• **Proposition** (Symmetric r.v.'s and Their Partial Maxima)

Let $\{X_n, n\geq 1\}$ be indep., symmetric r.v.'s. and set $S_n = \sum_{k=1}^{n} X_k$.

Then

$$P\left(\max_{1\leq k\leq n}|X_k| > 2x\right) \leq P\left(\max_{1\leq k\leq n}|S_k|>x\right) \leq 2P(|S_n|>x), \quad x>0.$$

Pf: $\because |X_n| = |S_n - S_{n-1}| \leq |S_n| + |S_{n-1}|$ by triangle inequality

$$\therefore \max_{1\leq k\leq n}|X_k| \leq 2\max_{1\leq k\leq n}|S_k|$$

$$\Rightarrow \left[\max_{1\leq k\leq n}|X_k| > 2x\right] \Rightarrow \left[2\max_{1\leq k\leq n}|S_k| > 2x\right]$$

Hence $P\left[\max_{1\leq k\leq n}|X_k|>2x\right] \leq P\left[\max_{1\leq k\leq n}|S_k|>x\right]$.

Since $\{X_n, n\geq 1\}$ are symmetric, the Lévy inequality implies that

$$P\left[\max_{1\leq k\leq n}|S_k|>x\right] \leq 2P[|S_n|>x].$$

#

• **Proposition** (Moments and Tails)

(i) Let $r>0$. Suppose that X is a non-negative r.v.. Then

$$E(X^r)<\infty \Rightarrow x^r P(X>x) \to 0 \text{ as } x\to\infty.$$

(ii) but not necessarily conversely.

Pf: (i) $$x^r P[X>x] = x^r \int_x^\infty dF_X(y)$$

$$\leq \int_x^\infty y^r\, dF_X(y) \to 0 \text{ as } x\to\infty$$

$$\because E(X^r) = \int_0^\infty y^r\, dF_X(y) < \infty.$$

(ii) We now show that the converse is not true.

$$f_X(x) = \frac{c}{x^{r+1}\log x}, \quad x>e$$

$$x^r P(X>x) \sim c x^r \frac{1}{x^r \log x} = \frac{c}{\log x} \to 0 \text{ as } x\uparrow\infty$$

$$\because \lim_{x\to\infty} \frac{P[X>x]}{\frac{1}{x^r\log x}} = \lim_{x\to\infty} \frac{\frac{-c}{x^{r+1}\log x}}{\cancel{\frac{-r}{x^{r+1}\log x}} - \cancel{\frac{1}{x^{r+1}(\log x)^2}}} = \frac{c}{r} = C$$

But $E(X^r) = c\int_e^\infty \frac{dx}{x\log x} = \infty$.

#

- **Theorem** (A Weak Law for Partial Maxima)

Let $\{X_n, n\geq 1\}$ be iid r.v.'s with $P(X_n=0)<1$, and set $Y_n=\max_{1\leq k\leq n}|X_k|$

If $\{b_n, n\geq 1\}$ is a seq s.t. $b_n\geq 0$ and $b_n\uparrow\infty$, then

$$\frac{Y_n}{b_n}\xrightarrow{P}0 \text{ as } n\to\infty \Leftrightarrow n\,P(|X_n|>b_n\varepsilon)\to 0 \text{ as } n\to\infty,\ \text{for } \varepsilon>0$$

In particular, for $r>0$,

$$\frac{Y_n}{n^{1/r}}\xrightarrow{P}0 \text{ as } n\to\infty \Leftrightarrow n\,P[|X_n|>n^{1/r}]\to 0 \text{ as } n\to\infty.$$

Pf: $P[|\frac{Y_n}{b_n}|>\varepsilon]=P[\frac{Y_n}{b_n}>\varepsilon]$

$$\leq P\left(\bigcup_{k=1}^{n}[\tfrac{|X_k|}{b_n}>\varepsilon]\right)$$

$$\leq n\,P[\tfrac{|X_n|}{b_n}>\varepsilon] \text{ by finite subadditivity and iid.}$$

$P[|\frac{Y_n}{b_n}|>\varepsilon]=P[\frac{Y_n}{b_n}>\varepsilon]$

$$=1-P[\tfrac{Y_n}{b_n}\leq\varepsilon]$$

$$=1-\left(P[\tfrac{|X_n|}{b_n}\leq\varepsilon]\right)^n \quad \text{by the } X_n \text{ are iid.}$$

$$=1-\left(1-P[\tfrac{|X_n|}{b_n}>\varepsilon]\right)^n$$

$$\geq(1-\delta)\,n\,P[\tfrac{|X_n|}{b_n}>\varepsilon], \text{ set } \delta=\tfrac{1}{2}$$

