Optimal and maximin procedures for multiple testing problems

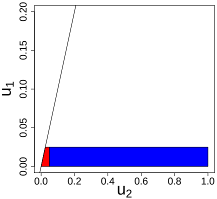
Saharon Rosset
Tel Aviv University

With: Ruth Heller, Amichai Painsky, Ehud Aharoni.

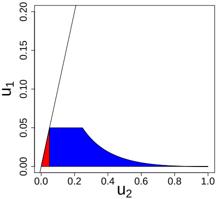
arxiv.org/abs/1804.10256
arxiv.org/abs/1902.00892

Two normal means, FWER control

Bonferroni-Holm

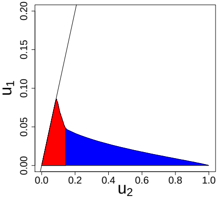


Closed testing using Stouffer 49

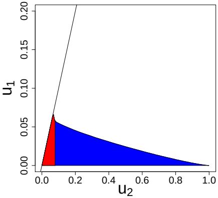


Optimal multiple test (OMT) for two false nulls:

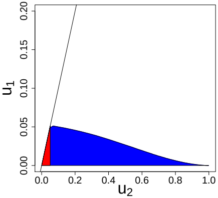
$\theta_0 = -0.5$



$\theta_0 = -1$



$\theta_0 = -2$



Hypothesis testing basics

Given some data X we want to test:

$$H_0: \; X \sim F_0$$
$$H_A: \; X \sim F_A$$

Assume F_0 and F_A have density f_0, f_a respectively, then Neyman-Pearson (NP) Lemma says that a most powerful (MP) test rejects H_0 at x iff $f_a(x)/f_0(x) \geq c$.

Different formulation in terms of *p-value*: We transform using the distribution of the likelihood ratio to get:

$$H_0: \; U = H(X) \sim U(0,1)$$
$$H_A: \; U \sim G$$

and G has density $g(u)$ that is a decreasing function. Now NP says MP test at level α rejects H_0 iff $U \leq \alpha$.

Hypothesis testing basics

Given some data X we want to test:

$$H_0: \; X \sim F_0$$
$$H_A: \; X \sim F_A$$

Assume F_0 and F_A have density f_0, f_a respectively, then Neyman-Pearson (NP) Lemma says that a most powerful (MP) test rejects H_0 at x iff $f_a(x)/f_0(x) \geq c$.

Different formulation in terms of *p-value*: We transform using the distribution of the likelihood ratio to get:

$$H_0: \; U = H(X) \sim U(0,1)$$
$$H_A: \; U \sim G$$

and G has density $g(u)$ that is a decreasing function. Now NP says MP test at level α rejects H_0 iff $U \leq \alpha$.

Most powerful tests as an optimization problem

We can think of the MP problem as an optimization problem on an infinite set of variables:

$$\max_{D:[0,1]\to\{0,1\}} \quad \int_0^1 D(u)g(u)du$$
$$\text{s.t.} \quad \int_0^1 D(u)du \le \alpha$$

This (integer, infinite) problem happens to have the simple solution structure implied by the NP Lemma (basically a continuous knapsack problem), because it has just one constraint.

Most powerful tests as an optimization problem

We can think of the MP problem as an optimization problem on an infinite set of variables:

$$\max_{D:[0,1]\to\{0,1\}} \quad \int_0^1 D(u)g(u)du$$
$$\text{s.t.} \quad \int_0^1 D(u)du \leq \alpha$$

This (integer, infinite) problem happens to have the simple solution structure implied by the NP Lemma (basically a continuous knapsack problem), because it has just one constraint.

Moving to multiple testing setup

In a multiple testing problem, we are given K pairs of hypotheses:

$$H_{0k}: \; U_k \sim U(0,1)$$
$$H_{Ak}: \; U_k \sim G$$

(assume for now all alternatives are the same).

In the paper we deal with (exchangeable) dependence, here we also assume U_j, U_k are independent for $j \neq k$.

We seek to design good tests that give *high power* while *controlling type-I error (level)*.

Moving to multiple testing setup

In a multiple testing problem, we are given K pairs of hypotheses:

$$H_{0k}: \ U_k \sim U(0,1)$$
$$H_{Ak}: \ U_k \sim G$$

(assume for now all alternatives are the same).

In the paper we deal with (exchangeable) dependence, here we also assume U_j, U_k are independent for $j \neq k$.

We seek to design good tests that give *high power* while *controlling type-I error (level)*.

