A Large-Sample Approach to Controlling the False Discovery Rate

Christopher R. Genovese

Department of Statistics

Carnegie Mellon University

Larry Wasserman

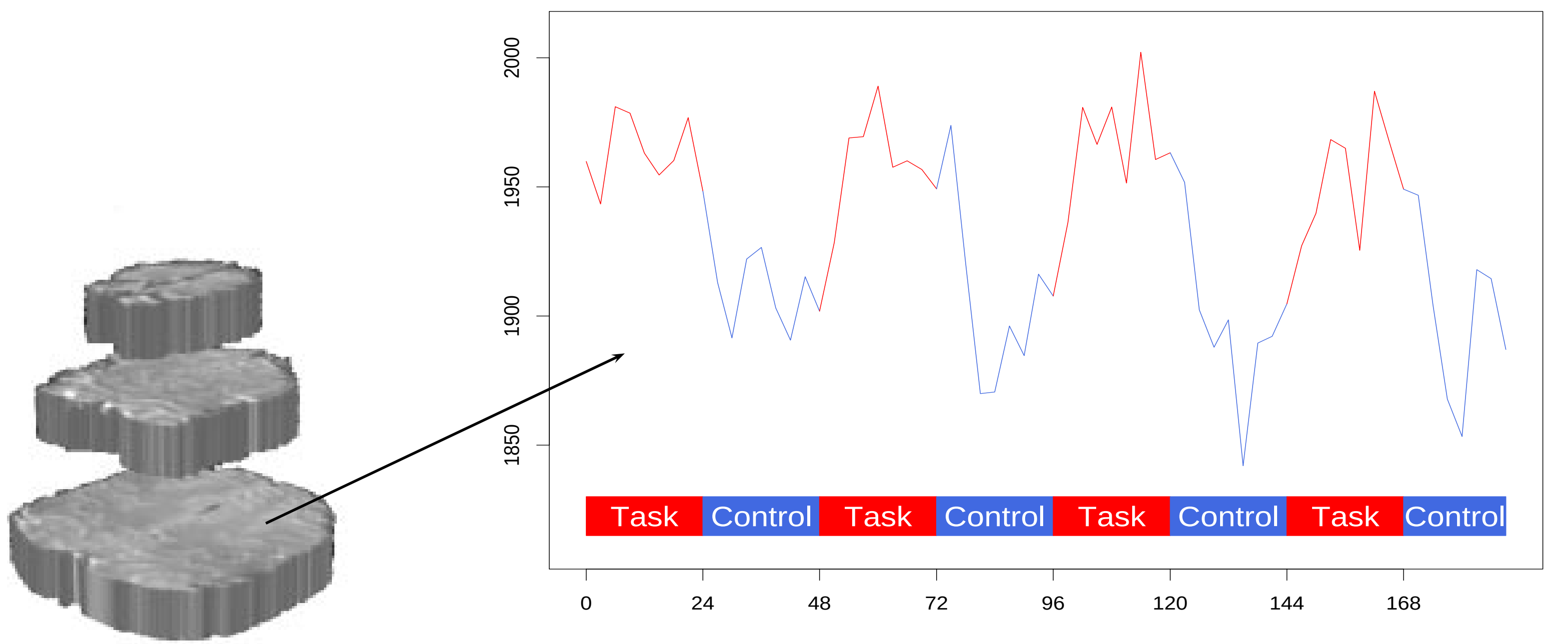
Department of Statistics

Carnegie Mellon University

This work partially supported by NSF Grant SES 9866147.

Motivating Example #1: fMRI

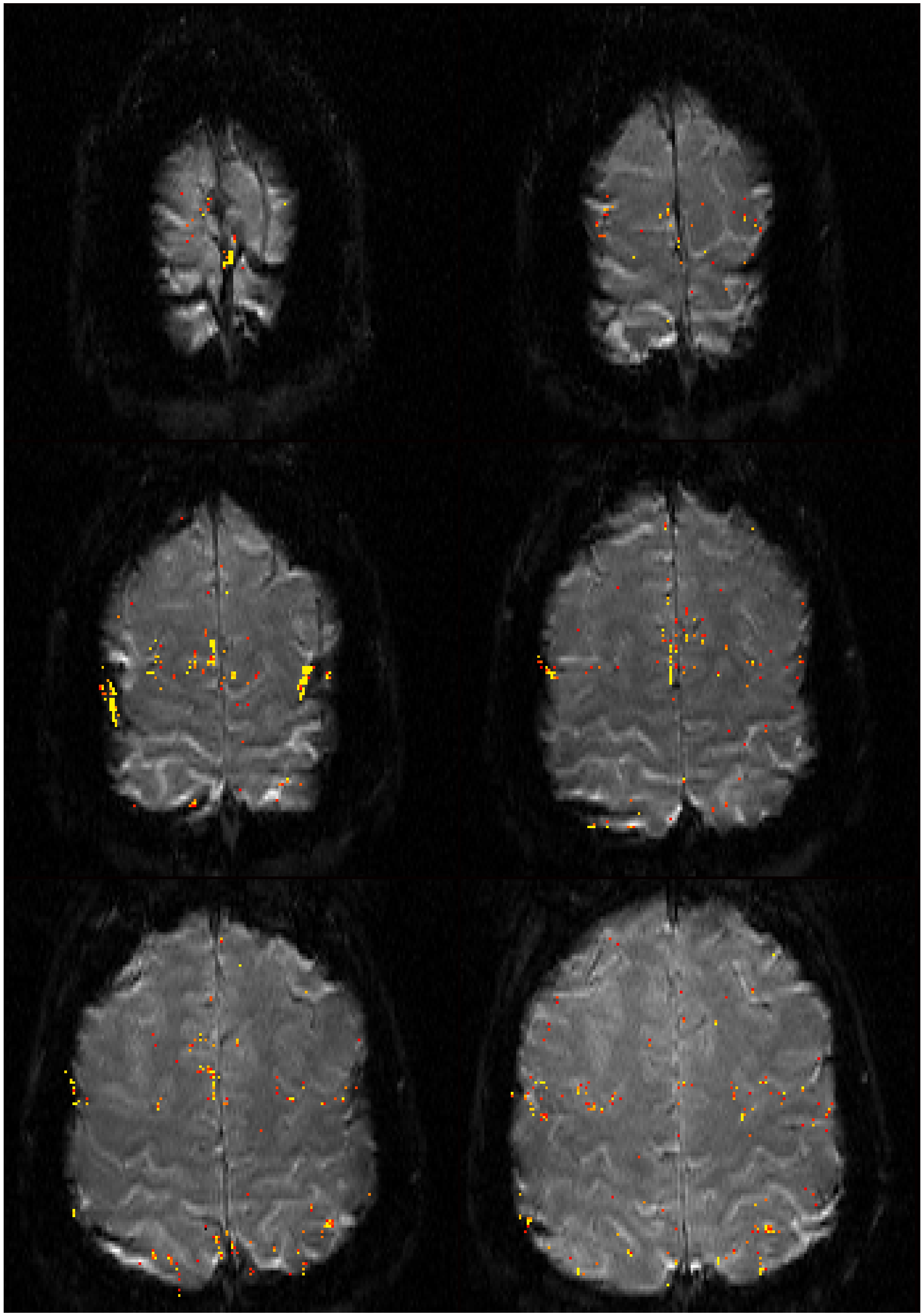
- fMRI Data: Time series of 3-d images acquired while subject performs specified tasks.



