



Leñitas Geométricas*

para el Fogón Matemático de los Festivales

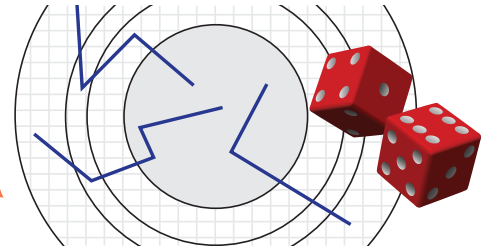
De OMA para Profesores y Maestros en actividad

5ª época ✂ N° 1
9 de marzo de 2023



"[...] más valioso que acumular información es interiorizarse de métodos para saber qué propósitos tienen, es decir, de dónde parten y adónde llevan. Los métodos tienen una orientación y una dinámica de las cuales los resultados y teoremas aislados carecen". *Dr. Alberto Calderón*

El Método Montecarlo



El casino de Montecarlo y las radiaciones en la física

Debe reconocerse que Stanislaw Ulam y John von Neumann tuvieron su momento de inspiración al bautizar este método como *Monte Carlo*. Y no era otra cosa que un método de resolución aproximada del transporte de partículas mediante el uso de números aleatorios. En realidad, el nombre fue sugerencia de su colega Nicholas Metropolis, quien relacionó el aspecto aleatorio del método con el hecho de que Ulam tuviera un tío que solía pedir dinero a la familia porque "tenía que ir al casino de Montecarlo", en el Principado de Mónaco. El nombre tuvo tanta aceptación que se generalizó a todo método de aproximación basado en el uso de números aleatorios y no solo a su aplicación en el transporte de partículas. Pero muchos pensarán "¿y quiénes son esos personajes?". Eran físicos estadounidenses de origen europeo que trabajaron conjuntamente en el famoso Proyecto Manhattan durante la Segunda Guerra Mundial y en los años posteriores.

Pero estrictamente hablando ellos no fueron los inventores del método. Quizás la primera documentación sobre el uso de un muestreo aleatorio para encontrar una solución a un problema fue la del naturalista francés Georges-Louis Leclerc, conde de Buffon.

Entre los científicos en general, a Buffon se lo conoce como un iconoclasta que, entre otras cosas, proponía como edad estimada de la Tierra unos 75.000 años, en lugar de la cifra generalmente aceptada de 6.000 años aproximadamente. Entre los matemáticos, Buffon es reconocido por dos contribuciones de distinto

**Los números complejos
en la geometría del plano.
Teorema de Ptolomeo.
Potencia.**



¡Hacé tu pedido!

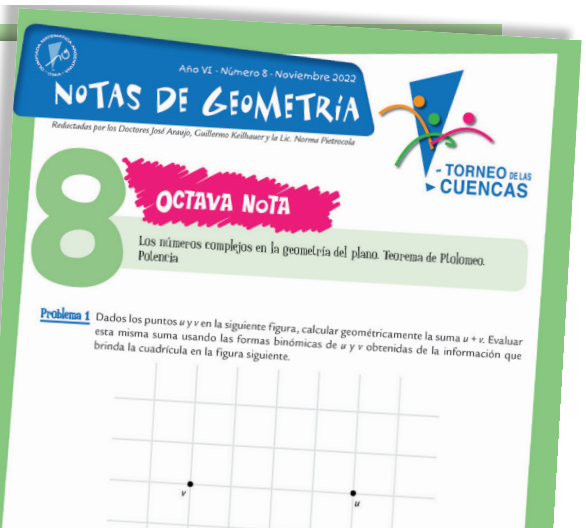
En la Red Olímpica realizamos envíos de todos nuestros títulos a todo el país bajo el sistema contra reembolso o delivery.

Publicación reciente

fenchu@oma.org.ar

☎ **11 4826 8976**

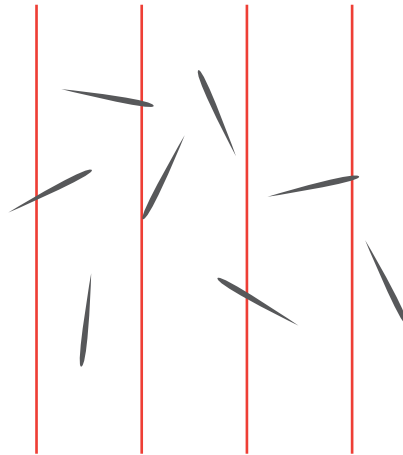
📞 **+54 9 11 5035 7537**



* Recordamos a los lectores que los temas editados en *Leñitas Geométricas* son material preparado y en gran parte desarrollado y sugerido por el doctor Miguel de Guzmán.

tipo: una traducción al francés del *Method of Fluxions*, de Isaac Newton, y por el denominado *problema de la aguja de Buffon* en teoría de probabilidades. Este matemático también se había visto impresionado por la paradoja de San Petersburgo, y en su *Essai d'arithmétique morale* –publicado en 1777 en el volumen cuarto de un suplemento a la *Histoire naturelle*–, daba varias razones para considerar el juego en cuestión como intrínsecamente imposible.

Buffon, en el mismo *Essai*, introducía también "lo que constituía esencialmente una nueva rama de la teoría de probabilidades: la que estudia los problemas probabilísticos basados en consideraciones geométricas". Allí proponía el problema siguiente: "Considérese un plano horizontal dividido en regiones por un haz de rectas paralelas equidistantes, sobre el que se lanza al azar una aguja de grosor despreciable. La probabilidad de que la aguja corte a una de las rectas paralelas aparece calculada correctamente por Buffon como $\frac{2l}{\pi d}$, donde d es la distancia común entre las paralelas y l la longitud de la aguja, con $l < d$.



El *Essai* también contiene una interesante colección de tablas de nacimientos, matrimonios y muertes correspondientes a París, que cubren los años 1709 a 1766, así como resultados obtenidos a partir de ellas relativos a esperanza de vida, que fueron criticados duramente por Jean D'Alembert.

Durante el siglo XVIII fue cuando se introdujo en Europa, procedente del este, la práctica de la vacunación contra la viruela, es decir, de la inoculación con una forma debilitada de viruela, con el objeto de desarrollar inmunidad contra dicha enfermedad. Esta costumbre provocó una controversia entre los matemáticos que trataban de aplicar la teoría de probabilidades a los asuntos de la vida social. En 1760 Daniel Bernoulli leyó ante la Académie des sciences de París un trabajo relativo a las ventajas de la inoculación, pero antes de que dicho trabajo fuese publicado en las *Mémoires de l'Académie D'Alembert* ya le había puesto objeciones: si bien no negaba las ventajas de la vacuna, sí sostenía que Bernoulli había exagerado tales ventajas.

Parte de la controversia se centró en la distinción que, según insistía D'Alembert, había que hacer entre la "vida media" y la "vida probable" de un individuo. La "vida probable" de un niño venía a ser entonces de unos 8 años (es decir, la mitad de los niños morían en esa época antes de cumplir los 8 años), mientras que

"Estas páginas servirán para animar a matemáticos y no matemáticos a meditar profundamente sobre el sentido mismo del quehacer matemático". *Miguel de Guzmán*

¿Ya lo tenés?

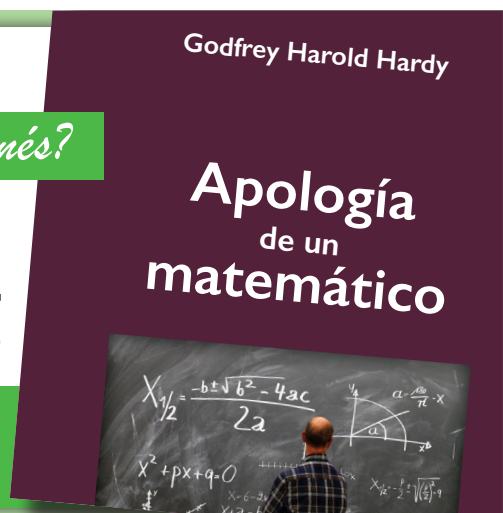


fenchu@oma.org.ar

☎ 11 4826 8976 📞 +54 9 11 5035 7537

¡Hacé tu pedido!

En la Red Olímpica realizamos envíos de todos nuestros títulos a todo el país bajo el sistema contra reembolso o delivery.



su “vida media” o duración de la vida humana en promedio era de unos 26 años. (Si comparamos estas cifras con las que corresponden a la época actual, se pone de manifiesto el horriblemente bajo nivel de los cuidados médicos de siglos pasados).

Las controversias en torno a la probabilidad de que la inoculación fuese ventajosa terminaron definitivamente cuando fue descubierta al fin la verdadera vacuna científica contra la viruela por el doctor Edward Jenner.

William Thomson, lord Kelvin (Belfast, 26 de junio de 1824-Largs, 17 de diciembre de 1907; Reino Unido) fue un físico y matemático británico que también usó el muestreo aleatorio para evaluar unas integrales de la teoría cinética de los gases. En una publicación de 1901, describía cómo había numerado trozos de papel que posteriormente extraía de un cuenco. Pero él mismo ponía de manifiesto lo poco que le convencía el método ante la dificultad de mezclar adecuadamente los trozos o, incluso, porque la electricidad estática provocaba que en ocasiones sacara varios trozos de una vez.

Existen otros ejemplos documentados sobre el uso del muestreo aleatorio antes de la década de 1940, pero podemos considerar a Ulam y von Neumann los creadores del método Montecarlo como lo conocemos hoy día. Ellos realizaron una descripción completa del método y le incluyeron un ingrediente que le confirió un potencial excepcional: la utilización de computadoras electrónicas.

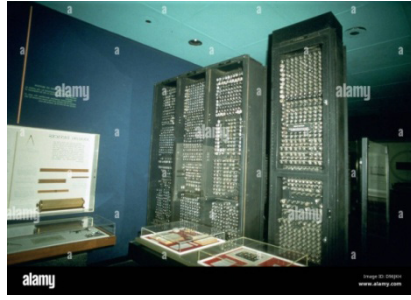
Al terminar la Segunda Guerra Mundial, en la Universidad de Pensilvania se construyó la primera computadora electrónica, ENIAC, con más de 17.000 válvulas de vacío y otras tantas resistencias, diodos y condensadores. Su objetivo inicial era realizar cálculos balísticos para el ejército estadounidense, que hasta ese momento eran efectuados por un ingente grupo de personas con calculadoras convencionales de la época. John von Neumann, que ejercía de consultor tanto para el ejército como para el Laboratorio Nacional de Los Álamos, implicado en el Proyecto Manhattan, convenció al ejército de ceder el poder de cálculo de ENIAC para comprobar modelos de reacciones nucleares.

¿En qué consistían esos cálculos?

Un ejemplo es la difusión de neutrones en un material fisionable. Los modelos iniciales consideraban una esfera de material fisionable rodeada de un material de alta densidad. Se asumía una distribución inicial de neutrones con diferentes velocidades. La idea era seguir el movimiento de un gran número de neutrones individuales que pudieran sufrir dispersiones, absorciones, fisiónar núcleos y escapar de la esfera.

En primer lugar, se toma un neutrón en concreto con una posición y velocidad dadas. Entonces hay que decidir en qué punto va a sufrir una colisión y de qué naturaleza será. Si se ha determinado que sea una fisión, por ejemplo, debe decidirse el número de neutrones que emergerán, los cuales deben ser seguidos como el que provocó la fisión. En cambio, si la colisión es una dispersión, debe determinarse su nueva velocidad. Una vez seguido el primer neutrón y todos los que produjera por fisiones hasta que sean absorbidos o abandonen la geometría, habrá concluido la “historia” de ese neutrón. Como resultado, el medio ha sufrido una cierta perturbación que habremos calculado. Esa perturbación puede ser, por ejemplo, la energía cedida al medio por los neutrones. Este mismo proceso se repite tantas veces como sea necesario hasta apreciar que el promedio de esa perturbación tiende a un valor estable.

Pero, ¿cómo se decide qué evento sufre cada neutrón y cuál será la longitud recorrida entre eventos consecutivos? Aquí es donde entran en juego los números aleatorios. Dado que conocemos las funciones de probabilidad de ocurrencia de los diferentes eventos posibles, tan solo se trata de muestrear aleatoriamente dichas funciones cada vez que se requiera determinar la longitud de una trayectoria, qué evento va a ocurrir y qué características van a tener los productos de dicho evento, si los hay. De la simulación de numerosas historias obtenemos un resultado promedio. Pero dicho resultado va cambiando conforme aumenta el número de historias. ¿Cómo podemos saber cuán cerca nos hallamos del valor que estamos buscando? Para ello nos ayudamos de la varianza. Cuanto menor sea esa varianza, más cerca del valor buscado nos encontraremos. Como regla general podemos decir que la varianza es inversamente proporcional al número de historias simuladas. Si queremos tener un resultado muy preciso, es necesario simular un alto número de historias.



Computadora ENIAC

La sigla ENIAC se obtiene de *Electronic Numerical Integrator and Computer* (Computador e Integrador Numérico Electrónico), y fue una de las primeras computadoras de propósito general. Era Turing-completa, digital y susceptible de ser reprogramada para resolver “una extensa clase de problemas numéricos”. Como ya se dijo, inicialmente fue diseñada para calcular tablas de tiro de artillería destinadas al Laboratorio de Investigación Balística del Ejército de los Estados Unidos.

Antes de disponer de ENIAC las simulaciones se efectuaban con calculadoras mecánicas. Es fácil intuir que se requería tanto tiempo para alcanzar resultados fiables que su abordaje era inviable. Enrico Fermi, que ya había hecho incursiones en este mundo de la estimación estadística en la década de 1930 –sin publicar sus métodos–, diseñó un artilugio que llamó FERMIAC. Consistía en una serie de ruedas que se ajustaban según el material en el que se encontrara el neutrón y permitía dibujar trayectorias sobre un gráfico bidimensional. Este dispositivo se utilizó durante un periodo en el que ENIAC no estuvo operativo debido a un traslado de centro de operaciones.

En los primeros años posteriores a la difusión del método Montecarlo asistido por computadoras, este tuvo una aplicación limitada. Eran muy pocos los científicos que disponían de dicho tipo de computadoras y su programación era muy compleja. En 1954, Evans Hayward y John H. Hubbell simularon la trayectoria de 67 fotones utilizando calculadoras mecánicas y una lista de números aleatorios. En la década de 1950 había una cierta insistencia en aplicar el método en la resolución de cualquier tipo de problema, a pesar de que en muchos casos resultaba menos eficiente que otras herramientas de análisis numérico, lo que contribuía a su desacreditación ante la comunidad científica.

Fue a partir de la década de 1960 cuando empezó a gozar de mayor consideración por varias razones. Se consiguió más reconocimiento en aquellos problemas en los que era la mejor –y en algunos casos la única– técnica disponible. También hubo un continuo progreso en la construcción de las computadoras electrónicas, lo que las hacía cada vez más asequibles y rápidas en sus cálculos.

