

継時測定データの時点毎の多重比較

岸本 淳司・小野 裕亮
(SAS インスティテュートジャパン)

Timewise Multiple Comparisons on Longitudinal Data

Junji Kishimoto and Yusuke Ono

SAS Institute Japan

e-mail address: jpnjak@jpn.sas.com

継時的に測定された 2 群のデータについて、時点毎に群間の比較を行ったときに生じる検定の多重性を調整する方法を考察した。多変量検定を実施すれば、そもそも多重性の問題は生じない。誤差の分散構造が複合対称型の場合、岸本 (1997) が報告しているように、多変量 t 分布の積分は簡単な計算で実行することが可能であり、解析的に正確な解を得ることができる。誤差の分散構造が複合対称型でない場合、Hsu の因子分析型の近似法で多変量 t 分布の積分の近似値を得る方法が考えられる。また、標本再抽出法 (MULTTEST) で近似 p 値を得ることもできる。

MIXED, PROBMCMC, MULTTEST, Multiple Comparisons, Repeated Measures

1 反復測定データ

2 群の個体を経時的に測定したデータを想定する。

表 1 経時的に測定した 2 群のデータ

	時点	t_1	t_2	\cdots	t_k
群 1	個体 ₁	y_{11}	y_{12}	\cdots	y_{1k}
	\vdots	\vdots	\vdots	\vdots	\vdots
	個体 _m	y_{m1}	y_{m2}	\cdots	y_{mk}
	平均	μ_{11}	μ_{12}	\cdots	μ_{1k}
群 2	個体 _{m+1}	y_{m+11}	y_{m+12}	\cdots	y_{m+1k}
	\vdots	\vdots	\vdots	\vdots	\vdots
	個体 _n	y_{n1}	y_{n2}	\cdots	y_{nk}
	平均	μ_{21}	μ_{22}	\cdots	μ_{2k}

ただし、各測定値はすべて処置後のものとし、ベースラインの値は測定されていないものとする。次の例題は、遅効性薬剤の実薬群と溶媒群とを4時点で測定したデータである。

```
data A(keep=group id time y)
  B(keep=group id y1-y4);
  array yt{4} y1-y4;
  input group $ id $ @;
  do time=1 to 4;
    input y @;
    yt{time} = y;
    output A;
  end;
  output B;
  cards;
C 01  3.25 3.55 3.44 1.78
C 02  2.59 2.80 4.18 2.63
C 03  3.09 3.29 2.05 1.28
C 04  2.66 2.86 2.27 1.82
C 05  2.84 2.43 2.36 2.30
C 06  2.01 2.86 1.84 2.50
C 07  3.22 3.36 2.80 2.26
C 08  3.16 3.85 3.42 1.93
C 09  2.07 1.58 1.84 2.06
A 10  1.86 2.09 1.76 2.43
A 11  2.84 2.13 1.62 1.62
A 12  3.30 2.03 2.88 2.60
A 13  1.73 1.67 1.15 1.21
A 14  3.40 2.16 3.20 2.24
A 15  2.79 1.82 2.38 1.29
A 16  2.73 3.38 1.93 2.61
;
```

このような実験データは、通常次のような MIXED プロシジャによる反復測定モデルを用いて解析される。

```
proc mixed data=A;
  class group time id;
  model y = group time group*time / ddfm=satterth;
  repeated time / subject=id(group) type=cs;
run;
```

Type 3 Tests of Fixed Effects				
Effect	Num DF	Den DF	F Value	Pr > F
group	1	14	2.76	0.1189
time	3	42	5.22	0.0037
group*time	3	42	1.82	0.1581

2 多変量型の検定法

各時点を別々の変数と考えた多変量検定を GLM プロシジャを用いて実行することも考えられる。

```
proc glm data=B;
  class group;
  model y1-y4 = group / nouni;
  manova h=group;
run;
```

MANOVA Test Criteria and Exact F Statistics for
the Hypothesis of No Overall group Effect
H = Type III SSCP Matrix for group
E = Error SSCP Matrix
S=1 M=1 N=4.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.58758408	1.93	4	11	0.1755
Pillai's Trace	0.41241592	1.93	4	11	0.1755
Hotelling-Lawley Trace	0.70188408	1.93	4	11	0.1755
Roy's Greatest Root	0.70188408	1.93	4	11	0.1755

上の解析とよく似ているが、REPEATED ステートメントを使って希望する比較を実行させるには注意が必要である。

```
proc glm data=B;
  class group;
  model y1-y4 = group / nouni;
  repeated time 4 identity / printe;
run;
```

Manova Test Criteria and Exact F Statistics
for the Hypothesis of no time*group Effect
H = Type III SSCP Matrix for time*group
E = Error SSCP Matrix
S=1 M=1 N=4.5

