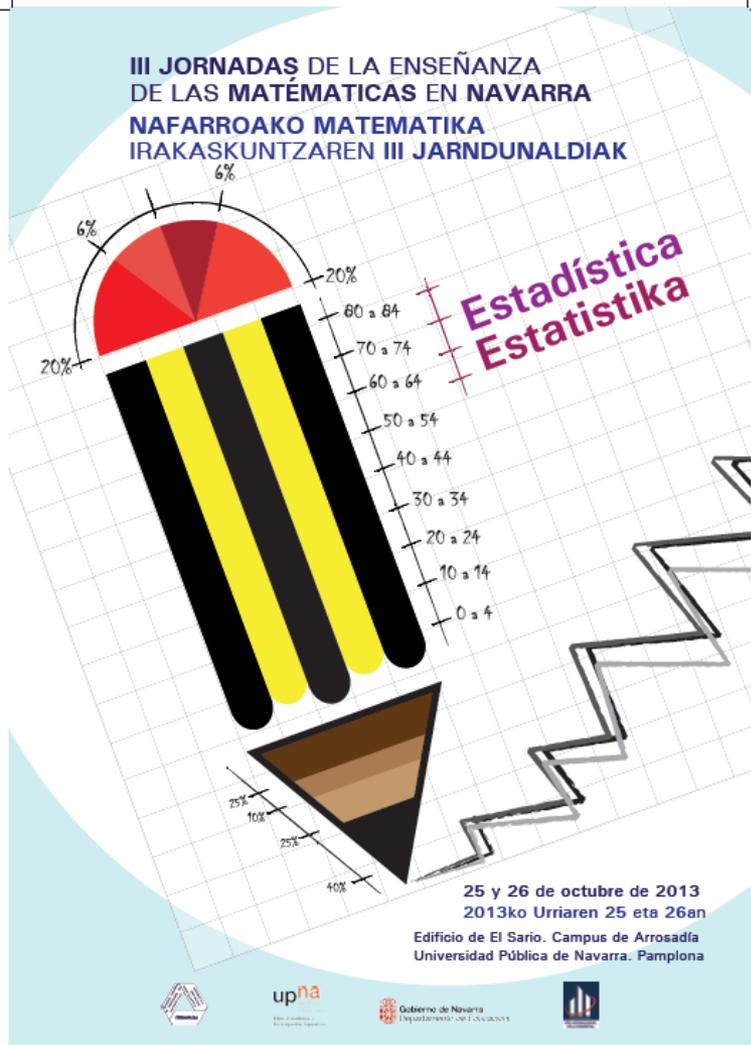


¿ Es predecible el azar?



Santiago Fernández
Asesor de matemáticas
Berritzegune Nagusia-Bilbao
Pamplona, 26, Octubre, 2013

*El año 2013
es el Año Internacional
de la Estadística*



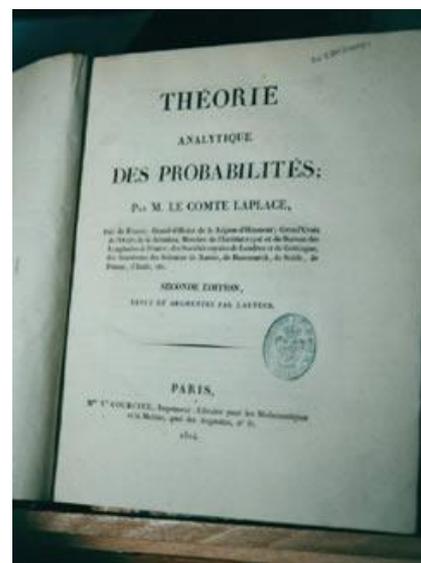
*y no es la típica
desviación de la moda.*



J. Bernoulli (1654-1705)

Ars Conjectandi (Arte de la conjetura),
1713

En **1812**, con la *Teoría Analítica de las Probabilidades*, expone los principios y las aplicaciones de lo que él llama "*geometría del azar*".



"En el fondo, la teoría de las probabilidades es sólo sentido común expresado con números".

"Hoy nuestra única certeza es la incertidumbre"



Zygmunt Bauman

Una aparente paradoja

Aunque la estadística y el azar se enseña en TODOS los niveles educativos la mayoría de los estudiantes finalizan los cursos sin aplicar ni conocer los conceptos y procedimientos estadísticos

Dos reflexiones:

- 1.- La relación entre la estadística y la “matemática” no es biunívoca: la estadística aplica los conocimientos matemáticos para el desarrollo de sus métodos, en cambio la matemática no usa conceptos estadísticos.
- 2.- La estadística tiene su propio razonamiento, que es necesario enseñar a los estudiantes.

El tratamiento de la información es un campo fundamental en la enseñanza de las matemáticas.

Nos permite:

- 1.- **Ampliar** el currículo matemático
- 2.- **Profundizar** en aspectos históricos
- 3.- **Desarrollar** intuiciones respecto al azar
- 4.- **Ser Críticos**, con fundamento, ante una avalancha de datos. Desmontar falsas intuiciones.
- 5.- **Realizar simulaciones** y proponer conjeturas razonables

- 6.- **Aplicar procedimientos** originales
- 7.- **Encontrar modelos** para resolver problemas
- 8.- **Utilizar las TICs**
- 9.- **Disfrutar** del pensamiento probabilístico

Datos

Tablas

Frecuencia

Variables

Parámetros

Estadística

Gráficas

Sucesos

Combinatoria

Probabilidad

Media

Laplace

Condiciona

Independencia

Binomial

Normal

Desviación típica

Distribuciones

Inferencia

Bayes

Muestreo

Población

Correlación-Regresión

Causalidad

Fiabilidad

¿ Cómo trabajar la Estadística y la probabilidad?

Generalmente , las clases de esta disciplina, se centran en presentar algunos conceptos y la descripción de algún tipo de procedimientos para posteriormente proponer problemas en el que se aplique dicho procedimiento

Esta estructura **no ayuda** a promover el razonamiento estadístico entre los estudiantes ¿ qué hacer?

PROPONER BUENOS PROBLEMAS y DISCUTIRLOS

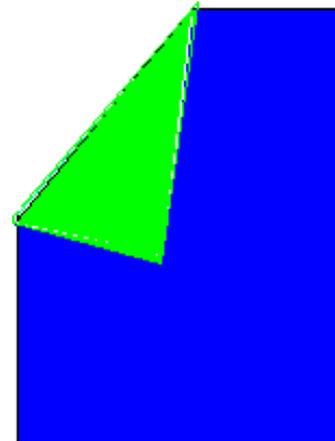
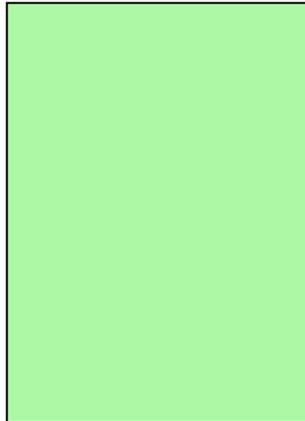
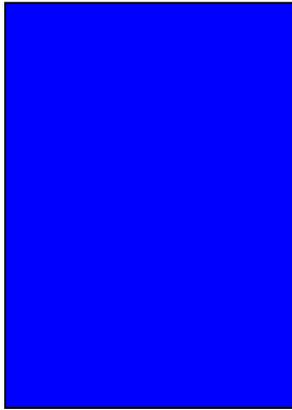
Para abrir Boca

A casi todo el mundo le sorprende que en familias de cuatro hijos lo más **probable** es que haya tres hijos de un sexo y uno del otro.



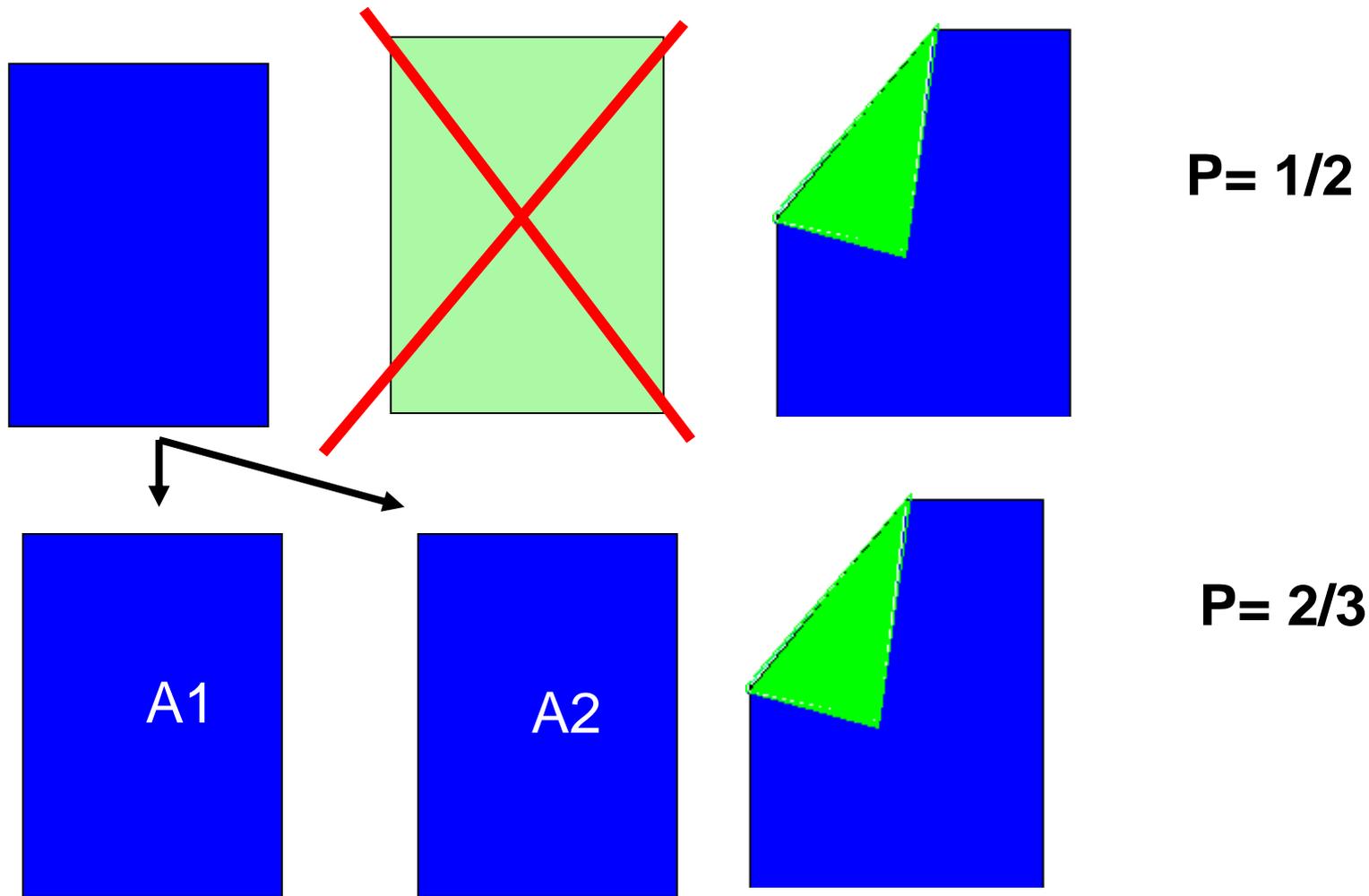
De aperitivo

Tenemos tres cartas: una con los dos caras azules, otra con las dos verdes y la tercera con una cara azul y otra verde:



