Leñitas Geométricas*

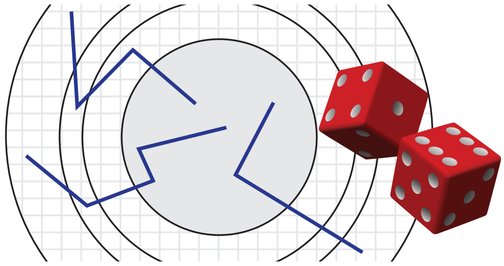
para el Fogón Matemático de los Festivales

De OMA para Profesores y Maestros en actividad

5ª época ✈ Nº 3
6 de abril de 2023

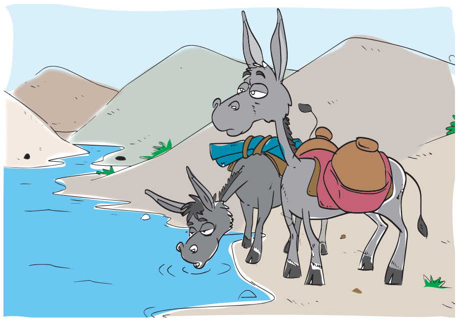
"[...] más valioso que acumular información es interiorizarse de métodos para saber qué propósitos tienen, es decir, de dónde parten y adónde llevan. Los métodos tienen una orientación y una dinámica de las cuales los resultados y teoremas aislados carecen". *Dr. Alberto Calderón*

El Método Montecarlo



Una aventura a la matemática de las chances

Probabilidades finitas y esperanzas. Probabilidad: concepto y orígenes



Resulta revelador –y curioso– el hecho muchas veces pasado por alto de que la moderna teoría de la probabilidad, con su alto grado de abstracción y su creciente aplicabilidad, deba su origen casi por completo a problemas planteados en el ámbito del juego. Habiendo ya hablado de la vida y la obra de Girolamo Cardano (véase *Leñitas Geométricas* Nº 2, 5ª época), describiremos a continuación dos problemas relativos al juego que suelen ser citados como el inicio de la ciencia de la probabilidad.

El filósofo francés Antoine Gombaud, conocido como el Caballero de Méré, fue un gentleman del siglo XVII con bastante experiencia en juegos de azar y, aparentemente, con buen sentido para intuir las diferentes posibilidades ofrecidas por las distintas apuestas en el juego de los dados. Luego de construir a lo largo de los años una fortuna considerable apostando a obtener un seis en cuatro lanzamientos de un dado, De Méré razonó, siguiendo un argumento común en su época, que apostar a sacar un doble seis en 24 lanzamientos con dos dados sería igual de provechoso.

* 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.

Hay que señalar, en descargo del caballero, que después de dolorosos ensayos llegó a la conclusión de que esta última apuesta no le reportaba ningún beneficio. Por lo tanto, en 1645 De Méré retó a su célebre amigo Blaise Pascal a ofrecerle una explicación.

En una correspondencia mantenida entre Pascal y Pierre de Fermat se resolvió el problema propuesto por De Méré y, de paso, también surgieron de allí la idea formal de probabilidad, el famoso triángulo de Pascal y la omnipresente distribución binomial.

Dejemos los cálculos que muestran que apostar por un doble seis en 24 lanzamientos de dos dados no es una decisión inteligente para los ejercicios y festivales. Mientras tanto, vayamos calentando motores probando que es matemáticamente ventajoso apostar a sacar un seis en cuatro lanzamientos.

Usando un truco muy común en el cálculo de probabilidades, daremos la vuelta a la pregunta tratando de hallar la probabilidad de no obtener ningún seis en cuatro lanzamientos. En cada uno de los lanzamientos, la probabilidad de no obtener seis es 5/6. Decimos que los diferentes lanzamientos son independientes si el resultado de uno cualquiera no influye en los resultados de los demás.

Pronto precisaremos el hecho de que las probabilidades de sucesos independientes pueden ser multiplicadas para obtener la probabilidad de todos los sucesos. Puesto que los diferentes lanzamientos de un dado son sucesos claramente independientes, la probabilidad de que no salga un seis en cuatro lanzamientos es

$$\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}=\frac{5^4}{6^4}=\frac{625}{1296}$$

por lo tanto, la probabilidad de sacar un seis en cuatro lanzamientos es

$$1-\frac{625}{1296}=\frac{671}{1296},$$

una probabilidad algo mayor que 1/2. Así pues, la apuesta al seis simple es ganadora.

Un segundo problema solucionado en la correspondencia entre Pascal y Fermat es el "problema de los puntos". Estudiaremos seguidamente una versión adaptada de este problema. Jimmy y Walter están en plena partida de un juego en el que se tiene que conseguir cinco puntos para ganar, y cada uno de los jugadores tiene las mismas oportunidades para vencer en una ronda y llevarse un punto. Jimmy está ganando por 4 a 3, cuando llega la policía y se interrumpe la partida.

Asumiendo que el vencedor tiene que recibir alguna recompensa, ¿cómo deberán repartirse las apuestas depositadas? La respuesta "correcta" a esta pregunta depende de una gran cantidad de suposiciones, tanto morales como sociológicas y matemáticas, que debe hacer quien intente solucionar el problema, por lo que puede haber más de una solución admisible. La esencia de una buena solución es que, bajo presupuestos claros y razonados, la respuesta no sea solamente plausible sino correcta en forma demostrable. Ilustraremos esto último resolviendo el problema de los puntos.

Nuestro objetivo consiste en encontrar la probabilidad de que Jimmy ganara el juego si este hubiera llegado hasta el final, y dividir las apuestas de acuerdo con esta probabilidad. Es fácil ver que de haberse jugado dos rondas más, alguno de los dos habría ganado. Dado un punto en juego, denotemos con J si lo gana Jimmy y con W si es para Walter; puesto que quedaban dos rondas por jugarse, tendremos cuatro posibles resultados (JJ, JW, WJ y WW), cada uno de los cuales, debido a lo que hemos asumido, son igualmente probables.

Punto 8	Punto 9	Ganador
J	J	J
J	W	J
W	J	J
W	W	W

Una simple inspección de la tabla anterior conduce a la conclusión de que la probabilidad de que Jimmy ganase es 3/4 mientras que la de Walter es 1/4. Por lo tanto, el dinero debería repartirse a razón de 3 a 1 a favor de Jimmy.

La solución anterior y el ejemplo de los dados ilustran la naturaleza ingeniosa, pero a la vez simple, del cálculo de probabilidades. Frecuentemente, para resolver un problema se puede escoger entre un gran número de argumentos, pero solo aquellos que obedezcan a las muy bien definidas leyes de la teoría de la probabilidad conducirán a soluciones válidas y consistentes. En lo que sigue, en primer lugar calcularemos algunas probabilidades, con la esperanza de aclarar la idea intuitiva de probabilidad; y más adelante presentaremos las definiciones esenciales de la teoría; para terminar formalizando las reglas del cálculo de probabilidades.

Dados, cartas y probabilidades

Consideremos un par de dados no cargados, en los que cada una de las seis caras tiene la misma probabilidad (a saber 1/6) de aparecer tras un lanzamiento. Si los dos dados son lanzados simultáneamente (o secuencialmente, que para el caso es lo mismo), existen 36 resultados posibles y equiprobables, listados en la figura de abajo. Para distinguir los dados, aunque los resultados no dependen del par concreto de dados no trucados que se utilicen, supongamos que uno es verde y el otro, blanco.

					↓					↓					↓					↓					↓											
Verde	1	1	1	1	1	1	2	2	2	2	2	2	3	3	3	3	3	3	4	4	4	4	4	4	5	5	5	5	5	5	6	6	6	6	6	6
Blanco	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6

Es importante darse cuenta de que, para sacar un 6 y un 1, podemos hacerlo de dos formas distintas (Verde 1, Blanco 6 y Verde 6, Blanco 1), por lo que la probabilidad de este resultado es el doble de la probabilidad de sacar un doble 3.

Sabiendo que cada uno de los 36 posibles resultados de lanzar dos dados tiene una probabilidad de 1/36, es sencillo calcular la probabilidad de que la suma de los puntos de los dados sea un determinado número entre 2 y 12; basta con contar el número de maneras de conseguirlo. Así, puesto que hay 5 resultados diferentes que totalizan 6 puntos (obsérvese la figura anterior), la probabilidad de este suceso es de 5/36. En el cuadro siguiente se muestra un análisis completo de las probabilidades para este problema.

Total	2	3	4	5	6	7	8	9	10	11	12
Modos de lograrlo	1	2	3	4	5	6	5	4	3	2	1
Probabilidad	1/36	1/18	1/12	1/9	5/36	1/6	5/36	1/9	1/12	1/18	1/36

El estudio que hemos realizado de las probabilidades en los juegos de dados no es exhaustivo; quedan aún muchas cuestiones por contestar que dejaremos para más adelante. Por ahora debemos familiarizarnos con el manejo de los argumentos de enumeración que se han expuesto. Para ello continuemos con algunos cálculos relativos a juegos de cartas.

Una baraja normal de póker o *bridge* tiene 52 cartas, 13 de cada palo. Si asumimos que una baraja se ha barajado perfectamente y que el reparto es honesto, entonces cada carta tiene la misma probabilidad de ser repartida. De este modo, si un jugador A no conoce las cartas que se han repartido, la probabilidad que tiene de recibir una carta de un determinado palo es 13/52 = 1/4; la probabilidad de recibir una carta de un rango determinado (como un rey o un tres) es 4/52 = 1/13; y la probabilidad de recibir una carta determinada (como la reina de picas) es 1/52.

El razonamiento que conduce a estas conclusiones es análogo al del par de dados, donde había 36 posibles resultados equiprobables. Una observación, que puede incomodar a algunos, es que las probabilidades anteriores no se ven alteradas por el hecho de que algunas cartas hayan sido ya repartidas a otros jugadores, siempre que estas no les sean mostradas al jugador A. Se requiere convencerse de esto y de que, desde el punto de vista de la probabilidad, no importa el orden en el que se repartan las cartas en una mano; claro está: ¡siempre y cuando no sea para hacer trampas!

Cuando algunas de las cartas repartidas le son mostradas a un jugador, surge una importante diferencia respecto del problema de los dados. Las cartas son repartidas sin reemplazamiento, por lo que si un corazón (o incluso un trébol) ha sido extraído y mostrado, eso afecta a la probabilidad de que al realizar una segunda extracción la carta obtenida sea de corazones.

Este hecho debe ser confrontado con lo que sucede al lanzar dos dados. Aquí, obtener en una tirada un seis doble no afecta –a pesar de las teorías acerca de que “los dados están calientes” y “está al caer”– a la probabilidad de obtener otro doble seis en el siguiente lanzamiento. Volviendo a nuestra baraja, esperemos que, después de razonar sobre resultados equiprobables y de hacer algunos cálculos, se convencerá de que:

a) La probabilidad de formar una pareja al extraer dos cartas de una baraja (dos cartas de igual rango) es de 3/51.

b) La probabilidad de obtener una sota o un diamante en una extracción es de 16/52.

c) Si el jugador A ha visto 9 cartas, una de las cuales es un siete, la probabilidad de que la siguiente carta sea un siete es de 3/43.

d) La probabilidad de obtener 5 picas (en el póker, obtener color) extrayendo 5 cartas es

$$\frac{13}{52}\cdot\frac{12}{51}\cdot\frac{11}{50}\cdot\frac{10}{49}\cdot\frac{9}{48}=0{,}000495,$$

es decir, algo menor que 1/2 000.

Indicación. Obtener una pica en la primera extracción y obtener una pica en la segunda sabiendo que la primera fue una pica son sucesos independientes. La probabilidad de obtener un par de picas en dos extracciones sucesivas es el producto de las probabilidades de estos sucesos independientes, e.d. (13/52) (12/51). Continúe con este razonamiento.

Ruleta, probabilidad y expectativas

Como demostraron los ejemplos anteriores, la probabilidad de un determinado suceso es un número entre 0 y 1 que mide la expectativa de realización de ese suceso. En los casos que ya hemos visto, cada suceso podía ser descompuesto en varios acontecimientos elementales equiprobables (el número total de sucesos elementales es 36 para el experimento "lanzar un par de dados" y 52 para el experimento "extraer una carta de una baraja completa"). La probabilidad de un suceso puede ser calculada, después de una enumeración cuidadosa, como sigue:

$$\text{Probabilidad del suceso } E = \frac{\text{Número de sucesos elementales en } E}{\text{Número de sucesos elementales igualmente probables}}.$$

Hay muchos experimentos aleatorios en los que los sucesos no pueden ser descompuestos en sucesos elementales equiprobables (por ejemplo, el suceso "nevada sobre San Francisco el día de Navidad"). Cuando en un experimento aleatorio cabe la posibilidad de efectuar un número indefinido de pruebas, se define desde un punto de vista empírico la probabilidad de un determinado suceso *E*, relacionado con este experimento, como:

$$\text{Probabilidad del suceso } E \approx \frac{\text{Número de pruebas con resultado de } E}{\text{Número total de pruebas}}$$

donde el símbolo ≈ significa "aproximadamente igual a", y la aproximación es tanto mejor en tanto aumenta el número de pruebas.

Aunque ambas definiciones son intuitivas e importantes para la simulación y el análisis de los juegos de azar, deberíamos caer en la cuenta de que las dos nociones de probabilidad enunciadas son conceptualmente distintas. Estas diferencias conceptuales son el centro de la controversia filosófica en torno de los fundamentos de la teoría de la probabilidad.

Afortunadamente, los dos enfoques están vinculados por un resultado decisivo de la teoría de la probabilidad, que establece que las dos definiciones coinciden en el límite cuando el número de pruebas tiende al infinito. El estudio en profundidad de este resultado queda para más adelante.

