Curs 6: Variabile aleatoare continue. Media şi dispersia unei variabile aleatoare continue.

1.1 Densitatea de probabilitate

Variabilele aleatoare discrete se asociază preponderent experimentelor aleatoare ce constau în numărarea unor rezultate. Există însă variabile aleatoare ce pot lua orice valoare dintr-un interval sau din întreg $\mathbb{R}$. De exemplu, timpul de execuţie a unui program, durata de viaţă a unei componente electronice, frecvenţa de acces în traficul pe WEB, dimensiunea pachetelor de date în FTP (File Transfer Protocol), etc.

O variabilă aleatoare discretă este perfect determinată de mulţimea valorilor $D = \{x_1, x_2, \ldots, x_n\}$ şi de distribuţia de probabilitate, adică lista $(p_1, p_2, \ldots, p_n)$, unde $p_k = P(X = x_k)$.

În cazul în care mulţimea valorilor unei variabile aleatoare nu este discretă, variabila este studiată şi caracterizată într-un context diferit. În acest caz distribuţia de probabilitate nu se poate specifica printr-un tablou.

Definiţia 1.1.1 *O funcţie $f : \mathbb{R} \to \mathbb{R}$ cu proprietăţile:*
1) $f(x) \geq 0$, $\forall\, x \in \mathbb{R}$.
2) f este integrabilă pe $\mathbb{R}$ şi $\int\limits_{-\infty}^{\infty} f(x)\, dx = 1$
se numeşte densitate de probabilitate.

O variabilă aleatoare continuă este o variabilă aleatoare pentru care distribuţia de probabilitate este definită de o densitate de probabilitate, f_X. Şi anume, pentru o variabilă aleatoare X evenimentele de interes sunt $(X \in I)$, unde I poate fi oricare din intervalele mărginite:

$$[a, b], (a, b), (a, b], (a, b)$$

sau nemărginite:

$$(-\infty, b], (-\infty, b), [a, \infty), (a, \infty)$$

iar probabilitatea ca X să ia valori într-un interval I este integrală pe intervalul I din densitatea sa de probabilitate:

$$P(X \in I) = \int_I f_X(x)dx$$

Din punct de vedere geometric, $\int_I f_X(x)\,dx$ este aria domeniului având "frontiera superioară" un arc din graficul densităţii f_X, iar cea inferioară, intervalul I (inclus în axa $x'Ox \equiv \mathbb{R}$ (Fig.1, Fig.3)).

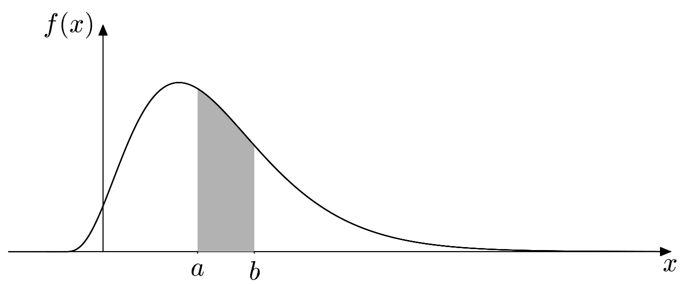

Fig.1: Aria domeniului haşurat reprezintă $P(a \leq X \leq b)$

Observaţia 1.1.1 *Pentru o variabilă aleatoare continuă de densitate de probabilitate f_X, probabilitatea ca X să ia o valoare fixată $a \in \mathbb{R}$ este $P(X = a) = 0$.*

Într-adevăr, $P(X = a) = P(X \in [a, a]) = \int_a^a f_X(x)dx = 0$.

Precizăm că densitatea de probabilitate a unei variabile aleatoare, X, este fie o funcţie continuă, fie o funcţie cu un număr finit de puncte de discontinuitate de speţa întâi. În figura Fig.1 este ilustrat graficul unei densităţi de probabilitate continue, iar în Fig.2 graficul unei densităţi de probabilitate cu un punct de discontinuitate de speţa întâi.

