13.12.2021.

~CURS 10~

Repartiția normală (Gaussiană)

Def: Spunem că o v.a. X este repartizată normal (sau Gaussian) de parametrii μ și σ^2, notăm $X \sim N(\mu, \sigma^2)$, dacă admite densitatea de repartiție

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

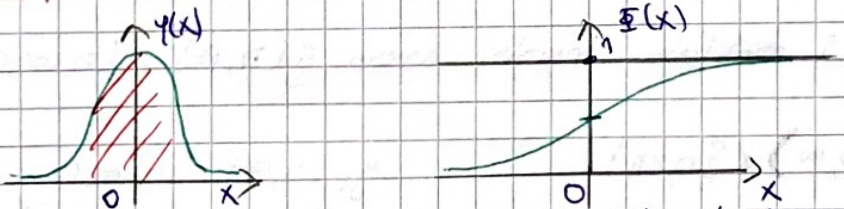
În cazul particular în care $\mu = 0$ și $\sigma^2 = 1$, spunem că v.a. X este repartizată normal standard, $X \sim N(0,1)$.

În acest caz densitatea de repartiție se notează:

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}$$

iar funcția de repartiție

Nu are formă compactă! → $$\Phi(x) = \int_{-\infty}^{x} \varphi(t)\,dt = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$



Obs. că φ este simetrică față de 0: $\varphi(z) = \varphi(-z), \forall z \in \mathbb{R}$ pt că φ este fct pară

Să verificăm că φ este o densitate de probabilitate:

a) $\varphi(x) \geq 0, \forall x \in \mathbb{R}$ (adev)

b) $\int_{-\infty}^{+\infty} \varphi(x)\,dx = 1$

Să se arate că $\int_{-\infty}^{+\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}$. Fie $I = \int_{-\infty}^{+\infty} e^{-x^2/2}\,dx$

Calculăm $I^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2/2}\,dx\right)^2 = \left(\int_{-\infty}^{+\infty} e^{-x^2/2}\,dx\right) \times \left(\int_{-\infty}^{+\infty} e^{-y^2/2}\,dy\right)$

$$\overset{(\text{Fubini})}{=} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}}\,dx\,dy$$

Transformăm în coordonate polare.

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$$

$g : (0,\infty) \times [0, 2\pi] \to \mathbb{R}^2$

$(r,\theta) \longrightarrow (x,y)$

$g(r,\theta) = (r\cos\theta, r\sin\theta) = (x,y)$

(T) Fie $G \subseteq \mathbb{R}^n$ o mulțime deschisă și $g: G \to \mathbb{R}^n$ o funcție diferențială și cu derivate parțiale continue. Presupunem că funcția g este injectivă și determinantul matricii Jacobiene a lui g este nenul.

$$J_g = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}$$

$g = (g_1, \dots g_n)$

$g(r,\theta) = (\underbrace{r\cos\theta}_{g_1}, \underbrace{r\sin\theta}_{g_2})$

Atunci pentru f a.î. f pe G este pozitivă sau integrabilă.

$$\int_{g(G)} f(y)\,dy = \int_G f(g(x))\;|\det J_g(x)|\,dx$$

În cazul problemei noastre avem $g(r,\theta) = (r\cos\theta, r\sin\theta)$

$G = [0,\infty) \times [0,2\pi)$ $\qquad g_1(r,\theta) = r\cos\theta$

$g(G) = \mathbb{R}^2$ $\qquad g_2(r,\theta) = r\sin\theta$

$$J_g(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial r} & \frac{\partial g_1}{\partial \theta} \\ \frac{\partial g_2}{\partial r} & \frac{\partial g_2}{\partial \theta} \end{pmatrix}$$

