Sharp Tail Bounds for Functionals of Markov Chains
–
Journées de Probabilité - Lille 2008

Stéphan Clémençon

Telecom ParisTech
-
LTCI UMR CNRS/Telecom ParisTech No 5141

Joint work with Patrice Bertail (Modal'x - Université Paris X)

1. Overall purpose

2. A little Markov chain theory
   - Markov with regeneration times
   - General Harris Markov chains

3. Probabilistic study based on renewal theory
   - Limit behavior
   - First order results

4. The regenerative method
   - Sharper results
   - Tail bounds through the regenerative approach

Overall purpose

- **(Bennett, '62)** Let $Y_1, \ldots, Y_n$ be i.i.d. r.v.'s such that $\mathbb{E}[Y_i] = 0$, $\mathbb{E}[Y_i^2] = \sigma^2 < \infty$ and $|Y_i| \leq M$ a.s. .

$$\mathbb{P}\{\sum_{i=1}^{n} Y_i \geq x\} \leq \exp\left(-\frac{n\sigma^2}{M^2} H(\frac{Mx}{n\sigma^2})\right),$$

  where $H(x) = (1+x)\log(1+x) - x$.

- **(Fuk & Nagaev, '71)** Relaxing the boundedness constraint: $\forall M,\ x > 0$,

$$\mathbb{P}\{\sum_{i=1}^{n} Y_i \geq x\} \leq \exp\left(-\frac{n\sigma_M^2}{M^2} H(\frac{Mx}{n\sigma_M^2})\right) + n\mathbb{P}\{Y_1 > M\},$$

  with $\sigma_M^2 = \mathbb{E}[Y_i^2 \mathbb{I}\{|Y_i| \leq M\}]$.

- **Our goal:** extension to instantaneous functionals of a Markov chain, *i.e.* $Y_i = f(X_i)$, by means of the **regenerative method**.

Markov chains with (pseudo-) regeneration times

- **The Markov property:** $X = (X_n)_{n \in \mathbb{N}}$ is a chain with state space $(E, \mathcal{E})$, trans. prob. $\Pi(x, dy)$ and initial distr. ν iff $X_0 \sim \nu$ and

$$\mathbb{P}(X_{n+1} \in dx \mid X_0, \ldots, X_n) = \Pi(X_n, dx).$$

Markov chains with (pseudo-) regeneration times

- **The Markov property:** $X = (X_n)_{n\in\mathbb{N}}$ is a chain with state space $(E, \mathcal{E})$, trans. prob. $\Pi(x, dy)$ and initial distr. ν iff $X_0 \sim \nu$ and

$$\mathbb{P}(X_{n+1} \in dx \mid X_0, \ldots, X_n) = \Pi(X_n, dx).$$

- **Strong version:**
  - "Shift operator" θ: $X_{n+1} = X_n \circ \theta$,
  - $\mathcal{F}_n = \sigma(X_0, \ldots, X_n)$, τ stopp. time, $H = H(x_1, x_2, \ldots)$ bounded,

$$\mathbb{E}_\nu[H \circ \theta^\tau \mid \mathcal{F}_\tau] = \mathbb{E}_{X_\tau}[H] \text{ on } \{\tau < \infty\}.$$

A little Markov chain theory

- **Communication/ Stochastic stability**
  - **Irreducibility**: $\Phi/\ \forall B \in \mathcal{E}$,

$$\Phi(B) > 0 \Rightarrow \forall x,\ \sum_{n \geq 1} \Pi_n(x, B) > 0,$$

    $\Rightarrow$ existence of a maximal irreducibility measure Ψ.
  - **Periodicity**: X Ψ-irreduc., $\exists d' \geq 1,\ D_1, \ldots, D_{d'},\ \Psi(D_i) > 0$, $D_i \cap D_j = \emptyset$ if $i \neq j$ and
    1. $\Pi(x, D_{i+1}) = 1,\ x \in D_i$,
    2. $\Psi(\cup\{D_i\}^c) = 0$,

