Primjer 7-8

U uzorku od 500 vlasnika kuća 175 planira da proda kuće i odseli se. Odrediti 98% interval povjerenja za one koji žele da prodaju kuće i odsele se.

- Imamo n=500, p=175/500=.35, i z=2.33
- 98% interval povjerenja je

$$.35 \pm 2.33\sqrt{\frac{(.35)(.65)}{500}} \text{ ili } .35 \pm .0497$$

Korektivni faktor za konačnu populaciju

- Konačne populacije
- Kod konačne populacije koja ima veličinu N a veličina uzorka je n, neophodno je napraviti sledeće prilagođavanje, ako je n/N veće ili jednako od 5%:
- Standardna greška je onda:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \sqrt{\frac{N - n}{N - 1}}$$

Korektivni faktor za konačnu populaciju

– Standardna greška za proporciju:

$$\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}\sqrt{\frac{N-n}{N-1}}$$

Primjer 7-9

- Napraviti 95% interval povjerenja za prosječan broj radnih sati u nedjelji ako ima 500 studenata u kampusu. Izabran je uzorak od 49 studenata koji su u prosjeku radili 24 sata a stand. dev. populacije je 4 sata.
- Pošto je n/N = 49/500 = .098>.05, moramo koristiti korektivni faktor za konačnu populaciju.

$$24 \pm 1.96\left(\frac{4}{\sqrt{49}}\right)\left(\sqrt{\frac{500-49}{500-1}}\right) = [22.9352,\ 25.0648]$$

Minimalna veličina uzorka: prosjek i proporcija

Minimalno zahtijevana veličina uzorka kod ocjene prosjeka populacije, μ:

$$n = \frac{z_{\frac{\alpha}{2}}^2 \sigma^2}{E^2}$$

Margina ocjene:

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

Minimalno zahtijevana veličina uzorka kod ocjene proporcije populacije, p

$$n = \frac{z_{\frac{\alpha}{2}}^2 pq}{E^2}$$

Određivanje veličine uzorka (Primjer 7-10)

Firma za marketing istraživanje želi da sprovede istraživanje za procjenu prosječnog iznosa koji troši svaka osoba na zabavu pri posjeti popularnog hotelskog kompleksa. Ljudi koji planiraju istraživanje žele da odrede da prosječan iznos koji potroše svi ljudi koji posjete kompleks bude do \$120 od prave vrijednosti, sa 95% povjerenja. Na osnovu prethodnih podataka, procjena standardne devijacije populacije je σ = \$400. Koja je minimalna neophodna veličina uzorka?

$$n = \frac{z_{\frac{\alpha}{2}}^2 \sigma^2}{E^2}$$

$$= \frac{(1.96)^2 (400)^2}{120^2}$$

$$= 42.684 \approx 43$$

Određivanje veličine uzorka (Primjer 7-11)

Proizvođači sportskih auta žele da ocjene proporciju ljudi u datoj klasi prihoda koji su zainteresovani za dati model. Kompanija želi da zna proporciju populacije, p, da bude do 0.1 od prave vrijednosti sa 99% povjerenja. Tekući podaci kompanije pokazuju da proporcija p može biti oko 0.25. Kolika je minimalna zahtijevana veličina uzorka za ovo istraživanje?

$$n = \frac{z_{\frac{\alpha}{2}}^2 pq}{E^2}$$

$$= \frac{2.576^2(0.25)(0.75)}{0.10^2}$$

$$= 124.42 \approx 125$$

8 Testiranje hipoteza

- Testiranje statističkih hipoteza
- Dvostrani test velikog uzorka za prosjek populacije
- Dvostrani test malog uzorka za prosjek populacije
- Dvostrani test velikog uzorka za proporciju populacije
- Jednostrani testovi
- *p*-vrijednost
- Vjerovatnoća tipa greške II i moć testa

8-2 Testiranje statističkih hipoteza

- **Nulta hipoteza**, u oznaci H_0, je tvrdnja o jednom ili više parametara populacije. To je tvrdnja za koju vjerujemo da je tačna dok ne dobijemo dovoljno statističkih dokaza da zaključimo drugačije.
  - H_0: $\mu=100$
- **Alternativna hipoteza**, u oznaci H_1, je tvrdnja o svim situacijama koje *ne* pokriva nulta hipoteza.
  - H_1: $\mu \neq 100$

- H_0 i H_1 su:
  - Međusobno isključive
    - Samo jedna može biti tačna.
  - Iscrpne
    - Zajedno pokrivaju *sve* mogućnosti, tako da ili jedna ili druga *mora* biti tačna.

