EL ORDEN DEL DESORDEN

J. M. Sanz-Serna
Universidad de Valladolid

I: EL EXTRAÑO CASO DE LAS COMPAÑÍAS RIVALES

EL PROCESO:

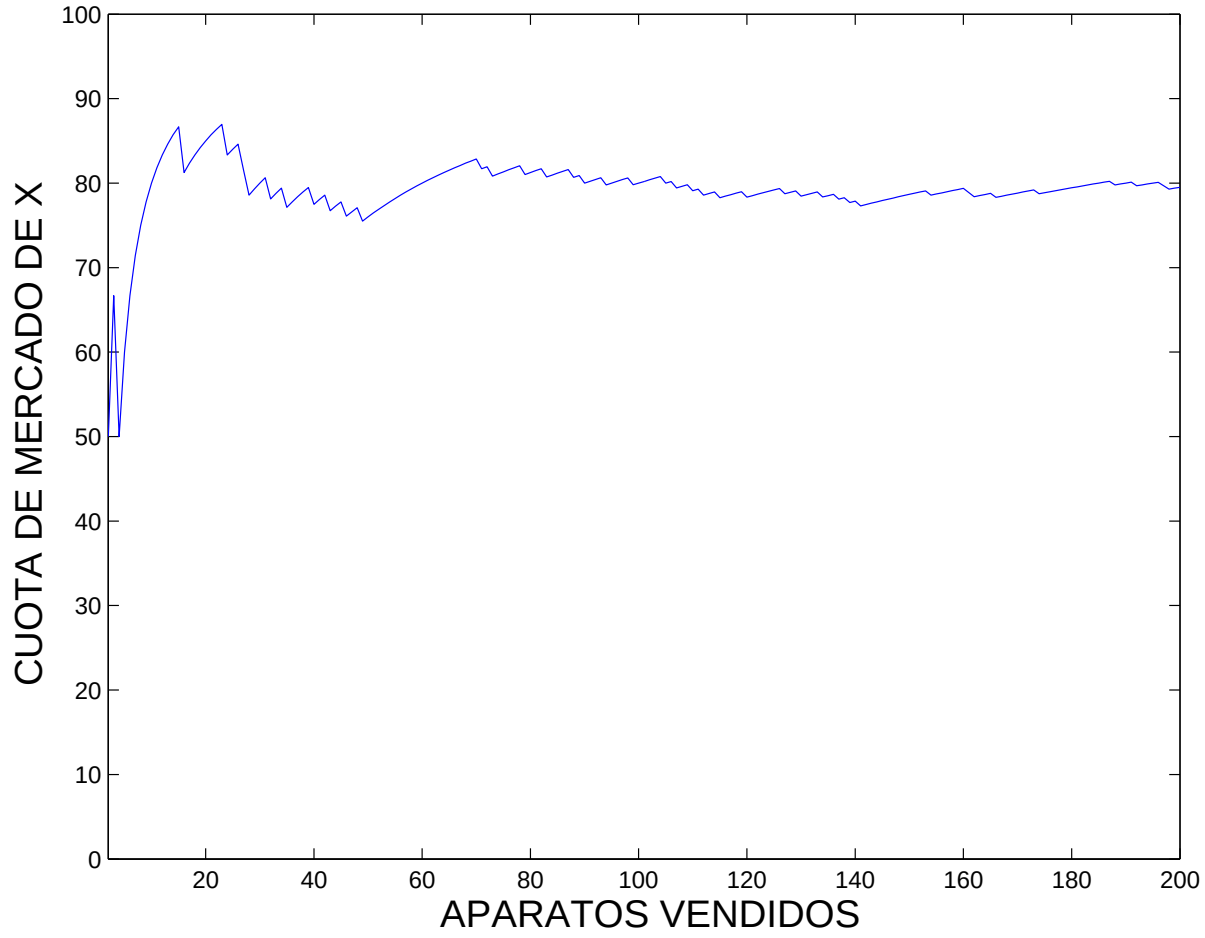
• Dos compañías rivales X, Y lanzan simultáneamente sus E-LOT de cuarta generación. No efectúan publicidad, confiando en que la satisfacción de los usuarios sea el mejor anuncio.

