UNIVERSIDAD NACIONAL DE COLOMBIA

Tratamiento de la Probabilidad y la Estadística para Principiantes

Jaime Enrique Niño Bernal

Universidad Nacional de Colombia
Departamento de Matemáticas
Bogotá, Colombia
2012

Tratamiento de la Probabilidad y la Estadística para Principiantes

Jaime Enrique Niño Bernal
Tesis presentada como requisito parcial para optar al título de:
Magister en Ciencias-Matemáticas

Directora:
Dr. rer. nat Liliana Blanco Castañeda

Universidad Nacional de Colombia
Departamento de matemáticas
Bogotá, Colombia
2012

Dedicatoria

A Dios por permitir conseguir este logro,
A Sergio y Gabriela que son mis motores,
A Patricia por su apoyo incondicional y paciencia.

"Llegará un día en el que el razonamiento estadístico será tan necesario para el ciudadano como ahora lo es la habilidad de leer y escribir"
H.G. Wells (1866-1946)

Agradecimientos

A la profesora Liliana Blanco Castañeda, por su apoyo, colaboración y confianza.
A Patricia por haberme apoyado en todo momento, por sus consejos, por la motivación constante.
A mis padres por ser el pilar fundamental en todo lo que soy, en toda mi educación, tanto académica, como de la vida.
A mi familia, por el apoyo frecuente y fuente de inspiración.
A mis compañeros de trabajo y amigos, quienes siempre tuvieron una voz de aliento.
A la Universidad Nacional de Colombia, por permitirme seguir desarrollando como profesional y como persona.

Resumen

El estudio de la estadística ha tomado fuerza a nivel mundial, la inclusión de ésta área en los programas educativos hace que los investigadores se preocupen por la didáctica envuelta en dichos procesos, por esta razón, y motivados por la falta de textos de estadísticas dirigidos a estudiantes de educación media, se elabora un texto de Probabilidad y Estadística para principiantes. Para el desarrollo del trabajo se consideran tres momentos principales: 1. La discusión con docentes del área sobre las debilidades presentes en la enseñanza de la estadística y el análisis de textos dirigidos al estudio de la estadística en bachillerato. 2. El análisis de los temas a incluir en el texto y el estudio de las investigaciones realizadas relacionadas con la educación estadística. 3. Organización de los temas seleccionados y redacción de las notas. Este texto presenta de manera clara los temas que abarcan los estándares curriculares de matemáticas, en relación con el pensamiento aleatorio y sistemas de datos, que establece el Ministerio de Educación Nacional de Colombia.

Palabras clave: probabilidad, estadística descriptiva, didáctica, educación media.

Abstract

The study of statistics has taken strength around the world, the inclusion of this area into the educative programs makes that researchers be concerned about the didactic involved in these process. For this reason, and motivated for the absence of texts about statistics focused to students of high school, we took on the task of developing a text Probability and Statistics for beginners. For the development of the worksheet, it has taken three main stages: 1. The discussion with the area teachers about the present weaknesses in the teaching of statistics, and the analysis of text focused in studies about statistic in high school. 2. The analysis of the topics in order to be included in the text and the study of research in relation with statistics education. 3. Organization of selected topics and construct the text. This text shows in a clear form the topics involved by the math curriculum standards, in relation with the random thought and data systems established by the Ministerio de Educación Nacional de Colombia.

Key words: probability, descriptive statistics, didactic, high school.

Contenido

Lista de figuras

Introducción

En los últimos 20 años se hace notoria la importancia que se le ha dado a la enseñanza de la estadística en niveles escolares básicos, muestra de ello son los proyectos realizados en diferentes partes del mundo, por ejemplo, en Inglaterra el *Schools Council Project in Statistical Education* [23] realizó el proyecto *Statistics in Your World*, en el que se unió la estadística a las diferentes asignaturas enseñadas en la escuela, como la geografía, las humanidades, las ciencias sociales, las ciencias y la vida, fue dirigido a estudiantes entre 11 y 16 años; en Estados Unidos el *Data Driven Curriculum Strand for High School Mathematics* [8], creado por la *National Science Fundation*, y dirigido a estudiantes de los grados 9 a 12; en Chile el programa de extensión en ciencia y tecnología en probabilidad y estadística llamado *explora*, coordinado por la *Comisión Chilena de la investigación en Ciencia y Tecnología* dedica una parte al proyecto *Azar, Ciencia y Sociedad* dirigido a estudiantes de 15 a 17 años.

En Colombia se realiza el primer acercamiento hacia la enseñanza de la estadística a nivel medio en el año 1978. En ese año se realizó la llamada renovación curricular, en la cual se incorporaron tópicos de estadística descriptiva al currículo de educación básica. En el año 1996 se elaboraron los lineamientos curriculares para el área de matemáticas, dentro de ellos se evidencian cinco tipos de pensamiento matemático, en el que se incluye el pensamiento aleatorio y sistemas de datos. Para el 2006 el Ministerio de Educación Nacional elabora los estándares básicos de competencias en el cual establece los conocimientos mínimos que deben alcanzar los estudiantes de los diferentes niveles académicos en educación básica y media, esto justificado con la necesidad de que todo ciudadano debe tener conocimiento matemático básico para desempeñarse en forma activa y crítica en su vida social y política y para interpretar la información necesaria en la toma de decisiones (MEN 2006).
Con relación al pensamiento aleatorio y sistemas de datos el MEN expresa textualmente: "pretende buscar soluciones razonables a problemas en los que no hay una solución clara y segura, abordándolos con un espíritu de exploración y de investigación mediante la construcción de modelos de fenómenos físicos, sociales o de juegos de azar y la utilización de estrategias como la exploración de sistemas de datos, la simulación de experimentos y la realización de los conteos"

En la actualidad la estadística está siendo difundida a nivel mundial. Cada vez es mayor la insistencia de académicos de diversas naciones en la necesidad que los ciudadanos sean estadísticamente cultos (Cuevas, 2008), la estadística permite a los ciudadanos ser críticos, asumir una actitud participativa en la sociedad, ser ciudadanos informados. Con esta necesidad se crearon asociaciones dedicadas al estudio de la educación estadística. Una de las más importantes, la IASE (*International Association for Statistical Education*), se dedica a promover la educación para el mejoramiento de la educación estadística. Esta asociación realiza cada cuatro años las conferencias ICONTS (*International congresses in Mathematical Education*) en las cuales uno de los temas principales es la educación estadística, también cobija el proyecto ISLP (*International Statistical Literacy Project*) que tiene como misión apoyar, crear y participar en actividades de alfabetización estadística y promoción en todo el mundo [15].

Un ejemplo claro de investigación en la enseñanza de la estadística es el realizado por la profesora española Carmen Batanero del departamento de matemáticas de la Universidad de Granada, quien con su grupo de investigación ha realizado importantes aportes a la educación estadística. Dentro de los contenidos analizados por este grupo se destacan los siguientes:

- Análisis exploratorio de datos: Gráficos estadísticos, medidas de posición central, asociación y correlación
- Probabilidad: Aleatoriedad, probabilidad, probabilidad condicional, variable aleatoria, distribución normal, teorema central del límite
- Inferencia: Intervalos de confianza, contraste de hipótesis, inferencia bayesiana, diseño de experimentos, uso de la inferencia en la investigación educativa

En la actualidad la sociedad exige ciudadanos críticos, participativos, preparados para tomar decisiones acertadas en su vida social y política, capaces de desarrollar acciones que, colectivamente, puedan transformar a la sociedad; la estadística es fundamental para desarrollar estas capacidades en cada individuo, la estadística es el lenguaje universal para argumentar y sustentar posiciones.

En países como España, México y Argentina, se han desarrollado trabajos en la misma dirección. Por ejemplo, en el año 2009 en el programa de doctorado de didáctica de las matemáticas de la Universidad de Granada España, la estudiante Silvia Azucena Mayén Galicia, realizo bajo la dirección de las Dras. Carmen Batanero Bernabeu y Carmen Díaz Batanero la tesis titulada comprensión de las medidas de tendencia central en estudiantes mexicanos de educación secundaria, cuyo objetivo principal era evaluar el significado personal que los estudiantes mexicanos de educación secundaria asignan a las medidas de tendencia central [18]. Otro trabajo realizado en el año 2003, fue elaborado por la profesora Belén Cobo Merino, quien en su tesis doctoral dirigida por la Dra. Carmen Batanero investigó el significado de las medidas de tendencia central para los estudiantes de secundaria [8], en el 2009 el docente del Departamento de Estadística e Investigación Operativa de la Universidad de Granada, José Miguel Contreras García, realizó un trabajo de investigación relacionado con la enseñanza de la probabilidad condicionada usando recursos en internet [9], y este año las profesoras argentinas de la Universidad Nacional de Catamarca (Unca) Norma Rodríguez, Graciela Montañés e Ilda Rojas, desarrollaron un estudio sobre las dificultades de los estudiantes universitarios en la comprensión de los contenidos de estadística inferencial [24].

De acuerdo a todo lo anterior y teniendo en cuenta la experiencia docente, se observa que es necesario disponer de herramientas dirigidas a estudiantes de educación básica y media académica, que al mismo tiempo orienten a los docentes encargados de dirigir la asignatura de estadística en estos niveles escolares. Los siguientes capítulos pretenden ofrecer esta herramienta, exponiendo de manera clara los temas correspondientes a los grados noveno, décimo y undécimo de la educación básica y media académica establecidos por el MEN en los estándares básicos de competencias [19].

1.BOSQUEJO HISTÓRICO

La estadística es una ciencia antigua, que tiene aplicación práctica en diversos campos científicos, tales como la biología, la medicina, la ingeniería, la psicología, etc. Su objetivo principal es reunir información, organizarla y presentarla de manera simplificada y fácil de entender.

Los comienzos de la estadística se remontan a épocas muy antiguas. El historiador y geógrafo griego Heródoto cuenta en los nueve libros de la historia, que los egipcios realizaban registros topográficos y demográficos, dando evidencia del uso de la estadística descriptiva en ese entonces [12]. En la Biblia, en el libro Números del antiguo testamento, escrito en el año 1400 a. C, se encuentran reportes de censos realizados por Moisés al pueblo de Israel después de la salida de Egipto. Moisés realizó un censo por familias y por linajes, describiendo los nombres de todos los varones aptos para el servicio de armas en Israel.

La excelente organización política, jurídica y administrativa de los romanos favoreció el uso de la estadística en su sociedad. Con el objetivo de conocer la cantidad de habitantes, y tener un registro de sus bienes, ellos realizaban un censo cada cinco años. Desde la caída de su imperio, no se encuentran registros estadísticos importantes durante un largo tiempo. En los años 758 y 762 Pipino el Breve y Carlo Magno, usan la estadística para realizar censos con el fin para conocer la extensión y el valor de las tierras pertenecientes a la Iglesia [6]. En el año 1085 el Rey Guillermo I de Inglaterra ordenó realizar un registro de las propiedades y ganado de los ingleses. La recopilación fue llamada *Domesday Book* o libro del día del juicio final [16], este registro fue muy parecido a los censos nacionales realizados en la actualidad.

En América Latina también hay muestras de uso de la estadística; el Imperio Incaico, utilizó la estadística, en los aspectos demográficos y económicos. Una prueba de esto es el uso de una cuerda hecha de lana, cáñamo, o algodón, llamada quipu. Estaba construida con una cuerda transversal de la cual colgaban cuerdas más delgadas que tenían una serie de nudos. En los nudos de los quipus se representaba la historia de los incas, relatando nacimientos, guerras, conquistas, número de habitantes, cantidad de animales, tierras, etc. [22]

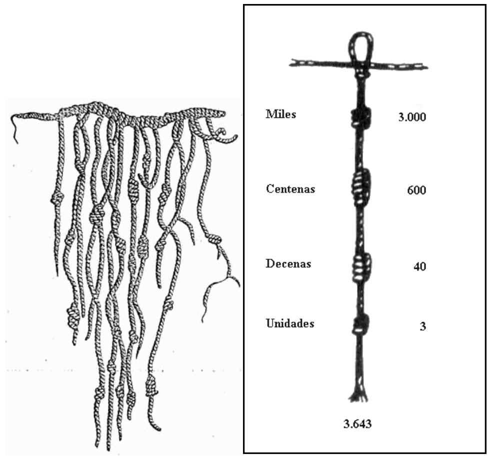


Figura 1.1. Quipu: instrumento usado por los Incas para registrar información[1]

En el siglo XVII el demógrafo londinense John Graunt (1620 - 1674), estudia la cantidad de nacimientos y defunciones ocurridas entre los años 1604 y 1661 en Londres. En el año 1662 publica el libro *Observaciones naturales y políticas a partir de los datos de mortalidad.* En este libro diseña una tabla de mortalidad, que expresaba las posibilidades de supervivencia para cada edad. En 1710 el médico, estadista inglés John Arbuthnot (1667 - 1735) escribió un artículo en el cual intentó probar la existencia de Dios [26], dando argumentos probabilísticos. Arbuthnot hace un análisis de las cifras de nacimientos de hombres y mujeres entre los años 1629 y 1710, y concluye que en todos estos años nacieron más hombres que mujeres, Arbuthnot argumentaba que esto no era posible por simple azar, puesto que la probabilidad de que naciera un hombre o una mujer debía ser la misma; por consiguiente el hecho de nacieran más varones se debía a una planificación previa, y puesto que, según él, era más beneficioso que nacieran más hombres que mujeres, ya que ellos realizaban trabajos más duros y por ende morían más rápido que las mujeres, concluye que debe haber una intervención divina que controla las proporciones de los sexos, lo cual demuestra la existencia de Dios.

La demostración de Arbuthnot es el primer ejemplo conocido de inferencia estadística.

Figura 1.2. Arbuthnot, Retrato por Godfrey Kneller

Hacia 1760 el economista Godofredo Anchenwall (1719 - 1772) nacido en Prusia, utiliza por primera vez, la palabra estadística. Anchenwall originalmente designó esta palabra

[1] Imagen http://es.wikipedia.org/wiki/Quipu

para el análisis de los datos de un gobierno, definiéndola como la “Ciencia del Estado”. La palabra estadística se deriva del término italiano statista, la cual a su vez procede de la raíz latina “status” que significa estado o situación.

Cuando hablamos de estadística, es inevitable hablar de probabilidad. La probabilidad nace con el fin de analizar los juegos de azar, principalmente los que involucraban dados y cartas. Los juegos de azar son tan antiguos como la misma humanidad. Se ha encontrado en sitios arqueológicos sumerios y asirios, una especie de dado, llamado astrágalo, el cual es un hueso extraído del talón de los mamíferos. Éste hueso se tallaba de tal forma que al lanzarlo podía caer en cuatro posiciones distintas, por el lado plano valía un punto, por el quebrado seis, por el cóncavo tres, y por el convexo cuatro.

Figura 1.3. Astrágalo con perforaciones indicativas de su valor (Alejandro Ramos Folqués: Tabas y dados)

El primero en analizar y tratar de resolver matemáticamente problemas relacionados con este tipo de juegos fue el italiano Girolamo Cardano (1501 - 1576), quien hacia 1560 escribió el libro *Liber de ludo aleae* (El libro de los juegos de azar). Este libro, es en principio, un manual para jugadores, pues describe las formas en que un jugador puede prevenir las trampas de los rivales. En el desarrollo de tales estrategias dedica buena parte de su libro al estudio del azar. Cardano establece que después de lanzarse muchas veces un dado, todas las caras tienden a aparecer la misma cantidad de veces, es decir que cada cara del dado tiene la misma oportunidad de aparición.

Durante el siglo XVII, los franceses Blaise Pascal (1623 - 1662) y Pierre de Fermat (1601 - 1665), intercambian una serie de cartas con el fin de tratar de solucionar algunos problemas de juegos de dados propuestos por el caballero de Meré (filósofo y jugador profesional de la época), quien era amigo de Blaise Pascal. Estas cartas son consideradas por muchos como el comienzo del cálculo de probabilidades [21]. En el año 1640 Pascal publica su tratado sobre el triángulo aritmético, este libro empieza con el conocido triángulo de Pascal, que sería la herramienta de conteo más útil hasta ese momento. Christian Huygens (1629 - 1694) continúa con los trabajos de Pascal y Fermat, y en el año 1656 publica un tratado sobre los cálculos de los juegos de azar, llamado *De Ratiociniis in Ludo Aleae* (Sobre el razonamiento de los juegos de azar). Durante el siglo XVIII la probabilidad tuvo grandes avances, Jacob Bernoulli (1654 - 1705) y Abraham de Moivre (1667 - 1754), fueron quienes más contribuyeron en este periodo, Jacob Bernoulli, escribió el tratado *Ars Conjectandi* (El arte de la previsión) sobre el cálculo de probabilidades, publicado en 1713, ocho años después de su muerte. Las contribuciones más significativas de Abraham De Moivre, se encuentran en su obra *Doctrine of Changes*, también llamada *A*

Method of Calculating the Probabilities of Events in Play. En esta obra el De Moivre introdujo el concepto de distribución normal como aproximación a la distribución binomial, concepto que trataremos más adelante.

Pierre Simon Laplace (1749 - 1827) escribió varias memorias relacionadas con la probabilidad, estas fueron incluidas en su obra principal Théorie Analytique des probabilités (La teoría analítica de las probabilidades). Esta obra es publicada en París en el año 1812 y en ella Laplace formaliza la teoría clásica de la probabilidad. --"Los temas tratados por Laplace cambian notablemente respecto a los de sus antecesores. Mientras que los matemáticos anteriores a Laplace investigan fundamentalmente problemas de juegos, es a partir de él cuando comienza a hacerse una formalización de la teoría de la probabilidad" [17] –

Karl Friedrich Gauss (1777 - 1855) uno de los matemáticos más grandes de la historia, conocido como "*El príncipe de las matemáticas*" [5]. Desde su infancia mostro ser un niño prodigio. Gauss se propuso calcular la órbita de un asteroide, para esto el demostró que los datos obtenidos experimentalmente, podían representarse por medio de una curva en forma de campana, que hoy se conoce como campana de Gauss, para esto uso el método de mínimos cuadrados.

En Moscú el matemático Chebyshev (1821 - 1894) contribuyó en el desarrollo de la teoría de la estadística y la probabilidad. En 1846 realiza su tesis de maestría titulada *An essay on the elementary analysis of the theory of probability* (Un ensayo sobre el análisis elemental de la teoría de la probabilidad), que se convirtió en un manual usado por las aseguradoras en Moscú y en San Petersburgo, pues les servía para analizar la posibilidad de ocurrencia de muertes, incendios, destrucción de mercancías, pérdida de objetos, de quienes iban a asegurar. Andrei Kolmogorov (1903-1987) en 1933 publica *Grundbegriffe der Wahrscheinlichkeitrechnung* (Fundamentos de Probabilidad), en donde axiomatiza dicha teoría, transformando el carácter del cálculo de probabilidades y convirtiéndolo en una disciplina matemática. La obra de Kolmogorov fue fundamental para el posterior desarrollo de la probabilidad y los procesos estocásticos.

2.ESTADÍSTICA DESCRIPTIVA

Pedro compra su periódico todos los días. Dentro de una de sus lecturas él se encontró con el siguiente gráfico:

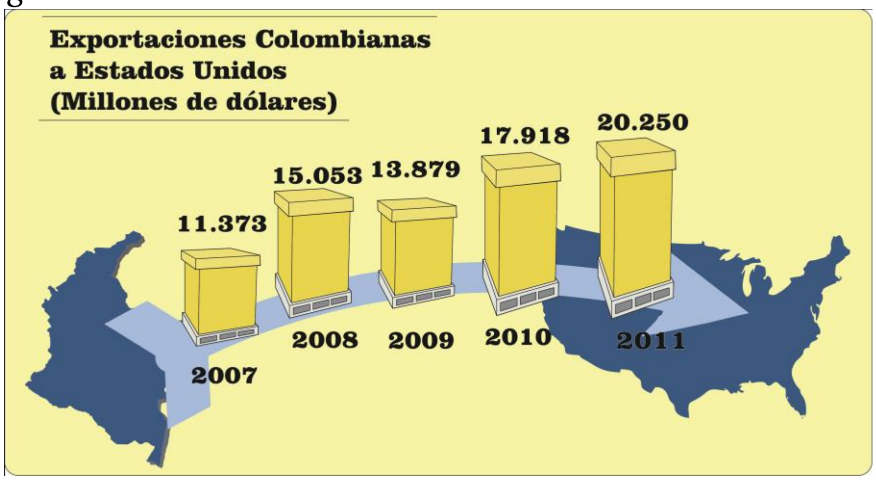


Figura 2.1

¿Cómo debe Pedro interpretar la información presentada en este dibujo?

La llamada estadística descriptiva permite analizar este tipo de información y sirve para representar los datos obtenidos en una investigación de forma resumida y fácil de entender.

2.1. Conceptos Preliminares

Para el estudio de la estadística descriptiva, se hace necesario, conocer algunos conceptos básicos como: población, muestra y variables.

Antes de dar estos conceptos de manera forma, veamos un ejemplo.

En las instituciones educativas públicas de Colombia el instituto de bienestar familiar (ICBF) ofrece el servicio de restaurante escolar. Con el fin de verificar el buen funcionamiento del programa, el ICBF realiza un seguimiento a los estudiantes que usan el restaurante revisando su talla y peso. Si el ICBF quiere tener mayor veracidad en la información recolectada, ¿a cuántos estudiantes les debe hacer el seguimiento?

El ICBF podría analizar la situación de dos maneras. La primera revisando la talla y el peso de todos los estudiantes que tienen acceso al restaurante, para poder concluir sobre el funcionamiento del programa. La segunda eligiendo al azar un grupo de estudiantes, analizar sus tallas y pesos, y con esta información concluir sobre la buena marcha del servicio de restaurante.

En el primer caso el análisis se realiza a todos los estudiantes, es decir al conjunto completo, el cual llamaremos ***población.***

Población (N): *Es el conjunto conformado por todos los elementos, individuos, objetos que vamos a estudiar.*

En el segundo caso, solamente analizamos una parte del conjunto completo o población. Este grupo se llama ***muestra***.

Muestra (n): *Es una parte de la población que tomamos para realizar un estudio estadístico.*

Cuando realizamos una investigación debemos tener especial cuidado en cómo elegir la muestra, por ejemplo si ejecutáramos una encuesta sobre las preferencias políticas, en un sector en el cual hay un dirigente político sobresaliente, seguramente el favoritismo se inclinará hacia su partido. Para que una muestra represente toda la población debe ser tomada al azar, ya sea utilizando sorteo, números aleatorios, o cualquier otro método de azar. Este tipo de muestra se llama probabilística, de otra manera la muestra se llamará no probabilística y es aquella en la que el investigador toma la muestra según su conveniencia.

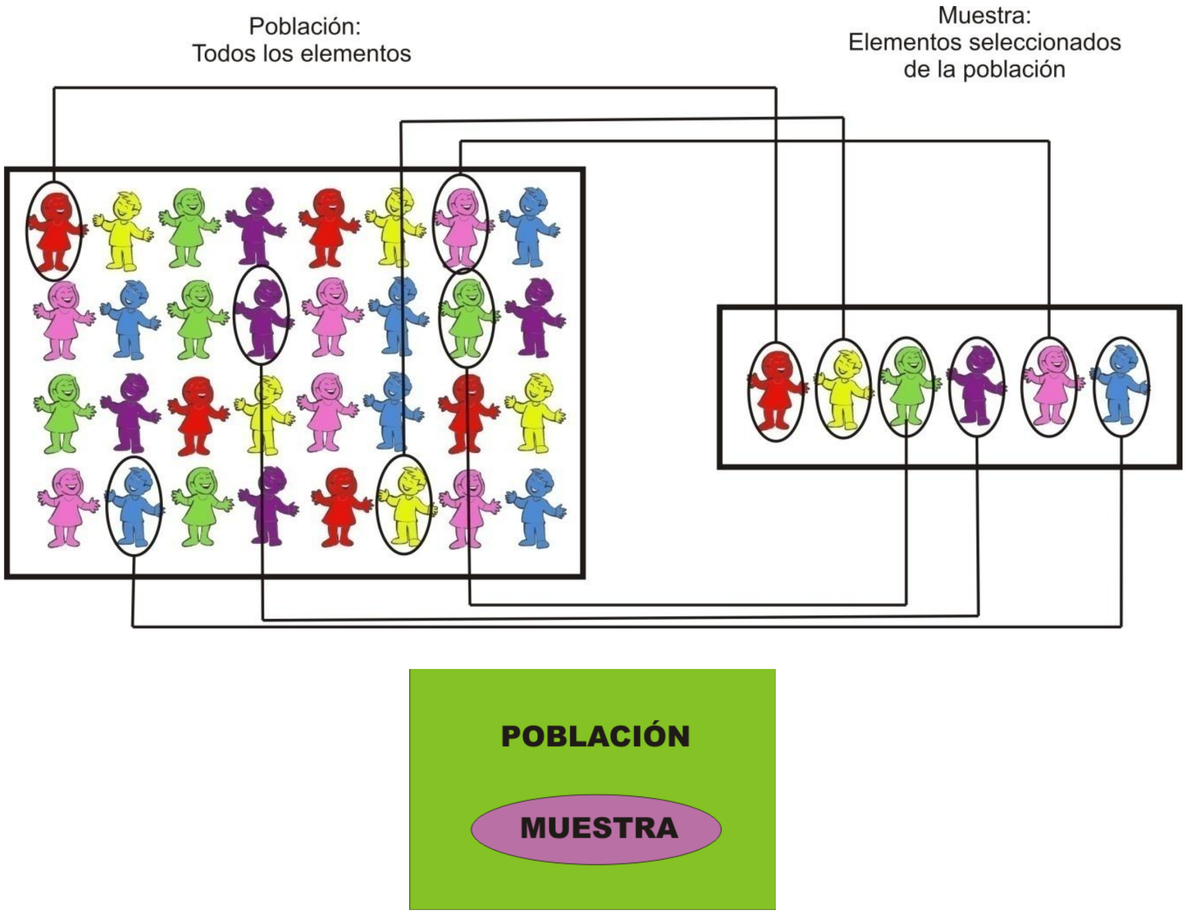


Figura 2.2. Representación de la población y la muestra

Ejemplo,
Un grupo de docentes de una institución educativa desean realizar un estudio sobre el nivel académico que presentan los estudiantes en el área de matemáticas. La institución cuenta con 2000 estudiantes desde el grado sexto al grado undécimo. Los profesores pensando en el costo que implicaría realizar la prueba a todos los estudiantes, deciden aplicarla únicamente a 7 estudiantes por grado, escogiéndolos por sorteo. Finalmente aplican la prueba a 399 estudiantes.

En este caso la población son los estudiantes de la institución, es decir los 2000 escolares (N = 2000) y la muestra son los 399 estudiantes que presentaron la prueba, (n = 399).

Ejemplo,
Se realizó una encuesta a los estudiantes de grado 9º sobre sus edades, de la cual se obtuvieron los siguientes resultados:

Tabla 2.1. Edad de un grupo de estudiantes

Edad (años)	Nº de estudiantes
13	3
14	15
15	10
16	4
Total	32

En este caso la población son los estudiantes de grado 9º. El total de estudiantes es 32 (N=32). Puesto que a todos los estudiantes se les realizó la encuesta, el tamaño de la muestra y la población es el mismo (n=32).

Cuando efectuamos un estudio estadístico, analizamos características específicas de los individuos que conforman la población o muestra, por ejemplo su edad, su talla, su color favorito, etc. Estas características reciben el nombre de variables. Las variables son aquellas que representan lo que estudiamos en cada individuo (sexo, edad, estatura, estrato, etc.), comúnmente se representan con las letras x e y. Los datos que se recogen tras una investigación son los valores que toma la variable en cada caso, por ejemplo si nuestra variable es el sexo, los posibles valores son: masculino o femenino.

Las variables se pueden dividir en dos:

Variables Cualitativas
Este tipo de variables representan una cualidad o atributo, como por ejemplo sexo, deporte favorito, lugar de residencia, estrato, etc. En el proceso de medición de estas variables, se pueden utilizar dos escalas:
Escalas nominales: estas variables no mantienen una relación de orden entre sí tales como: color favorito, sexo, profesión, etc.

Escalas ordinales: existe un cierto orden o jerarquía entre los valores que toma la variable, por ejemplo: estrato, nivel de estudios, rango militar, etc.

Variables Cuantitativas
Son las variables que pueden medirse, cuantificarse o expresarse numéricamente. Las variables cuantitativas pueden ser de dos tipos:
Variables cuantitativas discretas, suelen tomar solamente valores no fraccionables, por ejemplo: número de hijos, número de partos, número de hermanos, etc.
Variables cuantitativas continuas, admiten tomar cualquier valor numérico como: edad, peso, talla.

2.2 TABLAS DE FRECUENCIA

Cuando realizamos un estudio estadístico, es importante observar el comportamiento de los datos, es decir, qué datos aparecen en la investigación y cuántas veces aparecen dichos datos. Este número de apariciones se llama *frecuencia absoluta* y se representa con f_i. Dicha información se resume en la llamada *tabla de frecuencias*. Ésta nos permite ordenar y presentar los datos de manera clara.

Veamos el siguiente ejemplo:
Realizamos una encuesta a 200 estudiantes, en la que preguntamos acerca de sus deportes favoritos. Los resultados obtenidos fueron: fútbol: 65 estudiantes, baloncesto: 56 estudiantes, voleibol: 37 estudiantes, tenis: 29 estudiantes, otros deportes: 10 estudiantes y no le gusta el deporte a 3 estudiantes.
Esta información la podemos representar en una tabla. La variable investigada es cualitativa "deporte favorito" y la frecuencia absoluta representa el número de estudiantes que prefieren un deporte.

La tabla que podemos construir es:

Tabla 2.2. Deporte favorito de un grupo de estudiantes

Deporte Favorito (x_i)	Nº de estudiantes (f_i)
Fútbol	65
Baloncesto	56
Voleibol	37
Tenis	29
Otros	10
Ninguno	3
TOTAL	200

Los datos de esta tabla proporcionan exactamente el número de estudiantes que prefieren un determinado deporte.

