Lp 8
Rezumat

Capitolul 6. PROBLEMA TESTĂRII IPOTEZELOR

În ştiinţele experimentale verificarea *ipotezelor ştiinţifice* cere testări statistice. Testăm, deci, anumite *ipoteze statistice*.

6.1. Ipoteză ştiinţifică - Ipoteză statistică

6.1.1. Ipoteză ştiinţifică

Sn **ipoteză ştiinţifică** tentativa de a explica una sau mai multe observaţii. Ea trebuie să fie în concordanţă cu datele şi, dacă e falsă, să permită dovedirea acestui lucru prin experiment. Dacă e adevărată, nu se va putea dovedi niciodată că este falsă, dar nici că este adevărată. Deci, vom considera adevărată o ipoteză atât timp cât nu putem dovedi (prin experiment) că este falsă [13].

De exemplu, vom CONSIDERA că un anumit zar sau o anumită monedă sunt nemăsluite dacă, făcând EXPERIMENTUL aruncării cu zarul sau moneda respectivă de multe ori, vom obţine fiecare faţă cam de acelaşi număr de ori (vezi problemele 12 şi 13 în Lp 9). În caz contrar, spunem că am demonstrat experimental că zarul, respectiv moneda, sunt măsluite (vezi problema 14 din Lp 9).

6. 1. 2. Ipoteză statistică

Ipotezele ştiinţifice sunt emise întotdeauna din observaţii parţiale, dar vizează proprietăţi generale. Pentru susţinere au nevoie de verificări prin aşa-numitele ipoteze statistice.
Sn **ipoteză statistică** o aserţiune cu privire la una sau mai multe populaţii statistice. Se prezintă întotdeauna sub forma unui cuplu: ipoteză **nulă** (sau **de nul**) H_0 - ipoteză **alternativă** H_A. Exemple pentru una sau două populaţii statistice:

H_0: $\mu_1 = \mu_0$ (sau $\mu_1 - \mu_0 = \mathbf{0}$) H_A: $\mu_1 \neq \mu_0$ (sau $\mu_1 - \mu_0 \neq 0$) **test bilateral**	H_0: $\mu_1 = \mu_0$ (sau $\mu_1 - \mu_0 = \mathbf{0}$)* H_A: $\mu_1 > \mu_0$ (sau $\mu_1 - \mu_0 > 0$) **test unilateral dreapta**	H_0: $\mu_1 = \mu_0$ (sau $\mu_1 - \mu_0 = \mathbf{0}$) H_A: $\mu_1 < \mu_0$ (sau $\mu_1 - \mu_0 < 0$) **test unilateral stânga**

în care μ_1 şi μ_0 sunt medii de populații din care s-au extras aleator unul sau două eşantioane.

În 6. 2. se presupune că μ_0 este cunoscut şi s-a extras doar un eşantion de medie m_1 dintr-o populație cu media μ_1 necunoscută. Mai puțin riguros (decât formularea de mai sus), putem spune că se testează dacă media unui eşantion m_1 este conformă cu media unei populații μ_0. Mai general, se testează dacă o statistică este c o n f o r m ă cu un anumit parametru. De aceea, acest tip de test sn **test de conformitate**. În cazul mediei se întâlneşte şi sub forma "compararea unei medii empirice cu o medie teoretică".

6.2. Etapele construcției şi aplicării unui test statistic

Etapele aplicării unui test statistic	Observații:
1. Formularea clară a problemei pentru care se doreşte o decizie.	Cade exclusiv în sarcina biologului, ecologului, medicului, agronomului..
2. Identificarea: **a.** variabilei ca tip şi scală	Etapă esențială. Eventual cu ajutorul biometricianului.
b. eşantionului • ca mod de extragere • ca volum	 Randomizare pentru valabilitatea testului. Calculabil asistat de biostatistician.
c. informațiilor despre distribuția variabilei în populație (gaussianitate, posibilitatea cvasigaussianizării printr-o anumită transformare).	Din literatură şi/sau cu asistența biometricianului specializat pe domeniul biologic respectiv.
3. **a.** Stabilirea distribuției de eşantionaj care impune **b.** statistica testului.	De regulă, se preiau din literatura biostatistică sau de statistică matematică aplicată.
4. Bazat pe 1-3 **a.** formularea cuplului ipoteza nulă - ipoteza alternativă, care determină **b.** tipul de test (bilateral, unilateral stânga, unilateral dreapta)	Abilitate obligatorie pentru a se putea apela, eventual, la programele informatice.
5. Stabilirea nivelului de semnificație α sub care vom respinge ipoteza nulă (etapă denumită impropriu† **regula de decizie**).	De competența esențială a biologului.

* vezi Anexa 8.

