Exponential inequalities for self-normalized martingales

B. BERCU and A. TOUATI

University of Toulouse, France

Conference on Probability with Applications,
The University of Hong Kong, December 2005

Dedicated to Professor T. L. Lai on the occasion of his sixtieth birthday

Outline

1. Classical exponential inequalities
   - Azuma-Hoeffding's inequality
   - Freedman's inequality
   - De la Peña's inequalities
2. Main results
   - Heavy on left or on right
   - A keystone lemma
   - New exponential inequalities
3. Statistical application

Outline

Azuma-Hoeffding's inequality

Let (M_n) be a square integrable martingale adapted to $\mathbb{F} = (\mathcal{F}_n)$ with $M_0 = 0$. The **predictable** and the **total** quadratic variations of (M_n) are given by

$$< M >_n = \sum_{k=1}^{n} \mathbb{E}[\Delta M_k^2 | \mathcal{F}_{k-1}], \qquad [M]_n = \sum_{k=1}^{n} \Delta M_k^2$$

$\Delta M_n = M_n - M_{n-1}$.

Theorem (Azuma-Hoeffding's inequality)

Assume that for each $1 \leqslant k \leqslant n$, $a_k \leqslant \Delta M_k \leqslant b_k$ *a.s. for some constants* $a_k < b_k$. *Then,* $\forall x \geqslant 0$,

$$\mathbb{P}(|M_n| \geqslant x) \leqslant 2 \exp\left(-\frac{2x^2}{\sum_{k=1}^{n}(b_k - a_k)^2}\right).$$

Outline

Freedman's inequality

Theorem (Freedman's inequality)

Assume that for each $1 \leqslant k \leqslant n$, $|\Delta M_k| \leqslant c$ *a.s. for some constant* $c > 0$*. Then,* $\forall x, y > 0$,

$$\mathbb{P}(M_n \geqslant x, <M>_n \leqslant y) \leqslant \exp\left(-\frac{x^2}{2(y+cx)}\right).$$

Theorem

Freedman's inequality also holds under the Bernstein moment condition: $\forall n \geqslant 1$, $p \geqslant 2$ *and for some constant* $c > 0$

$$\sum_{k=1}^{n} \mathbb{E}[|\Delta M_k|^p | \mathcal{F}_{k-1}] \leqslant \frac{p!}{2} c^{p-2} <M>_n \qquad a.s.$$

Outline

De la Peña's inequalities

Definition. We shall say that (M_n) is **conditionally symmetric** if, $\forall n \geqslant 1$, $\mathcal{L}(\Delta M_n | \mathcal{F}_{n-1})$ is symmetric.

Theorem (De la Peña)

Assume that (M_n) is conditionally symmetric. Then, $\forall x, y > 0$

$$\mathbb{P}(M_n \geqslant x, [M]_n \leqslant y) \leqslant \exp\left(-\frac{x^2}{2y}\right).$$

Self-normalized martingales

Theorem (De la Peña)

Assume that (M_n) is conditionally symmetric. Then, $\forall x, y > 0$ and $\forall a \geqslant 0$, $b > 0$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x\Big) \leqslant \sqrt{\mathbb{E}\Big[\exp\Big(-x^2\Big(ab + \frac{b^2}{2}[M]_n\Big)\Big)\Big]},$$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x, [M]_n \geqslant \frac{1}{y}\Big) \leqslant \exp\Big(-x^2\Big(ab + \frac{b^2}{2y}\Big)\Big).$$

Goal. Avoid the symmetric condition on (M_n).

Self-normalized martingales

Theorem (De la Peña)

Assume that (M_n) *is conditionally symmetric. Then,* $\forall x, y > 0$ *and* $\forall a \geqslant 0$, $b > 0$

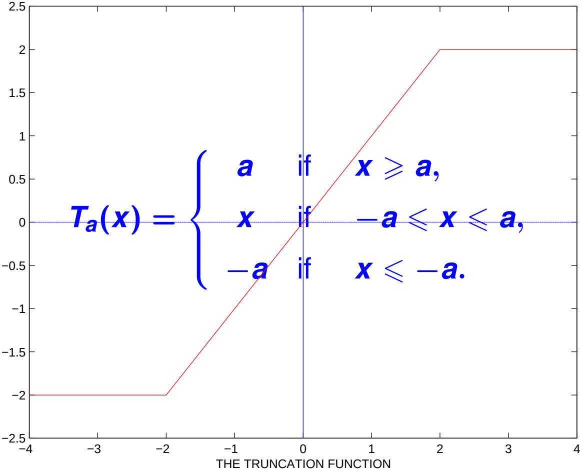
$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x\Big) \leqslant \sqrt{\mathbb{E}\Big[\exp\Big(-x^2\Big(ab + \frac{b^2}{2}[M]_n\Big)\Big)\Big]},$$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x, [M]_n \geqslant \frac{1}{y}\Big) \leqslant \exp\Big(-x^2\Big(ab + \frac{b^2}{2y}\Big)\Big).$$

Goal. Avoid the symmetric condition on (M_n).

