Prólogo

La presente recopilación es el material para el curso trimestral de Probabilidad y Estadística impartido por José Antonio Climent Hernández a los alumnos de las licenciaturas en ingeniería de la División de Ciencias Básicas e Ingeniería (DCBI) de la Universidad Autónoma Metropolitana, Unidad Azcapotzalco (UAM–A) a partir del trimestre 2018-O. Las notas de curso incluyen temas fundamentales de un curso de Probabilidad y Estadística del tronco general (TG).

Las notas del curso de Probabilidad y Estadística están dirigidas a los alumnos de las licenciaturas en Ingeniería (Ambiental, Civil, Computación, Eléctrica, Electrónica, Física, Industrial, Mecánica, Metalúrgica o Química) con los objetivos de interpretar y aplicar resultados teóricos.

Los requisitos incluyen conocimientos adquiridos en las unidades de enseñanza aprendizaje (UEA): Taller de Matemáticas, Complementos de Matemáticas, Introducción al Cálculo, Cálculo Diferencial y Cálculo Integral.

Las notas de curso son una recopilación de temas que el autor estudió en la licenciatura de Actuaría en los cursos: Probabilidad I y II, Estadística I y II, Cálculo Actuarial I y II, la revisión bibliográfica para elaborar la tesis de licenciatura Climent Hernández (2001), la tesis de maestría Climent Hernández (2004), las notas de curso Climent Hernández (2005) y la tesis de doctorado Climent Hernández (2013), los cursos de Estadística I y II, Valuación de Opciones, Matemáticas Actuariales del Seguro de Personas I, Productos Financieros Derivados I para la licenciatura de Actuaría de la Facultad de Ciencias de Universidad Nacional Autónoma de México (UNAM) durante los semestres 2001-2 al 2009-2, Probabilidad y Estadística para la licenciatura de Ingeniería de la Universidad Tecnológica de México (UNITEC) en el trimestre 2014-1, Estadística Aplicada a la Administración I para la licenciatura de Administración de la UAM–A durante el trimestre 2014-I, Investigación de Operaciones II, Análisis de Decisiones II, Ingeniería Financiera y Probabilidad y Estadística de las licenciaturas en Ingeniería de la DCBI de la UAM–A en los trimestres 2016-I y 2016-O, 2016-I, 2017-I y 2017-O, 2018-O, 2019-P y 2020-I, respectivamente, además Introducción a la Demografía y a la Actuaría del diplomado en Sistemas de Pensiones impartido en el Centro Interamericano de Seguridad Social (CISS) durante 2012 y 2013, Taller de Probabilidad y Estadística y Laboratorio de Probabilidad y Estadística de la maestría en Ciencias Económicas de la Escuela Superior de Economía (ESE) del Instituto Politécnico Nacional (IPN) en los semestres 2013-B y 2014-A.

La revisión bibliográfica está integrada por Hoel, Port y Stone (1971), Mood y Graybill (1978), Feller (1983), Canavos (1988), Devore (1998), Hernández Arellano (2003), Hernández del Valle y Hernández Lerma (2003), Devore (2008), Rincón Solís (2008), Hildebran y Ott (1998), Levin y Rubin (2004), Wackerly, Mendelhall y Sheaffer (2008), Kazmier (1998), DeGroot y Schervish (2012), Rincón Solís (2017), Walpole, Myers, Myers y Ye (2012), Johnson (2012) y Montgomery y Runger (2003).

Las notas están dirigidas para alumnos de licenciatura que están cursando Probabilidad o Estadística y el contenido está presentado con un enfoque didáctico y práctico. Los resultados teóricos son presentados sin demostración formal y con ejemplos prácticos para favorecer el aprendizaje con aplicaciones.

La UEA de Probabilidad y Estadística, con la clave 1153001, es una UEA obligatoria (OBL) del TG de las diez licenciaturas en ingeniería que son impartidas en la DCBI de la UAM–A. La UEA de Probabilidad y Estadística tiene características que son presentadas en la tabla 1.

Tabla 1: Programa de estudios de la UEA de Probabilidad y Estadística.

Clave	1153001
Créditos	9
Tipo	TG
Horas teóricas	4.5
Horas prácticas	0.0
Seriación	1112029

La tabla 1 indica que el programa de estudios de la UEA de Probabilidad y Estadística es de 9 créditos, 4.5 horas teóricas y está seriada con la UEA de Cálculo Integral (1112029).

Los objetivos generales están centrados en conocimientos, habilidades y actitudes para que al finalizar el curso los alumnos sean capaces de:

1. Identificar la aleatoriedad de experimentos.
2. Organizar e interpretar datos y propiedades de eventos probabilistas y utilizar modelos comunes.
3. Analizar, plantear y resolver problemas de estimación y de pruebas de hipótesis.

Los objetivos generales son alcanzados a través de temas incluidos en el contenido sintético siguiente:

1. Introducción a la probabilidad y teoría de conjuntos.
2. Probabilidad axiomática, análisis combinatorio y probabilidad condicional.
3. Variables aleatorias, funciones de probabilidad, esperanza matemática y varianza.
4. Distribuciones de probabilidad discretas y continuas.
5. Introducción a la estadística, variables y escalas de medición.
6. Estadística descriptiva, gráficas, e intervalos de clase.
7. Estimación puntual, por intervalos de confianza y tamaño de la muestra.
8. Pruebas de hipótesis.

Las modalidades de conducción hacen referencia a que la UEA de Probabilidad y Estadística es una clase teórica que se refleja en las características presentadas en la tabla 1 con 4.5 horas teóricas y cero horas prácticas, la cual es impartida en modalidad de sistema de aprendizaje individualizado (SAI).

Como parte de las modalidades de conducción del proceso de enseñanza–aprendizaje:

1. Es requisito que los alumnos con apoyo del profesor, revisen y analicen al menos un texto técnico, científico o de difusión en inglés que contribuya a alcanzar los objetivos del programa de estudios.
2. Se procura que los alumnos participen en la presentación oral de trabajos, tareas u otras actividades académicas desarrolladas durante el curso.

La referencia bibliográfica necesaria que está indicada en el contenido sintético y que el autor ha revisado es Johnson (2012).

Las referencias bibliográficas recomendables que están indicadas en el contenido sintético y que el autor ha revisado son: Canavos (1988), Devore (1998, 2008), Wackerly, Mendelhall y Sheaffer (2008), Montgomery y Runger (2003) y Walpole, Myers, Myers y Ye (2012).

La UEA de Probabilidad y Estadística del quinto trimestre es OBL del TG, es cursada por alumnos de las diez licenciaturas en Ingeniería porque pertenece al TG y tiene como requisito las UEA siguientes:

1. Taller de Matemáticas[1] del primer trimestre del tronco de nivelación académica (TNA).
2. Introducción al Cálculo del segundo trimestre del TG.
3. Cálculo Diferencial del tercer trimestre del TG.
4. Cálculo Integral del cuarto trimestre del TG.

Los conocimientos adquiridos en estas UEA permite plantear y resolver problemas de interés en Ingeniería aplicando Aritmética, Álgebra, Trigonometría, Geometría, Álgebra Lineal, Cálculo Diferencial e Integral para adquirir conocimientos, habilidades y actitudes para identificar el comportamiento estocástico de eventos y aplicar modelos probabilistas con las distribuciones Bernoulli, binomial, geométrica, Pascal, Huygens, Poisson, exponencial, Gauss y gama, para organizar, resumir y analizar información, estimar parámetros y realizar pruebas de hipótesis.

La versión del 28 de noviembre de 2018 comenzó con 99 páginas.

La versión del 14 de noviembre de 2019 incluyó 8 capítulos, 4 apéndices, 146 páginas, 95 ejemplos, 80 ejercicios, 21 tablas y 54 gráficas.

La versión del 7 de diciembre de 2021 contiene 8 capítulos, 4 apéndices, 218 páginas, 105 ejemplos, 90 ejercicios, 4 aplicaciones (análisis combinatorio, distribución binomial, distribución Poisson y distribución Gauss), 29 tablas, 81 gráficas, 86 referencias bibliográficas e índice analítico.

Las notas de curso revisan temas relevantes para el curso, pero existen otros temas relevantes como: esperanza condicional, varianza condicional, transformaciones de variables aleatorias, convergencia débil, convergencia fuerte, convergencia en probabilidad, convergencia es distribución, función generadora, función generadora de momentos, función característica, ley débil de los grandes números, ley fuerte de los grandes números y las definiciones, proposiciones y teoremas relacionados.

Las notas están escritas en el sistema LaTeX y el autor agradece comentarios, sugerencias o correcciones para mejorar el contenido y es posible enviarlos a la dirección jach@azc.uam.mx.

[1] Es posible acreditar por examen.

Índice general

Índice de tablas

Índice de gráficas

Parte I

Probabilidad

Capítulo 1

Introducción

El estudio de la teoría de probabilidad comenzó en 1654 con el problema de la apuesta del juego de dados. Gambaud envió a Pascal (1654) preguntas acerca del problema de los puntos. Pascal envió los problemas a Fermat. Las soluciones se basaron en el principio de enumeración de casos con probabilidades equivalentes de ocurrencia. Huygens (1657a) escribió un tratado sobre los juegos de azar. Cardano (1663) publicó un libro con temas de la probabilidad de los juegos de azar.

Problema 1. Cada uno de los dos jugadores selecciona un número del uno al seis donde ambos números son diferentes, los jugadores apuestan una cantidad en doblones de oro y el ganador es el jugador que selecciona el número que aparece tres ocasiones antes que el número seleccionado por el adversario aparezca tres ocasiones al lanzar un dado sucesivamente. Suponiendo que el número seleccionado por uno de los jugadores ha aparecido dos ocasiones y el número del otro jugador sólo una vez, si se suspende el juego, ¿cómo es dividida la apuesta de 64 doblones de oro?

Entre 1654 y 1657 Pascal y Fermat plantean soluciones al problema 1 y a los problemas siguientes:

Problema 2. ¿Cuántas veces es necesario lanzar un par de dados para que sea más favorable obtener un par de seises que no obtenerlo?

Problema 3. Los jugadores P y Q lanzan alternadamente un par de dados, el juego comienza con el jugador P lanzando el par de dados y si la suma de los números es igual a seis, entonces el jugador P gana el juego. En caso contrario el juego continúa y el jugador Q lanza el par de dados y si la suma es igual a siete gana el juego. En caso contrario el juego continúa con el jugador P lanzando el par de dados bajo las condiciones iniciales. ¿Cuáles son las probabilidades respectivas que cada jugador tiene de ganar el juego?

Problema 4. Los jugadores A y B poseen doce fichas cada uno y juegan a lanzar tres dados de forma sucesiva, se establece que el jugador A entrega una ficha al jugador B cada vez que obtenga una suma igual a once al lanzar los tres dados, mientras que el jugador B entrega una ficha al jugador A cada vez que obtenga una suma igual a catorce. Si el ganador del juego es el primer jugador que tiene las veinticuatro fichas, ¿cuáles son las probabilidades respectivas que cada jugador tiene de ganar el juego?

Huygens (1657a) encontró soluciones para los problemas 1 y 2 y publicó las soluciones en *De Ratiociniis in Aleae Ludo* que es la base para el cálculo de probabilidades.

Bernoulli (1713) proporcionó una definición que estableció las condiciones necesarias para estudiar la teoría de probabilidad.

Kolmogorov (1931) y Kolmogorov (1933a) propuso los postulados para la teoría de probabilidad clásica.

1.1. Experimento aleatorio

En la naturaleza existen fenómenos o experimentos que son clasificados como:

1. Deterministas.
2. Aleatorios.

Definición 1 (experimento). Un experimento es un proceso que conduce a un resultado específico.

Definición 2 (experimento determinista). Un experimento determinista es un proceso en donde las condiciones en las que es realizado el experimento conducen a que el resultado esté determinado de forma única.

Ejemplo 1 (movimiento uniforme). El movimiento con velocidad constante:

$$v = \frac{d}{t},$$

donde v es la velocidad, d es la distancia recorrida y t es el tiempo necesario para recorrer la distancia. $\square$

Definición 3 (experimento aleatorio). Un experimento aleatorio $\mathscr{E}$ es un proceso en donde las condiciones en las que es realizado el experimento conducen a que el resultado no esté determinado de forma única.

Ejemplo 2 (lanzamiento de un dado). El lanzamiento de un dado tiene el conjunto de resultados:

$$\Omega = \{\text{⚀},\text{⚁},\text{⚂},\text{⚃},\text{⚄},\text{⚅}\},$$

donde el conjunto Ω contiene todos los resultados posibles del experimento aleatorio $\mathscr{E}$. $\square$

Definición 4 (realización de un experimento aleatorio). La realización de un experimento aleatorio es la repetición del experimento aleatorio $\mathscr{E}$.

Ejemplo 3 (lanzamiento de un dado dos o más veces). El lanzamiento de un dado dos o más veces es la realización del experimento aleatorio $\mathscr{E}$. $\square$

Definición 5 (evento). Un evento es un subconjunto del espacio muestra.

Ejemplo 4 (lanzamiento de un dado). El lanzamiento de un dado tiene al conjunto $\Omega = \{\text{⚀},\text{⚁},\text{⚂},\text{⚃},\text{⚄},\text{⚅}\}$ como resultados posibles y dos eventos relativos al experimento aleatorio del lanzamiento de un dado son:

1. $A = \{x \mid \text{es un número par}\} = \{\text{⚁},\text{⚃},\text{⚅}\}$, es decir $A \subset \Omega$.
2. $B = \{x \mid \text{es un número non}\} = \{\text{⚀},\text{⚂},\text{⚄}\}$, es decir $B \subset \Omega$. $\square$

La problemática en el cálculo de probabilidades es medir la ocurrencia de un evento en la realización de un experimento aleatorio $\mathscr{E}$.

Definición 6 (probabilidad). La probabilidad es una rama de las matemáticas que estudia los experimentos aleatorios.

1.2. Espacio muestra

Definición 7 (espacio muestra). El espacio muestra es el conjunto Ω que está formado por todos los resultados posibles del experimento aleatorio $\mathscr{E}$.

Ejemplo 5 (espacio muestra del lanzamiento de un dado). El espacio muestra del lanzamiento de un dado es el conjunto:

$$\Omega = \{⚀, ⚁, ⚂, ⚃, ⚄, ⚅\} = \{1,2,3,4,5,6\}\,.$$

Algunas referencias bibliográficas denotan al espacio muestra con $\mathcal{S}$ por *sample space*. □

1.3. Conjuntos

La teoría de conjuntos proporciona los fundamentos para la teoría de la probabilidad.

Definición 8 (conjunto). Un conjunto es una colección de objetos.

Los objetos son entendidos en el sentido más amplio de la palabra y no sólo como elementos físicos.

Definición 9 (elemento). Un elemento es un objeto x que pertenece a un conjunto X, es decir $x \in X$.

Definición 10 (pertenencia). Si x es un elemento del conjunto X, entonces $x \in X$.

Definición 11 (no pertenencia). Si x es un elemento que no pertenece al conjunto X, entonces $x \notin X$ o $\neg(x \in X)$.

La cantidad de elementos que pertenecen a un conjunto X clasifica al conjunto como:

1. Finito.
2. Infinito.

Definición 12 (cardinalidad de un conjunto). La cardinalidad de un conjunto es la cantidad de elementos que tiene un conjunto X, es decir $\#X$ es la cardinalidad del conjunto X.

1.3.1. Definición de conjuntos

Los conjuntos son definidos con los métodos siguientes:

1. Extensión (listar cada uno de los elementos).
2. Comprensión (enunciar las propiedades de los elementos).

Ejemplo 6 (conjunto de los números dígitos definido con el método de extensión). El conjunto de los números dígitos definido con el método de extensión:

$$\mathbb{D} = \{0,1,2,3,4,5,6,7,8,9\}\,.$$

La cardinalidad del conjunto $\mathbb{D}$ es $\#\mathbb{D} = 10$, es decir, el conjunto $\mathbb{D}$ tiene 10 elementos. □

Ejemplo 7 (conjunto de los números dígitos definido con el método de comprensión). El conjunto de los números dígitos definido con el método de comprensión:

$$Y = \{y \mid y \text{ es un dígito}\} = \{y \in \mathbb{Z} \mid 0 \leq y \leq 9\} = \{y \in \mathbb{Z} \mid -1 < y < 10\},$$

donde el conjunto $Y = \mathbb{D}$. □

Nota 1 ($Y = \mathbb{D}$). El conjunto Y definido por comprensión está expresado de tres formas diferentes que tienen el mismo significado, la misma cantidad de elementos y la misma cardinalidad que el conjunto $\mathbb{D}$.

1.3.2. Conjunto finito

Definición 13 (conjunto finito). Un conjunto es finito si tiene una cantidad conocida de elementos.

Ejemplo 8 (conjunto finito). Sea $X = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 = 0\right\}$. Calcular la cardinalidad de X, es decir, $\#X$.

Solución. La ecuación $x^2 + 3x + 2 = (x+2)(x+1) = 0$, entonces las soluciones son $x + 2 = 0$ o $x + 1 = 0$, por lo tanto $x_1 = -2$ o $x_2 = -1$, $X = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 = 0\right\} = \{-2, -1\}$, es decir X es un conjunto finito, la cardinalidad de X es dos y $\#X = 2$. □

1.3.3. Conjunto infinito

Definición 14 (conjunto infinito). Un conjunto es infinito si es un conjunto no finito.

Ejemplo 9 (conjunto infinito). Un inversionista tiene un capital de \$ 15,000 y tiene la oportunidad de invetir una parte del capital a una tasa de interés del 7 % por unidad temporal y el resto a una tasa de interés del 8 % por unidad temporal. Al final del periodo de inversión recibe más de \$ 1,125 de intereses, entonces calcular el capital a invertir a la tasa del 7 % y el resto al 8 %. Sea X la cantidad del capital que el inversionista invierte a la tasa del 7 %, calcular la cardinalidad de X.

Solución. El problema es planteado como una desigualdad lineal de una variable, donde x es la cantidad del capital que es invertido al 7 % y $1 - x$ es la parte complementaria que es invertida al 8 %. La solución de la desigualdad lineal $15,000\,(0.07x + 0.08\,(1-x)) > 1,125$ está indicando al inversionista que cantidad x del capital tiene invertir al 7 % y después deducir que cantidad $1 - x$ invertir al 8 %. Resolviendo la desigualdad lineal de una variable $1,050x + 1,200\,(1-x) > 1,125$, se tiene que $-150x > -75$, entonces $x < \frac{1}{2}$, por lo tanto la cantidad x del capital de \$ 15,000 que es invertida al 7 % es el conjunto $X = \left\{x \in \mathbb{R} \mid 0 \leq x < \frac{1}{2}\right\}$ que es el intervalo semicerrado $\left[0, \frac{1}{2}\right)$, la cantidad del capital invertida al 7 % es $x \in \left[0, \frac{1}{2}\right)$, es decir, el capital a invertir al 7 % es menor que \$ 7,500 y la cardinalidad del conjunto X es no finita y no numerable, es decir, $\#X = \infty$. □

Nota 2 (conjunto infinito). Si el espacio muestra Ω considera únicamente inversiones con centésimos, entonces el espacio muestra es un conjunto finito porque $X = \{x \in \mathbb{R} \mid 0.00 \leq x < 0.50\}$ y $\#X = 50$, por lo tanto el espacio muestra no es único porque depende del objeto de estudio y del observador.

1.3.4. Conjunto vacío

Definición 15 (conjunto vacío). Un conjunto vacío no tiene elementos y es denotado por $\emptyset$ o por Ø.

Ejemplo 10 (conjunto vacío). Sea $X = \left\{x \in \mathbb{R} \mid x^2 + 1 = 0\right\}$. Calcular la cardinalidad de X, es decir, $\#X$.
Solución. La ecuación $x^2 + 1 = 0$ tiene las soluciones complejas $x_1 = \imath = \sqrt{-1}$ y $x_2 = -\imath = -\sqrt{-1} = \bar{x}_1$, es decir, $X = \{\imath, -\imath\} \in \mathbb{C}$, entonces $X = \{\imath, -\imath\} \notin \mathbb{R}$ y no tiene solución en los números reales, por lo tanto X es un conjunto vacío, $X = \emptyset$ y la cardinalidad es cero, es decir, $\#X = 0$. □

1.3.5. Conjunto universo

Definición 16 (conjunto universo). El conjunto universo contiene únicamente a los elementos de la unión de la familia de conjuntos considerados y es denotado por $\mathcal{U}$.

Ejemplo 11 (conjunto universo). El conjunto de números complejos $\mathbb{C}$ contiene al conjunto de números reales $\mathbb{R}$ que contiene al conjunto de números racionales $\mathbb{Q}$ que contiene al conjunto de números enteros $\mathbb{Z}$ que contiene al conjunto de números naturales $\mathbb{N}$. El conjunto $\mathbb{C}$ y el conjunto $\mathbb{R}$ contienen al conjunto de los números irracionales[1] $\mathbb{Q}^c$, entonces el conjunto universo $\mathcal{U}$ de los números es representado por $\mathbb{C}$. Lo que significa que $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$, es decir, los conjuntos $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{Q}^c, \mathbb{R}$ son subconjuntos propios del conjunto $\mathbb{C}$ y la unión de los conjuntos $\mathbb{Q}$ y $\mathbb{Q}^c$ representada por el conjunto $\mathbb{Q} \bigcup \mathbb{Q}^c = \mathbb{R} \subset \mathbb{C}$.
Solución. El conjunto de los números naturales $\mathbb{N} = \{1, 2, 3, \dots\}$ no es un conjunto cerrado con respecto a la suma. Sea $x + a = 0$ donde $a \in \mathbb{N}$, entonces $x = -a \notin \mathbb{N}$, por lo tanto es necesario un conjunto que contenga números naturales $\mathbb{N}$, números negativos y al número cero.

El conjunto de los números enteros $\mathbb{Z} = \{\dots, -3, -2 - 1, 0, 1, 2, 3, \dots\}$ no es un conjunto cerrado con respecto al inverso multiplicativo. Sea $qx = p$ donde $p, q \in \mathbb{Z}$ y $q \neq 0$, de tal forma que p y q son primos relativos entre sí, es decir, $(p, q) = 1$, entonces $x = \frac{p}{q} \notin \mathbb{Z}$, por lo tanto es necesario un conjunto que contenga números enteros y números racionales $\mathbb{Q}$.

El conjunto de los números racionales $\mathbb{Q} = \left\{r = \frac{p}{q} \mid p, q \in \mathbb{Z} \text{ donde } q \neq 0\right\}$ no es cerrado con respecto a los exponentes. Sea $r = \frac{p}{q} \in \mathbb{Q}$ donde $p, q \neq 0$ y $(p, q) = 1$, entonces $x^2 = 2 \notin \mathbb{Q}$, por lo tanto $x_1 = \sqrt{2} \notin \mathbb{Q}$ y $x_2 = -\sqrt{2} \notin \mathbb{Q}$.
Proposición 1 ($\sqrt{2} \notin \mathbb{Q}$). El número $\sqrt{2}$ no pertenece al conjunto $\mathbb{Q}$.

Demostración. Suponiendo que $\sqrt{2} \in \mathbb{Q}$ y que $(p, q) = 1$, entonces $\sqrt{2} = \frac{p}{q} \Rightarrow p^2 = 2q^2$, por lo tanto p^2 es un número par, es decir, $p^2 = 2n$ y p también es un número par, es decir, $p = 2n$, por lo tanto $\forall n \in \mathbb{N} \Rightarrow p = 2n$, entonces $4n^2 = 2q^2 \Rightarrow q^2 = 2n^2$, por lo tanto q^2 es un número par y q también es un número par, entonces p y q son números pares, es decir, una contradicción $\perp$ porque si $(p, q) = 1$, entonces p y q no tienen factores comunes, por lo tanto $\sqrt{2} \notin \mathbb{Q}$. □

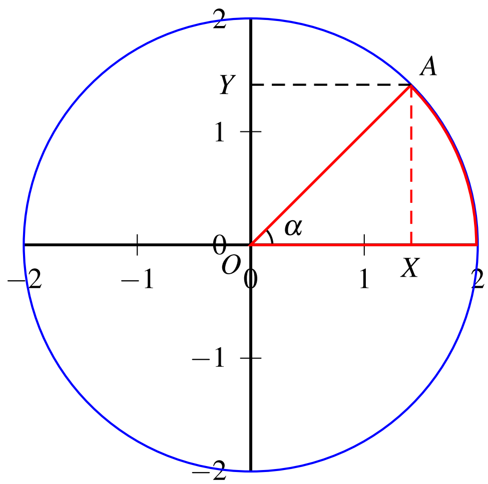
Los números racionales $\mathbb{Q}$ tienen una representación decimal que es alguno de los tipos siguientes:

1. Exacta.
2. Periódica pura.
3. Periódica mixta.

[1] Números que son imposibles de representar como el cociente de dos números enteros.

Sean $\frac{2}{5}, \frac{3}{7}, \frac{1}{12} \in \mathbb{Q}$, entonces $\frac{2}{5} = 0.4$ tiene una representación decimal exacta, $\frac{3}{7} = 0.\overline{428571}$ tiene una representación decimal periódica pura y $\frac{1}{12} = 0.08\overline{3}$ tiene una representación decimal periódica mixta.

Sea $C = \left\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 4\right\}$ el conjunto que representa una circunferencia con centro en el origen del plano cartesiano con radio de dos unidades. El conjunto C es presentado en la gráfica 1.1.



Gráfica 1.1: $\sqrt{2} \notin \mathbb{Q}$.

La gráfica 1.1 indica que la distancia $\mathrm{d}(\cdot,\cdot)$ entre los puntos $O = (0,0)$ y $A = (X,Y)$ es equivalente al radio de la circunferencia $r = \mathrm{d}(O,A) = 2$ si $\measuredangle\alpha = 45°$, entonces $\cos(\alpha) = \frac{X}{H}$, donde la hipotenusa H del triángulo rectángulo con vértices OXA y semejante al triángulo rectángulo con vértices OAY es equivalente al radio de la circunferencia, es decir, $r = H = 2$, por lo tanto $x = 2\cos(45) = 1.41421356237$.

Por el teorema de Pitágoras: $H^2 = X^2 + Y^2$ y si además $X = \mathrm{d}(O,X) = \mathrm{d}(O,Y) = Y$, entonces $4 = 2X^2$, por lo tanto $X^2 = 2$ y $A = \left(\sqrt{2}, \sqrt{2}\right) \in \mathbb{R}^2$.

El número $\sqrt{2}$ no es posible de representar como un número racional de forma exacta, periódica pura o mixta porque el teorema del binomio de Newton indica lo siguiente:

Teorema 1 (binomio de Newton). Sean $a, b, n \in \mathbb{R}$, entonces:

$$(a+b)^n = \begin{cases} \displaystyle\sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} & \text{si } n \in \mathbb{N} \\ \displaystyle\sum_{k=0}^{\infty} \binom{n}{k} a^k b^{n-k} & \text{si } n \in \mathbb{R} \end{cases}. \tag{1.1}$$

Si $n \in \mathbb{R}$ y $k \in \mathbb{Z}$, entonces el coeficiente binomial está dado por:

$$\binom{n}{k} = \begin{cases} \displaystyle\frac{1}{k!}\prod_{m=0}^{k-1} n - m = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!\,(n-k)!} & \text{si } k > 0 \\ 1 & \text{si } k = 0 \\ 0 & \text{si } k < 0 \end{cases}. \tag{1.2}$$

Si $n \in \mathbb{R}$, entonces el factorial está dado por:

$$n! = \begin{cases} \displaystyle\prod_{k=0}^{n-1} n-k = \prod_{k=1}^{n} k = n(n-1)\cdots 2\cdot 1 & \text{si } n \in \mathbb{N} \\ 1 & \text{si } n = 0 \\ \displaystyle\int_0^{\infty} t^n \exp(-t)\,\mathrm{d}t = \Gamma(n+1) & \text{si } n \in \mathbb{R} \end{cases} . \tag{1.3}$$

Por lo tanto $(a+b)^n$ es un binomio de Newton. □

Por el teorema 1 (binomio de Newton) $\sqrt{2} = (1+1)^{\frac{1}{2}} = 1 + \frac{1}{2} - \frac{1}{8} + \frac{1}{16} - \frac{5}{128} + \frac{7}{256} - \frac{21}{1{,}024} + \frac{33}{2{,}048} - \frac{429}{32{,}768} + \cdots$ es una serie (suma infinita) que oscila y converge a $\sqrt{2} \approx 1.414213562373095$ con quince dígitos decimales.

El conjunto de números que no es posible expresar como el conjunto de números racionales $\mathbb{Q}$ es el conjunto de los números irracionales $\mathbb{Q}^{\mathrm{c}}$, es decir, el conjunto de los números irracionales es la negación del conjunto de los números racionales $\neg\mathbb{Q}$ o el complemento $\mathbb{Q}^{\mathrm{c}}$ de éste.

La unión del conjunto de los números racionales $\mathbb{Q}$ y del conjunto de los números irracionales $\mathbb{Q}^{\mathrm{c}}$ es denotado mediante $\mathbb{Q} \bigcup \mathbb{Q}^{\mathrm{c}}$ y es el conjunto de los números reales $\mathbb{R} = \mathbb{Q} \bigcup \mathbb{Q}^{\mathrm{c}}$.

Las soluciones anteriores indican que el conjunto de números naturales $\mathbb{N}$ está contenido en el conjunto de números enteros $\mathbb{Z}$, el conjunto de números enteros $\mathbb{Z}$ está, a su vez, contenido en el conjunto de números racionales $\mathbb{Q}$ y el conjunto de números racionales $\mathbb{Q}$ está contenido en el conjunto de números reales $\mathbb{R}$, lo que simbólicamente es denotado $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$, por lo tanto $\forall n (n \in \mathbb{N} \to n \in \mathbb{Z} \wedge n \in \mathbb{Q} \wedge n \in \mathbb{R})$, además se tiene que $\forall z (z \in \mathbb{Z} \wedge z \in \mathbb{Q} \wedge z \in \mathbb{R})$ y finalmente $\forall r (r \in \mathbb{Q} \wedge r \in \mathbb{R})$.

Nota 3 ($\mathbb{Q}$ y $\mathbb{Q}^{\mathrm{c}}$ son conjuntos ajenos). El conjunto de los números racionales $\mathbb{Q}$ y el conjunto de los números irracionales $\mathbb{Q}^{\mathrm{c}}$ no tienen elementos en común, entonces la intersección del conjunto $\mathbb{Q}$ y del conjunto $\mathbb{Q}^{\mathrm{c}}$ es un conjunto vacío $\mathbb{Q} \bigcap \mathbb{Q}^{\mathrm{c}} = \emptyset$, es decir, son conjuntos ajenos.

Por lo tanto, el conjunto de los números irracionales $\mathbb{Q}^{\mathrm{c}}$ resuelve la ecuación $x^2 = 2$ porque $x = \pm\sqrt{2} \in \mathbb{Q}^{\mathrm{c}}$.

El conjunto de los números reales $\mathbb{R} = \mathbb{Q} \bigcup \mathbb{Q}^{\mathrm{c}}$ no es un conjunto cerrado con respecto a las potencias pares de números negativos. Sea $x^2 + 1 = 0$, entonces $x^2 = -1 \notin \mathbb{R}$ porque $\left(x + \sqrt{-1}\right)\left(x - \sqrt{-1}\right) = 0$, entonces las soluciones conjugadas $x_1 = \sqrt{-1} \notin \mathbb{R}$ y $x_2 = -\sqrt{-1} \notin \mathbb{R}$, es decir, las soluciones conjugadas $x_1 = \imath \notin \mathbb{R}$ y $x_2 = -\imath \notin \mathbb{R}$, donde $\imath = \sqrt{-1}$ es la raíz imaginaria, por lo tanto es necesario un conjunto que contenga números reales y números imaginarios, es decir, el conjunto de los números complejos $\mathbb{C}$.

El conjunto de los números complejos $\mathbb{C} = \{z = (x,y) \mid z = x + y\imath\}$ donde $x, y \in \mathbb{R}$ e $\imath^2 = -1$ o $\imath = \sqrt{-1}$ resuelve la ecuación $x^2 + 1 = 0$, porque $z = i = \sqrt{-1}$ y el número conjugado $\bar{z} = -i = -\sqrt{-1}$ son soluciones de la ecuación $x^2 + 1 = 0$, por lo tanto $X = \{\imath, -\imath\} \in \mathbb{C}$. □

1.3.6. Axiomas de la teoría de conjuntos

La teoría es presentada a partir de axiomas y siguiendo un razonamiento sistemático. Por lo tanto, los objetivos son la generalidad en el planteamiento y la precisión de la demostración.

Definición 17 (axioma). Una axioma es una proposición lógica que es admitida sin demostración porque es una premisa para deducir consecuencias y presentar resultados.

Las teorías son construidas con razonamiento lógico partiendo de axiomas que son las premisas para deducir consecuencias.

Axioma 1 (extensión). Los conjuntos A y B son iguales si tienen los mismos elementos:

$$\forall x\,(x \in A \leftrightarrow x \in B) \to A = B.$$

Ejemplo 12 (conjuntos iguales). En el ejemplo 7 los conjuntos $\mathbb{D}$ y Y son iguales porque por el axioma 1 se tiene que $\forall d\,(d \in \mathbb{D} \leftrightarrow d \in Y)$, es decir, los conjuntos $\mathbb{D}$ y Y tienen los mismos elementos, entonces $\mathbb{D} = Y$.

Solución. Se comprueba que para todo elemento, si un elemento pertenece al conjunto $\mathbb{D}$ es equivalente a que el elemento pertenece al conjunto Y, por lo tanto, los conjuntos $\mathbb{D}$ y Y son iguales, es decir, $\mathbb{D} = Y$. $\square$

Axioma 2 (comprensión). Sea A un conjunto y sea $P(x)$ una proposición abierta en el conjunto A, entonces existe el conjunto B, cuyos elementos son objetos en el conjunto A y que satisfacen la proposición $P(x)$.

Ejemplo 13 (axioma de comprensión). Sea $X = \{x \in \mathbb{Z} \mid 0 \leq x \leq 9\} = \{0,1,2,3,4,\ldots,9\}$ el conjunto de los números dígitos, sean $P(x) = \{x \in X \mid x < 5\}$ y $Q(x) = \{x \in X \mid x \equiv 0 \pmod{2}\}$ proposiciones abiertas sobre el conjunto X, entonces el conjunto $R = \{x \in X \mid x < 5\} = \{0,1,2,3,4\}$ satisface la proposición $P(x)$ y el conjunto $S = \{x \in X \mid x \equiv 0 \pmod{2}\} = \{0,2,4,6,8\}$ satisface la proposición $Q(x)$.

Solución. Los elementos del conjunto X que satisfacen la proposición abierta $P(x)$ son los elementos del conjunto $R = \{0,1,2,3,4\}$ y los elementos del conjunto X que hacen verdadera la proposición abierta $Q(x)$ son los elementos del conjunto $S = \{0,2,4,6,8\}$. El conjunto R es un subconjunto del conjunto X que tiene la notación $R \subset X$, el conjunto X no es un subconjunto del conjunto R y es utilizada la notación $\neg(X \subset R)$. Los elementos $\{0,2,4\}$ son elementos en común de los conjuntos R y S, es decir, son elementos que pertenecen a la intersección de los conjuntos R y S y es utilizada la notación $R \bigcap S = S \bigcap R = \{0,2,4\}$. $\square$

1.3.7. Subconjunto

Definición 18 (contención). El conjunto A es un subconjunto del conjunto B si y sólo si todos los elementos del conjunto A son elementos del conjunto B y la notación es $A \subseteq B$.

$$\forall x\,(x \in A \to x \in B).$$

Ejemplo 14 (contención). Sea $A = \{a, n \in \mathbb{N} \mid a = 2n\}$ y sea $B = \{b, n \in \mathbb{N} \mid b = 2^n\}$, entonces $B \subseteq A$.

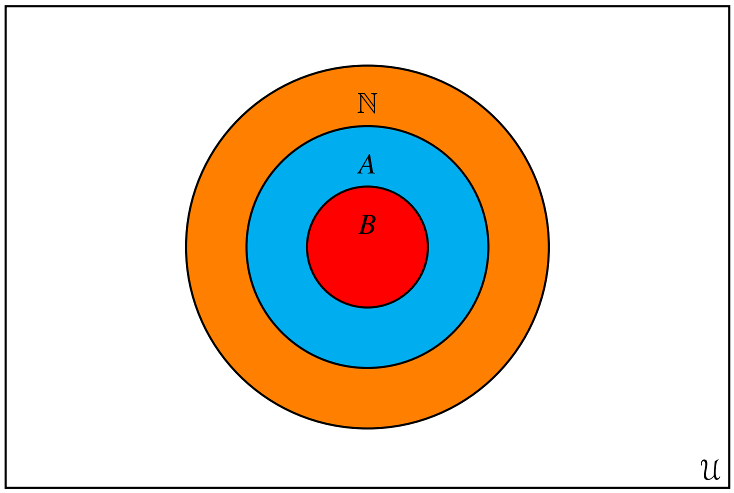
Solución. El conjunto $A = \{2,4,6,8,10,12,\ldots\}$ y el conjunto $B = \{2,4,8,16,\ldots\}$, entonces la definición 18 (contención) se satisface y $\forall x\,(x \in B \to x \in A)$, por lo tanto $B \subseteq A$. $\square$

1.3.8. Subconjunto propio

Definición 19 (contención estricta). El conjunto A es un subconjunto propio del conjunto B si y sólo si $A \subseteq B$ con $A \neq B$ y la notación es $A \subset B$.

Ejemplo 15 (contención estricta). Sea $A = \{a, n \in \mathbb{N} \mid a = 2n\}$ y sea $B = \{b, n \in \mathbb{N} \mid b = 2^n\}$, entonces $B \subset A$.
Solución. El conjunto $A = \{2, 4, 6, 8, 10, 12, \dots\}$ y el conjunto $B = \{2, 4, 8, 16, \dots\}$, entonces se satisface la definición 19 (contención estricta) y $\exists x(x \in A \to x \notin B)$, lo que significa $\neg(A \subseteq B)$, entonces $A \neq B$, por lo tanto $B \subset A$.

El conjunto B está contenido estrictamente en el conjunto A, es decir, $B \subset A$, $A \subset \mathbb{N}$, entonces $B \subset A \subset \mathbb{N}$ y la contención estricta es presentada en la gráfica 1.2.



Gráfica 1.2: $B \subset A \subset \mathbb{N}$.

La gráfica 1.2 indica que la definición por el método de extensión del conjunto A representa a los números naturales pares $A = \{2, 4, 6, 8, 10, 12 \dots\}$ y que la definición por el método de extensión del conjunto B representa a las potencias naturales del número dos $B = \{2, 4, 8, 16, 32, \dots\}$, entonces B es un subconjunto propio de A, es decir, $B \subset A$ y A es un subconjunto propio de $\mathbb{N}$, es decir, $A \subset \mathbb{N}$ y por lo tanto $B \subset A \subset \mathbb{N}$. $\square$

1.3.9. Operaciones con conjuntos

Dados los conjuntos A y B son definidos los conjuntos siguientes:

1. La unión de los conjuntos A y B tiene la notación $A \bigcup B$.
2. La intersección de los conjuntos A y B tiene la notación $A \bigcap B$.
3. La diferencia de los conjuntos A y B tiene la notación $A \setminus B$.
4. El complemento del conjunto A tiene la notación $A^c = A' = \mathcal{U} \setminus A$.

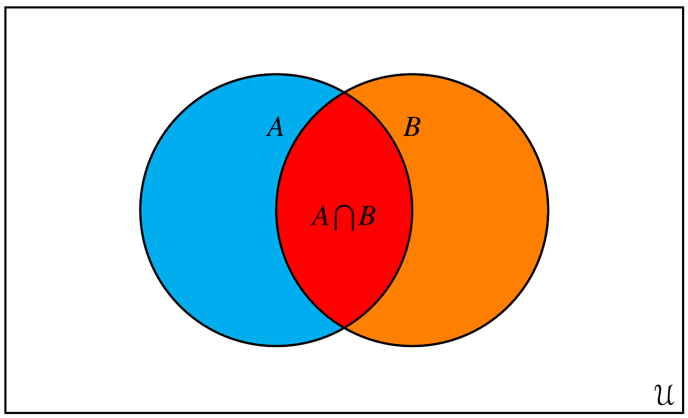
1.3.10. Unión de conjuntos

Definición 20 (unión de conjuntos). Sean A y B conjuntos diferentes del conjunto vacío $\emptyset$, entonces la unión de los conjuntos A y B es el conjunto:

$$A \bigcup B = \{x \mid x \in A \vee x \in B\} = \{x \mid x \in A \text{ o } x \in B\},$$

donde se satisface que $\forall x\,(x \in A \bigcup B \leftrightarrow (x \in A \vee x \in B))$ o $\forall x\,(x \in A \bigcup B \leftrightarrow (x \in A \text{ o } x \in B))$.

Ejemplo 16 (unión de conjuntos). La unión de los conjuntos A y B es el conjunto $A \bigcup B$ y si la intersección de los conjuntos A y B es diferente del conjunto vacío $\emptyset$, es decir, $A \bigcap B \neq \emptyset$, entonces el conjunto $A \bigcup B$ cuando $A \bigcap B \neq \emptyset$ es el conjunto representado por la gráfica 1.3.

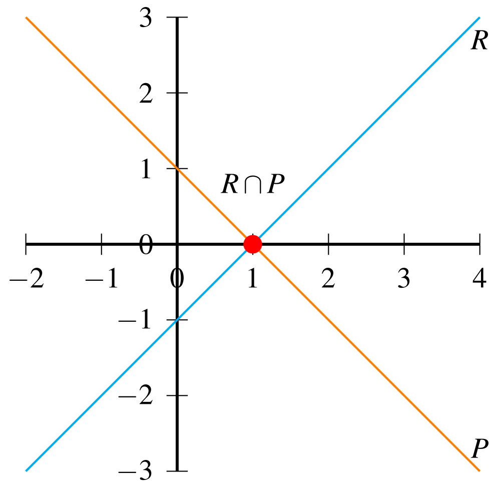


Gráfica 1.3: $A \bigcup B$ cuando $A \bigcap B \neq \emptyset$.

La gráfica 1.3 representa la unión de los conjuntos A y B cuando $A \bigcap B \neq \emptyset$, es decir, los conjuntos A y B tienen elementos en común, representados por el conjunto $A \bigcap B \neq \emptyset$. El conjunto $A \bigcup B$ está formado por los elementos que están en el conjunto A o en el conjunto B, es decir, por los elementos que están sólo en el conjunto A, por los elementos que están sólo en el conjunto B o por los elementos que están en los conjuntos A y B simultáneamente $A \bigcap B$. $\square$

Ejemplo 17 ($A \cup B$ cuando $A \cap B \neq \emptyset$). Sea $R = \left\{(x,y) \in \mathbb{R}^2 \mid y = x - 1\right\}$ el conjunto que representa la línea recta R en el plano cartesiano $\mathbb{R}^2$ y sea $P = \left\{(x,y) \in \mathbb{R}^2 \mid y = 1 - x\right\}$ el conjunto que representa la línea recta P en el plano cartesiano $\mathbb{R}^2$, encontrar el conjunto $R \bigcup P$ que representa la unión de los conjuntos R y P e indicar si el conjunto $R \bigcap P \neq \emptyset$.

Solución. El conjunto $R \bigcup P$ representa a dos líneas rectas en el plano cartesiano $\mathbb{R}^2$, $R = x - 1 = 1 - x = P$, entonces $x = 1$ e $y = 0$, por lo tanto el conjunto $R \bigcap P = \left\{(x,y) \in \mathbb{R}^2 \mid x = 1 \wedge y = 0\right\} = \{(1,0)\}$, entonces el conjunto $R \bigcup P$ tiene elementos $\left\{(x,y) \in \mathbb{R}^2 \mid (x,y) \in R \wedge (x,y) \notin P\right\}$ que pertenecen sólo al conjunto R, tiene elementos $\left\{(x,y) \in \mathbb{R}^2 \mid (x,y) \in R \vee (x,y) \in P\right\}$ que pertenecen, de forma simultanea, a los conjuntos R y P y también tiene elementos $\left\{(x,y) \in \mathbb{R}^2 \mid (x,y) \notin R \wedge (x,y) \in P\right\}$ que pertenecen sólo al conjunto P. Las líneas rectas representadas por los conjuntos R y P son perpendiculares (ortogonales), es decir, $R \perp P$ porque la línea R tiene pendiente $m_R = 1$ y la línea P tiene pendiente $m_P = -1$, por lo tanto las pendientes son recíprocas y de signo contrario $m_R m_P = -1$. El conjunto $R \bigcup B$ cuando $R \cap P \neq \emptyset$ está representado en la gráfica 1.4.



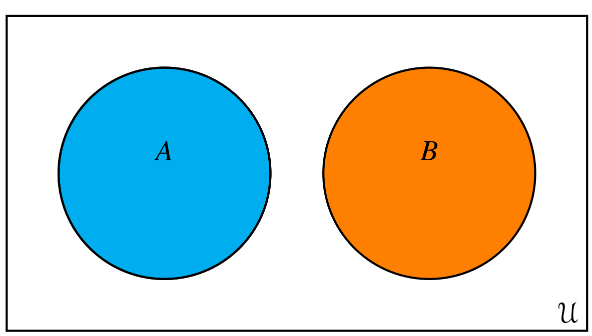
Gráfica 1.4: $R\bigcup P$ cuando $R\bigcap P \neq \emptyset$.

La gráfica 1.4 representa la unión de los conjuntos R y P cuando $P\bigcap R \neq \emptyset$, es decir, los conjuntos P y R tienen un elemento en común representado por el conjunto $A\bigcap B = \{(1,0)\}$. El conjunto $R\bigcup P$ está formado por elementos que están en el conjunto R o en el conjunto P, es decir, por elementos que están sólo en el conjunto R representados por la línea ascendente, por los elementos que están sólo en el conjunto P representados por la línea descendente y por el elemento que está simultáneamente en el conjunto R y en el conjunto P, es decir, en el conjunto $R\bigcap P$ que está representado por el punto $(1,0) \in \mathbb{R}^2$. $\square$

1.3.11. Unión de conjuntos ajenos

Definición 21 (conjuntos ajenos). Sean A, B conjuntos diferentes del conjunto $\emptyset$, entonces A y B son conjuntos ajenos si $A\bigcap B = \emptyset$.

La unión de los conjuntos A y B es el conjunto $A\bigcup B$ y si la intersección de los conjuntos A y B es el conjunto vacío $\emptyset$, es decir, $A\bigcap B = \emptyset$, entonces el conjunto $A\bigcup B$ cuando $A\bigcap B = \emptyset$ es el conjunto representado por la gráfica 1.5.



Gráfica 1.5: $A\bigcup B$ cuando $A\bigcap B = \emptyset$.

La gráfica 1.5 representa la unión de los conjuntos ajenos A y B, es decir, los conjuntos A y B no tienen elementos en común, entonces $A \bigcap B = \emptyset$.

Ejemplo 18 ($A \bigcup B$ cuando $A \bigcap B = \emptyset$). Sea $X = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 \geq 0\right\}$, entonces el conjunto X tiene dos conjuntos ajenos que son solución de la desigualdad $x^2 + 3x + 2 \geq 0$, los conjuntos ajenos están representados por el conjunto $A = \{x \in \mathbb{R} \mid -\infty < x \leq -2\} = (-\infty, -2]$ y el conjunto $B = \{x \in \mathbb{R} \mid -1 \leq x < \infty\} = [-1, \infty)$, por lo tanto el conjunto $X = A \bigcup B$ cuando $A \bigcap B = \emptyset$ es la unión de los conjuntos disjuntos A y B, es decir, los conjuntos A y B no tienen elementos en común.

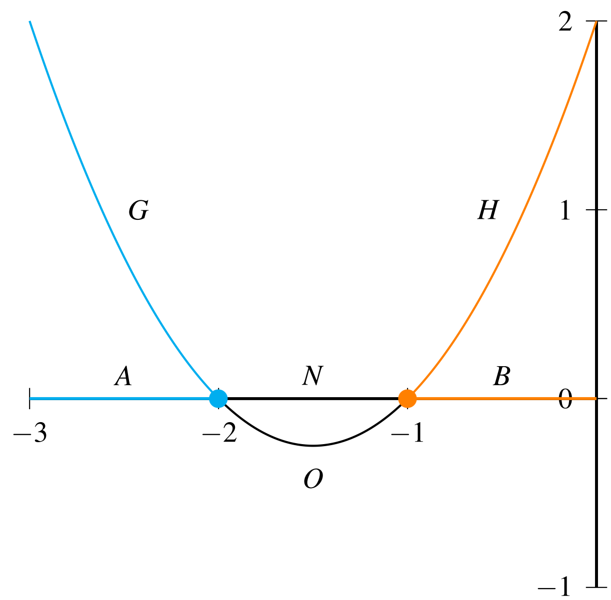
Solución. Los elementos del conjunto X que satisfacen la proposición abierta $P(x)$ en el conjunto $\mathbb{R}$ son los elementos del conjunto $X = A \bigcup B$ donde $A \bigcap B = \emptyset$. Resolviendo la igualdad $y = f(x) = 0$, es decir, los elementos del conjunto W que satisfacen la proposición abierta $Q(x) = y = 0$, entonces $W = \{-2, -1\}$ y el eje de las abscisas X está dividido en tres conjuntos:

1. $A = \{x \in \mathbb{R} \mid -\infty < x \leq -2\} = (-\infty, -2]$.
2. $N = \{x \in \mathbb{R} \mid -2 < x < -1\} = (-2, -1)$.
3. $B = \{x \in \mathbb{R} \mid -1 \leq x < \infty\} = [-1, \infty)$.

Entonces $\forall x \in A \subset \mathbb{R} \to f(x) \geq 0$, $\forall x \in N \subset \mathbb{R} \to f(x) < 0$ y $\forall x \in B \subset \mathbb{R} \to f(x) \geq 0$. Evaluando un elemento de cada conjunto A, N y B se tiene que:

1. $-3 \in A \to f(-3) = 2 \geq 0$.
2. $-\frac{3}{2} \in N \to f\left(-\frac{3}{2}\right) = -\frac{1}{4} < 0$.
3. $0 \in B \to f(0) = 2 \geq 0$.

Por lo tanto, $X = A \bigcup B = (-\infty, -2] \bigcup [-1, \infty)$ es el conjunto de números reales que satisface la proposición abierta $P(x) = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 \geq 0\right\}$ en $\mathbb{R}$. El conjunto $A \bigcup B$ cuando $A \cap B \neq \emptyset$ está representado en la gráfica 1.6.



Gráfica 1.6: $A \bigcup B$ cuando $A \bigcap B = \emptyset$.

La gráfica 1.6 presenta los conjuntos $A=\{x\in\mathbb{R}\mid -\infty<x\leq -2\}$ y $G=\left\{(x,y)\in\mathbb{R}^2\mid x\in A\wedge y\geq 0\right\}$ y los conjuntos $B=\{x\in\mathbb{R}\mid -1\leq x<\infty\}$ y $H=\left\{(x,y)\in\mathbb{R}^2\mid x\in B\wedge y\geq 0\right\}$ que satisfacen la proposición $P(x)$ en el conjunto $\mathbb{R}$. Los conjuntos abiertos $N=\{x\in\mathbb{R}\mid -2<x<-1\}$ y $O=\left\{(x,y)\in\mathbb{R}^2\mid x\in N\wedge y<0\right\}$ no satisfacen la proposición abierta $P(x)$. El conjunto $G\bigcup H$ está representando a los elementos de la función $f(x)=x^2+3x+2\geq 0$ y el conjunto $A\bigcup B$ está representando a los elementos del dominio de la función $f(x)$ que satisfacen la la proposición abierta $P(x)$. Los conjuntos A y B son conjuntos ajenos, igual que los conjuntos G y H. Por lo tanto, los conjuntos $A\bigcup B$ y $G\bigcup H$ son conjuntos disjuntos. □

Sean $A\neq\emptyset, B\neq\emptyset$ y $C\neq\emptyset$ conjuntos, entonces el conjunto $A\bigcup B$ tiene las propiedades siguientes:

Propiedad 1 (conmutativa). Sean A y B conjuntos, entonces $A\bigcup B=B\bigcup A$.

Propiedad 2 (idempotencia). Sea A un conjunto, entonces $A\bigcup A=A$.

Propiedad 3 (asociativa). Sean A,B y C conjuntos, entonces $(A\bigcup B)\bigcup C=A\bigcup(B\bigcup C)$.

Propiedad 4. Sea A un conjunto, entonces $A\bigcup\emptyset=\emptyset\bigcup A=A$.

Propiedad 5. Sean A y B conjuntos, entonces $A\subseteq(A\bigcup B)$.

Propiedad 6. Sean A y B conjuntos, entonces $A\subseteq B\Leftrightarrow A\bigcup B=B$.

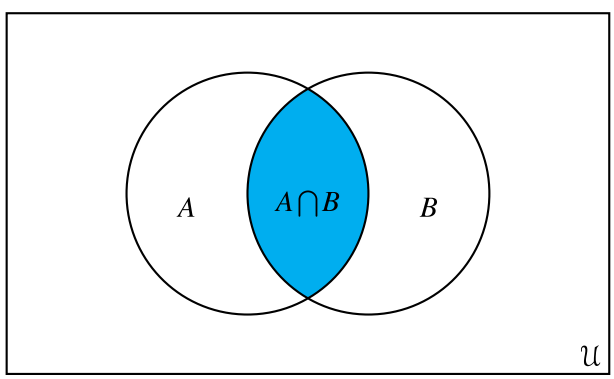
1.3.12. Intersección de conjuntos

Definición 22 (intersección de conjuntos). Sean A y B conjuntos diferentes del conjunto vacío $\emptyset$, entonces la intersección de los conjuntos A y B es el conjunto:

$$A\bigcap B=\{x\mid x\in A\wedge x\in B\}=\{x\mid x\in A \text{ y } x\in B\},$$

donde se satisface que $\forall x(x\in A\bigcap B\leftrightarrow(x\in A\wedge x\in B))$ o $\forall x(x\in A\bigcap B\leftrightarrow(x\in A \text{ y } x\in B))$.

La intersección de los conjuntos A y B es el conjunto $A\bigcap B$ que está representado por la gráfica 1.7.

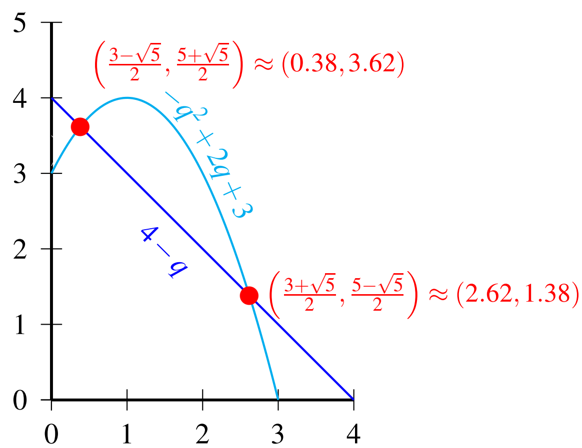


Gráfica 1.7: $A\bigcap B$.

La gráfica 1.7 representa la intersección de los conjuntos A y B, es decir, los elementos que tienen en común los conjuntos A y B. El conjunto $A\bigcap B$ está formado por los elementos que pertenecen simultáneamente al conjunto A y al conjunto B.

Ejemplo 19 ($Q \bigcap P$). Una aplicación de la teoría de conjuntos a la ingeniería se presenta al encontrar el punto de equilibrio de mercado cuando la función de la demanda es $q_D + p_D - 4 = 0$, donde q_D representa la cantidad demandada, p_D es el precio por unidad del bien o servicio y un proveedor oferta un bien o servicio a través de la función de oferta $q_O^2 - 2q_O - 3p_O = 0$, donde q_O es la cantidad ofertada y p_O es el precio por unidad del bien o servicio. Encontrar la cantidad q y el precio p de equilibrio (q^*, p^*).

Solución. El punto de equilibrio de mercado se presenta cuando la cantidad demandada q_D es igual a la cantidad ofertada q_O, es decir, cuando $q_D = q_O$ y esto ocurre en las intersecciones de las curvas de oferta y demanda. Las soluciones para las cantidades de equilibrio de mercado son obtenidas igualando la función de la oferta $p_O = -q_O^2 + 2q_O + 3$ y la función de la demanda $p_D = 4 - q_D$, entonces es obtenida la ecuación de segundo grado $q^2 - 3q + 1 = 0$ con el conjunto solución $Q = \left\{q \in \mathbb{R}_{++} \mid q^2 - 3q + 1 = 0\right\} = \left\{\frac{3-\sqrt{5}}{2}, \frac{3+\sqrt{5}}{2}\right\}$. Los precios de equilibrio de mercado son obtenidos con la sustitución de las soluciones del conjunto Q en las funciones de oferta y demanda, entonces el conjunto solución de los precios de equilibrio de mercado es $P = \left\{p \in \mathbb{R}_{++} \mid p = -q^2 + 2q + 3 \vee p = 4 - q\right\} = \left\{\frac{5+\sqrt{5}}{2}, \frac{5-\sqrt{5}}{2}\right\}$. Por lo tanto, el conjunto de equilibrio de mercado es $Q \bigcap P = \left\{(q^*, p^*) \in \mathbb{R}^2_{++} \mid p_D = p_O\right\} = \left\{\left(\frac{3-\sqrt{5}}{2}, \frac{5+\sqrt{5}}{2}\right), \left(\frac{3+\sqrt{5}}{2}, \frac{5-\sqrt{5}}{2}\right)\right\}$. El conjunto de equilibrio de mercado $Q \bigcap P$ está representado en la gráfica 1.8.



Gráfica 1.8: $Q \bigcap P$.

La gráfica 1.8 presenta la línea $4 - q$ que representa al conjunto de la función de la demanda $p_D = 4 - q_D$, la parábola $-q^2 + 2q + 3$ representa al conjunto de la función de la oferta $p_O = -q_O^2 + 2q_O + 3$ y los puntos están representando a los elementos del conjunto $Q \bigcap P$. Los dos elementos del conjunto $P \bigcap Q$ son factibles porque ambos son números reales estrictamente positivos, es decir, los dos puntos de equilibrio están ubicados en el primer cuadrante. Entonces con un precio $p^* = 3.62$ existe una demanda por la cantidad $q^* = 0.38$ de bienes o servicios con ingresos de \$ 1.38, con un precio de $p^* = 1.38$ existe una demanda por la cantidad equivalente a $q^* = 2.62$ de bienes o servicios con ingresos de \$ 3.62. El oferente tiene ingresos mayores por ofertar 2.62 bienes o servicios a un precio de \$ 1.38, que por ofertar 0.38 bienes a un precio de \$ 3.62, es decir, a un precio de \$ 1.38 existen más demandantes dispuestos a consumir los bienes o servicios que al precio de \$ 3.62. Por lo tanto, el conjunto $Q \bigcap P = \{(0.38, 3.62), (2.62, 1.38)\}$ representa a los elementos que satisfacen simultáneamente la demanda y la oferta. $\square$

Sean $A \neq \emptyset, B \neq \emptyset$ y $C \neq \emptyset$, entonces el conjunto $A \bigcap B$ tiene las propiedades siguientes:

Propiedad 7 (conmutativa). Sean A y B conjuntos, entonces $A \bigcap B = B \bigcap A$.

Propiedad 8 (idempotencia). Sea A un conjunto, entonces $A \bigcap A = A$.

Propiedad 9 (asociativa). Sean A, B y C conjuntos, entonces $(A \bigcap B) \bigcap C = A \bigcap (B \bigcap C)$.

Propiedad 10. Sea A un conjunto, entonces $A \bigcap \emptyset = \emptyset \bigcap A = \emptyset$.

Propiedad 11. Sean A y B conjuntos, entonces $(A \bigcap B) \subseteq A$.

Propiedad 12. Sean A y B conjuntos, entonces $A \subseteq B \Leftrightarrow A \bigcap B = A$.

Demostración. Para demostrar que $A \subseteq B \Rightarrow A \bigcap B = A$, entonces por hipótesis $A \subseteq B \Rightarrow \exists z \in A \bigcap B$, por la definición 22 (intersección de conjuntos) $z \in A \wedge z \in B$, entonces como $z \in A \Rightarrow A \bigcap B \subseteq A$. Si $z \in A$, entonces por hipótesis $A \subseteq B \Rightarrow z \in A \wedge z \in B \Rightarrow z \in A \bigcap B \Rightarrow A \subseteq A \bigcap B \Rightarrow A \bigcap B = A$. Para demostrar que $A \bigcap B = A \Rightarrow A \subseteq B$, entonces por hipótesis $A \bigcap B = A \Rightarrow \exists z \in A \Rightarrow z \in A \bigcap B$, por la definición 22 (intersección de conjuntos) $z \in A \wedge z \in B \Rightarrow A \subseteq B$. Por lo tanto $A \subseteq B \Leftrightarrow A \bigcap B = A$. ☐

Sean $A \neq \emptyset, B \neq \emptyset$ y $C \neq \emptyset$, entonces los conjuntos $A \bigcup B$ y $A \bigcap B$ tienen las propiedades siguientes:

Propiedad 13 (absorción). Sean A y B conjuntos, entonces $A \bigcup (A \bigcap B) = A$.

Propiedad 14 (absorción). Sean A y B conjuntos, entonces $A \bigcap (A \bigcup B) = A$.

Propiedad 15 (distributiva). Sean A, B y C conjuntos, entonces $A \bigcup (B \bigcap C) = (A \bigcup B) \bigcap (A \bigcup C)$.

Propiedad 16 (distributiva). Sean A, B y C conjuntos, entonces $A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$.

Demostración. Por la definición 22 (intersección de conjuntos) $A \bigcap (B \bigcup C) = \{x \in \mathcal{U} \mid x \in A \text{ y } x \in B \bigcup C\}$ y por la definición 20 (unión de conjuntos) $A \bigcap (B \bigcup C) = \{x \in \mathcal{U} \mid x \in A \text{ y } x \in B \text{ o } x \in C\}$, entonces por las definiciones 20 y 22:

$$A \bigcap \left(A \bigcup B\right) = \{x \in \mathcal{U} \mid x \in A \text{ y } x \in B \text{ o } x \in A \text{ y } x \in C\}.$$
$$A \bigcap \left(A \bigcup B\right) = \left\{x \in \mathcal{U} \mid x \in A \bigcap B \text{ o } x \in A \bigcap C\right\} = \left(A \bigcap B\right) \bigcup \left(A \bigcap C\right).$$

Por lo tanto, $A \bigcap (B \bigcup C) = (A \bigcap B) \bigcup (A \bigcap C)$. ☐

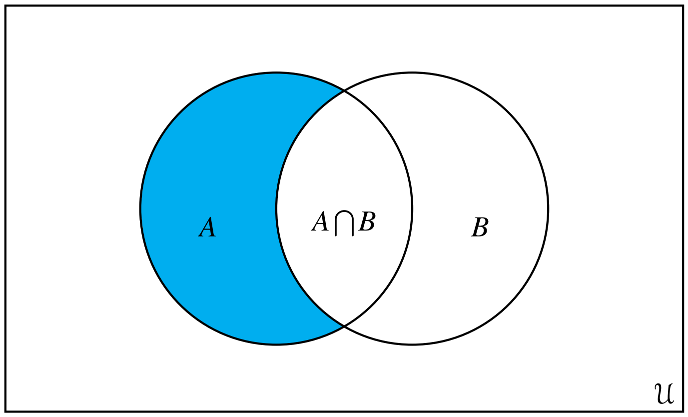
1.3.13. Diferencia de conjuntos

Definición 23 (diferencia de conjuntos). Sean A y B conjuntos diferentes del conjunto vacío $\emptyset$, entonces la diferencia de los conjuntos A y B es el conjunto:

$$A \setminus B = \{x \mid x \in A \wedge x \notin B\} = \{x \mid x \in A \text{ y } x \notin B\},$$

donde se satisface que $\forall x (x \in A \setminus B \leftrightarrow x \in A \wedge \neg (x \in B))$ o $\forall x (x \in A \setminus B \leftrightarrow x \in A \text{ y } \neg (x \in B))$.

La diferencia de los conjuntos A y B es el conjunto $A \setminus B$ representado por la gráfica 1.9.



Gráfica 1.9: $A \setminus B$.

La gráfica 1.9 representa la diferencia de los conjuntos A y B, es decir, el conjunto $A \setminus B$ representa a los elementos que pertenecen al conjunto A y no pertenecen al conjunto B.

Ejemplo 20 ($N = \mathbb{R} \setminus X$). Sea $N = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 < 0\right\}$, entonces el conjunto N es la diferencia entre el conjunto de los números reales $\mathbb{R}$ y el conjunto $X = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 \geq 0\right\}$, es decir, $N = \mathbb{R} \setminus X$ es el conjunto que satisface la proposición abierta $P(x) = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 < 0\right\}$, $N = \{x \in \mathbb{R} \mid -2 < x < -1\}$ es el conjunto que representa al intervalo abierto $(-2, -1)$ que es la diferencia entre el conjunto de los números reales $\mathbb{R}$ y el conjunto $X = A \bigcup B$, donde $A = \{x \in \mathbb{R} \mid -\infty < x \leq -2\}$ y $B = \{x \in \mathbb{R} \mid -1 \leq x < \infty\}$ son los conjuntos que satisfacen la proposición abierta $Q(x) = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 \geq 0\right\}$.

Solución. El conjunto $N = \mathbb{R} \setminus X$ es el intervalo abierto $(-2, -1)$ porque $\mathbb{R} \setminus X = \mathbb{R} \setminus (A \bigcup B)$, entonces el conjunto $N = \mathbb{R} \setminus ((-\infty, -2] \bigcup [-1, \infty)) = (-2, -1)$ satisface la proposición $P(x) = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 < 0\right\}$, como se observa en la gráfica 1.6 donde el conjunto N está representado por el intervalo $(-2, -1)$ y además el conjunto de los números reales $\mathbb{R} = (A \bigcup B) \bigcup N$, es decir, $\mathbb{R} = X \bigcup N$. $\square$

Sean $A \neq \emptyset$ y $B \neq \emptyset$, entonces el conjunto $A \setminus B$ tiene las propiedades siguientes:

Propiedad 17 (unicidad). Sean A y B conjuntos, entonces $A \setminus B = C$ donde el conjunto C es único.

Propiedad 18 (neutro). Sea A un conjunto, entonces $A \setminus \emptyset = A$ y $\emptyset \setminus A = \emptyset$.

Propiedad 19. Sea A un conjunto, entonces $A \setminus A = \emptyset$.

Propiedad 20. Sean A y B conjuntos, entonces $A \setminus B = \emptyset \Leftrightarrow A \subseteq B$.

Propiedad 21. Sean A y B conjuntos, entonces $A \setminus B = A \Leftrightarrow A \bigcap B = \emptyset$.

Propiedad 22. Sean A y B conjuntos, entonces $(A \setminus B) \bigcap (A \bigcap B) = \emptyset$.

Propiedad 23. Sean A y B conjuntos, entonces $(A \setminus B) \bigcup (A \bigcap B) = A$.

Propiedad 24. Sean A y B conjuntos, entonces $A \setminus (A \setminus B) = A \bigcap B$.

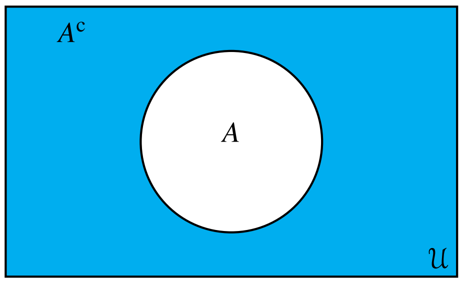
1.3.14. Conjunto complemento

Definición 24 (conjunto complemento). Sean $\mathcal{U}$ y A conjuntos diferentes del conjunto vacío $\emptyset$, entonces el complemento del conjunto A es la diferencia entre el conjunto $\mathcal{U}$ y el conjunto A:

$$A^{\mathrm{c}} = A' = \mathcal{U} \setminus A = \{x \mid x \notin A\},$$

donde se satisface que $\forall x\,(x \in \mathcal{U} \setminus A \leftrightarrow x \in \mathcal{U} \wedge \neg (x \in A))$ o $\forall x\,(x \in \mathcal{U} \setminus A \leftrightarrow x \in \mathcal{U} \text{ y } \neg (x \in A))$.

El complemento del conjunto A es el conjunto $A^{\mathrm{c}} = A' = \mathcal{U} \setminus A$ representado por la gráfica 1.10.



Gráfica 1.10: A^{c}.

La gráfica 1.10 representa al complemento del conjunto A, es decir, el conjunto A^{c} representa a los elementos que no pertenecen al conjunto A.

Ejemplo 21 (A^{c}). Sea $N = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 < 0\right\}$, entonces el conjunto $N = \{x \in \mathbb{R} \mid -2 < x < -1\}$ es el conjunto que satisface la proposición abierta $P(x) = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 < 0\right\}$ en el conjunto de los números reales $\mathbb{R}$, entonces el conjunto complemento del conjunto N es el conjunto $N^{\mathrm{c}} = \mathbb{R} \setminus N$, es decir, es el conjunto que satisface la proposición abierta $Q(x) = \left\{x \in \mathbb{R} \mid x^2 + 3x + 2 \geq 0\right\}$ en el conjunto de los números reales $\mathbb{R}$, por lo tanto $N^{\mathrm{c}} = \mathbb{R} \setminus (-2, -1) = (-\infty, -2] \bigcup [-1, \infty)$.

Solución. El conjunto $N^{\mathrm{c}} = \mathbb{R} \setminus N$ es el conjunto complemento del intervalo abierto $(-2, -1)$, entonces es la unión de los conjuntos ajenos $A = \{x \in \mathbb{R} \mid -\infty < x \leq -2\}$ y $B = \{x \in \mathbb{R} \mid -1 \leq x < \infty\}$, es decir, el conjunto complemento de N es el conjunto $N^{\mathrm{c}} = X = (-\infty, -2] \bigcup [-1, \infty)$, como se observa en la gráfica 1.6, donde el conjunto N^{c} está representado por el conjunto $A \bigcup B$ en el eje de las abscisas (los elementos que no pertenecen al conjunto N representan al conjunto N^{c}). $\square$

Sean $A \neq \emptyset$ y $B \neq \emptyset$, entonces el conjunto A^{c} tiene las propiedades siguientes:

Propiedad 25 (involutiva). Sea A un conjunto, entonces $(A^{\mathrm{c}})^{\mathrm{c}} = (A^{\mathrm{c}})^{\mathrm{c}} = A$.

Propiedad 26. Sea el conjunto $\emptyset$, entonces $\emptyset^{\mathrm{c}} = \mathcal{U}$.

Propiedad 27. Sea el conjunto $\mathcal{U}$, entonces $\mathcal{U}^{\mathrm{c}} = \emptyset$.

Propiedad 28. Sea A un conjunto, entonces $A \bigcup A^{\mathrm{c}} = \mathcal{U}$.

Propiedad 29. Sea A un conjunto, entonces $A \bigcap A^{\mathrm{c}} = \emptyset$.

Propiedad 30. Sean A y B conjuntos, entonces $A \bigcap B^{\mathrm{c}} = A \setminus B$.

Propiedad 31. Sean A y B conjuntos, entonces $A^{\mathrm{c}} \subseteq B^{\mathrm{c}} \Leftrightarrow B \subseteq A$.

1.3.15. Leyes de De Morgan

De Morgan (1847) demostró que si A y B son conjuntos, entonces los conjuntos $(A\bigcup B)^{\mathrm{c}}$ y $(A\bigcap B)^{\mathrm{c}}$ tienen las propiedades siguientes:

Propiedad 32 (primera ley de De Morgan). Sean A y B conjuntos, entonces $(A\bigcup B)^{\mathrm{c}} = A^{\mathrm{c}}\bigcap B^{\mathrm{c}}$.

Demostración. Por las definiciones 22 (intersección de conjuntos) y 24 (conjunto complemento):

$$\left(A\bigcup B\right)^{\mathrm{c}} = \left\{x\in\mathcal{U} \mid x\notin A\bigcup B\right\} = \{x\in\mathcal{U} \mid x\notin A \text{ y } x\notin B\} = \{x\in\mathcal{U} \mid x\in A^{\mathrm{c}} \text{ y } x\in B^{\mathrm{c}}\} = A^{\mathrm{c}}\bigcap B^{\mathrm{c}}.$$

Por lo tanto, $(A\bigcup B)^{\mathrm{c}} = A^{\mathrm{c}}\bigcap B^{\mathrm{c}}$. □

Propiedad 33 (segunda ley de De Morgan). Sean A y B conjuntos, entonces $(A\bigcap B)^{\mathrm{c}} = A^{\mathrm{c}}\bigcup B^{\mathrm{c}}$.

1.3.16. Conjunto potencia

Definición 25 (conjunto potencia). El conjunto potencia del conjunto Ω, denotado por 2^{Ω}, es el conjunto que está constituido por todos los subconjuntos posibles del conjunto Ω y si $\#\Omega < \infty$, entonces:

$$\#2^{\Omega} = 2^{\#\Omega}, \tag{1.4}$$

donde $\#2^{\Omega}$ es la cardinalidad del conjunto potencia y $\#\Omega$ es la cardinalidad del espacio muestra.

Ejemplo 22 (conjunto potencia). Sea $\Omega = \{\heartsuit, \diamondsuit, \clubsuit\}$, encontrar:

1. El conjunto potencia del conjunto Ω, denotado por 2^{Ω}.
2. Calcular la cardinalidad del conjunto potencia, es decir, $\#2^{\Omega}$.

Solución. El conjunto potencia 2^{Ω} es obtenido comenzando por el conjunto $\emptyset\subset\Omega$, después, los elementos son ordenados individualmente, por pares y así sucesivamente hasta obtener el conjunto Ω, entonces:

$$2^{\Omega} = \{\emptyset, \{\heartsuit\}, \{\diamondsuit\}, \{\clubsuit\}, \{\heartsuit,\diamondsuit\}, \{\heartsuit,\clubsuit\}, \{\diamondsuit,\clubsuit\}, \Omega\}.$$

La cardinalidad del conjunto 2^{Ω} es $\#2^{\Omega} = 2^{\#\Omega} = 2^3 = 8$, por lo tanto el conjunto 2^{Ω} tiene ocho elementos: el conjunto vacío $\emptyset$, tres conjuntos con elementos ordenados individualmente, tres conjuntos con elementos ordenados por pares y un conjunto con elementos ordenados por ternas, que es el conjunto Ω. □

1.4. Relaciones

Definición 26 (par ordenado). Sean $x, y\in A$, entonces se tiene el conjunto de la forma:

$$(x,y) = \{\{x\}, \{x,y\}\},$$

donde x es el primer componente y y es el segundo componente del par ordenado.

Los pares ordenados tienen las propiedades siguientes:

Propiedad 34 (conjunto singular). Si $x = y$, entonces:

$$\forall x, y\,(x = y \to (x,y) = (x,x) = \{\{x\},\{x,x\}\} = \{\{x\},\{x\}\} = \{\{x\}\}).$$

Propiedad 35 (conjunto identidad). Si $x, y, \upsilon, \nu \in A$, entonces:

$$\forall x, y, \upsilon, \nu\,((x,y) = (\upsilon,\nu) \leftrightarrow (x = \upsilon \wedge y = \nu) = (x = \upsilon \text{ y } y = \nu)).$$

1.4.1. Producto cartesiano

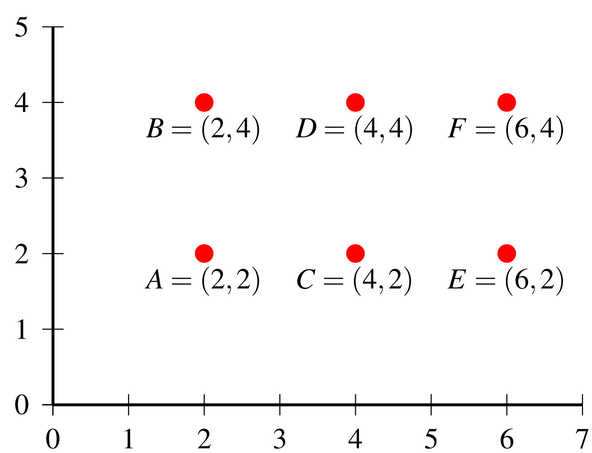
Definición 27 (producto cartesiano). Sean X y Y conjuntos, entonces el conjunto $X \times Y$ de pares ordenados denotado por:

$$X \times Y = \{(x,y) \mid x \in X \wedge y \in Y\} = \{(x,y) \mid x \in X \text{ y } y \in Y\}.$$

Ejemplo 23 (producto cartesiano). Sean $X = \{x \in \mathbb{N} \mid x = 2n\ \forall n = 1,2,3\}$ y $Y = \{y \in \mathbb{N} \mid y = 2n\ \forall n = 1,2\}$ conjuntos. Describir a los conjuntos X, Y y $X \times Y$ por el método de extensión, calcular $\#(X \times Y)$ y presentar la gráfica del conjunto $X \times Y$ en el plano cartesiano.

Solución. Los conjuntos X, Y y $X \times Y$ descritos con del método de extensión son $X = \{2,4,6\}$, $Y = \{2,4\}$ y $X \times Y = \{(2,2),(2,4),(4,2),(4,4),(6,2),(6,4)\}$. La cardinalidad del conjunto $X \times Y$ es $\#(X \times Y) = 6$. Los seis pares ordenados están representados por el conjunto $P = \{(x,y) \mid x \in X \wedge y \in Y\}$, donde el eje X es el eje de las abscisas y el eje Y es el eje de las ordenadas.

La gráfica del producto cartesiano consta de seis puntos que están ubicados en el primer cuadrante $(\forall x, y \geq 0)$ del plano cartesiano, representados por el conjunto $X \times Y$ que es presentado en la gráfica 1.11.



Gráfica 1.11: $X \times Y$.

La gráfica 1.11 presenta los seis puntos $A = (2,2), B = (2,4), C = (4,2), D = (4,4), E = (6,2), F = (6,4)$ que son los seis elementos del conjunto $X \times Y = \{(2,2),(2,4),(4,2),(4,4),(6,2),(6,4)\}$ que representa el producto cartesiano de los conjuntos $X = \{2,4,6\}$ y $Y = \{2,4\}$. $\square$

Sean $A \neq \emptyset$ y $B \neq \emptyset$, entonces $A \times B$ tiene las propiedades siguientes:

Propiedad 36 (no conmutatividad). Sean A, B conjuntos, entonces $\forall a \in A, b \in B$ con $a \neq b \rightarrow A \times B \neq B \times A$.

Propiedad 37 (principio de multiplicación). Sean A, B conjuntos, si $\#A = m$ y $\#B = n$, entonces la cardinalidad del producto cartesiano es $\#(A \times B) = \#A\#B = mn$.

Nota 4 (principio de multiplicación). Sean A, B conjuntos, entonces $\#(A \times B) = \#A\#B = \#B\#A = \#(B \times A)$.

El principio de multiplicación es aplicado en los procesos de conteo (propuesto por Bernoulli 1713).

