CUADERNOS DE TRABAJO

FACULTAD DE ESTUDIOS ESTADÍSTICOS

Jugando con la Estadística (y la Probabilidad)

Gloria Cabrera Gómez
Departamento de Matemática Aplicada
Facultad de Estudios Estadísticos
Universidad Complutense de Madrid

Cuaderno de Trabajo número 04/2014

UNIVERSIDAD
COMPLUTENSE
MADRID

Los Cuadernos de Trabajo de la Facultad de Estudios Estadísticos constituyen una apuesta por la publicación de los trabajos en curso y de los informes técnicos desarrollados desde la Facultad para servir de apoyo tanto a la docencia como a la investigación.

Los Cuadernos de Trabajo se pueden descargar de la página de la Biblioteca de la Facultad www.ucm.es/BUCM/est/ y en la sección de investigación de la página del centro www.ucm.es/centros/webs/eest/

Biblioteca de la Facultad de Estudios Estadísticos

Universidad Complutense de Madrid

Av. Puerta de Hierro, S/N

28040 Madrid

Tlf. 913944035

buc_est@buc.ucm.es

Los trabajos publicados en la serie Cuadernos de Trabajo de la Facultad de Estudios Estadísticos no están sujetos a ninguna evaluación previa. Las opiniones y análisis que aparecen publicados en los Cuadernos de Trabajo son responsabilidad exclusiva de sus autores.

ISSN: 2341-2550

Jugando con la Estadística

(y la Probabilidad)

Gloria Cabrera Gómez cabrera@ucm.es
Universidad Complutense de Madrid
Facultad de Estudios Estadísticos
Departamento de Matemática Aplicada

0. Introducción

Los siguientes son un conjunto de ejercicios recopilados para la las conferencias impartidas con el mismo título, durante la Semana de la Ciencia 2013, con el objeto de divulgar, motivar y atraer hacia la Estadística y la Probabilidad.

La primera recopilación y exposición de esta serie de atractivos ejercicios empezó con el taller de Juegos y Experimentos Didácticos de Estadística y Probabilidad, impartido en la Facultad de Estudios Estadísticos con motivo del día Mundial de la Estadística, celebrado el día 20 de Octubre de 2.010. Y fue publicado como cuaderno de trabajo con el mismo título.

Durante la conferencia Jugando con la Estadística hubo dos formas de exposición de los ejercicios: algunos fueron presentados como enunciados llamativos, mientras que otros se expusieron como actividades o juegos de magia, explicando en un segundo momento, el contenido probabilístico subyacente.
Para evitar repeticiones, algunos de los enunciados que se expusieron en estas conferencias no aparecen en esta publicación porque ya estaban recogidos en el cuaderno de trabajo del primer taller.

Para dar continuidad a este serie de ejercicios se recomienda al lector interesado visitar la publicación Juegos y Experimentos Didácticos de Estadística y Probabilidad en http://estudiosestadisticos.ucm.es/data/cont/docs/12-2013-02-06-CT01_2011.pdf

Empezamos esta exposición con unos ejercicios que ilustran que a menudo los números contradicen nuestra intuición

1. El diagnóstico médico (Teorema de Bayes)

Nos hacen una prueba para averiguar si padecemos una grave enfermedad que afecta a una de cada 200 personas.
Damos positivo en la prueba. ¿Debemos asustarnos?

Los aciertos de una prueba diagnóstica son de dos tipos:

Sensibilidad es la probabilidad de acertar sobre los enfermos.

Y la especificidad es la probabilidad de acertar sobre los no enfermos.

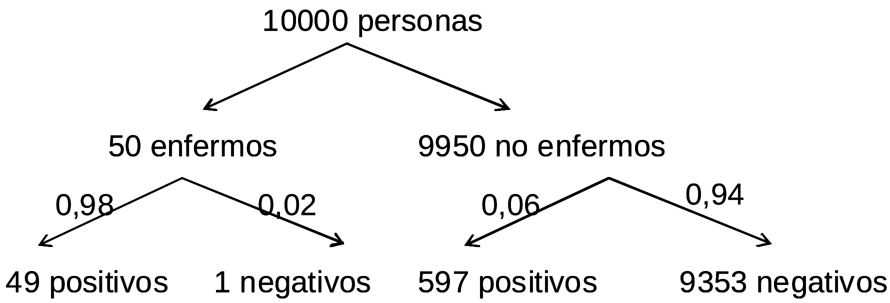
Supongamos que esta prueba tiene una sensibilidad del 98% y una especificidad del 94%

De cada 10.000 personas, unas 50 tendrán la enfermedad. De ellas, 49 obtendrán un resultado positivo en la prueba y una dará negativo.

