

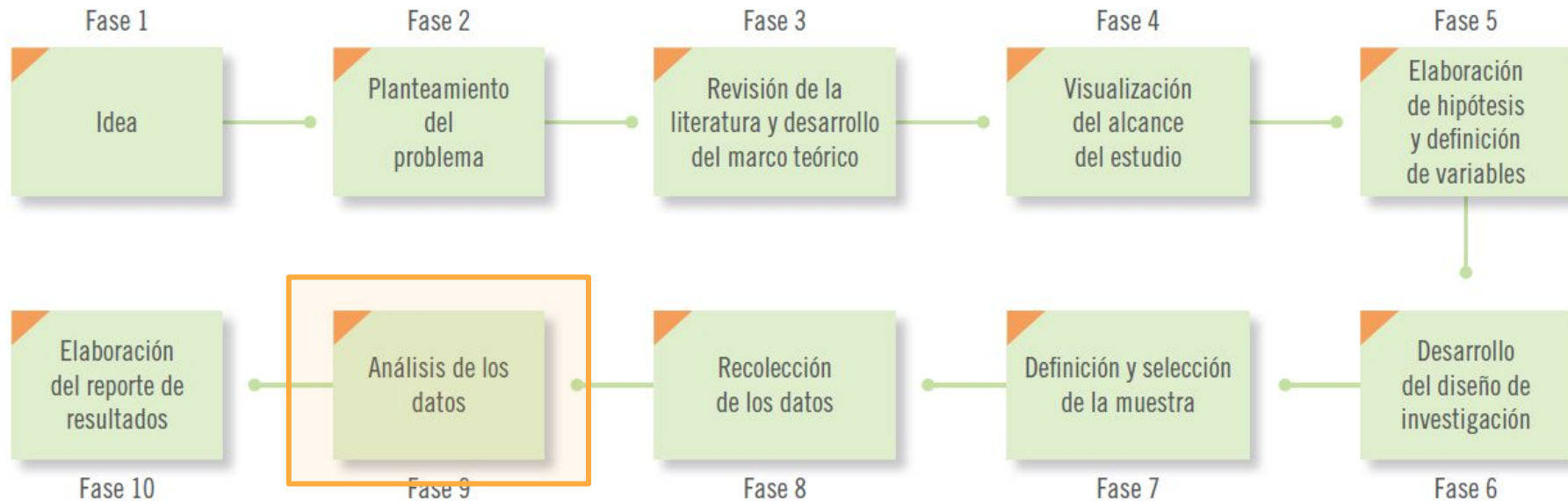


Investigación cuantitativa



Proceso cuantitativo

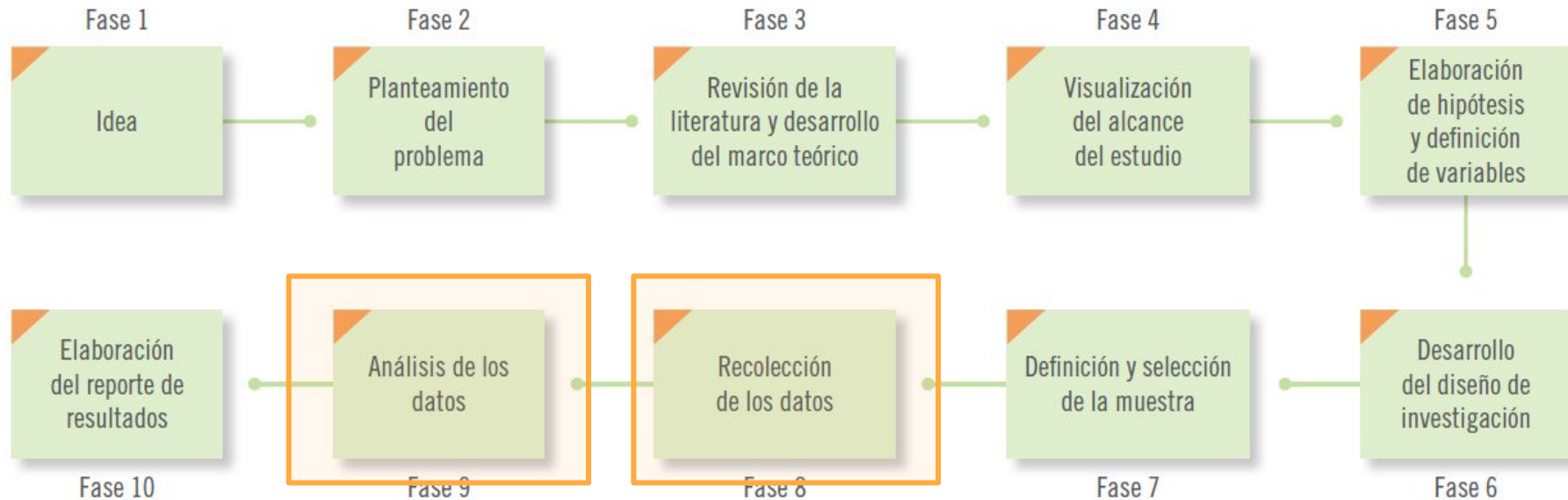
Lo visto tiene relación con solo una de las fases de la **investigación cuantitativa**.





