NEUROCIENCIA Y TOMA DE DECISIONES

EXTRA:
NANO-REPASO DE ESTADÍSTICA

¿QUÉ ES EL P-VALOR?

La mayoría de los fenómenos que podemos medir en la naturaleza siguen una cierta distribución estadística. Prácticamente NUNCA es gaussiana.

Sin embargo…

¿QUÉ ES EL P-VALOR?

La mayoría de los fenómenos que podemos medir en la naturaleza siguen una cierta distribución estadística. Prácticamente NUNCA es gaussiana.

Sin embargo…

$$\sum \begin{matrix} variables\ aleatorias\ (ind)\ con \\ distribuciones\ diferentes \end{matrix} = \color{red}{\begin{matrix} variable\ aleatoria\ con \\ distribución \sim gaussiana \end{matrix}}$$

¿QUÉ ES EL P-VALOR?

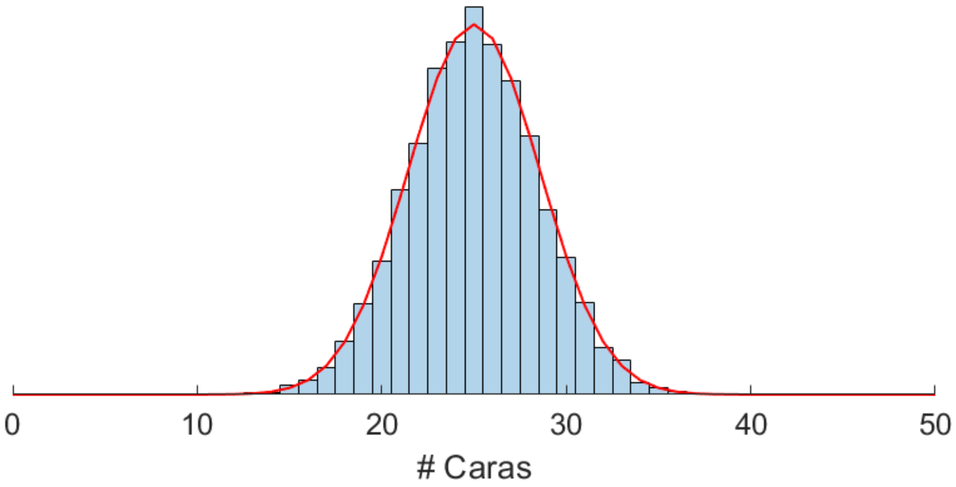
La mayoría de los fenómenos que podemos medir en la naturaleza siguen una cierta distribución estadística. Prácticamente NUNCA es gaussiana.

Sin embargo…

$$\sum \begin{matrix} variables\ aleatorias\ (ind)\ con \\ distribuciones\ diferentes \end{matrix} = \color{red}{\begin{matrix} variable\ aleatoria\ con \\ distribución \sim gaussiana \end{matrix}}$$

Pese a eso, fenómenos **puros** suelen seguir otras distribuciones, como por ejemplo…

BINOMIAL



¿QUÉ ES EL P-VALOR?

La mayoría de los fenómenos que podemos medir en la naturaleza siguen una cierta distribución estadística. Prácticamente NUNCA es gaussiana.

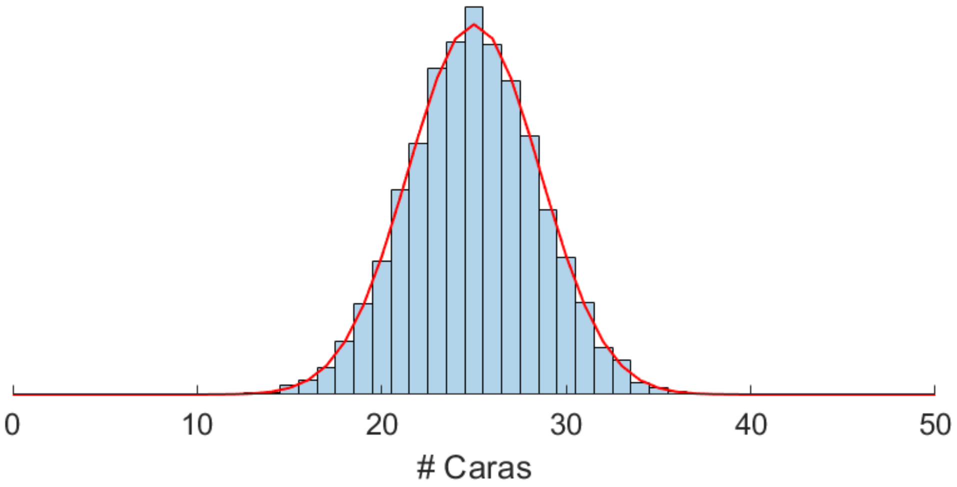
Sin embargo…

$$\sum \begin{matrix} variables\ aleatorias\ (ind)\ con \\ distribuciones\ diferentes \end{matrix} = \color{red}{\begin{matrix} variable\ aleatoria\ con \\ distribución \sim gaussiana \end{matrix}}$$

Pese a eso, fenómenos **puros** suelen seguir otras distribuciones, como por ejemplo…

Como las monedas…

BINOMIAL



¿QUÉ ES EL P-VALOR?

La mayoría de los fenómenos que podemos medir en la naturaleza siguen una cierta distribución estadística. Prácticamente NUNCA es gaussiana.

