Inference for the mode: is it possible?

Jon A. Wellner

University of Washington, Seattle

6th IMS-China Meeting, Nanning, June 28 - July 1, 2017

Invited Lecture

6th IMS-China Meeting

Nanning, China

Based on joint work with Charles Doss, University of Minnesota

Outline

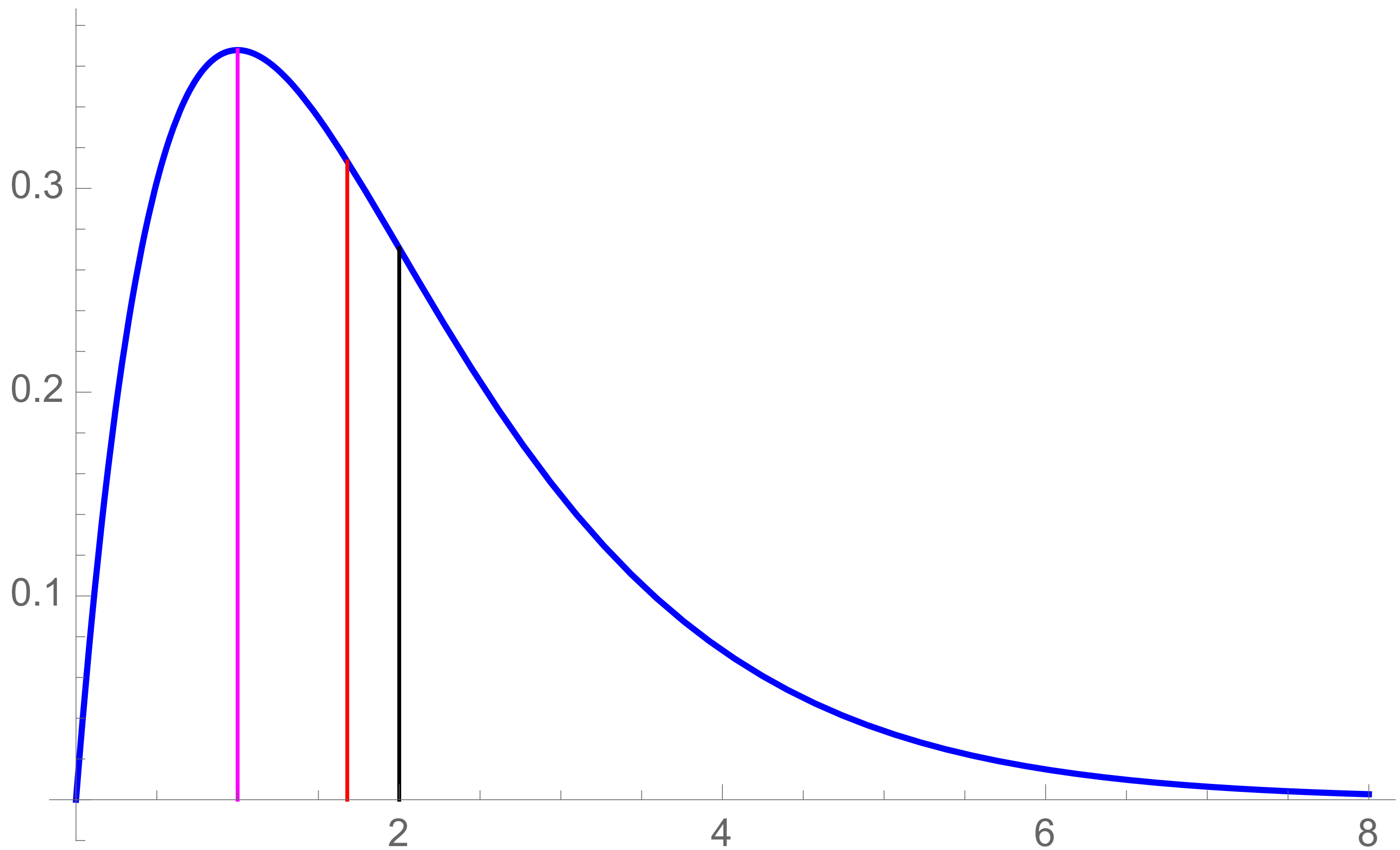
Outline:

- I: Three parameters: mean, median, mode.
- II: Problem: inference for the mode
- III: Estimation of log-concave densities
  - ▷ Log-concave MLE: unconstrained
  - ▷ Log-concave MLE: mode constrained
- IV: Inference for the mode via likelihood ratio tests
  - ▷ Likelihood Ratio test statistics for the mode.
  - ▷ Null hypothesis and curvature $r = 2$ assumption.
  - ▷ Alternative hypothesis (consistency).
  - ▷ Less curvature $r > 2$ holds?

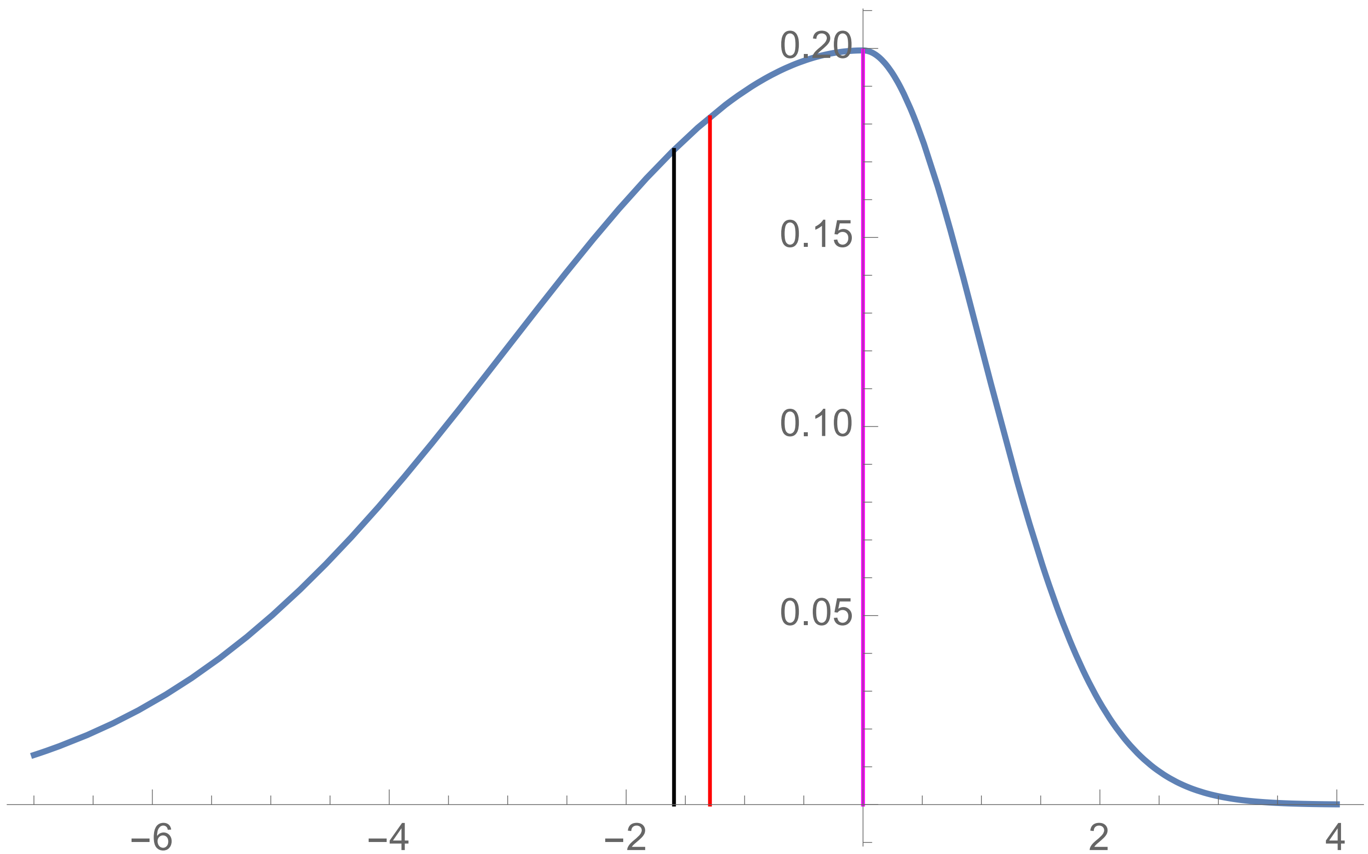
I: Three parameters: mean, median mode

Let X be a real-valued random variable with density f, distribution function F.

- Three parameters:
  - ▷ Mean: $\mu(F) := E_f(X) = \int_{\mathbb{R}} x dF(x) = \int_{\mathbb{R}} x f(x) dx$.
  - ▷ Median: $m(F) := F^{-1}(1/2)$.
  - ▷ Mode: $M(f) := \text{argmax}_{x \in \mathbb{R}} f(x)$.
- Three estimators:
  Let $X_1, \ldots, X_n$ be i.i.d. F, $\mathbb{F}_n(x) = n^{-1} \sum_{i=1}^{n} 1_{[X_i \le x]}$.
  - ▷ sample mean: $\overline{X}_n := \int_{\mathbb{R}} x d\mathbb{F}_n(x)$
  - ▷ sample median: $m(\mathbb{F}_n) := \mathbb{F}_n^{-1}(1/2)$.
  - ▷ sample mode? need to estimate f.
- Confidence intervals for μ, m, or M?



Gamma$(2, 1)$ density: mean (black), median (red), and mode (magenta)



Fechner$(3, 1)$ density: mean (black), median (red), and mode (magenta)

II: Problem: inference for the mode

- f = a unimodal density on $\mathbb{R}$.
- Let $\theta \equiv m \equiv M(f) \equiv$ the mode of f
- Observe: $X_1, \ldots, X_n$ i.i.d. f, empirical d.f. $\mathbb{F}_n$
- **Question 1:** Estimate $\theta = M(f)$ "nonparametrically".
  - ▷ $\widehat{\theta}_{n,h_n} = M(\widehat{f}_{n,h_n})$ where

$$\widehat{f}_{n,h_n}(x) \equiv \int_{\mathbb{R}} \frac{1}{h_n} k\left(\frac{x-y}{h_n}\right) d\mathbb{F}_n(y)$$

    Parzen (1962)
  - ▷ $\widehat{\theta}_n = M(\widehat{f}_n\ )$ where

$$\widehat{f}_n \equiv \text{argmax}_{f \in \mathcal{P}} l_n(f) \equiv \text{Maximum Likelihood Estimator}$$

    for some class of unimodal densities $\mathcal{P}$.
- **Question 2:** Find $1-\alpha$ confidence intervals for $\theta = M(f)$.

Selected previous / related work:

- Mode estimation via kernel density estimators.
  - ▷ Parzen (1962); $V(K, f) \equiv (f(m)/[f''(m)]^2) \cdot \int [k'(x)]^2 dx$
  - ▷ Romano (1988a, 1988b); (bootstrap confidence intervals)
  - ▷ Donoho and Liu (1991); Pfanzagl (1998; 2000)
- Lower bounds: Has'minskii (1979); Donoho and Liu (1991); Balabdaoui, Rufibach, & W (2009)
- Weak curvature or "acuteness" hypotheses: Ehm (1996); Hermann and Ziegler (2004).
- Multiscale methods:
  Schmidt-Hieber, Munk and Dümbgen (2013);
  Dümbgen and Walther (2008).

Our approach here: Let $m_0 \in \mathbb{R}$ be fixed. Consider testing

$$H \ : \ M(f) = m_0 \quad \text{versus} \quad K \ : \ M(f) \neq m_0$$

where f is log-concave under both H and K. This entails finding the maximum likelihood estimator $\widehat{f}_n^0$ of a log-concave density f

subject to the constraint $M(f) = m_0$.

Let

- f_0 the "true" density,
- $\widehat{f}_n$ the (unconstrained) log-concave MLE,
- $\widehat{f}_n^0$ the mode-constrained MLE.

and then $\varphi_0 \equiv \log f_0$; $\widehat{\varphi}_n \equiv \log \widehat{f}_n$; $\widehat{\varphi}_n^0 \equiv \log \widehat{f}_n^0$.

The log-likelihood ratio test statistic for testing H versus K is

$$2\log\lambda_n \;\equiv\; 2n\mathbb{P}_n \log\left(\frac{\widehat{f}_n}{\widehat{f}_n^0}\right) = 2n\mathbb{P}_n\left(\widehat{\varphi}_n - \widehat{\varphi}_n^0\right).$$

Limiting distribution of $2\log\lambda_n$ under H?

If we have $2\log\lambda_n \to_d \mathbb{D}$, then form confidence intervals for $m_0 = M(f_0)$ by inverting the tests.

III: Estimation of log-concave densities

A. Log concave MLE, unconstrained **MLE of f and φ:** Let $\mathcal{C}$ denote the class of all concave functions $\varphi : \mathbb{R} \to [-\infty, \infty)$. The estimator $\widehat{\varphi}_n$ based on $X_1, \ldots, X_n$ i.i.d. as f_0 is the maximizer of the "adjusted criterion function"

$$\ell_n(\varphi) \ = \ \int \log f_\varphi d\mathbb{F}_n - \int f_\varphi(x) dx = \int \varphi d\mathbb{F}_n - \int e^{\varphi(x)} dx$$

over $\varphi \in \mathcal{C}$ where $f_\varphi = e^\varphi$.

