Probabilităţi şi Statistică - Curs 12

18 mai 2026

Table of contents

Teste de semnificaţie

Paşii care trebuie urmaţi pentru a întreprinde un test de semnificaţie:

1-2. Se formulează cele două ipoteze: H_0 şi H_a: H_a va fi acceptată dacă H_0 este respinsă.
3. Se alege un nivel de semnificaţie α - cât de semnificative trebuie să fie evidenţele pentru a respinge H_0.
4. Se calculează statistica sau scorul testului.
5. Se determină valoarea critică.
6. Se compară scorul cu valoarea critică şi, dacă este cazul, se respinge H_0 şi se acceptă H_a, altfel nu se acceptă H_a.

Testul Z

- Testul Z este un test statistic asupra ipotezelor utilizat pentru statistici care urmează o distribuţie normală atunci când ipoteza nulă este adevărată.
- Datorită Teoremei Limite Centrale (TLC) putem folosi testul Z şi atunci când populaţia este aproximativ normală, dar doar pentru eşantioane mari ($n \geqslant 30$).
- Am utilizat deja un test de tip Z în cazul testului proporţiei(folosind teorema de Moivre-Laplace care e o consecinţă a TLC).
- Testul Z se bazează pe distribuţia normală; pentru eşantioane mici, acest test de semnificaţie poate fi utilizat atunci când eşantionul provine dintr-o populaţie normală sau foarte aproape de una normală.

Testul Z - Inferenţă asupra mediei unei populaţii (σ cunoscută)

- Vom considera o populaţie statistică a cărei dispersie (σ^2) este cunoscută.
- Populaţia este normal distribuită şi dorim să testăm o ipoteză asupra mediei populaţiei.
- Testul poate fi întreprins chiar dacă populaţia nu este distribuită normal, dacă utilizăm eşantioane destul de mari.
- Dacă μ_0 este media populaţiei (presupusă în ipoteza nulă), atunci următoarea statistică $\dfrac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}}$ este distribuită normal standard: $N(0,1)$.
- Testul Z are loc astfel:

Testul Z - Inferenţă asupra mediei unei populaţii (σ cunoscută)

1. Formulăm *ipoteza nulă*, care susţine că media populaţiei ia o anumită valoare:

$$\boxed{H_0 : \mu = \mu_0}$$

2. Formulăm *ipoteza alternativă* conform informaţiilor obţinute din eşantion. Putem avea trei tipuri de ipoteză alternativă

$$\boxed{H_a : \mu < \mu_0} \quad (\textit{asimetrică la stânga}) \text{ sau}$$

$$\boxed{H_a : \mu > \mu_0} \quad (\textit{asimetrică la dreapta}) \text{ sau}$$

$$\boxed{H_a : \mu \neq \mu_0} \quad (\textit{simetrică}).$$

Ipotezele asimetrice se mai numesc *one-tailed*, iar cea simetrică *two-tailed*.

Testul Z - Inferenţă asupra mediei unei populaţii (σ cunoscută)

3. Alegem un nivel de semnificaţie $\alpha \in \{1\%, 5\%\}$.
4. Calculăm *scorul z* (*statistica* testului)

$$\boxed{z = \frac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}}}$$

5. Determinăm valoarea critică corespunzătoare lui α:

$\boxed{z^* = qnorm(\alpha)}$ pentru H_a asimetrică la stânga ($z^* < 0$),

$\boxed{z^* = qnorm(1-\alpha)}$ pentru H_a asimetrică la dreapta ($z^* > 0$),

$\boxed{z^* = -qnorm(\alpha/2) = qnorm(1-\alpha/2)}$ pentru H_a simetrică ($z^* > 0$).

Testul Z - Inferenţă asupra mediei unei populaţii (σ cunoscută)

6. Comparăm valoarea critică cu scorul z; dacă scorul z aparţine *zonei de respingere*, atunci acceptăm H_a şi respingem H_0. Zonele de respingere sunt:

$$(-\infty, z^*] \text{ pentru } H_a \text{ asimetrică la stânga,}$$

$$[z^*, +\infty) \text{ pentru } H_a \text{ asimetrică la dreapta,}$$

$$(-\infty, -|z^*|] \cup [|z^*|, +\infty) \text{ pentru } H_a \text{ simetrică.}$$

Dacă scorul z nu aparţine zonei de respingere spunem că *nu există suficiente dovezi cu nivelul de semnificaţie α* pentru a respinge ipoteza nulă (*încercarea de a respinge H_0 eşuează*).

Testul Z - Inferenţă asupra mediei unei populaţii - Exemplu

Exemplu.

- O colonie de şoareci de laborator constă din câteva mii de animale. Greutatea lor urmează o lege normală cu deviaţia standard $\sigma = 5g$ şi se crede că greutatea lor medie este $30g$.
- Pentru un eşantion de 25 de şoareci se găseşte o medie de $32g$; este aceasta valoare semnificativă statistic (cu 5% nivel de semnificaţie)? dar cu 1% nivel de semnificaţie?

Soluţie.

- Se pare că adevărata medie de greutate a întregii populaţii este diferită de cea afirmată ($\mu_0 = 30g$).
- Ştiind că populaţia urmează o distribuţie normală şi că deviaţia standard a populaţiei este cunoscută putem întreprinde un test Z asupra mediei.

Testul Z - Inferenţă asupra mediei unei populaţii - Exemplu

- Colectăm informaţiile legate de populaţie şi de eşantion: $\mu_0 = 30$, $\sigma = 5$, $n = 25$, $\overline{x}_n = 32$.