Statistika testa i pravilo odlučivanja

- **Statistika testa** je statistika uzorka izračunata na osnovu podataka iz uzorka. Vrijednost statistike testa se koristi u u određivanju da li možemo ili ne odbaciti nultu hipotezu.
- **Pravilo odlučivanja** testa stat. hipoteza je pravilo koje određuje uslove pod kojima se nulta hipoteza može odbaciti.

Stvarno stanje, odluka, i dvije moguće greške

Mogući ishodi testiranja statističkih hipoteza

	Stvarno stanje	
Odluka	**H_0 tačna**	**H_0 netačna**
Ne odbacimo H_0	Ispravna odluka	Greška II vrste ili β greška
Odbacimo H_0	Greška I vrste ili α greška	Ispravna odluka

Elementi dvostranog standardizovanog testa, velikog uzorka, za prosjek populacije

Nulta Hipoteza	$H_0: \mu = \mu_0$
Alternativna Hipoteza	$H_1: \mu \neq \mu_0$
Nivo značajnosti testa	α (često 0.05 ili 0.01)
Statistika testa	$z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ (ako σ nije poznata, inače se mijenja σ sa s)
Kritične vrijednosti	$\pm z_{\frac{\alpha}{2}}$
Pravilo odlučivanja	Odbaciti nultu hipotezu ako je $z > z_{\frac{\alpha}{2}}$ ili $z < -z_{\frac{\alpha}{2}}$

Kritične vrijednosti *z*

α	$\frac{\alpha}{2}$	$z_{\frac{\alpha}{2}}$
0.01	0.005	2.576
0.02	0.010	2.326
0.05	0.025	1.960
0.10	0.050	1.645
0.20	0.100	1.282

Primjer 8-1

Istraživač želi da testira nultu hipotezu da je prosječna težina ručnog prtljaga po osobi $\mu_0=12$ funti, naspram alternativne hipoteze da prosječna težina nije 12 funti. Analitičar želi da testira nultu hipotezu sa $\alpha = 0.05$. Podaci su dati na sl. slajdu.

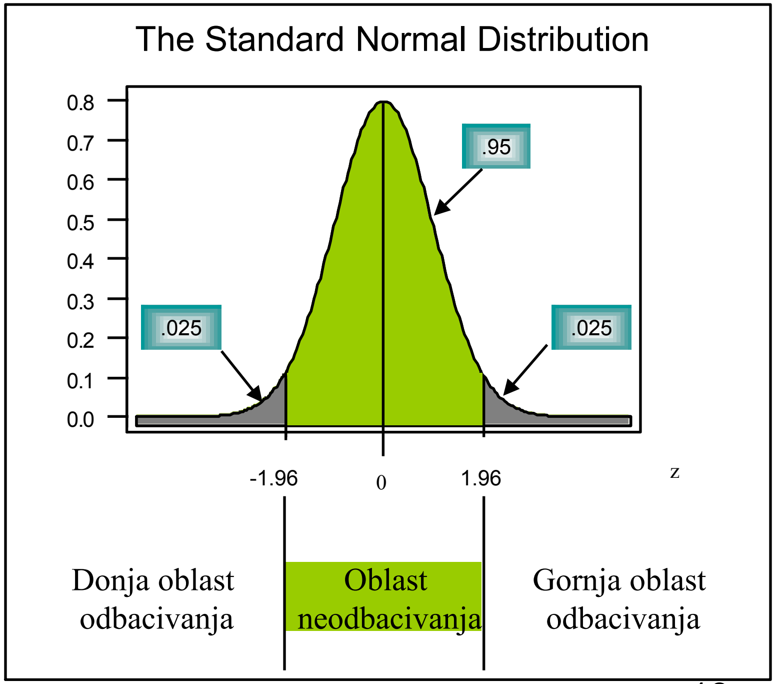
H_0: $\mu = 12$
H_1: $\mu \neq 12$

Za $\alpha = 0.05$, kritične vrijednosti z su ± 1.96

Statistika testa je: $$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Ne odbaciti H_0 ako: $[-1.96 \leq z \leq 1.96]$

