

Curs 12

Cristian Niculescu

1 Comparație între deducțiile frecvenționistă și Bayesiană

1.1 Scopul învățării

Să poată să explice diferența dintre valoare p și probabilitate a posteriori unui doctor.

1.2 Introducere

Am aflat despre cele 2 școli de deducție statistică: Bayesiană și frecvenționistă. Ambele abordări permit evaluarea dovezilor despre ipoteze concurente.

1.3 Formula lui Bayes ca piatră de încercare

Formula lui Bayes este o propoziție abstractă despre probabilități condiționate ale evenimentelor:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Am început deducția Bayesiană reinterpretând evenimentele din formula lui Bayes:

$$P(\mathcal{H}|\mathcal{D}) = \frac{P(\mathcal{D}|\mathcal{H})P(\mathcal{H})}{P(\mathcal{D})}.$$

Acum \mathcal{H} este o ipoteză și \mathcal{D} sunt datele care pot da dovezi pentru sau împotriva lui \mathcal{H} . Fiecare termen din formula lui Bayes are un nume și un rol. A priori $P(\mathcal{H})$ este probabilitatea ca \mathcal{H} să fie adevărată înainte de considerarea datelor.

A posteriori $P(\mathcal{H}|\mathcal{D})$ este probabilitatea ca \mathcal{H} să fie adevărată după considerarea datelor.

Verosimilitatea $P(\mathcal{D}|\mathcal{H})$ este dovada despre \mathcal{H} furnizată de datele \mathcal{D} .

$P(\mathcal{D})$ este probabilitatea totală a datelor luând în considerare toate ipotezele

posibile.

Dacă a priori și verosimilitatea sunt cunoscute pentru toate ipotezele, atunci formula lui Bayes calculează a posteriori exact. Acesta a fost cazul când am aruncat un zar selectat aleator dintr-o urnă [al cărei conținut îl știam](#). Numim aceasta logica deductivă a teoriei probabilităților și ea dă un mod direct de a compara ipotezele, a trage concluzii și a lua decizii.

În cele mai multe experimente, probabilitățile a priori ale ipotezelor nu sunt cunoscute. În acest caz, recurgem la arta deducției statistice: ori inventăm o a priori (Bayesianii), ori facem tot ce putem folosind doar verosimilitatea (frecvenționiștii).

Scoala Bayesiană modelează incertitudinea printr-o repartiție de probabilitate peste ipoteze. Abilitatea cuiva de a face deducții depinde de gradul de încredere în a priori aleasă și robustețea constatărilor pentru a alterna repartițiile a priori poate fi relevantă și importantă.

Scoala frecvenționistă folosește doar repartiții condiționate ale datelor cunoscând ipoteze specifice. Presupunerea este că o ipoteză (parametru specificând repartiția condiționată a datelor) este adevărată și că datele observate provin din acea repartiție. În particular, abordarea frecvenționistă nu depinde de o a priori subiectivă care poate varia de la un investigator la altul.

Aceste 2 școli pot fi comparate mai departe după cum urmează:

Deducția Bayesiană

folosește probabilități atât pentru ipoteze cât și pentru date;
depinde de a priori și verosimilitatea datelor observate;
cere să se știe sau să se construiască o "a priori subiectivă";
a dominat practica statistică înainte de secolul 20;
poate fi intens calculată din cauza integrării peste mulți parametri.

Deducția frecvenționistă (NHST)

nu folosește niciodată probabilitatea unei ipoteze (nu a priori sau a posteriori);
depinde de verosimilitatea $P(\mathcal{D}|\mathcal{H})$ atât pentru datele observate cât și pentru cele neobservate;
nu cere o a priori;
practica statistică dominantă de-a lungul secolului 20;
tinde să fie mai puțin intensă din punctul de vedere al calculelor.

Măsurile frecvenționiste ca valori p sau intervale de încredere continuă să domine cercetarea, în special în științele vieții. Totuși, în era actuală a computerelor puternice și datelor mari, metodele Bayesiene au avut o renaștere enormă în domenii ca învățarea automată și genetica. Acum sunt un număr de studii clinice mari, în curs de desfășurare folosind protocoale Bayesiene, ceea ce ar fi fost greu de imaginat cu o generație în urmă.

În timp ce împărțirile profesionale rămân, consensul care se formează printre

statisticienii de top este că cea mai eficace abordare a problemelor complexe adesea folosește cele mai bune perspective din ambele școli lucrând împreună.

1.4 Critici și apărări

1.4.1 Critica deducției Bayesiene

1. Principala critică a deducției Bayesiene este că o a priori este subiectivă. Nu există o singură metodă de a alege a priori, deci persoane diferite pot produce a priori diferite și de aceea pot ajunge la a posteriori și concluzii diferite.
2. Mai departe, sunt obiecții filozofice la atribuirea de probabilități ipotezelor, deoarece ipotezele nu sunt rezultate ale experimentelor repetabile în care se poate măsura frecvența pe termen lung. Mai degrabă, o ipoteză este ori adevărată ori falsă, indiferent dacă se știe care este cazul. O monedă este ori corectă, ori incorectă; tratamentul 1 este ori mai bun, ori mai rău ca tratamentul 2; soarele va răsări sau nu va răsări mâine.

1.4.2 Apărarea deducției Bayesiene

1. Probabilitatea ipotezelor este exact ce ne trebuie pentru a lua decizii. Când doctorul îmi spune că un test a ieșit pozitiv vreau să știu care este probabilitatea că aceasta înseamnă că sunt bolnav. Adică, vreau să știu probabilitatea ipotezei "Sunt bolnav".
2. Folosirea teoremei lui Bayes este riguroasă logic. Odată ce avem a priori toate calculele noastre au certitudinea logicii deductive.
3. Încercând diferite a priori, putem vedea cât de sensibile sunt rezultatele noastre de alegerea a priori.
4. Este ușor de comunicat un rezultat încadrat în termeni de probabilitățile ipotezelor.
5. Chiar dacă a priori pot fi subiective, se pot specifica presupunerile folosite pentru a ajunge la ele, care permit altor oameni să le folosească sau să încerce alte a priori.
6. Dovezile derivate din date sunt independente de noțiuni despre "datele mai extreme", care depind de schema experimentală exactă.
7. Datele pot fi folosite pe măsură ce vin. Nu există o cerință ca fiecare contingent să fie planificat înainte.

1.4.3 Critica deducției frecvenționiste

1. Este ad-hoc și nu are forța logicii deductive. Noțiuni ca "date mai extreme" nu sunt bine definite. Valoarea p depinde de schema experimentală exactă.

2. Experimentele trebuie complet specificate înainte. Aceasta poate conduce la rezultate aparent paradoxale. Vezi "istoria voltmetrului" în:

http://en.wikipedia.org/wiki/Likelihood_principle.

3. Valoarea p și nivelul de semnificație sunt notoriu predispuse la interpretare greșită. Un nivel de semnificație de 0.05 înseamnă că probabilitatea unei erori de tipul I este 5%. Adică, **dacă ipoteza 0 este adevărată**, atunci în 5% din timp va fi respinsă din cauza caracterului aleatoriu. Mulți oameni gândesc eronat că o valoare p de 0.05 înseamnă că probabilitatea ipotezei 0 este 5%. Ați putea argumenta că aceasta nu este o critică a deducției frecvenționiste ci, mai degrabă, o critică a ignoranței populare. Totuși, subtilitatea ideilor contribuie sigur la problemă.

1.4.4 Apărarea deducției frecvenționiste

1. Este obiectivă: toți statisticienii vor fi de acord cu valoarea p . Orice individ poate apoi decide dacă valoarea p garantează respingerea ipotezei 0.
2. Testarea ipotezei folosind testarea semnificației frecvenționistă este aplicată în analiza statistică a investigațiilor științifice, evaluând puterea dovezilor împotriva unei ipoteze 0 cu date. Interpretarea rezultatelor este lăsată utilizatorului testelor. Diferiți utilizatori pot aplica diferite nivele de semnificație pentru a determina semnificația statistică. Statistica frecvenționistă nu pretinde că dă un mod de a alege un nivel de semnificație; mai degrabă descrie explicit compromisul între erorile de tipul I și de tipul II.
3. Proiectarea experimentului frecvenționistă cere o descriere atentă a experimentului și metodelor de analiză înainte de startul lui. Aceasta ajută controlul pentru părtinirea experimentatorului.
4. Abordarea frecvenționistă a fost folosită de peste 100 de ani și am văzut progres științific extraordinar. Cu toate că frecvenționistul însuși n-ar pune o probabilitate pe încrederea că metodele frecvenționiste sunt valoroase, n-ar trebui ca această istorie să dea Bayesianului o încredere a priori puternică în utilitatea metodelor frecvenționiste?

1.5 Fiți atenți la p -urile voastre

Facem un test t cu 2 selecții pentru medii egale, cu $\alpha = 0.05$ și obținem o valoare p de 0.04. Care sunt șansele ca cele 2 selecții să provină din repartiții cu aceeași medie?

(a) 19/1 (b) 1/19 (c) 1/20 (d) 1/24 (e) necunoscute.

Răspuns. (e) necunoscute. Metodele frecvenționiste dau doar probabilități ale statisticilor condiționate de ipoteze. Ele nu dau probabilități ale ipotezelor.

1.6 Reguli de oprire

Când derulăm o serie de încercări, avem nevoie de o regulă despre când să ne oprim. 2 reguli uzuale sunt:

1. Facem exact n încercări și ne oprim.
2. Facem încercări până vedem un rezultat și apoi ne oprim.

În acest exemplu vom considera 2 experimente de aruncarea monedei.

Experimentul 1: Aruncăm moneda exact de 6 ori și raportăm numărul aversurilor.