Moving to multiple testing setup

In a multiple testing problem, we are given K pairs of hypotheses:

$$H_{0k}: \; U_k \sim U(0,1)$$
$$H_{Ak}: \; U_k \sim G$$

(assume for now all alternatives are the same).

In the paper we deal with (exchangeable) dependence, here we also assume U_j, U_k are independent for $j \neq k$.

We seek to design good tests that give *high power* while *controlling type-I error (level)*.

Some notation

$h \in \{0,1\}^K$ is the true state of all hypotheses:
$h_k = 1 \ \Leftrightarrow \ H_{Ak}$ holds.

$D : [0,1]^K \to \{0,1\}^K$ is the decision function:
Rejects H_{0k} at $u \in [0,1]^K \ \Leftrightarrow \ D_k(u) = 1$.

$R(D)(u) = \sum_{k=1}^{K} D(u)$ is the number of rejected nulls at u according to D.
$V(D)(u) = \sum_{k=1,h_k=0}^{K} D(u)$ is the number of type-I errors at u according to D.

We only consider *symmetric* D functions: $\sigma(D(u)) = D(\sigma(u))$ for any permutation σ.

Some notation

$h \in \{0,1\}^K$ is the true state of all hypotheses:
$h_k = 1 \Leftrightarrow H_{Ak}$ holds.

$D : [0,1]^K \to \{0,1\}^K$ is the decision function:
Rejects H_{0k} at $u \in [0,1]^K \Leftrightarrow D_k(u) = 1$.

$R(D)(u) = \sum_{k=1}^{K} D(u)$ is the number of rejected nulls at u according to D.
$V(D)(u) = \sum_{k=1, h_k=0}^{K} D(u)$ is the number of type-I errors at u according to D.

We only consider *symmetric* D functions: $\sigma(D(u)) = D(\sigma(u))$ for any permutation σ.

Some notation

$h \in \{0,1\}^K$ is the true state of all hypotheses:
$h_k = 1 \Leftrightarrow H_{Ak}$ holds.

$D : [0,1]^K \to \{0,1\}^K$ is the decision function:
Rejects H_{0k} at $u \in [0,1]^K \Leftrightarrow D_k(u) = 1$.

$R(D)(u) = \sum_{k=1}^K D(u)$ is the number of rejected nulls at u according to D.
$V(D)(u) = \sum_{k=1,h_k=0}^K D(u)$ is the number of type-I errors at u according to D.

We only consider *symmetric* D functions: $\sigma(D(u)) = D(\sigma(u))$ for any permutation σ.

Some notation

$h \in \{0,1\}^K$ is the true state of all hypotheses:
$h_k = 1 \;\Leftrightarrow\; H_{Ak}$ holds.

$D : [0,1]^K \to \{0,1\}^K$ is the decision function:
Rejects H_{0k} at $u \in [0,1]^K \;\Leftrightarrow\; D_k(u) = 1$.

$R(D)(u) = \sum_{k=1}^{K} D(u)$ is the number of rejected nulls at u according to D.
$V(D)(u) = \sum_{k=1, h_k=0}^{K} D(u)$ is the number of type-I errors at u according to D.

We only consider *symmetric* D functions: $\sigma(D(u)) = D(\sigma(u))$ for any permutation σ.

Generalizations of power and level

The best known notions of type-I error for multiple testing:

$$FWER = \mathbb{P}(V > 0) = \mathbb{P}\left((1-h)^t D(U) > 0\right),$$
$$FDR = \mathbb{E}\frac{V}{R} = \mathbb{E}\frac{(1-h)^t D(U)}{1^t D(U)}.$$

Popular generalized notions of power we consider:

Average power for L false nulls:

$$\Pi_L(D) = \frac{1}{L}\int_{[0,1]^K}\left(\sum_{l=1}^{L} D_l(u)\right)\prod_{l=1}^{L} g(u_l)du$$

Minimal power for K false nulls:

$$\Pi_{any}(D) = \int_{[0,1]^K}\mathbb{I}\left\{\sum_{l=1}^{K} D_l(u) > 0\right\}\prod_{l=1}^{K} g(u_l)du$$

Generalizations of power and level

The best known notions of type-I error for multiple testing:

$$FWER = \mathbb{P}(V > 0) = \mathbb{P}\left((1-h)^t D(U) > 0\right),$$
$$FDR = \mathbb{E}\frac{V}{R} = \mathbb{E}\frac{(1-h)^t D(U)}{1^t D(U)}.$$

Popular generalized notions of power we consider:

Average power for L false nulls:

$$\Pi_L(D) = \frac{1}{L}\int_{[0,1]^K}\left(\sum_{l=1}^{L} D_l(u)\right)\prod_{l=1}^{L} g(u_l)du$$