Proceso cuantitativo

Hablemos un poco de la fase previa.





Recolección de datos

En el contexto de la investigación cuantitativa, una **medición** es el proceso que vincula conceptos abstractos con indicadores, utilizando un recurso que registra **datos observables** que **representan conceptos o variables** que se quieren medir al que se le llama **instrumento de medición**.

Un instrumento de medición debe cumplir tres requisitos:

1. **validez**
2. **confiabilidad**
3. **objetividad**

--- Hernández, Fernández y Baptista (2014)



Instrumentos de medición

La **validez** se refiere al grado en que un instrumento de medición mide realmente las variables que pretende medir.

Se puede obtener **tres tipos de evidencia** para la validez cuantitativa:

1. de la **validez de contenido**, mostrando que los ítems presentes en el instrumento cubren las principales dimensiones de la variable estudiada
2. de la **validez de criterio**, mostrando que los resultados de la aplicación del instrumento concuerdan con los resultados esperados (de un criterio externo)
3. de la **validez de constructo**, mostrando que las mediciones están vinculadas a las mediciones de conceptos correlacionados teóricamente



Instrumentos de medición

La **confiabilidad** se refiere al grado en que la aplicación repetida de un instrumento de medición, en los mismos casos, produce resultados iguales.

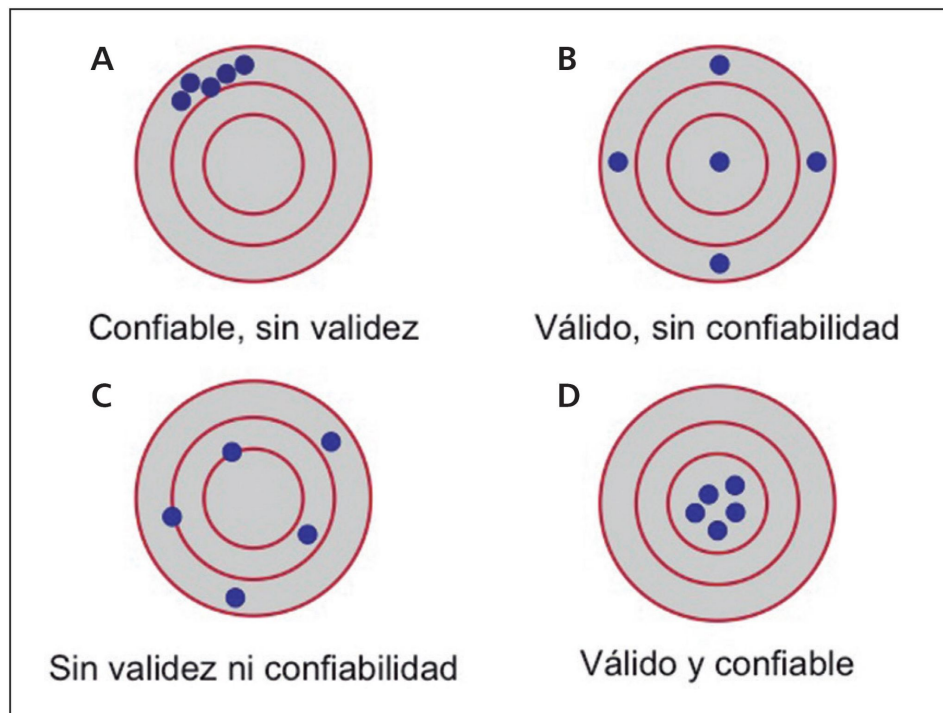
La confiabilidad cuantitativa se determina al calcular el **coeficiente de fiabilidad**.