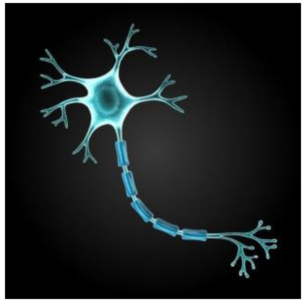
Sin embargo…

$$\sum \begin{matrix} variables\ aleatorias\ (ind)\ con \\ distribuciones\ diferentes \end{matrix} = \begin{matrix} variable\ aleatoria\ con \\ distribución \sim gaussiana \end{matrix}$$

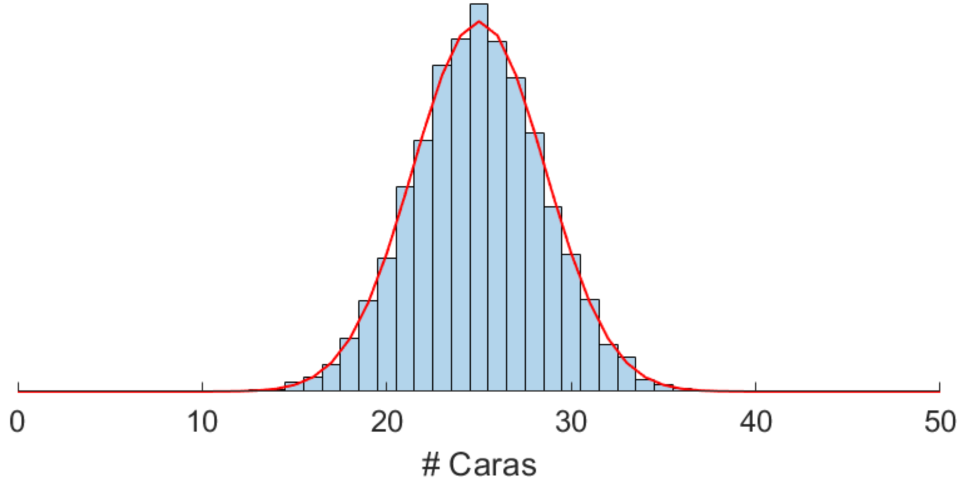
Pese a eso, fenómenos **puros** suelen seguir otras distribuciones, como por ejemplo…

Como las monedas…

O ¡¡las neuronas!!



BINOMIAL



¿QUÉ ES EL P-VALOR?

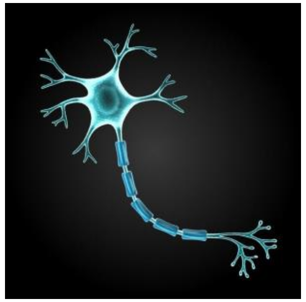
Definiendo el problema: supongamos que tenemos una neurona, y queremos saber si dispara frente a un cierto estímulo específico.

¿QUÉ ES EL P-VALOR?

Definiendo el problema: supongamos que tenemos una neurona, y queremos saber si dispara frente a un cierto estímulo específico.

H_0 = La neurona dispara aleatoriamente (p=0.5)

"Hipótesis Nula"

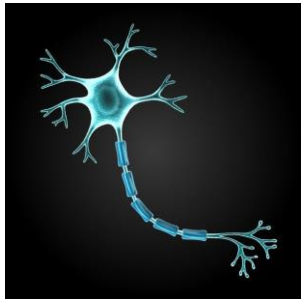


¿QUÉ ES EL P-VALOR?

Definiendo el problema: supongamos que tenemos una neurona, y queremos saber si dispara frente a un cierto estímulo específico.

H_0 = La neurona dispara aleatoriamente ($p=0.5$)

"Hipótesis Nula"



H_1 = La neurona responde al estímulo ($p \neq 0.5$)

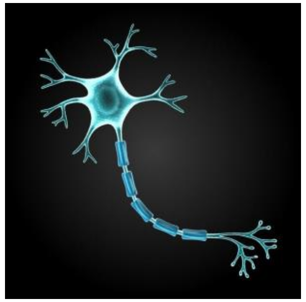
"Hipótesis Alternativa"

¿QUÉ ES EL P-VALOR?

Definiendo el problema: supongamos que tenemos una neurona, y queremos saber si dispara frente a un cierto estímulo específico.

H_0 = La neurona dispara aleatoriamente (p=0.5)

"Hipótesis Nula"



H_1 = La neurona responde al estímulo (p≠0.5)

"Hipótesis Alternativa"

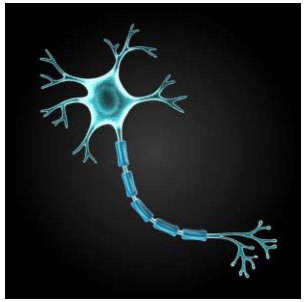
Necesitamos algo que sea sensible a la diferencia entre ambas hipótesis. Un "estadístico".

¿QUÉ ES EL P-VALOR?

Definiendo el problema: supongamos que tenemos una neurona, y queremos saber si dispara frente a un cierto estímulo específico.

H_0 = La neurona dispara aleatoriamente (p=0.5)

"Hipótesis Nula"



H_1 = La neurona responde al estímulo (p≠0.5)

"Hipótesis Alternativa"

Necesitamos algo que sea sensible a la diferencia entre ambas hipótesis. Un "estadístico".

Por ejemplo:

$$T = \frac{\text{\# total de disparos con el estímulo}}{\text{\# total de presentaciones del estímulo}}$$

Supongamos que presentamos el estímulo 10 veces, y 8 veces disparó. ¿Y ahora?

¿QUÉ ES EL P-VALOR?

$$\tau = \frac{\text{\# total de disparos con el estímulo}}{\text{\# total de presentaciones del estímulo}} = \frac{8}{10}$$

Y ahora sí…

¿QUÉ ES EL P-VALOR?

$$\tau = \frac{\text{\# total de disparos con el estímulo}}{\text{\# total de presentaciones del estímulo}} = \frac{8}{10}$$

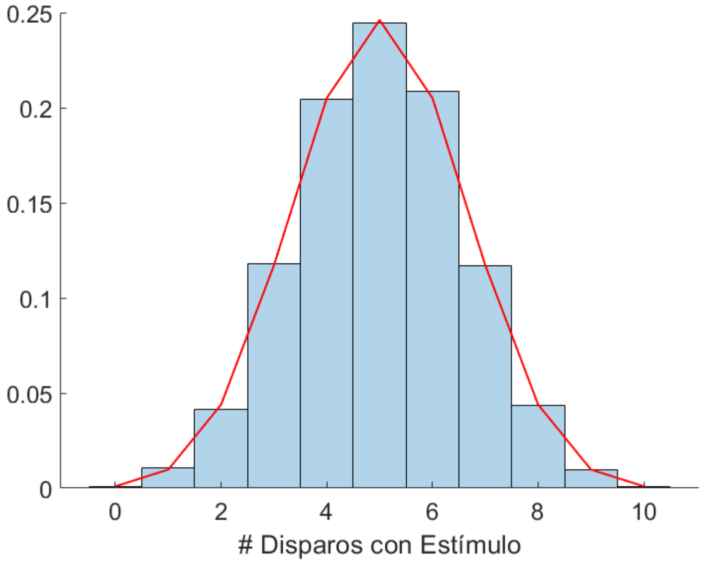
Y ahora sí…

El p-valor es la probabilidad de medir ese valor, o uno "más raro".