Propiedad 38 (principio de multiplicación para n conjuntos). Sean $A_k \neq \emptyset$ $\forall k = 1, \cdots, n$ donde $\#A_k < \infty$, entonces:

$$\#(A_1 \times A_2 \times \cdots \times A_n) = \prod_{k=1}^{n} \#A_k. \tag{1.5}$$

Ejemplo 24 (principio de multiplicación). Un tenista que está sirviendo tiene dos oportunidades:

1. Primer servicio.
2. Segundo servicio.

En el primer servicio tiene las resultados siguientes siguientes:

1. Un servicio as (sin respuesta) en el primer servicio y gana un punto.
2. Un servicio con devolución del adversario (es posible ganar o perder el punto).
3. Un servicio fuera de la zona (falta) y tiene oportunidad del segundo servicio.
4. Un servicio dentro de la zona permitida pero que tocó la red y repite el primer servicio.

Si el tenista comete falta en el primer servicio, entonces en el segundo servicio es posible obtener uno de los cuatro resultados siguientes:

1. Un servicio as (sin respuesta) en el segundo servicio y gana un punto.
2. Un servicio con devolución del adversario (es posible ganar o perder el punto).
3. Un servicio fuera de la zona (falta) en el segundo servicio y pierde un punto.
4. Un servicio dentro de la zona permitida pero que tocó la red y repite el segundo servicio.

Indicar cuántas formas posibles tiene el tenista para:

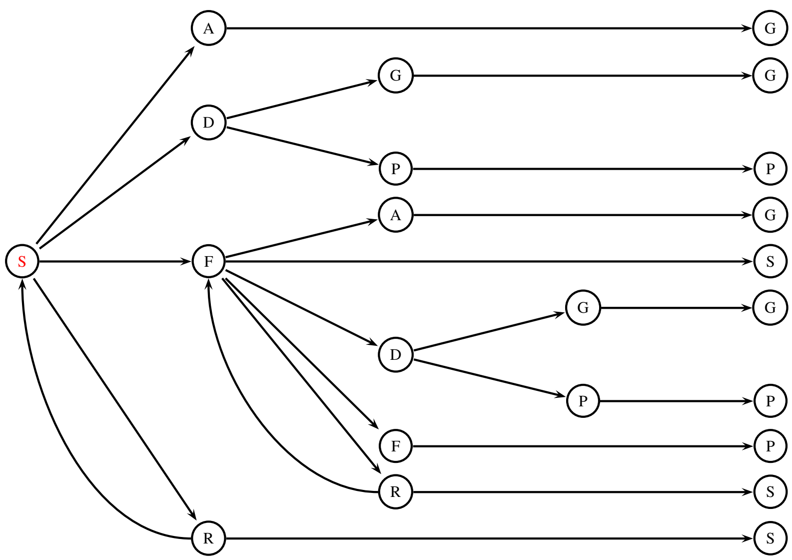
1. Ganar un punto.
2. Perder un punto.
3. Realizar un servicio sin ganar o perder el punto.

Solución. Las formas posibles que el tenista tiene para:

1. Ganar un punto son:
   a) Obtener un servicio as en el primer servicio.
   b) Obtener un servicio con respuesta en el primer servicio y ganar el punto.
   c) Cometer falta en el primer servicio y obtener un servicio as en el segundo servicio.
   d) Cometer falta en el primer servicio, obtener un servicio con respuesta en el segundo servicio y ganar el punto.

2. Perder un punto son:
   a) Obtener un servicio con respuesta en el primer servicio y perder el punto.
   b) Cometer falta en el primer servicio, obtener un servicio con respuesta en el segundo servicio y perder el punto.
   c) Cometer falta en el primer servicio y cometer falta en el segundo servicio.
3. Realizar un servicio sin ganar o perder el punto.
   a) Cometer falta en el primer servicio.
   b) Realizar un servicio dentro de la zona permitida pero que tocó la red y repetir el primer servicio.
   c) Realizar un servicio dentro de la zona permitida pero que tocó la red y repetir el segundo servicio.

Las resultados que el tenista tiene son presentados en el diagrama de árbol de la gráfica 1.12.



Gráfica 1.12: Diagrama de árbol.

Por lo tanto, es posible que el tenista gane un punto de cuatro formas, que pierda un punto de tres formas y que realice un servicio sin ganar o perder el punto de tres formas. □

1.5. Ejercicios

Ejercicio 1. Determinar un espacio muestra para los experimentos aleatorios siguientes:

1. La posición de una partícula que se mueve en un espacio tridimensional en un instante dado.
2. Lanzar un dado hasta obtener un ⚂.
3. Lanzar tres monedas distinguibles (©,®,®) al mismo tiempo.
4. Lanzar tres monedas indistinguibles (©,©,®) al mismo tiempo.

5. Lanzar dos dados distinguibles (⚄,⚃) y considerar los eventos siguientes:
    a) $A = \{(x,y) \mid x+y=7\}$.
    b) $B = \{(x,y) \mid$ un dado obtiene un número par y el otro obtiene un número non$\}$.
    c) $C = A \bigcap B$.
    d) $D = B^{\mathrm{c}}$.

Ejercicio 2. Sea $A = \left\{x \in \mathbb{R} \mid x^2+2x+1=0\right\}$, calcular la cardinalidad del conjunto A, es decir, $\#A$.

Ejercicio 3. Sea $B = \left\{x \in \mathbb{C} \mid x^2+1=0\right\}$, calcular la cardinalidad del conjunto B, es decir, $\#B$.

Ejercicio 4. Sea $R = \left\{(x,y) \in \mathbb{R}^2 \mid (1,1)+t\,(1,1)\,\forall t \in \mathbb{R}\right\}$ y sea $S = \left\{(x,y) \in \mathbb{R}^2 \mid (0,1)+t\,(1,1)\,\forall t \in \mathbb{R}\right\}$.

1. Encontrar el conjunto T que representa la intersección de R y S denotado por $T = R \bigcap S$.
2. Calcular la cardinalidad del conjunto T, es decir, $\#T$.

Ejercicio 5. Dibujar un diagrama de Venn–Euler para cada caso:

1. Si $A \bigcap B = \emptyset \Rightarrow A \subseteq B^{\mathrm{c}}$.
2. Si $A \subseteq B \Rightarrow A \bigcup (B \setminus A) = B$.
3. Si $(A^{\mathrm{c}} \bigcap B) \bigcup (A \bigcap B^{\mathrm{c}}) = (A \bigcap B)^{\mathrm{c}}$.

Ejercicio 6. Demostrar que $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$.

Ejercicio 7. Demostrar que $(A \cap B) \cap (A \cap B^{\mathrm{c}}) = (A \cap B) \cap (A \setminus B) = \emptyset$.

Ejercicio 8. Demostrar que $A = (A \cap B) \cup (A \cap B^{\mathrm{c}}) = (A \cap B) \cup (A \setminus B)$.

Ejercicio 9. Suponiendo que un juego con 20 cartas contiene 10 cartas rojas numeradas del 1 al 10, es decir, el conjunto $\{1_r, \ldots, 10_r\}$ y 10 cartas azules numeradas del 1 al 10, es decir, el conjunto $\{1_a, \ldots, 10_a\}$ y son seleccionados los conjuntos siguientes:

1. $A = \{x \mid x = 2n\ \forall 5 \geq n \in \mathbb{N}\} = \{2_a, 4_a, 6_a, 8_a, 10_a, 2_r, 4_r, 6_r, 8_r, 10_r\}$.
2. $B = \{x \mid x \text{ es azul}\} = \{1_a, 2_a, 3_a, 4_a, 5_a, 6_a, 7_a, 8_a, 9_a, 10_a\}$.
3. $C = \{x \mid x < 5\} = \{1_a, 2_a, 3_a, 4_a, 1_r, 2_r, 3_r, 4_r\}$.

Describir el espacio muestra Ω y describir cada uno de los eventos siguientes:

1. $A \cap B \cap C$.
2. $B \cap C^{\mathrm{c}}$.
3. $A \cup B \cup C$.
4. $A \cap (B \cup C)$.
5. $A^{\mathrm{c}} \cap B^{\mathrm{c}} \cap C^{\mathrm{c}}$.

Ejercicio 10. Una persona:

1. Tiene 8 camisas, 5 pantalones y 2 pares de zapatos. Calcular las formas diferentes como la persona se puede vestir combinando estas prendas (sin mezclar los pares de zapatos).
2. Tiene 6 blusas, 4 faldas y 3 pares de zapatos. Calcular las formas diferentes como la persona se puede vestir combinando estas prendas (sin mezclar los pares de zapatos).

Capítulo 2

Cálculo de probabilidades

La probabilidad de un evento E es un número real en el intervalo cerrado $[0,1]$ que representa la frecuencia de la ocurrencia del evento E cuando el experimento aleatorio $\mathscr{E}$ es realizado, es denotada por $\mathscr{P}(E)$ y es una medida tal que $0 \leq \mathscr{P}(E) \leq 1$.

Son presentadas cuatro definiciones de probabilidad:

1. Probabilidad clásica (Laplace 1812).
2. Probabilidad geométrica (Arbuthnot 1692).
3. Probabilidad frecuentista (Bernoulli 1713).
4. Probabilidad subjetiva (Ramsey 1931).
5. Probabilidad axiomática (Kolmogorov 1933a).

2.1. Probabilidad clásica

Definición 28 (probabilidad clásica). Sea $E \subset \Omega$ donde $\#E < \infty$, entonces:

$$\mathscr{P}(E) = \frac{\#E}{\#\Omega}, \tag{2.1}$$

donde $\mathscr{P}(E)$ es la probabilidad del evento E, $\#E$ es la cardinalidad del conjunto E y $\#\Omega$ es la cardinalidad del conjunto Ω.

La probabilidad clásica es conocida como probabilidad de Laplace, quien estableció formalmente los principios y las propiedades para el cálculo de probabilidades en un espacio muestra finito y con probabilidades equivalentes.

Ejemplo 25 (seleccionar bolas de una urna). Un experimento aleatorio consiste en seleccionar una bola de una urna regresando la bola a la urna. La urna contiene 2 bolas negras y 3 bolas blancas. Calcular la probabilidad de que la bola seleccionada sea blanca y la probabilidad de que la bola seleccionada sea negra.

Solución. El experimento aleatorio tiene el espacio muestra $\Omega = \{\bullet, \bullet, \circ, \circ, \circ\}$, entonces $\#\Omega = 5$ y $\forall \omega \in \Omega$ se tiene que $\mathscr{P}(\omega) = \frac{1}{5}$. Sean $B = \{b \mid b \text{ es una bola blanca}\}$ y $N = \{n \mid n \text{ es una bola negra}\}$, entonces $\#B = 3$ y $\#N = 2$, por lo tanto, sustituyendo en la ecuación (2.1) se tiene que $\mathscr{P}(B) = \frac{3}{5}$ y $\mathscr{P}(N) = \frac{2}{5}$. □

La definición 28 (probabilidad clásica) satisface las propiedades siguientes:

1. $\mathscr{P}(\Omega) = 1$.
2. $\mathscr{P}(A) \geq 0$ y $\mathscr{P}(B) \geq 0$.
3. $\mathscr{P}(A \bigcup B) = \mathscr{P}(A) + \mathscr{P}(B)$ cuando $\mathscr{P}(A \bigcap B) = 0$.

2.2. Probabilidad geométrica

Definición 29 (probabilidad geométrica). Si un experimento aleatorio $\mathscr{E}$ tiene un espacio muestra $\Omega \subset \mathbb{R}^n$ con una medida definida y finita $H(\Omega) < \infty$, entonces la probabilidad geométrica de un evento $E \subset \Omega$ es:

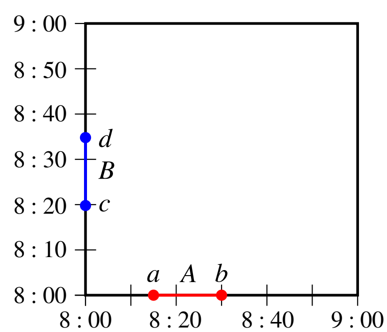
$$\mathscr{P}(E) = \frac{H(E)}{H(\Omega)}, \tag{2.2}$$

donde $\mathscr{P}(E)$ es la probabilidad geométrica del evento E, $H(E)$ es la medida del conjunto E y $H(\Omega)$ es la medida del conjunto Ω (la aplicación geométrica fue propuesta por Arbuthnot 1692).

La ecuación (2.2) supone que el espacio muestra tiene probabilidades equivalentes; en el contexto de que la probabilidad de observar la ocurrencia de un evento E depende de la medida $H(E)$ y no del conjunto.

Ejemplo 26 (encuentro de dos personas). Las personas X y Y se citan en un lugar entre las 8 y las 9 horas y acuerdan que el tiempo de espera es a lo más de 15 minutos y hasta las 9 horas. Calcular la probabilidad de que las personas coincidan durante la hora acordada.

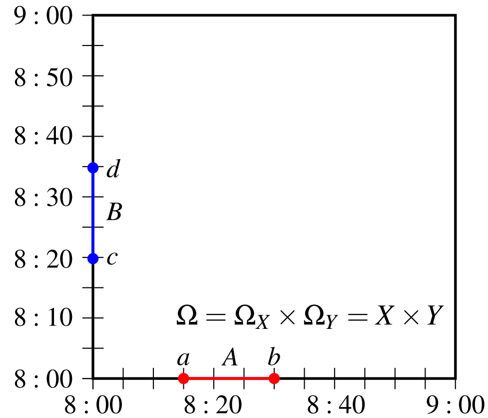
Solución. El tiempo de llegada de X es $8 \leq X \leq 9$, entonces $\Omega_X = \{X \mid 0 \leq X \leq 60\}$ y el tiempo de llegada de Y es $8 \leq Y \leq 9$, entonces $\Omega_Y = \{Y \mid 0 \leq Y \leq 60\}$. Por lo tanto, la longitud de Ω_X y Ω_Y es $\mathrm{d}(0, 60) = 60$. Sea $A = \{X \mid 0 \leq a \leq X \leq b \leq 60\}$ el evento que representa que X llega entre las a y las b horas, entonces sustituyendo en la ecuación (2.2), $\mathscr{P}(A) = \frac{b-a}{60}$. Sea $B = \{Y \mid 0 \leq c \leq Y \leq d \leq 60\}$ el evento que representa que Y llega entre las c y las d horas, entonces $\mathscr{P}(B) = \frac{d-c}{60}$, como es presentado en la gráfica 2.1.



Gráfica 2.1: Tiempos de espera de A y B.

La gráfica 2.1 presenta la longitud de A que representa al evento X, llegar entre las 8:15 y las 8:30 horas, es decir $\mathrm{d}(a,b) = b - a$ que representa 15 minutos y la longitud del conjunto $\Omega = X$ es 60 minutos; la longitud de B representa al evento Y, llegar entre las 8:20 y las 8:35 horas, es decir, $\mathrm{d}(c,d) = b - a$ que representa 15 minutos y la longitud del conjunto $\Omega = Y$ es 60 minutos.

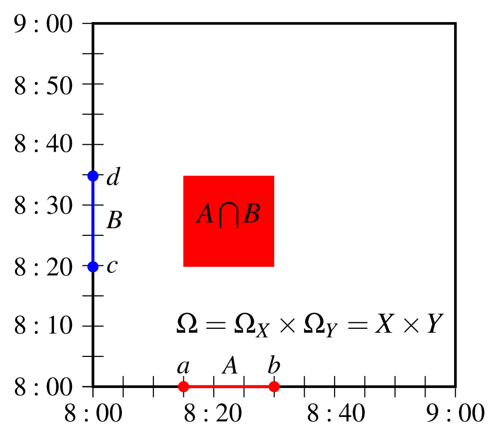
Los tiempos de llegada a la cita están representados por la pareja (X,Y) donde $8 \leq X \leq 9$ es el tiempo de llegada de X y $8 \leq Y \leq 9$ es el tiempo de llegada de Y, entonces los tiempos de llegada están representados por el cuadrado $\Omega = X \times Y$ que tiene un área de 60 minutos cuadrados que es presentado en la gráfica 2.2.



Gráfica 2.2: Tiempos de espera $A \times B \subset \Omega$.

La gráfica 2.2 presenta la medida de $\Omega = \Omega_X \times \Omega_Y = X \times Y$, es decir, $H(\Omega)$ que tiene un área de 60^2.

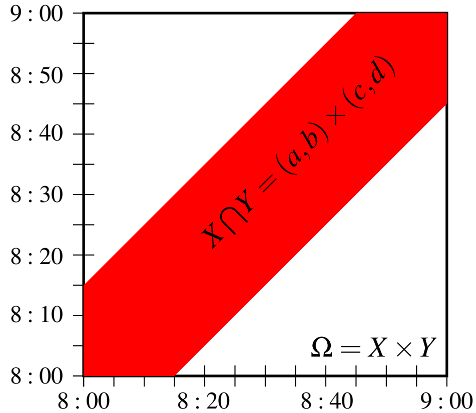
Si los tiempos de llegada son independientes, entonces $\mathscr{P}(A \bigcap B) = \frac{(b-a)(d-c)}{60^2}$. Por lo tanto, si los tiempos de espera de X y Y están ubicados en cuadrilátero $A \times B = (a,b) \times (c,d)$, sin la restricción de los 15 minutos tiene el área $\mathrm{d}(a,b) \cdot \mathrm{d}(c,d)$ y que ambas lleguen entre 8 y 9 está representado por el conjunto $\Omega = X \times Y$ que tiene el área 60^2 como es presentado en la gráfica 2.3.



Gráfica 2.3: Tiempos de espera $A \bigcap B \subset \Omega$.

La gráfica 2.3 presenta la medida de $\Omega = \Omega_X \times \Omega_Y = X \times Y$, es decir, $H(\Omega)$ que tiene un área de 60^2 y la medida de $A \bigcap B$, es decir, $H(A \bigcap B)$ que tiene un área de 15^2, entonces $\mathscr{P}(A \bigcap B) = \frac{15^2}{60^2} = \frac{1}{16} = 0.0625\ \%$.

Si X y Y coinciden en el lugar de la cita se satisface la condición necesaria y suficiente $|X-Y| \leq 15$, entonces el encuentro de X y Y se encuentra en el área $(b-a)(b+a) = b^2 - a^2$ porque[1] si X llega en el instante $0 \leq a \leq 60$, espera a Y hasta el instante b y después llega Y en el instante $b+a \leq 60$.



Gráfica 2.4: Tiempos de espera $X \bigcap Y \subset \Omega$.

La gráfica 2.4 presenta la región en donde X y Y coinciden entre las 8 y las 9 horas y la medida $H(X \bigcap Y)$ es el área $60^2 - 45^2$, por lo tanto $\mathscr{P}(X \bigcap Y) = \frac{60^2 - 45^2}{60^2} = \frac{1{,}575}{3{,}600} = \frac{7}{16} = 0.4375 = 43.75\ \%$. $\square$

La definición 29 (probabilidad geométrica) satisface las propiedades siguientes:

1. $\mathscr{P}(\Omega) = 1$.
2. $\mathscr{P}(A) \geq 0$ y $\mathscr{P}(B) \geq 0$.
3. $\mathscr{P}(A \bigcup B) = \mathscr{P}(A) + \mathscr{P}(B)$ cuando $\mathscr{P}(A \bigcap B) = 0$.

2.3. Probabilidad frecuentista

Si un experimento aleatorio $\mathscr{E}$ es realizado n veces, registrando el número de ocurrencias del evento E como la frecuencia absoluta f_E, entonces $\mathscr{P}(E)$ es calculada en función del número de ocurrencias acumuladas del evento E, donde el número de ocurrencias acumuladas es la frecuencia acumulada F_E.

[1] El resultado es equivalente a resolver:

$$\int_0^{60-x} 15 + t\,\mathrm{d}t - \int_x^{60} t - 15\,\mathrm{d}t + 60\int_{60-x}^{60} \mathrm{d}t = \left.\frac{t(30+t)}{2}\right|_0^{45} - \left.\frac{t(t-30)}{2}\right|_{15}^{60} + \left.60t\right|_{45}^{60} = \frac{3{,}375}{2} - \frac{2{,}025}{2} + 900 = 1{,}575,$$

donde x es el tiempo de espera acordado por X y Y.

Definición 30 (frecuencia absoluta). La frecuencia absoluta del evento E es el número de veces que se presenta el evento E en las n realizaciones del experimento aleatorio y es denotada por f_E, la k–ésima frecuencia absoluta del evento E es denotada por $f_k = 1_k$ y es una función indicadora que indica si se presenta o no el evento E en la k–ésima realización del experimento aleatorio, entonces $f_k = 1$ si se presenta el evento E y $f_k = 0$ si no se presenta el evento E, entonces:

$$f_E = \sum_{k=1}^{n} f_k. \tag{2.3}$$

Si $\Omega = \bigcup_{j=1}^{m} E_j$ donde $\forall i \neq j$ se tiene que $E_j \bigcap E_i = \emptyset$, entonces $\sum_{j=1}^{m} f_j = n$.

Ejemplo 27 (frecuencia absoluta). La UEA de Probabilidad y Estadística del trimestre 2018-O (grupo CSI02) inició con 44 asistentes, 42 de alguna de las 10 licenciaturas en ingeniería y 2 de la licenciatura en diseño industrial. El experimento aleatorio $\mathscr{E}$ consistió en preguntar a cada alumno a qué licenciatura pertenece y registrar la respuesta. Los resultados de las 44 preguntas son presentados en la tabla 2.1.

Tabla 2.1: Licenciaturas del grupo CSI02 del trimestre 2018-O.

Alumno	Licenciatura	Alumno	Licenciatura
1	Electrónica	23	Diseño
2	Metalúrgica	24	Física
3	Civil	25	Electrónica
4	Computación	26	Metalúrgica
5	Ambiental	27	Eléctrica
6	Civil	28	Computación
7	Química	29	Diseño
8	Industrial	30	Civil
9	Física	31	Metalúrgica
10	Química	32	Mecánica
11	Civil	33	Física
12	Metalúrgica	34	Computación
13	Industrial	35	Eléctrica
14	Electrónica	36	Ambiental
15	Física	37	Mecánica
16	Física	38	Mecánica
17	Metalúrgica	39	Eléctrica
18	Computación	40	Computación
19	Metalúrgica	41	Civil
20	Física	42	Metalúrgica
21	Metalúrgica	43	Física
22	Física	44	Civil

Calcular la frecuencia absoluta para cada una de las once licenciaturas.

Solución. El espacio muestra por clave de plan es $\Omega = \{1,2,3,4,5,6,7,8,9,19,122\}$, que corresponde a las licenciaturas: Ambiental, Civil, Eléctrica, Física, Industrial, Mecánica, Metalúrgica, Química, Electrónica, Diseño y Computación con los eventos $A, C, E, F, I, M, T, Q, L, D$ y O, entonces las frecuencias absolutas son:

$$
\begin{array}{llllll}
f_A = 2, & f_C = 6, & f_E = 3 & f_F = 8, & f_I = 2, & f_M = 3, \\
f_T = 8, & f_Q = 2, & f_L = 3, & f_D = 2, & f_O = 5. &
\end{array}
$$

La cardinalidad del espacio muestra es $\#\Omega = 11$, entonces:

$$\sum_{j=1}^{11} f_j = 2+6+3+8+2+3+8+2+3+2+5 = 44.$$

Por lo tanto, las licenciaturas con mayor frecuencia absoluta en el grupo CSI02 de Probabilidad y Estadística del trimestre 2018-O son Física con $f_F = 8$ y Metalúrgica $f_T = 8$. □

Definición 31 (frecuencia relativa). La frecuencia relativa del evento E es el cociente de la frecuencia absoluta del evento E y las n realizaciones del experimento aleatorio y es denotada por fr_E, entonces:

$$fr_E = \frac{f_E}{n}. \tag{2.4}$$

Si $\Omega = \bigcup\limits_{j=1}^{m} E_j$ donde $\forall i \neq j$ se tiene que $E_j \bigcap E_i = \emptyset$, entonces $\sum\limits_{j=1}^{m} fr_j = 1$.

Definición 32 (frecuencia acumulada). La frecuencia acumulada del evento E es la suma de la frecuencia absoluta del evento E hasta el k–ésimo término de las n realizaciones del experimento aleatorio:

$$F_k = \sum_{j=1}^{k} f_j, \tag{2.5}$$

donde F_k es la frecuencia acumulada del evento E hasta el k–ésimo término y f_j es la frecuencia absoluta del evento E en la j–ésima realización del experimento aleatorio.

Definición 33 (frecuencia relativa acumulada). La frecuencia relativa acumulada del evento E considerando hasta el k–ésimo término de las n realizaciones del experimento aleatorio es el cociente entre la frecuencia acumulada del evento E y las n realizaciones del experimento aleatorio y es denotada por Fr_k, entonces:

$$Fr_k = \frac{F_k}{n}. \tag{2.6}$$

Si $\Omega = \bigcup\limits_{j=1}^{m} E_j$ donde $\forall i \neq j$ se tiene que $E_j \bigcap E_i = \emptyset$, entonces $\sum\limits_{j=1}^{m} Fr_j = 1$.

Definición 34 (probabilidad frecuentista). Sea F_E el número de ocurrencias acumuladas del evento E en las n realizaciones del experimento aleatorio, entonces:

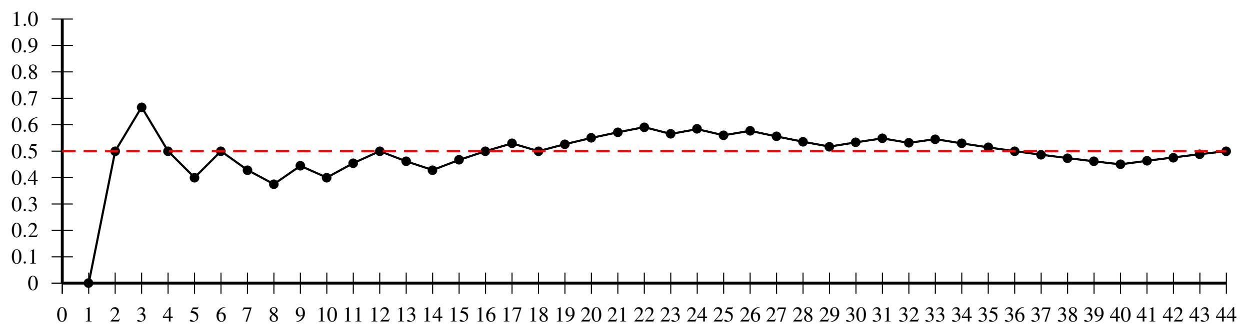
$$\mathscr{P}(E) = \lim_{n \to \infty} \frac{F_E}{n}, \tag{2.7}$$

donde $\mathscr{P}(E)$ es la probabilidad frecuentista del evento E, F_E es la frecuencia acumulada del evento E y n es el número de realizaciones del experimento aleatorio.

La definición 34 (probabilidad frecuentista) presenta una aproximación intuitiva de la probabilidad frecuentista del evento E, es decir, $\mathscr{P}(E) \approx \frac{F_E}{n}$ o equivalentemente $\mathscr{P}(E) \approx Fr_E$. El concepto de probabilidad frecuentista fue presentado por Bernoulli (1713).

Ejemplo 28 (probabilidad frecuentista del evento X). Utilizando la información de la tabla 2.1, ejemplo 27, calcular la probabilidad del evento $X = \{x \in X \mid x$ es alumno de Ingeniería Civil, Física o Metalúrgica$\}$ para cada pregunta y presentar la gráfica del experimento aleatorio para la probabilidad del evento X_k.

Solución. Calculando la frecuencia relativa acumulada del evento X se obtiene la gráfica 2.5.



Gráfica 2.5: Probabilidad frecuentista del evento X.

La gráfica 2.5 muestra que la probabilidad frecuentista del evento X se aproxima a $\frac{1}{2}$. $\square$

La definición 34 (probabilidad frecuentista) satisface las propiedades siguientes:

1. $\mathscr{P}(\Omega) = 1$.
2. $\mathscr{P}(A) \geq 0$ y $\mathscr{P}(B) \geq 0$.
3. $\mathscr{P}(A \bigcup B) = \mathscr{P}(A) + \mathscr{P}(B)$ cuando $\mathscr{P}(A \bigcap B) = 0$.

2.4. Probabilidad subjetiva

Definición 35 (probabilidad subjetiva). La probabilidad subjetiva del evento E depende del conocimiento que el observador tiene acerca del experimento aleatorio $\mathscr{E}$ (interpretación propuesta por Ramsey 1931 y formalizada por De Finetti 1937).

2.5. σ–álgebra

Una σ–álgebra es un conjunto que agrupa a todos los eventos de un experimento aleatorio $\mathscr{E}$ para los que son definidas o calculadas probabilidades y que es cerrado bajo complementos y uniones numerables.

Definición 36 (σ–álgebra). Sea $\mathscr{F} \neq \emptyset$ una colección de subconjuntos de Ω que satisface:

1. $\Omega \in \mathscr{F}$.
2. Si $A \in \mathscr{F} \Rightarrow A^{\mathrm{c}} \in \mathscr{F}$.
3. Si $A_{k \in \mathbb{N}} \in \mathscr{F} \Rightarrow \bigcup_{k=1}^{\infty} A_k \in \mathscr{F}$.

La primera condición indica que el espacio muestra pertenece a la colección $\mathscr{F}$, la segunda condición asegura que si A es un evento, entonces el complemento de este conjunto también es un evento y la tercera condición indica que si existe una colección infinita numerable de eventos, entonces la unión de todos estos también es un evento. Cada experimento aleatorio es asociado con una pareja $(\Omega, \mathscr{F})$ que está compuesta por el espacio muestra y una σ–álgebra de eventos.

Ejemplo 29 (σ–álgebra). Sea $\Omega \neq \emptyset$, entonces los conjuntos $\mathscr{F}_k$ para $k = 1, 2, 3$ son subconjuntos del espacio muestra Ω, donde cada $\mathscr{F}_k$ es una σ–álgebra.

Solución. Las σ–álgebras son:

1. $\mathscr{F}_1 = \{\emptyset, \Omega\}$, la σ–álgebra mínima.
2. $\mathscr{F}_2 = \{\emptyset, \{A\}, \{A^{\mathrm{c}}\}, \Omega\}$ donde $A^{\mathrm{c}} \in \Omega$.
3. $\mathscr{F}_3 = 2^{\Omega}$, el conjunto potencia es la σ–álgebra máxima.

Las $\mathscr{F}_k$ satisfacen la definición 36 (σ–álgebra), por lo tanto cada una del las $\mathscr{F}_k$ es una σ–álgebra. □

Ejemplo 30 (σ–álgebra). Sea $\Omega = \{\heartsuit, \diamondsuit, \spadesuit\}$ y sea $E = \{e \in \Omega \mid e = \diamondsuit\}$, encontrar la σ–álgebra mínima, la σ–álgebra para el evento E y la σ–álgebra máxima.

Solución. La σ–álgebra mínima es $\mathscr{F}_1 = \{\emptyset, \Omega\}$ y satisface la definición 36 (σ–álgebra), la σ–álgebra para el evento E es $\mathscr{F}_2 = \{\emptyset, \{\heartsuit\}, \{\diamondsuit\}, \{\spadesuit\}, \Omega\}$ y satisface la definición 36 (σ–álgebra) porque $E^{\mathrm{c}} \in \mathscr{F}_2$ y la σ–álgebra máxima es el conjunto $\mathscr{F}_3 = 2^{\Omega} = \{\emptyset, \{\heartsuit\}, \{\diamondsuit\}, \{\spadesuit\}, \{\heartsuit, \diamondsuit\}, \{\heartsuit, \spadesuit\}, \{\diamondsuit, \spadesuit\}, \Omega\}$ y satisface la definición 36 (σ–álgebra) porque la realización infinita del experimento aleatorio $\mathscr{E}$ es un elemento del conjunto potencia 2^{Ω}. □

La definición 36 (σ–álgebra) indica que $\mathscr{P}(\Omega) = 1$, si $\exists \mathscr{P}(E)$, entonces $\exists \mathscr{P}(E^{\mathrm{c}})$ y $\mathscr{P}(E^{\mathrm{c}}) = 1 - \mathscr{P}(E)$ y si $\exists \mathscr{P}(E_{k \in \mathbb{N}})$, entonces $\exists \mathscr{P}(E_1), \mathscr{P}(E_2), \ldots$

Una σ–álgebra agrupa a todos los subconjuntos de interés de Ω para calcular probabilidades y tal colección es el dominio sobre el que está definida una medida de probabilidad.

Definición 37 (conjunto medible). Un conjunto medible es un conjunto $E \in \mathscr{F}$ tal que E es un evento del experimento aleatorio $\mathscr{E}$.

Definición 38 (espacio medible). Un espacio medible es la pareja ordenada $(\Omega, \mathscr{F})$, donde Ω es un espacio muestra y $\mathscr{F}$ es una σ–álgebra.

2.6. Probabilidad axiomática

El estudio de la teoría de probabilidad fue formalizado cuando Kolmogorov (1933a) formuló un modelo axiomático que prevalece y que es el inicio de la teoría de la medida y de los procesos estocásticos.

Definición 39 (medida de probabilidad). Sea $(\Omega, \mathscr{F})$ un espacio medible y sea $\mathscr{P} : \mathscr{F} \to [0,1]$ una función que satisface los postulados siguientes:

$$\mathscr{P}(\Omega) = 1. \tag{Axioma 1}$$

$$\mathscr{P}(E) \geq 0 \quad \forall E \in \mathscr{F}. \tag{Axioma 2}$$

Si $\forall E_{k\in\mathbb{N}} \in \mathscr{F}$ y $\forall i \neq j$ se tiene que $E_i \bigcap E_j = \emptyset$, entonces:

$$\mathscr{P}\left(\bigcup_{k=1}^{\infty} E_k\right) = \sum_{k=1}^{\infty} \mathscr{P}(E_k). \tag{Axioma 3}$$

Las definiciones de probabilidad (clásica, geométrica y frecuentista) satisfacen los axiomas de la definición 39 (medida de probabilidad).

Definición 40 (espacio de probabilidad). Un espacio de probabilidad es la terna ordenada $(\Omega, \mathscr{F}, \mathscr{P})$, donde Ω es un espacio muestra, $\mathscr{F}$ es una σ–álgebra y $\mathscr{P}$ es una medida de probabilidad.

2.6.1. Propiedades de la probabilidad

Los tres axiomas de la medida de probabilidad son el fundamento para demostrar las propiedades siguientes:

Proposición 2 (probabilidad del evento imposible).

$$\mathscr{P}(\emptyset) = 0. \tag{2.8}$$

Proposición 3 (principio de aditividad con eventos ajenos). Sea E_k una colección finita de eventos mutuamente excluyentes dos a dos, es decir, $E_i \bigcap E_j = \emptyset$ donde $n = 1, 2, \ldots, n$, entonces:

$$\mathscr{P}\left(\bigcup_{k=1}^{n} E_k\right) = \sum_{k=1}^{n} \mathscr{P}(E_k). \tag{2.9}$$

Nota 5 (principio de aditividad con eventos ajenos). Sean A y B eventos tales que $A \bigcap B = \emptyset$, entonces:

$$\mathscr{P}\left(A \bigcup B\right) = \mathscr{P}(A) + \mathscr{P}(B). \tag{2.10}$$

Ejemplo 31 (principio de aditividad con eventos ajenos). Los eventos A y B son mutuamente excluyentes, es decir, $A \cap B = \emptyset$ y presentan las probabilidades $\mathscr{P}(B) = 0.4$ y $\mathscr{P}(A \cap B^{c}) = 0.3$. Calcular $\mathscr{P}(A \cup B)$.

Solución. Por hipótesis $A \cap B = \emptyset$, entonces por la proposición 3 (principio de aditividad con eventos ajenos) es aplicada la ecuación (2.10) y $\mathscr{P}(A) = \mathscr{P}(A \cap B^{c})$, por lo tanto $\mathscr{P}(A \cup B) = 0.4 + 0.3 = 0.7$. □

Proposición 4 (probabilidad complementaria). Sea E un evento, entonces:

$$\mathscr{P}(E^{c}) = 1 - \mathscr{P}(E). \tag{2.11}$$

Ejemplo 32 (probabilidad complementaria). El evento E presenta la probabilidad $\mathscr{P}(E) = \frac{5}{22}$. Calcular la probabilidad complementaria del evento E, es decir, $\mathscr{P}(E^{c})$.

Solución. Por la proposición 4 (probabilidad complementaria) y aplicando la ecuación (2.11), se tiene el resultado siguiente: $\mathscr{P}(E^{c}) = 1 - \frac{5}{22} = \frac{17}{22} = 0.7\overline{72} \approx 77.27\ \%$. $\square$

Proposición 5 (probabilidad de la contención). Sean $A \subseteq B$ eventos, entonces:

$$\mathscr{P}(A) \leq \mathscr{P}(B). \tag{2.12}$$

Proposición 6 (probabilidad de la diferencia). Sean $A \subseteq B$ eventos, entonces:

$$\mathscr{P}(B \setminus A) = \mathscr{P}(B) - \mathscr{P}(A). \tag{2.13}$$

Ejemplo 33 (probabilidad de la diferencia). Los eventos $D \subset G$ presentan las probabilidades $\mathscr{P}(D^{c}) = \frac{21}{22}$ y $\mathscr{P}(G) = 1$. Calcular $\mathscr{P}(G \setminus D)$.

Solución. Por la proposición 4 (probabilidad complementaria) y aplicando la ecuación (2.11), se tiene el resultado siguiente: $\mathscr{P}(D) = 1 - \frac{21}{22} = \frac{1}{22}$; por la proposición 6 (probabilidad de la diferencia) y aplicando la ecuación (2.13), $\mathscr{P}(G \setminus D) = \mathscr{P}(G) - \mathscr{P}(D) = 1 - \frac{1}{22} = \frac{21}{22} = 0.95\overline{45} \approx 95.45\ \%$. $\square$

Proposición 7 (probabilidad de un evento). Sea E un evento, entonces:

$$0 \leq \mathscr{P}(E) \leq 1. \tag{2.14}$$

Proposición 8 (principio de aditividad). Sean A y B eventos, entonces:

$$\mathscr{P}\left(A \bigcup B\right) = \mathscr{P}(A) + \mathscr{P}(B) - \mathscr{P}\left(A \bigcap B\right). \tag{2.15}$$

Nota 6 (principio de aditividad). Sean A, B y C, entonces:

$$\mathscr{P}(A \cup B \cup C) = \mathscr{P}(A) + \mathscr{P}(B) + \mathscr{P}(C) - \mathscr{P}(A \cap B) - \mathscr{P}(A \cap C) - \mathscr{P}(B \cap C) + \mathscr{P}(A \cap B \cap C). \tag{2.16}$$

Ejemplo 34 (principio de aditividad). Los eventos A y B presentan las probabilidades $\mathscr{P}(A) = 0.40$, $\mathscr{P}(B) = 0.50$ y $\mathscr{P}(A \bigcap B) = 0.20$. Calcular la probabilidad de que:

1. Ocurra exactamente uno de los dos eventos A o B.
2. No ocurra ninguno de los dos eventos A o B.

Solución. La probabilidad de que sólo ocurra A es $\mathscr{P}(A \setminus (A \cap B))$, la probabilidad de que sólo ocurra B es $\mathscr{P}(B \setminus (A \cap B))$ y la probabilidad de que no ocurran A y B es $\mathscr{P}(A^{c} \cap B^{c})$, entonces:

1. $\mathscr{P}(A \setminus (A \cap B)) + \mathscr{P}(B \setminus (A \cap B)) = (0.4 - 0.2) + (0.5 - 0.2) = 0.5$.
2. $\mathscr{P}(A^{c} \cap B^{c}) = \mathscr{P}((A \cup B)^{c}) = 1 - \mathscr{P}(A \cup B) = 1 - (\mathscr{P}(A) + \mathscr{P}(B) - \mathscr{P}(A \cap B)) = 1 - 0.7 = 0.3$.

Por lo tanto, sustituyendo la ecuación (2.13) de la proposición 6 (probabilidad de la diferencia) y la ecuación (2.10) de la proposición 3 (principio de aditividad con eventos ajenos) es obtenida la probabilidad de eventos mutuamente excluyentes $\mathscr{P}(A \setminus (A \cap B)) \cup \mathscr{P}(B \setminus (A \cap B))$, entonces aplicando la proposición 32 (primera ley de De Morgan) es obtenida $\mathscr{P}(A^{c} \cap B^{c})$. $\square$

Ejemplo 35 (principio de multiplicación). La primera urna contiene tres corazones y cuatro espadas, la segunda urna contiene siete corazones y seis espadas y la tercera urna contiene cuatro corazones y cinco espadas. El experimento aleatorio $\mathscr{E}$ consiste de dos partes:

1. Seleccionar una urna al azar.
2. Seleccionar al azar un elemento de la urna seleccionada en la parte inicial.

Calcular la probabilidad de seleccionar un corazón.

Solución. $U_1 = \{\heartsuit_1, \heartsuit_2, \heartsuit_3, \spadesuit_1, \spadesuit_2, \spadesuit_3, \spadesuit_4\}$, $U_2 = \{\heartsuit_1, \heartsuit_2, \heartsuit_3, \heartsuit_4, \heartsuit_5, \heartsuit_6, \heartsuit_7, \spadesuit_1, \spadesuit_2, \spadesuit_3, \spadesuit_4, \spadesuit_5, \spadesuit_6\}$ y $U_3 = \{\heartsuit_1, \heartsuit_2, \heartsuit_3, \heartsuit_4, \spadesuit_1, \spadesuit_2, \spadesuit_3, \spadesuit_4, \spadesuit_5\}$, entonces $\#U_1 = 7, \#U_2 = 13, \#U_3 = 9$ y $\#\Omega = 29$. Por hipótesis son tres urnas, entonces la probabilidad de seleccionar una de las tres urnas es $\mathscr{P}(U_k) = \frac{1}{3}$ para $k = 1, 2, 3$, entonces por la la ecuación (1.5) de la propiedad 38 (principio de multiplicación) y la ecuación (2.9), la probabilidad de seleccionar un corazón es $\mathscr{P}(\heartsuit) = \frac{1}{3}\left(\frac{3}{7} + \frac{7}{13} + \frac{4}{9}\right) = \frac{1{,}156}{2{,}457} = 0.\overline{470492} \approx 47.05\ \%$[2]. $\square$

2.7. Análisis combinatorio

El experimento aleatorio $\mathscr{E}$ en un espacio muestra Ω que es equiprobable, donde $\#\Omega = n$, entonces $\forall \omega_{k\in\mathbb{N}} \in \Omega$ tienen probabilidades equivalentes, es decir, $\mathscr{P}(\omega_k) = \frac{1}{n}$, por lo tanto la probabilidad del evento E es:

$$\mathscr{P}(E) = \frac{\#E}{\#\Omega}, \tag{2.17}$$

donde $\mathscr{P}(E)$ es la probabilidad del evento E, $\#E$ es la cardinalidad del evento E y $\#\Omega$ es la cardinalidad del espacio muestra (análisis realizado por Pascal 1654, Huygens 1657a y Bernoulli 1713).

2.7.1. Muestras aleatorias ordenadas sin reemplazo

Definición 41 (permutaciones de n en k elementos). Las permutaciones de n en k elementos son el número de arreglos ordenados de k elementos tomados sin reemplazo que son formados a partir de n elementos donde $0 \leq k < n \in \mathbb{N}$, entonces:

$$P(n,k) = \prod_{k=0}^{k-1}(n-k) = n\cdot(n-1)\cdots(n-k+1) = \frac{n!}{(n-k)!}, \tag{2.18}$$

donde $P(n,k)$ son las permutaciones de n en k elementos, n es el número de elementos del espacio muestra, k es el número de elementos de los arreglos y $\forall n \in \mathbb{N}$ es definido el factorial[3] de n como $n!$.

[2] Los denominadores son primos relativos entre si, es decir, $(7,9,13) = 1$, entonces el mínimo común múltiplo de $7,9$ y 13 es $\operatorname{mcm}(7,9,13) = 7\cdot 3^2\cdot 13 = 819$, por lo tanto el denominador es 2,457.

[3] El factorial de n es definido como $n! = \prod_{k=0}^{n-1}(n-k) = n(n-1)\cdots(2)(1) = \prod_{k=1}^{n} k$, donde el factorial de cero es $0! = 1$.

Ejemplo 36 (permutaciones de n en k elementos). Considerando cuatro cartas en $\Omega = \{\heartsuit, \diamondsuit, \clubsuit, \spadesuit\}$. Calcular de cuántas formas diferentes son repartidas dos cartas (ordenadas y sin repetición).

Solución. Al repartir la primera carta existen cuatro posibilidades y al repartir al segunda existen sólo tres posibilidades porque la primera carta no es reemplazada, entonces doce arreglos ordenados de dos elementos son formados a partir de las cuatro cartas. Sustituyendo en la ecuación (2.18):

$$P(4,2) = \frac{4!}{(4-2)!} = \frac{4!}{2!} = \frac{4\cdot 3\cdot 2\cdot 1}{2\cdot 1} = 4\cdot 3 = 12.$$

$$P = \{\{\heartsuit,\diamondsuit\},\{\heartsuit,\clubsuit\},\{\heartsuit,\spadesuit\},\{\diamondsuit,\heartsuit\},\{\diamondsuit,\clubsuit\},\{\diamondsuit,\spadesuit\},\{\clubsuit,\heartsuit\},\{\clubsuit,\diamondsuit\},\{\clubsuit,\spadesuit\},\{\spadesuit,\heartsuit\},\{\spadesuit,\diamondsuit\},\{\spadesuit,\clubsuit\}\}.$$

Por lo tanto, los doce arreglos de dos cartas están en el evento P. □

2.7.2. Muestras aleatorias exhaustivas ordenadas sin reemplazo

Definición 42 (permutaciones de n elementos). Las permutaciones de n elementos son el número de arreglos ordenados de n elementos sin reemplazo que son formados a partir de n elementos donde $n \in \mathbb{N}$, entonces:

$$P(n) = \prod_{k=0}^{n-1}(n-k) = n\cdot(n-1)\cdot\dots\cdot 2\cdot 1 = n!, \tag{2.19}$$

donde $P(n)$ son las permutaciones de n elementos, n es el número de elementos del espacio muestra y $\forall n \in \mathbb{N}$ está definido $n!$.

Ejemplo 37 (permutaciones de n elementos). Considerando tres cartas en $\Omega = \{\heartsuit, \diamondsuit, \clubsuit\}$. Calcular de cuántas formas diferentes son repartidas las tres cartas (exhaustivamente, ordenadas y sin repetición).

Solución. Sustituyendo en la ecuación (2.19):

$$P(3) = 3! = 6.$$
$$D = \{\{\heartsuit,\diamondsuit,\clubsuit\},\{\heartsuit,\clubsuit,\diamondsuit\},\{\diamondsuit,\clubsuit,\heartsuit\},\{\diamondsuit,\heartsuit,\clubsuit\},\{\clubsuit,\diamondsuit,\heartsuit\},\{\clubsuit,\heartsuit,\diamondsuit\}\}.$$

Por lo tanto, las seis formas diferentes de repartir tres cartas están en el evento D. □

2.7.3. Muestras aleatorias ordenadas con reemplazo

Definición 43 (ordenaciones con repetición de n en k elementos). Las ordenaciones con repetición de n en k elementos son el número de arreglos ordenados de k elementos con reemplazo que son formados a partir de n elementos donde $k \in \mathbb{N}$, entonces:

$$O(n,k) = n^k, \tag{2.20}$$

donde $O(n,k)$ son las ordenaciones con repetición de n en k elementos, n es el número de elementos del espacio muestra y k es el número de elementos de los arreglos (k puede ser menor, igual o mayor que n).

Ejemplo 38 (ordenaciones con repetición de n en k elementos). Si un byte está compuesto por ocho bits y cada bit tiene dos estados, entonces un byte es un arreglo ordenado de bits. Calcular el número de caracteres que es posible representar con un byte.

Solución. El espacio muestra es $\Omega = \{0,1\}, \#\Omega = 2$ y cada byte tiene ocho bits, entonces sustituyendo en la ecuación (2.20):

$$O(2,8) = 2^8 = 256.$$

Por lo tanto, con un byte de ocho bits es posible representar 256 caracteres. □

2.7.4. Muestras aleatorias no ordenadas sin reemplazo

Definición 44 (combinaciones de n en k elementos). Las combinaciones de n en k elementos son el número de arreglos no ordenados de k elementos sin reemplazo que son formados con n elementos donde $0 \leq k \leq n \in \mathbb{N}$, entonces:

$$C(n,k) = \binom{n}{k} = \frac{n!}{k!\,(n-k)!}, \tag{2.21}$$

donde $C(n,k)$ son las combinaciones no ordenadas sin repetición de n en k elementos, n es el número de elementos del espacio muestra y k es el número de elementos de los arreglos.

Ejemplo 39 (combinaciones de n en k elementos). Hay cuatro jugadores de tenis. Calcular cuántos equipos de dobles (dos jugadores) diferentes existen.

Solución. El espacio muestra es $\Omega = \{1,2,3,4\}, \#\Omega = 4$ y cada equipo cuenta con dos jugadores, entonces sustituyendo en la ecuación (2.21):

$$C(4,2) = \binom{4}{2} = \frac{4!}{2!\,(4-2)!} = \frac{24}{4} = 6.$$
$$E = \{\{1,2\},\{1,3\},\{1,4\},\{2,3\},\{2,4\},\{3,4\}\}.$$

Por lo tanto, con cuatro jugadores existen seis equipos de dobles en el evento E. □

2.7.5. Coeficiente multinomial

Definición 45 (coeficiente multinomial). El coeficiente multinomial de n elementos en m grupos es el número de arreglos no ordenados de m grupos tomados sin reemplazo que son formados a partir de n elementos donde $n = \sum_{j=1}^{m} k_j$ y los grupos $1 \leq j \leq n \in \mathbb{N}$, entonces:

$$M(n,k_j) = \binom{n}{k_1\; k_2 \cdots\; k_m} = \frac{n!}{\prod_{j=1}^{m} k_j!} = \frac{n!}{k_1!k_2!\cdots k_m!}, \tag{2.22}$$

donde $M(n,k_j)$ es el número de combinaciones no ordenadas sin repetición de n elementos en m grupos, n es el número de elementos del espacio muestra, m es el número de grupos en los que está dividido el espacio muestra y k_j es el número de elementos del j–ésimo grupo.

Ejemplo 40 (coeficiente multinomial). Una organización cuenta con treinta miembros y tres comités, los dos primeros comités están formados por doce miembros cada uno y el tercer comité está formado por seis miembros. Calcular el número de comités diferentes que existen.

Solución. Sustituyendo en la ecuación (2.22):

$$M(30, k_j) = \binom{30}{12\ 12\ 6} = \frac{30!}{12!12!6!} = 1'605,660,228,900.$$

Por lo tanto, existen un billón seiscientos cinco mil seiscientos sesenta millones doscientos veintiocho mil novecientos comités distintos. □

2.7.6. Aplicación del análisis combinatorio

Suponiendo que cinco cartas de una baraja con trece números $N = \{2,3,4,5,6,7,8,9,10,J,Q,K,A\}$ y cuatro símbolos $S = \{\heartsuit, \diamondsuit, \clubsuit, \spadesuit\}$ son repartidas, entonces es posible describir el espacio muestra con los eventos posibles y calcular la probabilidad de cada uno de los eventos. La baraja está representada en la tabla 2.2.

Tabla 2.2: Baraja con trece números y cuatro símbolos.

$2\heartsuit$	$3\heartsuit$	$4\heartsuit$	$5\heartsuit$	$6\heartsuit$	$7\heartsuit$	$8\heartsuit$	$9\heartsuit$	$10\heartsuit$	$J\heartsuit$	$Q\heartsuit$	$K\heartsuit$	$A\heartsuit$
$2\diamondsuit$	$3\diamondsuit$	$4\diamondsuit$	$5\diamondsuit$	$6\diamondsuit$	$7\diamondsuit$	$8\diamondsuit$	$9\diamondsuit$	$10\diamondsuit$	$J\diamondsuit$	$Q\diamondsuit$	$K\diamondsuit$	$A\diamondsuit$
$2\clubsuit$	$3\clubsuit$	$4\clubsuit$	$5\clubsuit$	$6\clubsuit$	$7\clubsuit$	$8\clubsuit$	$9\clubsuit$	$10\clubsuit$	$J\clubsuit$	$Q\clubsuit$	$K\clubsuit$	$A\clubsuit$
$2\spadesuit$	$3\spadesuit$	$4\spadesuit$	$5\spadesuit$	$6\spadesuit$	$7\spadesuit$	$8\spadesuit$	$9\spadesuit$	$10\spadesuit$	$J\spadesuit$	$Q\spadesuit$	$K\spadesuit$	$A\spadesuit$

La tabla 2.2 presenta las 52 cartas de la baraja y si son repartidas cinco cartas es posible obtener alguno de los eventos siguientes:

1. Quimera (cinco cartas con números diferentes que no son consecutivos y con símbolos diferentes).
2. Un par (dos cartas con el mismo número y las otras tres con números diferentes).
3. Dos pares (dos cartas con el mismo número, otras dos con otro número y otra diferente a las anteriores).
4. Tercia (tres cartas con el mismo número y las otras dos con números diferentes).
5. Escalera (cinco cartas con números consecutivos y símbolos diferentes).
6. Figura (cinco cartas con el mismo símbolo y números que no son consecutivos).
7. Completo (una tercia y un par).
8. Póker (cuatro cartas con el mismo número y la otra con un número diferente).
9. Botón (cinco cartas con números consecutivos del A al K y con el mismo símbolo).
10. Imperial (cinco cartas con números consecutivos del 10 a A y con el mismo símbolo).

Por ejemplo: Quimera $Q = \{2\heartsuit, 4\diamondsuit, 5\clubsuit, 6\spadesuit, 7\heartsuit\}$, es decir, un par, dos pares o tercia de un símbolo no son eventos considerados en el juego. Un par $U = \{3\clubsuit, 3\spadesuit, 2\heartsuit, 5\diamondsuit, A\spadesuit\}$. Dos pares $D = \{3\clubsuit, 3\spadesuit, 5\heartsuit, 5\diamondsuit, A\spadesuit\}$. Tercia $T = \{3\heartsuit, 3\clubsuit, 3\spadesuit, 5\diamondsuit, A\spadesuit\}$. Escalera $E = \{3\heartsuit, 4\clubsuit, 5\spadesuit, 6\diamondsuit, 7\spadesuit\}$. Figura $F = \{3\heartsuit, 5\heartsuit, 6\heartsuit, 10\heartsuit, Q\heartsuit\}$. Completo $C = \{9\heartsuit, 9\diamondsuit, 9\spadesuit, 3\heartsuit, 3\clubsuit\}$. Póker $P = \{10\heartsuit, 10\diamondsuit, 10\clubsuit, 10\spadesuit, 5\heartsuit\}$. Sin considerar comodines no existen quintillas. Botón $B = \{4\spadesuit, 5\spadesuit, 6\spadesuit, 7\spadesuit, 8\spadesuit\}$. Imperial $I = \{10\diamondsuit, J\diamondsuit, Q\diamondsuit, K\diamondsuit, A\diamondsuit\}$.

El juego es ganado por el jugador que recibe el evento con menor probabilidad y en caso de que varios jugadores tengan el mismo evento, el criterio de desempate es el evento con el mayor valor en función de los números, donde $2 < 3 < \ldots < 10 < J < Q < K < A$.

Ejemplo 41 (póker). Suponiendo que son repartidas cinco cartas de una baraja con trece números y cuatro símbolos. Calcular:

1. La cantidad de manos que es posible repartir, es decir, el número de casos totales.
2. La cantidad de formas posibles de repartir cada uno de los eventos $Q, U, D, T, E, F, C, P, B, I$.
3. La probabilidad de repartir cada uno de los eventos $Q, U, D, T, E, F, C, P, B, I$.

Solución. Los cálculos son realizados suponiendo que son repartidas cinco cartas de las 52 de la baraja.

1. La cantidad de manos que es posible repartir es el número de casos totales o la cardinalidad del espacio muestra, es decir, #Ω:
$$\#\Omega = C(52,5) = 2'598,960.$$
Entonces existen 2’598,960 formas de repartir cinco cartas de una baraja con 52 cartas.
2. La cantidad de formas posibles de repartir cada uno de los eventos $Q, U, D, T, E, F, C, P, B, I$.

   La cantidad de formas posibles de repartir una quimera es obtenida con cinco números diferentes $C(13,5)$ y cuatro símbolos para cada uno de los cinco números $C^5(4,1)$ menos las escaleras (cinco números consecutivos y símbolos diferentes) $E = C(10,1)C^5(4,1) - (B+I)$, las figuras (cinco cartas con el mismo símbolo y números no consecutivos) $F = C(13,5)C(4,1) - (B+I)$, los botones (cinco números consecutivos del A al K y con el mismo símbolo) $B = C(9,1)C(4,1)$ y los imperiales (cinco números consecutivos del 10 al A y con el mismo símbolo) $I = C(4,1)$:
$$\#Q = C(13,5)C^5(4,1) - \big(C(10,1)C^5(4,1) + (C(13,5) - C(9,1) - 1)C(4,1)\big) = 1'302,540.$$
   Es decir, existen 1’302,540 formas posibles de repartir una quimera.

   La cantidad de formas posibles de repartir un par es obtenida con uno de los trece números $C(13,1)$ y dos de los cuatro símbolos $C(4,2)$, de los doce números restantes son seleccionados tres $C(12,3)$ y cuatro símbolos para cada uno de los tres números $C^3(4,1)$, entonces:
$$\#U = C(13,1)C(4,2)C(12,3)C^3(4,1) = 1'098,240.$$
   Es decir, existen 1’098,240 formas posibles de repartir un par.

   La cantidad de formas posibles de repartir dos pares es obtenida con dos de los trece números $C(13,2)$ y dos de los cuatro símbolos para cada uno de los dos número $C^2(4,2)$, de los once números restantes es elegido un número $C(11,1)$ y uno de los cuatro símbolos $C(4,1)$, entonces:
$$\#D = C(13,2)C^2(4,2)C(11,1)C(4,1) = 123,552.$$
   Es decir, existen 123,552 formas posibles de repartir dos pares.

   La cantidad de formas posibles de repartir una tercia es obtenida con uno de los trece números $C(13,1)$ y tres de los cuatro símbolos $C(4,3)$, de los doce números restantes son elegidos dos números $C(12,2)$ y uno de los cuatro símbolos para cada uno de los dos números $C^2(4,1)$, entonces:
$$\#T = C(13,1)C(4,3)C(12,2)C^2(4,1) = 54,912.$$

Es decir, existen 54,912 formas posibles de repartir una tercia.

La cantidad de formas posibles de repartir una escalara $E = \{\{A-5\}, \{2-6\}, \ldots, \{9-K\}, \{10-A\}\}$ es obtenida con las combinaciones $C(10,1)$ y con uno de los cuatro símbolos para cada uno de los cinco números $C^5(4,1)$ menos los botones $B = C(9,1)C(4,1)$ y los imperiales $I = C(4,1)$, entonces:

$$\#E = C(10,1)C^5(4,1) - (C(9,1)C(4,1) + C(4,1)) = 10,200.$$

Es decir, existen 10,200 formas posibles de repartir una tercia.

La cantidad de formas posibles de repartir una figura es obtenida con cinco de los trece números con el mismo símbolo $C(13,5)$ y los cuatro símbolos $C(4,1)$ menos los botones $B = C(9,1)C(4,1)$ y los imperiales $I = C(4,1)$, entonces:

$$\#F = C(13,5)C(4,1) - (C(9,1)C(4,1) + C(4,1)) = 5,108.$$

Es decir, existen 5,108 formas posibles de repartir una figura.

La cantidad de formas posibles de repartir un completo es obtenida con una tercia, es decir, con uno de los trece números $C(13,1)$ y tres de los cuatro símbolos $C(4,3)$ y un par, es decir, de los doce números restantes es elegido uno $C(12,1)$ y dos de los cuatro símbolos $C(4,2)$, entonces:

$$\#C = C(13,1)C(4,3)C(12,1)C(4,2) = 3,744.$$

Es decir, existen 3,744 formas posibles de repartir un completo.

La cantidad de formas posibles de repartir un póker es obtenida con un cuarteto, es decir, con uno de los trece números $C(13,1)$ y cuatro de los cuatro símbolos $C(4,4)$, de los doce números restantes es elegido uno $C(12,1)$ y uno de los cuatro símbolos $C(4,1)$, entonces:

$$\#P = C(13,1)C(4,4)C(12,1)C(4,1) = 624.$$

Es decir, existen 624 formas posibles de repartir un póker.

La cantidad de formas posibles de repartir un botón es obtenida con una escalera con mismo símbolo, es decir, $E = \{\{A-5\}, \{2-6\}, \ldots, \{8-Q\}, \{9-K\}\}$ con las combinaciones $C(9,1)$ y uno de los cuatro símbolos $C(4,1)$, entonces:

$$\#B = C(9,1)C(4,1) = 36.$$

Es decir, existen 36 formas posibles de repartir un botón.

La cantidad de formas posibles de repartir un imperial es obtenida con la escalera $E = \{10-A\}$ y uno de los cuatro símbolos $C(4,1)$, entonces:

$$\#I = C(4,1) = 4.$$

Es decir, existen 4 formas posibles de repartir un imperial.

3. La probabilidad de repartir cada uno de los eventos es presentada en la tabla 2.3.

Tabla 2.3: Probabilidad de repartir una mano de póker.

Evento	Cardinalidad	Probabilidad
Q	1,302,540	0.501177
U	1,098,240	0.422569
D	123,552	0.047539
T	54,912	0.021128
E	10,200	0.003925
F	5,108	0.001965
C	3,744	0.001441
P	624	0.000240
B	36	0.000014
I	4	0.000002
Ω	2,598,960	1.000000

La tabla 2.3 presenta los eventos posibles al repartir cinco cartas de una baraja con trece números y cuatro símbolos, es decir, con 52 cartas, la cardinalidad de cada uno de los eventos y la probabilidad de cada evento, donde $\Omega = Q \cup U \cup D \cup T \cup E \cup F \cup C \cup P \cup B \cup I$.

Si X es alguno de los eventos posibles, entonces es posible comprobar que la ocurrencia de alguno de los eventos satisface los axiomas de la definición 39 (medida de probabilidad):

$$\mathscr{P}(\Omega) = \mathscr{P}\left(\bigcup_{k=1}^{10} X_k\right) = \sum_{k=1}^{10} \mathscr{P}(X_k) = 1.$$

$$\mathscr{P}(X_k) \geq 0 \qquad \forall X_k \in \mathscr{F}.$$

$$\mathscr{P}(\Omega) = \mathscr{P}(Q) + \mathscr{P}(U) + \mathscr{P}(D) + \mathscr{P}(T) + \mathscr{P}(E) + \mathscr{P}(F) + \mathscr{P}(C) + \mathscr{P}(P) + \mathscr{P}(B) + \mathscr{P}(I) = 1.$$

Las soluciones representan la probabilidad de cada evento al repartir las cinco cartas y es posible extender el juego cambiando hasta cinco cartas para obtener una mejor mano, problema que es una aplicación de la probabilidad condicional. □

2.8. Probabilidad condicional

Si un experimento aleatorio $\mathscr{E}$ está compuesto por experimentos aleatorios que son realizados sucesivamente, entonces los resultados de los experimentos sucesivos son:

1. Independientes.
2. Dependientes.

Definición 46 (probabilidad condicional). Sean A y B eventos donde $\mathscr{P}(A) > 0$, entonces:

$$\mathscr{P}(B \mid A) = \frac{\mathscr{P}(A \cap B)}{\mathscr{P}(A)}, \tag{2.23}$$

donde $\mathscr{P}(B \mid A)$ es la probabilidad condicional de B dado A, $\mathscr{P}(A \cap B)$ es la probabilidad de $A \cap B$, es decir, la probabilidad de que A y B ocurran simultáneamente y $\mathscr{P}(A)$ es la probabilidad de A.

Ejemplo 42 (probabilidad condicional). La probabilidad condicional $\mathscr{P}(B \mid A)$ es una medida de probabilidad, entonces satisface los axiomas de la definición 39 (medida de probabilidad):

Solución. Es suficiente comprobar que:

1. $\mathscr{P}(\Omega \mid A) = \frac{\mathscr{P}(\Omega \cap A)}{\mathscr{P}(A)} = \frac{\mathscr{P}(A)}{\mathscr{P}(A)} = 1$.
2. $\mathscr{P}(B \mid A) \geq 0$.
3. Si $B_1 \cap B_2 = \emptyset$, entonces $\mathscr{P}(B_1 \cup B_2 \mid A) = \mathscr{P}(B_1 \mid A) + \mathscr{P}(B_2 \mid A)$. □

2.8.1. Propiedades de la probabilidad condicional

La probabilidad condicional $\mathscr{P}(B \mid A)$ satisface las propiedades siguientes:

Propiedad 39 (probabilidad condicional del evento imposible).

$$\mathscr{P}(\emptyset \mid A) = 0. \tag{2.24}$$

Propiedad 40 (probabilidad condicional complementaria).

$$\mathscr{P}(B \mid A) = 1 - \mathscr{P}(B^{\mathrm{c}} \mid A). \tag{2.25}$$

Propiedad 41 (principio de aditividad).

$$\mathscr{P}(B_1 \cup B_2 \mid A) = \mathscr{P}(B_1 \mid A) + \mathscr{P}(B_2 \mid A) - \mathscr{P}(B_1 \cap B_2 \mid A). \tag{2.26}$$

Ejemplo 43 (probabilidad condicional). Sea E un evento donde $\mathscr{P}(E) > 0$ y A y B dos eventos tales que, dado que ocurre E, la probabilidad de que ocurra A es $\frac{3}{8}$, es decir, $\mathscr{P}(A \mid E) = \frac{3}{8}$, la probabilidad de que ocurra B es $\frac{1}{2}$, es decir, $\mathscr{P}(B \mid E) = \frac{1}{2}$ y la probabilidad de que no ocurran A y B ($A^{\mathrm{c}} \cap B^{\mathrm{c}}$) es $\frac{1}{4}$, es decir, $\mathscr{P}(A^{\mathrm{c}} \cap B^{\mathrm{c}} \mid E) = \frac{1}{4}$. Calcular las probabilidades de ocurrencia de que:

1. Al menos uno de los eventos A o B dado que ocurre E, es decir, $\mathscr{P}(A \cup B \mid E)$.
2. Exactamente los dos eventos A y B dado que ocurre E, es decir, $\mathscr{P}(A \cap B \mid E)$.
3. Exactamente uno de los eventos A o B dado que ocurre E, es decir, $\mathscr{P}((A \cup B) \setminus (A \cap B) \mid E)$.

Solución. Aplicando la ecuación (2.23) de la definición 46 (probabilidad condicional):

1. Por la propiedad 40 (probabilidad condicional complementaria) $\mathscr{P}(A \cup B \mid E) = 1 - \mathscr{P}((A \cup B)^{\mathrm{c}} \mid E)$, entonces por la propiedad 32 (primera ley de De Morgan) $1 - \mathscr{P}((A \cup B)^{\mathrm{c}} \mid E) = 1 - \mathscr{P}(A^{\mathrm{c}} \cap B^{\mathrm{c}} \mid E)$, entonces $\mathscr{P}(A \cup B \mid E) = 1 - \frac{1}{4} = \frac{3}{4}$.
2. Por la propiedad 41 (principio de aditividad) $\mathscr{P}(A \cap B \mid E) = \mathscr{P}(A \mid E) + \mathscr{P}(B \mid E) - \mathscr{P}(A \cup B \mid E)$, entonces $\mathscr{P}(A \cap B \mid E) = \frac{3}{8} + \frac{1}{2} - \frac{3}{4} = \frac{1}{8}$.

3. $\mathscr{P}((A\cup B)\setminus(A\cap B)\mid E)=\mathscr{P}(A\setminus(A\cap B)\mid E)\cup\mathscr{P}(B\setminus(A\cap B)\mid E)$, entonces por la propiedad 41 (principio de aditividad) $\mathscr{P}((A\cup B)\setminus(A\cap B)\mid E)=\mathscr{P}(A\setminus(A\cap B)\mid E)+\mathscr{P}(B\setminus(A\cap B)\mid E)$ y entonces $\mathscr{P}((A\cup B)\setminus(A\cap B)\mid E)=\left(\frac{3}{8}-\frac{1}{8}\right)+\left(\frac{1}{2}-\frac{1}{8}\right)=\frac{5}{8}$. □

Ejemplo 44. Un profesor realiza dos exámenes, 90 % aprueba el primero y 70 % aprueba ambos exámenes. Calcular la probabilidad de seleccionar un alumno que aprueba ambos exámenes, es decir, $\mathscr{P}(B\mid A)$.

Solución. Sea A el evento aprobar el primer examen, B el evento aprobar el segundo examen y $B\cap A$ el evento aprobar ambos exámenes, entonces $\mathscr{P}(A)=0.9, \mathscr{P}(A\cap B)=0.7$. Aplicando la ecuación (2.23) de la definición 46 (probabilidad condicional):

$$\mathscr{P}(B\mid A)=\frac{\mathscr{P}(B\cap A)}{\mathscr{P}(A)}=\frac{0.7}{0.9}=0.\overline{7}\approx 77.78\ \%.$$

Por lo tanto, la probabilidad de seleccionar un alumno que aprueba el segundo examen dado que aprobó el primer examen es 77.78 %. □

Ejemplo 45 (probabilidad condicional de eventos independientes). Un experimento aleatorio $\mathscr{E}$ consiste en lanzar un dado, los eventos son $A=\{$⚂$\}$ y $B=\{$⚀,⚂,⚄$\}$. Calcular las probabilidades condicionales siguientes:

1. $\mathscr{P}(B\mid A)$.
2. $\mathscr{P}(A\mid B)$.

Solución. Aplicando la ecuación (2.23) de la definición 46 (probabilidad condicional), entonces son obtenidas $\mathscr{P}(B\cap A)=\mathscr{P}(A)\mathscr{P}(B\mid A)=\frac{\frac{1}{6}\left(\frac{3}{6}\right)\left(\frac{1}{6}\right)}{\frac{1}{6}}=\frac{1}{12}$ y $\mathscr{P}(A\cap B)=\mathscr{P}(B)\mathscr{P}(A\mid B)=\frac{\frac{3}{6}\left(\frac{1}{6}\right)\left(\frac{3}{6}\right)}{\frac{3}{6}}=\frac{1}{12}$.

1. La probabilidad de que caiga un número impar dado que cayó un tres es:

$$\mathscr{P}(B\mid A)=\frac{\mathscr{P}(B\cap A)}{\mathscr{P}(A)}=\frac{\frac{1}{12}}{\frac{1}{6}}=\frac{1}{2}.$$

2. La probabilidad de que caiga un tres dado que cayó un un número impar es:

$$\mathscr{P}(A\mid B)=\frac{\mathscr{P}(A\cap B)}{\mathscr{P}(B)}=\frac{\frac{1}{12}}{\frac{3}{6}}=\frac{1}{6}.$$

Por lo tanto, la probabilidad de que caiga un número impar dado que cayó un tres es un medio y la probabilidad de que caiga un tres dado que cayó un un número impar es un sexto. □

2.9. Independencia estocástica de eventos

La probabilidad condicional de un evento B dada la ocurrencia de un evento A es calculada en un espacio muestra reducido, es decir, para el espacio muestra conformado por los resultados posibles del experimento aleatorio en los que ocurre el evento A (propuesta por De Moivre 1710).

Definición 47 (independencia estocástica de dos eventos). Los eventos A y B son independientes $(A\perp B)$ si y sólo si:

$$\mathscr{P}(A\cap B)=\mathscr{P}(A)\mathscr{P}(B), \tag{2.27}$$

donde $\mathscr{P}(A\cap B)$ es la probabilidad de $A\cap B$, $\mathscr{P}(A)$ es la probabilidad de A y $\mathscr{P}(B)$ es la probabilidad de B.

Afirmación (independencia estocástica de dos eventos). Si A y B son eventos independientes $(A \perp B)$, entonces:

$$\begin{aligned} \mathscr{P}(B \mid A) &= \mathscr{P}(B) && \text{si } \mathscr{P}(A) > 0. \\ \mathscr{P}(A \mid B) &= \mathscr{P}(A) && \text{si } \mathscr{P}(B) > 0. \end{aligned}$$

Es decir, la ocurrencia del evento A no afecta a la probabilidad del evento B y, análogamente, la ocurrencia del evento B no afecta a la probabilidad del evento A.

Ejemplo 46 (independencia estocástica de dos eventos). Si A y B son eventos estocásticamente independientes, no significa que los eventos son mutuamente excluyentes.
Solución. Es necesario demostrar que si $\mathscr{P}(A \cap B) = \mathscr{P}(A)\mathscr{P}(B) \nRightarrow A \cap B = \emptyset$. Sea $A \neq \emptyset \Rightarrow A \cap \Omega = A$, entonces $\mathscr{P}(A \cap \Omega) = \mathscr{P}(A)\mathscr{P}(\Omega) = \mathscr{P}(A)$ y $A \cap \Omega = A \neq \emptyset$, es decir:

$$\mathscr{P}(A \cap \Omega) = \mathscr{P}(A)\mathscr{P}(\Omega) \nRightarrow A \cap \Omega = \emptyset.$$

Por lo tanto, la independencia estocástica de A y B no significa que los eventos son ajenos. □

Ejemplo 47 (independencia estocástica de dos eventos). Si A y B son eventos mutuamente excluyentes, no significa que A y B son estocásticamente independientes.
Solución. Es necesario demostrar que si $A \cap B = \emptyset \nRightarrow \mathscr{P}(A \cap B) = \mathscr{P}(A)\mathscr{P}(B)$. Sea $\mathscr{E}$ lanzar una moneda, el evento A es © y A^{c} es ®, entonces $A \cap A^{c} = \emptyset$ y $\mathscr{P}(A \cap A^{c}) = 0$, pero $\mathscr{P}(A)\mathscr{P}(A^{c}) = \left(\frac{1}{2}\right)^2$, es decir:

$$A \cap A^{c} = \emptyset \Rightarrow \mathscr{P}(A \cap A^{c}) = 0 \neq \left(\frac{1}{2}\right)^2 = \mathscr{P}(A)\mathscr{P}(A^{c}) = \mathscr{P}(A \cap A^{c}).$$

Por lo tanto, si A y B son ajenos no significa que los eventos son estocásticamente independientes. □

Ejemplo 48 (independencia estocástica de dos eventos). Las estadísticas indican que la probabilidad de que un estudiante apruebe Química es 65 %, la probabilidad de que apruebe Física es 70 % y la probabilidad de que apruebe ambas es 40 %. Sea Q el evento aprobó Química y F el evento aprobó física. Indicar si Q y F son eventos estocásticamente independientes.
Solución. Por hipótesis $\mathscr{P}(Q) = 0.65, \mathscr{P}(F) = 0.70$ y $\mathscr{P}(Q \cap F) = 0.40$, entonces aplicando la ecuación (2.27) de la definición 47 (independencia estocástica de dos eventos) se tiene que:

$$\mathscr{P}(Q \cap F) = 0.40 \neq 0.455 = 0.65\,(0.70) = \mathscr{P}(Q)\mathscr{P}(F).$$

Por lo tanto, los eventos Q y F no son estocásticamente independientes. □

Definición 48 (independencia estocástica mutua de eventos). El conjunto de eventos $A_{k \in \mathbb{N}}$ son mutuamente independientes si satisfacen:

$$\mathscr{P}\left(\bigcap_{k=1}^{2} A_k\right) = \prod_{k=1}^{2} \mathscr{P}(A_k). \tag{2.28}$$

$$\mathscr{P}\left(\bigcap_{k=1}^{3} A_k\right) = \prod_{k=1}^{3} \mathscr{P}(A_k). \tag{2.29}$$

$$\vdots$$

$$\mathscr{P}\left(\bigcap_{k=1}^{n} A_k\right) = \prod_{k=1}^{n} \mathscr{P}(A_k). \tag{2.30}$$

Existe una igualdad de la ecuación (2.28), si se satisface la igualdad, entonces los eventos son independientes dos a dos. Existen $C(n,3)$ igualdades de la ecuación (2.29), si se satisfacen todas las igualdades, entonces los eventos son independientes tres a tres. Existe una igualdad de la ecuación (2.30), si se satisface la igualdad, entonces los n eventos son independientes.

La independencia estocástica de n eventos es probada comprobando las $2^n - n - 1$ igualdades[4] generadas por las ecuaciones (2.28)–(2.30).

Es decir, si se satisface alguna de estas igualdades no significa que se satisface otra y los pares ordenados del conjunto $I = \left\{(n,i) \in \mathbb{N}^2 \mid n \text{ es el número de eventos e } i \text{ es el número de igualdades por satisfacer}\right\}$ tienen el comportamiento siguiente: $I = \{(2,1),(3,4),(4,11),(5,26),(6,57),(7,120),(8,247),(9,502)\}$, entonces para probar la independencia de 10 eventos es necesario comprobar 1,013 igualdades.

2.10. Probabilidad total

Definición 49 (partición finita). Sea Ω el espacio muestra de un experimento aleatorio $\mathscr{E}$, entonces la colección de eventos $\{E_1,\ldots,E_n\}$ es una partición finita de Ω si $\forall k = 1,\ldots,n$ son satisfechas las condiciones siguientes:

1. Si $E_k \neq \emptyset$, entonces $\mathscr{P}(E_k) > 0$.
2. $\forall i \neq j$, entonces $E_i \cap E_j = \emptyset$, es decir, son conjuntos mutuamente excluyentes (conjuntos disjuntos).
3. $\bigcup\limits_{k=1}^{n} E_k = \Omega$.

Teorema 2 (probabilidad total). Sea $\{E_1,\ldots,E_n\}$ una partición finita del espacio muestra Ω de tal forma que $\forall k = 1,\ldots,n$ se tiene que $\mathscr{P}(E_k) > 0$, entonces para todo evento B:

$$\mathscr{P}(B) = \sum_{k=1}^{n} \mathscr{P}(B \mid E_k)\,\mathscr{P}(E_k), \tag{2.31}$$

donde $\mathscr{P}(B)$ es la probabilidad del evento B, $\mathscr{P}(B \mid E_k)$ es la probabilidad condicional del evento B dado el evento E_k y $\mathscr{P}(E_k)$ es la probabilidad del evento A_k. □

[4] La suma de las combinaciones de n en k elementos $\sum\limits_{k=0}^{n} C(n,k) = 2^k$ y restando las combinaciones de n en k elementos para $k = 0,1$, es decir, restando $\sum\limits_{k=0}^{1} C(n,k) = 1+n$, entonces $\sum\limits_{k=2}^{n} C(n,k) = 2^n - n - 1$.

Ejemplo 49 (probabilidad total). Una población con 51 % de mujeres y 49 % de hombres tiene 44 % de hombres que realizan deporte y 31 % de mujeres que realizan deporte. Una persona es elegida al azar, calcular la probabilidad de que la persona elegida realice deporte.

