

Cuprins

1 Statistică inferențială

- Estimarea parametrilor
 - Estimarea punctuală
 - Estimarea cu intervale
 - Intervale de încredere pentru medie

2 Teste de semnificație

- Testarea ipotezelor statistice
- Erori, nivel de semnificație și putere a testului
- Nivelul de semnificație și valoarea P
- Teste parametrice și neparametrice
- Testul proporției
- Teste unilaterale și bilaterale

3 Bibliografie

Statistică inferențială

- Statistica inferențială are scopul să trage concluzii relativ la o populație folosind rezultate ale teoriei probabilităților și statistici obținute din eșantioane.
- Fără utilizarea teoriei probabilităților (legea numerelor mari, teorema limită centrală ș. a.) am putea considera drept sistematic un comportament care este de fapt datorat hazardului sau, din contră, un comportament sistematic ar putea trece neobservat.
- Exemple de inferențe statistice: *intervale de încredere* pentru estimarea parametrilor sau *teste de semnificație*.
- Tehnicile de inferență au la bază distribuțiile eșantioanelor.

Estimarea parametrilor

- Distribuția unei anumite populații poate fi necunoscută în întregime; de aceea ne putem dori să aflăm cel puțin media, dispersia sau alți parametri ai ei.
- Acești parametri ai unei populații pot fi estimați folosind statistici calculate din eșantion.
- Există două tipuri de estimare: *estimare punctuală* și *estimare cu intervale*.

Definition 1

Estimarea punctuală a unui parametru constă în determinarea unui număr, de obicei valoarea unei statistici corespunzătoare, desemnat să estimeze acel parametru.

Estimarea parametrilor

Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică

Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică

Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică

Definition 2

Estimarea cu un interval a unui parametru constă în determinarea unui interval ale cărui limite sunt statistici calculate din eșantioane.

Nivelul de încredere $(1 - \alpha)$ este proporția acelor intervale (de încredere) care conțin parametrul estimat.

Un **interval de încredere** este un interval care are un anumit nivelul de încredere (prescris).

Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică

Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică

Probabilități și Statistică
 Probabilități și Statistică
 Probabilități și Statistică

Estimarea punctuală

- Exemple de estimatori punctuali sunt media de selecție \bar{x}_n , dispersia eșantionului s^2 , sau deviația standard a eșantionului s .
- Pentru un parametru dat putem avea mai mulți estimatori punctuali: media populație poate fi estimată prin media de selecție, mediană, mod.
- Se pot ridica anumite întrebări legate de calitatea estimatorilor punctuali.
- Cât de exact este un estimator (i.e., *încrederea*) - este în mod frecvent mai mare (*supra-estimator*) sau mai mic (*sub-estimator*) în raport cu parametrul estimat?
- Din acest punct de vedere se preferă estimatorii *nedepășăți*.
- Care este variabilitatea unui estimator punctual (văzut ca o variabilă aleatoare) - aceasta este *acuratețea*.

Caracteristici ale estimării punctuale

- De exemplu dispersia mediei de selecție, care este σ^2/n : cu cât este mai mare eșantionul cu atât mai mică este variabilitatea mediei de selecție.

Definition 3

Fie θ un anumit parametru al unei populații, x un eșantion (de dimensiune n) al acestei populații și $\hat{\theta}_n = \hat{\theta}_n(x)$ un estimator punctual al lui θ .

$\hat{\theta}_n$ este o **statistică nedeplasată** dacă $\mathbb{E}[\hat{\theta}_n] = \theta$. Altfel $\hat{\theta}_n$ este numită statistică **deplasată**.

Un estimator nedeplasat cu dispersie minimă, dacă există, se numește **statistică eficientă**.

Distribuția mediei de selecție

- Fie \mathcal{P} o populație formată din N indivizi ale căror valori ale atributului sunt a_1, a_2, \dots, a_N .
- Variabila care reprezintă atributul populației este notată cu X .
- Pentru această populație, deci și pentru X , media și dispersia sunt

$$\frac{1}{N} \sum_{i=1}^N a_i = \mu, \text{ respectiv } \frac{1}{N} \sum_{i=1}^N (a_i - \mu)^2 = \sigma^2.$$

- Pentru a estima μ și σ^2 folosim eșantioane de dimensiune n .

Distribuția mediei de selecție

- Pentru un eșantion dat, $\mathbf{x}^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$, media de selecție este

$$\bar{x}_n^{(k)} = \frac{1}{n} \sum_{j=1}^n x_j^{(k)}.$$

- Fiecare $x_j^{(k)}$ este o valoare a unei variabile aleatoare cu aceeași distribuție ca X .