6. **a.** Efectuarea calculelor	
b. obținerea deciziei statistice prin: ***I)*** calcul manual: dacă statistica testului este "mai excentrică"[‡] decât α-cuantila corespunzătoare (citită în tabela corespunzătoare) $\Rightarrow$ se respinge H_0.	Etapă cu rol instructiv esențial.
II) calcul automat: dacă riscul de respingere a lui H_0 corespunzător statisticii testului (risc notat α_c) < α fixat la pct. 5 $\Rightarrow$ se respinge H_0.	De preferat nu numai pentru uşurinţă ci şi pentru posibilitatea publicării rezultatului împreună cu α_c (nivelul cu care putem respinge H_0) chiar şi atunci când acceptăm H_0, ori când α_c < 1‰ – vezi 6.4..

1° Ce înseamnă grade de libertate ?

Numărul de grade de libertate = volumul eşantionului - numărul de relații (legături) utilizate în construcția statisticii testului, pe baza datelor din eşantion.

2° Regiune de acceptare şi regiune de respingere

Orice test statistic oferă legea de distribuție a statisticii sale atunci când ipoteza de nul (H_0) este considerată adevărată. De regulă, aceste distribuții au cel puțin o coadă nemărginită şi valorile sale centrale sunt cele mai probabile (atunci când H_0 este adevărată). De aceea, decizia statistică se va lua prin împărțirea în două zone a domeniului de variație al statisticii, adică a zonei de sub distribuția respectivă: o zonă de respingere a H_0 care acoperă coada sau cozile şi o zonă de acceptare a H_0 plasată pe valorile cele mai probabile. Zona de respingere va avea aria relativă α, adică nivelul de semnificație ales.

6.2.1. Diferențe semnificative versus nesemnificative în cazul testelor bilaterale

Dacă se respinge ipoteza nulă pentru **α = 0,05 (= 5%)**, se spune că testul a evidențiat o diferență *semnificativă* şi se marchează prin semnul "*". În caz de acceptare, diferența este considerată *nesemnificativă*. Deoarece orice test statistic conduce la decizia: semnificativ / nesemnificativ, testele statistice sn şi **teste de semnificație** (în sens larg).

[†] Mai adecvată ar fi fost, poate, denumirea de nivel de decizie, căci regula este aceeaşi la orice test: alegem ipoteza alternativă dacă statistica testului "cade" în zona de respingere a testului (vezi punctul 6. b. I. din tabel şi în continuare).

[‡] În rezumatul din acest volum referitor la subparagraful 6. 7. 1. se explicitează această exprimare.

O *diferență nesemnificativă* înseamnă (în cazul testelor statistice) o diferență suficient de mică, încât putem considera că a apărut doar ca rezultat al fluctuațiilor de eşantionaj, adică al inerentelor deosebiri între eşantion şi populația specificată în ipoteza nulă.
O *diferență semnificativă* înseamnă, în mod contrar, o diferență suficient de mare ca să nu o atribuim fluctuațiilor de eşantionaj, ci să presupunem că provine dintr-un motiv semnificativ, nu din întâmplare, şi anume pentru că eşantionul provine dintr-o populație diferită de populația din ipoteza nulă.

6.2.2. Diferențe foarte semnificative în cazul testelor bilaterale

Dacă se respinge ipoteza nulă pentru $\boldsymbol{\alpha}$ **= 0,01 (= 1%)**, se spune că testul a evidențiat o diferență *foarte semnificativă* şi se marchează prin semnul "**".

6.2.3. Diferențe înalt semnificative în cazul testelor bilaterale

Dacă se respinge ipoteza nulă pentru $\boldsymbol{\alpha}$ **= 0,001 (= 1‰)**, se spune că testul a evidențiat o diferență *înalt semnificativă* şi se marchează prin semnul "***".

6.3. Niveluri de semnificație standard în cazul testelor bilaterale, diagnosticele statistice corespunzătoare şi zona de incertitudine

În cazul calculului automat (care produce α_c) sau al utilizării tabelelor directe (care oferă p):

Praguri pentru α_c sau p	**Diagnostic statistic:** Explicit	codificat
$\alpha_c < 0{,}001$	diferență înalt semnificativă	***
$0{,}001 \le \alpha_c < 0{,}01$	diferență foarte semnificativă	**
$0{,}01 \le \alpha_c < 0{,}05$	diferență semnificativă	*
$0{,}05 \le \alpha_c$	diferență nesemnificativă	n.s. §

Aceleaşi diagnostice se pot obține şi prin calcul manual şi consultarea tabelelor indirecte. Despre p-ul de respingere nu vom şti însă decât că se află intre două niveluri tabelate succesiv. De exemplu, dacă vom obține respingerea lui H_0 pentru $\alpha = 0{,}01$ şi acceptarea pentru $\alpha = 0{,}001$, vom şti doar că p se află undeva în intervalul [0,001; 0,01). De aceea se spune că s-a respins H_0 cu (un) $p < 0{,}01$. (Prin convenție se subînțelege că $p \ge 0{,}001$). În cazul testelor bilaterale, intervalul **[0,01; 0,05)** pentru α_c sau p, sn **zonă de incertitudine.** În consecință, dacă într-un experiment sau observație se obține o diferență doar

§ **n.s.** = **n**on **s**emnificativ.

semnificativă (adică în zona de incertitudine) se recomandă repetarea experimentului sau a observației pe un număr superior de cazuri, luându-se decizia finală pe eşantionul de volum mai mare. Această regulă se bazează pe insuficienta repetabilitate a respingerii ipotezei nule (50%) în cazul în care diferența este doar semnificativă ($p < 0,05$) – vezi problemele 2, 3 şi 7.