Experimentul 2: Aruncăm moneda până apare primul revers și raportăm numărul aversurilor.

Jon este îngrijorat că moneda lui este deplasată spre aversuri, deci, înainte s-o utilizeze la cursuri, el îi testează corectitudinea. El face un experiment și raportează lui Jerry că secvența lui de aruncări a fost $HHHHHT$. Dar Jerry nu este atent și uită ce experiment a făcut Jon să producă datele.

Abordarea frecvenționistă

Deoarece a uitat ce experiment a făcut Jon, Jerry frecvenționistul decide să calculeze valorile p pentru ambele experimente cunoscând datele lui Jon.

Fie θ probabilitatea aversului. Avem ipotezele 0 și alternativă unilaterală

$$H_0 : \theta = 0.5, H_A : \theta > 0.5.$$

Experimentul 1: Repartiția 0 este binomială(6, 0.5) deci valoarea p unilaterală este probabilitatea a 5 sau 6 aversuri în 6 aruncări. Folosind R obținem

$$p = 1 - \text{pbinom}(4, 6, 0.5) = 0.109375.$$

Experimentul 2: Repartiția 0 este geometrică(0.5) deci, valoarea p unilaterală este probabilitatea a 5 sau mai multe aversuri înaintea primului revers. Folosind R obținem

$$p = 1 - \text{pgeom}(4, 0.5) = 0.03125.$$

Folosind nivelul de semnificație tipic de 0.05, aceleași date duc la concluzii diferite! Am respinge H_0 în experimentul 2, dar nu în experimentul 1.

Frecvenționistul nu are nicio problemă cu aceasta. Mulțimea rezultatelor este diferită pentru experimentele diferite, deci noțiunea de date extreme și de aceea, de valoare p , este diferită. De exemplu, în experimentul 1 am considera $THHHHH$ a fi la fel de extrem ca $HHHHHT$. În experimentul 2 n-am vedea niciodată $THHHHH$ deoarece experimentul s-ar termina după primul revers.

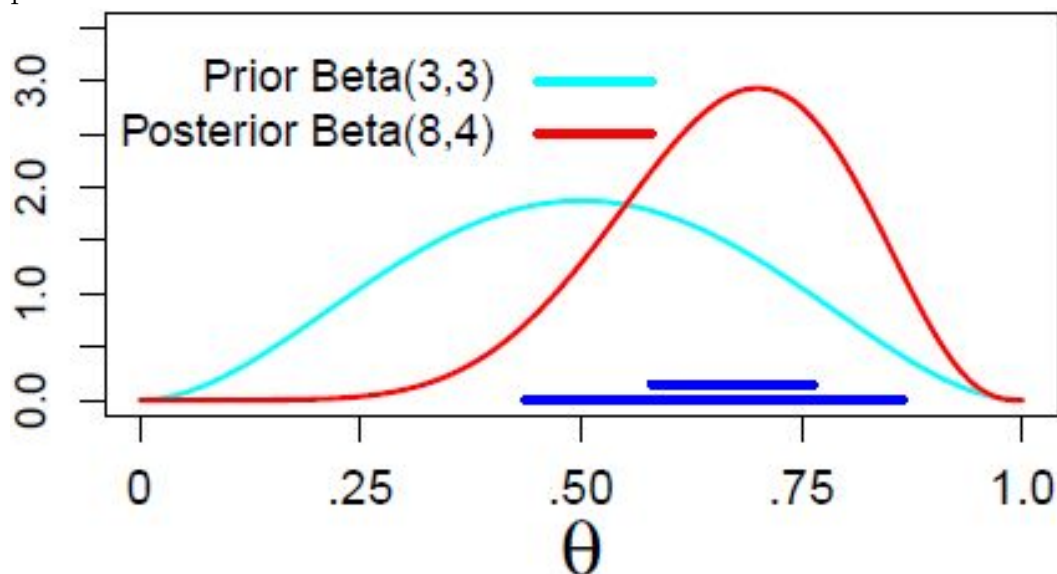
Abordarea Bayesiană.

Jerry Bayesianul știe că nu contează care dintre cele 2 experimente a fost

făcut de Jon, deoarece funcțiile (coloanele) de verosimilitate binomială și geometrică pentru datele $HHHHHT$ sunt proporționale. În orice caz, el trebuie să inventeze o a priori și alege $\text{Beta}(3, 3)$. Aceasta este o a priori relativ plată concentrată peste intervalul $0.25 \leq \theta \leq 0.75$.

Vezi <http://mathlets.org/mathlets/beta-distribution>.

Deoarece repartițiile beta și binomială (sau geometrică) formează o pereche conjugată, actualizarea Bayesiană este simplă. Datele de 5 aversuri și 1 revers dau o repartiție a posteriori $\text{Beta}(8,4)$. Iată graficul a priori și a posteriori. Liniile albastre de jos sunt intervale de probabilitate 50% și 90% pentru a posteriori.



Repartițiile a priori și a posteriori cu intervale de probabilitate 0.5 și 0.9

Iată calculele relevante din R:

Interval de probabilitate 50% a posteriori:

`qbeta(c(0.25,0.75),8,4) = [0.579529, 0.7635974]`.

Interval de probabilitate 90% a posteriori:

`qbeta(c(0.05,0.95),8,4) = [0.4356258, 0.8649245]`.

$P(\theta > 0.5 | \text{date}) = 1 - \text{pbeta}(0.5, 8, 4) = 0.8867188$.

Plecând de la a priori $\text{Beta}(3,3)$, probabilitatea a posteriori ca moneda să fie deplasată spre avers este 0.8867188.

1.7 Luarea deciziilor

Destul de des scopul deducției statistice este de a ajuta la luarea deciziilor, de exemplu dacă sau nu să fie supus unei intervenții chirurgicale, cât de mult să investească într-o acțiune, dacă sau nu să mergă să termine școala, etc.

În teoria deciziilor statistice, consecințele acțiunilor sunt măsurate printr-o funcție de utilitate. Funcția de utilitate atribuie o pondere fiecărui rezultat posibil; în limbajul probabilităților este pur și simplu o variabilă aleatoare.

De exemplu, în investițiile mele aș putea atribui o utilitate de d rezultatului unui câștig de d lei pe acțiune (dacă $d < 0$, utilitatea mea este negativă). Pe de altă parte, dacă toleranța mea pentru risc este mică, voi atribui mai multă utilitate negativă pentru pierderi decât pentru câștiguri (să zicem, $-d^2 - 1$ dacă $d < 0$ și d dacă $d \geq 0$).

O regulă de decizie combină utilitatea medie cu dovezile pentru fiecare ipoteză cunoscând datele (de exemplu, valori p sau repartiții a posteriori) într-un cadru statistic formal pentru luarea deciziilor.

În acest cadru, frecvenționistul va considera media utilității dată fiind o ipoteză

$$E(U|\mathcal{H}),$$

unde U este variabila aleatoare reprezentând utilitatea. Există metode frecvenționiste pentru combinarea utilității medii cu valori p ale ipotezelor pentru a ghida deciziile.

Bayesianul poate combina $E(U|\mathcal{H})$ cu a posteriori (sau a priori dacă este înaintea colectării datelor) pentru a crea o regulă de decizie Bayesiană.

În orice cadru, 2 persoane considerând aceeași investiție pot avea funcții de utilitate diferite și lua decizii diferite. De exemplu, o acțiune riscantă (cu un potențial mai mare de a fluctua) va fi mai atrăgătoare în raport cu prima funcție de utilitate de mai sus decât în raport cu a 2-a.

Un rezultat teoretic semnificativ este că pentru orice regulă de decizie există o regulă de decizie Bayesiană care este, într-un sens precis o regulă cel puțin la fel de bună.

2 Intervale de încredere bazate pe date normale

2.1 Scopurile învățării

1. Să poată să determine dacă o expresie definește o statistică de interval validă.

2. Să poată calcula intervale de încredere z și t pentru media datelor normale.
3. Să poată calcula intervalul de încredere χ^2 pentru dispersia datelor normale.
4. Să poată defini nivelul de încredere al unui interval de încredere.
5. Să poată explica relația dintre intervalul de încredere z (și nivelul de încredere) și regiunea de nerespingere z (și nivelul de încredere) în NHST.

2.2 Introducere

Intervalele de încredere sunt unelte ale statisticii frecvenționiste. Presupunem că avem un model (repartiție de probabilitate) pentru date observate cu un parametru necunoscut. Am văzut cum NHST folosește datele pentru a testa ipoteza că parametrul necunoscut are o valoare particulară.

Am văzut de asemenea cum estimări punctuale ca MLE folosesc datele pentru a da o estimare a parametrului necunoscut. Pe cont propriu, o estimare punctuală ca $\bar{x} = 2.2$ nu poartă informație despre acuratețea ei; este doar un singur număr, indiferent dacă este bazat pe 10 date sau 1000000 de date.

Din acest motiv, statisticienii măresc estimările punctuale la intervale de încredere. De exemplu, pentru a estima o medie necunoscută μ am putea spune că cea mai bună estimare a mediei este $\bar{x} = 2.2$ cu un interval de 95% încredere $[1.2, 3.2]$. Alt mod de a descrie intervalul este $\bar{x} \pm 1$.

Luare împreună, lungimea intervalului și nivelul de încredere dau o măsură a puterii dovezilor care sprijină ipoteza că μ este aproape de estimarea noastră \bar{x} . Nivelul de încredere al unui interval este analog nivelului de semnificație al unui NHST. Nu este un accident că adesea vedem nivelul de semnificație $\alpha = 0.05$ și nivelul de încredere $0.95 = 1 - \alpha$.

