Chernoff's distribution is log-concave

And more

Jon A. Wellner

University of Washington, Seattle, visiting Heidelberg

Seminaire Point de vue, Université Paris-Diderot Paris 7

23 January 2012

Seminaire Point de vue, Paris

Part 2

Based on joint work with:

- Fadoua Balabdaoui

Outline

- 1 Two limit theorems
  (how does Chernoff's distribution f appear?)
- 2 Graphical evidence for log-concavity of f
- 3 Chernoff and Groeneboom's formula
- 4 Chernoff's density f is a PF_2 density; i.e. log-concave
- 5 Is Chernoff's density f "Strongly log-concave"?
- 6 Relations with other classes: PF_∞ and HM_∞?
- 7 Summary; conjectures and open problems.

1. Two limit theorems: how does Chernoff's distribution appear?

First Limit Theorem:

- Suppose $X_1, \ldots, X_n$ are i.i.d. $EX_1 = \mu$, $E(X^2) < \infty$, $\sigma^2 = Var(X)$.

- Then, by the classical Central Limit Theorem:

$$\sqrt{n}(\overline{X}_n - \mu) \rightarrow_d \sigma Z_1 \sim N(0, \sigma^2).$$

  The limit random variable $\sigma Z_1 \sim N(0, \sigma^2)$ has density

$$\phi_\sigma(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{x^2}{2\sigma^2}\right) = e^{-V(x)},$$

$$V(x) = -\log \phi_\sigma(x) = \frac{x^2}{2\sigma^2} + \log(\sqrt{2\pi}\sigma)$$

$$V''(x) = (-\log \phi_\sigma)''(x) = \frac{1}{\sigma^2} > 0.$$

1. Two limit theorems: how does Chernoff's distribution appear?

Second Limit Theorem

Monotone regression problem: suppose that

- $r(x)$ is increasing for $x \in [0, 1]$.
- For $i \in \{1, \ldots, n\}$, $x_i = i/(n+1)$, and ϵ_i are i.i.d. with $E(\epsilon_i) = 0$, $\sigma^2 = E(\epsilon_i^2) < \infty$.
- $Y_i = r(x_i) + \epsilon_i \equiv \mu_i + \epsilon_i$, for $i \in \{1, \ldots, n\}$.
- Observe $(x_1, Y_1), \ldots, (x_n, Y_n)$. Estimate μ_j and $\underline{\mu}$:

$$\widehat{\mu}_j = \max_{i \le j} \min_{k \ge j} \left\{ \frac{\sum_{l=i}^{k} Y_l}{k - i + 1} \right\},$$

$$\underline{\widehat{\mu}} = (\widehat{\mu}_1, \ldots, \widehat{\mu}_n) \equiv T\underline{Y}$$

$$= \text{least squares projection of } \underline{Y} \text{ onto } K_n,$$

$$K_n = \{y \in \mathbb{R}^n : \quad y_1 \le \cdots \le y_n\}.$$

- $\widehat{r}_n(x_0) \equiv \widehat{r}_n(x_j) = \widehat{\mu}_j$ for $x_j \le x_0 < x_{j+1}$.

1. Two limit theorems: how does Chernoff's distribution appear?

Brunk (1970):
If $r'(x_0) > 0$, r' continuous in a neighborhood of x_0, then

$$n^{1/3}(\widehat{r}_n(x_0) - r(x_0)) \rightarrow_d (\sigma^2 r'(x_0)/2)^{1/3}(2Z_2).$$

where

$$\begin{aligned} 2Z_2 &= \text{slope at zero of the greatest convex minorant of } W(t) + t^2 \\ &\stackrel{d}{=} 2 \cdot \text{argmin}\{W(t) + t^2\} \end{aligned}$$

where W is two-sided Brownian motion started at 0.
The density f of Z_2 is called Chernoff's density.

1. Two limit theorems: how does Chernoff's distribution appear?

- Nonparametric estimation of a monotone function
- Four problems:
  - ▷ Estimation of a monotone regression function:
    van Eeden (1957), Brunk (1970)
  - ▷ Estimation of a monotone decreasing density:
    Grenander (1956), Prakasa Rao (1969)
  - ▷ Estimation of a monotone hazard function:
    Grenander (1956), Prakasa Rao (1970)
  - ▷ Interval censoring problems:
    Ayer, Brunk, Ewing, Reid, Silverman (1955),
    Groeneboom (1988; 1992)

1. Two limit theorems: how does Chernoff's distribution appear?

- In each case:
  - ▷ There is a monotone function m to be estimated
  - ▷ There is a natural nonparametric estimator $\widehat{m}_n$.
  - ▷ If $m'(x_0) \neq 0$ and m' continuous at x_0, then

$$n^{1/3}(\widehat{m}_n(x_0) - m(x_0)) \to_d C(m, x_0) 2Z_2$$

where

$$\begin{aligned} 2Z_2 &= 2 \cdot \text{argmin}\{W(t) + t^2\} \\ &\stackrel{d}{=} \text{slope at zero of greatest convex minorant} \\ &\qquad \text{of } \ W(t) + t^2. \end{aligned}$$

1. Two limit theorems: how does Chernoff's distribution appear?

First appearance of Z_2:
Chernoff (1964), *Estimation of the mode*:

- $X_1, \ldots, X_n$ i.i.d. with density f and distribution function F.
- Fix $a > 0$; $\widehat{x}_a \equiv$ center of the interval of length $2a$ containing the most observations.
- $x_a \equiv$ center of the interval of length $2a$ maximizing $F(x+a) - F(x-a)$.
- Chernoff shows:
  - ▷ $n^{1/3}(\widehat{x}_a - x_a) \to_d \left(\frac{8f(x_a+a)}{c}\right)^{1/3} Z_2$ where
  $$c \equiv f'(x_a - a) - f'(x_a + a).$$
  - ▷ $f(z) \equiv f_{Z_2}(z) = \frac{1}{2} g(z) g(-z)$ where
  $$g(t) \equiv \lim_{x \nearrow t^2} \frac{\partial}{\partial x} u(t,x) = \lim_{x \nearrow t^2} u_x(t,x),$$

