Anuška Ferligoj, Katja Lozar Manfreda, Aleš Žiberna: OSNOVE STATISTIKE NA PROSOJNICAH

Študijsko gradivo pri predmetu Statistika. Fakulteta za družbene vede, Univerza v Ljubljani

Ljubljana, 2011

4 INFERENČNA STATISTIKA

4.1 UVOD

Inferenčna (sklepna) statistika = področje statistike, ki se ukvarja z metodami statističnega sklepanja (angl. *to infer* = sklepati).

Statistično sklepanje (indukcija) (angl. *statistical inference*) = sklepanje iz vzorca na populacijo.

V statistiki množične pojave praviloma raziskujemo na podlagi podatkov iz verjetnostnih vzorcev. Pri tem uporabljamo metode statističnega sklepanja.

Metode statističnega sklepanja:

- **statistično ocenjevanje** = postopek, s katerim iz vzorca pridobimo oceno statistične karakteristike celotne populacije,
- **preizkušanje hipotez** (preverjanje domnev) = postopek, v katerem na osnovi podatkov iz vzorca preverjamo veljavnost domnev na populaciji.

Spomnimo se:

- **populacija** – množica vseh proučevanih elementov, vseh enot,
- **vzorec** – podmnožica populacije, na osnovi katere sklepamo o lastnostih cele populacije. Pomembno je, da je vzorec verjetnosten.

4.2 VZORČENJE

Literatura:

G. Kalton in V. Vehovar (2001): Vzorčenje v anketah. Ljubljana: FDV.

Primeri uporabe:

- testiranje znanstvenih teorij (sociol., polit., ekonom. itd.),
- zdravstvo (npr. testiranje zdravil),
- vladne institucije (demografski podatki, podatki o brezposelnosti, dohodkih, življenjskih stroških, izobraževanju itd.),
- tržno raziskovanje,
- javnomnenjske raziskave.

Zgodovina: Začelo v 30-ih letih 20. stol.. Razmah in splošno sprejeto v letih pred 2. sv. vojno.

Osnovni pojmi

- **Populacija**: množica vseh proučevanih elementov.
- **Vzorec**: podmnožica populacije, na osnovi katere sklepamo o lastnostnih cele populacije.
- **Vzorčenje**: Proces izbire dela populacije, ki bo vključen v raziskavo. Pomembno: elemente iz populacije slučajno izbiramo v vzorec – verjetnostno oz. slučajnostno vzorčenje.
- **Vzorčni okvir**: seznam elementov – enot –ciljne populacije, iz katerega izbiramo elemente v vzorec.
- V**erjetnostni vzorci** (angl. *probability samples*): vsak element v populaciji ima vnaprej znano in neničelno verjetnost, da se pojavi v vzorcu.
- **Neverjetnostni vzorci** (angl. *nonprobability samples*): ne zadostijo zgornjemu pogoju.

Pomembno: Samo pri verjetnostnih vzorcih lahko uporabimo statistično sklepanje iz vzorca na populacijo. Zato jih pogosto imenujemo tudi znanstveni vzorci.

Načini verjetnostnega vzorčenja:

- enostavno slučajno vzorčenje
- sistematično vzorčenje
- stratificirano vzorčenje
- večstopenjsko vzorčenje
- sorazmerno vzorčenje
- dvofazno vzorčenje
- vzorčenje z replikacijami
- panelni vzorci

Načini neverjetnostnega vzorčenja:

- priložnostno vzorčenje
- ekspertna izbira
- kvotno vzorčenje

4.2.1 Enostavno slučajno vzorčenje

- Najpreprostejše, osnova vsem drugim metodam.
- Angl. *simple random sampling* – SRS vzorci.

Kdaj je vzorec SRS vzorec?

1. Vsaka enota ima neničelno, vnaprej znano verjetnost izbora. → Verjetnostni vzorec.
2. Vsaka enota ima enako verjetnost izbora v vzorec. → EPSEM vzorec.
3. Vsi možni vzorci so enako verjetni. → SRS vzorec.

- Vzorčenje s ponavljanjem – izbrane enote pred ponovnih izbiranjem vrnemo v populacijo. Število vseh možnih vzorcev je N^n, če je N število enot v populaciji in n število enot v vzorcu.
- Vzorčenje brez ponavljanja – izbrane enote pred ponovnih izbiranjem ne vrnemo v populacijo. Število slučajnih vzorcev brez ponavljanja pa je $\binom{N}{n}$ (ne upoštevamo vrstnega reda izbranih enot v vzorcu).

PRIMER:

Vzorčimo 100 študentov med 500 rednimi študenti 1. letnika FDV v štud. letu 2003/2004.

Populacija: 500 rednih študentov FDV v 2003/2004 N=500

Vzorec: 100 študentov n=100

Vzorčni okvir: seznam študentov (iz Referata), urejen po identifikacijskih številkah od 001 do 500.

Načini izbire vzorca: npr. loterija, vnaprej pripravljena tabela slučajnih števil, slučajno računalniško generiran seznam števil.

4.2.2 Sistematično vzorčenje

- Bolj praktična, enostavna metoda.
- Angl. *systematic sampling*.
- Iz seznama – vzorčnega okvira – izberemo v vzorec vsak *k-ti* element, pri čemer smo predhodno izbrali določen slučajni začetek.
- Verjetnostni in EPSEM vzorec, ni pa SRS vzorec.

PRIMER:

Vzorčimo 100 študentov med 500 rednimi študenti 1. letnika FDV v štud. letu 2003/2004.

Načini izbire vzorca: sistematično vzorčenje.

Delež vzorca v celotni populaciji: 100/500=0.2 oz. razmerje 1:5. Sistematični vzorec velikosti 100 dobimo tako, da določimo slučajno številko med 1 in 5 in s tem določimo prvega študenta v vzorcu. Nato pa v vzorec vključimo vsakega petega študenta. Če bi npr. slučajno izbrali začetno številko 4, bi bili v vzorcu študenti s števili 4, 9, 14, 19, itd.

Zakaj sistematično vzorčenje ni SRS?

Npr. vzorec z elementoma 2 in 4 je nemogoč. Verjetnost, da izberemo vzorec z elementoma 4 in 9 pa je 1/5. Torej, vsi vzorci niso enako verjetni.

4.2.3 Stratificirano vzorčenje

Pomagamo si z vnaprej poznanimi informacijami o populaciji.

Pri stratificiranem vzorčenju populacijo na osnovi nekih vnaprej poznanih informacij razdelimo na podpopulacije oz. **stratume**. V postopku izbiranja vzorca nato opravimo vzorčenje ločeno in neodvisno v vsakem stratumu posebej.

Proporcionalna stratifikacija: velikosti vzorca v stratumih so sorazmerne ustrezni velikosti populacije.

Disproporcionalna stratifikacija: velikosti vzorca v stratumih niso sorazmerne ustrezni velikosti populacije.

PRIMER:

Vzorčimo 100 študentov med 500 rednimi študenti (350 ženskami, 150 moškimi , torej 70% žensk in 30% moških) 1. letnika FDV v štud. letu 2003/2004.

Stratificirano vzorčenje

Populacijo 500 študentov razdelimo na dva stratuma: $N_1=350$ (ženske) in $N_2=150$ (moški).

a) V vzorcu želimo 50 študentk in 50 študentov → DISPROPORCIONALNA STRATIFIKACIJA.

Znotraj žensk vzorčimo 50 enot in znotraj moških vzorčimo 50 enot.

b) V vzorcu želimo 70% (70) študentk in 30% (30) študentov → PROPORCIONALNA STRATIFIKACIJA

Znotraj žensk vzorčimo 70 enot in znotraj moških vzorčimo 30 enot. Torej razmerje ženske:moški je 7:3 na populaciji in v vzorcu.

4.2.4 Vzorčenje v skupinah in večstopenjsko vzorčenje

Pomagamo si z vnaprej poznanimi informacijami o populaciji o tem, kako se enote združujejo v neke skupine, npr.:

- osebe v gospodinjstva, gospodinjstva v naselja, naselja v občine, občine v regije,
- dijaki v razrede, razredi v šole.

Postopek vzorčenja:

Na prvi stopnji slučajno izberemo vzorec skupin.

Na drugi stopnji znotraj izbranih skupin

a) izberemo v vzorec vse enote (**vzorčenje v skupinah**) ali

b) enote vzorčimo (**dvostopenjsko vzorčenje**).

Možno je tudi vzorčenje na več kot dveh stopnjah - večstopenjsko vzorčenje.

PRIMER

Redni študenti 1. letnika FDV so vpisani v 13 programov.

Vzorčenje v skupinah: Na prvi stopnji slučajno izberemo 3 izmed 13 programov, nato pa znotraj izbranih programov (skupin) izberemo vse študente. Pričakujemo vzorec velikosti cca. 150 enot.

Dvostopenjsko vzorčenje: Na prvi stopnji slučajno izberemo 3 izmed 13 programov (skupin), nato pa znotraj izbranih skupin slučajno izberemo še vzorec 20 študentov, da dobimo vzorec velikosti 60.

Primerjava: stratificirano vzorčenje - vzorčenje v skupinah oz. večstopenjsko vzorčenje

Stratificirano vzorčenje	***Vzorčenje v skupinah oz. večstopenjsko vzorčenje***
Vsi stratumi so vključeni v vzorec. Povečuje natančnost ocen. Zaželena je notranja homogenost stratumov.	V vzorec je vključen le vzorec skupin. Zmanjšuje natančnost ocen, zato je pomembno, da izbrane skupine predstavljajo tudi ostale skupine. Zaželena je notranja homogenost skupin. Cenejše.

4.3 VZORČNE STATISTIKE IN NJIHOVE PORAZDELITVE

4.3.1 Osnovni pojmi in oznake

Statistične karakteristike imenujemo:

- **parametri**, če so izračunane ali ocenjene za populacijo (fiksna števila, običajno neznana),
- **statistike**, če so izračunane na vzorcu (variabilna števila, se spreminjajo od vzorca do vzorca).

Oznake:

Statistične karakteristike

- populacije (torej parametre) označujemo z grškimi črkami, npr.
  - število enot N,
  - aritmetična sredina μ,
  - standardni odklon σ,
  - delež π,
- vzorca (torej statistike) označujemo z latinskimi črkami, npr.
  - število enot n,
  - aritmetična sredina $\overline{X}$,
  - standardni odklon s,
  - delež p.

Večino zakonitosti inferenčne statistike je bilo izpeljanih pod predpostavko, da izbrane enote pred ponovnih izbiranjem vrnemo v populacijo (vzorec s ponavljanjem). Če je velikost vzorca v primerjavi s populacijo majhna, se ne pregrešimo preveč, če imamo za slučajni vzorec tudi vzorec, ki nastane s slučajnim izbiranjem brez ponavljanja.

Vsa pravila za izračun statističnih karakteristik na populaciji (torej parametrov), ki jih že poznamo, veljajo tudi za izračun statističnih karakteristik na vzorcu (torej statistik), le da v njih uporabimo druge oznake. Izjema sta varianca in iz nje izračunani standardni odklon.

	Populacija	*Vzorec*
Aritmetična sredina	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$
Varianca	$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$	$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2$
Standardni odklon	$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}$	$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2}$
Delež enot z določeno lastnostjo (relativna frekvenca vrednosti x_i)	$\pi = \frac{f_i}{N}$	$p = \frac{f_i}{n}$

4.3.2 Porazdelitve vzorčnih statistik

Zakaj nas zanimajo porazdelitve vzorčnih statistik?

Ker so osnova za dva osnovna postopka sklepanja iz vzorca na populacijo:

1. statistično ocenjevanje značilnosti populacije (s točkovnimi ocenami ali intervali zaupanja) in
2. preverjanje domnev (testiranje hipotez).

Gre za zakonitost zvez med osnovno populacijo in populacijo vseh možnih vzorcev iz te populacije, temeljijo pa na verjetnostnem računu.

Definirajmo ...

Denimo, da imamo v populaciji *N* enot in da iz te populacije slučajno izbiramo (s ponavljanjem) *n* enot v enostavni slučajni vzorec ali na kratko slučajni vzorec (vsaka enota ima enako verjetnost, da bo izbrana v vzorec, in vsi vzorci so enako verjetni).

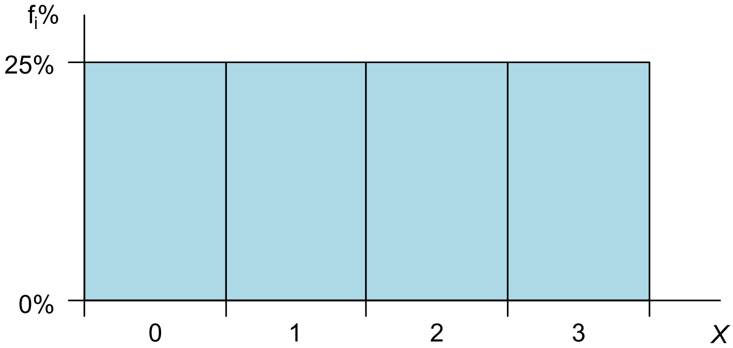
Predstavljamo si, da smo iz populacije zbrali vse možne vzorce. Dobili smo **populacijo vseh možnih vzorcev**.

Primer:

Vzemimo populacijo z $N = 4$ enotami, ki imajo naslednje vrednosti spremenljivke X:

0, 1, 2, 3

Grafično si lahko porazdelitev spremenljivke *X* predstavimo s histogramom:



Izračunamo lahko populacijsko aritmetično sredino in standardni odklon:

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i = \frac{3}{2}$$

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2 = \frac{5}{4}$$

$$\sigma = 1.12$$

Sedaj pa tvorimo vse možne vzorce velikosti $n = 2$ s ponavljanjem (vzorcev je $M = N^n = 4^2 = 16$) in na vsakem izračunajmo vzorčno aritmetično sredino $\overline{X}$:

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

vzorci	$\overline{X}$	vzorci	$\overline{X}$
0, 0	0	2, 0	1
0, 1	0.5	2, 1	1.5
0, 2	1	2, 2	2
0, 3	1.5	2, 3	2.5
1, 0	0.5	3, 0	1.5
1, 1	1	3, 1	2
1, 2	1,5	3, 2	2.5
1, 3	2	3, 3	3

enote

spremenljivka: vzorčna aritmetična sredina

vrednosti spremenljivke

Vse možne vzorce lahko obravnavamo kot neko novo populacijo, aritmetično sredino na vzorcih pa kot slučajno spremenljivko.

Enota: 1 vzorec

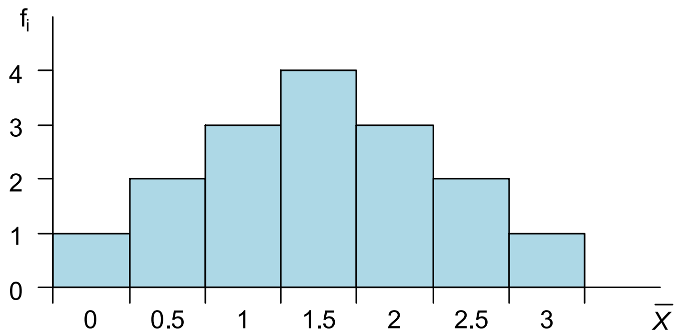
Velikost populacije: $M = N^n = 4^2 = 16$ – vsi možni vzorci

Spremenljivka: $\overline{X}$ vzorčna aritmetična sredina/povprečje (aritmetična sredina, izračunana za vsak vzorec)

Za spremenljivko $\overline{X}$ vzorčna aritmetična sredina/povprečje lahko zapišemo frekvenčno porazdelitev:

$\overline{X}_i$ Vzorčna aritmetična sredina	f_i Število vseh vzorcev	f_i% Delež vseh vzorcev – relativna frekvenca
0	1	1/16
0,5	2	2/16
1	3	3/16
1,5	4	4/16
2	3	3/16
2,5	2	2/16
3	1	1/16
	M=16	1

Spremenljivko vzorčna aritmetična sredina $\overline{X}$ lahko grafično predstavimo s histogramom:



Izračunamo lahko aritmetično sredino (imenovali jo bomo matematično upanje) in varianco (imenovali jo bomo disperzija) slučajne spremenljivke vzorčna aritmetična sredina $\overline{X}$.

Matematično upanje (angl. *expected value*) – aritmetična sredina slučajne spremenljivke.
Iz verjetnostne teorije: matematično upanje slučajne spremenljivke *X* (*E(X)*) je pričakovana vrednost slučajne spremenljivke, če poskus velikokrat ponovimo. Predstavlja povprečje realizacij spremenljivke v velikem številu ponovitev poskusa.
V našem primeru: matematično upanje slučajne spremenljivke vzorčna aritmetična sredina $\bar{X}$ ($E(\bar{X})$) predstavlja pričakovano vrednost vzorčne aritmetične sredine, ki jo izračunamo tako, da izračunamo povprečje vzorčnih aritmetičnih sredin na vseh možnih vzorcih.

$$E(\overline{X}) = \mu_{\overline{X}} = \frac{1}{M}\sum_{i=1}^{M} f_i \overline{X}_i = \frac{1}{16}(1\cdot 0 + 2\cdot 0{,}5 + 3\cdot 1 + 4\cdot 1{,}5 + 3\cdot 2 + 2\cdot 2{,}5 + 1\cdot 3) = 1{,}5$$

Disperzija (angl. *dispersion*) – varianca slučajne spremenljivke.
Iz verjetnostne teorije: disperzija ali varianca meri variabilnost slučajne spremenljivke *X* okoli svoje aritmetične sredine, torej okoli matematičnega upanja *E(X).* Izračunamo jo kot povprečni kvadratni odklon vrednosti od aritmetične sredine.

V našem primeru: disperzija ali varianca slučajne spremenljivke vzorčna aritmetična sredina $\bar{X}$ meri variabilnost vzorčnih aritmetičnih sredin okoli svojega povprečja, torej okoli matematičnega upanja $E(\bar{X})$.

Koren iz disperzije je **standardni odklon.**

$$D(\overline{X}) = \sigma^2_{\overline{X}} = \frac{1}{M}\sum_{i=1}^{m} f_i(\overline{X}_i - \mu_{\overline{X}})^2 = \frac{1}{16}(1\cdot(0-1{,}5)^2 + 2\cdot(0{,}5-1{,}5)^2 + \ldots + 1\cdot(3-1{,}5)^2 = \frac{5}{8} = 0{,}625$$

$$\sqrt{D(\overline{X})} = \sigma_{\overline{X}} = \sqrt{0{,}625} = 0{,}791$$

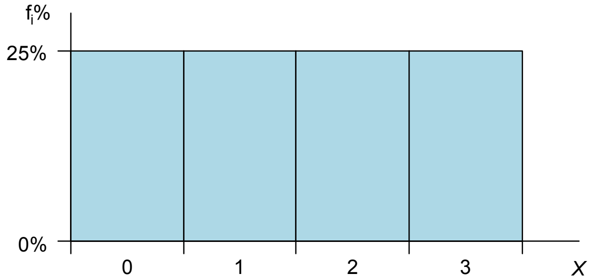
S tem primerom smo pokazali, da je statistika $\overline{X}$ "vzorčna aritmetična sredina" slučajna spremenljivka s svojo porazdelitvijo, s svojo aritmetično sredino in svojim standardnim odklonom.

Povzetek:

"Originalna" populacija

$N = 4$, spremenljivka X

Enakomerna porazdelitev.



$$\mu_x = \frac{3}{2}$$

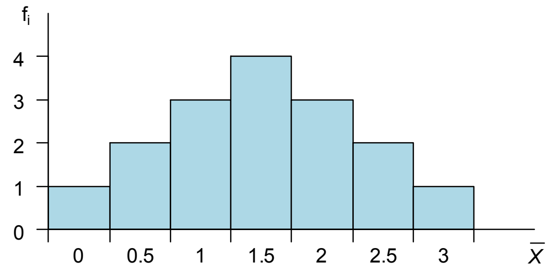
$$\sigma^2_X = \frac{5}{4}$$

$$\sigma_X = \sqrt{\frac{5}{4}}$$

Populacija vseh vzorcev

$M = N^n = 16$, spremenljivka $\bar{X}$

Simetrična, unimodalna porazdelitev.



$$E(\overline{X}) = \mu_{\overline{X}} = \frac{3}{2} = \mu_X$$

$$\sqrt{D(\overline{X})} = \sigma_{\overline{X}} = \sqrt{\sigma^2_{\overline{x}}} = \sqrt{\frac{5}{8}} = \sqrt{\frac{5}{4\cdot 2}} = \sqrt{\frac{\sigma_X{}^2}{n}} = \frac{\sigma_X}{\sqrt{n}}$$

Katere značilnosti porazdelitve vzorčnih statistik lahko opazimo iz tega primera?

1. Porazdelitev vzorčnih statistik se ne ujema s porazdelitvijo originalne spremenljivke na populaciji.
2. Povprečje (matematično upanje) vseh vzorčnih aritmetičnih sredin je ravno enako povprečju vrednosti na populaciji. $\mu_{\bar{X}} = \mu_X$

3. Standardni odklon vzorčnih aritmetičnih sredin je ravno enak standardnemu odklonu vrednosti na populaciji, deljeno s korenom iz velikosti vzorca. Imenovali ga bomo **standardna napaka**. Meri razpršenost vzorčnih statistik, v tem konkretnem primeru vzorčnih aritmetičnih sredin. $\sigma_{\overline{X}} = \frac{\sigma_X}{\sqrt{n}}$

Pogledali smo si, kaj se zgodi, če iz neke populacije tvorimo vse možne vzorce in za vsak možen vzorec izračunamo aritmetično sredino. V tem primeru lahko vzorčne aritmetične sredine obravnavamo kot novo slučajno spremenljivko, lahko jo grafično predstavimo in zanjo izračunamo matematično upanje in disperzijo.

Na enak način bi lahko na vseh možnih vzorcih izračunali katerokoli statistiko (poleg aritmetične sredine tudi mediano, modus, kvantile, varianco, delež itd.) in zanjo zapisali porazdelitev.

Če torej iz neke populacije naredimo vse možne vzorce velikosti *n* in če za vsak vzorec izračunamo neko statistiko, se ta vzorčna statistika razlikuje od vzorca do vzorca. Na ta način dobimo neko **porazdelitev vzorčne statistike**, ki jo imenujemo **vzorčna porazdelitev** (angl. *sampling distribution*). Npr. porazdelitev vzorčnih aritmetičnih sredin, deležev itd.

Splošen zapis:

V splošnem bomo porazdelitve vzorčnih statistik zapisali na naslednji način:

1. *g* je vzorčna statistika, s katero ocenjujemo populacijski parameter γ.
2. Porazdelitev vzorčne statistike npr.

$$\boxed{g: N\,(\underbrace{E(g),\ SE(g)})}$$

vzorčna statistika   oblika porazdelitve   parametri porazdelitve

3. Oblika porazdelitve je lahko normalna (*N*), Studentova t (*t*), hi-kvadrat (χ^2) itd. Kakšna bo oblika porazdelitve, je odvisno od tega, za katero vzorčno statistiko gre in kako velike vzorce delamo.