¿QUÉ ES EL P-VALOR?

$$T = \frac{\#\ \text{total de disparos con el estímulo}}{\#\ \text{total de presentaciones del estímulo}} = \frac{8}{10}$$

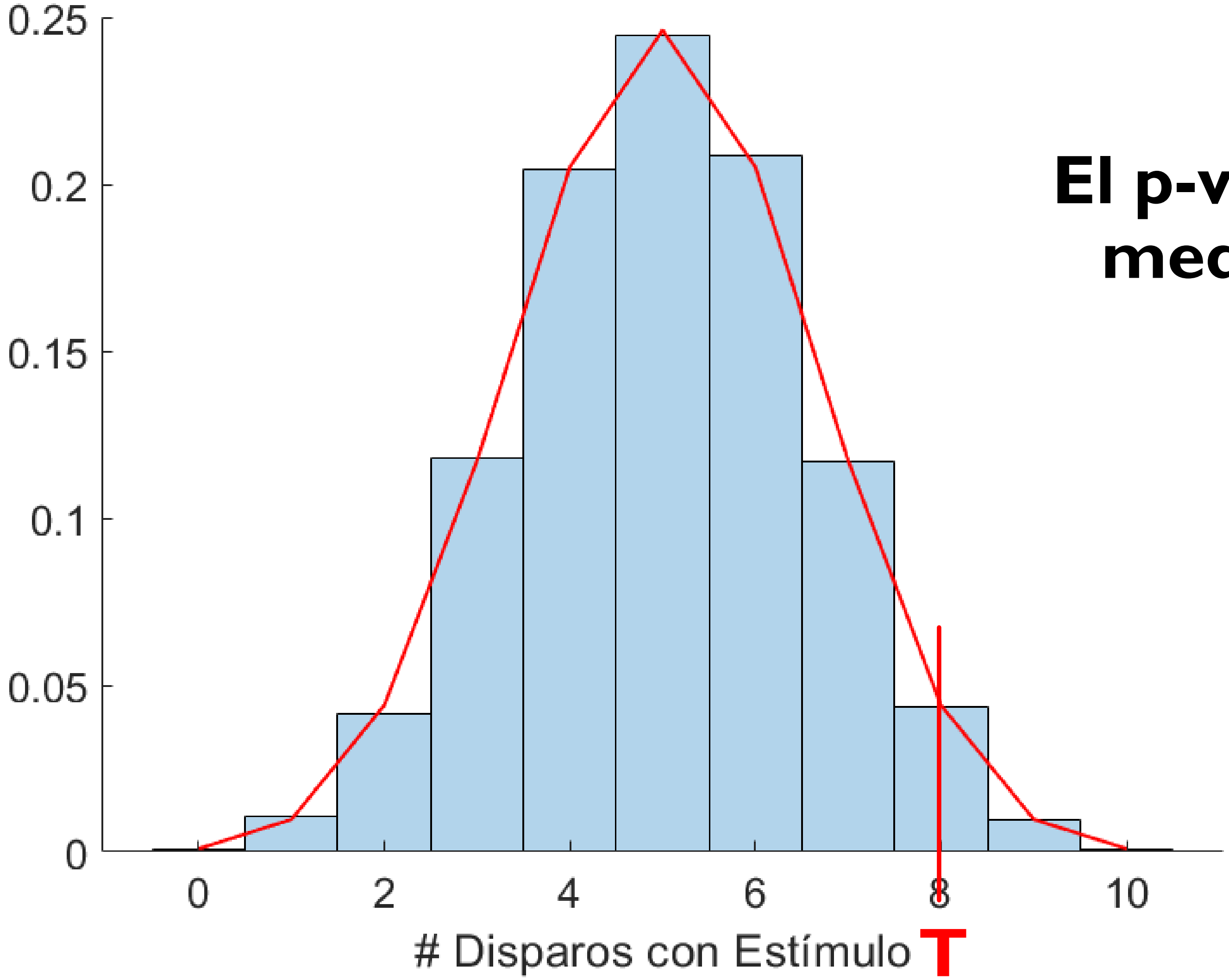
Si H_0 es verdadera (p=0.5)



