

Test de hipótesis

Curso de análisis de datos con R
Asociación Argentina de Bioinformática y Biología Computacional

Octubre 2020



Levi-Strauss en el Cerro Uritorco

Levi-Strauss y su equipo fueron invitados al Cerro Uritorco para investigar unos yacimientos.



De la web

Un esqueleto ¿humano?

En el yacimiento encontraron un esqueleto de 1.70 metros.



<https://www.lavanguardia.com/local/valencia/20150813/54435799231/esqueleto-humano-yacimiento-iberico-ademuz.html> y history.com (?)

Otro esqueleto ¿humano?

Siguieron buscando en el yacimiento y encontraron un esqueleto de 1.85 metros.



<https://elperiodicodevillena.com/> y [history.com](https://www.history.com) (?)

Un último esqueleto ¿humano?

Finalmente buscando en el yacimiento encontraron un esqueleto de ¡6 metros!.



National Geographic y history.com (?)

¿Podemos concluir que es extraterrestre?

¿Qué podemos concluir del esqueleto de 4 metros?
¿No es humano? ¿Es extraterrestre?



National Geographic y history.com (?)

¿Habrá dos personas que cumplan años el mismo día en el curso?



Paradoja del cumpleaños

Paradoja del cumpleaños – Wikipedia, la enciclopedia libre – Mozilla Firefox

W Paradoja del cumpleaños: X +

https://es.wikipedia.org/wiki/Paradoja_del_cumpleaños

No has accedido | Discusión | Contribuciones | Crear una cuenta | Acceder

Artículo | Discusión

Leer | Editar | Ver historial | Buscar en Wikipedia

Paradoja del cumpleaños

Este artículo o sección necesita **referencias** que aparezcan en una **publicación acreditada**.
Este aviso fue puesto el 3 de agosto de 2011.

El **problema del cumpleaños**, también llamado **paradoja del cumpleaños**, establece que de un conjunto de 23 personas, hay una probabilidad del 50.7% de que al menos dos personas de ellas **cumplan años** el mismo día. Para 57 o más personas la probabilidad es mayor del 99.666%. En sentido estricto esto no es una **paradoja** ya que no es una contradicción **lógica**: sin embargo, es una verdad matemática que contradice la común **intuición**. Mucha gente piensa que la probabilidad es mucho más baja, y que hacen falta muchas más personas para que se alcance la probabilidad del 50,666%. Si una habitación tuviera 367 personas, por el **Principio del palomar** sabemos que habría al menos dos personas cumpliendo años en la misma fecha, ya que un año normal tiene 365 días, y uno bisiesto tiene 366.

Índice [ocultar]

- 1 Estimación de la probabilidad
- 2 Programas informáticos
 - 2.1 Ruby
 - 2.2 Go
 - 2.3 Julia
 - 2.4 Pascal
 - 2.5 Python
 - 2.6 Perl
 - 2.7 C++
 - 2.8 C#
 - 2.9 Java
 - 2.10 PHP
 - 2.11 PL/SQL
 - 2.12 MATLAB
 - 2.13 Microsoft SQL Server
 - 2.14 R
 - 2.15 Swift
- 3 Enlaces externos

Estimación de la probabilidad ([editar](#))

Calcular esta probabilidad es el problema del cumpleaños. La teoría fue descrita en *American Mathematical Monthly* en 1938 en la teoría de Estimación del total de población de peces en un lago de **Zoe Emily Schnabel**, bajo el nombre de captura-recaptura estadística.

La clave para entender la paradoja del cumpleaños es pensar que hay muchas probabilidades de encontrar parejas que cumplan años el mismo día. Específicamente, entre 23 personas, hay $23 \times 22 / 2 = 253$ pares, cada uno de ellos un candidato potencial para cumplir la paradoja. Hay que entender que si una persona entrase en una habitación con 22 personas, la probabilidad de que cualquiera cumpla años el mismo día que quien entra, no es del 50%, es mucho más baja.



Número de personas	Probabilidad de al menos una coincidencia
1	0.0000
10	0.1078
20	0.3770
23	0.5073
30	0.7083
40	0.8912
50	0.9703
60	0.9941
70	0.9999
80	1.0000
90	1.0000
100	1.0000

https://es.wikipedia.org/wiki/Paradoja_del_cumpleaños

Resultados de un experimento aleatorio

Tiremos dos dados, ¿cuáles son los posibles resultados?

$\{(1,1),(1,2),(1,3),(1,4),(1,5),(1,6), (2,1),(2,2),(2,3),(2,4),(2,5),(2,6),$
 $(3,1),(3,2),(3,3),(3,4),(3,5),(3,6), (4,1),(4,2),(4,3),(4,4),(4,5),(4,6),$
 $(5,1),(5,2),(5,3),(5,4),(5,5),(5,6), (6,1),(6,2),(6,3),(6,4),(6,5),(6,6)\}$



Si ahora estamos interesados en la suma de los dos dados, ¿qué valores posibles obtengo?

¿De cuántas formas puedo construir el 2? ¿Y el 7? ¿La probabilidad de obtener cada uno es la misma?

Simulemos esto en R



Abrimos dia_4.R

Test de hipótesis y errores

Veamos qué tipo de situaciones se pueden dar al poner a prueba nuestras hipótesis.

	Declara inocente	Declara culpable
Es inocente	Bien	Mal
No es inocente	Mal	Bien

	No rechazamos H_0	Rechazamos H_0
H_0 es cierta	Ok	Error de Tipo I
H_0 no es cierta	Error de Tipo II	Ok

Los tests de hipótesis nos permiten controlar los errores de tipo I. Para eso tenemos que decidir previamente el porcentaje de veces que estamos dispuestos a cometer ese error. Se suele usar 5