$$

En consecuencia

$$K = \frac{\lambda^{s+t}}{\Gamma(s+t)} = \frac{\lambda^{s+t}}{(s+t-1)!}.$$

Sustituyendo K en la expresión que se dedujo para f_{X+Y}

$$f_{X+Y}(a) = \lambda e^{-\lambda a} \frac{(\lambda a)^{s+t-1}}{\Gamma(s+t)} \implies X+Y \sim \Gamma(s+t, \lambda).$$

La siguiente proposición se obtiene como corolario del teorema 4.1.

Proposición 5.3 Sean X e Y variables aleatorias independientes y sean g y h funciones reales con suficiente regularidad. Entonces se cumple que

$$E[g(X)h(Y)] = E[g(X)]E[h(Y)].$$

En particular $E[XY] = E[X]E[Y]$.

Demostración.

1) Caso discreto. Sean X e Y discretas con función de probabilidad $p(x, y)$. Entonces

$$\begin{aligned} E[g(X)h(Y)] &= \sum_i \sum_j g(x_i)h(y_j)p(x_i, y_j) = \sum_i \sum_j g(x_i)h(y_j)p_X(x_i)p_Y(y_j) \\ &= \sum_i g(x_i)p_X(x_i) \sum_j h(y_j)p_Y(y_j) = E[X]E[Y]. \end{aligned}$$

2) Caso continuo. Supongamos que X e Y son variables aleatorias conjuntamente continuas con densidad conjunta $f(x, y)$. Se tiene que

$$\begin{aligned} E[g(X)h(Y)] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f(x, y)dxdy = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x)h(y)f_X(x)f_Y(y)dxdy \\ &= \int_{-\infty}^{\infty} g(x)f_X(x)dx \int_{-\infty}^{\infty} h(y)f_Y(y)dy = E[g(X)]E[h(Y)]. \end{aligned}$$

Definición 5.2 Sean X e Y variables aleatorias dadas. La **covarianza** entre X e Y , denotada $Cov[X, Y]$, se define como

$$Cov[X, Y] = E[(X - E[X])(Y - E[Y])].$$

Una forma equivalente de esta definición se obtiene mediante el siguiente cálculo

$$\begin{aligned} Cov[X, Y] &= E[(X - E[X])(Y - E[Y])] = E[XY - XE[Y] - E[X]Y + E[X]E[Y]] \\ &= E[XY] - 2E[X]E[Y] + E[X]E[Y] = E[XY] - E[X]E[Y]. \end{aligned}$$

De esta expresión se observa que si X e Y son independientes entonces $Cov[X, Y] = 0$. El recíproco de esta propiedad no es en general cierto.

Ejemplos 5.3

1. Sea X una variable aleatoria con función de probabilidad

$$P(X = 0) = P(X = 1) = P(X = -1) = \frac{1}{3}.$$

Sea Y la variable aleatoria definida como

$$Y = \begin{cases} 0 & \text{si } X \neq 0 \\ 1 & \text{si } X = 0. \end{cases}$$

Entonces $XY = 0$ por lo que $E[XY] = 0$; además $E[X] = \frac{-1}{3} + \frac{1}{3} + 0\frac{1}{3} = 0$ por lo que $Cov[X, Y] = 0$. Sin embargo, X e Y no son independientes ya que:

$$P(X = 0, Y = 0) = 0 \neq P(X = 0)P(Y = 0) = \frac{1}{3} \times \frac{2}{3}.$$

La noción de covarianza permite obtener una expresión para la varianza de la suma de variables aleatorias:

$$\begin{aligned} Var[X + Y] &= E\left[\left(X + Y - E[X + Y]\right)^2\right] = E\left[\left(X - E[X] + Y - E[Y]\right)^2\right] \\ &= E\left[\left(X - E[X]\right)^2 + \left(Y - E[Y]\right)^2 + 2\left(X - E[X]\right)\left(Y - E[Y]\right)\right] \\ &= Var[X] + Var[Y] + 2Cov[X, Y]. \end{aligned}$$

Inductivamente puede verificarse que para variables aleatorias $X_1, X_2, \dots, X_n$

$$Var\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n Var[X_i] + 2 \sum_{i < j} Cov[X_i, X_j].$$

De estas expresiones se observa que, para variables aleatorias independientes la varianza de la suma es la suma de las varianzas:

$$Var\left[\sum_{i=1}^n X_i\right] = \sum_{i=1}^n Var[X_i].$$

Definición 5.3 Dadas dos variables aleatorias X e Y el **coeficiente de correlación ó correlación de X e Y** , denotado por $\rho(X, Y)$, se define como

$$\rho(X, Y) = \frac{Cov[X, Y]}{\sqrt{Var[X]Var[Y]}} \quad \text{siempre que } Var[X] \neq 0 \text{ y } Var[Y] \neq 0.$$

Proposición 5.4 Dadas variables aleatorias X e Y , se cumple que

$$-1 \leq \rho(X, Y) \leq 1.$$

Demostración. Sean $\sigma_X^2 = \text{Var}[X]$ y $\sigma_Y^2 = \text{Var}[Y]$, entonces

$$0 \leq \text{Var}\left[\frac{X}{\sigma_X} + \frac{Y}{\sigma_Y}\right] = \frac{\text{Var}[X]}{\sigma_X^2} + \frac{\text{Var}[Y]}{\sigma_Y^2} + 2\frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y} = 1 + 1 + 2\rho(X, Y);$$

por lo tanto $\rho(X, Y) \geq -1$. Por otra parte

$$0 \leq \text{Var}\left[\frac{X}{\sigma_X} - \frac{Y}{\sigma_Y}\right] = \frac{\text{Var}[X]}{\sigma_X^2} + \frac{\text{Var}[Y]}{\sigma_Y^2} - 2\frac{\text{Cov}[X, Y]}{\sigma_X \sigma_Y} = 1 + 1 - 2\rho(X, Y);$$

de donde $\rho(X, Y) \leq 1$.

5.2. Funciones de Varias Variables Aleatorias.

El Teorema de la función inversa. Preliminares.

Sea $S \subseteq \mathbf{R}^n$ abierto y $g : S \rightarrow \mathbf{R}^n$; $g = (g_1, g_2, \dots, g_n)$. Diremos que $g \in C^1$ si todas las componentes de g tienen derivadas parciales continuas; es decir, para cada $i \in \{1, \dots, n\}$, $\frac{\delta g}{\delta x_j}$ existe y es continua para $j \in \{1, \dots, n\}$.

Si $x = (x_1, x_2, \dots, x_n)$, se define la *matriz jacobiana de g en x* , como la matriz $n \times n$

$$D_g(x) = \left(\frac{\delta g}{\delta x_j}(x) \right)_{n \times n} = \begin{pmatrix} \frac{\delta g_1}{\delta x_1}(x) & \frac{\delta g_1}{\delta x_2}(x) & \dots & \frac{\delta g_1}{\delta x_n}(x) \\ \frac{\delta g_2}{\delta x_1}(x) & \frac{\delta g_2}{\delta x_2}(x) & \dots & \frac{\delta g_2}{\delta x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta g_n}{\delta x_1}(x) & \frac{\delta g_n}{\delta x_2}(x) & \dots & \frac{\delta g_n}{\delta x_n}(x) \end{pmatrix}_{n \times n} \quad (5.1)$$

El determinante de la matriz jacobiana se denomina *determinante jacobiano de g* (o simplemente jacobiano de g y se denota por

$$J_g(x) = \begin{vmatrix} \frac{\delta g_1}{\delta x_1} & \frac{\delta g_1}{\delta x_2} & \dots & \frac{\delta g_1}{\delta x_n} \\ \frac{\delta g_2}{\delta x_1} & \frac{\delta g_2}{\delta x_2} & \dots & \frac{\delta g_2}{\delta x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\delta g_n}{\delta x_1} & \frac{\delta g_n}{\delta x_2} & \dots & \frac{\delta g_n}{\delta x_n} \end{vmatrix} \quad (5.2)$$

Ejemplos 5.4

Sea $g : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ definida por

$$g(x, y) = (e^{x+2y}, \sin(y+2x)) = (g_1(x, y), g_2(x, y)).$$

El jacobiano de g en el punto $(\frac{\pi}{2}, 0)$ es el determinante

$$J_g\left(\frac{\pi}{2}, 0\right) = \begin{vmatrix} e^{\frac{\pi}{2}} & 2e^{\frac{\pi}{2}} \\ 2\cos(\pi) & \cos(\pi) \end{vmatrix} = 3e^{\frac{\pi}{2}} \quad (5.3)$$

Teorema 5.1 (de la función inversa) Sea $S \subseteq \mathbf{R}^n$ abierto y $g : S \rightarrow \mathbf{R}^n$; $g = (g_1, g_2, \dots, g_n) \in C^1$ y sea $T = g(S)$. Si $a \in S$ es un punto tal que el determinante jacobiano de g en a es no nulo ($J_g(a) \neq 0$), entonces existen conjuntos abiertos $U \subseteq S$ y $V \subseteq T$ y una función h univocamente determinada tales que

- i) $a \in U$ y $g(a) \in V$.
- ii) $g(U) = V$.
- iii) g es inyectiva en U .
- iv) h está definida en V , $h(V) = U$ y $h(g(x)) = x$ para todo $x \in U$.
- v) $h \in C^1$ en V .

Teorema 5.2 (de cambio de variables) Sea $V \subseteq \mathbf{R}^n$ abierto y sea $h : V \subseteq \mathbf{R}^n \rightarrow \mathbf{R}^n$ tal que

- i) $h \in C^1$ en U .
 - ii) h es inyectiva en V .
 - iii) El determinante jacobiano $J_h(y) \neq 0$ para todo $y \in V$.
- Sea además $f : h(V) \subseteq \mathbf{R}^n \rightarrow \mathbf{R}$ una función tal que existe

$$\int_{h(V)} f(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots x_n = \int_{h(V)} f(x) dx.$$

Entonces

$$\int_{h(V)} f(x) dx = \int_V f(h(y)) |J_h(y)| dy = \int_V f(h(y_1, \dots, y_n)) |J_h(y_1, \dots, y_n)| dy_1 \dots dy_n.$$

Una aplicación directa de estos teoremas se tiene en la siguiente situación.

Sean $X_1, X_2, \dots, X_n$ variables aleatorias conjuntamente continuas con función de densidad $f : \mathbf{R}^n \rightarrow \mathbf{R}$ y sea $g : \mathbf{R}^n \rightarrow \mathbf{R}^n$ una función en C^1 que satisface las condiciones del teorema de la función inversa sobre $U \subseteq \mathbf{R}^n$, donde U es tal que $f|_{U^c} \equiv 0$. Definamos

$$Y = g(X_1, X_2, \dots, X_n) = g(X) = (g_1(X_1, X_2, \dots, X_n), \dots, g_n(X_1, X_2, \dots, X_n)),$$

y sea $h : V \rightarrow U$ la inversa local de g (dada por el teorema de la función inversa). La variable aleatoria Y puede interpretarse como un cambio de coordenadas aplicado a X y podemos obtener su función de densidad en vista de que

$$P(Y \in V) = P(g(X) \in V) = P(X \in h(V)) = \int_{h(V)} f(x) dx = \int_V f(h(y)) |J_h(y)| dy.$$

Por lo tanto

$$f_Y(y_1, \dots, y_n) = \begin{cases} f(h_1(y_1, \dots, y_n), \dots, h_n(y_1, \dots, y_n)) |J_h(y_1, \dots, y_n)| & \text{si } Y \in V \\ 0 & \text{si } Y \notin V \end{cases} \quad (5.4)$$

Ejemplos 5.5

1. Sean X_1, X_2 variables aleatorias conjuntamente continuas con densidad conjunta f_{X_1, X_2} . Sean $Y_1 = X_1 + X_2$ y $Y_2 = X_1 - X_2$. Para obtener la densidad conjunta de (Y_1, Y_2) procedemos como sigue. Sean

$$y_1 = g_1(x_1, x_2) = x_1 + x_2 \quad y \quad y_2 = g_2(x_1, x_2) = x_1 - x_2.$$

Entonces

$$y_1 + y_2 = x_1 + x_2 + x_1 - x_2 = 2x_1 \quad \implies \quad x_1 = \frac{y_1 + y_2}{2} = h_1(y_1, y_2) \quad y$$

$$y_1 - y_2 = x_1 + x_2 - x_1 + x_2 = 2x_2 \quad \implies \quad x_2 = \frac{y_1 - y_2}{2} = h_2(y_1, y_2).$$

Además tenemos que

$$J_h(y_1, y_2) = \left(J_g(x_1, x_2) \right)^{-1} = \begin{vmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} \end{vmatrix}^{-1} = \begin{vmatrix} 1 & 1 \\ 1 & -1 \end{vmatrix}^{-1} = \frac{-1}{2}. \quad (5.5)$$

Por lo tanto

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} \frac{1}{2} f_{X_1, X_2} \left(\frac{y_1 + y_2}{2}, \frac{y_1 - y_2}{2} \right) & \text{si } y_1 = g_1(x_1, x_2) \text{ y } y_2 = g_2(x_1, x_2) \\ 0 & \text{en otro caso} \end{cases} \quad (5.6)$$

En particular si X_1 y X_2 son independientes y ambas uniformemente distribuidas sobre $(0, 1)$, tendremos que

$$\begin{aligned} f_{Y_1, Y_2}(y_1, y_2) &= \begin{cases} \frac{1}{2} & \text{si } 0 < \frac{y_1 + y_2}{2} < 1 \text{ y } 0 < \frac{y_1 - y_2}{2} < 1 \\ 0 & \text{en otro caso} \end{cases} \\ &= \begin{cases} \frac{1}{2} & \text{si } 0 < y_1 + y_2 < 2 \text{ y } 0 < y_1 - y_2 < 2 \\ 0 & \text{en otro caso} \end{cases} \end{aligned}$$

Si X_1 y X_2 son independientes con densidad exponencial con parámetros λ_1 y λ_2 respectivamente, entonces

$$f_{Y_1, Y_2}(y_1, y_2) = \begin{cases} \frac{\lambda_1 \lambda_2}{2} e^{\frac{\lambda_1}{2}(y_1 + y_2)} e^{\frac{\lambda_2}{2}(y_1 - y_2)} & \text{si } 0 < y_1 + y_2 \text{ y } 0 < y_1 - y_2 \\ 0 & \text{en otro caso} \end{cases} \quad (5.7)$$

2. Sean X e Y variables aleatorias independientes, cada una con distribución gamma con parámetros (α, λ) y (β, λ) respectivamente. Hallaremos la densidad conjunta de U y V cuando $U = X + Y$ y $V = \frac{X}{X+Y}$.