$\because$ $0\leq a_n<\delta<1$, $\forall n\in\mathbb{N}$,

$(1-a_n)^n\to 1$ as $n\to\infty \Leftrightarrow na_n\to 0$ as $n\to\infty$, and

$na_n<\delta(1+\delta)<1$ and

$(1-\delta)na_n\leq 1-(1-a_n)^n\leq\frac{na_n}{1+\delta}$

Hence,

$$\tfrac{1}{2}n\,P(|X_n|>b_n\varepsilon)\leq P(|\tfrac{Y_n}{b_n}|>\varepsilon)\leq n\,P(|X_n|>b_n\varepsilon)$$

$$\frac{Y_n}{b_n}\xrightarrow{P}0\Leftrightarrow n\,P(|X_n|>b_n\varepsilon)\to 0 \text{ as } n\to\infty,\ \text{for } \varepsilon>0.$$

#

- **Theorem**

Let $\{X_n, n\geq 1\}$ be indep r.v.'s, then

$$\frac{Y_n}{b_n}\xrightarrow{P}0 \text{ as } n\to\infty \Leftrightarrow \sum_{k=1}^{n}P(|X_k|>b_n\varepsilon)\to 0 \text{ as } n\to\infty,\ \text{for } \varepsilon>0.$$

Pf: By $(1-a_1)(1-a_2)\cdots(1-a_n)\to 1$ as $n\to\infty \Leftrightarrow \sum_{i=1}^{n}a_i\to 0$ as $n\to\infty$. #

• Theorem (The Law of Large Numbers: $\frac{S_n - E[S_n]}{n} = \bar{X}_n - E(\bar{X}_n) \xrightarrow[a.s.]{L_2, P} 0$; $\bar{X}_n \xrightarrow[a.s.]{L_2, P} E(\bar{X}_n)$)

If $\{X_j, j \geq 1\}$ are (independent or pairwise indep or) uncorrelated, and their second moments have a common bound; $E[|X_n|^2] < M' \ \forall n$,

(a) $\frac{S_n - E(S_n)}{n} \to 0$ in L_2 and hence in pr.

(b) $\frac{S_n - E(S_n)}{n} \to 0$ a.s.

Pf: (a) $E\left(\frac{S_n - E(S_n)}{n}\right)^2$

$$= \frac{1}{n^2} E(S_n - E(S_n))^2 = \frac{1}{n^2} E\left(\sum_1^n (X_j - E(X_j))\right)^2$$

$$= \frac{1}{n^2}\left\{\sum_1^n E(X_j - E(X_j))^2 + 2 \sum_{1 \leq j < k \leq n} \overbrace{E(X_j - E(X_j))(X_k - E(X_k))}^{0}\right\}$$

$\because$ (indep or pairwise indep or) uncorrelated

$$= \frac{1}{n^2} \sum_1^n \sigma^2(X_j) \leq \frac{nM}{n^2} \quad \because \sigma^2(X_j) \leq M \ \forall j$$

$\to 0$ as $n \to \infty$.

Thus $\frac{S_n - E(S_n)}{n} \to 0$ in L_2 and hence in pr.

(b) • WLOG, assume $E(X_j) = 0$ for each j.

$E(S_n^2) = \sigma^2(S_n) \leq nM$ from above

• We now want to show that $\frac{S_n}{n} \to 0$ a.s.

$$P\left[\frac{|S_n|}{n} > \varepsilon\right] = P[|S_n| > n\varepsilon] \leq \frac{E(S_n^2)}{n^2 \varepsilon^2} \leq \frac{nM}{n^2 \varepsilon^2}.$$

If we sum this over n, the resulting series on the right diverges.

• However, if we confine ourselves to the subsequence $\{n^2\}$, then

$$\sum_n P\left[\frac{|S_{n^2}|}{n^2} > \varepsilon\right] \leq \sum_n \frac{M}{n^2 \varepsilon^2} < \infty$$ ⇒ "method of subsequences"!

Hence by the Borel-Cantelli lemma we have

$P\left(\left[\frac{|S_{n^2}|}{n^2} > \varepsilon\right] \text{ i.o.}\right) = 0$. Hence $\frac{S_{n^2}}{n^2} \to 0$ a.s.

• We must show S_k does not differ from the nearest S_{n^2} to to make any real difference.

(next pg. cont.)