    $\Rightarrow$ period$= GCD\{d' \geq 1/\ 1.\ and\ 2.\}$, aperiodic when $= 1$.
  - **Positive recurrence**: $\exists!$ prob. $\mu/\ \mu(dy) = \int_{x \in E} \mu(dx)\Pi(x, dy)$ and

$$"\Pi_n(x, dy) = \mathbb{P}(X_n \in dy \mid X_0 = x) \rightarrow \mu(dy)" \text{ as } n \rightarrow \infty.$$

Regenerative Markov chains

- A set $A \in \mathcal{E}$ is an **accessible atom** if $\Psi(A) > 0$ and

$$\forall (x,y) \in A^2, \;\; \Pi(x,.) = \Pi(y,.).$$

- **"Dividing sample paths into regeneration cycles"**
  Suppose A is Harris recurrent. Let $\mathbb{E}_A[.] = \mathbb{E}[. \mid X_0 \in A]$ and

$$\begin{aligned} \tau_A(1) &= \tau_A = \inf\{k \geq 1 : \; X_k \in A\}, \\ \tau_A(j) &= \inf\{k \geq 1 + \tau_A(j-1) : \; X_k \in A\}, \; j > 1. \end{aligned}$$

Strong Markov $\Rightarrow \{\tau_A(j)\}_{j \geq 1}$ is a (possibly delayed) **renewal process**

$$\ldots, \underbrace{X_{1+\tau_A(1)}, \ldots, X_{\tau_A(2)}}, \ldots, \underbrace{X_{1+\tau_A(j)}, \;\ldots,\; X_{\tau_A(j+1)}}, \ldots$$

i.i.d. "regenerative segments" of random size

General Harris Markov chains - Pseudo-renewal

General Harris Markov chains - Pseudo-renewal

- A meas. set $S \subset E$ is **small** if $\exists m \geq 1$, $\delta > 0$, Φ proba. s.t. $\Phi(S) = 1$

$$\forall x \in E, \ \ \Pi_m(x, dy) \geq \delta \mathbb{I}_{\{x \in S\}} \Phi(dy).$$

- **The Nummelin technique** $(m = 1)$. Construct a bivar. chain (X, Y) with state sp. $E \times \{0, 1\}$ by **randomizing each time X_n hits S**:
  - if $x \notin S$, $\Pi^*((x, y), B \times \{1\}) = \delta \Pi(x, B)$,
  - if $x \in S$, then

$$\begin{aligned} \Pi^*((x,0), B \times \{1\}) &= \delta \frac{\Pi(x, B) - \delta \Phi(B)}{1 - \delta}, \\ \Pi^*((x,1), B \times \{1\}) &= \delta \Phi(B), \end{aligned}$$

$\Rightarrow S \times \{1\}$ is an atom for the *split chain*.

Stationary/asymptotic behavior

Stationary/asymptotic behavior

- Suppose X is Harris recurrent. Let A be an accessible atom.

The asymptotic behavior is ruled by the renewal process $\{\tau_A(j)\}_{j\geq 1}$.

Stationary/asymptotic behavior

- Suppose X is Harris recurrent. Let A be an accessible atom.

The asymptotic behavior is ruled by the renewal process $\{\tau_A(j)\}_{j\geq 1}$.

Theorem (Kac's theorem)

The chain X is positive recurrent iff $\alpha = \mathbb{E}_A[\tau_A] < \infty$.
If X is positive recurrent, then μ is the Pitman's **occupation measure**:

$$\mu(B) = \frac{1}{\mathbb{E}_A[\tau_A]}\mathbb{E}_A[\sum_{i=1}^{\tau_A} \mathbb{I}_{\{X_i \in B\}}].$$

Stationary/asymptotic behavior

- Suppose X is Harris recurrent. Let A be an accessible atom.

The asymptotic behavior is ruled by the renewal process $\{\tau_A(j)\}_{j\geq 1}$.