Odbaciti H_0 ako: $[z < -1.96]$ ili $[z > 1.96]$



Primjer 8-1: Rješenje

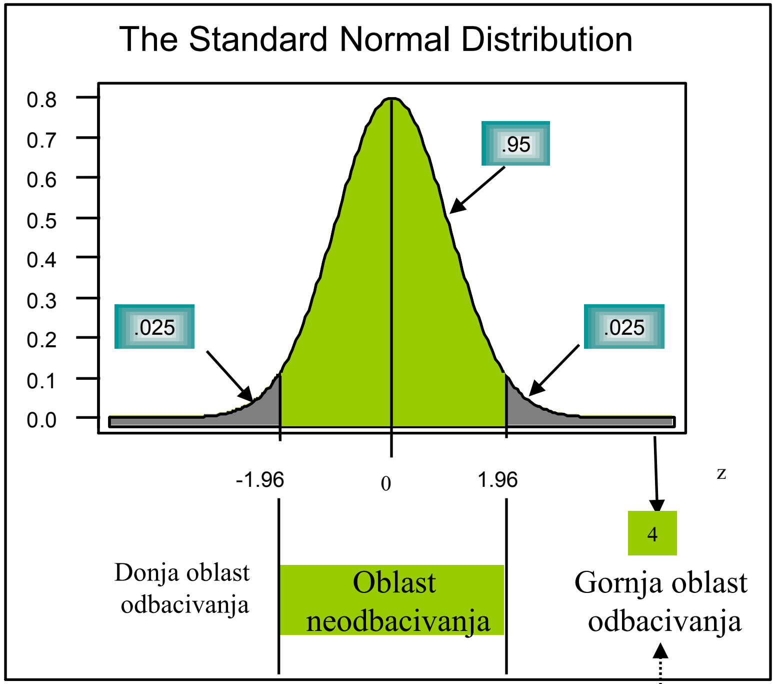
n = 144

x = 14.6

s = 7.8

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{14.6\text{-}12}{\frac{7.8}{\sqrt{144}}}$$

$$= \frac{2.6}{0.65} = 4$$



Pošto statistika testa upada u gornju oblast odbacivanja, H_0 se odbacuje, i možemo zaključiti da je prosječni ručni prtljag više od 12 funti.

8-2. zadatak

U slučajnom uzorku od 28 posmatranja, prosjek iznosi 32. Poznato je da osnovni skup ima normalnu raspodjelu sa $\sigma = 4$. Uz rizik greške 1%, ako se vrši testiranje hipoteze: $H_0 : \mu = 20$, $H_1 : \mu \neq 20$, kritična vrijednost za test je:

$$\alpha = 0.01$$

$$z_{\frac{\alpha}{2}} = z_{0,005} = 2,58$$

Primjer 8-3

Osiguravajuća kuća vjeruje da je, u nekoliko zadnjih godina, prosječno obavezno osiguranje po mjestu u upravnom odboru u "malim kompanijama" bilo \$2000. Koristeći $\alpha = 0.01$, testirati ovu hipotezu uz date podatke.

