ZASTOSOWANIA MATEMATYKI
APPLICATIONES MATHEMATICAE
XIV, 2 (1974)

ALGORITHM 33

ANNA BARTKOWIAK (Wrocław)

TEST FOR DIFFERENCE IN TRENDS IN A TWO-FACTOR EXPERIMENT WITH REPEATED MEASUREMENTS

1. Procedure declaration. Let a two-factor experiment $A\times B$ with levels $p\times q$ and r_i repetitions for each class A $(i = 1, \ldots, p)$ be given. The levels of factor B are regarded as steps along a continuum. Successive levels of factor B are interpreted as chosen points $x_1, \ldots, x_q$ of some continuous variable (e.g. time). We assume that between the observed variable y and different levels of factor B there exists the dependence

$$y = b_0 + b_1 p_1(x) + \ldots + b_k p_k(x) + e, \tag{1}$$

where e denotes the random error and $p_j(x)$ are orthogonal polynomials on the set $x_1, \ldots, x_q$.

If it is known that the coefficients $b_0, \ldots, b_k$ occurring in (1) are different for different levels of factor A (which follows from the significance of the interaction $A\times B$), then procedure *difftrend* allows an exact investigation which of the components of (1) are significantly different for different levels of factor A.

The interaction variation *ssi* is subdivided into trend components associated with difference in linear, quadratic, cubic etc. trends. The components are extracted as long as the following conditions hold:

a. the actual number of components already extracted is less than k (procedure parameter),

b. the residual of interaction variation diminished by the sum of extracted components is statistically significant,

c. the error indicator of *orthonw* (procedure parameter) does not equal 1 or 3.

Data:

p – number of levels of factor A,
q – number of levels of factor B,
$na[1:p]$ – number of repetitions in the A-classes,

$x[1:q]$ — values of levels of factor B,
$ab[1:p, 1:q]$ — mean values in the classes $A \times B$,
$b[1:q]$ — weighted means in the B-classes,
seb — residual variance (error),
df — degrees of freedom of *seb*,
ssi — sum of squares due to interaction,
k — the greatest number of components to be extracted,
alfa — significance level,
orthonw — procedure described in [1],
Ftest — function described in [2], with heading

```
real procedure Ftest(f, df1, df2, maxn);
    value f, df1, df2, maxn;
    integer df1, df2, maxn;
    real f;
```

which calculates the probability that a random variable with the F-distribution having *df1* and *df2* degrees of freedom exceeds the value f (*maxn* is an indicator showing when to use the normal approximation; usually *maxn* = 500).

Results:

k — number of extracted trend-difference components,
$ssdif[1:k]$ — sum of squares due to differences in trend of the i-th degree ($i = 1, \ldots, k$); each component has $p-1$ degrees of freedom.

2. Method used. The interaction sum *ssi* of squares with $(p-1)(q-1)$ degrees of freedom can be decomposed into at last $q-1$ components, $p-1$ degrees of freedom each, due to differences in trends of the 1-st, 2-nd, ..., $(q-1)$-st degree. The components *ssdif* are extracted as long as they appear statistically significant.

Let us assume that we have already j trend-difference components. The component $j+1$ is extracted if at least one of the following conditions hold:

(a) $ssi - \sum_{i=1}^{j} ssdif[i] > 0$,

(b) $j < k$ — the maximum number of trend-difference components to be extracted,

(c) the residual

$$R = ssi - \sum_{i=1}^{j} ssdif[i]$$

is statistically significant.

```
procedure difftrend(p,q,na,x,ab,b,seb,df,ssi,alfa,k,ssdif,
 orthonw,Ftest);
 value p,q,df,seb,alfa;
 integer k,p,q,df;
 real alfa,ssi,seb;
 array x,ab,b,ssdif;
 integer array na;
 procedure orthonw;
 real procedure Ftest;
 begin
  array pn[1:q],pt[1:q+6],wg[1:q];
  integer n,r,i,j,q1,fault;
  real s,g,s1;
  for i:=1 step 1 until q do
   wg[i]:=1.0;
  n:=0;
  for i:=1 step 1 until p do
   n:=n+na[i];
  q1:=q-1;
  orthonw(0,q,wg,x,pn,pt,fault);
  s1:=0.0;
  g:=ssi;
  for r:=1 step 1 until q1 do
   begin
    g:=g-s1;
    if g≤0∨r>k∨Ftest(g/((p-1)×(q-r))/seb,(p-1)×(q-r),df,80)>
     alfa
     then go to ettrendab;
    s1:=0.0;
    orthonw(r,q,wg,x,pn,pt,fault);
```

```
            if fault=1Vfault=3
             then go to dttrendab
             else
             begin
              for j:=1 step 1 until p do
               begin
                s:=0.0;
                for i:=1 step 1 until q do
                 s:=s+ab[j,i]×pn[i];
                s1:=s1+s↑2×na[j];
               end j;
             s:=0.0;
             for i:=1 step 1 until q do
              s:=s+b[i]×pn[i];
             ssdif[r]:=s1:=s1-s×s×n
            end fault=0Vfault=2
          end r;
        ettrendab:
          k:=r-1;
         end difftrend
```