Basic properties of $\widehat{f}_n$, $\widehat{\varphi}_n$:

- The (nonparametric) MLE $\widehat{f}_n$ exists for $n \geq 2$ (Walther, Rufibach, Dümbgen and Rufibach).
- $\widehat{f}_n$ can be computed: R-package "logcondens" (Dümbgen and Rufibach)
- $\widehat{\varphi}_n$ is piecewise linear with knots (or kinks) at a subset of the order statistics.
- $\widehat{\varphi}_n = -\infty$ on $\mathbb{R} \setminus [X_{(1)}, X_{(n)}]$; so $\widehat{f}_n = 0$ on $\mathbb{R} \setminus [X_{(1)}, X_{(n)}]$

- In contrast, the (nonparametric) MLE for the class of unimodal densities on $\mathbb{R}$ does not exist. Birgé (1997) and Bickel and Fan (1996) consider alternatives to maximum likelihood for the (large!) class of unimodal densities.

- $\widehat{f}_n$ is characterized by: $\widehat{\varphi}_n$ is the MLE of $\log f_0 = \varphi_0 \in \mathcal{K}_m$ if and only if

$$\widehat{H}_n(x) \begin{cases} \le \mathbb{Y}_n(x), & \text{for all } x \ge X_{(1)}, \\ = \mathbb{Y}_n(x), & \text{if } x \text{ is a knot.} \end{cases}$$

where

$$\widehat{F}_n(x) = \int_{X_{(1)}}^x \widehat{f}_n(y)dy, \qquad \widehat{H}_n(x) = \int_{X_{(1)}}^x \widehat{F}_n(y)dy,$$

$$\mathbb{F}_n(x) = \int_{[X_{(1)},x]} d\mathbb{F}_n(y), \qquad \mathbb{Y}_n(x) = \int_{-\infty}^x \mathbb{F}_n(y)dy.$$

Thus

$$D_n(x) \equiv \widehat{H}_n(x) - \mathbb{Y}_n(x) \le 0 \quad \text{with equality at the knots.}$$

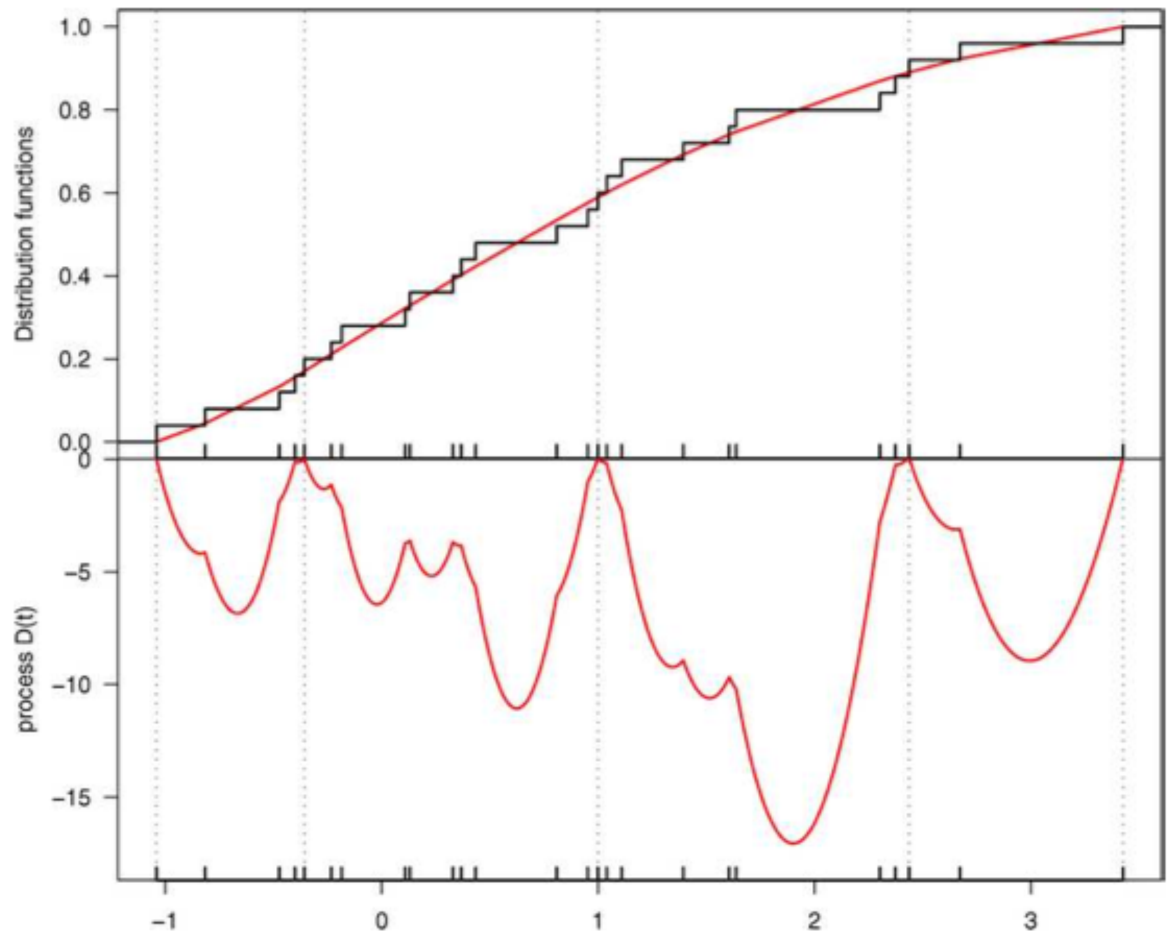
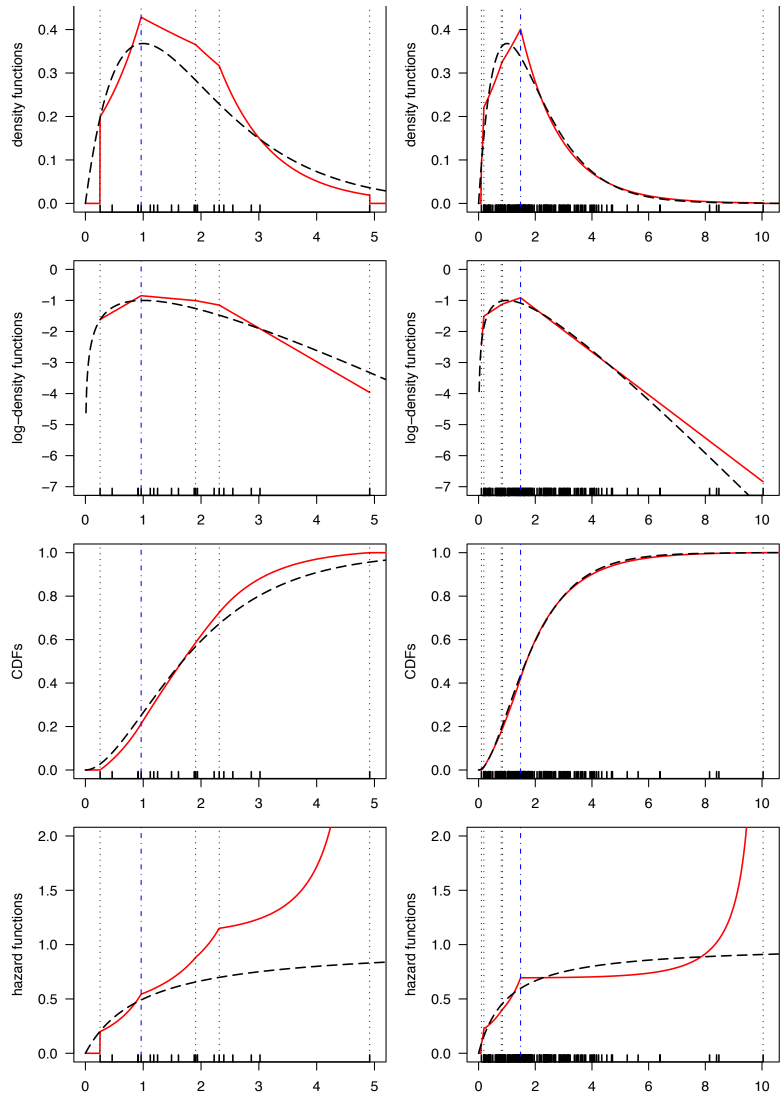


Figure 1. Distribution functions and the process $D(t)$ for a Gumbel sample.


density functions
log−density functions
CDFs
hazard functions

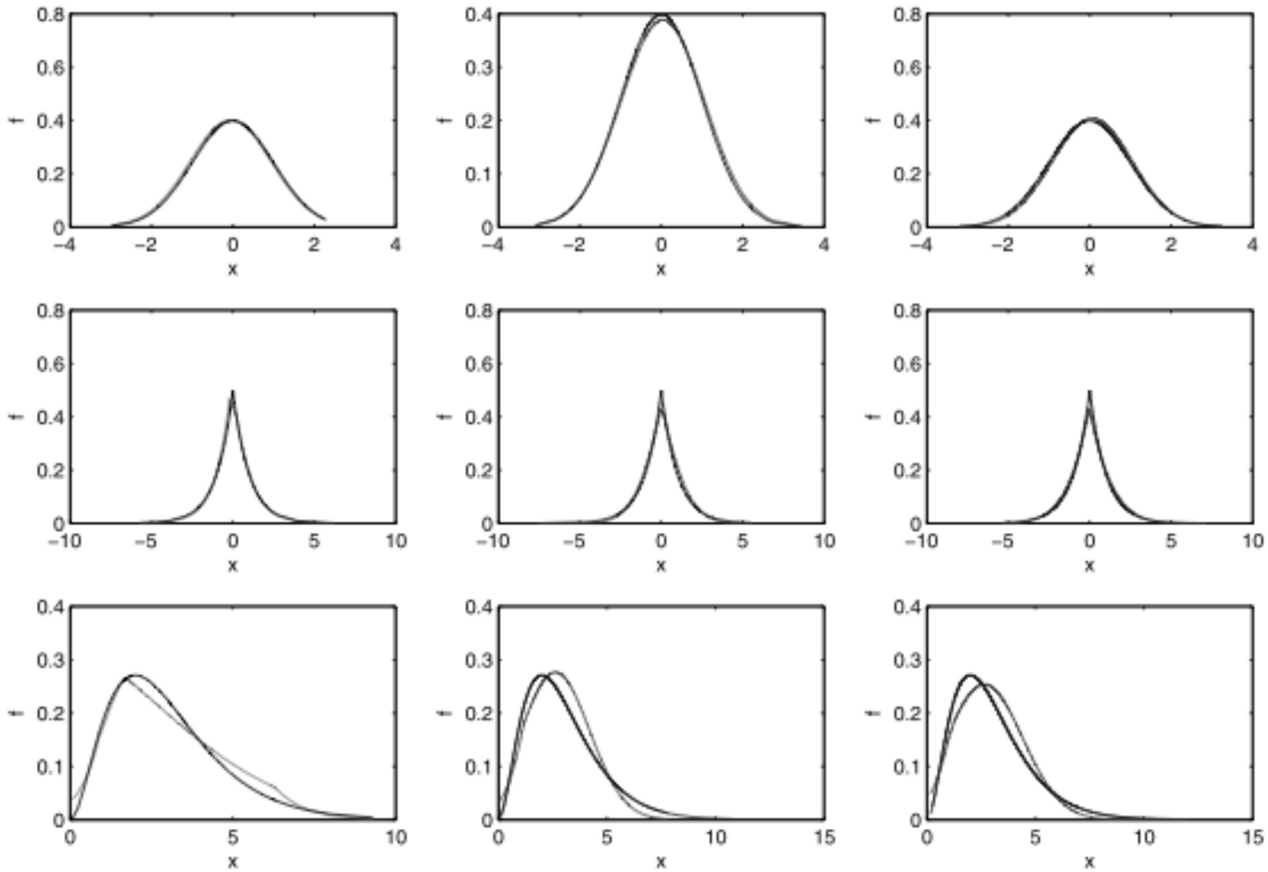


FIG 2. *The estimated log-concave density for different simulation examples. The sample sizes are 50,100 and 200 respectively for first, second and third columns. The three rows correspond to simulations from a Normal(0,1), a double-exponential and a Gamma(3,2) density. The bold one corresponds to the true density and the dotted one is the estimator.*

