Statistică fizică

1. Mărimi măsurate direct : $y = f(x)$

U(mV)	nr. impulsuri	$\overline{U}$ (mV)	D_{U_i} (mV)	$\sigma_{\overline{U}}$ (mV)	interval de încredere
5.10	10		5.1 − 5.6 = −0.5		
5.20	90	5.6.	5.2 − 5.6 = −0.4	0.20	5.6 ± 0.2
5.30	250		5.3 − 5.6 = −0.3		[5.4; 5.8]
6.20	450		6.2 − 5.6 = 0.6		

x (input), y (output), media (avg.), deviația (față de medie), abaterea standard medie, intervalul de încredere

$\underbrace{\text{x, y}}_{\text{datele problemei / măsurători}}$ $\underbrace{\text{media, deviația, abaterea, intervalul}}_{\text{interpretarea statistică a datelor}}$

Media : $\overline{x} = \dfrac{\sum_{i=1}^{n} x_i}{n}$ u.m.

Deviația (față de medie) : $d_i = x_i - \overline{x}$ u.m., $(\forall)\ i = \overline{1,n}$

Abaterea standard medie : $\sigma_{\overline{x}} = \sqrt{\dfrac{\sum_{i=1}^{n} d_i^2}{n(n-1)}} = \sqrt{\dfrac{\sum_{i=1}^{n}(x_i - \overline{x})^2}{n(n-1)}}$

Intervalul de încredere : $x = \overline{x} \pm \sigma_{\overline{x}}$ u.m.

n = numărul de puncte = cardinalul mulțimii de intrare

u.m. = unitatea de măsură

Media, deviația, abaterea standard și intervalul de încredere se calculează de obicei pe baza valorilor de intrare. Aceste măsuri se referă la modul în care datele de intrare sunt distribuite și variază, înainte de a fi supuse unei anumite operații sau analize.

interval de încredere = interval de confidență

$= \text{media} \pm (\text{Z score} \cdot \text{abatere standard})$

Z score → nivel de încredere

2. Mărimi măsurabile indirect (prin calcul)

$z = f(x, y)$

x → măsurată direct

y → măsurată direct

z → măsurată indirect

$\overline{x} = ?$ (media valorilor lui x)

$\overline{y} = ?$ (media valorilor lui y)

$$\sigma_{\overline{z}} = \sqrt{\left(\frac{\partial f}{\partial x}\Big/_{\substack{x=\overline{x}\\ y=\overline{y}}}\right)^2 \sigma_{\overline{x}}^2 + \left(\frac{\partial f}{\partial y}\Big/_{\substack{x=\overline{x}\\ y=\overline{y}}}\right)^2 \sigma_{\overline{y}}^2}$$

se înlocuiește în funcția $\frac{\partial f}{\partial x}$
- x cu $\overline{x}$
- y cu $\overline{y}$

se înlocuiește în funcția $\frac{\partial f}{\partial y}$
- x cu $\overline{x}$
- y cu $\overline{y}$

$$\frac{\partial f}{\partial x} = \frac{\partial}{\partial x} f(x,y) \overset{\text{not.}}{=\!=} f_1(x,y)$$

$$\frac{\partial f}{\partial y} = \frac{\partial}{\partial y} f(x,y) \overset{\text{not.}}{=\!=} f_2(x,y)$$

$$\sigma_{\overline{z}} = \sqrt{f_1^2(\overline{x}, \overline{y}) \cdot \sigma_{\overline{x}}^2 + f_2^2(\overline{x}, \overline{y}) \cdot \sigma_{\overline{y}}^2}$$

$z = \overline{z} \pm \sigma_{\overline{z}}$ (u. m.) → interval de încredere

3. Dependența ⇒ grafic (G_f): regresia (o dreaptă care să treacă printre puncte)

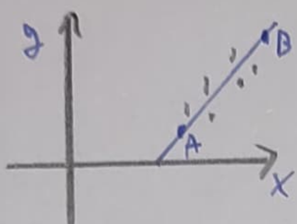
pentru grafic, unitățile vor fi: $1\cdot10^K$, $2\cdot10^K$, $5\cdot10^K$; $K\in\mathbb{Z}$