Solución. Sea $E_1 = M$ el evento elegir una mujer, $E_2 = H$ el evento elegir un hombre y D el evento elegir un deportista, entonces por el teorema 2 (probabilidad total) y sustituyendo en la ecuación (2.31):

$$\begin{aligned}
\mathscr{P}(D) &= \sum_{k=1}^{2} \mathscr{P}(D \mid E_k)\, \mathscr{P}(E_k) = \mathscr{P}(D \mid E_1)\, \mathscr{P}(E_1) + \mathscr{P}(D \mid E_2)\, \mathscr{P}(E_2). \\
\mathscr{P}(D) &= \frac{31}{100}\left(\frac{51}{100}\right) + \frac{44}{100}\left(\frac{49}{100}\right) = 0.3737.
\end{aligned}$$

Por lo tanto, la probabilidad de que la persona elegida al azar realice deporte es $\mathscr{P}(D) = 37.37\ \%$. □

Ejemplo 50 (probabilidad total). La urna uno tiene dos bolas negras y tres bolas blancas y la urna dos tiene cuatro bolas negras y una bola blanca. Una bola es extraída de la urna uno y es transferida a la urna dos y después una bola es extraída de la urna dos. Calcular la probabilidad de que la bola extraída de la urna dos sea negra.

Solución. Sea E_1 el evento la bola transferida de la urna uno a la urna dos es negra, E_2 el evento la bola transferida de la urna uno a la urna dos es blanca y N el evento la bola extraída de la urna dos es negra, entonces por el teorema 2 (probabilidad total) y sustituyendo en la ecuación (2.31):

$$\begin{aligned}
\mathscr{P}(N) &= \sum_{k=1}^{2} \mathscr{P}(N \mid E_k)\, \mathscr{P}(E_k) = \mathscr{P}(N \mid E_1)\, \mathscr{P}(E_1) + \mathscr{P}(N \mid E_2)\, \mathscr{P}(E_2). \\
\mathscr{P}(N) &= \left(\frac{4+1}{4+1+1}\right)\left(\frac{2}{2+3}\right) + \left(\frac{4}{4+1+1}\right)\left(\frac{3}{2+3}\right) = \frac{11}{15} = 0.7\overline{3} \approx 73.33\ \%.
\end{aligned}$$

Por lo tanto, la probabilidad de que la bola extraída de la urna dos sea negra es $\mathscr{P}(N) = 73.33\ \%$. □

Definición 50 (partición). Sea Ω el espacio muestra de un experimento aleatorio $\mathscr{E}$, entonces la colección de eventos $\{E_1, E_2, \ldots,\}$ es una partición de Ω si $\forall k = 1, \ldots,$ son satisfechas las condiciones siguientes:

1. Si $E_k \neq \emptyset$, entonces $\mathscr{P}(E_k) > 0$.
2. $\forall i \neq j$, entonces $E_i \cap E_j = \emptyset$, es decir, son conjuntos mutuamente excluyentes (conjuntos disjuntos).
3. $\bigcup_{k=1}^{\infty} E_k = \Omega$.

Teorema 3 (probabilidad total). Sea $\{E_1, E_2, \ldots,\}$ una partición del espacio muestra Ω de tal forma que $\forall k = 1, 2, \ldots,$ se tiene que $\mathscr{P}(E_k) > 0$, entonces para todo evento B:

$$\mathscr{P}(B) = \sum_{k=1}^{\infty} \mathscr{P}(B \mid E_k)\, \mathscr{P}(E_k), \tag{2.32}$$

donde $\mathscr{P}(B)$ es la probabilidad del evento B, $\mathscr{P}(B \mid E_k)$ es la probabilidad condicional del evento B dado el evento E_k y $\mathscr{P}(E_k)$ es la probabilidad del evento E_k (propuesto por Laplace 1812). □

2.11. Teorema de Bayes

Teorema 4 (Bayes 1763). Sea $(\Omega, \mathscr{F}, \mathscr{P})$ un espacio de probabilidad y sea $\{E_1, E_2, \ldots,\}$ una partición del espacio muestra Ω de tal forma que $\forall k = 1, \ldots,$ se tiene que $\mathscr{P}(E_k) > 0$, entonces para todo evento B tal que $\mathscr{P}(B) > 0$ y $\forall m = 1, \ldots,$ fijo:

$$\mathscr{P}(E_m \mid B) = \frac{\mathscr{P}(B \mid E_m)\, \mathscr{P}(E_m)}{\sum_{k=1}^{\infty} \mathscr{P}(B \mid E_k)\, \mathscr{P}(E_k)}, \tag{2.33}$$

donde $\mathscr{P}(E_m \mid B)$ es la probabilidad condicional del evento E_m dado el evento B, $\mathscr{P}(B \mid E_m)$ es la probabilidad del evento B dado el evento E_m, $\mathscr{P}(E_m)$ es la probabilidad del evento E_m, $\mathscr{P}(B \mid E_k)$ es la probabilidad del evento B dado el evento E_k y $\mathscr{P}(E_k)$ es la probabilidad del evento E_k. □

Nota 7 (teorema de Bayes). La partición $\{E_1, E_2, \ldots,\}$ representa causales mutuamente excluyentes y exhaustivas del evento B, es decir, una vez observado el evento B se determina cuál es la causa más probable, donde $\mathscr{P}(E_m)$ es la probabilidad de la causa E_m y $\mathscr{P}(B \mid E_m)$ es la probabilidad del evento B suponiendo que E_m es la causa (Bayes 1763).

Ejemplo 51 (teorema de Bayes). Una población con 51 % de mujeres y 49 % de hombres tiene el 44 % de hombres que realizan deporte y 31 % de mujeres que realizan deporte. Una persona que realiza deporte es elegida al azar, calcular la probabilidad de que la persona elegida sea:

1. Mujer.
2. Hombre.

Solución. Sea $E_1 = M$ el evento elegir una mujer, $E_2 = H$ el evento elegir un hombre y D el evento elegir un deportista, entonces por el teorema 4 (Bayes) y sustituyendo en la ecuación (2.33) es calculada la probabilidad de que la persona que realiza deporte elegida al azar sea:

1. Mujer.

$$\mathscr{P}(E_1 \mid D) = \frac{\mathscr{P}(D \mid E_1)\, \mathscr{P}(E_1)}{\sum_{k=1}^{2} \mathscr{P}(D \mid E_k)\, \mathscr{P}(E_k)} = \frac{\frac{31}{100}\left(\frac{51}{100}\right)}{\frac{31}{100}\left(\frac{51}{100}\right) + \frac{44}{100}\left(\frac{49}{100}\right)} = 0.\overline{472839175809} \approx 47.28\ \%.$$

   Por lo tanto, la probabilidad de que la persona elegida al azar sea mujer dado que realiza deporte es $\mathscr{P}(M \mid D) = 47.28\ \%$.

2. Hombre.

$$\mathscr{P}(E_2 \mid D) = \frac{\mathscr{P}(D \mid E_2)\, \mathscr{P}(E_2)}{\sum_{k=1}^{2} \mathscr{P}(D \mid E_k)\, \mathscr{P}(E_k)} = \frac{\frac{44}{100}\left(\frac{49}{100}\right)}{\frac{31}{100}\left(\frac{51}{100}\right) + \frac{44}{100}\left(\frac{49}{100}\right)} = 0.\overline{576933369012} \approx 57.69\ \%.$$

   Por lo tanto, la probabilidad de que la persona elegida al azar sea hombre dado que realiza deporte es $\mathscr{P}(H \mid D) = 57.69\ \%$. □

Ejemplo 52 (teorema de Bayes). Una prueba para detectar una enfermedad presenta probabilidades de un falso negativo y de un falso positivo de $\alpha = 5\ \%$ donde $0 < \alpha < 1$. Suponiendo que la probabilidad de que una persona elegida al azar está enferma es $0 \leq p \leq 1$ y el resultado de realizar la prueba es positivo. Calcular la probabilidad de que la persona está enferma.

Solución. Sea E_1 el evento la persona elegida está enferma, E_2 el evento la prueba resulta positiva, entonces por hipótesis la probabilidad de que la prueba resulte negativa dado que la persona está enferma es $\mathscr{P}(E_2^c \mid E_1)$, la probabilidad de que la prueba resulte positiva dado que la persona no está enferma es $\mathscr{P}(E_2 \mid E_1^c)$ y la probabilidad de que la persona está enferma es $\mathscr{P}(E_1) = p$.

Entonces por el teorema 4 (Bayes), la probabilidad de que la persona está enferma dado que la prueba es positiva es:

$$\mathscr{P}(E_1 \mid E_2) = \frac{\mathscr{P}(E_2 \mid E_1)\,\mathscr{P}(E_1)}{\mathscr{P}(E_2)} = \frac{\mathscr{P}(E_2 \mid E_1)\,\mathscr{P}(E_1)}{\mathscr{P}(E_2 \mid E_1)\,\mathscr{P}(E_1) + \mathscr{P}(E_2 \mid E_1^c)\,\mathscr{P}(E_1^c)}.$$

$$\mathscr{P}(E_1 \mid E_2) = \frac{(1-\alpha)\,p}{(1-2\alpha)\,p+\alpha} = \frac{0.95p}{0.95p+0.05\,(1-p)} = \frac{0.95p}{0.90p+0.05}.$$

Si la probabilidad p se aproxima a cero, dado que la prueba resulta positiva, la probabilidad de que la persona esté enferma también se aproxima a cero. Si $p = 0.006$, entonces $\mathscr{P}(E_1 \mid E_2) = 0.1029$, es decir, existe menos de 11 % de probabilidad de que la persona esté enferma si la prueba es positiva. La probabilidad de que la persona esté enferma es elevada para valores de $p > 0.10$. Si $p = 0.2$, entonces $\mathscr{P}(E_1 \mid E_2) = 0.8261$, es decir, existe más de 82 % de probabilidad de que la persona esté enferma si la prueba es positiva.

Las probabilidades de que la persona esté enferma dado que el resultado de la prueba es positivo en función de la probabilidad $0 \leq p \leq 1$ de que la persona elegida al azar está enferma es presentada en la tabla 2.4.

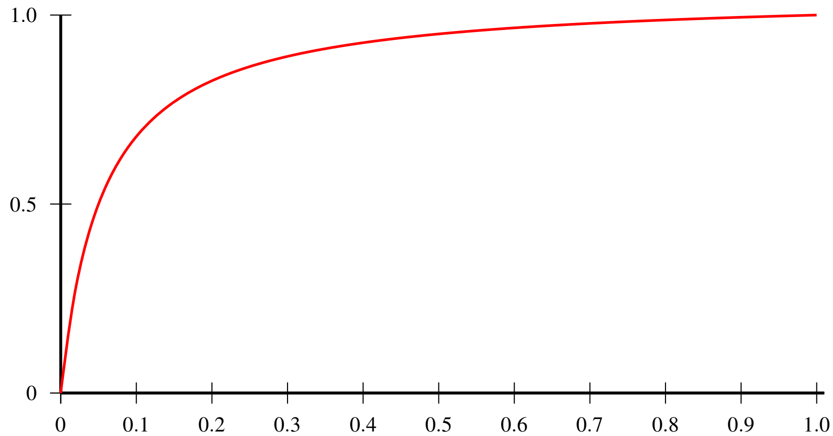
Tabla 2.4: Probabilidad de que una persona está enferma dado que la prueba es positiva.

p	$\mathscr{P}(E_1 \mid E_2)$
0.0000	0.0000
0.0010	0.0187
0.0020	0.0367
0.0030	0.0541
0.0040	0.0709
0.0050	0.0872
0.0060	0.1029
0.0070	0.1181
0.0080	0.1329
0.0090	0.1472
0.0100	0.1610

p	$\mathscr{P}(E_1 \mid E_2)$
0.00	0.0000
0.01	0.1610
0.02	0.2794
0.03	0.3701
0.04	0.4419
0.05	0.5000
0.06	0.5481
0.07	0.5885
0.08	0.6230
0.09	0.6527
0.10	0.6786

p	$\mathscr{P}(A_1 \mid A_2)$
0.00	0.0000
0.10	0.6786
0.20	0.8261
0.30	0.8906
0.40	0.9268
0.50	0.9500
0.60	0.9661
0.70	0.9779
0.80	0.9870
0.90	0.9942
1.00	1.0000

La tabla 2.4 presenta las probabilidades de que la persona esté enferma dado que el resultado de la prueba es positivo en función de la probabilidad $0 \leq p \leq 1$ y las probabilidades son monótonas crecientes y asintóticas a la unidad. Las probabilidades de que la persona esté enferma dado que el resultado de la prueba es positiva en función de la probabilidad $0 \leq p \leq 1$ de que la persona elegida al azar esté enferma es presentada en la gráfica 2.6.



Gráfica 2.6: Probabilidad de que una persona esté enferma dado que la prueba es positiva.

La gráfica 2.6 muestra que las probabilidades de que la persona esté enferma dado que el resultado de la prueba es positivo en función de la probabilidad $0 \leq p \leq 1$ son monótonas crecientes y asintóticas a la unidad. $\square$

2.12. Ejercicios

Ejercicio 11. Una moneda con © y ® es lanzada tres ocasiones consecutivas. Calcular la probabilidad de que el número de veces que:

1. Cae © es estrictamente mayor que el número de veces que cae ®, es decir, $\mathscr{P}(© > ®)$.
2. Cae ® es estrictamente mayor que el número de veces que cae ©, es decir, $\mathscr{P}(® > ©)$.

Ejercicio 12. Una carta es extraída de una baraja[5] con 52 cartas. Calcular la probabilidad de extraer:

1. Un as (A).
2. Un rey (*king*) de corazones $(K\heartsuit)$.
3. Un tres de tréboles $(3\clubsuit)$ o un seis de diamantes $(6\diamondsuit)$.
4. Un corazón $(\heartsuit)$.

Ejercicio 13. Sean A y B eventos ajenos donde $\mathscr{P}(A) = 0.2$ y $\mathscr{P}(B) = 0.5$. Calcular:

1. $\mathscr{P}(A \cup B)$.
2. $\mathscr{P}(A^{c} \cap B^{c})$.

[5]La baraja tiene trece números $2,3,4,5,6,7,8,9,10,J,Q,K,A$ y cuatro símbolos $\heartsuit, \diamondsuit, \clubsuit, \spadesuit$.

Ejercicio 14. Sean A y B eventos donde $\mathscr{P}(A) = p, \mathscr{P}(B) = q$ y $\mathscr{P}(A \cap B) = r$. Calcular las probabilidades:

1. $\mathscr{P}(A \cap B^{c})$.
2. $\mathscr{P}(A^{c} \cap B)$.
3. $\mathscr{P}(A^{c} \cap B^{c})$.

Ejercicio 15. Un alumno es seleccionado al azar, B es el evento donde el alumno practica basquetbol y N es el evento donde el alumno practica natación. Las estadísticas indican que $\mathscr{P}(B) = 0.6, \mathscr{P}(N) = 0.3$ y $\mathscr{P}(B \cap N) = 0.2$. Calcular la probabilidad de que el alumno seleccionado al azar:

1. Practique al menos uno de los dos deportes, es decir, $\mathscr{P}(B \cup N)$.
2. No practique ningún de los dos deportes, es decir, $\mathscr{P}((B \cup N)^{c})$.
3. Practique solamente basquetbol, es decir, $\mathscr{P}(B \cup N^{c})$.

Ejercicio 16. Una compañía ofrece una tarifa si el consumo de electricidad es menor que 200 kWh durante un periodo mensual. Si A es el evento donde el usuario no excede el consumo de 200 kWh durante diciembre y B es el evento donde el usuario no excede el consumo de 200 kWh durante marzo. Las estadísticas indican que $\mathscr{P}(A) = 0.7, \mathscr{P}(B) = 0.8$ y $\mathscr{P}(A \cup B) = 0.9$. Calcular la probabilidad de que el usuario:

1. No exceda el consumo de 200 kWh en diciembre y marzo, es decir, $\mathscr{P}(A \cap B)$.
2. Exceda el consumo de 200 kWh en sólo uno de los meses, es decir, $\mathscr{P}((A \cup B) \setminus (A \cap B))$.

Ejercicio 17. La función $\mathscr{P}$ asigna una probabilidad a cada evento $E \subset \mathbb{R}$, de la forma siguiente:

$$\mathscr{P}(E) = \int_E f(x)\,\mathrm{d}x, \qquad \text{donde } f(x) = \begin{cases} 2x & \text{si } 0 < x < 1 \\ 0 & \text{cualquier otro caso} \end{cases}.$$

Calcular la probabilidad del evento $E_1 = \left\{x \mid \frac{1}{3} \leq x \leq \frac{2}{3}\right\}$.

Ejercicio 18. Suponiendo que no existen empates. Calcular las formas en que son clasificados los tres primeros lugares de una carrera con ocho corredores.

Ejercicio 19. Calcular las formas posibles para sentar a cinco personas si sólo existen tres lugares.

Ejercicio 20. La cava tiene cinco botellas de Zinfandel de fábricas diferentes. Si desea servir dos botellas de Zinfandel y el orden de servicio es importante. Calcular las formas posibles que existen para servir las botellas de Zinfandel.

Ejercicio 21. Calcular las formas posibles que existen para sentar a siete personas en una mesa circular con siete lugares.

Ejercicio 22. Un librero tiene tres libros de Probabilidad. Calcular las formas posibles de ordenar los tres libros de Probabilidad.

Ejercicio 23. Calcular la cantidad de números enteros de tres cifras que es posible formar con el conjunto de los números dígitos.

Ejercicio 24. Si Beethoven escribió nueve sinfonías y Mozart escribió veintisiete conciertos y el locutor de una estación de radio desea iniciar con una sinfonía de Beethoven y luego un concierto de Mozart. Calcular las formas posibles que tiene el locutor para organizar las sinfonías y los conciertos.

Ejercicio 25. Calcular las formas posibles que existen de elegir una comisión de tres personas a partir de un conjunto de ocho personas.

Ejercicio 26. En un grupo de siete ingenieros, hay tres civiles y cuatro industriales que forman un comité de dos civiles y un industrial. Calcular las formas posibles de integrar el comité.

Ejercicio 27. La cava tiene cinco botellas de Zinfandel, cuatro de Merlot y siete de Cabernet de diferentes fábricas. Calcular las formas posibles que existen de servir seis de las dieciséis botellas utilizando dos botellas de cada cepa.

Ejercicio 28. Una selección de basquetbol tiene doce jugadores y sólo quedarán diez jugadores. Calcular las formas que existen para seleccionar a los diez jugadores.

Ejercicio 29. Una urna tiene tres bolas rojas, dos blancas y tres azules. Calcular las formas que existen de extraer las bolas de la urna.

Ejercicio 30. Calcular la probabilidad de que la suma de lanzar dos dados sea siete, por ejemplo: {⚄,⚁}.

Ejercicio 31. Cien productos son elaborados y veinte presentan un peso diferente al deseado, diez de los cien productos son pesados. Calcular la probabilidad de encontrar exactamente dos productos con peso incorrecto.

Ejercicio 32. Cinco cartas de una baraja son repartidas. Calcular la probabilidad de extraer tres ases y dos reyes, es decir, un completo con tres ases y dos reyes, por ejemplo: $C = \{A\heartsuit, A\clubsuit, A\spadesuit, K\diamondsuit, K\spadesuit\}$.

Ejercicio 33. Un librero tiene siete libros de Ingeniería y dos de Actuaría. Calcular la probabilidad de que tres libros de Ingeniería estén juntos.

Ejercicio 34. Si $\mathscr{P}(A) = \frac{1}{4}, \mathscr{P}(B \mid A) = \frac{1}{2}$ y $\mathscr{P}(A \mid B) = \frac{1}{2}$. Calcular $\mathscr{P}(A \mid B) + \mathscr{P}(A \mid B^{c})$.

Ejercicio 35. Dos dados son lanzados. Calcular la probabilidad de que la suma sea diferente de seis u ocho, por ejemplo: {⚃,⚂}.

Ejercicio 36. Dos cartas de una baraja con 52 cartas son repartidas. Calcular la probabilidad de que ambas cartas sean ases, por ejemplo: $\{A\heartsuit, A\spadesuit\}$.

Ejercicio 37. Si $\mathscr{P}(A) = \frac{2}{5}$ y $\mathscr{P}(B \cup A) = \frac{7}{10}$. Calcular $\mathscr{P}(B \mid A)$ suponiendo que:

1. A y B son estocásticamente independientes, es decir, suponiendo que $A \perp B$.
2. A y B son mutuamente excluyentes, es decir, suponiendo que $A \cap B = \emptyset$.
3. La probabilidad de A dado B es un medio, es decir, suponiendo que $\mathscr{P}(A \mid B) = \frac{1}{2}$.

Ejercicio 38. Cuatro monedas con © y ®, donde para la k–ésima moneda $\mathscr{P}(©) = \frac{k}{5}$ para $k = 1, 2, 3, 4$. Si una de las cuatro monedas es seleccionada y es lanzada. Calcular $\mathscr{P}(®)$.

Ejercicio 39. Un estudiante contesta un examen de opción múltiple donde cada pregunta tiene dos opciones y sólo una es correcta. Si el estudiante conoce la respuesta correcta, la selecciona, en caso contrario, selecciona una de las opciones al azar. Suponiendo que el estudiante conoce la respuesta correcta de cualquiera de las diez preguntas con probabilidad 0.5. Calcular la probabilidad de que el estudiante:

1. Responda correctamente una de las preguntas, es decir, $\mathscr{P}(C)$.
2. Responda correctamente más de cinco preguntas, es decir, $\mathscr{P}(A)$.

Ejercicio 40. Una persona selecciona uno de los números 1 o 2, con probabilidades idénticas, después lanza un dado las veces indicadas por el número seleccionado y suma los resultados de los lanzamientos. Calcular la probabilidad de seleccionar el número dos dado que la suma de los lanzamientos es cinco, es decir, $\mathscr{P}(2 \mid 5)$.

Capítulo 3

Variables aleatorias

Una variable aleatoria es una función del espacio muestra Ω a los números reales, entonces el análisis del experimento aleatorio $\mathscr{E}$ es trasladado al análisis de las variables aleatorias y sus características, es decir, el análisis de un experimento aleatorio $\mathscr{E}$ en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathscr{P})$.

Definición 51 (σ–álgebra de Borel en $\mathbb{R}$). Una σ–álgebra de Borel en $\mathbb{R}$ es la mínima σ–álgebra generada por los intervalos $(-\infty, x]$ donde $x \in \mathbb{R}$:

$$\mathscr{B}(\mathbb{R}) = \sigma\{(-\infty, x] \mid x \in \mathbb{R}\}, \tag{3.1}$$

donde $\mathscr{B}(\mathbb{R})$ es una σ–álgebra de Borel en $\mathbb{R}$ y son conjuntos de Borel medibles o borelianos de $\mathbb{R}$.

3.1. Variable aleatoria

El valor obtenido depende del resultado específico del experimento aleatorio $\mathscr{E}$, entonces las características de interés son representadas con una función X que depende del resultado del experimento aleatorio $\mathscr{E}$ con una función de probabilidad definida y dado un conjunto de Borel medible $\mathscr{B}(\mathbb{R})$ es posible calcular la probabilidad del conjunto $(B \in \mathscr{B}(\mathbb{R}))$ si éste es un elemento de la σ–álgebra $\mathscr{F}$.

Definición 52 (variable aleatoria). Una variable aleatoria es una función $X \colon \Omega \to \mathbb{R}$, donde $\forall B \in \mathscr{B}(\mathbb{R})$ se satisface que el conjunto $X^{-1}B \in \mathscr{F}$, entonces $\forall x \in \mathbb{R}$:

$$(X \in B) = \{\omega \in \Omega \mid X(\omega) \in B\} \in \mathscr{F}, \tag{3.2}$$

donde $(X \in B)$ es un conjunto de Borel medible respecto a la σ–álgebra $\mathscr{F}$ y $X^{-1}B \in \mathscr{F}$ es la imagen inversa de $B \in \mathscr{B}(\mathbb{R})$.

Es decir, $\forall x \in \mathbb{R}$ se satisface que $(X \in B) = \{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathscr{F}$ es un conjunto Borel medible.

Las variables aleatorias son clasificadas en:

1. Discretas.
2. Continuas.

La clasificación es realizada por el conjunto de valores que tiene la variable aleatoria. Por ejemplo, una variable aleatoria discreta tiene valores en un conjunto numerable (finito o infinito) y una variable aleatoria continua tiene una función de distribución continua.

3.1.1. Variable aleatoria discreta

Definición 53 (variable aleatoria discreta). Una variable aleatoria es discreta si existe una colección finita o infita numerable $V_x = \{x_1, x_2, \ldots\}$ donde $\forall x_k$:

$$f_X(x_k) = \mathscr{P}(X = x_k) \geq 0 \qquad \text{si } x_k = x_1, x_2, \ldots, \tag{3.3}$$

$$F_X(x) = \sum_{k \in V_x} f_X(x_k) = \sum_{k \in V_x} \mathscr{P}(X = x_k) = 1 \qquad \text{si } x_k = x_1, x_2, \ldots, \tag{3.4}$$

donde $f_X(x)$ es la función densidad de probabilidad de la variable aleatoria X, $\mathscr{P}(X = x_k)$ es la probabilidad de que la variable aleatoria X tenga el valor x_k y $V_x = \{x_1, x_2, \ldots, \}$ es el conjunto de valores de la variable aleatoria.

3.1.2. Variable aleatoria continua

Definición 54 (variable aleatoria continua). Una variable aleatoria es continua si $\forall x \in \mathbb{R}$ existe una función no negativa $f_X(x) : \mathbb{R} \to \mathbb{R}_+$ e integrable que satisface:

$$f_X(x_k) \geq 0, \qquad \forall x \in \mathbb{R}, \tag{3.5}$$

$$\lim_{x \to \infty} F_X(x) = \mathscr{P}(X \leq x) = \lim_{x \to \infty} \int_{-\infty}^{x} f_U(u)\, \mathrm{d}u = 1, \tag{3.6}$$

donde $f_X(x)$ es la función densidad de la variable aleatoria X, $F_X(x)$ es la función de distribución de la variable aleatoria X y $\mathscr{P}(X \leq x)$ es la probabilidad de que la variable aleatoria X tenga un valor menor o igual que x.

3.1.3. Propiedades de las variables aleatorias

Las operaciones con variables aleatorias en $(\Omega, \mathscr{F}, \mathscr{P})$ son variables aleatorias. Las propiedades son:

Propiedad 42. Si X es una variable aleatoria y $c \in \mathbb{R}$ es una constante, entonces $X = c$ es una variable aleatoria.

Propiedad 43. Si X es una variable aleatoria y $c \in \mathbb{R}$ es una constante, entonces cX es una variable aleatoria.

Propiedad 44. Si X y Y son variables aleatorias, entonces $Z = X + Y$ es una variable aleatoria.

Propiedad 45. Si X y Y son variables aleatorias, entonces $Z = XY$ es una variable aleatoria.

Propiedad 46. Si X y Y son variables aleatorias con $Y \neq 0$, entonces $Z = \frac{X}{Y}$ es una variable aleatoria.

3.2. Función de densidad

Definición 55 (función de densidad discreta). La función de densidad discreta es una función $f_X(x) : \mathbb{R} \to \mathbb{R}$ que $\forall x \in \mathbb{R}$ satisface:

$$f_X(x) = \mathscr{P}(X = x). \tag{3.7}$$

$$F_X(x) = \sum_{x_k \in V_x}^{x_n} f_X(x_k) = \mathscr{P}(X \leq x_n). \tag{3.8}$$

Definición 56 (función de densidad continua). La función de densidad continua es una función no negativa $f_X(x) : \mathbb{R} \to \mathbb{R}$ que $\forall x \in \mathbb{R}$ satisface:

$$f_X(x) \geq 0. \tag{3.9}$$

$$F_X(x) = \int_{-\infty}^{x} f_U(u)\,\mathrm{d}u = \mathscr{P}(X \leq x). \tag{3.10}$$

3.3. Función de distribución

Definición 57 (función de distribución). Sea $X \colon \Omega \to \mathbb{R}$ una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathscr{P})$, entonces la función de distribución es una función $F_X(x) : \mathbb{R} \to [0,1]$ que satisface:

$$F_X(x) = \mathscr{P}(X \leq x) = \mathscr{P}(\{\omega \in \Omega \mid X(\omega) \leq x\}), \tag{3.11}$$

donde $F_X(x)$ es la función de distribución y $\mathscr{P}(X \leq x)$ es la probabilidad de que la variable aleatoria X tenga un valor menor o igual que x.

Definición 58 (función de distribución discreta). Sea $X \colon \Omega \to \mathbb{R}$ una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathscr{P})$, entonces la función de distribución discreta es una función $F_X(x) : \mathbb{R} \to [0,1]$ que satisface:

$$F_X(x) = \mathscr{P}(X \leq x) = \sum_{x=0}^{x} f_X(x), \tag{3.12}$$

donde $F_X(x)$ es la función de distribución discreta, $\mathscr{P}(X \leq x)$ es la probabilidad de que la variable aleatoria X tenga un valor menor o igual que x y $f_X(x)$ es la función de densidad discreta.

Definición 59 (función de distribución continua). Sea $X \colon \Omega \to \mathbb{R}$ una variable aleatoria definida en un espacio de probabilidad $(\Omega, \mathscr{F}, \mathscr{P})$, entonces la función de distribución continua es una función no negativa $F_X(x) : \mathbb{R} \to [0,1]$ que satisface:

$$F_X(x) = \mathscr{P}(X \leq x) = \int_{-\infty}^{x} f_U(u)\,\mathrm{d}u, \tag{3.13}$$

donde $F_X(x)$ es la función de distribución continua, $\mathscr{P}(X \leq x)$ es la probabilidad de que la variable aleatoria X tenga un valor menor o igual que x y $f_X(x)$ es la función de densidad continua.

3.3.1. Propiedades de la función de distribución

La función de distribución $F_X(x)$ contiene información de la variable aleatoria X y de la medida de probabilidad. Las propiedades son:

Proposición 9 (función de distribución). Si X es una variable aleatoria, entonces:

$$\lim_{x\to-\infty} F_X(x) = 0, \tag{3.14}$$

$$\lim_{x\to+\infty} F_X(x) = 1, \tag{3.15}$$

$$\text{si } x_1 \leq x_2 \Rightarrow F_X(x_1) \leq F_X(x_2). \tag{3.16}$$

Proposición 10 (función de distribución). Si X es una variable aleatoria, entonces:

$$\mathscr{P}(X = x) = F_X(x) - F_X(x_-), \tag{3.17}$$

$$\mathscr{P}(X < x) = F_X(x_-), \tag{3.18}$$

$$\mathscr{P}(a < X \leq b) = F_X(b) - F_X(a), \tag{3.19}$$

$$\mathscr{P}(a \leq X \leq b) = F_X(b) - F_X(a_-), \tag{3.20}$$

$$\mathscr{P}(a < X < b) = F_X(b_-) - F_X(a), \tag{3.21}$$

$$\mathscr{P}(a \leq X < b) = F_X(b_-) - F_X(a_-), \tag{3.22}$$

donde $F_X(x_-) = \lim_{h\to 0} F_X(x-h)$ es el límite por la izquierda y $F_X(x) = \lim_{h\to 0} F_X(x+h)$ es el límite por la derecha.

3.4. Esperanza matemática

Los conceptos de esperanza y probabilidad surgieron en los inicios del cálculo de probabilidades. Pascal (1654) y Fermat resolvieron problemas de probabilidad con métodos que son fundamentos de la teoría de probabilidad. Huygens (1657b) resolvió los problemas que resolvieron Pascal y Fermat y otros sin utilizar probabilidades de eventos, las soluciones están basadas en el cálculo de esperanzas.

Definición 60 (esperanza matemática). Sea X una variable aleatoria discreta con función de densidad $f_X(x)$ y $g(x) : \mathbb{R} \to \mathbb{R}$ una función $\mathscr{B}(\mathbb{R})$–medible tal que $\sum_{x_k\in V_x} |g(x_k)| < \infty$, entonces $g(x)$ tiene esperanza finita si sólo si $\sum_{x_k\in V_x} |g(x_k)| f_X(x) < \infty$, por lo tanto:

$$\mathrm{E}(g(x)) = \sum_{x_k\in V_x} g(x_k) f_X(x_k), \tag{3.23}$$

donde $\mathrm{E}(g(x))$ es la esperanza matemática de la función $g(x)$.

Nota 8 (esperanza matemática de una variable aleatoria discreta). Si $g(x) = X$, entonces:

$$\mathrm{E}(X) = \sum_{x=0}^{\infty} x f_X(x).$$

Definición 61 (esperanza matemática). Sea X una variable aleatoria continua con función de densidad $f_X(x)$ y $g(x) : \mathbb{R} \to \mathbb{R}$ una función $\mathscr{B}(\mathbb{R})$–medible tal que $\int\limits_{-\infty}^{\infty} |g(x)| \,\mathrm{d}x < \infty$, entonce, $g(x)$ tiene esperanza finita si sólo si $\int\limits_{-\infty}^{\infty} |g(x)| f_X(x) \,\mathrm{d}x < \infty$, por lo tanto:

$$\mathrm{E}(g(x)) = \int_{-\infty}^{\infty} g(x) f_X(x) \,\mathrm{d}x, \tag{3.24}$$

donde $\mathrm{E}(g(x))$ es la esperanza matemática de la función $g(x)$.

Nota 9 (esperanza matemática de una variable aleatoria continua). Si $g(x) = X$, entonces:

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x) \,\mathrm{d}x.$$

La esperanza matemática de la variable aleatoria X representa el promedio de los valores que toma la variable aleatoria cuando el experimento aleatorio $\mathscr{E}$ es repetido varias ocasiones. La esperanza matemática de X es el valor que es esperado en promedio.

3.4.1. Propiedades de la esperanza matemática

Las propiedades básicas de la esperanza matemática son:

Proposición 11 (esperanza matemática). Si X y Y son variables aleatorias y $c \in \mathbb{R}$ es una constante, entonces:

$$\mathrm{E}(c) = c. \tag{3.25}$$

$$\mathrm{E}(cX) = c\,\mathrm{E}(X). \tag{3.26}$$

$$\text{Si } X \geq 0 \Rightarrow \mathrm{E}(X) \geq 0. \tag{3.27}$$

$$\text{Si } X \leq Y \Rightarrow \mathrm{E}(X) \leq \mathrm{E}(Y). \tag{3.28}$$

$$\mathrm{E}(X+Y) = \mathrm{E}(X) + \mathrm{E}(Y). \tag{3.29}$$

3.5. Varianza

Los valores de la variable aleatoria X difieren de la esperanza $\mathrm{E}(X)$ que es aproximadamente el promedio de los valores de X, entonces $\mathrm{E}\left((X - \mathrm{E}(X))^2\right)$ establece una medida para la diferencia entre $\mathrm{E}(X)$ y los valores de X, donde $\mathrm{E}(X)$ es la aproximación óptima de los valores de X.

Definición 62 (varianza). Sea X una variable aleatoria discreta con $f_X(x)$ donde $\mathrm{E}(X) < \infty$, entonces:

$$\mathrm{Var}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right) = \sum_{x_k \in V_x} (x_k - \mathrm{E}(X))^2 f_X(x) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X), \tag{3.30}$$

donde $\mathrm{Var}(X)$ es la varianza de X y es una medida de dispersión con respecto a $\mathrm{E}(X)$.

Definición 63 (varianza). Sea X una variable aleatoria continua con $f_X(x)$ donde $\mathrm{E}(X) < \infty$, entonces:

$$\mathrm{Var}(X) = \mathrm{E}\left((X - \mathrm{E}(X))^2\right) = \int_{-\infty}^{\infty} (x - \mathrm{E}(X))^2 f_X(x) \,\mathrm{d}x = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X), \tag{3.31}$$

donde $\mathrm{Var}(X)$ es la varianza de X y es una medida de dispersión con respecto a $\mathrm{E}(X)$.

3.5.1. Propiedades de la varianza

Las propiedades básicas de la varianza son:

Proposición 12 (varianza). Si X es una variable aleatoria y $c \in \mathbb{R}$ es una constante, entonces:

$$\operatorname{Var}(c) = 0. \tag{3.32}$$
$$\operatorname{Var}(X) \geq 0. \tag{3.33}$$
$$\operatorname{Var}(cX) = c^2 \operatorname{Var}(X). \tag{3.34}$$
$$\operatorname{Var}(X + c) = \operatorname{Var}(X). \tag{3.35}$$
$$\operatorname{Var}(X) = \operatorname{E}\left(X^2\right) - \operatorname{E}^2(X). \tag{3.36}$$

Ejemplo 53 (variable aleatoria discreta). Sea X una variable aleatoria discreta y $f(x) = c2^x$ para $x = 0, 1, 2, 3$.

1. Encontrar la constante c para que $f(x)$ sea una función de densidad.
2. Presentar la gráfica de $f_X(x)$ y $F_X(x)$.
3. Calcular $\operatorname{E}(X)$ y $\operatorname{Var}(X)$.

Solución. Si $f(x) = 2^x$, entonces $f(x) = \left\{(x, f(x)) \in \mathbb{R}^2 \mid (0,1), (1,2), (2,4), (3,8)\right\}$, por lo tanto $f(x) \geq 0$. Si la constante $c \geq 0$, entonces se satisface la ecuación (3.3) de la definición 53 (variable aleatoria discreta).

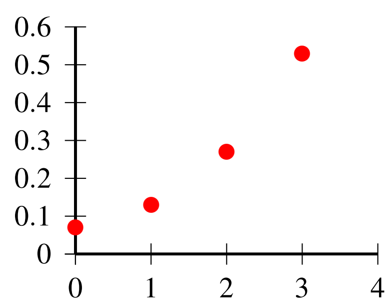
1. Para encontrar la constante c es necesario satisfacer la ecuación (3.4) de la definición 53 (variable aleatoria discreta), entonces:

$$c\sum_{x=0}^{3} 2^x = 1 \Leftrightarrow c(1+2+4+8) = 1 \Leftrightarrow \frac{c\left(2^4 - 1\right)}{2-1} = 1 \Leftrightarrow 15c = 1 \Leftrightarrow c = \frac{1}{15}.$$

   Entonces[1] $c \geq 0$, por lo tanto $f_X(x) = \frac{2^x}{15}$ es una función de densidad de probabilidad.

$$f_X(x) = \begin{cases} \frac{2^x}{15} & \text{si } x = 0,1,2,3 \\ 0 & \text{cualquier otro caso} \end{cases}.$$

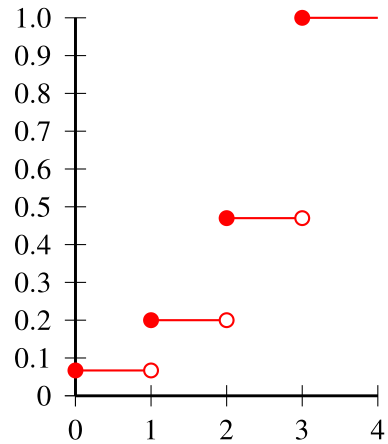
2. La gráfica de $f_X(x) = \frac{2^x}{15}$ es presentada en la gráfica 3.1.



Gráfica 3.1: $f_X(x) = \frac{2^x}{15}$.

[1] Si la razón geométrica $r \neq 1$, la serie $S = \sum_{x=0}^{n} r^x$, entonces $rS = \sum_{x=1}^{n+1} r^x$ y $S - rS = 1 - r^{n+1}$, por lo tanto $S = \frac{r^{n+1}-1}{r-1}$.

La gráfica de $F_X(x) = \sum_{x=0}^{3} \frac{2^x}{15}$ es presentada en la gráfica 3.2.



Gráfica 3.2: $F_X(x) = \sum_{x=0}^{3} \dfrac{2^x}{15}$.

3. Sea $g(x) = X$, entonces sustituyendo en la ecuación (3.23) de la definición 60 (esperanza matemática):

$$\mathrm{E}(X) = \frac{1}{15}\sum_{x=0}^{3} x2^x = \frac{0+2+8+24}{15} = \frac{34}{15} = 2.2\overline{6}.$$

Sea $g(x) = (X - \mathrm{E}(X))^2$, entonces sustituyendo en la ecuación (3.30) de la definición 62 (varianza):

$$\operatorname{Var}(X) = \sum_{x=0}^{3}\left(x - \frac{34}{15}\right)^2 \frac{2^x}{15} = \frac{1}{15}\left(\left(0-\frac{34}{15}\right)^2 + 2\left(1-\frac{34}{15}\right)^2 + 2^2\left(2-\frac{34}{15}\right)^2 + 2^3\left(3-\frac{34}{15}\right)^2\right).$$
$$\operatorname{Var}(X) = \frac{194}{225} = 0.86\overline{2}.$$

Por lo tanto $c = \frac{1}{15} = 0.0\overline{6}$, $\mathrm{E}(X) = \frac{34}{15} = 2.2\overline{6}$ y $\operatorname{Var}(X) = \frac{194}{225} = 0.86\overline{2}$. □

Ejemplo 54 (variable aleatoria continua). Sea X una variable aleatoria continua y $f(x) = c2^x$ para $0 \leq x \leq 3$.

1. Encontrar la constante c para que $f(x)$ sea una función de densidad.
2. Presentar la gráfica de $f_X(x)$ y $F_X(x)$.
3. Calcular $\mathrm{E}(X)$ y $\operatorname{Var}(X)$.

Solución. Si $f(x) = 2^x$, entonces $f(x) = \left\{(x, f(x)) \in \mathbb{R}^2 \mid (0,1), (1,2), (2,4), (3,8)\right\}$, por lo tanto $f(x) \geq 0$. Si la constante $c \geq 0$, entonces se satisface la ecuación (3.5).

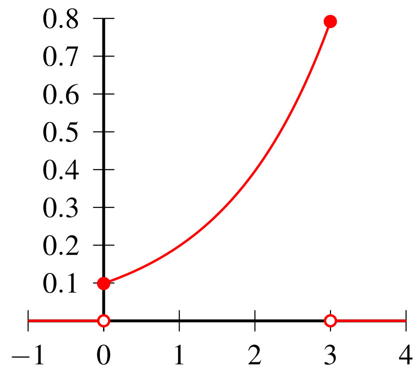
1. Para encontrar la constante c es necesario satisfacer la ecuación (3.6), entonces:

$$c\int_0^3 2^x \,\mathrm{d}x = 1 \Leftrightarrow \left.\frac{c2^x}{\ln(2)}\right|_0^3 = 1 \Leftrightarrow \frac{7c}{\ln(2)} = 1 \Leftrightarrow c = \frac{\ln(2)}{7} \approx 0.099021.$$

Entonces $c \geq 0$, por lo tanto $f_X(x) = \frac{2^x \ln(2)}{7}$ es una función de densidad de probabilidad.

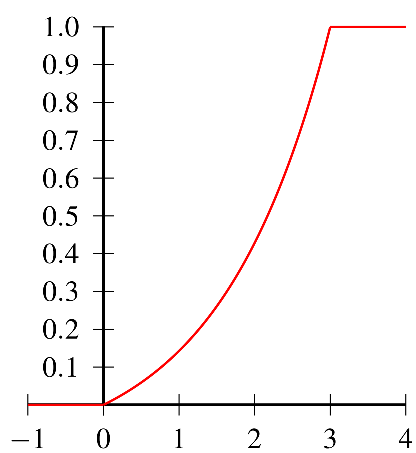
$$f_X(x) = \begin{cases} \frac{2^x \ln(2)}{7} & \text{si } 0 \leq x \leq 3 \\ 0 & \text{cualquier otro caso} \end{cases}.$$

2. La gráfica de $f_X(x) = \frac{2^x \ln(2)}{7}$ es presentada en la gráfica 3.3.



Gráfica 3.3: $f_X(x) = \frac{2^x \ln(2)}{7}$.

La gráfica[2] de $F_X(x) = \frac{2^x - 1}{7}$ es presentada en la gráfica 3.4.



Gráfica 3.4: $F_X(x) = \frac{2^x - 1}{7}$.

3. Sea $g(x) = X$, entonces sustituyendo[3] en la ecuación (3.24) de la definición 61 (esperanza matemática):

$$\mathrm{E}(X) = \frac{\ln(2)}{7} \int_0^3 x 2^x \,\mathrm{d}x = \left.\frac{2^x (x \ln(2) - 1)}{7 \ln(2)}\right|_0^3 = \frac{8((3\ln(2) - 1) + 1}{7\ln(2)} \approx 1.985876.$$

[2] $F_X(x) = \frac{\ln(2)}{7} \int\limits_0^x 2^u \,\mathrm{d}u = \left.\frac{2^u}{7}\right|_0^x = \frac{2^x - 1}{7}$.

[3] Integrando por partes $\int u \,\mathrm{d}v = uv - \int v \,\mathrm{d}u$. Sean $u = x$ y $\mathrm{d}v = 2^x \,\mathrm{d}x$, entonces $\mathrm{d}u = \mathrm{d}x$, $\int \mathrm{d}v = \int 2^x \,\mathrm{d}x$ y $v = \frac{2^x}{\ln(2)}$.

Sea $g(x) = (X - \mathrm{E}(X))^2$, entonces sustituyendo[4] en la ecuación (3.31) de la definición 63 (varianza):

$$\mathrm{Var}(X) = \frac{\ln(2)}{7}\int_0^3 (x - \mathrm{E}(X))^2 2^x \,\mathrm{d}x.$$
$$\mathrm{Var}(X) = \frac{2^x}{7}\left(\frac{x^2\ln^2(2) - 2x\ln(2) + 2}{\ln^2(2)} + \frac{2\mathrm{E}(X)(1 - x\ln(2))}{\ln(2)} + \mathrm{E}^2(X)\right)\Bigg|_0^3 \approx 0.611981.$$

Por lo tanto $c \approx 0.099021, \mathrm{E}(X) \approx 1.985876$ y $\mathrm{Var}(X) \approx 0.611981$. □

Ejemplo 55 (variable aleatoria discreta). Sea X una variable aleatoria discreta y $f(x) = cx$ para $x = 0,1,2,3$.

1. Encontrar la constante c para que $f(x)$ sea una función de densidad.
2. Presentar la gráfica de $f_X(x)$ y $F_X(x)$.
3. Calcular $\mathrm{E}(X)$ y $\mathrm{Var}(X)$.

Solución. Si $f(x) = cx$, entonces $f(x) = \left\{(x, f(x)) \in \mathbb{R}^2 \mid (0,0), (1,c), (2,2c), (3,3c)\right\}$, entonces $f(x) \geq 0$. Si la constante $c \geq 0$, entonces se satisface la ecuación (3.3) de la definición 53 (variable aleatoria discreta).

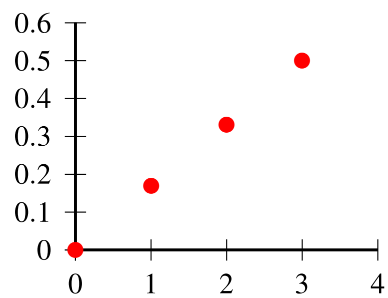
1. Para encontrar la constante c es necesario satisfacer la ecuación (3.4) de la definición 53 (variable aleatoria discreta), entonces:

$$c\sum_{x=0}^{3} x = 1 \Leftrightarrow c(0+1+2+3) = 1 \Leftrightarrow 6c = 1 \Leftrightarrow c = \frac{1}{6}.$$

Entonces $c \geq 0$, por lo tanto $f_X(x) = \frac{x}{6}$ es una función de densidad de probabilidad.

$$f_X(x) = \begin{cases} \frac{x}{6} & \text{si } x = 0,1,2,3 \\ 0 & \text{cualquier otro caso} \end{cases}.$$

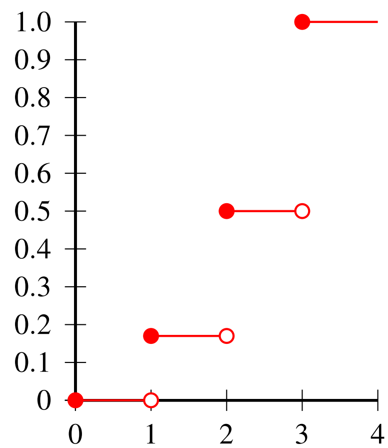
2. La gráfica de $f_X(x) = \frac{x}{6}$ es presentada en la gráfica 3.5.



Gráfica 3.5: $f_X(x) = \frac{x}{6}$.

[4] Sean $u = 2^x$ y $\mathrm{d}u = 2^x\ln(2)\,\mathrm{d}x$, entonces $\mathrm{Var}(X) = \frac{1}{7\ln^2(2)}\int(\ln(u) - \mathrm{E}(X)\ln(2))^2\,\mathrm{d}u$. Desarrollando el binomio al cuadrado $\mathrm{Var}(X) = \frac{\mathrm{E}^2(X)u}{7} - \frac{2\mathrm{E}(X)}{7\ln(2)}\int\ln(u)\,\mathrm{d}u + \frac{1}{7\ln^2(2)}\int\ln^2(u)\,\mathrm{d}u$. Las integrales son resueltas por partes y $u = 2^x$ es sustituida para obtener el resultado.

La gráfica de $F_X(x) = \sum_{x=0}^{3} \frac{2^x}{15}$ es presentada en la gráfica 3.6.



Gráfica 3.6: $F_X(x) = \sum_{x=0}^{3} \frac{x}{6}$.

3. Sea $g(x) = X$, entonces sustituyendo en la ecuación (3.23) de la definición 60 (esperanza matemática):

$$\mathrm{E}(X) = \frac{1}{6}\sum_{x=0}^{3} x^2 = \frac{0+1+4+9}{6} = \frac{7}{3} = 2.\overline{3}.$$

Sea $g(x) = (X - \mathrm{E}(X))^2$, entonces sustituyendo en la ecuación (3.30) de la definición 63 (varianza):

$$\mathrm{Var}(X) = \sum_{x=0}^{3}\left(x - \frac{7}{3}\right)^2 \frac{x}{6} = \frac{1}{6}\left(0\left(0-\frac{7}{3}\right)^2 + 1\left(1-\frac{7}{3}\right)^2 + 2\left(2-\frac{7}{3}\right)^2 + 3\left(3-\frac{7}{3}\right)^2\right).$$
$$\mathrm{Var}(X) = \frac{5}{9} = 0.\overline{5}.$$

Por lo tanto $c = \frac{1}{6} = 0.1\overline{6}, \mathrm{E}(X) = \frac{7}{3} = 2.\overline{3}$ y $\mathrm{Var}(X) = \frac{5}{9} = 0.\overline{5}$. □

Ejemplo 56 (variable aleatoria continua). Sea X una variable aleatoria continua y $f(x) = cx^2$ para $-1 \leq x \leq 1$.

1. Encontrar la constante c para que $f(x)$ sea una función de densidad.
2. Presentar la gráfica de $f_X(x)$ y $F_X(x)$.
3. Calcular $\mathrm{E}(X)$ y $\mathrm{Var}(X)$.

Solución. Si $f(x) = cx^2$, entonce $f(x) = \left\{(x, f(x)) \in \mathbb{R}^2 \mid (-1,c),(0,0),(1,c)\right\}$, por lo tanto $f(x) \geq 0$. Si la constante $c \geq 0$, entonces se satisface la ecuación (3.5) de la definición 54 (variable aleatoria continua).

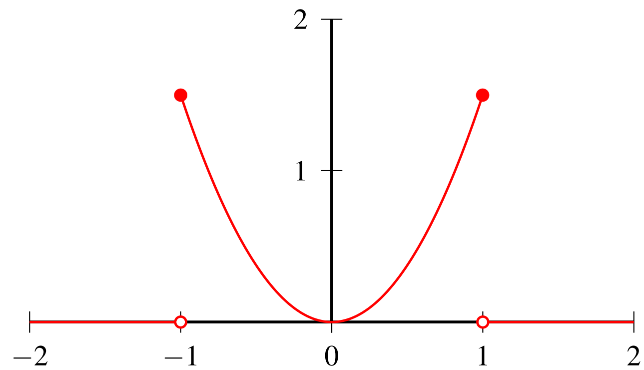
1. Para encontrar la constante c es necesario satisfacer la ecuación (3.6) de la definición 54 (variable aleatoria continua), entonces:

$$c\int_{-1}^{1} x^2 = 1 \Leftrightarrow \left.\frac{cx^3}{3}\right|_{-1}^{1} = 1 \Leftrightarrow \frac{2c}{3} = 1 \Leftrightarrow c = \frac{3}{2}.$$

Entonces $c \geq 0$ y por lo tanto $f_X(x) = \frac{3x^2}{2}$ es una función de densidad de probabilidad.

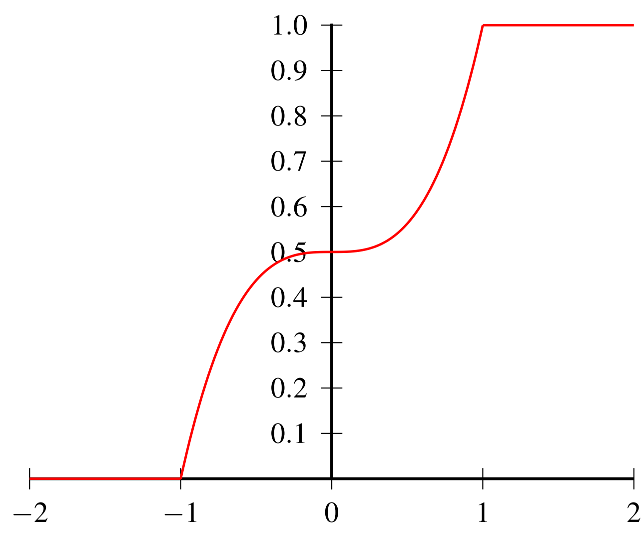
$$f_X(x) = \begin{cases} \frac{3x^2}{2} & \text{si } -1 \leq x \leq 1 \\ 0 & \text{cualquier otro caso} \end{cases}.$$

2. La gráfica de $f_X(x) = \frac{3x^2}{2}$ es presentada en la gráfica 3.7.



Gráfica 3.7: $f_X(x) = \frac{3x^2}{2}$.

La gráfica[5] de $F_X(x) = \frac{x^3+1}{2}$ es presentada en la gráfica 3.8.



Gráfica 3.8: $F_X(x) = \frac{x^3+1}{2}$.

[5] $F_X(x) = \frac{3}{2}\int_{-1}^{x} u^2 \,\mathrm{d}u = \left.\frac{u^3}{2}\right|_{-1}^{x} = \frac{x^3+1}{2}$.

3. Sea $g(x) = X$, entonces sustituyendo en la ecuación (3.24) de la definición 61 (esperanza matemática):

$$\mathrm{E}(X) = \frac{3}{2}\int_{-1}^{1} x^3 \,\mathrm{d}x = \left.\frac{3x^4}{8}\right|_{-1}^{1} = \frac{3(1-1)}{8} = 0.$$

Sea $g(x) = (X - \mathrm{E}(X))^2$, entonces sustituyendo en la ecuación (3.31) de la definición 63 (varianza):

$$\mathrm{Var}(X) = \frac{3}{2}\int_{-1}^{1} (x - \mathrm{E}(X))^2 x^2 \,\mathrm{d}x = \frac{3}{2}\int_{-1}^{1} (x-0)^2 x^2 \,\mathrm{d}x = \frac{3}{2}\int_{-1}^{1} x^4 \,\mathrm{d}x = \left.\frac{3x^5}{10}\right|_{-1}^{1} = \frac{3}{5}.$$

Por lo tanto $c = \frac{3}{2} = 1.5, \mathrm{E}(X) = 0$ y $\mathrm{Var}(X) = \frac{3}{5} = 0.6$. □

3.6. Ejercicios

Ejercicio 41. Un par de dados son lanzados y la variable aleatoria X es la suma de los puntos:

1. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
2. Calcular la esperanza $\mathrm{E}(X)$.
3. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de obtener una suma par.
5. Calcular la probabilidad de obtener una suma menor que 8.

Ejercicio 42. Las estadísticas indican que 65 % de las afinaciones son para automóviles de cuatro cilindros, 30 % de seis cilindros y 5 % de ocho cilindros. Si la variable aleatoria X es el número de cilindros:

1. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
2. Calcular la esperanza $\mathrm{E}(X)$.
3. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de afinar autos de cuatro y ocho cilindros.
5. Calcular la probabilidad de afinar autos de seis y ocho cilindros.

Ejercicio 43. El departamento de planeación necesita presentar uno, dos, tres o cuatro formatos para solicitar un permiso. Si la variable aleatoria Y es el número de formatos y las estadísticas indican que la probabilidad de presentar los formatos es directamente proporcional a Y, es decir, $f(y) = cy$ para $y = 1,2,3,4$:

1. Calcular el valor de la constante $c \in \mathbb{R}$ para que $f(y)$ sea una función de densidad discreta.
2. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
3. Calcular la esperanza $\mathrm{E}(X)$.
4. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
5. Calcular la probabilidad de necesitar al menos 3 formatos.
6. Calcular la probabilidad de necesitar 2 o 3 formatos.

Ejercicio 44. Una lotería reparte 20 premios de \$ 100, 10 premios de \$ 200 y un premio de \$ 1,000. La variable aleatoria X es la cantidad de dinero que es pagada por los boletos premiados y suponiendo que son vendidos 5,000 boletos:

1. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
2. Calcular la esperanza $\mathrm{E}(X)$.
3. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
4. Calcular el precio justo de venta de cada uno de los 5,000 boletos.

Ejercicio 45. Una persona aseguró su automóvil, la variable aleatoria Y es el número de infracciones cometidas por el asegurado y la función de densidad es:

$$f_Y(y) = \begin{cases} \frac{7}{10} & \text{si } y = 0 \\ \frac{2}{10} & \text{si } y = 1 \\ \frac{1}{10} & \text{si } y = 2 \end{cases}.$$

Suponiendo que cada infracción asciende a \$ 500:

1. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
2. Calcular la esperanza $\mathrm{E}(X)$.
3. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de cometer al menos una infracción.

Ejercicio 46. La función de distribución de una variable aleatoria X es:

$$F_X(x) = \begin{cases} 1 - \exp(-2x) & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}.$$

1. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
2. Calcular la esperanza $\mathrm{E}(X)$.
3. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de que $3 \leq x \leq 5$.
5. Calcular la probabilidad de que $x \geq 2$.

Ejercicio 47. Un profesor concluye la cátedra un minuto después de la hora, la variable aleatoria X es el tiempo que transcurre entre el final de la hora y el final de la cátedra. Si la función es $f(x) = cx^2$:

1. Calcular el valor de la constante $c \in \mathbb{R}$ para que $f(x)$ sea una función de densidad discreta.
2. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
3. Calcular la esperanza $\mathrm{E}(X)$.
4. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
5. Calcular la probabilidad de que la cátedra continúe al menos 50 segundos.
6. Calcular la probabilidad de que la cátedra continúe entre 10 y 40 segundos.
7. Calcular la probabilidad de que la cátedra continúe mas de 20 segundos.

Ejercicio 48. La variable aleatoria X es el tiempo que un libro reservado durante dos horas es solicitado en préstamo, suponiendo que X tiene la función de densidad:

$$f_X(x) = \begin{cases} \frac{x}{2} & \text{si } 0 \leq x \leq 2 \\ 0 & \text{cualquier otro caso} \end{cases}.$$

1. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
2. Calcular la esperanza $\mathrm{E}(X)$.
3. Calcular la varianza $\mathrm{Var}(X)$ y la desviación estándar $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de que $X < 1$.
5. Calcular la probabilidad de que $\frac{1}{2} < X < \frac{3}{2}$.
6. Calcular la probabilidad de que $X > \frac{1}{3}$.

Ejercicio 49. El peso de una pastilla de estéreo ajustado a tres gramos es una variable aleatoria X con una función $f(x) = c\left(1 - (x-3)^2\right)$ para $2 \leq x \leq 4$:

1. Calcular el valor de la constante $c \in \mathbb{R}$ para que $f(x)$ sea una función de densidad discreta.
2. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
3. Calcular la esperanza $\mathrm{E}(X)$, la varianza $\mathrm{Var}(X)$ y $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de que la pastilla pese más de 3 gramos.
5. Calcular la probabilidad de que la pastilla pese entre 2.9 y 3.1 gramos.
6. Calcular la probabilidad de que la pastilla pese menos de 2.8 gramos.

Ejercicio 50. La demanda semanal de gas natural es una variable aleatoria X con función $f(x) = c\left(1 - \frac{1}{x^2}\right)$ para $1 \leq x \leq 2$:

1. Calcular el valor de la constante $c \in \mathbb{R}$ para que $f(x)$ sea una función de densidad continua.
2. Presentar analítica y gráficamente las funciones de densidad $f_X(x)$ y de distribución $F_X(x)$.
3. Calcular la esperanza $\mathrm{E}(X)$, la varianza $\mathrm{Var}(X)$ y $\sqrt{\mathrm{Var}(X)}$.
4. Calcular la probabilidad de que la demanda sea mayor que $\frac{3}{2}$.
5. Calcular la probabilidad de que la demanda se encuentre entre $\frac{5}{4}$ y $\frac{7}{4}$.
6. Calcular la probabilidad de que la demanda sea menor que $\frac{3}{2}$.

Capítulo 4

Distribuciones de probabilidad

4.1. Distribuciones discretas

Las características de las funciones de distribución Bernoulli, binomial, geométrica, Pascal, Huygens y Poisson son presentadas.

4.1.1. Distribución Bernoulli

Un ensayo Bernoulli es un experimento aleatorio $\mathscr{E}$ con dos resultados posibles:

1. Éxito con probabilidad $0 < p < 1$.
2. Fracaso con probabilidad complementaria $1 - p$.

Suponiendo un experimento aleatorio $\mathscr{E}$ donde la probabilidad de éxito es p, entonces el espacio muestra Ω consiste en las sucesiones $\Omega = \{0,1\}$ donde $\#\Omega = 2$.

Definición 64 (distribución Bernoulli). Si $0 < p < 1$, entonces la variable aleatoria X se distribuye Bernoulli y la función de densidad es:

$$f_X(x,p) = \begin{cases} p^x(1-p)^{1-x} & \text{si } x = 0,1 \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.1}$$

donde $f_X(x,p)$ es la probabilidad de obtener x éxitos en un ensayo Bernoulli, p es la probabilidad de obtener un éxito y $1-p$ es la probabilidad de obtener un fracaso. La esperanza y la varianza son:

$$\mathrm{E}(X) = \sum_{x=0}^{1} xp^x(1-p)^{1-x} = p. \tag{4.2}$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = p(1-p). \tag{4.3}$$

Proposición 13 (maximizar). Si $X \sim Ber(p)$, entonces la probabilidad de obtener x éxitos alcanza el valor máximo cuando:

$$\underset{p}{\text{maximizar}}\, f_X(x,p) = x = \begin{cases} 0 & \text{si } p < \frac{1}{2} \\ 1 & \text{si } p > \frac{1}{2} \\ 0,1 & \text{si } p = \frac{1}{2} \end{cases}, \tag{4.4}$$

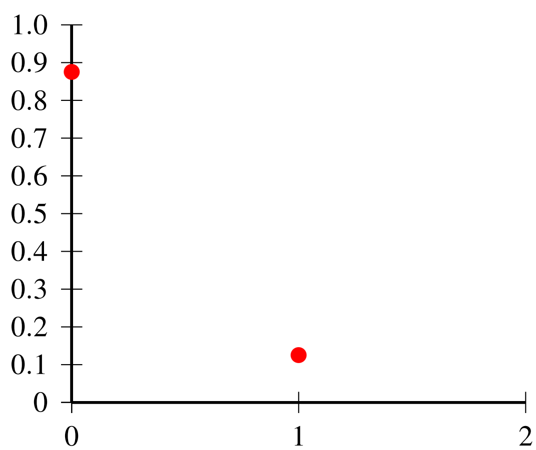
donde $f_X(x,p)$ es la probabilidad de obtener x éxitos y p es la probabilidad de obtener un éxito.

Ejemplo 57 (distribución Bernoulli). Si es realizada una competencia con ocho corredores y se apuesta a un corredor, es decir, la probabilidad de ganar la apuesta es $p = \frac{1}{8}$.

1. Calcular la probabilidad de perder la apuesta, es decir, $\mathscr{P}(X=0)$.
2. Calcular la esperanza matemática $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.
3. Presentar la gráfica de la función de densidad $f_X(x)$.
4. Presentar la gráfica de la función de distribución $F_X(x)$.

Solución. La probabilidad de ganar la apuesta es $p = \frac{1}{8}$, entonces:

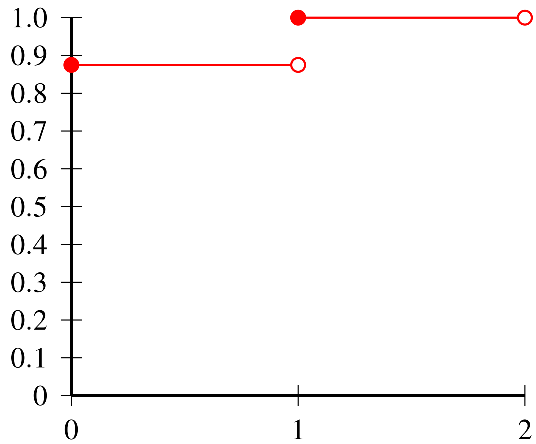
1. $\mathscr{P}(X=0) = 1-p = \frac{7}{8} = 0.875$.
2. $\mathrm{E}(X) = p = \frac{1}{8} = 0.125$ y $\mathrm{Var}(X) = p(1-p) = \frac{1}{8}\left(\frac{7}{8}\right) = \frac{7}{64} = 0.109375$.
3. La gráfica de $f_X(x,p) = \left(\frac{1}{8}\right)^x \left(\frac{7}{8}\right)^{1-x}$ es presentada en la gráfica 4.1.



Gráfica 4.1: $f_X(x,p) = \left(\frac{1}{8}\right)^x \left(\frac{7}{8}\right)^{1-x}$.

Por la proposición 13 (maximizar) y aplicando la ecuación (4.4), la probabilidad de obtener x éxitos en el ensayo Bernoulli alcanza el valor máximo cuando $x = 0$ porque $p < \frac{1}{2}$.

4. La gráfica de $F_X(x,p) = \sum_{k=0}^{x} \left(\frac{1}{8}\right)^k \left(\frac{7}{8}\right)^{1-k}$ es presentada en la gráfica 4.2.



Gráfica 4.2: $F_X(x,p) = \sum_{k=0}^{x} \left(\frac{1}{8}\right)^k \left(\frac{7}{8}\right)^{1-k}$.

Por lo tanto $\mathscr{P}(X=0) = \frac{7}{8}, \mathrm{E}(X) = \frac{1}{8}$ y $\mathrm{Var}(X) = \frac{7}{64}$. $\square$

4.1.2. Distribución binomial

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en n ensayos Bernoulli donde la probabilidad de éxito en cada ensayo es p, entonces el espacio muestra Ω consiste en las n sucesiones $\Omega = \{0,1\}$ donde $\#\Omega = 2^n$.

Sea X una variable aleatoria que indica el número de éxitos en cada muestra, esto es:

$$
\begin{aligned}
X(1_1, 1_2, \ldots, 1_n) &= n, \\
X(0_1, 1_1, \ldots, 1_{n-1}) &= n-1, \\
&\vdots \\
X(0_1, 0_2, \ldots, 0_n) &= 0,
\end{aligned}
$$

entonces, como los ensayos son independientes, por el principio de multiplicación la probabilidad de obtener x éxitos en n ensayos Bernoulli, al margen de ocurrencia, es:

$$\underbrace{pp\cdots p}_{x}\underbrace{(1-p)(1-p)\cdots(1-p)}_{n-x} = p^x(1-p)^{n-x}.$$

El evento C donde resultan x éxitos y $n-x$ fracasos en n ensayos Bernoulli está representado por muestras aleatorias sin orden y sin reemplazo, es decir, por las combinaciones de n ensayos con x éxitos y $n-x$ fracasos.

Definición 65 (combinaciones de n en x elementos). Las combinaciones de n ensayos en x éxitos son el número de arreglos no ordenados de x éxitos sin reemplazo obtenidos en n ensayos donde $0 \leq x \leq n \in \mathbb{N}$, entonces:

$$C(n,x) = \binom{n}{x} = \frac{n!}{x!\,(n-x)!}, \tag{4.5}$$

donde $C(n,x)$ son las combinaciones no ordenadas y sin repetición de n en x éxitos, n es el número de ensayos del espacio muestra Ω y x es el número de éxitos.

Definición 66 (distribución binomial). Si $x = 0,1,2,\ldots,n$ y $0 < p < 1$, entonces la variable aleatoria X se distribuye binomial y la función de densidad es:

$$f_X(x,n,p) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{si } x = 0,1,\ldots,n \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.6}$$

donde $f_X(x,n,p)$ es la probabilidad de obtener x éxitos en n ensayos, p es la probabilidad de obtener un éxito y $1-p$ es la probabilidad de obtener un fracaso. La esperanza y la varianza son:

$$\mathrm{E}(X) = \sum_{x=0}^{n} \binom{n}{x} x p^x (1-p)^{n-x} = np. \tag{4.7}$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = np(1-p). \tag{4.8}$$

Proposición 14 (maximizar). Si $X \sim Bin(n,p)$, entonces la probabilidad de obtener x éxitos en n ensayos alcanza el valor máximo cuando:

$$\underset{n,p}{\mathrm{maximizar}}\, f_X(x,n,p) = \{x \mid x = \lceil (n+1)p - 1 \rceil \wedge x = \lfloor (n+1)p \rfloor\}, \tag{4.9}$$

donde $f_X(x,n,p)$ es la probabilidad de obtener x éxitos en n ensayos, p es la probabilidad de obtener un éxito, $\lceil \cdot \rceil$ es la función entero superior y $\lfloor \cdot \rfloor$ es la función entero inferior.

Ejemplo 58 (distribución binomial). Si se presenta un examen de opción múltiple con las opciones verdadero y falso y un alumno contesta aleatoriamente las diez preguntas, es decir, la probabilidad de acertar es $p = \frac{1}{2}$.

1. Calcular la probabilidad de aprobar el examen, es decir, $\mathscr{P}(X \geq 6)$.
2. Calcular la esperanza matemática $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.
3. Presentar la gráfica de la función de densidad $f_X(x)$.
4. Presentar la gráfica de la función de distribución $F_X(x)$.

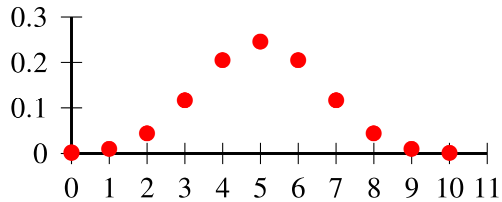
Solución. La probabilidad de acertar la respuesta es $p = \frac{1}{2}$, entonces:

1. $\mathscr{P}(X \geq 6) = 1 - \sum_{x=0}^{5} \binom{10}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x} = \sum_{x=6}^{10} \binom{10}{x} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{10-x}$, entonces:

$$\mathscr{P}(X \geq 6) = 1 - (0.0010 + 0.0098 + 0.0439 + 0.1172 + 0.2051 + 0.2461) = 0.3770.$$
$$\mathscr{P}(X \geq 6) = 0.2051 + 0.1172 + 0.0439 + 0.0098 + 0.0010 = 0.3770.$$

2. $\mathrm{E}(X) = np = 10\left(\frac{1}{2}\right) = 5$ y $\mathrm{Var}(X) = np(1-p) = 10\left(\frac{1}{2}\right)^2 = \frac{5}{2} = 2.5$.

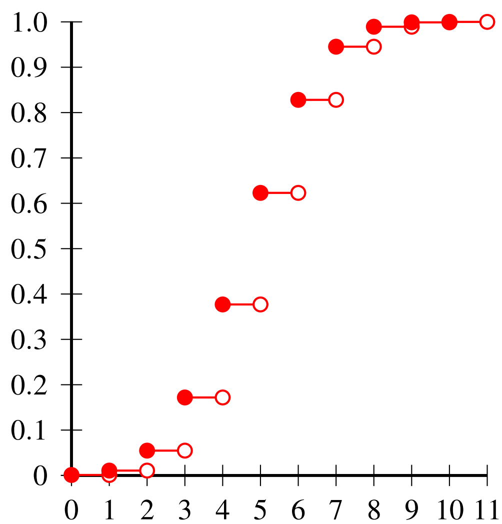
3. La gráfica de $f_X(x,n,p) = \binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}$ es presentada en la gráfica 4.3.



Gráfica 4.3: $f_X(x,n,p) = \binom{10}{x}\left(\frac{1}{2}\right)^x\left(\frac{1}{2}\right)^{10-x}$.

Por la proposición 14 (maximizar) y aplicando la ecuación (4.9), la probabilidad de obtener x éxitos en los diez ensayos Bernoulli alcanza el máximo cuando $x = 5$ porque $\left\lceil \frac{11}{2} - 1 \right\rceil = \left\lfloor \frac{11}{2} \right\rfloor = 5$.

4. La gráfica de $F_X(x,n,p) = \sum_{k=0}^{x}\binom{10}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{10-k}$ es presentada en la gráfica 4.4.



Gráfica 4.4: $F_X(x,n,p) = \sum_{k=0}^{n}\binom{10}{k}\left(\frac{1}{2}\right)^k\left(\frac{1}{2}\right)^{10-k}$.

Por lo tanto $\mathscr{P}(X \geq 6) = 0.377, \mathrm{E}(X) = 5$ y $\mathrm{Var}(X) = \frac{5}{2}$. $\square$

Proposición 15 (distribución binomial). Si $X_1, X_2, \ldots, X_n \sim Ber(p)$, entonces $X = \sum_{k=1}^{n} X_k \sim Bin(n,p)$.

Proposición 16 (distribución binomial). Si $X \sim Bin(n,p)$, entonces:

$$f_X(x,n,p) = \begin{cases} \frac{\Gamma(n+1)p^x(1-p)^{n-x}}{\Gamma(x+1)\Gamma(n-x+1)} & \text{si } x = 0,1,\ldots,n \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.10}$$

donde $f_X(x,n,p)$ es la probabilidad de obtener x éxitos en n ensayos, p es la probabilidad de obtener un éxito, $1-p$ es la probabilidad de obtener un fracaso y $\Gamma(\cdot)$ es la función gama (propuesta por Euler 1738).

4.1.3. Aplicación de la distribución binomial

Suponiendo que el precio subyacente tiene dos valores posibles derivados del estado anterior, entonces el precio subyacente es modelado con una caminata aleatoria y el valor de la opción es calculado a través de la programación dinámica estocástica que por medio del principio de Bellman maximiza el flujo de efectivo en valor presente, cuantificando el riesgo que otorga la cobertura de la opción. Las trayectorias del precio subyacente consideran la inexistencia de oportunidades de arbitraje para los inversionistas, entonces al dividir el tiempo de cobertura T en n intervalos homogéneos existen $n+1$ valores probables del precio subyacente y $n+1$ precios de la opción en la fecha de vencimiento.

Definición 67 (opción). Contrato en el que el comprador, mediante el pago de la prima, adquiere del vendedor el derecho, pero no la obligación, de comprar o vender el bien subyacente al precio pactado en una fecha futura y el vendedor se obliga a vender o comprar, según corresponda, el bien subyacente al precio convenido.

El problema para valuar opciones es resuelto empleando un modelo basado en la secuencia de decisiones que satisface que las decisiones futuras son decisiones óptimas con fundamento en las decisiones precedentes, entonces la decisión presente es óptima y tiene una definición recursiva.

Proposición 17 (activos del portafolio réplica). Sea $V\left(t,M_t\right)$ el valor de la opción en el instante $0\leq t\leq T$ cuando el precio subyacente es M_t, $V\left(t,M_t^a\right)$ el valor de la opción en el instante $0\leq t\leq T$ cuando el precio subyacente aumenta de M_{t-1} a $M_t^a=M_{t-1}a$ y $V\left(t,M_t^d\right)$ el valor de la opción en el instante $0\leq t\leq T$ cuando el precio subyacente disminuye de M_{t-1} a $M_t^d=M_{t-1}d$, entonces la estrategia óptima Δ_t en el instante t es:

$$\Delta_t=\frac{V\left(t,M_t^a\right)-V\left(t,M_t^d\right)}{M_t^a-M_t^d}. \tag{4.11}$$

El portafolio réplica es:

$$\Pi=\begin{cases}\Delta M\exp\left(-r\tau\right) & \text{posición larga}\\ -c\left(0,M_0\right) & \text{posición corta}\end{cases}. \tag{4.12}$$

Por lo tanto, el número de activos que tiene el portafolio réplica para estar libre de riesgo es Δ_t.

Proposición 18 (martingalas). Sea $\mathrm{E}\left(X\right)=\exp\left(\left(i-r\right)\delta\tau\right)=a\pi+d\left(1-\pi\right)$ el valor esperado de los rendimientos y $\mathrm{Var}\left(X\right)=\left(a-d\right)^2\pi\left(1-\pi\right)=\left(u-d\right)\left(a-u\right)$ la varianza de los rendimientos, entonces las probabilidades neutrales al riesgo son:

$$\pi=\frac{u-d}{a-d}\qquad \text{y}\qquad 1-\pi=\frac{a-u}{a-d}, \tag{4.13}$$

donde $\delta=\frac{1}{n}$, $0<\pi<1$, $0<d<1\leq u<a$, $u=\exp\left(\left(i-r\right)\delta\tau\right)$ y $\tau=T-t$.

Proposición 19 (valor del pago contingente). La definición recursiva para n intervalos bajo el supuesto de valuación en un mundo neutral al riesgo es:

$$V\left(\eta,M_\eta^{a^kd^{\eta-k}}\right)=\begin{cases}\left(V\left(\eta+1,M_{\eta+1}^{a^{k+1}d^{\eta-k}}\right)\pi+V\left(\eta+1,M_{\eta+1}^{a^kd^{\eta-k+1}}\right)\left(1-\pi\right)\right)\exp\left(-i\delta\tau\right) & \text{si } 0\leq k\leq\eta<n\\ V\left(\eta,M_\eta^{a^kd^{\eta-k}}\right)=\begin{cases}\left(Ma^kd^{\eta-k}-S\right)_+ & \text{si } 0\leq k\leq\eta=n\\ \left(S-Ma^kd^{\eta-k}\right)_+ & \text{si } 0\leq k\leq\eta=n\end{cases}\end{cases}, \tag{4.14}$$

donde $\delta=\frac{1}{n}$, $0<\pi<1$, $0<d<1\leq u<a$, $u=\exp\left(\left(i-r\right)\delta\tau\right)$ y $\tau=T-t$.

Proposición 20 (fórmula general para valuar opciones europeas). La fórmula general para valuar opciones europeas es:

$$V(0,M_0)=\begin{cases}\left(\sum_{k=0}^{n}\binom{n}{k}\pi^k(1-\pi)^{n-k}V\left(n,M_n^{a^kd^{n-k}}\right)\right)\exp(-iT) & \text{si } 0\leq k\leq n\\ V\left(\eta,M_\eta^{a^kd^{\eta-k}}\right)=\begin{cases}\left(Ma^kd^{\eta-k}-S\right)_+ & \text{si } 0\leq k\leq n\\ \left(S-Ma^kd^{\eta-k}\right)_+ & \text{si } 0\leq k\leq n\end{cases}\end{cases}, \tag{4.15}$$

donde $\delta=\frac{1}{n}$, $0<\pi<1$, $0<d<1\leq u<a$, $u=\exp((i-r)\delta\tau)$ y $\tau=T-t$.

