

第 2 回 続高橋セミナー

平均への回帰を考慮した投与前値の調整

2012 年 8 月 17 日

高橋 行雄
BioStat 研究所(株)

要約

平行群間試験でランダム化による群分けを行なっても, 処置前の群平均が等しくなることはまれで, 統計的に有意な差が出てしまうこともある. 処置前値を共変量とする ANCOVA は, 平均への回帰現象を調整する方法として知られている. 線形混合モデルを用いて処置後の全ての時点を反復測定, 処置前値を共変量とする反復測定 ANCOVA (R-ANCOVA) により推定された主要な時点での群間差は, ANCOVA よりもさらに平均への回帰現象を調整することができ, 検出力も高い. 群ごとに処置前値を設定する一般的な反復測定 ANOVA は, 平均への回帰現象を調整することができない. しかしながら, 処置前値を ANCOVA と同様な考え方で群によらず共通とする (R-C-ANOVA) は, R-ANCOVA よりもさらに平均への回帰現象の調整に優れ, さらに検出力も高い. R-C-ANCOVA の平均への回帰を減弱される能力は, 処置前と処置後の 2 時点でも ANCOVA に比べ高い. これらの結果から, 経時データの解析において処置後の特定の時点に対する ANCOVA に変えて, 測定時点を全て用いた R-ANCOVA あるいは R-C-ANOVA を使うことを推奨する.

目次

1. はじめに -----	1
2. 検討方法 -----	3
3. 共分散分析の頑健性 -----	7
4. R-ANCOVA と R-C-ANOVA の比較 (時点数=5) -----	10
5. 時点数を変えた場合 -----	12
6. 考察-----	15
SASプログラム-----	18
[付録 1]～[付録 8] 引用文献の概要とコメント -----	21～36

1. はじめに

**経時データにおける
平均への回帰現象を減弱させる
統計モデルの比較**

BioStat研究所(株)
高橋 行雄

2012.5.26 高橋行雄 1

群間で処置前の平均値が異なる

- ◆ 平行群間試験で数週間ごとに反復測定されるような臨床検査データを対象とする。
- ◆ 群ごとに投与前の観測データが存在する。
- ◆ 症例をランダムに割り付ける方法では、群ごとの処置前の平均値に何らかの差が生ずる。
- ◆ 処置後値を反応 Y , 処置前値 X を共変量とする **ANCOVA** の適用が半世紀前から推奨されている。

2012.5.26 高橋行雄 2

平行群間試験で反復測定されるような経時データに対し、処置前と処置後のある時点での差(DIFF)を用いた群間比較は、時点間の相関が低くなるにつれ DIFF の分散が増大し、検出力の低下を招く。

そのため、処置前値を共変量とした共分散分析(ANCOVA)の適用が半世紀前から推奨されている[文献・付録 1]。ランダム化による群分けを行なった場合、処置前の群平均が等しくなることはまれで、統計的に有意な差が出てしまうこともある。

平均への回帰現象により、処置前の平均が大きい群は、処置後に全体の平均に近づき、処置前の平均が小さい群も、処置後に全体の平均に近づき、DIFF による群間比較にはバイアスが生ずる。

このバイアスは、処置前値を X , 処置後値を Y , X と Y の分散は等しいとした場合の相関を ρ , 処置前値の群間差を δ としたときに、DIFF の群間差のバイアスは、 $\delta(1-\rho)$ となる[文献・付録 2]。ANCOVA は、このようなバイアスを除去するための標準的な方法として用いられている。

線形混合モデルの適用

- ◆ 処置前値を共変量とし、処置後の複数時点のデータを用いた**反復測定 ANCOVA(R-ANCOVA)**を考える。
- ◆ **R-ANCOVA**で処置後のある時点(最後時点)の群間比較を考え、**ANCOVA** に対する検出力を比較する。
- ◆ 時点数の増加に伴い **R-ANCOVA** のある時点での検出力の増大が期待できる。

2012.5.26 高橋行雄 3

R-ANCOVA の計画行列の一部

群	症例	時点	切片	前値	群	時点	群×時点	反応				
i	k	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y	
P	j	1	1	y_{1j0}	1	1	0	0	1	0	0	y_{1j1}
		2	1	y_{1j0}	1	0	1	0	0	1	0	y_{1j2}
		3	1	y_{1j0}	1	0	0	1	0	0	1	y_{1j3}
		4	1	y_{1j0}	1	0	0	0	0	0	0	y_{1j4}
A	j'	1	1	$y_{2j'0}$	0	1	0	0	0	0	0	$y_{2j'1}$
		2	1	$y_{2j'0}$	0	0	1	0	0	0	0	$y_{2j'2}$
		3	1	$y_{2j'0}$	0	0	0	1	0	0	0	$y_{2j'3}$
		4	1	$y_{2j'0}$	0	0	0	0	0	0	0	$y_{2j'4}$

$j = 1, 2, \dots, 10$ 固定効果 $X(80 \times 9)$, 変量効果 $Z(80 \times 20)$

2012.5.26 高橋行雄 4

線形混合モデルを用いることにより, 処置後の複数時点のデータを同一症例からの反復測定とみなし, 処置前値を共変量とした反復測定 ANCOVA (R-ANCOVA) は, ANCOVA よりも時点数の増加に比例して同じ傾きを持つ回帰直線をあてはめるため, 回帰係数 β の推定値の分散が小さくなると期待される。

そのために, 処置後のある時点のみを用いた ANCOVA の群間比較に比べ, R-ANCOVA で同一時点での群間差の分散が小さくなり, 検出力が大きくなると期待される[文献・付録 3]。

反復測定 ANOVA

- ◆ 処置前も処置後も全ての時点を用いた反復測定 ANOVA (R-ANOVA) を考える。
- ◆ R-ANOVA では, 群×時点が有意な差である場合に, 時点ごとの群間比較を引き続き行なうことが妥当と思われる。
- ◆ これは, 処置後のある時点 (POST) での群間比較を行なうことと同等であり, 平均への回帰現象によるバイアスを受ける。

5

R-ANOVA は, 処置前も処置後も全ての時点を反復測定とした解析方法である。R-ANOVA では, 群×時点が有意な差であることをもって, 群間に経時変化パターンに差があるとみなし, 時点ごとに輪切りの ANOVA による群間比較を引き続き行なうことが妥当とされている。しかしながら, この方法は, 処置後のある時点 (POST) での群間比較を行なうことと同等であり, 平均への回帰現象によるバイアスを受ける。

処置前値を各群で共通

- ◆ 処置前値を各群で共通とする反復測定 ANOVA (R-C-ANOVA) を考える。
- ◆ 全体の症例数を N , 時点数を p とした場合
 - R-ANCOVA の場合データ数は, $N(p-1)$
 - R-C-ANOVA の場合データ数は Np と N 人多くなる。
- ◆ 処置後のある時点での群間比較を考える
- ◆ R-C-ANOVA の検出力が, R-ANCOVA よりも, 解析に用いるデータ数の増加に伴ない大きくなること期待される。

6

R-C-ANOVA の計画行列の一部

群		症例	時点	前	群×時点									反応
i	j	k	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y		
P	$i=1$	0	1	0	0	0	0	0	0	0	0	$y'_{1/0}$		
		1	0	1	0	0	0	0	0	0	0	$y'_{1/1}$		
		2	0	0	1	0	0	0	0	0	0	0	$y'_{1/2}$	
		3	0	0	0	1	0	0	0	0	0	0	$y'_{1/3}$	
A	$i=2$	4	0	0	0	0	1	0	0	0	0	$y'_{1/4}$		
		0	1	0	0	0	0	0	0	0	0	$y'_{2/0}$		
		1	0	0	0	0	0	1	0	0	0	$y'_{2/1}$		
		2	0	0	0	0	0	0	1	0	0	$y'_{2/2}$		
		3	0	0	0	0	0	0	0	1	0	$y'_{2/3}$		
		4	0	0	0	0	0	0	0	0	1	$y'_{2/4}$		

7

R-ANOVA では, ランダム化前の処置前値を群ごとに分離して解析しているが, ANCOVA では, 処置前値を共変量とすることにより群で共通な変数とし, これが処置前値の群間差を補正する役割を担うことになっている。

同様な考え方で, 処置前値を各群で共通とする反復測定 ANOVA (R-C-ANOVA) [文献・付録 4] が提唱されている. R-C-ANOVA は, 処置前値も応答変数 Y に含めて解析することが, R-ANCOVA と異なる.

全体の症例数を N , 時点数を p とした場合に, R-ANCOVA の場合データ数が $N \times (p-1)$ であるのに対し, R-C-ANOVA の場合は $N \times p$ となり, 解析に用いるデータ数が常に N 人分多くなる.

これにより, 症例間の分散ならびに症例内の誤差分散の揺らぎが小さくなり, また誤差分散の自由度も大きくなることから, 処置後の群間差について R-C-ANOVA の検出力が R-ANCOVA よりも, 大きくなるのが期待される [文献・付録 5].

時点数が 2 の場合については, POST, DIFF, ANCOVA について, 平均への回帰現象が第 1 種の過誤におよぼす影響, 検出力におよぼす影響を明らかにし, 処置後の時点数を増やした場合の R-ANCOVA および R-C-ANOVA の検出力を SAS V9 によるシミュレーションにより検討することにした.

2. 検討方法

シミュレーション条件					
◆ 時点:	0	1	2	3	4
2 時点	○				○
3 時点	○		○		○
4 時点	○		○	○	○
5 時点	○	○	○	○	○
◆ 平均: P 群	100	97.5	95	92.5	90
A 群	100	95	90	85	80 (α エラー用 90)
◆ 分散	全体の分散を 144 とし, 時点間の相関 ρ に応じて, 症例間と症例内の分散に分配				
◆ 時点間の相関構造:	CS type				
◆ P 群:	10 例, A 群 10 例, 10,000 組				
2012.5.26 高橋行雄					
					8

- 1) 処置は 2 群 (P 群, A 群) とし, 正規乱数の発生により, 1 群当たり 10 症例, 全体で 20 症例を 1 組とした.

- 2) 時点数は処置前も含め, 2, 3, 4, 5 時点とし, 時点間の相関構造を複合同称型とした.
- 3) 症例間と症例内の全体の分散を $\sigma_B^2 + \sigma_W^2 = 144$ と一定とし, 時点間の相関 ρ によって症例間の分散 σ_B^2 と症例内の分散 σ_W^2 に配分した.
- 4) 同一症例内では共通の σ_B^2 が, それぞれの時点には, 互いに独立な σ_W^2 が付加されるようにした.
- 5) なお, 時点間の相関 ρ は, 0.05, 0.1, 0.2, . . . , 0.9, 0.95 とし, それぞれの相関で 10,000 組作成した.

処置前値の群間差での区分

◆ 処置前値について対応のない t 検定を行い, 得られた両側確率で 0.4 となる t 統計量により 3 区分とした.

- 1 (P>A) P群の平均値がA群より大きい場合
- 2 (P=A) 同じとみなす場合
- 3 (P<A) P群の平均値がA群より小さい場合

2012.5.26 高橋行雄 9

処置前値について対応のない t 検定を行い, 得られた t 統計量により

- 1) P 群の平均値が A 群より大きい場合 (P>A),
- 2) 同じとみなす場合 (P=A),
- 3) P 群の平均値が A 群より小さい場合 (P<A)

の 3 区分とした. なお, 区分けは両側確率で 0.4 となるような t 値とした.

時点数が 2 の場合について, P 群 A 群ともに処置前の平均値が 100, 処置後は 90 と設定し, 3 区分のそれぞれに対する POST, DIFF, ANCOVA の第 1 種の過誤を求めた.

検出力の算出のために, 処置後は P 群が 90, A 群は 80 と, A 群の平均値が 10 小さくなるように設定した. これにより, 時点間の相関に関連した平均への回帰現象に対する ANCOVA の頑健性がどの程度であるかを検討した.