That is, for $n^2 \le k < (n+1)^2$,

$$\frac{|S_k|}{k} = \frac{|S_k - S_{n^2} + S_{n^2}|}{k} \le \frac{|S_k - S_{n^2}|}{n^2} + \frac{|S_{n^2}|}{n^2}$$

- Set $D_n = \max_{n^2 \le k < (n+1)^2} |S_k - S_{n^2}|$, then

$$\frac{|S_k|}{k} \le \frac{|S_{n^2}| + D_n}{n^2} \quad \text{for } n^2 \le k < (n+1)^2.$$

We now show that $\frac{D_n}{n^2} \to 0$ a.e.

$$E(D_n^2) \le 2n\, E(|S_{(n+1)^2} - S_{n^2}|^2) = 2n \sum_{j=n^2+1}^{(n+1)^2} \sigma^2(X_j) \le 4n^2 M.$$

$\because$ $(n+1)^2 - 1 - n^2 = 2n$ terms

$$P\left[\frac{|D_n|}{n^2} > \varepsilon\right] = P[|D_n| > n^2 \varepsilon] \le \frac{4n^2 M}{n^4 \varepsilon^2} = \frac{4M}{n^2 \varepsilon^2}$$

$$\sum_n P\left[\frac{|D_n|}{n^2} > \varepsilon\right] \le \sum_n \frac{4M}{n^2 \varepsilon^2} < \infty$$

Thus $P\left(\left[\frac{D_n}{n^2} > \varepsilon\right] \text{ i.o.}\right) = 0$, Hence $\frac{D_n}{n^2} \to 0$ a.s.

The theorem is proved.

#

- <u>Definition</u> (Equivalence)

Two sequences $\{X_n\}$ and $\{X_n'\}$ are "equivalent" if

$$\sum_n P[X_n \neq X_n'] < \infty.$$

In order to drop any assumption on the second moment, we often check $\{X_n\}$ and $\{X_n 1_{[|X_n| \le n]}\}$ are equivalent sequences, since the later are bounded & have finite moments.

- <u>Theorem</u> (Equivalence).

Suppose the two sequences $\{X_n\}$ and $\{X_n'\}$ are equivalent. Then

(1) $\sum_n (X_n - X_n')$ converges a.s.

(2) The two series $\sum_n X_n$, $\sum_n X_n'$ converge a.s. together or diverge a.s. together; that is

$$\sum_n X_n \text{ converges a.s. iff } \sum_n X_n' \text{ converges a.s.}$$

(3) If there exists a sequence $\{a_n\}$ s.t. $a_n \uparrow \infty$ and if there exists a r.v. X s.t. $\frac{1}{a_n} \sum_{k=1}^{n} X_k' \xrightarrow{a.s.} X$,

then also $\frac{1}{a_n} \sum_{k=1}^{n} X_k \xrightarrow{a.s.} X$ (to the same r.v. X).

Pf: (1). Since $\{X_n\}$ and $\{X_n'\}$ are equivalent $\Rightarrow \sum_n P[X_n \neq X_n'] < \infty$,

the Borel-Cantelli lemma implies that $P([X_n \neq X_n'] \text{ i.o.}) = 0$,

$\Rightarrow P(\liminf_{n\to\infty} [X_n = X_n']) = 1$.

So for $\omega \in \liminf_{n\to\infty} [X_n = X_n']$, we have $X_n(\omega) = X_n'(\omega)$ for $n \ge N(\omega)$

Thus $\sum_n (X_n - X_n')$ converges a.s.

(2). Note $\sum_{n=N}^{\infty} X_n(\omega) = \sum_{n=N}^{\infty} X_n'(\omega)$.

Thus $\sum_n X_n$ converges a.s. iff $\sum_n X_n'$ converges a.s.

(3) $\because \sum_{k=1}^{n} (X_k - X_k')$ converges a.s. and $a_n \uparrow \infty$,

$$\therefore \frac{1}{a_n} \sum_{k=1}^{n} X_k = \frac{1}{a_n} \sum_{k=1}^{n} (X_k - X_k') + \frac{1}{a_n} \sum_{k=1}^{n} X_k' \xrightarrow{a.s.} 0 + X = X.$$

#

• <u>Theorem</u> (The General WLLN without Moment Assumption)

Let $\{X_n\}$ be a sequence of "independent" r.v.'s, $S_n = \sum_{k=1}^{n} X_k$.

Let $\{b_n\}$ be a sequence s.t. $b_n \uparrow \infty$.

Suppose that we have

(i) $\sum_{k=1}^{n} P[|X_k| > b_n] \to 0$; i.e. $\sum_{k=1}^{n} \int_{[|X_k|>b_n]} dP = o(1)$.

(ii) $\sum_{k=1}^{n} \frac{1}{b_n^2} E[X_k^2 1_{[|X_k| \le b_n]}] \to 0$; i.e. $\sum_{k=1}^{n} \frac{1}{b_n^2} \int_{[|X_k| \le b_n]} X_k^2 dP = o(1)$.