4.3.3 Porazdelitev vzorčnih aritmetičnih sredin

Denimo, da se spremenljivka X na populaciji porazdeljuje normalno $N(\mu, \sigma)$. Iz populacije tvorimo vse možne vzorce velikosti n in na vsakem vzorcu izračunamo vzorčno aritmetično sredino $\overline{X}$. Dokazati se da, da je porazdelitev vzorčnih aritmetičnih sredin na vseh možnih vzorcih normalna, kjer je:

- matematično upanje vzorčnih aritmetičnih sredin (t.j. povprečje vzorčnih aritmetičnih sredin preko vseh vzorcev) enako aritmetični sredini spremenljivke na populaciji (torej populacijskemu parametru): $E(\overline{X}) = \mu_{\overline{X}} = \mu$
- standardni odklon vzorčnih aritmetičnih sredin (im. ga standardna napaka) enak

  $SE(\overline{X}) = \sqrt{D(\overline{X})} = \sqrt{\dfrac{\sigma^2}{n}} = \dfrac{\sigma}{\sqrt{n}}$ če tvorimo vzorce iz končne populacije s ponavljanjem

  $SE(\overline{X}) = \dfrac{\sigma}{\sqrt{n}}\sqrt{\dfrac{N-n}{N-1}}$ če tvorimo vzorce iz končne populacije brez ponavljanja (če je vzorec v primerjavi s populacijo zelo majhen, je popravek $\cong 1$, torej ni razlike med vzorčenjem z ali brez ponavljanja)

Torej: $\overline{X} : N(E(\overline{X}), SE(\overline{X}))$

$\overline{X} : N(\mu, \dfrac{\sigma}{\sqrt{n}})$ ⟵ **porazdelitev vzorčnih aritmetičnih sredin**

Ta zakonitost velja za zadosti velike vzorce ($n \geq 30$), tudi če se spremenljivka X na populaciji ne porazdeljuje normalno.

Primer:
Denimo, da se spremenljivka inteligenčni kvocient (X) na populaciji porazdeljuje normalno z aritmetično sredino μ=100 in standardnim odklonom σ = 15.

X: N (100, 15)

Iz populacije tvorimo vzorce s ponavljanjem velikosti 225 enot (n=225).

a) Kako se porazdeljujejo vzorčne aritmetične sredine $\bar{X}$?
b) Grafično predstavite obe porazdelitvi (porazdelitev spremenljivke X – inteligenčni kvocient in spremenljivke $\bar{X}$ – vzorčne aritmetične sredine).
c) Kolikšne vzorčne aritmetične sredine ima 90% vseh vzorcev (simetrično na povprečje)?
d) Kakšen je ta interval, če povečamo velikost vzorca na n = 2500?
e) Primerjajte porazdelitve vzorčnih aritmetičnih sredin pri različno velikih vzorcih: n = 225, 625, 2500.

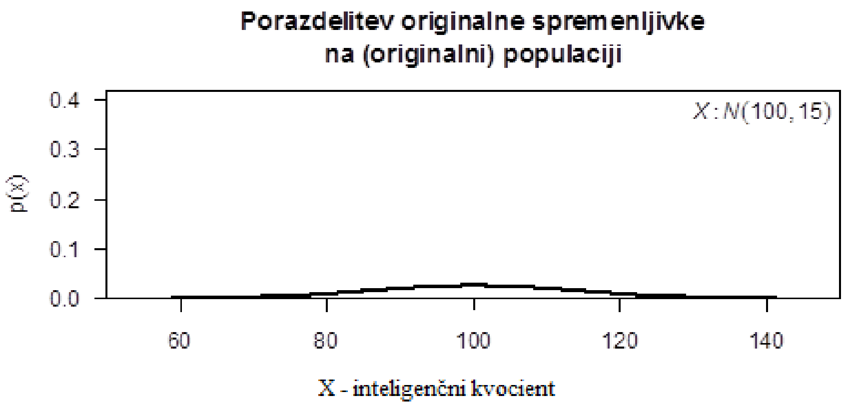
a) Porazdelitev vzorčnih aritmetičnih sredin (n = 225)
Vzorčne aritmetične sredine se porazdeljujejo normalno z naslednjima matematičnim upanjem in standardno napako:

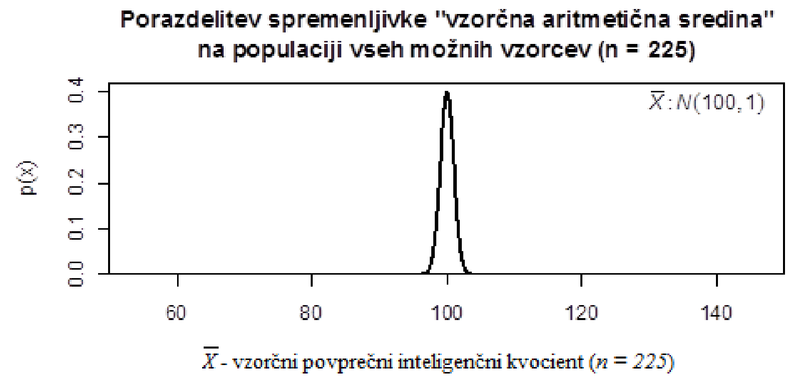
$$E(\overline{X}) = 100 \qquad SE(\overline{X}) = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{225}} = 1$$

$$\overline{X} : N(E(\overline{X}),\ SE(\overline{X}))$$
$$\overline{X} : N(\mu, \frac{\sigma}{\sqrt{n}})$$
$$\overline{X} : N(100,\ 1)$$

b) Porazdelitvi obeh spremenljivk

Populacija: *X: N (100, 15)* Populacija vseh vzorcev: $\overline{X}$ *: N(100, 1)*

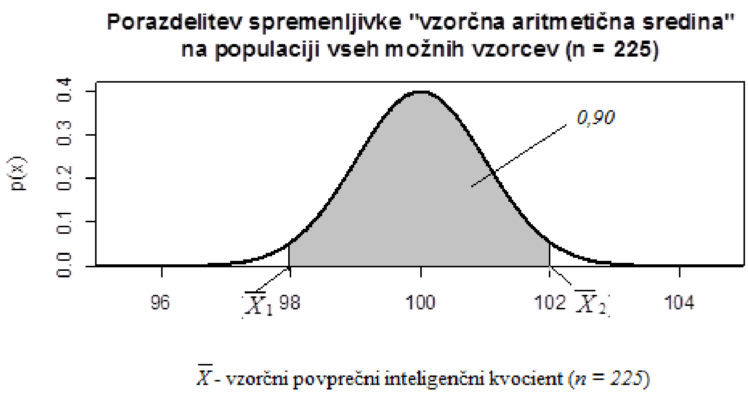




c) Vzorčne aritmetične sredine 90% vzorcev (simetrično na povprečje)

Iščemo takšne vrednosti spremenljivke "vzorčna aritmetična sredina" $\overline{X}$, da bo veljalo:

$$P(\overline{X}_1 < \overline{X} < \overline{X}_2) = 0{,}90$$



Pomagamo si s standardizirano slučajno spremenljivko Z, za katero velja:

$$z_i = \frac{\overline{X}_i - E(\overline{X})}{SE(\overline{X})} = \frac{\overline{X}_i - \mu}{\frac{\sigma}{\sqrt{n}}}$$

Izračunamo meje 90% intervala za spremenljivko Z in nato izračunamo še meje intervala za spremenljivko $\overline{X}$:

$$P(z_1 < Z < z_2) = 0{,}90$$
$$H(z_2) = 0{,}90 + (1 - 0{,}90)/2 = 0{,}95 \Rightarrow z_2 = +1{,}65$$
$$H(z_1) = (1 - 0.90)/2 = 0{,}05 \Rightarrow z_1 = -1{,}65$$

$$\overline{X}_1 = z_1 \cdot SE(\overline{X}) + E(\overline{X}) = -1{,}65 \cdot 1 + 100 = 98{,}35$$
$$\overline{X}_2 = z_2 \cdot SE(\overline{X}) + E(\overline{X}) = 1{,}65 \cdot 1 + 100 = 101{,}65$$

$$P(98{,}35 < \overline{X} < 101{,}65) = 0{,}90$$

90% vseh slučajnih vzorcev 225 enot s ponavljanjem bo imelo vzorčna povprečja za inteligenčni kvocient od 98,35 do 101,65.

d) Vzorčne aritmetične sredine 90% vzorcev (simetrično na povprečje) v primeru večjega vzorca (n=2500)

V primeru večjega vzorca je standardna napaka porazdelitve vzorčnih aritmetičnih sredin manjša.

$$E(\overline{X}) = 100 \qquad SE(\overline{X}) = \frac{\sigma}{\sqrt{n}} = \frac{15}{\sqrt{2500}} = 0{,}3$$

$$\overline{X} : N(100{,}0.3)$$

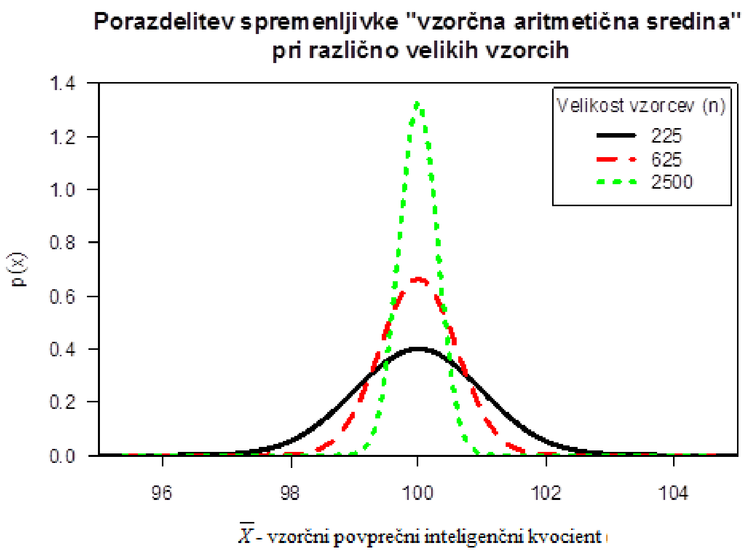
Zaradi manjše variabilnosti vzorčnih aritmetičnih sredin je ožji tudi tak interval:

$$\overline{X}_1 = -z_1 \cdot SE(\overline{X}) + E(\overline{X}) = -1{,}65 \cdot 0{,}3 + 100 = 99{,}5$$
$$\overline{X}_2 = z_1 \cdot SE(\overline{X}) + E(\overline{X}) = 1{,}65 \cdot 0{,}3 + 100 = 100{,}5$$

$$P(99{,}5 < \overline{X} < 100{,}5) = 0{,}90$$

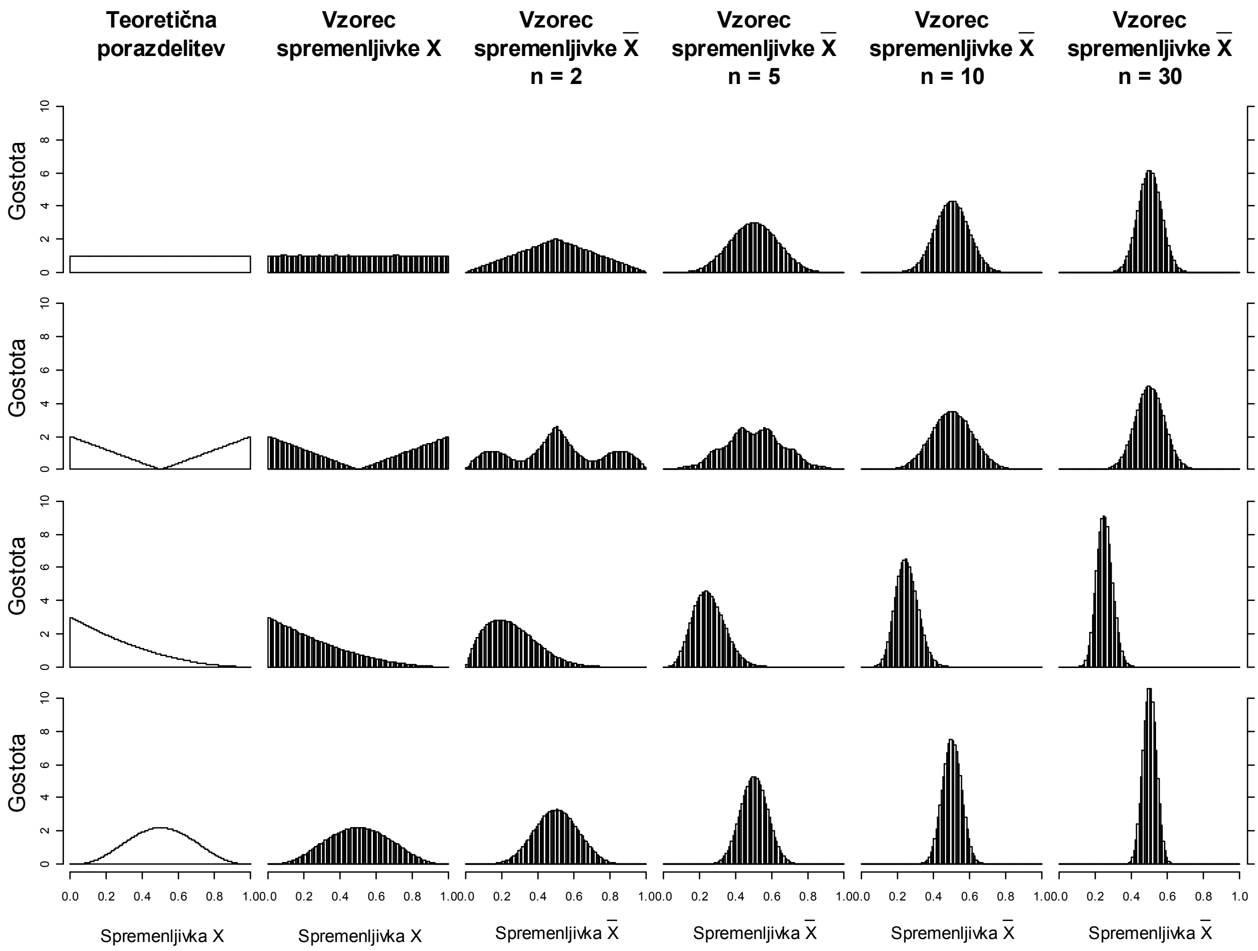
e) Porazdelitev vzorčnih aritmetičnih sredin pri različno velikih vzorcih

n = 225	$\overline{X} : N(100,\ 1)$
n = 625	$\overline{X} : N(100{,}0.6)$
n = 2500	$\overline{X} : N(100{,}0.3)$



NEKAJ ZAKONITOSTI ...

1. Za dovolj velike vzorce ($n \geq 30$) se vzorčne aritmetične sredine porazdeljujejo približno normalno (z zgoraj definiranima matematičnim upanjem in standardno napako), ne glede na obliko porazdelitve na populaciji.
2. Z večanjem vzorca se zmanjšuje variabilnost vzorčnih aritmetičnih sredin (zmanjšuje se $SE(\overline{X}) = \frac{\sigma}{\sqrt{n}}$).



ŠE NEKAJ OSNOVNIH POJMOV ...

Standardni odklon statistike imenujemo **standardna napaka**.

- Standardni odklon meri variabilnost vrednosti spremenljivke X na (originalni) populaciji $\rightarrow$ mera variabilnosti posameznih vrednosti spremenljivke X.
- Standardna napaka meri variabilnost vzorčnih statistik (npr. vzorčnih aritmetičnih sredin $\overline{X}$ na populaciji vseh možnih vzorcev $\rightarrow$ mera variabilnosti vzorčnih ocen (statistik). Je mera za natančnost vzorčne ocene iz slučajnega vzorca.

Vzorčna aritmetična sredina $\overline{X}_i$, izračunana na i-tem vzorcu, je ocena parametra populacijske aritmetične sredine μ. To je le ena od vrednosti, ki jo lahko zavzame slučajna spremenljiva $\overline{X}$. Vzorčno aritmetično sredino $\overline{X}$ imenujemo tudi **cenilka** populacijske aritmetične sredine μ.

Cenilka je slučajna spremenljivka, katere vrednosti so odvisne od vrednosti na vzorcu. Je namreč statistična karakteristika, ki je izračunana na vzorcu. Namenjena je ocenjevanju parametra.

Najbolj pomembna merila za kakovost cenilke so:

1. **nepristranskost:** cenilka je nepristranska, če je povprečje vseh vzorčnih ocen (matematično upanje cenilke) enako ocenjevanemu parametru;
2. **doslednost:** cenilka je dosledna, če se z večanjem vzorca vzorčna ocena bliža parametru,
3. **učinkovitost**: cenilka je učinkovita, kadar je v primerjavi s kako drugo cenilko učinkovitejša, če ima manjšo varianco,
4. **zadostnost**: cenilka je zadostna, če uporabi vse informacije vzorca, ki so pomembne za oceno parametra.

Najpomembnejša lastnost je nepristranskost. Le z nepristranskimi cenilkami lahko ocenjujemo populacijske parametre. Vzorčna aritmetična sredina je nepristranska cenilka populacijske aritmetične sredine, ker velja:

$$E(\overline{X}) = \mu$$

4.3.4 Porazdelitev vzorčnih deležev

Denimo, da želimo na populaciji oceniti delež enot π z določeno lastnostjo. Zato na vsakem vzorcu poiščemo vzorčni delež p. Pokazati se da, da se za dovolj velike slučajne vzorce s ponavljanjem (za deleže okoli 0.5 je dovolj 20 enot ali več) vzorčni deleži p porazdeljujejo približno normalno

- z matematičnim upanjem vzorčnih deležev (t.j. povprečjem vzorčnih deležev preko vseh vzorcev), ki je enako deležu na populaciji (torej populacijskemu parametru):

$$E(p) = \mu_p = \pi$$

- s standardnim odklonom vzorčnih deležev (standardno napako), ki je enak

$$SE(p) = \sqrt{D(p)} = \sqrt{\frac{\pi(1-\pi)}{n}}$$

Torej: $p : N(E(p), SE(p))$

$$p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}}) \longleftarrow$$ **porazdelitev vzorčnih deležev**

Vzorčni delež je nepristranska cenilka populacijskega deleža, ker velja $E(p) = \pi$.

Primer

V izbrani populaciji prebivalcev je polovica žensk (π = *0.5*). Če tvorimo vzorce po n = 25 enot, nas zanima, kakšna je verjetnost, da je v vzorcu več kot 55% žensk?

To pomeni, da iščemo verjetnost, da slučajna spremenljivka p "vzorčni delež" zavzame vrednosti, večje od 0.55:

$$P(p > 0.55)=?$$

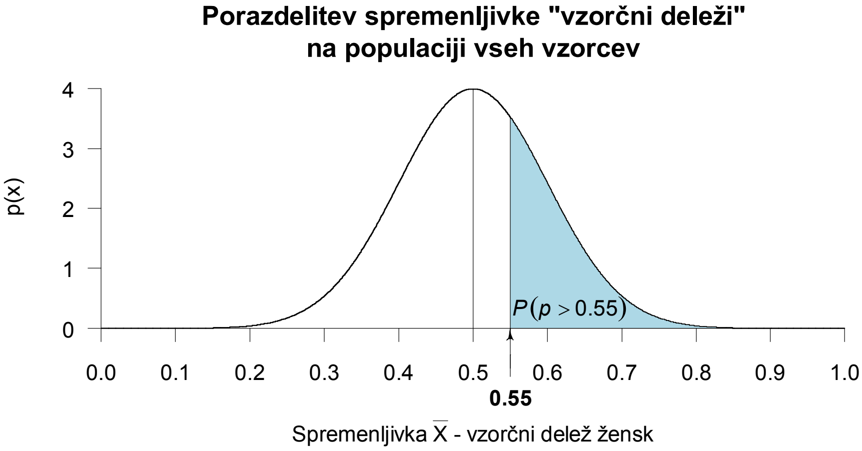
Vemo, da se vzorčni deleži p na vseh možnih vzorcih porazdeljujejo približno normalno z naslednjima matematičnim upanjem in standardno napako:

$$p : N(E(p), SE(p))$$

$$p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$$

$$p : N(0.5, \sqrt{\frac{0.5(1-0.5)}{25}})$$

$$p : N(0.5, 0.1)$$



Ker lahko tudi vzorčne deleže standardiziramo

$$z_i = \frac{p_i - E(p)}{SE(p)} = \frac{p_i - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}}$$

lahko verjetnost iskanega intervala računamo s pomočjo standardizirane spremenljivke Z:

$$P(p > 0.55) = P(Z > \frac{0.55 - 0.5}{0.1}) = P(Z > 0.5) = 1 - H(0.5) = 1 - 0.6915 = 0.3085$$

Če bi slučajno izbrali en vzorec, bi bila verjetnost, da je v njem delež žensk večji od 0.55, približno 31%. Pričakujemo lahko, da bo pri približno 31% vzorcev delež žensk večji od 0.55.

Poglejmo, kakšna je verjetnost, če bi tvorili vzorce velikost $n = 2500$. V tem primeru je porazdelitev vzorčnih deležev naslednja:

$$p : N(E(p), SE(p))$$

$$p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$$

$$p : N(0.5, \sqrt{\frac{0.5(1-0.5)}{2500}})$$

$$p : N(0.5, 0.01)$$

Verjetnost iskanega intervala je:

$$P(p > 0.55) = P(Z > \frac{0.55 - 0.5}{0.01}) = P(Z > 5) = 1 - H(5) = 1 - 1 \approx 0 \; (2.9 \cdot 10^{-7})$$

Če je delež žensk na populaciji 0.5, potem ni verjetno, da bi na vzorcu 2500 enot dobili več kot 55% žensk.

4.3.5 Porazdelitev razlik vzorčnih aritmetičnih sredin

Denimo, da imamo dve populaciji velikosti N_1 in N_2 in se spremenljivka X na prvi populaciji porazdeljuje normalno $X: N(\mu_1, \sigma_1)$, na drugi populaciji pa $X: N(\mu_2, \sigma_2)$. V vsaki od obeh populacij tvorimo neodvisno slučajne vzorce (s ponavljanjem) velikosti n_1 in n_2. Na vsakem vzorcu prve populacije izračunamo vzorčno aritmetično sredino $\overline{X}_1$ in podobno na vsakem vzorcu druge populacije $\overline{X}_2$. Dokazati se da, da je (za dovolj velike vzorce, $n \geq 30$) porazdelitev razlik vzorčnih aritmetičnih sredin ($\overline{X}_1 - \overline{X}_2$) približno normalna, kjer je

- matematično upanje razlik vzorčnih aritmetičnih sredin (t.j. povprečje razlik vzorčnih aritmetičnih sredin preko vseh parov vzorcev iz obeh populacij) enako razliki populacijskih aritmetičnih sredin (torej populacijskemu parametru)

$$E(\overline{X}_1 - \overline{X}_2) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2$$

- standardni odklon razlik vzorčnih aritmetičnih sredin (standardna napaka) enak

$$SE(\overline{X}_1 - \overline{X}_2) = \sqrt{D(\overline{X}_1 - \overline{X}_2)} = \sqrt{D(\overline{X}_1) + D(\overline{X}_2)} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Torej:

$$\overline{X}_1 - \overline{X}_2 : N(E(\overline{X}_1 - \overline{X}_2), SE(\overline{X}_1 - \overline{X}_2))$$

$$\overline{X}_1 - \overline{X}_2 : N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}) \quad \longleftarrow$$

porazdelitev razlik vzorčnih aritmetičnih sredin

Primer:
Dvema populacijama študentov na neki univerzi (tehnikom in družboslovcem) so izmerili neko sposobnost s povprečjema $\mu_d = 80$ in $\mu_t = 70$ točk ter standardnim odklonom, ki je na obeh populacijah enak, $\sigma_d = \sigma_t = 7$ točk. Kolikšna je verjetnost, da je aritmetična sredina slučajnega vzorca družboslovcev (velikost $n_d = 36$ enot) večja za več kot 12 točk od aritmetične sredine slučajnega vzorca tehnikov (velikost $n_t = 64$ enot)?

