

# FDR を制御する多重比較法の性能評価

堀内 賢太郎 松田 眞一

最近の多重比較では FWER ではなく FDR を制御の対象とする方法の研究が盛んである。本論文では、FDR を制御する多重比較法について述べ、いくつかの方法に対して R 上でのプログラミングを行い、それらの性能について研究する。結果として、データが相関構造を持つ場合での BH 法の安定性が確認された。

## 1 はじめに

堀内 (2004) では、多重比較に関する第 1 種の過誤確率にあたる FWER (Familywise error rate) を制御する方法のうち、すべての対比較を同時に検定する Tukey の方法と対比較を行う際に step-down 法を用いる Tukey-Welsch の方法を R 上でプログラム化した。一方、最近の多重比較法について調査すると、FDR (False Discovery Rate) を制御する方法が多く見られる。これらの方法は、従来の FWER を制御する方法とは異なり、棄却された仮説のうち誤って棄却された真の帰無仮説の割合 (FDR) を制御しているところが大きな特徴である。この考え方は Benjamini and Hochberg (1995) によって提案されたものであり、それに付随して考え出された手法がいくつか存在している。

本論文では、その中から BH 法、adaptive BH 法、BY 法を取り上げ、R 上でプログラミングした上で、それらの性能のシミュレーション評価を行う。

## 2 FDR (False Discovery Rate) の定義

前もって与えられた  $m$  個の仮説を同時に検定する問題を考える。そのうち  $m_0$  個は未知である真の帰無仮説、 $m_1$  個は未知である偽の帰無仮説とする。 ( $m = m_0 + m_1$  となる。) また、 $R$  は検定によって棄却される仮説の数を表す確率変数とする。 (太字で確率変数を表す記号法は Benjamini and Hochberg (1995) の考え方に従っている。) 表 1 は、状況を伝統的な形式にまとめたものである。 $R$  は、観測可能な確率変数であるが、 $U$ ,  $V$ ,  $S$ ,  $T$  は、観測不可能な確率変数である。それぞれの帰無仮説を有意水準  $\alpha$  で別々に検定すると、棄却されてはいけない真の仮説を多く棄却してしまう。この場合、FWER は有意水準  $\alpha$  を超えてしまう。なお、これらの確率変数に関して、FWER は  $P(V \geq 1)$  と表せる。

FWER を保証するための 1 つの方法として、有意水準  $\alpha/m$  でそれぞれの仮説を検定するボンフェローニの不等式に基づく多重比較法 (Bonferroni の方法) がある。

帰無仮説を誤って棄却することによって犯される過誤の割合は、確率変数  $Q = \frac{V}{(V+S)}$  によって調べることができる。 $V+S=0$  (間違っただけの棄却の過誤を犯すことができない、つまり 1 つも仮説が棄却されない) の時は、 $Q=0$  と定義する。なお、 $m_0$  個の真の帰無仮説は未知であるのだから、 $Q$  は観測できない確率変数であり、データを収集して検定を行った後でも、その実現値  $q$  は求めることができない値である。そこで Benjamini and

表 1:  $m$  個の帰無仮説の検定結果の分割表

	検定		計
	採択	棄却	
真の帰無仮説	U	V	$m_0$
偽の帰無仮説	T	S	$m - m_0$
計	$m - R$	R	$m$

Hochberg(1995) は、 $Q$  の期待値をとった以下の  $Q_e$  を定義した。

$$Q_e = E(Q) = E\left(\frac{V}{V+S}\right) = E\left(\frac{V}{R}\right)$$

この  $Q_e$  が狭い意味での FDR(False Discovery Rate) の定義である。手法として守るべき基準はこの FDR のあらゆる仮説の真偽における最大値であり、それが広い意味での FDR となる。

FWER と FDR の間には次のような関係がある。(Benjamini and Hochberg(1995) 参照)

- (a) 全ての帰無仮説が真であるならば、FDR は FWER に等しい
- (b)  $m_0 < m$  の時、FDR は FWER より小さいか等しい

その結果、FWER を制御するどんな手法においても、FDR は制御されている。しかし、FDR だけを制御する手法においては、FWER を制御する方法よりも、真の帰無仮説が誤って棄却される確率をゆるめている。つまり、検出力の増加を期待できる。特に、真でない帰無仮説の数が多い時は、S がより大きな値を取るようになるので、検出力を増加させる可能性はより大きいものとなる。

### 3 FDR を制御する多重比較法

#### 3.1 BH 法

仮説  $H_1, H_2, \dots, H_m$  はそれぞれの仮説に対応した  $p$  値 (有意確率)  $P_1, P_2, \dots, P_m$  を持つと仮定する。さらに、昇順に並べ替えられた  $p$  値を  $P_{(1)} \leq P_{(2)} \leq \dots \leq P_{(m)}$  と表し、 $P_{(i)}$  に相当する帰無仮説は  $H_{(i)}$  と表す。また、FWER での有意水準に対応する、FDR の割合を全体としてどれぐらいに抑えたいのかという基準となる値を  $q^*$  (通常は 0.05) とする。このとき、BH 法は以下のような手順である。(Benjamini and Hochberg(1995) 参照)