Intervalele de încredere z și t pentru medie și χ^2 pentru dispersie sunt ușor de calculat. Provocarea cu intervalele de încredere nu este calculul lor, ci mai degrabă interpretarea lor corectă și a ști cum să le folosim în practică.

2.3 Statistici interval

Reamintim definiția noastră de lucru: o statistică este orice se poate calcula din date. În particular, formula pentru o statistică nu poate include cantități necunoscute.

Exemplul 1. Presupunem că x_1, \dots, x_n provin din $N(\mu, \sigma^2)$, unde μ și σ sunt necunoscute.

- (i) \bar{x} și $\bar{x} - 5$ sunt statistici.
- (ii) $\bar{x} - \mu$ nu este o statistică deoarece μ este necunoscută.
- (iii) Dacă μ_0 este o valoare cunoscută, atunci $\bar{x} - \mu_0$ este o statistică. Acest

caz apare când considerăm ipoteza $\mu = \mu_0$. De exemplu, dacă ipoteza H_0 este $\mu = 5$, atunci statistica $\bar{x} - \mu_0$ este chiar $\bar{x} - 5$ de la (i).

Putem juca același joc cu intervale pentru a defini **statistici interval**.

Exemplul 2. Presupunem că x_1, \dots, x_n provin din $N(\mu, \sigma^2)$ unde μ este necunoscută.

(i) Intervalul $[\bar{x} - 2.2, \bar{x} + 2.2] = \bar{x} \pm 2.2$ este o statistică interval.

(ii) Dacă σ este cunoscută, atunci $\left[\bar{x} - \frac{2\sigma}{\sqrt{n}}, \bar{x} + \frac{2\sigma}{\sqrt{n}}\right]$ este o statistică interval.

(iii) Pe de altă parte, dacă σ este necunoscută, atunci $\left[\bar{x} - \frac{2s}{\sqrt{n}}, \bar{x} + \frac{2s}{\sqrt{n}}\right]$ **nu** este o statistică interval.

(iv) Dacă s^2 este dispersia de selecție, atunci $\left[\bar{x} - \frac{2s}{\sqrt{n}}, \bar{x} + \frac{2s}{\sqrt{n}}\right]$ este o statistică interval deoarece s^2 este calculată din date.

Tehnic, o statistică interval nu este mai mult decât o pereche de statistici punctuale care dau marginile inferioară și superioară ale intervalului.

Observații.

1. Intervalul este aleator - noi date aleatoare vor produce un nou interval.
2. Intervalul nu depinde de valoarea unui parametru necunoscut.
3. Pdf pentru unele repartiții este funcție pară, pentru altele nu.

Exemplul 3. Pdf-urile pentru $N(0, 1)$ și pentru t sunt funcții pare, i.e. graficul este simetric față de axa Oy , dar pentru χ^2 nu.

2.4 Intervale de încredere z pentru medie

Presupunem că avem date normal repartizate:

$$x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2)$$

2.4.1 Definiția intervalelor de încredere z pentru medie

Definiție. Presupunem că datele $x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2)$, cu medie necunoscută μ și dispersie cunoscută σ^2 . Intervalul de încredere $(1 - \alpha)$ pentru μ este

$$\left[\bar{x} - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \bar{x} + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}\right], \quad (1)$$

unde $z_{\alpha/2}$ este **valoarea critică din dreapta**, pentru care $P(Z > z_{\alpha/2}) = \alpha/2$. De exemplu, dacă $\alpha = 0.05$, atunci $z_{\alpha/2} = \text{qnorm}(1 - 0.05/2) \approx 1.96$, deci intervalul de încredere 0.95 (sau 95%) este

$$\left[\bar{x} - \frac{1.96\sigma}{\sqrt{n}}, \bar{x} + \frac{1.96\sigma}{\sqrt{n}}\right].$$

Următoarea aplicație generează date normale și arată intervalele de încredere z sau t pentru medie.

<http://mathlets.org/mathlets/confidence-intervals/>.

Exemplul 4. Presupunem că colectăm 100 de date dintr-o repartiție $N(\mu, 3^2)$ și media de selecție este $\bar{x} = 12$. Dați intervalul de încredere 95% pentru μ .

Răspuns. Folosind formula, intervalul de încredere 95% pentru μ este

$$\left[\bar{x} - \frac{1.96\sigma}{\sqrt{n}}, \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \right] = \left[12 - \frac{1.96 \cdot 3}{10}, 12 + \frac{1.96 \cdot 3}{10} \right] = [11.412, 12.588].$$

2.4.2 Explicarea definiției partea 1: regiuni de respingere

Explicăm definiția (1) plecând de la cunoștințele noastre despre regiuni de respingere/nerespingere. Explicarea ”regiune de nerespingere” nu este frumoasă, dar o vom folosi în locul exprimării incorecte ”regiune de acceptare”.

Exemplul 5. Presupunem că $n = 12$ date provin din $N(\mu, 5^2)$, unde μ este necunoscută. Faceți un test de semnificație bilateral pentru $H_0: \mu = 2.71$ folosind statistica \bar{x} la nivelul de semnificație $\alpha = 0.05$. Descrieți regiunile de respingere și nerespingere.

Răspuns. Sub ipoteza 0 avem $\mu = 2.71$, $x_i \sim N(2.71, 5^2)$ și deci

$$\bar{x} \sim N(2.71, 5^2/12),$$

unde $5^2/12$ este dispersia $\sigma_{\bar{x}}^2$ a lui \bar{x} . Știm că semnificația $\alpha = 0.05$ corespunde la o regiune de respingere în afara a 1.96 deviații standard ale mediei ipotetice. Adică, regiunile de nerespingere și de respingere sunt separate de valorile critice $2.71 \pm 1.96\sigma_{\bar{x}}$.

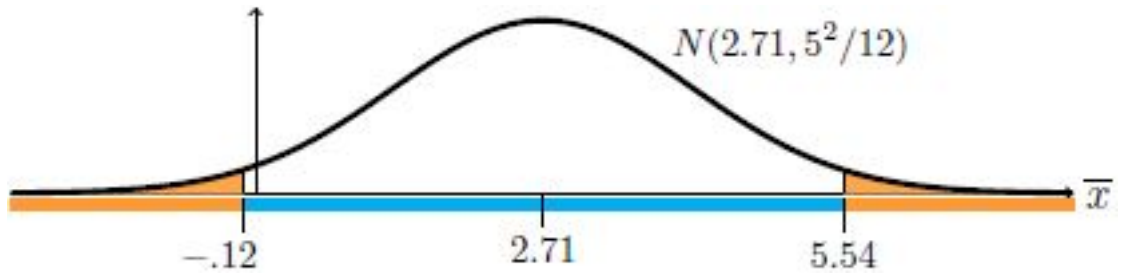
Regiunea de nerespingere:

$$\left[2.71 - \frac{1.96 \cdot 5}{\sqrt{12}}, 2.71 + \frac{1.96 \cdot 5}{\sqrt{12}} \right] \approx [-0.12, 5.54].$$

Regiunea de respingere:

$$\left(-\infty, 2.71 - \frac{1.96 \cdot 5}{\sqrt{12}} \right) \cup \left(2.71 + \frac{1.96 \cdot 5}{\sqrt{12}}, \infty \right) \approx (-\infty, -0.12) \cup (5.54, \infty).$$

Următoarea figură arată regiunile de respingere și nerespingere pentru \bar{x} . Regiunile reprezintă domenii pentru \bar{x} , deci sunt reprezentate de bare colorate pe axa \bar{x} . Aria regiunii colorate de sub grafic este nivelul de semnificație.



Regiunile de respingere (portocalie) și de nerespingere (albastră) pentru \bar{x} .

Exemplul 6. Presupunem că n date provin din $N(\mu, \sigma^2)$, unde μ este necunoscută și σ este cunoscută. Faceți un test de semnificație bilateral al lui $H_0: \mu = \mu_0$ folosind statistica \bar{x} la nivelul de semnificație $\alpha = 0.05$. Descrieți regiunile de respingere și nerespingere.

Răspuns. Sub ipoteza $H_0: \mu = \mu_0$ avem $x_i \sim N(\mu_0, \sigma^2)$ și deci

$$\bar{x} \sim N(\mu_0, \sigma^2/n),$$

unde σ^2/n este dispersia $\sigma_{\bar{x}}^2$ a lui \bar{x} și μ_0, σ și n sunt toate valori cunoscute.

Fie $z_{\alpha/2}$ valoarea critică: $P(Z > z_{\alpha/2}) = \alpha/2$. Regiunile de nerespingere și de respingere sunt separate de valorile lui \bar{x} care sunt la distanță de $z_{\alpha/2} \cdot \sigma_{\bar{x}}$ de media ipotetică. Deoarece $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$, avem:

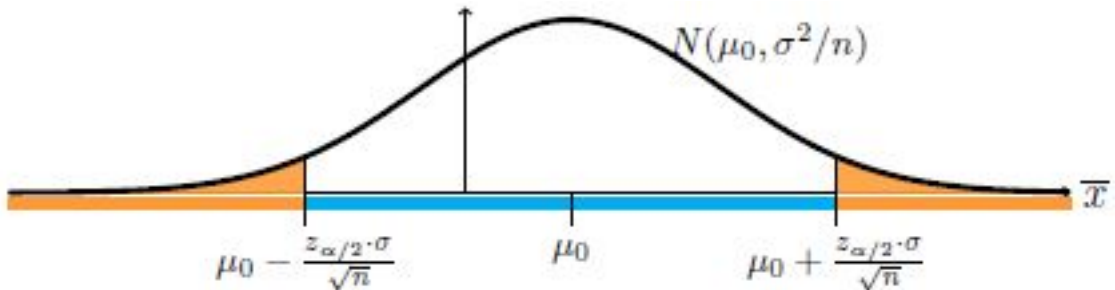
Regiunea de nerespingere:

$$\left[\mu_0 - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \mu_0 + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right]. \quad (2)$$

Regiunea de respingere:

$$\left(-\infty, \mu_0 - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right) \cup \left(\mu_0 + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \infty \right).$$

Obținem aceeași figură ca mai sus, cu numerele explicite înlocuite de valori simbolice.



