

# Aplikace matematiky

---

Pavel Bureš

Algoritmy. 8.  $t$ -TEST. Využití  $t$ -rozdělení o  $f$  stupních volnosti

*Aplikace matematiky*, Vol. 12 (1967), No. 4, 327–330

Persistent URL: <http://dml.cz/dmlcz/103107>

## Terms of use:

© Institute of Mathematics AS CR, 1967

Institute of Mathematics of the Czech Academy of Sciences provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This document has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://dml.cz>

## ALGORITMY

8. *t*-TESTVYUŽITÍ *t*-ROZDĚLENÍ O *f* STUPNÍCH VOLNOSTI

PAVEL BUREŠ, prom. mat., Ústav výpočtení techniky ČSAV a ČVUT, Praha 2, Horská 3.

```

procedure ttest (n1, n2, a1, a2); value n1, n2; integer n1, n2; real array a1, a2;
comment n1 resp. n2 jsou počty naměřených veličin v souborech, a1 resp. a2 jsou
identifikátory jednorozměrných polí, obsahující jako složky jednotlivé naměřené
veličiny indexované 1, 2, ..., n1 resp. 1, 2, ..., n2;
begin integer f; real p1, p2, o1, o2, r1, r2, t, pst, pr;
procedure produce (p, o, r, n, ar);
value n; integer n; real p, o, r; real array ar;
begin integer i, j; real ab, ac, med;
comment zde pro kontrolu můžeme tisknout n a vstupní data;
ab := ac := 0;
for i := 1 step 1 until n do
  begin ab := ar [i] + ab;
    ac := ar[i]↑2 + ac end;
p := ab/n;
o := ac - ab↑2/n;
r := o/(n - 1);
for i := n - 1 step -1 until 1 do
  for j := 1 step 1 until i do
    if ar[j] > ar[j + 1] then

```

```

begin  $ab := ar[j + 1]$ ;
       $ar[j + 1] := ar[j]$ ;
       $ar[j] := ab$  end;
 $med :=$  if  $entier((n - 1)/2) = entier(n/2)$ 
      then  $ar[(n + 1) \div 2]$  else  $(ar[n \div 2] + ar[n \div 2 + 1])/2$ ;
comment nyní lze tisknout průměr  $p$ , rozptyl  $r$ , směrodatnou odchylku  $sqrt(r)$ ,
čtverec odchylky průměru  $r/n$ , odchylku průměru  $sqrt(r/n)$  a medián  $med$ ;
end produce;
real procedure  $prvd(f, t)$ ;
value  $f, t$ ;
integer  $f$ ; real  $t$ ;
begin real  $u, v, w, x, y, z$ ;
      integer  $m, k$ ;
      boolean  $b$ ;
 $u := t/sqrt(f)$ ;
 $v := arctan(u)$ ;
 $b :=$  if  $entier((f - 1)/2) = entier(f/2)$  then true else false;
 $v :=$  if  $b$  then  $2/3.14159265 \times v$  else  $sin(v)$ ;
if  $f < 3$  then goto  $uk$ ;
 $x := 1/(1 + u^2)$ ;
      if  $b$  then begin  $y := 2 \times u/3.14159265$ ;
         $z := 0$ ;  $w := 1$ ;  $m := (f - 1) \div 2$  end
      else begin  $y := u \times sqrt(x)$ ;
         $z := 1$ ;  $w := 2$ ;  $m := (f - 2) \div 2$  end;
for  $k := 1$  step 1 until  $m$  do
      begin  $y :=$  if  $z = 0$  then  $y \times x/w$  else  $y \times x \times z/w$ ;
         $v := v + y$ ;  $z := z + 2$ ;  $w := w + 2$ 
      end;
 $uk : prvd := v$ ;
end prvd, která vypočítává integrál od  $-t$  do  $t$  z  $g_f(x)$ ;
 $produce(p1, o1, r1, n1, a1)$ ;
 $produce(p2, o2, r2, n2, a2)$ ;

