Probabilidades y Estadística

Jorge Soto Andrade

0. Introducción

En esta monografía nos proponemos tratar de manera integrada contenidos y metodologías de enseñanza. En lugar de una árida exposición de contenidos matemáticos *per se* o de un discurso didáctico que gire en el vacío, por falta de encarnación en contenidos y situaciones problemáticas específicas, queremos ejemplificar y discutir interactivamente metodologías, al momento de tratar contenidos específicos e introducir nuevas ideas y métodos matemáticos. Intentamos ejemplificar, con casos concretos, cómo el arte de enseñar consiste, en particular, en presentar y motivar conceptos y métodos matemáticos en forma contextualizada e inmediata, cortocircuitando al máximo los prerrequisitos y preliminares.

Así el lector podrá tener una experiencia directa de lo que significa el abordaje experimental de un problema probabilista, de los méritos de una simulación, o de la eficacia de un abordaje metafórico, entre otros. También esperamos mostrar cuán lejos se puede llegar, en la resolución de problemas estadísticos y probabilistas, procediendo de manera "artesanal", con un mínimo de herramientas, armados sólo de nuestro sentido común y una sana actitud experimental.

Las metodologías que queremos ilustrar, tienen particularmente en cuenta los avances recientes en ciencias cognitivas, que revelan el rol fundamental de la actividad sensorimotora en la construcción de las ideas y conceptos matemáticos, la importancia del inconsciente cognitivo en el aprendizaje, así como la relevancia del pensamiento metafórico y analógico en matemáticas. De hecho, a menudo podemos facilitar el aprendizaje gracias a metáforas que crecen en el terreno fértil de las experiencias sensorimotoras de la primera infancia. Así por ejemplo, más adelante, en el caso de las variables aleatorias o de las colecciones de datos, podremos abordar la esperanza o la media apoyándonos en la experiencia previa del juego del balancín, y la desviación estándar, apoyándonos en la experiencia previa del juego del giro sobre si mismo…

1. Paseos al azar: un abordaje pedestre al cálculo de probabilidades.

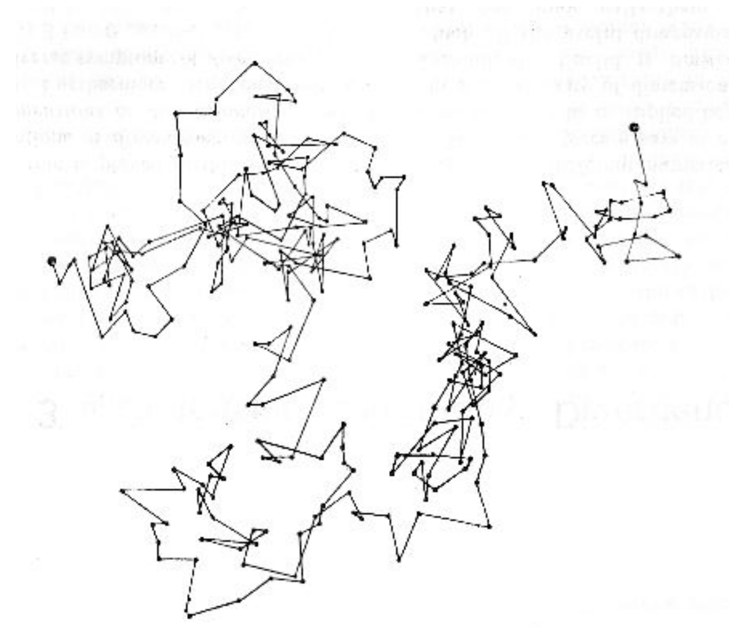
Esta primera sección es una invitación a descubrir la relevancia y ubicuidad de los paseos al azar en distintas áreas de la matemática, en otras ciencias y en la vida cotidiana.

Haciendo saltar una pulga por los vértices de un triángulo, o dejando caminar una hormiga por los alambres de una malla, podemos aprender mucho de geometría y álgebra, además de probabilidades y combinatoria. En particular, podemos redescubrir el teorema del binomio y el triángulo de Pascal.

Comenzamos con un ejemplo histórico en el contexto biológico.

1.1. ¿Por qué bailan los granos de polen?

En 1827 un biólogo inglés, llamado Robert Brown, notó algo que lo dejó perplejo: los granos de polen que estaban en una suspensión acuosa bajo el lente de su microscopio bailaban en todas direcciones, siguiendo caminos zigzageantes como el de la figura de más abajo (para applets de simulación en el computador, ver es.com/piratord//browni/Brown.html, http://xanadu.math.utah.edu/java/brownianmotion/1/). Pensó primero que los granitos bailaban por impulso propio, pero al probar con otros granos de polen que habían estado almacenados durante un siglo, ¡vio que bailaban igual!



El primero propuso una buena explicación del baile de los granos de polen, que pasó a llamarse desde entonces "movimiento Browniano", fue el físico francés Desaulx en 1877, quien dijo: "En mi opinión, este fenómeno es simplemente un resultado de la agitación térmica de las moléculas de agua".

Efectivamente, cada partícula en suspensión en una solución acuosa, que no está tan quieta como parece, es bombardeada aleatoriamente, sin cesar, desde todos lados por las moléculas de agua. Si la partícula es suficientemente pequeña, estos impactos la propulsarán en una dirección, y en seguida en otra, en forma errática e imprevisible.:molecular Estos pequeños y aleatorios saltos generan entonces el movimiento browniano. Vale la pena notar que esta invisible agitación de las moléculas de agua se manifiesta también en la forma en que se "difunde" una cucharada de tinta en un vaso con agua, en lugar de irse directamente al fondo… ¿Lo habían notado alguna vez?

La primera teoría matemática del movimiento browniano fue propuesta por Albert Einstein en 1905. Por este trabajo, recibió el premio Nobel de Física (ver: galileo.phys.virginia.edu/classes/109N/more_stuff/Applets/brownian/brownian.html)

Hoy en día, las aplicaciones de los modelos matemáticos del movimiento browniano son ubicuas, en el tratamiento de imágenes médicas, la robótica, economía (estudio de mercados bursátiles), construcciones fractales, simulaciones gráficas, teoría de colas, manufacturas y servicios, ecología, toma de decisiones, propagación de aerosoles, etc.

Matemáticamente, podemos mirar el movimiento browniano como un paseo al azar en que la partícula paseante da en cualquier instante un salto de dirección y magnitud arbitraria. En lo que sigue, estaremos estudiando, a mano, un "análogo bebé" discreto de este tipo de movimiento, en que nuestra partícula da saltos de la misma magnitud a sitios prescritos de su espacio ambiente (secciones 1 y 2).. En particular, consideraremos el caso de un espacio con sólo tres sitios, y veremos que incluso este fenómeno matemático permite modelar sistemas interesantes de la vida cotidian

1.2. Paseos al azar: Memorias de una pulga

Uno podría introducir el cálculo de probabilidades “paseando al azar”, es decir interesándose en el devenir de un ser u objeto que se pasea al azar por algún espacio o estructura. Uno de los ejemplos más sencillos imaginables de paseo al azar es el siguiente, que proponemos **como una actividad dirigida al alumno:**

MODULO DE TRABAJO PARA EL ALUMNO: El paseo de la pulga…

Una pulga se pasea alegremente por los vértices de un triángulo equilátero, saltando cada vez desde un vértice a cualquiera de los otros dos, con igual probabilidad.
¿Si la pulga reside inicialmente en el vértice superior del triángulo, dónde estará después de un salto, dos saltos, tres saltos, ...10 saltos, 100 saltos, muchísimos saltos? ¿A qué vértice(s) conviene apostar a la larga?

¿Quieres experimentar un poco para ver que sucede?

Cada compañero tuyo podría, en lugar de cazar una pulga y amaestrarla, simularla con ayuda de una moneda. Podemos tomar nota en seguida de dónde está la pulga de cada alumno después de 3, 4, 5, ... 10, 20, 30... saltos. ¿Qué observas, si tabulas y graficas?

Variante “determinista” de la pulga aleatoria:

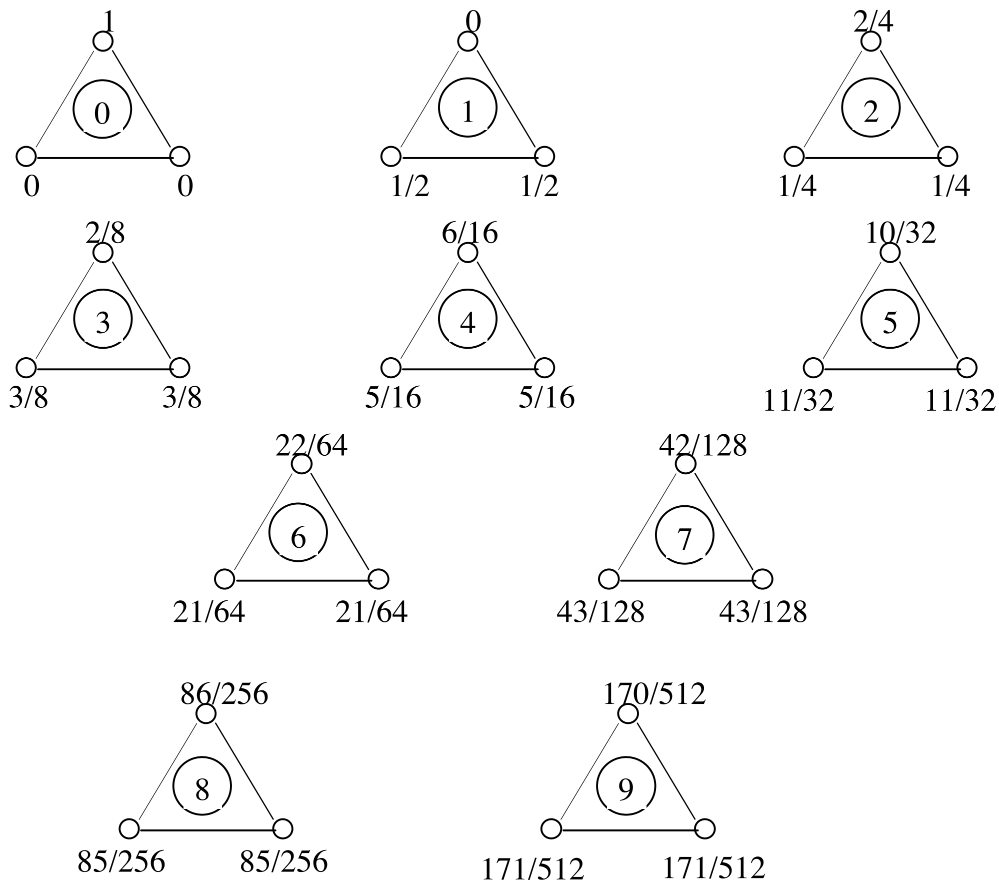
Imagina que en lugar de la pulga saltarina y aleatoria te encuentras con la siguiente situación, en la que no se ve azar, sino que un futuro totalmente determinado:

Una partícula, de masa 1, se encuentra en uno de los vértices de nuestro triángulo y, de pronto, se parte (“se fisiona”, en lenguaje elegante) en dos mitades iguales, que van a parar, cada una, a uno de los otros dos vértices. En seguida, cada mitad de partícula sufre el mismo destino, partiéndose en dos mitades, que van a aterrizar a los otros vértices, y así sucesivamente. Este proceso de fisión continua indefinidamente.
¿Cuál es la repartición de masa que se va produciendo en los vértices del triángulo, instante tras instante. ¿Qué ocurre a la larga

¿Ves alguna similitud entre los dos problemas?

Podrías haber pensado que como la pulga no sabe a cuál de los otros dos vértices saltar, se divide en dos, de modo que cada media pulga va a cada uno de los otros vértices. Nota que las fracciones de pulga, o de partícula, que van “aterrizando” en los distintos vértices te dan exactamente las probabilidades de presencia en ellos la pulga en dichos vértices. ¿Estarás de acuerdo entonces que el hecho de visualizar media pulga en un vértice equivale al hecho que al observar muchas veces la pulga, después de un salto, la encontremos aproximadamente la mitad de las veces en ese vértice?

Como ves, si no te gusta pensar en términos probabilistas, puedes visualizar la situación como un problema determinista: el destino de cada pedacito de partícula está totalmente determinado. No nos costaría mucho comenzar a graficarlo, anotando en cada vértice del triángulo equilátero la porción de partícula allí presente, instante tras instante, como en la figura siguiente.

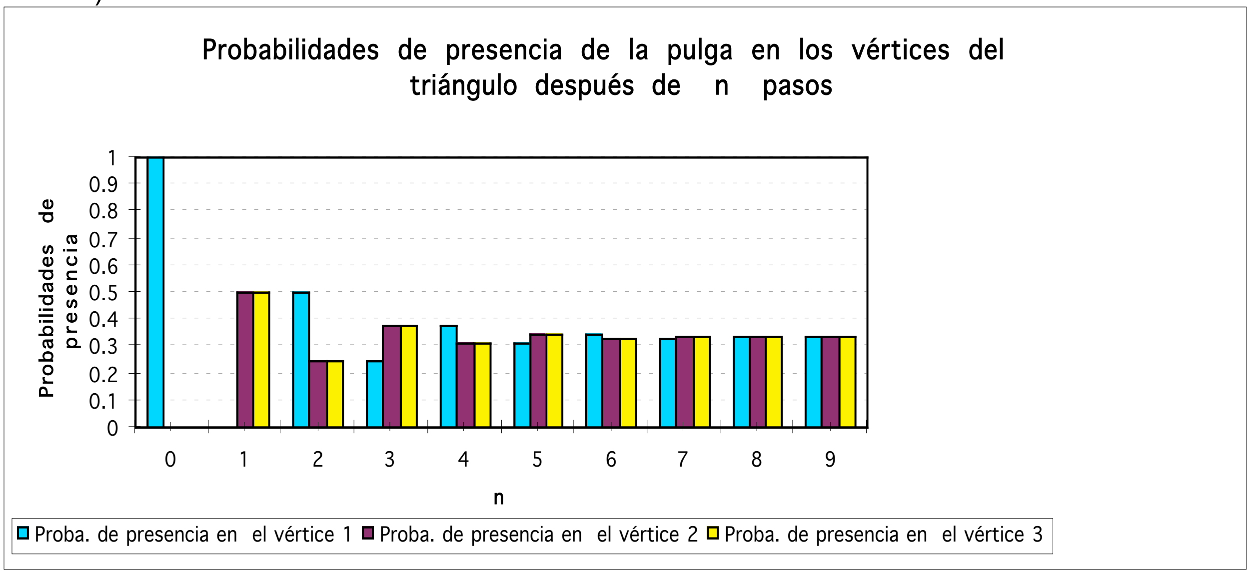


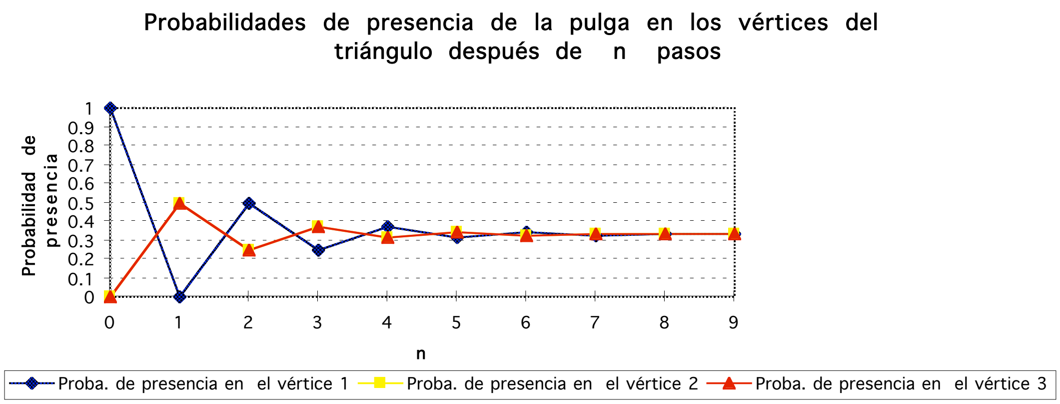
Si etiquetamos por 1 el vértice superior del triángulo, por 2 el inferior derecho y por 3 el inferior izquierdo, obtenemos la tabla siguiente, en que indicamos sólo la probabilidad de encontrar la pulga en los vértices 1 y 2, ya que la probabilidad de encontrarla en el vértice 3 es siempre igual a la de encontrarla en el vértice 2 (¡por simple simetría!).

Probabilidades de presencia de la pulga en los tres vértices del triángulo después de n saltos

n	Probabilidad que la pulga esté en el vértice 1	Probabilidad que la pulga esté en el vértice 2	Probabilidad que la pulga esté en el vértice 3
0	1	0	0
1	0	0,5	0,5
2	0,5	0,25	0,25
3	0,25	0,375	0,375
4	0,375	0,3125	0,3125
5	0,3125	0,34375	0,34375
6	0,34375	0,328125	0,328125
7	0,328125	0,3359375	0,3359375
8	0,3359375	0,33203125	0,33203125
9	0,33203125	0,333984375	0,333984375

Si graficamos estos datos, nos empezamos a dar cuenta que las probabilidades en cuestión, o bien la porción de partícula, en cada vértice tienden a equipararse, acercándose cada vez más al valor 0,33333.... = 1/3. Nota que la suma total de las tres probabilidades o de las tres porciones de partícula debe dar siempre 1 (esto se podría llamar "ley de conservación de la pulga": la pulga no se crea ni se destruye, sólo salta…)





Quizás tú estés de acuerdo en que puede ser entretenido interesarse en el destino de una pulga que salta alegremente por los vértices de un triángulo, un cuadrado, u otros polígonos regulares, pero también puede que te preguntes:

¿Y qué tiene que ver esto con los problemas contingentes de la vida cotidiana? ¿Para que puede servir?

En relación con tu pregunta, te sugerimos abordar el siguiente

Problema:
Dos marcas de dentífrico, llamadas A y B, compiten afanosamente por el mercado con grandes gastos publicitarios, que tienen el siguiente efecto:
Mes tras mes, los compradores de A se quedan con A o se cambian a B con igual probabilidad, igual a 1/2; por otra parte, los compradores de B se quedan con B con probabilidad 2/3 y se cambian a A con probabilidad 1/3.

Si al comienzo de la historia, A controlaba todo el mercado (hasta un mes antes que irrumpiera B), ¿cómo va cambiando la repartición mes tras mes? ¿ cómo se repartirán el mercado a la larga?, ¿será totalmente controlado por B?¿Corre A hacia la quiebra inminente, si no aumenta sus gastos publicitarios?

Según parece el gerente de A, bastante inquieto, solicitó los servicios de una matemática, en vez de recurrir a un vidente, para ver si podía decirle algo sobre el destino de su empresa, si las cosas seguían como iban.

La matemática, que tenía un cierto espíritu lúdico, se dijo: Investigar el devenir del mercado dividido entre estos dos productores de dentífrico se parece mucho a investigar el devenir de una partícula unitaria que, inicialmente está, entera, en la posición A, y en seguida comienza a fisionarse según la regla de más arriba: una mitad va a parar a la posición B, y luego, la mitad que se quedó en A se parte, en una mitad de la mitad, es decir 1/4 para A y otro 1/4 para B; y por otra parte, la mitad que se fue a B, se divide en 2/3 que se quedan en B y 1/3 que se va a A.

Pero, pensando un instante más, se dio cuenta que en realidad, daría lo mismo investigar el destino de su vieja amiga la pulga, que ahora estaría paseándose así:

Ella salta al azar entre dos posiciones A y B solamente, pero también puede ser que se quede en una de ellas. Más precisamente, la ley de probabilidad de su paseo es la siguiente:

Si está en la posición A , al segundo siguiente se queda allí con probabilidad 1/2 y salta a B con la misma probabilidad. Si está en B, al segundo siguiente, se queda allí con probabilidad 2/3 o bien salta a A, con probabilidad 1/3.

Si inicialmente está la pulga en A, que ocurre después de 1, 2, 3, 4,... segundos ? ¿Después de muchos segundos?

Resolución:

Como de costumbre, nada mejor que experimentar:

Fíjate que puedes simular este paseo, con ayuda de una moneda y un dado... ¿Cómo?

Puedes usar la moneda, para cuando la pulga esté en A, y el dado para cuando esté en B: si sale por ejemplo 1, 2, 3, o 4, la pulga se queda en B, pero si sale 5 o 6, salta a A.

Si cada alumno de tu curso simula su propia pulga, podremos tener unas 40 pulgas paseando simultáneamente en la sala de clases, y podremos ir haciendo la estadística correspondiente: anotar cuantas pulgas están en A y cuantas en B, después de un salto, dos saltos, ..., diez saltos, veinte saltos, cincuenta saltos... ¡Esto nos debería dar una buena idea de su destino!

Pero también podemos calcular teóricamente, pensando que en el segundo inicial, toda la masa probabilística, igual a 1, está concentrada en A. Después del primer mes,

A tendrá: 1/4 + 1/6 = 5/12

y B tendrá: 1/4 + 2/6 = 7/12.

Al mes siguiente, A tendrá: 5/24 + 7/36 = 29/72

y B tendrá: 5/24 + 14/36 = 43/72.

De nuevo, puedes verificar que la suma de 29/72 y 43/72 da 72/72 = 1, como debía ser.

¿Te animas a calcular la repartición del mercado para el mes siguiente?
Bueno, resulta que A se queda con

$$29/144 + 43/216 = 173/432$$

y B con

$$29/144 + 86/216 = 259/432.$$

Puedes verificar que la suma de las partes sigue dando 1, como corresponde.
Podemos graficar nuestros resultados en una tabla:

	A	B
Primer mes	1/2	1/2
Segundo mes	5/12	7/12
Tercer mes	29/72	43/72
Cuarto mes	173/432	259/432

Sin embargo, a estas alturas quizás estés algo mareado con este laborioso ejercicio de suma de fracciones y, para poder comparar mejor la repartición de un mes con la del siguiente, comenzarás a tener ganas de ponerlo todo en decimales, o mejor aún, en porcentajes...

