MS

(curs 9 - S9)

18.04.2024

O pereche (X, Y) sau, mai general, un n–uplu $(X_1, X_2, \dots, X_n)$ de variabile aleatoare continue se numește **vector aleator continuu**.
Vectori aleatori continui: în experimente în care se observă sau se măsoară simultan n caracteristici ale sistemului sau procesului supus studiului.

Ce ne interesează, relativ la vectorul aleator (X, Y)?

Răspuns:
$P((X, Y) \in D)$: evenimentul: "(X, Y) ia valori în domeniul $D \subset \mathbb{R}^2$".
Cele mai uzuale domenii sunt cele dreptunghiulare: $D = [a, b] \times [c, d]$, $D = [a, b) \times [c, d)$ etc.

Densitatea de probabilitate

$x \backslash y$	$\rightarrow y_i$
$\downarrow x_i$	p_{ij}

Densitatea de probabilitate a unui vector continu este o funcție $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$ ce verifică:

1) $f_{X,Y}(x, y) \geq 0, \;\; (x, y) \in \mathbb{R}^2$;
2) $f_{X,Y}$ este integrabilă pe $\mathbb{R}^2$
3) $\int_{\mathbb{R}^2} f_{X,Y}(x, y) dx\, dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dx\, dy = 1$.

Probabilitatea ca vectorul (X, Y) să ia valori în domeniul[1] $D \subset \mathbb{R}^2$ este integrala dublă pe D din densitate, adică

$$P((X, Y) \in D) = \iint_D f_{X,Y}(x, y)\, dx\, dy.$$

[1] Domeniul de integrare este în problemele practice o mulțime mărginită sau nemărginită pe care integrala există.

$$\int f(x)\,dx = 1 \quad \longleftrightarrow \quad \iint_{\mathbb{R}^2} f(x,y)\,dx\,dy = 1$$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dx\,dy = 1$$

! $P(X \in [a,b]) = \int_a^b f(x)\,dx = F_X(b) - F_X(a)$

$P((x,y) \in B) = \iint_B f(x,y)\,dx\,dy$, $B \subset \mathbb{R}^2$

Variabile aleatoare continue

- **Valori:** $I \subset \mathbb{R}$
- **p.d.f** : $f_X(x) \geq 0, \;\; x \in \mathbb{R}$, $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.
- **Calcul prob**: $P(X \in I) = \int_I f_X(x) dx$
- **Funcția de repartiție:** $F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f(t) dt$
- $P(X \in [a, b)) = F_X(b) - F_X(a)$,

Vectori aleatori continui

- $D \subset \mathbb{R}^2$
- p.d.f: $f : \mathbb{R}^2 \to \mathbb{R} : \; f(x, y) \geq 0$, $\int_{\mathbb{R}^2} f_{X,Y}(x, y) dx\, dy = 1$.
- $P((X, Y) \in D) = \int_D f_{X,Y}(x, y)\, dx\, dy$.
- $F_{X,Y}(x, y) = P(X \leq x, Y \leq y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{X,Y}(t, s)\, dt\, ds$.
- $P(x_1 < X \leq x_2, y_1 < Y \leq y_2) = F_{X,Y}(x_2, y_2) - F_{X,Y}(x_1, y_2) - F_{X,Y}(x_2, y_1) + F_{X,Y}(x_1, y_1)$.

Variabile aleatoare continue **Vectori aleatori continui**

- **Media**
  $M(X) = \int_{-\infty}^{\infty} x f_X(x)\,dx$

$$LOTUS : M(g(X,Y)) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y) f_{X,Y}(x,y)\,dx\,dy.$$

$$M(X+Y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y) f_{X,Y}(x,y)\,dx\,dy = M(X) + M(Y)$$
$$M(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\,y\,f_{X,Y}(x,y)\,dx\,dy \neq M(X)M(Y)$$

$\iint_D 1\,dx\,dy = \text{Aria}(D)$

$\iint_{Disc} 1\,dx\,dy = \text{Aria}(\text{Disc})$

$\text{Disc} = x^2 + y^2 \leq R^2$

? $\iint 1\,dx\,dy = \pi \cdot \quad = \text{Aria}(\text{Disc eliptic})$

Disc eliptic : $\frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1$

Idei principale legate de integrala dublă[2] :

- exemple de domenii $D \subset \mathbb{R}^2$: dreptunghi, triunghi, disc circular/eliptic
- dacă $f : D \subset \mathbb{R}^2 \to \mathbb{R}$, notăm $\int_D f(x,y)dx\,dy$-integrala dublă
- interpretare: $\int_D 1 dx\,dy = Aria(D)$
- cel mai simplu caz: $D = [a,b] \times [c,d]$

$$\iint_D f(x,y)dx\,dy = \int_c^d \left(\int_a^b f(x,y)dx\right)dy = \int_a^b \left(\int_c^d f(x,y)dy\right)dx$$

