A Kiefer - Wolfowitz theorem for Convex Densities

Jon A. Wellner

University of Washington, Statistics

$$\left.\frac{t_0)(1-F(t_0))f(t_0)}{g(t_0)}\right)^{1/3} \qquad n^{1/3}\left\{\hat{F}_n(t_0)-F(t_0)\right\}$$

$$/\mathrm{Ai}(c+is)| \sim 2\sqrt{\pi}|s|^{1/4}\exp\left\{-\frac{\sqrt{2}}{3}|s|^{3/2}+c\right.$$

entific
nators

- E.A. Cator, Delft
- G. Jongbloed, Amsterdam
- C. Kraaikamp, Delft
- H.P. Lopuhaä, Delft
- J.A. Wellner, Seattle

akers

- David Aldous
- Lucien Birgé
- Persi Diaconis
- Lutz Dümbgen
- Pablo Ferrari
- Sara van de Geer
- Richard Gill
- Piet Groeneboom
- Jian Huang
- Niels Keiding
- Steve Lalley
- Enno Mammen
- David Pollard
- Timo Seppäläinen
- Aad van der Vaart
- Jon Wellner

- joint work with Fadoua Balabdaoui,
  University of Paris, Dauphine
- Talk at Asymptotics: particles, processes and inverse problems,
  Leiden, The Netherlands, July 14, 2006
- *Email: jaw@stat.washington.edu*
  http: //www.stat.washington.edu/jaw/jaw.research.html

Outline

- Introduction: the Kiefer-Wolfowitz theorem and Piet

Outline

- Introduction: the Kiefer-Wolfowitz theorem and Piet
- The Kiefer-Wolfowitz theorem, monotone densities

Outline

- Introduction: the Kiefer-Wolfowitz theorem and Piet
- The Kiefer-Wolfowitz theorem, monotone densities
- Approximation theory: bounds for interpolation errors

Outline

- Introduction: the Kiefer-Wolfowitz theorem and Piet
- The Kiefer-Wolfowitz theorem, monotone densities
- Approximation theory: bounds for interpolation errors
- A Kiefer-Wolfowitz theorem, convex densities

Introduction: the Kiefer-Wolfowitz theorem

- Let f be a monotone density on $[0, \infty)$.

Introduction: the Kiefer-Wolfowitz theorem

- Let f be a monotone density on $[0, \infty)$.
- Thus $F(x) = \int_0^x f(y)dy$ is a concave distribution function on $[0, \infty)$.

Introduction: the Kiefer-Wolfowitz theorem

- Let f be a monotone density on $[0, \infty)$.
- Thus $F(x) = \int_0^x f(y)dy$ is a concave distribution function on $[0, \infty)$.
- Let $\mathbb{F}_n(x) = n^{-1} \sum_{i=1}^n 1\{X_i \leq x\}$ be the empirical d.f. of $X_1, \ldots, X_n$ i.i.d. with density f.

Introduction: the Kiefer-Wolfowitz theorem

- Let f be a monotone density on $[0, \infty)$.
- Thus $F(x) = \int_0^x f(y)dy$ is a concave distribution function on $[0, \infty)$.
- Let $\mathbb{F}_n(x) = n^{-1} \sum_{i=1}^n 1\{X_i \leq x\}$ be the empirical d.f. of $X_1, \ldots, X_n$ i.i.d. with density f.
- Let $\widehat{f}_n$ be the Grenander estimator of f;
  i.e. $\widehat{f}_n(x)$ is the slope at x of the least concave majorant $\widehat{F}_n$ of the empirical d.f. $\mathbb{F}_n$

- **Theorem 1. (Kiefer-Wolfowitz, 1976).** Suppose that:

  Then

$$\begin{aligned} \|\widehat{F}_n - \mathbb{F}_n\| &\equiv \sup_{t<\alpha_1(F)} |\widehat{F}_n(t) - \mathbb{F}_n(t)| \\ &= O((n^{-1}\log n)^{2/3}) \ \text{ almost surely.} \end{aligned}$$

- **Theorem 1. (Kiefer-Wolfowitz, 1976).** Suppose that:
  - $\alpha_1(F) = \inf\{t : F(t) = 1\} < \infty$.

Then

$$\begin{aligned} \|\widehat{F}_n - \mathbb{F}_n\| &\equiv \sup_{t<\alpha_1(F)} |\widehat{F}_n(t) - \mathbb{F}_n(t)| \\ &= O((n^{-1}\log n)^{2/3}) \text{ almost surely.} \end{aligned}$$

- **Theorem 1. (Kiefer-Wolfowitz, 1976).** Suppose that:
  - $\alpha_1(F) = \inf\{t : F(t) = 1\} < \infty$.
  - f' continuous on $[0, \alpha_1(F)]$.

Then

$$\begin{aligned} \|\widehat{F}_n - \mathbb{F}_n\| &\equiv \sup_{t<\alpha_1(F)} |\widehat{F}_n(t) - \mathbb{F}_n(t)| \\ &= O((n^{-1}\log n)^{2/3}) \ \text{almost surely.} \end{aligned}$$

- **Theorem 1. (Kiefer-Wolfowitz, 1976).** Suppose that:
  - $\alpha_1(F) = \inf\{t : F(t) = 1\} < \infty$.
  - f' continuous on $[0, \alpha_1(F)]$.
  - $\beta_1(F) \equiv \inf_{0<t<\alpha_1(F)}(-f'(t)/f^2(t)) > 0$.

  Then

$$
\begin{aligned}
\|\widehat{F}_n - \mathbb{F}_n\| &\equiv \sup_{t<\alpha_1(F)} |\widehat{F}_n(t) - \mathbb{F}_n(t)| \\
&= O((n^{-1}\log n)^{2/3}) \ \text{ almost surely.}
\end{aligned}
$$

- **Theorem 1. (Kiefer-Wolfowitz, 1976).** Suppose that:
  - $\alpha_1(F) = \inf\{t : F(t) = 1\} < \infty$.
  - f' continuous on $[0, \alpha_1(F)]$.
  - $\beta_1(F) \equiv \inf_{0<t<\alpha_1(F)}(-f'(t)/f^2(t)) > 0$.
  - $\gamma_1(F) \equiv \sup_{0<t<\alpha_1(F)}(-f'(t))/\inf_{0<t<\alpha_1(F)} f^2(t) < \infty$.

  Then

$$
\begin{aligned}
\|\widehat{F}_n - \mathbb{F}_n\| &\equiv \sup_{t<\alpha_1(F)} |\widehat{F}_n(t) - \mathbb{F}_n(t)| \\
&= O((n^{-1}\log n)^{2/3}) \text{ almost surely.}
\end{aligned}
$$

- **Theorem 2. (Wang (2000); Kulikov and Lopuhaa (2006)).** If $f(t_0) > 0$, $f'(t_0) < 0$, and f' is continuous in a neighborhood of t_0, then

$$n^{2/3}(\widehat{F}_n(t_0 + n^{-1/3}t) - \mathbb{F}_n(t_0 + n^{-1/3}t))$$
$$\Rightarrow \left(\frac{2f^2(t_0)}{-f'(t_0)}\right)^{1/3} \{\mathbb{C}(at) - (W(at) - a^2t^2)\}$$

in $D[-K, K]$ for each fixed $K > 0$ where W is two-sided standard Brownian motion starting from 0, $\mathbb{C}$ is the least concave majorant of $W(t) - t^2$, and $a \equiv ([f'(t_0)]^2/(4f(t_0)))^{1/3}$.

- Goal in these two talks: prove results similar to theorems 1 and 2 in the case when f is decreasing and convex.

- Goal in these two talks: prove results similar to theorems 1 and 2 in the case when f is decreasing and convex.
- Unfortunately, there is not yet an analogue of Marshall's lemma for the MLEs $\widehat{f}$ and $\widehat{F}_n$ in this case.

- Goal in these two talks: prove results similar to theorems 1 and 2 in the case when f is decreasing and convex.
- Unfortunately, there is not yet an analogue of Marshall's lemma for the MLEs $\widehat{f}$ and $\widehat{F}_n$ in this case.
- Good news: Dümbgen, Rufibach, Wellner have an analogue of Marshall's lemma for the Least Squares Estimators $\widetilde{f}_n$ and $\widetilde{F}_n$.

- Goal in these two talks: prove results similar to theorems 1 and 2 in the case when f is decreasing and convex.
- Unfortunately, there is not yet an analogue of Marshall's lemma for the MLEs $\widehat{f}$ and $\widehat{F}_n$ in this case.
- Good news: Dümbgen, Rufibach, Wellner have an analogue of Marshall's lemma for the Least Squares Estimators $\widetilde{f}_n$ and $\widetilde{F}_n$.
- Thus prove similar results for $\widetilde{F}_n$.

- Groeneboom (1989), page 104 says: “ We finally want to note that the process $\{V(a) : \; a \in \mathbb{R}\}$ not only describes the limiting global behavior of the Grenander maximum likelihood estimator of a (smooth and strictly decreasing) density (see Groeneboom (1985)), but also describes the limiting behavior of certain “isotonic” estimators of distribution functions and hazard functions. In particular, by using the properties of this process, a simple proof of results in Kiefer and Wolfowitz (1976) can be given, which at the same time clarifies the connection between these results (on the estimation of concave distribution functions) and results on the estimation of a monotone density.”
  To the best of my knowledge Piet has not yet written the “simple proof” he promised in 1989. We encourage him to do so soon!

- Theme 1: simplify proof of original K-W theorem using:

- Theme 1: simplify proof of original K-W theorem using:
  - Results from interpolation theory: de Boor, Hall and Meyer, ... for the “broken line” interpolation operator I_2.

- Theme 1: simplify proof of original K-W theorem using:
  - Results from interpolation theory: de Boor, Hall and Meyer, ... for the “broken line” interpolation operator I_2.
  - Results from empirical process theory involving the oscillation modulus of the uniform empirical process: Stute (1982), Mason, Shorack, Wellner (1983); Shorack and Wellner (1986).

- Theme 1: simplify proof of original K-W theorem using:
  - Results from interpolation theory: de Boor, Hall and Meyer, ... for the “broken line” interpolation operator I_2.
  - Results from empirical process theory involving the oscillation modulus of the uniform empirical process: Stute (1982), Mason, Shorack, Wellner (1983); Shorack and Wellner (1986).
- Theme 2: proof of analogue of the K-W theorem for $\widetilde{F}_n$ corresponding to f convex using

- Theme 1: simplify proof of original K-W theorem using:
  - Results from interpolation theory: de Boor, Hall and Meyer, ... for the "broken line" interpolation operator I_2.
  - Results from empirical process theory involving the oscillation modulus of the uniform empirical process: Stute (1982), Mason, Shorack, Wellner (1983); Shorack and Wellner (1986).
- Theme 2: proof of analogue of the K-W theorem for $\widetilde{F}_n$ corresponding to f convex using
  - Results from interpolation theory for the complete (cubic) spline interpolation operator I_4; c.f. de Boor (2001).

