Leñitas Geométricas*

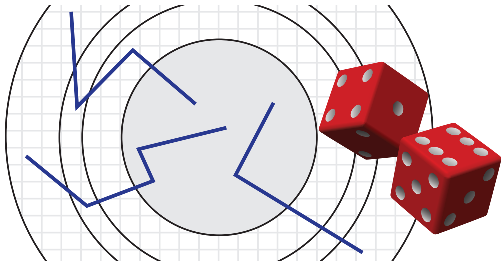
para el Fogón Matemático de los Festivales

De OMA para Profesores y Maestros en actividad

5ª época ✈ Nº 15
5 de octubre 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



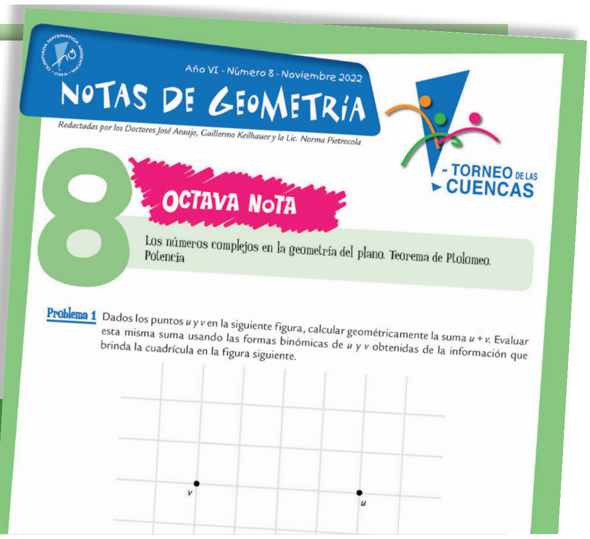
Continúa de *Leñitas Geométricas* Nº 14, 5ª época, Estadística complementaria al método Montecarlo

Nociones de metodología estadística

Con una cantidad limitada de observaciones (muestra) el estadístico se propone sacar conclusiones. El mecanismo general de esta operación es el siguiente: se formula, respecto de la población total, una hipótesis que es el resultado de consideraciones teóricas o que está sugerida por los datos mismos. ¿En qué medida esta hipótesis está fundada o, más exactamente, en qué medida no es invalidada por la muestra observada?

Es evidente que la certeza absoluta no puede alcanzarse nunca pues de lo contrario ello nos obligaría a conocer todos los casos posibles, lo cual es inalcanzable, cualquiera sea la importancia de la muestra. Pero el grado de certeza de una conclusión puede formularse en términos probabilísticos: en la práctica es legítimo considerar imposible una eventualidad que no tiene sino muy pocas probabilidades de producirse o, recíprocamente, tenerla por verosímil si su probabilidad alcanza un valor suficiente.

Es útil fijar de antemano el grado de certeza, es decir, la garantía que se desea obtener. Por ejemplo, se considerará poco verosímil toda hipótesis que solamente verifique los datos observados 5 veces de cada 100. Si la probabilidad sobrepasa el 5 % se podrá admitir la hipótesis, por lo menos provisionalmente, y bajo reserva hasta que se realicen comprobaciones más extensas. La probabilidad del 5 % se toma a menudo como la línea

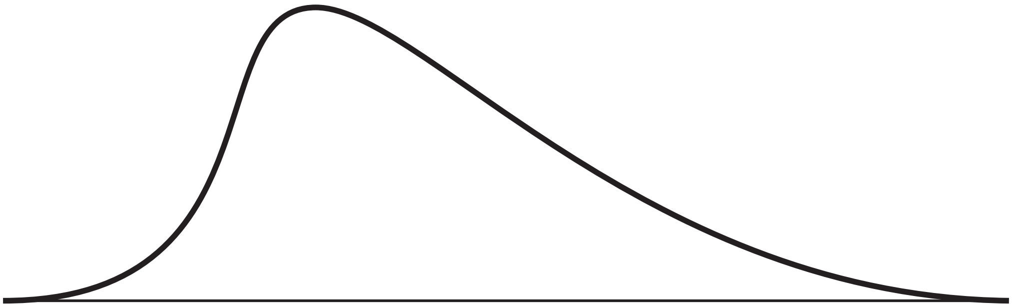


* Recordamos a los lectores que los temas editados en *Leñitas Geométricas* son material preparado y en gran parte desarrollado y sugerido por el doctor Miguel de Guzmán.

divisoria entre el grupo de las hipótesis admisibles y el de las inadmisibles; se la llama *nivel de significación* al 5 %. Naturalmente, es lícito mostrarse más o menos exigente, adoptando niveles tal vez más o menos elevados.

La interpretación estadística supone dos fases esenciales: una inductiva, que permite el pasaje de la muestra a la población total, y la otra deductiva, que sitúa a la muestra señalada con una tasa de probabilidad en el grupo de todas las eventualidades que pueden intervenir por extracción al azar en el seno de la población total. Si hay varias hipótesis admisibles, se elegirá aquella que conduzca, para la muestra, a la probabilidad máxima. La primera fase del trabajo supone una cierta intuición; la segunda, que se realiza con las normas matemáticas del cálculo de probabilidades, constituye la crítica de la primera.

Por ejemplo, mediante los datos observados es siempre posible calcular con las fórmulas dadas anteriormente para analizar los principios de interpretación, la media y la desviación típica de una población gaussiana completamente definida por estas características. Pero si la distribución experimental presenta la forma de la figura que sigue, la asimilación a la ley de Gauss es, por cierto, errónea.



Es infinitamente poco probable que se pueda extraer de una distribución total simétrica, no haciendo intervenir sino el azar, una muestra distribuida de manera tan asimétrica, por lo que es necesario pensar en orientarse hacia otras hipótesis. Sin duda, la evidencia no siempre es tan notable y conviene poseer métodos exactos –que constituyen las pruebas estadísticas–, para verificar la legitimidad de las hipótesis.

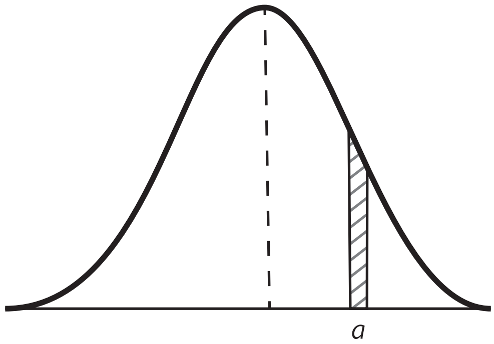
Es importante precisar con exactitud la significación del nivel que permite establecer la separación entre las distintas categorías de hipótesis.

Imaginemos una serie de 10 000 jugadas de "cara y ceca". La eventualidad más probable es la de ver aparecer 5 000 veces cara y 5 000 veces ceca; es ligeramente más probable que 4 999 caras y 5 001 cecas, y mucho más aún que 2 000 caras y 8 000 cecas. Hay un total de 10 001 eventualidades posibles y se sabe que sus probabilidades se distribuyen siguiendo la ley de Gauss; todas –aun las más extraordinarias, como obtener cara cada vez– acabarían por presentarse si se pudiera repetir infinidad de veces –digamos millones y millones de veces– la serie de 10 000 jugadas.

En una serie aislada, una eventualidad cualquiera fijada de antemano –incluso aquella que, al realizar la igualdad de caras y cecas, es la menos probable– tiene una probabilidad propia extremadamente pequeña y muy inferior al 5 %; sin embargo, en vista de un resultado tal como 4 990 caras y 5 010 cecas, nadie pretenderá argumentar que la moneda es falsa, o, dicho en lenguaje estadístico, que la hipótesis de una probabilidad idéntica para cara y para ceca debe rechazarse.

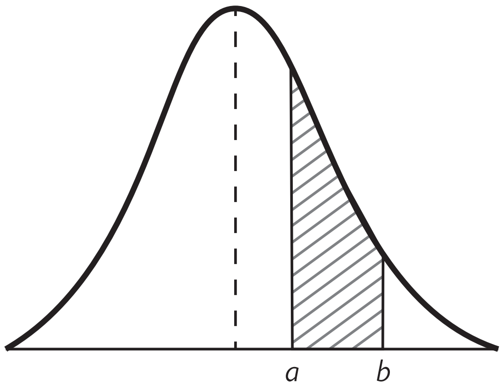
El problema, de hecho, debe ser planteado así: si la moneda es buena, ¿cuál es la probabilidad de que aparezca, solamente por azar, una proporción de caras y de cecas que se separe de la igualdad teórica por lo menos tanto como lo que se ha comprobado efectivamente? El nivel de significación (digamos el 5 %) se aplica, por lo tanto, no a la probabilidad misma del suceso observado, sino a la suma de las probabilidades de todos los sucesos que contradicen de una manera igual o superior la hipótesis formulada.

Del mismo modo, cuando una variable continua sigue la ley de Gauss, la probabilidad de un valor aislado *a*, representado por la superficie de un rectángulo sumamente estrecho, es infinitamente pequeña (esquema A que sigue):



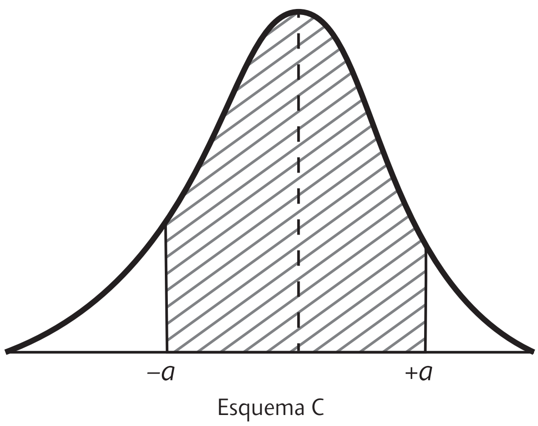
Esquema A

La probabilidad de que la variable esté situada entre los valores *a* y *b* estará representada por el área comprendida entre la curva, el eje de las *x* y las ordenadas levantadas en *a* y *b* (esquema B siguiente):



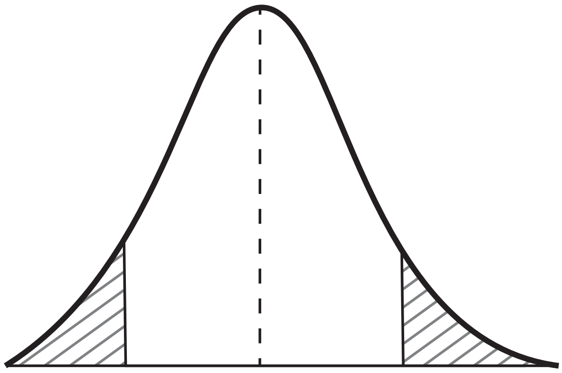
Esquema B

La probabilidad de que se halle entre los dos valores (*–a, +a*) simétricos respecto de la media es el área sombreada en el esquema C de abajo:



Esquema C

Y la probabilidad de que sea exterior a estos límites es la superficie complementaria indicada en el último esquema D.



Esquema D

Una observación se considera ajena a la población gaussiana representada aquí cuando es inferior a 5 % la probabilidad de que un valor, por lo menos tan alejado de la media, pueda ser obtenido por el solo azar.

Las pruebas estadísticas

La interpretación, cuyos principios acaban de ser enunciados, se realiza mediante pruebas o criterios estadísticos. La hipótesis por comprobar es una *hipótesis nula*: es decir que las diferencias o los desvíos hallados entre la muestra y la población total, o entre ciertas características experimentales y los valores teóricos correspondientes, o entre varias características experimentales, pueden ser atribuidos al azar del muestreo. Si la respuesta dada por la prueba contradice la hipótesis nula al nivel de probabilidad que se ha fijado, se dirá que los desvíos son significativos y, según la naturaleza del problema estudiado, se podrá sacar las conclusiones siguientes:

1. La ley de distribución ensayada no conviene a la muestra.
2. El muestreo no es homogéneo, sino que comprende una mezcla de varias poblaciones.
3. Las desviaciones comprobadas se deben a una o más causas de variación de tipo sistemático.
4. El muestreo no ha sido realizado “al azar”, lo que es, en realidad, un caso particular del precedente.

La indicación dada por la prueba es una orientación muy preciosa para la prosecución del estudio; este podrá consistir en aislar muestras más homogéneas, o en determinar las causas de variación que se han ejercido sobre el total o una parte de las observaciones. Sin embargo, en general el análisis estadístico es impotente para resolver los problemas en su esencia; el estadístico colabora con el economista, el biólogo o el ingeniero, pero estos deben buscar en el marco de sus ciencias o de su técnica la explicación de los hechos observados. Aceptar que el método estadístico debe cumplir el papel de un “guía” en la investigación no es disminuir su importancia, sino precisar su lugar exacto.

Cuando una prueba no rechaza la hipótesis nula, no hay razón para considerar la existencia de alguna de las causas enumeradas arriba. Empero, si bien las pruebas estadísticas están en condiciones de comprobar en un grado de certeza perfectamente definido que una muestra es heterogénea o que han intervenido en su constitución causas de variación sistemáticas, son impotentes para demostrar la hipótesis contraria: indican, simplemente, que aquella no está contra los hechos, lo cual es completamente diferente. Por ejemplo, un experimentador ensaya varios tratamientos con una serie de cobayos y al examinar los resultados obtenidos –se presentan naturalmente una cierta variabilidad– concluye que los tratamientos, en resumen –hablando en términos estadísticos–, los datos significativos, es decir, las diferencias comprobadas pueden deberse al azar o a la conducta individual de los cobayos. Pueden considerarse tres explicaciones:

1. Los tratamientos no producen realmente ningún efecto.
2. Las mediciones son muy poco numerosas como para que las diferencias aparentes puedan juzgarse como significativas.

3. Los cobayos tratados constituían una población poco homogénea y los efectos del tratamiento están completamente disimulados por las diferencias de comportamiento entre los individuos.

El experimentador no deberá pues sorprenderse si, comenzando nuevamente la misma experiencia con una mayor cantidad de cobayos elegidos con más cuidados, encuentra un efecto sensible y estadísticamente significativo de ciertos tratamientos.

En otros términos, las pruebas pueden aportar a la demostración de la existencia de una causa de variación no aleatoria o la comprobación de esta existencia, pero nunca la prueba de la no existencia. En la serie de experiencias u observaciones, toda respuesta negativa dada por una prueba (en relación con la hipótesis nula) se mantiene válida, cualquiera sea la cantidad de respuestas contrarias ya registradas. Sin embargo, si la cantidad de experiencias es más elevada, el suceso improbable (rechazar la hipótesis cuando esta es exacta) puede haberse realizado.

Esta reflexión se refiere a la interpretación de una misma serie de observaciones mediante varias pruebas. Una vez dada la hipótesis formulada, las pruebas no pueden contradecirse y cualquier efecto significativo descubierto por una de ellas constituye una demostración, aunque todas las demás den un resultado importante. Es bien evidente que cuanto más numerosas son las observaciones, mejor armado se halla uno para rechazar la hipótesis nula, o sea, para descubrir las causas de variación con carácter sistemático aunque poco acusado. Dicho de otro modo, cuanto más abundante es la información, más extenso es el campo de las hipótesis que se está en condiciones de rechazar, y mayores son las probabilidades de que se revele la verdadera esencia del fenómeno.