6.4. Modul de redactare a rezultatului unui test statistic

♦ În cazul calculului manual:

- dacă s-a respins H_0 spunem şi scriem că:

"s-a respins ipoteza nulă cu un risc $p < \alpha_{stabilit}$" sau că

"testul a evidențiat o diferență (/ foarte / înalt) semnificativă ($p < \alpha_{stabilit}$)". Calificativele "foarte" şi "înalt" se pun în funcție de valorile lui $\alpha_{stabilit}$ după convențiile de mai sus.

- dacă s-a acceptat H_0 spunem şi scriem că:

"s-a acceptat ipoteza nulă (** $p \geq \alpha_{stabilit}$)" sau că

"testul nu a evidențiat o diferență semnificativă ($p \geq \alpha_{stabilit}$)" sau că

"diferența este nesemnificativă ($p \geq \alpha_{stabilit}$)".

Notațiile folosite în redactare şi semnificațiile lor în cazul testelor bilaterale pot fi centralizate astfel:

Convenții clasice:	**Notații prescurtate:**	**Semnificație:**
$p < 0,001$	***	S-a respins H_0 cu un risc $p < 0,001$
$p < 0,01$	**	S-a respins H_0 cu un risc $p < 0,01$ dar $\geq 0,001$
$p < 0,05$	*	S-a respins H_0 cu un risc $p < 0,05$ dar $\geq 0,01$
$p \geq 0,05$	n.s.	S-a acceptat H_0, căci s-ar fi putut respinge doar cu un $p \geq 0,05$.

♦ În cazul calculului automat:

“Respingem H_0 cu $p < \alpha_c$“ - atunci când $\alpha_c \leq 0,05$ (sau, mai rar, 0,1) pentru biologi şi $\alpha_c \leq 0,1$ (sau 0,2) pentru ecologi, respectiv

“Acceptăm H_0 ($p < \alpha_c$)” – atunci când $\alpha_c > 0,05$† (sau, mai rar, 0,1) pentru biologi şi $\alpha_c > 0,1$ (sau 0,2) pentru ecologi.

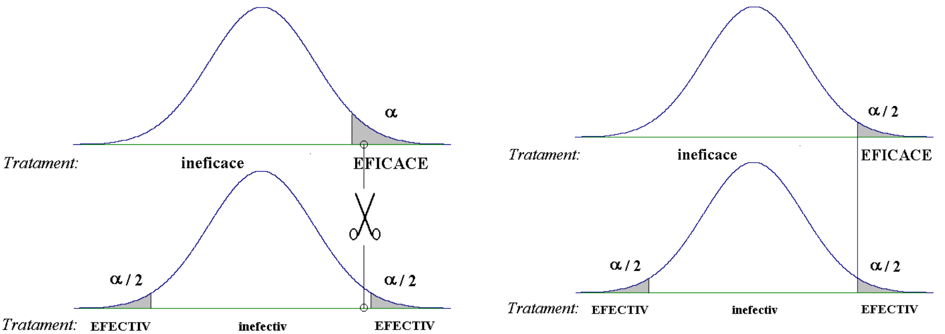
Observația 1: În [4] se afirmă că p (în calcul manual, respectiv α_c, în calcul automat) este adesea considerat drept un indice care reflectă gradul de discordanță între date şi ipoteza nulă. În consecință, atunci când este foarte mic ($p < 1‰$) se crede că, în cazul repetării experimentului, ipoteza nulă va fi

** Aici se gândeşte «deoarece», fără a se spune sau scrie.

† Se interpretează astfel: «Acceptăm H_0, deoarece p este mai mic doar decât un α_c mai mare ca 0,05».

respinsă aproape sigur, ceea ce este relativ eronat, probabilitatea repetării respingerii lui H_0 fiind doar de 91%. (A se vedea, mai jos, problemele 2, 3, 6 şi 7 inspirate din [4]). Rezultă, de aici, importanţa publicării α_c-ului de respingere cu atât mai mult, atunci când este mai mic decât 1‰. (Am calculat că o respingere cu un p < 0,1‰ garantează - în cazul distribuţiilor de eşantionaj gaussiene - o repetabilitate a respingerii ipotezei nule de circa 97%.) Este foarte important de publicat α_c şi atunci când depăşeşte cu puţin $\alpha_{stabilit}$. Justificarea acestei recomandări depăşeşte însă nivelul cadrului de faţă.