- Theme 1: simplify proof of original K-W theorem using:
  - Results from interpolation theory: de Boor, Hall and Meyer, ... for the “broken line” interpolation operator I_2.
  - Results from empirical process theory involving the oscillation modulus of the uniform empirical process: Stute (1982), Mason, Shorack, Wellner (1983); Shorack and Wellner (1986).
- Theme 2: proof of analogue of the K-W theorem for $\widetilde{F}_n$ corresponding to f convex using
  - Results from interpolation theory for the complete (cubic) spline interpolation operator I_4; c.f. de Boor (2001).
  - Oscillation theory of the uniform empirical process.

2. Proof of the Kiefer-Wolfowitz (1976) theorem

- **Lemma. (Marshall)** Let $\widehat{F}_n$ be the least concave majorant of $\mathbb{F}_n$, and let h be a concave function on $[0, \infty)$. Then

$$\|\widehat{F}_n - h\| \leq \|\mathbb{F}_n - h\|.$$

- Outline of the Proof of the KW theorem:

- Outline of the Proof of the KW theorem:
    - Step 1: Marshall’s lemma.

- Outline of the Proof of the KW theorem:
  - Step 1: Marshall's lemma.
  - Step 2: Construct $\mathbb{L}_n$ (convenient) satisfying

$$P_F(A_n) \equiv P_F(\mathbb{L}_n \text{ concave on } [0,\infty)) \geq 1 - n^{-2}$$

for all n sufficiently large.

- Outline of the Proof of the KW theorem:
  - Step 1: Marshall's lemma.
  - Step 2: Construct $\mathbb{L}_n$ (convenient) satisfying

$$P_F(A_n) \equiv P_F(\mathbb{L}_n \text{ concave on } [0,\infty)) \geq 1 - n^{-2}$$

    for all n sufficiently large.
  - Step 3. On A_n note that by Marshall's lemma we have

$$\begin{aligned}
\|\widehat{F}_n - \mathbb{F}_n\| &\leq \|\widehat{F}_n - \mathbb{L}_n + \mathbb{L}_n - \mathbb{F}_n\| \\
&\leq \|\widehat{F}_n - \mathbb{L}_n\| + \|\mathbb{L}_n - \mathbb{F}_n\| \\
&\leq \|\mathbb{F}_n - \mathbb{L}_n\| + \|\mathbb{L}_n - \mathbb{F}_n\| = 2\|\mathbb{F}_n - \mathbb{L}_n\| \\
&\leq 2\{\|\mathbb{F}_n - F - (\mathbb{L}_n - L_n)\| + \|F - L_n\|\} \\
&\equiv 2D_n + 2E_n.
\end{aligned}$$

- Outline of the Proof of the KW theorem:
  - Step 1: Marshall's lemma.
  - Step 2: Construct $\mathbb{L}_n$ (convenient) satisfying

$$P_F(A_n) \equiv P_F(\mathbb{L}_n \text{ concave on } [0,\infty)) \geq 1 - n^{-2}$$

    for all n sufficiently large.
  - Step 3. On A_n note that by Marshall's lemma we have

$$\begin{aligned} \|\widehat{F}_n - \mathbb{F}_n\| &\leq \|\widehat{F}_n - \mathbb{L}_n + \mathbb{L}_n - \mathbb{F}_n\| \\ &\leq \|\widehat{F}_n - \mathbb{L}_n\| + \|\mathbb{L}_n - \mathbb{F}_n\| \\ &\leq \|\mathbb{F}_n - \mathbb{L}_n\| + \|\mathbb{L}_n - \mathbb{F}_n\| = 2\|\mathbb{F}_n - \mathbb{L}_n\| \\ &\leq 2\{\|\mathbb{F}_n - F - (\mathbb{L}_n - L_n)\| + \|F - L_n\|\} \\ &\equiv 2D_n + 2E_n. \end{aligned}$$

  - Step 4. Fix $k \in \mathbb{N}$. Take $\mathbb{L}_n = I_2\mathbb{F}_n$, $L_n = I_2 F$ based on the knots $a_j \equiv F^{-1}(j/k)$, $j = 0, 1, \ldots, k-1$, $a_k = \alpha_1(F)$.

- From de Boor (2001), page 36, (18):

$$\|g - I_2 g\| \le \omega(g; |a|), \qquad |a| = \max_{1 \le j \le k} (a_j - a_{j-1}),$$

$$\omega(g; b) \equiv \sup\{|g(t) - g(s)| : \ |t - s| \le b\}.$$

- From de Boor (2001), page 36, (18):

$$\|g - I_2 g\| \le \omega(g; |a|), \qquad |a| = \max_{1\le j\le k}(a_j - a_{j-1}),$$

$$\omega(g; b) \equiv \sup\{|g(t) - g(s)| : \ |t-s| \le b\}.$$

- Thus with $R \equiv \max\{1, f(0)\}/f(\alpha_1(F))$,

$$\begin{aligned} D_n &= \|\mathbb{F}_n - F - I_2(\mathbb{F}_n - F)\| \le \omega(\mathbb{F}_n - F; |a|) \\ &\stackrel{d}{=} n^{-1/2}\omega(\mathbb{U}_n(F); |a|) \\ &\le n^{-1/2}\omega(\mathbb{U}_n; Rp_n), \quad p_n = 1/k_n \asymp C(n^{-1}\log n)^{1/3} \\ &\le n^{-1/2}\{(2+\epsilon)Rp_n \log(1/p_n)\}^{1/2} \ \text{a.s.} \\ &= O((n^{-1}\log n)^{2/3}) \ \text{a.s..} \end{aligned}$$

- For E_n, use the inequality of de Boor (2001), page 31, (2): if $\|g''\| < \infty$, then

$$\|g - I_2 g\| \le \frac{1}{8}|a|^2 \|g''\|.$$

- For E_n, use the inequality of de Boor (2001), page 31, (2): if $\|g''\| < \infty$, then

$$\|g - I_2 g\| \leq \frac{1}{8}|a|^2 \|g''\|.$$

- Taking $g = F$, this yields,

$$\begin{aligned}\|F - L_n\| &= \|F - I_2 F\| \leq \frac{1}{8}|a|^2 \|F''\| \\ &\leq \frac{1}{8}\gamma_1(F) p_n^2.\end{aligned}$$

- For E_n, use the inequality of de Boor (2001), page 31, (2): if $\|g''\| < \infty$, then

$$\|g - I_2 g\| \leq \frac{1}{8}|a|^2 \|g''\|.$$

- Taking $g = F$, this yields,

$$\begin{aligned} \|F - L_n\| &= \|F - I_2 F\| \leq \frac{1}{8}|a|^2 \|F''\| \\ &\leq \frac{1}{8}\gamma_1(F) p_n^2. \end{aligned}$$

- Thus we are done if we can prove “step 2” for $\mathbb{L}_n = I_2 \mathbb{F}_n$: i.e. $P_F(\mathbb{L}_n$ concave on $[0, \infty)) \geq 1 - n^{-2}$.

- **Lemma.** If $p_n \to 0$ and $\delta_n \to 0$, then for the uniform$(0,1)$ d.f. $F = I$,

$$Pr(|\mathbb{G}_n(p_n) - p_n| \geq \delta_n p_n) \leq 2\exp(-2^{-1} n p_n \delta_n^2 (1 + o(1)))$$

where $o(1)$ depends only on δ_n.

- **Lemma.** If $p_n \to 0$ and $\delta_n \to 0$, then for the uniform$(0,1)$ d.f. $F = I$,

$$Pr(|\mathbb{G}_n(p_n) - p_n| \geq \delta_n p_n) \leq 2\exp(-2^{-1} n p_n \delta_n^2 (1 + o(1)))$$

  where $o(1)$ depends only on δ_n.
- **Proof:** Hoeffding's inequality gives

$$Pr((\mathbb{G}_n(p_n) - p_n)^{\pm} \geq p_n \lambda) \leq \exp(-n p_n h(1 \pm \lambda))$$

  with $h(x) = x(\log x - 1) + 1$. Since $h(1 \pm \lambda) \sim 2^{-1}\lambda$ as $\lambda \searrow 0$, the lemma follows. $\square$

- **Lemma.** If $p_n \to 0$ and $\delta_n \to 0$, then for the uniform$(0,1)$ d.f. $F = I$,

$$Pr(|\mathbb{G}_n(p_n) - p_n| \geq \delta_n p_n) \leq 2\exp(-2^{-1} n p_n \delta_n^2 (1 + o(1)))$$

  where $o(1)$ depends only on δ_n.

- **Proof:** Hoeffding's inequality gives

$$Pr((\mathbb{G}_n(p_n) - p_n)^{\pm} \geq p_n \lambda) \leq \exp(-n p_n h(1 \pm \lambda))$$

  with $h(x) = x(\log x - 1) + 1$. Since $h(1 \pm \lambda) \sim 2^{-1}\lambda$ as $\lambda \searrow 0$, the lemma follows. $\square$

- **Lemma.** If $\beta_1(F) > 0$ and $\gamma_1(F) < \infty$, then for k_n large

$$Pr(A_n^c) \quad \leq \quad 4k_n \exp(-80^{-1}\beta_1^2(F) n p_n^3).$$

- **Proof:** Set $T_{n,j} \equiv \mathbb{F}_n(a_j) - \mathbb{F}_n(a_{j-1})$, $j = 1, \ldots, k = k_n$, and $\Delta_j a \equiv a_j - a_{j-1}$. Then

$$A_n = \bigcap_{j=1}^{k_n - 1} \left\{ \frac{T_{n,j}}{\Delta_j a} \ge \frac{T_{n,j+1}}{\Delta_{j+1} a} \right\} .$$

Thus

$$\begin{aligned} P(A_n^c) &\le \sum_{j=1}^{k_n-1} P\left\{ \frac{T_{n,j}}{\Delta_j a} < \frac{T_{n,j+1}}{\Delta_{j+1} a} \right\} \equiv \sum_{j=1}^{k_n-1} P(B_{n,j}) \\ &= \sum_{j=1}^{k_n-1} P\left(B_{n,j} \cap \{|T_{n,i} - 1/k_n| \le \delta_n/k_n\},\ i = j, j+1\right) \\ &\quad + \sum_{j=1}^{k_n-1} P\left(B_{n,j} \cap \{|T_{n,i} - 1/k_n| > \delta_n/k_n\},\ i = j, j+1\right) \\ &\equiv I_n + II_n . \end{aligned}$$

- If $|T_{n,i} - 1/k_n| \leq \delta_n/k_n$, $i = j, j+1$, and

$$\frac{\Delta_{j+1}a}{\Delta_j a} \geq 1 + 3\delta_n, \tag{1}$$

we have

$$\begin{aligned}
T_{n,j} &\geq \frac{1}{k_n} - \frac{\delta_n}{k_n} = \frac{1-\delta_n}{k_n}, \\
T_{n,j+1} &\leq \frac{1+\delta_n}{k_n}, \\
\frac{T_{n,j}}{\Delta_j a}\Delta_{j+1}a &\geq \frac{1-\delta_n}{k_n}(1+3\delta_n) \geq \frac{1-\delta_n}{k_n}\frac{1+\delta_n}{1-\delta_n} \geq T_{n,j+1}
\end{aligned}$$

if $\delta_n \leq 1/3$. Thus when (1) holds, the events in the sum I_n are empty and hence $I_n = 0$.