Y ahora sí…

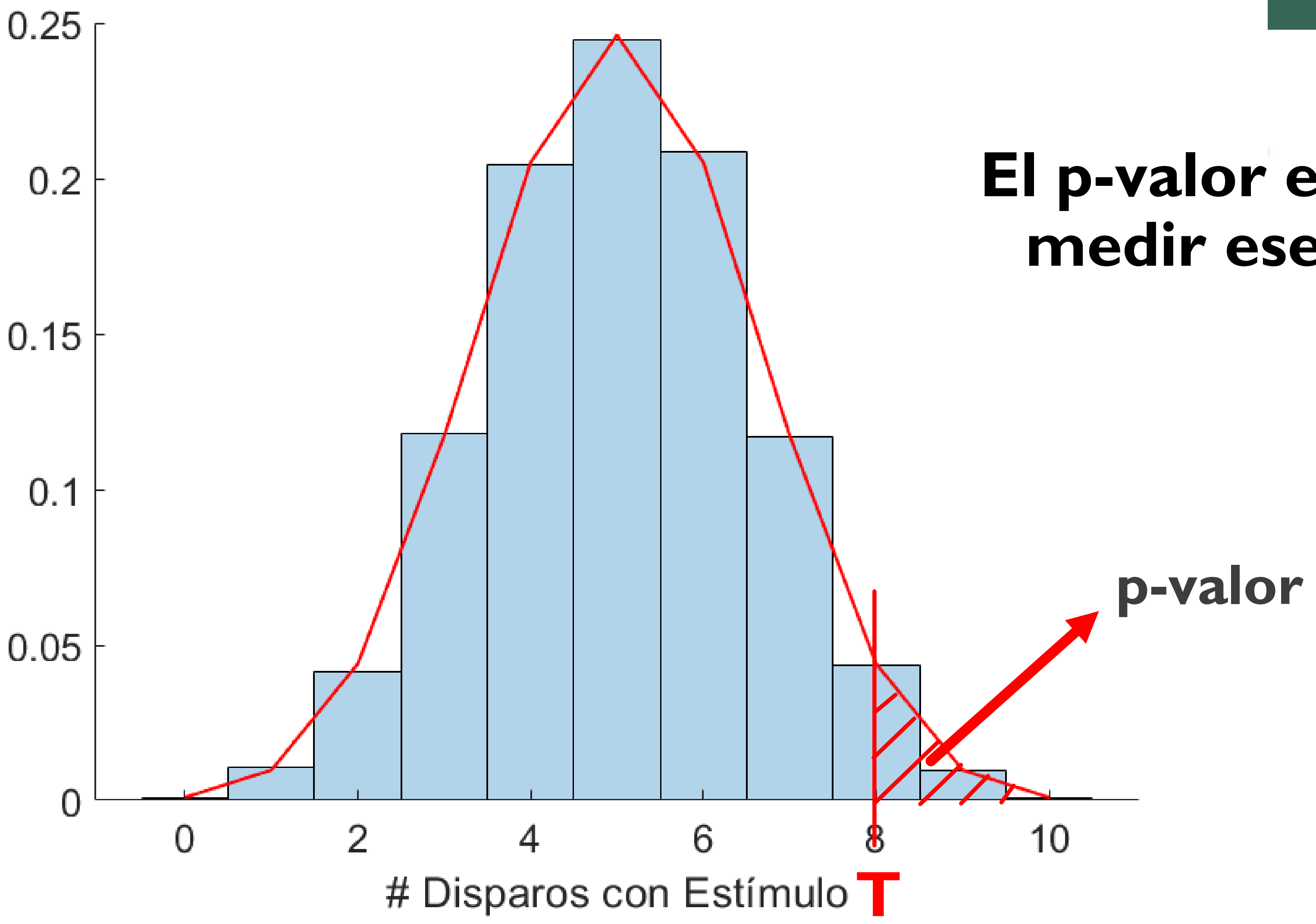
El p-valor es la probabilidad de medir ese valor, o uno "más raro".

¿QUÉ ES EL P-VALOR?



El p-valor es la probabilidad de medir ese valor, o uno "más raro".

¿QUÉ ES EL P-VALOR?



El p-valor es la probabilidad de medir ese valor, o uno "más raro".

¿QUÉ ES EL P-VALOR?

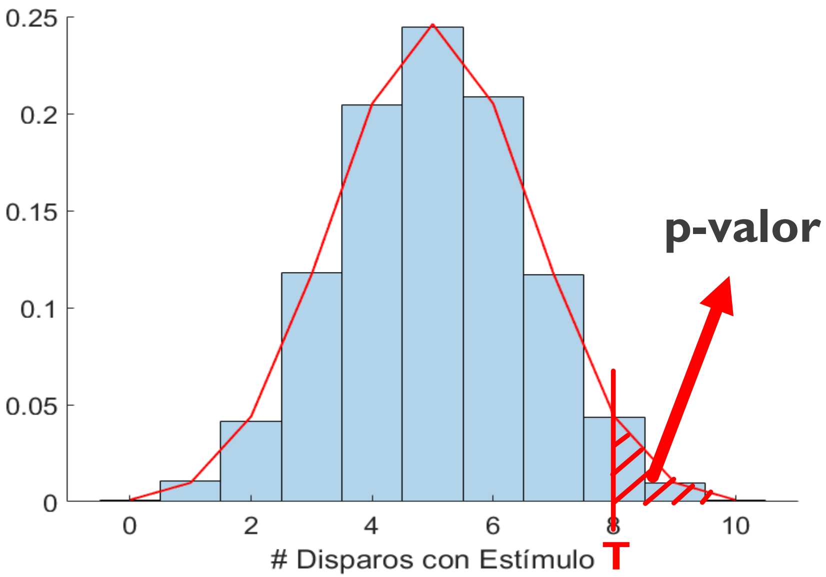
$$\mathbf{T} = \frac{\text{\# total de disparos con el estímulo}}{\text{\# total de presentaciones del estímulo}} = \frac{8}{10}$$

Y ahora sí…

El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Acá, el p-valor es: **p-val = p(8)+p(9)+p(10)= 0.0547**

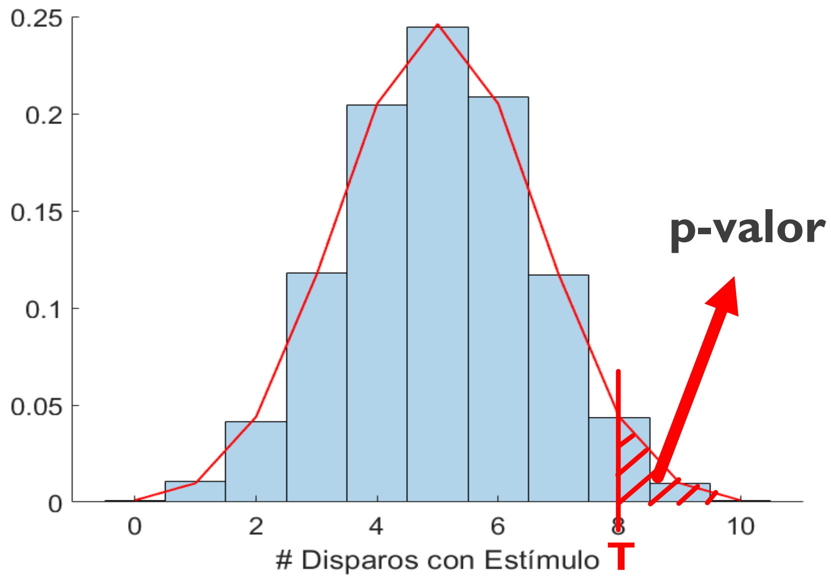
Si H_0 es verdadera (p=0.5)