Definition 4

Variabila aleatoare care are drept valori toate mediile de selecție posibile, $\bar{x}_n^{(k)}$, pentru eșantioane de dimensiune n , se numește distribuția mediei de selecție.

Distribuția mediei de selecție

- Se poate demonstra următorul rezultat

Theorem 1.1

Media și dispersia mediei de selecție, \bar{x}_n , sunt μ și σ^2/n :

$$\mathbb{E}[\bar{x}_n] = \mu, \text{Var}[\bar{x}_n] = \frac{\sigma^2}{n}.$$

În plus, pentru valori mari ale lui n (≥ 30), distribuția mediei de selecție este normală, i. e.

$$\bar{x}_n \sim N(\mu, \sigma^2/n).$$

Estimarea mediei și a dispersiei

- Bineînțeles, un estimator al mediei (necunoscute a) populației (μ) este media de selecție, \bar{x}_n .
- Un estimator pentru adevărata dispersie a populației (σ^2) este dispersia eșantionului s^2 - o statistică nedeplasată.
- Pentru deviația standard a mediei de selecție, σ/\sqrt{n} , un estimator este s/\sqrt{n} numită **eroarea standard a mediei** (the standard error of the mean - SEM).

Estimarea cu intervale

- Un estimator de tip interval al mediei populației poate fi obținut cu inegalitatea lui Cebâșev:

$$P(|\bar{x}_n - \mathbb{E}[\bar{x}_n]| \geq k \cdot \sqrt{\text{Var}[\bar{x}_n]}) \leq \frac{1}{k^2} \Leftrightarrow$$

$$P\left(|\bar{x}_n - \mu| \geq k \cdot \frac{\sigma}{\sqrt{n}}\right) \leq \frac{1}{k^2} \Leftrightarrow$$

$$P\left(|\bar{x}_n - \mu| < k \cdot \frac{\sigma}{\sqrt{n}}\right) \geq 1 - \frac{1}{k^2} \Leftrightarrow$$

$$P\left(\bar{x}_n - k \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x}_n + k \cdot \frac{\sigma}{\sqrt{n}}\right) \geq 1 - \frac{1}{k^2}.$$

- Astfel o estimare cu un interval a lui μ este

$$\left(\bar{x}_n - k \cdot \frac{\sigma}{\sqrt{n}} < \mu < \bar{x}_n + k \cdot \frac{\sigma}{\sqrt{n}}\right).$$

Intervale de încredere

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Definition 5

Un interval de încredere pentru un parametru θ cu nivel de încredere $(1 - \alpha)$ este definit cu ajutorul a două statistici L și U astfel ca

$$P(L \leq \theta \leq U) = (\geq) 1 - \alpha.$$

- L și U sunt de fapt variabile aleatoare ale căror valori sunt statistici: pentru diferite eșantioane au diferite valori.
- De obicei nivelul de încredere este o probabilitate mare, aproape de 1, cum ar fi 0.90, 0.95, sau 0.99 (care dau $\alpha \in \{0.10, 0.05, 0.01\}$).

Intervale de încredere pentru media unei populații cu dispersia cunoscută

- Să presupunem că avem un eșantion de dimensiune n și dorim să construim un interval de încredere pentru media μ cu un nivel de încredere $(1 - \alpha)$.
- Știm că media de selecție, \bar{x}_n , urmează (cu aproximație sau exact dacă populația este repartizată normal) o distribuție normală

$$N\left(\mu, \frac{\sigma^2}{n}\right).$$

- Prin standardizare variabila $Z = \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}}$ este distribuită $N(0, 1)$.
- Căutăm o valoare z^* , numită *valoare critică*, astfel ca o variabilă normal standard să acopere sub grafic o arie de $(1 - \alpha)$ pe intervalul centrat în medie și de lungime $2z^*$ deviații standard.