- To show that (1) holds, note that

$$
\begin{aligned}
\Delta_{j+1} a &= F^{-1}\left(\frac{j+1}{k}\right) - F^{-1}\left(\frac{j}{k}\right) \\
&= \frac{1}{k}\frac{1}{f(a_j)} + \frac{1}{2k_n^2}\frac{-f'}{f^3}(\xi), \quad a_j \le \xi \le a_{j+1} \\
\Delta_j a &\le \frac{1}{k}\frac{1}{f(a_j)},
\end{aligned}
$$

so

$$
\begin{aligned}
\frac{\Delta_{j+1} a}{\Delta_j a} &\ge 1 + \frac{1}{2k_n} f(a_j) \frac{-f'(\xi)}{f^3(\xi)} \ge 1 + \frac{1}{2k_n}\frac{-f'(\xi)}{f^2(\xi)} \\
&\ge 1 + \frac{1}{2k_n}\beta_1(F) \ge 1 + 3\delta_n
\end{aligned}
$$

if $\delta_n \equiv \beta_1(F)/(6k_n)$ (so $3\delta_n = \beta_1(F)/(2k_n)$).

- Then it follows from Lemma 1 that

$$
\begin{aligned}
II_n &\le 2\sum_{j=1}^{k_n} P(|T_{n,j} - 1/k_n| > \delta_n/k_n) \\
&\le 4\sum_{j=1}^{k_n} \exp(-2^{-1} n p_n \delta_n^2 (1+o(1))) \\
&\le 4k_n \exp(-2^{-1} n p_n^3 \frac{\beta_1^2(F)}{36}(1+o(1))) \\
&\le 4k_n \exp(-n p_n^3 \frac{\beta_1^2(F)}{80})
\end{aligned}
$$

for n sufficiently large. Choosing $1/k_n \asymp p_n = (7\log n/(3Kn))^{1/3}$ where $K = \beta_1^2(F)/80$ makes this bound smaller than $4n^{-2}$ for n sufficiently large. $\square$

3. The convex case, $k = 2$.

- Suppose that $\widetilde{f}_n = \widetilde{H}_n^{(2)}$ where:

  (a) $\widetilde{H}_n(x) \ge \mathbb{Y}_n(x) \equiv \int_0^x \mathbb{F}_n(s)ds$,

  (b) $\int_0^\infty (\widetilde{H}_n - \mathbb{Y}_n) d\widetilde{H}_n^{(3)} = 0$.

  Here we start with a local limit theorem which follows from Groeneboom, Jongbloed, and Wellner (2001).

3. The convex case, $k = 2$.

- Suppose that $\widetilde{f}_n = \widetilde{H}_n^{(2)}$ where:
  (a) $\widetilde{H}_n(x) \ge \mathbb{Y}_n(x) \equiv \int_0^x \mathbb{F}_n(s)ds$,
  (b) $\int_0^\infty (\widetilde{H}_n - \mathbb{Y}_n) d\widetilde{H}_n^{(3)} = 0$.
  Here we start with a local limit theorem which follows from Groeneboom, Jongbloed, and Wellner (2001).
- **Theorem.** If $f(t_0) > 0$, $f''(t_0) > 0$ and f and f'' are continuous in a neighborhood of t_0, then

$$\begin{pmatrix} n^{3/5}(\widetilde{F}_n(t_0 + n^{-1/5}t) - \mathbb{F}_n(t_0 + n^{-1/5}t)) \\ n^{4/5}(\widetilde{H}_n(t_0 + n^{-1/5}t) - \mathbb{Y}_n(t_0 + n^{-1/5}t)) \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} c_1(f, t_0)(H_2^{(1)}(at) - \mathbb{Y}_2^{(1)}(at) \\ c_2(f, t_0)(H_2(at) - \mathbb{Y}_2(at) \end{pmatrix}$$

where

- the processes $\mathbb{Y}_2$ and H_2 are described by

$$\mathbb{Y}_2(t) = \int_0^t W(s)ds + t^4,$$

$$H_2(t) \geq \mathbb{Y}_2(t) \ \text{ for all } \ t \in \mathbb{R},$$

$$\int_{-\infty}^{\infty} (H_2 - \mathbb{Y}_2) dH_2^{(3)} = 0,$$

- the processes $\mathbb{Y}_2$ and H_2 are described by

$$\mathbb{Y}_2(t) = \int_0^t W(s)ds + t^4,$$
$$H_2(t) \geq \mathbb{Y}_2(t) \;\; \text{for all} \;\; t \in \mathbb{R},$$
$$\int_{-\infty}^{\infty} (H_2 - \mathbb{Y}_2) dH_2^{(3)} = 0,$$

- the constants $c_j(f, t_0)$ are given by

$$c_1(f, t_0) = \left(24 \frac{f(t_0)^3}{f''(t_0)}\right)^{1/5}, \quad c_2(f, t_0) = \left(24^3 \frac{f(t_0)^4}{f''(t_0)^3}\right)^{1/5},$$

- the processes $\mathbb{Y}_2$ and H_2 are described by

$$\mathbb{Y}_2(t) = \int_0^t W(s)ds + t^4,$$
$$H_2(t) \geq \mathbb{Y}_2(t) \text{ for all } t \in \mathbb{R},$$
$$\int_{-\infty}^{\infty} (H_2 - \mathbb{Y}_2) dH_2^{(3)} = 0,$$

- the constants $c_j(f, t_0)$ are given by

$$c_1(f, t_0) = \left(24 \frac{f(t_0)^3}{f''(t_0)}\right)^{1/5}, \quad c_2(f, t_0) = \left(24^3 \frac{f(t_0)^4}{f''(t_0)^3}\right)^{1/5},$$

- and the constant a is defined by

$$a = \left(\frac{f''(t_0)^2}{24^2 f(t_0)}\right)^{1/5}.$$

- Global theory hypotheses:

- Global theory hypotheses:
    - **R1.** F has continuous 3rd derivative $F^{(3)} = f''(t) > 0$, $t \in [0, \tau]$, and

$$\beta_2(F, \tau) \equiv \inf_{0<t<\tau} \frac{f''(t)}{f^3(t)} > 0.$$

- Global theory hypotheses:
  - **R1.** F has continuous 3rd derivative $F^{(3)} = f''(t) > 0$, $t \in [0, \tau]$, and

$$\beta_2(F,\tau) \equiv \inf_{0<t<\tau} \frac{f''(t)}{f^3(t)} > 0.$$

  - **R2.** $\tilde{\gamma}_1(F,\tau) \equiv \sup_{0<t<\tau|}(-f'(t)/f^2(t)) < \infty$.

- Global theory hypotheses:
  - **R1.** F has continuous 3rd derivative $F^{(3)} = f''(t) > 0$, $t \in [0, \tau]$, and

$$\beta_2(F, \tau) \equiv \inf_{0<t<\tau} \frac{f''(t)}{f^3(t)} > 0.$$

  - **R2.** $\tilde{\gamma}_1(F, \tau) \equiv \sup_{0<t<\tau|}(-f'(t)/f^2(t)) < \infty$.
  - **R3.** $\gamma_2(F, \tau) \equiv \sup_{0<t<\tau} f''(t)/\inf_{0<t<\tau} f^3(t) < \infty$.

- Global theory hypotheses:
  - **R1.** F has continuous 3rd derivative $F^{(3)} = f''(t) > 0$, $t \in [0, \tau]$, and

$$\beta_2(F, \tau) \equiv \inf_{0<t<\tau} \frac{f''(t)}{f^3(t)} > 0.$$

  - **R2.** $\tilde{\gamma}_1(F, \tau) \equiv \sup_{0<t<\tau|}(-f'(t)/f^2(t)) < \infty$.
  - **R3.** $\gamma_2(F, \tau) \equiv \sup_{0<t<\tau} f''(t)/\inf_{0<t<\tau} f^3(t) < \infty$.
  - **R4.** $R \equiv \max\{1, f(0)\}/f(\tau) < \infty$.

- Global theory hypotheses:
  - **R1.** F has continuous 3rd derivative $F^{(3)} = f''(t) > 0$, $t \in [0, \tau]$, and

$$\beta_2(F,\tau) \equiv \inf_{0<t<\tau} \frac{f''(t)}{f^3(t)} > 0.$$

  - **R2.** $\tilde{\gamma}_1(F,\tau) \equiv \sup_{0<t<\tau|}(-f'(t)/f^2(t)) < \infty$.
  - **R3.** $\gamma_2(F,\tau) \equiv \sup_{0<t<\tau} f''(t)/\inf_{0<t<\tau} f^3(t) < \infty$.
  - **R4.** $R \equiv \max\{1, f(0)\}/f(\tau) < \infty$.
- **Theorem.** Suppose that **R1 - R4** hold. Then

$$\|\widetilde{F}_n - \mathbb{F}_n\| \equiv \sup_{0<t<\tau} |\widetilde{F}_n(t) - \mathbb{F}_n(t)| = O\left(\left(\frac{\log n}{n}\right)^{3/5}\right) \quad \text{a.s.}$$

- **Corollary.** If **R1-R4** hold, then

$$\sqrt{n}(\widetilde{F}_n - F) = \sqrt{n}(\mathbb{F}_n - F) + O(n^{-1/10}(\log n)^{3/5}) \text{ a.s.}$$

uniformly on $[0, \tau]$.

- **Corollary.** If **R1-R4** hold, then

$$\sqrt{n}(\widetilde{F}_n - F) = \sqrt{n}(\mathbb{F}_n - F) + O(n^{-1/10}(\log n)^{3/5}) \text{ a.s.}$$

  uniformly on $[0, \tau]$.
- Outline of the Proof of the KW theorem, convex case:

- **Corollary.** If **R1-R4** hold, then

$$\sqrt{n}(\widetilde{F}_n - F) = \sqrt{n}(\mathbb{F}_n - F) + O(n^{-1/10}(\log n)^{3/5}) \text{ a.s.}$$

  uniformly on $[0, \tau]$.

- Outline of the Proof of the KW theorem, convex case:
  - Step 1: Analogue of Marshall's lemma for the least squares estimator: for any h with convex second derivative,

$$\|\widetilde{\mathbb{H}}_n^{(1)} - h\| \leq 2\|\mathbb{F}_n - h\|.$$

- **Corollary.** If **R1-R4** hold, then

$$\sqrt{n}(\widetilde{F}_n - F) = \sqrt{n}(\mathbb{F}_n - F) + O(n^{-1/10}(\log n)^{3/5}) \text{ a.s.}$$

uniformly on $[0, \tau]$.

- Outline of the Proof of the KW theorem, convex case:
  - Step 1: Analogue of Marshall's lemma for the least squares estimator: for any h with convex second derivative,

$$\|\widetilde{\mathbb{H}}_n^{(1)} - h\| \le 2\|\mathbb{F}_n - h\|.$$

  - Step 2: Construct $\mathbb{H}_{n,k_n}$ (convenient), show that

$$P_F(A_n) \equiv P_F(\mathbb{H}_{n,k_n}^{(2)} \text{ is convex on } [0,\infty)) \ge 1 - n^{-2}$$

for all n sufficiently large if $k = k_n = (C_0\beta_2(F,\tau)n/\log n)^{1/5}$ for some C_0.

