

Limitni izreki

Verjetnost in statistika, študijsko leto 2018/19

doc. dr. Boštjan Kuzman

Univerza v Ljubljani, Pedagoška fakulteta

1. Kaj so limitni izreki
2. Neenakosti Markova in Čebiševa
3. Šibki zakon velikih števil
4. Centralni limitni izrek
5. Še nekaj računskih zgledov

Kaj so limitni izreki

S skupnim imenom **limitni izreki** poimenujemo različne izreke, ki opisujejo obnašanje neskončnega zaporedja slučajnih spremenljivk.

Najpomembnejše teoretične rezultate teorije verjetnosti predstavljata dva sklopa limitnih izrekov.

1. **Zakoni velikih števil** opisujejo, pod kakšnimi pogoji povprečje zaporedja slučajnih spremenljivk konvergira proti pričakovani vrednosti.
2. **Centralni limitni izreki** opisujejo, pod kakšnimi pogoji vsota zaporedja slučajnih spremenljivk konvergira proti normalni porazdelitvi.

V dokazovanju limitnih izrekov se uporabljajo različne neenakosti (npr. trikotniška, Markova, Čebiševa, Jensenova...), aditivnost pričakovane vrednosti $E(\sum X_i) = \sum E(X_i)$ in za neodvisne slučajne spremenljivke tudi aditivnost variance $\text{var}(\sum X_i) = \sum \text{var}(X_i)$.

Neenakosti Markova in Čebiševa

Izrek (Neenakost Markova)

Naj bo X slučajna spremenljivka, ki zavzame samo nenegativne vrednosti.

Potem za poljubno pozitivno število $a > 0$ velja

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Izrek (Neenakost Markova)

Naj bo X slučajna spremenljivka, ki zavzame samo nenegativne vrednosti. Potem za poljubno pozitivno število $a > 0$ velja

$$P(X \geq a) \leq \frac{E(X)}{a}.$$

Dokaz.

Naj bo $X \geq 0$ in $a > 0$. Potem je $X/a \geq 0$. Definiramo slučajno spremenljivko

$$I = \begin{cases} 1, & \text{če } 1 \leq X/a, \\ 0, & \text{če } 0 \leq X/a < 1. \end{cases}$$

Potem velja $I \leq X/a$ in zato $E(I) \leq E(X/a)$. Sledi

$$E(I) = 1 \cdot P(I = 1) + 0 \cdot P(I = 0) = P(I = 1) = P(X \geq a) \leq E(X)/a.$$



Izrek (Neenakost Čebiševa)

Naj bo X slučajna spremenljivka, tako da sta $\mu = E(X)$ in $\sigma^2 = \text{var}(X)$ končni vrednosti. Potem za vsako pozitivno število $k > 0$ velja

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$

Izrek (Neenakost Čebiševa)

Naj bo X slučajna spremenljivka, tako da sta $\mu = E(X)$ in $\sigma^2 = \text{var}(X)$ končni vrednosti. Potem za vsako pozitivno število $k > 0$ velja

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$

Dokaz.

Ker je $(X - \mu)^2 \geq 0$, po neenakosti Markova za $a = k^2 > 0$ dobimo

$$P(|X - \mu| \geq k) = P((X - \mu)^2 \geq k^2) \leq \frac{E((X - \mu)^2)}{k^2} = \frac{\sigma^2}{k^2}.$$

□

Neenakosti Markova in Čebiševa omogočata grobo oceno verjetnosti v primeru, ko o slučajni spremenljivki X poznamo le $E(X)$ in morda še $\text{var}(X)$.

Zgled

Rezultati mature so normalno porazdeljeni s povprečno oceno $\mu = 18$ točk in standardnim odklonom $\sigma = 3$. Izberemo naključnega kandidata. Kakšna je verjetnost, da je njegov rezultat X vsaj 24 točk?

- Neenakost Markova da oceno

$$P(X \geq 24) \leq \frac{\mu}{24} = \frac{18}{24} = 75\%$$

- S pomočjo neenakosti Čebiševa lahko dobimo oceno

$$P(X \geq 24) = P(X - 18 \geq 6) \leq P(|X - \mu| \geq 6) \leq \frac{3^2}{6^2} = 25\%$$

- Pravilo 68-95-99.7 za normalno porazdelitev da uporaben približek

$$P(X \geq 24) = P(X - 18 \geq 6) = P(|X - \mu| \geq 2\sigma)/2 \doteq 2.5\%$$

Neenakosti Markova in Čebiševa sta zelo splošni, saj veljata za poljubne slučajne spremenljivke, o katerih vemo zelo malo.

