Almost sure convergence rates in martingale convergence theorems

Michael A. Högele

joint work with

Luisa F. Estrada (University of Warwick, UK)

Alexander Steinicke (Montanuniversität Leoben, Austria)

LiBERA Online seminar
24/03/2025

I. Motivation: The Polya Urn process

- Urn with N balls
- B of them are **black**, and $N - B$ of them are **white**.

The game:

1. Pick a ball Z from the urn
2. If Z is

$$\begin{cases} \textbf{black}: & \text{put it } \textbf{back} \,\&\, \textbf{another black} \text{ ball into the urn} \\ \textbf{white}: & \text{put it } \textbf{back} \,\&\, \textbf{another white} \text{ ball into the urn} \end{cases}$$

3. Repeat step 1.

What about the (random) proportion of the black balls on the long run?

Formally:

- Urn with $\mathbf{N}$ balls
- $\mathbf{B}$ of them are **black**, and $\mathbf{N}-\mathbf{B}$ of them are **white**.

- Step 0: Initial value

$$\mathbf{X_0} := \frac{\mathbf{B}}{\mathbf{N}}$$

- Step 1: Given $\mathbf{X_0}$ sample $\mathbf{Y_1} \sim \mathcal{B}_{\mathbf{X_0}}$

$$\mathbf{X_1} := \frac{\mathbf{B}+\mathbf{Y_1}}{\mathbf{N}+\mathbf{1}}$$

- Step n: Given $\mathbf{X_{n-1}}$ sample $\mathbf{X_n} \sim \mathcal{B}_{\mathbf{X_{n-1}}}$

$$\mathbf{X_n} := \frac{\mathbf{B}+\sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathbf{Y_i}}{\mathbf{N}+\mathbf{n}}$$

Polya urns have “balanced” increments (= they are “martingales”)

$$\mathbb{E}[(\mathbf{X_{n+1}} - \mathbf{X_n}) \mid \sigma(\mathbf{X_n}, \ldots, \mathbf{X_1})] = \mathbf{0} \qquad \text{a.s.}$$

Easy check

$$(\mathbf{N} + \mathbf{n} + \mathbf{1})\mathbb{E}[\mathbf{X_{n+1}} - \mathbf{X_n} \mid \sigma(\mathbf{X_n}, \ldots, \mathbf{X_1})]$$

$$= \mathbb{E}[(\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n+1}} \mathbf{Y_i}) - \frac{\mathbf{N+n+1}}{\mathbf{N+n}}(\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{Y_i}) \mid \sigma(\mathbf{X_n}, \ldots, \mathbf{X_1})]$$

$$= \mathbb{E}[(\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n+1}} \mathbf{Y_i}) - \frac{\mathbf{N+n}}{\mathbf{N+n}}(\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{Y_i}) - \frac{\mathbf{1}}{\mathbf{N+n}}(\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{Y_i}) \mid \sigma(\mathbf{X_n}, \ldots, \mathbf{X_1}$$

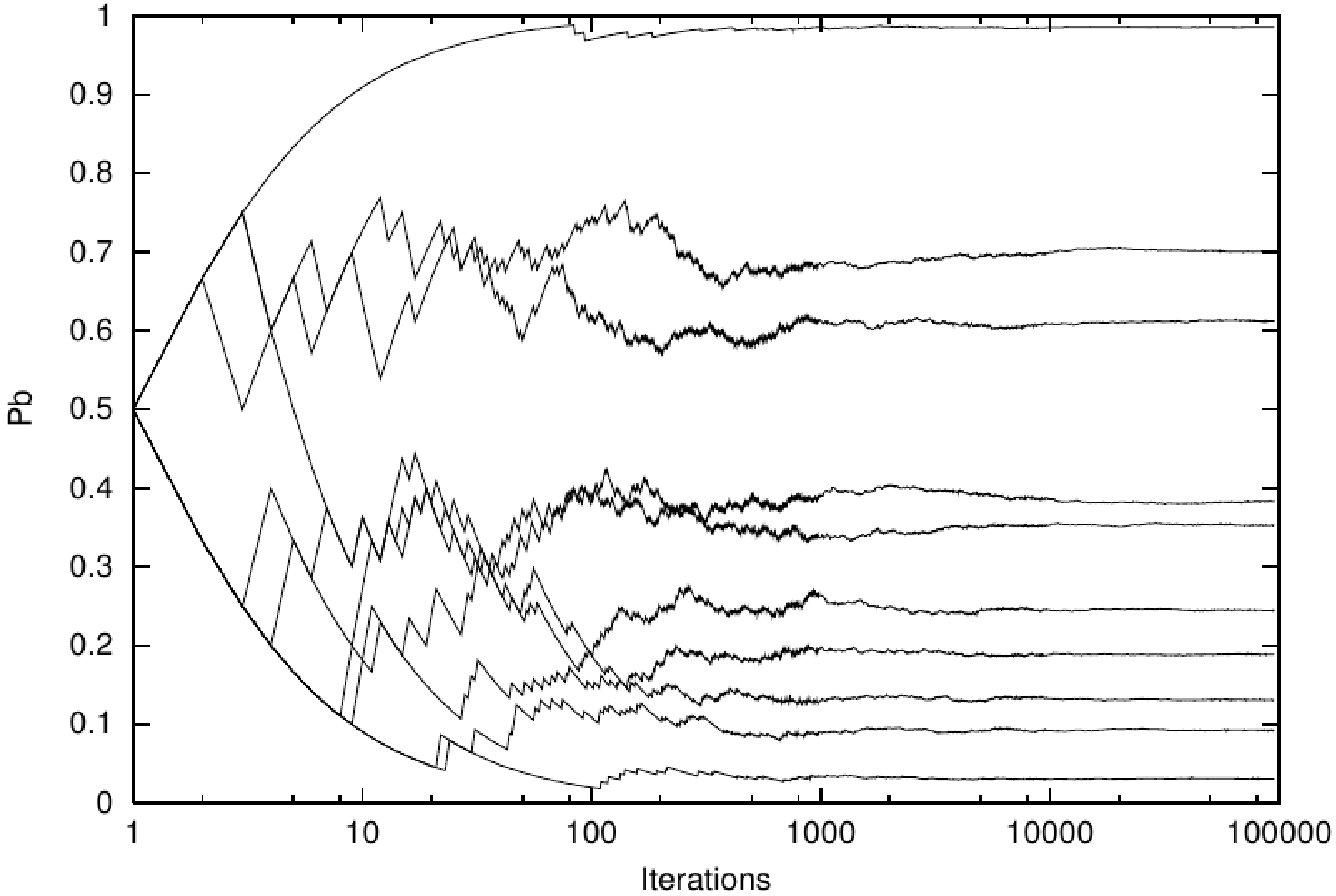
$$= \mathbb{E}[(\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n+1}} \mathbf{Y_i}) - (\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{Y_i}) - \mathbf{X_n} \mid \sigma(\mathbf{X_n}, \ldots,)]$$

$$= \mathbb{E}[\underbrace{\mathbf{Y_{n+1}}}_{\sim \mathcal{B}_{\mathbf{X_n}}} \mid \sigma(\mathbf{X_n}, \ldots, \mathbf{X_1})] - \mathbf{X_n} = \mathbf{X_n} - \mathbf{X_n} = \mathbf{0}.$$

Moreover their **increments decrease a.s. like $\mathbf{\frac{1}{n}}$**

$$
\begin{aligned}
|X_n - X_{n-1}| &= \left| \frac{S_n + B}{n + N} - \frac{S_{n-1} + B}{n - 1 + N} \right| \\
&= \frac{1}{n + N} \left| S_n + B - \left(1 + \frac{1}{n - 1 + N}\right) (S_{n-1} + B) \right| \\
&= \frac{1}{n + N} \left| Y_n + \frac{1}{n - 1 + N} (S_{n-1} + B) \right| \leqslant \frac{2}{n + N}
\end{aligned}
$$

Hence $\mathbf{X_n}$ should converge a.s. ... to a random variable $\mathbf{X_\infty}$