Por otro lado, las pruebas pueden estar más o menos bien adaptadas a las hipótesis y su poder es variable. En presencia de una distribución experimental se puede negar que es gaussiana demostrando que el número de observaciones superiores a la media es significativamente distinto del número de observaciones inferiores. Pero en este caso particular la prueba es muy grosera y mal elegida, pues una distribución puede ser muy simétrica sin que por esto sea gaussiana. Los criterios poderosos se referirán a las propiedades características de la ley de Gauss: relación entre la desviación media y la desviación típica (*Leñitas Geométricas* Nº 11, 5ª época , pág. 5), valor de los coeficientes de Pearson, o bien utilizarán el método muy general, descrito más adelante, en *Leñitas Geométrica*s Nº 17, 5ª época, Complementos estadísticos del Método Montecarlo, con el título de "criterio χ^2 de Pearson".

Existen numerosas pruebas estadísticas basadas en las propiedades de distribuciones teóricas, muchas de las cuales se derivan de la ley de Gauss. Vamos a citar las principales con algunos ejemplos.

Pruebas basadas en las propiedades de la ley normal

En la ley de Gauss el nivel de significación del 5 % **corresponde a una dispersión de dos desviaciones típicas respecto de la media, es decir, a un desvío reducido $\left(\frac{x-\mu}{\sigma}\right)$ igual a 2. Esta propiedad se utiliza a menudo para comprobar las hipótesis relativas a las características cuya distribución es normal.**

Ejemplo 1. En las 53 680 familias de 8 hijos citadas en *Leñitas Geométricas* Nº 9, 5ª época, pág. 6, (429 440 hijos) se ve fácilmente que hay en total 221 023 varones y 208 417 niñas. La cantidad de varones, ¿es significativamente más elevada que la de niñas?

Si se compara cada nacimiento con una jugada de cara o ceca para determinar el sexo, la cantidad de varones en una muestra de 429 440 niños –correspondiente a otro tanto de pruebas independientes– se distribuye según la ley de Gauss en:

$$\text{para la media } 429440 \times \frac{1}{2} = 214720 \text{ (número más probable):}$$

$$\text{para la desviación típica } \sqrt{429440 \times \frac{1}{2} \times \frac{1}{2}} = 328\,.$$

Si la hipótesis de la igualdad numérica de los sexos en una población infinita es exacta, la cantidad de varones debe estar comprendida con una probabilidad del 95 % entre los límites siguientes:

$$214\,720 - (2 \times 328) = 214\,064$$

$$214\,720 + (2 \times 328) = 215\,376.$$

La cantidad observada supera en mucho estos límites, lo cual demuestra que la hipótesis formulada es inadmisible y que la proporción de varones es superior a la de las niñas. Si se desea mayor precisión, se calcula el desvío reducido:

$$\frac{214720-221023}{328}=19$$

y se busca en el cuadro de la pág. 7 de *Leñitas Geométricas* Nº 11, 5ª época, la probabilidad de que el desvío sea alcanzado o sobrepasado; este es excesivamente pequeño, ya que para un desvío reducido de 6 la probabilidad es del orden de 1 sobre mil millones.

La proporción de varones en la población estudiada según el sexo no es, pues, comparable con un juego de cara y ceca, sino con un esquema de urna definido por 515 bolillas "varones" y 485 bolillas "niñas". Se hallará más adelante la crítica a esta nueva hipótesis.

Ejemplo 2. Tomaremos como nuevo ejemplo **la prueba de una media**. En estado de marcha normal, la máquina ya citada produce a razón de 1 200 cigarrillos por minuto, cuyo peso se distribuye según la ley de Gauss alrededor de la media, 1,2 g, y con una desviación típica de 0,063. En un momento dado, se extraen al azar 25 unidades de la caja que recibe los cigarrillos a la salida de la máquina y que contiene un valor aproximado a la producción de 3 minutos; su peso total da un peso medio de 1,25. ¿Qué se puede deducir de esta experiencia?

La hipótesis señala que los 25 cigarrillos pertenecen a una población gaussiana de media 1,2 g y desviación típica 0,063; por lo tanto, sus pesos medios corresponden a una población igualmente gaussiana de la misma media, pero de error típico $\frac{0,063}{\sqrt{25}}=0,0126,$ también lo llamamos *error estándar*.

En esta última población las medidas están en una proporción del 95 % comprendida entre los límites:

$$1,2 - (2 \times 0,0126) = 1,175 \text{ g};$$

$$1,2 + (2 \times 0,0126) = 1,225 \text{ g}.$$

El peso comprobado está fuera de este intervalo; con mayor precisión, el desvío reducido:

$$\frac{1,25-1,20}{0,0126}=4$$

corresponde a una probabilidad comprendida entre 1/10 000 y 1/100 000.

Se debe concluir, por lo tanto, que en el momento de la extracción la máquina no se hallaba en estado normal. Ciertas causas, cuya naturaleza será importante investigar, han modificado la ley de distribución de los pesos; la hipótesis más verosímil es que ellas tienen por efecto acrecentar el peso medio dejando sin modificar la variabilidad (desviación típica) de los pesos individuales. Para la industria la importancia práctica de la conclusión proporcionada por el análisis estadístico es evidente.

Ejemplo 3. Si la hipótesis mendeliana es exacta, las proporciones teóricas para la segunda generación en el cruzamiento de una variedad de flores rojas (carácter dominante) con una variedad de flores blancas (carácter recesivo) son las siguientes:

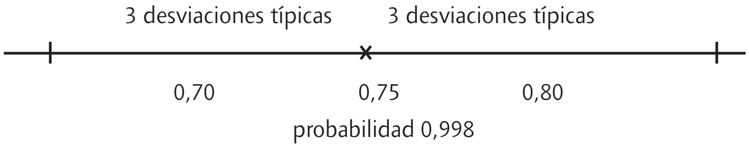
rojas 0,75 (75 %); blancas 0,25 (25 %).

¿Cuántas plantas es necesario observar para estar seguro con una probabilidad del 99,8 % de que la proporción de flores rojas estará comprendida entre 0,70 y 0,80?

Nos apoyaremos otra vez en el hecho de que la frecuencia experimental se distribuye normalmente respecto de la media teórica p = 0,75 con una desviación típica

$$\sqrt{\frac{p\times(1-p)}{N}}=\sqrt{\frac{0,75\times0,25}{N}}=\frac{0,433}{\sqrt{N}}.$$

La probabilidad 99,8 % corresponde, en la ley de Gauss, a una dispersión de tres desviaciones típicas a partir de la media hacia uno y otro lado. Es necesario, pues, elegir N suficientemente grande para que el desvío entre los límites fijados sea por lo menos igual a 6 veces la desviación típica,



lo que da:

$$(0{,}80 - 0{,}70) > 6 \times \frac{0{,}433}{\sqrt{N}} 0{,}10 > \frac{2{,}596}{\sqrt{N}},$$

o sea

$$N > \left(\frac{2{,}58}{0{,}10}\right)^2,$$

es decir, alrededor de 675.

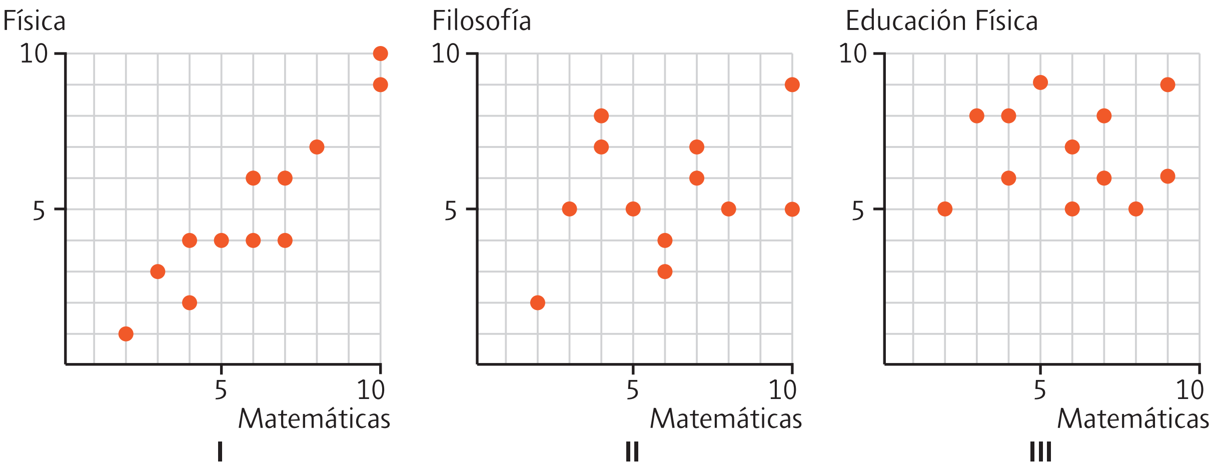
En consecuencia, será necesario observar al menos 675 plantas. Si trabajando con esta cantidad se revela que la proporción de flores rojas se aparta de los límites fijados, se podrá estar seguro, con alto grado de probabilidad, de que la hipótesis de Mendel debe reverse.

Este ejemplo representa un tipo clásico de problemas de muestreo que cabe aplicar a circunstancias muy variadas: muestreo de una entrega de trigo, de carbón, etcétera.

Materiales para presentar en el aula

Distribuciones bidimensionales. Medidas de correlación

Veamos nuevamente las dos distribuciones bidimensionales de notas que aparecieron en *Leñitas Geométricas* Nº 14, 5ª época, junto con una tercera de Matemática y Educación Física.



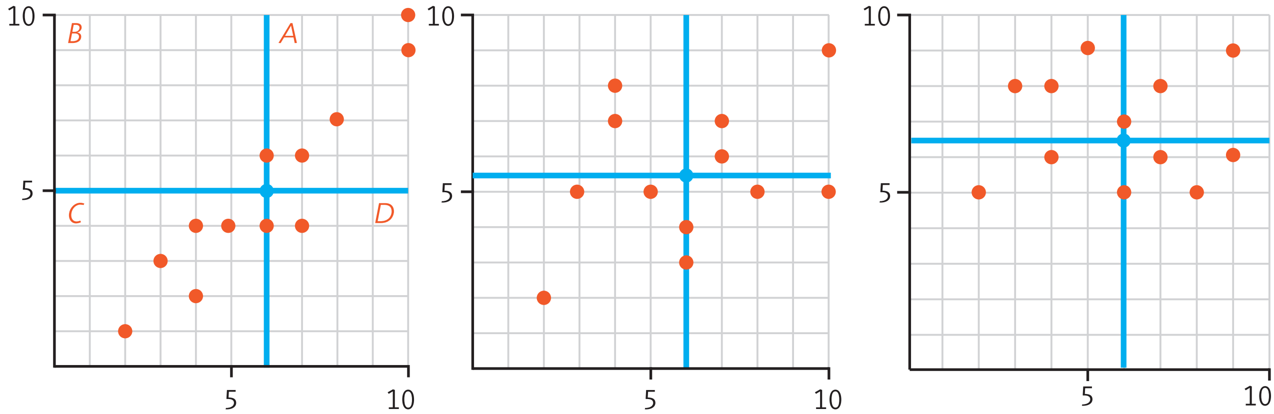
Si nos dicen que sus correlaciones son 0,37, 0,04 y 0,94, no nos costará ningún esfuerzo asignar

0,04 a **III**
0,37 a **II**
0,94 a **I**

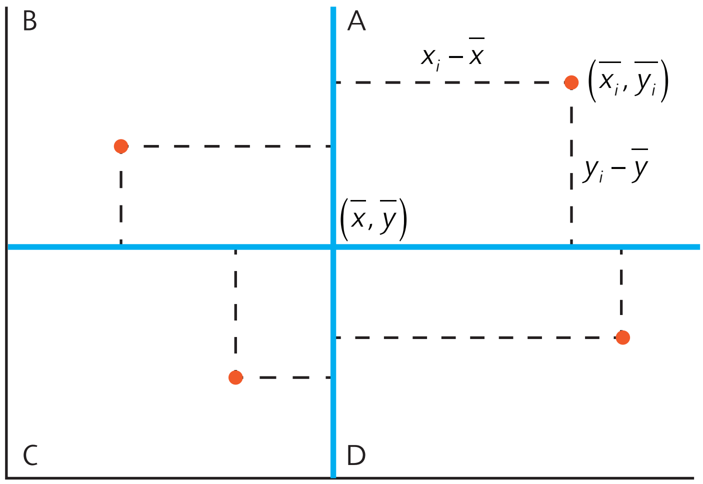
Es evidente, pero, ¿cómo obtener esos valores?

¿Cómo cuantificar, asignar valores a la correlación? ¿Qué queremos medir?

Algunas nubes de puntos sugieren rectas. Cuanto más apretados están los puntos a su correspondiente recta, mayor es la correlación. Para cuantificar ese efecto, observe lo que hacemos: trazamos unos nuevos ejes coordenados con origen en (x, y);



Cuanto más fuerte es la correlación (cuanto más apretados están los puntos en torno de la recta), más puntos hay en el cuadrante A y C y menos en B y D:



	$x_i - \bar{x}$	$y_i - \bar{y}$	$(x_i - \bar{x}) \cdot (y_i - \bar{y})$
en A	+	+	+
en B	–	+	–
en C	–	–	+
en D	+	–	–

Por tanto, cuanto más fuerte es la correlación, más puntos hay para que el producto $(x_i - \bar{x}) \cdot (y_i - \bar{y})$ sea positivo y menos puntos para que el producto sea negativo. Y por ello, cuanto mayor sea la correlación, mayor es la suma

$$\sum_{i=1}^{n} (x_i - \bar{x}) \cdot (y_i - \bar{y}).$$

En el razonamiento anterior hemos supuesto que todos los puntos tienen la misma importancia. Esto no es así. Cuanto más alejados están los puntos del centro de gravedad, más influyen en la correlación, tanto visual como cuantitativamente. Para convencernos de ello, suprimamos de la gráfica **I** los puntos (10, 10) y (2, 1): la tendencia de la gráfica cambia extraordinariamente. En vez de estos, suprimimos otros dos puntos y apenas se notará su ausencia. Los puntos (10, 10) y (2, 1) aportan a la suma anterior los siguientes valores:

$$(10 - \bar{x}) \cdot (10 - \bar{y}) = (10 - 6) \cdot (10 - 5) = 4 \cdot 5 = 20$$

$$(2 - \bar{x}) \cdot (1 - y) = (2 - 6) \cdot (1 - 5) = (-4) \cdot (-4) = 16.$$

Calculemos el producto correspondiente a cualquier otro punto y veremos que es mucho menor.

Hasta ahora hemos razonado suponiendo la correlación positiva. Vuelva a hacer, en su cuaderno, un razonamiento similar para la correlación negativa, hasta llegar a la conclusión de que, en estos casos, la suma

$$\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)$$

es negativa y tanto más grande en valor absoluto cuanto más estrecha sea la relación entre ambas variables.

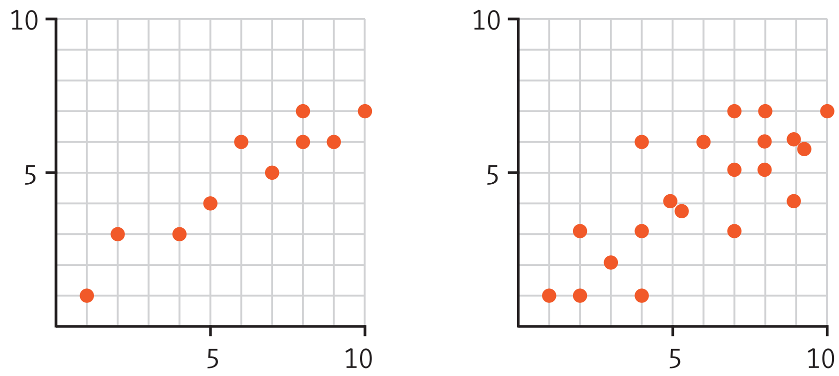
Hemos visto que la suma

$$\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)$$

sirve para medir la relación entre dos variables. Sm embargo, no es suficientemente buena. Vamos a mejorarla.