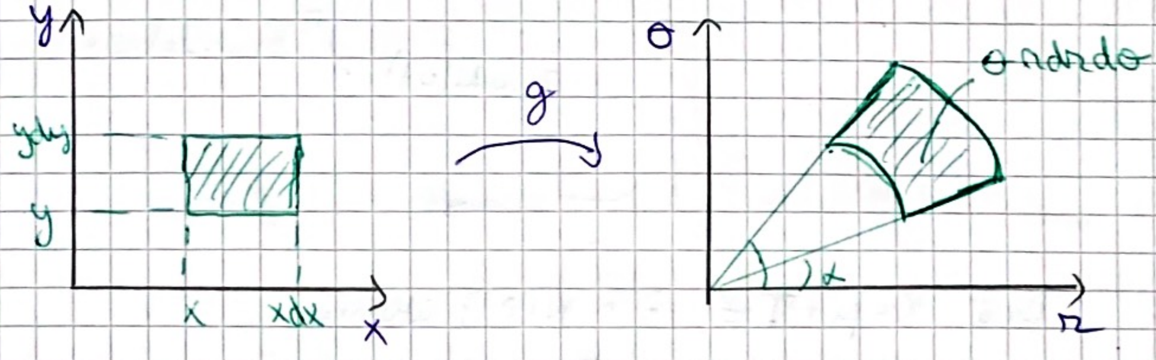
$$\det J_g(r,\theta) = r\cos^2\theta + r\sin^2\theta = r(\cos^2\theta + \sin^2\theta) = r$$

$$I^2 = \underbrace{\int_{\mathbb{R}^2}}_{g(G)}\int \underbrace{e^{-\frac{x^2+y^2}{2}}}_{f(x,y)}\,dx\,dy = \iint_{G=[0,\infty)\times[0,2\pi)} f(g(r,\theta))\;|\det J_g(r,\theta)|\,dr\,d\theta$$

$$= \iint_{[0,\infty)\times[0,2\pi)} f(r\cos\theta, r\sin\theta)\, r\,dr\,d\theta$$

$$= \int_0^\infty \int_0^{2\pi} e^{-r^2/2}\, r\,d\theta\,dr = \int_0^\infty 2\pi r e^{-r^2/2}\,dr = 2\pi \underbrace{\int_0^\infty r e^{-r^2/2}\,dr}_{1} = 2\pi$$

$$\Rightarrow I^2 = 2\pi \Rightarrow I = \sqrt{2\pi} \Rightarrow \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = 1$$



Dacă $X \sim N(0,1)$ atunci $\mathbb{E}\{X\}=0$ și $Var(X)=1$:

$$\mathbb{E}\{X\} = \int_{-\infty}^{+\infty} x\varphi(x)\,dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \underbrace{x e^{-x^2/2}}_{\text{funct. impară}}\,dx = 0$$

$$Var(X) = \mathbb{E}\{X^2\} - \mathbb{E}\{X\}^2 = \mathbb{E}\{X^2\}$$

$$\mathbb{E}\{X^2\} = \int_{-\infty}^{+\infty} x^2\varphi(x)\,dx = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} x^2 e^{-x^2/2}\,dx$$

Integrăm prin părți

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} x\left(-e^{-x^2/2}\right)'\,dx$$

$$= \frac{1}{\sqrt{2\pi}} \left\{ -x e^{-x^2/2} \Big|_{-\infty}^{+\infty} + \int_{-\infty}^{+\infty} \underbrace{x'}_{=1} e^{-x^2/2}\,dx \right\}$$

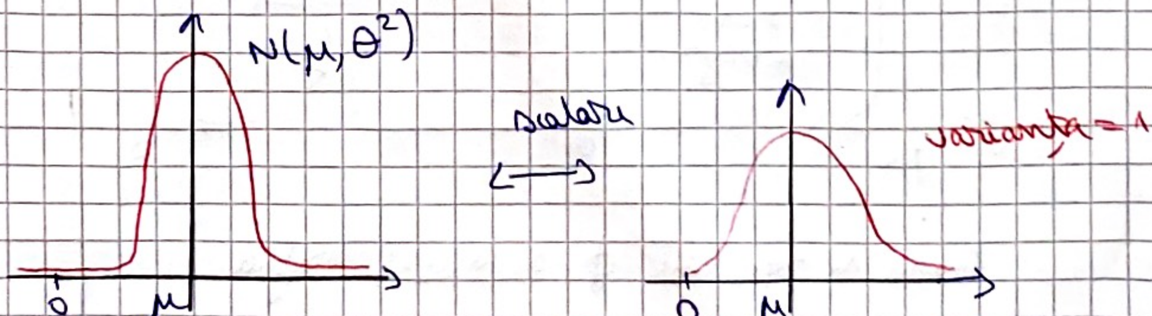
$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-x^2/2}\,dx = 1$$