Intervale de încredere pentru media unei populații cu dispersia cunoscută

- Fie $Z : N(0, 1)$, valoarea critică este aleasă așa încât

$$P(-z^* \leq Z \leq z^*) = 1 - \alpha.$$

- Definiții echivalente ale lui z^* :

$$P(Z \leq -z^*) = \alpha/2 \text{ sau } P(Z \geq z^*) = \alpha/2.$$

- Notăm cu $\Phi(a) = P(Z \leq a)$, funcția de repartiție normală standard.
- Astfel $z^* = -\Phi^{-1}(\alpha/2)$ - valoare care poate fi găsită în tabele sau aproximată în pachetele de prelucrări statistice uzuale (cum ar fi R, MiniTab, SPSS etc).
- Odată ce am determinat valoarea critică, știm că

$$P\left(-z^* \leq \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} \leq z^*\right) = 1 - \alpha.$$

Intervale de încredere pentru media unei populații cu dispersia cunoscută

- Echivalent,

$$P\left(\bar{x}_n - z^* \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x}_n + z^* \frac{\sigma}{\sqrt{n}}\right) = 1 - \alpha.$$

- Am demonstrat astfel

Theorem 1.2

Un interval de încredere cu nivelul de încredere $(1 - \alpha)$ pentru pentru media unei populații cu dispersia cunoscută este

$$\left(\bar{x}_n - z^* \frac{\sigma}{\sqrt{n}}, \bar{x}_n + z^* \frac{\sigma}{\sqrt{n}}\right),$$

unde z^ este valoarea critică asociată cu $\alpha/2$. Mai mult, acest interval este exact pentru o populație distribuită normal și aproximativ altfel, când eșantionul este suficient de mare ($n \geq 30$).*

Intervale de încredere pentru media unei populații cu dispersia cunoscută

Nivel de încredere	$\alpha/2$	z^*
90%	0.05	1.645
95%	0.025	1.960
99%	0.005	2.576

- $z^* \frac{\sigma}{\sqrt{n}}$ se numește *eroarea marginală*.
- Lungimea unui interval de încredere pentru medie este $2z^* \frac{\sigma}{\sqrt{n}}$.
- Dacă dorim o lungime anume pentru acest interval, w , atunci ne trebuie un eșantion de dimensiune $n = \frac{(2z^* \sigma)^2}{w^2}$. În practică această valoare poate fi nerealistă (dacă n este prea mare).
- Pe măsură ce n crește lungimea intervalului (sau eroarea marginală) scade, ceea ce este teoretic util, dar poate fi nepractic.

Intervale de încredere pentru media unei populații cu dispersia cunoscută

- Să ne amintim că $\frac{\sigma}{\sqrt{n}}$ este deviația standard a mediei de selecție.

Un interval de încredere poate fi văzut astfel

$$\text{estimatedMean} \pm z^* \text{StdDevofMean}$$

- Exemplu.** Un anumit medicament este analizat măsurându-i-se substanța activă de trei ori, rezultatele sunt 0.8403, 0.8636, și 0.8447 g/l. Se știe că această concentrație urmează o lege normală cu deviația standard $\sigma = 0.0068$ g/l. Să se determine un interval de încredere de 99% pentru adevărata medie a concentrației, μ .

Soluție:

$$\bar{x}_3 = 0.8495, \alpha = 0.01, \alpha/2 = 0.005, z^* = 2.576, z^* \frac{\sigma}{\sqrt{n}} = 0.0101$$

Intervalul de încredere este (0.8394, 0.8596).

Intervale de încredere pentru media unei populații cu dispersia necunoscută

- Să ne amintim că atunci când dispersia populației, σ^2 , este cunoscută, un interval de încredere de nivel $(1 - \alpha)$ pentru media populației μ , este

$$\left(\bar{x}_n - z^* \frac{\sigma}{\sqrt{n}}, \bar{x}_n + z^* \frac{\sigma}{\sqrt{n}} \right),$$

- Dacă dispersia este necunoscută, atunci putem folosi ca estimator al deviației standard a mediei de selecție eroarea standard a mediei: s/\sqrt{n} .
- Obținem o nouă statistică

$$T = \frac{\bar{x}_n - \mu}{s/\sqrt{n}}$$

cunoscută drept **statistica t a lui Student**.

- Aceasta deoarece noua statistică urmează o distribuție Student cu $(n - 1)$ *grade de libertate*, $t(n - 1)$ - dacă populația urmează o lege normală.

Intervale de încredere pentru media unei populații cu dispersia necunoscută

- Fie $T : t(n - 1)$, valoarea critică, t^* , pentru nivelul de încredere $(1 - \alpha)$ se alege astfel încât
$$P(-t^* \leq T \leq t^*) = 1 - \alpha, P(T \leq -t^*) = \alpha/2 \text{ sau } P(T \geq t^*) = \alpha/2.$$

Proposition 1

Un interval de încredere cu nivelul de încredere $(1 - \alpha)$ pentru pentru media unei populații normale cu dispersia necunoscută este

$$\left(\bar{x}_n - t^* \frac{s}{\sqrt{n}}, \bar{x}_n + t^* \frac{s}{\sqrt{n}} \right).$$

unde t^ este valoarea critică asociată cu $\alpha/2$.*

Intervale de încredere pentru media unei populații cu dispersia necunoscută - Exemplu

Exemplu.

