Nume și prenume: CONDRAT MIHAI Nota: __________
Grupa: 311

Examen

6 Februarie 2021

Timpul de rezolvare al problemelor este de 2h30. Pentru transmiterea soluțiilor în format PDF[1] în folderul vostru de pe Dropbox aveți 30 de minute timp suplimentar. Astfel, pentru dumneavoastră examenul începe la **ora 9 și 6 minute** și se termină la **ora 12 și 6 minute**.

Toate documentele, computerele personale, telefoanele mobile și/sau calculatoarele electronice de mână sunt autorizate. Orice modalitate de comunicare între voi este **strict interzisă**. Fiecare subiect valorează 10 puncte. Mult succes !

Exercițiul 1

10p

1. Considerăm densitatea de probabilitate

$$f(x) = \frac{1}{2\pi}\sqrt{4-x^2}\mathbf{1}_{[-2,2]}(x).$$

   a) Scrieți un cod `R` care să permită trasarea graficului funcției f pe intervalul $[-2, 2]$.
   b) Descrieți, folosind metoda respingerii, o procedură prin care să generați observații din f. Câte observații trebuie să generați în medie pentru a obține o realizare din f ?
   c) Scrieți un cod `R` care să genereze $n = 1000$ de observații i.i.d. din repartiția f și estimați valoarea medie a numărului de încercări necesare pentru a obține o realizare.

2. Considerăm cuplul de variabile aleatoare (X, Y) care este repartizat cu densitatea de repartiție

$$f(x, y) = \frac{1}{\sqrt{8\pi}}e^{-\frac{y^2 x}{2}-\sqrt{x}}\mathbf{1}_{x>0}.$$

   a) Determinați repartiția condiționată a lui Y la $X = x$
   b) Determinați repartiția lui $\sqrt{X}$
   c) Propuneți o metodă de simulare pentru o observație din densitatea $f(x, y)$ și scrieți un cod `R` care să permită acest lucru.

Exercițiul 2

10p

1. Numărul de clienți pe zi de la ghișeul unei bănci poate fi modelat ca o variabilă aleatoare $X \sim \mathcal{P}(\lambda)$. Pentru a îmbunătății serviciile oferite, banca vrea să estimeze parametrul λ atât prin metoda momentelor cât și prin metoda verosimilității maxime. Pentru aceasta dispune de următorul eșantion înregistrat pe parcursul a 3 săptămâni:

```
X: 19 23 30 25 25 23 22 20 22 21 25 20 25 28 27 24 32 26 24 17 32
```

   a) Determinați estimatorul obținut prin metoda momentelor $\tilde{\lambda}$ și estimatorul de verosimilitate maximă $\hat{\lambda}$ și verificați dacă aceștia sunt deplasați, consistenți și eficienți. Determinați repartiția lor limită.

   b) Găsiți estimatorul de verosimilitate maximă pentru $\mathbb{P}_\lambda(X_1 = 1 \,|\, X_1 > 0)$. Este acesta consistent ?

   c) Verificați dacă estimatorul aflat la punctul b) este sau nu nedeplasat.

[1]Pentru a transforma pozele în format PDF puteți folosi, de exemplu, programul CamScanner

2. Numărul de clienți care intră în magazinul Unirea zilnic poate fi modelat cu ajutorul unei variabile aleatoare repartizate Poisson de parametru λ, cunoscut. Odată intrat, un client cumpără produse în valoare de cel puțin 250 RON cu probabilitatea p. Pentru a estima p avem la dispoziție un eșantion $Y_1, Y_2, \ldots, Y_{30}$ pentru 30 zile, reprezentând numărul de clienți, zilnic, care au efectuat cumpărături de cel puțin 250 RON:

```
Y: 6 3 3 6 5 4 3 6 4 4 3 2 2 2 3 2 3 4 2 5 2 2 5 2 3 3 2 3 3 3
```

```
Y: 1 5 4 4 5 3 4 6 6 4 4 3 3 1 2 1 6 2 2 4 5 3 2 1 8 2 1 3 5 4
```

Propuneți un estimator pentru p, studiați proprietățile acestuia și dați o estimare plecând de la eșantionul dat (știind că $\lambda = 25$).

Exercițiul 3

10p

Fie $X_1, X_2, \ldots, X_n$ un eșantion de volum n din populația f_θ unde

$$f_\theta(x) = Axe^{-\theta x^2}\mathbf{1}_{\{x>0\}}$$

cu $\theta > 0$ parametru necunoscut și A o constantă (care depinde de θ).

a) Determinați constanta A și calculați funcția de repartiție $F_\theta(x)$ a lui X_1.

b) În cazul în care $\theta = 3$ dorim să generăm 3 valori aleatoare din repartiția lui X. Pentru aceasta dispunem de trei valori rezultate din repartiția uniformă pe $[0,\, 1]$: $u_1 = 0.923$, $u_2 = 0.4$ și $u_3 = 0.35$. Descrieți procedura și scrieți un cod `R` care să permită acest lucru.