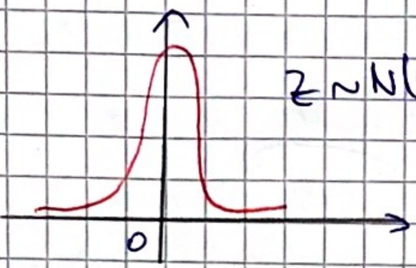
Astfel $Var(X)=1$

- Pt o $N(0,1)$ (normală standard) avem că $\mathbb{E}\{X\}=0$ și $Var(X)=1$.
- În general avem $X \sim N(\mu, \sigma^2)$ atunci: $\mathbb{E}\{X\}=\mu$ și $Var(X)=\sigma^2$.
- V.a. X se poate scrie sub forma

$$X = \mu + \sigma Z \quad \text{unde } Z \sim N(0,1)$$

scalare
translatare





Dacă $X=\mu+\sigma Z$, $Z\sim N(0,1)$ atunci

$$E\{X\}=E\{\mu+\sigma Z\}=\mu+\sigma \underbrace{E\{Z\}}_{=0}=\mu$$

$$Var(X)=Var(\mu+\sigma Z)=\sigma^2\underbrace{Var(Z)}_{=1}=\sigma^2$$

$$F(x)=P(X\le x)=P(\mu+\sigma Z\le x)$$
$$=P\left(Z\le \frac{x-\mu}{\sigma}\right)=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

Relația între $f(x)$ și $\varphi(x)$

$$(f\circ g)'=f'(g)\cdot g'$$

$$f(x)=\frac{d}{dx}F(x)=\frac{d}{dx}\Phi\left(\frac{x-\mu}{\sigma}\right)=\frac{d\Phi}{dx}\left(\frac{x-\mu}{\sigma}\right)\cdot\frac{d}{dx}\left(\frac{x-\mu}{\sigma}\right)$$
$$=\varphi\left(\frac{x-\mu}{\sigma}\right)\cdot\frac{1}{\sigma}$$

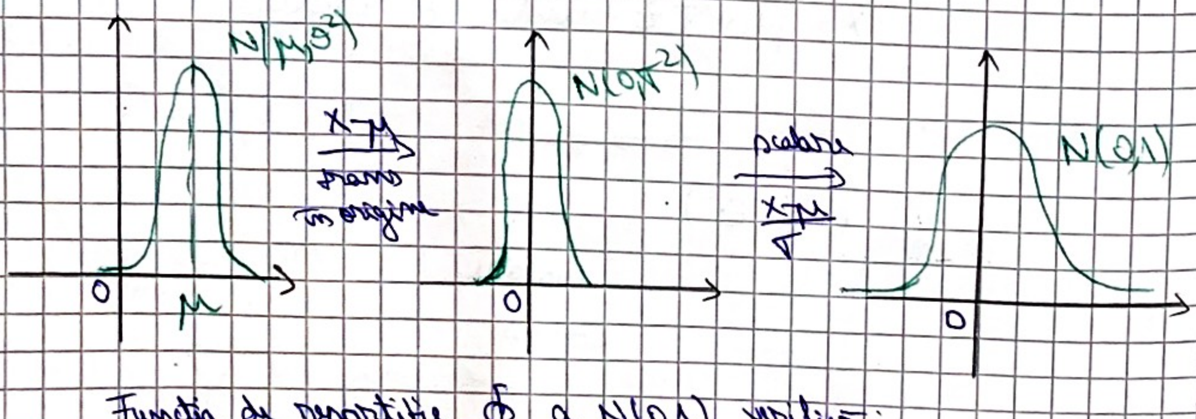
· Procedeul de standardizare: $X\sim N(\mu,\sigma^2)$ și vrem să calculăm:

$$P(a\le X\le b)=P(a-\mu\le X-\mu\le b-\mu)$$ (translatat în origine)

$$=P\left(\frac{a-\mu}{\sigma}\le\frac{X-\mu}{\sigma}\le\frac{b-\mu}{\sigma}\right)$$ (scalat ca var = 1)

