Bi-*s**-**Concave Distributions**



Jon A. Wellner (Seattle)

Shape Constrained Methods: Inference, Applications, and Practice

Shape Constrained Methods: Inference, Applications, and Practice Banff International Research Station

January 29 - February 2, 2018

Based on joint work with: Nilanjana Laha

- 0. A mysterious condition: quantile process theory
- 1. Log-concavity and *s*-concavity
- 2. Questions and examples.
- 3. Bi- s^* -concavity.
- 4. Improved confidence bands.
- 5. Open questions.

1. A Mysterious condition: quantile process theory

- Let X_1, \ldots, X_n be i.i.d. F, absolutely continuous with density f.
- Let \mathbb{F}_n denote the empirical distribution function of the X_i 's: $\mathbb{F}_n(x) = n^{-1} \sum_{i=1}^n \mathbb{1}\{X_i \leq x\}.$
- Let \mathbb{F}_n^{-1} denote the empirical quantile function, and let F^{-1} denote the population quantile function, where $F^{-1}(t) \equiv \inf\{x : F(x) \ge t\}, 0 < t < 1$.
- The standardized quantile process \mathbb{Q}_n is defined by

$$\mathbb{Q}_n(t) \equiv g(t)\sqrt{n}(\mathbb{F}_n^{-1}(t) - F^{-1}(t))$$
 for $0 < t < 1$.

where

$$g(t) \equiv f(F^{-1}(t)).$$

is the *density quantile function* or *isoperimetric profile function*.

Now suppose that $X_i \equiv F^{-1}(\xi_i)$ for $1 \le i \le n$ where:

- ξ_1, \ldots, ξ_n are i.i.d. Uniform(0, 1) random variables.
- \mathbb{G}_n is the empirical d.f. of the ξ_i 's.
- \mathbb{G}_n^{-1} is the empirical quantile function of the ξ_i 's.
- $\mathbb{V}_n(t) \equiv \sqrt{n}(\mathbb{G}_n^{-1}(t) t)$ is the uniform quantile process.

Csörgő and Révész (1978) imposed the following mysterious condition in their study of the asymptotic equivalence of \mathbb{V}_n and \mathbb{Q}_n^0 , the version of \mathbb{Q}_n with the X_i 's constructed in terms of the ξ_i 's as above.

Let $J(F) \equiv \{x \in \mathbb{R} : 0 < F(x) < 1\}$, and assume

$$\gamma(F) \equiv \sup_{x \in J(F)} F(x)(1 - F(x)) \frac{|f'(x)|}{f^2(x)} \le \text{some } M < \infty.$$
(1)

Then, for any (small)
$$r > 0$$

 $\|\mathbb{Q}_n^0 - \mathbb{V}_n\|_{\infty} = O\left(n^{-1/2} (\log \log n)^M (\log n)^{(1+r)(M-1)}\right)$ a.s.

Define the CR(x) and $CR_m(x)$ functions as follows:

$$CR(x) \equiv F(x)(1 - F(x)) \frac{|f'(x)|}{f^2(x)},$$

 $CR_m(x) \equiv \min\{F(x), 1 - F(x)\} \frac{|f'(x)|}{f^2(x)}.$

The condition (1) has also appeared in the study of transportation distances between the empirical measure and true measure \mathbb{P}_n and P on \mathbb{R} ; see e.g.

- del Barrio, E., Giné, E., and Utzet, F. (2005). Asymptotics for L2 functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli* 11, 131 - 189.
- Bobkov, S. and Ledoux, M. (2014). One Dimensional empirical measures, order statistics, and Kantorovich transport distances. *Memoirs of the American Mathematical Society*, to appear. (especially see p. 45)

If a density f on \mathbb{R}^d is of the form

$$f(x) \equiv f_{\varphi}(x) = \begin{cases} (\varphi(x))^{1/s}, & \varphi \ convex, \text{ if } s < 0\\ \exp(-\varphi(x)), & \varphi \ convex, \text{ if } s = 0\\ (\varphi(x))^{1/s}, & \varphi \ concave, \text{ if } s > 0, \end{cases}$$

then f is s-concave.