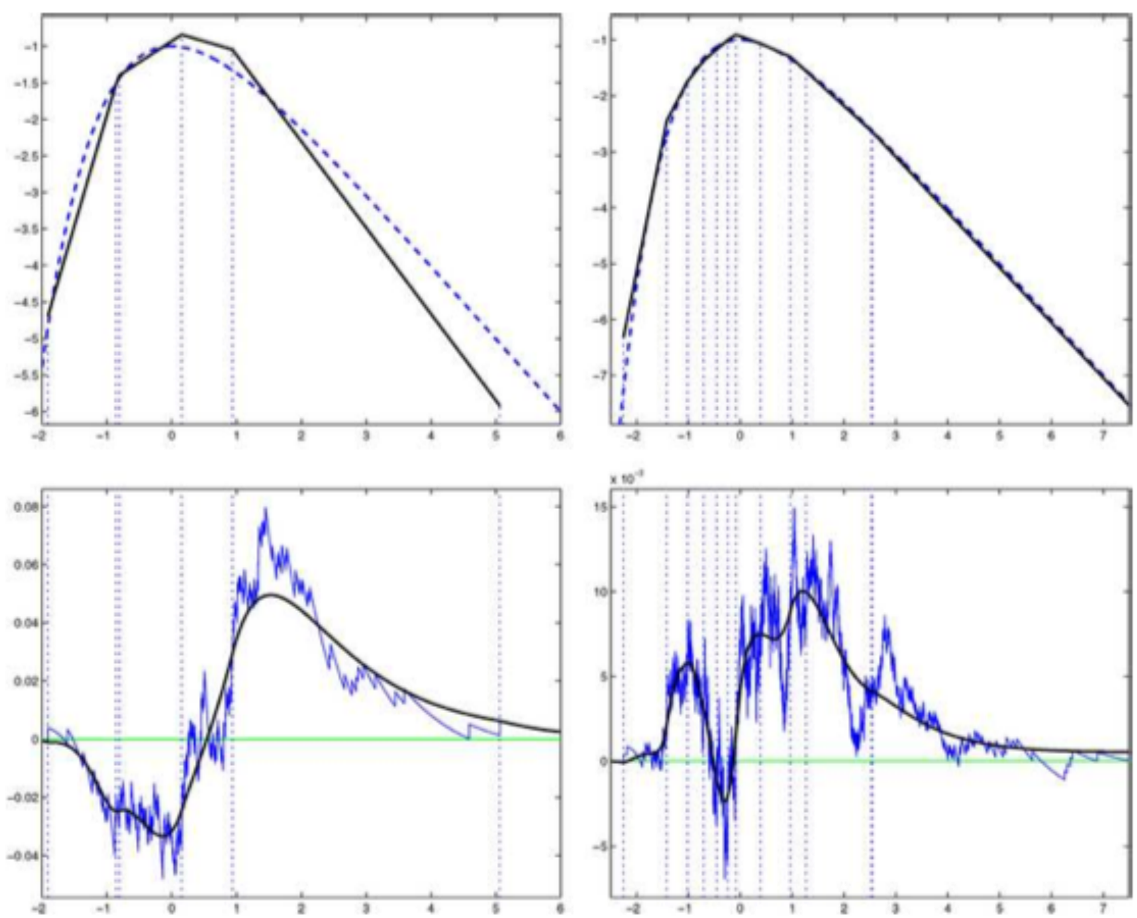


Figure 3. Density functions and empirical processes for Gumbel samples of size $n = 200$ and $n = 2000$.

- The MLE's are Hellinger and L_1- consistent: Pal, Woodroofe, & Meyer (2007).
- $H(\widehat{f}_n, f_0) = O_p(n^{-2/5})$ for $d = 1$: Doss & W (2016).
- The log-concave MLE's $\widehat{f}_{n,0}$ satisfy

$$\int e^{a|x|}|\widehat{f}_{n,0}(x) - f_0(x)|dx \to_{a.s.} 0.$$

  for $a < a_0$ where $f_0(x) \le \exp(-a_0|x| + b_0)$: Cule, Samworth, and Stewart (2010)
- Pointwise distribution theory for $\widehat{f}_n$:

Pointwise distribution theory for $\widehat{f}_n$:

Assumptions: • f_0 is log-concave, $f_0(x_0) > 0$.

- If $\varphi_0''(x_0) \neq 0$, then $k = 2$;
  otherwise, k is the smallest integer such that
  $\varphi_0^{(j)}(x_0) = 0$, $j = 2, \ldots, k-1$, $\varphi_0^{(k)}(x_0) \neq 0$.
- $\varphi_0^{(k)}$ is continuous in a neighborhood of x_0.

Example: $f_0(x) = C\exp(-x^4)$ with $C = \sqrt{2}\Gamma(3/4)/\pi$: $k = 4$.

Driving process: $Y_k(t) = \int_0^t W(s)ds - t^{k+2}$, W standard 2-sided Brownian motion.

Invelope process: H_k determined by limit Fenchel relations:

- $H_k(t) \leq Y_k(t)$ for all $t \in \mathbb{R}$
- $\int_{\mathbb{R}}(H_k(t) - Y_k(t))dH_k^{(3)}(t) = 0.$
- $H_k^{(2)}$ is concave.

Theorem. (Balabdaoui, Rufibach, & W, 2009)

- Pointwise limit theorem for $\widehat{f}_n(x_0)$:

$$\begin{pmatrix} n^{k/(2k+1)}(\widehat{f}_n(x_0) - f_0(x_0)) \\ n^{(k-1)/(2k+1)}(\widehat{f}_n'(x_0) - f_0'(x_0)) \end{pmatrix} \to_d \begin{pmatrix} c_k H_k^{(2)}(0) \\ d_k H_k^{(3)}(0) \end{pmatrix}$$

where

$$c_k \equiv \left(\frac{f_0(x_0)^{k+1} |\varphi_0^{(k)}(x_0)|}{(k+2)!} \right)^{1/(2k+1)},$$

$$d_k \equiv \left(\frac{f_0(x_0)^{k+2} |\varphi_0^{(k)}(x_0)|^3}{[(k+2)!]^3} \right)^{1/(2k+1)}.$$

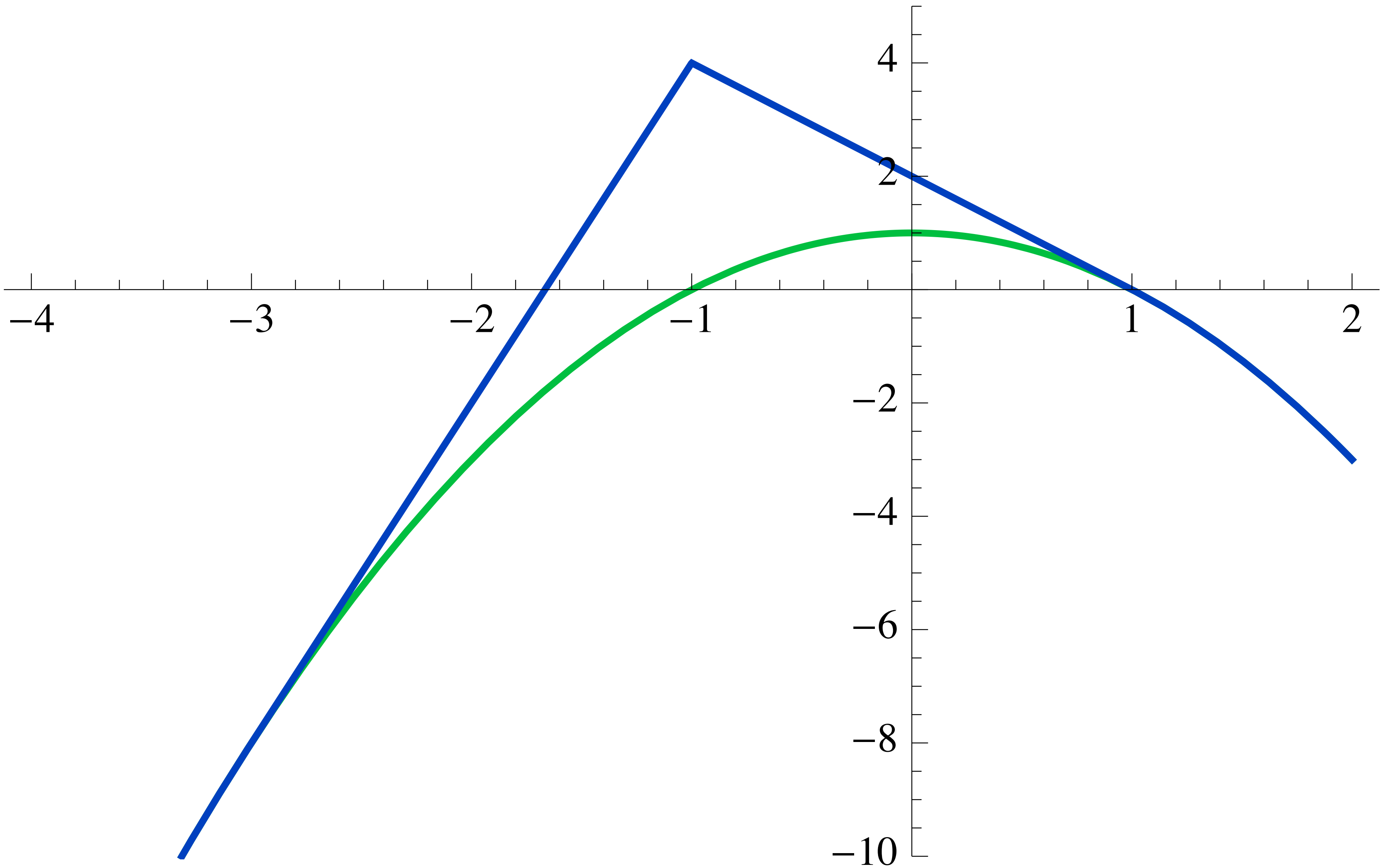
Mode estimation, log-concave density on $\mathbb{R}$

Let $m_0 = M(f_0)$ be the *mode* of the log-concave density f_0, recalling that $\mathcal{P} \subset \mathcal{P}_{unimodal}$. Lower bound calculations using Jongbloed's perturbation φ_ϵ of φ_0 yields:

Proposition. If $f_0 \in \mathcal{P}_0$ satisfies $f_0(m_0) > 0$, $f_0''(m_0) < 0$, and f_0'' is continuous in a neighborhood of x_0, and T_n is any estimator of the mode $m_0 \equiv M(f_0)$, then $f_n \equiv \exp(\varphi_{\epsilon_n})$ with $\epsilon_n \equiv \nu n^{-1/5}$ and $\nu \equiv 2 f_0''(x_0)^2/(5 f_0(x_0))$,

$$\liminf_{n\to\infty} n^{1/5} \inf_{T_n} \max\left\{E_n|T_n - M(f_n)|, E_0|T_n - M(f_0)|\right\}$$
$$\geq \frac{1}{4}\left(\frac{5/2}{10e}\right)^{1/5}\left(\frac{f_0(m_0)}{f_0''(m_0)^2}\right)^{1/5}.$$

Does the MLE $M(\widehat{f}_n)$ achieve this?



4
2
−2
−4
−6
−8
−10
−4
−3
−2
−1
1
2

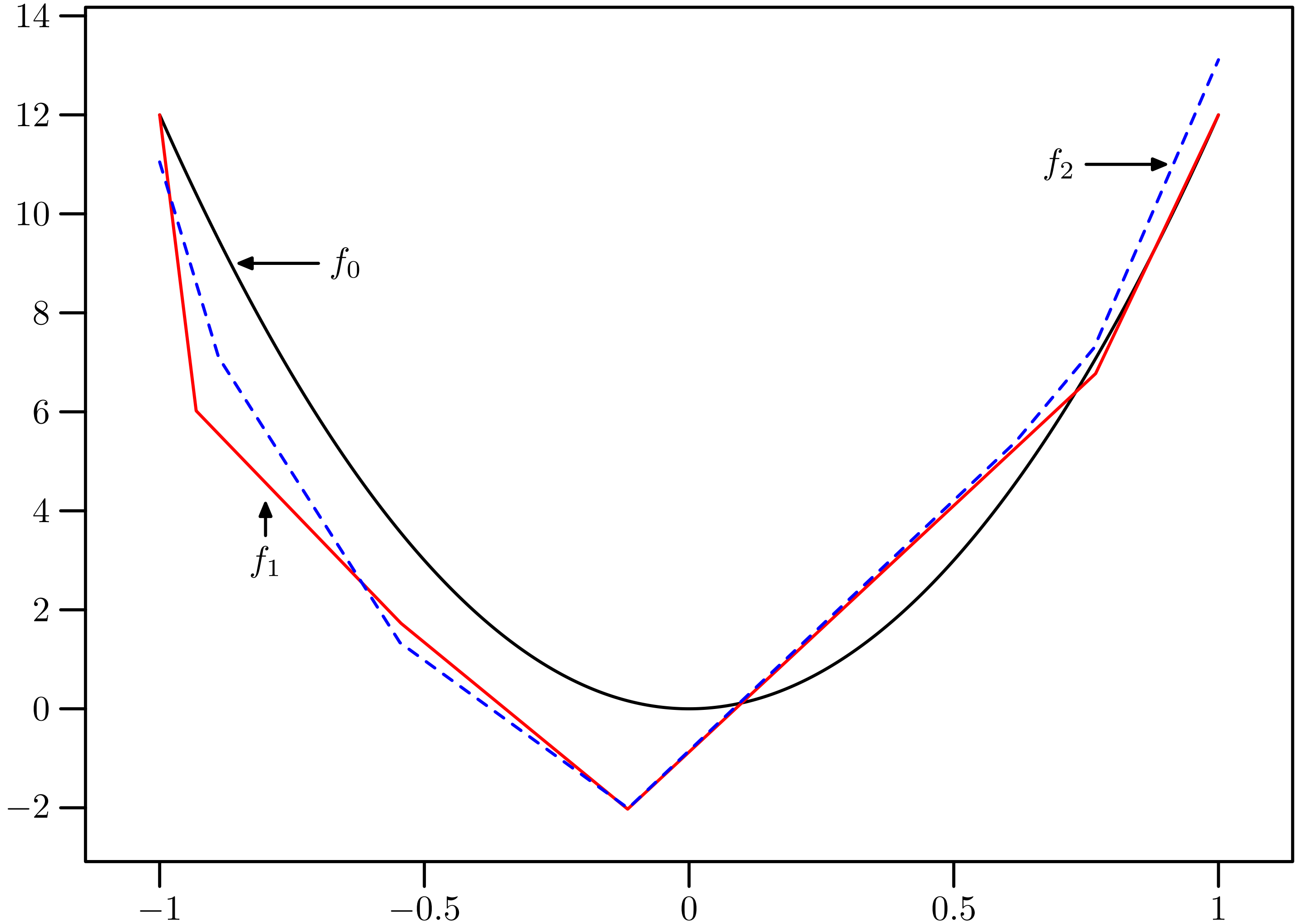
Proposition. (Balabdaoui, Rufibach, & W, 2009)
Suppose that $f_0 \in \mathcal{P}_0$ satisfies:

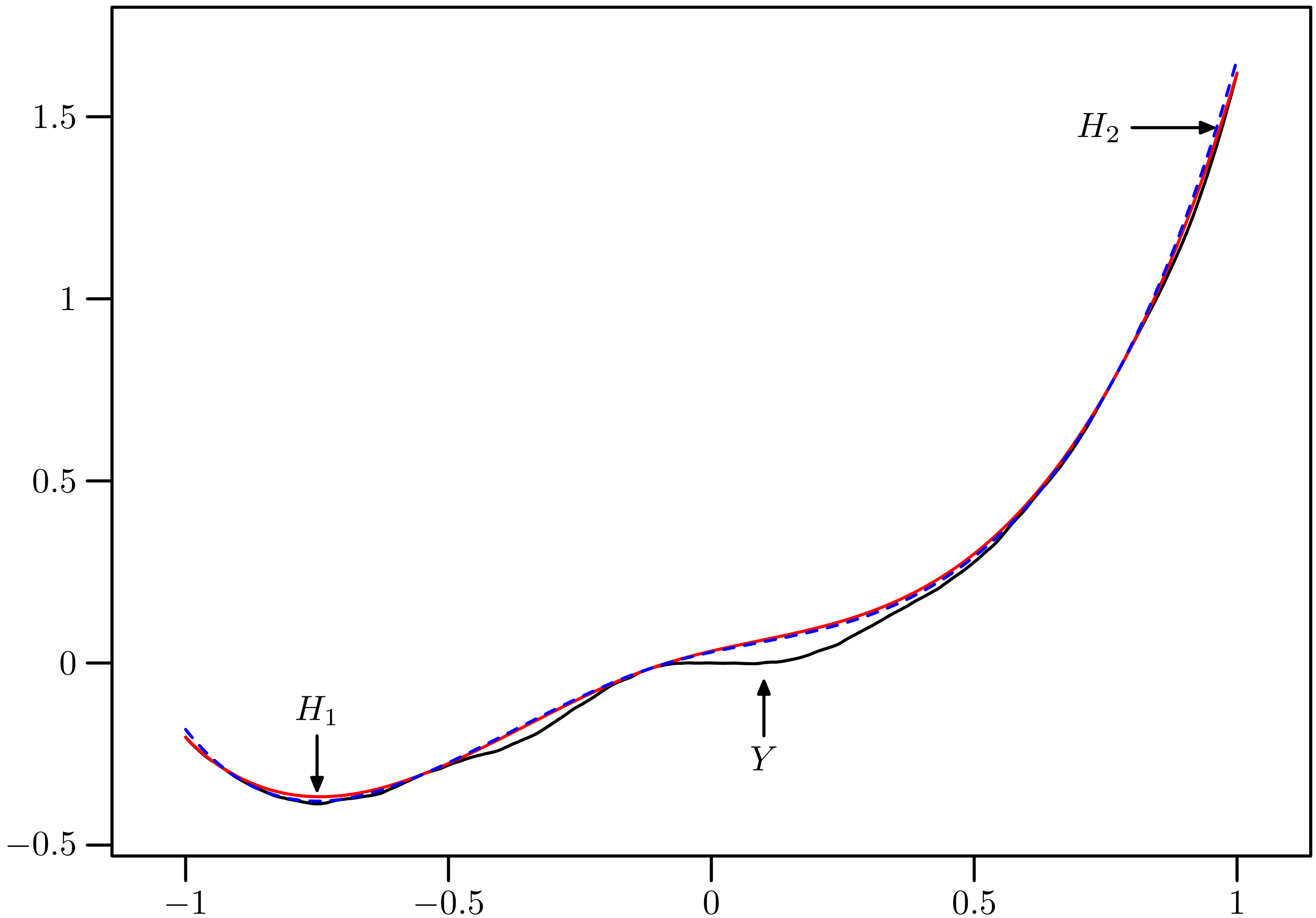
- $\varphi_0^{(j)}(m_0) = 0$, $j = 2, \ldots, k-1$; $\varphi_0^{(k)}(m_0) \neq 0$.
- $\varphi_0^{(k)}$ is continuous in a neighborhood of x_0.

Then $\widehat{M}_n \equiv M(\widehat{f}_n) \equiv \min\{u : \widehat{f}_n(u) = \sup_t \widehat{f}_n(t)\}$, satisfies

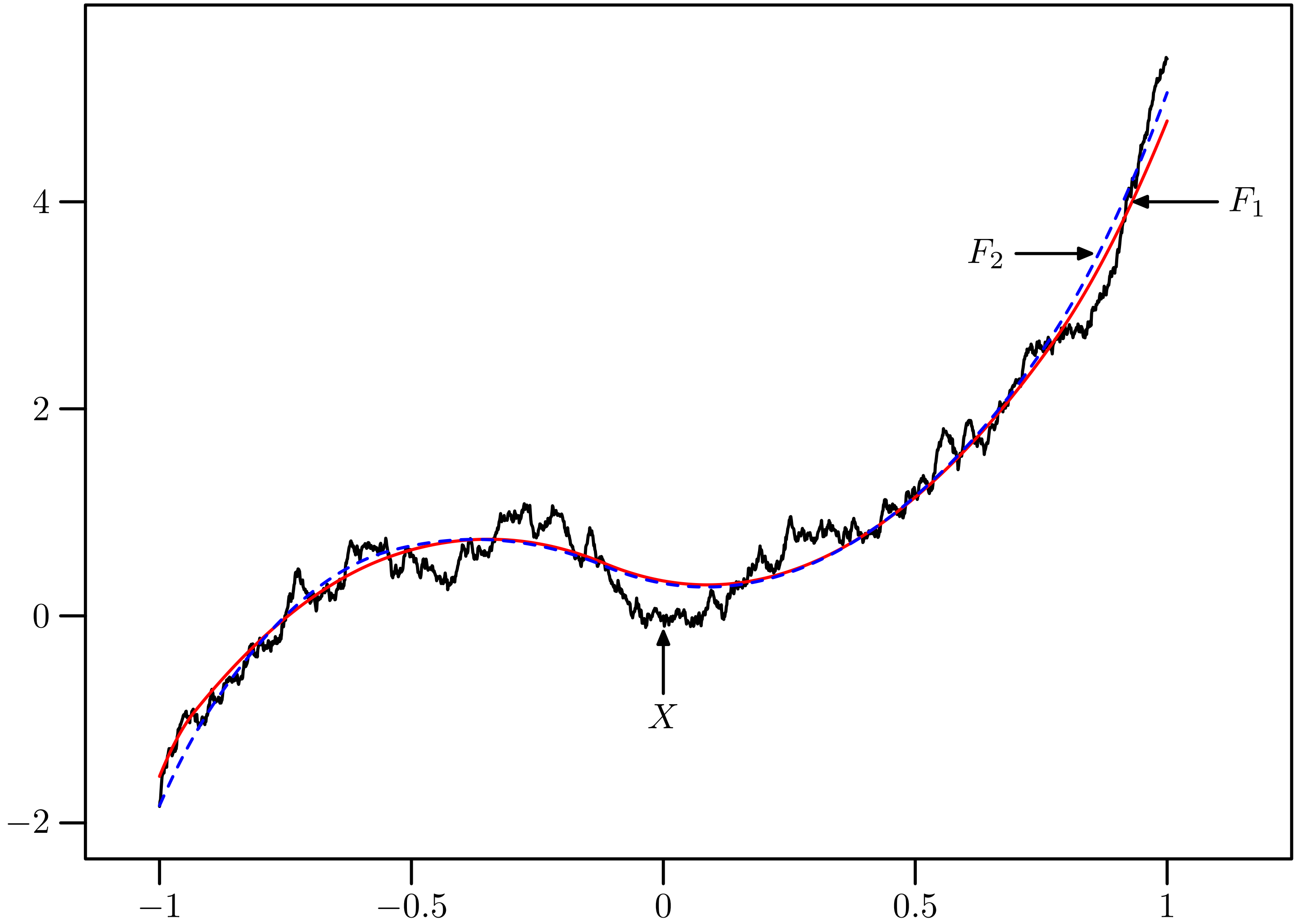
$$n^{1/(2k+1)}(\widehat{M}_n - M(f_0)) \to_d \left(\frac{((k+2)!)^2 f_0(m_0)}{f_0^{(k)}(m_0)^2} \right)^{1/(2k+1)} M(H_k^{(2)})$$

where $M(H_k^{(2)}) = \text{argmax}(H_k^{(2)})$ and $H_k^{(2)}$ is the LSE of $t \mapsto -(k+2)(k+1)|t|^k$ in the canonical Gaussian or white-noise model on $\mathbb{R}$. From Han and W (2016), the rate and dependence of the constant on $f_0(m_0)$ and $f_0^{(k)}(m_0)$ are optimal.