Prav zato nista dovolj natančni za računanje verjetnosti v konkretnih primerih.

Izjemno pomembni pa sta pri dokazovanju različnih limitnih izrekov.

Šibki zakon velikih števil

- Že **Gerolamo Cardano** je v 16. stoletju brez dokaza zapisal, da se natančnost statističnega povprečja izboljšuje s številom poskusov.
- **Jacob Bernoulli** je več kot 20 let porabil za dokaz izreka, ki ga je imenoval *Zlati zakon*, in je leta 1713 izšel v knjigi *Ars Conjectandi*, objavljeni 8 let po njegovi smrti.
- **Simeone Poisson** je leta 1837 v svoji knjigi Bernoullijev izrek poimenoval *Zakon velikih števil*.
- Nadaljne izboljšave izreka s šibkejšimi predpostavkami so v drugi polovici 19. in prvi polovici 20. stoletja prispevali še Čebišev, Markov, Borel, Cantelli, Kolmogorov in Khinchin.
- Danes običajno govorimo o dveh oblikah, ki ju imenujemo **Šibki** oziroma **Krepki zakon velikih števil** in govorita o različnem načinu konvergence povprečja zaporedja slučajnih spremenljivk k pričakovani vrednosti.

Izrek (Šibki zakon velikih števil)

Naj bodo slučajne spremenljivke X_1, X_2, \dots neodvisne in enako porazdeljene, tako da je $\mu = E(X_i) < \infty$ in $\sigma^2 = \text{var}(X_i) < \infty$.

Potem za vsak $\epsilon > 0$ velja

$$\lim_{n \rightarrow \infty} P \left(\left| \frac{X_1 + X_2 + \dots + X_n}{n} - \mu \right| \geq \epsilon \right) = 0.$$

Povedano drugače, verjetnost, da povprečna vrednost (neodvisnih in enako porazdeljenih) slučajnih spremenljivk odstopa od pričakovane vrednosti μ za več kot ϵ , z velikim številom spremenljivk konvergira proti 0.

Dokaz.

Povprečje prvih n sl. spremenljivk označimo z $\bar{X} = \frac{X_1 + \dots + X_n}{n}$. Potem je

$$E(\bar{X}) = E\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{E(X_1) + \dots + E(X_n)}{n} = \frac{n\mu}{n} = \mu,$$

in ker so X_i neodvisne, tudi

$$\text{var}(\bar{X}) = \text{var}\left(\frac{X_1 + \dots + X_n}{n}\right) = \frac{\text{var}(X_1) + \dots + \text{var}(X_n)}{n^2} = \frac{n\sigma^2}{n^2} = \sigma^2/n.$$

Po Čebiševu za slučajno spremenljivko \bar{X} sledi

$$P(|\bar{X} - \mu| \geq \epsilon) \leq \frac{\sigma^2/n}{\epsilon^2} = \frac{\sigma^2/\epsilon^2}{n}.$$

Izraz na desni očitno konvergira proti 0, ko gre n proti ∞ .



Bernoulli je obravnaval problem vlečenja naključne kroglice iz žare z r rdečimi in b belimi kroglicami brez vračanja.

Pri tem je ocenjeval verjetnost, da se pri nekem številu ponovitev poskusa povprečno število izvlečenih rdečih kroglic od pričakovane vrednosti $\frac{r}{r+b}$ razlikuje za več kot $\epsilon = \frac{1}{r+b}$.

Bernoulli je znal dokazati, da gre ta verjetnost proti 0, ko gre število ponovitev poskusa proti neskončno, ni pa znal odgovoriti, koliko ponovitev zadošča, da bo verjetnost zagotovo manjša od vnaprej izbrane vrednosti. Morda je tudi zato odlašal z objavo svoje knjige vse do smrti.

Dokaz Šibkega ZVŠ z uporabo neenakosti Čebiševa je veliko krajši in elegantnejši od izvirnega Bernoullijevega dokaza.