Una vez que la probabilidad de un suceso *E*, $p(E)$ ha sido definida, es común referirse a la expectativa (o expectativa a favor) de *E* y a la expectativa en contra de *E*. Estos conceptos se pueden definir formalmente como:

$$\text{Expectativa de } E = \frac{p(E)}{1-p(E)},$$

$$\text{Expectativa en contra de } E = (\text{Expectativa de } E)^{-1}.$$

Estos números se suelen expresar (siempre que sea posible) como el cociente de dos números enteros. Así, si la probabilidad de *E* es 1/4 tenemos,

$$\text{expectativa de } E = \frac{1/4}{3/4} = \frac{1}{3}.$$

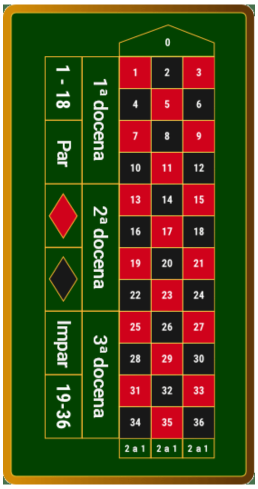
Recíprocamente, si la expectativa de un suceso *E* es a/b, no es difícil probar que:

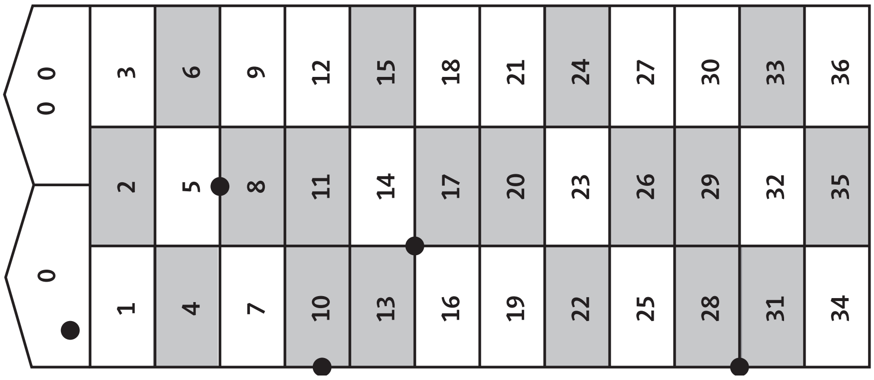
$$p(E) = \frac{a}{(a+b)}.$$

Usaremos la notación 1:3 (léase 1 a 3) para referirnos a la expectativa de *E*. Como era de esperar, la expectativa en contra de *E* es la inversa de la expectativa de *E*; así, en el ejemplo anterior, donde la expectativa de *E* era 1:3, la expectativa en contra de *E* será 3:1. Por lo tanto, si un suceso *E* tiene como expectativa en contra 7:3, tendrá como expectativa 3:7 y una probabilidad de 3/(3 + 7) = 3/10 = 0,3.

Todo esto se complica –a los jugadores entusiastas poco les importa la complicación pues asimilan estos conceptos con facilidad– con la costumbre de los apostadores de hablar más de las expectativas en contra que de las expectativas a favor. En este caso, el numerador de la expectativa en contra indica el beneficio que se puede obtener cuando se apuesta la cantidad que refleja el denominador. Por ejemplo, si apostamos 2 dólares estando la expectativa (en contra) 7:2 y ganamos, embolsaremos 7 dólares (más la devolución de lo apostado, 2 dólares).

A fin de clarificar los conceptos anteriores, los utilizaremos en el estudio de la ruleta de Las Vegas. Esta ruleta está dividida en 38 sectores iguales (y se supone que equiprobables) numerados del 1 al 36, más el 0 y 00. Los sectores del 1 al 36 están divididos en 18 sectores rojos y 18 sectores negros (los sectores 0 y 00 son verdes).





El esquema de arriba muestra el tapete, o mesa de apuestas de la ruleta, e ilustra algunos ejemplos de apuestas válidas. Los números sombreados indican que sus correspondientes en la ruleta son negros. Además de las apuestas explícitamente señaladas en el tapete, hay otras posibles que están indicadas con pequeños círculos negros. Estas apuestas significan (de arriba hacia abajo en el esquema):

- una apuesta al número 0 es admisible (cualquier otra apuesta a un solo número también es admisible);
- una apuesta al par de números contiguos (en el tapete) 5 y 8 (dos posibilidades para ganar);
- una apuesta a los números 10, 11, 12, que están en una misma fila (tres posibilidades para ganar);
- una apuesta a los números 13, 14, 16, 17;
- una apuesta a los números del 28 al 33 (dos filas, seis posibilidades de ganar).

Sea *E* el suceso "salir 10, 11 o 12 en la ruleta"; usando nuestra primera noción de probabilidad, obtenemos que la probabilidad de *E* es 3/38. Por lo tanto, las expectativas a favor y en contra de *E* serán 3:35 y 35:3, respectivamente.

Expectativas reales y expectativas para la banca en una ruleta de Las Vegas

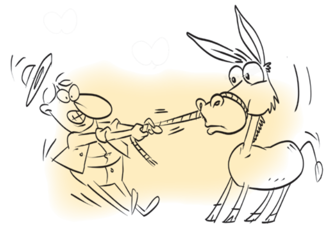
Tipo de apuesta*	Expectativas reales	Expectativas de la banca
Color (Rojo o Negro)	20:18	1:1
Paridad (Par o impar)	20:18	1:1
18# (1-18 o 19-36)	20:18	1:1
12# (columnas o docenas)	26:12	2:1
6# (dos filas)	32:6	5:1
4#	34:4	8:1
3# (cualquier fila)	35:3	11:1
2# (números adyacentes)	36:2	17:1
1#	37:1	35:1

* *El símbolo # denota número o números.*

En la tabla anterior se presenta una lista completa de las posibles apuestas en una ruleta de Las Vegas junto con sus expectativas reales (en contra), y con las llamadas *expectativas de la banca*. Las expectativas de la banca son todas de la forma *r*:*l*, donde *r* representa la cantidad de dólares que un jugador ganaría si apostase 1 dólar.

Después de comparar las expectativas reales con las de la banca, no debería extrañarnos el cariño que los dueños de los casinos sienten por la ruleta. Podremos comprobarlo en los ejemplos que desarrollaremos acerca de este tema.

Probabilidades compuestas: las reglas del juego



Luego de estudiar varios ejemplos del cálculo de probabilidades mediante la enumeración de sucesos simples, pasamos a presentar las reglas para calcular las probabilidades de sucesos compuestos. En lo que sigue necesitaremos la siguiente terminología:

- El suceso no *A* ocurre cuando el suceso *A* no ocurre durante un determinado experimento.
- El suceso *A* o *B* ocurre, durante un experimento dado, si *A* ocurre o *B* ocurre u ocurren ambos. Si es imposible que *A* y *B* se den simultáneamente, se dice que son sucesos disjuntos.
- El suceso *A* y *B* ocurre cuando en sucesivos experimentos independientes –es decir, cuando el resultado del primer experimento no influye en el resultado del segundo–, *A* ocurre en el primer experimento y *B* en el segundo. [Ejemplo: A = 1ª carta de una baraja es una pica; B = 1ª carta de una baraja es roja; cada experimento consiste en extraer la primera carta de una baraja completa de 52 naipes].

Reglas probabilísticas, explicaciones y ejemplos

Regla	Explicaciones y ejemplos
1. Para todo suceso *E* $0 \le p(E) \le 1$	$0 \le$ nº de sucesos $\le$ nº total de tiradas, así $0 \le \frac{\text{nº de sucesos}}{\text{nº de tiradas}} \le 1$
2. p(suceso imposible) = 0 p(suceso seguro) = 1	p(sumar 13 lanzando 2 dados) = 0 p(extraer un as con 49 extracciones de una baraja de *bridge*) = 1
3. $p(\text{no } E) = 1 - p(E)$	p(sacar al menos un 6 en el lanzamiento de dos dados) = 1 – p (ningún 6 en el lanzamiento de dos dados) = 1 – 25/36 = 11/36
4. Si *A* y *B* son sucesos disjuntos $p(A \text{ o } B) = p(A) + p(B)$	A = repartir un as rojo B = repartir una carta negra $p(A \text{ o } B) = 2/52 + 26/52 = 28/52 = 7/13$
5. Si *A* y *B* son sucesos en dos experimentos sucesivos e independientes, $p(A \text{ y } B) = p(A) \cdot p(B)$	A = extraer un as o un rey de una baraja B = al lanzar dos dados sacar en cada uno de ellos los mismos puntos $P(A \text{ y } B) = \frac{8}{52} \cdot \frac{1}{6} = \frac{1}{39}$

En la tabla anterior resumimos las reglas básicas de la teoría de la probabilidad (llamadas *axiomas* en exposiciones más formales). Dado un suceso *E*, denotamos con $p(E)$ a la probabilidad de que se verifique u ocurra el suceso *E*. De manera análoga, $p(A \wedge B)$ denota la probabilidad del suceso compuesto *A* y *B*.

Estas reglas representan una parte importante de la matemática de los juegos y las apuestas. Aunque son sencillas de establecer y fáciles de manejar, en casos no muy complicados, su aplicación correcta en general es difícil y requiere entender claramente la naturaleza de los sucesos que se estudian y qué suposiciones podemos hacer sobre ellos.

De hecho, la correcta aplicación de estas reglas es fuente de frustración, aunque también de deleite, para muchos estudiantes de la teoría elemental de la probabilidad. Aunque no dependeremos de la vaga capacidad de razonar con las probabilidades, las emplearemos cuando sea necesario, por lo que animamos a dominar su uso, al menos de forma elemental. En particular, los conceptos de sucesos disjuntos y experimentos independientes deberían ser completamente entendidos. Los ejercicios ayudarán a alcanzar este objetivo.

Esperanza matemática y sus aplicaciones

Como ya dijimos, la idea más importante a la hora de tomar decisiones racionales en situaciones de incertidumbre –una tarea muy común a la hora de apostar a algo y también en los juegos– es el concepto de *esperanza matemática*. En este apartado volvemos sobre este concepto con nuevos ejemplos, para después enunciar una formulación general que se aplicará en diversas situaciones.

Uno de nuestros resultados más significativos (en este caso, negativo) es que el tipo de apuesta depositada en la mesa de una ruleta de Las Vegas no influye en las ganancias o las pérdidas previstas –se probará que el tamaño de la apuesta no tiene ningún efecto en la esperanza matemática del jugador por unidad de dinero apostado–.

Esperanza, ejemplo 1. Le ofrecen participar de un juego que consiste en lanzar dos monedas al aire (no trucadas). Los pagos son:

Dos caras: gana \$2

Dos cruces: gana \$3

Una cruz y una cara: pierde \$4

¿Debería jugar?, y si juega, ¿cuál será su ganancia media esperada?

Solución. En promedio, durante una partida de cuatro lanzamientos saldrán una vez "dos caras", "dos cruces", otra; y en dos ocasiones, "una de cada". En esta sucesión su ganancia neta será de

$$1 \times \$2 + 1 \times \$3 + 2 \times \$(-4) = \$-3.$$

Es decir, cada cuatro jugadas perderá \$3. Así que no debería jugar, pero si lo hace, deberá enfrentar una pérdida en media de (1/4) × \$2 = 75 centavos en cada lanzamiento. Para obtener este resultado más directamente, obsérvese que $p(\text{ambas caras}) = \frac{1}{4}$, $p(\text{ambas cruz}) = \frac{1}{4}$, $p(\text{"una de cada"}) = \frac{1}{2}$ y

$$\frac{1}{4}(\$2) + \frac{1}{4}(\$3) + \frac{1}{2}(-\$4) = \$-0{,}75.$$

Por lo que llegamos al mismo resultado multiplicando las probabilidades de cada uno de los sucesos por sus correspondientes pagos (las pérdidas van acompañadas de signos negativos).

Esperanza, ejemplo 2. Al lanzar un dado, la banca de un casino le paga 12 fichas si saca un seis, 8 fichas si saca un número impar y nada en otro caso. Para que el juego fuese equilibrado, ¿cuántas fichas debería pagar a la casa en cada lanzamiento?

Solución. Usando el método del ejemplo anterior, tenemos que

$$p(\text{seis}) = \frac{1}{6},\ p(\text{"impar"}) = \frac{1}{2},\ p(\text{"ni seis ni impar"}) = \frac{1}{3}.$$

En un lanzamiento, su esperanza es de

$$\frac{1}{6}\times 12 + \frac{1}{2}\times 8 + \frac{1}{3}\times 0 = 6.$$

Así pues, para que el juego fuese justo, antes de cada lanzamiento debería pagar 6 fichas. Si se introduce el pago de 6 fichas, como tarifa para jugar, en la lista de pagos, y si se repiten los cálculos, se obtendrá una esperanza de

$$\frac{1}{6}\times 6 + \frac{1}{2}\times 2 + \frac{1}{3}\times(-6) = 0 \text{ fichas},$$

lo que indica que (al menos en teoría) el juego es justo para la banca y para el jugador.

Seguidamente estableceremos el procedimiento general para calcular la esperanza matemática.

Sean

$$E_1, E_2, E_3, \ldots, E_n$$

n sucesos disjuntos dos a dos (ningún par puede darse simultáneamente) con probabilidades

$$p_1, p_2, p_3, \ldots, p_n$$

y con pagos

$$r_1, r_2, r_3, \ldots, r_n.$$

Entonces, la "ganancia esperada", o *esperanza matemática X* de un experimento en el que el resultado debe ser uno de estos sucesos, está definida por

$$X = p_1 r_1 + p_2 r_2 + p_3 r_3 + \ldots + p_n r_n.$$

Basándonos en los ejemplos anteriores y en la idea que subyace tras esta fórmula, es conveniente hacer las siguientes puntualizaciones respecto de la *esperanza matemática* de un experimento aleatorio.