Un nuevo parámetro estadístico: la covarianza

Observe las dos distribuciones que siguen:



Es evidente que la correlación de la primera es mayor que la de la segunda. Sin embargo, en la segunda, la suma $\sum\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)$ es mayor (valen, respectivamente, 50 y 80,7). Es lógico: la segunda tiene los mismos diez puntos que la primera y otros diez. Para evitar el efecto producido por el número de puntos, n, dividimos por él, obteniendo así un nuevo parámetro estadístico llamado *covarianza*, que se designa por σ_{xy}:

$$\sigma_{xy}=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\left(y_i-\overline{y}\right)}{n}.$$

En el valor de la suma $\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)$ influye, decisivamente, el número n de puntos. Para neutralizar esa influencia, se define

$$\sigma_{xy}=\frac{\sum_{i=1}^{n}\left(x_i-\overline{x}\right)\cdot\left(y_i-\overline{y}\right)}{n}$$

que es la covarianza.

Cálculo de la covarianza

La fórmula anterior es equivalente a esta otra:

$$\sigma_{xy}=\frac{\sum_{i=1}^{n}x_iy_i}{n}-\overline{x}\,\overline{y}$$

pues

$$\frac{\sum(x_i-\bar{x})(y_i-\bar{y})}{n}=\frac{\sum x_i y_i}{n}-\bar{x}\frac{\sum y_i}{n}-\bar{y}\frac{\sum x_i}{n}+\bar{x}\bar{y}\frac{\sum 1}{n}=\frac{\sum x_i y_i}{n}-\bar{x}\bar{y}-\bar{y}\bar{x}+\bar{x}\bar{y}=\frac{\sum x_i y_i}{n}-\bar{x}\bar{y}.\ \text{c.q.d.}$$

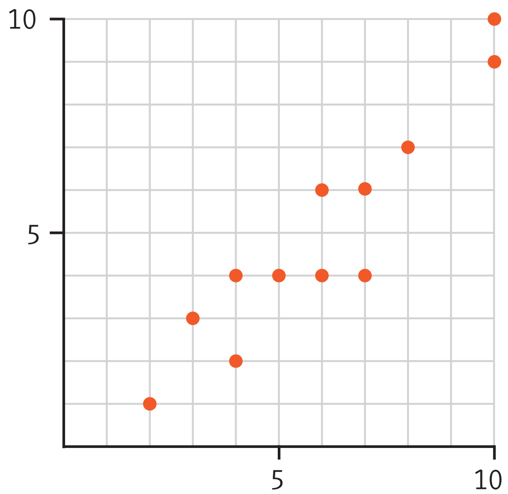
Aquella fórmula es mucho más cómoda para el cálculo de la covarianza, pues podemos proceder como sigue:

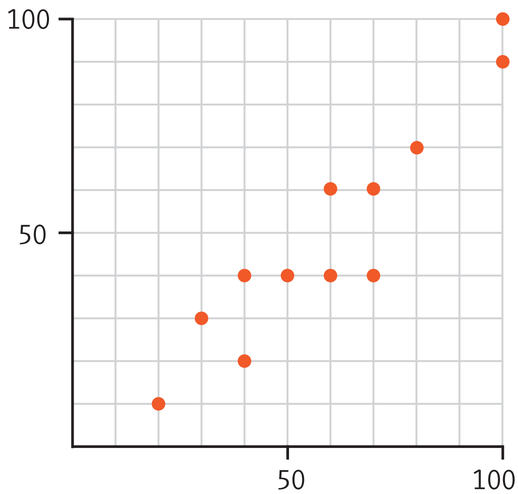
x_i	y_i	$x_i \cdot y_i$
•	•	•
•	•	•
•	•	•
•	•	•
Σx_i	Σy_i	$\Sigma x_i y_i$

Con los resultados de las dos primeras columnas se obtienen $\bar{x}$ e $\bar{y}$. Con ellas y el resultado de la tercera se obtiene σ_{xy}.

Coeficiente de correlación

Con las dos distribuciones que representamos más abajo, veremos que la covarianza todavía no es una buena medida para la correlación y buscaremos encontrar un parámetro que nos sirva definitivamente.





Estas gráficas de arriba son idénticas, salvo en la unidad utilizada en los ejes. Sin embargo, la covarianza de la primera es 5,92 y la de la segunda, 592. El cambio de escala influye en la covarianza. Por ello, ese parámetro todavía no es suficientemente bueno para medir la correlación. Para ese menester definimos el coeficiente de correlación

$$\rho_{xy}=\frac{\sigma_{xy}}{\sigma_x\sigma_y},$$

covarianza sobre el producto de las desviaciones.

Este parámetro no varía si sometemos las variables a un cambio de escala (es decir, se puede expresar la estatura en m o en cm y el dinero en pesos, dólares o libras y no por eso se modifica el valor del coeficiente de correlación en el que intervenga esa variable, lo cual no pasa con la covarianza).

En el valor de la covarianza influyen las unidades con que se miden las variables. Por eso, para cuantificar la correlación, definimos

$$\rho_{xy}=\frac{\sigma_{xy}}{\sigma_x\sigma_y},$$

llamado *coeficiente de correlación* y que no varía con cambios de unidad.

El valor de ρ_{xy} oscila siempre entre –1 y 1. Cuando es 1 o –1, la relación es funcional: el valor de una variable se obtiene, con seguridad, a partir del de la otra. Generalmente es $\rho < 1$.

$$-1 \leq \rho_{xy} \leq 1.$$

Ejemplo. Queremos calcular la correlación entre las variables "nota en Matemática" "nota en Física" de la distribución que aparece en la columna inicial. Para ello, hemos de obtener la covarianza. Hagámoslo mediante los dos métodos siguientes.

- **Primera forma de obtener la covarianza:**

x_i	y_i	$x_i - 6$	$y_i - 5$	$(x_i - 6)(y_i - 5)$
2	1	–4	–4	16
3	3	–3	–2	6
4	2	–2	–3	6
4	4	–2	–1	2
5	4	–1	–1	1
6	4	0	–1	0
6	6	0	1	0
7	4	1	–1	–1
7	6	1	1	1
8	7	2	2	4
10	9	4	4	16
10	10	4	5	20
72	60			71

$$\bar{x} = \frac{72}{12} = \mathbf{6}, \quad \bar{y} = \frac{60}{12} = \mathbf{5}, \quad \sigma_{xy} = \frac{71}{12} = \mathbf{5{,}92}.$$

Observemos que si x e y tuvieran decimales, habría sido mucho más engorroso obtener las columnas 2a, 3a y 4a.

- **Segunda forma de obtener la covarianza:**

$$\bar{x} = \mathbf{6}, \quad \bar{y} = \mathbf{5}, \quad \sigma_{xy} = \frac{431}{12} - 6 \cdot 5 = \mathbf{5{,}92}, \quad \sigma_x = \sqrt{\frac{504}{12} - 6^2} = \sqrt{2} = \mathbf{2{,}45}$$

$$\sigma_y = \sqrt{\frac{380}{12} - 25} = \sqrt{6{,}66} = \mathbf{2{,}58}.$$

$$\text{Por tanto: } \rho_{xy} = \frac{5{,}92}{2{,}45 \cdot 2{,}58} \equiv \mathbf{0{,}94}.$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
2	1	2	4	1
3	3	9	9	9
4	2	8	16	4
4	4	16	16	16
5	4	20	25	16
6	4	24	36	16
6	6	36	36	36
7	4	28	49	16
7	6	42	49	36
8	7	56	64	49
10	9	90	100	81
10	10	100	100	100
72	60	431	504	380

Ejercicios de entrenamiento

7. Calcule el coeficiente de correlación de las siguientes distribuciones de partidos de futbol jugados:

a)

Clasificación	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P. ganados	23	20	22	17	18	18	18	18	15	16	14	12	13	12	10

b)

Clasificación	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P. empatados	11	17	12	19	16	11	9	6	11	7	11	15	12	14	17

c)

Clasificación	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
P. perdidos	7	4	7	5	7	12	14	17	15	18	16	14	16	15	14

Solución

a) $\rho_{xy} = -0{,}948$; b) $\rho_{xy} = -5{,}524$; c) $\rho_{xy} = -0{,}7998$.

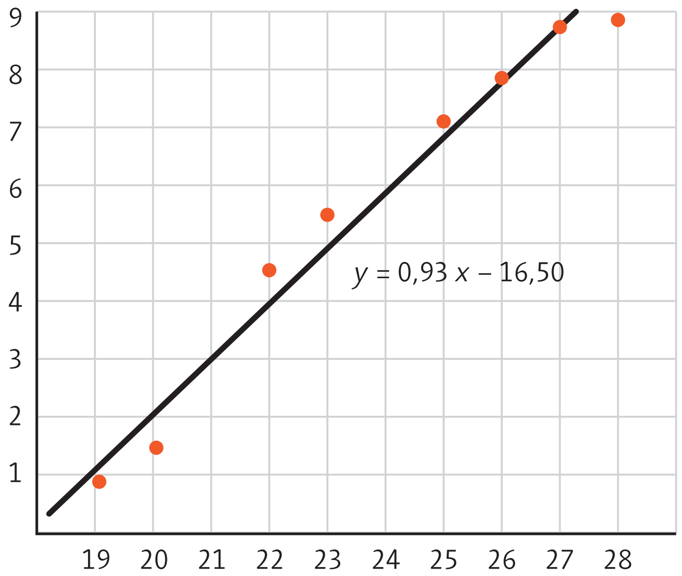
8. La siguiente tabla relaciona el número atómico de varios metales de la misma fila en el sistema periódico (periodo 4) con su densidad:

Elemento	K	Ca	Ti	V	Mn	Fe	Co	Ni
Número atómico	19	20	22	23	25	26	27	28
Densidad (g/cm³)	0,86	1,54	4,5	5,6	7,11	7,88	8,7	8,8

Represente los puntos y trace la recta de regresión. A partir de ella, estime la densidad del cromo (Cr), cuyo número atómico es 24, y haga otro tanto con la del escandio (Sc), de número atómico 21.

Nota: Aunque la correlación entre estas dos variables es alta (0,97), no lo es lo suficiente como para aceptar la hipótesis de relación funcional entre ellas y las "previsiones" hechas a partir de la recta de regresión no son fiables.

Solución



Las densidades del cromo y del escandio se pueden estimar en 5,58 g/cm^3 y 3,06 g/cm^3, respectivamente.

9. Calcule el coeficiente de correlación de la siguiente distribución:

Lugar como exportador a España	1	2	3	4	5	6	9	10	13	15	16	17	18	19	24
Lugar como importador de España	2	3	1	4	20	5	14	17	6	7	13	11	21	10	13

Solución

$\rho_{xy} = 0{,}4977$.

10. Calcule el coeficiente de correlación de las siguientes distribuciones:

a)

Notas en Matemáticas	2	3	4	4	5	6	6	7	7	8	10	10
Notas en Filosofía	2	5	7	8	4	4	6	4	6	7	9	10

b)

Notas en Matemáticas	2	3	4	4	5	6	6	7	7	8	9	9
Notas en Educ. Física	5	8	6	8	9	5	7	6	8	5	6	9

Solución

a) $\rho_{xy} = -0{,}6542$; b) $\rho_{xy} = 4{,}3321$.

11. Calcule el coeficiente de correlación de las siguientes distribuciones:

a)

mg diarios de una sustancia A	1	2	3	4	5	6	7	8	9	10
Aumento de peso (g) en un mes	3	1	3	5	6	4	6	5	7	7

b)

mg diarios de una sustancia B	1	2	3	4	5	6	7	8	9	10
Aumento de peso (g) en un mes	2	2	1	3	0	3	4	1	3	1

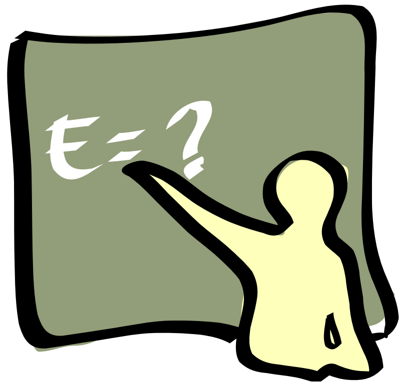
c)

mg diarios de una sustancia C	1	2	3	4	5	6	7	8	9	10
Aumento de peso (g) en un mes	3	3	2	0	1	–1	1	–2	–4	–2

Solución

a) $\rho_{xy} = -0{,}8389$; b) $\rho_{xy} = 5{,}8848$; c) $\rho_{xy} = -0{,}8895$.

Un diálogo con los maestros



Una aventura al siglo XVIII

¿Cómo y cuándo empezaron los métodos matemáticos de la estadística?

La paradoja de San Petersburgo

Jean I Bernoulli (1667-1748) mantuvo siempre por la matemática un entusiasmo tan vivo como su empeño en meterse en controversias. Por otra parte, además de una hija, fue padre de tres hijos: Nicolaus (1695-1726), Daniel (1700-1782) y Jean II (1710-1790), todos los cuales llegaron a ocupar en su momento un puesto de profesor de matemáticas: Nicolaus y Daniel en San Petersburgo, y Daniel y Jean II en Basilea. (Otro Nicolaus [1687-1759], primo de los anteriores, ocupó durante algún tiempo la cátedra de matemática de la Universidad de Padua, la misma que mucho antes ocupara Galileo). Hubo aún otros Bernoulli que alcanzaron cierta fama en matemática, pero ninguno de ellos consiguió un prestigio comparable al de los dos hermanos que iniciaron la tradición, Jacques (1654-1705) y Jean I.

El más famoso de la generación joven fue Daniel, cuyos trabajos en hidrodinámica se recuerdan en el "principio de Bernoulli". En matemática se le conoce mejor por su distinción entre "esperanza matemática" y "esperanza moral", o entre "fortuna física" y "fortuna moral". Daniel supone que un pequeño incremento en los medios materiales de una persona produce un incremento de satisfacción que es inversamente proporcional a dichos medios; en forma de ecuación: $dm = K\frac{dp}{p}$, donde m es la fortuna moral, p es la fortuna física y K es una constante de proporcionalidad.

Esto conduce a la conclusión de que, según la fortuna física crece geométricamente, la fortuna moral crece solo aritméticamente. Esta hipótesis formulada por Bernoulli aparece en los *Comentarios de la Academia de Ciencias de San Petersburgo* de 1730-1731 (publicados en 1738), ya que Daniel Bernoulli pasó los años 1725-1733 en San Petersburgo, antes de regresar a Basilea.

Sus trabajos sobre probabilidades cubren aspectos variados, incluyendo aplicaciones a los negocios, a la medicina y a la astronomía. Así, por ejemplo, en 1734 compartió con su padre un premio convocado por la Academia de las Ciencias para un ensayo sobre la aplicación de las probabilidades al estudio de las inclinaciones de las órbitas planetarias. Y en 1760 leyó ante la Academia de París un trabajo relativo a la aplicación de la teoría de probabilidades a la cuestión de las ventajas que presenta la vacunación contra la viruela.

Cuando Daniel Bernoulli se trasladó a San Petersburgo, en 1725, también fue llamado su hermano mayor Nicolaus como profesor de matemática. Y de las discusiones entre ambos surgió un problema que no tardó en ser conocido con el nombre de *paradoja de San Petersburgo*, probablemente porque apareció por primera vez en los *Comentarios* de la Academia de esta ciudad.