- Proof outline, continued:

- Proof outline, continued:
    - Step 3. On A_n note that by the analogue of Marshall's lemma we have

$$
\begin{aligned}
\|\widetilde{F}_n - \mathbb{F}_n\| &\le \|\widetilde{H}_n^{(1)} - \mathbb{H}_{n,k_n}^{(1)} + \mathbb{H}_{n,k_n}^{(1)} - \mathbb{F}_n\| \\
&\le \|\widetilde{H}_n^{(1)} - \mathbb{H}_{n,k_n}^{(1)}\| + \|\mathbb{H}_{n,k_n}^{(1)} - \mathbb{F}_n\| \\
&\le 2\|\mathbb{F}_n - \mathbb{H}_{n,k_n}^{(1)}\| + \|\mathbb{H}_{n,k_n}^{(1)}\| - \mathbb{F}_n\| \\
&= 3\|\mathbb{F}_n - \mathbb{H}_{n,k_n}^{(1)}\| \\
&\le 3\{\|\mathbb{F}_n - F - (\mathbb{H}_{n,k_n}^{(1)} - H_{k_n}^{(1)})\| + \|F - H_{k_n}^{(1)}\|\} \\
&\equiv 3D_n + 3E_n.
\end{aligned}
$$

- Proof outline, continued:
  - Step 3. On A_n note that by the analogue of Marshall's lemma we have

$$
\begin{aligned}
\|\widetilde{F}_n - \mathbb{F}_n\| &\le \|\widetilde{H}_n^{(1)} - \mathbb{H}_{n,k_n}^{(1)} + \mathbb{H}_{n,k_n}^{(1)} - \mathbb{F}_n\| \\
&\le \|\widetilde{H}_n^{(1)} - \mathbb{H}_{n,k_n}^{(1)}\| + \|\mathbb{H}_{n,k_n}^{(1)} - \mathbb{F}_n\| \\
&\le 2\|\mathbb{F}_n - \mathbb{H}_{n,k_n}^{(1)}\| + \|\mathbb{H}_{n,k_n}^{(1)}\| - \mathbb{F}_n\| \\
&= 3\|\mathbb{F}_n - \mathbb{H}_{n,k_n}^{(1)}\| \\
&\le 3\{\|\mathbb{F}_n - F - (\mathbb{H}_{n,k_n}^{(1)} - H_{k_n}^{(1)})\| + \|F - H_{k_n}^{(1)}\|\} \\
&\equiv 3D_n + 3E_n .
\end{aligned}
$$

  - Step 4. Fix $k \in \mathbb{N}$. Take $\mathbb{H}_{n,k_n} = I_4 \mathbb{Y}_n$, $H_n = I_4 Y$ where $I_4 g$ is the "complete spline interpolant" of g based on the knots $a_j \equiv F^{-1}((j/k)F(\tau))$, $j = 0, 1, \ldots, k$. Thus $\mathbb{H}_{n,k_n}^{(1)} = (I_4 \mathbb{Y}_n)^{(1)}$, $H_n = (I_4 Y)^{(1)}$.

- Interpolation theory bounds:

- Interpolation theory bounds:
    - If g has bounded 4th derivative $g^{(4)}$, then

$$\|g^{(1)} - (I_4 g)^{(1)}\| \leq \frac{1}{24}|a|^3\|g^{(4)}\|.$$

- Interpolation theory bounds:
  - If g has bounded 4th derivative $g^{(4)}$, then

$$\|g^{(1)} - (I_4 g)^{(1)}\| \leq \frac{1}{24}|a|^3 \|g^{(4)}\|.$$

  - Applying this to $Y(t) = \int_0^t F(s)ds$ yields

$$\begin{aligned} E_n &= \|Y^{(1)} - H_{k_n}^{(1)}\| \leq \frac{1}{24}|a|^3 \|Y^{(4)}\| \\ &\leq \frac{1}{24}\gamma_2(F,\tau)p_n^3 = O((n^{-1}\log n)^{3/5}). \end{aligned}$$

- Interpolation theory bounds:
  - If g has bounded 4th derivative $g^{(4)}$, then

$$\|g^{(1)} - (I_4 g)^{(1)}\| \le \frac{1}{24}|a|^3 \|g^{(4)}\|.$$

  - Applying this to $Y(t) = \int_0^t F(s)ds$ yields

$$\begin{aligned} E_n &= \|Y^{(1)} - H_{k_n}^{(1)}\| \le \frac{1}{24}|a|^3 \|Y^{(4)}\| \\ &\le \frac{1}{24}\gamma_2(F,\tau)p_n^3 = O((n^{-1}\log n)^{3/5}). \end{aligned}$$

- To handle the random term D_n, use de Boor's (2001) bounds together with empirical process oscillation theory:

- The random term D_n is bounded as follows:

$$
\begin{aligned}
D_n &= \|\mathbb{F}_n - F - (\mathbb{H}^{(1)}_{n,k_n} - H^{(1)}_{k_n})\| \\
&= \|(\mathbb{Y}_n - Y)^{(1)} - (I_4(\mathbb{Y}_n - Y))^{(1)}\| \\
&\le (19/4)\text{dist}((\mathbb{Y}_n - Y)^{(1)}, \$_3) \\
&\le (19/4)\text{dist}((\mathbb{Y}_n - Y)^{(1)}, \$_2) \\
&\le (19/4)\|(\mathbb{Y}_n - Y)^{(1)} - [I_2(\mathbb{Y}_n - Y))^{(1)}]\| \\
&\le (19/4)\|\mathbb{F}_n - F - I_2(\mathbb{F}_n - F)\| \\
&\le (19/4)\omega(\mathbb{F}_n - F; |a|) \le (19/4)n^{-1/2}\omega(\mathbb{U}_n; Rp_n) \\
&\le O(n^{-1/2}\sqrt{p_n \log(1/p_n)}) \quad \text{a.s.} \\
&= O((n^{-1}\log n)^{3/5})
\end{aligned}
$$

if we choose $p_n = (Mn^{-1}\log n)^{1/5}$ for some M; cf. de Boor (2001), pages 56 and 36.

- It remains to prove that step 2 holds: i.e. show that

$$P_F(A_n) \equiv P_F(\mathbb{H}^{(2)}_{n,k_n} \text{ is convex on } [0,\infty)) \geq 1 - n^{-2}$$

for n sufficiently large.

- It remains to prove that step 2 holds: i.e. show that

$$P_F(A_n) \equiv P_F(\mathbb{H}^{(2)}_{n,k_n} \text{ is convex on } [0,\infty)) \geq 1 - n^{-2}$$

  for n sufficiently large.
- Let

$$\mathbb{S}_{n,k} \equiv \mathbb{H}_{n,k_n} = I_4 \mathbb{Y}_n \equiv \mathcal{C}\mathbb{Y}_n,$$
$$S_{n,k} = H_{k_n} = I_4 Y \equiv \mathcal{C}Y.$$

- It remains to prove that step 2 holds: i.e. show that

$$P_F(A_n) \equiv P_F(\mathbb{H}^{(2)}_{n,k_n} \text{ is convex on } [0,\infty)) \geq 1 - n^{-2}$$

for n sufficiently large.

- Let

$$\mathbb{S}_{n,k} \equiv \mathbb{H}_{n,k_n} = I_4 \mathbb{Y}_n \equiv \mathcal{C}\mathbb{Y}_n,$$
$$S_{n,k} = H_{k_n} = I_4 Y \equiv \mathcal{C}Y.$$

- The slope of $\mathbb{S}^{(2)}_{n,k}$ on $[a_{j-1}, a_j]$ is

$$B_j \equiv B_j(CS) = \frac{12}{(\Delta_j a)^3} \left\{ \frac{\mathbb{S}^{(1)}_{n,k}(a_{j-1}) + \mathbb{S}^{(1)}_{n,k}(a_j)}{2} \Delta_j a - \Delta_j \mathbb{Y}_n \right\}$$
$$\equiv \frac{12}{(\Delta_j a)^3} T_{n,j}$$

- The slope of the Hermite interpolant on $[a_{j-1}, a_j]$ is

$$\begin{aligned}
\tilde{B}_j \equiv B_j(Herm) &= \frac{12}{(\Delta_j a)^3}\left\{\frac{\mathbb{F}_n(a_{j-1}) + \mathbb{F}_n(a_j)}{2}\Delta_j a - \Delta_j \mathbb{Y}_n\right\} \\
&\equiv \frac{12}{(\Delta_j a)^3} R_{n,j}
\end{aligned}$$

- The slope of the Hermite interpolant on $[a_{j-1}, a_j]$ is

$$\begin{aligned}\tilde{B}_j \equiv B_j(Herm) &= \frac{12}{(\Delta_j a)^3}\left\{\frac{\mathbb{F}_n(a_{j-1}) + \mathbb{F}_n(a_j)}{2}\Delta_j a - \Delta_j \mathbb{Y}_n\right\} \\ &\equiv \frac{12}{(\Delta_j a)^3} R_{n,j}\end{aligned}$$

- Corresponding to the random $T_{n,j}$ and $R_{n,j}$ the corresponding population values are given by

$$t_{n,j} = \frac{(\mathcal{C}[Y])^{(1)}(a_{j-1}) + (\mathcal{C}[Y])^{(1)}(a_j)}{2}\Delta_j a - \Delta_j Y,$$

$$r_{n,j} = \frac{Y^{(1)}(a_{j-1}) + Y^{(1)}(a_j)}{2}\Delta_j a - \Delta_j Y.$$

- We write

$$
\begin{aligned}
T_j - r_j &= T_j - t_j + t_j - r_j \\
&= R_j - r_j + (T_j - t_j - (R_j - r_j)) + t_j - r_j \\
&\equiv R_j - r_j + A_j + b_j.
\end{aligned}
$$

- We write

$$\begin{aligned} T_j - r_j &= T_j - t_j + t_j - r_j \\ &= R_j - r_j + (T_j - t_j - (R_j - r_j)) + t_j - r_j \\ &\equiv R_j - r_j + A_j + b_j. \end{aligned}$$

- For $0 \le s < t < \infty$, set

$$h_{s,t}(x) = \left(x - \frac{s+t}{2}\right) 1_{[s,t]}(x).$$

- We write

$$\begin{aligned} T_j - r_j &= T_j - t_j + t_j - r_j \\ &= R_j - r_j + (T_j - t_j - (R_j - r_j)) + t_j - r_j \\ &\equiv R_j - r_j + A_j + b_j. \end{aligned}$$

- For $0 \le s < t < \infty$, set

$$h_{s,t}(x) = \left(x - \frac{s+t}{2} \right) 1_{[s,t]}(x).$$

- Then:

$$Ph_{s,t} = \int h_{s,t}(x) dF(x) = \frac{1}{2}(F(t) + F(s))(t-s) - \int_s^t F(u)du,$$

$$\mathbb{P}_n h_{s,t} = \int h_{s,t}(x) d\mathbb{F}_n(x) = \frac{1}{2}(\mathbb{F}_n(t) + \mathbb{F}_n(s))(t-s) - \int_s^t \mathbb{F}_n(u)du,$$

$$r_{n,j} = Ph_{a_{j-1},a_j}, \quad R_{n,j} = \mathbb{P}_n h_{a_{j-1},a_j}.$$

- **Lemma.**

$$
\begin{aligned}
Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\
&\leq 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).
\end{aligned}
$$