V.a. $\frac{X-\mu}{\sigma}\sim N(0,1)$ (versiunea standardizată a lui X)

$$P(a\le X\le b)=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)$$



Funcția de repartiție Φ a $N(0,1)$ verifică:

$$\Phi(z)=1-\Phi(-z),\ \forall z\in\mathbb{R}$$

$$\Phi(-z) = \int_{-\infty}^{-z} \varphi(t)\,dt \underset{u=-t}{\overset{\text{sch. var.}}{=}} \int_{+\infty}^{z} \varphi(-u)\,(-du) =$$

$$= \int_{z}^{+\infty} \varphi(-u)\,du \underset{\substack{\varphi \text{ simetric} \\ \varphi(u)=\varphi(-u)}}{=} \int_{z}^{\infty} \varphi(u)\,du = 1 - \int_{-\infty}^{z} \varphi(u)\,du$$

$$= 1 - \Phi(z)$$

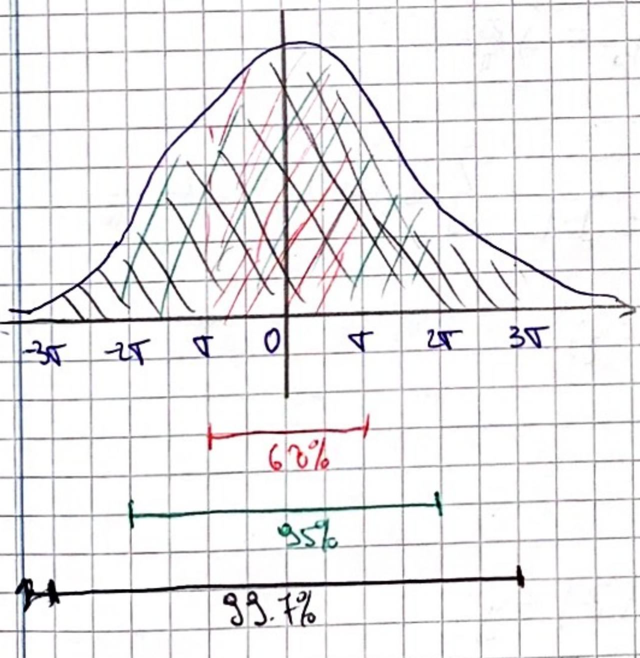
Mai mult, dacă $Z \sim N(0,1)$ atunci $-Z \sim N(0,1)$.

(T) Dacă $X \sim N(\mu, \sigma^2)$ atunci (Regula 68-95-99.7%)

$$P(|X-\mu| \leq \sigma) \simeq 0.68$$

$$P(|X-\mu| \leq 2\sigma) \simeq 0.95$$

$$P(|X-\mu| \leq 3\sigma) \simeq 0.997$$



Exp: $X \sim N(-1, 4)$, $\mu = -1$, $\sigma = 2$

$P(|X| < 3) = ?$

Etapa de standardizare:

$$P(|X| < 3) = P(-3 < X < 3)$$

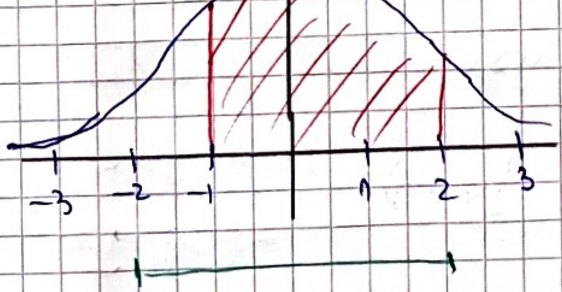
$$= P\left(\frac{-3-(-1)}{2} < \frac{X-(-1)}{2} < \frac{3-(-1)}{2}\right)$$

$$= P\left(-1 < \frac{X+1}{2} < 2\right)$$

$$= \Phi(2) - \Phi(-1)$$

~~$= P\left(-1 < \frac{X+1}{2} < 2\right)$~~

V.a. $\frac{X+1}{2} \sim N(0,1)$



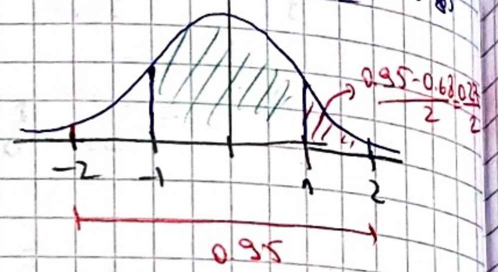
Știm că:

$$P(-1 < Z < 1) \simeq 0.68$$

$$P(-2 < Z < 2) \simeq 0.95$$

$P(-1 < Z < 2) = P(-1 < Z < 1) + P(1 \le Z < 2) \simeq 0.68 + \frac{0.27}{2} \simeq 0.85$

