Series of lectures on Bayesian selective inference
Lecture 1: Selective inference

Daniel Yekutieli

Statistics and OR
Tel Aviv University

Spring School of Research Unit “Structural Inference in Statistics”,
17-21 March 2014, Konigs Wusterhausen

Lectures on Bayesian selective inference

1. Selective inference
2. Bayesian FDR controlling testing procedure
3. Optimal exact tests for complex alternative hypotheses
4. Selection-Adjusted Bayesian Inference

Plan for this talk

1. Benjamini and Hochberg '95
   - FDR and the BH procedure
2. Selective inference
   - simultaneity, post-hoc analysis and FDR control
   - False Coverage-statement Rate control and BH procedure
3. Bayesian FDR
   - Connection to selective inference, frequentist FDR and BH procedure
   - How do we control the Bayesian FDR?
   - Connection to BH procedure

J. R. Statist. Soc. B (1995)
57, *No.* 1, *pp.* 289–300

Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing

By YOAV BENJAMINI† and YOSEF HOCHBERG

Tel Aviv University, Israel

[Received January 1993. Revised March 1994]

SUMMARY

The common approach to the multiplicity problem calls for controlling the familywise error rate (FWER). This approach, though, has faults, and we point out a few. A different approach to problems of multiple significance testing is presented. It calls for controlling the expected proportion of falsely rejected hypotheses—the false discovery rate. This error rate is equivalent to the FWER when all hypotheses are true but is smaller otherwise. Therefore, in problems where the control of the false discovery rate rather than that of the FWER is desired, there is potential for a gain in power. A simple sequential Bonferroni-type procedure is proved to control the false discovery rate for independent test statistics, and a simulation study shows that the gain in power is substantial. The use of the new procedure and the appropriateness of the criterion are illustrated with examples.

Keywords: BONFERRONI-TYPE PROCEDURES; FAMILYWISE ERROR RATE; MULTIPLE-COMPARISON PROCEDURES; *p*-VALUES

- “When pursuing multiple inferences, researchers tend to select the (statistically) significant ones for emphasis, discussion and support of conclusions.”
- “. . . To control this multiplicity (selection) effect , classical multiple-comparison proceudres aim to control the probability of committing any type I error . . .”
- “ . . . In this work we suggest a new point of view on the problem of multiplicity. . . . a desirable error to control may be the expected proportion of errors among the rejected hypotheses, which we term the false discovery rate (FDR).”

BH ‘95 multiple testing framework

- m tested null hypotheses $H_1 \cdots H_m$
- m_0 null hypotheses ($P_i \sim U[0,1]$),
  $m_1 = m - m_0$ false null hypotheses ($P_i \leq U[0,1]$)
- Rejecting a null hypothesis is a discovery, a true discovery is rejecting a false null hypothesis and a false discovery is erroneously rejecting a true null hypothesis
- R is the number of discoveries and V is the number of false discoveries

$$FDR = EQ, \quad Q = \begin{cases} 0 & \text{if } R = 0 \\ V/R & \text{if } R > 0 \end{cases}$$

- cf. $FWE = \Pr(V > 0)$

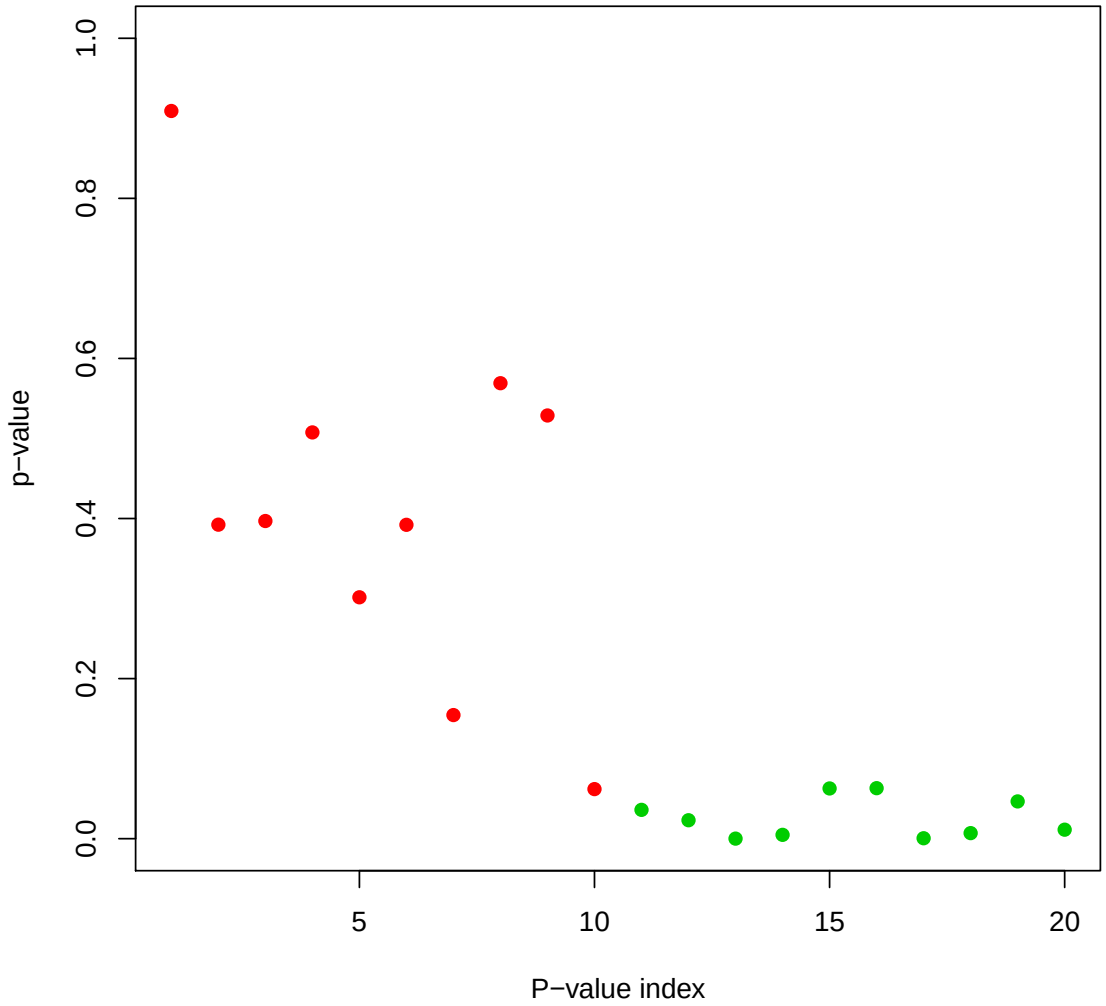
Level q BH procedure

1. Sort the p-values $P_{(1)} \leq P_{(2)} \leq \cdots \leq P_{(m)}$
2. Compare $P_{(i)}$ with $i \cdot q/m$
3. Let $r = \max\{i : P_{(i)} \leq i \cdot q/m\}$
4. Reject $H_{(1)} \cdots H_{(r)}$

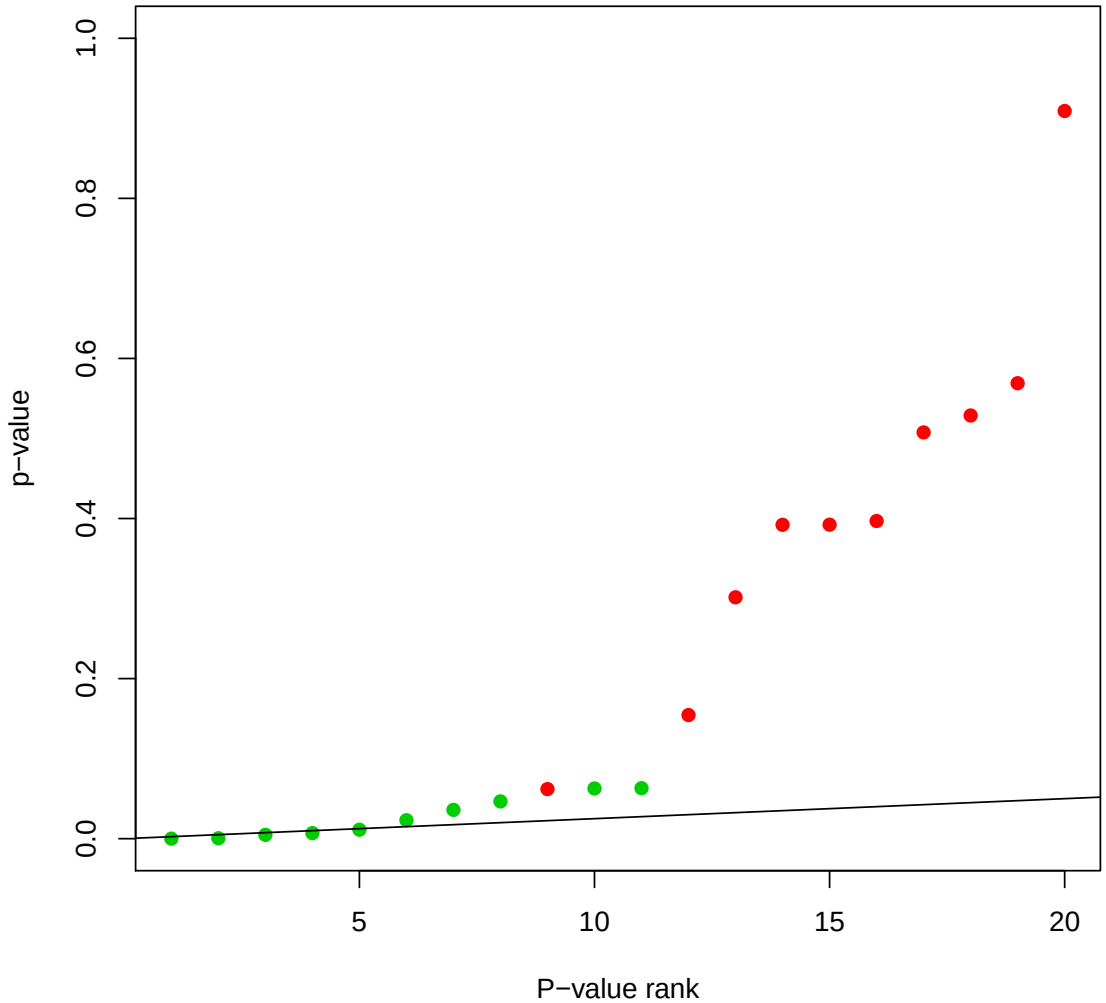
True null and False null p-values simulation

- Effect vector $\boldsymbol{\mu} = (\mu_1 \cdots \mu_m)$
- for $i = 1 \cdots m_0$, $\mu_i = 0$
- for $i = m_0 + 1 \cdots m$, $\mu_i = 3$
- Vector of effect Estimators $\boldsymbol{Z} = (Z_1 \cdots Z_m)$, $Z_i \overset{iid}{\sim} N(\mu_i, 1)$.
- P-value vector $\boldsymbol{P} = (P_1 \cdots P_m)$, $P_i = 1 - \Phi(Z_i)$,
  testing $H_{0i} : \ \theta_i = 0$ vs. $H_{1i} : \ \theta_i > 0$.