- Goal: Characterize task-related signal changes caused (indirectly) by neural activity. [See, for example, Genovese (2000), *JASA* 95, 691.]

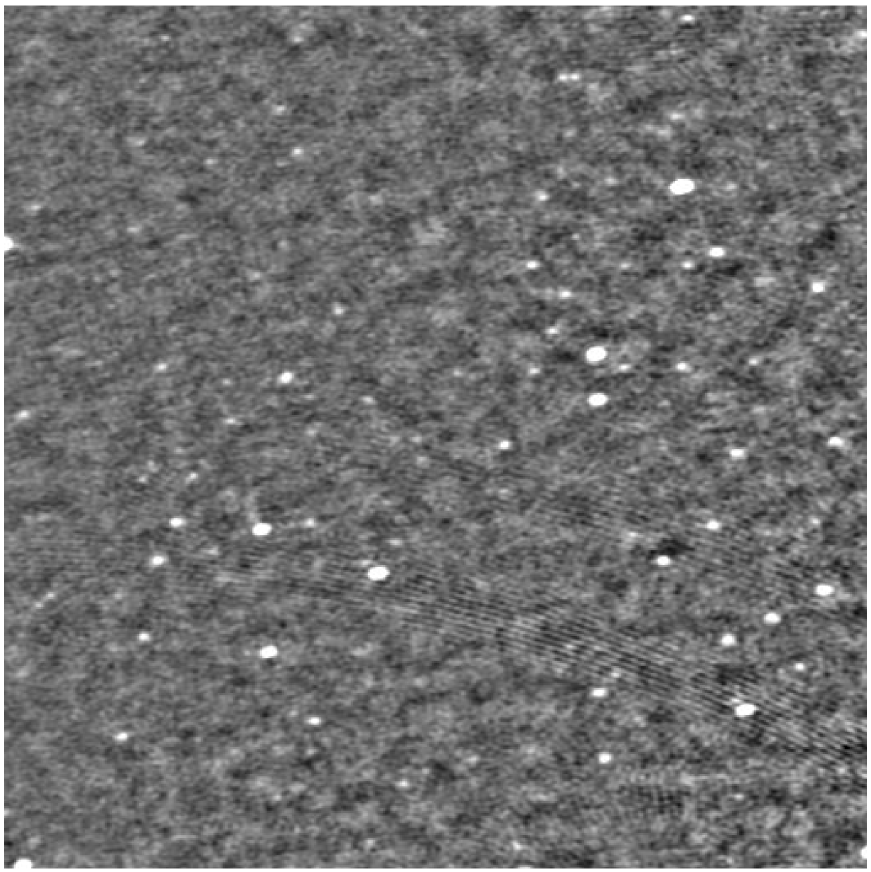
fMRI (cont'd)

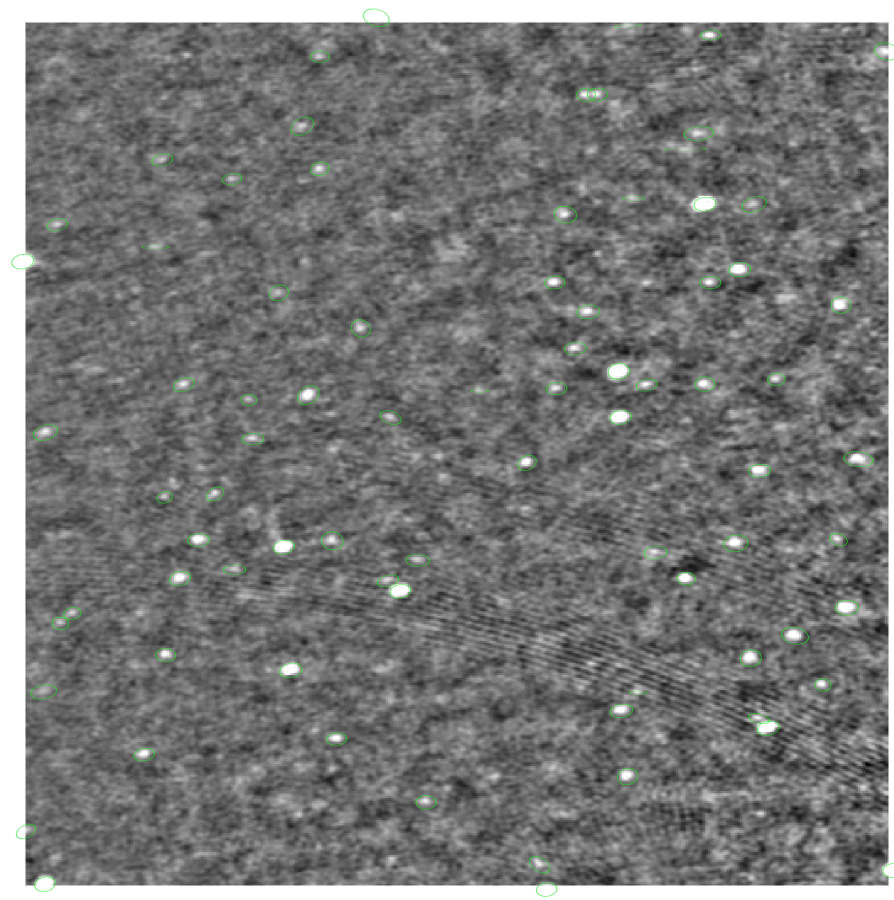
Perform hypothesis tests at many thousands of volume elements to identify loci of activation.