Minimal power for K false nulls:

$$\Pi_{any}(D) = \int_{[0,1]^K}\mathbb{I}\left\{\sum_{l=1}^{K} D_l(u) > 0\right\}\prod_{l=1}^{K} g(u_l)du$$

Resulting optimization problem for strong control

$$\max_{D:[0,1]^K \to \{0,1\}^K} \quad \Pi(D)$$
$$\text{S.t.} \quad Err_L(D) \leq \alpha, \;\; 0 \leq L < K,$$

where Π is the chosen power measure, *Err* is the chosen type-I error measure, and we have K and not $2^K - 1$ constraints because of the symmetry

"Minor" problems:

- D defines a continuum of variables
- The problem is integer
- The problem is not linear in D

Resulting optimization problem for strong control

$$\max_{D:[0,1]^K \to \{0,1\}^K} \quad \Pi(D)$$
$$\text{S.t.} \quad Err_L(D) \leq \alpha, \;\; 0 \leq L < K,$$

where Π is the chosen power measure, *Err* is the chosen type-I error measure, and we have K and not $2^K - 1$ constraints because of the symmetry

"Minor" problems:

- D defines a continuum of variables
- The problem is integer
- The problem is not linear in D

Monotonicity and linearity

Lemma

The optimal solution is always weakly monotone*:*

$$u_i \leq u_j \Rightarrow D_i^*(u) \geq D_j^*(u).$$

Given weak monotonicity, it turns out FDR_L, $FWER_L$, Π_L, Π_{any} can all be written as linear functionals of D, for example:

$$\Pi_{any}(D) = K! \int_Q D_1(u) \prod_{l=1}^{K} g(u_l) du$$

$$FWER_L(D) = L!(K-L)! \int_Q \sum_k D_k(u) \sum_{i \in \binom{K}{L}, \bar{i}_{min}=k} \prod_{l \in i} g(u_l) du,$$

where $Q = \left\{u \in [0,1]^K : u_1 \leq u_2 \leq \ldots \leq u_K\right\}$ is the ordered "corner", and i enumerates over possible combinations of L false nulls.

Monotonicity and linearity

Lemma

The optimal solution is always weakly monotone*:*

$$u_i \leq u_j \Rightarrow D_i^*(u) \geq D_j^*(u).$$

Given weak monotonicity, it turns out FDR_L, $FWER_L$, Π_L, Π_{any} can all be written as linear functionals of D, for example:

$$\Pi_{any}(D) = K! \int_Q D_1(u) \prod_{l=1}^{K} g(u_l) du$$
$$FWER_L(D) = L!(K-L)! \int_Q \sum_k D_k(u) \sum_{i \in \binom{K}{L}, \bar{i}_{min}=k} \prod_{l \in i} g(u_l) du,$$

where $Q = \left\{u \in [0,1]^K : u_1 \leq u_2 \leq \ldots \leq u_K\right\}$ is the ordered "corner", and i enumerates over possible combinations of L false nulls.

Relaxing to linear program

We found out everything is linear, next we relax the integer requirement, and end up with an infinite linear program:

$$\max_{D:Q\to[0,1]^K} \int_Q \left(\sum_{i=1}^{K} a_i(u)D_i(u)\right) du \tag{1}$$

$$\text{s.t.} \quad \int_Q \left(\sum_{i=1}^{K} b_{L,i}(u)D_i(u)\right) du \le \alpha \,,\; 0 \le L < K.$$

$$0 \le D_K(u) \le \ldots \le D_1(u) \le 1 \,,\; \forall u \in Q,$$

where $a_i, \; i = 1, \ldots, K$ and $b_{L,i}, \; i = 1, \ldots, K, \; L = 0, \ldots, K-1$ are fixed non-negative integrable functions over Q.

Remaining problems:

- How do we solve this infinite linear program?
- We still need an integer solution!

Relaxing to linear program

We found out everything is linear, next we relax the integer requirement, and end up with an infinite linear program:

$$\max_{D:Q\to[0,1]^K} \quad \int_Q \left(\sum_{i=1}^{K} a_i(u)D_i(u)\right) du \tag{1}$$

$$\text{s.t.} \quad \int_Q \left(\sum_{i=1}^{K} b_{L,i}(u)D_i(u)\right) du \le \alpha \ , \ 0 \le L < K.$$

$$0 \le D_K(u) \le \ldots \le D_1(u) \le 1 \ , \ \forall u \in Q,$$

where $a_i, \ i = 1, \ldots, K$ and $b_{L,i}, \ i = 1, \ldots, K, \ L = 0, \ldots, K-1$ are fixed non-negative integrable functions over Q.

