12.

Testiranje hipoteza

Sadržaj poglavlja

1. Hipoteze. Kriteriji i pogreške odlučivanja
2. Testiranje parametarskih hipoteza
3. Usporedbe dviju populacija
4. Testovi prilagodbe razdiobama

12.1. Hipoteze. Kriteriji i pogreške odlučivanja

Razmatranje u ovom poglavlju započet ćemo sljedećim primjerom

Primjer 12.1.

U specifikacijama pisača navodi se brzina ispisa od 50 stranica u minuti. Pisač je testiran na ispisu različitih materijala. Vrijeme ispisa u minutama podijeljeno je s brojem stranica, da se dobije prosječan ispis u minuti.

Na uzorku od 30 ispisa, dobivena je srednja vrijednost 46 stranica, uz standardnu devijaciju $s = 6$.

Odgovaraju li svojstva pisača specifikaciji?

► Ovdje se radi o provjeri **hipoteze**. Tvrdnju proizvođača bismo prihvatili ako mjerenja pokažu da za srednju brzinu ispisa vrijedi $a = 50$ (ili više).

Na temelju rezultata dobivenih u uzorku ne možemo nikad biti sasvim sigurni je li ponuđena hipoteza ispravna ili ne. Svaki naš zaključak može biti donesen samo s određenom vjerojatnošću, koja je iskazana kroz **nivo pouzdanosti**.

Pravilo je sljedeće: da bismo odbacili neku hipotezu, vjerojatnost pogrešne odluke mora bili malena. Obično se kao prag tolerancije za pogrešku uzima $\alpha = 0.05$ ili čak $\alpha = 0.01$.

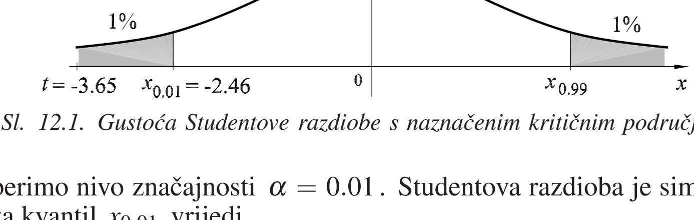
U uvodnom primjeru, dobivena srednja vrijednost 46 razlikuje se od specifikacije za $46-50 = -4$. Je li ta razlika zanemariva posljedica slučajnosti, ili je ona statistički značajna (signifikantna)?

Odgovorimo na ovo pitanje. Da bismo donijeli kriterij odlučivanja, nužno nam je definirati **test**.

U tu svrhu promotrit ćemo slučajnu varijablu

$$T = \frac{\overline{X} - a}{S/\sqrt{n}}.$$

Ona ima Studentovu razdiobu s $n-1$ stupnjeva slobode. Graf gustoće te funkcije dan je na sljedećoj slici



Sl. 12.1. Gustoća Studentove razdiobe s naznačenim kritičnim područjem.

Izaberimo nivo značajnosti $\alpha = 0.01$. Studentova razdioba je simetrična, pa za kvantil $x_{0.01}$ vrijedi

$$x_{0.01} = -x_{0.99}.$$

Iz tablica Studentove razdiobe iščitava se vrijednost ovog kvantila, za 29 stupnjeva slobode: $x_{0.99} = 2.46$.

To znači da za slučajnu varijablu T vrijedi

$$\boldsymbol{P}(T > -2.46) = 0.99,$$

ukoliko je hipoteza $a = 50$ istinita.

U konkretnom primjeru dobivamo sljedeću vrijednost varijable T:

$$t = \frac{\bar{x} - a}{\hat{s}/\sqrt{n}} = \frac{46-50}{6/\sqrt{30}} = -3.65$$

Ova je razlika veća (po apsolutnom iznosu) od dopuštene vrijednosti $x_{0.01} = -2.46$. Dobivena realizacija varijable T nalazi se izvan 99%-tnog intervala povjerenja. Zato je razumno *odbaciti hipotezu* $a \geqslant 50$. Vjerojatnost pogreške koju smo pri tom možda načinili manja je od nivoa signifikantnosti $\alpha = 0.01$.

Da se slučajno zbilo $t > x_{0.01}$ tad hipotezu o brzini ispisa $a = 50$ *ne bismo mogli odbaciti* (uz taj nivo pouzdanosti). ◄

12.1.1. Snaga testa

Neka je Λ skup svih mogućih parametara ϑ. **Hipoteza** se najopćenitije definira kao *podskup* od Λ.

Prava vrijednost parametra nije nam poznata. Zato ne možemo sa sigurnošću utvrditi je li hipoteza o vrijednosti tog parametra istinita ili ne.

Test će pružiti kriterij odlučivanja. Idealan test bi utvrdio da je hipoteza istinita onda i samo onda kad parametar zaista pripada području pokrivenom tom hipotezom.

Snaga testa

Snaga testa je funkcija $S : \Lambda \to [0,1]$ definirana ovako:

$$S(\vartheta) = \boldsymbol{P}(\{\text{prihvaćena je alternativa } H_1\})$$

Da bismo objasnili značenje ovog pojma, ilustrirajmo ga na primjerima.

Neka je A događaj čija je vjerojatnost p nepoznata. (Nepoznati parametar ϑ ovdje je označen s p.) Postavili smo hipotezu:

$$H_0 \quad \ldots \quad \text{vjerojatnost } p \text{ manja je od } 0.2\,.$$

Alternativa ovoj hipotezi može biti npr.

$$\textbf{1.} \qquad H_1 \quad \ldots \quad \text{vrijedi } p \geqslant 0.2\,.$$

Alternativa ne mora biti komplementna hipotezi. Tako npr., ona može glasiti

$$\textbf{2.} \qquad H_1 \quad \ldots \quad \text{vrijedi } p \geqslant 0.5\,,$$

ili čak se može "preklapati" s hipotezom:

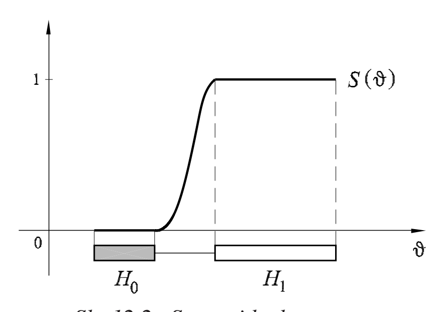
$$\textbf{3.} \qquad H_1 \quad \ldots \quad \text{vrijedi } p \geqslant 0.1\,,$$

Prava vrijednost parametra p nije nam poznata. Moguće vrijednosti su svi brojevi iz intervala $[0,1]$. Zato je $\Lambda = [0,1]$

Idealan test u ovoj situaciji imao bi sljedeću snagu:

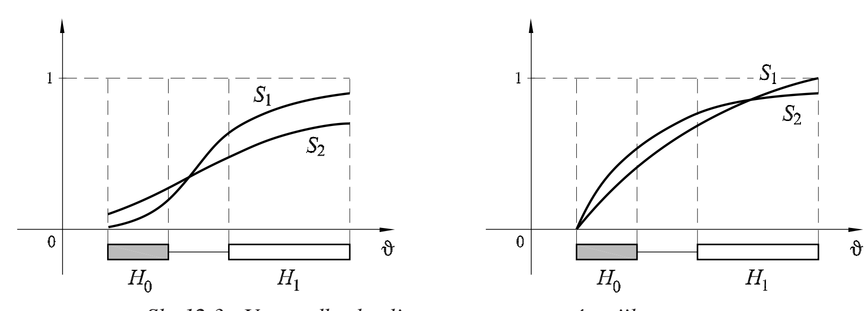
$$S(p) = \begin{cases} 0, & p \in H_0, \\ 1, & p \in H_1 \end{cases}$$

U slučaju **1.** snaga idealnog test bila bila bi stepenasta funkcija. U slučaju **3.** idealni test ne postoji, jer se hipotreze dijelom preklapaju. U slučaju **2.**, primjer snage idealnog testa dan je na sljedećoj slici:



Sl. 12.2. Snaga idealnog testa

U praksi, ovakav je test nemoguće uspostaviti. U odnosu na njihovu snagu, ponekad je moguće uspoređivati kvalitetu testova.



Sl. 12.3. Usporedba kvalitete testova pomoću njihove snage

Na slici lijevo prikazane su snage dvaju testova od kojih je prvi bolji od drugog. Naime, funkcija snage manja je na području prihvaćanja hipoteze H_0 a veća je na području prihvaćanja alternative H_1.

Testovi čije su snage prikazane na slici desno nisu usporedljivi.

Snaga testa određuje i pogreške koje se mogu načiniti prihvaćanjem pogrešne odluke.

Pogreška prve i druge vrste

Definirajmo

$$\alpha = \sup_{\vartheta \in H_0} S(\vartheta).$$

Ovaj broj označava maksimalnu vjerojatnost da je prihvaćena alternativa H_1, iako je istinita H_0. Broj α naziva se **pogreška prve vrste**.

Na isti način, broj

$$\beta = \sup_{\vartheta \in H_1} [1 - S(\vartheta)]$$

definira **pogrešku druge vrste**. Ona daje maksimalnu vjerojatnost da je prihvaćena hipoteza H_0, iako je istinita H_1.

Pogreška prve vrste nastoji se ograničiti. U praktičnim primjerima, zadaje se obično $\alpha = 0.05$.

Primjer 12.2. ■ **Snaga testa**

Populacija se sastoji od velikog broja istovjetnih proizvoda. Svaki od njih može biti loš, s vjerojatnošću p, ili dobar, s vjerojatnošću $1-p$.

Ta nam je vjerojatnost nepoznata. Postavljamo hipotezu:

$$H_0 \quad \ldots \quad p \leqslant 0.2\,.$$

Neka njoj alternativa bude

$$H_1 \quad \ldots \quad p \geqslant 0.6\,.$$

Da bismo provjerili je li nulta hipoteza istinita, moramo definirati test. Izbor različitih testova praktički je neograničen. Opredijelili smo se za ovaj:

Test 1. Izaberemo dva proizvoda iz populacije. Ako su oba dobra, prihvaća se hipoteza H_0. Ako su oba loša, odbacuje se H_0 (t.j. prihvaća se alternativa H_1). Ako je jedan dobar i jedan loš, izabire se novi proizvod i prihvaća H_0 ukoliko je on dobar. U protivnom se prihvaća H_1.

Odredi snagu ovog testa i pogreške prve i druge vrste.

► Snaga testa jednaka je vjerojatnosti da ćemo prihvatiti alternativu H_1, ukoliko je vrijednost parametra jednaka p.