La información que proporciona la tabla es útil, pero no es suficiente, decir que en este grupo hay 37 estudiantes a quienes les gusta el voleibol, proporciona poca información sobre si el número de escolares que prefieren este deporte es muy significativo, respecto al total de la población encuestada. Para valorar la representatividad de cada componente respecto al total de datos se calcula la *frecuencia relativa* (f_r), dividiendo la frecuencia absoluta (f_i) por el número total de observaciones (n), es decir,

$$\boldsymbol{f_r = \frac{f_i}{n}}$$

Por consiguiente la frecuencia relativa de la componente deporte favorito voleibol es:

$$\boldsymbol{f_r = \frac{37}{200}}$$

Lo que significa que 37 de cada 200 estudiantes de este grupo, prefieren el voleibol.

Otra medida importante a la hora de analizar datos y que nos proporciona información complementaria es el porcentaje. Este se calcula multiplicando la frecuencia relativa por 100.

Es decir el porcentaje de los estudiantes que prefieren el voleibol es:

$$\text{Porcentaje} = \boldsymbol{f_r \times 100 = 0.185 \times 100 = 18.5\%}$$

Podemos decir que el 18.5% de los estudiantes de este grupo prefieren el voleibol.

Realizando los procesos descritos anteriormente tabla correspondiente a los deportes favoritos es:

Tabla 2.3. Frecuencias del deporte favorito de un grupo de estudiantes

Deporte Favorito (x_i)	Frecuencia absoluta (f_i)	Frecuencia relativa (f_r)	%
Futbol	65	$\frac{65}{200} = 0{,}325$	0,325 x 100 = 32,5
Baloncesto	56	$\frac{56}{200} = 0{,}28$	0,28 x 100 = 28
Voleibol	37	$\frac{37}{200} = 0{,}185$	0,185 x 100 = 18,5
Tenis	29	$\frac{29}{200} = 0{,}145$	0,145 x 100 = 14,5
Otros	10	$\frac{10}{200} = 0{,}05$	0,05 x 100 = 5
Ninguno	3	$\frac{3}{200} = 0{,}015$	0,015 x 100 = 1,5
TOTAL	200	1	100

Veamos otro ejemplo:

En una encuesta realizada a 40 familias en un conjunto residencial, se les pregunto cuántas personas constituían su núcleo familiar.

Los datos registrados en la siguiente tabla corresponden a la información recopilada en la encuestas.

Tabla 2.4. Número de integrantes por familia en un conjunto residencial

Nº de personas (x_i)	N° Familias (f_i)
2	4
3	8
4	14
5	10
6	4
TOTAL	40

En este último ejemplo puede ser de interés conocer cuántas familias tienen menos de 4 integrantes, o, cuántas familias tienen más de 4 integrantes. Para tener esta información usaremos la llamada frecuencia acumulada $(\boldsymbol{F})$. La frecuencia acumulada se calcula sumando la frecuencia absoluta $(\boldsymbol{f_i})$ del valor que toma la variable y las frecuencias absolutas de los valores anteriores.
Veamos la tabla.

Tabla 2.5 Frecuencias del número de personas por familia en un conjunto residencial

Nº de integrantes (x_i)	Frecuencia absoluta (f_i)	Frecuencia acumulada (F)	Frecuencia relativa (f_r)	%
2	4	4	$\frac{4}{40}$	10
3	8	8+4 = 12	$\frac{8}{40}$	20
4	14	14+8+4 = 26	$\frac{14}{40}$	35
5	10	10+14+8+4=36	$\frac{10}{40}$	25
6	4	4+10+14+8+4 = 40	$\frac{4}{40}$	10
TOTAL	40	//	1	100

En la tabla podemos ver que la frecuencia acumulada $(\boldsymbol{F})$ correspondiente a la familia que tiene 4 integrantes es 26, esto quiere decir que hay 26 familias que tienen 4 integrantes o menos y las otras 14 familias tienen más de 4 integrantes.

Veamos otra tabla en la que calculemos la frecuencia acumulada ($\boldsymbol{F}$)

Tabla 2.6. Edades de un grupo de estudiantes de grado 9

Edad (años) $\boldsymbol{x_i}$	$\boldsymbol{f_i}$	$\boldsymbol{F}$	$\boldsymbol{f_r}$	**%**
13	3	3	$\frac{3}{32} = 0{,}09375$	0,9375 x 100 = 9,375
14	15	3+15=18	$\frac{15}{32} = 0{,}46875$	0,46875 x 100 = 46,875
15	10	3+15+10=28	$\frac{10}{32} = 0{,}3125$	0,3125 x 100 = 31,25
16	4	3+15+10+4=32	$\frac{4}{32} = 0{,}125$	0,125 x 100 = 12,5
Total	32	//	1	100

De la información contenida en la tabla, podemos obtener conclusiones como las siguientes:
Hay 18 estudiantes menores de 15 años y por consiguiente 14 estudiantes de 15 años o más. El 12,5 % de los estudiantes encuestados tienen 16 años, también podemos observar que 15 de cada 32 estudiantes tienen 14 años.

Supongamos ahora que se les preguntó la edad cumplida a 50 habitantes de un barrio y que obtuvimos los siguientes datos:

1	30	80	11	21	54	42	76	65	89
7	23	40	81	22	50	46	70	69	82
9	32	42	40	48	51	16	10	35	32
3	43	40	54	44	75	13	18	15	28
9	34	22	47	33	60	71	84	66	65

Nos podemos dar cuenta, que la mayoría de las personas encuestadas en este barrio tienen edades diferentes, esto hace que la tabla de frecuencia sea extensa y no resuma los datos, como es el objetivo. Por esta razón la posibilidad de encontrar tantos valores diferentes obliga a organizar los datos por grupos, estos son llamados intervalos o clases. Es decir en vez de analizar cuántos habitantes tienen un año de edad, cuántos habitantes tienen dos años y así sucesivamente, se puede analizar cuántos habitantes tienen entre 0 y 5 años de edad, o cuántos habitantes tienen entre 0 y 10 años.

Una manera de representar estos datos en una tabla de frecuencia, haciendo grupos es:

Tabla 2.7. Edad de 50 habitantes de un barrio

Edad (x_i)	**Nº de habitantes** (f_i)
De 1 a 10 años	6
De 11 a 20 años	5
De 21 a 30 años	6
De 31 a 40 años	8
De 41 a 50 años	8
De 51 a 60 años	4
De 61 a 70 años	5
De 71 a 80 años	4
De 81 a 90 años	4
TOTAL	50

Veamos otro ejemplo,
Los siguientes datos corresponden a los resultados obtenidos en un examen de matemáticas realizado a un grupo de estudiantes de grado 8°,

0.5	4.5	6.5	1.5	2.1
4.2	1.8	6.1	8.0	1.0
5.2	7.5	2.0	5.8	9.9
4.0	7.9	9.0	7.6	3.5
5.3	8.5	2.5	9.5	6.0
6.7	3.0	4.9	6.4	4.4
5.9	8.2	6.6	3.8	7.0
4.1	6.3	10.0	7.5	5.5

Podemos ver que sucede algo similar al caso analizado anteriormente, por consiguiente lo recomendable es organizar los datos en intervalos.

La siguiente tabla de frecuencia, nos muestra una manera de organizar la información suministrada.

Tabla 2.8. Resultados obtenidos en un examen de matemáticas, realizado a un grupo de estudiantes de grado 8°

Notas (x_i)	Nº estudiantes (f_i)
De 0 a 2,0	5
De 2,1 a 4,0	6
De 4,1 a 6,0	11
De 6,1 a 8,0	12
De 8,1 a 10,0	6
Total	40

Existen diferentes formas para determinar el número de intervalos que necesitamos para construir la tabla, en algunos casos los investigadores determinan los intervalos a conveniencia. Por esta razón enunciaremos los pasos recomendados, para organizar los datos en intervalos:

Paso 1.
Hallar la longitud del intervalo que contiene todos los datos que estamos analizando, para ello usamos el dato mayor o máximo $(Xmax)$ y el dato menor o mínimo $(Xmin)$, esta medida recibe el nombre de rango. Formalmente lo podemos expresar como:

Rango (R): Es la diferencia entre el valor máximo que toman los datos y el valor mínimo.

$$R = Xmax - Xmin$$

Paso 2.
Hallar el número de intervalos, o número de clases $(\#C)$ que se usarán en la tabla. El número de intervalos es arbitrario, depende del número total de observaciones. Se recomienda que no haya más de 18 intervalos ni menos de 5. Si hay pocos no se pueden observar las características importantes de los datos y si hay muchos no se obtiene un resumen apropiado de los mismos. Existen varias formas que permiten calcular el número de clases, pero ninguna es exacta. Uno de los criterios más usados es en el cual se obtiene calculando la raíz cuadrada al número de datos, es decir,

$$\#C = \sqrt{n}$$

En donde n corresponde a la cantidad de datos que vamos a agrupar.

Paso 3.
Determinar el tamaño del intervalo, o amplitud (A) usando la siguiente razón

$$A = \frac{R}{\#C}$$

Ejemplo
Presentamos las alturas en centímetros de un grupo de estudiantes de una institución educativa.

Datos de estatura en centímetros de un grupo de estudiantes

132	160	160	139	143	135	150	163	144	133
168	173	145	162	153	137	165	154	161	176
150	141	152	177	140	173	155	147	180	169
162	138	134	142	149	161	180	178	172	167
158	143	154	163	155	137	168	140	153	170

Para presentar los datos de una manera ordenada y de tal forma que se pueda tener una visión general de esta información, podemos realizar una tabla de frecuencia.

Paso 1.
Calculemos el rango:
Para esto buscamos el dato más grande y el dato más pequeño del conjunto de datos

$$Xmax = 180;\ Xmin = 132$$

Luego calculamos la diferencia entre los dos

$$R = Xmax - Xmin = 180 - 132 = 48$$

Paso 2.
Calculamos el número de intervalos correspondiente:

$$\#C = \sqrt{n} = \sqrt{50} = 7{,}071 \cong 7$$

Paso 3.
Por último calculamos la amplitud de los intervalos

$$A = \frac{R}{\#C} = \frac{48}{7{,}071} = 6.788$$

Para la amplitud siempre aproximamos al siguiente entero en caso de que haya decimal (aproximación por exceso)

$$A \cong 7$$

Con la información obtenida podemos realizar la tabla, el primer intervalo debe empezar en el dato menor ($Xmin$), al que debemos sumar la amplitud (A). Así el primer intervalo empezará en 132 y terminará en 139

Si el extremo superior de cada clase coincide con el inferior de la siguiente, los intervalos se suponen semiabiertos por la derecha. Es decir, en cada clase se incluyen los valores de la variable que sean mayores o iguales al extremo inferior del intervalo, pero estrictamente menores que el extremo superior.
Por ejemplo en el intervalo [132; 139) incluimos el número de estudiantes con una estatura de 132 cm, pero no el número de estudiantes con estatura de 139 cm, este última se debe incluir en el intervalo [139; 146).

Para construir los demás intervalos, sumamos la amplitud al extremo superior, de los intervalos anteriores

Intervalo 1	[132; 139)
Intervalo 2	[139; 146)
Intervalo 3	[146; 153)
Intervalo 4	[153; 160)
Intervalo 5	[160; 167)
Intervalo 6	[167; 174)
Intervalo 7	[174; 181)

Los demás cálculos de las frecuencias se realizan de la misma forma que en los ejemplos anteriores, por consiguiente, la tabla de frecuencia para la estatura del grupo de estudiantes corresponde a:

Tabla 2.9. Estatura de un grupo de estudiantes

$Estatura\ (cm)$	fi	F	fr	%
[132; 139)	7	7	0.14	14
[139; 146)	9	16	0.18	18
[146; 153)	5	21	0.1	10
[153; 160)	7	28	0.14	14
[160; 167)	9	37	0.18	18
[167; 174)	8	45	0.16	16
[174; 181)	5	50	0.1	10
Total	50	/	1	100

Veamos alguna información que podemos obtener de la tabla

El 14% de los estudiantes tienen una estatura que se encuentra entre 132 y 139 centímetros, solamente el 10% de los estudiantes tienen una estatura que supera los 174 centímetros y 7 de cada 50 estudiantes tienen una estatura que se encuentra entre los 153 y 160 centímetros.

2.2.1 ACTIVIDADES

1. Determine cuáles de las siguientes variables son continuas y cuáles son discretas.
   a. Número de integrantes de una familia
   b. m^3 de agua consumidos en un apartamento
   c. Número de galones de gasolina vendidos en una estación de servicio.
   d. Producción diaria de huevos en un galpón.
   e. Candidatos a representantes estudiantiles en una institución.
   f. Edad de las personas en años cumplidos

2. Busque en el periódico noticias que presenten estudios estadísticos, e identifique la población, la muestra y el tipo de variable.

3. Las siguientes son las edades de un grupo de 96 personas que viven en el conjunto residencial "las flores".

42 60 60 38 10 63 21 66 56 57 51 57 12 44 45 35
30 35 47 53 49 5 50 49 8 45 28 41 47 42 53 32
54 38 40 63 48 33 35 8 61 47 41 55 53 27 20 21
42 21 39 39 34 45 39 28 54 33 35 10 12 48 48 27
53 30 29 53 38 52 13 54 27 27 43 28 63 41 23 58
56 11 59 13 40 24 20 58 35 62 27 30 9 40 52 60

a) Construya una tabla de frecuencias de tipo continuo, en la que se incluya: frecuencia absoluta, frecuencia acumulada y frecuencia relativa.
b) ¿Qué porcentaje de personas tiene más de 40 años?
c) ¿Cuántas personas tienen menos de 30 años?

4. Estos son datos sobre ocupación en una ciudad, por sectores económicos:

Sector	N° de personas
Agricultura	135.220
Industria	62.350
Construcción	55.000
Otros	202.325

a) Construya una tabla de frecuencias en la que se incluya: frecuencia absoluta, frecuencia acumulada y frecuencia relativa.
b) ¿Qué porcentaje de habitantes se dedican a la construcción?

5. Las calificaciones obtenidas por los 40 estudiantes de grado 8º en una prueba de Matemáticas vienen dadas por la siguiente tabla:

Nota	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
Alumnos	**1**	**2**	**6**	**8**	**6**	**6**	**6**	**4**	**1**

a) Construya una tabla de frecuencias en la que se incluya: frecuencia absoluta, frecuencia acumulada y frecuencia relativa.
b) ¿Qué porcentaje de estudiantes aprueba el examen?
c) ¿Qué porcentaje de estudiantes obtiene más de 7 puntos?
d) ¿Qué porcentaje de estudiantes obtiene menos de 6 puntos?

2.3 GRÁFICOS

Aun cuando las tablas de frecuencia nos proporcionan valiosa información, se hace necesario usar herramientas que nos permitan interpretar los datos de manera fácil y rápida. Los gráficos resultan ser instrumentos muy importantes para la presentación de información.

Es muy común encontrar en los diarios gráficos como el siguiente:

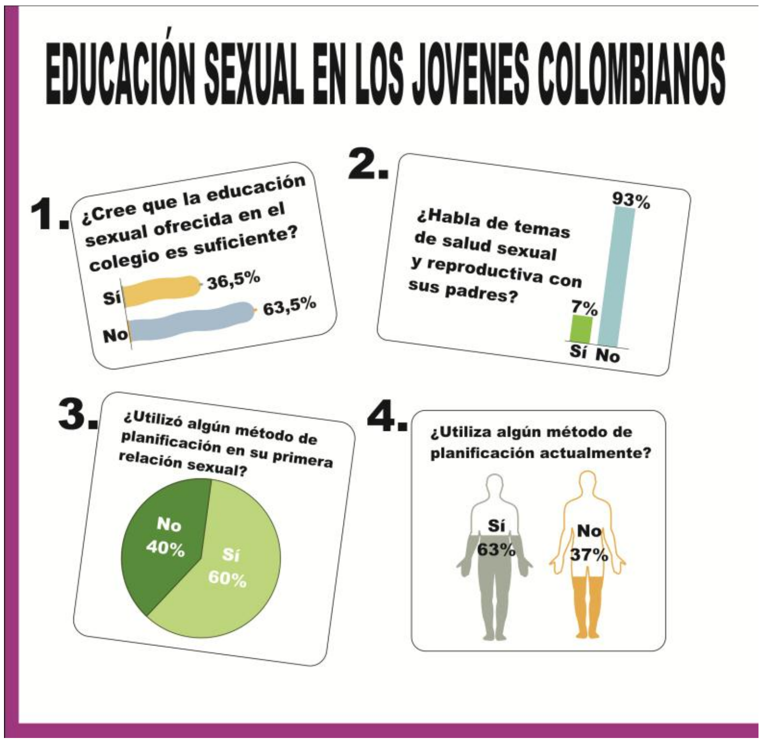


Figura 2.3.

En esta gráfica podemos observar tres tipos de diagramas estadísticos. El primero que corresponde a las preguntas 1 y 4, recibe el nombre de pictograma, y es sencillamente un dibujo que representa de manera clara información correspondiente a una variable. El segundo es el diagrama circular corresponde a la pregunta 3, generalmente se usa para resaltar la diferencia entre los porcentajes. El tercero es el diagrama de barras que corresponde a la pregunta 2. Es uno de los gráficos más usados para la representación de datos, por su facilidad de construcción y facilidad de comprensión. Se utiliza generalmente en variables cualitativas. Normalmente se usan la barras verticales, en donde en el eje vertical se ubican las frecuencias y en el eje horizontal se ubican las variables; sin embargo podemos encontrar barras horizontales, en donde el eje vertical representa las variables y el eje horizontal las frecuencias.

Veamos el ejemplo de las preferencias deportivas.

Tabla 2.10. Deporte favorito de un grupo de estudiantes

Deporte Favorito (x_i)	Nº de estudiantes (f_i)
Fútbol	65
Baloncesto	56
Voleibol	37
Tenis	29
Otros	10
Ninguno	3
TOTAL	200

La longitud de la barra representa la frecuencia de la variable. Para el caso de fútbol, la longitud de la barra debe ser 65 en la escala, para tenis debe ser 14, y así los demás valores que toma la variable.

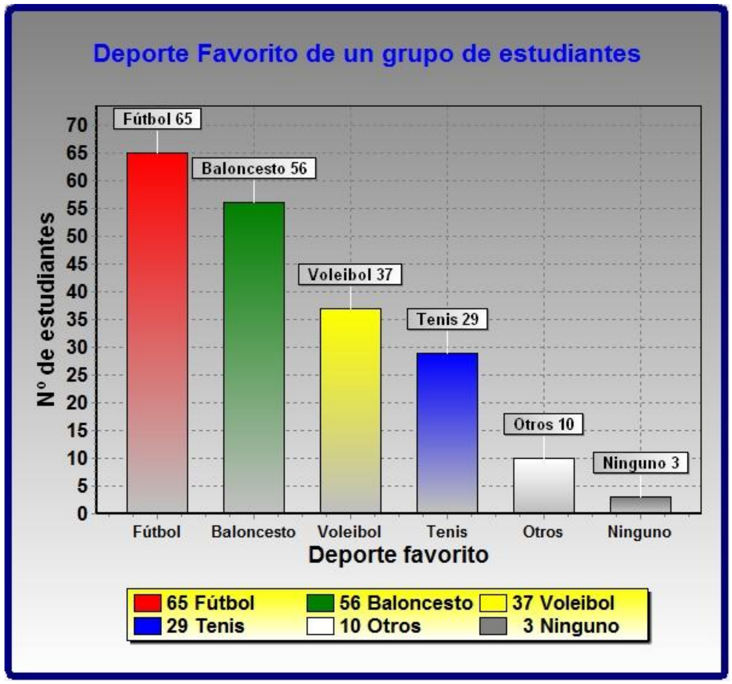


Figura 2.4.

Algunas veces resulta más conveniente presentar la información, en términos de porcentaje.

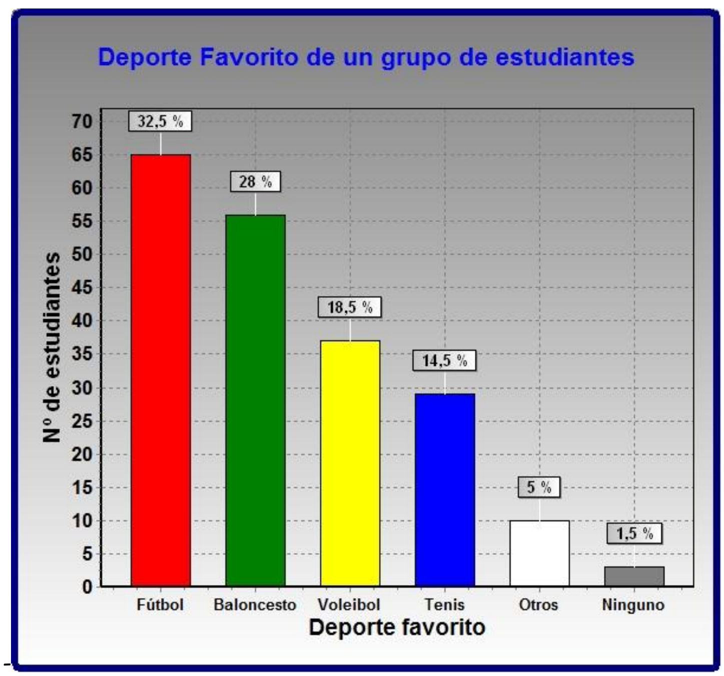


Figura 2.5.

Veamos otro ejemplo:

Tabla 2.11. Resultados obtenidos en examen de matemáticas, realizado a un grupo de estudiantes

Notas (x_i)	Nº estudiantes (f_i)	%
De 0 a 2,0	5	12,5
De 2,0 a 4,0	6	15
De 4,0 a 6,0	11	27,5
De 6,0 a 8,0	12	30
De 8,0 a 10,0	6	15
Total	40	100

El procedimiento de elaboración de un diagrama de barras horizontal es el mismo que el diagrama vertical.

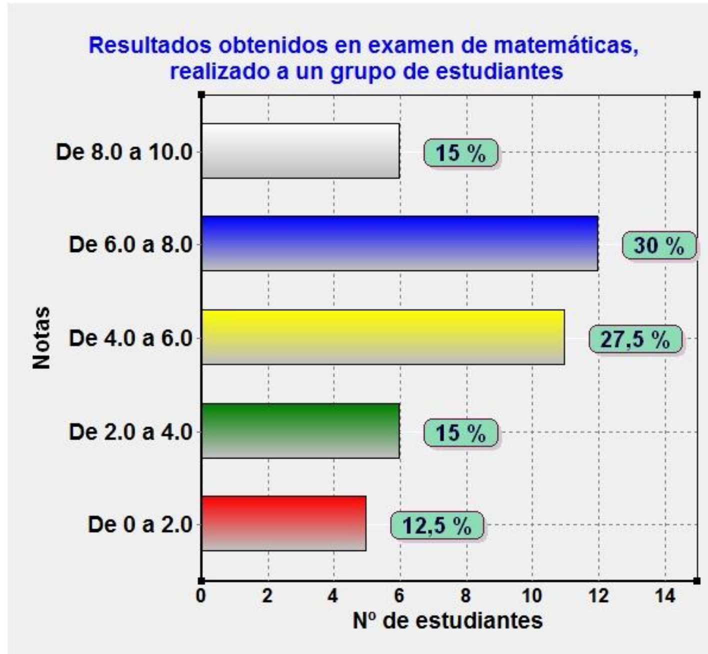


Figura 2.6.

En el caso que los datos estén organizados en intervalos resulta útil hacer la representación gráfica a través de los llamados histogramas de frecuencia. Estos son un conjunto de rectángulos en los que su base es del ancho del intervalo y su alto representa la frecuencia. Veamos el histograma correspondiente a las notas de los estudiantes del ejemplo anterior.

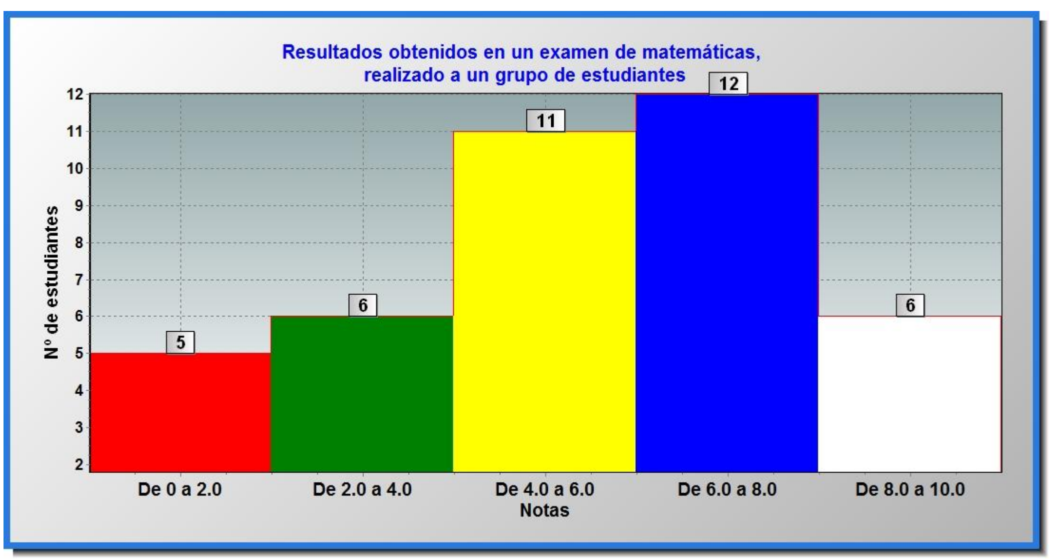


Figura 2.7.

Veamos otro ejemplo

Tabla 2.12. Estatura de un grupo de estudiantes

$Estatura$ (cm)	fi
[132; 139)	7
[139; 146)	9
[146; 153)	5
[153; 160)	7
[160; 167)	9
[167; 174)	8
[174; 181)	5
Total	50

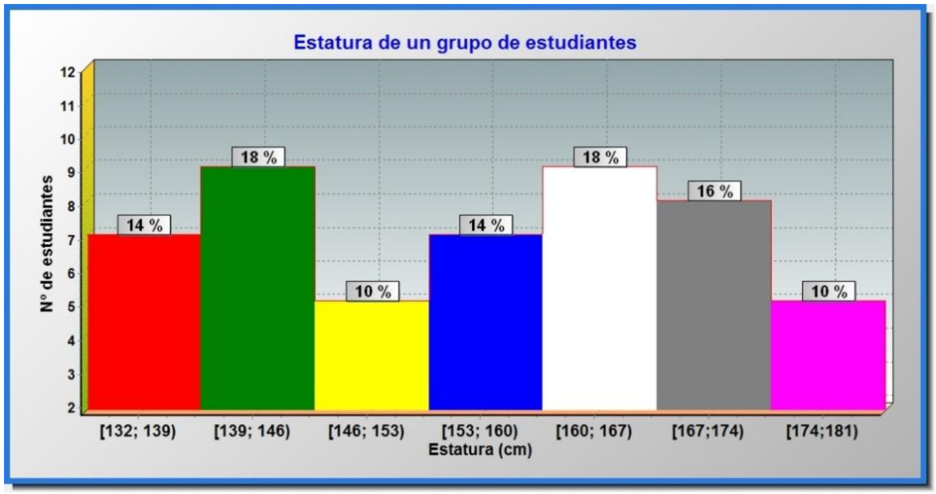


Figura 2.8.

Cuando se presenta la información en porcentajes, es usual representarla haciendo uso del llamado diagrama circular, pastel, o gráfico de sectores. Generalmente se usa para resaltar la diferencia entre los porcentajes.

Como su nombre lo indica su representación es una circunferencia. El proceso de elaboración de un diagrama circular es el mismo sin importar el tipo de variable que estemos analizando. Una manera fácil de construirlo es multiplicando la frecuencia relativa ($\boldsymbol{fr}$) por 360, de esta manera se obtiene el ángulo que corresponde al porcentaje.

$$Ángulo = \boldsymbol{fr} \times \mathbf{360}$$

Tabla 2.13. Deporte favorito de un grupo de estudiantes

Deporte Favorito (x_i)	Frecuencia absoluta (f_i)	Frecuencia relativa (f_r)	Ángulo
Fútbol	65	0,325	0,325 x 360 = 117
Baloncesto	56	0,28	0,28 x 360 = 100,8
Voleibol	37	0,185	0,185 x 360 = 66,6
Tenis	29	0,145	0,145 x 360 = 52,2
Otros	10	0,05	0,05 x 360 = 18
Ninguno	3	0,015	0,015 x 360 = 5,4
TOTAL	200	1	360

Después de calcular el ángulo correspondiente a cada porcentaje, con ayuda del transportador se procede a ubicarlos en la circunferencia.
Se recomienda que cada porción del pastel sea pintada de un color diferente, esto facilita la comprensión del diagrama y es importante que cada porción este marcada.

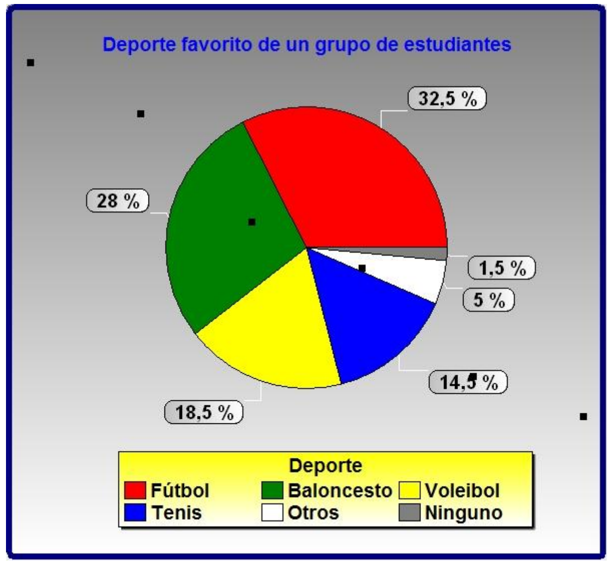


Figura 2.9.

