

Chapter 2: One big jump phenomenon

1) A maximal inequality 2) A local estimate 3) A one-big jump principle

Here we identify a regime where atypical events of the form  $\sum_{n \to \infty} \sum_{n \to \infty} \sum_{n$ 

1) A maximal inequality We start with pooving an inequality for future use.

Theorem [Field Nagaeu '71, Denisor-Dieler - Schweer 108] Assume that X is a R-valued rv with  $\mathbb{E}[X^2] < \infty$ ,  $\mathbb{E}[X] = 0$ . Let  $(X_i)_{i>1}$  beind having some law as X. Set  $S_n = X_1 + \dots + X_n$ . There  $\exists K>0$  s.t.  $\forall r>1, z>0$  and c>1: $\mathbb{P}(S_n > x \sqrt{n}, X_1 \leq \sqrt{n}, \dots, X_n \leq \sqrt{n}) \leq K \exp(-\frac{x}{c})$ 

 $\begin{array}{l} \displaystyle \Pr_{\text{reg}} : \text{ The idea is to introduce the frameworkd random when } & \sum_{i=1}^{n} X_i \mathbf{1}_{X_i \text{ scrit}} \text{ inded}, \\ \displaystyle \mathbb{P}(S_n > 2 \, \sqrt{n}, X_i \le (\sqrt{n}, \cdots, X_n \le (\sqrt{n})) \le \mathcal{D}(S_n \geqslant 2 \, \sqrt{n}). \\ \displaystyle \mathbb{T}_0 \text{ bound this probability we use the "exponential Markov" inequality: \\ \displaystyle \mathbb{P}(S_n \geqslant 2 \, \sqrt{n}) = \mathbb{P}(e^{\frac{2\pi}{2}m} \geqslant e^{\frac{\pi}{2}}) \le e^{\frac{\pi}{2}} \mathbb{E}[e^{\frac{2\pi}{2}m}]^n \text{ with } X = \times \mathbf{1}_{X \le (\sqrt{n})} \quad (A \times \text{ depends on } n] \\ \displaystyle \mathbb{N}e \text{ show that } \mathbb{E}[e^{\frac{2\pi}{2}m}] = 1 + \mathcal{O}(\frac{1}{n}) \text{ and the result will follow.} \\ \displaystyle \mathbb{T}e \text{ idea (s to write } e^{\frac{\pi}{2}} = 2 + 2 + 2^2 2(x) \text{ with } n(x) = \frac{e^2 - 1 - 2}{2x^2}, \text{ so that} \\ \displaystyle \mathbb{E}[e^{\frac{2\pi}{2}(x_n)}] = 1 + \mathbb{E}[\frac{2\pi}{2}] + \mathbb{E}[\frac{2\pi}{2}(x_n)] \\ \displaystyle \mathbb{N}n \quad \mathbb{N}n \quad \mathbb{P}(x_n) = \mathcal{O}(\frac{1}{n}) \\ \displaystyle \mathbb{N}e \text{ show that } m_n = \mathcal{O}(\frac{1}{n}) \text{ and } s_n = \mathcal{O}(\frac{1}{n}) \\ \displaystyle \mathbb{E}[e^{\frac{2\pi}{2}x_n} : since n \text{ is bounded on } (-\infty, 2], we have \quad S_n = \mathcal{O}(\mathbb{E}[\sum_{n=1}^{2}]) (here we rese (s,1)) \\ \displaystyle \mathbb{E}[x_n] \in \mathbb{E}[x_n]^2 : \mathbb{E}[x_n]^2, \quad \mathbb{E}[x_n] = \mathcal{O}(\frac{1}{n}). \end{array}$ 

$$\begin{array}{l} \overline{F_{x}} & m_{n}: \text{ write } m_{n} = \frac{1}{c\sqrt{n}} \mathbb{E}[X \downarrow_{X \leq c\sqrt{n}}] = -\frac{1}{c\sqrt{n}} \mathbb{E}[X \downarrow_{X > c\sqrt{n}}] \text{ because } \mathbb{E}[X] = 0 \\ \hline \\ \overline{P_{u}} \downarrow_{t} \mathbb{E}[X \downarrow_{t} \downarrow_{t}] = \frac{1}{c\sqrt{n}} \mathbb{E}[X \downarrow_{t} \downarrow_{t}] = \frac{1}{c\sqrt{n}} \mathbb{E}[X^{2}] \\ \hline \\ \overline{Thus} m_{n} = \mathcal{O}(\frac{1}{n}) \quad (howe we was c \geq 1) \\ \end{array}$$

