

# 1. Statistika

**Primjer 1.** Pri mjerenju vremena reakcije na neki događaj psiholozi su utvrdili da je standardna devijacija za taj događaj 0.05 sekundi. Koliko velik uzorak moramo uzeti tako da s pouzdanošću

(a) 95%,

(b) 99%,

budemo sigurni da greška u procjeni neće prijeći 0.01?

Za pouzdanost 95% je  $k = 1.96$ , a za 99% je  $k = 2.58$ . Za proporcije vrijedi

$$p = \bar{p} \pm k \frac{\sigma}{\sqrt{n}}.$$

Želimo da za grešku procjene vrijedi

$$k \frac{\sigma}{\sqrt{n}} \leq 0.01.$$

Odatle za  $n$  izlazi

$$n \geq \left(k \frac{\sigma}{0.01}\right)^2 = \left(k \frac{0.05}{0.01}\right)^2 = 25k^2,$$

pa za zadane  $k$ -ove imamo

(a)  $n \geq 96.04$ , tj. mora biti  $n \geq 97$ ,

(b)  $n \geq 166.41$ , tj. mora biti  $n \geq 167$ .

**Primjer 2.** Pri mjerenju vremena reakcije na neki događaj psiholozi su testirali slučajni uzorak od 256 ljudi. Kolika najviše smije biti utvrđena standardna devijacija tog događaja, tako da s pouzdanošću

(a) 95%,

(b) 99%,

budemo sigurni da greška u procjeni neće prijeći 0.01?

Za pouzdanost 95% je  $k = 1.96$ , a za 99% je  $k = 2.58$ . Za proporcije vrijedi

$$p = \bar{p} \pm k \frac{\sigma}{\sqrt{n}}.$$

Želimo da za grešku procjene vrijedi

$$k \frac{\sigma}{\sqrt{n}} \leq 0.01.$$

Odatle dobivamo

$$\sigma \leq \frac{0.01 \cdot \sqrt{n}}{k} = \frac{0.16}{k},$$

pa uvrštavanjem  $k$ -ova izlazi

- (a)  $\sigma \leq 0.081633$ ,
- (b)  $\sigma \leq 0.062016$ .

**Primjer 3.** Slučajni uzorak od 50 kolokvija iz matematike, od ukupno 200 kolokvija, pokazao je očekivanje 75 i standardnu devijaciju 10.

- (a) Koji je 95% interval pouzdanosti za procjenu očekivanja svih 200 kolokvija?
- (b) Koji je stupanj pouzdanosti ako možemo reći da se sredina svih 200 zadaća nalazi u intervalu  $75 \pm 1$ ?

Riješite primjer na dva načina:

- ako pretpostavljamo da je populacija dovoljno velika prema uzorku,
- ako pretpostavljamo da je populacija konačna i uzorak relativno velik prema populaciji.

Za 95% pouzdanost,  $k = 1.96$ .

- (a) Ako je populacija ‘dovoljno velika’, onda imamo

$$p = \bar{p} \pm k \frac{\sigma}{\sqrt{n}} = 75 \pm 1.96 \frac{10}{\sqrt{50}} = 75 \pm 2.77185858,$$

a ako je uzorak velik prema populaciji, onda imamo

$$p = \bar{p} \pm k \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{n_p - n}{n_p - 1}} = 75 \pm 1.96 \frac{10}{\sqrt{50}} \cdot \sqrt{\frac{150}{199}} = 75 \pm 2.40652380.$$

- (b) Ako je populacija ‘dovoljno velika’, onda mora biti greška procjene

$$k \frac{\sigma}{\sqrt{n}} \leq 1,$$

tj. mora vrijediti

$$k \leq \frac{\sqrt{n}}{\sigma} = \frac{\sqrt{50}}{10} = 0.70710678.$$

Sad se prisjetimo kako iz  $k$  očitavamo pouzdanost iz tablica normalne slučajne varijable za  $\Phi_0$ . Za pouzdanost vrijedi da je  $2\Phi_0(k) = 2 \cdot 0.26114 = 0.52228$  ili 52.228%.

A ako je uzorak velik prema populaciji, onda imamo

$$k \frac{\sigma}{\sqrt{n}} \cdot \sqrt{\frac{n_p - n}{n_p - 1}} \leq 1,$$

pa slijedi

$$k \leq \frac{\sqrt{n}}{\sigma} \cdot \sqrt{\frac{n_p - 1}{n_p - n}} = \frac{\sqrt{50}}{10} \sqrt{\frac{199}{150}} = 0.81445278.$$

Ponovno je pouzdanost  $2\Phi_0(k) = 2 \cdot 0.29103 = 0.58206$  ili 58.206%.

**Primjer 4.** U nekom okrugu slučajni uzorak od 100 glasača pokazao je da kandidat A dobiva izbore s 57% glasova. Nađite

- (a) 95%,
- (b) 99%,
- (c) 99.73%

pouzdanost za postotak glasova svih glasača u tom okrugu za kandidata A.

Riješimo zadatak (c). Za pouzdanost 99.73%,  $k = 3$ .

- (c) Prvo moramo procijeniti  $\sigma$ .

$$\sigma \approx s = \sqrt{\bar{p}(1 - \bar{p})} = \sqrt{0.57 \cdot 0.43} = 0.49507575.$$

$$\begin{aligned} p &= \bar{p} \pm k \frac{\sigma}{\sqrt{n}} \\ &= 57 \pm 3 \frac{0.49507575}{\sqrt{100}} = 57 \pm 0.148522725. \end{aligned}$$

**Primjer 5.** Koliko veliki uzorak glasača moramo uzeti u prethodnom primjeru, tako da s pouzdanošću

- (a) 95%,
- (b) 99%,
- (c) 99.73%

budemo sigurni da je kandidat izabran?

Riješimo zadatak (c). Za pouzdanost 99.73%,  $k = 3$ .

- (c) Da bi kandidat bio izabran mora biti  $p > 50\%$ . Prema tome, greška procjene ne smije biti veća od 7%, odnosno 0.07. Dakle, imamo

$$k \frac{\sigma}{\sqrt{n}} \leq 0.07,$$

pa je

$$n \geq \left( k \frac{\sigma}{0.07} \right)^2 = \left( 3 \frac{0.49507575}{0.07} \right)^2 = 450.18.$$

Dakle, moramo uzeti uzorak od 451 glasača.

**Primjer 6.** Vjeruje se da je u nekom okrugu da je razlika broj glasova za jednog, odnosno drugog kandidata vrlo malena. Koliko velik slučajni uzorak moramo odabrati da bismo s 80% pouzdanošću bili sigurni u pobjedu jednog od kandidata.

Za 80% pouzdanost,  $k = 1.28$ . Budući da kandidati dobivaju približno jednak broj glasova, onda je  $\bar{p} = 0.5$ ,

$$\sigma \approx s = \sqrt{\bar{p}(1 - \bar{p})} = 0.5.$$

Greška koju smo napravili je

$$1.28 \frac{0.5}{\sqrt{n}}$$

i ona mora biti manja od pola posto, tj. mora biti

$$1.28 \frac{0.5}{\sqrt{n}} \leq 0.005.$$

Odatle dobivamo

$$n \geq \left( \frac{1.28 \cdot 0.5}{0.005} \right)^2 = 16384.$$