時点数が 5 の場合に $t = 0, 1, 2, 3, 4$ とし,

P 群の固定効果にはそれぞれ

100, 97.5, 95, 92.5, 90

を割り当て, A 群は

100, 95, 90, 85, 80

を割り当て経時的变化を直線的な減少とした.

時点数が 4 の場合は $t = 0, 2, 3, 4$ と $t = 1$ を除き,
 時点数が 3 の場合は $t = 0, 2, 4$ とさらに $t = 3$ を除き,
 時点数が 2 の場合は $t = 0, 4$ とさらに $t = 2$

を順次除くことにした. これにより, 処置後の最終時点の群間差の算出を $t = 4$ 時点目とすると, 時点数が変化しても, P 群に対し A 群平均値が 10 減少したときの検出力の比較が行えるようにした.

処置後 4 時点の R-ANCOVA では, 表 1 に示すようなデザイン行列,

- 切片 x_0 : 全て 1,
- 前値 x_1 : y_{ij0} または x_{ij} , $i=1,2, j=1,2,\dots,10$,
- 群 x_2 : P 群の場合 1, A 群の場合 0,
- 時点 x_3 : $t=1$ の場合 1, それ以外 0
- 時点 x_4 : $t=2$ の場合 1, それ以外 0
- 時点 x_5 : $t=3$ の場合 1, それ以外 0
- 群 \times 時点 x_6 : $x_2 \times x_3$
- 群 \times 時点 x_7 : $x_2 \times x_4$
- 群 \times 時点 x_8 : $x_2 \times x_5$

を固定効果とし,

症例 _{ij} ($i=1,2, j=1,2,\dots,10$)

を变量効果とした REML 法による線形混合効果モデル (MIXED モデル) を適用した. 20 症例の固定効果のデザイン行列 \mathbf{X} の行方向は 20 症例 \times 4 時点 = 80, 列方向は, 9 変数の行列となる. 变量効果のデザイン行列 \mathbf{Z} の行方向は 20 症例 \times 4 時点 = 80, 列方向は 20 症例それぞれについて次に示すような標示型のダミー変数 z_1, z_2, \dots, z_{20} となる.

变量効果のデザイン行列 \mathbf{Z}

群	症例	時点	群 \times 症例			
			z_1	z_2	\dots	z_{20}
i	j	t				
1	1	1	1	0		0
1	1	2	1	0		0
1	1	3	1	0		0
1	1	4	1	0		0
\vdots	\vdots	\vdots	\vdots	\vdots		\vdots
2	10	1	0	0		1
2	10	2	0	0		1
2	10	3	0	0		1
2	10	4	0	0		1

推定された $\hat{\sigma}_b^2$ と $\hat{\sigma}_w^2$ を合成した分散については satterthwaite の方法により合成分散の自由度を求め, $t=4$ 時点での 2 群間の差について t 検定の p 値を求め $p < 0.05$ を有意差ありとした.

表 1 R-ANCOVA のデザイン行列 X (20 例中 2 例分)

群	症例	時点	切片	前値	群	時点					群×時点	反応
i	j	k	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
P	j	1	1	y_{1j0}	1	1	0	0	1	0	0	y_{1j1}
		2	1	y_{1j0}	1	0	1	0	0	1	0	y_{1j2}
		3	1	y_{1j0}	1	0	0	1	0	0	1	y_{1j3}
		4	1	y_{1j0}	1	0	0	0	0	0	0	y_{1j4}
A	j'	1	1	$y_{2j'0}$	0	1	0	0	0	0	0	$y_{2j'1}$
		2	1	$y_{2j'0}$	0	0	1	0	0	0	0	$y_{2j'2}$
		3	1	$y_{2j'0}$	0	0	0	1	0	0	0	$y_{2j'3}$
		4	1	$y_{2j'0}$	0	0	0	0	0	0	0	$y_{2j'4}$

群間差

$j = 1, 2, \dots, 10$ 固定効果 $X (80 \times 9)$, 変量効果 $Z (80 \times 20)$

時点数が 2 以上の R-C-ANOVA は, 表 2 に示すように,

- 前値 x_0 : P 群, A 群共に $t=0$ の場合 1, それ以外 0,
- 群×時点 x_1 : P 群の $t=1$ の場合 1, それ以外 0,
- 群×時点 x_2 : P 群の $t=2$ の場合 1, それ以外 0,
- :
- 群×時点 x_8 : A 群の $t=4$ の場合 1, それ以外 0,

を固定効果とし,

表 2 R-C-ANOVA のデザイン行列 X (20 例中 2 例分)

群	症例	時点	前	群×時点								反応
i	j	k	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
P	j	0	1	0	0	0	0	0	0	0	0	y_{1j0}
		1	0	1	0	0	0	0	0	0	0	y_{1j1}
		2	0	0	1	0	0	0	0	0	0	y_{1j2}
		3	0	0	0	1	0	0	0	0	0	y_{1j3}
A	j'	0	1	0	0	0	0	0	0	0	0	$y_{2j'0}$
		1	0	0	0	0	0	1	0	0	0	$y_{2j'1}$
		2	0	0	0	0	0	0	1	0	0	$y_{2j'2}$
		3	0	0	0	0	0	0	0	1	0	$y_{2j'3}$
A	j'	4	0	0	0	0	0	0	0	0	1	$y_{2j'4}$

群間差

$j = 1, 2, \dots, 10$ 固定効果 $X (100 \times 9)$ 変量効果 $Z (100 \times 20)$

症例 $_{ij}$ ($i=1,2, j=1,2,\dots,10$)

を変量効果とした MIXED モデルを適用した. R-ANCOVA と同様に, $t=4$ 時点での 2 群間の差の推定値について t 検定を行なった.

3. 共分散分析の頑健性

3.1. 第1種の過誤の比較

P 群と A 群の処置前値について t 検定を行い $p<0.4$ となる t 値により 3 区分化した. 全体で 10,000 組なので, 1:P>A 区分は, 2,000 組内外となり, おおよその P 群の平均は 103.72, A 群の平均は 96.28 となり, 3:P<A もおおよそ 2,000 組内外となる. 2:A=B となるのは, 残りの 60%, 6,000 組内外であった.

処置前の 3 区分ごとに, $\alpha=0.05$ とし, 処置前値を共変量, 時点 $t=4$ を処置後値とした ANCOVA による群間比較を行ない, 有意となる回数の割合を表 3 に第 1 種の過誤として示した. 時点間の相関に関係なく, 名目の有意水準を保っていた.

表 3 ANCOVA の第 1 種の過誤 (%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	4.9	4.8	5.3	4.6	4.3	5.2	5.0
2:P=A	5.0	4.9	5.0	4.9	4.9	4.9	4.7
3:P<A	5.6	4.8	4.8	5.6	4.7	4.6	4.1

POST の場合は, 時点間の相関が 0.30 以下であれば, 表 4 に示すように区分によらず名目の第 1 種の過誤を保っているかのように思われるが, 時点間の相関が 0.50 以上となると 2:P=A の場合 2.7%と 3%を切り, 時点間の相関が 0.9 の場合には 0%となってしまふ.

それに対し, 1:P>A, 3:P<A の場合に時点間の相関が 0.7 以上となると 10%を超え始める. もちろん, この区分を併合すれば, 時点間の相関にかかわらず名目の 5%となる.

表 4 POST の第 1 種の過誤 (%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	5.4	5.1	6.2	8.5	9.3	12.6	13.5
2:P=A	5.1	4.8	4.4	2.7	1.2	0.0	0.0
3:P<A	5.4	5.4	5.9	7.2	10.9	11.2	12.3

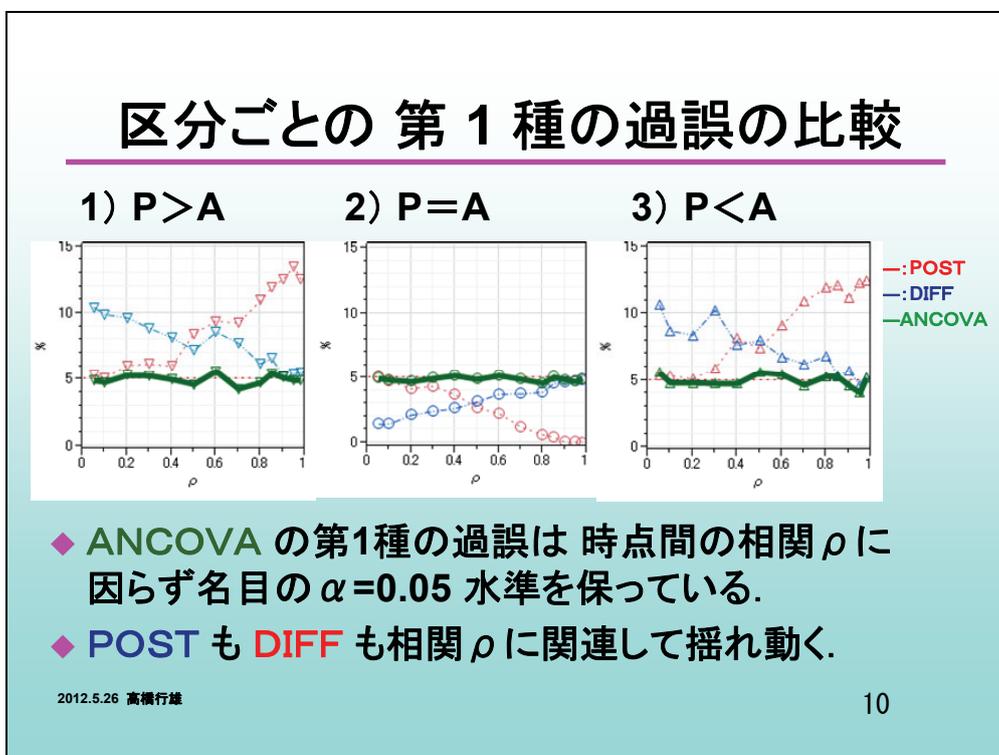
DIFF の場合は, 時点間の相関が 0.90 以上ならば, 表 5 に示すように区分によらず名目の第 1 種の過誤を保っているかのように思われるが, 時点間の相関が 0.70 以下となると 2:P=A の場

合 3.8%と 4%を切り, 時点間の相関が 0.05 の場合には 1.4%となってしまう。

それに対し, 1:P>A, 3:P<A の場合に時点間の相関が 0.5 以下となると 8%と 2:P=A の倍となり始める。もちろん, この区分を併合すれば, 時点間の相関にかかわらず名目の 5%となる。

表 5 DIFF の第 1 種の過誤 (%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	10.5	9.9	8.9	7.3	7.8	5.3	5.5
2:P=A	1.4	1.4	2.4	3.2	3.8	4.6	4.8
3:P<A	10.7	8.8	10.3	8.0	6.3	5.8	4.6



3.2. 検出力の比較

表 6 に示すように ANCOVA の検出力は, 時点間の相関が低くなれば低下するが, 区分によらず, ほぼ同等みなせる。

表 6 ANCOVA の検出力 (%)

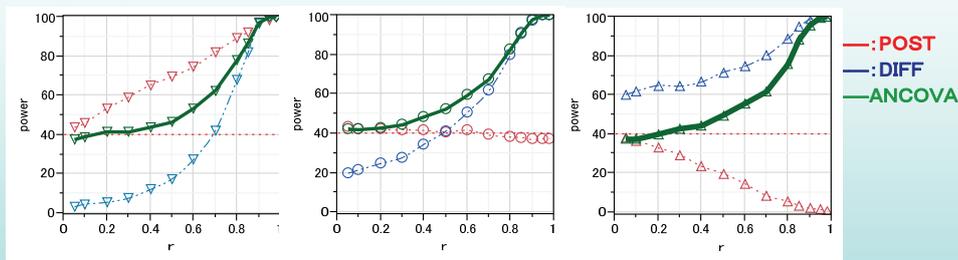
区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	37.5	39.0	41.4	46.4	62.6	97.0	99.9
2:P=A	42.2	42.0	44.4	52.4	67.5	97.8	100.0
3:P<A	37.3	37.4	42.7	49.6	61.7	95.8	99.9