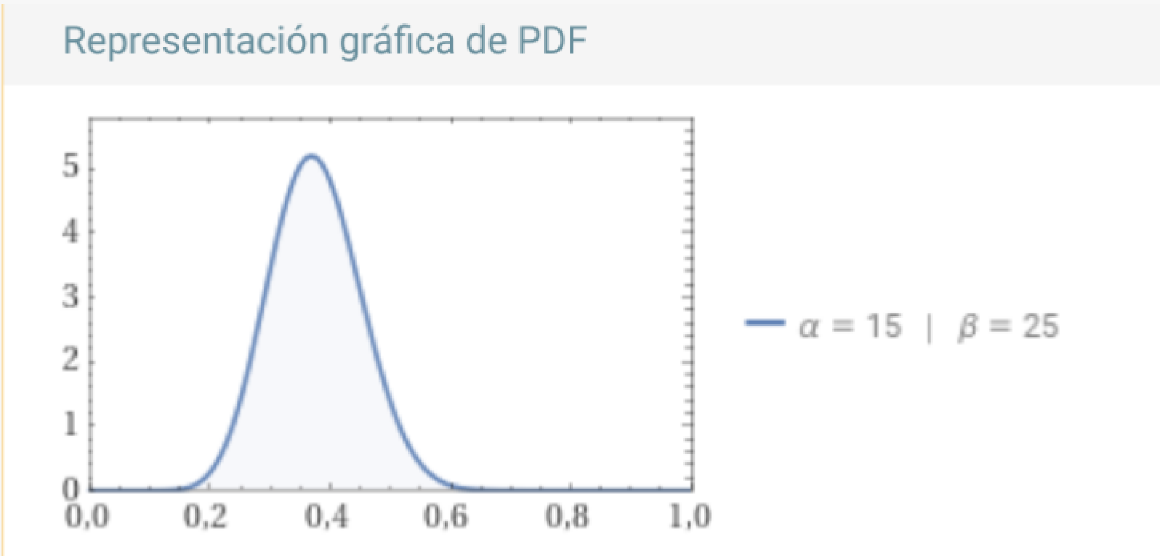


The **limiting distribution** is known

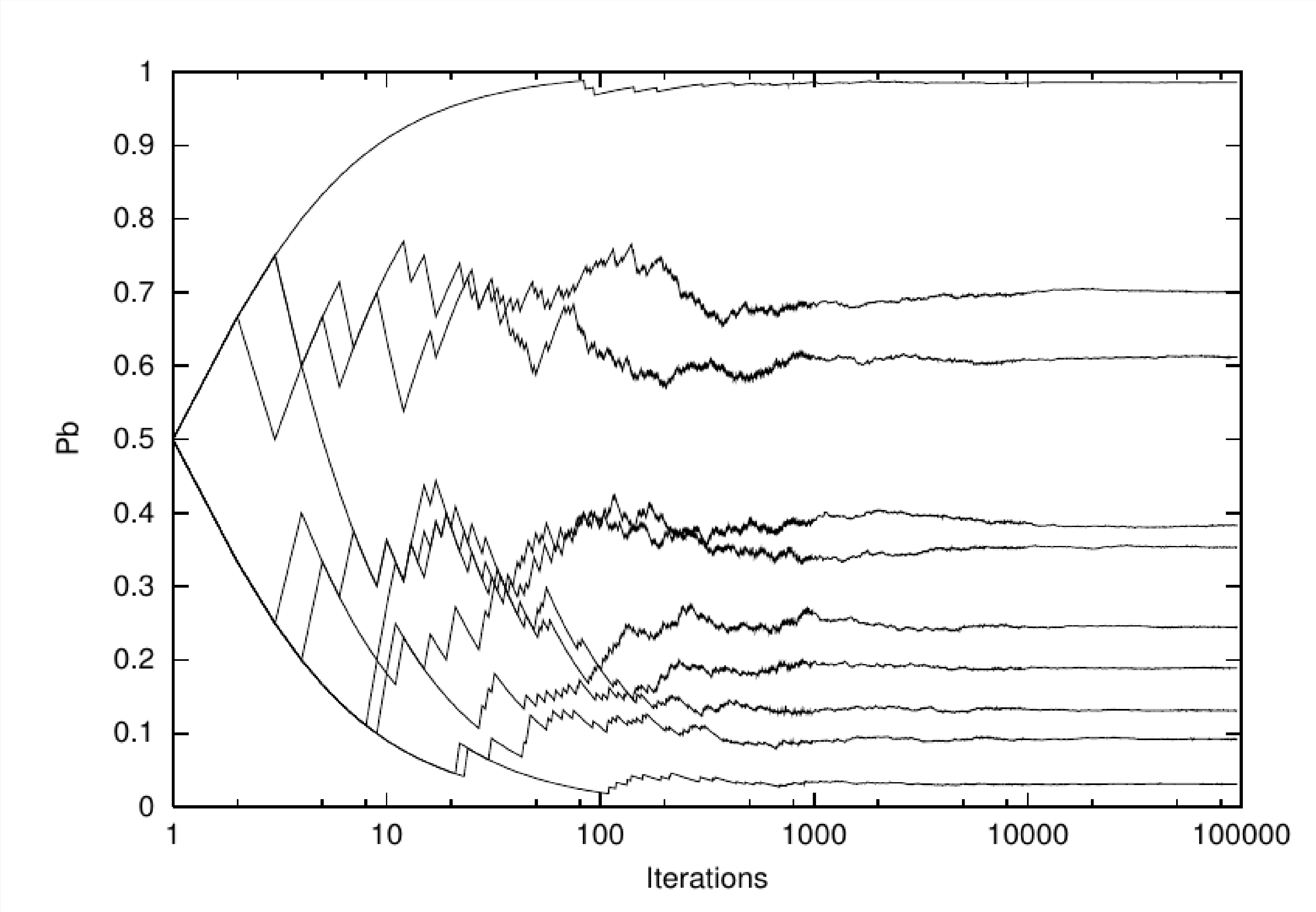
$$\mathbf{X_n} \to \mathbf{X_\infty} \sim \text{Beta}(\mathbf{B}, \mathbf{N} - \mathbf{B})$$

and the it depends on the initial distribution $\mathbf{B}$ vs $\mathbf{N} - \mathbf{B}$

$$\mathbf{f_{X_\infty}(x) = C \cdot x^{B-1} \cdot (1 - x)^{N-B-1}}, \qquad \mathbf{x} \in [\mathbf{0}, \mathbf{1}]$$



Note that its mean is given by the initial proportion $\mathbb{E}[\mathbf{X_\infty}] = \frac{\mathbf{B}}{\mathbf{N}}$
However, unclear **how fast converges** $\text{X}_\text{n} \to \text{X}_\infty$**.**



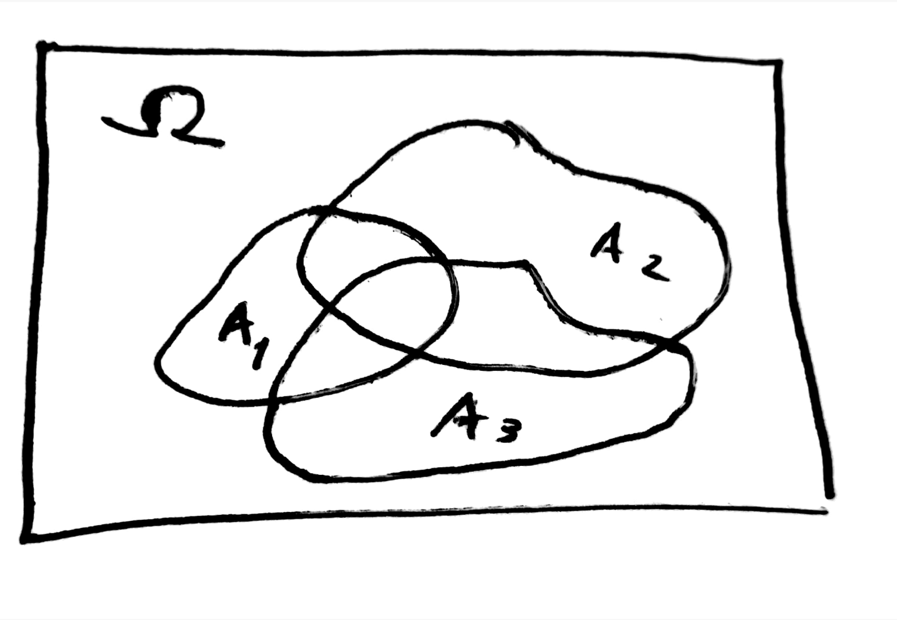
¿How fast does $X_n \to X_\infty$ converge ω-wise?

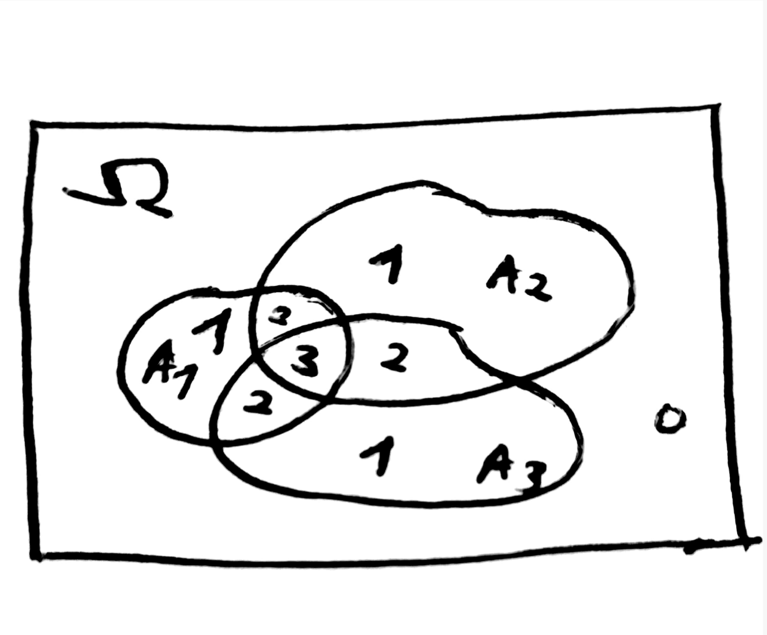
II. A device to prove a.s. convergence: the Borel-Cantelli Lemma

- $(\boldsymbol{\Omega}, \mathcal{A}, \mathbb{P})$ probability space
- $(\mathbf{A_n})_{\mathbf{n} \in \mathbb{N}}$, $\mathbf{A_n} \in \mathcal{A}$ family of events

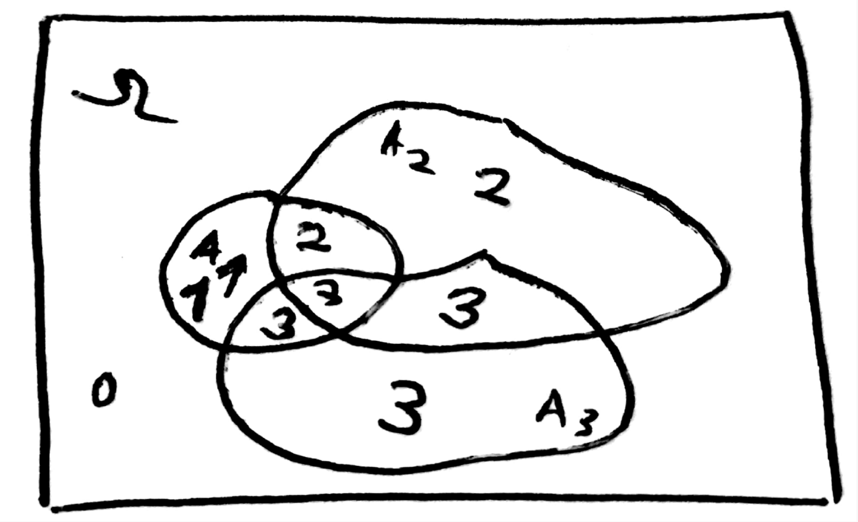
$$\mathcal{O}(\omega) := \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbf{1}(\mathbf{A_n})(\omega)$$ “overlap count”

$$\mathcal{M}(\omega) := \max\{n \in \mathbb{N} \mid \omega \in A_n\}$$ “last index”





“overlap count” $\mathcal{O}(\omega)$

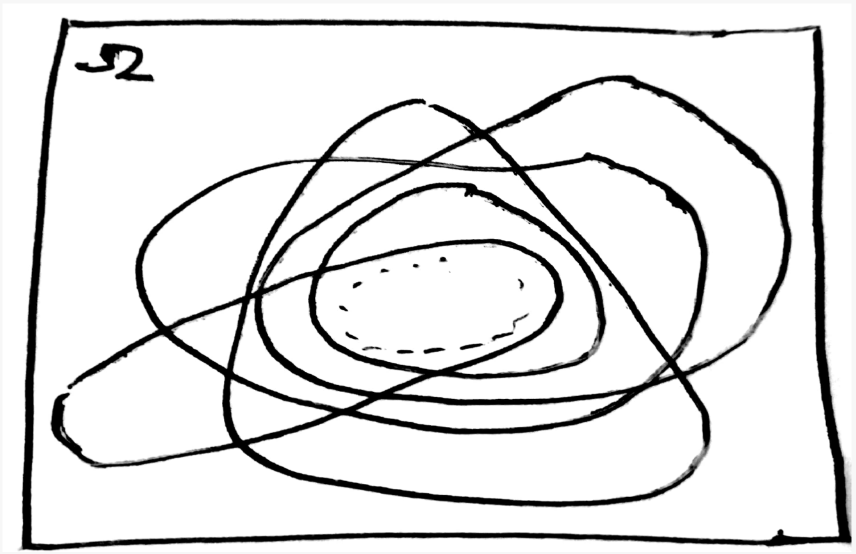


“last index” $\mathcal{M}(\omega)$

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$ probability space
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$, $\mathbf{A_n} \in \mathcal{A}$ family of events

$$\mathcal{O}(\omega) := \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbf{1}(\mathbf{A_n})(\omega) \qquad \text{“overlap count”}$$

$$\limsup_{\mathbf{n}\to\infty} \mathbf{A_n} = \{\omega \in \Omega \mid \mathcal{O}(\omega) = \infty\} \qquad \text{“infinite overlap”}$$



The first Borel-Cantelli lemma (1909 / 1917)

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$ probability space
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$, $\mathbf{A_n} \in \mathcal{A}$ family of events

Then

$$\sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$$

implies

$$\mathbb{P}(\limsup_{\mathbf{n}\to\infty} \mathbf{A_n}) = \mathbf{0}.$$

The first Borel-Cantelli lemma (1909 / 1917)

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$ probability space
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$, $\mathbf{A_n} \in \mathcal{A}$ family of events

Then

$$\sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$$

implies

$$\mathbb{P}(\limsup_{\mathbf{n}\to\infty} \mathbf{A_n}) = \mathbf{0}.$$

That is,

$$\mathcal{O} < \infty \qquad \mathbb{P} - \text{almost surely.}$$

"The overlap depth is finite a.s."