手順 1  $i = m$  とする。

手順 2  $P_{(i)} \leq \frac{i}{m} q^*$  を満たすならば、 $k = i$  として次へ進む。そうでなければ  $i - 1$  を改めて  $i$  とおき、この手順を繰り返す。ただし、 $i = 1$  まで達したのなら、どの仮説も棄却することなく終了する。

手順 3 仮説  $H_{(i)}; i = 1, 2, \dots, k$  をすべて棄却する。

手順 2 において  $p$  値の高いものから順に検討しており限界値が線形に推移するため、BH 法は The Linear Step-Up Procedure と呼ばれている。

### 3.2 Adaptive BH 法

最初の条件は BH 法と同じである。Adaptive BH 法 (以降 ABH 法とする) は以下のような手順である。(Benjamini and Hochberg(2000) 参照)

手順 1 基準値  $q^*$  で BH 法を実行する。もし 1 つも仮説を棄却することができないならば、終了する。それ以外ならば次へ進む。

手順 2 計算式  $S_i = \frac{1 - P_{(i)}}{m + 1 - i}$  によって、 $S_i$  ( $i = 1, 2, \dots, m$ ) を計算する。

手順 3  $i = 2$  とおく。

手順 4  $S_i \geq S_{i-1}$  ならば  $i + 1$  を改めて  $i$  とおき、この手順を繰り返す。

手順 5  $S_i < S_{i-1}$  と初めてなった時の  $S_i$  を  $S^*$  とおく。最後まで  $S_i \geq S_{i-1}$  ならば、 $S_m$  を  $S^*$  とおく。

手順 6  $\hat{m}_0 = \min \left( \left[ \frac{1}{S^*} + 1 \right], m \right)$  とおく。ここで  $[\ ]$  はガウス記号を表す。

手順 7  $i = m$  とする。

手順 8  $P_{(i)} \leq \frac{i}{\hat{m}_0} q^*$  を満たすならば、 $k = i$  として次へ進む。そうでなければ  $i - 1$  を改めて  $i$  とおき、この手順を繰り返す。(  $\hat{m}_0 \leq m$  なので必ず次に進む  $k$  が見つかる。)

手順 9 仮説  $H_{(i)}$ ;  $i = 1, 2, \dots, k$  をすべて棄却する。

### 3.3 BY (Benjamini and Yekutieli) 法

最初の条件は BH 法と同じである。Benjamini and Yekutieli 法 (以降 BY 法とする) は以下のような手順である。(Benjamini and Yekutieli(2001) 参照)

手順 1  $q^{**} = q^* / \sum_{j=1}^m \left( \frac{1}{j} \right)$  とする。

以下、 $q^*$  の代わりに  $q^{**}$  を用いて、BH 法を行う。

手順 2  $i = m$  とする。

手順 3  $P_{(i)} \leq \frac{i}{m} q^{**}$  を満たすならば、 $k = i$  として次へ進む。そうでなければ  $i - 1$  を改めて  $i$  とおき、この手順を繰り返す。ただし、 $i = 1$  まで達したのなら、どの仮説も棄却することなく終了する。

手順 4 仮説  $H_{(i)}$ ;  $i = 1, 2, \dots, k$  をすべて棄却する。

## 4 過去の比較研究

Benjamini and Hochberg(2000) では、FDR を制御する多重比較法のシミュレーション比較として、BH 法と ABH 法が比較されていた。仮説間が独立な場合、検出力が BH 法よりも高い手法として提案されていた ABH 法が、極端な仮説の配置を除いて勝っていた。

また、一番多くの手法をシミュレーション比較している Benjamini, Krieger and Yekutieli(2005) では、本論文で取り上げていないような手法がいくつか提案されていて、様々な状況下でのシミュレーションがなされていた。この論文では、真の仮説と偽の仮説の割合、仮説間の相関(正の相関)、仮説数をいろいろと変化させてシミュレーションを行っていた。結果として、仮説間が独立な場合、簡便な方法の中では ABH 法の検出力が高いことが示されており、様々な状況下では最も高い検出力が得られる手法として、Resampling を行う方法 Storey, Taylor and Siegmund(2004) が示されていた。

## 5 シミュレーションの設定

シミュレーションの目的は、様々な設定において BH 法、ABH 法、BY 法が FDR を基準値  $q^* = 0.05$  以下に保っているのかを確認することである。

### 5.1 シミュレーション方法

次に述べるシミュレーションパターン 1 から 10 (パターン 7 を除く) においては、分散 1 の正規乱数を用いて、1 標本検定を行い、 $P$  値を生成した。ここでは、帰無仮説を平均 0、対立仮説を平均 1 とした。