Tenemos que

$$u = g_1(x, y) = x + y \quad y \quad v = g_2(x, y) = \frac{x}{x + y};$$

por lo tanto

$$x = u \times v = h_1(u, v) \quad y \quad y = u - x = u - u \times v = u(1 - v) = h_2(u, v).$$

Por otra parte tenemos que

$$\begin{aligned} J_h(u, v) = \left(J_g(x, y) \right)^{-1} &= \begin{vmatrix} \frac{\delta g_1}{\delta x} & \frac{\delta g_1}{\delta y} \\ \frac{\delta g_2}{\delta x} & \frac{\delta g_2}{\delta y} \end{vmatrix}^{-1} = \begin{vmatrix} 1 & 1 \\ \frac{(x+y)-x}{(x+y)^2} & \frac{0-x}{(x+y)^2} \end{vmatrix}^{-1} = \begin{vmatrix} 1 & 1 \\ \frac{y}{(x+y)^2} & \frac{x}{(x+y)^2} \end{vmatrix}^{-1} \\ &= \left(\frac{-(x+y)}{(x+y)^2} \right)^{-1} = -(x+y) = -u. \end{aligned}$$

Entonces

$$\begin{aligned} f_{U,V}(u, v) &= \begin{cases} f_{X,Y}(uv, u(1-v))u & \text{si } 0 < v < 1 \text{ y } 0 \leq u \\ 0 & \text{en otro caso} \end{cases} \\ &= \begin{cases} \frac{\lambda e^{-\lambda uv} (\lambda uv)^{\alpha-1}}{\Gamma(\alpha)} \times \frac{\lambda e^{-\lambda u(1-v)} (\lambda u(1-v))^{\beta-1}}{\Gamma(\beta)} u & \text{si } 0 < v < 1 \text{ y } 0 \leq u \\ 0 & \text{en otro caso} \end{cases} \\ &= \begin{cases} \frac{\lambda e^{-\lambda u} (\lambda uv)^{\alpha+\beta-1}}{\Gamma(\alpha+\beta)} \times \frac{v^{\alpha-1} (1-v)^{\beta-1} \Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} u & \text{si } 0 < v < 1 \text{ y } 0 \leq u \\ 0 & \text{en otro caso.} \end{cases} \end{aligned}$$

Ejercicios 5.1

1. Un grupo de N hombres lanza sus sombreros en el centro de un salón. Los sombreros se mezclan y cada hombre selecciona uno al azar. a) Encuentre el número esperado de hombres que selecciona su propio sombrero. b) Calcule la varianza del número de hombres que selecciona su sombrero. Sugerencia: considere las variables aleatorias definidas como:

$$X_i = \begin{cases} 1 & \text{si el } i\text{-ésimo hombre selecciona su sombrero} \\ 0 & \text{si no} \end{cases} \quad (5.8)$$

para $i = 1, \dots, N$.

2. Si X e Y son idénticamente distribuidas, no necesariamente independientes, demuestre que $\text{Cov}[X + Y, X - Y] = 0$.
3. Demuestre que: a) $\text{Cov}[a + bX, c + dY] = bd\text{Cov}[X, Y]$.
b) $\text{Cov}[X + Y, Z] = \text{Cov}[X, Z] + \text{Cov}[Y, Z]$.
c) $\text{Cov}[\sum_{i=1}^n X_i, \sum_{j=1}^m Y_j] = \sum_{i=1}^n \sum_{j=1}^m \text{Cov}[X_i, Y_j]$.

4. Sean $X = \#$ de veces que aparece 1 en n lanzamientos de un dado simétrico y $Y = \#$ de veces que aparece 2 en n lanzamientos de un dado simétrico. Halle la $Cov[X, Y]$. Use la parte c) del ejercicio anterior.
5. Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes e idénticamente distribuidas con media μ y varianza σ^2 . Sea $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$. Demuestre que: a) $E[\bar{X}] = \mu$. b) $Var[\bar{X}] = \sigma^2/n$. c) $E[\sum_{i=1}^n (X_i - \bar{X})^2] = (n-1)\sigma^2$.

6. Sea $Y = a + bX$. Demuestre que

$$\rho(X, Y) = \begin{cases} 1 & \text{si } b > 0 \\ -1 & \text{si } b < 0 \end{cases} \quad (5.9)$$

7. Demuestre la desigualdad de Cauchy-Schwartz: $(E[XY])^2 \leq E[X^2]E[Y^2]$.
Sugerencia: Para cada $t \in \mathbf{R}$ considere $E[(tX + Y)^2] \geq 0$. Obtenga una ecuación cuadrática en t con a lo sumo una raíz real. Esto implica que el discriminante de esta ecuación es menor ó igual a cero.

8. Verifique que $Cov[X, Y] = 1/8$ si (X, Y) son variables aleatorias con densidad conjunta

$$f(x, y) = \begin{cases} \frac{2e^{-2x}}{x} & \text{si } 0 < x < \infty, 0 \leq y \leq x \\ 0 & \text{en otros casos} \end{cases} \quad (5.10)$$

9. La función de densidad conjunta de X e Y está dada por

$$f(x, y) = \begin{cases} \frac{1}{y} e^{-(y + \frac{x}{y})} & \text{si } 0 < x < \infty, 0 \leq y < \infty \\ 0 & \text{en otros casos} \end{cases} \quad (5.11)$$

Demuestre que $Cov[X, Y] = 1$.

10. Sean X_1, X_2, X_3, X_4 variables aleatorias no correlacionadas, es decir $\rho(X_i, X_j) = 0$ para $i \neq j$. Suponga que todas estas variables tienen esperanza cero y varianza uno. Calcule las correlaciones de: a) $X_1 + X_2$ y $X_2 + X_3$. b) $X_1 + X_2$ y $X_3 + X_4$.

11. Halle la convolución $f * g$ en cada uno de los casos siguientes:

$$a) f(x) = \begin{cases} 1 & \text{si } x \in [1, 2] \\ 0 & \text{en otros casos} \end{cases} \quad (5.12)$$

$$g(x) = \begin{cases} 1 & \text{si } x \in [-2, -1] \\ 0 & \text{en otros casos} \end{cases} \quad (5.13)$$

$$b) f(x) = g(x) = \begin{cases} 1 & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases} \quad (5.14)$$

¿Qué representan las convoluciones anteriores probabilísticamente?

12. Sean X e Y variables aleatorias binomiales con parámetros (n, p) y (m, p) respectivamente. Asuma que X e Y son independientes y calcule analíticamente la distribución de $X + Y$.

13. Use el ejercicio anterior y un razonamiento intuitivo para demostrar que

$$\binom{m+n}{k} = \sum_{i=0}^k \binom{n}{i} \binom{m}{k-i}$$

donde $k \leq n$ y $k \leq m$.

14. Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes todas ellas exponencialmente distribuidas con parámetro λ . Halle la distribución de $X_1 + X_2 + \dots + X_n$. Sugerencia: Observe la relación entre las variables gamma y exponencial.

15. Sean X_1 y X_2 variables aleatorias independientes con distribución de Poisson con parámetros λ_1 y λ_2 respectivamente. Demuestre que $X_1 + X_2$ tiene distribución de Poisson con parámetro $\lambda_1 + \lambda_2$.

16. Si X está uniformemente distribuida sobre $(0, 1)$ y Y está exponencialmente distribuida con parámetro $\lambda = 1$, encuentre la distribución de a) $Z = X + Y$, b) $Z = X/Y$. Asuma independencia.

17. Si X e Y son variables aleatorias independientes ambas uniformemente distribuidas sobre $(0, 1)$, encuentre la función de densidad conjunta de $R = \sqrt{X^2 + Y^2}$ y $\theta = \tan^{-1}(Y/X)$.

18. Si X e Y tienen función de densidad conjunta

$$f(x, y) = \frac{1}{x^2 y^2} \text{ para } x \geq 1, y \geq 1,$$

a) Calcule la función de densidad conjunta de $U = XY$, $V = X/Y$. b) Halle las densidades marginales.

19. Sean X e Y variables aleatorias independientes e idénticamente distribuidas con distribución uniforme sobre $(0, 1)$. Calcule la función de densidad conjunta de: a) $U = X+Y$, $V = X/Y$. b) $U = X$, $V = X/Y$. c) $U = X+Y$, $V = X/(X+Y)$.

20. Repita el ejercicio anterior en el caso en que X e Y tienen distribución exponencial con parámetro $\lambda = 1$.

21. Sean X_1 y X_2 variables aleatorias independientes e idénticamente distribuidas con distribución exponencial con parámetro λ . Encuentre la función de densidad conjunta de $Y_1 = X_1 + X_2$ y $Y_2 = e^{X_1}$.

22. Sean X , Y y Z variables aleatorias independientes con idéntica función de densidad dada por $f(x) = e^{-x}$, $0 < x < \infty$. Calcule la función de densidad conjunta de $U = X + Y$, $V = X + Z$ y $W = Y + Z$.
23. Se lanzan dos dados perfectos, uno rojo y uno verde. Sean X = resultado del dado rojo y Y = menor de los resultados. Halle la función de probabilidad condicional de X dado que $Y = y$.
24. La función de probabilidad conjunta de X e Y está dada por

$$f(x, y) = e^{-(x+y)}, \quad 0 \leq x < \infty, 0 \leq y < \infty.$$

Halle la función de densidad de $Z = \frac{X}{Y}$.

25. Se estima que 45 % de los clientes que entran a una tienda de televisores comprará un televisor normal, 15 % comprará un televisor con reproductor de video incorporado y 40 % no comprará nada. Si 5 clientes entran hoy, ¿cuál es la probabilidad de que se vendan exactamente 2 televisores normales y 1 con video incorporado?

5.3. Función Generadora de Momentos.

Sea X una variable aleatoria que toma sólo valores enteros no negativos y sea

$$a_j = P(X = j), \quad \text{para } j \geq 0,$$

su función de probabilidad. Sabemos que la función de probabilidad proporciona la información probabilística acerca de la variable aleatoria. Estudiaremos una forma muy útil de condensar esta información.

Para una variable ficticia $z \in \mathbf{R}$ consideremos la serie de potencias en z

$$g(z) = a_0 + a_1 z + a_2 z^2 + \dots = \sum_{i=0}^{\infty} a_i z^i.$$

Como $\sum_{i=0}^{\infty} a_i = 1$ tenemos que

$$|g(z)| \leq \sum_{i=0}^{\infty} |a_i| |z|^i \leq 1 \quad \text{siempre que } |z| \leq 1;$$

por lo que la serie de potencias $g(z)$ converge siempre que $|z| \leq 1$. Esto implica que para $|z| \leq 1$ la serie $g(z)$ puede ser diferenciada "término a término". Evaluando $g(z)$

y sus sucesivas derivadas en $z = 0$ se obtiene

$$\begin{aligned}
 g(0) &= a_0 \\
 g'(z) &= a_1 + 2a_2z + 3a_3z^2 + \dots \implies g'(0) = a_1 \\
 g''(z) &= 2a_2 + 3 \times 2a_3z + 4 \times 3a_4z^2 \dots \implies g''(0) = 2a_2 \\
 g'''(z) &= 3 \times 2a_3 + 4 \times 3 \times 2a_4z \dots \implies g'''(0) = 3 \times 2a_3 \\
 &\vdots \\
 g^{(n)}(z) &= n!a_n + n!(n+1)a_{n+1}z + n!(n+1)(n+2)z^2 + \dots \implies g^{(n)}(0) = n!a_n \\
 &\dots
 \end{aligned}$$

Por lo tanto, a partir de la función g se puede obtener la función de probabilidad de X evaluando en 0 las sucesivas derivadas de g ; esto significa que la función g en cierto modo caracteriza la función de distribución de X .

