

Prelucrarea datelor experimentale – problemă interdisciplinară de informatică pentru învățământul preuniversitar și universitar

Olga GROSU, Nicolae BALMUȘ, Ion BALMUȘ
Universitatea Pedagogică de Stat "Ion Creangă"
olga.grosu4@gmail.com; n_balmus@mail.md

Abstract — În lucrare se descrie un soft educațional, de concepție proprie, cu ajutorul căruia elevii și studenții însușesc mai profund bazele teoretice ale prelucrării datelor experimentale afectate de erori întâmplătoare.

Cuvinte cheie — calculare coeficienți Student; modelare date experimentale; experiment numeric;

INTRODUCERE

Orice experiment științific se bazează pe măsurători, în rezultatul cărora se obțin date experimentale afectate, în mod inevitabil, de erori: sistematice și aleatoare (statistice). Deoarece erorile sistematice pot fi reduse la minimum, se poate afirma că principala sursă de imprecizie a măsurătorilor o reprezintă erorile aleatoare [1].

Fie x_1, x_2, \dots, x_n - un set de date experimentale afectate de erori aleatoare. Problema principală a prelucrării acestor date experimentale constă în determinarea valorii medii \bar{x} și a intervalului de încredere Δx pentru nivelul de încredere α stabilit apriori.

$$x = \bar{x} \pm \Delta x, \quad \Delta x = t(\alpha, n)\sigma_{\bar{x}}, \quad (1)$$

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}, \quad \sigma_{\bar{x}} = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n(n-1)}}, \quad (2)$$

unde $\sigma_{\bar{x}}$ - eroarea standard a valorii medii, $t(\alpha, n)$ - coeficienții Student care, de regulă, se determină din tabele disponibile în bibliografia de specialitate [1,2] și Internet.

Scopul acestei lucrări constă în elaborarea unei aplicații cu ajutorul căreia utilizatorul va verifica în mod interactiv principalele ipoteze ale teoriei erorilor și va determina valoarea medie și intervalul de încredere introducând de la tastatură numai setul de date experimentale și nivelul de încredere. Coeficienții Student se vor calcula în interiorul aplicației.

I. CALCULAREA COEFICIENȚILOR STUDENT

Coeficienții Student $t(\alpha, n)$ se calculează, rezolvând următoarea ecuație neliniară [1]:

$$F(t) = \frac{\alpha}{C(n)} - \int_0^t \left(1 + \frac{x^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} dx = 0 \quad (3)$$

$$C(n) = \frac{2\Gamma\left(\frac{n+1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)\sqrt{\pi n}} = \begin{cases} \frac{1}{2\sqrt{n}} \prod_{i=1}^{\frac{n-1}{2}} \frac{i+0.5}{i}, & n - \text{par} \\ \frac{1}{\pi\sqrt{n}} \prod_{i=1}^{\frac{n-1}{2}} \frac{i}{i-0.5}, & n - \text{impar} \end{cases} \quad (4)$$

$\Gamma(n)$ - funcția Gama,

$$\Gamma(n) = (n-1)!, \quad \Gamma\left(n + \frac{1}{2}\right) = \frac{(2n-1)!!}{2^n} \sqrt{\pi} \quad (5)$$

Rezolvare numerică a ecuației neliniare (3) prin metoda coardelor sau înjumătățirii segmentului este neefectivă, fiindcă la fiecare pas al iterațiilor trebuie să calculăm integrala definită cu o precizie înaltă, minimum de două ori. Metoda Newton este foarte rapidă dar necesită calcularea funcției și derivatei ei la fiecare pas al iterațiilor.

$$t_c = t_p - F(t_p) / F'(t_p) \quad (6)$$

t_c, t_p - sunt soluțiile la pasul curent și precedent.