Veamos otro ejemplo

Tabla 2.14. Estatura de un grupo de estudiantes

Estatura (cm)	fi	fr	%	Ángulo
[132; 139)		0.14	14	0.14 × 360 = 50.4
[139; 146)	9	0.18	18	0.18 × 360 = 64.8
[146; 153)	5	[illegible].1	10	0.1 × 360 = 36
[153; 160)	7	0.14	14	0.14 × 360 =50.4
[160; 167)	9	0.18	18	0.18 × 360 =64.8
[167; 174)	8	0.16	16	0.16 × 360 = 57.6
[174; 181)	5	0.1	10	0.1 × 360 = 36
Total	50	1	100	360

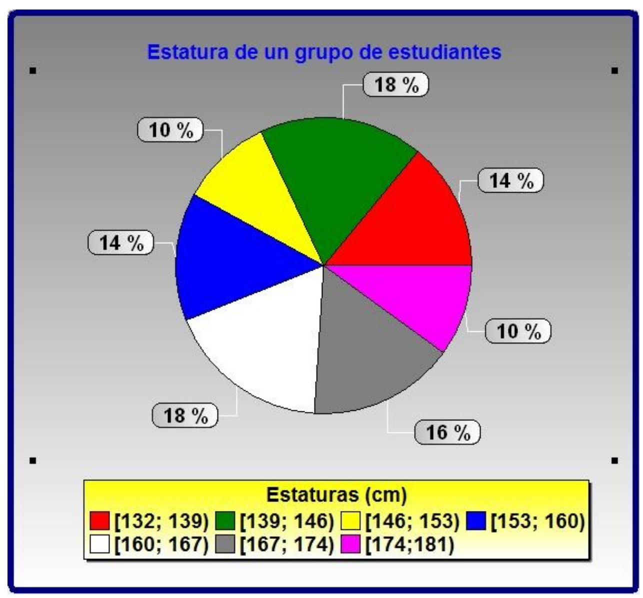


Figura 2.10.

2.3.1 ACTIVIDADES

1. La siguiente tabla representa las ventas, en miles de pesos, realizadas durante una semana en el supermercado "La Rebajita"

Producto	f_i	f_r
Elementos de aseo	525	0,265
Alimentos	850	0,430
Papelería	150	0,075
Otros	450	0,227
Total	1975	0,997

Realice un diagrama circular que represente la información dada en la tabla.

2. La siguiente tabla, representa el peso de los niños recién nacidos, durante una semana, en el hospital San José de Duitama.

Peso (en kg)	Número de niños
[2,5; 3)	8
[3; 3,5)	26
[3,5; 4)	14
[4; 4,5)	12
Total	60

Represente gráficamente la información contenida en la tabla

3. La siguiente tabla refleja las calificaciones de 30 alumnos en un examen de Matemáticas:

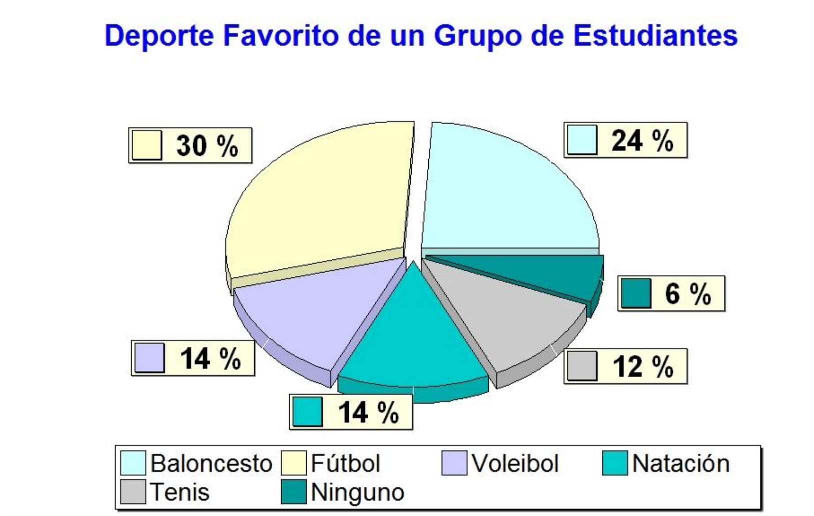
Nota	2	4	5	6	7	8	9	10
Nº alumnos	2	5	8	7	2	3	2	1

Dibuje un diagrama de barras con frecuencias absolutas acumuladas y un diagrama de barras con frecuencias absolutas usando como variable la nota.

4. En una encuesta realizada a 60 personas se les preguntó sobre las preferencias en géneros musicales. Los resultados se representaron en la siguiente gráfica:

a) Construya una tabla de frecuencias, usando la información de la gráfica
b) Realice un diagrama circular con la información de la gráfica.
c) ¿A qué porcentaje de las personas encuestadas les gustan la música clásica? ¿A qué porcentaje les gusta la música electrónica?

5. En una clase de 60 estudiantes, se les preguntó cuáles eran sus deportes favoritos. Los resultados se resumieron la siguiente gráfica.



a) Construya una tabla de frecuencias, usando la información de la gráfica
b) Realice un diagrama de barras con la información del gráfico.

2.4. MEDIDAS DE TENDENCIA CENTRAL

Si se quisiera premiar un estudiante por tener el mejor desempeño en álgebra. ¿De qué manera podemos determinar el ganador?
Supongamos que las notas obtenidas por tres estudiantes son las siguientes:

Estudiante 1: 3.2; 2.0; 4.1; 3.0
Estudiante 2: 4.2; 1.0; 2.0; 2.5
Estudiante 3: 3.2; 3.5; 3.2; 3.6

¿Cómo podemos comparar el rendimiento académico de los tres estudiantes? ¿Cuál de los tres estudiantes tuvo un mejor desempeño?

Lo primero que podemos pensar, es sumar las notas de cada estudiante y dividir el resultado por el número de notas. Esta medida recibe el nombre de *media aritmética* ($\bar{x}$).

Veamos cómo se calcula la media aritmética de las notas del estudiante **1** el primer dato es 3.2, el segundo 2.0, el tercero 4.1 y el cuarto 3.0 ($x_1 = 3.2;\ x_2 = 2.0;\ x_3 = 4.1;\ x_4 = 3.0$)

$$\bar{x} = \frac{3.2 + 2.0 + 4.1 + 3.0}{4} = 3.075$$

El valor 3.075 representa la nota promedio del estudiante **1**.

Calculemos las medias aritméticas ($\bar{x}$) de los estudiantes **2** y **3**

Estudiante 2: ($x_1 = 4.2;\ x_2 = 1.0;\ x_3 = 2.0;\ x_4 = 2.5$)

$$\bar{x} = \frac{4.2 +\ 1.0 + 2.0 +\ 2.5}{4} = 2.425$$

Estudiante 3: ($x_1 = 3.2;\ x_2 = 3.5;\ x_3 = 3.2;\ x_4 = 3.6$)

$$\bar{x} = \frac{3.2 +\ 3.5 +\ 3.2 +\ 3.6}{4} = 3.45$$

Con los valores de las medias aritméticas de cada estudiante podemos comparar sus rendimientos y concluir que el mejor desempeño corresponde al estudiante número 3.

La media o media aritmética nos permite comparar grupos de datos estadísticos correspondientes a la misma variable, se usa en las variables cuantitativas y se representa con el símbolo $\boldsymbol{\bar{x}}$.
De manera general se tiene que:

Media ($\boldsymbol{\bar{x}}$)*: Es la suma de todos los valores que toma la variable dividida en el número de valores observados.*

Supongamos que una variable cuantitativa x_i, toma los valores $x_1, x_2, x_3, \dots, x_n$ (pueden estar repetidos). Se define la media aritmética de los valores de x como:

$$\boldsymbol{\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n} = \frac{\sum x_i}{n}}$$

en donde:
$\boldsymbol{x_i}$ representa los valores que toma la variable, $\boldsymbol{n}$ representa el número total de observaciones, y la letra griega sigma ($\boldsymbol{\Sigma}$) se usa para indicar la suma.

Supongamos que un estudiante obtiene las siguientes calificaciones:

3.0	**3.5**	**4.0**	**3.5**	**4.0**	**3.5**	**4.0**	**2.0**	**3.5**	**3.0**

La media aritmética correspondiente será:

$$\bar{x} = \frac{3.0 + 3.5 + 4.0 + 3.5 + 4.0 + 3.0 + 4.0 + 2.0 + 3.5 + 3.0}{10} = 3.4$$

Como podemos ver algunos valores que toma la variable se repiten, por ejemplo la calificación 3.0 se repite 2 veces, la calificación 3.5, se repite 4 veces, la calificación 4.0 se repite 3 veces y la calificación 2.0 una vez. Por consiguiente podemos escribir la media de la siguiente manera:

$$\bar{x} = \frac{3.0(2) + 3.5(4) + 4.0(3) + 2.0}{10} = 3.4$$

Recordemos que el número de veces que se repite el dato lo llamamos frecuencia absoluta (f_i).

Por consiguiente, cuando queremos calcular la media ($\bar{x}$) a un grupo grande de datos resulta conveniente primero organizarlos en una tabla de frecuencias, como veremos a continuación:

Los siguientes datos, corresponden a la edad de un grupo de estudiantes de grado noveno:

12	**13**	**14**	**12**	**15**	**16**	**12**	**15**
15	**14**	**15**	**14**	**15**	**14**	**13**	**14**
14	**13**	**15**	**14**	**15**	**14**	**16**	**14**
13	**14**	**13**	**15**	**16**	**14**	**13**	**15**
12	**14**	**13**	**14**	**13**	**16**	**14**	**15**

Veamos la tabla correspondiente

Tabla 2.15. Edad de un grupo de estudiantes de grado noveno

Edad (años) (x_i)	**Nº de estudiantes** (f_i)
12	4
13	8
14	14
15	10
16	4
TOTAL	40

Calcular la media aritmética correspondiente a la edad de los estudiantes, implica sumar los 40 datos analizados y dividir en 40. Esto se puede abreviar multiplicando cada variable por el número de veces que se repite (frecuencia absoluta). Para esto podemos agregar una columna a la derecha de la tabla con este producto.

Tabla 2.16. Edad de un grupo de estudiantes grado 9

Edad (años) (x_i)	**Nº de estudiantes** (f_i)	$x_i f_i$
12	4	$12 \times 4 = 48$
13	8	$13 \times 8 = 104$
14	14	$14 \times 14 = 196$
15	10	$15 \times 10 = 150$
16	4	$16 \times 4 = 64$
TOTAL	40	562

Sumamos la columna $x_i f_i$ y la dividimos en el total de datos

$$\bar{x} = \frac{562}{40} = 14.05$$

Por consiguiente, la edad promedio que representa este grupo de estudiantes es 14.5 años.
En general si conocemos la frecuencia absoluta de cada dato, podemos calcular la media aritmética usando la expresión:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

En donde x_i representa los valores que toma la variable y f_i la frecuencia que tiene cada uno de estos valores x_i.

¿Cómo podemos calcular la media aritmética cuando tenemos una tabla organizada por intervalos?

Veamos la siguiente tabla:

Tabla 2.17 Edad de un grupo de personas

Edad (años) ($\boldsymbol{x_i}$)	**Nº de estudiantes** ($\boldsymbol{f_i}$)
[10; 12)	4
[12; 14)	8
[14; 16)	14
[16; 18)	10
[18; 20)	4
TOTAL	40

Para identificar cada intervalo, es necesario buscar un dato que represente todos los valores que se encuentran en él, dicho dato se conoce como ***marca de clase*** ($\boldsymbol{MC}$). La $\boldsymbol{MC}$ se define como el punto medio de cada uno de los intervalos, y se calcula conociendo los extremos del intervalo

Supongamos que tenemos el intervalo **[10; 12)**, la $\boldsymbol{MC}$ correspondiente es:

$$\boldsymbol{MC} = \frac{extremo\ inferior + extremo\ superior}{2} = \frac{10+12}{2} = 11$$

Las demás $\boldsymbol{MC}$ se calculan de la misma forma

Tabla 2.18. Edad de un grupo de personas

Edad (años) ($\boldsymbol{x_i}$)	$\boldsymbol{MC}$	**Nº de estudiantes** ($\boldsymbol{f_i}$)
[10; 12)	11	4
[12; 14)	13	8
[14; 16)	15	14
[16; 18)	17	10
[18; 20)	19	4
TOTAL	/	40

Para calcular la media aritmética en estos casos, se debe usar el dato representativo de cada intervalo, es decir la $\boldsymbol{MC}$, y realizar el mismo proceso usado en los primeros ejemplos, sólo que se emplea la $\boldsymbol{MC}$ en lugar de la variable $\boldsymbol{x}_i$, esto es,

$$\bar{x} = \frac{\sum(MC)f_i}{\sum f_i}$$

Por consiguiente, el valor de la media aritmética de las edades es:

$$\bar{x} = \frac{11(4)+13(8)+15(14)+17(10)+19(4)}{40} = \frac{604}{40} = 15.1\,,$$

es decir la edad promedio de este grupo de estudiantes es 15.1 años.

A pesar de que la media aritmética es una de las medidas de tendencia central más empleada, por usar todos los datos para su cálculo, no siempre proporciona información acertada. Veamos el siguiente caso

Un seleccionador de patinaje debe decidir entre dos corredores para representar su equipo en la siguiente competencia. Los tiempos registrados por los patinadores en sus últimas seis carreras fueron:

	Carrera 1	**Carrera 2**	**Carrera 3**	**Carrera 4**	**Carrera 5**	**Carrera 6**	**Carrera 7**
Patinador 1	12.2 s	12.3 s	16.1 s	12.2 s	12.2 s	12.2 s	12.4 s
Patinador 2	12.3 s	12.5 s	12.5 s	12.4 s	12.6 s	12.7 s	12.5

El seleccionador toma la decisión usando como referencia la media aritmética. Los resultados de $\bar{x}$ para cada participante son:

Patinador 1	Patinador 2
$\bar{x} = 12.85$	$\bar{x} = 12.5$

Por lo que el entrenador decide poner en la competencia el patinador 2.

¿Cambiaría la decisión del seleccionador, si se entera que en el comienzo de la carrera 3 el patinador 1 sufrió una caída?

Si analizamos carrera por carrera nos podemos dar cuenta, que el patinador 1 gano todas las carreras excepto la carrera 3, por consiguiente podríamos decir que es más conveniente que el patinador 1 participe en la siguiente competencia. En circunstancias como esta, resulta conveniente usar otra medida de tendencia central, puesto que la media aritmética es sensible a datos que estén separados del grupo, es decir extremadamente grandes o extremadamente pequeños respecto a los demás datos, como es el caso del tiempo que corresponde al patinador 1 en la carrera 3. Resultaría útil tomar un valor que sea intermedio a todos los datos. Este valor recibe el nombre de mediana ($\boldsymbol{Me}$).

Para calcular la mediana ($\boldsymbol{Me}$), lo primero que debemos hacer es ordenar los datos de menor a mayor. Si tenemos una cantidad impar de datos, debemos elegir el valor intermedio que separa en dos partes iguales el total de datos, en caso de tener una cantidad par de datos, elegimos los dos valores intermedios y calculamos la media aritmética entre estos.

Veamos cómo debemos calcular la mediana ($\boldsymbol{Me}$) a los dos patinadores

Patinador 1	12.2 s	12.3 s	16.1 s	12.2 s	12.2 s	12.2 s	12.4 s

Ordenamos los datos de menor a mayor y seleccionamos el valor que se encuentra en la mitad

$$12.2 \;\; 12.2 \;\; 12.2 \;\; \boxed{12.2} \;\; 12.3 \;\; 12.4 \;\; 16.1$$

El tiempo del patinador 1 que se encuentra en la centro corresponde a 12.2 s, $Me = 12.2$

Veamos que ocurre con el patinador 2

Patinador 2	12.3 s	12.5 s	12.5 s	12.4 s	12.6 s	12.7 s	12.5

$$12.3 \;\; 12.4 \;\; 12.5 \;\; \boxed{12.5} \;\; 12.5 \;\; 12.6 \;\; 12.7$$

El tiempo del patinador 2 que se encuentra en la centro corresponde a 12.5 s, $Me = 12.5$

Usando estos datos centrales ($\boldsymbol{Me}$), podemos ver que para el seleccionador sería conveniente elegir al patinador 1 para que participe en la siguiente competencia.

De manera general se tiene que:

Mediana ($\boldsymbol{Me}$): *Es el dato que divide en dos partes porcentualmente iguales el conjunto, es decir es el dato de la mitad.*

Ya mencionamos que para calcular la mediana, debemos ordenar los datos de menor a mayor y buscar el dato del centro, para esto usamos la expresión:

$$Me = D_t\left(\frac{n+1}{2}\right), D_t \; es \; la \; la \; abreviatura \; de \; dato$$

Veamos un ejemplo,
Consideremos las edades de 7 estudiantes

$$13, 18, 12, 15, 16, 17, 17.$$

Lo primero que debemos hacer es ordenar de menor a mayor los datos

$$12, 13, 15, 16, 17, 17, 18,$$

luego usando la expresión mencionada, buscamos el dato central

$$Me = D_t\left(\frac{n+1}{2}\right) = D_t\left(\frac{7+1}{2}\right) = D_t(4).$$

La mediana corresponde al dato número 4, ahora buscamos el dato que está ubicado en la posición cuatro

$$12, 13, 15, \boxed{16}, 17, 17, 18.$$

Por consiguiente la mediana es:

$$Me = 16.$$

Es decir el 50% de los estudiantes tiene 16 años o menos y el otro 50% tienen 16 años o más.
Cuando tenemos una cantidad impar de datos, el resultado de la mediana es inmediato, como ya lo vimos en el ejemplo anterior.
Veamos qué pasa cuando tenemos una cantidad par de datos.

Consideremos las estaturas de 10 estudiantes

$$162, 160, 155, 167, 170, 172, 158, 165, 155, 170.$$

Ordenamos los datos de menor a mayor

$$155, 155, 158, 160, 162, 165, 167, 170, 170, 172.$$

Usamos nuevamente la expresión

$$Me = D_t\left(\frac{n+1}{2}\right) = D_t\left(\frac{10+1}{2}\right) = D_t(5.5).$$

Esto quiere decir que la mediana se encuentra entre el dato 5 y el dato 6

$$155, 155, 158, 160, \boxed{162}, \boxed{165}, 167, 170, 170, 172.$$

Para determinar la mediana buscamos la media aritmética entre los datos 5 y 6

$$Me = \frac{162+165}{2} = 163.5$$

En conclusión el 50% de los estudiantes miden 163.5 cm o menos y el otro 50% miden 163.5 cm o más.

Al igual que en la media aritmética, cuando queremos calcular la mediana (Me) a un grupo grande de datos resulta conveniente primero organizarlos en una tabla de frecuencias en la que debemos incluir la frecuencia acumulada. Lo siguiente es buscar el dato central, que lo encontramos con la expresión $\left(\frac{n}{2}\right)$, por último nos deslizamos por la columna de la frecuencia acumulada ($\boldsymbol{F}$) hasta encontrar la primera frecuencia mayor o igual que $\left(\frac{n}{2}\right)$. El valor que toma la variable para esta frecuencia acumulada corresponde a la mediana, puesto que el dato central se encuentra allí.

Tabla 2.19. Edad de un grupo de estudiantes de grado noveno

Edad	f_i	F
12	4	4
13	8	12
14	14	26
15	10	36
16	4	40
Total	40	//

$$\frac{n}{2} = \frac{40}{2} = 20$$

La primera frecuencia acumulada mayor o igual que 20 es 26 y el valor que toma la variable para esta frecuencia es 14, por consiguiente

$$Me = 14$$

Otra medida que proporciona información rápida sobre los datos es la *Moda* ($\boldsymbol{Mo}$), ésta hace referencia al valor qué más se repite, es decir el valor que tiene una mayor frecuencia absoluta.

En la tabla, vemos que la mayor frecuencia es 14 y corresponde a la edad 14 años, por consiguiente

$$Mo = 14$$

"En el caso que todos los datos tengan la misma frecuencia, se dice que no hay moda."

2.5. MEDIDAS DE DISPERSIÓN

Veamos la siguiente información

Notas obtenidas en la clase de matemáticas

Roberto	**1.0**	**4.5**	**2.0**	**5.0**
Eduardo	**3.2**	**3.2**	**3.1**	**3.0**

¿Quién obtuvo mejor desempeño?

Ya vimos anteriormente que calculando la media aritmética podemos determinarlo

Roberto	Eduardo
$\bar{x} = 3.125$	$\bar{x} = 3.125$

¿Cómo podemos decidir en este caso?

La estadística nos proporciona herramientas útiles en este tipo de situaciones, estas reciben el nombre de medidas de dispersión o variación, son valores que nos proporcionan información sobre qué tan agrupados están los datos con respecto a las medidas de tendencia central, es decir si los datos están cercanos a la media, la dispersión es pequeña, si son lejanos a la media, la dispersión es grande, y si los datos son iguales a la media, significa que no hay dispersión. Las medidas de dispersión que calcularemos son la varianza (S^2) y la desviación estándar (S). El rango que ya calculamos en una sección anterior también es considerado como medida de dispersión.

La desviación de un dato es la distancia que hay entre el dato y la media, mediana, o moda. La podemos calcular mediante la expresión

$$D = x - \bar{x}$$

Las desviaciones positivas indican que el dato es mayor que la media, si la desviación es negativa significa que el dato es menor que la media.

Calculemos las desviaciones de las notas de Roberto

$$D = x - \bar{x}$$

$$D_1 = 1.0 - 3.125 = -2.125 \qquad D_2 = 4.5 - 3.125 = 1.375$$

$$D_3 = 2.0 - 3.125 = -1.125 \qquad D_4 = 5.0 - 3.125 = 1.875$$

Tenemos:

Nota	**1.0**	**4.5**	**2.0**	**5.0**
Desviación	-2.125	1.375	-1.125	1.875

Para calcular una desviación que represente todas las desviaciones, debemos hallar una media aritmética entre ellas, teniendo cuidado con los valores negativos, puesto que podemos encontrar desviaciones con la misma magnitud, pero con signo contrario, por ejemplo 1 y -1. La suma de estos valores es cero lo cual indica que no hay desviación, afirmación que es incorrecta. Por consiguiente para evitar valores negativos se eleva cada desviación al cuadrado, y luego se halla la media aritmética de los cuadrados de las desviaciones. Este cálculo nos da como resultado lo que llamaremos Varianza (S^2).

El cálculo de la varianza se realizará de la siguiente manera:

$$S^2 = \frac{D_1^2 + D_2^2 + D_3^2 + D_4^2}{4} = \frac{(-2.125)^2 + 1.375^2 + (-1.125)^2 + 1.875^2}{4}$$

$$S^2 = \frac{11.1875}{4} = 2.7968$$

La Varianza (S^2) de las notas obtenidas por Roberto en matemáticas es 2.7968

Ahora calculemos las desviaciones de las notas de Eduardo

$$D = x - \bar{x}$$

$$D_1 = 3.2 - 3.125 = 0.075$$

$$D_2 = 3.2 - 3.125 = 0.075$$

$$D_3 = 3.1 - 3.125 = -0.025$$

$$D_4 = 3.0 - 3.125 = 0.125$$

Tenemos:

Nota	**1.0**	**4.5**	**2.0**	**5.0**
Desviación	0.075	0.075	-0.025	0.125

Por consiguiente,

$$S^2 = \frac{D_1^2 + D_2^2 + D_3^2 + D_4^2}{4} = \frac{0.075^2 + 0.075^2 + (-0.025)^2 + 1.125^2}{4}$$

$$S^2 = \frac{1.2775}{4} = 0.3193$$

La Varianza (S^2) de las notas obtenidas por Eduardo en matemáticas es 0.3193.

En conclusión,

Notas Roberto	Notas Eduardo
$S^2 = 2.7968$	$S^2 = 0.3193$

Con el cálculo de la varianza vemos que las notas que tiene menos dispersión (tiene menor varianza) son las de Eduardo y por esta razón mejor rendimiento en el área de matemáticas.

De manera general se tiene que

Varianza (S^2): Es la media aritmética de los cuadrados de las deviaciones.

Podemos calcular la varianza en datos no agrupados mediante la expresión:

$$S^2 = \frac{\sum(x_i - \bar{x})^2}{n}.$$

En el caso de tener datos agrupados, usamos la expresión:

$$S^2 = \frac{\sum f_i(x_i - \bar{x})^2}{\sum f_i}$$

"Se recomienda incluir en la tabla de frecuencia dos columnas en las que calculemos $(x - \bar{x})^2 \ y \ f(x - \bar{x})^2$, esto para facilitar el cálculo de la varianza"

Ejemplo
Calculemos la varianza de los datos contenidos en la tabla

Tabla 2.20. Edades de un grupo de estudiantes

x_i	f_i
13	7
14	9
15	5
16	7
17	9
18	8
19	5
Total	50

Igual que en la desviación media, lo primero que hallamos es la media aritmética.

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

$$\bar{x} = \frac{796}{50} = 15.92$$

Ahora completemos la tabla incluyendo $(x_i - \bar{x})^2 \ y \ f_i(x_i - \bar{x})^2$

x_i	f_i	$(x_i - \bar{x})^2$	$f_i(x_i - \bar{x})^2$
13	7	8.5254	59.6848
14	9	3.6864	33.1776
15	5	0.8464	4.232
16	7	0.0064	0.0448
17	9	1.1664	10.4976
18	8	4.3264	34.6112
19	5	9.4864	47.432
Total	50	28.0438	189.68

De la tabla obtenemos:

$$\sum f_i = 50; \sum (x_i - \bar{x})^2 = 28.0438; \sum f_i (x_i - \bar{x})^2 = 189.68$$

Con esta información calculamos la varianza

$$s^2 = \frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i} = \frac{189.68}{50} = 3.7936.$$

Desviación estándar (S)

También llamada desviación típica, es la raíz cuadrada de la varianza, por tanto para datos no agrupados podemos calcularla usando la siguiente expresión:

$$S = \sqrt{\frac{\sum (x_i - \bar{x})^2}{\sum f_i}}$$

Para datos agrupados, usamos la expresión:

$$S = \sqrt{\frac{\sum f_i (x_i - \bar{x})^2}{\sum f_i}}$$

El proceso para el cálculo de la deviación estándar (S), es el mismo que el de la varianza (S^2), con la diferencia que al final del proceso debemos computar la raíz cuadrada.

Las varianzas calculadas para las notas de Roberto y Eduardo fueron:

Notas Roberto	Notas Eduardo
$S^2 = 2.7968$	$S^2 = 0.3193$

Por consiguiente las varianzas deben ser

Notas Roberto	Notas Eduardo
$S = \sqrt{S^2} = \sqrt{2.7968} = 1.6723$	$S = \sqrt{S^2} = \sqrt{0.3193} = 0.565$

Con el cálculo de la desviación estándar verificamos que las notas que tienen menor variación son las de Eduardo y por esta razón mejor rendimiento en el área de matemáticas.

En el caso de tener los datos agrupados, el proceso es el mismo del cálculo de la varianza y al final debemos calcular la raíz cuadrada.

La varianza de las edades del grupo de estudiantes fue:

$$S^2 = 3{,}7936$$

Por consiguiente su desviación estándar debe ser:

$$S = \sqrt{S^2} = \sqrt{3{,}7936} = 1{,}9477.$$

2.5.1. ACTIVIDADES

1. Se ha lanzado un dado 30 veces y se han obtenido los siguientes resultados:

   3 4 5 2 1 4 6 1 3 2
   5 5 3 2 4 4 1 2 5 6
   1 5 4 3 4 5 6 2 6 4

   a) Construya una tabla de frecuencias
   b) Calcule la media, la mediana y la moda.

2. Cristina juega baloncesto todos los fines de semana. En los últimos 5 juegos anoto 18, 11, 18, 23 y 15 puntos. ¿Cuál fue la media de su puntaje?

3. Un grupo de amigos están comparando el número de veces por mes que visitaron al médico el año pasado. La siguiente tabla presenta la información.

	Ene.	Feb.	Mar.	Abr.	Mayo	Junio	Julio	Agos.	Sept.	Oct.	Nov.	Dic.
Luis	1	3	2	5	2	3	1	4	2	3	2	1
Juliana	1	2	1	1	1	3	3	2	2	4	1	2
Sergio	1	3	2	2	1	4	5	3	2	2	1	3
Laura	2	2	1	1	3	2	4	1	3	2	3	2

   a) Según la media aritmética ¿qué persona fue más al médico por mes?
   b) Comparando sus medias, ordene a los amigos, de mayor a menor número de visitas al médico.
   c) Según la media aritmética ¿cuál es el mes con menos visitas al médico?
   d) ¿Cuál es la mediana correspondiente a las visitas de cada mes?
   e) Según la mediana ¿cuál es el mes con menos visitas al médico?

4. Los siguientes son resultados de exámenes de Química aplicados a 2 estudiantes. ¿Cuál de los dos estudiantes tuvo mejor rendimiento?

	Oscar	Andrés
Examen 1	4.0	3.2
Examen 2	1.0	3.0

Examen 3	3.0	2.8
Examen 4	2.5	3.0
Examen 5	4.5	3.0

5. Las calificaciones obtenidas por los 40 estudiantes de grado 8º en una prueba de Matemáticas vienen dadas por la siguiente tabla:

Nota	2	3	4	5	6	7	8	9	10
Alumnos	1	2	6	8	6	6	6	4	1

   a) Calcula, la nota media, la mediana y la moda.
   b) Calcula, la desviación media de las calificaciones, la varianza y la desviación estándar.

6. La siguiente gráfica representa los incrementos en el salario mínimo en Colombia en los últimos años. ¿Cuál es el incremento medio anual de 2003 a 2012?

7. Las siguientes son las edades de un grupo de habitantes en un conjunto residencial

42 60 60 38 60 63 21 66 56 57 51 57 44 45 35
30 35 47 53 49 50 49 38 45 28 41 47 42 53 32
54 38 40 63 48 33 35 61 47 41 55 53 27 20 21
42 21 39 39 34 45 39 28 54 33 35 43 48 48 27
53 30 29 53 38 52 54 27 27 43 28 63 41 23 58
56 59 60 40 24 20 58 35 62 27 30 39 40 52 60

   a) Calcule la edad media, la mediana y la moda
   b) Calcule la desviación media entre las edades.
   c) Calcule la varianza y desviación estándar de las edades.