La esperanza matemática no tiene por qué ser igual a ninguna de las ganancias o pérdidas particulares, sino que es una suma convenientemente ponderada, siendo las probabilidades los pesos de los pagos. El signo de un determinado pago refleja si es una ganancia o una pérdida. Se debe ser muy cuidadoso a la hora de asegurarnos de que las pérdidas contribuyan con términos negativos al cálculo de X. Finalmente, se dice que un juego o experimento aleatorio es justo o equilibrado si su esperanza global es 0; si un juego no es equilibrado, diremos que es un juego con ventajas.

Advirtamos nuevamente que para que la fórmula de la esperanza matemática sea correcta los sucesos E_1, E_2, E_3, …, E_n tienen que ser disjuntos dos a dos (ningún par puede darse simultáneamente) y constituir un conjunto exhaustivo de posibles resultados del experimento.

Concluiremos con unos ejemplos algo más realistas y con alguna aplicación significativa de la idea de esperanza matemática que, por otra parte, será un tema recurrente a lo largo de este apartado.

Esperanza, ejemplo 3. Se tiene la intención de asistir, a finales de enero, a una convención cuya tasa de matrícula es de \$15 (sin devolución) si se hace la reserva antes del 1 de enero y de \$20 si se paga después de esa fecha o durante la convención. El 28 de diciembre (último día si quiere estar seguro de que la comunicación por correo llegue antes del 1 de enero) estima que la probabilidad de que asista es p. ¿Para qué valores de p debería hacer la reserva con antelación y para cuáles le conviene posponer el pago hasta su llegada a la convención?

Solución.

$$X(\text{reserva anticipada}) = (1) \times (-15)$$

paga \$15, tanto si va, como si no.

$$X(\text{esperar}) = p \times (-20) + (1 - p) \times 0$$

(se paga \$20 solamente si se asiste).

Estas dos esperanzas serán iguales cuando

$$-15 = -20p \quad \text{o} \quad p = \frac{3}{4}.$$

Por lo tanto, si $p > 3/4$ se debería hacer la reserva anticipada; si $p < 3/4$ lo mejor es esperar; para $p = 3/4$ es indistinto, elija lo que quiera.

Esperanza, ejemplo 4. (La inutilidad de la ruleta). Supongamos que se ha hecho una apuesta de \$1 al rojo en una ruleta de Las Vegas. Entonces,

$$X(\text{rojo}) = \frac{18}{38} \times \$(1) + \frac{20}{38} \times \$(-1) = -\frac{2}{38} \times \$(1) = -5{,}26 \text{ centavos}.$$

Igualmente, si la apuesta es a la primera docena (primer tercio), tenemos

$$X(\text{primera docena}) = \frac{12}{38} \times \$(2) + \frac{26}{38} \times \$(-1) = -\frac{2}{38} \times \$(1) = -5{,}26 \text{ centavos.}$$

Si la apuesta es a un solo número, entonces

$$X(1 \text{ número}) = \frac{1}{38} \times \$(35) + \frac{37}{38} \times \$(-1) = -\frac{2}{38} \times \$(1) = -5{,}26 \text{ centavos.}$$

De hecho, se puede demostrar que toda apuesta, o combinación de apuestas, tiene la misma esperanza negativa: se pierde una media de 5,26 centavos por cada dólar apostado, sin importar el modo en el que se apueste.

Esperanza, ejemplo 5. (La apuesta de Pascal). En este ejemplo analizaremos, de manera algo frívola, el razonamiento filosófico que Blaise Pascal elaboró en 1660, mostrando que es probabilísticamente prudente creer en Dios.

Sea p la probabilidad de que Dios exista, y hagamos la razonable suposición de que, cualquiera sea su valor, $p > 0$ (la existencia de Dios no es un suceso imposible). Cada individuo debe tomar una decisión: creer o no creer en Dios. Suponiendo que la creencia en un Dios inexistente tiene una recompensa negativa (digamos $-z$), por la pérdida de tiempo, energía, amistades, etcétera (en la terminología de Pascal, el creyente "no disfrutará de placeres perniciosos, de la fama y el buen vivir"), y que en el caso de que Dios exista, el incrédulo tendrá un pago negativo infinito (la condenación eterna); obtenemos entonces la siguiente matriz de pagos, donde x e y son positivos y finitos.

	Existencia	No existencia
Creer	X	$-z$
No creer	∞	Y

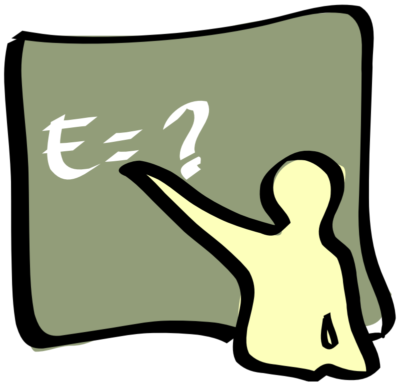
Las esperanzas para el creyente y el no creyente son, respectivamente:

$$X(\text{creer}) = px + (1 - p)(-z)$$

$$X(\text{no creer}) = p(-\infty) + (1 - p)y = -\infty.$$

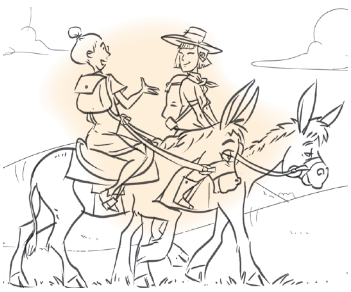
Se concluye que no importa cuán pequeña sea la probabilidad de la existencia de Dios: la esperanza es mayor (quizás menos negativa) si optamos por creer en Dios. Se puede tomar este ejemplo tan seriamente como la religión, la fe en los números y el sentido del humor nos lo permitan.

Un diálogo con los maestros



MATERIAL PARA PRESENTAR EN EL AULA

Irrupción del pensamiento estocástico



El azar se introdujo en el mundo de la matemática a través de consideraciones en torno de los problemas de juegos y las apuestas, a partir del siglo XVII. Desde entonces, el pensamiento determinista y el pensamiento estocástico (que significa aleatorio, azaroso y aparentemente caótico) jugaron sus papeles en el desarrollo de la ciencia, unas veces como rivales, otras como colaboradores.

Se puede decir que, en el siglo XX, el pensamiento aleatorio ha invadido no solo la ciencia, sino incluso la cultura y el modo ordinario de pensar, con la convicción profunda de que, en nuestros días, hemos de aprender a vivir en un mundo que es a la vez determinista y estocástico.

La estadística es el instrumento matemático para manejar lo estocástico. El diseño de experimentos y el análisis estadístico de los datos que de ellos resultan son las herramientas esenciales para los científicos y técnicos de muchas y diversas disciplinas del mundo moderno.

Desde el punto de vista del gran público, la estadística salta a la vista principalmente como la base científica para los muestreos adecuados de la población, a fin de proporcionar a las diversas empresas o a los partidos y al Gobierno el conocimiento necesario para evaluar la viabilidad o conveniencia de posibles futuras decisiones políticas, sociales o económicas.

Cuando hace unos años se produjo el accidente nuclear en la isla de las Tres Millas, en Estados Unidos, se nombró una comisión para que estudiase qué medidas se deberían tomar. Su presidente era John Kemeny, un gran experto en la teoría de la probabilidad.

Una buena proporción de las decisiones que nos afectan a todos está fundamentada en el índice de crecimiento de la producción, en el índice de precios al consumo, en el índice de inflación..., todos ellos siendo parámetros estadísticos indicativos de la marcha de la economía en el orden administrativo, tanto del gobierno como de las empresas.

¿Cómo trabaja?

Estructura de la estadística

La función de la estadística es triple. Ante los enormes colectivos de posible información que las ciencias de la naturaleza o las ciencias humanas le presentan a su consideración, debe ingeniárselas para proceder razonablemente a:

- recoger datos selectivamente;
- organizar, resumir y entender la masa de datos recogida;
- extraer conclusiones de la información así obtenida.

En estos tres pasos, la base matemática utilizada a fin de actuar debidamente es la teoría de la probabilidad, la disciplina que encuentra en la estadística su brazo aplicado.

El primer paso de la estadística, la acumulación adecuada de datos, que no es un problema superficial, constituye la teoría de muestras. A la hora de estudiar la distribución de la riqueza en un país, si recogemos los datos pertinentes parando a los coches que circulan por una autopista que une la gran ciudad con los barrios más ricos y preguntando a sus conductores, es claro que estamos introduciendo un importante sesgo en nuestra muestra.

La organización y el resumen de los datos mediante números representativos, como la media o la varianza –que ya hemos abordado en *Leñitas Geométricas* Nº 1, 5ª época–, constituyen la estadística descriptiva, que forma una base muy importante para el tercer paso.

Por último, se trata de extraer conclusiones de los datos así digeridos, lo que se hace mediante técnicas muy elaboradas matemáticamente, relativas a los campos que se detallan a continuación.

Inferencia estadística. Los colectivos estadísticos suelen ajustarse, muy frecuentemente, a algún tipo particular de serie no muy amplia de distribuciones específicas.

Un problema importante de la inferencia estadística consiste en averiguar, a partir del conocimiento de una muestra de datos tomada de un colectivo, a qué tipo de distribución se ajusta ese colectivo.

Estimación. De un cierto colectivo del que sabemos o suponemos que se ajusta a un tipo específico de distribución hemos extraído una muestra que proporciona datos numéricos $x_1, x_2, \dots, x_n$. La distribución viene caracterizada por un parámetro, un número θ que nos es desconocido.

El problema es ahora, ¿cómo procederemos a partir de $x_1, x_2, x_3, \dots, x_n$ para obtener una aproximación, una estimación fiable, de θ?

Test de hipótesis. Ante un cierto colectivo estadístico, elaboramos una hipótesis acerca del modelo de distribución al que se ajusta.

Recogemos datos del colectivo. ¿Es posible afirmar, razonablemente, que estos datos puedan resultar del colectivo si nuestra hipótesis de que sigue tal modelo es correcta?

Generalmente hay que tomar una decisión basada en los datos y se tolera un cierto grado de incertidumbre.

Estructura de datos. Los datos recogidos no provienen siempre de un mismo proceso. La naturaleza de este proceso es un punto muy esencial a la hora de realizar el análisis subsiguiente.

Diseño de experimentos. En esta fase se estudia, en cada circunstancia particular, cuál deberá ser el modo correcto de proceso de recolección de datos. Por otra parte, en cada observación de un elemento del colectivo se suelen tomar varias medidas y las relaciones entre ellas pueden ser de capital importancia.

Este aspecto se viene estudiando por el análisis de la regresión y correlación. En los procesos que se observan a lo largo del tiempo es muy interesante tratar de descubrir periodicidades y fenómenos semejantes. Tal es el objeto de estudio de las series temporales y su análisis espectral.

Problemas para resolver en grupos

Hacer estadística consiste en recoger datos, ordenarlos, analizarlos y obtener condiciones (interesantes) a partir de la información que se obtiene. ¿Se hace todo ello sin métodos ni guías que conduzcan a que la labor sea lo más fructífera posible? Aprenderemos a interpretar resultados convenientes para sacar conclusiones... y, además, podremos resolver problemas del tenor de los que siguen.

1. En un estanque de una piscifactoría se ha tomado una muestra de 3 000 truchas cuya longitud se ha medido en centímetros, resultando que se distribuyen según la $N(26, 7)$.

   ¿Cómo calcularía $\bar{x} - 3\sigma$, $\bar{x} - 2\sigma$, $\bar{x} - \sigma$, …? ¿Sabría establecer los intervalos de longitud que corresponden y repartir las 3 000 truchas de la muestra en esos intervalos?

2. Se colocan en un estante los seis tomos de cierta obra, de forma aleatoria. ¿Sabría indicar cuál es la probabilidad de que el orden en el que se ponen es el correcto?

3. En una gasolinera han anotado como sigue el número de vehículos que repostan a lo largo de un día:

Horas	0-4	4-8	8-12	12-16	16-20	20-24
Vehículos	5	20	120	132	161	35

   ¿Sabría representar estos datos en un histograma? ¿Y calcular la media y la desviación típica? ¿Qué significado tienen?

4. ¿Sabría calcular el coeficiente de correlación y la ecuación de la recta de regresión que liga las dos variables de esta tabla?

Altura sobre el nivel del mar	0	184	231	481	730	911	1 343	1 550	1 820	2 184
Presión atmosférica	760	745	740	720	700	630	630	650	610	580

¿Puede estimar qué presión atmosférica habrá a 2 000 m y a 3 000 m de altura?