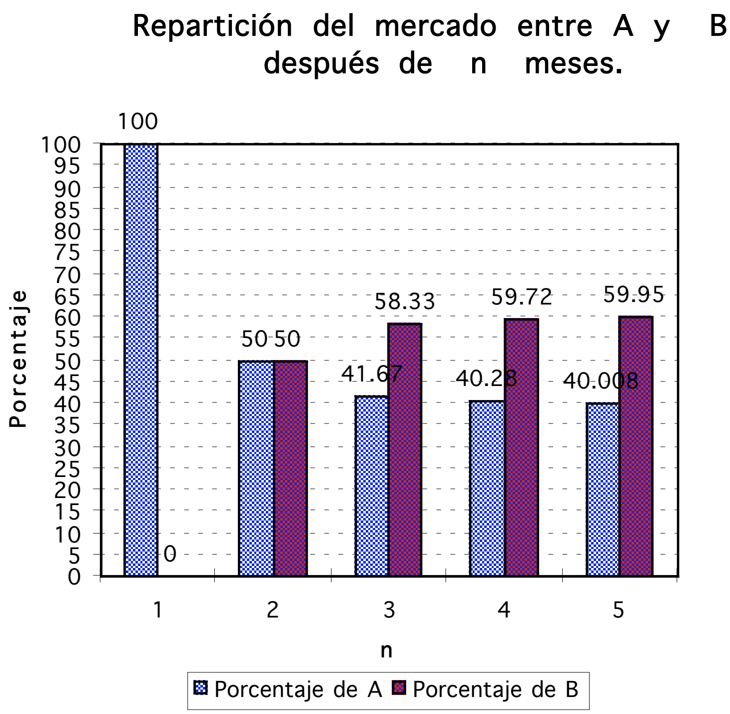
Si lo haces, con ayuda de la calculadora, obtendrás la siguiente tabla, en que dejamos pasar un mes más, para aprovechar el impulso:

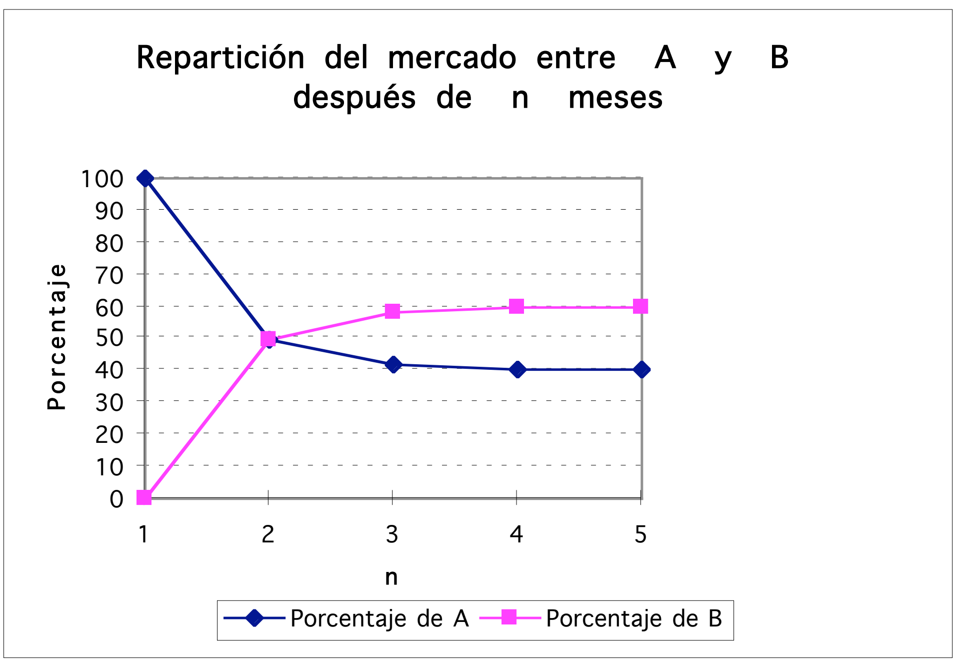
	A	B
Primer mes	50,00 %	50,00 %
Segundo mes	41,67 %	58,33 %
Tercer mes	40,28 %	59,72 %
Cuarto mes	40,05 %	59,95 %
Quinto mes	40,008 %	59,992 %

¿Ves ahora más claro que ocurre? ¿Te atreverás a apostar que después de muchos meses el porcentaje de A bajará muy, muy cerca de 40 % y el de B subirá muy, muy cerca del 60 %?

Fíjate que una consecuencia interesante de nuestro cálculo es que la marca A no será desplazada del mercado por B, si no que la repartición del mercado entre ambos tenderá a estabilizarse en torno a 40 % y 60 %. De modo que si A se conforma con su 40 % no es necesario que aumente su gasto en publicidad para estar a la altura del de B, que seguramente es mayor. Este sería el consejo de la matemática...

En forma gráfica vemos lo siguiente:





Ejercicio:¿Cómo se desarrolla la historia si al comienzo A y B se repartían el mercado por partes iguales? ¿Si sólo B estaba inicialmente presente en el mercado?

Un problema relacionado: El juego de los tres colores.

Te proponen el siguiente juego: De una bolsa que contiene una gran cantidad debolitas de 3 colores distintos, a saber rojo, azul y verde, en cantidades iguales, vas sacando sucesivamente bolitas para tratar de juntar los tres colores en su mano. Pero debes tener en cuenta que si repite un color, lo pierdes, es decir, cuando sacas una segunda bolita de un color que ya tenías, debes devolver las dos bolitas a la bolsa.

Como derecho de juego, te cobran 1 moneda de $ 100 cada vez que metes la mano en la bolsa y recibes el premio cuando logras juntar los tres colores en su mano.
¿Por qué valor del premio estarías dispuesto a jugar?

Sugerencias para un módulo de trabajo.

Comencemos jugando el juego, simplemente. He aquí un ejemplo de desarrollo (donde R= Rojo, V = Verde, A = Azul):

R V R V A A R V A

El juego duró 9 jugadas, al cabo de las cuales logramos juntar los tres colores en la mano, después de haberlos perdido varias veces. El contenido de la mano fue el siguiente, al cabo de cada jugada:

Jugada 1: R
Jugada 2: R V
Jugada 3: V (perdimos el Rojo)
Jugada 4: ø (mano vacía)
Jugada 5: A
Jugada 6: ø (mano vacía)
Jugada 7: R
Jugada 8: R V
Jugada 9: R V A (llegamos al Paraíso: los tres colores en la mano)

En este caso, habremos pagado en total 9 monedas por las 9 jugadas.
Ahora bien, un método bastante natural para poder estimar el premio justo, es simular "muchas" veces el juego, y hacer el promedio de lo que pagamos cada vez. Vale la pena observar de paso qué números obtenemos como costos posibles del juego...Pero dejamos al lector el placer de hacer su propia estimación de esta manera, a estas alturas deldesarrollo.

Ahora bien, si queremos registrar para la posteridad muchos desarrollos del juego, empezaremos quizás a sentirnos un poco insatisfechos de registrarlos cada vez como una larga palabra en las letras R, V y A. Quisiéramos entonces buscar otras maneras más geométricas de representar el desarrollo del juego que nos permitan mejor visualizar qué está ocurriendo. ¿Cómo hacerlo?

Podríamos, en lugar de hacer un video de todo el juego, tomar una foto de nuestra mano después de cada jugada, por ejemplo. Entonces el desarrollo del juego quedaría representado por una quizás larga galería de fotos.

Pregunta: ¿ Qué fotos aparecen en esta galería?

¡Notemos que la galería puede ser arbitrariamente larga, si uno tiene suficiente mala suerte! De hecho, a estas alturas, con un escalofrío, puede que nos demos cuenta que el juego podría durar muchísimo tiempo (¡más que el tiempo de vida estimado de nuestro sistema solar, por ejemplo!). Podría en teoría prolongarse indefinidamente (en la práctica, seguramente el jugador se muere antes...). Sin embargo, al jugarlo efectivamente, aunque sea con un computador, uno observa que el juego suele terminar, después de un cierto número de jugadas (que no es posible acotar de antemano). Pero, en todo caso, podemos observar que en nuestra galería de fotos hay sólo un número finito de fotos distintas, que se van eventualmente repitiendo. ¿Cuántas?

La siguiente sugerencia podría ser: Disponer este conjunto finito de fotos en la mesa o el pizarrón e ir describiendo como el desarrollo del juego nos hace partir de una cierta foto (de hecho, se trata de la foto con la mano vacía, antes de hacer la primera jugada), e ir después pasando de una foto a otra, hasta que, con un mínimo de suerte, en algún momento, llegamos al Paraíso, es decir a la foto que muestra los tres colores en la mano. ¿Nos recuerda esta representación a nuestra amiga la pulga?

Nos damos cuenta entonces que el desarrollo del juego se puede visualizar como un paseo al azar por nuestras ocho fotos, que comienza en la foto con la mano vacía y termina, en algún momento, esperamos, con la foto con la mano llena. Y no es posible transitar, en un salto, entre cualquier par de fotos. Seguramente es una buena idea entonces, conectar entre sí los pares de fotos entre las cuales se puede saltar en una jugada. Al hacer esto obtendremos un grafo, por el cual nos pasearemos al azar, al jugar el juego.

Pregunta: ¿Qué grafo se obtiene?

Ubicando convenientemente las 8 manos posibles y conectando por aristas aquellas entre las que se puede transitar en una jugada, se obtiene un cubo.

Redescubrimos de esta manera en fin de cuentas la realización combinatoria del cubo, cuyos vértices son las partes de un conjunto de 3 elementos (en este caso los 3 colores) y cuyas aristas conectan dos partes que difieren en sólo un elemento.

El desarrollo del juego se visualiza entonces como el paseo al azar canónico de la pulga por los vértices del cubo, que comienza desde el vértice correspondiente a la parte vacía y termina cuando llega (¡por primera vez!) al vértice correspondiente a la parte llena (el conjunto de los 3 colores).

Llamemos N el número de pasos que demora el paseo aleatorio en llegar a la mano llena si parte de la mano vacía. Notemos que N (¡que es un buen ejemplo de variable aleatoria!) toma los valores 3,5,7,9,11,13, …pero -según nuestra simulación- no con igual probabilidad.

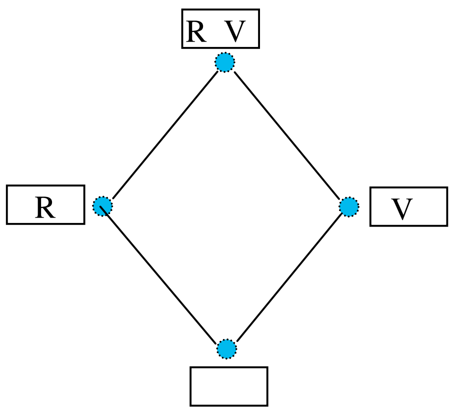
Preguntas:
- ¿A qué valor de N apostaría, si lo dejan elegir primero?
- ¿Cuál sospecha que puede ser el valor esperado de N?

Ejercicio: Calcular el número de caminos de *n* pasos que puede seguir la pulga para llegar al vértice lleno partiendo del vértice vacío y la probabilidad de cualquier camino específico de *n* pasos.

Complemento (dirigido al alumno):
El conjunto de todas las muestras de un conjunto dado: ¿cómo visualizarlo? ¡Como un hipercubo!

Si tienes una urna con dos bolitas de dos colores distintos, digamos rojo (R) y verde (V), cuáles son las posibles muestras que puedes sacar ?
Bueno, son
- la muestra vacía
- la muestra que consta de la bolita roja
- la muestra que consta de la bolita verde,
- la muestra completa, que consta de las dos bolitas.

Esto se puede representar por el diagrama de más abajo, en que hemos conectado de manera bastante natural algunas muestras, a saber aquellas que son “vecinas”. Esto último significa que uno puede pasar de una a otra añadiendo o quitando una bolita, nada más!



Quizás se ve más claro lo que ocurre, si metemos la mano en una urna con 3 bolitas de colores distintos, digamos rojo (R), azul (A) y verde (V)

Podemos disponer en un diagrama todas las muestras o partes del conjunto de los 3 colores, por jerarquías, o estratos, primero la muestra vacía, en la fila siguiente todas las muestras con 1 elemento, después las muestras con dos colores, luego la única con 3, a saber la muestra completa.

El número total de las muestras, de cualquier tamaño que sean, se obtiene entonces sumando la cantidad de muestras de cada estrato: 1 + 3 + 3 +1 , lo cual da, curiosamente 8.
Si hacemos lo mismo con 4 colores , nos queda 1+ 4 + 6 +4 + 1 = 16, lo que suma milagrosamente 16.
Nota que para dos colores, tenemos que el número total de muestras es

$$1 + 2 + 1 = 4.$$

¿ Qué te parece ? ¿Tienes la sensación que aquí hay gato encerrado, mejor dicho potencia de dos encerrada?
¿Se te ocurre como “ver” directamente estos hechos ? ¿Te atreves a conjeturar algo en general?
Si conjeturaste que el número de muestras que se puede sacar de un conjunto con n objetos es 2^n, estamos de acuerdo.

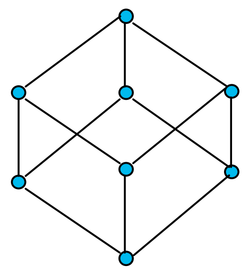
Una sugerencia: Un punto de vista periodístico.
Entrevistar a los colores, uno por uno, cada vez que se genera una muestra. Visualizar todas los posibles resultados de la encuesta o entrevista.
Si dibujas el árbol de posibilidades, verás que el número buscado es el número de descendientes de la n-ésima generación de un árbol genealógico binario.
Así vemos por qué vale la identidad

$$\binom{n}{0}+\binom{n}{1}+\ldots+\binom{n}{n-1}+\binom{n}{n}=2^n$$

Ejercicio ¿Podrías demostrar a ciegas “ esta identidad? ¡Haz la prueba para valores pequeños de n, por ejemplo n = 3, 4, 5, 6,…!

Pregunta: ¿Cómo visualizar de una buena manera este lote de muestras, en particular, estos estratos cuyo cardinal va cambiando, primero aumentando hasta llegar a un máximo y después disminuyendo, simétricamente.

Sugerencia: Trata de dibujar el diagrama (el grafo, en realidad) que resulta al conectar entre sí las muestras que son “vecinas inmediatas”, a saber, aquellas que difieren sólo en un color, añadido o quitado.
Para el caso de 3 colores, si dispusiste hábilmente tus muestras, encontrarás algo así:



¿ Te animarías a tratar de dibujar el diagrama o grafo correspondiente al caso de 4 colores? Tendrá ahora 16 vértices, que se obtienen sumando

1 + 4 + 6 + 4 + 1 .

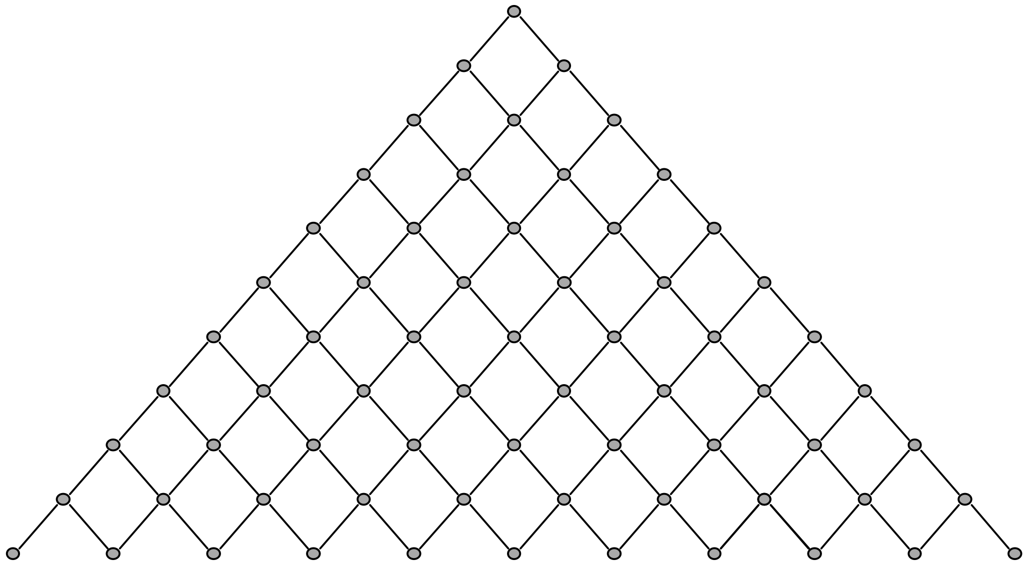
No es otra cosa que el famoso hipercubo, “aterrizado” en nuestro mundo 3D, como un grafo de alambre…

(fin del módulo de trabajo dirigido al alumno)

1.3. Sobre la migración de las hormigas hacia la miel, el triángulo de Pascal y las potencias del binomio.

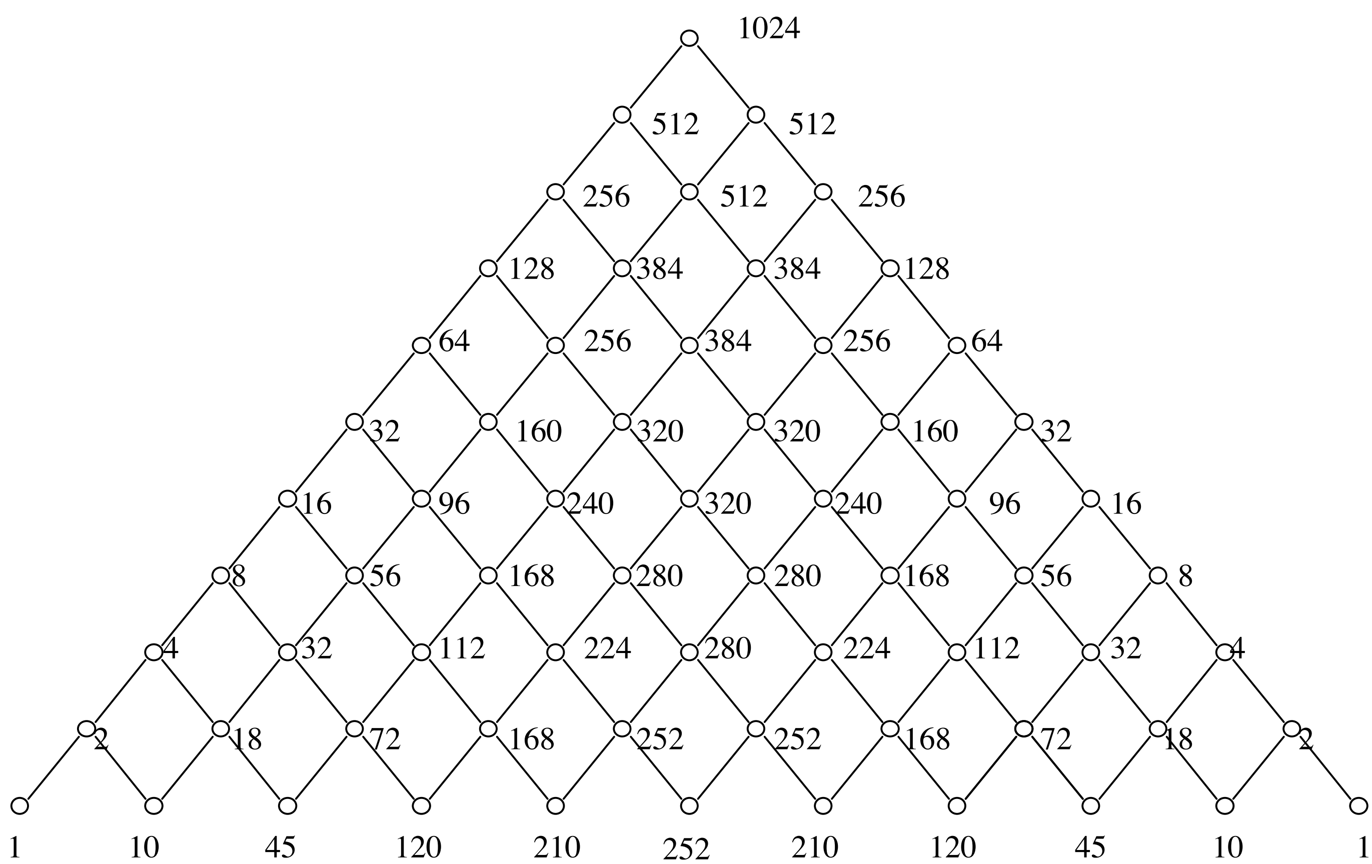
En esta subsección, nos proponemos enfatizar la interpretación combinatoria de los clásicos coeficientes binomiales y su relación con el conteo de caminos al azar en un reticulado o enrejado.

PROBLEMA MOTIVADOR: La migración de las hormigas hacia la miel
Al pie de la malla triangular de la figura de más abajo, alguien chorreó miel por inadvertencia. Y ocurre que en el nudo superior de la malla hay una gran cantidad de hormigas, digamos un millar (o mejor aún, 1024), que al sentir el olor de la miel, comienzan a migrar hacia la base de la malla. Como no tienen mayor razón para tomar derecha o izquierda en cada nudo, se dividen por mitades iguales.



PREGUNTA: ¿Cuántas hormigas llegan a cada uno de los nudos de la base de la malla de la figura

El alumno podrá ir calculando pacientemente, fila por fila, como se van repartiendo las hormigas e ir entonces obteniendo los números de hormigas de la figura de más abajo. ¡Notemos que los números de cada fila deben reconstituir el 1024 inicial!



Por supuesto, según sea el entusiasmo de los alumnos, se puede reemplazar esta migración de hormigas por una más corta, con una malla de sólo 7 o 5 filas, por ejemplo...

PROBLEMA MOTIVADOR BIS: El destino de una hormiga aleatoria.

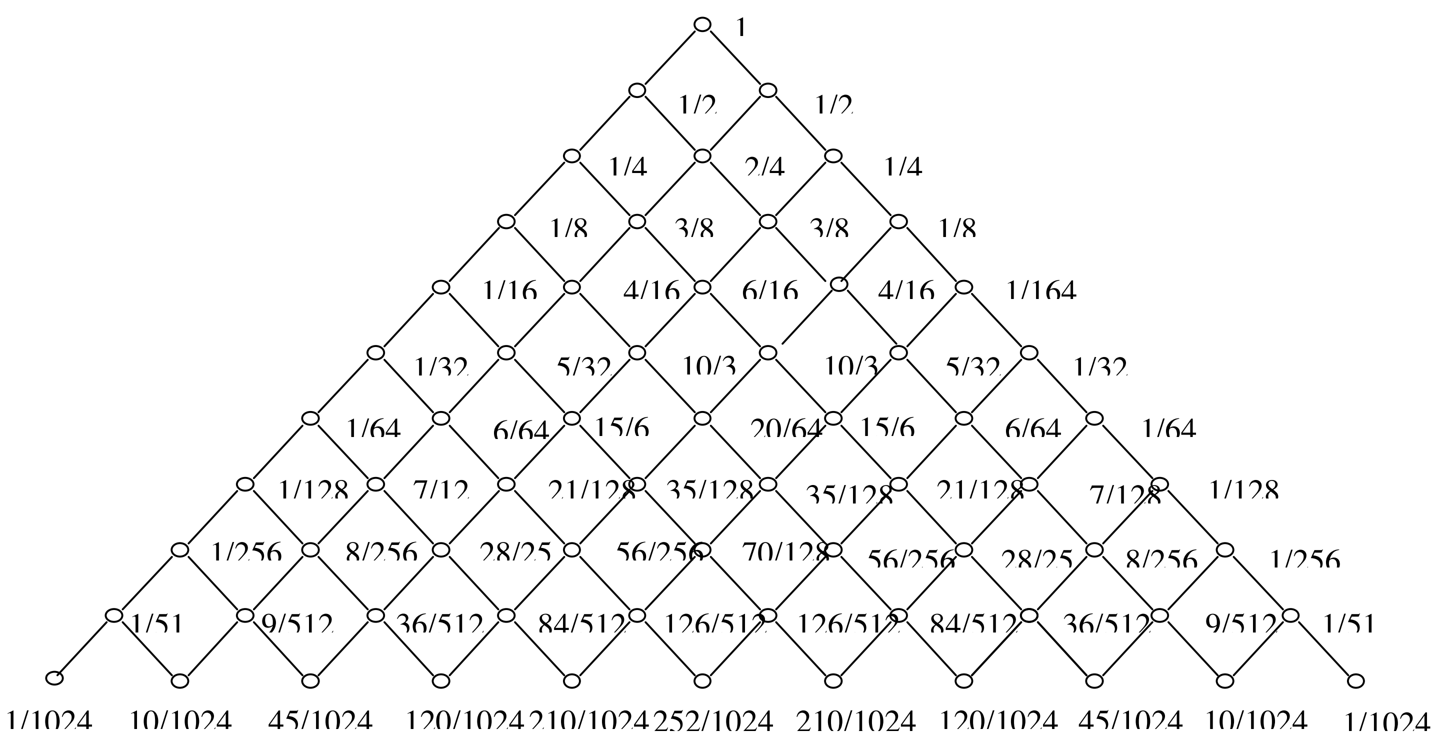
En la situación del problema anterior, vemos ahora llegar una hormiga atrasada al nodo superior de la malla. La hormiga huele la miel y lo que hicieron sus congéneres y tratará entonces apresuradamente de ir hacia la miel, eligiendo completamente al azar en cada nodo entre su derecha y su izquierda, como si fuera un cara o sello.