Se toma una carta y se ve que tiene color azul (una de las caras)

¿ cuál es la probabilidad de que la otra cara también sea azul?

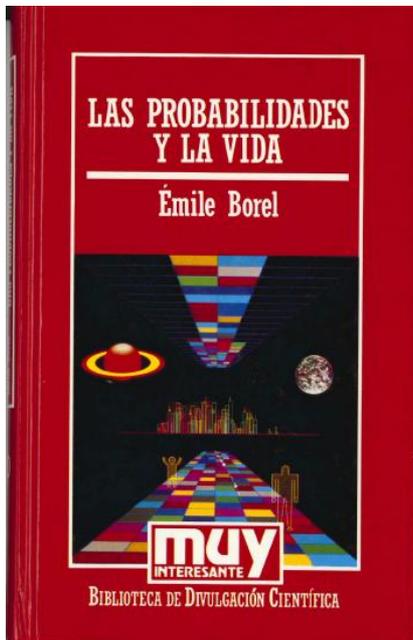


“La teoría de las probabilidades obedece a consideraciones tan delicadas y no es raro que, partiendo de los mismos datos, dos personas lleguen a resultados distintos”
Ensayo filosófico de Laplace (1814)

¿ Existen leyes del azar?

Parece evidente que la respuesta debería ser negativa, ya que precisamente el azar se define como característica de los fenómenos que no tienen ley, fenómenos cuyas causas son demasiado complejas para que podamos preverlas.

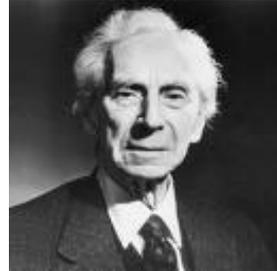
Sin embargo, a partir de Galileo, Pascal, Fermat, ... Se ha establecido una ciencia cuyo objeto es el estudio de las leyes del azar.



Emile Borel (1871-1956)



*¿Cómo osamos hablar de leyes del azar?
¿No es, acaso, el azar la antítesis de cualquier ley?*



Bertrand Russell

Es un hecho destacable que una ciencia que empezó analizando juegos de azar acabe convirtiéndose en el más importante objeto del conocimiento humano.



P.S. Laplace

¿Tiene leyes el azar?

Dos personas lanzan **150 veces** una moneda, los resultados obtenidos son los siguientes

:



<http://www.tirardado.com/1/10.html>

Clara:

C+C+ +CC+ +CC+C+C+ +C+ +C+CCC+ + +CCC+ + C+
+C+C+C+ +CC+CCC+C+C+CC+ + +CC+ +C+C+ +e C+C+
+CC+ C+ +CC+CC+ C+ + +C+ +CC+ +C+ +
C+C+CC+C+ +CC+C+C+ +CCC+CC+ +C+C+ +CC+ + +
C+++C+C++CCC++

Luisa

+CC+ + +C+ + + +C+CC+ + +CCCC+ + +CC+CCC+ + + C+ +
+ + + +C+ C+C+C+ + + +CCCCC+CCC+C+C
C+ CCCCC+CCC+ +CCC+ C+ ccccccccc + + c + ccccccc
+ + + + +CCCC+ +C+C+CC+CC+CC+ + + + + +C+CC+ +
+CCC+ +CCC

¿ Hizo trampas alguna de ellas al tirar la moneda?

Cuando el número de realizaciones de un experimento aleatorio crece mucho la frecuencia relativa del suceso se va acercando cada vez más hacia un cierto valor. Este valor se denomina PROBABILIDAD del suceso.



(1ª Ley de los grandes números)

Es muy POCO PROBABLE que si efectuamos un número suficientemente grande de experimentos, la frecuencia de un acontecimiento se aparte notablemente de su probabilidad.

(1ª Ley de los grandes números)

De las 150 veces que se ha tirado la moneda

Frecuencias de caras y cruces en las secuencias

	C	+
Clara	72	78
Luisa	67	83
Teórica	75	75

**Parece que la
Proporción
es buena**



Frecuencia simple

Clara:

C+C+ +CC+ +CC+C+C+ +C+ +C+CCC+ + +CCC+ + C+
+C+C+C+ +CC+CCC+C+C+CC+ + +CC+ +C+C+ +C+C+C+
+CC+ C+ +CC+CC+ C+ + +C+ +CC+ +C+ +
C+C+CC+C+ +CC+C+C+ +CCC+CC+ +C+C+ +CC+ + +
C+++C+C++CCC++

Luisa

+CC+ + +C+ + + +C+CC+ + +CCCC+ + +CC+CCC+ + +C+ +
+ + + +C+ C+C+C+ + + +CCCCC+CCC+C+C
C+ CCCCC+CCC+ +CCC+ C+ cccccccc+ + c+ ccccccc
+ + + + +CCCC+ +C+C+CC+CC+CC+ + + + + +C+CC+ +
+CCC+ +CCC

Ahora analizamos lanzamientos dobles (total 75)

	CC	C+	+C	++
Clara	12	30	18	15
Luisa	25	21	12	17
Teórica	18	18	18	18

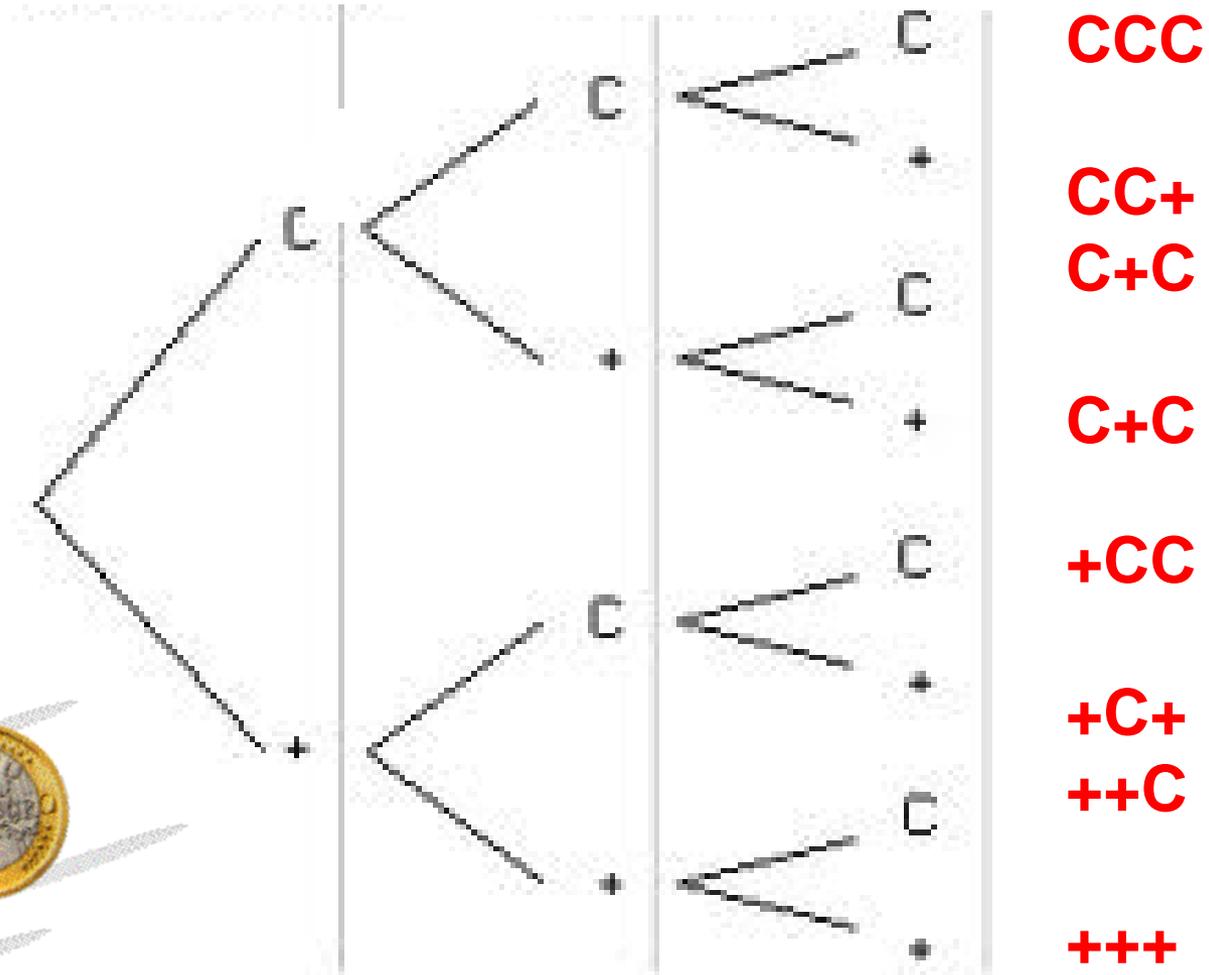
Parece que hay una ligera desviación

Primer

lanzamiento

Segundo

Tercero



Analizamos lanzamientos triples (50 lanzamientos)

Frecuencias de los diferentes resultados al lanzar las monedas 3 a 3

	CCC	CC+	C+C	C++	+CC	+C+	++C	+++
Clara	0	2	14	9	6	7	10	1
Luisa	8	11	6	4	6	4	5	6
Teórica	6	6	6	6	6	6	6	6

!!Parece que la que hace trampas es Clara!!