c) Determinați mediana $x_{1/2}$ repartiției lui X_1. Plecând de la aceasta deduceți un estimator $\tilde{\theta}_n$ a lui θ și determinați repartiția limită a lui $\sqrt{n}(\tilde{\theta}_n - \theta)$.

d) Determinați estimatorul de verosimilitate maximă $\hat{\theta}_n$ a lui θ.

e) Arătați că $X_1^2 \sim \text{Exp}(\theta)$ și calculați $\mathbb{E}[X_1^1]$ și $Var(X_1^2)$.

f) Determinați repartiția limită a lui $\sqrt{n}(\hat{\theta}_n - \theta)$.

g) Calculați informația lui Fisher și verificați dacă $\hat{\theta}_n$ este asimptotic eficient.

h) Pe care dintre cei doi estimatori îi preferați ? Ce puteți spune de estimatorul obținut prin metoda momentelor ?

Examen 6 feb 2021

Ex 1: $f(x) = \frac{1}{2\pi} \sqrt{4-x^2} \, \mathbb{1}_{[-2,2]}(x)$

a) Scrieți un cod R care să permită trasarea graficului funcției f pe $[-2;2]$.

```
# Scriem funcția f
f <- function(x){
    return(1/(2*pi)*sqrt(4-x^2))
# Discretizăm intervalul
interval <- seq(-2,2,0.001)
# Reprezentăm funcția f pe int [-2;2]
plot(interval, f(interval), type=l, col=red)
```

b) Descrieți, fol met respingerii, o procedură prin care să generați obs din f. Câte obs trebuie să generați în medie pt a obține o realizare din f?

Pas 1: Generez Y

Pas 2: Generez U

Pas 3: Dacă $U \le \frac{f(y)}{cg(y)}$ atunci $X=Y$ altfel repet pasul 1.

La noi $f(x) = \frac{1}{2\pi}\sqrt{1-x^2}\,\mathbb{1}_{[-2,2]}(x)$. Aleg $Y \sim Unif(-2,2)$, $g(x) = \frac{1}{4}$

$\hookrightarrow g(x) = \frac{1}{\underset{2}{b} - \underset{-2}{a}} = \frac{1}{4}$

Căutăm constanta c.

Calculăm $\frac{f(x)}{g(x)} = h(x) = \frac{2}{\pi}\sqrt{4-x^2}$

Verificăm dacă $h(x)$ are punct de maxim:

$$h'(x) = \frac{2}{\pi} \cdot \frac{1}{2\sqrt{4-x^2}} \cdot (-2x) = \frac{2}{\pi} \frac{-x}{\sqrt{4-x^2}}$$

$$h'(x) = 0 \Rightarrow x = 0$$

x	-2				0				2
$h'(x)$	$+$	$+$	$+$	$+$	0	$-$	$-$	$-$	$-$
$h(x)$	$\nearrow$				$h(0)$		$\searrow$		

$$h(0) = \frac{2}{\pi}\sqrt{4} = \frac{4}{\pi} = c$$

$$\Rightarrow \frac{f(x)}{c\,g(x)} = \frac{1/2\pi\,\sqrt{4-x^2}}{4/\pi \cdot 1/4} = \frac{1}{2}\sqrt{4-x^2}$$

c este val maximă ($h(0)$)

Pas 1: Generăm U_1 și U_2 ($\sim Unif(-2,2)$)

Pas 2: Dacă $U_2 \le \frac{1}{2}\sqrt{4-U_1^2}$ atunci $X = U_1$ și mergem la pas 1

! Cum nr de iterații pt a obține o val din X e o v.a. $\sim Geom$ cu medie c

$\Rightarrow c = \frac{4}{\pi} = 1{,}27 \rightarrow$ nr mediu de valori

c) $n = 1000$ obs i.i.d din rep f și estimați val medie a nr de încercări necesare pt a obține o realizare.

$c = \frac{4}{\pi}$

$$\overset{c}{\overset{\parallel}{\bar{X}}} = \frac{\sum x}{n} \Rightarrow \sum x = n \cdot c$$

Deci pt noi $c \cdot n = \frac{4}{\pi} \cdot 1000 = \simeq 1273{,}23.$

$\pi \simeq 3{,}1415$

media (câte încercări ca să ai 1000)

2)

Ex 2:

$X \sim P(\lambda)$

a) $\tilde{\lambda} = ?$; $\hat{\lambda} = ?$ deplasați, consistenți, eficienți?

