Testing for sparse normal means: an update

Jon A. Wellner

University of Washington, visiting Heidelberg

- joint work with Leah Jager,
  U. S. Naval Academy
- Talk at:
  Statisches Kolloquium
  Universität Dusseldorf

  Dusseldorf, Germany
  16 April, 2012
- *Email: jaw@stat.washington.edu*
  http: //www.stat.washington.edu/jaw/jaw.research.html

Outline

- Testing problems for normal means

Outline

- Testing problems for normal means
- Detection boundaries and Tukey's "higher criticism" statistic

Outline

- Testing problems for normal means
- Detection boundaries and Tukey's "higher criticism" statistic
- A new family of statistics via phi-divergences

Outline

- Testing problems for normal means
- Detection boundaries and Tukey's "higher criticism" statistic
- A new family of statistics via phi-divergences
- Beyond normality: generalized Gaussian distributions and ...

Outline

- Testing problems for normal means
- Detection boundaries and Tukey's "higher criticism" statistic
- A new family of statistics via phi-divergences
- Beyond normality: generalized Gaussian distributions and ...
- Donoho - Jin power heuristics

Outline

- Testing problems for normal means
- Detection boundaries and Tukey's "higher criticism" statistic
- A new family of statistics via phi-divergences
- Beyond normality: generalized Gaussian distributions and ...
- Donoho - Jin power heuristics
- Günther Walther's statistic(s)

Outline

- Testing problems for normal means
- Detection boundaries and Tukey's "higher criticism" statistic
- A new family of statistics via phi-divergences
- Beyond normality: generalized Gaussian distributions and ...
- Donoho - Jin power heuristics
- Günther Walther's statistic(s)
- Problems and Questions

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

$$H_{0,i} : X_i \sim N(0,1)$$

versus

$$H_{1,i} : \; X_i \sim N(\mu_i, 1) \; \text{ with } \; \mu_i > 0.$$

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

  $$H_{0,i} : X_i \sim N(0,1)$$

  versus

  $$H_{1,i} : \; X_i \sim N(\mu_i, 1) \;\; \text{with} \;\; \mu_i > 0.$$

- Sparsity: proportion $\epsilon_n \equiv n^{-1} \#\{i \le n : \; \mu_i > 0\}$ is small;
  $\epsilon_n \sim n^{-\beta}$ with $0 < \beta < 1$.

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

$$H_{0,i} : X_i \sim N(0,1)$$

  versus

$$H_{1,i} : \; X_i \sim N(\mu_i, 1) \; \text{ with } \; \mu_i > 0.$$

- Sparsity: proportion $\epsilon_n \equiv n^{-1} \#\{i \leq n : \; \mu_i > 0\}$ is small; $\epsilon_n \sim n^{-\beta}$ with $0 < \beta < 1$.
- Three questions (in increasing order of difficulty):

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

$$H_{0,i} : X_i \sim N(0,1)$$

  versus

$$H_{1,i} : \; X_i \sim N(\mu_i, 1) \;\text{ with }\; \mu_i > 0.$$

- Sparsity: proportion $\epsilon_n \equiv n^{-1}\#\{i \le n : \; \mu_i > 0\}$ is small;
  $\epsilon_n \sim n^{-\beta}$ with $0 < \beta < 1$.
- Three questions (in increasing order of difficulty):
  - Q1: Can we tell if at least one null hypothesis is false?

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

$$H_{0,i} : X_i \sim N(0,1)$$

  versus

$$H_{1,i} : \; X_i \sim N(\mu_i, 1) \;\text{ with }\; \mu_i > 0.$$

- Sparsity: proportion $\epsilon_n \equiv n^{-1}\#\{i \leq n : \; \mu_i > 0\}$ is small;
  $\epsilon_n \sim n^{-\beta}$ with $0 < \beta < 1$.
- Three questions (in increasing order of difficulty):
  - Q1: Can we tell if at least one null hypothesis is false?
  - Q2: What is the proportion of false null hypotheses?

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

$$H_{0,i} : X_i \sim N(0,1)$$

  versus

$$H_{1,i} : \; X_i \sim N(\mu_i, 1) \;\; \text{with} \;\; \mu_i > 0.$$

- Sparsity: proportion $\epsilon_n \equiv n^{-1} \#\{i \leq n : \;\mu_i > 0\}$ is small;
  $\epsilon_n \sim n^{-\beta}$ with $0 < \beta < 1$.
- Three questions (in increasing order of difficulty):
  - Q1: Can we tell if at least one null hypothesis is false?
  - Q2: What is the proportion of false null hypotheses?
  - Q3: Which null hypotheses are false?

1. Testing problems for sparse normal means

- Initial setting: multiple testing of normal means
  For $i = 1, \ldots, n$ consider testing

$$H_{0,i} : X_i \sim N(0,1)$$

  versus

$$H_{1,i} : \; X_i \sim N(\mu_i, 1) \;\; \text{with} \;\; \mu_i > 0.$$

- Sparsity: proportion $\epsilon_n \equiv n^{-1}\#\{i \le n : \; \mu_i > 0\}$ is small; $\epsilon_n \sim n^{-\beta}$ with $0 < \beta < 1$.
- Three questions (in increasing order of difficulty):
  - Q1: Can we tell if at least one null hypothesis is false?
  - Q2: What is the proportion of false null hypotheses?
  - Q3: Which null hypotheses are false?
- Main focus here: Q1.