Aunque tenemos la sospecha de que Clara hace trampas, serían necesarias nuevas pruebas para dejar a Luisa libre de toda sospecha. ¿ cuáles?

Contrastes Chi-cuadrado, etc.

En todo caso hablaríamos en términos de probabilidad.

Analizar los datos

Ejemplo:

Un taxi se ve envuelto en accidente nocturno, en una ciudad en la que hay un **15%** taxis **azules** y un **85 %** taxis **verdes**

- Una persona dice que es azul
 - Le hacen una prueba de visión y el **80%** de las veces contesta bien
- ¿ cuál es la probabilidad de que el taxi sea azul?

15 azules	12	3
85 verdes	17	68
	29	71
	Azul	Verde



$$P(\text{azul}) \text{ dado que el testigo decía azul} = 12 / 29 = 0,41$$

Falacia de la proporción básica (Segunda variante)

Ejemplo:

Un taxi se ve envuelto en accidente nocturno, en una ciudad en la que hay un **15%** taxis azules y un **85 %** taxis verdes

- Una persona dice que es azul

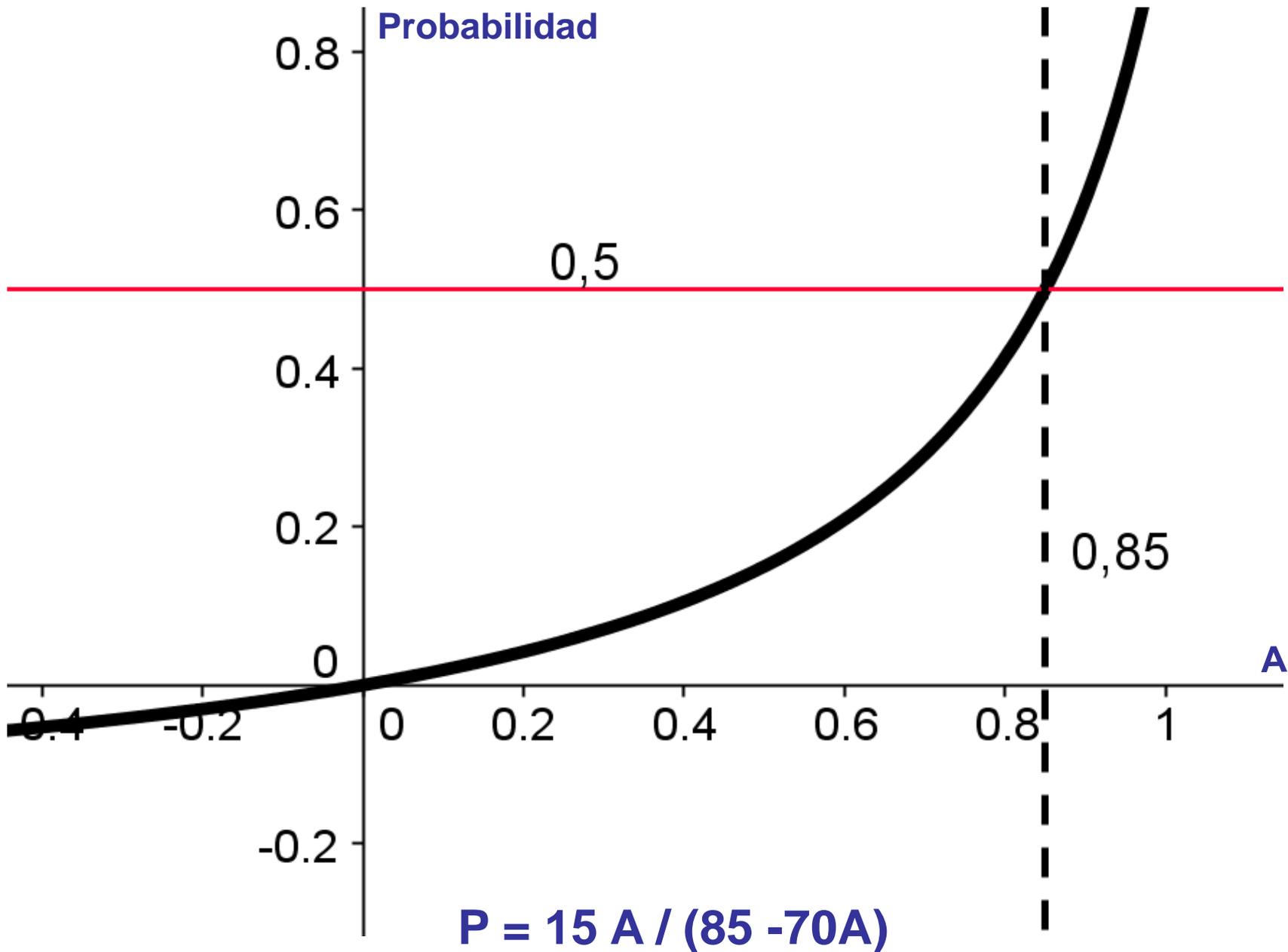
- Le hacen una prueba de visión y el **A %** de las veces contesta bien

¿ cuál es la probabilidad de que el taxi sea azul?

15 azules	15A	15(1-A)
85 verdes	85(1-A)	85A
	85-70A	15+70A
	Azul	Verde



$$P(\text{azul}) \text{ dado que el testigo decía azul} = \frac{15A}{85 - 70A}$$



Falacia de la proporción básica(tercera variante)

Ejemplo:

Un taxi se ve envuelto en accidente nocturno, en una ciudad en la que hay un $M\%$ taxis azules y un $(100-M)\%$ taxis verdes

- Una persona dice que es azul
- Le hacen una prueba de visión y el $A\%$ de las veces contesta bien

¿ cuál es la probabilidad de que el taxi sea azul?

M azules		
100-M verdes		
	Azul	Verde



$P(\text{azul})$ dado que el testigo decía azul = ???

¿ Cuales son nuestras percepciones respecto al azar?

100 personas con fuerte dolor abdominal

En estos casos se sabe que alrededor del 30% tiene apendicitis

Se realizan pruebas mediante ultrasonidos a TODOS

Las pruebas dan positivo en un 80% de los enfermos

Mientras que el 10% de los sanos dan positivo (falso positivo)

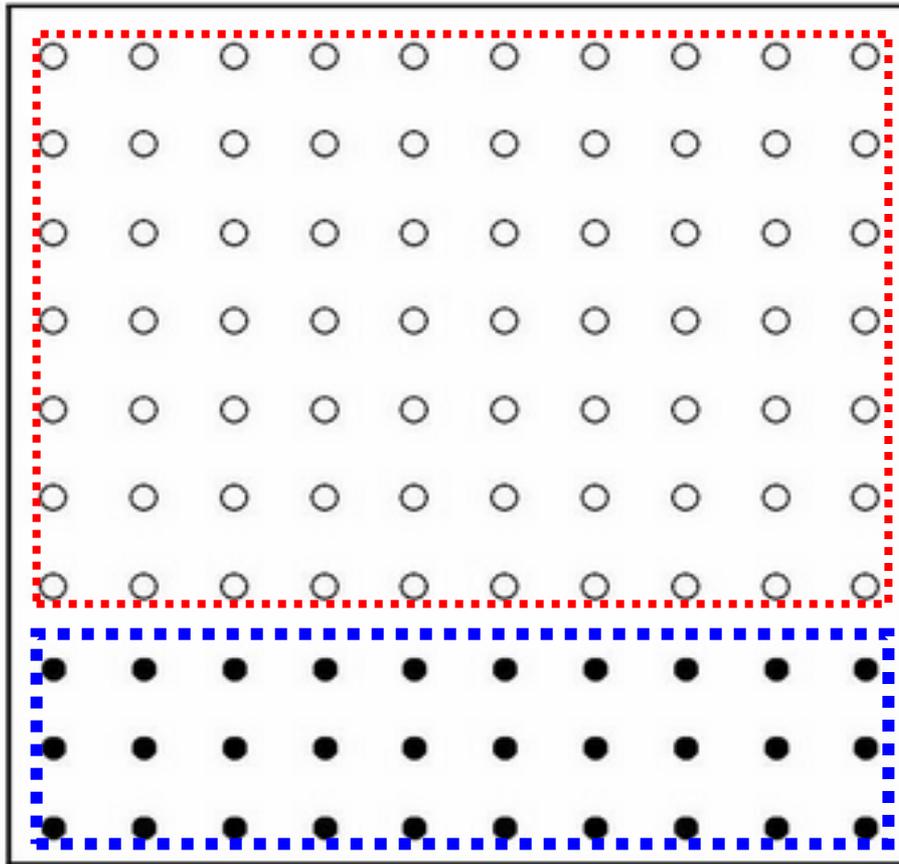
¿Cuál es la probabilidad de que una persona esté enferma si la prueba ha resultado positiva?

¿Cuál es la probabilidad de que una persona esté sana si la prueba ha resultado negativa?