¿QUÉ ES EL P-VALOR?

$$T = \frac{\text{\# total de disparos con el estímulo}}{\text{\# total de presentaciones del estímulo}} = \frac{8}{10}$$

Si H_0 es verdadera (p=0.5)



Y ahora sí…

El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Acá, el p-valor es: **p-val = p(8)+p(9)+p(10)= 0.0547**

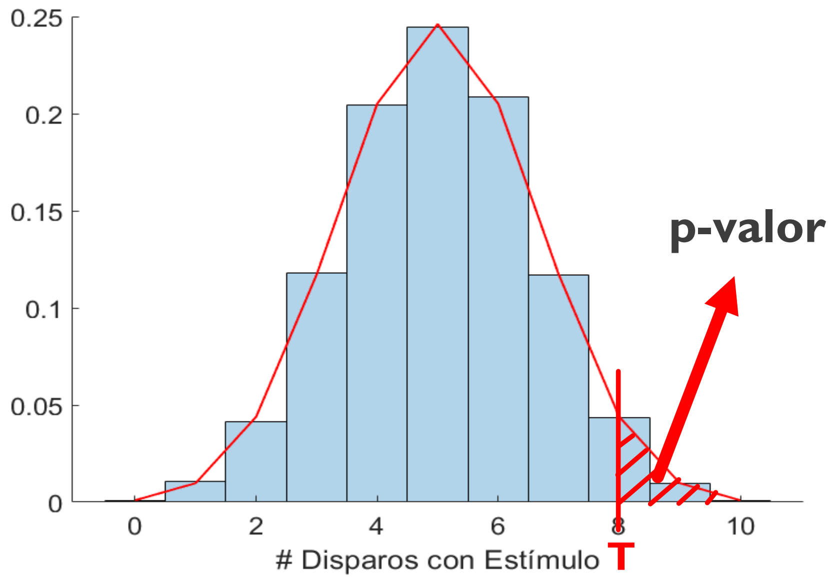
Entonces…

¿y qué?

¿QUÉ ES EL P-VALOR?

$$T = \frac{\#\ \text{total de disparos con el estímulo}}{\#\ \text{total de presentaciones del estímulo}} = \frac{8}{10}$$

Si H_0 es verdadera (p=0.5)



Y ahora sí…

El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Acá, el p-valor es: **p-val = p(8)+p(9)+p(10)= 0.0547**

Entonces…

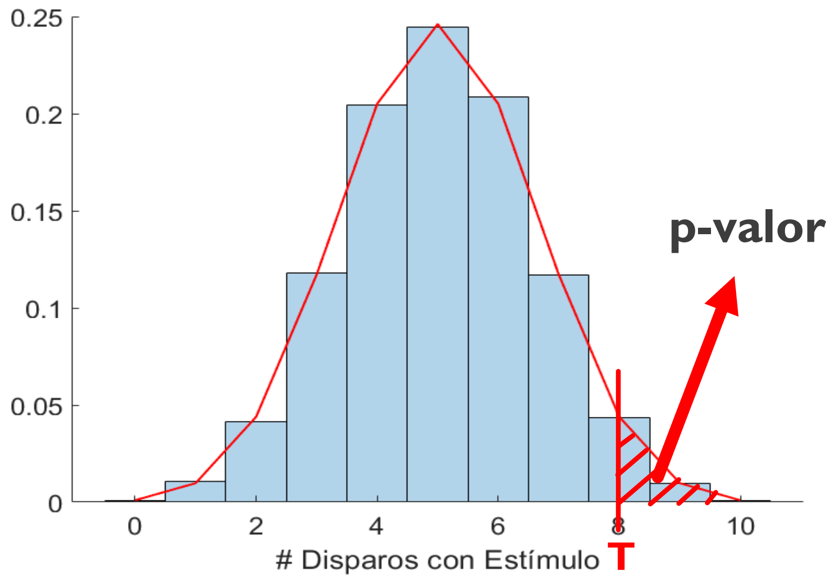
¿y qué?

Elegimos arbitrariamente un punto de corte:
α, el "**nivel de significancia**"

¿QUÉ ES EL P-VALOR?

$$T = \frac{\text{\# total de disparos con el estímulo}}{\text{\# total de presentaciones del estímulo}} = \frac{8}{10}$$

Si H_0 es verdadera (p=0.5)



Y ahora sí…

El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Acá, el p-valor es: **p-val = p(8)+p(9)+p(10)= 0.0547**