Regiunile de respingere (portocalie) și de nerespingere (albastră) pentru \bar{x} .

2.4.3 Manipularea intervalelor: pivotarea

Pivotarea este ideea că $\bar{x} \in [\mu_0 - a, \mu_0 + a] \iff \mu_0 \in [\bar{x} - a, \bar{x} + a], \forall a > 0$.

Exemplul 7. Presupunem că avem media de selecție \bar{x} și media ipotetică $\mu_0 = 2.71$. Presupunem de asemenea că repartiția 0 este $N(\mu_0, 3^2)$. Atunci, cu un nivel de semnificație de 0.05, avem:

$\mu_0 + 1.96\sigma = 2.71 + 1.96 \cdot 3 = 2.71 + 5.88$ este valoarea critică 0.025;

$\mu_0 - 1.96\sigma = 2.71 - 1.96 \cdot 3 = 2.71 - 5.88$ este valoarea critică 0.975.

Regiunea de nerespingere este centrată în $\mu_0 = 2.71$. Adică, nu respingem H_0 dacă \bar{x} este în intervalul

$$[\mu_0 - 1.96\sigma, \mu_0 + 1.96\sigma] = [2.71 - 5.88, 2.71 + 5.88].$$

Intervalul de încredere este centrat în \bar{x} . Intervalul de încredere 0.95 are aceeași lungime cu regiunea de nerespingere. Este intervalul

$$[\bar{x} - 1.96\sigma, \bar{x} + 1.96\sigma] = [\bar{x} - 5.88, \bar{x} + 5.88].$$

Aici este o simetrie:

$$\bar{x} \in [2.71 - 1.96\sigma, 2.71 + 1.96\sigma] \iff 2.71 \in [\bar{x} - 1.96\sigma, \bar{x} + 1.96\sigma].$$

Această simetrie este numită pivotare. Iată câteva exemple numerice de pivotare:

Exemplul 8. (i) $1.5 \in [0 - 2.3, 0 + 2.3] \implies 0 \in [1.5 - 2.3, 1.5 + 2.3]$.

(ii) Analog $1.5 \notin [0 - 1, 0 + 1] \implies 0 \notin [1.5 - 1, 1.5 + 1]$.

Simetria poate fi mai clară dacă vorbim în termeni de distanțe: afirmația

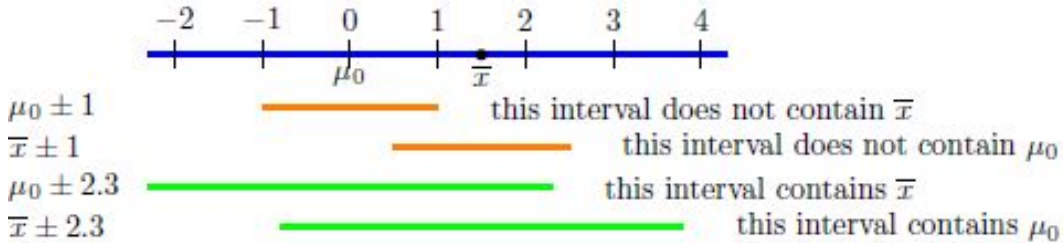
$$”1.5 \in [0 - 2.3, 0 + 2.3]”$$

spune că distanța de la 1.5 la 0 este cel mult 2.3. Analog, afirmația

$$”0 \in [1.5 - 2.3, 1.5 + 2.3]”$$

spune exact același lucru, i.e. distanța de la 0 la 1.5 este cel mult 2.3.

Iată o vizualizare a *pivotării* din intervalele din jurul lui μ_0 în intervalele din jurul lui \bar{x} .



Distanța dintre \bar{x} și μ_0 este 1.5. Deoarece $1 < 1.5$, $[\mu_0 - 1, \mu_0 + 1]$ nu se întinde destul de departe pentru a conține \bar{x} . Analog, intervalul $[\bar{x} - 1, \bar{x} + 1]$ nu se întinde destul de mult pentru a conține μ_0 . Deoarece $2.3 > 1.5$, $\bar{x} \in [\mu_0 - 2.3, \mu_0 + 2.3]$ și $\mu_0 \in [\bar{x} - 2.3, \bar{x} + 2.3]$.

2.4.4 Explicarea definiției, partea a 2-a: translatarea regiunii de nerespingere într-un interval de încredere

În exemplele precedente avem o ipoteză 0. Dar dacă **nu avem o ipoteză 0?** În acest caz, avem estimarea punctuală \bar{x} , dar încă vrem să folosim datele pentru a estima un interval pentru media necunoscută. Adică, vrem o statistică interval. Aceasta este dată de un interval de încredere.

Folosind numerele din exemplul 5, spunem că la nivelul de semnificație de 0.05, nu respingem H_0 dacă

$$\bar{x} \in \left[2.71 - \frac{1.96 \cdot 5}{\sqrt{12}}, 2.71 + \frac{1.96 \cdot 5}{\sqrt{12}} \right]. \quad (3)$$

Rolurile lui \bar{x} și 2.71 sunt simetrice. Nu respingem dacă

$$2.71 \text{ este în intervalul } \bar{x} \pm \frac{1.96 \cdot 5}{\sqrt{12}}. \quad (4)$$

Am ajuns la scopul nostru al unei statistici interval care estimează media necunoscută. Putem rescrie (4) astfel: la nivelul de semnificație de 0.05 nu respingem dacă

$$2.71 \in \left[\bar{x} - \frac{1.96 \cdot 5}{\sqrt{12}}, \bar{x} + \frac{1.96 \cdot 5}{\sqrt{12}} \right]. \quad (5)$$

Astfel, diferite valori ale lui \bar{x} generează diferite intervale.

Intervalul din (5) este exact **intervalul de încredere** definit în (1).

Facem observații despre intervalul de încredere:

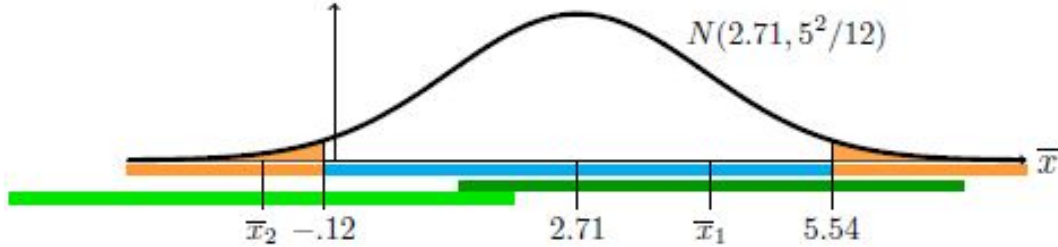
1. Depinde doar de \bar{x} , deci este o statistică.
2. Nivelul de semnificație $\alpha = 0.05$ înseamnă că, **presupunând că ipoteza 0 că $\mu = 2.71$ este adevărată**, datele aleatoare ne vor duce la respingerea ipotezei 0 în 5% din timp (o eroare de tipul I).
3. Din nou, **presupunând** că $\mu = 2.71$, atunci 5% din timp intervalul de încredere nu va conține 2.71 și reciproc, 95% din timp va conține 2.71.

Următoarea figură ilustrează cum nu respingem H_0 dacă intervalul de încredere centrat în \bar{x} conține μ_0 și respingem H_0 dacă intervalul de încredere nu conține μ_0 .

1. Am plecat cu figura de la exemplul 5, care arată repartiția 0 pentru $\mu_0 = 2.71$ și regiunile de respingere și nerespingere.
2. Am adăugat 2 valori posibile ale statisticii \bar{x} , \bar{x}_1 și \bar{x}_2 și intervalele lor de încredere. Observați că lungimea fiecărui interval este exact aceeași ca lungimea regiunii de nerespingere deoarece ambele folosesc $\pm \frac{1.96 \cdot 5}{\sqrt{12}}$.

Prima valoare, \bar{x}_1 este în regiunea de nerespingere și intervalul ei de încredere conține $\mu_0 = 2.71$. Aceasta ilustrează că **nerespingerea** lui H_0 corespunde faptului că intervalul de încredere **conține** pe μ_0 .

A 2-a valoare, \bar{x}_2 , este în regiunea de respingere și intervalul ei de încredere nu conține pe μ_0 . Aceasta ilustrează că **respingerea** lui H_0 corespunde faptului că intervalul de încredere **nu** conține pe μ_0 .



Regiunea de nerespingere (albastru) și cele 2 intervale de încredere (verde).

Alegerea noastră pentru ipoteza 0 $\mu = 2.71$ a fost complet arbitrară. Dacă înlocuim $\mu = 2.71$ cu altă ipoteză $\mu = \mu_0$, atunci intervalul (5) va fi același. Numim intervalul (5) un **interval de încredere** 95% deoarece, **presupunând** $\mu = \mu_0$, în medie, va conține μ_0 în 95% din experimentele aleatoare.