Observaţia 2: **ÎN CAZUL TESTELOR UNILATERALE** se lucrează fie cu aceleaşi niveluri α, fie cu jumătăţile acestora. Noi **recomandăm utilizarea nivelurilor α / 2** (vezi desenul din dreapta a figurii următoare). În acest fel se poate înlătura contradicţia semnalată în desenul din stânga figurii. În figura, sunt reprezentate zonele de acceptare şi de respingere ale unui test unilateral (vezi partea de sus a figurii), respectiv, bilateral (vezi partea de jos a figurii). Am ales intenţionat problema particulară a eficacităţii, respectiv, efectivităţii unui tratament, căci astfel, se observă mai bine contradicţia la care se ajunge dacă nu lucrăm cu varianta recomandată aici.
Amintim că “efectiv = cu efect” şi “eficace = cu efect (deci efectiv) în sensul dorit”. Vezi şi problemele 8 şi 19 din Lp 9.



În desenul din stânga figurii se observă că dacă păstrăm, în cazul testelor unilaterale, pragul pentru riscul α de la testele bilaterale, se produce contradicţia: test declarat eficace, dar inefectiv (vezi punctul marcat printr-un cerc). Contradicţia dispare dacă **în cazul testelor unilaterale lucrăm cu α = 0,025 în loc de 0,05, adică, în general, dacă jumătăţim nivelul α de la testele bilaterale**, ca în dreapta figurii anterioare. Soluţia provine dintr-un cadru mai general din [4].

6.5. Tipuri de erori asociate unui test statistic şi riscurile acestora

Probabilităţi din care **riscuri**:

	Accept H_0	Resping H_0
H_0 adevărat	$1 - \alpha$	**α sn riscul furnizorului (riscul de speţa I)**
H_0 fals	**β sn riscul beneficiarului (riscul de speţa a II-a)**	$1 - \beta = \pi$

π sn **puterea testului**.

$1 - \alpha$ corespunde, în teoria estimaţiei, *gradului de încredere*.

6.5.1. Explicarea terminologiei "riscul furnizorului", "riscul beneficiarului"

Acceptarea sau respingerea unui lot de produse se face după verificarea unui eşantion extras prin randomizare din acel lot. De regulă, furnizorul menţionează un standard minim de calitate (μ_0). Astfel, ipoteza nulă poate fi I_0: $\mu_1 \geq \mu_0$ (H_0: $\mu_1 = \mu_0$) vezi Anexa 8. În consecinţă, riscul α de respingere a ipotezei nule când aceasta este adevărată va fi riscul furnizorului, căci acesta este cel care pierde retrăgând de pe piaţă întregul lot când, de fapt acesta corespunde standardului afirmat în ipoteza nulă. Dacă, invers, este adevărată ipoteza alternativă (H_A: $\mu_1 < \mu_0$), adică lotul nu corespunde standardului afirmat şi totuşi este acceptat, atunci riscul β este al beneficiarului.

$\alpha\downarrow \Leftrightarrow \beta\uparrow$ (pentru eşantioane de acelaşi volum n).

Pentru micşorarea simultană a ambelor riscuri trebuie mărit volumul eşantionului, n.

Sn **puterea unui test** = $1 - \beta$. Creşte, pentru un α fixat, o dată cu volumul n după o curbă sigmoidă (de forma unui *s* alungit). Creşterea este cu atât mai rapidă cu cât α este mai mare.

Un **test** T_1 este **mai puternic** decât un alt test T_2 dacă respinge H_0 cu un volum $n_1 < n_2$, pentru acelaşi α fixat.

6.5.2. Justificarea nivelurilor de semnificație utilizate în medicină şi ecologie

Dacă admitem că, de regulă, ipoteza ştiințifică devine, într-un test statistic, ipoteză alternativă atunci se consideră:

1° α mic în medicină, biologie şi agricultură

α este riscul de a declara bolnav un pacient sănătos şi în consecință este riscul de a trata un individ sănătos, ceea ce îl poate îmbolnăvi din cauza efectelor secundare ale medicamentelor. Se ia α mic (1) pentru a respecta principiul hypocratic "primo non nocere, deinde vindecare", (2) din considerente economice, (3) pentru că riscul β chiar dacă ar fi mare afectează un număr mic de indivizi şi (4) β se minimizează prin strategia aplicării mai multor teste sau prin consultarea mai multor medici. În [15] se recomandă pragurile α = 0,001 şi 0,01 pentru medicină şi 0,05 pentru agricultură.

2° α mare în ecologie (α = 0,1 sau 0,2)

α este riscul de a declara în pericol un ecosistem atunci când acest lucru nu este adevărat. Se ia mare (1) pentru a se obține un β mic, β reprezentând probabilitatea de a nu observa o degradare atunci când ea există, ceea ce poate fi dezastruos pentru o întreagă comunitate. β trebuie să fie mic dintr-un singur test, pentru că (2) vinovații distrugerii respective limitează la maximum accesul la aplicarea mai multor teste.