Statistic	Value	F Value	Num DF	Den DF	Pr > F
Wilks' Lambda	0.58758408	1.93	4	11	0.1755
Pillai's Trace	0.41241592	1.93	4	11	0.1755
Hotelling-Lawley Trace	0.70188408	1.93	4	11	0.1755
Roy's Greatest Root	0.70188408	1.93	4	11	0.1755

正準判別分析のプログラムを使っても計算できる。

```
proc candisc data=B;
  class group;
  var y1-y4;
run;
```

Multivariate Statistics and Exact F Statistics						
	S=1	M=1	N=4.5			
Statistic	Value	F Value	Num DF	Den DF	Pr > F	
Wilks' Lambda	0.58758408	1.93	4	11	0.1755	
Pillai's Trace	0.41241592	1.93	4	11	0.1755	
Hotelling-Lawley Trace	0.70188408	1.93	4	11	0.1755	
Roy's Greatest Root	0.70188408	1.93	4	11	0.1755	

3 時点毎の比較

処置群間の効果についてより詳細に検討するため、各時点毎に対応のないt検定

$$H_{01} : \mu_{11} = \mu_{21}, H_{02} : \mu_{12} = \mu_{22}, \dots, H_{0k} : \mu_{1k} = \mu_{2k} \quad (1)$$

を反復して実施したいことがある。検定の多重性を無視するなら、この検定は MIXED プロシジャ LSMEANS ステートメントの SLICE= オプションあるいは CONTRAST ステートメントを用いて実行することができる。

```
proc mixed data=A;
  class group time id;
  model y = group time group*time / ddfm=satterth;
  repeated time / type=cs subject=id(group);
  lsmeans group*time / slice=time;
  contrast 'T=1' group 1 -1 time*group 1 0 0 0 -1 0 0 0;
  contrast 'T=2' group 1 -1 time*group 0 1 0 0 0 -1 0 0;
  contrast 'T=3' group 1 -1 time*group 0 0 1 0 0 0 -1 0;
  contrast 'T=4' group 1 -1 time*group 0 0 0 1 0 0 0 -1;
run;
```

Tests of Effect Slices						
Effect	time	Num DF	Den DF	F Value	Pr > F	
group*time	1	1	41.7	0.10	0.7506	
group*time	2	1	41.7	5.93	0.0193	
group*time	3	1	41.7	3.10	0.0855	
group*time	4	1	41.7	0.04	0.8451	

Contrasts				
Label	Num DF	Den DF	F Value	Pr > F
T=1	1	41.7	0.10	0.7506
T=2	1	41.7	5.93	0.0193
T=3	1	41.7	3.10	0.0855
T=4	1	41.7	0.04	0.8451

これら4本のコントラストについて、検定の多重性を調整する手段は MIXED プロシジャでは提供されていない。

4 誤差に複合対称型が仮定できる場合

時点毎の群間比較について多重性を調整した検定を行うためには、各比較のコントラスト間の相関をパラメタとして知る必要がある。現実には我々が知ることができるのは標本から計算された推定値であって、真の相関の値はわからない。また、任意の相関構造をもつ多変量t分布の数値積分は困難である。時点間の群間比較について多重性を調整することが流行らないのは、このような理論的困難が存在するからである。

時点間の相関構造が複合対称型に限定された場合、群間比較コントラスト間の相関構造も複合対称型となり、その多変量t分布の積分計算は容易である。しかも、PROBMC関数のDunnett型比較のための機能を使って実際に計算させることができる(岸本 1997a 1997b)。

この計算法を確認するため、シミュレーションを行った。時点数を4、サンプルサイズは各群3例計6例とし、公称有意水準を1%,2.5%,5%の3通り、母相関を0.1から0.9まで0.1刻みに設定した帰無仮説シミュレーションデータを1000件作成し、所定の計算法で有意な差が認められた割合を調べた。

各時点の平均間の真の相関は $cov/(cov + \sigma^2)$ である。ここで、 cov は同一被験者内の誤差間の共分散であり、 σ^2 は誤差分散である。相関を推定する際、 cov と σ^2 の推定値を直接計算式の cov と σ^2 に代入した。 cov が負の値になったときは、多変量t分布の積分計算が難しくなるので、 cov 推定時に下限をゼロに設定した。

シミュレーションの結果を表1に示す。帰無仮説が実際に棄却される確率が公称有意水準より大きい傾向があるようにも見受けられるが、一般に正確な結果が得られている。計算に利用したプログラムは付録に示す。

表1: 多変量t法による輪切検定のシミュレーション結果

α	母相関係数								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1.0%	0.8	1.0	1.0	1.3	1.3	1.7	1.4	1.3	1.0
2.5%	2.0	2.4	2.5	2.8	2.8	3.0	2.7	2.8	2.4
5.0%	5.4	5.7	5.4	5.2	5.9	5.6	5.3	5.3	4.8

