

Decizie și Estimare în Prelucrarea Informației

Capitolul I. Semnale aleatoare

Variabile aleatoare

- ▶ **Variabilă aleatoare** = o variabilă care denumește o valoare produsă printr-un fenomen aleator
 - ▶ Practic, reprezintă *un nume* atașat unei valori arbitrare
 - ▶ Prescurtat: v.a.
- ▶ Notăție uzuală: X , Y etc..
- ▶ Exemple:
 - ▶ Numărul obținut prin aruncarea unui zar
 - ▶ Voltajul măsurat într-un punct dintr=un circuit
- ▶ Opusul = o valoare constantă (de ex. $\pi = 3.1415\dots$)

- ▶ **Realizare** a unei v.a. = o valoare particulară rezultată în urma fenomenului aleator
- ▶ **Spațiul realizărilor** Ω = mulțimea valorilor posibile ale unei v.a.
 - ▶ = mulțimea tuturor realizărilor
- ▶ Exemplu: aruncarea unui zar
 - ▶ V.a. se notează X
 - ▶ Se poate obține o realizare $X = 3$
 - ▶ Dar s-ar fi putut obține orice valoare din spațiul realizărilor

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

V.a. discrete și continue

- ▶ V.a. **discretă**: dacă Ω este o mulțime discretă
 - ▶ Exemplu: Numărul obținut prin aruncarea unui zar
- ▶ V.a. **continuă**: dacă Ω este o mulțime compactă
 - ▶ Exemplu: Valoarea tensiunii măsurate într-un punct

- ▶ Fie o v.a. continuă X
- ▶ **Funcția de repartiție (FR)**: probabilitatea ca X să aibă valoarea mai mică sau egală cu x

$$F(x) = P\{X \leq x\}$$

- ▶ Derivata funcției de repartiție este **funcția densitate de probabilitate (FDP)**

$$w(x) = \frac{dF(x)}{dx}$$

$$F(x) = \int_{-\infty}^x w(t)dt$$

- FDP este probabilitatea ca valoarea lui X să fie într-o vecinătate ϵ mică în jurul lui x , raportat la ϵ

$$\begin{aligned}w(x) &= \frac{dF(x)}{dx} = \lim_{\epsilon \rightarrow 0} \frac{F(x + \epsilon) - F(x - \epsilon)}{2\epsilon} \\&= \lim_{\epsilon \rightarrow 0} \frac{P(X \in [x - \epsilon, x + \epsilon])}{2\epsilon}\end{aligned}$$

Probabilitatea unei valori anume

- ▶ Probabilitatea ca v.a. continuă X să fie **exact** egală cu un x este **zero**
- ▶ O v.a. continuă are o infinitate de realizări posibile
- ▶ Probabilitatea unei valori anume este practic 0
- ▶ FDP este probabilitatea de a fi **într-o vecinătate mică în jurul** unei valori x

- ▶ Fie o v.a. discretă X
- ▶ **Funcția de repartiție (FR)**: probabilitatea ca X să aibă valoarea mai mică sau egală cu x

$$F(x) = P\{X \leq x\}$$

- ▶ Exemplu: FR pentru un zar
- ▶ Pentru v.a. discrete, FR este de tip “treaptă”

- ▶ Nu putem defini densitatea de probabilitate
 - ▶ pentru că derivata în punctele de discontinuitate nu e definită
- ▶ **Funcția masă de probabilitate** (FMP) (*probability mass function*): probabilitatea ca X să aibă valoarea egală cu x

$$w(x) = P\{X = x\}$$

$$F(x) = \sum_{\text{toți } t \leq x} w(t)$$

- ▶ Exemplu: FMP pentru un zar?