To pomeni, da iščemo verjetnost, da slučajna spremenljivka $\bar{X}_d - \bar{X}_t$ "razlika vzorčnih aritmetičnih sredin" zavzame vrednosti, večje od 12:

$$P(\overline{X}_d - \overline{X}_t) > 12) = ?$$

Vemo, da se razlike vzorčnih aritmetičnih sredin na vseh možnih vzorcih porazdeljujejo približno normalno z naslednjima matematičnim upanjem in standardno napako:

$$\overline{X}_d - \overline{X}_t : N(E(\overline{X}_d - \overline{X}_t), SE(\overline{X}_d - \overline{X}_t))$$

$$\overline{X}_d - \overline{X}_t : N(\mu_d - \mu_t, \sqrt{\frac{\sigma_d^2}{n_d} + \frac{\sigma_t^2}{n_t}})$$

$$\overline{X}_d - \overline{X}_t : N(80 - 70, \sqrt{\frac{7^2}{36} + \frac{7^2}{64}})$$

$$\overline{X}_d - \overline{X}_t : N(10,\ 1.46)$$

Ker lahko tudi razlike vzorčnih aritmetičnih sredin standardiziramo

$$z_i = \frac{\overline{X}_1 - \overline{X}_2 - E(\overline{X}_1 - \overline{X}_2)}{SE(\overline{X}_1 - \overline{X}_2)} = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

lahko verjetnost iskanega intervala računamo s pomočjo standardizirane spremenljivke Z:

$$P(\overline{X}_d - \overline{X}_t > 12) = P(Z > \frac{12 - E(\overline{X}_d - \overline{X}_t)}{SE(\overline{X}_d - \overline{X}_t)}) = P(Z > \frac{12 - 10}{1.46}) = P(Z > 1.37) = 1 - H(1.37) = 1 - 0{,}9147 = 0.0853$$

Če bi slučajno izbrali po en vzorec iz vsake populacije študentov (en par vzorcev), bi bila verjetnost, da je povprečje družboslovcev za 12 točk večje od povprečja tehnikov, približno 8.5%.
8.5% parov vzorcev je torej takih, da je povprečje družboslovcev večje od povprečja tehnikov za 12 točk.

4.3.6 Porazdelitev razlik vzorčnih deležev

Podobno kot pri porazdelitvi razlik vzorčnih aritmetičnih sredin naj bosta dani dve populaciji velikosti N_1 in N_2 z deležema enot z neko lastnostjo π_1 in π_2. Iz prve populacijo tvorimo slučajne vzorce (s ponavljanjem) velikosti n_1 in na vsakem izračunamo delež enot s to lastnostjo p_1. Podobno naredimo tudi na drugi populaciji: (neodvisno od prvih vzorcev) tvorimo slučajne vzorce (s ponavljanjem) velikosti n_2 in na njih določimo deleže p_2. Pokazati se da, da je (za dovolj velike vzorce) porazdelitev razlik vzorčnih deležev $p_1 - p_2$ približno normalna, kjer je

- matematično upanje razlik vzorčnih deležev (t.j. povprečje razlik vzorčnih deležev preko vseh parov vzorcev iz obeh populacij) enako razliki populacijskih deležev (torej populacijskemu parametru) $E(p_1 - p_2) = E(p_1) - E(p_2) = \pi_1 - \pi_2$

- standardni odklon razlik vzorčnih deležev (standardna napaka) enak

$$SE(p_1 - p_2) = \sqrt{D(p_1 - p_2)} = \sqrt{D(p_1) + D(p_2)} = \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}$$

Torej: $p_1 - p_2 : N(E(p_1 - p_2), SE(p_1 - p_2))$

$$p_1 - p_2 : N(\pi_1 - \pi_2, \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}})$$ ← **porazdelitev razlik vzorčnih deležev**

Primer:
Za dve populaciji študentov na neki univerzi (tehniki in družboslovci) imamo podatek o tem, koliko jih redno uporablja internet. Izkazalo se je, da na obeh populacijah internet redno uporablja 95% vseh študentov. Recimo, da tvorimo vzorec študentov tehnike ($n_t=64$) in vzorec študentov družboslovja ($n_d=36$). Kolikšna je verjetnost, da bi dobili taka vzorca, da bi bil delež rednih uporabnikov interneta med tehniki večji kot med družboslovci?

To pomeni, da iščemo verjetnost, da slučajna spremenljivka p_t-p_d "razlika vzorčnih deležev" zavzame vrednosti, večje od 0:

$P(p_t > p_d) = P(p_t - p_d) > 0 = ?$

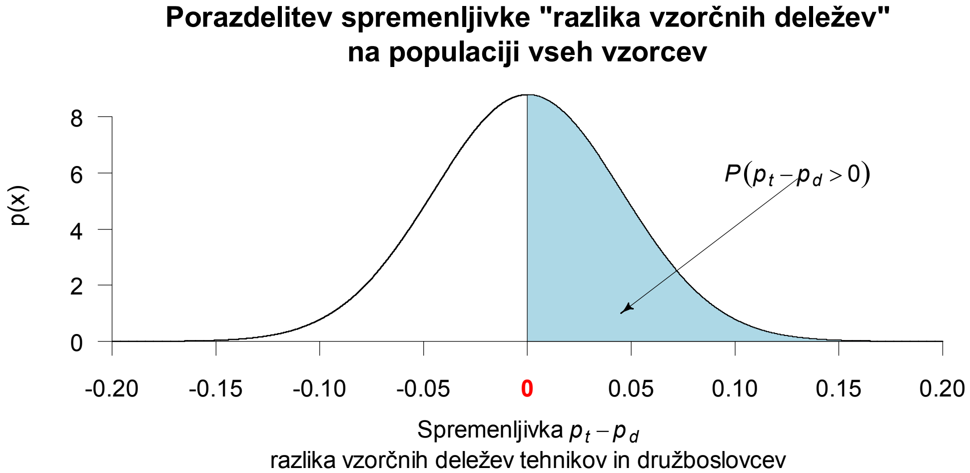
Vemo, da se razlike vzorčnih deležev na vseh možnih vzorcih porazdeljujejo približno normalno z naslednjima matematičnim upanjem in standardno napako:

$$p_t - p_d : N(E(p_t - p_d), SE(p_t - p_d))$$

$$p_t - p_d : N(\pi_t - \pi_d, \sqrt{\frac{\pi_t(1-\pi_t)}{n_t} + \frac{\pi_d(1-\pi_d)}{n_d}})$$

$$p_t - p_d : N(0.95 - 0.95, \sqrt{\frac{0.95(1-0.95)}{64} + \frac{0.95(1-0.95)}{36}})$$

$$p_t - p_d : N(0,\ 0.0454)$$



Ker lahko tudi razlike vzorčnih deležev standardiziramo

$$z_i = \frac{p_1 - p_2 - E(p_1 - p_2)}{SE(p_1 - p_2)} = \frac{p_1 - p_2 - (\pi_1 - \pi_2)}{\sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}}}$$

lahko verjetnost iskanega intervala računamo s pomočjo standardizirane spremenljivke Z:

$$P(p_t - p_d > 0) = P(Z > \frac{p_t - p_d - E(p_t - p_d)}{SE(p_t - p_d)} = (Z > \frac{0-0}{0.0454}) = P(Z > 0) = 1 - H(0) = 1 - 0.5 = 0.5$$

Če bi slučajno izbrali po en vzorec iz vsake populacije študentov (en par vzorcev), bi bila verjetnost, da je delež tehnikov, ki redno uporabljajo internet, večji od deleža družboslovcev, ki redno uporabljajo internet, 50%.
Kljub temu da na populacijah tehnikov in družboslovcev ni razlik v pogostosti uporabe interneta, bi v polovici primerov parov vzorcev dobili tak par, kjer bi bil delež rednih uporabnikov interneta med tehniki večji od deleža med družboslovci.

4.4 INTERVALI ZAUPANJA

4.4.1 Uvod

Pogledali smo si, kako lahko preko teorije vzorčenja pridobimo informacije o vzorcih, ki so bili slučajno izbrani iz znane populacije. V praksi pa je potrebno ravno obratno: sklepati o populaciji na osnovi vzorcev, ki so bili slučajno izbrani iz te (neznane) populacije. V tem primeru gre za **statistično sklepanje** (angl. *statistical inference*), s katerim se ukvarja področje **inferenčne statistike**.

Statistično sklepanje uporablja verjetnostno teorijo za to, da pove, koliko lahko zaupamo rezultatom, pridobljenim na verjetnostnem vzorcu.

Statistično sklepanje = sklepanje iz vzorca na populacijo:

1. ocena populacijskih parametrov na osnovi vzorčnih statistik – točkovne ocene, intervali zaupanja;
2. testiranje domnev o populaciji.

Točkovna ocena parametra – če je ocena parametra podana z eno samo vrednostjo.
Intervalna ocena parametra – če je ocena parametra podana z intervalom vrednosti.

Interval vrednosti kot oceno populacijskega parametra im. **interval zaupanja** – nakazuje na točnost ocene ter vsebuje informacijo o zanesljivosti te ocene (z določeno verjetnostjo "zaupamo", "smo gotovi" v pravilnost ocene).

Definirajmo ...

Denimo, da s slučajnim vzorcem ocenjujemo populacijski parameter γ. Poskušamo najti vzorčno statistko g, ki je nepristranska cenilka ($E(g) = \gamma$) in se na vseh možnih vzorcih vsaj približno normalno[*] porazdeljuje $g: N(E(g), SE(g))$. Nato poskušamo najti interval, v katerem se bo z dano gotovostjo ($1-\alpha$) nahajal ocenjevani parameter:

$$P(a < \gamma < b) = 1 - \alpha$$

a ... spodnja meja intervala zaupanja
b ... zgornja meja intervala zaupanja
$1-\alpha$ ... verjetnost gotovosti

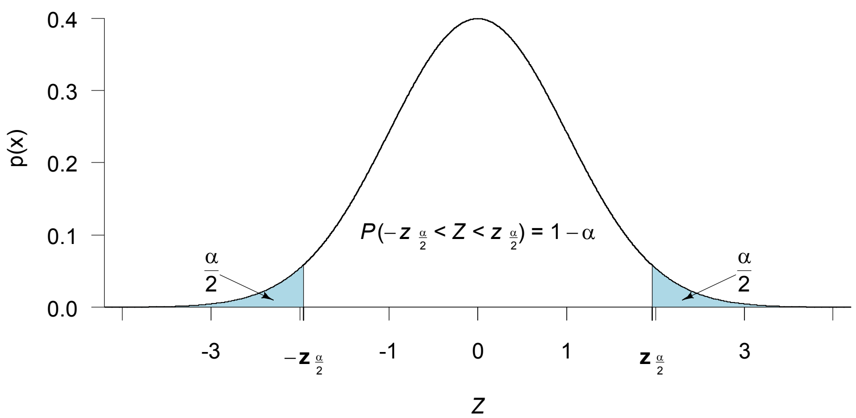
[*] Dejansko ni nujno, da se porazdeljuje po normalni porazdelitvi. Dovolj je že, da to porazdelitev poznamo (kakršna koli pač je). Postopek se ustrezno prilagodi.

Ta interval imenujemo **interval zaupanja**.

Interpretiramo: S tveganjem α (oz. z gotovostjo 1 - α) lahko trdimo, da se populacijski parameter γ nahaja v tem intervalu.

Kako konstruiramo interval zaupanja?

1. Statistiko g lahko glede na predpostavke o njeni porazdelitvi standardiziramo na naslednji način:
$$g \rightarrow Z = \frac{g - E(g)}{SE(g)} = \frac{g - \gamma}{SE(g)}$$
Dobljena slučajna spremenljivka Z se porazdeljuje standardizirano normalno $Z: N(0,1)$.
2. Interval s tveganjem α oz. gotovostjo $(1-\alpha)$ lahko zapišemo za standardizirano spremenljivko Z:
$$P(z_1 < Z < z_2) = 1 - \alpha$$
Ker tveganje α porazdelimo polovico na levo in polovico na desno na konce normalne porazdelitve (velja $H(z_1) = 1-H(z_2) = \alpha/2$), bomo z_1 in z_2 odslej pisali z indeksom $z_{\alpha/2}$. Torej lahko interval zapišemo kot: $P(-z_{\alpha/2} < Z < z_{\alpha/2}) = 1 - \alpha$
Grafično to predstavimo takole:



3. Namesto Z-ja v interval vstavimo obrazec, po katerem smo oblikovali standardizirano spremenljivko Z:
$$P(-z_{\alpha/2} < \frac{g - \gamma}{SE(g)} < z_{\alpha/2}) = 1 - \alpha$$

4. Izraz znotraj oklepaja ustrezno preuredimo:

$$-z_{\alpha/2} < \frac{g-\gamma}{SE(g)} < z_{\alpha/2} \qquad /\cdot SE(g)$$

$$-z_{\alpha/2}SE(g) < g-\gamma < z_{\alpha/2}SE(g) \qquad /-g$$

$$-z_{\alpha/2}SE(g)-g < -\gamma < z_{\alpha/2}SE(g)-g \qquad /\cdot(-1)$$

$$z_{\alpha/2}SE(g)+g > \gamma > -z_{\alpha/2}SE(g)+g$$

$$g-z_{\alpha/2}SE(g) < \gamma < g+z_{\alpha/2}SE(g)$$

in dobimo obrazec za interval zaupanja:

$$P(\underbrace{g - z_{\alpha/2}SE(g)}_{a} < \gamma < \underbrace{g + z_{\alpha/2}SE(g)}_{b}) = 1-\alpha$$

V tem obrazcu je $z_{\alpha/2}$ določen le s stopnjo tveganja α. Vrednosti $z_{\alpha/2}$ lahko razberemo iz tabele za verjetnosti za standardizirano normalno porazdelitev, ker velja $H(-z_{\alpha/2}) = 1-H(z_{\alpha/2}) = \alpha/2$.

$\alpha = 0.10,\ z_{\alpha/2} = 1.65$
$\alpha = 0.05,\ z_{\alpha/2} = 1.96$
$\alpha = 0.01,\ z_{\alpha/2} = 2.58$

Splošen obrazec za interval zaupanja s tveganjem α oz. gotovostjo $(1-\alpha)$:

$\gamma \leftarrow\ g:N(E(g), SE(g))$ oz. $g:N(\gamma, SE(g))$

$P(g - z_{\alpha/2}SE(g) < \gamma < g + z_{\alpha/2}SE(g))=1-\alpha$

4.4.2 Interval zaupanja za populacijsko aritmetično sredino

$$\mu \quad \leftarrow \quad \overline{X}:N(E(\overline{X}),SE(\overline{X})) \quad oz. \quad \overline{X}:N(\mu,\frac{\sigma}{\sqrt{n}})$$

Interval zaupanja za populacijsko aritmetično sredino μ s tveganjem α oz. gotovostjo $(1-\alpha)$:

$$P(\overline{X}-z_{\alpha/2}SE(\overline{X}) < \mu < \overline{X}+z_{\alpha/2}SE(\overline{X})) = 1-\alpha \qquad oz.$$

$$P(\overline{X}-z_{\alpha/2}\frac{\sigma}{\sqrt{n}} < \mu < \overline{X}+z_{\alpha/2}\frac{\sigma}{\sqrt{n}}) = 1-\alpha$$

Populacijskega standardnega odklona σ seveda običajno ne poznamo. Če lahko predpostavimo, da se spremenljivka X na populaciji porazdeljuje normalno in če imamo dovolj velik vzorec ($n > 30$), lahko namesto populacijskega standardnega odklona vzamemo oceno standardnega odklona s iz vzorca in interval zaupanja zapišemo:

$$P(\overline{X}-z_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X}+z_{\alpha/2}\frac{s}{\sqrt{n}}) = 1-\alpha$$

pri čemer vzorčni standardni odklon s izračunamo takole:

$$s = \sqrt{s^2} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2}$$

Tako definirana **vzorčna varianca** je namreč nepristranska cenilka populacijske variance ($E(s^2) = \sigma^2$)) in tako zadosti dobra ocena populacijske variance.

Primer: Na vzorcu velikosti n = 151 podjetnikov v majhnih podjetjih v Sloveniji, ki je bil izveden v okviru ankete "Drobno gospodarstvo v Sloveniji " (Prašnikar, 1993), so izračunali, da je povprečna starost anketiranih podjetnikov $\overline{X}$ = 40.4 let in standardni odklon s = 10.2 let. Pri 5% tveganju želimo z intervalom zaupanja oceniti povprečno starost podjetnikov v majhnih podjetjih v Sloveniji.

$$\mu \quad \leftarrow \quad \overline{X} : N(E(\overline{X}), SE(\overline{X})) \quad oz. \quad \overline{X} : N(\mu, \frac{\sigma}{\sqrt{n}})$$

$$P(\overline{X} - z_{\alpha/2} SE(\overline{X}) < \mu < \overline{X} + z_{\alpha/2} SE(\overline{X})) = 1 - \alpha \qquad oz.$$

$$P(\overline{X} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

oz. ker σ ne poznamo:

$$P(\overline{X} - z_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2} \frac{s}{\sqrt{n}}) = 1 - \alpha$$

Iz tabele za standardizirano porazdelitev Z razberemo:

$$a = 0.05 \rightarrow \alpha/2 = 0.025$$

$$H(-z_{\alpha/2}) = 1 - H(z_{\alpha/2}) = 0{,}025 \qquad \text{Iz tabele razberemo}: z_{\alpha/2} = \mp 1.96$$

V obrazec za interval zaupanja vstavimo potrebne podatke in izračunamo:

$$P(\overline{X} - z_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2} \frac{s}{\sqrt{n}}) = 1 - \alpha$$

$$40.4 - 1.96\frac{10.2}{\sqrt{151}} < \mu < 40.4 + 1.96\frac{10.2}{\sqrt{151}}$$

$$40.4 - 1.6 < \mu < 40.4 + 1.6$$

$$38.8 < \mu < 42.0$$

95% interval zaupanja za povprečno starost podjetnikov v majhnih podjetjih v Sloveniji je med 38.8 in 42.0 leti.
Z 95% gotovostjo (oz. 5% tveganjem) ocenjujemo, da je povprečna starost podjetnikov v majhnih podjetjih v Sloveniji med 38.8 in 42.0 leti.

4.4.3 Interval zaupanja za populacijski delež

$$\pi \quad \leftarrow \quad p : N(E(p), SE(p)) \quad oz. \quad p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$$

Interval zaupanja za populacijski delež π s tveganjem α oz. gotovostjo *(1-α)*:

$$P(p - z_{\alpha/2} SE(p) < \pi < p + z_{\alpha/2} SE(p)) = 1-\alpha \qquad oz.$$

$$P(p - z_{\alpha/2}\sqrt{\frac{\pi(1-\pi)}{n}} < \pi < p + z_{\alpha/2}\sqrt{\frac{\pi(1-\pi)}{n}}) = 1-\alpha$$

Seveda populacijskega deleža π ne poznamo, zato v standardni napaki *SE(p)* namesto π upoštevamo njegovo vzorčno oceno *p*. Če imamo dovolj velik vzorec ($n > 20$ pri deležih okoli 0.5, pri bolj "ekstremnih" deležih mora biti vzorec večji), lahko torej interval zaupanja zapišemo:

$$P(p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < \pi < p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}) = 1-\alpha$$

Primer: Na vzorcu (n = 151), ki je bil izveden v okviru ankete "Drobno gospodarstvo v Sloveniji ", so izračunali, da je delež obrtnih podjetij p = 0.50. Pri 10% tveganju želimo z intervalom zaupanja oceniti delež obrtnih majhnih podjetij v Sloveniji.

$$\pi \quad \leftarrow \quad p : N(E(p), SE(p)) \quad oz. \quad p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$$

$$P(p - z_{\alpha/2} SE(p) < \pi < p + z_{\alpha/2} SE(p)) = 1-\alpha \qquad oz.$$

$$P(p - z_{\alpha/2}\sqrt{\frac{\pi(1-\pi)}{n}} < \pi < p + z_{\alpha/2}\sqrt{\frac{\pi(1-\pi)}{n}}) = 1-\alpha$$

oz.:

$$P(p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < \pi < p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}) = 1-\alpha$$

Iz tabele za standardizirano porazdelitev Z razberemo:

$$a = 0.10 \rightarrow \alpha/2 = 0.05$$

$$H(-z_{\alpha/2}) = 1 - H(z_{\alpha/2}) = 0,05 \qquad \text{Iz tabele razberemo}: z_{\alpha/2} = \mp 1.65$$

V obrazec za interval zaupanja vstavimo potrebne podatke in izračunamo:

$$P(p - z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}} < \pi < p + z_{\alpha/2}\sqrt{\frac{p(1-p)}{n}}) = 1-\alpha$$

$$0.50 - 1.65\sqrt{\frac{0.50(1-0.50)}{151}} < \pi < 0.50 + 1.65\sqrt{\frac{0.50(1-0.50)}{151}}$$

$$0.50 - 0.07 < \pi < 0.50 + 0.07$$

$$0.43 < \pi < 0.57$$

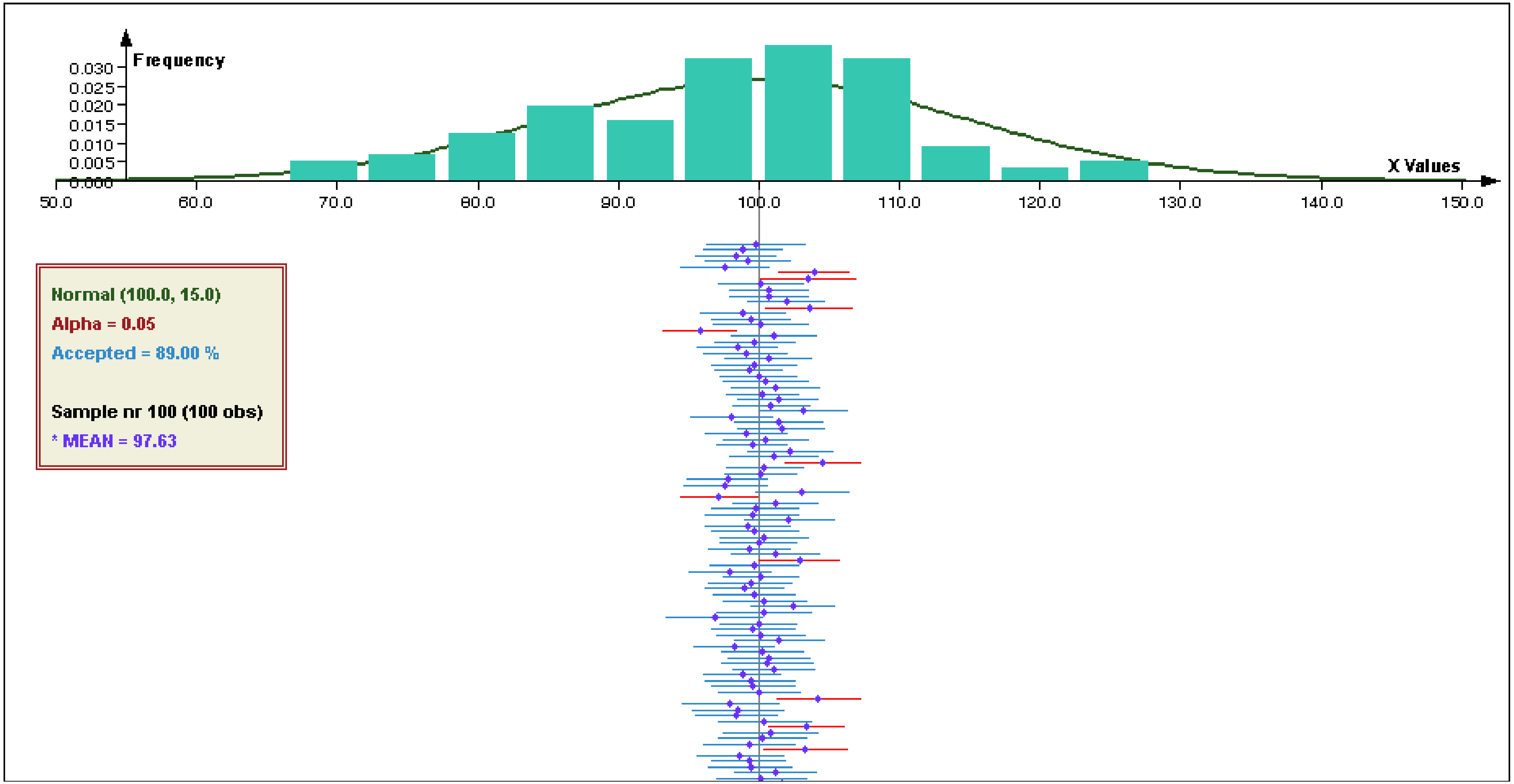
90% interval zaupanja za delež majhnih obrtnih podjetij med vsemi majhnimi podjetji v Sloveniji je med 0.43 in 0.57.
Z 90% gotovostjo (oz. 10% tveganjem) ocenjujemo, da je % majhnih obrtnih podjetij med vsemi podjetji v Sloveniji med 43% in 57%.