1. Two limit theorems: how does Chernoff's distribution appear?

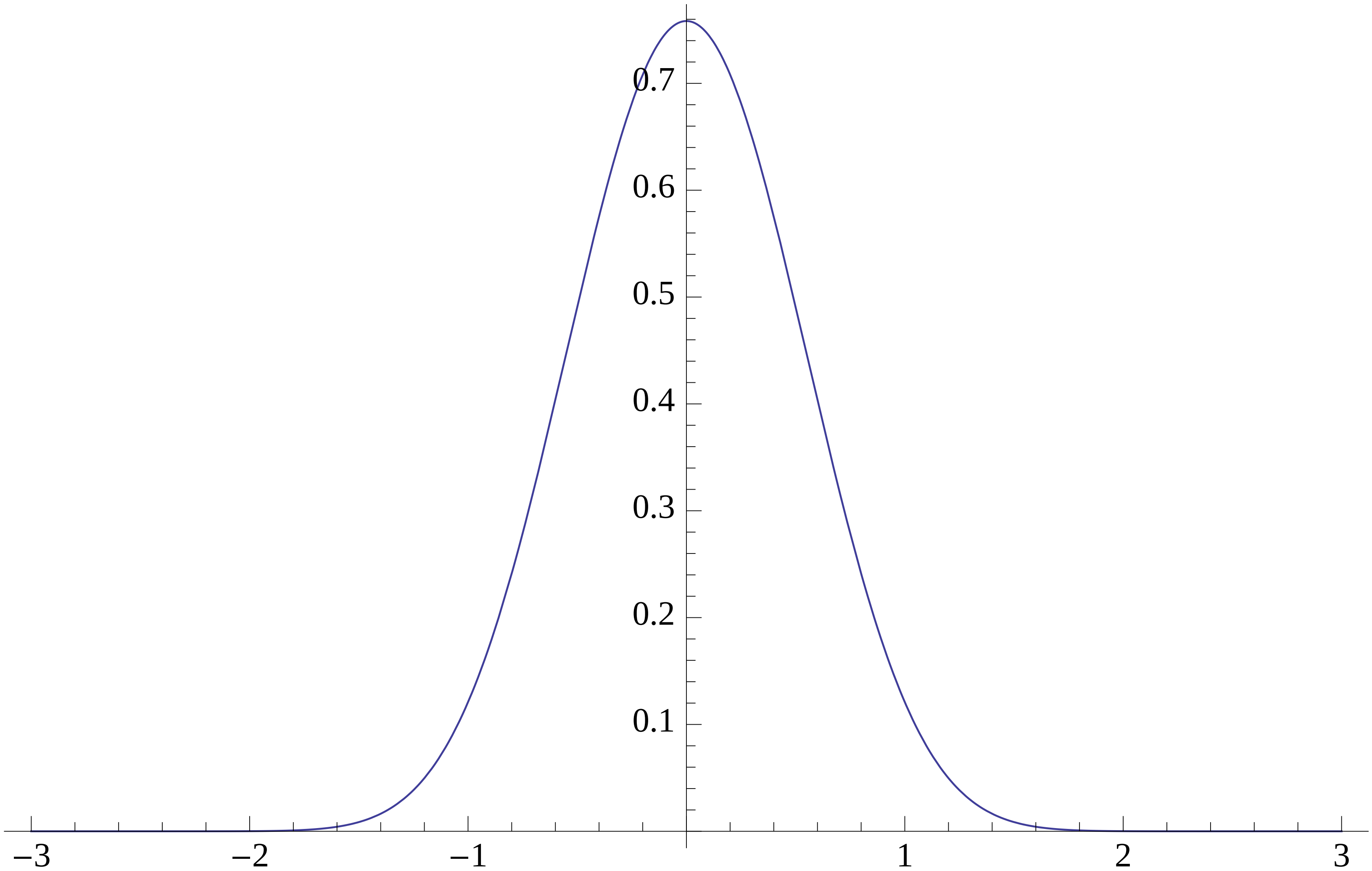
▷ $u(t,x) \equiv P^{(t,x)}(W(z) > z^2, \quad \text{for some } z \geq t)$ is a solution to the backward heat equation

$$\frac{\partial}{\partial t} u(t,x) = -\frac{1}{2} \frac{\partial^2}{\partial x^2} u(t,x)$$

under the boundary conditions

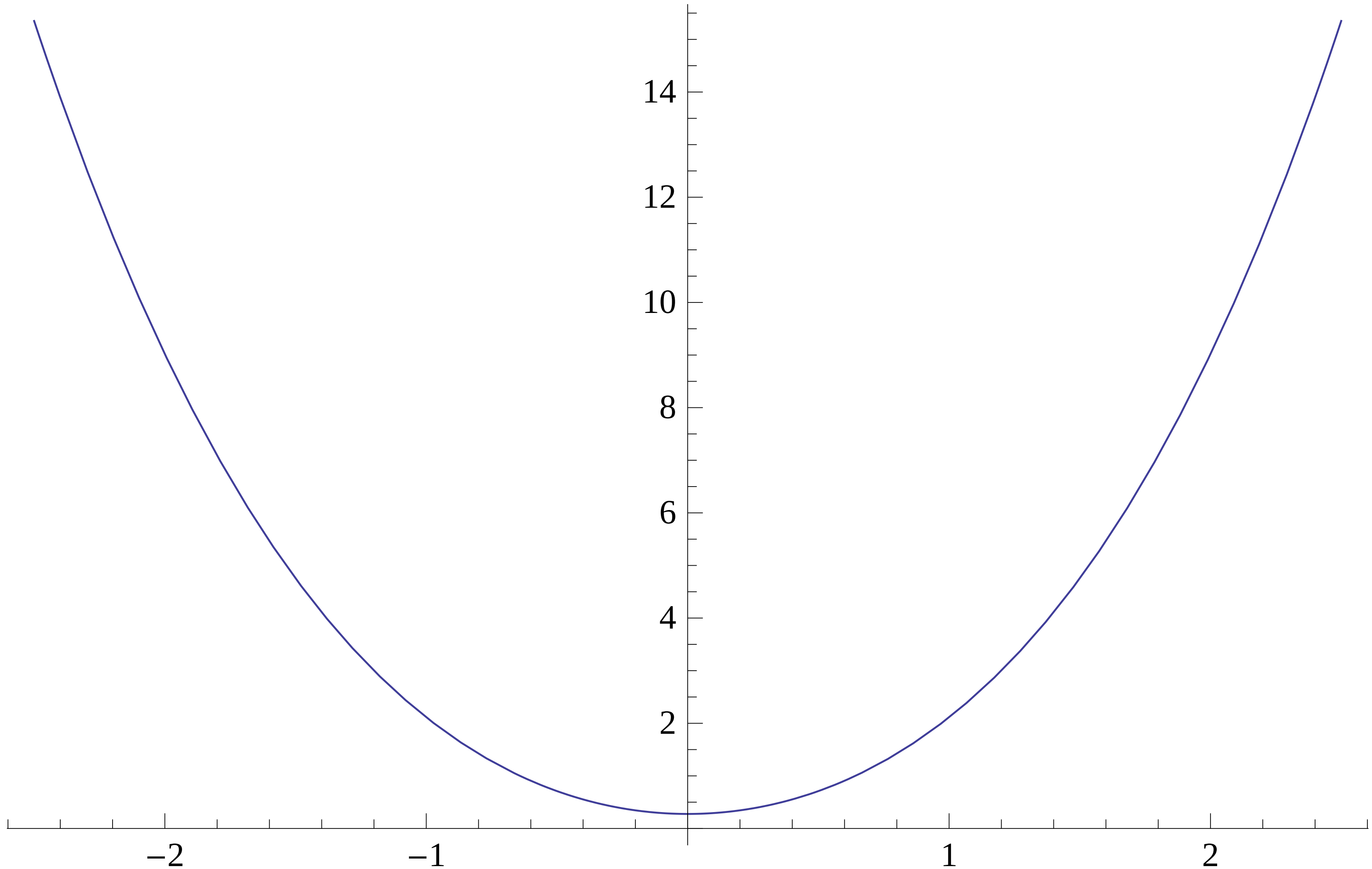
$$u(t,t^2) = \lim_{x \nearrow t^2} u(t,x) = 1, \qquad \lim_{x \to -\infty} u(t,x) = 0.$$