• Alicia compra un E-LOT X y Bernardo un Y. A partir de ese momento cada persona interesada habla con una sola, que encuentra al azar, de las que ya tienen E-LOT, y, de acuerdo con la información que recibe, adquiere su E-LOT X ó Y.

• Los siguientes compradores son Carlos X, Dario Y, Elena X, Fernanda X, Germán X, . . .

• El gráfico siguiente muestra la evolución de la cuota de mercado de X:

EVOLUCIÓN DE LA CUOTA DE MERCADO DE X:



OBSERVAMOS:

• Tras un periodo incial de fluctuaciones hasta las 25 ó 50 primeras ventas, la cuota de X se sitúa en el el entorno del 80%. Al incrementarse el número de aparatos vendidios las fluctuaciones van desapareciendo y la cuota se aproxima a un límite bien definido.

ANALIZAMOS LAS CAUSAS:

- El E-LOT X es notablemente superior al Y?
- El E-LOT X satisface a personas con mayor capacidad de persuasión?
- . . . . . ?

LA REALIDAD:

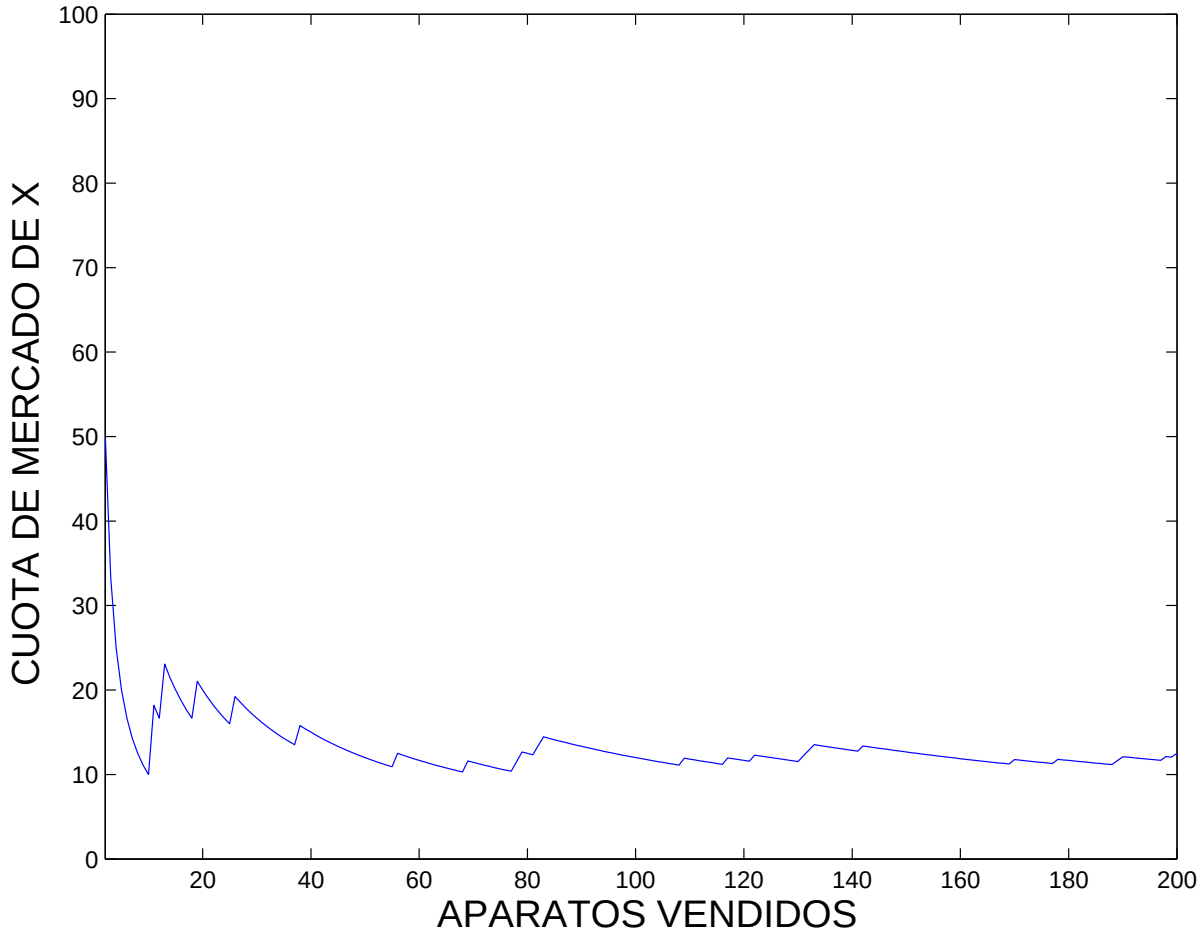
• Los productos de X e Y son indistinguibles y juegan papeles completamente simétricos. Todos los clientes han quedado igualmente satisfechos y han recomendado a quien les han preguntado la marca que ellos tenían.