Remaining problems:

- How do we solve this infinite linear program?
- We still need an integer solution!

Relaxing to linear program

We found out everything is linear, next we relax the integer requirement, and end up with an infinite linear program:

$$\max_{D:Q\to[0,1]^K} \quad \int_Q \left(\sum_{i=1}^K a_i(u)D_i(u)\right) du \tag{1}$$

$$\text{s.t.} \quad \int_Q \left(\sum_{i=1}^K b_{L,i}(u)D_i(u)\right) du \le \alpha \ , \ 0 \le L < K.$$

$$0 \le D_K(u) \le \ldots \le D_1(u) \le 1 \ , \ \forall u \in Q,$$

where $a_i, \ i = 1, \ldots, K$ and $b_{L,i}, \ i = 1, \ldots, K, \ L = 0, \ldots, K-1$ are fixed non-negative integrable functions over Q.

Remaining problems:

- How do we solve this infinite linear program?
- We still need an integer solution!

Optimality conditions for the infinite linear program

Using the theory of Euler-Lagrange, we can derive the following "KKT-like" necessary conditions for optimal solution to our problem, in addition to the (primal feasibility) original constraints:

$$a_i(u) - \sum_{L=0}^{K-1} \mu_L b_{L,i}(u) - \lambda_i(u) + \lambda_{i+1}(u) = 0, i = 1, \ldots, K. \quad (2)$$

$$\mu_L \left\{ \int_Q \left(\sum_{i=1}^{K} b_{L,i}(u) D_i(u) \right) du - \alpha \right\} = 0, L = 0, \ldots, K-1 \quad (3)$$

$$\lambda_{K+1}(u) D_K(u) = 0 \ \forall u \in Q \quad (4)$$

$$\lambda_j(u)(D_{j-1}(u) - D_j(u)) = 0 \ , \ \forall u \in Q, \ j = 2, \ldots, K \quad (5)$$

$$\lambda_1(u)(D_1(u) - 1) = 0 \ , \ \forall u \in Q., \quad (6)$$

μ_L and $\lambda_j(u)$ are non-negative Lagrange multipliers
condition (2) is the *stationarity* condition
conditions (3–6) are *complementary slackness* conditions.

Optimality conditions for the infinite linear program

Using the theory of Euler-Lagrange, we can derive the following "KKT-like" necessary conditions for optimal solution to our problem, in addition to the (primal feasibility) original constraints:

$$a_i(u) - \sum_{L=0}^{K-1} \mu_L b_{L,i}(u) - \lambda_i(u) + \lambda_{i+1}(u) = 0, i = 1, \ldots, K. \quad (2)$$

$$\mu_L \left\{ \int_Q \left(\sum_{i=1}^{K} b_{L,i}(u) D_i(u) \right) du - \alpha \right\} = 0, L = 0, \ldots, K-1 \quad (3)$$

$$\lambda_{K+1}(u) D_K(u) = 0 \ \forall u \in Q \quad (4)$$

$$\lambda_j(u)(D_{j-1}(u) - D_j(u)) = 0 \ , \ \forall u \in Q, \ j = 2, \ldots, K \quad (5)$$

$$\lambda_1(u)(D_1(u) - 1) = 0 \ , \ \forall u \in Q., \quad (6)$$

μ_L and $\lambda_j(u)$ are non-negative Lagrange multipliers
condition (2) is the *stationarity* condition
conditions (3–6) are *complementary slackness* conditions.

Optimality conditions for the infinite linear program

Using the theory of Euler-Lagrange, we can derive the following "KKT-like" necessary conditions for optimal solution to our problem, in addition to the (primal feasibility) original constraints:

$$a_i(u) - \sum_{L=0}^{K-1} \mu_L b_{L,i}(u) - \lambda_i(u) + \lambda_{i+1}(u) = 0, i = 1, \ldots, K. \quad (2)$$

$$\mu_L \left\{ \int_Q \left(\sum_{i=1}^{K} b_{L,i}(u) D_i(u) \right) du - \alpha \right\} = 0, L = 0, \ldots, K-1 \quad (3)$$

$$\lambda_{K+1}(u) D_K(u) = 0 \; \forall u \in Q \quad (4)$$

$$\lambda_j(u)(D_{j-1}(u) - D_j(u)) = 0 \;, \; \forall u \in Q, \; j = 2, \ldots, K \quad (5)$$

$$\lambda_1(u)(D_1(u) - 1) = 0 \;, \; \forall u \in Q., \quad (6)$$

μ_L and $\lambda_j(u)$ are non-negative Lagrange multipliers
condition (2) is the *stationarity* condition
conditions (3–6) are *complementary slackness* conditions.