¿En que lugares apostaríamos que encontramos la hormiga, cuando llega a los nodos de la base, donde está la miel?

Podemos simular el destino de esta hormiga con ayuda de una moneda. No sabremos con seguridad adonde llegó, pero tendremos una "nube de probabilidad de presencia de la hormiga", como ocurre para los átomos y otras partículas infra-atómicas en mecánica cuántica...

Con el llamado "método hidráulico", en el que imaginamos que dejamos escurrirse un litro de "fluido probabilístico" desde el ápice de la malla hacia abajo, con cañerías o mangueras en lugar de los alambres, podemos ir asignando fácilmente las probabilidades respectivas como corresponde, notando que en cada bifurcación del enrejado o malla, el fluido probabilista se reparte por mitades iguales entre el conducto o alambre de derecha y el de la izquierda y que además hay cantidades de líquido que confluyen a un mismo nudo, desde los nudos situados a la derecha e izquierda más arriba (es decir, desde sus nudos "antecesores").

Procediendo así obtenemos las probabilidades de presencia indicadas en la malla de más abajo:



Después de esto, podríamos preguntar a los alumnos si ven alguna relación entre los números de hormigas caminantes que llegan a los nudos de la base de la malla y las probabilidades de presencia de nuestra hormiga paseante en dichos nudos.

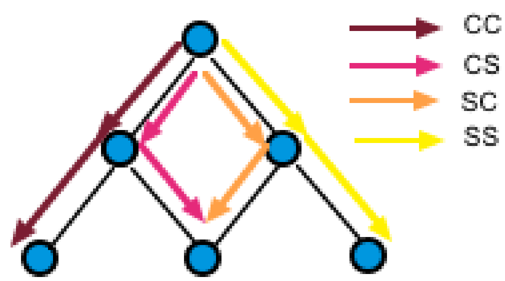
Así podrían darse cuenta que si dividen la cantidad de hormigas que llegó a cada nudo con miel por el total de hormigas que emprendió el viaje, ¡obtienen exactamente la probabilidad que nuestra hormiga aleatoria llegue a ese nudo!

En otros términos, la fracción o porcentaje del total de hormigas que llegó a cada nudo con miel es igual a la probabilidad que una hormiga aleatoria llegue a ese nudo.

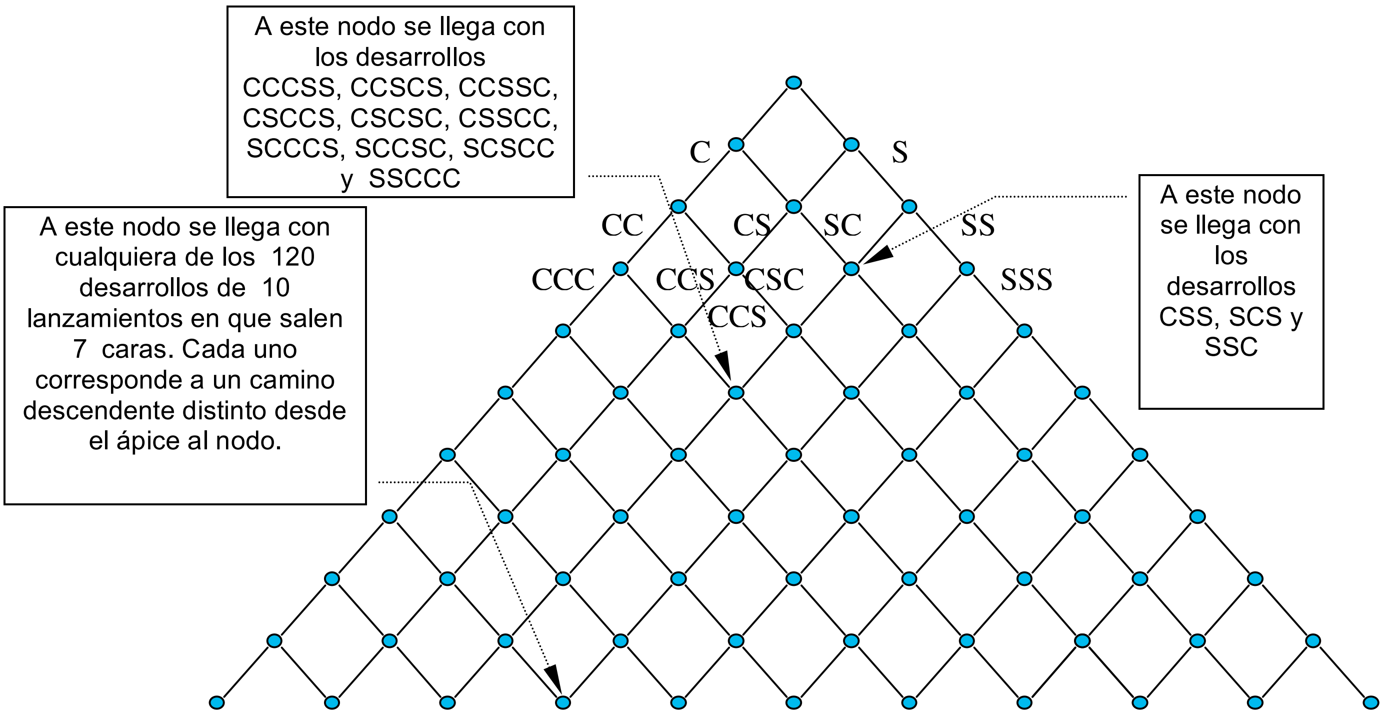
Pregunta: ¿ Cómo se relaciona esto con el lanzamiento reiterado de una moneda (experimento aleatorio con el que recomendamos iniciar el tópico de probabilidades).

Podríamos sugerir al alumno que imagine que nuestra hormiga tipo lanza una moneda cada vez que llega a un nudo para decidir si toma por el alambre de la derecha o el de la izquierda (digamos cara = izquierda, sello = derecha). Precisemos que se trata de nuestra derecha, al mirar nosotros el enrejado, no la derecha de la hormiga. Entonces a medida que la moneda va mostrando cara o sello, la hormiga va recorriendo un camino descendente por la malla, eligiendo la derecha o la izquierda al azar del cara o sello.

Los estudiantes podrán fijarse entonces que cuando la hormiga llega al segundo nudo de su viaje, puede estar en tres nudos posibles: el de la izquierda si salió CC, el del medio si salió CS o bien SC, y el de la derecha si salió SS. Y que puede llegar al nodo del medio por dos caminos distintos, mientras que a los nodos extremos sólo puede llegar por un camino. Graficamos esto en la figura siguiente:

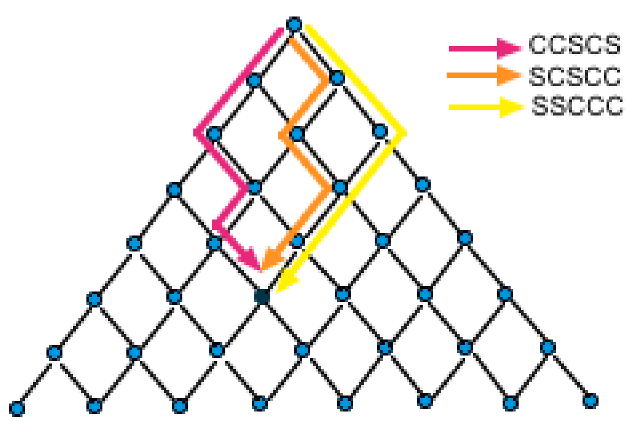


Al terminar la hormiga la tercera etapa de su viaje, tiene 4 nudos posibles de aterrizaje que corresponden a los desarrollos CCC , CCS o CSC o SCC, CSS o SCS o SSC, y SSS. Ilustramos esto en la figura siguiente:



¿A qué corresponde el nodo en que se encuentra la hormiga, en términos del cara o sello? O dicho de otra manera, ¿qué tienen en común todos los caminos descendentes que llevan a un mismo nodo?

Mirando con cuidado, se ve que tienen en común su largo, es decir el número de etapas, o nudos por que pasan, y además *el número de veces que la hormiga eligió tomar el alambre de la izquierda.*



Por ejemplo, en la figura de arriba, hemos dibujado (como podrían hacerlo los alumnos) tres caminos distintos que descienden desde el ápice al tercer nodo de la quinta fila. Vemos que todos ellos eligen 3 veces la izquierda y que esto corresponde exactamente al hecho que la hormiga obtuvo 3 veces cara al lanzar su moneda para elegir su camino.

Si nos fijamos bien, vemos entonces que, en general, la posición del nodo a que llega la hormiga después de n etapas está dada simplemente por el número de caras, por ejemplo, obtenido al lanzar n veces su moneda.

Al mirar los caminos descendentes como una sucesión de caras y sellos, a saber el resultado de lanzar n veces una moneda, cuando se trata de un camino de largo n, vemos que:

¡el número de veces que el camino elige la izquierda es exactamente el número de veces que la moneda salió cara!

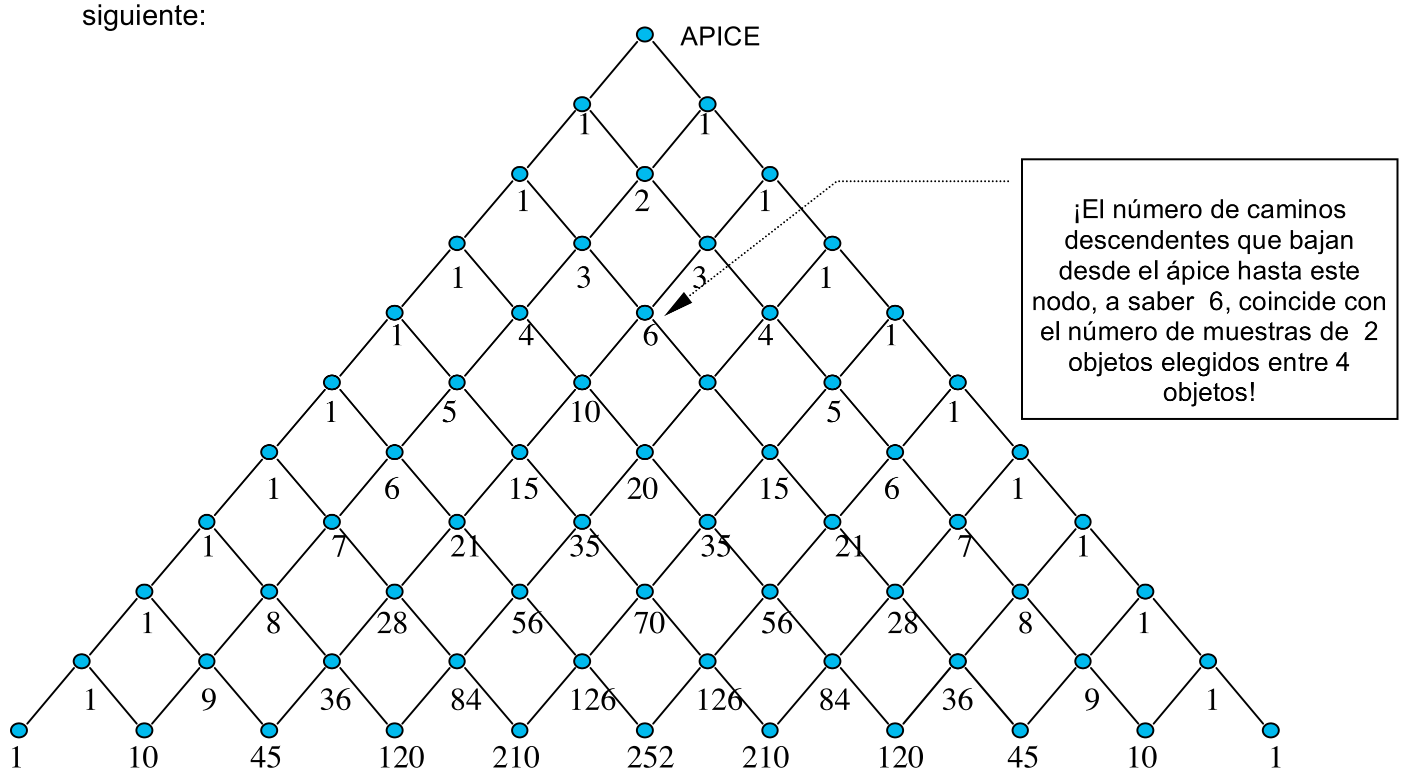
Si dejamos que los caminos se extiendan hasta el fin de la malla, es decir que consten de 10 etapas, veremos que el nudo en que termina el camino describe exactamente el número de caras (o el de sellos, si prefieren) obtenido al lanzar 10 veces la moneda.

Ejercicio: A estas alturas, podríamos invitar a los estudiantes a anotar frente a cada uno de los 11 nudos con miel a cuántas caras corresponde.

Contemos caminos descendentes...

En la misma malla anterior, ¿cuántos caminos podemos seguir para ir desde el ápice, hasta un nodo, o cruce de alambres, cualquiera?

Podríamos animar al alumno a ir anotando sucesivamente, fila por fila, estos números de caminos. Entonces encontrará los resultados indicados en el diagrama siguiente:



Número de caminos descendentes desde el ápice de una malla triangular

Ejemplo de Módulo de trabajo sobre el triángulo de Pascal (dirigido al alumno):

¿Te fijas que hay una relación estrecha entre el número de caminos descendentes que llega a un nodo y los números de caminos que llegan a sus dos "antecesores" (los dos nodos vecinos situados más arriba)?

Si la hormiga llegó afanosamente a uno de los dos antecesores de nuestro nodo, ya está casi al final de su viaje: le queda sólo un único tramo de alambre que recorrer en cada caso, para llegar a su nodo de destino. Así entonces:

El número de caminos descendentes que bajan desde el ápice a un nodo cualquiera se obtiene sumando los números de caminos que bajan desde el ápice a cada uno de los dos nodos antecesores.

Gracias a esta regla general que acabamos de descubrir, tú podrías calcular, pacientemente, paso a paso, el número de caminos que bajan hasta cualquier nodo de tu malla, por grande que sea ésta.

El triángulo de números así obtenido se llama históricamente **Triángulo de Pascal,** en honor a Blaise Pascal (matemático, físico y teólogo francés del siglo XVII), o también **Triángulo de Tartaglia** (matemático italiano del siglo anterior, artífice del método de resolución de la ecuación de tercer grado, atribuido hoy a su contemporáneo Cardano). Aparentemente fue descubierto en forma independiente por ambos.

Pero, ¿te suenan conocidos los números que vamos obteniendo así?

Por ejemplo, los números de caminos que bajan hasta los nudos de la cuarta "generación" o fila son, leyendo de izquierda a derecha (mira la figura de arriba): 1, 4, 6. 4. 1. Pero estos números son exactamente los llamados coeficientes binomiales,

$$\binom{4}{0}, \binom{4}{1}, \binom{4}{2}, \binom{4}{3}, \binom{4}{4}.$$

que cuentan el número de muestras de, respectivamente 0, 1, 2, 3, y 4 objetos que se puede elegir entre 4 objetos.

Nota que $\binom{4}{0} = 1 = \binom{4}{4}$, puesto que, por ejemplo, de una urna con 4 bolitas, hay una sola manera de elegir una muestra vacía, sin ninguna bolita, y también una sola manera de elegir la muestra completa, que contiene las 4 bolitas de la urna.

Puedes verificar que esto ocurre también para las otras generaciones, desde la primera en adelante. Seguramente, no es una coincidencia...¿Ves una explicación de este hecho?

¿Por qué el número de caminos que bajan, por ejemplo, hasta el nudo del medio de la cuarta generación, que es 6, ha de coincidir con $\binom{4}{2}$, que es el número de maneras de elegir 2 objetos entre 4?

Bueno, mira la figura de más arriba y piensa un poco....Cuando tú eliges uno de los 6 caminos descendentes posibles, estás eligiendo en realidad 2 objetos entre 4, sin darte cuenta... ¿Cuáles serán?

Si no te parece claro, pregúntale a la hormiga. ¡Ella lo debe tener bien claro! Te explicará que elegir su camino equivale a escoger, en este caso, **cuatro veces** entre el alambre de su derecha y el de su izquierda. Y que para llegar al nodo del medio de la cuarta generación, tiene que haber elegido, forzosamente, **dos veces** la derecha dentro de sus cuatro elecciones. Pero no importa si eligió la derecha en la primera y segunda bifurcación, o la segunda y la cuarta, por ejemplo. Esto cambia la trayectoria, pero no el nodo de llegada. Este último sólo depende del **número de veces** que eligió irse por la derecha...

Así, vemos que efectivamente, elegir un camino que baje hasta nuestro nodo corresponde a *elegir las dos bifurcaciones en que tomaremos la derecha. entre las cuatro que pasaremos*. Y esto último puede hacerse exactamente de $\binom{4}{2}$ = 6 maneras.

Este razonamiento se podría hacer más sistemático como sigue:

Mira, por ejemplo, la última generación de nuestra malla, que tiene 11 nudos (o nodos).. Cada uno de ellos se podría etiquetar por un número, a saber el número de veces que una diligente hormiga tiene que tomar el alambre de la derecha para bajar desde al ápice al nudo en cuestión. Al primer nudo de la fila, el de la extrema izquierda, le corresponde el número 0, porque para llegar a él la hormiga jamás tuvo que elegir la derecha, sino que sólo la izquierda, todo el tiempo. Al segundo, le corresponderá el número 1, porque para bajar hasta a él, la hormiga debe elegir sólo 1 vez la derecha y el resto de las veces la izquierda, no importando en que momento eligió el alambre de la derecha. Así sucesivamente, llegamos a etiquetar el nodo de la extrema derecha por el número 10, ya que para llegar a él la hormiga debe forzosamente elegir en las 10 bifurcaciones el alambre de la derecha.

Ahora bien, si estamos parados en el nodo número k entre los once nodos de la última generación, el conjunto de caminos que bajan del ápice hasta nuestro nodo está en correspondencia uno a uno con el conjunto de maneras de elegir k veces la derecha en un total de 10 elecciones. O también, si quieres pensar que la hormiga lanza una moneda para decidir cada vez, con el conjunto de todos los desarrollos posibles de 10 lanzamientos de una moneda en los que se obtuvo k caras.

¡El número de caminos que nos interesan es entonces $\binom{10}{k}$, nuestro antiguo conocido!

Más generalmente, imagina que estuviésemos parados en la n-ésima fila o “generación” de una malla de largo arbitrario, y que etiquetamos los nodos de esta fila con el método de más arriba. Entonces tendremos $\binom{n}{k}$ caminos distintos para bajar desde el ápice hasta el nudo de la n-ésima fila etiquetado por el número k, porque elegir un tal camino significa elegir en cuáles k bifurcaciones la hormiga tomó el alambre de la derecha.

Cosecha: Recuerda que habíamos notado que los caminos descendentes en nuestra malla se suman de una manera bien natural:

El número de caminos que bajan hasta un nodo cualquiera es la suma de los números de caminos que llegan a sus dos antepasados inmediatos, una fila más arriba.

Si nos fijamos en las etiquetas que acabamos de distribuir a todos los nodos, los antepasados del nodo etiquetado con el número **k** en la **n**-ésima fila son los nudos con las etiquetas **k** (el de arriba a la derecha) y **k - 1** (el de arriba a la izquierda). Entonces, esta propiedad de suma de número de caminos, clara para cualquier hormiga que camine por la malla, se puede escribir como una hermosa identidad de coeficientes binomiales, con lo cual parece más difícil que lo que en realidad es:

$$\boxed{\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}}$$

Ejercicio: ¿Podrías demostrar esta identidad de coeficientes binomiales “a ciegas”? ¿Es decir, calculando a partir de la definición numérica simplemente? , para soltar la mano, un ejemplo, primero, como

$$\binom{9}{4} + \binom{9}{5} = \binom{10}{5},$$

para soltar la mano..

(fin del Módulo de Trabajo dirigido al alumno)

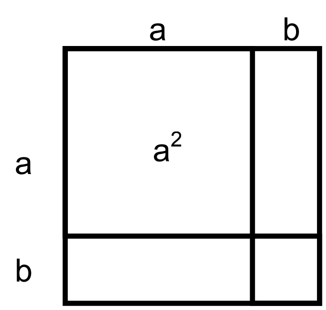
Caminar hacia las potencias del binomio:

Problema 1: No sólo calcular, sino que visualizar, $\mathbf{(a + b)^2}$.

Dinámica: El cálculo algebraico se hace sin mayor dificultad, pero seguramente parecerá algo árido y poco motivador a muchos alumnos. Suele suscitar mayor interés la pregunta de cómo visualizar el hecho que

$$\mathbf{(a + b)^2 = a^2 + ab + ba + b^2 = a^2 + 2ab + b^2\ \ ?}$$

Una buena respuesta a esa pregunta es el dibujo siguiente:



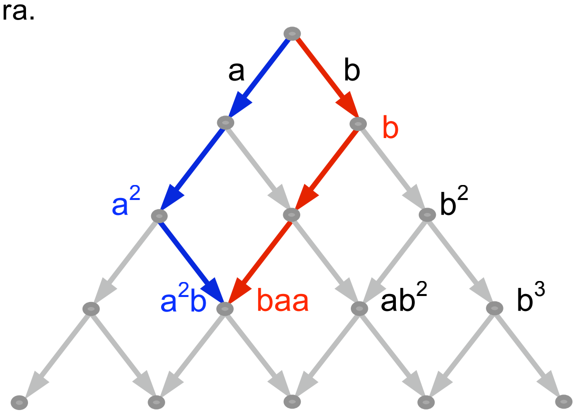
Problema 2:
Pero ahora, ¿ cómo visualizar que $\mathbf{(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3}$

Esto lleva a dibujar un cubo de lado **a + b** en perspectiva, que se descompone en subcubos de lado **a** y **b**, y además en subparalelepípedos cada uno de cuyos lados es **a** o **b**. En lugar de dibujar este cubo en perspectiva, se podría también abordar su construcción en 3 dimensiones, en alambre de colores, plasticina, cartón o madera coloreados. Esto puede ser el tema de una actividad interesante para los alumnos.