Entonces…

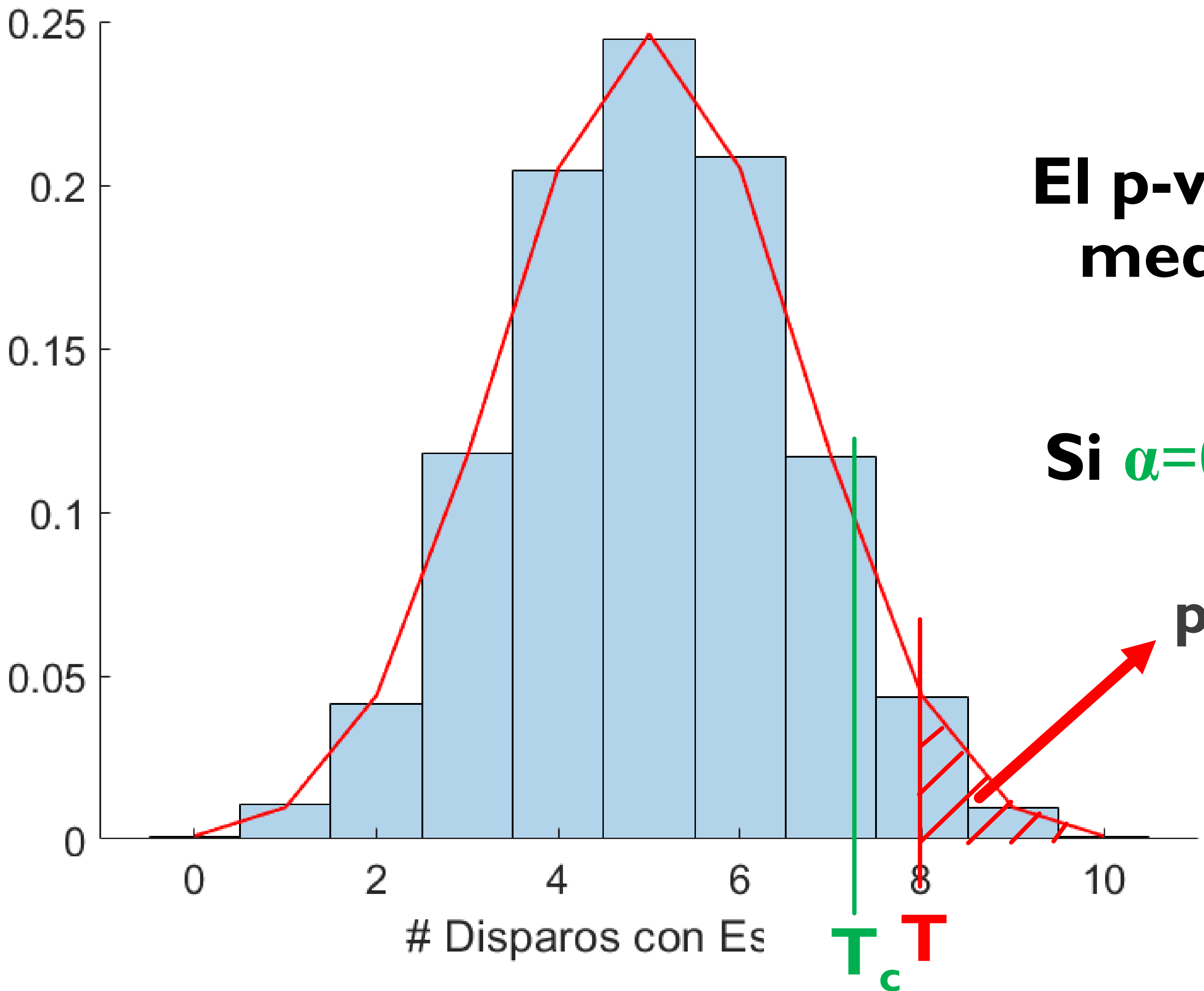
¿y qué?

Elegimos arbitrariamente un punto de corte: α, el "**nivel de significancia**"

Si **p< α**, decimos que es "demasiado raro" para venir de H_0 →

Rechazamos H_0

¿QUÉ ES EL P-VALOR?

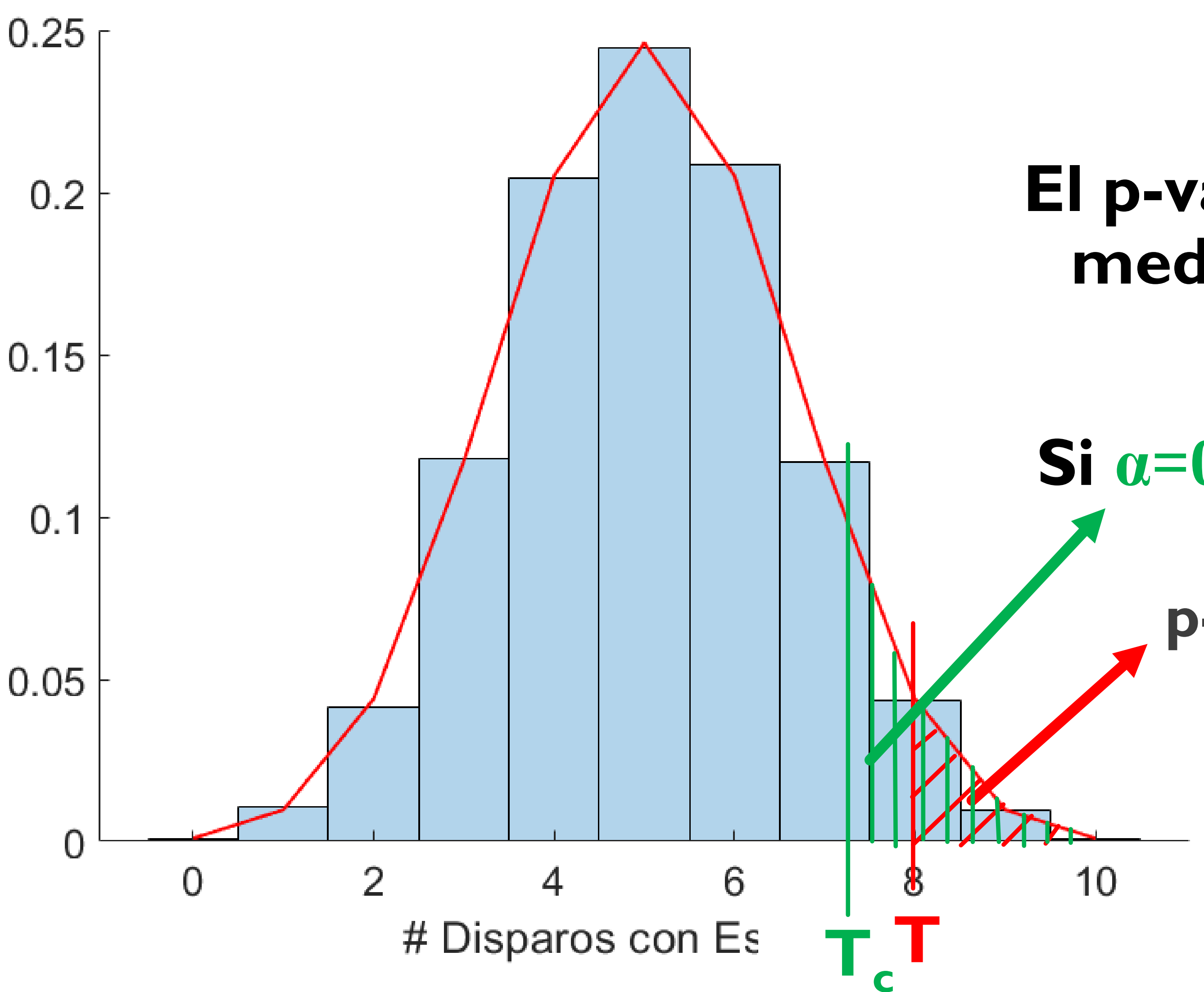


El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Si $\alpha=0.05$…

p-valor = 0.0547

¿QUÉ ES EL P-VALOR?

