Signal detection and goodness-of-fit: the Berk-Jones statistics revisited

Jon A. Wellner (Seattle)

INET Big Data Conference

INET Big Data Conference, Cambridge
September 29 - 30, 2015

Based on joint work with:

- Lutz Dümbgen (Bern)
- Leah Jager (U.S. Naval Academy)

Outline

- 1: Introduction: some history
- 2: Testing for sparse normal means: optimal detection boundary
- 3: Consequences and tradeoffs.
- 4: The LIL and strong LIL for Brownian motion and Brownian bridge
- 5: A new class of test statistics:
  LIL adjusted higher criticism & Berk-Jones.
- 6: Power properties and confidence bands.

1. Introduction: some history

- **Setting: classical "goodness - of - fit"**
- $X_1, \ldots, X_n$ i.i.d. with distribution function F
- $\mathbb{F}_n(x) = \frac{1}{n} \sum_{i=1}^{n} 1_{[X_i \leq x]}$
- Test $H: F = F_0$ versus $K: F \neq F_0$, F_0 continuous
- Without loss of generality $F_0(x) = x$, the $U(0,1)$ distribution
- Break hypotheses down into family of pointwise hypotheses: $H_x: \; F(x) = F_0(x)$ versus $K_x: \; F(x) \neq F_0(x)$
- $H = \cap_x H_x$, $K = \cup_x K_x$

1. Introduction: some history

- $n\mathbb{F}_n(x) \sim \text{Binomial}(n, F(x))$.
- Likelihood ratio statistic for testing H_x versus K_x:

$$
\begin{aligned}
\lambda_n(x) &= \frac{\sup_{F(x)} L_n(F(x))}{L_n(F_0(x))} = \frac{L_n(\mathbb{F}_n(x))}{L_n(F_0(x))} \\
&= \frac{\mathbb{F}_n(x)^{n\mathbb{F}_n(x)}(1-\mathbb{F}_n(x))^{n(1-\mathbb{F}_n(x))}}{F_0(x)^{n\mathbb{F}_n(x)}(1-F_0(x))^{n(1-\mathbb{F}_n(x))}} \\
&= \left(\frac{\mathbb{F}_n(x)}{F_0(x)}\right)^{n\mathbb{F}_n(x)} \left(\frac{1-\mathbb{F}_n(x)}{1-F_0(x)}\right)^{n(1-\mathbb{F}_n(x))}
\end{aligned}
$$

1. Introduction: some history

- Thus

$$\begin{aligned} \log\lambda_n(x) &= n\mathbb{F}_n(x)\log\left(\frac{\mathbb{F}_n(x)}{F_0(x)}\right) \\ &\qquad + n(1-\mathbb{F}_n(x))\log\left(\frac{1-\mathbb{F}_n(x)}{1-F_0(x)}\right) \\ &= nK(\mathbb{F}_n(x), F_0(x)) \end{aligned}$$

- $K(u,v) \equiv u\log\left(\frac{u}{v}\right) + (1-u)\log\left(\frac{1-u}{1-v}\right)$,
  Kullback - Leibler "distance"
  Bernoulli(u), Bernoulli(v)

- Berk-Jones (1979) test statistic:
  via S.N. Roy's union intersection principle,

$$R_n \equiv \sup_x n^{-1}\log\lambda_n(x) = \sup_x K(\mathbb{F}_n(x), F_0(x)) \,.$$

1. Introduction: some history

- **History:**
  - ▷ Berk and Jones (1979)
  - ▷ Groeneboom and Shorack (1981)
  - ▷ Shorack and Wellner (1986, p. 786)
  - ▷ Owen (1995): inversion of R_n to get confidence bands; finite - sample distribution via Noé's recursion
  - ▷ Einmahl and McKeague (2002): integral version of R_n
  - ▷ Donoho and Jin (2004): Tukey's "Higher-Criticism" statistic for testing "sparse normal means" model with comparisons to Berk - Jones statistic R_n

2. Testing 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?

- Main focus here: Q1.

2. Testing for sparse normal means

- Previous work: Q1: is there any signal?
  - ▷ Ingster (1997, 1999)
  - ▷ Jin (2004)
  - ▷ Donoho and Jin (2004)
  - ▷ Jager and Wellner (2007)
  - ▷ Hall and Jin (2007)
  - ▷ Cai and Wu (2014)

2. Testing for sparse normal means

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

2. Testing for sparse normal means

- 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 and Jin (2004) proposed

$$HC_n^* \equiv \sup_{X_{(1)} \le u \le X_{([n/2])}} \frac{\sqrt{n}(\mathbb{F}_n(u) - u)}{\sqrt{u(1-u)}}$$
$$\equiv \text{Tukey's ``higher criticism statistic''}$$

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.

- Let $K_2(u,v) = 2^{-1}(u-v)^2/(v(1-v))^2$; then

$$2^{-1}(HC_n^*)^2 = \sup_{X_{(1)} \le x \le X_{[n/2]}} K_2^+(\mathbb{F}_n(x), x)$$

where $K_2^+(u,v) = K_2(u,v)1_{[v \le u]}$.