$\frac{X+1}{2}$



Exp: $Y = |Z|$, $Z \sim N(0,1)$

a) $\mathbb{E}\{Y\}$ și $Var(Y)$

b) $F_Y(y)$ și $f_Y(y)$

Sol: a) $\mathbb{E}\{Y\} = \mathbb{E}\{|Z|\} = \int_{-\infty}^{+\infty} |z|\,\varphi(z)\,dz = \int_{-\infty}^{+\infty} |z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz$

$= 2\int_0^{\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\,dz = \sqrt{\frac{2}{\pi}} \int_0^{\infty} z e^{-z^2/2}\,dz$

$= \sqrt{\frac{2}{\pi}} \left(-e^{-z^2/2}\right) \Big|_0^{\infty} = \sqrt{\frac{2}{\pi}}$

$Var(Y) = \mathbb{E}\{Y^2\} - \mathbb{E}\{Y\}^2$

$Y^2 = Z^2 \Rightarrow \mathbb{E}\{Y^2\} = \mathbb{E}\{Z^2\} = 1$ $\Big\} \Rightarrow Var(Y) = 1 - \frac{2}{\pi}$

b) $Y = |Z|$, $F_Y(y) = P(Y \le y)$

Dacă $y \le 0 \Rightarrow P(Y \le y) = 0$

Fie $y > 0$

$F_Y(y) = P(Y \le y) = P(|Z| \le y)$

$= P(-y \le Z \le y) = \Phi(y) - \Phi(-y)$

$= \Phi(y) - (1 - \Phi(y)) = 2\Phi(y) - 1$

$$F_Y(y) = \begin{cases} 0, & y \le 0 \\ 2\Phi(y) - 1, & y > 0 \end{cases}$$

$$f_Y(y) = \frac{d}{dy} F_Y(y) \Rightarrow f_Y(y) = \begin{cases} 0, & y \le 0 \\ 2\varphi(y), & y > 0 \end{cases}$$

Repartiții comune, marginale și condiționate

Considerăm două variabile aleatoare X și Y. Dacă privim individual putem calcula probabilități de tipul:

$$P(X \in A),\ P(Y \in B)$$

În ansamblu, considerăm vectorul (X, Y) și vrem

$$P((X,Y) \in \underset{\mathbb{R}^2}{\underset{\cap}{C}}) \rightarrow \text{repartiție comună}$$

- repartiție marginală $(P(X \in A)$ sau $P(Y \in B))$
  (ne interesăm doar la repartiția unei v.a. ignorând cealaltă v.a.

- repartiția condiționată - rep cond a lui X știind că $Y \in B$

Cazul v.a. discrete

Fie $(\Omega, \mathcal{F}, P)$ un câmp de probabilitate și X și Y două v.a. discrete $X, Y : \Omega \rightarrow \mathbb{R}$

$X(\Omega) = \{x_1, x_2, \dots x_m\}$

$Y(\Omega) = \{y_1, y_2, \dots y_n\}$

Cuplul $(X, Y) : \Omega \rightarrow \mathbb{R}^2$, $(X,Y)(\Omega) = \{(x_i, y_j) \mid \substack{1 \le i \le m \\ 1 \le j \le n}\}$

Numim funcție de masă a vectorului (X, Y):

$$\boxed{p_{X,Y}(x,y) = P(X=x, Y=y)}$$

Uneori se poate nota cu $f_{X,Y}(x,y) = P(X=x, Y=y)$

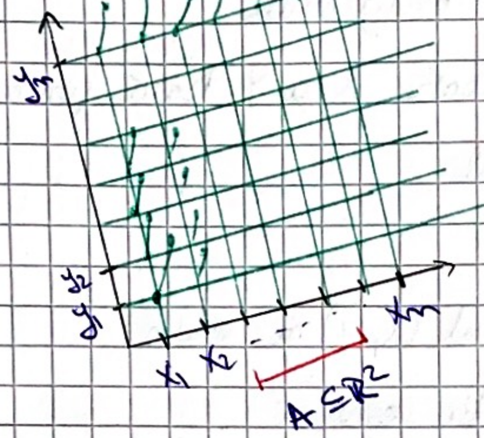
Funcția de repartiție a vectorului (X,Y):

$$F_{X,Y}(x,y) = P(X \le x, Y \le y)$$

(P) Funcția de masă $p_{X,Y}(x,y)$ (sau $f_{X,Y}(x,y)$) verifică:

a) $p_{X,Y}(x,y) \ge 0,\ \forall x, y$

b) $\sum_x \sum_y p_{X,Y}(x,y) = 1$ (masa totală $= 1$)