Din 1 în 1 / 10 în 10 / 100 în 100 / ... la fiecare centimetru

Din 2 în 2 / 20 în 20 / 200 în 200 / ... la fiecare centimetru

Din 5 în 5 / 50 în 50 / 500 în 500 / ... la fiecare centimetru

Regresie liniară (dreaptă)



$$m_d = \frac{y_B - y_A}{x_B - x_A} = \frac{\Delta y}{\Delta x}$$

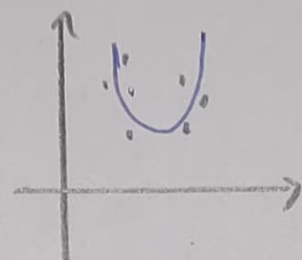
panta dreptei

$A, B \in d$ regresie

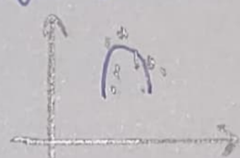
$d_{regresie}: y = m_d \cdot x + m$

Regresie polinomială de ordin II

→ polinom convex:

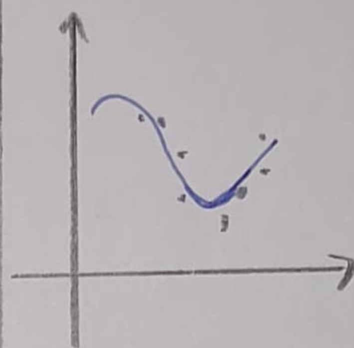


→ polinom concav:



$d_{regresie}: y = a\cdot x^2 + b\cdot x + c$

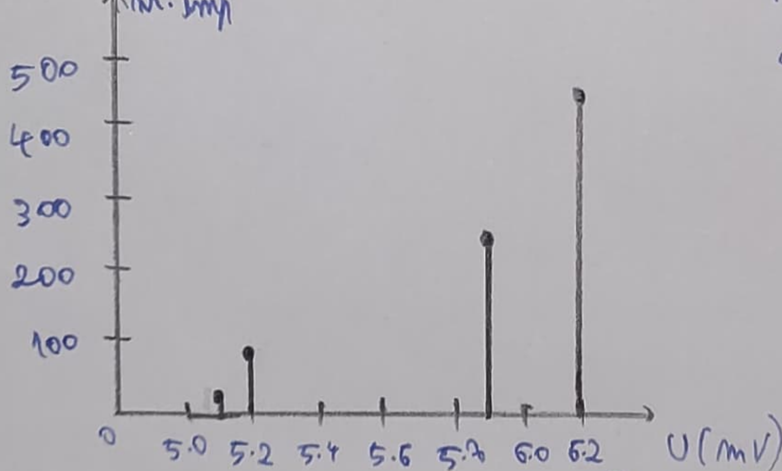
Regresie polinomială de ordin III



$d_{regresie}: y = a\cdot x^3 + b\cdot x^2 + c\cdot x + d$

Histograma (bare verticale) este o reprezentare grafică a distribuției frecvențelor într-un set de date.

Histograma numărului de impulsuri



Covariația: $Cov(x,y) = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{n}$

- $>0 \Rightarrow$ dacă o variabilă crește, este posibil ca cealaltă să crească
- $<0 \Rightarrow$ dacă o variabilă crește, este posibil ca cealaltă să scadă

Corelația: $Corr(x,y) = \dfrac{Cov(x,y)}{\sigma_x \cdot \sigma_y} = \dfrac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\dfrac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{N(N-1)}} \cdot \sqrt{\dfrac{\sum_{i=1}^{n} (y_i - \bar{y})}{n(n-1)}}}$

Corelația se mai notează: $Corr(x,y) = \rho(x,y) = r$

$r \rightarrow$ coeficientul de corelație

$r = Corr(x,y)$

- NU IMPLICĂ CAUZALITATE
- $\in [-1,1]$
- $>0 \Rightarrow$ variabilele sunt direct proporționale
  $\Rightarrow$ dreapta de ~~r~~ regresie are panta pozitivă
- $<0 \Rightarrow$ variabilele sunt indirect proporționale
  $\Rightarrow$ dreapta de regresie are panta negativă

regresie liniară $\Rightarrow$ estimare în sensul „celor mai mici pătrate" (C.M.M.P.)

$d_{regresie}: y = a \cdot x + b$

$a =$ "slope"

$b =$ "intercept"

puncte: $(x_1,y_1), (x_2,y_2), (x_3,y_3) \ldots$

$A = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix}; B = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

$X = \begin{pmatrix} b \\ a \end{pmatrix} = \begin{pmatrix} \text{"intercept"} \\ \text{"slope"} \end{pmatrix}$

$$\boxed{X = (A^t \cdot A)^{-1} \cdot A^t \cdot B}$$