区分ごとの検出力の比較

1) $P > A$

2) $P = A$

3) $P < A$



- ◆ ANCOVA の検出力は 時点間の相関 ρ に因らずどの区分でも安定している
- ◆ POST も DIFF も相関 ρ に関連して大きく揺れ動く。

2012.5.26 高橋行雄

11

ただし, 1: $P > A$ および 3: $P < A$ の場合, 2: $P = A$ に比べて時点間の相関が 0.5 の場合には $46.4 - 52.4 = -6.0\%$, $49.6 - 52.4 = -3.8\%$ のように, 検出力がわずかに低下傾向を示していて, 平均的には, -3.6% の検出力の低下となった。

区分ごとの ANCOVA の検出力

相関	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1: $P > A$	37.5	39.0	41.4	46.4	62.6	97.0	99.9
2: $P = A$	42.2	42.0	44.4	52.4	67.5	97.8	100.0
3: $P < A$	37.3	37.4	42.7	49.6	61.7	95.8	99.9

- ◆ 時点間の相関が低くなれば低下するが, 区分によらず, ほぼ同等みなせそうである。
- ◆ ただし, 1: $P > A$ および 3: $P < A$ の場合, 2: $P = A$ に比べて検出力が常に小さくなっている。
- ◆ 平均的には -3.6% の低下である。

2012.5.26 高橋行雄

12

表 6 再掲

POST の検出力を表 7 に示す. 処置前の区分がなければ, 当然ながら時点間の相関によらず, POST の検出力は 42% 前後でほぼ一定である。

表 7 POST の検出力 (%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	43.8	46.0	59.1	69.6	82.1	96.9	98.5
2:P=A	43.1	42.2	41.9	40.8	39.2	36.9	37.4
3:P<A	38.0	36.7	29.4	19.8	8.4	1.9	1.3

しかしながら, 1:P>A の場合に時点間の相関が高くなるにつれ, 検出力は増大する. これは, 1:P>A 区分の場合, P 群の処置前の平均は 103.72, A 群の平均は 96.28, 差は 7.54 なので, 時点間の相関を ρ とした場合に POST の検定の P 群と A 群の差 10.0 に $7.54 \times \rho$ が加算されて, 有意な差が出やすくなることによる.

逆に, 3:P<A 区分の場合には, P 群と A 群の差 10.0 から $7.54 \times \rho$ が差し引かれて有意な差が出にくくなっている.

DIFF の場合の検出力を表 8 に示す. 時点間の相関が 0.5 の場合, 1:P>A 区分では, 検出力が 17.5%と落ち込む. 3:P<A 区分では, 逆に 71.5%と増大する. 1:P>A 区分の場合 P 群の減少程度が増加し, A 群の減少は緩和し両群間の差が縮まり検出力が低下している.

表 8 DIFF の検出力 (%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	3.3	4.6	7.7	17.5	42.3	96.5	100.0
2:P=A	19.8	21.5	27.8	41.2	61.7	97.6	100.0
3:P<A	60.3	62.0	64.6	71.7	80.3	98.8	100.0

逆に, 3:P<A 区分の場合には, P 群と A 群の DIFF の群間差が拡大し検出力が増加している. 相関が大きい場合に検出力は 3 区分とも 100%なるが, これは, DIFF の群間差の分散が 0 に近づくことによる.

4. R-ANCOVAとR-C-ANOVAの比較(時点数=5)

表 6 の ANCOVA に比べて表 9 に示した R-ANCOVA の検出力は全般的に増大している. 相関が 0.5 の場合を例にとると, 1:P>A 区分では $54.8 - 46.4 = 8.4\%$, 3:P<A 区分でも $55.6 - 49.6 = 6.0\%$ と検出力が, 増加している.

さらに, 2:P=A 区分でも $55.0 - 52.4 = 2.6\%$ と検出力が増大している. 他の相関の場合にも共通しているが 1:P>A, 3:P<A 区分での検出力の伸びが ANCOVA に比べて大きい.

時点数が(5-1)のR-ANCOVA

- ◆ 群あたり10例, 2群で20例, 時点数が4なので, 全体でデータ数は80行となる.
- ◆ 時点が4の場合のP群とA群の差は, x_2 の推定値で得られる.

群	症例	時点	切片	前値	群	時点	群×時点	反応				
i	j	k	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
P	j'	1	1	$y'_{1/0}$	1	1	0	0	1	0	0	$y_{1/1}$
		2	1	$y'_{1/0}$	1	0	1	0	0	1	0	$y_{1/2}$
		3	1	$y'_{1/0}$	1	0	0	1	0	0	1	$y_{1/3}$
		4	1	$y'_{1/0}$	1	0	0	0	0	0	0	$y_{1/4}$
A	j'	1	1	$y'_{2/0}$	0	1	0	0	0	0	0	$y_{2/1}$
		2	1	$y'_{2/0}$	0	0	1	0	0	0	0	$y_{2/2}$
		3	1	$y'_{2/0}$	0	0	0	1	0	0	0	$y_{2/3}$
		4	1	$y'_{2/0}$	0	0	0	0	1	0	0	$y_{2/4}$

2012.5.26 高橋行雄 13

時点数が(5-1)のR-ANCOVA

相関	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	43.4	43.5	46.9	54.8	68.7	98.1	99.9
2:P=A	44.7	45.8	48.3	55.0	72.0	98.8	100.0
3:P<A	44.9	44.7	45.6	55.6	69.9	97.7	99.9

- ◆ 区分による検出力の差が見えにくくなる
- ◆ 処置前・処置後の ANCOVA に比べて検出力が増加している.
- ◆ 相関が0.5の場合は: 1:P>A区分 8.4%の増加, 3:P<A区分 6.0%の増加, さらに, 2:P=A区分でも 2.6%の増加となっている.

2012.5.26 高橋行雄 14

表9 R-ANCOVA(5-1)時点の検出力(%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	43.4	43.5	46.9	54.8	68.7	98.1	99.9
2:P=A	44.7	45.8	48.3	55.0	72.0	98.8	100.0
3:P<A	44.9	44.7	45.6	55.6	69.9	97.7	99.9

表10に示すようにR-C-ANOVAでは, R-ANCOVAに比べて1:P>A, 3:P<A区分での検出力の伸びが約3%程度増加し, 2:P=A区分でも約1%の増加している. その結果, 2:P=A区分に比べて1:P>Aおよび3:P<A区分での検出力の低下がほとんど起きていない. これは, 平均への回帰現象の影響をほとんど受けていないことになる.

表10 R-C-ANOVA(5時点)の検出力(%)

区分	時点間の相関						
	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	45.3	45.7	50.5	59.3	73.7	98.5	100.0
2:P=A	45.1	47.0	49.2	56.6	73.3	99.1	100.0
3:P<A	47.4	47.7	49.0	59.4	73.4	98.9	100.0

時点数が5のR-C-ANOVA

- ◆ 群あたり10例, 2群で20例, 時点数が5なので, 全体でデータ数は100行となる.
- ◆ 時点が4の場合のP群とA群の差は, x_4 と x_8 の推定値の差で得られる.

群	症例	時点	前	群×時点	反応							
i	j	k	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	y
P	j'	1	0	1	0	0	0	0	0	0	0	$y_{1/0}$
		1	0	1	0	0	0	0	0	0	0	$y_{1/1}$
		2	0	0	1	0	0	0	0	0	0	$y_{1/2}$
		3	0	0	0	1	0	0	0	0	0	$y_{1/3}$
A	j'	4	0	0	0	0	1	0	0	0	0	$y_{1/4}$
		1	0	0	0	0	0	0	0	0	0	$y_{2/0}$
		1	0	0	0	0	0	1	0	0	0	$y_{2/1}$
		2	0	0	0	0	0	0	1	0	0	$y_{2/2}$
A	j'	3	0	0	0	0	0	0	1	0	0	$y_{2/3}$
		3	0	0	0	0	0	0	0	1	0	$y_{2/4}$
		4	0	0	0	0	0	0	0	0	1	$y_{2/5}$
		4	0	0	0	0	0	0	0	0	1	$y_{2/6}$

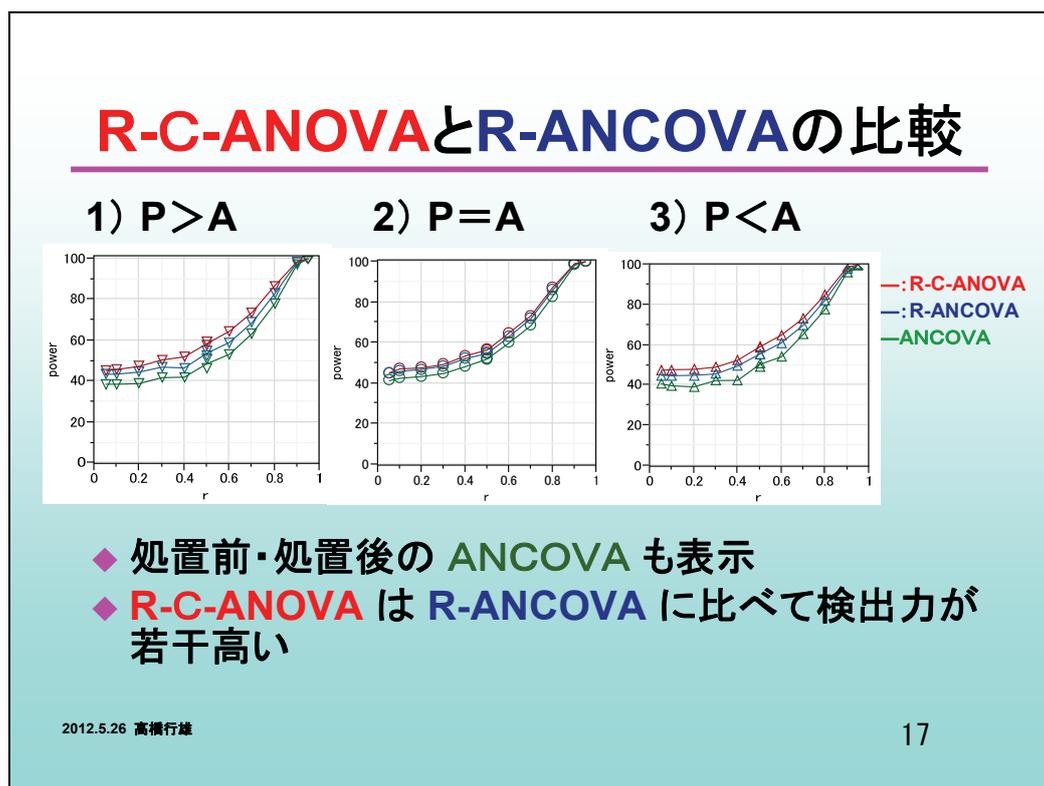
2012.5.26 高橋行雄 15

時点数が5のR-C-ANOVA

相関	0.05	0.10	0.30	0.50	0.70	0.90	0.95
1:P>A	45.3	45.7	50.5	59.3	73.7	98.5	100.0
2:P=A	45.1	47.0	49.2	56.6	73.3	99.1	100.0
3:P<A	47.4	47.7	49.0	59.4	73.4	98.9	100.0

- ◆ 処置前値の3区分分けの影響を全く受けていない.
- ◆ R-ANCOVAに比べても検出力は若干高い.