The classes of all densities f on \mathbb{R}^d of these forms are called the classes of s-concave densities, \mathcal{D}_s . The following inclusions hold: if $-\infty < s < 0 < r < \infty$, then

$$\mathcal{D}_{-\infty} \supset \mathcal{D}_s \supset \mathcal{D}_0 \supset \mathcal{D}_r \supset \mathcal{D}_\infty$$

Properties of *s***-concave densities:**

- Every s-concave density f is quasi-concave.
- The Student t_{ν} density, $t_{\nu} \in \mathcal{P}_s$ for $s \leq -1/(1+\nu)$. Thus the Cauchy density $(= t_1)$ is in $\mathcal{P}_{-1/2} \subset \mathcal{P}_s$ for $s \leq -1/2$.
- The classes \mathcal{D}_s have interesting closure properties under convolution and marginalization which follow from the Borell-Brascamp-Lieb inequality: let $0 < \lambda < 1$, $-1/d \leq s \leq \infty$, and let $f, g, h : \mathbb{R}^d \to [0, \infty)$ be integrable functions such that

$$h((1-\lambda)x+\lambda y)\geq M_s(f(x),g(x),\lambda)$$
 for all $x,y\in \mathbb{R}^d$

where

$$M_s(a,b,\lambda) = ((1-\lambda)a^s + \lambda b^s)^{1/s}, \quad M_0(a,b,\lambda) = a^{1-\lambda}b^{\lambda}.$$

Then

$$\int_{\mathbb{R}^d} h(x) dx \ge M_{s/(sd+1)} \left(\int_{\mathbb{R}^d} f(x) dx, \int_{\mathbb{R}^d} g(x) dx, \lambda \right).$$

If p is an s-concave density with corresponding probability measure P, then for Borel sets A, B

$$P(\lambda A + (1 - \lambda)B) = \int_{\lambda A + (1 - \lambda)B} p(x)dx$$

$$\geq M_{s/(sd+1)} \left(\int_{A} p(x)dx, \int_{B} p(x)dx, \lambda \right)$$

$$= M_{s/(sd+1)} \left(P(A), P(B), \lambda \right)$$

When d = 1, let $s^* \equiv s/(1+s)$ for $s \in (-1,\infty)$. Thus $0^* = 0$.

Definition: Dümbgen, Kolesnyk, and Wilke (2017) A distribution function F on \mathbb{R} is bi-log-concave if both logF and log(1 - F) are concave functions from \mathbb{R} to $[-\infty, 0]$.

- DKW (2017) noted that if F has log-concave density f = F', then F is bi-log-concave. This follows from the Borell-Brascamp-Lieb inequality with d = 1, s = 0, and considering sets of the form $(-\infty, x]$ or (x, ∞) .
- But ... bi-log-concavity of F is a much weaker constraint:
 - ▷ While any log-concave density is unimodal,
 - \triangleright a bi-log-concave distribution function F may have a density with an arbitrary number of modes; e.g.

$$f_{k,a}(x) = \left(1 + \frac{a\sin(2\pi kx)}{k\pi}\right) \mathbf{1}_{[0,1]}(x) / C(k,a).$$

for a small, is bi-log-concave











Multi-modal perturbed uniform density, $f_{k,a}$ with k = 4, a = 0.795



C-R functions, multi-modal perturbed uniform density: $f_{k,a}$ with k = 4, a = 0.0795

Let

$$J(F) \equiv \{ x \in \mathbb{R} : 0 < F(x) < 1 \}.$$

Then a distribution function F is non-degenerate if $J(F) \neq \emptyset$.

Theorem 1. (DKW-2017).