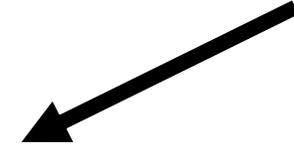
El 30% enfermos

Se sabe que alrededor del 30% tiene apendicitis



PACIENTES SANOS

PACIENTES ENFERMOS



10% falso positivo

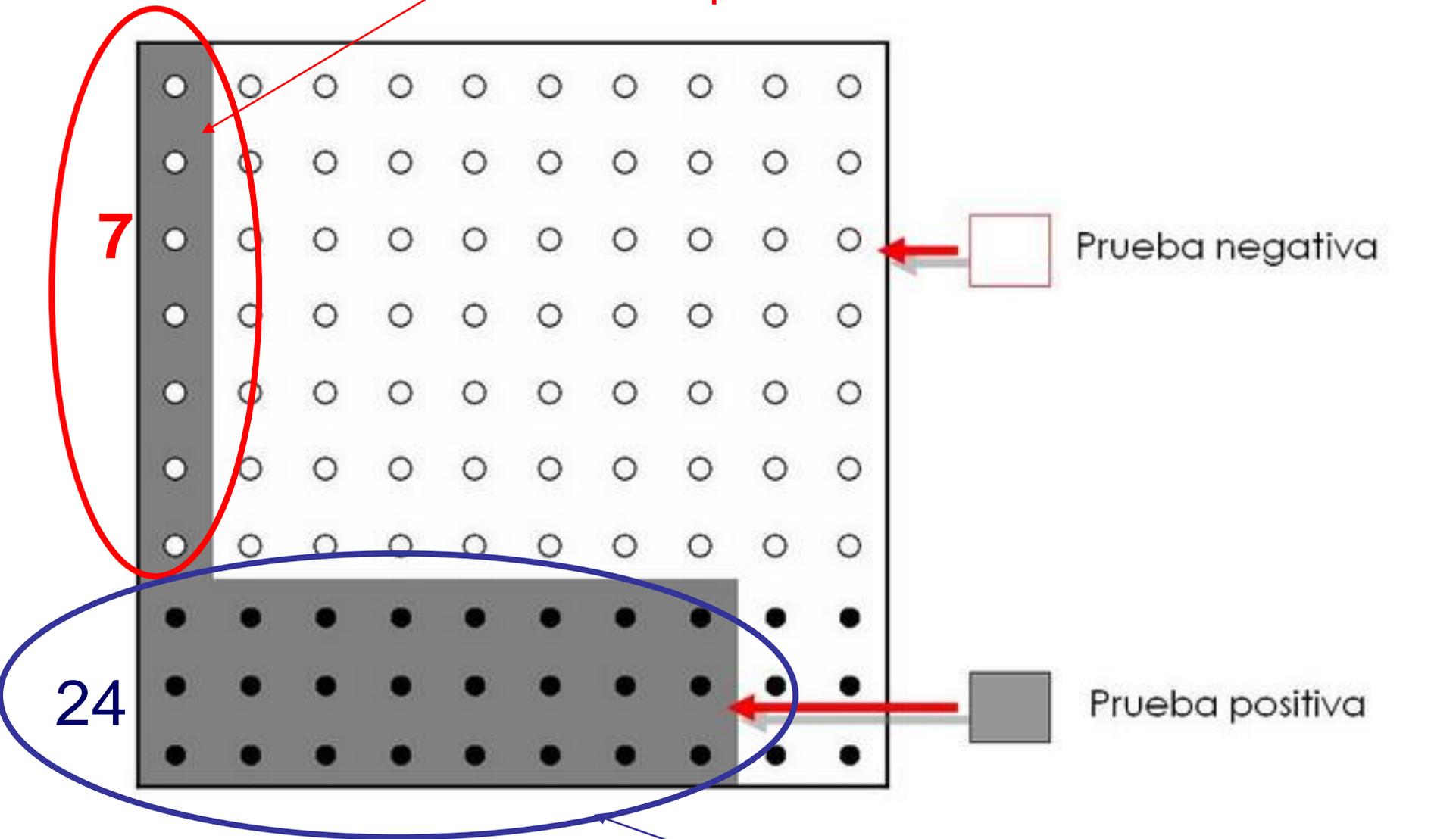
7

24

Prueba negativa

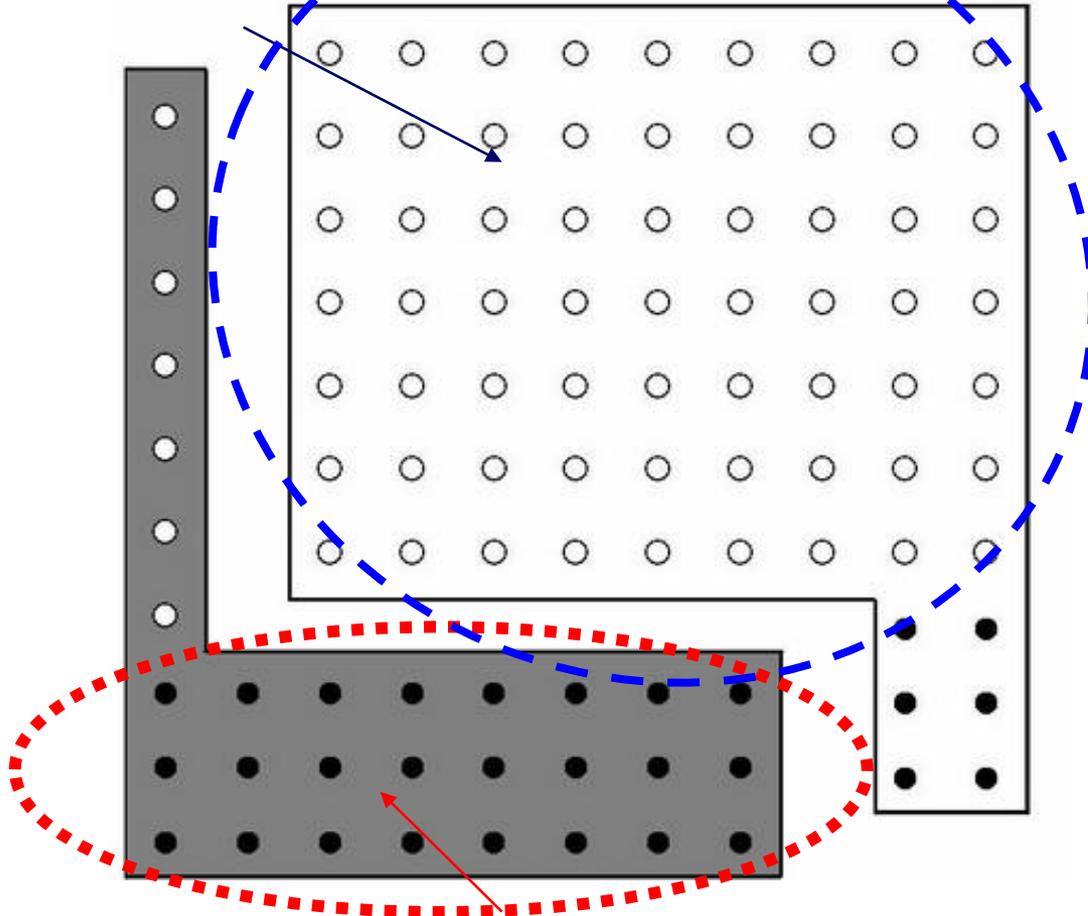
Prueba positiva

80% positivos de los enfermos



Un nuevo universo

Sanas



*¿Cuál es la probabilidad de que una persona esté **enferma** si la prueba ha resultado positiva?*

$$24/31 = 77\%$$

*¿Cuál es la probabilidad de que una persona esté **sana** si la prueba ha resultado negativa?*

$$63/69 = 91\%$$

Enfermas

Hubo una vez un Rayo que cayó dos veces en el mismo sitio; pero encontró que ya la primera había hecho suficiente daño, que ya no era necesario, y se deprimió mucho.

Augusto Monterroso



Un hombre que viajaba mucho estaba preocupado por la posibilidad de que hubiera una bomba en su avión. Calculó la probabilidad exacta de que fuera así y, aunque ésta era muy baja, no lo era lo suficiente como para dejarlo tranquilo.

Desde entonces lleva siempre una bomba en la maleta. Según él, la probabilidad de que haya dos bombas a bordo es infinitesimal.

J. A. Paulos

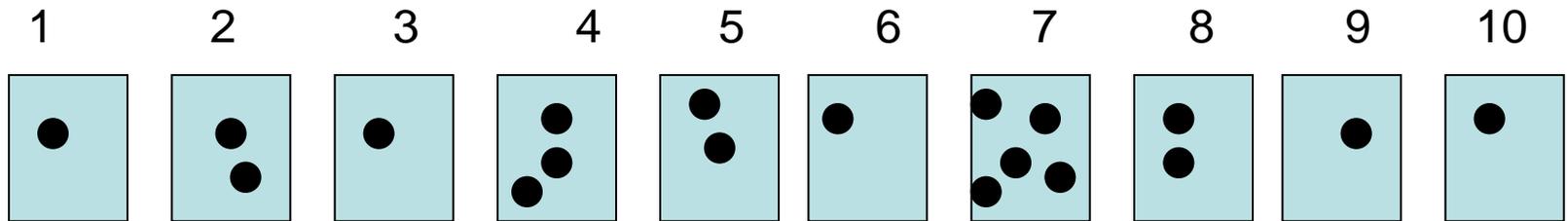


Colección de Cromos ¿ cuántos?



Colección de 10 cromos en total

Nº cromo: 2, 6, 4, 5, 7, 4, 7, 7, 7, 4, 9, 5, 2, 7, 0, 8, 1, 8, 3



Hemos tenido que comprar 19 cromos para completar la colección de 10 cromos
¿ son pocos? ¿ es lo esperable?

ESQUEMA DIDÁCTICO PARA AFRONTAR EL MUNDO DEL AZAR

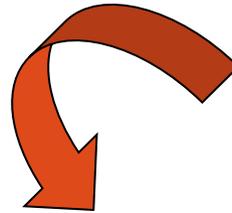
1.- EXPERIMENTACIÓN



2.- DESCUBRIR REGULARIDADES



3.-CONJETURAR



4. REPRESENTAR Y COMUNICAR



5.- FORMALIZAR Y COMPROBAR

Un acercamiento formal al problema.

Supongamos que la colección tiene 50 cromos.