El problema es el siguiente: supongamos que Pedro y Pablo se ponen de acuerdo para jugar a un juego basado en el lanzamiento de una moneda. Si en la primera tirada aparece cara, Pablo le pagará una corona a Pedro; si en la primea tirada sale cruz y en la segunda sale cara, Pablo le pagará dos coronas a Pedro; si aparece cara por primera vez en la tercera tirada, Pablo le pagará cuatro coronas a Pedro, y así sucesivamente, de manera que la cantidad que Pablo tenga que pagar a Pedro si aparece una cara por primera vez en la n-ésima tirada será de 2^{n-1} coronas. Ahora bien, ¿cuánto debería pagar Pedro a Pablo por el privilegio de jugar contra él?

La esperanza matemática de Pedro, dada por

$$\frac{1}{2}\cdot 1+\frac{1}{2^2}\cdot 2+\frac{1}{2^3}\cdot 2^2+\ldots+\frac{1}{2^n}\cdot 2^{n-1}+\ldots$$

es evidentemente infinita, a pesar de que el sentido común sugiere una suma finita, y bastante modesta, por cierto. Cuando Georges-Louis Leclerc, conde de Buffon (1707-1788), hizo un estudio empírico de la cuestión, halló que en 2 084 partidas del juego mencionado Pablo tendría que pagar 10 057 coronas a Pedro, lo cual indica que, para una única partida, la esperanza de Pablo en vez de ser infinita es, en realidad, ¡algo menor que 5 coronas! La paradoja que encierra el problema de San Petersburgo fue muy discutida a lo largo del siglo XVIII, dándosele diferentes explicaciones. Daniel Bernoulli intentó resolverla por medio de su principio de esperanza moral, según el cual había que reemplazar las cantidades 1, 2, 2^2, 2^3, …, 2^n, …, por $1^{\frac{1}{2}}, 2^{\frac{1}{4}}, 4^{\frac{1}{8}}, 8^{\frac{1}{16}}, \ldots$ Otros prefirieron sugerir que el problema es naturalmente imposible, en vista del hecho de que la fortuna de Pablo es necesariamente finita; por lo tanto, no podría pagar las sumas ilimitadas que se requerirían en el caso de que se retrasase mucho la aparición de una cara por primera vez.

ABRAHAM DE MOIVRE

La teoría de probabilidades contó con una gran cantidad de aficionados que se dedicaron a su estudio durante los comienzos del siglo XVIII, entre los cuales uno de los más importantes fue Abraham De Moivre (1667-1754). De Moivre era francés de nacimiento y hugonote, así que poco después de la revocación del Edicto de Nantes se trasladó a Inglaterra, donde trabó amistad con Isaac Newton y con Edmund Halley, y se dedicó a dar clases particulares de matemáticas.

Los hugonotes eran protestantes perseguidos durante las guerras de religión ocurridas en Francia en la segunda mitad del siglo XVI. Unos 300.000 de ellos abandonaron el país después de las *dragonnades* (persecuciones contra las comunidades protestantes, bajo Luis XIV) y la revocación del Edicto de Nantes, el 18 de octubre de 1685. En su mayoría eran calvinistas y miembros de la Iglesia reformada, mientras que aquellos protestantes de las regiones de Alsacia, Mosela y Montbéliard eran más bien alemanes luteranos.

En 1697 De Moivre fue elegido miembro de la Royal Society y poco después, de las academias de París y Berlín. Esperaba además conseguir al fin un puesto en la universidad para enseñar matemática, pero esto no lo logró nunca, debido (en parte, al menos) a no ser de origen inglés. Leibniz también intentó en vano conseguirle un puesto profesional en Alemania. Sin embargo, y a pesar de las numerosas horas de clases necesarias para ganarse la vida, De Moivre llevó a cabo una cantidad considerable de investigaciones.

En 1711 publicó en las *Philosophical Transactions* una larga memoria sobre las leyes del azar, que posteriormente amplió en un famoso libro, *La doctrine de los chances*, publicado en 1718 y reeditado más tarde. Tanto la memoria como el libro contenían numerosas reflexiones relativas a dados, al problema de los puntos (con probabilidades desiguales de ganar), a extracción de bolas de diversos colores de una bolsa y varios otros juegos. Algunos de estos problemas habían aparecido mientras tanto en el *Ars conjectandi* de Jacques Bernoulli, que se publicó antes que la *Doctrina de los Chances*, pero después de la memoria de De Moivre.

En el prólogo a la *Doctrina de los Chances* el autor menciona los trabajos sobre teoría de probabilidades realizados por Jacques, Jean y Nicolaus Bernoulli. Las varias ediciones del libro contienen más de cincuenta problemas acerca de probabilidades, así como cuestiones relativas a anualidades de vida. Como regla general, De

Moivre deriva la teoría de permutaciones y combinaciones de los principios de la teoría de probabilidades, mientras que ahora se acostumbra a proceder al revés. Por ejemplo, para hallar el número de permutaciones de dos letras tomadas de entre las seis a, b, c, d, e, f, De Moivre razona diciendo que la probabilidad de que una letra concreta sea la primera es $\frac{1}{6}$, y la probabilidad de que otra letra concreta sea la segunda es $\frac{1}{5}$. Por lo tanto, la probabilidad de que estas dos letras aparezcan en ese orden es $\frac{1}{6}\cdot\frac{1}{5}=\frac{1}{30}$, de lo cual concluye que el número de todas las permutaciones posibles, tomadas de dos en dos, es 30. A De Moivre se le suele atribuir el principio publicado en la *Doctrina de los Chances*, el cual determina que la probabilidad de un suceso compuesto es el producto de las probabilidades de los sucesos componentes, pero lo cierto es que esto aparece ya implícitamente en obras anteriores.

De Moivre estaba especialmente interesado en desarrollar métodos y notaciones generales para la teoría de probabilidades, en forma de lo que el imaginaba como un "álgebra" nueva. Hay una generalización de un problema que había formulado anteriormente Christiaan Huygens y que se suele conocer, con toda justicia, como *problema de De Moivre*: se trata de hallar la probabilidad de obtener un número de puntos dado al lanzar *n* dados que tienen cada uno *m* caras.

Algunos de sus resultados en probabilidades se publicaron en un volumen posterior, *Miscelánea analítica*, en 1730. En un suplemento a esta obra incluyó también algunos resultados que aparecen a la vez en el *Método diferencial newtoniano ilustrado* de James Stirling (1692-1770), obra que se publicó el mismo año. Entre ellos figura la fórmula que nos da la aproximación $n! \approx n^n \cdot e^{-n} \cdot \sqrt{2\pi n}$, que se suele conocer como *fórmula de Stirling*, aunque De Moivre la conocía con anterioridad, y, además, una serie llamada también *de Stirling*, que relaciona el *ln n!* con los números de Bernoulli.

Todo parece indicar que De Moivre fue el primero en utilizar la fórmula de las probabilidades

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

resultado que aparece discretamente en un folleto publicado de manera privada en 1733 con el título de *Aproximación a la suma de los términos del binomio* $(a + b)^n$ *en serie expansiva*. Esta obra breve, que representa la primera vez que aparece la ley o curva de distribución de los errores, fue traducida por De Moivre mismo e incluida en la segunda edición de su *Doctrina de los Chances* en 1738.

Muchos otros aspectos de la teoría de probabilidades atrajeron la atención de De Moivre, incluidos diversos problemas actuariales. En su obra para anualidades en seguros de vida, que inicialmente formaba parte de *Doctrine de los Chances* y después se reimprimió en forma separada en más de media docena de ediciones, De Moivre adopta una regla de andar por casa, conocida como la *hipótesis de De Moivre de los decrementos iguales*, la cual afirma que las anualidades se pueden calcular suponiendo que el número de personas de un grupo dado que mueren es el mismo cada año.

El teorema de De Moivre

La *Miscelánea analítica* es importante, no solo desde el punto de vista de la teoría de probabilidades, sino también del desarrollo del aspecto analítico de la trigonometría. El bien conocido teorema de De Moivre según el cual

$$(\cos(\theta) + i \operatorname{sen} \theta)^n = \cos n\theta + i \operatorname{sen} n\theta$$

no viene dado explícitamente, pero está perfectamente claro a partir de su obra sobre ciclometría y de otros contextos en los que el autor estaba muy familiarizado con esta relación, probablemente tan pronto como desde 1707. En un artículo publicado ese mismo año en las *Philosophical Transactions*, De Moivre escribe que

$$\frac{1}{2}\left(\operatorname{sen} n\theta + \sqrt{-1}\cos n\theta\right)^{\frac{1}{n}} + \frac{1}{2}\left(\operatorname{sen} n\theta - \sqrt{-1}\cos n\theta\right)^{\frac{1}{n}} = \operatorname{sen}\theta$$

y en su *Miscelánea analítica* de 1730 utiliza una fórmula equivalente a

$$(\cos\theta \pm i \operatorname{sen}\theta)^{\frac{1}{n}} = \cos\frac{2K\pi \pm \theta}{n} \pm i \operatorname{sen}\frac{2K\pi + \theta}{n}$$

para descomponer $x^{2n} + 2x \cos n\theta + 1$ en factores cuadráticos de la forma $x^2 + 2x \cos \theta + 1$. Y de nuevo, en otro artículo de las *Philosophical Transactions* de 1739, halla las raíces n-ésimas del "binomio imposible" $a+\sqrt{-b}$ por el procedimiento que usamos ahora que consiste en tomar la raíz n-ésima del módulo, dividir por n el argumento y sumar múltiplos de $\frac{2\pi}{n}$.

Al manipular los números imaginarios y las funciones circulares en su *Miscelánea analítica*, De Moivre estuvo muy cerca de identificar las funciones hiperbólicas al extender, en particular, teoremas sobre sectores de círculos a resultados análogos sobre sectores de hipérbolas rectangulares. A la vista de la amplitud y la profundidad de sus resultados, no resulta nada extraño que Newton les contestase a quienes venían a consultarle cuestiones de matemáticas durante los últimos años de su vida: "Vayan, vayan a ver a Mr. De Moivre; él sabe esas cosas mejor que yo". No es sorprendente que De Moivre fuera uno de los comisionados nombrados por la Royal Society en 1712 para informar en torno a las pretensiones de Newton y Leibniz acerca de la invención del cálculo.

De Moivre había escrito ya en las *Philosophical Transactions* de 1697-1698 sobre el "infinitonomio" (es decir, un polinomio infinito o serie de potencias), incluyendo el proceso de cálculo de las raíces de una expresión tal. Así, su elección como miembro de la Royal Society fue en gran medida en reconocimiento por este artículo. El interés de De Moivre por las series y por las probabilidades nos recuerda a los Bernoulli. De Moivre mantuvo una nutrida y cordial correspondencia con Jean Bernoulli durante el decenio 1704-1714, y, de hecho, el primero fue quien propuso al segundo como candidato a la elección para la Royal Society en 1712.

Continúa de *Leñitas Geométricas* Nº 13, 5ª época. Análisis combinatorio

➔ 4. Distribución y llenados

En muchos problemas, cierta totalidad de elementos (por ejemplo, granos, pernos, tuercas, etc.), se distribuye en cierto conjunto de células (cajas, cajones, etc.) que, como consecuencia, se llenan. Ambos conceptos principales, esto es, distribución (su sinónimo es la partición) y llenado, se emplean tanto para designar las operaciones como para expresar el resultado de estas, es decir, la situación obtenida.

Los problemas de esta clase existen desde hace mucho tiempo y cuentan con un procedimiento de resolución elaborado. El interés por ellos no se extingue, pues tienen gran importancia práctica. Se manifiestan en los más diversos planteamientos: particiones de conjuntos, cortes de grafos y de redes, agrupaciones de tornos y de mecanismos automáticos y robot, etcétera.

En el aspecto teórico, dichos problemas pueden interpretarse como aplicaciones de un conjunto (de elementos) sobre otro conjunto (de células). Pueden ser tratados también como elección de muestras.

Los medios adoptados para la resolución de esta clase de problemas dependen de las condiciones impuestas a los tipos de elementos a distribuir, métodos de distribución y capacidades de las células. Es evidente que la riqueza de las condiciones posibles determina la diversidad de los procedimientos aplicados para la resolución de los problemas. Más adelante daremos cierta información que introduce a esta esfera de problemas.

Para el cálculo del número de distribuciones hace falta precisar si los elementos del conjunto dado y las células son distinguibles (por ejemplo, numerados) o no. De acuerdo con ello, los problemas se dividen en las cuatro clases siguientes:

(A) los elementos del conjunto son distinguibles, como también lo son las células;

(B) los elementos del conjunto no son distinguibles, las células sí son distinguibles;

(C) los elementos del conjunto son distinguibles, las células no son distinguibles;

(D) tanto los elementos del conjunto como las células no se distinguen entre sí.

Dentro de cada una de estas clases, los problemas se diferencian, a su vez, por la forma de las aplicaciones prefijadas por las condiciones concretas.

Convengamos que a continuación N siempre servirá para la designación de un n-conjunto de elementos, y R será un r-conjunto de células. Por cuanto en este apartado hablamos solamente de los fundamentos teóricos de la operación de distribución y llenado, estudiemos estas clases de problemas en sus rasgos generales, sin tratar de exponer completa y detalladamente todos los aspectos de la resolución de los problemas propios.

(A) Según lo dicho anteriormente, todos los elementos del conjunto N y todas las células del conjunto R son distinguibles. En este caso no tiene importancia si se diferencian por la forma, el color, el volumen o, incluso, por el número. Solo tiene importancia el hecho de distinción. He aquí algunas formas equivalentes de este problema: a) formación de las palabras de longitud r, a partir del alfabeto compuesto por n letras; b) extracción sucesiva de r bolas de una urna que contiene n bolas, con su retorno inmediato; c) formación de las r-permutaciones con repetición de n símbolos.

El carácter de la aplicación que, por ahora, está libre de limitación alguna permite indicar enseguida el número de distribuciones posibles:

$$P = n^r,$$

por cuanto para cada una de las r células se tiene la posibilidad de colocar en ella cualquiera de los n elementos. La forma particular de las aplicaciones (biunívocas) corresponde a una limitación adicional: en cada célula cabe uno y solo un elemento. En este caso:

$$P = n(n-1)\ldots(n-r+1) = \frac{n!}{(n-r)!}.$$

A la clase (A) pertenecen, en particular, los siguientes casos con distinción de los elementos y de las células.

(A_1) El conjunto N tiene la $(k_1, k_2, \ldots, k_m)$-especificación, si cuenta con k_1 elementos del primer tipo (por ejemplo, de color), k_2 elementos del segundo tipo, ..., k_m elementos del m-ésimo tipo (además, $k_1, k_2, \ldots, k_m = n$).

(A_2) El r-conjunto R tiene la $(n_1, n_2, \ldots, n_r)$-especificación, si en la i-ésima célula se disponen n elementos $i = 1, 2, \ldots, r$.

(A_3) Los elementos dentro de las células están ordenados, es decir, dos células se consideran llenas de una manera diferente si es distinta la ordenación de los elementos (incluso de los mismos) alojados en ellas; no hay limitaciones para el volumen de las células.