Proof 1: $\mathcal{O} < \infty$ $\mathbb{P}-$a.s.

For all $\mathbf{N} \in \mathbb{N}$

$$\mathbb{P}(\limsup_{\mathbf{n}\to\infty} \mathbf{A_n}) = \mathbb{P}(\bigcap_{\mathbf{n}\in\mathbb{N}} \bigcup_{\mathbf{m}\geqslant\mathbf{n}} \mathbf{A_m}) \leqslant \mathbb{P}(\bigcap_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \bigcup_{\mathbf{m}\geqslant\mathbf{n}} \mathbf{A_m}) = \mathbb{P}(\bigcup_{\mathbf{m}=\mathbf{N}}^{\infty} \mathbf{A_m})$$

Proof 1: $\mathcal{O} < \infty$ $\mathbb{P}-$a.s.

For all $\mathbf{N} \in \mathbb{N}$

$$\mathbb{P}(\limsup_{\mathbf{n}\to\infty} \mathbf{A_n}) = \mathbb{P}(\bigcap_{\mathbf{n}\in\mathbb{N}} \bigcup_{\mathbf{m}\geqslant\mathbf{n}} \mathbf{A_m}) \leqslant \mathbb{P}(\bigcap_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \bigcup_{\mathbf{m}\geqslant\mathbf{n}} \mathbf{A_m}) = \mathbb{P}(\bigcup_{\mathbf{m}=\mathbf{N}}^{\infty} \mathbf{A_m}),$$

and $\mathbb{P}(\bigcup_{\mathbf{m}=\mathbf{N}}^{\infty} \mathbf{A_m}) \leqslant \sum_{\mathbf{m}=\mathbf{N}}^{\infty} \mathbb{P}(\mathbf{A_m}) < \infty$ by hypothesis.

$$\Longrightarrow \quad \mathbf{0} \leqslant \mathbb{P}(\limsup_{\mathbf{n}\to\infty} \mathbf{A_n}) \leqslant \sum_{\mathbf{m}=\mathbf{N}}^{\infty} \mathbb{P}(\mathbf{A_m}) \quad \text{for all } \mathbf{N} \in \mathbb{N}$$

$$\Longrightarrow \quad \mathbf{0} \leqslant \mathbb{P}(\limsup_{\mathbf{n}\to\infty} \mathbf{A_n}) \leqslant \lim_{\mathbf{N}\to\infty} \sum_{\mathbf{m}=\mathbf{N}}^{\infty} \mathbb{P}(\mathbf{A_m}) = \mathbf{0}$$

Proof 2: $\mathbb{E}[\mathcal{O}] < \infty$

By monotone convergence (Beppo-Levi)

$$\begin{aligned}
\mathbb{E}[\mathcal{O}] &= \mathbb{E}[\lim_{\mathbf{N}\to\infty} \sum_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \mathbf{1}_{\mathbf{A_n}}] \\
&= \lim_{\mathbf{N}\to\infty} \mathbb{E}[\sum_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \mathbf{1}_{\mathbf{A_n}}] \\
&= \lim_{\mathbf{N}\to\infty} \sum_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \mathbb{E}[\mathbf{1}_{\mathbf{A_n}}] \\
&= \lim_{\mathbf{N}\to\infty} \sum_{\mathbf{n}=\mathbf{1}}^{\mathbf{N}} \mathbb{P}(\mathbf{A_n}) = \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty \qquad \text{by hypothesis.}
\end{aligned}$$

This **implies** $\mathcal{O} < \infty$ with probability 1.

Observations:

- The distribution of $\mathcal{O}$ is well-known as **Schuette-Nesbitt formula** (Gerber 1979)

$$\mathbb{P}(\mathcal{O} = \mathbf{k}) = \sum_{\substack{\mathbf{J} \subset \{\mathbf{1},\ldots,\mathbf{N}\} \\ |\mathbf{J}| = \mathbf{k}}} \mathbb{P}(\bigcap_{\mathbf{j} \in \mathbf{J}} \mathbf{A_j})$$

  The **probabilities on the right-hand side** are **rarely at hand**.

- However: many applications with $\mathbb{P}(\mathbf{A_n}) \searrow \mathbf{0}$ **much faster, than merely summable**, **with unknown** $\mathbb{P}(\bigcap_{\mathbf{j} \in \mathbf{J}} \mathbf{A_j})$!!

- Seems natural **to translate the rate** $\mathbb{P}(\mathbf{A_n}) \searrow \mathbf{0}$ into **higher moments of** $\mathcal{O}$ **(and** $\mathcal{M}$**)**.

Example:

$$\mathbb{P}(\mathbf{A_n}) \leqslant \frac{\mathbf{1}}{\mathbf{n^q}}, \qquad \text{given} \quad \mathbf{q} > \mathbf{1}$$

Hence

$$\mathbb{E}[\mathcal{O}] = \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbb{P}(\mathbf{A_n}) \leqslant \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \frac{\mathbf{1}}{\mathbf{n^q}} < \infty.$$

Note: **larger values of $\mathbf{q}$ yield smaller values of** $\mathbb{E}[\mathcal{O}]$.

$\longrightarrow$ Instead: is such a relation also **true under the expectation**?
$\longrightarrow$ That is, what about **higher moments** of $\mathcal{O}$?

For instance, given $\mathbf{q} = \mathbf{5}$, for which $\mathbf{p} > \mathbf{0}$ do we get

$$\mathbb{E}[\mathcal{O}^{\mathbf{1}+\mathbf{p}}] < \infty \qquad ?$$

Questions:

i) Given **only the rate of convergence** $\mathbb{P}(\mathbf{A_n}) \searrow \mathbf{0}$ as $\mathbf{n} \to \infty$,

what can be said about **higher moments** of the **overlap**

$$\mathcal{O}(\omega) := \sum_{\mathbf{n}} \mathbf{1}(\mathbf{A_n}) \qquad ?$$

ii) How can the results of i) be improved by the **monotonicity** (nestedness) of the events $\mathbf{A_n} \supset \mathbf{A_{n+1}}$?

iii) How about **higher moments** of the **(random) last index**

$$\mathcal{M}(\omega) := \max\{\mathbf{n} \in \mathbb{N} \mid \omega \in \mathbf{A_n}\} \qquad ?$$

Lemma: Borel-Cantelli moment equation for nested events [1] [2]:

- $(\boldsymbol{\Omega}, \mathcal{A}, \mathbb{P})$
- $(\mathbf{A_n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ **nested events**: $\mathbf{A_n} \supset \mathbf{A_{n+1}}, \quad \mathbf{n} \geqslant \mathbf{n_0}$
- $\mathbf{a} = (\mathbf{a_n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ positive $\&$ nondecreasing

Then we have

$$\mathbb{E}[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{O}_{\mathbf{n_0}})] = \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{a_n} \cdot \mathbb{P}(\mathbf{A_n})$$

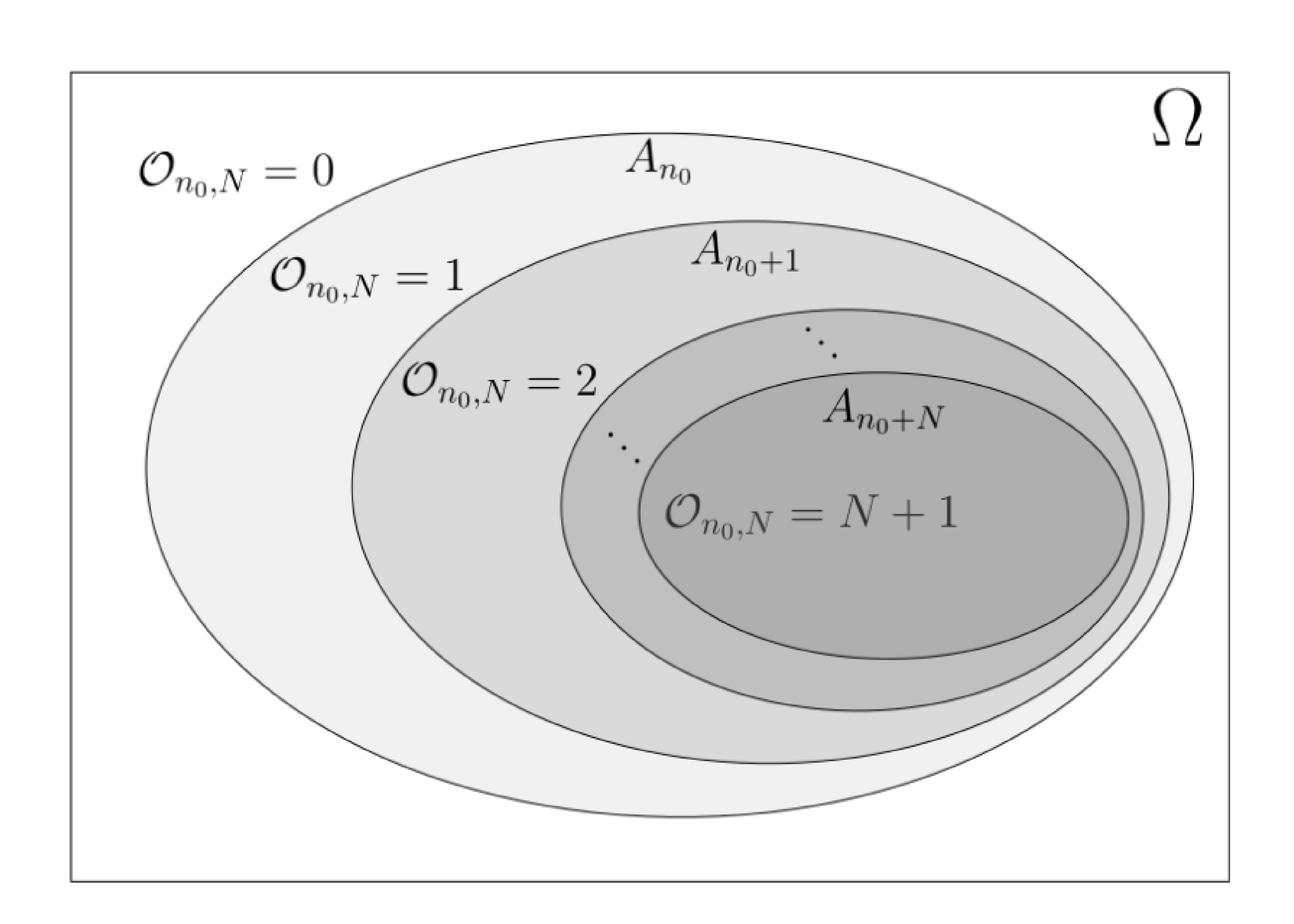
for

$$\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathbf{N}) := \sum_{\mathbf{n}=\mathbf{0}}^{\mathbf{N}-\mathbf{1}} \mathbf{a}_{\mathbf{n}+\mathbf{n_0}}, \qquad \mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathbf{0}) = \mathbf{0}$$