Es claro que el primero que compremos no lo tenemos, con el segundo la probabilidad de no tenerlo es de $49/50$,

Si ya tenemos 40 cromos diferentes la probabilidad de que al comprar uno nuevo no lo tengamos es de $10/50$, es por tanto esperable que tengamos que comprar 5 cromos (el inverso de $50/10$), al igual pasa en los otros casos:

Por tanto el número total de cromos a comprar es:

$$N = \frac{50}{50} + \frac{50}{49} + \dots + \frac{50}{10} + \dots + \frac{50}{3} + \frac{50}{2} + 50$$

$$N = 50 \cdot \left(\frac{1}{50} + \frac{1}{49} + \dots + \frac{1}{10} + \dots + \frac{1}{3} + \frac{1}{2} + 1 \right)$$

$$\mathbf{N = 50 (4.5) = 225 \text{ cromos}_}$$

Las Pastillas

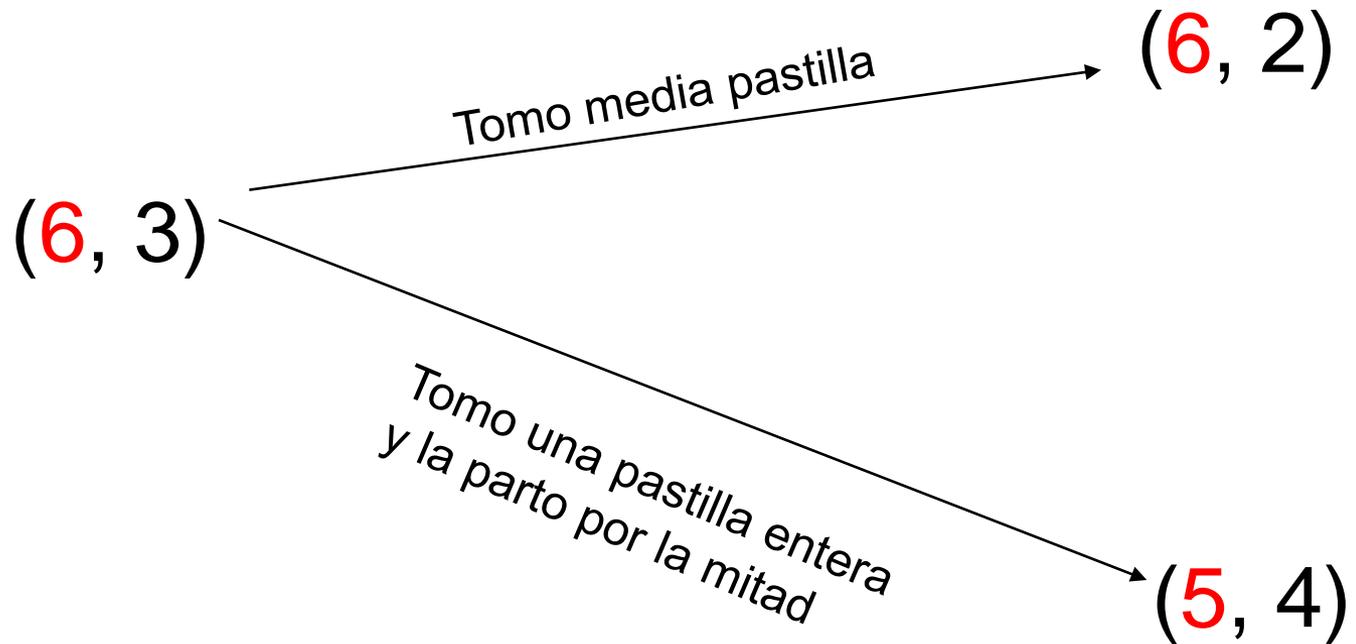
Una caja contiene inicialmente n pastillas, que se han de tomar en dosis mitad. Cada comprimido se retira de la caja al azar; si resulta ser una pastilla entera, se parte en dos y se deja una mitad en la caja; si se trata de media pastilla, se retira sin más.

Cuando ya no queda ninguna pastilla entera, nos preguntamos por las probabilidades de que el número de medias pastillas que restan sea 1, 2, 3, ..



Notación adecuada

Para seguir el proceso supongamos que en un momento hay 6 pastillas enteras y 3 medias pastillas. Este estado concreto lo denominaremos (6, 3)

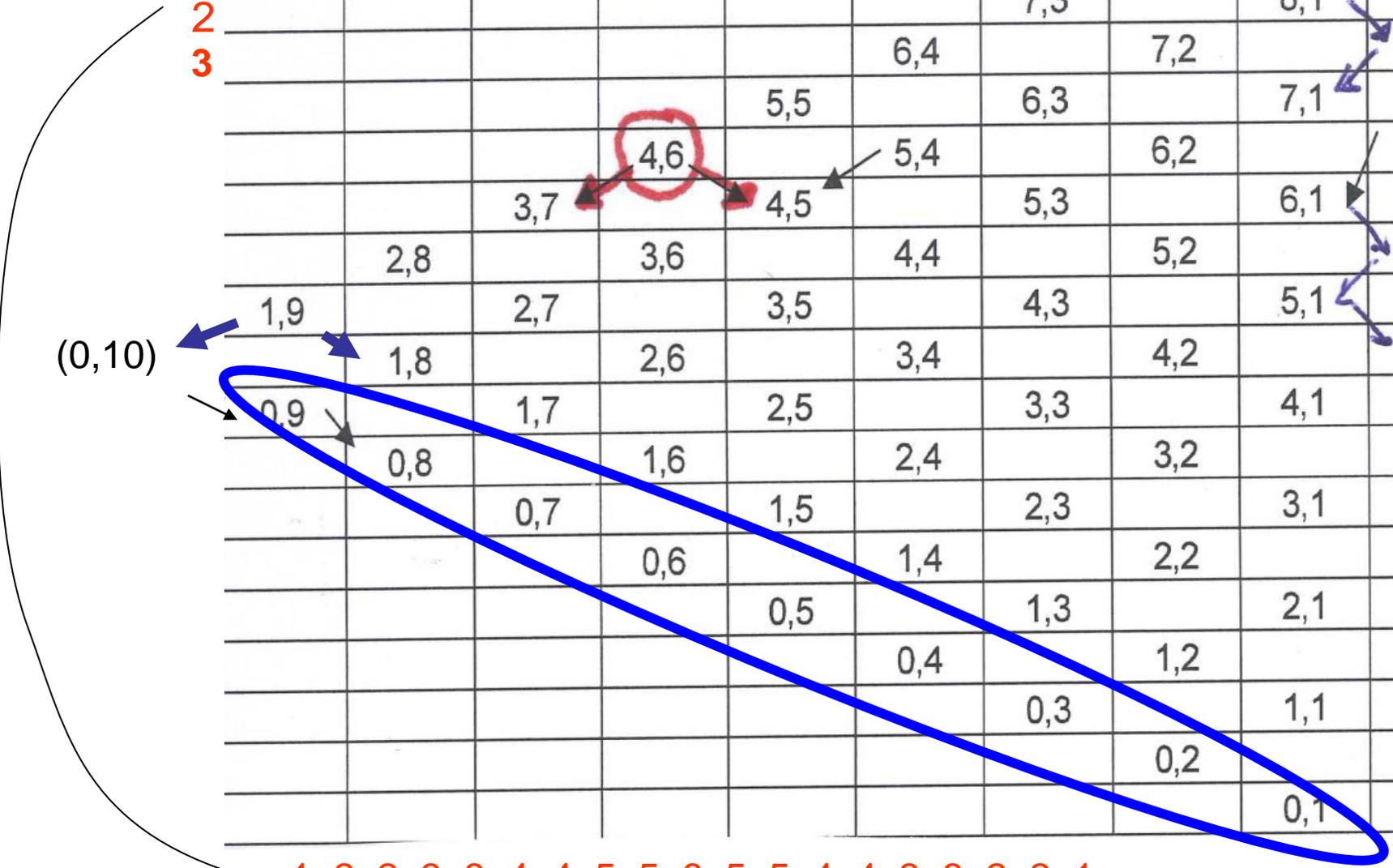


										10,0				
1									9,1					
2								8,2		9,0				
2							7,3		8,1					
3						6,4		7,2		8,0				
				5,5		6,3		7,1						
								6,2		7,0				
			3,7	4,6	4,5	5,4		6,1						
		2,8		3,6		4,4		5,2		6,0				
			1,9		2,7		3,5		4,3					
				1,8		2,6		3,4		4,2				
					1,7		2,5		3,3					
						1,6		2,4		3,2				
							1,5		2,3					
								1,4		3,1				
									1,3					
										2,1				
											2,0			
												1,1		
													1,0	
														0,1

1
2
2
3

(0,10)

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



Si calculamos el número máximo de estados posibles, si se exceptúan el estado inicial $(n, 0)$ y el de la caja vacía $(0, 0)$, es igual a