2012.5.26 高橋行雄 16



5. 時点数を変えた場合

R-ANCOVA は、処置後の時点数が 2 以上から適用可能であり、表 11 に示すように時点数が 3 の場合に、1: $P > A$, 3: $P < A$ 区分での検出力が 52.9%, 50.4%と $P = A$ 区分の 55.1%に比べて低くなる。

それに対し、R-C-ANOVA は、処置後の時点数が 1 の場合から適用可能であり、1: $P > A$, 3: $P < A$ 区分での検出力が 56.8%, 56.5%と、2: $P = A$ 区分の 56.5%に比べて低くなることなく、同程度の検出力を保ち、R-ANCOVA の処置後の時点数が 4 の場合の 55%前後の検出力に比べても検出力は 56%以上と 1~2%程度大きかった。

表 11 相関が 0.5 の場合の R-ANCOVA と R-C-ANOVA 検出力の比較

		時点数 (処置後の時点数)			
区分		2 (1)	3 (2)	4 (3)	5 (4)
R-ANCOVA	1: $P > A$	-	52.9	52.5	54.8
	2: $P = A$	-	55.1	55.2	55.0
	3: $P < A$	-	50.4	55.5	55.6
R-C-ANOVA	1: $P > A$	56.8	58.9	58.0	59.3
	2: $P = A$	56.5	57.3	57.1	56.6
	3: $P < A$	56.5	54.9	59.9	59.4

時点数を変えた場合

相関が 0.5 の場合の R-ANCOA と R-C-ANOVA 検出力の比較

		時点数 (処置後の時点数)			
区分		2(1)	3(2)	4(3)	5(4)
R-ANCOVA	1:P>A	-	52.9	52.5	54.8
	2:P=A	-	55.1	55.2	55.0
	3:P<A	-	50.4	55.5	55.6
R-C-ANOVA	1:P>A	56.8	58.9	58.0	59.3
	2:P=A	56.5	57.3	57.1	56.6
	3:P<A	56.5	54.9	59.9	59.4

- ◆ **R-ANCOVA**は、時点数が3の場合に、1:P>A, 3:P<A区分での検出力が52.9%, 50.4%とP=A区分の55.1%に比べて低くなる。
- ◆ **R-C-ANOVA**の時点数が2の場合、1:P>A, 3:P<A区分での検出力が56.8%, 56.5%と、2:P=A区分の56.5%と同程度である。

2012.5.26 高橋行雄

18

表 6 に示した ANCOVA の 3 区分の時点間の相関が 0.5 の場合の検出力は、それぞれ、46.4%, 52.4%, 49.6%であるので、R-C-ANOVA の検出力は、56.8-46.4=10.4%, 56.5-52.4=4.1%, 56.5-49.6=6.9%と高かった。

処置前, 処置後の 2 時点の場合でも R-C-ANOVA は ANCOVA に比べても平均への回帰現象に影響を受けにくいことは推測された。しかし、3 区分化した場合のシミュレーション回数は、10,000 組では不十分であるので、100,000 組と 10 倍を増やして ANCONA と R-C-ANOVA についての結果を表 12 に示すが、同様の結果であった。

表 12 相関が 0.5 の場合の 2 時点の場合の検出力の比較

	ANCOVA			R-C-ANOVA		
	p<0.05	n	検出力	p<0.05	n	検出力
1:P>A	9642	/19943	48.3	11387	/19943	57.1
2:P=A	31703	/60343	52.5	34176	/60343	56.6
3:P<A	9523	/19714	48.3	10987	/19714	55.7
All	50868	/100000	50.9	56550	/100000	56.6

処置前, 処置後の2時点の場合

相関が 0.5 の場合の 2 時点の場合の検出力の比較

	ANCOVA			R-C-ANOVA		
	$p < 0.05$	n	検出力	$p < 0.05$	n	検出力
1:P>A	9642	/19943	48.3	11387	/19943	57.1
2:P=A	31703	/60343	52.5	34176	/60343	56.6
3:P<A	9523	/19714	48.3	10987	/19714	55.7
All	50868	/100000	50.9	56550	/100000	56.6

- ◆ シミュレーション回数を 100,000組 にして, 2 時点の ANCONA と R-C-ANOVA を比較した.
- ◆ 3区分それぞれ $57.1-48.3=8.8\%$, $56.6-52.5=4.1\%$, $55.7-48.3=7.4\%$ と R-C-ANOVA の検出力が高かった.

2012.5.26 高橋行雄

19

結果の要約

- ◆ 測定されたすべての時点を用いた R-ANCOVA および R-C-ANOVA は, ANCOVA よりも平均への回帰現象の影響を減弱させる.
- ◆ R-ANCOVA では, 時点数が増えるに従い検出力の向上があり, 処置後3時点以上あれば影響を受けにくくなる.
- ◆ R-C-ANOVA も時点数が増えるに従い, 検出力の向上があるだけでなく, 時点数が2の場合から平均への回帰現象の影響を受けず, R-ANCONA の処置後4時点の場合よりも高い検出力となる.

2012.5.26 高橋行雄

20

6. 考察

多くの研究分野で得られる経時データに対して, 様々な統計手法が適用されている. 分割実験に対する分散分析の考え方をを用いた R-ANOVA は, 群と時間の交互作用に有意な差があれば, 時点ごとに多重比較法により輪切りの群間の平均値を比較する方法が多くの統計ソフトに実装されるようになってきた. この方法は, R-ANOVA を予備検定とし, その後の時点ごとの輪切りの ANOVA を正当化するための飾りのごとである.

Fleiss[文献・付録 6]は, 経時データに対して, R-ANOVA が, 統計学者たちの長年の批判にもかかわらず現実に使われていると批判している. また, Diggle ら[文献・付録 7]も経時データの解析に R-ANOVA を使うべきでないと述べている.

最もシンプルな経時データは, 処置前と処置後の 2 時点のみの場合である. この場合は, 処置前値を共変量にする ANCOVA が推奨されてはいるが, 処置前値にも測定誤差が含まれていることから ANOVA を無条件で推奨しがたいとの批判もある. そのために, 「私の実験データは, 時点間の相関が高いので, わずかな検出力の向上よりも, DIFF についての簡単な対応のない t 検定でも十分である」との反論を受け, 対応に苦慮する.

DIFF および POST は, 処置前値の群間差に関連して, 第 1 種の過誤および検出力が時点間の相関の大きさとともに揺れ動くことを示した. それに対して, ANCOVA は第 1 種の過誤, および, 検出力共に頑健であることを示した. ただし, 平均への回帰現象の影響が残り, 処置前値の群間差が広がるにつれて検出力の低下が起きる.

測定されたすべての時点を用いた R-ANCOVA および R-C-ANOVA は, ANCOVA よりも平均への回帰現象の影響を減弱させることができた. R-ANCOVA では, 時点数が増えるに従い検出力の向上があり, 処置後 3 時点以上あれば影響を受けにくくなった.

R-C-ANOVA も時点数が増えるに従い検出力の向上があるだけでなく, 時点数が 2 の場合から平均への回帰現象の影響を受けず, R-ANCONA の 5 時点の場合よりも高い検出力となった.

Frison ら[文献・付録 8]は, 経時データの解析で, 処置後の 1 時点の ANCOVA よりも複数時点の平均値を要約統計量とした解析を推奨した. しかしながら, 平均への回帰現象の影響は, 複数時点の平均値が処置後の 1 時点とみなされるので, ANCOVA と同様に平均への回帰現象が残ると思われる. また, 実験結果を示す統計量としては, 時点が特定できないことによる結果の提示に際して歯切れの悪さが残る.

要約統計量によるANCOVA

- ◆ Frisonらは, 経時データの解析で, 処置後の 1 時点の **ANCOVA** よりも複数時点の平均値を要約統計量とした解析を推奨した.
- ◆ しかしながら, 平均への回帰現象の影響は, 複数時点の平均値が処置後の 1 時点とみなされるので, **ANCOVA** と同様に平均への回帰現象の影響が残ると思われる.
- ◆ また, 実験結果を示す統計量としては, 時点が特定できないことによる結果の提示に際して歯切れの悪さが残る.

2012.5.26 高橋行雄

21

これらの結果から, 経時データの解析において処置後の特定の時点に対する ANCOVA に変えて, 測定時点を全て用いた R-ANCOVA あるいは R-C-ANOVA を使うことを推奨する.

Recommendation

以上の結果から, 経時データの解析において処置後の特定の時点に対する **ANCOVA** に変えて, 全ての測定時点をを用いた **R-ANCOVA** あるいは **R-C-ANOVA** を使うことを強く推奨する.
反復測定 ANOVA (**R-ANOVA**) の使用は全く推奨できない.

2012.5.26 高橋行雄

22

文 献

- [1] Snedecor.G.W.,Cochran.W.G.(1967), 畑村, 奥野, 津村共訳(1972), 統計的方法, 原著第6版, 岩波書店.
- [2] 岩崎学(2006), 単回帰分析, 東京図書.
- [3] 高橋行雄(2010), 経時データに対する投与前値を考慮した解析モデルの比較検討, SAS ユーザ会論文集.
- [4] Liang K.,Zeger S.L.(2000), Longitudinal Data Analysis of Continuous and Discrete Responses for Pre-Post Designs. The Indian J. of Statistics,62:134-148.
- [5] 高橋行雄, 渡辺琢也(2009), 投与前値を群間で共通とする経時データの解析, 統計関連学会連合大会抄録集.
- [6] Fleiss.J.L.,(1986), KR 研究会訳(2004), 臨床試験のデザインと解析, アーム.
- [7] Diggle.P.J.,Heagerty.P.,Liang.K-Y.,Zeger.S.L.,(2003), Analysis of Longitudinal Data 2ed., Oxford University Press.
- [8] Frison, S. and Pocock, S.J. (1992), Repeated Measures in Clinical Trials: Analysis Using Mean Summary Statistics and its Implications for Design. Statistics in Medicine, 11, 1685-1704.

連絡先 : 高橋 行雄, 〒105-0014 東京都港区芝 1-12-3, takahashi.stat@nifty.com , FAX : 03-342-8035

S A S プログラム

```

Title '平均への回帰 2012.03.28 高橋行雄';
Title2 '/*seed=20120331 n=1000 m1=10 m2=10 v=144*/';
Title3 '/* y10=100 y14=90 y20=100; y24=80; ';

data d01;
  array g_name[2] $4 g1 g2 ('G1:P' 'G2:A');
  array y_gt[2,5] y10 y11 y12 y13 y14
                y20 y21 y22 y23 y24;
  array g_time[2,5] $6 g10 g11 g12 g13 g14
                    g20 g21 g22 g23 g24
                ('G0' 'G1P_T1' 'G1P_T2' 'G1P_T3' 'G1P_T4'
                 'G0' 'G2A_T1' 'G2A_T2' 'G2A_T3' 'G2A_T4');
  array m[2] m1 m2;
  retain y0;

  seed=20120331;
  V=144; /* 全体の分散 */
  n=10000;
  m1=10; m2=10;
  y10=100; y11=97.5; y12=95; y13=92.5; y14=90;
  y20=100; y21=95; y22=90; y23=85; y24=80;

  do r = 0.05, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.95;
  sd1=sqrt(V*r); sd2=sqrt(V-sd1**2);
                /* 個体間 sd1, 個体内 sd2 */
  do i=1 to n;
    do l=1,2;
      g=g_name(l);
      do j=1 to m(l);
        sub=l*1000+j;
        e1=rannor(seed)*sd1;
        do t=0, 1, 2, 3, 4;
          e2=rannor(seed)*sd2;
          y=y_gt(l,t+1) + e1 + e2;
          gt=g_time(l,t+1);
          if t=0 then y0=y;
          if t>=1 then output;
        end;
      end;
    end; end; end; end;
  keep r i g sub t y0 y gt;

proc print data=d01;
  var r i g sub t y0 y gt;
  where i<=1 and sub<= 1002;
run;

proc tabulate data=d01 noseps;
  class r g t;
  var y y0;
  table r*g, n*f=7. (y0 y*t)*(mean std)*f=5.1 / rts=20;
run;

ods listing close;
proc means data=d01;
  class g;
  var y0;
  output out=m01;
  by r i t;
  where t=4;
run;

run;
ods listing;

proc print data=m01;
  where i<=1;
run;

data m01;
  set m01;
  if (_type_=1 and _stat_='MEAN') then output;
  keep r i t g y0;

proc print data=m01;
  where i<=2;
run;

title4 '/* 投与前の群平均の計算 */';
data m01;
  set m01;
  retain y0g1 y0g2;
  if (g='G1:P') then y0g1=y0;
  if (g='G2:A') then do; y0g2=y0; output; end;
  keep r i t y0g1 y0g2;

proc print data=m01;
  where i<=2;
run;

title4 '/* 前値の群間の y0 について平均値の比較 */';
ods listing close;
proc glm data=d01 outstat=s01 noprint;
  class g;
  model y0 = g / ss3;
  by r i t;
  where t=4;
run;
ods listing;

proc print data=s01;
  where i<=2;
run;

data s01pre;
  set s01;
  if prob<0.05 then pre05=1; else pre05=0;
  pre_t2 = sqrt(F);
  postp= prob;
  if (_SOURCE_ = 'g') then output;
  keep r i t pre_t2 prob pre05;
run;

proc print data=s01pre;
  where i<=2;
run;

proc tabulate data=s01pre noseps;
  class r t pre05;
  table r*t, (pre05 all)*f=10.;
run;

