Examen Statistică 11 Feb. 2023

1) X var. aleatoare cu funcția de repartiție F bijectivă

a) Cum putem genera cu ajutorul met. respingerii obs. din repartiția lui X condiționată la $X > a$? Ce se întâmplă at. când a este mare? Scrieți un cod R care să permită simularea unei obs. din repartiția lui X condiționată la $X > a$, unde var. al. X este rep. $Exp(1)$ iar $a = 4$

b) $U \sim \mathcal{U}[0,1]$ și T def. prin $T = F^{-1}(F(a) + (1 - F(a))U)$
Det. fct. de rep. a lui T și găsiți o met. de simulare a repartiției lui X condiționată la $X > a$. Care dintre cele 2 met. este de preferat?

c) Scrieți un cod R care să permită simularea unui eșantion de volum n din repartiția lui X condiționată la $X > a$, unde variabila aleatoare X este repartizată Laplace $(2,3)$ iar $a > 0$ i.e. $X \sim f$ unde
$$f(x) = \frac{1}{6} e^{-\frac{|x-2|}{3}} \mathbb{1}_{\mathbb{R}}(x).$$

2) Pt. $\theta \in \mathbb{R}$ definim fct. $f_\theta(x) = \begin{cases} 1-\theta & \text{dc. } -\frac{1}{2} < x \leq 0 \\ 1+\theta & \text{dc. } 0 < x \leq \frac{1}{2} \\ 0 & \text{altfel} \end{cases}$

a) Ce condiții tb. să verifice θ a.î. f_θ să fie o densitate de repartiție în raport cu măsura Lebesgue pe $\mathbb{R}$?

b) Pt. $\theta = \frac{1}{2}$ dorim să generăm un eșantion de volum n din pop. f_θ. Descrieți procedura și scrieți un cod R care să permită acest lucru. Pt. $n = 2000$ scrieți un cod R care să compare vizual repartiția eșantionului generat cu rep. teoretică.

c) Fie $X_1, \dots X_n$ un eșantion de volum n din populația f_θ. Verificați dc. estimatorul de verosimilitate maximă $\hat{\theta}_n$ a lui θ este bine definit.

d) Sub condițiile de la pct a) este $\hat{\theta}_n$ nedeplasat? consistent? asimp. normal

3) Fie $\theta > -1$ un parametru necunoscut și $X_1, \ldots, X_n$ un eșantion de vol. n din pop. f_θ cu $f_\theta(x) = a x^\theta \mathbb{1}_{(0,1)}(x)$

a) Det $a = ?$. Pt. $\theta = 2$ dorim să generăm 3 val. aleatoare din repartiția lui $X \sim f_\theta(x)$. Pt. aceasta dispunem de 3 val. rezultate din repartiția uniformă pe $[0,1]$: $u_1 = 0,479$; $u_2 = 0,178$ și $u_3 = 0,659$. Descrieți procedura și scrieți un cod R care să permită acest lucru.

b) Det. repartiția variabilei $Y = -\log(X)$ și calculați $E[Y]$ și $Var(Y)$.

c) Det. estimatorul de ver. max. $\hat{\theta}_n$ și verificați dacă este consistent și asimptotic normal.

d) Este estimatorul $\hat{\theta}_n$ asimptotic eficient?

e) Calculați deplasarea $b_\theta(\hat{\theta}_n)$.

Examen 2023 | ② Pt $\theta \in \mathbb{R}$ definim fct

$f_\theta(x) = \begin{cases} 1-\theta, & -\frac{1}{2} < x \le 0 \\ 1+\theta, & 0 < x \le \frac{1}{2} \\ 0, & \text{altfel} \end{cases}$

a) Ce condiții treb să verifice θ a.î. f_θ să fie o densitate de repartiție în raport cu măsura Lebesgue pe $\mathbb{R}$?

SOL: • $f \ge 0 = \begin{cases} 1+\theta \ge 0 \\ 1-\theta \ge 0 \end{cases} \Leftrightarrow \begin{cases} \theta \ge -1 \\ \theta \le 1 \end{cases} \Leftrightarrow \theta \in [-1,1]$

• $\int_{\mathbb{R}} f\, d\lambda = 1 \Leftrightarrow \int_{-\frac{1}{2}}^{0} 1+\theta\, dx + \int_{0}^{\frac{1}{2}} 1-\theta\, dx = 1 \Leftrightarrow$

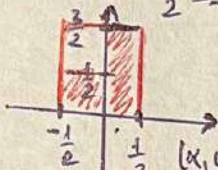
$\Leftrightarrow x+\theta x \Big|_{-\frac{1}{2}}^{0} + x-\theta x \Big|_{0}^{\frac{1}{2}} = 1 \Leftrightarrow \frac{1}{2} + \cancel{\frac{\theta}{2}} + \frac{1}{2} - \cancel{\frac{\theta}{2}} = 1 \Leftrightarrow$

$\Leftrightarrow \frac{2}{2} = 1 \Rightarrow 1 = 1$.