Proposición 21 (parámetros de las probabilidades libres de riesgo). Sean:

$$\mathrm{E}(X)=a\pi+d(1-\pi)=\exp((i-r)\delta\tau)=u. \tag{4.16}$$

$$\mathrm{Var}(X)=(a-d)^2\pi(1-\pi). \tag{4.17}$$

$$\pi=\frac{u-d}{a-d}. \tag{4.18}$$

$$1-\pi=\frac{a-u}{a-d}. \tag{4.19}$$

Sustituyendo las ecuaciones (4.18) y (4.19) en la ecuación (4.17), la varianza proporcional en el tiempo es:

$$\sigma^2\delta\tau=(u-d)(a-u). \tag{4.20}$$

Resolviendo el sistema de ecuaciones (4.16) y (4.20) suponiendo que $\pi=\frac{1}{2}$ como propuso Bachelier (1900):

$$a=\exp((i-r)\delta\tau)+\sigma\sqrt{\delta\tau}. \tag{4.21}$$

$$d=\exp((i-r)\delta\tau)-\sigma\sqrt{\delta\tau}. \tag{4.22}$$

Resolviendo el sistema de ecuaciones (4.16) y (4.20) suponiendo que $ad=1$ como propusieron Cox, Ross y Rubinstein (1979):

$$a=\exp\left(\sigma\sqrt{\delta\tau}\right). \tag{4.23}$$

$$d=\exp\left(-\sigma\sqrt{\delta\tau}\right). \tag{4.24}$$

Sustituyendo $ad=1$ en la ecuación (4.20) y despejando la variable a como es propuesto por Climent Hernández (2014):

$$a=\frac{\sqrt{B^2-4}-B}{2}, \tag{4.25}$$

$$d=-\frac{\sqrt{B^2-4}+B}{2}, \tag{4.26}$$

donde:

$$B=-\left(\sigma^2\delta\tau+u^2+1\right)\exp(-(i-r)\delta\tau). \tag{4.27}$$

con $\delta=\frac{1}{n}$, $0<\pi<1$, $0<d<1\leq u<a$, $u=\exp((i-r)\delta\tau)$ y $\tau=T-t$.

Ejemplo 59 (valuación de una opción europea de compra). Suponiendo que el precio de un activo subyacente es $M = 496$, la volatilidad anual es $\sigma = 0.15$, la tasa de interés libre de riesgo nacional instantánea anual es $i = 0.07$, la tasa de dividendos es nula $r = 0$, el precio de liquidación es $S = M\exp(iT) = 498.90$ y el tiempo de vigencia es un mes $T = \frac{1}{12}$. Calcular el precio de la opción europea de compra sobre el activo subyacente:

1. Con el modelo propuesto por Cox, Ross y Rubinstein (1979) con tres periodos.
2. Con el modelo propuesto por Climent Hernández (2014) con tres periodos.

Solución. Por la proposición 21 (parámetros de las martingalas):

1. Con el modelo propuesto por Cox, Ross y Rubinstein (1979):

$$
\begin{aligned}
\delta &= \frac{1}{n} = \frac{1}{3} \quad \text{y} \quad \tau = \frac{1}{12},\\
u &= \exp\left((i-r)\,\delta\tau\right) = \exp\left(0.07\left(\frac{1}{3}\right)\left(\frac{1}{12}\right)\right) = 1.001946,\\
a &= \exp\left(\sigma\sqrt{\delta\tau}\right) = \exp\left(0.15\sqrt{\left(\frac{1}{3}\right)\left(\frac{1}{12}\right)}\right) = 1.025315,\\
d &= \exp\left(-\sigma\sqrt{\delta\tau}\right) = \exp\left(-0.15\sqrt{\left(\frac{1}{3}\right)\left(\frac{1}{12}\right)}\right) = 0.975310,\\
\pi &= \frac{u-d}{a-d} = \frac{1.001946-0.975310}{1.025315-0.975310} = 0.532673,\\
1-\pi &= 1-0.532673 = 0.467327.
\end{aligned}
$$

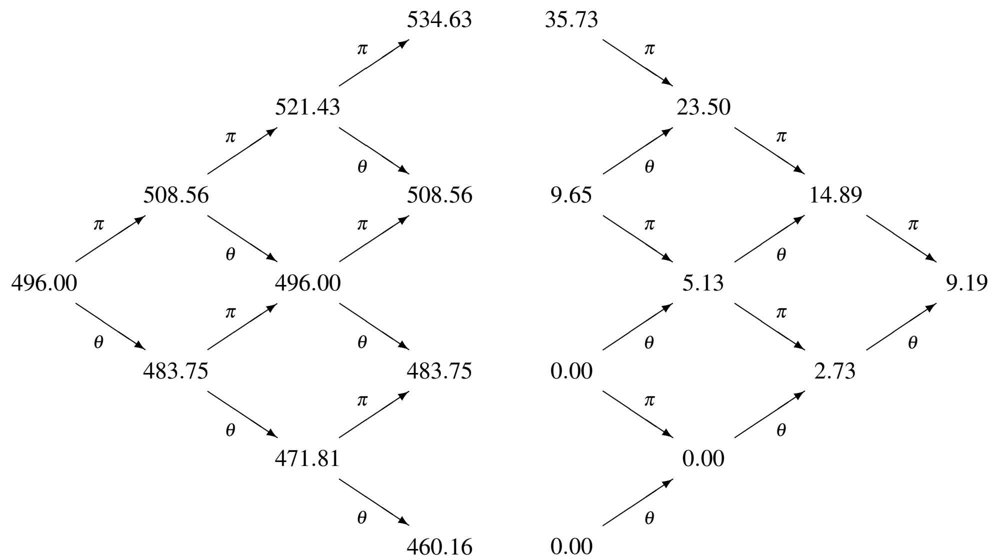
Por la proposición 20 (fórmula general para valuar opciones europeas) y sustituyendo en la ecuación (4.15):

$$
\begin{aligned}
c(0,M_0) &= \left(\sum_{k=0}^{3}\binom{3}{k}\pi^k(1-\pi)^{n-k}V\left(n,M_n^{a^kd^{n-k}}\right)\right)\exp(-iT).\\
c(0,M_0) &= \left(0+3\left(0+0.53^2 0.47^1(9.65)\right)+0.53^3(35.73)\right)\exp\left(-\frac{0.07}{12}\right) = 9.19.
\end{aligned}
$$

Por lo tanto, el precio de la opción europea de compra sobre el activo subyacente es $c(0,M_0) = 9.19$.

Por la proposición 19 (valor del pago contingente) y sustituyendo en la ecuación (4.14) es obtenido el valor de la opción europea de compra sobre el activo subyacente con el árbol binomial que modela el precio subyacente hacia el futuro y aplicando el principio de inducción regresiva.

El árbol binomial es presentado en la gráfica 4.5.



Gráfica 4.5: Árbol binomial de Cox, Ross y Rubinstein (1979).

Por lo tanto, el precio de la opción europea de compra sobre el activo subyacente es $c\,(0, M_0) = 9.19$.

2. Con el modelo propuesto por Climent Hernández (2014):

$$\delta = \frac{1}{3}, \quad \tau = \frac{1}{12}, \quad \text{y} \quad u = 1.001946,$$

$$B = -\left(\sigma^2\delta\tau + u^2 + 1\right)\exp\left(-i\delta\tau\right) = -\left(\frac{0.15^2}{36} + 1.0019^2 + 1\right)\exp\left(-\frac{0.07}{36}\right) = -2.0006.$$

$$a = \frac{\sqrt{B^2-4} - B}{2} = \frac{\sqrt{(-2.000638)^2 - 4} + 2.000638}{2} = 1.025367.$$

$$d = -\frac{\sqrt{B^2-4} + B}{2} = -\frac{\sqrt{(-2.00063)^2 - 4} - 2.00063}{2} = 0.975261.$$

$$\pi = \frac{u-d}{a-d} = \frac{1.001946 - 0.975261}{1.025367 - 0.975261} = 0.532582.$$

$$1 - \pi = 1 - 0.532673 = 0.467418.$$

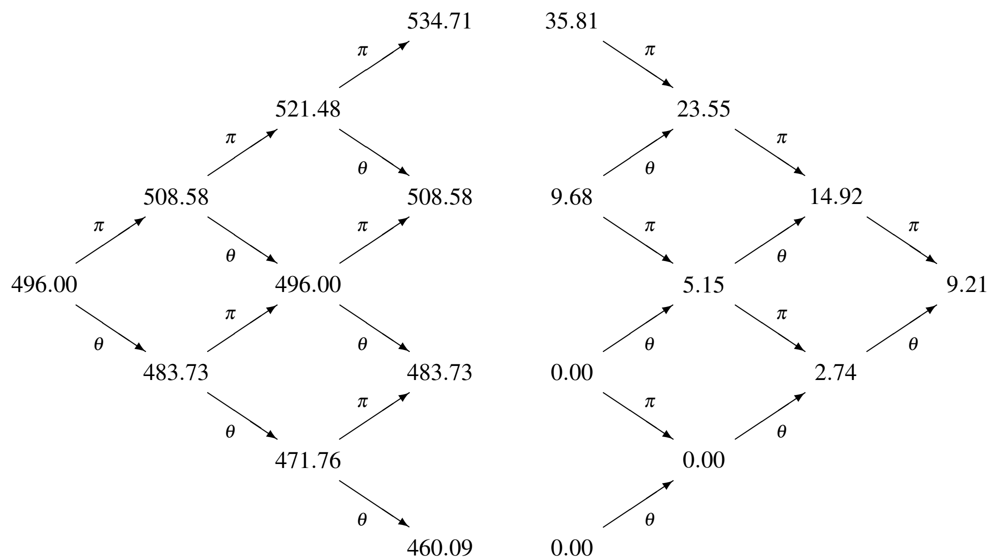
Por la proposición 20 (fórmula general para valuar opciones europeas) y sustituyendo en la ecuación (4.15):

$$c\,(0, M_0) = \left(\sum_{k=0}^{3}\binom{3}{k}\pi^k\,(1-\pi)^{n-k}\,V\left(n, M_n^{a^k d^{n-k}}\right)\right)\exp\left(-iT\right).$$

$$c\,(0, M_0) = \left(0 + 3\left(0 + 0.53^2 0.47^1\,(9.68)\right) + 0.53^3\,(35.81)\right)\exp\left(-\frac{0.07}{12}\right) = 9.21.$$

Por lo tanto, el precio de la opción europea de compra sobre el activo subyacente es $c(0, M_0) = 9.21$.

Por la proposición 19 (valor del pago contingente) y sustituyendo en la ecuación (4.14) es obtenido el valor de la opción europea de compra sobre el activo subyacente con el árbol binomial que modela el precio subyacente hacia el futuro y aplicando el principio de inducción regresiva. El árbol binomial es presentado en la gráfica 4.6.



Gráfica 4.6: Árbol binomial de Climent Hernández (2014).

Por lo tanto, el precio de la opción europea de compra sobre el activo subyacente es $c(0, M_0) = 9.21$ y el modelo de Cox, Ross y Rubinstein (1979) subestima el precio de la opción europea de compra. $\square$

4.1.4. Distribución geométrica

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en una sucesión infinita de ensayos independientes y con distribución Bernoulli, donde la probabilidad de éxito en cada ensayo es p, entonces el espacio muestra Ω consiste en las sucesiones infinitas $\Omega = \{0, 1\}$ donde $\#\Omega = \lim_{n\to\infty} 2^n = \infty$.

Sea X una variable aleatoria que indica el número de fracasos antes de obtener el primer éxito, es decir:

$$\begin{aligned}
X(1_1, 0_2, 1_3, 0_4, 1_5, \ldots) &= 0, \\
X(0_1, 1_2, 0_3, 1_4, 1_5, \ldots) &= 1, \\
X(0_1, 0_2, 1_3, 0_4, 0_5, \ldots) &= 2, \\
X(0_1, 0_2, 0_3, 1_4, 1_5, \ldots) &= 3, \\
&\vdots
\end{aligned}$$

Entonces como los ensayos son independientes, por el principio de multiplicación, la probabilidad de obtener x fracasos antes de obtener el primer éxito es:

$$\underbrace{(1-p)(1-p)\cdots(1-p)}_{x}p = (1-p)^x p.$$

Definición 68 (distribución geométrica). Si $x = 0,1,2,3,\ldots$ y $0 < p < 1$, entonces la variable aleatoria X se distribuye geométrica y la función de densidad es:

$$f_X(x,p) = \begin{cases} p(1-p)^x & \text{si } x = 0,1,2,3,\ldots \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.28}$$

donde $f_X(x,p)$ es la probabilidad de obtener x fracasos antes de obtener el primer éxito, p es la probabilidad de obtener un éxito y $1-p$ es la probabilidad de obtener un fracaso. La esperanza y la varianza son:

$$\mathrm{E}(X) = \sum_{x=0}^{\infty} xp(1-p)^x = \frac{1-p}{p}. \tag{4.29}$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = \frac{1-p}{p^2}. \tag{4.30}$$

Proposición 22 (maximizar). Si $X \sim Geo(p)$, entonces la probabilidad de obtener x fracasos antes de obtener el primer éxito alcanza el valor máximo cuando:

$$\underset{p}{\text{maximizar}}\, f_X(x,p) = \{x \mid x = 0\}, \tag{4.31}$$

donde $f_X(x,p)$ es la probabilidad de obtener x fracasos antes de obtener el primer éxito y p es la probabilidad de obtener un éxito.

La distribución geométrica es utilizada en la distribución de los tiempos de espera porque si los ensayos son realizados en periodos homogéneos, entonces la distribución geométrica proporciona el tiempo transcurrido hasta el primer éxito.

Ejemplo 60 (distribución geométrica). Si son repartidas 5 cartas de una baraja de 52 cartas cada minuto y la probabilidad de obtener una tercia es $p = \frac{13C(4,3)C(12,2)C^2(4,1)}{C(52,5)} = \frac{88}{4{,}165} \approx 0.021128$.

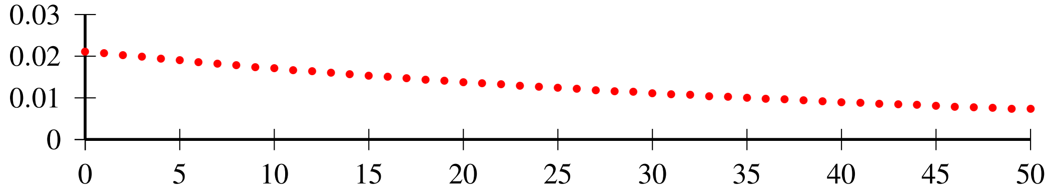
1. Calcular la probabilidad de tener a lo más diez fracasos antes de obtener la tercia, es decir, $\mathscr{P}(X \leq 10)$.
2. Calcular la esperanza matemática $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.
3. Presentar la gráfica de la función de densidad $f_X(x)$.
4. Presentar la gráfica de la función de distribución $F_X(x)$.

Solución. La probabilidad de obtener una tercia (tres cartas con el mismo número) es $p = \frac{88}{4{,}165}$, entonces:

1. $\mathscr{P}(X \leq 10) = \sum\limits_{k=0}^{10} \frac{88}{4{,}165}\left(1 - \frac{88}{4{,}165}\right)^k = 0.2094.$
2. $\mathrm{E}(X) = \frac{1-p}{p} = \frac{1-\frac{88}{4{,}165}}{\frac{88}{4{,}165}} = 46.33$ y $\mathrm{Var}(X) = \frac{1-p}{p^2} = \frac{1-\frac{88}{4{,}165}}{\left(\frac{88}{4{,}165}\right)^2} = 2{,}192.76.$

   El tiempo esperado para obtener la tercia es $\mathrm{E}(X+1) = \mathrm{E}(X) + 1 = \frac{1-p}{p} + 1 = 47.33$, es decir, 47 minutos y $19.77\overline{27}$ segundos.

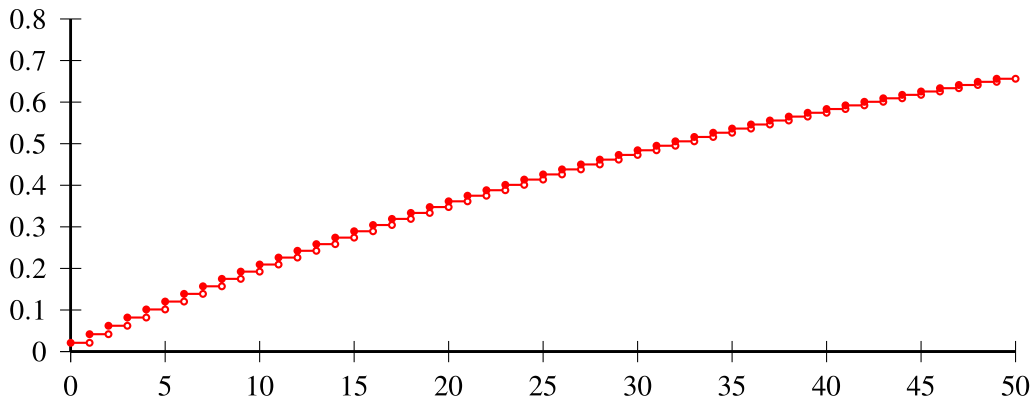
3. La gráfica de $f_X(x,p) = \frac{88}{4{,}165}\left(1-\frac{88}{4{,}165}\right)^x$ es presentada en la gráfica 4.7.



Gráfica 4.7: $f_X(x,p) = \frac{88}{4,165}\left(1-\frac{88}{4,165}\right)^x.$

Por la proposición 22 (maximizar) y aplicando la ecuación (4.31), la probabilidad de obtener x fracasos alcanza el máximo cuando $x = 0$.

4. La gráfica de $F_X(x,p) = \sum_{k=0}^{x} \frac{88}{4{,}165}\left(1-\frac{88}{4{,}165}\right)^x$ es presentada en la gráfica 4.8.



Gráfica 4.8: $F_X(x,p) = \sum_{k=0}^{x} \frac{88}{4,165}\left(1-\frac{88}{4,165}\right)^x.$

Por lo tanto $\mathscr{P}(X \leq 10) = 0.2094, \mathrm{E}(X) = 46.33$ y $\mathrm{Var}(X) = 2,192.76$. $\square$

Proposición 23 (distribución geométrica). Si $X_1, X_2, \ldots \sim Ber(p)$, entonces $X = \sum_{k=1}^{\infty} X_k \sim Geo(p)$.

4.1.5. Distribución Pascal

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en una sucesión infinita de ensayos independientes y con distribución Bernoulli donde la probabilidad de éxito en cada ensayo es p, entonces el espacio muestra Ω consiste en las sucesiones infinitas $\Omega = \{0,1\}$ donde $\#\Omega = \lim_{n\to\infty} 2^n = \infty$, por lo tanto, como los ensayos son independientes, por el principio de multiplicación, la probabilidad de obtener x fracasos antes de obtener el r–ésimo éxito es:

$$\underbrace{(1-p)(1-p)\cdots(1-p)}_{x}\underbrace{pp\cdots p}_{r} = (1-p)^x p^r.$$

Definición 69 (distribución Pascal). Si $r \in \mathbb{N}$ y $x = 0, 1, 2, 3, \ldots$ y $0 < p < 1$, entonces la variable aleatoria X se distribuye Pascal y la función de densidad es:

$$f_X(x, r, p) = \begin{cases} \binom{r+x-1}{x} p^r (1-p)^x & \text{si } x = 0, 1, 2, 3, \ldots \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.32}$$

donde $f_X(x, r, p)$ es la probabilidad de obtener x fracasos antes de obtener el r–ésimo éxito, p es la probabilidad de obtener un éxito y $1 - p$ es la probabilidad de obtener un fracaso. La esperanza y la varianza son:

$$\mathrm{E}(X) = \sum_{x=0}^{\infty} \binom{r+x-1}{x} x p^r (1-p)^x = \frac{r(1-p)}{p}. \tag{4.33}$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = \frac{r(1-p)}{p^2}. \tag{4.34}$$

Proposición 24 (maximizar). Si $r \in \mathbb{N}$ y $X \sim \mathit{Pascal}(r, p)$, entonces la probabilidad de obtener x fracasos antes de obtener el r–ésimo éxito alcanza el valor máximo cuando:

$$\underset{r,p}{\text{maximizar}}\, f_X(x, p) = \left\{ x \mid x = \left\lceil \frac{(r-1)(1-p)}{p} - 1 \right\rceil \wedge x = \left\lfloor \frac{(r-1)(1-p)}{p} \right\rfloor \right\}, \tag{4.35}$$

donde $f_X(x, r, p)$ es la probabilidad de obtener x fracasos antes de obtener el r–ésimo éxito y p representa la probabilidad de obtener un éxito.

Nota 10 (distribución Pascal). La distribución binomial negativa fue estudiada por De Montmort (1708) y descrita por Pascal (1654). La distribución geométrica es un caso particular de la distribución Pascal.

Ejemplo 61 (distribución Pascal). El número de anotaciones (*touch downs*) de seis puntos de 1967 en la NCAA[1] está descrito por una distribución Pascal con probabilidad de éxito $p = 0.6852$.

1. Calcular las probabilidades de tener $0 \leq x \leq 11$ fracasos antes de obtener $r = 6$ anotaciones.
2. Calcular la esperanza matemática $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.
3. Presentar la gráfica de la función de densidad $f_X(x)$.
4. Presentar la gráfica de la función de distribución $F_X(x)$.

[1] Acrónimo de National Collegiate Athletic Association.

Solución. La probabilidad de anotar es $p = 0.6852$ y el número de anotaciones necesario es $r = 6$, entonces:

1. Las probabilidades de tener $x = 0, 1, 2, \ldots, 11$ fracasos antes de obtener seis anotaciones son calculadas con la ecuación (4.32):

$$\begin{aligned}
\mathscr{P}(X = 0) &= \binom{5}{0} 0.6852^6 0.3148^0 = 0.1035,\\
\mathscr{P}(X = 1) &= \binom{6}{1} 0.6852^6 0.3148^1 = 0.1955,\\
\mathscr{P}(X = 2) &= \binom{7}{2} 0.6852^6 0.3148^2 = 0.2154,\\
&\vdots\\
\mathscr{P}(X = 11) &= \binom{16}{11} 0.6852^6 0.3148^{11} = 0.0014.
\end{aligned}$$

Las probabilidades acumuladas de tener $0 \leq x \leq 11$ fracasos antes de obtener seis anotaciones son:

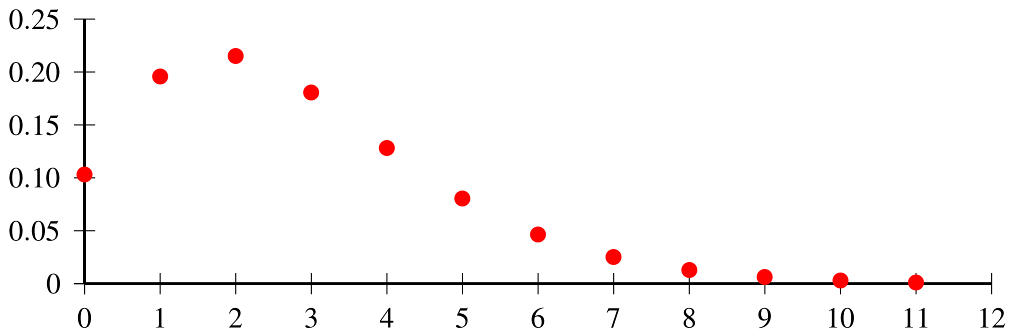
$$\begin{aligned}
\mathscr{P}(X \leq 0) &= \sum_{k=0}^{0} \binom{5}{0} 0.6852^6 0.3148^0 = 0.1035,\\
\mathscr{P}(X \leq 1) &= \sum_{k=0}^{1} \binom{6}{1} 0.6852^6 0.3148^1 = 0.2990,\\
\mathscr{P}(X \leq 2) &= \sum_{k=0}^{2} \binom{7}{2} 0.6852^6 0.3148^2 = 0.5144,\\
&\vdots\\
\mathscr{P}(X \leq 11) &= \sum_{k=0}^{11} \binom{16}{11} 0.6852^6 0.3148^{11} = 0.9989.
\end{aligned}$$

Las probabilidades son presentadas en la tabla 4.1.

Tabla 4.1: Probabilidades de obtener x fracasos antes de anotar seis veces.

x	$\mathscr{P}(X = x)$	$\mathscr{P}(X \leq x)$
0	0.1035	0.1035
1	0.1955	0.2990
2	0.2154	0.5144
3	0.1808	0.6952
4	0.1280	0.8233
5	0.0806	0.9039
6	0.0465	0.9504
7	0.0251	0.9755
8	0.0128	0.9883
9	0.0063	0.9946
10	0.0030	0.9976
11	0.0014	0.9989

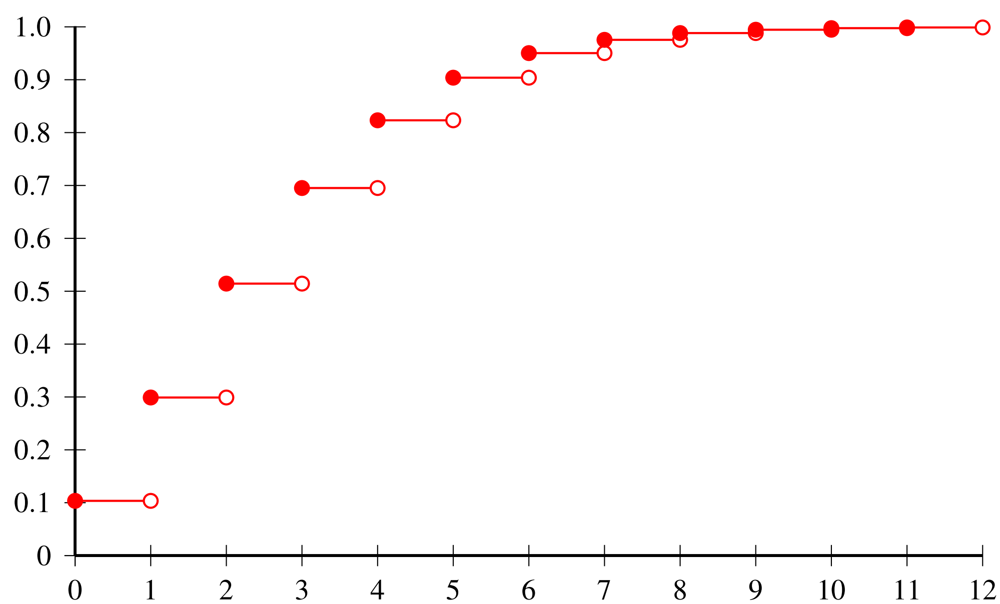
2. $\mathrm{E}(X) = \frac{r(1-p)}{p} = \frac{6(0.3148)}{0.6852} = 2.76$ y $\mathrm{Var}(X) = \frac{r(1-p)}{p^2} = \frac{6(0.3148)}{0.6852^2} = 4.02$.
3. La gráfica de $f_X(x,r,p) = 0.6852^6(0.3148)^x$ es presentada en la gráfica 4.9.



Gráfica 4.9: $f_X(x,r,p) = 0.6852^6(0.3148)^x$.

Por la proposición 24 (maximizar) y aplicando la ecuación (4.35), la probabilidad de obtener x fracasos alcanza el máximo cuando $x = 2$ porque $\left\lceil \frac{6(0.3148)}{0.6852} - 1 \right\rceil = \left\lfloor \frac{6(0.3148)}{0.6852} \right\rfloor = 2$.

4. La gráfica de $F_X(x,p) = \sum_{k=0}^{x} \frac{88}{4{,}165}\left(1 - \frac{88}{4{,}165}\right)^x$ es presentada en la gráfica 4.10.



Gráfica 4.10: $F_X(x,p) = \sum_{k=0}^{x} \frac{88}{4,165}\left(1 - \frac{88}{4,165}\right)^x$.

Por lo tanto $\mathrm{E}(X) = 2.76$ y $\mathrm{Var}(X) = 4.02$. □

Proposición 25 (distribución Pascal). Si $r \in \mathbb{N}$ y $X_1, X_2, \ldots \sim Geo(p)$, entonces $X = \sum_{k=1}^{\infty} X_k \sim Pascal(r,p)$.

Proposición 26 (distribución Pascal). Si $r = 1$, entonces $Pascal(1, p) = Geo(p)$.

Proposición 27 (distribución Pascal). Si $r \in \mathbb{R}_{++}$, entonces:

$$f_X(x, r, p) = \begin{cases} \frac{\Gamma(r+x)p^r(1-p)^x}{x!\Gamma(r)} & \text{si } x = 0, 1, 2, 3, \dots \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.36}$$

donde $f_X(x, r, p)$ es la probabilidad de obtener x fracasos antes de obtener el r–ésimo éxito, $\Gamma(\cdot)$ es la función gama (propuesta por Euler 1738) y p es la probabilidad de obtener un éxito.

Proposición 28 (distribución Pascal). Si $x \in \mathbb{R}_+$ y $r \in \mathbb{R}_{++}$, entonces:

$$f_X(x, r, p) = \begin{cases} \frac{\Gamma(r+x)p^r(1-p)^x}{\Gamma(x+1)\Gamma(r)} & \text{si } x \geq 0 \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.37}$$

donde $f_X(x, r, p)$ es la probabilidad de obtener x fracasos antes de obtener el r–ésimo éxito, $\Gamma(\cdot)$ es la función gama (propuesta por Euler 1738) y p es la probabilidad de obtener un éxito.

4.1.6. Distribución Huygens

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en una sucesión finita de N ensayos independientes y con distribución Bernoulli, donde la probabilidad de obtener uno de los k objetos en cada ensayo es $p = \frac{k}{N}$ y la probabilidad complementaria es $1 - p = \frac{N-k}{N}$, por lo tanto, si se toma una muestra aleatoria no ordenada y sin reemplazo de tamaño $n \leq N$, entonces el espacio muestra Ω consiste en las sucesiones finitas $\Omega = \{0, 1\}$ de tamaño n donde $\#\Omega = C(N, n)$.

Sea X una variable aleatoria que indica el número de los k objetos en cada muestra de tamaño n, esto es:

$$\begin{aligned} X(0_1, 0_2, 0_3, \dots, 0_{n-1}, 0_n) &= 0, \\ X(1_1, 0_2, 0_3, \dots, 0_{n-1}, 0_n) &= 1, \\ &\vdots \\ X(1_1, 1_2, 1_3, \dots, 1_{n-1}, 0_n) &= n-1, \\ X(1_1, 1_2, 1_3, \dots, 1_{n-1}, 1_n) &= n, \end{aligned}$$

entonces como los ensayos son independientes y por el principio de multiplicación, la probabilidad de obtener x de los k objetos, $n - x$ de los $N - k$ objetos y n de los N ensayos Bernoulli, al margen de ocurrencia es:

$$\underbrace{\underbrace{pp\cdots p}_{k}\underbrace{qq\cdots q}_{N-k}}_{N} = p^k q^{N-k},$$

$$\underbrace{\underbrace{\underbrace{pp\cdots p}_{x}\underbrace{qq\cdots q}_{n-x}}_{n}\underbrace{\underbrace{pp\cdots p}_{k-x}\ \underbrace{qq\cdots q}_{N-n-k+x}}_{N-n}}_{N} = p^x q^{n-x} p^{k-x} q^{N-n-k+x} = p^k q^{N-k}.$$

La probabilidad de obtener x de los k objetos y $n-x$ de los $N-k$ objetos de una muestra aleatoria no ordenada y sin reemplazo de tamaño n de los N ensayos Bernoulli es:

$$\underbrace{\underbrace{pp\cdots p}_{k}\underbrace{qq\cdots q}_{N-k}}_{N}=\underbrace{\underbrace{\underbrace{pp\cdots p}_{x}\underbrace{qq\cdots q}_{n-x}}_{n}\underbrace{\underbrace{pp\cdots p}_{k-x}\ \underbrace{qq\cdots q}_{N-n-k+x}}_{N-n}}_{N}=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},$$

donde $C(k,x)$ es el número de arreglos no ordenados y sin repetición de x elementos formados a partir de k elementos, $C(N-k,n-x)$ es el número de arreglos no ordenados y sin repetición de $n-x$ elementos formados a partir de $N-k$ elementos y $C(N,n)$ es el número de arreglos no ordenados y sin repetición de una muestra de tamaño n tomada a partir de N objetos (análisis realizado por Huygens 1657a).

Definición 70 (distribución Huygens). Si $N,k,n\in\mathbb{N}$ y $x=0,1,2,3,\ldots,n$, entonces la variable aleatoria X se distribuye Huygens y la función de densidad es:

$$f_X(x,N,k,n)=\begin{cases}\dfrac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} & \text{si } x=0,1,2,\ldots,n\\ 0 & \text{cualquier otro caso}\end{cases},\tag{4.38}$$

donde $f_X(x,N,k,n)$ es la probabilidad de obtener x de k objetos de una muestra aleatoria de tamaño n tomada de una población de N objetos. La esperanza y la varianza son:

$$\mathrm{E}(X)=\sum_{x=0}^{n}\frac{x\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}=\frac{nk}{N}.\tag{4.39}$$

$$\mathrm{Var}(X)=\mathrm{E}\left(X^2\right)-\mathrm{E}^2(X)=\frac{nk}{N}\left(\frac{N-k}{N}\right)\left(\frac{N-n}{N-1}\right).\tag{4.40}$$

Proposición 29 (maximizar). Si $r\in\mathbb{N}$ y $X\sim Huygens(N,k,n)$, entonces la probabilidad de obtener x de k objetos de una muestra de tamaño n tomada de una población de N objetos alcanza el valor máximo cuando:

$$\underset{N,k,n}{\text{maximizar}}\, f_X(x,N,k,n)=\left\{x\mid x=\left\lceil\frac{(n+1)(k+1)}{N+2}-1\right\rceil\wedge x=\left\lfloor\frac{(n+1)(k+1)}{N+2}\right\rfloor\right\},\tag{4.41}$$

donde $f_X(x,N,k,n)$ es la probabilidad de obtener x de k objetos de una muestra aleatoria de tamaño n tomada de una población de N objetos.

Nota 11 (distribución Huygens). La distribución Huygens (1657a) es conocida como hipergeométrica.

Ejemplo 62 (distribución Huygens). Una tienda tiene 500 artículos, 10 de estos artículos no tienen marcado el precio y un cliente ha seleccionado 50 artículos para comprar.

1. Calcular la probabilidad de seleccionar al menos un artículo sin la marca del precio, es decir, $\mathscr{P}(X \geq 1)$.
2. Calcular la esperanza matemática $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.
3. Presentar la gráfica de la función de densidad $f_X(x)$.
4. Presentar la gráfica de la la función de distribución $F_X(x)$.

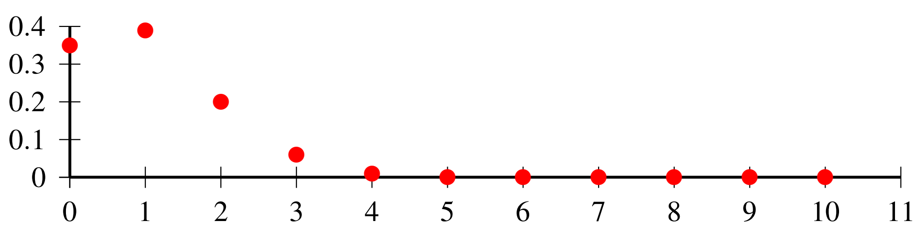
Solución. La probabilidad de seleccionar ningún artículo sin la marca del precio es $\mathscr{P}(X=0)$, sustituyendo en la ecuación (4.38) se tiene que $\mathscr{P}(X=0) = \dfrac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = \dfrac{\binom{10}{0}\binom{490}{50}}{\binom{500}{50}} = 0.3452$, entonces:

1. La probabilidad de seleccionar al menos un artículo sin la marca del precio es:

$$\mathscr{P}(X \geq 1) = \sum_{x=1}^{10} \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}} = 1 - \mathscr{P}(X=0) = 1 - 0.3452 = 0.6548.$$

2. $\mathrm{E}(X) = \frac{nk}{N} = \frac{50(10)}{500} = 1$ y $\mathrm{Var}(X) = \frac{nk}{N}\left(\frac{N-k}{N}\right)\left(\frac{N-n}{N-1}\right) = \frac{20(10)}{500}\left(\frac{490}{500}\right)\left(\frac{450}{499}\right) = 0.8838$.

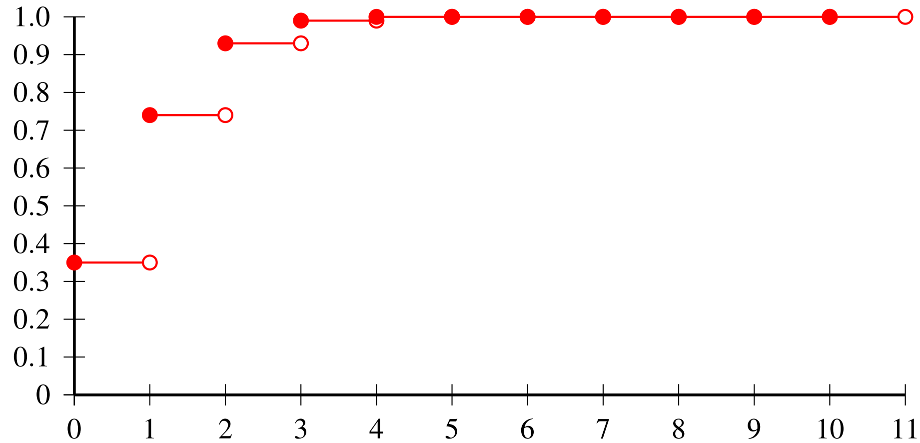
3. La gráfica de $f_X(x,N,k,n) = \dfrac{\binom{10}{x}\binom{490}{50-x}}{\binom{500}{50}}$ es presentada en la gráfica 4.11.



Gráfica 4.11: $\dfrac{\binom{10}{x}\binom{490}{50-x}}{\binom{500}{50}}$.

Por la proposición 29 (maximizar) y aplicando la ecuación (4.41), la probabilidad de obtener x fracasos alcanza el máximo cuando $x=1$ porque $\left\lceil \frac{51(11)}{502} - 1 \right\rceil = \left\lfloor \frac{51(11)}{502} \right\rfloor = 1$.

4. La gráfica de $F_X(x,N,k,n) = \sum_{k=0}^{x} \frac{\binom{10}{x}\binom{490}{50-x}}{\binom{500}{50}}$ es presentada en la gráfica 4.12.



Gráfica 4.12: $F_X(x,p) = \sum_{k=0}^{x} \frac{\binom{10}{x}\binom{490}{50-x}}{\binom{500}{50}}$.

Por lo tanto $\mathscr{P}(X \geq 1) = 0.6548, \mathrm{E}(X) = 1$ y $\mathrm{Var}(X) = 0.8838$. $\square$

Proposición 30 (distribución Huygens). Si $\forall N,k,n \in \mathbb{N}$ y $X \sim Huygens(N,k,n)$ donde $p = \frac{k}{N}$ es constante, entonces $\lim_{N\to\infty} Huygens(N,k,n) = Bin(n,p)$.

4.1.7. Distribución Poisson

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en observar el número de ocurrencias de un evento durante un periodo dado $0 \leq t \leq T$, por ejemplo, el número de alumnos que asisten a clase, el número de clientes que llegan a un cajero automático durante una hora o el número de reclamaciones que llegan a una aseguradora en un mes (distribución propuesta por De Moivre 1710 y aplicada por Poisson 1733).

La variable aleatoria X indica el número de ocurrencias de un evento durante un periodo dado $0 \leq t \leq T$ y toma los valores $x = 0,1,2,\ldots$. Suponiendo que la tasa de ocurrencia promedio del evento por unidad de tiempo es $\lambda > 0$, entonces la variable aleatoria X presenta una distribución Poisson, $X \sim Poisson(\lambda)$.

Definición 71 (distribución Poisson). Sea $\lambda > 0$, entonces la variable aleatoria X se distribuye Poisson y la función de densidad es:

$$f_X(x,\lambda) = \begin{cases} \frac{\lambda^x \exp(-\lambda)}{x!} & \text{si } x = 0,1,\ldots \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.42}$$

donde $f_X(x,\lambda)$ es la probabilidad de obtener x ocurrencias en el periodo $0 \leq t \leq T$ y λ es la tasa de ocurrencia promedio del evento por unidad de tiempo.

La esperanza y la varianza son:

$$\mathrm{E}\left(X\right) = \sum_{x=0}^{\infty} \frac{x\lambda^x \exp\left(-\lambda\right)}{x!} = \lambda. \tag{4.43}$$

$$\mathrm{Var}\left(X\right) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2\left(X\right) = \lambda. \tag{4.44}$$

La distribución Poisson[2] es aplicada al estudio de eventos que ocurren aleatoriamente dado un instante en el tiempo. Esta distribución fue aplicada por Poisson (1837) como el límite de la distribución binomial.

Proposición 31 (maximizar). Si $X \sim Poisson\left(\lambda\right)$, entonces la probabilidad de que ocurran x eventos durante el periodo $0 \leq t \leq T$ alcanza el valor máximo cuando:

$$\underset{\lambda}{\text{maximizar}}\, f_X\left(x, \lambda\right) = \left\{x \mid x = \lceil \lambda - 1 \rceil \wedge x = \lfloor \lambda \rfloor\right\}, \tag{4.45}$$

donde $f_X\left(x, \lambda\right)$ es la probabilidad de que ocurran x eventos durante el periodo $0 \leq t \leq T$ y λ es la tasa de ocurrencia promedio del evento por unidad de tiempo.

Teorema 5 (Poisson). Si $X \sim Bin\left(n, p\right)$ donde $0 < p < 1$ es una medida de probabilidad y $\lambda = np > 0$ es una constante, entonces $\forall x = 0, 1, 2, \ldots,$ se tiene que:

$$\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \frac{\lambda^x \exp\left(-\lambda\right)}{x!}. \tag{4.46}$$

Demostración. Si $X \sim Bin\left(n, p\right)$, entonces $\lim\limits_{n\to\infty} Bin\left(n, p\right) = Poisson\left(\lambda\right)$:

$$\begin{aligned}
\lim_{n\to\infty} Bin\left(n, p\right) &= \lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} = \lim_{n\to\infty} \frac{n!}{x!\left(n-x\right)!}\left(\frac{\lambda}{n}\right)^x \left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \prod_{k=0}^{x-1} \left(\frac{n-k}{n^x}\right)\left(1 - \frac{\lambda}{n}\right)^{n-x} \\
&= \frac{\lambda^x}{x!} \lim_{n\to\infty} \underbrace{\prod_{k=0}^{x-1} \left(1 - \frac{k}{n}\right)}_{\to 1} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{-x}}_{\to 1} \underbrace{\left(1 - \frac{\lambda}{n}\right)^{n}}_{\to \exp\left(-\lambda\right)} \\
&= \frac{\lambda^x \exp\left(-\lambda\right)}{x!}.
\end{aligned}$$

Por lo tanto, $\lim\limits_{n\to\infty} Bin\left(n, p\right) = Poisson\left(\lambda\right)$. □

Proposición 32 (suma de variables aleatorias). Sean $X_{k\in\mathbb{N}} \sim Poisson\left(\lambda_k\right)$ variables aleatorias independientes, entonces:

$$Z = \sum_{k=1}^{n} X_k \sim Poisson\left(\lambda\right), \tag{4.47}$$

donde $\lambda = \sum\limits_{k=1}^{n} \lambda_k$.

[2] $F_X\left(x\right)$ es una función de distribución porque $\exp\left(x\right) = \sum\limits_{k=0}^{\infty} \frac{x^k}{k!}$.

Proposición 33 (distribución Poisson). Si $X \sim \textit{Poisson}(\lambda)$, entonces:

$$f_X(x,\lambda) = \begin{cases} \frac{\lambda^x \exp(-\lambda)}{\Gamma(x+1)} & \text{si } x = 0,1,\ldots \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.48}$$

donde $f_X(x,\lambda)$ es la probabilidad de obtener x ocurrencias en el periodo $0 \leq t \leq T$, λ es la tasa de ocurrencia promedio del evento por unidad de tiempo y $\Gamma(\cdot)$ es la función gama (propuesta por Euler 1738).

Ejemplo 63 (distribución Poisson). El número de nacimientos promedio en una ciudad es de uno cada diez minutos, es decir, seis cada hora ($\lambda = 6$).

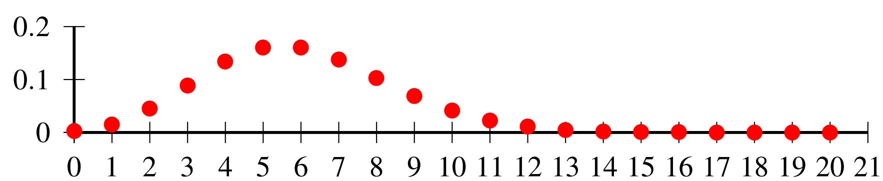
1. Calcular la probabilidad de que se presenten diez nacimientos diarios, es decir, $\mathscr{P}(X = 10)$.
2. Calcular la probabilidad de que se presenten al menos ocho nacimientos diarios, es decir, $\mathscr{P}(X \geq 8)$.
3. Calcular el número de nacimientos promedio anual, es decir, $\mathrm{E}(X)$.
4. Calcular la varianza anual, es decir, $\mathrm{Var}(X)$.
5. Presentar la gráfica de la función de densidad diaria, es decir, $f_X(x)$.
6. Presentar la gráfica de la función de distribución diaria, es decir, $F_X(x)$.

Solución. La tasa de ocurrencia promedio de nacimientos por hora es $\lambda = 6$, entonces:

1. $\mathscr{P}(X = 10) = \frac{6^{10}\exp(-6)}{10!} = \frac{60{,}466{,}176(0.002479)}{3{,}840} = 0.0413.$
2. $\mathscr{P}(X \geq 8) = 1 - \mathscr{P}(X \leq 7) = 1 - \sum_{x=0}^{7} \frac{6^x \exp(-6)}{x!}$, entonces:

   $\mathscr{P}(X \geq 8) = 1 - (0.0025 + 0.0149 + 0.0446 + 0.0892 + 0.1339 + 0.1606 + 0.1606 + 0.1377)\,.$
   $\mathscr{P}(X \geq 8) = 1 - 0.7440 = 0.2560.$

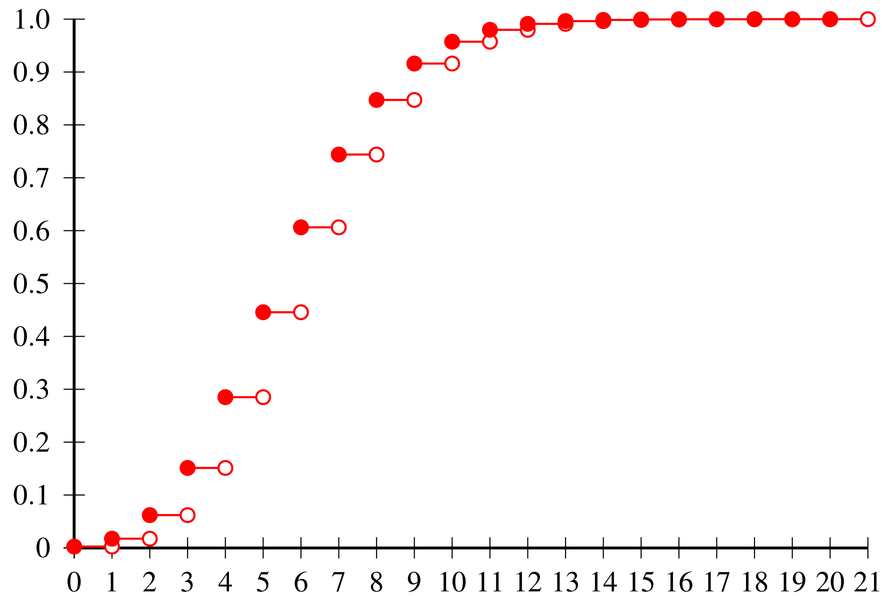
3. $\mathrm{E}(X) = \lambda(24)(365) = 6(24)(365) = 52{,}560$ anual o $\mathrm{E}(X) = 144$ diaria.
4. $\mathrm{Var}(X) = \lambda(24)(365) = 6(24)(365) = 52{,}560$ anual o $\mathrm{Var}(X) = 144$ diaria.
5. La gráfica de $f_X(x,\lambda) = \frac{6^x \exp(-10)}{x!}$ es presentada en la gráfica 4.13.



Gráfica 4.13: $f_X(x,\lambda) = \dfrac{6^x \exp(-6)}{x!}$.

Por la proposición 31 (maximizar) y aplicando la ecuación (4.45), la probabilidad de que ocurran x eventos durante una hora alcanza el máximo cuando $x = 5$ o $x = 6$ porque $\lceil 6-1 \rceil = 5$ y $\lfloor 6 \rfloor = 6$.

6. La gráfica de $F_X(x,\lambda) = \sum_{k=0}^{x} \frac{6^k \exp(-6)}{k!}$ es presentada en la gráfica 4.14.



Gráfica 4.14: $F_X(x,\lambda) = \sum_{k=0}^{x} \frac{6^k \exp(-6)}{k!}$.

Por lo tanto $\mathscr{P}(X = 10) = 0.0413, \mathscr{P}(X \geq 8) = 0.2560, \mathrm{E}(X) = 52,560$ y $\mathrm{Var}(X) = 52,560$. □

Ejemplo 64 (convergencia). En promedio, cuatro de cada diez mil productos está defectuoso. Calcular:

1. La probabilidad de encontrar menos de cinco productos defectuosos $\mathscr{P}(X < 5)$ aplicando $f_X(x,n,p)$.
2. La probabilidad de encontrar menos de cinco productos defectuosos $\mathscr{P}(X < 5)$ aplicando $f_X(x,\lambda)$.

Solución. La probabilidad de encontrar un producto defectuoso es $p = \frac{1}{2,500}$ y el número promedio de productos defectuosos es $\lambda = 4$, entonces:

1. La probabilidad de encontrar menos de cinco productos defectuosos es:

$$\mathscr{P}(X \leq 4) = \sum_{x=0}^{4} \binom{10,000}{x} \left(\frac{1}{2,500}\right)^x \left(\frac{2,499}{2,500}\right)^{10,000-x}.$$
$$\mathscr{P}(X \leq 4) = 0.0183 + 0.0732 + 0.1465 + 0.1954 + 0.1954 = 0.6288.$$

2. La probabilidad de encontrar menos de cinco productos defectuosos es:

$$\mathscr{P}(X \leq 4) = \exp(-4) \sum_{x=0}^{4} \frac{4^x}{x!} = 0.0183 + 0.0733 + 0.1465 + 0.1954 + 0.1954 = 0.6288.$$

Por lo tanto, cuando el número n de repeticiones de ensayos Bernoulli crece, el resultado se aproxima a la distribución Posisson. □

4.1.8. Aplicación de la distribución Poisson

La teoría de líneas de espera es para calcular el tiempo de espera con medidas descriptivas de eficiencia como la longitud promedio de la línea, el tiempo promedio de espera y la utilización promedio de las instalaciones (problema planteado por Erlang 1909).

Definición 72 (modelo generalizado de líneas de espera). El modelo está basado en el comportamiento de estado estable de la línea de espera y supone que la tasas promedio de llegada y salida dependen del estado, donde n es la cantidad de clientes en el sistema, λ_n es la tasa promedio de llegada, μ_n es la tasa promedio de salida y p_n es la probabilidad de estado estable de que existan n clientes en el sistema.

Para los estados $n > 0$ el estado n cambia a dos estados: $n-1$ cuando existe una salida con tasa μ_n o $n+1$ cuando existe una llegada con tasa λ_n. El estado $n = 0$ cambia al estado $n = 1$ cuando existe una llegada con tasa λ_0.

Nota 12 (tasa promedio de salida). La tasa promedio de salida μ_0 no está definida porque las salidas son inexistentes cuando el sistema está vacío.

Teorema 6 (probabilidad de estado estable). En el estado estable las tasas de entrada y salida son iguales y bajo el supuesto de que en el estado n cambia a los estados $n-1$ o $n+1$, entonces la probabilidad de estado estable es:

$$p_n = p_0 \prod_{n=1}^{n} \frac{\lambda_{n-1}}{\mu_n}, \tag{4.49}$$

donde:

$$p_0 = 1 - \sum_{n=1}^{\infty} p_n. \tag{4.50}$$

Demostración. Por hipótesis $\lambda_n = \mu_n$ y el estado n cambia a los estados $n-1$ o $n+1$, entonces las tasas promedio esperadas de llegada y de salida son:

$$\lambda_n = \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1} \qquad \text{si } n \geq 0. \tag{4.51}$$

$$\mu_n = (\lambda_n + \mu_n)\, p_n \qquad \text{si } n > 1. \tag{4.52}$$

La ecuación de balance es obtenida igualando las ecuaciones (4.51) y (4.52):

$$(\lambda_n + \mu_n)\, p_n = \lambda_{n-1} p_{n-1} + \mu_{n+1} p_{n+1}, \qquad \forall n > 0. \tag{4.53}$$

Entonces la ecuación de balance para $n = 0$ es:

$$\lambda_0 p_0 = \mu_1 p_1 \qquad \text{si } n = 0. \tag{4.54}$$

Entones la ecuación de estado estable para $n = 1$ es obtenida despejando p_1 de la ecuación (4.54):

$$p_1 = \frac{\lambda_0 p_0}{\mu_1}. \tag{4.55}$$

Las ecuaciones de balance son obtenidas recursivamente en función de p_0. Entonces:

$$(\lambda_1+\mu_1)\,p_1 = \lambda_0 p_0 + \mu_2 p_2 \qquad \text{si } n=1. \tag{4.56}$$

Por lo tanto:

$$p_2 = \frac{\lambda_0\lambda_1 p_0}{\mu_1\mu_2} \tag{4.57}$$

$$\vdots$$

$$p_n = \frac{\lambda_0\lambda_1\ldots\lambda_{n-1}p_0}{\mu_1\mu_2\ldots\mu_n} = p_0\prod_{n=1}^{n}\frac{\lambda_{n-1}}{\mu_n}. \tag{4.58}$$

Dado que $\sum\limits_{n=0}^{\infty} p_n = 1$, entonces:

$$p_0 = 1-\sum_{k=1}^{\infty} p_k. \tag{4.59}$$

Por lo tanto, $p_n = p_0\prod\limits_{n=1}^{n}\frac{\lambda_{n-1}}{\mu_n}$ donde $p_0 = 1-\sum\limits_{n=1}^{\infty} p_n$. □

Ejemplo 65 (modelo generalizado de líneas de espera). Una empresa opera con tres servidores y estima la cantidad de operadores en función de la cantidad de clientes de acuerdo con la planificación siguiente:

Clientes	Operadores
1–4	1
5–7	2
>7	3

La llegada de los clientes a las cajas de cobro presenta una distribución Poisson con una tasa promedio de llegada $\lambda_n = 12$ cada hora y el tiempo promedio de atención para cada cliente presenta una distribución exponencial con una tasa promedio de atención $\mu_n = 10$ minutos, es decir, el número promedio de clientes atendidos presenta una distribución Poisson con parámetro $\mu_n = 6$, es decir, seis clientes cada hora.

Con la información estadística anterior, calcular:

1. La probabilidad de que existan n clientes en las cajas, es decir, p_n.
2. El numero promedio esperado de servidores sin clientes, es decir, $\mathrm{E}(S \mid S=\emptyset)$.

Solución. La tasa promedio de llegada cada hora es $\lambda_n = 12$ y la tasa promedio de atención es $\mu_n = 10$ minutos, entonces la tasa de atención de clientes cada hora es:

$$\mu_n = \begin{cases} 6 & \text{si } n=1,2,3,4 \\ 12 & \text{si } n=5,6,7 \\ 18 & \text{si } n>7 \end{cases}.$$

1. Las probabilidades de estado estable para el primer servidor son:

$$p_1 = p_0 \prod_{n=1}^{1} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 p_0}{\mu_1} = \frac{12 p_0}{6} = 2p_0,$$
$$p_2 = p_0 \prod_{n=1}^{2} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 p_0}{\mu_1 \mu_2} = \left(\frac{12}{6}\right)^2 p_0 = 2^2 p_0,$$
$$p_3 = p_0 \prod_{n=1}^{3} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \lambda_2 p_0}{\mu_1 \mu_2 \mu_3} = \left(\frac{12}{6}\right)^3 p_0 = 2^3 p_0,$$
$$p_4 = p_0 \prod_{n=1}^{4} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \lambda_2 \lambda_3 p_0}{\mu_1 \mu_2 \mu_3 \mu_4} = \left(\frac{12}{6}\right)^4 p_0 = 2^4 p_0.$$

Las probabilidades de estado estable para el segundo servidor son:

$$p_5 = p_0 \prod_{n=1}^{5} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \lambda_2 \lambda_3 \lambda_4 p_0}{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5} = \left(\frac{12}{6}\right)^4 \frac{12 p_0}{12} = 2^4 p_0,$$
$$p_6 = p_0 \prod_{n=1}^{6} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5 p_0}{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5 \mu_6} = \left(\frac{12}{6}\right)^4 \left(\frac{12}{12}\right)^2 p_0 = 2^4 p_0,$$
$$p_7 = p_0 \prod_{n=1}^{7} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5 \lambda_6 p_0}{\mu_1 \mu_2 \mu_3 \mu_4 \mu_5 \mu_6 \mu_7} = \left(\frac{12}{6}\right)^4 \left(\frac{12}{12}\right)^3 p_0 = 2^4 p_0.$$

Las probabilidades de estado estable para el tercer servidor son:

$$p_{n>7} = p_0 \prod_{n=1}^{n} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \cdots \lambda_{n-1} p_0}{\mu_1 \mu_2 \cdots \mu_n} = 2^4 \left(\frac{12}{18}\right)^{n-7} p_0 = 2^4 \left(\frac{2}{3}\right)^{n-7} p_0.$$

Entonces:

$$p_8 = p_0 \prod_{n=1}^{8} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \cdots \lambda_7 p_0}{\mu_1 \mu_2 \cdots \mu_8} = 2^4 \left(\frac{2}{3}\right) p_0,$$
$$p_9 = p_0 \prod_{n=1}^{9} \frac{\lambda_{n-1}}{\mu_n} = \frac{\lambda_0 \lambda_1 \cdots \lambda_8 p_0}{\mu_1 \mu_2 \cdots \mu_9} = 2^4 \left(\frac{2}{3}\right)^2 p_0,$$
$$\vdots$$

Dado que $\sum_{n=0}^{n} p_n = 1$, entonces[3]:

$$1 = p_0 \left(1 + 2 + 4 + 8 + 16 + 16 + 16 + 16 \left(1 + \frac{2}{3} + \left(\frac{2}{3}\right)^2 + \cdots \right)\right) = p_0 \left(63 + 16 \sum_{k=0}^{\infty} \left(\frac{2}{3}\right)^k \right)^{3}.$$

$$p_0 = \frac{1}{111} = 0.\overline{009}.$$

[3]Si la razón geométrica $r < 1$, entonces $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$ y $\sum_{n=1}^{\infty} r^n = \frac{r}{1-r}$.

Entonces $p_0 = \frac{1}{111} = 0.\overline{009}$. Por lo tanto, las probabilidades de tener n clientes son:

$$\begin{array}{llll} p_0 = 0.0090, & p_1 = 0.0180, & p_2 = 0.0360, & p_3 = 0.0721, \\ p_4 = 0.1441, & p_5 = 0.1441, & p_6 = 0.1441, & p_7 = 0.1441, \\ p_8 = 0.0961, & p_9 = 0.0641, & p_{10} = 0.0427, & p_{11} = 0.0285, \\ p_{12} = 0.0190, & p_{13} = 0.0127, & p_{14} = 0.0084, & p_{15} = 0.0056. \end{array}$$

El número promedio esperado de clientes en el sistema es:

$$\mathrm{E}(n) = \sum_{n=0}^{\infty} n p_n = \frac{0 + 2 + 2^3 + 3 \cdot 2^3 + 2^6 + \cdots}{111} = 6.\overline{360}.$$

La varianza del número de clientes en el sistema es:

$$\mathrm{Var}(n) = \mathrm{E}\left(n^2\right) - \mathrm{E}^2(n) = 49.5315 - 6.\overline{360}^2 = 9.0773.$$

2. El numero promedio esperado de servidores sin clientes.

   La probabilidad de que los servidores estén funcionando es:

$$\begin{aligned} \mathscr{P}(S=1) &= p_0 \sum_{n=1}^{4} p_n = p_0 (2+4+8+16) = 30 p_0 = \frac{30}{111} = 0.27\overline{027}, \\ \mathscr{P}(S=2) &= p_0 \sum_{n=5}^{7} p_n = p_0 (16+16+16) = 48 p_0 = \frac{48}{111} = 0.\overline{4324}, \\ \mathscr{P}(S=3) &= p_0 \sum_{n=8}^{\infty} p_n = 32 p_0 = 0.\overline{2882} = p_0 \left(1 - \sum_{n=0}^{7} p_n\right) = 1 - \frac{79}{111} = \frac{32}{111} = 0.\overline{2882}. \end{aligned}$$

   Por lo tanto, la probabilidad de tener un servidor en operación es 27.03 %, la probabilidad de tener dos servidores en operación es 43.24 % y la probabilidad de tener tres servidores en operación es 28.82 %.

   El numero promedio esperado de servidores sin clientes es:

$$\mathrm{E}(S \mid S = \emptyset) = 0 \sum_{n=8}^{\infty} p_n + 1 \sum_{n=5}^{7} p_n + 2 \sum_{n=1}^{4} p_n + 3 p_0 = (0 + 48 + 60 + 3)\, p_0 = 1.$$

   Por lo tanto, el numero promedio esperado de servidores vacíos es uno. □

4.2. Distribuciones continuas

Las características de las funciones de distribución exponencial, Gauss y gama son presentadas.

4.2.1. Distribución exponencial

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en registrar el tiempo de espera transcurrido entre las ocurrencias consecutivas de eventos que ocurren en el periodo (Poisson 1837 y Erlang 1909).

La variable aleatoria X modela el tiempo de espera transcurrido $0 < x < \infty$ para la ocurrencia de un evento, suponiendo que el tiempo promedio de espera es $\lambda > 0$, entonces la variable aleatoria X tiene una distribución exponencial, es decir, $X \sim \exp(\lambda)$.

Definición 73 (distribución exponencial). Sea $\lambda > 0$, entonces la variable aleatoria X se distribuye exponencial y la función de densidad es:

$$f_X(x,\lambda) = \begin{cases} \lambda \exp(-\lambda x) & \text{si } x \geq 0 \\ 0 & \text{cualquier otro caso} \end{cases}, \tag{4.60}$$

donde $f_X(x,\lambda)$ es la frecuencia relativa de esperar x unidades de tiempo para la ocurrencia de un evento y λ el tiempo de espera promedio. La esperanza[4] y la varianza[5] son:

$$\begin{aligned}
\mathrm{E}(X) &= \lambda \int_0^\infty x\exp(-\lambda x)\,\mathrm{d}x = -x\exp(-\lambda x)\Big|_0^\infty + \int_0^\infty \exp(-\lambda x)\,\mathrm{d}x, \\
\mathrm{E}(X) &= -\lim_{x\to\infty} \cancelto{0}{x\exp(-\lambda x)} + \lim_{x\to 0} \cancelto{0}{x\exp(-\lambda x)} - \frac{\exp(-\lambda x)}{\lambda}\Big|_0^\infty, \\
\mathrm{E}(X) &= \frac{1}{\lambda}\left(-\lim_{x\to\infty} \cancelto{0}{\exp(-\lambda x)} + \lim_{x\to 0} \cancelto{1}{\exp(-\lambda x)}\right) = \frac{1}{\lambda}.
\end{aligned} \tag{4.61}$$

$$\begin{aligned}
\mathrm{Var}(X) &= \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = \lambda \int_0^\infty x^2 \exp(-\lambda x)\,\mathrm{d}x - \left(\frac{1}{\lambda}\right)^2, \\
\mathrm{Var}(X) &= -x^2\exp(-\lambda x)\Big|_0^\infty + 2\cancelto{\lambda^{-1}\mathrm{E}(X)}{\int_0^\infty x\exp(-\lambda x)\,\mathrm{d}x} - \frac{1}{\lambda^2}, \\
\mathrm{Var}(X) &= -\lim_{x\to\infty} x^2 \cancelto{0}{\exp(-\lambda x)} + \lim_{x\to 0} x^2 \cancelto{0}{\exp(-\lambda x)} + \frac{2\,\mathrm{E}(X)}{\lambda} - \frac{1}{\lambda^2} = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}, \\
\mathrm{Var}(X) &= \frac{1}{\lambda^2}.
\end{aligned} \tag{4.62}$$

Proposición 34 (función de distribución). Si $X \sim \exp(\lambda)$, entonces:

$$F_X(x,\lambda) = \lambda \int_0^x \exp(-\lambda t)\,\mathrm{d}t = 1 - \exp(-\lambda x). \tag{4.63}$$

$$1 - F_X(x,\lambda) = \exp(-\lambda x). \tag{4.64}$$

Demostración. Sea $u = \lambda t \Rightarrow \mathrm{d}u = \lambda\,\mathrm{d}t$, entonces:

$$\begin{aligned}
F_X(x,\lambda) &= \lambda \int_0^x \exp(-\lambda t)\,\mathrm{d}t = \int_0^x \exp(-u)\,\mathrm{d}u = -\exp(-u)\Big|_0^x = -\exp(-\lambda t)\Big|_0^x = 1 - \exp(-\lambda x). \\
1 - F_X(x,\lambda) &= \exp(-\lambda x).
\end{aligned}$$

Por lo tanto, $\mathscr{P}(X \leq x) = 1 - \exp(-\lambda x)$ y $\mathscr{P}(X \geq x) = \exp(-\lambda x)$. □

[4] Sea $u = x$ y $v = -\frac{\exp(-\lambda x)}{\lambda}$, entonces $\mathrm{d}u = \mathrm{d}x$ y $\mathrm{d}v = \exp(-\lambda x)$, por lo tanto $\mathrm{E}(X) = \int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$.

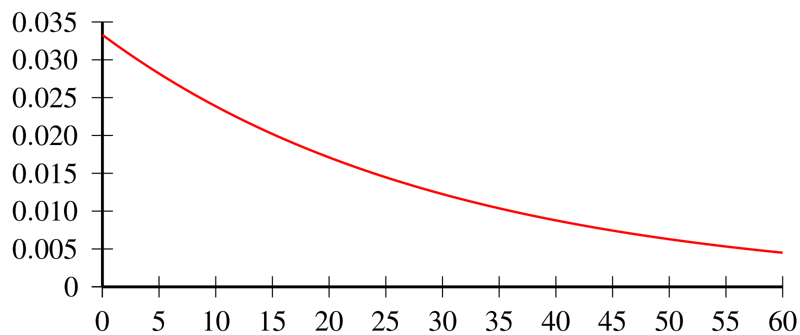
[5] Sea $u = x^2$ y $v = -\frac{\exp(-\lambda x)}{\lambda}$, entonces $\mathrm{d}u = 2x\mathrm{d}x$ y $\mathrm{d}v = \exp(-\lambda x)$, por lo tanto $\mathrm{E}\left(X^2\right) = \int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$.

Ejemplo 66 (distribución exponencial). El tiempo promedio para reparar un equipo es de treinta minutos y presenta una distribución exponencial, es decir, $X \sim \exp(\lambda)$.

1. Calcular la probabilidad de que el tiempo de reparación sea menor que 20 minutos $\mathscr{P}(X < 20)$.
2. Calcular la probabilidad de que el tiempo de reparación sea mayor que 40 minutos $\mathscr{P}(X > 40)$.
3. Calcular la probabilidad de que el tiempo de reparación sea entre 20 y 40 minutos $\mathscr{P}(20 \leq X \leq 40)$.
4. Calcular el tiempo promedio de reparación $\mathrm{E}(X)$.
5. Calcular la varianza $\mathrm{Var}(X)$.
6. Presentar la gráfica de la función de densidad $f_X(x,\lambda)$.
7. Presentar la gráfica de la función de distribución $F_X(x,\lambda)$.

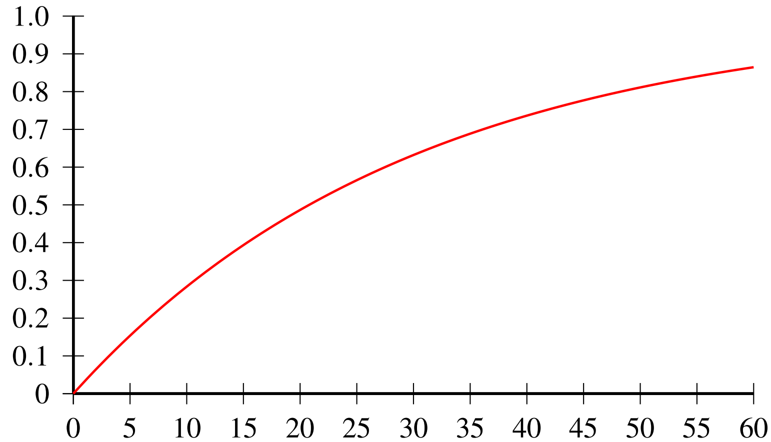
Solución. El tiempo promedio para reparar un equipo es $\lambda = \frac{1}{30} = 0.0\overline{3}$, entonces:

1. Por la ecuación (4.63) se tiene que $\mathscr{P}(X < 20) = 1 - \exp\left(-\frac{2}{3}\right) = 0.4866$.
2. Por la ecuación (4.64) se tiene que $\mathscr{P}(X > 40) = \exp\left(-\frac{4}{3}\right) = 0.2636$.
3. $\mathscr{P}(20 \leq X \leq 40) = \exp\left(-\frac{2}{3}\right) - \exp\left(-\frac{4}{3}\right) = 0.5134 - 0.2636 = 0.7364 - 0.4866 = 0.2498$.
4. Por la ecuación (4.61) se tiene que $\mathrm{E}(X) = \frac{1}{\lambda} = 30$ minutos.
5. Por la ecuación (4.62) se tiene que $\mathrm{Var}(X) = \frac{1}{\lambda^2} = 900$ minutos cuadrados, es decir, la desviación estándar es $\sigma_X = \sqrt{\mathrm{Var}(X)} = 30$ minutos.
6. La gráfica de $f_X(x,\lambda) = \frac{\exp\left(-\frac{x}{30}\right)}{30}$ es presentada en la gráfica 4.15.



Gráfica 4.15: $f_X(x,\lambda) = \dfrac{\exp\left(-\frac{x}{30}\right)}{30}$.

7. La gráfica de $F_X(x,\lambda) = \frac{1}{30}\int_0^x \exp\left(-\frac{t}{30}\right) \mathrm{d}t$ es presentada en la gráfica 4.16.



Gráfica 4.16: $F_X(x,\lambda) = \dfrac{1}{30}\displaystyle\int_0^x \exp\left(-\dfrac{t}{30}\right) \mathrm{d}t$.

Por lo tanto $\mathscr{P}(X < 20) = 0.4866$, $\mathscr{P}(X > 40) = 0.2636$, $\mathscr{P}(20 \leq X \leq 40) = 0.2498$, $\mathrm{E}(X) = 30$ y $\mathrm{Var}(X) = 900$. $\square$

Proposición 35 (relación entre las distribuciones Poisson y exponencial). Si $X \sim \mathit{Poisson}(\lambda)$ y $0 \leq t \leq T$ es el tiempo promedio esperado para la ocurrencia del evento próximo, entonces $X \sim \exp(\lambda)$.

4.2.2. Distribución Gauss

Suponiendo un experimento aleatorio $\mathscr{E}$, donde el número de ocurrencias promedio de un evento es $\mu \in \mathbb{R}$ constante y la varianza es $\sigma^2 > 0$ constante, entonces la variable aleatoria X presenta una distribución normal o Gauss y $X \sim \mathrm{N}(\mu, \sigma^2)$.

La variable aleatoria X modela la ocurrencia de un evento que tiene una ocurrencia esperada constante $\mu \in \mathbb{R}$ y una varianza constante $\sigma^2 > 0$.

Definición 74 (distribución Gauss). Sean $\mu \in \mathbb{R}$ y $\sigma^2 \in \mathbb{R}_{++}$ constantes, entonces la variable aleatoria X se distribuye normal y la función de densidad es:

$$f_X(x,\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \qquad -\infty < x < \infty, \tag{4.65}$$

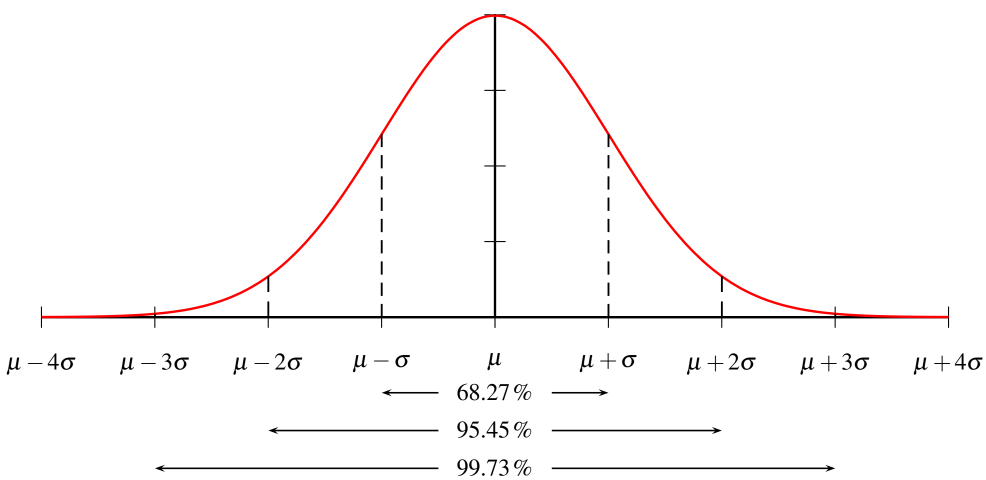
donde $f_X(x,\mu,\sigma^2)$ es la frecuencia relativa de un evento, μ es el número de ocurrencias promedio y σ^2 es la varianza del evento con respecto a μ. La esperanza y la varianza son:

$$\mathrm{E}(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{\infty}^{\infty} x \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \mathrm{d}x = \mu. \tag{4.66}$$

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - \mathrm{E}^2(X) = \sigma^2. \tag{4.67}$$

La distribución Gauss surgió como una distribución límite de la distribución binomial y fue propuesta por De Moivre (1733) que investigó las desviaciones entre la frecuencia relativa observada de un evento E en n ensayos Bernoulli y la probabilidad de ocurrencia del evento E, es decir, $\mathscr{P}(E)$. Laplace (1812) complementó el trabajo, por lo tanto es conocido como teorema de De Moivre y Laplace y que es un caso particular del teorema del límite central. Gauss (1809) aplicó la distribución suponiendo errores normales.

La función de densidad Gauss es presentada en la gráfica 4.17.



Gráfica 4.17: Función de densidad Gauss $f_X\left(x,\mu,\sigma^2\right)$.

La gráfica 4.17 indica que el parámetro μ es la media, la moda y la mediana de la función de densidad y el parámetro σ es la distancia de μ a los dos puntos de inflexión $\mu\pm\sigma$ de la función de densidad $f_X\left(x,\mu,\sigma^2\right)$.

La función de distribución de probabilidad $F_X\left(x,\mu,\sigma^2\right)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{x}\exp\left(-\frac{(u-\mu)^2}{2\sigma^2}\right)\mathrm{d}u$ es obtenida con métodos numéricos porque la integral indefinida $\int\exp\left(-x^2\right)\mathrm{d}x$ es imposible de expresar de forma cerrada, entonces la integral definida $\int_{-\infty}^{x}\exp\left(-u^2\right)\mathrm{d}u$ también es obtenida con métodos numéricos, por lo tanto existen tablas con valores para estas integrales (apéndice A).

Proposición 36 (cambio de variable). Si $X\sim\mathrm{N}\left(\mu,\sigma^2\right)$, entonces:

$$Z=\frac{X-\mu}{\sigma}\sim\mathrm{N}(0,1)\,. \tag{4.68}$$

Demostración.

$$F_Z(z)=\mathscr{P}(Z\leq z)=\mathscr{P}\left(\frac{X-\mu}{\sigma}\leq z\right)=\mathscr{P}(X\leq\mu+\sigma z)=F_X(\mu+\sigma z)\,.$$

$$f_Z(z)=\frac{\mathrm{d}F_Z(z)}{\mathrm{d}z}=\frac{\mathrm{d}F_X(\mu+\sigma z)}{\mathrm{d}z}=\sigma f_X(\mu+\sigma z)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{z^2}{2}\right).$$

Por lo tanto, $Z=\frac{X-\mu}{\sigma}\sim\mathrm{N}(0,1)$. $\square$

4.2.3. Distribución Guass estándar

Definición 75 (distribución Guass estándar). Sean $\mu = 0$ y $\sigma^2 = 1$, entonces la variable aleatoria Z se distribuye normal estándar y la función de densidad es:

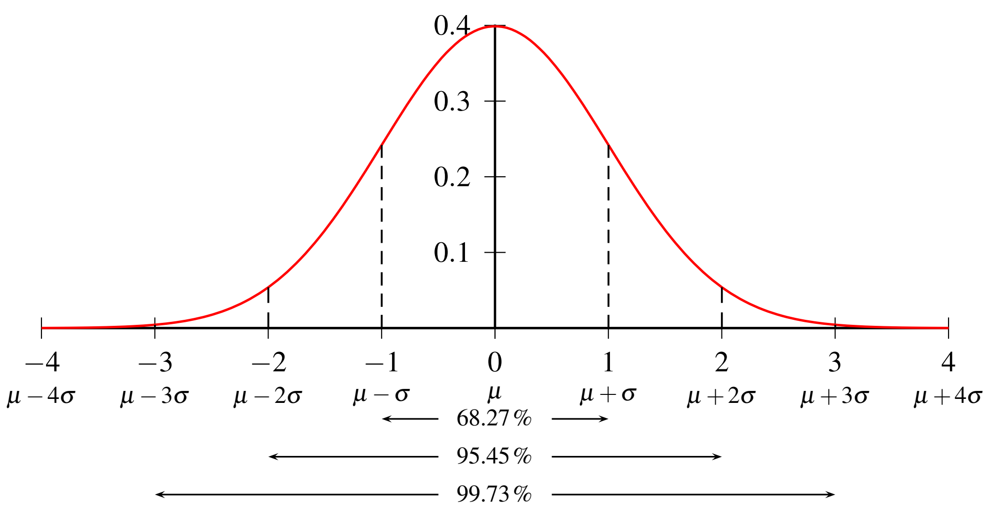
$$\phi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \qquad -\infty < z < \infty, \tag{4.69}$$

donde $\phi(z)$ es la frecuencia relativa de un evento. La esperanza y la varianza son:

$$\mathrm{E}(X) = \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\infty} z \exp\left(-\frac{z^2}{2}\right) \mathrm{d}z = 0. \tag{4.70}$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = 1. \tag{4.71}$$

La función de densidad Gauss estándar es presentada en la gráfica 4.18



Gráfica 4.18: Función de densidad Gauss estándar $f_Z(z, 0, 1)$.