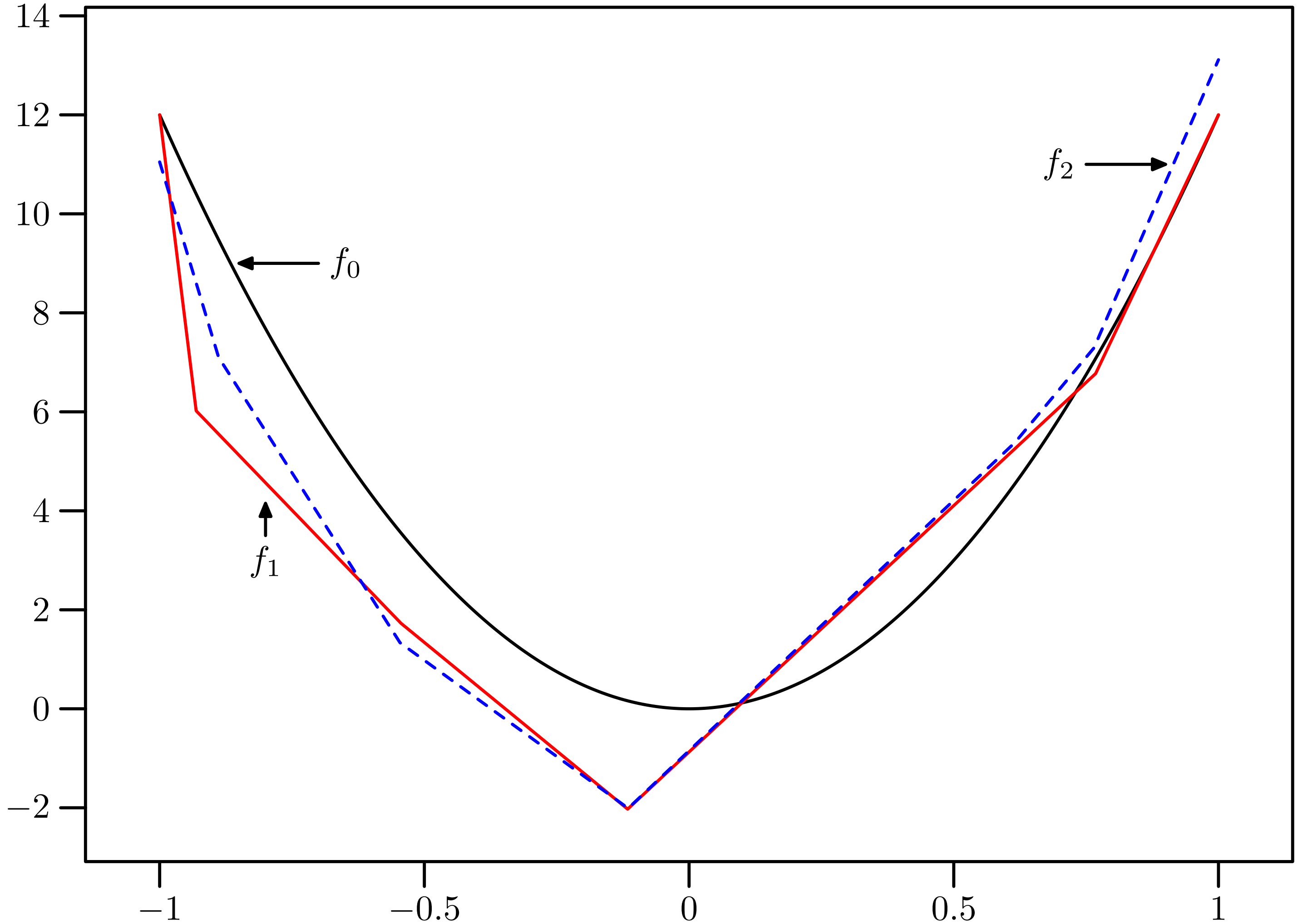

14
12
10
8
6
4
2
0
−2
−1
−0.5
0
0.5
1
f_0
f_1
f_2



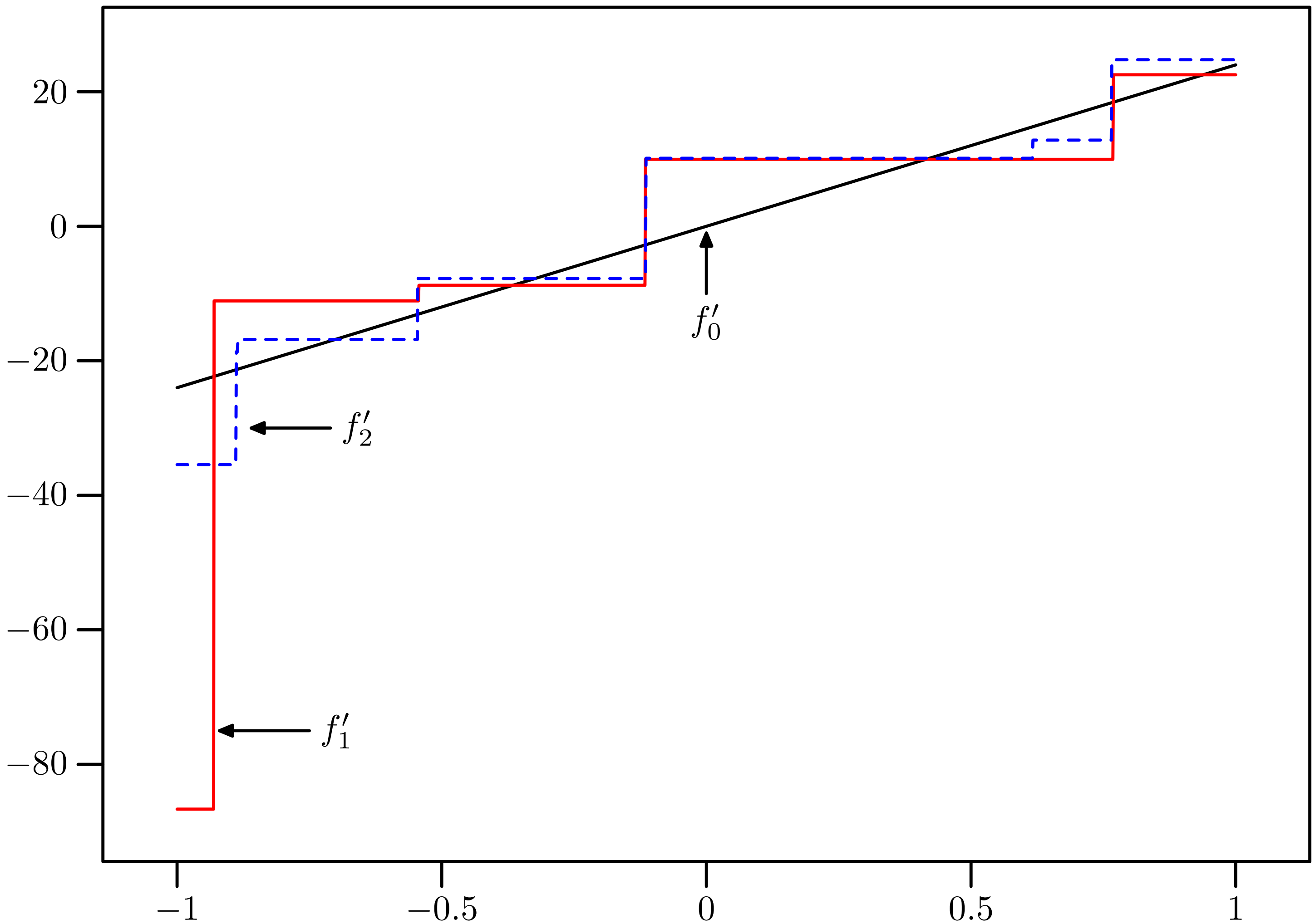
H_2
H_1
Y
1.5
1
0.5
0
−0.5
−1
−0.5
0
0.5
1



4
2
0
−2
−1
−0.5
0
0.5
1
F_1
F_2
X



f_0
f_1
f_2
14
12
10
8
6
4
2
0
−2
−1
−0.5
0
0.5
1



20
0
−20
−40
−60
−80
−1
−0.5
0
0.5
1
f_0'
f_1'
f_2'

B. Log concave MLE, mode constrained

Let $\mathcal{P}_m \equiv \{f \in \mathcal{P} : \quad M(f) = m\}$.
The mode constrained Maximum Likelihood Estimator is

$$\widehat{f}_n^0 \equiv \text{argmax}_{f \in \mathcal{P}_m} \int_{\mathbb{R}} \log f d\mathbb{F}_n.$$

Basic properties of $\widehat{f}_n^0$, $\widehat{\varphi}_n^0$:

- The mode constrained (nonparametric) MLE $\widehat{f}_n^0$ exists and is unique for each $m \in [X_{(1)}, X_{(n)}]$: Doss (2013).
- $\widehat{f}_n^0$ can be computed: Doss (2013); active set algorithm
- $\widehat{\varphi}_n^0 = \log \widehat{f}_n^0$ is piecewise linear with knots (or kinks) at a subset of the order statistics together with m.
- $\widehat{\varphi}_n^0 = -\infty$ on $\mathbb{R} \setminus [X_{(1)}, X_{(n)}]$; so $\widehat{f}_n^0 = 0$ on $\mathbb{R} \setminus [X_{(1)}, X_{(n)}]$ if $m \in [X_{(1)}, X_{(n)}]$.

- $\widehat{f}_n$ is characterized by: $\widehat{\varphi}_n^0$ is the MLE of $\log f_0 = \varphi_0 \in \mathcal{K}_m$ if and only if

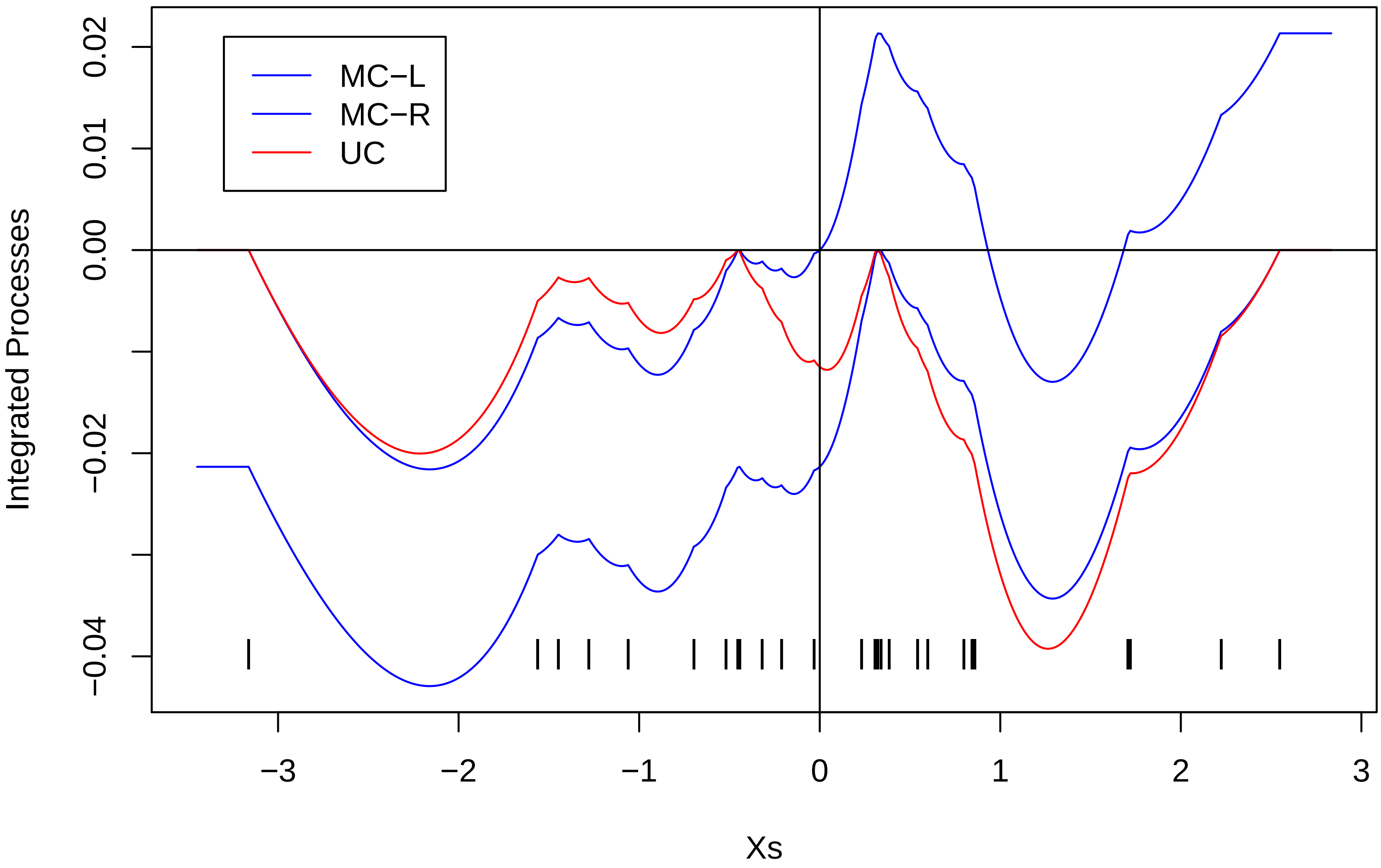
$$\widehat{H}^0_{n,L}(x) \le \mathbb{Y}_{n,L}(x) \quad \text{for all } X_{(1)} \le x \le m,$$

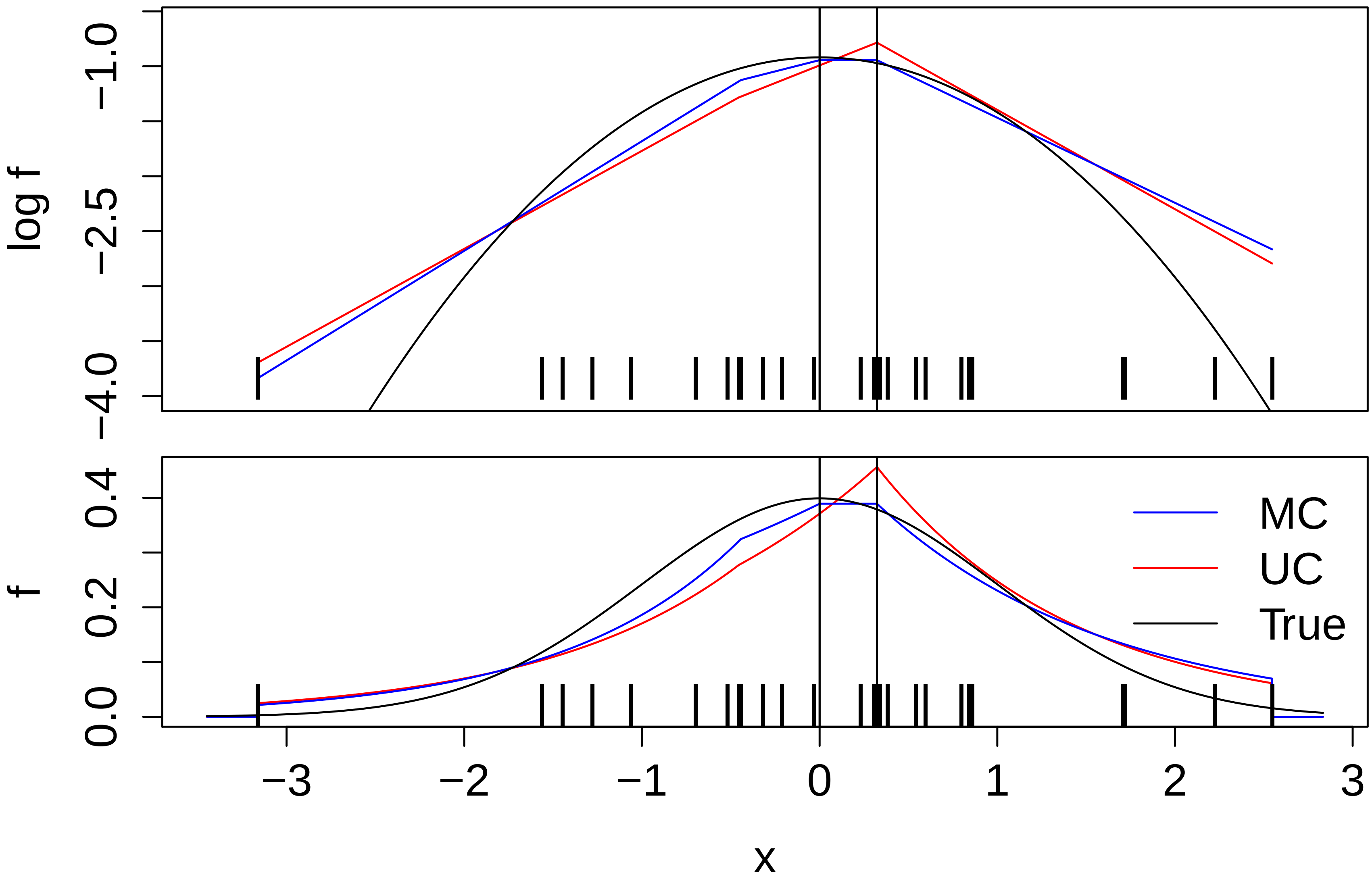
and

$$\widehat{H}^0_{n,R}(x) \le \mathbb{Y}_{n,R}(x) \quad \text{for all } m \le x \le X_{(n)},$$