La programación de las simulaciones también se facilitó considerablemente con el desarrollo de códigos Montecarlo que permiten la simulación del transporte de radiación con la mera introducción de unos parámetros que definen el problema en cuestión. Los usuarios de estos códigos no tienen que preocuparse de aspectos como la generación de números aleatorios ni de la definición de funciones de distribución de probabilidad de eventos. Entre esos códigos podemos encontrar el MCNP (desarrollado por el propio laboratorio de Los Álamos), EGS4, Geant4 y Fluka. ¡E incluso tenemos uno español: PENELOPE!, desarrollado en la Universidad de Barcelona, el cual goza de un gran prestigio sobre todo en el ámbito de la física médica. Aunque inicialmente había un gran interés por el transporte de neutrones, podemos encontrar que esos códigos pueden emplearse para el transporte de otros tipos de partículas como fotones, electrones, etcétera.

En la actualidad, los métodos Montecarlo son ampliamente utilizados en el estudio de problemas tan diversos como evaluación de integrales en matemáticas, crecimiento de bosques y estudios de contaminación en ingeniería ambiental, análisis de mercado en economía y un largo etcétera.

Por supuesto, podemos imaginar el impacto que ha tenido el desarrollo de códigos de simulación del transporte de radiación en el uso médico de las radiaciones ionizantes. Pero su aplicación en la medicina no solo se centra en el estudio de la interacción de la radiación con materia. En procesos en los que podemos encontrar una alta variabilidad, como es el caso de la respuesta de los tejidos y tumores en los tratamientos radioterápicos, obtener resultados a partir de los ensayos con pacientes requiere en muchos casos de un esfuerzo desmesurado. Es aquí donde la simulación Montecarlo aporta información muy valiosa que puede servir de apoyo en el avance científico.

Como curiosidad, en 1971 Luke Rhinehart publicó la novela *El hombre de los dados*. El método Montecarlo es llevado a su extremo por un psiquiatra que decide volverse un “hombre aleatorio” tomando todas sus decisiones a partir de los resultados del lanzamiento de un par de dados.

Caminos a la realidad: el método Montecarlo

Ilya Meyerovich Sobol (Lituania, 15 de agosto de 1926 [96 años]) es un matemático ruso, conocido por su trabajo sobre el método Montecarlo. Su investigación abarca varias aplicaciones, desde estudios nucleares hasta astrofísica, y ha contribuido significativamente al campo de la educación; es bien conocido su Seminario Elemental sobre este método.

Seguiremos aquí el contenido de su seminario. En su última propuesta se expresó sobre “probabilidad” y “variable aleatoria”; si bien cualquier persona pudo haberlas empleado alguna vez, la idea intuitiva de probabilidad (en tanto que frecuencia) corresponde más o menos al contenido del Seminario. En cambio, la idea intuitiva de la variable aleatoria la suponemos muy nueva y, como regla general, lejos de su noción matemática. Por eso, damos por conocido el concepto de probabilidad y explicaremos el concepto más complejo de variable aleatoria.



No se pretende aquí reemplazar un curso sobre teoría de las probabilidades, porque esta exposición es simplificada y no tiene demostraciones. Sin embargo, el desafío es lograr una idea clara sobre las variables aleatorias, o por lo menos suficiente para comprender los aspectos elementales del método Montecarlo.

El objetivo principal es sugerir su utilidad en las más diversas ramas del quehacer matemático en la vida cotidiana, es decir, en qué incide en nuestra cultura y, además, en qué resultó útil para atacar un problema. Los problemas que se analizarán son variados y bastante simples. Sin embargo, no pueden abarcar, claro está, todas las esferas de su aplicación.

Veamos un ejemplo cuyo desarrollo no trata para nada sobre su utilidad en la medicina; no obstante, estos métodos permiten evaluar las dosis en la radioterapia. Teniendo un programa de cálculo de los rayos absorbidos por distintos tejidos del cuerpo, se puede dosificar y orientar la radiación del modo más eficiente, cuidando de no dañar los tejidos sanos.

1. Generalidades sobre el método

El método Montecarlo es un método numérico que permite resolver problemas matemáticos mediante la simulación de variables aleatorias.

→ **1. Orígenes del método de Montecarlo.** Como ya lo señalamos, se considera como fecha de nacimiento del método de Montecarlo el año 1949, en el que apareció un artículo titulado “The Monte Carlo Method”, la redacción de este artículo sobre el método suele atribuirse a los matemáticos norteamericanos John von Neumann y Stanislaw Ulam.



John von Neumann



Stanislaw Ulam

Stanislaw Marcin Ulam (13 de abril de 1909-13 de mayo de 1984) fue un científico polaco-estadounidense en los campos de la matemática y la física nuclear. Participó en el Proyecto Manhattan, originó el proyecto Teller-Ulam para armas termonucleares, descubrió el concepto del autómatas celular, inventó el método de cálculo de Montecarlo y sugirió la propulsión de pulso nuclear. En matemática pura y aplicada, demostró algunos teoremas y propuso varias conjeturas.

John von Neumann (28 de diciembre de 1903-8 de febrero de 1957) fue un matemático, físico e informático húngaro-estadounidense, ingeniero y erudito. De él se dijo que había sido “el último representante de los grandes matemáticos que se sintieron igualmente cómodos tanto en matemáticas puras como aplicadas”. Integró las ciencias puras y aplicadas.

Von Neumann hizo contribuciones importantes en muchos campos, en la matemática (fundamentos de la matemática, teoría de la medida, análisis funcional, teoría ergódica, teoría de grupos, teoría de redes, teoría de la representación, álgebras de operadores, teoría de matrices, geometría y análisis numérico), en física (mecánica cuántica, hidrodinámica, balística, física nuclear y mecánica estadística cuántica), economía (teoría de juegos y teoría del equilibrio general), computación (arquitectura de Von Neumann, programación lineal, meteorología numérica, computación científica, máquinas autorreplicantes, computación estocástica) y estadística. Fue pionero en la aplicación de la teoría de operadores a la mecánica cuántica en el desarrollo del análisis funcional, y una figura clave en el desarrollo de la teoría de juegos y los conceptos de autómatas celular, el constructor universal y las computadoras digitales.

Von Neumann publicó más de 150 artículos en su vida: alrededor de 60 en matemáticas puras, 60 en matemáticas aplicadas, 20 en física y el resto sobre temas matemáticos especiales o no matemáticos. Su último trabajo, un manuscrito inacabado escrito mientras agonizaba en el hospital, se publicó más tarde en forma de libro con el título *La computadora y el cerebro*.

Su análisis de la estructura de la autorreplicación precedió al descubrimiento de la estructura del ADN. En una breve lista de hechos sobre su vida que envió a la Academia Nacional de Ciencias, escribió: “La parte de mi trabajo que considero más esencial es la de la mecánica cuántica, que se desarrolló en Göttingen en 1926 y, posteriormente, en Berlín en 1927-1929. Además, mi trabajo sobre varias formas de teoría de operadores, Berlín 1930 y Princeton 1935-1939; y sobre el teorema ergódico, Princeton, 1931-1932”.

Durante la Segunda Guerra Mundial, von Neumann trabajó en el Proyecto Manhattan con el físico-teórico Edward Teller, el matemático Stanislaw Ulam y otros, resolviendo problemas de pasos clave en la física nuclear involucrada en las reacciones termonucleares y la bomba de hidrógeno. Desarrolló los modelos matemáticos detrás de las lentes explosivas usadas en armas nucleares de implosión y acuñó el término “kilotón” (de TNT) como una medida de la fuerza explosiva. Durante ese tiempo y después de la guerra, asesoró a un gran número de organizaciones de Estados Unidos, incluida la Oficina de Investigación y Desarrollo Científico, el Laboratorio de Investigación Balística del Ejército, el Proyecto de Armas Especiales de las Fuerzas Armadas y el Laboratorio Nacional Oak Ridge. En el apogeo de su influencia en la década de 1950, fue presidente de varios comités críticos del Departamento de Defensa, incluido el Panel de Armas Nucleares de la Junta Asesora Científica de la Fuerza Aérea y el Comité Asesor Científico ICBM,

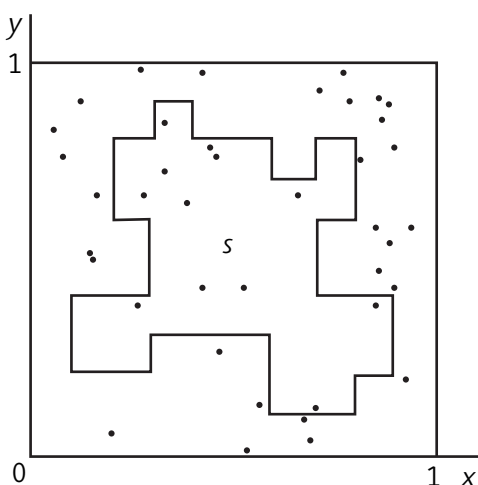
así como miembro destacado de la Comisión de Energía Atómica. Jugó un papel clave, junto a Bernard Schriever y Trevor Gardner, en el diseño y desarrollo de los primeros programas ICBM de Estados Unidos.

En honor a sus logros y contribuciones al mundo moderno, fue nombrado en 1999 la Persona del Siglo por el Financial Times, como representante del ideal característico del siglo en que el poder de la mente podía dar forma al mundo físico y del “poder intelectual”.

En la Unión Soviética los primeros artículos dedicados al método de Montecarlo aparecieron en 1955 y 1956. Las bases teóricas del método eran bien conocidas desde hacía mucho tiempo; es más, algunos problemas de la estadística se resolvían, a veces, empleando las muestras aleatorias, o sea, aplicando de hecho el método de Montecarlo. Sin embargo, hasta la aparición de las máquinas calculadoras electrónicas (MCE) este método no había encontrado aplicaciones suficientemente amplias, ya que la simulación a mano de variables aleatorias era un proceso muy laborioso. Es decir, la aparición del método de Montecarlo, en tanto que método numérico de gran universalidad, solo fue posible gracias a la creación de las MCE.

El nombre de Montecarlo se debe al de una población del Principado de Mónaco, célebre por su casa de juego. Resulta que uno de los aparatos mecánicos más sencillos que permite obtener variables aleatorias es la ruleta. Acerca de esto hablaremos más adelante. En cambio, conviene responder aquí a una pregunta frecuente: ¿ayuda o no el método de Montecarlo a ganar en el juego de la ruleta? No, no ayuda. E incluso no tiene nada que ver con este juego.

- **2. Veamos un ejemplo.** Para que nos hagamos una idea más clara de qué cuestiones vamos a tratar consideremos un ejemplo muy sencillo. Supongamos que queremos calcular el área de una figura plana S . Esta puede ser una figura arbitraria con frontera curvilínea, definida gráfica o analíticamente y compuesta de uno o varios pedazos. Supongamos que se trata del contorno representada en la figura siguiente y supongamos que toda la silueta está comprendida dentro de un cuadrado de dimensión 1.



Tomemos en el cuadrado un número N de puntos aleatorios. Sea N' el número de puntos que aparecen dentro de S . Es obvio, por razones geométrica, que el área de S es aproximadamente igual a la razón $\frac{N'}{N}$. Cuanto mayor sea N tanto mayor será la exactitud de esta estimación.

En el ejemplo representado en el contorno se ha escogido un total de $N = 40$ puntos, de estos, $N' = 12$ puntos aparecen dentro de S . Tenemos $\frac{N'}{N} = \frac{12}{40} = 0,30$, mientras que el valor exacto del área de S es 0,35.

En la práctica, el método Montecarlo no se aplica al cálculo de área de figuras planas ya que existen para ello otros métodos que, a pesar de ser más complejos, garantizan una exactitud mucho mayor. Sin embargo, el método Montecarlo expuesto en nuestro ejemplo permite calcular, con la misma facilidad, “volúmenes multidimensionales” de un cuerpo en un espacio multidimensional. Es más, sucede a menudo que el método Montecarlo es el único método numérico que permite resolver el problema.

- **3. Dos atributos del método Montecarlo.** El primero consiste en que su algoritmo tiene una estructura muy sencilla. Como regla, se elabora primero un programa para la realización de una prueba aleatoria (en

el ejemplo anterior, hay que escoger un punto aleatorio en el cuadrado dado y comprobar si pertenece o no a S); después, esta prueba se repite N veces, de modo que cada experimento sea independiente de los restantes y se tome la media de los resultados de todos los experimentos. Por esto el método Montecarlo se denomina a veces *método de pruebas estadísticas*.

La segunda peculiaridad consiste en que el error es, como regla general, proporcional a la magnitud $\sqrt{\frac{D}{N}}$, donde D es una constante y N es el número de pruebas. Esta fórmula permite ver que para disminuir el error en 10 veces (en otras palabras, para obtener en el resultado otra cifra decimal exacta) es preciso aumentar N (o sea, el volumen del trabajo) en 100 veces.

Se pone de manifiesto la imposibilidad de alcanzar por este camino una elevada exactitud. Por eso suele decirse que el método Montecarlo resulta especialmente eficaz en la solución de problemas en los cuales se necesita conocer el resultado con poca exactitud (del 5 al 10%).

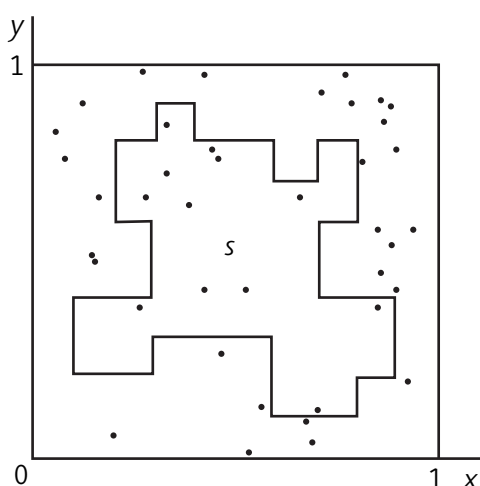
Sin embargo, un mismo problema puede ser resuelto aplicando distintas variantes del método Montecarlo a las que corresponden diferentes valores de D . En la literatura especial con frecuencia se suele hablar ahora de los métodos Montecarlo (en plural) para referirse al hecho de que un mismo problema se puede resolver mediante la simulación de distintas variables aleatorias. En numerosos problemas se logra elevar considerablemente la exactitud escogiendo un procedimiento de cálculo al que le corresponde un valor mucho menor de D .

- **4. Problemas que permite resolver el método Montecarlo.** En primer lugar, el método Montecarlo permite simular cualquier proceso cuya marcha dependa de factores aleatorios. En segundo lugar, en muchos problemas matemáticos, que no tienen la menor relación con cuestiones aleatorias, se puede inventar un modelo probabilístico artificial (e incluso más de un modelo) que permite resolver estos problemas. En realidad, esto es lo que hemos hecho en el ejemplo del punto 2.