Outline

Heavy on left or on right



$$T_a(x) = \begin{cases} a & \text{if} \quad x \geqslant a, \\ x & \text{if} \quad -a \leqslant x \leqslant a, \\ -a & \text{if} \quad x \leqslant -a. \end{cases}$$

Heavy on left or right

Definition. Let X be a random variable on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We shall say that

- X is heavy on left if, $\forall a > 0$, $\mathbb{E}[T_a(X)] \leqslant 0$,
- X is heavy on right if, $\forall a > 0$, $\mathbb{E}[T_a(X)] \geqslant 0$.

X is symmetric $\iff$ X is heavy on left and on right.

Heavy on left or right

Definition. Let X be a random variable on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$. We shall say that

- X is heavy on left if, $\forall a > 0$, $\mathbb{E}[T_a(X)] \leqslant 0$,
- X is heavy on right if, $\forall a > 0$, $\mathbb{E}[T_a(X)] \geqslant 0$.

X is symmetric $\Longleftrightarrow$ X is heavy on left and on right.

Heavy on left or right

Denote by F the distribution function of X and

$$H(a) = \int_0^a F(-x) - (1 - F(x))\, dx = -\mathbb{E}[T_a(X)].$$

- X is heavy on left if, $\forall a > 0$, $H(a) \geqslant 0$,
- X is heavy on right if, $\forall a > 0$, $H(a) \leqslant 0$.

Heavy on left or right

Denote by F the distribution function of X and

$$H(a) = \int_0^a F(-x) - (1 - F(x))\, dx = -\mathbb{E}[T_a(X)].$$

- X is heavy on left if, $\forall a > 0$, $H(a) \geqslant 0$,
- X is heavy on right if, $\forall a > 0$, $H(a) \leqslant 0$.

Outline

A keystone lemma

For all $t \in \mathbb{R}$, let

$$L(t) = \mathbb{E}\left[\exp\left(tX - \frac{t^2}{2}X^2\right)\right].$$

Lemma (Bercu-Touati)

Assume that $X \in \mathbb{L}^1(\mathbb{R})$ with $\mathbb{E}[X] = 0$.

- *X is heavy on left $\Longleftrightarrow \forall t \geqslant 0,\ L(t) \leqslant 1$,*
- *X is heavy on right $\Longleftrightarrow \forall t \leqslant 0,\ L(t) \leqslant 1$,*
- *X is symmetric $\Longleftrightarrow \forall t \in \mathbb{R},\ L(t) \leqslant 1$.*

A keystone lemma

For all $t \in \mathbb{R}$, let

$$L(t) = \mathbb{E}\left[\exp\left(tX - \frac{t^2}{2}X^2\right)\right].$$

Lemma (Bercu-Touati)

Assume that $X \in \mathbb{L}^1(\mathbb{R})$ with $\mathbb{E}[X] = 0$.

- *X is heavy on left $\Longleftrightarrow \forall t \geqslant 0,\ L(t) \leqslant 1$,*
- *X is heavy on right $\Longleftrightarrow \forall t \leqslant 0,\ L(t) \leqslant 1$,*
- *X is symmetric $\Longleftrightarrow \forall t \in \mathbb{R},\ L(t) \leqslant 1$.*

A keystone lemma

For all $t \in \mathbb{R}$, let

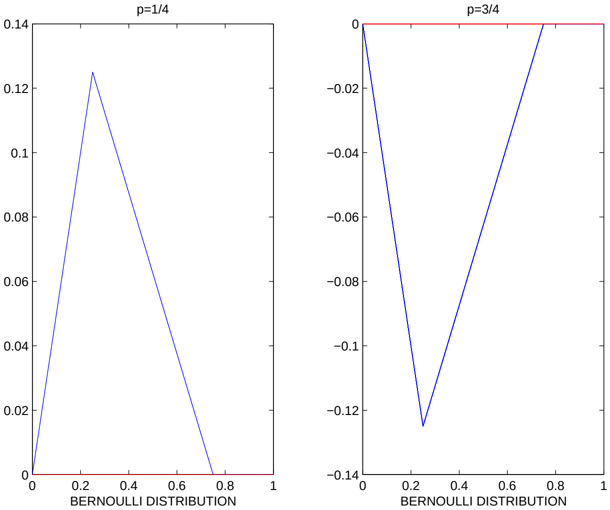
$$L(t) = \mathbb{E}\left[\exp\left(tX - \frac{t^2}{2}X^2\right)\right].$$