- Calculul probabilității pe baza FDP (v.a. continuă):

$$P\{A \leq X \leq B\} = \int_A^B w(x) dx$$

- Calculul probabilității pe baza FMP (v.a. discretă):

$$P\{A \leq X \leq B\} = \sum_{x=A}^B w_X(x)$$

- ▶ Probabilitatea ca X să fie între A și B este **suprafața de sub FDP**
 - ▶ adică integrala de la A la B
- ▶ Probabilitatea ca X să fie exact egal cu o valoare este zero
 - ▶ aria de sub un punct este nulă

Proprietățile FR/FDP/FMP

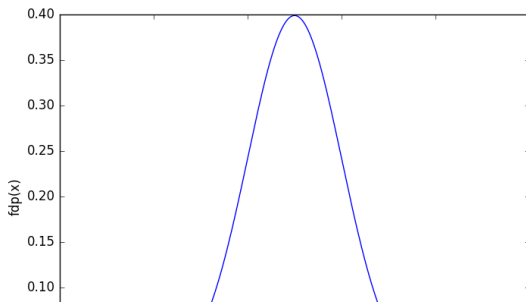
- ▶ FR este o funcție crescătoare
- ▶ FR / FDP / FMP sunt întotdeauna ≥ 0
- ▶ $FR(-\infty) = 0$ și $FR(\infty) = 1$
- ▶ Integrala FDP / suma FMP = 1

Distribuția normală

- Densitatea de probabilitate:

$$w(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

```
<type 'exceptions.UnicodeDecodeError'>  
'ascii' codec can't decode byte 0xc8 in position 8: ordinal n  
range(128)
```



Distribuția normală

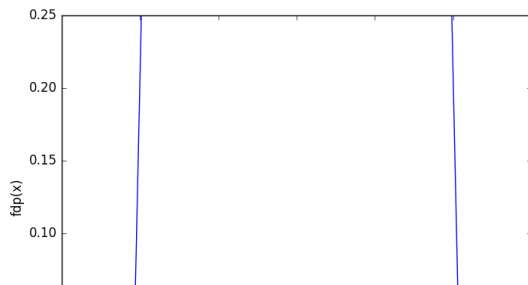
- ▶ Are doi parametri:
 - ▶ **Media** μ = “centrul” funcției
 - ▶ **Deviația standard** σ = cât de “lată” este funcția
- ▶ Constanta de la începutul expresiei asigură normalizarea (faptul că integrala = 1)
- ▶ Extrem de des întâlnită în practică
- ▶ Orice valoare reală este posibilă ($w(x) > 0, \forall x \in \mathbb{R}$)
- ▶ Se notează cu $\mathcal{N}(\mu, \sigma)$

Distribuția uniformă

- Densitatea de probabilitate = constantă între două limite

$$w(x) = \begin{cases} \frac{1}{b-a}, & x \in [a, b] \\ 0, & \text{elsewhere} \end{cases}$$

```
<type 'exceptions.UnicodeDecodeError'>  
'ascii' codec can't decode byte 0xc8 in position 8: ordinal not in  
range(128)
```



Distribuția uniformă

- ▶ Are doi parametri: limitele a și b ale intervalului
- ▶ “Înălțimea” funcției este $\frac{1}{b-a}$ pentru normalizare
- ▶ Foarte simplă
- ▶ Sunt posibile doar valori din intervalul $[a, b]$
- ▶ Se notează cu $\mathcal{U} [a, b]$

V.a. ca funcții de alte v.a

- ▶ O funcție aplicată unei v.a. produce o altă v.a.
- ▶ Exemple: dacă X este o v.a. distribuită $\mathcal{U} [0, 10]$, atunci
 - ▶ $Y = 5 + X$ este o altă v.a., distribuită $\mathcal{U} [5, 15]$
 - ▶ $Z = X^2$ este de asemenea o v.a.
 - ▶ $T = \cos(X)$ este de asemenea o v.a.
- ▶ Motivație: dacă X este aleatoare, și valorile Y , Z , T sunt aleatoare
- ▶ X , Y , Z , T nu sunt independente
 - ▶ O anumită valoare a uneia implică automat și valoarea celorlalte

Sisteme de mai multe variabile aleatoare

- ▶ Fie un sistem cu două v.a. continue X și Y
- ▶ Există funcția de repartiție comună:

$$F(x, y) = P\{X \leq x \cap Y \leq y\}$$

- ▶ Densitatea de probabilitate comună:

$$w(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

- ▶ FDP comună descrie probabilitatea ca X și Y să se găsească într-o vecinătate a lui x și y , **simultan**
- ▶ Similar pentru v.a. discrete: funcția masă de probabilitate comună

$$w(x, y) = P\{X = x \cap Y = y\}$$

Variabile independente

- ▶ Două v.a. X și Y sunt **independente** dacă valoarea uneia nu influențează în nici un fel valoarea celeilalte
- ▶ Pentru v.a. independente, probabilitatea ca $X = x$ și $Y = y$ este produsul celor două probabilități