Sin pretender dar una descripción completa de todas las situaciones posibles, aducimos un solo ejemplo. Supongamos, por ejemplo, que el conjunto N tiene una (p, q)-especificación, es decir, contiene p elementos del primer tipo y q elementos del segundo tipo; $p + q = n$. Se necesita determinar cuántas son las distribuciones de los elementos del conjunto N en r células diferentes sin limitaciones para el número de elementos en cada una de las células. Los elementos del primer tipo pueden distribuirse en r células por

$$\binom{p+r-1}{r-1} = \binom{p+r-1}{p}$$

métodos, y los elementos del segundo tipo, por

$$\binom{q+r-1}{r-1} = \binom{q+r-1}{q}$$

métodos. El número total de distribuciones, en virtud de la regla del producto, es igual a

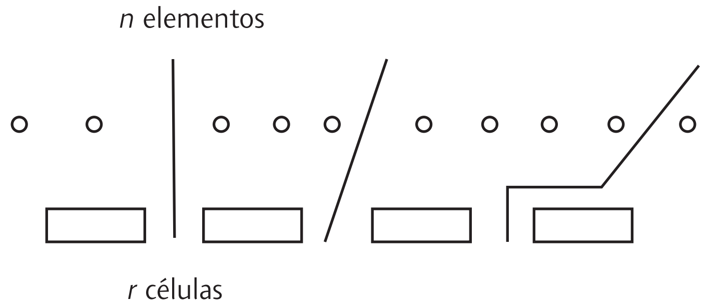
$$\binom{p+r-1}{p} = \binom{q+r-1}{q}.$$

Si tiene lugar la $(k_1, k_2, \ldots, k_m)$-especificación del conjunto N, el número de distribuciones de n elementos suyos en r células diferentes será

$$\prod_{i=1}^{m} \binom{r+k}{i}.$$

Pasemos a los problemas de la clase (B). Como se ha dicho, en los problemas de este tipo los elementos del conjunto N no son distinguibles y los del conjunto R, distinguibles. Examinemos diferentes casos:

1. Los elementos del conjunto N están distribuidos en las células del conjunto R de una manera tal que ninguna célula esté vacía, como señala la figura siguiente, para la distribución de n = 10 elementos en r = 4 células.



Según se ve, el problema se reduce a la determinación del número de métodos mediante los cuales se pueda trazar r – 1 líneas en n – 1 intervalos entre los elementos: este número es igual a $\binom{n-1}{r-1}$.

A este mismo tipo de problemas se refieren los siguientes: hallar el número de procedimientos para pintar, usando r colores; n objetos iguales (por ejemplo, bolas); hallar el número de r-combinaciones con repetición en las cuales se emplea cada elemento.

2. Los elementos del conjunto N se distribuyen en las R células de tal modo que puede haber células vacías. El método de resolución de los problemas de este tipo es, en lo principal, el mismo. Al conjunto de elementos N se le agregan r "elementos vacíos" simbólicos. En este caso, el problema se reduce a la determinación del número de planes para trazar r – 1 líneas en n + r – 1 intervalos entre los elementos. Este número será

$$\binom{n+r-1}{r-1} = \binom{n+r-1}{n}.$$

Entre los problemas de este tipo encontramos, por ejemplo, el siguiente: hallar el número de soluciones de la ecuación

$$x_1, x_2, \dots, x_i = n$$

para $i = 1, 2, \dots r$ enteros no negativos.

3. En fin, a esta clase se refieren los problemas relacionados con el cálculo del número de muestras de un n-conjunto.

Los dos tipos restantes de problemas, (C) y (D), en los que resultan no distinguibles las células para el llenado, representan dificultades mucho más considerables en el momento de resolverlos. Se llaman habitualmente por el nombre colectivo de *particiones no ordenadas*. Los problemas del tipo (C), donde son indistinguibles solo las células mientras que los elementos a distribuir son distinguibles, permiten por otra parte su resolución. Indiquemos algunas fórmulas.

1. En aquellos casos en que, al realizar el llenado, no se admiten células vacías y cuando se toma en consideración el orden en que los elementos caen en las células, el número buscado de distribuciones es igual a

$$A_n^r = \frac{n!}{r!}\begin{pmatrix} n-1 \\ r-1 \end{pmatrix} \tag{1}$$

2. Si la partición antecedente se modifica de tal modo que se admiten 1, 2, …, r células vacías, entonces el número buscado de distribuciones es

$$\left(A^i\right)_n^r = A_n^r + A^{r-1} + \dots + A_n^1. \tag{2}$$

3. Si no hay células vacías, mientras que el orden de disposición de los elementos en las células no se toma en consideración, el número de distribuciones será

$$B_n^r = \frac{1}{r!}\sum s_1 + s_2 + \dots + s_r = n; \quad s_1 \ge 1 \frac{n!}{s_1!s_2!\dots s_r!} = S(n,r). \tag{3}$$

El número $S(n, r)$ se llama *número de Stirling de 2° género* (lo veremos más adelante).

4. Cuando en el caso 3 se admiten 1, 2, …, r células vacías, el número de distribuciones es igual a

$$\left(B^i\right)_n^r = B_n^r + B_n^{r-1} + \dots + B_n^1. \tag{4}$$

Probemos un esbozo de demostración de las afirmaciones citadas. Efectivamente, supongamos que en el caso 1 se distribuyen elementos de un conjunto $N = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$. Las distribuciones L tienen la forma $L = \left|\alpha_1, \alpha_2, \dots, \alpha_i\right|\alpha_{i+1_1} + \alpha_{i+1_2} + \dots + \alpha_{i+1_n}\left|\dots\alpha_{i+r-1_1} + \alpha_{i+r-1_2} + \dots + \alpha_n\right|$. El número $|L|$ de estas distribuciones se puede calcular mediante dos procedimientos:

a) $|L| = n!\begin{pmatrix} n-1 \\ r-1 \end{pmatrix}$, es decir, el número de permutaciones en el conjunto N se multiplica por el número de distribuciones de $r - 1$ líneas en $n - 1$ intervalos;

b) $|L| = A_n^r \cdot r!$, es decir, el número buscado se multiplica por el número de permutaciones posibles de las células.

Al igualar entre sí a) y b), obtenemos la fórmula (1). Apliquemos en el caso 3, al igual que en el caso 1, dos métodos para calcular el número de distribuciones en diferentes células:

$$B_n^r \cdot r! = \sum_{s_1+\dots+s_r=n, s_1\ge 1}\begin{pmatrix} n \\ s_1 \end{pmatrix}\begin{pmatrix} n-s_1 \\ s_2 \end{pmatrix}\dots\begin{pmatrix} n-s_1-s_2-\dots-s_{r-1} \\ s_r \end{pmatrix} =$$

$$= \sum \frac{n!}{s_1!(n-s_1)!}\cdot\frac{(n-s_1)!}{s_2!(n-s_1-s_2)!}\dots\frac{(n-s_1-s_2-\dots-s_{r-1})}{s_r!(n-s_1-s_2-\dots-s_r)!} = \sum\frac{n!}{s_1!\,s_2!\dots s_r!}$$

de donde se deduce (3). Las fórmulas (2) y (4) se desprenden, evidentemente, de las (1) y (3), respectivamente.

Estudiemos, por fin, el caso (D), es decir, la clase de problemas respecto de las distribuciones, cuando ambos conjuntos N y R se componen de elementos indistinguibles. Dichos problemas resultan ser más difíciles y la teoría de su resolución aún no está elaborada completamente. La interpretación más conocida del caso dado la representa el problema teórico numérico sobre la partición de los números naturales en sumandos naturales.

Notemos, ante todo, que los problemas de la clase (D) no pueden equipararse con los de las clases anteriores, como se ha hecho más arriba: las conexiones resultan ser mucho más complejas y no se logra hallar una expresión analítica conveniente para obtener el número buscado.

Para calcular el número de particiones referentes a los problemas de tipo (D) sirve por ahora de medio principal el siguiente método recurrente.

Supongamos que un n-conjunto S se divide en k partes no vacías $a_1, a_2, \ldots, a_k$, con la particularidad de que $|a_i| \geq 1$, $i = 1, 2, \ldots, k$. Denotemos con $P_k(n)$ el número de tales particiones. Es obvio que

$$P_i(k) = 1;\ P_k(k) = 1;\ P_k(n) = 0 \text{ para } n < k.$$

$$\sum_{i=1}^{k} a_i = n. \qquad (5)$$

Admitamos que $|a_1| \geq |a_2| \geq \ldots \geq |a_k|$ (reenumeremos, si es necesario, las partes de la partición). Está claro que

$$\sum_{i=1}^{k} (a_i - 1) = n - k. \qquad (6)$$

Se ha obtenido la partición del $(n - k)$-conjunto en partes cuyo número es k (la igualdad tiene lugar si cualquier a_i contiene no menos de 2 elementos). Según la regla de la suma, el número de tales particiones es igual a

$$\sum_{i=1}^{k} P_i(n-k)$$

y, en virtud de la igualdad (6), dicho número es igual al número de particiones del n-conjunto en k partes, es decir, a

$$\sum_{i=1}^{k} P_i(n-k) = P_k(n) \qquad (7)$$

Teniendo en cuenta los valores (5), esta fórmula recurrente permite obtener sucesivamente los valores para $P_k(n)$, reuniéndolos, si es necesario, en la tabla siguiente.

k \ n	Números $P_1(n)$									
	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2		1	1	2	2	3	3	4	4	5
3			1	1	2	3	4	5	7	8
4				1	1	2	3	5	6	9
5					1	1	2	3	5	7
6						1	1	2	3	5
7							1	1	2	3
8								1	1	2
9									1	1
10										1
$P(n) = \sum P_2(n)$	1	2	3	5	7	11	15	22	30	42

Para los valores pequeños de k podemos obtener fórmulas de $P_k(n)$, por ejemplo

$$P_1(n) = 1;\ P_2(2) = 1;\ P_2(1) = 0;\ P_2(n) = P_2(n-2) + P_1(n-2) = P_2(n-2) + 1;$$

de donde

$$P_2(n) = \frac{n}{2},\ \text{si } n \text{ es par};\ P_2(n) = \frac{n-1}{2},\ \text{si } n \text{ es impar.}$$

Pero ya para $k = 3$ las fórmulas se vuelven bastante engorrosas. La investigación del comportamiento de las magnitudes

$$P_k(n) \quad y \quad P(n) = \sum_{k=1}^{n} P_k(n)$$

(el número de toda clase de particiones del número n) para grandes valores de k está ligada con dificultades considerables. Se ha hallado el valor aproximado

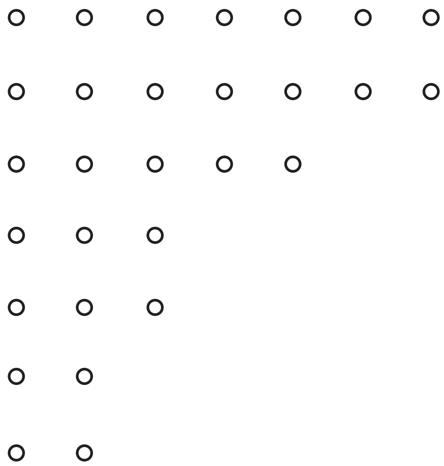
$$P_k(n) = \frac{1}{k!}\binom{n-1}{k-1}$$

el cual en la práctica resulta suficiente. Para $P(n)$ queda determinada una relación recurrente

$$P(n) = P(n-1) + P(n-2) - P(n-5) - P(n-7) + \ldots + (-1)^{k-1}\left(n - \frac{3k^2 \pm k}{2}\right) + \ldots$$

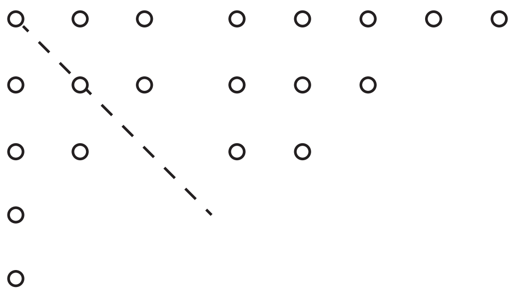
y están calculados los primeros valores sucesivos de esta magnitud. Todas estas cuestiones y otras semejantes las analizaremos al tratar la teoría de las particiones.

Aproximadamente desde la mitad del siglo XX, merced a los esfuerzos de Norman Ferrers, James Sylvester y, más tarde, de Percy Mac Mahon, en la teoría de las particiones entró su interpretación con ayuda de los grafos puntuales. Por ejemplo, la partición 29 = 7 + 7 + 5 + 3 + 3 + 2 + 2 está expuesta en la figura siguiente.

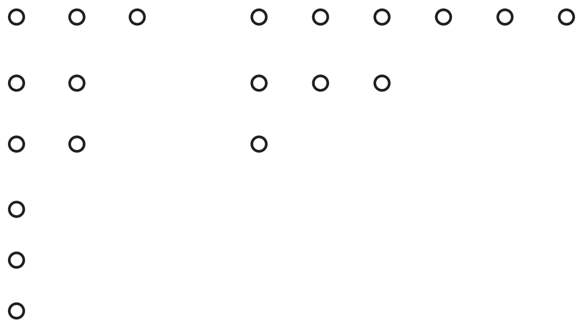


Las partes de las particiones se disponen, como regla, de arriba abajo en el orden de decrecimiento. Del análisis de los correspondientes grafos puntuales (llamados también *grafos de Ferrers*) pueden directamente obtenerse los resultados que se detallan a continuación.

1. El número de particiones de un n-conjunto en las que la mayor parte tiene k elementos es igual al número de partes, es decir, a $P_k(n)$. Este teorema se demuestra por transposición del grafo de Ferrers respecto de la diagonal principal; tales grafos se denominan *conjugados*. Así, por ejemplo, en la figura siguiente



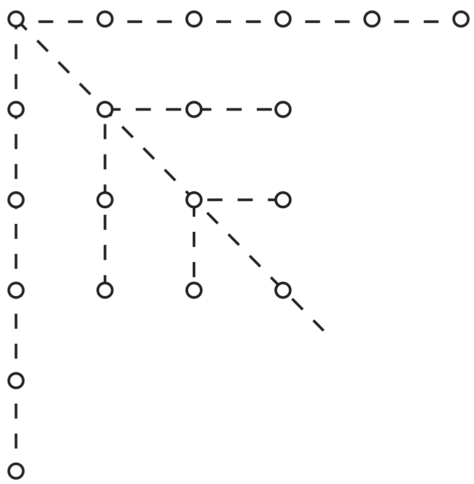
el conjunto de 10 elementos está partido en 5 partes: 10 = 3 + 3 + 2 + 1 + 1; al efectuar la transposición, obtenemos: 10 = 5 + 3 + 2, es decir, la partición del 10-conjunto en la cual la parte mayor contiene 5 elementos (una situación análoga, se muestra en la figura siguiente:



donde 10 = 3 + 2 + 2 + 1 + 1 + 1 antes de la transposición, y 10 = 6 + 3 + 1, después de esta).

2. El número de particiones autoconjugadas de un n-conjunto (una partición se llama *autoconjugada*, si el grafo de Ferrers que le corresponde es simétrico respecto de la diagonal principal) es igual al número de particiones del mismo conjunto en subconjuntos desiguales que se componen de un número impar de elementos. El teorema recíproco es también cierto.

   Así, por ejemplo, a la partición autoconjugada del 20-conjunto 20 = 6 + 4 + 4 + 4 + 1 + 1 (véase la figura siguiente)



   le corresponde biunívocamente la partición 20 = 11 + 5 + 3 + 1, si los puntos de la primera columna dispuestos debajo de la diagonal los trasladamos a la primera fila de la segunda columna, a la segunda fila, etcétera.

3. El número de particiones de un n-conjunto en partes diferentes es igual al número de particiones del mismo conjunto en partes compuestas por un número impar de elementos.