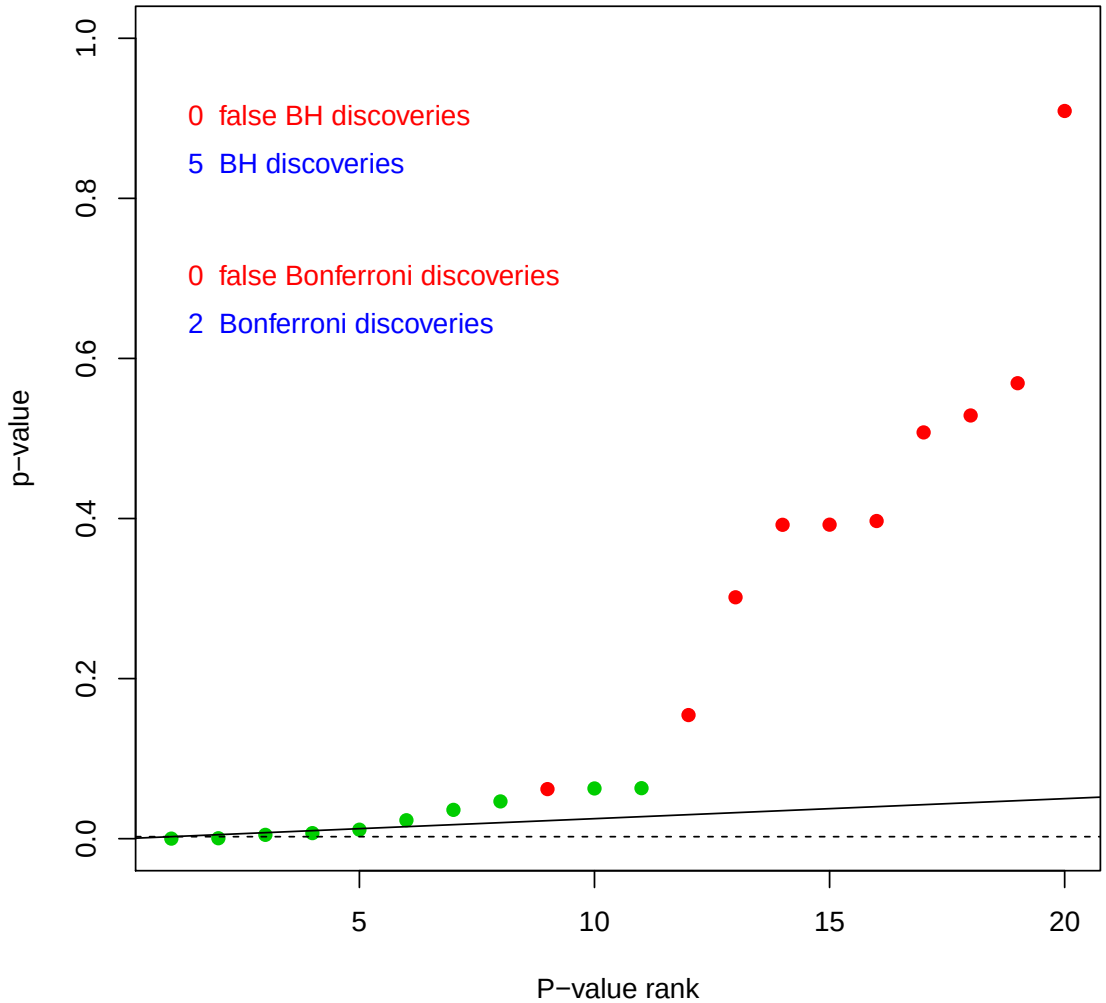
$P_1 \cdots P_{10}$ true null p-values, $P_{11} \cdots P_{20}$ false null p-values



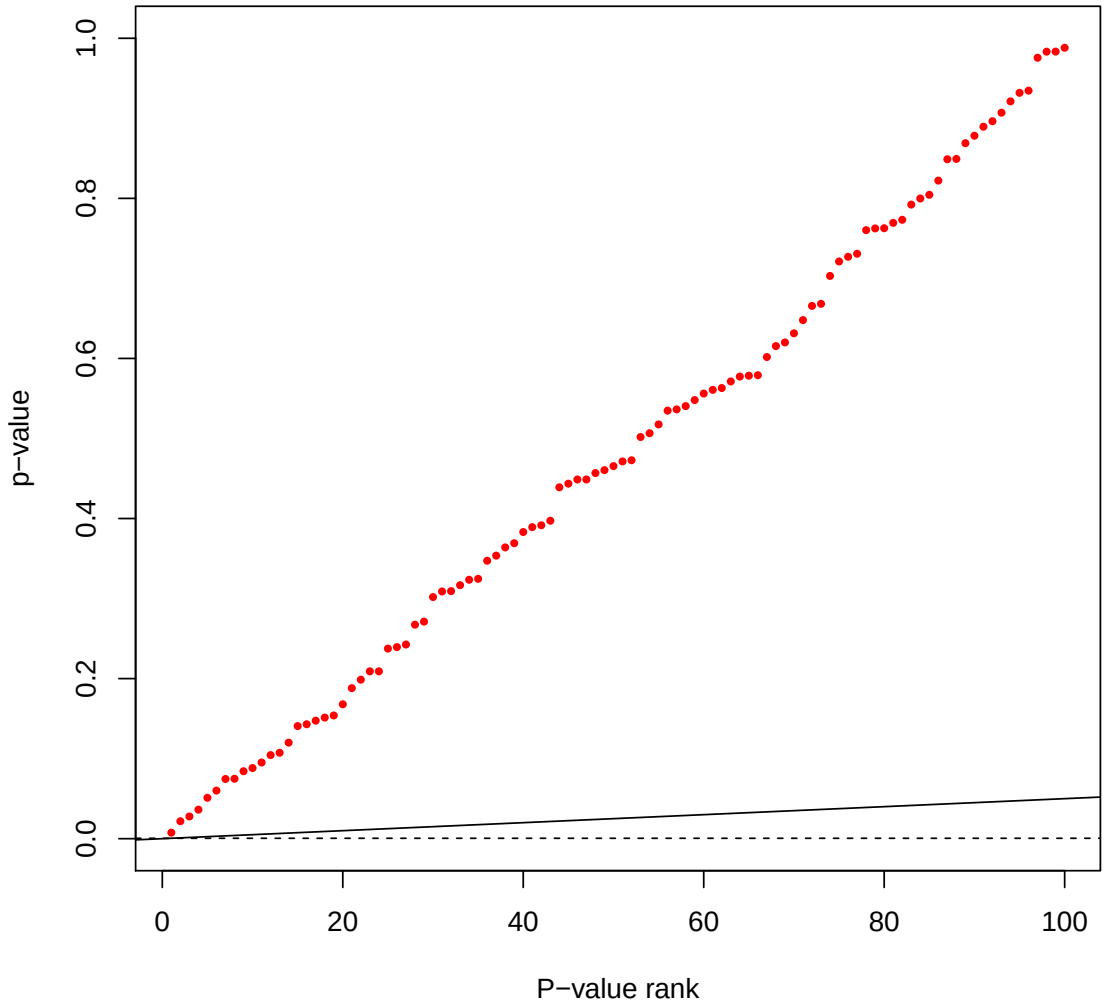
Level $q = 0.05$ BH procedure



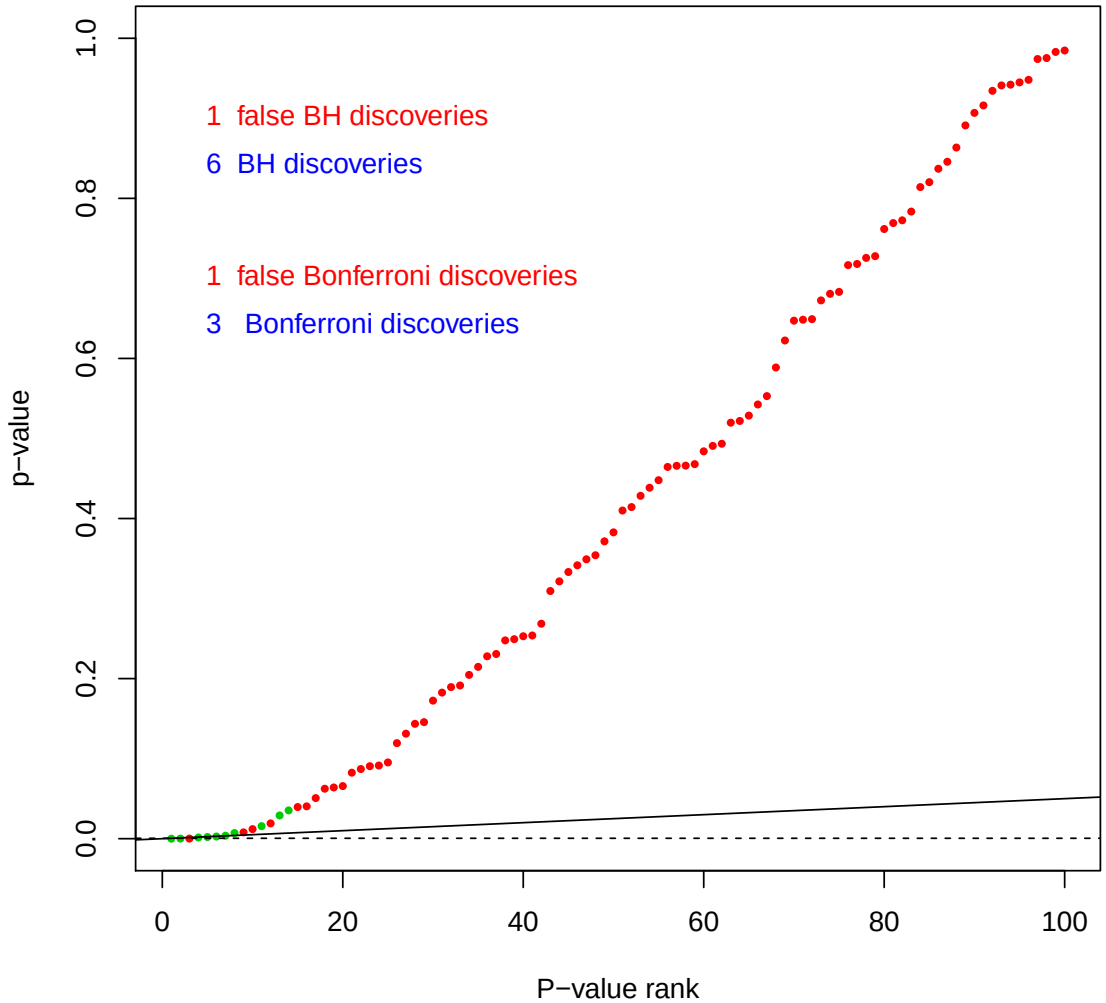
Compare with $\alpha = 0.05$ Bonferroni procedure