- ▶ V.a. discrete:

$$w(x, y) = w(x) \cdot w(y)$$

$$P\{X = x \cap Y = y\} = P\{X = x\} \cdot P\{Y = y\}$$

- ▶ Valabilă pentru FR / FDP / FMP, v.a. continue sau aleatoare
- ▶ Idem pentru mai mult de două v.a.

- ▶ V.a. sunt caracterizate prin medii statistice (“*momente*”)
- ▶ Valoarea medie (momentul de ordin 1)
- ▶ Pentru v.a. continue:

$$\bar{X} = E\{X\} = \int_{-\infty}^{\infty} x \cdot w(x) dx$$

- ▶ Pentru v.a. discrete:

$$\bar{X} = E\{X\} = \sum_{x=-\infty}^{\infty} x \cdot w(x)$$

- ▶ (Exemplu: entropia $H(X)$ = valoarea medie a informației)
- ▶ Notăție uzuală: μ

Proprietățile valorii medii

- ▶ Calculul valorii medii este o operație **liniară**
 - ▶ pentru că, la bază, integrala / suma este o operație liniară

- ▶ Liniaritate

$$E\{aX + bY\} = aE\{X\} + bE\{Y\}$$

- ▶ Sau:

$$E\{aX\} = aE\{X\}, \forall a \in \mathbb{R}$$

$$E\{X + Y\} = E\{X\} + E\{Y\}$$

- ▶ Fără demonstrație

Valoarea pătratică medie

- ▶ Valoarea pătratică medie = valoarea medie a pătratelor valorilor
- ▶ Momentul de ordin 2
- ▶ Pentru v.a. continue:

$$\overline{X^2} = E\{X^2\} = \int_{-\infty}^{\infty} x^2 \cdot w(x) dx$$

- ▶ Pentru v.a. discrete:

$$\overline{X^2} = E\{X^2\} = \sum_{-\infty}^{\infty} x^2 \cdot w(x)$$

- ▶ Interpretare: media pătratelor = energia medie a unui semnal

Dispersia (varianța)

- ▶ Dispersia (varianța) = valoarea pătratică medie a abaterii față de valoarea medie :)
- ▶ V.a. continue:

$$\sigma^2 = \overline{\{X - \mu\}^2} = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot w(x) dx$$

- ▶ V.a. discrete:

$$\sigma^2 = \overline{\{X - \mu\}^2} = \sum_{-\infty}^{\infty} (x - \mu)^2 \cdot w(x)$$

- ▶ Interpretare: cât de mult variază valorile în jurul mediei
 - ▶ σ^2 = mare: abateri mari față de medie
 - ▶ σ^2 = mic: valori concentrate în jurul mediei

Legătura între cele trei mărimi

- Legătura între medie, valoarea pătratică medie și dispersie:

$$\begin{aligned}\sigma^2 &= \overline{\{X - \mu\}^2} \\ &= \overline{X^2 - 2 \cdot X \cdot \mu + \mu^2} \\ &= \overline{X^2} - 2\mu\overline{X} + \mu^2 \\ &= \overline{X^2} - \mu^2\end{aligned}$$

Suma variabilelor aleatoare

- ▶ Suma a două sau mai multe v.a. **independente** este tot o v.a.
- ▶ Distribuția ei = **convoluția** distribuțiilor v.a. componente
- ▶ Dacă $Z = X + Y$

$$w(z) = w(x) \star w(y)$$

- ▶ Caz particular: dacă X și Y sunt v.a. normale, cu $\mathcal{N}(\mu_X, \sigma_X^2)$ și $\mathcal{N}(\mu_Y, \sigma_Y^2)$, atunci:
 - ▶ Z este tot o v.a. cu distribuție normală, $\mathcal{N}(\mu_Z, \sigma_Z^2)$, având:
 - ▶ media = suma mediilor: $\mu_Z = \mu_X + \mu_Y$
 - ▶ dispersia = suma dispersiilor: $\sigma_Z^2 = \sigma_X^2 + \sigma_Y^2$