- Previous work: Q1: is there any signal?
  - Ingster (1997, 1999)
  - Jin (2004)
  - Donoho and Jin (2004)
  - Jager and Wellner (2007)
  - Hall and Jin (2007)

- Previous work: Q1: is there any signal?
  - Ingster (1997, 1999)
  - Jin (2004)
  - Donoho and Jin (2004)
  - Jager and Wellner (2007)
  - Hall and Jin (2007)
- Previous work: Q2: What is the proportion of non-null hypotheses?
  - Swanepoel (1999)
  - Efron, Tibshirani, Storey, and Tusher (2001)
  - Meinshausen and Rice (2006)
  - Jin and Cai (2007)

- Previous work: Q1: is there any signal?
  - Ingster (1997, 1999)
  - Jin (2004)
  - Donoho and Jin (2004)
  - Jager and Wellner (2007)
  - Hall and Jin (2007)
- Previous work: Q2: What is the proportion of non-null hypotheses?
  - Swanepoel (1999)
  - Efron, Tibshirani, Storey, and Tusher (2001)
  - Meinshausen and Rice (2006)
  - Jin and Cai (2007)
- Previous work: Q3: Where is the signal and how big is it?
  - Benjamini and Hochberg (1995)
  - Efron, Tibshirani, Storey, and Tusher (2001)
  - Storey, Dai, and Leek (2005)
  - Donoho and Jin (2006)

2. Detection boundaries and

Tukey's "higher criticism statistic

- Change of setting: Ingster - Donoho - Jin testing problem

2. Detection boundaries and Tukey's "higher criticism statistic

- Change of setting: Ingster - Donoho - Jin testing problem
- Suppose $Y_1, \ldots, Y_n$ i.i.d. G on $\mathbb{R}$

2. Detection boundaries and

Tukey's "higher criticism statistic

- Change of setting: Ingster - Donoho - Jin testing problem
- Suppose $Y_1, \ldots, Y_n$ i.i.d. G on $\mathbb{R}$
- test $H: \; G = N(0,1)$ versus
  $H_1: G = (1-\epsilon)N(0,1) + \epsilon N(\mu,1)$, and, in particular, against

$$H_1^{(n)}: G = (1-\epsilon_n)N(0,1) + \epsilon_n N(\mu_n,1).$$

for $\epsilon_n = n^{-\beta}, \quad \mu_n = \sqrt{2r\log n}$
$0 < \beta < 1$, $0 < r < 1$.

2. Detection boundaries and

Tukey's "higher criticism statistic

- Change of setting: Ingster - Donoho - Jin testing problem
- Suppose $Y_1, \ldots, Y_n$ i.i.d. G on $\mathbb{R}$
- test $H: \;\; G = N(0,1)$ versus
  $H_1 : G = (1-\epsilon)N(0,1) + \epsilon N(\mu, 1)$, and, in particular, against

$$H_1^{(n)} : G = (1-\epsilon_n)N(0,1) + \epsilon_n N(\mu_n, 1).$$

  for $\epsilon_n = n^{-\beta}, \quad \mu_n = \sqrt{2r \log n}$
  $0 < \beta < 1$, $0 < r < 1$.
- Let $\Phi(z) \equiv P(Z \le z) = \int_{-\infty}^{z} (2\pi)^{-1/2} \exp(-x^2/2) dx$,
  $Z \sim N(0,1)$.

- transform to $X_i \equiv 1 - \Phi(Y_i) \in [0,1]$ i.i.d.

$$F = 1 - G(\Phi^{-1}(1 - \cdot)).$$

- transform to $X_i \equiv 1 - \Phi(Y_i) \in [0,1]$ i.i.d.

$$F = 1 - G(\Phi^{-1}(1 - \cdot)).$$

- Then the testing problem becomes: test

$$H_0 : F = F_0 = U(0,1) \qquad \text{versus}$$

$$\begin{aligned} H_1^{(n)} : F(u) &= u + \epsilon_n\{(1-u) - \Phi(\Phi^{-1}(1-u) - \mu_n)\} \\ &= (1-\epsilon_n)u + \epsilon_n\{1 - \Phi(\Phi^{-1}(1-u) - \mu_n)\} \end{aligned}$$

- transform to $X_i \equiv 1 - \Phi(Y_i) \in [0,1]$ i.i.d.

$$F = 1 - G(\Phi^{-1}(1 - \cdot)).$$

- Then the testing problem becomes: test

$$H_0 : F = F_0 = U(0,1) \qquad \text{versus}$$

$$H_1^{(n)} : F(u) = u + \epsilon_n\{(1-u) - \Phi(\Phi^{-1}(1-u) - \mu_n)\}$$

$$= (1-\epsilon_n)u + \epsilon_n\{1 - \Phi(\Phi^{-1}(1-u) - \mu_n)\}$$

- Test statistics: Donoho-Jin

$$HC_n^* \equiv \sup_{X_{(1)} \le u < X_{([n/2])}} \frac{\sqrt{n}(\mathbb{F}_n(u) - u)}{\sqrt{u(1-u)}}$$

$\equiv$ Tukey's "higher criticism statistic"

where $\mathbb{F}_n(u) \equiv n^{-1} \sum_{i=1}^{n} 1_{[0,u]}(X_i)$ = empirical distribution function of the X_i's.

- Optimal detection boundary $\rho^*(\beta)$ defined by:

$$\rho^*(\beta) = \begin{cases} \beta - 1/2, & 1/2 < \beta \leq 3/4 \\ (1 - \sqrt{1-\beta})^2, & 3/4 < \beta < 1 \end{cases}$$

- Optimal detection boundary $\rho^*(\beta)$ defined by:

$$\rho^*(\beta) = \begin{cases} \beta - 1/2, & 1/2 < \beta \leq 3/4 \\ (1 - \sqrt{1-\beta})^2, & 3/4 < \beta < 1 \end{cases}$$

- **Theorem 1:** (Donoho - Jin, 2004). For $r > \rho^*(\beta)$ the test based on HC_n^* is size and power consistent for testing H_0 versus $H_1^{(n)}$.