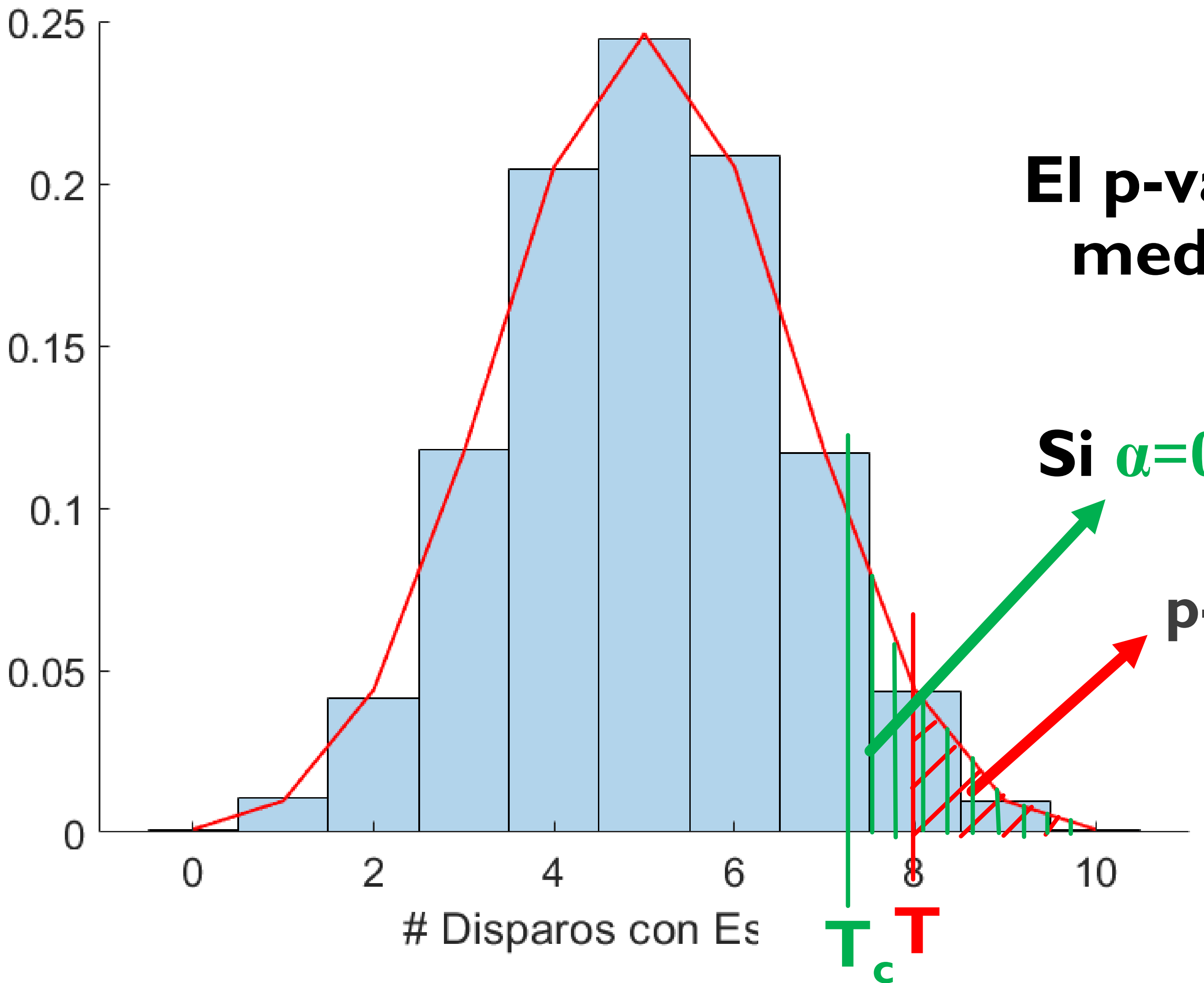


El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Si $\alpha=0.05$…

p-valor = 0.0547

¿QUÉ ES EL P-VALOR?



El p-valor es la probabilidad de medir ese valor, o uno "más raro".

Si α**=0.05…**

p-valor = 0.0547

$p>\alpha=0.05$ →

NO Rechazamos H_0

VARIABLES ALEATORIAS

Cuando realizamos un experimento para medir una variable (por ejemplo, temperatura), observamos diferentes valores, aún midiendo en igualdad de condiciones. Por distintos motivos, las mediciones siguen una distribución, y así, constituyen una **variable aleatoria**.

VARIABLES ALEATORIAS

Cuando realizamos un experimento para medir una variable (por ejemplo, temperatura), observamos diferentes valores, aún midiendo en igualdad de condiciones. Por distintos motivos, las mediciones siguen una distribución, y así, constituyen una **variable aleatoria**.

Cualquier variable que dependa de las mediciones, al depender de una variable aleatoria, son a su vez, variables aleatorias. Por ejemplo, el **p-valor** depende de la distribución del **estadístico** (dada H_0), que depende de las **mediciones**.

VARIABLES ALEATORIAS

Cuando realizamos un experimento para medir una variable (por ejemplo, temperatura), observamos diferentes valores, aún midiendo en igualdad de condiciones. Por distintos motivos, las mediciones siguen una distribución, y así, constituyen una **variable aleatoria**.

Cualquier variable que dependa de las mediciones, al depender de una variable aleatoria, son a su vez, variables aleatorias. Por ejemplo, el **p-valor** depende de la distribución del **estadístico** (dada H_0), que depende de las **mediciones**.

El p-valor es una variable aleatoria (y si H_0 es verdadera, sigue una **distribución uniforme;** si repetimos un mismo experimento muchas veces, y H_0 es V, veremos p-valores distribuidos uniforme-mente entre 0 y 1).