Exemplul 1. Funcţia (Fig.2)

$$f(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ \dfrac{1}{2}e^{-x/2} & \text{dacă } x \geq 0 \end{cases}$$

este o densitate de probabilitate, deoarece ia valori ≥ 0 şi

$$\begin{aligned} \int_{-\infty}^{\infty} f(x)\,dx &= \int_{-\infty}^{0} 0\,dx + \int_0^\infty \frac{1}{2}e^{-x/2}\,dx = 0 + \lim_{u\to\infty}\int_0^u \frac{1}{2}e^{-x/2}\,dx = \\ &= \lim_{u\to\infty}\left(-\int_0^u -\frac{1}{2}e^{-x/2}\,dx\right) = -\lim_{u\to\infty} e^{-x/2}\,|_0^u = \\ &= -\lim_{u\to\infty}(e^{-u/2} - 1) = 1 \end{aligned}$$

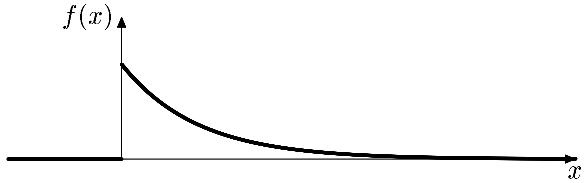

Fig.2: Densitate de probabilitate ce are un punct de discontinuitate de speţa întâi.

Funcţia de repartiţie a unei variabile aleatoare continue, X, având densitatea de probabilitate f_X este funcţia $F_X : \mathbb{R} \to \mathbb{R}$ definită astfel:

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)\, dx \tag{1}$$

Observaţia 1.1.2 *Funcţia de repartiţie F_X, a unei variabile aleatoare continue este continuă şi nedescrecătoare şi $\lim\limits_{x\to-\infty} F_X(x) = 0$, iar $\lim\limits_{x\to\infty} F_X(x) = 1$ (în timp ce pentru variabilele aleatoare discrete, funcţia de repartiţie este o funcţie nedescrescătoare, în scară, continuă la dreapta).*

În Fig.3 este ilustrată grafic semnificaţia geometrică a valorii $F_X(x)$.

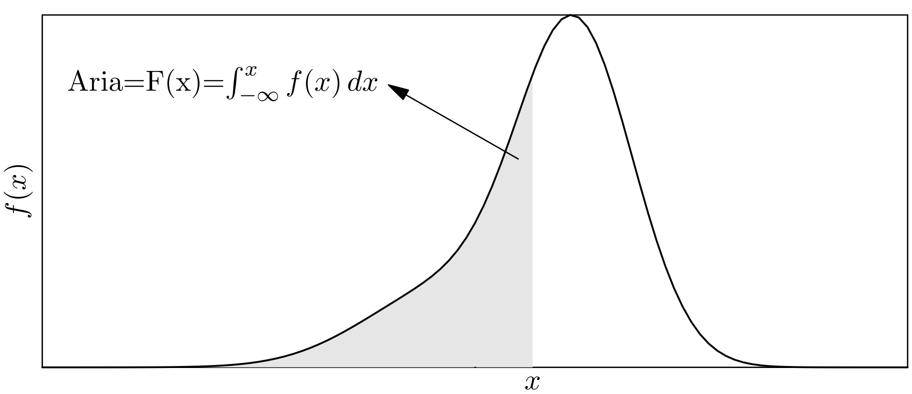

Fig.3: Ilustrarea semnificaţiei geometrice a valorii funcţiei de repartiţie într-un punct.

iar în Fig.8 graficul unei funcţii de repartiţie.

Să arătăm că putem calcula probabilitatea ca o variabilă aleatoare continuă să ia valori într-un interval I (I interval de capete $a < b$, cu a, b finite sau nu, şi aparţinând sau nu intervalului I) şi cu ajutorul funcţiei sale de repartiţie:

$$\begin{aligned} P(X \in I) &= \int_a^b f(x)\, dx = \int_{-\infty}^{b} f(x)\, dx - \int_{-\infty}^{a} f(x)\, dx = \\ &= F(b) - F(a) \end{aligned}$$

În cazul în care una dintre extremităţile intervalului este $\pm\infty$ notăm $F(-\infty) = \lim\limits_{x\to-\infty} F(x) = 0$, $F(\infty) = \lim\limits_{x\to\infty} F(x) = 1$.