Solving the infinite linear program

Lemma

Under non-redundancy assumptions, a solution that complies with the conditions (2)–(6) is integer almost everywhere on $[0,1]^K$*.*

Lemma

A solution that complies with these necessary conditions is in fact optimal.

The proof is based on convex duality arguments, which hold for infinite dimensional problems.

Solving the infinite linear program

Lemma

Under non-redundancy assumptions, a solution that complies with the conditions (2)–(6) *is integer almost everywhere on* $[0,1]^K$.

Lemma

A solution that complies with these necessary conditions is in fact optimal.

The proof is based on convex duality arguments, which hold for infinite dimensional problems.

Main theoretical result

Putting all of our lemmas together we conclude:

Theorem

Under mild regularity conditions, for any choice of power function from Π_{any}, Π_L and error measure FWER or FDR, the optimal procedure can be explicitly found by finding an integer solution which is feasible for Problem (1) *and complies with the optimality conditions.*

This in fact leads to an algorithm for finding the optimal solution, as follows.

Main ideas of the resulting algorithm

Investigating the optimality conditions we find that if we know the value of K Lagrange multipliers $\mu = (\mu_0, ...\mu_{K-1})$ we can infer the solution D^μ. If D^μ is feasible, then it is optimal.

Specifically an algorithm requires:

1. An approach for searching the space $(\mathbb{R}^+ \cup \{0\})^K$ of possible μ vectors for a solution μ^*.
2. An approach for efficiently calculating the coefficients b_{Li} in our integrals.
3. An approach for integration (exact or numerical), to calculate

$$\int_Q \left(\sum_{i=1}^{K} b_{L,i}(u) D_i^\mu(u) \right) du$$

for any given μ vector and asses the error relative to the optimality conditions.

This is very computationally demanding, but possible for low K.

Main ideas of the resulting algorithm

Investigating the optimality conditions we find that if we know the value of K Lagrange multipliers $\mu = (\mu_0, ...\mu_{K-1})$ we can infer the solution D^μ. If D^μ is feasible, then it is optimal.

Specifically an algorithm requires:

1. An approach for searching the space $(\mathbb{R}^+ \cup \{0\})^K$ of possible μ vectors for a solution μ^*.
2. An approach for efficiently calculating the coefficients b_{Li} in our integrals.
3. An approach for integration (exact or numerical), to calculate

$$\int_Q \left(\sum_{i=1}^{K} b_{L,i}(u) D_i^\mu(u) \right) du$$

   for any given μ vector and asses the error relative to the optimality conditions.

This is very computationally demanding, but possible for low K.

Example: Controlling FWER for $K = 3$ independent normal means

Given $X_k \sim N(\theta, 1)$, $k = 1, 2, 3$, testing:

$$H_{0k} \quad : \quad \theta = 0$$
$$H_{Ak} \quad : \quad \theta = \theta_A < 0$$

while (strongly) controlling FWER and seeking to maximize either Π_3 or Π_{any}

Standard solution: Bonferroni-Holm

$$u_1 = 0.000016 \qquad u_1 = 0.0166$$

Example: Controlling FWER for $K = 3$ independent normal means

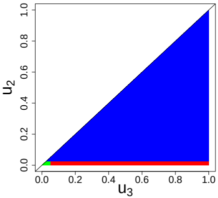
Given $X_k \sim N(\theta, 1)$, $k = 1, 2, 3$, testing:

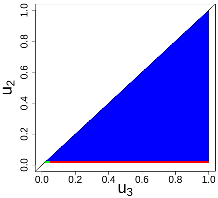
$$H_{0k} \quad : \quad \theta = 0$$
$$H_{Ak} \quad : \quad \theta = \theta_A < 0$$

while (strongly) controlling FWER and seeking to maximize either Π_3 or Π_{any}

Standard solution: Bonferroni-Holm

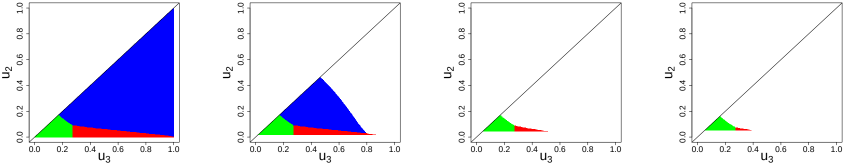
$u_1 = 0.000016$ $\qquad$ $u_1 = 0.0166$



