

Tests d'hipòtesi - Diferències entre Fisher i Neyman-Pearson

Pau Soler Valadés

November 2024

Encara que Neyman-Pearson parli d'una hipòtesi alternativa, la intenció tant de Fisher com de Neyman-Pearson era la mateixa: comprovar si la hipòtesi nul·la (o principal) és correcta. La diferència fonamental entre els dos és que Neyman-Pearson t'obliga a explicitar la hipòtesi alternativa, mentre que Fisher només et permet rebutjar o no la hipòtesi nul·la segons els resultats dels tests. Aleshores, l'elecció de la hipòtesi alternativa queda lligada a un esdeveniment plausible dins de les modificacions de la hipòtesi nul·la, és a dir, no pots escollir qualsevol cosa com a hipòtesi alternativa, sinó només la modificació de la hipòtesi nul·la corresponent. La filosofia d'aquesta elecció es pot resumir amb “*significatiu vs no significatiu*” o “ H_A (el canvi és significatiu) vs H_M (el canvi no és significatiu)”, que en essència és el mateix que fa Fisher (significatiu és rebutjo H_0 i no significatiu és no rebutjo H_0), però sense explicitar la hipòtesi alternativa.

Hipòtesi Plausible

H_0 Fisher	H_P Neyman-Pearson
No s'explicita en general	Concebuda a priori, a l'etapa de disseny
No té res a veure amb la mida de l'efecte	Conté el MSE (<i>minimum size effect</i>) a la formulació
Es busca rebutjar amb proves	S'escull entre H_P i entre H_A

Regió Crítica

Fisher	Neyman-Pearson
Depen del nivell de significació determinat a posteriori, i s'aprova amb el <i>p-value</i>	S'ha de trobar el valor crític α per poder delimitar la regió crítica.
Nivell de significància és variable a posteriori, canviant l'anàlisi del test.	A priori: immòbil i rígida
Es pot canviar	Non-gradable: només α

Rebutjar/No Rebutjar

Fisher	Neyman-Pearson
--------	----------------

O rebutjes H_0 o no rebutjes H_0 per falta de proves.	O acceptes H_P o acceptes H_A
Test de força (<i>significance test</i>): mesurem l'evidència per rebutjar H_0	Test d'acceptació (<i>acceptance test</i>): decidim entre hipòtesis H_P i H_A

Error de tipus I

Fisher	Neyman-Pearson
Tot l'experiment és <i>a posteriori</i> , per tant no importa si hi ha errors de tipus I.	No es pot saber amb només un intent del test, s'ha de considerar <i>in the long run</i> .
Es pot canviar a priori; recordar la Bonferroni quan volies fer diversos test.	S'ha de minimitzar durant el disseny de l'experiment (escollir la mínima α)

Tècnicament, Fisher no parla d'error tipus I, però la idea de no rebutjar H_0 quan s'havia de rebutjar és anàloga a l'error Tipus I de Fisher.

Nivell de Significació (α)

Fisher	Neyman-Pearson
Segons el resultat i els valors del <i>p-value</i>	S'escull a priori i depen de l'experiment.
Test de força (<i>test of strength</i>): donem un grau de rebuig d' H_0 segons el nivell de significació.	Test d'acceptació (<i>acceptation test</i>), binari: Si $\alpha \in R_c \Rightarrow$ acceptem H_A ; $\alpha \notin R_c \Rightarrow$ acceptem H_P

Per exemple, amb un *p-value* $p = 5.43 \times 10^{-4}$ amb un nivell de significació $s = 0.05$, podem dir que el grau de rebuig de H_0 és alt, és a dir, molt probablement les dades de l'experiment no segueixen la distribució H_0 . En canvi, no podem dir que H_P s'accepta “amb força” o “pels pels” ja que és un test d'acceptació.

Glossari de Termes

Terme	Definició
Hipòtesi nul · la (Null Hypothesis, H_0)	Hipòtesi que es vol contrastar o refutar. En Fisher, és la hipòtesi que es busca rebutjar amb evidència.
Hipòtesi alternativa (Alternative Hypothesis, H_A)	En Neyman-Pearson, és l'alternativa a la hipòtesi principal. Representa una situació diferent a la proposada per H_P .

Mida de l'efecte (Effect size)	Mesura que quantifica la magnitud d'un fenomen. En Neyman-Pearson, es considera explícitament en la formulació de les hipòtesis.
Error de tipus I (Type I error)	Error que es comet en rebutjar la hipòtesi nul · la sent aquesta vertadera (fals positiu).
Error de tipus II (Type II error)	Error que es comet en no rebutjar la hipòtesi nul · la sent aquesta falsa (fals negatiu).
Potència (Power)	Probabilitat de rebutjar correctament la hipòtesi nul · la quan és falsa (1 - probabilitat d'error tipus II).
Regió crítica (Critical region)	Conjunt de valors de l'estadístic que porten a rebutjar la hipòtesi nul · la.
Nivell de significació (Significance level, α)	Probabilitat màxima d'error tipus I que estem disposats a acceptar.
Test de força (Test of strength)	Enfocament de Fisher que mesura l'evidència contra la hipòtesi nul · la.
Test d'acceptació (Acceptance test)	Enfocament de Neyman-Pearson que decideix entre dues hipòtesis alternatives.
MSE (Minimum Size Effect)	Mínim efecte considerat significatiu en el context de l'estudi.