Deoarece necunoscuta t în ecuația (3) se află la limita de sus a integralei definite, problema calculării derivatei se rezolvă simplu:

$$F'(t) = \left(1 + \frac{t^2}{n}\right)^{-\left(\frac{n+1}{2}\right)} \quad (7)$$

Prin urmare implementarea metodei Newton (6) pentru rezolvarea numerică a ecuației (3) necesită calcularea integralei definite o singură dată la fiecare pas al iterațiilor.

În continuare prezentăm codul TPascal al funcției `cst(alfa,n)` care calculează numeric coeficienții Student.

```
function cst(alfa:extended;
            n:integer):extended;
var cstp,cstc,ri:extended;
function f(x:extended;n:integer):extended;
begin
    f:=1/power((1+sqr(x)/n),(n+1)/2);
end;
function c(n:integer):extended;
// se calculează funcția C(n) conform (4)
var i,nn:integer;p,cc:extended;
```

```

begin
  if odd(n)then begin
    p:=1;for i:=1 to (n-1)div 2 do
      p:=p*i/(i-0.5);
      c:=p/(pi*sqrt(n));
    end else
    begin
      p:=1;for i:=1 to (n div 2-1)
        dop:=p*(i+0.5)/i;
        c:=p/sqrt(n)/2;
      end;
    end;{end c }
  procedure simpson(a,b,eps:extended;var
s:extended);
// se calculează integrala definită,
metoda Simpson
  var nn,i:integer;
      h,s0,s2,s4,sh,sh2:extended;
  begin
    h:=(b-a)/2; nn:=1;s2:=0.0;
    s0:=f(a,n)+f(b,n);s4:=f(a+h,n);
    sh2:=h*(s0+4*s4)/3;
  repeat
    h:=h/2;sh:=sh2;nn:=nn+nn;
    s2:=s2+s4;
    s4:=0.0;
    for i:=1 to nn do
      s4:=s4+f(a+(2*i-1)*h,n);
      sh2:=h*(s0+2*s2+4*s4)/3;
    until (abs(sh2-sh)<eps);
    s:=sh2;
  end;{end simpson}
  begin
//se rezolvă ecuația neliniară (3), metoda
Newton
  cstc:=1; dec(n);
  repeat cstp:=cstc;
    simpson(0,cstp,0.000001,ri);
    cstc:=cstp-
      (ri-alfa/c(n)/2)/f(cstp,n);
  until (abs(cstc-cstp)<=0.000001);
  cst:=cstc;
end;{end cst}

```

I. APLICAȚIA STUDENT

În baza funcției *cst* a fost creată o aplicație "Student", care la momentul actual are trei pagini: I. Calculare coeficienți student; II. Prelucrare date experimentale măsurate direct; III. Generator de Subiecte.

În prima pagină, după introducerea nivelului de încredere (siguranță) și a numărului de date experimentale, utilizatorul apasă butonul OK pentru calcularea coeficientului student. În rezultat pe ecran se afișează valoarea coeficientului Student cu cel puțin 4 cifre semnificative corecte după virgulă. Tabelul cu coeficienții Student, prezent în fereastră, servește pentru verificarea corectitudinii calculului coeficienților Student.

Fig.1. Calcularea coeficienților Student.

Pagina a doua a aplicației "Prelucrare date experimentale măsurate direct" este preconizată pentru demonstrarea corectitudinii principalelor proprietăți ale erorilor întâmplătoare [1,2]: erorile mici, în valoare absolută, sunt mai frecvente decât erorile mari (**principiul cauzal**); toate erorile întâmplătoare sunt mai mici decât o anumită limită (**principiul limitativ**); dacă numărul de măsurători este suficient de mare se consideră că numărul erorilor negative este egal cu numărul erorilor pozitive (**principiul distributiv**); probabilitatea ca să avem o anumită eroare întâmplătoare, prin efectuarea unei măsurători, depinde de valoarea absolută a erorii (**principiul probabilistic**).