4.4.4 Pomen stopnje tveganja pri intervalih zaupanja

Za vsak slučajni vzorec lahko ob omenjenih predpostavkah izračunamo ob izbrani stopnji tveganja α interval zaupanja za parameter γ. Ker se podatki od vzorca do vzorca razlikujejo, se razlikujejo tudi izračunane vzorčne statistike (vzorčne ocene parametrov) in s tem izračunani intervali zaupanja za parameter γ.
Če bi na osnovi vseh možnih slučajnih vzorcev neke velikosti iz neke populacije izračunali intervale zaupanja s stopnjo tveganja $\alpha = 0.05$, bi 5% teh intervalov ne pokrilo iskanega parametra γ.

Primer: Ocenjujemo populacijski parameter aritmetična sredina μ spremenljivke inteligenčni kvocient. Denimo, da v tem primeru poznamo porazdelitev te spremenljivke na populaciji (*X: N(100, 15)*). Recimo, da naredimo 100 slučajnih vzorcev velikosti 100 enot in na osnovi vsakega vzorca izračunamo 95% interval zaupanja za to populacijsko aritmetično sredino. Pričakujemo lahko, da bo približno 95% teh intervalov pravilno pokrilo parameter μ.

S pomočjo računalniške simulacije (http://lstat.kuleuven.be/java/) smo izračunali 100 intervalov zaupanja (pri $\alpha = 0.05$) na osnovi slučajnih vzorcev (velikosti 100 enot) iz populacije s porazdelitvijo *X: N(100, 15)*. V zgornjem delu je prikazana porazdelitev spremenljivke *X* na populaciji (krivulja) ter porazdelitev vrednosti v enem izmed slučajnih vzorcev (histogram – stolpci). V spodnjem delu je prikazanih 100 intervalov zaupanja (pikica = vzorčna aritmetična sredina, daljica = meje intervala). Rdeče obarvani intervali ne pokrijejo (ne vsebujejo) populacijske aritmetične sredine $\mu = 100$. V tem primeru je takih intervalov 11% (v primeru vseh možnih vzorcev bi jih bilo 5%). Matematično upanje 100 vzorčnih aritmetičnih sredin je 97.63 (v primeru vseh možnih vzorcev bi bilo 100).


Frequency
0.030
0.025
0.020
0.015
0.010
0.005
0.000
X Values
50.0
60.0
70.0
80.0
90.0
100.0
110.0
120.0
130.0
140.0
150.0
Normal (100.0, 15.0)
Alpha = 0.05
Accepted = 89.00 %
Sample nr 100 (100 obs)
* MEAN = 97.63

4.4.5 Kako se spreminja interval zaupanja

1. Če spreminjamo stopnjo tveganja?

   Večji $\alpha \Rightarrow$ ožji interval

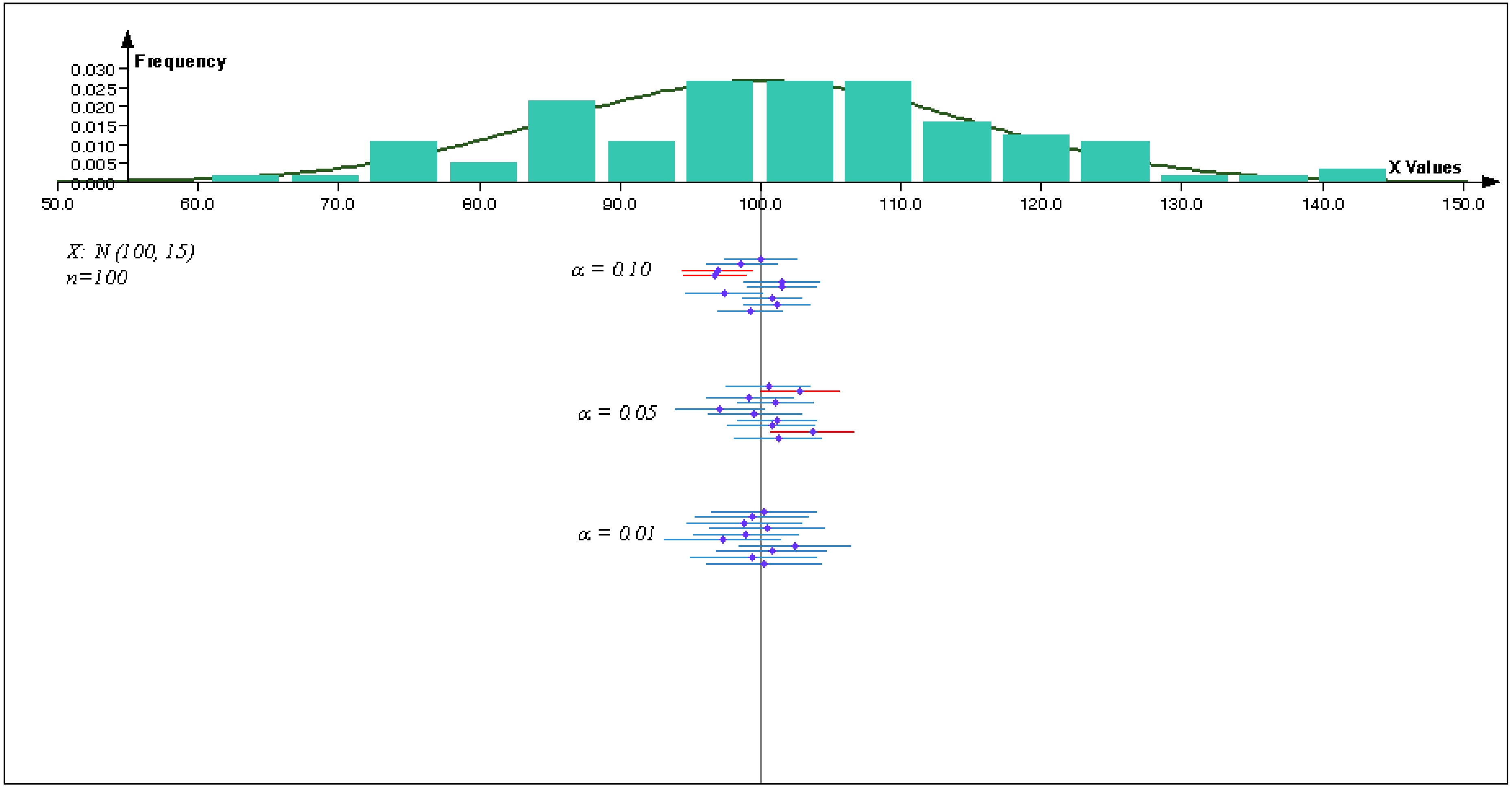
   Večje tveganje oz. manjša gotovost prinaša bolj natančno oceno.

2. Če spreminjamo velikost vzorca?

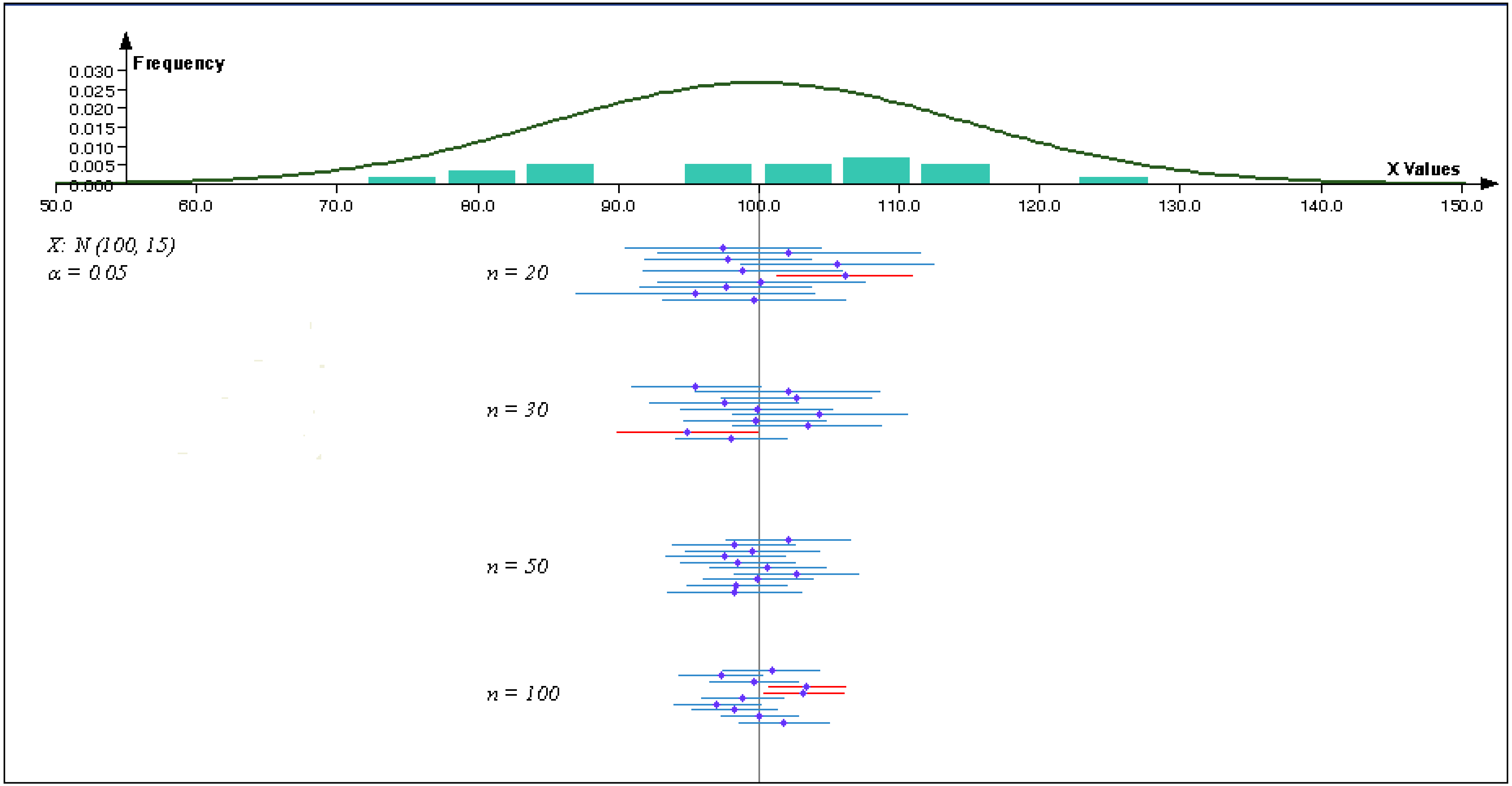
   Večji $n \Rightarrow$ ožji interval, ker se zmanjšuje *SE(g)* (npr., če je g aritmetična sredina velja: $SE(g) = st.odklon/\sqrt{n}$)

   Večji vzorec prinaša bolj natančno oceno.

Primer: Spreminjanje velikosti intervala zaupanja ob spreminjanju stopnje tveganja
(Računalniške simulacije, http://lstat.kuleuven.be/java/)



Primer: Spreminjanje velikosti intervala zaupanja ob spreminjanju velikosti vzorca
(Računalniške simulacije, http://lstat.kuleuven.be/java/)



4.5 PREVERJANJE DOMNEV

4.5.1 Uvod

Preverjanje domnev kot znanstvena metoda:

- opazovanje pojava,
- postavljanje vprašanj,
- formuliranje hipotez,
- zbiranje empiričnih podatkov (eksperiment, anketa itd.),
- sprejemanje oz. zavračanje hipotez in posledično, razvijanje teorij in zakonitosti.

Primer domneve o populaciji: Ženske so sposobne opravljati več nalog istočasno kot moški. (Ženske so bolj vsestranske.)

Reševanje problema: V eksperiment vključimo slučajni vzorec moških in slučajni vzorec žensk. V eksperimentu morajo te osebe opravili več enostavnih nalog istočasno. Na osnovi dobljenih rezultatov preverjamo domnevo: povprečno število nalog, ki so jih sposobne opravljati ženske istočasno, je večje od povprečnega števila nalog za moške:

$$H: (\mu_{Ž} - \mu_{M})>0$$

- **Znanstvena domneva (hipoteza)** = trditev o nekem dejstvu, ki zahteva potrditev.
- **Statistična domneva (hipoteza)** = domneva o enem ali več populacijskih parametrih, o populacijski porazdelitvi, o povezanosti dveh spremenljivk na populaciji, o odvisnosti dveh spremenljivk na populaciji itd., ki zahteva potrditev.

Statistične domneve običajno ne moremo preverjati na celotni populaciji, pač pa jo preverjamo na osnovi podatkov iz vzorca. Z metodami statističnega sklepanja preverjamo, ali zakonitost, ugotovljena na vzorcu, velja tudi za populacijo. Preverjamo, kako verjetno je, da se je zakonitost, ugotovljena na vzorcu, zgodila zgolj slučajno. Če je to zelo malo verjetno, predvidevamo da se je zgodila zato, ker velja ta zakonitost tudi na populaciji. To naredimo tako, da primerjamo oceno iz vzorca z domnevno populacijsko vrednostjo.

Preverjanje domnev im. tudi test (statistične) značilnosti (angl. *test of significance*).

V okviru našega predmeta bomo obravnavali naslednje vrste domnev:

- domneve o populacijskih parametrih,
- domneve o povezanosti dveh spremenljivk na populaciji,
- domneve o odvisnosti dveh spremenljivk na populaciji.

Nekaj osnovnih pojmov:

- ničelna, alternativna (osnovna) domneva
- statistična značilnost (signifikanca)
- *p* vrednost
- območje zavračanja (kritično območje), območje sprejemanja domneve
- testna statistika
- napaka I., II. vrste
- enostranski, dvostranski test

4.5.2 Način razmišljanja pri preverjanju domnev o populacijskih parametrih

1. Postavimo domnevo o vrednosti parametra, npr. π - delež enot z določeno lastnostjo na populaciji. Denimo, da je domneva

$$H: \pi = \pi_H = 0.36$$

Ob predpostavki, da je domneva pravilna, vemo, da se vzorčni deleži, izračunani na vseh možnih slučajnih vzorcih velikosti *n,* porazdeljujejo približno normalno

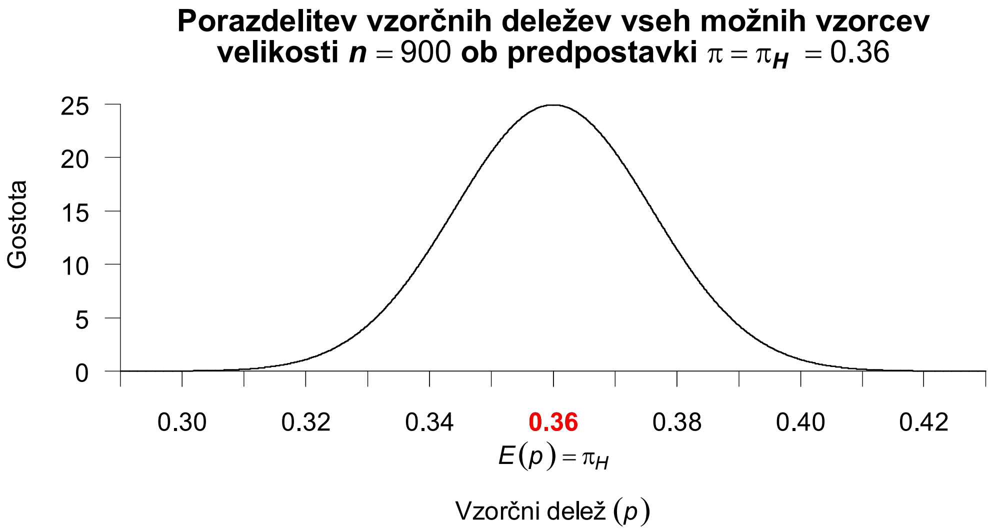
$$p : N(E(p), SE(p))$$

$$p : N(\pi_H, \sqrt{\frac{\pi_H(1-\pi_H)}{n}})$$

Če delamo npr. slučajne vzorce velikosti n = 900 in na vsakem vzorcu določimo vzorčni delež p, se ti vzorčni deleži porazdeljujejo na naslednji način:

$$p : N(0.36, \sqrt{\frac{0.36(1-0.36)}{900}})$$

$$p : N(0.36,\ 0.016)$$

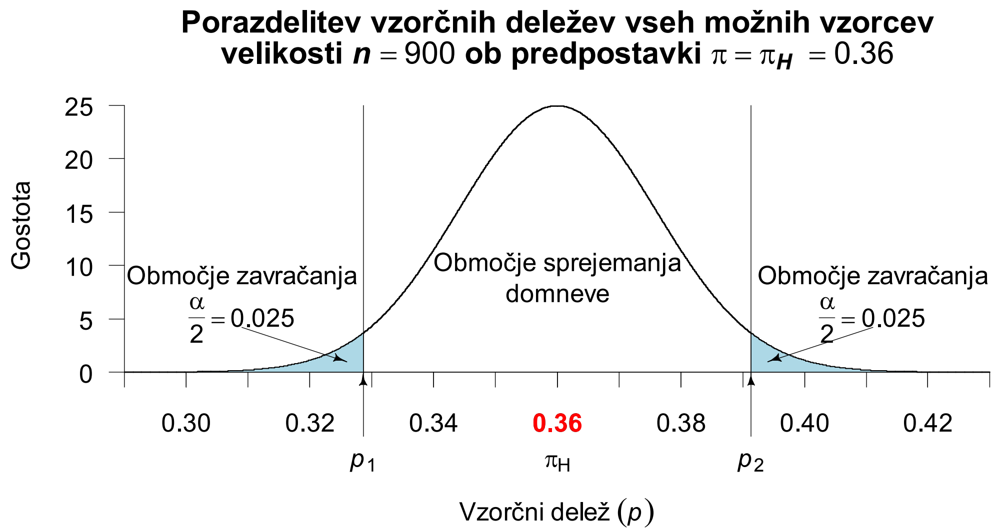


2. Vzemimo en slučajni vzorec velikosti n = 900 z vzorčnim deležem p. Ta se lahko bolj ali manj razlikuje od π_H. Če se zelo razlikuje, lahko podvomimo o resničnosti domneve o π_H.

   Npr. na nekem vzorcu določimo delež p = 0.35. Iz (hipotetične) porazdelitve vidimo, da je vzorcev, ki imajo tak vzorčni delež, kar nekaj. Torej bi bila domneva lahko resnična.
   Npr. na nekem vzorcu določimo delež p = 0.31. Vzorcev s takšnim deležem v zgornji porazdelitvi pa je zelo malo. Zato podvomimo o resničnosti domneve.

3. Kako določimo mejo – kdaj podvomimo o resničnosti domneve?

   Okoli π_H določimo območje sprejemanja domneve in izven tega območja območje zavračanja domneve (im. tudi kritično območje). Običajno je območje zavračanja določeno s 5% vzorcev, ki imajo ekstremne vrednosti deležev (2.5% levo in 2.5% desno), torej imajo zelo malo verjetne vzorčne deleže. (Včasih je takšno območje določeno tudi z 10% ali 1% vzorcev z ekstremnimi vrednostmi deležev.)



V našem primeru je tàko območje sprejemanja domneve pri vzorčnih deležih p med 0.33 in 0.39, območje zavračanja domneve pa pri vzorčnih deležih $p < 0.33$ in $p > 0.39$.

Kako smo določili meji 0.33 in 0.39?

$$P(p_1 < p < p_2) = 0.95$$
$$P(z_1 < Z < z_2) = 0.95$$
$$H(z_1) = (1-0.95)/2 = 0.256 \quad \rightarrow z_1 = -1.96$$
$$H(z_2) = 0{,}95 + 0{,}025 = 0.925 \quad \rightarrow z_2 = 1.96$$

$$z_i = \frac{p - E(p)}{SE(p)} = \frac{p_i - \pi_H}{\sqrt{\frac{\pi_H(1-\pi_H)}{n}}} \quad \Rightarrow \quad p_i = \pi_H + z_i\sqrt{\frac{\pi_H(1-\pi_H)}{n}}$$
$$p_1 = 0.36 - 1.96 \cdot 0.016 = 0.36 - 0.03 = 0.33$$
$$p_2 = 0.36 + 1.96 \cdot 0.016 = 0.36 + 0.03 = 0.39$$

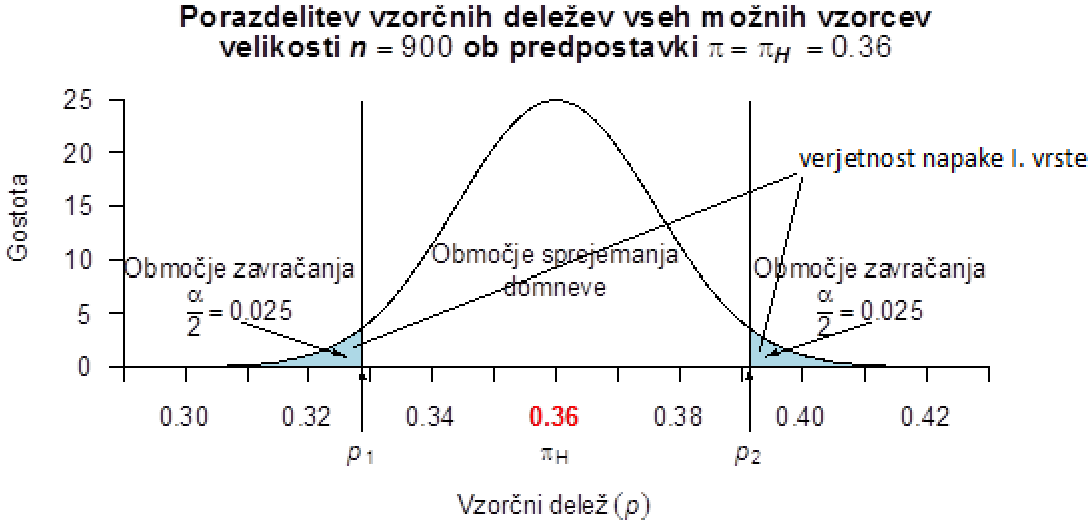
Če torej na vzorcu dobimo npr. delež p, ki je manjši od 0.33, bi bilo to pri domnevi, da je populacijski delež $\pi_H = 0.36$, tako malo verjetno, da podvomimo v pravilnost domneve.

4.5.3 Napaki I. in II. vrste

Sprejemanje ali zavračanje domnev po opisanem postopku je lahko napačno v dveh smislih:

1. Napaka I. vrste (α):

Če vzorčna vrednost deleža p pade v območje zavračanja, domnevo π_H zavrnemo. Pri tem pa vemo, da ob resnični domnevi π_H obstajajo vzorci, ki imajo vrednosti v območju zavračanja. Takih vzorcev je $\alpha\%$. α je verjetnost, da vzorčna vrednost pade v območje zavračanja ob predpostavki, da je domneva resnična. Zato je α verjetnost, da **zavrnemo pravilno domnevo**. To verjetnost imenujemo napaka I. vrste. Ta napaka je merljiva in jo lahko poljubno manjšamo.