- Într-un oraș există 10000 locuințe închiriate. O companie locală de imobiliare întreprinde un studiu asupra acestor locuințe: sunt intervievați 250 chiriași aleși la întâmplare. Chiria medie este găsită a fi 568\$, iar deviația standard a eșantionului este 385\$. (Se știe că chiria urmează o distribuție normală.)
- Determinați a un interval de încredere de 95% pentru chiria medie a tuturor celor 10000 de locuințe închiriate

Soluție.

- Datele colectate

$$\bar{x}_{250} = 568, s = 385, \alpha = 0.05, \alpha/2 = 0.025,$$

$$t^* = 1.9695, t^* \frac{s}{\sqrt{n}} = 48.0535$$

- Intervalul de încredere este (519.9465, 616.0535).

Intervale de încredere pentru medie - Exerciții

- I. Pentru un sondaj de opinie se alege un eșantion aleator simplu de 400 de persoane de cel puțin 25 de ani dintr-un anumit oraș din Appalachia. Numărul total de ani de școlarizare al membrilor eșantionului este 4635, iar deviația standard a eșantionului este 4.1 ani. Determinați a un interval de încredere de 95% pentru numărul mediu de ani de școlarizare al tuturor persoanelor de cel puțin 25 de ani din oraș. (Presupunem că perioada de școlarizare urmează o lege normală.)
- II. Primăria unui oraș vrea să cunoască venitul mediu al celor 25000 de familii din oraș. Pentru aceasta angajează un institut de sondare de opinie care interoghează 1000 de familii alese aleator. Venitul total al acestor familii este de 62396714\$, deviația standard a eșantionului fiind 53000\$. Determinați un interval de încredere de 99% pentru venitul mediu al unei familii din acest oraș. (Presupunem că venitul unei familii urmează o lege normală.)

Intervale de încredere pentru medie - Exerciții

- III. O universitate are 30000 studenți; în vederea unui sondaj, 900 dintre acești studenți sunt aleși aleator. Vârsta medie a eșantionului este de 22.3 ani cu o deviație standard a eșantionului de 4.5 ani. Determinați intervale de încredere de 90% și 95% pentru vârsta medie a studenților acestei universități. (Presupunem că vârsta unui student urmează o lege normală.)

Testarea ipotezelor statistice

- Deciziile trebuie luate în fiecare zi; unele dintre ele sunt mai semnificative decât altele, dar mecanismul luării unei decizii urmează același șablon.
- Avem două sau mai multe alternative și trebuie să alegem una dintre ele pe baza evidenței/convingerilor/contextului/informațiilor etc.
- Un test statistic de semnificație urmează un același patern cu excepția faptului că decizia este luată folosind informația statistică.
- Aceasta înseamnă că în timpul acestui proces se vor calcula anumite statistici și pe baza acestora se vor lua deciziile.
- Primul pas este acela de a identifica o situație care are un anumit grad de incertitudine și să formulăm două *ipoteze* legate de ea.

Testarea ipotezelor statistice

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Definition 6

Testarea ipotezelor statistice este un proces prin care se ia o decizie (se alege) între două ipoteze opuse.

Ipotezele statistice sunt formulate în așa fel încât întotdeauna una din ele este falsă iar cealaltă adevărată.

Una dintre ipoteze este testată sperând că se poate arăta că este puțin probabil ca ea să fie adevărată, ceea ce implică faptul că cealaltă ipoteză este probabil adevărată.

- Cele două ipoteze sunt numite **ipoteza nulă** și **ipoteza alternativă**.
- Un test statistic încearcă să dovedească faptul că ipoteza nulă este falsă.

Testarea ipotezelor statistice

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Definition 7

Ipoteza nulă, H_0 , este ipoteza status-quo-ului în ceea ce privește populația; formal este o afirmație legată de populație: spre exemplu că are o anumită medie sau dispersie, sau o anumită distribuție etc.

Ipoteza alternativă, H_a , este ipoteza de cercetare, și susține un lucru diferit despre obiectul ipotezei nule.

