

MC meetodid, praktikum 5

Teema 1: Lihtsa juhusliku suuruse genereerimine

Algoritm:

Vaatleme juhuslikku suurust, mille võimalikeks väärtusteks on $c_1 < c_2 < \dots < c_m$ olgu $p_1, p_2, \dots, p_m$ vastavad tõenäosused

1. Tekime ühtlase jaotusega pseudojuhuslikud arvud $u_1, \dots, u_n$
2. Kui $u_i \in (0, p_1]$, siis $x_i = c_1$,
kui $u_i \in (p_1, p_1 + p_2]$, siis $x_i = c_2$
kui $u_i \in (p_1 + p_2, p_1 + p_2 + p_3]$, siis $x_i = c_3$ jne

Harjutus 1

Vaatleme diskreetset juhuslikku suurust, mille jaotustabel on järgmine:

c_i	1	3	4.5	7
p_i	0.3	0.2	0.4	0.1

Tekita 1000 pseudojuhuslikku arvu sellest jaotusest, lähtudes käsuga `runif` genereeritud pseudojuhuslikest arvudest vahemikus $(0, 1)$. Kontrolli tulemuste vastavust soovitud jaotusele χ^2 -testiga

Vihjed Kasuks tulevad järgmised käsud: `cumsum`, `cut`, `as.numeric` ja teadmine, et vektori indeksisse võib kirjutada täisarvude jada, nt vektorite `a` ja `b=c(1,1,3,2)` korral `a[b]` annab tulemuseks vektori väärtustega `a[1]`, `a[1]`, `a[3]`, `a[2]`

Harjutus 2

Tegelikult on R-is mugav vahend selliste ülesannete lahendamiseks, nimelt käsk `sample`. Sõltumatu valimi saamiseks tuleb kasutada suvandit `replace = TRUE`. Vaata käsu spikrit (nt käsuga `?sample`). Leia ligikaudselt tõenäosus

$$P\left(\sum_{i=1}^5 X_i^2 < 100\right),$$

kus $X_i, i = 1, \dots, 5$ on eelmises ülesandes vaadeldud jaotusega sõltumatud juhuslikud suurused. Genereeri selleks 1000 summa väärtust, kusjuures vajaminevad pseudojuhuslikud arvud tekita `sample` käsku kasutades.

Teema 2: Juhusliku suuruse väärtuste genereerimine üldistatud pöördfunktsiooni kasutades.

Defineerime üldistatud pöördfunktsiooni valemiga

$$F^-(y) = \inf\{a : F(a) \geq y\}, \quad y \in (0, 1)$$

Teoreem: Olgu F jaotusfunktsioon ning olgu U ühtlase jaotusega $U(0, 1)$ juhuslik suurus. Siis juhusliku suuruse $X = F^-(U)$ jaotusfunktsiooniks on F

Harjutus 3

Kasuta üldistatud pöördfunktsiooni selleks, et käsuga `runif(1000)` genereeritud jaotusega $U(0, 1)$ pseudojuhuslikust arvust saada jaotusfunktsioonile

$$F(x) = \begin{cases} 0, & x < 1 \\ \frac{x^2-1}{4}, & 1 \leq x < 2 \\ 1, & x \geq 2 \end{cases}$$

vastavad pseudojuhuslikud arvud. Leia nende abil ligikaudselt vastava juhusliku suuruse keskväärts.

Teema 3: Valikumeetod etteantud tihedusfunktsiooniga pideva juhusliku suuruse väärtuste genereerimiseks.

Valikumeetodi aluseks olev tulemus:

Teoreem Eeldame, et tihedusfunktsioon f on 0 väljaspool lõiku $[a, b]$ ning et f on ülalt tõkestatud konstandiga c , st $f(x) \leq c \forall x$. Olgu $Y_1 \sim U(a, b)$ ja $Y_2 \sim U(0, c)$ sõltumatud juhusliud suurused. Defineerime juhusliku suuruse X järgmiselt:

$X = Y_1$ tingimusel, et $Y_2 \leq f(Y_1)$

Siis X on pidev juhuslik suurus tihedusfunktsiooniga f .

Algoritm lähtuvalt $U(0, 1)$ pseudojuhuslikest arvudest

- genereerime n paari (u_{i1}, u_{i2}) , $i = 1, \dots, n$ (vastavad sõltumatutele arvudele jaotusest $U(0, 1)$)
- arvutame $y_{i1} = a + (b - a) \cdot u_{i1}$, $y_{i2} = c \cdot u_{i2}$
- väljastame need y_{i1} väärtused, mille korral $y_{i2} \leq f(y_{i1})$.

Harjutus 4

Kasuta 8000 pseudojuhuslikku arvu jaotusest $U(0, 1)$, et saada valikumeetodi abil arve tihedusfunktsioonile

$$f(x) = \begin{cases} \frac{3}{4}\sqrt{2+x}, & -2 \leq x < -1, \\ \frac{3}{4}\sqrt{-x}, & -1 \leq x < 0, \\ 0 & \text{mujal} \end{cases}$$

vastavast jaotusest. Leia saadud arvude abil ligikaudselt väärtus, millest juhuslik suurus X on suurem tõenäosusega 0.05.

Vihje: Viimasele küsimusele vastuse saamisel võib abiks olla funktsioon `quintile`.

Tegelikult võib valikumeetodit rakendada ka siis, kui tihedusfunktsiooni asemel on antud sellega proportsionaalne (ehk mingi positiivse arvuga korrutatud) funktsioon.

Harjutus 5

Kasuta valikumeetodit juhul, kus tihedusfunktsiooni rollis on kasutatud funktsiooni

$$f(x) = \begin{cases} 2x, & 0 \leq x < 1, \\ 3 - x, & 1 \leq x < 3, \\ 0 & \text{mujal} \end{cases}$$

selleks, et lähtuvalt 4000 jaotusele $U(0, 1)$ vastavast pseudojuhuslikust arvust tekitada funktsioonile f vastavaid pseudojuhuslikke arve. Võrdle nende abil tekitatud histogrammi funktsioonile f vastava tihedusfunktsiooniga. Kas on usutav, et genereeritud arvud on õige jaotusega?