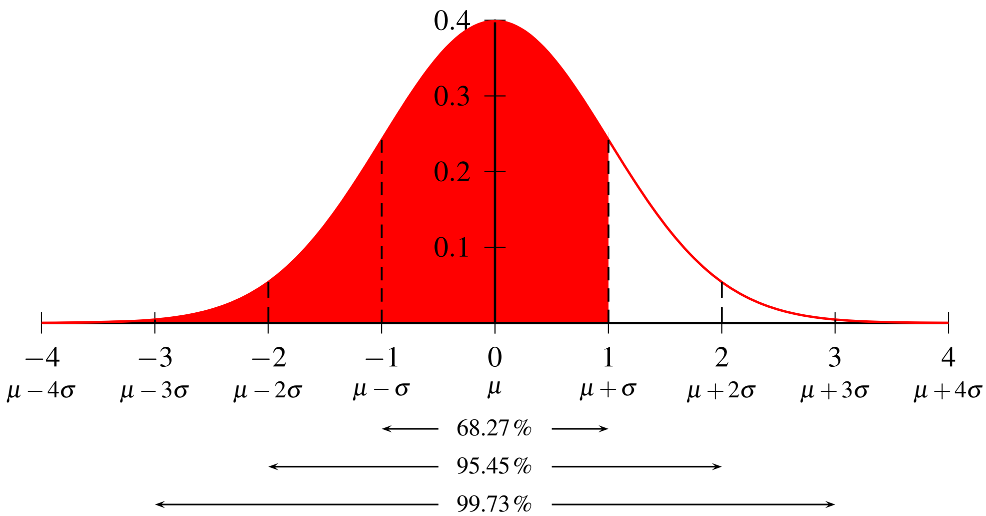
La gráfica 4.18 indica que el parámetro $\mu = 0$ es la media, la moda y la mediana de la función de densidad y el parámetro $\sigma = 1$ es la distancia de μ a los puntos de inflexión ± 1 de la función de densidad $f_Z(z)$.

Definición 76 (función de distribución Gauss estándar). Sean $\mu = 0$ y $\sigma^2 = 1$ constantes, entonces la variable aleatoria Z se distribuye normal estándar y la función de distribución es:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u \qquad -\infty < z < \infty, \tag{4.72}$$

donde $\phi(z)$ es la frecuencia relativa de un evento, $\mathrm{E}(X) = 0$ y $\mathrm{Var}(X) = 1$.

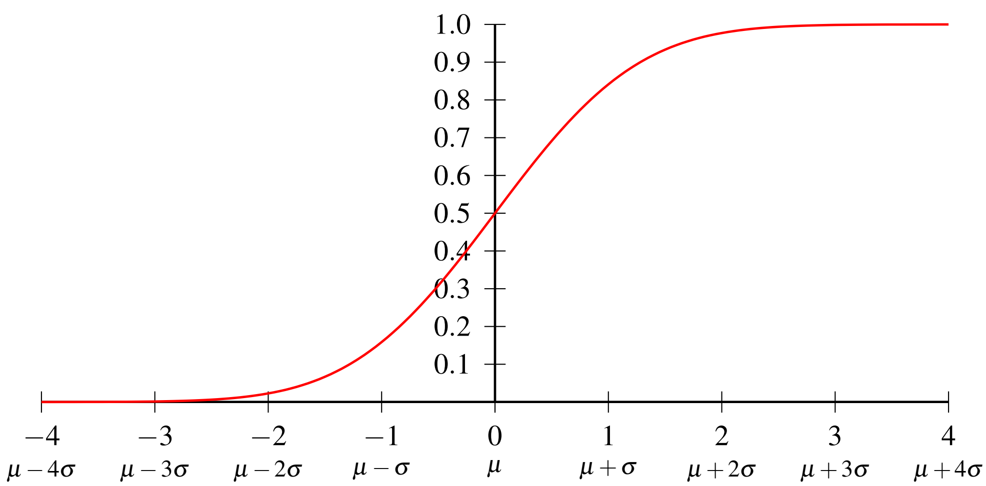
La interpretación gráfica de la función de distribución Gauss estándar es presentada en la gráfica 4.19.



Gráfica 4.19: Interpretación de la función de distribución Gauss estándar $F_Z(z,0,1)$.

La gráfica 4.19 indica que la función de distribución Gauss estándar $F_Z(z)$ es el área bajo la curva de la función de densidad $f_Z(z)$ y representa el 84.13 % (apéndice A página 196, en la celda que es la intersección de la fila con el valor 1.0 y la columna con el valor 0.00).

La función de distribución de una variable aleatoria $\Phi(z)$ es presentada en la gráfica 4.20.



Gráfica 4.20: Función de distribución Gauss estándar $F_Z(z,0,1)$.

La gráfica 4.20 muestra la simetría de la distribución de probabilidad Gauss estándar.

Ejemplo 67 (distribución Gauss). El tiempo para resolver un examen presenta una distribución Gauss con una media de cuarenta minutos ($\mu = 40$) y una desviación estándar de ocho minutos ($\sigma = 8$). Calcular:

1. El tiempo necesario para que el 95 % de alumnos resuelvan el examen.
2. El tiempo necesario para que al menos el 95 % de alumnos resuelvan el examen.
3. La probabilidad de resolver el examen en menos de 35 minutos.
4. La probabilidad de resolver el examen entre 30 y 50 minutos.

Solución. El apéndice A indica que el área bajo la curva de 95 % se presenta entre los puntos 1.64 y 1.66, es decir, $\Phi(1.6449) = 0.95$, entonces:

1. El tiempo necesario para que 95 % de alumnos resuelvan el examen es:

$$\mathscr{P}(X \leq x) = 0.95 \Rightarrow \mathscr{P}\left(Z \leq \frac{x-40}{8}\right) = 0.95 \Rightarrow \frac{x-40}{8} = 1.6449 \Rightarrow x = 53.1592.$$

2. El tiempo necesario para que al menos 95 % de alumnos resuelvan el examen es:

$$\mathscr{P}(X \leq x) \geq 0.95 \Rightarrow \mathscr{P}\left(Z \leq \frac{x-40}{8}\right) \geq 0.95 \Rightarrow \frac{x-40}{8} \geq 1.6449 \Rightarrow x \geq 53.1592.$$

3. $\mathscr{P}(X < 35) = \mathscr{P}\left(Z < -\frac{5}{8}\right) = 0.2660$ porque el apéndice A indica que $\Phi(-0.625) = 0.2660$.
4. $\mathscr{P}(30 \leq X \leq 50) = \mathscr{P}\left(-\frac{5}{4} \leq Z \leq \frac{5}{4}\right) = 0.7887$ porque el apéndice A indica que $\Phi(1.25) = 0.8944$ y $\Phi(-1.25) = 0.1056$.

Por lo tanto, el tiempo necesario para que al menos 95 % de alumnos resuelvan el examen es al menos de 53.1592 minutos, es decir, al menos 53 minutos 9.55 segundos (53'9.55"), la probabilidad de resolver el examen en menos de 35 minutos es 26.60 % y la probabilidad de resolver el examen entre 30 y 50 minutos es 78.87 %. □

4.2.4. Aplicación de la distribución Gauss

Si los rendimientos logarítmicos de un activo subyacente con una variable aleatoria $Y = \ln(X)$ presentan una distribución Gauss, entonces $\mathscr{P}(Y \leq y) = \mathscr{P}(\ln(X) \leq \ln(x)) = \mathscr{P}\left(Z \leq \frac{\ln(x)-\mu_Y}{\sigma_Y}\right) = \Phi\left(\frac{\ln(x)-\mu_Y}{\sigma_Y}\right)$.

Si $f(t,M) = \ln(M)$, entonces $f(t,M) \sim \mathrm{N}\left(\left(\mu - \frac{\sigma^2}{2}\right)\tau, \sigma^2\tau\right)$, por lo tanto $Z = \frac{\ln\left(\frac{M_T}{M_t} - \left(\mu - \frac{\sigma^2}{2}\right)\tau\right)}{\sigma\sqrt{\tau}}$, entonces se tiene que $\ln(M_T) \sim \mathrm{N}\left(\ln(M_t) + \left(\mu - \frac{\sigma^2}{2}\right)\tau, \sigma^2\tau\right)$, es decir, $\ln(M_T) = \ln(M_t) + \left(\mu - \frac{\sigma^2}{2}\right)\tau + \sigma\sqrt{\tau}Z_t$ y el precio del activo subyacente $M_T = M_t \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)\tau + \sigma\sqrt{\tau}Z_t\right)$, por lo tanto el precio del activo subyacente presenta una distribución log–gaussiana.

Black y Scholes (1973) y Merton (1973) propusieron una solución para valuar opciones europeas sobre activos subyacentes suponiendo que los rendimientos logarítmicos del activo subyacente presentan una distribución Gauss y una extensión de la solución propuesta por Black (1976) es:

$$c(t, M_t) = M_t \exp(-r\tau) \Phi(d_1) - S \exp(-i\tau) \Phi(d_2), \tag{4.73}$$

donde:

$$d_1 = \frac{\ln\left(\frac{M_t}{S}\right) + \left(i - r + \frac{\sigma^2}{2}\right)\tau}{\sigma\sqrt{\tau}}. \tag{4.74}$$

$$d_2 = d_1 - \sigma\sqrt{\tau}. \tag{4.75}$$

Ejemplo 68 (valuación de una opción europea de compra). Suponiendo que el precio de un activo subyacente es $M = 496$, la volatilidad anual es $\sigma = 0.15$, la tasa de interés libre de riesgo nacional instantánea anual es $i = 0.07$, la tasa de dividendos es nula $r = 0$, el precio de liquidación es $S = M \exp(iT) = 498.90$ y el tiempo de vigencia es un mes $T = \frac{1}{12}$. Calcular el precio de la opción europea de compra sobre el activo subyacente con la extensión del modelo de Black y Scholes (1973) y Merton (1973).

Solución. Sustituyendo en las ecuaciones (4.73)–(4.75):

$$c(t, M_t) = M_t \exp(-r\tau) \Phi(d_1) - S \exp(-i\tau) \Phi(d_1).$$
$$c(t, M_t) = 496\Phi(0.0217) - 498.90 \exp\left(-\frac{0.07}{12}\right) \Phi(-0.0217) = 8.57.$$

Por lo tanto, el precio de la opción europea de compra es $c(t, M_t) = 8.57$ que es el precio al que convergen los modelos de Cox, Ross y Rubinstein (1979) y Climent Hernández (2014) cuando $n \to \infty$. □

4.2.5. Distribución gama

Suponiendo un experimento aleatorio $\mathscr{E}$ que consiste en registrar el tiempo de espera transcurrido entre las ocurrencias consecutivas de eventos en el periodo $0 \leq t \leq T$, entonces si la variable aleatoria X es el número de veces que ocurre el evento durante el periodo $0 \leq t \leq T$, la variable aleatoria X presenta una distribución Poisson con parámetro λ, es decir, $X \sim Poisson(\lambda)$.

Si la variable aleatoria X modela el tiempo de espera transcurrido $0 \leq x < \infty$ para la ocurrencia de un evento, y suponiendo que el tiempo promedio de espera es $\lambda > 0$, entonces la variable aleatoria X presenta una distribución exponencial, es decir, $X \sim \exp(\lambda)$.

La variable aleatoria X modela el tiempo de espera transcurrido $0 \leq x < \infty$ para la ocurrencia de un evento y suponiendo que el tiempo promedio de espera es $\lambda > 0$, entonces $\forall n > 0$ la variable aleatoria X presenta una distribución gama, es decir, $X \sim \gamma(n, \lambda)$, donde n es el parámetro de forma y λ es el parámetro de escala (función de densidad obtenida por Bienaymé 1838 y 1852).

Definición 77 (distribución gama). Sean $n, \lambda > 0$, entonces la variable aleatoria X se distribuye gama y la función de densidad es:

$$f_X(x, n, \lambda) = \begin{cases} \frac{(\lambda x)^{n-1} \lambda \exp(-\lambda x)}{\Gamma(n)} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}, \tag{4.76}$$

donde $f_X(x, n, \lambda)$ es la frecuencia relativa de un evento, λ es el tiempo de espera promedio, n es el parámetro de forma y $\Gamma(n)$ es la función gama que tiene las propiedades siguientes:

Proposición 37 (función gama). Si $n > 0$, entonces la función gama (propuesta por Euler 1738) es:

$$\Gamma(n) = \int_0^\infty x^{n-1} \exp(-t) \,\mathrm{d}t \qquad \text{si } n \in \mathbb{R}_+. \tag{4.77}$$

$$\Gamma(n+1) = n\Gamma(n) \qquad \text{si } n \in \mathbb{R}_+. \tag{4.78}$$

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}. \tag{4.79}$$

$$\Gamma(n+1) = n! \qquad \text{si } n \in \mathbb{Z}_+. \tag{4.80}$$

La esperanza y la varianza son:

$$\mathrm{E}(X) = \frac{\lambda^n}{\Gamma(n)} \int_0^\infty x^n \exp(-\lambda x) \,\mathrm{d}x = \frac{n}{\lambda}. \tag{4.81}$$

$$\mathrm{Var}(X) = \mathrm{E}\left(X^2\right) - \mathrm{E}^2(X) = \frac{n}{\lambda^2}. \tag{4.82}$$

Proposición 38 (maximizar). Si $X \sim \gamma(n, \lambda)$, entonces la probabilidad de obtener el n–ésimo éxito durante el periodo $0 \leq t \leq T$ alcanza el valor máximo cuando:

$$\underset{n,\lambda}{\text{maximizar}} \, f_X(x, n, \lambda) = \left\{ x \mid x = \frac{n-1}{\lambda} \right\}, \tag{4.83}$$

donde $f_X(x, n, \lambda)$ es la probabilidad de obtener el n–ésimo éxito durante el periodo $0 \leq t \leq T$, n y λ es la tasa de ocurrencia promedio del evento por unidad de tiempo.

Ejemplo 69 (distribución gama). El tiempo de sobrevivencia de un ratón seleccionado al azar y expuesto a 240 rads de radiación presenta una distribución gama con $n = 8.5$ y $\lambda = \frac{1}{14.6} = 0.\overline{06849315}$.

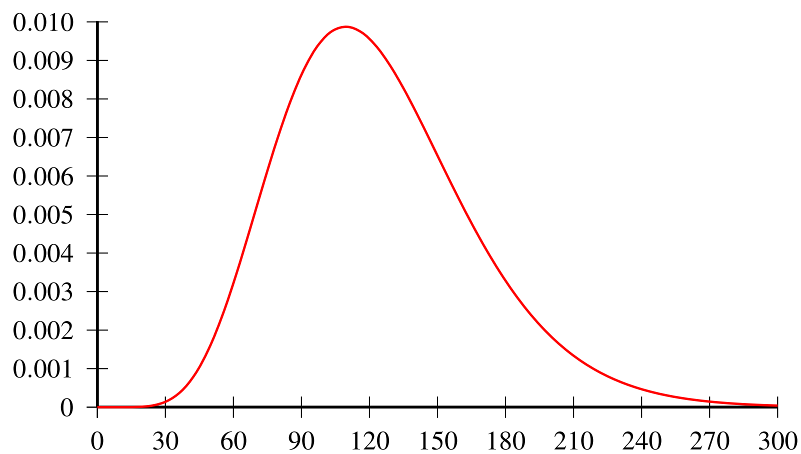
1. Calcular el tiempo esperado de sobrevivencia, es decir, $\mathrm{E}(X)$.
2. Calcular la varianza del tiempo esperado de sobrevivencia, es decir, $\mathrm{Var}(X)$.
3. La probabilidad de que un ratón sobreviva entre 60 y 120 semanas, es decir, $\mathscr{P}(60 \leq X \leq 120)$.
4. Presentar la gráfica de la función de densidad $f_X(x, n, \lambda)$ y de la
5. Presentar la gráfica de la función de distribución $F_X(x, n, \lambda)$.

Solución. El tiempo promedio de sobrevivencia es $\lambda = \frac{1}{14.6} = 0.\overline{06849315}$, es decir, 14.6 semanas, entonces:

1. El tiempo esperado de sobrevivencia es $\mathrm{E}(X) = \frac{n}{\lambda} = \frac{8.5}{14.6} = 124.1$ semanas.
2. La varianza es $\mathrm{Var}(X) = \frac{n}{\lambda^2} = \frac{8.5}{14.6^2} = 1,811.86$, es decir, la desviación estándar del tiempo de sobrevivencia es $\sigma_X = \sqrt{\mathrm{Var}(X)} = 42.5659$ semanas.

3. La probabilidad de que un ratón sobreviva entre 60 y 120 semanas es $\mathscr{P}(60 \leq X \leq 120)$, entonces se tiene que $\mathscr{P}(60 \leq X \leq 120) = \mathscr{P}(X \leq 120) - \mathscr{P}(X \leq 60) = 0.5070 - 0.0384 = 0.4686$.

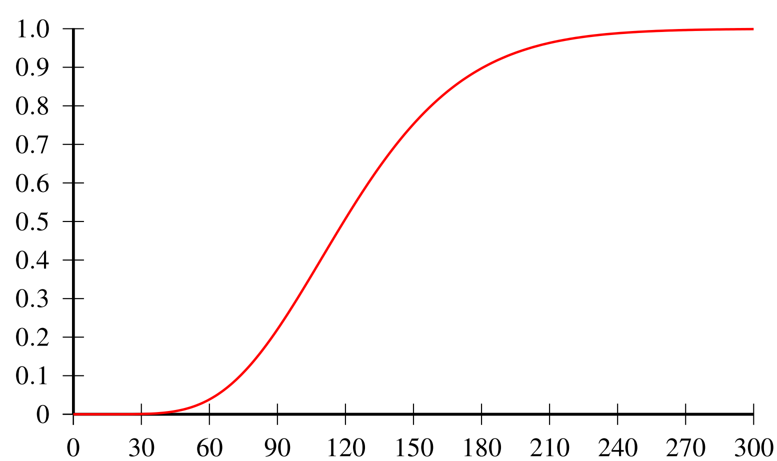
4. La gráfica de $f_X(x,n,\lambda) = \frac{x^{\frac{15}{2}} \exp\left(-\frac{x}{14.6}\right)}{14.6^{\frac{17}{2}}\Gamma\left(\frac{17}{2}\right)}$ es presentada en la gráfica 4.21.



Gráfica 4.21: $f_X(x,n,\lambda) = \dfrac{x^{\frac{15}{2}} \exp\left(-\frac{x}{14.6}\right)}{14.6^{\frac{17}{2}}\Gamma\left(\frac{17}{2}\right)}$.

Por la proposición 38 (maximizar) y aplicando la ecuación (4.83), la probabilidad de que el ratón sobreviva x semanas alcanza el máximo cuando $x = 109.5$ porque $7.5(14.6) = 109.5$.

5. La gráfica de $F_X(x,n,\lambda) = \dfrac{1}{14.6^{\frac{17}{2}}\Gamma\left(\frac{17}{2}\right)} \displaystyle\int_0^x t^{\frac{15}{2}} \exp\left(-\frac{t}{14.6}\right) \mathrm{d}t$ es presentada en la gráfica 4.22.



Gráfica 4.22: $F_X(x,n,\lambda) = \dfrac{1}{14.6^{\frac{17}{2}}\Gamma\left(\frac{17}{2}\right)} \displaystyle\int_0^x t^{\frac{15}{2}} \exp\left(-\frac{t}{14.6}\right) \mathrm{d}t$.

Por lo tanto $\mathrm{E}(X) = 124.1, \mathrm{Var}(X) = 1,811.86$ y $\mathscr{P}(60 \leq X \leq 120) = 0.4686$. $\square$

Proposición 39 (distribución gama). Si $X \sim \gamma(1,\lambda)$, entonces $X \sim \exp(\lambda)$.

Proposición 40 (distribución gama). Si $X \sim \mathrm{N}(0,1)$, entonces $X^2 \sim \gamma\left(\frac{1}{2},\frac{1}{2}\right)$.

Proposición 41 (distribución gama). Si $X \sim \chi^2(n)$, entonces $X \sim \gamma\left(\frac{n}{2},n\right)$.

Proposición 42 (distribución gama). Si $X_1,\ldots,X_n \sim \exp(\lambda)$, entonces $X = \sum_{k=1}^{n} X_k \sim \gamma(n,\lambda)$.

Proposición 43 (distribución gama). Si $X_1 \sim \gamma(n_1,\lambda),\ldots,X_m \sim \gamma(n_m,\lambda)$, entonces $X = \sum_{k=1}^{m} X_k \sim \gamma\left(\sum_{k=1}^{m} n_k,\lambda\right)$.

4.3. Ejercicios

Ejercicio 51. Si la variable aleatoria $X \sim Bin(n,p)$. Calcular los valores de los parámetros n y p cuando la esperanza es $\mathrm{E}(X) = 4$ y la varianza es $\mathrm{Var}(X) = 2$.

Ejercicio 52. Suponiendo que la probabilidad del nacimiento de un hombre es $\frac{1}{2}$ y que la familia tiene cuatro hijos, entonces $X \sim Bin\left(4,\frac{1}{2}\right)$. Calcular:

1. La probabilidad de que al menos un nacimiento sea hombre, es decir, $\mathscr{P}(X \geq 1)$.
2. La probabilidad de que al menos un nacimiento sea mujer, es decir, $\mathscr{P}(X \leq 3)$.
3. La probabilidad de que al menos un nacimiento sea mujer y uno sea hombre, es decir, $\mathscr{P}(1 \leq X \leq 3)$.
4. La esperanza $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.

Ejercicio 53. Si 10% del producto está defectuoso y $n = 5$, es decir, $X \sim Bin(5,0.1)$. Calcular:

1. La probabilidad de que ningún producto esté defectuoso, es decir, $\mathscr{P}(X = 0)$.
2. La probabilidad de que un producto esté defectuoso, es decir, $\mathscr{P}(X = 1)$.
3. La probabilidad de que menos de tres productos estén defectuosos, es decir, $\mathscr{P}(X < 3)$.
4. La esperanza $\mathrm{E}(X)$ y la varianza $\mathrm{Var}(X)$.

Ejercicio 54. El porcentaje de defectos de producción es 3%. Si la variable aleatoria X es el número de defectos en una muestra de tamaño $n = 50$, entonces $X \sim Bin(50,0.03)$. Calcular:

1. $\mathscr{P}(X = 0)$.
2. $\mathscr{P}(X \leq 3)$.
3. $\mathscr{P}(X \geq 4)$.
4. $\mathscr{P}(2 \leq X \leq 6)$.
5. $\mathrm{E}(X)$.
6. $\mathrm{Var}(X)$.

Ejercicio 55. Si 10% del producto está defectuoso y presenta una distribución geométrica. Calcular:

1. La probabilidad de aceptar 26 productos antes de rechazar el primer producto.
2. $\mathrm{E}(X)$.
3. $\mathrm{Var}(X)$.

Ejercicio 56. Si 10 % del producto está defectuoso y presenta una distribución Pascal. Calcular:

1. La probabilidad de aceptar 87 productos antes de rechazar el segundo producto.
2. $\mathrm{E}(X)$.
3. $\mathrm{Var}(X)$.

Ejercicio 57. Si 10 % del producto de una población con 100 productos está defectuoso y presenta una distribución Huygens. El departamento de control de calidad rechaza el lote de producto si encuentra cinco o más productos defectuosos. Calcular:

1. La probabilidad de rechazar menos de cinco productos de una muestra con 80 productos.
2. $\mathrm{E}(X)$.
3. $\mathrm{Var}(X)$.

Ejercicio 58. En un libro el número de errores por página es una variable aleatoria Poisson con media unitaria. Calcular:

1. La probabilidad de que una página no presente errores, es decir, $\mathscr{P}(X=0)$.
2. La probabilidad de que una página presente un error, es decir, $\mathscr{P}(X=1)$.
3. La probabilidad de que una página presente al menos tres errores, es decir, $\mathscr{P}(X \geq 3)$.
4. La probabilidad de que una página presente entre uno y cuatro errores, es decir, $\mathscr{P}(1 \leq X \leq 4)$.
5. $\mathrm{E}(X)$.
6. $\mathrm{Var}(X)$.

Ejercicio 59. El número de semillas en una naranja es una variable aleatoria Poisson con media tres. Calcular:

1. La probabilidad de que una naranja no tenga semillas, es decir, $\mathscr{P}(X=0)$.
2. La probabilidad de que una naranja tenga al menos dos semillas, es decir, $\mathscr{P}(X \geq 2)$.
3. La probabilidad de que una naranja tenga a los más tres semillas, es decir, $\mathscr{P}(X \leq 3)$.
4. La probabilidad de que una naranja tenga entre una y cinco semillas, es decir, $\mathscr{P}(1 \leq X \leq 5)$.
5. $\mathrm{E}(X)$.
6. $\mathrm{Var}(X)$.

Ejercicio 60. La probabilidad de presentar una reacción por un suero es 0.003. Calcular:

1. La probabilidad de que exactamente tres presenten una reacción si la muestra es de 1,200.
2. La probabilidad de que más de dos presenten una reacción si la muestra es de 1,200.
3. La esperanza de la variable aleatoria $X \sim Bin(1200, 0.003)$.
4. La varianza de la variable aleatoria $X \sim Bin(1200, 0.003)$.
5. La probabilidad de que tres presenten una reacción si la variable aleatoria $X \sim Poisson(36)$.
6. La probabilidad de que más de dos presenten una reacción si $X \sim Poisson(36)$.
7. La esperanza de la variable aleatoria $X \sim Poisson(36)$.
8. La varianza de la variable aleatoria $X \sim Poisson(36)$.

Ejercicio 61. El número de huracanes en un año es una variable aleatoria $X \sim \mathit{Poisson}(9)$. Calcular:

1. $\mathscr{P}(X < 5)$.
2. $\mathscr{P}(7 \leq X \leq 10)$.
3. $\mathscr{P}(X \geq 10)$.
4. $\mathrm{E}(X)$.
5. $\mathrm{Var}(X)$.
6. $\mathscr{P}\left(X \geq \mathrm{E}(X) + \sqrt{\mathrm{Var}(X)}\right)$.

Ejercicio 62. Si 10% del producto está defectuoso y presenta una distribución $\mathit{Poisson}(10)$. Control de calidad rechaza el lote de producto si encuentra cinco o más productos defectuosos. Calcular:

1. La probabilidad de rechazar menos de cinco productos.
2. $\mathrm{E}(X)$.
3. $\mathrm{Var}(X)$.

Ejercicio 63. El tiempo de revisión de un automóvil es una variable aleatoria $X \sim \exp(25)$, es decir, es una variable aleatoria exponencial con una media de 25 minutos. El costo de revisión tiene un costo fijo de \$ 400 más un costo variable de \$ 20 por el tiempo de la revisión. Calcular:

1. La probabilidad de que el tiempo de revisión sea menor que 15 minutos, es decir, $\mathscr{P}(X \leq 15)$.
2. La probabilidad de que el tiempo de revisión sea entre 15 y 45 minutos, es decir, $\mathscr{P}(15 \leq X \leq 45)$.
3. El tiempo de revisión para superar el 80% de los tiempos de revisión, es decir, $\mathscr{P}(X \leq x) = 0.8$.
4. El costo promedio esperado, es decir, $\mathrm{E}(C)$.
5. La varianza del costo, es decir, $\mathrm{Var}(C)$.

Ejercicio 64. El tiempo en minutos entre dos llegadas consecutivas a la ventanilla de un banco es una variable aleatoria $X \sim \exp(1)$. Calcular:

1. El tiempo esperado en minutos entre dos llegadas consecutivas, es decir, $\mathrm{E}(X)$.
2. La varianza del tiempo entre dos llegadas sucesivas, es decir, $\mathrm{Var}(X)$.
3. La probabilidad de que el tiempo entre dos llegadas sea menor que cinco minutos, es decir, $\mathscr{P}(X \leq 5)$.
4. La probabilidad de que el el tiempo entre dos llegadas sea mayor que seis minutos, es decir, $\mathscr{P}(X \geq 6)$.
5. La probabilidad de que el el tiempo sea entre tres y siete minutos, es decir, $\mathscr{P}(3 \leq X \leq 7)$.

Ejercicio 65. La distancia en metros que las ratas–canguro recorren desde la ubicación del nacimiento hasta el primer territorio vacante que encuentran es una variable aleatoria $X \sim \exp(0.0139)$. Calcular:

1. La distancia promedio esperada, es decir, $\mathrm{E}(X)$.
2. La varianza de la distancia recorrida, es decir, $\mathrm{Var}(X)$.
3. La probabilidad de que la distancia sea a lo más 100 metros, es decir, $\mathscr{P}(X \leq 100)$.
4. La probabilidad de que la distancia sea entre 70 y 95 metros, es decir, $\mathscr{P}(70 \leq X \leq 85)$.
5. La probabilidad de que la distancia supere la distancia promedio esperada por más de una desviación estándar, es decir, $\mathscr{P}\left(X \geq \mathrm{E}(X) + \sqrt{\mathrm{Var}(X)}\right)$.

Ejercicio 66. Calcular las probabilidades de una distribución Gauss estándar:

1. $\mathscr{P}(-1 \leq Z \leq 1)$.
2. $\mathscr{P}(0 \leq Z \leq 1.28)$.
3. $\mathscr{P}(-0.68 \leq Z \leq 1.65)$.
4. $\mathscr{P}(0.25 \leq Z \leq 1.04)$.
5. $\mathscr{P}(Z \leq 1.28)$.
6. $\mathscr{P}(Z \geq -2.33)$.

Ejercicio 67. El tiempo de vida en horas de un aparato es una variable aleatoria $X \sim \mathrm{N}\left(10000, 600^2\right)$. Calcular:

1. La probabilidad de que el aparato dure menos de 8,000 horas, es decir, $\mathscr{P}(X \leq 8,000)$.
2. La probabilidad de que dure entre de 9,000 y 12,000 horas, es decir, $\mathscr{P}(9,000 \leq X \leq 12,000)$.
3. La probabilidad de que el aparato dure más de 12,000 horas, es decir, $\mathscr{P}(X \geq 12,000)$.

Ejercicio 68. El peso en kilogramos de 1,000 estudiantes es una variable aleatoria $X \sim \mathrm{N}(70, 100)$. Calcular:

1. La probabilidad de que el peso sea menor que 50 kilogramos, es decir, $\mathscr{P}(X \leq 50)$.
2. La probabilidad de que el peso sea mayor que 80 kilogramos, es decir, $\mathscr{P}(X \geq 80)$.
3. El número de estudiantes que pesan entre 50 y 80 kilogramos, es decir, $1,000\mathscr{P}(50 \leq X \leq 80)$.
4. El número de estudiantes que pesan menos de 75 kilogramos, es decir, $1,000\mathscr{P}(X \leq 75)$.

Ejercicio 69. La resistencia en libras sobre pulgadas cuadradas del acero grado A36 es una variable aleatoria $X \sim \mathrm{N}(43, 20.25)$. Calcular:

1. La probabilidad de que la resistencia sea menor que 40 libras, es decir, $\mathscr{P}(X \leq 40)$.
2. La probabilidad de que la resistencia sea mayor que 50 libras, es decir, $\mathscr{P}(X \geq 50)$.
3. La probabilidad de que la resistencia esté entre 35 y 51 libras, es decir, $\mathscr{P}(35 \leq X \leq 51)$.

Ejercicio 70. Si $10\,\%$ del producto está defectuoso y presenta una distribución $\mathrm{N}(10, 9)$. Control de calidad rechaza el lote de producto si encuentra 5 o más productos defectuosos. Calcular:

1. La probabilidad de rechazar menos de 5 productos.
2. La probabilidad de rechazar entre 5 y 10 productos.
3. $\mathrm{E}(X)$.
4. $\mathrm{Var}(X)$.

Parte II

Estadística

Capítulo 5

Introducción

Una colección de datos de una población es de interés para el estado y las primeras colecciones de datos fueron llamadas estadísticas. Los primeros registros de una recolección sistemática de datos de una población surgieron en Venecia y Florencia durante el renacimiento. A mediados del siglo XVI, a través de parroquias, algunos gobiernos europeos registraban nacimientos, matrimonios y decesos para conocer las tasas de natalidad y de mortalidad por epidemias y plagas. La recolección de datos continuó durante los siglos XVII y XVIII para realizar censos.

En el siglo XVIII, De Moivre (1725) fundamentó la estadística demográfica. En el siglo XIX, Galton (1964) fue precursor para inferir información acerca de una población a partir de datos estadísticos. Los métodos y la teoría fueron desarrollados en el siglo XX con los trabajos de Pearson (1900), Fisher (1922), Neyman y Pearson (1928a, 1928b, 1933a, 1933b, 1936) y Neyman (1934, 1937).

Pearson (1900) desarrolló el estadístico χ^2 y encontró la distribución asintótica cuando los parámetros son conocidos y es considerada la epifanía de la estadística. Fisher (1922) introdujo los grados de libertad y la teoría de estimación por máxima verosimilitud que permite conocer la distribución asintótica de χ^2 con parámetros desconocidos.

El objetivo de los métodos estadísticos es la reducción de los datos que se logra con los datos disponibles de una muestra aleatoria de una población hipotéticamente infinita, cuya distribución con respecto a las características bajo discusión es especificada por el mínimo de parámetros y establece los tipos de problemas siguientes:

1. Especificación (seleccionar la distribución de la población).
2. Estimación (aplicar estadísticos para inferir los parámetros de la población).
3. Distribución (conocer la distribución de la muestra).

Fisher estableció tres criterios de estimación:

1. Consistencia.
2. Eficiencia.
3. Suficiencia.

Fisher (1922) utilizó los términos: parámetro, estadístico, varianza, verosimilitud, información, hipótesis nula, prueba de significación, nivel de significación, punto crítico, aleatorio y enfatizó la diferencia entre medidas descriptivas de la muestra y de la población, utilizó letras griegas para los parámetros de la población y letras latinas para los parámetros de la muestra. Fisher (1925) probó que los estimadores máximo verosímiles son asintóticos y de varianza mínima.

Neyman y Pearson (1928a, 1928b, 1933a, 1933b, 1936) describieron los métodos para seleccionar una región crítica para probar hipótesis simples en el contexto de un parámetro desconocido. Neyman (1934, 1937) inició la teoría de los intervalos de confianza.

Los métodos y conclusiones son utilizados para rechazar o no rechazar afirmaciones científicamente, es decir, la estadística explica las condiciones regulares de experimentos aleatorios.

5.1. Estadística

Definición 78 (estadística). La estadística es una ciencia que se encarga de recolectar, organizar, resumir, analizar e interpretar datos para obtener conclusiones a partir de éstos.

La estadística se divide en dos áreas:

1. Estadística descriptiva.
2. Inferencia estadística.

5.1.1. Estadística descriptiva

Definición 79 (estadística descriptiva). La estadística descriptiva es un conjunto de métodos para organizar, resumir y presentar datos.

5.1.2. Inferencia estadística

Definición 80 (inferencia estadística). La inferencia estadística es un conjunto de métodos para obtener, con un grado de confianza, información de una población con fundamento en una muestra.

Suponiendo que existe una población de interés, es decir, si existe un conjunto de objetos del que se desea conocer información y por la imposibilidad o la inconveniencia de obtener información de todos y cada uno de los elementos de la población es considerado un subconjunto y con fundamento en esta muestra se infiere la información de la población.

5.1.3. Población

Definición 81 (población). La población es un conjunto que contiene a todos y cada uno de los elementos sometidos a un estudio estadístico.

5.1.4. Muestra

Definición 82 (muestra). La muestra es un conjunto representativo de la población donde el número de elementos de la muestra es menor que el número de elementos de la población.

Los métodos estadísticos están divididos en:

1. Estadística paramétrica.
2. Estadística no paramétrica.

5.1.5. Estadística paramétrica

Definición 83 (estadística paramétrica). La estadística paramétrica es la rama de la inferencia estadística que comprende los procedimientos estadísticos que están fundamentados en la distribución de los datos y un número finito de parámetros.

La estadística paramétrica es aplicada cuando son satisfechas las condiciones siguientes:

1. La distribución de la variable dependiente es conocida.
2. La presencia de homocedasticidad (homogeneidad de la varianza).
3. La asignación y la selección de la muestra es aleatoria (muestreo aleatorio).
4. La variable dependiente está medida en escala de intervalo o de razón.

La estadística paramétrica resuelve tres tipos de problemas:

1. Estimación puntual (calcula un valor para el parámetro).
2. Estimación por intervalos (calcula un intervalo de confianza para el parámetro).
3. Contraste de hipótesis (contrasta información acerca de parámetros o de distribuciones).

5.1.6. Estadística no paramétrica

Definición 84 (estadística no paramétrica). La estadística no paramétrica es la rama de la estadística que estudia pruebas y modelos estadísticos con una distribución subyacente que no se ajusta a criterios paramétricos.

La estadística no paramétrica es aplicada cuando:

1. La distribución de los datos no es posible de definir a priori.
2. El nivel de medición empleado no es, como mínimo, la escala de intervalo.
3. El nivel de medición es ordinal o nominal.

5.2. Variables

Definición 85 (variable). Una variable es una característica de un elemento de la población.

Ejemplo 70 (variables). En una población humana se tienen variables de interés como:

1. Edad.
2. Peso.
3. Sexo.
4. Estatura.
5. Religión.
6. Grado académico.
7. Nacionalidad.
8. Estado civil.
9. Cédula única de registro poblacional (CURP). □

Las variables son clasificadas en:

1. Cualitativas.
   a) Nominal.
   b) Ordinal.
2. Cuantitativas.
   a) Discretas.
   b) Continuas.

5.3. Escalas de medición

De acuerdo con la relación de los valores de una variable son consideradas las escalas de medición siguientes:

1. Nominal.
2. Ordinal.
3. Razón.
4. Intervalo.

5.3.1. Escalas de medición de las variables cualitativas

Las variables cualitativas son clasificadas con las dos escalas siguientes:

1. Nominal.
2. Ordinal.

Definición 86 (variable nominal). Una variable cualitativa es nominal cuando es inexistente una relación de orden o magnitud entre valores y es imposible realizar operaciones aritméticas.

Ejemplo 71 (variable nominal). En una población humana son consideradas variables cualitativas nominales como:

1. Sexo.
2. Religión.
3. Nacionalidad. □

Definición 87 (variable ordinal). Una variable cualitativa es ordinal cuando existe una relación de orden entre valores y es imposible realizar operaciones aritméticas.

Ejemplo 72 (variable ordinal). En una institución educativa se tiene la variable cualitativa ordinal calificación global con los resultados siguientes:

1. No acreditado (NA).
2. Suficiente (S).
3. Bien (B).
4. Muy bien (MB).

Los valores NA, S, B y MB presentan el orden siguiente: $\mathrm{NA} < \mathrm{S} < \mathrm{B} < \mathrm{MB}$, pero es imposible realizar operaciones aritméticas como $\frac{\mathrm{B}}{\mathrm{NA}}$. □

5.3.2. Escalas de medición de las variables cuantitativas

Las variables cuantitativas son clasificadas de acuerdo con dos escalas:

1. Intervalo.
2. Razón.

Definición 88 (variable con escala de intervalo). Una variable cuantitativa se mide con escala de intervalo cuando existen relaciones de orden y distancia entre valores y es inexistente el cero absoluto, por lo tanto es imposible realizar operaciones aritméticas.

Ejemplo 73 (variable con escala de intervalo). La temperatura en grados Celsius es una variable cuantitativa con escala de intervalo porque la diferencia entre 10 °C y 12 °C es igual que la diferencia entre 15 °C y 17 °C y se tiene el orden siguiente: $10 < 12 < 15 < 17$, pero no se establece que 10 °C es equivalente a la mitad de 20 °C porque es inexistente el cero absoluto, por lo tanto es imposible realizar operaciones aritméticas como $\frac{10}{2}$. □

Definición 89 (variable con escala de razón). Una variable cuantitativa se mide con escala de razón cuando existen relaciones de orden y distancia entre valores y existe el cero absoluto, por lo tanto, es posible realizar operaciones aritméticas.

Ejemplo 74 (variable con escala de razón). En una población humana son consideradas variables cuantitativas con escala de razón como:

1. Edad.
2. Peso.
3. Estatura. □

La similitud entre variables con escalas de intervalo y de razón permite que variables con escala, relaciones de orden, distancia entre valores y con inexistencia o existencia del cero absoluto sean aplicadas con operaciones aritméticas como variables con escala de razón.

Ejemplo 75 (conjunto de datos). Un conjunto de datos es recopilado, organizado y presentado en formato de una tabla donde cada renglón representa una observación de 52 jugadores profesionales para los que son presentadas siete variables: número, posición, estatura (pies y pulgadas), peso (libras), edad y práctica profesional. Indicar si la variable es cualitativa o cuantitativa, nominal u ordinal, discreta o continua y si la escala es nominal, ordinal, razón o intervalo. Los datos[1] de los jugadores son presentado en la tabla 5.1.

Tabla 5.1: Conjunto de datos presentados en formato tabla.

#	Posición	Estatura	Peso	Edad	Práctica	#	Posición	Estatura	Peso	Edad	Práctica
1	WR	6’2”	195	26	4	53	LB	6’1”	230	26	2
2	K	6’2”	191	33	10	54	DE	6’4”	235	23	2
3	K	5’11”	205	31	1	55	LB	6’4”	256	24	4
4	QB	6´2”	238	28	6	57	LB	6’3”	243	26	3
5	P	6’3”	202	33	10	59	DE	6’5”	268	23	0
6	S	6’0”	204	26	3	63	C	6’3”	316	24	2
7	CB	6´2”	195	24	2	66	G	6’5”	308	24	3
10	QB	6’3”	225	28	4	68	G	6’5”	311	24	0
11	LB	6’3”	245	22	0	70	G	6’4”	315	31	8
13	WR	6’1”	198	25	4	71	T	6’4”	320	28	7
15	QB	6’1”	220	26	3	72	DT	6’3”	308	23	3
17	WR	6’1”	202	25	4	77	T	6’5”	320	30	11
18	S	6´2”	174	28	5	78	T	6’6”	310	24	2
20	RB	6’0”	209	24	3	79	T	6’8”	325	36	8
21	RB	6’0”	228	26	6	81	WR	6’3”	220	24	0
24	CB	6’1”	192	22	0	84	TE	6’5”	238	23	2
25	CB	6’4”	185	23	0	85	WR	6’2”	225	25	5
26	CB	5’10”	195	26	5	86	TE	6’5”	244	25	4
27	S	6’4”	215	27	6	87	TE	6’5”	252	27	5
28	S	6´2”	212	25	5	88	WR	6’2”	189	22	2
29	CB	6’3”	190	31	6	91	DE	6’3”	305	27	5
30	CB	5’11”	196	27	6	92	DE	6’4”	255	24	4
32	RB	5’10”	220	27	5	93	DE	6’4”	266	27	5
38	S	6’4”	205	21	0	97	DT	6’2”	280	23	0
42	LB	6’0”	211	26	6	98	DT	6’4”	360	22	0
52	G	6’5”	298	24	4	99	DT	6’2”	315	28	5

[1] www.dallascowboys.com/team/players-roster consultado el 29 de noviembre de 2021.

Solución. La variable número es cuantitativa, discreta y la escala es de intervalo. El número identifica de forma única a cada jugador, la tabla 5.1 está ordenada ascendentemente por el número y no es necesario realizar operaciones con los números que identifican a los jugadores.

La variable posición es cualitativa y es nominal. La posición está relacionada con las actividades de los jugadores. La tabla 5.1 presenta catorce posiciones (siete ofensivas, cinco defensivas y dos para patear): centro (C), guardia (G), mariscal de campo (QB), corredor (RB), bloqueador (T), cerrado externo (TE), receptor abierto (WR), esquinero (CB), defensivo externo (DE), atajador defensivo (DT), apoyador (LB), profundo (S), pateador inicial o de gol de campo (K) y pateador de despeje (P).

La variable estatura es cuantitativa, continua (pero está representada con una precisión en pulgadas) y la escala es de razón. La estatura mínima es 5 pies y 10 pulgadas (177.8 cm) y la estatura máxima es 6 pies y 8 pulgadas (203.2 cm).

La variable peso es cuantitativa, continua (pero está representada con una precisión en libras) y la escala es de razón. El peso mínimo es 174 libras (78.996 kg) y el peso máximo es 360 libras (163.440 kg).

La variable edad es cuantitativa, continua (pero está representada con una precisión en años) y la escala es de razón. La edad mínima es 21 años y la edad máxima es 36 años.

La variable práctica profesional es cuantitativa, continua (pero está representada con una precisión en años) y la escala es de razón. La experiencia mínima es 0 años y la experiencia máxima es 11 años.

Es posible estimar estadísticos descriptivos como: media, mediana, moda, varianza, desviación estándar, coeficiente de variación, coeficiente de asimetría y coeficiente de curtosis de la estatura, el peso, la edad y la práctica profesional. La estimación de la proporción, la media, la varianza y la desviación estándar es posible de realizar puntualmente y por intervalos de confianza y realizar pruebas de hipótesis acerca de parámetros como la proporción, la media, la varianza, la desviación estándar y pruebas de bondad de ajuste acerca de la distribución que presenta la variable analizada. □

5.4. Ejercicios

Ejercicio 71. Clasificar la variable como cualitativa o cuantitativa. Si la variable es cuantitativa indicar si es discreta o continua:

1. El tiempo promedio de vida útil de un foco.
2. El número de hijos de una familia.
3. La cantidad de agua en una presa.
4. El estado civil de una persona.
5. El tiempo promedio para realizar un trabajo.
6. La temperatura mínima de un observatorio en un día.
7. El número de mascotas de una familia.
8. El coeficiente intelectual de una persona.
9. El lugar de residencia de una persona.
10. La actividad profesional u oficio de una persona.

Ejercicio 72. Clasificar la variable cualitativa como nominal u ordinal:

1. El color de pelo de una persona.
2. El nivel de estudios de una persona de 21 años.
3. El país de de nacimiento de una persona.
4. La marca de un automóvil.
5. El mes de nacimiento de una persona.
6. El día del año más caluroso.
7. La estación del año más fría.
8. La raza de gatos de una ciudad.
9. La ocupación de una persona.
10. La respuesta a una pregunta en un examen.

Ejercicio 73. Clasificar la variable cuantitativa como discreta o continua y la escala como intervalo o razón.

1. La medida del calzado de una persona.
2. El peso en kilogramos de un recién nacido.
3. La altura de los árboles de una especie en un bosque.
4. El consumo de energía eléctrica de una casa en un periodo determinado.
5. El coeficiente intelectual de una persona.

Ejercicio 74. Clasificar la variable como discreta o continua:

1. El volumen de acciones que se venden diariamente en una bolsa de valores.
2. La temperatura registrada cada minuto en un observatorio.
3. El tiempo de vida promedio en horas de un proyector.
4. El ingreso promedio anual de los profesores universitarios.
5. La longitud de los pernos producidos en una fábrica.

Ejercicio 75. Para las poblaciones hipotéticas proporcionar una muestra de tamaño $n = 2$:

1. La distancia de lanzar un balón de fútbol americano.
2. El número de páginas de los libros publicados en el año 2020.
3. La magnitud (intensidad) de un temblor (escala de Richter) en México en el año 2020.

Capítulo 6

Estadística descriptiva

Suponiendo que se tiene un conjunto de datos $\{x_1,\ldots,x_N\}$ que representan valores de una variable de interés obtenidos en un experimento aleatorio $\mathscr{E}$, entonces la primera etapa de la estadística consiste en recolectar una muestra $\{x_1,\ldots,x_n\}$ donde $n < N$.

6.1. Muestreo

Suponiendo que se tiene una población de interés para inferir con respecto a la población, entonces debido a la imposibilidad o inconveniencia para obtener información de todos y cada uno de los elementos de la población, se toma una muestra para inferir la información de la población.

Definición 90 (muestreo). El muestreo es un procedimiento para estudiar una muestra e inferir con respecto a la población.

El muestreo tiene las razones siguientes:

1. Existen pruebas destructivas y es costoso destruir parte o toda la muestra.
2. Es más económico investigar con fundamento en una muestra.

Los tipos de muestreo son:

1. Aleatorio.
2. No aleatorio.

6.1.1. Muestreo aleatorio

Definición 91 (muestreo aleatorio). El muestreo aleatorio satisface las condiciones siguientes:

1. Cada uno de los elementos de la población tiene una probabilidad de ser seleccionado.
2. Cada selección es independiente de las otras (independencia).

El muestreo aleatorio aplica los métodos siguientes:

1. Métodos simples.
    a) Muestreo con reemplazo.
    b) Muestreo sin reemplazo.
2. Métodos modificados.
    a) Muestreo por conglomerados (la población es dividida en conglomerados).
    b) Muestreo estratificado (la población es dividida en estratos con igual número de elementos).
    c) Muestreo estratificado proporcional (la muestra de cada estrato es proporcional al estrato).
    d) Muestreo sistémico[1].
3. Métodos compuestos.
    a) Muestreo mixto (combina dos o más de los métodos simples anteriores).

6.1.2. Muestreo no aleatorio

Definición 92 (muestreo no aleatorio). El muestreo no aleatorio satisface la condición siguiente:

1. Cada elemento de la población es seleccionado sin azar.

El muestreo no aleatorio genera muestras representativas, pero no genera información acerca del grado de representatividad de la muestra.

El muestreo no aleatorio aplica los métodos siguientes:

1. Muestreo bola de nieve (ubica objetivos que conducen a otros).
2. Muestreo de conveniencia (incluye grupos típicos para obtener muestras representativas).
3. Muestreo por cuotas (conoce estratos de la población e individuos representativos).
4. Muestreo discrecional (los elementos son elegidos a criterio del investigador).

6.2. Clases

Las variables son dividas en categorías mutuamente excluyentes y exhaustivas $C_1, C_2, \ldots, C_n$, por lo tanto cada valor es clasificado en una (exhaustiva) y sólo una clase (exclusiva).

Definición 93 (clase). Una clase es un conjunto de categorías.

Si los datos de una variable están agrupados en clases, entonces los datos individuales son aproximados con las marcas de clase.

[1] La población de tamaño N es dividida por la muestra de tamaño n y es elegido un número aleatorio k para tomar los elementos $k, k+\mathrm{d}k, \ldots, k+(n-1)\,\mathrm{d}k$ donde $\mathrm{d}k = \frac{N}{n}$.

Definición 94 (marca de clase). Una marca de clase es un dato representativo de la clase:

$$\bar{C}_k = \frac{\text{mín}\,(C_k) + \text{máx}\,(C_k)}{2}, \tag{6.1}$$

donde $\bar{C}_k$ es la marca de clase, $\text{mín}\,(C_k)$ es el valor mínimo de la clase y $\text{máx}\,(C_k)$ es el valor máximo de la clase.

6.3. Gráficas

La información representada gráficamente está organizada, resumida y facilita la comprensión del experimento aleatorio.

La metodología para presentar las gráficas consiste en:

1. Definir las categorías (clases).
2. Ordenar la información.
3. Calcular:
    a) Las frecuencias absolutas f_k.
    b) Las frecuencias relativas fr_k.
    c) Las frecuencias acumuladas F_k.
    d) Las frecuencias relativas acumuladas Fr_k.
4. Resumir la información por categorías (clases).
5. Presentar las gráficas.
6. Analizar la información y las gráficas.

Las gráficas presentadas en esta sección son:

1. Gráfica de barras.
2. Histograma.
3. Polígono de frecuencias.
4. Ojiva.
5. Gráfica de pastel.

6.3.1. Gráfica de barras

Definición 95 (gráfica de barras). Una gráfica de barras representa las frecuencias (absolutas, relativas, acumuladas o relativas acumuladas) de las categorías de una variable a través de barras.

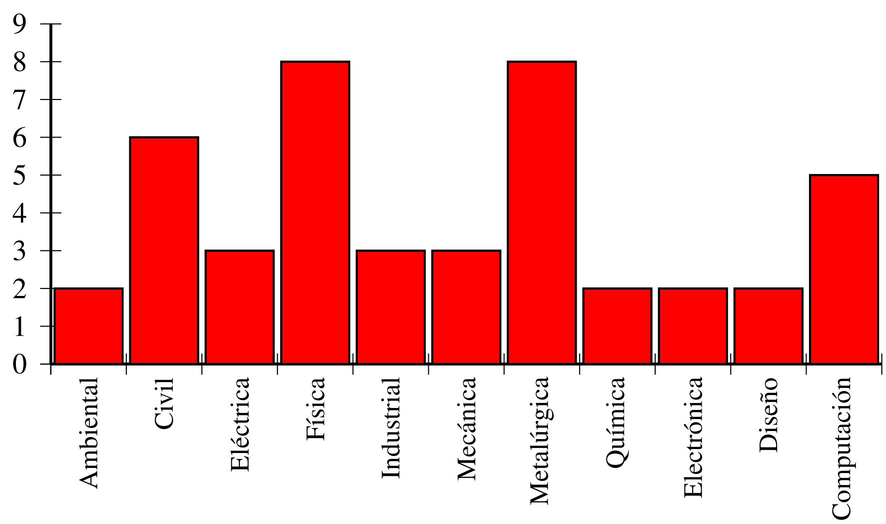
Ejemplo 76 (gráfica de barras). Con la información de la tabla 2.1 del ejemplo 27, presentar una gráfica de barras con las frecuencias absolutas de las once licenciaturas como categorías.

Solución. La información de la tabla 2.1 está resumida en la tabla 6.1.

Tabla 6.1: Frecuencias de las licenciaturas del grupo CSI02 del trimestre 2018-O.

Licenciatura	f	fr	F	Fr
Ambiental	2	0.0455	2	0.0455
Civil	6	0.1364	8	0.1818
Eléctrica	3	0.0682	11	0.2500
Física	8	0.1818	19	0.4318
Industrial	2	0.0455	21	0.4773
Mecánica	3	0.0682	24	0.5455
Metalúrgica	8	0.1818	32	0.7273
Química	2	0.0455	34	0.7727
Electrónica	3	0.0682	37	0.8409
Diseño	2	0.0455	39	0.8864
Computación	5	0.1136	44	1.0000

En la gráfica 6.1 es presentada la gráfica de barras de las frecuencias absolutas de las licenciaturas que estudian los alumnos del grupo CSI02 del trimestre 2018-O.



Gráfica 6.1: Gráfica de barras de las frecuencias absolutas del grupo CSI02 del trimestre 2018-O.

La variable es cualitativa y es ordinal porque el orden en el que están presentadas las licenciaturas es de acuerdo con el número de plan, pero la presentación de orden no es única porque es posible presentar las licenciaturas en otro orden, por ejemplo: alfabéticamente. □

6.3.2. Histograma

Definición 96 (histograma). Un histograma representa las frecuencias (absolutas, relativas, acumuladas o relativas acumuladas) de las categorías ordenadas de una variable a través de barras.

La gráfica 6.1 (gráfica de barras) es un histograma de las frecuencias absolutas.

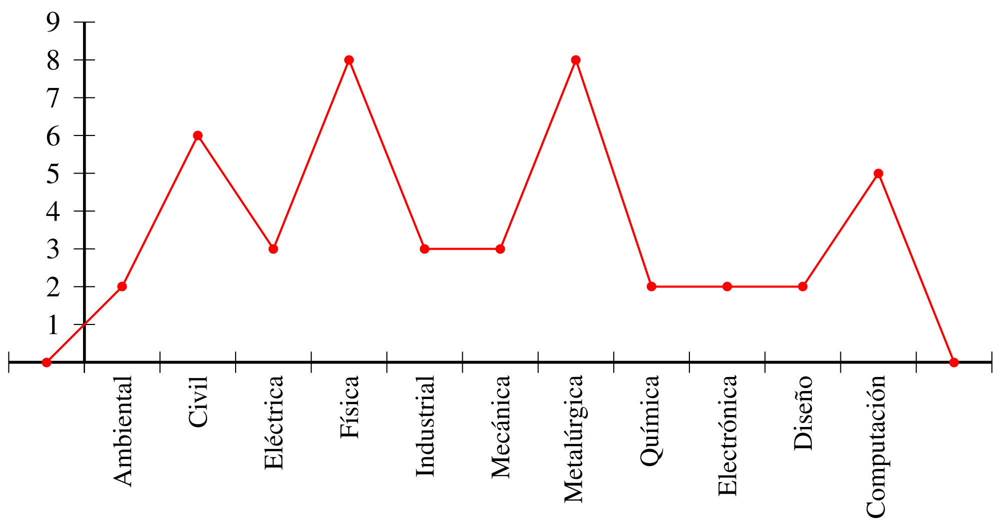
Los intervalos de clase de variables cuantitativas con escala de intervalo o escala de razón permiten presentar histogramas de frecuencias para identificar modelos de probabilidad.

6.3.3. Polígono de frecuencias

Definición 97 (polígono de frecuencias). Un polígono de frecuencias (absolutas, relativas, acumuladas o relativas acumuladas) representa las frecuencias de las categorías ordenadas de una variable a través de segmentos de líneas rectas que tienen vértices en las marcas de clase.

Ejemplo 77 (polígono de frecuencias). Con la información presente en la tabla 2.1, presentar un polígono de frecuencias absolutas de las once licenciaturas como categorías.

Solución. En la gráfica 6.2 es presentado el polígono de frecuencias absolutas de las licenciaturas que estudian los alumnos del grupo CSI02 del trimestre 2018-O.



Gráfica 6.2: Polígono de frecuencias absolutas del grupo CSI02 del trimestre 2018-O.

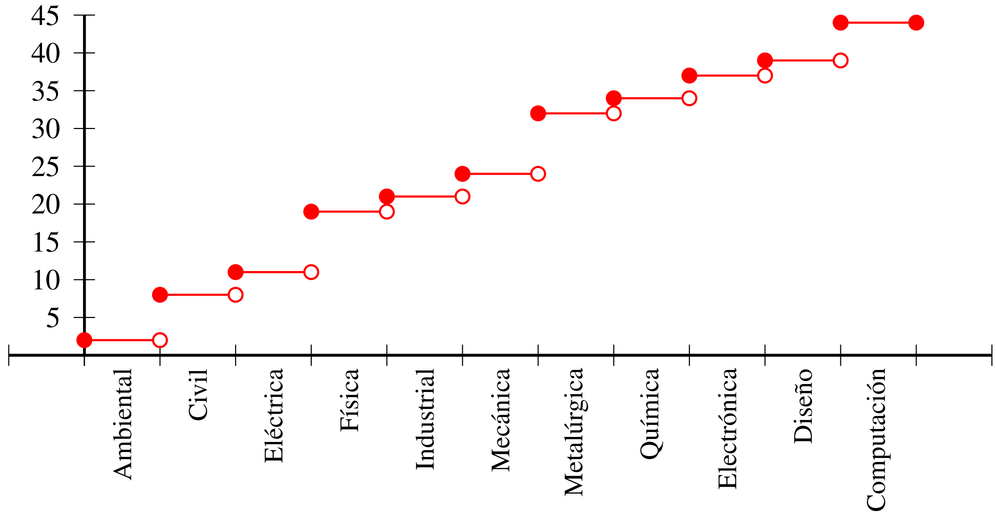
El polígono de frecuencias relativas es semejante, pero la escala del eje de las ordenadas cambia $2.\overline{27}$ % en cada uno de los nueve valores. □

6.3.4. Ojiva

Definición 98 (ojiva). Una ojiva representa las frecuencias absolutas acumuladas o las frecuencias relativas acumuladas de las categorías ordenadas de una variable a través de segmentos de líneas rectas.

Ejemplo 78 (ojiva). Con la información de la tabla 2.1, presentar una ojiva de frecuencias acumuladas de las once licenciaturas como categorías.

Solución. En la gráfica 6.3 es presentada la ojiva de frecuencias acumuladas de las licenciaturas en ingeniería y diseño que estudian los alumnos del grupo CSI02 del trimestre 2018-O.



Gráfica 6.3: Ojiva del grupo CSI02 del trimestre 2018-O.

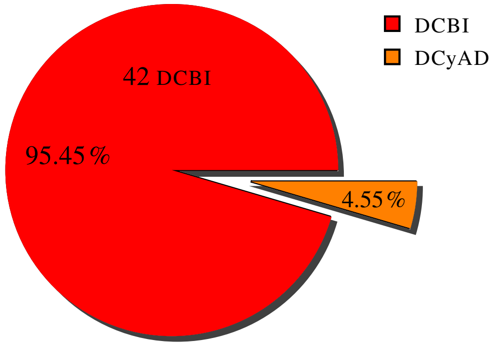
La ojiva de frecuencias absolutas es semejante, pero la escala del eje de las ordenadas cambia. □

6.3.5. Gráfica de pastel

Definición 99 (gráfica de pastel). Una gráfica de pastel es una gráfica circular que representa las frecuencias relativas de las categorías de una variable a través de arcos circulares.

Ejemplo 79 (gráfica de pastel). Con la información de la tabla 2.1, presentar una gráfica de pastel con los porcentajes de los alumnos de la División de Ciencias Básicas e Ingeniería (DCBI) y de la División de Ciencias y Artes para el Diseño (DCyAD).

Solución. En la gráfica 6.4 es presentada la gráfica de pastel de las licenciaturas en ingeniería de (DCBI) y diseño de (DCyAD) de los alumnos del grupo CSI02 del trimestre 2018-O.



Gráfica 6.4: Gráfica de pastel del grupo CSI02 del trimestre 2018-O.

Los porcentajes de la gráfica de pastel son equivalentes a las frecuencias relativas. □

Ejemplo 80 (gráficas). El grupo CSI02 de la Unidad de Enseñanza Aprendizaje (UEA) de Probabilidad y Estadística del trimestre 2018-O inició con 44 asistentes, 42 de alguna de las 10 licenciaturas en Ingeniería y 2 de la licenciatura en Diseño Industrial. El experimento aleatorio consistió en preguntar a cada alumno en qué trimestre se ha inscrito y registrar la respuesta. Después de las 44 preguntas fueron obtenidos los resultados que son presentados en la tabla 6.2.

Tabla 6.2: Trimestre de inscripción de los alumnos del grupo CSI02 del trimestre 2018-O.

Alumno	Trimestre	Alumno	Trimestre	Alumno	Trimestre	Alumno	Trimestre
1	5	12	5	23	7	34	5
2	5	13	3	24	10	35	7
3	6	14	9	25	5	36	4
4	7	15	7	26	8	37	8
5	7	16	5	27	7	38	5
6	7	17	5	28	7	39	7
7	7	18	5	29	7	40	7
8	7	19	7	30	7	41	9
9	7	20	7	31	5	42	8
10	7	21	11	32	9	43	4
11	7	22	7	33	7	44	4

Con la información presentada en la tabla 6.2:

1. Calcular la frecuencia absoluta para cada una de las 9 clases $3,4,\ldots,11$.
2. Calcular las marcas de clase de las 9 clases donde $C_1 = [3,4)\,, C_2 = [4,5)\,,\ldots, C_9 = [11,12)$.
3. Presentar un histograma de frecuencias absolutas.
4. Presentar un polígono de frecuencias absolutas.
5. Presentar una gráfica de pastel.

Solución. Con la información de la tabla 6.2 se tiene que:

1. Las frecuencias absolutas de las 9 clases son presentadas en la tabla 6.3.

Tabla 6.3: Frecuencia absoluta del trimestre de inscripción.

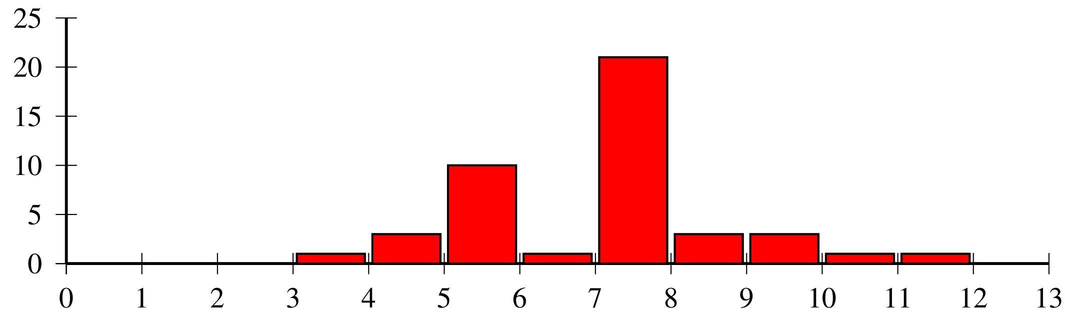
Categoría	$[3,4)$	$[4,5)$	$[5,6)$	$[6,7)$	$[7,8)$	$[8,9)$	$[9,10)$	$[10,11)$	$[11,12)$
f_k	1	3	10	1	21	3	3	1	1

2. Las marcas de clase de las 9 clases son presentadas en la tabla 6.4.

Tabla 6.4: Marcas de clase del trimestre de inscripción.

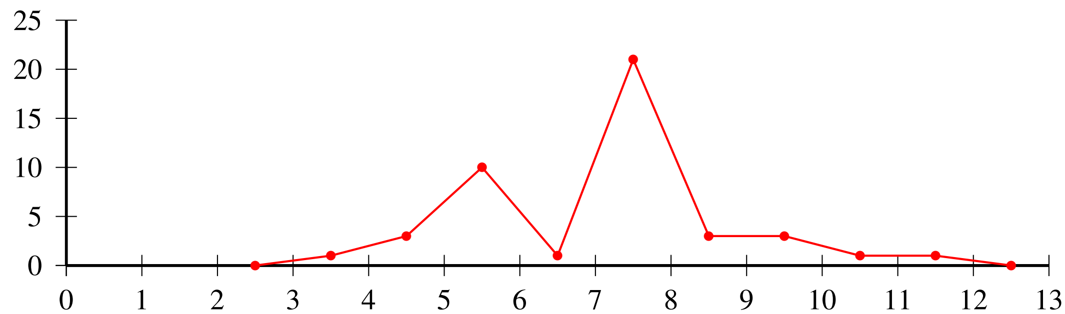
Categoría	$[3,4)$	$[4,5)$	$[5,6)$	$[6,7)$	$[7,8)$	$[8,9)$	$[9,10)$	$[10,11)$	$[11,12)$
$\bar{C}_k$	3.5	4.5	5.5	6.5	7.5	8.5	9.5	10.5	11.5

3. El histograma de frecuencias absolutas es presentado en la gráfica 6.5.



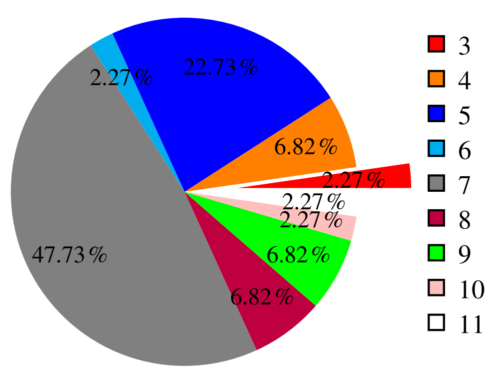
Gráfica 6.5: Histograma de frecuencias absolutas del trimestre de inscripción.

4. El polígono de frecuencias absolutas es presentado en la gráfica 6.6.



Gráfica 6.6: Polígono de frecuencias absolutas del trimestre de inscripción.

5. La gráfica de pastel es presentada en la gráfica 6.7.



Gráfica 6.7: Gráfica de pastel del trimestre de inscripción.

Los porcentajes de la gráfica de pastel son equivalentes a las frecuencias relativas. □

6.4. Intervalos de clase

Los intervalos de clase son para agrupar la muestra y el número de categorías depende de factores como:

1. El tamaño de la muestra.
2. La precisión.
3. Los métodos de inferencia.

El número de intervalos de clase es establecido en función de intereses concretos.

6.4.1. Método de Velleman

Definición 100 (método de Velleman 1976). El número de categorías en el que es dividida la muestra es:

$$k = \lceil \sqrt{n} \rceil, \tag{6.2}$$

donde k es el número de intervalos de clase (categorías), n es el número de elementos de la muestra y $\lceil \cdot \rceil$ es la función entero superior.

6.4.2. Método de Sturges

Definición 101 (método de Sturges 1926). El número de categorías en el que es dividida la muestra es:

$$k = \lceil 1 + \log_2 (n) \rceil, \tag{6.3}$$

donde k es el número de categorías, $\log_2 (n)$ es el logaritmo base dos, n es el número de elementos de la muestra y $\lceil \cdot \rceil$ es la función entero superior.

6.4.3. Método de Scott

Definición 102 (método de Scott 1979). El número de intervalos de clase en el que es dividida la muestra es:

$$k = \left\lceil \frac{2\sqrt[3]{n} \cdot \operatorname{ran}(X)}{7S_X} \right\rceil = \left\lceil \frac{2\sqrt[3]{n} \left(\operatorname{máx}(X) - \operatorname{mín}(X) \right)}{7S_X} \right\rceil, \tag{6.4}$$

donde k es el número de categorías, $\operatorname{ran}(X) = \operatorname{máx}(X) - \operatorname{mín}(X)$ es el rango de la muestra, n es el número de elementos de la muestra, S_X es la desviación estándar[2] de la muestra y $\lceil \cdot \rceil$ es la función entero superior.

[2] La varianza de la muestra es $S_X^2 = \frac{1}{n-1} \sum_{k=1}^{n} (x_k - \bar{x})^2$, donde $\bar{X} = \frac{1}{n} \sum_{k=1}^{n} X_k$ es la media de la muestra.

6.4.4. Método de Freedman y Diaconis

Definición 103 (método de Freedman y Diaconis 1981). El número de intervalos de clase en el que es dividida la muestra es:

$$k = \left\lceil \frac{\sqrt[3]{n} \cdot \operatorname{ran}(X)}{2(Q_3 - Q_1)} \right\rceil = \left\lceil \frac{\sqrt[3]{n}\,(\text{máx}(X) - \text{mín}(X))}{2(Q_3 - Q_1)} \right\rceil, \tag{6.5}$$

donde k es el número de categorías, $\operatorname{ran}(X) = \text{máx}(X) - \text{mín}(X)$ es el rango de la muestra, n es el número de elementos de la muestra, Q_3 es el tercer cuartil[3] de la muestra, Q_1 es el primer cuartil[4] de la muestra y $\lceil \cdot \rceil$ es la función entero superior.

Nota 13 (intervalos de clase). El número de categorías (intervalos de clase) en los que la muestra es dividida presenta el comportamiento siguiente:

$$\begin{aligned}
k_{FD} > k_{SC} \geq k_{ST} \geq k_V \quad & \text{si } 1 \leq n \leq 2,\\
k_{FD} \geq k_{ST} \geq k_{SC} = k_V \quad & \text{si } 3 \leq n \leq 9,\\
k_{FD} > k_{ST} \geq k_V > k_{SC} \quad & \text{si } 10 \leq n \leq 50,\\
k_{FD} \geq k_V > k_{ST} > k_{SC} \quad & \text{si } 51 \leq n \leq 324,\\
k_V > k_{FD} > k_{ST} \geq k_{SC} \quad & \text{si } 325 \leq n \leq 698,\\
k_V > k_{FD} > k_{SC} > k_{ST} \quad & \text{si } n \geq 699,
\end{aligned}$$

donde k_{FD} es el número de categorías por el método de Freedman y Diaconis, k_{SC} es el número de categorías por el método de Scott, k_{ST} es el número de categorías por el método de Sturges, k_V es el número de categorías por el método de Velleman y n es el tamaño de la muestra.

6.4.5. Límites de los intervalos de clase

El número de intervalos de clase k es para calcular los límites del k–ésimo intervalo de clase donde $1 \leq k$ a través de una progresión aritmética con la razón:

$$\mathrm{d}k = \frac{\text{máx}(X) + u - \text{mín}(X)}{k}, \tag{6.6}$$

donde $\mathrm{d}k$ es la razón de la progresión aritmética (amplitud de los intervalos de clase) y u es la unidad mínima utilizada. Entonces los límites de los intervalos de clase son:

$$\text{mín}(C_k) = \begin{cases} \text{mín}(X) & \text{si } k = 1 \\ \text{máx}(C_{k-1}) & \text{si } k > 1 \end{cases}, \tag{6.7}$$

$$\text{máx}(C_k) = \text{mín}(C_k) + k\,\mathrm{d}k, \tag{6.8}$$

donde $\text{mín}(C_k)$ es el mínimo del k–ésimo intervalo de clase, $\text{mín}(X)$ es el mínimo de la muestra, $\text{máx}(C_{k-1})$ es el máximo del k–ésimo intervalo de clase anterior y $\mathrm{d}k$ es la razón de la progresión aritmética que genera los k intervalos de clase mutuamente excluyentes y que contienen al rango de la muestra:

$$\operatorname{ran}(X) \subset \{[\text{mín}(C_1), \text{máx}(C_1)), [\text{mín}(C_2), \text{máx}(C_2)), \ldots, [\text{mín}(C_k), \text{máx}(C_k))\}.$$

[3] $Q_3 = \left\{ X_{(k)} \mid k = \frac{3(n+1)}{4} \right\}$. Si $k \notin \mathbb{N}$, entonces $Q_3 = X_{(\lfloor k \rfloor)} + \left(X_{(\lfloor k+1 \rfloor)} - X_{(\lfloor k \rfloor)} \right) (\lceil k \rceil - k)$, donde $\lfloor \cdot \rfloor$ es el entero inferior.

[4] $Q_1 = \left\{ X_{(k)} \mid k = \frac{n+1}{4} \right\}$. Si $k \notin \mathbb{N}$, entonces $Q_1 = X_{(\lfloor k \rfloor)} + \left(X_{(\lfloor k+1 \rfloor)} - X_{(\lfloor k \rfloor)} \right) (\lceil k \rceil - k)$, donde $\lfloor \cdot \rfloor$ es el entero inferior y $X_{(\lfloor k \rfloor)}$ es la k–ésima variable aleatoria en orden ascendente.

Ejemplo 81 (categorías). Las paridades del tipo de cambio peso mexicano y dólar estadounidense del periodo de 01/09/2018 a 08/11/2018 son publicadas en http://www.banxico.org.mx. Con las paridades del tipo de cambio:

1. Presentar la gráfica de los rendimientos logarítmicos $R_t = \ln\left(\frac{M_t}{M_{t-1}}\right)$.
2. Calcular el número de intervalos de clase para R_t por el método de Velleman (1976).
3. Presentar el histograma de frecuencias absolutas y relativas de R_t.
4. Presentar el polígono de frecuencias absolutas y relativas de R_t.
5. Presentar la ojiva de frecuencias absolutas y relativas de R_t.

Solución. Las paridades del tipo de cambio son presentadas en la tabla 6.5.

Tabla 6.5: Paridades del tipo de cambio peso mexicano y dólar estadounidense.

Fecha	**Paridad**	**Fecha**	**Paridad**	**Fecha**	**Paridad**
01/09/2018	19.1258	24/09/2018	18.8296	17/10/2018	18.8096
02/09/2018	19.1258	25/09/2018	18.8539	18/10/2018	18.7693
03/09/2018	19.1258	26/09/2018	18.8516	19/10/2018	18.8012
04/09/2018	19.1792	27/09/2018	18.9870	20/10/2018	19.1018
05/09/2018	19.2137	28/09/2018	18.8986	21/10/2018	19.1018
06/09/2018	19.3609	29/09/2018	18.8120	22/10/2018	19.1018
07/09/2018	19.4433	30/09/2018	18.8120	23/10/2018	19.2183
08/09/2018	19.3546	01/10/2018	18.8120	24/10/2018	19.3669
09/09/2018	19.3546	02/10/2018	18.7231	25/10/2018	19.4169
10/09/2018	19.3546	03/10/2018	18.6531	26/10/2018	19.4623
11/09/2018	19.2394	04/10/2018	18.7503	27/10/2018	19.4383
12/09/2018	19.2989	05/10/2018	18.8014	28/10/2018	19.4383
13/09/2018	19.2680	06/10/2018	19.1328	29/10/2018	19.4383
14/09/2018	19.0511	07/10/2018	19.1328	30/10/2018	19.4790
15/09/2018	18.8775	08/10/2018	19.1328	31/10/2018	19.8022
16/09/2018	18.8775	09/10/2018	18.9444	01/11/2018	20.0258
17/09/2018	18.8775	10/10/2018	18.9234	02/11/2018	20.3177
18/09/2018	18.8689	11/10/2018	19.0000	03/11/2018	20.3177
19/09/2018	18.8399	12/10/2018	19.1183	04/11/2018	20.3177
20/09/2018	18.7694	13/10/2018	19.0217	05/11/2018	20.3177
21/09/2018	18.7672	14/10/2018	19.0217	06/11/2018	20.1329
22/09/2018	18.8296	15/10/2018	19.0217	07/11/2018	19.9636
23/09/2018	18.8296	16/10/2018	18.9202	08/11/2018	19.8609

Con las 69 paridades del tipo de cambio son calculados los 68 rendimientos logarítmicos:

$$R_t = \ln\left(\frac{M_t}{M_{t-1}}\right) = \ln(M_t) - \ln(M_{t-1}),$$

donde R_t es el rendimiento logarítmico en el instante t, M_t es la paridad del tipo de cambio en el instante t y M_{t-1} es la paridad del tipo de cambio en el instante $t-1$.

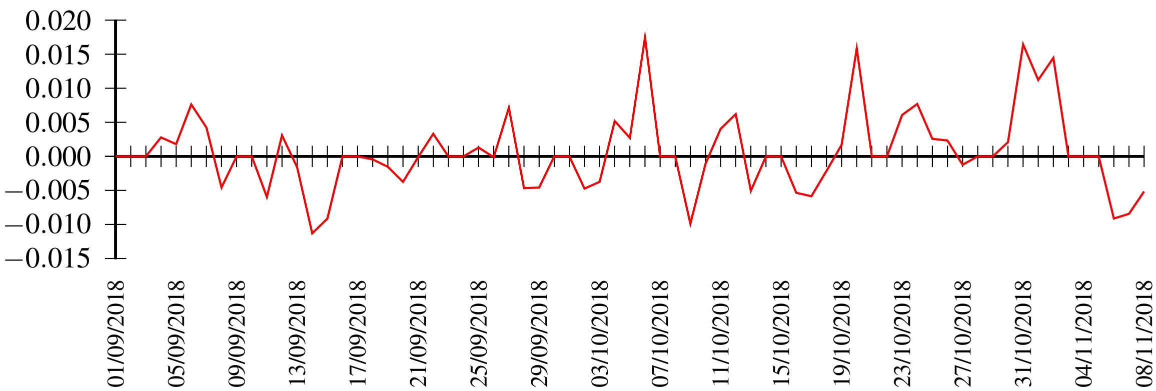
Los 68 rendimientos logarítmicos (con seis cifras significativas) son presentados en la tabla 6.6.