Situacija u kojima će se to dogoditi je sljedeća: oba proizvoda su loša, ili je jedan među njima loš i onda i treći izabrani ponovo loš. Vjerojatnost da se to dogodi je:

$$S(p) = p^2 + 2 \cdot p(1-p) \cdot p = p^2(3-2p).$$

Nacrtajmo graf ove funkcije. Promatramo je samo na intervalu $[0,1]$. Na tom je intervalu to rastuća funkcija. Njezini su ekstremi u točkama $x = 0$ (minimum) $m(0,0)$ i $p = 1$ (maksimum) $M(1,1)$. Točka infleksije je $(\frac{1}{2}, \frac{1}{4})$. Od bitne važnosti je i točka $A(0.2, 0.104)$.

Područje na kojim bi morala biti prihvaćena hipoteza H_0 je interval $[0, 0.2]$. Maksimum funkcije S na tom intervalu postiže se upravo u točki A. Zato je pogreška prve vrste $\alpha = 0.104$.

Pogreška druge vrste jednaka je

$$\beta = 1 - S(0.6) = 0.352.$$

Izbor testa mora voditi računa o veličini ovih pogrešaka. U ovom je testu pogreška druge vrste neprihvatljivo velika. Međutim, smanjenje jedne pogreške najčešće ima za posljedicu povećanje druge. ◄

Primjer 12.3.

Za situaciju opisanu u prethodnom primjeru, predloženi se test može napisati na ovaj način:

Test 1. Prihvaća se hipoteza H_0 ako među tri izabrana proizvoda bude najviše jedan loš. U protivnom, prihvaća se alternativa H_1.

Poopćimo taj test na sljedeći:

Test 2. Prihvaća se hipoteza H_0 ako među n izabranih proizvoda bude najviše jedan loš. U protivnom, prihvaća se alternativa H_1.

Odredimo n tako da odnos pogreški prve i druge vrste bude optimalan.

► Vjerojatnost prihvaćanja hipoteze H_0 je:

$$(1-p)^n + n(1-p)^{n-1}p = (1-p)^{n-1}(1+(n-1)p).$$

Zato je

$$S(p) = 1 - (1-p)^{n-1}(1+(n-1)p).$$

Derivacija ove funkcije je pozitivna na intervalu $[0,1]$, pa je ondje ta funkcija rastuća. Vrijedi $S(0) = 0$ i $S(1) = 1$.

Kritične vrijednosti za pogrešku prve i druge vrste dobivaju se u točkama $p = 0.2$ i $p = 0.6$:

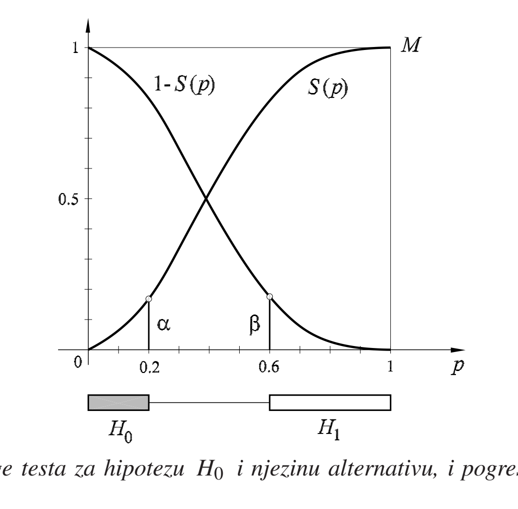
$$\alpha = S(0.2) = 1 - 0.8^{n-1}(1 + 0.2(n-1)),$$
$$\beta = 1 - S(0.6) = 0.4^{n-1}(1 + 0.6(n-1)).$$

Iznosi tih pogrešaka za nekoliko početnih vrijednosti broja n dane su u ovoj tablici:

n	α	β
2	0.040	0.640
3	0.104	0.352
4	0.181	0.179
5	0.263	0.087
6	0.345	0.041

Za test ovog tipa, izbor $n = 4$ daje najuravnoteženije pogreške prve i druge vrste.

Na sljedećoj slici nacrtan je graf funkcije snage testa za $n = 4$ i naznačene pogreške prve i druge vrste.



Sl. 12.4. Snage testa za hipotezu H_0 i njezinu alternativu, i pogreške obiju vrsta.

Nakon što upoznamo tehniku odlučivanja, u nastavku ćemo se ponovo vratiti na ovaj problem i ponuditi povoljniji test. ◄

12.2. Testiranje parametarskih hipoteza

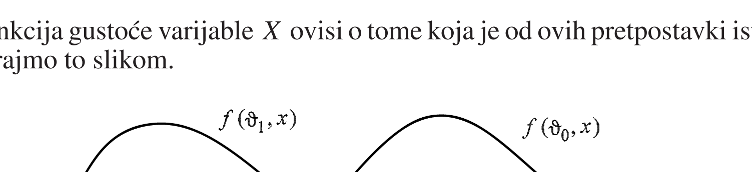
Pretpostavit ćemo da razdioba slučajne varijable ovisi o nepoznatom parametru ϑ. O tom parametru mogu postojati neke pretpostavke. Sa H_0 označit ćemo **primarnu hipotezu**

$$H_0 \quad \dots \quad \vartheta = \vartheta_0.$$

Njoj oprečna je **alternativa**, na primjer:

$$H_1 \quad \dots \quad \vartheta = \vartheta_1.$$

Funkcija gustoće varijable X ovisi o tome koja je od ovih pretpostavki istinita. Ilustrirajmo to slikom.



Sl. 12.5. Gustoća razdioba za svaku od hipoteza.

U ovisnosti o izabranom testu, treba prihvatiti neku od hipoteza.

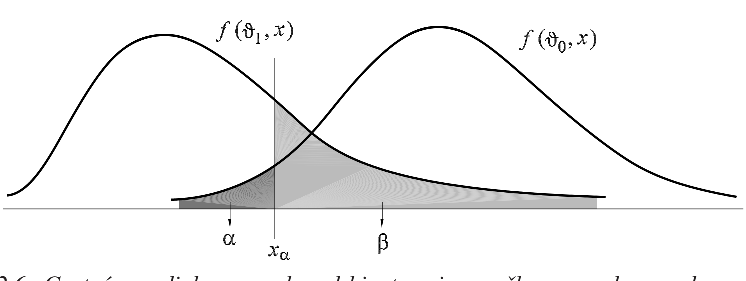
Mogu se zbiti četiri situacije:

1. Prihvaćena je istinita hipoteza H_0.
2. Odbačena je istinita hipoteza H_0 (a prihvaćena neistinita alternativa H_1.)
3. Prihvaćena je neistinita hipoteza H_0.
4. Prihvaćena je istinita hipoteza H_1.

Slučajevi **2.** i **3.** predstavljaju pogreške koje smo već analizirali. Iskazat ćemo ih s naglaskom na hipotezu H_0.

Pogreška I. vrste: Odbačena je istinita hipoteza H_0.

Pogreška II. vrste: Prihvaćena je neistinita hipoteza H_0



Sl. 12.6. Gustoća razdioba za svaku od hipoteza i pogreške prve odnosno druge vrste.

Odabir testa kod parametarskih hipoteza iznimno je jednostavan. U situaciji s prethodne slike test glasi ovako:

Test

Izabere se kritična vrijednost x_α.

Ukoliko je realizacija statistike veća od x_α, prihvaća se hipoteza H_0. (Pri tom je možda učinjena greška II. vrste maksimalnog iznosa β.)

Ukoliko je realizacija statistike manja od x_α, odbacuje se hipoteza H_0. (Pri tom je možda učinjena greška I. vrste maksimalnog iznosa α.)

O načinu izbora kritične vrijednosti x_α bit će riječi u nastavku.

Primjer 12.4.

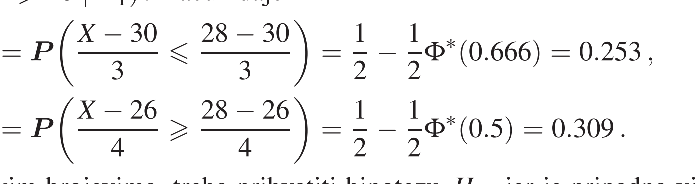
Prosječna masa jabuka prve kvalitete iznosi $a_0 = 30$ dag, uz standardnu devijaciju $s_0 = 3$ dag. Prosječna masa jabuka druge kvalitete je $a_1 = 26$ dag, uz standardnu devijaciju $s_1 = 4$ dag.

Izabrana je jabuka iz nasumce odabrane skupine. Njezina je masa $a = 28$ dag.

Kolike su vjerojatnosti pogrešaka prve i druge vrste ako prihvatimo hipotezu da jabuka pripada prvoj, odnosno drugoj skupini?

► Hipoteza i njezina alternativa su:

$$H_0 = \{\text{jabuka je prve kvalitete}\},$$
$$H_1 = \{\text{jabuka je druge kvalitete}\}.$$



Sl. 12.7. Gustoća razdioba za svaku od hipoteza.

Prirodno je pretpostaviti da je razdioba populacije normalna. Ukoliko je hippoteza H_0 istinita, riječ je o razdiobi $\mathcal{N}(30, 9)$. Ako je istinita alternativa, riječ je o razdiobi $\mathcal{N}(26, 16)$.

Sad je potrebno izračunati vjerojatnosti $\alpha = \boldsymbol{P}(X \leqslant 28 \mid H_0)$ i $\beta = \boldsymbol{P}(X \geqslant 28 \mid H_1)$. Račun daje

$$\alpha = \boldsymbol{P}\Big(\frac{X-30}{3} \leqslant \frac{28-30}{3}\Big) = \frac{1}{2} - \frac{1}{2}\Phi^*(0.666) = 0.253,$$
$$\beta = \boldsymbol{P}\Big(\frac{X-26}{4} \geqslant \frac{28-26}{4}\Big) = \frac{1}{2} - \frac{1}{2}\Phi^*(0.5) = 0.309.$$

Prema ovim brojevima, treba prihvatiti hipotezu H_1, jer je pripadna vjerojatnost veća. Dakako, hipoteza H_0 još uvijek može biti istinita. Pogreška prve vrste iznosi 0.253.

Da smo prihvatili hipotezu H_0, pogreška prve vrste bi bila 0.309. ◄

12.2.1. Izbori alternative. Utjecaj na grešku II. vrste

U parametarskim hipotezama zadana nam je vrijednost parametra koja odgovara temeljnoj hipotezi H_0:

$$H_0 \quad \dots \quad h = \vartheta_0.$$

Postoje tri uobičajena načina za odabir alternative:

1. $\vartheta = \vartheta_1 > \vartheta_0,$
2. $\vartheta = \vartheta_1 < \vartheta_0,$
3. $\vartheta = \vartheta_1 \neq \vartheta_0.$

Alternative **1.** i **2.** su **jednostrane alternative**, a **3.** je **dvostrana alternativa**. U testovima ovog tipa moguće je utvrditi pogrešku I. i II. vrste.