2.4.5 Explicarea definiției partea a 3-a: translatarea unei regiuni de nerespingere generale într-un interval de încredere

Valorile specifice ale lui σ și n din exemplul precedent pot fi înlocuite prin litere. În acest fel putem trece exemplul 6 prin aceiași pași ca exemplul 5. În cuvinte, relația (2) și figura corespunzătoare spune că nu respingem dacă

$$\bar{x} \in \left[\mu_0 - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \mu_0 + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right].$$

Aceasta este echivalent cu a spune că nu respingem dacă

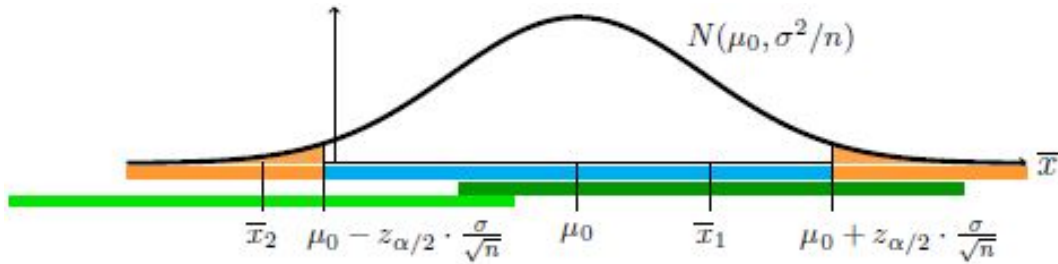
$$\mu_0 \text{ este în intervalul } \bar{x} \pm \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}. \quad (6)$$

Putem rescrie relația (6) astfel: la nivelul de semnificație α nu respingem dacă

$$\mu_0 \in \left[\bar{x} - \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}}, \bar{x} + \frac{z_{\alpha/2} \cdot \sigma}{\sqrt{n}} \right]. \quad (7)$$

Numim intervalul din (7) un **interval de încredere** $(1 - \alpha)$ deoarece, **presupunând** $\mu = \mu_0$, în medie va conține μ_0 în fracția de $(1 - \alpha)$ din experimentele aleatoare.

Figura următoare ilustrează că μ_0 este în intervalul de încredere $(1 - \alpha)$ centrat în $\bar{x} \iff \bar{x}$ este în regiunea de nerespingere (la nivelul de semnificație α) pentru $H_0: \mu = \mu_0$.



\bar{x}_1 este în regiunea de nerespingere pentru $H_0: \mu = \mu_0 \iff$ intervalul de încredere centrat în \bar{x}_1 conține pe μ_0 .

2.4.6 Exemplu de calcul

Exemplul 9. Presupunem că datele 2.5, 5.5, 8.5, 11.5 provin dintr-o repartiție $N(\mu, 10^2)$ cu medie necunoscută μ .

a) Calculați estimarea punctuală \bar{x} pentru μ și intervalele de încredere 95%, 80% și 50% corespunzătoare.

b) Considerăm ipoteza 0 $\mu = 1$. Ați respinge H_0 la $\alpha = 0.05$? $\alpha = 0.2$? $\alpha = 0.5$? Faceți aceasta în 2 moduri: întâi verificând dacă valoarea ipotetică a lui μ este în intervalul de încredere relevant și apoi construind o regiune de respingere.

Răspuns. a) $\bar{x} = \frac{2.5+5.5+8.5+11.5}{4} = 7$. Punctele critice sunt

$$z_{0.025} = \text{qnorm}(0.975) \approx 1.96, z_{0.1} = \text{qnorm}(0.9) \approx 1.28, z_{0.25} = \text{qnorm}(0.75) \approx 0.67.$$

Deoarece $n = 4$, avem $\bar{x} \sim N(\mu, 10^2/4)$, i.e. $\sigma_{\bar{x}} = 5$. Deci avem:

$$\text{interval de încredere 95\%} = [\bar{x} - z_{0.025}\sigma_{\bar{x}}, \bar{x} + z_{0.025}\sigma_{\bar{x}}] \approx [7 - 1.96 \cdot 5, 7 + 1.96 \cdot 5] = [-2.8, 16.8].$$

$$\text{interval de încredere 80\%} = [\bar{x} - z_{0.1}\sigma_{\bar{x}}, \bar{x} + z_{0.1}\sigma_{\bar{x}}] \approx [7 - 1.28 \cdot 5, 7 + 1.28 \cdot 5] = [0.6, 13.4].$$

$$\text{interval de încredere 50\%} = [\bar{x} - z_{0.25}\sigma_{\bar{x}}, \bar{x} + z_{0.25}\sigma_{\bar{x}}] \approx [7 - 0.67 \cdot 5, 7 + 0.67 \cdot 5] = [3.65, 10.35].$$

Fiecare dintre aceste intervale este o estimare interval pentru μ . Dacă nivelul de încredere este mai mare, atunci intervalul corespunzător este mai lung.

b) Deoarece $\mu = 1$ este în intervalele de încredere 95% și 80%, nu respingem ipoteza 0 la nivelele $\alpha = 0.05$ sau $\alpha = 0.2$. Deoarece $\mu = 1$ nu este în intervalul de încredere 50%, respingem H_0 la nivelul $\alpha = 0.5$.

Construim regiunile de respingere folosind aceleași valori critice ca la a). Diferența este că regiunile de respingere sunt complementare de intervale centrate în valoarea ipotetică pentru μ : $\mu_0 = 1$ și intervalele de încredere sunt centrate în \bar{x} . Iată regiunile de respingere:

$$\alpha = 0.05 \implies (-\infty, \mu_0 - z_{0.025}\sigma_{\bar{x}}) \cup (\mu_0 + z_{0.025}\sigma_{\bar{x}}, \infty) \approx (-\infty, -8.8) \cup (10.8, \infty).$$

$$\alpha = 0.2 \implies (-\infty, \mu_0 - z_{0.1}\sigma_{\bar{x}}) \cup (\mu_0 + z_{0.1}\sigma_{\bar{x}}, \infty) \approx (-\infty, -5.4) \cup (7.4, \infty).$$

$$\alpha = 0.5 \implies (-\infty, \mu_0 - z_{0.25}\sigma_{\bar{x}}) \cup (\mu_0 + z_{0.25}\sigma_{\bar{x}}, \infty) \approx (-\infty, -2.35) \cup (4.35, \infty).$$

Pentru a face NHST trebuie să verificăm dacă $\bar{x} = 7$ se află sau nu în regiunea de respingere.

$$\alpha = 0.05 : 7 \notin (-\infty, -8.8) \cup (10.8, \infty).$$

Nu respingem ipoteza că $\mu = 1$ la nivelul de semnificație de 0.05.

$$\alpha = 0.2 : 7 \notin (-\infty, -5.4) \cup (7.4, \infty).$$

Nu respingem ipoteza că $\mu = 1$ la nivelul de semnificație de 0.2.

$$\alpha = 0.5 : 7 \in (-\infty, -2.35) \cup (4.35, \infty).$$

Respingem ipoteza că $\mu = 1$ la nivelul de semnificație de 0.5.
Obținem aceleași răspunsuri cu orice metodă.

2.5 Intervale de încredere t pentru medie

În acest cadru σ este necunoscută, deci trebuie să facem următoarele înlocuiri:

1. Folosim $s_{\bar{x}} = \frac{s}{\sqrt{n}}$ în locul lui $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$. Aici s^2 este dispersia de selecție pe care am folosit-o mai înainte la teste t .
2. Folosim valori critice t în locul valorilor critice z .

2.5.1 Definiția intervalelor de încredere t pentru medie

Definiție. Presupunem că $x_1, \dots, x_n \sim N(\mu, \sigma^2)$, unde valorile mediei μ și deviației standard σ sunt ambele necunoscute. Intervalul de încredere $(1 - \alpha)$ pentru μ este

$$\left[\bar{x} - \frac{t_{\alpha/2} \cdot s}{\sqrt{n}}, \bar{x} + \frac{t_{\alpha/2} \cdot s}{\sqrt{n}} \right], \quad (8)$$

unde $t_{\alpha/2}$ este [valoarea critică din dreapta](#) a. î. $P(T > t_{\alpha/2}) = \alpha/2$ pentru $T \sim t(n - 1)$ și s^2 este dispersia de selecție a datelor.

2.5.2 Construirea intervalelor de încredere t

Presupunem că n date provin din $N(\mu, \sigma^2)$, unde μ și σ^2 sunt necunoscute. Sub ipoteza 0 $\mu = \mu_0$, avem $x_i \sim N(\mu_0, \sigma^2)$. Deci media Studentizată are o repartiție Student t cu $n - 1$ grade de libertate:

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \sim t(n - 1).$$

Fie $t_{\alpha/2}$ valoarea critică a. î. $P(T > t_{\alpha/2}) = \alpha/2$ pentru $T \sim t(n-1)$. Știm din testul t de o selecție că regiunea de nerespingere este dată de

$$|t| \leq t_{\alpha/2}.$$

Folosind definiția statisticii t pentru a scrie regiunea de respingere în termeni de \bar{x} , obținem: la nivelul de semnificație α nu respingem dacă

$$\frac{|\bar{x} - \mu_0|}{s/\sqrt{n}} \leq t_{\alpha/2} \iff |\bar{x} - \mu_0| \leq t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}.$$

Geometric, partea dreaptă spune că nu respingem dacă

$$\mu_0 \text{ este la cel mult } t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \text{ de } \bar{x}.$$

Aceasta este echivalent cu a spune că nu respingem dacă

$$\mu_0 \in \left[\bar{x} - \frac{t_{\alpha/2} \cdot s}{\sqrt{n}}, \bar{x} + \frac{t_{\alpha/2} \cdot s}{\sqrt{n}} \right].$$

Acest interval este intervalul de încredere definit în (8).

Exemplul 10. Presupunem că datele 2.5, 5.5, 8.5, 11.5 provin dintr-o repartiție $N(\mu, \sigma^2)$ cu μ și σ^2 ambele necunoscute.

Dați estimările interval pentru μ aflând intervalele de încredere 95%, 80% și 50%.

