Metode numerice și statistică
Curs 10

lect. Ciprian Deliu
✉ cdeliu@tuiasi.ro
moodle.deliu.ro

Universitatea Tehnică "Gh. Asachi" Iași
Facultatea de Hidrotehnică, Geodezie și Ingineria Mediului

2019

Repartiția discretă uniformă

Definiție

Spunem că variabila aleatoare discretă X are o repartiție **uniformă** *dacă distribuția ei este*

$$X : \begin{pmatrix} x_1 & x_2 & \dots & x_n \\ p_1 & p_2 & \dots & p_n \end{pmatrix}, \text{ unde } p_k = \frac{1}{n}, \forall k = 1, 2, \dots, n.$$

Teoremă

Media și dispersia unei variabile aleatoare discrete repartizate uniform sunt:

$$E(X) = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$Var(X) = \frac{1}{n}\sum_{i=1}^{n} x_i^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} x_i\right)^2$$

Repartiția binomială

Definiție

Spunem că variabila aleatoare discretă X are o repartiție **binomială** *cu parametrii $n \in \mathbb{N}$ și $p \in (0,1)$ dacă ia valorile $0, 1, 2, \dots, n$ cu probabilitățile*

$$P(X = k) = C_n^k p^k q^{n-k}, \ \forall k = 0, 1, 2, \dots, n,$$

unde $q = 1 - p$.

Teoremă

Media și dispersia unei variabile aleatoare discrete repartizate binomial sunt:

$$E(X) = np, \ Var(X) = npq.$$

Dacă A este un eveniment legat de o experiență iar proabilitatea ca A să se realizeze când efectuăm o singură dată experiența este $P(A) = p$, atunci variabila aleatoare care are ca valori numărul realizărilor lui A când efectuăm de n ori experiența are o repartiție binomială cu parametrii n și p.

Repartiția Poisson

Definiție

Spunem că variabila aleatoare discretă X are o repartiție **Poisson** *cu parametrul $\lambda > 0$ dacă poate lua orice valoare întreagă pozitivă cu probabilitățile*

$$P(X = k) = \frac{\lambda^k}{k!}e^{-\lambda}, \ \forall k = 0, 1, 2, \dots.$$

Teoremă

Media și dispersia unei variabile aleatoare discrete repartizate Poisson sunt:

$$E(X) = \lambda, \ Var(X) = \lambda.$$

- Dacă n este suficient de mare și p suficient de mic ($n \geq 30$ și $np < 5$), atunci putem aproxima repartiția binomială cu parametrii n și p prin repartiția Poisson de parametru $\lambda = np$. Din acest motiv repartiția Poisson se mai numește **legea evenimentelor rare**.
- Exemple de variabile aleatoare care au repartiții de tip Poisson: numărul erorilor de pe o pagină (sau grup de pagini) dintr-o carte, numărul persoanelor dintr-o comunitate care au peste 80 ani, numărul defectelor apărute la un dispozitiv complex într-un interval de timp, numărul clienților unui magazin pe parcursul unei zile, etc.

Repartiția continuă uniformă

Definiție

Spunem că variabila aleatoare X are o repartiție **uniformă** *pe intervalul $[a,b]$ dacă densitatea ei de repartiție este*

$$f(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b] \\ 0, & x \in (-\infty, a) \cup (b, \infty) \end{cases}$$

- Funcția de repartiție a unei variabile aleatoare uniforme este

$$F(x) = \begin{cases} 0, & x \leq a \\ \frac{x-a}{b-a}, & a < x \leq b \\ 1, & x > b \end{cases}.$$

- Media și dispersia unei variabile aleatoare uniforme sunt

$$E(X) = \frac{a+b}{2}$$

$$Var(X) = \frac{(a-b)^2}{12}.$$

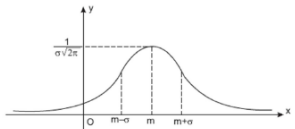
Repartiția normală

Definiție

Spunem că variabila aleatoare X are o repartiție **normală (gaussiană)** *de parametri m și σ dacă densitatea ei de repartiție este*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-m)^2}{2\sigma^2}}, \ m, x \in \mathbb{R}, \ \sigma > 0.$$

Graficul densității unei v. a. repartizate $N(m, \sigma^2)$ este



- dreapta $x = m$ este axă de simetrie;
- pentru $x = m$ se atinge valoarea maximă a densității $\frac{1}{\sigma\sqrt{2\pi}}$;
- punctele $x = m - \sigma$ și $x = m + \sigma$ sunt puncte de inflexiune.