Exemplul 2. Să determinăm funcţia de repartiţie a unei variabile aleatoare, X, ce are distribuţia de probabilitate dată de densitatea f, din Exemplul 1:

$$f(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ \frac{1}{2}e^{-x/2} & \text{dacă } x \geq 0 \end{cases}$$

Deoarece densitatea are o expresie pentru $x < 0$ şi o alta pentru $x \geq 0$, în calculul funcţiei de repartiţie, considerăm două cazuri:

$$F(x) = \int\limits_{-\infty}^{x} f(x)\,dx = \begin{cases} \int\limits_{-\infty}^{x} 0\,dx = 0 & \text{dacă } x < 0 \\ \int\limits_{-\infty}^{0} 0\,dx + \int\limits_{0}^{x} \frac{1}{2}e^{-x/2}\,dx = 1 - e^{-x/2} & \text{dacă } x \geq 0 \end{cases}$$

Astfel pentru o variabilă aleatoare ce are densitatea f, $P(X \in I)$, unde I este un interval arbitrar din $[0,\infty)$ se poate calcula în două moduri:

1) $P(X \in I) = \displaystyle\int_I \frac{1}{2}e^{-x/2}dx$;

2) $P(X \in I) = F(b) - F(a) = \dfrac{1}{2}(e^{-a/2} - e^{-b/2})$, unde $a < b$ sunt extremitătile intervalului I.

1.1.1 Distribuţia uniformă

Definiţia 1.1.2 *O variabilă aleatoare, X, continuă, ce are densitatea de probabilitate definită prin:*

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{dacă } x \in [a,b) \\ 0 & \text{dacă } x \notin (a,b) \end{cases} \tag{2}$$

se numeşte variabilă aleatoare uniform distribuită pe intervalul $[a,b)$ şi notăm $X \sim Unif[a,b)$.

Evident că alegerea intervalului $[a,b)$ este opţională, acesta putând fi înlocuit, de exemplu, cu $[a,b]$, $(a,b]$ sau (a,b).

Să calculăm funcţia de repartiţie asociată:

$$F(x) = P(X \leq x) = \int\limits_{-\infty}^{x} f(t)dt \tag{3}$$

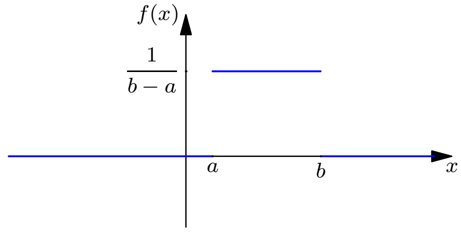

Fig.4: Densitatea de probabilitate pentru Unif$[a, b]$).

$$F(x) = \begin{cases} \int\limits_{-\infty}^{x} 0 \cdot dt = 0, \quad \text{dacă} \quad x < a \\ \int\limits_{-\infty}^{a} 0 \cdot dt + \int_a^x \frac{1}{b-a} dt = \frac{x-a}{b-a} \quad \text{dacă} \quad a \leq x < b \\ \int\limits_{-\infty}^{a} 0 \cdot dt + \int_a^b \frac{1}{b-a} dt + \int_b^x 0 \cdot dt = 1 \quad \text{dacă} \quad x \geq b \end{cases} \tag{4}$$

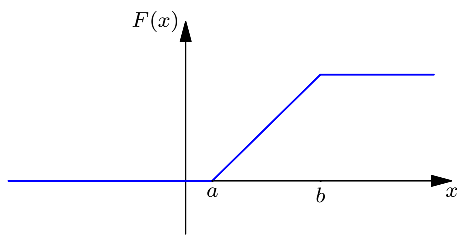

Fig.5: Funcţia de repartiţie pentru Unif$[a, b]$.