Răspuns. În R:

`x=c(2.5,5.5,8.5,11.5)`

`$\bar{x} = \text{mean}(x) = 7$`

`$s^2 = \text{var}(x) = 15$.`

Punctele critice sunt

`$t_{0.025} = \text{qt}(0.975, 3) = 3.182446$,`

`$t_{0.1} = \text{qt}(0.9, 3) = 1.637744$,`

`$t_{0.25} = \text{qt}(0.75, 3) = 0.7648923$.`

interval de încredere 95% = $[\bar{x} - t_{0.025} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{0.025} \cdot \frac{s}{\sqrt{n}}] = [0.8372198, 13.16278]$.

interval de încredere 80% = $[\bar{x} - t_{0.1} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{0.1} \cdot \frac{s}{\sqrt{n}}] = [3.828522, 10.17148]$.

interval de încredere 50% = $[\bar{x} - t_{0.25} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{0.25} \cdot \frac{s}{\sqrt{n}}] = [5.518792, 8.481208]$.

Toate aceste intervale de încredere dau estimări interval pentru valoarea μ . Din nou, dacă nivelul de încredere este mai mare, atunci intervalul corespunzător este mai lung.

2.6 Intervale de încredere χ^2 pentru dispersie

Definiție. Presupunem că datele x_1, \dots, x_n provin din $N(\mu, \sigma^2)$ cu media μ și deviația standard σ ambele necunoscute. Intervalul de încredere $(1 - \alpha)$ pentru dispersia σ^2 este

$$\left[\frac{(n-1)s^2}{c_{\alpha/2}}, \frac{(n-1)s^2}{c_{1-\alpha/2}} \right]. \quad (9)$$

Aici $c_{\alpha/2}$ este **valoarea critică din dreapta** a. î. $P(X^2 > c_{\alpha/2}) = \alpha/2$ pentru $X^2 \sim \chi^2(n-1)$ și s^2 este dispersia de selecție a datelor.

Deducerea acestui interval este aproape identică cu deducerile anterioare, acum plecând de la testul χ^2 pentru dispersie. Faptul de bază de care avem nevoie este că, pentru date provenite din $N(\mu, \sigma^2)$ cu σ cunoscută, statistica

$$\frac{(n-1)s^2}{\sigma^2}$$

are o repartiție χ^2 cu $n-1$ grade de libertate. Deci, dată fiind ipoteza 0 H_0 : $\sigma = \sigma_0$, statistica testului este $(n-1)s^2/\sigma_0^2$ și regiunea de nerespingere la nivelul de semnificație α este

$$c_{1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma_0^2} \leq c_{\alpha/2}.$$

Aceasta este echivalentă cu

$$\frac{(n-1)s^2}{c_{1-\alpha/2}} \geq \sigma_0^2 \geq \frac{(n-1)s^2}{c_{\alpha/2}}.$$

Aceasta spune că nu respingem dacă

$$\sigma_0^2 \in \left[\frac{(n-1)s^2}{c_{\alpha/2}}, \frac{(n-1)s^2}{c_{1-\alpha/2}} \right].$$

Acesta este intervalul nostru de încredere $(1 - \alpha)$.

3 Intervale de încredere: 3 puncte de vedere

3.1 Scopurile învățării

1. Să poată produce intervale de încredere z, t și χ^2 pe baza statisticilor standardizate corespunzătoare.
2. Să poată să utilizeze un test de ipoteză pentru a construi intervalul de încredere pentru un parametru necunoscut.
3. Să refuze să răspundă la întrebări de genul "dat fiind un interval de încredere, care este probabilitatea sau șansele ca el să conțină adevărata valoare a parametrului necunoscut?"

3.2 Introducere

Abordarea noastră a intervalelor de încredere a fost o combinație între statisticile standardizate și testarea ipotezelor. Azi vom considera fiecare dintre aceste perspective separat și vom introduce și un al 3-lea punct de vedere formal. Fiecare dă propria sa perspectivă.

1. **Statistica standardizată.** Cele mai multe intervale de încredere sunt bazate pe statistici standardizate cu repartiții cunoscute ca z, t sau χ^2 . Aceasta dă un mod direct de a construi și interpreta intervalele de încredere ca o estimare punctuală plus sau minus o eroare.

2. **Testarea ipotezei.** Intervale de încredere pot fi de asemenea construite din testele de ipoteză. În cazurile unde nu avem o statistică standardizată această metodă încă funcționează. Este în acord cu abordarea statisticii standardizate în cazurile unde se aplică ambele.

Această perspectivă leagă noțiunile de nivel de semnificație α pentru testarea ipotezei și nivel de încredere $(1 - \alpha)$ pentru intervale de încredere; în ambele cazuri α este probabilitatea de a face o eroare de "tipul 1". Aceasta dă o înțelegere pentru folosirea cuvântului "încredere". Această perspectivă ajută de asemenea la evidențierea naturii frecvenționiste a intervalelor de încredere.

3. **Formal.** Definiția formală a intervalelor de încredere este precisă și generală. Într-un sens matematic dă o perspectivă în funcționarea interioară a intervalelor de încredere. Totuși, deoarece este atât de generală, uneori duce la intervale fără proprietăți utile.

3.3 Intervale de încredere via statistici standardizate

Strategia de aici este în esență aceeași ca aceea anterioară. Presupunând date normale, avem ceea ce numim statistici standardizate, ca media standardizată, media Studentizată și dispersia standardizată. Aceste statistici au repartiții bine cunoscute care depind de valorile ipotetice ale lui μ și σ . Apoi folosim calcul algebric pentru a produce intervale de încredere pentru μ și σ .

Ideea subiacentă intervalelor de încredere este în esență simplă: plecăm cu o statistică standardizată (de exemplu z, t sau χ^2) și folosim ceva calcul algebric pentru a obține un interval care depinde doar de date și parametrii cunoscuți.

3.3.1 Intervale de încredere z pentru μ : date normale cu σ cunoscută

Intervalele de încredere z pentru media datelor normale sunt bazate pe [media standardizată](#), adică pe statistica z . Plecăm cu n date independente normale

$$x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2).$$

Presupunem că μ este parametrul necunoscut de interes și σ este cunoscută. Știm că media standardizată este normală standard:

$$z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \sim N(0, 1).$$

Pentru valoarea critică normală standard $z_{\alpha/2}$ avem:

$$P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1 - \alpha.$$

De aceea,

$$P\left(-z_{\alpha/2} \leq \frac{\bar{x} - \mu}{\sigma/\sqrt{n}} \leq z_{\alpha/2} \mid \mu\right) = 1 - \alpha.$$

Un pic de algebră pune aceasta în forma unui interval în jurul lui μ :

$$P\left(\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}} \mid \mu\right) = 1 - \alpha.$$

Putem sublinia că intervalul depinde doar de statistica \bar{x}, n și valoarea cunoscută σ scriind aceasta ca

$$P\left(\mu \in \left[\bar{x} - z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}\right] \mid \mu\right) = 1 - \alpha.$$

Acesta este intervalul de încredere z ($1 - \alpha$) pentru μ . Adesea îl scriem folosind prescurtarea

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}.$$

Gândiți-l ca $\bar{x} \pm$ eroarea.

3.3.2 Intervale de încredere t pentru μ : date normale cu μ și σ necunoscute

Intervalele de încredere t pentru media datelor normale sunt bazate pe [media Studentizată](#), adică pe statistica t . Plecăm cu n date independente normale

$$x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2).$$

Presupunem că μ este parametrul necunoscut de interes și σ este necunoscută. Știm că media Studentizată are o repartiție Student t cu $n - 1$ grade de libertate:

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \sim t(n - 1),$$

unde s^2 este dispersia de selecție.

Acum, tot ce avem de făcut este să înlocuim media standardizată cu [media Studentizată](#) și aceeași logică pe care am utilizat-o pentru z ne dă intervalul de încredere t : plecăm cu

$$P\left(-t_{\alpha/2} \leq \frac{\bar{x} - \mu}{s/\sqrt{n}} \leq t_{\alpha/2} | \mu\right) = 1 - \alpha.$$

Un pic de algebră pune aceasta în forma unui interval în jurul lui μ :

$$P\left(\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \leq \mu \leq \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} | \mu\right) = 1 - \alpha.$$

Putem sublinia că intervalul depinde doar de statisticile \bar{x} și s și de n scriind aceasta ca

$$P\left(\mu \in \left[\bar{x} - t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}, \bar{x} + t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}\right] | \mu\right) = 1 - \alpha.$$

Acesta este intervalul de încredere t ($1 - \alpha$) pentru μ . Adesea îl scriem folosind prescurtarea

$$\bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}}.$$

Gândiți-l ca $\bar{x} \pm$ eroarea.

3.3.3 Intervale de încredere χ^2 pentru σ^2 : date normale cu μ și σ necunoscute

Intervalele de încredere χ^2 pentru dispersia datelor normale sunt bazate pe [dispersia standardizată](#), adică pe statistica χ^2 .

Urmăm aceeași logică ca mai sus pentru a obține un interval de încredere χ^2 pentru σ^2 .

Presupunem că avem n date normale independente: $x_1, x_2, \dots, x_n \sim N(\mu, \sigma^2)$. Presupunem că μ și σ sunt ambele necunoscute. Dispersia standardizată este

$$X^2 = \frac{(n - 1)s^2}{\sigma^2} \sim \chi^2(n - 1).$$

Știm că statistica X^2 are o repartiție χ^2 cu $n - 1$ grade de libertate.

