A. Media şi abaterea pătratică medie

A1. Mărimea de interes în fizica atomică şi nucleară este numărul de pulsuri N care rezultă atunci când un detector este parcus de un fascicul de particule. Sistemul electronic numără aceste pulsuri într-un anumit timp (Δt) iar raportul dintre numărul de pulsuri şi intervalul de timp în care acestea au fost acumulate se numeşte viteză de numărare:

$$R = \frac{N}{\Delta t}$$

Atât N cât şi R sunt mărimi aleatoare, specifice fizicii cuantice, având o variabilitate intrinsecă proceselor atomice. Notând generic această variabilă aleatoare X_i, se poate calcula o valoare medie (aritmetică) a ei dacă sistemul care le generează este staționar. În acest caz valoarea medie va putea fi calculată cu relația:

$$X_{med} = \overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

unde n este numărul de măsurători care au condus la această medie.

Măsurătorile individuale au abateri de la această valoare. Abaterea va fi $\Delta X_i = X_i - X_{med}$ care are proprietatea ca suma tuturor abaterilor să fie nulă, dacă fluctuațiile de la valoarea medie sunt distribuite egal în jurul mediei (distribuție normală sau gaussiană). O măsură a împrăştierii acestor valori de la medie se poate calcula cu ajutorul relației care ne dă abaterea pătratică medie:

$$\sigma^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})$$

iar rădăcina pătrată a acestei valori se numeşte eroare standard (σ) a unei măsurători.

O măsurătoare se va prezenta sub forma:

$$X = X_i \pm \sigma$$

iar valoarea medie se va prezenta sub forma:

$$X = \overline{X} \pm \frac{1}{\sqrt{n}}\sigma$$

A2. Dacă pornim de la o distribuție cunoscută cum ar fi ce Poisson, atunci avem următoarele relații de bază pentru evaluarea erorilor.

Valoarea medie se calculează cu ajutorul relației:

$$m = \frac{\sum_{K=o}^{K_{max}} n_K \cdot K}{\sum_{K=o}^{K_{max}} n_K}$$

Frecvenţa ν_K de apariţie a K impulsuri se calculează pe baza distribuţiei teoretice considerate (Poisson în cazul proceselor atomice şi nucleară) cunoscând expresia probabilităţii de apariţie a K evenimente (impusuri):

$$N = \sum_{K=0}^{K_{max}} n_K \,, \qquad \nu_K = NP(K)$$

Abaterea pătratică medie se calculează utilizând relaţia:

$$\overline{D}^2 = \frac{1}{N}\sum_{i=1}^{N}\left(K_i - \overline{K}\right)^2 = \frac{1}{N}\sum_{K=0}^{K_{max}} n_K (K-m)^2$$

$$\sigma_K^{exp} = \sqrt{\overline{D}^2}$$

unde σ^{exp}_K reprezintă abaterea standard.

Pentru distribuţia Poisson, abaterea standard teoretică este dată de: $\sigma_K^t = \sqrt{m}$

B. Regresia liniară (Fitare liniară)

a. Pentru un set de *N* date experimentale de forma $y_i=\{x_i\}$ care se presupun a fi liniar corelate se poate căuta o funcţie de fitare de forma:

$$y(x)=a + bx$$

b. Pentru a obţine valorile cele mai bune pentru coeficienţii regresiei, se foloseşte funcţia χ^2 :

$$\chi^2 = \sum_i \left[\frac{1}{\sigma_i^2} (y_i - a - bx_i)^2 \right]$$

c. Procedura de fitare presupune minimizarea funcţiei χ^2 simultan în raport cu fiecare coeficient. Coeficienţii se află din rezolvarea sistemului rezultat şi au expresiile:

$$a = \frac{1}{\Delta}\left(\sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} - \sum \frac{x_i}{\sigma_i^2} \sum \frac{x_i\, y_i}{\sigma_i^2} \right)$$

$$b = \frac{1}{\Delta}\left(\sum \frac{1}{\sigma_i^2} \sum \frac{x_i\, y_i}{\sigma_i^2} - \sum \frac{x_i^2}{\sigma_i^2} \sum \frac{y_i}{\sigma_i^2} \right)$$

$$\Delta = \left(\sum \frac{1}{\sigma_i^2} \sum \frac{x_i^2}{\sigma_i^2} - \left(\sum \frac{x_i}{\sigma_i^2} \right)^2 \right)$$

d. Estimarea deviației standard se face conform relației:

$$\sigma^2 \cong s^2 = \frac{1}{N-2}\sum (y_i - a - bx_i)^2$$

e. Estimarea erorilor coeficienților de regresie se face cu formulele:

$$\sigma_a^2 \cong \frac{\sigma^2}{\Delta}\sum x_i^2$$

$$\sigma_b^2 \cong \frac{N}{\Delta}\sigma^2$$

f. Gradul de corelație liniară a acestor date experimentale se poate estima folosind expresia:

$$r = \frac{N\sum x_i y_i - \sum x_i \sum y_i}{\left[N\sum x_i^2 - \left(\sum x_i\right)^2\right]^{1/2}\left[N\sum y_i^2 - \left(\sum y_i\right)^2\right]^{1/2}}$$

g. Gruparea datelor în clase pentru a realiza o histogramă.Pentru majoritatea testelor de verificare a normalității repartiției, sau a descrierii repartiției, este necesar ca datele să fie grupate pe intervale de valori, pentru a le condensa ordonat. Aceste intervale sunt denumite clase.

Analiza statistică recomandă ca numărul claselor (k) să fie între 13 şi 20 care se determină cu relația Sturges:

$$k = 1 + 3{,}322 \log N,$$

unde *N* este numărul de date.

Când N este mai mic decât 250 este suficientă împărțirea în 10 clase. Pentru un număr redus de date (< 25) gruparea în clase are importanță redusă din punct de vedere al unui supliment de informație.

Amplitudinea intervalului unei clase este

$$\Delta x = (x_{max} - x_{min})/k.$$

Numărul n_i de date aparținând unei clase se numeşte frecvență.

Limitele claselor trebuie astfel alese încât să se identifice fără echivoc cărei clase îi aparține şirul de date.

Bibliografie:

P/R.Bevington, "*Data reduction and error analysis for the physical sciences*", McGraw-Hill, 1969;