Motivating Example #2: Source Detection

- Interferometric radio telescope observations processed into digital image of the sky in radio frequencies.
- Signal at each pixel is a mixture of source and background signals.





Motivating Example #3: DNA Microarrays

- New technologies allow measurement of gene expression for thousands of genes simultaneously.

Condition 1 (Subject)

Gene	1	2	3	...
1	X_{111}	X_{121}	X_{131}	$\ldots$
2	X_{211}	X_{221}	X_{231}	$\ldots$
3	$\vdots$	$\vdots$	$\vdots$	$\ldots$
4				
5				
6				
$\vdots$				

Condition 2 (Subject)

Gene	1	2	3	...
1	X_{112}	X_{122}	X_{132}	$\ldots$
2	X_{212}	X_{222}	X_{232}	$\ldots$
3	$\vdots$	$\vdots$	$\vdots$	$\ldots$
4				
5				
6				
$\vdots$				

- Goal: Identify genes associated with differences among conditions.
- Typical analysis: hypothesis test at each gene.

Recent Work on FDR

Abromovich, et al. (2000)

Benjamini & Hochberg (1995)

Benjamini & Liu (1999)

Benjamini & Hochberg (2000)

Benjamini & Yekutieli (2001)

Efron, et al. (2001)

Finner and Roters (2001, 2002)

Genovese & Wasserman (2001,2002)

Sarkar (2002)

Storey (2001,2002)

Storey & Tibshirani (2001)

Tusher, Tibshirani, Chu (2001)

The Multiple Testing Problem

- Perform m simultaneous hypothesis tests.

  Classify results as follows:

	H_0 Retained	H_0 Rejected	Total
H_0 True	$M_{0\|0}$	$M_{1\|0}$	M_0
H_0 False	$M_{0\|1}$	$M_{1\|1}$	M_1
Total	$m - R$	R	m

  Here, $M_{i|j}$ is the number of H_i chosen when H_j true.

  Only R and m are observed.

False Discovery and Nondiscovery Proportions

- Define the False Discovery Proportion (FDP) and the False Nondiscovery Proportion (FNP) as follows:

$$\text{FDP} = \begin{cases} \dfrac{M_{1|0}}{R} & \text{if } R > 0, \\ 0, & \text{if } R = 0. \end{cases} \qquad \text{FNP} = \begin{cases} \dfrac{M_{0|1}}{m - R} & \text{if } R < m, \\ 0, & \text{if } R = m. \end{cases}$$

- Then, the False Discovery Rate (FDR) and the False Nondiscovery Rate (FNR) are given by

$$\text{FDR} = \text{E}(\text{FDP}) \qquad \text{FNR} = \text{E}(\text{FNP}).$$

Road Map

1. Preliminaries
   - Models for FDP and FNP
   - FDP and FNP as stochastic processes
2. Plug-in Procedures
   - Asymptotic behavior of BH procedure
   - Optimal Thresholds
3. Confidence Thresholds
   - Controlling probability of exceeding specified proportion of false discoveries
4. Estimating the p-value distribution

Basic Models

- Let $P^m = (P_1, \ldots, P_m)$ be the p-values for the m tests.
- Let $H^m = (H_1, \ldots, H_m)$ where $H_i = 0$ (or 1) if the i^{th} null hypothesis is true (or false).
- We assume the following model:

$$
\begin{aligned}
H_1, \ldots, H_m &\text{ iid Bernoulli}\langle a \rangle \\
\Xi_1, \ldots, \Xi_m &\text{ iid } \mathcal{L}_{\mathcal{F}} \\
P_i \mid H_i = 0, \Xi_i = \xi_i &\sim \text{Uniform}\langle 0, 1 \rangle \\
P_i \mid H_i = 1, \Xi_i = \xi_i &\sim \xi_i.
\end{aligned}
$$

where $\mathcal{L}_{\mathcal{F}}$ denotes a probability distribution on a class $\mathcal{F}$ of distributions on $[0, 1]$.

Basic Models (cont'd)

- Marginally, $P_1, \ldots, P_m$ are drawn iid from

$$G = (1 - a)U + aF,$$

where U is the Uniform$\langle 0, 1\rangle$ cdf and

$$F = \int \xi \, d\mathcal{L}_{\mathcal{F}}(\xi).$$

- Typical examples:
  - Parametric family: $\mathcal{F}_\Theta = \{F_\theta\colon\ \theta \in \Theta\}$
  - Concave, continuous distributions

$$\mathcal{F}_C = \{F\colon\ F \text{ concave, continuous cdf with } F \geq U\}.$$

- Can also work under what we call the *conditional model* where $H_1, \ldots, H_m$ are fixed, unknown.

Multiple Testing Procedures

- A multiple testing procedure T is a map $[0,1]^m \longrightarrow [0,1]$, where the null hypotheses are rejected in all those tests for which $P_i \leq T(P^m)$. Often call T a *threshold*.
- Examples:

Uncorrected testing	$T_{\mathrm{U}}(P^m) = \alpha$
Bonferroni	$T_{\mathrm{B}}(P^m) = \alpha/m$
Fixed threshold at t	$T_t(P^m) = t$
First r	$T_{(r)}(P^m) = P_{(r)}$
Benjamini-Hochberg	$T_{\mathrm{BH}}(P^m) = \sup\{t\colon \widehat{G}(t) = t/\alpha\}$
Oracle	$T_{\mathrm{O}}(P^m) = \sup\{t\colon G(t) = (1-a)t/\alpha\}$
Plug In	$T_{\mathrm{PI}}(P^m) = \sup\{t\colon \widehat{G}(t) = (1-\widehat{a})t/\alpha\}$
Regression Classifier	$T_{\mathrm{Reg}}(P^m) = \sup\{t\colon \widehat{\mathsf{P}}\{H_1{=}1 \mid P_1{=}t\} > 1/2\}$