5 誤差に複合対称型が仮定できない場合

誤差に複合対称型が仮定できない場合は、その対処法を考慮しなければならない。1つの有力な方法は、Hsuの因子分析的近似法である。また、解析的な解を得ることをあきらめて、シミュレーションでp値を計算することも考えられる。MIXEDプロシジャを使って求めた最小二乗平均とその共分散を使ってシミュレーションで近似p値を求める方法がWestfall et. al.(1999)のpp.222 - 225に紹介されている。また、MULTTESTプロシジャを直接適用して、標本再抽出法により多重性を調整した検定を実施することもできる。

MULTTESTプロシジャによる標本再抽出法の能力を調べるため、シミュレーション研究を行った。各種設定は多変量t分布法のとおりとした。結果を表2に示す。

表 2: MULTTEST プロシジャによる輪切検定のシミュレーション結果

α	母相関係数								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1.0%	0.2	0.0	0.0	0.2	0.1	0.1	0.1	0.1	0.1
2.5%	1.0	1.0	0.8	0.6	0.7	0.9	0.7	0.8	0.7
5.0%	2.7	2.8	3.3	3.3	4.1	4.2	3.3	3.6	4.3

明らかに有意になりにくい傾向が読み取れる。

6 おわりに

反復測定データでの群間比較にはさまざまな方法がある。多変量検定のアプローチでは、一般に検出力が低い。また、GLM プロシジャの指定方法に注意しなければならない。t 検定の多重性を調整する方法は、検定間の相関を考慮する方法なら検出力も高い。MULTTEST プロシジャを直接利用する標本再抽出法は、汎用性は広いものの、検出力の点で多重性を調整した t 検定に劣る。

参考文献

- [1] Hsu, J.C. (1992), The Factor Analytic Approach to Simultaneous Inference in the General Linear Model, Journal of Computational Statistics and Graphics, 1, 151-168.
- [2] Westfall, P.H., Tobias, R.D., Rom, D., Wolfinger, R.D. and Hochberg, Y. (1999) Multiple Comparisons and Multiple Tests Using the SAS System, Cary, NC: SAS Institute Inc.
- [3] 岸本 淳司 (1997a) 経時データにおける時点毎の比較の多重性調整法, 日本統計学会第 65 回大会講演報告集.
- [4] 岸本 淳司 (1997b) 経時データの多重比較法, 日本 SAS ユーザー会論文集.

付録

```
*****
* Program of 1st Simulation
*****
%macro test(by=1000, ranstd=1, n1=3, n2=3, seed=12345);
options nonotes nocenter ls=64 ps=30;
data data1;
drop n1 n2;
array n[2] (&n1. &n2.);
do by=1 to &by;
do group=1 to 2;
do id=1 to n{group};
s=&ranstd*rannor(&seed);
do t=1 to 4;
y=s+rannor(&seed);
output;
end;
end;
end;
end;
run;
ods listing close;
```

```

***** Fit split-plot model *****
***** Negative estimate of CS covariance is fixed to zero *****
proc mixed data=data1;
  by by;
  class t group id ;
  model y=group t group*t / ddfm=satterth;
  random id(group);
  lsmeans group*t /pdiff;
  make 'diffs' out=out1(where=(t=_t));
  make 'covparms' out=out2;
run;
ods listing;
data out2;
  keep by rho estimate;
  set out2;
  lest=lag(estimate);
  if covparm='Residual' then do;
    rho=(lest)/(estimate+lest);
    output;
  end;
run;
data result;
  KEEP BY T StdErr DF tValue Probt P RHO one check;
  merge out1 out2;
  by by ;
  one=1;
  *****
  L1=SQRT(RHO); L2=SQRT(RHO); L3=SQRT(RHO); L4=SQRT(RHO);
  *****
  IF RHO^=0 THEN DO;
    IF ABS(ROUND(tValue,0.01))<=0.01 THEN P=1;
    ELSE P=1-PROBMC("DUNNETT2",ABS(tValue),,DF,4,OF L1-L4);
    check=0;
  END;
  ELSE DO;
    /** If rho=0 then... ***/
    P=1-(1-Probt)**4; /** DO Sidak Adjustment ***/
    check=1;
  END;
run;
*****
proc format;
  value pv (multilabel)
    0.05<-high='NonSig'
    low-0.05 ='Under 5.0%'
    low-0.025='Under 2.5%'
    low-0.01 ='Under 1.0%';
run;
*****
footnote "STD: &ranstd. **** n1:&n1. n2:&n2.";
title 'Per-Comparison Error';
proc means data=result n;
  class p /mlf;
  var one;
  format p pv.;
run;
*****
PROC MEANS DATA=RESULT NOPRINT ;
  BY BY;
  VAR P;
  ID ONE check;
  OUTPUT OUT=OUTMIN1 MIN=MINP;
RUN;
title 'FamilyWise Error';
proc means data=OUTMIN1 N;
  class minp /mlf;
  var one;
  format minp pv.;
run;
title; title2;
footnote;
%mend test;

```