Un abordaje diferente al problema de calcular las potencias del binomio **a + b** surge de la observación que para desarrollar estas potencias, por ejemplo **$(a + b)^3$**, hay que multiplicar "todos con todos", es decir elegir en cada factor **(a + b)** uno entre **a** y **b**, y hacer el producto correspondiente. La situación es entonces como un cara o sello, reiterado 3 veces: si sale cara, elijo **a** , si sale sello elijo **b**. Las posibilidades se pueden tabular como sigue, teniendo en cuenta la conmutatividad del producto:

Desarrollo del cara o sello (tres tiradas)	**Producto asociado a cada desarrollo**	**Resultado en la forma usual**
CCC	aaa	a^3
CCS	aab	a^2b
CSC	aba	a^2b
CSS	abb	ab^2
SCC	baa	a^2b
SCS	bab	ab^2
SSC	bba	ab^2
SSS	bbb	b^3

Pero esto se visualiza mejor aún en nuestra malla triangular. Como indicado en la figura de más abajo, la elección de **a** corresponde a un paso descendente hacia la izquierda, la **de b** a un paso descendente a la derecha, los productos **ab** y **ba** corresponden a los dos caminos que bajan desde el ápice al nodo medio de la segunda fila, definidos por las elecciones izquierda-derecha y derecha-izquierda, respectivamente. La conmutatividad del producto: **ab = ba**, se visualiza como el hecho que los dos caminos nos llevan al mismo nodo de la malla. Análogamente, para **$(a + b)^3$** , cada producto posible del desarrollo, por ejemplo, baa, corresponde a un camino descendente desde el ápice hasta uno de los nodos de la tercera fila, como indicado en la misma figura.



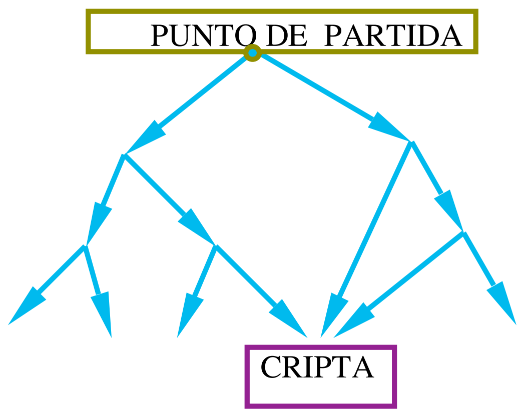
Esta representación gráfica y pedestre de los productos que aparecen al desarrollar las potencias del binomio es totalmente general, y debiera permitir al estudiante "ver", *paso a paso*, la siguiente descripción pedestre del triángulo de Pascal:

El coeficiente de cada monomio $a^{n-k} b^k$ en el desarrollo de la potencia n-ésima del binomio (a + b) es exactamente el número de caminos descendentes que bajan desde el ápice hasta el k - ésimo nodo de la fila n-ésima de nuestra malla.

Y este número coincide con el coeficiente binomial $\binom{n}{k}$, que cuenta el número de muestras de k objetos que se puede elegir entre n objetos.

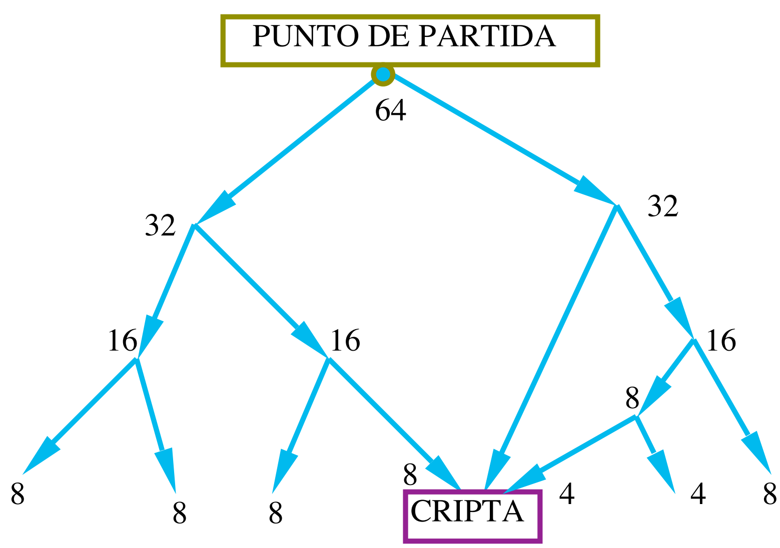
1.4 Una metáfora pedestre para calcular probabilidades: El peregrinaje.

Un grupo de 64 peregrinos ha perdido a su guía por el camino, pero de todas maneras quieren llegar a una cripta que era el objetivo de su viaje, siguiendo la red de caminos de la figura. Si en cada encrucijada se reparten por mitades iguales entre izquierda y derecha, ¿cuántos peregrinos llegan a la cripta?

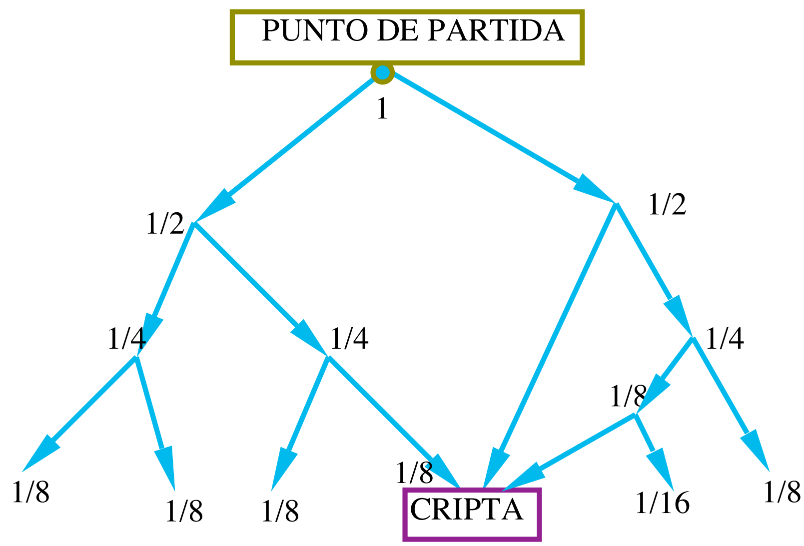


Después que el grupo de peregrinos había partido, llegó al punto de partida un peregrino atrasado, que se había quedado conversando por el camino. Este camina apresuradamente por la red de caminos, eligiendo al azar, como si lanzara una moneda, entre su derecha e izquierda, en cada encrucijada. ¿Cuál es la probabilidad que llegue a la cripta por puro azar?

Resolución: Calculando paso a paso obtenemos los números de peregrinos indicados en la figura:



y vemos así que llegan a la cripta 8 + 16 + 4 = 28 peregrinos. La fracción de peregrinos que llegan a buen puerto es entonces 28/64 ≈ 43,8%. En seguida, las probabilidades de paso del peregrino aleatorio por cada encrucijada se pueden ir asignando paso a paso por el “método hidráulico” de dejar escurrirse 1 litro de fluido probabilista desde el ápice del árbol de caminos, de modo que se reparta por partes iguales en cada bifurcación. Obtenemos así los números de la figura:



de donde se ve que la probabilidad que el peregrino aleatorio llegue a cripta por pura suerte es 1/8 + 1/4 + 1/16 = 7/16, igual a los 28/64 de más arriba.

Este ejemplo muestra entonces como se puede calcular probabilidades de manera pedestre, echando a caminar un grupo de peregrinos y contando que fracción de ellos llega a tal cual destino...

2. Al azar del juego: Estadística, probabilidades y combinatoria.

Nuestro punto de vista metodológico, de presentar los contenidos matemáticos de manera inductiva y constructivista, nos lleva naturalmente a motivar, por ejemplo, la combinatoria por las probabilidades: uno quiere contar de cuántas maneras se puede escoger algo para poder calcular su probabilidad de éxito en un cierto juego. Y desea calcular esta última para saber si el juego propuesto le conviene o no. al tener en cuenta el valor del derecho de juego y el premio ofrecido...

También enfatizamos el abordaje experimental de los problemas, es decir, en fin de cuentas, una actitud estadística. ¡ En términos clásicos podríamos decir que estamos mucho más cerca de Galileo que de Aristóteles!

Finalmente, al presentar cada noción o técnica, estamos pensando en cómo se podría llevar al aula. Por esto, aunque estas notas se dirigen primeramente al profesor, incluyen también algunos módulos o párrafos dirigidos directamente al alumno, además de sugerir actividades concretas realizables en grupos relativamente grandes de alumnos.

2.1. El rol de las simetrías en la vida cotidiana: El artesano y sus collares.

Un artesano quiere regalar collares, que él mismo fabrica, a algunas de sus amigas, de modo que nunca se pueda producir aquella catastrófica situación en que dos de ellas se encuentren llevando el "mismo" collar.

Si fabrica sus collares colocando piedras de lapislázuli (piedra azul originaria de Chile) y rodocrocita (piedra rosada, originaria de Córdoba, Argentina) ¿a cuántas de sus amigas puede impunemente regalar collares ?

Siendo el problema práctico de fácil comprensión, podemos animar a los estudiantes a tratar de resolverlo por intento y error, construyendo collares y descartando en seguida aquellos susceptibles de traerle problemas...

Una manera más sistemática de resolver este problema es mirar subrepticiamente por encima del hombro al artesano cuando fabrica los todos los collares hexagonales posibles, a saber $64 = 2^6$, usando las dos piedras indicadas y los pone sobre su mesa de trabajo.

En seguida se fija que dos collares distintos pueden verse iguales puestos al cuello si uno de ellos se obtiene a partir del otro por una rotación, o también *una reflexión* del hexágono. Es decir se da cuenta que tiene una relación de equivalencia muy natural entre sus collares:

"dos collares son equivalentes si podrían verse iguales puestos al cuello"

Entonces los agrupa ahora en clases de equivalencia, que son en realidad, en este caso, las órbitas de la acción natural del grupo de simetrías del hexágono regular en el conjunto de los 64 collares. Y encuentra de este modo la siguiente partición del conjunto de los 64 collares:

- 2 órbitas de 1 collar,
- 1 órbita de 2 collares,
- 2 órbitas de 3 collares,
- 7 órbitas de 6 collares,
- 1 órbita de 12 collares.

Es decir, en total sólo 13 clases de equivalencia. Podrá entonces regalar collares a sólo 13 amigas, si no quiere correr ningún riesgo...

Además se percatará de lo variable que es el tamaño de las órbitas (aunque este tamaño sea siempre un divisor de 12, el orden del grupo de simetrías, como corresponde), y se preguntará que propiedad geométrica de un collar se expresa en lo grande o pequeño que sea el cardinal de su órbita...

Este problema admite, por supuesto, una infinidad de variantes, tomando otros polígonos regulares como "molde" para fabricar collares, con más o menos vértices que un hexágono, usando eventualmente tres o cuatro clases de piedras en vez de dos, etc.

Una pregunta algo impertinente que se podría hacer más de algún alumno es: ¿cómo podría yo haber adivinado de antemano ese número fatal, 13, sin haber dibujado o construido laboriosamente todos los collares?
o bien:

¿No habrá alguna formulita mágica para calcularlo a partir de los datos del problema: 6 lugares donde poner piedrecitas, 2 tipos de piedrecitas?
o también:

¿Cuál será el correspondiente número crítico para collares pentagonales?
(*Respuesta*: 8).

Asimismo podría preguntarse uno si no habrá otra manera de acceder a este número crítico: 13 para el hexágono, 8 para el pentágono...

Y la verdad del asunto, es que sí hay otras maneras. Entre ellas una manera estadística, si el lector consiente en jugar el siguiente juego:

El juego de los collares hexagonales fijos:

El juego consiste en elegir totalmente al azar una de las 12 simetrías del hexágono y fijarse cuántos de los 64 collares hexagonales posibles deja fijos (o invariantes). Su premio en el juego es de tantos miles de pesos como collares dejó fijos su simetría.

Por ejemplo, si uno saca la identidad, que fija todos los collares, ganará 64 miles de pesos. Si por el contrario, tiene la mala suerte de extraer la rotación en un sexto de vuelta, que sólo fija los dos collares monocromáticos (todo azul o todo rosado), ganará sólo 2 mil pesos.

El valor del premio, es decir su ganancia en el juego (todavía no le cobran por jugar...) varía entonces entre 2 y 64 miles de pesos, pasando por otros valores intermedios.

La interrogante que surge naturalmente en relación a este juego es:

¿Cuánto puedo esperar ganar en él?

Esta pregunta, como muchas otras relativas a juegos de azar, se puede responder por un abordaje estadístico: simplemente jugando muchas veces el juego, y fijándose que ocurre. ¡Dicho sea de paso, esta actitud "galileana" frente a un problema seguramente merece ser estimulada en la enseñanza media!

Más precisamente, lo que trataríamos de hacer es jugar muchas veces el juego, anotar cuidadosamente cuanto ganamos cada vez y....¡hacer el promedio de nuestras ganancias!

Un problema práctico que los estudiantes deberían poder resolver, para hacer esto, es cómo simular la elección al azar de una de las 12 simetrías del hexágono (que conocen desde la educación básica, se supone), usando objetos de la vida cotidiana.

Bueno, usando una moneda y un dado, no debe ser muy difícil:

Lanzamos primero la moneda, para elegir entre las seis rotaciones y las seis reflexiones: si sale cara, nos quedamos con las rotaciones y si sale sello nos quedamos con las reflexiones. En seguida, si por ejemplo salió cara, elegimos entre las rotaciones en un sexto de vuelta, en dos sextos de vuelta, etc., hasta seis sextos de vuelta (es decir, la identidad), lanzando un dado: si sale 4, por ejemplo, elegimos la rotación en 4 sextos de vuelta. Análogamente podemos numerar las 6 reflexiones (arbitrariamente) de modo de elegir una de ellas tirando también un dado.

Ahora bien, si jugamos un buen número de veces este juego, por ejemplo 100 veces, ¡nuestra ganancia promedio será aproximadamente de 13 miles de pesos!

Miremos más de cerca los collares pentagonales, como ejemplo. En este caso, como ya dijimos, hay sólo 8 clases de equivalencia . Es decir el artesano puede regalar sólo 8 collares distintos (hechos con piedras de dos colores distintos), que nunca se verán iguales puestos al cuello. Si comenzamos ahora a "sacar" simetrías del pentágono, y a preguntarles cuántos de los 32 collares pentagonales posibles dejan fijos, observamos lo siguiente:

- Hay una simetría, a saber la identidad, que fija los 32 collares.
- Hay 4 simetrías, a saber las 4 rotaciones distintas de la identidad, que finan sólo 2 collares (los dos collares "tontos", es decir, monocromáticos).
- Hay 5 simetrías, a saber las 5 reflexiones, que tienen el mismo comportamiento, es decir fijan 8 collares. Cada reflexión tiene un cierto eje de reflexión (o de simetría), que pasa por un cierto vértice del pentágono, y fija los 8 collares que son simétricos respecto ese eje.

Al hacer entonces el promedio del número de collares fijos por cada una de las 10 simetrías del pentágono, obtenemos:

$$\frac{1\cdot 32+4\cdot 2+5\cdot 8}{10}=\frac{32+8+40}{10}=\frac{80}{10}=8\ .$$

Ejercicio:	Hacer el caso análogo de los collares hexagonales.

Es notable, entonces, como el artesano puede llegar a la misma conclusión, a saber que sólo puede regalar 8 collares sin correr riesgos, ¡tanto por un análisis geométrico sobre su mesa de trabajo como por una experimentación estadística!

Notemos también, que si nos entusiasmamos con el juego propuesto, de elegir al azar una simetría, del pentágono, por ejemplo, y recibir tantos miles de pesos como collares deje fijos la simetría, la pregunta crucial práctica que aparece es:

¿Cuál es el derecho de juego justo que debería pagar para que el juego sea equitativo?

Es decir, intuitivamente, ¿cuál es el valor del derecho de juego, que debería pagar, por cada jugada, para que al jugar muchas veces el juego yo, en promedio, no ganara ni perdiera?

El interés práctico de esta pregunta es que si tenemos claro cuál este valor, nos podremos dar cuenta inmediatamente si en la vida cotidiana nos proponen este juego con un precio de entrada demasiado alto, en cuyo caso el juego nos será desfavorable, o bien demasiado bajo, en cuyo caso el juego nos será favorable. También, los estudiantes podrían aplicar este conocimiento, para inventar o proponer un juego como éste, pero con un precio por jugada levemente más alto que el justo, para reunir fondos en una kermesse, por ejemplo.

Para nuestros ejemplos, en el caso de los collares pentagonales, si proponemos el juego con un premio de tantas monedas de cien pesos como collares fije la simetría elegida y con un derecho de juego de 9 monedas por jugada, deberíamos podernos enriquecer, con un poco de paciencia... Análogamente, si cobramos, digamos, 14 monedas por el juego con los collares hexagonales.

En este contexto vemos aparecer un ejemplo muy natural de "cantidad azarosa": el premio obtenido al jugar. A continuación, nos interesaremos más sistemáticamente en el comportamiento de estas cantidades azarosas, llamadas de costumbre "variables aleatorias", en lenguaje más elegante.

2.2. ¿Cómo se porta?: Introducción a las variables aleatorias

Entre los contenidos mínimos de 3º Medio aparece una introducción a la teoría de probabilidades (más modestamente, al cálculo de probabilidades) y a la estadística. Pero, si miramos bien, veremos que los verdaderos protagonistas de estos temas son las variables aleatorias: aquellas cantidades o magnitudes que varían azarosamente, y cuyos valores quedan determinados sólo después de haber realizado un experimento en que interviene el azar.

El objetivo de esta subsección es proponer una manera bastante experimental, inductiva y lúdica de tomar contacto con las variables aleatorias, averiguar cómo se portan y describir su comportamiento, aprovechando al máximo la intuición física de los estudiantes. En particular enfatizamos la interpretación física y visual de la esperanza y la desviación típica de una variable aleatoria, útiles indicadores suscintos de su comportamiento.

2.2.1 . Variables aleatorias por doquier.

En la vida cotidiana nos encontramos a cada rato con variables aleatorias. Nos parece entonces fundamental animar a los estudiantes a prestarles atención y a verbalizar múltiples ejemplos, de variadas índoles.

He aquí una pequeña lista:

- El número de minutos que un alumno espera la micro en su esquina de costumbre para venir al colegio.
- La nota que un estudiante del curso obtuvo en la última prueba. O bien la nota final que obtendrá en el curso.
- El número de hermanos que tiene un estudiante del curso.
- La estatura, medida en cm, de un alumno del curso.
- El número de caries no tratadas que tiene un alumno del curso.
- El número de años que vive un habitante de este país.
- El número de veces que uno sufre un corte de luz durante el año en su casa.
- El número de accidentes carreteros ocurridos en Chile durante un fin de semana.
- El número de partidos que gana su equipo de fútbol favorito durante el torneo anual.
- El ingreso mensual de su grupo familiar.
- El puntaje que obtendrá un alumno de su colegio, o de cualquier colegio de Chile, en la PAA este año.

- El número de amigos con que se encontró un alumno de su curso en la calle durante el mes pasado.

A estos ejemplos, se pueden añadir otros, que son más “artificiales”, más lúdicos o bien menos habituales:

- El número de puntos que muestra (la cara de arriba de) un dado lanzado al azar.
- La suma de puntos que muestran dos dados lanzados al azar.
- El número de caras que obtengo al lanzar 2 monedas al azar.
- El número de caras que obtengo al lanzar 10 monedas al azar.
- El número de caras que obtengo al lanzar 100 veces una moneda.
- El color de un naipe que saco al azar de un mazo de naipes inglés.
- El monto del premio que obtengo en un cierto juego.
- El número de vértices que deja fijos una simetría del hexágono, elegida al azar. Y análogamente para cualquier polígono regular...
- El número de veces que tendré que tirar una moneda hasta que me salga cara
- El número de veces que tendré que tirar un dado hasta que me salga seis.

Sugerimos comentar la (aparente) mayor o menor “utilidad” de estas variables aleatorias. Creemos que con demasiada frecuencia los estudiantes han recibido ejemplos sumamente “gratuitos” y artificiales de las nociones que les son impartidas, lo cual más bien bloquea el aprendizaje que lo facilita. Puede ser importante estimular la capacidad crítica del estudiante al respecto.

En la situación presente, el alumno podría decir: “Esta variable aleatoria me parece relevante o importante, pero esta otra no veo para que pueda servir, ni por qué interesarme en ella” .

Entre los ejemplos dados, posiblemente el número de minutos que debe esperar el micro en la esquina le parezca relevante (aunque quizás menos su número de años de vida...) pero el número de veces que debe lanzar una moneda para obtener cara quizás le interese mucho menos.

Después de discutir francamente este punto, el profesor podría sin embargo hacer notar a los estudiantes que las variables aleatorias aparentemente “artificiales”, “ociosas” o puramente lúdicas, sirven para modelar situaciones muy relevantes o importantes de la vida cotidiana. Y que se suele comenzar estudiando aquellas que son artificialmente simples.

A estas alturas, podríamos contarles la historia siguiente:

Vida y muerte de una mosca:

En el minuto 0 entra subrepticiamente una mosca a la sala de clases. El profesor empieza, poco compasivamente, a tratar de matarla, con un matamoscas, apenas se pose. La mosca vuela aproximadamente un minuto antes de hacerlo, ocasión en que el profesor le da con el matamoscas y la extermina, no con seguridad, sino que con probabilidad 1/2 , o 50%, si prefieren. Si falla, entonces la mosca sigue volando durante otro minuto, hasta que se posa en algún lugar de nuevo y sufre otro intento de asesinato, que tendrá éxito con probabilidad 1/2, una vez más. Y así continúa la historia, minuto tras minuto...

PREGUNTA (sumamente importante para la mosca, y también para las compañías aseguradoras de moscas...): ¿Cuánto tiempo vivirá la mosca, en estas difíciles circunstancias?

Esta pregunta, como muchas otras, admite un abordaje experimental:

Actividad sugerida:

Pedir a los alumnos del curso, que cada uno se ocupe de una mosca y anote su destino. Más precisamente, podría cada estudiante simular el destino de una mosca, con ... una moneda, lanzándola, de manera que la mosca muera si sale cara, por ejemplo, y sobreviva si sale sello. Y tirar al aire la moneda reiteradamente, desde el minuto 1, hasta obtener cara, es decir hasta ultimar la mosca.