• Por tanto han comprado X aquellas personas que han preguntado a alguien que tenía X (y analogamente para Y). [P. ej. Carlos compró X porque pregunto a Alicia, de haber preguntado a Bernardo hubiera adquirido Y.]

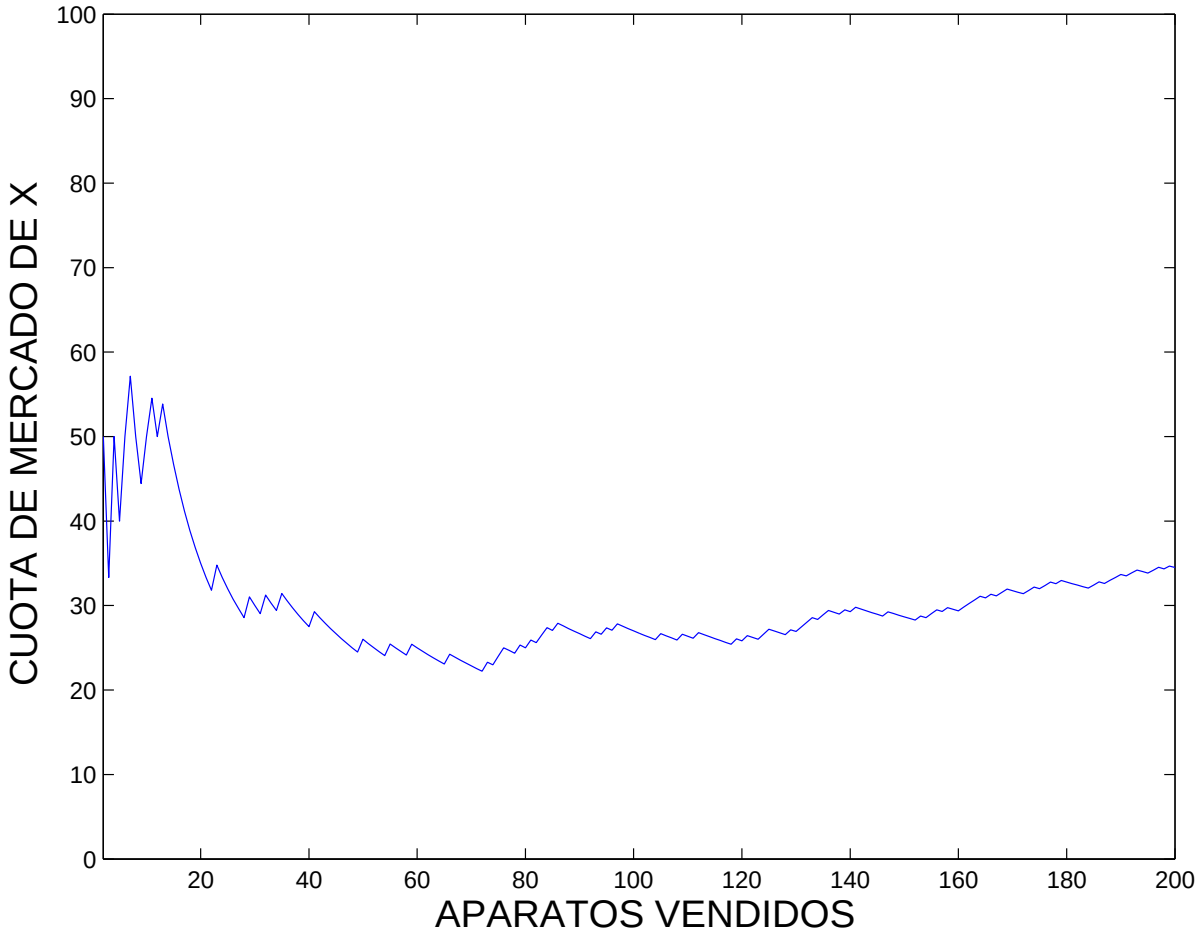
• La figura que hemos visto es por complete resultado del mero azar.

