

## 多重比較補正 **False Discovery Rate (FDR)** 法

FDR法による多重比較を実施する以下のMATLABスクリプトの (簡素な) 取扱い説明書です  
<https://jp.mathworks.com/matlabcentral/fileexchange/27418-fdr-bh>

### ◆注意点

- ・小林がメモ書き程度に作成したもので、ミスがあるかもです…。
- ・不確かなところには (?) をつけています。

### ◆使用方法

- 1, FDRのMATLABプログラム (fdr\_bh.m) を  
<https://jp.mathworks.com/matlabcentral/fileexchange/27418-fdr-bh>  
からダウンロードする。あるいは、上記サイトのFunctionsにあるスクリプトを新規スクリプトにコピーする
- 2, 次の【入力】変数 (**pvals, q, method, report**) をすべて入力していく。  
※入力はコマンドウィンドウ上で記述して、エンターキーで確定する。

-----  
▼**入力** (※**緑字**はスクリプトそのものからの引用分です。※**赤字**はデフォルトとして紹介されているものです)

### ・pvals

- : FDR補正を行いたいp値を入力する。(e.g., pvals = [0.001 0.01 0.05])  
入力順などに制限はないようなので, pvals = [0.6 0.03 0.15 0.10 0.001] とかでもOKなはず…。
- ※MATLABの使用上 [ ] でp値をくくらないとうまく動きません
- ※あくまでp値が対象なので, 「1以上」の数値はエラーがでます
- ※ちなみに, pvals = [ 0.08 0.12  
0.52 0.03] とかでもOKです。

```
% pvals - A vector or matrix (two dimensions or more) containing the  
%         p-value of each individual test in a family of tests.
```

### ・q

- : 有意水準を指定する。一般的には, **q = 0.05** でOK。
- ※qは, desired false discovery rateを入力する変数なのですが, 参考の表1を見て, 有意水準と同じような意味かなと判断しています (なので, 間違っているかもです…)
- % q - The desired false discovery rate. {default: 0.05}

## ・method

: FDR補正の方法を選択できるようです。詳しいことは実際のスクリプトを確認してみてください。

デフォルトで良い場合には, **method = 'pdep'** とコマンドウィンドウに記す。

※ 引用符 ' ' がないとエラーが起きるので, 必ず付けること。

```
% method - ['pdep' or 'dep'] If 'pdep,' the original Benjamini & Hochberg
%         FDR procedure is used, which is guaranteed to be accurate if
%         the individual tests are independent or positively dependent
%         (e.g., Gaussian variables that are positively correlated or
%         independent). If 'dep,' the FDR procedure
%         described in Benjamini & Yekutieli (2001) that is guaranteed
%         to be accurate for any test dependency structure (e.g.,
%         Gaussian variables with any covariance matrix) is used. 'dep'
%         is always appropriate to use but is less powerful than 'pdep.'
%         {default: 'pdep'}
```

## ・report

: 結果をコマンドウィンドウに出力するかどうかを選択できる。(が, ここで紹介している方法では, どちらを選んでも, 変化はない。おそらく, ▶実行 ができる場合には意味をなす変数だと思われます)

とりあえず, **report = 'yes'** とコマンドウィンドウに入力すればよいはずです。

```
% report - ['yes' or 'no'] If 'yes', a brief summary of FDR results are
%         output to the MATLAB command line {default: 'no'}
```

---

## 3, コマンドウィンドウに

**[h, crit\_p, adj\_ci\_cvrg, adj\_p]=fdr\_bh(pvals,q,method,report)**

と入力 (あるいはスクリプト本文からコピー) して, エンターキーを押す

※このとき, 必ず, 現在のフォルダーに fdr\_bh.m が存在するようにしてください。

現在のフォルダーに fdr\_bh.m がないとエラーが出るので, フォルダを指定してください。

※MATLAB2014aでは, ▶実行 を押してもエラーが出たため, こちらの方法を記してみました。

もしかすると, ▶実行 を押すだけでもOKなバージョンもあるのかも… (?)