En cuanto a la población sana (9.950 personas), 9.353 darán negativo y 497 positivo.

Luego de las personas diagnosticadas del mal en ese análisis (los 646 que han dado positivos) en realidad 597 estará sanos (el 92,5%). Solo estarán realmente enfermos el 7,5% (49 de 646)

Es decir el 92.5% son falsos positivos



Usando el teorema de Bayes:

E = Tener la Enfermedad	$P(E) = 0,005$	$P(\overline{E}) = 0,995$
+Dar Positivo	$P(+/E) = 0,98$	$P(-/E) = 0,02$
- Dar negativo	$P(-/\overline{E}) = 0,94$	$P(+/\overline{E}) = 0,06$

$$P(E/+) = \frac{P(+/E)\cdot P(E)}{P(+)} = \frac{P(+/E)\cdot P(E)}{P(+/E)\cdot P(E)+P(+/\overline{E})\cdot P(\overline{E})} = \frac{0,98\cdot 0,.005}{0,98\cdot 0,.005+0,06\cdot 0,995} = 0,0758$$

Es decir habiendo dado positivo en la prueba, la probabilidad de estar realmente enfermo es del 7,5%.

Y la de estar sano entonces es del 92,5%

A veces nos dejamos sorprender por las coincidencias, hechos que nos parecen increíblemente casuales. Un cierto conocimiento de la probabilidad de ocurrencia de estos sucesos haría que no nos maravilláramos por algunos y sí por otros

2. Dos barajas de cartas

Dos barajas de cartas separadas, bien barajadas cada una de ellas.

Vamos sacando cartas simultáneamente de uno y otro mazo.

¿Es probable o improbable que coincidan la misma carta de las dos barajas en la misma posición, es decir en el mismo momento?

¿Se trata de un truco de magia? O ¿se trata mas bien de sucesos muy probables?

Versión A: El experimento es el mismo que considerar una sola baraja española, numerar las cartas de 1 a 40 y analizar la probabilidad de que una vez barajadas al menos una de ellas ocupe su lugar de orden

$$P(\text{no coincidencias de lugar}) = \left(\tfrac{39}{40}\right)^{40} \simeq 36,3\%$$
$$P(\text{ al menos una coincidencia}) = 1 - \left(\tfrac{39}{40}\right)^{40} \simeq 63,7\%$$

Versión B: Y si las barajas son de póker

$$P(\text{no coincidencias de lugar}) = \left(\tfrac{51}{52}\right)^{52} \simeq 36,4\%$$
$$P(\text{ al menos una coincidencia}) = 1 - \left(\tfrac{51}{52}\right)^{52} \simeq 63,6\%$$

3. Los sombreros

30 personas van a una fiesta y dejan su sombrero en un perchero. A la salida, cada una toma uno sin fijarse bien si es el suyo. ¿Qué probabilidad hay de que ninguna acierte?

Nótese que es otro enunciado del mismo problema de las dos barajas

$$P(\text{ninguno lleve su sombrero}) = \left(\frac{29}{30}\right)^{30} \simeq 36,1\%$$
$$P(\text{ al menos una persona acierte}) = 1 - \left(\frac{29}{30}\right)^{30} \simeq 63,9\%$$

Por cierto, es aproximadamente la misma que la probabilidad de que solo una persona acierte con su sombrero.

$$P(\text{ solo una acierte con su sombrero}) = \binom{30}{1}\frac{1}{30}\left(\frac{28}{29}\right)^{29} \simeq 36,1\%$$

El siguiente va dedicado a los "locos de los números y las matrículas"

4. Matrículas de coches

¿Qué probabilidades hay de encontrar dos coches con los mismos números en sus matrículas?

Reflexionemos sobre la probabilidad de encontrar cerca (en la misma calle, en el mismo plano de una foto) dos coches con los mismos números en sus matrículas.