   Sea dada una partición del n-conjunto (por ejemplo, $n = 34$) en componentes impares (34 = 5 + 5 + 5 + 5 + + 3 + 3 + 3 + 1 + 1 + 1 + 1 + 1). Todos los componentes impares iguales se reúnen en grupos (4 de cinco, 3 de tres y 5 unidades) y se escriben los números de sus repeticiones (4, 3 y 5) en el sistema binario ($4 = 2^2$, $5 = 2^2 + 2^0$; $3 = 2^1 + 2^0$). Escribamos la nueva partición, teniendo en cuenta la representación binaria obtenida

$$(34 = 5 \cdot 4 + 3 \cdot 3 + 1 \cdot 5 = 5 \cdot 2^2 + 3 \cdot 2^1 + 3 \cdot 2^0 + 1 \cdot 2^2 + 1 \cdot 2^0 = 20 + 6 + 3 + 4 + 1).$$

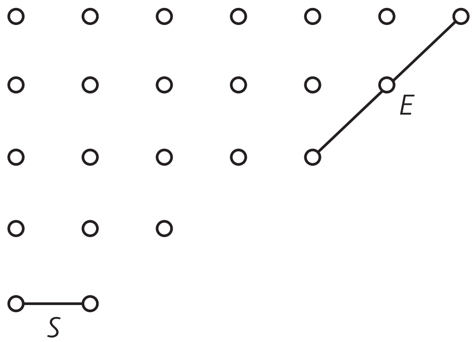
   Este procedimiento es siempre factible, pues cualquier número se escribe de un modo único en el sistema binario. Siempre se puede razonar también del modo inverso.

4. Si Q_n y Q'_n son los números de particiones del n-conjunto en un número par e impar, respectivamente, de las partes desiguales entre sí, entonces

$$Q_n \begin{cases} Q'_n \text{ si } n \neq \dfrac{k}{2}(3k \pm 1) \\ Q'_k + (-1)^k, \text{ si } n = \dfrac{n}{2}(3k \pm 1) \end{cases}$$

donde $k = 1, 2, \ldots$

Supongamos que el grafo de Ferrers de la partición del n-conjunto en partes desiguales tiene la forma mostrada en la figura siguiente.



S es la parte menor de la partición y E, un conjunto de puntos (línea) dispuestos a partir de la parte mayor, bajo un ángulo de 45°. Si $|S| \leq |E|$, entonces S se traslada al conjunto E; en cambio, si $|S| > |E|$, viceversa, E se traslada a la parte menor S. Como resultado de tales traslaciones, tendremos el cambio de la paridad del número de partes desiguales de la partición. A toda partición par se le pone en correspondencia una impar y, además, biunívocamente: $Q_n = Q'_n$. Realice dicha operación con el siguiente ejemplo:

$$7 + 6 + 5 + 3 + 2 \leftrightarrow 8 + 7 + 5 + 3.$$

Sin embargo, la operación no será posible si las líneas S y E se intersecan y

$$|S| = |E| \quad \text{o bien} \quad |S| = |E| + 1.$$

Sea $|E| = k$. Entonces, en el primero de los casos exclusivos tendremos

$$n = k + (k+1) + \ldots + (2k-1)\frac{k}{2}(3k-1);$$

y en el segundo:

$$n = (k+1) + (k+2) + \ldots + 2k = \frac{k}{2}(3k+1);$$

lo que demuestra nuestra afirmación.

11. Sistemas incompatibles

Hasta el momento nos hemos interesado preferentemente por aquellos sistemas de desigualdades que tienen por lo menos una solución (son compatibles). Con respecto a los sistemas incompatibles, su estudio representa, a primera vista, una ocupación nimia. Además, parece poco real que con sistemas semejantes se pueda crear una teoría sobre algo más o menos sustancial. De hecho, la situación es completamente otra: las propiedades de los sistemas incompatibles no solamente presentan interés como tales, sino que nos abren la posibilidad de comprender una serie de hechos importantes. Así, por ejemplo, el teorema fundamental de la programación lineal (teorema de la dualidad, lo veremos en el apartado 15) se demuestra, al fin de cuentas, mediante algunas propiedades de sistemas incompatibles.

Examinemos un sistema arbitrario de desigualdades lineales. Para comodidad de escritura, de momento consideraremos que el número de incógnitas es igual a tres, aunque todo lo expuesto se refiere en igual medida a sistemas con cualquier número de incógnitas. Así, pues, tenemos el siguiente sistema:

$$\left.\begin{array}{l} a_1x+b_1y+c_1z+d_1 \geq 0, \\ a_2x+b_2y+c_2z+d_2 \geq 0, \\ \dots\dots\dots\dots\dots\dots \\ a_mx+b_my+c_mz+d_m \geq 0. \end{array}\right\}. \qquad (1)$$

Multipliquemos los dos miembros de la primera desigualdad de (1) por un número cualquiera no negativo k_1, los dos miembros de la segunda, por un número no negativo k_2 y así sucesivamente; sumemos después las desigualdades obtenidas. Como resultado llegamos a la desigualdad

$$(k_1a_1 + k_2a_2 + \dots + k_ma_m)x + (k_1b_1 + k_2b_2 + \dots + k_mb_m)y + (k_1c_1 + k_2c_2 + \dots + k_mc_m)z +$$
$$+ k_1d_1 + k_2d_2 + \dots + k_md_m \geq 0, \qquad (2)$$

la cual llamaremos *combinación de las desigualdades* (1).

Puede resultar que cierta combinación de las desigualdades (1) representa una desigualdad de la forma

$$0 \cdot x + 0 \cdot y + 0 \cdot z + d \geq 0, \qquad (3)$$

siendo d un número estrictamente negativo (después de dividir por $|d|$ obtenemos la desigualdad $-1 \geq 0$). Está claro que esta desigualdad no podrá satisfacer ningún conjunto de valores de las incógnitas; por lo tanto, en el caso que analizamos el sistema (1) es incompatible (no tiene soluciones). Es interesante que también la afirmación inversa es justa, a saber: si el sistema (1) es incompatible, entonces cierta combinación de sus desigualdades tendrá la forma (3), siendo $d < 0$.

Ahora demostraremos esta afirmación en forma general (o sea, para sistemas con cualquier número de incógnitas), pero antes introduciremos la siguiente definición. Acordemos denominar la desigualdad

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

como *contradictoria*, si no la satisface ningún conjunto de valores de las incógnitas.

Es evidente que cualquier desigualdad contradictoria tiene la forma (3), siendo $d < 0$. La afirmación que debemos demostrar puede ser ahora formulada por el siguiente teorema.

Teorema sobre sistemas incompatibles de desigualdades

Si un sistema de desigualdades lineales es incompatible, entonces cierta combinación de estas desigualdades es una desigualdad contradictoria.

Demostración. La efectuamos por el método de inducción por n-ésimo número de incógnitas para nuestro sistema. Siendo $n = 1$, dicho sistema tiene la forma

$$\left.\begin{array}{l} a_1x+b_1 \geq 0, \\ a_2x+b_2 \geq 0, \\ \dots\dots\dots\dots \\ a_mx+b_m \geq 0. \end{array}\right\} \qquad (4)$$

Se puede considerar que todos los coeficientes $a_1, a_2, \dots, a_m$ son distintos de cero. En efecto, si, por ejemplo, $a_1 = 0$, entonces la primera desigualdad tendrá la forma $0 \cdot x + b_1 \geq 0$ siendo el número b_1 no negativo, esta desigualdad podrá ser suprimida; siendo b_1 un número negativo, la primera desigualdad de nuestro sistema será contradictoria y no tendremos qué demostrar.

Así, consideremos que ningún número $a_1, a_2, \ldots, a_m$ es igual a cero. Es fácil observar que entre estos números forzosamente tiene que haber otros, tanto positivos como negativos; en efecto, si todos los números indicados tuviesen un mismo signo, por ejemplo, fuesen positivos, entonces el sistema (4) se reduciría a la forma

$$\left.\begin{array}{l} x \geq -\dfrac{b_1}{a_1}, \\ x \geq \dfrac{b_2}{a_2}, \\ \cdots\cdots\cdots \\ x \geq \dfrac{b_m}{a_m}. \end{array}\right\}$$

y, por consiguiente, sería compatible.

Supongamos, para precisar, que los primeros k números de $a_1, a_2, \ldots, a_m$ son positivos, mientras que los restantes $m - k$ son negativos. Entonces, el sistema (4) es equivalente al sistema

$$\left.\begin{array}{l} \left.\begin{array}{l} x \geq -\dfrac{b_1}{a_1}, \\ \cdots\cdots\cdots\cdots \\ x \geq \dfrac{b_k}{a_k}. \end{array}\right) \\ \left.\begin{array}{l} x \leq -\dfrac{b_{k+1}}{a_{k+1}}, \\ \cdots\cdots\cdots\cdots \\ x \leq -\dfrac{b_m}{a_m}. \end{array}\right) \end{array}\right\} \qquad (5)$$

Entre los números $-\frac{b_1}{a_1}, \ldots, -\frac{b_k}{a_k}$ elegimos el máximo; supongamos, por ejemplo, que dicho número es $-\frac{b_1}{a_1}$. Entonces, en el sistema (5) las primeras k desigualdades pueden ser reemplazadas por la primera desigualdad. De forma análoga, entre los números $-\frac{b_{k+1}}{a_{k+1}}, \ldots, -\frac{b_m}{a_m}$ elegimos el mínimo; entonces, las restantes $m - k$ desigualdades del sistema (5) pueden ser reemplazadas por la última desigualdad.

Así, pues, el sistema (4) es equivalente a un sistema de dos desigualdades:

$$x \geq -\frac{b_0}{a_0},$$

$$x \leq -\frac{b_m}{a_m},$$

y su incompatibilidad significa que

$$-\frac{b_1}{a_1} > -\frac{b_m}{a_m^n}. \qquad (6)$$

De (6) se deduce que

$$b_m a_1 - b_1 a_m < 0 \qquad (7)$$

(hay que tener en cuenta que $a_1 > 0$ y $a_m < 0$). Si multiplicamos ahora la primera desigualdad de (4) por un número positivo $-a_m$ y la última, por un número positivo a_1, y después sumamos, tendremos la desigualdad

$$0 \cdot x + (b_m \cdot a_1 - b_1 a_m) \geq 0, \qquad (8)$$

la cual, según (7), es contradictoria. Así, pues, para sistemas con una incógnita el teorema es válido.

Supongamos ahora que la afirmación del teorema es válida para sistemas de desigualdades con $n - 1$ incógnitas y basándonos en esta suposición demostraremos que dicha afirmación es válida para el caso de n incógnitas.

Entonces, sea dado un sistema incompatible de desigualdades lineales con n incógnitas $x_1, x_2, \ldots, x_n$; ateniéndonos a las designaciones dadas en el apartado 8 (véase *Leñitas Geométricas* Nº 11, 5ª época), designamos este sistema por "sistema (S)". Formamos para él un sistema asociado (S'): este último será incompatible seguidamente después de (S). Como en el sistema (S') el número de incógnitas es igual a $n - 1$, con respecto a este sistema puede aplicarse la suposición de inducción. Esto significa que cierta combinación de desigualdades del sistema (S') representa una desigualdad contradictoria.

Pero no es difícil ver que cada desigualdad del sistema (S') es una combinación de desigualdades del sistema (S): en efecto, si simplemente sumamos las desigualdades $P_\alpha \geq x_n$ y $x_n \geq Q_\beta$ del sistema (S), obtenemos $P_\alpha + x_n \geq x_n + Q_\beta$, o bien, $P_\alpha \geq Q_\beta$, es decir, una de las desigualdades del sistema (S') [véase el apartado 8 ya mencionado, sistemas (T) y (S')]. De esto, a su vez, se deduce que cierta combinación de desigualdades del sistema inicial (S) es una desigualdad contradictoria. El teorema queda demostrado.

El teorema sobre sistemas incompatibles de desigualdades lineales es solamente una de las manifestaciones de la profunda analogía existente entre las propiedades de sistemas de desigualdades lineales y de sistemas de ecuaciones lineales. Si probamos cambiar, en la formulación del teorema, la palabra "desigualdad" por "ecuación" tendremos la siguiente afirmación: **si un sistema de ecuaciones lineales es incompatible, entonces cierta combinación de dichas ecuaciones es una ecuación contradictoria**.

Resulta que esta afirmación también es válida. Con una formulación algo distinta, dicha afirmación se llama *teorema de Kronecker-Capelli* y se demuestra en el curso de álgebra lineal (así se llama una parte de la matemática en la cual se estudian las operaciones lineales, o sea, operaciones semejantes a la adición de puntos y multiplicación de puntos por un número en un espacio n-dimensional). Además, para que lo dicho sea comprendido correctamente, es preciso hacer algunas precisiones respecto del concepto de "combinación".

Una combinación de ecuaciones se confecciona de la misma forma que una combinación de desigualdades, con la única diferencia de que se permite multiplicar estas ecuaciones por números cualesquiera y no solamente por números no negativos. Lo mismo que para desigualdades, la ecuación que no tiene soluciones se llama *contradictoria*. No será difícil demostrar que las ecuaciones contradictorias forzosamente deben presentar la forma

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n + b \geq 0 \qquad (9)$$

siendo b un número diferente de cero (después de dividir los dos miembros por b obtenemos la "ecuación" $1 = 0$).

Sumamente importante es un caso particular del teorema de sistemas incompatibles de desigualdades, precisamente cuando el sistema dado contiene las desigualdades

$$x_1 \geq 0, \qquad x_2 \geq 0, \qquad \ldots, x_n \geq 0. \qquad (10)$$

Designando por (S) la parte restante del sistema, se puede decir que el problema consiste en hallar todas las soluciones no negativas del sistema (S) [o sea, soluciones que satisfagan las condiciones (10)]. Si este problema no tiene soluciones, entonces, conforme al teorema que acabamos de demostrar, cierta combinación de desigualdades del sistema (S),

$$a_1x_1 + a_2x_2 + \ldots + a_nx_n + a \geq 0, \qquad (11)$$

en suma, con cierta combinación de las desigualdades (10)

$$k_1x_1 + k_2x_2 + \ldots + k_nx_n \geq 0 \ (k_1, k_2, \ldots, k_n \text{ no son negativos}).$$

dará una desigualdad contradictoria

$$0 \cdot x_1 + 0 \cdot x_2 + \ldots + 0 \cdot x_n + c \geq 0,$$

siendo c un número negativo. Por consiguiente,

$$a_1 = -k_1 \leq 0, \qquad a_2 = -k_2 \leq 0, \qquad \ldots, \qquad a_n = -k_n \leq 0, \ a < 0.$$

Formulemos el resultado obtenido como una afirmación particular.

Corolario del teorema sobre sistemas incompatibles:

Si un sistema de desigualdades carece de soluciones no negativas, entonces cierta combinación de dichas desigualdades es una desigualdad de la forma (11), en la que todos los coeficientes $a_1, a_2, \ldots, a_n \le 0$ y el término independiente $a < 0$.

12. Conos poliedros convexos recíprocamente conjugados

En el apartado 4 (véase *Leñitas Geométricas* Nº 5, 5ª época) hemos expuesto algunos hechos referentes a conos poliedros convexos; en el presente apartado examinaremos hechos complementarios de la teoría de los conos convexos.

Cada cono poliedro convexo con el vértice en el origen de las coordenadas en un espacio tridimensional, conforme se demuestra anteriormente, es una región de soluciones de cierto sistema de desigualdades homogéneas lineales con tres incógnitas:

$$\left.\begin{array}{r} a_1x+b_1y+c_1z \ge 0, \\ a_2x+b_2y+c_2z \ge 0, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ a_mx+b_my+c_mz \ge 0. \end{array}\right\} \qquad (1)$$

Supongamos que, además de este sistema, se da una desigualdad aparte:

$$ax + by + cz \ge 0. \qquad (2)$$

Consideramos que la desigualdad (2) es consecuencia del sistema (1), si cualquier conjunto de valores de las incógnitas x, y, z que satisface el sistema (1) satisface también la desigualdad (2).