with equality at the knot points where

$$\begin{aligned}
\widehat{F}^0_{n,L}(x) &= \int_{X_{(1)}}^x \widehat{f}^0_n(y)dy, & \widehat{F}^0_{n,R}(x) &= \int_x^{X_{(n)}} \widehat{f}^0_n(y)dy,\\
\widehat{H}^0_{n,L}(x) &= \int_{X_{(1)}}^x \widehat{F}^0_n(y)dy, & \widehat{H}^0_{n,R}(x) &= \int_x^{X_{(n)}} \widehat{F}^0_{n,R}(y)dy,\\
\mathbb{F}_{n,L}(x) &= \int_{[X_{(1)},x]} d\mathbb{F}_n(y), & \mathbb{F}_{n,R}(x) &= \int_{[x,X_{(n)}]} d\mathbb{F}_n(y),\\
\mathbb{Y}_{n,L}(x) &= \int_{-\infty}^x \mathbb{F}_{n,L}(y)dy, & \mathbb{Y}_{n,R}(x) &= \int_{[x,X_{(n)}]} \mathbb{F}_{n,R}(y)dy.
\end{aligned}$$


MC−L
MC−R
UC
Integrated Processes
0.02
0.01
0.00
−0.02
−0.04
−3
−2
−1
0
1
2
3
Xs


log f
−1.0
−2.5
−4.0
f
0.4
0.2
0.0
MC
UC
True
−3
−2
−1
0
1
2
3
x

IV. Inference for the mode of a log-concave density

From Balabdaoui, Rufibach, and W (2009): suppose that

- f_0 is log-concave;
- $f_0'' \equiv f_0^{(2)}$ is continuous in a neighborhood of $m_0 \equiv M(f_0)$;
- $f_0^{(2)}(m_0) < 0$. Then $\widehat{M}_n \equiv M(\widehat{f}_n)$ satisfies

$$n^{1/5}(\widehat{M}_n - m_0) \to_d \left(\frac{(4!)^2 f_0(m_0)}{f_0^{(2)}(m_0)^2} \right)^{1/5} M(H_2^{(2)})$$

Key difficulties:

(1) to get (Wald-type) confidence intervals for $M(f_0) = m_0$,

need to estimate $f_0^{(2)}(m_0)$ or $f_0(m_0)/f_0^{(2)}(m_0)^2$?!

(2) The intervals rely on the assumption $f_0^{(2)}(m_0) < 0$.

For example, if $f_0(x) = c_4 \exp(-|x|^4/4)$, then $k = 4$, the rate of convergence is $n^{1/9}$ and the limit distribution becomes

$$\left(\frac{(6!)^2}{f_0(m_0)|\varphi_0^{(4)}(m_0)|^2} \right)^{1/9} M(H_4^{(2)}).$$

Can we avoid estimation of $f_0^{(2)}(m_0)$ or $\varphi_0^{(k)}(m_0)$ via

likelihood ratio methods?

Let $m_0 \in \mathbb{R}$ be fixed. Consider testing

$$H \; : \; M(f) = m_0 \quad \text{versus} \quad K \; : \; M(f) \neq m_0$$

where f is log-concave under both H and K. This entails finding the maximum likelihood estimator $\widehat{f}_n^0$ of a log-concave density f

subject to the constraint $M(f) = m_0$.

Let

- f_0 the "true" density,
- $\widehat{f}_n$ the (unconstrained) log-concave MLE,
- $\widehat{f}_n^0$ the mode-constrained MLE.

and then $\varphi_0 \equiv \log f_0$; $\widehat{\varphi}_n \equiv \log \widehat{f}_n$; $\widehat{\varphi}_n^0 \equiv \log \widehat{f}_n^0$.

The log-likelihood ratio test statistic for testing H versus K is

$$
\begin{aligned}
2\log\lambda_n &\equiv 2n\mathbb{P}_n\log\left(\frac{\widehat{f}_n}{\widehat{f}_n^0}\right) \\
&= 2n\mathbb{P}_n\left(\widehat{\varphi}_n - \widehat{\varphi}_n^0\right) \\
&= 2n\mathbb{P}_n\left(\widehat{\varphi}_n - \widehat{\varphi}_n^0\right) - \int_{\mathbb{R}}\left(\widehat{f}_n(u) - \widehat{f}_n^0(u)\right)du \\
&= 2n\int_{[X_{(1)},X_{(n)}]}\left(\widehat{\varphi}_n d\widehat{F}_n - \widehat{\varphi}_n^0 d\widehat{F}_n^0\right) \\
&\qquad - \int_{[X_{(1)},X_{(n)}]}\left(e^{\widehat{\varphi}_n(u)} - e^{\widehat{\varphi}_n^0}\right)du
\end{aligned}
$$

Limiting distribution of $2\log\lambda_n$ under H?

If we have $2\log\lambda_n \to_d \mathbb{D}$, then form confidence intervals for $m_0 = M(f_0)$ by inverting the tests.

Theorem. A. (At $x_0 \neq m_0$) If $\varphi_0^{(2)}(x_0) < 0$ and $f_0(x_0) > 0$, then

$$\begin{pmatrix} n^{2/5}(\widehat{\varphi}_n(x_0) - \varphi_0(x_0)) \\ n^{2/5}(\widehat{\varphi}_n^0(x_0) - \varphi_0(x_0)) \end{pmatrix} \to_d \begin{pmatrix} \mathbb{V} \\ \mathbb{V} \end{pmatrix}$$

where $\mathbb{V} = C(x_0, \varphi_0) H_2^{(2)}(0)$, $C(x_0, \varphi_0) = \left(|\varphi_0^{(2)}(x_0)/(4! f_0(x_0)) \right)^{1/5}$.

Consequently

$$n^{2/5}\left(\widehat{\varphi}_n(x_0) - \widehat{\varphi}_n^0(x_0)\right) \to_p 0.$$

B. (In $n^{-1/5}-$neighborhoods of m_0) If $\varphi_0^{(2)}(m_0) < 0$ and $f_0(m_0) > 0$, then the processes

$$\mathbb{X}_n(t) \equiv n^{2/5}(\widehat{\varphi}_n(m_0 + n^{-1/5}t) - \varphi_0(m_0)),$$
$$\mathbb{X}_n^0(t) \equiv n^{2/5}(\widehat{\varphi}_n^0(m_0 + n^{-1/5}t) - \varphi_0(m_0)),$$

converge finite-dimensionally in distribution (jointly) to the finite dimensional distributions of the processes $(\mathbb{X}(t), \mathbb{X}^0(t)) \equiv (\widehat{\varphi}(t/\gamma_2), \widehat{\varphi}^0(t/\gamma_2))/(\gamma_1 \gamma_2^2)$ where $\gamma_1 = (a^{3/5}/\sigma^{8/5})$, $\gamma_2 = (\sigma/a)^{2/5}$, and $\sigma \equiv 1/\sqrt{f_0(m_0)}$, $a = |\varphi_0^{(2)}(m_0)|/4!$

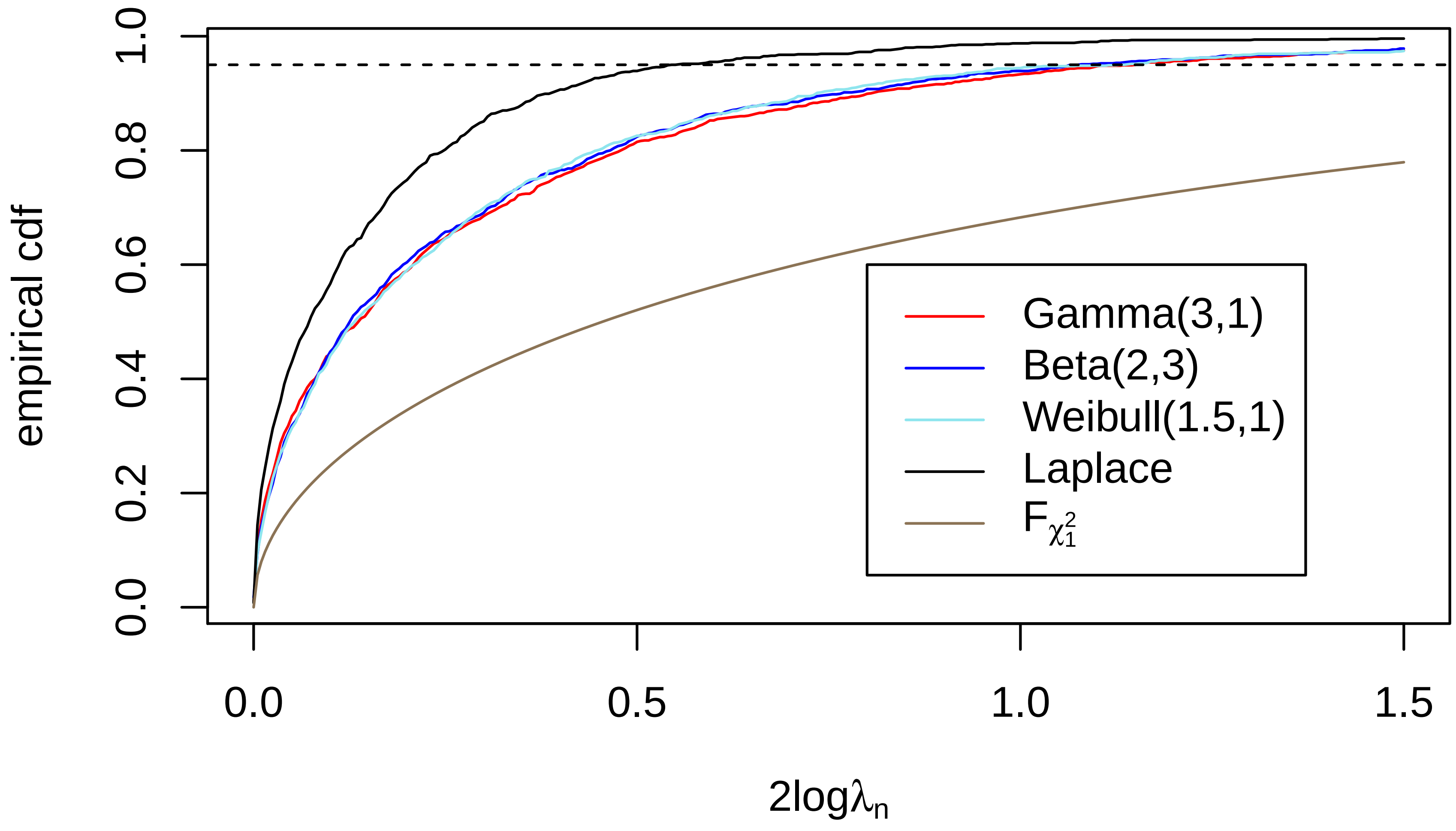
$$2\log\lambda_n = n\int_{D_n} f_0(m_0)\left\{(\widehat{\varphi}_n(u)-\varphi_0(m_0))^2 - \left(\widehat{\varphi}_n^0(u)-\varphi_0(m_0)\right)^2\right\}du + R_n.$$