Ukoliko vrijednost parametra koja odgovara alternativi nije određena, tada alternativa npr. glasi

$$\vartheta > \vartheta_0$$

i pogrešku druge vrste nije moguće jednostavno utvrditi. U testovima s ovakvim alterbativama zadaje se samo pogreška prve vrste.

Prihvaćanje odnosno odbacivanje hipoteze H_0 temelji se na sljedećem testu:

Testiranje parametarskih hipoteza — jednostrani testovi

Zadaje se nivo značajnosti α. Na temelju poznate razdiobe uz pretpostavljenu istinitost hipoteze H_0 odrede se

1. Kvantil $x_{1-\alpha}$ (za jednostranu alternativu $\vartheta > \vartheta_0$), ili

2. Kvantil x_α za jednostranu alternativu $\vartheta < \vartheta_0$,

Hipoteza H_0 se odbacuje ukoliko vrijednost $\widehat{x}$ varijable izračunate iz uzorka padne van intervala povjerenja reda $1-\alpha$:

$$\widehat{x} > x_{1-\alpha},$$

u prvom slučaju, odnosno

$$\widehat{x} < x_\alpha,$$

u drugom slučaju. U protivnom se hipoteza ne može odbaciti (t.j. prihvaća se).

Za dvostrane hipoteze, način izbora kvantila i područja odbacivanja se mijenja:

Testiranje parametarskih hipoteza — dvostrani test

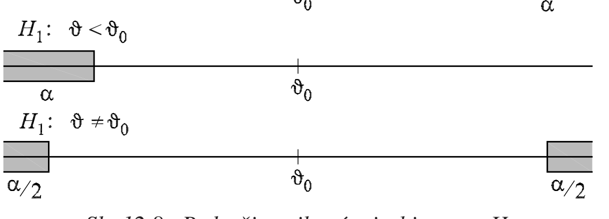
Zadaje se nivo značajnosti α. Na temelju poznate razdiobe uz pretpostavljenu istinitost hipoteze H_0 odrede se kvantili $x_{\alpha/2}$ i $x_{1-\alpha/2}$. za dvostranu alternativu $\vartheta \neq \vartheta_0$.

Hipoteza H_0 se odbacuje, ukoliko vrijednost $\widehat{x}$ varijable izračunate iz uzorka padne van intervala povjerenja reda $1-\alpha$:

$$\widehat{x} > x_{1-\alpha/2} \qquad \text{ili} \qquad \widehat{x} < x_{\alpha/2}.$$

U protivnom se hipoteza ne može odbaciti (t.j. prihvaća se).

Ove su situacije ilustrirane na sljedećoj slici:



Sl. 12.8. Područje prihvaćanja hipoteze H_0.

Račun koji treba načiniti pri provjeri hipoteza nalikuje onom u određivanju intervala povjerenja. Uvjerimo se u to na primjeru testova koji slijede.

12.2.2. U-test: nepoznato očekivanje i poznata disperzija

Razdioba populacije je normalna, s poznatom disperzijom σ^2. Hipoteza se odnosi na vrijednost očekivanja:

$$H_0 \quad \dots \quad a = a_0$$

Statistika na temelju koje će se napraviti test je

$$U = \frac{\overline{X} - a_0}{\sigma/\sqrt{n}}$$

Ako je hipoteza H_0 istinita, onda je razdioba ove statistike normalna, $U \sim \mathcal{N}(0,1)$.

U-test

Zadaje se nivo značajnosti α. Na temelju njega, odredi se kvantil

1. $u_{1-\alpha}$, u slučaju jednostranih alternativa, te

2. $u_{1-\alpha/2}$, u slučaju dvostranih alternativa.

Izračuna se vrijednost $\widehat{u}$ statistike dobivena iz uzorka. Test glasi:

1. Ako je $\widehat{u} > u_{1-\alpha}$ (za desnu alternativu), ili $\widehat{u} < -u_{1-\alpha}$ (za lijevu alternativu), hipoteza H_0 se odbacuje.

2. Ako je $|\widehat{u}| > u_{1-\alpha/2}$ (za dvostranu alternativu), hipoteza H_0 se odbacuje.

U protivnom se ta hipoteza ne može odbaciti.

Primjer 12.5.

Gradska je uprava dobila ponudu za nabavu jeftinijih žarulja za javnu rasvjetu. Srednji vijek trajanja postojećih žarulja je 1200 sati, uz standardnu devijaciju od 150 sati. Gradska uprava će odabrati na novi tip žarulja, osim ako se uz nivo značajnosti $\alpha = 0.05$ pokaže da su lošije kvalitete.

Testirano je 100 žarulja i dobivena je srednja vrijednost $\widehat{x} = 1160$. Pretpostavlja se da je standardna devijacija nepromijenjena. Hoće li se uprava odlučiti za kupovinu novih žarulja?

► Hipoteza koju testiramo je

$$H_0 \quad \dots \quad a = a_0 = 1200.$$

Njoj je alternativa

$$H_1 \quad \dots \quad a = a_1 < a_0.$$

Riječ je o jednostranom testu sa zadanom pogreškom prve vrste $\alpha = 0.05$. Razdioba koja odgovara hipotezi H_0 je $\mathcal{N}(1200, 150^2)$. U tom slučaju, slučajna varijabla

$$\frac{\overline{X} - a_0}{\sigma/\sqrt{n}}$$

ima jediničnu normalnu razdiobu.

Za vrijednost dobivenu iz uzorka je

$$\widehat{u} = \frac{\widehat{x} - a_0}{\sigma/\sqrt{n}} = \frac{1160 - 1200}{150}\sqrt{100} = -2.67.$$

Kritična vrijednost za zadani nivo značajnosti je $u_\alpha = -u_{1-a} = -u_{0.95} = -1.645$.

Dobivena vrijednost je ispod kritične, pa se hipoteza H_0 treba odbaciti. Nove su žarulje lošije kvalitete. ◄

Primjer 12.6.

Uređaj za pakiranje šećera radi sa standardnom devijacijom 2 grama. Povremeno se obavlja kontrola s ciljem korekcije uređaja ukoliko neto sadržaj ne odgovara nominali. Uzorak od 100 paketa dao je srednju vrijednost $\overline{x} = 999$ grama. Provjeri uz $\alpha = 0.01$ hipotezu o ispravnosti uređaja.

► Hipoteza o ispravnosti je

$$H_0 \quad \dots \quad a = a_0 = 1000.$$

Alternativa je dvostrana:

$$H_1 \quad \dots \quad a \neq a_0.$$

Zato računamo kritični kvantil

$$u_{1-\alpha/2} = u_{0.995} = 2.58.$$

Vrijednost statistike je

$$\widehat{u} = \frac{\overline{x} - a_0}{\sigma/\sqrt{n}} = \frac{999 - 1000}{2}\sqrt{100} = -5 < u_{\alpha/2}.$$

Hipoteza o ispravnosti mora se odbaciti. ◄

12.2.3. T-test: nepoznato očekivanje i nepoznata disperzija

Razdioba populacije je normalna, s nepoznatom disperzijom σ^2. Hipoteza se odnosi na vrijednost očekivanja:

$$H_0 \quad \dots \quad a = a_0$$

Statistika na temelju koje će se napraviti test je

$$T = \frac{\overline{X} - a_0}{S/\sqrt{n}}$$

Ako je hipoteza H_0 istinita, onda je razdioba ove statistike Studentova, s $n-1$ stupnjem slobode.

Studentova je razdioba simetrična, pa su i kvantili simetrični:

$$t_\alpha = -t_{1-\alpha}.$$

Zato je oblik testa identičan onom za normalnu razdiobu, s tim da se kvantili u_α normalne razdiobe zamijene s kvantilima t_α Studentove razdiobe s $n-1$ stupnjem slobode.

Primjer 12.7.

Centar za medicinska istraživanja najavio je da raspolaže s takvim načinom liječenja visokog krvnog pritiska kojim se ovaj može smanjiti za 20 jedinica. Liječnik je isprobao postupak na 10 pacijenata i ustanovio da je prosječno smanjenje pritiska $\overline{x} = 18$ jedinica, uz odstupanje $\widehat{s} = 4.2$. Provjeri uz nivo značajnosti $\alpha = 0.05$ je li tvrdnja o postupku istinita.

► Radi se o jednostranoj hipotezi uz nepoznatu disperziju. Temeljna je hipoteza

$$H_0 \quad \dots \quad a = a_0 = 20,$$

a njezina alternativa

$$H_1 \quad \dots \quad a < a_0.$$

Slučajna varijabla

$$T = \frac{\overline{X} - a_0}{S}\sqrt{n}$$

ima Studentovu razdiobu s 9 stupnjeva slobode. Odgovarajući kvantil je

$$t_{0.05} = -t_{0.95} = -1.83.$$

Vrijednost dobivena iz uzorka, uz istinitost hipoteze H_0, je

$$\widehat{t} = \frac{\overline{x} - a_0}{\widehat{s}}\sqrt{n} = \frac{18 - 20}{4.2}\sqrt{10} = -1.51.$$

Ova je vrijednost veća od kritične. Zato se hipoteza ne može odbaciti. Ako bismo odbacili hipotezu, učinjena pogreška mogla bi biti veća od 0.05. ◄

12.2.4. Hipoteza o proporciji

Nepoznata vjerojatnost p nekog svojstva populacije pokušava se odrediti na temelju uzorka. Ako u uzorku od n elemenata m od njih ima to svojstvo, tada za p uzimamo vrijednost m/n.

Osnovna hipoteza je

$$H_0 \quad \ldots \quad p = p_0$$

a alternative biramo, na primjer, u obliku

$$H_1 \quad \ldots \quad p = p_1 > p_0$$

(desna) jednostrana alternativa.

Vrijednost m ima binomnu razdiobu $\mathscr{B}(n,p) \approx \mathscr{N}(np, npq)$. Zato slučajna varijabla $X = \dfrac{m}{n}$ ima razdiobu $X \approx \mathscr{N}\Big(p, \dfrac{pq}{n}\Big)$.

Pretpostavimo da je H_0 istinita i odaberimo statistiku

$$U = \frac{\dfrac{m}{n} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} = \Big(\frac{m}{n} - p_0\Big)\sqrt{\frac{n}{p_0 q_0}}.$$

Ukoliko je H_0 istinita, ova statistika ima približno razdiobu $\mathscr{N}(0,1)$.

Hipotezu H_0 ćemo odbaciti, ukoliko se ostvari $\widehat{u} > u_{1-\alpha}$.

Za slučaj lijeve jednostrane hipoteze, ili dvostrane hipoteze, ovaj se test korigira na istovjetan način kao u U-testu.

Primjer 12.8.