- **Lemma.**

$$\begin{aligned} Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\ &\leq 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))). \end{aligned}$$

- **Proof.** Bernstein's inequality. $\square$

- **Lemma.**

$$\begin{aligned} Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\ &\leq 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))). \end{aligned}$$

- **Proof.** Bernstein's inequality. $\square$
- **Lemma.** With $A_j \equiv T_{n,j} - t_{n,j} - (R_{n,j} - r_{n,j})$,

$$Pr(|A_j| > \delta_n p_n^3) \leq 4\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).$$

- **Lemma.**

$$
\begin{aligned}
Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\
&\leq 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).
\end{aligned}
$$

- **Proof.** Bernstein's inequality. $\square$
- **Lemma.** With $A_j \equiv T_{n,j} - t_{n,j} - (R_{n,j} - r_{n,j})$,

$$
Pr(|A_j| > \delta_n p_n^3) \leq 4\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).
$$

- **Proof.** Interpolation theory bounds and Bernstein's inequality. $\square$

- **Lemma.**

$$\begin{aligned} Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\ &\leq 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1+o(1))). \end{aligned}$$

- **Proof.** Bernstein's inequality. $\square$
- **Lemma.** With $A_j \equiv T_{n,j} - t_{n,j} - (R_{n,j} - r_{n,j})$,

$$Pr(|A_j| > \delta_n p_n^3) \leq 4\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1+o(1))).$$

- **Proof.** Interpolation theory bounds and Bernstein's inequality. $\square$
- **Lemma.** $\max_{1 \leq j \leq k} |t_{n,j} - r_{n,j}|/(\Delta_j a)^4 = o(1)$.

- **Lemma.**

$$\begin{aligned} Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\ &\le 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))). \end{aligned}$$

- **Proof.** Bernstein's inequality. □
- **Lemma.** With $A_j \equiv T_{n,j} - t_{n,j} - (R_{n,j} - r_{n,j})$,

$$Pr(|A_j| > \delta_n p_n^3) \le 4\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).$$

- **Proof.** Interpolation theory bounds and Bernstein's inequality. □
- **Lemma.** $\max_{1\le j\le k} |t_{n,j} - r_{n,j}|/(\Delta_j a)^4 = o(1)$.
- **Lemma.**

$$Pr(|T_{n,j} - r_{n,j}| > 3\delta_n p_n^3) \le 6\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).$$

- **Lemma.**

$$\begin{aligned} Pr(|R_{n,j} - r_{n,j}| > \delta_n p_n^3) &= Pr(|(\mathbb{P}_n - P)h_{s,t}| > \delta_n p_n^3) \\ &\leq 2\exp(-3n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))). \end{aligned}$$

- **Proof.** Bernstein's inequality. $\square$
- **Lemma.** With $A_j \equiv T_{n,j} - t_{n,j} - (R_{n,j} - r_{n,j})$,

$$Pr(|A_j| > \delta_n p_n^3) \leq 4\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).$$

- **Proof.** Interpolation theory bounds and Bernstein's inequality. $\square$
- **Lemma.** $\max_{1\leq j\leq k} |t_{n,j} - r_{n,j}|/(\Delta_j a)^4 = o(1)$.
- **Lemma.**

$$Pr(|T_{n,j} - r_{n,j}| > 3\delta_n p_n^3) \leq 6\exp(-100^{-1} n\delta_n^2 p_n^3 f^2(a_j^*)(1 + o(1))).$$

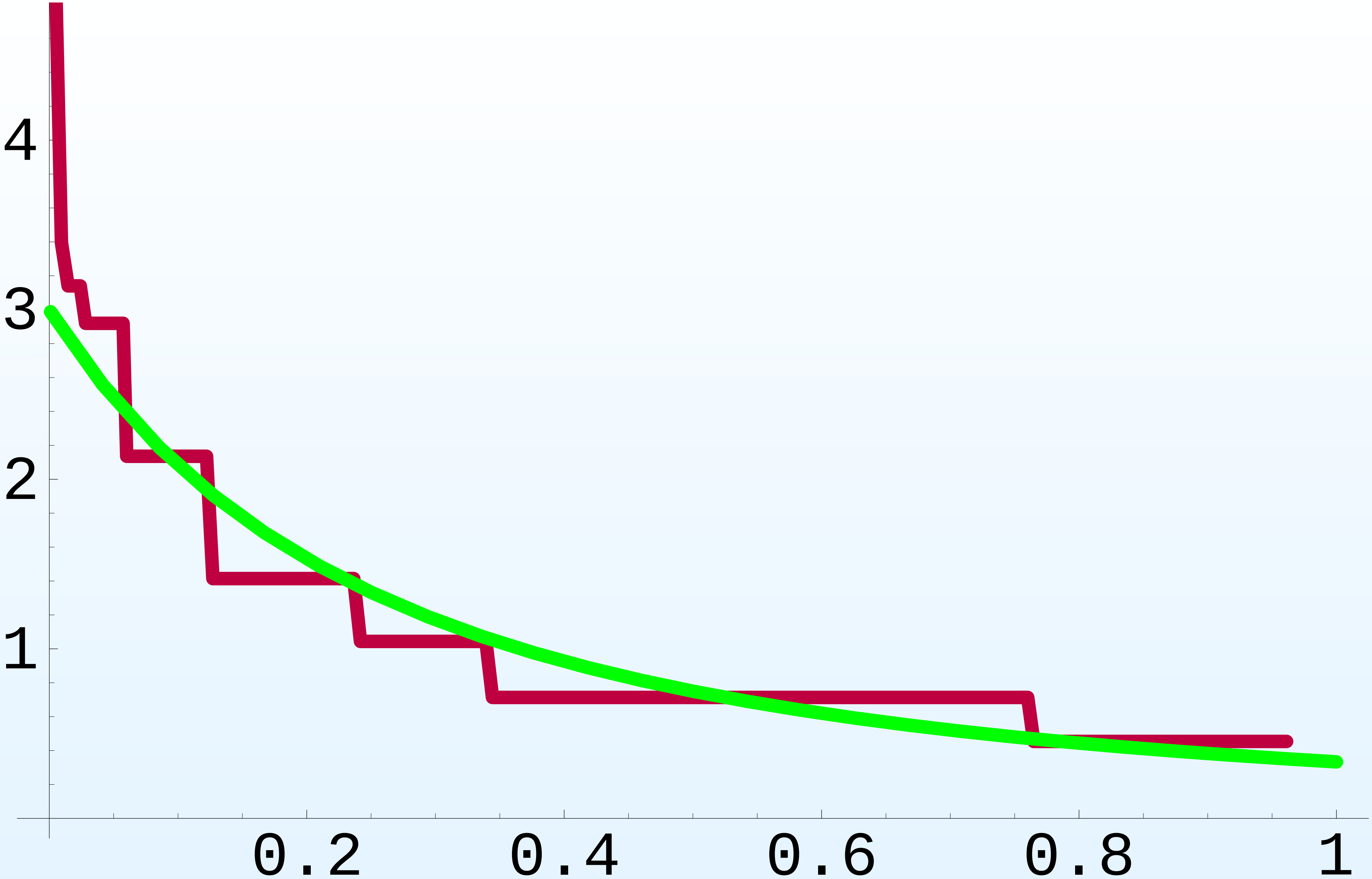
- **Proof.** Combine the previous 3 lemmas $\square$

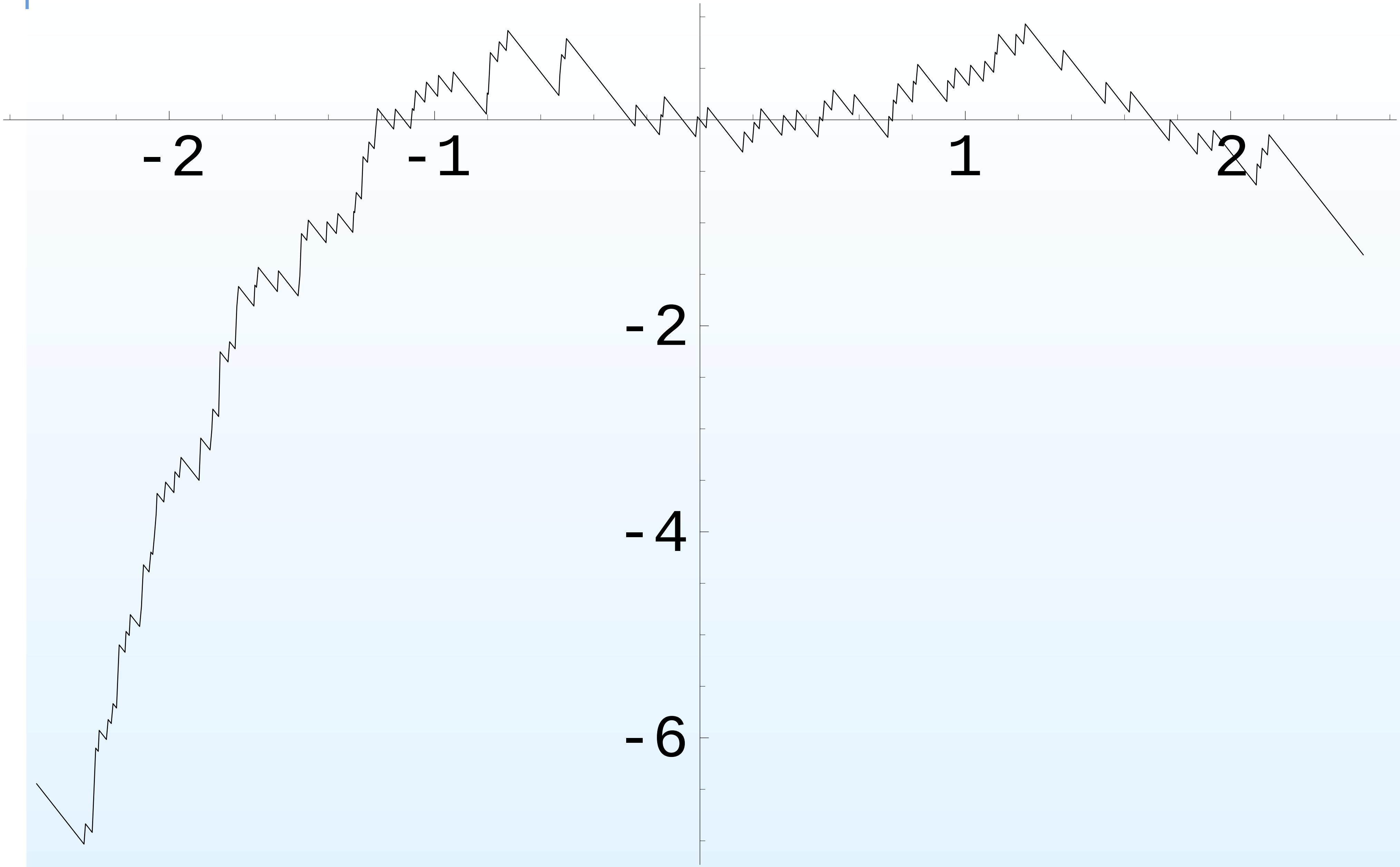
- **Lemma.** (Final exponential bound for step 2.)