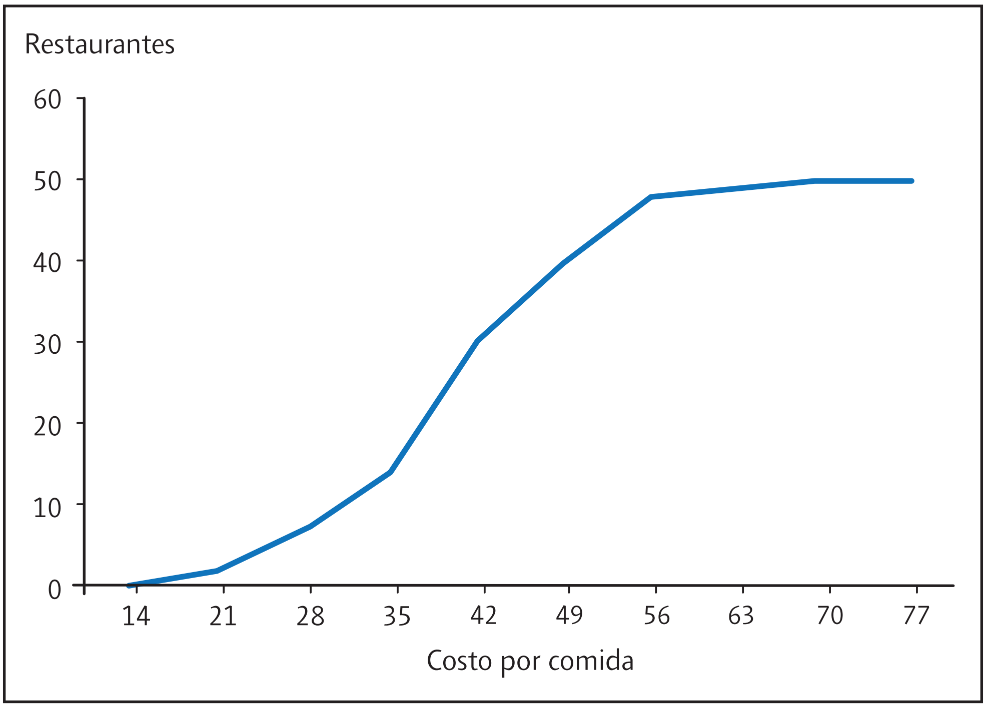
Volvemos a los promedios

Parámetros de dispersión

Hemos discutido varias formas de medir la tendencia central de las distribuciones (véase *Leñitas Geométricas* Nº 1, 5ª época, pp. 15-23), y observado que tales medidas son características de la distribución de una cantidad cualquiera, de modo que las diversas poblaciones se distinguen entre sí por medio de los diferentes valores de estas medidas. Por ejemplo, el valor medio de las estaturas de las mujeres difiere del correspondiente a los hombres.

Las características numéricas de las poblaciones se llaman *parámetros*. Hemos tratado ya con parámetros de tendencia central; lo haremos ahora con los no menos importantes parámetros de dispersión. De acuerdo con el memorándum N° 20 del *Medical Research Council* (El físico de los jóvenes en el Reino Unido), la estatura de los hombres jóvenes con edades de entre 20 y 21 años tiene un valor medio de 5 pies $7\frac{1}{2}$ pulgadas (1,72 m, aproximadamente). Esto es una información que podríamos volcar en un gráfico.

En estadística, una ojiva es un gráfico que muestra la curva de una función de distribución acumulativa dibujada a mano o en *software* de computadora. Los puntos trazados son el límite de la clase superior y la frecuencia acumulativa correspondiente. Se ve, por ejemplo, en la figura siguiente.

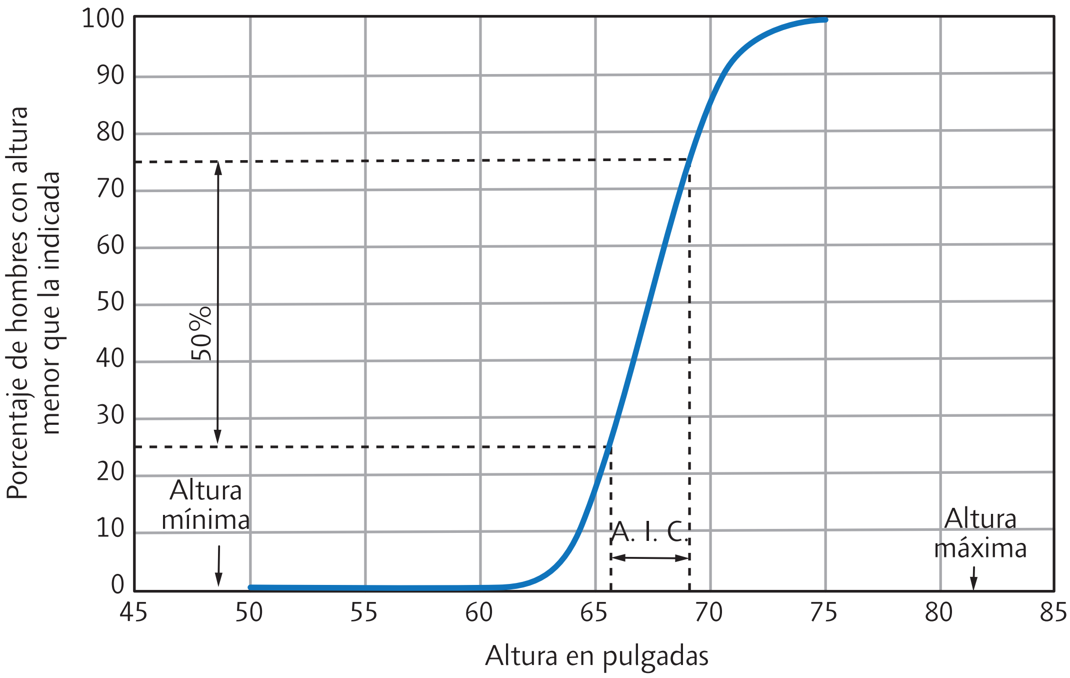


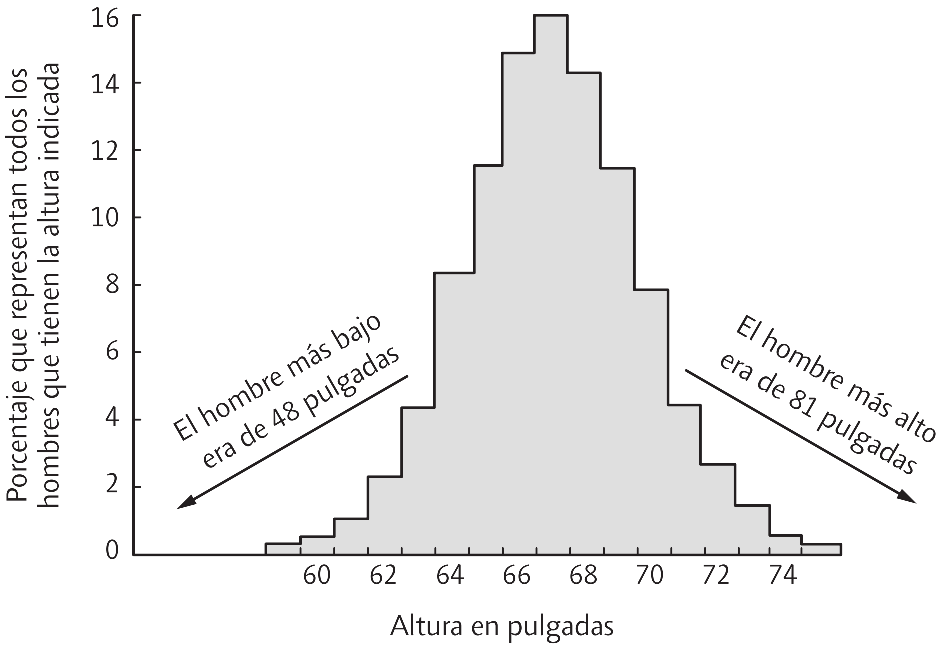
La ojiva para la distribución normal se asemeja a un lado de un arco arabesco u ojival. Asimismo, el término se puede usar para referirse a la función empírica de distribución acumulativa. Este es un tipo de gráfico de frecuencia y también se denomina *polígono de frecuencia acumulada*. Sirve para dar el número (o la proporción) de observaciones más pequeño o igual a un valor particular.

Una ojiva se construye sobre un sistema de ejes perpendiculares. Colocamos en el eje horizontal los límites de los intervalos de clase determinados previamente. En base a cada uno de estos valores límite, determinamos la ordenada de la altura igual a la frecuencia acumulada correspondiente a este valor. Uniendo por segmentos de línea los puntos sucesivos así determinados, obtenemos una línea llamada *ojiva*. Normalmente colocamos la frecuencia acumulada en el eje vertical izquierdo y el porcentaje de frecuencia acumulada en el eje vertical derecho.

Las ojivas son especialmente útiles para estimar los percentiles en una distribución. Por ejemplo, podemos conocer el punto central para que el 50% de las observaciones esté por debajo de este punto y el 50% por encima. Para hacer esto, dibujamos una línea desde el punto del 50% en el eje de porcentaje hasta que se cruza con la curva. Luego proyectamos verticalmente la intersección en el eje horizontal. La última intersección nos da el valor deseado. El polígono de frecuencia y la ojiva se usan para comparar dos conjuntos estadísticos cuyo número podría ser diferente.

Pero si deseamos conocer más sobre la estadística de los jóvenes británicos, ya que es evidente que no todos tendrían esta estatura, la ojiva del primer gráfico estadístico de abajo muestra los porcentajes de hombres de estaturas menores que las indicadas sobre un total de 91 163 que fueron medidos, y la segunda figura muestra los datos dispuestos en forma de histograma.





Es evidente que existe una gran variación, puesto que mientras la gran mayoría de los hombres registrados difieren relativamente poco de la estatura media, no son de ningún modo raras las estaturas muy notables alejadas de esta media. ¿Cómo podremos obtener una medida de la variabilidad respecto del valor medio?

La forma más sencilla es indicar la estatura del hombre más alto y la del más bajo; así, el más alto mide 6 pies, 9 pulgadas; la estatura promedio es de 5 pies, $7\frac{1}{2}$ pulgadas; y el más bajo mide 4 pies, 0 pulgada. Podíamos también haber establecido la amplitud, es decir, la diferencia entre el más alto y el más bajo, a saber: 6 pies, 9 pulgadas menos 4 pies, 0 pulgadas = 2 pies, 9 pulgadas. ¡Este no es un buen método!

Un momento de reflexión nos hará ver claramente que hubiera sido muy fácil no encontrar estas dos estaturas extremas. Podría muy bien haber ocurrido que el más alto de los hombres fuera de 6 pies, 6 pulgadas y el más bajo, de 4 pies, 4 pulgadas. Esto nos daría una amplitud de 6 pies, 6 pulgadas menos 4 pies, 4 pulgadas = 2 pies, 2 pulgadas, resultado bastante diferente del anterior.

Además, podría haber ocurrido que entre los hombres examinados en este grupo para el servicio militar se encontrasen el gigante y el enano de algún circo. Si suponemos que el gigante fuera de 9 pies, 7 pulgadas y el enano, de 3 pies, 2 pulgadas, hubiéramos obtenido una amplitud de 6 pies, 5 pulgadas. Es evidentemente un inconveniente adoptar una medida que dependa totalmente de cualquier rareza que pueda surgir.

Es imposible que una medida basada en una rareza sea considerada representativa de la población ordinaria. La amplitud, entonces, aunque usada en ciertas circunstancias, no es ideal como medida de dispersión. Será mejor disponer de un parámetro menos susceptible de ser influido por los valores extremos.

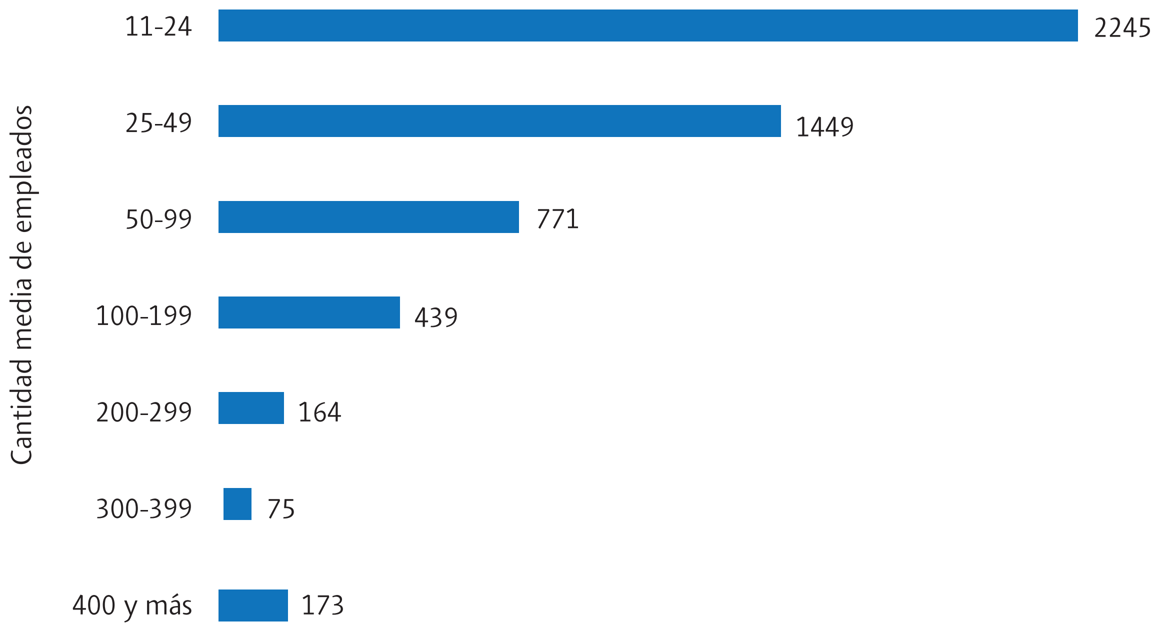
Podemos atacar este problema creando una medida de dispersión dentro de la línea que nos llevó a adoptar la mediana cuando discutimos las medidas de tendencia central. La mediana era el valor por encima del cual se encontraba el 50% de la población y debajo del cual se encontraba el otro 50%. Supongamos ahora dividir la población en cuatro grupos iguales, luego de ordenarla por tamaño.

Llamaremos *cuartilla superior* al valor por sobre el cual se halla solo el 25% de la población, y *cuartilla inferior* a aquel por debajo del cual se encuentra solo el 25% de la misma. Es evidente que el 50% de la población se encuentra comprendido entre los valores de la cuartilla superior y los de la cuartilla inferior. Podemos verificar que las cuartillas superior e inferior de la tabla de estaturas que estamos usando como ejemplo son, aproximadamente, de 5 pies, 9 pulgadas y 5 pies, 6 pulgadas, respectivamente.