Los coeficientes de fiabilidad cuantitativa varían entre 0 (nula confiabilidad) y 1 (total confiabilidad).

Los métodos más conocidos para calcular la confiabilidad son: medida de estabilidad (**test-retest**), formas alternativas o **paralelas**, mitades **partidas** y **consistencia interna**.

Instrumentos de medición

Gráficamente:



--- Manterola et al. (2018)
Revista chilena de infectología



Instrumentos de medición

Los factores que pueden afectar la confiabilidad y la validez son:

1. la **improvisación**, pensar que elegir un instrumento o desarrollarlo es algo que puede tomarse a la ligera (e.g. encuestas de usabilidad)
2. utilizar instrumentos desarrollados en el extranjero y que no han sido validados en el **contexto local**
3. poca o nula **empatía** con los participantes
4. **condiciones** en las que se aplica el instrumento de medición

No hay medición perfecta, pero el error de medición debe reducirse a límites tolerables.



Instrumentos de medición

La **objetividad** se refiere al grado en que el instrumento es o no permeable a la influencia de los sesgos y tendencias del investigador o investigadores que lo administran, califican e interpretan.

La objetividad es **difícil de lograr**, particularmente en las ciencias sociales. Se refuerza:

1. **estandarizando** la aplicación del instrumento (mismas instrucciones y condiciones para todos los participantes) y la evaluación de los resultados
2. empleando **personal capacitado** y experimentado en el instrumento



Instrumentos de medición

Los estudios cuantitativos buscan que la influencia de las características y las tendencias del investigador se reduzca al mínimo posible.

Pero esto es un ideal, pues la investigación siempre es realizada por seres humanos.

La validez, la confiabilidad y la objetividad **no** deben tratarse de forma separada. Sin alguna de las tres, el instrumento **no es útil** para llevar a cabo un estudio.



Instrumentos de medición

Los primeros pasos para elaborar un instrumento de medición son:

1. Redefiniciones fundamentales sobre propósitos, definiciones operacionales y participantes
2. Revisar la literatura, particularmente la enfocada en los instrumentos utilizados para medir las variables de interés
3. Identificar el conjunto o dominio de conceptos o variables a medir e indicadores de cada variable
4. Tomar decisiones en cuanto a: tipo y formato; utilizar uno existente, adaptarlo o construir uno nuevo, así como el contexto de aplicación
5. Construir la primera versión del instrumento



Instrumentos de medición

Los siguientes pasos para elaborar un instrumento de medición son:

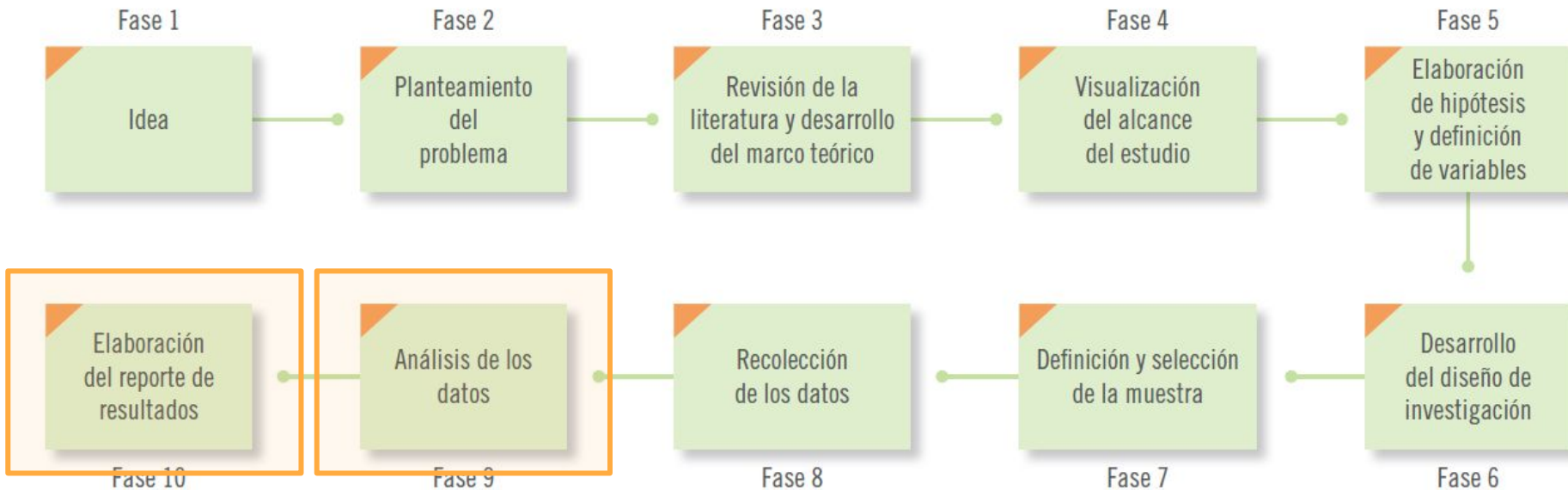
6. Aplicar la prueba piloto (para calcular la confiabilidad y validez iniciales)
7. Desarrollar la versión definitiva del instrumento
8. Entrenar al personal que va a aplicarlo
9. Obtener autorizaciones para aplicarlo
10. Aplicar el instrumento