Tengamos en cuenta que, para empezar a la mayoría de nosotros nos resultará prácticamente imposible; no estamos acostumbrados a ir por la calle leyendo los números de las matrículas, a lo que hay que añadir que nuestra capacidad de retener y recordar números ya leídos es escasa con lo que el ejercicio en realidad se limita a ¿en cada grupo de 10 o 15 coches que abarquemos visualmente en el mismo campo, con qué probabilidad encontraremos dos coches con los mismos números?

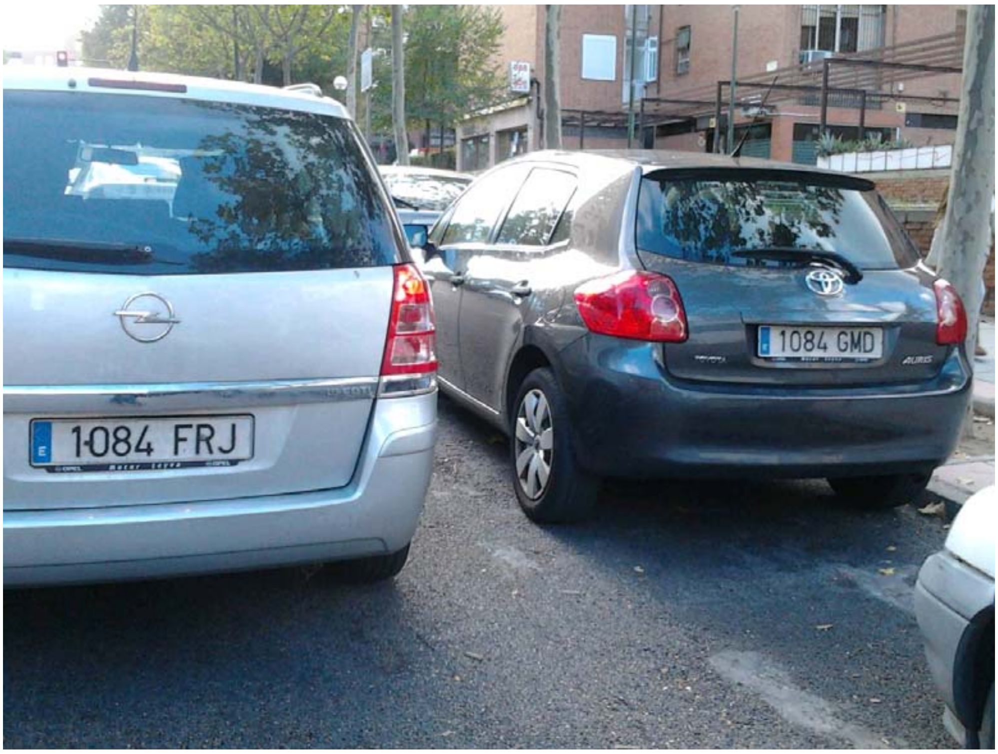
Dejamos ahí el plantemanento, sin entrar a detallar los cálculos que serían muy vagos e imprecisos.

Ilustremos las siguientes situaciones para reflexionar sobre ello; y sobre el hecho de que los sucesos pueden ser raros, puedesn ser poco probables, pero por muy pequeña que sea una probabilidad, el suceso puede darse.

1) Probabilidad de esto suceda una vez.

2) Probabilidad de esto suceda dos veces (y en la misma calle). Otros dos coches distintos con sus dos números de matrículas iguales.

3) Probabilidad de esto suceda siendo uno de los coches el tuyo



Invito a reflexionar también sobre las similitudes entre estas tres circunstancias y las versiones A y B del famoso ejercicio de las coincidencias en las fechas de cumpleaños que aparece en Juegos y Experimentos Didácticos de Estadística y Probabilidad (*)

A veces el enunciado es engañoso

5. El chico de las dos novias

Un chico tiene dos novias una en cada punta de la ciudad.

Cada día va a visitar a una, pero le es indiferente visitar a una o a otra.

Como durante todo el día pasan trenes en ambas direcciones y con la misma frecuencia, el chico ha decidido tomar cada vez el primer tren que llegue, y así el azar decide a qué novia visita cada día.

¿Cómo es posible que poco después una de las novias, realmente enamorada, se queje de que solo ha acudido a una cuarta parte de las citas, mientras la otra que ha empezado a aburrirse, se queje de que la ha visitado tres veces de cada cuatro?

La respuesta puede ser simple.