6.6. Observații importante pentru aplicații

6.6.1. Observații terminologice şi de interpretare statistică

- Dacă H_0 este respinsă, spunem că:
  "**avem o diferență** (sau **diferențe**, în funcție de test) **semnificativă**(**e**), **foarte semnificativă**(**e**), **respectiv înalt semnificativă**(**e**)",
  în funcție de valoarea lui α şi de tipul testului (bi sau uni lateral).
- Dacă H_0 nu este respinsă, spunem că:
  "***H₀*** **POATE FI adevărată ori că NU AVEM SUFICIENTE DATE pentru a fi, eventual, respinsă**".

6.6.2. Observații de filozofia științei

- **Ipoteza nulă** este *analogul* unei **ipoteze ştiinţifice adevărate**, amândouă neputând fi dovedite niciodată a fi adevărate.
- **Ipoteza alternativă** este *omologul* unei **ipoteze ştiinţifice de testat** care, prin acest mecanism, devine sustenabilă (verificabilă) statistic.
- **Mecanismul de testare a unei ipoteze nule** este o *demonstrație prin reducere la absurd* **într-o logică nuanțată**: respingem ipoteza nulă atunci când o consecință logică din aceasta este (nu chiar absurdă, ci doar) improbabilă.

6.6.3. Comparații între testele binare şi cele statistice

- **Testele statistice** sunt o specie de **teste binare**. Dacă stabilim drept clasă căutată X, mulțimea eşantioanelor posibil a fi extrase aleator din populația specificată în ipoteza nulă H_0, atunci: $Se = 1- \alpha$, iar $Sp = \pi = 1- \beta$.
- **Testele statistice** se aplică asupra unor *distribuții teoretice*.

Lp 8 Teste, exerciții și probleme

TG8. Durata 200'' pe calculator.

Intre cele doua riscuri existente intr-un test statistic este adevarata propozitia:
1. Daca alfa creste, creste si beta.
2. Alfa si beta sunt independente.
3. Daca alfa creste, beta scade.

Alegeti afirmatia corecta:
1. Daca respingem ipoteza nula ne asumam riscul beta.
2. Daca respingem ipoteza nula ne asumam riscul alfa.
3. Daca acceptam ipoteza nula ne asumam riscul alfa.

Gradul de incredere asociat unui interval de incredere corespunde in testarea ipotezelor statistice:
1. lui 1 - alfa, alfa fiind riscul de speta I;
2. lui alfa, alfa fiind riscul de speta I;
3. lui 1 - beta, beta fiind riscul de speta a II-a;

Un alfa calculat ne conduce la afirmarea inexistentei unor deosebiri semnificative daca:
1. alfa = 0,05
2. alfa >= 0,05
3. alfa >= 0,05%

O ipoteza statistica este:
1. o asertiune cu privire la o ipoteza stiintifica
2. o asertiune cu privire la o populatie statistica
3. o asertiune cu privire la una sau mai multe populatii statistice

Ce afirmatie este corecta:
1. Nivelul de semnificatie este riscul de speta a II-a iar puterea unui test este 1 - riscul de speta I.

2. Nivelul de semnificatie este riscul de speta I, iar puterea unui test este 1 - riscul de speta a II-a.
Ce afirmatie este corecta:
1. Nivelul de semnificatie = puterea testului
2. Nivelul de semnificatie = gradul de incredere
3. Nivelul de semnificatie = riscul de speta a doua
4. Nivelul de semnificatie = 1 - gradul de incredere

Puterea π a unui test statistic:
1. creste o data cu cresterea volumului n al esantionului
2. scade o data cu cresterea volumului n al esantionului
3. nu depinde de volumul n al esantionului

Atunci cand nu respigem ipoteza nula H0 spunem ca:
1. H0 este adevarata sau nu avem suficiente date pentru a fi respinsa.
2. H0 poate fi adevarata sau nu avem suficiente date pentru a fi respinsa.
3. H0 poate fi adevarata.
4. H0 este adevarata.

TC8. Durata 5'.

1. Ca şi în cazul estimării, verificarea ipotezelor statistice este un tip de ______________statistică.
2. Încercarea de a explica ştiinţific una sau mai multe observaţii se numeşte ___________________.
3. Care este afirmaţia corectă:
   a) o ipoteză falsă nu poate fi dovedită niciodată a fi falsă
   b) o ipoteză adevărată nu poate fi dovedită niciodată a fi adevărată
   c) o ipoteză adevărată poate fi dovedită întotdeauna a fi adevărată
4. Ipoteza statistică este o afirmaţie cu privire la una sau mai multe ___________________.
5. La aplicarea unui test statistic, decizia statistică de acceptare a H_0 se ia atunci când α_c este ______________ decât α fixat.
6. În cazul situaţiei de incertitudine a deciziei statistice se recomandă refacerea experimentului utilizând un eşantion cu ________________________________.
7. Când obţinem α_c < 0,001 înseamnă că testul a evidenţiat o diferenţă _______________.