• Si los encuentros se producen de modo diverso . . .

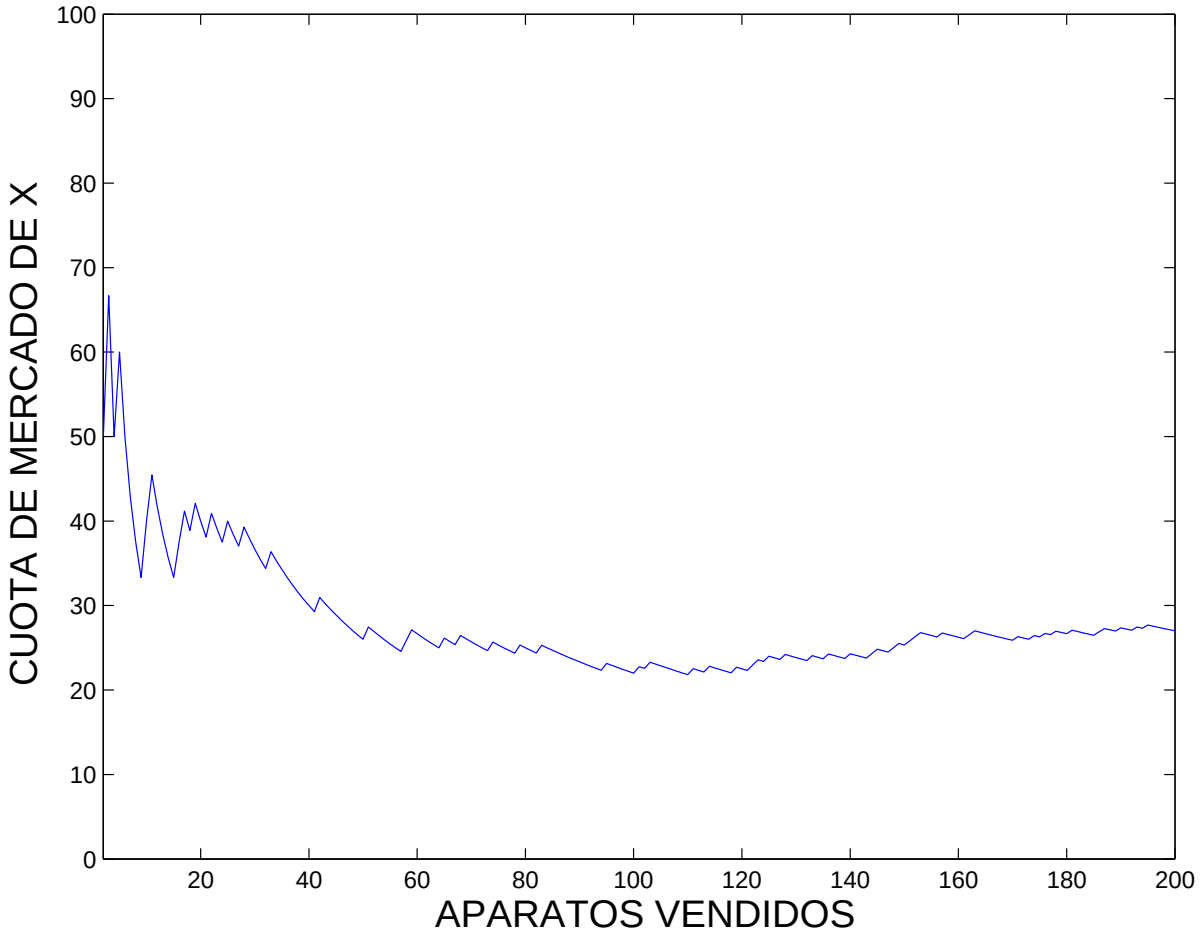
LAS COSAS PODRÍAN HABER SIDO ASÍ:



O ASÍ:


CUOTA DE MERCADO DE X
APARATOS VENDIDOS
0
10
20
30
40
50
60
70
80
90
100
20
40
60
80
100
120
140
160
180
200

O ASÍ:


CUOTA DE MERCADO DE X
APARATOS VENDIDOS
0
10
20
30
40
50
60
70
80
90
100
20
40
60
80
100
120
140
160
180
200

EL CALCULO DE PROBABILIDADES PRUEBA QUE:

• En (casi) todas las repiticiones del experimento la cuota de mercado tiene límite.

• El valor concreto de ese límite varía de repitición a repetición y todos los valores son 'uniformente probables'. (Es decir la probabilidad de obtener un valor del límite entre 0% y 50% es la misma que la de obtener un valor entre 50% y 100% y el doble de la obtener un límite entre 0% y 25% o entre 10% y 35%, . . .)

UNA CONCLUSIÓN MATEMÁTICA Y OTRA PSICOLÓGICA:

• En un fenómeno aleatorio la predicción de un caso singular es, por definición, imposible; sin embargo las *colecciones* de eventos aleatorios a menudo exhiben gran regularidad.

• Nuestra mente nos lleva siempre a buscar explicaciones *causales* y, por el contrario, encuentra de modo invariable dificultades con lo aleatorio. (La identificación de causas tiene ventajas evolutivas obvias.)

II: AL ORDEN POR EL DESORDEN

LOS EJEMPLOS DE REGULARIDAD ESTADÍSTICA NOS RODEAN:

• En los fenómenos de difusión (de animales en un habitat homogéneo, de un soluto en una disolvente, etc.), los movimientos individuales (de los animales, moléculas, etc.) son enteramente aleatorios. Y fruto de la aleatoriedad se llega a una distribución uniforme . . . que nosotros tendemos a concebir como resultado de la voluntad de los individuos de alejarse entre sí.

• A la misma categoría pertencen las mezclas y diversos fenómenos térmicos.

LEY [DÉBIL] DE LOS GRANDES NÚMEROS:

- Jakob Bernoulli (1654-1705), Ars Conjectandi 1713.

- Resultado matemático más antiguo y simple de los que nos permiten garantizar el compartamiento regular de de una colectividad de eventos aleatorios.

LA LEY EN EL EJEMPLO MÁS SENCILLO:

• Se arroja una moneda N veces y se obtienen C caras.

• La probabilidad de que la fracción C/N quede en el intervalo entre 0.4 y 0.6 (es decir con error inferior o igual a ± 0.10 frente al valor esperado 0.5) o en le intervalo entre 0.45 y 0.55 (es decir con error inferior o igual a ± 0.05) resulta ser:

N		
40	84.61%	57.04%
80	94.33%	68.57%
160	99.11%	82.12%
320	99.97%	93.51%
640	$\cdots$	98.99%

LA LEY EN GENERAL:

• X_1, X_2, ..., X_n, ... variables aleatorias que recogen el resultado de repeticiones sucesivas e independientes de un experimento.

• E el valor esperado de cada una de las X_n.

• $\epsilon > 0$ margen de error aceptable.

• La probabilidad de que el valor del promedio $(X_1 + \cdots + X_n)/n$ se aparte de E en una cantidad inferior o igual a ϵ se acerca al 100% cuando n crece.

• La variable $(X_1 + \cdots + X_n)/n$, para n grande, es aproximadamente constante e igual a E.