Remark The result is folse with "c > 0" instead of " $c \ge 1$ ". Indeed, take  $B(X_1 = \pm 1) = \frac{1}{2}$ , x = 1,  $c = n^{-1/4}$ . Then  $B(S_n \ge x \sqrt{n}, X_1 \le c \sqrt{n}, ..., X_n \le c \sqrt{n}) = B(S_n \ge \sqrt{n}) \xrightarrow{n \to \infty} B(N(Q_1) \ge 1)$ but  $K \exp(-\frac{x}{c}) \xrightarrow{n \to \infty} 0$ 

## 2) A local estimate

We introduce a gramework which will allow to freat "Sn > an" and "Sn = an" at the same time. Fix  $T \in (0, +\infty]$  and set  $\Delta = EO, T$ . For mGB we set  $\Delta_m = m + \Delta = Em, m + T$ .

• If 
$$T = \infty$$
,  $B(X_1 \in \Delta u) = B(X_1 \ge u) \sim \frac{C}{u \ge \omega}$   
• If  $T = \infty$ ,  $B(X_1 \in \Delta u) = B(X_1 \ge u) \sim \frac{C}{u \ge \omega}$   
• If  $T < \infty$ ,  $B(X_1 \in \Delta u) = B(X_1 \in [u, u+T)) \sim \frac{C}{u \ge \omega}$ 

Example If X, is Z-valued and if  $B(X_1 = n) \sim \frac{C}{n^{1+p}}$  then X, satisfies (HA) for T=1 and T=  $\infty$ , but not T=1.5 (indeed,  $B(X_1 \in En, n+3.5)$ )  $\sim \frac{2C}{n^{1+p}}$  and  $B(X_1 \in En+0, 1, n+3.6)$ ]  $\sim \frac{C}{n^{1+p}}$ ) Remark One can check that if X, satisfies (HA) then X, has finite variance (because p>2)

Setting Assume that for (>0, T>0, B>2, 
$$B(X_1 \in EU, U+T)) \sim \frac{C}{u + \omega u^{1+p}}$$
 then  $B(X_1 \ge u) \sim \frac{C}{BT} \cdot \frac{1}{u^p}$   
In particular, if  $X_1$  satisfies (Ha) for  $T \le +\infty$  then it satisfies (Ha) for  $T = \infty$ 

Proof: Exercise for vert week End of lecture 5

We will use several times the following fact  

$$\begin{bmatrix}
Fat (t) Under (H_{\Delta}), B(X_{i} \in \Delta u) & B((X_{i} \in \Delta u + y) uniformly in 1y] \leq \frac{u}{u + \omega}, \\
\stackrel{ie}{=} \sup_{\substack{X \in \mathcal{U} \\ Y \in \mathcal{U} \\ enum}} \left| \frac{B(X_{i} \in \Delta u + y)}{B(X_{i} \in \Delta u + y)} - 1 \right| \xrightarrow{u \to \omega}$$