De ce se spune că o astfel de variabilă are distribuţia uniformă? Alegând orice subinterval $[x, x+\ell)$ de lungime ℓ din $[a, b)$, probabilitatea ca variabila aleatoare să ia valori într-un astfel de interval este:

$$P(x \leq X < x + \ell) = \int\limits_{x}^{x+\ell} f(x)\, dx = \int\limits_{x}^{x+\ell} \frac{1}{b-a} = \frac{\ell}{b-a}, \tag{5}$$

adică această probabilitate nu depinde de capetele intervalului, ci doar de lungimea lui. Deci valorile lui X sunt "uniform distribuite" în subintervalele din $[a, b]$ de aceeaşi lungime (Fig.6).

De interes pentru algoritmii de generare de numere (pseudo)aleatoare sunt variabilele aleatoare U, uniform distribuite pe $[0, 1)$, $U \sim$Unif[0,1). Densitatea de probabilitate a unei astfel de variabile este:

$$f_U(x) = \begin{cases} 1 & \text{dacă } x \in [0, 1) \\ 0 & \text{în rest,} \end{cases}$$

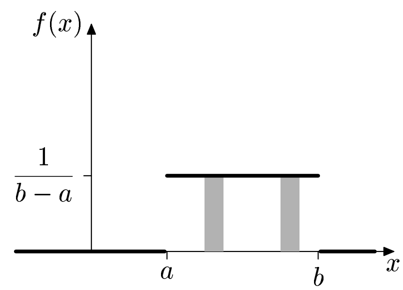

Fig.6: Distribuţia uniformă a valorilor în intervale de aceeaşi lungime.

iar funcţia de repartiţie

$$F_U(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ x & \text{dacă } x \in [0, 1) \\ 1 & \text{dacă } x \geq 1 \end{cases}$$

Probabilitatea ca o variabilă aleatoare U, uniform distribuită pe $[0, 1)$ să ia valori într-un subinterval $[c, d) \subset [0, 1)$ este egală cu lungimea, $d - c$, a intervalului:

$$P(c \leq U < d) = F(d) - F(c) = d - c$$

1.1.2 Distribuţia exponenţială

Definiţia 1.1.3 *O variabila aleatoare X ce are densitatea de probabilitate f, definită prin:*

$$f(x) = \begin{cases} 0 & \textit{dacă } x < 0 \\ \dfrac{1}{\theta} e^{-x/\theta} & \textit{dacă } x \geq 0 \end{cases}, \quad \theta > 0 \tag{6}$$

se numeşte variabilă aleatoare cu distribuţie exponenţială, de parametru θ.

Notăm cu $X \sim \mathrm{Exp}(\theta)$.

Funcţia de repartiţie a unei variabile $X \sim \mathrm{Exp}(\theta)$ este:

$$F(x) = \begin{cases} 0 & \text{dacă } x < 0 \\ 1 - e^{-x/\theta} & \text{dacă } x \geq 0 \end{cases}, \tag{7}$$

Variabilele aleatoare exponenţial distribuite se folosesc ca modele pentru:

- Durata servirii unui client, de către un server dintr-un sistem coadă;
- Intervalul de timp dintre două sosiri consecutive a clienţilor la o coadă ;
- Durata de viată a componentelor electronice;

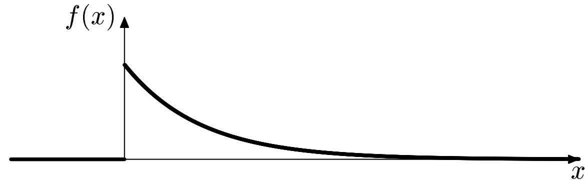

Fig.7: Densitatea de probabilitate exponenţială cu parametrul θ,

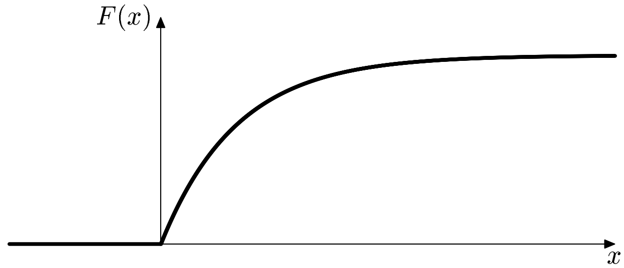

Fig.8: Funcţia de repartiţie a unei variabile aleatoare exponenţială de parametru θ,