[1] Luisa F. Estrada, Michael A. Högele: Moment estimates in the first Borel-Cantelli Lemma with applications to mean deviation frequencies. Statistics and Probability Letters 190 (2022) 109636, https://doi.org/10.1016/j.spl.2022.109636

[2] Luisa F. Estrada, Michael A. Högele, Alexander Steinicke: On the tradeoff between almost sure error tolerance versus mean deviation frequency in martingale convergence, https://arxiv.org/abs/2310.09055

Proof:



Ω
$\mathcal{O}_{n_0,N} = 0$
A_{n_0}
$\mathcal{O}_{n_0,N} = 1$
A_{n_0+1}
$\mathcal{O}_{n_0,N} = 2$
A_{n_0+N}
$\mathcal{O}_{n_0,N} = N + 1$

- $\{\mathcal{O}_{\mathbf{n_0},\mathbf{N}} = \mathbf{0}\} = \mathbf{A}^{\mathbf{c}}_{\mathbf{n_0}}$

- nestedness: $\{\mathcal{O}_{\mathbf{n_0},\mathbf{N}} = \mathbf{k}\} = \mathbf{A}_{\mathbf{n_0}+\mathbf{k}-\mathbf{1}} \setminus \mathbf{A}_{\mathbf{n_0}+\mathbf{k}}$ for $\mathbf{k} = \mathbf{1}, \dots, \mathbf{N}$

- $\{\mathcal{O}_{\mathbf{n_0},\mathbf{N}} = \mathbf{N}+\mathbf{1}\} = \mathbf{A}_{\mathbf{n_0}+\mathbf{N}}$

- For $\mathbf{p_k} = \mathbb{P}(\mathbf{A_k})$ we have

$$\mathbb{E}[\mathcal{S}_{a,n_0}(\mathcal{O}_{n_0,N})]$$
$$= \mathcal{S}_{a,n_0}(0)\mathbb{P}(\Omega \setminus A_{n_0}) + \sum_{k=1}^{N} \mathcal{S}_{a,n_0}(k)\mathbb{P}(\mathcal{O}_{n_0,N} = k) + \mathcal{S}_{a,n_0}(N+1)\mathbb{P}(A_{N+n_0})$$

Summation by parts:

$$\sum_{k=0}^{N} f_k g_k = f_N \sum_{k=0}^{N} g_k - \sum_{j=0}^{N-1} (f_{j+1} - f_j) \sum_{\ell=0}^{j} g_k.$$

$$f_k = p_{n_0+k}, \qquad g_k = a_{n_0+k} \qquad p_k = \mathbb{P}(A_k)$$

$$\sum_{k=0}^{N} a_{n_0+k} p_{n_0+k} = p_{n_0+N} \sum_{k=0}^{N} a_{n_0+k} + \sum_{j=0}^{N-1} (p_{n_0+j} - p_{n_0+j+1}) \sum_{\ell=0}^{j} a_{n_0+\ell}.$$

$$
\begin{aligned}
&\sum_{k=0}^{N} a_{n_0+k}\mathbb{P}(A_{n_0+k}) \\
&= \mathbb{P}(A_{n_0+N})\sum_{k=0}^{N} a_{n_0+k} + \sum_{j=0}^{N-1}(\mathbb{P}(A_{n_0+j}) - \mathbb{P}(A_{n_0+j+1}))\sum_{\ell=0}^{j} a_{n_0+\ell} \\
&= \mathbb{P}(A_{n_0+N})\sum_{k=0}^{N} a_{n_0+k} + \sum_{j=0}^{N-1}\Big(\sum_{\ell=0}^{j} a_{n_0+\ell}\Big)\mathbb{P}(\mathcal{O}_{n_0,N} = j+1) \\
&= \mathbb{P}(A_{n_0+N})\sum_{k=0}^{N} a_{n_0+k} + \sum_{j=1}^{N}\Big(\sum_{\ell=0}^{j-1} a_{n_0+\ell}\Big)\mathbb{P}(\mathcal{O}_{n_0,N} = j)
\end{aligned}
$$

$$
\begin{aligned}
&\mathbb{E}[\mathcal{S}_{a,n_0}(\mathcal{O}_{n_0,N})] \\
&= \mathcal{S}_{a,n_0}(0)\mathbb{P}(\Omega \setminus A_{n_0}) + \sum_{k=1}^{N}\mathcal{S}_{a,n_0}(k)\mathbb{P}(\mathcal{O}_{n_0,N} = k) + \mathcal{S}_{a,n_0}(N+1)\mathbb{P}(A_{N+n_0})
\end{aligned}
$$

for

$$\mathbf{\mathcal{S}_{a,n_0}(N) := \sum_{n=0}^{N-1} a_{n+n_0}, \qquad \mathcal{S}_{a,n_0}(0) = 0.}$$

Sending $N \to \infty$ we have by monotone convergence

$$\mathbb{E}[\mathcal{S}_{a,n_0}(\mathcal{O}_{n_0})] = \sum_{k=0}^{\infty} a_{n_0+k}\mathbb{P}(A_{n_0+k}) = \sum_{\ell=n_0}^{\infty} a_\ell \cdot \mathbb{P}(A_\ell).$$

Lemma: Borel-Cantelli moment estimate for general events

- $(\Omega, \mathcal{A}, \mathbb{P})$
- $(\mathbf{A_n})_{\mathbf{n \geqslant n_0}}$ **general events (not nested)**
- $(\mathbf{a_n})_{\mathbf{n \geqslant n_0}}$ positive $\&$ nondecreasing

Then we have

$$\mathbb{E}[\mathcal{S}_{\mathbf{a,n_0}}(\mathcal{O}_{\mathbf{n_0}})] \leqslant \mathbb{E}[\mathcal{S}_{\mathbf{a,n_0}}(\mathcal{M}_{\mathbf{n_0}})] \leqslant \sum_{\mathbf{n=n_0}}^{\infty} \mathbf{a_n} \cdot \sum_{\mathbf{m=n}}^{\infty} \mathbb{P}(\mathbf{A_m}),$$

for

$$\mathcal{O}_{\mathbf{n_0}} := \sum_{\mathbf{n=n_0}}^{\infty} \mathbf{1}(\mathbf{A_n})$$

$$\mathcal{M}_{\mathbf{n_0}} := \max\{\mathbf{n} \geqslant \mathbf{n_0} \mid \omega \in \mathbf{A_n}\} = \sum_{\mathbf{n=n_0}}^{\infty} \mathbf{1}\Big(\bigcup_{\mathbf{m=n}}^{\infty} \mathbf{A_m}\Big)$$

and

$$\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathbf{N}) := \sum_{\mathbf{n}=\mathbf{0}}^{\mathbf{N}-\mathbf{1}} \mathbf{a}_{\mathbf{n}+\mathbf{n_0}}, \qquad \mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathbf{0}) = \mathbf{0}$$

Proof:

- Fix
$$\mathbf{A_n} \subset \tilde{\mathbf{A}}_\mathbf{n} := \bigcup_{\mathbf{m}=\mathbf{n}}^{\infty} \mathbf{A_m}.$$
Note that $(\tilde{\mathbf{A}}_\mathbf{n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ is nested. Then by construction
$$\mathcal{O}_{\mathbf{n_0}} = \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{1}(\mathbf{A_n}) \leqslant \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{1}(\tilde{\mathbf{A}}_\mathbf{n}) = \mathcal{M}_{\mathbf{n_0}}.$$
- The nestedness of $(\tilde{\mathbf{A}}_\mathbf{n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ allows to apply our Lemma for nested events:
$$\mathbb{E}[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{O}_{\mathbf{n_0}})] \leqslant \mathbb{E}[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{M}_{\mathbf{n_0}})] = \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{a_n} \mathbb{P}(\tilde{\mathbf{A}}_\mathbf{n}) \leqslant \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{a_n} \sum_{\mathbf{m}=\mathbf{n}}^{\infty} \mathbb{P}(\mathbf{A_n})$$

Lemma: Moment version of the first Borel-Cantelli lemma

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$, $\mathbf{n_0} \in \mathbb{N}$
- $(\mathbf{A_n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ **family of events**
- $(\mathbf{a_n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ positive $\&$ nondecreasing

Then we have

$$\mathbb{E}\Big[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{O}_{\mathbf{n_0}})\Big] \leqslant \mathbb{E}\Big[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{M}_{\mathbf{n_0}})\Big] \leqslant \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{a_n} \sum_{\mathbf{m}=\mathbf{n}}^{\infty} \mathbb{P}(\mathbf{A_m}).$$

If the sequence $(\mathbf{A_n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ is **nested**, we have

$$\mathbb{E}\Big[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{O}_{\mathbf{n_0}})\Big] = \mathbb{E}\Big[\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathcal{M}_{\mathbf{n_0}})\Big] = \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{a_n}\mathbb{P}(\mathbf{A_n}).$$

Example 1: Polynomical probability decay $\boxed{\mathbb{P}(\mathbf{A_m}) \leqslant \mathbf{cm^{-q}}}$

Then for $\mathbf{0 < p < q - 2}$:

$$\mathbb{E}[\mathcal{O}_{\mathbf{n_0}}^{\mathbf{1+p}}] \leqslant \mathbb{E}[\mathcal{M}_{\mathbf{n_0}}^{\mathbf{1+p}}] \leqslant \mathbf{cq}\zeta(\mathbf{q-p-1};\mathbf{n_0}), \qquad \zeta(\mathbf{z};\mathbf{n_0}) = \sum_{\mathbf{n=n_0}}^{\infty} \frac{\mathbf{1}}{\mathbf{n^z}}$$

$$\mathbb{P}(\mathcal{O}_{\mathbf{n_0}} \geqslant \mathbf{k}) \leqslant \mathbb{P}(\mathcal{M}_{\mathbf{n_0}} \geqslant \mathbf{k}) \leqslant \mathbf{cq} \cdot \mathbf{k}^{-(\mathbf{p+1})} \cdot \zeta(\mathbf{q-p-1};\mathbf{n_0}) \qquad \mathbf{k} \geqslant \mathbf{1}.$$

This answers our **pink question:** For instance, given $\mathbf{q = 5}$, for which $\mathbf{p > 0}$ do we get

$$\mathbb{E}[\mathcal{O}^{\mathbf{1+p}}] < \infty \qquad ?$$

How to calculate this:

$$\sum_{n=n_0}^{\infty} n^p \sum_{m=n}^{\infty} cm^{-q} \leqslant c \sum_{n=n_0}^{\infty} n^p \left(n^{-q} + \int_n^{\infty} x^{-q} dx\right)$$

$$= c\zeta(q-p; n_0) + \frac{c}{q-1}\zeta(q-p-1; n_0)$$

$$\leqslant \frac{cq}{q-1}\zeta(q-p-1; n_0)$$

$$\mathcal{S}_{a,n_0}(N) = \sum_{n=n_0}^{N+n_0-1} n^p \geqslant \sum_{n=1}^{N} n^p \geqslant \int_0^N x^p dx = \frac{N^{p+1}}{p+1}$$

Example 2: Exponential probability decay $\boxed{\mathbb{P}(\mathbf{A_m}) \leqslant \mathbf{cb^m}}$