Definición 5.4 Dada una variable aleatoria X cuyos valores son enteros no negativos, se define la **función generadora de momentos ó función generatriz** de X como la serie de potencias

$$g(z) = \sum_{i=0}^{\infty} a_i z^i.$$

Definición 5.5 Dada una variable aleatoria X y un entero $k > 0$, la esperanza $E[X^k]$ se denomina **momento de orden k ó k -ésimo momento** de X . Así mismo la esperanza $E[(X - E[X])^k]$ se denomina **momento centrado de orden k de X** .

Si evaluamos las derivadas de $g(z)$ en $z = 1$ tenemos que

$$g'(z) = a_1 + 2a_2z + 3a_3z^2 + \dots \implies g'(1) = \sum_{i=1}^{\infty} i a_i = E[X]$$

$$g''(1) = \sum_{i=0}^{\infty} i(i-1)a_i = \sum_{i=0}^{\infty} i^2 a_i - \sum_{i=0}^{\infty} i a_i = E[X^2] - E[X]$$

$$\Rightarrow E[X^2] = g''(1) + E[X]$$

$$\begin{aligned}
 g'''(1) &= \sum_{i=0}^{\infty} i(i-1)(i-2)a_i = \sum_{i=0}^{\infty} i(i^2 - 3i + 2)a_i \\
 &= \sum_{i=0}^{\infty} i^3 a_i - 3 \sum_{i=0}^{\infty} i^2 a_i + 2 \sum_{i=0}^{\infty} i a_i = E[X^3] - 3E[X^2] + 2E[X]
 \end{aligned}$$

$\vdots$

Esto significa que también pueden obtenerse los momentos de cualquier orden evaluando las derivadas de la función generadora de momentos en $z = 1$

Ejemplos 5.6

1. Sea X una variable aleatoria binomial, $X \sim b(n, p)$. Entonces la función generadora de momentos de X es igual a

$$g(z) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} z^k = \sum_{k=0}^n \binom{n}{k} (zp)^k (1-p)^{n-k} = (pz + 1 - p)^n.$$

Por lo que

$$\begin{aligned} g'(z) &= np(pz + 1 - p)^{n-1} \implies g'(1) = np = E[X] \\ g''(z) &= n(n-1)p^2(pz + 1 - p)^{n-2} \implies g''(1) = n(n-1)p^2 \end{aligned}$$

En este caso

$$E[X^2] = n(n-1)p^2 + np = n^2p^2 - np^2 + np \implies \text{Var}[X] = E[X^2] - (E[X])^2 = np(1-p).$$

2. Sea X una variable aleatoria geométrica con parámetro p . La función generadora de momentos de X es

$$g(z) = \sum_{k=1}^{\infty} p(1-p)^{k-1} z^k = \frac{p}{1-p} \sum_{k=1}^{\infty} [(1-p)z]^k = \frac{pz}{1 - (1-p)z}.$$

Teorema 5.3 La distribución de probabilidad de una variable aleatoria con valores enteros no negativos está unívocamente determinada por su función generatriz.

Demostración. Por una parte sabemos que dada la función generatriz de X , está dada por

$$g_X(z) = \sum_{i=0}^{\infty} a_i z^i,$$

la función de probabilidad de X se obtiene como

$$P(X = j) = \frac{g_X^{(j)}(0)}{j!}.$$

Supongamos ahora que Y es una variable aleatoria con valores enteros no negativos y con función de probabilidad dada por $P(Y = j) = b_j$, $j \geq 0$. Si Y tiene la misma función generatriz que X entonces

$$g_X(z) = \sum_{i=0}^{\infty} a_i z^i = g_Y(z) = \sum_{i=0}^{\infty} b_i z^i \implies d(z) = \sum_{i=0}^{\infty} (b_i - a_i) z^i = 0.$$

Esto implica que $d^{(j)}(z) = 0$, en particular $d^{(j)}(0) = j!(b_j - a_j) = 0$; es decir, $a_j = b_j$ para todo $j \geq 0$, por lo que X e Y están idénticamente distribuidas.

Proposición 5.5 Sean $X_1, \dots, X_n$ variables aleatorias independientes con valores enteros no negativos y sean $g_1, \dots, g_n$ sus funciones generatrices. Entonces la función generatriz de la suma $X_1 + \dots + X_n$ es el producto de las funciones generatrices $g_1 \times g_2 \times \dots \times g_n$.

Demostración 1. Supongamos que $n = 2$. Sean $\bar{a} = (a_j)_{j \geq 1}$ y $\bar{b} = (b_j)_{j \geq 1}$, donde para cada $j \geq 1$, $a_j = P(X_1 = j)$ y $b_j = P(X_2 = j)$. Entonces

$$\begin{aligned} P(X_1 + X_2 = k) &= \sum_{j=0}^k P(X_1 = j, X_2 = k - j) = \sum_{j=0}^k P(X_1 = j)P(X_2 = k - j) \\ &= c_k = \sum_{j=0}^k a_j b_{k-j}; \end{aligned}$$

donde $c_l = (a * b)(l)$ es la convolución (discreta) de $\bar{a}$ y $\bar{b}$. Entonces se tiene que la función generatriz de la suma es

$$s(z) = \sum_{k=0}^{\infty} P(X_1 + X_2 = k)z^k = \sum_{k=0}^{\infty} c_k z^k;$$

y

$$g(z)h(z) = \left(\sum_{j=0}^{\infty} a_j z^j \right) \left(\sum_{k=0}^{\infty} b_k z^k \right) = \sum_{k=0}^{\infty} \sum_{j=0}^{\infty} a_j b_k z^{j+k} = \sum_{l=0}^{\infty} c_l z^l.$$

Por lo tanto el resultado es válido para $n = 2$. Para $n \in \mathbf{N}$ cualquiera el resultado se obtiene por inducción sobre n .

Demostración 2. Podemos escribir para $z \in \mathbf{R}$:

$$g(z) = \sum_{j=0}^{\infty} a_j z^j = \sum_{j=0}^{\infty} P(X = j) z^j = E[z^X],$$

por lo que la función generatriz de la variable aleatoria $X_1 + \dots + X_n$ es igual a

$$E[z^{X_1 + \dots + X_n}] = E[z^{X_1} z^{X_2} \dots z^{X_n}] = E[z^{X_1}] E[z^{X_2}] \dots E[z^{X_n}] = g_1(z) \dots g_n(z)$$

donde g_i es la función generatriz de X_i .

Ejemplos 5.7

1. Sea X la variable aleatoria binomial negativa con parámetros (n, p) , es decir $X = \#$ ensayos requeridos para obtener n éxitos siendo p la probabilidad de éxito en cada ensayo. Entonces $X = X_1 + \dots + X_n$ donde $X_i = \#$ ensayos requeridos, después del éxito $i-1$, para obtener el i -ésimo éxito. Las variables X_i son independientes y tienen distribución geométrica con parámetro p y función generatriz $g(z)$; por lo que la función generadora de momentos de X es igual a

$$h(z) = g(z)^n = \left[\frac{pz}{1 - (1-p)z} \right]^n.$$

Tenemos además que

$$E[X] = h'(1) = n \left(\frac{p^2 + p - p^2}{p^2} \right) = \frac{n}{p}.$$

5.3.1. Cálculo de Momentos.

La representación de la función generadora de momentos utilizada en la segunda demostración de la proposición anterior motiva la siguiente generalización.

Sea $z = e^{-t}$ para $t \in (0, \infty)$ y sea $X > 0$ una variable aleatoria. Se define la función generadora de momentos de X como

$$\phi(t) = E[z^X] = E[e^{-tX}] = \begin{cases} \sum_{\{x: P(X=x) > 0\}} e^{-xt} P(X=x) & \text{si } X \text{ es discreta} \\ \int_{-\infty}^{\infty} e^{-tx} f(x) dx & \text{si } X \text{ es continua} \end{cases}$$

La función ϕ se denomina *transformada de Laplace* de X (ó de la función de densidad f).

Comentarios.

- La función generatriz existe siempre que X tome valores no negativos, en el caso continuo, esto significa que $f(u) = 0$ para todo $u < 0$. De lo contrario la esperanza de Z^X podría no existir.

- Con esta definición de la función generatriz no es tan sencillo obtener los valores de la función de probabilidad o la función de densidad; sin embargo, al igual que en el primer caso estudiado ($X > 0$ toma sólo valores enteros), la distribución de la variable aleatoria X está determinada en forma única por su función generatriz.

- Es posible usar la función generatriz para obtener momentos de cualquier orden. De manera análoga a lo hecho para g se calcula la derivada de ϕ y se evalúa en $t = 0$. En el caso discreto se puede calcular la derivada de la serie correspondiente derivando término a término. De manera similar en el caso continuo la derivada de la integral se calcula como la integral de la derivada del integrando. Esto significa que

$$\phi'(t) = \frac{d}{dt} E[e^{-tX}] = E\left[\frac{d}{dt}(e^{-tX})\right] = E[-Xe^{-tX}].$$

En el caso en que X es una variable aleatoria *discreta* se tiene que

$$E[-Xe^{-tX}] = \sum_{\{x: P(X=x)>0\}} -xe^{-xt} P(X=x).$$

En el caso en que X es una variable aleatoria continua

$$E[-Xe^{-tX}] = \int_{-\infty}^{\infty} -xe^{-tx} f(x) dx.$$

Por lo tanto

$$\begin{aligned}\phi'(0) &= -E[X] \\ \phi''(0) &= E[X^2] \\ &\dots \\ \phi^{(n)}(0) &= (-1)^n E[X^n]\end{aligned}$$

Ejemplos 5.8 Distribución exponencial con parámetro λ

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

$$\phi(t) = E[e^{-tX}] = \int_0^{\infty} e^{-tx} \lambda e^{-\lambda x} dx = \lambda \int_0^{\infty} e^{-(\lambda+t)x} dx = \frac{\lambda}{\lambda+t}$$

$$E[X] = -\phi'(0) = \frac{1}{\lambda}$$

$$E[X^2] = \phi''(0) = \frac{2}{\lambda^2}$$

$$\Rightarrow \text{var}[X] = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

La siguiente proposición establece una propiedad análoga a la enunciada en la proposición 1.

Proposición 5.6 Sean $X_1, X_2, \dots, X_n$ variables aleatorias independientes con funciones generatrices $\phi_1, \phi_2, \dots, \phi_n$. Entonces la función generatriz de $X_1 + X_2 + \dots, X_n$ es igual a $\phi_1 \times \phi_2 \times \dots \times \phi_n$.

Demostración. Para $i \in \{1, 2, \dots, n\}$ se tiene que $\phi_{X_i}(t) = E[e^{-tX_i}]$, por lo que

$$\begin{aligned}\phi_{X_1+X_2+\dots+X_n}(t) &= E[e^{-t(X_1+X_2+\dots+X_n)}] = E[e^{-tX_1}e^{-tX_2}\dots e^{-tX_n}] \\ &= E[e^{-tX_1}]E[e^{-tX_2}]\dots E[e^{-tX_n}] = \phi_{X_1}(t) \times \phi_{X_2}(t) \times \dots \times \phi_{X_n}(t).\end{aligned}$$

Definición General de la Función Generadora de Momentos.

Si se consideran variables aleatorias con valores reales (no necesariamente no negativas), la función generadora de momentos se puede definir de dos maneras:

i) $\phi(t) = E[e^{tX}]$, para $t \in \mathbf{R}$, siempre que la esperanza exista.

Ejemplos 5.9

Sea X una variable aleatoria normalmente distribuida con media y varianza (μ, σ^2) , se tiene que

$$\begin{aligned}\phi_X(t) &= E[e^{tX}] = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{tx} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{2tx}{2}} e^{-\frac{x^2+2x\mu-\mu^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{2tx\sigma^2}{2\sigma^2}} e^{-\frac{x^2+2x\mu-\mu^2}{2\sigma^2}} dx = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\mu^2}{2\sigma^2}} e^{-\frac{x^2+2x\mu+2tx\sigma^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\mu^2}{2\sigma^2}} e^{-\frac{x^2+2x(\mu+t\sigma^2)-(\mu+t\sigma^2)^2+(\mu+t\sigma^2)^2}{2\sigma^2}} dx \\ &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-\mu^2+(\mu+t\sigma^2)^2}{2\sigma^2}} e^{-\frac{x^2-2x(\mu+t\sigma^2)+(\mu+t\sigma^2)^2}{2\sigma^2}} dx \\ &= e^{\frac{-\mu^2+(\mu+t\sigma^2)^2}{2\sigma^2}} \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{(x-(\mu+t\sigma^2))^2}{2\sigma^2}} dx = e^{\mu t + \frac{t^2\sigma^2}{2}}\end{aligned}$$

En particular si $\mu = 0$ $\sigma = 1$ se tiene que $\phi_X(t) = e^{\frac{t^2}{2}}$

Con esta definición de la función generadora de momentos es claro que se cumple la propiedad enunciada en las proposiciones 1 y 2. Sigue siendo igualmente válida la propiedad de que la función generadora de momentos de una variable aleatoria determina de manera univoca su distribución. Estos hechos se combinan para obtener la importante propiedad que sigue.