Za primerjavo skupaj navedimo še Šibki in Krepki ZVŠ (brez dokaza).

Izrek (Šibki in Krepki ZVŠ)

Naj bodo slučajne spremenljivke X_1, X_2, \dots neodvisne in enako porazdeljene, tako da je $\mu = E(X_i) < \infty$. Potem za povprečje $\bar{X} = \sum_{i=1}^n X_i$ velja

- Šibki ZVŠ:

$$\forall \epsilon > 0: \lim_{n \rightarrow \infty} P(|\bar{X} - \mu| \geq \epsilon) = 0.$$

- Krepki ZVŠ:

$$P(\lim_{n \rightarrow \infty} \bar{X} = \mu) = 1.$$

Centralni limitni izrek

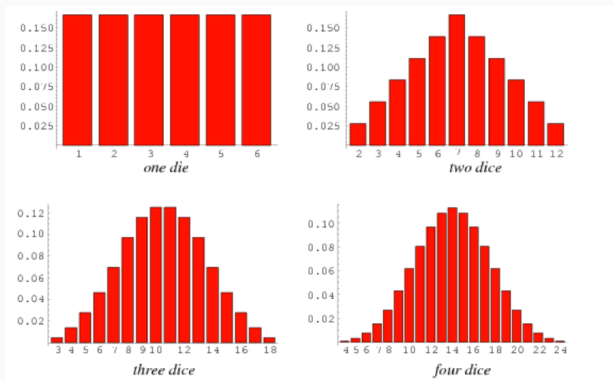
Centralni limitni izrek (CLI) je izjemno pomemben izrek teorije verjetnosti. Pravzaprav gre za skupino podobnih izrekov z različno strogimi predpostavkami.

Ohlapno povedano, CLI pravi, da je vsota velikega števila neodvisnih, enako porazdeljenih slučajnih spremenljivk približno normalno porazdeljena.

CLI omogoča preprosto aproksimacijo verjetnosti za vsote takih spremenljivk in pojasni, zakaj se v naravi tako pogosto pojavlja normalna porazdelitev (oziroma histogram v obliki Gaussove krivulje).

CLI je tudi osnova za testiranje hipotez v statistiki. Z njim namreč ocenimo verjetnost, da izmerjeni podatki potrjujejo pravilnost hipoteze.

Vsota pik na n kockah ustreza vsoti $X_1 + \dots + X_n$, kjer so diskretne slučajne spremenljivke X_i enakomerno porazdeljene na $\{1, 2, 3, 4, 5, 6\}$.



Centralni limitni izrek pove, da postaja porazdelitev vsote pik z večanjem števila kock vse bolj podobna normalni porazdelitvi.

(Vir slike: spletna stran mathworld.wolfram.com.)

Izrek (Centralni limitni izrek, standardna verzija)

Naj bodo X_1, X_2, \dots neodvisne, enako porazdeljene slučajne spremenljivke s pričakovano vrednostjo $\mu = E(X_i)$ in varianco $\sigma^2 = \text{var}(E_i)$, in naj bo $S_n = X_1 + \dots + X_n$ njihova vsota, torej je $E(S_n) = n\mu$ in $\text{var}(S_n) = n\sigma^2$. Potem je normalizirana vsota $\frac{S_n - n\mu}{\sigma\sqrt{n}}$ porazdeljena približno standardno normalno:

$$P\left(a \leq \frac{S_n - n\mu}{\sigma\sqrt{n}} \leq b\right) \approx \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx = \Phi(b) - \Phi(a),$$

kjer je Φ porazdelitvena funkcija standardne normalne spremenljivke.

Dokaz je zahteven in ga ne bomo navedli. Obstajajo številne verzije CLI z različno močnimi predpostavkami in zaključki (npr. natančnejši opis konvergence).

Zgled (Astronomske meritve (vir: S. Ross))

Astronom meri razdaljo do neke zvezde. Meritve motijo vremenski pojavi, atmosferski vplivi, sevanja in podobno, zato namerava astronom izvesti večje število meritev in izračunati njihovo povprečje. Predpostavimo, da so meritve neodvisne in enako porazdeljene slučajne spremenljivke s pričakovano vrednostjo d in standardnim odklonom 2 svetlobni leti.