Theorem (Doney 183, Nagaen 157)  
Assume that X, satisfies (14a) and that ETX, 3:0. Fix so. Then, caniformly in 
$$m \ge n$$
,  
 $\mathbb{P}(S_n \in \Delta_m) \underset{n \to \infty}{\longrightarrow} n \mathbb{P}(X_1 \in \Delta_m)$   
That is, sup  $\left| \frac{\Re(S_n \in \Delta_m)}{n \Re(X_1 \in \Delta_m)} - 1 \right| \xrightarrow{n \to \infty} 0$   
Intuition: Sue  $\Delta_m$  Supredly hoppers when one of the usuaps is in  $\Delta_n$  (this will be made precise laber)  
 $\mathbb{P}_{rood}$  Set  $\overline{m} = \frac{m}{m_n}$ . Write  $n \mathbb{P}_n^m \in \mathbb{R}(S_n \in \Delta_m) \le n \mathbb{P}_n^m + \mathbb{P}_n^{n,n} + \mathbb{P}_n^{n,n}$  with  
 $\mathbb{P}_n^{S,n} \in \mathbb{C}$  Such  $m_n = \frac{m}{m_n}$ . Write  $n \mathbb{P}_n^m \in \mathbb{R}(S_n \in \Delta_m) \le n \mathbb{P}_n^{n,n} + \mathbb{P}_n^{n,n} + \mathbb{P}_n^{n,n}$  with  
 $\mathbb{P}_n^{S,n} = \mathbb{R}(S_n \in \Delta_n, X_n \geqslant \overline{m}, \max_{n \ge n \le n}, X_n \ge \overline{m})$   
 $\mathbb{P}_n^{S,n} = \mathbb{R}(S_n \in \Delta_n, x_n \geqslant \overline{m}, \max_{n \ge n \le n}, X_n \ge \overline{m})$   
Note show that antifuely in  $m \ge n$ . Notice  $\mathbb{R}(X_n \in \Delta_n)$ ,  $\mathbb{P}_n^{n,n} = n(n\mathbb{R}(X_n \in \Delta_n))$ ,  $\mathbb{P}_n^{n,n} = n(n\mathbb{R}(X_n \in \Delta_n))$   
Recall that  $\mathbb{R}(X_n \in \Delta_n) \le \frac{1}{m} \frac{1}{(8\pi - \infty)} = \frac{n^2}{m^2}$  when  $\frac{1}{m} \frac{1}{m}$  for some two by the linea.  
 $\mathbb{E}_n \mathbb{P}_n^{S,n} \le \mathbb{R}(n) \xrightarrow{(n)} \mathbb{P}(X_n \ge \overline{m}, X_n \ge \overline{m})$   
 $\mathbb{E}_n \mathbb{R}_n^{S,n} \le \mathbb{R}(S_n \ge M_n, \max_{n \ge n}, X_n \ge \overline{m}) = \frac{n^2}{m^2} \mathbb{R}(n) \xrightarrow{(n \ge n \le N_n \log n)} \le C(\frac{n}{m^2} \log(n) \stackrel{(n \ge n \le N_n \log n)}{m^{N-1}} = n(n\mathbb{R}(X_n \in M_n))$   
 $\mathbb{E}_n \mathbb{R}_n^{S,n} \le \mathbb{R}(S_n \ge M_n, \max_{n \ge N_n \log n)} \xrightarrow{(n \ge n \le N_n \log n)} = n(\frac{n}{m^{N-1}})$   
 $\mathbb{E}_n \mathbb{R}_n^{S,n} \le \mathbb{R}(S_n \ge M_n, \max_{n \ge N_n \log n)} \le K \exp(-\frac{m}{m}) = K \exp(-\ln(n)^3) = n(\frac{n}{m^{N-1}})$   
 $\mathbb{E}_n \mathbb{R}_n^{S,n} \le \mathbb{R}(S_n \ge M_n, \max_{n \ge N_n \log n} X_n \le \mathbb{R}) \le \mathbb{R}_n \mathbb{R}_n \mathbb{R}_n^{S,n} = \mathbb{R}(n)^{S,n} = \mathbb{R}(-\frac{n}{m^{N-1}})$   
 $\mathbb{E}_n \mathbb{R}_n^{S,n}$  This is none delically we will be consider cases according to the walk of  $S_{n-1}$ .

Since 
$$\frac{Sn}{m}$$
 converges in distribution (CLT),  $\frac{Sn}{n^{3/4}} \xrightarrow{\mathbb{P}} \partial$ . Then write  $P_{\Delta}^{m,n} = Q_{\Delta}^{m,n} + Q_{2}^{m,n} + Q_{3}^{m,n}$  with

$$\begin{split} & \mathbb{Q}_{1}^{Nn} = \mathbb{B}\left(\sum_{n \in \Delta_{m}} \sum_{N_{n} \neq m} \frac{1}{1} \sum_{\substack{n \in \Delta_{m} \neq m}} \frac{1}{1} \sum_{\substack{n \in \Delta$$

Remark There has been quite some work to find the "best" sequence  $m_n$  such that the theorem helds unifounly for  $m \ge m_n$  (we have chown that  $m_n \ge \varepsilon_n$  works).

We beep the previous gramework:  $T \in (0, +\infty]$ , we set  $\Delta = EO, T$ ),  $\Delta u = Eu, u+T$ ) and essume that  $X_1$  satisfies  $(H_A)$  and that  $E[X_1]=0$ .

$$\frac{N_{otahon}}{N_{otahon}} \cdot \text{Set } V_n = \min\{2 \leq j \leq n : X_j = \max(X_1, \dots, X_n) \}$$

$$\cdot \text{Set } (\widehat{X}_1, \dots, \widehat{X}_{n-1}) = (X_{11} \dots, X_{N_{n-1}}, X_{N_{n+1}} \dots, X_n)$$

Fix a sequence (In) such that lines in so.