パターン 1 については、それぞれ独立に 10 個のデータを用いた。仮説間に正の相関構造を持たせる場合、パターン 2 についてはコレスキー分解を用いて等相関の変数を 10 個ずつ生成し、パターン 3,4 については独立な正規乱数列に対して長さが 10 の移動平均を用いて仮説間に相関構造を持たせた。さらに、負の相関構造を持たせる場合、パターン 5,6 については移動平均を取ったものの符号を反転させることによって生成した。パターン 8,9,10 についてはそれぞれパターン 3,4,6 と同じ生成方法である。

一方、パターン 7 については対比較の設定で、分散 1 の正規乱数を各群 10 個ずつ用いて 2 標本検定を行い、 $P$  値を生成した。帰無仮説を「平均が等しい」、対立仮説を「平均の差が 1 である」とした。

### 5.2 シミュレーションパターン

1 < 仮説間が独立な場合 >

- 正の相関

2 < 仮説間がすべて同じ正の相関である場合 >

3 < 偶数番号の変数に 1 を加え、それを偽の仮説と設定したもの >

- 4 <後半の番号の変数に 1 を加え、それを偽の仮説と設定したもの>
  - 負の相関のある場合
- 5 <偶数番号の変数を反転し 1 を加えたものを偽の仮説と設定したもの>
- 6 <偶数番号の変数を反転させ、後半の番号の変数に 1 を加え、それを偽の仮説と設定したもの>
- 7 <対比較の設定で P 値を計算したもの [対比較の場合]>
  - 仮説数を拡張したもの
- 8 <パターン 3 を偶数番号ではなく、もっと広げた間隔で行ったもの [偽の仮説数は 5 に固定]>
- 9 <パターン 4 を偽の仮説数 5 に固定して、仮説数を増やしたもの>
- 10 <パターン 6 を偽の仮説数 5 に固定して、仮説数を増やしたもの>

以上のような独立 (パターン 1)、正の相関 (パターン 2,3,4,8,9)、負の相関 (パターン 5,6,7,10) の計 10 個の状況において、シミュレーションを行った。シミュレーション回数はいずれのパターンもすべて 10000 回である。

### 5.3 シミュレーションの正当性

BH 法、ABH 法については、特殊な場合でのシミュレーション評価が書かれている Benjamini, Krieger and Yekutieli(2005) に掲載されていたものと同じ設定のシミュレーションを行い、シミュレーションの正当性を確認する。ここでのシミュレーション回数は 20000 回である。表 2 がその比較状況である。Benjamini, Krieger and Yekutieli(2005) には数値が記されていないため表の数値は図から読み取ったものである。結果として、BH 法、ABH 法ともに非常に近い値が得られている。

表 2: 正当性のための FDR の比較

$m_0 : m_1$	本研究の FDR		$m_0 : m_1$	BKY(2005) の FDR	
	BH	ABH		BH	ABH
0:8	0	0	0:8	0	0
2:6	0.01231	0.03488	2:6	0.012	0.034
4:4	0.02446	0.04144	4:4	0.024	0.042
6:2	0.03683	0.04327	6:2	0.036	0.044
8:0	0.05080	0.05080	8:0	0.050	0.050

BY 法については、手順からも分かるとおり  $q^{**} = q^* / \sum_{j=1}^m (\frac{1}{j})$  とし、 $q^*$  の代わりに  $q^{**}$  を用いて、BH 法を行っている点に注目してほしい。これは、仮説数に依存して、基準値である  $q^*$  を 0.05 からそれよりも小さな値になるように調整しているため、保守的な方法であることに間違いは無い。従って、BY 法はかなり厳しく FDR を制御している方法なので

ある。また、佐藤ら (2004) でも BY 法はかなり保守的な方法であると書かれていたことと一致している。

## 6 シミュレーションによる評価と考察

### 6.1 <パターン 1> (独立)

表 3: 独立な場合の FDR と FWER

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
5: 5	0.02349	0.04159	0.00877	0.0980	0.1760	0.0270
15: 5	0.03747	0.04513	0.01087	0.1220	0.1566	0.0237
25: 5	0.04145	0.04489	0.00977	0.1178	0.1343	0.0183
10:10	0.02515	0.04572	0.00752	0.1676	0.3123	0.0298
20:10	0.03271	0.04524	0.00835	0.1855	0.2731	0.0258
5:15	0.01289	0.04278	0.00363	0.1418	0.4190	0.0242
15:15	0.02518	0.04706	0.00619	0.2288	0.4252	0.0299
10:20	0.01635	0.04456	0.00424	0.2131	0.5233	0.0297
5:25	0.00838	0.03886	0.00196	0.1496	0.5624	0.0195