Uprava velike pivovare razmatra prijedlog da se na tržište plasira veleboca piva sadržine 2.5 L. U proizvodnju će se krenuti ako najmanje 60% potrošača to želi. Napravljena je anketa među 200 potrošača, od kojih se 95 izjasnilo u prilog novoj ambalaži. Uz nivo značajnosti $\alpha = 0.05$ utvrdite hoće li se pokrenuti ta proizvodnja.

► Ovdje je $n = 200$, $m = 95$, $p_0 = 0.6$. Alternativa je *lijeva jednostrana*, $p < p_0$. Statistika poprima vrijednost:

$$U = \Big(\frac{m}{n} - p_0\Big)\sqrt{\frac{n}{p_0 q_0}} = -3.61.$$

Kritična vrijednost kvantila je

$$u_\alpha = -u_{1-\alpha} = -u_{0.95} = -1.645\,.$$

Hipoteza H_0 mora se odbaciti. Veleboca će pričekati. ◄

12.2.5. * Snaga U-testa

Ponovimo osnovne pretpostavke o U-testu.

Osnovna populacija ima normalnu razdiobu $\mathscr{N}(a, \sigma^2)$, disperzija σ^2 je poznata. Osnovna hipoteza je

$$H_0 \quad \ldots \quad a = a_0.$$

Statistika testa:

$$U = \frac{\overline{X} - a_0}{\sigma/\sqrt{n}}.$$

Razdioba statistike je $\mathscr{N}(0,1)$, ukoliko je hipoteza H_0 istinita. Za alternativu

$$H_1 \quad \ldots \quad a > a_0$$

područje odbacivanja hipoteze H_0 glasi: $\widehat{u} > u_{1-\alpha}$, gdje je α zadana pogreška prve vrste.

Odredimo snagu ovog jednostranog testa. Ona ovisi o nepoznatom parametru a. Prema definiciji snage testa,

$$\begin{aligned} S(a) &= \boldsymbol{P}(\{\text{prihvaća se } H_1, \text{ ako je očekivanje jednako } a\}) \\ &= \boldsymbol{P}(U > u_{1-\alpha} \mid a) \end{aligned}$$

S je očito rastuća funkcija. Za $a < a_0$ vrijedi stoga $S(a) < S(a_0)$ pa za pogrešku prve vrste α vrijedi

$$\alpha = \sup_{a \in H_0} S(a) = S(a_0).$$

Pretpostavimo sad da je prava vrijednost očekivanja jednaka a. Očekivanje statistike U u ovom slučaju je

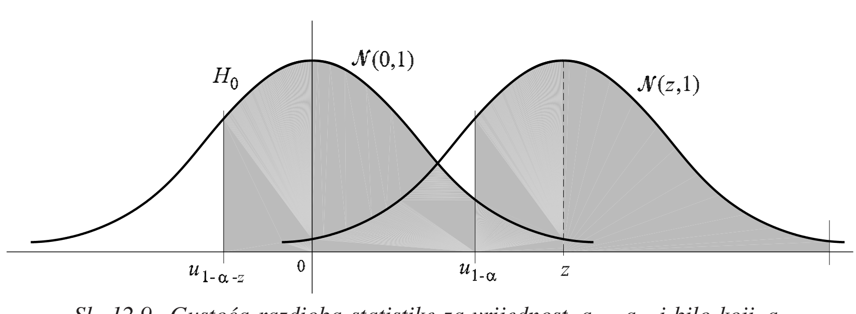
$$\boldsymbol{E}(U) = \boldsymbol{E}\Big(\frac{\overline{X} - a_0}{\sigma/\sqrt{n}}\Big) = \frac{a - a_0}{\sigma/\sqrt{n}}.$$

Označimo broj zdesna sa z, $\boldsymbol{E}(U) = z$. razdioba statistike je sada

$$U = \frac{\overline{X} - a_0}{\sigma/\sqrt{n}} \sim \mathscr{N}(z, 1).$$

Alternativa H_1 bit će prihvaćena ako statistika poprimi vrijednost veću od kritične $u_{1-\alpha}$.

Na sljedećoj slici nacrtane su gustoće razdioba $\mathscr{N}(0,1)$, koja odgovara hipotezi H_0, i razdiobe $\mathscr{N}(z,1)$, koja odgovara alterantivi $\boldsymbol{E}(X) = a$.



Sl. 12.9. Gustoća razdioba statistike za vrijednost $a = a_0$ i bilo koji a.

Vjerojatnost događaja $\{U > u_{1-\alpha}\}$, uz uvjet $\boldsymbol{E}(X) = a$ vidi se iz grafa gustoće razdiobe $\mathscr{N}(z,1)$. To je naznačeno područje ispod grafa te funkcije, desno od $u_{1-\alpha}$. Tom području odgovara istovjetno isod grafa gustoće razdiobe $\mathscr{N}(0,1)$, desno od $u_{1-\alpha} - z$. To ćemo koristiti u sljedećem računu:

$$\begin{aligned} S(a) &= \boldsymbol{P}(U > u_{1-\alpha}) = \boldsymbol{P}(U - z > u_{1-\alpha} - z) = \boldsymbol{P}(\mathscr{N}(0,1) > u_{1-\alpha} - z) \\ &= 1 - \Phi(u_{1-\alpha} - z) = 1 - \Phi\left(u_{1-\alpha} - \frac{a - a_0}{\sigma/\sqrt{n}}\right). \end{aligned}$$

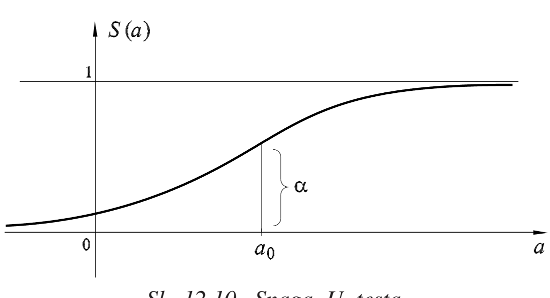
Ovdje je Φ funkcija razdiobe jedinične normalne slučajne varijable.

Na primjer, za $a = a_0$ dobivamo:

$$S(a_0) = 1 - \Phi(u_{1-\alpha}) = 1 - (1 - \alpha) = \alpha,$$

što je u skladu s prije izvedenim.

Graf ove snage dan je na sljedećoj slici:



Sl. 12.10. Snaga U-testa.

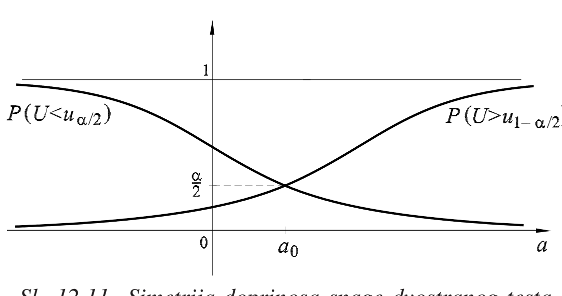
Snaga dvostranog U-testa može se izračunati ovako:

$$S(a) = \boldsymbol{P}(|U| > u_{1-\alpha/2} \mid a) = \boldsymbol{P}(U < u_{\alpha/2} \mid a) + \boldsymbol{P}(U > u_{1-\alpha/2} \mid a).$$

Posljednju vjerojatnost znamo:

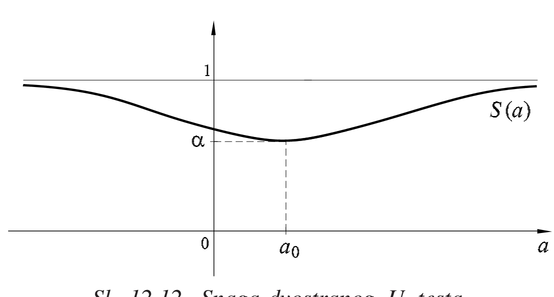
$$\boldsymbol{P}(U > u_{1-\alpha/2} \mid a) = 1 - \Phi\left(u_{1-\alpha/2} - \frac{a - a_0}{\sigma/\sqrt{n}}\right)$$

Zbog svojstva simetrije normalne razdiobe, prvu pribrojnik dobivamo zrcaljenjem ove funkcije oko pravca $a = a_0$:



Sl. 12.11. Simetrija doprinosa snage dvostranog testa.

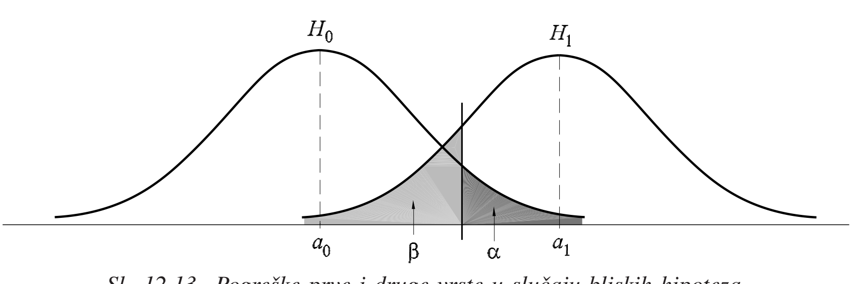
Zbrajanjem ove dvije funkcije dobiva se snaga za dvostrani U-test.



Sl. 12.12. Snaga dvostranog U-testa.

12.2.6. * Pogreške prve i druge vrste i veličina uzorka

U statističkim testovima obično se zadaje maksimalna veličina pogreške prve vrste α. To može imati za posljedicu veliku pogrešku druge vrste β. Ova situacija vidi se na sljedećoj slici:

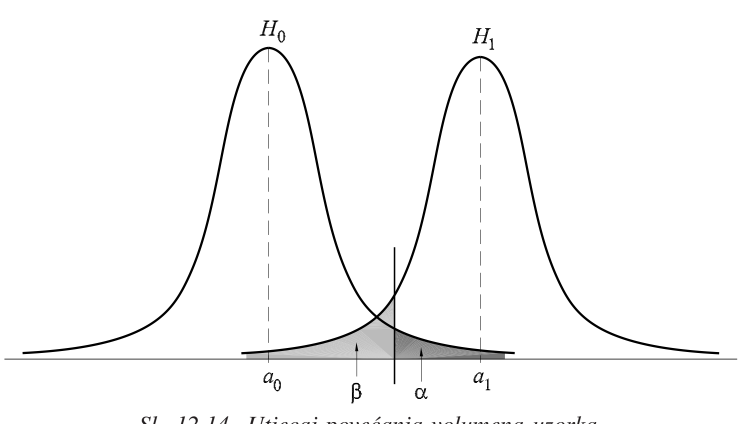


Sl. 12.13. Pogreške prve i druge vrste u slučaju bliskih hipoteza.

Pretpostavimo da se radi o uzorku iz normalne populacije.