Theorem (Kac's theorem)

The chain X is positive recurrent iff $\alpha = \mathbb{E}_A[\tau_A] < \infty$.
If X is positive recurrent, then μ is the Pitman's **occupation measure**:

$$\mu(B) = \frac{1}{\mathbb{E}_A[\tau_A]}\mathbb{E}_A[\sum_{i=1}^{\tau_A} \mathbb{I}_{\{X_i \in B\}}].$$

- Let $f : E \to \mathbb{R}$ μ-integrable

$$\mu(f) = \int_{x\in E} f(x)\mu(dx) = \frac{1}{\mathbb{E}_A[\tau_A]}\mathbb{E}_A[\sum_{i=1}^{\tau_A} f(X_i)].$$

The regenerative method - First order results

The regenerative method - First order results

- **Law of Large Numbers** - Set $\alpha = \mathbb{E}_A[\tau_A]$. Count the nb of renewals up to time n: $l_n = \sum_{i=1}^{n} \mathbb{I}_{\{X_i \in A\}} \sim n/\alpha$. Write

$$\sum_{i=1}^{n} f(X_i) = \sum_{i=1}^{\tau_A} f(X_i) + \sum_{j=1}^{l_n - 1} S_j(f) + \sum_{i=1+\tau_A(l_n)}^{n} f(X_i),$$

with $S_j(f) = \sum_{i=1+\tau_A(j)}^{\tau_A(j+1)} f(X_i)$. The $S_j(f)$'s are **i.i.d. r.v.'s**:

$$\Rightarrow \hat{\mu}_n = \frac{1}{n} \sum_{i=1}^{n} f(X_i) \to \mu(f) = \alpha^{-1} \mathbb{E}[S_1(f)] \text{ as } n \to \infty.$$

The regenerative method - First order results

- **Law of Large Numbers** - Set $\alpha = \mathbb{E}_A[\tau_A]$. Count the nb of renewals up to time n: $l_n = \sum_{i=1}^n \mathbb{I}_{\{X_i \in A\}} \sim n/\alpha$. Write

$$\sum_{i=1}^n f(X_i) = \sum_{i=1}^{\tau_A} f(X_i) + \sum_{j=1}^{l_n - 1} S_j(f) + \sum_{i=1+\tau_A(l_n)}^{n} f(X_i),$$

with $S_j(f) = \sum_{i=1+\tau_A(j)}^{\tau_A(j+1)} f(X_i)$. The $S_j(f)$'s are **i.i.d. r.v.'s**:

$$\Rightarrow \hat{\mu}_n = \frac{1}{n} \sum_{i=1}^n f(X_i) \to \mu(f) = \alpha^{-1}\mathbb{E}[S_1(f)] \text{ as } n \to \infty.$$

- **Central Limit Theorem** - Suppose that $\mathbb{E}_A[(\sum_{i=1}^{\tau_A} f(X_i))^2] < \infty$,

$$\sqrt{n}(\hat{\mu}_n - \mu(f)) \Rightarrow \mathcal{N}(0, \sigma_f^2) \text{ as } n \to \infty,$$

with $\sigma_f^2 = \alpha^{-1}\mathbb{E}_A[(\sum_{i=1}^{\tau_A} \{f(X_i) - \mu(f)\})^2] \neq Var_\mu[f(X)]$.

Non asymptotic bounds and second order results

- Given a sample path of length n, form $l_n + 1$ blocks:

$$\mathcal{B}_0 = (X_1, \ldots, X_{\tau_A}), \; \mathcal{B}_1 = (X_{1+\tau_A(1)}, \ldots, X_{\tau_A(2)}), \; \ldots$$

with $l_n = \sum_{i \leq n} \mathbb{I}\{X_i \in A\}$.
In general, they are **not independent**! (their lengths sum up to n)

Non asymptotic bounds and second order results

- Given a sample path of length n, form $l_n + 1$ blocks:

$$\mathcal{B}_0 = (X_1, \ldots, X_{\tau_A}),\ \mathcal{B}_1 = (X_{1+\tau_A(1)}, \ldots, X_{\tau_A(2)}),\ \ldots$$

  with $l_n = \sum_{i \leq n} \mathbb{I}\{X_i \in A\}$.
  In general, they are **not independent**! (their lengths sum up to n)
- The method comprises **3 steps**. Let $s_j = \tau_A(j+1) - \tau_A(j)$.