H_0: $\mu = 2000$
H_1: $\mu \neq 2000$

Za $\alpha = 0.01$, kritične vrijednosti z su ± 2.576

Statistika testa je: $z = \dfrac{\bar{x} - \mu_0}{\dfrac{s}{\sqrt{n}}}$

Ne odbaciti H_0 ako: $[-2.576 \le z \le 2.576]$

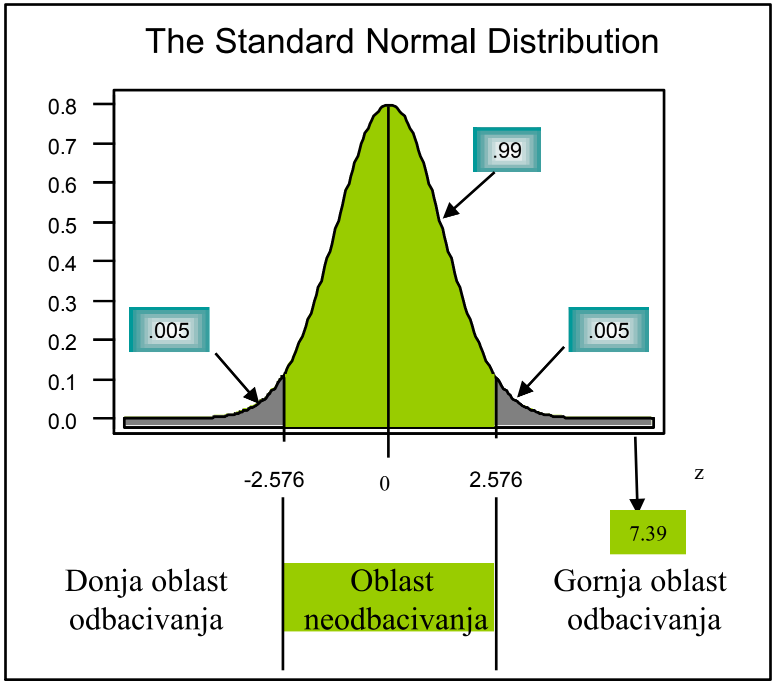
Odbaciti H_0 ako: $[z < -2.576]$ ili $[z > 2.576]$

$n = 100$
$\bar{x} = 2700$
$s = 947$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{2700 - 2000}{\frac{947}{\sqrt{100}}}$$

$$= \frac{700}{94.7} = 7.39 \Rightarrow \text{Odbaciti } H_0$$

Primjer 8-3: nastavak



Pošto statistika testa upada u gornju oblast odbacivanja, H_0 se odbacuje, i možemo zaključiti da je dato prosječno obavezno osiguranje u "malim kompanijama" više od $2000.

Primjer 8-4

Vjeruje se da je prosječno vrijeme koje treba kompjuteru da obavi određeni zadatak 3.24 sekunde. Odlučeno je da se testira statistička hipoteza da je prosječno funkcionalno vrijeme koristeći novi algoritam isto, protiv alternativne da prosječno vrijeme više nije isto, uz 0.05 nivo značajnosti.

H_0: $\mu = 3.24$
H_1: $\mu \neq 3.24$

Za $\alpha = 0.05$, kritične vrijednosti z su ± 1.96

Statistika testa je: $$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Ne odbaciti H_0 ako: $[-1.96 \leq z \leq 1.96]$

Odbaciti H_0 ako: $[z < -1.96]$ ili $[z > 1.96]$

n = 200
$\bar{x} = 3.48$
s = 2.8

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{3.48 - 3.24}{\frac{2.8}{\sqrt{200}}}$$

$$= \frac{0.24}{0.20} = 1.21 \Rightarrow \text{Ne odbaciti } H_0$$

8-4 Dvostrani standardizovani test, malih uzoraka za prosjek populacije

Kada je populacija normalna, stand. devijacija populacije, σ, nije poznata i veličina uzorka je mala, test hipoteza se bazira na *t* raspodjeli, sa (n-1) stepeni slobode, umjesto standardne normalne raspodjele.

Statistika testa t za male uzorke za prosjek μ:

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

Primjer 8-5

Prema nekoj agenciji za nekretnine, prosječna cijena zemlje u centralnom Tokiju porasla je 49% u prvih šest mjeseci 1995. Međunarodna kompanija za investiranje u nekretnine želi da testira ovu tvrdnju naspram alternativne da prosječna cijena nije rasla 49%, uz 0.01 nivo značajnosti.