VARIABLES ALEATORIAS

Cuando realizamos un experimento para medir una variable (por ejemplo, temperatura), observamos diferentes valores, aún midiendo en igualdad de condiciones. Por distintos motivos, las mediciones siguen una distribución, y así, constituyen una **variable aleatoria**.

Cualquier variable que dependa de las mediciones, al depender de una variable aleatoria, son a su vez, variables aleatorias. Por ejemplo, el **p-valor** depende de la distribución del **estadístico** (dada H_0), que depende de las **mediciones**.

El p-valor es una variable aleatoria (y si H_0 es verdadera, sigue una **distribución uniforme;** si repetimos un mismo experimento muchas veces, y H_0 es V, veremos p-valores distribuidos uniformemente entre 0 y 1).

Cuando estimamos **la media** o **el desvío** de un conjunto de mediciones (variables aleatorias), esos estimativos **son también variables aleatorias** (y van a cambiar si repetimos el experimento más veces). Y si usamos esa media y desvío para armar otra variable, también será una variable aleatoria.

INTERVALOS DE CONFIANZA

Un **intervalo de confianza** es un rango elaborado con una **receta** específica, tal que, si se repite el mismo experimento y se utiliza esa misma receta, una X cantidad de veces el intervalo contendrá al valor real. Ese X se llama "**nivel de confianza**". Por ejemplo, si X=95, entonces 95 de cada 100 intervalos de confianza armados de esa forma contendrán al valor real, y 5 de ellos no. Los límites de los intervalos de confianza son **variables aleatorias**.

INTERVALOS DE CONFIANZA

Un **intervalo de confianza** es un rango elaborado con una **receta** específica, tal que, si se repite el mismo experimento y se utiliza esa misma receta, una X cantidad de veces el intervalo contendrá al valor real. Ese X se llama "**nivel de confianza**". Por ejemplo, si X=95, entonces 95 de cada 100 intervalos de confianza armados de esa forma contendrán al valor real, y 5 de ellos no. Los límites de los intervalos de confianza son **variables aleatorias**.

Un intervalo de confianza de nivel de confianza X **NO** es un intervalo que tiene X% de probabilidad de contener al valor real. Esta definición es FALSA porque considera al intervalo como algo fijo. Pero en realidad, se obtiene un intervalo de confianza DISTINTO cada vez que se repite el experimento. Ese intervalo puede NO contener al valor real (e incluso puede no tener sentido). La receta para armar intervalos de confianza correctamente es bastante compleja, pero si el N no es muy chico (por ejemplo, mayor a 10), se puede usar la simplificación:

INTERVALOS DE CONFIANZA

Un **intervalo de confianza** es un rango elaborado con una **receta** específica, tal que, si se repite el mismo experimento y se utiliza esa misma receta, una X cantidad de veces el intervalo contendrá al valor real. Ese X se llama "**nivel de confianza**". Por ejemplo, si X=95, entonces 95 de cada 100 intervalos de confianza armados de esa forma contendrán al valor real, y 5 de ellos no. Los límites de los intervalos de confianza son **variables aleatorias**.

Un intervalo de confianza de nivel de confianza X **NO** es un intervalo que tiene X% de probabilidad de contener al valor real. Esta definición es FALSA porque considera al intervalo como algo fijo. Pero en realidad, se obtiene un intervalo de confianza DISTINTO cada vez que se repite el experimento. Ese intervalo puede NO contener al valor real (e incluso puede no tener sentido). La receta para armar intervalos de confianza correctamente es bastante compleja, pero si el N no es muy chico (por ejemplo, mayor a 10), se puede usar la simplificación:

X=66%: $[\mu-\sigma, \mu+\sigma]$, X=95%: $[\mu-2\sigma, \mu+2\sigma]$

INTERVALOS DE CONFIANZA Y P-VALORES

A priori, los intervalos de confianza y los p-valores **no tienen nada que ver**. Tampoco tienen nada que ver el nivel de significancia y el nivel de confianza. Son conceptos totalmente diferentes. Sin embargo…

INTERVALOS DE CONFIANZA Y P-VALORES

A priori, los intervalos de confianza y los p-valores **no tienen nada que ver**. Tampoco tienen nada que ver el nivel de significancia y el nivel de confianza. Son conceptos totalmente diferentes. Sin embargo…

Cuando realizamos un T-test para diferenciar entre dos muestras independientes, lo que se hace es evaluar si las restas entre los datos de las dos muestras dan una distribución con media 0 (como se esperaría si H_0 es correcta), o no. Al hacer esto, es posible armar un intervalo de confianza para el estadístico T, y ver si este intervalo contiene al 0 o no, con un dado nivel de confianza. Si uno en cambio busca el p-valor del estadístico T obtenido, uno puede rechazar H_0 con un dado nivel de significancia. Sólo en este caso particular, el nivel de confianza de ese intervalo puede coincidir con el nivel de significancia de ese estadístico T. Esto **NO** significa que conceptualmente sean lo mismo.

INTERVALOS DE CONFIANZA Y P-VALORES

A priori, los intervalos de confianza y los p-valores **no tienen nada que ver**. Tampoco tienen nada que ver el nivel de significancia y el nivel de confianza. Son conceptos totalmente diferentes. Sin embargo…

Cuando realizamos un T-test para diferenciar entre dos muestras independientes, lo que se hace es evaluar si las restas entre los datos de las dos muestras dan una distribución con media 0 (como se esperaría si H_0 es correcta), o no. Al hacer esto, es posible armar un intervalo de confianza para el estadístico T, y ver si este intervalo contiene al 0 o no, con un dado nivel de confianza. Si uno en cambio busca el p-valor del estadístico T obtenido, uno puede rechazar H_0 con un dado nivel de significancia. Sólo en este caso particular, el nivel de confianza de ese intervalo puede coincidir con el nivel de significancia de ese estadístico T. Esto **NO** significa que conceptualmente sean lo mismo.

Cuando hablamos de p-valor, hablamos de nivel de significancia. Cuando hablamos de intervalo de confianza, hablamos de nivel de confianza. No conviene mezclar ambos conceptos, en ningún caso.