8. Se numeşte zonă de incertitudine pentru α_C intervalul ________________.
9. Riscul de speța I se notează cu α şi se numeşte "riscul ________________" sau „riscul __________________".
10. ______________ testului se notează cu π şi este egală cu __________, în care _____ este riscul de speța a II-a.
11. Cu cât valoarea α este mai _______, cu atât mai mare este considerat nivelul de semnificație respectiv.
12. Corectați într-un singur loc afirmația următoare:

Riscul furnizorului, notat α, apare atunci când H_0 este adevărată şi o accept.

13. În medicină se preferă un risc β ________, iar în ecologie un risc β ________.
14. Dacă vrem să micşorăm riscul beneficiarului fără a mări riscul furnizorului, atunci trebuie să _________________________________.

Exerciții sau probleme rezolvate

Recomandare: Înaintea parcurgerii acestor probleme este indicat să se studieze Anexa 8.

1.

S-au făcut experimente de inginerie genetică pe un lot dintr-o specie de sfeclă de zahăr cu scopul de a obține un soi cu dimensiuni mai mari. Din lotul experimental s-a extras apoi un eşantion aleator simplu de 26 de sfecle şi s-au măsurat diametrele acestor exemplare obținându-se media de 17 cm şi abaterea standard de 2 cm. Ştiind că media în specia originară este de 15 cm, să se testeze dacă noua generație de sfeclă de zahăr constituie un soi cu diametrul mai mare decât al soiului originar, adică dacă s-a reuşit obținerea un soi mai productiv.

Rezolvare:
Deoarece trebuie să comparăm media unui eşantion cu media unei populații statistice, vom aplica un test de conformitate după următoarele etape:

1. **Formularea clară a problemei pentru care se doreşte o decizie:** Provine sau nu eşantionul de medie m_1 = 17 cm şi volum n = 26 dintr-o populație

cu media mai mare de 15 cm şi abaterea standard estimată prin abaterea standard [corectată $s_c = s \cdot \sqrt{\frac{n}{n-1}}$] a eşantionului ?

2. **Identificarea:**
   a) **variabilei:** tip măsurătoare pe scală raport;
   b) **eşantionului:** eşantion aleator simplu cu volum mic (n = 26 < 30);
   c) **informațiilor despre distribuția variabilei în populație:** Se admite în literatura biostatistică faptul că distribuția celei mai mari dimensiuni (aici diametrul rădăcinii) pentru un anumit soi este gaussiană.

3. a) **Stabilirea distribuției de eşantionaj:** $t = \frac{m-\mu}{s/\sqrt{n-1}} \in t_{n-1}$.

b) **Statistica testului:** $t_1 = \frac{m_1-\mu}{s/\sqrt{n-1}}$.

4. a) **Formularea cuplului de ipoteze logice, respectiv, statistice** (vezi Anexa 8):

$$\text{ipoteze logice}\begin{cases} I_0 : \mu \le 15cm \\ I_A : \mu > 15cm \end{cases} \qquad \text{ipoteze statistice}\begin{cases} H_0 : \mu = 15cm \\ H_A : \mu > 15cm \end{cases}$$

b) **Tipul de test:** test unilateral dreapta (zona de respingere este dată de $t > t_{n-1,\alpha}$.

5. **Regula de decizie:** Stabilim nivelul de semnificație α = 2,5 %, - vezi obs. 2 de la 6. 4.

6. a) **Efectuarea calculelor:**

$$t_1 = \frac{m_1-\mu}{s/\sqrt{n-1}} = \frac{17-15}{2/\sqrt{25}} = \frac{2}{2/5} = \frac{2\cdot 5}{2} = \frac{10}{2} = 5$$

b) **Obținerea deciziei statistice prin calcul manual:**

Valoarea tabelată: $t_{n-1,\alpha} = t_{25;\,0,025}$ = 2,060

t_1 = 5 > 2,060 ⇒ se respinge H_0 pentru α = 2,5 %.

Observație: Următoarea „rețetă practică” este de preferat. Consultăm tabela Student (vezi Anexa 3) pe linia corespunzătoare numărului de grade de libertate (25 aici) şi căutăm cele două „valori critice” între care se plasează ca mărime valoarea calculată $|t_1|$ (5 aici). Vom găsi doar valoarea 3,725. Citim nivelul de semnificație α corespunzătoare testului unilateral (vezi linia de sus a tabelei).

Acesta este 0,5‰. În concluzie, putem respinge ipoteza nulă (egalitatea mediei cu valoarea 15) cu un risc $p < 0{,}005$. Altfel spus, putem afirma, cu un risc sub 0,5‰, că media eşantionului diferă semnificativ de 15,. Dacă luăm în considerație pragurile standard (0,05; 0,01; 0,001) putem afirma existența unei diferențe înalt semnificative ($p < 0{,}001$) sau (***).