Proposición 5.7 Si X y Y son variables aleatorias independientes normalmente distribuidas con media y varianza (μ_1, σ_1^2) y (μ_2, σ_2^2) respectivamente se tiene que la suma $X + Y$ también está normalmente distribuida con parámetros $(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

$$\phi_{X+Y}(t) = E[e^{t(X+Y)}] = \phi_X(t)\phi_Y(t) = e^{\mu_1 t + \frac{t^2 \sigma_1^2}{2}} e^{\mu_2 t + \frac{t^2 \sigma_2^2}{2}} = e^{(\mu_1 + \mu_2)t + \frac{t^2(\sigma_1^2 + \sigma_2^2)}{2}}$$

que es la función generadora de momentos de una variable aleatoria normal con media $\mu_1 + \mu_2$ y varianza $\frac{\sigma_1^2 + \sigma_2^2}{2}$, esto equivale a afirmar que $X + Y \sim N(\mu_1 + \mu_2, \frac{\sigma_1^2 + \sigma_2^2}{2})$.

Esta propiedad puede extenderse por inducción a la suma de n variables aleatorias independientes y normalmente distribuidas.

ii) En el caso en que $E[e^{tX}]$ no existe, se define la función generadora de momentos de X como

$$\phi(t) = E[e^{itX}], \quad \text{donde } i = \sqrt{-1}. \quad (5.15)$$

Observemos que utilizando la fórmula de De Moivre: $e^{iu} = \cos(u) + i \sin(u)$ es evidente que $|e^{itX}| = 1$ por lo que la esperanza anterior siempre es finita.

La transformada dada en (4.1) se denomina *transformada de Fourier de X* (ó de la densidad f)

Ejercicios 5.2

Dada una variable aleatoria X , considere la función generadora de momentos de X $\phi(t) = E[e^{tX}]$, $t \in \mathbf{R}$ y la transformada de Laplace de X como $L(t) = E[e^{-tX}]$, $t \geq 0$.

1. Halle la función generadora de momentos de las variables aleatorias siguientes:
 - a) Binomial con parámetros (n, p) , $0 \leq p \leq 1$.
 - b) Poisson con parámetro $\lambda > 0$.
 - c) Geométrica con parámetro $0 \leq p \leq 1$.
 - d) Binomial negativa con parámetros (r, p) , $0 \leq p \leq 1$.
 - e) Uniforme sobre (a, b) .
 - f) Gamma con parámetro $\lambda > 0$.

En cada uno de los casos anteriores derive para verificar las fórmulas de la esperanza y la varianza.

2. Si $Y = aX + b$, donde a y b son constantes, exprese la función generadora de momentos de Y en términos de la función generadora de momentos de X .

3. Halle la transformada de Laplace para cada una de las siguientes densidades:
 - a) $f(u) = \frac{1}{c}$, $u \in [0, c]$.
 - b) f es la densidad uniforme sobre (a, b) .
 - c) $f(u) = \frac{2u}{c^2}$, $u \in (0, c)$, $c > 0$.
 - d) f es la densidad gamma con parámetros (n, λ) , $\lambda > 0$.
4. Sea X una variable aleatoria normal con parámetros (μ, σ^2) .
 - a) Calcule $E[(X - \mu)^3]$.
 - b) Use a) para determinar $E[X^3]$.
 - c) Verifique el resultado obtenido en b) usando la transformada de Laplace o la función generadora de momentos de X .
5. Use la función generadora de momentos para determinar la distribución de $X_1 + X_2 + \dots + X_n$, donde $X_1, \dots, X_n$ son variables aleatorias independientes e idénticamente distribuidas exponencialmente con parámetro λ .
6. La función generadora de momentos de una variable aleatoria X está dada por $g_X(z) = e^{2z-2}$ y la de Y por $g_Y(z) = (\frac{1}{4})^{10}(3z + 1)^{10}$. Si X e Y son independientes, calcule: a) $P(X + Y = 2)$; b) $P(XY = 0)$; c) $E[XY]$.
7. Suponga que la función generadora de momentos de una variable aleatoria X está dada por $g(z) = e^{3(z-1)}$. Calcule $P(X = 0)$. Halle la función de distribución de X .
8. Se lanzan 3 dados simétricos e idénticos. Halle la función generatriz de la suma de los tres dados. Halle la probabilidad de que la suma sea igual a 3.
9. Considere n sucesos independientes A_j , $j = 1, \dots, n$, con $P(A_j) = p_j$ $j = 1, \dots, n$. Denotemos por N el número de estos eventos que ocurren. Halle la función generatriz de N y calcule $E[N]$ usando esta función.

Capítulo 6

Teoremas Limite

En este capítulo, consideraremos algunos resultados relacionados con propiedades asintóticas de sucesiones de variables aleatorias. Comenzaremos con algunos de los resultados fundamentales en la teoría de las probabilidades.

6.1. Leyes de Grandes Números.

6.1.1. Introducción

Hasta ahora hemos definido la probabilidad de un evento, en situaciones discretas, como la frecuencia relativa de ocurrencia del evento. De esta forma se pueden obtener algunas distribuciones empíricas ampliamente usadas. Por ejemplo, con esta definición, nos resulta aceptable el asignar el valor de $1/2$ a la probabilidad de que se obtenga una cara al lanzar una moneda legal. Los teoremas de esta sección constituyen el fundamento teórico de esta práctica y por ello se encuentran entre los resultados fundamentales de la teoría de las probabilidades.

La idea es la siguiente: Suponga que se tiene una sucesión $X_1, X_2, \dots$ de variables aleatorias independientes, idénticamente distribuidas y con esperanza finita

$$E[X_i] = \mu \quad \text{para } i \geq 1,$$

(por ejemplo, variables asociadas con ensayos de Bernoulli).

Sea $S_n = X_1 + X_2 + \dots + X_n$, $n \geq 1$, suma de las n primeras X_i 's.

Estudiaremos el comportamiento de los promedios $\frac{S_n}{n}$ cuando $n \rightarrow \infty$.

La Ley de los Grandes Números establece que estos promedios **se acercan** a $E[X_i] = \mu$, de acuerdo con una noción de proximidad que se expondrá más adelante.

Ahora bien, en términos de las frecuencias relativas, consideremos ensayos independientes en los cuales un evento A ocurre con probabilidad p . Si definimos

$$X_i = \begin{cases} 1 & \text{si } A \text{ ocurre en } i\text{-ésimo ensayo} \\ 0 & \text{si } A \text{ no ocurre en } i\text{-ésimo ensayo} \end{cases}$$

entonces S_n es una variable aleatoria de Bernoulli con parámetros (n, p) y el promedio $\frac{S_n}{n}$, en este caso, es la frecuencia relativa es de casos en que ocurre A .

La Ley de los Grandes Números postula que, para n grande

$$\frac{S_n}{n} \sim E[X_i] = p = P(A);$$

es decir, $P(A) \approx$ frecuencia relativa de A , lo que permite estimar p por $\frac{S_n}{n}$ para n "suficientemente grande". Esto fundamenta los criterios básicos empleados en el cálculo de probabilidades.

En esta sección, estudiaremos la Ley Débil de Grandes Números, la cual se establece para sucesiones de variables aleatorias independientes, idénticamente distribuidas y con valores reales.

6.1.2. Desigualdad de Markov

Proposición 6.1 Si X es una variable aleatoria con valores no negativos, entonces para $a > 0$

$$P(X \geq a) \leq \frac{E[X]}{a}$$

Demostración 1. Caso discreto. Sea X una variable aleatoria discreta con valores $\{x_j\}_{j \geq 1}$ y función de probabilidad dada por $p_j = P(X = x_j)$ para $j \geq 1$, tal que $\sum_j p_j = 1$.

En este caso $E[X] = \sum_{j \geq 1} x_j p_j$. Sea $A = \{j : x_j \geq a\}$ entonces: $P(X \geq a) = \sum_{j \in A} p_j$ y

$$E[X] = \sum_{j \in A} x_j p_j + \sum_{j \in A^c} x_j p_j \geq \sum_{j \in A} x_j p_j \geq \sum_{j \in A} a p_j = a P(X \geq a).$$

Por lo tanto,

$$P(X \geq a) \leq \frac{E[X]}{a}.$$

Caso continuo: Sea X una variable aleatoria continua con función de densidad f . Entonces

$$P(X \geq a) = \int_a^\infty f(x) dx,$$

y se cumple que

$$E[X] = \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^a xf(x)dx + \int_a^{\infty} xf(x)dx \geq \int_a^{\infty} xf(x)dx \geq a \int_a^{\infty} f(x)dx.$$

Esto implica que

$$E[X] \geq aP(X \geq a).$$

Demostración 2. Utilizando la notación de la demostración anterior, sea

$$A = \{j : x_j \geq a\}.$$

Como $X \geq 0$ se tiene que

$$a\chi_A \leq X\chi_A \leq X.$$

Utilizando la monotonía de la esperanza se obtiene

$$E[a\chi_A] \leq E[X\chi_A] \leq E[X].$$

Esto implica que

$$aP(A) = aP(X \geq a) \leq E[X] \implies P(X \geq a) \leq \frac{E[X]}{a}.$$

Proposición 6.2 (Desigualdad de Chebyshev)

Si X es una variable aleatoria con media finita μ y varianza σ^2 , entonces para cualquier $k > 0$

$$P\{|X - \mu| \geq k\} \leq \frac{\sigma^2}{k^2} = \frac{E[(X - \mu)^2]}{k^2}$$

Demostración. Como $(X - \mu)^2 \geq 0$ se puede usar la desigualdad de Markov con $Y = (X - \mu)^2$, $a = k^2 > 0$. Se obtiene:

$$P(|X - \mu| \geq k) = P((X - \mu)^2 \geq k^2) \leq \frac{E[(X - \mu)^2]}{k^2} = \frac{\sigma^2}{k^2}.$$

Observe que, en la proposición anterior, X puede tomar valores negativos. Por otra parte X debe tener esperanza ($E[X] < \infty$).

Nota: Las desigualdades de las dos proposiciones anteriores dan cotas de la probabilidad en casos en que no se tiene información precisa acerca de la distribución y sólo la esperanza ó la esperanza y la varianza son conocidas.

Ejemplos 6.1

1. Suponga que se sabe que el número de artículos producidos por una fábrica durante una semana es una variable aleatoria con media = 50

(a) ¿Qué puede decirse acerca de la probabilidad de que se produzcan más de 75 artículos en 1 semana?

(b) Si la varianza de la producción en 1 semana es igual a 25, entonces ¿qué puede decirse acerca de la probabilidad de que esta semana se produzcan entre 40 y 60 artículos?

Solución

(a) $X = \#$ artículos a la semana

$$P(X > 75) \leq \frac{E[X]}{75} = \frac{50}{75} = \frac{2}{3}.$$

(b)

$$\begin{aligned} P(40 \leq X \leq 60) &= P(-10 < X - 50 < 10) = P(|X - 50| < 10) \\ &= 1 - P(|X - 50| \geq 10). \end{aligned}$$

Entonces,

$$P(|X - 50| \geq 10) \leq \frac{25}{10^2} = \frac{1}{4} \Rightarrow P(40 \leq X \leq 60) \geq 1 - \frac{1}{4} = \frac{3}{4}.$$

2. Si X está uniformemente distribuida sobre $(0, 10)$, entonces

$$E[X] = 5, \quad \text{Var}[X] = \frac{25}{3}$$

y por la desigualdad de Chebyshev:

$$P(|X - 5| > 4) \leq \frac{25}{3 \times 16} \cong 0,52.$$

Note que el resultado correcto es: $P(|X - 5| > 4) = 0,2$

Aplicaciones de la desigualdad de Chebyshev

El siguiente resultado establece que sólo una variable aleatoria degenerada (constante con probabilidad 1) puede tener varianza 0.

Proposición 6.3 Sea X una variable aleatoria tal que $\text{Var}X = 0 = \sigma^2$, entonces $P(X = E[X]) = 1$.