- in cazul D-disc circular/eliptic se face schimbare de variabile
- $f(x,y) = c$, atunci $\int_D f(x,y)dx\,dy = c \cdot Aria(D)$

Fie (X, Y) un vector aleator ce are densitatea

$$f_{X,Y}(x,y) = \begin{cases} 2-y, & \text{dacă } x \in [0,2],\ y \in [1,2], \\ 0, & \text{în rest.} \end{cases}$$

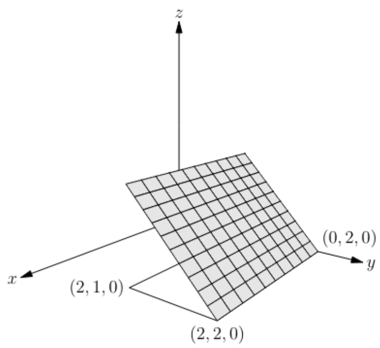


Figure: Graficul unei densități de probabilitate cu suport mărginit.

$G = [0,2] \times [1,2] = \mathrm{supp}(f_{x,y}) \longrightarrow \mathbb{R}^2 \qquad f_{x,y} \neq 0$

f - densitate de probabilitate pentru că

$$f_{x,y}(x,y) \geq 0, \quad \forall\, (x,y) \in \mathbb{R}^2$$

$$\int_{\mathbb{R}^2} f_{x,y}(x,y)\, dx dy = \int_1^2 \int_0^2 2-y \, dx\, dy = \int_1^2 \left(2x - yx\right)\Big|_0^2 dy =$$

$$\int_1^2 (4-2y) dy = (4y - y^2)\Big|_1^2 = 8-4-4+1 = 1$$

Fie discul $D = \{(x,y) \mid (x-1)^2 + (y-1.5)^2 \leq 0.5^2\} \subset G$.
Dorim să calculăm
$P((X,Y) \in D) = \int_D f_{X,Y}(x,y)\, dx\, dy = \int_D (2-y)\, dx\, dy$.
Facem schimbarea de variabile:

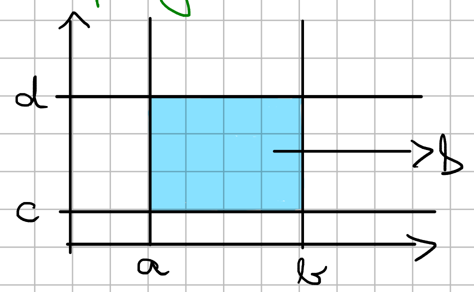
$$\begin{cases} x &= 1 \quad + \rho \cos\theta, \\ y &= 1.5 + \rho\ \sin\theta, \end{cases}$$

unde $\rho \in (0, 0.5]$ și $\theta \in [0, 2\pi)$, avem

$$\int\int_D (2-y)\, dx\, dy = \int_0^{0.5} \left(\int_0^{2\pi} (2 - 1.5 - \rho \sin\theta)\rho\, d\theta \right) d\rho = \int_0^{0.5} \pi\rho\, d\rho.$$

Obținem $P((X,Y) \in D) = \frac{\pi}{8}$.

Dreptunghi



$$\text{Aria}(D) = (b-a)\cdot(d-c) =$$

$$D = [a,b]\times[c,d]$$

$$= \int_c^d \left(\int_a^b 1\,dx\right)dy = \int_c^d (b-a)\,dy =$$

$$= (b-a)\int_c^d 1\,dy = (b-a)(d-c)$$

Vectori discreti

- **Distribuții marginale**: distrb lui X si Y din tablou, prin însumare linii/coloane
- **Funcțiile de repartiție marginală** F_X si $F_Y(y)$ se obțin din tabloul de lui X si al lui Y;
- X, Y independente dacă $F_{(X,Y)} := F_X(x)F_Y(y)$ $p_{X,Y}(x_i, y_j) = p_X(x_i)p_Y(y_j)$

Vectori continui

- **densitatea marginală a lui X** $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dy$,
- **densitatea marginală a lui Y** $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)\,dx$.
- $F_X(x) = P(X \le x) = P(X \le x, Y \in \mathbb{R}) = \int_{-\infty}^{x} f_X(t)\,dt$.
- X, Y independente dacă $f_{(X,Y)} = f_X(x)f_Y(y)$

Varb. cond. discrete

- **Distribuția variabilei condiționate**: $(X|Y = y_j)$ se bazează pe determinarea $P(X = x_i|Y = y_j) = \frac{p_{ij}}{P(Y=y_j)}$.
- **Cond. de independență**: X, Y independente dacă $P(X = x_i|Y = y_j) = P(X = x_i)$ sau $P(Y = y_j|X = x_i) = P(Y = y_j)$.

