Hommel's Method for False Discovery Proportions

Jelle Goeman

Joint work with: Aldo Solari, Rosa Meijer

Van Dantzig, 2016-02-26

Data analysis in genomics

Top differential expression

Gene	p-value
XDH	5.5e-10
NEK3	6.7e-7
TAF5	7.1e-7
CYP2A7	1.6e-6
NAT2	1.8e-6
ZNF19	2.6e-6
SKP1	2.7e-6
NAT1	3.1e-6
GDF3	2.0e-5
CCDC25	2.1e-5
⋮	⋮

Familywise error control

95% conf.: no false positives

Data analysis in genomics

Top differential expression

Gene	p-value
XDH	5.5e-10
NEK3	6.7e-7
TAF5	7.1e-7
CYP2A7	1.6e-6
NAT2	1.8e-6
ZNF19	2.6e-6
SKP1	2.7e-6
NAT1	3.1e-6
GDF3	2.0e-5
CCDC25	2.1e-5
⋮	⋮

Familywise error control

95% conf.: no false positives

False discovery rate control

Expected prop. of false positives $< 5\%$

Data analysis in genomics

Top differential expression

Gene	p-value
XDH	5.5e-10
NEK3	6.7e-7
TAF5	7.1e-7
CYP2A7	1.6e-6
NAT2	1.8e-6
ZNF19	2.6e-6
SKP1	2.7e-6
NAT1	3.1e-6
GDF3	2.0e-5
CCDC25	2.1e-5
⋮	⋮

Familywise error control
95% conf.: no false positives

False discovery rate control
Expected prop. of false positives $< 5\%$

Practice
Genes chosen for validation

Data analysis in genomics

Top differential expression

Gene	p-value
XDH	5.5e-10
NEK3	6.7e-7
TAF5	7.1e-7
CYP2A7	1.6e-6
NAT2	1.8e-6
ZNF19	2.6e-6
SKP1	2.7e-6
NAT1	3.1e-6
GDF3	2.0e-5
CCDC25	2.1e-5
⋮	⋮

Familywise error control
95% conf.: no false positives

False discovery rate control
Expected prop. of false positives $< 5\%$

Practice
Genes chosen for validation

Question
How many false positives to expect?

Set-up

Hypotheses

$H_1, \ldots, H_m$

True hypotheses

$T \subseteq \{1, \ldots, m\}$ indices of true hypotheses

Rejections

$R \subseteq \{1, \ldots, m\}$ set of rejected hypotheses (usually random)

Type I errors

$T \cap R \subseteq \{1, \ldots, m\}$

FWER, FDR, k-FWER

User role

Before seeing the data choose error rate to be controlled

$$\text{FWER: } P(T \cap R \neq \emptyset) \qquad \text{FDR: } E\Big(\frac{\#(T \cap R)}{\#R \vee 1}\Big)$$

Procedure

Chooses R that controls the chosen error rate

FWER, FDR, k-FWER

User role

Before seeing the data choose error rate to be controlled

$$\text{FWER: } P(T \cap R \neq \emptyset) \qquad \text{FDR: } E\left(\frac{\#(T \cap R)}{\#R \vee 1}\right)$$

Procedure

Chooses R that controls the chosen error rate

Problem

- R is often too small or too large
- R based on p-values only
- "Take it or leave it"

Alterative: simultaneous control

Role of the user

The user selects collection of hypotheses R freely and post hoc

Alterative: simultaneous control

Role of the user
The user selects collection of hypotheses R freely and post hoc

Role of the multiple testing procedure
Inform user of the number/proportion of false rejections incurred

Alterative: simultaneous control

Role of the user
The user selects collection of hypotheses R freely and post hoc

Role of the multiple testing procedure
Inform user of the number/proportion of false rejections incurred

Number of false rejections
$= \#(T \cap R)$
$=$ function of the model parameters
$=$ something we can estimate or make a confidence interval for

Alterative: simultaneous control

Role of the user
The user selects collection of hypotheses R freely and post hoc

Role of the multiple testing procedure
Inform user of the number/proportion of false rejections incurred

Number of false rejections
$= \#(T \cap R)$
$=$ function of the model parameters
$=$ something we can estimate or make a confidence interval for