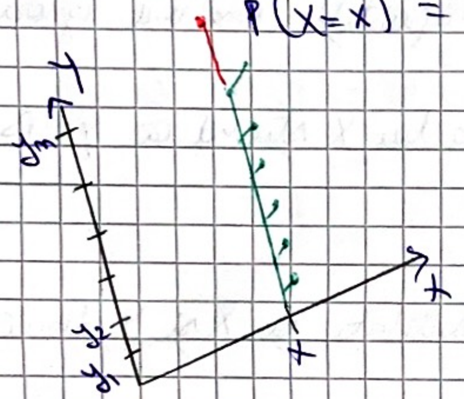
$X \in \{x_1, \dots x_m\}$

$Y \in \{y_1, \dots y_m\}$

$$P((X,Y) \in A) = \sum_{(x,y) \in A} P(X=x, Y=y)$$

• **Def:** Se numește repartiție marginală a lui X:

$$P(X=x) = \sum_y P(X=x, Y=y)$$



De ce?

$$P(X=x) = P(X=x \cap \Omega)$$
$$= P(X=x, Y \in \mathbb{R})$$
$$= P(X=x, \bigcup_y \{Y=y\})$$
$$= P(\bigcup_y \{X=x, Y=y\})$$
$$= \sum_y P(X=x, Y=y)$$

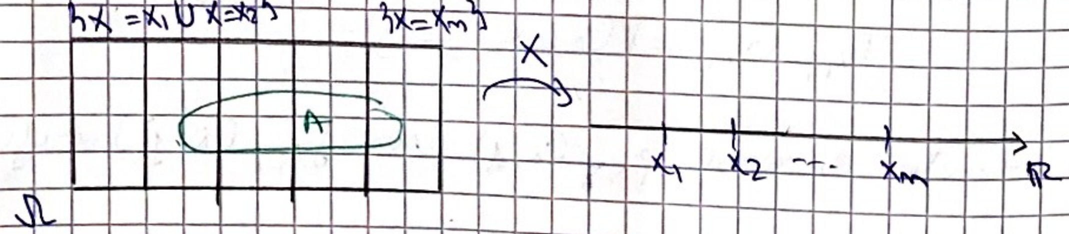
În mod similar se definește resp. marginală a lui Y:

$$P(Y=y) = \sum_x P(X=x, Y=y)$$

Fie A un eveniment, $A \in \mathcal{F}$ cu $P(A) > 0$. Atunci funcția de masă a lui X condiționat la A:

$$P_{X|A}(x) = P(X=x \mid A) = \frac{P(\{X=x\} \cap A)}{P(A)}$$

$(f_{X|A}(x))$



Cum $\{X=x\} \cap A$ sunt disjuncte două câte două avem că

$$P(A) = \sum_x P(\{X=x\} \cap A) \quad | : P(A) > 0$$

$$1 = \sum_x \frac{P(\{X=x\} \cap A)}{P(A)} = \sum_x P_{X|A}(x)$$

Exp: X este rezultatul aruncării cu zarul

A – rezultatul este par

$$p_{X|A}(x) = P(X=x \mid par), \quad x \in \{1,2,3,4,5,6\}$$

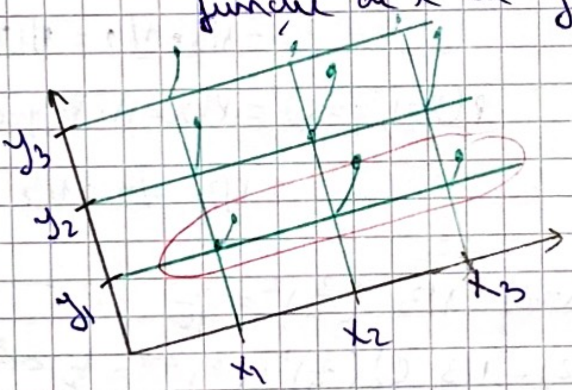
$$= \frac{P(X=x,\ par)}{P(par)} = \begin{cases} \frac{1/6}{1/2} = 1/3, & x=2,4,6 \\ 0, & x=1,3,5 \end{cases}$$

