

Testimi i hipotezave dhe intervalet e besimit

Nga evidenca e kampionit te përfundimet e kujdesshme për popullatën

Ratiomera Statistics

Përmbajtja

Përmbledhja e faqes së kursit	2
Arsyetimi kryesor	2
Zgjidhe procedurën nga dizajni	3
Listë kontrolli për interpretimin	3
Referencë e zgjeruar	4
Qëllimi dhe bazat	4
Idetë kryesore	4
Udhëzuesi i formulave	5
Si lexohet figura shpjeguese	7
Lista e kontrollit për interpretim	8
Si lidhet kjo temë me të tjerat	8

Përmbledhja e faqes së kursit

Inferenca përdor një kampion për të mësuar rreth një popullate, duke e mbajtur të dukshme pasigurinë. Fillo me pyetjen kërkimore dhe dizajnin, emërto madhësinë e popullatës dhe vetëm pastaj zgjidh intervalin e besimit ose testin e hipotezës që përputhet me strukturën e të dhënave.

Arsyetimi kryesor

Tabela 1: Inferenca e besueshme lidh pyetjen kërkimore, dizajnin, llogaritjen dhe interpretimin.

Hapi	Pyetja që duhet bërë	Ideja kryesore
1. Përcakto	Cila madhësi e popullatës po studiohet?	Shkruaj parametrin dhe hipotezën zero e alternative para llogaritjes
2. Përputh	A janë vrojtimet të pavarura, të çiftuara apo numërime kategorike?	Dizajni përcakton gabimin standard dhe shpërndarjen referuese
3. Llogarit	Sa larg është vlerësimi nga vlera zero, matur në njësi të gabimit standard?	Statistika e testit e standardizon mospërputhjen e vrojtuar
4. Vlerëso	Sa e pazakontë do të ishte kjo statistikë nëse modeli zero do të ishte i vërtetë?	Vlera p është probabilitet i kushtëzuar nën modelin zero
5. Përfundo	Çfarë mbështet evidenca në nivelin e zgjedhur?	Shpreh pohimin për popullatën, pasigurinë, drejtimin dhe kufizimet e studimit

Për mesataren e kampionit me devijim standard të popullatës të vlerësuar, statistika për një kampion është

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}},$$

me $n - 1$ shkallë lirie. Intervali i besimit përdor të njëjtin vlerësim dhe të njëjtin gabim standard:

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}}.$$

Intervali jep një gamë mesataresh të popullatës që përputhen me të dhënat dhe modelin në nivelin e zgjedhur të besimit. Kuptimi i tij në kampione të përsëritura lidhet me metodën: në shumë kampione të krahasueshme, përqindja e deklaruar e intervaleve do ta përmbante parametrin fiks të popullatës.

Zgjidhe procedurën nga dizajni

Tabela 2: Numra që duken njësoj mund të kërkojnë procedura të ndryshme kur ndryshon struktura e varësisë mes tyre.

Struktura e të dhënave	Pyetja	Procedura që përdoret këtu
Një kampion sasior	A ndryshon mesatarja e një popullate nga një vlerë referuese?	Procedura t për një kampion
Dy grupe sasiore të pavarura	A ndryshojnë dy mesataret e popullatave?	Procedura t për kampione të pavarura
Dy matje në të njëjtat raste	A ndryshon nga zero mesatarja e popullatës e dallimeve brenda rastit?	Procedura t e çiftuar
Dy ndryshore kategorike	A janë ndryshoret të pavarura në popullatë?	Testi hi-katror i pavarësisë

Listë kontrolli për interpretimin

- Vlera p është probabiliteti i të dhënave të paktën kaq të papajtueshme me modelin zero, duke supozuar atë model dhe kushtet e tij. Nuk është probabiliteti që hipoteza zero është e vërtetë.
- Vendimi për ta refuzuar nuk është provë, ndërsa vendimi për të mos e refuzuar nuk është provë e barazisë.
- Domethënia statistikore nuk tregon madhësinë e efektit, rëndësinë praktike, cilësinë e matjes apo shkakësinë.
- Gabimi i llojit I do të thotë refuzim i një hipoteze zero të vërtetë. Gabimi i llojit II do të thotë mosrefuzim i një hipoteze zero të rreme. Fuqia është probabiliteti i refuzimit të hipotezës zero kur alternativa e përcaktuar është e vërtetë.
- Një kampion më i madh mund ta zvogëlojë gabimin standard dhe ta rrisë fuqinë, por nuk korrigjon kampionimin e anuar, matjen e dobët, varësinë që modeli e shpërfill ose një dizajn të dobët.