1.2 Determinarea densităţii de probabilitate din functia de repartiţie

Dacă f_X este densitatea de probabilitate a variabilei aleatoare X, atunci funcţia sa de repartiţie este:

$$F_X(x) = \int_{-\infty}^{x} f_X(x)dx$$

Reciproc avem următoarea propoziţie:

Propoziţia 1.2.1 *Dacă densitatea de probabilitate, f, a unei variabile aleatoare, X, este continuă, atunci funcţia sa de repartiţie, F, este derivabilă şi $F'(x) = f(x)$, oricare ar fi $x \in \mathbb{R}$.*

Demonstraţie: Să arătăm că există limita de definiţie a derivatei într-un punct x_0 şi

$$\lim_{h\to 0} \frac{F(x_0+h) - F(x_0)}{h} = f(x_0)$$

Într-adevăr,

$$\begin{aligned} \lim_{h\to 0} \frac{F(x_0+h) - F(x_0)}{h} &= \lim_{h\to 0} \frac{1}{h} \left(\int_{-\infty}^{x_0+h} f(x)\,dx - \int_{-\infty}^{x_0} f(x)\,dx \right) = \\ &= \lim_{h\to 0} \frac{1}{h} \int_{x_0}^{x_0+h} f(x)\,dx \end{aligned}$$

Funcţia f fiind continuă admite primitive. Fie G o primitivă a sa, adică $G'(x) = f(x)$, $\forall\, x \in \mathbb{R}$. Astfel:

$$\lim_{h\to 0}\frac{F(x_0+h)-F(x_0)}{h} = \lim_{h\to 0}\frac{1}{h}G(x)|_{x_0}^{x_0+h} = \lim_{h\to 0}\frac{G(x_0+h)-G(x_0)}{h} =$$

$$= G'(x_0) = f(x_0)$$

Deci funcţia de repartiţie F este derivabilă în x_0 şi $F'(x_0) = f(x_0)$, $\forall\, x_0 \in \mathbb{R}$. □

Exploatând această proprietate avem:

$$f(x) = \lim_{h\to 0}\frac{F(x+h)-F(x)}{h} = \lim_{h\to 0}\frac{P(x < X \leq x+h)}{h} \tag{8}$$

Această egalitate ne permite să scriem $f(x)h = P(x < X \leq x+h) + O(h)$, cu $\lim_{h\to 0}\dfrac{O(h)}{h} = 0$, adică $f(x)h \simeq P(x < X \leq x+h)$ ($\simeq$ se citeşte "aproximativ egal"). În inginerie se foloseşte scrierea dx pentru un h foarte mic: $f(x)dx \simeq P(x < X \leq x+dx)$, se interpretează ca probabilitatea ca variabila aleatoare X, f–distribuită, să ia valori în intervalul $(x, x+dx)$ este aproximată de $f(x)dx$.

Observaţia 1.2.1 *În cazul în care densitatea de probabilitate are un număr finit, $x_1, x_2, \ldots, x_n$, de puncte de discontinuitate de speţa întâi, atunci funcţia sa de repartiţie este continuă pe $\mathbb{R}$ şi derivabilă pe porţiuni, adică este derivabilă pe intervalele $(-\infty, x_1)$, (x_1, x_2), $\ldots$, (x_{n-1}, x_n), (x_n, ∞), notate în continuare J şi derivata funcţiei de repartiţie restricţionată la J coincide cu densitatea de probabilitate, restricţionată la intervalul J:*

$$(F_{|J})' = f_{|J}$$

Bazat pe această observaţie putem determina densitatea de probabilitate f a unei variabile aleatoare X, ce are funcţia de repartiţie F continuă şi derivabilă pe porţiuni, şi anume:

$$\begin{array}{l} f_{|(-\infty,x_1)} = F'_{|(-\infty,x_1)} \\ f_{|(x_j,x_{j+1})} = F'_{|(x_j,x_{j+1})}, j = \overline{0, n-1} \\ f_{|(x_n,\infty)} = F'_{|(x_k,\infty)} \end{array}$$

În punctele x_j, $j = \overline{1, n}$, funcţia f se prelungeşte astfel încât să fie ori continuă la stânga ori la dreapta.