H_0: $\mu = 49$
H_1: $\mu \neq 49$
n = 18
Za $\alpha = 0.01$ i (18-1) = 17 s.s.,
Kritične vrijednosti t su ± 2.898

Statistika testa je: $$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Ne odbaciti H_0 ako: $[-2.898 \leq t \leq 2.898]$

Odbaciti H_0 ako: $[t < -2.898]$ ili $[t > 2.898]$

$n = 18$
$\bar{x} = 38$
$s = 14$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{38 - 49}{\frac{14}{\sqrt{18}}}$$

$$= \frac{-11}{3.3} = -3.33 \Rightarrow \text{Odbaciti } H_0$$

Primjer 8-6

Canon, Inc,. je proizveo kopir mašinu koja ima mogućnost kopiranja u boji. Prosječna brzina standardnog kopir aparata je 27 kopija u minuti. Kompanija želi da testira da li novi kopir aparat ima istu prosječnu brzinu kao standardni. Sprovesti test uz nivo značajnosti $\alpha = 0.05$.

H_0: $\mu = 27$
H_1: $\mu \neq 27$
$n = 24$
Za $\alpha = 0.05$ i $(24-1) = 23$ s.s. ,
Kritične vrijednosti t su ± 2.069

Statistika testa je: $$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Ne odbaciti H_0 ako: $[-2.069 \leq t \leq 2.069]$
Odbaciti H_0 ako: $[t < -2.069]$ ili $[t > 2.069]$

$n = 24$
$\overline{x} = 24.6$
$s = 7.4$

$$t = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{24.6 - 27}{\frac{7.4}{\sqrt{24}}}$$

$$= \frac{-2.4}{1.51} = -1.59 \Rightarrow \text{Ne odbaciti } H_0$$

8-5 Dvostrani test velikih uzoraka za proporciju populacije

Kada je veličina uzorka velika (*np*> *5* i *nq* > *5*), raspodjela uzoračke proporcije se približava normalnoj raspodjeli sa prosjekom *p* i varijansom *pq*.

Statistika testa velikih uzoraka za proporciju populacije, p :

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

gdje je $q_0 = (1 - p_0)$

Elementi dvostranog testa velikih uzoraka za proporciju populacije

Nulta Hipoteza	H_0: $p=p_0$
Alternativna Hipoteza	H_1: $p \neq p_0$
Nivo značajnosti testa	α (često 0.05 ili 0.01)
Statistika testa	$z = \dfrac{\bar{p} - p_0}{\sqrt{\dfrac{p_0 q_0}{n}}}$ [gdje je $q_0 = (1-p_0)$]
Kritične vrijednosti	Granice $\pm z_{\frac{\alpha}{2}}$ koje zahvataju prostor $(1-\alpha)$
Pravilo odlučivanja	Odbaciti nultu hipotezu ako je $z > z_{\frac{\alpha}{2}}$ ili $z < -z_{\frac{\alpha}{2}}$

Kritične tačke z

α	$\frac{\alpha}{2}$	$z_{\frac{\alpha}{2}}$
0.01	0.005	2.576
0.02	0.010	2.326
0.05	0.025	1.960
0.10	0.050	1.645
0.20	0.100	1.282

Primjer 8-7

Investicioni analitičar želi da testira hipotezu britanskih eksperata obveznica da 70% svih stranih investitora na britanskom tržištu čine Amerikanci. Sakupio je sl. uzorak od 210 računa stranih investitora u Londonu i pronašao da je 130 američkih. Pri $\alpha = 0.05$ nivou značajnosti, da li se može odbaciti mišljenje britanskih eksperata obveznica?

H_0: $p = 0.70$

H_1: $p \neq 0.70$

n = 210

Za $\alpha = 0.05$ kritične vrijednosti z su ± 1.96

Statistika testa je:
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}}$$

Ne odbaciti H_0 ako: $[-1.96 \leq z \leq 1.96]$

Odbaciti H_0 ako: $[z < -1.96]$ ili $[z > 1.96]$

$$n = 210$$
$$\hat{p} = \frac{130}{210} = 0.619$$
$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0 q_0}{n}}} = \frac{0.619 - 0.70}{\sqrt{\frac{(0.70)(0.30)}{210}}}$$
$$= \frac{-0.081}{0.0316} = -2.5614 \Rightarrow \text{Odbaciti } H_0$$

8-6 Jednostrani testovi

U **jednostranom testu**, pitanje od interesa je da li je parametar populacije veći od (ili manji od) hipotetičke vrijednosti. U kontroli kvaliteta, ima više smisla testirati nultu hipotezu da je proporcija defektnih proizvoda manja od ili jednaka 0.10 naspram alternativne da je proporcija defektnih proizvoda veća od 0.10.

$$H_0\!: p \le 0.10$$
$$H_1\!: p > 0.10$$

Ovo vodi **desno-stranom testu**, pošto je cijela oblast odbacivanja na desnom kraju raspodjele.