- Ipoteza nulă este punctul de plecare al studiului și, în mod conservator, susține că "nu există nici o diferență" sau că "nimic nu se întâmplă" (are tendința de a se opune oricărei schimbări).
- Într-un anumit fel ipoteza alternativă are rolul de a ataca ipoteza nulă și observă o diferență acolo unde cea nulă nu găsește nimic.

Testarea ipotezelor statistice - Exemplul I

- O companie de produse lactate cumpără lapte de la mai mulți distribuitori. Există un dubiu asupra calității laptelui astfel cumpărat.
- Temperatura de îngheț a laptelui urmează o lege normală cu media $\mu_0 = -0.545^\circ C$ și deviația standard $\sigma = 0.008^\circ C$. Dacă se adaugă apă în lapte temperatura de îngheț crește.
- Se măsoară temperatura de îngheț a laptelui pentru cinci loturi diferite primite de la unul dintre distribuitori și se determină o temperatură medie de $\bar{x}_5 = -0.538^\circ C$. Este aceasta o dovadă ca distribuitorul adaugă apă în lapte? Sau este doar o valoare datorată hazardului?

Testarea ipotezelor statistice - Exemplul I

- Ipoteza de lucru, adică ipoteza nulă, este aceea că media populației este $\mu = \mu_0$.
- Ipoteza de cercetare, adică ipoteza alternativă, este că media populației este mai mare: $\mu > \mu_0$.
- Relativ la ipoteza alternativă o întrebare care se poate formula este: în condiții obișnuite care sunt șansele ca $\mu > \mu_0$?
- Fie X variabila aleatoare asociată temperaturii de îngheț a laptelui; $X : N(-0.545, (0.008)^2)$.

Testarea ipotezelor statistice - Exemplit I

● Arunci

$$P(X > -0.538) = P\left(\frac{X - (-0.545)}{0.008/\sqrt{5}} > \frac{(-0.538) - (-0.545)}{0.008/\sqrt{5}}\right) = \\ = P\left(\frac{X - (-0.545)}{0.008/\sqrt{5}} > 1.95655948\right).$$

- Această probabilitate este $P(Z > 1.95655948)$, unde $Z : N(0, 1)$.
- Pentru variabila aleatoare normală standard avem $P(Z > 1.95655948) \sim 0.0250$. Astfel, cu probabilitate 0.025, putem spune că diferența observată în eșantion se datorează unei variații aleatoare așteptate.

Testarea ipotezelor statistice

- Dacă, presupunând că ipoteza nulă este adevărată, găsim că o anumită statistică obținută dintr-un eșantion diferă mult de rezultatul așteptat, atunci spunem că diferența este *semnificativă*.
- În acest caz putem fi tentați să *respingem ipoteza nulă*.
- Dacă statistica nu diferă semnificativ de rezultatul așteptat în condițiile ipotezei nule, atunci spunem că *am eșuat în respingerea ipotezei nule*.
- *Testarea ipotezelor* este o procedură care ne permite să verificăm dacă anumite statistici diferă semnificativ de valoarea așteptată în condițiile ipotezei nule.

Testarea ipotezelor statistice

- Dacă decizia este de a respinge H_0 , atunci concluzia testului trebuie să fie: "*Există suficiente dovezi pentru nivelul de semnificație α pentru ca ...*" (și aici urmează afirmațiile din ipoteza alternativă).
- Dacă decizia este de a nu respinge H_0 , atunci concluzia testului trebuie să fie: "*Nu există suficiente dovezi pentru nivelul de semnificație α pentru ca ...*" (și aici urmează afirmațiile din ipoteza alternativă).

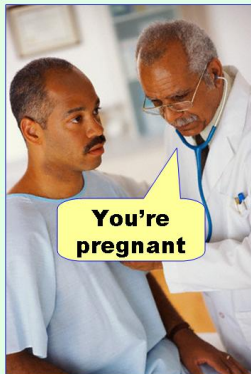
Erori de tipul I și de tipul II

- Evident că decizia aceasta, luată pe baza unor probabilități, poate fi greșită. Acesta este locul unde pot apărea erorile.
- Există două tipuri de erori: primul tip apare atunci când respingem o ipoteză care este adevărată, iar al doilea tip apare atunci când acceptăm o ipoteză care este falsă și ar trebui respinsă.

		validitatea ipotezei H_0	
		adevărată	falsă
Decizia asupra H_0	se respinge	Eroare de tip I (fals pozitiv)	Corect (adevărat pozitiv)
	nu se respinge	Corect (adevărat negativ)	Eroare de tip II (fals negativ)

Erori de tipul I și de tipul II

Type I error
(false positive)



Type II error
(false negative)



Figure: Erori de tip I și II.