真の仮説と偽の仮説の配置  $m_0:m_1$  が 5:5, 15:5, 25:5, 10:10, 20:10, 5:15, 15:15, 10:20, 5:25 の 9 通りを考えシミュレーションを行った。表 3 は方法別 FDR と方法別 FWER を示したものである。この表から言えることは、すべての場合において ABH 法の FDR 値がその他の方法の FDR 値よりも勝っていることである。また、データ間に相関がある場合に用いると記されていた BY 法は、相関がない場合は FDR 値はかなり低く、かなり保守的な方法であることが分かった。さらに、BH 法において、真の仮説の数を 5 に固定して表を見た場合、FDR 値は減少しているが、逆に偽の仮説の数を 5 に固定して表を見た場合、FDR 値は増加していた。それに比べて、ABH 法の FDR 値においては、いずれの場合も安定した FDR 値を取っていた。

また、FDR を制御する方法は (誤って棄却された真の帰無仮説の数)/(棄却された仮説の数) の期待値を 0.05 以下に保つことを保証しているので、FWER 値は 0.05 を上回ってもよい。この表より BH 法と ABH 法においてはいずれの場合も FWER 値は 0.05 を超えているが、BY 法においては、いずれの場合も FWER 値は 0.05 以下に制御されていることが分かった。さらに、この表から BH 法では仮説の数が少ない場合、FWER 値は比較的小さな値を取っているが、逆に仮説の数が増加すると、それに伴って FWER 値も増加している傾向が確認できた。それに対し、ABH 法は偽の仮説の数に伴って FWER 値が増加していた。

### 6.2 <パターン 2> (正の相関)

変数間のすべての相関係数が  $\text{cor}=0.1, 0.5, 0.9$  の 3 パターンにおいて、仮説の配置  $m_0:m_1$  が 5:5, 15:5, 25:5, 10:10, 20:10, 5:15, 15:15, 10:20, 5:25 の 9 通りを考えシミュレーションを行った。検定は両側と片側の両方を考えた。この表 4、表 5 から言えることは、以下のよ

うなことである。

表 4: 仮説間に同じ正の相関 (0.1,0.5,0.9) がある場合の FDR と FWER [両側]

$m_0:m_1$	cor	FDR			FWER		
		BH	ABH	BY	BH	ABH	BY
5: 5	0.1	0.02321	0.04158	0.00893	0.0976	0.1753	0.0269
5: 5	0.5	0.02595	0.04858	0.00887	0.0915	0.1643	0.0277
5: 5	0.9	0.01886	0.06021	0.00663	0.0489	0.1466	0.0181
15: 5	0.1	0.03692	0.04408	0.01003	0.1207	0.1529	0.0227
15: 5	0.5	0.03692	0.05463	0.01001	0.1011	0.1370	0.0254
15: 5	0.9	0.02669	0.07553	0.00780	0.0480	0.1041	0.0159
25: 5	0.1	0.04116	0.04604	0.00906	0.1189	0.1393	0.0177
25: 5	0.5	0.03792	0.05224	0.01064	0.0999	0.1207	0.0271
25: 5	0.9	0.02371	0.06676	0.00607	0.0387	0.0818	0.0116
10:10	0.1	0.02645	0.04618	0.00747	0.1710	0.5271	0.0310
10:10	0.5	0.02410	0.04653	0.00660	0.1255	0.3640	0.0326
10:10	0.9	0.01665	0.04510	0.00503	0.0524	0.1577	0.0161
20:10	0.1	0.03243	0.04460	0.00782	0.1808	0.2638	0.0249
20:10	0.5	0.03012	0.04998	0.00828	0.1327	0.1921	0.0318
20:10	0.9	0.02253	0.06322	0.00571	0.0569	0.1010	0.0163
5:15	0.1	0.01293	0.04395	0.00354	0.1397	0.4134	0.0238
5:15	0.5	0.01257	0.04173	0.00343	0.1075	0.3174	0.0279
5:15	0.9	0.01055	0.04214	0.00283	0.0561	0.1868	0.0171
15:15	0.1	0.02524	0.04566	0.00619	0.2188	0.4019	0.0312
15:15	0.5	0.02386	0.04946	0.00570	0.1574	0.2802	0.0352
15:15	0.9	0.01890	0.05435	0.00462	0.0640	0.1220	0.0174
10:20	0.1	0.01721	0.04545	0.00392	0.2167	0.5268	0.0300
10:20	0.5	0.01578	0.04456	0.00344	0.1538	0.3660	0.0311
10:20	0.9	0.01315	0.04457	0.00302	0.0686	0.1541	0.0189
5:25	0.1	0.00858	0.03861	0.00223	0.1492	0.5530	0.0209
5:25	0.5	0.00830	0.03727	0.00188	0.1150	0.4377	0.0234
5:25	0.9	0.00716	0.03126	0.00161	0.0596	0.2159	0.0146

両側検定の場合は、BH 法と BY 法は、FDR 値が 0.05 以下に保たれていて、ABH 法においては、真の仮説が総仮説数の 1/2 以上存在する場合で、仮説間の相関が 0.5 以上のものはほとんど FDR 値が 0.05 を上回っていた。それ以外の場合では、0.05 を守りながら ABH 法が一番高い FDR 値を取ることが分かった。

片側検定の場合は、両側検定同様、BH 法と BY 法は、FDR 値が 0.05 以下に保たれていて、ABH 法においては、ほとんどの場合で、FDR 値が 0.05 を超えていた。