Povećanje volumena uzorka ima za posljedicu smanjenje odgovarajuće disperzije. Standardna devijacija smanjuje se proporcionalno korijenu veličine uzorka.

Prema tome, povećanje volumena ima za posljedicu jače razgraničenje odgovarajućih gustoća pojedinih hipoteza.



Sl. 12.14. Utjecaj povećanja volumena uzorka.

Promotrimo dvije situacije u kojima se pojavljuje ovo pitanje.

Primjer 12.9.

U uzorku iz normalne populacije pretpostavljena vrijednost sredine je $H_0 \ldots a = a_0$. Alternativa ovome je $H_1 \ldots a = a_1 > a_0$.

Vjerojatnosti pogrešaka prve i druge vrste su zadane i iznose α i β. Koliko velik mora biti uzorak da bi se te pogreške mogle poštovati?

Kako će glasiti test, ukoliko je $a_0 = 50$, $a_1 = 52$, $\sigma = 5$ te $\alpha = 0.01$ i $\beta = 0.05$?

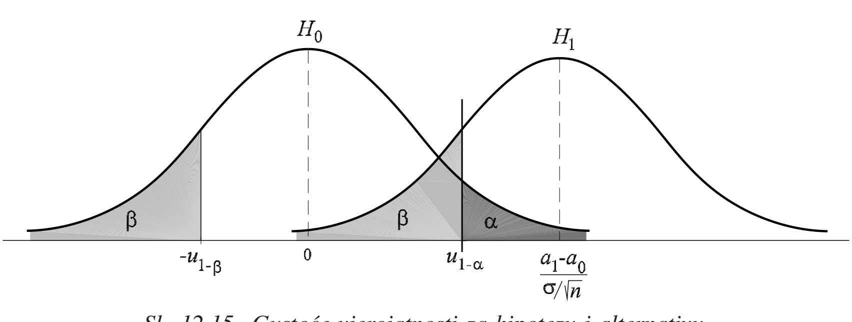
► Ukoliko vrijedi hipoteza $a = a_0$, sredina uzorka $\overline{X}$ ima normalnu razdiobu $\mathscr{N}(a_0, \sigma^2/n)$. To znači da statistika

$$\Theta = \frac{\overline{X} - a_0}{\sigma/\sqrt{n}}$$

ima jediničnu normalnu razdiobu. Međutim, ako je istinita alternativa $a = a_1$, onda je razdioba te statistike također normalna s istom disperzijom, ali joj je očekivanje

$$\boldsymbol{E}(\Theta \mid a = a_1) = \frac{\boldsymbol{E}(\overline{X} \mid a = a_1) - a_0}{\sigma/\sqrt{n}} = \frac{a_1 - a_0}{\sigma/\sqrt{n}}.$$

Grafovi gustoća ovih dviju razdoba dani su na slici. Graf desno, koji odgovara alternativi H_1, dobiven je iz grafa koji odgovara hipotezi H_0 translacijom za iznos $\dfrac{a_1 - a_0}{\sigma/\sqrt{n}}$.



Sl. 12.15. Gustoće vjerojatnosti za hipotezu i alternativu.

Sa slike vidimo da je veza među kvantilima sljedeća:

$$u_{1-\alpha} - \frac{a_1 - a_0}{\sigma/\sqrt{n}} = -u_{1-\beta}.$$

Odavde dobivamo minimalnu vrijednost za volumen n:

$$n = \left(\frac{u_{1-\alpha} + u_{1-\beta}}{a_1 - a_0} \cdot \sigma\right)^2, \tag{1}$$

koju treba zaokružiti na veći prirodni broj.

U zadanom primjeru, vrijednosti kvantila su $u_{1-\alpha} = u_{0.99} = 2.33$ i $u_{1-\beta} = u_{0.95} = 1.64$ pa vrijedi

$$n \geqslant \left(\frac{2.33 + 1.64}{52 - 20} \cdot 5\right)^2 = 98.5\,.$$

Minimalan broj pokusa mora biti 99. Hipoteza H_0 će se prihvatiti ukoliko bude

$$\frac{\overline{x} - a_0}{\sigma/\sqrt{n}} < u_{1-\alpha}$$

t.j.

$$\overline{x} < a_0 + \frac{\sigma}{\sqrt{n}} \cdot u_{1-\alpha} = 51.17\,.$$ ◄

Primjer 12.10.

Želimo testirati ispravnost bacanja novčića. Temeljna hipoteza je da je način bacanja ispravan: $p = 0.5$.

1. Vjerojatnost odbacivanja te hipoteza, kad je ona ispravna, mora biti najviše 0.05.

2. Formiramo alternativu: $|p - 0.5| > 0.1$. Vjerojatnost prihvaćanja temeljne hipoteze, kad je alternativa istinita, mora biti najviše 0.05.

Koliki je minimalni broj bacanja potreban da bi se poštivali ovi uvjeti? Kako će glasiti test?

► Ovdje se radi o *dvostranom testu*. Neka je X broj pisama koji će se pojaviti u n bacanja. Onda je $X \sim \mathscr{B}(np, npq)$. Vjerojatnost ćemo računati pomoću relativne frekvencije

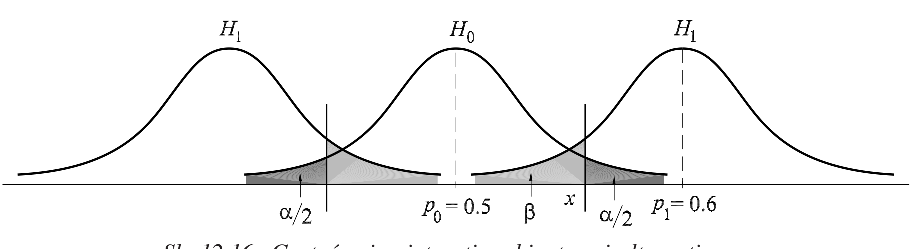
$$\widehat{p} = \frac{X}{n}.$$

Ova varijabla ima približno normalnu razdiobu $\widehat{p} \sim \mathscr{N}\Big(p, \dfrac{pq}{n}\Big)$.

Neka je p_0 vjerojatnost koja odgovara hipotezi H_0, a p_1 ona koja odgovara alternativi.

Ako je istinita H_0, onda vrijedi $\widehat{p} \sim \mathscr{N}\Big(p_0, \dfrac{p_0 q_0}{n}\Big)$. Ako je istinita H_1, onda je $\widehat{p} \sim \mathscr{N}\Big(p_1, \dfrac{p_1 q_1}{n}\Big)$.

Nacrtajmo grafove gustoća ovih razdioba:



Sl. 12.16. Gustoće vjerojatnosti za hipotezu i alternativu.

Sa slike vidimo:

$$\frac{x - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}} = u_{1-\alpha/2}, \qquad \frac{x - p_1}{\sqrt{\dfrac{p_1 q_1}{n}}} = -u_{1-\beta}.$$

Eliminacijom varijable x dobivamo vezu:

$$x = p_0 + u_{1-\alpha/2}\sqrt{\frac{p_0 q_0}{n}} = p_1 - u_{1/\beta}\sqrt{\frac{p_1 q_1}{n}}.$$

Odavde je

$$n = \left(\frac{u_{1-\alpha/2}\sqrt{p_0 q_0} + u_{1-\beta}\sqrt{p_1 q_1}}{p_1 - p_0}\right)^2$$

U konkretnom je primjeru $p_0 = 0.5$, $p_1 = 0.6$, $u_{1-\alpha/2} = u_{0.975} = 1.96$, $u_{1-\beta} = u_{0.95} = 1.64$. Uvrštavanjem dobivamo $n = 318.06$. Minimalni broj bacanja je $n = 319$. Za taj broj dobivamo

$$\widehat{p} = p_0 \pm u_{1-\alpha/2}\sqrt{\frac{p_0 q_0}{n}} = 0.5 \pm 0.0549\,.$$

Sada je $n\widehat{p} = 159.5 \pm 17.5$. Prema tome, prihvatit ćemo hipotezu ako broj pisama bude između 142 i 177. ◄

12.3. Usporedbe dviju populacija

Preciznim uređajem mjerena su svojstva nekog elementa. Radi važnosti ispitivanja, ta su ista svojstva mjerena drugim uređajem. Tako su dobivena dva niza podataka.

Kako ćemo analizirati te podatke? Na primjer, kako ćemo odrediti sredinu i disperziju populacije? Ako se pojavila razlika u sredinama pri ta dva mjerenja, je li ona ukazuje na različita svojstva mjernih uređaja, ili je posljedica slučajnosti koja se treba tolerirati?

Na ta i na još neka pitanja odgovorit ćemo u ovom poglavlju.

12.3.1. Hipoteza o sredinama, uz poznatu disperziju

Teorijski ćemo model postaviti ovako. Zadane su dvije slučajne varijable, X s razdiobom F_1 i Y s razdiobom F_2. Pretpostavljat ćemo da su te varijable *nezavisne*.

Te varijable generiraju uzorke:

$$\begin{aligned} &X_1, X_2, \ldots, X_n, && \text{s razdiobom } F_1,\\ &Y_1, Y_2, \ldots, Y_m, && \text{s razdiobom } F_2. \end{aligned}$$

Prvo pitanje na koje ćemo djelomično odgovoriti jest: kako se može provjeriti jesu li neki parametri razdioba F_1 i F_2 identični?

Promotrit ćemo sljedeću situaciju: $X \sim \mathcal{N}(\mu, \sigma_X^2)$, $Y \sim \mathcal{N}(\nu, \sigma_Y^2)$. Osnovna hipoteza je:

$$H_0 \qquad \ldots \qquad \mu = \nu.$$

U postupku analize koristit ćemo statistiku $\overline{X} - \overline{Y}$.