Then for all $\mathbf{b} \in (\mathbf{0}, \mathbf{1}), \mathbf{c} > \mathbf{0}$, $\mathbf{p} \in (\mathbf{0}, \mathbf{1})$:

$$\mathbb{E}[\mathbf{b}^{-\mathbf{p}\mathcal{O}_{\mathbf{n_0}}}] \leqslant \mathbb{E}[\mathbf{b}^{-\mathbf{p}\mathcal{M}_{\mathbf{n_0}}}] \leqslant \mathbf{1} + \frac{\mathbf{cb^{n_0-1}}}{\mathbf{1} - \mathbf{b^{1-p}}}$$

and for k$\geqslant$1

$$\mathbb{P}(\mathcal{O}_{\mathbf{n_0}} \geqslant \mathbf{k}) \leqslant \mathbb{P}(\mathcal{M}_{\mathbf{n_0}} \geqslant \mathbf{k}) \leqslant \mathbf{2}^{\mathbf{9/8}}[\mathbf{k}(\mathbf{cb^{n_0-1}} + \mathbf{1}) + \mathbf{1}] \cdot \mathbf{b^k}.$$

Example 3: **Weibull decay** $\boxed{\mathbb{P}(\mathbf{A_m}) \leqslant \mathbf{c}\mathbf{b}^{\mathbf{m}^{\alpha}}}$ $\mathbf{b}, \alpha, \in (\mathbf{0}, \mathbf{1}), \mathbf{c} > \mathbf{0}$

Then for any $\mathbf{p} \in (\mathbf{0}, \mathbf{1}) \ \exists \, \mathbf{K} = \mathbf{K}(\mathbf{p}, \mathbf{c}, \mathbf{b}, \alpha) > \mathbf{0}$:

$$\mathbb{E}[\mathbf{b}^{-\mathbf{p}(\mathcal{O}_{\mathbf{n_0}} + \mathbf{n_0} - \mathbf{1})}] \leqslant \mathbb{E}[\mathbf{b}^{-\mathbf{p}(\mathcal{M}_{\mathbf{n_0}} + \mathbf{n_0} - \mathbf{1})}] \leqslant \mathbf{K}$$

and

$$\mathbb{P}(\mathcal{O}_{\mathbf{n_0}} \geqslant \mathbf{k}) \leqslant \mathbb{P}(\mathcal{O}_{\mathbf{n_0}} \geqslant \mathbf{k}) \leqslant \mathbf{b}^{\mathbf{p}(\mathbf{k}-\mathbf{1})^{\alpha}}\mathbf{K}$$

Recall:

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$
- $(\mathbf{X_n})_{\mathbf{n}\in\mathbb{N}}$, $\mathbf{X_n} : \mathbf{\Omega} \to \mathbb{R}$
- $\mathbf{X} : \mathbf{\Omega} \to \mathbb{R}$

1)

$$\mathbf{X_n} \stackrel{\mathbf{n}\to\infty}{\longrightarrow} \mathbf{X} \qquad \textbf{in probability}$$

if

$$\lim_{\mathbf{n}\to\infty} \mathbb{P}(|\mathbf{X_n} - \mathbf{X}| > \varepsilon) = \mathbf{0} \qquad \forall \varepsilon > \mathbf{0}$$

2)

$$\mathbf{X_n} \stackrel{\mathbf{n}\to\infty}{\longrightarrow} \mathbf{X} \qquad \mathbb{P} - \text{a.s.}$$

if $\exists\, \tilde{\mathbf{\Omega}} \in \mathcal{A}$ con $\mathbb{P}(\tilde{\mathbf{\Omega}}) = \mathbf{1}$ s.th.

$$\forall \omega \in \tilde{\mathbf{\Omega}} : \qquad \lim_{\mathbf{n}\to\infty} \mathbf{X_n}(\omega) = \mathbf{X}(\omega).$$

Tradeoff Lemma of a.s. convergence[3]:

- $(\mathcal{X}, \mathcal{B}(\mathcal{X}), \mathbf{d})$ Polish space
- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$ with $\mathbf{X_n}, \mathbf{X} : \mathbf{\Omega} \to \mathcal{X}$ r.v., $\qquad \mathbf{n} \geqslant \mathbf{n_0}$
- $\mathbf{p}(\delta, \mathbf{n}) := \mathbb{P}(\mathbf{d}(\mathbf{X_n}, \mathbf{X}) > \delta) \to \mathbf{0}$, as $\mathbf{n} \to \infty \quad \forall\, \delta > \mathbf{0}$ (conv. in $\mathbb{P}$)

Then for any $\epsilon := (\varepsilon_\mathbf{n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ $(> 0, \searrow)$ and $\mathbf{a} = (\mathbf{a_n})_{\mathbf{n} \geqslant \mathbf{n_0}}$ $(> 0, \nearrow)$ s.th.

$$\mathbf{K}(\mathbf{a}, \epsilon, \mathbf{n_0}) := \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{a_n} \sum_{\mathbf{m}=\mathbf{n}}^{\infty} \mathbf{p}(\varepsilon_\mathbf{m}, \mathbf{m}) < \infty$$

1) we have the **a.s. asymptotic rate**

$$\mathbf{d}(\mathbf{X_n}, \mathbf{X}) \leqslant \varepsilon_\mathbf{n} \qquad \text{a.s. for all } \mathbf{n} \geqslant \mathcal{M}_{\varepsilon, \mathbf{n_0}}$$

2) we have the integrability of the **overshoot / modulus of convergence**

$$\mathbb{E}[\mathcal{S}_{\mathbf{a}, \mathbf{n_0}}(\mathcal{O}_{\epsilon, \mathbf{n_0}})] \leqslant \mathbb{E}[\mathcal{S}_{\mathbf{a}, \mathbf{n_0}}(\mathcal{M}_{\epsilon, \mathbf{n_0}})] \leqslant \mathbf{K}(\mathbf{a}, \epsilon, \mathbf{n_0})$$

[3] Luisa F. Estrada, Michael A. Högele, Alexander Steinicke: On the tradeoff between almost sure error tolerance versus mean deviation frequency in martingale convergence, https://arxiv.org/abs/2310.09055

and in particular

$$\mathbb{P}(\mathcal{M}_{\epsilon,\mathbf{n_0}} \geqslant \ell) \leqslant \frac{\mathbf{K}(\mathbf{a}, \epsilon, \mathbf{n_0})}{\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\ell)}$$

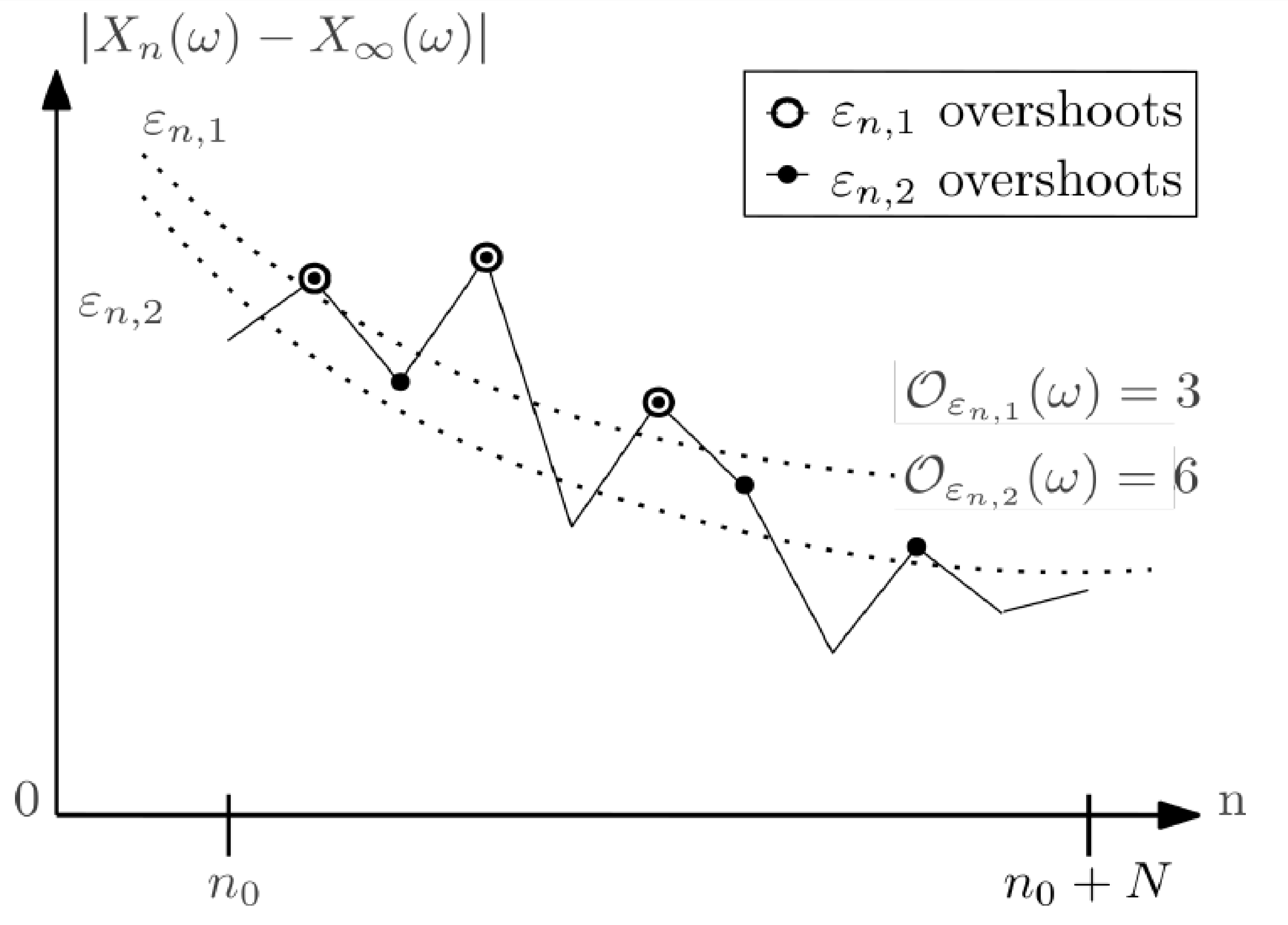
where

$$\mathcal{O}_{\varepsilon,\mathbf{n_0}} = \sum_{\mathbf{n}=\mathbf{n_0}}^{\infty} \mathbf{1}\{\mathbf{d}(\mathbf{X_n}, \mathbf{X}) > \varepsilon_{\mathbf{n}}\},$$

$$\mathcal{M}_{\varepsilon,\mathbf{n_0}} = \max\{\mathbf{n} \geqslant \mathbf{n_0} \mid \mathbf{d}(\mathbf{X_n}, \mathbf{X}) > \varepsilon_{\mathbf{n}}\}$$

and

$$\mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathrm{N}) = \sum_{\mathbf{n}=\mathbf{0}}^{\mathbf{N}-\mathbf{1}} \mathbf{a}_{\mathbf{n}+\mathbf{n_0}} \qquad \text{with} \qquad \mathcal{S}_{\mathbf{a},\mathbf{n_0}}(\mathbf{0}) = \mathbf{0}$$



|Xn(ω) − X∞(ω)|
εn,1 overshoots
εn,2 overshoots
εn,1
εn,2
O εn,1 (ω) = 3
O εn,2 (ω) = 6
0
n
n0
n0 + N