Naturalmente, cualquier desigualdad, siendo una combinación lineal de las desigualdades (1), es consecuencia del sistema (1). Pero, ¿es válido lo inverso? Resulta que sí.

Teorema 1. La desigualdad homogénea (2), siendo consecuencia del sistema homogéneo (1), puede ser representada como una combinación lineal de las desigualdades (1).

Demostración. Para comodidad de escritura, en lo sucesivo designaremos los primeros miembros de la primera, segunda, .. ., m-ésima desigualdades del sistema (1) por P_1, P_2, ..., P_m, respectivamente, y el primer miembro de la desigualdad (2), por P. Así, pues, tenemos dado el sistema

$$\left.\begin{array}{r} P_1 \ge 0, \\ P_2 \ge 0, \\ \ldots\ldots \\ P_m \ge 0. \end{array}\right\} \qquad (1)$$

y también la desigualdad

$$P \ge 0, \qquad (2)$$

que es consecuencia del sistema. Es preciso demostrar que esta desigualdad puede ser expresada como una combinación lineal de las desigualdades (1).

Puesto que la desigualdad $P \ge 0$ es una consecuencia del sistema (1), entonces la desigualdad $P \le -1$, por el contrario, es incompatible con este sistema; es decir, el sistema mixto

$$\left.\begin{array}{r} P_1 \ge 0, \\ P_2 \ge 0, \\ \ldots\ldots\ldots \\ P_m \ge 0, \\ -P-1 \ge 0. \end{array}\right\} \qquad (3)$$

es incompatible.

Conforme al teorema de sistemas incompatibles, cierta combinación lineal de las desigualdades (3) es una desigualdad contradictoria. En otras palabras, existen unos números no negativos $k_1, k_2, \ldots, k_m, k_{m+1}$ tales que la desigualdad

$$k_1P_1 + k_2P_2 + \ldots, k_mP_m + k_{m+1}(-P - 1) \geq 0$$

(después de una reducción de los términos semejantes) tendrá la forma

$$0 \cdot x + 0 \cdot y + \ldots + 0 \cdot z + d \geq 0,$$

donde d es un número negativo. Por consiguiente,

$$k_1P_1 + k_2P_2 + \ldots + k_mP_m - k_{m+1}P = 0 \cdot x + 0 \cdot y + 0 \cdot z,$$

y $-k_{m+1} = d$ es un numero negativo.

De ello resulta que

$$P = \frac{k_1}{k_{m-1}}P_1 + \ldots + \frac{k_m}{k_{m+1}}P_m;$$

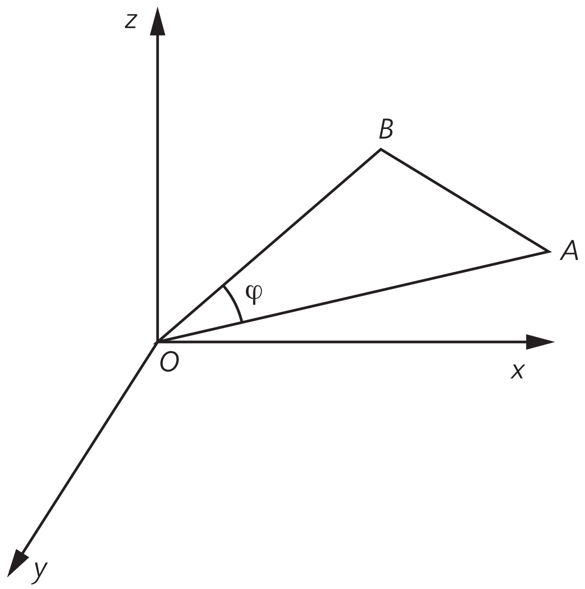
además, los factores de $P_1, \ldots, P_m$ no son negativos. Por lo visto, esto significa que la desigualdad (2) es una combinación lineal de las desigualdades (1); esto era preciso demostrar.

El teorema demostrado presenta interés como tal, pero más curioso es su contenido geométrico. Para darlo a conocer, tendremos que valernos de algunos hechos de la geometría analítica; por cierto, estos son tan elementales como los pocos hechos de la geometría analítica que venimos utilizando hasta aquí.

Sean

$$A(x_A, y_A, z_A), B(x_B, y_B, z_B)$$

dos puntos en el espacio diferentes del origen O de las coordenadas. Apliquemos con relación al triángulo OAB (figura siguiente)



el así llamado *teorema de los cosenos*, conforme al cual el cuadrado de cualquier lado de un triángulo es igual a la suma de los cuadrados de los otros dos lados menos el producto doble de estos lados por el coseno del ángulo entre ellos. En el caso dado:

$$\overline{AB}^2 = \overline{OA}^2 + \overline{OB}^2 - 2 \cdot \overline{OA} \cdot \overline{OB} \cdot \cos\varphi, \qquad (4)$$

siendo φ el ángulo entre los segmentos $\overline{OA}$ y $\overline{OB}$. Pero

$$\overline{OA}^2 = x_A^2 + y_A^2 + z_A^2$$

$$\overline{OB}^2 = x_B^2 + y_B^2 + z_B^2$$

$$\overline{AB}^2 = (x_A - x_B)^2 + (y_A - y_B)^2 + (z_A - z_B)^2.$$

O sea que después de reducir los términos semejantes, obtenemos

$$-2(x_A x_B + y_A y_B + z_A z_B) = -2 \cdot OA \cdot OB - \cos\varphi. \qquad (5)$$

El ángulo φ será no agudo en aquel y solo en aquel caso, cuando $\cos\varphi \leq 0$. Por lo tanto, también de (5) se deduce que el ángulo entre los segmentos $\overline{OA}$ y $\overline{OB}$ será no agudo en aquel y solo en aquel caso, cuando

$$x_A x_B + y_A y_B + z_A z_B \leq 0.$$

Acordaremos en adelante designar la expresión que forma el primer miembro de esta desigualdad, brevemente, por (A, B):

$$(A, B) = x_A x_B + y_A y_B + z_A z_B.$$

Entonces, lo dicho anteriormente se puede parafrasear de la siguiente forma.

El ángulo entre los segmentos $\overline{OA}$ y $\overline{OB}$ será no agudo en aquel y solo en aquel caso, cuando

$$(A, B) \leq 0.$$

Precisamente, estos son los conocimientos de la geometría analítica que necesitamos en lo sucesivo. Para finalizar señalaremos una propiedad de la magnitud (A, B):

$$(k_1 A_1 + k_2 A_2, B) = k_1(A_1, B) + k_2(A_2, B), \qquad (6)$$

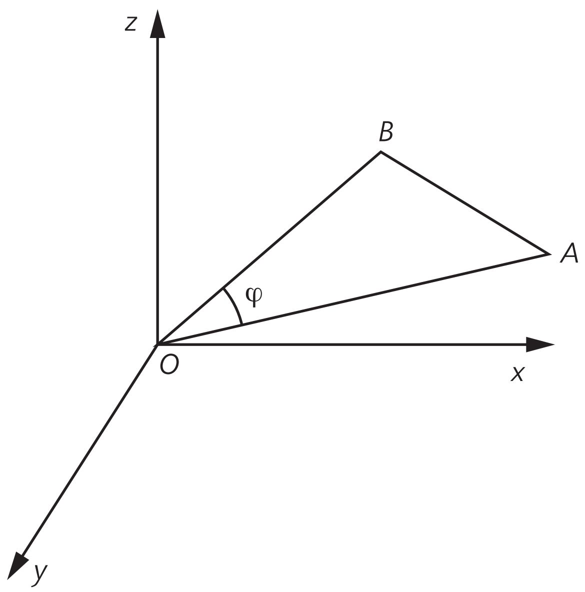
sean cuales sean los números k_1 y k_2. La demostración es casi evidente: puesto que el punto $k_1 A_1 + k_2 A_2$ tiene las coordenadas

$$k_1 x_{A_1} + k_2 x_{A_2}, k_1 y_{A_1} + k_2 y_{A_2}, k_1 z_{A_1} + k_2 z_{A_2}$$

entonces

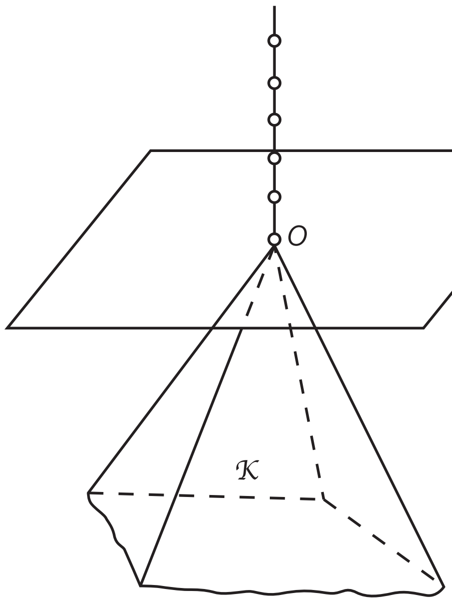
$$(k_1 A_1 + k_2 A_2, B) = \left(k_1 x_{A_1} + k_2 x_{A_2}\right) x_B + \left(k_1 y_{A_1} + k_2 y_{A_2}\right) y_B + \left(k_1 z_{A_1} + k_2 z_{A_2}\right) z_B =$$

$$= k_1\left(x_{A_1} x_B + y_{A_1} y_B + z_{A_1} z_B\right) + k_2\left(x_{A_2} x_B + y_{A_2} y_B + z_{A_2} z_B\right) = k_1(A_1, B) + k_2(A_2, B).$$



Por fin nos dirigimos al tema principal de este apartado, o sea, a los conos poliedros convexos situados en el espacio. Según la definición dada en el apartado 4 ya mencionado, cono poliedro convexo con el vértice en el punto S se llama la intersección de un número finito de semiespacios, cuyos planos confines pasan por S. El ejemplo más típico de cono poliedro convexo, como es sabido, es una pirámide convexa ilimitada. En lo que queda de este apartado, se supone que el vértice S coincide con el origen de las coordenadas O.

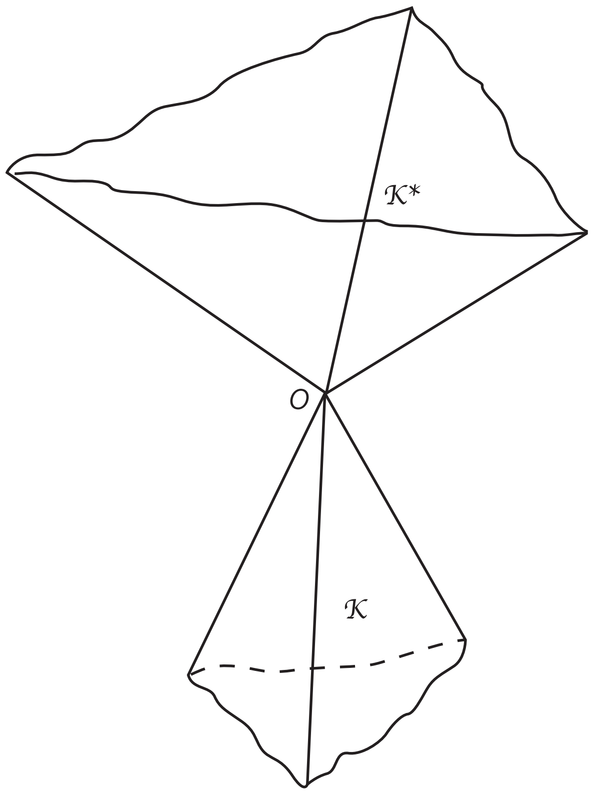
Supongamos que el punto B, distinto del origen de las coordenadas, tiene la siguiente propiedad: el segmento $\overline{OB}$ forma un ángulo no agudo con cualquiera de los segmentos $\overline{OA}$, siendo A un punto arbitrario del cono considerado $\mathcal{K}$. Un punto tal B siempre puede hallarse: para ello es suficiente, por ejemplo, trazar un plano π a través del vértice del cono, de tal forma que todo el cono se sitúe en uno de los dos semiespacios determinados por este plano, como se muestra en la figura siguiente:



Entonces, una perpendicular al plano π, trazada en el otro semiespacio, estará compuesta totalmente por puntos B de la forma requerida.

Examinemos el conjunto de todos los puntos B que poseen la propiedad antes indicada; completemos dicho conjunto con un punto más –el origen de las coordenadas– y designemos el conjunto obtenido por $\mathcal{K}^*$. Demostraremos, primeramente, que es válido el siguiente lema.

Lema. $\mathcal{K}^*$ es nuevamente un cono convexo poliedro, como se sugiere en la ilustración siguiente:



Demostración. Conforme al teorema 2 del apartado 4, cualquier cono convexo poliedro $\mathcal{K}$ es un conjunto de la forma $(A_1, A_2, \ldots, A_m)$ (en el teorema indicado las designaciones son algo distintas: en lugar de $A_1, A_2, \ldots, A_m$, en él se da $B_1, B_2, \ldots, B_q$).

Esto significa que cualquier punto $A \in \mathcal{K}$ se expresa de la forma

$$A = t_1A_1 + t_2A_2 + \ldots + t_mA_m,$$

siendo $t_1, t_2, \ldots, t_m$ números no negativos. Si el punto B pertenece al conjunto $\mathcal{K}^*$, entonces el ángulo entre el segmento $\overline{OB}$ y cualquiera de los segmentos $\overline{OA}$ ($A \in \mathcal{K}$) no es agudo, es decir:

$$(A, B) \leq 0, \quad \text{para todos } A \in \mathcal{K}.$$

Puesto que

$$(A, B) = t_1(A_1, B) + t_2(A_2, B) + \ldots + t_m(A_m, B)$$

[véase la fórmula (6)], entonces tenemos en este caso

$$t_1(A_1, B) + t_2(A_2, B) + \ldots + t_m(A_m, B) \leq 0, \tag{7}$$

sean cuales sean los números no negativos $t_1, t_2, \ldots, t_m$. En particular,

$$(A_1, B) \leq 0, \quad (A_2, B) \leq 0, \quad \ldots, \quad (A_m, B) \leq 0. \tag{8}$$

Y también a la inversa: si son válidas las desigualdades (8), entonces también será válida (con cualesquiera no negativos $t_1, t_2, \ldots, t_m$) la desigualdad (7), es decir, el punto $B \in \mathcal{K}^*$. Así, pues, el cumplimiento de las desigualdades (8) es lo necesario y suficiente para que el punto B pertenezca a $\mathcal{K}^*$.

Designemos las coordenadas del punto A_i por $a_i, b_i, \ldots, c_i$, ($i = 1, 2, \ldots, m$) y las coordenadas del punto B, por x, y, z. Entonces, las condiciones (8) se escribirán de la siguiente forma

$$\left.\begin{aligned} a_1x + b_1y + c_1z &\leq 0, \\ a_2x + b_2y + c_2z &\leq 0, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ a_mx + b_my + c_mz &\leq 0. \end{aligned}\right\} \tag{9}$$

Para que B pertenezca al conjunto $\mathcal{K}^*$ es necesario y suficiente que sus coordenadas x, y, z satisfagan todas las desigualdades (9). Es decir, que $\mathcal{K}^*$ sea la región de soluciones del sistema (9).

Como el sistema (9) es homogéneo, la región de sus soluciones es cierto cono poliedro convexo, situado en el espacio. Por consiguiente, $\mathcal{K}^*$ es un cono poliedro convexo; esto queríamos demostrar.