1-2. Formulăm ipoteza nulă şi o ipoteză alternativă simetrică

$$H_0 : \mu = 30 \quad H_a : \mu \neq 30.$$

3. $\alpha = 0.05$.
4. Scorul z

$$z = \frac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}} = \frac{32 - 30}{5/\sqrt{25}} = 2.$$

5. Valoarea critică este $z^* = -qnorm(\alpha/2) = 1.9599$, pentru $\alpha = 5\%$.
6. Cum $|z| > |z^*|$, putem respinge ipoteza nulă, şi să acceptăm că adevărata medie a populaţiei nu este $\mu_0 = 30g$.

Testul Z - Inferenţă asupra mediei unei populaţii - Exemplu

- Reluăm ultimii doi paşi pentru celălalt nivel de semnificaţie:

5'. Pentru $\alpha = 1\%$ valoarea critică este $z^* = -qnorm(\alpha/2) = 2.5758$.

6'. Cum $|z| < |z^*|$, încercarea de a respinge ipoteza nulă eşuează cu 1% nivel de semnificaţie (nu există suficiente dovezi pentru nivelul de 1% pentru a susţine că adevărata medie a populaţiei este diferită de $30g$).

Testul Z - Inferenţă asupra mediei unei populaţii - Exemplu revăzut

- Privind din nou informaţiile din eşantion putem observa că media de selecţie, $\overline{x}_n$, este mai mare decât media presupusă a populaţiei.
- Într-un astfel de caz am putea formula o ipoteză alternativă asimetrică la dreapta.

1-2. Noile ipoteze sunt

$$H_0 : \mu = 30 \quad H_a : \mu > 30.$$

3. $\alpha = 0.05$.
4. Scorul z

$$z = \frac{\overline{x}_n - \mu_0}{\sigma/\sqrt{n}} = \frac{32 - 30}{5/\sqrt{25}} = 2.$$

Testul Z - Inferenţă asupra mediei unei populaţii - Exemplu revăzut

5. Valoarea critică: $z^* = qnorm(1-\alpha) = 1.6448$, pentru $\alpha = 5\%$.
6. Deoarece $z > z^*$, putem respinge ipoteza nulă şi să acceptăm că media adevărată a populaţiei este mai mare decât $\mu_0 = 30g$.

- Pentru celălalt nivel de semnificaţie (1%):

5'. Valoarea critică este $z^* = qnorm(1-\alpha) = 2.3263$.

6'. Deoarece $z < z^*$, nu putem respinge ipotez nulă cu 1% nivel de semnificaţie (nu există dovezi suficiente cu 1% nivel de semnificaţie pentru a susţine că adevărata media a populaţiei este mai mare decât $30g$).

Testul Z - Inferenţă asupra mediei unei populaţii - Observaţii

- Merită observat că pentru diverse niveluri de semnificaţie putem avea concluzii diferite: ipoteza nulă poate fi respinsă pentru un anumit nivel de semnificaţie iar cu un alt nivel încercarea poate eşua.
- Dacă ipoteza nulă este respinsă cu 1% nivel de semnificaţie, atunci va fi respinsă şi cu 5%; altfel spus, dacă ipoteza nulă nu poate fi respinsă pentru 5%, nu va putea fi respinsă nici pentru 1%.
- Ipoteza alternativă trebuie formulată conform datelor din eşantion: când $\mu_0 << \overline{x}_n$ putem formula o ipoteză alternativă asimetrică la dreapta, $H_a : \mu > \mu_0$, când $\mu_0 >> \overline{x}_n$ putem formula o ipoteză alternativă asimetrică la stânga, $H_a : \mu < \mu_0$.
- Dacă media de selecţie nu este nici atât de mare şi nici atât de mică prin comparaţie cu media presupusă a populaţiei, putem presupune că media este doar diferită de valoarea din ipoteza nulă, $H_a : \mu \neq \mu_0$.

Testul Z - Exerciţii

I. Reluaţi testul Z pentru exerciţiul anterior cu $\overline{x}_n = 27g$ şi $\sigma = 6$. (Folosiţi ambele nivele de semnificaţie.)

II. Se afirmă că studenţii unei anumite universităţi vor obţine la un anumit test o medie de 35 de puncte cu $\sigma = 4$. Are această afirmaţie susţinere ştiind că pentru un eşantion aleator se obţin următoarele rezultate la test: $33, 42, 38, 37, 30, 42$? Întreprindeţi un test corespunzător pentru $\alpha = 5\%$. Se presupune că rezultatele la acest test sunt normal distribuite.

III. Conform National Center for Health Statistics, înălţimea medie a femeilor din SUA (care este normal distribuită) este de 63.7 in cu o deviaţie standard cunoscută de $\sigma = 2.75$ in. Pentru un eşantion aleator simplu de 50 femei care lucrează în domeniul serviciilor medicale se găseşte o înăţime de 65.2. Testaţi ipoteza că media de înăţime a femeilor care lucrează în sănătate este diferită de 63.7 in. 5% nivel de semnificaţie.

Testul T

- Testul T este folosit pentru ipotezele statistice pentru statistici care urmează o distribuţie Student.
- Testul T se foloseşte atunci când deviaţia standard a populaţiei (distribuită normal) este necunoscută.
- Un test T este de asemeni potrivit când avem de-a face cu eşantioane mici ($n < 30$) pentru populaţii care sunt aproximativ normale.
- Urmând aceste observaţii putem spune că un test T pentru media unei populaţii este complementar unui test Z pentru medie.
- În secţiunea următoare descriem testul T pentru media unei populaţii cu dispersia necunoscută.