Tema 3 jep gjuhën inferenciale që përdoret në çdo temë të mëvonshme. Korrelacioni, regresioni, korrelacioni i pjesshëm, regresioni i shumëfishtë dhe ANOVA e ndryshojnë parametrin dhe modelin, por ruajnë të njëjtën disiplinë: përcakto pyetjen, respekto dizajnin, mat pasigurinë dhe mbaje përfundimin brenda shtrirjes së evidencës.

Referencë e zgjeruar

Qëllimi dhe bazat

Inferenca statistikore përdor informacionin nga një kampion për të mësuar rreth një popullate, duke pranuar njëkohësisht pasigurinë e kampionimit. **Popullata** është grupi i plotë i rasteve që përfshin pyetja kërkimore. **Kampioni** është nëngrupi që është vrojtuar. **Parametri** është një veçori numerike e popullatës, si mesatarja e popullatës μ . **Statistika** është madhësia përkatëse që llogaritet nga kampioni, si mesatarja e kampionit $\bar{x}$. Statistika dihet pasi mbledhen të dhënat; parametri zakonisht mbetet i panjohur.

Nëse do të merrej një kampion i ri i rastësishëm nga e njëjta popullatë, statistika e tij zakonisht do të ndryshonte. Shpërndarja hipotetike e asaj statistike nëpër kampione të përsëritura quhet **shpërndarja e kampionimit**. Devijimi i saj standard quhet **gabimi standard**. Gabimi standard mat ndryshueshmërinë e një vlerësimi nga një kampion në tjetrin. Ai nuk mat shpërhapjen e vrojtimeve individuale, e cila është detyra e devijimit standard të zakonshëm.

Inferenca varet nga më shumë se një formulë. Kampioni duhet të lidhet në mënyrë të besueshme me popullatën, vrojtimit duhet t'u përshtaten supozimeve të metodës për varësinë dhe matja duhet ta përfaqësojë ndryshoren e synuar. Një gabim i vogël standard nuk mund ta korrigjojë anshmërinë e përzgjedhjes, matjen e dobët ose një dizajn të papërshtatshëm. Fillo me pyetjen kërkimore dhe dizajnin e studimit, pastaj shqyrto statistikat përshkruese dhe vetëm atëherë zgjidh një procedurë inferenciale.

Elementi	Gjuha e kampionit	Gjuha e popullatës
Qendra	Mesatarja e kampionit $\bar{x}$	Mesatarja e popullatës μ
Përpjesëtimi	Përpjesëtimi i kampionit $\hat{p}$	Përpjesëtimi i popullatës p
Ndryshueshmëria e pikëzimeve	Devijimi standard i kampionit s	Devijimi standard i popullatës σ
Pasiguria e një vlerësimi	Gabimi standard i vlerësuar	Devijimi standard i shpërndarjes së kampionimit

Idetë kryesore

Intervali i besimit jep një diapazon vlerash të parametrut që pajtohen me vlerësimin dhe pasigurinë e tij të kampionimit sipas modelit. Niveli i besimit përshkruan si funksionon procedura në afat të gjatë. Nëse e njëjta metodë kampionimi dhe intervali do të përsëritej shumë herë, përpjesëtimi i deklaruar i intervaleve të krijuara do ta përmbante parametrin e pandryshueshëm të popullatës. Pasi llogaritet një interval i vetëm, parametri nuk lëviz mes vlerave; ajo që ka ndryshuar përmes kampionimit është intervali.

Testi i hipotezës fillon me një **hipotezë zero** H_0 , një pohim të saktë referues për një parametër të popullatës. **Hipoteza alternative** H_1 tregon drejtimin ose dallimin me interes përmbajtësor. Një statistikë testi mat sa larg gjendet vlerësimi i vrojtuar nga vlera zero, në njësi të gabimit standard. **Vlera p** është probabiliteti që, duke supozuar hipotezën zero dhe të gjitha kushtet e modelit, të merret një statistikë testi të paktën po aq e papajtueshme me hipotezën zero sa ajo e vrojtuar. Ajo nuk është probabiliteti që hipoteza zero të jetë e vërtetë.

Niveli i rëndësisë statistikore α është një prag vendimmarrjeje që zgjidhet para se të shihet rezultati. Nëse vlera p është më e vogël ose e barabartë me α , rezultati quhet statistikisht i rëndësishëm dhe H_0 hidhet poshtë. Nëse vlera p është më e madhe se α , analiza nuk e hedh poshtë H_0 . Moshedhja poshtë nuk provon se nuk ka efekt. Të dhënat mund të japin një vlerësim me saktësi të ulët, pra me pasiguri të lartë, efekti i vërtetë mund të jetë i vogël ose dizajni mund të ketë fuqi të kufizuar.