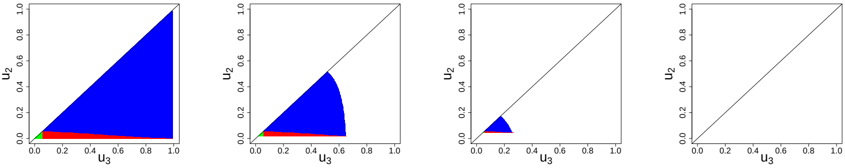


$u_1 = 0.000016$ $\quad$ $u_1 = 0.0166$ $\quad$ $u_1 = 0.044$ $\quad$ $u_1 = 0.054$

$\theta_A = -0.5$:



$\theta_A = -2$:



Example: Controlling FDR for $K = 3$ independent normal means

Given $X_k \sim N(\theta, 1)$, $k = 1, 2, 3$, testing:

$$\begin{aligned} H_{0k} &: \quad \theta = 0 \\ H_{Ak} &: \quad \theta = \theta_A < 0 \end{aligned}$$

while (strongly) controlling FDR and seeking to maximize Π_3

Standard solution: (MA)BH (Solari & Goeman 17)

$$u_1 = 0.000016 \quad u_1 = 0.0173 \quad u_1 = 0.0438$$

Example: Controlling FDR for $K = 3$ independent normal means

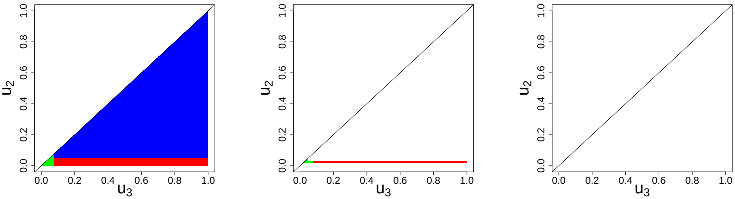
Given $X_k \sim N(\theta, 1)$, $k = 1, 2, 3$, testing:

$$H_{0k} \;:\; \theta = 0$$
$$H_{Ak} \;:\; \theta = \theta_A < 0$$

while (strongly) controlling FDR and seeking to maximize Π_3

Standard solution: (MA)BH (Solari & Goeman 17)

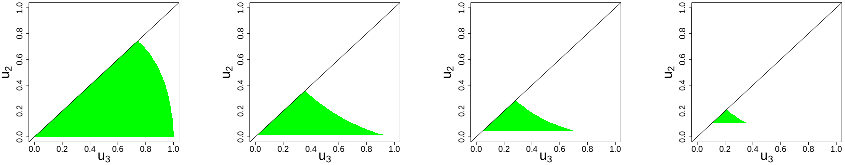
$u_1 = 0.000016$ $\qquad$ $u_1 = 0.0173$ $\qquad$ $u_1 = 0.0438$



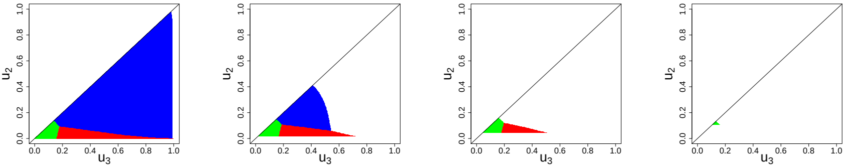
FDR OMT solutions for Π_3

$u_1 = 0.000016$ $\quad$ $u_1 = 0.0166$ $\quad$ $u_1 = 0.0438$ $\quad$ $u_1 = 0.106$

$\theta_A = -0.35$:



$\theta_A = -2$:



FWER, FDR OMT power gains for $\Pi_{\theta,3}$

FWER		
θ_A	Bonferroni-Holm	OMT policy
-0.5	0.0547	0.111
-1.33	0.241	0.363
-2	0.530	0.633

FDR			
θ_A	Benjamini-Hochberg	MABH	OMT policy
-0.35	0.042	0.045	0.150
-0.5	0.059	0.064	0.196
-2	0.574	0.633	0.799

Summary so far

We can in principle find optimal procedures for strong FDR or FWER control with simple, fixed alternative for any K, but computations are hard
In practice we demonstrate $K = 3$

Next steps:
Deal with complex alternatives — find maximin solutions
Design approximations for large K

Summary so far

We can in principle find optimal procedures for strong FDR or FWER control with simple, fixed alternative for any K, but computations are hard
In practice we demonstrate $K = 3$

