

Verificarea ipotezelor statistice, în general, este un procedeu care are 5 etape.

Etapa 1. Formularea ipotezei nule H_0

Prin ipoteză înțelegem o afirmație care susține că ceva este adevărat. În general, ipoteza nulă este o afirmație relativă la un parametru al unei populații și afirmă că parametrul are o valoare dată. Adesea expresia "nu diferă" este folosită în formularea ei, de aici vine numele de ipoteză nulă. (diferența este nulă)

Etapa 2. Formularea ipotezei alternative H_a

Ipoteza alternativă H_a este o afirmație relativă la același parametru al populației care apare în ipoteza nulă H_0 . în ipoteza H_a se afirmă că parametrul are o valoare diferită de cea susținută în H_0 .

Deciziile care se iau sunt prezentate în tabelul următor:

Decizia	Ipoteza H_0 este	
	Adevărată	Falsă
Nu respingem H_0 (acceptăm)	decizie corectă Tip A	eroare Tip II
Respingem H_0	eroare Tip I	decizie corectă Tip B

Ar fi foarte frumos ca de fiecare dată când luăm decizii să luăm decizii corecte, dar aceasta este statistic imposibil pentru că ne bazăm pe informații furnizate de eșantioane. Cel mai bun lucru la ce putem spera este **să controlăm riscul sau probabilitatea de a comite o eroare:**

Eroarea	Tipul de eroare	Probabilitate
Respingerea unei ipoteze adevărate	I	α
Acceptarea unei ipoteze false	II	β

Etapa 3 Metodologia de verificare a ipotezelor:

- 1) identificarea unui test statistic;
- 2) specificarea nivelului de semnificație (de neîncredere);
- 3) determinarea regiunii critice.

1. Testul statistic este determinat în funcție de parametrul care se studiază și se calculează statistica acestuia în funcție de datele din eșantion.
2. Frecvent valorile pentru nivelul de neîncredere se ia 0.05 sau 0.01.

3. Se determină regiunea critică (valorile critice) folosind α . Regiunea critică sau regiunea de respingere este definită cu ajutorul valorilor critice ("critical values") care sunt quantile ale repartiției pe care o urmează statistica testului ales. Dacă testul este bilateral, regiunea critică este de forma

$$(-\infty, -\text{valoarecritica}) \cup (\text{valoarecritica}, \infty),$$

iar dacă testul este unilateral, regiunea critică este de forma: $(-\infty, -\text{valoarecritica})$ sau $(\text{valoarecritica}, \infty)$.

Etapa 4. Formularea concluziei testului

Dacă valoarea statisticii se află în regiunea critică, H_0 se respinge; altfel H_0 nu se respinge.

P-valoarea unui test statistic este cea mai mică valoare a nivelului de semnificație α pentru care informația extrasă din eșantion este semnificativă (H_0 adevărată se respinge). Determinarea regiunii critice este înlocuită de calcul **p-valorii**.

P-valoarea este aria aflată între curba de densitate de probabilitate axa Ox și $z = z^*$. Există trei cazuri posibile: două unilaterale și unul bilateral. Direcția (sau semnul) în ipoteza H_a este indiciul:

Cazul 1. Dacă H_a este unilaterală la dreapta (" $>$ ") atunci $p = P(z > z^*)$ și aria este în dreapta lui z^* .

Cazul 2. Dacă H_a este unilaterală stânga (" $<$ "), atunci $p = P(z < z^*)$ este aria din stânga lui z^* .

Cazul 3. Dacă H_a este bilaterală (" \neq "), atunci $p = P(z < -|z^*|) + P(z > |z^*|) = 2 \cdot P(z > |z^*|)$

Decizia se ia comparând P -valoarea cu nivelul de semnificație α :

- a) Dacă $P \leq \alpha$ atunci H_0 se respinge;
- b) Dacă $P > \alpha$ atunci H_0 se acceptă.