Non asymptotic bounds and second order results

- Given a sample path of length n, form $l_n + 1$ blocks:

$$\mathcal{B}_0 = (X_1, \ldots, X_{\tau_A}),\ \mathcal{B}_1 = (X_{1+\tau_A(1)}, \ldots, X_{\tau_A(2)}),\ \ldots$$

  with $l_n = \sum_{i \leq n} \mathbb{I}\{X_i \in A\}$.
  In general, they are **not independent**! (their lengths sum up to n)
- The method comprises **3 steps**. Let $s_j = \tau_A(j+1) - \tau_A(j)$.
  1. **Partition** the space according to all possible ways for X to regenerate

$$U_{r,l,m} = \{\tau_A = r,\ \sum_{j=1}^{m-1} s_j = n - r - l,\ \tau_A(m+1) - \tau_A(m) > l\}.$$

Non asymptotic bounds and second order results

- Given a sample path of length n, form $l_n + 1$ blocks:

$$\mathcal{B}_0 = (X_1, \ldots, X_{\tau_A}),\ \mathcal{B}_1 = (X_{1+\tau_A(1)}, \ldots, X_{\tau_A(2)}),\ \ldots$$

with $l_n = \sum_{i \leq n} \mathbb{I}\{X_i \in A\}$.
In general, they are **not independent**! (their lengths sum up to n)

- The method comprises **3 steps**. Let $s_j = \tau_A(j+1) - \tau_A(j)$.
  1. **Partition** the space according to all possible ways for X to regenerate

$$U_{r,l,m} = \{\tau_A = r,\ \sum_{j=1}^{m-1} s_j = n - r - l,\ \tau_A(m+1) - \tau_A(m) > l\}.$$

  2. On each subset $U_{r,l,m}$, establish the required result for the i.i.d. sequence of 1-**lattice random vectors** $\{(S_j(f), L_j)\}_{j \geq 1}$.

Non asymptotic bounds and second order results

- Given a sample path of length n, form $l_n + 1$ blocks:

$$\mathcal{B}_0 = (X_1, \ldots, X_{\tau_A}),\ \mathcal{B}_1 = (X_{1+\tau_A(1)}, \ldots, X_{\tau_A(2)}),\ \ldots$$

  with $l_n = \sum_{i \leq n} \mathbb{I}\{X_i \in A\}$.
  In general, they are **not independent**! (their lengths sum up to n)
- The method comprises **3 steps**. Let $s_j = \tau_A(j+1) - \tau_A(j)$.
  1. **Partition** the space according to all possible ways for X to regenerate

$$U_{r,l,m} = \{\tau_A = r,\ \sum_{j=1}^{m-1} s_j = n - r - l,\ \tau_A(m+1) - \tau_A(m) > l\}.$$

  2. On each subset $U_{r,l,m}$, establish the required result for the i.i.d. sequence of 1-**lattice random vectors** $\{(S_j(f), L_j)\}_{j\geq 1}$.
  3. **Sum up** the results obtained so as to **identify** the global bound/limit.

Assumptions and notations

Block-moment conditions

(i) $\alpha = \mathbb{E}_A[\tau_A] < \infty$ and $0 < \sigma_A^2 = \mathbb{E}_A[\tau_A^2] < \infty$,

(ii) $\mathbb{E}_A[S_A(f)] = 0$ and $0 < \sigma_f^2 = \mathbb{E}_A[S_A(f)^2] < \infty$.