Podemos así ver inmediatamente que alrededor del 50% de la población difiere en talla de no más de tres pulgadas, a pesar del hecho de que el hombre más alto observado tenía no menos de 2 pies, 9 pulgadas por sobre el más bajo. Esto es, por supuesto, una consecuencia del modo en que la gran mayoría de las estaturas se agrupan estrechamente alrededor del promedio. Existe un efecto muy común. Los cocientes de inteligencia se comportan del mismo modo. La mayor parte de la gente se aparta poco del promedio de inteligencia, pero los genios, así como los retrasados, tienden a aparecer en magnífico aislamiento. (Recordaremos aquí que el valor modal ["de moda"] tiende a coincidir con la media aritmética cuando la distribución es bastante simétrica).

Luego, la *amplitud intercuartílica*, es decir, la diferencia entre los valores de la cuartilla superior y los de la cuartilla inferior, resulta ser una buena medida de dispersión. Ella es inmune a las perturbaciones ocasionadas por la incidencia de los valores extremos, es fácil de calcular, y su significado es sencillo y claro, ya que nos da el intervalo de variabilidad que alcanza para contener el 50% de la población.

La amplitud intercuartílica se usa con frecuencia en estadísticas económicas y comerciales por otras razones. A menudo, los datos son recogidos de tal modo que existen intervalos indeterminados en uno o en ambos extremos del cuadro.



La figura anterior muestra un ejemplo tomado de cantidad de firmas con un determinado número de empleados en los rubros de alimento, bebidas y tabaco en el Reino Unido (Censo de 1930). El grupo superior está clasificado "400 y más". Esto resulta vago y sería evidentemente imposible realizar un cálculo preciso de cualquier medida que dependiera de procesos aritméticos con inclusión de los valores verdaderos de la ilimitada clase superior (mostraremos más adelante cómo se trata una cierta vaguedad en las otras clases ilimitadas). La mediana y la amplitud intercuartílica nos proporcionan, en tales casos, medidas de tendencia central y de dispersión, respectivamente.

La mediana y las cuartillas son sencillos casos especiales de un esquema completamente general que divide una distribución en *cuartillas*. Así, podemos ordenar nuestra distribución por tamaño y dividirla en diez grupos que contengan igual número de valores. Los valores de la variable en los cuales caen las divisiones se conocen como el primer, segundo, tercer, etc., *deciles*. Esta idea es usada por los psicólogos educacionales para dividir los alumnos en "10% superior, segundo 10%, tercer 10%", etc., todo esto respecto de la inteligencia y la posibilidad de medirla por tests.

Otra medida de dispersión que depende de todas las mediciones es la *desviación media*. Para obtener este parámetro, en primer lugar calculamos la media aritmética de las cantidades de la distribución. Hallamos luego las diferencias entre cada uno de los valores y este promedio, considerándolas a todas como positivas. Sumamos entonces las diferencias así obtenidas y encontramos su promedio dividiendo por el número de ellas. Así, la desviación media es el promedio de las diferencias entre los diversos valores y su media aritmética.

En forma matemática, tenemos:

$$D_m = \frac{\sum |x - \bar{x}|}{n}$$

donde, como ya dijimos en *Leñitas Geométricas* Nº 1, 5ª época, p. 18, el símbolo $\bar{x}$ representa la media aritmética de los diversos valores de x. El signo $|x - \bar{x}|$, como sabemos, indica que debemos encontrar las diferencias entre x y el promedio de los valores de x, sin tener en cuenta el signo. Como también sabemos, el signo Σ significa "súmense todos los términos de la forma".

Ejemplo. Encuéntrese la media aritmética y la desviación media del conjunto de números: 11, 8, 6, 7, 8. Hay aquí $n = 5$ cantidades. Como ya vimos, su promedio es

$$\bar{x} = \frac{\sum x}{n} = \frac{11+8+6+7+8}{5} = \frac{40}{5} = 8.$$

Para obtener la diferencia media, calculamos las desviaciones entre los valores y su promedio 8, y las sumamos como sigue:

$$= \overbrace{3}^{|11-8|} + \overbrace{0}^{|8-8|} + \overbrace{2}^{|6-8|} + \overbrace{1}^{|7-8|} + \overbrace{0}^{|8-8|} = 6.$$

Calculamos entonces la desviación media dividiendo este total de las desviaciones por $n = 5$ y encontramos así $D_5 = 1{,}2$. La desviación media se encuentra frecuentemente en estadísticas económicas.

Las medidas presentadas son muy usadas en el trabajo elemental por ser fáciles de calcular y de comprender. No se usan, sin embargo, en trabajos avanzados, porque son de muy complejo manejo en la teoría de las muestras, de la que dependen tantos trabajos del nivel superior. La medida de dispersión más importante es la *desviación típica*, que es algo más difícil de calcular y cuyo significado es menos evidente a primera vista.

Cálculo e interpretación, sin embargo, pronto se hacen cómodos con un poco de práctica, y entonces la desviación típica se convierte en el más útil de todos los parámetros de dispersión. La desviación típica les es familiar a los ingenieros electricistas y a los matemáticos como *desviación cuadrática media*.

Los pasos del cálculo de la desviación σ típica de un conjunto de valores son los que siguen.

1. Calcúlese la media aritmética del conjunto de valores.
2. Calcúlese las diferencias entre los diversos valores y su media aritmética.
3. Calcúlese los cuadrados de estas diferencias.
4. Calcúlese la suma de los cuadrados de las diferencias para obtener una cantidad conocida como *suma de los cuadrados de la muestra*.
5. Divídase esta "suma de los cuadrados de la muestra" por el número de elementos n del conjunto de valores. Esto da una cantidad conocida como varianza de la muestra.
6. Sáquese la raíz cuadrada de la varianza y obténgase así la desviación típica.

Esto parece mucho más complicado de lo que es en realidad. Resolvamos un ejemplo paso a paso.

Ejemplo. Encuéntrese la desviación típica σ del conjunto de valores 11, 8, 6, 7, 8.

1. Se calculó previamente la media aritmética como $\bar{x} = 8$.
2. Las diferencias entre los valores y este promedio (sin tener en cuenta el signo) son: 3, 0, 2, 1, 0.
3. Los cuadrados de estas diferencias son: $3 \times 3 = 9$; $0 \times 0 = 0$; $2 \times 2 = 4$; $1 \times 1 = 1$; $0 \times 0 = 0$.
4. La suma de los cuadrados de la muestra es: $9 + 0 + 4 + 1 + 0 = 14$.
5. Si dividimos la suma de los cuadrados de la muestra por el número de valores $n = 5$, obtenemos *la varianza* de la muestra como $s^2 = \frac{14}{5} = 2{,}8$ (s^2 es el símbolo aceptado para designar la varianza σ de la muestra).
6. La desviación típica se encuentra sacando la raíz cuadrada de la varianza s^2 de la muestra: $s = \sqrt{2{,}8} = 1{,}7$.

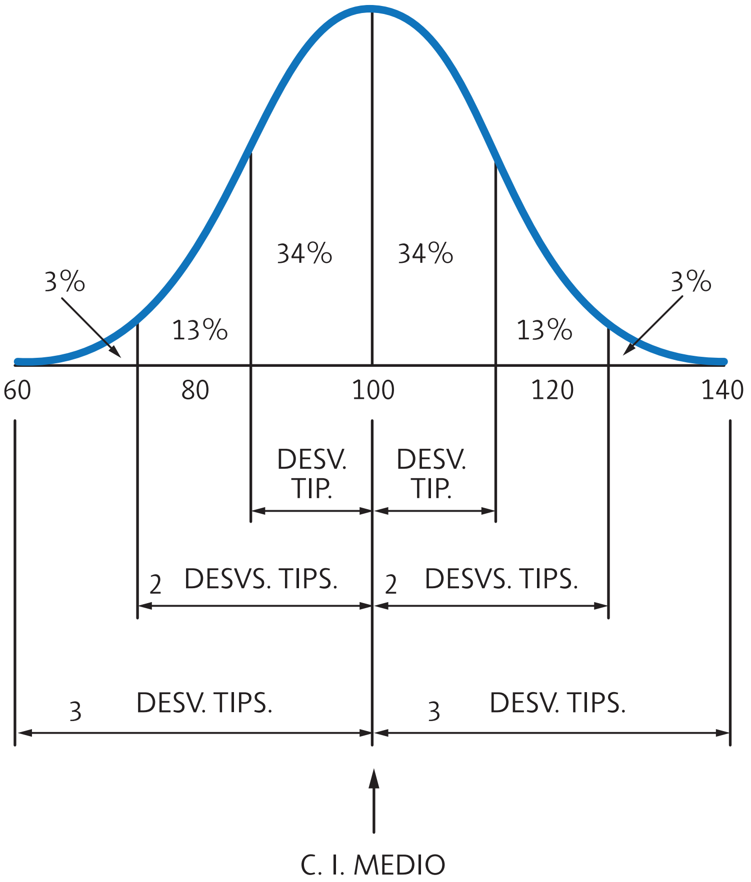
La fórmula de la desviación típica es:

$$s = \frac{\overline{\sum\left(x-\bar{x}\right)^2}}{n}.$$

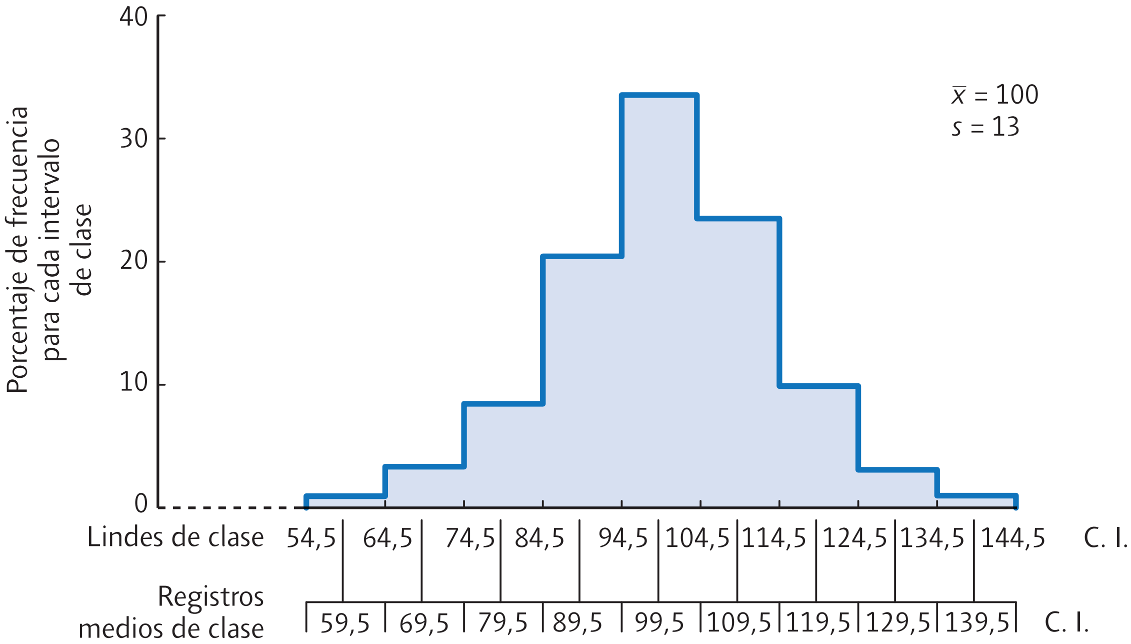
Nos encontraremos más adelante con la cantidad denominada *varianza*. Es una cantidad muy importante, usada en lo que precisamente llamaremos *análisis de la varianza*. Por ahora basta con recordar que la varianza es el cuadrado de la desviación típica. Se calcula exactamente como esta última, omitiendo la extracción de la raíz cuadrada del último paso.

Vimos cómo se calcula la desviación típica. ¿Qué uso se le da en el análisis? Es realmente fácil imaginarlo. Si nos dan una distribución cualquiera que sea razonablemente simétrica respecto de su promedio y unimodal (es decir, que tiene una sola prominencia en el centro, como en el último histograma), entonces encontramos que se comete un error muy pequeño suponiendo que dos tercios de la distribución caen a menos de una desviación típica de la media; que el 95% de la distribución se ubica a menos de dos desviaciones de ella; y que menos del 1% de la distribución cae fuera de tres desviaciones típicas a partir del valor. Esta es una regla aproximada, lógicamente, pero se ha comprobado que funciona muy bien en la práctica.

Supongamos, por ejemplo, que solamente se nos dice que la distribución de la inteligencia, medida con cocientes –el cociente intelectual (CI) de una persona se define como $\frac{\text{Edad mental}}{\text{Edad cronológica}} \times 100$ –, tiene un valor medio: $\bar{x} = 100$, con desviación $s = 13$.



Podríamos entonces trazar la distribución como algo parecido al gráfico de la figura expuesta arriba. No olvidemos comparar el gráfico aproximado así logrado del simple conocimiento de las dos medidas $\bar{x}$ y s con el histograma de la figura de abajo, que se basa en resultados obtenidos por Lewis M. Terman, vertidos en su libro *The Measurement of Intelligence* y citados por J. F. Kenney.



Este uso es típico de las medidas de tendencia central y de dispersión en el trazado esquemático de distribuciones completas (supuestas razonablemente simétricas y unimodales) con los dos valores $\bar{x}$ y s. Tales medidas, puede decirse propiamente, representan la distribución para la cual fueron calculadas.

Las medidas de dispersión tratadas se expresan en las mismas unidades que la variable. A veces ocurre que deseamos saber si una distribución es relativamente más variable que la otra; por ejemplo, encontramos un valor medio de 67 pulgadas para las estaturas de los hombres en las islas británicas, con una desviación típica de 2,5 pulgadas, y para los españoles la estatura media es de 64 pulgadas, con desviación típica de 2,4 pulgadas.