$$E = (n^2 + 3n - 2) / 2$$

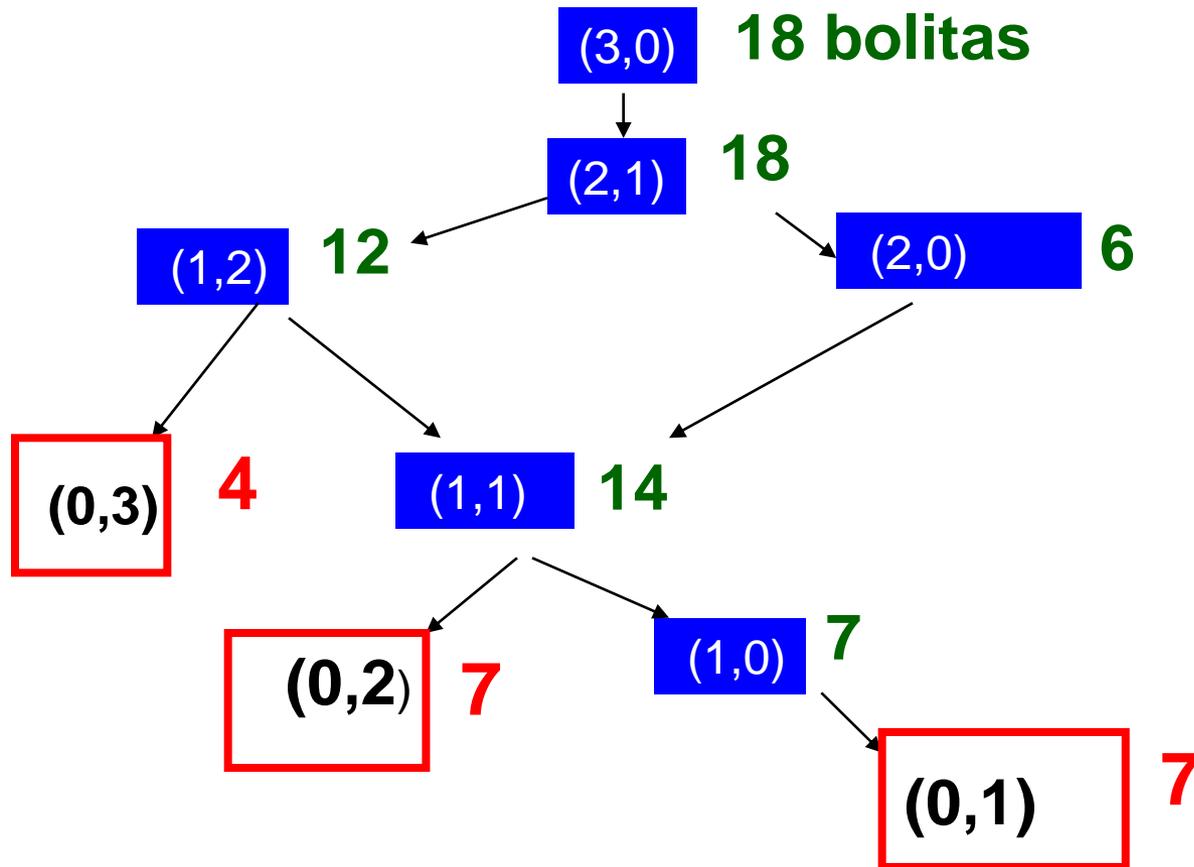
n	Estados
1	1
2	5
7	34
10	64
100	5148

Tomemos 3 pastillas enteras

Estado 0				(3,0)	
Estado 1			(2,1)		
Estado 2		(1,2)		(2,0)	
Estado 3	(0,3)		(1,1)		
Estado 4		(0,2)		(1, 0)	
Estado 5			(0,1)		

Realmente estamos interesados en las probabilidades de los estados del tipo $(0,m)$, por tanto para su cálculo habremos de considerarlos estados finales (hojas terminales del árbol).

Para ello basta con eliminar las transiciones $(0,m) \rightarrow (0, m-1)$. De este modo obtendremos las probabilidades que nos interesan



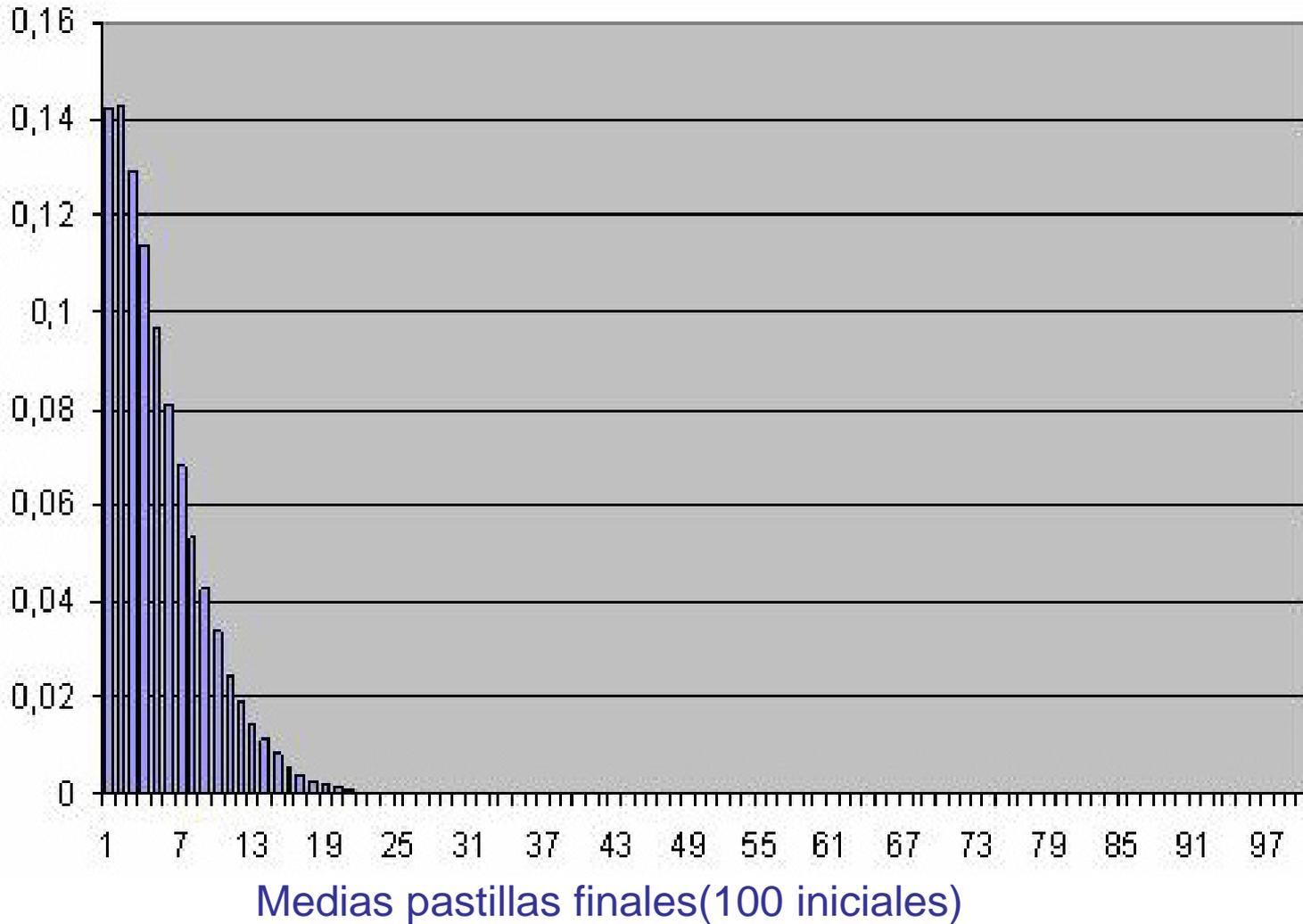
Es evidente que la suma de todas las probabilidades de las medias pastillas ha de ser 1,

$$p(0,n) + p(0,n-1) + \dots + p(0,2) + p(0,1) = 1$$

Simulación mediante un ordenador

Pastillas

Medias finales	Probabilidad calculada	Frecuencia ensayo
1	0,238557094	0,238189
2	0,238557094	0,238793
3	0,198144990	0,198143
4	0,144262185	0,143843
5	0,092546956	0,093087
6	0,051628725	0,051639
7	0,024315122	0,024281
8	0,009153107	0,009254
9	0,002471847	0,002405
10	0,000362880	0,000366



De acuerdo a la simulación se aprecia la concentración de las frecuencias significativas en los valores pequeños del número final de medias pastillas. El descenso de las frecuencias tiene la apariencia de ser de tipo exponencial,

El problema de las coincidencias, de los sombreros, de Euler,

Cuatro señores, cada uno con su sombrero, van a la ópera y al entrar dejan los sombreros en el guardarropa. A la salida cada uno toma al azar un sombrero. ¿Cuál es la probabilidad de que ninguno de los señores reciba su sombrero?

Este problema fue estudiado, en un caso particular, por el científico francés **Pierre Remond de Montmort** en 1708



			Posición original
			
			←
			
			
			
			←

Con 3 sombreros en **2 casos** no hay ninguna coincidencia.

Efectivamente, existe tal relación, encontrada por L. Euler

$$P(n) = (n-1) \cdot [P(n-1) + P(n-2)]$$

y también obtuvo que :

$$P(n) = n \cdot P(n-1) + (-1)^n$$

Nº de permutaciones en las que ninguna ocupa su posición original, $P(n)$

$$P(1) = 0$$

$$P(2) = 1$$

$$P(3) = 2$$

$$P(4) = 9$$

$$P(5) = 44$$

¿ Podemos encontrar alguna relación de recurrencia entre los distintos valores de $P(n)$?

$$\lim(P_n) = \lim\left(\frac{P(n)}{n!}\right) \dots > 1/e \approx 0,3678$$

$$P(n) = n! \left[1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \dots + \frac{(-1)^n}{n!} \right]$$

Por tanto la probabilidad de **NO coincidencia** es :

$$P_n = \frac{P(n)}{n!}$$

$$P_3 = 0,333$$

$$P_4 = 0,375$$

$$P_5 = 0,366$$

$$P_6 = 0,368$$

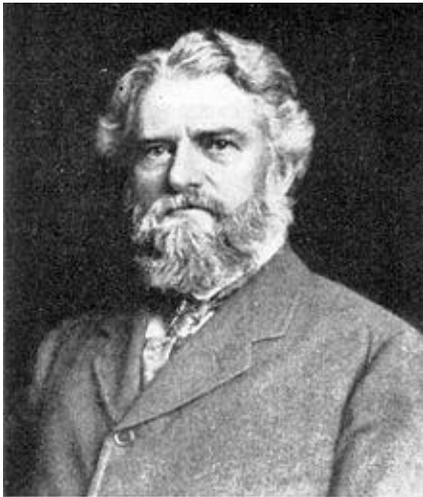
$$P_{23} = 0,3678$$

$$P_{24} = 0,3678$$

Parece que se estabiliza en torno a 0,3678

Ley de Benford

Estamos en los primeros meses del año **1881**, el astrónomo norteamericano **S. Newcomb** estaba usando un libro de logaritmos, cuando observó que las páginas del libro estaban más viejas y usadas cuanto más cercanas estaban del principio



S. Newcomb
(1835-1909)

Actualmente equivaldría a examinar el desgaste de la **tecla "1"** en calculadoras u ordenador ¿A qué se debía? ¿Había alguna razón?

Sólo podía tener una explicación: a lo largo de los años había consultado mucho más el logaritmo de los números que comenzaban por 1 que de los que comenzaban por números más altos.