2. Graphical evidence for log-concavity of f



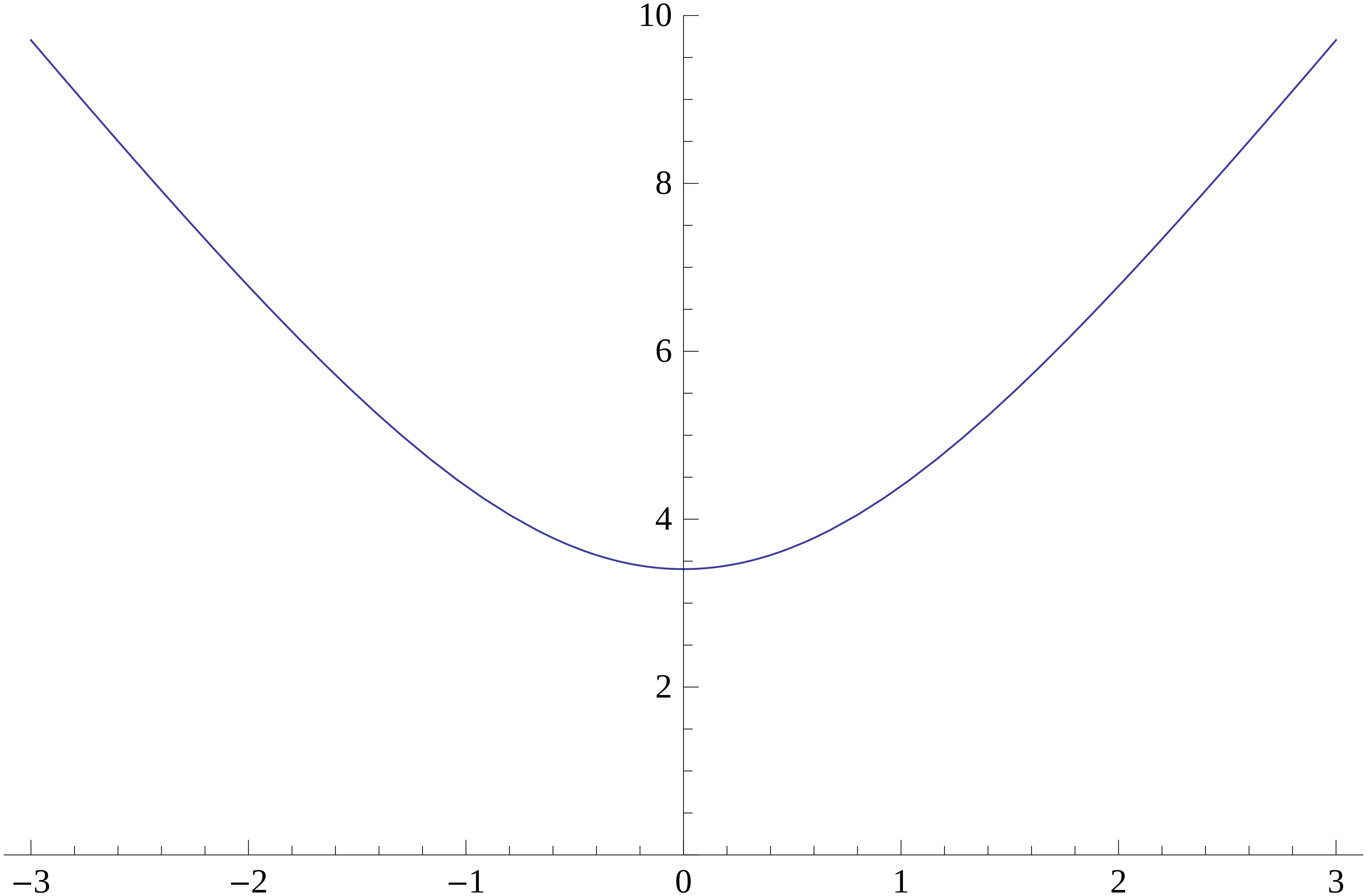
The density f_Z

2. Graphical evidence for log-concavity of f



$-\log f_Z$

2. Graphical evidence for log-concavity of f



$(-\log f_Z)''$

2. Graphical evidence for log-concavity of f

Recall that f is log-concave if and only if, for all $x, y \in \mathbb{R}$, $\lambda \in [0,1]$,

$$
\begin{aligned}
& \log f(\lambda x + (1-\lambda)y) \ge \lambda \log f(x) + (1-\lambda) f(y) \\
\text{iff} \quad & f(\lambda x + (1-\lambda)y) \ge f(x)^{\lambda} f(y)^{1-\lambda} \\
\text{iff} \quad & f\left(\frac{x+y}{2}\right) \ge f(x)^{1/2} f(y)^{1/2} \qquad (Sierpinski, Jensen) \\
\text{iff} \quad & f\left(\frac{x+y}{2}\right)^2 \ge f(x) f(y) \\
\text{iff} \quad & f\left(\frac{x+y}{2}\right)^2 - f(x) f(y) \ge 0 \\
\text{iff} \quad & D_f(t,s) \equiv f(t)^2 - f(t+s) f(t-s) \ge 0 \quad \text{for all} \quad t, s.
\end{aligned}
$$

The plots on the following pages show the function $D_f(t,s)$ for $f = \phi$, the standard normal density and for $f = f_Z$, Chernoff's density.

2. Graphical evidence for log-concavity of f

Plot deleted because of size!

The function $D_f(t,s)$, $f = \phi$, standard normal density

2. Graphical evidence for log-concavity of f

Plot deleted because of size!

The function $D_f(t,s)$, $f = f_Z$, Chernoff's density

3. Groeneboom's formula

Let $W(t)$ be two-sided Brownian motion starting from 0. Groeneboom (1985, 1989) analyzed the process

$$V_c(a) \equiv \sup\{t \in \mathbb{R} : \ W(t) - c(t-a)^2 \ \text{ is maximal}\},$$

and along the way showed that $Z_c \equiv V_c(0)$ has density

$$f_{Z_c}(t) = \frac{1}{2} g_c(t) g_c(-t)$$

where g_c has Fourier transform given by

$$\widehat{g}_c(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda s} g_c(s) ds = \frac{2^{1/3} c^{-1/3}}{Ai(i(2c^2)^{-1/3}\lambda)};$$

here $Ai(z) \equiv U(z)$ is the Airy function satisfying $U''(z) - zU(z) = 0$. Thus Chernoff's $Z \equiv Z_1, \ \ g \equiv g_1$, and

$$g_c(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-iux} 2^{1/3} c^{-1/3}}{Ai(i(2c^2)^{-1/3} u)} du.$$