- Optimal detection boundary $\rho^*(\beta)$ defined by:

$$\rho^*(\beta) = \begin{cases} \beta - 1/2, & 1/2 < \beta \leq 3/4 \\ (1 - \sqrt{1-\beta})^2, & 3/4 < \beta < 1 \end{cases}$$

- **Theorem 1:** (Donoho - Jin, 2004). For $r > \rho^*(\beta)$ the test based on HC_n^* is size and power consistent for testing H_0 versus $H_1^{(n)}$.
- With $h_n(\alpha_n) = \sqrt{2 \log\log(n)}(1 + o(1))$

$$P_{H_0}(HC_n^* > h_n(\alpha_n)) = \alpha_n \to 0, \qquad \text{and}$$
$$P_{H_1^{(n)}}(HC_n^* > h_n(\alpha_n)) \to 1, \qquad \text{as } n \to \infty.$$

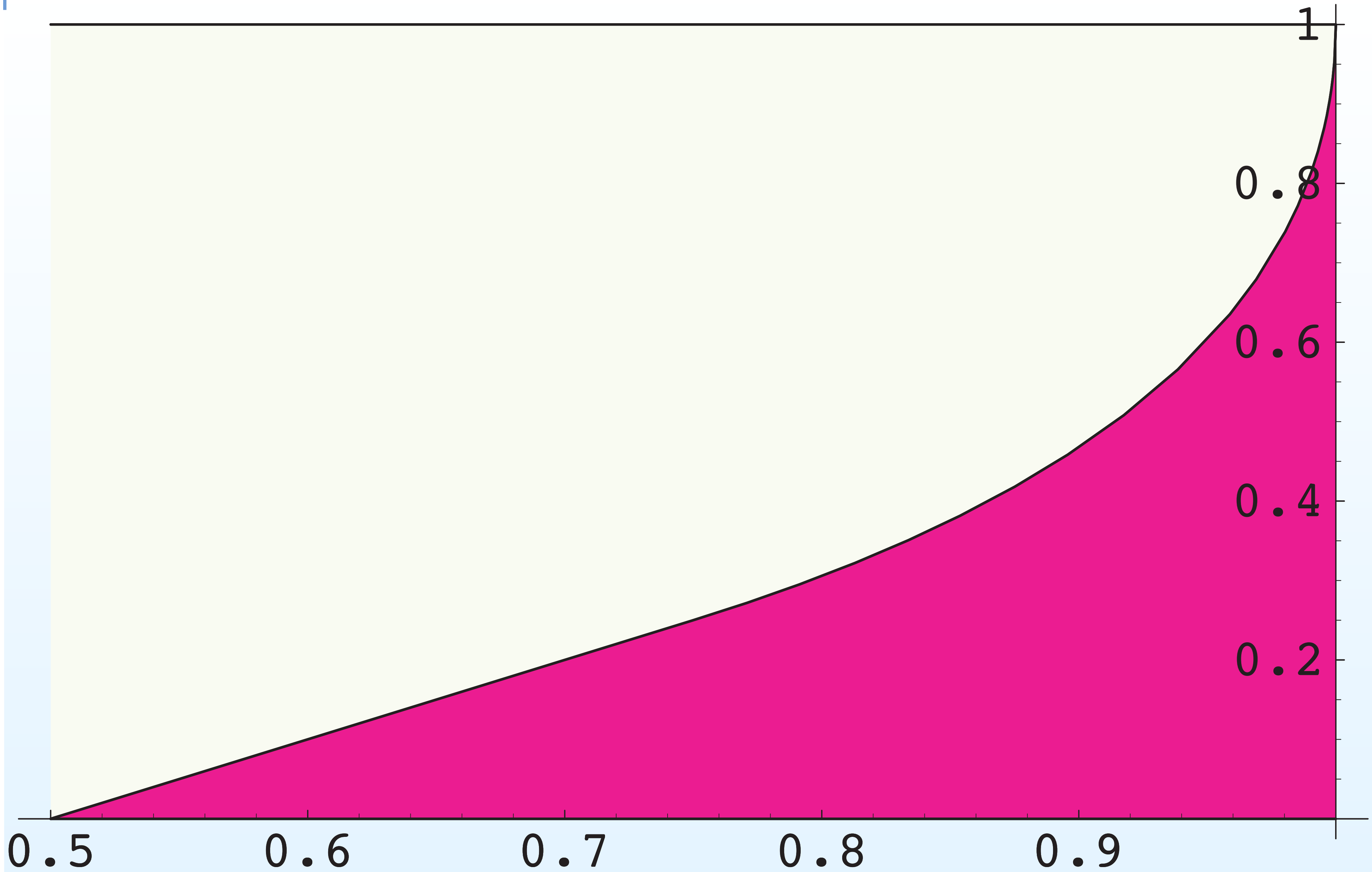


Figure 1. Detection boundary: $r > \rho^*(\beta)$ detectable

Some alternative statistics:

- Berk-Jones (1979) test statistic:

$$R_n \equiv \sup_x \log \lambda_n(x) = \sup_x K(\mathbb{F}_n(x), F_0(x)) \quad \text{with}$$

$$K(u,v) \equiv u \log\left(\frac{u}{v}\right) + (1-u)\log\left(\frac{1-u}{1-v}\right)$$

Some alternative statistics:

- Berk-Jones (1979) test statistic:

$$R_n \equiv \sup_x \log \lambda_n(x) = \sup_x K(\mathbb{F}_n(x), F_0(x)) \text{ with}$$

$$K(u,v) \equiv u \log\left(\frac{u}{v}\right) + (1-u)\log\left(\frac{1-u}{1-v}\right)$$

- Adaptation to one-sided $p-$value setting:

$$BJ_n^+ \equiv n \sup_{X_{(1)} \le u \le 1/2} K^+(\mathbb{F}_n(u), u)$$

where

$$K^+(u,v) \equiv \begin{cases} K(u,v), & \text{if } 0 < v < u < 1, \\ 0, & \text{if } 0 \le u \le v \le 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

- **Theorem 2:** (Donoho - Jin, 2004). For $r > \rho^*(\beta)$ the test based on BJ_n^+ is size and power consistent for testing H_0 versus $H_1^{(n)}$; i.e. with $h_n(\alpha_n) = \sqrt{2 \log \log(n)}(1 + o(1))$

$$P_{H_0}(BJ_n^+ > h_n(\alpha_n)) = \alpha_n \to 0, \qquad \text{and}$$
$$P_{H_1^{(n)}}(BJ_n^+ > h_n(\alpha_n)) \to 1, \qquad \text{as } n \to \infty.$$

3. A new family of statistics via phi-divergences

A family of test statistics connecting "Higher criticism" and Berk-Jones:

- For $s \in \mathbb{R}$, $x \geq 0$ define

$$\phi_s(x) = \begin{cases} \frac{1-s+sx-x^s}{s(1-s)}, & s \neq 0,1 \\ x \log x - x + 1, & s = 1 \\ -\log x + x - 1, & s = 0 \, . \end{cases}$$

3. A new family of statistics via phi-divergences

A family of test statistics connecting "Higher criticism" and Berk-Jones:

- For $s \in \mathbb{R}$, $x \geq 0$ define

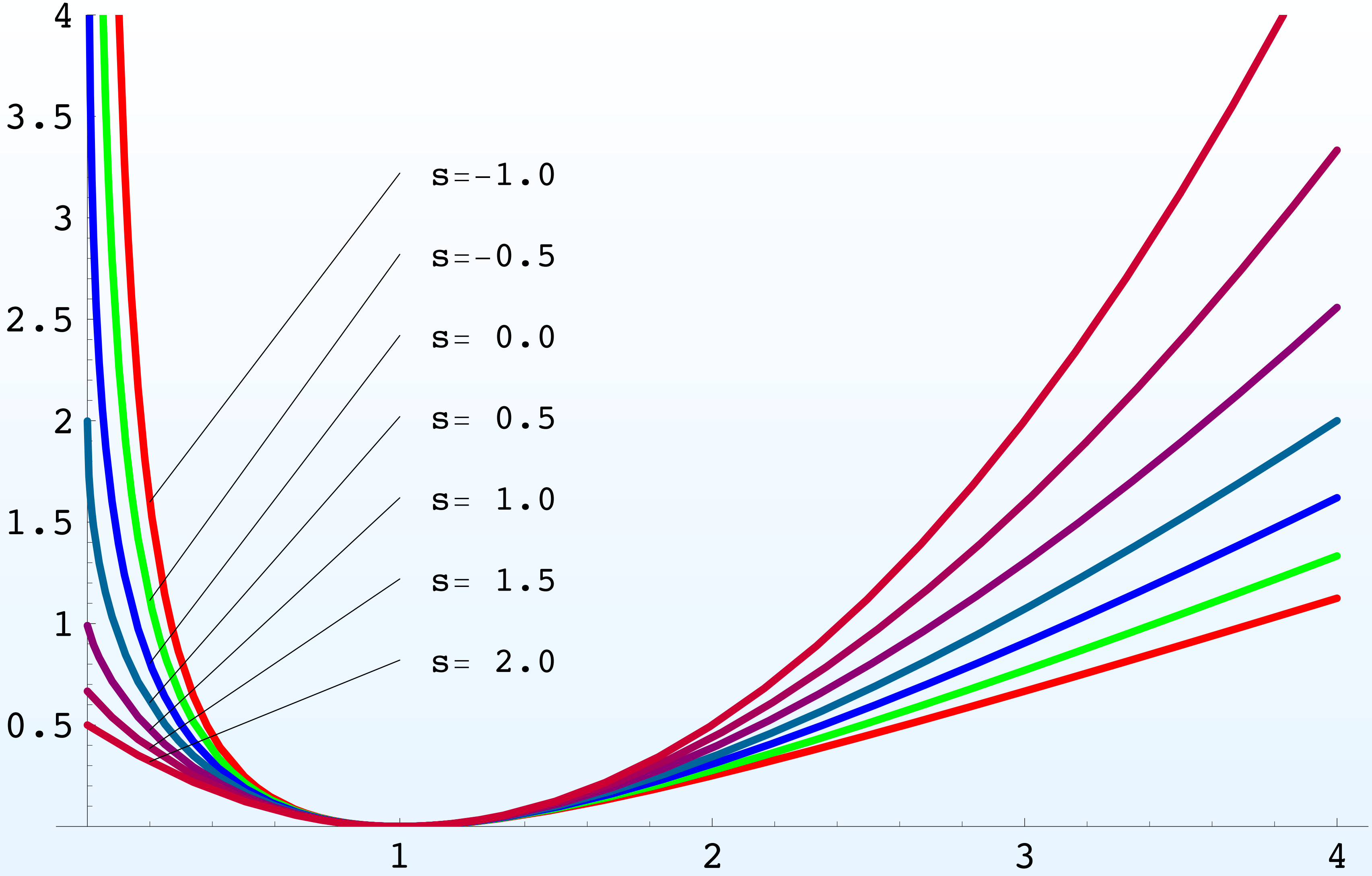
$$\phi_s(x) = \begin{cases} \frac{1-s+sx-x^s}{s(1-s)}, & s \neq 0,1 \\ x \log x - x + 1, & s = 1 \\ -\log x + x - 1, & s = 0\,. \end{cases}$$