Pentru $m = 0$ și $\sigma = 1$ se obține repartiția **normală standard** $N(0,1)$ având densitatea de repartiție

$$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}.$$

Repartiția normală

- Funcția $\Phi : \mathbb{R} \to \mathbb{R}$ dată prin

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}} dt$$

se numește **funcția integrală a lui Laplace** și are proprietățile
  1. $\Phi(0) = 0$,
  2. $\Phi(\infty) = \frac{1}{2}$,
  3. $\Phi(-x) = -\Phi(x), \ \forall x > 0$.
- Funcția de repartiție a unei variabile aleatoare $N(0,1)$ este

$$F(x) = \frac{1}{2} + \Phi(x).$$

- Dacă variabila aleatoare X este repartizată $N(m, \sigma^2)$, atunci variabila aleatoare $Y = \frac{1}{\sigma}(X - m)$ este repartizată $N(0,1)$ și obținem pentru funcția de repartiție a lui X:

$$F(x) = \frac{1}{2} + \Phi\left(\frac{x-m}{\sigma}\right).$$

Repartiția normală

Teoremă

Dacă variabila aleatoare X are o repartiție normală cu parametrii m și σ, atunci valoarea medie și dispersia sa sunt

$$E(X) = m, \ Var(X) = \sigma^2.$$

Propoziție

Dacă variabila aleatoare X are o repartiție normală cu parametrii m și σ, iar $a, b, k \in \mathbb{R}$, $k > 0$ atunci

a) $P(a < X < b) = \Phi\left(\frac{b-m}{\sigma}\right) - \Phi\left(\frac{a-m}{\sigma}\right)$;

b) $P(|X - m| < k\sigma) = 2\Phi(k)$.

Folosind tabelele de valori ale funcției Φ găsim:

- $P(|X - m| < \sigma) = 2\Phi(1) = 0.6826$, adică cel puțin 68% din valorile lui X se află în intervalul $(m - \sigma, m + \sigma)$.
- $P(|X - m| < 2\sigma) = 2\Phi(2) = 0.9544$, adică cel puțin 95% din valorile lui X se află în intervalul $(m - 2\sigma, m + 2\sigma)$.
- $P(|X - m| < 3\sigma) = 2\Phi(3) = 0.9974$, adică cel puțin 99% din valorile lui X se află în intervalul $(m - 3\sigma, m + 3\sigma)$.

Repartiția exponențială

Definiție

Spunem că variabila aleatoare X are o repartiție **exponențială** *de parametru $\lambda > 0$ dacă densitatea ei de repartiție este*

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0. \end{cases}$$

- Funcția de repartiție este

$$F(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-\lambda x}, & x > 0. \end{cases}$$

- Media și dispersia sunt:

$$E(X) = \frac{1}{\lambda}, Var(X) = \frac{1}{\lambda^2}.$$

- Dacă variabila aleatoare care dă numărul de apariții ale unui eveniment în intervalul $[0,t]$ este repartizată Poisson de parametru λt, atunci variabila aleatoare care dă lungimea intervalului de timp dintre două apariții succesive ale evenimentului este repartizată exponențial de parametru λ.

Repartiția Gamma

Definiție

Spunem că variabila aleatoare X are o repartiție **Gamma** *de parametru $a > 0$ dacă densitatea ei de repartiție este*

$$f(x) = \begin{cases} \frac{\lambda^a x^{a-1} e^{-\lambda x}}{\Gamma(a)}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

unde $\lambda > 0, a > 0$, iar $\Gamma(a) = \int_0^\infty x^{a-1} e^{-x} dx$.

$\Gamma(a)$ se numește **integrala Gamma a lui Euler** și are proprietățile:

- $\Gamma(a)$ este convergentă pentru $a > 0$ și divergentă pentru $a \leq 0$;
- $\Gamma(a+1) = a\Gamma(a)$;
- $\Gamma(1) = 1$ și $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$;
- $\Gamma(n) = (n-1)!, \ \forall n \in \mathbb{N}^*$.

Repartiția Gamma

- Funcția de repartiție este

$$F(x) = \begin{cases} 0, & x \leq 0 \\ \frac{1}{\Gamma(a)} \int_0^x \lambda^a t^{a-1} e^{-\lambda t} dt, & x > 0. \end{cases}$$

- Media și dispersia sunt

$$E(X) = \frac{a}{\lambda}, \ Var(X) = \frac{a}{\lambda^2}.$$

- Pentru $a = 1$ se obține repartiția exponențială.
- Pentru $a = 2, 3, \dots$ repartiția Gamma corespunzătoare se numește repartiție **Erlang**.
- Dacă variabila aleatoare X este repartizată normal de parametri m și σ, atunci variabila aleatoare $Y = \frac{1}{2\sigma^2}(X-m)^2$ are o repartiție Gamma cu parametrii $\lambda = 1$ și $a = \frac{1}{2}$.