Post hoc
If we make a simultaneous CI, post hoc choice of R is allowed

Closed Testing: ingredients

Marcus, Peritz and Gabriel (1976)
Fundamental principle of FWER control

Intersection hypothesis
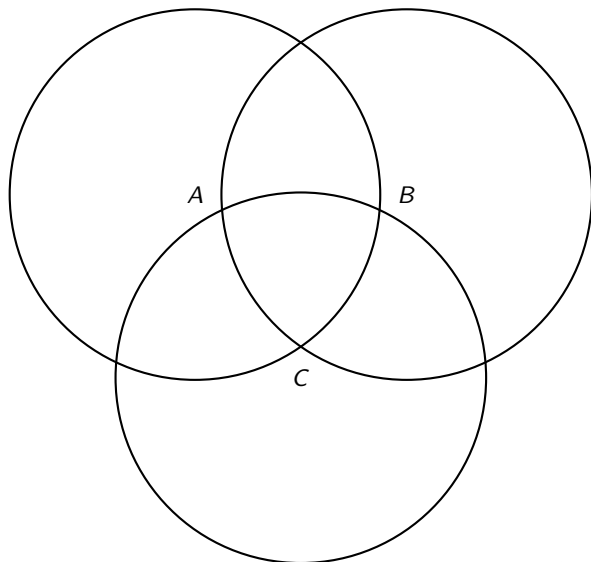$H_C = \bigcap_{i \in C} H_i$, for $C \subseteq \{1, \ldots, m\}$

Closure
Collection of all intersection hypotheses
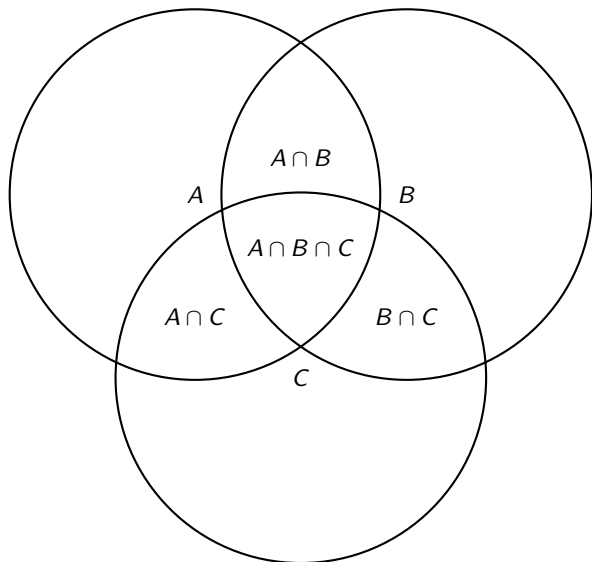$\mathcal{C} = \{H_C : C \subseteq \{1, \ldots, m\}\}$

Local test
Valid α-level test for every intersection hypothesis

Closed testing (graphically)

Closed testing (graphically)

Closed testing: procedure

Raw rejections

Hypotheses $\mathcal{U} \subseteq \mathcal{C}$ rejected by the local test

Multiplicity-rejected rejections

Reject $H \in \mathcal{C}$ if $J \in \mathcal{U}$ for every $J \subseteq H$

Closed testing: procedure

Raw rejections
Hypotheses $\mathcal{U} \subseteq \mathcal{C}$ rejected by the local test

Multiplicity-rejected rejections
Reject $H \in \mathcal{C}$ if $J \in \mathcal{U}$ for every $J \subseteq H$

Statement

$$\mathrm{P}(\mathcal{R} \cap \mathcal{T} = \emptyset) \geq 1 - \alpha$$

with $\mathcal{R} = \{C \in \mathcal{C} : C \text{ rejected}\}$ and $\mathcal{T} = \{C \in \mathcal{C} : C \text{ true}\}$

Closed testing: procedure

Raw rejections
Hypotheses $\mathcal{U} \subseteq \mathcal{C}$ rejected by the local test

Multiplicity-rejected rejections
Reject $H \in \mathcal{C}$ if $J \in \mathcal{U}$ for every $J \subseteq H$