Lemma (Bercu-Touati)

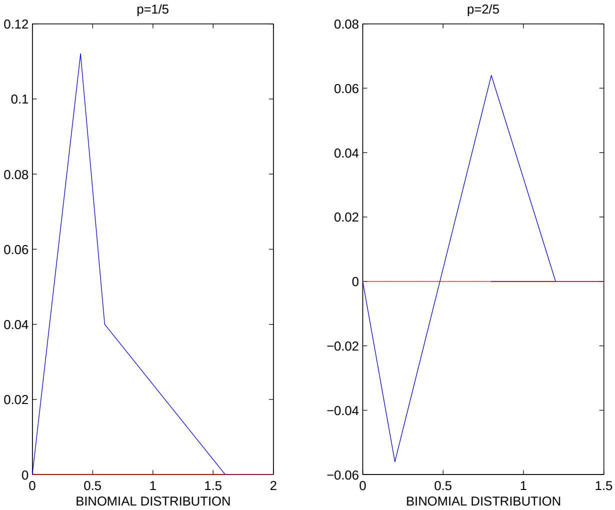
Assume that $X \in \mathbb{L}^1(\mathbb{R})$ *with* $\mathbb{E}[X] = 0$.

- X *is heavy on left* $\Longleftrightarrow \forall t \geqslant 0,\ L(t) \leqslant 1$,
- X *is heavy on right* $\Longleftrightarrow \forall t \leqslant 0,\ L(t) \leqslant 1$,
- X *is symmetric* $\Longleftrightarrow \forall t \in \mathbb{R},\ L(t) \leqslant 1$.