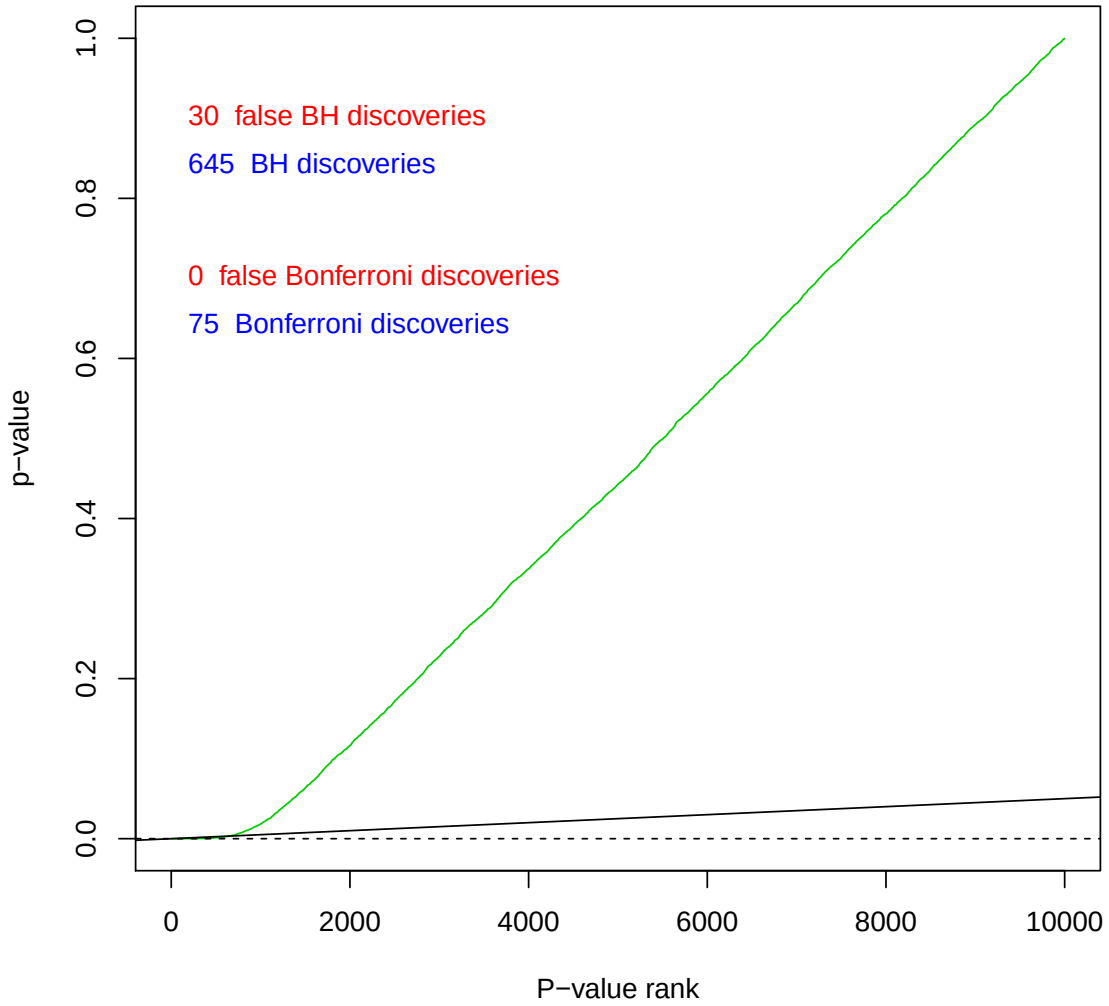
$P_1 \cdots P_{100}$ true null p-values



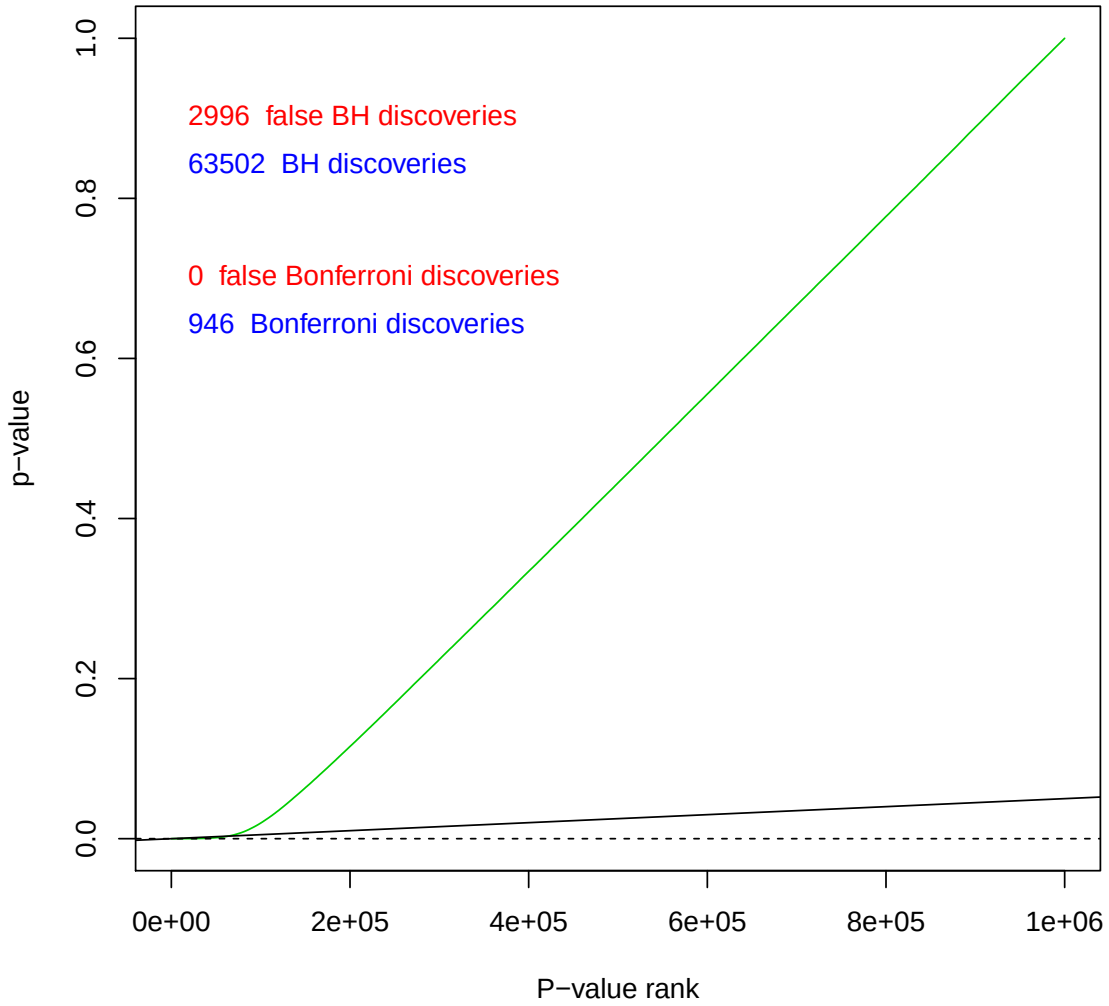
$P_1 \cdots P_{90}$ true nulls, $P_{91} \cdots P_{100}$ false nulls



Same simulation ($m_0/m = 0.9$) but now $m = 10^4$



Same simulation but now $m = 10^6$



Is the BH procedure new?

Simes '86

1. Sort the p-values $P_{(1)} \leq P_{2)} \leq \cdots \leq P_{(m)}$
2. Compare $P_{(i)}$ with $i \cdot q/m$
3. If $\exists i : P_{(i)} \leq i \cdot q/m$ then reject the global null that $H_1 \cdots H_m$ are true null hypotheses.

BH procedure $\rightarrow$ FDR control

BH '95:

- Independence $FDR \leq m_0 \cdot q/m$

Benjamini and Yekutieli '01:

- Independence $FDR = m_0 \cdot q/m$
- Positive dependence $FDR \leq m_0 \cdot q/m$
- Geneal dependence $FDR \leq (1 + 1/2 + \cdots + 1/m) \cdot m_0 \cdot q/m$

Simulations:

- Robustness to dependence

Interpretation of level 0.05 FDR control

All noise regime ($m_0 = m$)

- Any discovery is false – $FDR \equiv FWE$
- 0.95 probability of not making any false discovery

Signal and noise regime ($m_0 < m$)

- Many discoveries $\approx$ 0.05 false – $FDR < FWE$
- A randomly selected discovery is true with prob. 0.95

Selective vs. simultaneous inference

Benjamini and Yekutieli '05: two types of problems can arise when providing inferences for multiple parameters . . .

- *Selective inference* – need to provide valid marginal inferences for parameters that are selected after viewing the data (e.g. microarray analysis) – solution FDR control
- *Simultaneous inference* – need to provide valid inferences for all of the the parameters (e.g. subgroup analysis) – solution FWE control

Selective inference a new idea?

- Ioannidis, Plos Medicine ‘05
  “Why Most Published Research Findings Are False”

- Soric, JASA ‘89
  “ . . . It is mainly the discoveries that are reported and included into science . . . unless the proportion of false discoveries is kept small there is danger that a large part of science is untrue”

- Tukey and Scheffe ‘53
  post-hoc analysis

Post-hoc analysis – Scheffe's method

- $\boldsymbol{\mu} = (\mu_1 \cdots \mu_k)$ is a vector of k treatment effects, mean response in i'th treatment group is $\hat{\mu}_i \sim N(\mu_i, \sigma^2/n)$
- After viewing the data (and ANOVA) a contrast, $\boldsymbol{a\mu} = a_1\mu_1 + \cdots a_k\mu_k$ with $a_1 + \cdots + a_k = 0$, is selected

Selective inference problem: how do we use data to select $\boldsymbol{a\mu}$, test its significance, and construct a confidence interval for it?

Solution: base inference on confidence interval

$$CI_{Scheffe}(\boldsymbol{a\mu}, \alpha) := \boldsymbol{a\hat{\mu}} \pm \frac{\hat{\sigma} \cdot \|\boldsymbol{a}\|}{\sqrt{n}} \cdot \sqrt{(k-1) \cdot F_{1-\alpha,k-1,N-k}}$$

offering simultaneous coverage for all contrasts

$$\Pr_{\boldsymbol{\mu}}\{\forall \boldsymbol{a} : \ \boldsymbol{a\mu} \in CI_{Scheffe}(\boldsymbol{a\mu}, \alpha)\} \geq 1 - \alpha.$$

Post-hoc analysis – Tukey's method

Same setting, but instead of all contrasts
consider only pairwise comparisons ...

Solution: base inference on confidence interval

$$CI_{Tukey}(\mu_i - \mu_j, \alpha) := \hat{\mu}_i - \hat{\mu}_j \pm \frac{\hat{\sigma} \cdot 2}{\sqrt{n}} \cdot \sqrt{q_{1-\alpha,k-1,N-k}/2}$$

offering simultaneous coverage for all pairwise comparisons

$$\Pr_{\boldsymbol{\mu}}\{\forall i \neq j : \ \mu_i - \mu_j \in CI_{Tukey}(\mu_i - \mu_j, \alpha)\} \geq 1 - \alpha.$$

FDR control – same problem slightly different objective

- Post hoc analysis is concerned with valid inference for a single contrast (possibly the most significant contrast) that is specified according to the data
- Modern applications (microarrays / GWAS / fMRI / nonparametric regression) are concerned with marginal inferences for multiple parameters that are selected after first considering m pre-specified parameters
- And indeed ... Williams, Jones and Tukey '99 suggest using the BH procedure for discovering state-to-state (pairwise) differences in educational achievement.