Por consiguiente se puede hablar del método Montecarlo como de un método universal en la solución de problemas matemáticos. Es de especial interés el hecho de que en algunos casos conviene abandonar la simulación del proceso aleatorio real y concentrarse en el análisis del modelo artificial. Más adelante consideraremos una situación de este tipo.

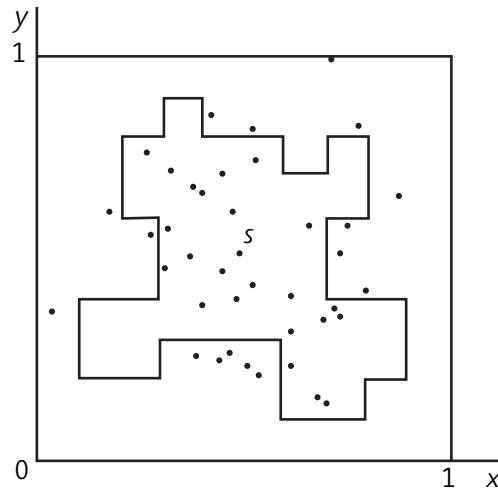
- **5. Algo más sobre el ejemplo.** Volvamos al ejemplo del punto 2. Para realizar el cálculo, debimos escoger en el cuadrado puntos aleatorios. ¿Cómo realizar esto?

Imaginemos el siguiente experimento. La figura anterior (aumentada proporcionalmente), con el contorno S y el cuadrado, se coloca sobre un muro que sirve de blanco para un tirador que, situado a cierta distancia, dispara N veces apuntando al centro del cuadrado.



Claro está que no todas las balas darán exactamente en el centro: perforarán el blanco en N puntos aleatorios. Aceptamos que el tirador está lejos de ser el campeón del mundo y se encuentra a distancia suficientemente grande del blanco. ¿Es posible estimar el área de S a partir de estos puntos?

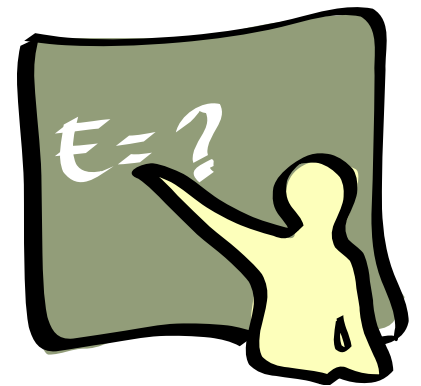
En la figura de abajo se ha representado el resultado de un experimento semejante.



En este experimento se tiene $N = 40$, $N' = 24$ y $\frac{N'}{N} = 0,60$ lo que representa casi el doble del valor exacto (0,35) del área. Por supuesto, tratándose de un tirador de gran puntería los resultados del experimento serán muy malos, ya que casi todas las balas pegarán cerca del centro y perforarán S. Más adelante explicaremos cómo han sido escogidos los puntos aleatorios en las figuras de arriba.

Es fácil persuadirse de que el método que hemos aplicado para calcular el área tendrá “validez” solo si los puntos aleatorios, además de ser aleatorios, están “distribuidos uniformemente” sobre el cuadrado. Para dar un sentido exacto a estas palabras es preciso estudiar la definición de variable aleatoria y algunas de sus propiedades. Estos conocimientos se exponen, sin demostración, en el próximo apartado 2, del método Montecarlo.

Un diálogo con los maestros



MATERIALES PARA PRESENTAR EN EL AULA

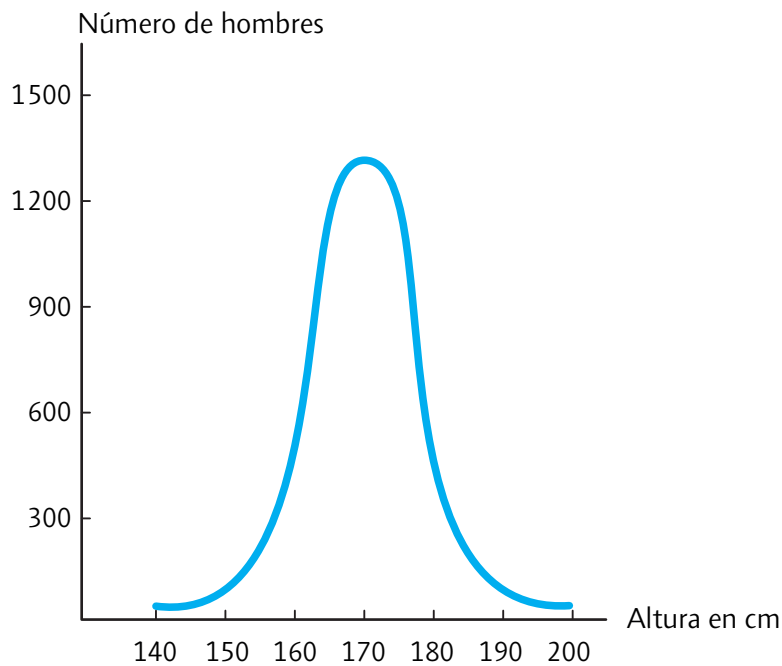
Los promedios en estadística

La estadística es, en esencia, inducción. La ciencia, en general, avanza por dos caminos fundamentales: deducción e inducción. Con la deducción, a partir de ciertos principios básicos y mediante razonamientos lógicamente correctos, se va tratando de obtener consecuencias y proposiciones (inferir) que constituyen la teoría. La inducción científica (que no hay que confundir con la inducción matemática o inducción completa) procede por otro camino bien diferente: partiendo de hechos y observaciones experimentales, trata de elevarse a conclusiones generales sobre el objeto que estudia.

Así obtenemos la distribución de alturas de 8555 hombres.



Curva típica de distribución normal



Thomas Bayes, en 1763, fue el primero en introducir elementos matemáticos en este proceso inductivo, dando así los pasos iniciales en lo que ha llegado a ser la estadística actual.

La estadística es el estudio de los mejores modos de acumular y analizar datos y de establecer conclusiones acerca del colectivo del que se han recogido tales datos.

Todos los datos, ¿para qué? Si, por ejemplo, se quiere estudiar las características fisiológicas (estatura, peso etc.) de la población de España que tiene entre 14 y 17 años de edad, está claro que no sería posible, ni ayudaría mucho, recolectar los datos de todos y cada uno de los individuos en esa edad. Es necesario restringirse a escoger una muestra.

A medida que avancemos veremos la importancia de ciertos valores: la *media* y la *desviación típica* son los parámetros que manejaremos principalmente. Calcularemos, entre otras, las de las estaturas de los jugadores de varios equipos deportivos como el básquetbol.



Algunos campos de la estadística. La estadística considera las siguientes cuestiones fundamentales:

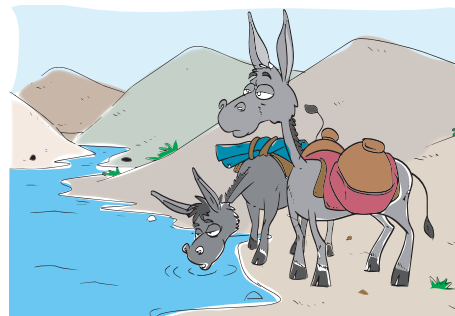
1. **¿Cómo describir mejor la muestra?** Es claro que incluso una muestra de 1 000 individuos proporciona un montón inmanejable de datos. Hay que ordenarlos, agruparlos, determinar medias y otros números que describan la muestra de modo resumido y transparente (cantidades a las que llamaremos parámetros). Esta es la finalidad de la estadística descriptiva o elemental, de la que, en estos próximos temas, veremos algo interesante.
2. **¿Qué conclusiones se pueden inferir acerca de la población total? ¿Hasta qué punto son fiables estas conclusiones?** Esta es la parte que llamaremos *inferencia estadística* y que actualmente tiene un enorme aparato matemático basado, fundamentalmente, en la teoría de la probabilidad.

3. **¿Cómo tomar las muestras para que nos proporcionen una información más fiable?** A esta pregunta responde la parte de la estadística que se denomina *diseño de experimentos*.

Una aventura al siglo XVII

¿Cómo y cuándo empezó todo esto?

Desde la antigüedad, reyes y emperadores se preocuparon de conseguir datos abundantes sobre sus posesiones. Ya el emperador Augusto (coetáneo de Cristo) mandó realizar una gran encuesta sobre las riquezas del Imperio Romano: soldados, navíos, recursos, rentas, etc. Pero hasta comienzos del siglo XVII la estadística era puramente descriptiva, es decir, una enumeración sistemática y ordenada de datos.



John Graunt

John Graunt, un tendero londinense –persona ingeniosa y estudiosa, que se levantaba pronto por la mañana para sus estudios, antes de la apertura de la tienda–, inspirado por unas Tablas de Mortalidad que semanalmente se editaban en su parroquia, publicó un librito que puede considerarse el nacimiento de la Estadística como ciencia. Esto sucedía a mediados del siglo XVII.

El libro, además de valerle a su autor el ser nombrado por el Rey miembro de una importante sociedad científica, fomentó el estudio de las estadísticas de vida, tanto en su país como en otros países europeos.

Sin embargo, la palabra “estadística” para designar la obtención, el estudio y la interpretación de grandes masas de datos parece que fue utilizada por primera vez un siglo más tarde (a mediados del XVIII) en Alemania. Su nombre proviene del interés que tenía para los asuntos de Estado.

Bayes en la teoría de probabilidades de Laplace

El primero en utilizar elementos matemáticos de la teoría de la probabilidad, a fin de hacer predicciones basadas en experimentos previos, fue Thomas Bayes (1763), que estableció un teorema importante en el desarrollo de la teoría. Fue un avanzado en su tiempo; quizá por eso, olvidado hasta que lo rescató Pierre-Simon de Laplace.



Hasta el momento poco hemos hablado de Laplace, a quien sus contemporáneos tenían, como matemático, en tan alta consideración, como Joseph-Louis Lagrange. Hay dos razones para este aparente olvido. En primer lugar, Laplace no tomó parte prácticamente en actividades revolucionarias; parece haber tenido un alto sentido de la honradez intelectual en la ciencia, pero en política carecía de convicciones concretas. Esto no quiere decir que fuera timorato, puesto que se habría relacionado libremente con aquellos de sus colegas científicos que fueron sospechosos durante el período de crisis.

Se ha dicho que también él habría estado en peligro de ser enviado a la guillotina, de no ser por sus contribuciones a la ciencia, pero esta opinión parece bastante dudosa, ya que Laplace apareció a menudo como un descarado oportunista. Jugó un papel, aunque no de gran importancia, en el Comité de Pesos y Medidas, y fue también, naturalmente, profesor de la École Normale y de la École Polytechnique, pero no publicó sus lecciones, al contrario que Gaspard Monge y que Lagrange.

Sus publicaciones estuvieron dedicadas principalmente a la mecánica celeste, campo en el que destacó sin rivales de su talla desde la época de Isaac Newton. Laplace tuvo también su relación con la administración política unos años más tarde, cuando Napoleón, que sentía una gran admiración por los hombres de ciencia, lo nombró ministro del Interior, puesto que había ocupado también Nicolas Carnot por algún tiempo con Napoleón. Pero es bien sabido que Laplace, a diferencia de Carnot, no mostró tener aptitudes para el cargo, y Napoleón se burlaba de él diciendo que “aplicaba el espíritu de lo infinitamente pequeño a la dirección de los asuntos de Estado”.

Una segunda razón del escaso énfasis que hemos puesto en la obra de Laplace es la de que dicha obra no tuvo la influencia inmediata y duradera que se observa en la de otros miembros del de la Revolución Francesa. Sus obras principales representan, en cierto sentido, el final de una época más bien que el comienzo de un periodo nuevo, aunque tenemos que hacer una clara excepción con su obra sobre la teoría de probabilidades y sobre teoría del potencial. La teoría de probabilidades debe más a Laplace que a ningún otro matemático.

Desde 1774 escribió muchos artículos sobre el tema, y los resultados obtenidos los incorporó y organizó en su libro clásico *Théorie analytique des probabilités*, de 1812. Laplace consideró la teoría de probabilidades desde todos los puntos de vista y en todos los niveles y su *Essai philosophique des probabilités* de 1814 es en realidad una exposición introductoria para lectores no especializados. Escribía en ella que “en el fondo, la teoría de probabilidades es solo sentido común expresado con números”; pero su *Teoría analítica* muestra la mano magistral del analista consumado que domina el cálculo superior. El libro está lleno de integrales que incluyen las funciones beta y gamma.

Laplace fue uno de los primeros que demostró que la integral

$$\int_{-\infty}^{+\infty} e^{-x^2} dx$$

es decir, el área bajo la curva de las probabilidades, es igual a $\sqrt{\pi}$. El método que utilizó para obtener este resultado era un tanto artificioso, pero tampoco difiere demasiado del artificio moderno de transformar la integral

$$\int_0^{\infty} e^{-x^2} dx \cdot \int_0^{\infty} e^{-y^2} dy = \iint_0^{\infty} e^{-(x^2+y^2)} dx dy$$

a coordenadas polares como

$$\iint_0^{\frac{\pi}{2}} r e^{-r^2} dr d\theta$$

la cual se puede calcular ya fácilmente y da como resultado

$$\int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$$

Entre las muchas cosas a las que dedica su atención Laplace en su *Teoría analítica* está el cálculo de π por medio del problema de la aguja de Buffon, que había permanecido casi olvidado durante 35 años.

A este problema se le conoce a veces como el *problema de la aguja de Buffon-Laplace*, dado que Laplace extendió el problema original a una cuadrícula formada por dos haces de rectas paralelas equidistantes y perpendiculares el uno al otro. Si las distancias entre las rectas de cada uno de los haces son a y b , entonces la probabilidad de que una aguja de longitud l (menor que a y que b) lanzada al azar corte a una de estas rectas es

$$p \frac{2l(a+b) - l^2}{\pi ab}$$

Laplace rescató también del olvido la obra Bayes (1761) sobre probabilidades inversas. En este libro de Laplace aparece además la teoría de los mínimos cuadrados, inventada por Adrien-Marie Legendre, junto con la demostración formal que no había dado Legendre. La *Teoría analítica* contiene también la transformada de Laplace, que tan útil resultó para la teoría de ecuaciones diferenciales:

Si se tiene que

$$f(x) = \int_0^{\infty} e^{-xt} g(t) dt$$

entonces se dice que la función $f(x)$ es la transformada de Laplace de la función $g(x)$.

En el siglo XIX Carl Friedrich Gauss se enfrentó con el siguiente problema al efectuar una misma observación astronómica o geodésica repetidas veces; obtenía valores diferentes debido a errores humanos e instrumentales en cada observación. Si se obtienen, por ejemplo, 20 valores distintos, ¿qué número tomar como más probable? Gauss trabajó el método de los mínimos cuadrados, de gran importancia teórica y práctica.