Demostración. Aplicando la desigualdad de Chebyshev se tiene que, para todo $n \geq 1$,

$$P\left(|X - \mu| > \frac{1}{n}\right) \leq \sigma^2 n^2 = 0,$$

porque $\sigma^2 = 0$. Por otra parte

$$P(|X - \mu| \neq 0) = P\left(\bigcup_n \left\{|X - \mu| > \frac{1}{n}\right\}\right) \leq \sum_n P\left(|X - \mu| > \frac{1}{n}\right) = 0,$$

lo que implica que $P(|X - \mu| = 0) = 1$.

Teorema 6.1 Ley Débil de los Grandes Números

Sean $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes, idénticamente distribuidas cada una con media $E[X_i] = \mu$ y varianza σ^2 (ambas finitas). Entonces para cada $\varepsilon > 0$

$$P\left\{\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \varepsilon\right\} \rightarrow 0 \quad \text{cuando } n \rightarrow \infty.$$

Demostración. Puesto que $\{X_i\}_{i \geq 1}$ son independientes e idénticamente distribuidas, para cada n

$$E\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{1}{n} \sum E[X_i] = \frac{1}{n} \cdot n\mu = \mu$$

y

$$\text{Var}\left[\frac{X_1 + \dots + X_n}{n}\right] = \frac{1}{n^2} \sum \text{Var} X_i = \frac{1}{n} \sigma^2.$$

Aplicando la desigualdad de Chebyshev a $\frac{S_n}{n} = \frac{X_1 + \dots + X_n}{n}$ se tiene que

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \varepsilon\right) \leq \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0, \quad \text{cuando } n \rightarrow \infty.$$

Ejemplos 6.2

1. Sea A un evento que ocurre con probabilidad $p = P(A)$ en un experimento dado y $S_n = \#$ veces que A ocurre en n realizaciones del experimento, observe que S_n es una variable aleatoria binomial con parámetros (n, p) . Se tiene que $P\left(\left|\frac{S_n}{n} - p\right| \geq \varepsilon\right) \rightarrow 0$ cuando $n \rightarrow \infty \Rightarrow p_n = \frac{S_n}{n}$ se aproxima a p cuando n es suficientemente grande.

2. Supongamos que una falla en un sistema mecánico ocurre de acuerdo a una distribución de Poisson con parámetro $\lambda > 0$. Entonces el número de fallas en un intervalo $[0, t]$ está dado por $X(t)$ y es tal que

$$P(X(t) = k) = e^{-\lambda t} \frac{(\lambda t)^k}{k!}$$

Sea $\tilde{X}(t) = \frac{X(t)}{t} = \# \text{ fallas por unidad de tiempo}$.

Entonces

$$E[\tilde{X}(t)] = \frac{1}{t} E[X(t)] = \frac{\lambda t}{t} = \lambda$$

$$\text{Var}(\tilde{X}(t)) = \frac{1}{t^2} \text{Var}(X(t)) = \frac{\lambda t}{t^2} = \frac{\lambda}{t}$$

$$P(|\tilde{X}(t) - \lambda| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \frac{\lambda}{t} \longrightarrow 0 \quad \text{cuando } t \rightarrow \infty$$

$\tilde{X}(t)$ es aproximada por su esperanza λ cuando t es suficientemente grande.

En general se tiene que el promedio $\frac{X_1 + \cdots + X_n}{n} \cong E[X_i]$.

Ejercicios 6.1

1. Sea X una variable aleatoria tal que $E[X] = \mu$ y $\text{Var}[X] = \sigma^2$. Demuestre que $P(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}$.
2. Si X es una variable aleatoria tal que $E[X] = \mu$ y $\text{Var}[X] = \sigma^2$, el radio $r = \frac{|\mu|}{\sigma}$ se denomina medida del radio de la señal de ruido de X . La idea es que la variable aleatoria puede expresarse como $X = \mu + (X - \mu)$ donde μ representa la señal (ó equivalentemente μ es una medida de la señal aleatoria X y $X - \mu$ el ruido. Si definimos $D = \frac{|X - \mu|}{|\mu|}$ como la desviación relativa de X de la señal, demuestre que para $\alpha > 0$ $P(D \leq \alpha) \geq 1 - \frac{1}{r^2 \alpha^2}$.
3. Suponga que X es una variable aleatoria con media y varianza iguales a 20. ¿Qué puede decirse de $P(0 \leq X \leq 40)$?
4. El número de objetos producidos por una fábrica es una variable aleatoria X con media 100. (a) Dé una cota superior de la probabilidad de que se produzcan al menos 200 artículos. (b) Si la varianza de X es 25, dé una cota de la probabilidad de que se produzcan entre 80 y 120 artículos.

5. Se sabe que el número esperado de objetos defectuosos producidos en una fábrica en una semana es 18. (a) ¿Qué puede decirse de la probabilidad de que esta semana el número de objetos defectuosos sea mayor que 25?. (b) Si la varianza del número de objetos defectuosos en una semana es igual a 9, ¿qué puede decirse de la probabilidad de que esta semana se produzcan entre 11 y 25 objetos defectuosos?
6. Por experiencia un profesor sabe que la nota del examen final de un estudiante es una variable aleatoria con media de 15 puntos y varianza igual a 5 puntos. a) Dé una cota superior para la probabilidad de que la nota de un estudiante sea mayor que 17. b) ¿Qué puede decirse acerca de la probabilidad de que la nota del estudiante esté entre 12 y 18 puntos?. c) ¿Cuántos estudiantes tendrán que tomar el examen final para asegurar, con probabilidad no menor a 0.9, que el promedio de la clase estará entre 12 y 18 puntos?
7. Sean $X_1, \dots, X_{20}$ variables aleatorias de Poisson, independientes con media igual a 1. D'e una cota de $P(X_1 + \dots + X_{20} > 15)$.
8. Sea X una variable aleatoria gamma con parámetros $(n, 1)$. ¿Qué tan grande se necesita n para que $P(|\frac{X}{n} - 1| > 0,01) < 0,01$.
9. Se toma una muestra al azar y con reposición, a efectos de estimar la fracción p de hembras en una población. Encuentre un tamaño de muestra que asegure que, con una probabilidad de al menos 0,99, la estimación se hará con un error menor a 0,005.
10. Se desea estimar la probabilidad de falla p en un proceso de producción mediante la observación de n objetos producidos, elegidos independientemente. Por información previa que se tiene acerca del proceso, se sabe que p está comprendida entre 0.1 y 0.3. Halle el tamaño n de la muestra para que probabilidad de que la frecuencia relativa de objetos defectuosos en la muestra difiera del verdadero valor p en más de 0,01 sea menor a 0,05.
11. Se observan n repeticiones independientes de un juego de ruleta. Halle n para que con probabilidad menor que 0.1, la frecuencia relativa del número de veces que ocurre la primera docena sea mayor que $\frac{1}{3}$.
Nota. Los resultados posibles en un juego de ruleta son 00, 0, 1, 2, $\dots$, 36 y la primera docena consta de los numeros 1, 2, $\dots$, 12.

6.2. Teorema Central del Límite

6.2.1. Convergencia de Variables Aleatorias.

En esta sección se presentan algunas de las nociones de convergencia de sucesiones de variables aleatorias más ampliamente usadas.

Definición 6.1 Una sucesión de variables aleatorias $X_1, X_2, \dots$ se dice que converge en probabilidad a la variable aleatoria X , si se verifica que para todo $\epsilon > 0$

$$P(|X_n - X| \geq \epsilon) \longrightarrow 0 \quad \text{cuando} \quad n \longrightarrow \infty.$$

En este caso se usa la notación $X_n \xrightarrow{(P)} X \quad (n \longrightarrow \infty)$.

Observemos que en términos de esta definición la ley débil de los grandes números puede enunciarse diciendo que la sucesión de variables aleatorias $Y_n = \frac{S_n}{n}$ converge en probabilidad a la variable aleatoria Y constantemente igual a μ .

Definición 6.2 Una sucesión de variables aleatorias $X_1, X_2, \dots$ converge casi seguramente (o con probabilidad 1) a la variable aleatoria X , si se verifica que

$$P(X_n \longrightarrow X \quad \text{cuando} \quad n \longrightarrow \infty) = 1. \quad (6.1)$$

En este caso se usa la notación $X_n \longrightarrow X \quad \text{c.s. (o c.p. 1)}$.

NOTA: La expresión (6.1) significa que el evento formado por los $\omega \in \Omega$ tales que

$$\lim_{n \rightarrow \infty} X_n(\omega) \neq X(\omega)$$

tiene probabilidad cero.

Definición 6.3 Sean $X_1, X_2, \dots$ y X variables aleatorias con funciones de distribución $F_1, F_2, \dots$ y F respectivamente. Se dice que (X_n) converge en distribución a X si se verifica que

$$\lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x)$$

siempre que x sea un punto de continuidad de F ($P(X = x) = 0$).

Para la convergencia en distribución usamos la notación $X_n \Longrightarrow X$.

6.2.2. Teorema Central del Límite

Teorema 6.2 Sea $X_1, \dots, X_n, \dots$ una sucesión de variables aleatorias independientes e idénticamente distribuidas con media μ y varianza σ^2 . Entonces la sucesión de variables aleatorias $Y_n = \frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}}$ converge en distribución a una variable aleatoria normal unitaria cuando $n \rightarrow \infty$. Es decir:

$$\lim_{n \rightarrow \infty} P(Y_n \leq a) = \lim_{n \rightarrow \infty} P\left(\frac{X_1 + \dots + X_n - n\mu}{\sigma\sqrt{n}} \leq a\right) = \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^a e^{-\frac{x^2}{2}} dx$$

Aplicaciones

Antes de presentar la demostración de una de las versiones más conocidas del teorema central del límite, el teorema de De Moivre-Laplace, veremos algunas de sus aplicaciones más comunes al cálculo de probabilidades.

Ejemplos 6.3

1. El número de estudiantes que se inscribe en una clase de psicología es una variable aleatoria Poisson con media $\lambda = 100$. Si el número de estudiantes inscritos es mayor que 120 entonces se dividirá la clase en 2 grupos ¿Cuál es la probabilidad de que esto pase?

$$X = \# \text{ estudiantes} \quad P(X > 120) = e^{-100} \sum_{k=120}^{\infty} \frac{(100)^k}{k!}.$$

En vista de que determinar el valor de la suma anterior puede ser una tarea engorrosa, se puede tratar de aplicar el Teorema Central del Límite para obtener un valor aproximado de la probabilidad anterior. Para esto, observemos que la variable aleatoria X se puede escribir como la suma de 100 variables aleatorias independientes de Poisson, con parámetros $\lambda = 1$; es decir,

$X = X_1 + \dots + X_{100}$ donde X_i Poisson $\lambda = 1$ para $i \geq 1$ variable aleatoria independientes. Se puede entonces usar el teorema central del límite para obtener una aproximación de la probabilidad deseada. Observe que $E[X_i] = 1 = \text{Var}[X_i]$.

$$\begin{aligned} P(X > 120) &= P(X_1 + \dots + X_{100} > 120) \\ &= P\left(\frac{X_1 + \dots + X_{100} - 100 \times 1}{1 \times \sqrt{100}} > \frac{120 - 100}{\sqrt{100}}\right) \\ &= P\left(\frac{X - 100}{10} > 2\right) \sim 1 - \Phi(2) = 0,0228. \end{aligned}$$

2. Si $X_i, i = 1, \dots, 10$ son variables aleatorias independientes uniformemente distribuidas sobre $(0,1)$. Calcule una aproximación a $P\left(\sum_{i=1}^{10} X_i > 6\right)$.

Se tiene que

$$E[X_i] = \int_0^1 x dx = \frac{x^2}{2} \Big|_0^1 = \frac{1}{2} \quad y \quad E[X_i^2] = \int_0^1 x^2 dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3};$$

por lo tanto $Var[X_i] = \frac{1}{3} - \frac{1}{4} = \frac{1}{12}$. Aplicando el teorema central del límite se tiene que

$$\begin{aligned} P\left(\sum_{i=1}^{10} X_i > 6\right) &= P\left(\frac{\sum_{i=1}^{10} X_i - 10 \times \frac{1}{2}}{\frac{1}{\sqrt{12}} \times \sqrt{10}} > \frac{6 - 5}{\frac{\sqrt{10}}{\sqrt{12}}}\right) = P\left(\frac{\sum_{i=1}^{10} X_i - 5}{\sqrt{\frac{10}{12}}} > \frac{\sqrt{12}}{\sqrt{10}}\right) \\ &\sim 1 - \Phi(\sqrt{1,2}) = 1 - \Phi(1,09545) \sim 0,16. \end{aligned}$$

6.2.3. Teorema de De Moivre-Laplace

En esta sección demostraremos el resultado conocido como teorema de De Moivre Laplace que proporciona una versión del teorema central del límite para variables aleatorias de Bernoulli.