2. Testing for sparse normal means

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

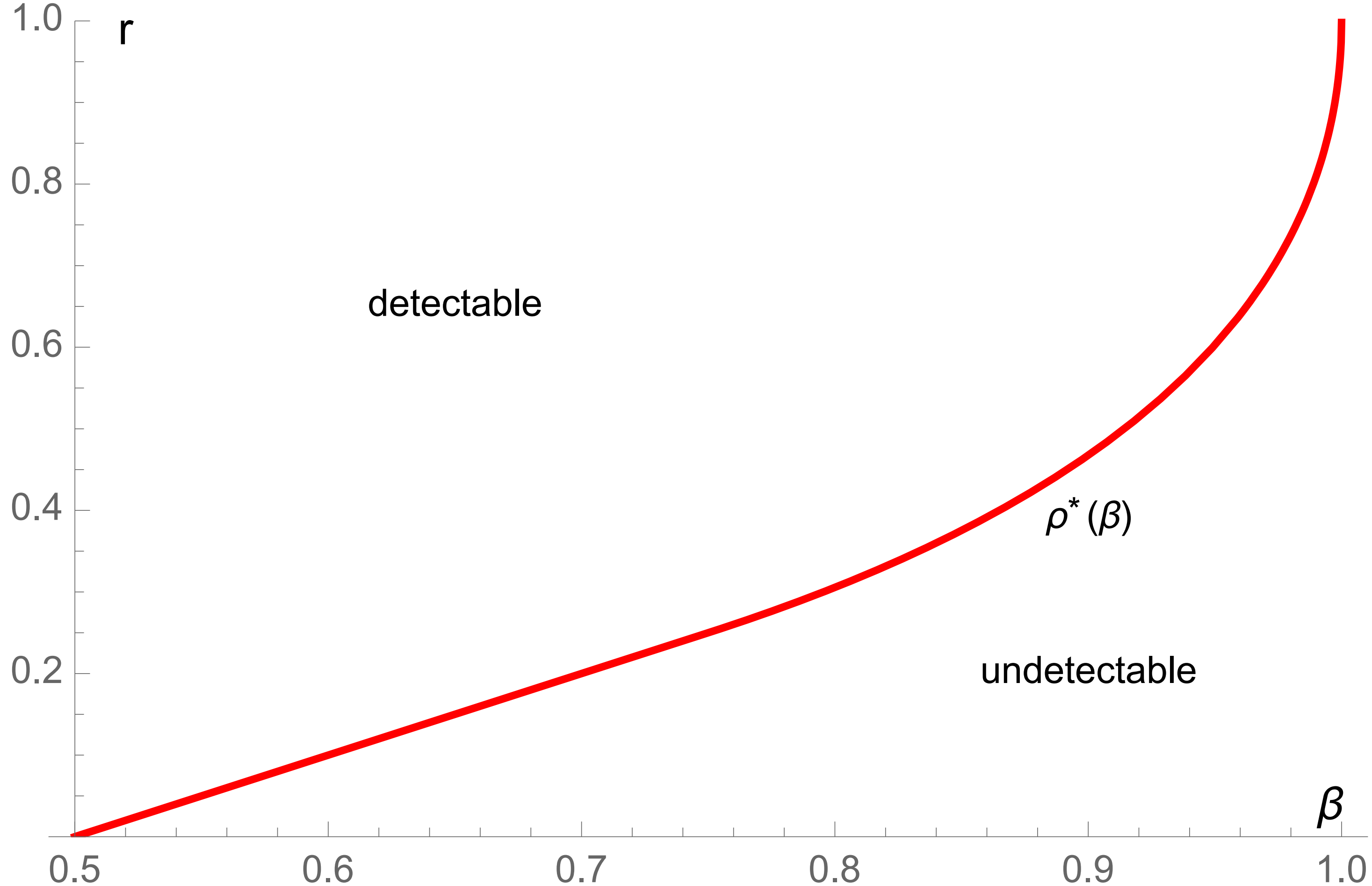
$$\rho^*(\beta) = \begin{cases} \beta - 1/2, & 1/2 < \beta \leq 3/4 \\ (1 - \sqrt{1-\beta})^2, & 3/4 < \beta < 1. \end{cases}$$

- **Theorem 1:** (Donoho - Jin, 2004). For $r > \rho^*(\beta)$ the tests T_n based on $2^{-1}(HC_n^*)^2$ or R_n^+ are both size and power consistent for testing H_0 versus $H_1^{(n)}$.

- With $t_n(\alpha_n) = \text{loglog}(n)(1 + o(1))$

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

2. Testing for sparse normal means



2. Testing for sparse normal means

- A family of statistics via phi-divergences
  - For $s \in \mathbb{R}$, $x \geq 0$ define

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

  - Then define

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

2. Testing for sparse normal means

- Special cases:

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

2. Testing for sparse normal means

- The $\phi-$divergence family of statistics (Jager & W, 2007):

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

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

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

$$S_n(1) = R_n = \text{Berk-Jones statistic}$$

$$S_n(1/2)$$

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

$$S_n(0) = \text{“reversed” Berk-Jones} \equiv \widetilde{R}_n$$

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

2. Testing for sparse normal means

- Null hypothesis distribution theory:
  - Owen (1995) and Jager (2006):
    finite sample critical points via Noé's recursion for $n \leq 3000$
  - For $n \geq 3000$, asymptotic theory via Jaeschke (1979) and Eicker (1979) (cf. SW p. 597 - 615), together with

$$K_s(u,v) \approx 2^{-1}(u-v)^2/[v(1-v)]$$

so

$$nK_s(\mathbb{F}_n(x), x) \approx \frac{1}{2}\frac{n(\mathbb{F}_n(x)-x)^2}{x(1-x)} \equiv \frac{1}{2}\mathbb{Z}_n(x)^2$$

where

$$\mathbb{Z}_n(x) \equiv \frac{\sqrt{n}(\mathbb{F}_n(x)-x)}{\sqrt{x(1-x)}} \to_{f.d.} \frac{\mathbb{U}(x)}{\sqrt{x(1-x)}} \equiv \mathbb{Z}(x)$$

with $\mathbb{U}$ a standard Brownian bridge process on $[0,1]$.

2. Testing for sparse normal means

- Let $r_n \equiv \text{loglog}(n) + (1/2)\text{logloglog}(n) - (1/2)\log(4\pi)$
  $= \text{loglog}(n)(1 + o(1))$.

- **Theorem 1.** If $F = F_0$, the uniform distribution on $[0, 1]$, then for $-1 \le s \le 2$

$$nS_n(s) - r_n \to_d Y_4$$

  where $P(Y_4 \le x) = \exp(-4\exp(-x))$.

- **Theorem 2.** If $F = F_n$, the Ingster - Donoho - Jin sparse normal means model, then for each $s \in [-1, 2]$

$$P_{H_0}(nS_n(s) > t_n(\alpha_n)) \to 0, \qquad \text{and}$$
$$P_{H_1^{(n)}}(nS_n(s) > t_n(\alpha_n)) \to 1 \quad \text{if} \quad r > \rho^*(\beta).$$

3. Consequences and tradeoffs: trouble in the middle!

Although the test statistics $nS_n(s)$ (and their one-sided, one-tailed counterparts) have excellent power behavior against sparse normal means and other "tail" alternatives, we have lost something in the middle:
if $\{F_n\}$ is a sequence of distribution functions satisfying

$$\sqrt{n}(f_n^{1/2} - f_0^{1/2}) \to 2^{-1} a f_0^{1/2}, \quad \text{in } L_2(\lambda)$$
$$\sqrt{n}(F_n - F_0) \to A \quad \text{uniformly}$$

where $A(x) = \int_{-\infty}^{x} a(y) dF_0(y)$, then it is not hard to see that for any $\kappa > 0$

$$P_{F_n}(nS_n(s) - r_n > \kappa) \to 0.$$

Can we find a new family of test statistics which have good power properties for alternatives of both the "tail" and "central" type?

4. The LIL and Strong LIL for Brownian Motion and Bridge

Standard Brownian motion $\mathbb{W} = (\mathbb{W}(t))_{t \geq 0}$

LIL for BM:

$$\limsup_{t \downarrow 0} \frac{\pm \mathbb{W}(t)}{\sqrt{2t \log\log(t^{-1})}} = 1 \quad \text{almost surely,}$$

$$\limsup_{t \uparrow \infty} \frac{\pm \mathbb{W}(t)}{\sqrt{2t \log\log(t)}} = 1 \quad \text{almost surely.}$$

Refined (upper) strong LIL for BM: For arbitrary constants $\nu > 3/2$,

$$\limsup_{t \to \{0,\infty\}} \left(\frac{\mathbb{W}(t)^2}{2t} - \log\log(t + t^{-1}) - \nu \cdot \log\log\log(t + t^{-1}) \right) < 0$$

almost surely.