Trabajando sobre una propuesta de Michael Joseph Moroney

Michael Joseph Moroney (Leicester, 1918-1990). Después de obtener en Londres un título externo en matemática y física en 1940, fue destinado como científico al servicio industrial en el laboratorio de Leicester de la Standard Telephones and Cables Ltd. En 1945 se incorporó al Departamento de Matemáticas de la Facultad de Tecnología y Comercio de Leicester, donde se convirtió en profesor titular de estadística industrial y dirigió un servicio de consultoría para firmas locales. En 1950 hizo una breve incursión en la política, aunque sin éxito, como candidato liberal de Leicester North-east. En el mismo año se sumó como uno de los primeros miembros de la Asociación de Estadísticos Incorporados (como se llamaba entonces), convirtiéndose más tarde en miembro del consejo, donde actuó como examinador en estadísticas industriales y jugó un papel muy activo en los primeros años de la Asociación.



En 1951 se publicó *Facts from Figures*, que se convirtió en un éxito de ventas inmediato. Más ediciones, con revisiones menores, siguieron en 1953, 1956 y 1965, y aparecieron muchas reimpresiones a lo largo de las décadas de 1950 y 1960. Ha sido traducido a nueve idiomas y continúa publicándose hasta el día de hoy.

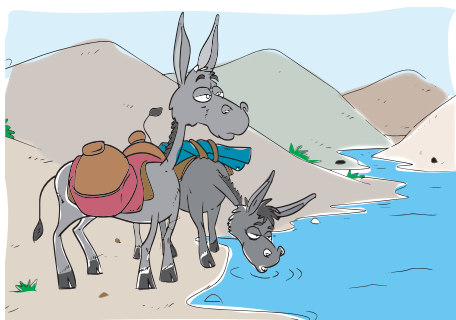


Michael Joseph Moroney

Moroney nunca se consideró a sí mismo como un académico. Por lo tanto, cuando a partir de la publicación de *Facts from Figures* se convirtió muy rápidamente en el más conocido estadístico de su época, no fue de extrañar que se atrajera la etiqueta de divulgador. No era un título que Moroney despreciara; estaba feliz de hacer por las estadísticas lo que W. W. Sawyer (y L. Hogben antes que él) habían hecho en matemáticas (Sawyer fue su jefe de Departamento en Leicester).

En 1954, Moroney dejó la vida académica para unirse a la oficina central de Unilever Ltd. Aparte de su desempeño como consultor privado durante tres años, permaneció en Unilever (Ltd. o N.V.) hasta su jubilación en 1981.

Importancia de los promedios



En otro tiempo, cuando los azares de los viajes por mar eran mucho más serios de lo que son hoy, cuando en los barcos azotados por las tormentas se echaba por la borda parte de su carga, se reconocía que aquellos cuyas mercancías fueran sacrificadas tenían derecho de reclamar indemnización a expensas de los otros, que las tenían a salvo. El valor de las mercancías perdidas era pagado por acuerdo entre todos aquellos cuyas cargas habían estado en el mismo barco.

Este daño marítimo a la carga en tránsito era conocido como *havaría*, y la palabra se aplicó luego naturalmente a la compensación monetaria con la que debía contribuir cada individuo. De esta palabra latina deriva la inglesa moderna *average* (promedio). Así, la idea de un promedio tiene sus raíces en el seguro primitivo. Naturalmente, con el desarrollo de la navegación el seguro se perfeccionó, de modo que el riesgo se repartiera no solo entre los poseedores de los bienes arriesgados en un viaje, sino entre grupos más grandes de negociantes. Con el tiempo, el cargar con tales riesgos se transformó en una profesión lucrativa y especializada. Esto implicaba el pago al asegurador de una suma de dinero que guardaba una relación evidente con el riesgo corrido.

La idea de un promedio es un bien común. A pesar de nuestros escasos conocimientos de aritmética, estamos muy acostumbrados a la idea de promedio de goles, de puntos, etc. Nos damos cuenta de que el objetivo de un promedio es representar un grupo de valores individuales de una manera simple y concisa, para que la mente pueda lograr una rápida comprensión del tamaño de los individuos en el grupo, libre de la impresión de variaciones fortuitas e irregulares.

Es de suma importancia comprender el hecho de que el promedio debe actuar como representante. En consecuencia, es el colmo del desatino recorrer toda la aritmética para calcular el promedio de un conjunto de cifras que no constituyen, en un sentido real, una familia singular. Supongamos que un próspero médico que



CASIO

fx-82LA PLUS
NATURAL-V.P.A.M. 2nd edition

$$\frac{2\sqrt{3}}{3} + \frac{6}{3\sqrt{3}} = \frac{4\sqrt{3}}{3}$$

fx-82LA PLUS
NATURAL-V.P.A.M. 2nd edition

fx-82LA PLUS
NATURAL-V.P.A.M. 2nd edition

Descubrí toda la línea CASIO en
www.calculadoras.ar

gana \$ 3 000 por año tiene mujer y dos niños, ninguno de los cuales tiene empleo rentado alguno, que este médico tiene a su servicio a una muchacha a quien paga \$ 150 por año, y que hay además un jardinero que recibe \$ 40 por año.

Podemos recorrer los procesos del cálculo de la renta promedio para este pequeño grupo. Seis personas ganan, entre todas, \$ 3 190 en el año. Si se divide el total de rentas por el número de personas, podría determinarse que la renta promedio del grupo es de \$ 531,66. Pero esta cifra no es más que un impostor con traje de promedio, que no representa a persona alguna del grupo.

Da una información sin ningún sentido, porque no se puede hacer una sola deducción legítima de ella. Este es un ejemplo extremo, pero promedios parecidos se calculan con gran despreocupación. Pocas personas se preguntan: "¿Qué conclusiones se sacarán de este promedio que voy a calcular? ¿Crearé una falsa impresión?".

La idea del promedio es tan accesible que no debe sorprender que se hayan inventado varias clases, de modo que se cubriera un campo tan vasto como fuera posible con un mínimo de falseamiento. Disponemos de una colección de promedios, y elegimos el que sea apropiado a nuestra información y a nuestro propósito. No debemos caer en el error de que, dada la facilidad de comprensión del concepto de promedio, no haya que decir nada más del asunto, pues los promedios pueden ser muy engañosos.

El promedio más sencillo es aquel que todos conocemos bien. Este promedio ordinario también es llamado *media*, palabra que indica "centro". (Todos los promedios son conocidos por los estadísticos como *medidas de tendencia central*, porque nos indican el punto alrededor del cual se agrupan los diferentes valores). La media aritmética o promedio de un conjunto de número se calcula tomando el total de todas las cantidades y dividiendo este total por el número de ellas. No es preciso decir nada más sobre este punto, salvo que las cantidades que se promedian deben ser de la misma clase. No podemos, por ejemplo, promediar el sueldo de un polígamo con el número de sus esposas.

Una segunda clase de promedio es la *media armónica*, que es la recíproca de la media aritmética de las recíprocas de los valores que se desea promediar. La recíproca de un número se encuentra dividiendo la unidad por ese número, por ejemplo: la recíproca de 4 = $\frac{1}{4} = 0,25$.

La media armónica es el promedio apropiado para usar cuando se trabaja con tasas y precios. Considérese el ejemplo clásico bien conocido del aeroplano que vuela alrededor de un cuadrado que tiene 100 millas de lado, recorriendo el primero de estos a 100 m.p.h., el segundo a 200 m.p.h., el tercero a 300 m.p.h. y el cuarto lado a 400 m.p.h. ¿Cuál es la velocidad media del aeroplano en su vuelo alrededor del cuadrado? Si promediamos las velocidades haciendo uso de la media aritmética, como de ordinario, obtenemos:

$$\text{Velocidad media} = \frac{100 + 200 + 300 + 400}{4} = 250 \text{ m.p.h.}$$

Pero este no es el resultado correcto, como fácilmente puede verse a continuación:

$$\begin{aligned} \text{Tiempo de vuelo sobre el primer lado} &= 1 \text{ hora} \\ \text{Tiempo de vuelo sobre el segundo lado} &= 30 \text{ minutos} \\ \text{Tiempo de vuelo sobre el tercer lado} &= 20 \text{ minutos} \\ \text{Tiempo de vuelo sobre el cuarto lado} &= 15 \text{ minutos} \\ \text{Tiempo total de vuelo sobre 400 millas} &= 2 \text{ hs } 5 \text{ minutos} \\ &= \frac{25}{12} \text{ horas} \end{aligned}$$

$$\text{De esto se deduce que la velocidad media es } \frac{400}{\frac{25}{12}} = 192 \text{ m.p.h.}$$

La media aritmética de las velocidades nos da, entonces, un resultado erróneo. Un indicio para descubrir la razón de ello es el hecho de que las diferentes velocidades no se mantienen durante el mismo tiempo sino solo a lo largo de la misma distancia. El promedio correcto que debe emplearse en tal caso es la media armónica.

Para dar la fórmula de esto introduciremos aquí un poco más de notación matemática, que nos será de gran utilidad más adelante. Cuando calculamos promedios debemos hacer la suma de una sucesión de cantidades que componen el conjunto del cual se desea el promedio. El matemático hace uso de un signo taquigráfico para indicarnos cuándo debemos sumar.

Se le llama *signo de suma*, y se usa la letra S del alfabeto griego, que se escribe Σ , y lleva el nombre de "sigma", para indicar cuándo debe sumarse una sucesión de términos. Se trata, efectivamente, de la sigma mayúscula. (Más adelante nos ocuparemos bastante de la sigma minúscula, que se escribe σ .) Cada uno de los números que han de tenerse en cuenta en nuestros cálculos se representa con la letra x . Si deseamos diferenciar las diversas cantidades, podemos entonces numerarlas: x_1, x_2, x_3, x_4 , etc., escribiendo los números distintivos con subíndices, de modo que no se confundan con las cifras que intervienen en el cálculo. (Esto parecerá tan confuso al aprendiz como aburrido al aficionado).

Tomemos como ejemplo el cálculo de la media aritmética de los cinco números 5, 6, 8, 7, 6. Si hubiera razón para ello, podríamos distinguirlos como sigue: $x_1 = 5$; $x_2 = 6$; $x_3 = 8$; $x_4 = 7$; $x_5 = 6$.

Ahora bien, la ventaja de usar notación algebraica (es decir, letras que pueden sustituirse por números cualesquiera de acuerdo con la naturaleza del problema) es que podemos escribir en forma muy abreviada las reglas necesarias para efectuar el cálculo que nos dé la respuesta correcta al tipo de problema que nos ocupa. De hecho, una fórmula no es sino la solución de todo problema de un tipo determinado. Resolvemos el problema de una vez por todas cuando encontramos la fórmula. La fórmula es la solución. Todo lo que hay que hacer es remplazar las letras por las cantidades concretas que aparecen en el problema.

Supongamos, ahora, que se indica el número de cantidades que se promedian en nuestro problema con la letra n (en el ejemplo que dimos, $n = 5$). Para calcular la media aritmética tenemos que sumar las cinco cantidades: $3 + 6 + 8 + 7 + 6 = 32$.

Esta parte del cálculo aparecerá en la forma algebraica como $x_1 + x_2 + x_3 + x_4 + x_5$. El próximo paso sería dividir el total por el número de cantidades que se promedian, es decir 5, dando como promedio 6,4. Con notación algebraica esto aparecería así:

$$\text{Promedio} = \frac{x_1 + x_2 + x_3 + x_4 + x_5}{n}$$

Esta forma de escribir la fórmula sería muy inconveniente si hubiera que promediar muchas cantidades; más aún, no hay necesidad de mantener las cantidades señaladas, ya que en el promedio se deja precisamente de lado las características individuales. Por eso se introduce el signo de suma, con lo cual la fórmula toma la forma muy abreviada:

$$\text{Promedio} = \frac{\sum x}{n}$$

La fórmula nos dice que, para obtener el promedio, debemos sumar todos los valores de x y dividir el total por el número de valores n .

Del mismo modo ahora la media armónica, que es, como dijimos, el promedio usado para calcular velocidades medias, etc., y que se define como la recíproca (la recíproca de un número x es igual a $\frac{1}{x}$) de la media aritmética de las recíprocas de los valores x que se quiera promediar, tiene por fórmula:

$$\text{Media armónica} = \frac{n}{\sum \left(\frac{1}{x}\right)}$$

Para ilustrar el uso de esta fórmula, apliquémosla a nuestro problema del aeroplano. Las cuatro velocidades, mantenidas cada una de ellas sobre la misma distancia, eran 100, 200, 300 y 400 m.p.h. Estos son nuestros valores de x . Como existen cuatro de ellos, el valor de n en nuestra fórmula es 4, y obtendremos:

$$\text{Media armónica} = \frac{n}{\sum \left(\frac{1}{x}\right)} = \frac{4}{\left(\frac{1}{100} + \frac{1}{200} + \frac{1}{300} + \frac{1}{400}\right)} = \frac{4}{\left(\frac{25}{1200}\right)} = \frac{4 \times 1200}{25} = 192 \text{ m.p.h.}$$

que es, como sabemos, la solución correcta.

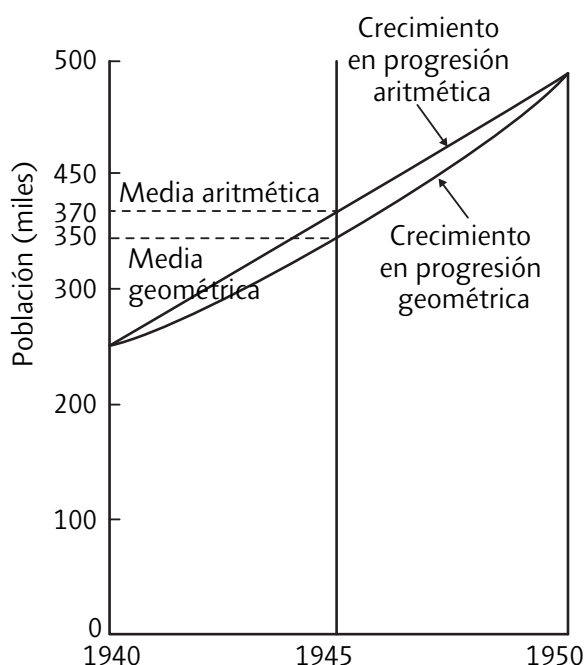
Debemos observar con cuidado que la media armónica es apropiada aquí porque los tiempos son variables, con distancias constantes. Si hubieran sido los tiempos constantes y las distancias variables, hubiera sido correcto usar la media aritmética ordinaria. El tipo de promedio que debe ser usado depende siempre de los términos del problema planteado. Las fórmulas no deben emplearse sin discriminación.