Anotará entonces el tiempo de vida "su" mosca y podrá comparar con los tiempos de vida de las moscas de sus compañeros. Desde un punto de vista práctico, si se trabaja con un curso numeroso, cuando todos han terminado de simular su mosca, se puede pedir simplemente que alcen la mano primero, todos aquellos cuya mosca vivió 1 minuto, en seguida todos aquellos cuya mosca vivió 2 minutos, etc. Esto permite establecer rápidamente la tabla de frecuencias absolutas de cada tiempo de vida. De aquí se puede pasar a la tabla de frecuencias relativas y preguntar cómo podríamos graficar estos datos.

En seguida, más de alguno propondrá hacer el promedio de los tiempos de vida y, quizás, venderle esa información a una compañía aseguradora de moscas...

Nótese que de esta manera los estudiantes estarán investigando el comportamiento de la variable aleatoria "tiempo de vida" (de una mosca que es ultimada minuto tras minuto con probabilidad 1/2) y calculando su valor promedio o valor esperado. Todo esto de manera bastante contextualizada y natural.

Observación: Si se quiere evitar un proceso potencialmente infinito (teóricamente, el profesor podría fallar indefinidamente en matar la mosca, o bien la moneda podría caer obstinadamente sello "hasta el fin de los tiempos"), se podría hacer notar que la paciencia del profesor es finita (más finita aún que su vida, o la vida del sistema solar...), así que si falla, digamos 10 veces, sacará un spray de su maletín con el cual ultimará a la mosca con seguridad al minuto siguiente (y de paso intoxicará a toda la clase y contribuirá a aumentar el agujero de ozono...). En todo caso, lo que observarán experimentalmente los alumnos es que las moscas rarísima vez viven mucho tiempo. En un curso entero, posiblemente ninguna mosca viva más de 10 minutos...

Esperamos que este ejemplo muestre como el juego, que algunos podrían tildar de "ocioso", de esperar cara al lanzar una moneda, sirve para modelar o simular problemas de vida o muerte (para las moscas en particular, pero también para muchos otros seres, sometidos a condiciones de vida precarias...), o también situaciones de éxito o fracaso muy variadas: éxito de estudiantes al aprobar un curso, "éxito" al contagiarse con un virus al exponerse reiteradamente a él, etc. Ejemplos de este último tipo, corresponden a una mosca que sería ultimada con probabilidad 1/6, o 1/10 , 1/100 o 1/1.000 , en lugar de 1/2.

Ahora bien, a estas alturas, habrán notado que ¡olvidamos recordar la *definición* de variable aleatoria! En lugar de eso, nos concentramos en dar *ejemplos* de variables aleatorias. Pero en realidad, desde el punto de vista de la compresión de los estudiantes, nos parece incluso más importante que sean capaces de dar ejemplos cotidianos y no cotidianos de variables aleatorias antes que repetir monocordemente una definición abstracta de variable aleatoria, que no les llega mayormente...

Desde el punto del proceso de evaluación, esto sugiere también que sería más sensato evaluar la capacidad de dar ejemplos del alumno, en lugar de su capacidad de recitar definiciones y teoremas. Asimismo, evaluar su capacidad de reconocer variables aleatorias en una situación contextual, en vez de recitar su definición formal.

De todas maneras, puede ser un ejercicio instructivo, preguntarles cómo definirían ellos una variable aleatoria, en general, después de todas las actividades realizadas.

Quizás redescubran algunas de las definiciones algo vagas pero intuitivas del siglo pasado, comenzando por la de "variable":

"Una variable es una cantidad o magnitud que no es constante, que es susceptible de variar (en el curso del tiempo, por ejemplo)".

"Una variable aleatoria es una variable cuyos valores son determinados por el resultado de un experimento aleatorio"

(recordemos que "aleatorio" significa "azaroso", "determinado por el azar", y proviene del latín "aleas" que significó originalmente "dados" y luego, por extensión, "azar".

2.2.2. ¿Cómo se portan las variables aleatorias?

El paso siguiente, sería preguntarle a los alumnos:

Si se encuentran a boca de jarro con una variable aleatoria en la vida cotidiana, que le preguntarían, para conocerla mejor?

Se puede hacer notar que en la práctica uno está constantemente haciendo apuestas sobre el comportamiento de las variables aleatorias. Por ejemplo, cuando un alumno parte de su casa en la mañana hacia el colegio, pretendiendo no llegar atrasado, está apostando algo sobre el comportamiento de la variable aleatoria "tiempo de espera del bus en la esquina". Más precisamente, está estimando el valor promedio o valor esperable de ese tiempo que tendrá que aguardar el bus. Y "sabe" que los valores de esta variable estarán entre 0 minutos (si el bus llega junto con él a la esquina) hasta ....posiblemente no más de 20 o 30 minutos, muy excepcionalmente.

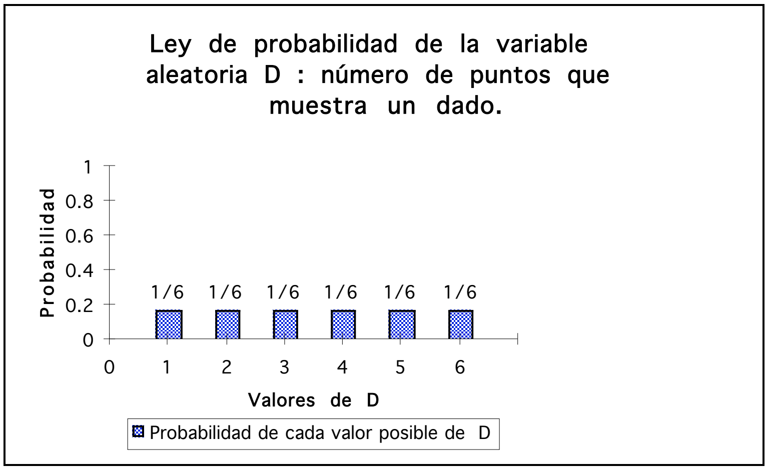
Para abreviar, entonces, podríamos decir que uno estará muy contento en la vida cotidiana si al encontrarse con una variable aleatoria uno averigua:

- qué valores toma,
- con qué probabilidad toma cada uno de esos valores,
- cuál es el valor promedio o valor esperado de la variable.

Es muy importante que podamos visualizar estos datos, para tener un retrato de la variable, es decir una imagen de su comportamiento, o más precisamente aún, un gráfico de su ley de probabilidad. Usualmente, ponemos los valores de la variable en el eje horizontal y las probabilidades respectivas como barritas encima de cada valor, en la vertical.

Como ejemplo, en los gráficos de más abajo ilustramos la ley de probabilidades de las variables aleatorias siguientes:

- D = número de puntos que muestra un dado,
- C_4 = número de caras obtenido al lanzar 4 veces una moneda
- T = tiempo de vida de una mosca que es ultimada con seguridad al minuto 5 si sobrevivió los 4 intentos de asesinato previos.


Ley de probabilidad de la variable aleatoria D : número de puntos que muestra un dado.
Probabilidad
1
0.8
0.6
0.4
0.2
0
1/6
1/6
1/6
1/6
1/6
1/6
0
1
2
3
4
5
6
Valores de D
Probabilidad de cada valor posible de D

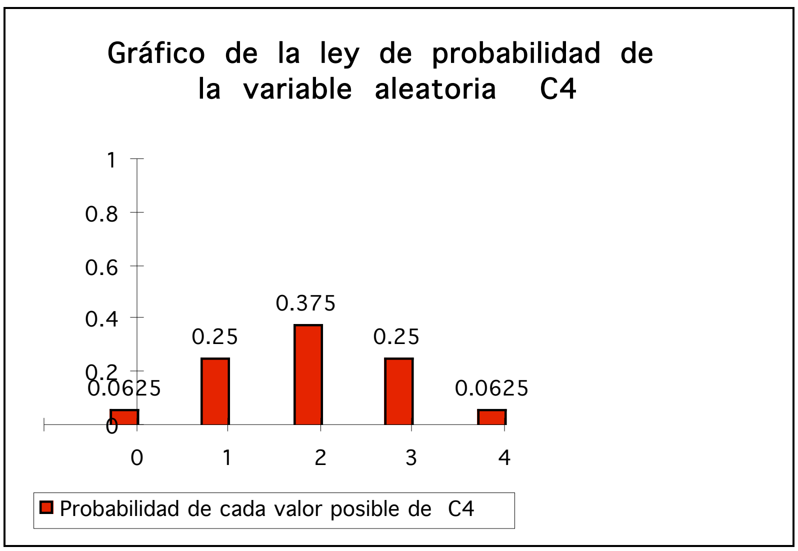


Gráfico de la ley de probabilidad de la variable aleatoria C4
1
0.8
0.6
0.4
0.2
0
0.0625
0.25
0.375
0.25
0.0625
0
1
2
3
4
Probabilidad de cada valor posible de C4

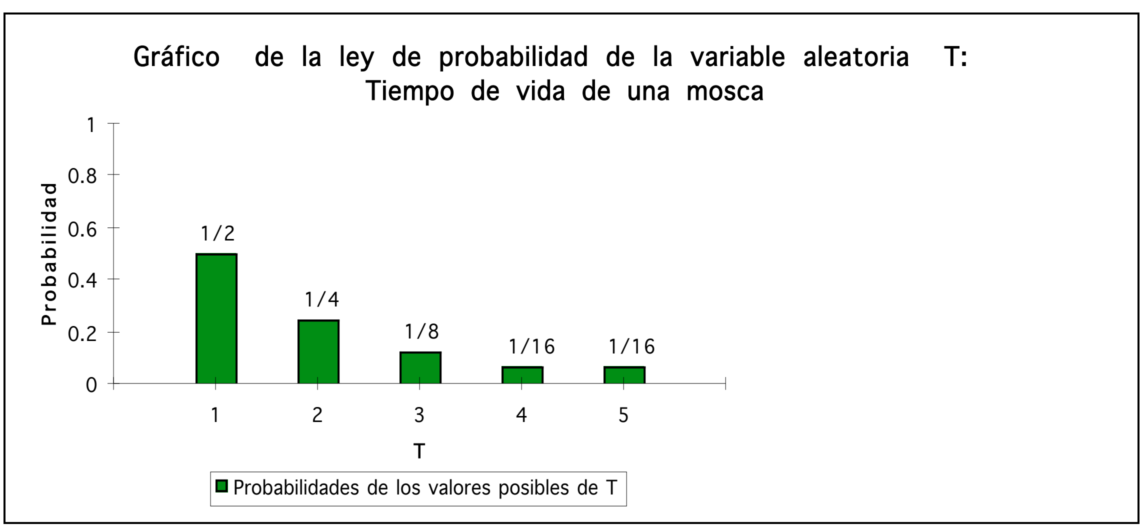


Gráfico de la ley de probabilidad de la variable aleatoria T: Tiempo de vida de una mosca
Probabilidad
1
0.8
0.6
0.4
0.2
0
1/2
1/4
1/8
1/16
1/16
1
2
3
4
5
T
Probabilidades de los valores posibles de T

Las probabilidades con que estas variables aleatorias toman sus valores posibles se calculan sin mayor dificultad, con la regla de Laplace, dividiendo el número de casos favorables por el número total de casos posibles (fijándose bien que estos últimos sean igualmente posibles).

En el caso de D, que es el número de puntos que marca el dado, los seis resultados posibles son claramente igualmente posibles (si el dado no está cargado), lo cual significa en fin de cuentas que la masa probabilística total 1 se reparte equitativamente entre los seis, tocándole 1/6 a cada uno, como indicado en el gráfico.

En el caso de la variable aleatoria C_4 , que es el número de caras obtenido al lanzar 4 veces una moneda, los resultados del experimento son los 16 desarrollos posibles de este cara o sello reiterado 4 veces, que son igualmente posibles, por simple simetría (si la moneda no está cargada). En seguida hay que fijarse cuántos de ellos nos dan 0 caras, 1 cara, 2 caras, 3 caras y 4 caras.

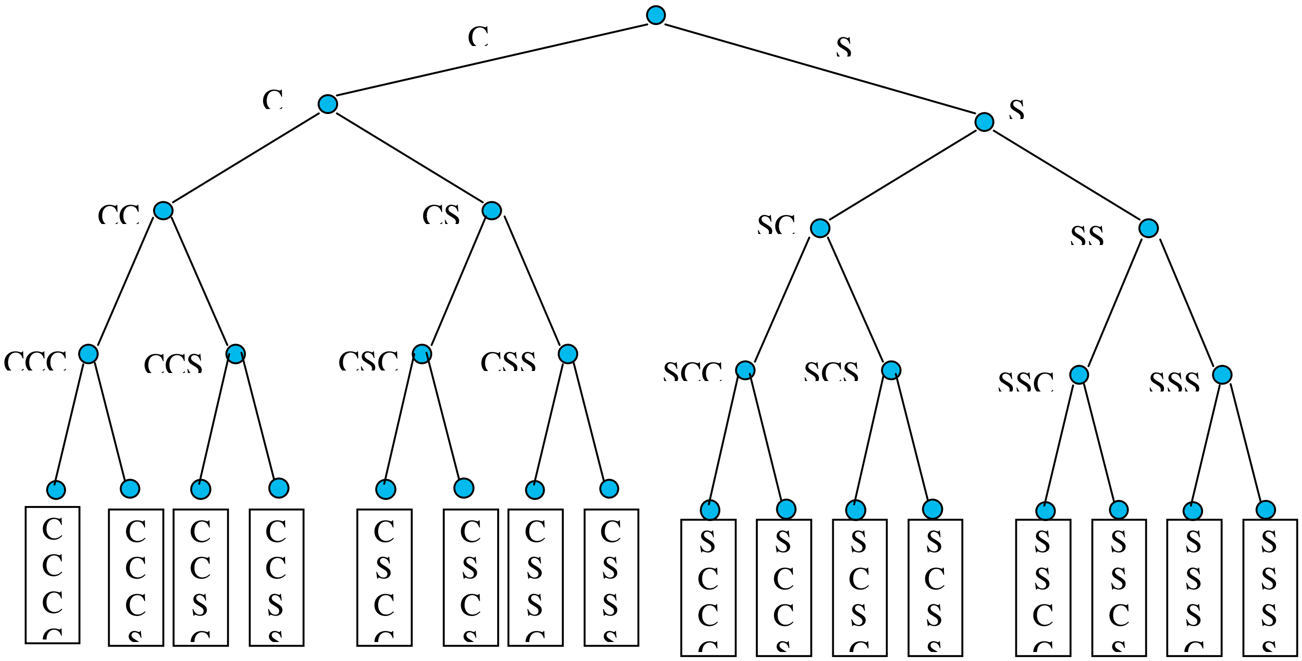
Actividades guiadas:

Creemos muy instructivo que, como primera estrategia para abordar el problema, los alumnos dibujen y visualicen el árbol de posibilidades que corresponde a este experimento, y trabajen en él, en vez de hacerlo de manera puramente numérica. Mirando el árbol pueden contar entonces cuántos desarrollos nos dan 0, 1, 2, 3 o 4 caras.

Estrictamente hablando, sin embargo, no deberíamos olvidar que hay una estrategia previa a la visualización del árbol de posibilidades, a saber la simple experimentación estadística: repetir muchas veces los 4 lanzamientos de la moneda y tomar nota de lo que pasa. En particular, ¿a qué cantidad de caras conviene más apostar?

Este abordaje estadístico puede ser realizado mediante una actividad colectiva en paralelo en el aula, en que cada alumno lance una moneda 4 veces y anote cuantas caras obtuvo. En seguida, basta que el profesor pida que levanten la mano los alumnos que no obtuvieron ninguna cara, los que obtuvieron una cara, dos caras, tres caras y, finalmente, los que obtuvieron cuatro caras. Al comparar, en un curso de unos cuarenta alumnos, las frecuencias relativas de estos cuatro sucesos (“ninguna cara”, “una cara”, “dos caras”, “tres caras” y “cuatro caras”), los alumnos podrán convencerse empíricamente que no son en absoluto igualmente posibles (o, como se suele decir, “equiprobables”), sino que “ninguna cara” y “cuatro caras” son bien poco probables, que “una cara” y “tres caras” son bastante más probables, y que el suceso al que conviene apostar entre todos estos, parece ser “dos caras”, como era de imaginar intuitivamente (¿o no?).

Arbol de posibilidades para 4 lanzamientos de una moneda

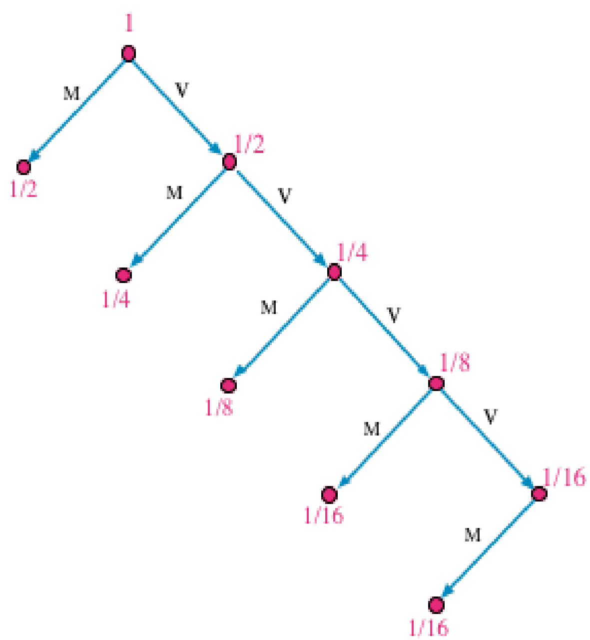


Ejercicio para los alumnos:

Ubicar y contar pacientemente los desarrollos de los 4 lanzamientos que realizan los sucesos “ninguna cara”, “una cara”, “dos caras”, “tres caras” y “cuatro caras”. Deducir así las probabilidades indicadas en el gráfico de más arriba para la variable C_4 .

En el caso de la mosca, obtenemos este árbol con probabilidades asignadas:

Arbol de probabilidades para el tiempo de vida T de la mosca que sufre los embates de un profesor impaciente.



La asignación de probabilidades se visualiza muy intuitivamente con ayuda del método "hidráulico" en que ponemos un litro de "fluido probabilista" en la raíz o ápice del árbol y lo dejamos simplemente escurrir hacia abajo, siguiendo la regla siguiente: en cada bifurcación se reparte simétricamente, por partes iguales, entre la rama o conducto de la derecha y el de la izquierda. Esta repartición simétrica corresponde al hecho que la mosca es exterminada con probabilidad 0,5 en cada intento del profesor, del primero al cuarto. Nótese que en el minuto 5 cuando el profesor utiliza alevosamente un spray, no hay bifurcación, sino que un solo conducto saliente del nodo, así que, naturalmente, todo el fluido restante se va por ahí...

En este árbol, se indica por M la muerte de la mosca y por V su (sobre)vida, cada una de las cuales acontece, minuto tras minuto, con probabilidad 1/2, salvo en el minuto 5 en que ocurre M con probabilidad 1. Nótese como se reconstituye el litro inicial si se suma todas las cantidades de fluido probabilista que terminan apozadas en los nodos de la izquierda, a saber:

$$1 = 1/2 + 1/4 + 1/8 + 1/16 + 1/16.$$

Esta observación sugiere como calcular cómoda e intuitivamente variadas sumas geométricas...

2.2.3. Esperanza o valor esperado de una variable aleatoria: interpretación física.

Se suele definir, algo secamente, la esperanza $E(X)$ de una variable aleatoria X, que toma valores $x_1, x_2, \ldots, x_n$, con probabilidades $p_1, p_2, \ldots, p_n$, por la fórmula:

$$\boxed{E(X) = p_1 x_1 + p_2 x_2 + \ldots + p_n x_n}$$

Creemos que esta presentación merece por lo menos dos contrapuntos:

1. La esperanza es simplemente el valor promedio de la variable X.

Es decir, con un abordaje estadístico, si uno realiza muchas veces el experimento aleatorio que determina el valor de X, toma nota de los valores obtenidos y hace el promedio de éstos, obtendrá una buena aproximación de $E(X)$.

2. La esperanza es el centro de masa de la distribución de masas probabilísticas.

Si miramos el gráfico de la ley de probabilidad de X como una distribución de masas, $p_1, p_2, \ldots, p_n$, (¡cuya suma total es 1, por supuesto!) ubicadas en los puntos $x_1, x_2, \ldots, x_n$, de una regla rígida graduada, entonces la esperanza $E(X)$ es aquel punto de la regla en que hay que poner un pivote (o el dedo) para dejar en equilibrio el sistema de masas (de modo que no se ladee ni hacia la derecha ni hacia la izquierda). En otras palabras, la esperanza $E(X)$ no es otra cosa que el centro de masa del sistema de masas $p_1, p_2, \ldots, p_n$.

Actividades sugeridas:

Comentemos estas estrategias para introducir la esperanza en el ejemplo, bien sencillo, en que la variable aleatoria X es igual a D: el número de puntos que muestra un dado.

En el abordaje estadístico, los alumnos se pondrían a lanzar el dado reiteradamente, unas 100 veces (por ejemplo, en grupos de a 5, lanzando 20 veces cada uno), anotarían los puntos que muestra cada vez y calcularían el promedio de todos los puntajes.

Obtendrán así un valor que muy probablemente estará bien cercano a 3.5 (que es exactamente el promedio aritmético de 1, 2, 3, 4, 5 y 6).

En este momento conviene hacer notar que el valor obtenido no es un valor posible de D, ya que el dado nunca muestra 3,5 puntos, pero esto no le quita en absoluto relevancia a este valor promedio en la vida cotidiana.

Una historia verídica:

Podríamos, como ejemplo concreto de la relevancia del valor promedio, contarles la historia de aquel generoso padre de familia, bastante patriarcal y algo perverso, que le da a su primogénito una mesada semanal de tantos miles de pesos como puntos muestre un dado que lanza ceremoniosamente todos los domingos a medio día. El niño planea economizar toda su mesada aleatoria para comprarse un lector de CD, y le interesa saber cuánto dinero podría esperar haber recibido después de, digamos, 10 semanas. La respuesta es . . . justamente 10 veces 3,5 miles de pesos, es decir 35.000 pesos, cifra bastante honorable para sus fines.

Si ahora miramos físicamente el problema, se trata de ver dónde está el centro de masa de la distribución de masas ilustrada en el gráfico de la variable aleatoria D de más arriba. Y se ve a simple vista que el lugar donde hay que poner el dedo para equilibrar el sistema de masas es el “punto medio”, a saber 3,5. Aquí ayuda a los alumnos su intuición psicomotora, desarrollada mediante la manipulación de objetos y los juegos que hacen intervenir la motricidad (y no precisamente por el trabajo con el computador....).

Vale la pena notar también que, al aplicar mecánicamente la formula que da la esperanza como promedio ponderado, no es improbable que los alumnos cometan pequeños errores de cálculo, que suelen tener catastróficas consecuencias sobre el valor final de la esperanza. Sin embargo, si la visualizan como el punto de equilibrio de la distribución de masas que representa la ley de probabilidad de la variable aleatoria, verán a simple vista dónde está ubicada, aproximadamente por lo menos, y esto *les permitirá darse cuenta si han obtenido antes un valor absurdo por simple error de cálculo*.

En el caso de la segunda variable aleatoria, el número de caras, llamado C_4, obtenido al lanzar 4 veces una moneda, se ve por simple simetría en el gráfico de su ley de probabilidad que el punto de equilibrio del sistema de masas correspondientes es 2 .