Pentru Z și t am folosit simetria repartițiilor pentru a înlocui $z_{1-\alpha/2}$ cu $-z_{\alpha/2}$

și $t_{1-\alpha/2}$ cu $-t_{\alpha/2}$. Deoarece repartiția χ^2 nu este simetrică trebuie să fim expliciti despre valorile critice atât la stânga cât și la dreapta. Adică,

$$P(c_{1-\alpha/2} \leq X^2 \leq c_{\alpha/2}) = 1 - \alpha,$$

unde $c_{\alpha/2}$ și $c_{1-\alpha/2}$ sunt valorile critice din [coada dreaptă](#). Deci

$$P\left(c_{1-\alpha/2} \leq \frac{(n-1)s^2}{\sigma^2} \leq c_{\alpha/2} | \sigma\right) = 1 - \alpha.$$

Un pic de algebră pune aceasta în forma unui interval în jurul lui σ^2 :

$$P\left(\frac{(n-1)s^2}{c_{\alpha/2}} \leq \sigma^2 \leq \frac{(n-1)s^2}{c_{1-\alpha/2}} | \sigma\right) = 1 - \alpha.$$

Putem sublinia că intervalul depinde doar de statistica s^2 și de n scriind aceasta ca

$$P\left(\sigma^2 \in \left[\frac{(n-1)s^2}{c_{\alpha/2}}, \frac{(n-1)s^2}{c_{1-\alpha/2}}\right] | \sigma\right) = 1 - \alpha.$$

Acesta este intervalul de încredere χ^2 $(1 - \alpha)$ pentru σ^2 .

3.4 Intervale de încredere via testarea ipotezei

Presupunem că avem date provenite dintr-o repartiție cu un parametru θ a cărui valoare este necunoscută. Un test de semnificație pentru valoarea θ are următoarea descriere scurtă:

1. Fixăm ipoteza 0 H_0 : $\theta = \theta_0$ pentru o valoare specială θ_0 , de exemplu adesea avem H_0 : $\theta = 0$.
2. Folosim datele pentru a calcula valoarea unei statistici a testului, s-o numim x .
3. Dacă x este destul de departe în coada repartiției 0 (repartiția [presupunând](#) ipoteza 0), atunci respingem H_0 .

În cazul unde nu este o valoare specială pentru a fi testată, putem dori încă să estimăm θ . Acesta este reversul testului de semnificație: în loc să vedem dacă ar trebui să respingem o anumită valoare a lui θ deoarece nu se potrivește cu datele, mai degrabă vrem să găsim un domeniu de valori ale lui θ care, într-un anumit sens, se potrivesc cu datele.

Aceasta ne dă următoarele definiții:

Definiție. Fiind dată o valoare x a statisticii testului, intervalul de încredere $(1-\alpha)$ conține toate valorile θ_0 care nu sunt respinse (la nivelul de semnificație α) când aceste valori sunt ipoteza 0.

Definiție. O eroare de tipul 1 CI apare când intervalul de încredere nu

conține adevărata valoare a lui θ .

Pentru un interval de încredere $(1 - \alpha)$ rata erorii de tipul 1 CI este α .

Exemplul 1. Iată un exemplu care leagă intervalele de încredere și testele ipotezei. Presupunem că datele x provin dintr-o repartiție binomială $(12, \theta)$ cu θ necunoscută. Fie $\alpha = 0.1$ și creăm intervalul de încredere $(1 - \alpha) = 90\%$ pentru fiecare valoare posibilă a lui x .

Strategia noastră este să considerăm pe rând câte o valoare posibilă a lui θ și să alegem regiunile de respingere pentru testul de semnificație cu $\alpha = 0.1$. Odată ce acest lucru este făcut, vom ști, pentru orice valoare a lui x , ce valori ale lui θ nu sunt respinse, i.e. intervalul de încredere asociat cu x .

Pentru a începe, facem un tabel de verosimilitate pentru binomiala $(12, \theta)$ (tabelul 1). Fiecare linie arată probabilitățile $p(x|\theta)$ pentru o valoare a lui θ .

Pentru a păstra mărimea rezonabilă arătăm θ doar în creșteri de 0.1.

$\theta \backslash x$	0	1	2	3	4	5	6	7	8	9	10	11	12
1.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
0.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.23	0.38	0.28
0.8	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.05	0.13	0.24	0.28	0.21	0.07
0.7	0.00	0.00	0.00	0.00	0.01	0.03	0.08	0.16	0.23	0.24	0.17	0.07	0.01
0.6	0.00	0.00	0.00	0.01	0.04	0.10	0.18	0.23	0.21	0.14	0.06	0.02	0.00
0.5	0.00	0.00	0.02	0.05	0.12	0.19	0.23	0.19	0.12	0.05	0.02	0.00	0.00
0.4	0.00	0.02	0.06	0.14	0.21	0.23	0.18	0.10	0.04	0.01	0.00	0.00	0.00
0.3	0.01	0.07	0.17	0.24	0.23	0.16	0.08	0.03	0.01	0.00	0.00	0.00	0.00
0.2	0.07	0.21	0.28	0.24	0.13	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00
0.1	0.28	0.38	0.23	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
0.0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Tabelul 1. Tabelul de verosimilitate pentru binomiala $(12, \theta)$

Tabelele 2-4 de mai jos arată regiunea de respingere (în portocaliu) și regiunea de nerespingere (în albastru) pentru diferite valori ale lui θ . Pentru a sublinia natura linie cu linie a procesului, tabelul 2 arată aceste regiuni pentru $\theta = 1$, apoi tabelul 3 adaugă regiunile pentru $\theta = 0.9$ și tabelul 4 le arată pentru toate valorile lui θ .

Imediat după tabele explicăm detaliat cum au fost alese regiunile de respingere/nerespingere. "significance" = "semnificație".

$\theta \backslash x$	0	1	2	3	4	5	6	7	8	9	10	11	12	significance
1.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	
0.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.23	0.38	0.28	0.000
0.8	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.05	0.13	0.24	0.28	0.21	0.07	
0.7	0.00	0.00	0.00	0.00	0.01	0.03	0.08	0.16	0.23	0.24	0.17	0.07	0.01	
0.6	0.00	0.00	0.00	0.01	0.04	0.10	0.18	0.23	0.21	0.14	0.06	0.02	0.00	
0.5	0.00	0.00	0.02	0.05	0.12	0.19	0.23	0.19	0.12	0.05	0.02	0.00	0.00	
0.4	0.00	0.02	0.06	0.14	0.21	0.23	0.18	0.10	0.04	0.01	0.00	0.00	0.00	
0.3	0.01	0.07	0.17	0.24	0.23	0.16	0.08	0.03	0.01	0.00	0.00	0.00	0.00	
0.2	0.07	0.21	0.28	0.24	0.13	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	
0.1	0.28	0.38	0.23	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Tabelul 2. Tabelul de verosimilitate pentru binomiala(12, θ) cu regiunile de respingere/nerespingere pentru $\theta = 1$

$\theta \backslash x$	0	1	2	3	4	5	6	7	8	9	10	11	12	significance
1.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.000
0.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.23	0.38	0.28	0.026
0.8	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.05	0.13	0.24	0.28	0.21	0.07	
0.7	0.00	0.00	0.00	0.00	0.01	0.03	0.08	0.16	0.23	0.24	0.17	0.07	0.01	
0.6	0.00	0.00	0.00	0.01	0.04	0.10	0.18	0.23	0.21	0.14	0.06	0.02	0.00	
0.5	0.00	0.00	0.02	0.05	0.12	0.19	0.23	0.19	0.12	0.05	0.02	0.00	0.00	
0.4	0.00	0.02	0.06	0.14	0.21	0.23	0.18	0.10	0.04	0.01	0.00	0.00	0.00	
0.3	0.01	0.07	0.17	0.24	0.23	0.16	0.08	0.03	0.01	0.00	0.00	0.00	0.00	
0.2	0.07	0.21	0.28	0.24	0.13	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	
0.1	0.28	0.38	0.23	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
0.0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

Tabelul 3. Tabelul de verosimilitate cu regiunile de respingere/nerespingere arătate pentru $\theta = 1$ și 0.9

$\theta \backslash x$	0	1	2	3	4	5	6	7	8	9	10	11	12	significance
1.0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.000
0.9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.09	0.23	0.38	0.28	0.026
0.8	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.05	0.13	0.24	0.28	0.21	0.07	0.073
0.7	0.00	0.00	0.00	0.00	0.01	0.03	0.08	0.16	0.23	0.24	0.17	0.07	0.01	0.052
0.6	0.00	0.00	0.00	0.01	0.04	0.10	0.18	0.23	0.21	0.14	0.06	0.02	0.00	0.077
0.5	0.00	0.00	0.02	0.05	0.12	0.19	0.23	0.19	0.12	0.05	0.02	0.00	0.00	0.092
0.4	0.00	0.02	0.06	0.14	0.21	0.23	0.18	0.10	0.04	0.01	0.00	0.00	0.00	0.077
0.3	0.01	0.07	0.17	0.24	0.23	0.16	0.08	0.03	0.01	0.00	0.00	0.00	0.00	0.052
0.2	0.07	0.21	0.28	0.24	0.13	0.05	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.073
0.1	0.28	0.38	0.23	0.09	0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.026
0.0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.000

Tabelul 4. Tabelul de verosimilitate cu regiunile de respingere/nerespingere

Alegerea regiunilor de respingere și nerespingere în tabele Prima problemă cu care ne confruntăm este cum să alegem exact regiunea de respingere. Am folosit 2 reguli:

1. Probabilitatea totală a regiunii de respingere, i.e. semnificația, trebuie să fie mai mică sau egală cu 0.1. (Deoarece avem o repartiție discretă este imposibil să facem semnificația exact 0.1.)
2. Construim regiunea de respingere alegând pe rând valorile lui x , totdeauna luând valoarea nefolosită cu cea mai mică probabilitate. Ne oprim când valoarea următoare ar face semnificația mai mare ca 0.1.