If F is non-degenerate, the following four statements are equivalent:

(i) F is bi-log-concave, and we write

 $F \in \mathcal{F}_{blc} = \mathcal{F}_{bi-s^*}.$

(ii) F is continuous on \mathbb{R} and differentiable on J(F) with derivative f = F' such that

$$F(x+t) \begin{cases} \leq F(x) \exp\left(\frac{f(x)}{F(x)}t\right), \\ \geq 1 - (1 - F(x)) \exp\left(-\frac{f(x)}{1 - F(x)}t\right) \end{cases}$$

for all $x \in J(F)$ and $t \in \mathbb{R}$.

(iii) F is continuous on \mathbb{R} and differentiable on J(F) with derivative f = F' such that the hazard function f/(1-F) is non-decreasing and the reverse hazard function f/F in non-increasing on J(F).

(iv) F is continuous on \mathbb{R} and differentiable on J(F) with bounded and strictly positive derivative f = F'. Furthermore, f is locally Lipschitz continuous on J(F) with L^1 -derivative f' = F''satisfying

$$\frac{-f^2}{1-F} \le f' \le \frac{f^2}{F},$$

or, equivalently,

$$\frac{-f}{1-F} \le \frac{f'}{f} \le \frac{f}{F}.$$

Corollary. If F is bi-log-concave

$$ilde{\gamma}(F)\equiv \sup_{x\in\mathbb{R}}\{F(x)\wedge(1-F(x))\}rac{|f'(x)|}{f^2(x)}\leq 1,$$
 $\gamma(F)\equiv \sup_{x\in\mathbb{R}}F(x)(1-F(x))rac{|f'(x)|}{f^2(x)}\leq 1.$

If a density f on \mathbb{R}^d is of the form

$$f(x) \equiv f_{\varphi}(x) = \begin{cases} (\varphi(x))^{1/s}, & \varphi \ convex, \text{ if } s < 0\\ \exp(-\varphi(x)), & \varphi \ convex, \text{ if } s = 0\\ (\varphi(x))^{1/s}, & \varphi \ concave, \text{ if } s > 0, \end{cases}$$

then f is s-concave.

The classes of all densities f on \mathbb{R}^d of these forms are called the classes of s-concave densities, \mathcal{P}_s . The following inclusions hold: if $-\infty < s < 0 < r < \infty$, then

$$\mathcal{P}_{\infty} \subset \mathcal{P}_r \subset \mathcal{P}_0 \subset \mathcal{P}_s \subset \mathcal{P}_{-\infty}$$

Questions:

• Q1: What if the density f is s-concave with $s \neq 0$. In particular, what if $f \in \mathcal{D}_s$ with s < 0 where we know (Borell, Brascamp & Lieb, Rinott, ...)

$$\mathcal{D}_{-\infty} \supset \mathcal{D}_s \supset \mathcal{D}_0 \supset \mathcal{D}_r \supset \mathcal{D}_\infty$$

for $-\infty < s < 0 < r < \infty$?

- Q2: If $f \in D_s$, is there a class of bi- s^* -concave distribution functions F with the property that F and 1 F are s^* -concave?
- Q3: Is there an analogue of Theorem 1 including an analogue of Theorem 1(iv) with the corollary that γ(F) is bounded by some function of s?

From Borell, Brascamp, & Lieb, Rinott, we know that if $f \in \mathcal{D}_s$ for s > -1, then the measure $P_f(A) = \int_A f d\lambda$ for Borel sets A is t-concave with $t = s/(1+s) \equiv s^*$ for s > -1. Thus by taking $A = (-\infty, x]$, it follows that $x \mapsto F(x)$ is s^* -concave; similarly, taking $A = [x, \infty)$ it follows that $x \mapsto 1 - F(x)$ is s^* -concave.