Teorema 6.3 (Teorema de De Moivre-Laplace)

Sea $(X_n)_{n \geq 1}$ una sucesión de variables aleatorias independientes de Bernoulli con probabilidad p de éxito. Sea $S_n = \#$ éxitos en n ensayos de Bernoulli. Para $q = 1 - p$, y constantes $a < b$ cualesquiera se cumple que

$$\lim_{n \rightarrow \infty} P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) = \Phi(b) - \Phi(a) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx$$

Comentarios:

1. Para todo $i \geq 1$ $E[X_i] = p = \mu$ y $Var[X_i] = pq$.
2. Para cada $n \geq 1$, S_n es una variable aleatoria binomial con parámetro (n, p) . El teorema de De Moivre-Laplace establece que es posible aproximar una variable aleatoria binomial por una normal con parámetros (np, npq)

Preliminares

1. Usaremos la notación $\sim$ para indicar la relación asintóticamente igual. Es decir, diremos que $a_n \sim b_n$, si y sólo si

$$\lim_{n \rightarrow \infty} \frac{a_n}{b_n} = 1.$$

Cuando se verifique esta relación entre sucesiones (a_n) y (b_n) , para n suficientemente grande se podrán reemplazar los términos de la sucesión (a_n) por los de (b_n) .

2. La Fórmula de Stirling.

Lema 6.1 Cuando $n \rightarrow \infty$, se cumple que

$$n! \sim \left(\frac{n}{e}\right)^n \sqrt{2\pi n}.$$

Comentario. La fórmula de Stirling implica que

$$\begin{aligned} \binom{n}{k} &= \frac{n!}{k!(n-k)!} \sim \frac{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}}{\left(\frac{k}{e}\right)^k \sqrt{2\pi k} \left(\frac{n-k}{e}\right)^{n-k} \sqrt{2\pi(n-k)}} \\ &= \frac{n^k}{k^k} \cdot \frac{n^{n-k}}{(n-k)^{n-k}} \cdot \sqrt{\frac{n}{k} \cdot \frac{1}{2\pi(n-k)}} \\ \binom{n}{k} &\sim \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k} \sqrt{\frac{1}{2\pi} \frac{n}{k(n-k)}} \end{aligned}$$

3. **Lema 6.2** Sea S_n una variable aleatoria binomial con parámetros (n, p) y sea $q = 1 - p$, sea

$$x_{n,k} = x_k = \frac{k - np}{\sqrt{npq}} \quad 0 \leq k \leq n$$

Para $A > 0$ dado se define $S_A = \{k : |x_k| \leq A\}$. Entonces si $k \in S_A$ se cumple que

$$P(S_n = k) = \binom{n}{k} p^k q^{n-k} \sim \frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2},$$

para n suficientemente grande. Sobre S_A la convergencia es uniforme.

Demostración. Sea $k \in S_A$, por definición de x_k se tiene que $k = np + \sqrt{npq}x_k$ lo que implica que

$$\frac{k}{np} = 1 + \frac{\sqrt{npq}}{np} x_k = 1 + \frac{\sqrt{q}}{\sqrt{np}} x_k \rightarrow 1 \quad \text{cuando } n \rightarrow \infty;$$

es decir que,

$$k \sim np. \tag{6.2}$$

Por otra parte

$$\begin{aligned} n - k &= n - np - \sqrt{npq}x_k = nq - \sqrt{npq}x_k \\ \frac{n - k}{nq} &= 1 - \frac{\sqrt{npq}}{nq}x_k = 1 - \frac{\sqrt{p}}{\sqrt{nq}}x_k \longrightarrow 1 \quad \text{cuando } n \rightarrow \infty; \end{aligned}$$

por lo cual

$$n - k \sim nq \quad (6.3)$$

para n suficientemente grande.

De (6.2) y (6.3) se deduce que (para n suficientemente grande) $k(n - k) \sim npnq$.

Usando la formula de Stirling tenemos que (para n suficientemente grande)

$$\begin{aligned} \binom{n}{k} p^k q^{n-k} &\sim \left(\frac{n}{k}\right)^k \left(\frac{n}{n-k}\right)^{n-k} \sqrt{\frac{1}{2\pi} \frac{n}{k(n-k)}} p^k q^{n-k} \\ &= \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \sqrt{\frac{1}{2\pi} \frac{n}{k(n-k)}} \\ &\sim \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \sqrt{\frac{n}{2\pi npnq}} \\ &= \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \frac{1}{\sqrt{2\pi npq}}. \end{aligned}$$

Para demostrar el lema será entonces suficiente demostrar que

$$\left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} \sim e^{-x_k^2/2}.$$

Sabemos que $np = k - \sqrt{npq}x_k$ y $nq = n - k + \sqrt{npq}x_k$ lo que implica que

$$\frac{np}{k} = 1 - \frac{\sqrt{npq}x_k}{k} \quad \text{y} \quad \frac{nq}{n-k} = 1 + \frac{\sqrt{npq}x_k}{n-k}.$$

Entonces

$$\ln \left(\frac{np}{k}\right)^k \left(\frac{nq}{n-k}\right)^{n-k} = k \ln \left(1 - \frac{\sqrt{npq}x_k}{k}\right) + (n-k) \ln \left(1 + \frac{\sqrt{npq}x_k}{n-k}\right).$$

Recordemos que usando la fórmula de Taylor se obtiene la aproximación

$$\ln(1+x) \sim x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^{n-1} \frac{x^n}{n}.$$

Aplicando esta expresión a los términos del lado derecho de la última igualdad se obtiene

$$\begin{aligned} k \ln \left(1 - \frac{\sqrt{npq}x_k}{k} \right) &= k \left(-\frac{\sqrt{npq}x_k}{k} - \frac{npq(x_k)^2}{2k^2} + \dots \right) \\ (n-k) \ln \left(1 + \frac{\sqrt{npq}x_k}{n-k} \right) &= (n-k) \left(\frac{\sqrt{npq}x_k}{n-k} - \frac{npq(x_k)^2}{2(n-k)^2} + \dots \right) \end{aligned}$$

siempre que

$$\left| \frac{\sqrt{npq}x_k}{k} \right| < 1 \quad \text{y} \quad \left| \frac{\sqrt{npq}x_k}{n-k} \right| < 1$$

lo cual es cierto para n suficientemente grande ya que en este caso $k \sim np$ y $n-k \sim nq$. Se tiene entonces que

$$\begin{aligned} \ln \left(\frac{np}{k} \right)^k \left(\frac{nq}{n-k} \right)^{n-k} &\sim -\sqrt{npq}x_k - \frac{npq(x_k)^2}{2k} + \sqrt{npq}x_k - \frac{npq(x_k)^2}{2(n-k)} \\ &= -npq \frac{x_k^2}{2} \left(\frac{1}{k} + \frac{1}{n-k} \right) = -npq \frac{x_k^2}{2} \frac{n}{(n-k)k} \\ &= -\frac{x_k^2}{2} \frac{np}{k} \cdot \frac{nq}{n-k} \sim -\frac{x_k^2}{2}. \end{aligned}$$

Por lo tanto

$$\left(\frac{np}{k} \right)^k \left(\frac{nq}{n-k} \right)^{n-k} \sim e^{-x_k^2/2}.$$

Demostración del teorema de De Moivre-Laplace.

Cuando ocurre el evento $\{S_n = k\}$ se tiene $\frac{S_n - np}{\sqrt{npq}} = X_k = X_{n,k}$, entonces

$$\begin{aligned} P \left(a < \frac{S_n - np}{\sqrt{npq}} \leq b \right) &= P \left(\bigcup_{k=0}^n \left(a < \frac{S_n - np}{\sqrt{npq}} \leq b, S_n = k \right) \right) \\ &= \sum_{k=0}^n P \left(a < \frac{S_n - np}{\sqrt{npq}} \leq b | S_n = k \right) P(S_n = k). \end{aligned}$$

Observando que

$$P \left(a < \frac{S_n - np}{\sqrt{npq}} \leq b | S_n = k \right) = \begin{cases} 1 & \text{si } x_k \in (a, b] \\ 0 & \text{si } x_k \notin (a, b] \end{cases} \quad (6.4)$$

se tiene

$$\begin{aligned} P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) &= \sum_{\{a < x_k \leq b\}} P(S_n = k) = \sum_{\{a < x_k \leq b\}} \binom{n}{k} p^k q^{n-k} \\ &\sim \sum_{\{a < x_k \leq b\}} \frac{1}{\sqrt{2\pi npq}} e^{-x_k^2/2} = \frac{1}{\sqrt{2\pi}} \sum_{\{a < x_k \leq b\}} \frac{1}{\sqrt{npq}} e^{-x_k^2/2} \end{aligned}$$

Ahora bien

$$x_{k+1} - x_k = \frac{k+1-np}{\sqrt{npq}} - \frac{k-np}{\sqrt{npq}} = \frac{1}{\sqrt{npq}}.$$

Esto implica que

$$P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) \sim \frac{1}{\sqrt{2\pi}} \sum_{\{a < x_k \leq b\}} e^{-x_k^2/2} (x_{k+1} - x_k) \quad (6.5)$$

Por otra parte la correspondencia $k \longleftrightarrow x_k$ es 1-1 y cuando k varía de 0 a n $x_k = \frac{k-np}{\sqrt{npq}}$ varía desde $-\sqrt{\frac{np}{q}}$ hasta $\sqrt{\frac{nq}{p}}$. Tomando n suficientemente grande se tendrá que $(a, b] \subseteq \left(-\sqrt{\frac{np}{q}}, \sqrt{\frac{nq}{p}}\right]$ y los puntos x_k que caen dentro de $(a, b]$ forman una partición de $(a, b]$ en intervalos de longitud $\frac{1}{\sqrt{npq}}$.

Por lo tanto la suma en (4,6) es la suma (inferior) de Riemann de $\int_a^b e^{-x^2/2} dx$.

Es decir que

$$P\left(a < \frac{S_n - np}{\sqrt{npq}} \leq b\right) \sim \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

Teorema 6.4 Sea $\{X_n\}_{n \geq 1}$ una sucesión de variables aleatorias independientes, de Bernoulli, con probabilidad p de éxito y sea $q = 1 - p$. Sea $S_n = \sum_{i=1}^n X_i$ ($= \#$ éxitos en n ensayos de Bernoulli). Entonces se cumple que

$$\lim_{n \rightarrow \infty} P\left\{\frac{S_n - np}{\sqrt{npq}} \leq x\right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-u^2/2} du.$$

Demostración. Sea $\tilde{S}_n = \frac{S_n - np}{\sqrt{npq}}$. Para $\epsilon > 0$ dado elijamos números reales a_1 y a_2 tales que

- i) $a_1 < x < a_2$,
- ii) $\Phi(a_1) < \frac{\epsilon}{2}$ y
- iii) $1 - \Phi(a_2) < \frac{\epsilon}{2}$.

Se tiene entonces que

$$\{x < \tilde{S}_n \leq a_2\} \subseteq \{\tilde{S}_n > x\} \implies P(x < \tilde{S}_n \leq a_2) \leq P(\tilde{S}_n > x)$$

$$\implies 1 - P(x < \tilde{S}_n \leq a_2) \geq 1 - P(\tilde{S}_n > x) = P(\tilde{S}_n \leq x).$$

Por otra parte

$$\{a_1 < \tilde{S}_n \leq x\} \subseteq \{\tilde{S}_n \leq x\} \implies P(a_1 < \tilde{S}_n \leq x) \leq P(\tilde{S}_n \leq x) \leq 1 - P(x < \tilde{S}_n \leq a_2).$$

Aplicando el teorema de De Moivre-Laplace

$$1 - P(x < \tilde{S}_n \leq a_2) \longrightarrow 1 - (\Phi(a_2) - \Phi(x)) \quad \text{cuando } n \rightarrow \infty \text{ y}$$

$$P(a_1 < \tilde{S}_n \leq x) \longrightarrow \Phi(x) - \Phi(a_1) \quad \text{cuando } n \rightarrow \infty.$$

Por lo tanto existe N tal que si $n > N$:

$$1 - P(x < \tilde{S}_n \leq a_2) \leq \Phi(x) + (1 - \Phi(a_2)) + \frac{\epsilon}{2} \text{ y}$$

$$P(a_1 < \tilde{S}_n \leq x) \geq \Phi(x) - \Phi(a_1) - \frac{\epsilon}{2}$$

$$\implies \Phi(x) - \Phi(a_1) - \frac{\epsilon}{2} \leq P(\tilde{S}_n \leq x) \leq \Phi(x) + (1 - \Phi(a_2)) + \frac{\epsilon}{2}$$

$$\implies \Phi(x) - \frac{\epsilon}{2} - \frac{\epsilon}{2} \leq P(\tilde{S}_n \leq x) \leq \Phi(x) + \frac{\epsilon}{2} + \frac{\epsilon}{2}$$

$$\implies |P(\tilde{S}_n \leq x) - \Phi(x)| \leq \epsilon.$$

Acerca de la demostración del Teorema Central del Limite (para variables aleatorias independientes e idénticamente distribuidas).

A continuación presentaremos los lineamientos generales de la demostración del teorema 4.3, la cual está basada en el siguiente lema que enunciaremos sin demostración.