Concluzie: f_θ e fct de densitate $\Leftrightarrow \theta \in [-1,1]$

b) Pt $\theta = \frac{1}{2}$ dorim să generăm un eșantion de valoare n dintr-o populație f_θ. Descrieți procedura și scrieți un cod R care să permită acest lucru! Pt n=2000 scrieți un cod R care să compare vizual repartiția eșantionului generat cu repartiția teoretică. (→ nu știu)

SOL: $\theta = \frac{1}{2} \Rightarrow f_\theta(x) = \begin{cases} \frac{1}{2}, & -\frac{1}{2} < x \le 0 \\ \frac{3}{2}, & 0 < x \le \frac{1}{2} \\ 0, & \text{altfel} \end{cases} \Rightarrow G_{f_\theta} \subseteq [-\frac{1}{2}, \frac{1}{2}] \times [0, \frac{3}{2}]$



• Procedura: generăm un pct (x,y) din dreptunghiul $[-\frac{1}{2}, \frac{1}{2}] \times [0, \frac{3}{2}]$ și verificăm dacă $y \ge f(x)$. În caz afirmativ, adăugăm x în eșantion. Facem acest lucru până obținem n valori potrivite pt x.

• COD R:

```
esantion <- c(); n <- 2000
for (i in 1:n){
  x <- runif(-0.5, 0.5)
  y <- runif(0, 1.5)
  while(y > f(x)){
    x <- runif(-0.5, 0.5)
    y <- runif(0, 1.5)
  }
  append(esantion, x)
}.
hist(c) / barplot(esantion)
f <- function(x) {
if (x > -0.5 and x <= 0) {
  return 0.5
  }
if (x > 0 and x <= 0.5) {
return 1.5
}
return 0
}
```

append - adaugă val x la eșantion

hist - construiește o histogramă. la fel și barplot

→ afișează pct și cu val lui

→ afișează pt fiec x nr de apariții

→ aici am definit fct.

```
if (x>-0.5 and x<=0){
 return 0.5
}
if (x>0 and x<=0.5){
return 1.5
}
return 0
}
```

fct.

X v.a cu fct de repartiție F bijectivă. a) Cum putem genera cu ajutorul met. respingerii observații din repartiția lui X condiționată la $x>a$? Ce se întâmplă atunci când a e mai mare? Scrieți un cod R care să permită simularea unei observații din repartiția lui X condiționată la $x>a$, unde var. aleat. $X \sim Exp(1)$, iar $a=4$.

Dacă $U \sim Unif([0,1]) \Rightarrow F^{-1}(U) \sim X$

Folosim. met. respingerii pt a genera U până când $F^{-1}(U)>a$.

Când a e mai mare va crește timpul de execuție al algoritmului:

COD. R:

```
G <- function(x){        * log = ln.
 return ((-1)*log(x))
} // am construit G = F^-1
U <- runif(0,1)
while (G(U)<=4){
```

U = runif(0,1) //alegem U până când
} $G(U) > 4$, i.e $F(U) > 4$
U // afișăm U

b) $U \sim Unif([0,1])$ și T e definită prin $T = F^{-1}(F(a)+(1-F(a))U)$. Det. fct de rep. a lui T și găsiți o met de simulare condiționată la $x > a$. Care dintre cele două met e de preferat?

$$F(T \le t) = \mathbb{P}(F^{-1}(F(a)) + (1-F(a))U) \le t) =$$
$$= \mathbb{P}(F(a) + (1-F(a))U \le F(t)) =$$
$$= \mathbb{P}((1-F(a))U \le F(t) - F(a) \le F(t)) =$$
$$= \mathbb{P}\left(U \le \frac{F(t)-F(a)}{1-F(a)} \le \frac{F(t)}{1-F(a)} < F(t)\right) =$$
$$= \mathbb{P}(U \le F(t)) = \mathbb{P}(F^{-1}(U) \le t) = \mathbb{P}(X \le t) =$$
$= F_X(t) \Rightarrow$ rep lui T e aceeași cu cea a lui X

A doua met. e de preferat, deoarece generează direct variabile care se află între 0 și $F(1)$

③ Fie $\theta \ge 1$ un param. $n \in$ și $x_1, \dots, x_n$ un eșantion de vol n din pop f_θ cu $f_\theta = a x^\theta \mathbb{1}_{(0,1)}(x)$.
a) Det $a = ?$. Pt $\theta = 2$ dorim să găsim să generăm 3 valori aleatoare din rep. lui $X \sim f_\theta(x)$. Pt aceasta dispunem de 3 val rezultate din rep unif. pe $[0,1]$: $u_1 = 0,479$; $u_2 = 0,178$ și $u_3 = 0,659$. Descrieți procedura și scrieți un cod R care să permită acest lucru.

f_θ - dens de prob $\Rightarrow \begin{cases} f_\theta(x) \ge 1 \\ \int_{\mathbb{R}} f_\theta(x) \ge 0 \end{cases} \Leftrightarrow$