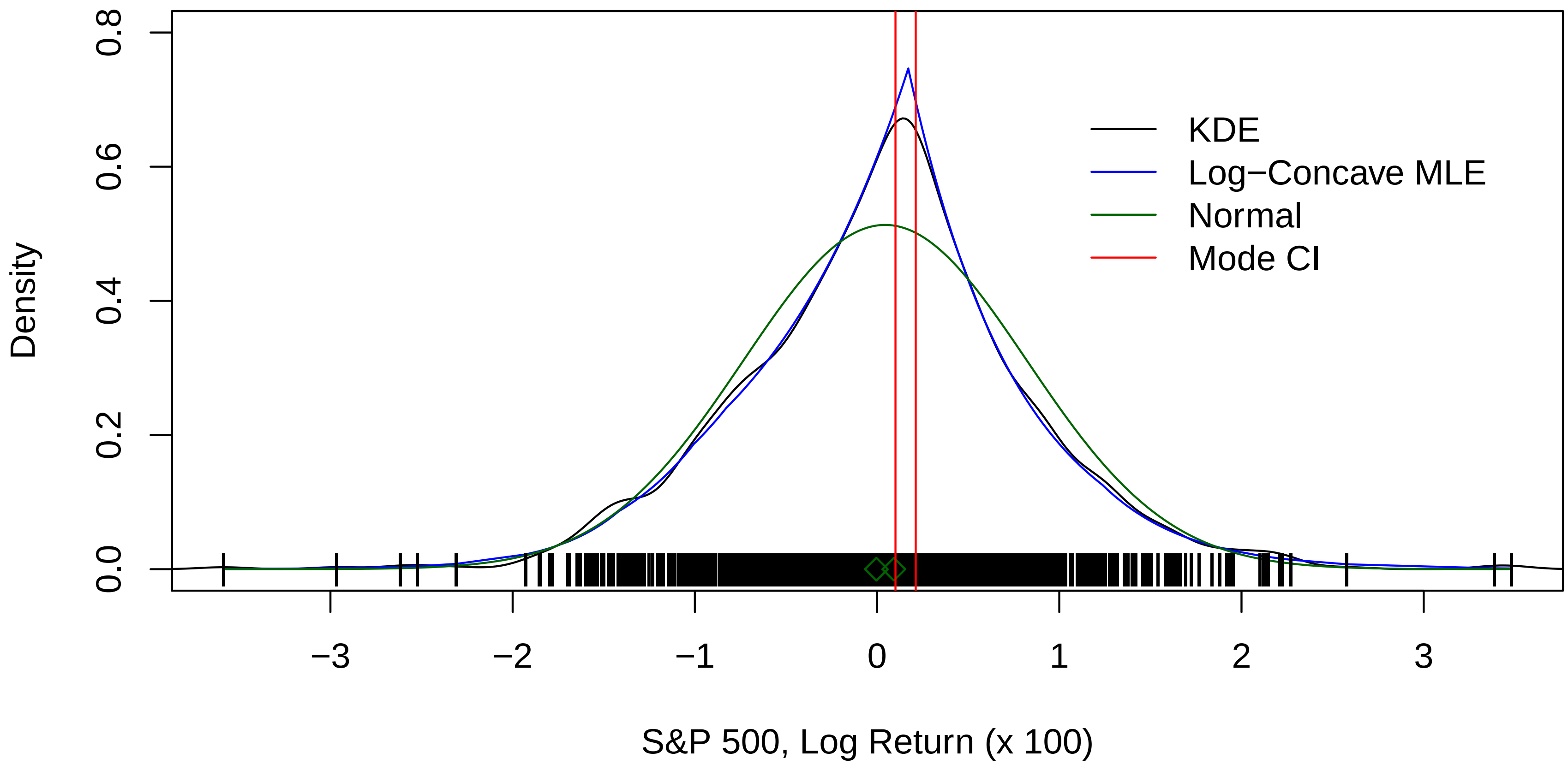
where $D_n = [m_0 - n^{-1/5}c_n, m_0 + n^{-1/5}d_n]$, $c_n \wedge d_n \to_p \infty$, $R_n = o_p(1)$ under H.

Consequently, under H and the assumption $\varphi_0^{(2)}(m_0) < 0$,

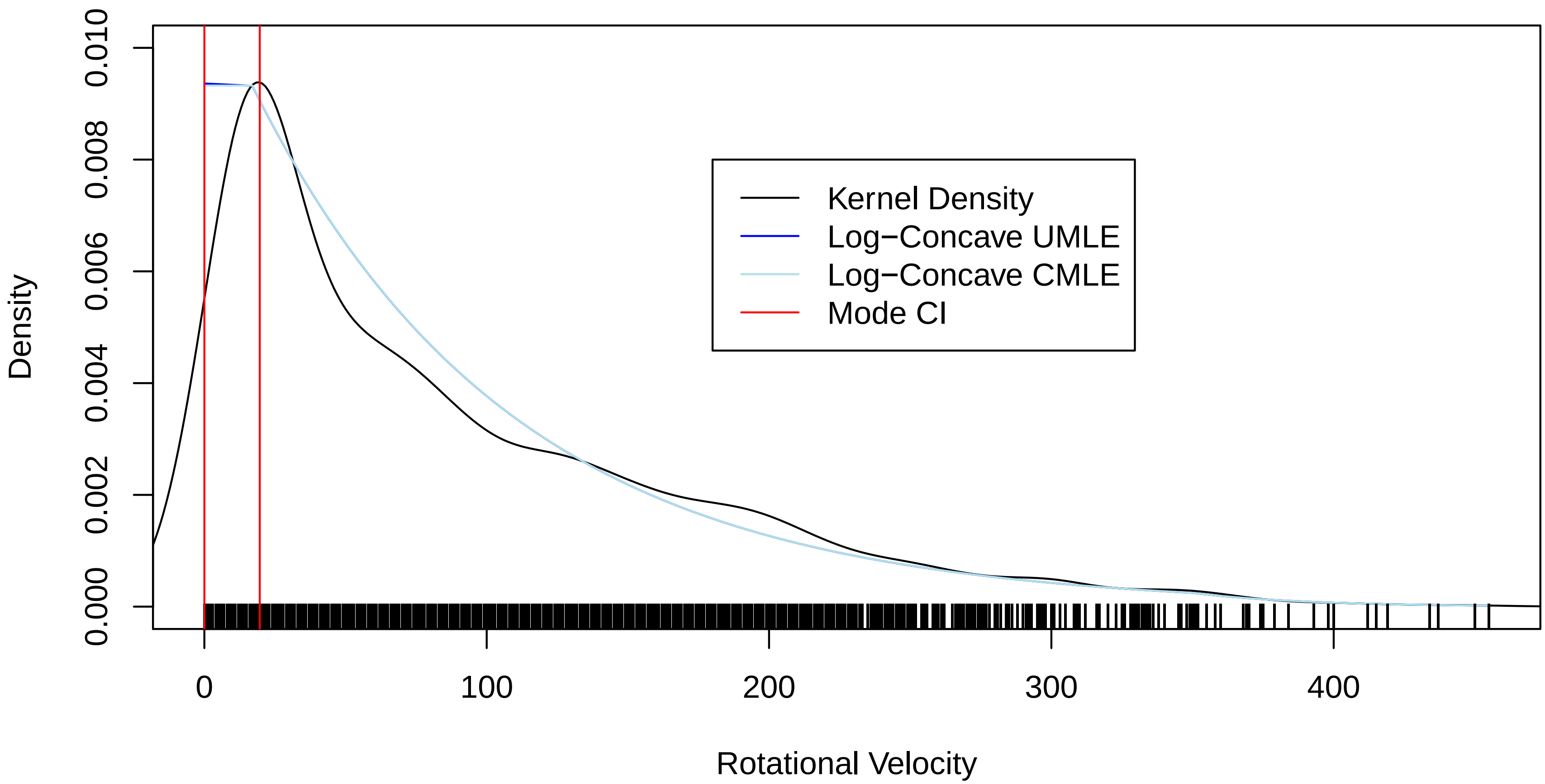
$$2\log\lambda_n \to_d \int_{\mathbb{R}}\left\{\widehat{\varphi}(v)^2 - \widehat{\varphi}^0(v)^2\right\}dv \equiv \mathbb{D}$$

which is free of the parameters $\varphi_0^{(2)}(m_0)$ and $f_0(m_0)$.

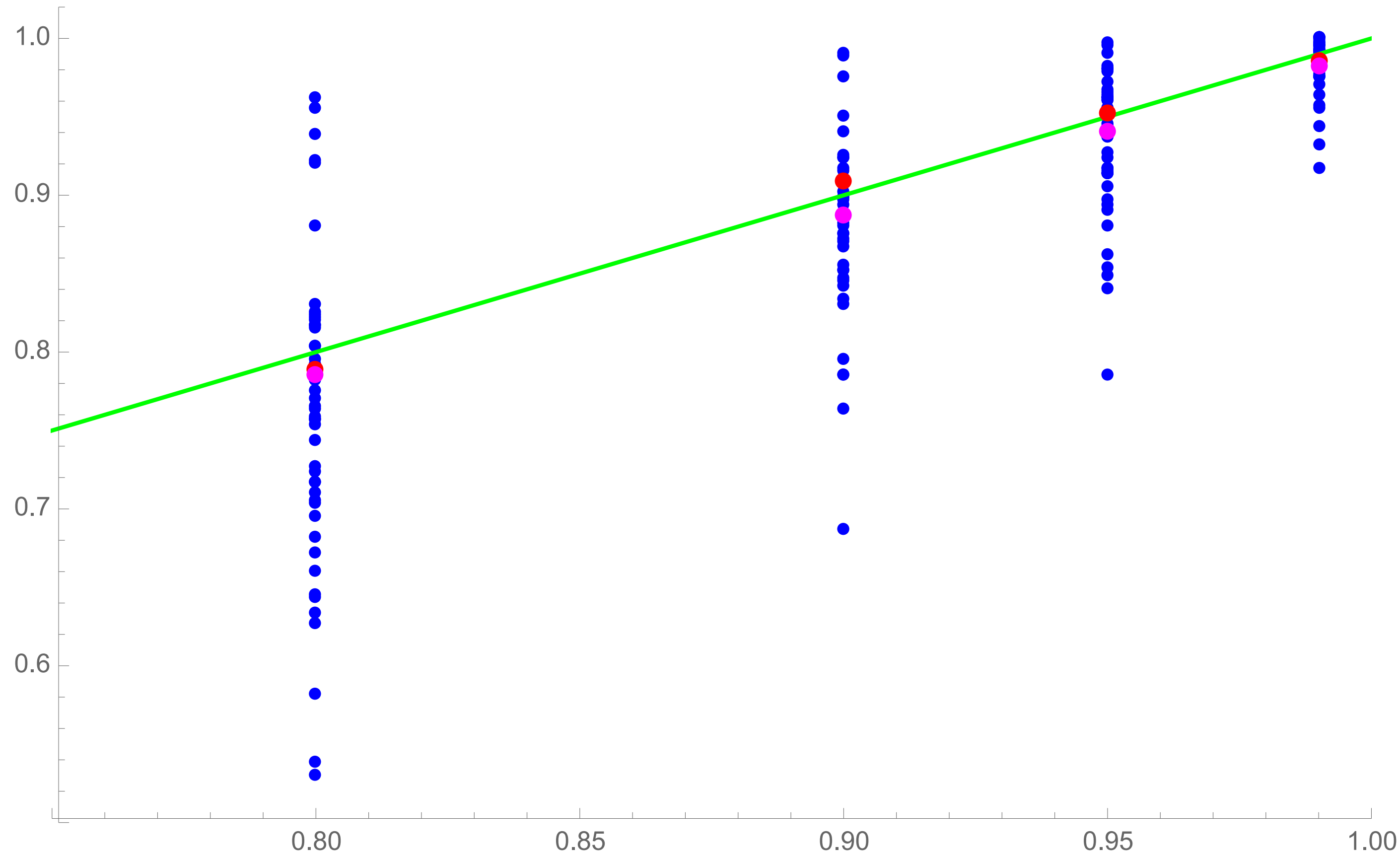

1.0
0.8
0.6
0.4
0.2
0.0
empirical cdf
0.0
0.5
1.0
1.5
$2\log\lambda_n$
Gamma(3,1)
Beta(2,3)
Weibull(1.5,1)
Laplace
$F_{\chi^2_1}$



$n = 1006$ S&P daily log returns, 2003 - 2006



Rotational velocity, $n = 3933$ stars with stellar magnitude > 6.5



Blue: observed coverages Romano bootstrap CI methods (all bandwidths)

Red: LR confidence coverages; chi-square 4 population

Magenta: LR coverages; Gaussian population

What's the difficulty?