Selective inference framework

Benjamini and Yekutieli '05:

- m parameters $\theta_1 \cdots \theta_m$ with corresponding estimators $T_1 \cdots T_m$
- There is a selection rule $\mathcal{S}(T_1 \cdots T_m) \subseteq \{1 \cdots m\}$
- Goal: to construct valid marginal confidence interval for the selected parameters: θ_i, $i \in \mathcal{S}(T_1 \cdots T_m)$

Continuous parameter-value simulation

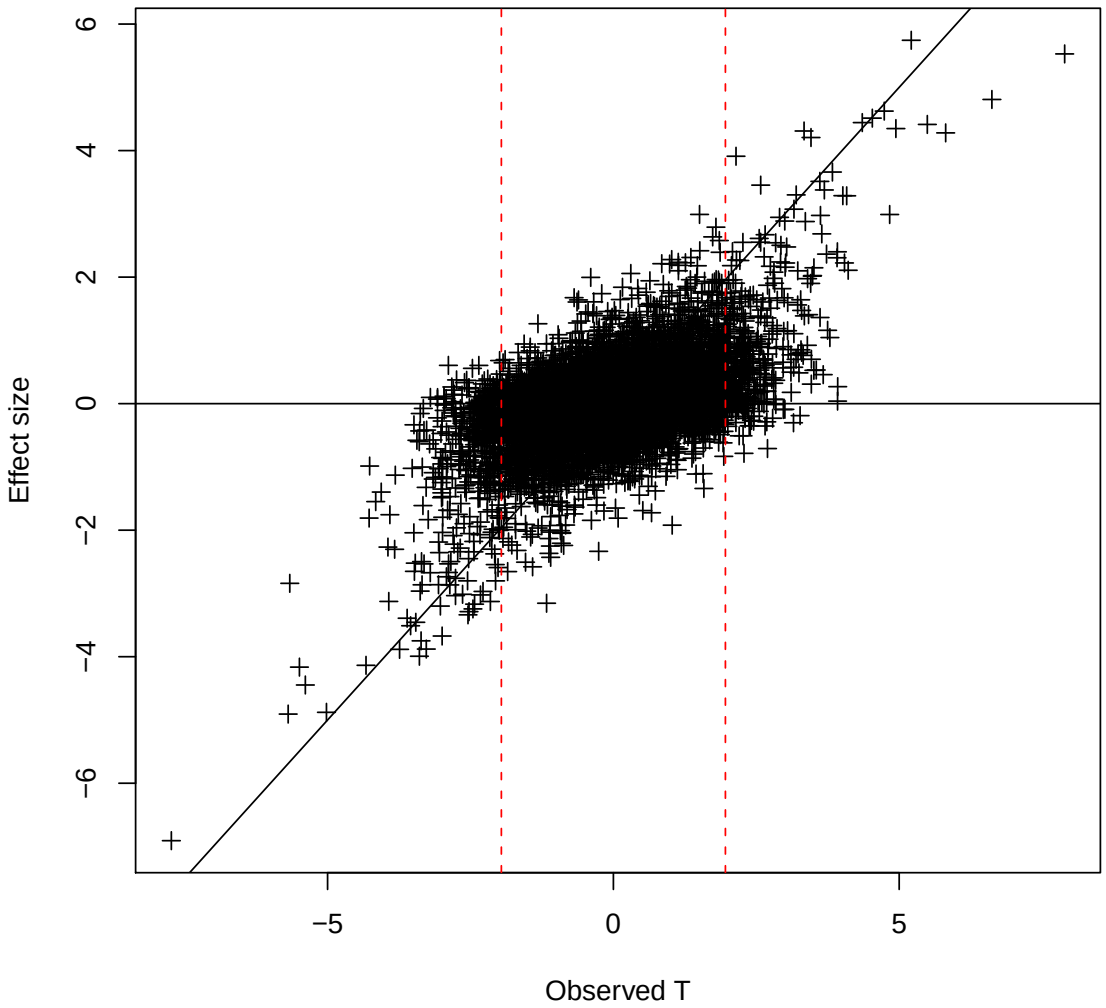
Generate $m = 10,000$ iid (θ_i, Y_i):

- Parameter $\theta_i \sim \pi(\theta_i)$

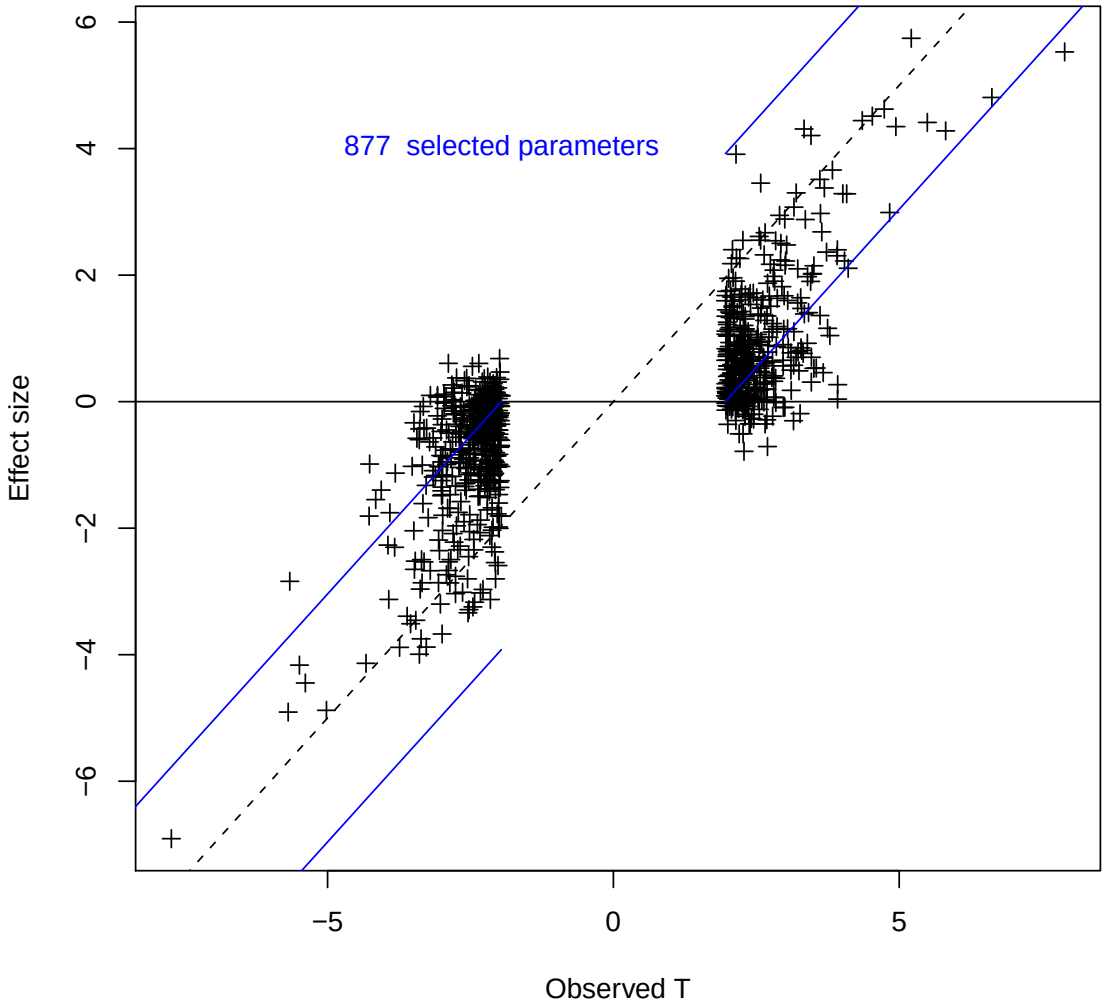
$$\pi(\theta_i) \;=\; 0.9 \cdot \frac{3 \cdot e^{-3 \cdot |\theta_i|}}{2} + 0.1 \cdot \frac{1 \cdot e^{-1 \cdot |\theta_i|}}{2} \tag{1}$$

- Observations $T_i \sim N(\theta_i, 1)$

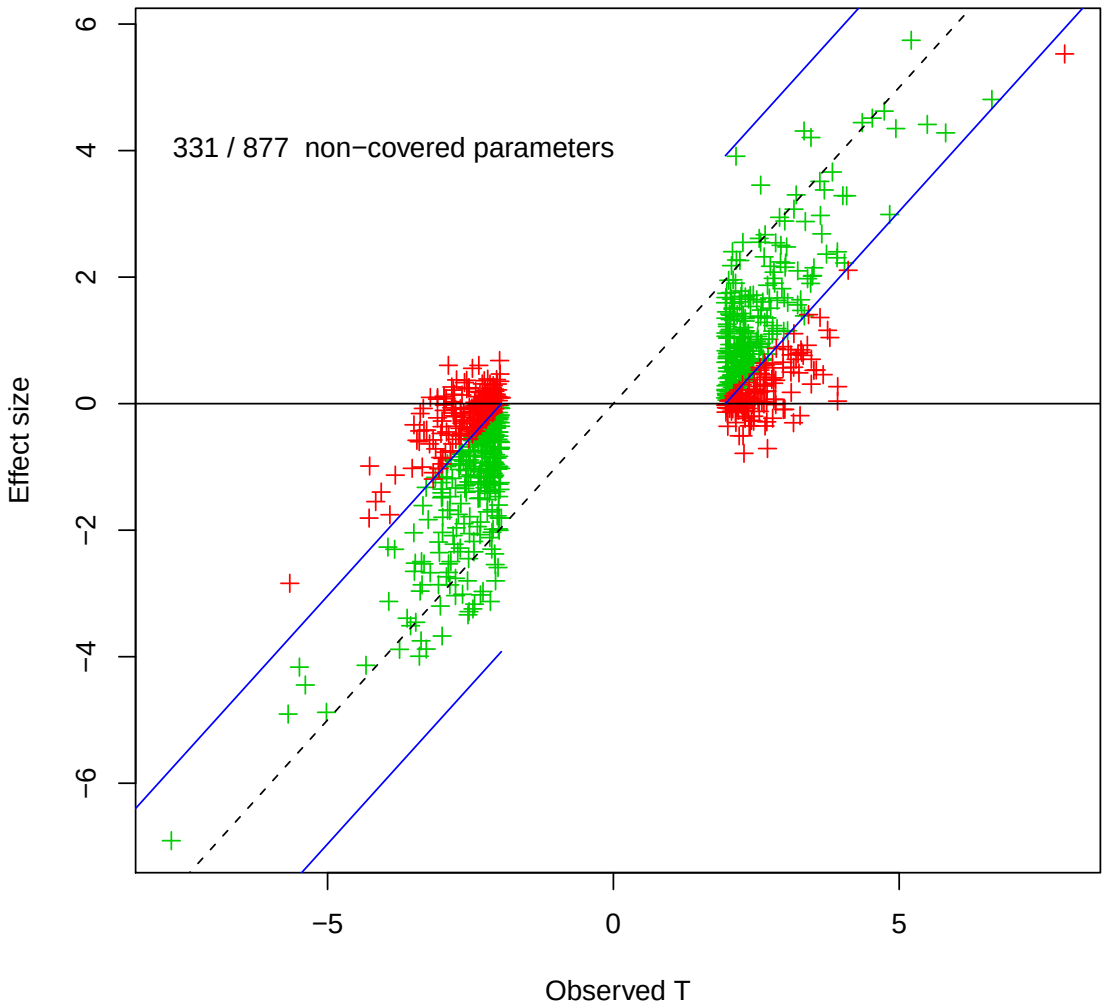
Entire data set



Marginal 0.95 CI's for selected parameters



CI's fail to cover 0.95 of the selected parameters



Valid marginal CI's for selected parameters

- Benjamini and Yekutieli '05 suggest the False Coverage-statement Rate as a measure for the validity of CI's constructed for the selected parameters