Exemplul 3. Funcţia de repartiţie a unei variabile aleatoare X este (Fig.9, stânga)

$$F(x) = \begin{cases} 0 & \text{dacă } x \in (-\infty, a) \\ \dfrac{x-a}{b-a} & \text{dacă } x \in [a, b) \\ 1 & \text{dacă } x \geq b \end{cases}$$

Funcţia de repartiţie este nederivabilă în a şi b. Deci densitatea sa de probabilitate va fi:

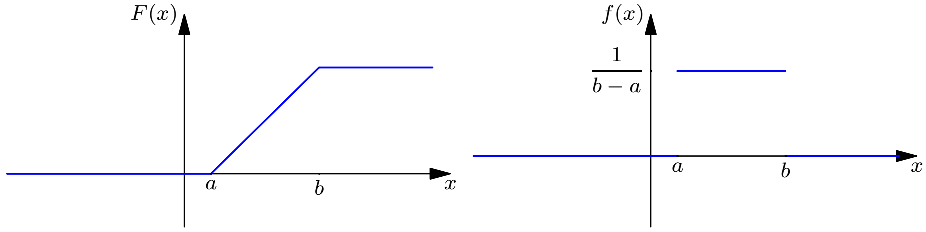

Fig.9: Funcţie de repartiţie derivabilă pe porţiuni (stânga) şi densitatea sa de probabilitate (dreapta)

$$f(x) = (0)' = 0 \text{ pentru } x \in (-\infty, a)$$
$$f(x) = 1/(b-a) \text{ pentru } x \in (a, b)$$
$$f(x) = (1') = 0 \text{ pentru } x \in (b, \infty)$$

În punctele a şi b densitatea se prelungeşte ”aranjând” să fie continuă la dreapta sau la stânga. De exemplu luăm $f(a) = 1/(b-a)$, şi $f(b) = 1/(b-a)$, caz în care f este densitatea unei variabile aleatoare, ce ia valori în $[a, b]$. Luând $f(a) = 1/(b-a)$, $f(b) = 0$, f este densitatea unei variabile aleatoare ce ia valori în $[a, b)$.

1.3 Media şi dispersia unei variabile aleatoare continue

Fie X o variabilă aleatoare continuă ce are densitatea de probabilitate f. Dacă funcţia g, $g(x) = xf(x)$, este integrabilă pe $\mathbb{R}$, atunci integrala sa $\int_{-\infty}^{\infty} xf(x)\,dx$ se numeşte *valoarea medie* a variabilei şi se notează $M(X)$ sau $E(X)$.

Exemplul 4. Fie X o variabilă aleatoare exponenţial distribuită de parametru, $\theta > 0$, ce are densitatea de probabilitate:

$$f(x) = \begin{cases} \dfrac{1}{\theta} e^{-x/\theta} & \text{dacă } x > 0 \\ 0 & \text{în rest} \end{cases} \tag{9}$$

Calculând integrala prin părţi, avem:

$$M(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{-\infty}^{0} x \cdot 0\,dx + \int_{0}^{\infty} x \frac{1}{\theta} e^{-x/\theta}\,dx = \theta. \tag{10}$$

Există şi variabile aleatoare care nu au valoare medie. De exemplu, pentru variabila aleatoare X, ce are distribuţia de probabilitate Cauchy, $f(x) = \dfrac{1}{\pi}\dfrac{1}{1+x^2}$, $\forall\, x \in \mathbb{R}$, integrala $\int_{-\infty}^{\infty} xf(x)\,dx$ este divergentă.