- Then define

$$K_s(u,v) = v\phi_s(u/v) + (1-v)\phi_s((1-u)/(1-v))\,.$$

- Special cases:

$$
\begin{aligned}
K_1(u,v) &= K(u,v) \\
&= u\log(u/v) + (1-u)\log((1-u)/(1-v)) \\
K_0(u,v) &= K(v,u) \\
K_2(u,v) &= \frac{1}{2}\frac{(u-v)^2}{v(1-v)} \\
K_{-1}(u,v) &= K_2(v,u) = \frac{1}{2}\frac{(u-v)^2}{u(1-u)} \\
K_{1/2}(u,v) &= 2\{(\sqrt{u}-\sqrt{v})^2 + (\sqrt{1-u}-\sqrt{1-v})^2\} \\
&= 4\{1-\sqrt{uv}-\sqrt{(1-u)(1-v)}\}.
\end{aligned}
$$



4
3.5
3
2.5
2
1.5
1
0.5
s=-1.0
s=-0.5
s= 0.0
s= 0.5
s= 1.0
s= 1.5
s= 2.0
1
2
3
4

- The new family of statistics:

$$S_n(s) = \begin{cases} \sup_{x\in\mathbb{R}} K_s(\mathbb{F}_n(x), F_0(x)), & s \geq 1 \\ \sup_{x\in[X_{(1)},X_{(n)})} K_s(\mathbb{F}_n(x), F_0(x)), & s < 1\,, \end{cases}$$

- The new family of statistics:

$$S_n(s) = \begin{cases} \sup_{x\in\mathbb{R}} K_s(\mathbb{F}_n(x), F_0(x)), & s \geq 1 \\ \sup_{x\in[X_{(1)},X_{(n)})} K_s(\mathbb{F}_n(x), F_0(x)), & s < 1\,, \end{cases}$$

- Thus, with $F_0(x) = x$,

$$S_n(1) = R_n, \qquad S_n(0) = \text{"reversed" Berk-Jones} \equiv \widetilde{R}_n$$

$$S_n(2) = \frac{1}{2}\sup_{x\in\mathbb{R}} \frac{(\mathbb{F}_n(x) - x)^2}{x(1-x)},$$

$$S_n(-1) = \frac{1}{2}\sup_{x\in[X_{(1)},X_{(n)})} \frac{(\mathbb{F}_n(x) - x)^2}{\mathbb{F}_n(x)(1-\mathbb{F}_n(x))}$$

$$S_n(1/2) = 4 \sup_{x\in[X_{(1)},X_{(n)})} \{1 - \sqrt{\mathbb{F}_n(x)x} - \sqrt{(1-\mathbb{F}_n(x))(1-x)}\}$$

- Version of the statistics for one-sided $p-$value setting:

$$S_n^+ \equiv n \sup_{X_{(1)} \leq u \leq 1/2} K_s^+(\mathbb{F}_n(u), u)$$

where

$$K_s^+(u,v) \equiv \begin{cases} K_s(u,v), & \text{if } 0 < v < u < 1, \\ 0, & \text{if } 0 \leq u \leq v \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

- Version of the statistics for one-sided $p-$value setting:

$$S_n^+ \equiv n \sup_{X_{(1)} \le u \le 1/2} K_s^+(\mathbb{F}_n(u), u)$$

where

$$K_s^+(u,v) \equiv \begin{cases} K_s(u,v), & \text{if } 0 < v < u < 1, \\ 0, & \text{if } 0 \le u \le v \le 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

- **Theorem:** (Jager - Wellner, 2007). For $r > \rho^*(\beta)$ the tests based on $S_n^+(s)$ with $-1 \le s \le 2$ are size and power consistent for testing H_0 versus $H_1^{(n)}$; i.e. With $s_n(\alpha_n) = \log\log(n)(1 + o(1))$

$$P_{H_0}(S_n^+ > s_n(\alpha_n)) = \alpha_n \to 0, \qquad \text{and}$$
$$P_{H_1^{(n)}}(S_n^+ > s_n(\alpha_n)) \to 1, \qquad \text{as } n \to \infty.$$