Next steps:
Deal with complex alternatives — find maximin solutions
Design approximations for large K

Beyond simple hypotheses: a maximin formulation

Maximize minimal power among all alternatives of interest $\theta \in \Theta_B \subseteq (0, -\infty)$, requiring validity for all one-sided alternatives:

$$\begin{aligned} \max_{D:[0,1]^K \to \{0,1\}^K} \quad & \min_{\theta \in \Theta_B} \Pi_\theta(D) \\ \text{s.t.} \quad & Err_{h,\theta}(D) \leq \alpha \; , \; \forall h \in \{0,1\}^K, \; \theta \in (0, -\infty)^K. \end{aligned} \tag{7}$$

Maximin solution

Theorem

Assume that we can find two values $\theta_O \in \Theta_B,\ \theta_A \leq 0$ such that:

1. *$D^*(\theta_O, \theta_A)$ is the optimal solution of a single objective problem at θ_O.*
2. *The power of this solution at other values is higher:*

$$\Pi_{\theta_O^K}\left(D^*(\theta_O, \theta_A)\right) \leq \Pi_\theta\left(D^*(\theta_O, \theta_A)\right)\ \forall \theta \in \Theta_B^K.$$

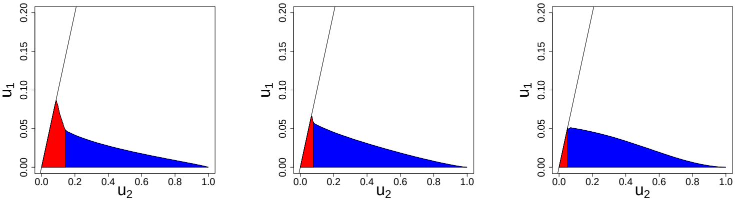
Then $D^*(\theta_O, \theta_A)$ is the solution to the maximin problem (7).

This is a sufficient condition — we don't know when it holds, but when it does we can confirm optimality.

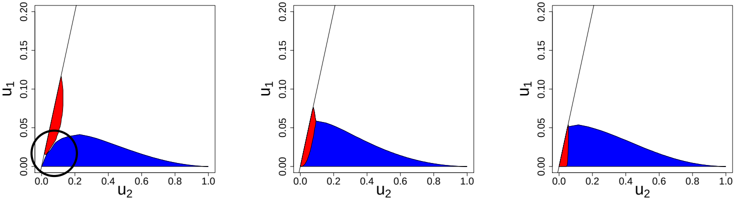
Two normal means, FWER control, $\Theta_B = \{\theta \leq \theta_0\}$

$\theta_0 = -0.5$ $\theta_0 = -1$ $\theta_0 = -2$

Simple optimal:



Maximin optimal:



Power comparison

	Strong FWER control			Strong FDR control		
θ_0	Bonf.-Holm	OMT	maximin	MABH	OMT	maximin
-0.5	0.076	0.118	0.099	0.086	0.174	0.129
-1	0.184	0.251	0.237	0.214	0.326	0.296
-2	0.581	0.637	0.636	0.660	0.734	0.733

Application to systematic reviews in the Cochrane library

For subgroup analyses with $K = 3$ subgroups, here is a summary of discoveries made by each rejection policy, for the 1321 outcomes from the Cochrane database that met our selection criteria[1].

	maximin	Holm	closed-Stouffer
Avg. no. discoveries	1.097	1.089	1.040
% at least one discovery	0.620	0.594	0.548

[1] We considered all the updated reviews up to 2017 in all domains. For subgroup analysis, we considered outcomes that satisfied the following criteria: the outcome was a comparison of means; the number of participants in each comparison group was more than ten; there were at least three subgroups.

Adding weak monotonicity constraint

A major point of concern (or interest?) is the surprising shape of the rejection regions and especially maximin

In our view, this is a property of the problem and error measure, not the solution

Still, we also solve the problem with a *weak monotonicity* requirement, that decreasing p-values u increases the rejection vector $D(u)$

Turns out, this can be solved with similar tools, a bit more complex — essentially adding a solution of an isotonic regression problem within the Lagrange multipliers search

Adding weak monotonicity constraint

A major point of concern (or interest?) is the surprising shape of the rejection regions and especially maximin

In our view, this is a property of the problem and error measure, not the solution

Still, we also solve the problem with a *weak monotonicity* requirement, that decreasing p-values u increases the rejection vector $D(u)$

Turns out, this can be solved with similar tools, a bit more complex — essentially adding a solution of an isotonic regression problem within the Lagrange multipliers search