Luego se debe preparar los datos para el análisis estadístico: codificarlos, limpiarlos, insertarlos en una matriz o base de datos.



Proceso cuantitativo

Hablemos brevemente ahora de la fase que sigue.





Reportes

La etapa final es el reporte. Existen varios tipos:

- tesis, ensayos, libros
- informes de investigación, resumen ejecutivo
- reportes técnicos, informe de estudio de casos
- ponencias o artículos
- presentaciones audiovisuales, artículos de periodísticos o divulgación científica

La elección del tipo de reporte depende de:

- la motivación de la investigación
- los usuarios del estudio
- el contexto donde se presenta



Reporte

En todos los tipos de reportes es importante **mostrar y discutir los resultados** del análisis.

Existen estándares para apoyarse.

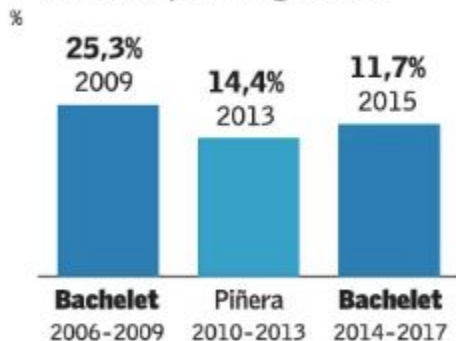
Se destaca el estándar APA de cómo reportar las pruebas realizadas. Detalles se “compran”, pero hay algunos resúmenes disponibles en la web. Por ejemplo:

- <https://www.apa.org/pubs/authors/jars.pdf>
- <http://users.sussex.ac.uk/~grahamh/RM1web/APA%20format%20for%20statistical%20notation%20and%20other%20things.pdf>
- http://evc-cit.info/psych018/Reporting_Statistics.pdf

Evitar

Impresiones “equivocadas” (o errores) en los gráficos.

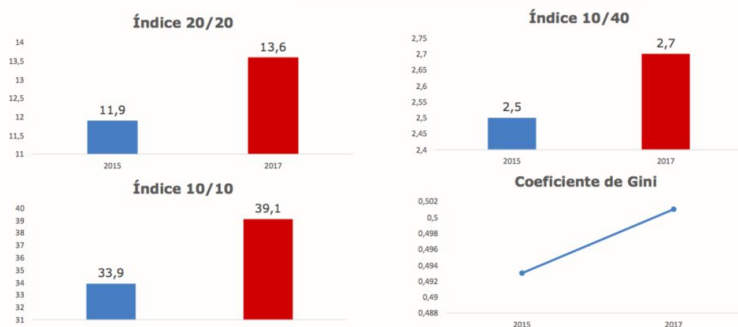
Pobreza por ingresos



Fuente Casen



Indicadores de la distribución del ingreso del trabajo de los hogares (2006-2017) (Porcentaje)



* Se excluye servicio doméstico puertas adentro y su núcleo familiar.