Este hecho fue olvidado durante años, hasta que en **1938**, **Frank Benford**, un físico de la compañía General Electric, se dio cuenta del mismo patrón.



Frank Benford
(1883-1948)

Entusiasmado por el descubrimiento, estudió 20.229 números provenientes de 20 muestras de todo tipo: constantes y magnitudes físicas, longitudes de ríos, estadísticas de béisbol, direcciones de personas... incluso cifras sacadas de portadas de revistas.

A partir de los datos extraídos del mundo real, comprobó que la probabilidad de que en una serie de muchos datos el primer dígito de un número sea 1 es del 30%, 17,6% para un 2, 12'5% para el 3 y así va decreciendo...

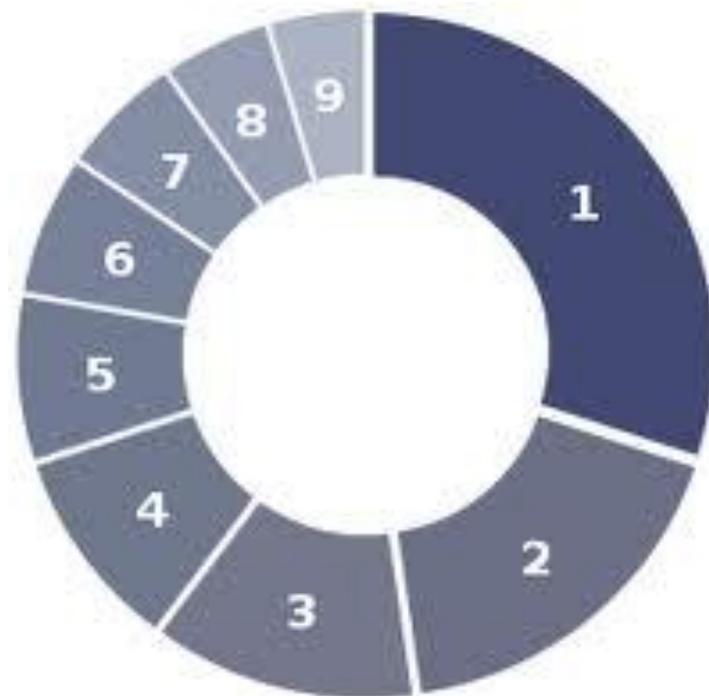
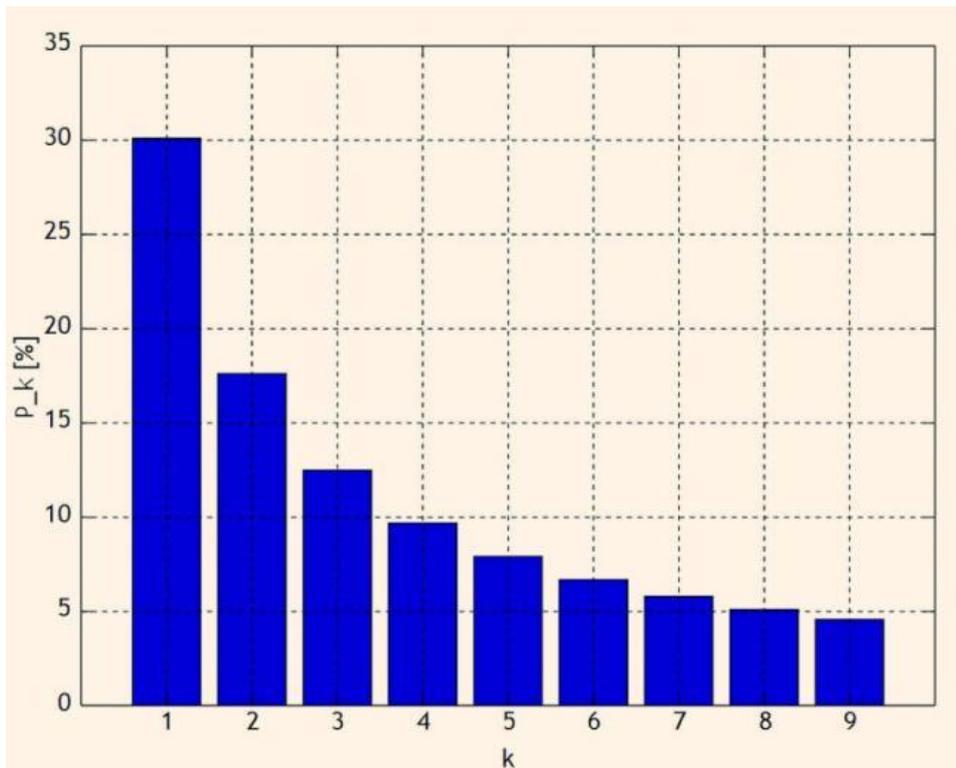
Primera cifra Probabilidad

1	30.1 %
2	17.6 %
3	12.5 %
4	9.7 %
5	7.9 %
6	6.7 %
7	5.8 %
8	5.1 %
9	4.6 %

La ley de Benford,
también conocida
como la **ley del
primer dígito**



Esta ley, asegura que, en los números que existen en la vida real, en la primera cifra, el 1 aparece con mucha más frecuencia que el resto de los números.



Ley de Benford, o del primer dígito

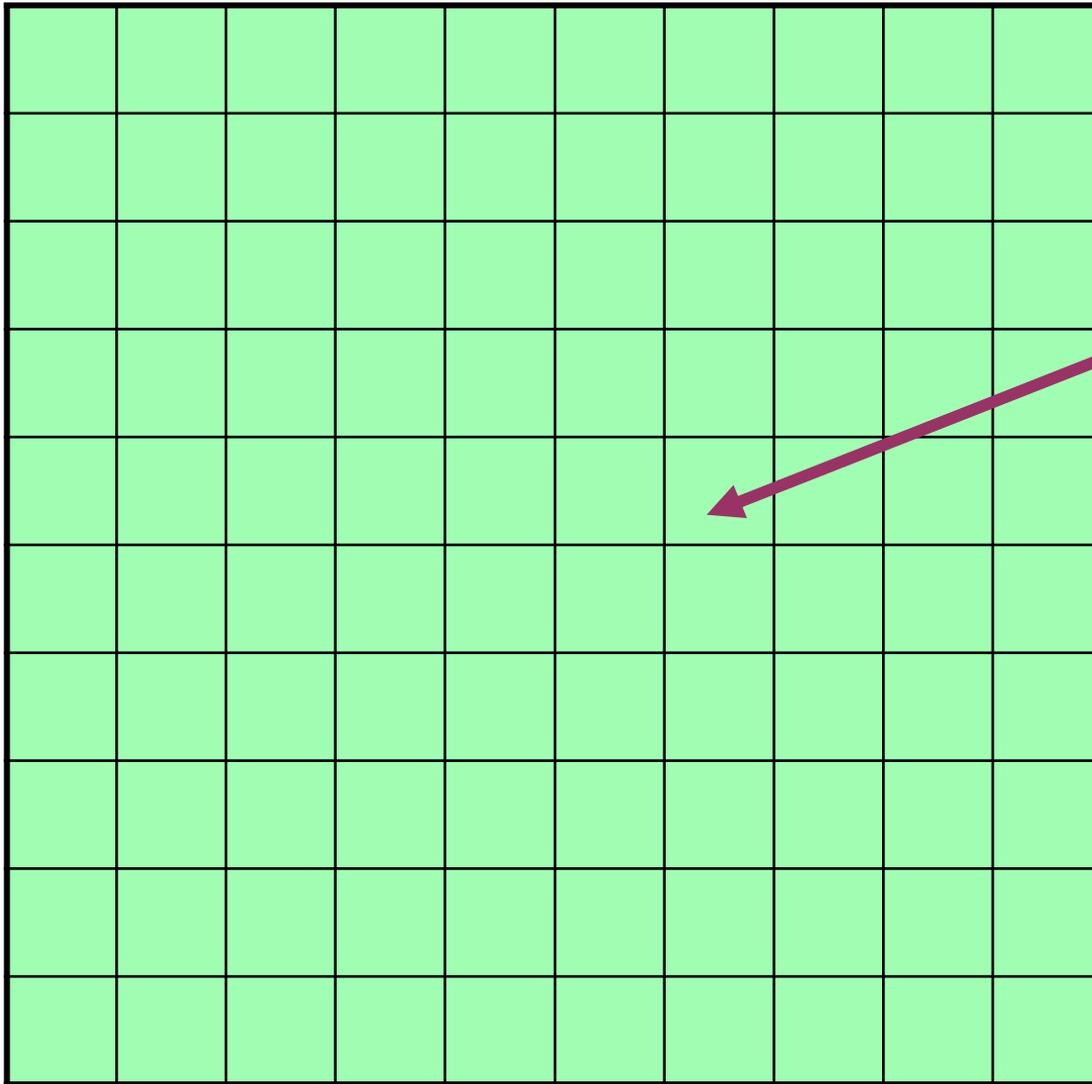
¿ Cuando se cumple la Ley de Benford?

Necesita datos que **no sean totalmente aleatorios** ni muy condicionados.

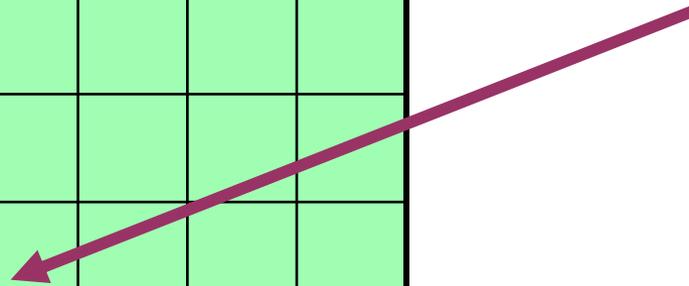
Los datos pueden ser de una gran variedad de fuentes y suelen ser el resultado típico de diversos procesos, con muchas influencias, como ocurre con la mayoría de datos extraídos de fenómenos naturales, sociales y económicos.

- Ejemplos:
- a) Número de habitantes de los Pueblos de Bizkaia
 - c) Primos
 - d) Números aleatorios (No)

¿ CASUALIDADES?