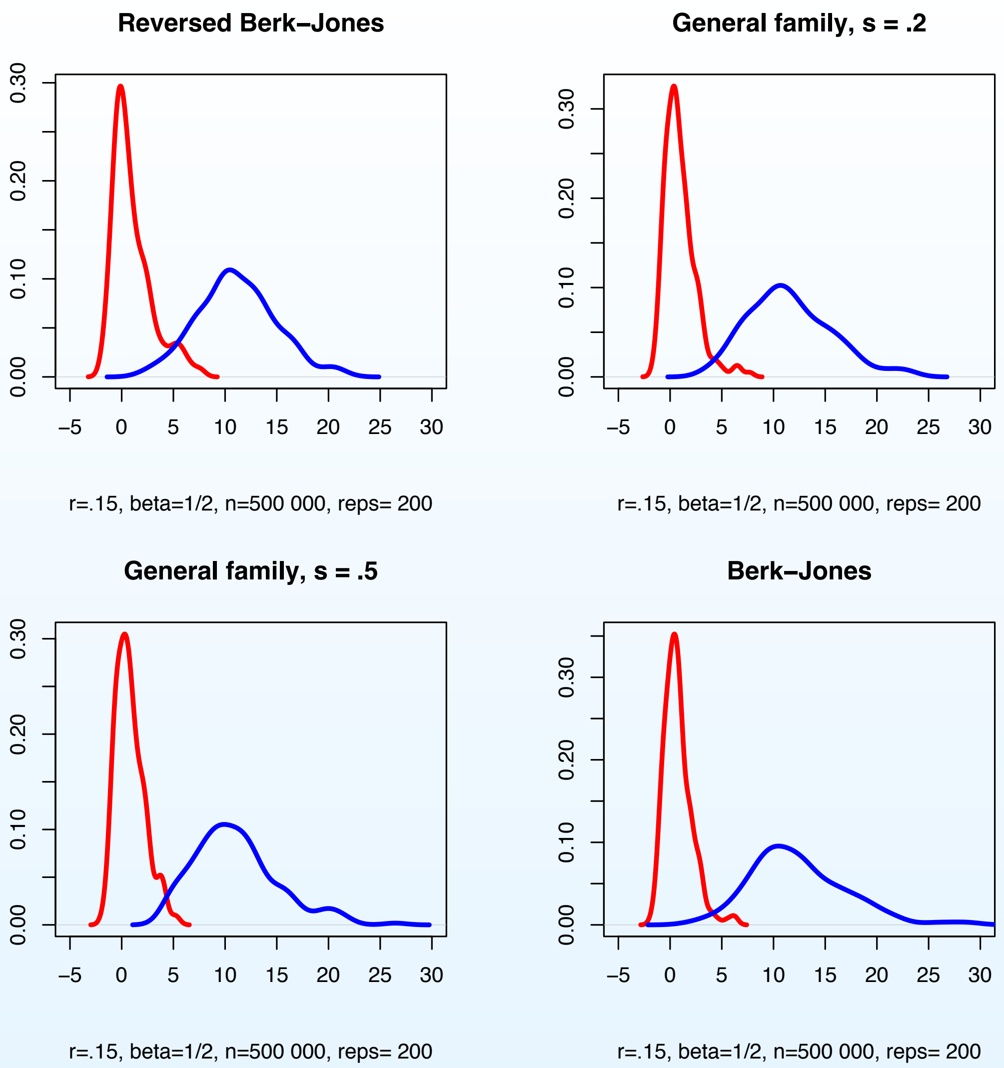


Figure 2. Separation plots: $n = 5 \times 10^5$, $r = .15$, $\beta = 1/2$
Smoothed histograms of reps $= 200$ of the statistics under the null hypothesis and the the alternative hypothesis

4. Beyond normality:

generalized Gaussian distributions, ...

- Donoho and Jin (2004) also computed detection boundaries for sparse mixtures of “Generalized Gaussian” or Subbotin distributions: $X \sim GN_\gamma(\mu)$ has density function

$$f_{\gamma,\mu}(x) = \frac{1}{C_\gamma} \exp\left(-\frac{|x-\mu|^\gamma}{\gamma}\right), \quad C_\gamma = 2\Gamma(1/\gamma)\gamma^{1/\gamma - 1}.$$

4. Beyond normality:

generalized Gaussian distributions, ...

- Donoho and Jin (2004) also computed detection boundaries for sparse mixtures of "Generalized Gaussian" or Subbotin distributions: $X \sim GN_\gamma(\mu)$ has density function

$$f_{\gamma,\mu}(x) = \frac{1}{C_\gamma} \exp\left(- \frac{|x-\mu|^\gamma}{\gamma} \right), \quad C_\gamma = 2\Gamma(1/\gamma)\gamma^{1/\gamma - 1}.$$

- Suppose $Y_1, \ldots, Y_n$ i.i.d. G on $\mathbb{R}$.

4. Beyond normality:

generalized Gaussian distributions, ...

- Donoho and Jin (2004) also computed detection boundaries for sparse mixtures of “Generalized Gaussian” or Subbotin distributions: $X \sim GN_\gamma(\mu)$ has density function

$$f_{\gamma,\mu}(x) = \frac{1}{C_\gamma} \exp\left(-\frac{|x-\mu|^\gamma}{\gamma}\right), \quad C_\gamma = 2\Gamma(1/\gamma)\gamma^{1/\gamma-1}.$$

- Suppose $Y_1, \ldots, Y_n$ i.i.d. G on $\mathbb{R}$.
- Test $H_0 : G = GN_\gamma(0)$ versus $H_1^{(n)} : \; G = (1-\epsilon_n)GN_\gamma(0) + \epsilon_n GN_\gamma(\mu_n)$ where

$$\epsilon_n = n^{-\beta}, \qquad \mu_{\gamma,n} = (\gamma r \log n)^{1/\gamma},$$

where $1/2 < \beta < 1, \quad 0 < r < 1$.

- Detection boundary for $1 < \gamma \leq 2$:

$$\rho_\gamma^*(\beta) = \begin{cases} (2^{1/(\gamma-1)} - 1)^{\gamma-1}(\beta - 1/2), & 1/2 < \beta \leq 1 - 2^{-\gamma/(\gamma-1)}, \\ (1 - (1-\beta)^{1/\gamma})^\gamma, & 1 - 2^{-\gamma/(\gamma-1)} \leq \beta < 1. \end{cases}$$

- Detection boundary for $1 < \gamma \leq 2$:

$$\rho^*_\gamma(\beta) = \begin{cases} (2^{1/(\gamma-1)} - 1)^{\gamma-1}(\beta - 1/2), & 1/2 < \beta \leq 1 - 2^{-\gamma/(\gamma-1)}, \\ (1 - (1-\beta)^{1/\gamma})^\gamma, & 1 - 2^{-\gamma/(\gamma-1)} \leq \beta < 1. \end{cases}$$

- Detection boundary for $0 < \gamma \leq 1$:

$$\rho^*_\gamma(\beta) = 2(\beta - 1/2), \qquad 1/2 < \beta < 1.$$

Note: The detection boundary is the same for all for $0 < \gamma \leq 1$!