Realiteti dhe vendimi	Mos e hidh poshtë H_0	Hidhe poshtë H_0
H_0 është e vërtetë	Vendim i saktë për moshedhje poshtë	Gabimi i llojit I, me probabilitet të kontrolluar nga α
H_0 është e rreme	Gabimi i llojit II, i shënuar me β	Zbulim i saktë, me probabilitet të quajtur fuqi $1 - \beta$

Fuqia është probabiliteti që një test ta hedhë poshtë H_0 kur një alternativë e përcaktuar është e vërtetë. Ajo rritet kur efekti i vërtetë është më i madh, pikëzimet janë më pak të ndryshueshme, kampioni është më i madh ose rregulli i rëndësisë bëhet më pak i rreptë. Këto ndikime përfshijnë shkëmbime mes përfitimeve dhe kostove. Prandaj planifikimi kërkon një madhësi efekti me kuptim përmbajtësor dhe një dizajn të mbrojtshëm, jo kërkim të rëndësisë statistikore pasi janë mbledhur të dhënat.

Zgjedhja e procedurës ndjek strukturën e pyetjes kërkimore. Një procedurë për mesataren e një kampioni krahason një grup me një vlerë referuese. Një procedurë për grupe të pavarura krahason grupe të ndara. Një procedurë me çifte analizon matje të lidhura, si të njëjtët pjesëmarrës para dhe pas një ndërhyrjeje, duke punuar me dallimet brenda çifteve. Një procedurë hi-katror për një tabelë kontingjence krahason numërimet kategorike të vrojtura me numërimet e pritura sipas modelit zero. Në çdo rast, duhen deklaruar njësia e analizës dhe struktura e varësisë.

Udhëzuesi i formulave

Për një kampion të rastësishëm të pavarur, gabimi standard i vlerësuar i mesatares së kampionit është devijimi standard i kampionit i pjesëtuar me rrënjën katrore të madhësisë së kampionit:

$$SE(\bar{x}) = \frac{s}{\sqrt{n}}$$

Rrënja katrore shpjegon pse pasiguria ulet më ngadalë sesa rritet madhësia e kampionit. Shumëzimi i n me katër e përgjysmon këtë gabim standard kur ndryshueshmëria mbetet e njëjtë.

Një interval besimi lidh një vlerësim $\hat{\theta}$ me gabimin e tij standard dhe me një vlerë kritike c të zgjedhur për nivelin e besimit:

$$\hat{\theta} \pm c \cdot SE(\hat{\theta})$$

Nëse devijimi standard i popullatës σ dihet dhe zbatohet modeli normal i deklaruar, përdor referencën normale standarde:

$$\bar{x} \pm z_{1-\alpha/2} \frac{\sigma}{\sqrt{n}}, \quad z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}}.$$

Hipoteza alternative përcakton sipërfaqen referuese. Një alternativë e dyanshme përdor të dy bishtat përtej $|z|$ ose $|t|$. Një alternativë e njëanshme përdor bishtin drejtimor të përcaktuar paraprakisht. Zgjedhja e drejtimit pasi është parë rezultati nuk përbën një test të njëanshëm të përcaktuar paraprakisht.

Për mesataren e një kampioni me devijim standard të vlerësuar të popullatës, intervali përdor një vlerë kritike nga shpërndarja t me $n - 1$ shkallë lirie:

$$\bar{x} \pm t_{1-\alpha/2, n-1} \frac{s}{\sqrt{n}}$$

Statistika përkatëse e testit për një kampion e krahason mesataren e vrojtuar të kampionit me vlerën zero μ_0 :

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}}$$

Numëruesi është dallimi i vrojtuar nga pohimi zero. Emëruesi e kthen atë dallim në njësi të gabimit standard. Për matje me çifte, llogarit fillimisht nga një dallim d_i për secilin çift dhe pastaj zbato të njëjtin arsyetim të një kampioni për mesataren e dallimeve $\bar{d}$:

$$t = \frac{\bar{d} - 0}{s_d/\sqrt{n}}$$

Kjo e ruan çiftimin. Trajtimi i matjeve si të palidhura do ta hidhte poshtë informacionin se cilat dy vrojtime i përkasin njëra-tjetrës.