3. Groeneboom's formula

Asymptotics of g_c and f_Z; Groeneboom (1989):

as t, $z \to \infty$,

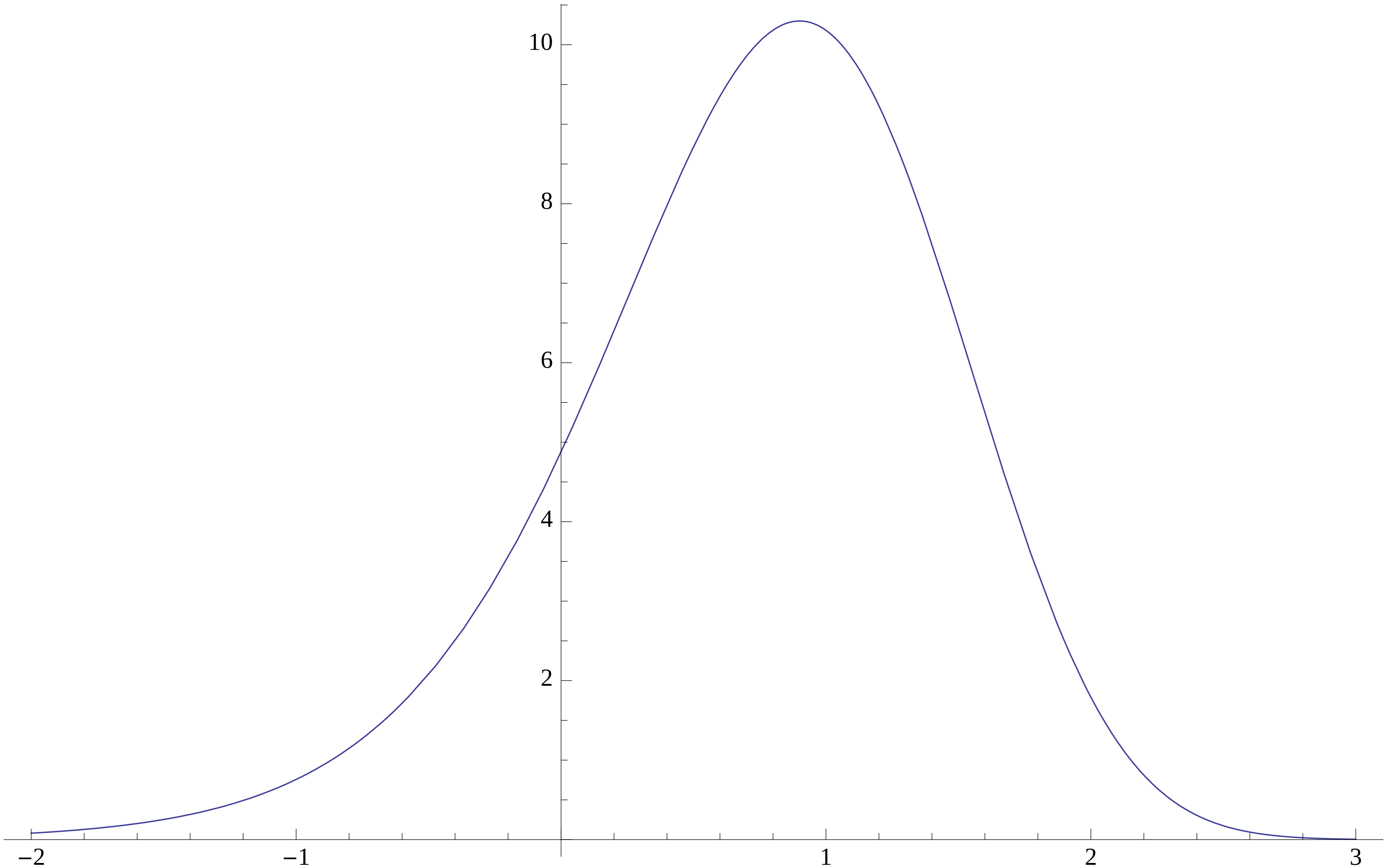
$$g_c(t) \sim 4ct\text{exp}\left(-\frac{2}{3}c^2t^3\right),$$

$$g_c(-t) \sim (4c)^{1/3}\text{exp}(-(2c^2)^{1/3}(-a_1)t)/Ai'(a_1), \quad \text{and}$$

$$f_Z(z) \sim \frac{(1/2)(4c)^{4/3}}{Ai'(a_1)}|z|\text{exp}\left(-\frac{2}{3}c^2|z|^3 + (2c^2)^{1/3}a_1|z|\right)$$

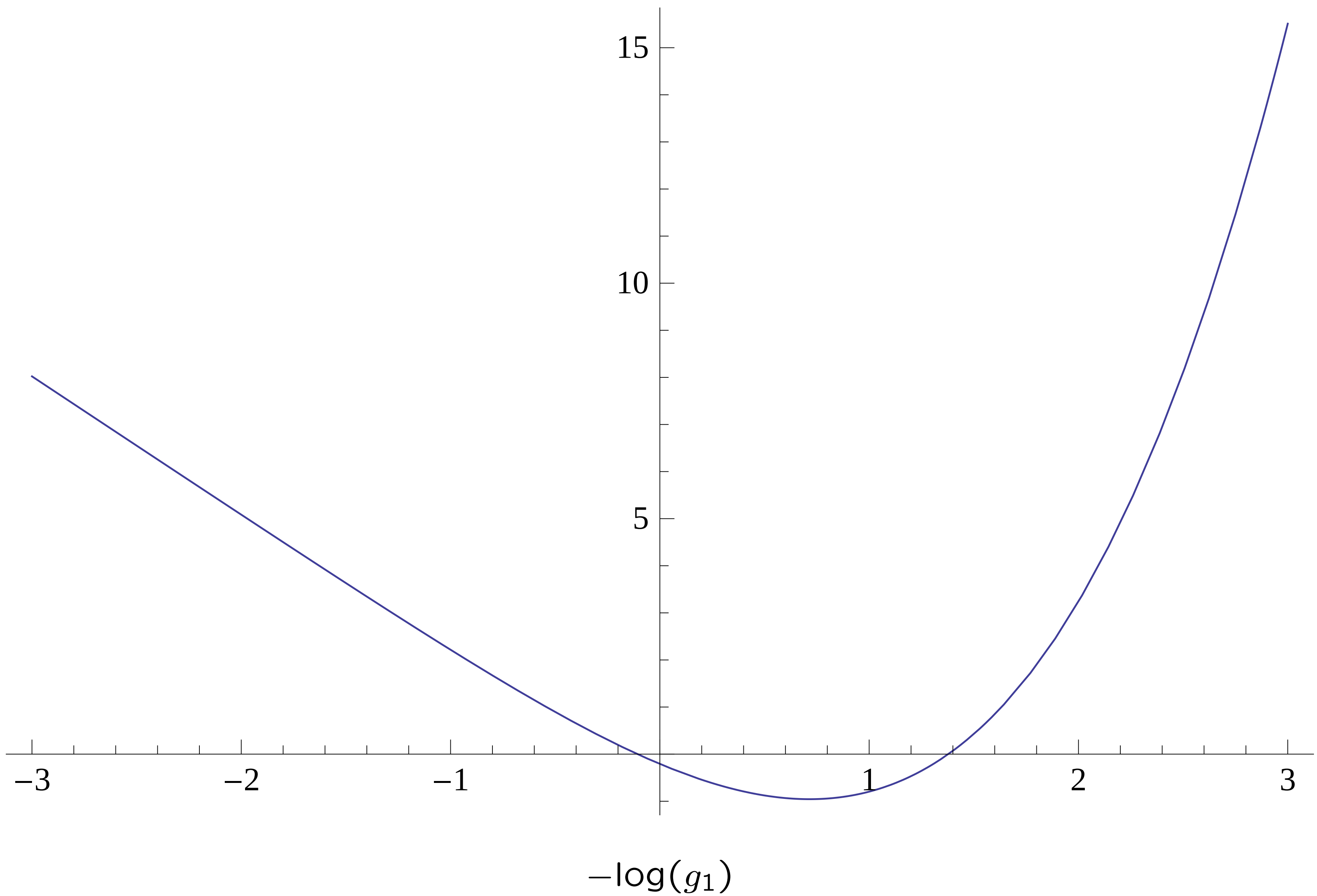
where $a_1 \approx -2.3381$ is the largest zero of the Airy function Ai and where $Ai'(a_1) \approx 0.7022$.