```

```

Title4 ' /* t0の群平均とp値との併合 */ ;
data t0mean ;
  retain t0p t0g ;
  merge m01 s01pre ;
  t0p=PROB ;
  t0g='2:P=A' ;
  pre_t=pre_t2 ;
  if ( y0g1>y0g2 and t0p<0.40) then t0g='1:P>A' ;
  if ( y0g1<y0g2 and t0p<0.40) then t0g='3:P<A' ;
  if ( y0g1>y0g2 ) then pre_t=pre_t2*(-1) ;
  by r i t ;
  keep r i t y0g1 y0g2 t0p t0g pre_t ;
run ;

proc print data=t0mean ;
  var r i t y0g1 y0g2 t0p t0g pre_t ;
  where i<=12 ;
run ;

proc tabulate data=t0mean noseps ;
  class r t0g ;
  var y0g1 y0g2 ;
  table r*(t0g all) t0g*all all , n*f=7.
         ((y0g1 y0g2)*mean (y0g1 y0g2)*std)*f=7.2/rts=20;
run ;

Title4 ' /* ANCOVA 投与前y0を共変量, */ ;
ods listing close ;
proc glm data=d01 outstat=s01 noprint ;
  class g ;
  model y=y0 g / ss3 ;
  by r i t ;
  where t=4 ;
run ;

ods listing ;
data s01ancova ;
  set s01 ;
  if prob<0.05 then ancova05=1 ; else ancova05=0 ;
  if prob=. then ancova05=0 ;
  ancovap=prob ;
  if (_SOURCE_ = 'g') then output ;
run ;

proc print data=s01ancova ;
  where i<=2 ;
run ;

proc tabulate data=s01ancova ;
  class r t ancova05 ;
  table r*t,(ancova05 all)*f=10. ;
run ;

Title4 ' ANCOVA ' ;
data s01pdac ;
  merge t0mean s01pre s01ancova ;
  by r i t ;
  keep r i t y0g1 y0g2 t0p t0g pre_t ancova05 ancovap ;

proc print data=s01pdac ;

```

```

var r i t y0g1 y0g2 t0p t0g pre_t ancova05 ancovap ;
where i<=12 ;
run ;

proc tabulate data=s01pdac noseps ;
  class r t0g ancova05 / missing ;
  var y0g1 y0g2 ;
  table r*t*(t0g all), (y0g1 y0g2 )*mean*f=7.1
         (ancova05 all)*f=7. / rts=25 ;
run ;

/* ===== MIXED =====*/
Title4 ' 4時点を込みにした実データ共分散分析' ;
/* ancrepy */
ods listing close ;
proc mixed data=d01 ;
  class g t sub ;
  model y=y0 g t g*t / DDFM=SATTERTH ;
  lsmeans g*t / pdiff ;
  random sub / ;
  make 'diffs' out=s01 ;
  by r i ;
run ;

ods listing ;
proc print data=s01 ;
  where i<=3 ;
run ;

data s01ancrepy ;
  set s01 ;
  if prob<0.05 then ancrepy05=1 ; else ancrepy05=0 ;
  if prob=. then ancrepy05=0 ;
  ancrepyp = probt ;
  if g='G1:P' and t=_t then output ;
  where t=4 ;

proc print data=s01ancrepy ;
  where i<=3 ;
run ;

proc tabulate data=s01ancrepy ;
  class r t ancrepy05 ;
  table r*t,(ancrepy05 all)*f=10. / rts=14 ;
run ;

/* ancova12 common pre */
data d02 ;
  set d01 ;
  t=0 ; y = y0 ; gt='G0' ;
  where t=4 ;
  keep r i g sub t y gt ;

proc print data=d02 ;
  where (i<=1 and sub<= 1002) ;
run ;

data d01 ;
  merge d02 d01 ;

```

```

by r i sub t ;
keep r i g sub t y gt ;

proc print data=d01 ;
  where (i<=1 and sub<= 1002) ;
run ;

Title4 ' 投与前を2群で共通にした共分散分析' ;
/* common pre */
ods listing close ;
proc mixed data=d01 noinfo NOPROFILE noprofile
NOITPRINT ;
* where i<=3 ;
class gt sub ;
model y=gt / noint DDFM=SATTERTH ;
lsmeans gt / pdiff ;
random sub ;
make 'diffs' out=s01 ;
by r i ;
run ;

ods listing ;
proc print data=s01 ;
  where i<=1 ;
run ;

data s01common ;
set s01 ;
group = input(substr( gt,2,1),1.) ;
_group = input(substr(_gt,2,1),1.) ;
t = input(substr( gt,6,1),1.) ;
_t = input(substr(_gt,6,1),1.) ;
if probt<0.05 then common05=1 ; else common05=0 ;
if probt=. then common05=0 ;
commonp = probt ;
if group=1 and t = _t then output ;

data s01common ;
set s01common ;
keep r i t commonp common05 ;
where t=4 ;

proc print data=s01common ;
  where i<=2 ;
run ;

proc tabulate data=s01common ;
class r t common05 ;
table r*t,(common05 all)*f=10. / rts=14 ;
run ;

/* ancova repeat common */
data s01arc ;
merge t0mean s01ancova s01ancrep
s01common ;
by r i t ;
keep r i t y0g1 y0g2 pre_t t0p t0g
ancova05 ancovap

```

```

ancrep05 ancrepyp common05 commonp ;

proc print data=s01arc ;
  where i<=12 ;
run ;

proc tabulate data=s01arc noseps ;
class r t t0g ancova05 ancrep05 common05/missing ;
table r*t*(t0g all),
(ancova05 ancrep05 common05 all)*f=5. /rts=15 ;
run ;

data arc_power ;
set s01arc ;
arc05_nm = '1:ancova05' ;
arc05 = ancova05 ; output ;
arc05_nm = '2:ancrep05' ;
arc05 = ancrep05 ; output ;
arc05_nm = '3:common05' ;
arc05 = common05 ; output ;
keep r i t0g arc05_nm arc05 ;

proc freq data=arc_power noprint ;
table r*t0g*arc05_nm*arc05 / list out=arc_power ;
run ;

proc transpose data=arc_power out=arc_power ;
var count ;
id arc05 ;
by r t0g arc05_nm ;
run ;

data arc_power ;
set arc_power ;
n0 = _0 ; if n0=. then n0 = 0 ;
n1 = _1 ; if n1=. then n1 = 0 ;
n = n0 + n1 ;
power = (n1/n) * 100 ;
keep r t0g arc05_nm n0 n1 n power ;

proc tabulate data=arc_power noseps ;
class r t0g arc05_nm / missing ;
var n1 n power ;
table arc05_nm*r, (t0g=' ')* ((n1 n)*
sum=' *f=5. power*mean=' *f=5.1)
all* ((n1 n)*sum=' *f=5.) / rts=25 ;
table (t0g=' ')* ((n1 n)*sum=' *f=7.
power*mean=' *f=7.1)
all* ((n1 n)*sum=' *f=7.)
,arc05_nm*r / rts=25 ;
run ;

data dd.arc_power ;
set arc_power ;
run ;

data dd.s01arc ;
set s01arc ;
run ;

```

[付録 1]

[1]

Snedecor.G.W.,Cochran.W.G.(1967), 畑村, 奥野, 津村共訳(1972), 統計的方法, 原著第6版, 岩波書店.
処置前値を共変量とした共分散分析(ANCOVA)の適用が半世紀前から推奨されている.

第14章 共分散分析

14.1 まえがき 共分散分析は分散分析と回帰の特徴をあわせた技術である. 1元分類の場合, i 番目の級の j 番目観測値 Y_{ij} に対する典型的な分散分析モデルは

$$Y_{ij} = \mu_i + e_{ij}$$

である, ただし μ_i はその級の母平均をあらわし, e_{ij} は残差である. しかし各単位について他の変数 X_{ij} をも測定しており, これは Y_{ij} と線型の関係があると仮定しよう. この場合にモデルを

$$Y_{ij} = \mu_i + \beta(X_{ij} - \bar{X}_{..}) + \varepsilon_{ij}$$

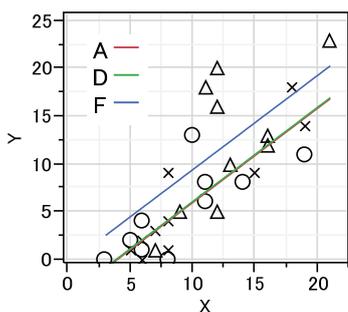
とするのが自然である, ただし β は Y の X に対する回帰係数である. これが共分散分析の典型的なモデルである. もし X と Y との関係が密接であれば, このモデルは元々の分散分析モデルよりよく Y_{ij} 値にあてはまると期待できよう. すなわち, 残差 ε_{ij} は一般に e_{ij} より小さくなるであろう. (略)

14.2 完全無作為化実験における共分散

表 14.2.1 ライ菌の点数, 処理前(X)と処理後(Y)

	薬					
	A		D		F	
	X	Y	X	Y	X	Y
	11	6	6	0	16	13
	8	0	6	2	13	10
	5	2	7	3	11	18
	14	8	8	1	9	5
	19	11	18	18	21	23
	6	4	8	4	16	12
	10	13	19	14	12	5
	6	1	8	9	12	16
	11	8	5	1	7	1
	3	0	15	9	12	20
計	93	53	100	61	129	123

JMPIによる追加解析(高橋)+文献での結果



分散分析				
要因	自由度	平方和	平均平方	F値
モデル	3	871.4974	290.499	18.1039
誤差	26	417.2026	16.046	p値(Prob>F)
全体(修正済み)	29	1288.7000		<.0001*

パラメータ推定値				
項	推定値	標準誤差	t値	p値(Prob> t)
切片	-2.695773	1.911085	-1.41	0.1702
薬[A]	-1.185037	1.060822	-1.12	0.2742
薬[D]	-1.076065	1.041298	-1.03	0.3109
X	0.9871838	0.164498	6.00	<.0001*

$$\begin{aligned}
 A: \bar{Y}_1 - b(\bar{X}_1 - \bar{X}_{..}) &= 5.3 - (0.988)(9.3 - 10.73) = 6.71 \\
 D: \bar{Y}_2 - b(\bar{X}_2 - \bar{X}_{..}) &= 6.1 - (0.988)(10.0 - 10.73) = 6.82 \\
 F: \bar{Y}_3 - b(\bar{X}_3 - \bar{X}_{..}) &= 12.3 - (0.988)(12.9 - 10.73) = 10.16
 \end{aligned}$$

最小2乗平均表			
水準	最小2乗平均	標準誤差	平均
A	6.714963	1.2884943	5.3000
D	6.823935	1.2724690	6.1000
F	10.161102	1.3159234	12.3000

14.2 節 最後のパラグラフ

(略)

この実験のように, X と Y とが同種の量(ライ菌に対する点数)の測定である場合に, 共分散に代わる方法として $(Y-X)$, 点数の《変化》, を処理効果の測度として用いることがある. $(Y-X)$ の誤差平均平方は表 14.2.1 から

$$\frac{E_{yy} - 2E_{xy} + E_{xx}}{2} = \frac{[995.1 - 2(585.4) + 592.9]}{2} = 15.45$$

として求まる. これは共分散の 17.03 と比較さるべきものである. この実験では, $(Y-X)$ の使用の方が共分散より, 計算が簡単なばかりか, 僅かに有効となっている. このデータを選び出した元のより大きな実験では, これは推奨さるべき変数である. しかし多くの実験では $(Y-X)$ は共分散より劣り, X, Y 間の相関が低いときには Y よりも劣ることもある.