El caso de la tercera variable aleatoria, a saber el tiempo de vida T de la mosca que será ultimada por el profesor impaciente, es más interesante.

Sugerimos el siguiente plan de trabajo:

1. **Experimentación estadística:**
2.

Primeramente experimentar, simulando una buena cantidad de moscas. Si cada alumno de un curso de 40 alumnos simula tan sólo 10 moscas, tendremos en muy poco tiempo la biografía de 400 moscas. Haciendo el promedio del tiempo de vida de todas estas moscas se obtendrá una muy buena aproximación del tiempo de vida esperado de una mosca que tiene una tal vida. Se puede aquí incitar a los alumnos a adivinar cuál es el valor “exacto” de esta esperanza de vida.

2. Estimación psicomotriz:

En seguida, animar a los alumnos a mirar el gráfico de la ley de probabilidad de T como una distribución de masas y estimar, apoyándose en su intuición motriz, dónde está el punto de equilibrio.

¿Estará cerca de 2? ¿Un poquito a la izquierda de 2? ¿O un poquito a la derecha de 2? En particular, como siempre algunos conjeturan que es exactamente 2, se los puede incitar a que verifiquen si el punto de equilibrio es 2 o no.

Esto los lleva a verificar la ecuación de equilibrio, *en que la contribución de cada masa es igual al producto de la masa por su brazo* (= distancia de la masa al punto de pivote). Esta ecuación se puede motivar mediante el ejemplo del juego de balancín entre un niño que pesa 40 k y otro que pesa 20 k. ¿A qué distancia del punto de pivote del balancín debiera ubicarse el niño que pesa 40 k si el otro está ubicado a 1 m, de modo que queden equilibrados si están en reposo?

En nuestro caso, la ecuación de equilibrio, para el valor candidato 2, es:

$$\boxed{\frac{1}{2}\cdot 1 = \frac{1}{8}\cdot 1 + \frac{1}{16}\cdot 2 + \frac{1}{16}\cdot 3}$$

que equivale a $\frac{8}{16} = \frac{7}{16}$, ecuación incorrecta, que muestra al mismo tiempo cómo habría que desplazar el punto de pivote para tener una igualdad ... El valor correcto del punto de equilibrio es, por lo demás,

$$\frac{31}{16} = 2 - \frac{1}{16}$$

es decir, la esperanza de vida de nuestra perseguida mosca es de dos minutos menos un dieciseisavo de minuto, ¡un poquito menos de 2 minutos!

Ejercicio para los alumnos más emprendedores:

¿Qué ocurre con la esperanza de vida de la mosca cuando el profesor se pone cada vez más paciente, y recién usa el fatal spray después de 10 fracasos, de 50 fracasos, de 100 fracasos...? Y si no sufren de vértigo, ¿cuánto sería la esperanza de vida de la mosca si el profesor tuviera una paciencia infinita? Abordar el problema de las diversas maneras sugeridas, y a lo mejor alguna otra...

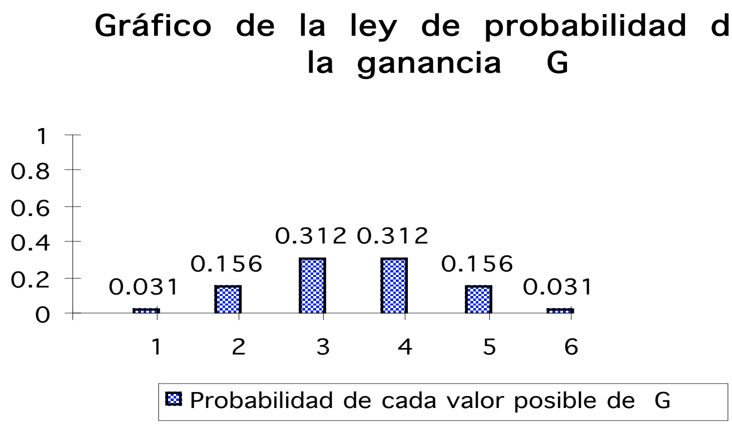
2.2.4. Desviación típica de una variable aleatoria

Este importante parámetro estadístico tiene también una interpretación motriz muy intuitiva, que podríamos motivar a través del ejemplo siguiente.

Miremos dos variables aleatorias distintas pero parecidas en alguna medida, a saber D (el número de puntos que muestra nuestro fiel dado) y G, que es la ganancia que se puede obtener en el siguiente juego:

El alumno recibe 5 monedas, las lanza al aire simultáneamente y se queda con todas aquellas que salieron cara. Y además le regalan una moneda suplementaria, por haber participado en el juego. Así se determina su ganancia G en el juego, que puede ser 1 moneda como mínimo y 6 monedas como máximo.

El gráfico de la ley de probabilidad de G es muy parecido al de la variable C_4 de más arriba:



Si los miramos como distribuciones de masas, ambos tienen masa total 1, por supuesto, y además se equilibran en el punto 3.5, es decir tienen el mismo centro de masa, o esperanza. La diferencia está en que la variable D tiene una distribución de masa más "desparramada" que G.

Pregunta 1: ¿Cómo podríamos detectar la diferencia entre estas dos distribuciones de masas probabilísticas, de una manera motriz, imaginando que las tenemos en la mano en una pieza oscura, ambas equilibradas en torno al valor 3.5? La idea es que no vemos los tamaños de las barritas y sólo podemos manipular discretamente ambas distribuciones de masas (que están puestas sobre una regla rígida), es decir no está permitido palpar cada masita por separado.

Por supuesto, ambas tienen una masa total de 1 kilo, y mientras las tengamos equilibradas estáticamente sobre el punto de equilibrio o pivote 3.5, no notaremos la diferencia entre ellas. Creemos que vale la pena dejar a los alumnos reaccionar frente a este problema.

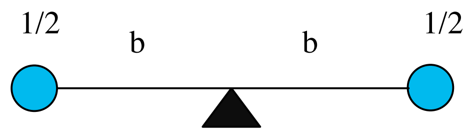
No es infrecuente que algunos de ellos imaginen rotar, o hacer girar, la distribución de masas en torno a su punto de equilibrio. Los que tienen un mayor grado de contacto con su motricidad tendrán claro que será más difícil hacer girar sobre si misma la distribución de masas más desparramada, porque tiene una mayor "inercia" al giro.

Para equiparar las experiencias motrices, se podría hacer una demostración, en la cual el profesor gire sobre sí mismo con dos pesados libros en las manos, con los brazos extendidos, y en seguida los repliegue... ¡La aceleración del movimiento de rotación que se produce es notable! Recíprocamente, si se comienza girando con los brazos replegados, al extenderlos uno se frena. Los alumnos podrían verificar este fenómeno por sí mismos, idealmente, usando un taburete giratorio, por ejemplo. Se podría señalar lo que algunos conocen bien, a saber, que en la danza, los bailarines aprovechan este fenómeno para girar vertiginosamente sobre sí mismos.

Se podría mencionar aún, que si uno hace mediciones precisas, como los físicos, puede constatar, por ejemplo, que al disminuir el brazo de los dos libros a la mitad, ¡la velocidad de rotación no aumenta al doble sino que al cuádruple! Y que esto corresponde al hecho que la contribución de cada masa a la inercia a la rotación depende del producto de la masa por el *cuadrado* de su brazo (distancia al pivote de giro). En seguida, se podría continuar con la siguiente

Pregunta 2: ¿Cuál es el sistema de masas más sencillo imaginable que se comportaría como el de nuestra variable, teniendo la misma masa total (1 kilo), el mismo centro de masa o punto de equilibrio y la misma inercia a la rotación en torno al punto de equilibrio usado como pivote?

Debe estar relativamente claro que este sistema prototipo debe constar por lo menos de dos masas, que pivotan sobre el punto medio entre ellas, algo así como una pesa, como las usadas en halterofilia:



Cada masa será de medio kilo, y tendrá el mismo brazo b respecto al punto de pivote.

Es intuitivamente claro que achicando o agrandando b, sin variar la masa de medio kilo en las puntas, podemos imitar cualquier inercia a la rotación de un sistema de masas de peso total 1 kilo que gire en torno a su punto de equilibrio.

Según lo que recordamos antes, la contribución a la inercia a la rotación de esta pesita será

$$\mathbf{(0.5)b^2 + (0.5)b^2 = b^2}$$

y si nuestro sistema de masas consta, por otro lado, como en nuestro caso, de masas p_1 , p_2 , . . . , p_6 , con brazos b_1 , b_2 , . . . , b_6 , relativos al punto de equilibrio, entonces su inercia a la rotación estará dada por

$$\mathbf{p_1\, b_1^2 + p_1\, b_1^2 + \ldots + p_6\, b_6^2 .}$$

Escribiendo las igualdad de las inercias a la rotación, se obtiene

$$\mathbf{b^2 = p_1\, b_1^2 + p_1\, b_1^2 + \ldots + p_6\, b_6^{2.}}$$

lo que determina exactamente el valor de **b** (¡que es positivo, por supuesto!). Y este brazo **b** es justamente la famosa desviación típica, designada usualmente por la letra griega σ (sigma), en lugar de **b** .

En este punto, sugerimos animar a los estudiantes a usar su intuición psicomotriz para estimar el brazo σ de la pesita tipo que simula perfectamente la inercia a la rotación (llamada “momento angular” por los físicos) de los sistemas de masas probabilísticas asociados a D y a G. Posiblemente estimarán “a ojo” que el σ de D es algo mayor que 1.5 y el de G es un poco mayor que 1...

Calculando con la fórmula de más arriba se obtiene que la desviación típica de D es 1.71 y la de G sólo 1.12 .

Se puede hacer notar, para terminar, que en el caso de D, dos tercios de la masa total quedan a
trapados en el intervalo definido por la pesita, que va desde E(D) - σ hasta E(D) + σ , y que en el caso de G, la cantidad de masa probabilística atrapada en el intervalo análogo es de
20/32 = 0.625. Es decir, vemos que aunque las distribuciones de masa son bastante distintas, en todo caso, si uno se permite un error de σ hacia la derecha o hacia la izquierda del valor esperado, atrapará entre 60% y 70% de la masa total.

En otros términos, es bastante probable que los valores de nuestra variable aleatoria no se aparten en más de una desviación típica de su valor esperado.

2.2.5. Relaciones transversales con otros temas: Variables aleatorías geométricas:

Sugerimos proponer el siguiente juego a los alumnos:

Dibujar un hexágono regular y elegir al azar una de sus simetrías (descubrir primero que son 6 rotaciones, incluyendo la identidad, y 6 reflexiones). El premio del juego es de tantos miles de pesos como vértices del hexágono fije la simetría elegida. Si sacamos la identidad recibiremos entonces $6.000, pero hay otras simetrías del hexágono que no fijan ningún vértice...

Pregunta: ¿Cuál es el precio justo que uno aceptaría pagar por jugar este juego?

El "precio justo" es aquel que hace que el juego sea equitativo, ni ventajoso ni desventajoso.

Actividades sugeridas:

Preguntar a los alumnos cuál es su estimación intuitiva, "a ojo", de este precio justo.
¿Será $10.000, $6.000, $3.000, $2.000, $1.000, $500, $100, por ejemplo? Darles la oportunidad para que lleguen a percibir que ese precio justo debiera ser exactamente lo que ellos esperarian ganar, *en promedio*, o *idealmente*, en este juego. Y esta ganancia promedio, o ganancia esperada, debe ser una cantidad comprendida entre $6.000 y $0.

En seguida, no olvidar el abordaje experimental del problema: simplemente jugar el juego muchas veces y ver qué pasa. Aunando esfuerzos en un curso de 40 alumnos, si cada uno juega tan sólo 10 veces el juego, ¡podemos disponer de 400 rrepeticiones del juego!

Pero para jugar el juego, se plantea un problema: ¿Cómo elegir totalmente al azar, de una manera cómoda, una simetría entre las doce que tiene el hexágono regular? Se podría hacer con papelitos en una urna, por supuesto, pero puede ser más interesante sugerir a los alumnos que traten de usar objetos cotidianos, como monedas o ... dados. Una manera entre otras, sería por ejemplo, usar una moneda para decidir entre rotaciones y .reflexiones, y enseguida el dado para escoger cuál rotación o cuál reflexión.

A partir de la experimentación estadística, haciendo el promedio de los valores observados, los alumnos deberían poder llegar a una buen estimación de la esperanza de la variable aleatoria *Número de puntos fijos de una simetría del hexágono regular,* que abreviaremos por N.

Por otro lado, deberíamos animarlos a hacer el cálculo teórico de las probabilidades con que N toma sus diversos valores.

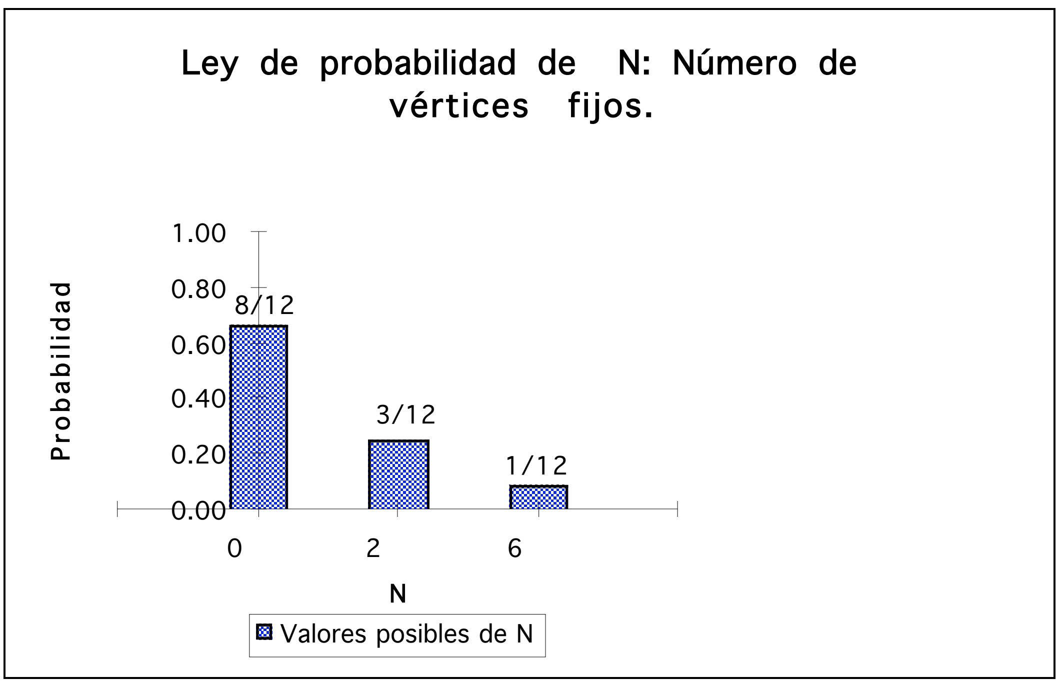
Teniendo en cuenta que hay una simetría que fija 6 vértices (la identidad), 3 reflexiones que fijan 2 vértices, 3 reflexiones y 5 rotaciones que no fijan ningún vértice, los alumnos podrían obtener sin mayor dificultad que:

- N vale 6 con probabilidad 1/12,
- N vale 2 con probabilidad 3/12,
- N vale 0 con probabilidad 8/12.

¡Notar que la suma de las probabilidades obtenidas da 1, como corresponde!

Como etapa siguiente, preguntar a los alumnos que harían. Esperamos que espontáneamente se interesen en hacer el "retrato" de N, es decir el gráfico de su ley de probabilidad, para "ver" mejor como se porta.

A partir de este gráfico podrán seguramente estimar "a ojo" el punto de equilibrio de la correspondiente distribución de pesos, y adivinar que está ubicado en 1, cosa que pueden verificar experimentalmente (disponiendo los pesos adecuados en una regla rígida) o también con el cálculo algebraico del promedio ponderado de los valores de N por sus respectivas probabilidades.



Preguntas ulteriores (Actividades suplementarias):

Creemos que conviene hacer notar a los alumnos que cada respuesta que encontramos para una pregunta, suscita a su vez otras preguntas, que hubiera sido difícil imaginar de antemano.

Así entonces, después de haber resuelto nuestro problema y haber encontrado el precio justo que pagar por este juego (a saber $1000), podríamos preguntar a los alumnos:

¿Y ahora qué? ¿Qué preguntas se les ocurre?

He aquí una lista, no exhaustiva, por supuesto, de preguntas:

- ¿Qué sucede si nos vamos a un pentágono, un cuadrado, u otro polígono regular?
- ¿Dependerá la ganancia esperada del número de vértices del polígono?
- ¿Qué sucede con el juego análogo en que nos paramos en un cuadrado y en lugar de mirar lo que le hacen las simetrías del cuadrado a los vértices, miramos lo que le hacen a las "líneas" del cuadrado, es decir sus aristas y diagonales. ¿Cuál será el número esperado de líneas fijas, entre aristas y diagonales?
- ¿Qué sucede si pasamos con la pregunta anterior a otros polígonos regulares, entendiendo que una "línea" pasa siempre por dos vértices del polígono.

Comentarios:

Para cualquier polígono regular el número esperado de vértices fijos de una de sus simetrías es igual a 1. En el caso de las líneas, el número esperado de líneas fijas es igual al número de "tipos" de líneas. Estos son 2 para el cuadrado (aristas y diagonales), 2 también para el pentágono, 3 para el hexágono y heptágono, y así sucesivamente. El resultado abstracto y general que estamos descubriendo en casos particulares, inductivamente, es que si un grupo de transformaciones actúa sobre un conjunto, el número esperado de puntos fijos que tiene una transformación del grupo es igual al número de órbitas del grupo en el conjunto (es decir, el número de "tipos de congruencia de los elementos del conjunto bajo la acción del grupo"¨).

Comentarios sobre evaluación:

Creemos que una alternativa interesante a los métodos más tradicionales puede ser proponer trabajos voluntarios, optativos, a los estudiantes, con un plazo limitado, que puedan ser resueltos por ellos y entregados para ser ¨pesados¨ en la evaluación del curso.

Esto puede estimular la iniciativa de los estudiantes, y su motivación, ya que trabajarían en un tema que es elegido por ellos, en alguna medida.

En nuestro caso, se podría proponer un abanico de variables aleatorias, entre las cuales los estudiantes podrían elegir, o incluso añadir otras por las que sientan más simpatía, para que investigaran su comportamiento.

Nos parece importante proponer problemas concretos y contextualizados y evaluar si el alumno logra llegar *por algún método* a una solución "honorable" al problema. Dejar abierto qué tipo de método emplee: podrá ser un método experimental, de simulación estadística, o teórico, construyendo un modelo adecuado. No desvalorizar la experimentación numérica, que a veces es el único método a nuestro alcance...

Un problema oral: ¿Cuántas caries?

Un examen dental sorpresivo hecho a un curso de Cuarto Medio arrojó los siguientes resultados, en cuanto a caries no tratadas:

- 3 alumnos con 0 caries,
- 8 alumnos con 1 caries,
- 14 alumnos con 2 caries,
- 11 alumnos con 3 caries,
- 6 alumnos con 4 caries,
- 2 alumnos con 5 caries,

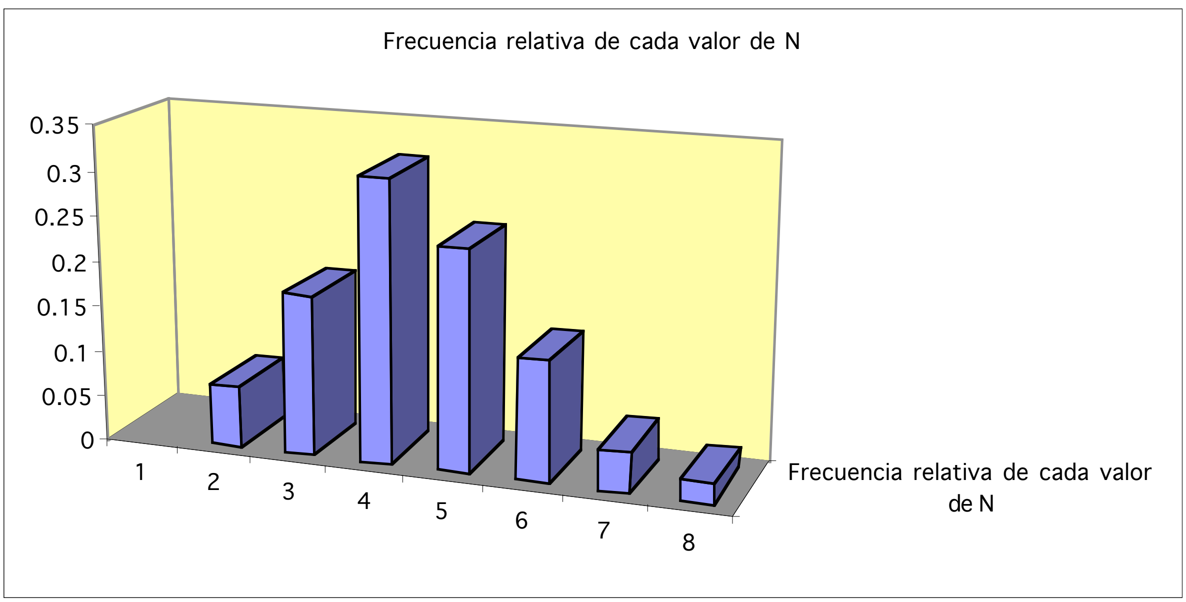
- 1 alumno con 6 caries.

a) Dibuja el gráfico de frecuencias relativas para estos datos.
b) ¿Cuál es la probabilidad que un alumno del curso elegido al azar tenga 2 caries?
c) ¿Cuál es el número esperado de caries que tiene un alumno del curso, elegido al azar?
d) ¿Cuál es la desviación típica del número de caries de un alumno del curso?
e) ¿Cuál es la probabilidad que el número de caries de un alumno elegido al azar no se aparte en más de una desviación típica del número esperado de caries?

Resolución:
a) Obtenemos las frecuencias relativas de cada número de caries dividiendo el número de alumnos con esa cantidad de caries por el número total de alumnos, a saber 45. Así debieras poder obtener la siguiente tabla, en forma decimal, por ejemplo, en que denotamos N el número de caries y f la frecuencia relativa:

N	0	1	2	3	4	5	6
f	0,0667	0,1778	0,3111	0,2444	0,1333	0,0444	0,0222

A partir de aquí, podemos dibujar el gráfico pedido.



b) ¿Te parece ahora que, además, este gráfico de frecuencias relativas se pueda interpretar como un gráfico de probabilidades? En otras palabras, ¿en qué sentido es el número de caries de un alumno de este curso una variable aleatoria? Bueno, recuerda que para definir o especificar una variable aleatoria siempre hay que precisar cuál es el experimento aleatorio que determina sus valores. En este caso, ¿cuál es ese experimento?