Zanima nas, koliko meritev je potrebnih, da bo verjetnost, da se izračunano povprečje za več kot 0.5 svetlobnega leta razlikuje od dejanske razdalje, manjša od 5%.

Naj bodo d dejanska razdalja in X_1, X_2, \dots vrednosti meritev.

Iščemo torej najmanjše celo število n , za katerega je

$$P(-0.5 < \frac{\sum X_i}{n} - d < 0.5) > 95\%.$$

Neenakosti v oklepaju pomnožimo z $\sqrt{n}/2$ in po preurejanju dobimo

$$P\left(-\frac{\sqrt{n}}{4} < \frac{\sum X_i - nd}{2\sqrt{n}} < \frac{\sqrt{n}}{4}\right) > 95\%.$$

Po CLI je leva stran približno enaka

$$\Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right) = \Phi\left(\frac{\sqrt{n}}{4}\right) - (1 - \Phi\left(\frac{\sqrt{n}}{4}\right)) = 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1.$$

Iščemo torej najmanjši n , da bo $\Phi\left(\frac{\sqrt{n}}{4}\right) > 0.975$ oziroma $\frac{\sqrt{n}}{4} > 1.96$, kot lahko odčitamo iz tabele za Φ .

Sledi $n > 61.47$, torej je potrebnih vsaj 62 meritev.

Normalna aproksimacija binomske porazdelitve je poseben primer CLI.

Izrek (De Moivre - Laplaceov limitni izrek)

Če je S_n število uspehov v n ponovitvah poskusa z verjetnostjo uspeha p (torej, $S_n \sim \text{Bin}(n, p)$), potem za vse $a < b$ in dovolj velike n velja

$$P\left(a \leq \frac{S_n - np}{\sqrt{npq}} \leq b\right) \approx \Phi(b) - \Phi(a).$$

Dokaz.

Binomska slučajna spremenljivka $S_n = X_1 + \dots + X_n$ je vsota Bernoullijevih slučajnih spremenljivk z verjetnostjo uspeha p , za katere velja $\mu = E(X_i) = p$ in $\sigma^2 = \text{var}(X_i) = pq$. Aproksimacija zdaj sledi neposredno iz CLI. \square

- Pri obravnavi velikega števila metov poštenega kovanca je preprosto verzijo CLI odkril in dokazal že De Moivre okoli 1730.
- Laplace je okoli 1810 De Moivrejev rezultat posplošil na poljuben p . To verzijo izreka danes imenujemo normalna aproksimacija binomske porazdelitve.
- Laplace je odkril tudi verzijo CLI, ki jo danes imenujemo standardna, vendar je ni znal pravilno dokazati.
- Korekten dokaz je prvi podal Ljapunov okoli 1902.

Še nekaj računskih zgledov

Zgled

*Vržemo 10 poštenih kock. Z uporabo CLI ocenite verjetnost, da bo vsota pik med vključno 30 in 40.*¹

¹Vsi zgledi v razdelku so povzeti po knjigi S. Ross: A First Course in Probability

Zgled

Vržemo 10 poštenih kock. Z uporabo CLI ocenite verjetnost, da bo vsota pik med vključno 30 in 40.¹

REŠITEV: Naj bodo X_1, \dots, X_{10} števila pik na vsaki kocki.

Potem je $E(X_i) = 3.5$ in $\text{var}(X_i) = 35/12$ za vse i . Za slučajno spremenljivko $X = X_1 + \dots + X_{10}$ velja $E(X) = 35$ in $\text{var}(X) = 350/12$ zaradi neodvisnosti. Po CLI sledi

$$\begin{aligned} P(29.5 \leq X \leq 40.5) &= P\left(\frac{29.5 - 35}{\sqrt{350/12}} \leq \frac{X - 35}{\sqrt{350/12}} \leq \frac{40.5 - 35}{\sqrt{350/12}}\right) \\ &\approx 2\Phi\left(\frac{5.5}{\sqrt{350/12}}\right) - 1 \\ &\approx 2\Phi(1.0184) - 1 \\ &\approx 0.692. \end{aligned}$$

¹Vsi zgledi v razdelku so povzeti po knjigi S. Ross: A First Course in Probability

Zgled

Kocko mečemo, dokler vsota padlih pik ne preseže 300. Z uporabo CLI ocenite verjetnost, da moramo vreči vsaj 80-krat?

