

1. Cererea de memorie pentru o aplicație, ca proporție din memoria ce poate fi alocată de un utilizator, este o variabilă aleatoare X ce are densitatea de probabilitate

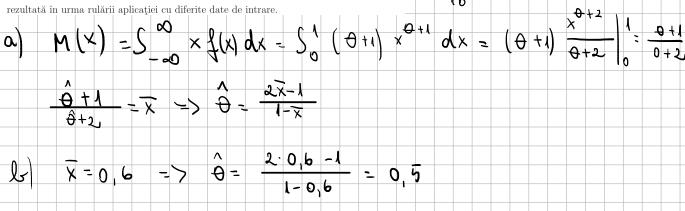
$$f(x) = \left\{ \begin{array}{ll} (\theta+1)x^{\theta}, & 0 < x < 1, \\ 0, & \text{in rest.} \end{array} \right.$$

- a) Să se determine media teoretică M(X) a variabilei aleatoare X și apoi să se estimeze
- θ în funcție de media de selecție \overline{x} a unei selecții aleatoare de volum n.
- b) Să se determine un estimator al parametrului θ din selecția următoare:

stimator al parametrului
$$\theta$$
 din selecția următoare:

0.2, 0.4, 0.5, 0.7, 0.8, 0.9, 0.9, 0.6, 0.6, 0.4, \longrightarrow $\overline{\chi} = 0.2 + 0.4 + 0.5 + \cdots = 0.4$
ii aplicației cu diferite date de intrare.

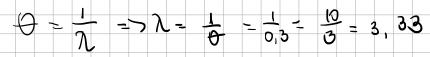
rezultată în urma rulării aplicației cu diferite date de intrare.



2. Pentru a estima rata sosirii λ a cererilor de acces la o bază de date s-au monitorizat intervalele de timp dintre 10 cereri consecutive și s-au înregistrat valorile:

$$0.2, 0.1, 0.1, 0.05, 0.05, 0.2, 0.8, 0.2, 0.8, 0.2, 0.8, 0.2, 0.8, 0.2, 0.8$$

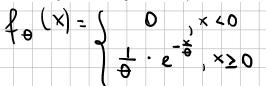
Care este estimatorul ratei sosirilor, $\hat{\lambda}$?



3. Fie populația \mathcal{P} formată dintr-un tip de circuite. Caracteristica ce dorim să o investigăm prin sondaj statistic este durata de viață a acestor circuite, știind că aceasta este exponențial distribuită, cu parametrul θ necunoscut. Măsurând timpul de viață (în ani) a 10 circuite, se obțin valorile:

 $0.8830, \ 1.96511, \ 1.9189, \ 4.8448, \ 0.9208 \ 3.4377, \ 1.7162, \ 4.2327, \ 5.9435, \ 8.3128.$

Să determinăm estimatorul de verosimilitate maximă pentru θ (adică pentru media du-



Astfel, funcția de verosimilitate este

$$L(\theta; x_1, x_2, \dots, x_n) = \prod_{i=1}^n \frac{1}{\theta} e^{-x_i/\theta} = \frac{1}{\theta^n} e^{-\sum_{i=1}^n x_i/\theta}.$$
 (14.2)

Funcția logaritmică cu bază mai mare ca 1 are derivata întâi pozitivă. Notând cu $h=\ln$ și cu $\ell(\theta)=h(L(\theta)),$ avem că funcția $\ell'(\theta)=h(L(\theta))L'(\theta)$ are același semn ca derivata lui L, deci ℓ și L au aceleași puncte de extrem și de aceeași natură. Pentru simplitatea calculelor, determinăm punctul de maxim absolut (dacă acesta există) pentru ℓ și acesta va fi punct de maxim absolut și pentru L:

$$\ell(\theta) = \ln L(\theta; x_1, x_2, \dots, x_n) = -n \ln \theta - \frac{\sum_{i=1}^{n} x_i}{\theta}.$$
 (14.3)

Avem

$$\ell'(\theta) = -\frac{n}{\theta} + \frac{\sum_{i=1}^{n} x_i}{\theta^2}.$$
(14.4)

Rezolvând ecuația $\ell'(\theta) = 0$ în raport cu θ , obținem punctul $\hat{\theta} = \frac{\sum_{i=1}^{n} x_i}{n} = \overline{x}$, care este maxim absolut pentru ℓ , deci și pentru L. Reamintim condiția suficientă de maxim: $l^{(n)}(x_0) < 0, n$ -par, unde x_0 este punct critic, i.e. $l'(x_0) = 0$. Aici, $l''(\overline{x}) = -\frac{n}{\overline{x}^2} < 0$.

În concluzie.

$$\operatorname{argmax} L(\theta; x_1, x_2, \dots, x_n) = \overline{x},$$

estimatorul de verosimilitate maximă a parametrului θ a distribuției exponențiale este media de selecție. În cazul exemplului dat, estimatorul verosimilității maxime a mediei de viață a circuitelor este media selecției:

$$\hat{\theta} = \frac{x_1 + x_2 + \ldots + x_{10}}{10} = 3.525.$$

Observație: Valorile de selecție au fost generate simulând o variabilă $X \sim Exp(\theta=3.5)$, deci estimatorul verosimilitații maxime $\hat{\theta}=3.525$ este "destul de bun".

4. Un simulator al distribuției Bernoulli

$$X = \left(\begin{array}{cc} 1 & 0 \\ p & 1 - p \end{array}\right),$$

de parametru p necunoscut, generează stringul de biți:

Să se determine estimatorul verosimilitătii maxime al parametrului p pe baza eșantionului de biti.

într-adevăr, $l''(\frac{12}{26}) < 0$. Se observă că $\hat{\theta}$ este egal cu numărul biţilor 1 din string supra numărul total de biţi. Acest estimator al lui p este de fapt probabilitatea intuitivă de a obţine bitul 1: numărul cazurilor favorabile din string supra numărul cazurilor posibile.

5. Variabila aleatoare X, care dă numărul de zone defecte ale unui CD de un anumit tip, are următoarea distribuție de probabilitate: p(x) = P(X = x), unde

X	p(x)	ײ	1				
0	0.75	0	0,75	_	+	-	
1	0.15	l	0, 15	M(x2)= 0,15+	ОЧ	- C	96
2	0.10	ч	0, (11(x/= 0)) '		7 77

- a) Să se calculeze media și abaterea standard a numărului de zone defecte ale CD-ului.
- b) Ce distribuție de probabilitate are media de selecție a unui eșantion de 400 de CD-uri din tipul investigat? Care este media și dispersia acestei distribuții?
- c) Care este probabilitatea ca media numărului de zone defecte/CD într-un lot de 400 de CD-uri să fie mai mică decât 0.3?

a)
$$M(x) = 0.0, 45 + 1.0, 15 + 2.0, 1 = 0,35$$

$$\nabla^{2}(x) = M(x^{2}) - M(x) = 0,55 - (0,35) = 0,4245$$

$$\nabla(x) = \sqrt{0,4245} = 0,6538$$

$$X_{400} \sim A_{0}N(m, \sqrt{1}/400)$$

$$M = 0,35$$

$$S^{2} = \frac{\sqrt{100}}{400} = 0,00106$$

$$S^{2} = \frac{\sqrt{100}}{400} = 0,00106$$

$$C) P(x_{100} < 0,3) = 7 = (0,3) - \sqrt{\frac{0,5-m}{b}} = \sqrt{(-1,5,625)} = 1 - \sqrt{(1,5625)} = \sqrt{\frac{0,5-m}{b}}$$

$$= 1 - 0,94 = 0,06$$