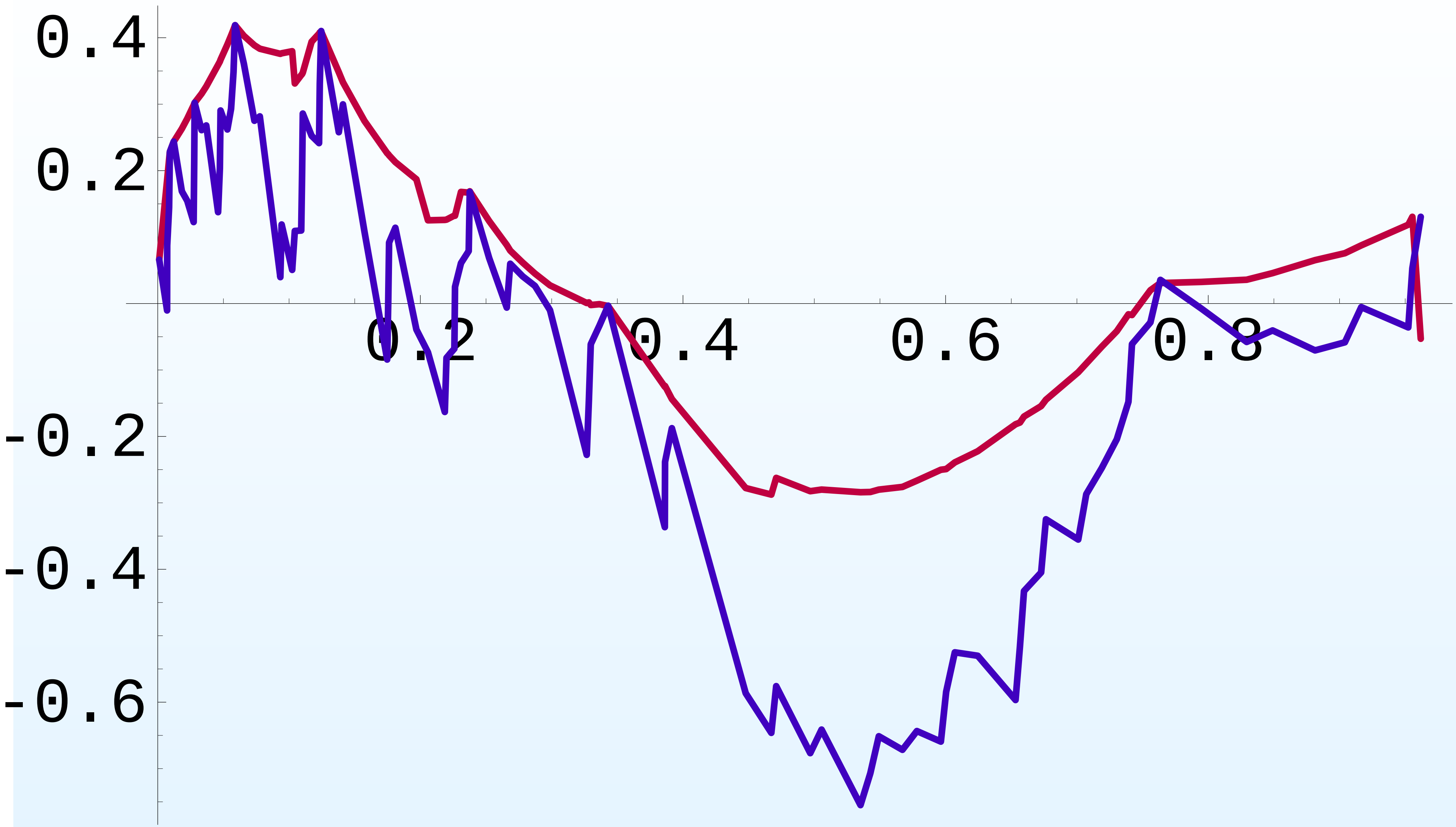
$$P_F(A_n^c) \le 12 k_n \exp(-K \beta_2^2(F, \tau) n p_n^5)$$

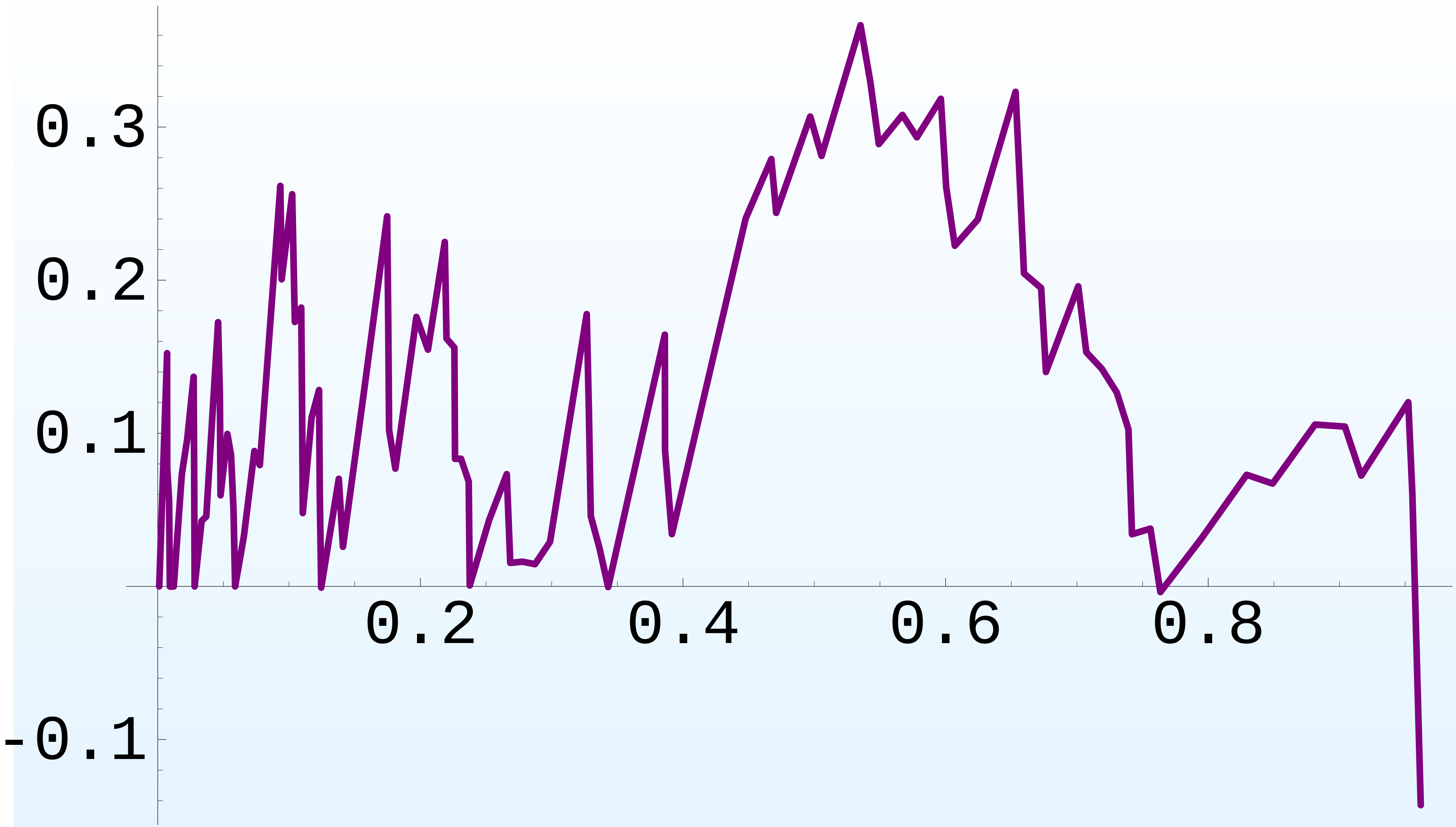
where $K^{-1} = 4246732800 \le 4.3 \times 10^9$.