Varb. cond. continue

- **Densitatea variabilei condiționate** $(X|Y = y_0)$ este $g(x|y_0) = \frac{f_{X,Y}(x,y_0)}{f_Y(y_0)}$
- **Densitatea variabilei condiționate** $(Y|X = x_0)$ este $h(y|x_0) = \frac{f_{X,Y}(x_0,y)}{f_X(x_0)}$
- dacă X, Y independente, atunci $g(x|y) = f_X(x), \forall x \in \mathbb{R}$; $h(y|x) = f_Y(y), \forall y \in \mathbb{R}$

Recomandare: de citit din notițele de curs formula lui Bayes pentru densități de probabilitate.

Varb. a. cont. unif.

- $I = [a,b)$
- **P.d.f.**: $f(x) = \begin{cases} \frac{1}{b-a} & x \in [a,b) \\ 0 & x \notin (a,b) \end{cases}$
- **Notatie**: $X \sim \text{Unif}[a,b)$

Vect. a cont. unif.

- $D = [a,b] \times [c,d]$
- $f_{X,Y}(x,y) = \begin{cases} \frac{1}{\text{aria}(D)} = \frac{1}{(b-a)(d-c)}, & (x,y) \in D, \\ 0, & \text{în rest.} \end{cases}$
- $(X,Y) \sim \text{Unif}([a,b] \times [c,d])$
- **densități marginale** $f_X(x) = \begin{cases} \frac{1}{b-a}, & \text{dacă } x \in [a,b], \\ 0, & \text{în rest,} \end{cases}$
- $f_Y(y) = \begin{cases} \frac{1}{d-c}, & \text{dacă } y \in [c,d], \\ 0, & \text{în rest.} \end{cases}$
- $f_{X,Y}(x,y) = f_X(x)\,f_Y(y)$

Proprietate

- dacă $(X,Y) \sim Unif([a,b] \times [c,d])$, atunci $X \sim Unif[a,b]$ $Y \sim Unif[c,d]$ si sunt independente.
- dacă $X \sim Unif[a,b]$ și $Y \sim Unif[c,d]$ sunt variabile aleatoare independente, atunci $(X,Y) \sim Unif[a,b] \times [c,d]$.

Metodă de generare de puncte (x,y) **uniform distribuite în dreptunghiul** $D = [a,b] \times [c,d]$

```
x ← a + (b − a) ∗ urand();
y ← c + (d − c) ∗ urand();
return (x,y);
```

Cum procedăm dacă dorim să generăm puncte aleatoare în domenii G mărginite, ne-dreptunghiulare și de arie nenulă?

- se fixează un dreptunghi $D = [a,b] \times [c,d]$ ce include domeniul G;
- se generează la întâmplare puncte în D
- se rețin doar cele care aparțin lui G:

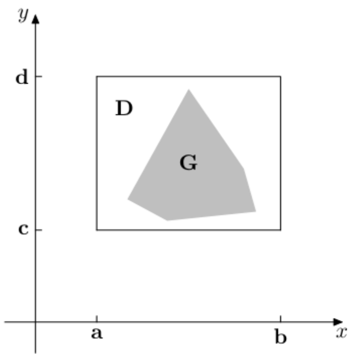


Figure: Regiunea G inclusă într-un dreptunghi $D = [a,b] \times [c,d]$.

```
do{
   generează un punct (x, y) în D;
  }while((x, y) nu aparține lui G);
return(x,y);
```

Cum trebuie să fie D pentru ca algoritmul să fie cât mai eficient?

- N numărul de parcurgeri ale buclei `do-while`
- $N \sim Geom(p)$
- $p \in (0,1)$ probabilitatea ca punctul generat să cadă în G.
- $p = P((X,Y) \in G) = \iint_G f_{X,Y}(x,y)\,dx\,dy = \frac{1}{\text{aria}(D)} \iint_G 1\,dx\,dy = \frac{\text{aria}(G)}{\text{aria}(D)}$.
- $M(N) = 1/p = \text{aria}(D)/\text{aria}(G)$,

Recomandare: să se folosească pentru generatorul de puncte din G acel dreptunghi D cu laturile paralele cu axele de coordonate și arie minimă, care include domeniul G.

Pentru a genera puncte uniform distribuite în discul eliptic

$$G = \left\{(x, y) \in \mathbb{R}^2 \,:\, \frac{x^2}{9} + \frac{y^2}{16} \leq 1\right\},$$

se alege dreptunghiul $D = [-3, 3] \times [-4, 4]$, de arie minimă, ce include discul eliptic G.
Generatorul de puncte aleatoare în G funcționează astfel:

```
do{
  x = -3 + 6 * urand();
  y = -4 + 8 * urand();
}while(x * x/9 + y * y/16 > 1);
return (x, y);
```