Procese aleatoare

- ▶ Un **proces aleator** = o secvență de variabile aleatoare indexate (înșiruite) în timp
- ▶ Proces aleator **în timp discret** $f[n]$ = o secvență de v.a. la momente de timp discrete
 - ▶ ex: o secvență de 50 aruncări de zar, cotația zilnică a unor acțiuni la bursă
- ▶ Proces aleator **în timp continuu** $f(t)$ = o secvență de v.a. la orice moment de timp
 - ▶ ex: un semnal tip zgomot de tensiune
- ▶ Fiecare eșantion dintr-un proces aleator este o v.a. de sine stătătoare
 - ▶ ex.: $f(t_0)$ = valoarea la momentul t_0 este o v.a.

Realizări ale proceselor aleatoare

- ▶ **Realizare** a unui p.a. = o secvență particulară de realizări ale v.a. componente
 - ▶ ex: un anumit semnal de zgomot măsurat cu un osciloscop; dar am fi putut măsura orice altă realizare
- ▶ Când considerăm un p.a., considerăm întregul set de realizări posibile
 - ▶ la fel ca atunci când considerăm o v.a.

Distribuții de ordin 1 ale proceselor aleatoare

- ▶ Fiecare eșantion $f(t_1)$ ($f[n_1]$) dintr-un p.a. este o v.a.
 - ▶ cu FR $F_1(x; t_1)$
 - ▶ cu FDP / FMP $w_1(x; t_1) = \frac{dF_1(x; t_1)}{dx}$
- ▶ Eșantionul de la momentul t_2 este o v.a. diferită, cu funcții posibil diferite
 - ▶ cu FR $F_1(x; t_2)$
 - ▶ cu FDP / FMP $w_2(x; t_2) = \frac{dF_1(x; t_2)}{dx}$
- ▶ Indicele w_1 arată că considerăm o singură v.a. din proces (distribuții *de ordin 1*)
- ▶ Similar pentru p.a. discrete

Distribuții de ordin 2

- ▶ O pereche de v.a. $f(t_1)$ and $f(t_2)$ dintr-un proces aleator $f(t)$ au:
 - ▶ FR comună $F_2(x_i, x_j; t_1, t_2)$
 - ▶ FDP / FMP comună $w_2(x_i, x_j; t_1, t_2) = \frac{\partial^2 F_2(x_i, x_j; t_1, t_2)}{\partial x_i \partial x_j}$
- ▶ Aceste funcții descriu cum sunt distribuite valorile perechilor (distribuții *de ordin 2*)
- ▶ Similar pentru p.a. discrete

Distribuții de ordin n

- ▶ Generalizare la n eșantioane ale unui p.a.
- ▶ Un set de n v.a. $f(t_1), \dots, f(t_n)$ dintr-un proces aleator $f(t)$ au:
 - ▶ FR comună $F_n(x_1, \dots, x_n; t_1, \dots, t_n)$
 - ▶ FDP / FMP comună $w_n(x_1, \dots, x_n; t_1, \dots, t_n) = \frac{\partial^2 F_n(x_1, \dots, x_n; t_1, \dots, t_n)}{\partial x_1 \dots \partial x_n}$
- ▶ Aceste funcții descriu cum sunt distribuite valorile seturilor de n valori (distribuții *de ordin n*)
- ▶ Similar pentru p.a. discrete

Hic sunt leones

Multiple random variables

- ▶ Consider a system with two random variables X and Y
- ▶ Joint cumulative distribution function:

$$F_{XY}(x_i, y_j) = P\{X \leq x_i \cap Y \leq y_j\}$$

- ▶ Joint probability density function:

$$w_{XY}(x_i, y_j) = \frac{\partial^2 P_{XY}(x_i, y_j)}{\partial x \partial y}$$

- ▶ The joint PDF gives the probability that the values of the two r.v. X and Y are in a **vicinity** of x_i and y_j simultaneously
- ▶ Similar definitions extend to the case of discrete random variables

Random process

- ▶ A **random process** = a sequence of random variables indexed in time
- ▶ **Discrete-time** random process $f[n]$ = a sequence of random variables