Lema 6.3 Sean $Z_1, Z_2, \dots$ y Z variables aleatorias con funciones de distribución $F_1, F_2, \dots$ y F y funciones generadoras de momentos $\varphi_1, \varphi_2, \dots$ y φ , respectivamente. Si $\varphi_n(t) \rightarrow \varphi(t)$ para todo t , cuando $n \rightarrow \infty$, entonces

$$F_n(t) \rightarrow F(t) \quad \text{para todo } t \text{ cuando } n \rightarrow \infty$$

.

Consecuencia del lema: Si $Z \sim N(0, 1)$ entonces $\varphi(t) = e^{t^2/2}$. El lema implica que si

$$\varphi_n(t) \rightarrow e^{t^2/2}, \quad n \rightarrow \infty, \quad \text{entonces } F_n(t) \xrightarrow{n \rightarrow \infty} \Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx.$$

Demostración del teorema 3. Asumiremos que para cada X_i , $i \geq 1$ la función generadora de momentos $\varphi_i(t) = E[e^{tX_i}]$ existe y es finita.

(I) Supongamos que $\mu = 0$ y $\sigma^2 = 1$. Denotemos por $\varphi_{i,n}$ a la función generadora de momentos de $\frac{X_i}{\sqrt{n}}$; es decir, $\varphi_{\frac{X_i}{\sqrt{n}}}$, se tiene que

$$\varphi_i(t) = E[e^{tX_i}] \implies \varphi_{i,n}(t) = E\left[e^{t\frac{X_i}{\sqrt{n}}}\right] = E\left[e^{\frac{t}{\sqrt{n}}X_i}\right] = \varphi_i\left(\frac{t}{\sqrt{n}}\right)$$

Por lo tanto la función generadora de momentos de $S_n = \sum_{i=1}^n \left(\frac{X_i}{\sqrt{n}}\right)$ está dada por

$$\varphi_{S_n}(t) = \prod_{i=1}^n \varphi_i\left(\frac{t}{\sqrt{n}}\right) = \left[\varphi_i\left(\frac{t}{\sqrt{n}}\right)\right]^n = \left[\varphi_1\left(\frac{t}{\sqrt{n}}\right)\right]^n.$$

Sea $L(t) = \ln \varphi_1(t)$. Para demostrar el teorema será suficiente demostrar que

$$\left[\varphi_1\left(\frac{t}{\sqrt{n}}\right)\right]^n \longrightarrow e^{t^2/2} \text{ cuando } n \longrightarrow \infty.$$

Esto equivale a demostrar que

$$\begin{aligned} n \ln \left(\varphi_1 \left(\frac{t}{\sqrt{n}} \right) \right) &\xrightarrow{n \rightarrow \infty} t^2/2 \quad \text{ó} \\ n L \left(\frac{t}{\sqrt{n}} \right) &\xrightarrow{n \rightarrow \infty} t^2/2 \end{aligned}$$

Observemos que, como L es continua

$$\lim_{n \rightarrow \infty} L \left(\frac{t}{\sqrt{n}} \right) = L(0) = 0 \implies \lim_{n \rightarrow \infty} n \ln \left(\varphi_1 \left(\frac{t}{\sqrt{n}} \right) \right) = \lim_{n \rightarrow \infty} \frac{L \left(\frac{t}{\sqrt{n}} \right)}{\frac{1}{n}} \left(\frac{0}{0} \right).$$

Pot lo tanto, podemos aplicar la regla de L'Hôpital para calcular el límite.

Diferenciando L y evaluando en cero las sucesivas derivadas obtenemos

$$L(0) = \log \varphi(0) = \ln(1) = 0$$

$$L'(0) = \frac{\varphi'(0)}{\varphi(0)} = \frac{\mu}{1} = \mu = 0$$

$$L''(0) = \frac{\varphi(0)\varphi''(0) - (\varphi'(0))^2}{(\varphi(0))^2} = E[X^2] - \mu^2 = E[X^2] = Var[X] = 1$$

Entonces,

$$\begin{aligned}\lim_{n \rightarrow \infty} n \ln \left(\varphi_1 \left(\frac{t}{\sqrt{n}} \right) \right) &= \lim_{n \rightarrow \infty} \left(\frac{L' \left(\frac{t}{\sqrt{n}} \right) \cdot \frac{t}{2}}{n^{-2} \cdot n^{3/2}} \right) = \lim_{n \rightarrow \infty} \left(\frac{L' \left(\frac{t}{\sqrt{n}} \right) \cdot t}{\frac{2}{\sqrt{n}}} \right) \\ &= \lim_{n \rightarrow \infty} \left(L'' \left(\frac{t}{\sqrt{n}} \right) \frac{t^2}{2} \right) = L''(0) \cdot \frac{t^2}{2} = \frac{t^2}{2}\end{aligned}$$

(II) Caso general. Sea $X_i^* = \frac{X_i - \mu}{\sigma}$. Entonces $E[X_i^*] = 0$ y $Var[X_i^*] = 1$. Entonces el resultado de la parte (I) vale para las X_i^* lo que significa que

$$\begin{aligned}P \left(\frac{X_1^* + \cdots + X_n^*}{\sqrt{n}} \leq a \right) &= P \left(\frac{\sum (X_i - \mu)}{\sigma \sqrt{n}} \leq a \right) \\ &= P \left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \leq a \right) \xrightarrow{n \rightarrow \infty} \Phi(a)\end{aligned}$$

Esto demuestra que para cada a $F_n(a) = P \left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma \sqrt{n}} \leq a \right) \xrightarrow{n \rightarrow \infty} \Phi(a)$

Nota: Puede demostrarse que para cada $\varepsilon > 0$, existe N tal que $|F_n(a) - \phi(a)| < \varepsilon \forall a$ siempre que $n \geq N$. Es decir que la convergencia es uniforme en a .

Teorema Central del Límite para variables aleatorias independientes no idénticamente distribuidas.

Enunciamos dos versiones más generales del teorema central del límite.

Teorema 6.5 Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes con medias $\mu_i = E[X_i]$ y varianzas $Var[X_i] = \sigma_i^2$, $i \geq 1$. Si

a) Las variables X_i , $i \geq 1$ están uniformemente acotadas es decir, si existe $M > 0$ tal que $P(|X_i| \leq M) = 1$ para todo $i \geq 1$.

b) Se cumple que $\sum_{i=1}^{\infty} \sigma_i^2 = \infty$, entonces

$$P \left(\frac{\sum_{i=1}^n (X_i - \mu_i)}{\sqrt{\sum_{i=1}^n \sigma_i^2}} \leq a \right) \xrightarrow{n \rightarrow \infty} \Phi(a)$$

Condición de Lindeberg. Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes con medias $\mu_i = E[X_i] = 0$ y sea $\varsigma_n^2 = \sum_{i=1}^n \sigma_i^2$. Para $\epsilon > 0$ se define $A_{i,n} = \{\omega : |X_i| > \epsilon \times \varsigma_n\}$, entonces, decimos que se cumple la condición de Lindeberg si:

$$\frac{1}{\varsigma_n^2} \sum_{i=1}^n E[X_i^2 \chi_{A_{i,n}}] \xrightarrow{n \rightarrow \infty} 0.$$

Teorema 6.6 (Lindeberg) Sea $X_1, X_2, \dots$ una sucesión de variables aleatorias independientes con medias $\mu_i = E[X_i] = 0$ y varianzas $\text{Var}[X_i] = \sigma_i^2 < \infty$ $i \geq 1$. Si $S_n = X_1 + \dots + X_n$, entonces $E[S_n] = 0$ y $\text{Var}[S_n] = \sum_{i=1}^n \sigma_i^2 = \varsigma_n^2$. Si se cumple la condición de Lindeberg entonces

$$P\left(\frac{S_n}{\varsigma_n} \leq a\right) \xrightarrow{n \rightarrow \infty} \Phi(a).$$

6.3. Ley Fuerte de los Grandes Números.

6.3.1. Preliminares.

Hemos visto que la ley débil de los grandes números establece que si $X_1, X_2, \dots$ son variables aleatorias independientes e idénticamente distribuidas con media μ y varianza σ^2 , entonces para cada $\epsilon > 0$

$$P\left(\left|\frac{X_1 + \dots + X_n}{n} - \mu\right| \geq \epsilon\right) \xrightarrow{n \rightarrow \infty} 0$$

¿Qué significa esto?. Esta ley afirma que si se toma n suficientemente grande, el evento $A_n = \left\{\omega : \left|\frac{X_1 + \dots + X_n}{n} - \mu\right| < \epsilon\right\}$ tiene una alta probabilidad de ocurrir.

Sin embargo esto no garantiza que si $\omega \in A_n$ la sucesión $\frac{S_n(\omega)}{n} = \frac{X_1(\omega) + \dots + X_n(\omega)}{n}$ se mantendrá cerca de μ para todo $m > n$.

Un resultado más categórico lo constituye la *ley fuerte de los grandes números*.

Teorema 6.7 (Ley Fuerte de los Grandes Números). Sea $(X_n)_{n \geq 1}$ una sucesión de variables aleatorias independientes, idénticamente distribuidas y tales que $E[X_i] = \mu$, $i = 1, \dots$. Entonces con probabilidad 1

$$\lim_{n \rightarrow \infty} \frac{X_1 + \dots + X_n}{n} = \mu;$$

es decir,

$$P\left(\left\{\omega : \lim_{n \rightarrow \infty} \frac{S_n(\omega)}{n} = \mu\right\}\right) = 1$$

Comentarios.

1. Si $A = \left\{\omega : \lim_{n \rightarrow \infty} \frac{S_n(\omega)}{n} = \mu\right\}$, para todo $\omega_1 \in A$ se cumple que $\frac{S_n(\omega_1)}{n}$ estará arbitrariamente próximo a μ cuando $n \rightarrow \infty$.
2. Esto en el caso de variables aleatorias de Bernoulli, el teorema afirma que para n grande $\frac{S_n}{n} - p$ es pequeño y con probabilidad 1 permanecerá así.

Demostraremos la ley fuerte de los grandes números para variables aleatorias de Bernoulli. Pero antes demostraremos un importante lema.

Lema de Borel-Cantelli

Definición 6.4 Sea $\{A_n\}_{n \geq 1}$ una sucesión de eventos en un espacio de probabilidad. Se define el límite superior de la sucesión de conjuntos $\{A_n\}_{n \geq 1}$, como el conjunto

$$\bar{A} = \{\omega \in \Omega : \omega \in A_n \text{ para infinitos valores de } n\}.$$

Lema 6.4 El límite superior $\bar{A}$, de una sucesión de eventos $\{A_n\}_{n \geq 1}$ en un espacio de probabilidad, es igual a

$$\bar{A} = \bigcap_{m=1}^{\infty} \bigcup_{n=m}^{\infty} A_n.$$

Demostración. Si definimos $B_m = \bigcup_{n=m}^{\infty} A_n$, demostrar el lema equivale a demostrar que

$$\bar{A} = \bigcap_{m=1}^{\infty} B_m.$$

Sea $\omega \in \bar{A}$, como ω pertenece a A_n para infinitos valores de n , se tiene que para todo $m \geq 1$ existe $n_0 \geq m$ tal que $\omega \in A_{n_0}$. Por lo tanto

$$\omega \in \bigcup_{n=m}^{\infty} A_n = B_m \implies \omega \in \bigcap_{m=1}^{\infty} B_m \implies \bar{A} \subseteq \bigcap_{m=1}^{\infty} B_m$$

Por otra parte si $\omega \in \bigcap_{m=1}^{\infty} B_m$ se tiene que $\omega \in B_m$ para todo $m \geq 1$; por lo que para todo $m \geq 1$ existe $n_1 \geq m$ tal que $\omega \in A_{n_1} \Rightarrow \omega \in$ infinitos $A_{n'_s}$, es decir $\omega \in \bar{A}$.

Esto demuestra la igualdad entre los conjuntos.

Lema 6.5 Dado $x \in \mathbb{R}$ se cumple que

$$1 - x \leq e^{-x}.$$

La demostración es sencilla y se deja como ejercicio.

Lema 6.6 (Lema de Borel-Cantelli.) Sea $\{A_n\}_{n \geq 1}$ una sucesión de eventos en un espacio de probabilidad. Sea Entonces

1. Si $\sum_{n=1}^{\infty} P(A_n) < \infty$ se cumple que $P(\bar{A}) = 0$.
2. Si $\sum_{n=1}^{\infty} P(A_n) = \infty$ y $\{A_n\}_{n \geq 1}$ es una sucesión de eventos independientes entonces se cumple que $\Rightarrow P(\bar{A}) = 1$.

Demostración.

1. Si $\sum_{n=1}^{\infty} P(A_n) < \infty$, de la definición de B_m deducimos que

$$P(B_m) \leq \sum_{n=m}^{\infty} P(A_n) \xrightarrow{m \rightarrow \infty} 0.$$

Pero como $\bar{A} \subseteq B_m$ para todo m entonces $P(\bar{A}) \leq P(B_m)$. Tomando $\lim_{m \rightarrow \infty}$ obtenemos que es $P(\bar{A}) \leq \lim_{m \rightarrow \infty} P(B_m) = 0 \Rightarrow P(\bar{A}) = 0$.