Un tercer tipo de promedio es la *media geométrica*. Este es el promedio que debe usarse cuando se desea promediar cantidades deducidas de una situación en la cual siguen lo que W. W. Sawyer, llama la *bárbara ley del crecimiento*, es decir, una progresión geométrica o la ley exponencial.

Muchas cantidades siguen esta forma de crecimiento. Por ejemplo, la población de una ciudad con tasas de natalidad y mortalidad estables, y sin migración, crecerá en proporción con el número de personas de la ciudad. Supóngase que cierta ciudad tuviese en el año 1940 una población de 250 000 personas y que esta fuera de 490 000 en 1950. Si quisiéramos estimar la población en el año 1945 (la estimación de poblaciones en varios momentos entre los sucesivos censos es materia importante en estadísticas de salud pública), entonces, como aproximación grosera deberíamos tomar el promedio de las poblaciones conocidas en las dos fechas dadas, es decir:

$$\text{Población en 1945} = \frac{250000 + 490000}{2} = 370.000$$

Este sería un método razonable solo en el caso de que pudiésemos suponer que la población aumentara la misma cantidad cada año. Esto no sería probable, sin embargo, ya que, a medida que la ciudad aumenta su tamaño, es lícito suponer que el aumento de habitantes se verifica con velocidad siempre creciente (ver figura siguiente).



Comparación de interpolaciones por medio de la media aritmética y de la media geométrica. La población de una ciudad crece a menudo de acuerdo con la ley exponencial. Esto sería cierto, en efecto, con tasas de natalidad y mortalidad estables y ausencia de migración. En estas condiciones, la media geométrica sería más apropiada que la aritmética para interpolar la población estimada en un momento dado entre dos fechas con datos conocidos.

Es probable obtener una estimación mejor, en circunstancias normales, calculando la media geométrica de las poblaciones en las dos fechas dadas.

Para calcular la media geométrica se efectúa el producto de todas las cantidades que se desee promediar. Luego, si hay n de tales cantidades, se hallará la raíz n -ésima del producto indicando nuestras n cantidades con: $x_1, x_2, x_3, \dots, x_n$, podríamos escribir la fórmula de la media geométrica como sigue:

$$\text{Media geométrica} = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$$

Si aplicamos esto al problema dado más arriba, donde deseabamos estimar la población de una ciudad en 1945, supuesto que en 1940 la población fuese de 250 000 y en 1950 de 490 000, tendríamos $n = 2$, cantidades para promediar, y resulta:

$$\text{Media geométrica} = \sqrt{250000 \cdot 490000} = 350000$$

como estimación nuestra de la población de 1945. Puede observarse que este resultado es sensiblemente inferior al que obtuvimos haciendo uso de la media aritmética (370 000) y si observamos la figura, notaremos que es la estimación más aceptable.

Si resumimos aquí nuestros tres diferentes promedios, tendremos:

Media aritmética (indicada a menudo con: \bar{x})

$$\bar{x} = \frac{\sum x}{n}$$

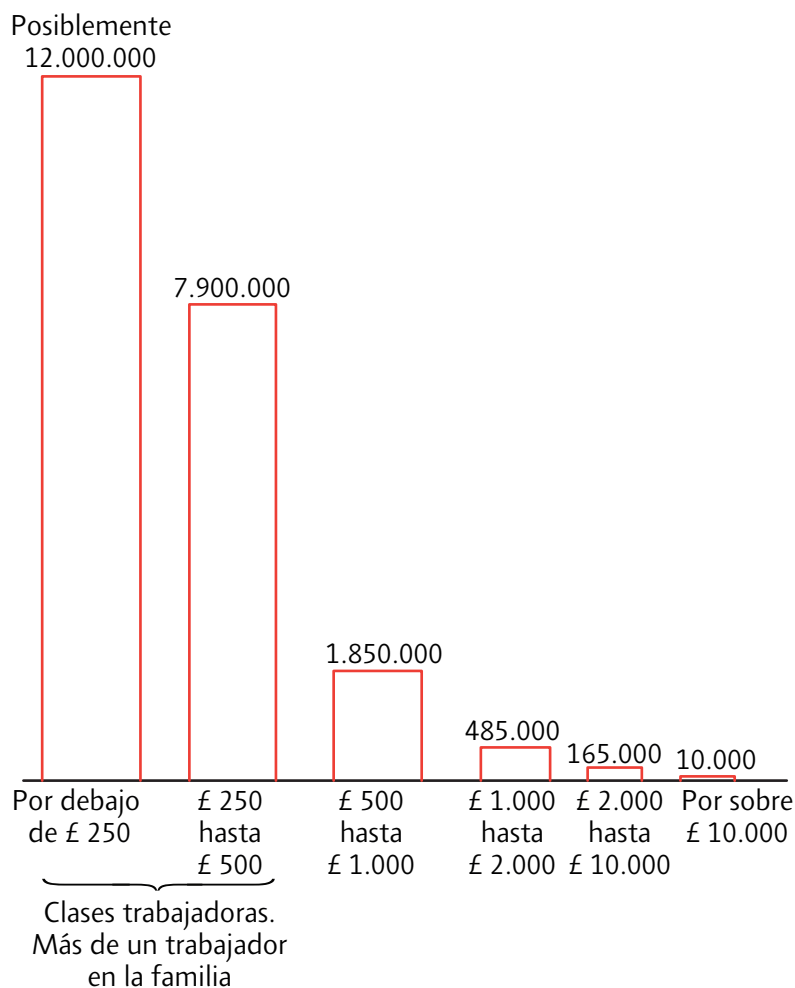
Media armónica (indicada de ordinario con H)

$$H = \frac{n}{\sum \left(\frac{1}{x}\right)}$$

$$\text{Media geométrica } G = \sqrt[n]{x_1 \cdot x_2 \cdot x_3 \cdot \dots \cdot x_n}$$

Cada una de estas medidas de tendencia central tiene sus aplicaciones especiales propias. Todas ellas se obtienen por procesos aritméticos sencillos que toman en cuenta la magnitud de cada dato individual.

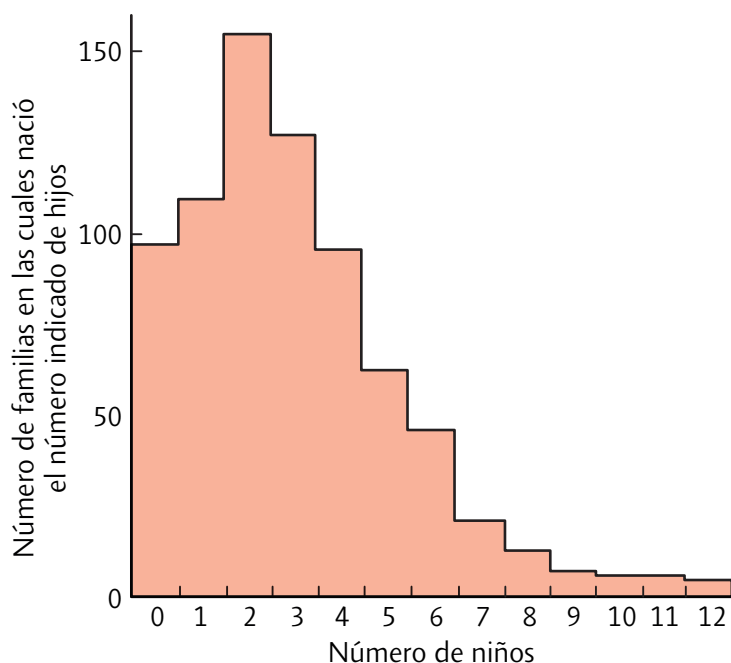
Recalamos el importante concepto de que cualquier promedio o medida de tendencia central representa a un grupo homogéneo cuyos miembros se reconocen como similares. Muchas distribuciones, aunque son homogéneas en el sentido de que existe cierta continuidad entre los diversos miembros del grupo, son, sin embargo, tales que existen grandes diferencias entre los miembros más grandes y los más pequeños y muestran, además, una notable falta de simetría, con tendencia a agruparse más bien en un extremo que en el otro.



Número de personas con diferentes niveles de rédito formando una distribución positivamente asimétrica. Compárese con otras formas figuradas de representación respecto de la forma de la distribución, observando

a) la alta renta combinada de un número cada vez menor en los niveles superiores y b) el efecto del impuesto.

La figura de arriba es un ejemplo típico. Muestra la forma en que se distribuye el rédito anual. Es cierto que hay continuidad, pero las pequeñas rentas son lo ordinario. Apreciamos que el cálculo de promedios para distribuciones de este tipo usando la media aritmética sería muy engañoso. La relativamente pequeña cantidad de gente con rentas extremadamente altas elevaría el promedio en forma apreciable, de modo que no representaría realmente a la población en general.



La figura de arriba, que nos muestra la frecuencia relativa de los diferentes tamaños de familia, presenta la misma dificultad. Algunas familias están muy lejos del problema de los niños, y el cálculo de una media aritmética sería irrazonable, particularmente si nuestro propósito fuera puramente descriptivo.

Es evidente que lo que necesitamos en tales casos es una medida de tendencia central que no se vea afectada por los relativamente pocos valores extremos de la “cola” de la distribución. Dos ideas se manifiestan. La primera es que, si tomásemos todas nuestras familias y las ordenáramos en una larga columna comenzando con la más pequeña y llegando hasta la mayor, podríamos entonces usar la medida de la que se encuentra a mitad de camino entre ambos extremos como nuestra medida de tendencia central. Esta medida se llama *mediana* (significa “cantidad del medio”). La mitad de todas las familias tendrían un tamaño no menor que el de la mediana y la otra mitad de ellas serían de tamaño no mayor.

Obsérvese que de este modo no tomamos en cuenta para nada el número real de niños, excepto para fines de ordenamiento. Es evidente que el número de niños en la familia más numerosa podría aumentarse hasta 50 000 sin perturbar para nada nuestra medida de tendencia central, que sería siempre la cantidad del medio.

Un segundo método para obtener una medida de tendencia central, en la cual no influyen los valores extremos de la distribución, es tomar el valor que se presenta más frecuentemente. Este es el que está de moda. Se le llama *moda* o *valor modal*. Por ejemplo, en la última figura puede verse que el valor modal para tamaños de familia es dos niños. Este es realmente un valor típico, y así nos parece comparado con la media aritmética que, para el caso, da 2,96. Es difícil imaginar 2,96 niños.

Obsérvese que la media aritmética es influida notablemente por las relativamente pocas familias numerosas. ¿Cuál es el correcto? Ninguno y ambos. Los dos promedios sirven a un propósito. La moda sería una base muy pobre para cualquier operación aritmética posterior, ya que en ella se excluyó a propósito la precisión matemática por el interés en la presentación de un resultado típico. La media aritmética, por otra parte, excelente como es para propósitos numéricos, sacrificó sus aspiraciones de ser típica en favor de la exactitud numérica. Por eso es a menudo deseable incluir ambas medidas de tendencia central. Mejor aún es ir más allá, presentando un histograma de la distribución, como en la figura comentada.

Un problema frecuente se presenta cuando queremos hacer una estimación del valor de la mediana de una distribución y no disponemos de los valores de las cantidades individuales, aunque sí del número de ellas en

determinados niveles. Nos ocuparemos de este asunto más adelante, cuando consideramos la cuestión de los límites de clase y lindes.

Diremos ahora algunas palabras acerca de las distribuciones de frecuencia. Si tenemos un grupo grande de elementos, cada uno de ellos en correspondencia con un valor que indica su magnitud variable de un elemento a otro (como, por ejemplo, cuando consideramos alturas de hombres o el monto del impuesto a la renta pagado por cada uno de ellos), y si hacemos una tabla o gráfico que muestre la frecuencia relativa con la cual los miembros del grupo muestran los varios valores posibles de la cantidad variable (por ejemplo, proporción de hombres con cada una de las diferentes alturas, o proporciones de la población contenidas en los diversos grupos del impuesto sobre la renta), entonces tendremos lo que se conoce como una distribución de frecuencia para la cantidad variable en cuestión.

A menudo se le llama sencillamente *distribución*. Así, tenemos distribuciones para alturas, pesos, capacidad torácica, rentas, cantidad de habitaciones por persona, etc. Del mismo modo tenemos distribuciones para el número de muertes que producen las diferentes enfermedades según la edad, número de áreas de gobierno local con determinadas tasas de natalidad y mortalidad, etc. La cantidad que varía (altura, tasa de natalidad, renta, etc.) se llama *variable*.

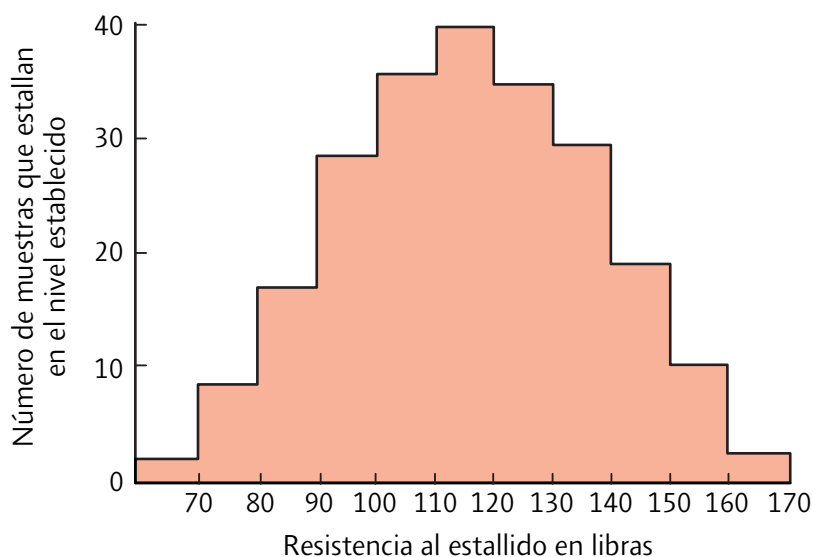
Algunas variables son continuas, es decir, pueden tomar absolutamente cualquier valor dentro de un cierto intervalo. Renta, altura, tasa de natalidad y variables similares, son continuas. Otras variables se conocen como discontinuas, porque pueden tomar solo valores aislados. Por ejemplo, el número de niños de una familia puede ser solo entero, ya que las fracciones son imposibles. Las familias crecen por saltos enteros. Un aumento en la familia es un acontecimiento. Los goles registrados en partidos de fútbol, los objetos perdidos en ómnibus, el número de pétalos de una flor; todas estas cantidades variables son discontinuas.

Cuando se recogen datos con propósitos de análisis estadístico, es raro que se disponga de información acerca de todos los individuos del grupo. Los datos del censo son, quizás, lo más cercano a la perfección en este sentido; pero, aun en este caso, la información pierde actualidad a medida que se registra. Si bien decimos que el censo tomado en cierto día, en un país determinado, da la cifra 43 574 205, no podremos seguir empleando sin reparos, por ejemplo, el último cinco durante los diez años siguientes.