Statement

$$\mathrm{P}(\mathcal{R} \cap \mathcal{T} = \emptyset) \geq 1 - \alpha$$
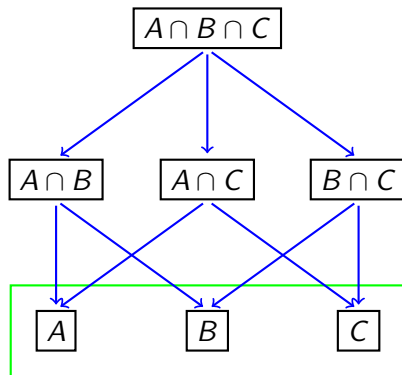
with $\mathcal{R} = \{C \in \mathcal{C} : C \text{ rejected}\}$ and $\mathcal{T} = \{C \in \mathcal{C} : C \text{ true}\}$

Proof

$$\{\mathcal{R} \cap \mathcal{T} = \emptyset\} \supseteq \{H_{\mathcal{T}} \notin \mathcal{U}\}$$
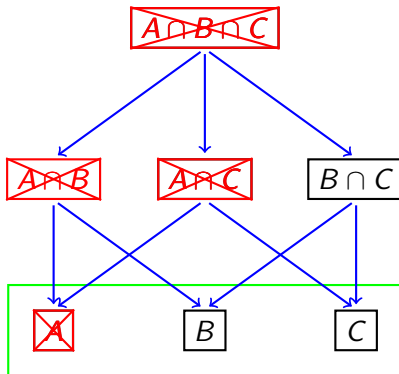
Consonance

Traditionally, only rejection of elementary hypotheses is of interest

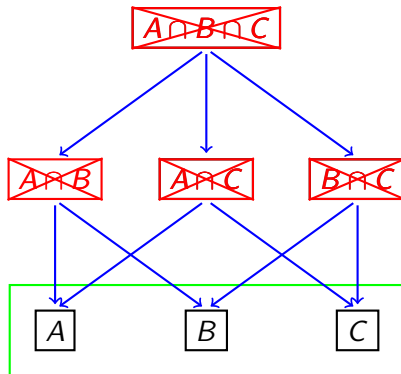The closed graph of hypotheses A, B and C

Consonance

Traditionally, only rejection of elementary hypotheses is of interest

Consonant rejections

Consonance

Traditionally, only rejection of elementary hypotheses is of interest

Non-consonant rejections of $A \cap B$, $A \cap C$, $B \cap C$

Parameter, confidence bound and coverage

Parameter
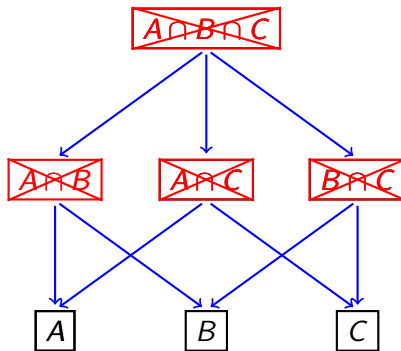
$\tau(R) = \#(T \cap R)$ for a fixed set R

Closed testing

Let $\mathcal{X}$ be the collection of hypotheses rejected

Confidence bound

$t_\alpha(R) = \max(\#C : C \subseteq R, H_C \notin \mathcal{X}\}$

In the example

$t_\alpha(\{A, B, C\}) = 1$

Coverage

Coverage statement

$\mathrm{P}(\tau(R) \leq t_\alpha(R)) \geq 1 - \alpha$

Coverage

Coverage statement

$\mathrm{P}(\tau(R) \leq t_\alpha(R)) \geq 1 - \alpha$

Proof

$\{\tau(R) \leq t_\alpha(R)\} \subseteq \{H_\mathcal{T} \notin \mathcal{U}\}$

Coverage

Coverage statement

$\mathrm{P}(\tau(R) \leq t_\alpha(R)) \geq 1 - \alpha$

Proof

$\{\tau(R) \leq t_\alpha(R)\} \subseteq \{H_T \notin \mathcal{U}\}$

Confidence set

- Trivial lower bound $\tau(R) \geq 0$: confidence set $\{0, \ldots, t_\alpha(R)\}$
- Confidence set for $\phi(R) = \#R - \tau(R)$ immediate
- Confidence set for $\mathrm{FDP} = \phi(R)/\#R$ immediate