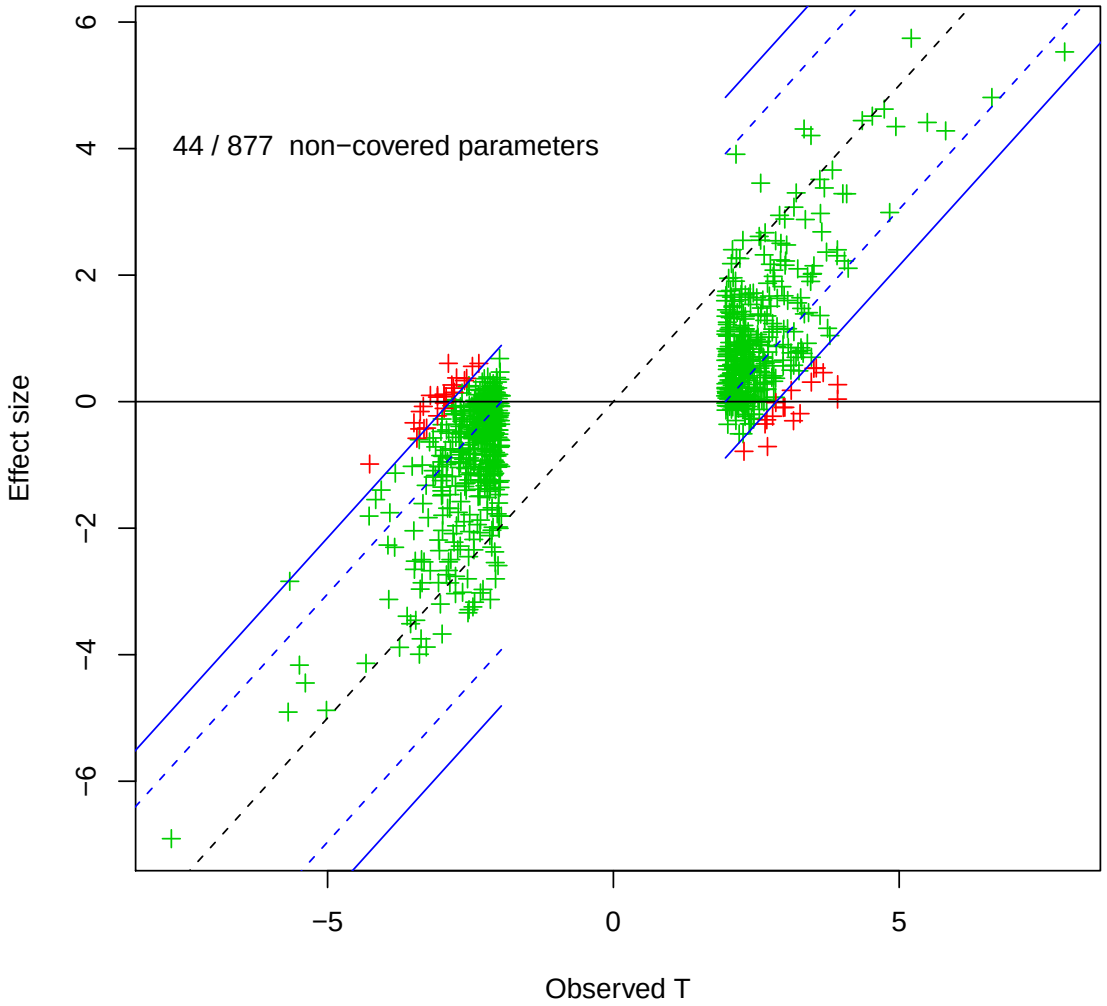
$$FCR = E\{V/\max(R,1)\}$$

where $R = |\mathcal{S}(T_1 \cdots T_m)|$ and V is the number of non-covering CI's

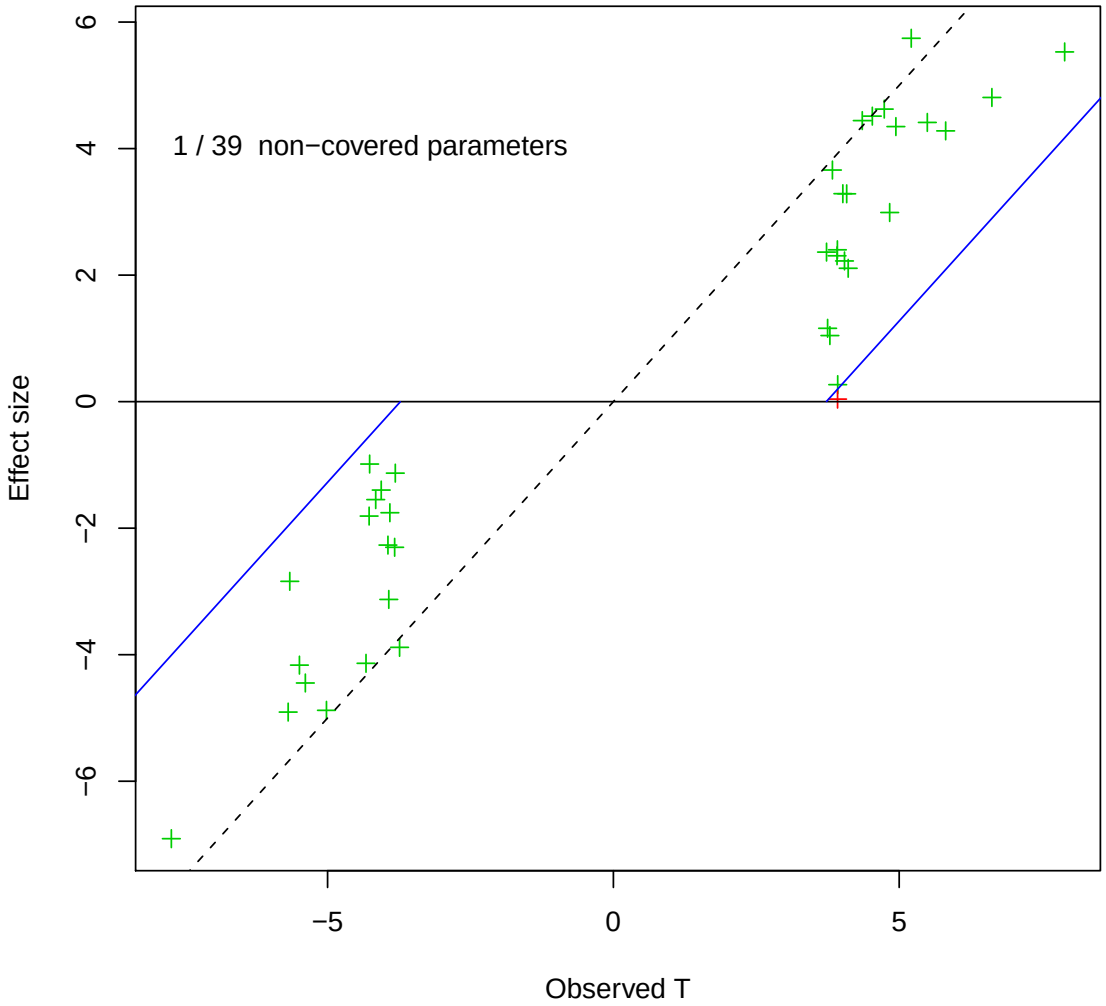
- Main result: for independent $\boldsymbol{T}$ and any selection rule $\mathcal{S}$ constructing marginal $1 - R \cdot q/m$ CI's for each selected parameter ensures $FCR \leq q$

(= *FCR adjusted CI's*)

FCR adjusted CI's for selected parameters



$\Rightarrow$ level 0.05 BH procedure



Characterizing the BH procedure by FCR adjusted CI's

BH procedure is most liberal testing procedure that ensures sign determination of all FCR-adjusted CI's

- Another BY '05 result: for independent $\boldsymbol{T}$ and any selection rule $\mathcal{S}$ that is at least as liberal as the BH procedure constructing FCR-adjusted CI's ensures $q/2 \leq FCR \leq q$

- Q: what is the FDR of the BH procedure in this example? and what is the directional FDR of the BH procedure? (Answers: 0, $\leq q$)

Summary

- The FDR/FCR are frequentist measures for validity of conditional statistical inference following selection
- the BH procedure provides is a method for making statistical discoveries that ensures that the discoveries are marginal true
- Next: a Bayesian perspective

The two group mixture model

Introduced in Efron et al. '01, see also Efron '10:

- Hypotheses vector $\boldsymbol{H} = (H_1, \cdots, H_m)$, statistic vector $\boldsymbol{Z} = (Z_1, \cdots, Z_m)$
- $H_i \overset{iid}{\sim} Bernouli(1-\pi_0)$
- $H_i = 0$ corresponds to a true null hypothesis, in which case $Z_i \sim f_0$ where usually $f_0 = N(0,1)$
- $H_i = 1$ corresponds to a false null hypothesis in which case $Z_i \sim f_1$
- If the statistics are p-values then $Z_i = \Phi^{-1}(P_i) \quad \rightarrow \quad f_0 = N(0,1)$

Multiple testing

- Multiple testing procedure

$$\mathcal{R}(\boldsymbol{Z}) = (R_1(\boldsymbol{Z}), \cdots, R_m(\boldsymbol{Z})) \in \{0,1\}^m$$

- $R_i = 1$ corresponds to declaring that $H_i = 1$, i.e. rejecting the null hypothesis that $H_i = 0$ and making a statistical discovery.

- $V_i = 0$ if the discovery is true, i.e. $R_i = 1 \wedge H_i = 1$

- $V_i = 1$ if the discovery is false (= type I error), i.e. $R_i = 1 \wedge H_i = 0$

- Number of discoveries $R = R_1 + \cdots + R_m$,
  number of false discoveries $V = V_1 + \cdots + V_m$

Natural measure for type I error rate?

For $m = 1$ consider a $\alpha = 0.05$ significance test: (= level 0.05 BH proc.)

$$R_1 \;=\; I(1.96 \leq |Z_1|)$$

first let's look at the BH '95 FDR

$$FDR = E_{\boldsymbol{Z}|\boldsymbol{H}}\{V/max(R,1)\}$$

- For $H_1 = 0$, $FDR = \Pr(R_1 = 0) = 0.05$ and for $H_1 = 1$, $FDR = 0$
- Thus for $\pi_0 = 0.90$ and $\Pr(R_1 = 1|H_1 = 1) = 0.50$

$$Pr(H_1 = 0|R_1 = 1) = \frac{0.90 \cdot 0.05}{0.90 \cdot 0.05 + 0.10 \cdot 0.50} = 0.474$$

Bayesian FDR

For "marginal" multiple testing procedures $R_i(\boldsymbol{Z}; T, \delta) = I\{T(Z_i) \leq \delta\}$

- Efron et al '01 (= pFDR in Storey '02-'03):

$$Fdr := Pr(H_i = 0 | R_i = 1) = \frac{Pr(H_i = 0, R_i = 1)}{Pr(R_i = 1)} = \frac{EV_i}{ER_i} = \frac{EV}{ER}$$

- As for $R > 0$, $V \sim Binom(R, Fdr)$:

$$Fdr = E_{\boldsymbol{H},\boldsymbol{Z}}(V/R | R > 0)$$

Relation between Bayesian FDR and the BH FDR

Expressing

$$\begin{aligned} FDR &= E_{\boldsymbol{Z}|\boldsymbol{H}}\{V/max(R,1), R>0\} + E_{\boldsymbol{Z}|\boldsymbol{H}}\{V/max(R,1), R=0\} \\ &= E_{\boldsymbol{Z}|\boldsymbol{H}}\{V/R|\ R>0\} \cdot \Pr_{\boldsymbol{Z}|\boldsymbol{H}}(R>0) \ + \ 0 \end{aligned}$$

Yields

$$Fdr = E_{\boldsymbol{H}}\{E_{\boldsymbol{Z}|\boldsymbol{H}}(V/R \mid R>0)\} = E_{\boldsymbol{H}}\{FDR \ / \ \Pr_{\boldsymbol{Z}|\boldsymbol{H}}(R>0)\}$$