III: A VUELTAS CON LO ALEATORIO

NO TODOS LOS RESULTADOS ALEATORIOS NOS PARECEN IGUALMENTE ALEATORIOS:

• Un determinado día (digamos el 15 de enero de 2012) nacen ocho bebés en un mismo hospital, cuatro son varones V y cuatro hembras H. ¿Cuál de los órdenes siguentes es más probable?

$VVVVMMMM$

$VMVMVMVM$

$VVMVMMVM$

• ¡Las dos primeras sucesiones no nos parecen aleatorias!

RACHAS:

- He aquí 100 símbolos 0, 1 generados aleatoriamente (prob. 50%):

0101010111000101110 0

0101110001010000111 0

1101000000001100010 0

1000001110001000111 0

1010100110110000111 1

- ¡Notemos una racha de ocho ceros y otra de cinco!

MÁS RACHAS:

• Cuando una persona trata de escribir una sucesión aleatoria de ceros y unos el resultado es, por la ausencia de rachas, claramente distinguible de una sucesión realmente aleatoria.

• Estudios (Gilovich, Vallone, Tversky 1985) prueban que los periodos de buen o mal juego de los deportistas se explican por las fluctuaciones del azar.

• Durante los bombardeos de Londres en la segunda Guerra Mundial se observaron zonas (rachas) libres de bombas. Se buscaron las causas cuando no hay otra explicación que el azar. (W. Feller: 'To the untrained eye randomness appears as regularity or tendency to cluster.')

CÁNCER I:

• (A. Gelman & D. Nolan, Teaching Statistics: A Bag of Tricks) Un estudio de la incidencia del cáncer de riñón en la década 1980–1990 en EEUU muestra que, de los 3141 cóndados, aquellos con en la enfermedad fue más infrecuente eran rurales, de baja población, tradicionalmente republicanos y situados en el centro o sur del país.

• Vida sana, alimentos no contaminados, . . . ?

CÁNCER II:

- El mismo estudio muestra que, de los 3141 cóndados, aquellos con en la enfermedad fue más frecuente eran rurales, de baja población, tradicionalmente republicanos y situados en el centro o sur del país.

- Tabaquismo, ausencia de atención médica adecuada,. . . ?

LA LEY DE LOS PEQUEÑOS NÚMEROS:

• Muchas repeticiones del mismo experimento hacen que el comportamiento promedio de la muestra se acerque a la verdadera media (esperanza matemática) (Ley Fuerte).

• Por ello las grandes desviaciones aparecen cuando el número de repeticiones es bajo (condados con baja población en los que un sólo caso de cáncer les hace ponerse muy por encima de la media y cero casos muy por debajo de la media).

• De la información proporcionada tendemos a no dar importancia precisamente al dato más relevante: que los condados eran de baja población.

TAXIS I:

- (Tversky & Kahneman 1980) Un taxi atropella a una persona de noche y se da a la fuga.

- Hay dos empresas de taxis, Verde y Azul, 85% de los taxis son verdes y el 15% azules.

- Un testigo afirma que el culpable es azul. Se comprueba que el testigo identifica correctamente el color en el 80% de casos.

- ¿Cuál es la probabilidad de que el taxi fuera realmente azul?

TAXIS II:

- La respuesta típica es 80%.

- La verdadera 41% (Teorema de Bayes).

- Ignoramos el dato relevante de que la mayoría de taxis son verdes.