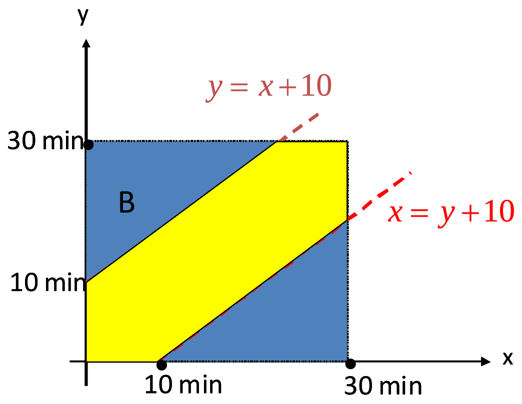
Puede ser que los trenes sean igual de frecuentes todo el día, pero no separados en el tiempo por intervalos iguales.

Suponiendo que pasen trenes cada 8 minutos puede ser que la distribución horaria sea:
A: (Hacia una de las citas) 12:00 - 12:08 - 12:16 – etc.
B: (Hacia la otra cita) 12:02 – 12:10 - 12:18 - …
Por tanto, de manera aleatoria, el chico llegará a la estación con probabilidad ¼ cuando ha pasado el tren para A y tomará el de B.
Y con probabilidad ¾ llegará a la estación en el intervalo de espera al tren para A.

6. Un encuentro

En un intervalo de tiempo de 30 minutos llegan a un mismo punto de encuentro y de forma aleatoria dos personas.
¿Qué probabilidad existe de que una de las dos personas espere a la otra al menos 10 minutos?

Este es una variante del poblema de la cita publicado en Juegos y Experimentos Didácticos de Estadística y Probabilidad (*)



$$E = \{(x,y) : 0 \le x \le 30,\ 0 \le y \le 30\}$$

$$B = \{(x,y) : |y-x| \ge 10\} =$$
$$= \{(x,y) : y-x \ge 10\} \cup \{(x,y) : x-y \ge 10\}$$

$$P(B) = \frac{Area(B)}{Area(E)} = \frac{20^2}{30^2} = \frac{4}{9} \simeq 44\%$$

Reflexiones sobre la probabilidad condicionada

7. El dilema del prisionero

En una cárcel, tres prisioneros de historiales similares, solicitan el indulto a un tribunal.

Poco después se sabe que el indulto ha sido concedido a dos de los tres presos.

Uno de los prisioneros conoce a uno de los miembros del tribunal y sabe que si le pregunta podrá obtener algo de información. Podrá preguntarle por el nombre de uno de los indultados, pero no podrá preguntar si él es uno de ellos.

Reflexionando concluye que

Si no pregunta, la probabilidad de ser uno de los indultados es 2/3.

- Mientras que si pregunta, obtendrá respuesta, y entonces la probabilidad de ser el otro de los indultados es ½.

Por tanto concluye que será mejor no preguntar, pues que eso solo le servirá para disminuir su probabilidad de ser uno de los indultados.

¿Es correcto su razonamiento? ¿Dónde está el error?

¿Es similar este problema al del concurso o al problema de las tres puertas?

A diferencia del problema del concurso (véase Juegos y Experimentos Didácticos de Estadística y Probabilidad *) en este enunciado no hay probabilidad condicionada. El conocimiento de la ocurrencia de un suceso (nombre de un idultado) no condiciona la probabilidad de que nuestro preso sea o no indultado. La probabilidad es la misma a priori que a posteriori, no ha cambiado.

Un enunciado curioso para un ejercicio sencillo

8. Las bodas de Manchuria

Cuenta la leyenda que en Manchuria, cuando una mujer quería casarse, su padre la sometía a la siguiente prueba.

Toma tres cuerdas, dobladas por la mitad y sujetas de modo que no se vea la doblez, con los 6 extremos colgando.

La candidata a esposa tiene que tomar los cabos de dos en dos y atarlos.

Solo obtendrá permiso para casarse si consigue atarlos formando una gran circunferencia.

¿Hay muchas mujeres solteras en Manchuria?