3. Groeneboom's formula ...

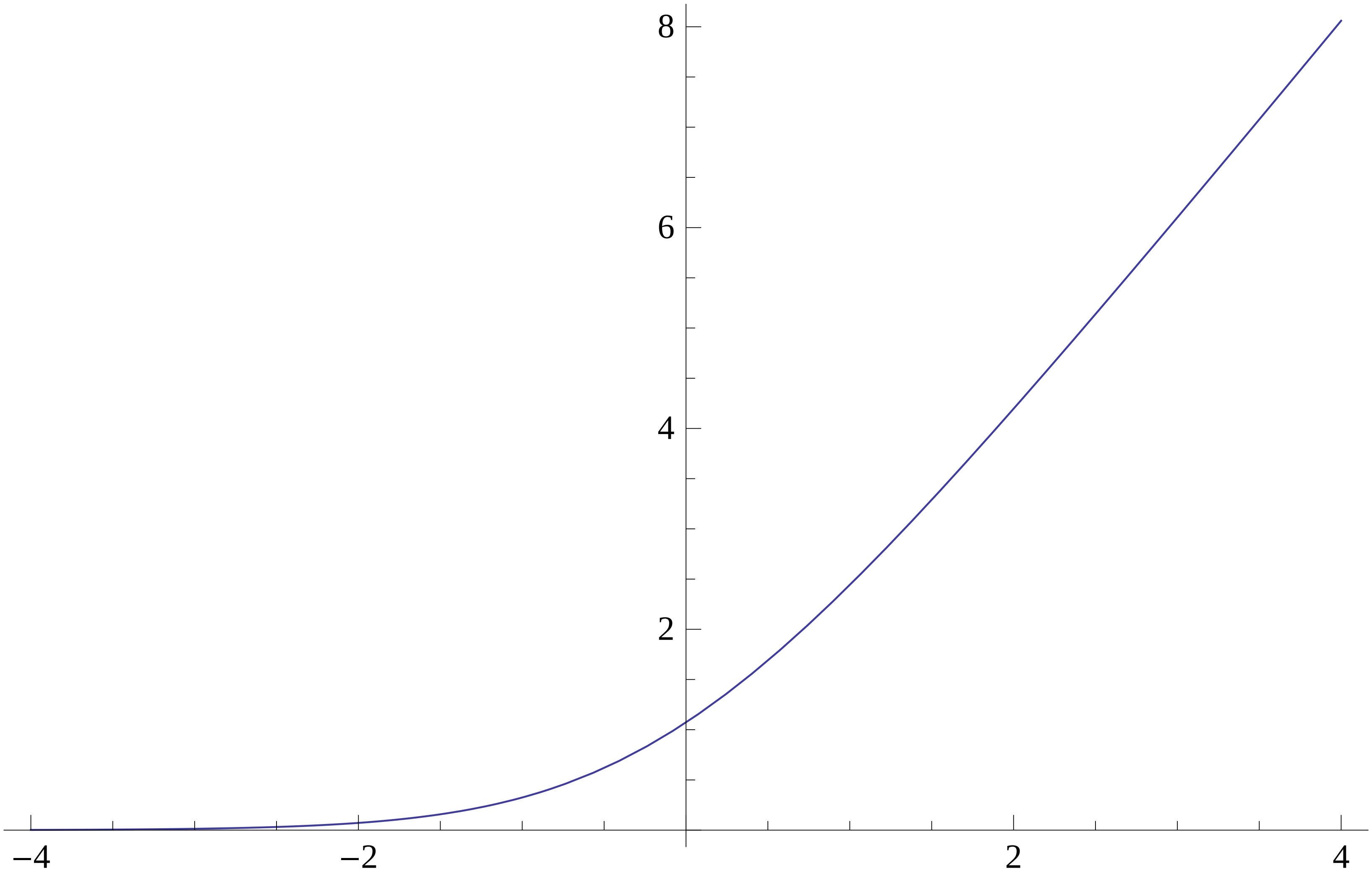


The function g_1

3. Groeneboom's formula



3. Groeneboom's formula



Second derivative, $(-\log(g_1))''(x)$

4. Chernoff's density f is in PF_2, i.e. log-concave

From Chernoff (1964)

$$f(z) = \frac{1}{2} g(z) g(-z) \equiv \frac{1}{2} g_1(z) g_1(-z).$$

From Groeneboom (1885, 1989), Daniels and Skyrme (1985)

$$\widehat{g}_c(\lambda) = \int_{-\infty}^{\infty} e^{i\lambda s} g_c(s) ds = \frac{2^{1/3} c^{-1/3}}{Ai(i(2c^2)^{-1/3}\lambda)}.$$

4. Chernoff's density f is in PF_2, i.e. log-concave

Theorem. Schoenberg (1951)
A necessary and sufficient condition that a (density) function $g(x)$, $-\infty < x < \infty$, be a PF_∞ (density) function is that the reciprocal of its bilateral Laplace transform (i.e. Fourier) be an entire function of the form

$$\psi(s) \equiv \frac{1}{\widehat{g}(s)} = Ce^{-\gamma s^2+\delta s} s^k \prod_{j=1}^{\infty} (1 + b_j s)\exp(-b_j s)$$

where $C > 0$, $\gamma \geq 0$, $\delta \in \mathbb{R}$, $k \in \{0, 1, 2, \ldots\}$, $b_j \in \mathbb{R}$, $\sum_{j=1}^{\infty} b_j^2 < \infty$. (For the subclass of densities, the if and only if statement holds for $1/\widehat{g}$ of this form with $k = 0$ and $\psi(0) = C = 1$.)

See also Karlin (1968), Theorem 7.3.2, page 345.

