

TESTAREA IPOTEZELOR DE CERCETARE PRIN STATISTICA BAYESIANĂ

Maria PAVEL, dr., conf. univ., Catedra ITI, UPSC

<https://orcid.org/0000-0003-4803-6398>

Dorin PAVEL, dr., conf. univ., Catedra ITI, UPSC

<https://orcid.org/0000-0002-9600-1360>

Rezumat. În lucrare se scot în evidență restricțiile statisticii clasice de testare a ipotezelor de cercetare formulate de cercetătorii din domeniul științelor sociale, inclusiv din domeniul psihopedagogic și se descrie alternativa modelului de testare a semnificației ipotezei nule, ca un nou model de statistică - statistica Bayesiană. Acest model se numește „Null Hypothesis Bayesian Testing” (NHBT) și presupune utilizarea factorilor Bayes în loc de valorile semnificației p .

Cuvinte cheie: cercetare pedagogică, ipoteză de cercetare, prag de semnificație, model NHST, model NHBT, statistică Bayesiană.

Abstract. The paper highlights the restrictions of classical statistics for testing research hypotheses formulated by researchers in the field of social sciences, including the psychopedagogical field, and describes the alternative to the model of testing the significance of the null hypothesis, as a new statistical model - Bayesian statistics. This model is called "Null Hypothesis Bayesian Testing" (NHBT) and involves using Bayes factors instead of significance p values..

Keywords: pedagogical research, research hypothesis, significance threshold, NHST model, NHBT model, Bayesian statistics.

Introducere

Metodele de cercetare a diverselor variabile ale fenomenului educațional, printre care și experimentul pedagogic, furnizează date ce necesită o analiză statistică pentru a formula careva concluzii. Această analiză și prelucrare statistică a datelor, de obicei argumentează soluționarea problemei științifice înaintate, inclusiv și prin interpretarea semnificației ipotezelor de cercetare.

Statistica inferențială tradițională utilizează criteriile parametrice (teste t) și neparametrice (coeficienți z) pentru a testa semnificația ipotezei nule H_0 , cu scopul de a o respinge și de a accepta ipoteza alternativă H_1 . Acest model, cunoscut sub numele „Null Hypothesis Significance Testing” (NHST) se utilizează de foarte mult timp de statisticieni, însă literatura de specialitate scoate în evidență în ultimul timp un șir de neajunsuri ale acestuia. În primul rând, se testează de fapt semnificația ipotezei nule, a cărei adevăr nu se demonstrează niciodată: fie se respinge și se acceptă ipoteza alternativă, fie că nu poate fi respinsă. Prin urmare, ipoteza H_1 nu se testează de fapt, ci se acceptă. O altă limită a modelului statistic de testare a semnificației ipotezei nule, se referă la valoarea de raportare a pragului de semnificație, care a fost stabilită aleatoriu la 0,05 de către Ronald A. Fisher [1] și creează confuzii în rândul statisticienilor în cazul când semnificația unei ipoteze este

apropiată de această valoare (de ex. $p=0,059$, ce se întâmplă? Se acceptă sau nu ipoteza alternativă?). Cu atât mai mult că pentru $p=0,05$ ambele ipoteze au aceeași probabilitate. De asemenea, în cazul când pragul de semnificație este mai mic decât 0,05 se respinge ipoteza nulă și se confirmă ipoteza alternativă, dar aceasta nu înseamnă că este adevărată. Totodată mărimea semnificației ipotezei testate depinde de valoarea cât mai mică a lui p . Și nu în ultimul rând, atunci când se verifică semnificația ipotezei alternative repetat, aceeași ipoteză nulă, testată anterior, rămâne ca punct de referință, fie că s-a respins sau nu anterior.

În acest context, un șir de statisticieni (D. V. Lindley, W. W. Rozeboom, J. Tendeiro, H. Kiers ș.a.) scot în evidență incertitudinea, caracterul dogmatic, de rutină a modelului NHST. Printre soluțiile propuse se menționează fie eliminarea pragului semnificației (Cohen, Benjamin, Berger, McShane, Gal, Ioannidis ș.a.), micșorarea acestuia la 0,005 sau chiar 0,001 (Benjamin, Berger, Johnson), sau propunerea unor noi criterii de interpretare a pragului de semnificație, fie introducerea unei noi statistici alternative modelului NHST [2].

Această din urmă alternativă propusă, se numește „Null Hypothesis Bayesian Testing” (NHBT, Testarea Bayesiană a ipotezei nule) și presupune utilizarea factorilor Bayes în loc de valorile semnificației p .

Principiile Statisticii Bayesiane

Sunt bine-cunoscute contribuțiile semnificative ale lui Thomas Bayes, matematician englez (1702 – 1761) în domeniul teoriei probabilității, în special pentru dezvoltarea teoremei cu același nume, care a avut un impact major în statistică și în analiza datelor.