Zgled

Kocko mečemo, dokler vsota padlih pik ne preseže 300. Z uporabo CLI ocenite verjetnost, da moramo vreči vsaj 80-krat?

R: Naj bo X_i vrednost v i -tem metu in $S_n = \sum X_i$. Potem je $E(X_i) = 7/2$ in $\text{var}(X_i) = 35/12$. Sledi

$$\begin{aligned} P\left(\sum_{i=1}^{79} X_i \leq 300\right) &= P\left(\sum_{i=1}^{79} X_i \leq 300.5\right) \\ &\approx P\left(Z \leq \frac{300.5 - 79 \cdot 7/2}{\sqrt{79 \cdot 35/12}}\right) \\ &= P(Z \leq 1.58) = 94.29\%, \end{aligned}$$

kjer je $Z \sim N(0, 1)$ in smo uporabili tabelo za Φ .

Zgled

Povprečni rezultat NPZ iz matematike je 75 točk, standardni odklon pa 5 točk.

- 1. Z neenakostjo Markova podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel vsaj 85 točk.*

Zgled

Povprečni rezultat NPZ iz matematike je 75 točk, standardni odklon pa 5 točk.

- 1. Z neenakostjo Markova podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel vsaj 85 točk.*

$$R: P(X \geq 85) \leq E(X)/85 = 15/17 \doteq 88.2\%$$

Zgled

Povprečni rezultat NPZ iz matematike je 75 točk, standardni odklon pa 5 točk.

- 1. Z neenakostjo Markova podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel vsaj 85 točk.*

$$R: P(X \geq 85) \leq E(X)/85 = 15/17 \doteq 88.2\%$$

- 2. Z neenakostjo Čebiševa podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel med vključno 65 in 85 točk.*

Zgled

Povprečni rezultat NPZ iz matematike je 75 točk, standardni odklon pa 5 točk.

1. Z neenakostjo Markova podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel vsaj 85 točk.

$$R: P(X \geq 85) \leq E(X)/85 = 15/17 \doteq 88.2\%$$

2. Z neenakostjo Čebiševa podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel med vključno 65 in 85 točk.

$$R: P(65 \leq X \leq 85) = 1 - P(|X - 75| > 10) \geq 1 - 25/100 = 75\%$$

Zgled

Povprečni rezultat NPZ iz matematike je 75 točk, standardni odklon pa 5 točk.

1. Z neenakostjo Markova podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel vsaj 85 točk.

$$R: P(X \geq 85) \leq E(X)/85 = 15/17 \doteq 88.2\%$$

2. Z neenakostjo Čebiševa podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel med vključno 65 in 85 točk.

$$R: P(65 \leq X \leq 85) = 1 - P(|X - 75| > 10) \geq 1 - 25/100 = 75\%$$

3. Na dva načina (Čebišev, CLI) ocenite, koliko naključnih študentov bi morali izbrati, da bi bila verjetnost, da je povprečje njihovih rezultatov med vključno 70 in 80 točk, vsaj 90%.

Zgled

Povprečni rezultat NPZ iz matematike je 75 točk, standardni odklon pa 5 točk.

1. Z neenakostjo Markova podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel vsaj 85 točk.

$$R: P(X \geq 85) \leq E(X)/85 = 15/17 \doteq 88.2\%$$

2. Z neenakostjo Čebiševa podajte zgornjo mejo za verjetnost, da je naključno izbrani kandidat dosegel med vključno 65 in 85 točk.

$$R: P(65 \leq X \leq 85) = 1 - P(|X - 75| > 10) \geq 1 - 25/100 = 75\%$$

3. Na dva načina (Čebišev, CLI) ocenite, koliko naključnih študentov bi morali izbrati, da bi bila verjetnost, da je povprečje njihovih rezultatov med vključno 70 in 80 točk, vsaj 90%.

$$R1 \text{ (Čebišev): } P(|\sum X_i/n - 75| > 5) \leq \frac{25}{25n} = 1/n < 10\%, \text{ torej } n \geq 10.$$

$$R2 \text{ (CLI): } P(|\sum X_i/n - 75| > 5) \approx P(|Z| > \sqrt{n}) < 10\%, \text{ kjer je } Z \sim N(0, 1), \text{ z uporabo tabele za } \Phi \text{ sledi } n \geq 3.$$

Zgled

Povprečni rezultat standardiziranega preizkusa znanja je $\mu = 74$ točk, s standardnim odklonom $\sigma = 14$.