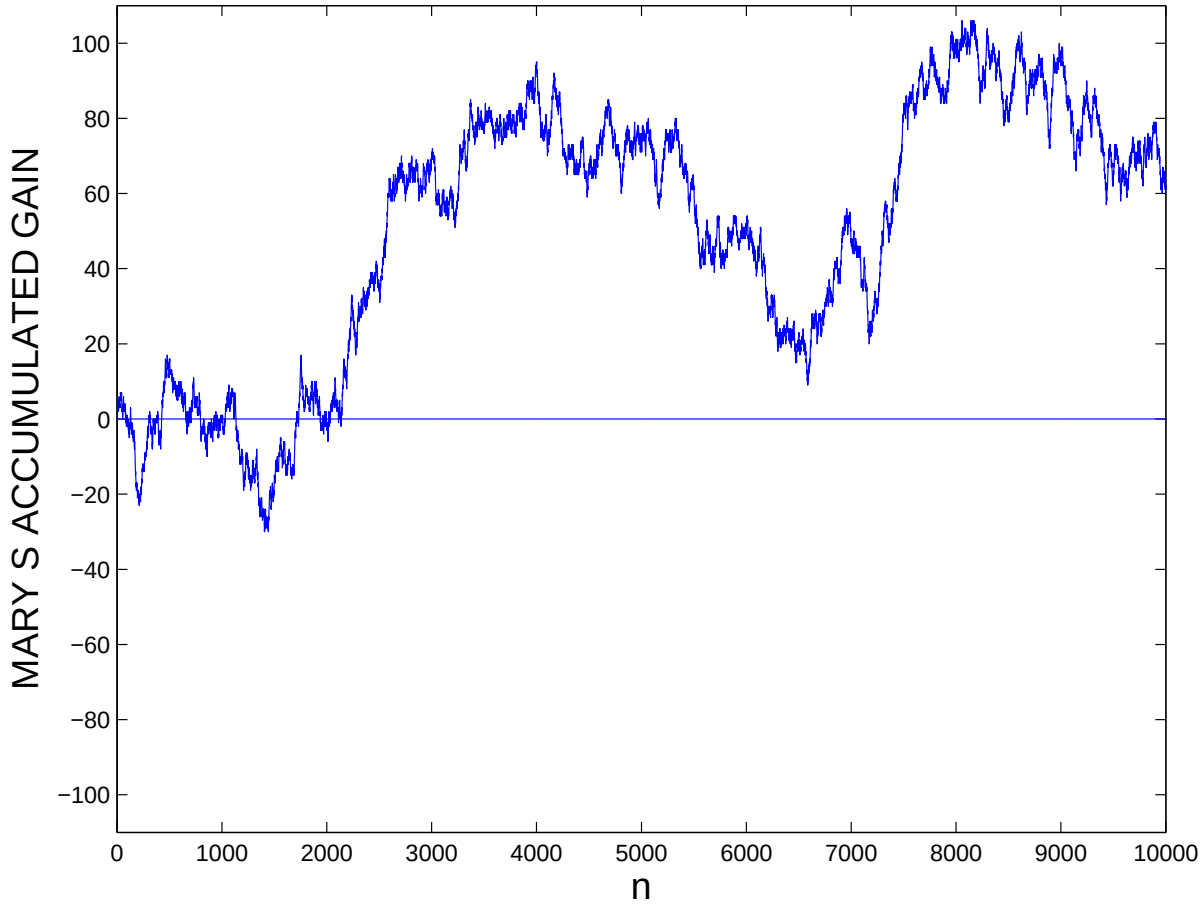
- Si se cambia la formulación del problema suprimiendo el dato 'el 85% de los taxis son verdes' y reemplazandolo por 'los taxis verdes originan el 85% de los accidentes,' la respuesta verdadera sigue siendo 41%, pero la respuesta típica es muy inferior a 80% y se acerca a la verdadera. (Pensamos que los verdes son culpables.)

CARA Y CRUZ:

- Mary y Pedro juegan a cara o cruz. Si sale cara Mary entrega un euro a Pedro y si cruz Pedro a Mary.

- He aquí un resultado típico pero inesperado:

GANANCIAS ACULMULADAS DE MARY:



OBSERVAMOS:

• Mary ha ido ganando la mayor parte del tiempo en la partida. Es falso que porque el juego sea simétrico cada jugador vaya por delante la mitad del tiempo en una partida dada. Típicamente en cada partida uno de los jugadores va por delante la mayor parte del tiempo.

• La simetría conduce a que en la mitad de las partidas sea Mary quien vaya por delante casi siempre y en la otra mitad sea Pedro.

PARA ACABAR, DOS ACERTIJOS:

- La gráfica contiene una ilustración del cumplimiento de la Ley de los Grandes Números . . .

- Y una ilustración de su incumplimiento . . .