2. Napaka II. vrste (β):

Vzorčna vrednost p lahko pade v območje sprejemanja domneve π_H, čeprav je ta napačna. Verjetnost, da se to zgodi, označimo z β. Gre za verjetnost, da **sprejmemo napačno domnevo** in jo imenujemo napaka II. vrste.

V primeru, ki ga obravnavamo, naj bo prava vrednost deleža na populaciji $\pi = 0.40$. Tedaj je porazdelitev vzorčnih deležev dejansko

$$p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$$

$$p : N(0.40, \sqrt{\frac{0.40(1-0.40)}{900}})$$

$$p : N(0.40,\ 0.0163)$$

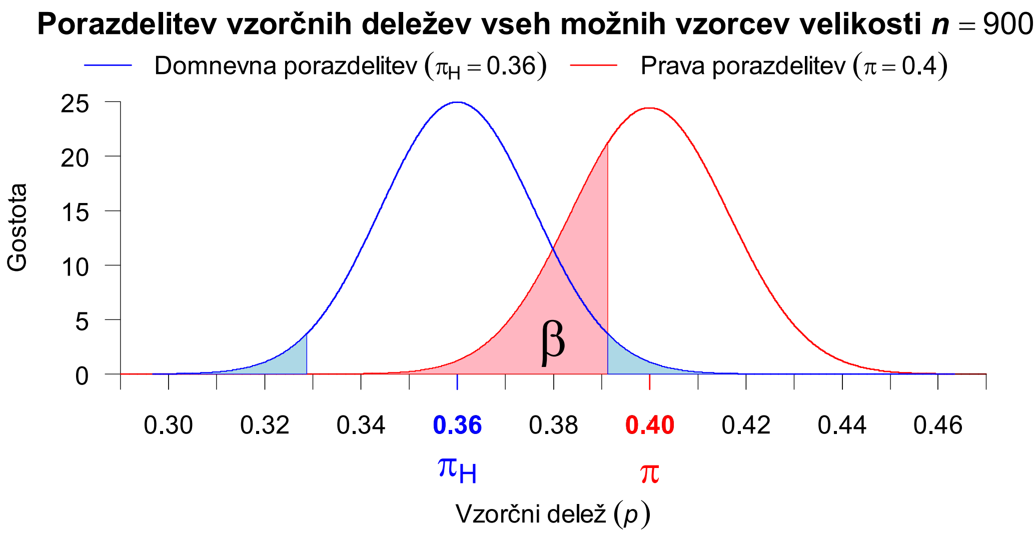
(in ne *p: N (0.36, 0.016)).*

Ker je območje sprejemanja domneve v intervalu $0.33 < p < 0.39$, lahko izračunamo verjetnost β, da bomo sprejeli napačno domnevo, tako da poiščemo verjetnost, da vzorčni deleži iz prave porazdelitve vzorčnih deležev zavzamejo vrednosti med 0.33 in 0.39.

$$\beta = P(0.33 < p < 0.39) = P(\frac{0.33 - E(p)}{SE(p)} < Z < \frac{0.39 - E(p)}{SE(p)}) =$$

$$== P(\frac{0.33 - 0.40}{0.0163} < Z < \frac{0.39 - 0.40}{0.0163}) = P(-4.29 < Z < -0.61) =$$

$$= H(-0.61) = 0.2709$$



Napako II. vrste lahko izračunamo le, če imamo znano resnično vrednost parametra π. Ker pa ga ponavadi ne poznamo, tudi ne poznamo verjetnosti napake II. vrste. Zato je sprejemanje domnev problematično.

V statistiki je bil oblikovan poseben postopek preverjanja domnev, v katerem domneve, ki jo neposredno preverjamo (ničelne domneve), nikoli ne sprejemamo, pač pa jo le zavrnemo ali pa zaključimo, da nimamo dokazov, da bi jo zavrnili.

4.5.4 Postopek preverjanja domnev o populacijskih parametrih

1. Ničelna in alternativna domneva

Postavimo dve nasprotni si domnevi o parametru, ki ju imenujemo **ničelna** in **alternativna** (osnovna) domneva (angl. *null* in *alternative hypothesis*).
Ker zaradi napake II. vrste domnev ne moremo sprejemati, si postavimo ničelno domnevo, ki jo neposredno preverjamo, ter njej nasprotno alternativno domnevo, ki je ne preverjamo neposredno. Če v postopku preverjanja domnev ničelno domnevo lahko zavrnemo, to pomeni, da ostaja alternativna domneva kot pravilna domneva.

Običajno je ničelna domneva tista, ki jo želimo zavrniti, ker bi radi dokazali pravilnost alternativne domneve. Od tod tudi poimenovanje alternativne domneve osnovna domneva.

- Ničelna domneva je vedno: H_0: $\gamma = \gamma_H$

- Alternativna domneva pa je lahko:

H_1: $\gamma \neq \gamma_H$	dvostranski test
H_1: $\gamma > \gamma_H$	enostranski test
H_1: $\gamma < \gamma_H$	enostranski test

V nadaljevanju **preverjamo ničelno domnevo** (s ciljem, da bi jo zavrnili).

Postavimo torej ničelno domnevo o populacijskem parametru, za katero predpostavljamo, da je pravilna. Če na slučajnem vzorcu dobimo rezultat, ki se znatno razlikuje od rezultata, ki bi ga pričakovali pri taki domnevi, potem rečemo, da je opažena razlika **statistično značilna** (angl. *statistically significant*) in zavrnemo ničelno domnevo. Kako določimo, kdaj je razlika statistično značilna?

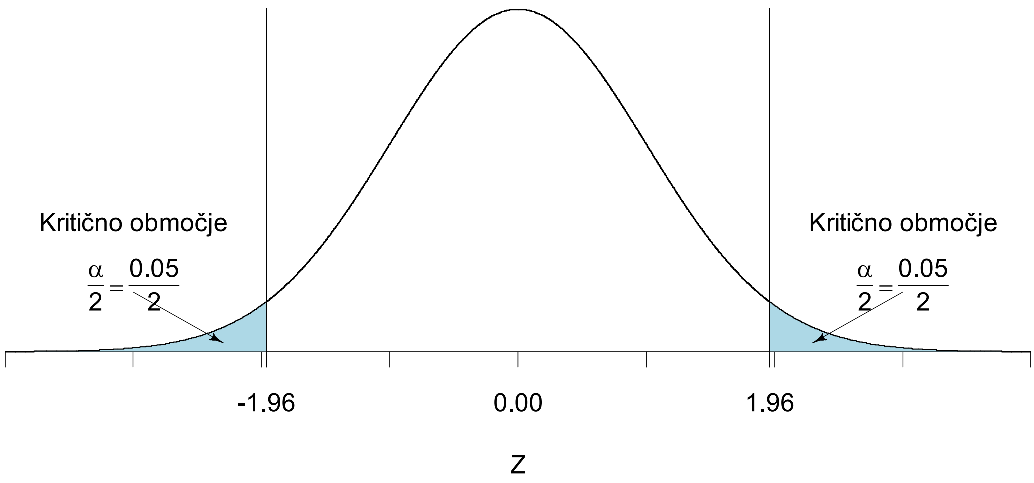
Odločamo se na osnovi porazdelitve vzorčne statistike, ki je kar se da dobra cenilka (npr. nepristranska) populacijskega parametra. To porazdelitev standardiziramo in zanjo določimo območje zavračanja oz. kritično območje. Če standardizirana vrednost vzorčne statistike pade v kritično območje, rečemo, da je razlika med vzorčno oceno in domnevno vrednostjo parametra statistično značilna in ničelno domnevo zavrnemo.

2. Stopnja značilnosti in kritično območje

Odločimo se za maksimalno verjetnost α, ob kateri smo pripravljeni tvegati napako I. vrste (napako, da zavrnemo pravilno domnevo). To verjetnost imenujemo **stopnja značilnosti** (angl. *level of significance*). Običajno se odločimo za 5% stopnjo značilnosti ($\alpha = 0.05$), lahko pa tudi 1% ($\alpha = 0.01$) ali 10% ($\alpha = 0.10$). Iz tabele razberemo vrednost ustrezno standardizirane spremenljivke ob izbranem α in s tem **določimo kritično območje**, torej kdaj bomo zavrnili ničelno domnevo. Pri tem upoštevamo, ali delamo dvostranski ali enostranski test.

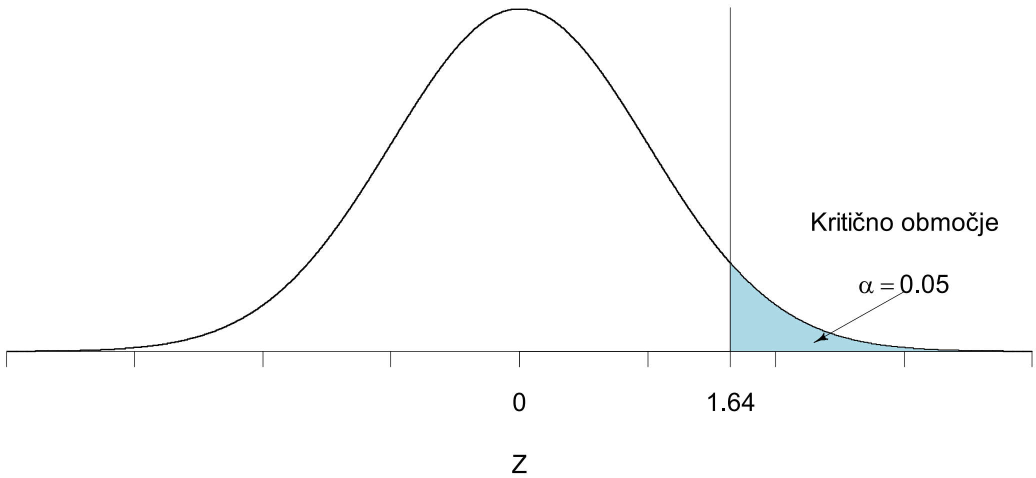
Dvostranski test: H_1: $\gamma \neq \gamma_H$

Iščemo $z_{\alpha/2}$. Velikost posameznega dela kritičnega območja je v tem primeru $\alpha/2$. Kritično območje se nahaja na obeh koncih porazdelitve (vsak del na enem koncu) in ima skupno velikost α.



Enostranski test: H_1: $\gamma > \gamma_H$ ali H_1: $\gamma < \gamma_H$

Iščemo z_α. Velikost kritičnega območja je enako α in se nahaja skrajno levo ($\gamma < \gamma_H$) ali skrajno desno ($\gamma > \gamma_H$) v porazdelitvi.



3. Testna statistika in njena eksperimentalna vrednost

Za parameter poiščemo ustrezno statistiko, ki je kar se da dobra cenilka (npr. nepristranska) parametra, in njeno porazdelitev. To porazdelitev standardiziramo. Standardizirano cenilko im. **testna statistika.**

V naših primerih bomo uporabljali z (kasneje tudi t) testno statistiko**,** odvisno od tega:

- kateri populacijski parameter ocenjujemo,
- kako se porazdeljuje vzorčna statistika, ki je nepristranska cenilka tega parametra.

$$z = \frac{g - E(g)}{SE(g)} = \frac{g - \gamma_H}{SE(g)} \qquad\qquad t = \frac{g - E(g)}{SE(g)} = \frac{g - \gamma_H}{SE(g)}$$

V izrazu nastopa γ_H, ki je hipotetična vrednost populacijskega parametra ob predpostavki, da je pravilna ničelna domneva.

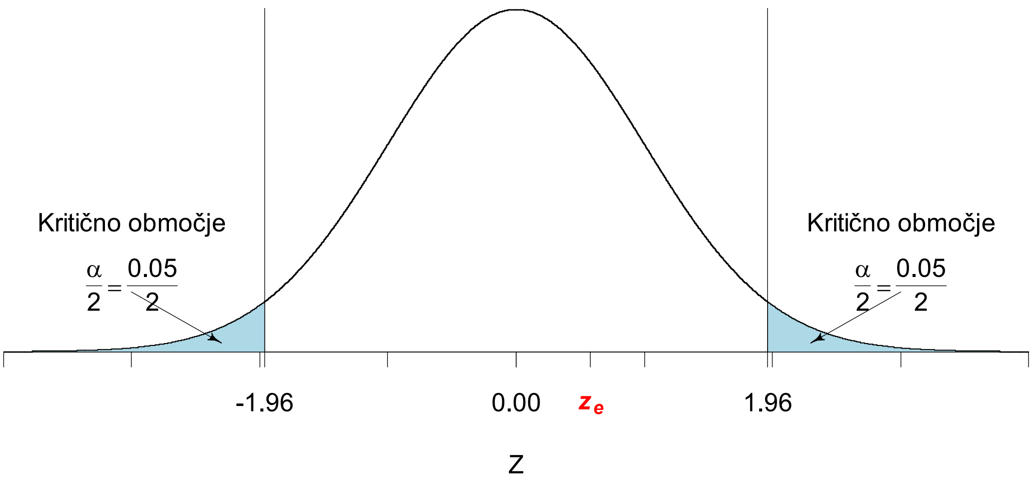
Na osnovi podatkov iz vzorca izračunamo **eksperimentalno vrednost testne statistike** (z_e ali t_e), ki je standardizirana vrednost vzorčne ocene.

4. Sklep

Primerjamo eksperimentalno vrednost testne statistike z vrednostjo testne statistike pri izbrani stopnji značilnosti.

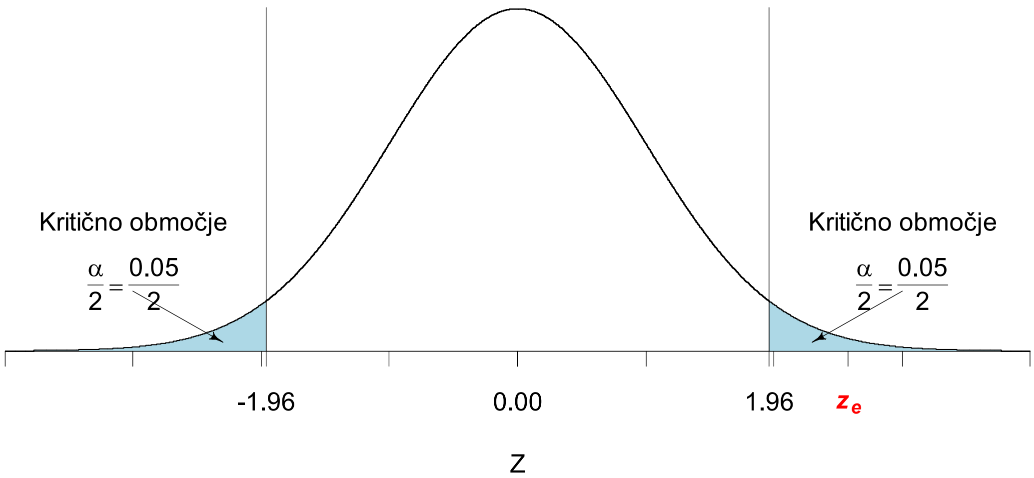
- ***H_0* ne zavrnemo**, če je
  - **enostranski test:** $|z_e| < |z_\alpha|$
  - **dvostranski test:** $|z_e| < |z_{\alpha/2}|$.

  Če eksperimentalna vrednost testne statistike na pade v kritično območje, ničelne domneve ne zavrnemo. Rečemo, da vzorčni podatki kažejo na statistično neznačilne razlike med (hipotetičnim) parametrom (γ_H) in vzorčno oceno (g) pri stopnji značilnosti α.



- ***H_0* zavrnemo**, če
  - **enostranski test:** $|z_e| \geq |z_\alpha|$,
  - **dvostranski test:** $|z_e| \geq |z_{\alpha/2}|$.

  Če eksperimentalna vrednost testne statistike pade v kritično območje, ničelno domnevo zavrnemo in kot pravilno sprejmemo alternativno domnevo. Rečemo, da vzorčni podatki kažejo na statistično značilne razlike med (hipotetičnim) parametrom (γ_H) in vzorčno oceno (g) pri stopnji značilnosti α.



V postopku preverjanja domnev torej najprej postavimo ničelno domnevo o populacijskem parametru in to domnevo preverjamo. Postavimo tudi njej nasprotno alternativno (osnovno) domnevo. Nato vrednost vzorčne statistike, ki jo dobimo na slučajnem vzorcu, primerjamo z rezultatom, ki bi ga pričakovali pri pravilni ničelni domnevi. Zanima nas, ali je razlika med njima statistično značilna (angl. *statistically significant*). O tem, ali je razlika statistično značilna, se odločamo na osnovi porazdelitve vzorčne statistike, ki je kar se da dobra cenilka (npr. nepristranska) populacijskega parametra. To porazdelitev standardiziramo in zanjo - na podlagi izbrane stopnje značilnosti - določimo območje zavračanja oz. kritično območje. Če na slučajnem vzorcu dobimo rezultat, ki se znatno razlikuje od rezultata, ki bi ga pričakovali pri pravilni ničelni domnevi, potem rečemo, da je opažena razlika statistično značilna in zavrnemo ničelno domnevo ter sprejmemo alternativno (osnovno) domnevo.

Ponovimo:

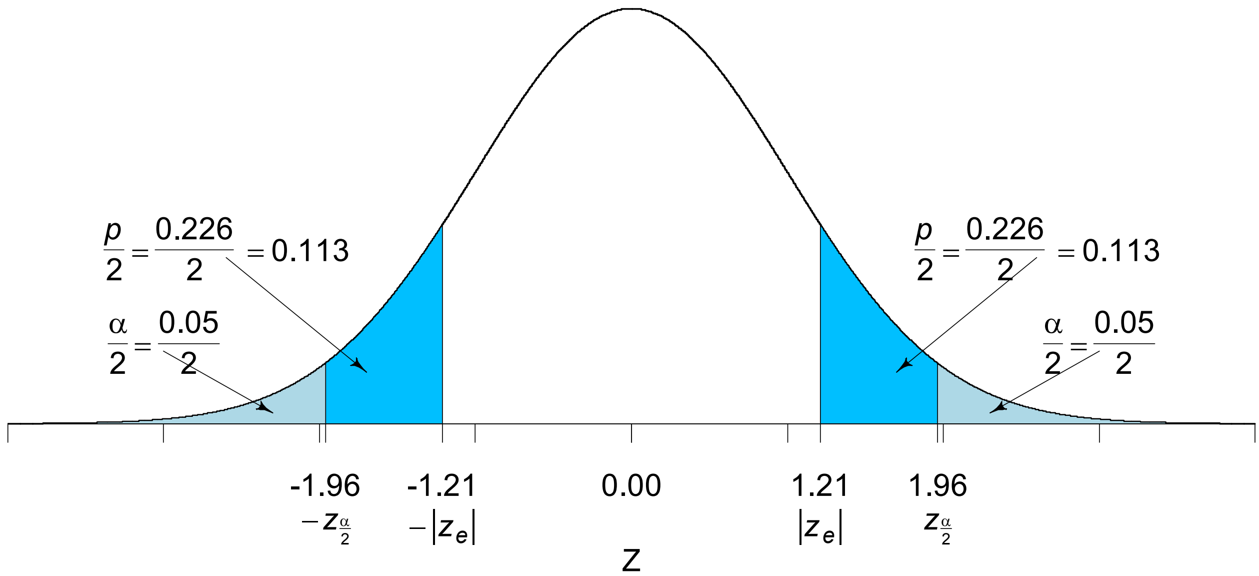
1. Postavimo ničelno in njen nasprotno alternativno domnevo.
2. Poiščemo ustrezno testno statistiko in njeno porazdelitev.
3. Izračunamo eksperimentalno vrednost testne statistike (standardizirana vrednost) na osnovi podatkov iz vzorca.
4. Na podlagi izbrane stopnje značilnosti določimo območje zavračanja ničelne domneve (kritično območje).

5. Če eksperimentalna vrednost testne statistike pade v kritično območje, ničelno domnevo zavrnemo in sprejmemo osnovno domnevo. Če eksperimentalna vrednost testne statistike ne pade v kritično , rečemo, da pri izbrani stopnji značilnosti ničelne domneve ne moremo zavrniti.

4.5.5 Izpis programov za statistično analizo podatkov

Programi za statistično analizo podatkov (npr. SPSS, Excel) izpišejo eksperimentalno vrednost testne statistike in podajo t.im. ***p*** **vrednost** (angl. *P-value)*, t.j. verjetnost, da ima standardizirana spremenljivka bolj ekstremno vrednost od eksperimentalne vrednosti.

Pozor: Programi običajno (pri testih, kjer je to smiselno) vrnejo dvostransko *p*-vrednost (kot na sliki). Če želimo dobiti enostransko (iz dvostranske), jo enostavno delimo z 2. Seveda moramo biti pozorni, če je z_e na pravi strani.



Gre torej za verjetnost, da ob pravilni ničelni domnevi dobimo eksperimentalno vrednost testne statistike, ki je tako ekstremna ali bolj ekstremna kot vrednost, ki smo jo dejansko dobili na vzorcu. Manjša kot je ta verjetnost (manjši kot je *p*), z večjo gotovostjo lahko zavrnemo ničelno domnevo, saj pomeni, da bi bila takšna vrednost, kot smo jo dobili na vzorcu, zelo malo verjetna ob pravilni ničelni domnevi.

p vrednost lahko interpretiramo tudi kot verjetnost napake I. vrste, t.j. napake, da zavrnemo pravilno domnevo. Če je verjetnost te napake velika, potem ne bomo zavrnili ničelne domneve. Če pa je verjetnost te napake majhna, bomo zavrnili ničelno domnevo.

Običajno pravilo:

- Če je p (*sign.*) ≤ 0.05, ničelno domnevo zavrnemo.
- Če je p (*sign.*) > 0.05, ničelne domneve ne zavrnemo.

4.5.6 Primeri

Primer 1: Domneva o populacijski aritmetični sredini

V anketi SJM 2002/1 (n = 1115) je bilo povprečno število članov v slovenskih gospodinjstvih 3.27 in varianca 2.032. Ali lahko ob 10% stopnji značilnosti trdimo, da je povprečno število članov gospodinjstva v Sloveniji statistično značilno različno od 3?

- Spremenljivka X - število članov gospodinjstva
- Populacijski parameter, ki ga ocenjujemo: populacijska aritmetična sredina μ (povprečno število članov gospodinjstva)
- Podatki iz vzorca: $n = 1115$, $\overline{X} = 3{,}27$, $s^2 = 2.032$, $s = 1.425$

1. Domnevi:

H_0: $\mu = 3$ (v obliki $\gamma = \gamma_H$)

H_1: $\mu \neq 3$ (v obliki $\gamma \neq \gamma_H$) $\rightarrow$ dvostranski test (Ker sprašujemo, ali je povprečno število članov *različno* od 3.)