Example 4: Law of large numbers (Baum, Katz, 1965)

$(\mathbf{X_i})_{\mathbf{i}\in\mathbb{N}}$ centered i.i.d. Then are equivalent:

1. $\mathbb{E}[|\mathbf{X_1}|^{\mathbf{p}}] < \infty$ for $\mathbf{p} > \mathbf{1}$
2. For any $\frac{\mathbf{p}}{\mathbf{2}} < \alpha \leqslant \mathbf{p}$ and any $\mathbf{c} > \mathbf{0}$ we have

$$\sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbf{n}^{\alpha-\mathbf{2}} \cdot \mathbb{P}\left(|\bar{\mathbf{X}}_{\mathbf{n}}| > \frac{\mathbf{c}}{\mathbf{n}^{\mathbf{1}-\frac{\alpha}{\mathbf{p}}}}\right) < \infty$$

For any $\frac{\mathbf{p}}{\mathbf{2}} < \alpha \leqslant \mathbf{p}$ and any $\mathbf{c} > \mathbf{0}$ we have

$$\sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbf{n}^{\alpha-\mathbf{2}} \cdot \mathbb{P}\Big(|\bar{\mathbf{X}}_{\mathbf{n}}| > \frac{\mathbf{c}}{\mathbf{n}^{\mathbf{1}-\frac{\alpha}{\mathbf{p}}}}\Big) < \infty$$

Kronecker's lemma: $(\mathbf{c_n})_{\mathbf{n}\in\mathbb{N}}$, $(\mathbf{b_n})_{\mathbf{n}\in\mathbb{N}}$, both $> \mathbf{0}$, $\mathbf{b_n} \to \infty$

$$\sum_{n=1}^{\infty} \frac{c_n}{b_n} < \infty \qquad \text{implies} \quad \lim_{n\to\infty} \frac{1}{b_n} \sum_{i=1}^{n} c_i = 0$$

$$\Rightarrow \mathbb{P}\Big(|\bar{\mathbf{X}}_{\mathbf{n}}| > \frac{\mathbf{c}}{\mathbf{n}^{\mathbf{1}-\frac{\alpha}{\mathbf{p}}}}\Big) \cdot \sum_{\mathbf{i}=\mathbf{1}}^{\mathbf{n}} \mathbf{i}^{\alpha-\mathbf{2}} \to \mathbf{0} \quad \text{and} \quad \mathbb{P}\Big(|\bar{\mathbf{X}}_{\mathbf{n}}| > \frac{\mathbf{c}}{\mathbf{n}^{\mathbf{1}-\frac{\alpha}{\mathbf{p}}}}\Big) = \mathbf{o}(\mathbf{n}^{\alpha-\mathbf{1}})$$

Combining Lemma 2 + Example 1 + $\mathbb{P}\left(|\bar{\mathbf{X}}_{\mathbf{n}}| > \frac{\mathbf{c}}{\mathbf{n}^{1-\frac{\alpha}{\mathbf{p}}}}\right) = \mathbf{o}(\mathbf{n}^{\alpha-1})$:

We have the tradeoff for moments $\mathbf{p} > \mathbf{3}$!

For $\mathbf{p}, \alpha > \mathbf{3}$ with $\frac{\mathbf{p}}{\mathbf{2}} < \alpha \leqslant \mathbf{p}$ and $\mathbf{0} < \tilde{\mathbf{p}} \leqslant \alpha - \mathbf{3}$

$$\mathbb{E}[\mathcal{O}^{\mathbf{1}+\tilde{\mathbf{p}}}] \leqslant \mathbb{E}[\mathcal{M}^{\mathbf{1}+\tilde{\mathbf{p}}}] \leqslant \mathbf{C}(\alpha - \mathbf{1})\zeta(\alpha - \mathbf{2} - \tilde{\mathbf{p}}, \mathbf{n_0}) < \infty,$$

and

$$\limsup_{\mathbf{n}\to\infty} |\bar{\mathbf{X}}_{\mathbf{n}}| \cdot \mathbf{n}^{\mathbf{1}-\frac{\alpha}{\mathbf{p}}} \leqslant \mathbf{1} \qquad \text{a.s.}$$

Example 5: Cramér's theorem

For $(\mathbf{X_i})_{\mathbf{i}\in\mathbb{N}}$ centered i.i.d. with $\mathbb{E}[\mathbf{e}^{\lambda|\mathbf{X_1}|}] < \infty$ for some $\lambda > \mathbf{0}$

Then for any vecinity $\bar{\mathbf{A}} \ni \mathbf{0}$

$$\mathbb{P}(\bar{\mathbf{X}}_\mathbf{n} \in \mathbf{A}^\mathbf{c}) \leqslant \mathbf{2}\exp\Big(-\mathbf{n}\inf_{\mathbf{x}\in\mathbf{A}^\mathbf{c}} \mathcal{I}(\mathbf{x})\Big), \qquad \mathbf{n} \geqslant \mathbf{1}$$

where $\mathcal{I}(\mathbf{x})$ is the Fenchel-Legendre transform of $\mathbf{X_1}$
($\mathcal{I}(\mathbf{x}) \geqslant \mathbf{0}$, convex, $\mathcal{I}(\mathbf{0}) = \mathbf{0}$)

Large deviations principle (LDP)

In particular, for $\epsilon = (\varepsilon_{\mathbf{n}})_{\mathbf{n}\in\mathbb{N}}$ ($> \mathbf{0}$, $\searrow$) and $n \geqslant 1$ we have that

$$\mathbb{P}(\bar{\mathbf{X}}_{\mathbf{n}} \in \mathbf{B}^{\mathbf{c}}_{\varepsilon_{\mathbf{n}}}(\mathbf{0}) \leqslant \mathbf{2}\,\exp\Big(-\mathbf{n} \inf_{|\mathbf{x}|>\varepsilon_{\mathbf{n}}} \mathcal{I}(\mathbf{x})\Big) \approx \exp\Big(-\mathbf{n}\frac{\varepsilon_{\mathbf{n}}^{\mathbf{2}}}{\mathbf{2}}(\mathbf{D}^{\mathbf{2}}\mathcal{I}(\mathbf{0}))\Big)$$

Combining Lemma 2 + Example 3 + $\mathbb{P}(\bar{\mathbf{X}}_{\mathbf{n}} \in \mathbf{B}^{\mathbf{c}}_{\varepsilon}(\mathbf{0}))) = \mathbf{O}(\mathbf{e}^{\mathbf{n}\varepsilon_{\mathbf{n}}^{\mathbf{2}}\frac{(\mathbf{D}^{\mathbf{2}}\mathcal{I}(\mathbf{0}))}{\mathbf{2}}})$:

We have the tradeoff for $\varepsilon_{\mathbf{n}} = \mathbf{n}^{-\rho}$, $\rho \in [\mathbf{0}, \frac{\mathbf{1}}{\mathbf{2}})$!

$$\mathbb{E}[\mathbf{e}^{\tilde{\mathbf{p}}\mathcal{O}_{\epsilon}^{\mathbf{1}-\mathbf{2}\rho}}] \leqslant \mathbb{E}[\mathbf{e}^{\tilde{\mathbf{p}}\mathcal{M}_{\epsilon}^{\mathbf{1}-\mathbf{2}\rho}}] \leqslant \mathbf{K}(\rho, \tilde{\mathbf{p}}, \epsilon, \mathbf{D}^{\mathbf{2}}\mathcal{I}(\mathbf{0})) < \infty$$

y

$$\limsup_{\mathbf{n}\to\infty} |\bar{\mathbf{X}}_{\mathbf{n}}| \cdot \varepsilon_{\mathbf{n}}^{-\mathbf{1}} \leqslant \mathbf{1} \qquad \text{a.s.}$$

Obvious applications in any context with an **LDP**:

- The Glivenko-Cantelli theorem
- The Sanov theorem (i.i.d. + MC)
- Excursion frequencies of rare sequences for random walks

Applications with **sums of independent increments**:

- Quantifying the a.s. version of the CLT, Gaal-Koksma strong law
- A.s. rates of convergence of statistical M-estimators for bounded r.v.

III. Returning to the initial equation:

The rate of convergence of the Polya urn

Recall:

- Urn with $\mathbf{N}$ balls
- $\mathbf{B}$ of them are **black**, and $\mathbf{N}-\mathbf{B}$ of them are white.

- Step 0: Initial value

$$\mathbf{X_0} := \frac{\mathbf{B}}{\mathbf{N}}$$

- Step n: Given $\mathbf{X_{n-1}}$ sample $\mathbf{X_n} \sim \mathcal{B}_{\mathbf{X_{n-1}}}$

$$\mathbf{X_n} := \frac{\mathbf{B} + \sum_{\mathbf{i=1}}^{\mathbf{n}} \mathbf{Y_i}}{\mathbf{N}+\mathbf{n}}$$

1. It is a martingale ✓
2. It has increments, which are bounded by $\frac{\mathbf{1}}{\mathbf{n}}$

Theorem: Azuma-Hoeffding inequality

- Martingale $\mathbf{X} = (\mathbf{X_n})_{\mathbf{n} \in \mathbb{N}}$
- The increments of $\mathbf{X}$ are a.s. bounded by $(\mathbf{c_n})_{\mathbf{n} \in \mathbb{N}}$ that is

$$|\mathbf{X_n} - \mathbf{X_{n-1}}| \leqslant \mathbf{c_n} \qquad \text{a.s. for all } \mathbf{n} \in \mathbb{N}.$$

Then for $\mathbf{m} \leqslant \mathbf{n}$

$$\mathbb{P}(\mathbf{X_n} - \mathbf{X_m} \geqslant \varepsilon) \leqslant \exp\left(-\frac{\mathbf{1}}{\mathbf{2}} \frac{\varepsilon^{\mathbf{2}}}{\sum_{\mathbf{i}=\mathbf{m}+\mathbf{1}}^{\mathbf{n}} \mathbf{c_i^2}}\right).$$

Theorem: a.s. rates via Azuma-Hoeffding closure

- Martingale $\mathbf{X} = (\mathbf{X_n})_{\mathbf{n}\in\mathbb{N}}$
- The increments of $\mathbf{X}$ are a.s. bounded by $(\mathbf{c_n})_{\mathbf{n}\in\mathbb{N}}$ that is
- $$|\mathbf{X_n} - \mathbf{X_{n-1}}| \leqslant \mathbf{c_n} \qquad \text{a.s. for all } \mathbf{n} \in \mathbb{N}.$$ $$\sum_{\mathbf{n=1}}^{\infty} \mathbf{c_n^2} < \infty \qquad \text{and set } \mathbf{r}(\mathbf{n}) := \sum_{\mathbf{k=n+1}}^{\infty} \mathbf{c_k^2}$$

Then there exists a r.v. $\mathbf{X}_\infty$ such that $\mathbf{X_n} \to \infty$ a.s. as $\mathbf{n} \to \infty$. If