Así pues, hemos aprendido a relacionar cada cono poliedro convexo $\mathcal{K}$ con otro cono poliedro convexo $\mathcal{K}^*$. Este último está compuesto por todos los puntos B en los cuales el segmento $\overline{OB}$ forma ángulos no agudos con cualquiera de los segmentos $\overline{OA}$, en los que $A \in \mathcal{K}$. *Diremos que el cono $\mathcal{K}^*$ es conjugado con $\mathcal{K}$.*

Surge una pregunta lógica: ¿qué representa un cono conjugado con $\mathcal{K}^*$?, o sea, ¿qué se puede decir con respecto al conjunto $(\mathcal{K}^*)^*$? De la definición dada al cono $\mathcal{K}^*$ se desprende directamente que el conjunto $(\mathcal{K}^*)^*$ debe contener el conjunto inicial $\mathcal{K}$ (¿por qué?).

Pero no está del todo claro: ¿coincidirán estos dos conjuntos o no? Más aún, intentando comprobar este hecho mediante razonamientos geométricos, nos convencemos de que esto no es tan simple. Sea como fuere, para demostrar la coincidencia entre $(\mathcal{K}^*)^*$ y $\mathcal{K}$ nosotros elegiremos el camino algebraico. Basado de hecho en el teorema 1, como veremos por la demostración dada a continuación, el contenido geométrico del teorema 1, para el caso de un sistema con tres incógnitas, consiste precisamente en la igualdad $(\mathcal{K}^*)^* = \mathcal{K}$.

Así pues, demostremos el siguiente teorema.

Teorema 2. Sea $\mathcal{K}$ un cono poliedro convexo. Entonces, el conjunto $(\mathcal{K}^*)^*$ coincide con $\mathcal{K}$.

Lo mismo se puede decir de otra forma quizás más clara. Designemos el cono $\mathcal{K}$ por $\mathcal{K}_1$ y el cono $\mathcal{K}^*$ por $\mathcal{K}_2$. Está claro que el teorema afirma lo siguiente:

$$\text{si } \mathcal{K}_1^* = \mathcal{K}_2, \text{ entonces } \mathcal{K}_2^* = \mathcal{K}_1,$$

o bien, si un cono es conjugado de otro, entonces el segundo también es el conjugado del primero, o sea, la relación de conjugación es recíproca.

Demostración. Sea $C(a, b, c)$ un punto arbitrario del conjunto $(\mathcal{K}^*)^*$. Para cada punto $B(x, y, z) \in \mathcal{K}$ debe cumplirse la desigualdad $(B, C) \leq 0$, es decir,

$$ax + by + cz \leq 0. \tag{10}$$

Pero la propia pertenencia del punto B al conjunto $\mathcal{K}^*$ significa, como decíamos antes, el cumplimiento de la desigualdad (9). De tal forma, cada solución (x, y, z) del sistema (9) deberá satisfacer también la desigualdad (10). En otras palabras, la desigualdad (10) es consecuencia del sistema (9).

Por el teorema 1, demostrado anteriormente, esto es posible solamente en el caso cuando la desigualdad (10) es una combinación lineal de la desigualdad (9), es decir, cuando

$$(a, b, c) = t_1(a_1, b_1, c_1) + t_2(a_2, b_2, c_2) + \dots + t_m(a_m, b_m, c_m),$$

siendo $t_1, t_2, \dots, t_m$ ciertos números no negativos.

Mas esta última igualdad significa que

$$C = t_1A_1 + t_2A_2 + \dots + t_mA_m,$$

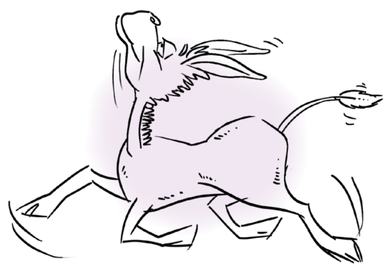
o sea, el punto C pertenece al cono $\mathcal{K}$. Asi pues, cualquier punto C perteneciente a $(\mathcal{K}^*)^*$ pertenece también a $\mathcal{K}$. Pero antes hemos señalado lo inverso, o sea que $\mathcal{K}$ pertenece a $(\mathcal{K}^*)^*$. Por consiguiente $\mathcal{K} = (\mathcal{K}^*)^*$. El teorema queda demostrado.

Utilizando los hechos ya conocidos, ahora no será difícil demostrar el teorema 2 dado en el apartado 7 (véase *Leñitas Geométricas* Nº 9, 5ª época), relativo a la estructura de los conos poliedros convexos. Recordaremos el contenido de este teorema: cualquier conjunto de la forma $(A_1, A_2, \dots, A_m)$ es un cono poliedro convexo.

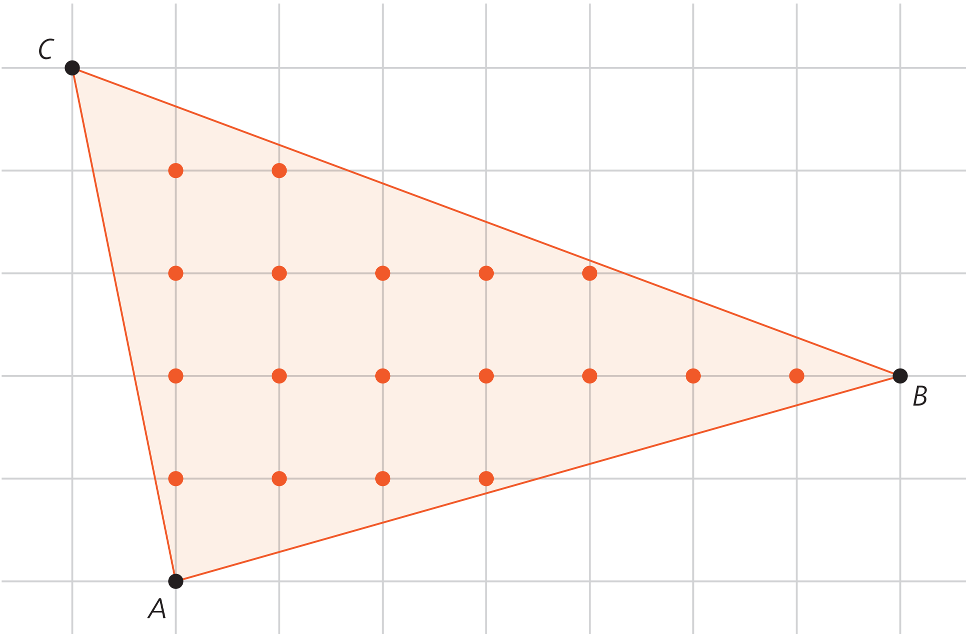
Para realizar la demostración examinaremos el conjunto: $\mathcal{K} = (A_1, A_2, \dots, A_m)^*$, compuesto (en conformidad con la definición) por aquellos puntos B en los cuales el vector $\overline{OB}$ forma ángulos no agudos con todos los vectores $\overline{OA}$ siendo A cualquier punto de $(A_1, A_2, \dots, A_m)$. El conjunto se determina por las desigualdades (8) y, por lo tanto, es un cono poliedro convexo. Como cualquier otro cono poliedro convexo, es cono conjugado con relación a cierto cono $\mathcal{L}$ [precisamente, con relación a $(\mathcal{K}^*)$]. Pero, de acuerdo con lo demostrado anteriormente (véase el final del apartado 9 en *Leñitas Geométricas* Nº 11, 5ª época,), el cono $\mathcal{L}$ se expresa de la forma $(B_1, B_2, \dots, B_s)$. O sea, $\mathcal{K} = \mathcal{L}^*$ o, de otra forma,

$$(A_1, A_2, \dots, A_m)^* = (B_1, B_2, \dots, B_s)^* \qquad (11)$$

Basándose en el teorema 1 del presente apartado, de la relación (11) es fácil deducir lo siguiente: cada uno de los puntos A_i pertenece al conjunto $(B_1, B_2, \dots, B_s)$ y cada uno de los puntos B_j a $(A_1, A_2, \dots, A_m)$. De ello se deduce la coincidencia de los conjuntos $(A_1, A_2, \dots, A_m)$ y $(B_1, B_2, \dots, B_s)$. Por lo tanto, el conjunto $(A_1, A_2, \dots, A_m)$ es el cono $\mathcal{L}$. El teorema está demostrado.

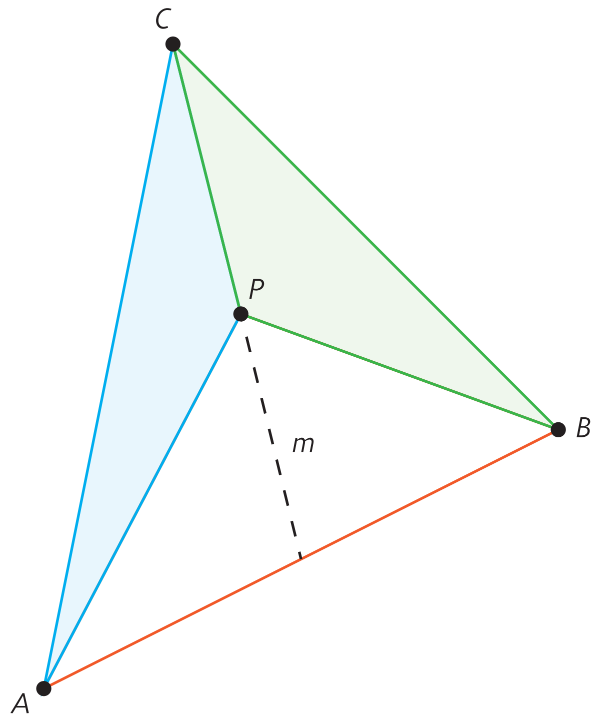


Mostrar que cualquiera sea *P*, un punto de la cuadrícula en el interior del triángulo *ABC*, los triángulos *ABP*, *BCP* y *CAP* tienen áreas distintas.



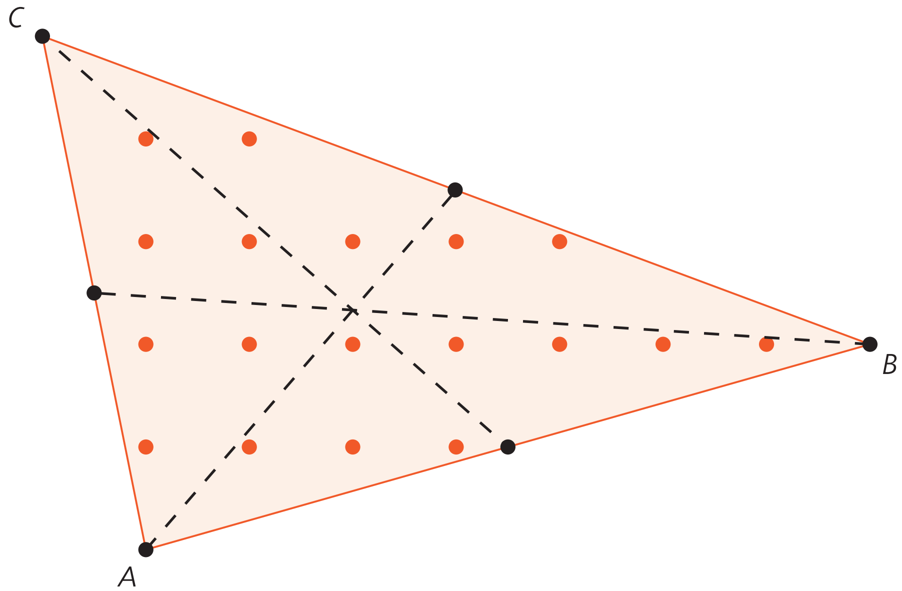
Solución

El lugar geométrico de los puntos *P* en el interior del triángulo *ABC* tales que los triángulos *ABP* y *BCP* tengan igual área son exactamente los puntos en la mediatriz del triángulo que parte del vértice *C*. (Ver Problema semanal 31 de *Leñitas Geométricas* 4ª época, 2022).



Luego, si un punto *P* no está sobre alguna mediana del triángulo, las áreas de los triángulos *ABP*, *BCP* y *CAP* serán todas de distinto valor.

Trazando las medianas en el triángulo del problema, se puede observar que ningún punto de la cuadrícula, en el interior del triángulo, está sobre alguna mediatriz del triángulo.

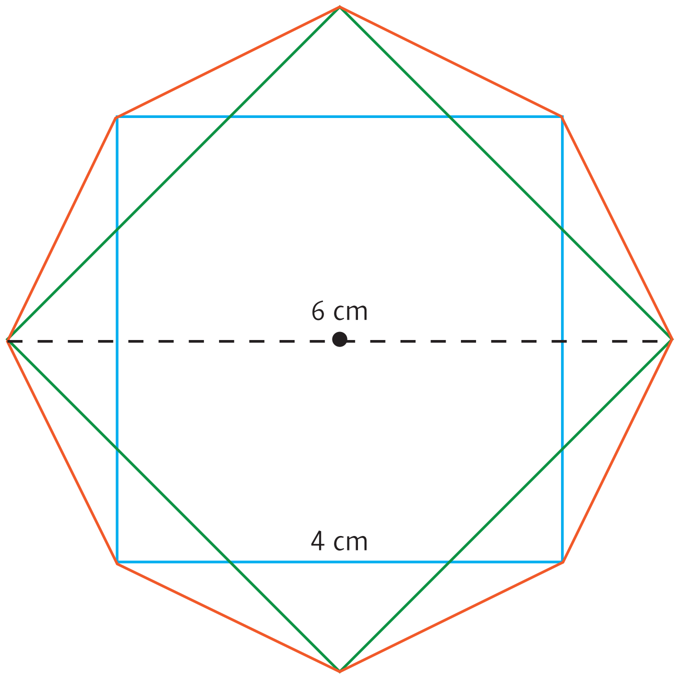


PARA RESOLVER

con imaginación e inteligencia

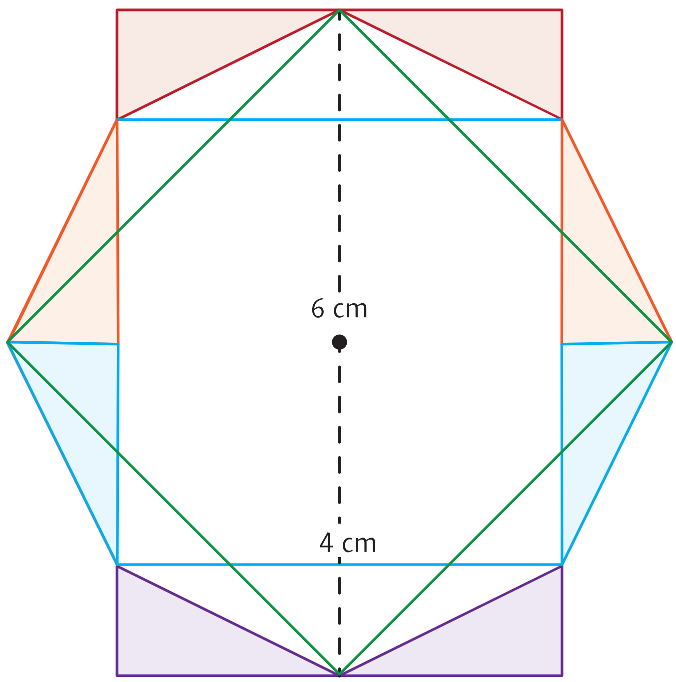
Respuesta del Nº 29

En la figura, la diagonal de un cuadrado mide 6 cm y es paralela a un lado del otro cuadrado, que mide 4 cm. Hallar el área del octógono que tiene por vértices a los vértices de los dos cuadrados.



Solución

Es claro que todos los triángulos sombreados en la figura son congruentes.



Se puede formar un rectángulo de 4 cm × 6 cm, recortando del octógono los triángulos naranjas y celestes y ubicándolos sobre los rojos y azules, respectivamente. El área del octógono es 24 cm^2.