Testarea ipotezelor statistice - Exemplul II

- Se bănuie că un detergent foarte cunoscut este mai bun decât detergentul vândut sub marca de casă a unui hipermarket și se dorește să se testeze calitatea celor două tipuri de detergent, deoarece un cumpărător va dori să cumpere calitate la un preț convenabil. Formulăm cele două ipoteze.
- Bănuiala, "Detergentul cunoscut este mai bun decât detergentul de casă al hipermarketului," este motivul pentru care se întreprinde testul și devine, deci, ipoteză alternativă.

H_0 : "Nu există nicio diferență între cele două tipuri de detergent."

H_a : "Detergentul cunoscut este mai bun decât celălalt."

- Testul este conceput în speranța de a respinge ipoteza nulă, dar pentru consumator speranța este de a nu respinge ipoteza nulă (din rațiuni bugetare).

Testarea ipotezelor statistice - Exemplul II

	H_0 este adevărată	H_0 este falsă
se respinge	<p>Eroare de tipul I</p> <p>Situația reală: Nu există diferență. Concluzie: Detergentul cunoscut e mai bun. Acțiune: Clientul plătește mult cu aceleași rezultate.</p>	<p>Decizie corectă</p> <p>Situația reală: Detergentul cunoscut e mai bun. Concluzie: Detergentul cunoscut e mai bun. Acțiune: Clientul plătește mult cu rezultate bune.</p>
nu se respinge	<p>Decizie corectă</p> <p>Situația reală: Nu există diferență. Concluzie: Nu există diferență. Acțiune: Clientul plătește puțin cu rezultate similare.</p>	<p>Eroare de tipul II</p> <p>Situația reală: Detergentul cunoscut e mai bun. Concluzie: Nu există diferență. Acțiune: Clientul plătește puțin cu rezultate slabe.</p>

Testarea ipotezelor statistice - Exemplit II

- Am descris cele patru tipuri de rezultate posibile și acțiunile core-spunzătoare pentru un test statistic.
- Situația de fapt (reală) nu este cunoscută înainte de a lua decizia, de a trage concluziile și de a întreprinde acțiunea. Adevărul despre ipoteza nulă, H_0 , poate rămâne necunoscut pentru totdeauna.
- Rezultatul unei erori de tipul II este adesea ceea ce se numește o "oportunitate pierdută"; se pierde șansa de a utiliza un produs care are o mai bună calitate (în cazul expus).

Dimensiunea efectului, nivelul de semnificație și puterea testului

- Regula după care se ia decizia trebuie să minimizeze erorile descrise.
- Exisă patru factori care influențează o decizie luată pe baza unui test statistic (ignorând metoda de eșantionare): *dimensiunea efectului*, *dimensiunea eșantionului*, *nivelul de semnificație* și *puterea testului*.

Definition 8

Dimensiunea efectului este magnitudinea diferenței descoperite în eșantionul aleator (dacă există).

Nivelul de semnificație, α , este probabilitatea (condiționată) maximă pe care ne-o asumăm drept risc de face o eroare de tipul I.

Puterea testului, este 1 minus probabilitatea de a face o eroare de tipul II.

Dimensiunea efectului, nivelul de semnificație și puterea testului

- Nivelul de semnificație este specificat de obicei înaintea eșantionării, astfel rezultatele testului nu vor influența eșantionarea.
- De obicei nivelul de semnificație este 0.05 sau 0.01; de exemplu 0.05 (sau 5% nivel de semnificație) este utilizat pentru o regula de luare a deciziei care dă cel mult 5 șanse din 100 de a respinge ipoteza nulă când ea este adevărată.
- Puterea testului este probabilitatea ca testul să detecteze o diferență atunci când există cu adevărat o asemenea diferență de detectat.
- Dacă puterea testului este mare, atunci probabilitatea de a face o eroare de tipul II (probabilitatea ca testul să nu detecteze o diferență atunci când există vreuna) este scăzută.

Nivelul de semnificație și valoarea P

- Există două moduri de a desfășura un test de semnificație: folosind *scorul* și *valoarea critică* sau folosind *valoarea P* .

Definition 9

Valoarea P este probabilitatea de a obține un rezultat cel puțin la fel de neobișnuit (extrem) ca rezultatul obținut din eșantion. Când valoarea P este "mică" putem respinge H_0 .