Sunt și alte moduri de a alege regiunea de respingere din care ar putea rezulta mici diferențe.

Tabelul 2 arată regiunile de respingere (portocalie) și nerespingere (albastră) pentru $\theta = 1$. Acesta este un caz special deoarece cele mai multe probabilități de pe această linie sunt 0. Ne mutăm la tabelul următor și parcurgem procesul pentru acesta.

În tabelul 3, parcurgem pașii pentru a afla regiunile pentru $\theta = 0.9$.

Cea mai mică probabilitate este când $x = 0$ și aceasta este mai mică decât 0.1, deci $x = 0$ este în regiunea de respingere.

Următoarea cea mai mică probabilitate este când $x = 1$ și suma celor 2 probabilități (pentru $x = 0$ și $x = 1$) este mai mică decât 0.1, deci $x = 1$ este în regiunea de respingere.

Continuăm cu $x = 2, 3, \dots, 8$. În acest punct probabilitatea totală din regiunea de respingere este 0.026.

Următoarea cea mai mică probabilitate este când $x = 9$. Adunarea acestei probabilități (0.09) la 0.026 ar face probabilitatea totală peste 0.1. Deci lăsăm $x = 9$ în afara regiunii de respingere și oprim procesul.

Observăm 3 lucruri pentru linia $\theta = 0.9$:

1. Niciuna din probabilitățile de pe această linie nu este cu adevărat 0, totuși unele dintre ele sunt suficient de mici astfel încât aproximarea lor la 2 zecimale este 0.
2. Arătăm semnificația pentru această valoare a lui θ în marginea din dreapta. Mai precis, arătăm nivelul de semnificație al NHST cu ipoteza $0 \theta = 0.9$ și regiunea de respingere dată.
3. Regiunea de respingere constă din valori ale lui x . Când spunem că regiunea de respingere este arătată în portocaliu, în realitate ne referim la faptul că regiunea de respingere constă din valorile lui x corespunzătoare probabilităților evidențiate cu portocaliu.

Gândiți: Întoarceți-vă la linia $\theta = 1$ și asigurați-vă că înțelegeți de ce regiunea de respingere este $x \in \{0, 1, \dots, 11\}$ și semnificația este 0.

Exemplul 2. Folosind tabelul 4, determinați intervalul de încredere 0.9 când $x = 8$.

Răspuns. Intervalul de încredere 90% constă din toate acele θ care nu ar fi respinse de un test al ipotezei cu $\alpha = 0.1$ când $x = 8$. Privind la tabel, intrările albastre (nerespinse) din coloana $x = 8$ corespund la $0.5 \leq \theta \leq 0.8$; intervalul de încredere este $[0.5, 0.8]$.

Observație. Scopul acestui exemplu este de a arăta cum sunt legate intervalele de încredere și testele ipotezei. Deoarece tabelul 4 are doar un număr finit de valori pentru θ , răspunsul nostru este aproape, dar nu exact. Pentru această problemă am putea utiliza R pentru a afla că, corect la 2 zecimale, intervalul de încredere este $[0.42, 0.84]$.

Exemplul 3. Explicați de ce rata erorii CI de tipul 1 va fi cel mult 0.092, cu condiția ca adevărata valoare a lui θ să fie în tabel.

Răspuns. Răspunsul scurt este că acesta este maximul semnificațiilor pentru orice θ din tabelul 4. Mai pe larg: facem o eroare CI de tipul 1 dacă intervalul de încredere nu conține adevărata valoare a lui θ , s-o numim $\theta_{\text{adevărată}}$. Aceasta se întâmplă exact când data x este în regiunea de respingere pentru $\theta_{\text{adevărată}}$. Probabilitatea ca aceasta să se întâmple este semnificația pentru

$\theta_{\text{adevărată}}$ și aceasta este cel mult 0.092.

Observație. Scopul acestui exemplu este să arate cum nivelul de încredere, rata erorii CI de tipul 1 și semnificația pentru fiecare ipoteză sunt legate. Ca în exemplul precedent, putem folosi R pentru a calcula semnificația pentru mult mai multe valori ale lui θ . Când facem asta aflăm că semnificația maximă pentru orice θ este 0.1 apărând când $\theta = 0.04524072$.

Observații de sinteză:

1. Începem cu o statistică de test x . Intervalul de încredere este aleator deoarece depinde de x .
2. Pentru fiecare valoare ipotetică a lui θ facem un test de semnificație cu nivelul de semnificație α alegând regiunile de respingere.
3. Pentru o valoare specifică a lui x intervalul de încredere asociat pentru θ constă din toți θ care nu au fost respinși pentru acea valoare, i.e. toți θ care au x în regiunile lor de nerespingere.
4. Deoarece repartiția este discretă nu putem atinge totdeauna nivelul exact de semnificație, deci intervalul de încredere al nostru este în realitate un "interval de încredere cel puțin 90%".

Exemplul 4. Deschideți aplicația

<http://mathlets.org/mathlets/confidence-intervals/>. Jucați-vă cu aplicația pentru a înțelege natura aleatoare a intervalelor de încredere și sensul încrederii ca (1 - rata erorii CI de tipul 1).

- a) Citiți ajutorul. Este scurt și vă va ajuta să vă orientați în aplicație. Jucați-vă cu diferite setări ale parametrilor pentru a vedea cum afectează ele mărimea intervalelor de încredere.
- b) Stabiliți numărul experimentelor la $N = 1$. Apăsăți butonul "Run N trials" repetat și vedeți că de fiecare dată când datele sunt generate intervalele de încredere sar împrejur.
- c) Acum puneți nivelul de încredere la $c = .5$. Când apăsați butonul "Run N trials" ar trebui să vedeți că aproximativ 50% din intervalele de încredere includ adevărata valoare a lui μ . Valorile "Z correct" și "t correct" ar trebui să se schimbe în consecință.
- d) Acum puneți numărul încercărilor la $N = 100$, cu $c = .8$. Butonul "Run N trials" va face acum 100 de încercări o dată. Doar ultimul interval de încredere va fi arătat în grafic, dar încercările sunt făcute toate și statistica "percent correct" va fi actualizată pe baza tuturor celor 100 de încercări. Apăsăți butonul "Run N trials" repetat. Urmăriți cum ratele corecte încep să converge la nivelul de încredere. Pentru a converge și mai rapid, puneți $N = 1000$.

3.5 Viziune formală a intervalelor de încredere

Reamintim: O statistică interval este un interval I_x calculat din datele x . Un interval este determinat de marginile lui inferioară și superioară și acestea sunt aleatoare deoarece x sunt aleatoare.

Presupunem că x provin dintr-o repartiție cu pdf $f(x|\theta)$, unde parametrul θ este necunoscut.

Definiție: Un interval de încredere $(1 - \alpha)$ pentru θ este o statistică interval I_x astfel încât

$$P(\theta_0 \in I_x | \theta = \theta_0) = 1 - \alpha$$

pentru toate valorile posibile ale lui θ_0 .

Această definiție este un mod de a cântări dovezile produse de datele x .

Nivelul de încredere a unei statistici interval este o probabilitate privind un interval aleator și o valoare ipotetică θ_0 pentru parametrul necunoscut. Exact, este probabilitatea ca intervalul aleator (calculat din datele aleatoare) să conțină valoarea θ_0 , **dat fiind că parametrul modelului este cu adevărat θ_0** . Deoarece adevărata valoare a lui θ este necunoscută, statisticianul frecvenționist definește intervalele de încredere 95% astfel încât probabilitatea 0.95 este validă **indiferent care valoare ipotetică a parametrului este de fapt adevărată**.

3.6 Comparatie cu intervalele de probabilitate Bayesiene

Intervalele de încredere sunt o noțiune frecvenționistă și frecvenționistii nu atribuie probabilități ipotezelor, de exemplu valorii unui parametru necunoscut. Mai degrabă ei calculează verosimilități; adică, probabilități despre date sau statistici asociate dată fiind o ipoteză (observați condiția $\theta = \theta_0$ din viziunea formală a intervalelor de încredere). Construcția intervalelor de încredere provine în întregime din tabelul complet de verosimilitate.

În contrast, intervalele de probabilitate a posteriori Bayesiene dau într-adevăr probabilitatea ca valoarea parametrului necunoscut să fie în domeniul raportat. Adăugăm avertismentul uzual că aceasta depinde de alegerea specifică a unei a priori Bayesiene (posibil subiectivă).

Această distincție între cele 2 este subtilă deoarece intervalele de probabilitate a posteriori Bayesiene și intervalele de încredere frecvenționiste împart următoarele proprietăți:

1. Ele pleacă de la un model $f(x|\theta)$ pentru datele observate x cu parametrul necunoscut θ .
2. Date fiind datele x , ele dau un interval $I(x)$ specificând un domeniu de valori pentru θ .
3. Ele vin cu un număr (să zicem 0.95) care este probabilitatea a ceva.

În practică, mulți oameni interpretează greșit intervalele de încredere ca intervale de probabilitate Bayesiene, uitând că frecvenționiștii nu pun **nicio-dată** probabilități peste ipoteze (aceasta este analoaga greșelii că valoarea p în NHST este probabilitatea că H_0 este falsă). Dauna acestei interpretări greșite este cumva atenuată de faptul că, fiind date destule date și o a priori rezonabilă, intervalele Bayesian și frecvenționist adesea se dovedesc a fi destul de similare.

Pentru un exemplu amuzant ilustrând cât de diferite pot fi, vezi primul răspuns de aici (care implică fursecuri cu ciocolată!):

<http://stats.stackexchange.com/questions/2272/whats-the-difference-between-a->.

Acest exemplu folosește definiții formale și este în realitate despre mulțimi de încredere în locul intervalelor de încredere.