Se podrán obtener 3 configuraciones posibles:

A= Un gran circunferencia $P(A) = \frac{8}{C_{6.2}} = \frac{8}{15}$

B=Una pequeña y una mediana $P(B) = \frac{6}{15}$

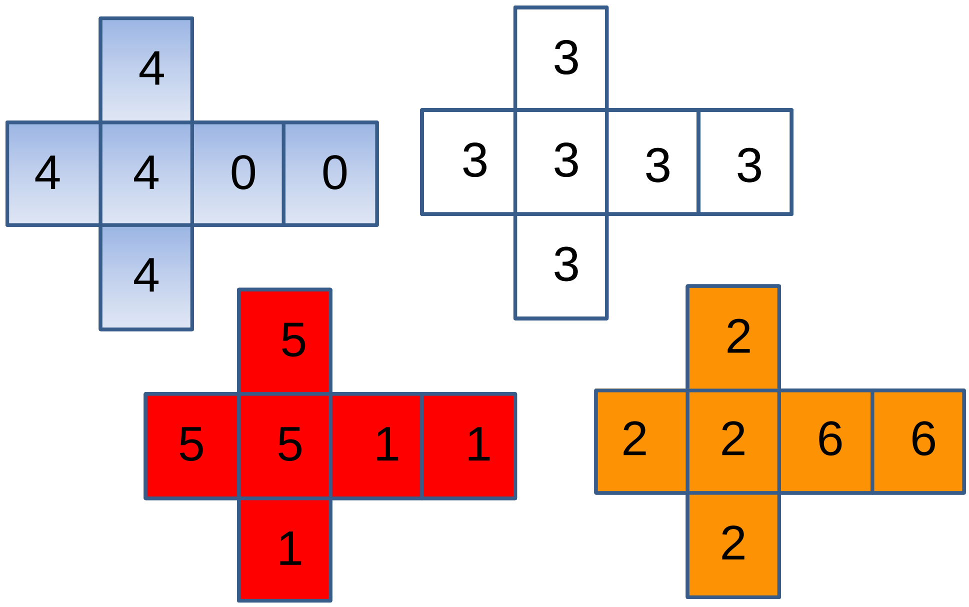
C=Tres pequeñas $P(C) = \frac{1}{15}$

Mas del 50% de las mujeres obtendrán el permiso de boda

A continuación unos juegos con dados, con monedas y con cartas

9. Probabilidad ventajosa

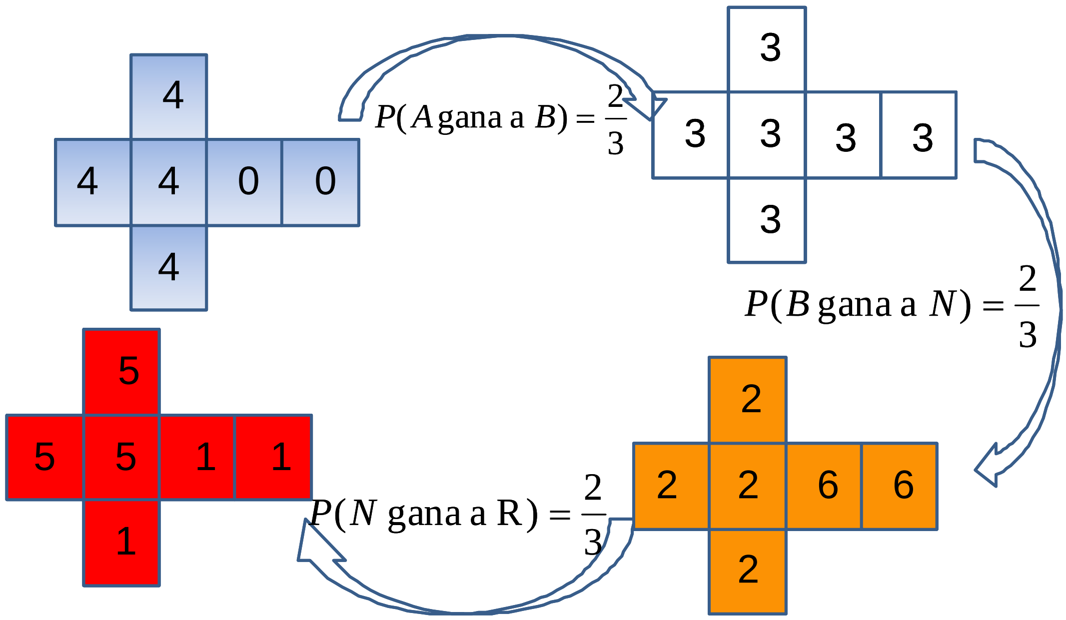
Dos jugadores eligen un dado cado uno, de entre los 4 dados siguientes.
¿Quién tiene ventaja? ¿El jugador que elija dado el primero o el que lo elija el segundo?