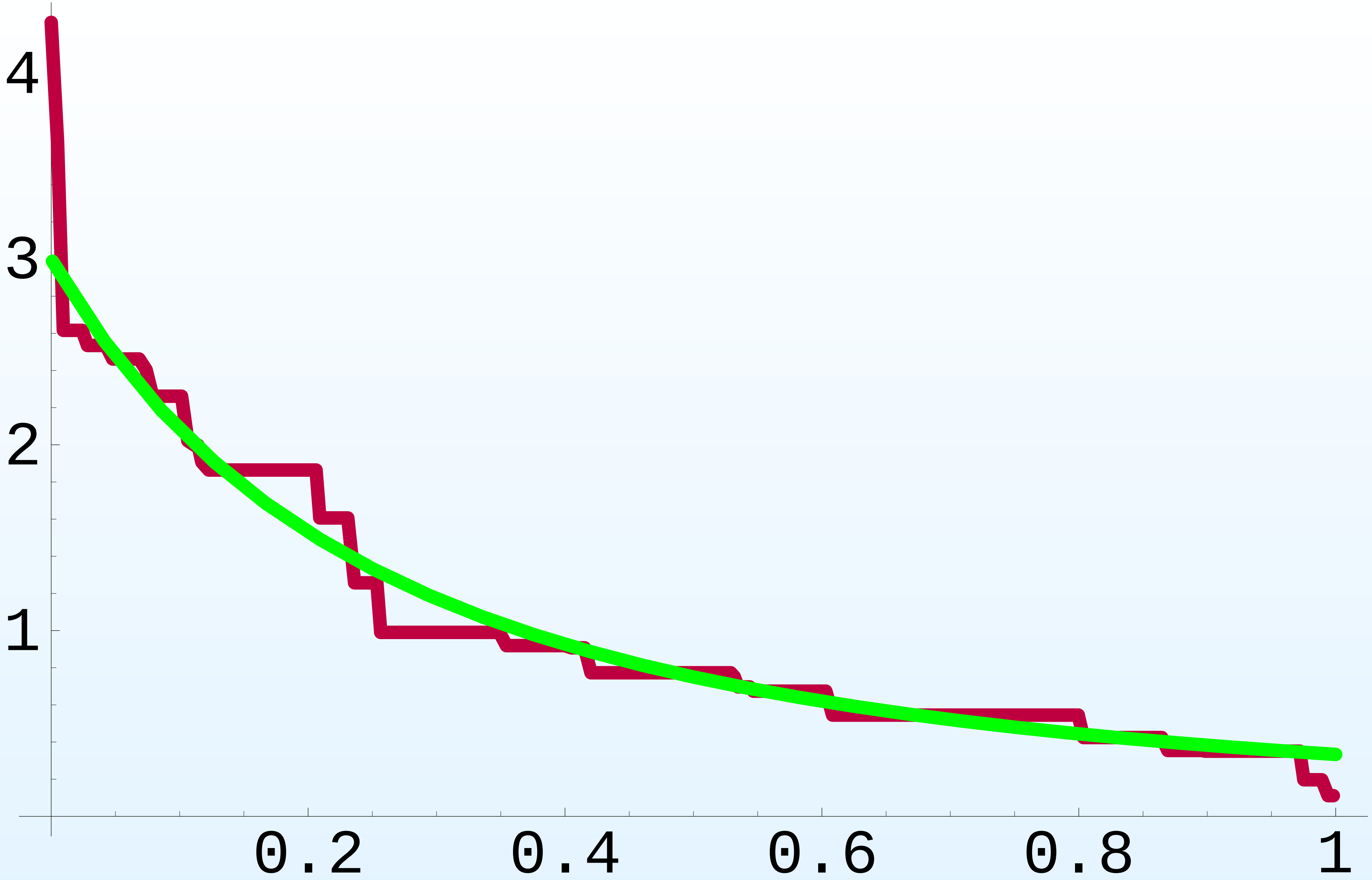
4. Some illustrative plots

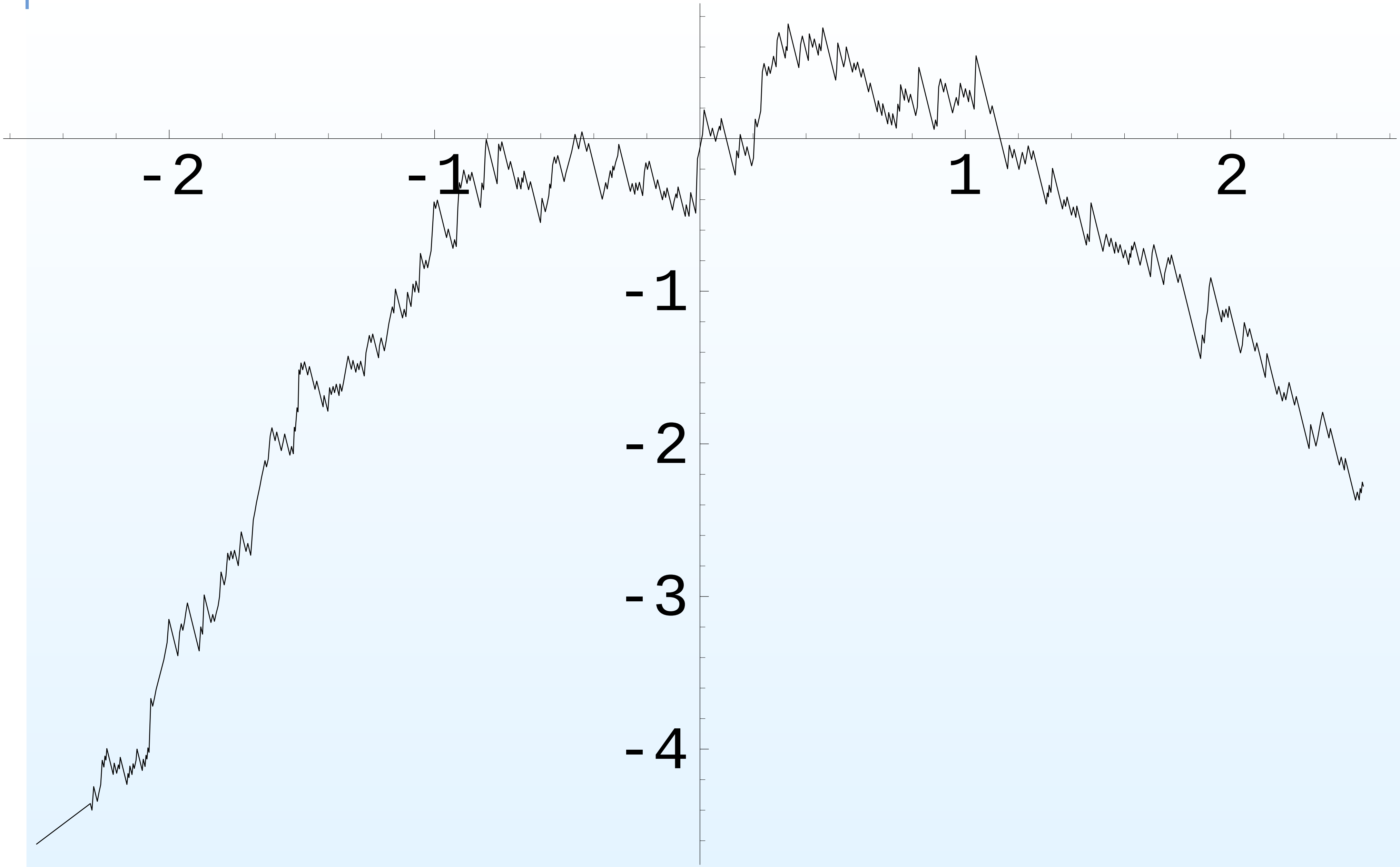

4
3
2
1
0.2
0.4
0.6
0.8
1

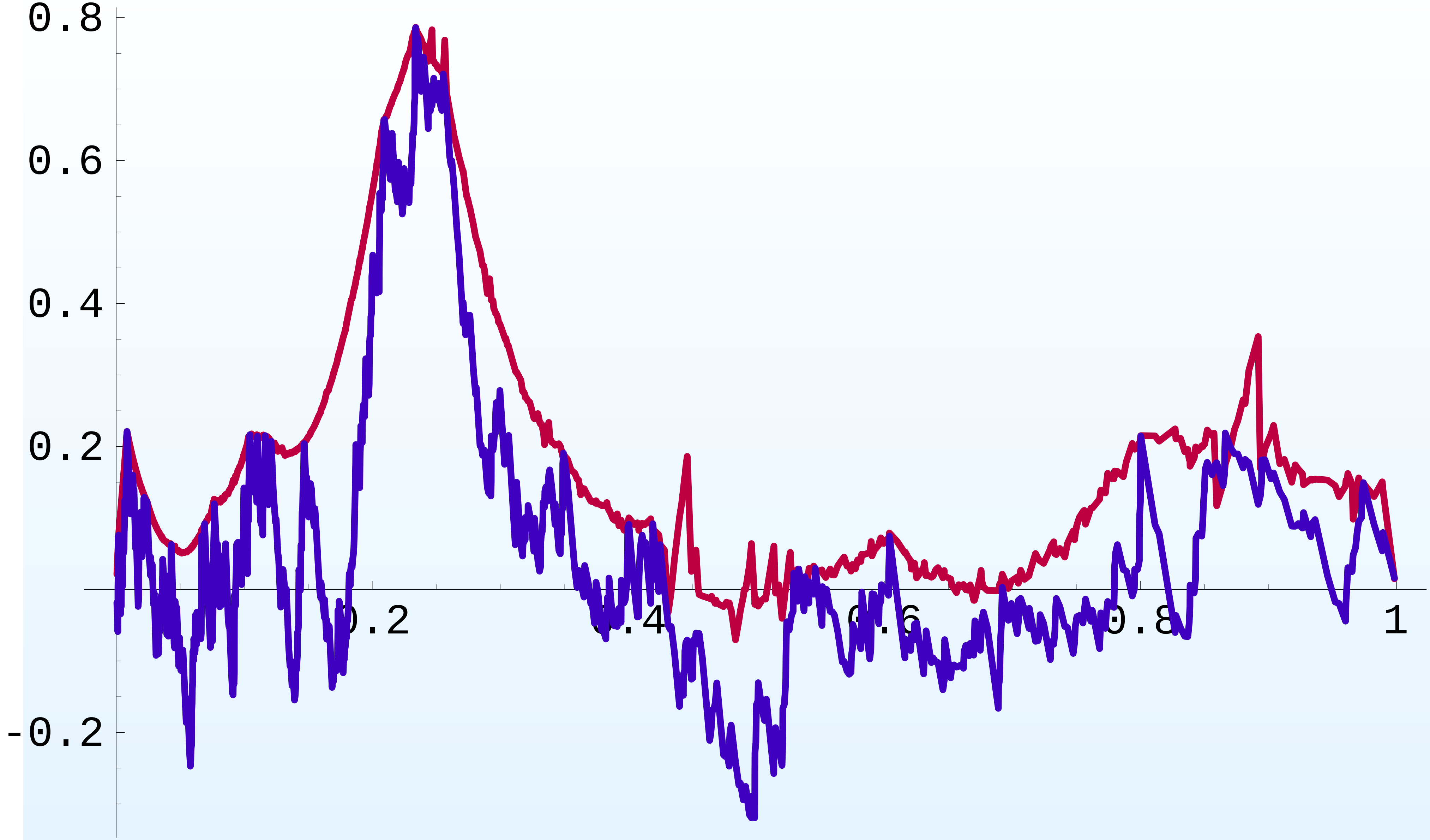

-2
-1
1
2
-2
-4
-6

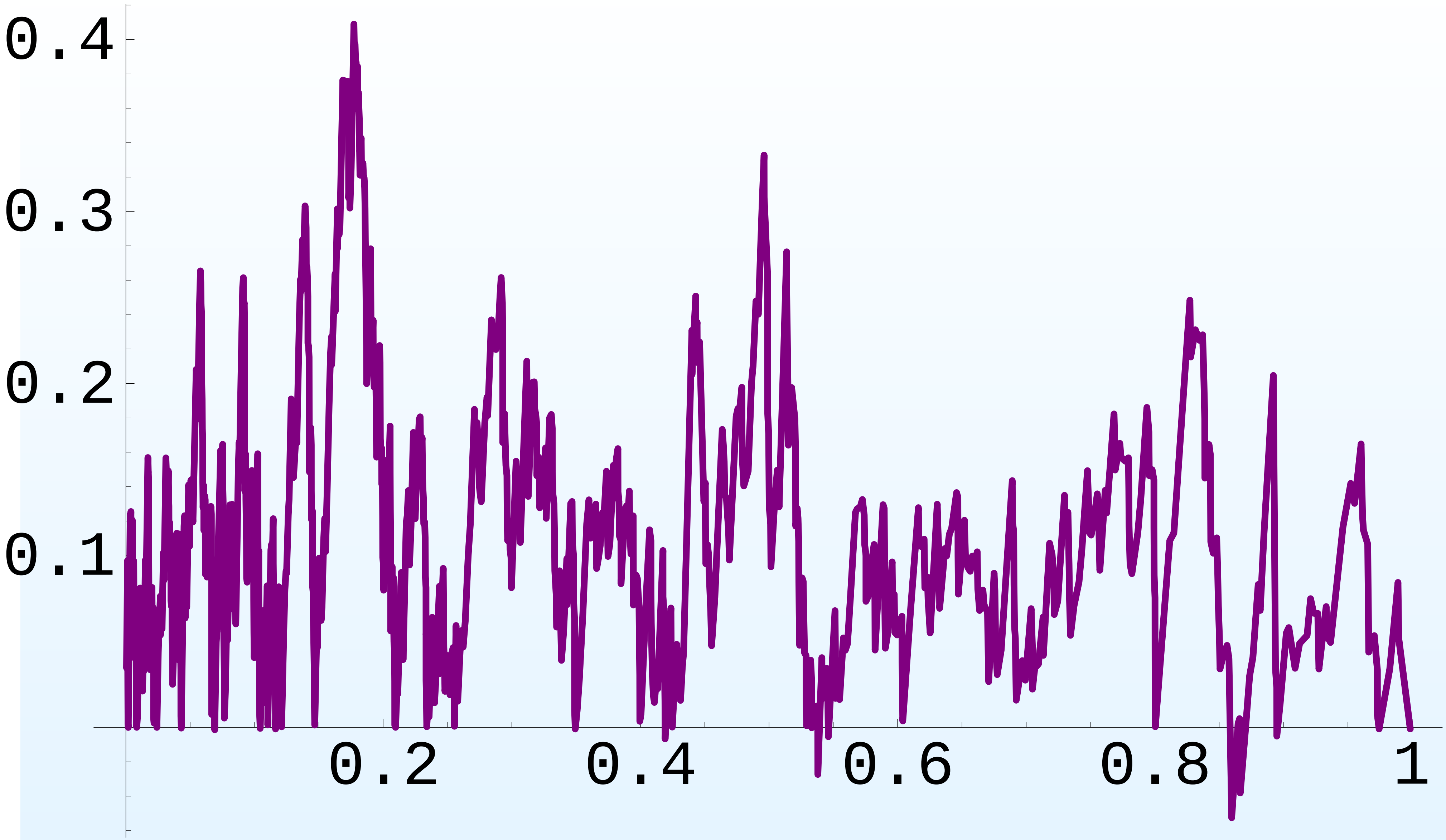

0.4
0.2
-0.2
-0.4
-0.6
0.2
0.4
0.6
0.8

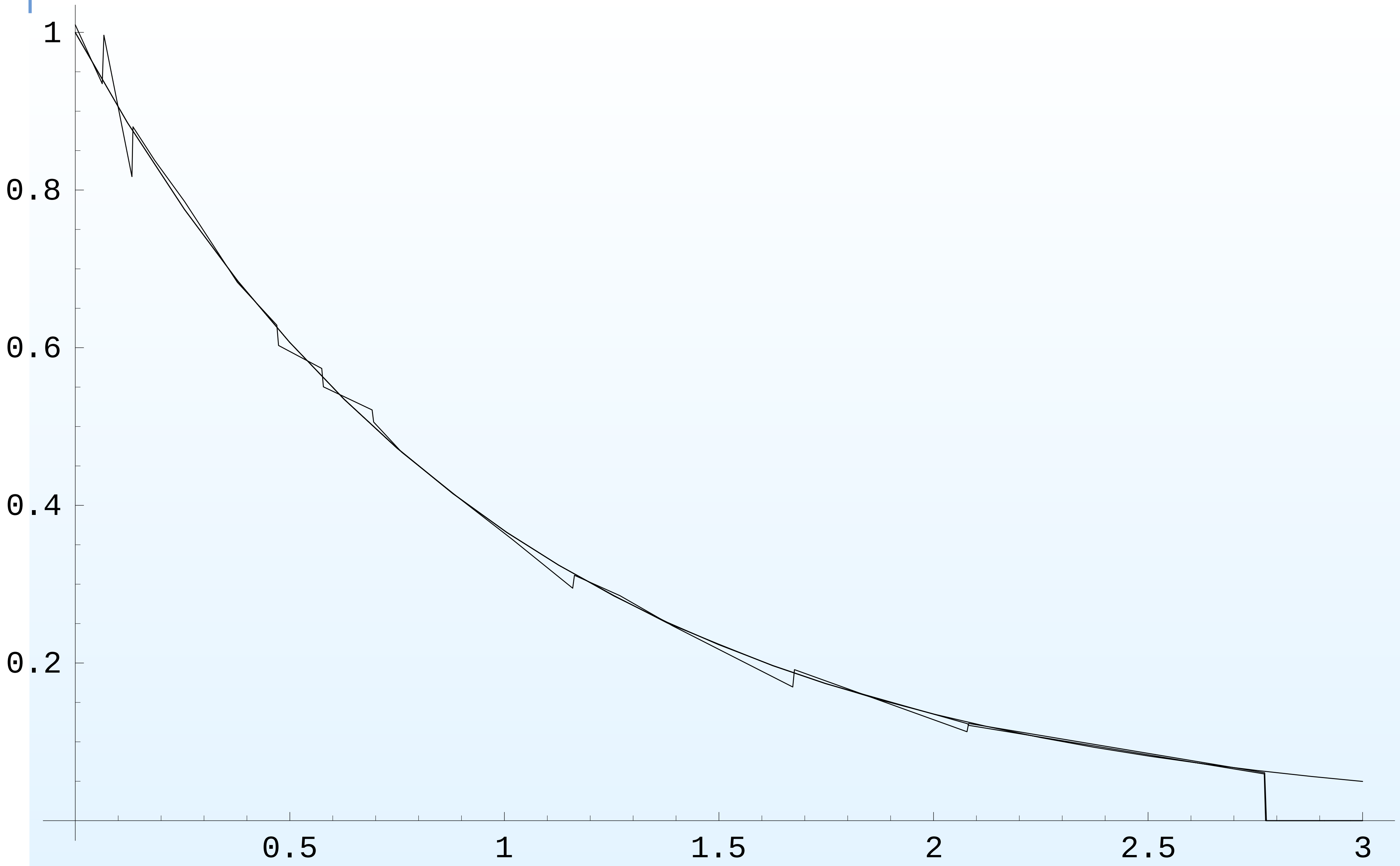

0.3
0.2
0.1
-0.1
0.2
0.4
0.6
0.8

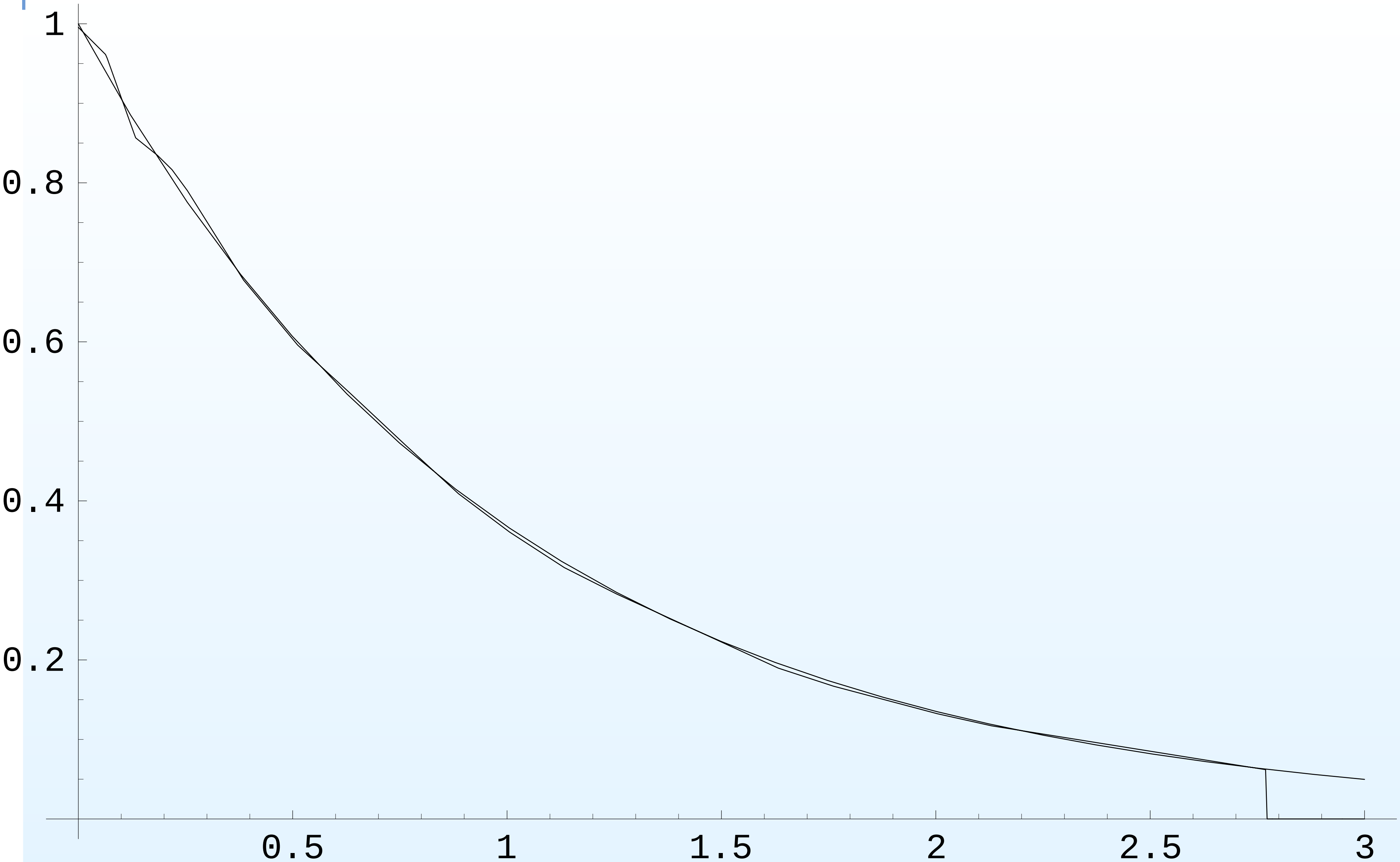

4
3
2
1
0.2
0.4
0.6
0.8
1


-2
-1
1
2
-1
-2
-3
-4


0.8
0.6
0.4
0.2
-0.2
0.2
0.4
0.6
0.8
1


0.4
0.3
0.2
0.1
0.2
0.4
0.6
0.8
1


1
0.8
0.6
0.4
0.2
0.5
1
1.5
2
2.5
3


1
0.8
0.6
0.4
0.2
0.5
1
1.5
2
2.5
3

5. Selected references

- Kiefer, J. and Wolfowitz, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34**, 73 - 85.

5. Selected references

- Kiefer, J. and Wolfowitz, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34**, 73 - 85.
- de Boor, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York.

5. Selected references

- Kiefer, J. and Wolfowitz, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34**, 73 - 85.
- de Boor, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York.
- Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Th. Rel. Fields* **81**, 79 - 109.

5. Selected references

- Kiefer, J. and Wolfowitz, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34**, 73 - 85.
- de Boor, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York.
- Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Th. Rel. Fields* **81**, 79 - 109.
- Groeneboom, P., Jongbloed, G. and Wellner, J. A. (2001). Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist.* **29**, 1653 - 1698.

5. Selected references

- Kiefer, J. and Wolfowitz, J. (1976). Asymptotically minimax estimation of concave and convex distribution functions. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **34**, 73 - 85.
- de Boor, C. (2001). *A Practical Guide to Splines*. Springer-Verlag, New York.
- Groeneboom, P. (1989). Brownian motion with a parabolic drift and Airy functions. *Probab. Th. Rel. Fields* **81**, 79 - 109.
- Groeneboom, P., Jongbloed, G. and Wellner, J. A. (2001). Estimation of a convex function: characterizations and asymptotic theory. *Ann. Statist.* **29**, 1653 - 1698.
- Dümbgen, L., Rufibach, K., and Wellner, J. A. (2006). Marshall's lemma for convex density estimation. Technical report, University of Bern.