Për dy kampione të pavarura sipas modelit me varianca të barabarta të popullatave që mësohet këtu, fillimisht bashkoi dy variancat e kampioneve:

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}.$$

Pastaj llogarit

$$SE(\bar{x}_1 - \bar{x}_2) = s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}},$$

$$t = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{1/n_1 + 1/n_2}}, \quad df = n_1 + n_2 - 2.$$

Intervali përkatës i dyanshëm e zëvendëson numëruesin me

$$(\bar{x}_1 - \bar{x}_2) \pm t_{1-\alpha/2, n_1+n_2-2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$

Kushti i variancave të barabarta, emri i procedurës dhe llogaritja duhet të përputhen. Të dhënat me çifte kërkojnë procedurën me pikëzime të dallimeve.

Për një pyetje planifikimi me një kampion dhe me σ të njohur, përkufizo dallimin e standardizuar të popullatës

$$\delta = \frac{\mu - \mu_0}{\sigma}, \quad \text{Fuqia} = 1 - \beta.$$

Në modelin e njëanshëm z për planifikim që përdoret në materialin e dhënë, madhësia e kampionit që nevojitet për nivelin e rëndësisë α dhe fuqinë e synuar $1 - \beta$ është

$$n = \left(\frac{z_{1-\alpha} + z_{1-\beta}}{\delta} \right)^2.$$

Rrumbullakojë rezultatin lart. Kjo formulë i përket modelit të deklaruar dhe nuk është një rregull universal për madhësinë e kampionit. Fuqia rritet me madhësinë e kampionit dhe madhësinë e efektit, ndërsa ulet kur një nivel më i rreptë i rëndësisë e zhvendos kufirin e refuzimit më larg në bisht.

Për dy ndryshore kategorike, numërimi i pritur nën pavarësi në rreshtin i dhe kolonën j është

$$m_{ij} = \frac{n_{i.} n_{.j}}{n}.$$

Statistika hi-katror dhe shkallët e lirisë janë

$$\chi^2 = \sum_i \sum_j \frac{(n_{ij} - m_{ij})^2}{m_{ij}}, \quad df = (k-1)(l-1).$$

Për një tabelë dy me dy, madhësia e koeficientit phi është

$$|\phi| = \sqrt{\frac{\chi^2}{n}}.$$

Përafrimi që përdoret në këtë rrjedhë mësimore kërkon një kampion të thjeshtë të rastësishëm dhe numërime të pritshme më të mëdha se 5 në secilën qelizë. Një χ^2 e madhe numëron kundër pavarësisë. Ajo nuk është evidencë për pavarësi.

Struktura e pyetjes	Procedura e paraqitur	Madhësia që analizohet
Një kampion kundrejt një reference	Procedura z ose t për një kampion	Mesatarja e një kampioni
Dy grupe të ndara	Procedura t me variancë të përbashkët për kampione të pavarura	Dallimi mes mesatareve të grupeve
Dy matje të lidhura	Procedura t me çifte	Mesatarja e dallimeve brenda çifteve
Grupe të pavarura me një pyetje të bazuar në rangje	Procedura e shumës së rangjeve të Wilcoxon-it	Rangjet relative mes grupeve
Vrojtime me çifte me një pyetje të bazuar në rangje	Procedura e rangjeve me shenjë të Wilcoxon-it	Rangjet me shenjë të dallimeve të çiftuara
Dy ndryshore kategorike	Procedura hi-katror e pavarësisë	Numërimet e vrojtueshme kundrejt atyre të pritshme në qeliza

Si lexohet figura shpjeguese

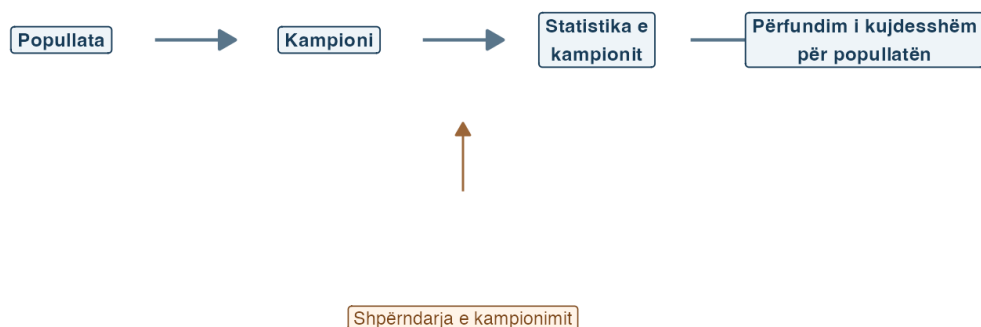


Figura 1: Një rrjedhë horizontale kalon nga popullata te kampioni, te statistika e kampionit dhe pastaj te një përfundim i kujdesshëm për popullatën, ndërsa shpërndarja e kampionimit ushqen statistikën.