Testul T - Inferenţă asupra mediei unei populaţii (σ necunoscută)

- Considerăm o populaţie statistică a cărei dispersie (σ^2) este necunoscută.
- Populaţia este normal distribuită şi dorim să testăm o ipoteză asupra adevăratei medii a populaţiei.
- Testul poate fi întreprins dacă populaţia are o distribuţie foarte apropiată de cea normală, când eşantionul la dispoziţie este mic.
- Dacă μ_0 este media populaţiei (presupusă în ipoteza nulă), atunci următoarea statistică $\dfrac{\overline{x}_n - \mu_0}{s/\sqrt{n}}$ este distribuită Student cu $(n-1)$ grade de libertate: $T(n-1)$.
- O diferenţă faţă de un test Z constă în înlocuirea deviaţiei standard a populaţiei, σ, cu deviaţia standard a eşantionului s.
- Testul T decurge astfel:

Testul T - Inferenţă asupra mediei unei populaţii (σ necunoscută)

1. Formulăm mai întâi *ipoteza nulă*, care susţine că media populaţiei ia o anumită valoare:

$$\boxed{H_0 : \ \mu = \mu_0}$$

2. Formulăm *ipoteza alternativă* conform informaţiilor obţinute din eşantion. Putem avea trei tipuri de ipoteză alternativă

$$\boxed{H_a : \ \mu < \mu_0} \quad (\textit{asimetrică la stânga}) \text{ sau}$$

$$\boxed{H_a : \ \mu > \mu_0} \quad (\textit{asimetrică la dreapta}) \text{ sau}$$

$$\boxed{H_a : \ \mu \neq \mu_0} \quad (\textit{simetrică}).$$

Ipotezele asimetrice se mai numesc *one-tailed*, iar cea simetrică *two-tailed*.

Testul T - Inferenţă asupra mediei unei populaţii (σ necunoscută)

3. Alegem un nivel de semnificaţie $\alpha \in \{1\%, 5\%\}$.
4. Calculăm *scorul t* (*statistica* testului)

$$\boxed{t = \frac{\overline{x}_n - \mu_0}{s/\sqrt{n}}}$$

5. Determinăm valoarea critică corespunzătoare lui α:

$\boxed{t^* = qt(\alpha, n-1)}$ pentru H_a asimetrică la stânga ($t^* < 0$),

$\boxed{t^* = qt(1-\alpha, n-1)}$ pentru H_a asimetrică la dreapta ($t^* > 0$),

$\boxed{t^* = -qt(\alpha/2, n-1) = qt(1-\alpha/2, n-1)}$ pt. H_a simetrică ($t^* > 0$).

Testul T - Inferenţă asupra mediei unei populaţii (σ necunoscută)

6. Comparăm valoarea critică cu scorul t; dacă scorul t aparţine *zonei de respingere*, atunci acceptăm H_a şi respingem H_0. Zonele de respingere sunt:

$$(-\infty, t^*] \text{ pentru } H_a \text{ asimetrică la stânga,}$$

$$[t^*, +\infty) \text{ pentru } H_a \text{ asimetrică la dreapta,}$$

$$(-\infty, -|t^*|] \cup [|t^*|, +\infty) \text{ pentru } H_a \text{ simetrică.}$$

Dacă scorul t nu aparţine zonei de respingere vom spune că *nu există suficiente dovezi cu nivelul de semnificaţie α* pentru a respinge ipoteza nulă (*încercarea de a respinge H_0 eşuează*).

Testul T - Inferenţă asupra mediei unei populaţii - Exemplu

Exemplu

- Concentraţia de CO (monixid de carbon) se măsoară cu un aparat numit spectrofotometru care are o precizie de aporape 100 ppm. Aceste aparate trebuie calibrate zilnic măsurând concentraţia de CO din eşantioane de gaz industrial care au o concentraţie controlată de 70 ppm. Dacă aparatul dă valori "apropiate" de 70 ppm poate fi folosit, dacă nu, trebuie ajustat.
- Presupunem că această concentraţie urmează o distribuţie normală dar deviaţia standard este necunoscută. Într-o anumită zi se obţin următoarele valori

$$58 \ 71 \ 67 \ 64 \ 62.$$

- Patru dintre aceste valori sunt mai mici decât 70; putem explica aceasta doar pe seama întâmplării? Sau aceasta arată că aparatul trebuie ajustat?

Testul T - Inferenţă asupra mediei unei populaţii - Exemplu

Solution

- Este posibil ca aparatul să necesite o ajustare, deci vom testa ipoteza că media este diferită de $\mu_0 = 70$.
- Cum populaţia urmează o distribuţie normală, iar deviaţia standard a populaţiei este necunoscută putem întreprinde un test T asupra mediei.
- Informaţiile privind populaţia şi eşantionul: $\mu_0 = 70$, $s = 4.9295$, $n = 5$, $\overline{x}_n = 64.4$.

1-2. Putem formula ipoteza nulă şi o ipoteză alternativă simetrică

$$H_0 : \mu = 70 \quad H_a : \mu \neq 70.$$

Testul T - Inferenţă asupra mediei unei populaţii - Exemplu

3. $\alpha = 0.05$.
4. Scorul t
$$t = \frac{\overline{x}_n - \mu_0}{s/\sqrt{n}} = \frac{64.4 - 70}{4.9295/\sqrt{5}} = -2.5402.$$
5. Valoarea critică este $t^* = -qt(\alpha/2, 4) = 2.7764$, for $\alpha = 5\%$.
6. Cum $|t| < |t^*|$, încercarea de a respinge ipoteza nulă eşuează (nu există suficiente dovezi pentru nivelul de 5% pentru a susţine că adevărata medie a populaţiei este diferită de 70 ppm).