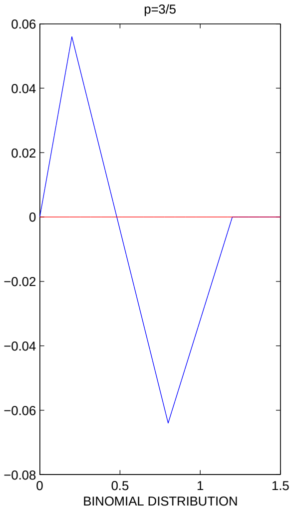
Centered Bernoulli $\mathcal{B}(p)$

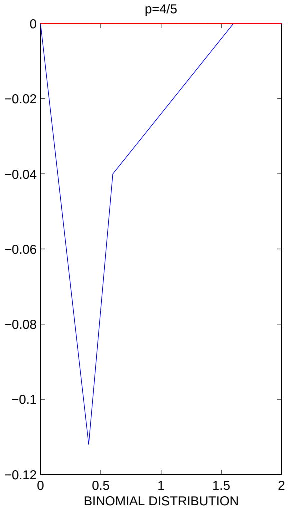


Centered Binomial $\mathcal{B}(2, p)$

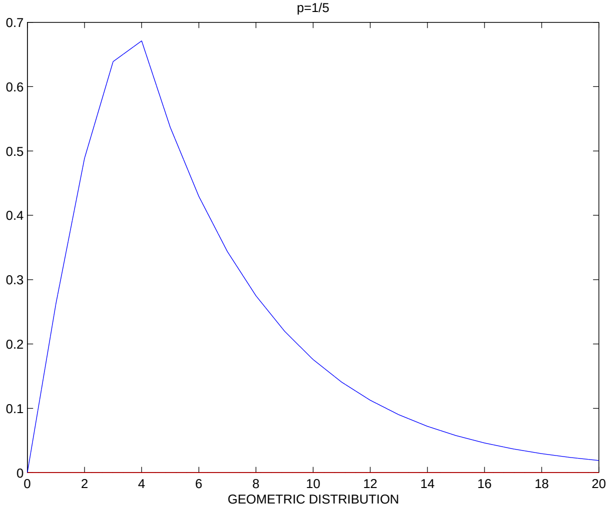


Centered Binomial $\mathcal{B}(2, p)$

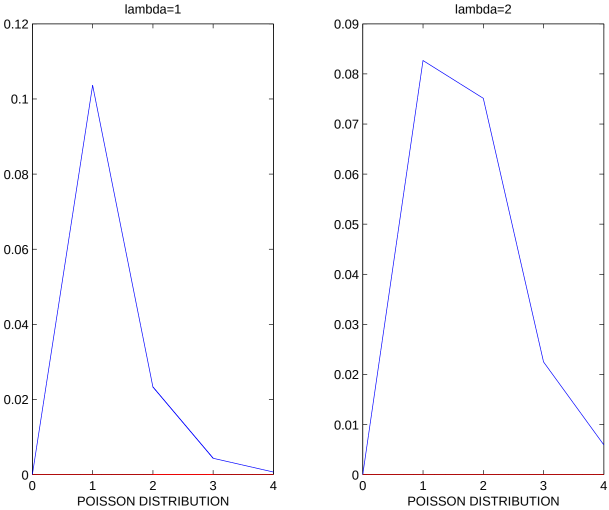




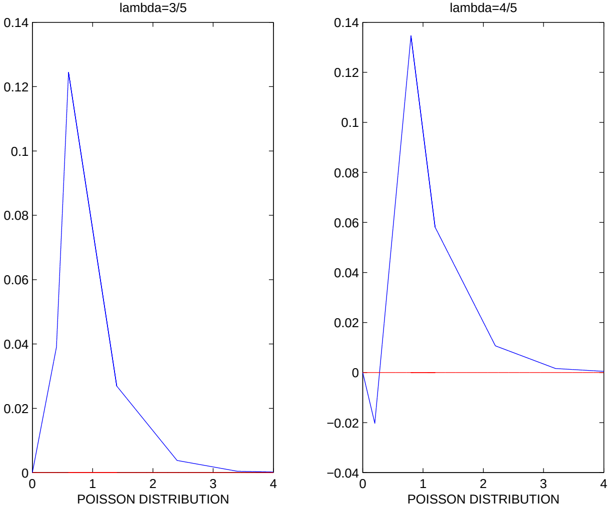
Centered Geometric $\mathcal{G}(p)$



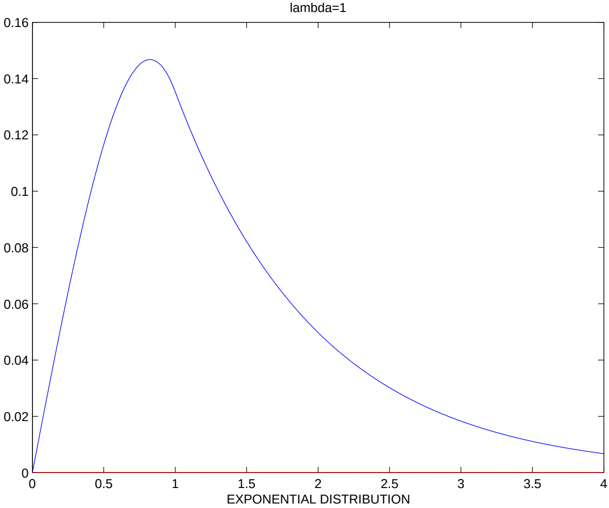
Centered Poisson $\mathcal{P}(\lambda)$



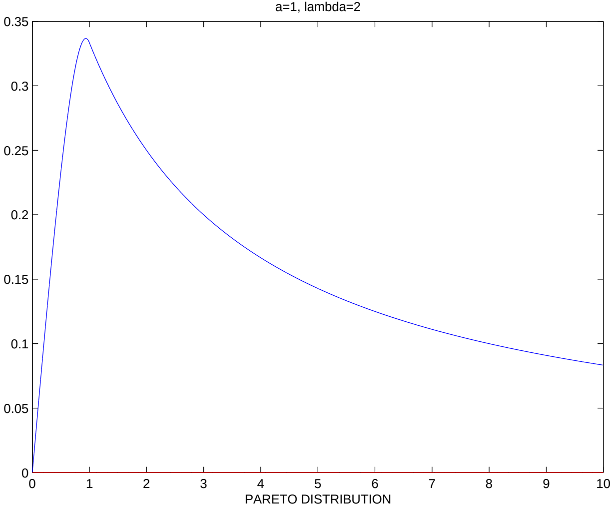
Centered Poisson $\mathcal{P}(\lambda)$



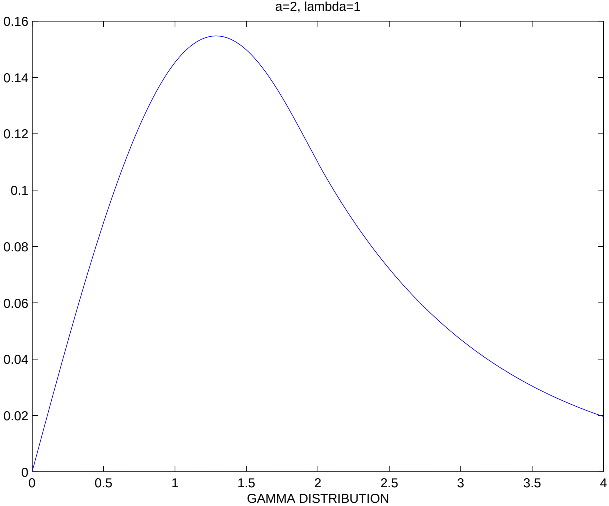
Centered Exponential $\mathcal{E}(\lambda)$



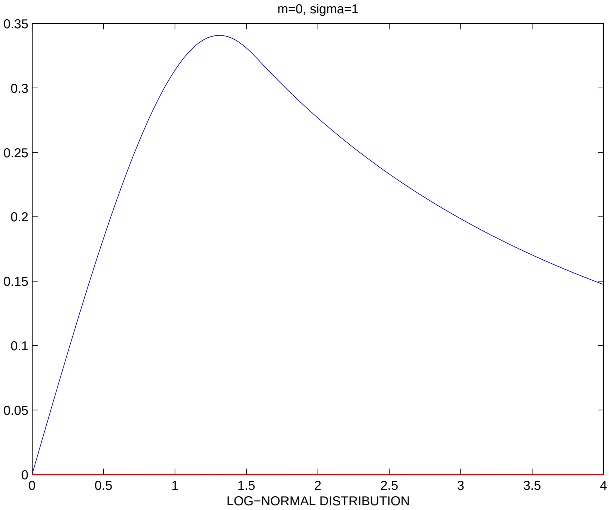
Centered Pareto $\mathcal{P}(a, \lambda)$



Centered Gamma $\mathcal{G}(a, \lambda)$



Centered Log-Normal $\mathcal{L}(m, \sigma^2)$



Outline

Martingales heavy on left or right

Definition. We shall say that (M_n) is **conditionally heavy on left** if, $\forall n \geqslant 1$ and $\forall a > 0$,

$$\mathbb{E}[T_a(\Delta M_n)|\mathcal{F}_{n-1}] \leqslant 0 \qquad \text{a.s.}$$

(M_n) is conditionally heavy on right if $(-M_n)$ is conditionally heavy on left.