- Revenim la primul exemplu. Am calculat $P(X > -0.538) \sim 0.0250$; aceasta este valoarea P a testului.
- Dacă 2.5% este considerată o probabilitate mică, atunci putem respinge ipoteza nulă.

Nivelul de semnificație și valoarea P

- Dacă însă 2.5% nu este de ajuns de mică, atunci nu respingem H_0 .
- De exemplu $2.5\% < 5\%$, astfel, cu 5% nivel de semnificație H_0 este respinsă și este acceptată H_a .
- Cu 1% nivel de semnificație încercarea de a respinge ipoteza nulă eșuează, datele noastre nu sunt suficient de semnificative pentru a accepta ipoteza alternativă, H_a .

Nivelul de semnificație și valoarea P

Definition 10

Scorul testului este statistica ce corespunde valorii P ; **valoarea critică**, pe de altă parte, corespunde nivelului de semnificație. Concluzia testului depinde de rezultatul comparării celor două valori.

- Calculăm statistica

$$z = \frac{\bar{x}_n - \mu}{\sigma/\sqrt{n}} = \frac{(-0.538) - (-0.545)}{0.008/\sqrt{5}} = 1.95655948.$$

- z este scorul testului.
- Valoarea critică, z^* , pentru 1% nivel de semnificație îndeplinește condiția

$$P(Z > z^*) = 1 - \alpha = 0.99 \Rightarrow z^* = 2.32, \text{ unde } Z : N(0, 1).$$

Nivelul de semnificație și valoarea P

- Cu 1% nivel de semnificație $z < z^*$, deci H_0 nu poate fi respinsă
- Pentru 5% nivel de semnificație, valoarea critică, z^* , se calculează astfel ca
$$P(Z > z^*) = 1 - \alpha = 0.95 \Rightarrow z^* = 1.64,$$
- Cu 5% nivel de semnificație $z > z^*$, deci H_0 poate fi respinsă; acceptăm H_a .
- Acest exemplu arată că pentru diferite nivele de semnificație concluziile testului pot fi diferite.

Teste parametrice și neparametrice

- În exemplul anterior am presupus că populația urmează o distribuție normală.
- Câteodată este dificil de aflat dacă populația urmează o distribuție anume (normală cel mai adesea).
- Relativ la distribuția populației avem două tipuri de teste de semnificație: parametrice și neparametrice.

Definition 11

Un test parametric presupune că populația urmează o anumită distribuție și inferează asupra parametrilor acelei distribuții.

Teste parametrice și neparametrice

- Testele neparametrice inferează mai degrabă asupra distribuției decât asupra parametrilor populației.
- De obicei testele parametrice presupun că populația urmează o distribuție normală. (Unii autori consideră că orice altfel de test este neparametric).

Definition 12

Un test neparametric numit și distribution-free sau parameter-free se bazează pe puține fapte - de obicei distribuția și parametrii săi (medie, dispersie) nu sunt cunoscuți.

Testul proporției

- Una dintre cele mai utilizate inferențe este cea asupra *parametrului binomial* p , probabilitatea succesului.
- În multe situații ne interesează doar ceea ce se "realizează" sau "nu se realizează"; există doar două rezultate posibile - aceasta este condiția de bază a unui experiment binomial.
- Proporția indivizilor dintr-o populație care au o anumită trăsătură poate fi considerată probabilitatea succesului.
- Exemple: proporția indivizilor care fumează, a cetățenilor care votează cu un anumit politician, a șoferilor care accelerează la lumina galbenă a semaforului, a tinerilor care își întemeiază o familie etc.

Testul proporției

- Dacă $X : B(n, p)$ (numărul de succese), atunci parametrii populației sunt

$$\mu = np, \sigma^2 = np(1 - p).$$

- Dacă dimensiunea eșantionului este n și x este numărul de succese din eșantion, atunci **frecvența** sau **probabilitatea binomială a eșantionului** este

$$p' = \frac{x}{n}.$$

- Pentru $n \geq 20$ și $np \geq 5$, p' poate fi aproximată cu o distribuție normală.

Testul proporției

- Parametrii lui p' sunt

$$\mathbb{E}[p'] = \frac{\mathbb{E}[X]}{n} = p, D^2[p'] = \frac{D^2[X]}{n^2} = \frac{p(1-p)}{n}.$$

- Deci

$$p' \sim N\left(p, \frac{p(1-p)}{n}\right).$$

- Astfel următoarea statistică urmează o distribuție normală standard

$$z = \frac{p' - p}{\sqrt{p(1-p)/n}}.$$

- Aceasta va fi statistica sau scorul testului (unde $p' = x/n$).