14.4 2元分類による共分散

【要約】 24人の被験者にたいして, (モルヒネ, ヘロイン, プラセーボ)の3薬剤を皮下注射し, 精神活動を投与前 X 値と2時間後 Y に測定した. 実験は8人ごとのラテン方格で割り付けた.

共分散分析で傾きを求めると, $b = -16/199$ と負になるので, 2時間後 Y のみを用いた解析が最もよい. ライ麦の例とは異なり, $(X-Y)$ を用いることは, 誤差平方和より大きくなるので, 薬の効果の測度として用いるのは賢明でない.

【要約】 トウモロコシ6品種の収量 Y を比較した. それぞれ4ブロックに割り付けた. 共変量は, ブロックごとのトウモロコシの本数である. 注)これは, 処置前のデータではない.

14.6 回帰直線の比較

【要約】 異なる環境で実験された, 年齢 X とコレステロール Y の関係に対し, 結果を要約する場合は, それぞれの研究の回帰曲線が同一か検討する. 高橋の注)これも, 処置前のデータではない

14.8 重共分散

【要約】 40頭のブタのを10頭ごとに4種の処理をし, 初期の年齢 X_1 , 初期の体重 X_2 , 一日当りの体重増加率 Y を得た. 前値を共変量とした解析と同様の事例である. 高橋の注)前値を共変量とすることの意義については記述が見出せない.

14.9 2元表における重共分散

【要約】 小麦の収量 Y とし, 小麦の出穂期の草丈 X_1 , 分けつ期の個体数 X_2 を, 6ヶ所の試験場で3年間間した実験結果が示されている. 高橋の注)これも, 処置前のデータではない.

[付録 2]

[2]

岩崎学(2006), 単回帰分析, 東京図書.

このバイアスは, 処置前値を X , 処置後値を Y , X と Y の分散は等しいとした場合の相関を ρ , 処置前値の群間差を δ としたときに, DIFF の群間差のバイアスは, $\delta(1-\rho)$ となる. ANCOVA は, このようなバイアスを除去するための標準的な方法として用いられている.

第3章 平均への回帰

3.1 平均への回帰とは

説明変数 x と目的変数 y が同じ種類の測定値であるとする. ある処置の効果を調べるため, 同一個体に対してその処置を施す前と後で測定するいわゆる処置前後データ (pretest-posttest data) の場合や, 1-5 節で扱った両親と子供の身長がその例である. このような場合には平均への回帰 (regression to the mean, regression towards the mean) とよばれる現象に注意を払わなくてはならない. (略)

3.2 平均への回帰の源泉

平均への回帰の生じる理由をモデルにより考察する. 2 回の観測値を表わす確率変数を X および Y とし,

$$\begin{aligned} \mu_X &= E[X], \quad \mu_Y = E[Y], \quad \sigma_X^2 = V[X], \quad \sigma_Y^2 = V[Y], \\ \sigma_{XY} &= \text{Cov}[X, Y], \quad \rho_{XY} = R[X, Y] = \sigma_{XY} / (\sigma_X \sigma_Y) \end{aligned}$$

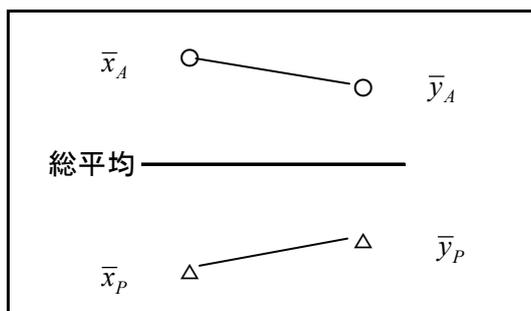
とする. このとき,

$$E[Y | X = x] = \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2}(x - \mu_X), \quad V[Y | X = x] = \sigma_Y^2 + \frac{\sigma_{XY}^2}{\sigma_X^2} \quad (1)$$

であり, $\frac{\sigma_{XY}}{\sigma_X^2}(x - \mu_X)$ が平均への回帰分である. (略)

(高橋のコメント)

処置前 X の期待値を μ_X とし, A 群の処置前の平均値を \bar{x}_A , P 群は \bar{x}_P とする. 処置後 Y の期待値を μ_Y とし, A 群の処置後の平均値を \bar{y}_A , P 群は \bar{y}_P とする.



処置後 Y の分散が, 処置前 X の分散と等しいと仮定すると, これらの相関は,

$$\rho_{XY} = \sigma_{XY} / (\sigma_X \sigma_Y) = \sigma_{XY} / \sigma_X^2$$

となる. 式(1)から,

$$E[Y | X = \bar{x}_A] = \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2} (\bar{x}_A - \mu_x) = \mu_Y + \rho_{XY} (\bar{x}_A - \mu_x) \quad (2)$$

$$E[Y | X = \bar{x}_P] = \mu_Y + \frac{\sigma_{XY}}{\sigma_X^2} (\bar{x}_P - \mu_x) = \mu_Y + \rho_{XY} (\bar{x}_P - \mu_x) \quad (3)$$

処置後の A 群と P 群の平均値の期待値の差は,

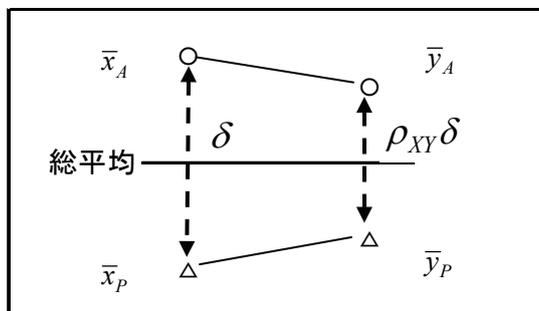
$$\begin{aligned} & E[Y | X = \bar{x}_A] - E[Y | X = \bar{x}_P] \\ &= (\mu_Y + \rho_{XY} (\bar{x}_A - \mu_x)) - (\mu_Y + \rho_{XY} (\bar{x}_P - \mu_x)) \\ &= \rho_{XY} (\bar{x}_A - \bar{x}_P) \end{aligned}$$

となる. 処置前の A 群と P 群の平均値の差を $\delta = \bar{x}_A - \bar{x}_P$ とおくと, 処置後の A 群と P 群の平均値の期待値の差は, $\rho_{XY} \delta$ となり, 処置前後の相関が低下するにつれて縮小する. 相関が 0.7 の場合は, 0.7δ となり, 相関が 0.4 の場合は 0.4δ と縮小する.

処置前値の平均値の群間差 $\delta = \bar{x}_A - \bar{x}_P$ があつた場合であっても, A 群と P 群の処置前後の差の平均値を比較する際に δ の影響を受けなくなるとされている. しかしながら, 平均への回帰現象により,

$$\text{処置前の群間差} - \text{処置後の群間差} = \delta - \rho_{XY} \delta = \delta(1 - \rho_{XY})$$

とバイアスが必ず生じてしまう.



(高橋のコメント・終り)

[付録 3]

[3]

高橋行雄 (2010), 経時データに対する投与前値を考慮した解析モデルの比較検討, SAS ユーザ会論文集.

処置後のある時点のみを用いた ANCOVA の群間比較に比べ, R-ANCOVA で同一時点での群間差の分散が小さくなり, 検出力が大きくなると期待される.

<p>経時データに対する投与前値を考慮した解析モデルの比較検討</p> <p>中外製薬 高橋 行雄</p> <p>2010.7.26 高橋行雄</p> <p>1</p>	<p>群間で投与前の平均値が異なる</p> <ul style="list-style-type: none"> ◆ 平行群間試験で数週間ごとに反復測定される臨床検査データを対象 ◆ 投与前の平均値が群間で均一となることは極めてまれ ◆ 群間で有意な差となる場合の扱い ◆ 投与後の経時データを群間で比較する際に, この差をどのように考慮するのか. <p>2010.7.26 高橋行雄</p> <p>2</p>
--	---

平均への回帰の例示

実データ

差のデータ

○: プラセボP
×: 実薬R

- ◆ 平均への回帰は投与前と時点1で観察される.
- ◆ 時点2では, 実データと差のデータで薬剤間の差の解釈が異なる.

2010.7.26 高橋行雄

アウトプット 4 の平均値をプロット

4

解析方法をめぐる衝突

Aさん: 単純な前後差 派

Bさん: ANCOVA 派 (共分散分析 派)

Cさん: (群と時点の) 交互作用 派

Dさん: 反復測定 ANCOVA 派

Eさん: CPMP派 (差のデータ反復測定 ANCOVA)

Fさん: 投与前値 共通 派

Gさん: SASプログラマー

2010.7.26 高橋行雄

10

実データで反復測定・共分散分析

- ◆ 実データで T1 と T2 を反復測定とした場合の共分散分析
- ◆ 交互作用 $g * t$ の検定結果には興味がない. 時点ごとの群間比較が出来るようするために入れてある.

```
proc mixed data=d01 ;  
  where (t=1 or t=2) ;  
  class g t sub ;  
  model y = y0 g t g*t / ddfm=satterth ;  
  lsmeans g*t / pdiff ;  
  random sub / ;  
  make 'diffs' out=s01 ;  
  by i ;  
run ;
```

2010.7.26 高橋行雄

15

Gさんは、MIXED プロシジャに詳しくないので Dさんにサンプルプログラムを示してもらったことにした。Dさんは、(スライド 15) MIXED プロシジャで、Random ステートメントで症例 sub を指定し、同じ症例の時点 T1 と T2 の対応付けを行い、lsmeans ステートメントの pdiff で時点ごとの群間比較を行い、この結果をファイルに出力し必要な情報を data ステップで取り出し方を示した。

アウトプット 5. 時点 $T1$ と $T2$ を込みにした実データの共分散分析 ($\rho=0.38$)

	ancropy05		All
	0	1	
t			
1	956	44	1000
2	500	500	1000

$T1$ 時点は α 危険率となっていて 44/1000 : 4.4%
 $T2$ 時点は検出力となっていて 500/1000 : 50.0%

投与前 T0 と 投与後 T2 の比較

2群の時期別の期待値 ($\rho=0.38$)

群	G1:P薬	G2:R薬	差
投与前T0	100.0	100.0	0.0
投与後T2	90.0	80.0	10.0

2群間の1000回中の有意差検定の結果 ($\rho=0.38$)

	NS	*	全体	検出力	
投与後T2	553	447	1000	44.7%	2位
前後差 T2-T0	698	368	1000	36.8%	3位
ANCOVA	533	467	1000	46.7%	1位

時点間の相関が 0.5 より小さければ, 実データでの比較が差のデータよりも検出力が高くなることは良く知られている. $n=1000$ 程度のシミュレーションでも簡単にできる.

2010.7.26 高橋行雄

12

反復測定 ANCOVA (R-ANCOVA) の検出力 50.0% に対して,
 ANCOVA の検出力は, 46.7% と小さい.

[付録 4]

[4]

Liang K., Zeger S.L.(2000), Longitudinal Data Analysis of Continuous and Discrete Responses for Pre-Post Designs. The Indian J. of Statistics, 62:134-148.