2. Testna statistika in njena eksperimentalna vrednost

Za cenilko populacijske aritmetične sredine (parameter) vzamemo vzorčno aritmetično sredino (statistika), ki se porazdeljuje:

$$\overline{X} : N(E(\overline{X}), SE(\overline{X})) \quad \text{oz.} \quad \overline{X} : N(\mu, \frac{\sigma}{\sqrt{n}})$$

Izračunajmo hipotetično porazdelitev ob predpostavki, da velja ničelna domneva. Pri tem populacijski standardni odklon, ki ga ne poznamo, ocenimo z vzorčnim standardnim odklonom.

$$\overline{X} : N(\mu_H, \frac{s}{\sqrt{n}}) \qquad \overline{X} : N(3, \frac{1.425}{\sqrt{1115}}) \qquad \overline{X} : N(3, 0.0427)$$

Testna statistika je v tem primeru standardizirana vzorčna aritmetična sredina:

$$z = \frac{\overline{X} - E(\overline{X})}{SE(\overline{X})} = \frac{\overline{X} - \mu_H}{\frac{s}{\sqrt{n}}}$$

Izračunamo eksperimentalno vrednost testne statistike:

$$z_e = \frac{\overline{X} - E(\overline{X})}{SE(\overline{X})} = \frac{\overline{X} - \mu_H}{\frac{s}{\sqrt{n}}} = \frac{3.27 - 3}{0.0427} = 6.32$$

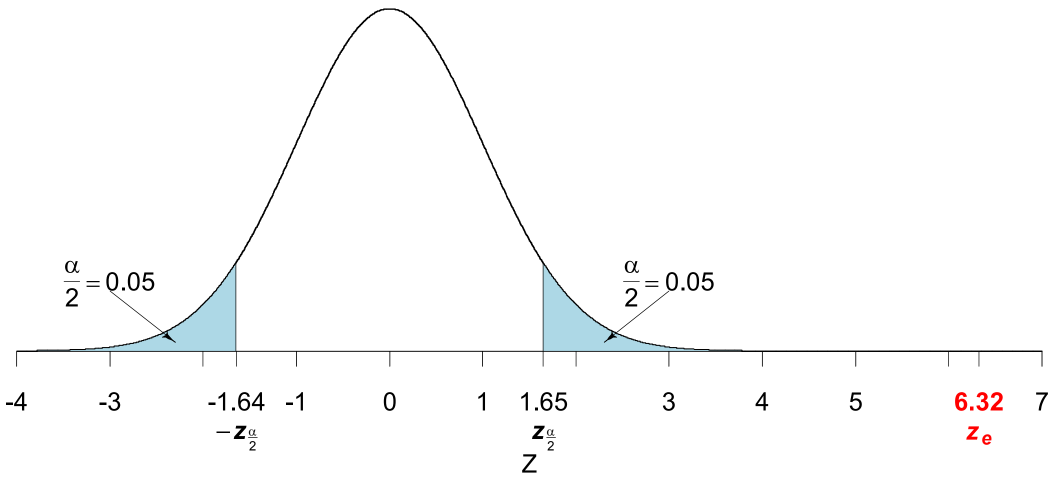
3. Stopnja značilnosti α in kritično območje

$\alpha = 0.10 \rightarrow z_{\alpha/2} = \pm 1.65$

4. Sklep

$6.32 > 1.65$ oz. $|z_e| > |z_{\alpha/2}|$

Eksperimentalna vrednost testne statistike pade v kritično območje. $\rightarrow$ H_0 zavrnemo, sprejmemo H_1.



(Izpis iz stat. programa: sign =.000, kar beremo kot "p vrednost (dvostranska) je manjša od 0.0005" in pomeni $P(|Z| > 6.32) < 0.0005$. Torej je zagotovo manjša od 0.10 in ničelno domnevo lahko zavrnemo.)

Interpretacija

Ničelno domnevo zavrnemo ob 10% stopnji značilnosti. Povprečno število članov gospodinjstva na populaciji je statistično značilno različno od 3 pri 10% stopnji značilnosti.

Pomen stopnje značilnosti

V tem primeru smo 90% gotovi, da je sprejeta odločitev o zavrnitvi ničelne domneve pravilna. Oziroma, obstaja 10% možnosti, da smo zavrnili pravilno domnevo (napaka I. vrste). Ko rečemo, da domnevo zavrnemo ob 10% stopnji značilnosti, pomeni, da obstaja 10% možnosti, da se motimo. Vendar, ker smo ugotovili, da je vrednost p < 0.0005, je možnost te napake dejansko še manjša.

Primer 2: Domneva o populacijskem deležu

V telefonski anketi (n=510) RIS – Raba interneta v Sloveniji (http://www.ris.org) decembra 2002 so ugotovili, da je delež uporabnikov, ki internet uporabljajo skoraj vsak dan, 49%. S 5% stopnjo značilnosti preverimo domnevo, da v Sloveniji manj kot polovica prebivalcev dnevno uporablja internet.

- Spremenljivka X – dnevna uporaba interneta (da, ne)
- Populacijski parameter, ki ga ocenjujemo: populacijski delež π (delež dnevnih uporabnikov interneta).
- Podatki iz vzorca: $n = 510$, $p = 0.49$

1. Domnevi:

H_0: $\pi = 0.5$ (v obliki $\gamma = \gamma_H$)

H_1: $\pi < 0.5$ (v obliki $\gamma < \gamma_H$) $\rightarrow$ enostranski test (Ker sprašujemo, ali je delež dnevnih uporabnikov interneta *manjši* od 0.5.)

2. Testna statistika in njena eksperimentalna vrednost

Za cenilko populacijskega deleža (parameter) vzamemo vzorčni delež (statistika), ki se porazdeljuje:

$$p : N(E(p), SE(p)) \quad \text{oz.} \quad p : N(\pi, \sqrt{\frac{\pi(1-\pi)}{n}})$$

Izračunamo hipotetično porazdelitev ob predpostavki, da je pravilna ničelna domneva:

$$p : N(\pi_H, \sqrt{\frac{\pi_H(1-\pi_H)}{n}}) \qquad p : N(0.5, \sqrt{\frac{0.5(1-0.5)}{510}}) \qquad p : N(0.5, 0.0221)$$

Testna statistika je v tem primeru standardizirani vzorčni delež:

$$z = \frac{p - E(p)}{SE(p)} = \frac{p - \pi_H}{\sqrt{\frac{\pi_H(1-\pi_H)}{n}}}$$

Izračunamo eksperimentalno vrednost testne statistike:

$$z_e = \frac{p - E(p)}{SE(p)} = \frac{p - \pi_H}{\sqrt{\dfrac{\pi_H(1-\pi_H)}{n}}} = \frac{0.49 - 0.5}{0.0221} = -0.452$$

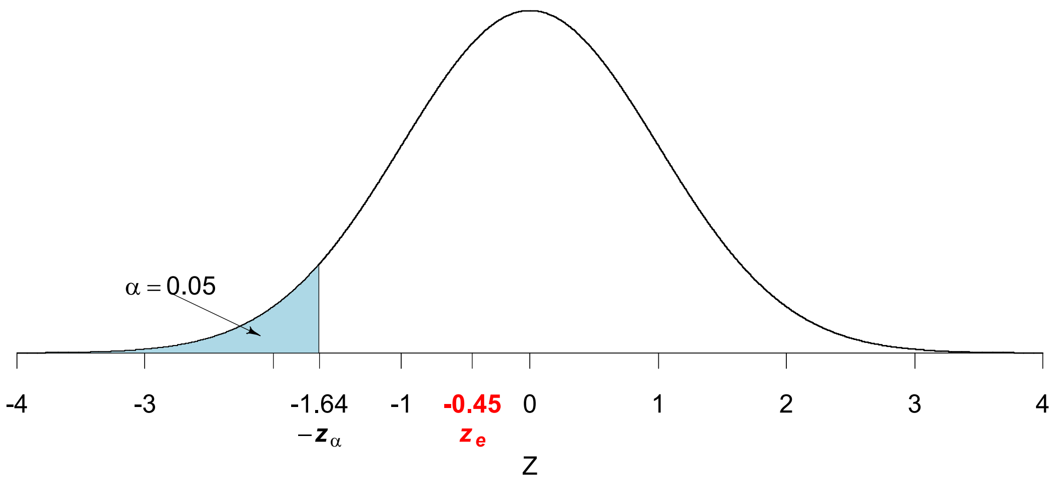
3. Stopnja značilnosti α in kritično območje

$\alpha = 0.05 \;\rightarrow\; z_\alpha = -1.65$

4. Sklep

$|-0.452| < |-1.65|$ oz. $|z_e| < |z_\alpha|$

Eksperimentalna vrednost testne statistike ne pade v kritično območje. → H_0 ne moremo zavrniti.



(Izpis iz stat. programa: sign =.6527, kar beremo kot "*p* vrednost (dvostranska) je 0.6527" in pomeni $P(|Z| > |-0.45|) = 0.6527$. Enostranska p vrednost – tista, ki je ustrezna v tem primeru - pa je 0.6527/2 = 0.3264 in pomeni $P(Z < -0.45) = 0.\ 3264$. Ker ni manjša od 0.05, ničelne domneve ne moremo zavrniti.)

Interpretacija

Ničelne domneve ob 5% stopnji značilnosti ne moremo zavrniti. Delež dnevnih uporabnikov interneta pri 5% stopnji značilnosti ni statistično značilno manjši od 0.5. Nimamo dokazov, da bi rekli, da manj kot polovica prebivalcev dnevno uporablja internet.

Pomen

Vzorčni delež 0.49 ne spada med 5% najmanjših možnih vzorčnih deležev pri populacijskem deležu 0.50 in zato ne moremo podvomiti v pravilnost ničelne domneve. Kar 33% vseh možnih vzorcev (p = 0.3264) ali vsak tretji možni vzorec ima ob populacijskem deležu 0.50 vzorčni delež 0.49 ali manj, zato ni razloga, da bi mislili, da domneva o populacijskem deležu 0.5 ni pravilna.

Primer 3: Domneva o razliki populacijskih aritmetičnih sredin

V neki manjši anketi so anketirance spraševali, koliko se strinjajo s trditvijo "Zakon je zastarela institucija". Odgovarjali so na lestvici od 1 (sploh se ne strinjam) do 10 (zelo se strinjam). Preveriti želimo domnevo, da se moški v povprečju bolj strinjajo s to trditvijo kot ženske. Med anketiranci je bilo 36 moških, ki so imeli na tej lestvici povprečje 7.2, standardni odklon pa 1. Med anketiranci je bilo še 36 žensk, ki so na tej lestvici imele povprečje 7.0 in standardni odklon 1. Domnevo preverimo pri 5% stopnji značilnosti.

- Spremenljivka X – strinjanje s trditvijo "Zakon je zastarela institucija" na lestvici od 1 do 10.
- Populacijski parameter, ki ga ocenjujemo: razlika v aritmetični sredini moških in žensk μ_M - $\mu_Ž$ (razlika v povprečnem strinjanju moških in žensk).
- Podatki iz vzorcev:
  - moški: $n_M = 36$, $\overline{X}_M = 7.2$, $s_M = 1$
  - ženske: $n_Ž = 36$, $\overline{X}_Ž = 7.0$, $s_Ž = 1$

1. Domnevi:

H_0: $\mu_M = \mu_Ž$ *oz.* $\mu_M - \mu_Ž = 0$ (v obliki $\gamma = \gamma_H$)

H_1: $\mu_M > \mu_Ž$ *oz.* $\mu_M - \mu_Ž > 0$ (v obliki $\gamma > \gamma_H$) $\rightarrow$ enostranski test

(Ker sprašujemo, ali se moški v povprečju bolj strinjajo kot ženske, torej, ali je razlika med moškimi in ženskami *večja* od nič.)

2. Testna statistika in njena eksperimentalna vrednost

Za cenilko razlike populacijskih aritmetičnih sredin (parameter) vzamemo razliko vzorčnih aritmetičnih sredin (statistika), ki se porazdeljuje na naslednji način:

$$\overline{X}_M - \overline{X}_Ž : N(E(\overline{X}_M - \overline{X}_Ž), SE(\overline{X}_M - \overline{X}_Ž))$$

$$\overline{X}_M - \overline{X}_Ž : N(\mu_M - \mu_Ž, \sqrt{\frac{\sigma_M^2}{n_M} + \frac{\sigma_Ž^2}{n_Ž}})$$

Izračunamo hipotetično porazdelitev ob predpostavki, da velja ničelna domneva. Pri tem za oceno populacijskih varianc, ki jih ne poznamo, vzamemo vzorčni varianci kot nepristranski cenilki populacijskih varianc:

$$\overline{X}_M - \overline{X}_Ž : N((\mu_M - \mu_Ž)_H, \sqrt{\frac{s_M^2}{n_M} + \frac{s_Ž^2}{n_Ž}})$$

$$\overline{X}_M - \overline{X}_Ž : N(0, \sqrt{\frac{1^2}{36} + \frac{1^2}{36}})$$

$$\overline{X}_M - \overline{X}_Ž : N(0,\ 0.236)$$

Testna statistika je v tem primeru standardizirana razlika vzorčnih aritmetičnih sredin:

$$z = \frac{\overline{X}_M - \overline{X}_Ž - E(\overline{X}_M - \overline{X}_Ž)}{SE(\overline{X}_M - \overline{X}_Ž)} = \frac{\overline{X}_M - \overline{X}_Ž - (\mu_M - \mu_Ž)_H}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_Ž^2}{n_Ž}}}$$

Izračunamo eksperimentalno vrednost testne statistike:

$$z_e = \frac{\overline{X}_M - \overline{X}_Ž - E(\overline{X}_M - \overline{X}_Ž)}{SE(\overline{X}_M - \overline{X}_Ž)} = \frac{\overline{X}_M - \overline{X}_Ž - (\mu_{DM} - \mu_Ž)_H}{\sqrt{\frac{s_M^2}{n_M} + \frac{s_Ž^2}{n_Ž}}} = \frac{7.2 - 7 - 0}{0.236} = 0.847$$

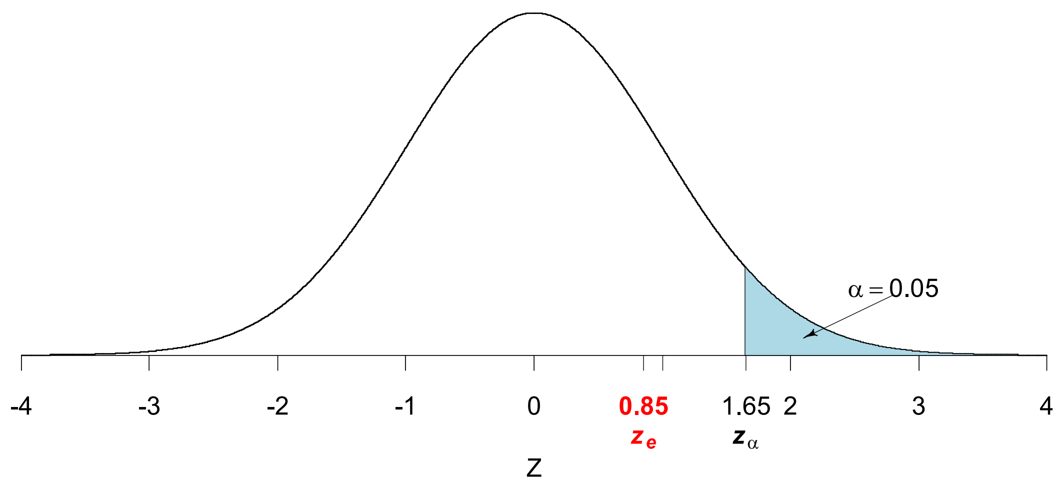
3. Stopnja značilnosti α in kritično območje

$\alpha = 0.05 \ \rightarrow \ z_\alpha = \ 1.65$

4. Sklep

$0.847 < 1.65$ oz. $|z_e| < |z_\alpha|$

Eksperimentalna vrednost testne statistike ne pade v kritično območje. → H_0 ne moremo zavrniti.



(Izpis iz stat. programa: sign =.3953, kar beremo kot "*p* vrednost (dvostranska) je 0.3953" in pomeni $P(|Z| > |0.85|) = 0.3953$. Enostranska p vrednost – ki je ustrezna v tem primeru - je 0.3953/2 = 0.1977 in pomeni $P(Z > 0.85) = 0.1977$. Ker ni manjša od 0.05, ničelne domneve ne moremo zavrniti.)

Interpretacija

Ničelne domneve ob 5% stopnji značilnosti ne moremo zavrniti. Pri 5% stopnji značilnosti ugotavljamo, da se moški v povprečju s trditvijo, da je zakon zastarela institucija, ne strinjajo statistično značilno bolj kot ženske.

Pomen

Razlika vzorčnih aritmetičnih sredin 0.2 ne spada med 5% največjih možnih razlik vzorčnih aritmetičnih sredin pri populacijski razliki 0. Zato ne moremo podvomiti v nepravilnost ničelne domneve. Kar 20% vseh možnih parov vzorcev ($p = 0.1977$) ali vsak peti možni par vzorcev ima ob populacijski razliki 0 razliko vzorčnih aritmetičnih sredin 0.2 ali več, zato ni razloga, da bi mislili, da domneva o razliki populacijskih aritmetičnih sredin 0 ni pravilna.

Primer 4: Domneva o razliki populacijskih deležev

Želimo preveriti, ali je predsedniški kandidat različno priljubljen med mestnimi in vaškimi prebivalci. Zato smo slučajni vzorec prebivalcev povprašali, ali bi glasovali za predsedniškega kandidata. Od 300 slučajno izbranih mestnih prebivalcev (n_1) jih je 90 glasovalo za kandidata. Od 200 slučajno izbranih vaških prebivalcev (n_2) pa je za kandidata glasovalo 50 prebivalcev. Domnevo, da je kandidat različno priljubljen v teh dveh območjih, preverimo pri 10% stopnji značilnosti.

- Spremenljivka X – ali bi glasovali za kandidata (DA, NE)
- Populacijski parameter, ki ga ocenjujemo: razlika v deležu med mestnimi in vaškimi prebivalci $\pi_1 - \pi_2$ (razlika v deležu tistih, ki bi glasovali za kandidata, med mestnimi in med vaškimi prebivalci).
- Podatki iz vzorcev:
    - mestni prebivalci: $n_1 = 300$, $p_1 = 90/300 = 0.30$,
    - vaški prebivalci: $n_2 = 200$, $p_2 = 50/200 = 0.25$

1. Domnevi:

H_0: $\pi_1 = \pi_2$ *oz.* $\pi_1 - \pi_2 = 0$ (v obliki $\gamma = \gamma_H$)

H_1: $\pi_1 \neq \pi_2$ *oz.* $\pi_1 - \pi_2 \neq 0$ (v obliki $\gamma \neq \gamma_H$) $\rightarrow$ dvostranski test (Ker sprašujemo, ali je kandidat *različno* priljubljen.)

2. Testna statistika in njena eksperimentalna vrednost

Za cenilko razlike populacijskih deležev (parameter) vzamemo razliko vzorčnih deležev (statistika), ki se porazdeljuje:

$$p_1 - p_2 : N(E(p_1 - p_2), SE(p_1 - p_2))$$

$$p_1 - p_2 : N(\pi_1 - \pi_2, \sqrt{\frac{\pi_1(1-\pi_1)}{n_1} + \frac{\pi_2(1-\pi_2)}{n_2}})$$

Izračunamo hipotetično porazdelitev ob predpostavki, da velja ničelna domneva. V tem primeru je matematično upanje razlike vzorčnih deležev hipotetična vrednost razlike deležev, ki je v našem primeru enaka 0. Problem pa je, kako oceniti standardni odklon. Ker predpostavljamo, da velja domneva $\pi_1 = \pi_2 = \pi$, je potrebno oceniti populacijski delež π kot uteženo povprečje obeh vzorčnih deležev:

$$p = \frac{n_1 p_1 + n_2 p_2}{n_1 + n_2} = \frac{300 \cdot 0.30 + 200 \cdot 0.25}{300 + 200} = \frac{90 + 50}{300 + 200} = 0.28$$

Izračunajmo torej hipotetično porazdelitev:

$$p_1 - p_2 : N((\pi_1 - \pi_2)_H, \sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}})$$

$$p_1 - p_2 : N(0, \sqrt{\frac{0.28(1-0.28)}{300} + \frac{0.28(1-0.28)}{200}})$$

$$p_1 - p_2 : N(0,\ 0.0410)$$

Testna statistika je v tem primeru standardizirana razlika vzorčnih deležev in njena eksperimentalna vrednost je:

$$z = \frac{p_1 - p_2 - E(p_1 - p_2)}{SE(p_1 - p_2)} = \frac{p_1 - p_2 - (\pi_1 - \pi_2)_H}{\sqrt{\frac{p(1-p)}{n_1} + \frac{p(1-p)}{n_2}}} = \frac{0.30 - 0.25 - 0}{0.0410} = 1.22 = z_e$$

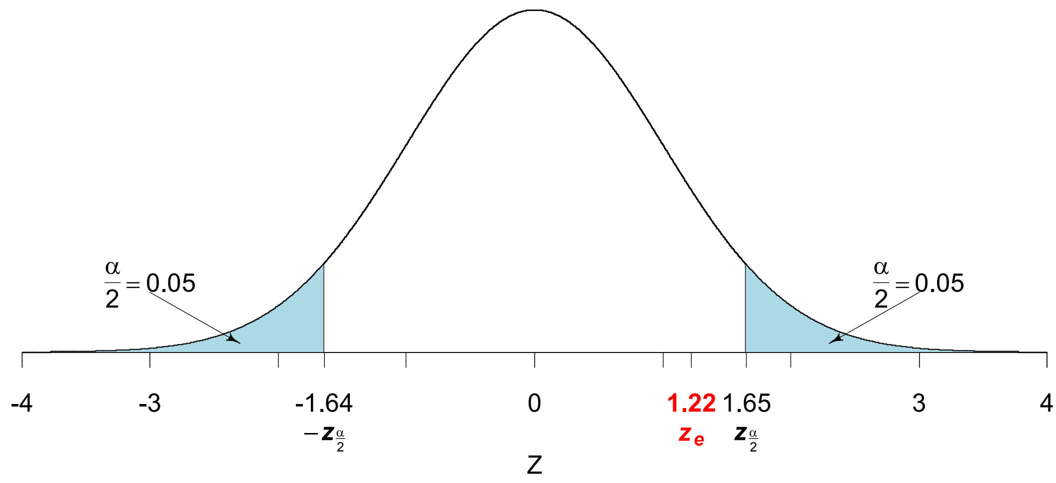
3. Stopnja značilnosti α in kritično območje

$\alpha = 0.10 \rightarrow z_{\alpha/2} = \pm 1.65$

4. Sklep

$1.22 < 1.65$ oz. $|z_e| < |z_{\alpha/2}|$

Eksperimentalna vrednost testne statistike ne pade v kritično območje. → H_0 ne moremo zavrniti.



(Izpis iz stat. programa: sign =.2224, kar beremo kot "*p* vrednost (dvostranska) je manjša od 0.2224 " in pomeni $P(|Z| > 1.22) < 0.2224$. Ker ni manjša od 0.10, ničelne domneve ne moremo zavrniti.)

Interpretacija

Ničelne domneve ob 10% stopnji značilnosti ne moremo zavrniti. Delež mestnih prebivalcev, ki je glasoval za kandidata, pri 10% stopnji značilnosti ni statistično značilno različen od deleža vaških prebivalcev, ki je glasoval za kandidata. Pri 10% stopnji značilnosti ne moremo trditi, da je priljubljenost predsedniškega kandidata statistično značilno različna med mestnimi in vaškimi prebivalci.

Pomen

Razlika vzorčnih deležev 0.05 ne spada med 10% največjih ali najmanjših možnih razlik vzorčnih deležev pri populacijski razliki 0. Zato ne moremo podvomiti v pravilnost ničelne domneve. Kar 22% vseh možnih parov vzorcev (p = 0.2224) ali približno vsak peti možni par vzorcev ima ob populacijski razliki 0 razliko vzorčnih deležev 0.05 ali več (v katerokoli smer), zato ni razloga, da bi mislili, da domneva o razliki populacijskih deležev 0 ni pravilna.

4.6 TEORIJA MALIH VZORCEV

Za velike vzorce (n ≥ 30) smo ugotovili, da se cenilke obravnavanih parametrov (μ, π, μ_1-μ_2, π_1-π_2) porazdeljujejo normalno ali približno normalno, ne glede na obliko porazdelitve na populaciji. Z večanjem vzorca postajajo porazdelitve vedno bolj normalne.