The components *ssdif* expressing the difference in polynomial regression of degree k have the following general form:

$$ssdif[k] = \sum_{i=1}^{p} \frac{\left(\sum_{j=1}^{q} n_{ij} y_{ij.} p_k(x_j)\right)^2}{\sum_{j=1}^{q} n_{ij} [p_k(x_j)]^2} - \frac{\left[\sum_{j=1}^{q} \left(\sum_{i=1}^{p} n_{ij}\right) y_{.j.} p_k(x_j)\right]^2}{\sum_{j=1}^{q} \left(\sum_{i=1}^{p} n_{ij}\right) [p_k(x_j)]^2}.$$

The $p_k(x_j)$ should be taken with weights equal to appropriate counts n_{ij}.

Assuming equal numbers of repetitions in each class of B,

$$n_{i1} = n_{i2} = \ldots = n_{iq} = n_i, \quad \sum_{i=1}^{p} n_i = N,$$

we have

$$ssdif[k] = \frac{\sum_{i=1}^{p} n_i (\sum_{j=1}^{q} y_{ij.} p_k'(x_j))^2}{\sum_{j=1}^{q} [p_k'(x_j)]^2} - \frac{N(\sum_{j=1}^{q} y_{.j.} p_k'(x_j))^2}{\sum_{j=1}^{q} [p_k'(x_j)]^2},$$

where $p_k'(x_j)$ are generated with weights $w_1 = \ldots = w_q = 1$. Using procedure *orthonw*, we have

$$\sum_{j=1}^{q} w_j [p_k(x_j)]^2 = 1,$$

so, finally,

$$ssdif[k] = \sum_{i=1}^{p} n_i \Big(\sum_{j=1}^{q} y_{ij.} p_k'(x_j)\Big)^2 - N\Big(\sum_{j=1}^{q} y_{.j.} p_k'(x_j)\Big)^2,$$

which is used in *difftrend*.

3. Certification. Algorithm *difftrend* has been verified on the Odra 1204 computer. The values of parameters of the procedure are the following:

$$\begin{aligned} p &= 2, \\ q &= 5, \\ na &= [3, 3], \\ x &= [1, 2, 3, 4, 5], \\ ab &= \begin{bmatrix} .667 & 2 & 4 & 8 & 16 \\ 63.00 & 32 & 16 & 4 & 4 \end{bmatrix}, \\ b &= [31.834 \quad 17 \quad 10 \quad 6 \quad 10], \\ seb &= .5333, \\ df &= 16, \\ ssi &= 5428.53, \\ k &= 4, \\ alfa &= .05. \end{aligned}$$

Hence we get the following results:

$$k = 4,$$

$$ssdif = [5115.3,\ 297.19,\ 8.07,\ 8.01].$$

4. Example of application. Let us consider the time-reaction curves for a stimulator A acting on some objects (animals). The stimulator A is administrated in two dilutions ($p = 2$). The time-reaction curve is measured for $t = 1, 2, 3, 5, 10$ minutes, representing $q = 5$ levels of factor B (time). The observed values of the characteristics measured are the following:

A1	object 1	.6	.7	1.2	2.4	5.2
	object 2	.6	.7	1.4	2.6	5.0
A2	object 3	1.0	2.2	3.8	5.5	7.2
	object 4	.8	2.3	3.7	5.6	7.3
	object 5	1.0	2.4	3.6	5.7	7.4

First, we compute a two-way analysis of variance for a partially hierarchical design (details of such an analysis can be found in Winer [3]).