Efectivamente, es *elegir un alumno del curso al azar (tan posiblemente uno como otro),* por ejemplo sacando papelitos con los nombres de una urna. Una vez que esto está claro, vemos que la probabilidad que N valga 2, por ejemplo, es el cuociente entre el número de "casos favorables", a saber 14, y el "número total de casos", a saber 45, es decir 14/45 = 0,3111....

c) Mirando el gráfico de más arriba, posiblemente estimaste que el valor esperado E(N) de N, mirado como punto de equilibrio del sistema de pesos correspondiente, está más o menos . . . a mitad de camino entre 2 y 3. Calculándolo como promedio ponderado de sus valores posibles por sus probabilidades (expresadas todas como fracciones de denominador 45) se obtiene 109 / 45 = 2,4222,,, = 2,4 aproximadamente.
d) Mirando de nuevo nuestro sistema de pesos probabilistas te podrías atrever a estimar qué brazo debiera tener la pesita que tiene su misma inercia al giro en torno al punto de equilibrio 2,42 como pivote. Recuerda que ese brazo es la desviación típica σ.
¿Por lo menos, dirías que anda entre 1 y 1,5?
Veamos, calculando abnegadamente los respectivos brazos y sus cuadrados, resulta

$$\begin{aligned}\sigma^2 &= (3/45)\cdot(2{,}42)^2 + (8/45)\cdot(1{,}42)^2 + (14/45)\cdot(0{,}42)^2 + (11/45)\cdot(3-2{,}42)^2 + \\ &\quad (6/45)\cdot(4-2{,}42)^2 + (2/45)\cdot(5-2{,}42)^2 + (1/45)\cdot(6-2{,}42)^2 \\ &= (3/45)\cdot 5{,}856 + (8/45)\cdot 2{,}016 + (14/45)\cdot 0{,}176 + (11/45)\cdot 0{,}336 + \\ &\quad (6/45)\cdot 2{,}496 + (2/45)\cdot 6{,}656 + (1/45)\cdot 12{,}816 \\ &= 80{,}960/45 = 1{,}799,\end{aligned}$$

es decir, $\sigma = \sqrt{1{,}799} = 1{,}341 = 1{,}34$ aproximadamente.
¡No andaba tan mal la estimación “a ojo”!

e) Conviene fijarse bien hasta donde se alcanza si uno se para en el punto de equilibrio 2,42 y suma y resta un σ es decir, se corre en un σ hacia la derecha y en un σ hacia la izquierda:

hasta 2,42 + 1,34 = 3,8 y 2,42 - 1,34 = 1,08.

Entonces, los valores de N que caen en este “intervalo de tolerancia” son sólo 2 y 3. El valor 1 queda afuera por un pelito... Pero estos dos valores ya acaparan mal que mal un peso probabilista de

14/45 + 11/45 = 25/45 = 0,555... = 0,56 aproximadamente,

que es la respuesta a la pregunta.

Nota que si preguntaras qué valores de N quedan comprendidos entre E(N) + 2σ y E(N) – 2σ la respuesta sería . . . ¡todos, menos 6, y la probabilidad que el número de caries de un alumno no se aparte en más de 2σ del valor esperado es nada menos que 44/45 = 0,98 aproximadamente

Un problema viral: test de hipótesis en salud pública.
He aquí un ejemplo de problema concreto, que podríamos proponer a los alumnos para que puedan experimentar por sí mismos la a veces difícil lectura de los datos estadísticos. Es susceptible además de, por lo menos, un abordaje estadístico y uno probabilista…

ACTIVIDAD: ¿Cómo tratar el virus?

¿Qué haría Ud. para tomar una decisión frente al siguiente problema?

En un cierto lejano país, ha hecho su aparición un alevoso virus, que mata a los seres humanos que infecta con probabilidad 1/2, es decir, tan posiblemente sí como no.

A lo largo del país, frente a lo intratable del virus por los métodos tradicionales, se ha recurrido a diversos métodos de medicina alternativa para tratar de salvar a los pacientes infectados : acupuntura, visualización, homeopatía y hierbas, entre otros.
He aquí los datos estadísticos que el Ministro de Salud Pública ha podido recoger hasta ahora en relación con la presunta eficacia de estos métodos:

Con un tratamiento de acupuntura, de 45 infectados, sanaron 30.
Con un método de visualización positiva, de 120 infectados sanaron 70.
Con un tratamiento homeopático, de 600 sanaron 324.
Con un tratamiento con hierbas autóctonas, de 1800 infectados, sanaron 945.

¿Qué les dicen estos datos?
A la luz de ellos, ¿se puede concluir en la eficacia de alguno de los métodos empleados?
¿Hay unos más eficaces que otros?
¿Cómo podemos discernir objetivamente cuáles merecen ser promovidos?

Queremos aprender cómo tomar decisiones sensatas en situaciones de este tipo.
Formemos grupos de a 5 en promedio y discutamos cómo podríamos abordar este problema, sin mayores conocimientos de estadística, armados de nuestro solo sentido común.
¿ De qué manera podríamos estimar la eficacia relativa de los 4 métodos alternativos de más arriba.?
¿Cuáles métodos cree cada uno que son más eficaces, a la luz de los datos recogidos?
¿Qué criterios usan para justificar su creencia?
Un par de sugerencias:
- Comparar con lo que ocurre al lanzar una moneda.
- Considerar casos "extremos" como el siguiente: Con un quinto tratamiento recién aparecido, de 2 infectados, ¡sanaron todos! ¿Cómo funcionan los criterios que propuso cada uno en este caso? ¿Les parecen razonables las conclusiones a que llegan?

3. ¿Cuántos de cada uno? Estadística, historia, ciencias naturales y sociales.

3.1. Al comienzo, eran los censos…

Desde antes que existieran los Estados sobre la faz de la tierra, el hombre hacía "estadísticas", contando animales, personas, u objetos diversos, con ayuda de símbolos dibujados sobre las paredes de las cavernas , maderas, pieles…Ya en el tercer milenio a.C., los babilonios, con su sistemático espíritu, hacían tablas de datos de producción agrícola o trueque de mercancías, con ayuda de tablillas en arcilla. Por ese tiempo, también los egipcios y los chinos se ocupaban de reunir datos demográficos y pecuniarios. En la Biblia, hay un libro que por algo lleva el nombre de "Números": contiene dos censos de la población de Israel.
Los griegos del 600 a.C., utilizaban la información obtenida en los censos de población para recaudar tributos. Más tarde, el imperio romano, muy interesado en impuestos y tributos, continuó en esta dirección, recopilando una multitud de datos sobre la población, superficie y pecunio de sus extensos dominios. Durante la Edad Media, en Europa, se llevó a cabo algunos censos acuciosos , como aquellos de las propiedades de la Iglesia, ordenados en 758 y 762por los reyes Pepino el Breve y Carlomagno. A comienzos del segundo milenio, Guillermo I de Inglaterra ordena realizar un censo nacional, que será publicado en el Domesday Book en 1086.
¿Ves ahora por qué la Estadística lleva ese nombre? ¡Las primeras estadísticas eran "las cuentas del Estado"!
Hasta comienzos del siglo XVII, sin embargo, esta Estadística seguía siendo puramente *descriptiva .* Consistía en enumerar y tabular datos, ordenada y

sistemáticamente., nada más. Por lo demás, es lo que comenzamos haciendo en la unidad sobre Estadística, sólo que utilizando mucho más los recursos gráficos que los estadísticos de aquel tiempo. ¿No te parece?

3.2. Luego, se comenzó a inferir…

A mediados del siglo XVII aparece alguien que ve más lejos: un pequeño comerciante londinense, llamado John Graunt, quien publicó un interesante librito de comentarios sobre las partidas de defunción ("Bills of Mortality") que publicaba semanalmente su parroquia. Este texto le valió ser promovido miembro de una respetada sociedad científica por el rey y fue el germen del desarrollo de las estadísticas vitales (registro sistemático de nacimientos y defunciones) en Inglaterra y otros países europeos. Así, poco después, el famoso astrónomo inglés Edmund Halley calcul ó la primera tabla de mortalidad, apoyándose en una estadística de defunciones realizada en Breslau (Alemania) en 1691.
El gran pionero del desarrollo de la estadística fue Thomas Bayes, quien logró aplicar la naciente teoría de probabilidades para hacer predicciones apoyándose en experimentos estadísticos previos. Es conocido hoy día sobretodo por su famoso Teorema sobre "las probabilidades de las causas ", que data de 1768.

¡HAZLO TU!
Te proponen el siguiente juego: eliges al azar, tirando una moneda, entre dos urnas, de idéntico aspecto, una de las cuales contiene una bolita roja y dos blancas, y la otra, una bolita blanca y dos rojas. Luego, sacas una bolita al azar de la urna seleccionada. Ganas si sale roja.
Pregunta: Si ganaste, ¿cuán probable es que extrajiste tu bolita de la urna que contiene dos rojas y una blanca? Es decir, si ganaste, ¿cuál es la probabilidad que la" causa" haya sido esa urna?
Sugerencia: Si no ves claro como razonar al estilo de Bayes, ¡haz una simulación estadística!
Así podrás calcular el porcentaje de veces que la bolita roja provino de la urna con mayoría roja, cuando ganaste

Las reflexiones de Bayes, contienen ya en germen el modo de pensar que habría de constituir el mayor avance de la estadística en el próximo siglo: aquel de *la inferencia estadística*, que va más lejos que la mera *descripción estadística*.

Se trata ahora, de inferir, o hacer predicciones, aunque sea estimativas, sobre la composición de una población o universo, a partir de observaciones estadísticas sobre muestras extraídas de ella.

La situación lúdica de recién, con esas dos urnas, de distinta composición puede servir de modelo para una situación de la vida cotidiana en que tu tuvieras que elegir entre dos alternativas para la posible composición de la población de peces en un lago, por ejemplo. A saber, aquella en que priman los peces rojos sobre los blancos en la proporción 2:1, o aquella en que priman los blancos sobre los rojos en la misma razón 2:1. Tu has pescado un pez del lago, que resultó ser un pez rojo.
Te preguntas entonces, ¿cuán probable es que la composición del lago sea la primera, en lugar de la segunda alternativa?
La respuesta a que debes haber llegado, es que esa probabilidad es 2/3.

Miremos un ejemplo con números un poco más cotidianos:

EJERCICIO RESUELTO

¿Qué porcentaje de rojas hay?

Tienes una urna con 100 bolitas, de las cuales 50 son rojas y 50 blancas. ¿Cómo podrían tus compañeros y compañeras inferir la composición de la urna, si tu los dejas sacar muestras pequeñas , pero sin mirar dentro de la urna?

Resolución :

Supón que sacan, por ejemplo, *muestras de 10 bolitas*. Llama R el número de bolitas rojas de la muestra. ¿Podrías calcular la ley de probabilidad de la variable aleatoria R? Bueno, recordando como contabas muestras en Segundo y Tercero Medio, puedes descubrir con ayuda de la calculadora, redondeando a 4 decimales, que:

$$P[R=5]=\frac{\binom{50}{5}\binom{50}{5}}{\binom{100}{10}}=0,2593,$$

en tanto que

$$P[R=4]=\frac{\binom{50}{4}\binom{50}{6}}{\binom{100}{10}}=0,2114\,.$$

Fíjate que no necesitas trabajar de nuevo para calcular P[R = 6], porque coincide P[R = 4], ¿cierto?.

Así podrás ver - ¡hazlo tú! - que la probabilidad que R valga *k* va decreciendo a medida que el entero *k* se aparta de 5 hasta llegar a los casos extremos e inverosímiles en que ninguna, o todas, las bolitas de la muestra son rojas, que tienen probabilidad :

P[R = 0] = P[R = 10] = 0,0006.

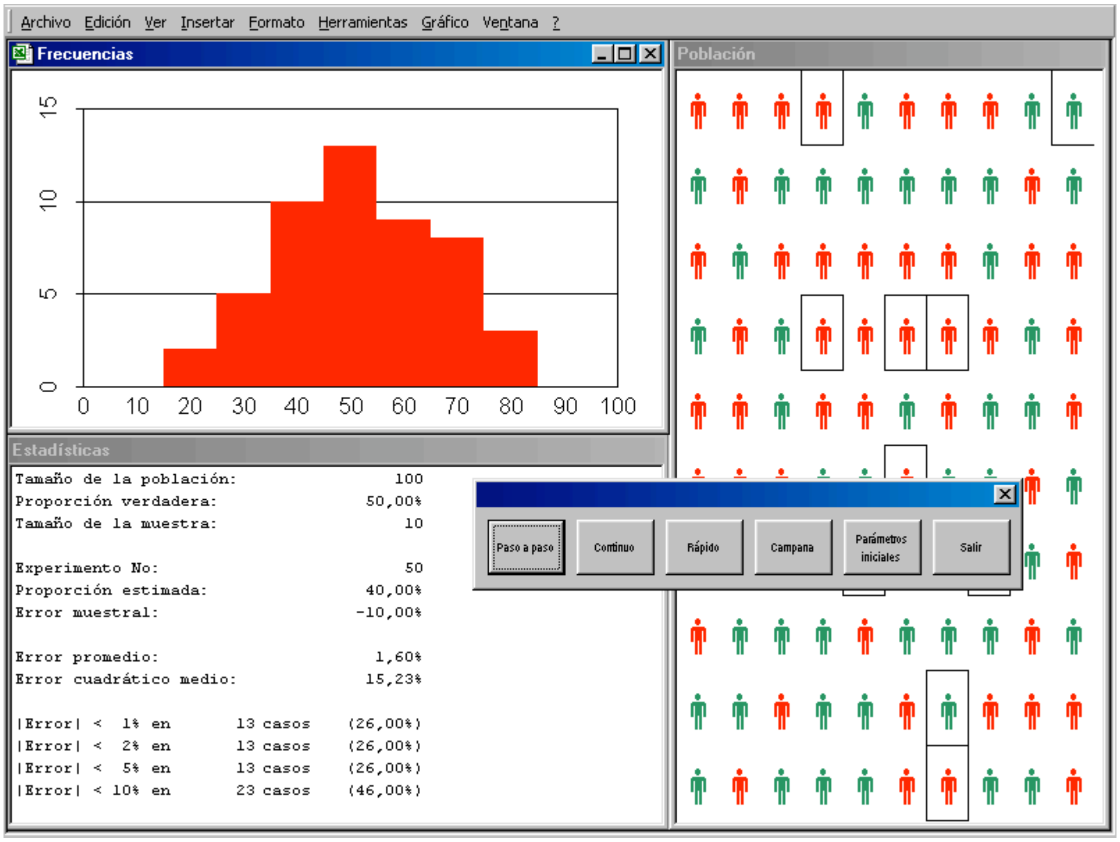
Podemos concluir entonces que la muestra más probable es aquella que contiene el mismo porcentaje de bolitas rojas que la población entera, lo cuál es bastante alentador para tus compañeros.

Sin embargo, seguramente – como está reclamando ya Segismundo – sería un poco aventurado que hicieran una predicción sobre el porcentaje de rojas en la urna luego de haber extraído sólo *una* muestra. Porque la probabilidad de tener 50% de rojas en la muestra no es abrumadoramente mayor que la de tener 40% o 60%, por ejemplo.

Para que "se note" realmente que es más probable sacar 50% de rojas que 40%, ¿qué deberíamos hacer?

¡Repetir bastantes veces el experimento, como de costumbre en estos casos!

Te podemos contar que cuando los 50 alumnos de un curso sacaron una muestra cada uno, y tabularon y graficaron el porcentaje de rojas en la muestra, obtuvieron lo siguiente:

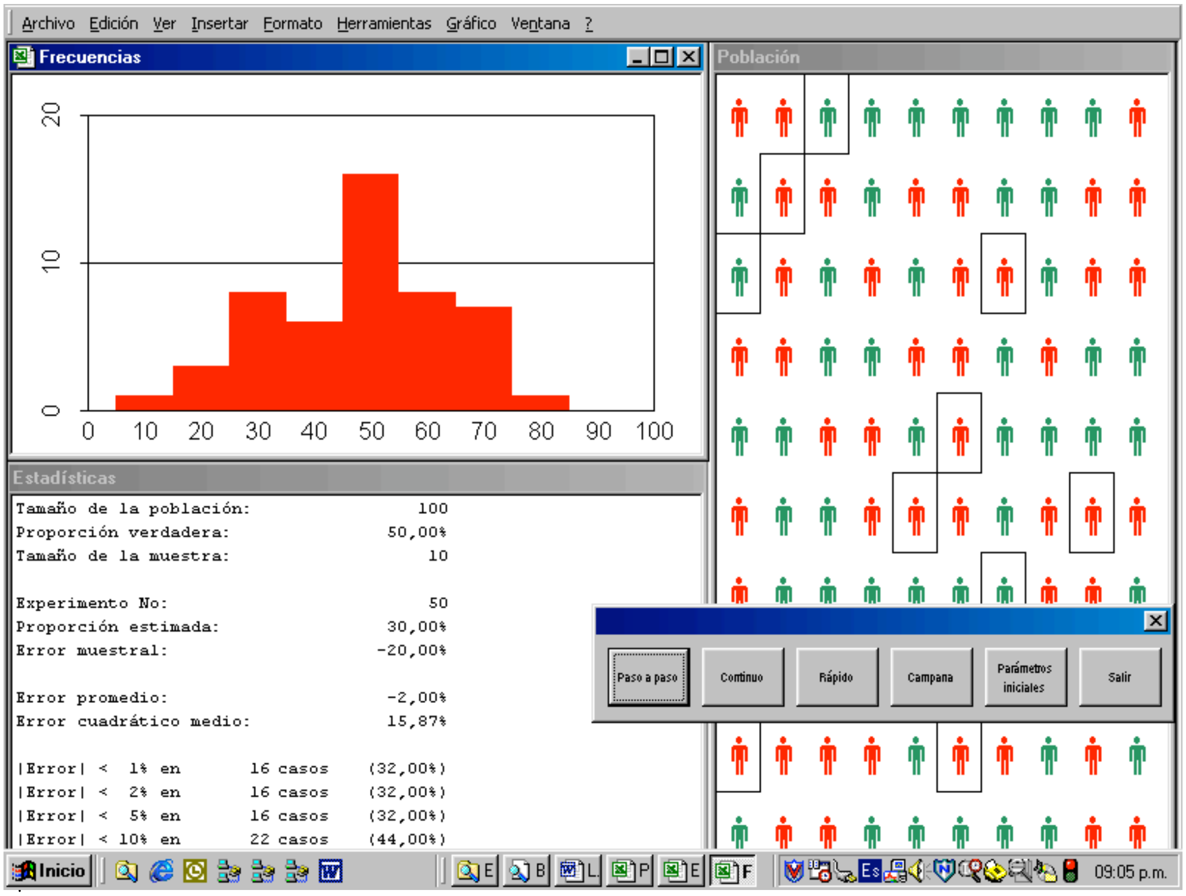


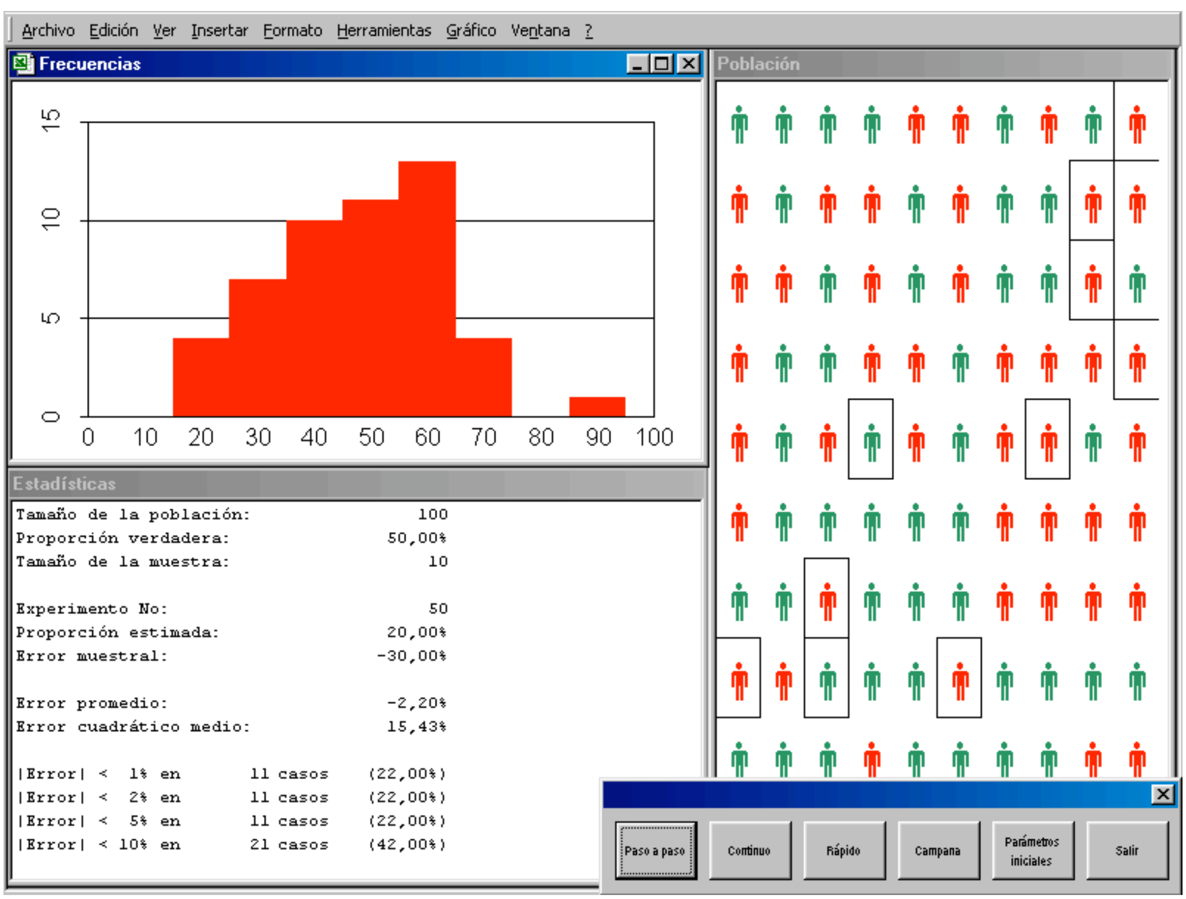
Como ves, por un lado el valor más frecuente del porcentaje de rojas en la muestra es justamente 50% y por otra parte, si calculamos el promedio de los porcentajes de rojas en las 50 muestras, vemos que obtendremos el porcentaje correcto con un error del 1,6% apenas.

OBSERVA
Fíjate que el error del promedio de los porcentajes de rojas de las 50 muestras es lo mismo que el promedio de los errores de que "cometen" los porcentajes de rojas de cada una, cuando tratan de adivinar el porcentaje de rojas en la población total. Nuestro programa de simulación nos da cada vez este último *error promedio.*

HAZLO TU
Sea "a mano" o con un programa de simulación, como "sampling" (que puedes encontrar en el sitio WWW.UNAP.CL) , saca tú mismo 50 muestras de 10 bolitas de la urna y ve que pasa…

Además otros cursos de 50 alumnos, cuando supieron, se entusiasmaron y empezaron también a sacar muestras.. En las pantallas de más abajo puedes ver cómo les fue:


Archivo Edición Ver Insertar Formato Herramientas Gráfico Ventana ?
Frecuencias
Población
Estadísticas
Tamaño de la población: 100
Proporción verdadera: 50,00%
Tamaño de la muestra: 10
Experimento No: 50
Proporción estimada: 30,00%
Error muestral: -20,00%
Error promedio: -2,00%
Error cuadrático medio: 15,87%
|Error| < 1% en 16 casos (32,00%)
|Error| < 2% en 16 casos (32,00%)
|Error| < 5% en 16 casos (32,00%)
|Error| < 10% en 22 casos (44,00%)
Paso a paso
Continuo
Rápido
Campana
Parámetros iniciales
Salir
Inicio
09:05 p.m.