BH 法は、仮説間が独立である場合に用いられる方法論であるが、今回のシミュレーションから、仮説間にどのような (正の) 相関構造があったとしても、BH 法の FDR 値は 0.05 以下に保たれていることが分かった。また、仮説間に相関のある場合を考慮したと記されていた方法論 BY 法においては、いずれの場合もかなり保守的な結果が得られた。FWER 値もほとんどの場合 0.05 以下に保たれていることから、もしかすると、BY 法は FWER を制御する方法論であるのかと思われるほどだった。BH 法や ABH 法の FWER 値は、ほとんどすべての配置で 0.05 を上回っていた。

表 5: 仮説間に同じ正の相関 (0.1,0.5,0.9) がある場合の FDR と FWER [片側]

$m_0:m_1$	cor	FDR			FWER		
		BH	ABH	BY	BH	ABH	BY
5: 5	0.1	0.02413	0.05227	0.00823	0.1200	0.2240	0.0343
5: 5	0.5	0.02567	0.08086	0.00867	0.1157	0.2543	0.0424
5: 5	0.9	0.02204	0.10388	0.00799	0.0649	0.2456	0.0266
15: 5	0.1	0.03677	0.05411	0.00960	0.1496	0.2032	0.0298
15: 5	0.5	0.03424	0.09708	0.00899	0.1233	0.1995	0.0357
15: 5	0.9	0.02811	0.11152	0.00819	0.0558	0.1504	0.0198
25: 5	0.1	0.04320	0.05535	0.01038	0.1602	0.1948	0.0278
25: 5	0.5	0.03927	0.10171	0.01088	0.1299	0.1871	0.0399
25: 5	0.9	0.02738	0.09554	0.00698	0.0468	0.1155	0.0143
10:10	0.1	0.02579	0.05308	0.00716	0.2023	0.5429	0.0460
10:10	0.5	0.02437	0.07100	0.00676	0.1615	0.3885	0.0521
10:10	0.9	0.01986	0.06536	0.00603	0.0703	0.2071	0.0256
20:10	0.1	0.03238	0.05266	0.00819	0.2281	0.3282	0.0427
20:10	0.5	0.02944	0.08867	0.00720	0.1656	0.2505	0.0490
20:10	0.9	0.02572	0.08881	0.00681	0.0674	0.1363	0.0246
5:15	0.1	0.01247	0.04805	0.00343	0.1594	0.4318	0.0363
5:15	0.5	0.01257	0.06288	0.00371	0.1395	0.3787	0.0457
5:15	0.9	0.01146	0.06463	0.00334	0.0790	0.2683	0.0273
15:15	0.1	0.02439	0.05283	0.00635	0.2604	0.4435	0.0529
15:15	0.5	0.02427	0.08785	0.00614	0.2038	0.3377	0.0611
15:15	0.9	0.02080	0.07506	0.00517	0.0825	0.1564	0.0253
10:20	0.1	0.01658	0.05199	0.00432	0.2511	0.5358	0.0518
10:20	0.5	0.01625	0.07162	0.00391	0.1952	0.3977	0.0576
10:20	0.9	0.01475	0.06434	0.00384	0.0910	0.2026	0.0299
5:25	0.1	0.00820	0.04220	0.00214	0.1704	0.5554	0.0351
5:25	0.5	0.00830	0.05033	0.00201	0.1544	0.4643	0.0424
5:25	0.9	0.00813	0.04655	0.00193	0.0897	0.3021	0.0254

### 6.3 <パターン 3,4 > (正の相関)

パターン 2 の結果から片側検定の方が手法の特徴がよく出するため、以下のシミュレーションではすべて片側検定で行った。

表 6、表 7 から以下のようなことが分かる。

パターン 3 の BH 法と ABH 法では、仮説の配置 (真と偽) が異なっても FDR 値は 2 手法とも安定して保たれていた。このシミュレーションの設定は、隣り合う仮説間には強い正の相関が見られ、離れている仮説間には弱い正の相関が見られる構造なので、全体的に見るとそれほど強い相関があるとは言えない。以前の結果から、ABH 法は仮説間に正の相関関係 (cor=0.5,0.9) がある場合、FDR 値は 0.05 を超えていた。しかし、今回は BH 法とほぼ同じ値を取り、FDR 値はいずれの場合も 0.05 以下に保たれていた。

パターン 4 の BH 法と ABH 法では、いずれの場合も FDR 値が 0.05 以下に保たれており、真の仮説が多いと高い FDR 値を取っており、偽の仮説が多いと低い FDR 値を取っていた。さらに、仮説数が増加するにつれて、FDR 値は減少していた。その違いは顕著に見られる。

また、FWER 値も、若干ではあるが、FDR 値と同じように真と偽の仮説の影響を受け

表 6: 偶数番号の変数に 1 を加え、それを偽の仮説と設定したもの [パターン 3]