Thus for large m if data is not pure noise (i.e. $\forall \boldsymbol{H},\ \Pr_{\boldsymbol{Z}|\boldsymbol{H}}(R>0)=1$)

$$Fdr \approx FDR$$

Bayes rule for classification in the two group mixture model

- The loss function for incorrectly classifying H_i is

$$L_{\mathcal{R},i}(\boldsymbol{H},\boldsymbol{Z}) = \lambda_1 \cdot I(H_i = 0, R_i(\boldsymbol{Z}) = 1) + \lambda_2 \cdot I(H_i = 1, R_i(\boldsymbol{Z}) = 0)$$

- The average risk is

$$\begin{aligned} & E_{\boldsymbol{H},\boldsymbol{Z}}\{\lambda_1 \cdot I(H_i = 0, R_i(\boldsymbol{Z}) = 1) + \lambda_2 \cdot I(H_i = 1, R_i(\boldsymbol{Z}) = 0)\} \\ = \;& E_{\boldsymbol{Z}}[E_{\boldsymbol{H}|\boldsymbol{Z}}\{\lambda_1 \cdot I(R_i(\boldsymbol{Z}) = 1) \cdot I(H_i = 0)\} \\ & \quad + E_{\boldsymbol{H}|\boldsymbol{Z}}\{\lambda_2 \cdot I(R_i(\boldsymbol{Z}) = 0) \cdot I(H_i = 1)\}] \\ = \;& E_{\boldsymbol{Z}}\{\underline{\lambda_1 \cdot I(R_i(\boldsymbol{Z}) = 1) \cdot \Pr(H_i = 0|\boldsymbol{Z})} \\ & \quad \underline{+ \lambda_2 \cdot I(R_i(\boldsymbol{Z}) = 0) \cdot \Pr(H_i = 1|\boldsymbol{Z})}\} \end{aligned}$$

- And $\forall \boldsymbol{z}$ the Bayes rule is $R_i(\boldsymbol{z})$ that minimizes the underlined expression

Bayes rule for classification .... (cont.)

- Note that for $j = 0, 1$

$$\begin{aligned}\Pr(H_i = j|\mathbf{Z}) &= \frac{f(\mathbf{Z}|H_i = j) \cdot \Pr(H_i = j)}{f(\mathbf{Z})} \\ &= \frac{\Pi_{k \neq i} f(Z_k) \cdot f(Z_i|H_i = j) \cdot \pi_j}{f(\mathbf{Z})} \propto f_j(Z_i) \cdot \pi_j\end{aligned}$$

- Thus the Bayes rule can be expressed as $R_i(\mathbf{z})$ minimizing

$$\lambda_1 \cdot I(R_i(\mathbf{z}) = 1) \cdot f_0(z_i) \cdot \pi_0 + \lambda_2 \cdot I(R_i(\mathbf{z}) = 0) \cdot f_1(z_i) \cdot \pi_1$$

which is

$$R_i(\mathbf{z}) = I\{ \frac{\pi_0 \cdot f_0(z_i)}{\pi_1 \cdot f_1(z_i)} \leq \frac{\lambda_2}{\lambda_1} \}$$

Bayes rule for classification .... (cont.)

- For

$$f(z_i) = \pi_0 \cdot f_0(z_i) + \pi_1 \cdot f_1(z_i)$$

defining the local FDR (Efron '01)

$$fdr(z_i) = \pi_0 \cdot f_0(z_i)/f(z_i) = \Pr(H_i = 0|z_i)$$

As

$$\frac{\pi_0 \cdot f_0(z_i)}{\pi_1 \cdot f_1(z_i)} = \frac{fdr(z_i)}{1 - fdr(z_i)}$$

- The Bayes rule for classifying H_i can also be expressed

$$R_i(z_i;\ fdr,\ \delta(\lambda_1, \lambda_2)) = I\{\ fdr(z_i) \leq \frac{\lambda_2}{\lambda_1 + \lambda_2}\}$$

$Fdr = q$ classification/selection/rejection/discovery rules

- For any $R_i(Z_i; T, \delta)$

$$\begin{aligned} Fdr = \Pr(H_i = 0 | R_i = 1) &= E_{Z_i|R_i=1} \Pr(H_i = 0|Z_i) \\ &= E_{Z_i|R_i=1} fdr(Z_i) \end{aligned}$$

- A $Fdr = q$ classification procedure is $R_i(Z_i; T, \delta(q))$ for which

$$Fdr = \Pr(H_i = 0 | R_i = 1) = q.$$

- In particular, the Bayes classifier can be specified by its Fdr level, instead of by λ_2 and λ_1. i.e. the $Fdr = q$ Bayes classifier is

$$R_i(Z_i;\ fdr,\ \delta(q)) = I\{\ fdr(z_i) \leq \delta(q)\}$$

with $Fdr = q$

The $Fdr = q$ Bayes classifier is optimal

Of all $R_i(Z_i; T, \delta)$ with $Fdr = q$, the Bayes classifier has

- Maximum power to make discoveries

$$\Pr(R_i = 1)$$

- and minimum type II error

$$Fnr = \Pr(H_i = 1 | R_i = 0)$$

(Storey '07; Sun and Cai '07; Efron '10; Heller and Yekutieli '13)

Is Bayesian FDR control valid?

Bearable sufficient scenario:

- The data and parameters are a single realization of $(\boldsymbol{H}, \boldsymbol{Z})$
- Regard (H_i, Z_i) for $i = 1 \cdots m$ exchangeable

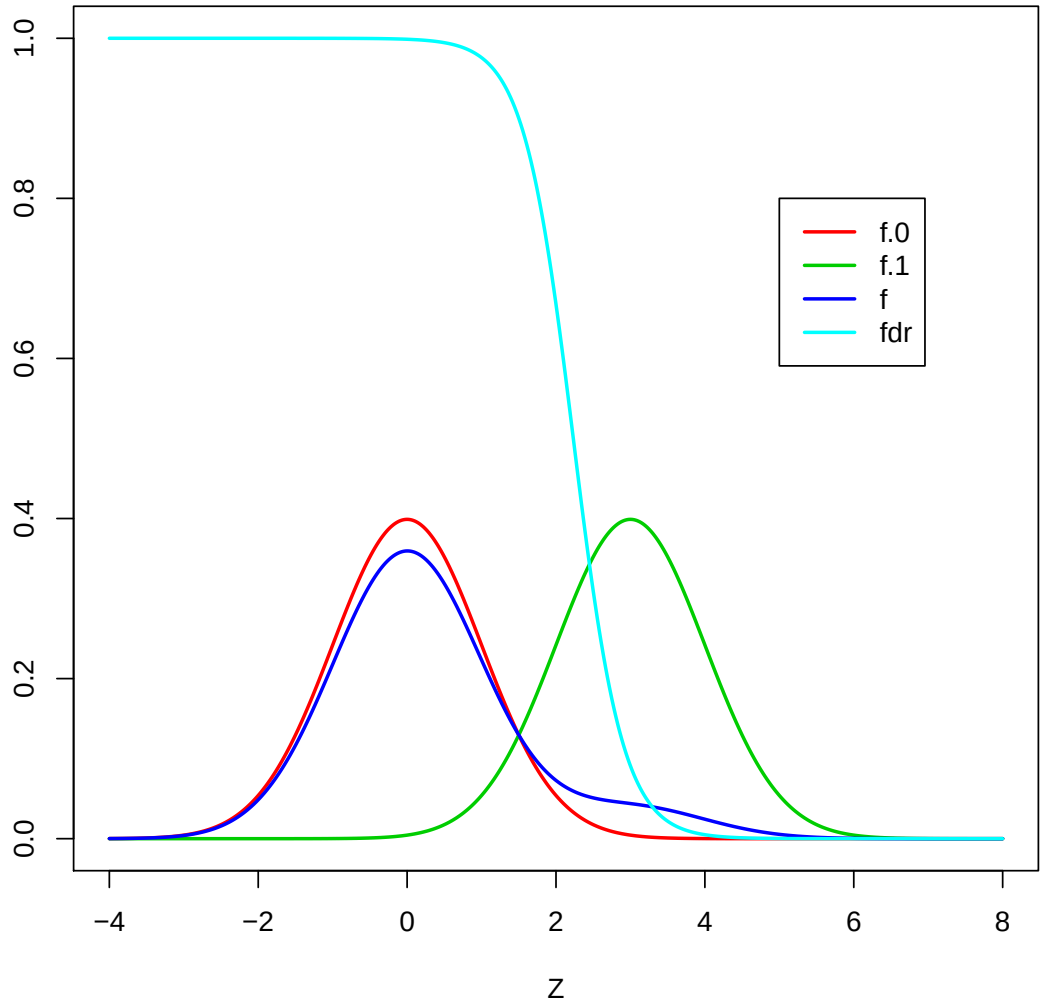
Is exchangeability necessary?

1. Finner: FDR control doesn't make sense without exchangeability
2. The two group model holds for a randomly chosen (H_i, Z_i) component

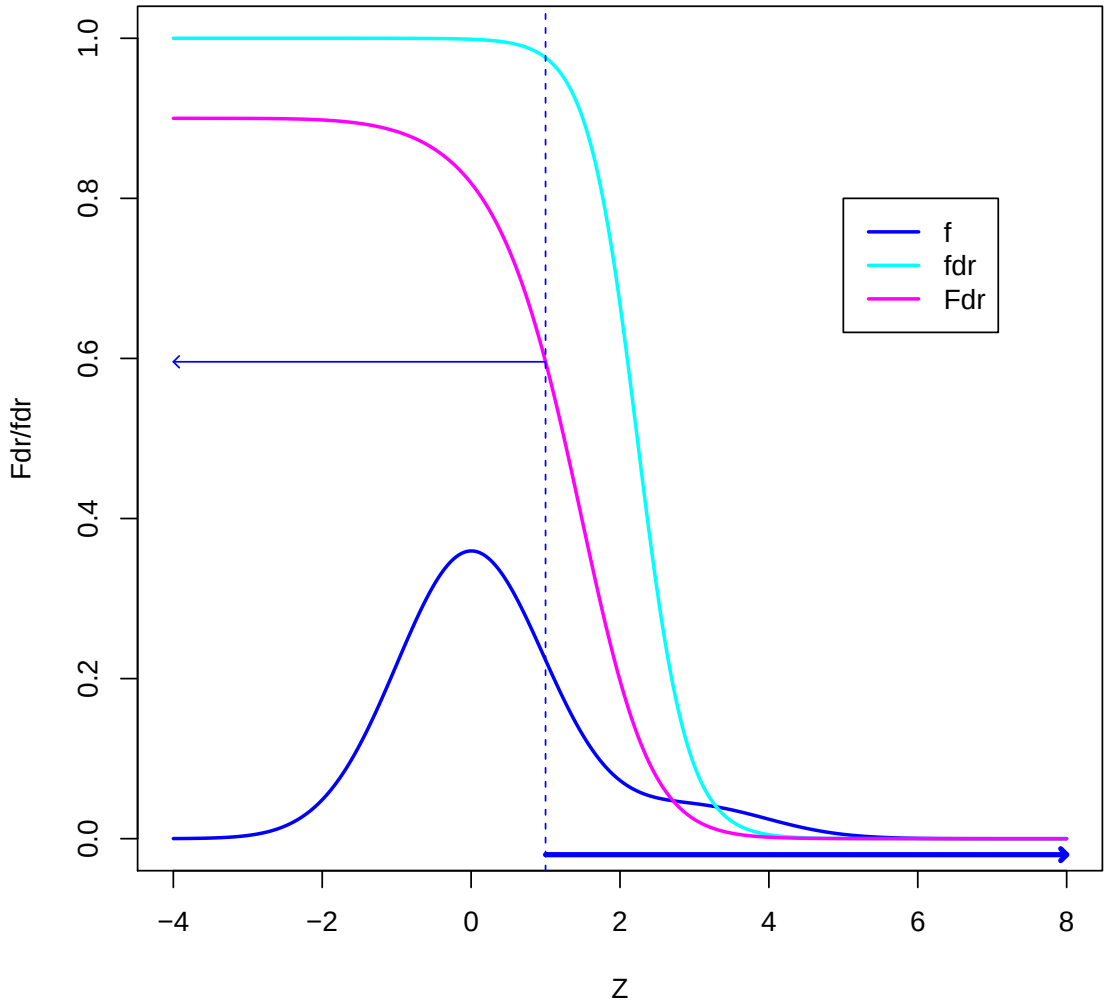
Return to True null and False null p-values simulation

- $m = 10^5$, $m_0/m = 0.9$
- for $i = 1 \cdots m_0$, $\mu_i = 0$
- for $i = m_0 + 1 \cdots m$, $\mu_i = 3$
- Vector of effect Estimators $\boldsymbol{Z} = (Z_1 \cdots Z_m)$, $Z_i \overset{iid}{\sim} N(\mu_i, 1)$.
- P-value vector $\boldsymbol{P} = (P_1 \cdots P_m)$, $P_i = 1 - \Phi(Z_i)$,
  testing $H_{0i} : \ \theta_i = 0$ vs. $H_{1i} : \ \theta_i > 0$.