FDP and FNP as Stochastic Processes

- Inherent difficulty: FDP, FNP, and a general threshold all depend on the same data.
- Define the FDP and FNP processes, respectively, by

$$\mathsf{FDP}(t) \equiv \mathsf{FDP}(t; P^m, H^m) = \frac{\sum_i 1\Big\{P_i \le t\Big\}(1 - H_i)}{\sum_i 1\Big\{P_i \le t\Big\} + 1\Big\{\text{all } P_i > t\Big\}}$$

$$\mathsf{FNP}(t) \equiv \mathsf{FNP}(t; P^m, H^m) = \frac{\sum_i 1\Big\{P_i > t\Big\} H_i}{\sum_i 1\Big\{P_i > t\Big\} + 1\Big\{\text{all } P_i \le t\Big\}}.$$

- For procedure T, the FDP and FNP are obtained by evaluating these processes at $T(P^m)$.

FDP and FNP as Stochastic Processes (cont'd)

- Both these processes converge to Gaussian processes outside a neighborhood of 0 and 1 respectively.
- For example, define

$$Z_m(t) = \sqrt{m}\left(\mathsf{FDP}(t) - Q(t)\right), \quad \delta \le t \le 1,$$

  where $0 < \delta < 1$ and $Q(t) = (1-a)U/G$.

- Let Z be a mean 0 Gaussian process on $[\delta, 1]$ with covariance kernel

$$K(s,t) = a(1-a)\,\frac{(1-a)stF(s \wedge t) \;+\; aF(s)F(t)(s \wedge t)}{G^2(s)\,G^2(t)}.$$

- Then, $Z_m \rightsquigarrow Z$.

Plug-in Procedures

- Let $\widehat{G}_m$ be the empirical cdf of P^m under the mixture model. Ignoring ties, $\widehat{G}_m(P_{(i)}) = i/m$, so BH equivalent to

$$T_{\mathrm{BH}}(P^m) = \max\left\{t\colon\ \widehat{G}_m(t) = \frac{t}{\alpha}\right\}.$$

  as Storey (2002) first noted.

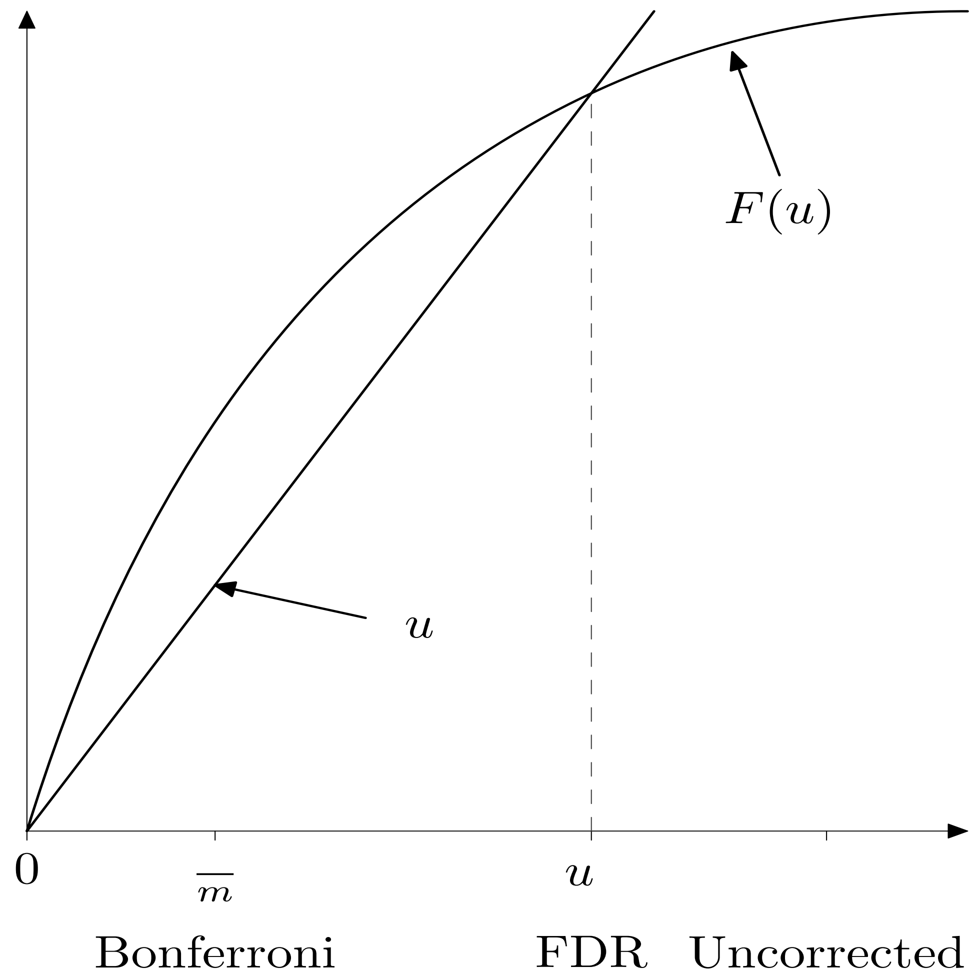
- One can think of this as a plug-in procedure for estimating

$$\begin{aligned} u^*(a, G) &= \max\left\{t\colon\ G(t) = \frac{t}{\alpha}\right\} \\ &= \max\left\{t\colon\ F(t) = \beta t\right\}, \end{aligned}$$

  where $\beta = (1 - \alpha + \alpha a)/\alpha a$.

Asymptotic Behavior of BH Procedure

This yields the following picture:



Optimal Thresholds

- Under the mixture model and in the continuous case,

$$\mathsf{E}(\mathsf{FDP}(T_{\mathrm{BH}}(P^m))) = (1-a)\alpha.$$

- The BH procedure overcontrols FDR and thus will not in general minimize FNR.
- This suggests using T_{PI}, the plug-in estimator for

$$\begin{aligned} t^*(a,G) &= \max\left\{t\colon\ G(t) = \frac{(1-a)t}{\alpha}\right\} \\ &= \max\left\{t\colon\ F(t) = (\beta - 1/\alpha)t\right\}, \end{aligned}$$

where $\beta - 1/\alpha = (1-a)(1-\alpha)/a\alpha$.
- Note that $t^* \geq u^*$.