Lijevi, desni, i dvostrani testovi

desno-strani test: $H_0: \mu \leq 50$
$H_1: \mu > 50$

lijevo-strani test: $H_0: \mu \geq 50$
$H_1: \mu < 50$

Dvostrani test: $H_0: \mu = 50$
$H_1: \mu \neq 50$

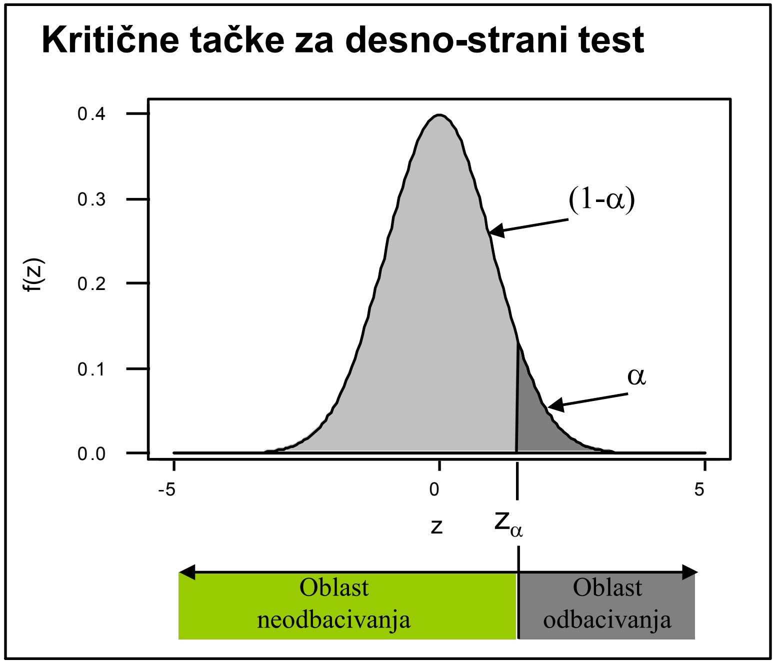
Elementi desno-stranog standadizovanog testa velikih uzoraka za prosjek populacije

Nulta Hipoteza	H_0: $\mu \leq \mu_0$
Alternativna Hipoteza	H_1: $\mu > \mu_0$
Nivo značajnosti testa	α (često 0.05 ili 0.01)
Statistika testa	$z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$ (pretpostav. da je σ nepoznata, inače σ umjesto s)
Kritične vrijednosti	Granica z_α koja obuhvata oblast α nadesno
Pravilo odlučivanja	Odbaciti nultu hipotezu ako $z > z_\alpha$

Kritične tačke *z*
(Jedno-strani test)

α	z_α
0.005	2.576
0.010	2.326
0.025	1.960
0.050	1.645
0.100	1.282

Desno-strani test



Primjer 8-8

EPA postavlja granice za koncentraciju polutanata koje emituju različite industrije. Pretpostavimo da je gornja dozvoljena granica emisije vinil hlorida postavljena na prosjek od 55 ppm. Da bi se provjerilo poštovanje ovog pravila, EPA sakuplja slučajan uzorak od 100 čitanja. Uzoračka prosječna koncentracija je 60 ppm i uzoračka standardna devijacija je 20 ppm. Da li se može zaključiti da se krši dato pravilo?

H_0: $\mu \leq 55$

H_1: $\mu > 55$

n = 100

Za $\alpha = 0.01$, kritična vrijednost z je 2.326

Statistika testa je: $z = \dfrac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$

Ne odbaciti H_0 ako: [$z \leq 2.326$]

Odbaciti H_0 ako: [$z > 2.326$]

$$n = 100$$
$$\bar{x} = 60$$
$$s = 20$$

$$z = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}} = \frac{60 - 55}{\frac{20}{\sqrt{100}}}$$

$$= \frac{5}{2} = 2.5 \Rightarrow \text{Odbaciti } H_0$$