3.PROBABILIDAD

3.1. Experimento Aleatorio, Espacio Muestral, Evento

Es muy común encontrarnos con situaciones cotidianas imprevisibles; como por ejemplo, saber si nuestro equipo de fútbol favorito ganará el próximo partido, saber cómo será el clima del día en que tenemos planeado un paseo, saber cuál cara de un dado quedará arriba después de lanzarlo. Este tipo de experiencias son llamadas experimentos aleatorios, en ellas está presente el azar, y su resultado es incierto, es decir no lo conoceremos sino después de realizar dicha experiencia. Lanzar una moneda corriente, escoger el número ganador en una lotería, sacar una bola de una urna que tiene bolas de distintos colores sin ver su interior, todos estos son experimentos aleatorios.

Experimento aleatorio: es aquel que realizándolo bajo las mismas condiciones iniciales, puede tener diferentes resultados. Este resultado sólo se conoce después de realizar el experimento.

Experimento Determinista: Es aquel para el que realizándolo bajo las mismas condiciones iniciales, siempre se obtiene el mismo resultado.

A pesar de que en los experimentos aleatorios no podemos saber qué resultado obtendremos, si tenemos una serie de posibles resultados. Por ejemplo cuando lanzamos un dado tenemos seis posibles resultados, cuando lanzamos una moneda corriente tenemos dos. Todos los posibles resultados de un experimento aleatorio conforman lo que llamaremos espacio muestral.

Espacio muestral: es el conjunto compuesto por todos los posibles resultados que pueda tener el experimento. En adelante el espacio muestral lo representaremos con la letra Ω.

Ejemplos:

Si el experimento consiste en lanzar una moneda corriente y ver cuál lado queda en la parte superior, el espacio muestral será:

$$\Omega = \{cara, sello\}$$

Si el experimento consiste en lanzar un dado y ver cuál número sale en la cara superior, el espacio muestral correspondiente es:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Si el experimento consiste en lanzar dos monedas y observar cuáles lados quedan arriba, el espacio muestral es:

$$\Omega = \{(cara, cara), (cara, sello), (sello, cara), (sello, sello)\}$$

Si el experimento consiste en sacar una carta de una baraja española, sin tener en cuenta la pinta. El espacio muestral será:

$$\Omega = \{as, dos, tres, cuatro, cinco, seis, siete, sota, caballo, rey\}$$

Algunas veces podemos no estar interesados en conocer todos y cada uno de los posibles resultados del experimento aleatorio, sino sólo en obtener alguna información en particular relacionada con este, por ejemplo si lanzamos una moneda corriente al aire tres veces seguidas, podemos estar interesados en determinar el conjunto de todos aquellos resultados en los que se obtuvo 2 o más caras en los tres lanzamientos. Esto es, nos interesa

$$A = \{(C, C, C), (C, C, S), (C, S, C), (S, C, C)\}$$

Los resultados obtenidos en estos hechos conforman un conjunto, que llamaremos evento.

Evento o suceso: es un subconjunto del espacio muestral. Se representa usualmente con las primeras letras mayúsculas del abecedario A, B, C,…

En el lanzamiento de un dado vemos que el espacio muestral es:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Algunos ejemplos de evento de este espacio muestral son:

$$B = "Salir\ un\ número\ par" = \{2, 4, 6\}$$
$$C = "Salir\ un\ número\ impar" = \{1, 3, 5\}$$
$$D = "Salir\ un\ número\ menor\ que\ 3" = \{1,2\}$$

Hay algunos eventos que tienen un nombre en particular.
Si pensamos nuevamente en el lanzamiento del dado, y definimos un evento E, como obtener el número 7, sabemos que este resultado no es posible de obtener.
Este evento ciertamente no tiene elementos. Todos los eventos que tengan esta característica reciben el nombre de "*Eventos Imposibles" o "eventos vacíos"*

$$E = \emptyset = \{\quad\}$$

Ahora si en el mismo lanzamiento, definimos el evento F obtener un número menor que siete, es claro que $F = \{1,2,3,4,5,6\}$este evento tiene los mismos elementos que el espacio muestral. Cuando ocurre esto, este evento es llamado *"Evento Seguro".*
Si el evento tiene únicamente un elemento, se llama *"evento unitario o simple"*, por ejemplo el evento G, obtener el número 5 al lanzar un dado.

$$G = \{5\}$$

3.1.1 Intersección de Eventos

Si tenemos dos eventos A, B, la intersección de ellos $(A \cap B)$, es el conjunto cuyos elementos pertenecen al evento A y al evento B simultáneamente.

Ejemplo:

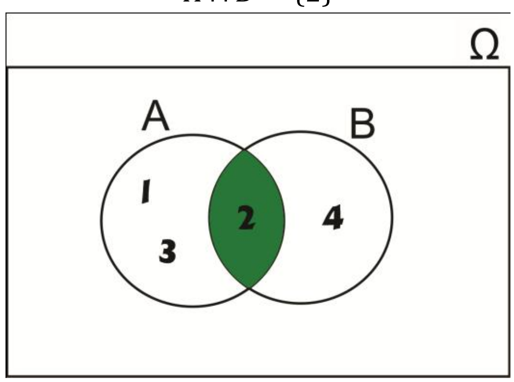
Se lanza un dado y se definen los eventos:

$$A = \text{"}Obtener\ un\ número\ menor\ que\ 4\text{"} = \{1, 2, 3\}$$

$$B = \text{"}Obtener\ un\ número\ par\text{"} = \{2, 4, 6\}$$

La intersección de los eventos está constituida por todos aquellos números que son pares y menores que 4, por consiguiente

$$A \cap B = \{2\}$$



$\boldsymbol{A \cap B}$

3.1.2. Unión de Eventos

Si tenemos dos eventos A, B, la unión de los dos $(A \cup B)$, es el conjunto cuyos elementos pertenecen al evento A, o al evento B, o a los dos.

Ejemplo:

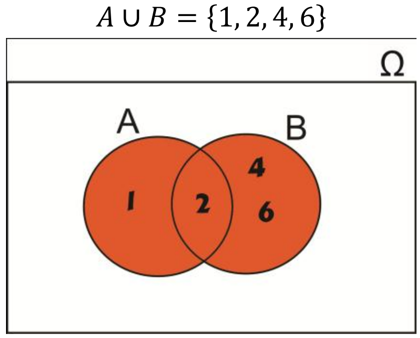
Se lanza un dado y se definen los eventos

$$A = \text{"}obtener\ un\ número\ menor\ que\ 3\text{"} = \{1,2\}$$

$$B = \text{"}obtener\ un\ número\ múltiplo\ de\ 2\text{"} = \{2, 4, 6\}$$

La unión de los eventos será:

$$A \cup B = \{1, 2, 4, 6\}$$



$\boldsymbol{A \cup B}$

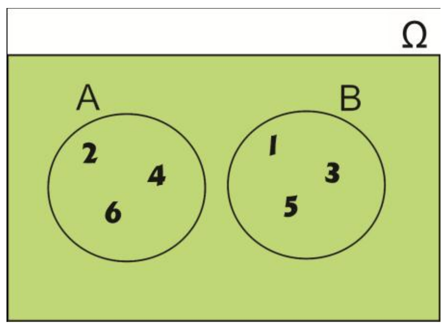
3.1.3. Eventos Disjuntos o mutuamente excluyentes

Dos eventos son A y B son disjuntos si no pueden suceder simultáneamente, es decir si la intersección entre los eventos A y B es vacía ($A \cap B = \emptyset$).

Ejemplo:
Se lanza un dado, y se definen los eventos

$$A = "obtener\ un\ número\ par" = \{2, 4, 6\}$$
$$B = \text{obtener un número impar} = \{1, 3, 5\}.$$

Estos dos eventos no tienen elementos en común, su intersección es vacía ($A \cap B = \emptyset$), por consiguiente estos eventos son mutuamente excluyentes, es decir no pueden suceder el mismo tiempo.



$\boldsymbol{A \cap B = \emptyset}$

3.1.4. Diferencia de Eventos

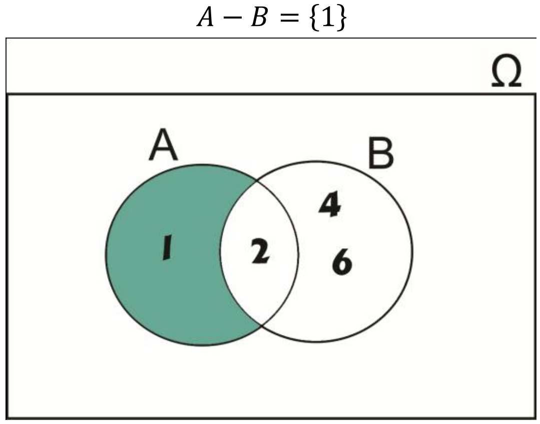
Si tomamos dos eventos A y B, llamaremos diferencia de eventos $(A - B)$ a los elementos que pertenecen al evento A, pero que no pertenece al evento B.

Ejemplo:
Se lanza un dado y se definen los eventos

$$A = "obtener\ un\ número\ menor\ que\ 3" = \{1,2\}$$
$$B = "obtener\ un\ número\ par" = \{2, 4, 6\}$$

La diferencia de eventos será:

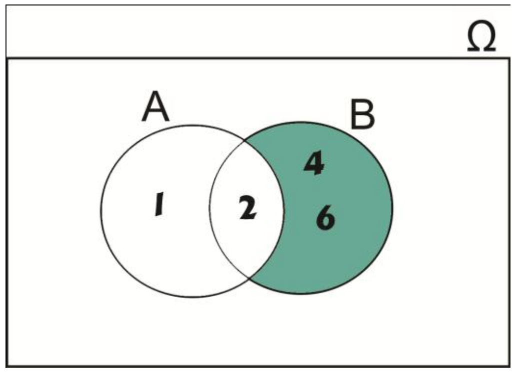
$$A - B = \{1\}$$



$\boldsymbol{A - B}$

Cabe anotar que esta operación no es conmutativa, es decir $A - B \neq B - A$. Por ejemplo para la situación descrita en este ejemplo se tiene

$$B - A = \{4, 6\}$$



$\boldsymbol{B - A}$

3.1.5. Complemento de los Eventos

Si tenemos un evento A, el complemento de A se representa con los símbolos A^c, A', o, $\bar{A}$. Este conjunto está constituido por los elementos que le faltan al evento A para convertirse en el espacio muestral, es decir todos los elementos que están fuera de A.

Ejemplo:

Se lanza un dado, y se define el evento:

$$A = \text{"}obtener\ un\ número\ par\text{"} = \{2, 4, 6\}$$

Ya sabemos que el espacio muestral para este experimento aleatorio es

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

Luego los elementos de A^c, serán los elementos del espacio muestral que no están en A.

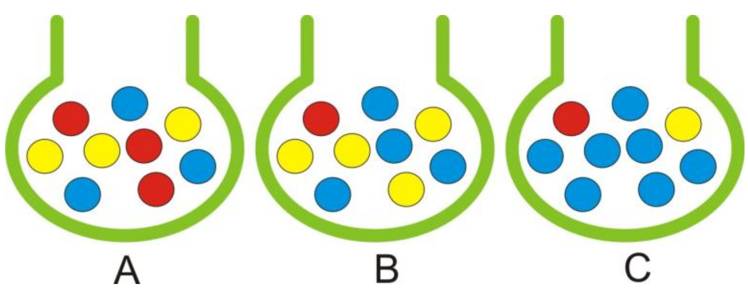
$$A^c = \{1, 3, 5\}$$

3.2. Definición de Probabilidad

Supóngase que se tienen tres urnas, como se muestra en la gráfica 3.1.

Juan saca una bola al azar de la primera urna, Pedro saca una bola a azar de la segunda urna y María una de la tercera.

¿Si se ofrece un premio al que saque una balota azul, quien tiene más chance de ganar el premio?



Gráfica 3.1. Urnas

Por simple intuición, podemos afirmar que existe más posibilidad de sacar una balota azul de la urna **C**. Podemos afirmar que el chance de sacar la balota azul depende del número de balotas azules, respecto al total de balotas, es decir entre más balotas azules haya en proporción al total de balotas, hay más chance de sacar una balota de este color. Por consiguiente podemos establecer una relación entre el número de posibilidades de que ocurra el evento, (sacar una balota azul) y el total de posibles resultados (sacar una balota de cualquier color).

Ahora veamos la siguiente situación. Lanzamos una moneda corriente 50, 100, 150, 200, 250, 300, 350, 400, 450 y 500 veces y registramos el número de veces que aparece "cara" y el número de veces que aparece "sello". Obteniéndose los siguientes resultados:

Eventos	**Número de lanzamientos**									
	50	100	150	200	250	300	350	400	450	500
Salir cara	19	42	72	115	138	146	176	201	226	254
Proporción	$\frac{19}{50} = 0.38$	$\frac{42}{100} = 0.42$	$\frac{72}{150} = 0.48$	$\frac{115}{200} = 0.575$	$\frac{138}{250} = 0.552$	$\frac{146}{300} = 0.49$	$\frac{176}{350} = 0.504$	$\frac{201}{400} = 0.502$	$\frac{226}{450} = 0.502$	$\frac{254}{500} = 0.508$
Salir sello	31	58	78	85	112	154	174	199	224	246
Proporción	$\frac{31}{50} = 0.62$	$\frac{58}{100} = 0.58$	$\frac{78}{150} = 0.52$	$\frac{85}{200} = 0.425$	$\frac{112}{250} = 0.448$	$\frac{154}{300} = 0.51$	$\frac{174}{350} = 0.496$	$\frac{199}{400} = 0.498$	$\frac{224}{450} = 0.498$	$\frac{246}{500} = 0.492$

Tabla 3.1. Proporción de caras obtenidas en el lanzamiento de una moneda

Obsérvese que la proporción de caras y sellos obtenidos, corresponde a la definición de frecuencia relativa dada en el capítulo anterior.

Es decir

$$f_r(\text{"}cara\text{"}) = \frac{N^{\underline{o}}\ de\ caras\ obtenidas}{N^{\underline{o}}\ de\ lanzamientos}$$

$$f_r(\text{"}sello\text{"}) = \frac{N^{\underline{o}}\ de\ sellos\ obtenidos}{N^{\underline{o}}\ de\ lanzamientos}$$

Veamos gráficamente las proporciones, en relación al número de lanzamientos

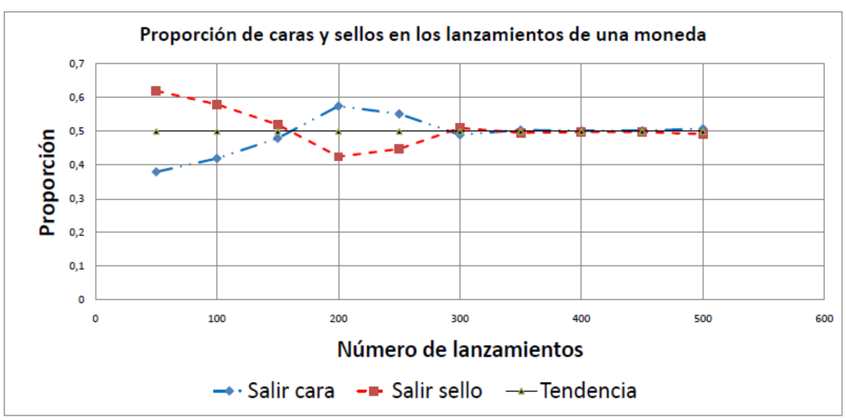


Figura 3.2.

En la gráfica podemos ver que cuando el número de lanzamientos se aumenta, la frecuencia tiende a estabilizarse alrededor de $^1/_2$.

Veamos una situación similar. Lanzamos un dado corriente 40, 80, 120, 160, 200, 240, 280, 320, 360 y 400 veces y registramos el número de veces que aparece cada cara del dado (1, 2, 3, 4, 5 y 6). Veamos los resultados:

Eventos	**Número de lanzamientos**									
	40	80	120	160	200	240	280	320	360	400
Salir un 1	6	12	24	30	36	41	55	55	65	60
Proporción	$\frac{6}{40} = 0.15$	$\frac{12}{80} = 0.15$	$\frac{24}{120} = 0.2$	$\frac{30}{160} = 0.018$	$\frac{36}{200} = 0.18$	$\frac{41}{240} = 0.17$	$\frac{55}{280} = 0.196$	$\frac{55}{320} = 0.171$	$\frac{65}{360} = 0.18$	$\frac{60}{400} = 0.15$
Salir un 2	9	13	12	17	30	48	47	54	51	66
Proporción	$\frac{9}{40} = 0.225$	$\frac{13}{80} = 0.162$	$\frac{12}{120} = 0.1$	$\frac{17}{160} = 0.106$	$\frac{30}{200} = 0.15$	$\frac{48}{240} = 0.2$	$\frac{47}{280} = 0.167$	$\frac{54}{320} = 0.168$	$\frac{51}{360} = 0.141$	$\frac{66}{400} = 0.165$
Salir un 3	8	20	19	22	42	38	43	66	65	69
Proporción	$\frac{8}{40} = 0.2$	$\frac{20}{80} = 0.25$	$\frac{19}{120} = 0.159$	$\frac{22}{160} = 0.137$	$\frac{42}{200} = 0.21$	$\frac{38}{240} = 0.158$	$\frac{43}{280} = 0.153$	$\frac{66}{320} = 0.206$	$\frac{65}{360} = 0.18$	$\frac{69}{400} = 0.172$
Salir un 4	3	14	18	34	31	32	51	43	62	71
Proporción	$\frac{3}{40} = 0.075$	$\frac{14}{80} = 0.175$	$\frac{18}{120} = 0.15$	$\frac{34}{160} = 0.215$	$\frac{31}{200} = 0.155$	$\frac{32}{240} = 0.133$	$\frac{51}{280} = 0.182$	$\frac{43}{320} = 0.134$	$\frac{62}{360} = 0.172$	$\frac{71}{400} = 0.177$
Salir un 5	7	12	21	32	34	44	43	49	58	71
Proporción	$\frac{7}{40} = 0.175$	$\frac{12}{80} = 0.15$	$\frac{21}{120} = 0.175$	$\frac{32}{160} = 0.2$	$\frac{34}{200} = 0.17$	$\frac{44}{240} = 0.183$	$\frac{43}{280} = 0.153$	$\frac{49}{320} = 0.153$	$\frac{58}{360} = 0.161$	$\frac{71}{400} = 0.177$
Salir un 6	7	9	26	25	27	37	41	53	59	63
Proporción	$\frac{7}{40} = 0.175$	$\frac{9}{80} = 0.112$	$\frac{26}{120} = 0.216$	$\frac{25}{160} = 0.156$	$\frac{27}{200} = 0.135$	$\frac{37}{240} = 0.154$	$\frac{41}{280} = 0.146$	$\frac{53}{320} = 0.165$	$\frac{59}{360} = 0.163$	$\frac{63}{400} = 0.157$

Tabla 3.2. Proporción de lados en el lanzamiento de un dado

Nuevamente la proporción de caras obtenidas, corresponde a la definición de frecuencia relativa dada en el capítulo anterior. Por ejemplo vemos que:

$$f_r(\text{"}cinco\text{"}) = \frac{Nº\ de\ cinco\ obtenidos}{Nº\ de\ lanzamientos}$$

Veamos gráficamente las proporciones, en relación al número de lanzamientos

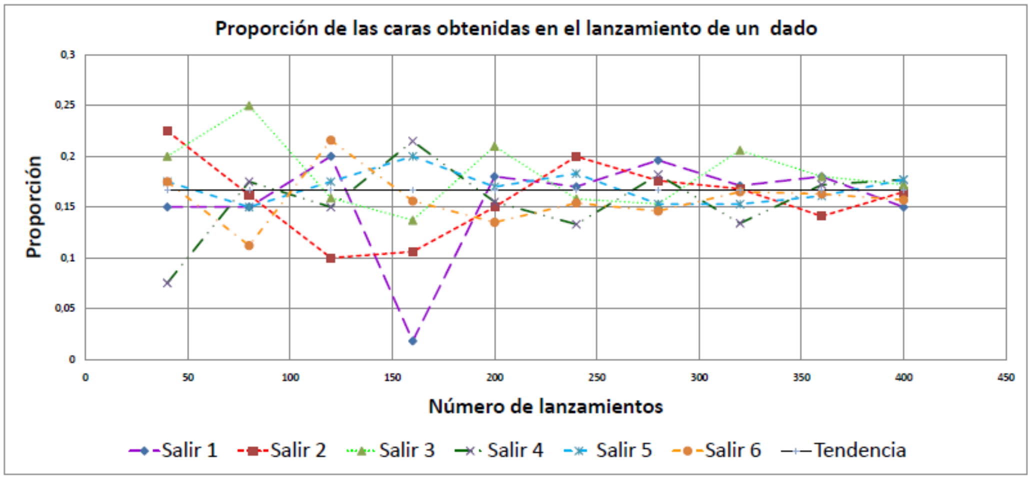


Figura 3.3.

En la gráfica podemos ver que cuando el número de lanzamientos se aumenta, la frecuencia tiende a estabilizarse alrededor de $1/_6$.

Podríamos pensar en definir el chance que tiene un evento de ocurrir como la frecuencia relativa del evento, pero esto tiene el inconveniente que el valor dependería del número de veces que se repita el experimento. Sin embargo, se observa que cuando el experimento se repite en condiciones similares, un número "grande" de veces esa frecuencia relativa tiende a estabilizarse alrededor de un número entre 0 y 1, el cual llamaremos probabilidad del evento A ($P(A)$).

Observemos que si A y B son eventos mutuamente excluyentes entonces

$$f_r(A \cup B) = f_r(A) + f_r(B)$$

en consecuencia

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

Supongamos que nuestro experimento es de tipo laplaciano, esto es, tiene solo un número finito de posibles resultados y que todos los resultados tienen el mismo chance de ocurrir, es decir se asume que el "chance" de un evento simple es igual a $\frac{1}{n}$ siendo n el número de posibles resultados del experimento.

Observemos que en ese caso, si A es un evento con m elementos, digamos

$$A = \{w_1, w_2, \ldots, w_m\}$$

entonces

$$P(A) = (\{w_1\} \cup \{w_2\} \cup \ldots \cup \{w_m\}) = P(\{w_1\}) + P(\{w_2\}) + \cdots + P(\{w_m\})$$

$$= \underbrace{\frac{1}{n} + \frac{1}{n} + \cdots + \frac{1}{n}}_{m\ veces} = \frac{m}{n} = \frac{\#\ de\ elementos\ de\ A}{\#\ de\ elementos\ de\ \Omega}$$

Si lo vemos de otra manera el número de elementos de A se refiere al número de casos favorables, y el número de elementos de Ω se refiere al número de casos posibles, por consiguiente

$$P(A) = \frac{número\ de\ casos\ favorables}{número\ de\ casos\ posibles}$$

Esta definición, es conocida como probabilidad de Laplace, y en ella es fundamental la condición de que los eventos unitarios son equiprobables.

Ejemplo:

Queremos saber cuál es la probabilidad de obtener un número par cuando lanzamos un dado.

Sabemos que el espacio muestral para este experimento es:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

El evento es

$$A = "obtener\ un\ número\ par" = \{2, 4, 6\}$$

El espacio muestral tiene 6 elementos

$$\#\Omega = 6$$

El evento tiene 3 elementos

$$\#A = 3$$

La probabilidad de que ocurra el evento A *es*

$$P(A) = \frac{\#A}{\#\Omega} = \frac{3}{6} = \frac{1}{2}$$

Cuando lanzamos un dado tenemos un 50% de posibilidad de obtener un número par.

Ejemplo, En una encuesta realizada a un grupo de estudiantes, se les pidió que eligieran un deporte favorito cada uno, los resultados se presentan en la siguiente tabla.

	NIÑOS	NIÑAS
BALONCESTO	5	12
FUTBOL	10	2
VOLEIBOL	3	4

Si elegimos al azar un estudiante, algunas de las probabilidades que podemos calcular son:

La probabilidad de elegir un niño

$$P(niño) = \frac{18}{36}$$

La probabilidad de elegir un estudiante que prefiera el voleibol

$$P(Voleibol) = \frac{7}{36}$$

La probabilidad de elegir una niña que le guste el fútbol

$$P(niña\ \ fútbol) = \frac{2}{36}$$

Para realizar el cálculo de algunas probabilidades, es importante conocer las propiedades de la probabilidad de ocurrencia de un evento.

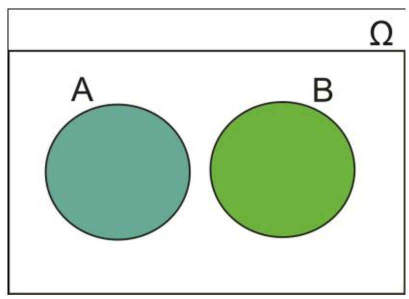
Dado un evento A, la probabilidad de ocurrencia es siempre mayor o igual a cero y menor e igual que uno, es decir la probabilidad de un evento está entre 0 y 1.

$$0 \leq P(A) \leq 1$$

Si el evento A es el mismo espacio muestral (Ω), la probabilidad de ocurrencia es 1, es decir es un evento seguro

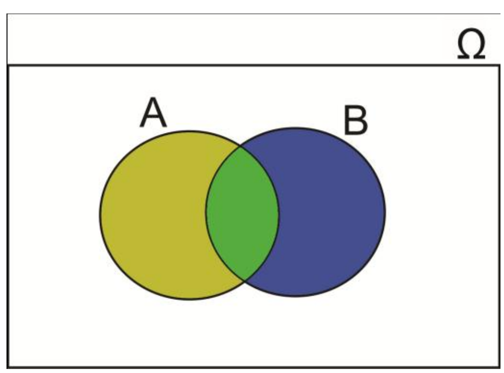
$$P(\Omega) = 1.$$

Si tenemos dos eventos A y B disjuntos ($A \cap B = \emptyset$), la probabilidad de su unión es igual a la suma de sus probabilidades.



$$P(A \cup B) = P(A) + P(B), si\ A \cap B = \emptyset$$

Si tenemos dos eventos A y B que se intersecan ($A \cap B \neq \emptyset$), entonces se cumple que la probabilidad de su unión es igual a la suma de sus probabilidades menos la probabilidad de su intersección.



$$P(A \cup B) = P(A) + P(B) - P(A \cap B), si\ A \cap B \neq \emptyset$$

"Para el caso de tener tres conjuntos que se intersequen se cumple:

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

Si tenemos un evento A y su complemento A^c si los unimos tenemos el espacio muestral (Ω)

$$A \cup A^c = \Omega,$$

Por lo tanto la probabilidad de dos eventos complementarios suma 1

$$P(A) + P(A^c) = 1.$$

De esto podemos decir que:

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

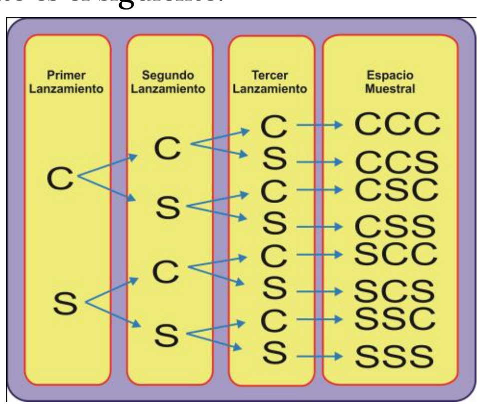
Como ya vimos, si queremos aplicar la definición de probabilidad, es necesario contar los elementos de cada evento y de Ω. En algunos casos es fácil hacer una lista y por consiguiente es fácil contarlos; pero es muy común encontrar situaciones en las que es complicado hacer un listado de todos y cada uno de los posibles elementos del evento o del espacio muestral. Por esta razón abordaremos algunas técnicas que permiten contar los elementos que contiene cada conjunto. Estos métodos reciben el nombre de técnicas de conteo.

Una manera de contar los elementos de un evento es usando los *diagramas de árbol.* Para la construcción de un diagrama de árbol se empieza poniendo una rama para cada uno de los valores iniciales. Estas ramas reciben el nombre de ramas de primera generación. Una vez construidas las ramas de primera generación se construye un nudo del cual parten nuevas ramas. Este proceso se repite sobre las nuevas ramas hasta ubicar todos los valores que aparecen en el experimento.

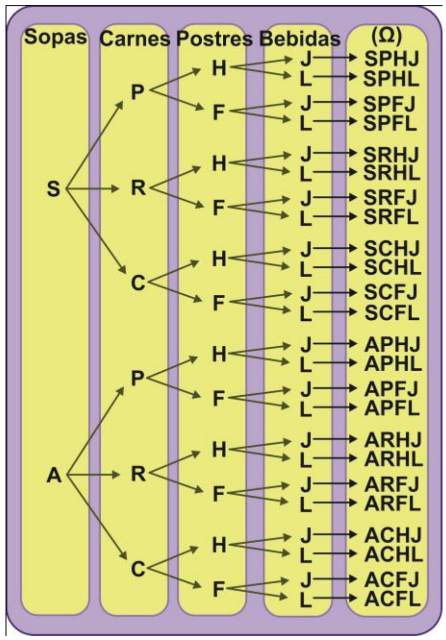
Ejemplo:

Lanzar una moneda tres veces y apuntar los datos obtenidos (cara(C), sello(S)): Para construir el diagrama de árbol verificamos los posibles casos en cada lanzamiento.

El diagrama correspondiente es el siguiente:



Ejemplo:
En un restaurante a la hora del almuerzo, ofrecen dos tipos de sopa: sancocho (S), y ajiaco (A); tres tipos de carne: pollo (P), res (R) y cerdo (C); dos postres: helado (H) y flan (F), y dos bebidas: jugo de fresa (J) y limonada (L). ¿Cuántos almuerzos completos puedo formar con estos componentes?
Para nuestro diagrama de árbol podemos empezar con las sopas, las ramas corresponden a las carnes, los postres y las bebidas.
El árbol correspondiente es:



En consecuencia podemos formar 24 almuerzos completos con estas opciones.