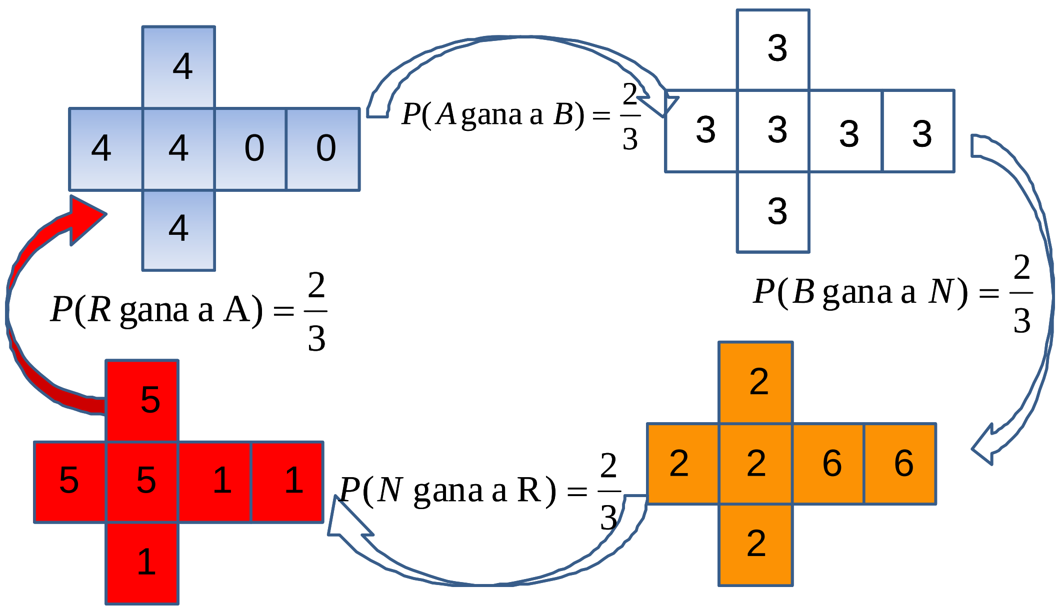
Podemos afirmar, sin dudas que si uno de los dados gana a todos los demás, el primer jugador en elegir tiene ventaja.

Podemos comprobar fácilmente que A gana a B; B gana a N y N gana a R.

¿Es razonable imaginar que entonces A gana a R. ¿Sucede entonces que A es el dado ganador?
Si es cierto, el jugador aventajado es el primero, puesto que ganará facilmente (o con mayor probabilidad que el segundo jugador) sin mas que elegir el dado A.



Sin embargo, **La probabilidad no es transitiva**.
Sucede que R gana a A
Por tanto el que tiene ventaja es el jugador que elija en segundo lugar.



---

10. Cómo jugar limpiamente con una moneda defectuosa

Dos jugadores deciden establecer un turno o un ganador entre los dos, usando para ello como criterio el azar determinado por el lanzamiento de una moneda. Cada uno de ellos elige cara o cruz, y gana el que acierte el resultado.

¿Qué sucede si la moneda es defectuosa?, ¿o se tienen dudas sobre ello?

Puede usarse este mecanismo (ideado por John Von Neumann)
Tírese la moneda dos veces:
- Si sale Cara-Cara o Cruz-Cruz se vuelve a tirar
- Si sale Cara –Cruz gana el jugador 1
- Si sale Cruz-Cara gana el jugador 2

Este método sí proporciona igualdad de condiciones para los dos jugadores puesto que $P(CX) = P(XC)$

Los juegos con cartas, habitualmente se disfrazan de magia. En algunas ocasiones se trata de hablidad. En otras son juegos probabilísticos como el siguiente.
Un juego se llama determinista cuando repitiendo un determinado algoritmo se obtiene siempre el mismo resultado. Puede entonces hacerse "magia" ante un público que desconoce que el resultado al que vamos es un suceso seguro.
El juego se llama probabilístico si el suceso final al que queremos llegar es un suceso muy probable, pero no seguro.
Y a veces puede incluso intervenir en la probabilidad final, la habilidad del "mago", como en el siguiente juego, donde la forma de mezclar las cartas es importante.