Fuente: Ministerio de Desarrollo Social, Encuesta Casen 2006-2017.

www.desarrollosocial.cl

Ministerio de Desarrollo Social

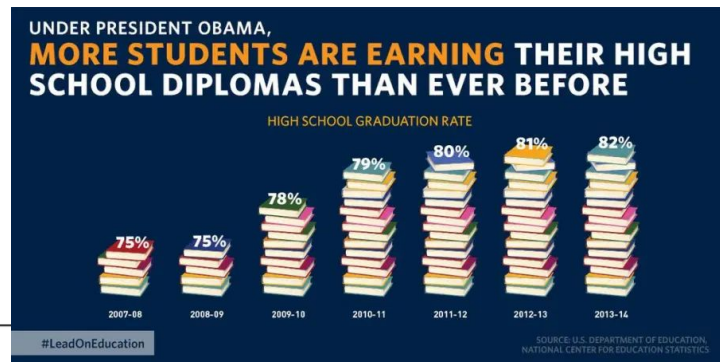
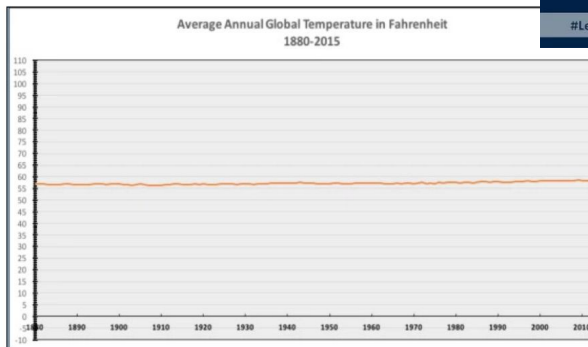
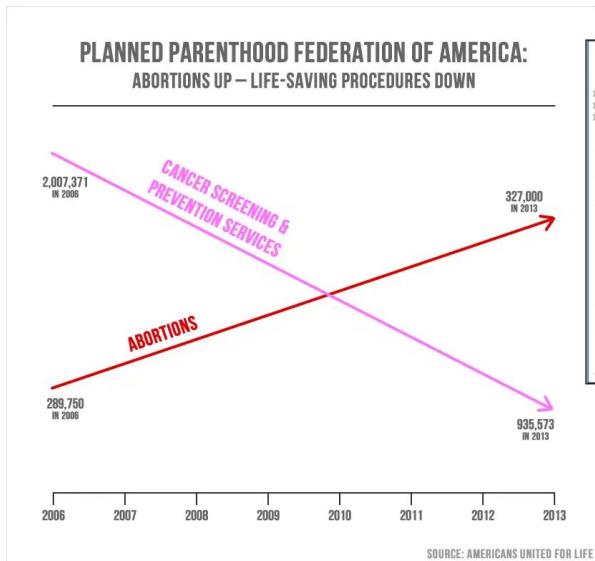
Observatorio Social

Casen
2017



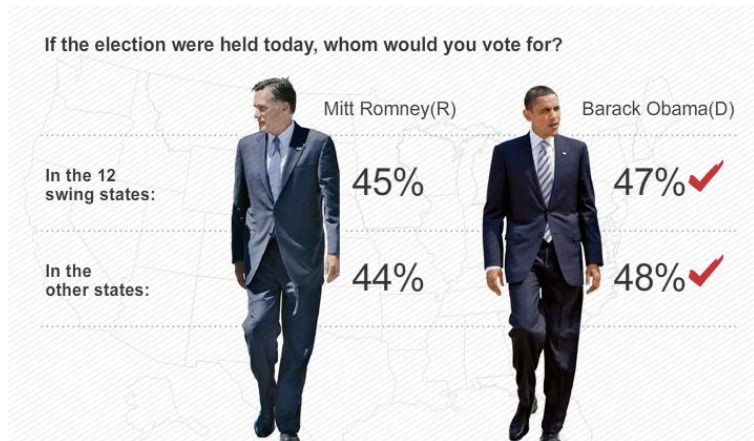
Evitar

Impresiones “equivocadas”.





Declaraciones sin sustento estadístico.



EN VIVO

PAÍS

Piñera: "Supe después que baja en pobreza era margen de error de encuesta Casen"

El Presidente se refirió así a los resultados del estudio, asegurando que "hemos tenido una reducción significativa de la pobreza extrema, que me llena de orgullo, y una reducción menos significativa en la pobreza".

Por El Dínamo, 20 de Septiembre de 2012





¿y en el futuro?



Regla del $n > 30$

Como hemos visto, esta regla tiene varias suposiciones:

- La población es infinita (o muy grande)
- Los datos fueron obtenidos con muestreo aleatorio simple
- La población tiene una distribución “razonable”

Sin embargo, estas **condiciones suelen olvidarse**, al igual que para muchas otras reglas y métodos clásicos.

Peor, debemos recordar que los métodos clásicos de inferencia se basan en **resultados asintóticos**.

--- Hesterberg (2008) Proc. of the American Statistical Association



Regla del $n > 30$

Como el teorema del límite central y la ley de los grandes números.