57



Lanzamos dardos sobre una cuadrícula de 10x10 .
¿Cuántos dardos habrá que lanzar por término medio hasta que caigan dos dardos en la misma cuadrícula?

La aparente paradoja del cumpleaños

Usted entra en su clase y pregunta por el día del cumpleaños de todos y cada uno de sus alumnos y alumnas. Si en su clase hay 30 alumnos seguro que espera pocas coincidencias. Razona de la manera siguiente. El año tiene 365 días y solamente hay 30 alumnos, por tanto la posibilidad de coincidencia es muy pequeña ¿es esto cierto?

Haga la experiencia en su clase y seguro que se quedará sorprendido.

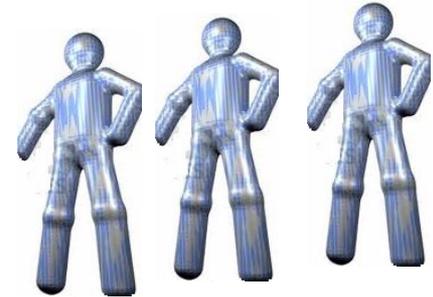
La probabilidad de que los dos alumnos no coincidan es:

$$\frac{365}{365} \cdot \frac{364}{365} = 0,9973$$



Si razonamos de manera análoga, la probabilidad de que tres hayan nacido en días distintos será:

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} = 0,9918$$

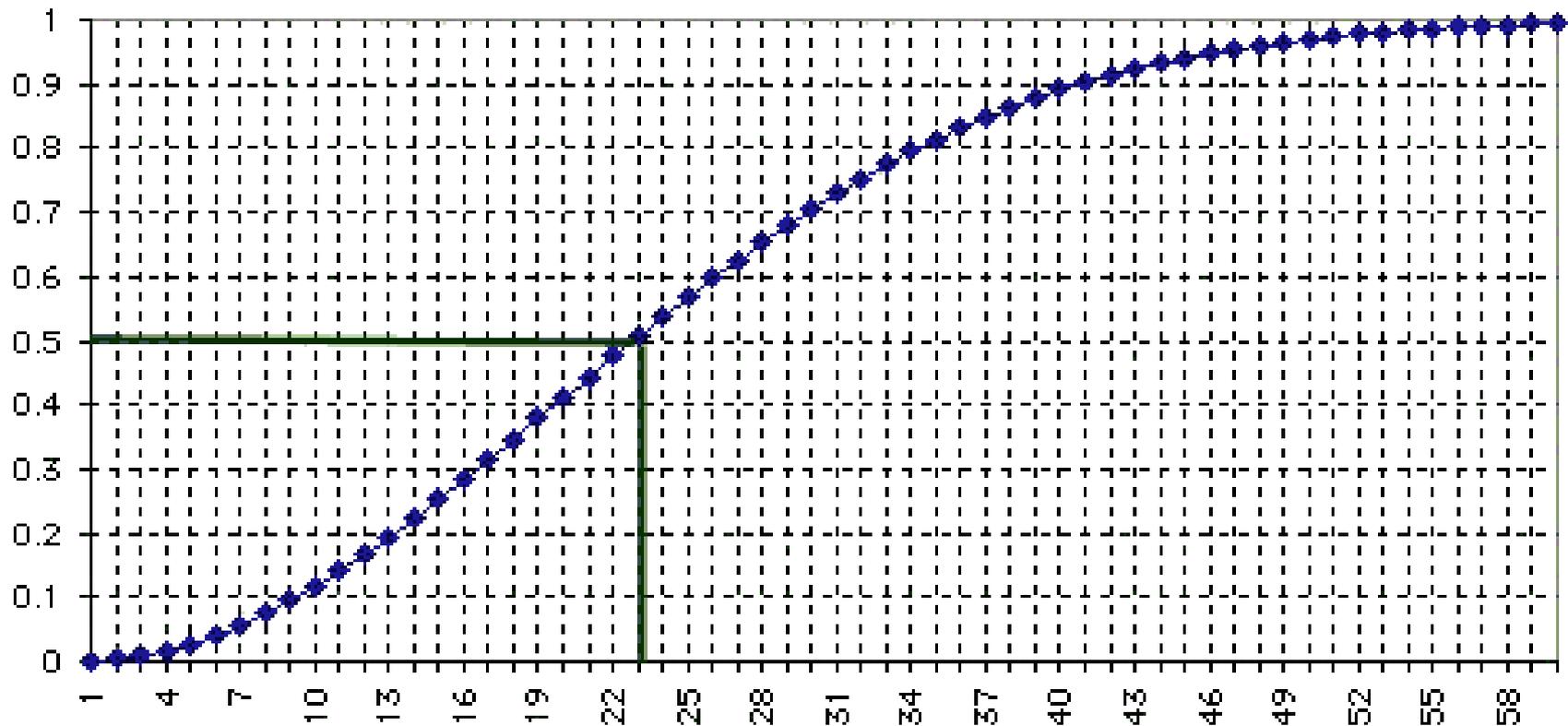


Y la probabilidad de que los 30 hayan nacido en días distintos (no existan coincidencias entre ellos) será:

$$\frac{365}{365} \cdot \frac{364}{365} \cdot \frac{363}{365} \cdots \frac{336}{365} = 0,2937$$

La probabilidad **de que sí exista coincidencia**, esto es que coincidan al menos dos personas en su cumpleaños, será igual a:

$$1 - 0,2937 = 0,7063$$



P(coincidencia entre n personas) =
$$\frac{365 \times 364 \times 363 \times \dots (365 - n + 1)}{365^n}$$



El 22 de agosto de 2002 la combinación ganadora fue la siguiente **13, 21, 24, 26, 32 y 34**. Mientras que el 10 de Diciembre de 2009 fue la misma ¡qué casualidad!



Número total de posibles combinaciones: 13.983.816

Sorteos de la primitiva celebrado en estos años 2.245 sorteos.

$$1 - \frac{13.983.816}{13.983.816} \cdot \frac{13.983.815}{13.983.816} \cdots \frac{13.981.573}{13.983.816} \cdot \frac{13.981.572}{13.983.816} = 0,16$$

Anotamos los dos números finales de las matrículas de coches y apostamos a que se repiten las dos últimas cifras de la matrícula en quince automóviles anotados al azar ¿es razonable mi apuesta?



Es equivalente a un año de 100 días y en este caso las personas son 15. Por tanto la probabilidad de coincidencia será igual a:

$$1 - \frac{100}{100} \cdot \frac{99}{100} \cdots \frac{86}{100} = 0,67$$

Precisión en el lenguaje, fallos_



SOLUCIÓN !!

Lanzamos un dado

¿ En qué intento es más fácil de que salga el número 5 ?

En el primer intento

En el segundo

En el tercero

En el cuarto,.....

En el N-esimo

No ha salido en el primero = $(5/6)(1/6)$

No ha salido ni en el 1º ni en 2º = $(5/6)(5/6)(1/6)$

No ha salido ni el 1º, ni 2º,.. Ni $(N-1)º = (5/6)^N(1/6)$

Los sabios también se equivocan...

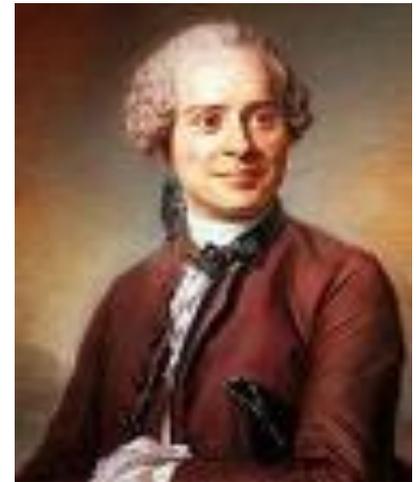
Problema: Al tirar una moneda dos veces,
¿cuál es la probabilidad de obtener, por lo menos, una cara ?

C+, CC, +C, ++

Solución **-errónea-** de D Alembert

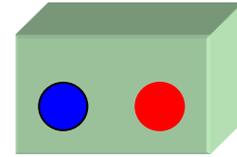
Dudaba de que la probabilidad fuese $3/4$,
razonando que si una cara aparecía en la primera
tirada, el juego habría terminado, pues no era
necesario continuar con una segunda.
Enumerando sólo tres casos posibles,

C , +C, ++ llegó a la probabilidad $2/3$.



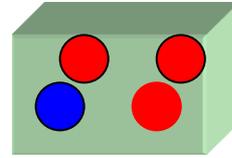
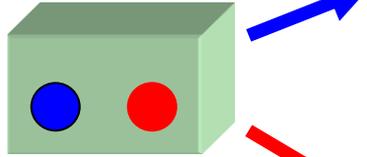
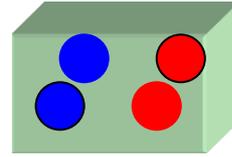
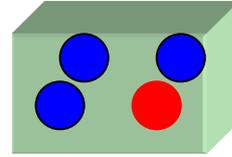
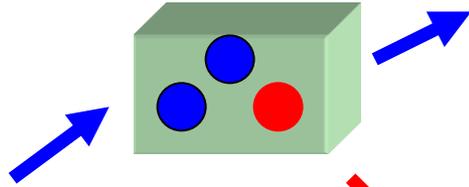
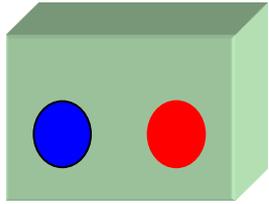
D Alembert (1717-83)

Urna de Polya



Disponemos de una urna, con 2 bolas (1 bola azul y 1 bola roja), acordamos el siguiente procedimiento: extraemos al azar una bola de la urna y dependiendo del color de dicha bola añadimos otra bola del mismo color que la extraída.,... y actuamos así sucesivamente. De manera que poco a poco se van añadiendo bolas a la urna.

El asunto consiste en estudiar la proporción de bolas azules y rojas a medida que el experimento avanza.





*El incierto mañana nunca nos pertenece
Goza del hoy. Y bebe a la luz de la luna,
de esa luna que en vano, milenio tras milenio,
nos buscará fielmente para darnos su brillo*

Omar Jayyam

Gracias