

LABORATOR #12

EX#1 Fie setul de date (x_1, x_2, \dots, x_n) din fișierul atașat `sample.Normal.npy`. Creați un fișier în Python® prin care

- (a) să se afișeze într-o figură histograma datelor;
- (b) să se afișeze într-o figură graficul funcției log-verosimilitate

$$\log L(x_1, x_2, \dots, x_n, \theta_1, \theta_2) = -n \log(\sqrt{2\pi\theta_2}) - \frac{1}{2\theta_2} \sum_{i=1}^n (x_i - \theta_1)^2$$

- corespunzătoare setului de date (x_1, x_2, \dots, x_n) , pentru $(\theta_1, \theta_2) \in (-1, 1) \times (0, 0.1)$;
- (c) să se determine estimarea $(\hat{\theta}_1, \hat{\theta}_2)$ a parametrilor distribuției care maximizează funcția de log-verosimilitate corespunzătoare asociată setului de date, unde $\hat{\theta}_1 = \frac{1}{n} \sum_{i=1}^n x_i$, $\hat{\theta}_2 = \frac{1}{n} \sum_{i=1}^n (x_i - \theta_1)^2$;
- (d) să se afișeze în figura de la (b) graficul punctului $(\hat{\theta}_1, \hat{\theta}_2, \log L(x_1, x_2, \dots, x_n, \hat{\theta}_1, \hat{\theta}_2))$.

EX#2 Fie seturile de date atașate

- (i) `sample.NegativBinom.npy`;
- (ii) `sample.Gamma.npy`;

Creați un fișier în Python® prin care, pentru fiecare set de date (x_1, x_2, \dots, x_n) de la (i)-(ii):

- (a) să se afișeze într-o figură histograma datelor;
- (b) să se afișeze într-o figură graficul funcției log-verosimilitate $\log L(x_1, x_2, \dots, x_n, \theta_1, \theta_2)$ corespunzătoare setului de date (x_1, x_2, \dots, x_n) :

(i) $\log L(x_1, x_2, \dots, x_n, \theta_1, \theta_2)$

$$= \sum_{i=1}^n \log \Gamma(x_i + \theta_1) - n \log \Gamma(\theta_1) - \sum_{i=1}^n \log \Gamma(x_i + 1) + \log(1 - \theta_2) \sum_{i=1}^n x_i + n \theta_1 \log \theta_2,$$

pentru $(\theta_1, \theta_2) \in (0, 10) \times (0, 1)$;

(ii) $\log L(x_1, x_2, \dots, x_n, \theta_1, \theta_2)$

$$= n \theta_1 \log \theta_2 - n \log \Gamma(\theta_1) + (\theta_1 - 1) \sum_{i=1}^n \log x_i - \theta_2 \sum_{i=1}^n x_i,$$

pentru $(\theta_1, \theta_2) \in (0, 10) \times (0, 5)$;

- (c) să se determine o aproximare numerică pentru estimarea $(\hat{\theta}_1, \hat{\theta}_2)$ a parametrilor distribuției care maximizează funcția de log-verosimilitate $\log L(x_1, x_2, \dots, x_n, \theta_1, \theta_2)$ corespunzătoare setului de date (x_1, x_2, \dots, x_n) ;
- (d) să se afișeze în figura de la (c) graficul punctului $(\hat{\theta}_1, \hat{\theta}_2, \log L(x_1, x_2, \dots, x_n, \hat{\theta}_1, \hat{\theta}_2))$.

Indicații Python®: `numpy`, `scipy.stats`, `scipy.optimize.minimize`, `scipy.optimize.fsolve`, `scipy.special.gamma`, `scipy.special.digamma`, `matplotlib.pyplot`, `matplotlib.pyplot.hist`, `3D plotting`