Es evidente que los británicos resultan más altos que los españoles y presentan una estatura ligeramente más variable. ¿Cómo podemos comparar la variabilidad relativa teniendo en cuenta que los españoles son más bajos que los británicos? La medida más usada en la práctica para este caso es el *coeficiente de variación de Pearson*.

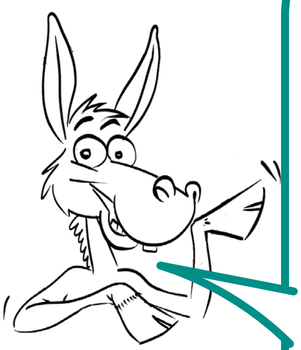
Se lo define como: $v = \dfrac{100s}{\bar{x}}$.

Si calculamos el coeficiente de variación de Pearson para nuestros dos casos obtenemos:

$$\text{Británicos: } v = \frac{100 \times 2{,}5}{67} = 3{,}73\%,$$

$$\text{Españoles: } v = \frac{100 \times 2{,}4}{64} = 3{,}75\%.$$

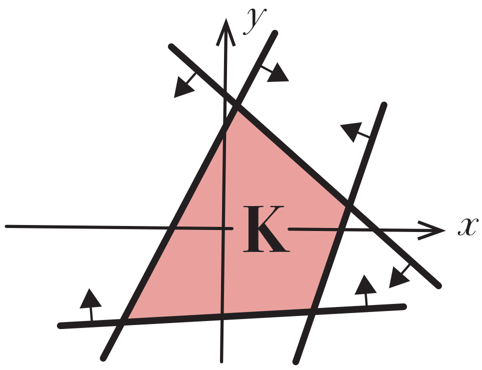
Concluimos que, aunque los británicos son más variables en un sentido absoluto, la variabilidad de los españoles, expresada como porcentaje de la estatura media, es levemente mayor. Más adelante, cuando consideremos la "forma típica" de la llamada *curva normal*, nos concentraremos en otro excelente método de comparación de distribuciones, no enfocado en su variabilidad relativa, sino en su asimetría y otras características.



Disponemos ahora de todas las ideas básicas necesarias respecto de medidas de tendencia central y de dispersión. Tenemos una buena selección de medidas a nuestro alcance, cada una de ellas apropiada para algún tipo común de problema de representación numérica. Este conocimiento, junto con lo que ya sabemos acerca de los métodos de representación gráfica, es suficiente para conducirnos a través de la mayor parte de aquello que se conoce como estadística descriptiva, es decir, la clase de estadística con la que el aficionado se encontrará más a menudo. Con el material precedente razonablemente bien entendido, seremos capaces de seguir crítica e inteligentemente la mayor parte del material descriptivo. Existe una masa inmensa de este tipo de estadísticas y muchas de ellas con relevante información para los interesados en leerlas.

Sin embargo, las ideas hasta aquí presentadas a consideración no son suficientes para que uno mismo las utilice en el manejo de una masa de datos no elaborados. Existe una gigantesca cantidad de aritmética procedente de muchos casos analizados, y si tratáramos de hacer los cálculos necesarios de acuerdo con las definiciones elementales y los procedimientos explicados, el más sencillo agrupamiento de datos resultaría una carga insoportable. Pero el estadístico ve facilitada su tarea por el uso de métodos abreviados especiales de cálculo como los que estudiaremos en entregas posteriores.

No olvidemos lo fundamental



Continúa de *Leñitas Geométricas* N° 1, 5ª época, p. 23. Los sistemas de desigualdades lineales.

2. Representación geométrica de desigualdades de primer grado con dos o tres incógnitas.

Examinemos la ecuación de primer grado con dos incógnitas x e y:

$$ax + by + c = 0. \qquad (1)$$

Considerando a x e y como coordenadas de un punto en un plano, es lógico que surja la pregunta: ¿qué conjunto forman en este plano los puntos cuyas coordenadas satisfacen la ecuación (1)? O, más breve: ¿qué conjunto de puntos determina la ecuación (1)?

Daremos respuesta, aunque quizás sea conocida para algunos: la multitud de puntos determinados por la ecuación (1) es una línea recta en un plano. En efecto, si $b \neq 0$, entonces la ecuación (1) se reduce a la forma

$$y = kx + p,$$

y, como es sabido, esta ecuación determina una recta. Si, por el contrario, $b = 0$, entonces la ecuación se reduce a la forma

$$x = h$$

y determina una recta paralela al eje de las coordenadas.

Semejante pregunta surge de igual modo con relación a la desigualdad

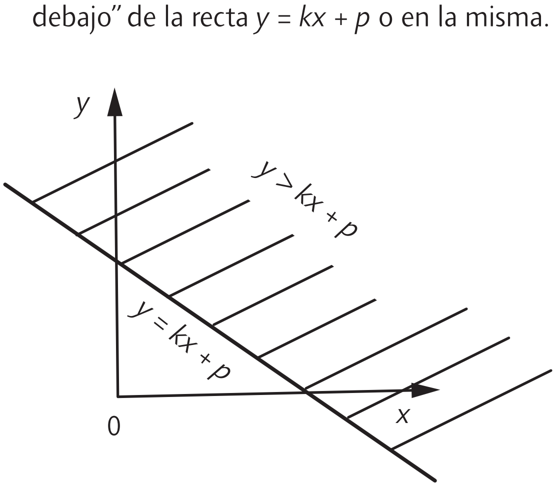
$$ax + by + c \geq 0. \qquad (2)$$

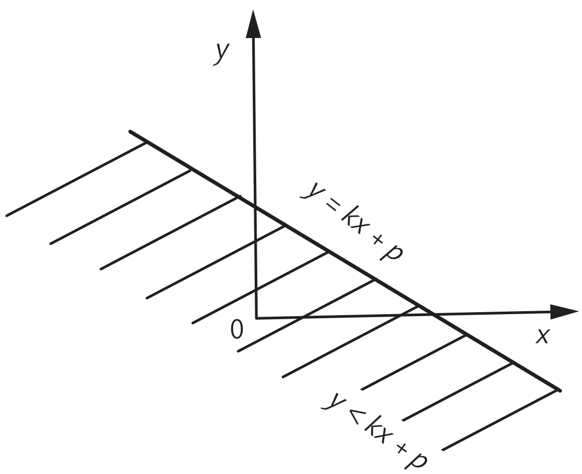
¿Qué conjunto de puntos determina en un plano la desigualdad (2)?

También aquí la respuesta es simple. Si $b \neq 0$, entonces la desigualdad dada se reduce a una de las dos formas

$$y \geq kx + p \quad \text{o} \quad y \leq kx + p.$$

No es difícil observar que a la primera de estas dos desigualdades la satisfacen todos los puntos situados "por encima" de la recta $y = kx + p$ o en ella, mientras que a la segunda, todos los puntos situados "por debajo" de la recta $y = kx + p$ o en la misma.

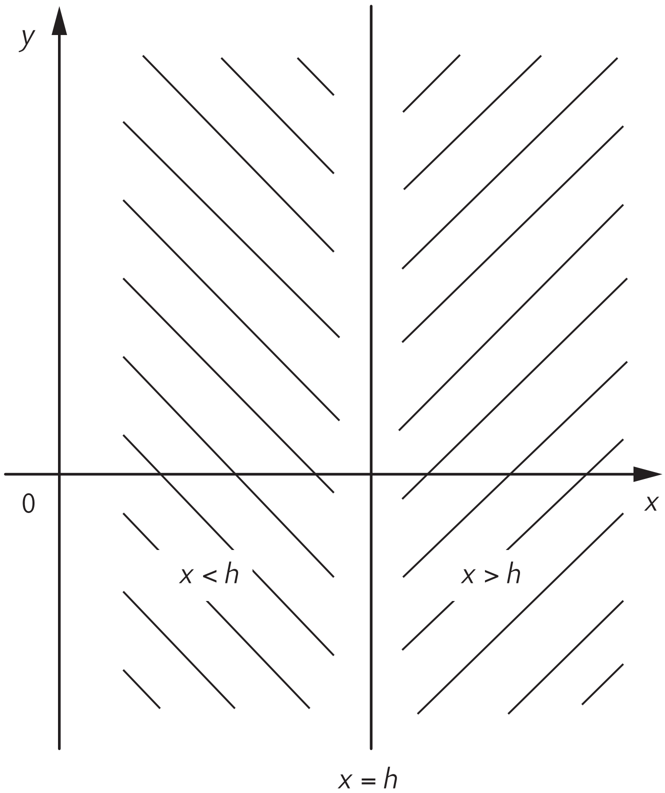




Siendo $b = 0$ la desigualdad (2) se reduce a una de las formas

$$x \geq h \qquad \text{o} \qquad x \leq h;$$

a la primera de ellas la satisfacen todos los puntos situados "a la derecha" de la recta $x = h$ o en esta; a la segunda, todos los puntos situados "a la izquierda" de la recta $x = h$ o en dicha recta.



Así, pues, la ecuación (1) determina en un plano de coordenadas una línea recta y la desigualdad (2), uno de los dos semiplanos en que esta recta corta el plano (consideramos que esta recta pertenece a cualquiera de los dos semiplanos determinados por ella misma).

Ahora resolveremos cuestiones análogas con relación a la ecuación

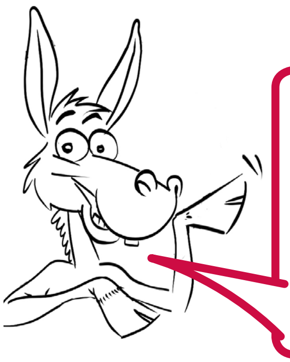
$$ax + by + cz + d = 0, \qquad (3)$$

y a la desigualdad

$$ax + by + cz + d \geq 0. \qquad (4)$$

Claro está que para ello x, y, z deberán considerarse como coordenadas de un punto en el espacio.

El resultado, como no es difícil prever, será el siguiente.



Teorema. La ecuación (3) determina cierto plano en el espacio, y la desigualdad (4), uno de los dos semiespacios en los que dicho plano corta a todo el espacio (el mismo plano se considera perteneciente a cualquiera de los dos semiespacios determinados por él).

Demostración. De tres números a, b y c, por lo menos uno es diferente de cero; sea, por ejemplo, $c \neq 0$. Entonces, la ecuación (3) se reduce a la forma

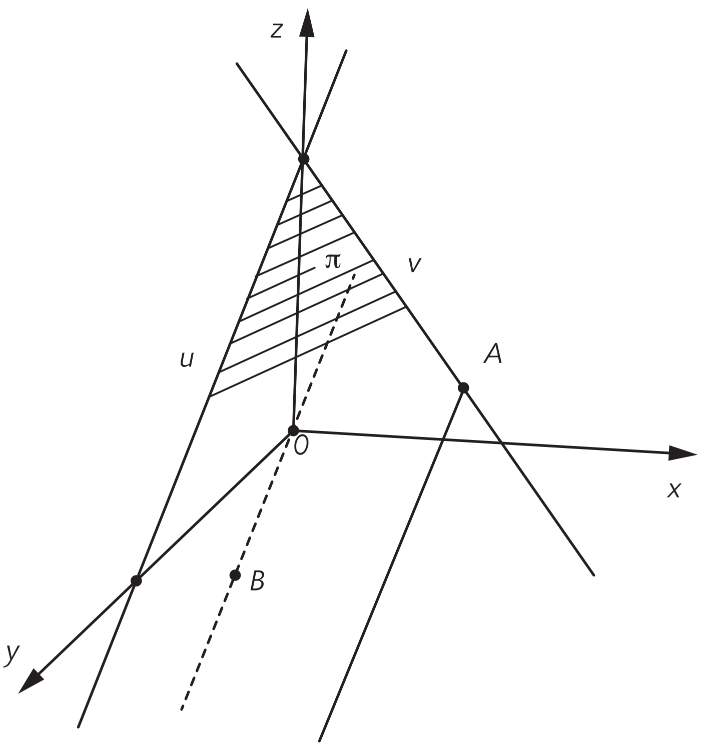
$$z = kx + ly + p. \qquad (5)$$

Designemos por $\mathcal{L}$ el conjunto de puntos $M(x, y, z)$ que verifican la ecuación (5). Nuestro objetivo es demostrar que $\mathcal{L}$ es un plano.

Veamos qué puntos de $\mathcal{L}$ pertenecen al plano de coordenadas yOz. Para ello consideramos que en la ecuación (5) $x = 0$. Tenemos entonces

$$z = ly + p. \qquad (6)$$

Por lo tanto, la intersección de $\mathcal{L}$ con el plano yOz es la recta u, determinada en este plano por la ecuación (6) (figura de abajo).



De la misma forma hallamos que la intersección de $\mathcal{L}$ con el plano xOz es la recta v determinada en este plano por la ecuación

$$z = kx + p. \qquad (7)$$

Las dos rectas u y v pasan por el punto $P(0, 0, p)$.

Designemos por π el plano que contiene las rectas u y v. Demostremos que π pertenece al conjunto $\mathcal{L}$. Para ello será suficiente establecer el siguiente hecho: la recta que pase por cualquier punto $A \in v$ paralela a u pertenece a $\mathcal{L}$.

En primer lugar, hallaremos cualquier punto B, tal que $\overline{OB} \parallel u.$ En el plano yOz la ecuación $z = ly + p$ determina la recta u; por consiguiente, la ecuación $z = ly$ determina una recta paralela a u, que pasa por el origen de las coordenadas (en la figura de arriba se da en línea punteada). En calidad de B se puede tomar un punto con coordenadas $y = 1$, $z = 1$, yaciente en esta recta.

El punto arbitrario $A \in v$ tiene las coordenadas x, 0, $kx + p$. El punto B elegido por nosotros tiene las coordenadas 0, 1, l.