処置前値を各群で共通とする反復測定 ANOVA (R-C-ANOVA) が提唱されている。R-C-ANOVA は、処置前値も応答変数 Y に含めて解析することが、R-ANCOVA と異なる。

SUMMARY. The availability of biomarkers has led to the increasing adoption of the pre-and post-randomization designs in clinical trials. In this paper we discuss the use of random effects models when the primary objective of the trial is to assess the efficacy of newly developed treatments in terms of progression of biomarkers over time. One issue of particular interest is how to best utilize the pre-randomization responses. We discuss the pros and cons of several inferential procedures including the conditional and full likelihood approaches. Throughout, these issues are illustrated by an analysis of data from a schizophrenic trial by the Janssen Research Foundation.

(高橋のコメント)

一般的に Longitudinal Data Analysis に対する各種の書物では、処置後のデータを扱い、処置前値を含めないのが通例である。この論文は、処置前値を含めた解析法を論じている。この論文では、統合失調症を対象とした臨床試験データの一部を用いて、処置前値を含めた各種の解析法を論じている。

解析の結果として、1) 処置前値を群間で別々に扱うモデル、2) 処置前値を群間で共通と知るモデルが示されているが、唯一つの臨床試験の解析結果なので、この論文を基に、解析方法の性能を論ずることはできない。

(高橋のコメント終り)

[付録 5]

[5]

高橋行雄, 渡辺琢也(2009), 投与前値を群間で共通とする経時データの解析, 統計関連学会連合大会抄録集.

症例間の分散ならびに症例内の誤差分散の揺らぎが小さくなり, また誤差分散の自由度も大きくなることから, 処置後の群間差について R-C-ANOVA の検出力が R-ANCOVA よりも, 大きくなるのが期待される.

1. はじめに

平行群間試験で反復測定される臨床検査データにおいて, 投与前の平均値が群間で均一となることは極めてまれであり, 有意な差となってしまう場合もときには生ずる. 投与後の経時変化を検討する際に, 群間の投与前値の違いをどのように考慮したらよいのだろうか.

ランダム化が適切に行なわれた場合には, 投与前値が高ければ平均への回帰の現象により投与後は低めに, 低ければ高めになりがちである. 平均への回帰現象を考慮した場合には, 群間の投与前値は群全体で共通とすることも考えられる.

2. 解析方法

表1に経時測定データの一部を示す. 対照群の投与前値が 30 mg/kg に比べて低めになっていることも影響し 13 週後の赤血球数に有意な差が見出せない. 投与前値を共変量とした解析では, 平均への回帰を考慮しないために群間の差を過大に評価しがちになる.

Liang, Zegerr (2000) が提唱している投与前値を両群に共通とした反復測定分散分析を試み, 投与前値を共変量とした解析結果と比較検討する. 両者の統計的な性能評価を日本製薬工業協会 (JPMA) で集積された大動物の (15 施設 47 試験 1339 頭 42 変数) の反復毒性試験データを用いて行なったので, その結果を報告する.

表 1 赤血球数の経時変化 (CN=8, TN=5, VAR=H01 の一部)

番号	対照群 : 0 mg/kg			番号	中用量群 : 30 mg/kg		
	投与前	4 週後	13 週後		投与前	4 週後	13 週後
1	807	815	810	9	730	670	722
2	646	667	717	10	826	766	780
3	695	774	780	11	785	772	771
4	672	742	769	12	653	606	612

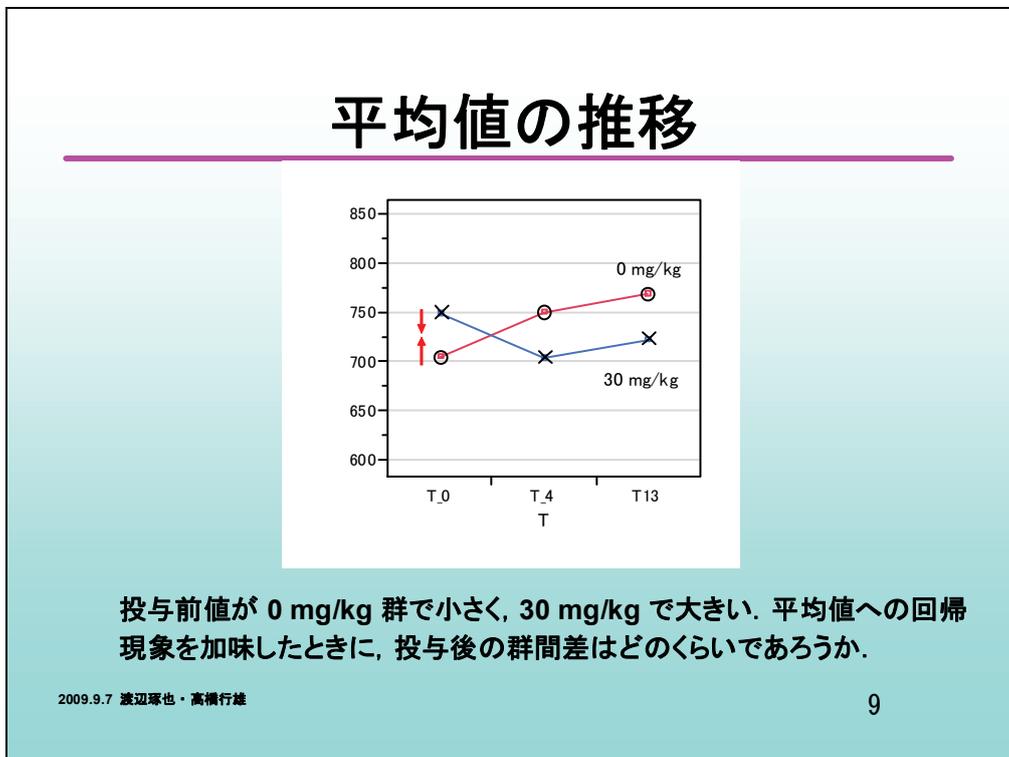
3. 結果

表 2 に示すように, 投与前を共変量とした反復測定分散分析では, 13 週目で p 値は 0.0095 であるが, 投与前値を共通とする反復測定分散分析では差が p 値は 0.0007 と小さくなった. 他のデータについても同様に p 値を求め, 両者の統計的な性能評価を行った結果を報告する.

表 2 平均値と 2 つの推定値の比較

	平均値		投与前値を共変量			投与前値を共通		
	0mg/kg	30mg/kg	0mg/kg	30mg/kg	差	0mg/kg	30mg/kg	差
投与前	705.0	748.5	726.75			726.75		
4 週後	749.5	703.5	762.98	681.45	-81.53	769.28	683.72	-85.56
13 週後	769.0	721.3	782.48	699.20	-83.28	788.78	701.47	-87.31
13 週後 SE					21.38			19.96
p 値					0.0095			0.0007

文献: Liang K., Zeger S.L.(2000), Longitudinal Data Analysis of Continuous and Discrete Responses for Pre-Post Designs. The Indian J. of Statistics, 62:134-148.



0mg/kg 群の投与前値が 30mg/kg 群よりも小さく, 平均への回帰現象により, 13 週目 (T13) の実データでの群間比較にバイアスの混入が避けられない.

投与前値を共通・解析用データ

投与量 mg/kg	No	時点 週	投与前	0mg/gk 群		30mg/kg 群		反応
			x0	x11	x12	x21	x22	v
0	1	0	1	0	0	0	0	807
	1	4	0	1	0	0	0	815
	1	13	0	0	1	0	0	810
	:	:	:	:	:	:	:	:
	4	0	1	0	0	0	0	672
	4	4	0	1	0	0	0	742
	4	13	0	0	1	0	0	769
30	5	0	1	0	0	0	0	730
	5	4	0	0	0	1	0	670
	5	13	0	0	0	0	1	722
	:	:	:	:	:	:	:	:
	8	0	1	0	0	0	0	653
	8	4	0	0	0	1	0	606
	8	13	0	0	0	0	1	612

2009.9.7 渡辺琢也・高橋行雄

10

投与前値を群間で共通とするためのデザイン行列である。両群の 0 週のデータに対してダミー変数が $x0=1$, それ以外は $x0=0$ としている。

解析モデル

$$y_{ijk} = \beta_0 x_{ij0} + \beta_{jk'} x_{ijk'} + \varepsilon_{ij} + \varepsilon_{ijk}$$

症例番号: $i = 1, 2, \dots, n_j$

薬剤: $j = 1, 2$

時期: $k = 0, 1, 2, \quad k' = 1, 2$

投与前: $x_{ij0} = 1, \text{ for } k = 0$

$x_{ij0} = 0, \text{ for } k = 1, 2$

薬剤: $d_1 = 1; \quad d_2 = 0, \text{ for } j = 1$

$d_1 = 0; \quad d_2 = 1, \text{ for } j = 2$

時期: $t_1 = 1; \quad t_2 = 0, \text{ for } k = 1$

$t_1 = 0; \quad t_2 = 1, \text{ for } k = 2$

薬剤×時期: $x_{ijk'} = d_j t_{k'}$

個人間誤差: $\varepsilon_{ij} \sim N(0, \sigma_{\text{between}}^2)$

個人内誤差: $\varepsilon_{ijk} \sim N(0, \sigma_{\text{within}}^2)$

2009.9.7 渡辺琢也・高橋行雄

11

線形混合モデルを適用するためのモデルを示す。

投与前値を共通とした場合の推定値

群	投与前	4 週	13 週
0 mg/kg	726.75	769.28	788.78
30 mg/kg		683.72	701.47
推定値の差	0	-85.56	-87.31

個体間分散 = 64.77^2 , 個体内分散 = 20.43^2

0 mg/kg と 30 mg/kg の 13 週目の群間の推定値の差は **-87.3** と投与前値からの差から求めた **-91.3** より小さめに推定されている。

2009.9.7 渡辺琢也・高橋行雄

12

群で共通の投与前値 726.75 は, 投与前値を共変量とした解析と同じになる。これは, 2 群の投与前値の算術平均に一致する。4 週目および 13 週目の推定値は, 共分散分析と若干異なる。

時期ごとの群間比較

◆ 投与前値を共変量 (ANCOVA)

比較	Estimate	Standard Error	DF	t Value	Pr > t
4週目 0mg/kg vs 30mg/kg	-81.9972	22.7513	6.6	-3.60	0.0096
13週目 0mg/kg vs 30mg/kg	<u>-83.7472</u>	22.7513	<u>6.6</u>	-3.68	<u>0.0087</u>

◆ 投与前値を共通とするモデル

比較	Estimate	Standard Error	DF	t Value	Pr > t
4週目 0mg/kg vs 30mg/kg	-85.5647	19.9601	13.1	-4.29	0.0009
13週目 0mg/kg vs 30mg/kg	<u>-87.3147</u>	19.9601	<u>13.1</u>	-4.37	<u>0.0007</u>

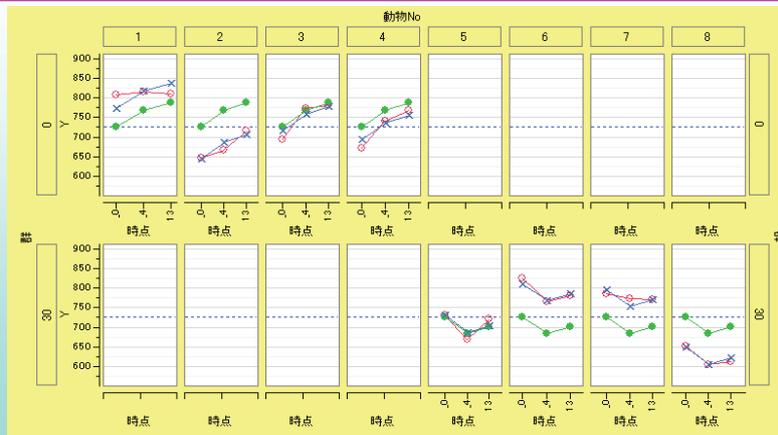
推定値の差は, この例では ANCOVA より大きくなっている。自由度も大きくなり, P 値は ANCOVA よりも小さくなっている。

2009.9.7 渡辺琢也・高橋行雄

13

4 週目および 13 週目の群間の差の推定量は, SAS/MIXED プロシジャの Estimate ステートメントで推定した結果である。

個体間の変動・個体内の変動



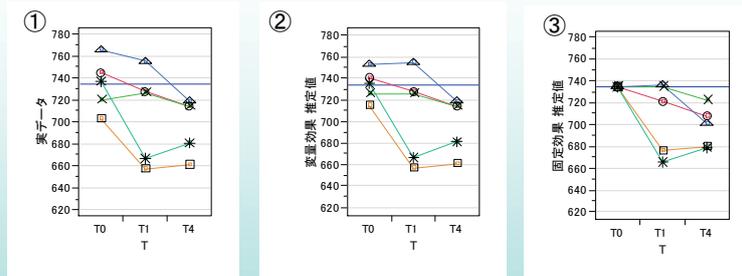
○: 実データ y_{ijk} , ●: 固定効果, ×: 固定効果+個体間変動

2009.9.7 渡辺琢也・高橋行雄

14

投与前値を共通とした推定値を ● で示してある. 群ごと, 症例ごとの経時推移 × は, これと同じ経時パターンとなり, 変量効果(個体間変動)を現している. 実データは ○ で示してあり, × から ○ までの距離が個体内の誤差なる.