Optimal Thresholds (cont'd)

- For each $0 \leq t \leq 1$,

$$\mathsf{E}(\mathsf{FDP}(t)) = \frac{(1-a)\,t}{G(t)} + O\left((1-t)^m\right)$$

$$\mathsf{E}(\mathsf{FNP}(t)) = a\,\frac{1-F(t)}{1-G(t)} + O\left((a+(1-a)t)^m\right).$$

- Ignoring $O()$ terms and choosing t to minimize E(FNP(t)) subject to E(FDP(t)) $\leq \alpha$, yields $t^*(a, G)$ as the optimal threshold.
- GW (2002) show that

$$\mathsf{E}(\mathsf{FDP}(t^*(\widehat{a}, \widehat{G}))) \leq \alpha + O(m^{-1/2}).$$

Confidence Thresholds

- In practice, it would be useful to have a procedure T_C that guarantees

$$\mathsf{P}_G\Big\{\mathsf{FDP}(T_C) > c\Big\} \leq \alpha$$

  for some specified c and α.

  We call this a $(1-\alpha, c)$ *confidence threshold procedure*.

- Four approaches: (i) an asymptotic Bootstrap threshold, (ii) an asymptotic closed-form threshold, (iii) an exact (small-sample) threshold requiring numerical search, and (iv) a Bayesian threshold.

- Here, I'll discuss the case where a is known.

  In general, all of this works using an estimator, but this introduces additional complexity.

Bootstrap Confidence Thresholds

- First guess: Choose T such that

$$\mathsf{P}_{\widehat{G}}\Big\{\mathsf{FDP}^*(T) \le c\Big\} \ge 1-\alpha.$$

- This fails. The problem is an additional bias term:

$$\begin{aligned} 1-\alpha &= \mathsf{P}_{\widehat{G}}\Big\{\mathsf{FDP}^*(T) \le c\Big\} \\ &\approx \mathsf{P}_{G}\Big\{\mathsf{FDP}(T) \le c + (Q(T) - \widehat{Q}(T))\Big\} \\ &\neq \mathsf{P}_{G}\Big\{\mathsf{FDP}(T) \le c\Big\}, \end{aligned}$$

  where $Q = (1-a)U/G$ and $\widehat{Q} = (1-a)U/\widehat{G}$.

- Can fix this with double bootstrap (harder) or DKW correction (easier).

Bootstrap Confidence Thresholds (cont'd)

- Let $\beta = \alpha/2$ and $\epsilon_m \equiv \epsilon_m(\beta) = \sqrt{\frac{1}{2m} \log\left(\frac{2}{\beta}\right)}$.
- Procedure
  1. Draw $H_1^* \ldots, H_m^*$ iid Bernoulli$\langle a \rangle$
  2. Draw $P_i^* | H_i^*$ from $(1 - H_i^*)U + H_i^* \widehat{F}$.
  3. Define $\Omega_c^*(t) = \sum_i 1\left\{P_i^* \leq t\right\}(1 - H_i^* - c)$.
  4. Use threshold defined by

$$T_C = \max\left\{t\colon \mathsf{P}_{\widehat{G}}\left\{\Omega_c^*(t) \leq -c\,\epsilon_m\right\} \geq 1 - \beta\right\}.$$

- Then,

$$\mathsf{P}_G\left\{\mathsf{FDP}(T_C) \leq c\right\} \geq 1 - \alpha + O\left(\frac{1}{\sqrt{m}}\right).$$

Closed-Form Asymptotic Confidence Thresholds

- Let

$$t_0 = Q^{-1}(c) \qquad \hat{t}_0 = \widehat{Q}^{-1}(c).$$

- Then define

$$T_C = \hat{t}_0 + \frac{\widehat{\Delta}_{m,\alpha}}{\sqrt{m}},$$

where $\widehat{\Delta}_{m,\alpha}$ is depends on a density estimate of $g = G'$.

- Then,

$$\mathsf{P}_G\Big\{\mathsf{FDP}(T_C) \leq c\Big\} \geq 1 - \alpha + o(1).$$

Closed-Form Asymptotic Confidence Thresholds

- Details:

$$\widehat{\Delta}_{m,\alpha} = \frac{z_{\alpha/2}\left(\sqrt{\widehat{K}_{Q^{-1}}(\widehat{t_0},\widehat{t_0})} + \widehat{g}(\widehat{t_0})\right) + 2\sqrt{\log m}}{1 - \widehat{a} - c\widehat{g}(\widehat{t_0})}$$

$$\widehat{K}_{Q^{-1}}(s,t) = \frac{\widehat{K}_Q(\widehat{Q}^{-1}(s), \widehat{Q}^{-1}(t))}{\widehat{Q'}(\widehat{Q}^{-1}(s))\widehat{Q'}(\widehat{Q}^{-1}(t))}$$

$$\widehat{K}_Q(s,t) = \frac{(1-\widehat{a})^2 st}{\widehat{G}^2(s)\widehat{G}^2(t)}\left[\widehat{G}(s \wedge t) - \widehat{G}(s)\widehat{G}(t)\right].$$

- This requires no bootstrapping but does require density estimation.

  This is analogous to the situation faced when estimating the standard error of a median.

Exact Confidence Thresholds

- Let $\mathcal{M}_\beta$ be a $1-\beta$ confidence set for M_0, derived from the Binomial$\langle m, 1-a\rangle$.
- Define

$$S(t; h^m, p^m) = \frac{\sum_i 1\{p_i \le t\}(1-h_i)}{\sum_i (1-h_i)} \qquad \text{[EDF of null p-values]}$$

$$\mathcal{U}_\beta(p^m) = \left\{ h^m \colon \sum_i (1-h_i) \in \mathcal{M}_\beta \text{ and } \| S(\cdot; h^m, p^m) - U \|_\infty \le \epsilon_{m_0}(\beta) \right\},$$

where $m_0 = \sum_i (1-h_i)$ and $\epsilon_{m_0}(\beta) = \sqrt{\log(2/\beta)/2m_0}$.