4. Chernoff's density f is log-concave

Connection? The following Hadamard product formula for $Ai(z)$ is due to Merkes and Salmassi (1997):

Proposition. (Merkes and Salmassi) Let $\{-a_k\}$ be the zeros of the Airy function Ai (so that $a_k > 0$ for each k). The Hadamard representation of Ai is given by

$$Ai(z) = Ai(0)e^{-\nu z} \prod_{j=1}^{\infty} (1 + z/a_j)\exp(-z/a_j)$$

where $Ai(0) = c_1 = 1/(3^{2/3}\Gamma(2/3)) \approx 0.35503$, and

$$\nu = -Ai'(0)/Ai(0) = 3^{1/3}\Gamma(2/3)/\Gamma(1/3) \approx .729011\ldots.$$

Since it is well-known that $a_j \sim ((3/8)\pi(4j-1))^{2/3}$ as $k \to \infty$, we have

$$\sum_1^\infty b_j^2 = \sum_1^\infty (1/a_j^2) < \infty, \qquad \sum_1^\infty b_j = \sum_1^\infty (1/a_j) = \infty.$$

4. Chernoff's density f is Polya frequency ...

Conclusion: Schoenberg's characterization holds for $g = g_1$ with $\gamma = 0$, $k = 0$, $\delta = -(2c^2)^{1/3}\nu = (2c^2)^{1/3} Ai'(0)/Ai(0)$, and $b_j = 1/a_j$, $j \geq 1$.

Theorem. The functions $x \mapsto g_c(x)$ are in $PF_\infty \subset PF_2$ for every $c > 0$. (In particular, they are log-concave.) Thus $f(x) = (1/2)g(x)g(-x)$ is log-concave:

$$(-\log f)''(x) = (-\log g)''(x) + (-\log g)''(-x) \geq 0 + 0 = 0.$$

5. Is Chernoff's density f "Strongly log-concave"?

If $f \in C^2$, the following are equivalent:

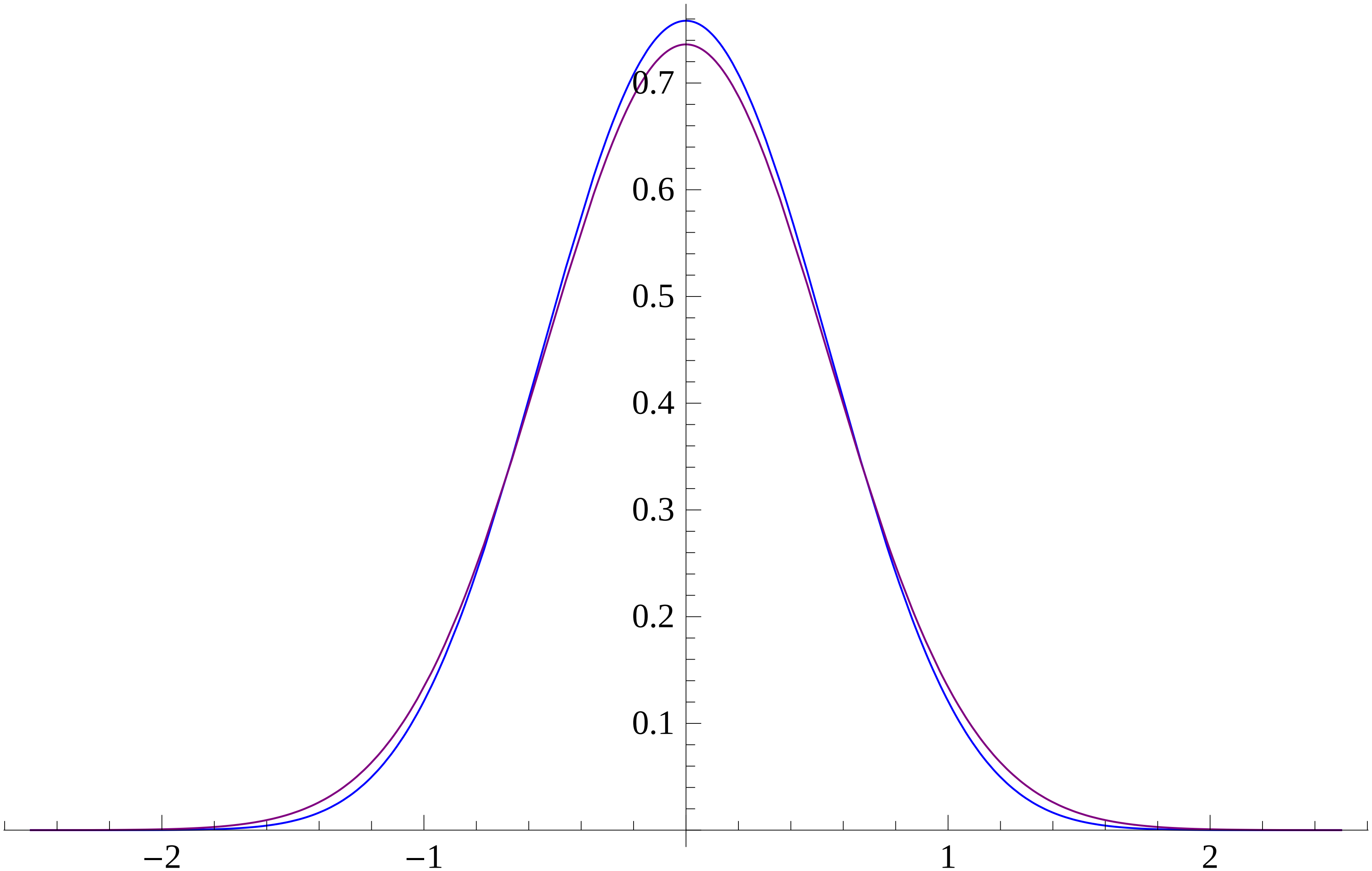
- $(-\log f)''(x) \geq 1/\sigma^2 > 0$.
- $f(x) = \rho(x)\phi_\sigma(x) = \rho(x)\phi(x/\sigma)/\sigma$ where ρ is log$-$concave.

Definition:

- f on $\mathbb{R}$ is strongly log-concave if and only if $f = \rho\phi_\sigma$ for some $\sigma > 0$ and ρ log-concave.
- f on $\mathbb{R}^d$ is strongly log-concave if and only if $f = h\phi_\Sigma$ for some $\Sigma > 0$ and h log-concave where ϕ_Σ is the density of $N_d(0, \Sigma)$.

