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 (?)

Un último esqueleto ¿humano?

Finalmente buscando en el yacimiento encontraron un esqueleto de ¡4 metros!.



National Geographic y history.com (?)

¿Podemos concluir que es extraterrestre?

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

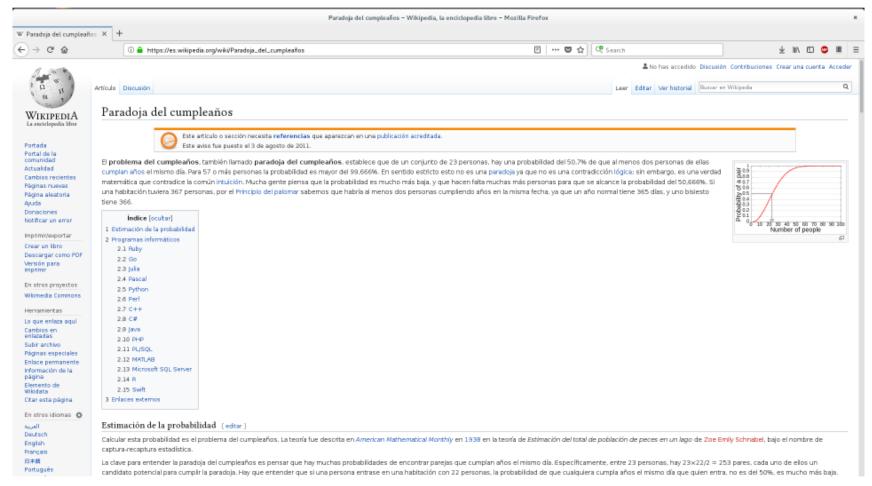


National Geographic y history.com (?)

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



Paradoja del cumpleaños



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 H0	Rechazamos H0
H0 es cierta	Ok	Error de Tipo I
H0 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