Tabla 6.6: Rendimientos logarítmicos de la paridad del tipo de cambio.

n	Rendimiento	n	Rendimiento	n	Rendimiento	n	Rendimiento
1	0.000000	18	-0.001538	35	0.017473	52	0.006080
2	0.000000	19	-0.003749	36	0.000000	53	0.007702
3	0.002788	20	-0.000117	37	0.000000	54	0.002578
4	0.001797	21	0.003319	38	-0.009896	55	0.002335
5	0.007632	22	0.000000	39	-0.001109	56	-0.001234
6	0.004247	23	0.000000	40	0.004040	57	0.000000
7	-0.004572	24	0.001290	41	0.006207	58	0.000000
8	0.000000	25	-0.000122	42	-0.005066	59	0.002092
9	0.000000	26	0.007157	43	0.000000	60	0.016456
10	-0.005970	27	-0.004667	44	0.000000	61	0.011228
11	0.003088	28	-0.004593	45	-0.005350	62	0.014471
12	-0.001602	29	0.000000	46	-0.005863	63	0.000000
13	-0.011321	30	0.000000	47	-0.002145	64	0.000000
14	-0.009154	31	-0.004737	48	0.001698	65	0.000000
15	0.000000	32	-0.003746	49	0.015862	66	-0.009137
16	0.000000	33	0.005197	50	0.000000	67	-0.008445
17	-0.000456	34	0.002722	51	0.000000	68	-0.005158

La tabla 6.6 presenta los rendimientos del tipo de cambio peso mexicano y dólar estadounidense del periodo de 01/09/2018 a 08/11/2018. Entonces:

1. Los rendimientos logarítmicos son presentados en la gráfica 6.8.



Gráfica 6.8: Rendimientos logarítmicos de la paridad peso y dólar.

2. El número de categorías[5] para R_t por el método de Velleman (1976) es obtenido sustituyendo en la ecuación (6.2):

$$k = \lceil \sqrt{68}\, \rceil = \lceil 8.2462 \rceil = 9.$$

[5]Por el método de Sturges $k = 8$, por el método de Scott $k = 6$ y por el método de Freedman y Diaconis $k = 14$.

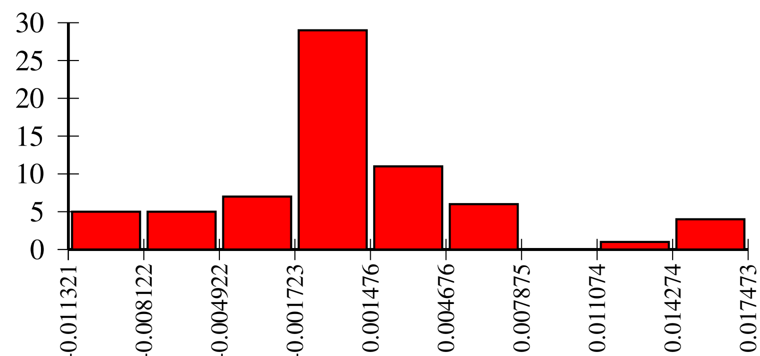
3. Los histogramas tienen nueve intervalos de clase que son construidos con una progresión aritmética donde la razón es obtenida sustituyendo en la ecuación (6.6):

$$\mathrm{d}k = \frac{\text{máx}(X) + u - \text{mín}(X)}{k} = \frac{0.017473 + 1 \times 10^{-16} + 0.011321}{9} = 0.003199,$$

donde $u = 1 \times 10^{-16}$. Entonces el rango está contenido en los intervalos de clase $C = \{C_1, C_2, \ldots, C_9\}$, donde:

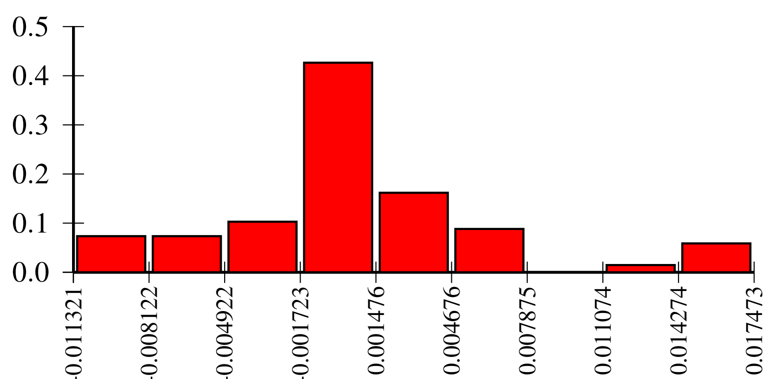
$$\begin{aligned} &C_1 = [-0.011321, -0.008122), && C_2 = [-0.008122, -0.004922), && C_3 = [-0.004922, -0.001723),\\ &C_4 = [-0.001723, 0.001476), && C_5 = [0.001476, 0.004676), && C_6 = [0.004676, 0.007875),\\ &C_7 = [0.007875, 0.011074), && C_8 = [0.011074, 0.014274), && C_9 = [0.014274, 0.017473]. \end{aligned}$$

El histograma de frecuencias absolutas de los rendimientos logarítmicos de la paridad peso y dólar es presentado en la gráfica 6.9.



Gráfica 6.9: Histograma de frecuencias absolutas de los rendimientos logarítmicos.

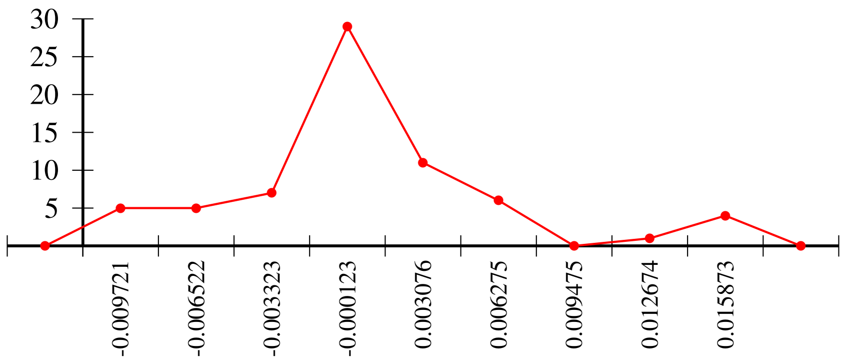
El histograma de frecuencias relativas de los rendimientos logarítmicos de la paridad peso y dólar es presentado en la gráfica 6.10.



Gráfica 6.10: Histograma de frecuencias relativas de los rendimientos logarítmicos.

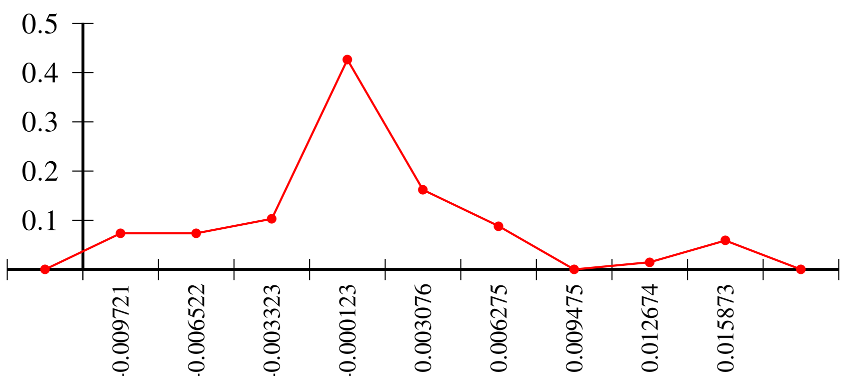
Las gráficas 6.9 y 6.10 presentan las frecuencias absolutas y relativas de los rendimientos logarítmicos de la paridad peso y dólar y se observa que son similares, la diferencia es la escala del eje de las ordenadas y las frecuencias relativas representan la probabilidad de ocurrencia de cada uno de los intervalos de clase.

4. El polígono de frecuencias absolutas de los rendimientos logarítmicos de la paridad peso y dólar es presentado en la gráfica 6.11.



Gráfica 6.11: Polígono de frecuencias absolutas de los rendimientos logarítmicos.

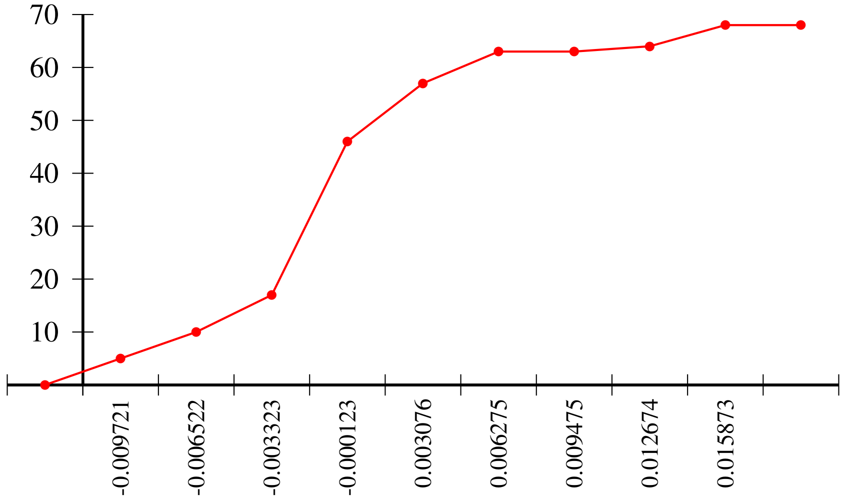
El polígono de frecuencias relativas de los rendimientos logarítmicos de la paridad peso y dólar es presentado en la gráfica 6.12.



Gráfica 6.12: Polígono de frecuencias relativas de los rendimientos logarítmicos.

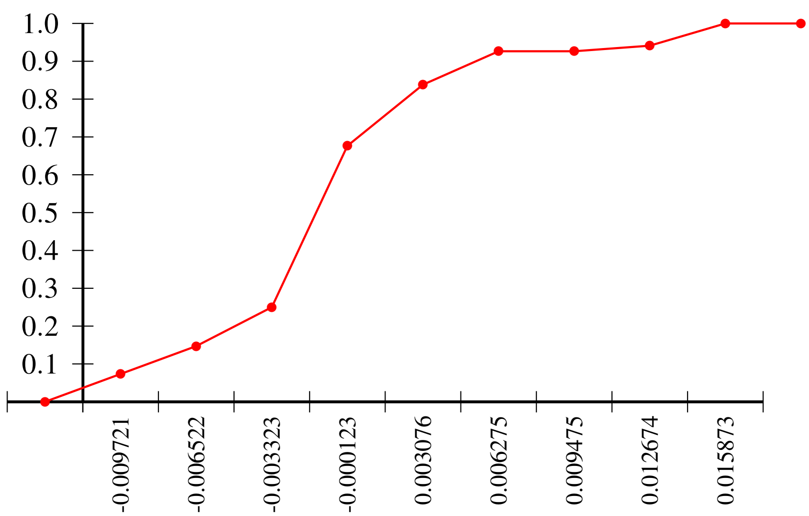
Las gráficas 6.11 y 6.12 presentan las frecuencias absolutas y relativas de los rendimientos logarítmicos de la paridad peso y dólar y se observa que son similares, la diferencia es la escala del eje de las ordenadas y las frecuencias relativas representan la probabilidad de ocurrencia de cada uno de los intervalos de clase.

5. La ojiva de frecuencias absolutas de los rendimientos logarítmicos de la paridad peso y dólar es presentada en la gráfica 6.13.



Gráfica 6.13: Ojiva de frecuencias absolutas de los rendimientos logarítmicos.

La ojiva de frecuencias relativas de los rendimientos logarítmicos de la paridad peso y dólar es presentada en la gráfica 6.14.



Gráfica 6.14: Ojiva de frecuencias relativas de los rendimientos logarítmicos.

Las gráficas 6.13 y 6.14 presentan las frecuencias absolutas y relativas acumuladas de los rendimientos logarítmicos de la paridad peso y dólar y se observa que son similares, la diferencia es la escala del eje de las ordenadas y las frecuencias relativas acumuladas representan la probabilidad de ocurrencia acumulada de cada uno de los intervalos de clase, es decir, la función de distribución observada. □

6.5. Ejercicios

Ejercicio 76. El grupo CTG82 de Probabilidad y Estadística tiene 24 hombres y 14 mujeres. Presentar:

1. Una gráfica de barras de las frecuencias absolutas.
2. Una gráfica de barras de las frecuencias relativas.
3. Un polígono de frecuencias absolutas.
4. Un polígono de frecuencias relativas.
5. Una ojiva de frecuencias absolutas.
6. Una ojiva de frecuencias relativas.
7. Una gráfica de pastel.

Ejercicio 77. Las frecuencias relativas de los grupos sanguíneos son presentadas en la tabla siguiente:

Grupo	**A**	**B**	**AB**	**O**
fr	0.20	0.15	0.03	0.62

Presentar:

1. Una gráfica de barras de las frecuencias absolutas suponiendo una muestra de 100 pacientes.
2. Una gráfica de barras de las frecuencias relativas.
3. Un polígono de frecuencias absolutas suponiendo una muestra de 100 pacientes.
4. Un polígono de frecuencias relativas.
5. Una ojiva de frecuencias absolutas suponiendo una muestra de 100 pacientes.
6. Una ojiva de frecuencias relativas.
7. Una gráfica de pastel.

Ejercicio 78. Cinco monedas son lanzadas simultáneamente 500 veces. Los resultados de los 500 lanzamientos son presentados en la tabla siguiente:

Caras	0	1	2	3	4	5
f	19	72	171	144	82	12

Presentar:

1. Un histograma de las frecuencias absolutas.
2. Un histograma de barras de las frecuencias relativas.
3. Un polígono de frecuencias absolutas.
4. Un polígono de frecuencias relativas.
5. Una ojiva de frecuencias absolutas.
6. Una ojiva de frecuencias relativas.
7. Una gráfica de pastel.

Ejercicio 79. Las distancias, en kilómetros, recorridas por 50 personas de su casa al trabajo son presentadas en la tabla siguiente:

6.82	10.40	11.52	12.32	8.00
11.20	13.92	14.08	3.84	16.43
12.80	1.44	12.48	12.85	19.68
6.24	1.48	7.84	12.88	6.04
5.92	20.16	3.20	7.36	6.08
13.44	6.40	4.80	2.24	10.56
4.16	16.48	6.72	3.52	3.24
1.60	16.00	5.28	3.04	2.56
25.12	9.92	7.64	5.12	7.04
6.29	1.76	7.04	7.68	6.88

1. Calcular el número de clases por los métodos:
   a) Velleman (1976).
   b) Sturges (1926).
   c) Scott (1979).
   d) Freedman y Diaconis (1981).
2. Para cada uno de los métodos (Velleman, Sturges, Scott y Freedman y Diaconis) presentar en una tabla y para cada una de las categorías:
   a) Los límites de los intervalos de clase.
   b) Las marcas de clase.
   c) Las frecuencias absolutas.
   d) Las frecuencias relativas.
   e) Las frecuencias absolutas acumuladas.
   f) Las frecuencias relativas acumuladas.
3. Para el método con el número máximo de intervalos de clase presentar:
   a) Un histograma de las frecuencias absolutas.
   b) Un histograma de barras de las frecuencias relativas.
   c) Un polígono de frecuencias absolutas.
   d) Un polígono de frecuencias relativas.
   e) Una ojiva de frecuencias absolutas.
   f) Una ojiva de frecuencias relativas.
   g) Una gráfica de pastel.

Ejercicio 80. Las densidades de maderas utilizadas en la construcción son presentadas en la tabla siguiente:

0.31	0.35	0.36	0.36	0.37	0.38	0.40	0.40	0.40
0.41	0.41	0.42	0.42	0.42	0.42	0.42	0.43	0.44
0.45	0.46	0.46	0.47	0.48	0.48	0.48	0.51	0.54
0.54	0.55	0.58	0.62	0.66	0.66	0.67	0.68	0.75

1. Calcular el número de clases por los métodos:
   a) Velleman (1976).
   b) Sturges (1926).
   c) Scott (1979).
   d) Freedman y Diaconis (1981).
2. Para cada uno de los métodos (Velleman, Sturges, Scott y Freedman y Diaconis) presentar en una tabla y para cada una de las categorías:
   a) Los límites de los intervalos de clase.
   b) Las marcas de clase.
   c) Las frecuencias absolutas.
   d) Las frecuencias relativas.
   e) Las frecuencias absolutas acumuladas.
   f) Las frecuencias relativas acumuladas.
3. Para el método con el número máximo de intervalos de clase presentar:
   a) Un histograma de las frecuencias absolutas.
   b) Un histograma de barras de las frecuencias relativas.
   c) Un polígono de frecuencias absolutas.
   d) Un polígono de frecuencias relativas.
   e) Una ojiva de frecuencias absolutas.
   f) Una ojiva de frecuencias relativas.
   g) Una gráfica de pastel.

Capítulo 7

Estimación

La estimación de parámetros a partir de estadísticas es un problema de la inferencia estadística.

Definición 104 (espacio de la muestra). El espacio de la muestra es el conjunto $\mathfrak{X}$ que está formado por los valores de la muestra aleatoria $X_1, \ldots, X_n$.

Definición 105 (estadística). Una estadística es una función $T(X_1, \ldots, X_n)$ de una muestra aleatoria $X_1, \ldots, X_n$ que no depende de parámetros desconocidos, por lo tanto la estadística $T(\underline{X})$ es una variable aleatoria.

Definición 106 (espacio paramétrico). El espacio paramétrico es el conjunto Θ que está formado por los valores del parámetro θ.

Definición 107 (parámetro). Un parámetro θ es un valor que expresa una característica de una población.

La estimación de un parámetro θ es realizada a través de:

1. La estimación puntual.
2. La estimación por intervalos.

7.1. Estimación puntual

Sea X una variable aleatoria que tiene una distribución de probabilidad con función de densidad $f_X(x, \theta)$ y $\mathscr{E}$ un experimento aleatorio donde θ es el conjunto de parámetros de la distribución.

Definición 108 (estimador puntual). Un estimador puntual $\hat{\theta}$ es una función $T(\underline{X})$ de una muestra aleatoria $X_1, \ldots, X_n$ que es utilizada para estimar al conjunto Θ de parámetros θ:

$$\hat{\theta} = T(\underline{X}) = f(x_1, \ldots, x_n), \tag{7.1}$$

donde $\hat{\theta}$ es el estimador del espacio de parámetros Θ de la distribución.

Los estimadores puntuales son calculados con:

1. El método de momentos de Pearson (1900).
2. El método de máxima verosimilitud de Fisher (1922).

7.1.1. Método de momentos

Definición 109 (r–ésimo momento de la población). Sea $f_X(x, \theta)$ la función de densidad de una variable aleatoria X que depende de un parámetro θ, entonces el r–ésimo momento poblacional de X es:

$$\mu_r = \mathrm{E}\left(X^r\right) \tag{7.2}$$

cuando $\mathrm{E}\left(X^r\right) < \infty$.

Definición 110 (r–ésimo momento de la muestra). Sea $f_X(x, \theta)$ la función de densidad de una variable aleatoria X que depende de un parámetro θ, entonces el r–ésimo momento muestral de X es:

$$M_r = \frac{1}{n} \sum_{k=1}^{n} X_k^r. \tag{7.3}$$

El método de momentos propuesto por Pearson (1900) consiste en igualar los momentos de la muestra con los momentos de la población y resolver el sistema de ecuaciones para el conjunto de parámetros θ y la solución es el estimador de momentos.

7.1.2. Método de máxima verosimilitud

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población con función de densidad $f_X(x, \theta)$, entonces las variables de la muestra aleatoria tienen una función de densidad $f_X(x)$ que depende del conjunto de parámetros θ.

Definición 111 (función de verosimilitud). Sea $X_1, \ldots, X_n$ una muestra aleatoria, entonces la función $\mathscr{L}(\theta)$ es la función de densidad conjunta:

$$\mathscr{L}(\theta) = f_{X_1,\ldots,X_n}(x_1, \ldots, x_n, \theta) = \prod_{k=1}^{n} f_{X_k}(x_k, \theta). \tag{7.4}$$

La función $\mathscr{L}$ proviene del término en inglés *likelihood* que es traducido como verosimilitud o credibilidad. El método de máxima verosimilitud propuesto por Fisher (1922) consiste en obtener el valor de θ para maximizar la función de verosimilitud y encontrar el conjunto de valores de θ, de tal forma que los datos observados maximizan la probabilidad de ocurrencia. El conjunto de valores de θ donde se maximiza $\mathscr{L}$ es el estimador de máxima verosimilitud.

Definición 112 (estimador de máxima verosimilitud). Sea $X_1, \ldots, X_n$ una muestra aleatoria con la función de densidad $f_X(x, \theta)$ y $\mathscr{L}(\theta)$ la función de verosimilitud, entonces $\hat{\theta} = T(\underline{X})$ es el estimador máximo verosímil de θ si $\forall \theta \in \Theta$ se satisface $\mathscr{L}\left(\hat{\theta}\right) \geq \mathscr{L}(\theta)$.

7.2. Propiedades de los estimadores puntuales

Fisher (1922) propuso que los estimadores son:

1. Insesgados.
2. Eficientes.
3. Consistentes.
4. Suficientes.

7.2.1. Estimador insesgado

Definición 113 (estimador insesgado). Un estimador $\hat{\theta}$ de un parámetro $\theta \in \Theta$ es insesgado si:

$$\mathrm{E}\left(\hat{\theta}\right) = \theta. \tag{7.5}$$

Un estimador insesgado es adecuado porque si el objetivo es estimar el valor del parámetro, entonces es adecuado saber que el valor esperado del estimador coincide con el valor a estimar.

7.2.2. Estimador eficiente

Definición 114 (estimador eficiente). Si $\hat{\theta}_1$ y $\hat{\theta}_2$ son estimadores insesgados de un parámetro $\theta \in \Theta$, entonces $\hat{\theta}_1$ es un estimador más eficiente que $\hat{\theta}_2$ si:

$$\mathrm{E}\left(\hat{\theta}_1 - \theta\right)^2 = \mathrm{Var}\left(\hat{\theta}_1\right) < \mathrm{Var}\left(\hat{\theta}_2\right) = \mathrm{E}\left(\hat{\theta}_2 - \theta\right)^2. \tag{7.6}$$

Un estimador insesgado con menor varianza es preferido sobre un estimador insesgado con mayor varianza.

7.2.3. Estimador consistente

Definición 115 (estimador consistente). Un estimador $\hat{\theta}_n$ de un parámetro $\theta \in \Theta$ es consistente si $\forall \varepsilon > 0$:

$$\lim_{n \to \infty} \mathscr{P}\left(\left|\hat{\theta}_n - \theta\right| > \varepsilon\right) = 0. \tag{7.7}$$

Un estimador consistente converge al parámetro si converge en probabilidad (medida): $\lim_{n \to \infty} \hat{\theta}_n \xrightarrow{\mathscr{P}} \theta$.

7.2.4. Estimador suficiente

Definición 116 (estimador suficiente). Un estimador $\hat{\theta} = T$ de un parámetro $\theta \in \Theta$ es suficiente si la distribución conjunta $F_{X_1,\ldots,X_n}\left(x_1,\ldots,x_n \mid T = t\right)$ de la muestra aleatoria $X_1,\ldots,X_n$ no depende del parámetro θ.

Un estimador suficiente dado un valor t de la estadística T a partir de una muestra aleatoria $\underline{X}$ no contiene información adicional acerca del parámetro θ que la proporcionada por la estadística T.

7.3. Medidas de tendencia central

Las medidas de localización de una variable aleatoria son:

1. La media (aritmética, ponderada, geométrica y armónica).
2. La mediana.
3. La moda.

La medida de tendencia central más utilizada es la media aritmética, representa el promedio de la variable aleatoria y es el centro de gravedad o el punto de equilibrio.

7.3.1. Media aritmética

La media aritmética es el promedio aritmético.

Definición 117 (media aritmética de la población). La media aritmética de la población es:

$$\mu = \frac{1}{N}\sum_{k=1}^{N} X_k = \frac{1}{N}\sum_{k=1}^{N} X_k f_k = \sum_{k=1}^{N} X_k fr_k, \tag{7.8}$$

donde μ es la media de la población, N es el tamaño de la población, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 118 (media aritmética de la población agrupada). La media aritmética de la población agrupada es:

$$\mu = \frac{1}{N}\sum_{j=1}^{k} \bar{C}_j f_j = \sum_{j=1}^{k} \bar{C}_j fr_j, \tag{7.9}$$

donde μ es la media de la población agrupada, N es el tamaño de la población, f_j es la frecuencia absoluta del j–ésimo intervalo de clase, $\bar{C}_j$ es la j–ésima marca de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Definición 119 (media aritmética de la muestra). La media aritmética de la muestra es:

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{1}{n}\sum_{k=1}^{n} X_k f_k = \sum_{k=1}^{n} X_k fr_k, \tag{7.10}$$

donde $\bar{X}$ es la media de la muestra, n es el tamaño de la muestra, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 120 (media aritmética de la muestra agrupada). La media aritmética de la muestra agrupada es:

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{k} \bar{C}_j f_j = \sum_{j=1}^{k} \bar{C}_j fr_j, \tag{7.11}$$

donde $\bar{X}$ es la media de la muestra agrupada, n es el tamaño de la muestra, f_j es la frecuencia absoluta del j–ésimo intervalo de clase, $\bar{C}_j$ es la j–ésima marca de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

7.3.2. Media ponderada

La media ponderada es el promedio ponderado donde cada variable aleatoria tiene una ponderación de acuerdo con la relevancia relativa.

Definición 121 (media ponderada de la población). La media ponderada de la población es:

$$\mu_w = \frac{\sum_{k=1}^{N} w_k X_k}{\sum_{k=1}^{N} w_k} = \frac{\sum_{k=1}^{N} w_k X_k f_k}{\sum_{k=1}^{N} w_k f_k} = \frac{\sum_{k=1}^{N} w_k X_k fr_k}{\sum_{k=1}^{N} w_k fr_k}, \tag{7.12}$$

donde μ_w es la media ponderada de la población, w_k es la ponderación de la k–ésima variable aleatoria, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 122 (media ponderada de la población agrupada). La media ponderada de la población agrupada es:

$$\mu_w = \frac{\sum_{j=1}^{k} w_j \bar{C}_j f_j}{\sum_{j=1}^{k} w_j f_j} = \frac{\sum_{j=1}^{k} w_j \bar{C}_j fr_j}{\sum_{j=1}^{k} w_j fr_j}, \tag{7.13}$$

donde μ_w es la media ponderada de la población agrupada, w_j es la ponderación de la j–ésimo intervalo de clase, $\bar{C}_j$ es la j–ésima marca de clase, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Definición 123 (media ponderada de la muestra). La media ponderada de la muestra es:

$$\bar{X}_w = \frac{\sum_{k=1}^{n} w_k X_k}{\sum_{k=1}^{n} w_k} = \frac{\sum_{k=1}^{n} w_k X_k f_k}{\sum_{k=1}^{n} w_k f_k} = \frac{\sum_{k=1}^{n} w_k X_k fr_k}{\sum_{k=1}^{n} w_k fr_k}, \tag{7.14}$$

donde $\bar{X}_w$ es la media ponderada de la muestra, w_k es la ponderación de la k–ésima variable aleatoria, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 124 (media ponderada de la muestra agrupada). La media ponderada de la muestra agrupada es:

$$\bar{X}_w = \frac{\sum_{j=1}^{k} w_j \bar{C}_j f_j}{\sum_{j=1}^{k} w_j f_j} = \frac{\sum_{j=1}^{k} w_j \bar{C}_j fr_j}{\sum_{j=1}^{k} w_j fr_j}, \tag{7.15}$$

donde $\bar{X}_w$ es la media ponderada de la muestra agrupada, w_j es la ponderación de la j–ésimo intervalo de clase, $\bar{C}_j$ es la j–ésima marca de clase, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

7.3.3. Media geométrica

La media geométrica es el promedio del cambio porcentual de una variable aleatoria estrictamente positiva.

Definición 125 (media geométrica de la población). Si $X_k > 0$, entonces la media geométrica de la población es:

$$\mu_G = \left(\prod_{k=1}^{N} X_k\right)^{\frac{1}{N}} = \left(\prod_{k=1}^{N} X_k^{f_k}\right)^{\frac{1}{N}} = \prod_{k=1}^{N} X_k^{fr_k}, \tag{7.16}$$

donde μ_G es la media geométrica de la población, N es el tamaño de la población, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 126 (media geométrica de la población agrupada). Si $\bar{C}_j > 0$, entonces la media geométrica de la población agrupada es:

$$\mu_G = \left(\prod_{j=1}^{k} \bar{C}_j^{f_j}\right)^{\frac{1}{N}} = \prod_{j=1}^{k} \bar{C}_j^{fr_j}, \tag{7.17}$$

donde μ_G es la media geométrica de la población agrupada, $\bar{C}_j$ es la j–ésima marca de clase, f_j es la frecuencia absoluta del j–ésimo intervalo de clase, N es el tamaño de la población y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Definición 127 (media geométrica de la muestra). Si $X_k > 0$, entonces la media geométrica de la muestra es:

$$\bar{X}_G = \left(\prod_{k=1}^{n} X_k\right)^{\frac{1}{n}} = \left(\prod_{k=1}^{n} X_k^{f_k}\right)^{\frac{1}{n}} = \prod_{k=1}^{n} X_k^{fr_k}, \tag{7.18}$$

donde $\bar{X}_G$ es la media geométrica de la muestra, n es el tamaño de la muestra, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 128 (media geométrica de la muestra agrupada). Si $\bar{C}_j$, entonces la media geométrica de la muestra agrupada es:

$$\bar{X}_G = \left(\prod_{j=1}^{k} \bar{C}_j^{f_j}\right)^{\frac{1}{n}} = \prod_{j=1}^{k} \bar{C}_j^{fr_j}, \tag{7.19}$$

donde $\bar{X}_G$ es la media geométrica de la muestra agrupada, $\bar{C}_j$ es la j–ésima marca de clase, f_j es la frecuencia absoluta del j–ésimo intervalo de clase, n es el tamaño de la muestra y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

7.3.4. Media armónica

La media armónica es el inverso del promedio de los inversos de las variables aleatorias, es decir, es el inverso de la media aritmética de los inversos de las variables aleatorias.

Definición 129 (media armónica de la población). Si $X_k \neq 0$, entonces la media armónica de la población es:

$$\mu_A = \frac{N}{\sum\limits_{k=1}^{N} X_k^{-1}} = \frac{N}{\sum\limits_{k=1}^{N} X_k^{-1} f_k} = \frac{1}{\sum\limits_{k=1}^{N} X_k^{-1} fr_k}, \tag{7.20}$$

donde μ_A es la media armónica de la población, N es el tamaño de la población, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 130 (media armónica de la población agrupada). Si $\bar{C}_j \neq 0$, entonces la media armónica de la población agrupada es:

$$\mu_A = \frac{N}{\sum\limits_{j=1}^{k} \bar{C}_j^{-1} f_j} = \frac{1}{\sum\limits_{j=1}^{k} \bar{C}_j^{-1} fr_j}, \tag{7.21}$$

donde μ_A es la media armónica de la población agrupada, N es el tamaño de la población, $\bar{C}_j$ es la j–ésima marca de clase, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Definición 131 (media armónica de la muestra). Si $X_k \neq 0$, entonces la media armónica de la muestra es:

$$\bar{X}_A = \frac{n}{\sum\limits_{k=1}^{n} X_k^{-1}} = \frac{n}{\sum\limits_{k=1}^{n} X_k^{-1} f_k} = \frac{1}{\sum\limits_{k=1}^{n} X_k^{-1} fr_k}, \tag{7.22}$$

donde $\bar{X}_A$ es la media armónica de la muestra, N es el tamaño de la población, X_k es la k–ésima variable aleatoria, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 132 (media armónica de la muestra agrupada). Si $\bar{C}_j \neq 0$, entonces la media armónica de la muestra agrupada es:

$$\bar{X}_A = \frac{n}{\sum\limits_{j=1}^{k} \bar{C}_j^{-1} f_j} = \frac{1}{\sum\limits_{j=1}^{k} \bar{C}_j^{-1} fr_j}, \tag{7.23}$$

donde $\bar{X}_A$ es la media armónica de la muestra agrupada, N es el tamaño de la población, $\bar{C}_j$ es la j–ésima marca de clase, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Nota 14 (relación entre la media aritmética, geométrica y armónica). La relación de la media aritmética, la media geométrica y la media armónica satisface:

$$\begin{aligned} \mu &\geq \mu_G \geq \mu_A, \\ \bar{X} &\geq \bar{X}_G \geq \bar{X}_A, \end{aligned}$$

donde μ es la media aritmética de la población, μ_G es la media geométrica de la población, μ_A es la media armónica de la población, $\bar{X}$ es la media aritmética de la muestra, $\bar{X}_G$ es la media geométrica de la muestra y $\bar{X}_A$ es la media armónica de la muestra.

7.3.5. Mediana

La mediana divide a la función de densidad de la variable aleatoria en dos partes con áreas iguales.

Definición 133 (mediana de la población). La mediana de la población es:

$$\tilde{\mu} = \begin{cases} X_{\left(\frac{N+1}{2}\right)} & \text{si } N \text{ es un número impar} \\ \frac{1}{2}\left(X_{\left(\frac{N}{2}\right)} + X_{\left(\frac{N}{2}+1\right)}\right) & \text{si } N \text{ es un número par} \end{cases}, \tag{7.24}$$

donde $\tilde{\mu}$ es la mediana de la población, N es el tamaño de la población y $X_{(k)}$ es la k–ésima variable aleatoria en orden ascendente.

Definición 134 (mediana de la población agrupada). La mediana de la población agrupada es:

$$\tilde{\mu} = \text{mín}\left(C_{\tilde{\mu}}\right) + \left(\frac{\frac{N}{2} - F_{\tilde{\mu}-1}}{f_{\tilde{\mu}}}\right) \mathrm{d}k, \tag{7.25}$$

donde $\tilde{\mu}$ es la mediana de la población agrupada, $\text{mín}\left(C_{\tilde{\mu}}\right)$ es el mínimo del intervalo de clase que contiene a la mediana, N es el tamaño de la población, $F_{\tilde{\mu}-1}$ es la frecuencia absoluta acumulada del intervalo de clase anterior al que contiene a la mediana, $f_{\tilde{\mu}}$ es la frecuencia absoluta del intervalo de clase que contiene a la mediana y $\mathrm{d}k$ es la amplitud de los intervalos de clase.

Definición 135 (mediana de la muestra). La mediana de la muestra es:

$$\tilde{X} = \begin{cases} X_{\left(\frac{n+1}{2}\right)} & \text{si } n \text{ es un número impar} \\ \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right) & \text{si } n \text{ es un número par} \end{cases}, \tag{7.26}$$

donde $\tilde{X}$ es la mediana de la muestra, n es el tamaño de la muestra y $X_{(k)}$ es la k–ésima variable aleatoria en orden ascendente.

Definición 136 (mediana de la muestra agrupada). La mediana de la muestra agrupada es:

$$\tilde{X} = \text{mín}\left(C_{\tilde{X}}\right) + \left(\frac{\frac{n}{2} - F_{\tilde{X}-1}}{f_{\tilde{X}}}\right) \mathrm{d}k, \tag{7.27}$$

donde $\tilde{X}$ es la mediana de la muestra agrupada, $\text{mín}\left(C_{\tilde{X}}\right)$ es el mínimo del intervalo de clase que contiene a la mediana, n es el tamaño de la muestra, $F_{\tilde{X}-1}$ es la frecuencia absoluta acumulada del intervalo de clase anterior al que contiene a la mediana, $f_{\tilde{X}}$ es la frecuencia absoluta del intervalo de clase que contiene a la mediana y $\mathrm{d}k$ es la amplitud de los intervalos de clase.

7.3.6. Moda

La moda es la variable aleatoria que se presenta con mayor frecuencia.

Definición 137 (moda de la población). La moda de la población es:

$$\hat{\mu} = \left\{ X_k \mid X_k = \max_N (f_k) \right\}, \tag{7.28}$$

donde $\hat{\mu}$ es la moda de la población, $\max_N (f_k)$ es la variable aleatoria X_k con la frecuencia máxima y N es el tamaño de la población.

Definición 138 (moda de la población para datos agrupados). La moda de la población para datos agrupados es:

$$\hat{\mu} = \min \left(C_{\hat{\mu}} \right) + \left(\frac{f_{\hat{\mu}} - f_{\hat{\mu}-1}}{f_{\hat{\mu}} - f_{\hat{\mu}-1} + f_{\hat{\mu}} - f_{\hat{\mu}+1}} \right) \mathrm{d}k, \tag{7.29}$$

donde $\hat{\mu}$ es la moda de la población agrupada, $\min \left(C_{\hat{\mu}} \right)$ es el mínimo del intervalo de clase que contiene a la moda, $f_{\hat{\mu}}$ es la frecuencia absoluta del intervalo de clase que contiene a la moda, $f_{\hat{\mu}-1}$ es la frecuencia absoluta del intervalo de clase anterior al que contiene a la moda, $f_{\hat{\mu}+1}$ es la frecuencia absoluta del intervalo de clase posterior al que contiene a la moda y $\mathrm{d}k$ es la amplitud de los intervalos de clase.

Definición 139 (moda de la muestra). La moda de la muestra es:

$$\hat{X} = \left\{ X_k \mid X_k = \max_n (f_k) \right\}, \tag{7.30}$$

donde $\hat{X}$ es la moda de la población, $\max_n (f_k)$ es la variable aleatoria X_k con la frecuencia máxima y n es el tamaño de la muestra.

Definición 140 (moda de la muestra para datos agrupados). La moda de la muestra para datos agrupados es:

$$\hat{X} = \min (C_{\hat{X}}) + \left(\frac{f_{\hat{X}} - f_{\hat{X}-1}}{f_{\hat{X}} - f_{\hat{X}-1} + f_{\hat{X}} - f_{\hat{X}+1}} \right) \mathrm{d}k, \tag{7.31}$$

donde $\hat{X}$ es la moda de la muestra agrupada, $\min (C_{\hat{X}})$ es el mínimo del intervalo de clase que contiene a la moda, $f_{\hat{X}}$ es la frecuencia absoluta del intervalo de clase que contiene a la moda, $f_{\hat{X}-1}$ es la frecuencia absoluta del intervalo de clase anterior al que contiene a la moda, $f_{\hat{X}+1}$ es la frecuencia absoluta del intervalo de clase posterior al que contiene a la moda y $\mathrm{d}k$ es la amplitud de los intervalos de clase.

La moda puede ser inexistente. Si la variable aleatoria con la frecuencia máxima es única, entonces la distribución es unimodal. Si son dos variables aleatorias con frecuencia máxima, entonces la distribución es bimodal y si son más de dos variables aleatorias con frecuencia máxima, entonces la distribución es multimodal.

Ejemplo 82 (medidas de tendencia central). Con la información presentada en la tabla 7.1.

Tabla 7.1: Créditos, calificaciones y frecuencias absolutas.

C_k	Créditos	$\bar{C}_k$	f_k
[6.0,6.5)	10	6	1
[6.0,6.5)	18	6	1
[6.5,7.5)	8	7	2
[6.5,7.5)	10	7	1
[7.5,8.5)	8	8	3
[7.5,8.5)	10	8	2
[7.5,8.5)	12	8	1
[7.5,8.5)	18	8	2
[8.5,9.5)	10	9	2
[9.5,10.0)	6	10	2
[9.5,10.0)	8	10	3
[9.5,10.0)	10	10	16
[9.5,10.0)	18	10	1

Calcular:

1. La media aritmética de las calificaciones obtenidas en la licenciatura.
2. La media ponderada de las calificaciones considerando los créditos de cada asignatura.
3. La media geométrica de las calificaciones.
4. La media armónica de las calificaciones.
5. La mediana de las calificaciones (sin agrupar en categorías).
6. La moda de las calificaciones (sin agrupar en categorías).

Solución. El tamaño de la muestra es $n = 37$ y los datos sin considerar los créditos son:

X_k	6	7	8	9	10
f_k	2	3	8	2	22

1. La media aritmética de las calificaciones calculada con la ecuación (7.10):

$$\bar{X} = \frac{1}{37}\sum_{k=1}^{37} X_k f_k = \frac{12+21+64+18+220}{37} = \frac{335}{37} = 9.0541.$$

   La media aritmética de las calificaciones calculada con la ecuación (7.11):

$$\bar{X} = \frac{1}{37}\sum_{j=1}^{5} \bar{C}_j f_j = \frac{12+21+64+18+220}{37} = \frac{335}{37} = 9.0541.$$

2. La media ponderada de las calificaciones considerando los créditos, calculada con la ecuación (7.14):

$$\bar{X}_w = \frac{\sum_{k=1}^{37} w_k X_k f_k}{\sum_{k=1}^{n} w_k f_k} = \frac{3,406}{380} = 8.9632.$$

La media ponderada de las calificaciones considerando los créditos, calculada con la ecuación (7.15):

$$\bar{X}_w = \frac{\sum_{j=1}^{13} w_k \bar{C}_j f_j}{\sum_{j=1}^{13} w_j f_j} = \frac{3,406}{380} = 8.9632.$$

3. La media geométrica de las calificaciones con la ecuación (7.18):

$$\bar{X}_G = \left(\prod_{k=1}^{37} X_k^{f_k} \right)^{\frac{1}{37}} = \left(6^2 7^3 8^8 9^2 10^{22}\right)^{\frac{1}{37}} = \left(1.6780 \times 10^{35}\right)^{\frac{1}{37}} = 8.9541.$$

La media geométrica de las calificaciones con la ecuación (7.19):

$$\bar{X}_G = \left(\prod_{j=1}^{5} \bar{C}_j^{f_k} \right)^{\frac{1}{37}} = \left(6^2 7^3 8^8 9^2 10^{22}\right)^{\frac{1}{37}} = \left(1.6780 \times 10^{35}\right)^{\frac{1}{37}} = 8.9541.$$

4. La media armónica de las calificaciones con la ecuación (7.22):

$$\bar{X}_A = \frac{37}{\sum_{k=1}^{37} X_k^{-1} f_k} = \frac{37}{\frac{2}{6} + \frac{3}{7} + \frac{8}{8} + \frac{2}{9} + \frac{22}{10}} = 8.8429.$$

La media armónica de las calificaciones con la ecuación (7.23):

$$\bar{X}_A = \frac{n}{\sum_{j=1}^{k} \bar{C}_j^{-1} f_j} = \frac{37}{\frac{2}{6} + \frac{3}{7} + \frac{8}{8} + \frac{2}{9} + \frac{22}{10}} = 8.8429.$$

Por lo tanto, $\bar{X} \geq \bar{X}_G \geq \bar{X}_A$.

5. La mediana de las calificaciones con la ecuación (7.30):

$$\tilde{X} = X_{(19)} = 10.$$

Por lo tanto, la mediana de las calificaciones es $\tilde{X} = 10$ porque la décima novena variable aleatoria en orden ascendente es $X_{(19)} = 10$.

6. La moda de las calificaciones con la ecuación (7.30):

$$\hat{X} = \left\{ X_k \mid X_k = \max_n (f_k) \right\} = \left\{ X_k \mid X_k = \max_n (2,3,8,2,22) \right\} = 10.$$

Por lo tanto, la moda de las calificaciones es $\hat{X} = 10$ porque la frecuencia absoluta máxima pertenece a la calificación 10, es decir, $f_{10} = 22$. $\square$

Ejemplo 83 (medidas de tendencia central). Con los rendimientos logarítmicos de las paridades del tipo de cambio peso mexicano y dólar estadounidense del periodo de 01/09/2018 a 08/11/2018, presentados en la tabla 6.6 del ejemplo 81, calcular:

1. La media muestral $\bar{X}$ y la media muestral para datos agrupados $\bar{X}$.
2. La mediana muestral $\tilde{X}$ y la mediana muestral para datos agrupados $\tilde{X}$.
3. La moda muestral $\hat{X}$ y la moda muestral para datos agrupados $\hat{X}$.

Solución. Los rendimientos logarítmicos son presentados en la tabla 6.6 y las medidas de tendencia central para la muestra de los 68 rendimientos son:

1. La media de la muestra es calculada substituyendo en la ecuación (7.10):

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{0.000000 + \cdots - 0.005158}{68} = \frac{0.037715}{68} = 0.000555.$$

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k f_k = \frac{-0.011321 + \cdots + 21\,(0.000000) + \cdots + 0.017473}{68} = \frac{0.037715}{68} = 0.000555.$$

$$\bar{X} = \sum_{k=1}^{n} X_k fr_k = -0.000166 + \cdots + 0.000257 = 0.000555.$$

La media de la muestra para datos agrupados es calculada substituyendo en la ecuación (7.11):

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{k} \bar{C}_j f_j = \frac{1}{68}\left(-0.048606 - 0.032610 + \cdots + 0.063493\right) = \frac{0.039604}{68} = 0.000582.$$

$$\bar{X} = \sum_{j=1}^{k} \bar{C}_j fr_j = -0.000715 - 0.000480 + \cdots + 0.000582 = 0.000582.$$

2. La mediana de la muestra es calculada substituyendo en la ecuación (7.26):

$$\tilde{X} = \frac{1}{2}\left(X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}\right) = \frac{1}{2}\left(X_{(34)} + X_{(35)}\right) = \frac{0+0}{2} = 0.$$

La mediana de la muestra para datos agrupados es calculada substituyendo en la ecuación (7.27):

$$\tilde{X} = \text{mín}\left(C_{\tilde{X}}\right) + \left(\frac{\frac{n}{2} - F_{\tilde{X}-1}}{f_{\tilde{X}}}\right) \mathrm{d}k = -0.001723 + 0.003199\left(\frac{\frac{68}{2} - 17}{29}\right) = 0.000152.$$

3. La moda muestral es calculada substituyendo en la ecuación (7.30):

$$\hat{X} = \left\{X_k \mid X_k = \max_{n}\left(f_k\right)\right\} = \left\{X_k \mid X_k = \max_{68}\left(f_k\right) = 21\right\} = \left\{X_k \mid X_{(25-45)} = 0\right\}.$$

La moda muestral para datos agrupados es calculada substituyendo en la ecuación (7.31):

$$\hat{X} = \text{mín}\left(C_{\hat{X}}\right) + \left(\frac{f_{\hat{X}} - f_{\hat{X}-1}}{f_{\hat{X}} - f_{\hat{X}-1} + f_{\hat{X}} - f_{\hat{X}+1}}\right) \mathrm{d}k.$$

$$\hat{X} = -0.001723 + 0.003199\left(\frac{29-7}{29-7+29-11}\right) = 0.000037.$$

Por lo tanto, las medidas de localización varían cuando la información es agrupada. □

7.4. Medidas de dispersión

Las medidas de dispersión muestran la variación de la distribución de una variable aleatoria con respecto a la media. Las medidas de dispersión son:

1. El rango.
2. La varianza.
3. La desviación estándar

7.4.1. Rango

La dispersión de una variable aleatoria es la diferencia entre el valor máximo y el valor mínimo.

Definición 141 (rango de la población). El rango de la población es:

$$\operatorname*{ran}_N(X) = \operatorname*{máx}_N(X_k) - \operatorname*{mín}_N(X_k) = X_{(N)} - X_{(1)}, \tag{7.32}$$

donde $\operatorname*{ran}_N(X)$ es el rango poblacional, $\operatorname*{máx}_N(X_k)$ es el valor máximo de las N variables aleatorias, $\operatorname*{mín}_N(X_k)$ es el valor mínimo de las N variables aleatorias, $X_{(N)}$ es el valor máximo y $X_{(1)}$ es el valor mínimo con las variables aleatorias ordenadas ascendentemente.

Definición 142 (rango de la población para datos agrupados). El rango poblacional para datos agrupados es:

$$\operatorname*{ran}_N(X) = \operatorname*{máx}_k(C_k) - \operatorname*{mín}_k(C_1), \tag{7.33}$$

donde $\operatorname*{ran}_N(X)$ es el rango poblacional para datos agrupados, $\operatorname*{máx}_k(C_k)$ es el valor máximo del k–ésimo intervalo de clase y $\operatorname*{mín}_k(C_1)$ es el valor mínimo del primer intervalo de clase.

Definición 143 (rango de la muestra). El rango de la muestra es:

$$\operatorname*{ran}_n(X) = \operatorname*{máx}_n(X_k) - \operatorname*{mín}_n(X_k) = X_{(n)} - X_{(1)}, \tag{7.34}$$

donde $\operatorname*{ran}_n(X)$ es el rango muestral, $\operatorname*{máx}_n(X_k)$ es el valor máximo de las n variables aleatorias, $\operatorname*{mín}_n(X_k)$ es el valor mínimo de las n variables aleatorias, $X_{(n)}$ es el valor máximo y $X_{(1)}$ es el valor mínimo con las variables aleatorias ordenadas ascendentemente.

Definición 144 (rango de la muestra para datos agrupados). El rango de la muestra para datos agrupados es:

$$\operatorname*{ran}_n(X) = \operatorname*{máx}_k(C_k) - \operatorname*{mín}_k(C_1), \tag{7.35}$$

donde $\operatorname*{ran}_n(X)$ es el rango muestral para datos agrupados, $\operatorname*{máx}_k(C_k)$ es el valor máximo del k–ésimo intervalo de clase y $\operatorname*{mín}_k(C_1)$ es el valor mínimo del primer intervalo de clase.

7.4.2. Varianza

La varianza es el segundo momento con respecto a la media, es el promedio de la distancia cuadrática de cada una de las observaciones de la variable aleatoria a la media.

Definición 145 (varianza de la población). La varianza de la población es:

$$\sigma_X^2 = \frac{1}{N}\sum_{k=1}^{N}(X_k-\mu)^2 = \frac{1}{N}\sum_{k=1}^{N}X_k^2-\mu^2, \tag{7.36}$$

$$\sigma_X^2 = \frac{1}{N}\sum_{k=1}^{N}(X_k-\mu)^2 f_k = \frac{1}{N}\sum_{k=1}^{N}X_k^2 f_k-\mu^2, \tag{7.37}$$

$$\sigma_X^2 = \sum_{k=1}^{N}(X_k-\mu)^2 fr_k = \sum_{k=1}^{N}X_k^2 fr_k-\mu^2, \tag{7.38}$$

donde σ_X^2 es la varianza poblacional, N es el tamaño de la población, X_k es la k–ésima variable aleatoria, μ es la media poblacional, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 146 (varianza de la población para datos agrupados). La varianza poblacional para datos agrupados:

$$\sigma_X^2 = \frac{1}{N}\sum_{j=1}^{k}\left(\bar{C}_j-\mu\right)^2 f_j = \frac{1}{N}\sum_{j=1}^{k}\bar{C}_j^2 f_j-\mu^2 = \sum_{j=1}^{k}\left(\bar{C}_j-\mu\right)^2 fr_j = \sum_{j=1}^{k}\bar{C}_j^2 fr_j-\mu^2, \tag{7.39}$$

donde σ_X^2 es la varianza poblacional para datos agrupados, N es el tamaño de la población, $\bar{C}_j$ es la marca de clase del j–ésimo intervalo de clase, μ es la media poblacional, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Nota 15 (varianza de la población). La varianza de la población es un estimador sesgado, es decir:

$$\mathrm{E}\left(\hat{\sigma}_X^2\right) = \frac{(n-1)\,\sigma_X^2}{n}$$

Definición 147 (varianza de la muestra). La varianza de la muestra es:

$$S_X^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2-n\bar{X}^2\right), \tag{7.40}$$

$$S_X^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 f_k = \frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2 f_k-n\bar{X}^2\right), \tag{7.41}$$

$$S_X^2 = \frac{n}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 fr_k = \frac{n}{n-1}\left(\sum_{k=1}^{n}X_k^2 fr_k-\bar{X}^2\right), \tag{7.42}$$

donde S_X^2 es la varianza muestral, n es el tamaño de la muestra, X_k es la k–ésima variable aleatoria, $\bar{X}$ es la media muestral, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 148 (varianza de la muestra para datos agrupados). La varianza de la muestra para datos agrupados es:

$$S_X^2 = \frac{1}{n-1}\sum_{j=1}^{k}\left(\bar{C}_j - \bar{X}\right)^2 f_j = \frac{1}{n-1}\left(\sum_{j=1}^{k}\bar{C}_j^2 f_j - n\bar{X}^2\right), \tag{7.43}$$

$$S_X^2 = \frac{n}{n-1}\sum_{j=1}^{k}\left(\bar{C}_j - \bar{X}\right)^2 fr_j = \frac{n}{n-1}\left(\sum_{j=1}^{k}\bar{C}_j^2 fr_j - \bar{X}^2\right), \tag{7.44}$$

donde S_X^2 es la varianza muestral para datos agrupados, n es el tamaño de la muestra, $\bar{C}_j$ es la marca de clase del j–ésimo intervalo de clase, $\bar{X}$ es la media muestral, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Nota 16 (varianza de la muestra). La varianza de la muestra es un estimador insesgado de la varianza de la población, es decir, $\mathrm{E}\left(\hat{S}_X^2\right) = \sigma_X^2$.

7.4.3. Desviación estándar

Las unidades de la varianza son cuadráticas, entonces es imposible comparar las unidades de la varianza con las unidades de la variable aleatoria, por lo tanto la desviación estándar permite comparar las unidades de la variable aleatoria.

Definición 149 (desviación estándar de la población). La desviación estándar de la población es:

$$\sigma_X = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(X_k - \mu\right)^2} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}X_k^2 - \mu^2}, \tag{7.45}$$

$$\sigma_X = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left(X_k - \mu\right)^2 f_k} = \sqrt{\frac{1}{N}\sum_{k=1}^{N}X_k^2 f_k - \mu^2}, \tag{7.46}$$

$$\sigma_X = \sqrt{\sum_{k=1}^{N}\left(X_k - \mu\right)^2 fr_k} = \sqrt{\sum_{k=1}^{N}X_k^2 fr_k - \mu^2}, \tag{7.47}$$

donde σ_X es la desviación estándar poblacional, N es el tamaño de la población, X_k es la k–ésima variable aleatoria, μ es la media poblacional, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 150 (desviación estándar de la población para datos agrupados). La desviación estándar de la población para datos agrupados es:

$$\sigma_X = \sqrt{\frac{1}{N}\sum_{j=1}^{k}\left(\bar{C}_j - \mu\right)^2 f_j} = \sqrt{\frac{1}{N}\sum_{j=1}^{k}\bar{C}_j^2 f_j - \mu^2}, \tag{7.48}$$

$$\sigma_X = \sqrt{\sum_{j=1}^{k}\left(\bar{C}_j - \mu\right)^2 fr_j} = \sqrt{\sum_{j=1}^{k}\bar{C}_j^2 fr_j - \mu^2}, \tag{7.49}$$

donde σ_X es la desviación estándar poblacional para datos agrupados, N es el tamaño de la población, $\bar{C}_j$ es la marca de clase del j–ésimo intervalo de clase, μ es la media poblacional, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Definición 151 (desviación estándar de la muestra). La desviación estándar de la muestra es:

$$S_X = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2} = \sqrt{\frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2-n\bar{X}^2\right)}, \tag{7.50}$$

$$S_X = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2 f_k} = \sqrt{\frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2 f_k-n\bar{X}^2\right)}, \tag{7.51}$$

$$S_X = \sqrt{\frac{n}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2 fr_k} = \sqrt{\frac{n}{n-1}\left(\sum_{k=1}^{n}X_k^2 fr_k-\bar{X}^2\right)}, \tag{7.52}$$

donde S_X^2 es la desviación estándar de la muestra, n es el tamaño de la muestra, X_k es la k–ésima variable aleatoria, $\bar{X}$ es la media muestral, f_k es la frecuencia absoluta de la k–ésima variable aleatoria y fr_k es la frecuencia relativa de la k–ésima variable aleatoria.

Definición 152 (desviación estándar de la muestra para datos agrupados). La desviación estándar de la muestra para datos agrupados es:

$$S_X = \sqrt{\frac{1}{n-1}\sum_{j=1}^{k}\left(\bar{C}_j-\bar{X}\right)^2 f_j} = \sqrt{\frac{1}{n-1}\left(\sum_{j=1}^{k}\bar{C}_j^2 f_j-n\bar{X}^2\right)}, \tag{7.53}$$

$$S_X = \sqrt{\frac{n}{n-1}\sum_{j=1}^{k}\left(\bar{C}_j-\bar{X}\right)^2 fr_j} = \sqrt{\frac{n}{n-1}\left(\sum_{j=1}^{k}\bar{C}_j^2 fr_j-\bar{X}^2\right)}, \tag{7.54}$$

donde S_X^2 es la desviación estándar muestral para datos agrupados, n es el tamaño de la muestra, $\bar{C}_j$ es la marca de clase del j–ésimo intervalo de clase, $\bar{X}$ es la media muestral, f_j es la frecuencia absoluta del j–ésimo intervalo de clase y fr_j es la frecuencia relativa del j–ésimo intervalo de clase.

Ejemplo 84 (medidas de dispersión). Con los rendimientos logarítmicos de las paridades del tipo de cambio peso mexicano y dólar estadounidense del periodo de 01/09/2018 a 08/11/2018, presentados en la tabla 6.6 del ejemplo 81, calcular:

1. La varianza muestral S_X^2 y la varianza muestral para datos agrupados S_X^2.

2. La desviación estándar muestral S_X^2 y la desviación estándar muestral para datos agrupados S_X^2.

Solución. Las medidas de dispersión para la muestra de los 68 rendimientos son:

1. La varianza muestral es calculada substituyendo en las ecuaciones (7.40)–(7.42):

$$S_X^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 = \frac{0.002250}{67} = 0.000034.$$

$$S_X^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2 - n\bar{X}^2\right) = \frac{1}{67}\left(0.002271 - 68\,(0.000555)^2\right) = 0.000034.$$

$$S_X^2 = \frac{1}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 f_k = \frac{0.002250}{67} = 0.000034.$$

$$S_X^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2 f_k - n\bar{X}^2\right) = \frac{1}{67}\left(0.002271 - 68\,(0.000555)\right) = 0.000034.$$

$$S_X^2 = \frac{n}{n-1}\sum_{k=1}^{n}(X_k-\bar{X})^2 fr_k = \frac{68}{67}0.000033 = 0.000034.$$

$$S_X^2 = \frac{n}{n-1}\left(\sum_{k=1}^{n}X_k^2 fr_k - \bar{X}^2\right) = \frac{68}{67}\left(0.000033 - 0.000555^2\right) = 0.000034.$$

La varianza muestral para datos agrupados es calculada substituyendo en las ecuaciones (7.43) y (7.44):

$$S_X^2 = \frac{1}{n-1}\sum_{j=1}^{k}\left(\bar{C}_j-\bar{X}\right)^2 f_j = \frac{0.002249}{67} = 0.000034.$$

$$S_X^2 = \frac{1}{n-1}\left(\sum_{j=1}^{k}\bar{C}_j^2 f_j - n\bar{X}^2\right) = \frac{0.002272 - 68\,(0.000582)^2}{67} = 0.000034.$$

$$S_X^2 = \frac{n}{n-1}\sum_{j=1}^{k}\left(\bar{C}_j-\bar{X}\right)^2 fr_j = \frac{68}{67}0.000033 = 0.000034.$$

$$S_X^2 = \frac{n}{n-1}\left(\sum_{j=1}^{k}\bar{C}_j^2 fr_j - \bar{X}^2\right) = \frac{68}{67}\left(0.000033 - 0.000582^2\right) = 0.000034.$$

Por lo tanto, la varianza de la muestra es mayor que la varianza de la muestra agrupada en intervalos de clase por 0.0000000251.

2. La desviación estándar de la muestra es calculada con las ecuaciones (7.50)–(7.52):

$$S_X = 0.005795.$$

La desviación estándar de la muestra para datos agrupados es calculada con las ecuaciones (7.53) y (7.54):

$$S_X = 0.005793.$$

Por lo tanto, la desviación estándar de la muestra es mayor que la desviación estándar de la muestra agrupada en intervalos de clase por 0.0000021676 pesos mexicanos. ☐

7.5. Dispersión relativa

El coeficiente de variación de Pearson (1900) es para comparar la dispersión de variables aleatorias.

7.5.1. Coeficiente de variación de Pearson

Definición 153 (coeficiente de variación de la población). El coeficiente de variación de la población es:

$$v_X = \frac{\sigma_X}{\mu_X}, \tag{7.55}$$

donde v_X es el coeficiente de variación de la población, σ_X es la desviación estándar de la población y μ_X es la media aritmética de la población.

Definición 154 (coeficiente de variación de la muestra). El coeficiente de variación de la muestra es:

$$V_X = \frac{S_X}{\bar{X}}, \tag{7.56}$$

donde V_X es el coeficiente de variación de la muestra, S_X es la desviación estándar de la muestra y $\bar{X}$ es la media aritmética de la muestra.

Ejemplo 85 (dispersión relativa). Con los rendimientos logarítmicos de las paridades del tipo de cambio peso mexicano y dólar estadounidense del periodo de 01/09/2018 a 08/11/2018, presentados en la tabla 6.6 del ejemplo 81, calcular:

1. El coeficiente de variación de la muestra.
2. El coeficiente de variación de la muestra para datos agrupados.

Solución. El coeficiente de variación muestral se calcula sustituyendo en la ecuación (7.56):

1. El coeficiente de variación de la muestra:

$$V_X = \frac{S_X}{\bar{X}} = \frac{0.005795}{0.000555} = 10.4492.$$

2. El coeficiente de variación de la muestra para datos agrupados:

$$V_X = \frac{S_X}{\bar{X}} = \frac{0.005793}{0.000555} = 9.9470.$$

Entonces el coeficiente de variación de los rendimientos es 10.4492 y el coeficiente de variación de la muestra para datos agrupados es 9.9470. Por lo tanto, la muestra con menor dispersión es la de datos agrupados. □

7.6. Medidas de forma

Las medidas de forma de una variable aleatoria propuestas por Pearson (1900) y Fisher (1922) son:

1. Coeficiente de asimetría.
2. Coeficiente de curtosis.

7.6.1. Coeficiente de asimetría de Fisher

El coeficiente de asimetría propuesto por Fisher (1922) es una medida de forma de la distribución de una variable aleatoria que describe la forma. Existen tres tipos de asimetría:

1. Asimetría negativa: $\mu < \tilde{\mu} < \hat{\mu}$ o $\bar{X} < \tilde{X} < \hat{X}$.
2. Asimetría nula (simetría): $\mu = \tilde{\mu} = \hat{\mu}$ o $\bar{X} = \tilde{X} = \hat{X}$.
3. Asimetría positiva: $\mu > \tilde{\mu} > \hat{\mu}$ o $\bar{X} > \tilde{X} > \hat{X}$.

El coeficiente de asimetría indica el sesgo de la simetría que presenta la distribución de la variable aleatoria.

Definición 155 (coeficiente de asimetría de la población). El coeficiente de asimetría de la población es:

$$\gamma_1 = \frac{\mu_3}{\sigma_X^3} = \frac{1}{N\sigma_X^3} \sum_{k=1}^{N} (X_k - \mu)^3 , \tag{7.57}$$

donde γ_1 es el coeficiente de asimetría de la población, μ_3 es el tercer momento de la población con respecto a la media de la población y σ_X^3 es la desviación estándar de la población al cubo.

Definición 156 (coeficiente de asimetría de la muestra). El coeficiente de asimetría de la muestra es:

$$g_1 = \frac{M_3}{S_X^3} = \frac{1}{nS_X^3} \sum_{k=1}^{n} (X_k - \bar{X})^3 , \tag{7.58}$$

donde g_1 es el coeficiente de asimetría de la muestra, M_3 es el tercer momento de la muestra con respecto a la media de la muestra y S_X^3 es la desviación estándar de la muestra al cubo.

Si el coeficiente de asimetría $\gamma_1 > 0$ o $g_1 > 0$, entonces la distribución presenta sesgo positivo, si $\gamma_1 = 0$ o $g_1 = 0$, entonces la distribución es simétrica y si $\gamma_1 < 0$ o $g_1 < 0$, entonces la distribución presenta sesgo negativo.

Definición 157 (coeficiente de asimetría ajustado). El coeficiente de asimetría ajustado de la muestra es:

$$g_1^* = \frac{n^2 M_3}{(n-1)(n-2) S_X^3} = \frac{n}{(n-1)(n-2) S_X^3} \sum_{k=1}^{n} (X_k - \bar{X})^3 . \tag{7.59}$$

donde g_1^* es el coeficiente de asimetría ajustado, n es el tamaño de la muestra, M_3 es el tercer momento de la muestra con respecto a la media de la muestra y S_X^3 es la desviación estándar de la muestra al cubo.

7.6.2. Coeficiente de curtosis de Fisher

Pearson (1905) utilizó el término curtosis (platicúrtica, mesocúrtica y leptocúrtica) para indicar la forma de la distribución de una variable aleatoria. El coeficiente de curtosis es una medida de forma de la distribución de una variable aleatoria que describe la forma. Existen tres tipos de curtosis:

1. Leptocurtosis: presenta en los extremos una frecuencia mayor que $\mathrm{N}\left(\mu, \sigma^2\right)$.
2. Mesocurtosis: presenta en los extremos una frecuencia igual que $\mathrm{N}\left(\mu, \sigma^2\right)$.
3. Platocurtosis: presenta en los extremos una frecuencia menor que $\mathrm{N}\left(\mu, \sigma^2\right)$.

El coeficiente de curtosis indica la forma de los extremos de la distribución de una variable aleatoria con respecto a la distribución Gauss.

Definición 158 (coeficiente de curtosis de la población). El coeficiente de curtosis de la población es:

$$\gamma_2 = \frac{\mu_4}{\sigma_X^4} = \frac{1}{N\sigma_X^4}\sum_{k=1}^{N}(X_k-\mu)^4, \tag{7.60}$$

donde γ_2 es el coeficiente de curtosis de la población, μ_4 es el cuarto momento de la población con respecto a la media de la población y σ_X^4 es la varianza de la población al cuadrado.

Definición 159 (coeficiente de curtosis de la muestra). El coeficiente de curtosis de la muestra es:

$$g_2 = \frac{M_4}{S_X^4} = \frac{1}{nS_X^4}\sum_{k=1}^{n}(X_k-\bar{X})^4, \tag{7.61}$$

donde g_2 es el coeficiente de curtosis de la muestra, M_4 es el cuarto momento de la muestra con respecto a la media de la muestra y S_X^4 es la varianza de la muestra al cuadrado.

Si el coeficiente de curtosis $\gamma_2 > 3$ o $g_2 > 3$, entonces la distribución de la variable aleatoria es leptocúrtica, si $\gamma_2 = 3$ o $g_2 = 3$, entonces la distribución de la variable aleatoria es mesocúrtica y si $\gamma_2 < 3$ o $g_2 < 3$, entonces la distribución de la variable aleatoria es platicúrtica. Si el coeficiente de curtosis es $\gamma_2^\star = \gamma_2 - 3$ o $g_2^\star = g_2 - 3$, entonces si el coeficiente de curtosis $\gamma_2^\star > 0$ o $g_2^\star > 0$ la distribución de la variable aleatoria es leptocúrtica, si $\gamma_2^\star = 0$ o $g_2^\star = 0$ la distribución de la variable aleatoria es mesocúrtica y si $\gamma_2^\star < 0$ o $g_2^\star < 0$ la distribución de la variable aleatoria es platicúrtica.

Ejemplo 86 (medidas de forma). Con los rendimientos logarítmicos de las paridades del tipo de cambio peso mexicano y dólar estadounidense del periodo de 01/09/2018 a 08/11/2018, presentados en la tabla 6.6 del ejemplo 81, calcular:

1. El coeficiente de asimetría de la muestra.
2. El coeficiente de asimetría de la muestra para datos agrupados por el método de Velleman (1976).
3. El coeficiente de curtosis de la muestra.
4. El coeficiente de curtosis de la muestra para datos agrupados por el método de Velleman (1976).

Solución. El coeficiente de asimetría es calculado sustituyendo en la ecuación (7.58) y el coeficiente de curtosis es calculado sustituyendo en la ecuación (7.61):

1. El coeficiente de asimetría de la muestra:
$$g_1 = \frac{M_3}{S_X^3} = \frac{0.00000016}{0.005795^3} = 0.810883.$$
2. El coeficiente de asimetría de la muestra para datos agrupados:
$$g_1 = \frac{M_3}{S_X^3} = \frac{0.00000014}{0.005793^3} = 0.741264.$$
3. El coeficiente de curtosis de la muestra:
$$g_2 = \frac{M_4}{S_X^4} = \frac{0.0000000048}{0.005795^3} = 4.273136.$$

4. El coeficiente de curtosis de la muestra para datos agrupados:

$$g_2 = \frac{M_4}{S_X^4} = \frac{0.0000000047}{0.005795^3} = 4.206924.$$

Entonces los rendimientos presentan asimetría positiva y leptocurtosis, es decir, $g_1 > 0$, $g_2 > 3$ y $g_2^* > 0$. Por lo tanto, la distribución de los rendimientos es leptocúrtica y asimétrica positiva. □

Byers (2000) demostró que para una muestra de tamaño n, el valor máximo de g_2 es $n-2+\frac{1}{n-1}$, entonces el valor del estimador del coeficiente de curtosis ajustado propuesto por Byers (2000) es:

Definición 160 (coeficiente de curtosis ajustado). El coeficiente de curtosis ajustado de la muestra es:

$$\begin{aligned} g_2^* &= \frac{n^2\,(n+1)\,M_4}{(n-1)\,(n-2)\,(n-3)\,S_X^4} - \frac{3\,(n-1)^2}{(n-2)\,(n-3)}, \\ g_2^* &= \frac{n\,(n+1)}{(n-1)\,(n-2)\,(n-3)\,S_X^4}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^4 - \frac{3\,(n-1)^2}{(n-2)\,(n-3)}, \end{aligned} \tag{7.62}$$

donde g_2^* es el coeficiente de curtosis ajustado, n es el tamaño de la muestra, M_4 es el cuarto momento de la muestra con respecto a la media de la muestra y S_X^4 es la varianza de la muestra al cuadrado.

7.7. Distribuciones de las muestras

Las distribuciones de las muestras son fundamentales porque si son conocidas, entonces las probabilidades de ocurrencia de los eventos para obtener intervalos de confianza para un conjunto de parámetros también son conocidas y es factible realizar contrastes de hipótesis o calcular la potencia de los contrastes de hipótesis.

7.7.1. Media de la muestra

Definición 161 (media de la muestra). Sea $X_1,\ldots,X_n$ una muestra aleatoria, entonces la función $\bar{X}$ es el estadístico siguiente:

$$\bar{X} = \frac{1}{n}\sum_{k=1}^{n} X_k, \tag{7.63}$$

donde $\bar{X}$ es la media de la muestra, n es el tamaño de la muestra y X_k es la k–ésima variable aleatoria.

7.7.2. Proporción de la muestra

Definición 162 (proporción de la muestra). Sea $X_1,\ldots,X_n$ una muestra aleatoria, entonces la función P es el estadístico siguiente:

$$P = \frac{X}{n}, \tag{7.64}$$

donde X es el número de éxitos en la muestra y n es el tamaño de la muestra.

7.7.3. Varianza de la muestra

Definición 163 (varianza de la muestra). Sea $X_1,\ldots,X_n$ una muestra aleatoria, entonces la función S^2 es el estadístico siguiente:

$$S_X^2 = \frac{1}{n-1}\sum_{k=1}^{n}\left(X_k-\bar{X}\right)^2 = \frac{1}{n-1}\left(\sum_{k=1}^{n}X_k^2 - n\bar{X}^2\right), \tag{7.65}$$

donde S_X^2 es la varianza de la muestra, n es el tamaño de la muestra, X_k es la k–ésima variable aleatoria y $\bar{X}$ es la media de la muestra.

7.7.4. Teorema del límite central

El teorema del límite central tiene aplicaciones en el cálculo de probabilidades para aproximar distribuciones de probabilidad. De Moivre (1733) realizó la primera formulación y posteriormente las generalizaciones de Laplace (1814), Bessel (1838), Chebychev (1887), Markov (1898), Lyapunov (1901), Lindeberg (1922), Bernstein (1926) y Feller (1943).

Teorema 7 (teorema del límite central). Sea $X_1,X_2,\ldots,$ una sucesión infinita de variables aleatorias independientes e idénticamente distribuidas con media $\mathrm{E}\left(X_n\right)=\mu_X$ y varianza finita $\mathrm{Var}\left(X_n\right)=\sigma_X^2<\infty$, entonces:

$$\lim_{n\to\infty} Z_n = \lim_{n\to\infty}\frac{1}{\sigma_X\sqrt{n}}\left(\sum_{k=1}^{n}X_k - n\mu_X\right)\xrightarrow{\mathscr{D}}\mathrm{N}\left(0,1\right), \tag{7.66}$$

donde Z_n es una variable aleatoria estandarizada, X_k es la k–ésima variable aleatoria, n es el tamaño de la muestra, μ_X es la media, σ_X^2 es la varianza, σ_X es la desviación estándar y $\mathrm{N}\left(0,1\right)$ es una distribución Gauss estándar.

El teorema 7 (teorema del límite central) fue demostrado por Lyapunov (1901), donde la hipótesis acerca de una distribución particular es inexistente, es decir, la sucesión puede presentar cualquier distribución independiente e idéntica con media μ_X, varianza σ_X^2 y la variable aleatoria Z_n converge en distribución (ley) a una distribución Gauss estándar.

7.7.5. Distribución $\chi^2(n)$

Pearson (1900) propuso la distribución siguiente:

Definición 164 (distribución χ^2). Sea $n\in\mathbb{Z}_{++}$, entonces la variable aleatoria X se distribuye χ^2 con n grados de libertad y la función de densidad es:

$$f_X\left(x,n\right)=\begin{cases}\dfrac{\left(\frac{1}{2}\right)^{\frac{n}{2}}x^{\frac{n}{2}-1}\exp\left(-\frac{x}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} & \text{si } x>0\\ 0 & \text{si } x\leq 0\end{cases}, \tag{7.67}$$

donde $f_X\left(x,n\right)$ es al frecuencia relativa, n son los grados de libertad. La esperanza y la varianza son:

$$\mathrm{E}\left(X\right)=n. \tag{7.68}$$

$$\mathrm{Var}\left(X\right)=2n. \tag{7.69}$$

La distribución $\chi^2(n)$ tiene las propiedades siguientes:

Proposición 44 (distribución χ^2). Si $X \sim \mathrm{N}(0,1)$, entonces $X^2 \sim \chi^2(1)$.

Proposición 45 (distribución χ^2). Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes que presentan una distribución $\chi^2(n_k)$, entonces la variable aleatoria $X = \sum_{k=1}^{n} X_k \sim \chi^2\left(\sum_{k=1}^{k} n_k\right)$.

Proposición 46 (distribución χ^2). Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes que presentan una distribución $\mathrm{N}(0,1)$, entonces la variable aleatoria $X = \sum_{k=1}^{n} X_k^2 \sim \chi^2(n)$.

Proposición 47 (distribución χ^2). Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes que presentan una distribución $\mathrm{N}(\mu, \sigma^2)$, entonces:

$$\frac{(n-1)S_X^2}{\sigma_X^2} \sim \chi^2(n-1), \tag{7.70}$$

donde n es el tamaño de la muestra, S_X^2 es la varianza de la muestra y σ_X^2 es la varianza de la población.

7.7.6. Distribución $t(n)$

Student (1908) propuso la distribución siguiente:

Definición 165 (distribución t). Sea $n \in \mathbb{Z}_{++}$, entonces la variable aleatoria X se distribuye t con n grados de libertad y la función de densidad es:

$$f_X(x,n) = \frac{\Gamma\left(\frac{n+1}{2}\right)\left(1+\frac{x^2}{n}\right)^{-\frac{n+1}{2}}}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \qquad -\infty < x < \infty, \tag{7.71}$$

donde $f_X(x,n)$ es la frecuencia relativa y n son los grados de libertad. La esperanza y la varianza son:

$$\mathrm{E}(X) = 0. \tag{7.72}$$

$$\mathrm{Var}(X) = \frac{n}{n-1} \qquad \forall n > 2. \tag{7.73}$$

La distribución $t(n)$ tiene las propiedades siguientes:

Proposición 48 (distribución t). Si $X \sim \mathrm{N}(0,1)$ y $Y \sim \chi^2(n)$ son variables aleatorias independientes, entonces $\frac{X}{\sqrt{\frac{Y}{n}}} \sim t(n)$.

Proposición 49 (distribución t). Si $X_1, X_2, \ldots, X_n$ son variables aleatorias independientes e idénticamente distribuidas con distribución $\mathrm{N}(\mu, \sigma^2)$, entonces:

$$\frac{\bar{X} - \mu_X}{\frac{S_X}{\sqrt{n}}} \sim t(n-1), \tag{7.74}$$

donde $\bar{X}$ es la media de la muestra, μ_X es las media de la población, S_X es la desviacón estándar de la muestra y n es el tamaño de la muestra.

7.7.7. Distribución $F(n,m)$

Fisher (1922) propuso la distribución siguiente:

Definición 166 (distribución F). Sean $n,m \in \mathbb{Z}_{++}$, entonces la variable aleatoria X se distribuye F con n y m grados de libertad y la función de densidad es:

$$f_X(x,n,m) = \begin{cases} \dfrac{\Gamma\left(\frac{n+m}{2}\right)\left(\frac{n}{m}\right)^{\frac{n}{2}} x^{\frac{n}{2}-1}\left(1+\frac{nx}{m}\right)^{-\frac{n+m}{2}}}{\Gamma\left(\frac{n}{2}\right)\Gamma\left(\frac{m}{2}\right)} & \text{si } x>0 \\ 0 & \text{si } x \leq 0 \end{cases}, \tag{7.75}$$

donde $f_X(x,n)$ es al frecuencia relativa y n y m son los grados de libertad. La esperanza y la varianza son:

$$\mathrm{E}(X) = \frac{m}{m-2} \qquad \forall m>2. \tag{7.76}$$

$$\mathrm{Var}(X) = \frac{2m^2(n+m-2)}{n(m-2)^2(m-4)} \qquad \forall m>4. \tag{7.77}$$

La distribución $F(n,m)$ tiene las propiedades siguientes:

Proposición 50 (distribución F). Si $X \sim \chi^2(n)$ y $Y \sim \chi^2(m)$ son variables aleatorias independientes, entonces $\frac{\frac{X}{n}}{\frac{Y}{m}} \sim F(n,m)$.

Proposición 51 (distribución F). Si $X \sim t(m)$, entonces $X^2 \sim F(1,m)$.

7.8. Estimación por intervalos de confianza

Definición 167 (estimación por intervalos de confianza). Sea $\alpha \in (0,1)$, entonces un intervalo de confianza para un conjunto de parámetros θ de una distribución de probabilidad es un intervalo aleatorio de la forma $\left(\hat{\theta}_1, \hat{\theta}_2\right)$ donde $\hat{\theta}_1$ y $\hat{\theta}_2$ son estadísticos que satisfacen:

$$\mathscr{P}\left(\hat{\theta}_1 < \theta < \hat{\theta}_2\right) = 1-\alpha, \tag{7.78}$$

donde $\mathscr{P}\left(\hat{\theta}_1 < \theta < \hat{\theta}_2\right)$ es la probabilidad de que el intervalo aleatorio $\left(\hat{\theta}_1, \hat{\theta}_2\right)$ contenga al conjunto de parámetros θ, $\hat{\theta}_1$ es límite inferior del intervalo de confianza, $\hat{\theta}_2$ es el límite superior del intervalo de confianza, $1-\alpha$ es el nivel de confianza y α es el nivel de significación.