4. The LIL and Strong LIL for Brownian Motion and Bridge

- Reformulation for standard Brownian bridge $\mathbb{U} = (\mathbb{U}(t))_{t \in (0,1)}$

$$(0,1) \ni t \mapsto \text{logit}(t) := \log\left(\frac{t}{1-t}\right) \in \mathbb{R},$$

$$\mathbb{R} \ni x \mapsto \ell(x) := \frac{e^x}{1+e^x} \in (0,1).$$

$$C(t) := \log\sqrt{1 + \text{logit}(t)^2/2} \approx \text{loglog}(1/t) \quad \text{as } t \downarrow 0,$$

$$D(t) := \log\sqrt{1 + C(t)^2/2} \approx \text{logloglog}(1/t) \quad \text{as } t \downarrow 0.$$

$$\limsup_{t \to \{0,1\}} \left(\frac{\mathbb{U}(t)^2}{2t(1-t)} - C(t) - \nu \cdot D(t) \right) < 0$$

almost surely.

5: A new class of test statistics:
LIL adjusted higher criticism & Berk-Jones.

This suggests a new class of test statistics as follows:

- Fix $s \in [1,2]$ and $\nu > 3/2$.
- Define $X_{n,s}(t) \equiv nK_s(\mathbb{G}_n(t), t)$.
- Set $T_n(s,\nu) \equiv \sup_{0<t<1} \{X_{n,s}(t) - C(t) - \nu D(t)\}$.

If $s \in [-1,1)$ replace the supremum over $(0,1)$ by the sup over $[X_{(1)}, X_{(n)})$.

Theorem. For all $s \in [-1,2]$ and $\nu > 3/2$, if H_0 holds then

$$T_n(s,\nu) \to_d T_\nu \equiv \sup_{0<t<1} \left\{ \frac{\mathbb{U}^2(t)}{2t(1-t)} - C(t) - \nu D(t) \right\}$$

where T_ν is finite almost surely.

Proof: Careful use of strong approximation methods: Csörgő, Csörgő, Horvath, and Mason (1986).

In particular for the LIL adjusted higher criticism & Berk-Jones statistics:

$$T_n(2,\nu) \equiv \sup_{0<t<1} \left\{ \frac{\mathbb{U}_n^2(t)}{2t(1-t)} - C(t) - \nu D(t) \right\} \rightarrow_d T_\nu,$$
$$T_n(1,\nu) \equiv \sup_{0<t<1} \{ nK_1(\mathbb{G}_n(t), t) - C(t) - \nu D(t) \} \rightarrow_d T_\nu.$$

where $K_1(u,v) \equiv K(u,v) = u\log(u/v) + (1-u)\log((1-u)/(1-v))$.

- What about power?
  - ▷ Tail alternatives, e.g. sparse normal means?
  - ▷ Central (or contiguous) alternatives?
- Widths of confidence bands?

6. Confidence bands and power properties

Let $U_1, U_2, \ldots, U_n$ be i.i.d. $\sim$ Unif$[0,1]$. Auxiliary function K : $[0,1] \times (0,1) \to [0,\infty]$,

$$K(x,p) \ := \ x\log\Bigl(\frac{x}{p}\Bigr) + (1-x)\log\Bigl(\frac{1-x}{1-p}\Bigr)$$

i.e. Kullback-Leibler divergence between Bin$(1,x)$ and Bin$(1,p)$.

Two key properties:

$$K(x,p) \ = \ \frac{(x-p)^2}{2p(1-p)}\bigl(1+o(1)\bigr) \quad \text{as } x \to p.$$

$$K(x,p) \le c \quad \text{implies} \quad |x-p| \le \begin{cases} \sqrt{2c\,p(1-p)} + c \\ \sqrt{2c\,x(1-x)} + c \end{cases}$$

6. Confidence bands and power properties

Uniform order statistics

$$0 < U_{n:1} < U_{n:2} < \cdots < U_{n:n} < 1.$$

$$\mathcal{T}_n := \{t_{n1}, t_{n2}, \ldots, t_{nn}\} \quad \text{with} \quad t_{ni} := \mathbb{E}(U_{n:i}) = \frac{i}{n+1}.$$

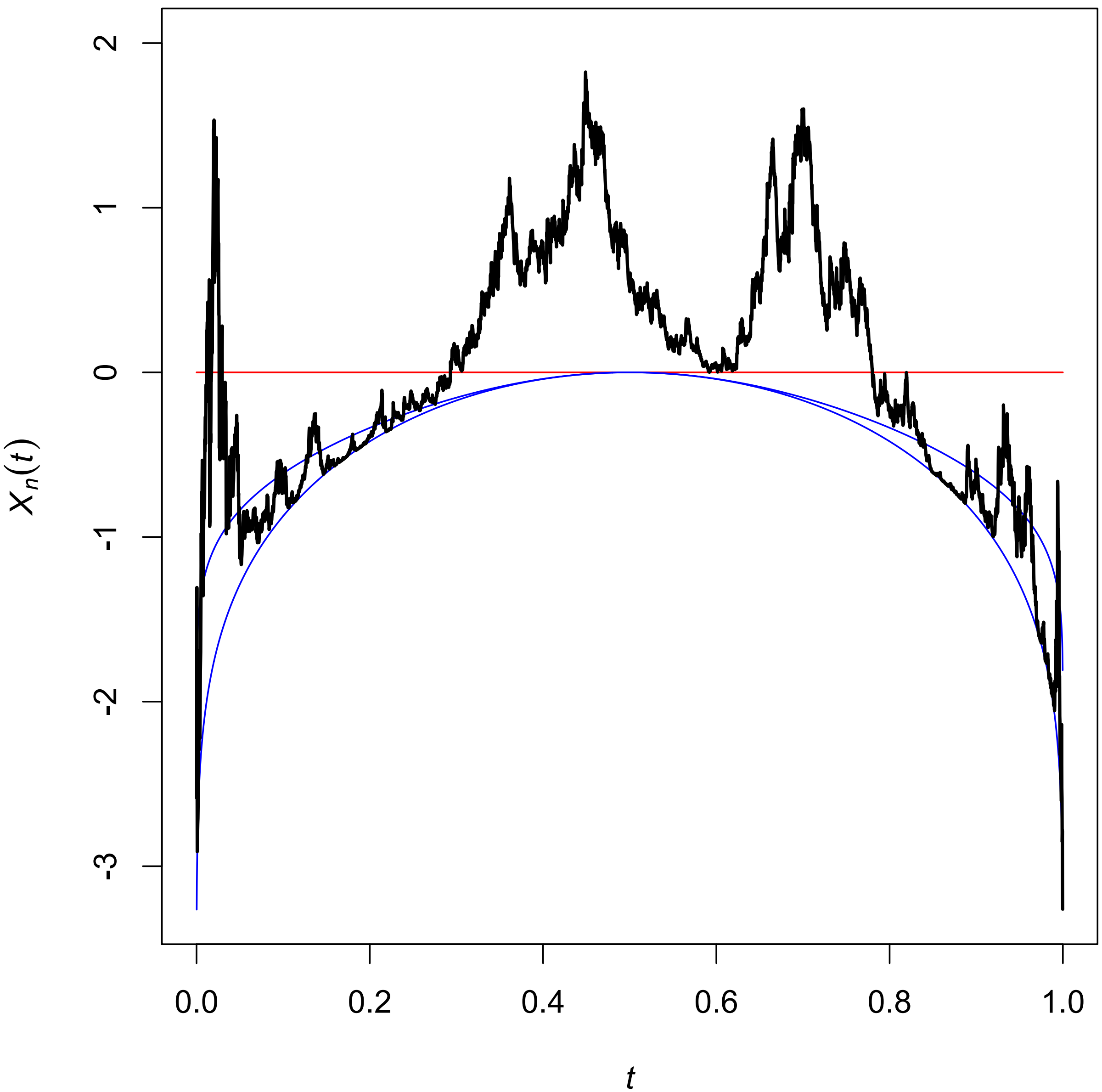
Theorem 2. For the process $\tilde{X}_n = (\tilde{X}_n(t))_{t \in \mathcal{T}_n}$ with