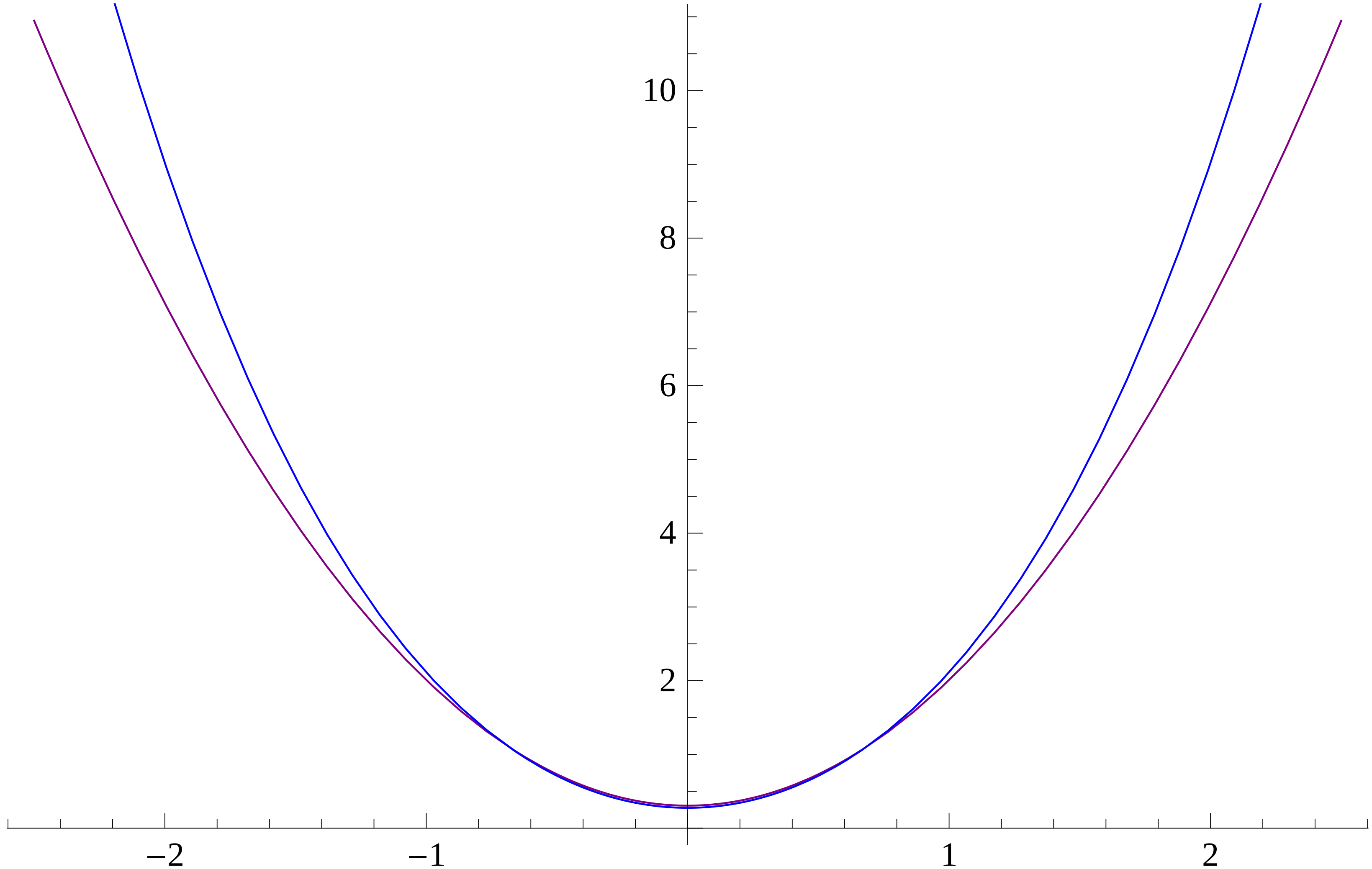
Alternative possible terminologies:
(a) Super Gaussian?
(b) log-concave perturbation of Gaussian?
(c) Schoenberg PF_2?
(d) Gaussian product log-concave?

5. Is Chernoff's density f "Strongly log-concave"?



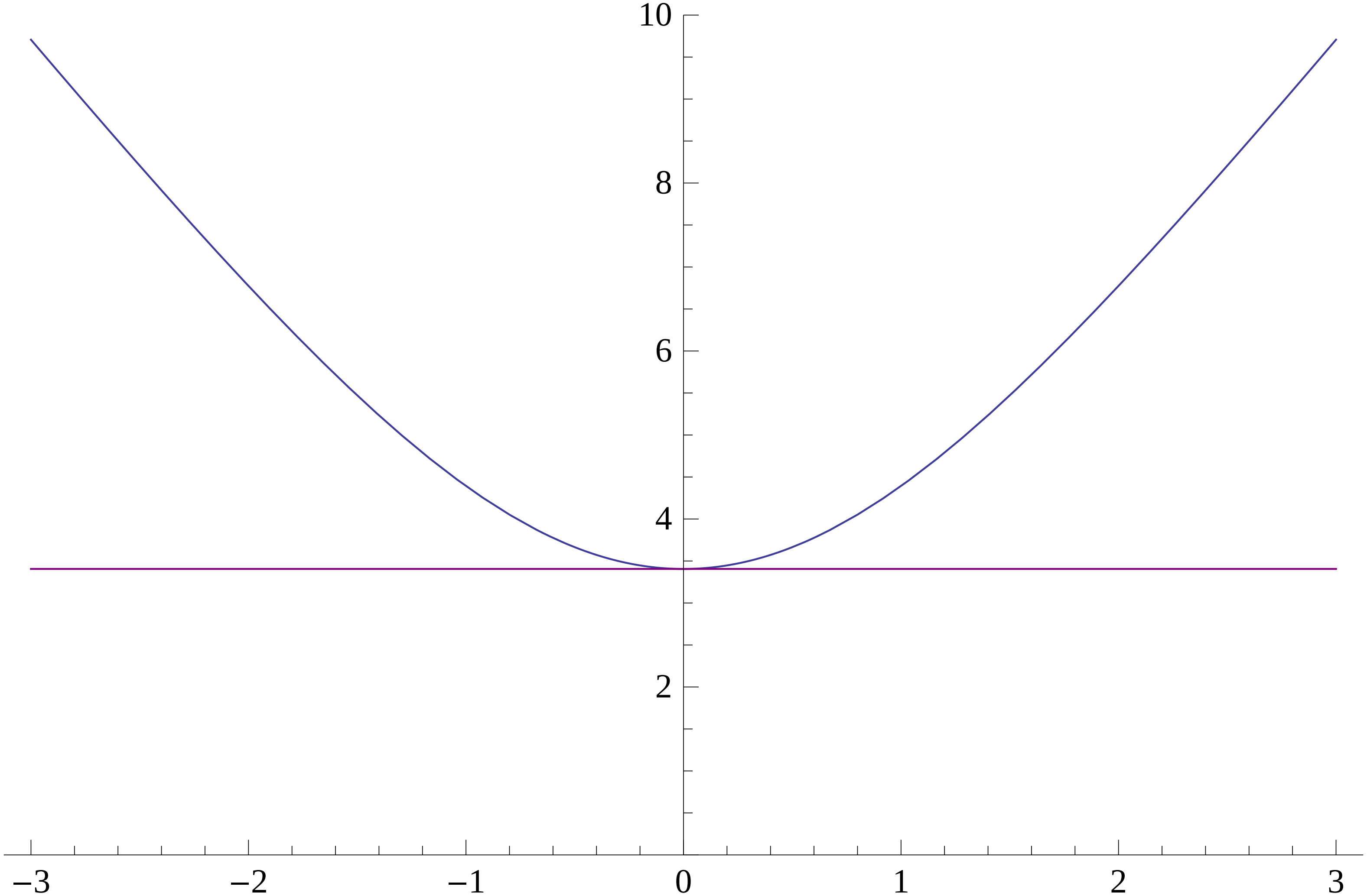
Chernoff's f (blue) and $N(0, \sigma^2)$, $\sigma = .541912...$ (purple)