Testul T - Inferenţă asupra mediei unei populaţii - Exemplu

- Reluăm ultimii doi paşi pentru celălalt nivel de semnificaţie:

5'. Pentru $\alpha = 1\%$ valoarea critică este $t^* = -qt(\alpha/2, 4) = 4.6040$.

6'. Deoarece $|t| < |t^*|$, încercarea de a respinge ipoteza nulă eşuează (nu există suficiente dovezi pentru nivelul de 1% pentru a susţine că adevărata medie a populaţiei este diferită de 70 ppm).

- Merită observat că reluarea ultimilor doi paşi cu un α mai mic nu era necesară, deoarece, ştim deja, ipoteza nulă nu a putut fi respinsă cu 5% nivel de semnificaţie ($|t^*|$ creşte dacă α scade).

Testul T - Inferenţă asupra mediei unei populaţii - Exemplu revăzut

- Privind din nou la datele din eşantion putem observa că media de selecţie, $\overline{x}_n$, este mai mică decât media presupusă a populaţiei.
- Într-o astfel de situaţie putem formula o ipoteză alternativă asimetrică la stânga.

1-2. Noile ipoteze sunt

$$H_0 : \mu = 70 \quad H_a : \mu < 70.$$

3. $\alpha = 0.05$.
4. Scorul t

$$t = \frac{\overline{x}_n - \mu_0}{s/\sqrt{n}} = \frac{64.4 - 70}{4.9295/\sqrt{5}} = -2.5402.$$

Testul T - Inferenţă asupra mediei unei populaţii - Exemplu revăzut

5. Valoarea critică este $t^* = qt(\alpha, 4) = -2.1318$, pentru $\alpha = 5\%$.
6. Deoarece $t < t^*$, putem respinge ipoteza nulă şi să acceptăm că media adevărată a populaţiei este mai mică decât 70 ppm.

- Cu celălalt nivel de semnificaţie (1%):

5'. Valoarea critică este $t^* = qt(\alpha, 4) = -3.7469$.

6'. Deoarece $t > t^*$, nu putem respinge ipotez nulă cu 1% nivel de semnificaţie (nu există dovezi suficiente cu 1% nivel de semnificaţie pentru a susţine că adevărata media a populaţiei este mai mică decât 70 ppm.

Testul T - Exerciţii

I. O asociaţie studenţească susţine că un student călătoreşte în fiecare zi, în medie, 25 de minute pe drumul către universitate. Universitatea determină un eşantion aleator obţinut de la 32 de studenţi. Eşantionul are o medie de selecţie de 19.4 minute şi o deviaţie standard de 9.6 minute. Datele acestea sunt suficiente statistic pentru a respinge afirmaţiile asociaţiei? Folosiţi $\alpha = 0.01$. (Presupunem că timpul unei călătorii urmează o distribuţie normală.)

II. Se ştie că tinerii adulţi cheltuie săptămânal 40$ pentru mâncare de tip fast food. Un studiu bazat pe un eşantion aleator fromat din 1000 de tineri adulţi (Greenfield Online - USA Today Snapshot) determină o medie săptămânală de 35$ cheltuiţi pe fast food cu o deviaţie standard de 14.50$. Presupunând ca aceste cheltuieli urmează o distribuţie normală întreprindeţi un test statistic potrivit asupra mediei acestor cheltuieli pentru întreaga populaţie.

Testul T - Exerciţii

III. Locuinţele dintr-un oraş au o valoare medie de 88950\$. Se presupune că locuinţele din apropierea universităţii au un preţ mediu mai mare. Pentru a testa aceasta, se foloseşte un eşantion aleator simplu format din 12 case alese în apropierea universităţii. Media lor de selecţie este 92460\$, cu o deviaţie standard de 5200\$. Întreprindeţi un test statistic pe baza acestor date cu $\alpha = 5\%$. (Se presupune că preţurile caselor sunt normal distribuite.)

Testul χ^2

- Câteodată frecvenţele dintr-un eşantion nu se potrivesc cu cele teoretice (conform distribuţiei populaţiei). De exemplu când aruncăm un zar ne aşteptăm ca toate feţele să aibă aceleaşi şanse de a apărea, dar în practică frecvenţele pot fi diferite.
- Testul χ^2 are scopul de a măsura "diferenţa" dintre frecvenţele observate şi cele aşteptate pentru date care pot fi asimilate cu cele categorice.
- Mai precis testul χ^2 este utilizat pentru a găsi (dacă există) o diferenţă semnificativă statistic între datele aşteptate şi cele observate.
- Există mai multe teste care utilizează distribuţia χ^2; toate testele a căror statistică (scor) urmează o distribuţie χ^2 sunt numite teste χ^2.
- Un alt test de χ^2 este folosit pentru a decide dacă există o dependenţă (legătură) între două variabile categorice.