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
5: 5	0.01911	0.01969	0.00717	0.0347	0.0373	0.0122
15: 5	0.01443	0.01458	0.00343	0.0280	0.0289	0.0065
25: 5	0.01525	0.01492	0.00393	0.0295	0.0296	0.0075
10:10	0.01473	0.01455	0.00416	0.0281	0.0290	0.0075
20:10	0.01432	0.01398	0.00365	0.0276	0.0278	0.0070
5:15	0.01506	0.01517	0.00486	0.0292	0.0302	0.0093
15:15	0.01252	0.01259	0.00368	0.0247	0.0250	0.0064
10:20	0.01385	0.01384	0.00361	0.0273	0.0276	0.0069
5:25	0.01365	0.01395	0.00346	0.0271	0.0277	0.0061

表 7: 後半の番号の変数に 1 を加え、それを偽の仮説と設定したもの [パターン 4]

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
5: 5	0.01760	0.01861	0.00794	0.0266	0.0326	0.0100
15: 5	0.02110	0.02097	0.00493	0.0254	0.0274	0.0059
25: 5	0.02474	0.02421	0.00694	0.0285	0.0289	0.0077
10:10	0.01560	0.01505	0.00497	0.0223	0.0264	0.0064
20:10	0.01838	0.01816	0.00480	0.0242	0.0261	0.0061
5:15	0.00750	0.00762	0.00278	0.0152	0.0261	0.0045
15:15	0.01340	0.01334	0.00329	0.0195	0.0230	0.0046
10:20	0.00894	0.00965	0.00244	0.0171	0.0234	0.0039
5:25	0.00495	0.00542	0.00137	0.0128	0.0226	0.0026

ていた。BY 法はどちらのパターンでも、保守的に FDR 値を保っていた。

#### 6.4 <パターン 5> (負の相関がある場合)

このシミュレーションの設定は、仮説間に正の相関と負の相関が交互に現れる相関構造がある場合である。従って、真の仮説の中身がすべて正の相関を持っていて、真の仮説と偽の仮説の間で負の相関を持つような構造である。このような場合、表 8 から分かるように、ABH 法はいずれの場合も、FDR 値が 0.05 を超えていた。それに対して、BH 法は ABH 法よりも安定した FDR 値を得ることができて、なおかつ FDR 値が 0.05 以下に保たれていることが分かった。さらに、BH 法は仮説の配置 (真と偽) が異なっても FDR 値を安定して保っていた。

また、FWER 値は、ABH 法を除くと、FDR 値同様 BH 法が一番高く、その次に BY 法の順になっている。この場合も BY 法の FWER 値はいずれも 0.05 以下に保たれていた。

表 8: 偶数番号の変数を反転し 1 を加えたもの (反転させたものを偽)

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
5: 5	0.02367	0.08502	0.00863	0.0938	0.2458	0.0380
15: 5	0.02263	0.08780	0.00607	0.1116	0.2346	0.0374
25: 5	0.02334	0.07857	0.00628	0.1290	0.2233	0.0444
10:10	0.02284	0.08829	0.00629	0.1161	0.2365	0.0381
20:10	0.02337	0.08161	0.00579	0.1302	0.2274	0.0404
5:15	0.02290	0.09072	0.00685	0.1143	0.2400	0.0412
15:15	0.02219	0.08254	0.00515	0.1270	0.2295	0.0379
10:20	0.02230	0.07875	0.00572	0.1234	0.2248	0.0414
5:25	0.02294	0.08039	0.00545	0.1274	0.2232	0.0409

表 9: 偶数番号の変数を反転させ、後半の番号の変数に 1 を加えたもの

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
5: 5	0.02500	0.02872	0.00910	0.1158	0.1353	0.0443
15: 5	0.03231	0.03231	0.00885	0.0424	0.0424	0.0351
25: 5	0.03959	0.03959	0.00928	0.0472	0.0472	0.0316
10:10	0.02255	0.02255	0.00678	0.0451	0.0451	0.0447
20:10	0.02953	0.02953	0.00840	0.0443	0.0443	0.0457
5:15	0.01226	0.01650	0.00343	0.1144	0.1562	0.0390
15:15	0.02308	0.02306	0.00526	0.0459	0.0459	0.0435
10:20	0.01367	0.01367	0.00361	0.0410	0.0410	0.0426
5:25	0.00793	0.01163	0.00198	0.1114	0.1665	0.0338

## 6.5 <パターン 6> (負の相関がある場合)

このシミュレーションの設定は、仮説間で適度に正と負が混じり合ったような相関構造を持つ場合である。

表 9 より、いずれの手法も FDR 値が 0.05 以下に保たれていて、BH 法と ABH 法は、ほとんど同じような FDR 値を取っていた。しかし、真の仮説の数が 5 の場合のみ、BH 法と ABH 法の FDR 値に差が見られた。値としては、ABH 法が BH 法よりも高い FDR 値を取っていた。また、いずれの手法も仮説数が増加するにつれて、FDR 値が減少していった。