$$\tilde{X}_n(t_{ni}) \equiv (n+1)K(t_{ni}, U_{n:i})$$

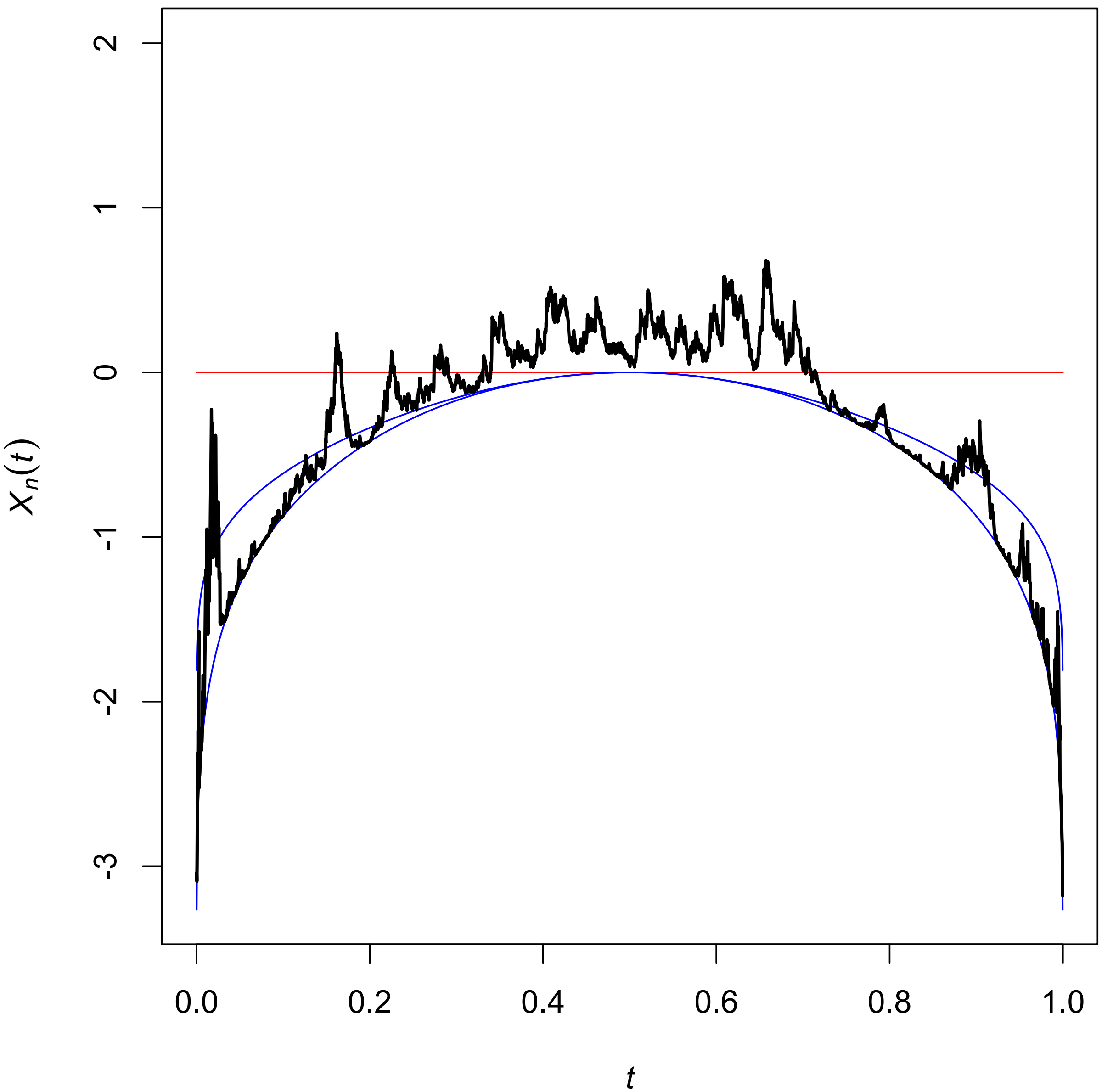
for $\nu > 3/2$ we have

$$\widetilde{T}_n(1,\nu) \equiv \sup_{\mathcal{T}_n} \left\{ \tilde{X}_n - C - \nu D \right\} \to_d T_\nu.$$