Pentru verificarea acestor principii este necesar să efectuăm un număr mare de măsurători (~ 1000-10000). Prin urmare, în condițiile laboratoarelor de instruire (preuniversitare și universitare) este imposibil de verificat riguros aceste principii. O soluție pentru rezolvarea acestei probleme este simularea asistată de calculator a rezultatelor măsurătorilor.

Karl Gauss, în anul 1821, a demonstrat că în rezultatul măsurătorilor se obțin mărimi pentru care densitatea de repartiție are forma:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, \quad (8)$$

unde \bar{x} este valoarea medie, iar σ – dispersia.

În literatura de specialitate [3] sunt descriși mai mulți algoritmi pentru generarea numerelor aleatoare cu densitatea de repartiție Gauss. În calitate de algoritm pentru generarea valorilor măsurătorilor experimentale poate servi relația

$$x = \bar{x} + \sigma \frac{\left(\sum_{i=1}^{12} r - 6 \right)}{12}, \quad (9)$$

Unde r este un număr aleatoriu uniform repartizat pe segmentul [0..1]. În majoritatea limbajelor de programare sunt predefinite funcții care generează astfel de numere. În Pascal și Delphi această funcție are numele **random**. În Delphi, în unitatea **math** este predefinită funcția

Function RandG (Mean, StdDev:Extended):Extended;, care generează numere aleatoare cu valoarea medie Mean și dispersia (abaterea standard) StdDev.

În baza acestor funcții a fost implementat codul paginii "Prelucrare date experimentale măsurate direct" Fig. 2

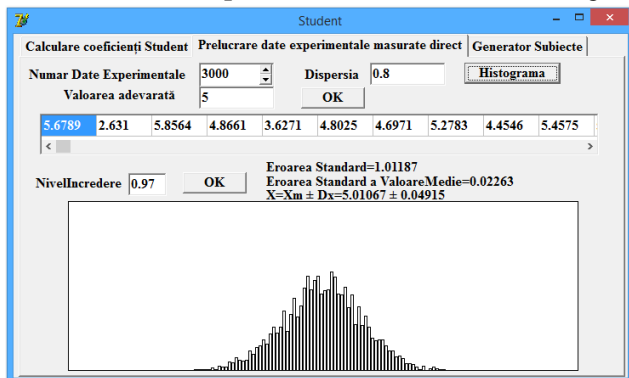


Fig.2. Prelucrarea datelor experimentale măsurate direct $\sigma=0.8$.

În pagina II (fig.2) utilizatorul setează numărul dorit de date experimentale, valoarea adevărată și dispersia. După apăsarea butonului **OK** în tabel se generează numărul respectiv de date experimentale. După apăsarea butonului **Histograma** pe ecran se afișează histograma de repartizare. La această etapă utilizatorul observă cum influențează dispersia datelor experimentale asupra densității lor de repartiție. În fig.3 este reprezentată histograma pentru dispersia 0.5. Comparând imaginile din figurile 2 și 3 utilizatorul observă că în cazul dispersiei 0.5 datele sunt mai bine concentrate în jurul valorii cu densitatea maximală de cât în cazul dispersiei 0.8, în conformitate cu proprietățile erorilor enumerate mai sus.

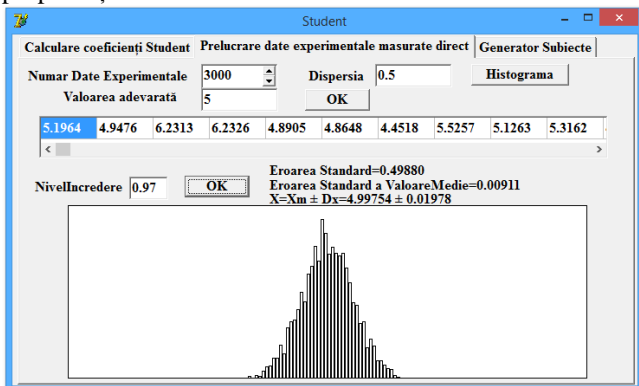


Fig.3. Prelucrarea datelor experimentale măsurate direct $\sigma=0.5$.