La recta que pasa por A paralelamente a u está compuesta por los puntos

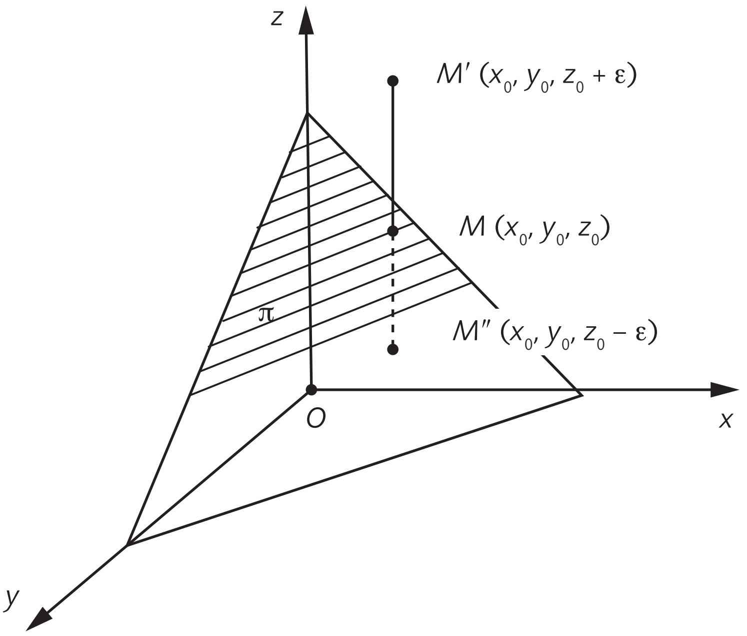
$$A + sB = (x, 0, kx + p) + s(0, 1, l) = (x, s, kx + p + sl),$$

siendo s un número arbitrario (ver la afirmación 3 de la página 29, en *Leñitas Geométricas* N° 1, 5ª época). Es fácil comprobar que las coordenadas del punto $A + sB$ verifican la ecuación (5), es decir, $A + sB \in \mathcal{L}$.

Con ello queda demostrado que el plano π pertenece totalmente al conjunto $\mathcal{L}$.

Nos queda por dar el último paso, que consiste en demostrar que $\mathcal{L}$ coincide con π, es decir que el conjunto $\mathcal{L}$ no contiene ningún punto fuera de π.

Para demostrarlo examinemos tres puntos: el punto $M(x_0, y_0, z_0)$, situado en el plano π; el punto $M'(x_0, y_0, z_0 + \varepsilon)$ situado "sobre" el plano π ($\varepsilon > 0$) y el punto $M''(x_0, y_0, z_0 - \varepsilon)$, situado "bajo" el plano π (figura de abajo).



Como $M \in \pi$, entonces $z_0 = kx_0 + ly_0 + p$ y, por consiguiente,

$$z_0 + \varepsilon > kx_0 + ly_0 + p,$$

$$z_0 - \varepsilon < kx_0 + ly_0 + p.$$

De esto se deduce que las coordenadas del punto M' satisfacen la desigualdad estricta

$$z > kx + ly + p,$$

y las coordenadas del punto M'', la desigualdad estricta

$$z < kx + ly + p.$$

Por lo tanto, M' y M'' no pertenecen a $\mathcal{L}$. Esto demuestra que $\mathcal{L}$ coincide con el plano π. Además, de nuestros razonamientos se deduce que el conjunto de todos los puntos que verifican la desigualdad

$$ax + by + cz + d \geq 0$$

forma uno de los dos semiespacios ("superior" o "inferior") en que el plano π divide todo el espacio.

2. Representación geométrica de un sistema de desigualdades lineales con dos o tres incógnitas

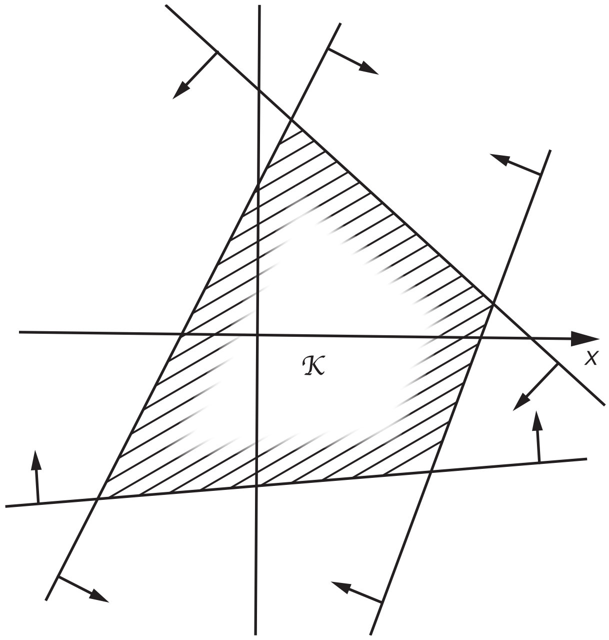
Sea dado el sistema de desigualdades

$$\left.\begin{array}{r} a_1x + b_1y + c_1 \geq 0, \\ a_2x + b_2y + c_2 \geq 0, \\ \dots \\ a_mx + b_my + c_m \geq 0, \end{array}\right\} \quad (1)$$

con dos incógnitas x e y.

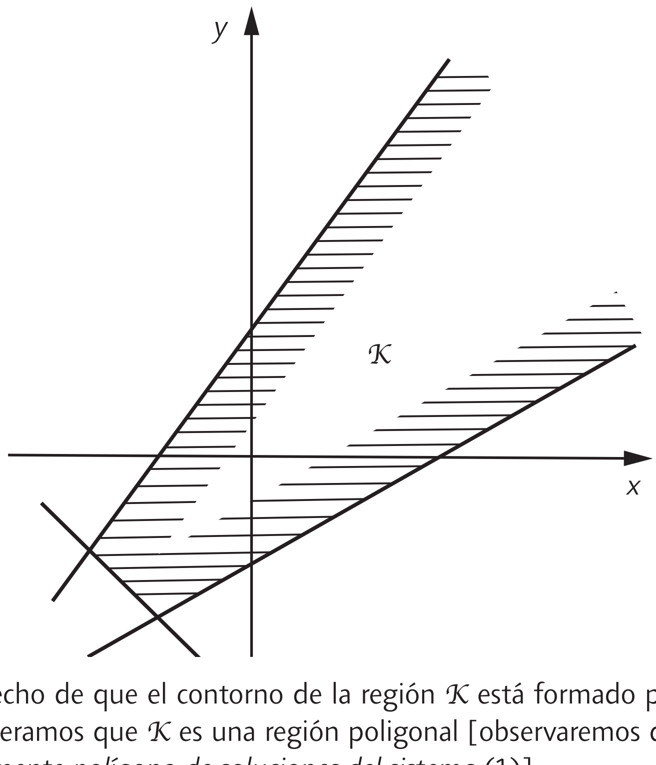
La primera desigualdad de este sistema determina en un plano de coordenadas xOy cierto semiplano Π_1; la segunda, el semiplano Π_2; y así sucesivamente. Si cualquier par de números x, y satisface todas las desigualdades (1), entonces el punto $M(x, y)$ pertenece a todos los semiplanos $\Pi_1, \Pi_2, \dots, \Pi_m$ al mismo tiempo.

En otros términos, el punto M pertenece a la intersección de dichos semiplanos. Es fácil ver que la intersección de una cantidad finita de semiplanos representa una región poligonal $\mathcal{K}$. En la figura de abajo se muestra una de estas posibles regiones. A lo largo del contorno de esta región se dan rayas dirigidas hacia el interior de la región.



Estas rayas indican a qué lado de cada recta yacen en el respectivo semiplano; y lo mismo indican las flechas. La región o recinto $\mathcal{K}$ se llama *región de soluciones del sistema* (1).

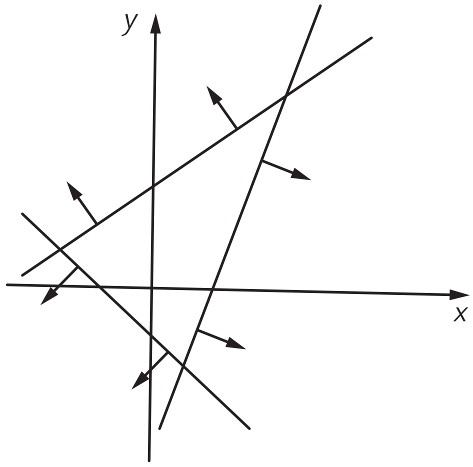
A su vez, indicamos que la región de soluciones no siempre es limitada: como resultado de la intersección de varios semiplanos puede surgir también una región ilimitada como, por ejemplo, la región dada en la figura siguiente.



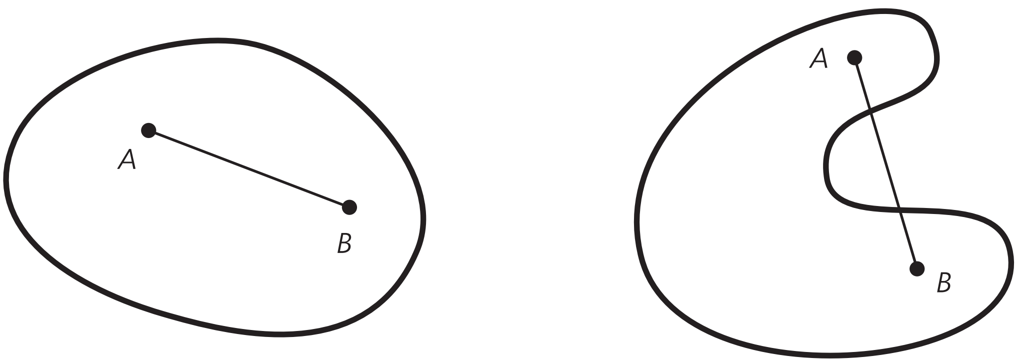
Teniendo en cuenta el hecho de que el contorno de la región $\mathcal{K}$ está formado por trozos de rectas (o por rectas completas), consideramos que $\mathcal{K}$ es una región poligonal [observaremos que cuando la región $\mathcal{K}$ es ilimitada se llama simplemente *polígono de soluciones del sistema* (1)].

Aquí debemos hacer una advertencia para evitar malentendidos. En el curso escolar de geometría, por "polígono" se entiende una línea cerrada compuesta por segmentos de rectas. Mientras tanto, en la literatura sobre desigualdades lineales este término significa, no la propia línea sino el conjunto de todos los puntos del plano abarcados por ella (es decir, yacientes dentro o en la misma línea). Más adelante el término "polígono" se interpreta precisamente en este último sentido.

Asimismo, también es posible el caso en que no hay ni un solo punto común a todos los semiespacios considerados, es decir, cuando la región $\mathcal{K}$ está "vacía". Esto significa que el sistema (1) es contradictorio. Este caso viene dado en la figura siguiente.



La región $\mathcal{K}$ de soluciones es siempre convexa. Recordemos que, según la definición general, el conjunto de puntos (en un plano o espacio) se llama *convexo* si dados dos puntos A y B cualesquiera del conjunto, este contiene todo el segmento $\overline{AB}$ que los une. En las siguientes figuras se muestra la diferencia entre conjuntos convexos y no convexos.

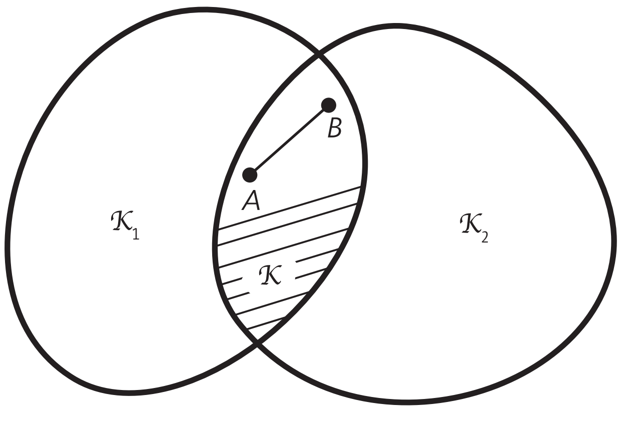


La convexidad de la región $\mathcal{K}$ de soluciones se deduce del propio procedimiento de formación de esta región pues esta última se ha obtenido mediante la intersección de varios semiplanos, cada uno de los cuales es un conjunto convexo.

Además, para que no quede ninguna duda con relación a la convexidad de $\mathcal{K}$, demostraremos el siguienle lema.

Lema. La intersección de cualquier número de conjuntos convexos es un conjunto convexo.

Demostración. Sean $\mathcal{K}_1$ y $\mathcal{K}_2$ dos conjuntos convexos y $\mathcal{K}$, su intersección. Examinemos dos puntos cualesquiera A y B pertenecientes a $\mathcal{K}$ (figura siguiente):



Como $A \in \mathcal{K}_1$ y $B \in \mathcal{K}_2$ y el conjunto $\mathcal{K}_1$ es convexo, entonces el segmento $\overline{AB}$ pertenece a $\mathcal{K}_1$. De forma análoga, el segmento $\overline{AB}$ pertenece a $\mathcal{K}_2$. O sea, el segmento $\overline{AB}$ pertenece al mismo tiempo a los dos conjuntos $\mathcal{K}_1$ y $\mathcal{K}_2$ y, por consiguiente, también a la intersección $\mathcal{K}$ de ambos.

Con esto queda demostrado que $\mathcal{K}$ es un conjunto convexo. Una reflexión similar nos demuestra que la intersección de cualquier número (no obligatoriamente de dos) de conjuntos convexos es un conjunto convexo.