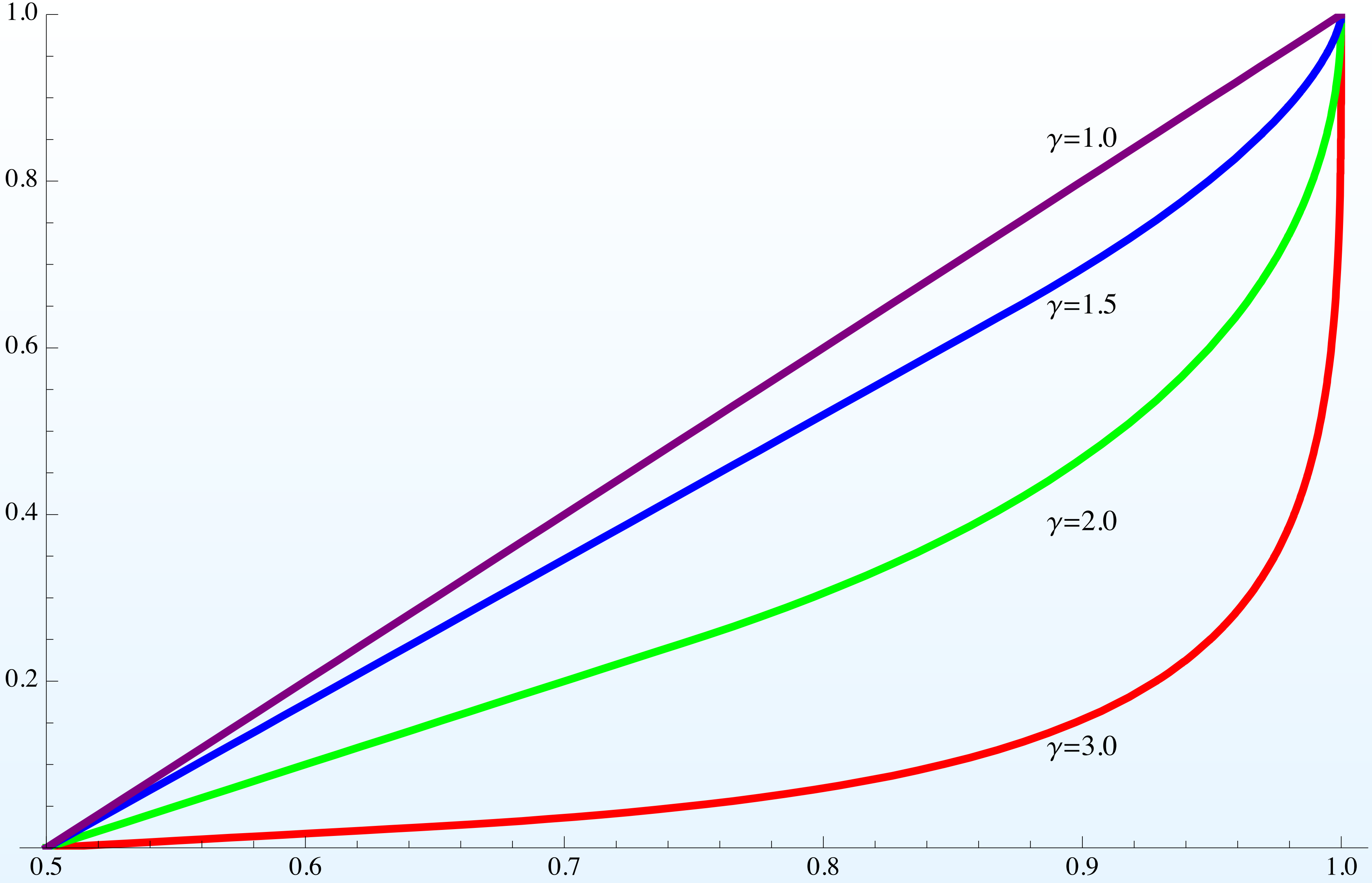


Figure 3. Detection boundaries for GN testing problem, $\gamma \in \{1, 1.5, 2, 3\}$.

- **Theorem:** (Donoho - Jin, 2004). For the higher criticism test statistic applied to the p-values $p_i \equiv P(GN_\gamma(0) > Y_i)$, $i = 1, \ldots, n$. Then the detection boundary $\rho_{HC,\gamma}$ for this procedure is the same as the efficient detection boundary:

$$\rho_{HC,\gamma}(\beta) = \rho^*_\gamma(\beta), \qquad 1/2 < \beta < 1.$$

- **Theorem:** (Donoho - Jin, 2004). For the higher criticism test statistic applied to the p-values $p_i \equiv P(GN_\gamma(0) > Y_i)$, $i = 1, \ldots, n$. Then the detection boundary $\rho_{HC,\gamma}$ for this procedure is the same as the efficient detection boundary:

$$\rho_{HC,\gamma}(\beta) = \rho^*_\gamma(\beta), \qquad 1/2 < \beta < 1.$$

- Similar theorem for χ^2_ν mixtures.

5. Donoho - Jin Power heuristics

What part of the sample contributes to the power?

- When $\beta \in [3/4, 1)$, the strongest evidence against H_0 is found near the maximum of the observations; i.e. at the smallest $p-$values.