Expliquemos antes de empezar que una **mezcla americana** consiste en unir las cartas de dos montones de forma que en la nueva configuación haya unas pocas que vienen de la derecha, unas pocas de la izquierda, una pocas de la derecha...y así sucesivamente.
La mezcla se llama **faro** cuando las cartas caen una a una, es decir, una de la derecha, una de la izquierda, una de la derecha y así sucesivamente.

10. Magia al distribuir cartas de una baraja francesa

Contamos con una baraja francesa, en la que previamente hemos separado las cartas rojas de las negras (sin que el público lo sepa) y hemos colocando el montón de las 26 rojas sobre las 26 negras (o al revés).
Repartimos todas las cartas en 4 montones boca abajo, soltándolas una a una sin que el orden tenga que ser riguroso, es decir no necesariamente hacemos ciclos exactos en el mismo orden. Solo cuidaremos de que haya mas o menos el mismo número de cartas en cada montón.
A continuación mezclamos mediante una mezcla america* dos de los cuatro montones; y entre ellos otra mezcla con los dos montones restantes. Finalmente con otra mezcla americana unimos los dos montones resultantes, en uno.

Y ahora vamos a construir 3 montones nuevos: En el primero colocamos 16 cartas, dejándolas de una en una; en el segundo montón colocamos 20 cartas y en el tercer montón colocamos las 16 cartas restantes.

Y ahora podemos sorprender al público afirmando que al descubrir las cartas de los tres montones, serán:

Todas del mismo color en los montones primero y tercero (en un montón todas serán rojas y en el otro todas serán negras). Mientras que en el montón central tendremos la mitad de las cartas de cada color.

Este se trata, efectivamente de un juego probabilístico en el que el resultado anunciado se consigue con una alta probabilidad, en función de la habilidad al realizar las mezclas.Y cuanta mas habilidad tenga " el mago" para acercarse a una mezcla faro mas seguro está del resultado definitivo. Y será mas vistoso.

Para profundizar en las explicaciones del juego remitimo al arículos Is it ratdom? de Fernando Blasco (http://www.seio.es/BEIO/files/BEIOVol28Num3_Opinion.pdf)

---

11. Bibliografía

- Alegria, P y Ruiz de Arcahuete, JC La magia desvelada.
  http://www.ehu.es/~mtpalezp/mates/lamat.pdf
- Aigner, Martin y Ziegler, Günter M. El libro de las demostraciones. Madrid : Nívola Libros, 2005
- Blasco, Fernando. Matemagia.
- Gonick, Larry y Smith, Woollcott. La estadística en cómic.Barcelona : Zendrera Zariquiey, 2006
- Haigh, John. Matemáticas y juegos de azar : jugar con la probabilidad. Barcelona : Tusquets, 2008
- Paulos,.John Allen. El hombre anumérico : el analfabetismo matemático y sus consecuencias .Barcelona : Tusquets Editores, 2000
- Salazar González, JJ y López Yurda; M. Ejercicios Resueltos de Probabilidad
- http://www.seio.es/BEIO/Is-it-random.html
- http://www.fblasco.com/
- http://magiaymatematicas.blogspot.com.es/
- DivulgaMat, El rincón matemágico:
  http://divulgamat2.ehu.es/divulgamat15/index.php?option=com_alphacontent&section=11
- Matemagia enredadora y enredos matemágicos - Arundanet:
  www.arundanet.com/matesxronda/matemagia.php