Testul proporției - Exemplu

- Conform unui sondaj Harris din august 2008, 68% dintre americanii adulți au un abonament la o bibliotecă. Să presupunem că se alege un eșantion aleator de 1000 de adulți pentru a testa $H_0 : p = 0.68$ versus $H_a : p < 0.68$, unde p reprezintă proporția adulților care au într-adevăr abonament.
- 651 din cei 1000 de indivizi au un abonament la o bibliotecă. Folosiți $\alpha = 0.01$.
 - a. Calculați valoarea statisticii (scorul) testului.
 - b. Aplicați testul folosind valoarea P .
 - c. Aplicați testul folosind abordarea clasică (cu valoarea critică).

Testul proporției - Exemplu

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

Probabilități și Statistică

- Cu valoarea P :

$$P = P(Z < z | H_0) = P(Z < -1.965) = 0.0246$$

- Cu $\alpha = 1\%$ nu putem respinge H_0 , datele nu sunt semnificative pentru aceasta.

- Cu valoarea critică:

$$z^* = -2.326, (P(Z < z^*) = 0.01).$$

- Deoarece $z \not< z^*$ nu putem respinge H_0 , datele nu sunt suficient de semnificative pentru a accepta ipoteza alternativă.

Testul proporției - Exemplu

- Reluăm acum exercițiul cu $\alpha = 5\%$.
- Cu P -value: $P = 0.0246 < 0.05$, deci putem respinge ipoteza nulă și să acceptăm H_a , adică faptul că proporția adulților care au abonament la bibliotecă este mai mic decât 0.68.
- Cu valoarea critică: $z^* = -1.644 > z$ cea ce înseamnă că H_0 poate fi respinsă.

Testul proporției - Exerciții

- I. O monedă este aruncată de 10000 de ori și de 5167 ori apare stema. Probabilitatea de a apărea stema este 0.5? Sau dimpotrivă, este semnificativ numărul mare de apariții ale stemei? Aplicați un test statistic corespunzător pentru a răspunde acestor întrebări.
- II. Rezolvați din nou exercițiul I când stema apare de 5067 ori.
- III. În 2009, într-un articol din USA Today se afirma că 58% dintre adulții americani accelerează la lumina galbenă a semaforului. Într-un oraș se desfășoară un sondaj cu 150 de adulți aleși aleator; se află că 71 dintre aceștia recunosc că accelerează la lumina galbenă a semaforului. Se poate trage concluzia că în acest oraș există o rată mai mică decât aceea de la nivel național? Folosiți 0.05 nivel de semnificație.

Teste unilaterale și bilaterale

- Există trei tipuri de ipoteze alternative relativ la probabilitatea binomială p .

$$H_a : p < p_0, H_a : p > p_0 \text{ sau } H_a : p \neq p_0.$$

- Primele două se numesc ipoteze alternative **unilaterale** sau ("one-tailed"), iar a treia este ipoteză alternativă **bilaterală** sau "two-tailed").
- Pentru fiecare din aceste cazuri valoarea critică se calculează în mod diferit.

Testul unilateral la proporției

- Pentru $H_0 : p = p_0$ and $H_a : p > p_0$ - *unilateral sau asimetric la dreapta - right tail.*

valoarea $P : P(Z > z)$,

valoarea critică $z^* > 0$, a. î. $P(Z > z^*) = \alpha = 1 - P(Z < -z^*)$.

în R $z^* = qnorm(1 - \alpha)$.

- Pentru $H_0 : p = p_0$ and $H_a : p < p_0$ - *unilateral sau asimetric la stânga - left tail.*

valoarea $P : P(Z < z)$,

valoarea critică $z^* < 0$, a. î. $P(Z < z^*) = \alpha = 1 - P(Z > -z^*)$.

în R $z^* = qnorm(\alpha)$.

Testul bilateral la proporție






- Pentru $H_0 : p = p_0$ and $H_a : p \neq p_0$ - *bilateral sau simetric - two-tailed.*

valoarea $P : P(Z > |z|) + P(Z < -|z|)$,

valoarea critică $z^* > 0$, a. î. $P(Z < -|z^*|) = \alpha/2 =$
 $= 1 - P(Z > |z^*|)$.

în $\mathbb{R} z^* = -qnorm(\alpha/2) = qnorm(1 - \alpha/2)$.

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.
-  <http://effectsizefaq.com/2010/05/31/i-always-get-confused-about-type-i-and-ii->