7.8.1. Intervalo de confianza para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 conocida

Sea $X_1,\ldots,X_n$ una muestra aleatoria de una población que se distribuye $\mathrm{N}\left(\mu,\sigma^2\right)$ con media μ desconocida y varianza σ^2 conocida, entonces $\bar{X} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Estandarizando:

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim \mathrm{N}(0,1).$$

Por lo tanto, $\forall\alpha \in (0,1)$:

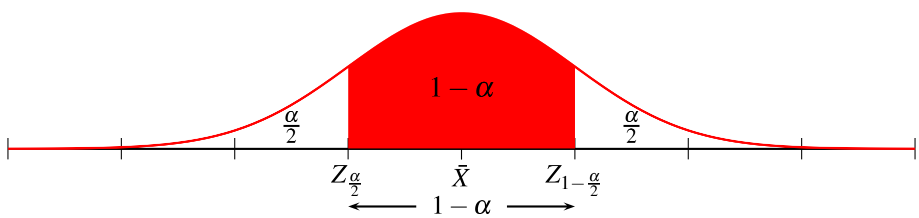
$$\mathscr{P}\left(-Z_{1-\frac{\alpha}{2}} < \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha.$$

Despejando el parámetro μ se tiene el resultado siguiente:

Proposición 52 (intervalo de confianza). Un intervalo de confianza para μ de una distribución Gauss con varianza σ^2 conocida es:

$$\mathscr{P}\left(\bar{X} - \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \mu < \bar{X} + \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1-\alpha. \tag{7.79}$$

La probabilidad de que el intervalo de confianza $\left(\bar{X} - \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X} + \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$ contenga a μ es $1-\alpha$ y está representada en la gráfica 7.1.



Gráfica 7.1: Intervalo de confianza para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 conocida.

La gráfica 7.1 indica que para el nivel de significación $0<\alpha<1$ son localizados los coeficientes de confianza o valores críticos $Z_{\frac{\alpha}{2}}$ y $Z_{1-\frac{\alpha}{2}}$ en la tabla A.1. Por ejemplo, dado el el nivel de significación $0<\alpha<1$, entonces es deducido el nivel de confianza $1-\alpha$.

Algunos valores críticos son presentados en la tabla 7.2.

Tabla 7.2: Valores críticos $Z_{1-\alpha}$.

$1-\alpha$	0.9950	0.9900	0.9750	0.9500	0.9000	0.1000	0.0500	0.0250	0.0100	0.0050
$Z_{1-\alpha}$	2.5758	2.3263	1.9600	1.6449	1.2816	-1.2816	-1.6449	-1.9600	-2.3263	-2.5758

Nota 17 (intervalo de confianza para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 desconocida y $n \geq 30$). El intervalo de confianza para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 desconocida y $n \geq 30$ es calculado con la ecuación (7.79) y sustituyendo σ^2 por S_X^2, pero formalmente es calculado con la ecuación (7.80).

Ejemplo 87 (intervalo de confianza para la media de la población Gauss con varianza conocida). Una muestra de 1,000 individuos de una población de 15,460 individuos presenta una media muestral $\bar{X} = 171.32$ y una desviación estándar poblacional $\sigma_X = 7.4422$ en centímetros con una distribución Gauss. Calcular el intervalo de confianza para la estatura de la población con un nivel de confianza de:

1. 95 %.
2. 99 %.

Solución. Por la proposición 52 (intervalo de confianza) y sustituyendo en la ecuación (7.79):

1. El intervalo de confianza al 95 % es:

$$\mathscr{P}\left(171.32-\frac{7.4422Z_{0.975}}{\sqrt{1000}}<\mu<171.32+\frac{7.4422Z_{0.975}}{\sqrt{1000}}\right)=1-0.05,$$

   donde $Z_{0.975}=1.96$ (tabla 7.2), entonces $170.86<\mu<171.78$, por lo tanto $\mu\in(170.86,171.78)$ con un nivel de confianza del 95 %.

2. El intervalo de confianza al 99 % es:

$$\mathscr{P}\left(171.32-\frac{7.4422Z_{0.995}}{\sqrt{1000}}<\mu<171.32+\frac{7.4422Z_{0.995}}{\sqrt{1000}}\right)=1-0.05.$$

   donde $Z_{0.975}=2.5758$ (tabla 7.2), entonces $170.71<\mu<171.93$, por lo tanto $\mu\in(170.71,171.93)$ con un nivel de confianza del 99 %. □

7.8.2. Intervalo de confianza para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 desconocida y $n<30$

Sea $X_1,\ldots,X_n$ una muestra aleatoria de una población que se distribuye $\mathrm{N}\left(\mu,\sigma^2\right)$ con media μ desconocida varianza σ^2 desconocida y $n<30$, entonces por la proposición 49 (distribución t):

$$T=\frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}\sim t\,(n-1)\,.$$

Por lo tanto, $\forall\alpha\in(0,1)$:

$$\mathscr{P}\left(-t_{1-\frac{\alpha}{2}}\,(n-1)<\frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}<t_{1-\frac{\alpha}{2}}\,(n-1)\right)=1-\alpha.$$

Despejando el parámetro μ se tiene el resultado siguiente:

Proposición 53 (intervalo de confianza). Un intervalo de confianza para la media μ de una distribución Gauss con varianza σ^2 desconocida y $n<30$ es:

$$\mathscr{P}\left(\bar{X}-\frac{S_X t_{1-\frac{\alpha}{2}}\,(n-1)}{\sqrt{n}}<\mu<\bar{X}+\frac{S_X t_{1-\frac{\alpha}{2}}\,(n-1)}{\sqrt{n}}\right)=1-\alpha. \tag{7.80}$$

La probabilidad de que el intervalo $\left(\bar{X}-\frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}},\bar{X}+\frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}}\right)$ contenga a μ es $1-\alpha$ y dado el el nivel de significación $0<\alpha<1$, entonces es deducido el nivel de confianza $1-\alpha$.

Algunos valores críticos de la distribución $t_{1-\alpha}\,(19)$ son presentados en la tabla 7.3.

Tabla 7.3: Valores críticos $t_{1-\alpha}\,(19)$.

$1-\alpha$	0.9950	0.9900	0.9750	0.9500	0.9000	0.1000	0.0500	0.0250	0.0100	0.0050
$t_{1-\alpha}\,(19)$	2.8609	2.5395	2.0930	1.7291	1.3277	-1.3277	-1.7291	-2.0930	-2.5395	-2.8609

Ejemplo 88 (intervalo de confianza para la media de la población Gauss con varianza desconocida y $n < 30$). Una muestra de 20 diámetros en centímetros tiene una media $\bar{X} = 10.9273$ y una desviación estándar (con 8 cifras significativas) $S_X = 0.12384731$. Calcular el intervalo de confianza para el diámetro de la población con un nivel de confianza de:

1. 95 %.
2. 99 %.

Solución. Por la proposición 53 (intervalo de confianza) y sustituyendo en la ecuación (7.80):

1. El intervalo de confianza al 95 % es:

$$\mathscr{P}\left(10.9273 - \frac{0.12384731 t_{0.975}(19)}{\sqrt{20}} < \mu < 10.9273 + \frac{0.12384731 t_{0.975}(19)}{\sqrt{20}}\right) = 1 - 0.05,$$

   donde $t_{0.975}(19) = 2.0930$ (tabla 7.3), entonces $10.87 < \mu < 10.99$, por lo tanto $\mu \in (10.87, 10.99)$ con un nivel de confianza del 95 %.

2. El intervalo de confianza al 99 % es:

$$\mathscr{P}\left(10.9273 - \frac{0.12384731 t_{0.995}(19)}{\sqrt{20}} < \mu < 10.9273 + \frac{0.12384731 t_{0.995}(19)}{\sqrt{20}}\right) = 1 - 0.05,$$

   donde $t_{0.995}(19) = 2.8609$ (tabla 7.3), entonces $10.85 < \mu < 11.01$, por lo tanto $\mu \in (10.85, 11.01)$ con un nivel de confianza del 95 %. □

7.8.3. Intervalo de confianza para μ de cualquier distribución y $n \geq 30$

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población que presenta una distribución cualquiera con media μ desconocida, varianza σ^2 desconocida y $n \geq 30$, entonces por el teorema 7 (teorema del límite central):

$$Z = \frac{\bar{X} - \mu}{\frac{S_X}{\sqrt{n}}} \sim \mathrm{N}(0,1).$$

Por lo tanto, $\forall \alpha \in (0,1)$:

$$\mathscr{P}\left(-Z_{1-\frac{\alpha}{2}} < \frac{\bar{X} - \mu}{\frac{S_X}{\sqrt{n}}} < Z_{1-\frac{\alpha}{2}}\right) = 1 - \alpha.$$

Despejando el parámetro μ se tiene el resultado siguiente:

Proposición 54 (intervalo de confianza). Un intervalo de confianza para la media μ de una distribución cualquiera con varianza σ^2 desconocida y $n \geq 30$ es:

$$\mathscr{P}\left(\bar{X} - \frac{S_X Z_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \mu < \bar{X} + \frac{S_X Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right) = 1 - \alpha. \tag{7.81}$$

La probabilidad de que el intervalo de confianza $\left(\bar{X} - \frac{S_X Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}, \bar{X} + \frac{S_X Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}\right)$ contenga a μ es $1 - \alpha$.

7.8.4. Intervalo de confianza para μ de cualquier distribución con $n < 30$

Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población que presenta una distribución cualquiera con media μ desconocida, varianza σ^2 desconocida y $n < 30$, entonces por la proposición 49 (distribución t):

$$T = \frac{\bar{X} - \mu}{\frac{S_X}{\sqrt{n}}} \sim t\,(n-1)\,.$$

Por lo tanto, $\forall \alpha \in (0,1)$:

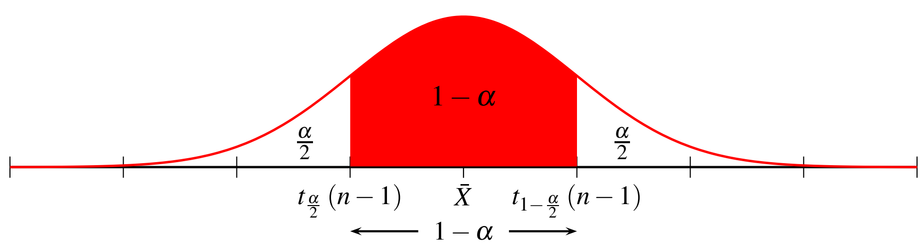
$$\mathscr{P}\left(-t_{1-\frac{\alpha}{2}}\,(n-1) < \frac{\bar{X} - \mu}{\frac{S_X}{\sqrt{n}}} < t_{1-\frac{\alpha}{2}}\,(n-1)\right) = 1 - \alpha.$$

Despejando el parámetro μ se tiene el resultado siguiente:

Proposición 55. (intervalo de confianza) Un intervalo de confianza para la media μ de una distribución cualquiera con varianza σ^2 desconocida y $n < 30$ es:

$$\mathscr{P}\left(\bar{X} - \frac{S_X t_{1-\frac{\alpha}{2}}\,(n-1)}{\sqrt{n}} < \mu < \bar{X} + \frac{S_X t_{1-\frac{\alpha}{2}}\,(n-1)}{\sqrt{n}}\right) = 1 - \alpha \tag{7.82}$$

La probabilidad de que el intervalo $\left(\bar{X} - \frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}}, \bar{X} + \frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}}\right)$ contenga a μ es $1 - \alpha$ y está representada en la gráfica 7.2.



Gráfica 7.2: Intervalo de confianza para μ de una distribución cualquiera con $n < 30$.

La gráfica 7.2 indica que para el nivel de significación $0 < \alpha < 1$ son localizados los valores críticos $t_{\frac{\alpha}{2}}\,(n-1)$ y $t_{1-\frac{\alpha}{2}}\,(n-1)$ en la tabla B.1.

7.8.5. Intervalo de confianza para p de una $Bin\,(n,p)$

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población que se distribuye $Bin\,(n,p)$ con proporción p desconocida y varianza $\sigma^2 = n\hat{P}\left(1 - \hat{P}\right)$ donde $n \geq 30$, entonces por el teorema 7 (teorema del límite central):

$$P = \frac{\hat{P} - p}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n}}} \sim \mathrm{N}\,(np, np\,(1-p))\,.$$

Por lo tanto, $\forall\alpha \in (0,1)$:

$$\mathscr{P}\left(-Z_{1-\frac{\alpha}{2}} < \frac{\hat{P}-p}{\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n}}} < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha.$$

Despejando el parámetro p se tiene el resultado siguiente:

Proposición 56 (intervalo de confianza). Un intervalo de confianza para la proporción p de una distribución binomial con una varianza $\sigma^2 = n\hat{P}(1-\hat{P})$ es:

$$\mathscr{P}\left(\hat{P}-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}\left(1-\hat{P}\right)}{n}} < p < \hat{P}+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right) = 1-\alpha. \tag{7.83}$$

La probabilidad de que el intervalo $\left(\hat{P}-Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}, \hat{P}+Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}\right)$ contenga a p es $1-\alpha$.

Ejemplo 89 (intervalo de confianza de la proporción). Una muestra de 1,000 electores indica que el 60 % están a favor del candidato del partido demócrata. Encontrar el intervalo de confianza para la proporción de los electores que están a favor del candidato demócrata con un nivel de confianza de:

1. 95 %.
2. 99 %.

Solución. Por la proposición 56 (intervalo de confianza) y sustituyendo en la ecuación (7.83):

1. El intervalo de confianza con un nivel de confianza de 95 %:

$$\mathscr{P}\left(0.6-Z_{0.975}\sqrt{\frac{0.6(0.4)}{1,000}} < p < 0.6+Z_{0.975}\sqrt{\frac{0.6(0.4)}{1,000}}\right) = 1-0.05,$$

donde $Z_{0.975} = 1.96$ (tabla 7.2). Por lo tanto, $0.5696 < p < 0.6304$.

2. El intervalo de confianza con un nivel de confianza de 99 %:

$$\mathscr{P}\left(0.6-Z_{0.995}\sqrt{\frac{0.6(0.4)}{1,000}} < p < 0.6+Z_{0.995}\sqrt{\frac{0.6(0.4)}{1,000}}\right) = 1-0.05,$$

donde $Z_{0.995} = 2.5758$ (tabla 7.2). Por lo tanto, $0.5601 < p < 0.6399$. □

7.8.6. Intervalo de confianza para σ^2 de una $\mathrm{N}\left(\mu,\sigma^2\right)$

Sea $X_1, X_2, \ldots, X_n$ una muestra aleatoria de una población que se distribuye $\mathrm{N}\left(\mu,\sigma^2\right)$ con varianza σ^2 desconocida, entonces por la proposición 47 (distribución χ^2):

$$X^2 = \frac{(n-1)\,S_X^2}{\sigma^2} \sim \chi^2\,(n-1)\,.$$

Por lo tanto, $\forall\alpha \in (0,1)$:

$$\mathscr{P}\left(\frac{1}{\chi^2_{1-\frac{\alpha}{2}}\,(n-1)} < \frac{\sigma^2}{(n-1)\,S_X^2} < \frac{1}{\chi^2_{\frac{\alpha}{2}}\,(n-1)}\right) = 1-\alpha.$$

Despejando el parámetro σ^2 se tiene el resultado siguiente:

Proposición 57 (intervalo de confianza). Un intervalo de confianza para la varianza σ^2 de una distribución Gauss basado en el tamaño de la muestra es:

$$\mathscr{P}\left(\frac{(n-1)S_X^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)} < \sigma^2 < \frac{(n-1)S_X^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right) = 1-\alpha. \tag{7.84}$$

La probabilidad de que el intervalo de confianza $\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}(n-1)}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}(n-1)}\right)$ contenga al parámetro σ^2 es $1-\alpha$ y dado el el nivel de significación $0<\alpha<1$, entonces es deducido el nivel de confianza $1-\alpha$.

Algunos valores críticos de la distribución $\chi^2_{1-\alpha}(19)$ son presentados en la tabla 7.4.

Tabla 7.4: Valores críticos $\chi^2_{1-\alpha}(19)$.

$1-\alpha$	0.9950	0.9900	0.9750	0.9500	0.9000	0.1000	0.0500	0.0250	0.0100	0.0050
$\chi_{1-\alpha}(9)$	23.5894	21.6660	19.0228	16.9190	14.6837	4.1682	3.3251	2.7004	2.0879	1.7349

Ejemplo 90 (intervalo de confianza para la varianza). Una muestra aleatoria de una población que tiene una distribución $\mathrm{N}(\mu,\sigma^2)$ tiene los valores del conjunto $X=\{6.8,5.9,8.1,9.1,8.4,7.7,7.9,8.8,6.9,7.8\}$. Estimar el intervalo de confianza para la varianza con un nivel de confianza de:

1. 95 %.
2. 99 %.

Solución. La estimación de la media muestral es $\bar{X}=\frac{1}{10}\sum_{k=1}^{10}X_k=7.74$ y la estimación de la varianza muestral es $S_X^2=\frac{1}{9}\sum_{k=1}^{10}(X_k-\bar{X})^2=0.949\bar{3}$. Entonces por la proposición 57 (intervalo de confianza) y sustituyendo en la ecuación (7.84):

1. El intervalo de confianza con un nivel de confianza de 95 %:

$$\mathscr{P}\left(\frac{9(0.949\bar{3})}{\chi^2_{0.975}(9)} < \sigma^2 < \frac{9(0.949\bar{3})}{\chi^2_{0.025}(9)}\right) = 1-0.05,$$

donde $\chi^2_{0.025}(9)=2.7004$ y $\chi^2_{0.975}(9)=19.0228$ (tabla 7.4), entonces $0.4491<\sigma^2<3.1640$, por lo tanto $\sigma^2\in(0.4491,3.1640)$.

2. El intervalo de confianza con un nivel de confianza de 99 %:

$$\mathscr{P}\left(\frac{9(0.949\bar{3})}{\chi^2_{0.975}(9)} < \sigma^2 < \frac{9(0.949\bar{3})}{\chi^2_{0.025}(9)}\right) = 1-0.05,$$

donde $\chi^2_{0.005}(9)=1.7349$ y $\chi^2_{0.975}(9)=23.5894$ (tabla 7.4), entonces $0.3622<\sigma^2<4.9247$, por lo tanto $\sigma^2\in(0.3622,4.9247)$. □

7.9. Tamaño de la muestra

La estimación del tamaño de la muestra en función del nivel del error ε para conocer tiempos y costos.

7.9.1. Tamaño de la muestra para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 conocida

Proposición 58 (tamaño de la muestra). Sea $\varepsilon = \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}}$, entonces el tamaño mínimo de la muestra es:

$$n \geq \left\lceil \left(\frac{\sigma Z_{1-\frac{\alpha}{2}}}{\varepsilon} \right)^2 \right\rceil, \tag{7.85}$$

donde n es el tamaño de la muestra, σ es la desviación estándar de la población, $Z_{1-\frac{\alpha}{2}}$ es el percentil de la distribución Gauss, ε es el nivel de error y $\lceil \cdot \rceil$ es la función entero superior.

Ejemplo 91 (tamaño de la muestra). Encontrar el tamaño de la muestra mínimo para una población $\mathrm{N}\left(\mu,\sigma^2\right)$, para que el error de estimación del intervalo de confianza para μ sea menor que 0.3 con un nivel de confianza de 95 % cuando $\sigma = 7.44$.

Solución. El tamaño de la muestra para la media es estimado de acuerdo con la proposición 58 (tamaño de la muestra) y sustituyendo en la ecuación (7.85):

$$n \geq \left\lceil \left(\frac{\sigma Z_{1-\frac{\alpha}{2}}}{\varepsilon} \right)^2 \right\rceil = \left\lceil \left(\frac{7.44\,(1.96)}{0.3} \right)^2 \right\rceil = \lceil 2,362.74 \rceil = 2,363.$$

Por lo tanto, el tamaño de la muestra es cuando menos de 2,363 elementos para que el error de la estimación de la media sea menor que 30 %. □

7.9.2. Tamaño de la muestra para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 desconocida y $n < 30$

Proposición 59 (tamaño de la muestra). Sea $\varepsilon = \frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\sqrt{n}}$, entonces el tamaño mínimo de la muestra es:

$$n \geq \left\lceil \left(\frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\varepsilon} \right)^2 \right\rceil, \tag{7.86}$$

donde n es el tamaño de la muestra, S_X es la desviación estándar de la muestra, $t_{1-\frac{\alpha}{2}}(n-1)$ es el percentil de la distribución t, ε es el nivel de error y $\lceil \cdot \rceil$ es la función entero superior.

Ejemplo 92 (tamaño de la muestra). Encontrar el tamaño de la muestra mínimo de una población $\mathrm{N}\left(\mu,\sigma^2\right)$, para que el error de estimación del intervalo de confianza para μ sea menor que 0.2 con un nivel de confianza de 95 % si una muestra de tamaño $n = 21$ presentó una desviación estándar $S_X = 8$ y una media $\bar{X} = 66$.

Solución. El tamaño de la muestra para la media es estimado de acuerdo con la proposición 59 (tamaño de la muestra) y sustituyendo en la ecuación (7.86):

$$n \geq \left\lceil \left(\frac{S_X t_{1-\frac{\alpha}{2}}(n-1)}{\varepsilon} \right)^2 \right\rceil = \left\lceil \left(\frac{8\,(2.0860)}{0.2} \right)^2 \right\rceil = \lceil 6,961.99 \rceil = 6,962.$$

Por lo tanto, el tamaño de la muestra es cuando menos de 6,962 elementos para que el error en la estimación de la media sea menor que 20 %. □

7.9.3. Tamaño de la muestra para p una $Bin(n,p)$

Proposición 60 (tamaño de la muestra). Sea $\varepsilon = Z_{1-\frac{\alpha}{2}}\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}$, entonces el tamaño mínimo de la muestra es:

$$n \geq \left\lceil \hat{P}\left(1-\hat{P}\right)\left(\frac{Z_{1-\frac{\alpha}{2}}}{\varepsilon}\right)^2 \right\rceil, \tag{7.87}$$

donde n es el tamaño de la muestra, $\hat{P}$ es el estimador de la proporción, $Z_{1-\frac{\alpha}{2}}$ es el percentil de la distribución Gauss, ε es el nivel de error y $\lceil\cdot\rceil$ es la función entero superior.

Ejemplo 93 (tamaño de la muestra). La cámara legislativa promulga una ley que prohíbe fumar en edificios públicos. Determinar el tamaño de la muestra para que el error del intervalo de confianza de la proporción sea menor que 0.01 con un nivel de confianza de 99 %.

Solución. El tamaño de la muestra para la proporción es estimado de acuerdo con la proposición 60 (tamaño de la muestra) y sustituyendo en la ecuación (7.87):

$$n \geq \left\lceil \hat{P}\left(1-\hat{P}\right)\left(\frac{Z_{1-\frac{\alpha}{2}}}{\varepsilon}\right)^2 \right\rceil = \left\lceil \left(\frac{0.5\,(2.5758)}{0.01}\right)^2 \right\rceil = \lceil 16,587.24 \rceil = 16,588.$$

Algunas proporciones son presentadas en la tabla 7.5.

Tabla 7.5: Proporciones para el tamaño de la muestra.

$\hat{P}$	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
$\hat{P}\left(1-\hat{P}\right)$	0.00	0.09	0.16	0.21	0.24	0.25	0.24	0.21	0.16	0.09	0.00

Por lo tanto, el tamaño de la muestra[1] es cuando menos de 16,588 ciudadanos para que el error en la estimación de la proporción sea menor que 1 %. □

[1]La proporción $\hat{P} = \frac{1}{2}$ es utilizada porque maximiza el tamaño de la muestra.

7.10. Ejercicios

Ejercicio 81. Un ingeniero mide los diámetros de diez cilindros y obtiene los valores (en centímetros) presentados en la tabla siguiente:

4.88	5.09	4.92	4.97	5.02	4.95	5.03	4.92	4.98	5.06

Calcular los estimadores siguientes:

1. Las medidas de tendencia central, es decir, $\bar{X}, \tilde{X}, \hat{X}$.
2. Las medidas de dispersión, es decir, $\operatorname{ran}(X), S_X^2, S_X$.
3. El coeficiente de variación, es decir, V_X.
4. Las medidas de forma, es decir, g_1^* y g_2^*.
5. El intervalo de confianza para μ al 95 %.
6. El intervalo de confianza para σ^2 al 95 %.
7. El tamaño de la muestra para μ al 95 % con $\varepsilon \leq 0.1$.

Ejercicio 82. Las 38 calificaciones de un estudiante y las frecuencias absolutas son presentadas en la tabla siguiente:

Calificación	6	7	8	9	10
Frecuencia	0	2	8	16	12

Calcular los estimadores siguientes:

1. Las medidas de tendencia central, es decir, $\bar{X}, \tilde{X}, \hat{X}$.
2. Las medidas de dispersión, es decir, $\operatorname{ran}(X), S_X^2, S_X$.
3. El coeficiente de variación, es decir, V_X.
4. Las medidas de forma, es decir, g_1^* y g_2^*.
5. El intervalo de confianza para μ al 99 %.
6. El intervalo de confianza para σ^2 al 99 %.
7. El tamaño de la muestra para μ al 99 % con $\varepsilon \leq 0.05$.

Ejercicio 83. Los flujos de las regaderas (litros sobre minuto) de una muestra de 120 casas son presentados en la tabla siguiente:

4.6	12.3	7.1	7.0	4.0	9.2	6.7	6.9	11.5	5.1	11.2	10.5
14.3	8.0	8.8	6.4	5.1	5.6	9.6	7.5	7.5	6.2	5.8	2.3
3.4	10.4	9.8	6.6	3.7	6.4	8.3	6.5	7.6	9.3	9.2	7.3
5.0	6.3	13.8	6.2	5.4	4.8	7.5	6.0	6.9	10.8	7.5	6.6
5.0	3.3	7.6	3.9	11.9	2.2	15.0	7.2	6.1	15.3	18.9	7.2
5.4	5.5	4.3	9.0	12.7	11.3	7.4	5.0	3.5	8.2	8.4	7.3
10.3	11.9	6.0	5.6	9.5	9.3	10.4	9.7	5.1	6.7	10.2	6.2
8.4	7.0	4.8	5.6	10.5	14.6	10.8	15.5	7.5	6.4	3.4	5.5
6.6	5.9	15.0	9.6	7.8	7.0	6.9	4.1	3.6	11.9	3.7	5.7
6.8	11.3	9.3	9.6	10.4	9.3	6.9	9.8	9.1	10.6	4.5	6.2

Calcular los estimadores siguientes:

1. Las medidas de tendencia central, es decir, $\bar{X}, \tilde{X}, \hat{X}$.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
2. Las medidas de dispersión, es decir, $\operatorname{ran}(X), S_X^2, S_X$.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
3. El coeficiente de variación, es decir, V_X.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
4. Las medidas de forma, es decir, g_1^* y g_2^*.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
5. El intervalo de confianza para μ al 99 %.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
6. El intervalo de confianza para σ^2 al 99 %.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
7. El tamaño de la muestra para μ al 99 % con $\varepsilon \leq 0.05$.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).

Ejercicio 84. Las alturas (centímetros) de una muestra de 100 árboles son presentadas en la tabla siguiente:

143	155	133	158	161	87	120	89	112	150
178	155	166	130	166	183	140	181	145	191
135	122	140	171	153	153	186	188	110	188
181	135	199	150	143	158	122	166	173	130
186	158	204	135	163	112	171	115	148	122
127	145	183	140	143	158	183	145	125	158
117	155	133	117	183	143	117	122	145	133
138	186	181	178	168	171	148	181	191	127
112	150	143	138	161	110	173	176	140	161
122	125	178	153	171	120	125	176	168	186

Calcular los estimadores siguientes:

1. Las medidas de tendencia central, es decir, $\bar{X}, \tilde{X}, \hat{X}$:
   a) Datos sin agrupar.
   b) Datos agrupados con el método Velleman (1976).
2. Las medidas de dispersión, es decir, $\operatorname{ran}(X), S_X^2, S_X$.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
3. El coeficiente de variación, es decir, V_X.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
4. Las medidas de forma, es decir, g_1^* y g_2^*.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
5. El intervalo de confianza para μ al 99 %.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
6. El intervalo de confianza para σ^2 al 99 %.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).
7. El tamaño de la muestra para μ al 99 % con $\varepsilon \leq 0.1$.
   a) Datos sin agrupar.
   b) Datos agrupados con el método de Velleman (1976).

Ejercicio 85. Una muestra de 20 estudiantes de un curso de Estadística tiene información acerca de las marcas de calculadoras que utilizaron (T=Texas Instruments, H=Hewlett Packard, C=Casio, S=Sharp). Las marcas utilizadas son presentadas en la tabla siguiente:

T	T	H	T	C	T	T	S	C	H
S	S	T	H	C	T	T	T	H	T

Calcular los estimadores siguientes:

1. La proporción para los estudiantes que utilizaron T, es decir p_T.
2. La proporción para los estudiantes que utilizaron H, es decir p_H.
3. La proporción para los estudiantes que utilizaron C, es decir p_C.
4. La proporción para los estudiantes que utilizaron S, es decir p_S.

Capítulo 8

Pruebas de hipótesis

Si el sistema supone que un acusado es inocente hasta probar la culpabilidad, entonces la conjetura de que el acusado es culpable es probada con evidencia. La sentencia con fundamento en la evidencia dicta inocente o culpable, pero existe la probabilidad de culpar a un inocente o absolver a un culpable, es decir, existe la probabilidad de falsos positivos o falsos negativos.

Definición 168 (hipótesis estadística). Una hipótesis estadística es una conjetura acerca de la distribución de un conjunto no vacío de variables aleatorias.

La hipótesis es la respuesta tentativa para la solución de preguntas de investigación.

Ejemplo 94 (hipótesis estadística). Si la variable aleatoria $X \sim Bin(n, p)$, entonces las afirmaciones $p = \frac{1}{2}$ y $p \leq \frac{1}{2}$ son hipótesis estadísticas. $\square$

Las hipótesis estadísticas son clasificadas en:

1. Hipótesis simples.

2. Hipótesis compuestas.

Definición 169 (hipótesis simple). Una hipótesis simple especifica completamente la función de distribución de probabilidad de un conjunto no vacío de variables aleatorias.

Ejemplo 95 (hipótesis simple). Si la variable aleatoria $X \sim \mathrm{N}(\mu, 1)$, entonces la afirmación $\mu = 0$ es una hipótesis simple porque especifica completamente a la función de distribución. $\square$

Definición 170 (hipótesis compuesta). Una hipótesis compuesta no especifica completamente la función de distribución de probabilidad de un conjunto no vacío de variables aleatorias.

Ejemplo 96 (hipótesis compuesta). Si la variable aleatoria $X \sim Poisson(\lambda)$, entonces la afirmación $\lambda > 10$ es una hipótesis compuesta porque no especifica completamente la función de distribución. $\square$

La hipótesis nula H_0 y la hipótesis alternativa H_1 son simples o compuestas, entonces son considerados cuatro contrastes de hipótesis:

1. Hipótesis simple H_0 e hipótesis simple H_1.
2. Hipótesis simple H_0 e hipótesis compuesta H_1.
3. Hipótesis compuesta H_0 e hipótesis simple H_1.
4. Hipótesis compuesta H_0 e hipótesis compuesta H_1.

La información para establecer una regla de decisión para no rechazar o rechazar una hipótesis proviene de una muestra aleatoria $X_1,\ldots,X_n$ de una distribución. Si la hipótesis no es rechazada, entonces no se afirma que esta es absolutamente cierta, simplemente es consistente con la muestra aleatoria, por lo tanto si la muestra cambia, la decisión de rechazar o no rechazar la hipótesis también cambia.

El modelo de decisión para inferir estadísticas tiene los elementos siguientes:

1. La hipótesis nula H_0.
2. La hipótesis alternativa H_1.
3. El nivel de significación α utilizado en la prueba estadística.
4. El tamaño de la muestra n.
5. La regla de decisión.

El modelo de decisión para inferir estadísticas tiene las etapas siguientes:

1. Plantear la hipótesis nula H_0.
2. Plantear la hipótesis alternativa H_1.
3. Especificar el nivel de significación α utilizado en la prueba estadística.
4. Determinar el tamaño n de la muestra.
5. Determinar la estadística de prueba.
6. Establecer los puntos críticos que determinan la región crítica $\mathscr{C}$.
7. Establecer la regla de decisión.
8. Determinar si la prueba estadística está en la región de rechazo o de no rechazo.
9. Determinar la decisión estadística.
10. Expresar la decisión estadística en términos del problema.

8.1. Pruebas de hipótesis

Definición 171 (prueba de hipótesis). Una prueba de hipótesis γ es una regla para decidir con fundamento en la muestra $\underline{X} \in \mathfrak{X}$ si la hipótesis nula H_0 es rechazada a favor de la hipótesis alternativa H_1 o si no es rechazada.

La prueba de hipótesis estadística γ es un procedimiento para tomar una decisión, bajo incertidumbre, sobre la validez de la hipótesis nula H_0 usando la evidencia de la muestra $\underline{X}$, entonces la determinación de la decisión estadística tiene una probabilidad de error.

Las pruebas de hipótesis son clasificadas como pruebas:

1. Unilaterales.
2. Bilaterales.

8.1.1. Prueba unilateral

Definición 172 (prueba unilateral). La región crítica $\mathscr{C}$ de la prueba γ es el complemento del extremo izquierdo o del extremo derecho de la función de distribución.

8.1.2. Prueba bilateral

Definición 173 (prueba bilateral). La región crítica $\mathscr{C}$ de la prueba γ es la unión del extremo izquierdo y del extremo derecho de la función de distribución.

Las clasificaciones de las pruebas estadísticas son resumidas en la tabla 8.1.

Tabla 8.1: Clasificaciones de las pruebas de hipótesis del conjunto de parámetros θ.

Pruebas unilaterales		Prueba bilateral
Extremo izquierdo	Extremo derecho	Ambos extremos
$H_0: \theta \geq \theta_0$	$H_0: \theta \leq \theta_0$	$H_0: \theta = \theta_0$
$H_1: \theta < \theta_0$	$H_1: \theta > \theta_0$	$H_1: \theta \neq \theta_0$

La tabla 8.1 presenta las dos clasificaciones de pruebas (unilaterales y bilaterales) y los tres tipos de pruebas (extremo izquierdo, extremo derecho y ambos extremos).

8.1.3. Región crítica

La información de la muestra aleatoria $\underline{X} = X_1, \ldots, X_n$ es resumida en una estadística de prueba $T(\underline{X})$ que define la región crítica $\mathscr{C}$ en función de la estadística de prueba utilizada, entonces el espacio de la muestra $\mathfrak{X}$ es dividido en las dos regiones $\mathscr{C}$ y $\mathscr{C}^{\mathrm{c}}$ que son mutuamente excluyentes, es decir, $\mathscr{C} \bigcap \mathscr{C}^{\mathrm{c}} = \emptyset$.

Definición 174 (región crítica). Una región crítica $\mathscr{C} \subset \mathfrak{X}$ es la región de rechazo de la hipótesis nula H_0 y es la probabilidad máxima de cometer un error de tipo I, es decir, es el nivel de significación α y es denotada por $\mathscr{C} = \{\underline{X} \in \mathfrak{X} \mid H_1\}$.

La región crítica $\mathscr{C} \subset \mathfrak{X}$ está definida con fundamento en la estadística de prueba $T(\underline{X})$. La regla de decisión indica que si la estadística de prueba $T(\underline{X})$ pertenece a la región de crítica $\mathscr{C}$, es decir, $T(\underline{X}) \in \mathscr{C}$, entonces la hipótesis nula H_0 es rechazada y si la estadística de prueba $T(\underline{X})$ no pertenece a la región crítica $\mathscr{C}$, es decir, $T(\underline{X}) \notin \mathscr{C}$ o equivalentemente $T(\underline{X}) \in \mathscr{C}^{\mathrm{c}}$, entonces la hipótesis nula H_0 no es rechazada.

8.1.4. Tipos de errores

Las pruebas de hipótesis están fundamentadas en una muestra aleatoria $\underline{X}$, por lo tanto la toma de decisiones tiene decisiones correctas e incorrectas.

Definición 175 (error de tipo I). Un error de tipo I significa que la hipótesis nula H_0 es rechazada cuando es verdadera con probabilidad $\alpha = \mathscr{P}\left(\underline{X} \in \mathscr{C} \mid H_0\right)$.

Definición 176 (error de tipo II). Un error de tipo II significa que la hipótesis nula H_0 no es rechazada cuando es falsa con probabilidad $\beta = \mathscr{P}\left(\underline{X} \notin \mathscr{C} \mid H_1\right) = \mathscr{P}\left(\underline{X} \in \mathscr{C}^{\mathrm{c}} \mid H_1\right)$.

8.1.5. Característica de operación

Definición 177 (característica de operación). La característica de operación es la probabilidad máxima β de no rechazar la hipótesis nula H_0 cuando es falsa, es decir, es la probabilidad cometer un error de tipo II.

Definición 178 (curva característica de operación). La curva característica de operación es la probabilidad máxima β de no rechazar la hipótesis nula H_0 cuando es falsa, es decir, es la probabilidad cometer un error de tipo II en función de los parámetros $\theta \in \Theta$, la hipótesis nula H_0, el nivel de significación α y el tamaño de la muestra n.

El nivel de significación α, la característica de operación β, el nivel de confianza $1-\alpha$ y la potencia de la prueba $1-\beta$ son resumidos en la tabla 8.2.

Tabla 8.2: Resultados de la prueba de hipótesis.

Decisión sobre H_0	Situación sobre H_0	
	H_0 es verdadera	H_0 es falsa (H_1 es verdadera)
No rechazar H_0	Decisión correcta $1-\alpha = \mathscr{P}\left(\underline{X} \notin \mathscr{C} \mid H_0\right)$ Nivel de confianza	Error tipo II $\beta = \mathscr{P}\left(\underline{X} \notin \mathscr{C} \mid H_1\right)$ Característica de operación
Rechazar H_0	Error tipo I $\alpha = \mathscr{P}\left(\underline{X} \in \mathscr{C} \mid H_0\right)$ Nivel de significación	Decisión correcta $1-\beta = \mathscr{P}\left(\underline{X} \in \mathscr{C} \mid H_1\right)$ Potencia de la prueba

La tabla 8.2 presenta los cuatro resultados de una prueba de hipótesis:

1. Decisión correcta, es decir, no rechazar H_0 cuando H_0 es verdadera con nivel de confianza $1-\alpha$.
2. Error de tipo II, es decir, no rechazar H_0 cuando H_0 es falsa con probabilidad β.
3. Error de tipo I, es decir, rechazar H_0 cuando H_0 es verdadera con probabilidad α.
4. Decisión correcta, es decir, rechazar H_0 cuando H_0 es falsa con probabilidad $1-\beta$.

8.1.6. Potencia de la prueba

El concepto de potencia es atribuido a Neyman y Pearson (1928a, 1928b, 1933a, 1933b, 1936). La potencia $\pi(\theta)$ mide la bondad de una prueba de hipótesis γ.

Definición 179 (potencia de la prueba). Dada una región crítica $\mathscr{C} \subset \mathfrak{X}$, entonces la potencia $\pi(\theta)$ de la prueba de hipótesis γ acerca de un parámetro θ desconocido es:

$$\pi(\theta) = \mathscr{P}(\underline{X} \in \mathscr{C} \mid \theta). \tag{8.1}$$

La función potencia óptima es:

$$\pi^*(\theta) = \mathscr{P}(\underline{X} \in \mathscr{C} \mid \theta) = \begin{cases} 0 & \text{si } \theta \in \Theta_0 \\ 1 & \text{si } \theta \in \Theta_1 \end{cases}. \tag{8.2}$$

Definición 180 (tamaño de la prueba). Sea $H_0\colon \theta \in \Theta_0$ y $H_1\colon \theta \in \Theta_1$, entonces el tamaño de la prueba es:

$$\underset{\theta \in \Theta_0}{\text{maximizar}}\, \pi(\theta). \tag{8.3}$$

El tamaño de la prueba es la probabilidad máxima de tomar una decisión incorrecta bajo el supuesto de que el valor del parámetro θ es verdadero.

Si γ es una prueba de hipótesis para $H_0\colon \theta = \theta_0$ y $H_1\colon \theta = \theta_1$ que son hipótesis simples donde $\Theta = \{\theta_0, \theta_1\}$ y $\pi(\theta)$ es la potencia de la prueba, entonces una prueba adecuada satisface:

1. $\pi(\theta_0) = \mathscr{P}(\underline{X} \in \mathscr{C} \mid \theta_0) = \alpha \to 0$.
2. $\pi(\theta_1) = \mathscr{P}(\underline{X} \in \mathscr{C} \mid \theta_1) = 1 - \beta \to 1$.

Por lo tanto, minimizar la característica de operación β es equivalente a maximizar la potencia de la prueba evaluada en la hipótesis alternativa H_1.

Definición 181 (prueba más potente). La prueba más potente γ^* de la prueba de hipótesis γ para $H_0\colon \theta = \theta_0$ y $H_1\colon \theta = \theta_1$ con $0 < \alpha < 1$ satisface:

1. $\pi^*(\theta_0) = \alpha$.
2. $\pi^*(\theta_1) \geq \pi(\theta_1) \quad \forall \gamma$ donde $\pi(\theta_0) = \alpha$.

Definición 182 (región crítica óptima). La región crítica óptima $\mathscr{C}^*$ de la prueba de hipótesis γ para las hipótesis $H_0\colon \theta = \theta_0$ y $H_1\colon \theta = \theta_1$ con $0 < \alpha < 1$ satisface:

1. $\mathscr{P}(\underline{X} \in \mathscr{C}^* \mid H_0) = \alpha$.
2. $\mathscr{P}(\underline{X} \in \mathscr{C}^* \mid H_1) \geq \mathscr{P}(\underline{X} \in \mathscr{C} \mid H_1) \quad \forall \mathscr{C}$ donde $\mathscr{P}(\underline{X} \in \mathscr{C} \mid H_0) = \alpha$.

Lema 1 (Neyman y Pearson). La región crítica óptima $\mathscr{C}^*$ de la prueba más potente γ^* para $H_0\colon \theta = \theta_0$ y $H_1\colon \theta = \theta_1$ con $0 < \alpha < 1$ de una muestra aleatoria $\underline{X}$ de una población con función de densidad $f_X(x, \theta)$ donde $\theta \in \Theta = \{\theta_0, \theta_1\}$ y $c > 0$ es:

$$\mathscr{C}^* = \left\{ \underline{X} \,\middle|\, \frac{\mathscr{L}(\theta_0)}{\mathscr{L}(\theta_1)} < c \right\}, \tag{8.4}$$

donde $\mathscr{L}(\theta)$ es la función de verosimilitud de una muestra aleatoria $\underline{X}$ y c es una constante que hace que la región crítica óptima $\mathscr{C}^*$ sea de tamaño α.

8.1.7. Nivel de significación descriptivo

Un enfoque complementario para las pruebas de hipótesis es el nivel de significación descriptivo que indica que el resultado es significativo cuando el nivel de significación descriptivo es menor o igual que el nivel de significación especificado, entonces las conclusiones están fundamentadas en la estadística de prueba y, por lo tanto, el resultado no es considerado una casualidad.

Definición 183 (nivel de significación descriptivo). El nivel de significación descriptivo $\mathscr{P}(\underline{X})$ es la mínima probabilidad α para que la estadística $T(\underline{X}) \in \mathscr{C}$ dado que la hipótesis nula H_0 es cierta, entonces la prueba γ es significativa, por lo tanto:

$$\mathscr{P}(\underline{X}) = \max_{\theta \in \Theta} \left(\mathscr{P}(T(\underline{X}) = t(\underline{x})) \right), \tag{8.5}$$

donde $t(\underline{x})$ es el valor de la estadística $T(\underline{X})$ para los valores $\underline{x} \in \mathfrak{X}$.

Definición 184 (prueba estadísticamente significativa). Si el nivel de significación descriptivo $\mathscr{P}(\underline{X}) \to 0$, entonces la prueba γ es estadísticamente significativa y existe evidencia a favor de la hipótesis alternativa H_1, es decir, $T(\underline{X}) \in \mathscr{C}$ y la hipótesis nula H_0 es rechazada si y sólo si el nivel de significación descriptivo $\mathscr{P}(\underline{X}) \leq \alpha$.

Si el nivel de significación es establecido como $\alpha = 0.10$, $\alpha = 0.05$ o $\alpha = 0.01$, entonces cuando el nivel de significación descriptivo $\mathscr{P}(\underline{X}) \leq \alpha$, la hipótesis nula H_0 es rechazada y la prueba γ es estadísticamente significativa. Cuando el nivel de significación descriptivo $\mathscr{P}(\underline{X}) > 0.10$, entonces la hipótesis nula H_0 no es rechazada y, por lo tanto, la prueba γ es estadísticamente no significativa.

Proposición 61 (media μ de una $\mathrm{N}(\theta, \sigma^2)$ con $n \geq 30$). Si $\underline{X} \in \mathfrak{X}$ es una muestra aleatoria de tamaño $n \geq 30$ de una población con una distribución $\mathrm{N}(\theta, \sigma^2)$ donde $\theta \in \Theta$ es el parámetro desconocido, entonces la prueba bilateral γ para la hipótesis nula $H_0\colon \theta = \theta_0$ y la hipótesis alternativa $H_1\colon \theta \neq \theta_0$ presenta la estadística de prueba $T(\underline{X})$, donde $Z = \frac{\sqrt{n}(\bar{X} - \theta_0)}{\sigma}$ y la región crítica $\mathscr{C} = \left\{ \underline{X} \in \mathfrak{X} \;\middle|\; |Z| \geq Z_{1-\frac{\alpha}{2}} \right\}$, por lo tanto la potencia $\pi(\theta)$ de la prueba γ para $\theta = \theta_1$ es:

$$\pi(\theta_1) = 1 - \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{1-\frac{\alpha}{2}} \right) + \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{\frac{\alpha}{2}} \right). \tag{8.6}$$

Demostración. Por hipótesis $\underline{X} \sim \mathrm{N}(\theta, \sigma^2)$, entonces $\bar{X} \sim \mathrm{N}\left(\theta, \frac{\sigma^2}{n}\right)$, $Z = \frac{\sqrt{n}(\bar{X} - \theta_0)}{\sigma} \sim \mathrm{N}(0,1)$ y la región crítica es $\mathscr{C} = \left\{ \underline{X} \in \mathfrak{X} \mid |Z| \geq Z_{1-\frac{\alpha}{2}} \right\}$, por lo tanto la regla de decisión es rechazar H_0 cuando $|Z| \geq Z_{1-\frac{\alpha}{2}}$, entonces la característica de operación para $\theta = \theta_1$ es $\beta(\theta_1) = \mathscr{P}(\underline{X} \notin \mathscr{C} \mid \theta = \theta_1)$, por lo tanto:

$$\beta(\theta_1) = \mathscr{P}\left(|Z| < Z_{1-\frac{\alpha}{2}} \mid \theta = \theta_1 \right) = \mathscr{P}\left(Z_{\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}} \mid \theta = \theta_1 \right).$$

$$\beta(\theta_1) = \mathscr{P}\left(\theta_0 - \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}} < \bar{X} < \theta_0 + \frac{\sigma Z_{1-\frac{\alpha}{2}}}{\sqrt{n}} \mid \theta = \theta_1 \right).$$

$$\beta(\theta_1) = \mathscr{P}\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} - Z_{1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\bar{X} - \theta_1)}{\sigma} < \frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{1-\frac{\alpha}{2}} \right).$$

$$\beta(\theta_1) = \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{1-\frac{\alpha}{2}} \right) - \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{\frac{\alpha}{2}} \right).$$

Por lo tanto, la potencia de la prueba para $\theta = \theta_1$ es:

$$\pi(\theta_1) = 1 - \beta(\theta_1) = 1 - \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{1-\frac{\alpha}{2}}\right) + \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{\frac{\alpha}{2}}\right).$$

$\square$

Proposición 62 (media μ de una $\mathrm{N}(\theta, \sigma^2)$ con $n \geq 30$). Si $\underline{X} \in \mathfrak{X}$ es una muestra aleatoria de tamaño $n \geq 30$ de una población con una distribución $\mathrm{N}(\theta, \sigma^2)$ donde el parámetro desconocido $\theta \in \Theta$, entonces la prueba unilateral γ para la hipótesis nula $H_0: \theta = \theta_0$ y la hipótesis alternativa $H_1: \theta < \theta_0$ presenta la estadística de prueba $T(\underline{X})$ donde $Z = \frac{\sqrt{n}(\bar{X} - \theta_0)}{\sigma}$, la región crítica $\mathscr{C} = \{\underline{X} \in \mathfrak{X} \mid Z \leq Z_\alpha\}$, por lo tanto la potencia $\pi(\theta)$ de la prueba γ para $\theta = \theta_1$ es:

$$\pi(\theta_1) = \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_\alpha\right). \tag{8.7}$$

Proposición 63 (media μ de una $\mathrm{N}(\theta, \sigma^2)$ con $n \geq 30$). Si $\underline{X} \in \mathfrak{X}$ es una muestra aleatoria de tamaño $n \geq 30$ de una población con una distribución $\mathrm{N}(\theta, \sigma^2)$ donde el parámetro desconocido $\theta \in \Theta$, entonces la prueba unilateral γ para la hipótesis nula $H_0: \theta = \theta_0$ y la hipótesis alternativa $H_1: \theta > \theta_0$ presenta la estadística de prueba $T(\underline{X})$ donde $Z = \frac{\sqrt{n}(\bar{X} - \theta_0)}{\sigma}$, la región crítica $\mathscr{C} = \{\underline{X} \in \mathfrak{X} \mid Z \geq Z_{1-\alpha}\}$, por lo tanto la potencia $\pi(\theta)$ de la prueba γ para $\theta = \theta_1$ es:

$$\pi(\theta_1) = 1 - \Phi\left(\frac{\sqrt{n}(\theta_0 - \theta_1)}{\sigma} + Z_{1-\alpha}\right). \tag{8.8}$$

Las pruebas de hipótesis son aplicadas para:

1. La media de una población con distribución $\mathrm{N}(\theta, \sigma^2)$ con $n \geq 30$.
2. La media de una población con distribución $\mathrm{N}(\theta_1, \theta_2^2)$ con $n \geq 30$.
3. La media de una población con distribución $\mathrm{N}(\theta, \theta_2^2)$ con $n < 30$.
4. La proporción de una población con distribución $Bin(n, \theta)$ con $n \geq 30$.
5. La varianza de una población con distribución $\mathrm{N}(\mu, \theta^2)$.
6. La diferencia de medias de poblaciones con $\mathrm{N}(\theta_1, \sigma_1^2)$ y $\mathrm{N}(\theta_2, \sigma_2^2)$ con $n_1 + n_2 \geq 30$.
7. La diferencia de medias de poblaciones con $\mathrm{N}(\theta_1, \vartheta_1^2)$ y $\mathrm{N}(\theta_2, \vartheta_2^2)$ con $n_1 + n_2 \geq 30$.
8. La diferencia de medias de poblaciones con $\mathrm{N}(\theta_1, \vartheta_1^2)$ y $\mathrm{N}(\theta_2, \vartheta_2^2)$ con $n_1 + n_2 \geq 30$.
9. La diferencia de medias de poblaciones con $\mathrm{N}(\theta_1, \vartheta_1^2)$ y $\mathrm{N}(\theta_2, \vartheta_2^2)$ con $n_1 + n_2 < 30$.
10. La diferencia de proporciones poblacionales con $Bin(n_1, \vartheta_1)$ y $Bin(n_2, \vartheta_2)$ con $\vartheta_1 \neq \vartheta_2$ y $n_1 + n_2 \geq 30$.
11. El cociente de varianzas de poblaciones con $\mathrm{N}(\theta_1, \vartheta_1^2)$ y $\mathrm{N}(\theta_2, \vartheta_2^2)$.

8.2. Pruebas de hipótesis para la media y la proporción

8.2.1. Prueba de hipótesis para μ de una $\mathrm{N}\left(\mu,\sigma^2\right)$ con σ^2 conocida y $n \geq 30$

Proposición 64 (estadístico de prueba). Sea $X_1,\ldots,X_n$ una muestra aleatoria de una población que presenta una distribución $\mathrm{N}\left(\mu,\sigma^2\right)$ con media μ desconocida y varianza σ^2 conocida, entonces:

$$Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}. \tag{8.9}$$

Ejemplo 97 (prueba de hipótesis bilateral para la media). Los pesos (en kilogramos) de los Vikingos de Minnesota[1] son presentados en la tabla 8.3.

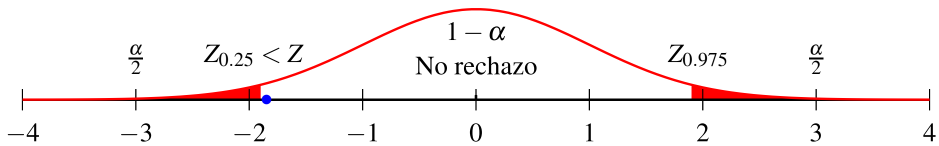
Tabla 8.3: Peso en kilogramos de los Vikingos de Minnesota 2019.

92.162	87.168	86.260	115.770	83.082	93.524	90.800	138.470
109.868	95.340	115.316	95.340	91.708	111.230	88.530	86.714
93.070	141.648	137.562	110.776	123.942	106.690	91.708	88.984
142.102	128.482	85.806	114.408	92.616	143.464	143.010	149.366
97.610	105.328	136.200	104.420	115.316	99.880	120.310	134.838
117.132	138.470	98.972	120.310	139.832	95.340	97.156	109.868
140.286	90.800	97.610	147.550	133.930	86.260	120.310	104.420

Indicar si la media de la muestra $\bar{X}$ es diferente que $\mu = 116$ con un nivel de confianza de 95 % cuando es conocido que $\sigma = 21$.
Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : \bar{X} = \mu$.
2. La hipótesis alternativa es $H_1 : \bar{X} \neq \mu$.
3. El nivel de significación es $\alpha = 0.05$.
4. El tamaño de la muestra es $n = 56$.
5. La estadística de prueba es $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$.
6. Los puntos críticos son $Z_{\frac{\alpha}{2}} = Z_{0.025} = -1.96$ y $Z_{1-\frac{\alpha}{2}} = Z_{0.975} = 1.96$.
7. La regla de decisión es: si $-1.96 \leq Z \leq 1.96$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.9): $Z = \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{111.1976-116}{\frac{21}{\sqrt{56}}} = -1.8521$ donde $-1.96 \leq -1.85 \leq 1.96$, entonces $Z \notin \mathscr{C}$, es decir, Z está en la región de no rechazo $Z \in \mathscr{C}^{c}$ como se muestra en la gráfica 8.1.

[1] www.nfl.com/teams/roster?team=MIN consultado el 11 de noviembre de 2019.



Gráfica 8.1: Región de no rechazo de la prueba bilateral para la media de los Vikingos.

9. La hipótesis nula $H_0 : \bar{X} = \mu$ no es rechazada.
10. Por lo tanto, el peso promedio de los Vikingos es igual que el peso promedio de la NFL y la diferencia es aleatoria. □

8.2.2. Prueba de hipótesis para μ de una $\mathrm{N}\left(\mu, \sigma^2\right)$ con σ^2 desconocida y $n \geq 30$

Proposición 65 (estadístico de prueba). Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población que presenta una distribución $\mathrm{N}\left(\mu, \sigma^2\right)$ con media μ desconocida y varianza σ^2 desconocida, entonces:

$$Z = \frac{\bar{X} - \mu}{\frac{S_X}{\sqrt{n}}}. \tag{8.10}$$

Ejemplo 98 (prueba de hipótesis unilateral para la media). Los pesos en kilogramos de los Pieles Rojas de Washington[2] son presentados en la tabla 8.4.

Tabla 8.4: Peso en kilogramos de los Pieles Rojas de Washington 2019.

136.200	115.770	93.070	140.740	111.230	114.408	138.470	108.960
143.010	98.972	86.714	113.500	112.592	91.708	92.162	149.820
103.512	102.150	106.690	97.610	111.230	104.874	111.230	138.016
108.960	93.070	108.052	140.740	86.260	97.610	120.310	143.010
96.248	95.340	92.616	82.628	149.820	96.248	90.800	145.280
143.010	99.880	143.010	90.800	90.800	109.868	81.720	93.070
141.648	143.010	139.832	86.260	94.432	96.702	113.954	115.770
82.628	111.230	118.948	89.438	88.530	104.420	99.880	104.420

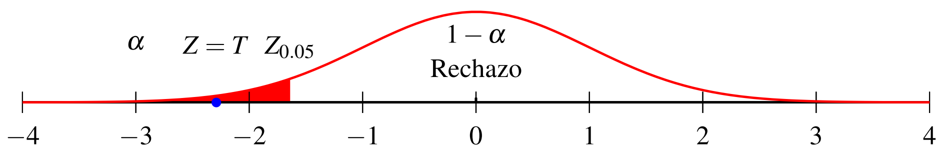
Indicar si la media $\bar{X}$ es menor que $\mu = 116$ con un nivel de confianza de 95 %.

Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : \bar{X} < \mu$.
2. La hipótesis alternativa es $H_1 : \bar{X} \geq \mu$.
3. El nivel de significación es $\alpha = 0.05$.
4. El tamaño de la muestra es $n = 64$.
5. La estadística de prueba es $Z = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}$ o alternativamente $T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}$.

[2]www.nfl.com/teams/roster?team=WAS consultado el 11 de noviembre de 2019.

6. El punto crítico es $Z_\alpha = Z_{0.05} = -1.6449$ o alternativamente $t_{0.05}(63) = -1.6694$.
7. La regla de decisión es: si $Z < -1.6449$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.10) o en la ecuación (8.11): $Z = T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}} = \frac{110.2014-116}{\frac{20.2622}{\sqrt{64}}} = -2.2894$, donde $-2.2894 \leq -1.6449$, entonces $Z \notin \mathscr{C}$, es decir, Z está en la región de no rechazo $Z \in \mathscr{C}^c$ como se muestra en la gráfica 8.2.



Gráfica 8.2: Región de no rechazo de la prueba unilateral para la media de los Pieles Rojas.

9. La hipótesis nula $H_0 : \bar{X} < \mu$ no es rechazada.
10. Por lo tanto, el peso promedio de los Pieles Rojas es menor que el promedio de la NFL. □

8.2.3. Prueba de hipótesis para μ de una $\mathrm{N}(\mu, \sigma^2)$ con σ^2 desconocida y $n < 30$

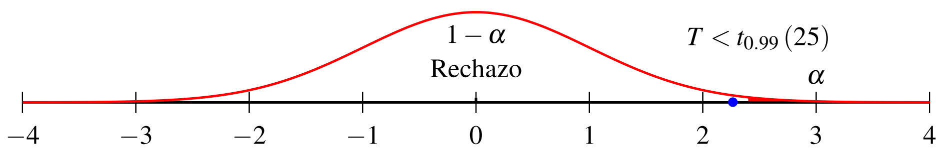
Proposición 66 (estadístico de prueba). Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población que presenta una distribución $\mathrm{N}(\mu, \sigma^2)$ con media μ desconocida y varianza σ^2 desconocida, entonces:

$$T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}. \tag{8.11}$$

Ejemplo 99 (prueba de hipótesis unilateral para la media). El cociente de inteligencia de una muestra 26 estudiantes de una una universidad tiene una media $\bar{X} = 108$ con una desviación estándar $S_X = 9$. Indicar si la media $\bar{X}$ es mayor que la media universitaria $\mu = 104$ con un nivel de confianza de 99 % y 95 %.

Solución. El modelo de decisión con un nivel de confianza de 99 % es:

1. La hipótesis nula es $H_0 : \bar{X} > \mu$.
2. La hipótesis alternativa es $H_1 : \bar{X} \leq \mu$.
3. El nivel de significación es $\alpha = 0.01$.
4. El tamaño de la muestra es $n = 26$.
5. La estadística de prueba es $T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}$.
6. El punto crítico es $t_{1-\alpha}(n-1) = t_{0.99}(25) = 2.4851$.
7. La regla de decisión es: si $T > 2.4851$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.11): $T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}} = \frac{108-104}{\frac{9}{\sqrt{26}}} = 2.2662$ donde $2.2662 < 2.4851$, entonces $T \in \mathscr{C}$, es decir, T está en la región de rechazo $T \notin \mathscr{C}^c$ como se muestra en la gráfica 8.3.

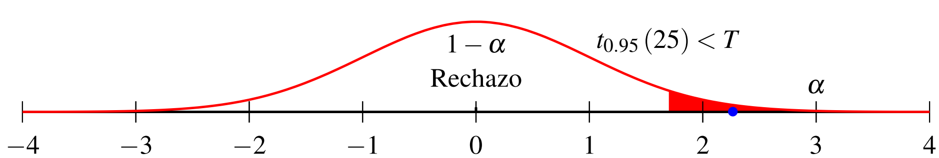


Gráfica 8.3: Región de rechazo de la prueba unilateral para la media de estudiantes.

9. La hipótesis nula $H_0 : \bar{X} > \mu$ es rechazada.
10. Por lo tanto, el cociente de inteligencia promedio de la universidad no es mayor que el promedio universitario con un nivel de confianza de 99 %.

El modelo de decisión con un nivel de confianza de 95 % es:

1. La hipótesis nula es $H_0 : \bar{X} > \mu$.
2. La hipótesis alternativa es $H_1 : \bar{X} \leq \mu$.
3. El nivel de significación es $\alpha = 0.05$.
4. El tamaño de la muestra es $n = 26$.
5. La estadística de prueba es $T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}}$.
6. El punto crítico es $t_{1-\alpha}\,(n-1) = t_{0.95}\,(25) = 1.7081$.
7. La regla de decisión es: si $T > 1.7081$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.11): $T = \frac{\bar{X}-\mu}{\frac{S_X}{\sqrt{n}}} = \frac{108-104}{\frac{9}{\sqrt{26}}} = 2.2662$ donde $2.2662 > 1.7081$, entonces $T \notin \mathscr{C}$, es decir, T está en la región de no rechazo $T \in \mathscr{C}^{\mathrm{c}}$ como se muestra en la gráfica 8.4.



Gráfica 8.4: Región de no rechazo de la prueba unilateral para la media de estudiantes.

9. La hipótesis nula $H_0 : \bar{X} > \mu$ no es rechazada.
10. Por lo tanto, el cociente de inteligencia promedio de la universidad es mayor que el promedio universitario con un nivel de confianza de 95 %.

Es decir, el nivel de significación descriptivo es $\mathscr{P}\,(\underline{X}) = 0.0162$, por lo tanto, la hipótesis nula $H_0 : \bar{X} > \mu$ no es rechazada con un nivel de significación mínimo de 1.62 % $(\alpha > 0.0162)$. $\square$

8.2.4. Prueba de hipótesis para p de una $Bin(n,p)$

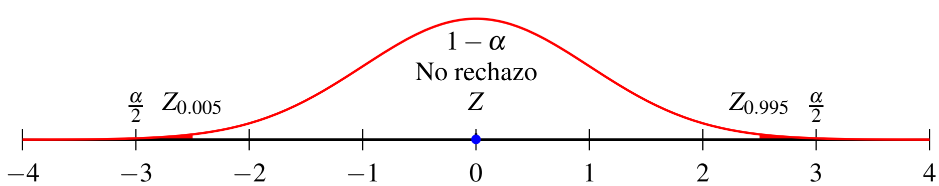
Proposición 67 (estadístico de prueba). Sea $X_1,\dots,X_n$ una muestra aleatoria de una población que presenta una distribución $Bin(n,p)$ con una proporción p desconocida y varianza $S_X^2 = n\hat{P}\left(1-\hat{P}\right)$ donde $\hat{P} = P_X$ y $n \geq 30$, entonces:

$$Z = \frac{\hat{P}-p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}}. \tag{8.12}$$

Ejemplo 100 (prueba de hipótesis bilateral para la proporción). Los pesos (en kilogramos) de los Vikingos son presentados en la tabla 8.3 del ejemplo 97. Indicar si la proporción $\hat{P}$ de jugadores profesionales con un mayor peso que $\tilde{X} = 108.279$ es igual que la proporción $p = \frac{1}{2}$ con un nivel de confianza de 99 % y 95 %.

Solución. El modelo de decisión con un nivel de confianza de 99 % es:

1. La hipótesis nula es $H_0 : \hat{P} = p$.
2. La hipótesis alternativa es $H_1 : \hat{P} \neq p$.
3. El nivel de significación es $\alpha = 0.01$.
4. El tamaño de la muestra es $n = 56$.
5. La estadística de prueba es $Z = \frac{\hat{P}-p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}}$.
6. Los puntos críticos son $Z_{\frac{\alpha}{2}} = Z_{0.005} = -2.5758$ y $Z_{1-\frac{\alpha}{2}} = Z_{0.995} = 2.5758$.
7. La regla de decisión es: si $-2.5758 < Z < 2.5758$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.12): $Z = \frac{\hat{P}-p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}} = \frac{\frac{1}{2}-\frac{1}{2}}{\sqrt{\frac{\frac{1}{2}\left(\frac{1}{2}\right)}{56}}} = 0$ donde $-2.5758 < 0 < 2.5758$, entonces $Z \notin \mathscr{C}$, es decir, Z está en la región de no rechazo $Z \in \mathscr{C}^c$ como se muestra en la gráfica 8.5.



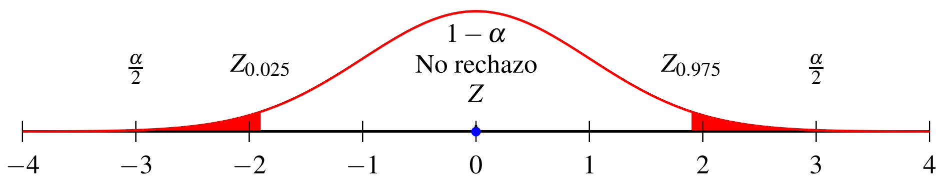
Gráfica 8.5: Región de no rechazo de la prueba bilateral de la proporción de los Vikingos al 99 %.

9. La hipótesis nula $H_0 : \hat{P} = p$ no es rechazada.
10. Por lo tanto, la proporción de jugadores con mayor peso que la mediana es igual que la proporción esperada con un nivel de confianza de 99 %.

El modelo de decisión con un nivel de confianza de 95 % es:

1. La hipótesis nula es $H_0 : \hat{P} = p$.
2. La hipótesis alternativa es $H_1 : \hat{P} \neq p$.
3. El nivel de significación es $\alpha = 0.05$.
4. El tamaño de la muestra es $n = 56$.
5. La estadística de prueba es $Z = \frac{\hat{P}-p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}}$.

6. Los puntos críticos son $Z_{\frac{\alpha}{2}} = -1.96$ y $Z_{1-\frac{\alpha}{2}} = 1.96$.
7. La regla de decisión es: si $-1.96 < Z < 1.96$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.12): $Z = \frac{\hat{P}-p}{\sqrt{\frac{\hat{P}(1-\hat{P})}{n}}} = \frac{\frac{1}{2}-\frac{1}{2}}{\sqrt{\frac{\frac{1}{2}\left(\frac{1}{2}\right)}{56}}} = 0$ donde $-1.96 < 0 < 1.96$, entonces $Z \notin \mathscr{C}$, es decir, Z está en la región de no rechazo $Z \in \mathscr{C}^c$ como se muestra en la gráfica 8.6.



Gráfica 8.6: Región de no rechazo de la prueba bilateral de la proporción de los Vikingos al 95 %.

9. La hipótesis nula $H_0 : \hat{P} = p$ no es rechazada.
10. Por lo tanto, la proporción de jugadores con mayor peso que la mediana es igual que la proporción esperada con un nivel de confianza de 95 %. □

8.3. Prueba de hipótesis para la varianza

8.3.1. Prueba de hipótesis para σ^2 de una $\mathrm{N}\left(\mu, \sigma^2\right)$

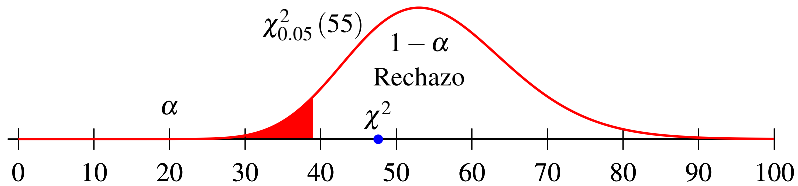
Proposición 68 (estadístico de prueba). Sea $X_1, \ldots, X_n$ una muestra aleatoria de una población que presenta una distribución $\mathrm{N}\left(\mu, \sigma^2\right)$ con media μ conocida y varianza σ^2 desconocida, entonces:

$$\chi^2 = \frac{nS_X^2}{\sigma^2}. \tag{8.13}$$

Ejemplo 101 (prueba de hipótesis unilateral para la varianza). Los pesos en kilogramos de los Vikingos son presentados en la tabla 8.3 del ejemplo 97. Indicar si la desviación estándar $S_X = 20.2826$ del peso de los jugadores es menor que la desviación estándar $\sigma = 22$ con un nivel de confianza de 95 %.

Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : S_X < \sigma$.
2. La hipótesis alternativa es $H_1 : S_X \geq \sigma$.
3. El nivel de significación es $\alpha = 0.05$.
4. El tamaño de la muestra es $n = 56$.
5. La estadística de prueba es $\chi^2 = \frac{nS_X^2}{\sigma^2}$.
6. El punto crítico es $\chi^2_{\alpha}(n-1) = \chi^2_{0.05}(55) = 38.9580$.
7. La regla de decisión es: si $\chi^2 < 38.9580$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.13): $\chi^2 = \frac{nS_X^2}{\sigma^2} = 56\left(\frac{20.2826}{22}\right)^2 = 47.5981$ donde $47.5981 > 38.9580$, entonces $\chi^2 \in \mathscr{C}$, es decir, χ^2 está en la región de rechazo $Z \notin \mathscr{C}^c$ como se muestra en la gráfica 8.7.



Gráfica 8.7: Región de rechazo de la prueba unilateral de la desviación estándar de los Vikingos.

9. La hipótesis nula $H_0 : S_X < \sigma$ es rechazada.
10. Por lo tanto, la desviación estándar de los pesos de los jugadores del equipo no es menor que la desviación estándar de la liga.

La prueba de hipótesis para la varianza donde $H_0 : S_X = \sigma$, $H_1 : S_X \neq \sigma$ con un nivel de confianza de 95 %, entonces $36.3981 < 47.5981 < 77.3805$, por lo tanto la desviación estándar de los pesos de los jugadores de los Vikingos es igual que la desviación estándar de los pesos de los jugadores de la liga. □

8.4. Prueba de hipótesis para la diferencia de medias y de proporciones

8.4.1. Prueba de hipótesis para $\mu_X - \mu_Y$ con σ_X^2, σ_Y^2 conocidas y $n_X + n_Y \geq 30$

Proposición 69 (estadístico de prueba). Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias con medias $\mu_X = \mu_Y$ desconocidas y varianzas σ_X^2, σ_Y^2 conocidas, entonces:

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}. \tag{8.14}$$

8.4.2. Prueba de hipótesis para $\mu_X - \mu_Y$ con σ_X^2, σ_Y^2 desconocidas y $n_X + n_Y \geq 30$

Proposición 70 (estadístico de prueba). Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias con medias $\mu_X = \mu_Y$ desconocidas y varianzas σ_X^2, σ_Y^2 desconocidas, entonces:

$$Z = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}}. \tag{8.15}$$

Ejemplo 102 (prueba de hipótesis bilateral para la diferencia de medias). Los pesos (en kilogramos) de los Tigres de la Universidad Estatal de Luisina[3] son presentados en la tabla 8.5.

[3]https://lsusports.net/sports/football/roster/2019 consultado el 11 de noviembre de 2019.

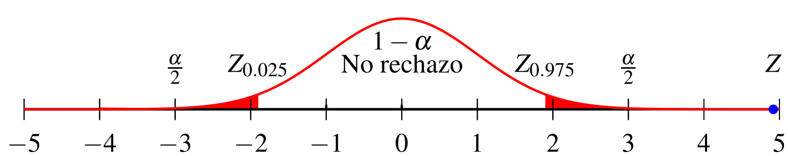
Tabla 8.5: Peso en kilogramos de los Tigres de la Universidad Estatal de Luisina 2019.

108.000	108.000	103.680	114.480	122.040	123.120	109.620	102.600	120.420	99.900
108.000	125.820	109.620	102.600	122.580	104.760	116.640	130.680	112.320	117.720
117.180	99.900	115.560	100.440	111.780	90.180	119.340	135.000	100.440	111.780
112.860	92.340	123.120	103.140	116.100	102.600	86.400	103.680	97.200	97.200
86.940	96.120	111.780	106.380	100.980	117.180	97.200	108.000	100.980	129.060
102.060	129.600	90.180	131.760	131.220	125.280	136.080	88.560	122.580	135.000
139.320	130.140	123.120	135.000	154.440	136.620	126.900	129.060	132.840	169.560
132.300	147.960	124.200	142.560	181.440	115.020	177.660	113.400	162.000	191.160
175.500	167.400	158.760	179.280	143.640	159.300	194.400	186.840	185.220	172.800
191.700	173.880	159.300	184.140	170.100	139.320	134.460	106.380	105.300	142.020
125.820	93.960	104.760	104.760	166.320	157.140	160.920	138.240	155.520	166.860

Indicar si la media $\bar{X}$ de los Tigres (equipo colegial) es diferente que la media $\bar{Y}$ de los Pieles Rojas (equipo profesional) del ejemplo 98 con un nivel de confianza del 95 %.

Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : \bar{X} = \bar{Y}$.
2. La hipótesis alternativa es $H_1 : \bar{X} \neq \bar{Y}$.
3. El nivel de significación es $\alpha = 0.05$.
4. Los tamaños de las muestras son $n_X = 64$ y $n_X = 110$.
5. La estadística de prueba es $Z = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_X^2}{n_X}+\frac{S_Y^2}{n_Y}}}$ o alternativamente $T = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_X^2}{n_X}+\frac{S_Y^2}{n_Y}}}$.
6. Los puntos críticos son $Z_{0.025} = -1.96$ y $Z_{0.975} = 1.96$ o $t_{0.025}(169) = -1.97$ y $t_{0.975}(169) = -1.97$.
7. La regla de decisión es: si $-1.96 < Z < 1.96$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.15): $Z = \frac{\bar{X}-\bar{Y}}{\sqrt{\frac{S_X^2}{n_X}+\frac{S_Y^2}{n_Y}}} = Z = \frac{128.3138-110.2014}{\sqrt{\frac{27.9781^2}{110}+\frac{20.2622^2}{64}}} = 4.92$ donde $4.92 > 1.96$, entonces $Z \in \mathscr{C}$, es decir, Z está en la región de rechazo $Z \notin \mathscr{C}^c$ como se muestra en la gráfica 8.8.



Gráfica 8.8: Región de rechazo de la prueba bilateral de la diferencia de medias.

9. La hipótesis nula $H_0 : \bar{X} = \bar{Y}$ es rechazada.
10. Por lo tanto, el peso promedio de los Tigres (equipo colegial) es diferente que el peso promedio de los Pieles Rojas (equipo profesional).

La prueba de hipótesis unilateral para la diferencia de medias donde $H_0 : \bar{X} > \bar{Y}$, $H_1 : \bar{X} \leq \bar{Y}$ con un nivel de confianza de 95 %, entonces $1.6449 < 4.9239$, por lo tanto el promedio de peso de los jugadores colegiales es mayor que el promedio de peso de los jugadores profesionales. □

8.4.3. Prueba de hipótesis para $\mu_X - \mu_Y$ con σ_X^2, σ_Y^2 desconocidas y $n_X + n_Y \geq 30$

Proposición 71 (estadístico de prueba). Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias con medias $\mu_X = \mu_Y$ desconocidas y varianzas $\sigma_X^2 = \sigma_Y^2$ desconocidas, entonces:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{(n_X-1)S_X^2 + (n_Y-1)S_Y^2}{n_X+n_Y-2}\right)\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}}. \tag{8.16}$$

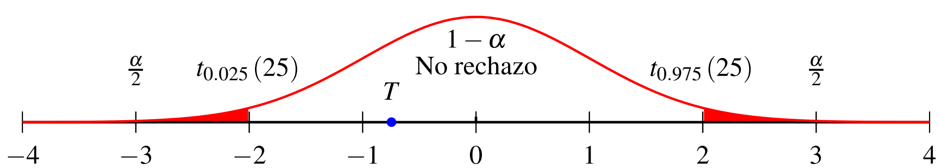
Ejemplo 103 (prueba de hipótesis bilateral para la diferencia de medias). El cociente de inteligencia de una muestra $n_X = 26$ estudiantes de una universidad tiene una media $\bar{X} = 108$ con una desviación estándar $S_X = 9$ y el cociente de inteligencia de una muestra $n_Y = 24$ estudiantes de otra universidad tiene una media $\bar{Y} = 110$ con desviación estándar $S_Y = 10$. Indicar si existe diferencia entre el cociente de inteligencia de ambas universidades con un nivel de significación de $5\,\%$.

Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : \bar{X} = \bar{Y}$.
2. La hipótesis alternativa es $H_1 : \bar{X} \neq \bar{Y}$.
3. El nivel de significación es $\alpha = 0.05$.
4. Los tamaños de las muestras son $n_X = 26$ y $n_Y = 24$.
5. La estadística de prueba es $T = \frac{\bar{X}-\bar{Y}}{\sqrt{\left(\frac{(n_X-1)S_X^2+(n_Y-1)S_Y^2}{n_X+n_Y-2}\right)\left(\frac{1}{n_X}+\frac{1}{n_Y}\right)}}$.
6. Los puntos críticos son $t_{\frac{\alpha}{2}}(n_X + n_Y - 2) = -2.01$ y $t_{1-\frac{\alpha}{2}}(n_X + n_Y - 2) = 2.01$.
7. La regla de decisión es: si $-2.01 < T < 2.01$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.17):

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{(n_X-1)S_X^2 + (n_Y-1)S_Y^2}{n_X+n_Y-2}\right)\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}} = \frac{108 - 110}{\sqrt{\left(\frac{25(9)^2+23(10)^2}{26+24-2}\right)\left(\frac{1}{26} + \frac{1}{24}\right)}} = -0.7443,$$

donde $-2.01 < -0.7443 < 2.01$, entonces $T \notin \mathscr{C}$, es decir, T no está en la región de rechazo $T \in \mathscr{C}^c$ como se muestra en la gráfica 8.9.