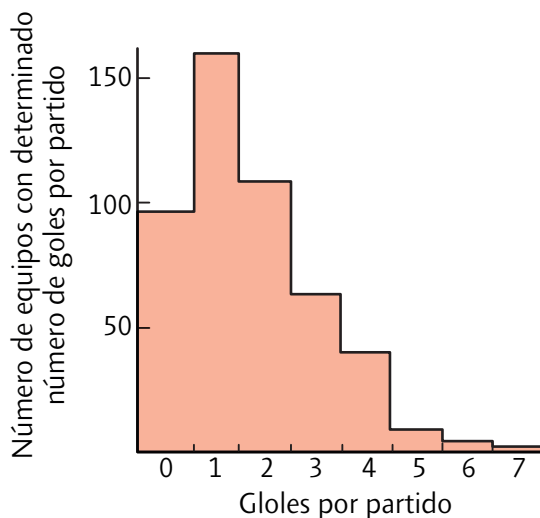
En realidad, ni siquiera durante los siguientes diez minutos. Tal exactitud sería espuria. No es posible, en general, la investigación de la población total y debemos conformarnos con una muestra. Se toma una muestra con la idea de lograr una deducción acerca de la población de la cual fue tomada, suponiendo, por ejemplo, que el promedio de una buena muestra se relaciona estrechamente con el promedio de la población total. Diremos más sobre muestras más adelante.

La palabra "población" se usa en estadística no solo para referirse a grupos de personas, sino también, por extensión natural, a grupos de medidas asociadas con conjuntos de objetos inanimados. Si sacamos una muestra con un número suficientemente grande de medidas, podremos llegar a una distribución de frecuencia para cualquier población. Las figuras siguientes ofrecen ejemplos de varios tipos de distribuciones.

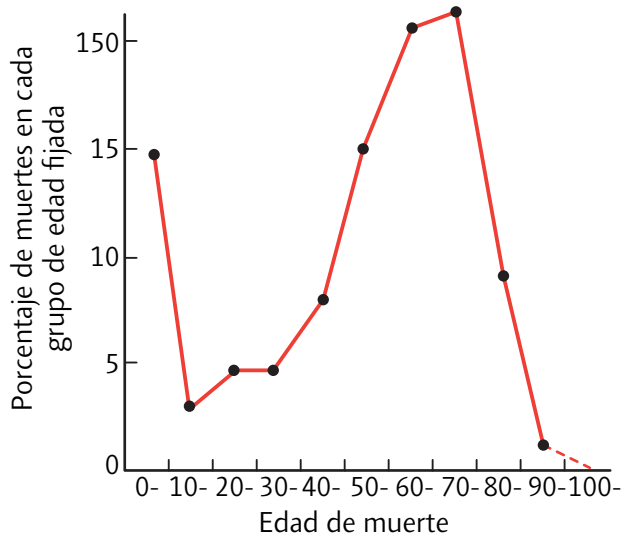
Un ejemplo de simetría. Distribución de la resistencia al estallido en muestra de nailon de cubiertas vinílicas.



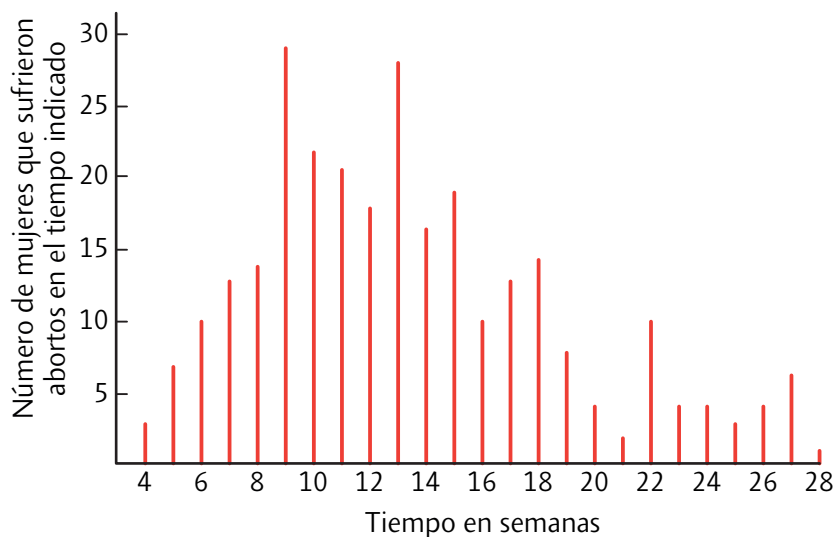
Distribución positivamente asimétrica de una variable discontinua. Número de goles logrados por equipo y por partido.



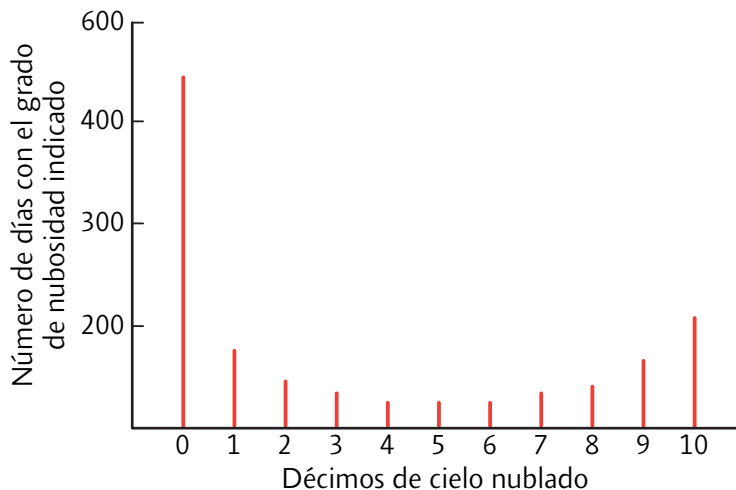
Ejemplo de distribución bimodal (de dos picos). El pico de los primeros años revela la seria pérdida de vida potencial debida a la mortalidad infantil. (Del *Registrar General's years 1930-32*, citado por Kendall en *Advanced Statistics*).



Aborto en mujeres. Datos de T. V. Pearce (1930) citado por M. G. Kendall, *Advanced Statistics*. Podríamos querer especular acerca de una posible periodicidad en los datos. ¿Sugieren razonablemente un ciclo aproximadamente mensual? ¿Qué otra conclusión puede deducirse?



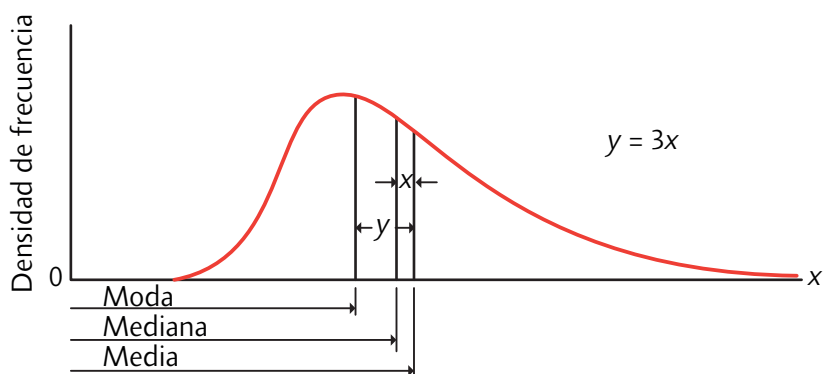
Distribución peculiar de la nubosidad en Greenwich. Basada en datos de Gertrude Pearse (1928) para el mes de julio 1890-1904 (excluido 1901) y citado por M. G. Kendall, *Advanced Statistics*, vol. I. Nótese la tendencia de cielo muy despejado o muy nublado.



Algunas distribuciones, como se verá en los diagramas, son *simétricas* respecto de su valor central. Otras tienen una asimetría notable y se conocen como *asimétricas*.

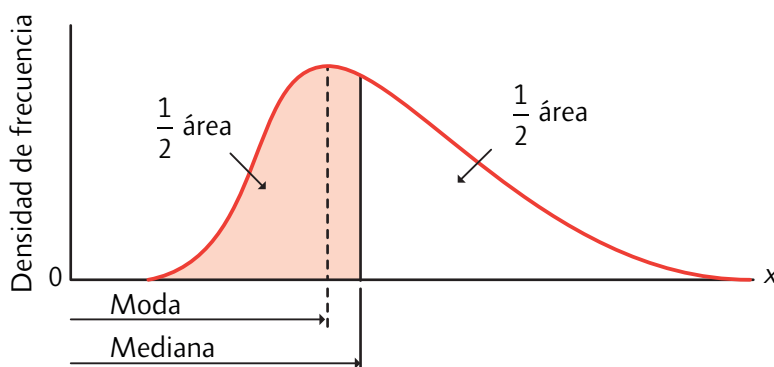
Estas se dividen en dos tipos. Si la “cola” de la distribución se extiende hacia los valores mayores de la variable, se trata de una distribución con asimetría positiva; si se dirige hacia los valores menores, la distribución tendrá asimetría negativa.

A continuación nos dedicaremos a la cuestión de la concentración de los elementos de una distribución alrededor de su valor central, ya que es un asunto de gran importancia poder medir el grado en que difieren entre sí los diversos elementos de la población.

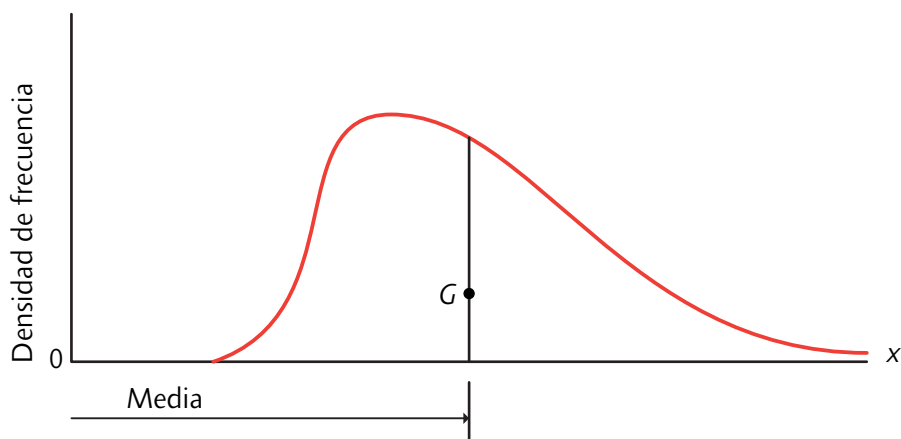


La figura ilustra una interesante relación que mantienen entre sí en forma aproximada la mediana, la moda y la media, de distribuciones moderadamente asimétricas. Las dos figuras que siguen ilustran interpretaciones geométricas de las tres medidas de tendencia central.

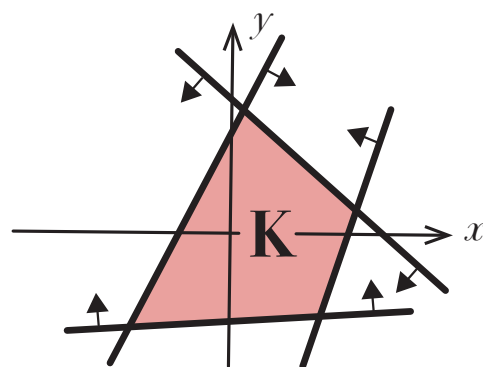
Interpretación geométrica de la moda y de la mediana. La línea vertical que pasa por el valor de la mediana divide el área bajo la curva de frecuencias en mitades (el área es proporcional a la frecuencia). La vertical que pasa por la moda lo hace por el pico de la curva, es decir, por el valor en el cual la densidad de la frecuencia es máximo.



Interpretación geométrica de la media. La vertical que pasa por la media, lo hará por el centro de gravedad de una hoja de grosor y densidad uniformes, recortada según la forma de la distribución. La media es la abscisa del centro de gravedad G .



Cerraremos sin entrar en consideraciones sobre números índices, que no son, en realidad, más que una clase especial de promedio.



LOS SISTEMAS DE DESIGUALDADES LINEALES

Nuestro objetivo. El material que presentamos trata sobre la relación entre sistemas de desigualdades lineales y poliedros convexos: contiene la descripción de los conjuntos de todas las soluciones para desigualdades lineales; estudia las cuestiones de compatibilidad e incompatibilidad; se da una introducción elemental a la programación lineal que, de hecho, es uno de los capítulos de la teoría de desigualdades lineales. Al final expondremos el método de solución del problema del transporte en la programación lineal.

Introducción al tema. Las desigualdades de primer grado o lineales son aquellas que tienen la forma

$$ax + by + c \geq 0$$

(para mayor simplicidad damos una desigualdad con dos incógnitas x e y). La teoría de sistemas de desigualdades lineales es una rama de la matemática, si bien no extensa, sumamente interesante. El interés hacia ella, en gran parte, es debido a la belleza de su contenido geométrico, pues, traducido al lenguaje de la geometría, la definición de un sistema de desigualdades lineales con dos o tres incógnitas significa la representación de una región o recinto convexo y poligonal en un plano; o, respectivamente, la representación de un cuerpo convexo y poliédrico en el espacio.

Así, por ejemplo, la teoría de los poliedros convexos, que es una parte de la geometría tan antigua como el mundo, se convierte en uno de los capítulos de la teoría de sistemas de desigualdades lineales. Esta teoría tiene partes que agradan al corazón del algebrista; entre ellas, por ejemplo, la formidable analogía entre las propiedades de sistemas de desigualdades lineales y las propiedades de sistemas de ecuaciones lineales (todo lo relacionado con las últimas ha sido estudiado hace mucho tiempo y con mucho detalle).

Hasta no hace mucho tiempo podía pensarse que las desigualdades lineales serían siempre objeto de creación puramente matemática. La situación cambió radicalmente a mediados de los años 40 del siglo xx, al

surgir una nueva rama de la matemática aplicada –la programación lineal– con importantes aplicaciones en la economía y en la técnica. A fin de cuentas, la programación lineal no es más que una de las partes (aunque muy importante) de la teoría de sistemas de desigualdades lineales.

Precisamente, este breve desarrollo tiene por objeto discutir el conocimientos sobre distintos aspectos de la teoría de sistemas de desigualdades lineales: sobre la parte geométrica de la cuestión y los métodos de solución de sistemas estrechamente relacionados con ella; sobre algunas propiedades puramente algebraicas de estos sistemas y sobre cuestiones de la programación lineal. Para su lectura no son precisos conocimientos que superen el curso escolar de matemática.