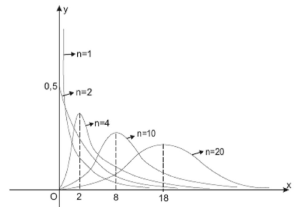
Repartiția χ^2

Definiție

Spunem că variabila aleatoare X are o repartiție χ^2 cu n grade de libertate dacă densitatea ei de repartiție este

$$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{1}{2}x}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Graficele densităților unor v. a. repartizate χ^2 pentru diverse grade de libertate:



- repartiția χ^2 este un caz particular al repartiției Gamma pentru $\lambda = \frac{1}{2}$ și $a = \frac{n}{2}, \ n \in \mathbb{N}$;
- media și dispersia sunt

$$E(X) = n, \ Var(X) = 2n;$$

- are multe aplicații în statistica inferențială, motiv pentru care funcția ei de repartiție este tabelată pentru diverse valori ale lui n.

Definiție

Spunem că variabila aleatoare X are o repartiție **Student (t)** *cu n grade de libertate dacă densitatea ei de repartiție este*

$$f(x) = \frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)} \left(1 + \frac{x^2}{n}\right)^{-\frac{n+1}{2}}, \ n \in \mathbb{N}, x \in \mathbb{R}.$$

- Media și dispersia unei variabile aleatoare repartizate Student sunt

$$E(X) = 0, \ Var(X) = \frac{n}{n-2}.$$

- Pentru $n \to \infty$ densitatea repartiției Student converge către densitatea repartiției normale standard.

Definiție

Spunem că variabila aleatoare X are o repartiție **Fisher-Snedecor** *cu parametrii $n_1, \ n_2 \in \mathbb{N}$ dacă densitatea ei de repartiție este*

$$f(x) = \begin{cases} \left(\frac{n_1}{n_2}\right)^{\frac{n_1}{2}} \frac{\Gamma\left(\frac{n_1+n_2}{2}\right)}{\Gamma(\frac{n_1}{2})\Gamma(\frac{n_2}{2})} x^{\frac{n_1}{2}-1} \left(1 + \frac{n_1}{n_2}x\right)^{-\frac{n_1+n_2}{2}}, & x > 0 \\ 0, & x \leq 0. \end{cases}$$

Legea numerelor mari

Definiție

Spunem că șirul de variabile aleatoare $(X_n)_{n\in\mathbb{N}}$ este **convergent în probabilitate** *la variabila aleatoare X dacă pentru orice $\varepsilon > 0$*

$$\lim_{n\to\infty} P(|X_n - X| > \varepsilon) = 0.$$

Teoremă

Fie un șir de variabile aleatoare $(X_n)_{n\in\mathbb{N}}$ independente și identic repartizate cu

$$E(X_n) = m \text{ și } Var(X_n) = \sigma^2, \ \forall n \in \mathbb{N}.$$

Atunci variabila aleatoare $Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k$ converge în probabilitate la m.

Legea numerelor mari este cunoscută și sub denumirea de **regularitate statistică**. Termenul de regularitate statistică se motivează în felul următor: dacă avem o caracteristică X și se fac n determinări independente $x_1, x_2, \dots, x_n$, apoi m determinări independente $y_1, y_2, \dots, y_m$ ale acesteia, atunci pentru n, m mari avem $\frac{1}{n}\sum_{k=1}^{n} x_k \simeq \frac{1}{m}\sum_{k=1}^{m} y_k$.

Teorema limită centrală

Teoremă

Fie șirul de variabile aleatoare independente $(X_n)_{n\in\mathbb{N}}$ cu aceeași repartiție, pentru care există $E(X_n)$ și $Var(X_n) \neq 0, \ \forall n \in \mathbb{N}$. Notăm cu

$$Y_n = \frac{1}{n}\sum_{k=1}^{n} X_k, \ Z_n = \frac{Y_n - E(Y_n)}{\sqrt{Var(Y_n)}}$$

și cu $(G_n)_{n\in\mathbb{N}}$ șirul funcțiilor de repartiție ale variabilelor aleatoare Z_n. În aceste condiții șirul (Z_n) converge în repartiție la o variabilă aleatoare cu distribuția $N(0,1)$, adică

$$\lim_{n\to\infty} G_n(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt.$$