Teorema lui Bayes descrie matematic modul în care putem actualiza probabilitățile bazate pe noile informații disponibile. Ea este fundamentală în statistica bayesiană, ce se dezvoltă ca o ramură a statisticii și se bazează pe aplicarea teoremei lui Bayes pentru a obține estimări și concluzii în contexte unde există incertitudine. Această teorie este utilizată într-o varietate de domenii, inclusiv în analiza datelor, în inteligența artificială, în învățarea automată și în multe alte științe și industrii.

Testarea ipotezelor de cercetare prin statistica Bayesiană este o abordare alternativă la testarea statistică tradițională (cum ar fi testele t și z) care se bazează pe principiile teoriei probabilității bayesiene. Statistica Bayesiană permite cercetătorilor să combine datele experimentale cu cunoștințele prealabile (cunoscute sub numele de distribuție apriori) pentru a obține o înțelegere mai completă a ipotezelor de cercetare.

Principalele componente ale statisticii bayesiene includ:

- Distribuția apriori (*Prior*) - cunoștințele sau convingerile preexistente cu privire la ipoteza de cercetare înainte de a colecta datele, ce presupune descrierea amănunțită a tehnicilor de analiză a datelor, cu precizarea volumului eșantionului de cercetare.

- Distribuția de probabilitate condiționată (*Likelihood*) - modul în care datele observate sunt legate de ipoteza de cercetare. Poate fi exprimat prin intermediul funcției de verosimilitate.
- Distribuția aposteriori (*Posterior*) - distribuția de probabilitate actualizată pentru ipoteza de cercetare după colectarea datelor. Această distribuție combină distribuția apriori cu distribuția de probabilitate condiționată și este obținută prin intermediul teoremei lui Bayes.

Statistica bayesiană este o abordare a analizei datelor care se bazează pe teorema lui Bayes pentru a evalua și actualiza probabilitățile în funcție de noile informații disponibile. Principiile statisticii bayesiene pot fi explicate în patru pași principali: specificarea distribuției anterioare, actualizarea cu date noi folosind teorema lui Bayes, calcularea distribuției posterioare și luarea deciziilor sau formularea concluziilor.

Modelul NHBT începe cu distribuția anterioară, cunoscută, cu privire la variabila sau parametrul de interes, ce se poate baza pe date prelabile, expertiză sau alte informații disponibile. De exemplu, dacă se dorește să se estimeze probabilitatea de succes a unei echipe de elevi într-un concurs internațional de matematică sau informatică, se poate identifica o distribuție anterioară bazată pe performanța echipei la concursurile anterioare sau pe informații din edițiile anterioare ale concursului. Distribuția apriori se caracterizează prin gradul de precizie, sau prin gradul de subiectivitate. După precizie, distribuțiile apriori pot fi: 1. *neinformative*, când cercetătorul presupune că nu cunoaște nimic, distribuția este plată; și 2. *informative*, atunci când cercetătorul își asumă o careva distribuție. După gradul de subiectivitate, distribuțiile pot fi, în general, *subiective*, *obiective*, cu diferite variații sau combinații, în dependență de autorul ce propune această clasificare (Z. Dienes, R. Heide, P. D. Grunwald ș.a.).

Următorul pas al modelului NHBT constă în actualizarea datelor, folosind teorema lui Bayes, ce spune că *probabilitatea* distribuției anterioare, față de observațiile noi, este proporțională cu produsul dintre distribuția anterioară și *verosimilitatea* datelor (probabilitatea datelor sub distribuția anterioară). De exemplu, dacă se observă că echipa a câștigat ultimele două concursuri, se actualizează distribuția anterioară cu aceste informații pentru a obține o distribuție posterioară a probabilității de succes. În acest caz *probabilitatea datelor* sau dovezilor (D) cercetării depind de o anumită ipoteză formulată (H), prin urmare probabilitatea este condiționată – $P(D|H)$. Formula matematică a teoremei lui Bayes poate fi scrisă astfel [3, p. 37, 2, p. 30]:

$$P(H|D) = \frac{P(D|H) \times P(H)}{P(D)},$$

care în cuvinte, se poate scrie astfel:

$$Posterior = \frac{Likelihood \times Prior}{\text{media Likelihood}}$$

În continuare, modelul NHBT recurge la calcularea distribuției posterioare, ce reprezintă estimarea actualizată a variabilei sau parametrului de interes, date fiind informațiile anterioare și datele noi. Această distribuție oferă o distribuție a probabilităților, în loc de un singur punct estimat. De exemplu, după actualizarea cu datele despre concursurile câștigate, se obține o distribuție nouă a probabilității de succes a echipei în concursul următor. Distribuția posteriori reprezintă rezultatul cercetării și este produsul dintre distribuția apriori și funcția de verosimilitate.

La final, se formulează concluzii informate și/sau se iau decizii, ce implică estimări punctuale sau intervale de încredere bazate pe distribuția posterioară, se pot lua în considerare și riscurile. De exemplu, pe baza distribuției posterioare a probabilității de succes a echipei la un concurs, se poate decide dacă să se investească în echipa respectivă sau să aleagă o altă strategie.