2. Supongamos que $\sum_{n=1}^{\infty} P(A_n) = \infty$. Basta demostrar que $P(\bar{A}^c) = 0$

$$\bar{A} = \bigcap_{m=1}^{\infty} B_m \Rightarrow \bar{A}^c = \bigcup_{m=1}^{\infty} B_m^c, \text{ donde } B_m^c = \bigcap_{n=m}^{\infty} A_n^c \subseteq \bigcap_{n=m}^l A_n^c,$$

Entonces $P(\bar{A}^c) \leq \sum_{m=1}^{\infty} P(B_m^c)$. Demostraremos que $P(B_m^c) = 0 \forall m$.

$P(B_m^c) = P\left(\bigcap_{n=m}^{\infty} A_n^c\right)$. Ahora bien, usando la independencia de los $A_{i's}$ tenemos que para cada $l \geq m$ $P\left(\bigcap_{n=m}^l A_n^c\right) = \prod_{n=m}^l P(A_n^c) = \prod_{n=m}^l (1 - P(A_n))$.

Usando la desigualdad: $1 - x \leq e^{-x}$ para $x = P(A_n)$ obentemos:

$$P(B_m^c) \leq P\left(\bigcap_{n=m}^l A_n^c\right) = \prod_{n=m}^l (1 - P(A_n)) \leq \prod_{n=m}^l e^{-P(A_n)} = e^{-\sum_{n=m}^l P(A_n)}.$$

Para m dado, cuando $l \rightarrow \infty$ se obtiene

$$0 \leq P(B_m^c) = P\left(\bigcap_{n=m}^{\infty} A_n^c\right) \leq e^{-\sum_m^{\infty} P(A_n)} = 0 \Rightarrow P(\bar{A}) = 1$$

Lema 6.7 Sean $\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$ la densidad normal, $\mu = 0, \sigma^2 = 1$, y sea $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-t^2/2} dt$ la función de distribución normal. Entonces, cuando $x \rightarrow \infty$, se cumple que $1 - \Phi(x) \sim \frac{\varphi(x)}{x}$.

Demostración. Para $t > 0$ es claro que

$$\left(1 - \frac{3}{t^4}\right) \varphi(t) < \varphi(t) < \left(1 + \frac{1}{t^2}\right) \varphi(t)$$

Entonces

$$\begin{aligned} \int_x^{\infty} \left(1 - \frac{3}{t^4}\right) \varphi(t) dt &< \int_x^{\infty} \varphi(t) dt < \int_x^{\infty} \left(1 + \frac{1}{t^2}\right) \varphi(t) dt \\ \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \left(1 - \frac{3}{t^4}\right) e^{-t^2/2} dt &< \int_x^{\infty} \varphi(t) dt < \frac{1}{\sqrt{2\pi}} \int_x^{\infty} \left(1 + \frac{1}{t^2}\right) e^{-t^2/2} (t) dt \\ \Rightarrow \left(\frac{1}{x} - \frac{1}{x^3}\right) \frac{e^{-x^2/2}}{\sqrt{2\pi}} &< \int_x^{\infty} \varphi(t) dt < \frac{1}{x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} \end{aligned}$$

Por lo tanto, para valores grandes de x se cumple que

$$1 - \Phi(x) \sim \frac{1}{x} \frac{e^{-x^2/2}}{\sqrt{2\pi}}.$$

6.3.2. Ley fuerte de los grandes números para variables aleatorias de Bernoulli.

Teorema 6.8 Sean $X_1, X_2, \dots$, una sucesión de variables aleatorias de Bernoulli independientes con probabilidad p de éxito. Sea $S_n = \sum_{i=1}^n X_i$.

Entonces con probabilidad 1

$$\lim_{n \rightarrow \infty} \frac{S_n}{n} = p$$

Demostración. Sea $A_k = \left\{ \omega : \left| \frac{S_k - kp}{\sqrt{kpq}} \right| \geq \sqrt{2a \log k} \right\}$, para $a > 1$. Entonces usando el teorema de De Moivre Laplace tenemos que $S_k^* = \frac{S_k - kp}{\sqrt{kpq}} \sim N(0, 1)$ para valores grandes de k . Usando este resultado y el lema 6.7

$$\begin{aligned} P(A_k) &= P(|S_k^*| > \sqrt{2a \log k}) = P(S_k^* > \sqrt{2a \log k}) + P(S_k^* < -\sqrt{2a \log k}) \\ &\sim 1 - \Phi(\sqrt{2a \log k}) + \Phi(-\sqrt{2a \log k}) \\ &= 1 - \Phi(\sqrt{2a \log k}) + 1 - \Phi(\sqrt{2a \log k}) = 2(1 - \Phi(\sqrt{2a \log k})) \\ &\sim \frac{2e^{\frac{-2a \log k}{2}}}{\sqrt{2\pi} \sqrt{2a \log k}} = \frac{2}{2\sqrt{\pi} \sqrt{a \log k}} e^{\log k^{-a}} < e^{\log k^{-a}} = \frac{1}{k^a}, \quad a > 1 \\ &\implies \sum_{k=1}^{\infty} P(A_k) < \sum_{k=1}^{\infty} \frac{1}{k^a} < \infty. \end{aligned}$$

Por el lema de Borel-Cantelli se tiene que $P(\bar{A}) = 0$. Por lo tanto, con probabilidad 1, $A_k = \{|S_k^*| > \sqrt{2a \log k}\}$ ocurre sólo para un número finito de k 's.

Por otra parte si $\left| \frac{S_n}{n} - p \right| > \varepsilon$, entonces $\left| \frac{S_n - np}{\sqrt{n} \sqrt{pq}} \right| > \frac{\varepsilon}{\sqrt{pq}} \sqrt{n} > \sqrt{2a \log n}$ para n suficientemente grande de donde se obtiene que si se cumple $\left| \frac{S_n}{n} - p \right| > \varepsilon$ para infinitos valores de n entonces $|S_n^*| > \sqrt{2a \log n}$ para infinitos n 's lo cual contradice el hecho de que $P(\bar{A}) = 0$. Por lo tanto $\forall \varepsilon > 0$ con probabilidad 1, $\left| \frac{S_n}{n} - p \right| > \varepsilon$ puede ocurrir sólo para finitos n 's. Esto implica que, con probabilidad 1

$$\frac{S_n}{n} \rightarrow p \quad (n \rightarrow \infty)$$

Teorema 6.9 (Ley fuerte de los grandes números para variables aleatorias independientes.) Sean $X_1, X_2, \dots$ variables aleatorias independientes con $E[X_i] = 0$ y $\text{Var}[X_i] = \sigma_i^2 < \infty$. Si $\sum_{i=1}^{\infty} \frac{\sigma_i^2}{i^2} < \infty$, entonces con probabilidad 1

$$\frac{X_1 + \dots + X_n}{n} \rightarrow 0, \quad \text{cuando } n \rightarrow \infty.$$

Ejercicios 6.2

1. Sean $X_1, \dots, X_{20}$ variables aleatorias de Poisson, independientes con media igual a 1. Use el Teorema Central del Límite para aproximar $P(X_1 + \dots + X_{20} > 15)$.
2. Sea $X = \#$ caras en 40 lanzamientos de una moneda perfecta. Halle una aproximación de $P(X = 20)$.
3. El tamaño ideal de una clase en una universidad es de 150 estudiantes. La universidad sabe por experiencia que sólo 30 % de los estudiantes admitidos asistirá, por lo tanto tiene la política de admitir 450 estudiantes para la mencionada clase. Halle la probabilidad de que el número de estudiantes que asista sea mayor o igual a 150.
4. Dos tipos de monedas son producidos por una fábrica: unas son perfectamente simétricas, otras arrojan cara en un 55 % de los casos. Tenemos una de esas monedas pero no sabemos de que tipo es. Para determinar si es perfecta o no hacemos la siguiente prueba estadística: lanzamos la moneda 1000 veces y si caen 525 ó más caras concluimos que no es perfecta. Si la moneda es perfecta, ¿Cuál es la probabilidad de que lleguemos a una conclusión equivocada? ¿Cuál si la moneda no es perfecta?
5. Considere una sucesión de experimentos independientes. Sea E un evento que ocurre con probabilidad $P(E)$ y $Y_n = \#$ veces que ocurre E en n ensayos. Demuestre que con probabilidad 1 $\frac{Y_n}{n} \rightarrow P(E)$ cuando $n \rightarrow \infty$.
6. Una moneda perfecta es lanzada en ensayos independientes sucesivos. Sea $S_n =$ el total de caras después de n ensayos. Dé un estimado de $P(\frac{S_n}{n} < x)$.

7. Un evento E ocurre con probabilidad $\frac{1}{4}$ en una realización de un experimento. Dé una aproximación de la frecuencia relativa de ocurrencia de E si el experimento se realiza de forma independiente n veces, para n suficientemente grande. Justifique su respuesta.

8. Un jugador gana o pierde 1 dólar con probabilidades p y q respectivamente. Si el puede (y lo hace) seguir jugando incluso después de haber perdido todo, a) ¿Qué puede decirse de la ganancia acumulada en un número grande de juegos? b) Demuestre que si el juego no es justo y $p < 1/2$, entonces a medida que el jugador continua jugando su deuda aumenta constantemente con una probabilidad abrumadora. Sugerencia: $an + b\sqrt{n} \rightarrow -\infty$ cuando $n \rightarrow \infty$, para $a < 0$ y $b > 0$.

9. Se tienen 100 componentes que se usan uno después de otro de manera secuencial, reemplazando de inmediato un componente cuando este falla. Si el tiempo de vida de cada componente está exponencialmente distribuido con media igual a 10 horas, independientemente uno de otro, estime la probabilidad de que después de 1200 horas todavía se encuentre algún componente funcionando.

10. En una máquina de juego se gana o pierde 1 punto en cada jugada. En un día especial el casino ofrece fichas gratis a los jugadores que acumulen más de 80 puntos en los primeros 100 juegos. Los jugadores experimentados saben que en la máquina roja la probabilidad de ganar un juego es $\frac{4}{5}$. (a) ¿Cuál es la probabilidad de que un jugador gane las fichas extra si juega en la máquina roja?. (b) Si cada jugador realiza sólo 100 jugadas, ¿cuál es la probabilidad de que los primeros 8 jugadores en la máquina roja no obtengan las fichas extra?

11. Se tienen 25 componentes electrónicos cuyos tiempos de vida son variables aleatorias independientes exponencialmente distribuidas con media igual a 1 año. Si se usa sólo uno de estos componentes a la vez y cada componente dañado se reemplaza inmediatamente; halle la probabilidad de que los 25 componentes tengan un tiempo total de vida que oscile entre 30 y 40 años.

12. *Se lanza un dado perfecto consecutivamente hasta que la suma de todos los valores obtenidos sea mayor que 260. Halle la probabilidad de que se requieran al menos 65 lanzamientos.*
13. *Una bolsa contiene 10 bolas numeradas del 1 al 10. Se saca una bola al azar, se anota el número y se coloca la bola de nuevo en la bolsa. La operación se repite sucesivamente hasta que la suma de los números obtenidos sea mayor que 500. Halle la probabilidad aproximada de que se requieran al menos 82 extracciones.*
14. *Un apostador tiene probabilidad $\frac{1}{3}$ de ganar en una jugada. Si él juega indefinidamente, dé un valor aproximado del número promedio de juegos que ganará. Justifique su respuesta.*
15. *Una computadora genera una secuencia de 10000 dígitos aleatorios. Halle la probabilidad de que el dígito 7 aparezca más de 968 veces.*
16. *Al lanzar una moneda 10.000 veces se obtuvieron 5.500 caras. ¿Se puede suponer que la moneda es legal?*
17. *Calcule una aproximación de la probabilidad de que el número de 6 obtenidos al lanzar un dado perfecto 12.000 veces esté comprendido entre 1900 y 2153.*
18. *Encuentre k tal que la probabilidad de obtener entre 490 y k veces cara en 1000 lanzamientos de una moneda perfecta sea $\frac{1}{2}$.*
19. *Un dado se lanza hasta que la suma de las caras sea mayor que 300. Halle la probabilidad de que se requieran al menos 80 lanzamientos.*

Bibliografía

- [1] K.L.Chung, *Teoría Elemental de la Probabilidad y de los Procesos Estocásticos*, Ed. Reverté (1983) .
- [2] W. Feller, *Introducción a la Teoría de Probabilidades y sus Aplicaciones*, Limusa-Wiley (1973).
- [3] M.M. Olivares *Notas de Probabilidades*. Escuela de Matemáticas, UCV (2003).
- [4] S. Ross, *A First Course in Probability*, 3rd. Ed. Macmillan (1988).