La apăsarea butonului **OK** din fereastra II, datele experimentale se prelucrează pentru nivelul de încredere setat. Pe ecranul calculatorului se afișează eroarea standard a seriei de măsurători, eroarea standard a valorii medii, valoarea medie și intervalul de încredere al rezultatului măsurătorilor. Din aceste rezultate, utilizatorul face următoarele concluzii:

- eroarea standard a seriei de măsurători practic nu depinde de numărul de măsurători în seri și este comparabilă cu valoarea 0.5 a dispersiei, utilizată în algoritmul de generare a datelor experimentale;
- eroarea standard a valorii medii se micșorează atunci când crește numărul de măsurători;

- Intervalul de încredere calculat conform formulei (1) se micșorează atunci când crește numărul de măsurători;
- Valoarea adevărată, cunoscută în cazul acestui experiment numeric pentru orice număr de date în seria de măsurători se află în interiorul intervalului de încredere.

Cu ajutorul aplicației din pagina II, utilizatorul poate prelucra o serie de măsurători reale realizată în laborator. Pentru aceasta în fereastra aplicației utilizatorul setează numărul de măsurători și introduce în celulele grilei valorile fiecărei măsurători. După aceasta setează nivelul de încredere dorit și apasă butonul **OK**. În rezultat valoarea medie și intervalul de încredere a mărimii fizice măsurate se afișează în fereastra aplicației.

Pagina III a aplicației "Generator Subiecte" este destinată profesorilor. În această pagină se generează variante de subiecte în care se cere de prelucrat manual sau cu ajutorul calculatorului un set de date experimentale generate de calculator. În fig.4 este reprodusă această pagină.

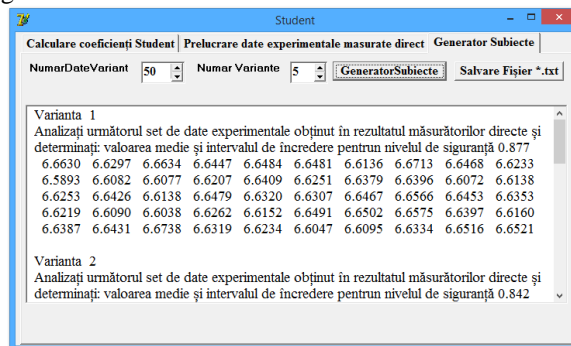


Fig.4 Generarea subiectelor de evaluare.

Pentru generarea unui set de variante, profesorul setează numărul de date experimentale care va fi creat în fiecare variantă și numărul de variante (în conformitate cu numărul de elevi/studenti). După apăsarea butonului "Generator subiecte", în fereastra aplicației apare numărul respectiv de variante de subiecte, generate în mod aleatoriu. Toate aceste subiecte pot fi vizualizate deplasând cursorul barei verticale de defilare. Cu ajutorul butonului "Salvare Fișier *.txt" se produce salvarea subiectelor într-un fișier cu extensia *.txt, care în continuare poate fi deschis și prelucrat cu orice procesor de texte.

II. CONCLUZII

Subiectele de prelucrare a datelor experimentale măsurate direct, generate de aplicația "Student" pot fi propuse elevilor și studenților pentru consolidarea cunoștințelor la fizică. Dacă numărul de date este relativ mic, rezolvarea poate fi realizată manual. În cazul unui număr mare de date problema poate fi propusă în calitate de subiect interdisciplinar la orele de informatică.

REFERINȚE

- [1] T. Crețu, V.Fălie, Prelucrarea datelor experimentale în fizică. București, Ed. Didactică, 1984.
- [2] Протасов К.В. Статистический анализ экспериментальных данных М. Мир., 2005.
- [3] F. Corunescu, A.Prodan, Modelare stochastică și simulare, Cluj-Napoca, Ed. Albastră, 2001.