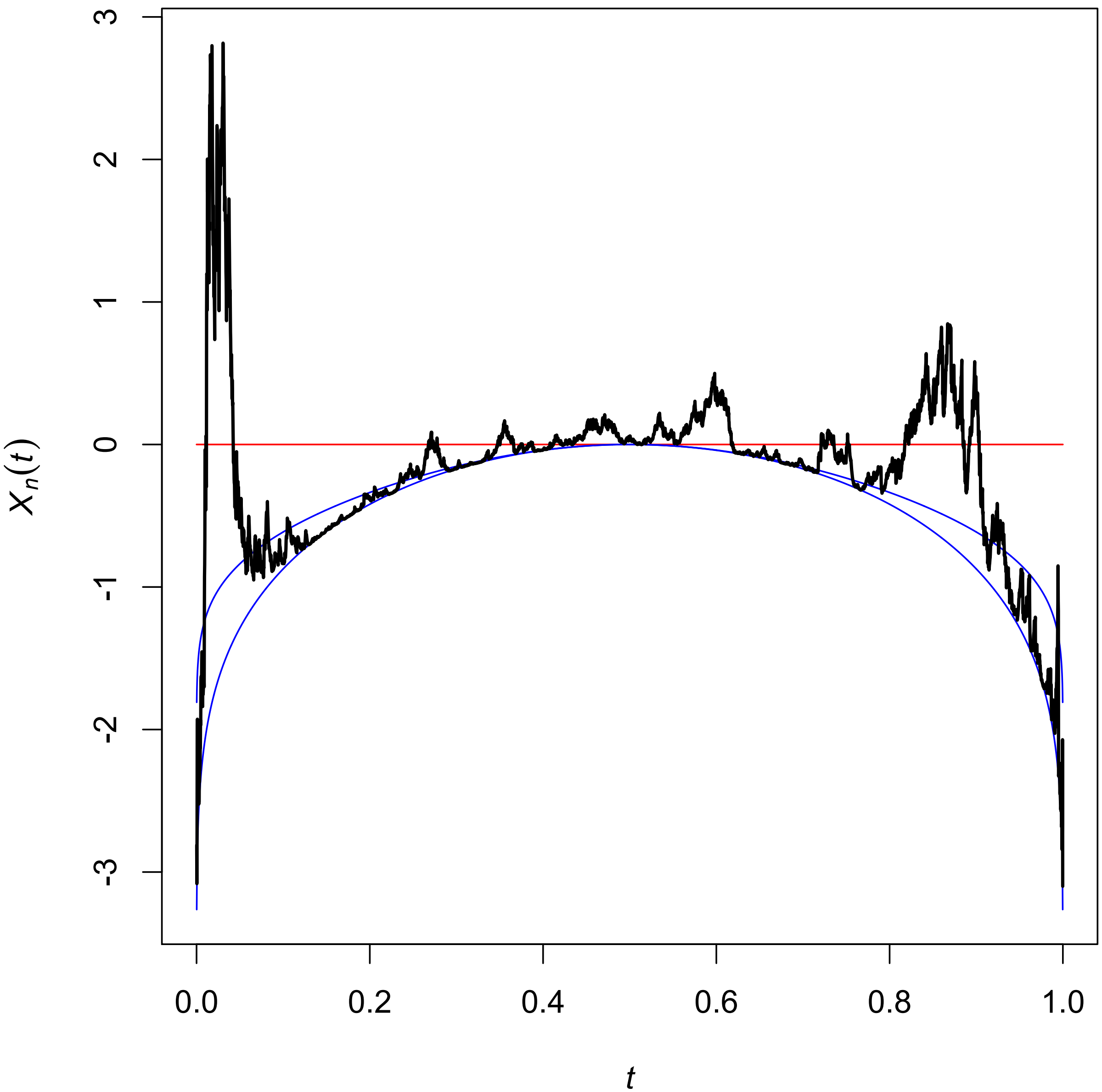
- Some realizations of $\tilde{X}_n - C - \nu D$ for $n = 5000$ and $\nu = 3$:



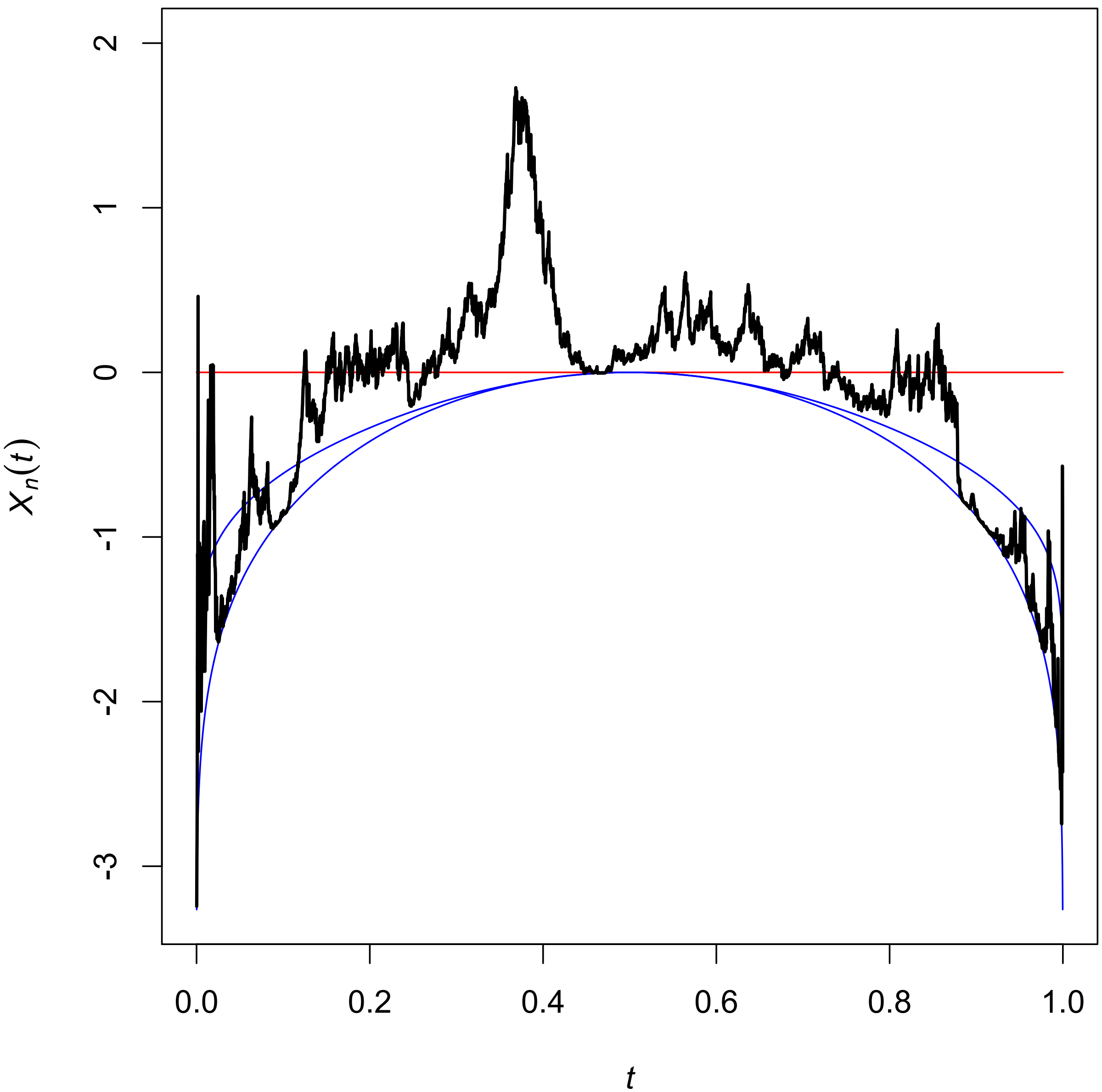
2
1
0
-1
-2
-3
$X_n(t)$
0.0
0.2
0.4
0.6
0.8
1.0
t