Decizie statistică finală: Se respinge H_0 cu (risc) $p < 0{,}5$ ‰. Testul a evidențiat o diferență înalt semnificativă (***).

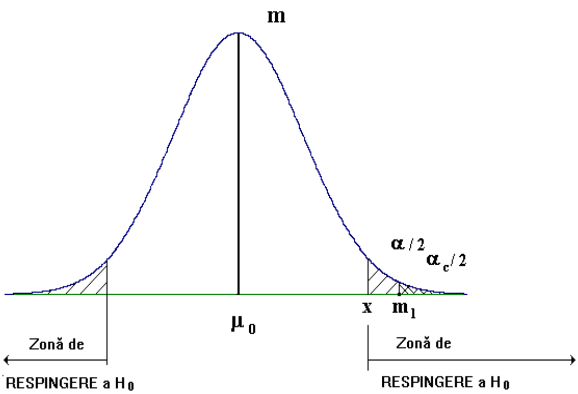
Decizie de specialitate: Putem afirma, cu un risc sub 0,5 ‰, că s-a obținut un soi nou!

2.

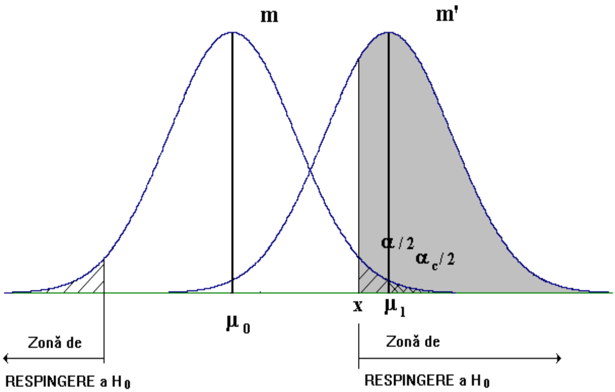
Să se calculeze probabilitatea repetării respingerii ipotezei nule de către un test de conformitate bilateral dacă respingerea s-a făcut cu $\alpha_c = 0{,}01$ şi media eşantionului, m_1, este tocmai media populației gaussiene de abatere standard σ, din care acesta a fost extras aleator.

Rezolvare:

Notând cu μ_1 media populației din care s-a extras aleator eşantionul cu media m_1, s-a testat cuplul de ipoteze: H_0: $\mu_1 = \mu_0$; H_A: $\mu_1 \neq \mu_0$. Deoarece $\alpha_c = 0{,}01 < \alpha = 0{,}05$, s-a respins ipoteza nulă. (Vezi figura alăturată în care media de eşantionaj, m, se distribuie gaussian, deoarece distribuția în populație este gaussiană şi cu σ cunoscut).



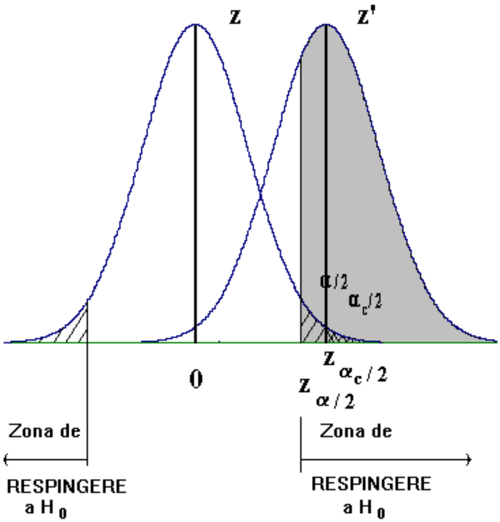
Dacă vom repeta, de multe ori, testarea ipotezei (nule) că media populației este egală μ_0 şi, de fapt, media populației, μ_1, este tocmai media eşantionului extras prima oară, m_1 (adică $\mu_1 = m_1 \neq \mu_0$) atunci media de eşantionaj va fi variabila aleatoare *m’* cu media în μ_1.



În această situație, ori de câte ori media unui eşantion va fi mai mare decât *x*, vom respinge ipoteza nulă. Deci proporția de respingeri este dată de proporția ariei de la dreapta lui *x* (aria gri) din aria de sub distribuția *m’*.

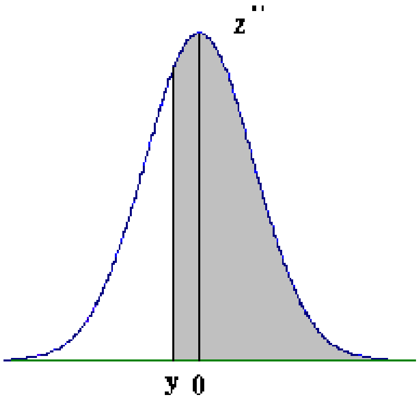
Să calculăm în continuare această proporție. Pentru aceasta, vom aplica transformarea $z = \dfrac{m - \mu_0}{\sigma / \sqrt{n}}$ variabilelor *m* şi *m’*. Variabila *m* va deveni distribuția normală standard.