Kadar pa imamo male vzorce (n < 30), porazdelitve vzorčnih statistik niso več približno normalne. V tem primeru moramo uporabiti primerne prilagoditve (razen če je standardna napaka znana). Kadar sklepamo iz vzorca na populacijo na osnovi malih vzorcev, so namreč vzorčne ocene manj zanesljive, saj je variabilnost vzorčnih statistik večja.

Področje statistike, ki se ukvarja s porazdelitvami vzorčnih statistik za male vzorce, se imenuje **teorija malih vzorcev** (angl. *small sampling theory*). Včasih jo imenujemo tudi **natančna teorija vzorcev** (angl. *exact sampling theory*), saj so postopki primerni tako za male kot tudi za velike vzorce.

Za male vzorce bomo obravnavali:

- porazdelitev vzorčnih aritmetičnih sredin $\overline{X}$,
- porazdelitev razlik vzorčnih aritmetičnih sredin $\overline{X}_1 - \overline{X}_2$.

Obe omenjeni vzorčni statistiki se v primeru malih vzorcev porazdeljujeta po **Studentovi *t* porazdelitvi**. Pogledali si bomo, kako v tem primeru izračunamo intervale zaupanja in preverjamo domneve.

Avtor Studentove *t* porazdelitve je *William Sealey Gossett*, ki jo je objavil l. 1908 pod psevdonimom *Student*. Odkril jo je, ko je iskal način ocenjevanja kvalitete piva v podjetju Guiness.

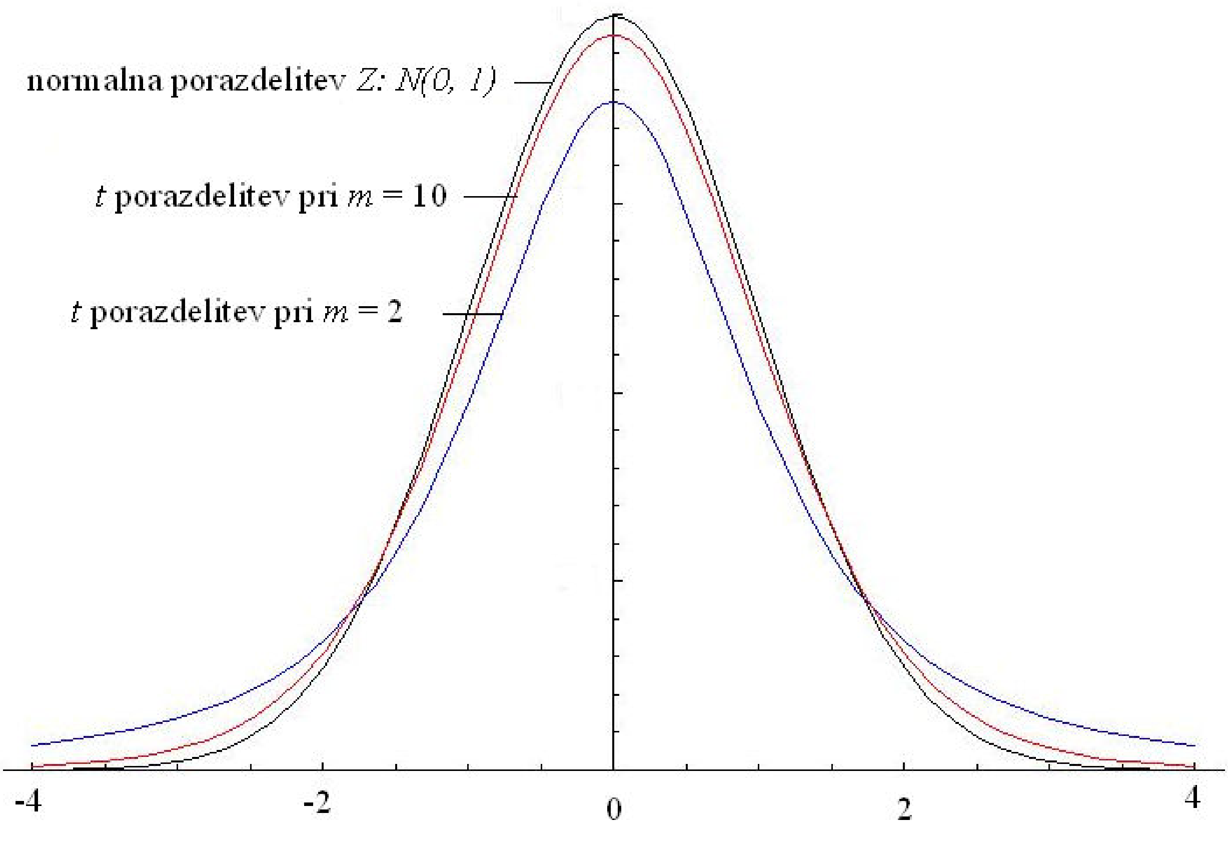
4.6.1 Studentova *t* porazdelitev

Studentova *t* porazdelitev je zelo podobna standardizirani normalni porazdelitvi. Matematično upanje slučajne spremenljivke, ki se porazdeljuje po *t* porazdelitvi, je enako 0 ($E(t) = 0$). Porazdelitev je unimodalna in simetrična. Od normalne porazdelitve se razlikuje v disperziji $D(t)$, saj je njena disperzija večja (ker je variabilnost vzorčnih statistik v primeru malih vzorcev večja).

Studentova t porazdelitev namreč upošteva, da namesto dejanske standardne napake uporabimo vzorčno oceno le-te.

Pri Studentovi t porazdelitvi gre za družino porazdelitev, ki se razlikujejo v disperziji. Disperzija je odvisna od t.im. prostostnih stopenj (angl. *degrees of freedom* ali *df*). Prostostne stopnje, ki jih bomo označevali kot m, so definirane kot velikost vzorca minus število populacijskih parametrov, ki jih moramo oceniti iz vzorca. Torej je disperzija t porazdelitve dejansko odvisna od velikosti vzorca. Večji kot je vzorec, bolj t porazdelitev postaja podobna normalni porazdelitvi Z. Pri prostostnih stopnjah okoli m = 30 sta si porazdelitvi že enaki.

Slika: Studentova t porazdelitev



4.6.2 Porazdelitev vzorčnih aritmetičnih sredin

PONOVIMO ...

Za velike vzorce (n ≥ 30) velja - ne glede na obliko porazdelitve spremenljivke X na populaciji - da se vzorčne aritmetične sredine $\overline{X}$ na vseh možnih vzorcih velikosti n porazdeljujejo normalno na naslednji način: $\overline{X} : N(\mu, \frac{\sigma}{\sqrt{n}})$

V primeru da se spremenljivka X na populaciji porazdeljuje normalno, lahko populacijski standardni odklon σ ocenimo z vzorčnim standardnim odklonom s. V takem primeru lahko:

- izračunamo interval zaupanja za populacijsko aritmetično sredino μ na naslednji način:

$$P(\overline{X} - z_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + z_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha$$

- pri preverjanju domneve o populacijski aritmetični sredini μ uporabimo naslednjo testno statistiko: $z = \frac{\overline{X} - \mu_H}{\frac{s}{\sqrt{n}}}$

ZA MALE VZORCE PA VELJA ...

Za male vzorce (n < 30) velja - ob predpostavki, da se spremenljivka X na populaciji porazdeljuje normalno - da se vzorčne aritmetične sredine $\overline{X}$ na vseh možnih vzorcih velikosti n porazdeljujejo po Studentovi t porazdelitvi z $m = n\text{-}1$ prostostnimi stopnjami.

V tem primeru lahko populacijski standardni σ odklon prav tako cenimo z vzorčnim standardnim odklonom s[1]. Potem lahko:

- izračunamo interval zaupanja za populacijsko aritmetično sredino μ na naslednji način:

$$P(\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha$$

  kjer vrednost $t_{\alpha/2}$ najdemo v ustrezni tabeli.

- pri preverjanju domneve o populacijski aritmetični sredini μ uporabimo naslednjo testno statistiko: $t = \frac{\overline{X} - \mu_H}{\frac{s}{\sqrt{n}}}$

[1] Če je populacijski standardni odklon znan, lahko tudi pri majhnih vzorcih uporabimo Z porazdelitev.

Primer: interval zaupanja za μ za mali vzorec

Vzemimo, da se spremenljivka *X* (število ur branja dnevnih časopisov na teden) porazdeljuje normalno *N(μ, σ)*. Na osnovi podatkov za 7 slučajno izbranih oseb ocenimo interval zaupanja za aritmetično sredino pri 10% tveganju. Podatki za 7 oseb: 5 7 9 7 6 10 5

Enota: osebe.

Spremenljivka: število ur branja dnevnih časopisov na teden.

Populacijski parameter, ki ga ocenjujemo: populacijska aritmetična sredina μ (povprečno število ur branja dnevnih časopisov na teden).

Ker imamo podatke za mali vzorec, bomo interval zaupanja za populacijsko aritmetično sredino izračunali po obrazcu:

$$P(\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}) = 1-\alpha$$

V ta namen moramo izračunati vzorčno aritmetično sredino in vzorčni standardni odklon ter iz ustrezne tabele razbrati vrednost $t_{\alpha/2}$.

x_i	$x_i - \overline{X}$	$(x_i - \overline{X})^2$
5	-2	4
7	0	0
9	2	4
7	0	0
6	-1	2
10	3	9
5	2	4
Σ49	Σ0	Σ22

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{49}{7} = 7$$

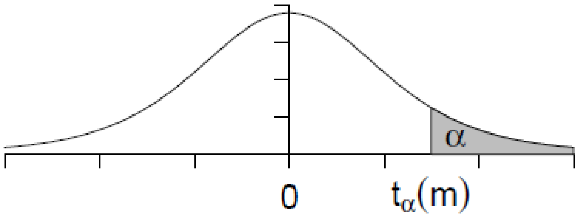
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \overline{X})^2 = \frac{22}{6} = 3.67$$

$$s = \sqrt{s^2} = 1.91$$

Iščemo vrednost *t* pri *α/2* in (*n-1*) prostostnih stopnjah. Iz tabele za *t* porazdelitev razberemo:

$$t_{\alpha/2}(n-1) = t_{0.10/2}(7-1) = t_{0.05}(6) = 1.94$$

Slika: Tabela za *t* porazdelitev (del tabele), pri čemer je $\alpha = P(t > t(m))$.



m \ α	0.100	0.050	0.025	0.010	0.005	0.001
1	3.08	6.31	12.71	31.82	63.66	318.31
2	1.89	2.92	4.30	6.96	9.92	22.33
3	1.64	2.35	3.18	4.54	5.84	10.21
4	1.53	2.13	2.78	3.75	4.60	7.17
5	1.48	2.02	2.57	3.36	4.03	5.89
6	1.44	$\boxed{1.94}$	2.45	3.14	3.71	5.21
7	1.41	1.89	2.36	3.00	3.50	4.79
8	1.40	1.86	2.31	2.90	3.36	4.50

Interval zaupanja je:

$$P(\overline{X} - t_{\alpha/2}\frac{s}{\sqrt{n}} < \mu < \overline{X} + t_{\alpha/2}\frac{s}{\sqrt{n}}) = 1 - \alpha$$

$$7 - 1.94\frac{1.91}{\sqrt{7}} < \mu < 7 + 1.94\frac{1.91}{\sqrt{7}}$$

$$7 - 1.4 < \mu < 7 + 1.4$$

$$5.6 < \mu < 8.4$$

Z 90% gotovostjo (oz. 10% tveganjem) ocenjujemo, da na populaciji osebe v povprečju berejo dnevne časopise od 5.6 do 8.4 ure na teden.

Primer: preverjanje domneve o μ za mali vzorec

Za slučajni vzorec 16 odraslih Slovencev smo izračunali povprečno število in varianco priznanih let šolanja: $\overline{X} = 9$ in $s^2 = 9$. Predpostavljamo, da se spremenljivka na populaciji porazdeljuje normalno. Ali lahko sprejmemo domnevo, da imajo odrasli Slovenci v povprečju več kot osemletko, pri 5% stopnji značilnosti?

Domnevi: H_0: $\mu = 8$

H_1: $\mu > 8$

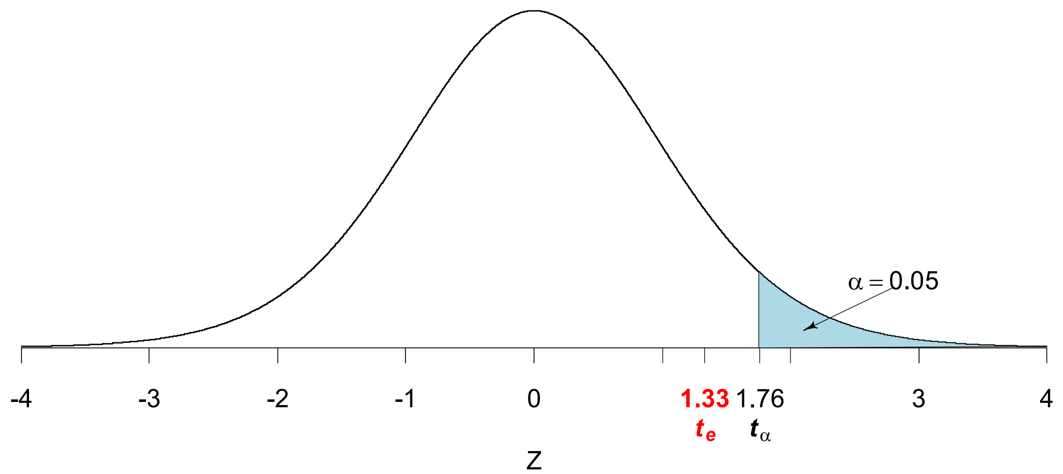
Ker imamo podatke za mali vzorec, upoštevamo, da se vzorčne aritmetične sredine pri malih vzorcih porazdeljujejo po Studentovi t porazdelitvi (z $m = n - 1$ prostostnimi stopnjami), zato je **testna statistika** naslednja: $t = \dfrac{\overline{X} - \mu_H}{\dfrac{s}{\sqrt{n}}}$

Izračunajmo njeno **eksperimentalno vrednost** (torej standardizirano vrednost vzorčne aritmetične sredine):

$$t_e = \frac{\overline{X} - \mu_H}{\dfrac{s}{\sqrt{n}}} = \frac{9-8}{\dfrac{3}{\sqrt{16}}} = 1.33$$

Ker gre za enostranski test, je glede na osnovno domnevo kritično območje na desni strani porazdelitve in **kritična vrednost** t_α naslednja:

$$t_\alpha(n\text{-}1) = t_{0.05}(16\text{-}1) = t_{0.05}(15) = 1.75$$



Sklep:

1.3 < 1.75 oz. $|t_e| < |t_\alpha|$ ... Eksperimentalna vrednost testne statistike ne pade v kritično območje. → H_0 ne zavrnemo.

(Izpis iz stat. programa: sig =.2034, kar beremo kot "p vrednost (dvostranska) je 0.2034" in pomeni $P(|t| > |1.33|) = 0.2034$. Enostranska p vrednost je 0.2034/2 = 0.1017 in pomeni $P(t > 1.33) = 0.1017$. Ker ni manjša od 0.05, ničelne domneve ne moremo zavrniti.)

Interpretacija: Ker eksperimentalna vrednost ne pade v kritično območje, ničelne domneve pri 5% stopnji značilnosti ne moremo zavrniti. Povprečno število let šolanja pri 5% tveganju ni statistično značilno večje od 8. Pri 5% tveganju ne moremo torej sprejeti domneve, da imajo odrasli Slovenci v povprečju več kot osemletko.

.

Obrazložitev: Vzorčna aritmetična sredina 0.9 (ki smo jo dobili na našem vzorcu) ali celo večja se lahko zgodi v kar 10% vseh možnih vzorcih (*p = 0.1045*) pri pravilni ničelni domnevi, zato ne moremo podvomiti v pravilnost ničelne domneve. Verjetnost napake, če bi ničelno domnevo zavrnili, bi bila 10.45%, kar je več kot dopustnih 5%.

4.6.3 Porazdelitev razlik vzorčnih aritmetičnih sredin

PONOVIMO ...

Za velike vzorce ($n_1, n_2 \geq 30$) velja - ne glede na obliko porazdelitve spremenljivke X na dveh obravnavnih populacijah - da se razlike vzorčnih aritmetičnih sredin $\overline{X}_1 - \overline{X}_2$ na vseh možnih parih vzorcev velikosti n_1 in n_2 porazdeljujejo normalno na naslednji način:

$$\overline{X}_1 - \overline{X}_1 : N(\mu_1 - \mu_2, \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}})$$

V primeru da se spremenljivka X na obeh populacijah porazdeljuje normalno, lahko populacijska standardna odklona σ_1 in σ_1 ocenimo z vzorčnima standardnima odklonoma s_1 in s_2. V takem primeru lahko:

- pri preverjanju domneve o razliki populacijskih aritmetičnih sredin μ_1 - μ_2 uporabimo naslednjo testno statistiko:

$$z = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)_H}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

ZA MALE VZORCE PA VELJA ...

Za male vzorce ($n_1, n_2 < 30$) velja - ob predpostavki, da se spremenljivka X na dveh obravnavanih populacijah porazdeljuje normalno in da sta populacijska standardna odklona enaka $\sigma_1 = \sigma_2 = \sigma$ - da se razlike vzorčnih aritmetičnih sredin $\overline{X}_1 - \overline{X}_2$ na vseh možnih parih vzorcev velikosti n_1 in n_2 porazdeljujejo po Studentovi t porazdelitvi z $m = n_1 + n_2 - 2$ prostostnimi stopnjami.

V tem primeru lahko populacijski standardni odklon σ oz. varianco σ^2 ocenimo s pomočjo obeh vzorčnih varianc, in sicer kot uteženo povprečje obeh vzorčnih varianc:

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

Nato lahko:

- pri preverjanju domneve o razliki populacijskih aritmetičnih sredin μ_1 - μ_2 na dveh populacijah uporabimo naslednjo testno statistiko:

$$t = \frac{\overline{X}_1 - \overline{X}_2 - (\mu_1 - \mu_2)_H}{s\sqrt{\frac{n_1 + n_2}{n_1 \cdot n_2}}}$$

Primer: preverjanje domneve o μ_1-μ_2 za male vzorce

V eksperiment, s katerim so ugotavljali razlike med moškimi in ženskami, je bilo vključenih 10 slučajno izbranih moških in 10 slučajno izbranih žensk. V okviru eksperimenta so morale te osebe opraviti več enostavnih nalog istočasno. Moški so v povprečju uspešno opravili 1.2 nalogi istočasno (s standardnim odklonom 0.5), ženske pa 2.1 nalog istočasno (s standardnim odklonom 0.8). Ali lahko na osnovi tega eksperimenta (ob 5% stopnji značilnosti) zaključimo, da so ženske sposobne opravljati več nalog istočasno kot moški?

(Predpostavljamo, da je populacijska varianca na obeh populacijah enaka in da se spremenljivka na obeh populacijah porazdeljuje normalno.)

Spremenljivka: X – število nalog, ki so jih osebe v eksperimentu sposobne opraviti istočasno.

Populacijski parameter, ki ga ocenjujemo: razlika v aritmetični sredini žensk in moških $\mu_Ž$ - μ_M (razlika med ženskami in moškimi v povprečnem številu nalog, ki so jih sposobni opravljati istočasno).

Podatki iz vzorcev:

- moški: $n_M = 10$, $\overline{X}_M = 1.2$, $s_M = 0.5$
- ženske: $n_Ž = 10$, $\overline{X}_Ž = 2.1$, $s_Ž = 0.8$

1. **Domnevi**:

H_0: $\mu_Ž = \mu_M$ *oz.* $\mu_Ž - \mu_M = 0$ (v obliki $\gamma = \gamma_H$)

H_1: $\mu_Ž > \mu_M$ *oz.* $\mu_Ž - \mu_M > 0$ (v obliki $\gamma > \gamma_H$) $\rightarrow$ enostranski test

2. Stopnja značilnosti α in **kritično območje**

$t_\alpha(n_Ž+n_M-2) = t_{0.05}(10+10-2) = t_{0.05}(18) = 1.73$

3. **Testna statistika** je standardizirana razlika vzorčnih aritmetičnih sredin in upoštevamo, da se v primeru majhnih vzorcev $\overline{X}_Ž - \overline{X}_M$ porazdeljujejo po Studentovi t porazdelitvi:

$$t = \frac{\overline{X}_Ž - \overline{X}_M - (\mu_Ž - \mu_M)_H}{s\sqrt{\dfrac{n_Ž + n_M}{n_Ž \cdot n_M}}}$$

Izračunajmo s, ki je ocena populacijske variance:

$$s^2 = \frac{(n_Ž - 1)s_Ž^2 + (n_M - 1)s_M^2}{n_Ž + n_M - 2} = \frac{(10-1)0.8^2 + (10-1)0.5^2}{10+10-2} = 0.445$$

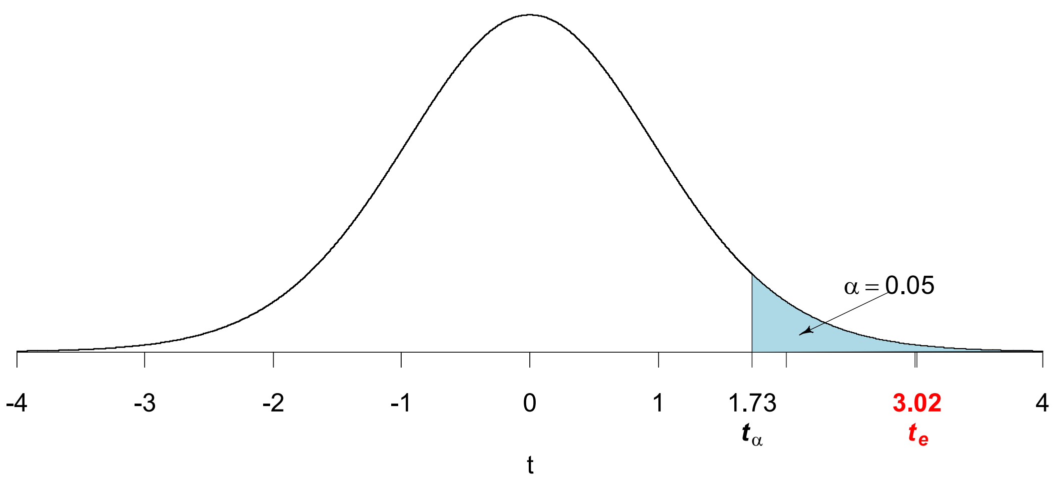
$$s = \sqrt{s^2} = \sqrt{0.445} = 0.667$$

Izračunamo eksperimentalno vrednost testne statistike:

$$t_e = \frac{\overline{X}_Ž - \overline{X}_M - (\mu_Ž - \mu_M)_H}{s\sqrt{\dfrac{n_Ž + n_M}{n_Ž \cdot n_M}}} = \frac{2.1 - 1.2 - 0}{0.667\sqrt{\dfrac{10+10}{10 \cdot 10}}} = 3.02$$

4. Sklep

3.02 > 1.73 oz. $|t_e| > |t_\alpha|$ Eksperimentalna vrednost testne statistike pade v kritično območje.
$\rightarrow$ H_0 zavrnemo.



(Izpis iz stat. programa: sig =.0074, kar beremo kot "p vrednost (dvostranska) je 0.0074" in pomeni $P(|t| > |3.02|) = 0.0074$. Enostranska p vrednost je 0.0074/2 = 0.0037 in pomeni $P(t > 3.02) = 0.0037$. Ker ni manjša od 0.05, ničelne domneve ne moremo zavrniti.)

Interpretacija: Ženske v povprečju opravijo statistično značilno (5% tveganje) več nalog istočasno kot moški.