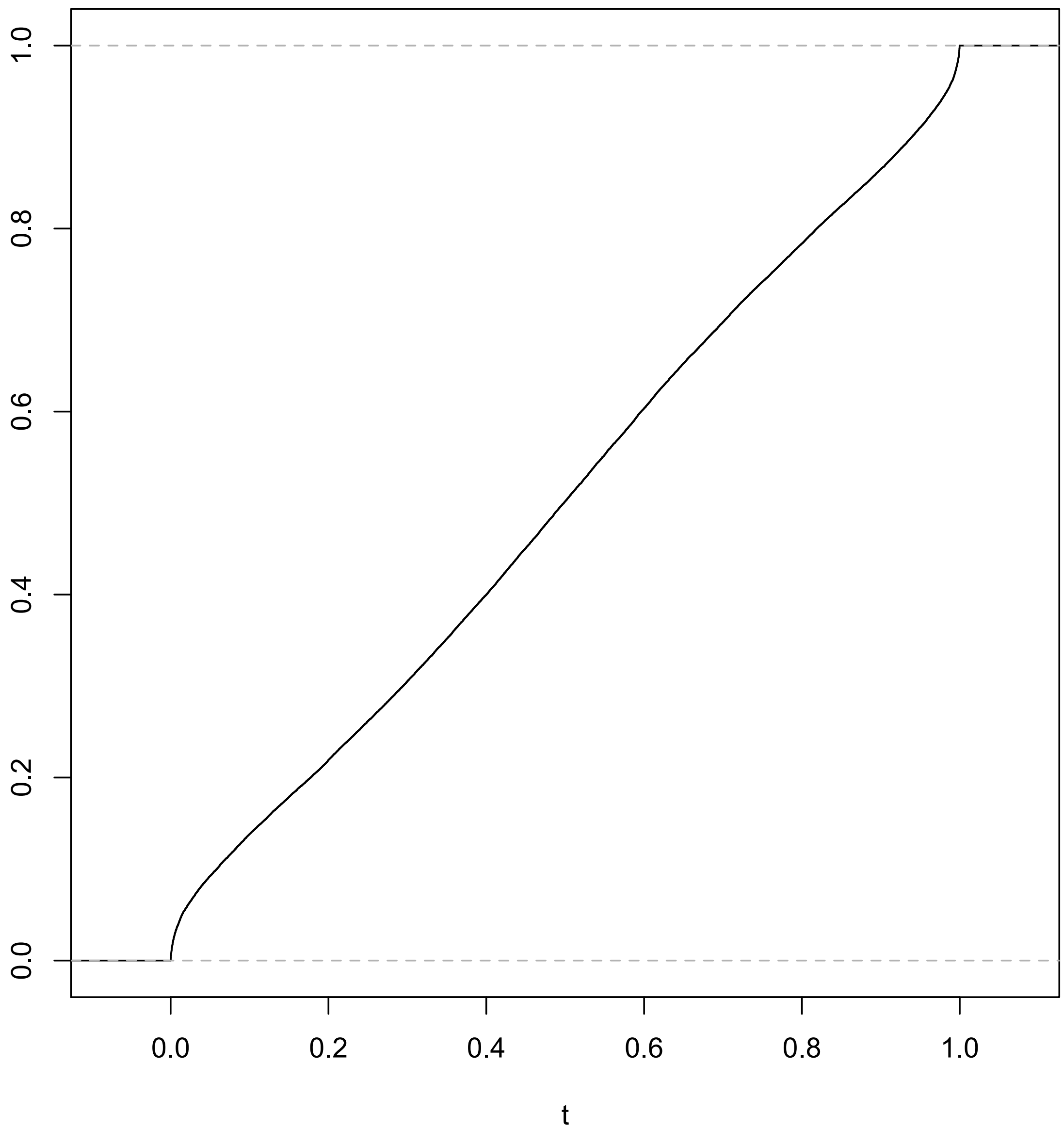
2
1
0
-1
-2
-3
$X_n(t)$
0.0
0.2
0.4
0.6
0.8
1.0
t



3
2
1
0
-1
-2
-3
$X_n(t)$
0.0
0.2
0.4
0.6
0.8
1.0
t


2
1
0
-1
-2
-3
$X_n(t)$
0.0
0.2
0.4
0.6
0.8
1.0
t

Distribution function of $\arg\max_t \tilde{X}_n(t)$:



Let $X_1, X_2, \ldots, X_n$ be i.i.d. with unknown c.d.f. F on $\mathbb{R}$.

- Empirical df: $\mathbb{F}_n(x) \;:=\; \frac{1}{n}\sum_{i=1}^n 1_{[X_i \le x]}$.
- Testing problem: $H_0 : F \equiv F_0$ versus $K : F \not\equiv F_0$.
- Berk-Jones statistic: $R_n(F_0) \;:=\; \sup_{\mathbb{R}} n\, K(\mathbb{F}_n, F_0)$.
- Critical value:

$$\begin{aligned} \kappa_{n,\alpha}^{\mathrm{BJ}} \;&\equiv\; (1-\alpha) - \text{quantile of } \sup_{t\in(0,1)} n\, K_n(\mathbb{G}_n(t), t) \\ &=\; \operatorname{loglog}(n) + O\Big(\operatorname{logloglog}(n)\Big). \end{aligned}$$

- New proposal:

$$T_n(F_0) \;\equiv\; \sup_{\mathbb{R}} \Big(n\, K(\mathbb{F}_n, F_0) - C(F_0) - \nu D(F_0)\Big).$$

... with critical value

$$\begin{aligned}
\kappa_{n,\alpha}^{\text{new}} &\equiv (1-\alpha) - \text{quantile of} \\
&\qquad \sup_{t\in(0,1)} \Big(n\, K(\mathbb{G}_n(t), t) - C(t) - \nu D(t) \Big) \\
&\to (1-\alpha) - \text{quantile of} \\
&\qquad \sup_{t\in(0,1)} \Big(\frac{\mathbb{U}(t)^2}{2t(1-t)} - C(t) - \nu D(t) \Big).
\end{aligned}$$

Lemma. For any critical value $\kappa > 0$ there exists a constant B_κ such that

$$\mathbb{P}_F\Big(T_n(F_o) \le \kappa\Big) \;\le\; B_\kappa\, \Delta_n(F, F_o)^{-4/5}$$

where

$$\Delta_n(F, F_o) \;:=\; \sup_{\mathbb{R}} \frac{\sqrt{n}\,|F - F_o|}{\sqrt{\Gamma(F_o)\,F_o(1 - F_o)} + \Gamma(F_o)/\sqrt{n}}$$

and $\Gamma(\cdot) := C(\cdot) + 1$.

Note:

$$\sqrt{\Gamma(t)\,t(1-t)} \;\to\; 0 \quad \text{as } t \to \{0, 1\}.$$

Special case: Detecting heterogeneous Gaussian mixtures (Donoho–Jin 2004)

$$\begin{aligned} F_o &:= \Phi, \\ F_n &:= (1-\varepsilon_n)\,\Phi + \varepsilon_n\,\Phi(\cdot - \mu_n), \qquad \varepsilon_n \in (0,1),\ \mu_n > 0. \end{aligned}$$

Setting 1 (Donoho–Jin, 2004):

$$\varepsilon_n \;=\; n^{-\beta+o(1)} \quad \text{for some } \beta \in (1/2, 1).$$

Setting 2:

$$\varepsilon_n \;=\; n^{-1/2+o(1)} \quad \text{but} \quad \pi_n := n^{1/2}\varepsilon_n \;\to\; 0.$$

Theorem A. For any fixed $\kappa > 0$,

$$\mathbb{P}_{F_n}\Big(T_n(F_o) > \kappa\Big) \;\to\; 1$$

provided that μ_n satisfies the following conditions:

Setting 1 $(\varepsilon_n = n^{-\beta+o(1)},\ \beta \in (1/2, 1))$:

$$\mu_n = \sqrt{2r\log(n)} \quad \text{with} \quad r > \begin{cases} \beta - 1/2 & \text{if } \beta \le 3/4, \\ \left(1 - \sqrt{1-\beta}\right)^2 & \text{if } \beta \ge 3/4. \end{cases}$$