Transformări ale unei variabile aleatoare

Pornind de la o variabilă aleatoare putem defini variabile asociate în felul următor:

Propoziţia 1.3.1 *Fie X o variabilă aleatoare ce are densitatea de probabilitate f. Dacă $g : \mathbb{R} \to \mathbb{R}$ este o funcţie continuă sau cu un număr finit de puncte de discontinuitate de speţa întâi, atunci $Y = g(X)$ este o variabilă aleatoare. Dacă $\int_{-\infty}^{\infty} |g(x)|f(x)dx < \infty$ atunci variabila aleatoare $Y = g(X)$ are media*

$$M(g(X)) = \int_{-\infty}^{\infty} g(x)f(x)\,dx \tag{11}$$

Dispersia (varianţa) unei variabile aleatoare continue X ce are media $m = M(X)$ este ca şi în cazul variabilelor discrete, numărul notat $\sigma^2(X)$ sau $D^2(X)$, şi egal cu:

$$\sigma^2(X) = M((X - m)^2) \tag{12}$$

Notând cu g funcţia definită prin $g(X) = (X - m)^2$, rezultă că dacă f este densitatea de probabilitate a variabilei X atunci dispersia sa se calculează astfel:

$$\sigma^2(X) = \int_{-\infty}^{\infty} (x - m)^2 f(x)\,dx.$$

Ca şi în cazul mediei, dacă integrala $\int_{-\infty}^{\infty}(x - m)^2 f(x)\,dx$ este divergentă, variabila aleatoare X nu are dispersie. Teoria probabilităţilor studiază preponderent variabile aleatoare cu dispersie finită. În ultimii ani s-a constatat însă că traficul neregulat al datelor prin reţelele de comunicaţie, poate fi modelat de variabile aleatoare cu medie finită şi dispersie infinită. Astfel vom studia câteva particularităţi şi ale acestei clase de variabile aleatoare.

Abaterea standard a unei variabile aleatoare, X, de dispersie finită este $\sigma(X) = \sqrt{\sigma^2(X)}$.

Reguli de calcul a mediei şi dispersiei

1. $M(aX + b) = aM(X) + b$, $\forall\ a, b \in \mathbb{R}$.
Într-adevăr luând $g(x) = ax + b$, $M(aX + b) = \int_{-\infty}^{\infty}(ax + b)f(x)\,dx = a\int_{-\infty}^{\infty} xf(x)\,dx + b\int_{-\infty}^{\infty} f(x)\,dx = aM(X) + b$.

2. $\sigma^2(aX + b) = a^2\sigma^2(X)$, $\forall\ a, b \in \mathbb{R}$.
Notând $m = M(X)$, variabila aleatoare $aX + b$ are media $am + b$ şi deci:

$$\sigma^2(aX + b) = M((aX + b - am - b)^2) = M(a^2(X - m)^2) = a^2\sigma^2(X).$$

3. $\sigma^2(X) = M(X^2) - [M(X)]^2$ (proprietate folosită în exerciţii)

Definiţia 1.3.1 *Fie X o variabilă aleatoare de medie m şi abatere standard σ. Variabila aleatoare asociată, $Y = \dfrac{X - m}{\sigma}$, se numeşte variabilă aleatoare standardizată.*

Media variabilei aleatoare standardizate, asociate variabilei aleatoare X este 0, iar dispersia 1. Într-adevăr, media este $M(Y) = \frac{1}{\sigma}M(X) - \frac{m}{\sigma} = 0$, iar dispersia $D^2(Y) = \frac{1}{\sigma^2}D^2(X-m) = \frac{1}{\sigma^2}D^2(X) = 1$.

1.4 Distribuţia Pareto

Opţional. Distribuţia Pareto şi distribuţia discretă Zipf se numesc distribuţiile internetului, pentru că aceste legi de probabilitate guvernează multe caracteristici aleatoare ale reţelei internet şi a celei WWW. Distribuţia Pareto poartă numele economistului şi sociologului italian, Vilfredo Pareto, care a identificat-o încercând să găsească un model pentru venitul indivizilor într-o economie.