Notations

$$\begin{aligned}
\tilde{\sigma}_f^2 &= Var_A(S_A(f)\mathbb{I}_{\{|S_A(f)|\leq M_1\}}),\\
\tilde{\sigma}_A^2 &= Var_A(\tau_A \mathbb{I}_{\{\tau_A \leq M_2\}}),\\
\tilde{\rho} &= (\tilde{\sigma}_f \tilde{\sigma}_A)^{-1} Cov_A(S_A(f)\mathbb{I}_{\{|S_A(f)|\leq M_1\}}, \tau_A \mathbb{I}_{\{\tau_A \leq M_2\}}),\\
\tilde{\sigma} &= \sqrt{\tilde{\sigma}_f^2 \tilde{\sigma}_A^2 / (\tilde{\sigma}_f^2 + \tilde{\sigma}_A^2)},
\end{aligned}$$

A probability inequality

Theorem (Bertail & Clémençon, 2007)

Under the previous assumptions, for any vector $M = (M_1, M_2) \in \mathbb{R}_+^{*2}$ of euclidian norm $\|M\| = (M_1^2 + M_2^2)^{1/2}$, there exists a constant C_M depending on σ_A^2 and σ_f^2 such that , for any $n \geq \alpha$,

$$\begin{aligned} \mathbb{P}_\nu(\sum_{j=1}^{l_n-1} S_j(f) \geq x) &\leq C_M \exp\left\{ -\frac{n(1+|\tilde{\rho}|)\tilde{\sigma}^2}{2\|M\|^2} H\left(\frac{\|M\|\sqrt{2}}{(1+|\tilde{\rho}|)\tilde{\sigma}\tilde{\sigma}_f} x\right)\right\} \\ &+ (n-1)\mathbb{P}_A(|S_A(f)| \geq M_1) + (n-1)\mathbb{P}_A(\tau_A \geq M_2), \end{aligned}$$

- For $A = E$, f bounded and M properly chosen, one recovers Bennett's inequality (1962).

Sketch of Proof - Step 1 "Truncation"

- Let

$$\tilde{S}_j(f) = S_j(f)\mathbb{I}\{|S_j(f)| \leq M_1\} \text{ and } \tilde{s}_j(f) = s_j(f)\mathbb{I}\{|s_j| \leq M_2\}$$

- From the union bound, get:

$$\begin{aligned} \mathbb{P}_\nu\left(\sum_{j=1}^{l_n-1} S_j(f) \geq x\right) &\leq \mathbb{P}_\nu\left(\left\{\sum_{j=1}^{l_n-1} \tilde{S}_j(f) \geq x\right\} \cap \{s_j = \tilde{s}_j, 1 \leq j \leq l_n - 1\}\right) \\ &+ (n-1)\mathbb{P}_A(\{|S_A(f)| > M_1\} \cup \{\tau_A > M_2\}) \end{aligned}$$

- Set

$$P_{\nu,n}(x) = \mathbb{P}_\nu\left(\left\{\sum_{j=1}^{l_n-1} \tilde{S}_j(f) \geq x\right\} \cap \{s_j = \tilde{s}_j, 1 \leq j \leq l_n - 1\}\right)$$

Step 2 "Partitioning"

- consider the collection of events: for $1 \leq r,\ l \leq n$ and $2 \leq m \leq n$,

$$U_{r,l,m} = \{\tau_A = r,\ \sum_{j=1}^{m} s_j = n - r - l,\ s_{m+1} > l\},$$

- By the formula of total probability and the strong Markov property:

$$P_{\nu,n}(x) = \sum_{m=2}^{n}\sum_{r=1}^{n}\sum_{l=1}^{n} \mathbb{P}_\nu(\tau_A = r) \times p_{r,l,m}(x) \times \mathbb{P}_A(\tau_A > l),$$

where $\tilde{\alpha} = \mathbb{E}[\tilde{s}_j]$

$$p_{r,l,m}(x) \quad = \quad \mathbb{P}(\sum_{1\leq j\leq m} \tilde{S}_j(f) \geq x,\ \sum_{1\leq j\leq m} \{\tilde{s}_j - \mathbb{E}[\tilde{s}_j]\} = n - r - l - m\tilde{\alpha})$$

Step 2 "Partitioning"