Setting 2 $(\varepsilon_n = n^{-1/2+o(1)},\ \pi_n = n^{1/2}\varepsilon_n \to 0)$:

$$\mu_n = \sqrt{2s\log(1/\pi_n)} \quad \text{with} \quad s > 1.$$

Setting 2' (contiguous alternatives): For fixed $\pi, \mu > 0$,

$$\varepsilon_n = \pi n^{-1/2} \quad \text{and} \quad \mu_n = \mu.$$

Optimal test of F versus F_n has asymptotic power

$$\Phi\Big(\Phi^{-1}(\alpha) + \frac{\pi^2(\exp(\mu^2) - 1)}{4}\Big).$$

Theorem B. As $\pi \downarrow 0$ and $\mu = \sqrt{2s\log(1/\pi)}$ for fixed $s > 0$,

$$\Phi\Big(\Phi^{-1}(\alpha) + \frac{\pi^2(\exp(\mu^2) - 1)}{4}\Big) \to \begin{cases} \alpha & \text{if } s < 1, \\ 1 & \text{if } s > 1, \end{cases}$$

$$\limsup_{n\to\infty} \mathbb{P}_{F_n}\big(T_n(F_0) > \kappa_{n,\alpha}\big) \to 1 \quad \text{if } s > 1.$$

- Confidence Bands

Owen (1995) proposed $(1-\alpha)$-confidence band

$$\Big\{F : \sup_{\mathbb{R}} n\, K(\mathbb{F}_n, F) \leq \kappa^{\mathsf{BJ}}_{n,\alpha}\Big\}.$$

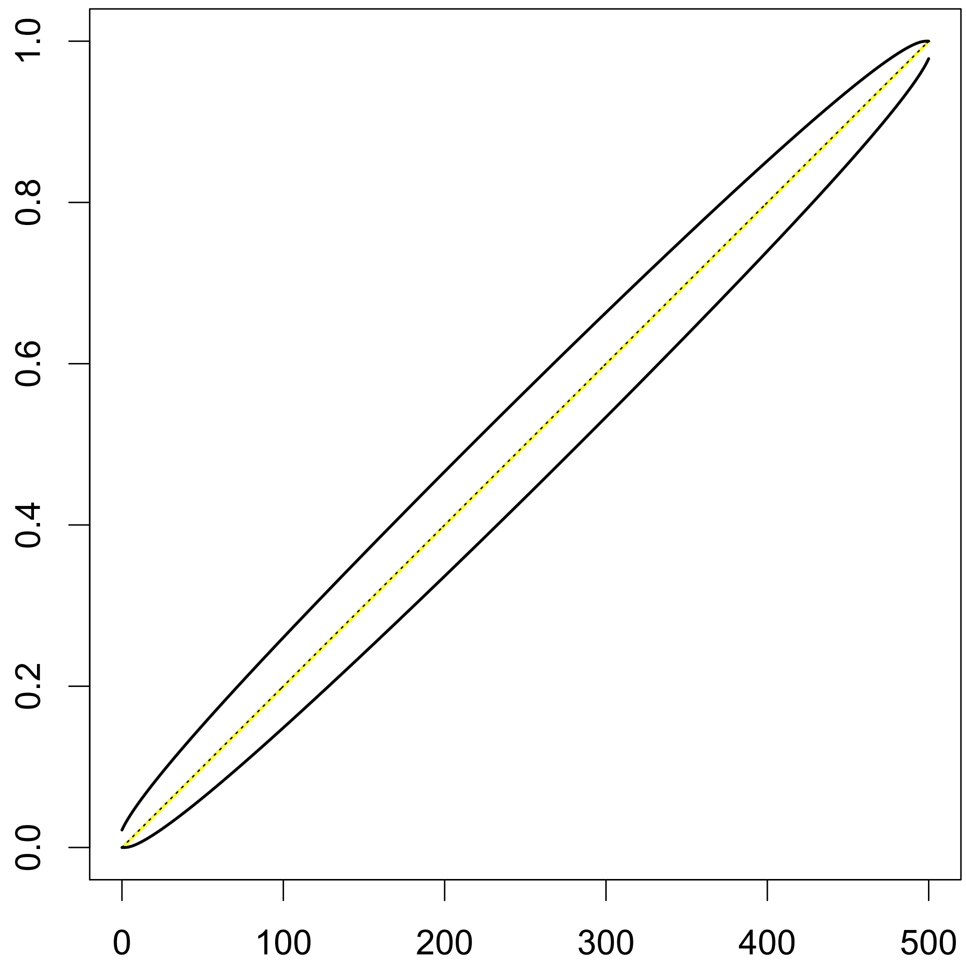
New proposal: With order statistics $X_{n:1} \leq X_{n:2} \leq \cdots \leq X_{n:n}$,

$$\Big\{F : \max_{1\leq i\leq n}\Big((n+1)K(t_{ni}, F(X_{n:i})) - C(t_{ni}) - \nu D(t_{ni})\Big) \leq \tilde{\kappa}_{n,\alpha}\Big\}$$

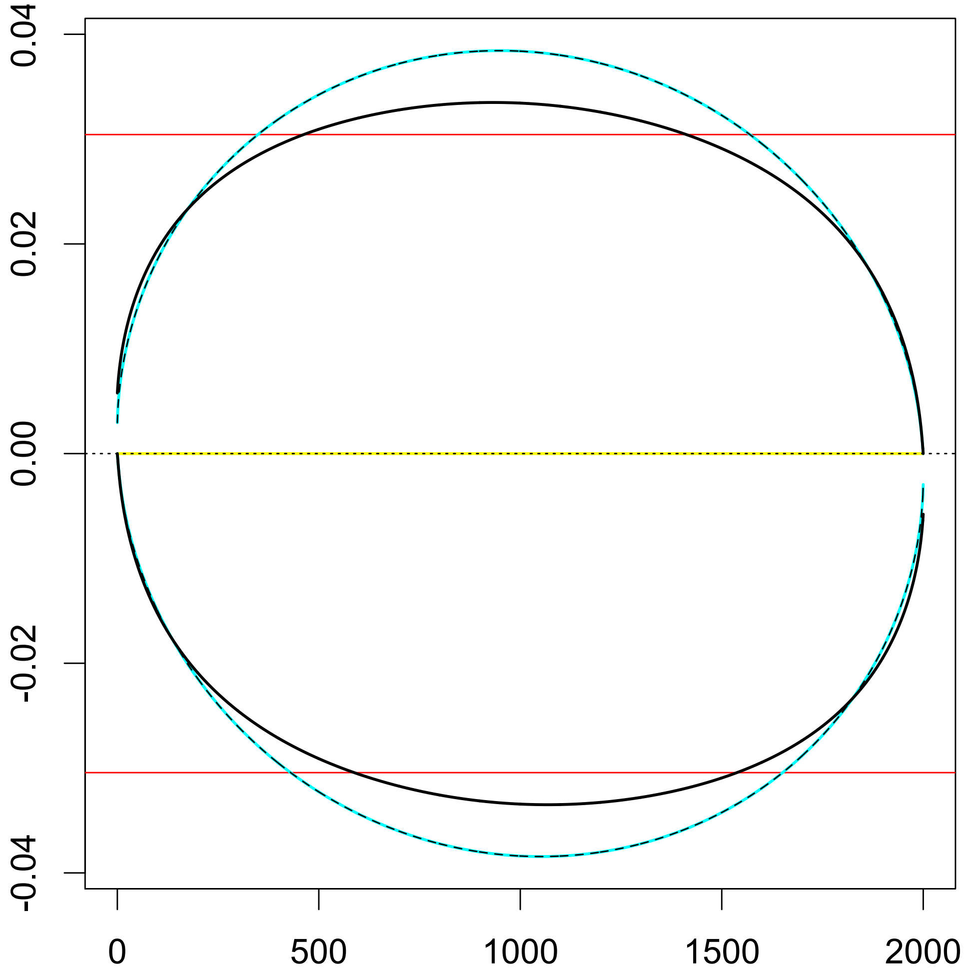
Resulting bounds for $F(x)$: with confidence $1-\alpha$, on $[X_{n:i}, X_{n:i+1})$, $0 \leq i \leq n$,