雌についての5群比較



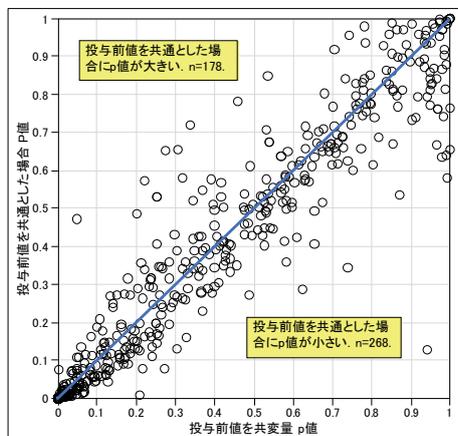
- ① 実データの群ごとの時期ごとの平均値 .
- ② 個体内変動を取り除いた推定値の平均値 .
- ③ 5群の平均値 734.0 へ平行移動 ,
- ③の推定値についての群間比較を行う.

2009.9.7 渡辺琢也・高橋行雄

17

5 群全体に対して, 投与前値を共通とする線形混合モデルでの推定値を線グラフで示した.

P値による性能比較



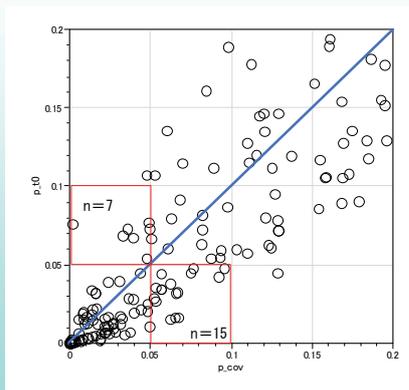
- ◆ ある反復投与毒性試験のすべての血液生化学検査・血液学検査データを対象.
- ◆ 時期ごとに0mg/kgと投与群間の差の検定のP値を求め散布図とした.
- ◆ 投与前値を共通とする解析でP値が小さくなる検定の数が多いことが示されている.

2009.9.7 渡辺琢也・高橋行雄

19

投与前値を共変量とする共分散分析との性能比較を行なった. 投与前を共変量とした場合に p 値が小さくなる回数が多いようであり, この解析モデルの優位性が示唆され, 検出力の比較など更なる検討をする価値が見出された.

P値の 0.0~0.2 の範囲での比較



- ◆ P値が<0.1から<0.05へ変化した検定結果
- ◆ 投与前値を共通とする解析:
n = 15
- ◆ 投与前値を共変量とする解析:
n = 7

2009.9.7 渡辺琢也・高橋行雄

20

P 値が 0.2 以下の場合について, 拡大して, 共分散分析で $0.05 < p < 0.1$ となるが, 投与前を共通とするモデルで $p < 0.05$ となる検定の回数は 15 件, 逆の場合は 7 件となり, 投与前を共通とするモデルの優位性が示唆された.

[付録 6]

[6]

Fleiss.J.L.,(1986), KR 研究会誌(2004), 臨床試験のデザインと解析, アーム.

Fleiss は, 経時データに対して, R-ANOVA が, 統計学者たちの長年の批判にもかかわらず現実に使われていると批判している.

第 8 章 反復測定試験

これまではただ 1 時点における処理の比較にしか注目してこなかった。しかし、処理に対する反応が投与中あるいは投与後に複数回測定されることがよくある。それは、反応の平均レベルだけでなく、処理効果の時間経過に伴う傾向についても処理間で比較を行うためである。長年にわたる批判 (Danford, Hughes, and McNee, 1960; Greenhouse and Geisser, 1959) にも関わらず、反復測定試験 (repeated measurements study) においてなされるような経時的な推移の解析は、分散分析を単純に適用することによって行われてきたが、それは一般に不適切なものである (Aitkin, 1981; Wallenstein, 1982; Yates, 1982)。8.1 節では、反復測定試験に対する分散分析表を示し、解析の基礎となる理論をまとめる。そして、時間要因についての種々の仮説を検定するために用いられる平均平方の比の分布について提案された幾つかの近似法を紹介する。反復測定試験のデータの多変量解析については 8.2 節に述べ、時間要因を含む多重比較の問題は 8.3 節で述べる。Friedman, Furberg and DeMets (1981, pp.83-85) Schlesselman (1973) は反復測定試験において時間要因の主たる対比が直線傾向 1 つのみである場合のサンプルサイズの設定法を示している。(略)

(高橋のコメント)

投与後に複数回測定 (経時データ) された実験データに対して、時間を分散分析の一つの因子として加えた分散分析は、同一症例内で互いにランダムに各時点のデータが得られたと仮定する「分割実験」タイプである。この解析モデルは、時点間の相関が 0 と仮定しているが、複数回測定した時点間のデータには、正の相関となることが指摘されていて、単純な「分割実験」タイプの分散分析を行うことは不適切である。この本には、時点間の相関を考慮し、要因の自由度を割り引く Huynh and Feldt などの近似法も示されていて、SAS の GLM プロシジャで計算することもできた。

フライスは、分散分析を単順に適用する代わりに、時点を多変量とみなした多変量解析 (多変量分散分析, MANOVA) を本書で推奨している。1980 年後半には SAS の GLM プロシジャで MANOVA の解析をすることは可能となっていたのであるが、全ての時点のデータが揃っている必要性があり、また、経時的なプロファイルの変化を直接的に捉えにくい、などの制約があり一般的に普及はしなかった。この当時は、まだ線形混合モデルは使える状況ではなかった。

(高橋のコメント終了)

[付録 7]

[7]

Diggle.P.J.,Heagerty.P.,Liang.K-Y.,Zeger.S.L.,(2003), Analysis of Longitudinal Data 2ed., Oxford University Press.

Diggle らも経時データの解析に R-ANOVA を使うべきでないと述べている.

6 Analysis of variance methods

6.1 Preliminaries

In this chapter we describe how simple methods for the analysis of data from designed experiments, in particular the analysis of variance (ANOVA) can be adapted to longitudinal studies. ANOVA has limitations which prevent its recommendation as a general approach for longitudinal data. (略)

6.4 Repeated measures

A repeated measures ANOVA can be regarded as a first attempt to provide a single analysis of a complete longitudinal data set. The rationale for the analysis is to regard time as a factor on n levels in a hierarchical design with units as sub-plots. In agricultural research, this type of experiment is usually called a split-plot experiment. However, the usual randomization justification for the split-plot analysis is not available because there is no sense in which the allocation of times to the n observations within each unit can be randomized. (略)

(高橋のコメント)

Diggle らは, 解析に分割実験 (split-plot) での仮定している時点についてのランダムマイゼーションは, 経時データの場合には満たされていないと指摘している.

分割実験モデルを用いた A repeated measures ANOVA は, 線形混合モデルによる解析が手軽にできるようになる以前の「主要」な解析方法であり, これに慣れ親しんだ人達(高橋も含めて)にとって「なぜいけないのか」との疑問を持つ. 多少の不具合があったとしても, 変わるべき方法がなければ, それを知りつつ結果の解釈を慎重にすることで対処するしかなかった.

実験データの統計解析を実践している人達にとっては, 自らが慣れ親しんだ統計の教科書を超える新たな解析法を再度学習することは難儀であるに違いない. 問題があるとの指摘があったとしても, それに変わる実用的な解析法が示されなければ対応できない.

(高橋のコメント終り)

[付録 8]

[8]

Frison, S. and Pocock, S.J. (1992), Repeated Measures in Clinical Trials: Analysis Using Mean Summary Statistics and its Implications for Design. *Statistics in Medicine*, 11, 1685-1704.

Frison らは、経時データの解析で、処置後の 1 時点の ANCOVA よりも複数時点の平均値を要約統計量とした解析を推奨した。

SUMMARY

This paper explores the use of simple summary statistics for analysing repeated measurements in randomized clinical trials with two treatments. Quite often the data for each patient may be effectively summarized by a pre-treatment mean and a post-treatment mean. Analysis of covariance is the method of choice and its superiority over analysis of post-treatment means or analysis of mean changes is quantified, as regards both reduced variance and avoidance of bias, using a simple model for the covariance structure between time points. Quantitative consideration is also given to practical issues in the design of repeated measures studies: the merits of having more than one pre-treatment measurement are demonstrated, and methods for determining sample sizes in repeated measures designs are provided. Several examples from clinical trials are presented, and broad practical recommendations are made. The examples support the value of the compound symmetry assumption as a realistic simplification in quantitative planning of repeated measures trials. The analysis using summary statistics makes no such assumption. However, allowance in design for alternative non-equal correlation structures can and should be made when necessary.

(高橋のコメント)

2 群の投与前値を含む経時データについて、投与後値のみを使う群間比較 POST, 投与前から投与後の差についての群間比較 CHANGE, 投与前値を共変量とした ANCOVA について、時点間の相関が変化した場合に、POST vs ANCOVA, CHANGE vs ANCOVA について比較し、ANCOVA が、時点間の相関にかかわらず常に群間比較の際の分散が小さくなるので、標準的に使うことを薦めている。

$$\text{POST:} \quad \text{variance} = \left(\frac{1}{n_A} + \frac{1}{n_B} \right) \frac{\sigma^2}{r} [1 + (r-1)\rho]$$

$$\text{CHANGE:} \quad \text{variance} = \left(\frac{1}{n_A} + \frac{1}{n_B} \right) \frac{\sigma^2}{r} [1 + (r-1)\rho + r(1-2\rho)]$$

$$\text{ANCOVA:} \quad \text{variance} \approx \left(\frac{1}{n_A} + \frac{1}{n_B} \right) \frac{\sigma^2}{r} [1 + (r-1)\rho - r\rho^2]$$

ANCOVA の分散は, POST または CHANGE の分散よりも常に小さくなる事が確認できる.

投与後の時点数 r が増えると分散が小さくなるだけでなく, CHANGE との比較で相対的に分散が小さくなる事がわかる. 相関が 0.7 の場合に $r = 1$ の場合に分散が相対的に 85%であるが, $r = 5$ の場合には, 75%とさらに減少している.

CHANGE vs ANCOVAの分散の大きさの比較					
r	ρ	POST	CHANGE	ANCOVA	CHANGE vs ANCOVA
1	0.001	0.800	1.598	0.800	0.50
1	0.3	0.800	1.120	0.728	0.65
1	0.5	0.800	0.800	0.600	0.75
1	0.7	0.800	0.480	0.408	0.85
1	0.999	0.800	0.002	0.002	1.00
1	0.7	0.800	0.480	0.408	0.85
2	0.7	0.680	0.360	0.288	0.80
3	0.7	0.640	0.320	0.248	0.78
4	0.7	0.620	0.300	0.228	0.76
5	0.7	0.608	0.288	0.216	0.75
$n_A = 10$		$\sigma^2 = 4$			
$n_B = 10$					

(高橋のコメント終り)