Testul χ^2 de concordanţă (goodness-of-fit)

- Presupunem că indivizii dintr-un anumit eşantion pot fi grupaţi în k categorii $C_1, C_2, \ldots, C_k$ cu frecvenţele (absolute) observate $o_1, o_2, \ldots, o_k$. Pe de altă parte se aşteaptă ca frecvenţele (absolute) teoretice să fie $e_1, e_2, \ldots, e_k$.
- Statistica χ^2 este

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - e_i)^2}{e_i} = \sum_{i=1}^{k} \frac{o_i^2}{e_i} - m, \tag{1}$$

unde m este frecvenţa totală (absolută): $m = \sum_{i=1}^{k} o_i$.

- Evident, $\chi^2 = 0$ dacă şi numai dacă frecvenţele (absolute) observate şi cele aşteptate (teoretice) sunt egale două câte două.
- Cu cât este mai mare valoarea statisticii χ^2, cu atât este mai mare diferenţa dintre frecvenţele observate şi cele teoretice.

Testul χ^2 de concordanţă (goodness-of-fit)

- Se consideră că statistica de mai sus urmează foarte fidel o distribuţie χ^2 dacă frecvenţele teoretice sunt egale cel puţin cu 5.
- Ca test de concordanţă (goodness-of-fit) testul χ^2 îşi propune să determine cât de bine o distribuţie teoretică se potriveşte cu una empirică (obţinută din eşantion). În acest caz $df = k - 1 - s$, dacă frecvenţele teoretice pot fi calculate estimând parametri ai s populaţiei din eşantion.
- $df = k - 1$ dacă frecvenţele aşteptate (teoretice) pot fi calculate fără a estima vreun parametru al populaţiei deoarece cunoscând $(n - 1)$ dintre frecvenţele teoretice şi frecvenţa totală, ultima frecvenţă poate fi dedusă.

Testul χ^2 de concordanţă (goodness-of-fit) - exemplu

- Zacariah Labby de la Universitatea din Chicago a repetat în 2009 următorul experiment al lui Walter Weldon (din 1894): a aruncat de $315,672$ ori un zar[1].
- Experimentul a fost făcut pentru a decide dacă feţele cu numerele 5 şi 6 au frecvenţe mai mari (deoarece un zar obişnuit are punctele gravate pe feţe).
- Rezultatele sunt mai jos:

	Feţe					
	1	2	3	4	5	6
Observate	$53,222$	$52,118$	$52,465$	$52,338$	$52,224$	$52,285$
Aşteptate	$52,612$	$52,612$	$52,612$	$52,612$	$52,612$	$52,612$

[1] De fapt a aruncat doisprezece zaruri de $26,306$ ori pentru a accelera experienţa.

Testul χ^2 de concordanţă (goodness-of-fit) - exemplu

- Întrebarea este dacă aceste rezulate oferă dovezi suficiente pentru a a dovedi că cele şase feţe nu sunt echiprobabile.

$$H_0: \text{ nu există vreo diferenţă între cele două distribuţii}$$

$$H_a: \text{ există diferenţe între cele două distribuţii}$$

- Folosind (1) statistica χ^2 este

$$\chi^2 = 24.73$$

- Valoarea critică pentru numărul de grade de libertate cerut ($df = 6 - 1 = 5$) este:

$$\chi^{2*} = 11.07 \text{ pentru } \alpha = 0.05 \text{ şi}$$

$$\chi^{2*} = 15.08 \text{ pentru } \alpha = 0.01.$$

Testul χ^2 de concordanţă (goodness-of-fit) - exemplu

- Ipoteza nulă este respinsă dacă $\chi^2 > \chi^{2*}$, ceea ce se întâmplă pentru ambele niveluri de semnificaţie. Astfel, putem trage concluzia că există o diferenţă: datele observate ne spun că feţele 1 şi 6 (care sunt opuse pe zar) apar cu frecvenţă mai mare (decât celelalte două perechi: $3-4$ şi $2-5$).
- Testul χ^2 este asimetric la dreapta: ipoteza nula (H_0) se respinge dacă scorul testului (statistica χ^2) est mai mare decât valoarea critică.
- Valoarea critică, pentru un nivel de semnificaţie α, poate fi calculată în R cu $\chi^{2*} = qchisq(1-\alpha, df)$.
- Este bine de ştiut: cu cât mai mare este numărul de grade de libertate, cu atât mai mai apropiate sunt χ^{2*} şi $\chi^2(df)$.
- O regulă empirică spune că df ar trebui să fie cel puţin 5 pentru a utiliza testul de concordanţă.

Testul χ^2 de concordanţă - distribuţia multinomială

- În exemplul de mai sus am testat dacă cele şase categorii sunt uniform distribuite (cu probabilitate $1/6$). Mai general putem testa dacă datele observate (care sunt valori ale unor variabile aleatoare) aparţin unor anumite clase de distribuţie.
- Considerăm o experienţă aleatoare care are $k \geqslant 2$ rezultate posibile: $A_1, \ldots, A_k$, cu probabilităţi cunoscute, $P(A_i) = p_i$, $i = \overline{1,k}$. Efectuăm experienţa în mod independent de m ori.
- Fie X_i variabila aleatoare care numără de câte ori apare A_i. Variabilele X_i are nu sunt independente, de fapt $\sum_{i=1}^{k} X_i = m$.
- Vectorul $X = (X_1, X_2, \ldots, X_k)$ urmează o distribuţie multinomială cu parametrii $n, p_1, \ldots, p_k$: $M(m, p_1, \ldots, p_k)$.
- Funcţia de masă de probabilitate a lui X este

$$P(X = (n_1, \ldots, n_k)) = \frac{m!}{n_1! \cdot \ldots \cdot n_k!} p_1^{n_1} \cdot \ldots \cdot p_k^{n_k}.$$