Coverage

Coverage statement
$\mathrm{P}(\tau(R) \le t_\alpha(R)) \ge 1 - \alpha$

Proof
$\{\tau(R) \le t_\alpha(R)\} \subseteq \{H_T \notin \mathcal{U}\}$

Confidence set
- Trivial lower bound $\tau(R) \ge 0$: confidence set $\{0, \dots, t_\alpha(R)\}$
- Confidence set for $\phi(R) = \#R - \tau(R)$ immediate
- Confidence set for FDP $= \phi(R)/\#R$ immediate

Simultaneous control over all R
Consequence: coverage robust against post hoc selection of R

Reject hypotheses

R	confidence set for $\tau(R)$	confidence set for $\phi(R)$
$\{A\}$	$\{0,1\}$	$\{0,1\}$
$\{B\}$	$\{0,1\}$	$\{0,1\}$
$\{C\}$	$\{0,1\}$	$\{0,1\}$
$\{A,B\}$	$\{0,1\}$	$\{1,2\}$
$\{A,C\}$	$\{0,1\}$	$\{1,2\}$
$\{B,C\}$	$\{0,1\}$	$\{1,2\}$
$\{A,B,C\}$	$\{0,1\}$	$\{2,3\}$

Bonus: an estimate

Point estimate of FDP
Take confidence bound at $\alpha = 1/2$

Property (immediate)
FDP overestimated at most with probability 0.5

Reporting (classical!)
FDP estimate and confidence bound

Single hypothesis
Estimated false if $p < 0.5$; confidently false if $p < 0.05$

Shortcuts

General

Procedure can be used for any local test

Number of intersection hypotheses

$2^m - 1$: computationally prohibitive above $\approx$20 hypotheses
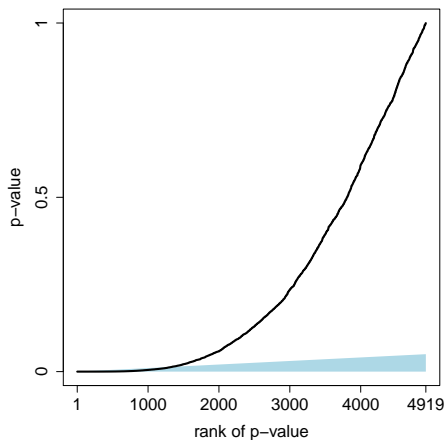
Concept: shortcut

Smart choice of local test to save calculations

Smart choice of local test

Also crucial for the power properties of the procedure

Simes' inequality

Sorted *p*-value curve and lower confidence bound

Local test based on Simes' inequality

Simes' inequality

With probability $\geq 1 - \alpha$, we have $p_{(i:T)} > \frac{i\alpha}{\#T}$ for all
$i = 1, \ldots, \#T$.
where $p_{(i:I)}$ is the Ith smallest p-value among p_i, $i \in I$.

Use Simes as local test

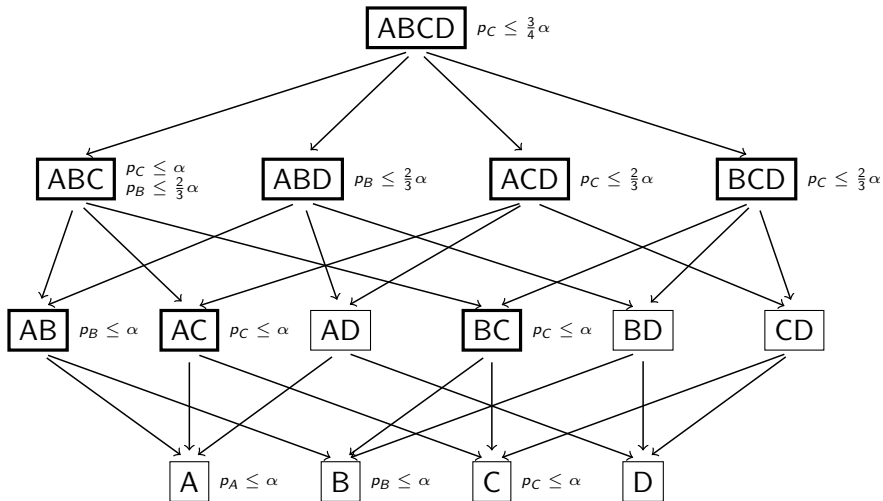Reject if any $p_{(i:I)} \leq \frac{i\alpha}{\#I}$