5. Is Chernoff's density f "Strongly log-concave"?



Log of f (Chernoff) (blue) and log of $N(0, \sigma^2)$, $\sigma = .541912...$, purple

5. Is Chernoff's density f "Strongly log-concave"?



$(-\log f)''(x)$ (Chernoff) (blue) and $(-\log \phi_\sigma)''(x)$ for $\sigma = .541912...$, purple

Back to Schoenberg's representation:
recall that g is a PF_∞ (density) if and only if $X \sim g$ has

$$\begin{aligned} 1/Ee^{sX} &\equiv \frac{1}{\widehat{g}(s)} \equiv \psi(s) \\ &= Ce^{-\gamma s^2+\delta s} s^k \prod_{j=1}^{\infty} (1+b_j s)\text{exp}(-b_j s) \end{aligned}$$

For a density function $C = 1$ and $k = 0$; then

- $e^{\gamma s^2}$ is a "Gaussian component;
- $e^{-\delta s}$ is a "shift component" (deterministic);
- $\prod_{j=1}^{\infty} \text{exp}(b_j s)/(1+b_j s)$ is a "centered independent exponentials" component; i.e. with $X_i \sim \text{exp}(1/b_i)$ independent and

$$Y \equiv -\sum_{i=1}^{\infty} (X_i - b_i)$$

we have

$$\begin{aligned} Ee^{sY} &= \prod_{i=1}^{\infty} Ee^{-s(X_i - b_i)} = \prod_{i=1}^{\infty} \frac{e^{b_i s}}{1 + b_i s} \\ &= \frac{1}{\prod_{i=1}^{\infty}(1 + b_i s)e^{-b_i s}}. \end{aligned}$$

Note that

$$Y_m \equiv -\sum_{i=1}^{m}(X_i - b_i) \to_{a.s.} -\sum_{i=1}^{\infty}(X_i - b_i) \equiv Y$$

if $\sum_{i=1}^{\infty} b_i^2 = \sum_{i=1}^{\infty} Var(X_i) < \infty$. For Chernoff's distribution, with $\{-a_j\}$ the zeros of the Airy function Ai,

$$b_j = \frac{1}{(2c^2)^{1/3} a_j}, \qquad a_j \sim ((3/8)\pi(4k-1))^{2/3}, \qquad \text{so}$$

$$\sum_{i=1}^{\infty} b_i = \sum_{i=1}^{\infty} \frac{1}{a_j} = \infty, \qquad \sum_{i=1}^{\infty} b_i^2 = \sum_{i=1}^{\infty} \frac{1}{a_j^2} < \infty.$$

Recall that Chernoff's density $f \equiv f_1$ has the form

$$f(z) = \frac{1}{2} g(z) g(-z).$$

Let h_∞ denote the density of Y.
The preceding yields $h = g / \int g(y) dy$!

Then

$$\begin{aligned}
-\log f(z) &= -\log g(z) + (-\log g(-z)), \\
(-\log f)''(z) &= (-\log g)''(z) + (-\log g)''(-z) \\
&= (-\log h)''(z) + (-\log h)''(-z) \\
&\equiv v(z) + v(-z) \\
&\geq \begin{cases} 2v(0), & \text{if } v \text{ is convex,} \\ v(0), & \text{if } v(z) \geq v(0) \text{ for } z \geq 0. \end{cases} \\
&\geq v(0) > 0 \quad \text{for all } z
\end{aligned}$$

So, strong-log concavity of Chernoff's density holds if

$$v = (-\log h)'' \quad \text{is convex}$$

or (with a weaker constant) if

$$v(z) \geq v(0) \quad \text{for all} z \geq 0.$$

(Enough to show these for h_m the density of Y_m?!)

6. Relations with other classes: PF_∞ and HM_∞

- Other elements $g \in PF_\infty$ and corresponding $f(x) = g(x)g(-x)/C$? $Y = -\sum_{j=1}^\infty (X_j - b_j)$ with $b_j \sim (C/j)^\gamma$, $\gamma \in (1/2, 1)$?
- Connections to $\log(HM_1)$ and $\log(HM_\infty)$? For a density f on $(0, \infty)$ set

$$H(w) = f(uv)f(u/v) \quad \text{where} \quad w = v + 1/v.$$

Definitions: (Bondesson)

- $f \in HM_\infty$ if $H(w)$ is a completely monotone function of w for each real u.
- $f \in HM_k$ if $(-1)^j H^{(j)}(w) \geq 0$, $j \in \{0, \ldots, k-1\}$ and $(-1)^{k-1}H^{(k-1)}$ is right-continuous and decreasing.

Let $\log HM_k \equiv \{x \mapsto e^x f(e^x) : \quad f \in HM_k\}$ for $k \geq 1$.

Theorem 1. (Bondesson) log$HM_1 = PF_2 =$ log-concave

Proposition 1. (Bondesson) log$HM_\infty \not\subset PF_\infty$,
and $PF_\infty \not\subset$ logHM_∞.

7. Summary; conjectures and open problems.

Summary: Chernoff's density is log-concave!

- Conjecture 1: Chernoff's density f is strongly log-concave for $\sigma \geq \sigma_0 \equiv \{(-\log f)''(0)\}^{-1/2} = .541912\ldots > \sqrt{Var(Z)} = 0.513381\ldots$.
- Conjecture 2: In the monotone regression problem with Gaussian errors, all the one-dimensional marginal distributions of $\underline{\widehat{\mu}}_n$ are log-concave if $\underline{\mu} \in K_n$.
- Problem 1: Characterize the subset of $\log(HM_\infty)$ corresponding to strongly log-concave densities.
- Problem 2: Prove an analogue of the Berry-Esseen theorem corresponding to Brunk's theorem for monotone regression.
- Problem 3. Prove an analogue of the Linnik-Brown-Barron entropy CLT corresponding to Brunk's theorem.

7. Summary; conjectures and open problems.

- Problem 4: Extend the development here to the case when the convex cone K_n is replaced by the convex cone in $\mathbb{R}^n$ corresponding to convex regression functions: see e.g. Groeneboom, Jongbloed, and Wellner (2001).

8. Selected references.

- Brunk, H. D. (1970). Estimation of isotonic regression. In *Nonparametric Techniques in Statistical Inference (Proc. Sympos., Indiana Univ., Bloomington, Ind., 1969).* Cambridge Univ. Press, London, 177 - 197.

- Chernoff, H. (1964). Estimation of the mode. *Ann. Inst. Statist. Math.* **16**, 31-41.

- Daniels, H. E. and Skyrme, T. H. R. (1985). The maximum of a random walk whose mean path has a maximum. *Adv. in Appl. Probab.* **17**, 85-99.

- Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Theory Related Fields* **81**, 79-109.

- Karlin, S. (1968). *Total Positivity.* Vol. I. Stanford University Press, Stanford, Calif.

- Merkes, E. P. and Salmassi M. (1997). On univalence of certain infinite products. *Complex Variables Theory Appl.* **33**, 207-215.

- Schoenberg, I. J. (1951). On Pólya frequency functions. I. The totally positive functions and their Laplace transforms. *J. Analyse Math.* **1**, 331-374.

Merci!