- la distribución de un estimador **tiende** a una distribución normal (o t) a medida que el tamaño de la muestra **crece a infinito**...

Estas **aproximaciones teóricas** fueron desarrolladas por Ronald Fisher, Jerzy Neyman y Egon Pearson entre 1920 y 1935 aproximadamente.

Mucho antes de que existieran los computadores.

¿Se justifica seguir con esta regla?



Regla del $n > 30$

Muchas personas piensan que no.

Por ejemplo, Hesterberg (2008) mostró que:

- existen alternativas basadas en **remuestreo** que hoy son viables y fáciles de utilizar
- **diagnósticos** con remuestreo son más efectivos que la regla clásica
- se logran **inferencias más precisas**, aunque a veces se necesitan muchas repeticiones (>60.000) y mejores métodos para muestras muy pequeñas

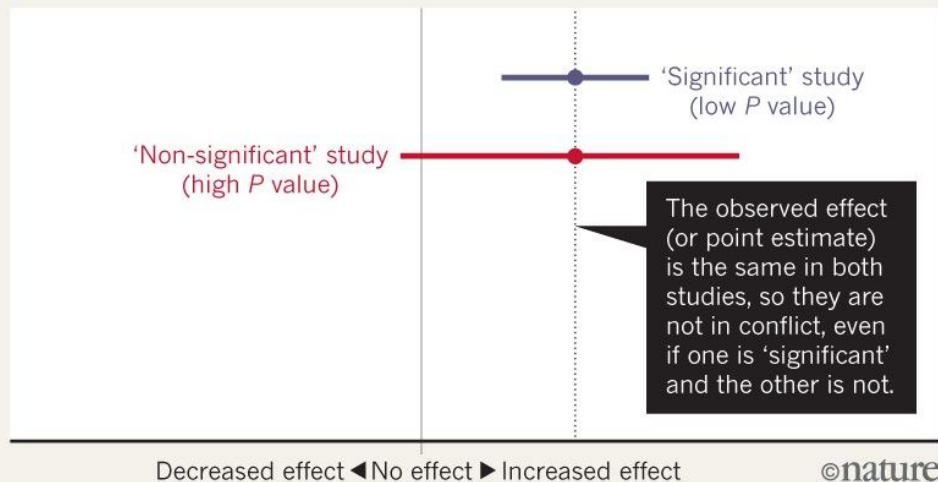


Críticas al p-valor

Centrarse en p-valores ha sido y todavía es **controversial**.

BEWARE FALSE CONCLUSIONS

Studies currently dubbed 'statistically significant' and 'statistically non-significant' need not be contradictory, and such designations might cause genuine effects to be dismissed.



“After 4 decades of severe criticism, the ritual of null hypothesis significance testing—mechanical dichotomous decisions around a sacred .05 criterion—still persists.”

--- Cohen, J. (1994). The Earth is round ($p < .05$). American psychologist, 49(12).

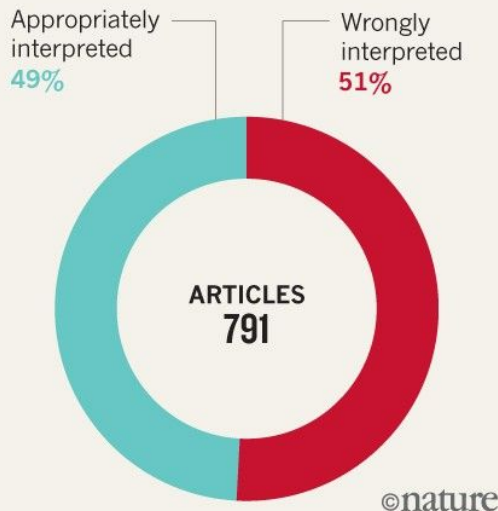
Críticas al p-valor

Nunca se debe concluir que “**no hay diferencia** o **asociación**” solamente porque un p-valor es menor que un umbral como 0,05.

WRONG INTERPRETATIONS

An analysis of 791 articles across 5 journals* found that around half mistakenly assume non-significance means no effect.

*Data taken from: P. Schatz *et al. Arch. Clin. Neuropsychol.* **20**, 1053–1059 (2005); F. Fidler *et al. Conserv. Biol.* **20**, 1539–1544 (2006); R. Hoekstra *et al. Psychon. Bull. Rev.* **13**, 1033–1037 (2006); F. Bernardi *et al. Eur. Sociol. Rev.* **33**, 1–15 (2017).