Adding weak monotonicity constraint

A major point of concern (or interest?) is the surprising shape of the rejection regions and especially maximin

In our view, this is a property of the problem and error measure, not the solution

Still, we also solve the problem with a *weak monotonicity* requirement, that decreasing p-values u increases the rejection vector $D(u)$

Turns out, this can be solved with similar tools, a bit more complex — essentially adding a solution of an isotonic regression problem within the Lagrange multipliers search

Adding weak monotonicity constraint

A major point of concern (or interest?) is the surprising shape of the rejection regions and especially maximin

In our view, this is a property of the problem and error measure, not the solution

Still, we also solve the problem with a *weak monotonicity* requirement, that decreasing p-values u increases the rejection vector $D(u)$

Turns out, this can be solved with similar tools, a bit more complex — essentially adding a solution of an isotonic regression problem within the Lagrange multipliers search

Comparing results without and with weak monotonicity

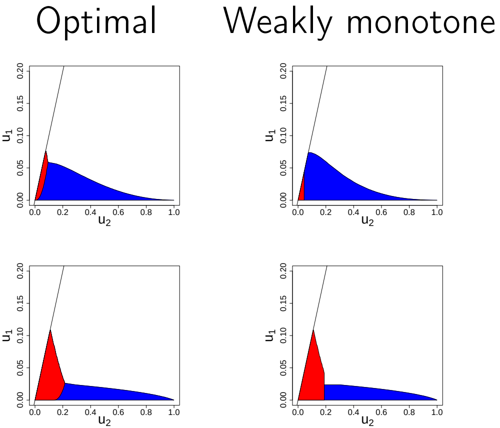


Figure: Top row: maximin for FWER control with $\Theta_B = (-\infty, -1]$. Bottom row: OMT for FDR control with $\theta = -1$. The power loss is minimal: from 0.237 to 0.231 in the first row, and from 0.326 to 0.325 in the second.

Discussion: Computation and Approximations

We currently solve problems up to $K = 3$

We believe with improved computation we can solve $K = 10$ or possibly $K = 100$

But for K in thousands as in modern domains like genetics need a different approach

In Ruth Heller's talk we discuss the two-group model, where we can apply our thinking to solve such large problems

Conclusions

- Attaining high power while controlling type-I error is the primary criterion for designing good tests. This issue becomes more critical as the number of tests increases
- This leads to optimal multiple testing problems that are inherently (hard) optimization problems
- We demonstrate that they can be solved, leading to novel and more powerful procedures than existing methods
- We encounter computational and theoretical challenges
- The maximin approach and the two-group model demonstrate two distinctly different directions that we can take to overcome challenges and produce practically useful tools

Conclusions

- Attaining high power while controlling type-I error is the primary criterion for designing good tests. This issue becomes more critical as the number of tests increases
- This leads to optimal multiple testing problems that are inherently (hard) optimization problems
- We demonstrate that they can be solved, leading to novel and more powerful procedures than existing methods
- We encounter computational and theoretical challenges
- The maximin approach and the two-group model demonstrate two distinctly different directions that we can take to overcome challenges and produce practically useful tools

Conclusions

- Attaining high power while controlling type-I error is the primary criterion for designing good tests. This issue becomes more critical as the number of tests increases
- This leads to optimal multiple testing problems that are inherently (hard) optimization problems
- We demonstrate that they can be solved, leading to novel and more powerful procedures than existing methods
- We encounter computational and theoretical challenges
- The maximin approach and the two-group model demonstrate two distinctly different directions that we can take to overcome challenges and produce practically useful tools

Conclusions

- Attaining high power while controlling type-I error is the primary criterion for designing good tests. This issue becomes more critical as the number of tests increases
- This leads to optimal multiple testing problems that are inherently (hard) optimization problems
- We demonstrate that they can be solved, leading to novel and more powerful procedures than existing methods
- We encounter computational and theoretical challenges
- The maximin approach and the two-group model demonstrate two distinctly different directions that we can take to overcome challenges and produce practically useful tools

Conclusions

- Attaining high power while controlling type-I error is the primary criterion for designing good tests. This issue becomes more critical as the number of tests increases
- This leads to optimal multiple testing problems that are inherently (hard) optimization problems
- We demonstrate that they can be solved, leading to novel and more powerful procedures than existing methods
- We encounter computational and theoretical challenges
- The maximin approach and the two-group model demonstrate two distinctly different directions that we can take to overcome challenges and produce practically useful tools

Thanks!

saharon@tauex.tau.ac.il

arxiv.org/abs/1804.10256
arxiv.org/abs/1902.00892