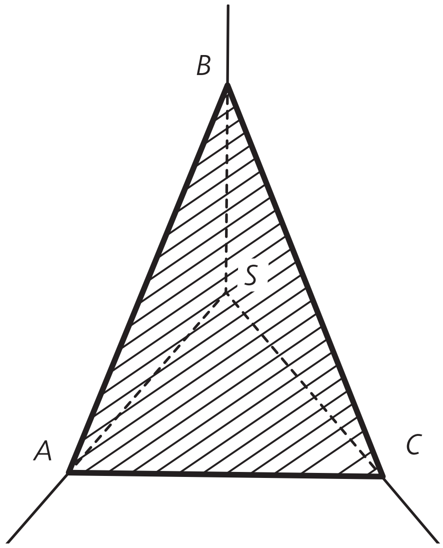
Así, pues, el lugar geométrico de los puntos cuyas coordenadas satisfacen todas las desigualdades (1), o la región de soluciones del sistema (1), es la región convexa y poligonal $\mathcal{K}$. Esta región surge como resultado de la intersección de todos los semiplanos que corresponden a las desigualdades del sistema dado.

Examinemos el caso de tres incógnitas. Ahora tenemos dado el sistema

$$\left.\begin{array}{c} a_1x+b_1y+c_1z+d_1\geq 0,\\ a_2x+b_2y+c_2z+d_2\geq 0,\\ \cdots\\ a_mx+b_my+c_mz+d_m\geq 0,\end{array}\right\} \qquad (2)$$

Como es sabido por lo dicho en la *Leñitas Geométricas* 1, 5ª época, cada una de estas desigualdades determina cierto semiespacio.

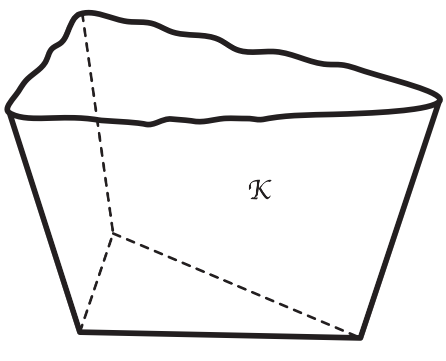
Por lo tanto, la región determinada por este sistema representará la intersección (parte común) de m semiespacios. Pero la intersección de una cantidad finita de semiespacios es cierta región convexa y poliédrica $\mathcal{K}$.



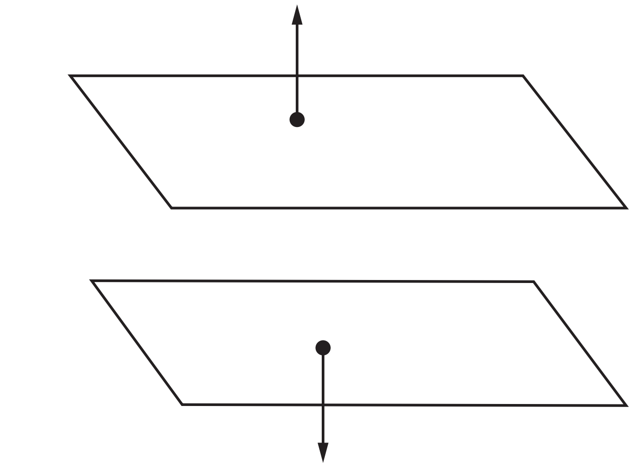
En la figura de arriba se da un ejemplo de tal región, siendo $m = 4$. En este ejemplo, la región $\mathcal{K}$ representa un simple tetraedro (más exactamente, $\mathcal{K}$ está compuesta por todos los puntos que yacen dentro y en las superficies del tetraedro). Y, en general, no es difícil comprobar que cualquier poliedro convexo puede obtenerse como consecuencia de la intersección de una cantidad finita de semiespacios.

Este punto necesita una aclaración semejante a la dada anteriormente. La cuestión es que, en el curso escolar de geometría, por "poliedro" se entiende una superficie cerrada, compuesta por caras planas. Nosotros vamos a depositar en este término un contenido más amplio, denominando como *poliedro*, no a la superficie propia, sino al conjunto de todos los puntos del espacio abarcados por este (naturalmente, este conjunto incluye también la misma superficie, pero solamente como parte).

Claro que el caso también se presenta cuando la región $\mathcal{K}$ no es limitada (se extiende infinitamente); un ejemplo de tal región se da en la figura siguiente.



Por último, puede suceder que no existan, en general, puntos que satisfagan a todas las desigualdades consideradas [si el sistema (2) fuera contradictorio]; entonces, la región $\mathcal{K}$ estaría vacía. Un caso semejante está representado en la figura siguiente.



Especialmente, debemos detenernos en el caso en el que, entre las desigualdades del sistema (2) hay dos:

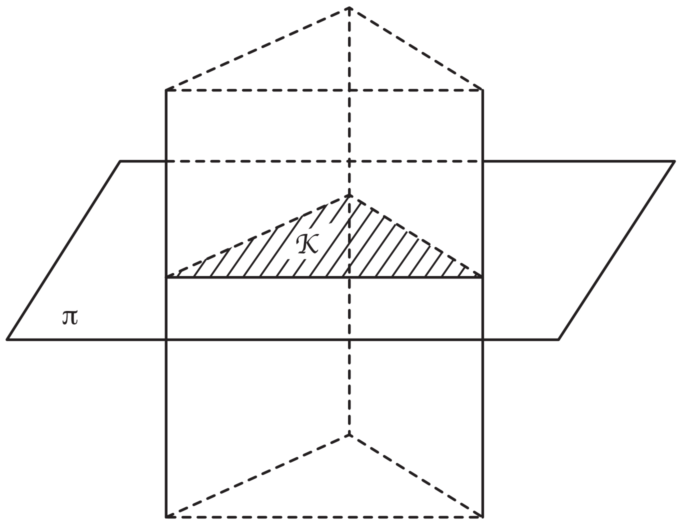
$$ax + by + cz + d \geq 0,$$

$$-ax - by - cz - d \geq 0,$$

tales que cambian para ser sustituidas por una sola ecuación:

$$ax + by + cz + d = 0.$$

Esta última determina cierto plano π en el espacio. Las otras desigualdades del sistema (2) determinan en el plano cierta región convexa y poligonal $\mathcal{K}$, la cual sirve como región de soluciones del sistema (2). Vemos, que un caso particular de región convexa y poliédrica en el espacio puede servir una región convexa y poligonal en un plano.



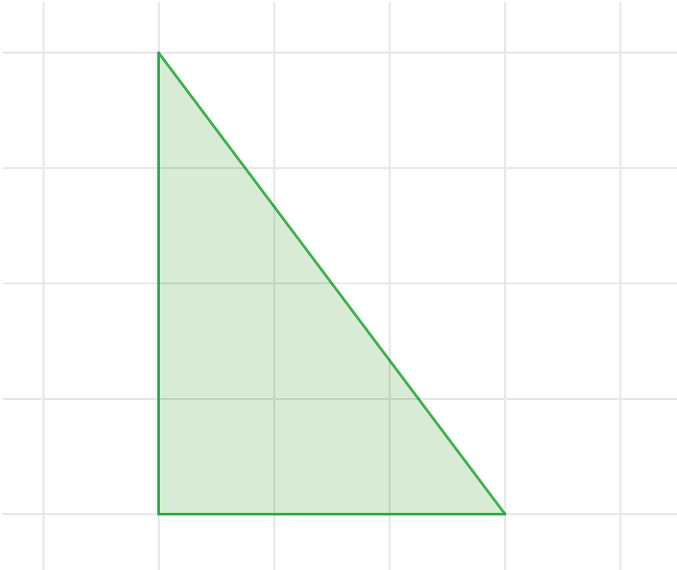
En la figura anterior, la región $\mathcal{K}$ es un triángulo, constituido por la intersección de cinco semiespacios: dos de ellos están limitados por un plano "horizontal" π y los otros tres se forman en la intersección con un prisma triedro "vertical".

Por ser un caso análogo al de las desigualdades con dos incógnitas, denominaremos la región $\mathcal{K}$ como región de soluciones del sistema (2). Subrayamos una vez más la circunstancia de que la región $\mathcal{K}$ por ser la intersección de un cierto número de semiespacios, es forzosamente convexa.

O sea, el sistema (2) determina en el espacio una región $\mathcal{K}$ convexa y poliédrica. Esta última es el resultado de la intersección de todos los semiespacios que satisfacen las desigualdades del sistema dado.

Si $\mathcal{K}$ es una región limitada, entonces se denomina simplemente *poliedro de soluciones del sistema* (2).

Marcar el incentro del triángulo con vértices en la cuadrícula, usando solo un lápiz.

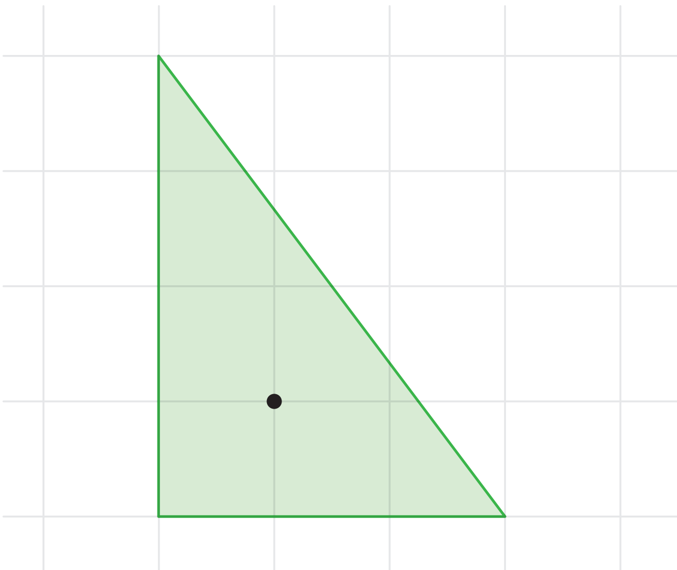


Nota: El incentro de un triángulo es el punto donde concurren las bisectrices de sus ángulos y es el centro de la circunferencia inscripta en el triángulo.

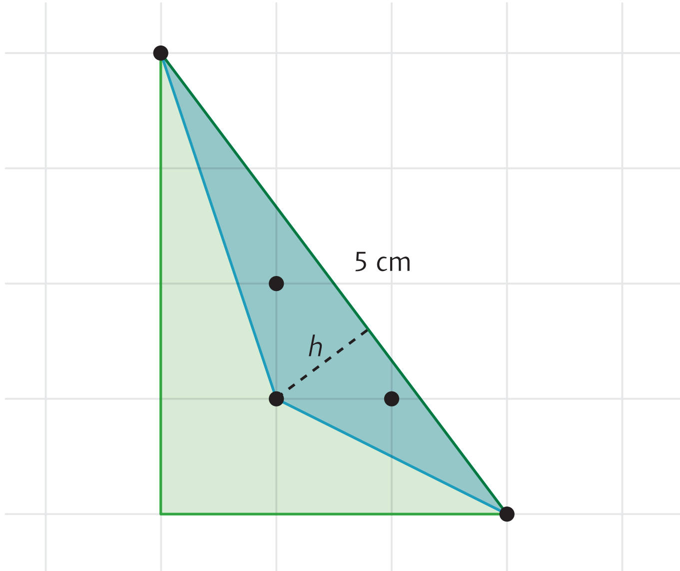
Solución

Podemos pensar que los cuadrados de la cuadrícula son de 1 cm por 1 cm. Por ser el triángulo rectángulo, sus lados miden 3 cm, 4 cm y 5 cm.

El incentro, por estar en las bisectrices, es el punto que equidista de los lados del triángulo. El punto que se indica en la figura a continuación equidista de los catetos del triángulo, estando a 1 cm de cada cateto.



La distancia desde este punto a la hipotenusa es la altura h indicada del siguiente triángulo sombreado:



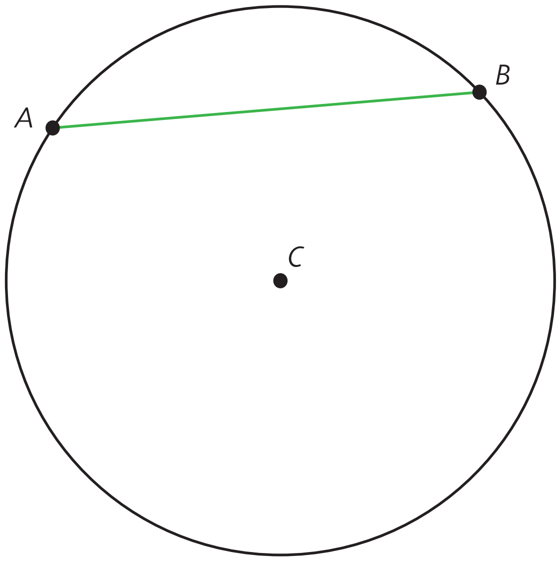
Por la fórmula de Pick, el área de este triángulo es $2+\frac{3}{2}-1=\frac{5}{2}$, pero este valor debe coincidir con $\frac{5\times h}{2}=\frac{5}{2}\times h$, en consecuencia, $h = 1$ y el punto indicado al comienzo es el incentro.

Nota: Este problema también puede resolverse usando el Teorema de Poncelet que establece lo siguiente:

Si r es el radio de la circunferencia inscripta en el triángulo rectángulo de catetos a y b y de hipotenusa c, entonces se cumple la igualdad:

$$a\ \ + b = b + 2r.$$

Usando solo una escuadra y un lápiz, trazar la mediatriz de la cuerda *AB* de la circunferencia con centro *C*, según muestra la figura.



Solución

La mediatriz es el lugar geométrico formado por los puntos del plano que equidistan de *A* y *B*, de modo que *C* está en la mediatriz. Además, la mediatriz es la recta perpendicular al segmento *AB* que pasa por el punto medio de *AB*.

Como sabemos que pasa por *C* y es perpendicular a *AB*, podemos trazar la mediatriz usando una escuadra.