Pomen: Razlika vzorčnih aritmetičnih sredin 0.9 spada med 5% največjih možnih razlik vzorčnih aritmetičnih sredin pri populacijski razliki 0, zato podvomimo v pravilnost ničelne domneve. Natančneje, razlika 0.9 ali več se lahko zgodi le pri 0.4% vseh možnih vzorcih (p = 0.0037), torej je zelo malo verjetna v primeru, da bi bila na populaciji razlika 0. Zato lahko zavrnemo domnevo, da je razlika populacijskih aritmetičnih sredin enaka 0.

4.7 DOLOČANJE VELIKOSTI VZORCA

Pomen velikosti vzorca

Velikost vzorca vpliva na razpršenost vzorčnih statistik, saj je standardna napaka vzorčnih statistik odvisna tudi od velikosti vzorca. Običajno velja, da če se vzorec poveča za faktor f, se standardna napaka zmanjša za faktor $\sqrt{f}$.

Pri večjih vzorcih je torej razpršenost vzorčnih statistik na vseh možnih vzorcih manjša. Posledično so ožji tudi intervali zaupanja, saj sta meji intervala odvisni od standardne napake $SE(g)$ in s tem tudi od velikosti vzorca n.

Če torej raziskavo izvedemo na večjem vzorcu, bo izračunani interval zaupanja za nek populacijski parameter ožji in s tem ocena bolj natančna.

Opis problema

Raziskovalci navadno vedo, kako natančno želijo na osnovi vzorčnih podatkov oceniti parametre, ki jih potrebujejo (npr. aritmetično sredino neke spremenljivke ali delež neke lastnosti na populaciji).

Npr. v javnomnenjskih raziskavah pred volitvami želijo oceniti delež volivcev, ki bodo volili določeno stranko, na *1%* natančno (torej bi bili zadovoljni z intervalno oceno npr. *15%±1%*, kar je interval (*14%* do *16%*)).

Raziskovalec glede na neke kriterije določi, kolikšna je lahko največja razlika med iskanim populacijskih parametrom γ (npr. deležem volivcev, ki bodo volili neko stranko, na populaciji) in njegovo vzorčno oceno g (deležem volivcev te stranke na vzorcu).

To razliko lahko označimo z E, kar izhaja iz angleške besede *error*, saj pomeni ravno napako, ki jo raziskovalec še dovoljuje pri ocenjevanju populacijskega parametra na osnovi vzorca. Zapišemo kot:

$$\gamma - g = E\ .$$

Gre torej za razliko med pravo populacijsko vrednostjo in oceno iz vzorca, torej za neko napako. Ta razlika nastopa tudi v intervalu zaupanja.

Če predpostavljamo, da se cenilka (vzorčne ocene) porazdeljujejo normalno in če se populacijski parameter nahaja v intervalu zaupanja, torej med mejama $a=g-z_{\alpha/2}\cdot SE(g)$ ter $b=g+z_{\alpha/2}\cdot SE(g)$, je torej odstopanje od ocene iz vzorca do mej intervala zaupanja ($z_{\alpha/2}\cdot SE(g)$) ravno dovoljena napaka E. Odstopanje od ocene vzorca do mej intervala zaupanja mora torej biti manjše od dovoljene razlike oz. napake:

$$z_{\alpha/2}\cdot SE(g) < \gamma - g = \mathrm{E}$$

Če v takšno neenačbo vstavimo formulo za $SE(g)$, ki praviloma vključuje n, lahko na podlagi te neenačbe izpeljemo obrazec za izračun najmanjše še primerne velikosti vzorca.

4.7.1 Določanje velikosti vzorca, ko ocenjujemo aritmetično sredino

Ko želimo določiti potrebno velikost vzorca za zaželeno natančno oceno populacijske aritmetične sredine, v zgornjo neenačbo vpišemo obrazec za standardno napako vzorčnih aritmetičnih sredin ter iz neenačbe izpeljemo n (velikost vzorca).

$$z_{\alpha/2} se(g) < E$$

$$z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < E$$

$$n > \left(\frac{z_{\alpha/2} \sigma}{E} \right)^2$$

Primer: Določanje velikosti vzorca, ko ocenjujemo aritmetično sredino

Z vzorčno raziskavo želimo pri 5% tveganju oceniti povprečno starost podjetnikov malih podjetij na 1 leto natančno. Dovoljujemo torej, da bo razlika med populacijsko aritmetično sredino (pravo vrednostjo) ter ocenjeno aritmetično sredino iz vzorca (vzorčno oceno) največ eno leto. Pri tem ocenjujemo, da je populacijski standardni odklon za povprečno starost σ = 10 let. Izberemo 5% stopnjo tveganja. Potrebno velikost vzorca določimo na naslednji način:

$$n > \left(\frac{z_{\alpha/2} \sigma}{E} \right)^2$$

$$n > \left(\frac{1.96 \cdot 10}{1} \right)^2$$

$$n > 384.2$$

Če želimo doseči postavljeno natančnost ocenjevanja pri izbranem tveganju (takšno oceno, da se pri 5% tveganju ne zmotimo za več kot 1 leto), potrebujemo slučajni vzorec velikosti vsaj 385 enot.

4.7.2 Določanje velikosti vzorca, ko ocenjujemo delež

Podobno lahko ocenimo velikost vzorca, če želimo ocenjevati z določeno natančnostjo populacijski delež. V splošno neenačbo vstavimo standardno napako za vzorčne deleže in po preureditvi potrebno velikost vzorca določimo na naslednji način:

$$z_{\alpha/2} se(g) < E$$

$$z_{\alpha/2} \sqrt{\frac{\pi(1-\pi)}{n}} < E$$

$$n > \frac{z_{\alpha/2}^2 \pi(1-\pi)}{E^2}$$

Primer: Določanje velikosti vzorca, ko ocenjujemo delež

S predvolilno anketo, ki jo izvajamo en teden pred volitvami, bi želeli oceniti odstotek volivcev, ki bodo volili SDS, na 1% natančno pri 5% tveganju. Na podlagi preteklih raziskav ocenjujemo, da bo delež okoli 25%. Dovoljujemo torej, da bo razlika med populacijskim deležem (pravo vrednostjo) ter ocenjenim deležem (vzorčno oceno) največ 0.01. Pri napovedi bi se torej želeli zmotiti za manj kot 1%.

$$n > \frac{z_{\alpha/2}^2 \pi(1-\pi)}{E^2}$$

$$n > \frac{1.96^2 \cdot 0.25(1-0.25)}{0.01^2}$$

$$n > 7203$$

Če želimo doseči postavljeno natančnost ocenjevanja (takšno oceno, da se ne zmotimo za več kot 1 %) pri 5% tveganju, potrebujemo pri oceni deleža (če predpostavljamo, da je okoli 25%) zelo velik vzorec, in sicer vsaj 7203 enot.

4.8 VAJE

4.8.1 Vzorčenje, porazdelitve vzorčnih statistik: uvod, veliki vzorci

1. Ferligoj: Naloge iz statistike: 7.1, 7.2, 7.3, 7.5

2. Populacijo sestavljajo N=4 enote. Vrednosti spremenljivke X so 0, 1, 2 in 3. Spremenljivka se na populaciji torej porazdeljuje enakomerno.
   a) Poiščite populacijsko aritmetično sredino μ in populacijski standardni odklon σ spremenljivke X.
   b) Določite vse možne enostavne slučajne vzorce brez ponavljanja, velikosti n=2.
   c) Za vsak vzorec izračunajte vzorčno aritmetično sredino $\overline{X}_i$.
   d) Vzorce obravnavajte kot novo populacijo, vzorčno aritmetično sredino pa kot novo slučajno spremenljivko $\overline{X}$. Zapišite verjetnostno shemo te spremenljivke ter jo tudi grafično predstavite s histogramom.
   e) Poiščite matematično upanje in standardno napako slučajne spremenljivke vzorčna aritmetična sredina $\overline{X}$.
   f) Kakšna je povezava med populacijsko aritmetično sredino μ in standardnim odklonom σ spremenljivke X ter matematičnim upanjem $E(\overline{X})$ in standardno napako $SE(\overline{X})$ slučajne spremenljivke $\overline{X}$.

3. Predpostavimo, da je število prijateljev za redne študente FDV normalno porazdeljena slučajna spremenljivka X z aritmetično sredino *5* in standardnim odklonom *2* (X: $N(5,2)$).
   a) Koliko prijateljev imajo tisti študenti, ki so med 40% študentov z najmanj prijatelji?
   b) Koliko izmed 300 študentov ima med 2 in 7 prijateljev?
   c) Koliko izmed 300 študentov ima več kot 7 prijateljev?
   d) Iz populacije študentov tvorimo enostavne slučajne vzorce s ponavljanjem velikosti 50 enot. Kako se porazdeljujejo vzorčne aritmetične sredine za število prijateljev?
   e) Kolikšno bo povprečno število prijateljev za 40% tistih vzorcev, kjer bo to povprečno število najmanjše?
   f) Kolikšen del vzorcev bo imelo povprečno število prijateljev med 2 in 7?
   g) Kolikšen del vzorcev bo imel povprečno število prijateljev večje od 7?

4. Denimo, da se indeks besedne agresivnosti med slovenskimi študenti porazdeljuje približno normalno s parametroma $\mu=7.6$ in $\sigma=2.7$.
   a) Denimo, da smo opravili anketo, v katero je bil vključen enostavni slučajni vzorec s ponavljanjem velikosti n = 30 slovenskih študentov. V katerih mejah se nahaja 90% aritmetičnih sredin teh vzorcev?
   b) Izračunajte interval, v katerih bi se nahajajo 90% aritmetičnih sredin teh vzorcev, če bi bil vzorec velik n = 100 / 1000 / 3000. Kaj se dogaja s tem intervalom, če povečujemo velikost vzorca? Pojasnite.
   c) Kolikšen povprečni indeks besedne agresivnosti ima 1% najbolj pristranskih vzorcev (vzorcev z najmanjšim ali z največjim povprečnim indeksom) velikosti n = 30 / 100 / 1000 /3000? Pojasnite rezultate.

5. Anketarska družba želi ugotoviti, kolikokrat tedensko slovenska gospodinjstva nakupujejo v prehrambenih trgovinah. Denimo, da rezultate ankete izpred enega leta vzamemo kot prave (populacijske) vrednosti za to slučajno spremenljivko, in sicer *X: N(5, 2)*. Anketarska družba namerava izvesti anketo na enostavnem slučajnem vzorcu brez ponavljanja velikost n=1000 gospodinjstev.
   a) Koliko različnih enostavnih slučajnih vzorcev brez ponavljanja velikost 1000 gospodinjstev lahko tvori, če upoštevamo, da je v Sloveniji 500.000 gospodinjstev?
   b) V katerih mejah se po pričakovanjih nahaja 95% vseh aritmetičnih sredin teh vzorcev?

6. V neki evropski državi imajo 100 nogometnih klubov. Zadolženost teh nogometnih klubov se porazdeljuje približno normalno s povprečno zadolženostjo 230.000 evrov ter standardnim odklonom 110.000 evrov.
   a) Kolikšna je verjetnost, da bo izračunana povprečna zadolženost klubov na enostavnem slučajnem vzorcu s ponavljanjem velikosti 36 enot manjša od 200.000 evrov?
   b) V katerih mejah se nahaja 99% aritmetičnih sredin vseh enostavnih slučajnih vzorcev s ponavljanjem velikosti 36 enot?

7. Delež med vernimi volivci, ki so na zadnjih predsedniških volitvah volili nekega kandidata, je bil 35%. Ta delež med nevernimi volivci pa je bil 55%. Na obeh populacijah delamo vzorce s ponavljanjem velikosti 300 enot. Kolikšen delež vseh možnih parov vzorcev nevernih in vernih volivcev bi imel absolutno razliko med tema dvema deležema manjšo od 10%? Zanima nas torej verjetnost, da je $|p_1 - p_2| < 0.10$.

8. V populaciji neke države je polovica prebivalcev uporabnikov interneta.
   a) Recimo, da iz populacije tvorimo enostavne slučajne vzorce s ponavljanjem velikosti 25 enot. Kakšna je verjetnost, da bo na izbranem vzorcu uporabnikov interneta več kot 60%?
   b) Kolikšna je ta verjetnost, če bi tvorili vzorec velikosti 100 enot?
   c) V katerih mejah se nahaja 95% vseh vzorčnih deležev za vzorce velikosti 25 (100) enot, simetrično na populacijski delež?

9. Za populacijo študentov iz vasi in mesta imamo naslednje podatke glede pogostosti političnega udejstvovanja, merjeno s številom volitev ali referendumov, ki so se jih udeležili v zadnjih petih letih. Študenti iz mesta so se v povprečju udeležili 3,5 takih dogodkov v zadnjih petih letih, st. odklon pa je bil 1. Študenti iz vasi so se v povprečju udeležili 3 takih dogodkov, st. odklon pa je bil prav tako 1.
   a) Recimo, da smo med študenti iz mesta izbrali enostavni slučajni vzorec s ponavljanjem 36 študentov, med študenti iz vasi pa vzorec velikosti 50 študentov. Kolikšna je verjetnost, da se bodo za ta dva slučajno izbrana vzorca študenti iz mesta pogosteje politično udejstvovali kot študenti iz vasi?
   b) V katerih mejah se nahaja 99% razlik vzorčnih aritmetičnih sredin, če bi tvorili vse možne pare vzorcev iz obeh populacij?

10. Povprečna življenjska doba sijalk proizvajalca A je 1400 ur, s standardnim odklonom 200 ur. Povprečna življenjska doba sijalk proizvajalca B pa je 1200 ur, s standardnim odklonom 100 ur. Če preizkusimo slučajna vzorca 125 sijalk obeh proizvajalcev, kakšna je verjetnost, da bodo sijalke proizvajalca A iz vzorca imele v povprečju za vsaj 160 ur (in vsaj 250 ur) daljšo življenjsko dobo kot sijalke proizvajalca B iz vzorca?

11. Primeri izpitnih vprašanj:

Katera od naštetih metod vzorčenja **ni** verjetnostna metoda?
a) enostavno slučajno vzorčenje
b) stratificirano vzorčenje
c) vzorčenje v skupinah
d) ekspertna izbira

V populaciji imamo 600 žensk in 400 moških. V vzorec smo izbrali 50 žensk in 50 moških. Kakšen tip vzorčenja smo uporabili?
a) proporcionalno stratifikacijo
b) disproporcionalno stratifikacijo
c) vzorčenje v skupinah
d) večstopenjsko vzorčenje

Če želimo pri ocenjevanju populacijske aritmetične sredine zmanjšati velikost standardne napake, moramo
a) zmanjšati velikost vzorca
b) povečati velikost populacije
c) zmanjšati stopnjo tveganja
d) povečati velikost vzorca

Slučajna spremenljivka je porazdeljena po zakonu $X: N(2, 1)$. Iz tega sledi, da se vzorčne aritmetične sredine vzorcev velikosti 100 enot porazdeljujejo
a) $\overline{X} : N(2, 0.1)$
b) $\overline{X} : N(2, 0)$
c) $\overline{X} : N(0, 0.1)$
d) $\overline{X} : N(2, 0.01)$

4.8.2 Intervali zaupanja

1. Ferligoj: Naloge iz statistike: 7.4, 7.6, 7.7, 7.9, 7.20

2. Dopolnitev naloge 6 iz 4.8.1.
V neki evropski državi imajo 100 nogometnih klubov. Zadolženost teh nogometnih klubov v neki evropski državi se porazdeljuje približno normalno s povprečno zadolženostjo 230.000 evrov ter standardnim odklonom 110.000 evrov.
   a) Recimo, da smo izbrali enostavni slučajni vzorec velikost 36 enot. Povprečna zadolženost teh 36-ih klubov je 220.000 evrov, vzorčni standardni odklon pa 100.000 evrov. Na osnovi teh podatkov izračunajte 99% interval

zaupanja za povprečno zadolženost nogometnih klubov. (Pri tem predpostavljamo, da se zadolženost klubov na populaciji porazdeljuje približno normalno in da populacijskega standardnega odklona ne poznamo.)

b) Ali izračunani interval zaupanja vsebuje dejansko populacijsko aritmetično sredino? Pojasnite odgovor.

3. V raziskavi *Slovensko javno mnenje* v letu 2001 je med drugim 293 slučajno izbranih oseb izrazilo svoje strinjanje oz. nestrinjanje s trditvijo "*Članstvo v Evropski uniji bi Sloveniji omogočilo vsestranski razvoj in napredek*". Odgovor so izbirali na lestvici od 1 "*sploh se ne strinjam*" do 5 "*popolnoma se strinjam*". Povprečno strinjanje s to trditvijo je bilo 2.65, standardni odklon pa 1.03.

   a) Izračunajte interval, v katerem se z 10% tveganjem nahaja populacijsko povprečje. Pri tem predpostavljamo, da se spremenljivka na populaciji porazdeljuje normalno.

   b) Izračunajte interval tudi s 5% in 1% tveganjem. Kaj se dogaja z velikostjo intervala, če tveganje zmanjšujemo?

4. Eno izmed vprašanj, ki je kontinuirano vključeno v raziskavo *Slovensko javno mnenje*, je tudi vprašanje "*Ali ste član politične stranke?*". V raziskavi leta 2002 je na slučajnem vzorcu na to vprašanje pritrdilno odgovorilo 50 anketirancev izmed 1121 anketirancev.

   a) Ocenite 99% interval zaupanja za populacijski delež oseb, ki so člani politične stranke.

   b) Izračunajte, kakšen bi bil ta interval, če bi bil vzorec velik le 100 enot. Pojasnite.

5. Primeri izpitnih vprašanj:

Če namesto 100 enot v vzorec vzamemo 10.000 enot, se interval zaupanja za aritmetično sredino

a) zoži
b) razširi
c) ne spremeni
d) zoži za 100-krat

Stopnja tveganja pri intervalu zaupanja je verjetnost

a) da z oceno zgrešimo pravo vrednost parametra
b) nasprotnega dogodka
c) gotovosti
d) da z oceno zgrešimo pravo vrednost statistike

Na vzorcu (s ponavljanjem) 500 zaposlenih je bilo 250 takih, ki niso zadovoljni s svojim delom. Izračunajte interval zaupanja za delež zaposlenih, ki niso zadovoljni s svojim delom, med vsemi zaposlenimi, pri stopnji tveganja 0,05.

a) [0.48, 0.52]
b) [1.223, 1.227]
c) [228, 272]
d) [0.46, 0.54]

4.8.3 Preverjanje domnev

1. Ferligoj: Naloge iz statistike: 8.1, 8.3, 8.4, 8.5, 8.7, 8.8, 8.11.
2. Primeri izpitnih vprašanj:

Pri preverjanju domnev se napaki II. vrste popolnoma izognemo tako, da

a) ničelne domneve samo sprejemamo
b) vzamemo večji vzorec
c) povečamo stopnjo tveganja
d) ničelne domneve samo zavračamo

Statistika, ki jo uporabljamo za preverjanje domneve o razliki aritmetičnih sredin za velike vzorce, se porazdeljuje

a) enakomerno
b) normalno
c) diskretno
d) po hi-kvadrat porazdelitvi

Pri preverjanju domneve, da je večina Slovencev *za* vstop v NATO, je alternativna domneva:

a) $\pi < 0,5$
b) $\mu < 0,5$
c) $\pi > 0,5$
d) $\pi = 0,5$

Povprečno število klicev na dan z mobilnim telefonom, ki jih opravi 100 slučajno (s ponavljanjem) izbranih študentov FDV, je 9, standardni odklon pa 3. Preveriti želimo domnevo, da študenti FDV v povprečju opravijo na dan z mobilnim telefonom manj kot 10 klicev. Na podlagi zgornjih podatkov lahko pri stopnji značilnosti 10% ugotovimo:

a) $z_e = -3.33$, $z_a = -1.28$, v povprečju opravijo manj kot 10 klicev na dan
b) $z_e = -3.33$, $z_{\alpha/2} = -1.65$, v povprečju opravijo manj kot 10 klicev na dan
c) $z_e = 10$, $z_{\alpha/2} = -1.65$, v povprečju ne opravijo manj kot 10 klicev na dan
d) $z_e = -3.33$, $z_\alpha = -1.28$, v povprečju opravijo 10 klicev na dan

4.8.4 Teorija malih vzorcev

1. Ferligoj: Naloge iz statistike: 7.11, 7.13, 7.22, 8.6, 8.12, 8.13.
2. Na enostavnem slučajnem vzorcu s ponavljanjem 15 študentov FDV smo dobili naslednje podatke o številu najboljših prijateljev: 1 2 3 5 1 1 3 2 1 5 11 2 2 4 2
   (Predpostavljamo, da se spremenljivka na populaciji porazdeljuje približno normalno.)
   a) Kaj je v tem primeru enota, vzorec in spremenljivka?
   b) Ocenite interval, v katerem se s stopnjo gotovosti 95% nahaja populacijska aritmetična sredina. Rezultat interpretirajte.
   c) Recimo, da bi enako vzorčno aritmetično sredino in vzorčni standardni odklon dobili na vzorcu 100 študentov. Kakšen bi bil v tem primeru interval zaupanja za populacijsko aritmetično sredino? Pojasnite rezultat.

   Rezultati: b) [1.57, 4.43], c) [2.49, 3.51]

3. Primeri izpitnih vprašanj:

Pri določanju intervala zaupanja za aritmetično sredino za vzorce, manjše od 20 enot, si pomagamo z naslednjo porazdelitvijo:

a) standardizirano normalno
b) hi-kvadrat
c) Studentovo
d) binomsko

Na vzorcu 15 žensk in 20 moških smo ugotovili, da so bile v preteklem letu ženske v kinu v povprečju 3,6-krat, ocena standardnega odklona $s_Ž$ pa je bila 0.4 Pri moških pa je bilo povprečje v enakem obdobju 4,5, ocena standardnega odklona s_M pa je bila 0.5. Predpostavljamo, da se spremenljivka »*število obiskov kinopredstav v preteklem letu*« na obeh populacijah porazdeljuje normalno in da sta populacijska standardna odklona enaka. Pri preverjanju ničelne domneve, da je povprečno število obiskov kinopredstav v preteklem letu pri moških in ženskah na populaciji enaka, proti alternativni domnevi, da sta povprečji na populaciji različni, upoštevajoč 10% stopnjo značilnosti, ugotovimo naslednje:

a) Eksperimentalna vrednost testne statistike t_e = 5.57, meja kritičnega območja pa je določena z $t_{\alpha/2} = \pm 1.69$. Populacijski aritmetični sredini sta si statistično značilno različni.
b) Eksperimentalna vrednost testne statistike t_e = 5.57, meja kritičnega območja pa je določena z $t_\alpha = 2.44$. Ničelne domneve ne moremo zavrniti.
c) Eksperimentalna vrednost testne statistike z_e = 5.91, meja kritičnega območja pa je določena z $z_{\alpha/2} = 1.65$. Ničelno domnevo lahko zavrnemo.
d) Eksperimentalna vrednost testne statistike t_e = 5.57, meja kritičnega območja pa je določena z $t_{\alpha/2} = \pm 1.69$. Populacijski aritmetični sredini se statistično značilno ne razlikujeta med seboj.

4.8.5 Določanje velikosti vzorca

1. Ferligoj: Naloge iz statistike: 7.14, 7.15, 7.16, 7.17, 7.18, 7.19.
2. Primer izpitnega vprašanja

Oceniti želimo težo vojakov tako, da bo razlika med populacijskih povprečjem in ocenjenim povprečjem iz vzorca manjša od 2 kg. Kako velik vzorec potrebujemo, če predpostavljamo, da je standardni odklon na populaciji 10 kg in izberemo 5% tveganje?

a) več kot 193 enot
b) 10 enot ali več
c) več kot 166 enot

d) več kot 96 enot