Example 1. t_r densities: s = -1/(1+r); $s^* = s/(1+s) = -1/r$. Suppose that

$$f_r(x) = \frac{C_r}{\left(1 + \frac{x^2}{r}\right)^{(r+1)/2}}$$

where $C_r = \Gamma((r+1)/2)/(\sqrt{\pi}\Gamma(r/2))$. Then f_r is *s*-concave for all $s \leq -(1+r)^{-1}$. By the Borell-Brascamp-Lieb-Rinott correspondence between *s*-concave densities we know that

$$x\mapsto F_r(x)^{s^*}$$
 and $x\mapsto (1-F_r(x))^{s^*}$

are convex.

Here are some plots, for $r \in \{1/8, 1/4, 1/2, 1, 4, 16\}$, and hence $s \in \{-8/9, -4/5, -2/3, -1/2, -1/3, -1/5\}$:

•
$$f_r$$
,

•
$$f_r^s$$
, $s = -1/(1+r)$.

•
$$f_r/(1-F_r)^{1-s^*}$$
, $s^* = s/(1+s) = -1/r$.

• $CRm(x, f) \equiv \min\{F(x), 1 - F(x)\}f'(x)/f(x)^2$ for $f = f_r$. $CR(x, f) \equiv F(x)(1 - F(x))f'(x)/f(x)^2$.



 f_r , $r \in \{1/8, 1/4, 1/2, 1, 4, 16\}$ or $s \in \{-8/9, -4/5, -2/3, -1/2, -1/3, -1/5\}$



 f_r^s , $r \in \{1/8, 1/4, 1/2, 1, 4, 16\}$ or $s \in \{-8/9, -4/5, -2/3, -1/2, -1/3, -1/5\}$





$$CR_m(x, f) \equiv \min\{F(x), 1 - F(x)\}f'(x)/f(x)^2 \text{ for } f = f_r.$$

Here the black bounding lines at the top and bottom are given by

$$1 - s^* = \frac{1}{1+s} = \frac{1}{1-8/9} = 9$$
 since $s = -\frac{1}{1+1/8} = -\frac{8}{9}$.

Example 2. (Mixtures of t_r) Suppose that

$$f(x) = f(x; r, \delta) \equiv \frac{1}{2}g_r(x - \delta) + \frac{1}{2}g_r(x + \delta)$$

where g_r is the t_r -density in Example 1 and where $\delta > 0$ is not too large. For example here are Figures 2 - 5 of Laha & W (2017). Showing g_r with r = 1 and $\delta = 1.3$





f', black; bounds f^2/F and $f^2/(1-F)$ from bi $-s^*$ -concave characterization



Example 3. (symmetric beta densities). Now consider the family of s-concave densities with s > 0 given for any $r \in (0, \infty)$ by

$$f_r(x) = \sqrt{r}C_r(1-x^2)^{r/2}\mathbf{1}_{[-1,1]}(x)$$

where $C_r \equiv \Gamma((3+r)/2)/(\sqrt{\pi r}\Gamma(1+r/2))$. Then $f_r \in \mathcal{P}_s$. with s = 2/r. Here are some plots, for $r \in \{1/8, 1/2, 2, 4, 8, 16\}$, and hence $s \in \{16, 4, 1, 1/2, 1/4, 1/8\}$:

•
$$f_r$$
,

- f_r^s , s = 2/r.
- $f_r/(1-F_r)^{1-s^*}$, $s^* = s/(1+s) =$.
- $CRm(x, f) \equiv \min\{F(x), 1 F(x)\}f'(x)/f(x)^2$ for $f = f_r$. $CR(x, f) \equiv F(x)(1 - F(x))f'(x)/f(x)^2$.



Symmetrized beta densities f_r with $r \in \{1/8, 1/4, 1/2, 2, 4, 16\}$







CRm(x) for f_r symmetrized Beta, $r \in \{1/8, 1/4, 1/2, 2, 4, 16\}$ Here the black bounding lines at the top and bottom are given by the bound for the biggest class, namely for r = 16, so s = 1/8and

$$1 - s^* = \frac{1}{1+s} = \frac{1}{1+1/8} = \frac{8}{9}$$
 since $s = \frac{2}{16} = \frac{1}{8}$.