Znamo da $\overline{X}$ ima normalnu razdiobu $\sim \mathcal{N}\Big(\mu, \dfrac{\sigma_X^2}{n}\Big)$. Također, $\overline{Y}$ ima normalnu razdiobu $\sim \mathcal{N}\Big(\nu, \dfrac{\sigma_Y^2}{m}\Big)$. Slučajne varijable $X_1, \ldots, X_n$ i $Y_1, \ldots, Y_m$ su nezavisne, pa vrijedi

$$\boldsymbol{E}(\overline{X} - \overline{Y}) = \boldsymbol{E}(\overline{X}) - \boldsymbol{E}(\overline{Y}) = \mu - \nu,$$

$$\boldsymbol{D}(\overline{X} - \overline{Y}) = \boldsymbol{D}(\overline{X}) + \boldsymbol{D}(\overline{Y}) = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}.$$

Označimo zbog kratkoće ovu zajedničku disperziju s

$$\sigma_z^2 = \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}.$$

Zbroj nezavisnih normalnih varijabla ima normalnu razdiobu. Zato vrijedi:

$$\overline{X} - \overline{Y} \sim \mathcal{N}(\mu - \nu, \sigma_z^2).$$

U testu ćemo koristiti statistiku

$$U = \frac{\overline{X} - \overline{Y}}{\sigma_z}.$$

Njezina je teorijska razdioba

$$U \sim \mathcal{N}\Big(\frac{\mu - \nu}{\sigma_z}, 1\Big)$$

Ako je hipoteza H_0 istinita, tada je $U \sim \mathcal{N}(0, 1)$. Zato se u testu može primjenjivati uobičajeni U-test. Za moguće alternative, kriterij za odbacivanje hipoteze glasi:

$$\begin{aligned} &H_1 && \ldots && \nu < \mu && \text{(desna)} && \widehat{u} > u_{1-\alpha},\\ &H_1 && \ldots && \nu > \mu && \text{(lijeva)} && \widehat{u} < -u_{1-\alpha},\\ &H_1 && \ldots && \nu \neq \mu && \text{(obostrana)} && |\widehat{u}| > u_{1-\alpha/2}, \end{aligned}$$

Da bi rezultati bili smisleni, oba uzorka moraju biti istog reda veličine.

Primjer 12.11.

Uzorak od $n = 40$ čeličnih žica proizvođača A imao je srednju čvrstoću na kidanje $\overline{x} = 1190$ kg. Uzorak od $m = 50$ čeličnih žica proizvođača B dao je srednju čvrstoću $\overline{y} = 1220$ kg. Standardna devijacija za oba uzorka je poznata i iznosi 90 kg. Postoji li razlika u čvrstoći ovih žica? Ocjenu treba dati uz nivo značajnosti $\alpha = 0.05$.

► Postupimo prema uputama. Primjenit ćemo test o dvostranoj alternativi, zanemarujući rezultate dobivene uzorkom.[1]

Standardno odstupanje zajedničkog uzorka je

$$\sigma_z = \sigma\sqrt{\frac{1}{n} + \frac{1}{m}} = 19.09$$

Vrijednost statistike, uz pretpostavljenu uspravnost hipoteze o jednakim očekivanjem, je

$$\widehat{u} = \frac{\overline{x} - \overline{y}}{\sigma_z} = -1.57\,.$$

Kritična vrijednost kvantila je $u_{1-\alpha/2} = u_{0.975} = 1.96$ Budući da je $|\widehat{u}| < 1.96$, hipoteza o jednakosti ne može se odbaciti. ◄

12.3.2. Hipoteza o sredinama, uz nepoznatu disperziju

Pretpostavimo sad da su uzorci $X_1, \ldots, X_n$ i $Y_1, \ldots, Y_m$ nezavisni s normalnom razdiobom uz jednaku disperziju čiji iznos nije poznat.

U tom slučaju računamo procjenu disperziju iz uzorka:

$$S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2, \qquad S_Y^2 = \frac{1}{m-1}\sum_{j=1}^{m}(Y_j - \overline{Y})^2.$$

Zajednička disperzija uzorka računa se kao težinska sredina ovih disperzija:

$$S_Z^2 = \frac{1}{n+m-2}\Big[(n-1)S_X^2 + (m-1)S_Y^2\Big].$$

Slučajna varijabla

$$U = \frac{(\overline{X} - \mu) - (\overline{Y} - \nu)}{\sigma\sqrt{\dfrac{1}{n} + \dfrac{1}{m}}}$$

ima jediničnu normalnu razdiobu.

Slučajna varijabla

$$W^2 = \frac{(n-1)S_X^2 + (m-1)S_Y^2}{\sigma^2}$$

ima hi kvadrat razdiobu s $m + n - 2$ stupnjeva slobode. Zato slučajna varijabla

$$\frac{U\sqrt{n+m-2}}{W}$$

ima Studentovu razdiobu s $m + n - 2$ stupnjeva slobode.

Time smo pokazali da je distribucija slučajne varijable

$$\frac{(\overline{X} - \overline{Y}) - (\mu - \nu)}{\sqrt{(n-1)S_X^2 + (m-1)S_Y^2}}\sqrt{\frac{nm(n+m-2)}{n+m}}$$

Studentova razdioba s $m + n - 2$ stupnjeva slobode.

Pretpostavimo da je hipoteza H_0 o jednakosti očekivanja istinita. U tom slučaju možemo koristiti statistiku

$$T = \frac{\overline{X} - \overline{Y}}{S_Z}\sqrt{\frac{nm}{n+m}}.$$

Njezina je razdioba Studentova, s $n + m - 2$ stupnjeva slobode.

12.3.3. Hipoteza o jednakosti proporcija

U uzorcima veličina n_1 i n_2 broj objekata s danim svojstvima iznosi m_1 odnosno m_2. Time dobivamo relativne frekvencije

$$\widehat{p}_1 = \frac{m_1}{n_1}, \qquad \widehat{p}_2 = \frac{m_2}{n_2}.$$

Želimo odrediti test za hipotezu o jednakosti ovih proporcija.

Varijabla m_1 ima binomnu razdiobu $\mathcal{B}(n_1, p')$, varijabla m_2 ima binomnu razdiobu $\mathcal{B}(n_2, p'')$. Temeljna hipoteza jest da je vjerojatnost realizacije u oba uzorka jednaka:

$$H_0 \qquad \ldots \qquad p' = p''$$

Pretpostavimo da je ta hipoteza istinita i označimo s p zajedničku vjerojatnost. Neka je $\sigma^2 = p(1-p)$. Sada imamo

$$\frac{m_1}{n_1} \sim \frac{1}{n_1}\mathcal{B}(n_1, p) \approx \mathcal{N}\Big(p, \frac{\sigma^2}{n_1}\Big),$$

$$\frac{m_2}{n_2} \sim \frac{1}{n_2}\mathcal{B}(n_2, p) \approx \mathcal{N}\Big(p, \frac{\sigma^2}{n_2}\Big).$$

Zato za razliku proporcija vrijedi

$$\widehat{p}_1 - \widehat{p}_2 \sim \mathcal{N}\Big(0, \frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}\Big).$$

Time možemo definirati statistiku

$$U = \frac{\widehat{p}_1 - \widehat{p}_2}{\sqrt{\dfrac{\sigma^2}{n_1} + \dfrac{\sigma^2}{n_2}}}.$$

Njezina je distribucija $\mathcal{N}(0, 1)$.

Vrijednost disperzije σ^2 nije poznata, pa je aproksimiramo pomoću cjelokupnog uzorka:

$$p = \frac{m_1 + m_2}{n_1 + n_2}, \qquad \sigma^2 = p(1-p).$$

Primjer 12.12.

Od 100 anketiranih muškaraca 30 je gledalo televizijsku emisiju. Tu istu emisiju gledalo je 45 od 120 anketiranih žena. Uz nivo značajnosti $\alpha = 0.05$ provjerite hipotezu da je jednak postotak muškaraca i žena koji su gledali tu emisiju.

► Izračunajmo potrebne veličine. Zadano je $m_1 = 30$, $n_1 = 100$, $m_2 = 45$, $n_2 = 120$ pa je $\widehat{p}_1 = 0.3$ i $\widehat{p}_2 = 0.375$. Nadalje

$$p = \frac{m_1 + m_2}{n_1 + n_2} = \frac{75}{220} = 0.34, \qquad \sigma = \sqrt{p(1-p)} = 0.474\,.$$

Vrijednost statistike je

$$U = \frac{\widehat{p}_1 - \widehat{p}_2}{\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}} = \frac{0.3 - 0.375}{0.474\sqrt{\dfrac{1}{100} + \dfrac{1}{120}}} = -1.16\,.$$

Kritična vrijednost kvantila za dvostrani test je $u_{1-\alpha/2} = u_{0.975} = 1.96$.

Hipoteza o jednakosti proporcija ne može se odbaciti. ◄

Primjer 12.13.

Poljoprivredni institut ispituje utjecaj gnojiva na novi nasad. U tu svrhu, 150 parcela od 400 ostalo je nezagnojeno. Na 77 među njima prinos je bio zadovoljavajući. Preostalih 250 je gnojeno, a zadovoljavajući prinos se dobio na 158 parcela.

Uz nivo značajnosti $\alpha = 0.05$ testirajmo hipotezu da gnojenje ne utječe na prinos te kulture.

► Sada je $m_1 = 77$, $n_1 = 150$, $m_2 = 158$, $n_2 = 250$. Računamo:

$$\widehat{p}_1 = 0.513, \qquad \widehat{p}_2 = 0.632, \qquad p = 0.588, \qquad \sigma = 0.492.$$

Vrijednost statistike je

$$\widehat{u} = \frac{\widehat{p}_1 - \widehat{p}_2}{\sigma}\sqrt{\frac{n_1 n_2}{n_1 + n_2}} = -2.34$$

Riječ je o jednostranoj (lijevoj) hipotezi, jer se pretpostavlja da gnojenje povećava prinos. Zato je kvantil $u_\alpha = -u_{1-\alpha} = -u_{0.95} = -1.645$.

Rezultat pokazuje da se hipoteza mora odbaciti. ◄

12.4. Testovi prilagodbe razdiobama

U svim dosadašnjim razmatranjima određivali smo procjene ili intervale povjerenja za nepoznate parametre *poznate* razdiobe.

Ako razdioba slučajne varijable nije poznata, možemo li na temelju vrijednosti uzorka otkriti o kojoj se razdiobi radi? Kolika je pouzdanost takvog zaključka?

Na ta pitanja najbolji odgovor daje **χ^2-test** (**Pearsonov test**).

12.4.1. χ^2-test

Pretpostavimo da slučajna varijabla X ima nepoznatu razdiobu F. Cilj hi kvadrat testa jest provjeriti hipotezu o vrsti te razdiobe. Ideja testa temelji se na sljedećem.

Područje vrijednosti slučajne varijable X razbije se na disjunktne intervale A_k, $k = 1, \dots, m$. Uz pretpostavku da je hipoteza o razdiobi istinita, odrede se **teorijske vjerojatnosti**

$$p_k = \boldsymbol{P}(X \in A_k).$$

Zbroj ovih vjerojatnosti iznosi

$$p_1 + \dots + p_m = 1.$$

Na temelju realizacija slučajne varijable X, odredi se broj realizacija n_k koji pripada pojedinom razredu. Zbroj svih tih realizacija jednak je volumenu uzorka:

$$n_1 + \dots + n_m = n.$$

Označimo s Y_k slučajnu varijablu koja mjeri broj realizacija unutar razreda A_k. To je binomna slučajna varijabla $Y_k \sim \mathscr{B}(n, p_k)$. Njezina očekivana vrijednost je np_k. na taj način možemo uspoređivati dvije vjerojatnosti:

p_k teorijska vjerojatnost k-tog razreda,

$\dfrac{n_k}{n}$ vjerojatnost k-tog razreda dobivena iz uzorka.