- Set $\lambda_{r,l,m} = (n - r - l - \widetilde{\alpha} m)/(\tilde{\sigma}_A \sqrt{m})$ and denote by $\tilde{p}_{r,l,m}(x)$ the probability

$$\mathbb{P}\left(\frac{\sum_{j=1}^{m}\{\tilde{S}_j(f) - \mathbb{E}[\tilde{S}_j(f)]\}}{\tilde{\sigma}_f \sqrt{m}} \geq \frac{x}{\tilde{\sigma}_f \sqrt{m}}, \frac{\sum_{j=1}^{m}\{\tilde{s}_j - \mathbb{E}[\tilde{s}_j]\}}{\tilde{\sigma}_A \sqrt{m}} = \lambda_{r,l,m}\right)$$

- Since $\mathbb{E}[\tilde{S}_j(f)] \leq 0$, we have $p_{r,l,m}(x) \leq \tilde{p}_{r,l,m}(x)$
- From limit theorems for 1-lattice random vectors (Dubinskaite, '82):

$$\tilde{p}_{r,l,m}(x) \sim m^{-1/2} h \int_{t=x}^{\infty} \phi_{\Sigma}(t/\sqrt{m}, \lambda_{r,l,m}) dt$$

with $\Sigma = Var((\tilde{\sigma}_f^{-1} \tilde{S}_j(f), \tilde{\sigma}_A^{-1} \tilde{s}_j))$

- Set $\tilde{\rho} = (\tilde{\sigma}_f \; \tilde{\sigma}_A)^{-1} Cov(\tilde{S}_j(f), \tilde{s}_j)$

Step 3 Exponential inequality for 1-lattice bounded r.v.'s

Lemma

Let $(S_j^*, L_j^*)_{1\leq j\leq m}$ be i.i.d. centered and square integrable bivariate r.v.'s with covariance $\Sigma = \binom{1\ \rho}{\rho\ 1}$. Assume that the L_j^*'s are lattice r.v.'s with minimal span $h > 0$ and there exists finite constants B_1 and B_2 such that $|S_j^*| \leq B_1$ and $|L_j^*| \leq B_2$ for $j = 1, ..., m$. Set $B^2 = B_1^2 + B_2^2$, then there exists a universal constant c s.t. for all $m \geq 1$ and $y \geq 0$,

$$\mathbb{P}(m^{-\frac{1}{2}} \sum_{j=1}^{m} S_j^* \geq y,\ m^{-\frac{1}{2}} \sum_{j=1}^{m} L_j^* = \lambda) \leq e^{\frac{3}{2}}(\frac{h}{2\pi\sqrt{m}} + 4c\frac{B}{m}) \exp\left\{-\frac{y^2+\lambda^2}{4B^2}\right\}$$

Step 3 - Application: control of $\tilde{p}_{r,l,m}(x)$

- In this case, the $L_j^{*\prime}s$ are lattice with minimal span $h = \tilde{\sigma}_A^{-1}$

$$\begin{aligned}\tilde{p}_{r,l,m}(x) &\leq e^{\frac{3}{2}}\left(\frac{1}{2\pi\tilde{\sigma}_A\sqrt{m}} + 8c\frac{||M||}{\tilde{\sigma}m}\right)\\ &\times e^{-\frac{m(1+|\tilde{\rho}|)\tilde{\sigma}^2}{2||M||^2}\left\{H\left(\frac{||M||x\sqrt{2}}{(1+|\tilde{\rho}|)\tilde{\sigma}_f\tilde{\sigma}m}\right)+H\left(\frac{||M|||\lambda_{r,l,m}|\sqrt{2}}{(1+|\tilde{\rho}|)\tilde{\sigma}\sqrt{m}}\right)\right\}}.\end{aligned}$$