Assumptions (Sarkar, Yekutieli and others)

Generally assumed valid for two-sided asymptotically normal tests

Variant without assumptions (conservative)

Reject if any $p_{(i)} \leq \frac{i\alpha}{kb(k)}$ with $b(k) = \sum_{s=1}^{k} 1/s$

Example: $\alpha/2 < p_A \leq p_B \leq p_C \leq 2\alpha/3$ **and** $p_D > \alpha$

Shortcut

Lemma

H_I is rejected in closed testing with Simes local tests at level α iff there is an $i \in I$ such that

$$p_{(i:I)} \leq \frac{i\alpha}{j(\alpha)}$$

Crucial quantity $j(\alpha)$

All H_I with $|I| > j(\alpha)$ are rejected
At least one H_I with $|I| = j(\alpha)$ is not rejected

$$j(\alpha) = \max\{s \in 1, \ldots, m : p_{(m-s+k)} > k\alpha/s \text{ for } k = 1, \ldots, s\}$$

Calculating $j(\alpha)$

Steps of $j(\alpha)$

$j(\alpha)$ jumps from s to $s - 1$ $(s = 1, \ldots, m)$ at

$$\alpha_s = \min_{j=1,\ldots,s} \frac{s \cdot p_{m-s+k}}{k}$$

Naive calculation of α_s, $s = 1, \ldots, m$

Order m^2 steps

Use lemma (next slide)

Reduce to $m \log(m)$ steps

Calculating α_s

To calculate

$\alpha_m, \ldots, \alpha_1$ are minima of columns of matrix

$$
M = \begin{pmatrix}
p_1 & & & \\
p_2/2 & p_2 & & \\
p_3/3 & p_3/2 & p_3 & \\
\vdots & \vdots & & \ddots \\
p_m/m & p_m/(m-1) & & \cdots & p_m
\end{pmatrix}
$$

Lemma

Row location of minimum is non-increasing

Find minima in $m \log(m)$ time

By starting in the middle column

Calculating the confidence bound $t_\alpha(R)$

Category

Find category $c_i = \left\lceil \frac{j(\alpha)}{\alpha} p_i \right\rceil$ for all $i \in R$

Then $(1 - \alpha)$ confidence lower bound for $\tau(R)$

$t_\alpha(R) = \#R - \max_{r=1,\dots,\#R} \left\{ 1 - r + \#\{c_i \le r\} \right\}$

Computation

Linear complexity

Relationship with Hommel

Hommel's procedure

- FWER control
- Uniformly better than Hochberg's procedure
- Also based on closed testing plus Simes

Relationship with Hommel

R rejected by Hommel $\rightarrow$ bound $t_\alpha(R) = 0$

Improvements

- Better bounds by exploiting non-consonant rejections
- Faster algorithm (order $m \log(m)$ instead of classical m^2

Relationship with Benjamini/Hochberg

Assumptions

$\approx$ same assumptions; same weak FWER control

Lemma

Let R with $r = |R|$ and $mp_{(r:R)}/r = q \leq \alpha$. Then $t_\alpha(R)/r \leq \frac{j(\alpha)q}{m\alpha}$.

Colloquially

Set R with maximal FDR-corrected p-value q has
$(1 - \alpha)$-confidence of FDP $\leq q/\alpha$

Consequences

- FDR rejected set R has FDP estimate < 0.10
- FDR rejected set R has $(1 - \alpha)$-confidence of FDP < 1

Scalability

Assume non-vanishing alternative

$\# T/m \to \mathrm{const} < 1$ as $m \to \infty$

FWER methods as $m \to \infty$: not scalable

- Rejected set $\to \emptyset$
- Adjusted p-values $\to 1$

FDR methods as $m \to \infty$: scalable (under condition)

- Rejected set R has $\#R \to \mathrm{const} > 0$
- Adjusted p-value $\tilde{p}_{(cm)} \to \mathrm{const} < 1$

If FDR scales, FDP confidence scales too

$\exists R$ with $\#R/m \to \mathrm{const} > 0$ so that $t_\alpha(R)/\#R \to \mathrm{const} < 1$

Data analysis in genomics

Top differential expression

Gene	p-value
XDH	5.5e-10
NEK3	6.7e-7
TAF5	7.1e-7
CYP2A7	1.6e-6
NAT2	1.8e-6
ZNF19	2.6e-6
SKP1	2.7e-6
NAT1	3.1e-6
GDF3	2.0e-5
CCDC25	2.1e-5
⋮	⋮

How many false positives to expect?