- Take $\beta = 1 - \sqrt{1-\alpha}$.

Exact Confidence Thresholds (cont'd)

- Let

$$T_C = \text{sup}\left\{t: \text{ FDP}(t; h^m, P^m) \leq c \text{ and } h^m \in \mathcal{U}_\beta(P^m)\right\}$$
$$\mathcal{G} = \left\{\text{FDP}(\cdot; h^m, P^m)\text{: } h^m \in \mathcal{U}_\beta(P^m)\right\}.$$

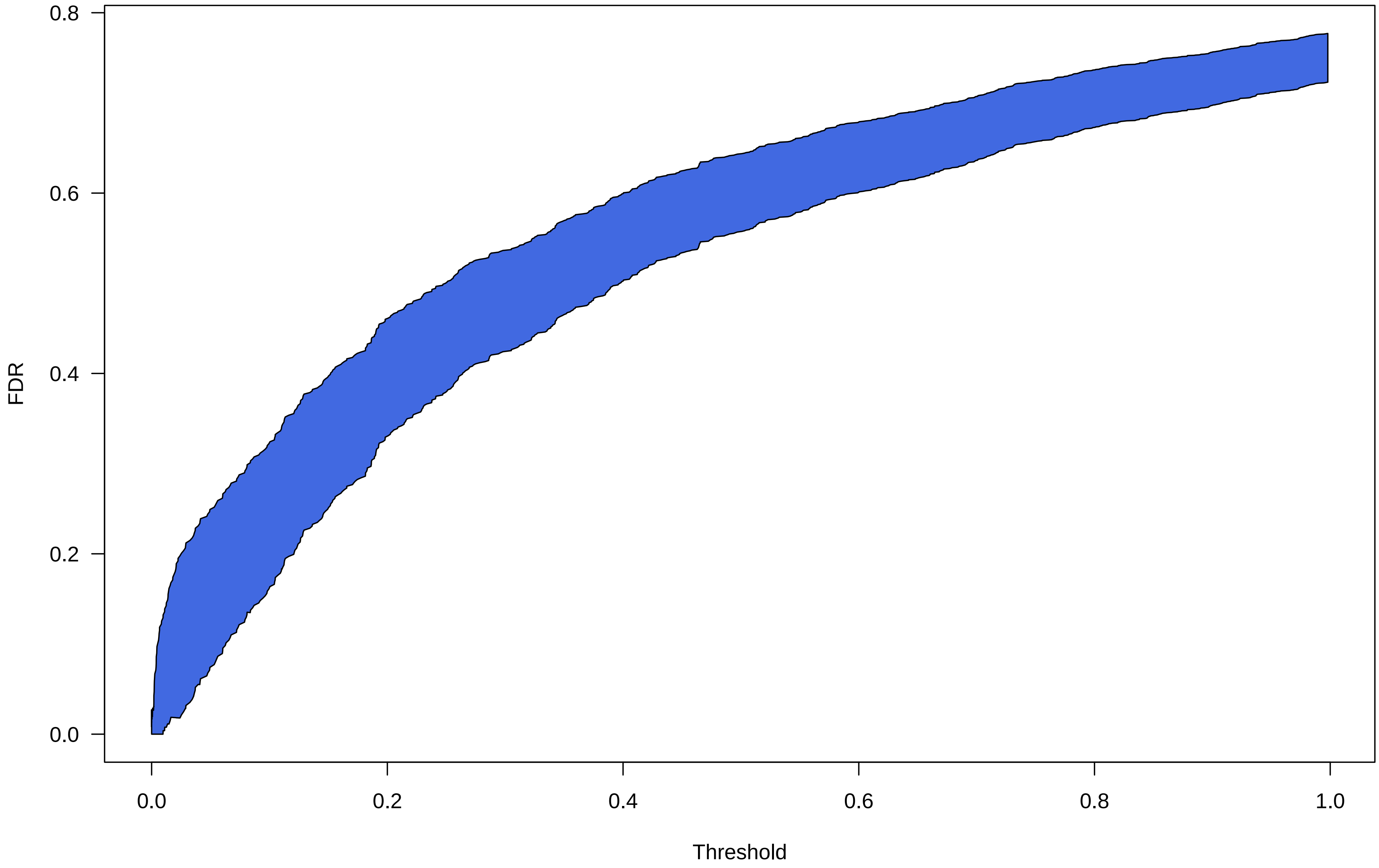
- Then,

$$\text{P}_G\left\{H^m \in \mathcal{U}_\beta(P^m)\right\} \geq 1 - \alpha,$$
$$\text{P}_G\left\{\text{FDP}(\cdot; H^m, P^m) \in \mathcal{G}\right\} \geq 1 - \alpha,$$
$$\text{P}_G\left\{\text{FDP}(T_C) \leq c\right\} \geq 1 - \alpha.$$

Hence, T_C is a $(1 - \alpha, c)$ confidence threshold procedure.

Exact Confidence Thresholds (cont'd)

$\mathcal{G}$ gives a confidence envelope for FDP(t) sample paths.



Estimating a and F

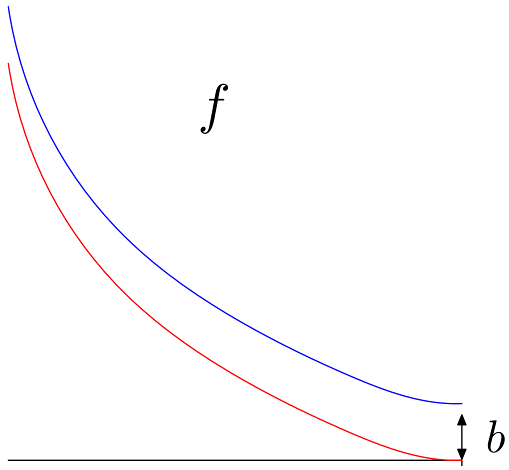
- Recall that the p-value distribution $G = (1 - a)U + aF$ where a and F are unknown.
- We need a good estimate of a for plug-in estimates,

$$T_{\mathrm{PI}}(P^m) = \max\left\{t\colon\ \widehat{G}(t) = \frac{(1-\widehat{a})t}{\alpha}\right\},$$

  that approximate the optimal threshold.
- We need good estimates of a and F for confidence thresholds.