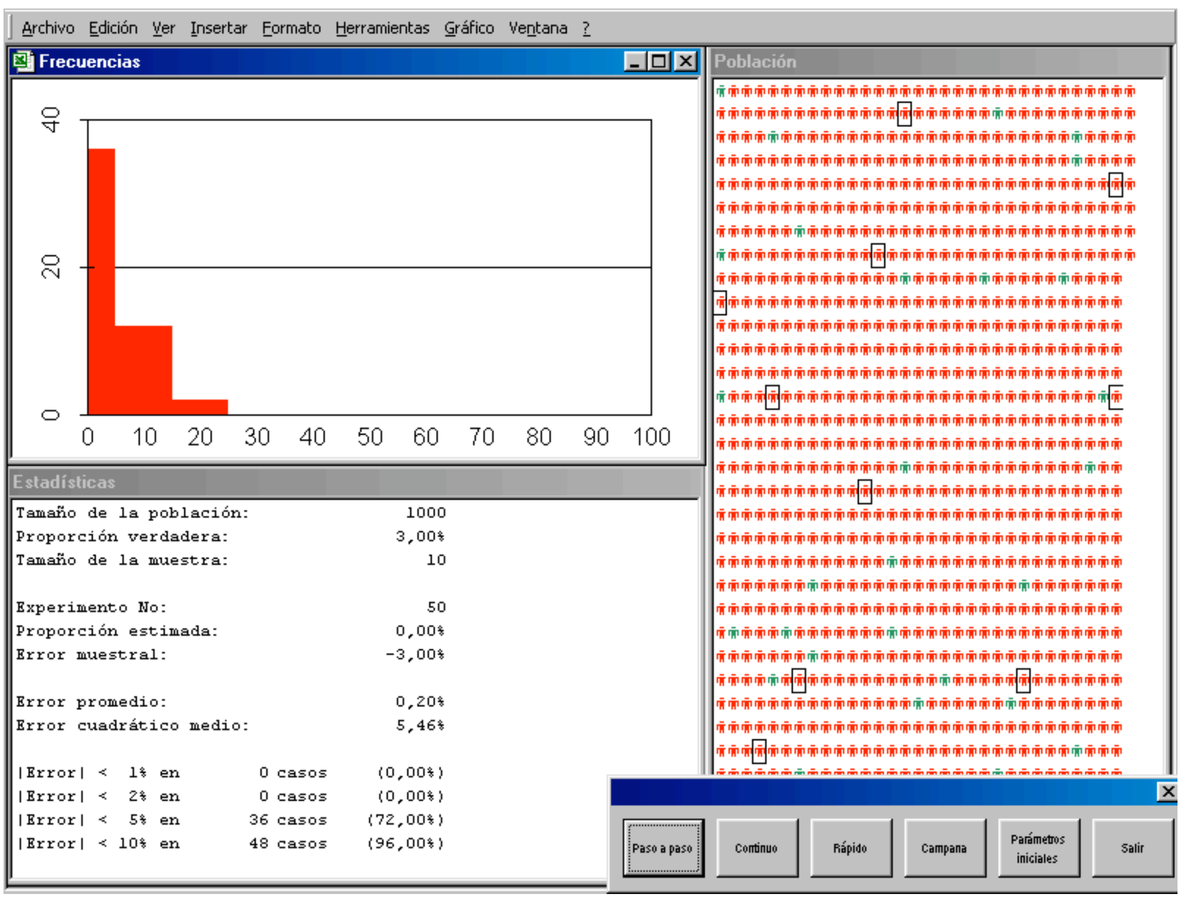

Archivo Edición Ver Insertar Formato Herramientas Gráfico Ventana ?
Frecuencias
Población
Estadísticas
Tamaño de la población: 100
Proporción verdadera: 50,00%
Tamaño de la muestra: 10
Experimento No: 50
Proporción estimada: 20,00%
Error muestral: -30,00%
Error promedio: -2,20%
Error cuadrático medio: 15,43%
|Error| < 1% en 11 casos (22,00%)
|Error| < 2% en 11 casos (22,00%)
|Error| < 5% en 11 casos (22,00%)
|Error| < 10% en 21 casos (42,00%)
Paso a paso
Continuo
Rápido
Campana
Parámetros iniciales
Salir

¿Notas que aunque el porcentaje correcto no sea la moda, de todas maneras al hacer el promedio de los porcentajes de rojas de las muestras, se obtiene una aproximación cada vez mejor de l porcentaje real de bolitas rojas en la población, cuanto más muestras se tome. ¡Mira el promedio de errores correspondiente!
Por otra parte , también importa el tamaño de la muestra: Con 50 muestras de 20 bolitas se obtiene una mejor aproximación del porcentaje real que con 50 muestras de 10. ¡Experimenta!

ACTIVIDAD

¿Y ahora, que porcentaje de rojas hay?

Dividan el curso en dos grupos de igual tamaño. El primer grupo prepara una urna con 1000 bolitas (u otros objetos similares entre sí), de los cuales un pequeño porcentaje, por ejemplo el 3%, son de un color , digamos rojo, y el resto de otro, digamos blanco. El segundo grupo, sin conocer la composición de la urna, trata de inferirla extrayendo muestras de 10 bolitas. Gana, si logra adivinar la composición en menos de media hora. En seguida, los grupos intercambian los roles. El segundo grupo prepara ahora la urna, con un porcentaje eventualmente distinto de bolitas rojas y el primero trata de adivinarlo, infiriendo estadísticamente.
En las figuras que siguen puedes apreciar un ejemplo de muestreo para un porcentaje de rojas (= porcentaje de individuos que poseen un cierto atributo) del 3%, simulado con el programa "sampling" disponible en WWW.UNAP.CL.



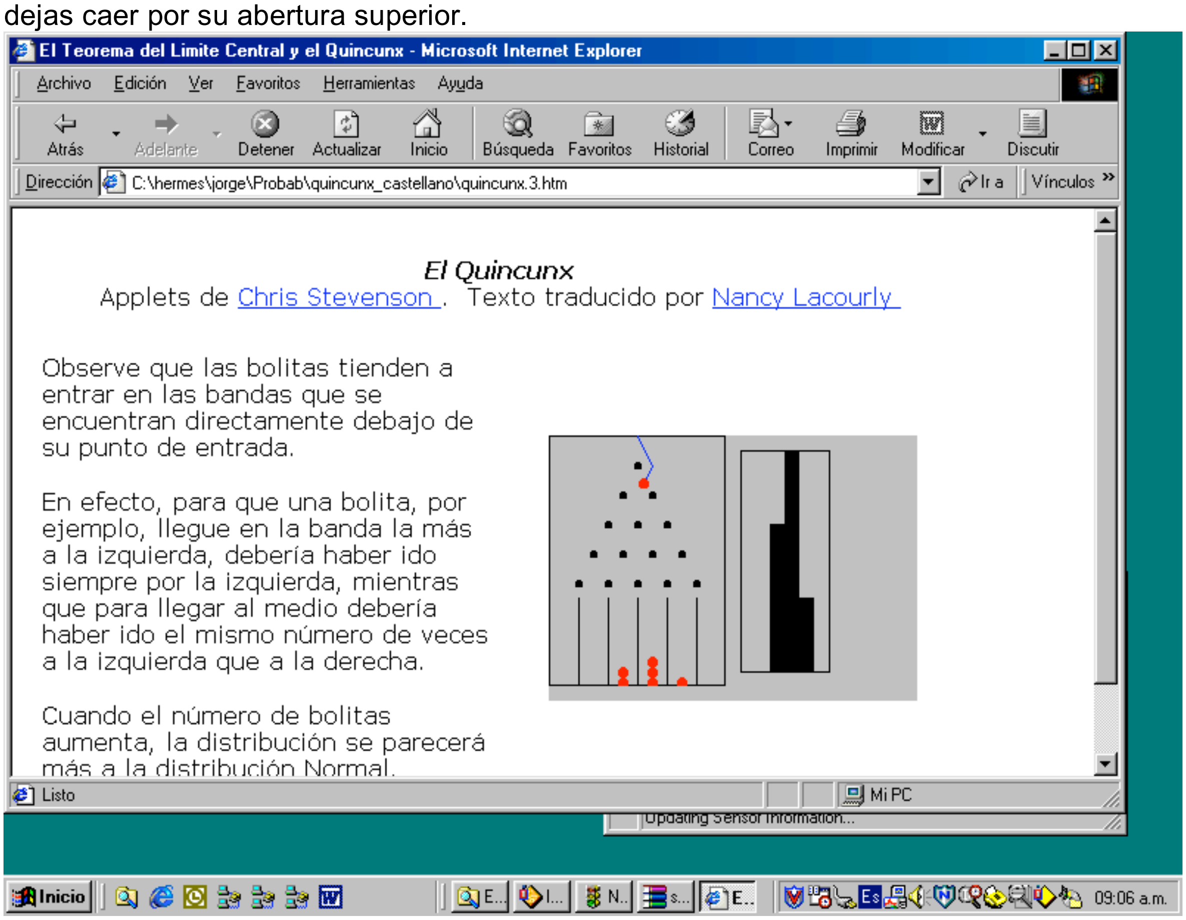
3.3. Los inevitables errores…

Un siglo más tarde, Carl Friedrich Gauss, el príncipe de las matemáticas, se percató que al realizar varias veces una misma observación astronómica o geodésica, obtenía valores distintos, debido – según parecía – a errores humanos o instrumentales.
¿Cómo proceder para elegir un solo valor, que fuese lo más confiable posible, pensando que la fluctuación del error de medición es irremediablemente aleatoria?
Gauss resolvió este problema de "estadística de errores" inventando su famoso método de los *mínimos cuadrados*, en que se busca un valor "conciliador", cuya suma de los cuadrados de las diferencias con los valores observados sea lo más pequeña posible.
Estudiando la distribución de los errores de medida, Gauss encontró la curva que hoy día se llama familiarmente "campana de Gauss"

ACTIVIDAD

Una aproximación física a la campana de Gauss.

Para aproximarse concreta y físicamente a la famosa campana de Gauss, te sugerimos armar un artilugio como el de la figura, llamado "plancha de Galton", en madera, por ejemplo, proveerlo de una tapa de vidrio, ponerlo vertical y "alimentarlo" con bolitas que dejas caer por su abertura superior.

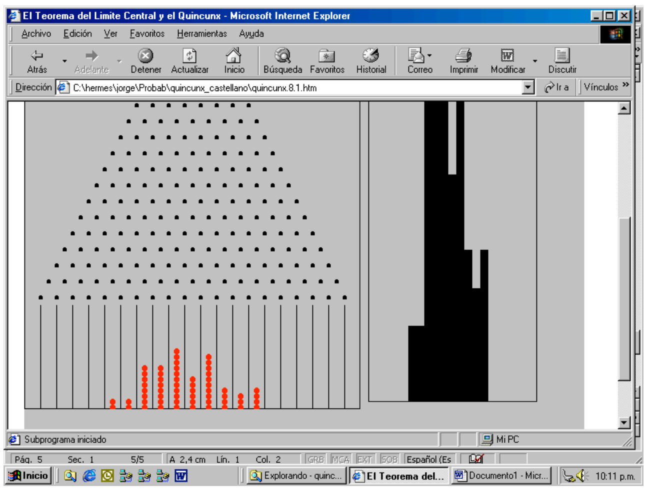


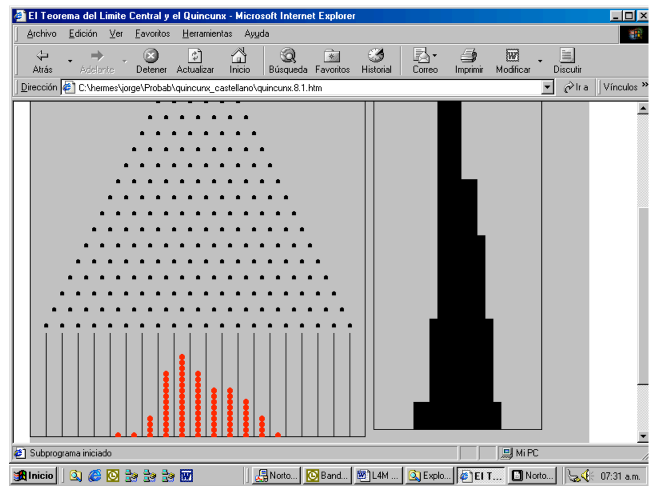
¿Puedes adivinar , por anticipado, cómo se distribuirán las bolitas en los casilleros? ¡Seguramente no en forma pareja!
Fíjate que se trata de un *paseo al azar*, como los que estudiamos en Segundo y Tercero Medio.

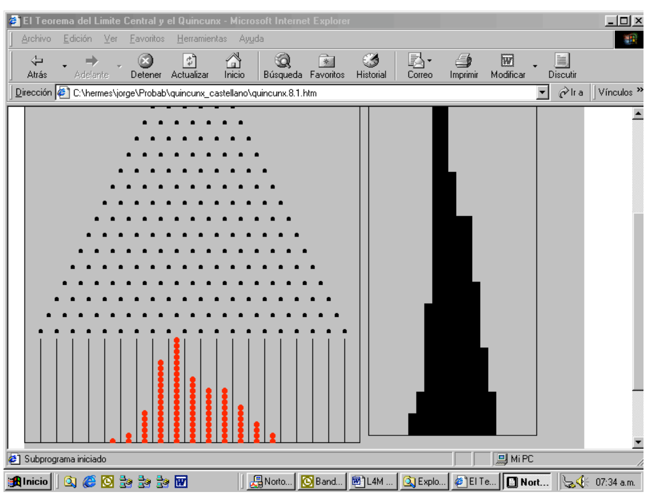
Podrías calcular paso a paso la probabilidad con que una bolita aleatoria aterriza en cada uno de los casilleros de la base. Por ejemplo, soltando un grupo de 64 bolitas, imaginando que "idealmente" se van dividiendo por partes iguales en cada bifurcación, contando cuántas llegan a cada casillero de la base, y calculando el porcentaje del total que eso significa… Ahora deja caer un gran número de bolitas y observa que pasa.

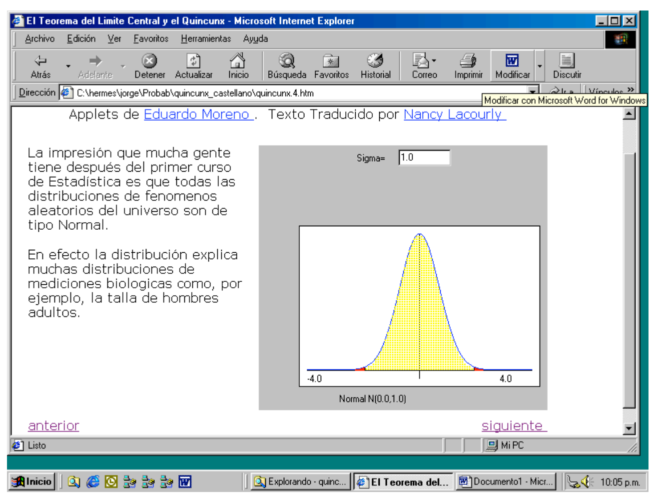
La distribución teórica o "ideal" de probabilidades así obtenida , se llama *distribución binomial.*

Cuando el número de casilleros de tu plancha de Galton aumenta, te vas acercando al perfil de la "famosa campana de Gauss" o curva normal (mira la Tabla 1, en anexo). He aquí algunos ejemplos, de simulaciones realizadas en el computador.









Si tienes la posibilidad, entra al sitio WWW.UNAP.CL y utiliza el material allí disponible (bajo el nombre Quincunx = quinconcio) para simular la plancha de Galton.

¿Qué podemos inferir? ¿Cómo tomar las muestras? : Inferencia estadística y diseño de experimentos.

En el siglo XX vemos ya establecerse la estadística como ciencia autónoma con las contribuciones enjundiosas de los británicos Karl Pearson (1857 - 1936), escritor y político además de matemático, y Sir. Ronald Fisher (1890 – 1962), distinguido genetista, además de estadístico.

El primero investigó sobre los modelos matemáticos de los mecanismos de la evolución biológica y la herencia e inventó uno de los tests más utilizados en estadística hoy día: el llamado test de Pearson o test de .

El segundo comenzó ocupándose, en los años 20, del diseño de experimentos en agricultura para llegar a desarrollar una teoría genera l de diseño de experimentos: ¿Cómo tomar muestras de modo que nos proporcionen una información lo más fiable posible sobre la población que queremos estudiar?

ACTIVIDAD

Júntate con 4 otros compañeros o compañeras y discutan cómo deberían diseñar la toma de muestras para

a) Hacer una estadística de los ingresos de los habitantes de Santiago

b) Hacer una estadística de los ingresos de los habitantes de todo Chile.

c) Hacer una estadística de la opinión de los ciudadanos en edad de votar, sobre la gestión educacional del gobierno.

d) Hacer una encuesta sobre la opinión de los habitantes de Santiago sobre la utilidad del metro.

e) Averiguar si los habitantes de Chile estiman que el país está excesivamente centralizado en Santiago.

f) Estimar la cantidad de peces de una cierta especie en vuestro lago favorito del Sur

g) Estimar la cantidad total de peces del lago.

Alfonso estuvo tomando muestras de peces en su lago favorito y descubrió – según él – que no había peces de menos de 30 cm de largo. ¿Le creerían?

Teresa encuestó 100 personas para realizar d) y descubrió que 85 pensaban que el metro es poco útil. Cuando le preguntamos cómo eligió a las personas, nos dijo que los eligió al azar entre los pasajeros de los microbuses que toma durante el día. ¿Qué les parece?

Resumiendo: Las tres caras de la estadística

Hoy en día la estadística juega un rol de apoyo fundamental para las ciencias naturales, humanas y sociales . Cuando éstas tratan de validar sus teorías y modelos a partir de datos experimentales, se enfrentan a las preguntas siguientes, para las cuales la estadística aporta algunas respuestas bastante eficaces:

- ¿Cómo recoger selectivamente los datos relevantes?
  Respuestas: Teoría de muestras y diseño de experimentos.
- ¿Cómo organizar, resumir e interpretar la montaña de datos así reunida?
  Respuesta: Estadística descriptiva
- ¿Qué conclusiones podemos extraer de nuestros datos ?
  Respuestas: Estadística inferencial, test de hipótesis o docimasia.

Para responder la tercera pregunta , se hace un uso amplio y profundo de la teoría de probabilidades.

Habrás notado que en la unidad previa, hemos practicado sobretodo la Estadística Descriptiva. El desarrollo de los métodos generales de la Estadística Inferencial sobrepasa los objetivos de este año. Por esto, vemos ahora sólo algunos ejemplos y actividades ilustrativas. como las anteriores y las siguientes.

¡Infiere tú mismo!
Para que puedas experimentar por ti mismo el sabor de la inferencia estadística, te proponemos la siguiente

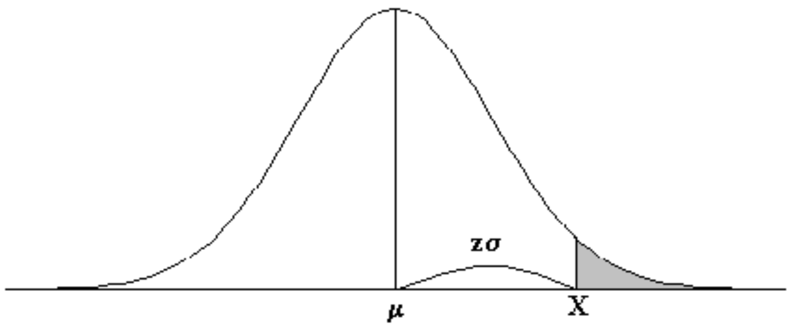
ACTIVIDAD
¿Cuál es la composición de la población?
Júntense de a parejas. Uno llena una urna con 100 bolitas, de dos colores distintos, por ejemplo: blancas y rojas. Sin que el otro vea, elige el porcentaje de bolitas rojas que pondrá en la urna, de modo que tenga un número exacto de decenas. Por ejemplo, 20%, 70%, 50%, sin excluir las opciones 0% y 100%. El otro jugador trata de adivinar el porcentaje de rojas, sacando muestras de la urna. Puede sacar el número de muestras que quiera, pero pequeñas, de no más de 10 bolitas cada una. Tiene 15 minutos de tiempo para experimentar. Luego emite su hipótesis sobre el porcentaje real de bolitas rojas en la urna, teniendo en cuenta que es un múltiplo de 10. Gana premio si acierta y recibe una mención honrosa si yerra en un 10%, no más. En seguida, invierten los roles.
Esta actividad probablemente les permita comenzar apreciar la efectividad del "muestreo" para estimar la composición de una población grande, en particular, el porcentaje de individuos portadores de un cierto atributo.
¿Pueden imaginar una docena situaciones de la vida cotidiana que puedan ser modeladas por un muestreo de este tipo?

REFERENCIAS

- P. González – J. Soto Andrade Matemática Activa, Segundo Medio, Mineduc - Editorial Marenostrum, Santiago 2002
- P. González – J. Soto Andrade Matemática Activa, Tercero Medio, Mineduc - Editorial Marenostrum, Santiago 2002
- P. González – J. Soto Andrade Matemática Activa, Cuarto Medio, Mineduc - Editorial Marenostrum, Santiago 2002
- N. Lacourly, J. Soto Andrade, "¿Qué dicen los datos?: Introducción a la estadística", Curso Taller PPF sobre Estadística, U. de Chile, Santiago, 2002
(introducción sistemática a la estadística, más avanzada que estas notas)
- J. Soto Andrade, "La cognición hecha cuerpo florece en metáforas…¨, en A. Ibañez, & D. Cosmelli, (Eds.), « Nuevos Enfoques de la Cognición: Redescubriendo la dinámica de la acción, la intención y la intersubjetividad"intersubjetividad », Universidad Diego Portales, Santiago, 2007.
- J. Soto Andrade, "Al azar del cara o sello: Introducción al Cálculo de Probabilidades", Módulo de Probabilidades para los GPT, 80 p., Mineduc 2001

TABLA 1: DISTRIBUCIÓN NORMAL

Áreas bajo la curva normal



Ejemplo:

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

$P\,[Z > 1] = 0.1587$

$P\,[Z > 1.96] = 0.0250$

Desv. normal x	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641
0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3121
0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
0.7	0.2420	0.2389	0.2358	0.2327	0.2296	0.2266	0.2236	0.2206	0.2177	0.2148
0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010