Obsérvese que este valor se obtiene multiplicando el número de sopas, el número de carnes, el número de postres y el número de bebidas, es decir

$$2 \times 3 \times 2 \times 2 = 24$$

El resultado obtenido se basa en el llamado *"Principio Fundamental del Conteo (PFC)", el cual establece que si un evento* A *ocurre de* n *maneras y un evento* B *ocurre de* m *maneras, entonces el número total de formas diferentes en que ambos eventos pueden ocurrir es igual a* $m \times n$ *maneras*

Ejemplo,
Una lotería colombiana tiene un sistema el cual juega con cuatro números y un signo zodiacal. ¿Cuántos resultados se pueden generar en este sorteo?

Para cada número tenemos 10 posibilidades $\{0,1,2,3,4,5,6,7,8,9\}$ y para el signo zodiacal tenemos 12 posibilidades,

$$\left\{\begin{matrix} aries, tauro, géminis, cáncer, leo, virgo, escorpión, sagitario, \\ capricornio, acuario, piscis, libra \end{matrix}\right\}$$

por consiguiente el total de sucesos son:

$$10 \times 10 \times 10 \times 10 \times 12 = 120.000$$

Ejemplo,

Un estudiante tiene un libro de Religión, uno de Matemáticas, uno de Estadística, uno de Física y uno de Química. ¿De cuántas maneras puede organizar los libros en una repisa?

Inicialmente el estudiante tiene 5 libros para ubicar en la repisa, después de ubicar el primer libro le quedan 4 libros para ubicar uno en la segunda posición y así sucesivamente, hasta que solamente le queda un libro, por consiguiente el número de posibilidades de ordenamiento será:

$$5 \times 4 \times 3 \times 2 \times 1 = 120\ formas\ de\ ordenar\ los\ libros.$$

El proceso anterior es llamado *permutación*, y corresponde a las maneras que se tiene de ordenar los elementos de un conjunto. Por ejemplo si tenemos el conjunto $\{a,b,c\}$, las formas en las que podemos ordenar los tres elementos son:

$$abc, acb, bac, bca, cab, cba$$

De manera general podemos representar el número total de permutaciones de n elementos diferentes (P_n) como:

$$(P_n) = n!$$

En donde $n!$ se lee n factorial y se define como el producto de los números menores e iguales a n.

$$n! = n(n-1)(n-2)(n-3)\cdots 3 \cdot 2 \cdot 1$$

Y definiremos

$$0! = 1$$

Por ejemplo

$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120$$
$$6! = 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 720$$

Ejemplo,

¿De cuántas maneras puedo organizar a 10 estudiantes en una fila?

$$P_{10} = 10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1 = 3'628.800$$

Hay 3'628.800 formas de organizar la fila.

Ahora supongamos que tenemos 10 libros y disponemos de una repisa donde sólo caben 5 libros ¿de cuántas maneras puedo ubicar los 5 libros en la repisa?

Inicialmente tenemos 10 posibilidades para ubicar en la repisa el primer libro, para el segundo lugar tenemos 9 libros para elegir, para el tercer lugar tenemos 8 libros, para el cuarto lugar tenemos 7 libros, y para el quinto y último lugar tendríamos 6 libros para elegir, por consiguiente usando el principio fundamental del conteo, las posibilidades de ordenamiento son:

$$10 \times 9 \times 8 \times 7 \times 6 = 30.240$$

De manera general, el número de permutaciones de un grupo de n objetos, ordenados en un grupo más pequeño de r objetos lo podemos definir como

$$nPr = \frac{n!}{(n-r)!}.$$

Ejemplo,

De un grupo de 30 estudiantes se desea conformar un equipo de trabajo compuesto por un tesorero, un secretario y un vocal. ¿De cuántas maneras se puede conformar el equipo?

En este caso tenemos 30 personas $(n = 30)$ y queremos ordenarlas en grupos de 3 $(r = 3)$, por consiguiente

$$_{30}P_3 = \frac{30!}{(30-3)!} = \frac{30!}{27!} = 30 \times 29 \times 28 = 24360$$

Hay 24360 formas de conformar el equipo de trabajo.

Ejemplo,

María, Pedro, Luis, Lorena y Fernanda, conforman un grupo dedicado a la conservación del medio ambiente. El gobierno decidió enviar a tres de ellos a participar en un simposio fuera del país. ¿De cuántas maneras pueden conformar el equipo que participará en el evento?

Podemos pensar en una permutación $_5P_3$, pero como no importa el orden en que sean elegidos los integrantes los siguientes conjuntos representan el mismo equipo:

$$\{María, Pedro, Luis\}, \{María, Luis\ Pedro\}, \{Luis, Pedro, María\},$$
$$\{Luis, María, pedro\}, \{Pedro, María, Luis\}\ y\ \{Pedro, Luis, Maria\}$$

Obsérvese que debemos disminuir el número de ordenamientos que no se deben contar, que corresponde a $6 = 3!$

Así que si en este caso estamos hablando de una permutación, de 5 elementos en un grupo de 3 elementos, dividida en 3! por consiguiente,

$$\frac{_5P_3}{3!} = \frac{5!}{(5-3)!\,3!}$$

$$\frac{5 \times 4 \times 3}{3!} = \frac{60}{6} = 10$$

Hay 10 formas de conformar el grupo que participará en la conferencia.

Cuando no tenemos en cuenta el orden en el que son tomados los elementos del conjunto, deja de ser *permutación* y recibe el nombre de *combinación*. Las combinaciones son agrupaciones de los objetos que conforman un conjunto, en los que no tenemos en cuenta el orden de los elementos. Para la combinación debemos disminuir el número de ordenamientos que tenemos en cuenta en la permutación. De manera general, el número de combinaciones de un grupo de n objetos, agrupados en un conjunto más pequeño de r objetos lo podemos definir como

$$nCr = \binom{n}{r} = \frac{nPr}{r!} = \frac{n!}{(n-r)!\,r!}$$

El número nCr se llama "n combinado r" y se denota por $\binom{n}{r}$.

Ejemplo,

En una clase de 20 estudiantes, se quiere elegir aleatoriamente un comité formado por 5 estudiantes ¿Cuántos comités se pueden formar?

En este caso no importa el orden en el que escojamos lo estudiantes, por consiguiente es una combinación, donde $n = 30$ y $r = 5$, usando la expresión tenemos:

$$\binom{20}{5} = \frac{20!}{(20-5)!\,5!} = \frac{20!}{15!\,5!} = \frac{1'860.480}{120} = 15.504$$

Podemos formar 15.504 comités.

Ejemplo,

Un entrenador de baloncesto dispone de 15 jugadores para conformar el equipo titular ¿de cuántas maneras puede escoger los jugadores?

En este caso no importa el orden en el que se eligen los jugadores, por consiguiente son combinaciones, usando la expresión:

$$\binom{15}{5} = \frac{15!}{(15-5)!\,5!} = \frac{15!}{10!\,5!} = \frac{360.360}{120} = 3.003$$

El entrenador puede formar el equipo titular de 3.003 formas diferentes.

Con estos métodos de conteo claros, podemos calcular fácilmente la probabilidad de ocurrencia de un evento.

Veamos algunos ejemplos

Ejemplo,

El Instituto de tránsito colombiano marca las placas de las motos con tres letras, dos números y otra letra, por ejemplo ADU25B. Si se va a entregar la placa de la moto de una persona, ¿Cuál es la probabilidad de que la placa empiece con la letra C o con la letra W y termine con una vocal?

Tenemos nuestro evento

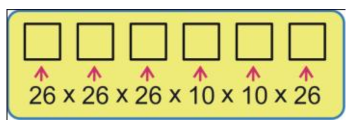
$$A = seleccionar\ una\ placa\ que\ empiece\ en\ la\ letra\ C\ o\ W, y\ termine\ en\ una\ vocal$$

Sabemos que

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

Veamos cuántos elementos tiene Ω

En este caso para los lugares 1,2 y 6 hay 26 posibilidades para cada uno correspondientes a cada letra del alfabeto y para las posiciones 4 y 5 hay 10 posibilidades para cada una correspondientes a los dígitos. Usando el principio de multiplicación se tiene

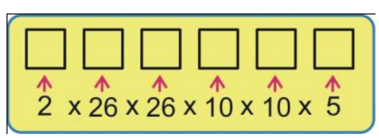


$$26 \times 26 \times 26 \times 10 \times 10 \times 26 = 45697600$$

En total se pueden hacer 45697600 placas para las motos

Ahora veamos cuantos elementos tiene el evento A

Para el evento A la posición 1 solamente puede ser ocupada por las letras C o W, los lugares 2 y 3 tienen 26 posibilidades para cada uno correspondientes a cada letra del alfabeto, para las posiciones 4 y 5 hay 10 posibilidades para cada una correspondientes a los dígitos y para la sexta posición hay 5 posibilidades correspondientes al número de vocales. Usando el principio de multiplicación se tiene



$$2 \times 26 \times 26 \times 10 \times 10 \times 5 = 676000$$

Hay 676000 placas que cumplen las condiciones del evento A

Por consiguiente la probabilidad de ocurrencia del evento A es

$$P(A) = \frac{\#A}{\#\Omega} = \frac{676000}{45697600} \cong 0.01479$$

Ejemplo

En un aula de clase postularon a Luis, Pedro, Juan, Lucia, María y Luisa para ser elegidos como representante, tesorero y monitor. Si son elegidos al azar por medio de sorteo, ¿cuál es la probabilidad de que Pedro sea el representante?

Tenemos nuestro evento

$$C = seleccionar\ un\ grupo\ en\ el\ que\ Pedro\ sea\ \ representante$$

Verifiquemos cuántos elementos tiene Ω

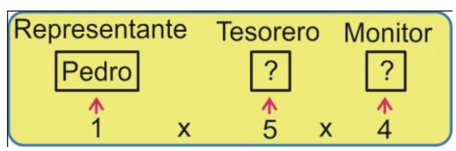
En este caso tenemos que importa el orden en el que seleccionamos a los estudiantes, por consiguiente es una permutación de 6 elementos en grupos de 3 elementos

$${}_6P_3 = \frac{6!}{(6-3)!} = 120$$

En total hay 120 posibilidades de ordenamiento.

Ahora veamos cuántos elementos tiene el evento C.

Para este evento Pedro será el representante, por consiguiente hay 5 persona que pueden ocupar el cargo de tesorero, después de ser elegido el personero, hay 4 estudiantes que pueden ocupar el cargo de monitor, en efecto usando el principio de multiplicación, se tiene



$$1 \times 5 \times 4 = 20$$

Hay 20 grupos en los que Pedro sea representante, en consecuencia

$$P(C) = \frac{\#C}{\#\Omega} = \frac{20}{120} = \frac{1}{6}$$

Ejemplo

El Baloto es un juego de lotería de Colombia, en el cual se eligen seis números diferentes entre los números 1 a 45. ¿Cuál es la probabilidad de ganarse el baloto, con la compra de un boleto?

Veamos cuántos elementos tiene Ω

En este caso no importa el orden de los números por consiguiente es una combinación de 45 elementos en grupos de 6 elementos

$$\binom{45}{6} = \frac{45!}{(45-6)!\,6!} = 8145060$$

En total hay 8145060 posibles números para jugar.

En consecuencia la probabilidad de ganar el baloto con la compra de un boleto "$P(B)$" es:

$$P(B) = \frac{\#B}{\#\Omega} = \frac{1}{8145060}.$$

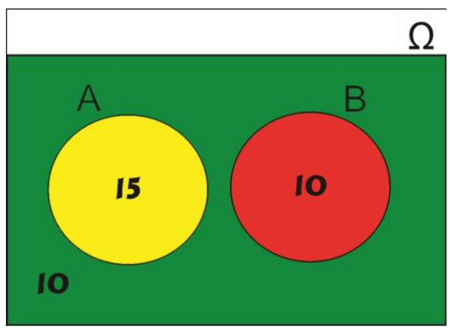
Ejemplo

Si un experimento aleatorio consiste en sacar una balota de una bolsa que contiene 10 balotas verdes, 15 balotas amarillas y 10 balotas rojas. ¿Cuál es la probabilidad que saque una balota amarilla o una balota roja?

Los eventos son:

$$A = "sacar\ una\ balota\ amarilla"$$
$$B = "sacar\ una\ balota\ roja"$$

Veamos en un diagrama de venn



Los eventos A y B no pueden ocurrir al mismo tiempo, son disjuntos, por lo tanto

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

Las probabilidades de cada evento son:

$$P(A) = \frac{\#A}{\#\Omega} = \frac{15}{35}; P(B) = \frac{\#B}{\#\Omega} = \frac{10}{35}$$

La probabilidad de que saque una balota roja o una balota amarilla es:

$$P(A \cup B) = P(A) + P(B) = \frac{15}{35} + \frac{10}{35} = \frac{25}{35}$$

Ejemplo

Si elegimos al azar un estudiante del grupo representado en la tabla ¿Cuál es la probabilidad que el estudiante tenga 14, 16 o 17 años?

Edades de un grupo de estudiantes

Edad	*fi*	*F*	*fr*	*Fr*	%
13	10	10	0.25	0.25	25
14	9	19	0.225	0.475	22.5
15	5	24	0.125	0.6	12.5
16	6	30	0.15	0.75	15
17	10	40	0.25	1	25
Total	40	/	1	/	100

Los eventos para este caso son:

$$A = "elegir\ un\ estudiante\ de\ 14\ años"$$
$$B = "elegir\ un\ estudiante\ de\ 16\ años"$$
$$C = "elegir\ un\ estudiante\ de\ 17\ años"$$

Los eventos $A, B\ y\ C$ no pueden ocurrir al tiempo (son disjuntos), por consiguiente

$$P(A \cup B \cup C) = P(A) + P(B) + P(C)$$

La probabilidad de cada evento es:

$$P(A) = 0.225;\ P(B) = 0.15; P(C) = 0.25$$

La probabilidad que el estudiante tenga 14, 16 o 17 años es:

$$P(A \cup B \cup C) = P(A) + P(B) + P(C) = 0.225 + 0.15 + 0.25 = 0.625.$$

Ejemplo:

El presidente de un club con canchas de golf y de tenis quiere averiguar qué canchas usan los socios y realiza una encuesta. Los resultados indican que el 80% de los socios usan regularmente las canchas de golf, el 60% usa regularmente las canchas de tenis y el 95% de los socios usan al menos una de las dos. Si uno de los miembros del club es elegido aleatoriamente, ¿cuál es la probabilidad de que use los dos tipos de canchas?

Los eventos en este caso son:

$$A = "elegir\ un\ socio\ que\ use\ las\ canchas\ de\ golf"$$
$$B = "elegir\ un\ socio\ que\ use\ las\ canchas\ de\ tenis"$$

Los porcentajes nos indican la probabilidad de ocurrencia de cada evento

$$P(A) = 0.8;\ P(B) = 0.6;\ P(A \cup B) = 0.95$$

Usamos

$P(A \cup B) = P(A) + P(B) - P(A \cap B)$, de donde despejamos $P(A \cap B)$,

$$P(A \cap B) = P(A) + P(B) - P(A \cup B) = 0.8 + 0.6 - 0.95 = 0.45$$

Podemos concluir que hay una probabilidad del 45% que los socios usen los dos tipos de canchas.

Ejemplo,

Una empresa de fármacos realizó un estudio para valorar el efecto de un medicamento para el alivio de los síntomas de la gripe. Para la investigación se seleccionaron 300 personas que presentaban síntomas que incluían dolor de cabeza y escalofrío. Los 300 pacientes recibieron la nueva medicina. Los resultados del estudio son los siguientes: 200 de los pacientes tratados se aliviaron del dolor de cabeza, 160 pacientes se aliviaron del escalofrío y a pacientes 100 les pasó tanto el dolor de cabeza como el escalofrío. ¿Cuál es la probabilidad de que un paciente que toma el medicamento sienta alivio al menos en uno de los dos síntomas?

Los eventos son:

$$A = "Experimentar\ alivio\ del\ dolor\ de\ cabeza"$$
$$B = "Experimentar\ alivio\ en\ el\ escalofrío"$$

Se tiene que

$$P(A) = \frac{2}{3};\ P(B) = \frac{8}{15};\ \ P(A \cap B) = \frac{1}{3}$$

por lo tanto

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{2}{3} + \frac{8}{15} - \frac{1}{3} = \frac{13}{15}.$$

La probabilidad de que un paciente que toma el medicamento sienta alivio al menos en uno de los dos síntomas es $\frac{13}{15}$.

Ejemplo,

Si se toma una carta de una baraja francesa ¿cuál es la probabilidad de que la carta seleccionada sea diferente a un as?

Una baraja de este tipo tiene 52 cartas, divididas en cuatro grupos “4 palos”: picas, corazones, diamantes y tréboles.

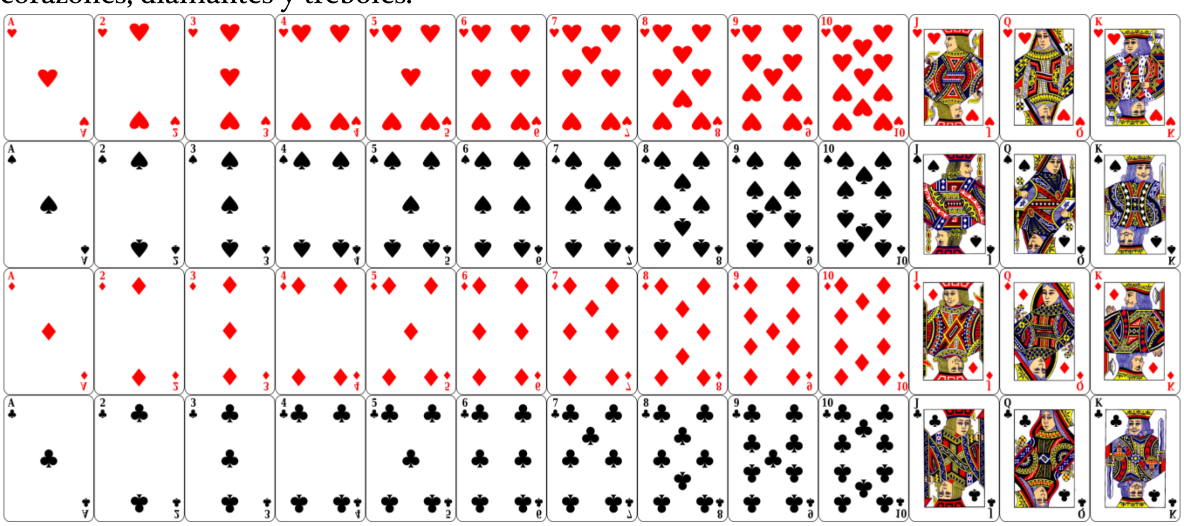


Figura 3.4. Baraja francesa

Sean el evento

$$A = "Sacar\ una\ carta\ que\ no\ sea\ As"$$

Es más fácil contar las cartas que son ases, que las que no son ases, por consiguiente resulta útil ver A^c que corresponde al evento de sacar un as

$$A^c = "Sacar\ una\ carta\ que\ se\ a\ As"$$

Por tanto

$$P(A) = 1 - p(A^c) = 1 - \frac{4}{52} = \frac{12}{13}$$

La probabilidad de no sacar un as es de $^{12}/_{13}$

Ejemplo,
Suponga que en un grupo de 500 estudiantes de Universidad se encuentra que 322 fuman, 300 consumen bebidas alcohólicas, 122 fuman y consumen bebidas alcohólicas. Si se selecciona al azar un miembro de este grupo, encuentre la probabilidad de que el estudiante:
No consuma bebidas alcohólicas
No fumen
Solo fume, o solo consuma bebidas alcohólicas

Los eventos definidos en el problema son:

$$A = \text{"}el\ estudiante\ fuma\text{"}$$
$$B = \text{"}el\ estudiante\ consume\ bebidas\ alcohólicas\text{"}$$

Las probabilidades para cada evento son:

$$P(A) = \frac{322}{500} = 0.644; P(B) = \frac{300}{500} = 0.6; P(A \cap B) = \frac{122}{500} = 0.244$$

El evento B^ccorresponde a los estudiantes que no consumen bebidas alcohólicas, por tanto

$$P(B^c) = 1 - P(B) = 1 - 0.6 = 0.4$$

Hay una probabilidad de 40% que el estudiante elegido no consuma bebidas alcohólicas

El evento A^ccorresponde a los estudiantes que no fuman, por tanto

$$P(A^c) = 1 - P(A) = 1 - 0.644 = 0.356$$

La probabilidad de elegir a un estudiante no fumador es 0.356.

El evento que un estudiante sólo fume, o sólo consuma bebidas alcohólicas, es el complemento del evento que un estudiante fume y consuma bebidas alcohólicas, por consiguiente

$$P(A \cap B)^c = 1 - P(A \cap B) = 1 - 0.244 = 0.756$$

La probabilidad de que el estudiante realice únicamente una de las dos actividades es 0.756

3.3. Probabilidad condicional

Veamos la información contenida en la tabla:

Tabla 3.3. Grupo de estudiantes con sus preferencias deportivas

	Fútbol (F)	Baloncesto (B)	Voleibol (V)	Ninguno (N)	Total
Mujeres (M)	4	9	4	2	19
Hombres (H)	10	3	4	1	18
Total	14	12	8	3	37

Si es seleccionado al azar un hombre de este grupo ¿Cuál será la probabilidad de que le guste el baloncesto?, o si elegimos al azar un estudiante que le guste el voleibol ¿Cuál será la probabilidad de que sea mujer?

Veamos cómo podemos solucionar estas preguntas de manera intuitiva:

De acuerdo a la información contenida en la tabla verificamos que hay 18 hombres en total, y que de estos hombres a 3 les gusta el baloncesto, por consiguiente la probabilidad de que a un hombre le guste el baloncesto es $\frac{3}{18} = \frac{1}{6}$.

Ahora la segunda pregunta razonando de la misma forma:

Sabemos que a 8 estudiantes en total les gusta el voleibol, de los cuales 4 son mujeres, por consiguiente la probabilidad de que a un estudiante que le guste el voleibol sea mujer es $\frac{4}{8} = \frac{1}{2}$

Las anteriores probabilidades corresponden a lo que llamaremos probabilidad condicional, ésta se presenta cuando la probabilidad de ocurrencia de un evento está sujeta a la ocurrencia de otro y se cumple para todo par de eventos A y B contenidos en el mismo espacio muestral.

Se simboliza $P(A/B)$ y se lee "probabilidad condicional de A dado B", también se puede leer como "probabilidad de que ocurra A sabiendo que ocurrió B".

La expresión que nos permite calcular la probabilidad condicional es:

$$P(A/B) = \frac{P(A \cap B)}{P(B)}, \qquad en\ donde\ \ P(B) > 0.$$

Veamos otro ejemplo con la misma información de la tabla

¿Cuál será la probabilidad de que el estudiante elegido le guste el fútbol dado que es mujer?

Definamos lo eventos

$$F = "Al\ estudiante\ elegido\ le\ gusta\ el\ fútbol"$$
$$M = "El\ estudiante\ elegido\ es\ mujer"$$

De acuerdo a la información contenida en la tabla tenemos

$$P(F) = \frac{14}{37}, \qquad P(M) = \frac{19}{37}, \qquad P(F \cap M) = \frac{4}{37}$$

Por consiguiente

$$P(F/M) = \frac{P(F \cap M)}{P(M)} = \frac{\frac{4}{37}}{\frac{19}{37}} = \frac{4}{19}$$

La probabilidad de que el estudiante elegido le guste el fútbol dado que es mujer es de 0,21.

Veamos otro ejemplo,

Nuestro experimento aleatorio consiste en sacar cartas de un mazo de cartas francesas.

¿Cuál es la probabilidad de sacar un as, dado que es una carta negra?

Los eventos son:

$$A = "Sacar\ un\ as\ de\ la\ baraja"$$
$$B = "Sacar\ una\ carta\ negra"$$
$$A \cap B = "sacar\ un\ as\ negro"$$

Queremos calcular $P(A/B)$. La baraja tiene 52 cartas, de las cuales 26 son negras, cuatro son ases, y hay dos ases rojos y dos negros, por consiguiente las probabilidades son:

$$P(A) = \frac{4}{52} = \frac{1}{13}; \; P(B) = \frac{26}{52} = \frac{1}{2}; \quad P(A \cap B) = \frac{2}{52} = \frac{1}{26}$$

Por consiguiente

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{1}{26}}{\frac{1}{2}} = \frac{1}{13}.$$

¿Cuál es la probabilidad de sacar una carta impar, dado que es un corazón?

Los eventos son:

$$A = Sacar\ una\ carta\ par$$
$$B = Sacar\ un\ corazón$$
$$A \cap B = sacar\ una\ carta\ par\ que\ sea\ corazón$$

Queremos calcular $P(A/B)$, La baraja tiene 5 cartas pares en cada palo (4 palos), 13 cartas son corazones, por consiguiente las probabilidades son:

$$P(A) = \frac{5 \times 4}{52} = \frac{5}{13}; \; P(B) = \frac{13}{52} = \frac{1}{4}; \quad P(A \cap B) = \frac{5}{52}$$

Por consiguiente

$$P\left(\frac{A}{B}\right) = \frac{P(A \cap B)}{P(B)} = \frac{\frac{5}{52}}{\frac{1}{4}} = \frac{5}{13}.$$

¿Cuál es la probabilidad de sacar un corazón, dado que es una carta impar?
Los eventos son los mismos que en el ejercicio anterior, por consiguiente

$$P\left(\frac{B}{A}\right) = \frac{P(B \cap A)}{P(A)} = \frac{\frac{5}{52}}{\frac{5}{13}} = \frac{1}{4}.$$

3.4. Independencia

Puede suceder que la ocurrencia de un evento no afecte la probabilidad de ocurrencia de otro evento, por ejemplo si se lanzan de manera consecutiva una moneda corriente y un dado al aire y consideramos el evento B que consiste en que la moneda salga cara y el evento A que consiste en que resultado del dado sea un número par, obsérvese que el hecho de caer cara en el lanzamiento de la moneda no afecta la probabilidad de ocurrencia de que el resultado del lanzamiento del dado sea par. Esto es $P(A|B) = P(A)$, en este caso decimos que el evento A es independiente del evento B. Observamos de la definición de probabilidad condicional que se tiene

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

Formalmente se dice que A y B son independientes, si

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

Retomando el ejemplo, si queremos calcular la probabilidad que la moneda caiga cara y el resultado del lanzamiento del dado sea par, será

$$P(A \cap B) = P(A) \times P(B) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}.$$

Ejemplo,
Si se toman dos cartas con reemplazo de una baraja francesa, determina la probabilidad de que la primera carta sea un as y la segunda un diamante.
Este problema se hace con reemplazo; es decir que cuando se toma la primera carta, ésta se regresa al mazo. De manera que cuando se saca la segunda carta, están disponibles nuevamente las 52 cartas.

Veamos los eventos

$$A = \text{"}elegir\ un\ as\text{"}$$
$$B = \text{"}elegir\ un\ diamante\text{"}$$
$$P(A) = \frac{4}{52}, \qquad P(B) = \frac{13}{52}$$

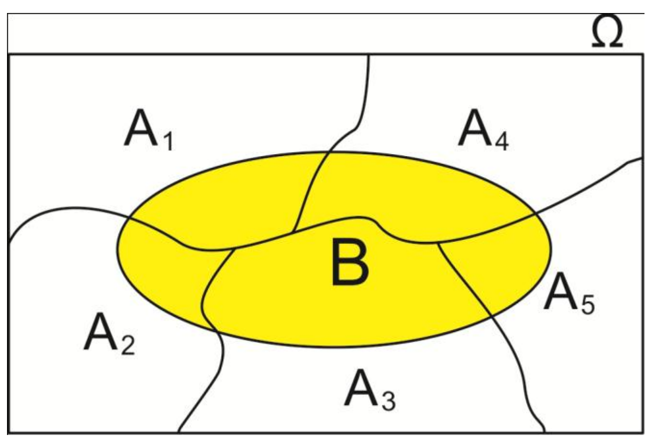
Los eventos A y B son independientes. Por consiguiente

$$P(A \cap B) = \frac{4}{52} \times \frac{13}{52} = \frac{1}{52}.$$

3.5. Teorema de probabilidad total

Supongamos que un experimento aleatorio tiene 5 eventos que llamaremos $A_1, A_2, A_3, A_4\ y\ A_5$, de tal forma que estos eventos no se intersecan entre sí, es decir son disjuntos ($A_i \cap A_j = \emptyset,\ para\ todo\ i \neq j$), por otro lado la unión de los eventos forman el espacio muestral ($A_1 \cup A_2 \cup A_3 \cup A_4 \cup A_5 = \Omega$) y además $P(A_i) > 0\ para\ todo\ i$

Queremos calcular la probabilidad de ocurrencia de un evento B dentro del mismo experimento



Podemos observar que

$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + P(A_4 \cap B) + P(A_5 \cap B).$$

Usando la definición de probabilidad condicional, se tiene

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) +$$
$$P(B|A_3) \cdot P(A_3) + P(B|A_4) \cdot P(A_4) + P(B|A_5) \cdot P(A_5).$$

De manera general tenemos *el teorema de probabilidad total*
Si tenemos $A_1, A_2, A_3, \dots, A_n$ eventos de un experimento aleatorio de tal forma que

$\left(A_i \cap A_j = \emptyset, para\ todo\ i = j\right)$ y $A_1 \cup A_2 \cup A_3 \cup A_4 \cup \dots \cup A_n = \Omega$,

entonces para cualquier evento B se cumple,

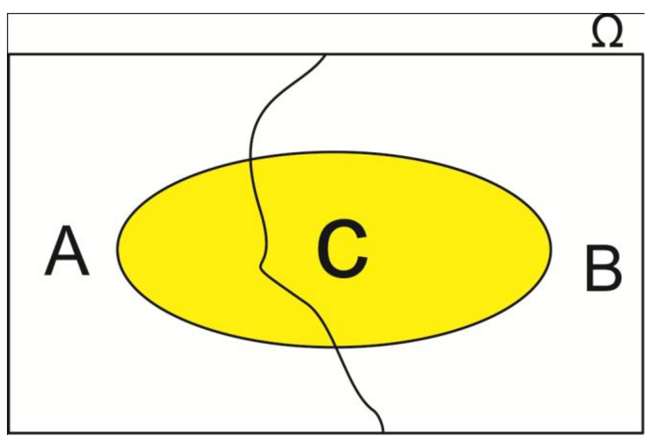
$$P(B) = P(A_1 \cap B) + P(A_2 \cap B) + P(A_3 \cap B) + \cdots + P(A_n \cap B)$$
$$= P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + \cdots + P(B|A_n) \cdot P(A_n).$$

Veamos algunos ejemplos

Ejemplo,
En un colegio, la probabilidad de que un estudiante elegido al azar sea hombre y le guste el fútbol es 0.6 y la probabilidad de que el estudiante elegido al azar sea mujer y le guste el fútbol es 0.2. ¿Cuál es la probabilidad de que una persona elegida al azar le guste el fútbol?
En este ejercicio aparecen tres eventos

$$A = \text{"}Elegir\ una\ mujer\text{"}$$
$$B = \text{"}Elegir\ un\ hombre\text{"}$$
$$C = \text{"}Elegir\ un\ estudiante\ que\ le\ guste\ el\ fútbol\text{"}$$

Como el estudiante es hombre o es mujer tenemos que
$$A \cup B = \Omega$$
Y como el estudiante no puede ser hombre y mujer al mismo tiempo, los eventos A y B son disjuntos, $A \cap B = \emptyset$
Veamos gráficamente



Tenemos
$$P(C) = P(A \cap C) + P(B \cap C) = 0.2 + 0.6 = 0.8$$
La probabilidad de elegir un estudiante que le guste el fútbol es 0.8

Ejemplo,

Una empresa de taxis tiene tres líneas en la ciudad. El 50% de los vehículos cubren el servicio de la primera línea, el 30% cubren la segunda y el 20% cubren la tercera línea. Se sabe que la probabilidad de que, diariamente, un taxi tenga una falla mecánica es del 3%, 5% y 2%, respectivamente, para cada línea. Determine la probabilidad de que, en un día, un taxi sufra una falla mecánica.