Testul χ^2 de concordanţă - distribuţia multinomială

- Testul multinomial este un test de semnificaţie statistic care inferează asupra probabilităţilor unei distribuţii multinomiale: $p_1, \ldots, p_k$.
- Considerăm ca mai sus o experienţă aleatoare cu $k \geqslant 2$ rezultate posibile: $A_1, \ldots, A_k$; probabilităţile necunoscute ale acestor rezultate sunt $\pi_1, \ldots, \pi_k$.
- Eşantionăm m indivizi pentru a obţine un eşantion aleator simplu; fie $o_1, \ldots, o_k$ frecvenţele (absolute) observate ale celor k rezultate posibile.
- Dorim să decidem dacă aceste observaţii sunt suficient de semnificative pentru a susţine că probabilităţile rezultatelor posibile sunt diferite de $(p_1, \ldots, p_k)$.

$$H_0 : (\pi_1, \ldots, \pi_k) = (p_1, \ldots, p_k)$$

$$H_a : (\pi_1, \ldots, \pi_k) \neq (p_1, \ldots, p_k)$$

Testul χ^2 de concordanţă - distribuţia multinomială

- Variabila care numără apariţiile lui A_i, X_i, este distribuită $B(m, \pi_i)$, deci are media $e_i = \mathbb{E}[X_i] = m \cdot \pi_i$, pentru $i = \overline{1, k}$.
- (1) devine

$$\chi^2 = \sum_{i=1}^{k} \frac{(o_i - m\pi_i)^2}{m\pi_i} = \sum_{i=1}^{k} \frac{o_i^2}{m\pi_i} - m, \tag{2}$$

deoarece $m = \sum_{i=1}^{k} m\pi_i$.

- Pentru restul testului calculăm valoarea critică $\chi^{2*} = \chi^{2*}(\alpha) = qchisq(1 - \alpha, k - 1)$ pentru nivelul respectiv de semnificaţie; dacă $\chi^2 > \chi^{2*}$ putem respinge ipoteza nulă (H_0) şi accepta ipoteza alternativă (H_a).

Testul χ^2 de concordanţă - distribuţia multinomială - exemplu

- Poziţiile de pe o ruletă[2] sunt împărţite în şapte culori: 6 roşii, 5 negre, 7 portocalii, 10 albastre, 8 galbene, 7 maro şi 2 verzi. Dacă ruleta este corect construită şansele ca bila să ajungă în orice slot sunt de 1/45.
- Pentrua teste dacă ruleta este corect construită este învârtită de 1000 de ori şi se obţin următoarele frecvenţe absolute (am calculat şi frecvenţele aşteptate).

	Culoarea slot-ului						
	roşu	negru	portocaliu	albastru	galben	maro	verde
observate	119	117	140	204	165	136	19
aşteptate	120	100	140	200	160	140	40

- Sunt aceste valori statistic semnificative pentru a susţine că ruleta nu este corect construită?

[2] Ipotetică.

Testul χ^2 de concordanţă - distribuţia multinomială - exemplu

- Întreprindem un test multinomial folosind testul χ^2 al lui Pearson:

$$H_0 :$$

$$(\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7) = (6/45, 5/45, 7/45, 10/45, 8/45, 7/45, 2/45)$$

$$H_a :$$

$$(\pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7) \neq (6/45, 5/45, 7/45, 10/45, 8/45, 7/45, 2/45)$$

- Statistica χ^2 calculată folosind (2) este 14.273.
- Valorile critice pentru 5% şi 1% nivel de semnificaţie sunt

$$\chi^{2*}_{0.05} = qchisq(0.95, 6) = 12.591,$$

$$\chi^{2*}_{0.01} = qchisq(0.99, 6) = 16.811.$$

- Cu 5% nivel de semnificaţie putem trage concluzia că ruleta nu este corect construită, dar cu 1% nu pute respinge ipoteza nulă.

Testulχ^2 pentru testarea independenţei statistice

- Pentru testarea independenţei statistice testul χ^2 inferează asupra independenţei a două variabile categorice. Valorile observate ale acestor două variabile sunt date într-un aşa numit tabel de contingenţă.

	Y				
X	o_{11}	o_{12}	$\ldots$	o_{1r}	$o_{1,}$
	o_{21}	o_{22}	$\ldots$	o_{2r}	$o_{2,}$
	$\vdots$	$\vdots$	$\ldots$	$\vdots$	
	o_{p1}	o_{p2}	$\ldots$	o_{pr}	$o_{p,}$
	$o_{,1}$	$o_{,2}$	$\ldots$	$o_{,r}$	m

- o_{ij} este numărul de observaţii aparţinând simultan categoriei i a lui X şi j a lui Y. Frecvenţel (absolute) aşteptate când cele două variabile sunt independente (în ipoteza nulă) sunt $e_{ij} = \dfrac{o_{i,} \cdot o_{,j}}{m}$,

Testulχ^2 pentru testarea independenţei statistice

unde

$$p_{i,}=\frac{o_{i,}}{m}=\sum_{j=1}^{r}\frac{o_{ij}}{m};\, p_{,j}=\frac{o_{,j}}{m}=\sum_{i=1}^{p}\frac{o_{ij}}{m}.$$

- Statistica testului este

$$\chi^2=\sum_{i=1}^{p}\sum_{j=1}^{r}\frac{(o_{ij}-e_{ij})^2}{e_{ij}}=m\sum_{i=1}^{p}\sum_{j=1}^{r}\frac{(o_{ij}/m-p_{i,}p_{,j})^2}{p_{i,}p_{,j}} \qquad (3)$$

- Numărul de grade de libertate este $df=(p-1)(r-1)$, dacă frecvenţele aşteptate pot fi calculate fără a estima vreun alt parametru al populaţiei (din eşantion).