O variabilă aleatoare, X, a cărei densitate de probabilitate, depinzând de parametrii $\alpha, \beta > 0$, este:

$$f(x;\alpha,\beta) = \begin{cases} 0 & \text{pentru } x < \beta \\ \frac{\alpha}{\beta}\left(\frac{\beta}{x}\right)^{\alpha+1} & \text{pentru } x \geq \beta, \end{cases} \tag{13}$$

se numeşte variabilă aleatoare Pareto. Clasa variabilelor aleatoare Pareto se notează Pareto(α,β). Pentru parametrul β fixat, densitatea corespunzătoare unui parametru α mai mare

Funcţia de repartiţie a unei variabile aleatoare $X \sim$ Pareto(α,β) este:

$$F(x;\alpha,\beta) = \begin{cases} 0 & \text{dacă } x < \beta \\ 1 - \left(\frac{\beta}{x}\right)^{\alpha} & \text{dacă } x \geq \beta \end{cases}, \tag{14}$$

Parametrul α al distribuţiei Pareto nu are dimensiune (unitate de măsură), în timp ce parametrul β se exprimă în aceleaşi unităţi de măsură ca şi valorile variabilei $X \sim$ Pareto(α,β). Parametrul α este un parametru de alură, adică influenţează alura graficului densităţii de probabilitate. În versiunea în care a fost introdusă de Pareto, dacă X dă venitul mediu al unui individ într-o economie, atunci β reprezintă venitul minim pe economie şi deci o variabilă Pareto ia valori mai mari sau egale cu parametrul β.

Să calculăm media şi dispersia unei variabile Pareto:

$$\begin{aligned} M(X) &= \int_{-\infty}^{\infty} xf(x)\,dx = \alpha\beta^{\alpha}\int_{\beta}^{\infty} x^{-\alpha}\,dx = \frac{\alpha\beta^{\alpha}}{-\alpha+1}x^{-\alpha+1}|_{\beta}^{\infty} = \\ &= \frac{\alpha\beta^{\alpha}}{-\alpha+1}\left(\lim_{x\to\infty}\frac{1}{x^{\alpha-1}} - \beta^{-\alpha+1}\right) \end{aligned}$$

Observăm că limita $\lim_{x\to\infty}\frac{1}{x^{\alpha-1}}$ există doar pentru $\alpha > 1$. Prin urmare o variabilă aleatoare Pareto de parametru $\alpha > 1$ are media:

$$M(X) = \frac{\alpha\beta}{\alpha-1}, \quad \alpha > 1, \beta > 0 \tag{15}$$

Să calculăm şi dispersia: $\sigma^2(X) = M(X^2) - (M(X))^2$.

$$M(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \frac{\alpha\beta^2}{\alpha - 2}, \quad \alpha > 2, \beta > 0$$

Deci o variabilă Pareto are dispersie finită doar pentru $\alpha > 2$ şi aceasta este:

$$\sigma^2(X) = \frac{\alpha\beta^2}{\alpha - 2} - \frac{\alpha^2\beta^2}{(\alpha - 1)^2} = \frac{\alpha\beta^2}{(\alpha - 1)^2(\alpha - 2)} \tag{16}$$

Distribuţia Pareto a fost identificată ca fiind distribuţia unor variabile ce caracterizează performanţa sistemelelor de calcul, traficul internet şi alţi parametri:

- Durata serviciului procesorului pentru procesele sistemului de operare UNIX, s-a dovedit prin experimentare că are distribuţia Pareto.
- Dimensiunea fişierelor stocate pe serverele WEB are distribuţia Pareto de parametru $\alpha \in [1.1, 1.3]$.
- Lungimea intervalului de timp, X, între două pachete succesive de informaţie în traficul neregulat al datelor, în sistemele de comunicaţie, este o variabilă Pareto de parametru $\alpha \in (1, 2]$ (deci o variabilă de medie finită şi dispersie infinită). Parametrul β reprezintă intervalul de timp minim între două pachete, iar α caracterizează intensitatea utilizării reţelei.
- Distribuţia Pareto se foloseşte ca model pentru simularea traficului internet. Această simulare este necesară pentru a studia parametrii reţelei internet în condiţii de trafic intens.