Definition.

- For $s \in (-1, \infty)$, let $s^* \equiv s/(1+s) \in (-\infty, 1]$.
- For $s \in (-1,0)$, a distribution function F on \mathbb{R} is bi- s^* -concave if both $x \mapsto F^{s^*}(x)$ and $x \mapsto (1-F)^{s^*}(x)$ are convex functions of $x \in J(F)$.
- For s ∈ (0,∞), F on R is bi-s*-concave if x → F^{s*}(x) is concave for x ∈ (inf J(F),∞) and and x → (1 F)^{s*}(x) is concave for x ∈ (-∞, sup J(F)).
- For s = 0, F on \mathbb{R} is bi-0-concave or bi-log-concave if both $x \mapsto \log F(x)$ and $x \mapsto \log(1 F(x))$ are concave functions of $x \in J(F)$.

Theorem 2. (Bi- s^* -characterization theorem) Let $s \in (-1, \infty]$. For a non-degenerate distribution function F the following four statements are equivalent:

(i) F is bi- s^* -concave.

(ii) F is continuous on \mathbb{R} and differentiable on J(F) with derivative f = F'. Moreover when $s \leq 0$,

$$F(x+t) \begin{cases} \leq F(x) \cdot \left(1 + s^* \frac{f(x)}{F(x)} t\right)_+^{1/s^*} \\ \geq 1 - (1 - F(x)) \cdot \left(1 - s^* \frac{f(x)}{1 - F(x)} t\right)_+^{1/s^*} \end{cases}$$
(2)

for all $x \in \mathbb{R}$ and $t \in \mathbb{R}$. When s > 0,

$$F(x+t) \begin{cases} \leq F(x) \cdot \left(1 + s^* \frac{f(x)}{F(x)} t\right)_+^{1/s^*}, & \text{for } t \in (a-x,\infty) \\ \geq 1 - (1 - F(x)) \cdot \left(1 - s^* \frac{f(x)}{1 - F(x)} t\right)_+^{1/s^*}, & \text{for } t \in (-\infty, b-x) \end{cases}$$
(3)

for all $x \in J(F)$.

(iii) F is continuous on \mathbb{R} and differentiable on J(F) with derivative f = F' such that the s^* -hazard function $f/(1-F)^{1-s^*}$ is non-decreasing, and the reverse s^* -hazard function f/F^{1-s^*} is non-increasing on J(F).

(iv) F is continuous on \mathbb{R} and differentiable on J(F) with bounded and strictly positive derivative f = F'. Furthermore, f is locally Lipschitz-continuous on J(F) with L^1 -derivative f' = F''satisfying

$$-(1-s^*)\frac{f^2}{1-F} \le f' \le (1-s^*)\frac{f^2}{F}.$$
 (4)

Corollary.

Suppose that F is bi-s^{*}-concave for $s \in (-1, \infty]$. Then

$$\gamma(F) = \sup_{x \in J(F)} F(x)(1 - F(x)) \frac{|f'(x)|}{f^2(x)} = 1 - s^* = \frac{1}{1 + s},$$

 $\quad \text{and} \quad$

$$\tilde{\gamma}(F) = \sup_{x \in J(F)} \min\{F(x), 1 - F(x)\} \frac{|f'(x)|}{f^2(x)} \le 1 - s^* = \frac{1}{1+s}.$$

Suppose F is bi-logconcave. Given valid nonparametric confidence bands for F, can we improve them under the assumption of bi-log-concavity?