5. Donoho - Jin Power heuristics

What part of the sample contributes to the power?

- When $\beta \in [3/4, 1)$, the strongest evidence against H_0 is found near the maximum of the observations; i.e. at the smallest $p-$values.
- When $\beta \in (1/2, 3/4]$ other $p-$values beyond the smallest contribute to the power.

5. Donoho - Jin Power heuristics

What part of the sample contributes to the power?

- When $\beta \in [3/4, 1)$, the strongest evidence against H_0 is found near the maximum of the observations; i.e. at the smallest $p-$values.
- When $\beta \in (1/2, 3/4]$ other $p-$values beyond the smallest contribute to the power.
- Since the higher criticism statistic HC_n^* gives more weight to the smaller $p-$values, we expect it to have higher power for alternatives with $\beta \in [3/4, 1)$

5. Donoho - Jin Power heuristics

What part of the sample contributes to the power?

- When $\beta \in [3/4, 1)$, the strongest evidence against H_0 is found near the maximum of the observations; i.e. at the smallest $p-$values.
- When $\beta \in (1/2, 3/4]$ other $p-$values beyond the smallest contribute to the power.
- Since the higher criticism statistic HC_n^* gives more weight to the smaller $p-$values, we expect it to have higher power for alternatives with $\beta \in [3/4, 1)$
- Since the Berk-Jones (supremum of pointwise likelihood ratios) statistic BJ_n^+ gives less weight to the very smallest $p-$values, we expect that it might have higher power for $\beta \in (1/2, 3/4]$.

6. Walther's weighted likelihood ratio statistic

Let

$$\log LR_n(t) = \begin{cases} n\{\mathbb{F}_n(t)\log\frac{\mathbb{F}_n(t)}{t} + (1-\mathbb{F}_n(t))\log\frac{1-\mathbb{F}_n(t)}{1-t}, & \text{if } 0 < t < \mathbb{F}_n(t) \\ 0, & \text{otherwise.} \end{cases}$$

Thus

$$BJ_n^+ = \max_{1\le i\le n/2} \log LR_{n,i}$$

where

$$\begin{aligned} \log LR_{n,i} &\equiv \log LR_n(p_{(i)}) \\ &= \left\{ i\log\left(\frac{i}{np_{(i)}}\right) + (n-i)\log\left(\frac{1-i/n}{1-p_{(i)}}\right)\right\} 1\{p_{(i)} < i/n\}. \end{aligned}$$

Start with a uniform prior on $\beta \in [1/2, 1)$. Since the smallest $p-$value has most of the information for $\beta \in [3/4, 1)$, collapse the weight for this interval to weight $1/2$ on the interval $(0, p_{(1)}]$. For $\beta \in (1/2, 3/4)$, the most promising interval to detect alternatives with r close to the detection boundary $\rho^*(\beta) = \beta - 1/2$ is the interval $(0, n^{-4r}]$. Thus given such a β we will use the LR test on the interval $(0, t]$ with $t = n^{-4(\beta-1/2)}$. If $\beta \sim U(1/2, 3/4)$, then $t = n^{-4(\beta-1/2)}$ has density proportional to $1/t$ on $(1/n, 1]$.

Approximation of the resulting posterior integral with the corresponding weighted sum of the LR at the $p_{(i)}$'s, normalized by

$$\sum_{i=2}^{n/2} i^{-1} \approx \log(n/3)$$

yields the Average Likelihood Ratio Statistic

$$ALR_n = \frac{1}{2} LR_{n,1} + \frac{1}{2} \sum_{i=2}^{n/2} \frac{1}{i \log(n/3)} LR_{n,i}$$

where

$$LR_{n,i} = \begin{cases} \left(\frac{i}{np_{(i)}}\right)^i \left(\frac{1-i/n}{1-p_{(i)}}\right)^{n-i}, & \text{if } p_{(i)} < i/n, \\ 1, & \text{if otherwise.} \end{cases}$$

Proposition. (Walther) ALR_n attains the optimal detection boundary for the sparse normal means problem.

7. Further problems and challenges

- Exact contiguity results for HC_n and the phi-diverence statistics under (some refinement of) the exact optimal boundary

7. Further problems and challenges

- Exact contiguity results for HC_n and the phi-diverence statistics under (some refinement of) the exact optimal boundary
- Do all the phi-divergence statistics achieve the optimal detection regions for the Generalized Gaussian (Subbotin) sparse mixture model?

7. Further problems and challenges

- Exact contiguity results for HC_n and the phi-diverence statistics under (some refinement of) the exact optimal boundary
- Do all the phi-divergence statistics achieve the optimal detection regions for the Generalized Gaussian (Subbotin) sparse mixture model?
- Limiting null distribution of Walther's Average Likelihood Ratio Statistic

7. Further problems and challenges

- Exact contiguity results for HC_n and the phi-diverence statistics under (some refinement of) the exact optimal boundary
- Do all the phi-divergence statistics achieve the optimal detection regions for the Generalized Gaussian (Subbotin) sparse mixture model?
- Limiting null distribution of Walther's Average Likelihood Ratio Statistic
- Average versions of the Jager-Wellner divergence family statistics?

7. Further problems and challenges

- Exact contiguity results for HC_n and the phi-diverence statistics under (some refinement of) the exact optimal boundary
- Do all the phi-divergence statistics achieve the optimal detection regions for the Generalized Gaussian (Subbotin) sparse mixture model?
- Limiting null distribution of Walther's Average Likelihood Ratio Statistic
- Average versions of the Jager-Wellner divergence family statistics?
- More systematic study of power properties of all these tests.

Vielen Dank!