Fie $X_1, \dots, X_n$ un eșantion de talie n dintr-o pop Poisson de param $\lambda > 0$

$X \sim Pois(\lambda)$

Din teorie știm că $P(X=x) = \frac{\lambda^x e^{-\lambda}}{x!}$, $x>0$ discretă

Pas 1: scriu funcția de verosimilitate:

$$L(x|\lambda) = \prod_{i=1}^{n} P(X=x_i) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = (e^{-\lambda})^n \prod_{i=1}^{n} \frac{\lambda^{x_i}}{x_i!} = e^{-\lambda n} \cdot \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!}$$

Pas 2: Logaritmez:

$$\ln L(x|\lambda) = \ln e^{-\lambda n} + \ln \lambda^{\sum_{i=1}^{n} x_i} - \ln \prod_{i=1}^{n} x_i! = -\lambda n + \sum_{i=1}^{n} x_i \ln \lambda - \sum_{i=1}^{n} \ln (x_i!)$$

Pas 3+4: Derivez; egalez cu 0 și rezolvăm ec de verosimilitate

$$\frac{\partial \ln L(x|\lambda)}{\partial \lambda} = 0 \Rightarrow \frac{\partial \ln L(x|\lambda)}{\partial \lambda} = -n + \sum_{i=1}^{n} x_i \cdot \frac{1}{\lambda}$$

$$\frac{-n\lambda + \sum_{i=1}^{n} x_i}{\lambda} = 0 \Rightarrow -n\lambda + \sum_{i=1}^{n} x_i = 0 \Rightarrow n\lambda = \sum_{i=1}^{n} x_i$$

$$\lambda = \frac{1}{n} \sum_{i=1}^{n} x_i \Rightarrow \hat{\lambda} = \overline{X}_n$$

$\tilde{\lambda} = ?$ (Met Mom)

$E[X] = \overline{X}_n$ Din teorie știu că $E[X] = \lambda$

$\Rightarrow \lambda = \overline{X}_n \Rightarrow \tilde{\lambda} = \overline{X}_n$

Deoarece $\hat{\lambda}$ este obținut prin M.V.M. $\Rightarrow \hat{\lambda}$ este asim eficient și consistent $\Rightarrow$ și $\tilde{\lambda}$ este la fel pt că $\hat{\lambda} = \tilde{\lambda} = \overline{X}_n$

$$E[\hat{\lambda}] = E[\overline{X}_n] = E\left[\frac{1}{n} \sum_{i=1}^{n} x_i\right] = \frac{1}{n} E\left[\sum_{i=1}^{n} x_i\right] = \frac{1}{n} \cdot n \underbrace{E[X]}_{= \lambda \text{ (din teorie)}} = \lambda$$

$\Rightarrow \hat{\lambda}$ e nedeplasat $\Rightarrow \tilde{\lambda}$ nedeplasat pt că $\hat{\lambda} = \tilde{\lambda}$

b) $\hat{\lambda} = ?$ pt $P_\lambda(X_1 = 1 | X_1 > 0)$

$X_1 \sim P(\lambda)$

$$P_\theta(X_1 = 1 | X_1 > 0) = \frac{P(X_1 = 1) \cap P(X_1 > 0)}{P(X_1 > 0)}$$

$$= \frac{P(X_1 = 1)}{1 - P(X_1 \le 0)} \longrightarrow \text{asta e intersecția pt că } X>0 \text{ în Poisson}$$

$$= \frac{P(X_1 = 1)}{1 - P(X = 0)} = \frac{\lambda^1 e^{-\lambda}}{1!} \cdot \frac{1}{1 - \frac{\lambda^0 e^{-\lambda}}{0!}} = \frac{\lambda e^{-\lambda}}{1 - \frac{e^{-\lambda}}{1}} = \frac{\lambda e^{-\lambda}}{1 - e^{-\lambda}}$$

$$g(\lambda) \overset{not}{=} \frac{e^{-\lambda}}{1 - e^{-\lambda}}$$

Se cere un estimator de verosim. maximă pt $g(\lambda)$

! Avem o teoremă: Dacă $\hat{\lambda}$ este EVM pt λ atunci pt $(\forall)$ funcție g avem că $g(\hat{\lambda})$ este EVM pt $g(\lambda)$

$\Rightarrow \hat{\lambda}$ EVM pt $\lambda \Rightarrow g(\hat{\lambda})$ EVM pt $g(\lambda)$

$$g(\hat{\lambda}) = \frac{\hat{\lambda} \cdot e^{-\hat{\lambda}}}{1-e^{-\hat{\lambda}}} = \frac{\bar{X}_n \cdot e^{-\bar{X}_n}}{1-e^{-\bar{X}_n}} \overset{not}{=} \widehat{g(\lambda)}$$ (este estimatorul lui $g(\lambda)$)

Teorema aplicațiilor continue (T.A.C.)

$X_n \xrightarrow{P} X$ atunci $g(X_n) \xrightarrow{P} g(X)$

$X_n \xrightarrow{\mathcal{D}} X$ și $X_n \xrightarrow{a.s} X$

$\Rightarrow$ dacă $\bar{X}_n \xrightarrow{P} \lambda$ atunci și $g(\bar{X}_n) \xrightarrow{P} g(\lambda)$

adică $g(\hat{\lambda}) \xrightarrow{P} g(\lambda) \Rightarrow \widehat{g(\lambda)}$ e consistent