- $\epsilon = (\varepsilon_\mathbf{n})_{\mathbf{n}\in\mathbb{N}}$ positive $\&$ nonincreasing $\varepsilon_\mathbf{n} \to \mathbf{0}$
- $\mathbf{a} = (\mathbf{a_n})_{\mathbf{n}\in\mathbb{N}}$ positive $\&$ nondecreasing

such that

$$\mathbf{K}(\mathbf{a}, \epsilon) := \mathbf{2} \sum_{\mathbf{n=1}}^{\infty} \mathbf{a_n} \sum_{\mathbf{m=n}}^{\infty} \exp\Big(- \frac{\mathbf{1}}{\mathbf{2}} \frac{\varepsilon_\mathbf{m}^\mathbf{2}}{\mathbf{r}(\mathbf{m})} \Big) < \infty,$$

then

$$\limsup_{\mathbf{n}\to\infty} |\mathbf{X_n} - \mathbf{X}_\infty| \cdot \varepsilon_\mathbf{n}^{-\mathbf{1}} \leqslant \mathbf{1} \quad \text{a.s.} \qquad \&$$

$$\mathbb{E}[\mathcal{S}_{\mathbf{a},\mathbf{1}}(\mathcal{O}_\varepsilon)] \leqslant \mathbb{E}[\mathcal{S}_{\mathbf{a},\mathbf{1}}(\mathcal{M}_\varepsilon)] \leqslant \mathbf{K}(\mathbf{a}, \varepsilon).$$

The rate of convergence of the Polya urn:

For any $\mathbf{p} \in (\mathbf{0}, \mathbf{1})$ and

$$\epsilon = (\varepsilon_{\mathbf{n}})_{\mathbf{n} \in \mathbb{N}} \qquad \varepsilon_{\mathbf{n}} := \sqrt{\frac{\mathbf{2}}{\mathbf{3n^p}}}, \qquad \mathbf{n} \geqslant \mathbf{1}$$

by the Corollary we have

1. We have a.s.

$$\limsup_{\mathbf{n} \to \infty} |\mathbf{X_n} - \mathbf{X_\infty}| \cdot \mathbf{n}^{\frac{\mathbf{p}}{\mathbf{2}}} \leqslant \sqrt{\frac{\mathbf{3}}{\mathbf{2}}}$$

2. For any $\mathbf{q} \in (\mathbf{0}, \mathbf{1})$ we have

$$\mathbb{E}[\mathbf{e}^{\mathbf{q}\mathcal{O}_{\varepsilon,\mathbf{1}}^{\mathbf{1-2p}}}] \leqslant \mathbb{E}[\mathbf{e}^{\mathbf{q}\mathcal{M}_{\varepsilon,\mathbf{1}}^{\mathbf{1-2p}}}] \leqslant \mathbf{K}((\mathbf{e}^{\mathbf{qn^{1-2p}}})_{\mathbf{n}}, \varepsilon, \mathbf{1})$$

Note:

- We only used the Azuma-Hoeffding inequality for **martingales with a.s. square summably bounded increments**.

- More inmediate examples:

  - **Generalized Polya urns** with more colors and more general replacement rules

  - **Excursion frequencies for different heights** for the martingales associated to the **supercritical branching process**

IV. Another type of applications: Brownian path properties approximations

- More than 1 century of hiding the approximations of Brownian sample paths in order to extract precise path properties
  $\longrightarrow$ rough path theory

- **Idea:** reverse engineering of this path abstraction, terms of a.s. convergence with higher order MDF.

- Use the approximations of Brownian path properties in the literature and quantify those.

Theorem 6 (Paley, Wiener, Zygmund). *The event*

$$\{\omega \in \Omega \mid \text{ for each } t \in [0,1] \text{ either } D^+W_t(\omega) = \infty \text{ or } D_+W_t(\omega) = -\infty\}$$

contains an event $E \in \mathcal{A}$ *with* $\mathbb{P}(E) = 1$.

- Clearly, for any finite time step discretization $(\mathbf{W_{t_n}})_\mathbf{n}$ this is false.
- However we can quantify, how fast, this property emerges a.s.

Theorem: (Paley, Wiener, Zygmund, quantitative)

For $c_\pi := \frac{2^{10}}{\pi^2}$ *and any* $b \in (1, 2^{1/4})$ *we have*

$$\mathbb{P}\Big(\#\Big\{n \in \mathbb{N} \mid \exists\, s \in [0,1] : \sup_{t \in [s-2^{-n}, s+2^{-n}] \cap [0,1]} \frac{|W(s) - W(t)|}{2^{-n}} \leqslant b^n\Big\} \geqslant k\Big)$$

$$\leqslant 2e^{\frac{9}{8}} \cdot \Big[k\big(\frac{2c_\pi}{b^4} + 1\big) + 1\Big] \cdot \Big(\frac{b^4}{2}\Big)^k, \quad k \geqslant 1.$$

Sketch of proof (Koshnevisa, Karatzas / Shreve)

$$E_\lambda^n := \{\exists\, s \in [0,1] \mid \sup_{t\in[s-2^{-n},s+2^{-n}]\cap[0,1]} \frac{|W(s)-W(t)|}{2^{-n}} \leqslant \lambda\}$$

$$\mathbb{P}(E_\lambda^n) \leqslant 2^n\Big(\int_{-\lambda 2^{-n/2+2}}^{\lambda 2^{-n/2+2}} \frac{e^{-x^2/2}}{\sqrt{2\pi}}dx\Big)^4 \leqslant 2^n\Big(\frac{2}{\sqrt{2\pi}}\lambda 2^{-n/2+2}\Big)^4$$
$$= 2^n\lambda^4\Big(\frac{8}{\sqrt{2\pi}}\Big)^4 2^{-2n} = \Big(\frac{8}{\sqrt{2\pi}}\Big)^4\lambda^4 2^{-n} = \Big(\frac{1024}{\pi^2}\Big)\lambda^4 2^{-n}$$

$$\mathcal{O}_\lambda := \textstyle\sum_{n=0}^\infty \mathbf{1}_{E_\lambda^n}$$

$$\mathbb{E}[e^{r\mathcal{O}_\lambda}] \leqslant 1 + \frac{2c_\pi\lambda^4}{1-e^p\frac{1}{2}}$$

for $c_\pi := \frac{1024}{\pi^2}$

$$\mathbb{P}(\mathcal{O}_\lambda \geqslant k) \leqslant 2e^{\frac{9}{8}}\cdot[k(2c_\pi\lambda^4+1)+1]\cdot 2^{-k}, \qquad k \geqslant 1$$

For the special case of $\lambda = \lambda_n = b^n$ for some $1 < b < 2^{1/4}$ we have

$$\mathbb{P}(E_\lambda^n) \leqslant c_\pi (b^4/2)^n, \qquad n \in \mathbb{N},$$

$\mathcal{O} := \sum_{n=0}^\infty \mathbf{1}_{E_{\lambda_n}^n}$ and any $0 < r < \ln(2/b^4)$

$$\mathbb{E}[e^{r\mathcal{O}}] \leqslant \frac{2}{b^4} \frac{c_\pi}{(1 - e^r b^4/2)} + 1$$

$$\mathbb{P}\left(\#\left\{ n \in \mathbb{N} \mid \exists\, s \in [0,1] : \sup_{t \in [s-2^{-n}, s+2^{-n}] \cap [0,1]} \frac{|W(s) - W(t)|}{2^{-n}} \leqslant b^n \right\} \geqslant k \right)$$
$$\leqslant 2e^{\frac{9}{8}} \cdot [k(\frac{2c_\pi}{b^4} + 1) + 1] \cdot \left(\frac{b^4}{2}\right)^k.$$

More Brownian path property approximations:

1. A.s. uniform convergence of Lévy’s construction

2. Kolmogorov-Chentsov continuity theorem

3. Lévy’s modulus of continuity

4. Loss of path monotonicity

5. Laws of the iterated logarithm (Khinchin, Chung’s “other” law, Strassen)

V. Extention of qBC1 to independent events

Motivation: Kolmogorov's 3 series theorem

- $(\Omega, \mathcal{A}, \mathbb{P})$,
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events, $\mathcal{O} := \sum_{\mathbf{n=1}}^{\infty} \mathbf{1}(\mathbf{A_n})$

Hence

$$\mathcal{O} < \infty \quad \mathbb{P}-\text{c.s.} \quad \Leftrightarrow \quad \text{Var}(\mathcal{O}) < \infty \quad \Leftrightarrow \quad \mathbf{C_1} = \sum_{\mathbf{n=1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty.$$

since

$$\begin{aligned}
\mathbb{E}[\mathcal{O}^2] = \mathbb{E}[(\sum_{\mathbf{n=1}}^{\infty} \mathbf{1}_{\mathbf{A_n}})^2] &= \mathbb{E}[\sum_{\mathbf{n=1}}^{\infty} \mathbf{1}_{\mathbf{A_n}} + \sum_{\mathbf{n\neq m}} \mathbf{1}_{\mathbf{A_n}}\mathbf{1}_{\mathbf{A_m}}] \\
&= \sum_{\mathbf{n=1}}^{\infty} \mathbb{E}[\mathbf{1}_{\mathbf{A_n}}] + \sum_{\mathbf{n\neq m}} \mathbb{E}[\mathbf{1}_{\mathbf{A_n}}\mathbf{1}_{\mathbf{A_m}}] \\
&\leqslant \sum_{\mathbf{n=1}}^{\infty} \mathbb{P}(\mathbf{A_n}) + \sum_{\mathbf{n,m=1}}^{\infty} \mathbb{P}(\mathbf{A_n})\mathbb{P}(\mathbf{A_m}) = \mathbf{C_1}(\mathbf{1} + \mathbf{C_1})
\end{aligned}$$

Preliminary result: Freedman's universal bound

- $(\boldsymbol{\Omega}, \mathcal{A}, \mathbb{P})$,
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_1} := \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbb{P}(\mathbf{E_n}) < \infty$

Then for all $r > 0$

$$\mathbb{E}[\mathbf{e}^{\mathbf{r}\mathcal{O}}] \leqslant \mathbf{e}^{\mathbf{C_1}(\mathbf{e}^{\mathbf{r}}-\mathbf{1})}$$

$$\mathbb{P}(\mathcal{O} \geqslant k) \leqslant \inf_{r>0} \exp(-kr + C_1(e^r - 1)) = \exp(-k\ln(k) + k(\ln(C_1) + 1) - C_1)$$

Proof (sketch):

$$\text{Bernoulli inequality } 1 + x \leqslant e^x,\ x \in \mathbb{R}$$

$$\mathbb{E}\big[e^{r\mathcal{O}}\big] = \prod_{n=1}^{\infty} \mathbb{E}[e^{r\mathbf{1}_{E_n}}] = \prod_{n=1}^{\infty} \big(e^r\mathbb{P}(E_n) + 1 - \mathbb{P}(E_n)\big) = \prod_{n=1}^{\infty} \exp\big(\ln(1 + e^r\mathbb{P}(E_n) - \mathbb{P}(E_n))\big)$$

$$= \exp\Big(\sum_{n=1}^{\infty} \ln(1 + (e^r - 1)\mathbb{P}(E_n))\Big) \leqslant \exp\Big((e^r - 1)\sum_{n=1}^{\infty} \mathbb{P}(E_n)\Big) = \exp\big(C_1(e^r - 1)\big).$$

The second Borel-Cantelli lemma:

- $(\boldsymbol{\Omega}, \mathcal{A}, \mathbb{P})$,
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_1} := \sum_{\mathbf{n=1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$

$$1.\quad \sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \quad \Longrightarrow \quad \mathcal{O} = \infty \quad a.s.$$

$$2.\quad \sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty \quad \Longrightarrow \quad \mathcal{O} < \infty \quad a.s.\ \ with \quad \mathbb{E}[e^{r\mathcal{O}}] \leqslant \exp\Big((e^r - 1) \cdot \sum_{n=1}^{\infty} \mathbb{P}(E_n)\Big) \leqslant \infty$$

The second Borel-Cantelli lemma:

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$,
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_1} := \sum_{\mathbf{n=1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$

1. $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \quad \Longrightarrow \quad \mathcal{O} = \infty \quad a.s.$
2. $\displaystyle\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty \quad \Longrightarrow \quad \mathcal{O} < \infty \quad a.s. \ with \quad \mathbb{E}[e^{r\mathcal{O}}] \leqslant \exp\Big((e^r - 1) \cdot \sum_{n=1}^{\infty} \mathbb{P}(E_n)\Big) \leqslant \infty$

How can this result be quantified, where the **rate of convergence** of $(\mathbb{P}(\mathbf{A_n}))_{\mathbf{n}\in\mathbb{N}}$ appears, instead of only the **value** $\mathrm{C}_{\mathbf{1}}$?