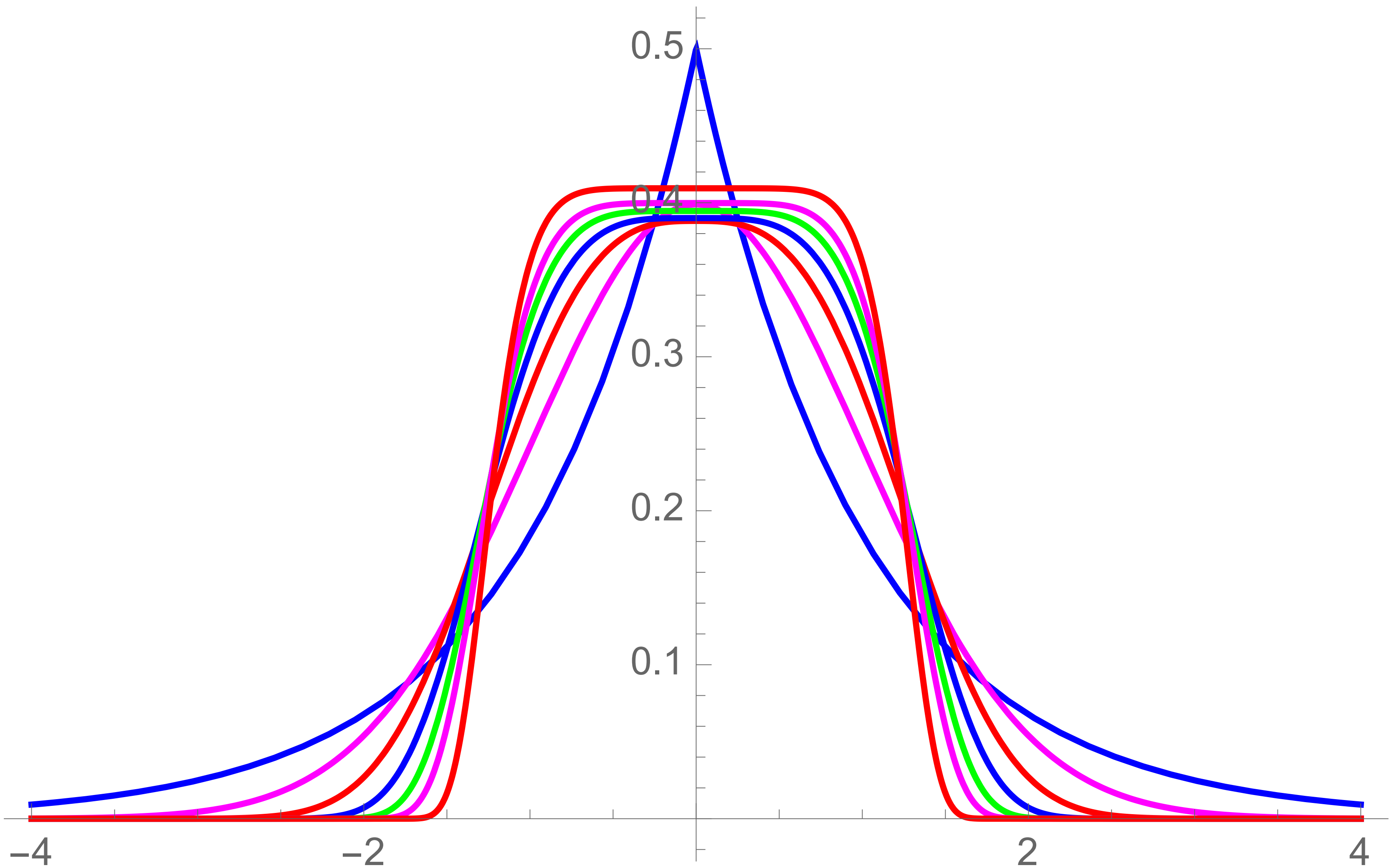
Different limits for different curvature at m_0!?

Suppose that for some $r \geq 1$

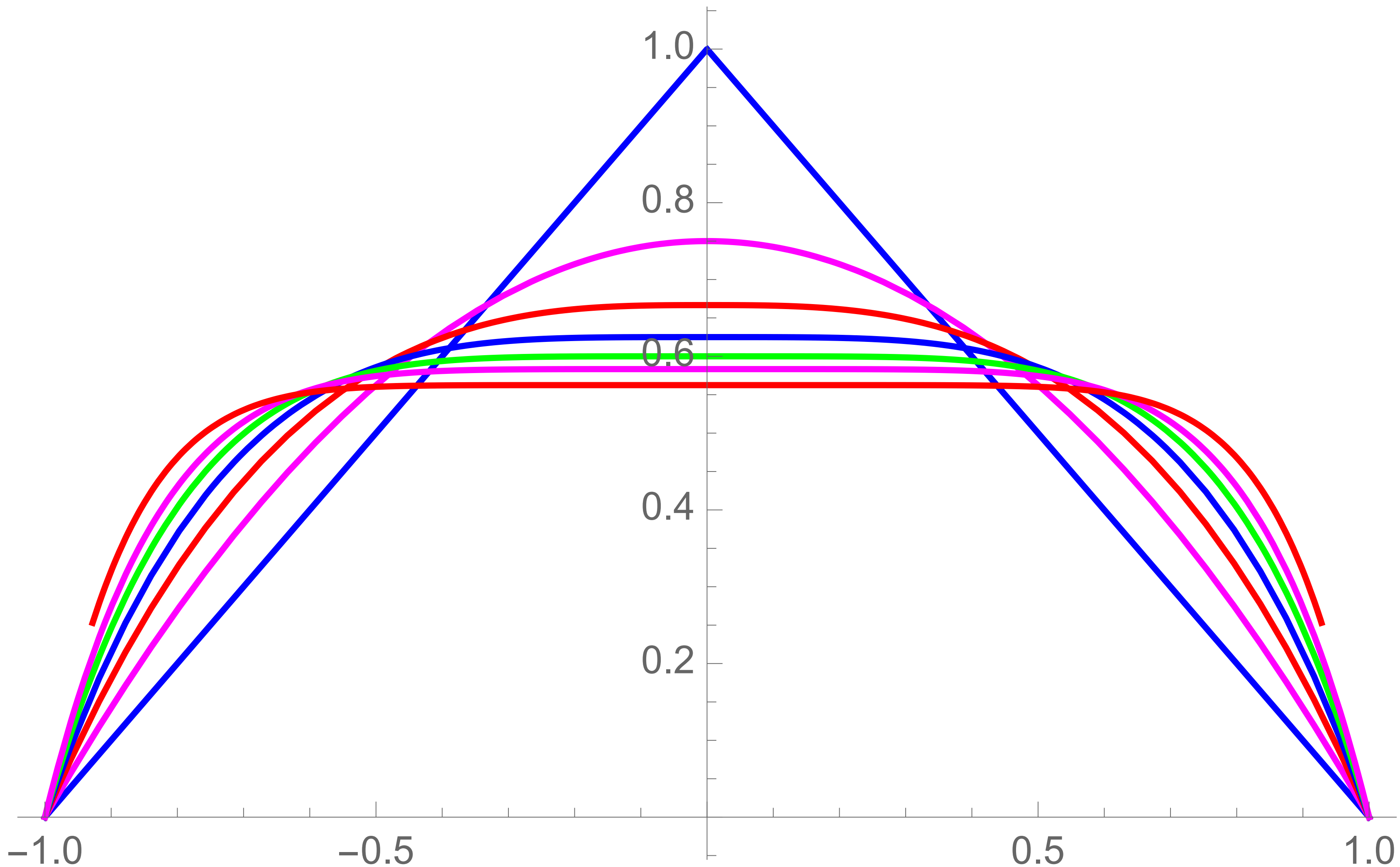
$$f_0(x) = f_0(m_0) - (C + o(1))|x - m_0|^r \quad \text{as} \quad x \to m_0,$$

For example,

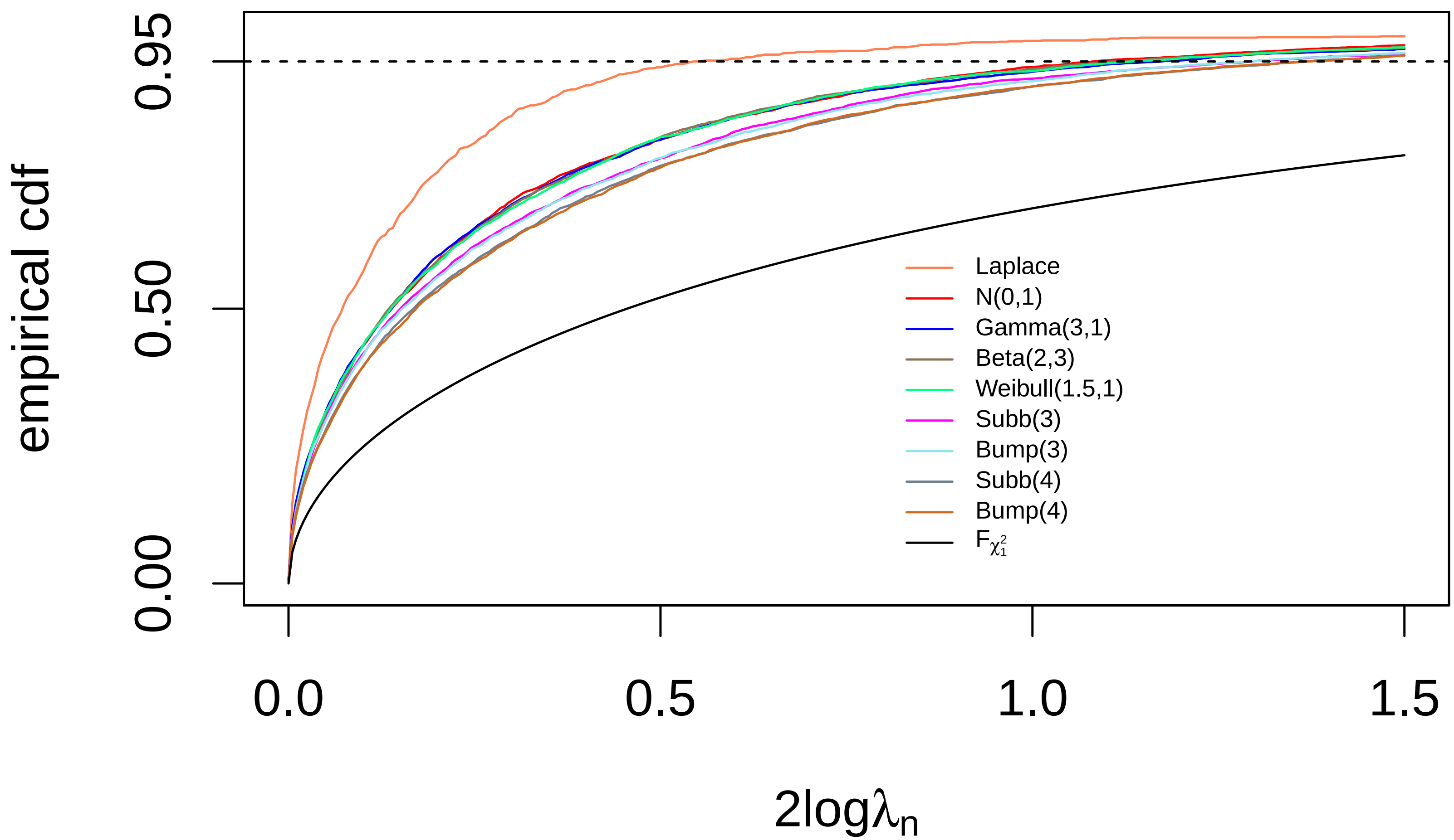
$$\begin{aligned} f_0(x) &= c_r \exp\left(-\frac{|x|^r}{r}\right) && \equiv \text{Subbotin}(r) \\ f_0(x) &= c_r(1 - |x|^r) 1_{[-1,1]}(x) && \equiv \text{Bump}(r) \end{aligned}$$



Subbotin(r), $r \in \{1, 2, 3, 4, 5, 6, 8\}$



Bump(r), $r \in \{1, 2, 3, 4, 5, 6, 8\}$



based on $n = 10000$ and $M = 5000$ Monte Carlo replications
except Laplace: $M = 1200$ Monte Carlo replications

This suggests that

$$2\log\lambda_n \ \to_d \ \int_{\mathbb{R}} \left\{ \widehat{\varphi}_r(v)^2 - \widehat{\varphi}_r^0(v)^2 \right\} dv \equiv \mathbb{D}_r$$

where $r \geq 1$ is unknown.

Two possible approaches:

Approach 1: Find an (upper bound) estimator $\widehat{r}$ of r. (This seems a bit like "tail index estimation" in extreme value theory; we might call it "modal index estimation".) Carry out the test as "reject H if $2\log\lambda_n > d_{\alpha,\widehat{r}}$" where $d_{\alpha,r}$ satisfies $P(\mathbb{D}_r > d_{\alpha,r}) = \alpha \in (0,1)$.

Approach 2: The plots on the previous page suggest that the distributions of $\{\mathbb{D}_r : \ r \geq 1\}$ are tight and may have a limit $\mathbb{D}_\infty$. (This might correspond to the limit white noise model $dX(t) = -f_\infty(t)dt + dW(t)$ where $f_\infty(t) = 0 \cdot 1_{[-1,1]}(t) + \infty \cdot 1_{(1,\infty)}(|t|)$.) If this holds we could carry out our test conservatively as "reject H if $2\log\lambda_n > d_{\alpha,\infty}$" where $d_{\alpha,\infty}$ satisfies $P(\mathbb{D}_r > d_{\alpha,\infty}) = \alpha \in (0,1)$.

Problems

- Limit theory relevant for approaches 1 and 2?
- Does the limit theory for log-concave densities extend to $s-$concave densities with $s \in (-1, 0)$.
- Extend the likelihood ratio approach to inference for modes in $\mathbb{R}^d$ with $d \geq 2$?
  - ▷ If $f = e^{-\varphi}$ with Hess$(\varphi)(m) > 0$, do we have

$$n^{1/(4+d)}(\widehat{M}_n - m) \rightarrow_d \text{something?}$$

  - ▷ Likelihood ratio for testing $M(f) = m_0$? Does

$$2\log\lambda_n \rightarrow_d \quad \text{something universal?}$$

IV. Selected references

- Dümbgen and Rufibach (2009), *Bernoulli*.
- Balabdaoui, Rufibach, & W (2009), *Ann. Statist.* **37**, 1299-1331.
- Han & W (2016): *Ann. Statist.* **44**, 1332 - 1359.
- Doss & W (2016): *Ann. Statist.* **44**, 954-981.
- Doss & W (2016): Mode constrained estimation of a log-concave density. arXiv:1611.10335.
- Doss & W (2016): Inference for the mode of a log-concave density. arXiv:1611.10348.

Merci beaucoup!