Așa cum s-a menționat anterior testarea ipotezelor în cazul modelului statistic NHBT se face pe baza *factorilor lui Bayes* (notați cu BF_{01} și BF_{10}). Aceștia se calculează în urma evaluării probabilității de susținere a celor două ipoteze (modele) - nulă și alternativă:

- modelul pentru H_0 :

$$P(H_0|D) = \frac{P(D|H_0) \times P(H_0)}{P(D)},$$

- modelul pentru H_1 :

$$P(H_1|D) = \frac{P(D|H_1) \times P(H_1)}{P(D)}.$$

Dacă se va nota raportul de șansă pentru distribuția apriori (verosimilitatea relativă a ipotezei alternative față de ipoteza nulă până la obținerea datelor) astfel:

$$\text{\textit{șansa Prior}} = \frac{P(H_1)}{P(H_0)},$$

iar raportul de șansă pentru distribuția posteriori (verosimilitatea relativă a ipotezei alternative față de ipoteza nulă după obținerea datelor) astfel:

$$\text{\textit{șansa Posterior}} = \frac{P(H_1|D)}{P(H_0|D)},$$

atunci factorii Bayes se vor calcula după formulele:

$$BF_{10} = \frac{\text{\textit{șansa Posterior}}}{\text{\textit{șansa Prior}}} \text{ și } BF_{01} = \frac{\text{\textit{șansa Prior}}}{\text{\textit{șansa Posterior}}}.$$

Acești factori pot lua valori între 1 și ∞ și cuantifică gradul de suport a unei ipoteze sau a celeilalte și nu semnifică mărimea efectului (*Effect Size*), ci estimează cât de verosimil este acesta în condițiile datelor obținute [2 - 5]. În tabelul 1 este prezentată scala de măsură a suportului ipotezei alternative față de ipoteza nulă și invers și interpretarea valorilor factorilor Bayes.

Tabelul 1. Interpretarea valorii factorilor Bayes [6]

<i>Valoare</i>	<i>Interpretare</i>
> 100	dovezi extreme, decisive pentru ipoteza alternativă
30-100	dovezi foarte puternice pentru ipoteza alternativă
10-30	dovezi puternice pentru ipoteza alternativă
3-10	dovezi moderate pentru ipoteza alternativă
1-3	dovezi „anecdotice” pentru ipoteza alternativă
1	ambele ipoteze au aceeași verosimilitate
$\frac{1}{3} - 1$	dovezi „anecdotice” pentru ipoteza nulă
$\frac{1}{10} - \frac{1}{3}$	dovezi moderate pentru ipoteza nulă
$\frac{1}{30} - \frac{1}{10}$	dovezi puternice pentru ipoteza nulă
$\frac{1}{100} - \frac{1}{30}$	dovezi foarte puternice pentru ipoteza nulă
$< \frac{1}{100}$	dovezi extreme, decisive pentru ipoteza nulă

Concluzii

Deși statistica Bayesiană nu este perfectă și sunt cercetători care au identificat limitele acesteia, totuși din cele expuse mai sus, pot fi evidențiate avantajele modelului NHBT:

- permite actualizarea cunoștințelor anterioare cu date noi într-un mod consistent;
- poate oferi distribuții de probabilitate pentru parametrii modelului, ceea ce furnizează mai multă informație decât valorile punctuale estimate;
- abordează în mod natural problemele cu eșantionări mici sau date incomplete, deoarece poate să facă uz de distribuții apriori informative.

Articol realizat în cadrul proiectului de cercetări științifice „Metodologia implementării TIC în procesul de studiere a științelor reale în sistemul de educație din Republica Moldova din perspectiva inter/transdisciplinarității (concept STEAM)”, inclus în „Program de stat” (2020-2023), Prioritatea IV: Provocări societale, cifrul 20.80009.0807.20, cu suportul financiar oferit de Agenția Națională pentru Dezvoltare și Cercetare

Bibliografie

1. FISHER, R. A. *Statistical methods for research workers*. Disponibil online: <https://psychclassics.yorku.ca/Fisher/Methods/chap4.htm>
2. POPA, M.; LUCA, G.; BRAN, G. *Inferența Bayesiană*. București, 2020. Disponibil online: <https://osf.io/v2m7g/download>
3. McELREATH, R. *Statistical Rethinking: A Bayesian Course with Examples in R and Stan*. Chapman and Hall/CRC, 2020. 469 p. ISBN 978-036713991-9.
4. KRUSCHKE, J. K. *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Elsevier Academic Press, 2015. 759 p. ISBN: 978-0-12-405888-0.
5. GELMAN, A.; CARLIN, J. B.; STERN, H. S.; DUNSON, D. B.; VEHTARI, A.; RUBIN, D. B. *Bayesian Data Analysis*. 677 p. 2013.
6. <https://www.statology.org/bayes-factor/>