Testulχ^2 pentru testarea independenţei statistice - exemplu

- Să presupunem că dorim să ştim dacă culoarea părului şi cea a ochilor sunt legate (dependente). În [Snee74] s-au găsit următoarele date

		Ochi				
		negri	albaştri	căprui	verzi	
Păr	negru	68	20	15	5	108
	şaten	119	84	54	29	286
	roşu	26	17	14	14	71
	blond	7	94	10	16	127
		220	215	93	64	592

- Întreprindem un test χ^2 (de tip Pearson) pentru testarea independenţei celor două variabile ("culoarea ochilor" şi "culoarea părului"):

$$H_0 : \text{ variabilele sunt independente}$$

$$H_a : \text{ variabilele sunt dependente}$$

Testulχ^2 pentru testarea independenţei statistice - exemplu

- În ipoteza nulă statistica (scorul) testului este $\chi^2 = 138.29$.
- Numărul de grade de libertate este $(4-1)(4-1) = 9$, astfel valorile critice pentru niveluri uzuale de semnificaţie sunt

$$\chi^{2*}_{0.05} = qchisq(0.95, 9) = 16.918,$$

$$\chi^{2*}_{0.01} = qchisq(0.99, 9) = 21.665.$$

- Deoarece scorul este mai mare decât valoarea critică (în ambele cazuri), putem respinge ipoteza nulă şi să acceptăm că cele două variabile sunt dependente pentru ambele niveluri de semnificaţie.
- Notă: atunci când respingem ipoteza nulă şi acceptăm ipoteza alternativă, deşi dependenţa există, nu putem spune nimic despre "relaţia" dintre cele două variabile.

Testul F - inferenţă asupra raportului dintre dispersii

- Procedura de inferenţă prezentată aici va fi un test de semnificaţie pentru deviaţiile standard (sau dispersiile) a două populaţii normale.
- Pentru acest test prezumţia de normalitate este foarte importantă.
- Alegem două eşantioane aleatoare independente (de dimensiuni n_1 şi n_2, respectiv) cu deviaţiile standard s_1 şi s_2.
- Considerăm că, în condiţile ipotezei nule, adevăratele deviaţii standard σ_1 şi σ_2 sunt egale.
- Următoarea statistică este distribuită Fisher

$$F = \frac{s_1^2}{s_2^2}.$$

Distribuţia Fisher $F(r_1, r_2)$

- Există o familie de distribuţii Fisher. Fiecare distribuţie F se identifică prin două numere de grade de libertate.
- Proprietăţi ale distribuţiei Fisher:
  - F are valori nenegative;
  - F este asimetrică.
- Pentru inferenţele din această secţiune numerele de grade de libertate sunt $r_1 = n_1 - 1$, $r_2 = n_2 - 1$.

Testul F - Inferenţu a asupra raportului dintre dispersii

- Testul se desfăşoară astfel:

1. Formulăm *ipoteza nulă*:

$$\boxed{H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1}$$

2. Formulăm *ipoteza alternativă* conform informaţiilor obţinute din eşantion. Putem avea două[3] tipuri de ipoteză alternativă

$$\boxed{H_a : \frac{\sigma_1}{\sigma_2} > 1}$$ (*asimetrică la dreapta*) pentru un *un test one-taile*

$$\boxed{H_a : \frac{\sigma_1}{\sigma_2} \neq 1}$$ (*simetrică*) pentru un *un test two-tailed*.

[3] Este recomandat ca dispersia mai mare sau care e de aşteptat sa fie mai mare să fie pusă la numărător.

Testul F - Inferenţă asupra raportului dintre dispersii

3. Alegem nivelul de semnificaţie $\alpha \in \{1\%, 5\%\}$.
4. Calculăm *scorul F* (*statistica* testului)

$$\boxed{F = \frac{s_1^2}{s_2^2}}$$

5. Determinăm valoarea critică corespunzătoare pentru α

$\boxed{F^* = qf(1-\alpha, n_1 - 1, n_2 - 1)}$ pentru H_a asimetrică la dreapta ,

$$\boxed{F_s^* = qf(\alpha/2, n_1 - 1, n_2 - 1), F_d^* = qf(1-\alpha/2, n_1 - 1, n_2 - 1)}$$

pentru H_a simetrică.

Testul F - Inferenţă asupra a două dispersii

6. Comparăm valoarea critică cu scorul F; dacă scorul F aparţine *zonei de respingere*, atunci acceptăm H_a şi respingem H_0. Zonele de respingere sunt:

$$[F^*, +\infty) \text{ pentru } H_a \text{ asimetrică la dreapta,}$$

$$(0, F_s^*] \cup [F_d^*, +\infty) \text{ pentru } H_a \text{ simetrică.}$$

Dacă scorul F nu aparţine zonei de respingere spunem că *nu există suficiente dovezi cu nivelul de semnificaţie α* pentru a respinge ipoteza nulă (*încercarea de a respinge H_0 eşuează*).

Testul F - Inferenţă asupra raportului dintre dispersii - Exemplu

Exemplu.