Theorem (Bercu-Touati)

Assume that (M_n) is conditionally heavy on left. Then, $\forall x, y > 0$

$$\mathbb{P}(M_n \geqslant x, [M]_n \leqslant y) \leqslant \exp\left(-\frac{x^2}{2y}\right).$$

Self-normalized martingales

Theorem (Bercu-Touati)

Assume that (M_n) *is conditionally heavy on left. Then,* $\forall x, y > 0$ *and* $\forall a \geqslant 0$, $b > 0$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x\Big) \leqslant \sqrt{\mathbb{E}\Big[\exp\Big(-x^2\Big(ab + \frac{b^2}{2}[M]_n\Big)\Big)\Big]},$$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x, [M]_n \geqslant \frac{1}{y}\Big) \leqslant \exp\Big(-x^2\Big(ab + \frac{b^2}{2y}\Big)\Big).$$

Self-normalized martingales

Theorem (Bercu-Touati)

Assume that (M_n) is conditionally heavy on left. Then, $\forall x, y > 0$ and $\forall a \geqslant 0$, $b > 0$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x\Big) \leqslant \sqrt{\mathbb{E}\Big[\exp\Big(-x^2\Big(ab + \frac{b^2}{2}[M]_n\Big)\Big)\Big]},$$

$$\mathbb{P}\Big(\frac{M_n}{a + b[M]_n} \geqslant x, [M]_n \geqslant \frac{1}{y}\Big) \leqslant \exp\Big(-x^2\Big(ab + \frac{b^2}{2y}\Big)\Big).$$