First formulation: C_1

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_1} := \sum_{\mathbf{n}=\mathbf{1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$
- $\boxed{\mathbf{C_1} < \mathbf{e^{-r}}}$ for some $\mathbf{r} > \mathbf{0}$

Then

$$\mathbb{E}\left[e^{r\mathcal{O}}\right] \leqslant (1 - C_1 e^r)^{-1} \qquad \textit{for all } r < |\ln(C_1)|.$$

Proof (sketch): The distribution of $\mathcal{O}_{\mathrm{N}}$ is known

$$G_k^N := \{\mathcal{O}_N = k\} \text{ and } \mathcal{O}_N := \sum_{n=1}^N \mathbf{1}_{E_n}$$

$$\sum_{k=0}^{N} a_k \,\mathbb{P}(G_k^N) = \sum_{n=0}^{N} \mathcal{Q}_n^N (a_n - a_0), \qquad \text{where} \quad \mathcal{Q}_n^N = \sum_{\substack{J \subset \{1,\ldots,N\} \\ |J|=n}} \mathbb{P}(\bigcap_{j \in J} E_j).$$

$$\mathcal{Q}_n^N = \sum_{\substack{J \subset \{1,\ldots,N\} \\ |J|=n}} \prod_{j \in J} \mathbb{P}(E_j) = \sum_{i_1=1}^{N-n} \sum_{i_2=i_1+1}^{N-n+1} \cdots \sum_{i_{n-1}=i_{n-2}+1}^{N-1} \sum_{i_n=i_{n-1}+1}^{N} \prod_{\ell=1}^{n} \mathbb{P}(E_{i_\ell}) \leqslant \Big(\sum_{i=1}^{N} \mathbb{P}(E_i)\Big)^n \leqslant C_1^n$$

$$\sum_{k=0}^{\infty} \mathbb{P}(G_k^N) \cdot e^{rk} = \sum_{k=0}^{N} \mathbb{P}(G_k^N) \cdot e^{rk} = \sum_{n=0}^{N} \mathcal{Q}_n^N (e^{rn} - 1) \leqslant \sum_{n=0}^{N} e^{rn} \mathcal{Q}_n^N \leqslant \sum_{n=0}^{N} e^{rn} C_1^n \leqslant (1 - C_1 e^r)^{-1}$$

Passing to the limit $\mathrm{N} \to \infty$ we conclude.

Second formulation:

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_m} := \sum_{\mathbf{n}=\mathbf{m}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty, \mathbf{m} \in \mathbb{N}$
- $\mathbf{N}_\delta(\mathbf{r}) := \inf\{\mathbf{m} \in \mathbb{N} \mid \mathbf{C_m} < \mathbf{e}^{-\mathbf{r}}/\delta\}$ for $\mathbf{r} > \mathbf{0}$ and $\delta > \mathbf{1}$

Then for all $\mathbf{r} > \mathbf{0}$ we have

$$\mathbb{E}[\mathbf{e}^{\mathbf{r}\mathcal{O}}] \leqslant \frac{\mathbf{e}^{\mathbf{rm}}}{\mathbf{1} - \mathbf{C_m}\mathbf{e}^{\mathbf{r}}} \qquad \text{for all } \mathbf{m} \geqslant \mathbf{N}_\delta(\mathbf{r}), \delta > \mathbf{1}$$

Second formulation:

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_m} := \sum_{\mathbf{n}=\mathbf{m}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$
- $\mathbf{N}_\delta(\mathbf{r}) := \inf\{\mathbf{m} \in \mathbb{N} \mid \mathbf{C_m} < \mathbf{e}^{-\mathbf{r}}/\delta\}$ for $\mathbf{r} > \mathbf{0}$ and $\delta > \mathbf{1}$

Then for all $\mathbf{r} > \mathbf{0}$ we have

$$\mathbb{E}[\mathbf{e}^{\mathbf{r}\mathcal{O}}] \leqslant \frac{\mathbf{e}^{\mathbf{rm}}}{\mathbf{1} - \mathbf{C_m}\mathbf{e}^{\mathbf{r}}} \qquad \text{for all } \mathbf{m} \geqslant \mathbf{N}_\delta(\mathbf{r}), \delta > \mathbf{1}$$

which we can optimize

$$\mathbb{E}\left[e^{r\mathcal{O}}\right] \leqslant \inf_{\delta>1} \inf_{m \geqslant N_r(\delta)} e^{rm}(1 - C_m e^r)^{-1} = \inf_{\delta>1} \frac{\delta}{\delta - 1} e^{r \cdot N_r(\delta)}$$

Third formulation:

- $(\boldsymbol{\Omega}, \mathcal{A}, \mathbb{P})$,
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- Sea $\sum_{\mathbf{n=1}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty$
- $\mathbf{C_m} := \sum_{\mathbf{n=m}}^{\infty} \mathbb{P}(\mathbf{A_n})$
- For $L : (0, \infty) \to (0, \infty)$ non increasing, invertible such that

$$\mathbf{L}(\mathbf{m}) = \mathbf{C_m}$$

Then for all $\mathbf{r} > \mathbf{0}$ we get

$$\mathbb{E}\left[e^{r\mathcal{O}}\right] \leqslant \inf_{\delta>1} \frac{\delta}{\delta - 1} \exp\left(r \cdot L^{-1}(e^{-r}/\delta)\right)$$

The quantitative second Borel-Cantelli lemma:

- $(\mathbf{\Omega}, \mathcal{A}, \mathbb{P})$,
- $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent events
- $\mathbf{C_m} := \sum_{\mathbf{n}=\mathbf{m}}^{\infty} \mathbb{P}(\mathbf{A_n}) < \infty, \quad \mathbf{m} \in \mathbb{N}$
- $\mathbf{L} : (\mathbf{0}, \infty) \to (\mathbf{0}, \infty)$ non-increasing, invertible with $\boxed{\mathbf{L}(\mathbf{m}) = \mathbf{C_m}}$

Then

1. $\sum_{n=1}^{\infty} \mathbb{P}(E_n) = \infty \quad \Longrightarrow \quad \mathcal{O} = \infty \quad a.s.$
2. $\sum_{n=1}^{\infty} \mathbb{P}(E_n) < \infty \quad \Longrightarrow \quad \mathcal{O} < \infty \quad a.s.$

$$\mathbb{E}\left[e^{r\mathcal{O}}\right] \leqslant \inf_{\delta>1} \frac{\delta}{\delta - 1} \exp(r \cdot L^{-1}(e^{-r}/\delta)) < \infty$$

Example: $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent $\mathbb{P}(\mathbf{E_n}) \leqslant \mathbf{c}/\mathbf{n^p}$, $\mathbf{p} > \mathbf{1}$.

$$\mathbb{E}\left[e^{r\mathcal{O}}\right] \leqslant \inf_{\delta>1} \frac{\delta}{\delta - 1} \exp\left((\delta c)^{1/p} \cdot r\, e^{r/p}\right) \leqslant 2 \exp\left((2c)^{1/p} \cdot r\, e^{r/p}\right)$$

By Markov's inequality

$$\mathbb{P}(\mathcal{O} \geqslant k) \leqslant \inf_{r>0} 2 \exp\left(-kr + (2c)^{1/p} r e^{r/p}\right)$$

$$\mathbb{P}(\mathcal{O} \geqslant k) \leqslant \mathcal{K} \cdot \exp(-pk[\ln(k) - \ln(\ln(k))]),\ k > e^2$$

Example: $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent and $\mathbb{P}(\mathbf{A_n}) \leqslant \mathbf{cb^n}$, $\mathbf{b} \in (\mathbf{0},\mathbf{1})$

$$\mathbb{E}\big[e^{r\mathcal{O}}\big] \leqslant \inf_{\delta>1} \frac{\delta}{\delta-1} \exp\big(\big[r^2 + r\cdot \ln(\delta c)\big]/|\ln(b)|\big)$$

$$\leqslant 2\exp\big(\big[r^2 + r\cdot \ln(2c)\big]/|\ln(b)|\big)$$

Hence

$$\mathbb{P}(\mathcal{O} \geqslant k) \leqslant 2 \inf_{r>0} \exp\big(\big(r^2 + r\cdot [\ln(2c) - k|\ln(b)|]\big)/|\ln(b)|\big)$$

$$= 2\exp\big(- (|\ln(b)|/4)\big[k - (\ln(2C)/|\ln(b)|)\big]^2\big)$$

Example: $(\mathbf{A_n})_{\mathbf{n}\in\mathbb{N}}$ independent and $\mathbb{P}(\mathbf{A_n}) \leqslant \mathbf{b^{n^2}}$, $\mathbf{b} \in (\mathbf{0}, \mathbf{1})$

$$\mathbb{E}[\mathbf{e^{r\mathcal{O}}}] \leqslant \mathbf{2}\exp(\frac{\sqrt{\mathbf{r^3} - \mathbf{r^2}\ln(\mathbf{2})}}{\sqrt{|\ln(\mathbf{b})|}}).$$

and

$$\mathbb{P}(\mathcal{O} \geqslant \mathbf{k}) \leqslant \mathbf{2}\exp\Big(-((\frac{\mathbf{2}|\ln(\mathbf{b})|}{\mathbf{3}})^{\mathbf{2}} \cdot \frac{\mathbf{k^3}}{\mathbf{3}} + \sqrt{\mathbf{2}})/\sqrt{|\ln(\mathbf{b})|}\Big) \qquad \mathbf{k} \geqslant \mathbf{1}.$$

More applications:

- Random graphs (Coloring numbers, clique numbers)
- A.s. invariance principles (A.s. versions of the CLT)

Literature:

1. L.F. Estrada, M.A.H.:
   Moment estimates in the first Borel-Cantelli Lemma with applications to mean deviation frequencies
   Statistics and Probability Letters 190 (2022) 109636

2. L.F. Estrada, M.A.H., A. Steinicke:
   On the tradeoff between almost sure error tolerance versus mean deviation frequency in martingale convergence
   `https://arxiv.org/abs/2310.09055`

3. M.A.H., A. Steinicke:
   Deviation frequencies of Brownian path property approximations
   `https://arxiv.org/abs/2302.04115`