În schimb, *m'* va deveni o distribuție normală standard deplasată, $z' = \dfrac{m' - \mu_0}{\sigma/\sqrt{n}}$ (vezi figura alăturată), deoarece aplicând translația $m' - \mu_0$, în loc de $m' - \mu_1$, nu s-a mai obținut centrarea. S-a obținut doar reducerea, prin comprimarea cu $\sigma/\sqrt{n}$. Prin standardizare μ_1 va deveni α-cuantila bilaterală superioară $z_{\alpha_c/2}$ deoarece lasă la dreapta sa aria $\alpha_c/2$ (ca şi μ_1), iar punctul de separare a zonei de respingere, notat *x* în figurile anterioare, va deveni, din acelaşi motiv, $z_{\alpha/2}$. Astfel, proporția căutată a devenit proporția reprezentată de aria gri din aria de sub curba *z'*.



Această proporție se poate determina după ce translatăm în origine curba *z*’. Vom obține curba *z*’’ = *z*’ - $z_{\alpha_c/2}$.

Evident $z_{\alpha_c/2}$ s-a deplasat în origine, iar $z_{\alpha/2}$ a devenit punctul $z_{\alpha/2}$ - $z_{\alpha_c/2}$ pe care îl vom nota *y*.



În consecință, va trebui să determinăm aria relativă aflată la dreapta acestui punct *y* (= $z_{\alpha/2}$ - $z_{\alpha_c/2}$), aria gri din figura de deasupra. În problema noastră, deoarece α = 0,05, $z_{\alpha/2}$ = 1,96 şi, deoarece α_c = 0,01, $z_{\alpha_c/2}$ = 2,58. Deci *y* = 1,96 - 2,58 = - 0,62.. În anexa 2 citim (pe linia lui – 0,6 şi coloana lui – 0,02) aria relativă aflată la stânga punctului – 0,62. Aceasta este 0,2676. Prin urmare aria la dreapta punctului *y* este complementul față de 1 al valorii 0,2676. Deci aria căutată are valoarea 1 – 0,2676 = 0,7324. În concluzie, probabilitatea repetării respingerii ipotezei nule, este, în acest caz, de circa 73%.

3.

Să se calculeze probabilitatea repetării respingerii ipotezei nule de către un test de conformitate bilateral dacă respingerea s-a făcut cu α_c = 0,05 şi media eşantionului, m_1, este tocmai media populației gaussiene de abatere standard σ, din care acesta a fost extras aleator.

Rezolvare:
Utilizăm rezolvarea generală din problema anterioară. $z_{\alpha/2}$ = $z_{\alpha_c/2}$= 1,96 pentru α = α_c = 0,05. Deci – *y* = $z_{\alpha_c/2}$ - $z_{\alpha/2}$ = 1,96 – 1,96 = 0.. Consultând Anexa 1 pe linia lui 0 şi coloana lui 0 găsim valoarea 0,5. În concluzie probabilitatea repetării respingerii ipotezei nule, este, în acest caz, de 50%.
Observație: Acest rezultat explică de ce atunci când efectuăm o respingere a ipotezei nule doar cu un *p* < 0,05, repetarea experimentului ne va conduce la repetarea respingerii doar în circa 50% din cazuri.

Exerciții sau probleme propuse

4.

Se specifică în rețeta unui anumit produs alimentar că pentru cacao concentrația minimă este de 10 g în fiecare unitate din acest produs. S-a verificat concentrația de cacao pe o linie de 50 de unități extrase aleator dintr-un lot cu acest produs şi s-a obținut media 9,6 g. Considerând că abaterea standard de 1 g specificată în rețetă este corectă, se cere să se testeze dacă lotul verificat prin acest sondaj conține cacao sub limita declarată.

5.

Pentru o specie de țestoase se cunoaşte din literatură că lungimea carapacei la indivizii adulți este în medie de 25 cm cu o abatere standard de 3 cm. Dintr-o populație de țestoase s-a extras un eşantion aleator simplu de 20 exemplare, având media 23,5 cm. Se cere să se testeze dacă populația biologică investigată este, mai degrabă, dintr-o altă specie, sau putem considera că este formată din indivizi dintr-o specie cu lungimea medie a carapacei de 25 cm, eventual specia citată în literatură?

6.

Să se calculeze probabilitatea repetării respingerii ipotezei nule de către un test de conformitate bilateral dacă respingerea s-a făcut cu α_c = 0,001 şi media eşantionului, m_1, este tocmai media populației gaussiene de abatere standard σ, din care acesta a fost extras aleator.

7.

Să se centralizeze într-un tabel probabilitățile repetării respingerii ipotezei nule de către un test de conformitate bilateral dacă respingerea s-a făcut cu una din valorile α_c convenționale (5%, 1%, 1‰) şi valoarea 0,1‰, media primului eşantion, m_1, fiind tocmai media populației gaussiene de abatere standard σ, din care acesta a fost extras aleator.