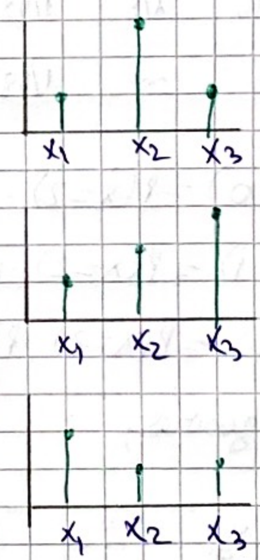
În particular, dacă $A = \{Y=y\}$ unde Y este o v.a. discretă cu $Y(\Omega) = \{y_1, \dots, y_n\}$ atunci putem vorbi de funcția de masă condiționată a lui X la $Y=y$:

$$p_{X|Y}(x|y) = P(X=x \mid Y=y).$$

Avem

$$\underbrace{p_{X|Y}(x|y)}_{\text{funcție de } x \text{ cu } y \text{ fixat}} = \frac{P(X=x, Y=y)}{P(Y=y)} = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$





$p_{X|Y}(x|y_1)$

$p_{X|Y}(x|y_2)$

$p_{X|Y}(x|y_3)$

Și păstrează forma dar este necesar să normalizăm (suma=1)

Obs: $p_{X,Y}(x,y) = p_X(x)\, p_{Y|X}(y|x)$

$$= p_Y(y)\, p_{X|Y}(x|y)$$

Similar cu situația de la evenim. condiționate:

$$P(A \cap B) = P(A) \cdot P(B|A) = P(B) \cdot P(A|B)$$

$A = \{X=x\}$

$B = \{Y=y\}$

Exp: Presupunem că avem un profesor care răspunde incorect în 25% din cazuri, independent de întrebare.

Presupunem, de asemenea, că pe parcursul unui curs putem avea 0, 1 sau 2 întrebări cu prob 1/3.

Fie X v.a care dă nr. de întrebări

Y v.a. care dă nr de răspunsuri greșite

$(X,Y): \Omega \to \mathbb{R}^2$ $\quad X(\Omega) = \{0,1,2\}$

$Y(\Omega) = \{0,1,2\}$

$P(X=x, Y=y) = ?$

$(x,y) \in \{(i,j) \mid i \in \{0,1,2\},\ j \in \{0,1,2\}\}$

X\Y	0	1	2	Σ
0	1/3	0	0	1/3
1	1/4	1/12	0	1/3
2	3/16	1/8	1/48	1/3
Σ	~	~	1/48	

$P(X=0, Y=0) = 1/3$

$P(X=1, Y=1) = P(X=1) \cdot P(Y=1 \mid X=1)$

$= 1/3 \cdot 1/4 = 1/12$

$P(X=1, Y=0) = P(X=1) \cdot P(Y=0 \mid X=1)$

$= 1/3 \cdot 3/4 = 1/4$

$P(X=2, Y=0) = P(X=2) \cdot P(Y=0 \mid X=2) = 1/3 \cdot \left(\frac{3}{4}\right)^2 = \frac{3}{16}$

$P(X=2, Y=1) = P(X=2) \cdot P(Y=1 \mid X=2) = 1/3 \cdot C_2^1 \left(\frac{1}{4}\right)^1 \cdot \left(\frac{3}{4}\right)^{2-1} = \frac{1}{\cancel{3}} \cdot 2 \cdot \frac{\cancel{3}}{16_8} = \frac{1}{8}$

$P(X=2, Y=2) = P(X=2) \cdot P(Y=2 \mid X=2) = 1/3 \cdot \left(\frac{1}{4}\right)^2 = \frac{1}{3} \cdot \frac{1}{16} = \frac{1}{48}$

În general,

X\Y	y_1	...	y_j	...	y_m	Σ
x_1						
⋮						
x_i			$P_{X,Y}(x_i, y_j)$			$P(X=x_i)$
⋮						
x_n						
Σ			$P(Y=y_j)$			

$$P(X=x_i) = \sum_{j=1}^{m} P(X=x_i, Y=y_j) = \sum_{j=1}^{m} P_{X,Y}(x_i, y_j)$$

$$P(Y=y_j) = \sum_{i=1}^{m} P(X=x_i, Y=y_j)$$

$$P(X=x_i \mid Y=y_j) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)}$$