Veamos los eventos

$$A = \text{"}sufrir\ una\ falla\ mecánica\text{"}$$
$$B = \text{"}Seleccionar\ un\ taxi\ de\ la\ linea\ 1\text{"}$$
$$C = \text{"}Seleccionar\ un\ taxi\ de\ la\ linea\ 2\text{"}$$
$$D = \text{"}Seleccionar\ un\ taxi\ de\ la\ linea\ 3\text{"}$$

$$P(B) = 0.5, \qquad P(C) = 0.3, \qquad P(D) = 0.2, \qquad P(A|B) = 0.03, \qquad P(A|C) = 0.05,$$
$$P(A|D) = 0.02$$

Usando probabilidad total tenemos

$$P(A) = P(A|B) \cdot P(B) + P(A|C) \cdot P(C) + P(A|D) \cdot P(D)$$
$$P(A) = 0.03 \cdot 0.5 + 0.05 \cdot 0.3 + 0.02 \cdot 0.2 = 0.034$$

La probabilidad de que un taxi sufra de una falla mecánica es 0.034.

Ejemplo,

Una empresa de alimentos elabora sus productos en cuatro lugares. La producción total que se fabrica en cada uno es del 35%, 30%, 20% y 15%, respectivamente. El porcentaje de envasado incorrecto en cada lugar es del 2%, 3%, 5% y 6% correspondientemente. Si tomamos un producto de la empresa al azar. ¿Cuál es la probabilidad de que se encuentre mal envasado?

Los eventos

$$A_1 = \text{"}elaborar\ el\ producto\ en\ el\ lugar\ 1\text{"}$$
$$A_2 = \text{"}elaborar\ el\ producto\ en\ el\ lugar\ 2\text{"}$$
$$A_3 = \text{"}elaborar\ el\ producto\ en\ el\ lugar\ 3\text{"}$$
$$A_4 = \text{"}elaborar\ el\ producto\ en\ el\ lugar\ 4\text{"}$$
$$B = \text{"}Envasar\ de\ manera\ incorrecta\ el\ producto\text{"}$$

$$P(B) = P(B|A_1) \cdot P(A_1) + P(B|A_2) \cdot P(A_2) + P(B|A_3) \cdot P(A_3) + P(B|A_4) \cdot P(A_4)$$
$$P(B) = 0.02 \cdot 0.35 + 0.03 \cdot 0.3 + 0.05 \cdot 0.2 + 0.06 \cdot 0.15 = 0.035$$

La probabilidad de encontrar un producto mal envasado es de 0.035

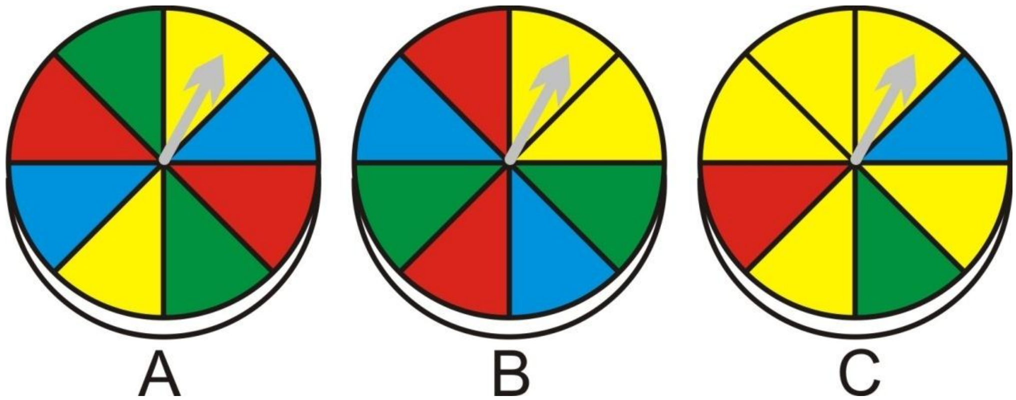
3.6. Actividades

1. Un experimento aleatorio consiste en lanzar un dado y verificar el número que sale en la cara superior. Se definen 3 eventos

$$A = \text{"}obtener\ un\ número\ mayor\ que\ 4\text{"}$$
$$B = \text{"}obtener\ un\ número\ primo\text{"}$$
$$C = \text{"}obtener\ un\ número\ impar\text{"}$$

   a) Determine los elementos que conforman los siguientes eventos: $\Omega, A, B, C, A^C, A \cup B, B \cap C,\ A - C, A \cup C^C$.
   b) ¿El evento A y el evento B, son mutuamente excluyentes? Explique.
   c) ¿El evento B y el evento C, son mutuamente excluyentes? Explique.
2. Escriba el espacio muestral en cada uno de los siguientes experimentos aleatorios:
   a. Se lanza una moneda corriente
   b. Se lanzan dos monedas corriente
3. En una caja hay cuatro papeletas rojas y una papeleta verde. Se van sacando papeletas (sin volver a introducirlas en la caja) hasta que aparezca la verde.
4. A una fiesta llegan Sonia, Gabriela, Laura, Sergio y Andrés. Se eligen dos personas al azar para un premio. Determine el espacio muestral de este experimento.
5. Una caja contiene tres esferas azules y dos amarillas, y otra contiene dos esferas azules y tres amarillas. Se toma, al azar, una esfera de cada caja. Escriba el espacio muestral del experimento.
6. ¿Cuántos números de cuatro cifras se pueden formar con los dígitos 4, 5, 6, 7, 8 y 9 si no se permite la repetición?
7. De una clase de 20 niñas se escogerán 6 para ir a un paseo. ¿Cuántos posibles grupos de 6 se pueden formar?
8. Una persona olvidó su clave de acceso a un sistema de computadoras, la clave está formada por 4 números, determine cuántas formas diferentes puede tener la clave si no se permite los números que empiecen, o terminen en cero.
9. El Instituto de tránsito colombiano marca las placas de los carros con tres letras y tres números. Si se van a hacer las placas para los carros de una ciudad, de tal forma que deben iniciar con la letra C o con la letra B. ¿Cuántas placas se podrán hacer allí? Si se elige al azar un carro de esta ciudad ¿cuál es la probabilidad que su placa empiece con la letra B y su número tercer número sea par?
10. En una ciudad, los números telefónicos deben comenzar con 2, 3 ó 6 y deben constar de siete dígitos, con la condición que no pueden terminar en 0000. ¿Cuántos números telefónicos generar en esta ciudad? Si la empresa de telefonía desea rifar un carro en esta ciudad ¿Cuál es la probabilidad que el número telefónico ganador termine en 77?
11. Juana quiere abrir una cuenta en el banco “La República”, allí los caracteres asignados a las cuentas de los clientes constan de dos letras, seguidas por cuatro números y luego dos letras más. ¿Cuál es la probabilidad que la cuenta de Juana empiece en la letra A o Z, el primer número sea par y termine en una vocal?

12. Calcule las probabilidades que la flecha se detenga en el color amarillo en cada una de las ruletas.



13. Si elegimos al azar un estudiante graduado de secundaria, un año atrás. Existe una probabilidad del 60% de que este trabaje, una probabilidad del 60% de que siga estudiando y una probabilidad de 50% de hacer las dos cosas. ¿Cuál es la probabilidad que el bachiller ni estudie ni trabaje?
14. Una bolsa tiene 100 monedas de las cuales, 40 son de \$500, 30 son de \$200, 20 son de \$100, y 10 son de \$50. Si saco una moneda al azar ¿Cuál es la probabilidad de no sacar una moneda de \$50, ni una moneda de \$200?
15. Si se toma una carta de una baraja francesa, determine la probabilidad que la carta sea un as de diamantes.
16. En una ciudad el 45% de los habitantes consume pan integral, el 30% consume pan de mantequilla y el 20% consume ambos. Se selecciona un habitante al azar, ¿Cuál es la probabilidad que coma pan de mantequilla, si se sabe que consume pan integral?
17. De una urna que contiene 4 balotas negras y 6 blancas se extraen dos balotas sin reposición, ¿cuál es la probabilidad que la segunda balota sea blanca, sabiendo que la primera es negra?
18. De una investigación a un grupo de personas se obtuvieron los siguientes resultados: 2 de cada 7 son morenas, y 2 de cada 8 tienen los ojos azules. Si se elige al azar una de las personas investigadas ¿Cuál es la probabilidad que una persona sea morena y tenga los ojos azules?
19. En una clase, un 50% de estudiantes aprobaron Biología, y un 45% matemáticas. Se sabe que la probabilidad de aprobar Biología si se ha aprobado matemáticas es 0.6. ¿Qué porcentaje de alumnos aprobaron ambas asignaturas?
20. Se tienen tres cajas con bombillos. La primera contiene 20 bombillos, de las cuales hay cinco fundidos; en la segunda hay 12 bombillos buenos y tres fundidos, y la tercera hay siete bombillos fundidos de 20 que hay en la caja. ¿Cuál es la probabilidad que al tomar un bombillo al azar de cualquiera de las cajas, esté fundido?

4. Distribuciones de probabilidad

4.1 Variable aleatoria

Se lanza una moneda corriente cuatro veces consecutivas. Se quiere analizar el número de sellos que se pueden obtener.
El espacio muestral correspondiente a este experimento es:

$$\Omega = \begin{Bmatrix} CCCC, CCCS, CCSS, CCSC, CSSS, CSSC, CSCS, CSCC, \\ SSSC, SSCC, SSCS, SCCC, SCCS, SCSC, SCSS, SSSS \end{Bmatrix}$$

Veamos cuántos sellos hay en cada una de las posibilidades

Espacio Muestral (S)	# de sellos
$CCCC$	0
$CCCS, CCSC, SCCC, CSCC$	1
$CCSS, CSSC, SCCS, SSCC, SCSC, CSCS$	2
$CSSS, SSSC, SCSS, SSCS$	3
$SSSS$	4

Nos podemos dar cuenta que a todos los eventos simples del espacio muestral se les asigna un número real, que corresponde al número de sellos obtenidos después de los cuatro lanzamientos de la moneda, de esta manera, hemos definido una función, cuyo dominio es el espacio muestral y que toma valores en el conjunto de los números reales. Una aplicación de esta forma es lo que llamaremos *Variable aleatoria real.*

Variable aleatoria*: es una función que asigna un número real a cada elemento del espacio muestral. Habitualmente se representan con las letras mayúsculas* X, Y *y* Z.

Las variables aleatorias al igual que las variables estadísticas, se clasifican en discretas y continuas. Por ejemplo el número de preguntas en un examen, el número obtenido al lanzar un dado, el número de personas que se suben en un bus de servicio público, entre otras son variables discretas. El tiempo de duración de un bombillo, la cantidad de agua consumida en un mes, la estatura de los miembros de un grupo de personas, son variables continuas.

Ejemplo

Considérese el lanzamiento de un dado corriente dos veces consecutivas. Sea X la variable aleatoria que asigna a cada pareja de resultados su suma. Observemos que la variable toma un número finito de valores que son $2, 3, 4, 5, 6, 7, 8, 9, 10, 11\ y\ 12$. Esta es una variable aleatoria discreta.

Ejemplo,
Pensemos en el caso que se lanza una moneda tantas veces sea necesario hasta obtener por primera vez sello. Sea X la variable aleatoria que asigna a cada elemento del espacio muestral el número de caras obtenidas antes de obtener un sello. Es fácil ver que la variable toma un número infinito de valores que son $0, 1, 2, \dots, \infty$, obsérvese que si la variable aleatoria toma el valor 0, el sello se obtuvo en el primer lanzamiento, si la variable aleatoria toma el valor 5, indica que en los primeros cinco lanzamientos se obtuvo cara y en el sexto se obtuvo sello, el valor "∞" indica que nunca se obtiene sello. Esta variable aleatoria se considera discreta, puesto que se puede establecer una correspondencia uno a uno entre el conjunto de valores que toma la variable y el conjunto de los números naturales, es decir se pueden contar sus elementos.

Variable aleatoria discreta *es aquella que toma valores en un conjunto con un número finito de elementos, o en un conjunto infinito pero contable.*

Los resultados de algunos experimentos aleatorios no son ni finitos, ni contables. Supongamos que se quiere elegir al azar un estudiante de un salón de clase, y sea X la variable aleatoria que asigna al estudiante su peso. Si utilizamos una báscula de precisión es claro que tenemos una cantidad infinita de valores no contables, este tipo de variable será considerada continua.

Variable aleatoria continua *es aquella que toma valores en un conjunto infinito no contable.*

Ejemplo,
Con el fin de realizar un control de calidad en una fábrica de bombillos, se mide el tiempo de duración de bombillos elegidos al azar y sea X la variable aleatoria que asigna a cada bombillo su tiempo de duración. Esta es una variable aleatoria continua.

4.2. Función de distribución acumulada de una variable aleatoria discreta

Una variable aleatoria discreta se describe por medio de las probabilidades de ocurrencia de cada uno de los valores que toma la variable. $P(X = x)$ representa la probabilidad de ocurrencia de un valor que toma la variable aleatoria, por ejemplo la probabilidad de que salgan dos sellos cuando se lanzan las cuatro monedas es $P(X = 2) = {}^{3}/_{8}$. Esta asignación de probabilidades a las variables aleatorias discretas reciben el nombre de función de probabilidad o función de masa de probabilidad $f(x)$, y se define como:

$$f(x) = P(X = x_i), \qquad si\ x = x_1, x_2, \dots\ son\ valores\ de\ X;$$

La distribución de probabilidad se puede representar por medio de una tabla, un diagrama de puntos o un diagrama de barras. Veamos cómo se puede representar la función de probabilidad para el lanzamiento de la moneda cuatro veces.

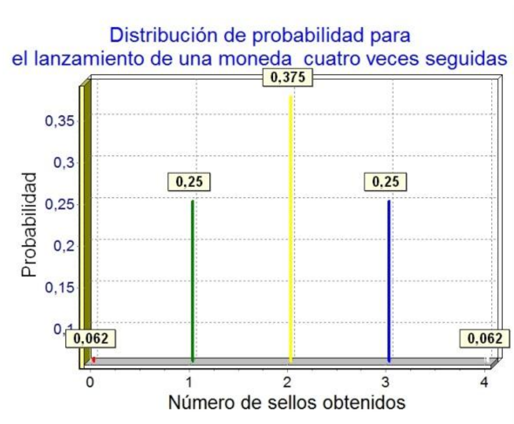


Figura 4.1.

$X = x$	0	1	2	3	4
$P(X = x)$	1/16	1/4	3/8	¼	1/16

Tabla 4.1. Distribución de probabilidad
Para el lanzamiento de una moneda
Cuatro veces seguidas

Veamos otro ejemplo,

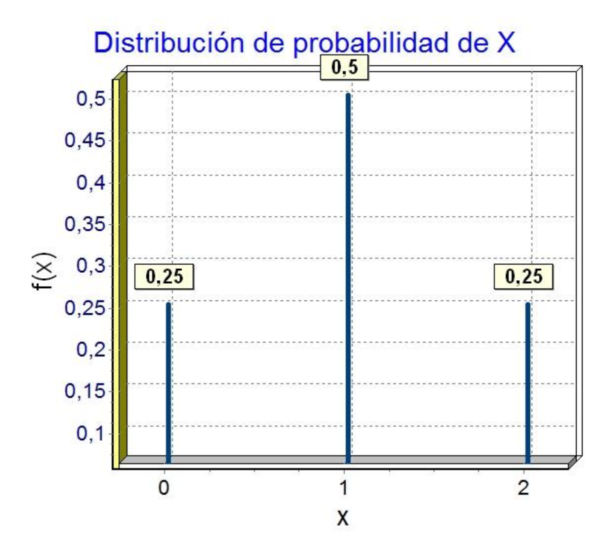
Dentro de una urna hay cuatro bolas, dos rojas y dos azules. Si se extraen al azar dos de ellas consecutivamente sin reposición. Determinar la distribución de probabilidad de la variable aleatoria X: número de balotas rojas extraídas.

El espacio muestral correspondiente a este experimento es

$$\Omega = \{(R,R), (R,A), (A,R), (A,A)\}$$

y los valores que toma X son 0, 1 y 2, en consecuencia la distribución de probabilidad es

$$f(x) = P(X = x) = \begin{cases} 1/_4 \; si \; x = 0 \\ 1/_2 \; si \; x = 1 \\ 1/_4 \; si \; x = 2 \end{cases}$$



Si se conoce la función de masa de probabilidad de la variable aleatoria X podemos calcular fácilmente las probabilidades de ocurrencia de ciertos eventos de interés, por ejemplo la probabilidad que se extraiga al menos una bola roja.

$$P(X > 0) = P(X = 1) + P(X = 2) = 1/_4 + 1/_2 = 3/_4$$

Si queremos calcular probabilidades como $(P(X \leq x))$, usaremos lo que se conoce como función de distribución acumulada ($F(x)$), esta se define como

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

Por ejemplo, cuando se lanzan cuatro monedas de los ejemplos anteriores, podemos calcular la probabilidad de obtener dos o menos sellos

$$P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 1/_{16} + 1/_4 + 3/_8 = 11/_{16}$$

Veamos cómo obtener la función de distribución acumulada de la variable aleatoria X: número de sellos obtenidos en el lanzamiento de las cuatro monedas.

Pensemos los casos posibles

El primer caso, es en el cual no se obtiene ningún sello ($X = 0$)

$$F(0) = P(X \leq 0) = 1/_6$$

Segundo caso en el cual se obtiene máximo un sello ($X \leq 1$)

$$F(1) = P(X \leq 1) = P(X = 0) + P(X = 1) = 1/_{16} + 1/_4 = 5/_{16}$$

De la misma forma verificamos el caso en el que se obtiene máximo dos sellos ($X \leq 2$)

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2) = 1/_{16} + 1/_4 + 3/_8 = 11/_{16}$$

Ahora el caso en el cual se obtienen máximo tres sellos ($X \leq 3$)

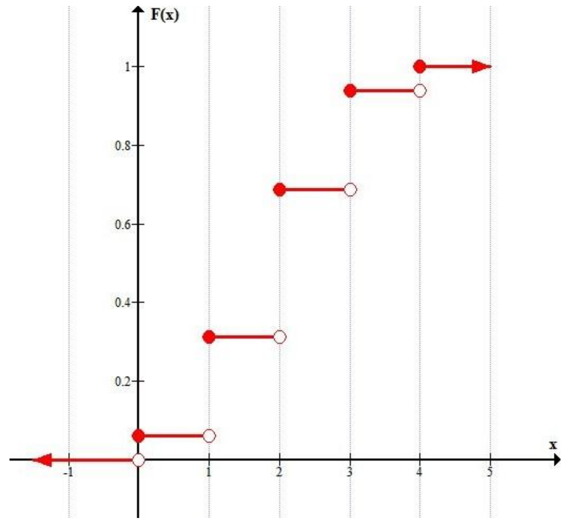
$$F(3) = P(X \leq 3) = 1/_{16} + 1/_4 + 3/_8 + 1/_4 = 15/_{16}$$

Y por último el caso en el cual se obtiene máximo cuatro sellos

$$F(4) = P(X \leq 4) = 1/_{16} + 1/_4 + 3/_8 + 1/_4 + 1/_{16} = 1$$

Por consiguiente nuestra función de distribución acumulada $(F(x))$ es

$$F(x)=\begin{cases} 0 & si \quad x<0 \\ {}^{1}/_{16} & si\ 0\le x<1 \\ {}^{5}/_{16} & si\ 1\le x<2 \\ {}^{11}/_{16} & si\ 2\le x<3 \\ {}^{15}/_{16} & si\ 3\le x<4 \\ 1 & si \quad x\ge 4 \end{cases}$$



Veamos otro ejemplo,

Dentro de una urna hay cuatro bolas, dos rojas y dos azules y el experimento consiste en extraer al azar dos de ellas consecutivamente sin reposición. La variable aleatoria X: número de balotas rojas extraídas. La distribución de probabilidad es

$\boldsymbol{X=x}$	**0**	**1**	**2**
$\boldsymbol{P(X=x)}$	${}^{1}/_{4}$	${}^{1}/_{2}$	${}^{1}/_{4}$

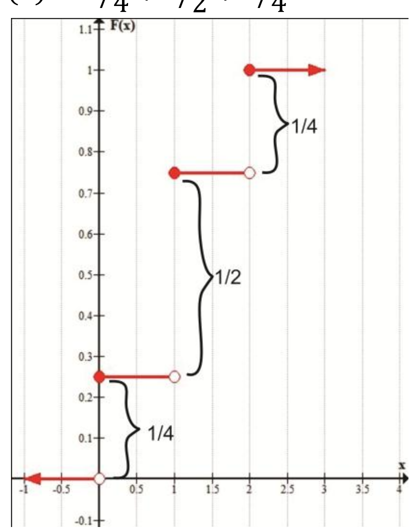
La función de distribución acumulada correspondiente es:

En este experimento la variable X no puede tomar valores menores a 0, por consiguiente $P(X<0)=0$

Para los demás valores se tiene

$$F(0)={}^{1}/_{4};\ F(1)={}^{1}/_{4}+{}^{1}/_{2}={}^{3}/_{4};\ F(2)={}^{1}/_{4}+{}^{1}/_{2}+{}^{1}/_{4}=1$$

$$F(x)=\begin{cases} 0 & si \quad x<0 \\ {}^{1}/_{4} & si\ 0\le x<1 \\ {}^{3}/_{4} & si\ 1\le x<2 \\ 1 & si \quad x\ge 2 \end{cases}$$



Obsérvese que las magnitudes de los saltos en los valores de $x = 0{,}1\ y 2$ son $1/4$, $1/2\ y\ 1/4$ respectivamente y corresponden a las probabilidades asignadas por $P(X = x)$. Es decir $f(0) = 1/4, f(1) = 1/2\ y\ f(2) = 1/4$.

A partir de la observación anterior, es claro que podemos obtener la distribución de probabilidad de una variable aleatoria discreta usando la función de distribución acumulada. Podemos notar que

$$f(x) = P(X = x) = F(x) - F(x^-)$$

donde $F(x^-)$ representa el límite por izquierda de F en x. Esto es, $P(X = x)$ es la magnitud del "salto" de la función en el punto x.

Veamos un ejemplo de cómo a partir de la función de distribución acumulada podemos construir la función de masa de probabilidad de la variable

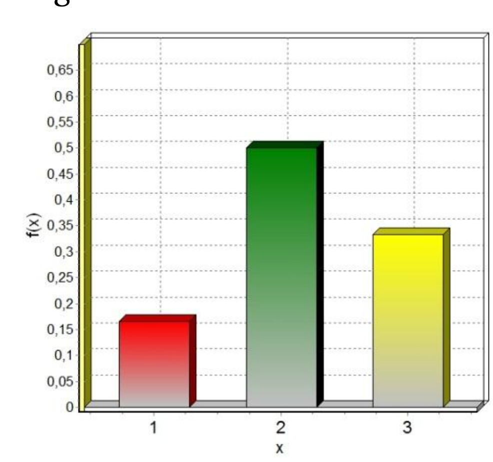
$$F(x) = \begin{cases} 0 & si \quad x < 1 \\ 1/6 & si \quad 1 \le x < 2 \\ 4/6 & si \quad 2 \le x < 3 \\ 1 & si \quad x \ge 3 \end{cases}$$

Para cada valor de la variable tenemos

$$P(X = 1) = 1/6$$
$$P(X = 2) = 4/6 - 1/6 = 1/2$$
$$P(X = 3) = 1 - 4/6 = 1/3$$

Por consiguiente la función de distribución y su diagrama son:

$$f(x) = \begin{cases} 1/6 & si\ x = 1 \\ 1/2 & si\ x = 2 \\ 1/3 & si\ x = 3 \end{cases}$$



4.2.1. Esperanza matemática, o valor esperado de una variable aleatoria discreta

En un salón de clase 10 estudiantes tienen 14 años, 20 estudiantes tienen 15 años, 15 estudiantes tienen 16 años y 5 estudiantes tienen 17 años. Si el experimento consiste en elegir un estudiante al azar, y la variable aleatoria X es la edad del estudiante elegido, entonces tenemos la siguiente distribución

$\boldsymbol{X = x}$	**14**	**15**	**16**	**17**
N° estudiantes	10	20	15	5
$\boldsymbol{P(X = x)}$	$\frac{10}{50} = 0.2$	$\frac{20}{50} = 0.4$	$\frac{15}{50} = 0.3$	$\frac{5}{50} = 0.1$

Si nos interesa la edad promedio de estos estudiantes, es decir el valor promedio de X, lo podemos calcular de la siguiente forma

$$\frac{14 \cdot 10 + 15 \cdot 20 + 16 \cdot 15 + 17 \cdot 5}{50} = 15.3$$

Utilizando algunas propiedades aritméticas, esta expresión se puede rescribir de la siguiente forma

$$14 \cdot \frac{10}{50} + 15 \cdot \frac{20}{50} + 16 \cdot \frac{15}{50} + 17 \cdot \frac{5}{50} = 15.3$$

Nos podemos dar cuenta que las fracciones corresponden a las probabilidad de ocurrencia de cada uno de los valores que toma la variable aleatoria. Este valor recibe el nombre de valor esperado o esperanza matemática (E).

Lo anterior motiva a la siguiente definición:

Si X es una variable aleatoria que toma sólo un número finito de valores $x_1, x_2, \dots, x_n$ con probabilidad $P(X = x_i)$, el valor esperado o esperanza de X se define como:

$$E(X) = \mu = \sum_{i=1}^{n} x_i P(X = x_i) = \sum_{i=1}^{n} x_i f(x_i)$$

La esperanza se interpreta como el valor promedio de los valores que toma la variable.

Ejemplo,

Sea X es la variable aleatoria que mide el número de personas que entran a un café internet en una hora, cuya función de distribución es

$\boldsymbol{X = x}$	**0**	**1**	**2**	**3**	**4**
$\boldsymbol{P(X = x)}$	0.05	0.1	0.2	0.5	0.15

determinemos cuántas personas se espera que entren al café internet en una hora

$$E(X) = 0 \cdot 0.05 + 1 \cdot 0.1 + 2 \cdot 0.2 + 3 \cdot 0.5 + 4 \cdot 0.15 = 2.6$$

Se espera que entren 2.6 personas, es decir esperamos que entren 2 o 3 personas en una hora al café internet.

Como ya tenemos un promedio entre los valores que toma la variable aleatoria discreta ($E(X)$), es natural buscar un valor que nos permita calcular la dispersión entre estos valores. Al igual que en el primer capítulo, este valor recibe el nombre de varianza, y se define de la siguiente manera:

Sea X una variable aleatoria discreta con función de distribución $f(x)$, se define la varianza de X, usualmente se simboliza con $V(X)$, o σ^2 como:

$$V(X) = E(X^2) - (E(X))^2$$

Calculemos la varianza para el ejemplo anterior del café internet.