Ukoliko je hipoteza o razdiobi točna, tada razlike $\dfrac{n_k}{n} - p_k$ moraju biti malene. Zato je prirodno uzeti kao *mjeru odstupanja od teorijske razdiobe* sljedeću težinsku varijantu zbroja najmanjih kvadrata:

$$\sum_{k=1}^{m} t_k \left(\frac{n_k}{n} - p_k\right)^2.$$

Težinske faktore ćemo odabrati tako da bude $t_k = n/p_k$. Razlog za ovaj odabir jest što će ova suma dobiti oblik.

$$\sum_{k=1}^{m} \frac{(n_k - np_k)^2}{np_k}.$$

Pearson[1] je pokazao da ova slučajna varijabla (vrijednosti brojeva n_k ovise o realizaciji uzorka) **ima približno χ^2-razdiobu**, s $n-1$ stupnjeva slobode.

Dokaz te tvrdnje može se naslutiti iz oblika pribrojnika. Slučajna varijabla

$$\frac{Y_k - np_k}{\sqrt{np_k(1-p_k)}}$$

ima očekivanje 0 i disperziju 1, pa se može aproksimirati razdiobom $U \sim \mathscr{N}(0,1)$. To znači da kvadrat te varijable ima približno U^2 razdiobu, pa je onda i

$$\frac{(n_k - np_k)^2}{np_k} \approx U^2.$$

Zbroj ovakvih slučajnih varijabli ima χ^2 razdiobu. Broj stupnjeva slobode smanjen je za 1, jer varijable $Y_1, \dots, Y_m$ nisu nezavisne, njihov zbroj iznosi n. Nazivnik $np_k(1-p_k)$ zamijenjen je s np_k da bi se dobila *bolja aproksimacija*.

Dokaz ovog teorema o aproksimaciji ne možemo navesti na ovom mjestu. Umjesto toga, izvest ćemo transformaciju u najjednostavnijem slučaju $m = 2$. Tu je $n = n_1 + n_2$ i $1 = p_1 + p_2$ pa imamo

$$\begin{aligned}\frac{(n_1-np_1)^2}{np_1} + \frac{(n_2-np_2)^2}{np_2} &= \frac{(n_1-np_1)^2}{np_1} + \frac{[n - n_1 - n(1-p_1)]^2}{np_2} \\ &= \frac{(n_1-np_1)^2}{n}\left(\frac{1}{p_1} + \frac{1}{p_2}\right) = \frac{(n_1-np_1)^2}{np_1(1-p_1)}.\end{aligned}$$

Dobivena slučajna varijabla ima približno razdiobu U^2, što je i trebalo pokazati u ovom slučaju.

U primjenama vrlo često poznajemo samo *tip* razdiobe, ali ne i sve njezine parametre. Tada nepoznate parametre računamo iz uzorka. To ima za posljedicu *smanjenje stupnjeva slobode* odgovarajuće χ^2 razdiobe.

Opišimo sad algoritam.

Hi kvadrat test

1. Uzorak $\{x_1, x_2, \dots, x_n\}$ podijelimo u m razreda. Neka je n_j broj realizacija u pojedinom razredu, te p_j teorijska vjerojatnost pojedinog razreda. Minimalna teorijska frekvencija np_j pojedinog razreda treba biti barem 5, razrede za koje je $np_j < 5$ spajamo s njima susjednim.

2. Statistika χ^2-testa dana je s

$$\chi_q^2 = \sum_{j=1}^{m} \frac{(n_j - np_j)^2}{np_j}$$

Slučajna varijabla χ_q^2 ima približno χ^2 razdiobu s $f = m - r - 1$ stupnjeva slobode. Tu je r broj parametara razdiobe izračunatih iz uzorka.

3. U tablicama pročitamo kvantil $\chi^2_{\text{krit}} = \chi^2_{f,1-\alpha}$, za zadani nivo značajnosti α i broj stupnjeva slobode f.

4. Ako je $\chi_q^2 < \chi^2_{\text{krit}}$, prihvaća se hipoteza da se razdioba varijable X podvrgava dotičnom zakonu. U protivnom, ta se hipoteza odbacuje.

Primjer 12.14.

Na jednom strelištu gađano je u 100 meta, u svaku s po 10 metaka. Bilježen je broj pogodaka u svaku od meta:

j	0	1	2	3	4	5	6	7	8	9	10
n_j	0	2	4	10	22	26	18	12	4	2	0

S kojom vjerojatnošću se može tvrditi da se on ravna po binomnom zakonu?

► Prvi parametar binomne razdiobe je 10. Nepoznati parametar p binomne razdiobe odredit ćemo iz uzorka. U tu svrhu, iskoristit ćemo vezu parametra s očekivanjem: $\boldsymbol{E}(X) = 10p$. Vrijedi

$$\bar{x} = \frac{1}{n}\sum_{j=0}^{10} n_j \cdot j = 5.0\,.$$

Zato je $p = \frac{\bar{x}}{10} = 0.5$.

Prema tome, hipoteza koju testiramo glasi $X \sim \mathscr{B}(10, \frac{1}{2})$. Teorijske frekvencije pojedinog razreda su

$$p_j = \binom{10}{j}\left(\frac{1}{2}\right)^j\left(\frac{1}{2}\right)^{10-j} = \binom{10}{j}\frac{1}{2^{10}}.$$

Sad možemo popuniti tablicu:

j	n_j	p_j	$n_j - np_j$	$\frac{(n_j-np_j)^2}{np_j}$
0	0			
1	2	0.0546	0.540	0.053
2	4			
3	10	0.1172	$-1,720$	0.252
4	22	0.2051	1,490	0.108
5	26	0.2466	1.340	0.073
6	18	0.2051	-2.510	0.307
7	12	0.1172	0.280	0.007
8	4			
9	2	0.0546	0.540	0.053
10	0			
	100			$\chi_q^2 = 0.854$

Spojili smo prva tri i posljednja tri razreda. S obzirom da je u ovoj tablici preostalo 7 razreda, a jedan parametar je dobiven iz uzorka, broj stupnjeva slobode je $f = 7 - 1 - 1 = 5$.

Uvidom u tablicu kvantila hi kvadrat razdiobe s 5 stupnjeva slobode čitamo: $\chi^2_{0.025} = 0.75$, $\chi^2_{0.05} = 1.14$. Prema tome X se ravna po binomnoj razdiobi uz nivo značajnosti $\alpha = 0.95$. ◄

Primjer 12.15.

Provjeravaju se dimenzije iz uzorka načinjenog na preciznom tokarskom automatu, s točnošću od $1\ \mu$m. U tablici su dana odstupanja od nominalne vrijednosti, svrstani u intervale $[a_j, a_{j+1}]$ zadanih duljina. Provjeri hipotezu da se odstupanja ravnaju po normalnoj razdiobi, uz nivo značajnosti 5%.

a_j	a_{j+1}	n_j
$-\infty$	-15	9
-15	-10	12
-10	-5	14
-5	0	17
0	5	43
5	10	45
10	15	23
15	20	15
20	25	14
25	$+\infty$	8

► Volumen uzorka iznosi $n = 200$. Svi razredi su dovoljne veličine, pa je $m = 10$.

Parametre razdiobe izračunat ćemo iz uzorka. S obzirom da je varijabla neprekinutog tipa (razbijena u razrede), svaki ćemo razred predstaviti reprezentantom, brojem x_j koji leži u sredini intervala. Za prvi i posljednji interval uzet ćemo ekvidistantne brojeve $x_1 = -17.5$ i $x_{10} = 27.5$.

$$\bar{x} = \frac{1}{n}\sum n_j x_j = 5.375$$

Na isti način računamo i disperziju uzorka:

$$\overline{m}_2 = \frac{1}{n}\sum n_j x_j^2 = 148.25$$

Odavde je

$$\widehat{\sigma}^2 = \overline{m}_2 - \bar{x}^2 = 119.36,$$

pa je nepristrana procjena za disperziju

$$\widehat{s}^2 = \frac{n}{n-1}\widehat{\sigma}^2 = 119.96\,.$$

Odavde je $\widehat{s} = 10.95$.

Prema tome, testiramo hipotezu $X \sim \mathscr{N}(5.375, 10.95^2)$.

Da bismo odredili teorijske frekvencije razreda, izračunat ćemo i upisati u tablicu iznos brojeva z_j, koji odgovaraju granicama intervala a_j, ali za odgovarajuću jediničnu normalnu razdiobu Z:

$$z_j = \frac{a_j - \bar{x}}{\widehat{s}}.$$

Zatim se u tablicama potraže vrijednosti normalne razdiobe u tim točkama. te se vrijednosti upišu u sljedeći stupac tablice. One su potrebne da bi se izračunale teorijske vjerojatnosti:

$$p_{j+1} = \boldsymbol{P}(a_j \leqslant X \leqslant a_{j+1}) = \boldsymbol{P}(z_j \leqslant Z \leqslant z_{j+1}) = \frac{1}{2}[\Phi^*(z_{j+1}) - \Phi^*(z_j)]$$

Ostali stupci tablice popune se odgovarajućim vrijednostima.

a_j	n_j	x_j	z_j	$\Phi^*(z_j)$	p_j	$\frac{(n_j-np_j)^2}{np_j}$
$-\infty$			$-\infty$	-1		
-15	9	-17.5	-1.86	-0.937	0.031	1,17
-10	12	-12.5	-1.40	-0.840	0.049	0.52
-5	14	-7.5	-0.947	-0.657	0.092	1.02
0	17	-2.5	-0.491	-0.376	0.140	4.33
5	43	2.5	-0.034	-0.027	0.175	1.88
10	45	7.5	0.422	0.327	0.177	2.57
15	23	12.5	0.879	0.620	0.147	1.37
20	15	17.5	1.34	0.818	0.099	1.15
25	14	22.5	1.79	0.927	0.054	0.91
$+\infty$	8	27.5	$+\infty$	1	0.037	0.06
$\sum$	200				1	14.97

U ovoj tablici rub prvog intervala $a_0 = -20$ zamijenjen je s $-\infty$, a rub posljednjeg intervala a_{10} zamijenjen je s $+\infty$. Tako je vrijednost funkcije Φ^* u tim točkama jenaka -1 odnosno 1. To je učinjeno da bi zbroj svih teorijskih vjerojatnosti iznosio 1.

Broj stupnjeva slobode je $k = m - r - 1 = 7$. Kritična vrijednost kvantila χ^2 razdiobe je $\chi^2_{0.95} = 14.1$. Dobivena vrijednost je veća od granične, pa se hipoteza o normalnoj razdiobi mora odbaciti.

