

Seminarska naloga 1 - simulacije

Navodilo za oddajo

Opis problema, komentarje in rešitve zanj zapišite v datoteko `sem1_priimek.Rmd`. Prek spletne učilnice oddajte datoteko `sem1_priimek.html`, ki naj vsebuje vse komentarje, rešitve, grafe in vso kodo v R, ki ste jo za reševanje problema uporabili.

Celotna seminarska naloga je namenoma malce ohlapneje definirana.

Naloga naj bo smiselno in pregledno urejena, celota naj predstavlja zaključen izdelek.

Navodilo naloge

S pomočjo R in njenih vgrajenih porazdelitev bi radi simulirali izračun (oceno) konstante e (na 4 decimalke natančno je to 2.7183).

Simulirali ga bomo na dva načina:

- a. Vemo, da je verjetnost, da iz končne populacije n enot izberemo vrednost, ki ni a : $P(X \neq a) = (1 - \frac{1}{n})$. Hkrati pa vemo tudi

$$\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n}\right)^n = \frac{1}{e}.$$

- b. Vsota kvadratov dveh neodvisnih standardno normalno porazdeljenih spremenljivk je porazdeljena po eksponentni porazdelitvi s parametrom $\frac{1}{2}$. Torej:

$$\begin{aligned} X_1, X_2 &\sim_{iid} N(0, 1) \\ Y &= X_1^2 + X_2^2 \\ Y &\sim \text{Exp}\left(\lambda = \frac{1}{2}\right) \end{aligned}$$

Za eksponentno porazdeljeno spremenljivko s parametrom $\lambda = \frac{1}{2}$ pa velja

$$P(Y \geq y) = 1 - \left(1 - e^{-(y/2)}\right).$$

Za $y = 2$, je kumulativna verjetnost enaka $\frac{1}{e}$.

1. Zapišite, kako se boste lotili reševanja problema in vse skupaj podkrepite z grafičnim(i) prikazom(i).
2. Definirajte funkciji `ocenaA` in `ocenaB` za oceno konstante (\hat{e}) po načinu a. in b. Obvezni vhodni argument funkcije naj bo število ponovitev n , ki jih boste izvedli, da boste dobili \hat{e} .
3. Simulirajte najmanj 1000 ponovitev ocenjevanja konstante za 5 zelo različnih n -jev (določite jih smiselno na podlagi spodnjih alinej!). Za vsak n izračunajte in grafično prikažite porazdelitev izračunanih ocen.
 - V pregledno urejeni tabeli predstavite glavne opisne statistike za vsak posamezen n in vsakega izmed načinov.
 - Izračunajte intervale zaupanja za e (lahko so vključeni tudi v tabelo) in jih prikažite na enem grafu (razmislite tudi, kakšni *intervali zaupanja* so smiselni glede na obliko porazdelitve ocen).
 - Utemeljite spremembo širine intervalov zaupanja. Od česa je širina intervalov odvisna in kako se to kaže na vašem vzorcu. (V utemeljitev lahko dodate tudi formule.)