Lexoje vijën kryesore nga e majta në të djathtë. Popullata është synimi i pyetjes kërkimore. Kampioni është pjesa që bëhet e vrojtueshme. Një statistikë kampioni e përmbledh evidencën përkatëse, si

një mesatare kampioni, dallim, përpjesëtim ose lidhje. Kutia e fundit quhet qëllimisht përfundim i kujdesshëm, sepse kthimi nga kampioni te popullata nuk është kurrë automatik.

Shigjeta që ngrihet nga shpërndarja e kampionimit është ura që siguron probabiliteti. Ajo përfaqëson si do të ndryshonte statistika nëpër kampione të përsëritura sipas supozimeve të deklaruara. Intervali i besimit e përdor këtë ndryshueshmëri për të treguar saktësinë. Testi e krahason statistikën e vërtetuar me sjelljen e kampionimit që pritet nën H_0 . Figura nuk nënkupton se një kampion i madh garanton përgjithësim. Metoda e kampionimit, matja, të dhënat që mungojnë, varësia dhe dizajni i studimit vazhdojnë të përcaktojnë se cili përfundim për popullatën mund të mbrohet.

Ndarja mes «statistikës së kampionit» dhe «përfundimit për popullatën» është një pikë e dobishme ndalimi. Para se ta kalosh, pyet nëse gabimi standard pasqyron dizajnin real, nëse supozimet e procedurës janë të arsyeshme dhe nëse formulimi i përfundimit përputhet me atë që është testuar. Një rezultat mund të mbështetë një lidhje ose dallim pa vërtetuar një efekt shkakor.

Lista e kontrollit për interpretim

Trego popullatën, kampionin, parametrin, statistikën dhe njësinë e analizës. Përshkruaj si hynë rastet në kampion. Shqyrto të dhënat dhe identifikoj vlerat që mungojnë ose vërtetimet e pazakonta. Përshtate procedurën me shkallën e ndryshores së rezultatit dhe me të dhënat e pavarura, me çifte ose kategorike. Shkruaj H_0 dhe H_1 me fjalë dhe simbole. Raporto vlerësimin dhe intervalin e besimit krahas statistikës së testit, shkallëve të lirisë kur zbatohen, vlerës p dhe një interpretimi në kontekst.

Mos e kthe vlerën p në probabilitet se një hipotezë është e vërtetë. Mos e përdor rëndësinë statistikore si sinonim të rëndësisë praktike. Krahaso madhësinë dhe pasigurinë e vlerësimit me pyetjen kërkimore. Kur një rezultat nuk është statistikisht i rëndësishëm, diskuto intervalin dhe saktësinë e tij në vend që të deklaroj se grupet janë të barabarta. Kur kryhen disa teste, prano se mund të rritet mundësia për të paktën një gabim të llojit I dhe përdor një qasje të planifikuar për shumëfishësinë kur kërkohet.

Si lidhet kjo temë me të tjerat

Probabiliteti siguroi shpërndarjet e kampionimit që i bëjnë të mundshme intervalet e besimit dhe testet. I njëjti model inferencial tani shoqëron çdo koeficient të mëvonshëm. Korrelacioni ka gabim standard dhe test. Pjerrësia e regresionit ka vlerësim, interval dhe vlerë p . Korrelacioni i pjesshëm dhe secili koeficient i regresionit të shumëfishtë interpretohen në mënyrë të kushtëzuar. Analiza e variancës përdor një statistikë F për të krahasuar ndryshueshmërinë që lidhet me modelin me ndryshueshmërinë reziduale.

Prandaj inferenca nuk është një ritual i veçantë që i shtohet fundit të një analize. Ajo është një urë e disiplinuar nga evidenca përshkruese e kampionit të një pohim i kufizuar për popullatën. Mbajtja së bashku e vlerësimit, pasigurisë, dizajnit dhe kuptimit përmbajtësor e bën atë urë të besueshme.

ID-ja e dokumentit: topic-03-hypothesis-testing-summary-sq

Shënime 1 nga 3

Testimi i hipotezave dhe intervalet e besimit | Tema 3

Shënime 2 nga 3

Testimi i hipotezave dhe intervalet e besimit | Tema 3

Shënime 3 nga 3

Testimi i hipotezave dhe intervalet e besimit | Tema 3