Račun u ovoj tablici načinjen je programom Excell. Većina statističkih izračuna može se vrlo jednostavno računati uporabom tog programa. To se pogotovo odnosi na programe specijalizirane za primjenu u matematičkoj statistici. ◄

12.4.2. Relativne frekvencije i rekonstrukcija funkcije razdiobe

Pretpostavimo da nam je nepoznata funkcija razdiobe populacije X. Možemo li tu funkciju odrediti na temelju vrijednosti koje slučajna varijabla poprima?

Neka je x bilo koji realni broj. Definirajmo funkciju

$$F_n(x) = \frac{\mu_n(x)}{n}$$

pri čemu je

$$\mu_n(x) = \text{broj elemenata uzorka } \{X_1, \dots, X_n\} \text{ koji su manji od } x.$$

Primjer 12.16.

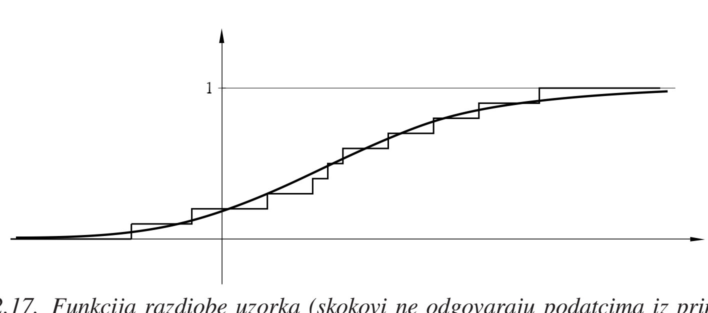
Nacrtajmo graf funkcije F_n, ako je uzorak poprimio vrijednosti

$$-2,\ 3,\ -1.5,\ 0.4,\ 2,\ 5,\ -1,\ 2.2,\ -0.3,\ 1.7\,.$$

► Volumen uzorka n iznosi 10. Vrijednosti u uzorku poredamo u rastućem poretku:

$$-2,\ -1.5,\ -1,\ -0.3,\ 0.4,\ 1.7,\ 2,\ 2.2,\ 3,\ 5\,.$$

Prema definiciji funkcije F_n, ona će biti stepenasta funkcija sa skokovima iznosa $\dfrac{1}{n} = \dfrac{1}{10}$ u ovim točkama. ◄



Sl. 12.17. Funkcija razdiobe uzorka (skokovi ne odgovaraju podatcima iz primjera).

Vrijednosti funkcije F_n ovise o realizaciji uzorka. Ona je slučajna varijabla tog uzorka.

Teorem 12.1.

Statistika $F_n(x)$ je nepristrana i valjana procjena za vrijednost slučajne varijable $F(x)$.

Dokaz. Označimo, za svaki k iz skupa $\{1, \dots, n\}$:

$$p = \boldsymbol{P}(X_k < x) = F(x),$$
$$q = \boldsymbol{P}(X_k \geqslant x) = 1 - F(x).$$

Tada slučajna varijabla $\mu_n(x)$ ima binomnu razdiobu $\mathscr{B}(n,p)$. Stoga vrijedi

$$\boldsymbol{E}(F_n(x)) = \frac{1}{n}\boldsymbol{E}(\mu_n(x)) = \frac{1}{n}\cdot np = p = F(x).$$

Nadalje, prema slabom zakonu velikih brojeva, vrijedi

$$\left|\frac{\mathscr{B}(n,p) - np}{n}\right| \xrightarrow{P} 0$$

pa dobivamo

$$|F_n(x) - F(x)| \xrightarrow{P} 0.$$

Dakle, $F_n(x)$ je valjana procjena za $F(x)$.

Iskoristimo li jaki zakon velikih brojeva, zaključit ćemo da u ovom slučaju vrijedi i mnogo jača tvrdnja:

Teorem 12.2. ■ **Glivenko-Cantelli**

Neka su (X_n) nezavisne kopije slučajne varijable X s razdiobom F. tada vrijedi

$$\lim_{n\to\infty} \boldsymbol{P}\left(\sup_x \left|\frac{\mu_n(x)}{n} - F(x)\right| = 0\right) = 1.$$

Dakle, $\dfrac{\mu_n(x)}{n}$ konvergira ka $F(x)$ skoro sigurno.

12.4.3. Kolmogorovljev kriterij

Ovim testom provjeravamo hipotezu o ravnanju podataka dobivenih uzorkom prema nekoj razdiobi s poznatom funkcijom razdiobe F.

Kriterij se temelji na teoremu Glivenko-Cantellija prema kojem niz $(F_n(x))$ funkcija razdioba dobivenih iz uzorka konvergira prema funkciji razdiobe $F(x)$. To znači da će maksimum razlike tih dviju funkcija težiti u nulu, ukoliko je $F(x)$ zaista funkcija razdiobe koja odgovara uzorku.

Definirajmo slučajnu varijablu

$$\lambda := \sqrt{n}\cdot\sup_x |F(x) - F_n(x)|.$$

Kolmogorov je odredio zakon razdiobe ove slučajne varijable. Njezini kvantili su izračunati i zapisani u tablicama.

Za zadani nivo pouzdanosti p, kvantil ove razdiobe λ_p pročita se u tablicama. Hipotezu o ravnanju uzorka prema teorijskoj razdiobi ćemo odbaciti, ukoliko je dobivena vrijednost veća od kritične:

$$\lambda > \lambda_p.$$

U protivnom, hipotezu prihvaćamo (ili je ne možemo odbaciti).

Primjer 12.17.

Znamenke 0,1,2,...,9 među prvih 800 decimala broja π pojavljuju se 74-92-83-79-80-73-77-75-76-91 puta. Kad bi te znamenke bile slučajne, s kojom vjerojatnošću bi mogli prihvatiti tvrdnju da se svaka među njima može pojaviti na nekom mjestu decimalnog zapisa s jednakom vjerojatnošću?

► Funkcije $F_n(x)$ i $F(x)$ su stepenaste, sa skokovima u točkama $0, 1, 2, \dots, 9$, Iznos skoka funkcije F iznosi 0.1, a iznos skoka funkcije F_n omjer je broja realizacija pojedine znamenke i ukupnog broja znamenaka. Eksplicitne formule su:

$$F(x) = \frac{j+1}{10}, \qquad j \leqslant x < j+1,$$
$$F_n(x) = \frac{1}{n}\sum_{i \leqslant j+1} n_i, \qquad j \leqslant x < j+1.$$

Vrijednosti tih funkcija iščitavaju se iz ove tablice:

x_j	n_j	$F_n(x)$	$F(x)$
0	74	0.093	0.1
1	92	0.208	0.2
2	83	0.311	0.3
3	79	0.410	0.4
4	80	0.510	0.5
5	73	0.601	0.6
6	77	0.698	0.7
7	75	0.791	0.8
8	76	0.886	0.9
9	91	1.000	1.0

Za ovu realizaciju, slučajna varijabla λ poprima vrijednost

$$\lambda = \sqrt{n}\cdot\max_x |F(x) - F_n(x)| = \sqrt{800}\cdot 0.014 = 0.396\,.$$

Iz tablica se očitavaju kvantili:

$$\lambda_{0.001} = 0.38, \qquad \lambda_{0.005} = 0.42.$$

Dakle, vjerojatnost prihvaćanja hipoteze po ovom kriteriju je veća od 0.995. ◄

12.4.4. Ispitivanje nezavisnosti slučajnih varijabla

Pretpostavit ćemo da su slučajne varijable X i Y diskretnog tipa, ili je područje njihovih vrijednosti razbijeno u disjunktne razrede, i to:

- vrijednosti varijable X u r razreda,
- vrijednosti varijable Y u s razreda.

Slučajne varijable definirane su na istom vjerojatnosnom prostoru i poprimaju istovremeno vrijednosti na svakom elementarnom događaju. Označimo:

$$k_{ij} = \text{ broj realizacija za koje je } X = x_i,\ Y = y_j$$

Marginalne frekvencije dobivamo zbrajanjem. Ukupan broj realizacija za koje je $X = x_i$ označavamo s:

$$k_{i0} = \sum_{j=1}^{s} k_{ij}.$$

Analogno, ukupan broj realizacija za koje je $Y = y_j$ označavamo s:

$$k_{0j} = \sum_{i=1}^{r} k_{ij}.$$

Ako su slučajne varijable nezavisne, tad bi frekvencija k_{ij} trebala biti proporcionalna s $k_{i0}\cdot k_{0j}$. Zato je prirodno kao mjeru za odstupanje od nezavisnosti definirati sumu

$$\chi_q^2 = \sum_{i=1}^{r}\sum_{j=1}^{s}\frac{(k_{ij} - n_{ij})^2}{n_{ij}}.$$

Tu je

$$n_{ij} = \frac{k_{i0}k_{0j}}{N}, \qquad N = \sum_{i,j} k_{ij}.$$

Slučajna varijabla χ_q^2 ima približno χ^2 razdiobu s

$$k = (r-1)(s-1)$$

stupnjeva slobode. Hipoteza o nezavisnosti će se prihvatiti ukoliko je izračunata vrijednost χ_q^2 manja od kritičnog kvantila za tu hi kvadrat razdiobu.

Primjer 12.18.

Izmjereno je 600 detalja, pri čemu su za svaki od njih provjeravane dimenzije, dužina X i širina Y.

	podbačaj	u granicama	prebačaj
podbačaj	6	48	8
u granicama	52	402	36
prebačaj	6	38	4

Pomoću χ^2-testa provjerimo da li su otkloni dimenzija X i Y međusobno nezavisni, uz nivo značajnosti $\alpha = 0.1$.

► Postupak računanja pratimo u sljedećim tablicama:

	podbačaj	u granicama	prebačaj	k_{i0}
podbačaj	6	48	8	62
u granicama	52	402	36	490
prebačaj	6	38	4	48
k_{0j}	64	488	48	600

n_{ij}	$j=1$	$j=2$	$j=3$
$i=1$	6.61	50.43	4.96
$i=2$	52.27	398.53	39.20
$i=3$	5.12	39.04	3.84

$\frac{(k_{ij}-n_{ij})^2}{n_{ij}}$	$j=1$	$j=2$	$j=3$
$i=1$	0.06	0.12	1.86
$i=2$	0.00	0.03	0.26
$i=3$	0.15	0.03	0.01

Zbroj elemenata u ovoj tablici daje $\chi_q^2 = 2.52$. Kritična vrijednost hi kvadrat razdiobe za $k = (r-1)(s-1) = 4$ stupnja slobode je $\chi^2_{0.9} = 7.78$.

Prema tome, hipoteza o nezavisnosti ne može se odbaciti. ◄