The table for the analysis of variance is the following:

Source	*SS* (sum of squares)	*df*	*MS* (mean square)	*F*	*alfa*
between subjects					
dilution	22.2723	1	22.2723	2637.51	.0000
res *A*	.0253	3	0.0084		
within subjects					
time	100.286	4	25.0714	2128.70	.0000
time-dilution interaction	5.185	4	1.29627	110.06	.0000
res *B*	.1413	12	0.01177		

The averages in groups $A \times B$ are given in the following table:

	B1	*B2*	*B3*	*B4*	*B5*
A1	.6	.7	1.3	2.5	5.1
A2	.933	2.3	3.7	5.6	7.3

The trend-difference components calculated by the use of procedure *difftrend* are as follows:

degree	*ssdif*	*df*	mean square	*F*	*alfa*	percent
1	1.4274	1	1.4274	121.19	.0000	27.528
2	3.6647	1	3.6647	311.16	.0000	70.679
3	.0924	1	.0924	7.84	.0160	1.781

The analysis of variance yields a highly significant dilution-time interaction term (sum of squares for interaction, i.e. $ssi = 5.18507$, with 4 degrees of freedom versus residual sum of squares $SS_{\text{res}} = .141333$

with 12 degrees of freedom). This means that the response curves for the two dilutions are different in shape. The essential features of their difference can be elucidated by the trend-difference component analysis. We decompose the interaction sum of squares ssi into 3 components, representing the difference in components calculated from over-all column averages and from group averages calculated separately for each dilution:

$$ssi = SS_1 + SS_2 + SS_3 + R.$$

Exactly, $5.1851 = 1.4274 + 3.6647 + 0.0924 + 0.0006$.

Only SS_1 and SS_2 are statistically significant which means that the essential difference between the shape of the response curve for two dilutions under investigation is associated with the linear and quadratic terms of regression.

References

[1] A. Bartkowiak, *Construction of polynomial values orthogonal on a given set of a one-dimensional variable*, this fascicle, p. 327-333.

[2] J. Morris, *Algorithm 346, F-test probabilities*, Comm. ACM 12 (1969), p. 184-185; see also: *Funkcja F-test*, Algol procedures library for the Odra 1204 computer, part 4, Elwro, Wrocław 1971.

[3] B. J. Winer, *Statistical principles in experimental design*, McGraw-Hill, New York 1970.

COMPUTING CENTRE
UNIVERSITY OF WROCŁAW
50-384 WROCŁAW

Received on 10. 3. 1973

ALGORYTM 33

ANNA BARTKOWIAK (Wrocław)

TESTOWANIE ISTOTNOŚCI RÓŻNIC SKŁADNIKÓW TRENDOWYCH W DOŚWIADCZENIU Z DWOMA CZYNNIKAMI I POWTÓRZENIAMI W PODKLASACH

STRESZCZENIE

Dane są wyniki doświadczenia $p \times q$ z dwoma czynnikami A i B. Czynnik A jest badany na p poziomach, a czynnik B — na q poziomach. Dla i-tego poziomu czynnika A mamy r_i powtórzeń w każdej podklasie czynnika B. Kolejne poziomy czynnika B interpretujemy jako wybrane punkty $x_1, \ldots, x_q$ pewnej ciągłej zmiennej (np. osi czasu). Zakładamy, że między obserwowaną zmienną y a różnymi poziomami

czynnika B istnieje zależność, dająca się przedstawić w postaci (1), gdzie e oznacza błąd losowy, a $p_j(x)$ są wielomianami ortogonalnymi na danym zbiorze $x_1, \ldots, x_q$.

Jeśli wiadomo, że współczynniki $b_0, \ldots, b_k$, występujące w zależności (1), są różne dla różnych poziomów czynnika A (co wynika z istotności interakcji $A \times B$), to procedura *difftrend* umożliwia dokładne zbadanie, które składniki zależności (1) są istotnie różne dla różnych poziomów czynnika A.

Dane:

p — liczba poziomów czynnika A,
q — liczba poziomów czynnika B,
na[*1* : *p*] — liczba powtórzeń (obiektów) w podklasach czynnika A,
x[*1* : *q*] — wartości poziomów czynnika B,
ab[*1* : *p*, *1* : *q*] — wartości średnie dla klas $A \times B$,
b[*1* : *q*] — ważone średnie w klasach czynnika B,
seb — wariancja rezydualna (błąd resztowy),
df — liczba stopni swobody *seb*,
ssi — interakcyjna suma kwadratów,
k — największa liczba czynników, które mają być wyróżnione,
alfa — poziom istotności.

Wyniki:

k — liczba wyróżnionych składników trendu,
ssdif[*1* : *k*] — suma kwadratów spowodowana różnicą w trendzie stopnia i dla $i = 1, \ldots, k$; każdy składnik ma $p-1$ stopni swobody.

Inne parametry:

Potrzebne są procedury *Ftest* (patrz [2]) i *orthonw* (patrz [1]).