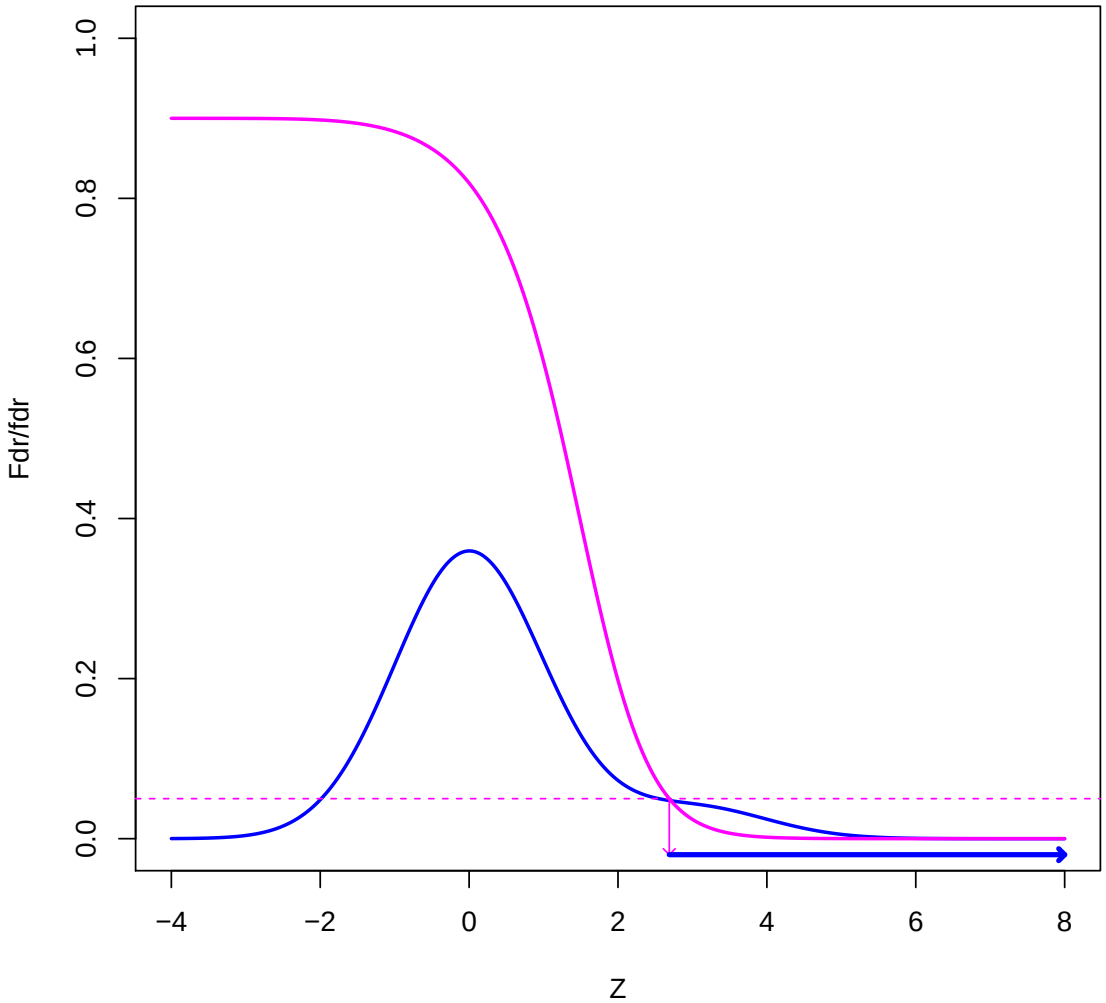
The densities of z_i and the *fdr*



$Fdr = 0.600$ for classifying rule $R_i = I(1 \leq Z_i)$



$R_i = I(2.689 \leq Z_i)$ is classification rule with $Fdr = 0.05$



How do we control Fdr?

With empirical Bayes:

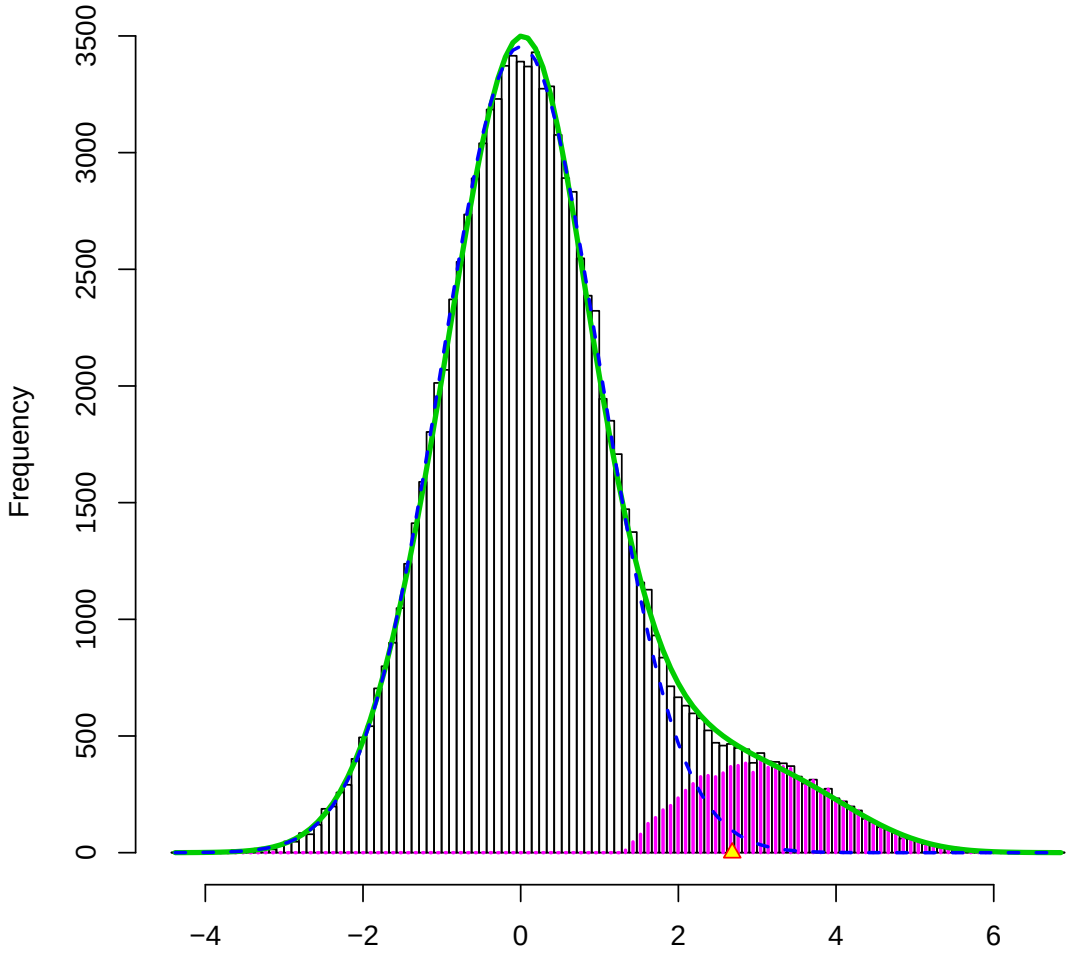
1. R *locfdr* package applied to $Z_1 \cdots Z_m$, that estimates π_0, f_0, f_1 and uses them to compute $Fdr(z)$, the Fdr of the rejection rule $R_i = I(z \leq Z_i)$
2. R *qvalue* package applied to $P_1 \cdots P_m$ to estimate $qvalue(p)$, the Fdr of the rejection rule $R_i = I(P_i \leq p)$, that equals

$$\Pr(H_i = 0 | P_i \leq p) = \frac{Pr(H_i = 0, P_i \leq p)}{Pr(P_i \leq p)} = \frac{Pr(P_i \leq p | H_i = 0) \cdot \pi_0}{Pr(P_i \leq p)}$$

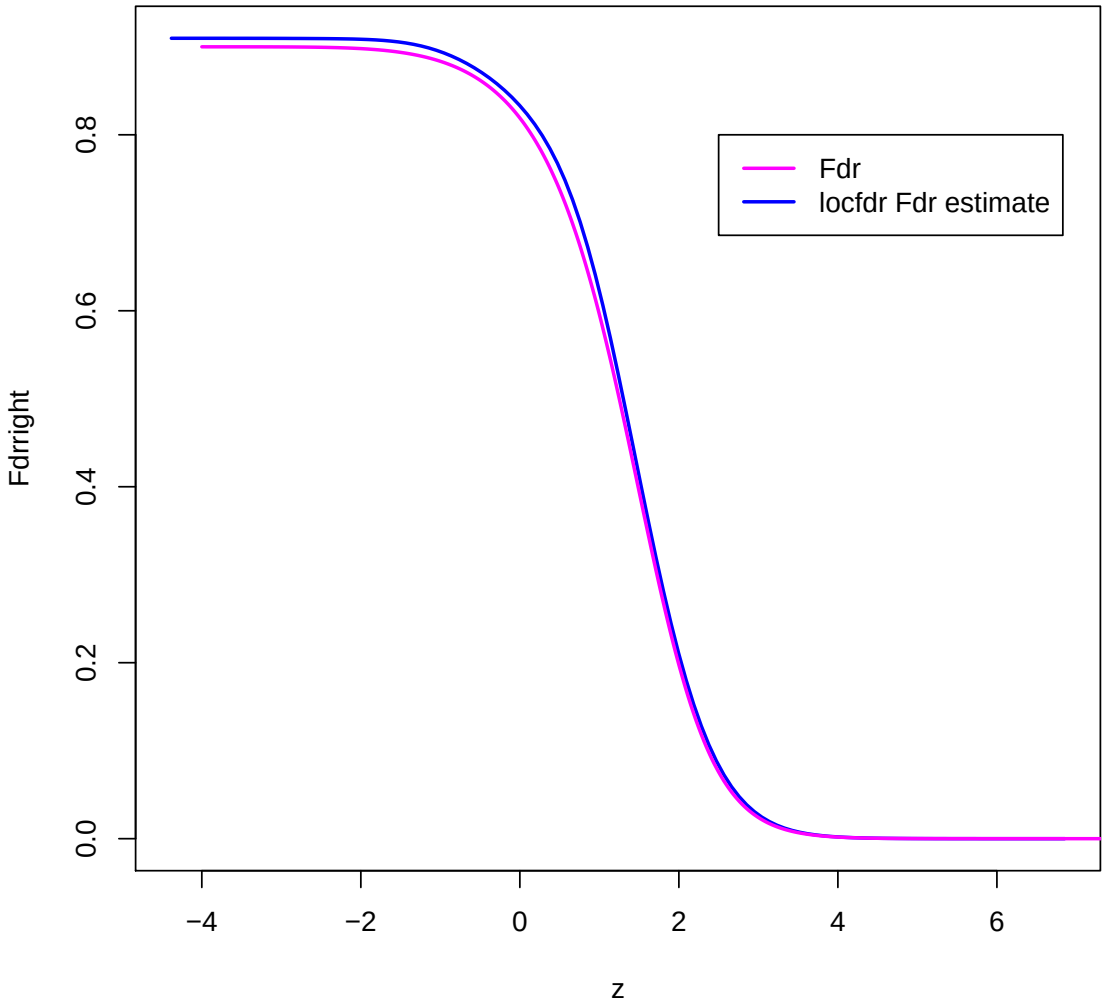
by

$$\widehat{qvalue}(p) = \frac{p \cdot \hat{\pi}_0}{\#\{i : \ P_i \leq p\}/m}, \quad with \ \ \hat{\pi}_0 = 2 \cdot \#\{i : \ 0.5 \leq P_i\}$$