Gráfica 8.9: Región de no rechazo de la prueba bilateral de la diferencia de medias

9. La hipótesis nula $H_0 : \bar{X} = \bar{Y}$ no es rechazada.
10. Por lo tanto, los promedios de los cocientes de inteligencia de ambas universidades son iguales con un nivel de confianza de $95\,\%$. □

8.4.4. Prueba de hipótesis para $\mu_X - \mu_Y$ con $\sigma_X^2 \neq \sigma_Y^2$ desconocidas y $n_X + n_Y < 30$

Proposición 72 (estadístico de prueba). Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias con medias $\mu_X = \mu_Y$ desconocidas y varianzas $\sigma_X^2 \neq \sigma_Y^2$ desconocidas, entonces:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)}}, \tag{8.17}$$

donde $\nu = \left\lceil \frac{\left(\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}\right)^2}{\frac{\left(\frac{S_X^2}{n_X}\right)^2}{n_X+1} + \frac{\left(\frac{S_Y^2}{n_Y}\right)^2}{n_Y+1}} - 2 \right\rceil$ para los puntos críticos $t_{\frac{\alpha}{2}}(\nu)$ y $t_{1-\frac{\alpha}{2}}(\nu)$.

8.4.5. Prueba de hipótesis para $p_X - p_Y$ de $Bin(n_X, p_X)$ y $Bin(n_Y, p_Y)$

Proposición 73 (estadístico de prueba). Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias de dos poblaciones con distribuciones $Bin(n_X, p_X), Bin(n_Y, p_Y)$, respectivamente, con proporciones $p_X = p_Y$ desconocidas y varianzas $\sigma_X^2 = n_X \hat{P}_X \left(1 - \hat{P}_X\right), \sigma_Y^2 = n_Y \hat{P}_Y \left(1 - \hat{P}_Y\right)$ donde $n \geq 30$, entonces:

$$Z = \frac{\bar{P}_X - \bar{P}_Y}{\sqrt{\bar{P}(1-\bar{P})\left(\frac{1}{n_X} + \frac{1}{n_Y}\right)}}, \tag{8.18}$$

donde $\bar{P} = \frac{n_X \hat{P}_X + n_Y \hat{P}_Y}{n_X + n_Y}$.

8.5. Prueba de hipótesis para el cociente de varianzas

8.5.1. Prueba de hipótesis para $\frac{\sigma_X^2}{\sigma_Y^2}$ de $\mathrm{N}\left(\mu_X, \sigma_X^2\right)$ y $\mathrm{N}\left(\mu_Y, \sigma_Y^2\right)$

Proposición 74 (estadístico de prueba). Sean $X_1, X_2, \ldots, X_n$ y $Y_1, Y_2, \ldots, Y_n$ dos muestras aleatorias con varianzas σ_X^2 y σ_Y^2 desconocidas, entonces:

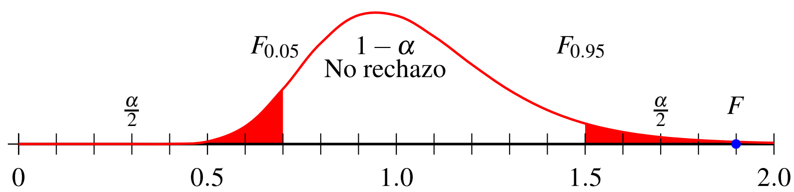
$$F = \frac{S_X^2}{S_Y^2}. \tag{8.19}$$

Ejemplo 104 (prueba de hipótesis bilateral para el cociente de varianzas). Los pesos (en kilogramos) de un equipo colegial (Tigres) y un equipo profesional (Vikingos) son presentados en la tabla 8.5 del ejemplo 102 y en la tabla 8.3 del ejemplo 97, respectivamente. Indicar si existe diferencia entre las desviaciones estándar de ambos equipos con un nivel de significación de 10%.

Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : \sigma_X = \sigma_Y$.
2. La hipótesis alternativa es $H_1 : \sigma_X \neq \sigma_Y$.
3. El nivel de significación es $\alpha = 0.10$.

4. Los tamaños de las muestras son $n_X = 110$ y $n_Y = 56$.
5. La estadística de prueba es $F = \frac{S_X^2}{S_Y^2}$.
6. Los puntos críticos[4] son $F_{\frac{\alpha}{2}}(n_X - 1, n_Y - 1) = F_{0.05}(109, 55) = 0.69$ y $F_{0.95}(109, 55) = 1.49$.
7. La regla de decisión es: si $0.69 < F < 1.49$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.19): $F = \frac{S_X^2}{S_Y^2} = \left(\frac{27.9781}{20.2826}\right)^2 = 1.90$ donde $1.90 > 1.49$, entonces $F \in \mathscr{C}$, es decir, F está en la región de rechazo $F \notin \mathscr{C}^c$ como se muestra en la gráfica 8.10.



Gráfica 8.10: Región de rechazo de la prueba bilateral para el cociente de varianzas

9. La hipótesis nula $H_0 : \sigma_X = \sigma_Y$ es rechazada.
10. Por lo tanto, las varianzas de ambos equipos son diferentes con un nivel de confianza de 95 %.

La prueba de hipótesis unilateral para el cociente de varianzas donde $H_0 : \sigma_X > \sigma_Y$, $H_1 : \sigma_X \leq \sigma_Y$ con un nivel de confianza de 95 %, entonces $1.90 > 1.49$, por lo tanto la varianza del peso de los jugadores colegiales es mayor que la varianza del peso de los jugadores profesionales, es decir, el nivel de significación descriptivo es $\mathscr{P}(\underline{X}) = 0.0046$, por lo tanto la prueba de hipótesis unilateral no es rechazada con un nivel de significación mínimo de 0.46 %. □

8.6. Pruebas de bondad de ajuste

Las pruebas de bondad de ajuste son utilizadas para comprobar si los datos de una muestra aleatoria provienen de una distribución de probabilidad. Las pruebas de bondad de ajuste utilizadas son:

1. Pearson (1900).
2. Cramer (1928) y Mises (1947).
3. Kolmogorov (1933b) y Smirnov (1933).
4. Anderson y Darling (1954).
5. Shapiro y Wilk (1965).
6. Shapiro y Francia (1972).
7. Jarque y Bera (1987).

[4] En la tabla D.1 son observados los valores para la distribución $F_{0.95}(\nu_X, \nu_Y)$ donde $\nu_X = n_X - 1$ y $\nu_Y = n_Y - 1$.

8.6.1. Prueba de bondad de ajuste Kolmogorov y Smirnov

La prueba de bondad de ajuste Kolmogorov y Smirnov plantea las hipótesis siguientes:

1. H_0: La muestra aleatoria presenta una distribución de probabilidad particular.
2. H_1: La muestra aleatoria no presenta una distribución de probabilidad particular.

La estadística de prueba es:

$$D = \max_{1\leq k\leq n} \left(\left|\hat{F}_X(X_k) - F_X(X_k)\right|\right), \tag{8.20}$$

donde D es el estadístico de prueba Kolmogorov y Smirnov, $\hat{F}_X(X_k) = \frac{F_{k-1}+1}{n+1}$ es la distribución acumulada observada de la k–ésima variable aleatoria, X_k es la k–ésima variable aleatoria (j–ésimo intervalo de clase) y $F_X(X_k) = \Phi(X_k) = \Phi(Z_k)$ es la distribución acumulada teórica de la k–ésima variable aleatoria (j–ésimo intervalo de clase) con $Z_k = \frac{X_k - \bar{X}}{S_X}$.

La regla de decisión es:

1. Si $D > D_\alpha$, entonces la hipótesis nula H_0 es rechazada.
2. Si $D \leq D_\alpha$, entonces la hipótesis nula H_0 no es rechazada.

El valor crítico de la estadística de prueba es:

$$D_\alpha = \frac{c_{1-\alpha}}{k(n)} = \begin{cases} \dfrac{1.224}{\sqrt{n}+0.12+\frac{0.11}{\sqrt{n}}} & \text{si } \alpha = 0.10 \\ \dfrac{1.358}{\sqrt{n}+0.12+\frac{0.11}{\sqrt{n}}} & \text{si } \alpha = 0.05 \\ \dfrac{1.629}{\sqrt{n}+0.12+\frac{0.11}{\sqrt{n}}} & \text{si } \alpha = 0.01 \end{cases}, \tag{8.21}$$

donde D_α es el valor crítico, α es el nivel de significación, $c_{1-\alpha} = \{1.224, 1.358, 1.629\}$ es una función del nivel de confianza y $k(n) = \sqrt{n} + 0.12 + \frac{0.11}{\sqrt{n}}$ es una función del tamaño de la muestra n.

Ejemplo 105 (prueba de bondad de ajuste). Las estaturas (en centímetros) de los Vaqueros de Dallas[5] son mostradas en la tabla 8.6.

Tabla 8.6: Estaturas en centímetros de los Vaqueros de Dallas 2019.

193.04	172.72	182.88	193.04	180.34	187.96	190.50	177.80
187.96	193.04	185.42	187.96	193.04	182.88	195.58	182.88
193.04	185.42	190.50	190.50	195.58	185.42	190.50	187.96
193.04	195.58	195.58	182.88	185.42	182.88	193.04	195.58
190.50	187.96	182.88	177.80	190.50	182.88	180.34	193.04
195.58	185.42	182.88	187.96	193.04	198.12	193.04	190.50
195.58	185.42	187.96	195.58	193.04	185.42	187.96	193.04
193.04	195.58	195.58	182.88	187.96	198.12	185.42	180.34

Indicar si las estaturas presentan una distribución $\mathrm{N}(\mu, \sigma^2)$ con un nivel de confianza de 95 %.

[5] www.nfl.com/teams/roster?team=DAL consultado el 11 de noviembre de 2019.

Solución. El modelo de decisión es:

1. La hipótesis nula es $H_0 : E \sim \mathrm{N}\left(\mu, \sigma^2\right)$.
2. La hipótesis alternativa es $H_1 : E \nsim \mathrm{N}\left(\mu, \sigma^2\right)$.
3. El nivel de significación es $\alpha = 0.05$.
4. El tamaño de la muestra es $n = 64$.
5. El estadístico de prueba es $D = \max\limits_{1 \leq k \leq n} \left(\left|\hat{F}_X\left(X_k\right) - F_X\left(X_k\right)\right|\right)$.
6. El punto crítico es $D_\alpha = 0.1670$.
7. La regla de decisión es: si $D < 0.1670$, entonces H_0 no es rechazada.
8. Sustituyendo en la ecuación (8.20): $D = \max\limits_{1 \leq k \leq n} \left(\left|\hat{F}_X\left(X_k\right) - F_X\left(X_k\right)\right|\right) = 0.1561$ donde $0.1561 < 0.1670$, entonces $D \notin \mathscr{C}$, es decir, D está en la región de no rechazo $F \in \mathscr{C}^{\mathrm{c}}$.
9. La hipótesis nula $H_0 : E \sim \mathrm{N}\left(\mu, \sigma^2\right)$ no es rechazada.
10. Por lo tanto, las estaturas presentan una distribución $\mathrm{N}\left(\mu, \sigma^2\right)$ con un nivel de confianza de 95 %.

El cálculo[6] del estadístico D es presentado en la tabla 8.7.

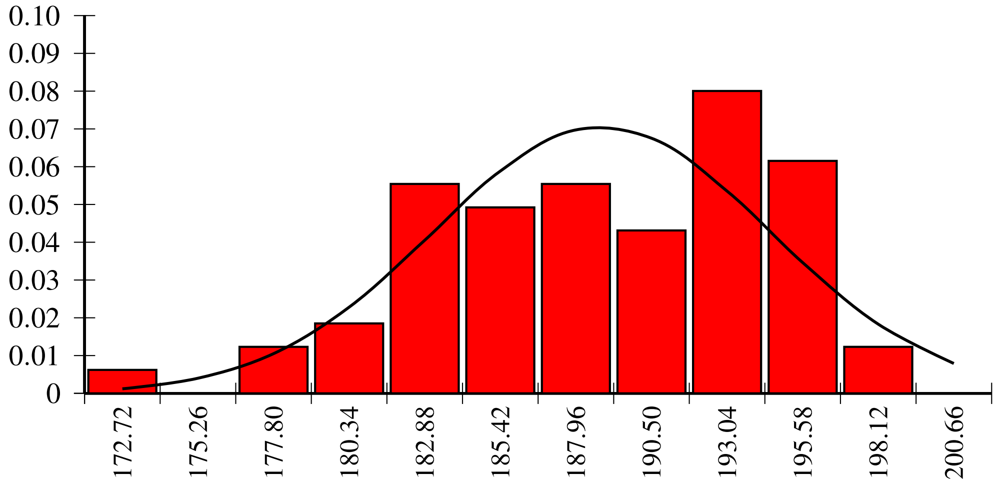
Tabla 8.7: Cálculo del estadístico D.

Estatura	f	F	fr	Fr	$\hat{F}_E(E)$	$F_E(E)$	D
172.720	1	1	0.0156	0.0156	0.0154	0.0022	0.0132
175.260	0	1	0.0000	0.0156	0.0308	0.0082	0.0226
177.800	2	3	0.0313	0.0469	0.0308	0.0256	0.0052
180.340	3	6	0.0469	0.0938	0.0615	0.0666	0.0051
182.880	9	15	0.1406	0.2344	0.1077	0.1463	0.0386
185.420	8	23	0.1250	0.3594	0.2462	0.2731	0.0270
187.960	9	32	0.1406	0.5000	0.3692	0.4387	0.0694
190.500	7	39	0.1094	0.6094	0.5077	0.6159	0.1082
193.040	13	52	0.2031	0.8125	0.6154	0.7715	0.1561
195.580	10	62	0.1563	0.9688	0.8154	0.8835	0.0681
198.120	2	64	0.0313	1.0000	0.9692	0.9497	0.0195

188.8331 31.9956 5.6565

La tabla 8.7 indica que el valor máximo de la columna D es 0.1561, por lo tanto la estatura de los Vaqueros presenta una distribución $\mathrm{N}\left(\mu, \sigma^2\right)$ donde $\hat{\mu} = 188.8331$ y $\hat{\sigma}^2 = 31.9956$ con un nivel de confianza de 95 %. El nivel de significación descriptivo es $\mathscr{P}\left(\underline{X}\right) = 0.0616$, entonces la hipótesis nula de la prueba de bondad de ajuste no es rechazada con un nivel de significación máximo de 6.16 %. El ajuste de la densidad Gauss a las estaturas de los Vaqueros es presentado en la gráfica 8.11.

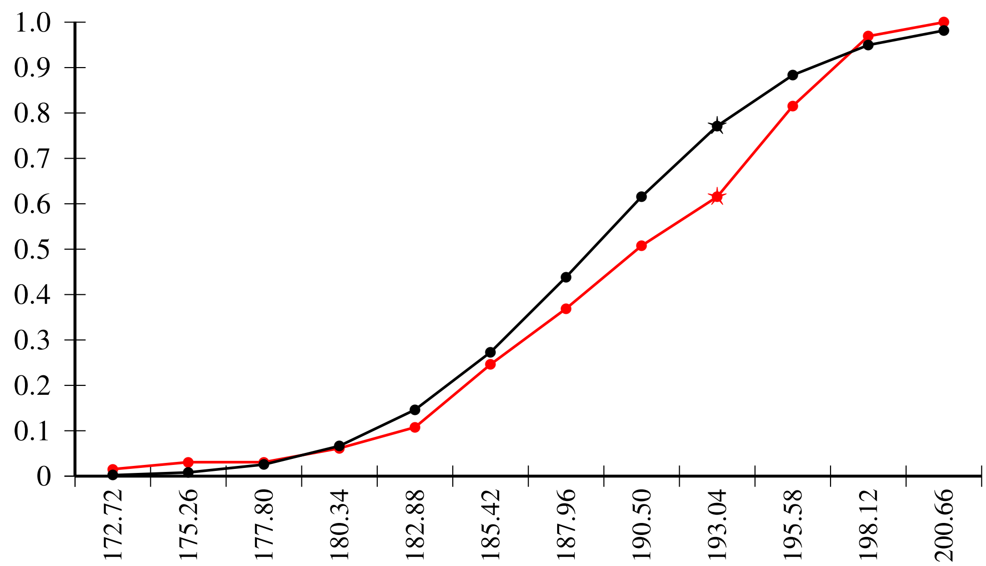
[6]El cálculo de la columna $\hat{F}_E(E)$ es $\hat{F}_E(E_k) = \frac{F_{k-1}+1}{n+1}$ y el cálculo de la columna $F_E(E)$, donde $Z_k = \frac{E-\bar{E}}{S_E}$ con $\bar{E} = 188.8331$ y $S_E = 5.6565$, entonces los valores para $F_Z(Z_k) = \Phi(Z_k)$ son obtenidos de la tabla A.1.



Gráfica 8.11: Ajuste de la densidad Gauss de las estaturas de los Vaqueros.

El histograma de frecuencias relativas de las estaturas es calculado como la densidad observada $\hat{f}_E(E) = \frac{f}{n\,\mathrm{d}k}$ donde $\mathrm{d}k = 2.54$ y la densidad teórica es calculada como $f_E(E) = \phi(E) = \frac{\exp\left(-\frac{(E-\bar{E})^2}{2S_E^2}\right)}{S_E\sqrt{2\pi}}$.

El ajuste de la distribución Gauss a las estaturas de los Vaqueros es presentado en la gráfica 8.12.



Gráfica 8.12: Ajuste de la distribución Gauss a las estaturas de los Vaqueros.

La diferencia máxima en el ajuste de la distribución Gauss a las estaturas de los Vaqueros es 0.1561, por lo tanto las estaturas presentan una distribución $\mathrm{N}\left(188.8331, 5.6565^2\right)$ con un nivel de confianza de 95 %. $\square$

La prueba de bondad de ajuste Kolmogorov y Smirnov es posible de realizar con los datos agrupados en intervalos de clase con los métodos de Velleman, Sturges, Scott o Freedman y Diaconis donde $Z_j = \frac{\bar{C}_j - \bar{E}}{S_E}$.

8.7. Ejercicios

Ejercicio 86. Una báscula es calibrada 60 veces con una pesa de 1,000 gramos. Las 60 lecturas tienen una media $\bar{X} = 1,000.65$ y una desviación estándar $S_X = 3$. Calcular el nivel de significación descriptivo $\mathscr{P}\left(\underline{X}\right)$ y realizar la prueba de la hipótesis nula $H_0\colon\ \mu = 1,000$ y la hipótesis alternativa $H_1\colon\ \mu \neq 1,000$.

Ejercicio 87. Una muestra de seis separadores de anillos, cuyos espesores (en milímetros) proviene de una población con una distribución Gauss y son:

39.031	38.998	39.012	39.018	39.009	39.005

Calcular el nivel de significación descriptivo $\mathscr{P}\left(\underline{X}\right)$ y realizar la prueba de la hipótesis nula $H_0\colon\ \mu = 39$ y la hipótesis alternativa $H_1\colon\ \mu \neq 39$.

Ejercicio 88. Un experimento para determinar la tasa de absorción que presenta una sustancia en la piel aplicó cantidades medidas de dos sustancias, después se midieron las cantidades absorbidas (en miligramos). La sustancia A presentó una varianza $S_A^2 = 2.4$ en siete aplicaciones y la sustancia B presentó una varianza $S_B^2 = 0.5$ en nueve aplicaciones. Realizar la prueba de hipótesis con la hipótesis nula $H_0\colon\ \mu = 39$ y la hipótesis alternativa $H_1\colon\ \mu \neq 39$ con un nivel de significación $\alpha = 0.05$.

Ejercicio 89. Los diámetros (en centímetros) de los ejes de una muestra aleatoria de una distribución Gauss son presentados en la tabla siguiente:

4.996	4.998	5.002	4.999	5.010
4.997	5.003	4.998	5.006	5.004
5.000	4.993	5.002	4.996	5.005
4.992	5.007	5.003	5.000	5.000

1. Realizar la prueba de bondad de ajuste Kolmogorov y Smirnov para la distribución Gauss con un nivel de significación $\alpha = 0.05$.
2. Realizar la prueba de la hipótesis nula $H_0\colon\ \sigma \leq 0.005$ y la hipótesis alternativa $H_1\colon\ \sigma > 0.005$ con un nivel de significación $\alpha = 0.05$.

Ejercicio 90. Las estaturas de una muestra aleatoria de 22 estudiantes son presentadas en la tabla siguiente:

172	167	170	170	170	168	154	155	173	168	168
160	170	190	169	174	154	179	165	174	165	166

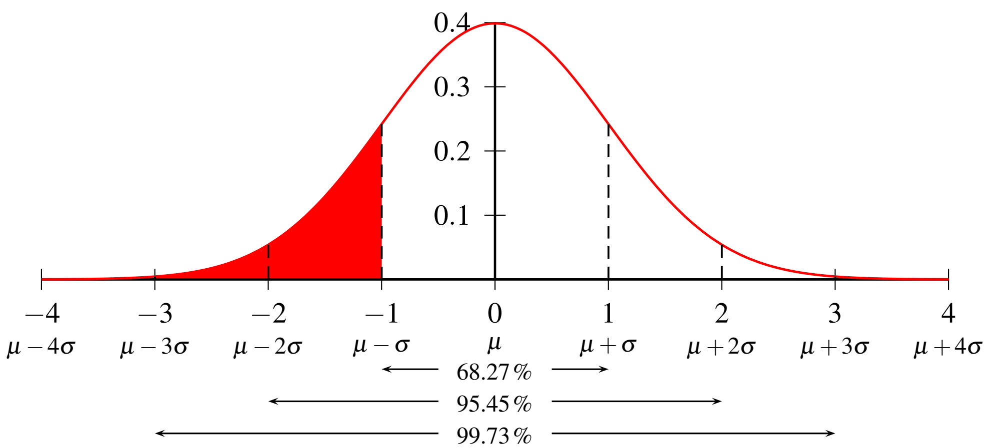
Calcular:

1. Los estadísticos descriptivos $\text{mín}\left(X\right), \text{máx}\left(X\right), \bar{X}, S_X^2, S_X, g_1^*, g_2^*$ agrupando los datos con el método de Sturges (1926).
2. Calcular el nivel de significación descriptivo $\mathscr{P}\left(\underline{X}\right)$ para la prueba γ con la hipótesis nula $H_0\colon\ \mu = 169$ y la hipótesis alternativa $H_1\colon\ \mu < 162$.
3. Realizar la prueba de bondad de ajuste Kolmogorov y Smirnov para la distribución Gauss con un nivel de significación $\alpha = 0.05$.

Apéndice A

Distribución Gauss estándar $\mathrm{N}(0,1)$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u \qquad -\infty < z < \infty.$$



Gráfica A.1: Distribución Gauss estándar $\mathrm{N}(0,1)$.

La gráfica A.1 presenta la distribución Gauss estándar en $z = -1$ y en la tabla A.1 se observa que el valor para $z = -1$ es $\Phi(-1) = 0.1587$.

La distribución Gauss estándar tiene la propiedad siguiente:

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = \frac{1}{\sqrt{2\pi}} \int_{-z}^{\infty} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \Phi(-z).$$

Por ejemplo:

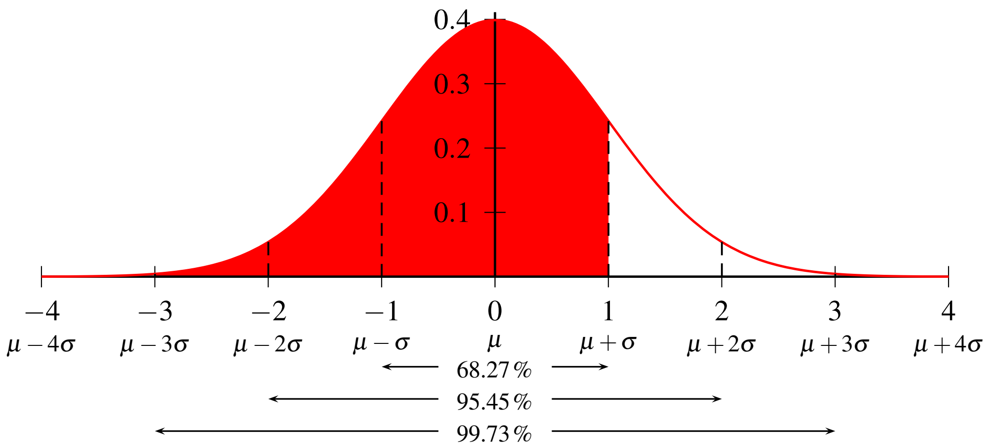
$$\Phi(-1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-1} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \frac{1}{\sqrt{2\pi}} \int_{1}^{\infty} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \Phi(1) = 0.1587.$$

La distribución Gauss estándar tiene la propiedad siguiente:

$$\Phi(-z) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \frac{1}{\sqrt{2\pi}} \int_{-z}^{\infty} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \Phi(z).$$

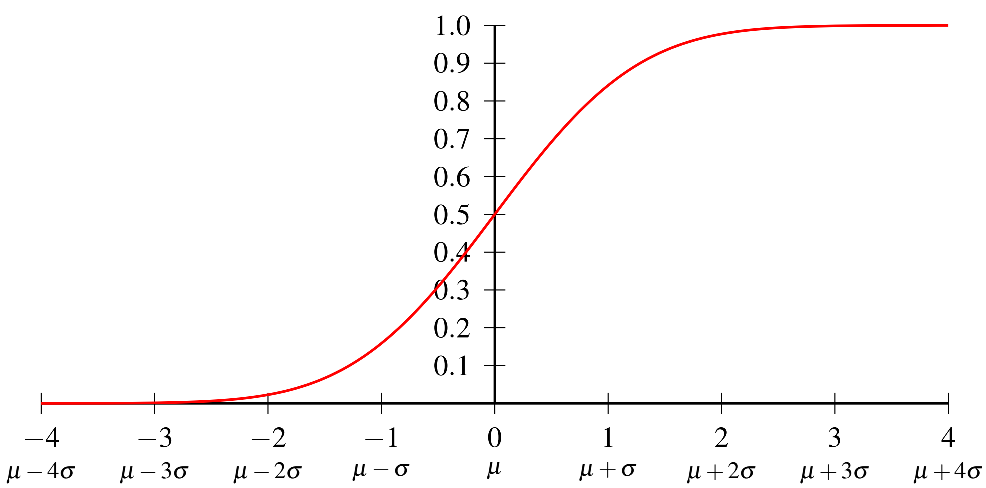
Por ejemplo:

$$\Phi(1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{1} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \frac{1}{\sqrt{2\pi}} \int_{-1}^{\infty} \exp\left(-\frac{u^2}{2}\right) \mathrm{d}u = 1 - \Phi(-1) = 0.8413.$$



Gráfica A.2: Distribución Gauss estándar $F_z(z)$.

La gráfica A.3 presenta la distribución Gauss estándar en $z = 1$ y en la tabla A.1 se observa que el valor para $z = -1$ es $\Phi(-1) = 0.8413$.



Gráfica A.3: Función de distribución Gauss estándar $F_Z(z)$.

La gráfica A.3 muestra la simetría de la distribución Gauss estándar.

Tabla A.1: Distribución Gauss estándar $N(0,1)$.

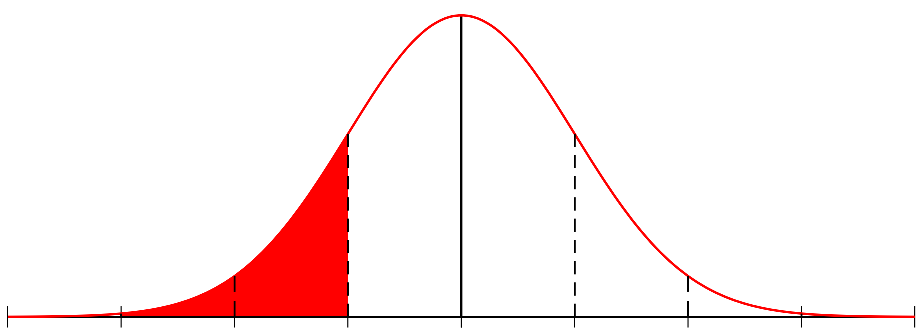
z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
-4.0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
-3.8	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.7	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.6	0.0002	0.0002	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
-3.5	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002	0.0002
-3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
-3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
-3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
-3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
-3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Tabla A.1: Distribución Gauss estándar.

z	**0.00**	**0.01**	**0.02**	**0.03**	**0.04**	**0.05**	**0.06**	**0.07**	**0.08**	**0.09**
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4.0	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Apéndice B

Percentiles de la distribución $t\,(\nu)$



Gráfica B.1: Distribución $t\,(\nu)$.

La gráfica B.1 presenta la distribución $t\,(\nu)$ en $t_\alpha\,(\nu)$, donde α es el α–ésimo percentil del extremo izquierdo.

La distribución $t\,(\nu)$ tiene la propiedad siguiente:

$$t_\alpha\,(\nu) = -t_{1-\alpha}\,(\nu)\,.$$

Por ejemplo:

$$t_{0.025}\,(29) = -t_{0.975}\,(29) = -2.0452.$$

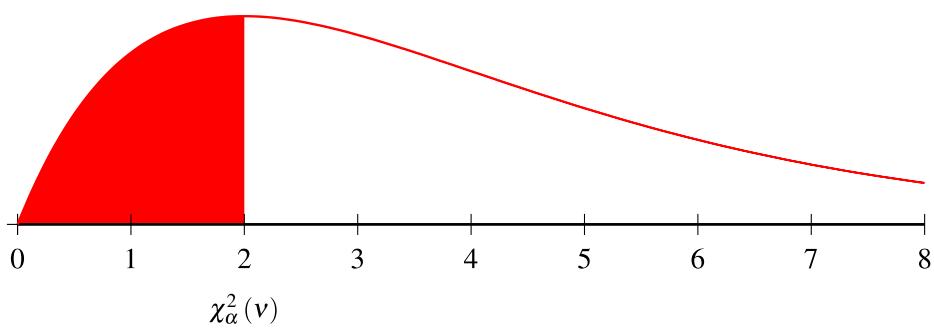
Por lo tanto, el área bajo la curva de la distribución $t\,(29)$ entre los puntos críticos -2.0452 y 2.0452 representa el 95 % de probabilidad.

Tabla B.1: Distribución $t\,(\nu)$.

ν	$t_{0.005}\,(\nu)$	$t_{0.010}\,(\nu)$	$t_{0.025}\,(\nu)$	$t_{0.050}\,(\nu)$	$t_{0.100}\,(\nu)$	$t_{0.200}\,(\nu)$	$t_{0.300}\,(\nu)$	$t_{0.400}\,(\nu)$
1	-63.6567	-31.8205	-12.7062	-6.3138	-3.0777	-1.3764	-0.7265	-0.3249
2	-9.9248	-6.9646	-4.3027	-2.9200	-1.8856	-1.0607	-0.6172	-0.2887
3	-5.8409	-4.5407	-3.1824	-2.3534	-1.6377	-0.9785	-0.5844	-0.2767
4	-4.6041	-3.7469	-2.7764	-2.1318	-1.5332	-0.9410	-0.5686	-0.2707
5	-4.0321	-3.3649	-2.5706	-2.0150	-1.4759	-0.9195	-0.5594	-0.2672
6	-3.7074	-3.1427	-2.4469	-1.9432	-1.4398	-0.9057	-0.5534	-0.2648
7	-3.4995	-2.9980	-2.3646	-1.8946	-1.4149	-0.8960	-0.5491	-0.2632
8	-3.3554	-2.8965	-2.3060	-1.8595	-1.3968	-0.8889	-0.5459	-0.2619
9	-3.2498	-2.8214	-2.2622	-1.8331	-1.3830	-0.8834	-0.5435	-0.2610
10	-3.1693	-2.7638	-2.2281	-1.8125	-1.3722	-0.8791	-0.5415	-0.2602
11	-3.1058	-2.7181	-2.2010	-1.7959	-1.3634	-0.8755	-0.5399	-0.2596
12	-3.0545	-2.6810	-2.1788	-1.7823	-1.3562	-0.8726	-0.5386	-0.2590
13	-3.0123	-2.6503	-2.1604	-1.7709	-1.3502	-0.8702	-0.5375	-0.2586
14	-2.9768	-2.6245	-2.1448	-1.7613	-1.3450	-0.8681	-0.5366	-0.2582
15	-2.9467	-2.6025	-2.1314	-1.7531	-1.3406	-0.8662	-0.5357	-0.2579
16	-2.9208	-2.5835	-2.1199	-1.7459	-1.3368	-0.8647	-0.5350	-0.2576
17	-2.8982	-2.5669	-2.1098	-1.7396	-1.3334	-0.8633	-0.5344	-0.2573
18	-2.8784	-2.5524	-2.1009	-1.7341	-1.3304	-0.8620	-0.5338	-0.2571
19	-2.8609	-2.5395	-2.0930	-1.7291	-1.3277	-0.8610	-0.5333	-0.2569
20	-2.8453	-2.5280	-2.0860	-1.7247	-1.3253	-0.8600	-0.5329	-0.2567
21	-2.8314	-2.5176	-2.0796	-1.7207	-1.3232	-0.8591	-0.5325	-0.2566
22	-2.8188	-2.5083	-2.0739	-1.7171	-1.3212	-0.8583	-0.5321	-0.2564
23	-2.8073	-2.4999	-2.0687	-1.7139	-1.3195	-0.8575	-0.5317	-0.2563
24	-2.7969	-2.4922	-2.0639	-1.7109	-1.3178	-0.8569	-0.5314	-0.2562
25	-2.7874	-2.4851	-2.0595	-1.7081	-1.3163	-0.8562	-0.5312	-0.2561
26	-2.7787	-2.4786	-2.0555	-1.7056	-1.3150	-0.8557	-0.5309	-0.2560
27	-2.7707	-2.4727	-2.0518	-1.7033	-1.3137	-0.8551	-0.5306	-0.2559
28	-2.7633	-2.4671	-2.0484	-1.7011	-1.3125	-0.8546	-0.5304	-0.2558
29	-2.7564	-2.4620	-2.0452	-1.6991	-1.3114	-0.8542	-0.5302	-0.2557
30	-2.7500	-2.4573	-2.0423	-1.6973	-1.3104	-0.8538	-0.5300	-0.2556
40	-2.7045	-2.4233	-2.0211	-1.6839	-1.3031	-0.8507	-0.5286	-0.2550
50	-2.6778	-2.4033	-2.0086	-1.6759	-1.2987	-0.8489	-0.5278	-0.2547
60	-2.6603	-2.3901	-2.0003	-1.6706	-1.2958	-0.8477	-0.5272	-0.2545
70	-2.6479	-2.3808	-1.9944	-1.6669	-1.2938	-0.8468	-0.5268	-0.2543
80	-2.6387	-2.3739	-1.9901	-1.6641	-1.2922	-0.8461	-0.5265	-0.2542
90	-2.6316	-2.3685	-1.9867	-1.6620	-1.2910	-0.8456	-0.5263	-0.2541
100	-2.6259	-2.3642	-1.9840	-1.6602	-1.2901	-0.8452	-0.5261	-0.2540
200	-2.6006	-2.3451	-1.9719	-1.6525	-1.2858	-0.8434	-0.5252	-0.2537
300	-2.5923	-2.3388	-1.9679	-1.6499	-1.2844	-0.8428	-0.5250	-0.2536
400	-2.5882	-2.3357	-1.9659	-1.6487	-1.2837	-0.8425	-0.5248	-0.2535
500	-2.5857	-2.3338	-1.9647	-1.6479	-1.2832	-0.8423	-0.5247	-0.2535
1000	-2.5808	-2.3301	-1.9623	-1.6464	-1.2824	-0.8420	-0.5246	-0.2534
∞	-2.5763	-2.3267	-1.9602	-1.6450	-1.2816	-0.8417	-0.5244	-0.2534

Apéndice C

Percentiles de la distribución $\chi^2(\nu)$



Gráfica C.1: Distribución $\chi^2(\nu)$.

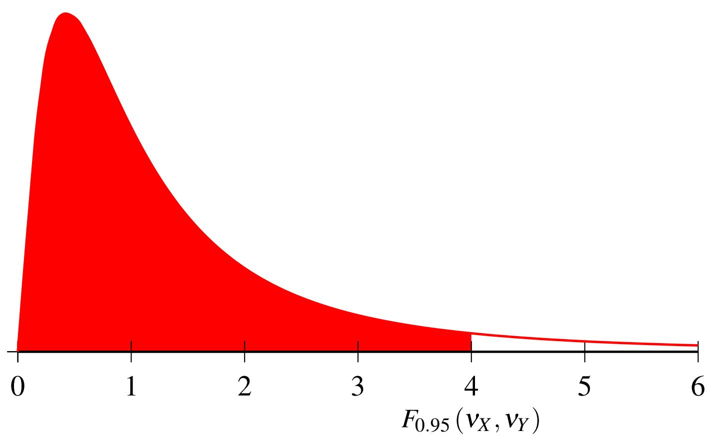
La gráfica C.1 presenta la distribución $\chi^2(\nu)$ en $\chi^2_\alpha(\nu)$, donde α es el α–ésimo percentil del extremo izquierdo.

Tabla C.1: Distribución $\chi^2(\nu)$.

ν	$\chi^2_{0.005}(\nu)$	$\chi^2_{0.010}(\nu)$	$\chi^2_{0.025}(\nu)$	$\chi^2_{0.050}(\nu)$	$\chi^2_{0.100}(\nu)$	$\chi^2_{0.900}(\nu)$	$\chi^2_{0.950}(\nu)$	$\chi^2_{0.975}(\nu)$	$\chi^2_{0.990}(\nu)$	$\chi^2_{0.995}(\nu)$
1	0.0000	0.0002	0.0010	0.0039	0.0158	2.7055	3.8415	5.0239	6.6349	7.8794
2	0.0100	0.0201	0.0506	0.1026	0.2107	4.6052	5.9915	7.3778	9.2103	10.5966
3	0.0717	0.1148	0.2158	0.3518	0.5844	6.2514	7.8147	9.3484	11.3449	12.8382
4	0.2070	0.2971	0.4844	0.7107	1.0636	7.7794	9.4877	11.1433	13.2767	14.8603
5	0.4117	0.5543	0.8312	1.1455	1.6103	9.2364	11.0705	12.8325	15.0863	16.7496
6	0.6757	0.8721	1.2373	1.6354	2.2041	10.6446	12.5916	14.4494	16.8119	18.5476
7	0.9893	1.2390	1.6899	2.1673	2.8331	12.0170	14.0671	16.0128	18.4753	20.2777
8	1.3444	1.6465	2.1797	2.7326	3.4895	13.3616	15.5073	17.5345	20.0902	21.9550
9	1.7349	2.0879	2.7004	3.3251	4.1682	14.6837	16.9190	19.0228	21.6660	23.5894
10	2.1559	2.5582	3.2470	3.9403	4.8652	15.9872	18.3070	20.4832	23.2093	25.1882
11	2.6032	3.0535	3.8157	4.5748	5.5778	17.2750	19.6751	21.9200	24.7250	26.7568
12	3.0738	3.5706	4.4038	5.2260	6.3038	18.5493	21.0261	23.3367	26.2170	28.2995
13	3.5650	4.1069	5.0088	5.8919	7.0415	19.8119	22.3620	24.7356	27.6882	29.8195
14	4.0747	4.6604	5.6287	6.5706	7.7895	21.0641	23.6848	26.1189	29.1412	31.3193
15	4.6009	5.2293	6.2621	7.2609	8.5468	22.3071	24.9958	27.4884	30.5779	32.8013
16	5.1422	5.8122	6.9077	7.9616	9.3122	23.5418	26.2962	28.8454	31.9999	34.2672
17	5.6972	6.4078	7.5642	8.6718	10.0852	24.7690	27.5871	30.1910	33.4087	35.7185
18	6.2648	7.0149	8.2307	9.3905	10.8649	25.9894	28.8693	31.5264	34.8053	37.1565
19	6.8440	7.6327	8.9065	10.1170	11.6509	27.2036	30.1435	32.8523	36.1909	38.5823
20	7.4338	8.2604	9.5908	10.8508	12.4426	28.4120	31.4104	34.1696	37.5662	39.9968
21	8.0337	8.8972	10.2829	11.5913	13.2396	29.6151	32.6706	35.4789	38.9322	41.4011
22	8.6427	9.5425	10.9823	12.3380	14.0415	30.8133	33.9244	36.7807	40.2894	42.7957
23	9.2604	10.1957	11.6886	13.0905	14.8480	32.0069	35.1725	38.0756	41.6384	44.1813
24	9.8862	10.8564	12.4012	13.8484	15.6587	33.1962	36.4150	39.3641	42.9798	45.5585
25	10.5197	11.5240	13.1197	14.6114	16.4734	34.3816	37.6525	40.6465	44.3141	46.9279
26	11.1602	12.1981	13.8439	15.3792	17.2919	35.5632	38.8851	41.9232	45.6417	48.2899
27	11.8076	12.8785	14.5734	16.1514	18.1139	36.7412	40.1133	43.1945	46.9629	49.6449
28	12.4613	13.5647	15.3079	16.9279	18.9392	37.9159	41.3371	44.4608	48.2782	50.9934
29	13.1211	14.2565	16.0471	17.7084	19.7677	39.0875	42.5570	45.7223	49.5879	52.3356
30	13.7867	14.9535	16.7908	18.4927	20.5992	40.2560	43.7730	46.9792	50.8922	53.6720
40	20.7065	22.1643	24.4330	26.5093	29.0505	51.8051	55.7585	59.3417	63.6907	66.7660
50	27.9907	29.7067	32.3574	34.7643	37.6886	63.1671	67.5048	71.4202	76.1539	79.4900
60	35.5345	37.4849	40.4817	43.1880	46.4589	74.3970	79.0819	83.2977	88.3794	91.9517
70	43.2752	45.4417	48.7576	51.7393	55.3289	85.5270	90.5312	95.0232	100.4252	104.2149
80	51.1719	53.5401	57.1532	60.3915	64.2778	96.5782	101.8795	106.6286	112.3288	116.3211
90	59.1963	61.7541	65.6466	69.1260	73.2911	107.5650	113.1453	118.1359	124.1163	128.2989
100	67.3276	70.0649	74.2219	77.9295	82.3581	118.4980	124.3421	129.5612	135.8067	140.1695
110	75.5500	78.4583	82.8671	86.7916	91.4710	129.3851	135.4802	140.9166	147.4143	151.9485
120	83.8516	86.9233	91.5726	95.7046	100.6236	140.2326	146.5674	152.2114	158.9502	163.6482
130	92.2225	95.4510	100.3313	104.6622	109.8110	151.0452	157.6099	163.4531	170.4231	175.2783
140	100.6548	104.0344	109.1369	113.6593	119.0293	161.8270	168.6130	174.6478	181.8403	186.8468
150	109.1422	112.6676	117.9845	122.6918	128.2751	172.5812	179.5806	185.8004	193.2077	198.3602
200	152.2410	156.4320	162.7280	168.2786	174.8353	226.0210	233.9943	241.0579	249.4451	255.2642
300	240.6634	245.9725	253.9123	260.8781	269.0679	331.7885	341.3951	349.8745	359.9064	366.8444
400	330.9028	337.1553	346.4818	354.6410	364.2074	436.6490	447.6325	457.3055	468.7245	476.6064
500	422.3034	429.3875	439.9360	449.1468	459.9261	540.9303	553.1268	563.8515	576.4928	585.2066

Apéndice D

Percentil 95 de la distribución $F(\nu_X, \nu_Y)$



Gráfica D.1: Distribución $F(\nu_X, \nu_Y)$.

La gráfica D.1 presenta la distribución $F(\nu_X, \nu_Y)$ en $F_\alpha(\nu_X, \nu_Y)$, donde α es el α–ésimo percentil del extremo izquierdo.

La distribución $F(\nu_X, \nu_Y)$ tiene la propiedad siguiente:

$$F_{1-\alpha}(\nu_X, \nu_Y) = \frac{1}{F_\alpha(\nu_Y, \nu_X)}.$$

Por ejemplo:

$$F_{0.95}(110, 55) = 1.49$$

Entonces:

$$F_{0.05}(110, 55) = \frac{1}{F_{0.95}(55, 110)} = \frac{1}{1.45} = 0.69.$$

Por lo tanto, el área bajo la curva de la distribución $F(110, 55)$ entre los puntos críticos 0.69 y 1.49 representa el 90 % de probabilidad.

Tabla D.1: Distribución $F_{0.95}(\nu_X, \nu_Y)$.

ν_Y \ ν_X	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
1	161.45	199.50	215.71	224.58	230.16	233.99	236.77	238.88	240.54	241.88
2	18.51	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.99	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95	2.90	2.85
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
13	4.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77	2.71	2.67
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70	2.65	2.60
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55	2.49	2.45
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48	2.42	2.38
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37	2.32	2.27
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
26	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22
27	4.21	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
29	4.18	3.33	2.93	2.70	2.55	2.43	2.35	2.28	2.22	2.18
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16
35	4.12	3.27	2.87	2.64	2.49	2.37	2.29	2.22	2.16	2.11
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08
45	4.06	3.20	2.81	2.58	2.42	2.31	2.22	2.15	2.10	2.05
50	4.03	3.18	2.79	2.56	2.40	2.29	2.20	2.13	2.07	2.03
55	4.02	3.16	2.77	2.54	2.38	2.27	2.18	2.11	2.06	2.01
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99
65	3.99	3.14	2.75	2.51	2.36	2.24	2.15	2.08	2.03	1.98
70	3.98	3.13	2.74	2.50	2.35	2.23	2.14	2.07	2.02	1.97
75	3.97	3.12	2.73	2.49	2.34	2.22	2.13	2.06	2.01	1.96
80	3.96	3.11	2.72	2.49	2.33	2.21	2.13	2.06	2.00	1.95
85	3.95	3.10	2.71	2.48	2.32	2.21	2.12	2.05	1.99	1.94
90	3.95	3.10	2.71	2.47	2.32	2.20	2.11	2.04	1.99	1.94
95	3.94	3.09	2.70	2.47	2.31	2.20	2.11	2.04	1.98	1.93
100	3.94	3.09	2.70	2.46	2.31	2.19	2.10	2.03	1.97	1.93
105	3.93	3.08	2.69	2.46	2.30	2.19	2.10	2.03	1.97	1.92
110	3.93	3.08	2.69	2.45	2.30	2.18	2.09	2.02	1.97	1.92
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83

Tabla D.1:Distribución $F_{0.95}(\nu_X, \nu_Y)$.
Continuación para los grados de libertad del numerador ν_X.

ν_Y \ ν_X	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
1	242.98	243.91	244.69	245.36	245.95	246.46	246.92	247.32	247.69	248.01
2	19.40	19.41	19.42	19.42	19.43	19.43	19.44	19.44	19.44	19.45
3	8.76	8.74	8.73	8.71	8.70	8.69	8.68	8.67	8.67	8.66
4	5.94	5.91	5.89	5.87	5.86	5.84	5.83	5.82	5.81	5.80
5	4.70	4.68	4.66	4.64	4.62	4.60	4.59	4.58	4.57	4.56
6	4.03	4.00	3.98	3.96	3.94	3.92	3.91	3.90	3.88	3.87
7	3.60	3.57	3.55	3.53	3.51	3.49	3.48	3.47	3.46	3.44
8	3.31	3.28	3.26	3.24	3.22	3.20	3.19	3.17	3.16	3.15
9	3.10	3.07	3.05	3.03	3.01	2.99	2.97	2.96	2.95	2.94
10	2.94	2.91	2.89	2.86	2.85	2.83	2.81	2.80	2.79	2.77
11	2.82	2.79	2.76	2.74	2.72	2.70	2.69	2.67	2.66	2.65
12	2.72	2.69	2.66	2.64	2.62	2.60	2.58	2.57	2.56	2.54
13	2.63	2.60	2.58	2.55	2.53	2.51	2.50	2.48	2.47	2.46
14	2.57	2.53	2.51	2.48	2.46	2.44	2.43	2.41	2.40	2.39
15	2.51	2.48	2.45	2.42	2.40	2.38	2.37	2.35	2.34	2.33
16	2.46	2.42	2.40	2.37	2.35	2.33	2.32	2.30	2.29	2.28
17	2.41	2.38	2.35	2.33	2.31	2.29	2.27	2.26	2.24	2.23
18	2.37	2.34	2.31	2.29	2.27	2.25	2.23	2.22	2.20	2.19
19	2.34	2.31	2.28	2.26	2.23	2.21	2.20	2.18	2.17	2.16
20	2.31	2.28	2.25	2.22	2.20	2.18	2.17	2.15	2.14	2.12
21	2.28	2.25	2.22	2.20	2.18	2.16	2.14	2.12	2.11	2.10
22	2.26	2.23	2.20	2.17	2.15	2.13	2.11	2.10	2.08	2.07
23	2.24	2.20	2.18	2.15	2.13	2.11	2.09	2.08	2.06	2.05
24	2.22	2.18	2.15	2.13	2.11	2.09	2.07	2.05	2.04	2.03
25	2.20	2.16	2.14	2.11	2.09	2.07	2.05	2.04	2.02	2.01
26	2.18	2.15	2.12	2.09	2.07	2.05	2.03	2.02	2.00	1.99
27	2.17	2.13	2.10	2.08	2.06	2.04	2.02	2.00	1.99	1.97
28	2.15	2.12	2.09	2.06	2.04	2.02	2.00	1.99	1.97	1.96
29	2.14	2.10	2.08	2.05	2.03	2.01	1.99	1.97	1.96	1.94
30	2.13	2.09	2.06	2.04	2.01	1.99	1.98	1.96	1.95	1.93
35	2.07	2.04	2.01	1.99	1.96	1.94	1.92	1.91	1.89	1.88
40	2.04	2.00	1.97	1.95	1.92	1.90	1.89	1.87	1.85	1.84
45	2.01	1.97	1.94	1.92	1.89	1.87	1.86	1.84	1.82	1.81
50	1.99	1.95	1.92	1.89	1.87	1.85	1.83	1.81	1.80	1.78
55	1.97	1.93	1.90	1.88	1.85	1.83	1.81	1.79	1.78	1.76
60	1.95	1.92	1.89	1.86	1.84	1.82	1.80	1.78	1.76	1.75
65	1.94	1.90	1.87	1.85	1.82	1.80	1.78	1.76	1.75	1.73
70	1.93	1.89	1.86	1.84	1.81	1.79	1.77	1.75	1.74	1.72
75	1.92	1.88	1.85	1.83	1.80	1.78	1.76	1.74	1.73	1.71
80	1.91	1.88	1.84	1.82	1.79	1.77	1.75	1.73	1.72	1.70
85	1.90	1.87	1.84	1.81	1.79	1.76	1.74	1.73	1.71	1.70
90	1.90	1.86	1.83	1.80	1.78	1.76	1.74	1.72	1.70	1.69
95	1.89	1.86	1.82	1.80	1.77	1.75	1.73	1.71	1.70	1.68
100	1.89	1.85	1.82	1.79	1.77	1.75	1.73	1.71	1.69	1.68
105	1.88	1.85	1.81	1.79	1.76	1.74	1.72	1.70	1.69	1.67
110	1.88	1.84	1.81	1.78	1.76	1.74	1.72	1.70	1.68	1.67
∞	1.79	1.75	1.72	1.69	1.67	1.64	1.62	1.60	1.59	1.57

Tabla D.1:Distribución $F_{0.95}(\nu_X, \nu_Y)$.
Continuación para los grados de libertad del numerador ν_X.

ν_Y \ ν_X	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
1	248.31	248.58	248.83	249.05	249.26	249.45	249.63	249.80	249.95	250.10
2	19.45	19.45	19.45	19.45	19.46	19.46	19.46	19.46	19.46	19.46
3	8.65	8.65	8.64	8.64	8.63	8.63	8.63	8.62	8.62	8.62
4	5.79	5.79	5.78	5.77	5.77	5.76	5.76	5.75	5.75	5.75
5	4.55	4.54	4.53	4.53	4.52	4.52	4.51	4.50	4.50	4.50
6	3.86	3.86	3.85	3.84	3.83	3.83	3.82	3.82	3.81	3.81
7	3.43	3.43	3.42	3.41	3.40	3.40	3.39	3.39	3.38	3.38
8	3.14	3.13	3.12	3.12	3.11	3.10	3.10	3.09	3.08	3.08
9	2.93	2.92	2.91	2.90	2.89	2.89	2.88	2.87	2.87	2.86
10	2.76	2.75	2.75	2.74	2.73	2.72	2.72	2.71	2.70	2.70
11	2.64	2.63	2.62	2.61	2.60	2.59	2.59	2.58	2.58	2.57
12	2.53	2.52	2.51	2.51	2.50	2.49	2.48	2.48	2.47	2.47
13	2.45	2.44	2.43	2.42	2.41	2.41	2.40	2.39	2.39	2.38
14	2.38	2.37	2.36	2.35	2.34	2.33	2.33	2.32	2.31	2.31
15	2.32	2.31	2.30	2.29	2.28	2.27	2.27	2.26	2.25	2.25
16	2.26	2.25	2.24	2.24	2.23	2.22	2.21	2.21	2.20	2.19
17	2.22	2.21	2.20	2.19	2.18	2.17	2.17	2.16	2.15	2.15
18	2.18	2.17	2.16	2.15	2.14	2.13	2.13	2.12	2.11	2.11
19	2.14	2.13	2.12	2.11	2.11	2.10	2.09	2.08	2.08	2.07
20	2.11	2.10	2.09	2.08	2.07	2.07	2.06	2.05	2.05	2.04
21	2.08	2.07	2.06	2.05	2.05	2.04	2.03	2.02	2.02	2.01
22	2.06	2.05	2.04	2.03	2.02	2.01	2.00	2.00	1.99	1.98
23	2.04	2.02	2.01	2.01	2.00	1.99	1.98	1.97	1.97	1.96
24	2.01	2.00	1.99	1.98	1.97	1.97	1.96	1.95	1.95	1.94
25	2.00	1.98	1.97	1.96	1.96	1.95	1.94	1.93	1.93	1.92
26	1.98	1.97	1.96	1.95	1.94	1.93	1.92	1.91	1.91	1.90
27	1.96	1.95	1.94	1.93	1.92	1.91	1.90	1.90	1.89	1.88
28	1.95	1.93	1.92	1.91	1.91	1.90	1.89	1.88	1.88	1.87
29	1.93	1.92	1.91	1.90	1.89	1.88	1.88	1.87	1.86	1.85
30	1.92	1.91	1.90	1.89	1.88	1.87	1.86	1.85	1.85	1.84
35	1.87	1.85	1.84	1.83	1.82	1.82	1.81	1.80	1.79	1.79
40	1.83	1.81	1.80	1.79	1.78	1.77	1.77	1.76	1.75	1.74
45	1.80	1.78	1.77	1.76	1.75	1.74	1.73	1.73	1.72	1.71
50	1.77	1.76	1.75	1.74	1.73	1.72	1.71	1.70	1.69	1.69
55	1.75	1.74	1.73	1.72	1.71	1.70	1.69	1.68	1.67	1.67
60	1.73	1.72	1.71	1.70	1.69	1.68	1.67	1.66	1.66	1.65
65	1.72	1.71	1.70	1.69	1.68	1.67	1.66	1.65	1.64	1.63
70	1.71	1.70	1.68	1.67	1.66	1.65	1.65	1.64	1.63	1.62
75	1.70	1.69	1.67	1.66	1.65	1.64	1.63	1.63	1.62	1.61
80	1.69	1.68	1.67	1.65	1.64	1.63	1.63	1.62	1.61	1.60
85	1.68	1.67	1.66	1.65	1.64	1.63	1.62	1.61	1.60	1.59
90	1.67	1.66	1.65	1.64	1.63	1.62	1.61	1.60	1.59	1.59
95	1.67	1.66	1.64	1.63	1.62	1.61	1.60	1.59	1.59	1.58
100	1.66	1.65	1.64	1.63	1.62	1.61	1.60	1.59	1.58	1.57
105	1.66	1.64	1.63	1.62	1.61	1.60	1.59	1.58	1.58	1.57
110	1.65	1.64	1.63	1.62	1.61	1.60	1.59	1.58	1.57	1.56
∞	1.56	1.54	1.53	1.52	1.51	1.50	1.49	1.48	1.47	1.46

Tabla D.1:Distribución $F_{0.95}(\nu_X, \nu_Y)$.
Continuación para los grados de libertad del numerador ν_X.

ν_Y \ ν_X	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	40
1	250.23	250.36	250.48	250.59	250.69	250.79	250.89	250.98	251.06	251.14
2	19.46	19.46	19.47	19.47	19.47	19.47	19.47	19.47	19.47	19.47
3	8.61	8.61	8.61	8.61	8.60	8.60	8.60	8.60	8.60	8.59
4	5.74	5.74	5.74	5.73	5.73	5.73	5.72	5.72	5.72	5.72
5	4.49	4.49	4.48	4.48	4.48	4.47	4.47	4.47	4.47	4.46
6	3.80	3.80	3.80	3.79	3.79	3.79	3.78	3.78	3.78	3.77
7	3.37	3.37	3.36	3.36	3.36	3.35	3.35	3.35	3.34	3.34
8	3.07	3.07	3.07	3.06	3.06	3.06	3.05	3.05	3.05	3.04
9	2.86	2.85	2.85	2.85	2.84	2.84	2.84	2.83	2.83	2.83
10	2.69	2.69	2.69	2.68	2.68	2.67	2.67	2.67	2.66	2.66
11	2.57	2.56	2.56	2.55	2.55	2.54	2.54	2.54	2.53	2.53
12	2.46	2.46	2.45	2.45	2.44	2.44	2.44	2.43	2.43	2.43
13	2.38	2.37	2.37	2.36	2.36	2.35	2.35	2.35	2.34	2.34
14	2.30	2.30	2.29	2.29	2.28	2.28	2.28	2.27	2.27	2.27
15	2.24	2.24	2.23	2.23	2.22	2.22	2.21	2.21	2.21	2.20
16	2.19	2.18	2.18	2.17	2.17	2.17	2.16	2.16	2.15	2.15
17	2.14	2.14	2.13	2.13	2.12	2.12	2.11	2.11	2.11	2.10
18	2.10	2.10	2.09	2.09	2.08	2.08	2.07	2.07	2.07	2.06
19	2.07	2.06	2.06	2.05	2.05	2.04	2.04	2.03	2.03	2.03
20	2.03	2.03	2.02	2.02	2.01	2.01	2.01	2.00	2.00	1.99
21	2.00	2.00	1.99	1.99	1.98	1.98	1.98	1.97	1.97	1.96
22	1.98	1.97	1.97	1.96	1.96	1.95	1.95	1.95	1.94	1.94
23	1.95	1.95	1.94	1.94	1.93	1.93	1.93	1.92	1.92	1.91
24	1.93	1.93	1.92	1.92	1.91	1.91	1.90	1.90	1.90	1.89
25	1.91	1.91	1.90	1.90	1.89	1.89	1.88	1.88	1.88	1.87
26	1.89	1.89	1.88	1.88	1.87	1.87	1.87	1.86	1.86	1.85
27	1.88	1.87	1.87	1.86	1.86	1.85	1.85	1.84	1.84	1.84
28	1.86	1.86	1.85	1.85	1.84	1.84	1.83	1.83	1.82	1.82
29	1.85	1.84	1.84	1.83	1.83	1.82	1.82	1.81	1.81	1.81
30	1.83	1.83	1.82	1.82	1.81	1.81	1.80	1.80	1.80	1.79
35	1.78	1.77	1.77	1.76	1.76	1.75	1.75	1.74	1.74	1.74
40	1.74	1.73	1.73	1.72	1.72	1.71	1.71	1.70	1.70	1.69
45	1.71	1.70	1.69	1.69	1.68	1.68	1.67	1.67	1.66	1.66
50	1.68	1.67	1.67	1.66	1.66	1.65	1.65	1.64	1.64	1.63
55	1.66	1.65	1.65	1.64	1.64	1.63	1.63	1.62	1.62	1.61
60	1.64	1.64	1.63	1.62	1.62	1.61	1.61	1.60	1.60	1.59
65	1.63	1.62	1.61	1.61	1.60	1.60	1.59	1.59	1.58	1.58
70	1.62	1.61	1.60	1.60	1.59	1.59	1.58	1.58	1.57	1.57
75	1.60	1.60	1.59	1.59	1.58	1.57	1.57	1.56	1.56	1.55
80	1.59	1.59	1.58	1.58	1.57	1.56	1.56	1.55	1.55	1.54
85	1.59	1.58	1.57	1.57	1.56	1.56	1.55	1.55	1.54	1.54
90	1.58	1.57	1.57	1.56	1.55	1.55	1.54	1.54	1.53	1.53
95	1.57	1.57	1.56	1.55	1.55	1.54	1.54	1.53	1.53	1.52
100	1.57	1.56	1.55	1.55	1.54	1.54	1.53	1.52	1.52	1.52
105	1.56	1.55	1.55	1.54	1.54	1.53	1.52	1.52	1.51	1.51
110	1.56	1.55	1.54	1.54	1.53	1.52	1.52	1.51	1.51	1.50
∞	1.45	1.44	1.44	1.43	1.42	1.42	1.41	1.40	1.40	1.39

Tabla D.1:Distribución $F_{0.95}\left(\nu_X, \nu_Y\right)$.
Continuación para los grados de libertad del numerador $\nu_X = n_X - 1$.

ν_Y \ ν_X	**50**	**55**	**60**	**65**	**70**	**80**	**90**	**100**	**110**	**∞**
1	251.77	252.00	252.20	252.36	252.50	252.72	252.90	253.04	253.16	254.31
2	19.48	19.48	19.48	19.48	19.48	19.48	19.48	19.49	19.49	19.50
3	8.58	8.58	8.57	8.57	8.57	8.56	8.56	8.55	8.55	8.53
4	5.70	5.69	5.69	5.68	5.68	5.67	5.67	5.66	5.66	5.63
5	4.44	4.44	4.43	4.43	4.42	4.41	4.41	4.41	4.40	4.37
6	3.75	3.75	3.74	3.73	3.73	3.72	3.72	3.71	3.71	3.67
7	3.32	3.31	3.30	3.30	3.29	3.29	3.28	3.27	3.27	3.23
8	3.02	3.01	3.01	3.00	2.99	2.99	2.98	2.97	2.97	2.93
9	2.80	2.79	2.79	2.78	2.78	2.77	2.76	2.76	2.75	2.71
10	2.64	2.63	2.62	2.61	2.61	2.60	2.59	2.59	2.58	2.54
11	2.51	2.50	2.49	2.48	2.48	2.47	2.46	2.46	2.45	2.40
12	2.40	2.39	2.38	2.38	2.37	2.36	2.36	2.35	2.34	2.30
13	2.31	2.30	2.30	2.29	2.28	2.27	2.27	2.26	2.26	2.21
14	2.24	2.23	2.22	2.22	2.21	2.20	2.19	2.19	2.18	2.13
15	2.18	2.17	2.16	2.15	2.15	2.14	2.13	2.12	2.12	2.07
16	2.12	2.11	2.11	2.10	2.09	2.08	2.07	2.07	2.06	2.01
17	2.08	2.07	2.06	2.05	2.05	2.03	2.03	2.02	2.02	1.96
18	2.04	2.03	2.02	2.01	2.00	1.99	1.98	1.98	1.97	1.92
19	2.00	1.99	1.98	1.97	1.97	1.96	1.95	1.94	1.93	1.88
20	1.97	1.96	1.95	1.94	1.93	1.92	1.91	1.91	1.90	1.84
21	1.94	1.93	1.92	1.91	1.90	1.89	1.88	1.88	1.87	1.81
22	1.91	1.90	1.89	1.88	1.88	1.86	1.86	1.85	1.84	1.78
23	1.88	1.87	1.86	1.86	1.85	1.84	1.83	1.82	1.82	1.76
24	1.86	1.85	1.84	1.83	1.83	1.82	1.81	1.80	1.79	1.73
25	1.84	1.83	1.82	1.81	1.81	1.80	1.79	1.78	1.77	1.71
26	1.82	1.81	1.80	1.79	1.79	1.78	1.77	1.76	1.75	1.69
27	1.81	1.79	1.79	1.78	1.77	1.76	1.75	1.74	1.74	1.67
28	1.79	1.78	1.77	1.76	1.75	1.74	1.73	1.73	1.72	1.65
29	1.77	1.76	1.75	1.75	1.74	1.73	1.72	1.71	1.70	1.64
30	1.76	1.75	1.74	1.73	1.72	1.71	1.70	1.70	1.69	1.62
35	1.70	1.69	1.68	1.67	1.66	1.65	1.64	1.63	1.63	1.56
40	1.66	1.65	1.64	1.63	1.62	1.61	1.60	1.59	1.58	1.51
45	1.63	1.61	1.60	1.59	1.59	1.57	1.56	1.55	1.55	1.47
50	1.60	1.59	1.58	1.57	1.56	1.54	1.53	1.52	1.52	1.44
55	1.58	1.56	1.55	1.54	1.54	1.52	1.51	1.50	1.49	1.41
60	1.56	1.55	1.53	1.52	1.52	1.50	1.49	1.48	1.47	1.39
65	1.54	1.53	1.52	1.51	1.50	1.49	1.47	1.46	1.46	1.37
70	1.53	1.52	1.50	1.49	1.49	1.47	1.46	1.45	1.44	1.35
75	1.52	1.50	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.34
80	1.51	1.49	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.32
85	1.50	1.48	1.47	1.46	1.45	1.44	1.43	1.42	1.41	1.31
90	1.49	1.48	1.46	1.45	1.44	1.43	1.42	1.41	1.40	1.30
95	1.48	1.47	1.46	1.45	1.44	1.42	1.41	1.40	1.39	1.29
100	1.48	1.46	1.45	1.44	1.43	1.41	1.40	1.39	1.38	1.28
105	1.47	1.46	1.44	1.43	1.42	1.41	1.40	1.39	1.38	1.28
110	1.47	1.45	1.44	1.43	1.42	1.40	1.39	1.38	1.37	1.27
∞	1.35	1.33	1.32	1.30	1.29	1.27	1.26	1.24	1.23	1.00

Referencias bibliográficas

Anderson, T. W. y Darling, D. A. (1954). A test of goodness of fit, *Journal of the American Statistical Association*, vol. 49, núm. 268, pp. 765–769.

Arbuthnot, J. (1692). *Of the laws of chance*, 1^a ed. London, Taylor.

Bienaymé, I. J. (1838). Mémoire sur la probabilité des résultats moyens des observations; démonstration directe de la règle de Laplace, *Mémoires de l'Académie de Sciences de l'Institut de France*, vol. 5: pp. 513–558.

Bienaymé, I. J. (1852). Mémoire sur la probabilité des erreurs d'après la methode des moindres carres, *Journal de Mathématiques Pures et Appliquées*, vol. 1, núm. 17: pp. 33–78.

Bachelier, L. J. B. A. (1900). *Théorie de la spéculation*, Tesis de doctorado, Sorbonne, Annales scientifiques de l'École Normale Supérieure.

Bayes, F. R. S. (1763). An essay towards solving a problem in the doctrine of chances, *Philosophical Transactions of the Royal Society*, vol. 53: pp. 370–418.

Bernoulli, J. (1713). *Ars conjectandi, opus posthumum: accedit tractatus de seriebus infinitis, et epistola gallicè scripta de ludo pilae reticularis*, Basileae impensis Thurnisiorum.

Bernstein, S. N. (1926). Sur l'extension du théorème limite du calcul des probabilities aux sommes des quantitésdépendants, *Mathematische Annalen*, vol. 97: pp. 1–59.

Bessel, F. W. (1838). Untersuchungen über Wahrscheinlichkeit der Berobachtungsfehler, *Revist Astronomische Nachrichtena*, vol. 15: pp. 358–369.

Black, F. (1976). The pricing of commodity contracts, *Journal of Financial Economics*, vol. 3, núm. 1–2, pp. 167–179.

Black, F. y Scholes, M. (1973). The pricing of options and corporate liabilities, *The Journal of Political Economy*, vol. 81, núm. 3, pp. 637–654.

Byers, R. H. (2000). On the maximum of the standardized fourth moment, *Interstat*, vol. 2, pp. 1–7.

Canavos, J. C. (1988). *Probabilidad y estadística: aplicaciones y métodos*, 2^a ed. McGraw Hill.

Cardano, G. (1663). *Liber de Ludo Aleae*, Vol. 1, Opera Omnia.

Chebychev, P. L. (1887). Sur deux théorèms relatives aux probabilités, *Acta Mathematica*, vol. 14: pp. 98–116.

Climent Hernández, J. A. (2001). *Análisis teórico práctico para la valuación de opciones*, Tesis de licenciatura, Universidad Nacional Autónoma de México.

Climent Hernández, J. A. (2004). *Sistema de información electrónica para valuación de opciones*, Tesis de maestría, Universidad Nacional Autónoma de México.

Climent Hernández, J. A. (2005). *Valuación de opciones. Vínculo Matemático No. 38*, Coordinación de Servicios Editoriales de la Facultad de Ciencias. Universidad Nacional Autónoma de México.

Climent Hernández, J. A. (2013). *Opciones europeas en mercados α–estables y el consumidor estocástico*, Tesis de doctorado, Instituto Politécnico Nacional.

Climent Hernández, J. A. (2014). La ecuación de segundo grado en la estimación de parámetros de la martingala y la valuación de opciones americanas a través de la programación dinámica estocástica, *Estocástica: Finanzas y Riesgo*, vol. 4, núm. 2, pp. 155–190.

Cox, J. C., Ross, S. A. y Rubinstein, M. (1979). Option pricing: a simplified approach, *Journal of Financial Economics*, vol. 7, núm. 1, pp. 229–263.

Cramer, H. (1928). On the composition of elementary errors, *Skand Aktuarietids*, vol. 11, pp. 13–74.

De Finetti, B. (1937). La prévision : ses lois logiques, ses sources subjectives, *Annales de l'institut Henri Poincaré*, vol. 7, núm. 1: pp. 1–68.

De Moivre, A. (1710). D mensura sortis seu: de probabilitate eventuum in ludis a casu fortuito pendentibus, *Philosophical Transactions of the Royal Society of London Series I*, vol. 27: pp. 213-264.

De Moivre, A. (1725). *Annuities upon lives: or, the valuation of annuities upon any number of lives; as also, of reversions*, 1^a ed. W. P. and sold by Francis Fayram and B. Motte and W. Pearson.

De Moivre, A. (1733). *Approximatio ad summam terminorum binomii $(a+b)^n$ in seriem expansi*, London.

De Montmort, P. R. (1708). *Essay d'analyse sur les jeuxde hazard*, 1^a ed. Jacques Quillau, Paris.

De Morgan, A. (1847). *Formal logic: or, the calculus of inference, necessary and probable*, Mondon: Taylor and Walton.

DeGroot, M. H. y Schervish, M. J. (2012). *Probability and statistics*, 4^a ed. Pearson Education.

Devore, J. L. (1998). *Probabilidad y estadística para ingeniería y ciencias*, 4^a ed. Thomson Editores.

Devore, J. L. (2008). *Probabilidad y estadística para ingeniería y ciencias*, 7^a ed. Cengage Learning.

Erlang, A. K. (1909). The theory of probabilities and telephone conversations, *Nyt Tidsskrift for Matematik B*, vol. 20: p. 33.

Euler, L. (1738). De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt, *Commentarii academiae scientiarum Petropolitanae*, vol. 5: pp. 36–57.

Feller, W. (1943). The general form of the so-called law of the iterated logarithm, *Transactions of the American Mathematical Society*, vol. 54: pp. 373–402.

Feller, W. (1983). *Introducción a la teoría de probabilidades y sus aplicaciones*, Vol. I, 1^a ed. Limusa.

Fisher, R. A. (1922). On the mathematical foundations of the theoretical statistics, *Philosophical Transactions of the Royal Society of London*, vol. A222, pp. 309–368.

Fisher, R. A. (1925). Theory of statistical estimation, *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 22, núm. 5, pp. 700–725.

Freedman, D. y Diaconis, P. (1981). On the histogram as a density estimator: L_2 theory, *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, vol. 57, pp. 453–476.

Galton, F. (1964). On Hereditary character and talent, *MacMillan's Magazine*, vol. 11: pp. 157–166.

Galton, F. (1964). On Hereditary character and talent, *MacMillan's Magazine*, vol. 11: pp. 318–327.

Gauss, C. F. (1809). *Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, 1^a ed. Hamburg: Friedrich Perthes and I. H. Besser.

Hernández Arellano, F. M. (2003). *Cálculo de probabilidades*, 2^a ed. Sociedad Matemática Mexicana.

Hernández del Valle, A. y Hernández Lerma, O. (2003). *Elementos de probabilidad y estadística*, 1^a ed. Sociedad Matemática Mexicana.

Hildebran, D. K. y Ott, R. L. (1998). *Estadística aplicada a la administración y a la economía*, 3^a ed. Addison Wesley, México.

Hoel, P. G., Port, S. C. y Stone, C. J. (1971). *Introduction to probability theory*, The Houghton Mifflin Series in Statistics, Boston: Houghton Mifflin.

Huygens, C. (1657a). *De Ratiociniis in Ludo Aleae*, Ex officinia J. Elsevirii.

Huygens, C. (1657b). *Du calcul dans les jeux de hasard, Oeuvres Complètes de Christiaan Huygens*, Vol. XIV de *Traducción de De Ratiociniis in Aleae Ludo*, Martinus Nijhoff 1920.

Jarque, C. y Bera, A. (1987). A test for normality of observations and regression residuals, *International Statistical Review*, vol. 55, núm. 2, pp. 163–172.

Johnson, R. A. (2012). *Probabilidad y estadística para ingenieros*, 8^a ed. Pearson.

Kazmier, L. J. (1998). *Estadística aplicada a la administración y economía*, 3^a ed. McGraw–Hill.

Kolmogorov, A. N. (1931). Über die analytischen methoden in der wahrscheinlichkeitsrechnung, *Mathematische Annalen*, vol. 104: pp. 415–458.

Kolmogorov, A. N. (1933a). *Grundbegriffe der Wahrscheinlichkeitsrechnung*, Springer, Berlin.

Kolmogorov, A. N. (1933b). Sulla determinazione empirica di una legge di distribuzione, *Giornale dell' Istituto Italiano degli Attuari*, vol. 4: pp. 83–91.

Laplace, P. S. (1812). *Théorie analytique des probabilités*, Livre I. Calcul des fonctions génératrices, Troisième edition, Courcier, Paris, 1820.

Laplace, P. S. (1814). *Essai philosophique sur les probabilités*, Oeuvres completes. Académie des Sciences.

Levin, R. I. y Rubin, D. S. (2004). *Estadística para administradores*, 7^a ed. Pearson.

Lindeberg, J. W. (1922). Eine neue Harleitung des Exponentialgesetzes in der Wahrscheinlichkeitsrechnung, *Mathematische Zeitschrift*, vol. 15: pp. 211–225.

Lyapunov, A. M. (1901). *Nouvelle forme du théorème sur la limite des probabilités*, (Mémoires de l'Académie Impériale de Sciences de St. Pétersbourg).

Markov, A. A. (1898). The law of large numbers and the method of least squares, *Bulletin de la société physico-mathématique de Kasan*, vol. 2, núm. 8: pp. 110–128.

Merton, R. C. (1973). Theory of rational option pricing, *The Bell Journal of Economics and Management Science*, vol. 4, núm. 1: pp. 141–183.

Mises, R. (1947). On the asymptotic distribution of differentiable statistical functions, *The Annals of Mathematical Statistics*, vol. 18, núm. 3: pp. 309–348.

Montgomery, D. C. y Runger, G. C. (2003). *Probabilidad y estadística aplicadas a la ingeniería*, 2^a ed. Limusa Wiley.

Mood, A. M. y Graybill, F. A. (1978). *Introducción a la teoría de la estadística*, 4^a ed. Aguilar.

Neyman, J. (1934). On the two different aspects of the representative method: The method of stratified sampling and the method of purposive selection, *Journal of the Royal Statistical Society*, vol. 97, núm. 4: pp. 558–625.

Neyman, J. (1937). Outline of a theory of statistical estimation based on the classical theory of probability, *Philosophical Transactions of the Royal Society of London Series A, Mathematical and Physical Sciences*, vol. 236, núm. 763: pp. 333–380.

Neyman, J. y Pearson, E. S. (1928a). On the use and interpretation of certain test criteria for purposes of statistical inference: Part I, *Biometrika*, vol. 20A, núm. 1–2: pp. 175–240.

Neyman, J. y Pearson, E. S. (1928b). On the use and interpretation of certain test criteria for purposes of statistical inference: Part II, *Biometrika*, vol. 20A, núm. 3–4: pp. 263–294.

Neyman, J. y Pearson, E. S. (1933a). On the problem of the most efficient tests of statistical hypotheses, *Philosophical Transactions of the Royal Society of London. Series A*, vol. 231: pp. 289–337.

Neyman, J. y Pearson, E. S. (1933b). The testing of statistical hypotheses in relation to probabilities a priori, *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 29, núm. 4: pp. 492–510.

Neyman, J. y Pearson, E. S. (1936). Contributions to the theory of testing statistical hypotheses, *Statistical Research Memoirs*, vol. 1: pp. 1–37.

Pascal, B. (1653). *Traité du triangle arithmétique, avec quelques autres petits traitez sur la mesme matière*, Desprez, Guillaume. Publicado en 1965.

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling, *Philosophical Magazine Series*, vol. 5, núm 302: pp. 157–175.

Pearson, K. (1905). Skew variation, a rejoinder, *Biometrika*, vol. 4: pp. 169–212.

Poisson, S. D. (1837). *Recherches sur la probabilité des jugements en matière criminelle et en matière civile, précedées des regles générales du Calcul des Probabilités*, Bachelier, Paris.

Ramsey, F. P. (1931). *The Foundations of Mathematics and Other Logical Essays*, 1^a ed. Braithwaite, Cambridge.

Rincón Solís, L. A. (2008). *Curso Intermedio de Probabilidad*, 2^a ed. Las Prensas de Ciencias. Universidad Nacional Autónoma de México.

Rincón Solís, L. A. (2017). *Estadística descriptiva*, 1^a ed. Las Prensas de Ciencias. Universidad Nacional Autónoma de México.

Scott, D. W. (1979). On optimal data based histograms, *Biometrika*, vol. 66, núm. 3: pp. 605–610.

Shapiro, S. y Francia, R. (1972). An approximate analysis of variance test for normality, *Journal of the American Statistical Association*, vol. 67, núm. 337: pp. 215–216.

Shapiro, S. y Wilk, M. (1965). An analysis of variance test for normality (complete samples), *Biometrika*, vol. 52: pp. 591–611.

Smirnov, N. V. (1933). Estimate of deviation between empirical distribution functions in two independent samples, *Bulletin Moscow University*, vol. 2: pp. 3–16.

Student (1908). The probable error of a mean, *Biometrika*, vol. 4, núm. 1: pp. 1–25. Gosset, William Sealy.

Sturges, H. (1926). The choice of a class interval, *Journal of the American Statistical Association*, vol. 21, núm. 153: pp. 65–66.

Velleman, P. F. (1976). Interactive computing for exploratory data analysis I: Display algorithms, 1975 *Proceedings of the Statistical Computing Section*, pp. 142–147.

Wackerly, D. D., Mendelhall, W. y Sheaffer, R. L. (2008). *Estadística matemática con aplicaciones*, 7^a ed. Cengage Learning.

Walpole, R. E., Myers, R. H., Myers, S. L. y Ye, K. (2012). *Probabilidad y Estadística para Ingeniería y Ciencias*, 9^a ed. Pearson.

Índice analítico