FWER 値は、真の仮説の数が 5 の場合のみ、BH 法と ABH 法で高くなっていた。ここでは、5:25 の状況から真の仮説の数を順番に増やしていくような配置を設定し、追加のシミュレーションを行った。その結果、真の仮説が 5 の場合から、真の仮説が 10 になるまで FWER 値は順番に減少していていることが確認できた。また、BY 法の FWER 値はパターン 5 に続き、いずれも 0.05 以下に保たれていた。

表 10: 対比較の絶対値なし (下側)

a:b	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
5:0	0.04230	0.04230	0.01530	0.0423	0.0423	0.0153
4:1	0.02745	0.04893	0.00924	0.0768	0.1334	0.0236
3:2	0.01933	0.04500	0.00655	0.0804	0.1843	0.0253
2:3	0.01999	0.04589	0.00650	0.0807	0.1888	0.0243
1:4	0.02855	0.04803	0.00994	0.0797	0.1334	0.0246
0:5	0.04240	0.04240	0.01500	0.0424	0.0424	0.0150

### 6.6 <パターン 7> (負の相関がある場合)

ここでは、全部で5群あるうち、平均0と平均1の配置 a:b が 5:0, 4:1, 3:2, 2:3, 1:4, 0:5 の計6通りを考えている。対比較は絶対値のない下側検定で行った。

表 10 より、全体的に正の相関や負の相関が混じっていて、仮説間の相関構造が 0.5 前後であっても、いずれの手法も FDR 値が 0.05 以下に保たれていることが分かった。また、ABH 法の FDR 値は、いずれの場合も安定しており、非常に高い FDR 値を取っていることが分かった。さらに、BH 法において、仮説がすべて真の場合 (5:0, 0:5) は FDR 値は高く、仮説に真と偽がほぼ均等に配置されている場合 (3:2, 2:3) は、他よりも厳しめに FDR 値を守っている。FWER 値はその逆で、仮説に真と偽がほぼ均等に配置されている場合 (3:2, 2:3) は高い値をとり、仮説がすべて真の場合 (5:0, 0:5) は厳しく FWER 値を守っている。ABH 法の FWER 値では、BH 法と同様のことが顕著に現れていた。

### 6.7 <パターン 8,9,10> (仮説数を増やしたもの)

最大仮説数を 100 まで増やしていったその動向を確認した。

表 11: パターン 3 を偶数番号ではなく間隔を広げたもの (パターン 8)

間隔	$m_0:m_1$	FDR			FWER		
		BH	ABH	BY	BH	ABH	BY
2	5:5	0.01911	0.01969	0.00717	0.0347	0.0373	0.0122
4	15:5	0.02229	0.02241	0.00641	0.0287	0.0295	0.0081
8	35:5	0.02276	0.02293	0.00460	0.0262	0.0262	0.0052
16	75:5	0.02502	0.02495	0.00508	0.0266	0.0266	0.0054
20	95:5	0.02417	0.02421	0.00422	0.0255	0.0255	0.0044

表 11、表 12、表 13 から以下のようなことが分かる。

パターン 8 では、偽の仮説の数を 5 に固定して間隔を 2, 4, 8, 16, 20 とし、真の仮説の数を増やしていった。若干仮説数による影響を受けているが、やはりパターン 3 と同様に、BH 法と ABH 法ともに安定して、FDR 値を 0.05 以下に保っていた。FWER 値は FDR 値と逆の結果となった。

パターン 9 では、偽の仮説の数を 5 に固定して、真の仮説の数を増やしていった。若干

表 12: 偽の仮説数を 5 に固定して、パターン 4 の拡張したもの (パターン 9)

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
25:5	0.02309	0.02296	0.00549	0.0267	0.0277	0.0063
45:5	0.02127	0.02094	0.00484	0.0233	0.0237	0.0053
95:5	0.02399	0.02379	0.00476	0.0253	0.0253	0.0051

表 13: 偽の仮説数を 5 に固定して、パターン 6 を拡張したもの (パターン 10)

$m_0:m_1$	FDR			FWER		
	BH	ABH	BY	BH	ABH	BY
25:5	0.03959	0.04007	0.00900	0.0472	0.0483	0.0304
45:5	0.04017	0.04014	0.00781	0.0451	0.0451	0.0225
95:5	0.04353	0.04353	0.00740	0.0458	0.0458	0.0172

仮説数による影響を受けているが、やはりパターン 4 と同様に、BH 法と ABH 法ともに安定して、FDR 値を 0.05 以下に保っていた。FWER 値は FDR 値と逆の結果となった。

パターン 10 では、偽の仮説の数を 5 に固定して、真の仮説の数を増やしていった。若干仮説数による影響を受けているが、やはりパターン 6 と同様に、BH 法と ABH 法ともに安定して、FDR 値を 0.05 以下に保っていた。FWER 値は FDR 値と逆の結果となった。