YES! DKW (2017)

Suppose that (L_n, U_n) is a $(1 - \alpha)$ -confidence band for F with $0 < \alpha \leq 1/2$. Thus $L_n = L_{n,\alpha}(\cdot|X_1, \ldots, X_n) < 1$ and $U_n = U_{n,\alpha}(\cdot|X_1, \ldots, X_n)$ are non-decreasing functions on \mathbb{R} with $L_n \leq U_n$ pointwise and

 $P_F(L_n(x) \leq F(x) \leq U_n(x) \text{ for all } x \in \mathbb{R}) = 1 - \alpha.$

Example: (Kolmogorov-Smirnov band). Let \mathbb{F}_n be the empirical distribution function as in Section 1. Then

$$[L_n(x), U_n(x)] \equiv \left[\mathbb{F}_n(x) - \frac{\kappa_{\alpha, n}^{KS}}{n^{1/2}}, \mathbb{F}_n(x) + \frac{\kappa_{\alpha, n}^{KS}}{n^{1/2}} \right] \cap [0, 1]$$

where $\kappa_{\alpha,n}^{KS}$ denotes the $(1-\alpha)$ -quantile of

$$\sup_{x\in\mathbb{R}}n^{1/2}|\mathbb{F}_n(x)-F(x)|;$$

cf. Shorack and W (1986) (and note that $\kappa_{\alpha,n}^{KS} \leq \sqrt{\log(2/\alpha)/2}$ by Massart's inequality).

If $F \in F_{blc}$, find a new band for F as follows:

$$L_n^0(x) \equiv \inf\{G(x) : G \in \mathcal{F}_{blc}, L_n \leq G \leq U_n\}, \\ U_n^0(x) \equiv \sup\{G(x) : G \in \mathcal{F}_{blc}, L_n \leq G \leq U_n\}.$$

If no bi-log-concave distribution function fits into the band (L_n, U_n) , set $L_n^0 \equiv 1$ and $U_n^0 \equiv 0$ and conclude with confidence $1 - \alpha$ that $F \notin \mathcal{F}_{blc}$. But if $F \in \mathcal{F}_{blc}$, this happens with probability at most α : by the construction of (L_n^0, U_n^0)

$$P_F(L_n^0) \leq F \leq U_n^0) = P(L_n \leq F \leq U_n) = 1 - \alpha$$
 if $F \in \mathcal{F}_{blc}$.

Similarly, if $F \in \mathcal{F}_{bi-s^*}$, find a new band for F by refining the given band.



CEO salary data; Woolridge 2002: n = 177; $X_i = \log_{10} Y_i$ $Y_i =$ annual salaries of CEO's in multiples of USD 95% unconstrained band (black); log-concave-constrained band (blue)



 $bi-s^*$ -constrained band, s = 0.10 (gray)



95% unconstrained band (black); log-concave-constrained band (blue); bi $-s^*$ -constrained band, s = 0.20 (green)



95% unconstrained band (black); log-concave-constrained band (blue); bi $-s^*$ -constrained band, s = 0.30 (red)



95% unconstrained band (black); log-concave-constrained band (blue); bi $-s^*$ -constrained band, s = 0.30 (red)



95%unconstrained band (black); log-concave-constrained band (blue);

Questions and further problems: Q1. What can be said about estimation of s^* (and s)?

- **Q2.** Can anything be said when f is s-concave with $s \leq -1$?
- **Q3.** Bi-log-concave or $bi-s^*$ -concave in higher dimensions?
- **Q4.** What are the "right" hypotheses for the study of transportation (Wasserstein) distances for empirical measures on \mathbb{R}^d with $d \ge 2$?

- Dümbgen, L., Kolesnyk, P., and Wilke, R. (2017). Bi-logconcave distribution functions. *J. Statist. Planning and Inference* 184, 1 - 17.
- del Barrio, E., Giné, E., and Utzet, F. (2005). Asymptotics for L₂ functionals of the empirical quantile process, with applications to tests of fit based on weighted Wasserstein distances. *Bernoulli* **11**, 131 - 189.
- Laha, N. and Wellner, J. A. (2017). Bi-s*-concave distributions. Available as arXiv:1705.00252.

Thank You!