Una aventura al principio del siglo xx

Breve historia de lo tratado en esta sección

Si bien, por su objeto, la teoría de los sistemas de desigualdades lineales forma, al parecer, una de las partes más importantes y básica de la matemática, hasta no hace mucho tiempo se le ha dedicado relativamente poco espacio.

A partir de los últimos años del siglo XIX comenzaron a aparecer trabajos que tratan de unas u otras propiedades de los sistemas de desigualdades lineales. A este respecto pueden mencionarse nombres de matemáticos tales como Hermann Minkowski, uno de los más grandes geómetras de finales del siglo XIX y comienzos del siglo XX, conocido particularmente por sus trabajos sobre variedades convexas y como autor de la *geometría de Minkowski*; Gueorgui Feodosievich Voronoy (uno de los fundadores, en Petersburgo, de la escuela de la teoría de los números); Alfréd Haar (matemático húngaro, conocido por sus trabajos sobre la integración por grupos); Hermann Weyl (uno de los más prominentes matemáticos de la primera mitad de siglo XX, sobre su vida y obra se puede leer en el opúsculo de I. M. Yaglom “Herman Weyl”, editorial “Znanie”, M., 1967). Algunos de los resultados obtenidos por ellos, de una u otra forma, están reflejados en esta propuesta (aunque sin mencionar a sus autores).



De hecho, la teoría de los sistemas de desigualdades lineales comienza a desarrollarse con intensidad a partir de los años 40 a 50 del siglo XX, cuando el desarrollo impetuoso de las asignaturas aplicadas (la programación lineal, convexa y otras variedades; la programación matemática; la así llamada *teoría de los juegos* y otras) hace indispensable el estudio profundo y sistemático de las desigualdades lineales.

A la fecha, una lista completa de artículos y libros sobre las desigualdades lineales contaría, seguramente, con centenares de distintos títulos y autores.

HERMANN MINKOWSKI

Nació el 22 de junio de 1864, Alexotas, Imperio Ruso (ahora Lituania). Murió el 12 de enero de 1909, Göttingen, Alemania. **Resumen:** Minkowski desarrolló una nueva visión del espacio y el tiempo y sentó las bases matemáticas de la teoría de la relatividad.

Su biografía

Los padres de Hermann Minkowski fueron Lewin Minkowski, un hombre de negocios, y Rachel Taubmann; Hermann fue su tercer hijo. El hermano mayor de Hermann, Max (1844-1930), se hizo cargo del negocio familiar, pero también era coleccionista de arte y cónsul francés en Königsberg. El segundo hermano Oskar (1858-1931) fue médico, mejor conocido por su trabajo sobre la diabetes, y padre del astrofísico Rudolph Minkowski (1895-1976). Además de Max y Oskar, Minkowski también tenía una hermana mayor, Fanny (1863-1954) y un hermano menor, Toby (1873-1906). Lewin y Rachel Minkowski eran alemanes, aunque su hijo Hermann nació mientras vivían en Rusia. Cuando Hermann tenía ocho años, la familia regresó a Alemania y se instaló en Königsberg, donde Lewin Minkowski dirigió su negocio.



Minkowski mostró por primera vez su talento para la matemática mientras estudiaba en el Gymnasium de Königsberg. Ya en esta etapa de su educación estaba leyendo la obra de Richard Dedekind, Peter G. L. Dirichlet y Carl Friedrich Gauss. Las habilidades sobresalientes que mostró en ese momento se notaron en una carta que Heinrich Weber, entonces en la Universidad de Königsberg, le escribió a Dedekind en 1881. Estudió en la Universidad de Königsberg, a donde ingresó en abril de 1880. Pasó tres semestres en la Universidad de Berlín, incluido el semestre de invierno 1882-1883. Se hizo amigo cercano de David Hilbert mientras estaba en Königsberg, quien era estudiante universitario al mismo tiempo que Minkowski.



Hermann Minkowski

En 1884, mientras estudiaba en Königsberg, Adolf Hurwitz fue nombrado miembro del personal. El estudiante Minkowski pronto se hizo muy amigo del recién nombrado académico Hurwitz. Recibió su doctorado en 1885 de Königsberg, por una tesis titulada *Investigaciones sobre formas cuadráticas, determinación del número de formas diferentes contenidas en un género determinado*. Minkowski se interesó en las formas cuadráticas al principio de sus estudios universitarios. En 1881 la Academia de Ciencias (París) anunció que el Gran Premio de Ciencias Matemáticas de 1883 se otorgaría a una solución para el problema del número de representaciones de un número entero como la suma de cinco cuadrados. Gotthold Eisenstein había dado una fórmula para el número de tales representaciones en 1847, pero no había dado una prueba de ese resultado. De hecho, la Academia de Ciencias planteó un problema para el Gran Premio que ya había sido resuelto por Henry Smith, quien había publicado un esquema de una prueba en 1867; sin embargo, la Academia desconocía las contribuciones de Smith cuando estableció el tema del premio.

Eisenstein había estado estudiando formas cuadráticas en variables con coeficientes enteros en el momento en que publicó su fórmula no probada en 1847, pero como ya estaba enfermo en ese momento, nunca se publicaron los detalles. Minkowski, aunque solo tenía 18 años, reconstruyó la teoría de las formas cuadráticas de Eisenstein y produjo una hermosa solución al problema del Grand Premio. Smith reelaboró su prueba anterior, agregó detalles y la envió a la Academia.

Se decidió que el premio fuera compartido entre Minkowski y Smith; de todos modos este fue un comienzo sorprendente para la carrera matemática de Minkowski. Así, el 2 de abril de 1883 la Academia otorgó el Gran Premio de Matemáticas conjuntamente al joven Minkowski, al comienzo de su carrera, y al anciano Smith, al final de la suya. La tesis doctoral de Minkowski, presentada en 1885, fue una continuación del trabajo premiado e implicaba su definición natural del género de una forma. Después de obtener su doctorado, continuó realizando investigaciones en Königsberg.

En 1887, quedó vacante una cátedra en la Universidad de Bonn, y Minkowski solicitó ese puesto; de acuerdo con las regulaciones de las universidades alemanas, tenía que presentar oralmente a la facultad un trabajo original. Minkowski presentó *La visión espacial y los mínimos definen formas cuadráticas*, que no se publicó hasta 1991. Jean Dieudonné escribió sobre el tema: "Esta conferencia es particularmente interesante, ya que contiene el primer ejemplo del método que Minkowski desarrollaría algunos años más tarde en su famosa geometría de los números".

Minkowski enseñó en Bonn desde 1887 y fue ascendido a profesor asistente en 1892. Dos años más tarde regresó a Königsberg, donde enseñó durante dos años antes de ser nombrado miembro del Politécnico Federal de Zúrich. Allí se convirtió en colega de su amigo Hurwitz, que había sido designado para ocupar el puesto de Georg Frobenius después de que partiera del Politécnico Federal de Zúrich hacia Berlín en 1892. Albert Einstein fue alumno de varios de los cursos que impartía y los dos se interesarían más tarde en problemas similares de la teoría de la relatividad.

Minkowski se casó con Auguste Adler en Estrasburgo en 1897; tuvieron dos hijas, Lily, nacida en 1898, y Ruth, nacida en 1902. La familia abandonó Zúrich el año en que nació su segunda hija porque Minkowski

aceptó una cátedra en la Universidad de Göttingen en 1902. Fue Hilbert quien dispuso que la silla se creara especialmente para Minkowski y la mantuvo por el resto de su vida. En Göttingen se interesó por la física matemática y se entusiasmó con Hilbert y sus colegas. Participó en un seminario sobre teoría de electrones en 1905 y aprendió los últimos resultados y teorías en electrodinámica.



Hermann Minkowski

Minkowski desarrolló una nueva visión del espacio y el tiempo y sentó las bases matemáticas de la teoría de la relatividad. En 1907 se dio cuenta de que el trabajo de Hendrik Lorentz y Einstein podría entenderse mejor en un espacio no euclidiano. Consideró que el espacio y el tiempo, que anteriormente se pensaba que eran independientes, estaban acoplados en un “continuo espacio-tiempo” de cuatro dimensiones. También elaboró un tratamiento de cuatro dimensiones de la electrodinámica. Sus principales obras en esta área son *Espacio y tiempo* (1907) y *Dos aspectos sobre las ecuaciones fundamentales de la electrodinámica* (1909). Morris Kline, revisando este tema, escribe:

“Un punto clave del documento es la diferencia en el enfoque de los problemas físicos adoptado por los físicos matemáticos en comparación con los físicos teóricos.

En un artículo publicado en 1908, Minkowski reformuló el artículo de Einstein de 1905 introduciendo la geometría no euclidiana de cuatro dimensiones (espacio-tiempo), un paso en el que Einstein no pensó mucho en ese momento. Pero más importante es la actitud o filosofía que Minkowski, Hilbert –con quien Minkowski trabajó unos años–, Felix Klein y Hermann Weyl consideraban: que las diferencias puramente matemáticas, incluyendo la armonía y la elegancia de las ideas, deberían dominar al adoptar nuevos hechos físicos.

La matemática, por así decirlo, debía ser maestra y la teoría física podía hacerse para inclinarse ante el maestro. Dicho de otro modo, la física teórica era un subdominio de la física matemática, que a su vez era una subdisciplina de la matemática pura. Respecto de este punto de vista, Minkowski siguió a Henri Poincaré, cuya filosofía era que la física matemática, a diferencia de la física teórica, proporcionara nuevos principios físicos. Esta filosofía parecería ser un remanente (modificado, por supuesto) de la visión del siglo XVIII de que el mundo está diseñado matemáticamente y, por lo tanto, que el mundo debe obedecer principios y leyes que los matemáticos descubren, como el principio de acción mínima de Maupertuis, Lagrange y Hamilton. Einstein fue un físico teórico y para él la matemática debe adaptarse a la física”.

[L. Pyenson, “Hermann Minkowski and Einstein’s. Special Theory of Relativity: With an appendix of Minkowski’s ‘Funktiontheorie’ manuscript”, *Arch. History Exact Sci.* 17 (1) (1977), 71-95].

Este continuo espacio-tiempo proporcionó un marco para todo el trabajo matemático ulterior en relatividad. Y estas ideas fueron utilizadas por Einstein para desarrollar la teoría general de la relatividad. De hecho, Minkowski tuvo una gran influencia en Einstein, como señala Leo Corry: “En los primeros años de su carrera científica, Albert Einstein consideraba la matemática como una mera herramienta al servicio de la intuición física. En años posteriores, llegó a considerar a la matemática como la fuente misma de la creatividad científica. Un motivo principal detrás de este cambio fue la influencia de dos destacados matemáticos alemanes: David Hilbert y Hermann Minkowski”. [L. Corry, “The influence of David Hilbert and Hermann Minkowski on Einstein’s views over the interrelation between physics and mathematics”, *Endeavor* 22 (3) (1998), 95-97].

Hemos mencionado varias veces en esta biografía que Minkowski y Hilbert eran amigos cercanos. Menos conocido es el hecho de que Minkowski le sugirió a Hilbert qué tema debía considerar para su famosa conferencia de 1900 en París. Minkowski, en una carta a Hilbert dirigida el 5 de enero de 1900, escribe: “Tendría un gran impacto si intentara dar un panorama del futuro, es decir, un esbozo de los problemas de los que deberían ocuparse los futuros matemáticos. De esta manera, tal vez podría asegurarse de que la

gente hable sobre su conferencia durante décadas". ¡El tiempo ciertamente ha demostrado que Minkowski tenía razón!

El primer Congreso Internacional de Matemáticos se celebró en Zúrich en 1897. Minkowski se unió al comité organizador en diciembre de 1896; es posible que no estuviera en Zúrich para la reunión preliminar de julio. Se unió al comité de diversiones y fue designado para el subcomité responsable de elegir a los oradores. Sugirió invitar a Hilbert a dar una charla en caso de que Klein no pudiera asistir, pero Klein sí asistió al congreso y Hilbert no.

Minkowski también se ofreció a dar una charla él mismo en una de las reuniones de la sección, pero por razones que no se explican en las actas, finalmente no lo hizo. En el congreso presidió la sección 1: Aritmética y Álgebra.

También actuó como uno de los secretarios en el Congreso Internacional de Matemáticos de 1900 en París y dio una charla en la sección I en el Congreso de 1904 en Heidelberg, titulada "Sobre la geometría de los números". En ese momento representó a la Universidad de Göttingen, al igual que en el Congreso de 1908 en Roma.

Los intereses matemáticos originales de Minkowski estaban en la matemática pura y pasó gran parte de su tiempo investigando formas cuadráticas y fracciones continuas. Su logro más original, sin embargo, fue su geometría de los números, que inició en 1890. *Geometría de los números* se publicó por primera vez en 1910, pero las primeras 240 páginas (de las 256) aparecieron como primera sección en 1896. Fue reimpresso en 1953 por Chelsea, Nueva York, y reimpresso nuevamente en 1968. Minkowski publicó *Aproximaciones diofánticas: una introducción a la teoría de números* en 1907.

Daba cuenta elemental de su trabajo sobre la geometría de los números y de sus aplicaciones a las teorías de la aproximación diofántica y de los números algebraicos. El trabajo sobre la geometría de los números condujo al trabajo sobre cuerpos convexos y a preguntas sobre problemas de empaquetamiento, las formas en que las figuras de una forma dada pueden colocarse dentro de otra figura dada.

Murió repentinamente a causa de la ruptura del apéndice cuando solo tenía 44 años.

1. Algunos hechos de la geometría analítica

→ **1. Operaciones con los puntos.** Tracemos en un plano un sistema rectangular de coordenadas. El punto M tiene en este sistema las coordenadas x e y , que escribimos de la siguiente forma:

$$M = (x \cdot y) \text{ o simplemente, } M = (x, y).$$

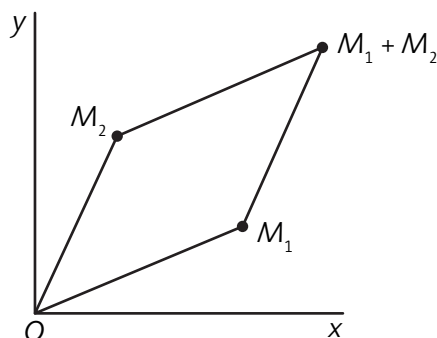
El sistema de coordenadas permite realizar con los puntos del plano algunas operaciones, tales como adición de puntos y multiplicación de puntos por un número.

La adición de puntos se define de la siguiente forma: si $M_1 = (x_1, y_1)$ y $M_2 = (x_2, y_2)$, entonces

$$M_1 + M_2 = (x_1 + x_2, y_1 + y_2).$$

Así, pues, la adición de puntos se reduce a la adición de sus coordenadas con el mismo nombre.