Lo primero que debemos hacer es calcular $E(X)$

$$E(X) = 2.6$$

Luego calculamos $E(X^2)$. De manera general podemos calcular $E(X^2)$ usando la siguiente expresión:

$$E(X^2) = \sum_{i=1}^{n} x_i{}^2 P(X = x_i) = \sum_{i=1}^{n} x_i{}^2 f(x_i),$$

Por consiguiente

$$E(X^2) = 0^2 \cdot 0.05 + 1^2 \cdot 0.1 + 2^2 \cdot 0.2 + 3^2 \cdot 0.5 + 4^2 \cdot 0.15 = 7.8$$

Por consiguiente el valor de la varianza es

$$\sigma^2 = (X^2) - \left(E(X)\right)^2 = 7.8 - (2.6)^2 = 1.04$$

Podemos calcular también la desviación estándar de X, que corresponde a la raíz cuadrada de la varianza. Usualmente se denota por σ, por consiguiente la desviación estándar es:

$$\sigma = \sqrt{\sigma^2} = \sqrt{1.04} = 1.019.$$

4.3. Distribución binomial

La distribución binomial es una distribución de probabilidad discreta. Comúnmente se usa en los experimentos que se repiten n veces de manera independiente y permiten describir experimentos en los cuales analizamos la ocurrencia o no ocurrencia de un evento. Será llamado éxito en el caso que ocurra el evento y se llamará fracaso en el caso de que no ocurra. Las probabilidades de éxito y fracaso son las mismas en todas las repeticiones, la probabilidad de éxito se representan con p y la probabilidad de fracaso $(1-p)$ se representa con q. La distribución binomial permite calcular la probabilidad de obtener cierto número de éxitos en una secuencia de n repeticiones del experimento. En este tipo de experimentos la variable aleatoria X representa el número de éxitos obtenidos en n repeticiones, y se dice que X tiene una distribución binomial de parámetro n y p, y se nota como $X\sim\mathcal{B}(n,p)$. Para estos casos la probabilidad de que el evento ocurra x veces en n repeticiones se puede calcular por medio de la función de masa de probabilidad:

$$f(x) = P(X = x) = \binom{n}{x} p^x q^{n-x} = \frac{n!}{x!\,(n-x)!} p^x q^{n-x}, \qquad si\ \ x = 0, 1, 2, \dots, n$$

Veamos algunos ejemplos

Ejemplo,

La probabilidad de éxito de un medicamento para aliviar el dolor de cabeza es 0.89. Si quince pacientes toman el medicamento. ¿Cuál es la probabilidad de que ninguno siga con el dolor de cabeza?

En este caso tenemos que el experimento se repite quince veces ($n = 15$), y considerando el evento A: éxito del medicamento, se tiene $P(A) = p = 0.89$ y $q = 1 - p = 0.11$. La variable X corresponde al número de pacientes en lo que tendrá éxito el medicamento. Por consiguiente X tiene una distribución binomial de parámetros 15 y 0.89 ($X\sim\mathcal{B}(15,\ 0.89)$).

Queremos calcular la probabilidad de que ninguno siga con dolor de cabeza, que es lo mismo que decir que el medicamento tenga éxito en los 15 pacientes, esto es

$$P(X = 15) = \binom{15}{15} \times 0.89^{15} \times 0.11^{0} = 0.174$$

La probabilidad de que ninguno de los quince pacientes sufra de dolor de cabeza después de tomado el medicamento es de 0.174.

Ejemplo,

La probabilidad de que le hagan un gol al arquero del equipo “Dorado” cada vez que un jugador del equipo contrario remata a su cancha es del 20%. Si en el partido le hacen 20 remates. ¿Cuál es la probabilidad de que le hagan cinco goles? ¿Cuál es la probabilidad de que no le hagan goles? ¿Cuál es la probabilidad de que le hagan algún gol?

En este caso tenemos que el experimento se repite veinte veces ($n = 20$), y considerando el evento A: hacer gol, se tiene $P(A) = p = 0.2$ y $q = 1 - p = 0.8$. La variable X corresponde al número de goles que le hacen al equipo Dorado. Por consiguiente X tiene una distribución binomial de parámetros 20 y 0.2 ($X\sim\mathcal{B}(20,\ 0.2)$).

La probabilidad de que le hagan cinco goles es:

$$P(X=5)=\binom{20}{5}\times 0.2^5\times 0.8^{15}=0.175$$

La probabilidad de que no le hagan goles es:

$$P(X=0)=\binom{20}{0}\times 0.2^0\times 0.8^{20}=0.012$$

La probabilidad de que le hagan algún gol es:

$$P(X\geq 1)=1-P(X=0)=1-\binom{20}{0}\times 0.2^0\times 0.8^{20}=0.988$$

Ejemplo,

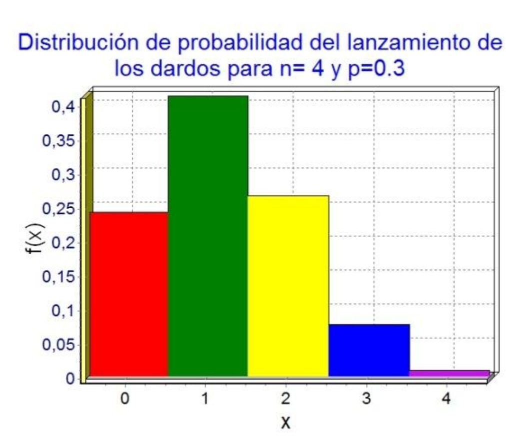
Se sabe que la probabilidad de que Juan acierte en un blanco con un dardo es 0.3. Sea X la variable aleatoria que representa el número de veces que Juan da en el blanco. Si Juan dispara cuatro dardos independientemente. Hallar la función de probabilidad de X.

La variable aleatoria X tiene una distribución binomial de parámetros 4 y 0.3 ($X\sim\mathcal{B}(4, 0.3)$). Por consiguiente

$$P(X=0)=\binom{4}{0}\times 0.3^0\times 0.7^4=0.2401$$
$$P(X=1)=\binom{4}{1}\times 0.3^1\times 0.7^3=0.4116$$
$$P(X=2)=\binom{4}{2}\times 0.3^2\times 0.7^2=0.2646$$
$$P(X=3)=\binom{4}{3}\times 0.3^3\times 0.7^1=0.0756$$
$$P(X=4)=\binom{4}{4}\times 0.3^4\times 0.7^0=0.0081$$

Tenemos

$$f(x)=P(X=x)=\begin{cases}0.2401 & si\ x=0\\ 0.4116 & si\ x=1\\ 0.2646 & si\ x=2\\ 0.0756 & si\ x=3\\ 0.0081 & si\ x=4\end{cases}$$



A partir de la siguiente expresión podemos calcular algunas probabilidades de la variable X con distribución binomial de parámetros n y p

$$F(x)=P(X\leq x)=\binom{n}{0}p^0q^n+\binom{n}{1}p^1q^{n-1}+\cdots+\binom{n}{x}p^xq^{n-x},\qquad para\ x\leq n$$

Usando esta definición calculemos la probabilidad de que Juan de en el blanco en dos o menos lanzamientos.

En este ejemplo tenemos parámetros 4 y 0.3, por consiguiente

$$F(2) = P(X \leq 2) = P(X = 0) + P(X = 1) + P(X = 2)$$
$$= 0.2401 + 0.4116 + 0.2646 = 0.9163$$

Otra manera de calcular esta probabilidad es usando la tabla T.1. que aparece en el apéndice A.

En esta tabla encontramos el número de repeticiones del experimento (n), el número de éxitos (x) y la probabilidad de éxito (p). Para encontrar la probabilidad que estamos buscando, lo primero que debemos hacer es ubicar en la tabla el número de repeticiones (n), luego para este valor n, ubicamos el número de éxitos (x), luego nos deslizamos horizontalmente hasta encontrarnos con la probabilidad de éxito p.

En este caso tenemos $n = 4,\ x = 2$ y $p = 0.3$.

Tabla 4.2. Parte de la tabla de la función distribución binomial acumulada

		p										
	X	0.05	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	0.95
$n=1$	0	0.9500	0.9000	0.8000	0.7000	0.6000	0.5000	0.4000	0.3000	0.2000	0.1000	0.0500
	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$n=2$	0	0.9025	0.8100	0.6400	0.4900	0.3600	0.2500	0.1600	0.0900	0.0400	0.0100	0.0025
	1	0.9975	0.9900	0.9600	0.9100	0.8400	0.7500	0.6400	0.5100	0.3600	0.1900	0.0975
	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$n=3$	0	0.8574	0.7290	0.5120	0.3430	0.2160	0.1250	0.0640	0.0270	0.0080	0.0010	0.0001
	1	0.9928	0.9720	0.8960	0.7840	0.6480	0.5000	0.3520	0.2160	0.1040	0.0280	0.0073
	2	0.9999	0.9990	0.9920	0.9730	0.9360	0.8750	0.7840	0.6570	0.4880	0.2710	0.1426
	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
$n=4$	0	0.8145	0.6561	0.4096	0.2401	0.1296	0.0625	0.0256	0.0081	0.0016	0.0001	0.0000
	1	0.9860	0.9477	0.8192	0.6517	0.4752	0.3125	0.1792	0.0837	0.0272	0.0037	0.0005
	2	0.9995	0.9963	0.9728	0.9163	0.8208	0.6875	0.5248	0.3483	0.1808	0.0523	0.0140
	3	1.0000	0.9999	0.9984	0.9919	0.9744	0.9375	0.8704	0.7599	0.5904	0.3439	0.1855
	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

$$F(2) = P(X \leq 2) = 0.9163$$

Ejemplo,

En una disquera, el proceso de grabación produce 5% de unidades defectuosas. Se seleccionan aleatoriamente con reemplazo 8 discos y todos tienen la misma posibilidad de ser seleccionados. ¿Cuál es la probabilidad de que se encuentren 4 ó más discos defectuosos?

Sea X la variable aleatoria que representa el número de discos defectuosos. Podemos observar que el problema se distribuye en forma binomial con parámetros 8 y 0.05

La probabilidad de encontrar 4 o más discos defectuosos es

$$P(X \geq 4) = 1 - P(X \leq 3)$$

Usando la tabla encontramos que $P(X \leq 3) = 0.9996$, por consiguiente

$$P(X \geq 4) = 1 - 0.9996 = 0.0004.$$

4.3.1. Valor esperado y varianza de la distribución binomial

Si tenemos una variable aleatoria X con distribución binomial de parámetros n y p. El valor esperado, la varianza y la desviación estándar de X son:

$$\mu = np; \ \sigma^2 = npq; \ \sigma = \sqrt{npq}$$

Ejemplo,

La probabilidad de que un estudiante obtenga el título de licenciado en matemáticas es 0.25. Si se matriculan 30 estudiantes determine el valor esperado y la desviación estándar de los estudiantes que se gradúan

$$\mu = np = 30 \times 0.25 = 7.5$$

$$\sigma = \sqrt{npq} = \sqrt{30 \times 0.25 \times 0.75} = 2.371$$

Se espera que se gradúen 7.5 estudiantes, con una desviación de 2.371

Ejemplo,

En una urna hay 20 balotas, 5 negras y el resto verdes. Se elige una balota al azar y el proceso se repite ocho veces, devolviendo la balota a la urna después de sacarla. Determine cuantas balotas negras se espera que salgan una vez realizado el proceso y calcule la desviación estándar.

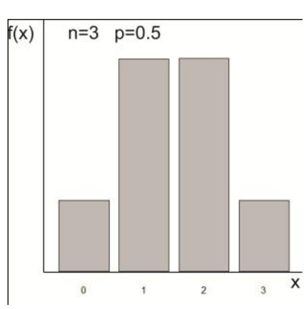
$$\mu = 8 \times \frac{1}{4} = 2$$

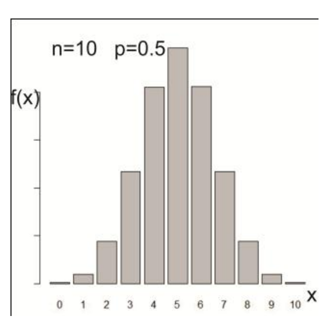
$$\sigma = \sqrt{8 \times \frac{1}{4} \times \frac{3}{4}} = 1.225$$

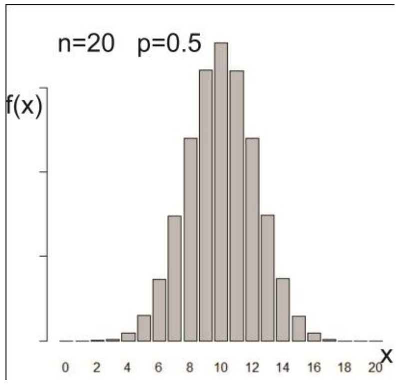
Se espera que se saquen 2 balotas negras, con una desviación 1.225

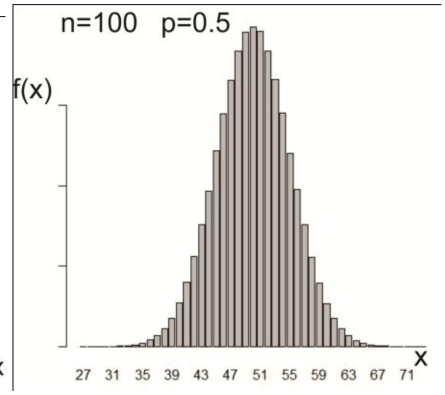
4.4. Distribución normal

Veamos el comportamiento de una variable aleatoria binomial X con probabilidad de éxito $p = 0.5$ a medida que se aumenta n

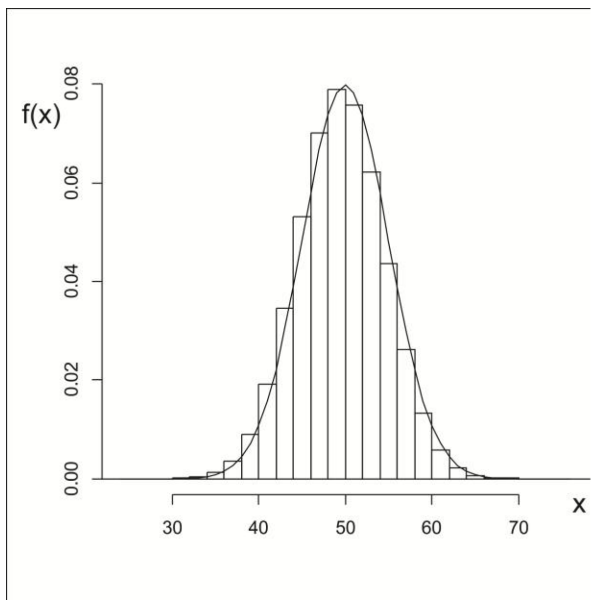








Podemos observar que a medida que n aumenta el diagrama de barras se asemeja a una curva en forma de campana, llamada curva normal o campana de Gauss. Vemos la siguiente gráfica

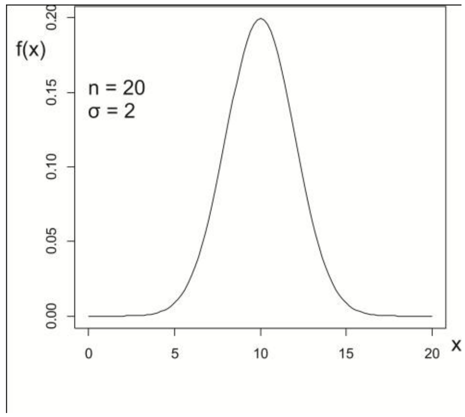


Esta curva corresponde a la llamada distribución normal de probabilidad. Esta distribución es una de las que más se usa en la teoría de probabilidad, puesto que un gran número de variables aleatorias observadas en la naturaleza poseen una distribución que se aproxima a la forma de campana. La distribución normal es continua, su función de densidad es

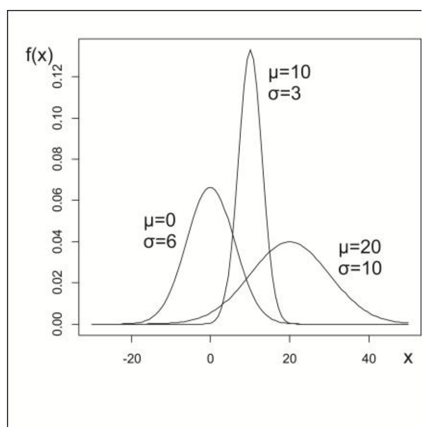
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]};\ x \in \mathbb{R}, \qquad \mu \in \mathbb{R}, \qquad \sigma > 0$$

donde μ y σ son los parámetros que representan la media y la desviación estándar de la población. Se escribe $X \sim N(\mu, \sigma)$ para mostrar que X tiene una distribución normal de parámetros μ y σ.

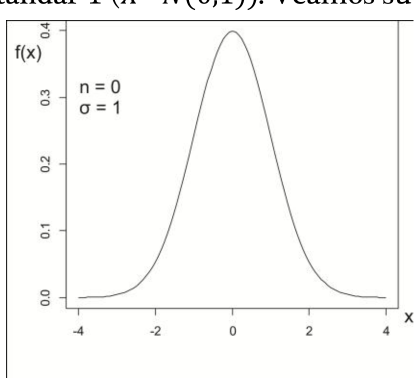
Por ejemplo la siguiente grafica representa una distribución normal con $\mu = 10$ y $\sigma = 2$



En las gráficas de una distribución normal con media μ y desviación estándar σ, la media localiza el centro de la distribución, lo cual implica que la gráfica es simétrica respecto a μ. El área bajo la curva de una distribución normal es igual a 1, por consiguiente a la derecha de la media se encuentra un área de 0.5 y a la izquierda de la media se encuentra un área de 0.5. La forma de la curva está determinada por σ, para los valores grandes de σ reducen la altura de la curva, mientras que los valores pequeños de σ aumenta la altura de la curva. Veamos algunas gráficas en las que cambiamos la media y cambiamos la desviación estándar.



Una de las distribuciones normales más usadas es la normal estándar. Esta distribución tiene media 0 y desviación estándar 1 ($X \sim N(0,1)$). Veamos su gráfica



Para hacer el cálculo de las probabilidades en una distribución normal estándar se usará la tabla T.2. que se encuentra en el apéndice A. Esta tabla está diseñada para encontrar las probabilidades en una distribución de media 0 y desviación estándar 1, por consiguiente para poder usarla en una distribución de parámetros diferentes, se debe realizar una estandarización, es decir definir una nueva variable que tenga media 0 ($\mu = 0$) y desviación estándar 1 ($\sigma = 1$).
La nueva variable aleatoria normal estandarizada, Z, se define como

$$Z = \frac{X - \mu}{\sigma}$$

Cuya función de distribución acumulada se denota como $\phi(Z) = P(X \leq x)$, y se calcula por medio da la tabla T.2. En esta tabla en la primera columna se encuentran los valores de Z con una cifra decimal, el segundo decimal aparece en la primera fila. Por ejemplo si $Z = 1.14$ entonces $\phi(1.14) = 0.87286$

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	$\boxed{0.87286}$	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147

Tabla 4.3. Parte de la tabla de la distribución normal acumulada

Ejemplo,
El peso en kilogramos de los habitantes de una determinada población sigue una distribución normal de media 60 kg y una desviación estándar de 5 Kg. Calcule la probabilidad de que un individuo de la población pese 70 Kg o menos.
La variable aleatoria X es el peso en kilogramos de los habitantes, y tiene una distribución normal de parámetros 60 y 5 ($X \sim N(60, 5)$)
Queremos calcular $P(X \leq 70)$

$$F(70) = P(X \leq 70) = \phi(Z) = \phi\left(\frac{70 - 60}{5}\right) = \phi(2) = 0.9772$$

Ejemplo,

Se estima que la temperatura máxima en verano en una ciudad, presenta una distribución normal, con media 35°C y desviación estándar 6°C. Calcule la probabilidad de que la temperatura se encuentre entre 29°C y 32°C.

La variable aleatoria X corresponde a la temperatura máxima, y la distribución normal tiene parámetros 35 y 6 ($X \sim N(35, 6)$)

por consiguiente

$$
\begin{aligned}
P(29 \leq X \leq 32) &= P\left(\frac{29-35}{6} \leq \frac{X-35}{6} \leq \frac{32-35}{6}\right) \\
&= P(-1 \leq Z \leq -0.5) \\
&= \phi(-0.5) - \phi(-1) \\
&= 0.3085 - 0.1587 \\
&= 0.1498
\end{aligned}
$$

Ejemplo,

Supongamos que lanzamos una moneda corriente 100 veces consecutivas y queremos calcular la probabilidad de obtener 45 caras o menos.

La variable aleatoria X corresponde al número de caras obtenidas. Es claro que esta variable tiene una distribución binomial de parámetros 100 y 0.5 ($X \sim \mathcal{B}(100,\ 0.5)$). Para calcular la probabilidad $P(X \leq 45)$ encontramos algunas dificultades. Si usamos la función de masa, estamos obligados a realizar cómputos extensos, y si utilizamos la tabla T1, las probabilidades para $n = 45$ no se encuentran, puesto que la tabla solamente tiene ciertos valores de n y p disponibles. Obsérvese que estas dificultades las encontramos cuando el número de repeticiones (n) aumenta. Como vimos en los diagramas de barras, al comienzo de esta sección, cuando el número de repeticiones (n) aumenta, las barras se aproximan a una curva normal, por consiguiente la distribución normal proporciona una aproximación a la distribución binomial. Cuando n sea alto y la probabilidad de éxito se aproxime a $1/2$, el comportamiento de una distribución binomial de parámetros n y p se aproxima a una distribución normal de media $\mu = np$ y desviación estándar $\sigma = \sqrt{npq}$. Esta aproximación es considerada como buena, si $np \geq 5$ y $nq \geq 5$.

Veamos cómo se usa la aproximación en el caso de los 100 lanzamientos de la moneda.

Tenemos una distribución binomial de parámetros 100 y 0.5

Como

$$\mu = np = 100 \times 0.5 = 50 \text{ y } \sigma = \sqrt{npq} = \sqrt{100 \times 0.5 \times 0.5} = 5$$

Entonces se aproxima a una distribución normal de parámetros 50 y 5, por consiguiente

$$F(45) = P(X \leq 45) = \phi(Z) = \phi\left(\frac{45-50}{5}\right) = \phi(-1) = 0.1587$$

La probabilidad de obtener 45 caras o menos es 0.1587.

Veamos otro ejemplo,

En una ciudad tres de cada cinco familias tienen lavadora. Se eligen al azar 120 familias. Calcule la probabilidad de que más de 50 familias tengan lavadora.

Tenemos una distribución binomial de parámetros 120 y ${}^{3}/_{5}$

Como

$$\mu = np = 120 \times {}^{3}/_{5} = 72 \text{ y } \sigma = \sqrt{npq} = \sqrt{120 \times {}^{3}/_{5} \times {}^{2}/_{5}} = 28.8$$

Entonces se aproxima a una distribución normal de parámetros 72 y 28.8, por consiguiente

$$P(X > 50) = 1 - P(X \leq 50) = 1 - \phi(Z) = 1 - \phi\left(\frac{50 - 72}{28.8}\right)$$
$$= 1 - \phi(-0.76) = 1 - 0.2236 = 0.7764$$

La probabilidad de que más de 50 familias tengan lavadora es 0.7764.

4.6. Actividades

1. Indique para cada una de las siguientes variables aleatorias si son discretas o continuas. Haga las aclaraciones que considere necesarias.
   a) El número que sale al tirar un dado.
   b) La cantidad de caras que salen al tirar 5 monedas.
   c) La cantidad de accidentes automovilísticos en una ciudad por mes.
   d) Peso de una naranja.
   e) Diámetro de una arandela.
   f) Edad de una persona.
2. La probabilidad de que un tirador de en el blanco es 0.3. Si dispara 8 veces ¿cuál es la probabilidad de que acierte exactamente 4 veces? ¿Cuál es la probabilidad de que acierte por lo menos en una vez?
3. Determine la probabilidad de que en una familia que tiene cinco hijos, tres de ellos sean niñas
4. En una fábrica de electrodomésticos, la probabilidad de que uno sea defectuoso, es 0.1. Si se revisan 10 electrodomésticos, calcule:
   a) La probabilidad de que ninguno sea defectuoso.
   b) La probabilidad de que haya más de 3 electrodomésticos defectuosos.
   c) Cuantos electrodomésticos se espera que salgan defectuosos.
5. Una persona apuesta que en 8 lanzamientos de una moneda corriente obtendrá 5 caras o más. ¿Cuál es su probabilidad de ganar?
6. Entre dos amigos se proponen el siguiente juego para elegir su comida. Lanzamos una moneda corriente 20 veces. Si aparece cara 9, 10 u 11 veces, comemos hamburguesa, de otra manera comemos pizza. ¿Cuál es la probabilidad de que los amigos coman pizza?

7. Un laboratorio farmacéutico asegura que uno de sus medicamentos causa efectos secundarios a 3 de cada 80 pacientes. Para verificar esta afirmación, otro laboratorio elige al azar a 5 pacientes, a los que se les proporciona el medicamento. ¿Cuál es la probabilidad de los siguientes eventos?
   a) Ningún paciente tenga efectos secundarios.
   b) Al menos dos tengan efectos secundarios.
   c) Si eligen 120 pacientes al azar. ¿Cuántos pacientes se espera que sufran efectos secundarios?
8. La media de los pesos de 450 personas en un conjunto residencial es 60 kg y la desviación estándar 4 kg. Suponiendo que los pesos tienen una distribución normal, hallar cuántos personas pesan:
   a) Entre 45 kg y 65 kg.
   b) Más de 70 kg.
   c) 50 kg o menos.
9. Una máquina dispensadora de gaseosa sirve un promedio de 150 mm por vaso. Si la cantidad de líquido se distribuye normalmente con una desviación estándar igual a 13 mm.
   a) ¿Cuál es la probabilidad de que un vaso contenga entre 140 mm y 160 mm?
   b) Si se utilizan vasos de 200 mililitros ¿Cuál es la probabilidad de que la gaseosa se derrame?
10. Por experiencia se sabe que los resultados de los exámenes de matemáticas realizados a un grupo de estudiantes siguen una distribución normal con media 70 y desviación estándar 40. ¿Cuál es la probabilidad de que un estudiante que presenta el examen obtenga más de 60?
11. Un profesor sale todos los días de su casa hacia el colegio. El tiempo promedio que gasta en su desplazamiento es 30 minutos, con una desviación estándar de 5 minutos. Si los tiempos de desplazamiento están distribuidos normalmente. ¿Cuál es la probabilidad de que se demore 20 minutos o menos en ir de su casa al colegio? Si sale de su casa a las 6:25 am y se empiezan clases a las 7: 00 am, ¿cuál es la probabilidad de que llegue tarde a su trabajo?
12. El 30% de los estudiantes de un colegio tienen gripe. Si se seleccionan 40 estudiantes al azar. Calcule la probabilidad de que:
    a) Exactamente 21 estudiantes tengan gripa
    b) Entre 15 y 25 estudiantes estén enfermos
13. Un determinado banco aprueba el 60% de los préstamos solicitados en un año. Si en un año llegan 2100 solicitudes, ¿cuál es la probabilidad de que se aprueben más de 1500 préstamos?

14. Un examen consta de 90 preguntas de selección múltiple con única respuesta, cada pregunta tiene cuatro respuestas. Calcule la probabilidad de que un estudiante que responde al azar acierte más de 65 preguntas.