- O companie care îmbuteliază băuturi răcoritoare deţine o maşină care umple sticle de 16 uncii. Compania trebuie să controleze deviaţia standard σ (sau dispersia σ^2) cantităţii de băutură din sticle. O medie corectă a cantităţii de băutură nu garantează faptul ca maşina este bine calibrată. Dacă dispersia este prea mare multe sticle vor avea prea multă sau prea puţină băutură.
- Astfel, maşina trebuie să aibă o deviaţie standard (sau dispersie) cât mai mică posibil.
- Compania doreşte să decidă dacă să instaleze o maşină de îmbuteliat mai modernă şi mai rapidă.
- Una dintre îngrijorări este aceea că o viteză mai mare de îmbuteliere corespunde unei dispersii mai mari, iar o asemenea creştere nu este acceptabilă.

Testul F - Inferenţă asupra raportului dintre dispersii - Exemplu

- Producătorul noii maşini de îmbuteliat susţine că dispersia nu este mai mare decât cea a vechii maşini.
- Un eşantion de 25 de sticle îmbuteliate de maşina nouă are o dispersie de 0.0018, în timp ce un eşantion de 22 sticle îmbuteliate de maşina curentă are o dispersie de 0.0008.

Soluţie.

- Vom testa amândouă tipurile de ipoteză alternativă: mai întâi faptul că dispersiile celor două maşini sunt diferite, iar apoi că noua maşină are o dispersie mai mare.
- Datele relativ la cele două populaţii/eşantioane: $n_1 = 25$, $s_1^2 = 0.0018$, $n_2 = 22$ şi $s_2^2 = 0.0008$.

Testul F - Inferenţă asupra raportului dintre dispersii - Exemplu

1-2. Formulăm ipoteza nulă şi o ipoteză alternativă simetrică

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \quad H_a : \frac{\sigma_1^2}{\sigma_2^2} \neq 1.$$

3. Alegem $\alpha = 5\%$.
4. Calculăm scorul F al testului

$$F = \frac{s_1^2}{s_2^2} = 2.2500$$

5. Valorile critice sunt $F_s^* = qf(\alpha/2, n_1 - 1, n_2 - 1) = 0.4327$, $F_d^* = qf(1 - \alpha/2, n_1 - 1, n_2 - 1) = 2.3675$, pentru $\alpha = 5\%$.

Testul F - Inferenţă asupra raportului dintre dispersii - Exemplu

6. Deoarece $F \in [F_s^*, F_d^*]$, încercarea de a respinge ipoteza nulă eşuează cu 5% nivel de semnificaţie (nu există suficiente dovezi pentru nivelul de 5% pentru a susţine că dispersiile celor două maşini sunt diferite): diferenţele sunt datorate hazardului.

- Reluăm ultimii doi paşi pentru celălalt nivel de semnificaţie:

5'. Pentru $\alpha = 1\%$ valorile critice vor da un interval mai mare: $F_s^* = qf(\alpha/2, n_1 - 1, n_2 - 1) = 0.3294$, $F_d^* = qf(1 - \alpha/2, n_1 - 1, n_2 - 1) = 3.1473$.

6'. Concluzia este aceeaşi: nu putem respinge ipoteza nulă.

- La fel ca şi în cazul testelor Z şi T, dacă H_0 nu se respinge pentru $\alpha = 5\%$ nu se va respinge nici pentru $\alpha = 1\%$ (pentru că zona de respingere în al doilea caz o înglobează pe prima).

Testul F - Inferenţă asupra raportului dintre dispersii - Exemplu revizuit

- Deoarece dispersia primului eşantion este mai mare decât cea de-a doua vom întreprinde un al doilea test cu o ipoteză alternativă asimetrică la dreapta.

1-2. Noile ipoteze sunt

$$H_0 : \frac{\sigma_1^2}{\sigma_2^2} = 1 \quad H_a : \frac{\sigma_1^2}{\sigma_2^2} > 1.$$

3. Alegem $\alpha = 5\%$.
4. Scorul testului va fi acelaşi

$$F = \frac{s_1^2}{s_2^2} = 2.2500$$

5. Valoarea critică este $F^* = qf(1-\alpha, n_1 - 1, n_2 - 1) = 2.0540$, pentru $\alpha = 5\%$.

Testul F - Inferenţă asupra raportului dintre dispersii - Exemplu revizuit

6. Deoarece $F \in [F^*, +\infty)$, putem să respingem ipoteza nulă şi să acceptăm ipoteza alternativă: dispersia noii maşini este mai mare decât a celei vechi.

- Pentru $\alpha = 1\%$

5'. Valoarea critică este $F^* = qf(1-\alpha, n_1 - 1, n_2 - 1) = 2.8010$.

6'. Cum $F < F^*$, încercarea de a respinge ipoteza nulă eşuează cu 1% nivel de semnificaţie (nu există suficiente dovezi pentru nivelul de 1% pentru a susţine că maşina nouă are o dispersie mai mare).

Sfârşit.

Bibliography

- Freedman, D., R. Pisani, R. Purves, *Statistics*, W. W. Norton & Company, 4th edition, 2007.
- Johnson, R., P. Kuby, *Elementary Statistics*, Brooks/Cole, Cengage Learning, 11th edition, 2012.
- Shao, J., *Mathematical Statistics*, Springer Verlag, 1998.
- Spiegel, M. R., L. J. Stephens, *Theory and Problems of Statistics*, Schaum's Outline Series, McGraw Hill, 3rd edition, 1999.
- Snee, R. D., *Graphical Display of Two-Way Contingency Tables*, The American Statistician, 28 (1), pp 9–12, 1974.