Estimating a and F (cont'd)

- Identifiability and Purity



If $\min f = b > 0$, can write $F = (1-b)U + bF_0$,
$\mathcal{O}_G = \{(\widetilde{a}, \widetilde{F}) : \widetilde{F} \in \mathcal{F}, G = (1-\widetilde{a})U + \widetilde{a}\widetilde{F}\}$
may contain more than one element.

If $f = F'$ is decreasing with $f(1) = 0$, then (a, F) is identifiable.

- In general, let $\underline{a} \leq a$ be the smallest mixing weight in the orbit: $\underline{a} = 1 - \min_t g(t)$. This is identifiable.

  Storey (2002) notes that $0 \leq \sup_{0<t<1} \dfrac{G(t) - t}{1 - t} \leq \underline{a} \leq a \leq 1$.

- $a - \underline{a}$ is typically small: $a - \underline{a} = ae^{-n\theta^2/2}$ in the two-sided test of $\theta = 0$ versus $\theta \neq 0$ in the Normal$\langle\theta, 1\rangle$ model.

Estimating a and F (cont'd)

- Parametric Case
  - Derived a $1-\beta$ one-sided conf. int. for $\underline{a}$ and thus a. (a,θ) typically identifiable even if $a > \underline{a}$; use MLE.
- Non-parametric case:
  - Derived a $1-\beta$ one-sided conf. int. for $\underline{a}$ and thus a.
  - When F concave, get $\widehat{a}_{\mathrm{LCM}} = \underline{a} + O_P(m^{-1/3})$.
  - When F smooth enough, get $\widehat{a}_{\mathrm{S}} = \underline{a} + O_P(m^{-2/5})$.
  - Consistent estimate for F_0 if $\widehat{a}$ consistent for $\underline{a}$:

$$\widehat{F}_m = \operatorname*{argmin}_{H \in \mathcal{F}} \|\widehat{G} - (1-\widehat{a})U - \widehat{a}H\|_\infty.$$

Estimating a and F (cont'd)

- $\widehat{a}_{\mathrm{S}}$ uses "spacings" estimator (Swanepoel, 1999) to estimate $\min g(t)$. This yields

$$\frac{m^{2/5}}{(\log m)^{\delta}}(\widehat{a} - \underline{a}) \rightsquigarrow \text{Normal}\langle 0, (1 - \underline{a})^2 \rangle$$

- In the concave case, take $\widehat{g} = G'_{LCM}$ and $\widehat{a}_{\mathrm{LCM}} = 1 - \widehat{g}(1)$. A $1 - \alpha$ confidence interval for a is

$$\widehat{a}_{\mathrm{LCM}} \pm 4 q_{\alpha} \, |\widehat{g}(1)|^{1/3} \, n^{-1/3}$$

where $\mathsf{P}\left\{ \text{argmax}_h (W(h) - h^2) \geq q_{\alpha} \right\} = \alpha$ and W_h is a 2-sided Brownian motion tied down at 0.

Estimating a and F (cont'd)

- Confidence interval for a given by

$$\mathcal{A}_m = \left[\max_t \frac{\widehat{G}_m(t) - t - \epsilon_m(\alpha)}{1-t}, 1\right],$$

where $\widehat{G}_m$ is EDF and $\epsilon_m(\alpha) = \sqrt{\log(2/\alpha)/2m}$.

Then,

$$1 - \alpha \leq \inf_{a,F} \mathsf{P}\left\{a \in A_m\right\} \leq 1 - \alpha + R_m$$

where

$$R_m = \sum_j (-1)^j \frac{\alpha^{j^2}}{2^{j^2-1}} + O\left(\frac{(\log m)^2}{\sqrt{m}}\right)$$