95% conf.: max. 1 false positive

Estimated number of false positives

no false positives

Data analysis in genomics

Top differential expression

Gene	p-value
XDH	5.5e-10
NEK3	6.7e-7
TAF5	7.1e-7
CYP2A7	1.6e-6
NAT2	1.8e-6
ZNF19	2.6e-6
SKP1	2.7e-6
NAT1	3.1e-6
GDF3	2.0e-5
CCDC25	2.1e-5
⋮	⋮

How many false positives to expect?

95% conf.: max. 1 false positive

Estimated number of false positives

no false positives

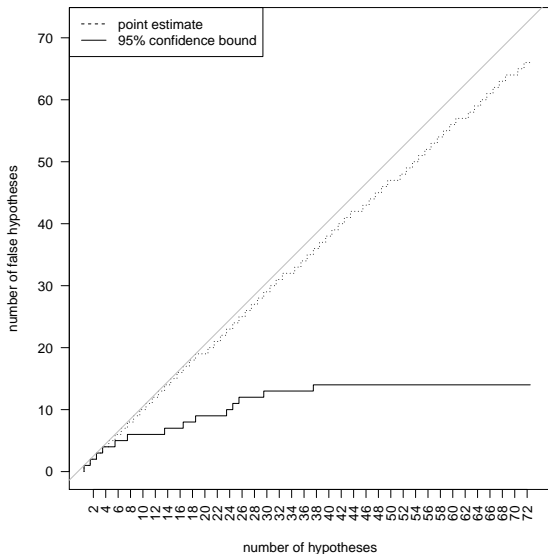# Example: Rosenwald DLBCL data

Data
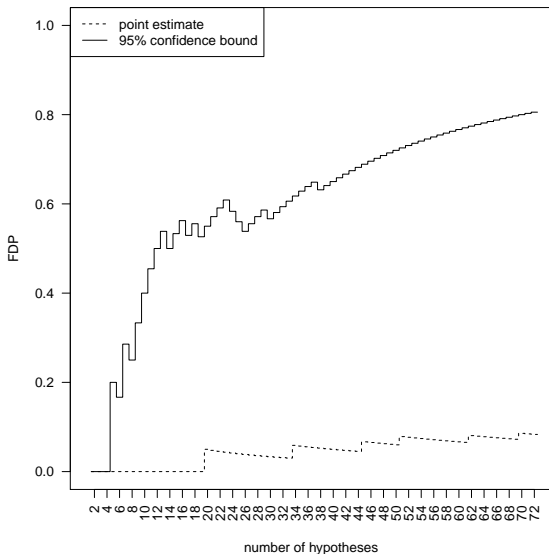
240 diffuse large B-cell lymphoma patients; 7399 hypotheses

Classical results

- Bonferroni, Holm, Hocherg, Hommel: 4 hypotheses

- Benjamini and Hochberg: 72 hypotheses

false hypotheses among top k p-values

FDP estimates and bounds: top k p-values

number of hypotheses

Conclusion

New method

Between weak and strong FWER control
Counting false positives: tail probabilities for FDP

Nothing new

Just closed testing and simultaneous confidence sets
But free additional statements relative to classical Hommel

Fast algorithms

Reduced from exponential to $m \log(m)$ complexity
Side effect: fast algorithm for Hommel's procedure

Simultaneous but still scalable

Rejections don't vanish when $m \to \infty$

Read more?

📄 Goeman JJ and Solari A (2011)
Multiple Testing for Exploratory Research.
Statistical Science 26:584–597 and 608–612

📄 Goeman JJ and Solari A (2014)
Tutorial in Biostatistics: Multiple Hypothesis Testing in Genomics.
Statistics in Medicine, 23 (11) 1946–1978

📄 Meijer RJ, Krebs T, Solari A and Goeman JJ (2015)
Extending Hommel's method
In preparation

🌐 Goeman JJ, Solari A, Meijer RJ
cherry R package
`cran.r-project.org`