最終的に、仮説数をなるべく多く増やしてみたが拡張前のパターン別結果と同様の結果が得られた。また、パターン 10 における BH 法と ABH 法の FDR 値は、他の 2 つ (パターン 8,9) の FDR 値と比較して、大きくなっている。これは拡張前と同様のことが言える。

## 7 各方法の性能のまとめ

今回は、総仮説数が 100 までしかシミュレーションを行っていないが、どのような状況下でも BH 法の FDR 値は 0.05 以下に保たれていることが分かった。また、ABH 法は、仮説間に正の相関と負の相関が交互に現れるもの、強い正の相関があるもの以外においては、FDR 値を 0.05 以下に保っていることが分かった。BY 法も BH 法と同様に保たれていた。BY 法については、ほとんどの場合で FWER が 0.05 以下に保たれていた。そのため、BY 法は、非常に保守的な方法であるといえる。

また、すべてのシミュレーションパターンの中で、BH 法の FDR 値が最大になる (0.05 に近くなる) 値を 3 つ取り上げてみると表 14 のようになる。

1 つ目は、仮説間に正の相関 0.1 があり、配置 (25:5) で片側検定を行った場合であった。これは相関なし ( $\text{cor}=0$ ) に近い場合、つまり独立に近い場合なので、このような設定の辺りで 0.05 を越えることはほとんど考えられない。次に、2 つ目、3 つ目は、対比較を行う場合である。配置は (0:5)(5:0) である。この場合は、どの群とどの群をとっても仮説は真であるので、すべてが真の仮説の場合である。すべてが真の仮説の場合はボンフェローニの方法と近い状況になり、越えることは考えにくい。また、シミュレーションでは 5 群で行っているため、2 倍の 10 群にまで増やした時でも総仮説数は 45 であり、今回のシミュレーション

表 14: BH 法で FDR が 0.05 に近い状況

仮説間に正の相関 (片側検定)		
$m_0:m_1$	cor	BH
25:5	0.1	0.04320
対比較の場合 (5 群)		
0:5		0.04240
5:0		0.04230

ンで仮説数 100 まで調べた関係からも、通常の対比較では BH 法の FDR 値は 0.05 を超えないと思われる。

最終的に以下のような結論が得られる。

- BH 法は仮説間にどのような相関関係があっても、安定した検定結果が得られる手法である。従って、仮説間が独立ではなく、相関関係 (正と負) があつたとしても、ある程度信頼性のある結果を得ることができる。
- ABH 法はある特殊な状況 (正の相関が強く、仮説に真の仮説が多く含んでいる場合や仮説間に正の相関と負の相関が交互に現れる場合) を除けば、BH 法と同様な安定した FDR 値を得ることができる手法である。さらに、ABH 法は仮説間に相関がない場合や低い場合、対比較を行う場合においては、BH 法に勝る。
- BY 法は BH 法と同様、仮説間にどのような相関構造があっても、用いることは可能である。しかし、シミュレーション結果を見てみると FWER を制御しているような振り舞いをしているので、FDR を制御する多重比較法の立場から見るとかなり保守的な方法である。

## 8 おわりに

本論文では、最近の多重比較法の主流である FDR を制御する多重比較法のプログラム化を行い、そのシミュレーション評価をすることができた。ここでは、BH 法の安定性、ABH 法が強い状況、などを確認することができた。

今後の課題として ABH 法の使える範囲の明確化があるが、今のところ、仮説間に未知の相関構造がある場合には、BH 法を適用するのがよいと言える。

## 謝辞

本研究に関連して活発に議論していただいた田中豊先生、木村美善先生、富田誠先生に深く感謝いたします。

## 参考文献

- Benjamini, Y., Hochberg, Y. (1995): Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing, *J. R. Statist. Soc. ser.B*, **57**(1), 289-300.
- Benjamini, Y., Hochberg, Y. (2000): On the Adaptive Control of the False Discovery Rate in Multiple Testing With Independent Statistics, *J. Edu. Behavioral Statistics*, **25**(1), 60-83.
- Benjamini, Y., Krieger, A. M. and Yekutieli, D. (2005): Adaptive Linear Step-up Procedures that control the False Discovery Rate, Unpublished paper, <http://www.math.tau.ac.il/~ybenja/MyPapers/bkymarch9.pdf> .
- Benjamini, Y. and Yekutieli, D. (2001): The Control of the False Discovery Rate in Multiple Testing under Dependency, *Annals of Statistics*, **29**(4), 1165-1188.
- 堀内賢太郎 (2004): “R による多重比較法の研究”, 南山大学数理情報学部数理科学科卒業論文.
- 佐藤亜香里・角谷伸一・田崎武信 (2004): “FDR コンセプトと FDR 法についての考察”, 塩野義製薬株式会社解析センター.
- Storey, J. D., Taylor, J. E., and Siegmund, D. (2004): Strong Control, Conservative Point Estimation, and Simultaneous Conservative Consistency of False Discovery Rates: a Unified Approach, *J. R. Statist. Soc. Ser.B*, **66**(1), 187-205.