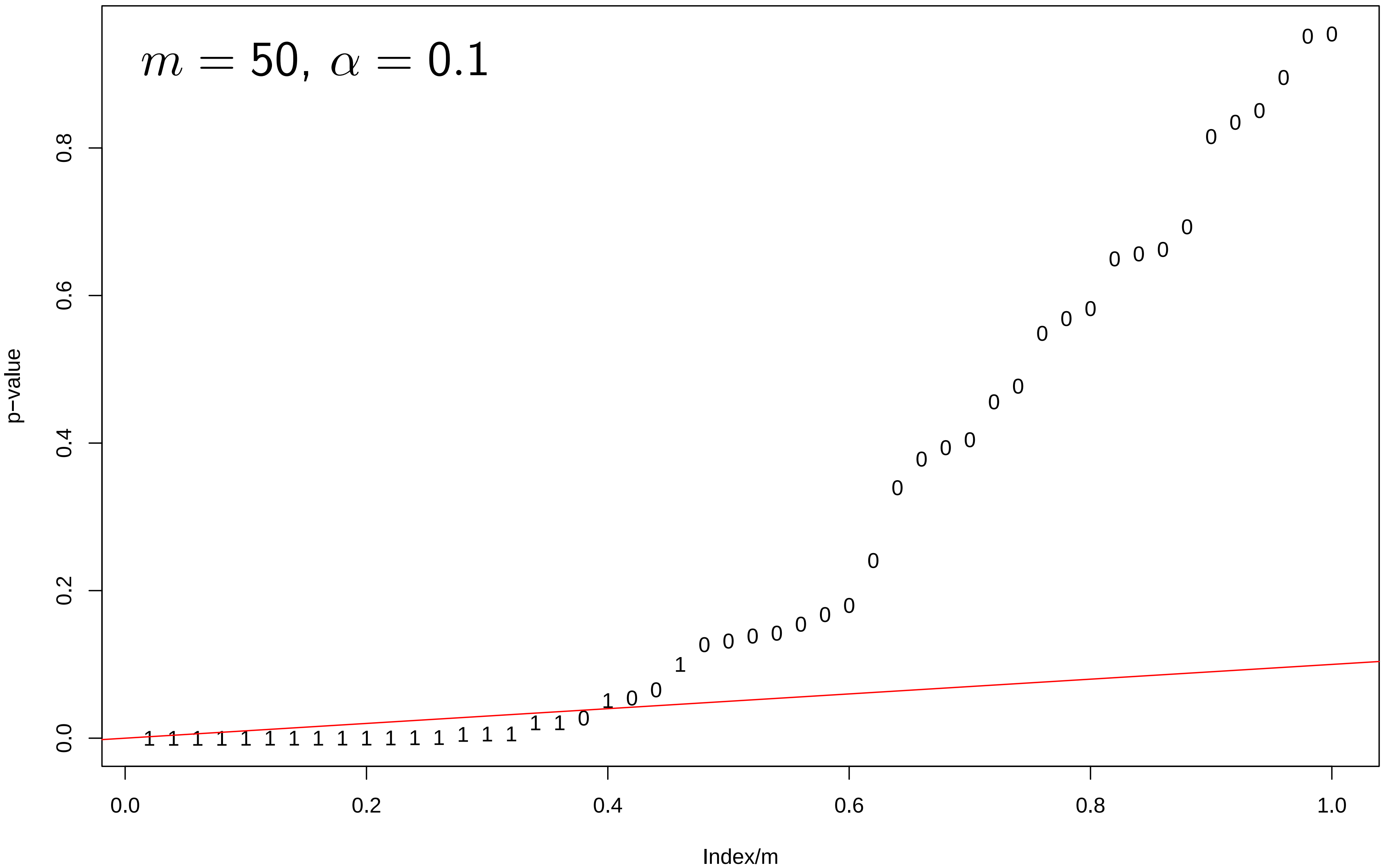
Take-Home Points

- Asymptotic view motivated by particular applications, but asymptotics appear to kick in rather quickly.
- Confidence thresholds address a question that collaborating scientists frequently raise.
- Helpful to think of FDP (FDR) and FNP (FNR) as stochastic processes.

  In general, the threshold and the FDP are coupled, and these correlations can have a large effect.
- Dependence

Recurring Notation

$m, M_0, M_{1\|0}$	# of tests, true nulls, false discoveries
a	Mixture weight on $a$lternative
$H^m = (H_1, \ldots, H_m)$	Unobserved true classifications
$P^m = (P_1, \ldots, P_m)$	Observed p-values
U	CDF of Uniform$\langle 0, 1\rangle$
F, f	Alternative CDF and density
$G = (1-a)U + aF$	Marginal CDF of P_i
$g = G'$	Marginal density of P_i
$\widehat{G}_m$	Estimate of G (e.g., empirical CDF of P^m)
$\epsilon_k(\beta) = \sqrt{\frac{1}{2k}\log\left(\frac{2}{\beta}\right)}$	DKW bound $1-\beta$ quantile of $\lVert \widehat{G}_k - G \rVert_\infty$



$m = 50,\ \alpha = 0.1$
p−value
Index/m
0.0
0.2
0.4
0.6
0.8
1.0

Bayesian Thresholds

- Bayesian Threshold bounds posterior FDR:

$$T_{\text{Bayes}} = \mathsf{sup}\{t : \ \mathsf{E}(\mathsf{FDP}(t) \mid P^m) \leq \alpha\}$$

- Similarly, can construct a posterior (c, α) confidence threshold $T_{\text{Bayes},c}$ by

$$T_{\text{Bayes},c} = \mathsf{sup}\{t : \ \mathsf{P}\Big\{\mathsf{FDP}(t) \leq c \mid P^m\Big\} \leq \alpha\}$$

EBT (Empirical Bayes Testing)

- Efron et al (2001) note that

$$\mathsf{P}\left\{H_i = 0 \mid P^m\right\} = \frac{(1-a)}{g(P_i)} \equiv q(P_i)$$

- Reject whenever $q(p) \leq \alpha$?
- For a, f unknown, $f \geq 0$ implies that

$$a \geq 1 - \min_p g(p) \Longrightarrow \widehat{a} = 1 - \min_p \widehat{g}(p).$$

- Then,

$$\widehat{q}(p) = \frac{1-\widehat{a}}{\widehat{g}(p)} = \frac{\min_s \widehat{g}(s)}{\widehat{g}(p)}$$

EBT versus FDR

- If we reject when $\mathsf{P}\Big\{H_i = 0 \mid P^m\Big\} \leq \alpha$, how many errors are we making?

- Under weak conditions, can show that

$$q(t) \leq \alpha \;\; \text{implies} \;\; Q(t) < \alpha$$

So EBT is conservative.

Behavior of $\widehat{q}$

- THEOREM. Let $\widehat{q}(t) = \frac{(1-a)}{\widehat{g}(t)}$. Suppose that

$$m^{\alpha}(\widehat{g}(t) - g(t)) \rightsquigarrow W$$

for some $\alpha > 0$, where W is a mean 0 Gaussian process with covariance kernel $\tau(v, w)$. Then

$$m^{\alpha}\left(\widehat{q}(t) - q(t)\right) \rightsquigarrow Z$$

where Z is a Gaussian process with mean 0 and covariance kernel

$$K_q(v, w) = \frac{(1-a)^2 \tau(v, w)}{g(v)^4 g(w)^4}.$$

Behavior of $\widehat{q}$ (cont'd)

- Parametric Case: $g \equiv g_\theta = (1-a) + a f_\theta(v)$ Then,

$$\text{rel}(v) = \frac{\widehat{\text{se}}(\widehat{q}(v))}{q(v)} \approx O\left(\frac{1}{\sqrt{m}}\right) \left|\frac{\partial \log g_\theta}{\partial d\theta}\right| = O\left(\frac{1}{\sqrt{m}}\right) |v-\theta| \quad \text{Normal case}$$

- Nonparametric Case

$$\widehat{g}(t) = \frac{1}{m}\sum_{i=1}^{m} \frac{1}{h_m} K\left(\frac{t-P_i}{h_m}\right)$$

$h_m = cm^{-\beta}$ where $\beta > 1/5$ (undersmooth). Then

$$\text{rel}_v = \frac{c}{m^{(1-\beta)/2}\sqrt{g(v)}}.$$