Stable autoregressive process

Consider the stable autoregressive process

$$X_{n+1} = \theta X_n + \varepsilon_{n+1}, \qquad |\theta| < 1$$

where (ε_n) is iid $\mathcal{N}(0, \sigma^2)$, $\sigma^2 > 0$ and X_0 is independent of (ε_n) with $\mathcal{N}(0, \sigma^2/(1-\theta^2))$ distribution. Denote by $\widehat{\theta}_n$ and $\widetilde{\theta}_n$ the **least squares** and the **Yule-Walker** estimators of θ

$$\widehat{\theta}_n = \frac{\displaystyle\sum_{k=1}^{n} X_k X_{k-1}}{\displaystyle\sum_{k=1}^{n} X_{k-1}^2}, \qquad \widetilde{\theta}_n = \frac{\displaystyle\sum_{k=1}^{n} X_k X_{k-1}}{\displaystyle\sum_{k=0}^{n} X_k^2}.$$

$$a = \frac{\theta - \sqrt{\theta^2 + 8}}{4} \qquad \text{and} \qquad b = \frac{\theta + \sqrt{\theta^2 + 8}}{4}.$$

Theorem (Bercu-Gamboa-Rouault)

- $(\widehat{\theta}_n)$ *satisfies an* **LDP** *with rate function*

$$J(x) = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{1+\theta^2-2\theta x}{1-x^2}\right) & \textit{if } \ x \in [a,b], \\ \log|\theta - 2x| & \textit{otherwise.} \end{cases}$$

- $(\widetilde{\theta}_n)$ *satisfies an* **LDP** *with rate function*

$$I(x) = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{1+\theta^2-2\theta x}{1-x^2}\right) & \textit{if } \ x \in]-1,1[, \\ +\infty & \textit{otherwise.} \end{cases}$$

$$a = \frac{\theta - \sqrt{\theta^2 + 8}}{4} \qquad \text{and} \qquad b = \frac{\theta + \sqrt{\theta^2 + 8}}{4}.$$

Theorem (Bercu-Gamboa-Rouault)

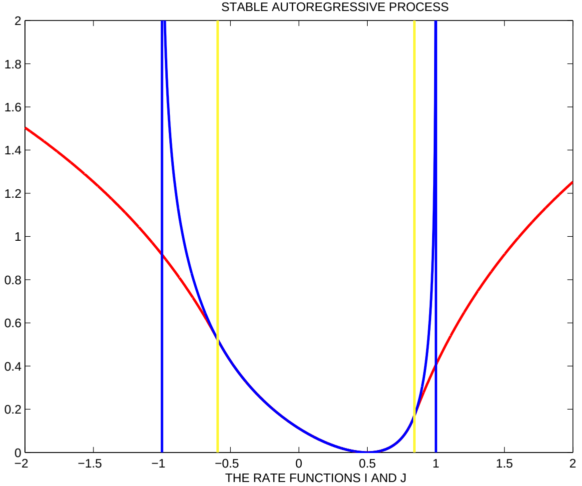
- $(\widehat{\theta}_n)$ *satisfies an* **LDP** *with rate function*

$$J(x) = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{1+\theta^2-2\theta x}{1-x^2}\right) & \text{if } x \in [a,b], \\ \log|\theta - 2x| & \text{otherwise.} \end{cases}$$

- $(\widetilde{\theta}_n)$ *satisfies an* **LDP** *with rate function*

$$I(x) = \begin{cases} \dfrac{1}{2}\log\left(\dfrac{1+\theta^2-2\theta x}{1-x^2}\right) & \text{if } x \in]-1,1[, \\ +\infty & \text{otherwise.} \end{cases}$$

Least squares and Yule-Walker



Theorem (Bercu-Touati)

Assume that X_0 is independent of (ε_n) with $\mathcal{N}(0, \tau^2)$ distribution where $\tau^2 \geqslant \sigma^2$. For all $\theta \in \mathbb{R}$, $n \geqslant 0$ and $x > 0$

$$\mathbb{P}(\widehat{\theta}_n - \theta \geqslant x) \leqslant 2 \exp\Big(-\frac{nx^2}{2(1+y_x)}\Big)$$

where y_x is the unique positive solution of

$$(1+y)\log(1+y) - y = x^2.$$

Remark. This inequality also holds for $\widetilde{\theta}_n$. In addition, for all $0 < x < 1/2$, $y_x < 2x$ so that

$$\mathbb{P}(\widehat{\theta}_n - \theta \geqslant x) \leqslant 2 \exp\Big(-\frac{nx^2}{2(1+2x)}\Big).$$