1. Preizkus piše 25 učencev. Ocenite verjetnost, da bo povprečni rezultat v razredu višji od 80 točk.

Zgled

Povprečni rezultat standardiziranega preizkusa znanja je $\mu = 74$ točk, s standardnim odklonom $\sigma = 14$.

1. Preizkus piše 25 učencev. Ocenite verjetnost, da bo povprečni rezultat v razredu višji od 80 točk.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{5}} > 15/\sqrt{5}\right) \approx P(Z > 2.14) \approx 1.6\%.$$

Zgled

Povprečni rezultat standardiziranega preizkusa znanja je $\mu = 74$ točk, s standardnim odklonom $\sigma = 14$.

1. Preizkus piše 25 učencev. Ocenite verjetnost, da bo povprečni rezultat v razredu višji od 80 točk.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{5}} > 15/\sqrt{5}\right) \approx P(Z > 2.14) \approx 1.6\%.$$

2. Ocenite podobno verjetnost še za razred z 64 učenci.

Zgled

Povprečni rezultat standardiziranega preizkusa znanja je $\mu = 74$ točk, s standardnim odklonom $\sigma = 14$.

1. Preizkus piše 25 učencev. Ocenite verjetnost, da bo povprečni rezultat v razredu višji od 80 točk.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{5}} > 15/\sqrt{5}\right) \approx P(Z > 2.14) \approx 1.6\%.$$

2. Ocenite podobno verjetnost še za razred z 64 učenci.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{8}} > 24/\sqrt{8}\right) \approx P(Z > 3.43) \approx 0.03\%.$$

(Verjetnost se z večjim številom učencev zmanjša!)

Zgled

Povprečni rezultat standardiziranega preizkusa znanja je $\mu = 74$ točk, s standardnim odklonom $\sigma = 14$.

1. Preizkus piše 25 učencev. Ocenite verjetnost, da bo povprečni rezultat v razredu višji od 80 točk.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{5}} > 15/\sqrt{5}\right) \approx P(Z > 2.14) \approx 1.6\%.$$

2. Ocenite podobno verjetnost še za razred z 64 učenci.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{8}} > 24/\sqrt{8}\right) \approx P(Z > 3.43) \approx 0.03\%.$$

(Verjetnost se z večjim številom učencev zmanjša!)

3. Ocenite verjetnost, da bo povprečna ocena v razredu z 25 učenci za vsaj 2.2 točki višja od povprečne ocene v razredu s 64 učenci.

Zgled

Povprečni rezultat standardiziranega preizkusa znanja je $\mu = 74$ točk, s standardnim odklonom $\sigma = 14$.

1. Preizkus piše 25 učencev. Ocenite verjetnost, da bo povprečni rezultat v razredu višji od 80 točk.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{5}} > 15/\sqrt{5}\right) \approx P(Z > 2.14) \approx 1.6\%.$$

2. Ocenite podobno verjetnost še za razred z 64 učenci.

$$R: P(\bar{X} > 80) = P\left(\frac{\bar{X}-74}{14/\sqrt{8}} > 24/\sqrt{8}\right) \approx P(Z > 3.43) \approx 0.03\%.$$

(Verjetnost se z večjim številom učencev zmanjša!)

3. Ocenite verjetnost, da bo povprečna ocena v razredu z 25 učenci za vsaj 2.2 točki višja od povprečne ocene v razredu s 64 učenci.

R: Zaradi neodvisnosti je

$$\text{var}(\bar{X} - \bar{Y}) = \text{var}(\bar{X}) + (-1)^2 \text{var}(\bar{Y}) = 196/64 + 196/25 \doteq 3.3^2. \text{ Sledi}$$

$$P(\bar{X} - \bar{Y} > 2.2) = P\left(\frac{\bar{X} - \bar{Y}}{3.3} > \frac{2.2}{3.3}\right) = P(Z > 0.67) \approx 25.14\%.$$