Sin embargo, se han encontrado **cientos de artículos científicos** cometiendo estos errores de interpretación en diferentes estudios y en **diferentes áreas** del saber.

--- Amrhein, Greenland & McShane (2019) NATURE



Críticas al p-valor

Esto ha llevado a **desperdiciar** muchos recursos innecesariamente.

Se recomiendan las **3R** (¡no el pisco!):

1. **Randomize** data production (muestras aleatorias, grupos aleatorios)
2. **Repeat** by simulation (hasta ver qué resultados son típicos y cuáles no)
3. **Reject** any model that puts your data in its tail (contradice los datos)

--- Cobb (2007) Technology innovations in statistics education



Críticas a la hipótesis nula

Pero esta interpretación errada tiene raíz en el **uso de hipótesis nulas**.

De acuerdo a Cohen (1994), las pruebas de hipótesis **no nos dicen** lo que queremos saber:

- **queremos**: dados estos datos, ¿cuál es la probabilidad que H_0 es verdadera?
- **nos dice**: si H_0 es verdadera, ¿con qué probabilidad veríamos estos datos o datos más extremos?

Y sabemos desde hace mucho que $P(D|H_0) \neq P(H_0|D)$.



Críticas a la hipótesis nula

Veamos un ejemplo:

La incidencia de esquizofrenia en adultos es $\approx 2\%$. Se estima que un test puede detectar $\approx 95\%$ de los casos positivos (sensibilidad) y $\approx 97\%$ de los negativos (especificidad).

Así: H_0 : el caso es normal, H_1 : el caso es esquizofrenia, T^+ : el test resulta positivo para esquizofrenia (dato).

Luego: $P(T^+|H_0) < .05$, por lo que se puede rechazar H_0 en favor de H_1 (con 95% confianza).

Pero, aunque a muchos le parezca así o quieran que sea así, esto **no nos dice** cuál es la probabilidad de observar un caso normal a pesar de tener un examen positivo: $P(H_0|T^+)$.

--- Cohen (1994)



Críticas a la hipótesis nula

Veamos un ejemplo:

Usando **estadística Bayesiana**, se puede estimar que $P(H_0|T^+) = \approx .607$! Lejos del p-valor obtenido.

Puede que sea más fácil ver esta probabilidad en una tabla con 1.000 datos:

	Normal	Esquiz.	Total
T^-	949	1	950
T^+	30	20	50
Total	979	21	1.000

¡30 falsos positivos en 50 tests!

Esto ocurre porque los casos de esquizofrenia son poco frecuentes:

$$P(H_0) = .98$$

--- Cohen (1994)



Críticas a la hipótesis nula

También se ha criticado porque muchas veces no tiene sentido.

Una verdadera H_0 : no hay diferencia, igual probabilidad, no hay correlación;
siempre es falsa.

“It is foolish to ask ‘Are the effects of A and B different?’ They are always different for some decimal place” (Tukey, 1991)

--- Tukey (1991) Statistical Science



Críticas más antiguas

Desde Berkson (1938) que hay personas que cuestionan el uso de la prueba de hipótesis en la ciencia...

Some Difficulties of Interpretation Encountered in the Application of the Chi-Square Test

Joseph Berkson

Journal of the American Statistical Association, Vol. 33, No. 203. (Sep., 1938), pp. 526-536.

Journal of the American Statistical Association is currently published by American Statistical Association.





Críticas más antiguas

La principal tiene que ver con la idea de **probabilidad \approx frecuencia**.

La teoría se formuló con monedas, dados, cartas y bolas de colores... que pueden generar **variables aleatorias** asociadas a **distribuciones de probabilidad**.

Luego, como vimos con la interpretación de los intervalos de confianza:

Si un parámetro tiene un **valor fijo**,
¿por qué se trata como una variable aleatoria?



¿Alternativas?

Varias, la más popular es la **estadística Bayesiana**.

Las 3R como alternativa a medio camino.

Pero no vamos a entrar en eso este semestre.

Última reflexión:

Question: Why, then, is the t-test the centerpiece of the introductory statistics curriculum? Answer: The t-test is what scientists and social scientists use most often.

Question: Why does everyone use the t-test? Answer: Because it's the centerpiece of the introductory statistics curriculum.

--- Cobb (2007)



¿Dudas?
¿Comentarios?