Cuadernos de Trabajo
Facultad de Estudios Estadísticos

CT03/2014	**Análisis estadístico de las consultas a la base de datos de ayudas e incentivos a empresas de la Dirección General de Industria y de la PYME.** *Adolfo Coello de Portugal Muñoz, Juana María Alonso Revenga*
CT02/2014	**Values of games with weighted graphs** *E. González-Arangüena, C. Manuel y M. del Pozo*
CT01/2014	**Estimación de la tasa de retorno de la carta del censo de los Estados Unidos a través del modelo de regresión lineal y técnicas de predicción inteligentes.** *José Luis Jiménez-Moro y Javier Portela García Miguel*
CT03/2013	**Provisión de siniestros de incapacidad temporal utilizando análisis de supervivencia.** *Ana Crespo Palacios y Magdalena Ferrán Aranaz*
CT02/2013	**Consumer need for touch and Multichannel Purchasing Behaviour.** *R. Manzano, M. Ferrán y D. Gavilán*
CT01/2013	**Un método gráfico de comparación de series históricas en el mercado bursátil.** *Magdalena Ferrán Aranaz*
CT03/2012	**Calculando la matriz de covarianzas con la estructura de una red Bayesiana Gaussiana** *Miguel A. Gómez-Villegas y Rosario Susi*
CT02/2012	**What's new and useful about chaos in economic science.** *Andrés Fernández Díaz, Lorenzo Escot and Pilar Grau-Carles*
CT01/2012	**A social capital index** *Enrique González-Arangüena, Anna Khmelnitskaya, Conrado Manuel, Mónica del Pozo*
CT04/2011	**La metodología del haz de rectas para la comparación de series temporales.** *Magdalena Ferrán Aranaz*
CT03/2011	**Game Theory and Centrality in Directed Social Networks** *Mónica del Pozo, Conrado Manuel, Enrique González-Arangüena y Guillermo Owen.*
CT02/2011	**Sondeo de intención de voto en las elecciones a Rector de la Universidad Complutense de Madrid 2011** *L.Escot, E. Ortega Castelló y L. Fernández Franco (coords)*
CT01/2011	**Juegos y Experimentos Didácticos de Estadística y Probabilidad** *G. Cabrera Gómez y Mª.J. Pons Bordería*
CT04/2010	**Medio siglo de estadísticas en el sector de la construcción residencial** *M. Ferrán Aranaz*
CT03/2010	**Sensitivity to hyperprior parameters in Gaussian Bayesian networks.** *M.A. Gómez-Villegas, P. Main, H. Navarro y R. Susi*

CT02/2010 **Las políticas de conciliación de la vida familiar y laboral desde la perspectiva del empleador. Problemas y ventajas para la empresa.**
R. Albert, L. Escot, J.A. Fernández Cornejo y M.T. Palomo

CT01/2010 **Propiedades exóticas de los determinantes**
Venancio Tomeo Perucha

CT05/2009 **La predisposición de las estudiantes universitarias de la Comunidad de Madrid a auto-limitarse profesionalmente en el futuro por razones de conciliación**
R. Albert, L. Escot y J.A. Fernández Cornejo

CT04/2009 **A Probabilistic Position Value**
A. Ghintran, E. González–Arangüena y C. Manuel

CT03/2009 **Didáctica de la Estadística y la Probabilidad en Secundaria: Experimentos motivadores**
A. Pajares García y V. Tomeo Perucha

CT02/2009 **La disposición entre los hombres españoles a tomarse el permiso por nacimiento. ¿Influyen en ello las estrategias de conciliación de las empresas?**
L. Escot, J.A. Fernández-Cornejo, C. Lafuente y C. Poza

CT01/2009 **Perturbing the structure in Gaussian Bayesian networks**
R. Susi, H. Navarro, P. Main y M.A. Gómez-Villegas

CT09/2008 **Un experimento de campo para analizar la discriminación contra la mujer en los procesos de selección de personal**
L. Escot, J.A. Fernández Cornejo, R. Albert y M.O. Samamed

CT08/2008 **Laboratorio de Programación. Manual de Mooshak para el alumno**
D. I. de Basilio y Vildósola, M. González Cuñado y C. Pareja Flores

CT07/2008 **Factores de protección y riesgo de infidelidad en la banca comercial**
J. Mª Santiago Merino

CT06/2008 **Multinationals and foreign direct investment: Main theoretical strands and empirical effects**
María C. Latorre

CT05/2008 **On the Asymptotic Distribution of Cook's distance in Logistic Regression Models**
Nirian Martín y and Leandro Pardo

CT04/2008 **La innovación tecnológica desde el marco del capital intelectual**
Miriam Delgado Verde, José Emilio Navas López, Gregorio Martín de Castro y Pedro López Sáez

CT03/2008 **Análisis del comportamiento de los indecisos en procesos electorales: propuesta de investigación funcional predictivo-normativa**
J. Mª Santiago Merino

CT02/2008 **Inaccurate parameters in Gaussian Bayesian networks**
Miguel A. Gómez-Villegas, Paloma Main and Rosario Susi

CT01/2008 **A Value for Directed Communication Situations.**
E. González–Arangüena, C. Manuel, D. Gómez, R. van den Brink

UNIVERSIDAD COMPLUTENSE