Step 4 - Control of the sum

- We deduce that

$$P_{n,\nu}(x) \leq e^{3/2} \exp\left\{-\frac{n(1+|\tilde{\rho}|)\tilde{\sigma}^2}{2||M||^2} H\left(\frac{||M||x\sqrt{2}}{(1+|\tilde{\rho}|)\tilde{\sigma}_f\tilde{\sigma}n}\right)\right\}\Gamma_n, \qquad (1)$$

where

$$\Gamma_n = \sum_{m=2}^{n}\sum_{r=1}^{n}\sum_{l=1}^{n} \mathbb{P}_\nu\left(\tau_A = r\right)\mathbb{P}_A\left(\tau_A > l\right)\left(\frac{c_1}{\sqrt{m}} + \frac{c_2}{m}\right)\gamma_{r,l,m},$$

with

$$\gamma_{r,l,m} \quad \leq \quad \exp\left\{-\frac{\lambda^2_{r,l,m}}{2(1+|\tilde{\rho}| + \frac{|\lambda_{r,l,m}|||M||\sqrt{2})}{3\tilde{\sigma}\sqrt{m}}}\right\}.$$

Step 4 - Control of the sum

- Recall that $\lambda_{r,l,m} = (n - r - l - \tilde{\alpha} m)/\tilde{\sigma}_A \sqrt{m}$ and consider the subdivision

$$a_{n,m} = (n - \tilde{\alpha} m)/(\tilde{\sigma}_A \sqrt{m}), \text{ for } 1 \leq m \leq n.$$

- Split the sum into two terms

$$\begin{aligned} \sum_{r,l} \sum_{m} &= \sum_{r+l > \tilde{\sigma}_A} \sum_{m} + \sum_{r+l \leq \tilde{\sigma}_A} \sum_{m} \\ &\leq \textit{constant} + \textit{Riemann sum} \end{aligned}$$

- *etc.*

Proof of the Lemma (Sketch of)

- Talagrand '95 "The missing factor in Hoeffding's inequalities"
- Exponential change of probability measure: $\mathcal{S}_m = m^{-1/2} \sum_{i \leq n} S_j^*$, $\mathcal{L}_m = m^{-1/2} \sum_{i \leq n} L_j^*$ and $\psi_m(u) = \log \mathbb{E}[\exp\{< u, (\mathcal{S}_m,\ \mathcal{L}_m) >\}]$

$$d\mathbb{P}_{u,m} = \exp\{\langle u, (\mathcal{S}_m, \mathcal{L}_m)\rangle - \psi_m(u)\} d\mathbb{P}$$

- One has

$$\begin{aligned} \mathbb{P}(\mathcal{S}_m \geq y, \mathcal{L}_m = \lambda) &= e^{-\langle u,(y,\lambda)\rangle - \psi_m(u)} \\ &\times \mathbb{E}_{u,m}[e^{-\langle u,(\mathcal{S}_m - y, \mathcal{L}_m - \lambda)\rangle} \mathbb{I}_{\{\mathcal{S}_m \geq y, \mathcal{L}_m = \lambda\}}] \end{aligned}$$

- Choose $u = u^*$ such that $\nabla \psi_m(u^*) = (y, \lambda)$:

$$u^* = \arg \sup_{u \in \mathbb{R}_+ \times \mathbb{R}} \{< u, (y, \lambda) > - \psi_m(u)\}$$

Proof of the Lemma (Sketch of)

- The quantity $\mathbb{E}_{u,m}[e^{-\langle u,(\mathcal{S}_m-y,\mathcal{L}_m-\lambda)\rangle}\mathbb{I}_{\{\mathcal{S}_m\geq y,\mathcal{L}_m=\lambda\}}]$ is bounded by

$$u_1\int_{s=0}^{\infty} e^{-u_1 s}\mathbb{P}_{u,m}(S_m-y\geq 0, L_m-\lambda=0) \\ -\quad \mathbb{P}_{u,m}(S_m-y\geq s, L_m-\lambda=0)ds$$

- Use a local Berry-Esseen bound (Bolthausen '80)

$$|\mathbb{P}_{u,m}(\mathcal{S}_m-y\geq s,\mathcal{L}_m-\lambda=0)-\frac{h}{\sqrt{m}}\int_s^{\infty}\phi_{W_u}(t,0)dt|\leq C_u m^{-1},$$

where $W_u = Var_{u,m}[(\mathcal{S}_m,\mathcal{L}_m)]$ and ϕ_{W_u} is the density of the bivariate gaussian distribution with covariance matrix W_u