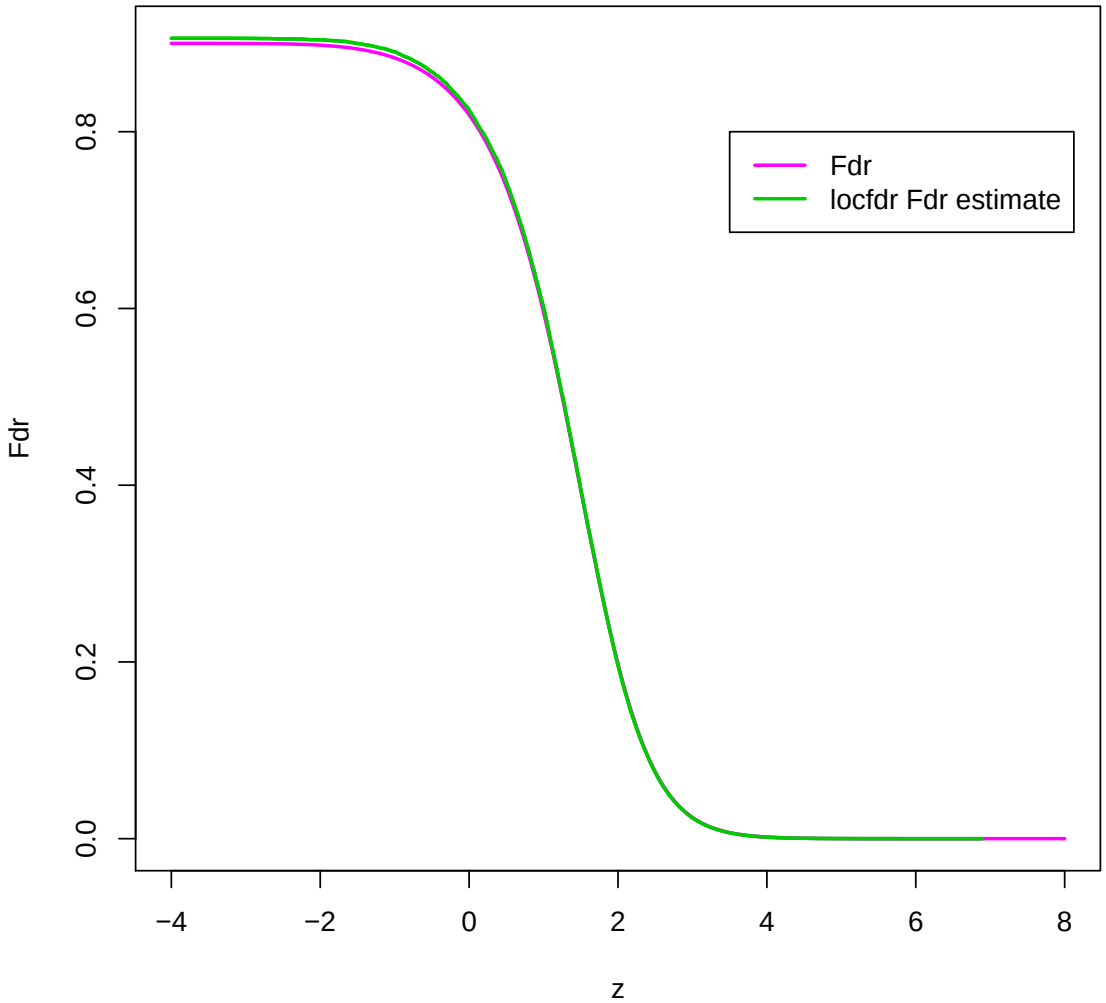
locfdr density estimates



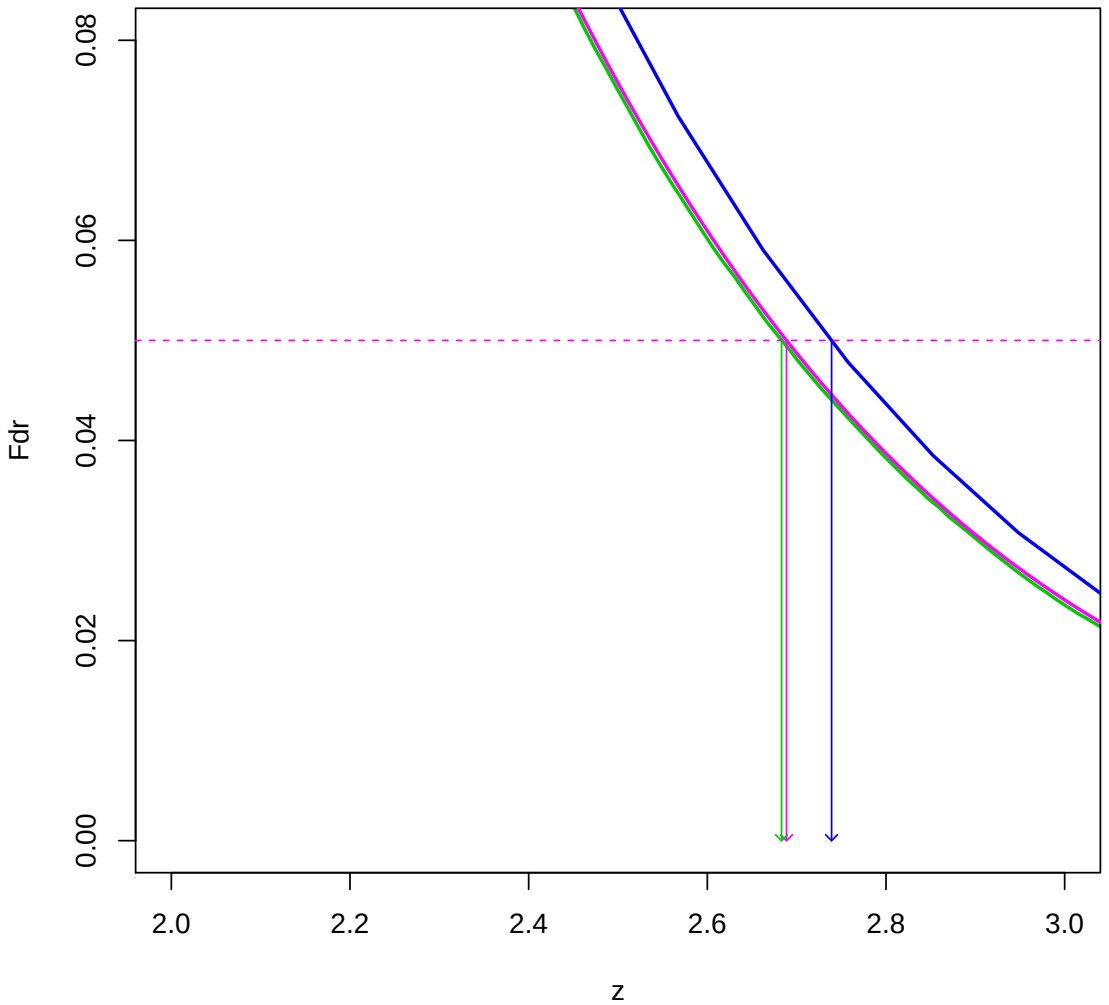
locfdr qvalue estimate



qvalue qvalue estimate



$Fdr = 0.05$ rejection rules



Results for $Fdr = 0.05$ rejection rules

- Bayes classifier:

$$z_{0.05} = 2.689 \;\rightarrow\; V/R = 338/6,568 = 0.0515$$

- Bayes classifier based on the *locfdr* estimate:

$$z_{0.05} = 2.739 \;\rightarrow\; V/R = 284/6,332 = 0.0445$$

- Bayes classifier based on the *qvalue* estimate:

$$z_{0.05} = 2.683 \;\rightarrow\; V/R = 346/6609 = 0.0524$$

Connection to BH procedure

From a Bayesian perspective, the BH procedure is a $Fdr = q$ classifier $R_i(Z_i; T, \hat{\delta}(q)) = I\{T(Z_i) \leq \hat{\delta}(q)\}$ for which

1. the test statistic is the p-value $T(Z_i) = P_i$
2. the critical value is $\hat{\delta}(z; q) = p_{(r)}$ is the level q BH critical value for which

$$\frac{p_{(r)}}{\#\{i:\ P_i \leq p_{(r)}\}/m} \ \leq \ q$$

Note that if we further apply the adaptive BH procedure that includes an estimate of π_0 we get the *qvalue* critical value:

$$\hat{p} \ \text{ such that } \ \frac{\hat{p} \cdot \hat{\pi}_0}{\#\{i:\ P_i \leq \hat{p}\}/m} \leq q$$

BH procedure results

We applied the BH procedure using the R *p.adjust* package

- Results for level $q = 0.05$ BH procedure:

$$z_{0.05} = 2.726 \;\rightarrow\; V/R = 300/6,402 = 0.0469$$

- $\hat{\pi}_0 = 2 \cdot \#\{i : 0.5 \leq P_i\}/100,000 = 0.901$
- Results for level $q = 0.05/0.901$ BH procedure:

$$z_{0.05} = 2.681 \;\rightarrow\; V/R = 348/6,616 = 0.0526$$

Some references

- Benjamini Y., Yekutieli D. (2005) "False Discovery Rate-Adjusted Multiple Confidence Intervals for Selected Parameters" *Journal of the American Statistical Association*, **100**, 71-81.
- Efron B., (2010) "Large-Scale Inference" Cambridge, United Kingdom
- Efron, B., Tibshirani, R., Storey, J. D. and Tusher, V. (2001) "Empirical Bayes analysis of a microarray experiment". *Journal of the American Statistical Association*, **96**, 1151-1160.
- Storey J. D. (2002) "A direct approach to false discovery rates" *Journal of the Royal Statistical Society: Series B*, **64** 479-498.
- Storey J. D., (2003) "The positive false discovery rate: A Bayesian interpretation and the q-value" *Annals of Statistics*, **31**, 2013-2035.