## 4, コマンドウィンドウ, およびワークスペースに結果が出力される。

出力内容については 次の ▼出力 のところに記しています。

---

▼出力 (※緑字はスクリプトそのものから、引用している部分です)

・h

: 出力が「1」であれば、FDR法によって、多重比較補正を行っても、有意であることを意味している。  
逆に、出力が「0」であれば、非有意であることを意味している。

```
% h      - A binary vector or matrix of the same size as the input "pvals."  
%         If the ith element of h is 1, then the test that produced the  
%         ith p-value in pvals is significant (i.e., the null hypothesis  
%         of the test is rejected).
```

・crit\_p

: 【入力】のpvalsが、このcrit\_p以下であれば、FDR補正後にも有意になりますよ、という値。

```
% crit_p - All uncorrected p-values less than or equal to crit_p are  
%         significant (i.e., their null hypotheses are rejected). If  
%         no p-values are significant, crit_p=0.
```

・adj\_ci\_cvrg

: FDR補正後に有意なp値について、confidence intervalsを算出するとき利用できる値 (?)

```
% adj_ci_cvrg - The FCR-adjusted BH- or BY-selected  
%              confidence interval coverage. For any p-values that  
%              are significant after FDR adjustment, this gives you the  
%              proportion of coverage (e.g., 0.99) you should use when generating  
%              confidence intervals for those parameters. In other words,  
%              this allows you to correct your confidence intervals for  
%              multiple comparisons. You can NOT obtain confidence intervals  
%              for non-significant p-values. The adjusted confidence intervals  
%              guarantee that the expected FCR is less than or equal to q  
%              if using the appropriate FDR control algorithm for the  
%              dependency structure of your data (Benjamini & Yekutieli, 2005).  
%              FCR (i.e., false coverage-statement rate) is the proportion  
%              of confidence intervals you construct  
%              that miss the true value of the parameter. adj_ci=NaN if no  
%              p-values are significant after adjustment.
```

・adj\_p

: FDR法による調整p値。この値が、【入力】のq (i.e., 設定した有意水準、大体的場合0.05) 以下であれば、FDR法による多重比較では、有意になることを意味している。

```
% adj_p  - All adjusted p-values less than or equal to q are significant  
%         (i.e., their null hypotheses are rejected). Note, adjusted  
%         p-values can be greater than 1.
```

---

#### ◆実施例

・テストの平均得点 (小数点以下は省略)

→A条件 55点

→B条件 58点

→C条件 60点

→D条件 50点

→E条件 56点

・A条件と他4条件の対比較を実施する

→A vs. B :  $p = .045$

→A vs. C :  $p = .008$

→A vs. D :  $P = .010$

→A vs. E :  $p = .694$

・入力

→pvals = [0.048 0.008 0.01 0.694]

→q=0.05

→method='pdep'

→report = 'yes'

・出力

→h = 0 1 1 0

→crit\_p = 0.01

→adj\_ci\_cvrg = 0.9750

→adj\_p = 0.0640 0.0200 0.0200 0.6940

#### ◆参考

：松田眞一(2008) FDRの概説とそれを制御する多重検定法の比較 計量生物学, 29, 125-139

：上述の▼入力の「q」では、「有意水準」を入力すると書くにあたって参考にしました。表1とその周辺の記述を踏まえて、「q」で入力の求められているdesired false discovery rateは、有意水準と同じような意味かなと判断しています。

➡<https://mumu.jpn.ph/forest/computer/2017/04/11/8581/> では、以下のように記されている。

false discovery rateとは、多重検定において、下した全てのpositive inference（有意差ありと判断した検定結果）の中で誤っているものの割合のことである。

FDR = 誤ったpositive inferenceの数 / 下した全positive inferenceの数

FDRは多重検定においてαエラーに相当する概念である。