```

```

f := n1 + n2 - 2;
pr := if r1 > r2 then r1/r2 else r2/r1;
t := abs(p1 - p2)/sqrt((o1 + o2) * (n1 + n2)/(f * n1 * n2));
pst := (1 - prvd(f, t)) * 100;
comment zde se tiskne hodnota t, kvantil v procentech pst a poměr rozptylů pr;
end ttest;

```

Algoritmus je určen pro testování hypotéz o rovnosti dvou výběrových průměrů za předpokladu stejných rozptylů ze získaných naměřených hodnot  $y_1 \dots y_{n_1}$  a  $y'_1 \dots y'_{n_2}$ . Přesnost výsledků závisí na počtu vstupních dat, jejich měřítkovém rozsahu a přesnosti zobrazení čísel v počítači. Pro běžnou praxi je přesnost na 3–4 platná místa dostačující. Procedura *prvd* je pomalá a méně přesná pro velká  $f$ . Program vypočítává pro každý soubor  $y_1 \dots y_{n_1}$  a  $y'_1 \dots y'_{n_2}$  (odpovídající identifikátorům  $a1$  a  $a2$ ): průměr, rozptyl, směrodatnou odchylku průměru, čtverec odchylky průměru, odchylku průměru a medián. Dále se počítají tyto veličiny

$$t = \frac{|\bar{y} - \bar{y}'|}{\sqrt{\frac{\left(\sum_{i=1}^{n_1} y_i^2 - \frac{1}{n_1} \left(\sum_{i=1}^{n_1} y_i\right)^2 + \sum_{i=1}^{n_2} y_i'^2 - \frac{1}{n_2} \left(\sum_{i=1}^{n_2} y_i'\right)^2 \cdot (n_1 + n_2)\right)}{(n_1 + n_2 - 2) \cdot n_1 \cdot n_2}}},$$

kvantil *pst* v procentech  $pst = 100 \cdot (1 - 2 \int_0^t g_f(x) dx)$ , kde

$$g_f(x) = \frac{\Gamma\left(\frac{f+1}{2}\right)}{\Gamma\left(\frac{f}{2}\right) \sqrt{(\pi f)}} \left(1 + \frac{x^2}{f}\right)^{-(f+1)/2} \quad \text{a kde } f = n_1 + n_2 - 2,$$

a poměr rozptylů:

$$pr = \frac{\left(\sum_{i=1}^{n_1} y_i^2 - \frac{1}{n_1} \left(\sum_{i=1}^{n_1} y_i\right)^2\right) (n_2 - 1)}{\left(\sum_{i=1}^{n_2} y_i'^2 - \frac{1}{n_2} \left(\sum_{i=1}^{n_2} y_i'\right)^2\right) (n_1 - 1)}$$

resp.  $1/pr$  tak, aby  $pr \geq 1$ .

Kontrolní příklad:

```

n1 = 7
n2 = 2
a1 = 5; 5,5; 4,5; 4,25; 4,75; 5,25; 5,75;
a2 = 0; 3,3596163;

```

	První soubor:	Druhý soubor:
průměr	5,0	1,6798081
rozptyl	0,29166667	5,6435108
směrodatná odchylka průměru	0,54006172	2,3756075
čtverec odchylky průměru	0,041666667	2,8217554
odchylka průměru	0,20412415	1,6798081
medián	5,0	1,6798081

$$t = 4,0293$$

kvantil  $pst$  v procentech = 0,501222 ~ 5 promile

poměr rozptylů = 19,349180

Příklad je pouze ilustrativní pro ověření výše uvedeného algoritmu (předpoklad o přibližné rovnosti rozptylů není v něm splněn).

Algoritmus byl odladěn v jazyku uralgol v ÚVT a užíván rovněž v rozšířeném tvaru jako standartní program ve strojním kódu URAL 2.

[1] *A. Hald*: Statistical Theory with Engineering Applications.

[2] *J. Janko*: Statistické tabulky kap. 3.