Apéndice A

Tablas estadísticas: T.1. Probabilidades Binomiales

n	x	0,050	0,100	0,150	0,200	0,250	0,300	0,400	0,450	0,500	0,600	0,700	0,750	0,800	0,900	0,950
1	0	0,950	0,900	0,850	0,800	0,750	0,700	0,600	0,550	0,500	0,400	0,300	0,250	0,200	0,100	0,050
2	0	0,903	0,810	0,723	0,640	0,563	0,490	0,360	0,303	0,250	0,160	0,090	0,063	0,040	0,010	0,003
	1	0,998	0,990	0,978	0,960	0,938	0,910	0,840	0,798	0,750	0,640	0,510	0,438	0,360	0,190	0,098
3	0	0,857	0,729	0,614	0,512	0,422	0,343	0,216	0,166	0,125	0,064	0,027	0,016	0,008	0,001	0,000
	1	0,993	0,972	0,939	0,896	0,844	0,784	0,648	0,575	0,500	0,352	0,216	0,156	0,104	0,028	0,007
	2	1,000	0,999	0,997	0,992	0,984	0,973	0,936	0,909	0,875	0,784	0,657	0,578	0,488	0,271	0,143
4	0	0,815	0,656	0,522	0,410	0,316	0,240	0,130	0,092	0,063	0,026	0,008	0,004	0,002	0,000	0,000
	1	0,986	0,948	0,891	0,819	0,738	0,652	0,475	0,391	0,313	0,179	0,084	0,051	0,027	0,004	0,001
	2	1,000	0,996	0,988	0,973	0,949	0,916	0,821	0,759	0,688	0,525	0,348	0,262	0,181	0,052	0,014
	3	1,000	1,000	1,000	0,998	0,996	0,992	0,974	0,959	0,938	0,870	0,760	0,684	0,590	0,344	0,186
5	0	0,774	0,591	0,444	0,328	0,237	0,168	0,078	0,050	0,031	0,010	0,002	0,001	0,000	0,000	0,000
	1	0,977	0,919	0,835	0,737	0,633	0,528	0,337	0,256	0,188	0,087	0,031	0,016	0,007	0,001	0,000
	2	0,999	0,991	0,973	0,942	0,897	0,837	0,683	0,593	0,500	0,317	0,163	0,104	0,058	0,009	0,001
	3	1,000	1,000	0,998	0,993	0,984	0,969	0,913	0,869	0,813	0,663	0,472	0,367	0,263	0,082	0,023
	4	1,000	1,000	1,000	1,000	0,999	0,998	0,990	0,982	0,969	0,922	0,832	0,763	0,672	0,410	0,226
6	0	0,735	0,531	0,377	0,262	0,178	0,118	0,047	0,028	0,016	0,004	0,001	0,000	0,000	0,000	0,000
	1	0,967	0,886	0,777	0,655	0,534	0,420	0,233	0,164	0,109	0,041	0,011	0,005	0,002	0,000	0,000
	2	0,998	0,984	0,953	0,901	0,831	0,744	0,544	0,442	0,344	0,179	0,071	0,038	0,017	0,001	0,000
	3	1,000	0,999	0,994	0,983	0,962	0,930	0,821	0,745	0,656	0,456	0,256	0,169	0,099	0,016	0,002
	4	1,000	1,000	1,000	0,998	0,995	0,989	0,959	0,931	0,891	0,767	0,580	0,466	0,345	0,114	0,033
	5	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,992	0,984	0,953	0,882	0,822	0,738	0,469	0,265
7	0	0,698	0,478	0,321	0,210	0,134	0,082	0,028	0,015	0,008	0,002	0,000	0,000	0,000	0,000	0,000
	1	0,956	0,850	0,717	0,577	0,445	0,329	0,159	0,102	0,063	0,019	0,004	0,001	0,000	0,000	0,000
	2	0,996	0,974	0,926	0,852	0,756	0,647	0,420	0,316	0,227	0,096	0,029	0,013	0,005	0,000	0,000
	3	1,000	0,997	0,988	0,967	0,929	0,874	0,710	0,608	0,500	0,290	0,126	0,071	0,033	0,003	0,000
	4	1,000	1,000	0,999	0,995	0,987	0,971	0,904	0,847	0,773	0,580	0,353	0,244	0,148	0,026	0,004
	5	1,000	1,000	1,000	1,000	0,999	0,996	0,981	0,964	0,938	0,841	0,671	0,555	0,423	0,150	0,044
	6	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,996	0,992	0,972	0,918	0,867	0,790	0,522	0,302
8	0	0,663	0,431	0,273	0,168	0,100	0,058	0,017	0,008	0,004	0,001	0,000	0,000	0,000	0,000	0,000
	1	0,943	0,813	0,657	0,503	0,367	0,255	0,106	0,063	0,035	0,009	0,001	0,000	0,000	0,000	0,000
	2	0,994	0,962	0,895	0,797	0,679	0,552	0,315	0,220	0,145	0,050	0,011	0,004	0,001	0,000	0,000
	3	1,000	0,995	0,979	0,944	0,886	0,806	0,594	0,477	0,363	0,174	0,058	0,027	0,010	0,000	0,000
	4	1,000	1,000	0,997	0,990	0,973	0,942	0,826	0,740	0,637	0,406	0,194	0,114	0,056	0,005	0,000
	5	1,000	1,000	1,000	0,999	0,996	0,989	0,950	0,912	0,856	0,685	0,448	0,322	0,203	0,038	0,006
	6	1,000	1,000	1,000	1,000	1,000	0,999	0,992	0,982	0,965	0,894	0,745	0,633	0,497	0,187	0,057
	7	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,998	0,996	0,983	0,942	0,900	0,832	0,570	0,337
9	0	0,630	0,387	0,232	0,134	0,075	0,040	0,010	0,005	0,002	0,000	0,000	0,000	0,000	0,000	0,000

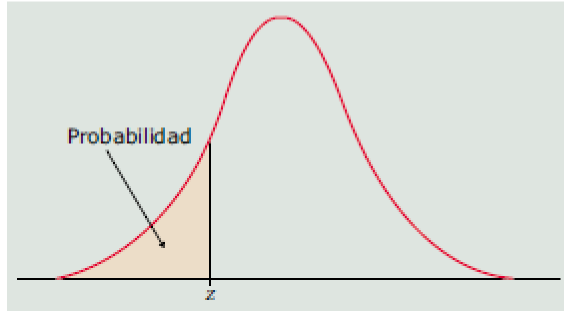
n	x	0,050	0,100	0,150	0,200	0,250	0,300	0,400	0,450	0,500	0,600	0,700	0,750	0,800	0,900	0,950
	1	0,929	0,775	0,600	0,436	0,300	0,196	0,071	0,039	0,020	0,004	0,000	0,000	0,000	0,000	0,000
	2	0,992	0,947	0,859	0,738	0,601	0,463	0,232	0,150	0,090	0,025	0,004	0,001	0,000	0,000	0,000
	3	0,999	0,992	0,966	0,914	0,834	0,730	0,483	0,361	0,254	0,099	0,025	0,010	0,003	0,000	0,000
	4	1,000	0,999	0,994	0,980	0,951	0,901	0,733	0,621	0,500	0,267	0,099	0,049	0,020	0,001	0,000
	5	1,000	1,000	0,999	0,997	0,990	0,975	0,901	0,834	0,746	0,517	0,270	0,166	0,086	0,008	0,001
	6	1,000	1,000	1,000	1,000	0,999	0,996	0,975	0,950	0,910	0,768	0,537	0,399	0,262	0,053	0,008
	7	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,991	0,981	0,930	0,804	0,700	0,564	0,225	0,071
	8	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,998	0,990	0,960	0,925	0,866	0,613	0,370
10	0	0,599	0,349	0,197	0,107	0,056	0,028	0,006	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,914	0,736	0,544	0,376	0,244	0,149	0,046	0,023	0,011	0,002	0,000	0,000	0,000	0,000	0,000
	2	0,989	0,930	0,820	0,678	0,526	0,383	0,167	0,100	0,055	0,012	0,002	0,000	0,000	0,000	0,000
	3	0,999	0,987	0,950	0,879	0,776	0,650	0,382	0,266	0,172	0,055	0,011	0,004	0,001	0,000	0,000
	4	1,000	0,998	0,990	0,967	0,922	0,850	0,633	0,504	0,377	0,166	0,047	0,020	0,006	0,000	0,000
	5	1,000	1,000	0,999	0,994	0,980	0,953	0,834	0,738	0,623	0,367	0,150	0,078	0,033	0,002	0,000
	6	1,000	1,000	1,000	0,999	0,997	0,989	0,945	0,898	0,828	0,618	0,350	0,224	0,121	0,013	0,001
	7	1,000	1,000	1,000	1,000	1,000	0,998	0,988	0,973	0,945	0,833	0,617	0,474	0,322	0,070	0,012
	8	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,996	0,989	0,954	0,851	0,756	0,624	0,264	0,086
	9	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,994	0,972	0,944	0,893	0,651	0,401
11	0	0,569	0,314	0,167	0,086	0,042	0,020	0,004	0,001	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,898	0,697	0,492	0,322	0,197	0,113	0,030	0,014	0,006	0,001	0,000	0,000	0,000	0,000	0,000
	2	0,985	0,910	0,779	0,617	0,455	0,313	0,119	0,065	0,033	0,006	0,001	0,000	0,000	0,000	0,000
	3	0,998	0,982	0,931	0,839	0,713	0,570	0,296	0,191	0,113	0,029	0,004	0,001	0,000	0,000	0,000
	4	1,000	0,997	0,984	0,950	0,885	0,790	0,533	0,397	0,274	0,099	0,022	0,008	0,002	0,000	0,000
	5	1,000	1,000	0,997	0,988	0,966	0,922	0,754	0,633	0,500	0,247	0,078	0,034	0,012	0,000	0,000
	6	1,000	1,000	1,000	0,998	0,992	0,978	0,901	0,826	0,726	0,467	0,210	0,115	0,050	0,003	0,000
	7	1,000	1,000	1,000	1,000	0,999	0,996	0,971	0,939	0,887	0,704	0,430	0,287	0,161	0,019	0,002
	8	1,000	1,000	1,000	1,000	1,000	0,999	0,994	0,985	0,967	0,881	0,687	0,545	0,383	0,090	0,015
	9	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,998	0,994	0,970	0,887	0,803	0,678	0,303	0,102
	10	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,980	0,958	0,914	0,686	0,431
12	0	0,540	0,282	0,142	0,069	0,032	0,014	0,002	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,882	0,659	0,444	0,275	0,158	0,085	0,020	0,008	0,003	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,980	0,889	0,736	0,558	0,391	0,253	0,083	0,042	0,019	0,003	0,000	0,000	0,000	0,000	0,000
	3	0,998	0,974	0,908	0,795	0,649	0,493	0,225	0,135	0,073	0,015	0,002	0,000	0,000	0,000	0,000
	4	1,000	0,996	0,976	0,927	0,842	0,724	0,438	0,304	0,194	0,057	0,010	0,003	0,001	0,000	0,000
	5	1,000	1,000	0,995	0,981	0,946	0,882	0,665	0,527	0,387	0,158	0,039	0,014	0,004	0,000	0,000
	6	1,000	1,000	0,999	0,996	0,986	0,961	0,842	0,739	0,613	0,335	0,118	0,054	0,019	0,001	0,000
	7	1,000	1,000	1,000	0,999	0,997	0,991	0,943	0,888	0,806	0,562	0,276	0,158	0,073	0,004	0,000
	8	1,000	1,000	1,000	1,000	1,000	0,998	0,985	0,964	0,927	0,775	0,508	0,351	0,205	0,026	0,002
	9	1,000	1,000	1,000	1,000	1,000	1,000	0,997	0,992	0,981	0,917	0,747	0,609	0,442	0,111	0,020

n	x	0,050	0,100	0,150	0,200	0,250	0,300	0,400	0,450	0,500	0,600	0,700	0,750	0,800	0,900	0,950
	10	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,980	0,915	0,842	0,725	0,341	0,118
	11	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,986	0,968	0,931	0,718	0,460
13	0	0,513	0,254	0,121	0,055	0,024	0,010	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,865	0,621	0,398	0,234	0,127	0,064	0,013	0,005	0,002	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,976	0,866	0,692	0,502	0,333	0,203	0,058	0,027	0,011	0,001	0,000	0,000	0,000	0,000	0,000
	3	0,997	0,966	0,882	0,747	0,584	0,421	0,169	0,093	0,046	0,008	0,001	0,000	0,000	0,000	0,000
	4	1,000	0,994	0,966	0,901	0,794	0,654	0,353	0,228	0,133	0,032	0,004	0,001	0,000	0,000	0,000
	5	1,000	0,999	0,993	0,970	0,920	0,835	0,574	0,427	0,291	0,098	0,018	0,006	0,001	0,000	0,000
	6	1,000	1,000	0,999	0,993	0,976	0,938	0,771	0,644	0,500	0,229	0,062	0,024	0,007	0,000	0,000
	7	1,000	1,000	1,000	0,999	0,994	0,982	0,902	0,821	0,710	0,426	0,165	0,080	0,030	0,001	0,000
	8	1,000	1,000	1,000	1,000	0,999	0,996	0,968	0,930	0,867	0,647	0,346	0,206	0,099	0,007	0,000
	9	1,000	1,000	1,000	1,000	1,000	0,999	0,992	0,980	0,954	0,831	0,579	0,416	0,253	0,034	0,003
	10	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,989	0,942	0,798	0,667	0,498	0,134	0,025
	11	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,987	0,936	0,873	0,766	0,379	0,135
	12	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,990	0,976	0,945	0,746	0,487
14	0	0,488	0,229	0,103	0,044	0,018	0,007	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,847	0,585	0,357	0,198	0,101	0,048	0,008	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,970	0,842	0,648	0,448	0,281	0,161	0,040	0,017	0,007	0,001	0,000	0,000	0,000	0,000	0,000
	3	0,996	0,956	0,854	0,698	0,521	0,355	0,124	0,063	0,029	0,004	0,000	0,000	0,000	0,000	0,000
	4	1,000	0,991	0,953	0,870	0,742	0,584	0,279	0,167	0,090	0,018	0,002	0,000	0,000	0,000	0,000
	5	1,000	0,999	0,989	0,956	0,888	0,781	0,486	0,337	0,212	0,058	0,008	0,002	0,000	0,000	0,000
	6	1,000	1,000	0,998	0,988	0,962	0,907	0,693	0,546	0,395	0,150	0,032	0,010	0,002	0,000	0,000
	7	1,000	1,000	1,000	0,998	0,990	0,969	0,850	0,741	0,605	0,308	0,093	0,038	0,012	0,000	0,000
	8	1,000	1,000	1,000	1,000	0,998	0,992	0,942	0,881	0,788	0,514	0,220	0,112	0,044	0,002	0,000
	9	1,000	1,000	1,000	1,000	1,000	0,998	0,983	0,957	0,910	0,721	0,416	0,259	0,130	0,009	0,000
	10	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,989	0,971	0,876	0,645	0,479	0,302	0,044	0,004
	11	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,998	0,994	0,960	0,839	0,719	0,552	0,158	0,030
	12	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,992	0,953	0,899	0,802	0,415	0,153
	13	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,993	0,982	0,956	0,771	0,512
15	0	0,463	0,206	0,087	0,035	0,013	0,005	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,829	0,549	0,319	0,167	0,080	0,035	0,005	0,002	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,964	0,816	0,604	0,398	0,236	0,127	0,027	0,011	0,004	0,000	0,000	0,000	0,000	0,000	0,000
	3	0,995	0,944	0,823	0,648	0,461	0,297	0,091	0,042	0,018	0,002	0,000	0,000	0,000	0,000	0,000
	4	0,999	0,987	0,938	0,836	0,687	0,516	0,217	0,120	0,059	0,009	0,001	0,000	0,000	0,000	0,000
	5	1,000	0,998	0,983	0,939	0,852	0,722	0,403	0,261	0,151	0,034	0,004	0,001	0,000	0,000	0,000
	6	1,000	1,000	0,996	0,982	0,943	0,869	0,610	0,452	0,304	0,095	0,015	0,004	0,001	0,000	0,000
	7	1,000	1,000	0,999	0,996	0,983	0,950	0,787	0,654	0,500	0,213	0,050	0,017	0,004	0,000	0,000
	8	1,000	1,000	1,000	0,999	0,996	0,985	0,905	0,818	0,696	0,390	0,131	0,057	0,018	0,000	0,000
	9	1,000	1,000	1,000	1,000	0,999	0,996	0,966	0,923	0,849	0,597	0,278	0,148	0,061	0,002	0,000

n	x	0,050	0,100	0,150	0,200	0,250	0,300	0,400	0,450	0,500	0,600	0,700	0,750	0,800	0,900	0,950
	10	1,000	1,000	1,000	1,000	1,000	0,999	0,991	0,975	0,941	0,783	0,485	0,314	0,164	0,013	0,001
	11	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,994	0,982	0,910	0,703	0,539	0,352	0,056	0,006
	12	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,973	0,873	0,764	0,602	0,184	0,036
	13	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,965	0,920	0,833	0,451	0,171
	14	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,987	0,965	0,794	0,537
16	0	0,440	0,185	0,074	0,028	0,010	0,003	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,811	0,515	0,284	0,141	0,064	0,026	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,957	0,789	0,561	0,352	0,197	0,099	0,018	0,007	0,002	0,000	0,000	0,000	0,000	0,000	0,000
	3	0,993	0,932	0,790	0,598	0,405	0,246	0,065	0,028	0,011	0,001	0,000	0,000	0,000	0,000	0,000
	4	0,999	0,983	0,921	0,798	0,630	0,450	0,167	0,085	0,038	0,005	0,000	0,000	0,000	0,000	0,000
	5	1,000	0,997	0,977	0,918	0,810	0,660	0,329	0,198	0,105	0,019	0,002	0,000	0,000	0,000	0,000
	6	1,000	1,000	0,994	0,973	0,920	0,825	0,527	0,366	0,227	0,058	0,007	0,002	0,000	0,000	0,000
	7	1,000	1,000	0,999	0,993	0,973	0,926	0,716	0,563	0,402	0,142	0,026	0,008	0,002	0,000	0,000
	8	1,000	1,000	1,000	0,999	0,993	0,974	0,858	0,744	0,598	0,284	0,074	0,027	0,007	0,000	0,000
	9	1,000	1,000	1,000	1,000	0,998	0,993	0,942	0,876	0,773	0,473	0,175	0,080	0,027	0,001	0,000
	10	1,000	1,000	1,000	1,000	1,000	0,998	0,981	0,951	0,895	0,671	0,340	0,190	0,082	0,003	0,000
	11	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,985	0,962	0,833	0,550	0,370	0,202	0,017	0,001
	12	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,989	0,935	0,754	0,595	0,402	0,068	0,007
	13	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,998	0,982	0,901	0,803	0,648	0,211	0,043
	14	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,997	0,974	0,937	0,859	0,485	0,189
	15	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,997	0,990	0,972	0,815	0,560
17	0	0,418	0,167	0,063	0,023	0,008	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,792	0,482	0,253	0,118	0,050	0,019	0,002	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,950	0,762	0,520	0,310	0,164	0,077	0,012	0,004	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	3	0,991	0,917	0,756	0,549	0,353	0,202	0,046	0,018	0,006	0,001	0,000	0,000	0,000	0,000	0,000
	4	0,999	0,978	0,901	0,758	0,574	0,389	0,126	0,060	0,025	0,003	0,000	0,000	0,000	0,000	0,000
	5	1,000	0,995	0,968	0,894	0,765	0,597	0,264	0,147	0,072	0,011	0,001	0,000	0,000	0,000	0,000
	6	1,000	0,999	0,992	0,962	0,893	0,775	0,448	0,290	0,166	0,035	0,003	0,001	0,000	0,000	0,000
	7	1,000	1,000	0,998	0,989	0,960	0,895	0,641	0,474	0,315	0,092	0,013	0,003	0,001	0,000	0,000
	8	1,000	1,000	1,000	0,997	0,988	0,960	0,801	0,663	0,500	0,199	0,040	0,012	0,003	0,000	0,000
	9	1,000	1,000	1,000	1,000	0,997	0,987	0,908	0,817	0,686	0,360	0,105	0,040	0,011	0,000	0,000
	10	1,000	1,000	1,000	1,000	0,999	0,997	0,965	0,917	0,834	0,552	0,225	0,107	0,038	0,001	0,000
	11	1,000	1,000	1,000	1,000	1,000	0,999	0,989	0,970	0,928	0,736	0,403	0,235	0,106	0,005	0,000
	12	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,991	0,976	0,874	0,611	0,426	0,242	0,022	0,001
	13	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,994	0,954	0,798	0,647	0,451	0,083	0,009
	14	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,988	0,923	0,836	0,690	0,238	0,050
	15	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,981	0,950	0,882	0,518	0,208
	16	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,993	0,978	0,833	0,582
18	0	0,397	0,150	0,054	0,018	0,006	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000

n	x	0,050	0,100	0,150	0,200	0,250	0,300	0,400	0,450	0,500	0,600	0,700	0,750	0,800	0,900	0,950
	1	0,774	0,450	0,224	0,099	0,040	0,014	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,942	0,734	0,480	0,271	0,135	0,060	0,008	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	3	0,989	0,902	0,720	0,501	0,306	0,165	0,033	0,012	0,004	0,000	0,000	0,000	0,000	0,000	0,000
	4	0,999	0,972	0,879	0,716	0,519	0,333	0,094	0,041	0,015	0,001	0,000	0,000	0,000	0,000	0,000
	5	1,000	0,994	0,958	0,867	0,718	0,534	0,209	0,108	0,048	0,006	0,000	0,000	0,000	0,000	0,000
	6	1,000	0,999	0,988	0,949	0,861	0,722	0,374	0,226	0,119	0,020	0,001	0,000	0,000	0,000	0,000
	7	1,000	1,000	0,997	0,984	0,943	0,859	0,563	0,392	0,240	0,058	0,006	0,001	0,000	0,000	0,000
	8	1,000	1,000	1,000	0,996	0,981	0,940	0,737	0,578	0,407	0,135	0,021	0,005	0,001	0,000	0,000
	9	1,000	1,000	1,000	0,999	0,995	0,979	0,865	0,747	0,593	0,263	0,060	0,019	0,004	0,000	0,000
	10	1,000	1,000	1,000	1,000	0,999	0,994	0,942	0,872	0,760	0,437	0,141	0,057	0,016	0,000	0,000
	11	1,000	1,000	1,000	1,000	1,000	0,999	0,980	0,946	0,881	0,626	0,278	0,139	0,051	0,001	0,000
	12	1,000	1,000	1,000	1,000	1,000	1,000	0,994	0,982	0,952	0,791	0,466	0,283	0,133	0,006	0,000
	13	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,995	0,985	0,906	0,667	0,481	0,284	0,028	0,002
	14	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,967	0,835	0,694	0,499	0,098	0,011
	15	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,992	0,940	0,865	0,729	0,266	0,058
	16	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,986	0,961	0,901	0,550	0,227
	17	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,994	0,982	0,850	0,603
19	0	0,377	0,135	0,046	0,014	0,004	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,755	0,420	0,199	0,083	0,031	0,010	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,934	0,705	0,441	0,237	0,111	0,046	0,006	0,002	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	3	0,987	0,885	0,684	0,455	0,263	0,133	0,023	0,008	0,002	0,000	0,000	0,000	0,000	0,000	0,000
	4	0,998	0,965	0,856	0,673	0,465	0,282	0,070	0,028	0,010	0,001	0,000	0,000	0,000	0,000	0,000
	5	1,000	0,991	0,946	0,837	0,668	0,474	0,163	0,078	0,032	0,003	0,000	0,000	0,000	0,000	0,000
	6	1,000	0,998	0,984	0,932	0,825	0,666	0,308	0,173	0,084	0,012	0,001	0,000	0,000	0,000	0,000
	7	1,000	1,000	0,996	0,977	0,923	0,818	0,488	0,317	0,180	0,035	0,003	0,001	0,000	0,000	0,000
	8	1,000	1,000	0,999	0,993	0,971	0,916	0,668	0,494	0,324	0,089	0,011	0,002	0,000	0,000	0,000
	9	1,000	1,000	1,000	0,998	0,991	0,967	0,814	0,671	0,500	0,186	0,033	0,009	0,002	0,000	0,000
	10	1,000	1,000	1,000	1,000	0,998	0,990	0,912	0,816	0,676	0,333	0,084	0,029	0,007	0,000	0,000
	11	1,000	1,000	1,000	1,000	1,000	0,997	0,965	0,913	0,820	0,512	0,182	0,078	0,023	0,000	0,000
	12	1,000	1,000	1,000	1,000	1,000	0,999	0,988	0,966	0,917	0,692	0,335	0,175	0,068	0,002	0,000
	13	1,000	1,000	1,000	1,000	1,000	1,000	0,997	0,989	0,968	0,837	0,526	0,332	0,163	0,009	0,000
	14	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,990	0,930	0,718	0,535	0,327	0,035	0,002
	15	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,977	0,867	0,737	0,545	0,115	0,013
	16	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,995	0,954	0,889	0,763	0,295	0,067
	17	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,990	0,969	0,917	0,580	0,245
	18	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,996	0,986	0,865	0,623
20	0	0,359	0,122	0,039	0,012	0,003	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	1	0,736	0,392	0,176	0,069	0,024	0,008	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000	0,000
	2	0,925	0,677	0,405	0,206	0,091	0,036	0,004	0,001	0,000	0,000	0,000	0,000	0,000	0,000	0,000

n	x	0,050	0,100	0,150	0,200	0,250	0,300	0,400	0,450	0,500	0,600	0,700	0,750	0,800	0,900	0,950
	3	0,984	0,867	0,648	0,411	0,225	0,107	0,016	0,005	0,001	0,000	0,000	0,000	0,000	0,000	0,000
	4	0,997	0,957	0,830	0,630	0,415	0,238	0,051	0,019	0,006	0,000	0,000	0,000	0,000	0,000	0,000
	5	1,000	0,989	0,933	0,804	0,617	0,416	0,126	0,055	0,021	0,002	0,000	0,000	0,000	0,000	0,000
	6	1,000	0,998	0,978	0,913	0,786	0,608	0,250	0,130	0,058	0,007	0,000	0,000	0,000	0,000	0,000
	7	1,000	1,000	0,994	0,968	0,898	0,772	0,416	0,252	0,132	0,021	0,001	0,000	0,000	0,000	0,000
	8	1,000	1,000	0,999	0,990	0,959	0,887	0,596	0,414	0,252	0,057	0,005	0,001	0,000	0,000	0,000
	9	1,000	1,000	1,000	0,997	0,986	0,952	0,755	0,591	0,412	0,128	0,017	0,004	0,001	0,000	0,000
	10	1,000	1,000	1,000	0,999	0,996	0,983	0,873	0,751	0,588	0,245	0,048	0,014	0,003	0,000	0,000
	11	1,000	1,000	1,000	1,000	0,999	0,995	0,944	0,869	0,748	0,404	0,113	0,041	0,010	0,000	0,000
	12	1,000	1,000	1,000	1,000	1,000	0,999	0,979	0,942	0,868	0,584	0,228	0,102	0,032	0,000	0,000
	13	1,000	1,000	1,000	1,000	1,000	1,000	0,994	0,979	0,942	0,750	0,392	0,214	0,087	0,002	0,000
	14	1,000	1,000	1,000	1,000	1,000	1,000	0,998	0,994	0,979	0,874	0,584	0,383	0,196	0,011	0,000
	15	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,994	0,949	0,763	0,585	0,370	0,043	0,003
	16	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,984	0,893	0,775	0,589	0,133	0,016
	17	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,996	0,965	0,909	0,794	0,323	0,076
	18	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,992	0,976	0,931	0,608	0,264
	19	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	1,000	0,999	0,997	0,989	0,878	0,642

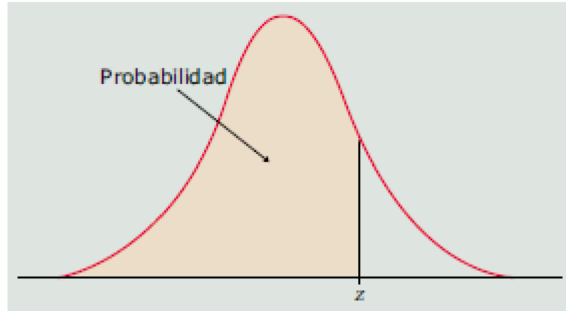


El valor de la tabla para z
es el área bajo la curva
de la normal estándar
a la izquierda de z

T.2. Función de distribución normal estándar acumulada

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
-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	0.5	0,4960	0,4920	0,4880	0,4840	0,4801	0,4761	0,4721	0,4681	0,4641

El valor de la tabla para z
es el área bajo la curva
de la normal estándar
a la izquierda de z



T.2. Función de distribución normal estándar acumulada

Z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0.5	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

Bibliografía

[1] BATANERO, Carmen, DÍAZ, Carmen. Estadística con Proyectos, Departamento de Didáctica de la Matemática, Universidad de Granada, Granada, 2011.

[2] BATANERO, Carmen, Enseñanza de la estadística en los niveles no universitarios: algunos retos para la investigación, Investigación en educación matemática XI, Sociedad Española de Investigación en Educación Matemática, SEIEM, Tenerife, 2007, p93 – 97.

[3] BATANERO, Carmen, GODINO, Juan D., Análisis de datos y su didáctica, Departamento de Didáctica de la Matemática, Universidad de Granada, Granada, 2001.

[4] BATANERO, Carmen, Significados de la probabilidad en la educación secundaria, Relime, México, 2005.

[5] BLANCO CASTAÑEDA, Liliana. Probabilidad, Universidad Nacional de Colombia, Bogotá, 2010, p4.

[6] BONNET JEREZ, José Luis. Lecciones de estadística: estadística descriptiva y probabilidad, Editorial club universitario, Alicante, 2003, p3.

[7] BURRILL, Gail, Data Driven Mathematics: A Curriculum Strand for High School Mathematics, Mathematics Teacher, tomo 89, número 6, Septiembre, 1996, p460-65, 540.

[8] COBO M., Belén, Significado de las medidas de posición central para los estudiantes de Secundaria, Tesis Doctoral, Universidad de Granada, Granada, 2003.

[9] CONTRERAS G., José Miguel, Recursos en Internet para la enseñanza de la probabilidad condicionada, Universidad de Granada, Granada 2009.

[10] CUEVAS, Jesús, IBÁÑEZ, Carlos: Estándares en educación estadística: Necesidad de conocer la base teórica y empírica que lo sustentan, Unión, número15, México, 2008, p33 - 45.

[11] DAUME, Peggy, Aktien und Optionen: Zur Integration von Inhalten der stochastischen Finanzmathematik in einem allgemeinbildenden und anwendungsorientierten Stochastikunterricht, Tesis Doctoral, Universidad de Humboldt de Berlin, 2008

[12] DE HALICARNASO, Heródoto. Los nueve libros de la historia, Edaf, Madrid, 2007.

[13] FERNÁNDEZ F., Santiago, CORDERO S., José María, CÓRDOBA L., Alejandro, Estadística descriptiva, ESIC, Madrid, 2002.

[14] GOROSTIZA, Luis G., La probabilidad en el siglo XX, Miscelánea Matemática, Sociedad Matemática Mexicana, número 33, México, 2001 p69–92.

[15] International Statistical Literacy Project, 2012
http://www.stat.auckland.ac.nz/~iase/islp/home

[16] LOYN, Henry R., Diccionario Alkal de Historia Medieval. Ed Akal, Madrid 1998, p143.

[17] MATEOS, Gregoria, MORALES, Aparicio. Historia de la probabilidad (desde sus orígenes a Laplace) y su relación con la historia de la teoría de la decisión, Historia de la Probabilidad y de la Estadística I, Ed. AC, Madrid, 2002, p14.

[18] MAYÉN, Silvia, Comprensión de las medidas de posición central en estudiantes mexicanos de bachillerato, Tesis doctoral, Universidad de Granada, Granada, 2009.
[19] MEN, documento N° 3, Estándares básicos de competencias en Lenguaje, Matemáticas, Ciencias y Ciudadanas, Colombia, 2006.
[20] MENDENHALL, William, BEAVER, Robert J, BEAVER, Barbara M. Introducción a la probabilidad y estadística, Internacional Thomson Editores, México, 2002.
[21] PÉREZ ABREU, Víctor M. Una Invitación a la Probabilidad, Departamento de Probabilidad y Estadística, Centro de Investigación en Matemáticas A. C., Guanajuato, 2010, p5.
[22] RADICATI, Carlos. Introducción al estudio de los Quipus, fondo editorial universidad Nacional de San Marcos, Lima, 2006, p100.
[23] RAMESH, Kapadia, Statistical Education 11 to 16 — The Schools Council Project, Teaching statistics, tomo 1, número 1, enero 1979, p11 – 14.
[24] RODRÍGUEZ, Norma, MONTAÑEZ, Graciela y ROJAS, Ilda: Dificultades en contenidos de estadística inferencial en alumnos universitarios. Revista Electrónica Iberoamericana de Educación en Ciencias y Tecnología, Volumen 2, Número 1, 2010, p57-73
[25] ROMERO, Juan de Jesús, Estadística y probabilidad II, Editorial Santillana, 2008.
[26] ROSS, Sheldon M., Introducción a la estadística, Reverté, Barcelona, 2007, p425.
[27] SPIEGEL, Murray R., SCHILLER, John, SRINIVASAN, R. Alu. Probabilidad y estadística, Mc Graw Hill, México, 2003.
[28] WILHELMI, Miguel R., Combinatoria y probabilidad. Departamento de didáctica de la Matemática, Universidad de Granada, Granada, 2004.