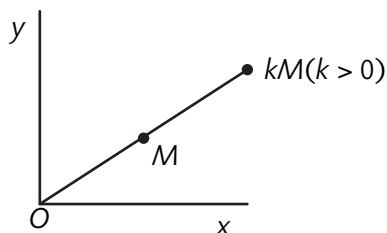
El sentido geométrico de la operación es simple: el punto $M_1 + M_2$ es el cuarto vértice del paralelogramo construido en los segmentos OM_1 y OM_2 como lados (O es origen de coordenadas).



Los otros tres vértices del paralelogramo son M_1 , O , M_2 .



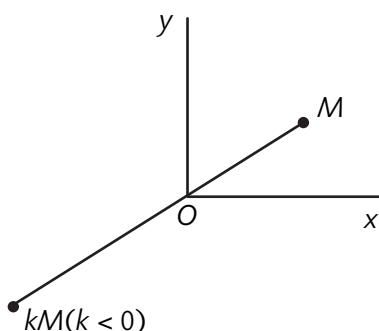
Esto mismo puede ser explicado de otra forma: el punto $M_1 + M_2$ se obtiene del punto M_2 , procediendo a la traslación paralela de este último en dirección al segmento $\overline{OM_1}$ y a una distancia igual a la longitud de este segmento.



La multiplicación del punto $M(x, y)$ por un número arbitrario k se realiza conforme a la regla siguiente

$$kM = (kx, ky).$$

El sentido geométrico de esta operación es aún más simple que el de la operación de adición: siendo $k > 0$, el punto $M' = kM$ yace en el rayo \overline{OM} , además $OM' = k \cdot OM$; siendo $k < 0$, el punto M' yace en la prolongación del rayo \overline{OM} más allá del punto O , además $OM' = |k| \cdot OM$.



La deducción de la interpretación geométrica que acabamos de dar a estas dos operaciones será un ejercicio efectivo e interesante para hacer, si se desconocen los fundamentos de la teoría de los vectores. Como es sabido, desde el punto de vista del cálculo vectorial, nuestras operaciones significan lo siguiente: el punto $M_1 + M_2$ es el extremo del vector $\overline{OM_1} + \overline{OM_2}$, y el punto kM , el extremo del vector $k \cdot \overline{OM}$ (con la condición de que el origen de este vector es el punto O).

Estas operaciones son muy cómodas para traducir hechos geométricos al lenguaje del álgebra. Daremos algunos ejemplos de tales traducciones.

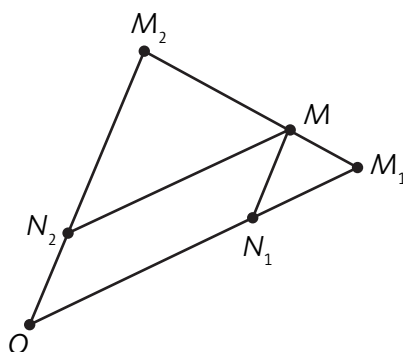
1) El segmento $\overline{M_1M_2}$ está compuesto por todos los puntos de la forma

$$s_1M_1 + s_2M_2, \tag{1}$$

siendo s_1 y s_2 dos número cualesquiera no negativos, cuya suma es igual a 1.

Aquí, un hecho puramente geométrico, la pertenencia de un punto al segmento $\overline{M_1M_2}$ se expresa en forma de relación algebraica $M = s_1M_1 + s_2M_2$, con las limitaciones para s_1 y s_2 establecidas anteriormente.

Para demostrarlo, veamos el punto arbitrario M situado en el segmento $\overline{M_1M_2}$. Trazando líneas rectas, a través del punto M , paralelas a $\overline{OM_2}$ y $\overline{OM_1}$, obtenemos el punto N_1 , situado en el segmento $\overline{OM_1}$, y el punto N_2 , situado en el segmento $\overline{OM_2}$.



Supongamos que

$$s_1 = \frac{\overline{M_2M}}{\overline{M_2M_1}}, \quad s_2 = \frac{\overline{M_1M}}{\overline{M_1M_2}}$$

los números s_1 y s_2 no son negativos y su suma se iguala a 1. De la semejanza de los respectivos triángulos hallamos

$$\frac{\overline{ON_1}}{\overline{OM_1}} = \frac{\overline{M_2M}}{\overline{M_2M_1}} = s_1, \quad \frac{\overline{ON_2}}{\overline{OM_2}} = \frac{\overline{M_1M}}{\overline{M_1M_2}} = s_2,$$

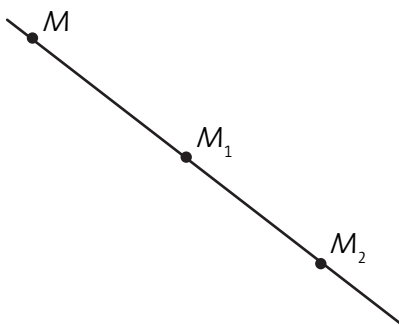
de donde se deduce: $N_1 = s_1M_1$; $N_2 = s_2M_2$. Pero $M = N_1 + N_2$, por lo tanto $M = s_1M_1 + s_2M_2$. Por último, observaremos que cuando el punto M recorre el segmento $\overline{M_1M_2}$ en dirección de M_1 a M_2 , el número s_2 adquiere todos los valores desde 0, hasta 1. La afirmación 1 queda demostrada.

2) Cualquier punto M en la recta $\overline{M_1M_2}$ se expresa de la forma

$$tM_1 + (1-t)M_2,$$

siendo t un cierto número.

Verdaderamente, si el punto M yace en el segmento $\overline{M_1M_2}$, entonces nuestra afirmación se deduce de lo anteriormente demostrado. Supongamos que M yace fuera del segmento $\overline{M_1M_2}$. Entonces, o el punto M_1 yace en el segmento $\overline{MM_2}$,



o M_2 yace en el segmento $\overline{MM_1}$. Supongamos, por ejemplo, que tiene lugar el primer caso. Entonces, conforme a lo ya demostrado,

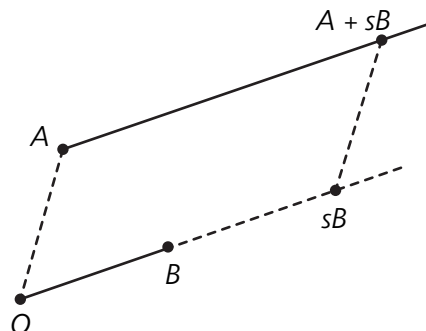
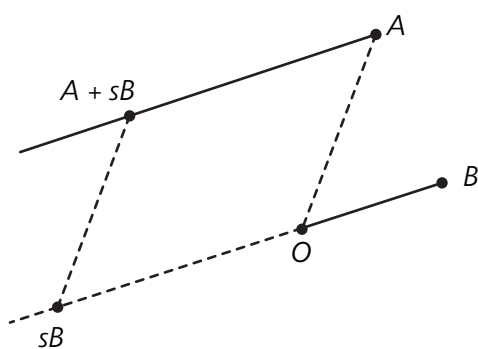
$$M_1 = sM + (1-s)M_2 \quad (0 < s < 1)$$

De donde

$$M = \frac{1}{s}M_1 - \frac{1-s}{s}M_2 = tM_1 + (1-t)M_2,$$

siendo $t = \frac{1}{s}$. Proponemos que se examine el caso en que M_2 yace en el segmento $\overline{MM_1}$.

3) Cuando el parámetro s crece desde 0 hasta ∞ , el punto sB recorre el rayo \overline{OB} (se supone que el punto B es distinto del de origen de coordenadas 0) y el punto $A + sB$, el rayo saliente de A en dirección a OB . Cuando s decrece desde 0 hasta $-\infty$, los puntos sB y $A + sB$ recorren rayos complementarios a los indicados anteriormente. Para demostrarlo basta con fijarse en las figuras de abajo.



De la afirmación 3) se deduce que cuando varía desde $-\infty$ hasta $+\infty$: el punto $A + sB$ recorre la recta que pasa por A paralelamente a \overline{OB} .

Está claro que las operaciones de adición y multiplicación por un número también pueden realizarse con puntos situados en el espacio. En este caso, conforme a la definición.

$$M_1 + M_2 = (x_1 + x_2, y_1 + y_2, z_1 + z_2).$$

$$kM = (kx, ky, kz).$$

Lógicamente todas las afirmaciones demostradas anteriormente serán justas también en el espacio.

Para dar fin a esta parte adoptaremos un acuerdo, que más adelante nos servirá para formular muchos hechos con más claridad y laconismo. Precisamente, si \mathcal{K} y \mathcal{L} son dos conjuntos cualesquiera de puntos (en un plano o espacio), acordaremos considerar por su "suma" $\mathcal{K} + \mathcal{L}$ el conjunto de todos los puntos de la forma $K + L$, siendo K un punto arbitrario de \mathcal{K} , y L , un punto arbitrario de \mathcal{L} .

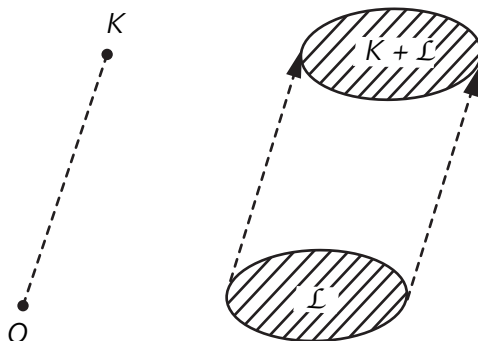
En la matemática hace ya tiempo que se utiliza una escritura especial para indicar la pertenencia de cualquier punto a un conjunto determinado. Concretamente, para indicar que el punto M pertenece al conjunto \mathcal{M} se escribe $M \in \mathcal{M}$ (el signo \in sustituye la palabra "pertenece"). O sea, $\mathcal{K} + \mathcal{L}$ es el conjunto de todos los puntos de la forma $K + L$, siendo $K \in \mathcal{K}$ y $L \in \mathcal{L}$.

Partiendo del sentido geométrico de la adición de puntos, podemos deducir una regla simple para adicionar los conjuntos puntuales \mathcal{K} y \mathcal{L} . Esta regla consiste en lo siguiente: para cada punto $K \in \mathcal{K}$ es preciso formar un conjunto que resulta de \mathcal{L} mediante su traslado paralelo al segmento \overline{OK} : a continuación todos los conjuntos obtenidos de esta forma se unen en uno. Este último será $\mathcal{K} + \mathcal{L}$.

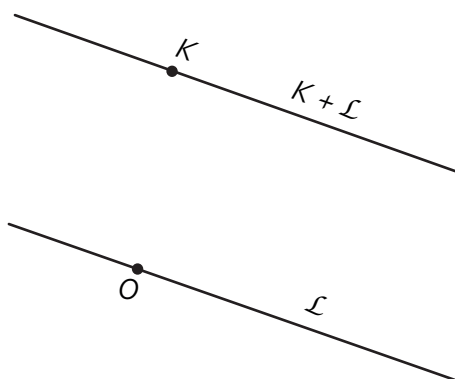
Daremos algunos ejemplos.

- I. Supongamos que el conjunto \mathcal{K} está compuesto por un punto K , mientras que \mathcal{L} contiene un conjunto de puntos cualquiera.

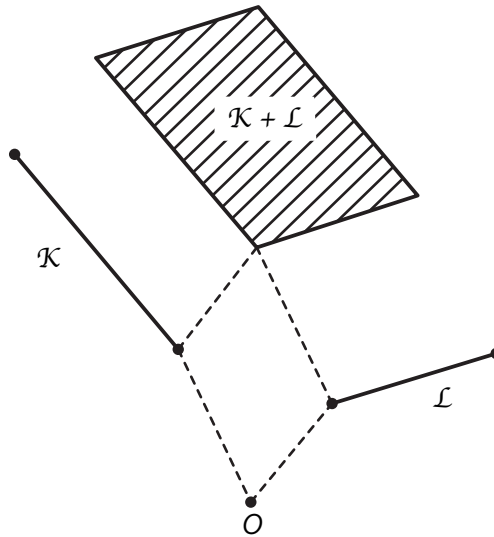
El conjunto $K + \mathcal{L}$ es el resultado de un traslado paralelo del conjunto \mathcal{L} al segmento OK .



En particular, si \mathcal{L} es una recta, entonces $K + \mathcal{L}$ será una recta paralela a \mathcal{L} . Si en este caso la recta \mathcal{L} pasa por el origen de las coordenadas, entonces $K + \mathcal{L}$ será una recta paralela a \mathcal{L} que pasa por el punto K .

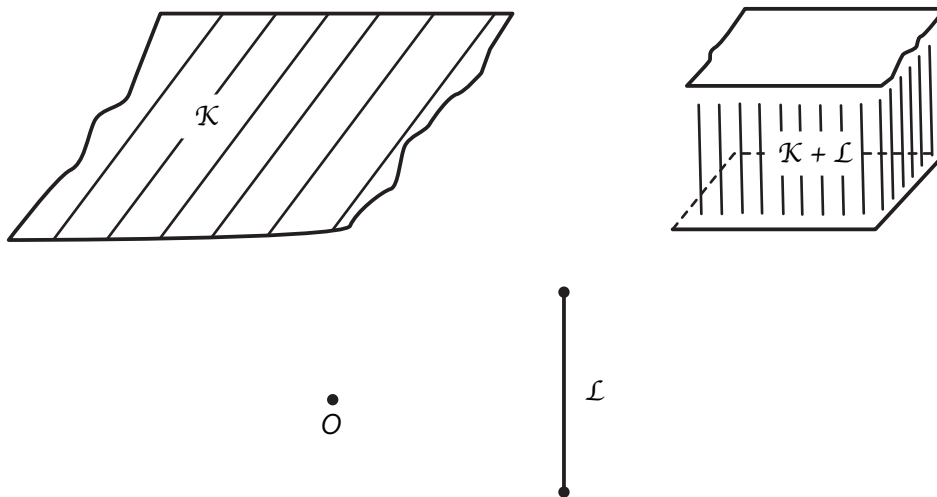


II. \mathcal{K} y \mathcal{L} son segmentos (en un plano o en el espacio) no paralelos entre sí.



Entonces el conjunto $\mathcal{K} + \mathcal{L}$ es un paralelogramo con lados iguales y paralelos a \mathcal{K} y \mathcal{L} (respectivamente).
 ¿Qué resulta cuando los segmentos \mathcal{K} y \mathcal{L} son paralelos?

III. \mathcal{K} es un plano y \mathcal{L} es un segmento no paralelo a este plano. El conjunto $\mathcal{K} + \mathcal{L}$ es una parte del espacio, comprendida entre dos planos paralelos a \mathcal{K} .



IV. \mathcal{K} y \mathcal{L} son círculos de radios r_1 y r_2 con sus respectivos centros P_1 y P_2 , que yacen en un mismo plano π .
 Entonces: $\mathcal{K} + \mathcal{L}$ es un círculo que yace en un plano paralelo a π con radio $r_1 + r_2$ y centro en el punto $P_1 + P_2$.