$$F \quad \in \quad \begin{cases} [a^{\mathsf{BJO}}_{ni}, b^{\mathsf{BJO}}_{ni}] & \text{with Owen's (1995) proposal,} \\ [a^{\mathsf{new}}_{ni}, b^{\mathsf{new}}_{ni}] & \text{with new proposal,} \end{cases} \text{while} \quad \mathbb{F}_n(X_{n:i}) =$$

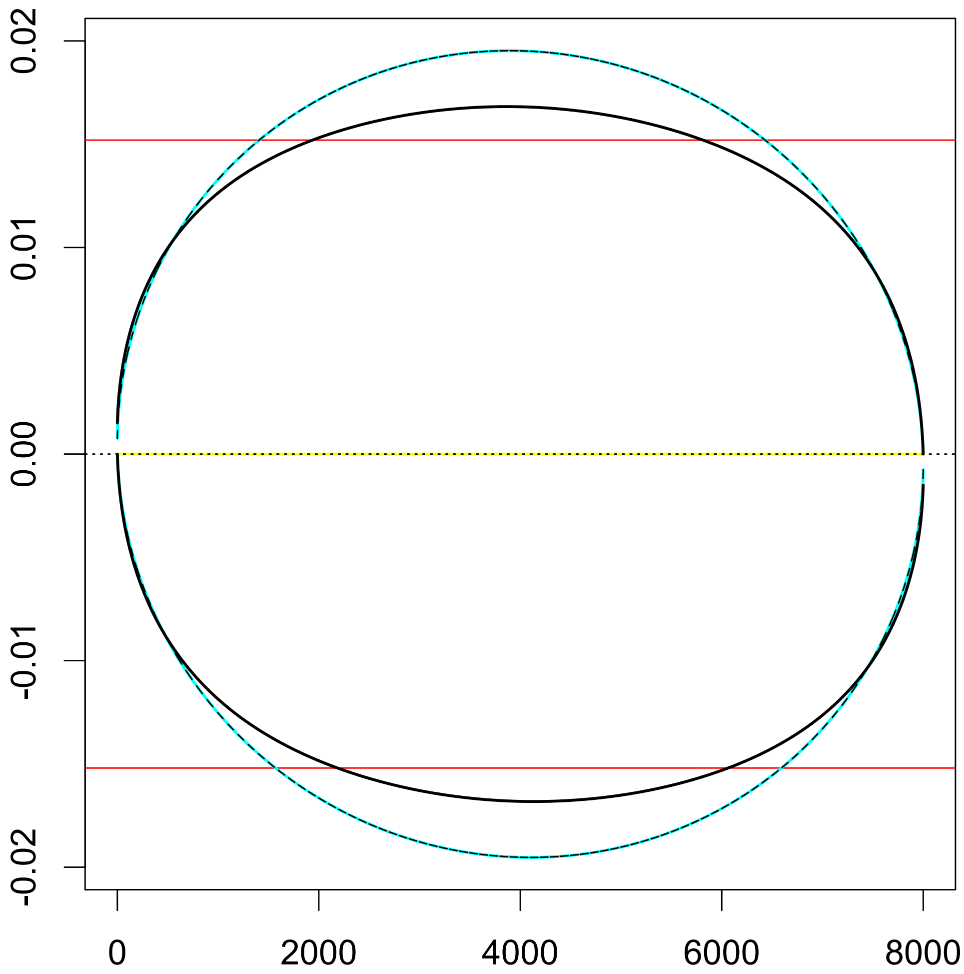
$n = 500$: $\quad i \mapsto a_{ni}^{\text{new}}, s_{ni}, b_{ni}^{\text{new}}$



$n = 2000$: $\quad i \mapsto a^*_{ni} - s_{ni}, b^*_{ni} - s_{ni}$



$n = 8000$: $\quad i \mapsto a^*_{ni} - s_{ni}, b^*_{ni} - s_{ni}$



Theorem C. For any fixed $\alpha \in (0,1)$,

$$\max_{0\le i\le n} \frac{b_{ni}^{\text{new}} - a_{ni}^{\text{new}}}{b_{ni}^{\text{BJO}} - a_{ni}^{\text{BJO}}} \;\to\; 1,$$

while

$$\max_{0\le i\le n} (b_{ni}^{\text{BJO}} - a_{ni}^{\text{BJO}}) \;=\; (1+o(1))\sqrt{\frac{2\text{loglog}n}{n}},$$

$$\max_{0\le i\le n} (b_{ni}^{\text{new}} - a_{ni}^{\text{new}}) \;=\; O(n^{-1/2}).$$

Final comments, extensions

- One can replace $K(u,v) = K_1(u,v)$ with more general 'ϕ-divergences' $K_s(u,v)$ as in Jager–Wellner (2007) under the null hypothesis.

- Power behavior of the family $T_n(s,\nu)$ for $s \notin \{1,2\}$ is still unknown.

- Numerical experiments of Walther (2013) and Siegmund and Li (2015) indicate that $K = K_1$ has the best small/moderate sample performance in the sparse normal means model of Donoho–Jin (2004).

- Results for more general mixture models: Cai and Wu (2014).

- Proof of Jaeschke - Eicker theorem for $\sup_{0<t<1} \frac{\mathbb{U}_n(t)}{\sqrt{t(1-t)}}$:

$$d_n = \frac{(\log n)^5}{n} < 1/2 \qquad \text{if} \quad n > 1010388 \approx 10^6$$

- Number theory: Littlewood showed that $Li(x) - \pi(x)$ changes sign infinitely often for x large.

- Skewes (1933): first sign change of $Li(x) - \pi(x)$ before

$$10^{10^{10^{34}}} \quad \text{if the Riemann hypothesis holds}$$

- Current estimate: first sign change of $Li(x) - \pi(x)$ before $10^{316} \approx e^{726.95133}$.