0 m

Theorem Love big jump principle, Armendariz & Loolabis '11) We have  $d_{\tau v} ((\hat{X}_{1}, ..., \hat{X}_{n-1}))$  under  $B[\cdot|S_n \in D_{xd}], (X_{1}, ..., X_{n-1})) \xrightarrow{n \to \infty} 0$ , that is step  $|B((\hat{X}_{1}, ..., \hat{X}_{n-1}) \in A|S_n \in D_{xn}) - B((X_{1}, ..., X_{n-1}) \in A)| \xrightarrow{n \to \infty} 0$  $A \in B(\mathbb{R}^{n-1})$ 

This means that under  $\mathbb{P}(\cdot | S_n \in S_{\infty_n})$ , once the biggest jump is removed, the remaining is are esymptotically identify with some law as  $X_i$ !

In practice, to show that a property holds with probability tranking to 0 or 1 for  $(\hat{X}_{1,-1}, \hat{X}_{m})$ under  $\mathbb{P}(\cdot | S_n \in \Delta_{x_n})$  one can show that it holds for  $(X_{1,-1}, X_{n-1})$  (which are iid!)

Before proving this, let us see a striking consequence.  
Cotollary Set 
$$D_n = \max(X_{11}, Y_n)$$
 and let  $D_n^{(2)}$  be the second largest element of  $(X_{11}, Y_n)$ . Then:  
(1)  $\forall u \ge 0$   $\mathbb{B}\left(\frac{D_n^{(2)}}{n^{1/p}} \le u \mid S_n \in \Delta_{X_n}\right) \xrightarrow{n \to \infty} e_X p\left(-\frac{c_0}{u^B}\right)$  with  $c_0 = \int c_1 g \tau = \infty$   
Now assume  $\tau \ge \infty$   
(2)  $D_n - X_n$  under  $\mathbb{B}\left(\cdot \mid S_n \in \Delta_{X_n}\right) \xrightarrow{(a)} \mathbb{N}(c_1)$  with  $\sigma^2 = \operatorname{Von}(X_1)$ 

To do this, set 
$$E_n = \begin{cases} a = (a_{1,...}, a_{n-1}) \in \mathbb{R}^{n-1}$$
:  $1 = (a_{1},...,a_{n-1}| \le n^{n-1})$  and  $\max_{1 \le n} a_{1 \le n} \le n^{n-1} \end{cases}$   
We decle  $D$   $\mu_n(E_n^{-1}) \le B(1S_{n-1}| > n^{n-1}) + B(100 \times X_j > n^{n-1})$   
. He first term is o(i) some  $\frac{S_{n-1}}{V_{\infty}}$  converges in distribution  
. The second term is  $1 - (1 - p(X_1 > n^{n+1}))^{n-1} \longrightarrow D$  is the  $\frac{n}{n+1} \longrightarrow D$  ( $p > 2$ )  
 $\sum_{i=1}^{n} \frac{1}{n}$   
No decle  $B: \text{ for } a = (a_{1,...,a_{n-1}}) \in \mathbb{R}^{n-1}$ , observe that  
 $E(X_{2,...,N}, X_{n-1}) = a, S_n \in A_{X_n} \le - \bigcup_{i=1}^{n} \le (X_{1,...,N}, X_{i+1}, ..., X_n) = a, X_i \in A_{X_n-a_{n-1}} \le a_{n-1} \end{cases}$   
and that the version is disjoict. Then  
 $\widehat{\mu}_n(a) = \mu_n(a) \times \frac{n \cdot B(X_1 \in A_{X_n} - a_{n-1} - a_{n-1})}{1 = 1 + \varepsilon(a)} \xrightarrow{\text{Prime}} 1 = 1 + \varepsilon(a)}$   
This entrals  $\sup_{A \in E_n} |\widehat{\mu}_n(A) - \mu_n(A)| = 1 + \varepsilon(a)$ .  
This entrals  $\sup_{A \in E_n} |\widehat{\mu}_n(A) - \mu_n(A)| \le \sum_{a \in E_n} |\widehat{\mu}_n(a)| \le 1 + \varepsilon(a)$ .  
 $\sum_{A \in E_n} |\widehat{\mu}_n(A) - \mu_n(A)| \le \sum_{a \in E_n} |\widehat{\mu}_n(a)| \le 1 + \varepsilon(a)$ .  
Eval of  $k$  the definition of  $A \in E_n$ .  
 $\sum_{A \in E_n} |\widehat{\mu}_n(A) - \mu_n(A)| \le \sum_{a \in E_n} |\widehat{\mu}_n(a)| \le 1 + \varepsilon(a)$ .