Then we have

$$\frac{S_n - a_n}{b_n} \to_p 0, \quad \text{where } a_n = \sum_{k=1}^{n} E[X_k \underbrace{1_{[|X_k| \le b_n]}}_{\text{truncated by } \{b_n\}}]$$

Pf: • Define $Y_{nk} = X_k 1_{[|X_k| \le b_n]}$, and $S_n' = \sum_{k=1}^{n} Y_{nk}$.

• The condition (i) means that

$$\sum_{k=1}^{n} [Y_{nk} \ne X_k] = \sum_{k=1}^{n} P[|X_k| > b_n] \to 0.$$

Hence $\{Y_{nk}\}$ and $\{X_k\}$ are equivalent if $\{X_k\}$ is truncated by $\{b_n\}$.

$$\because P[S_n' \ne S_n] = P[\bigcup_{k=1}^{n} (Y_{nk} \ne X_k)]$$
$$\le \sum_{k=1}^{n} P[Y_{nk} \ne X_k] \to 0$$

$$\therefore P[|S_n' - S_n| > \epsilon] \le P[S_n' \ne S_n] \to 0$$

$$\Rightarrow S_n' - S_n \to_p 0.$$

• The condition (ii) means that

$$\sum_{k=1}^{n} \frac{1}{b_n^2} E[X_k^2 1_{[|X_k| \le b_n]}] = \sum_{k=1}^{n} E[(\frac{Y_{nk}}{b_n})^2] \to 0.$$

Since $\{Y_{nk}, 1 \le k \le n\}$ are independent r.v.'s,

$$Var[\frac{S_n'}{b_n}] = \sum_{k=1}^{n} Var[\frac{Y_{nk}}{b_n}] \le \sum_{k=1}^{n} E[(\frac{Y_{nk}}{b_n})^2] \to 0. \text{ Then}$$

$$P[|\frac{S_n' - E[S_n']}{b_n}| > \epsilon] \le \frac{1}{\epsilon^2} Var[\frac{S_n'}{b_n}] \to 0.$$

$$\Rightarrow \frac{S_n' - E[S_n']}{b_n} \to_p 0.$$

• Since $\frac{S_n - E[S_n']}{b_n} = \frac{S_n - S_n'}{b_n} + \frac{S_n' - E[S_n']}{b_n} \to_p 0 + 0 = 0$, and

$$E[S_n'] = E[\sum_{k=1}^{n} Y_{nk}] = \sum_{k=1}^{n} E[X_k 1_{[|X_k| \le b_n]}] \equiv a_n,$$

$$\Rightarrow \frac{S_n - a_n}{b_n} \to_p 0.$$

\#

• <u>Theorem</u> (Khintchin's WLLN under the First Moment Assumption)
Let $\{X_n, n\geq 1\}$ be "iid" r.v.'s with "finite mean" μ, and let $S_n = \sum_{j=1}^{n} X_j$.
Then $\frac{S_n}{n} \xrightarrow{P} \mu$.

Pf: • By means of the General WLLN, we show that $(b_n = n)$

(1) $\sum_{j=1}^{n} P[|X_j| > n] \to 0$

(2) $\frac{1}{n^2} \sum_{j=1}^{n} E(X_j^2 1_{[|X_j| \leq n]}) \to 0$.

• For (1) observe that

$$n P[|X_j|>n] = E(n 1_{[|X_j|>n]}) \quad \text{by } \{X_n\} \text{ iid}$$
$$\leq E(|X_j| 1_{[|X_j|>n]}) \to 0$$

Since $|X_j| 1_{[|X_j|\leq n]} \leq |X_j| \in L_1$ and $|X_j| 1_{[|X_j| \leq n]} \to 0$ as $n\to\infty$,
the DCT implies that $E(|X_j| 1_{[|X_j|>n]}) \to 0$ as $n\to\infty$.

• For (2), for any $\epsilon > 0$,

$$\frac{1}{n^2} n E(X_1^2 1_{[|X_1|\leq n]}) \leq \frac{1}{n}\left(E(X_1^2 1_{[|X_1|\leq \epsilon\sqrt{n}]}) + E(X_1^2 1_{[\epsilon\sqrt{n} \leq |X_1| \leq n]})\right) \quad \text{by } \{X_n\} \text{ iid}$$
$$\leq \frac{\epsilon^2 n}{n} + \frac{1}{n} E(n|X_1| 1_{[\epsilon\sqrt{n} \leq |X_1| \leq n]})$$
$$\leq \epsilon^2 + E(|X_1| 1_{[\epsilon\sqrt{n} \leq |X_1|]}) \to \epsilon^2 \text{ as } n\to\infty,$$

• So applying the General WLLN we conclude that

$$\frac{S_n - n E(X_1 1_{[|X_1|\leq n]})}{n} \xrightarrow{P} 0. \quad \text{where } a_n = \sum_{k=1}^{n} E[X_k 1_{[|X_k| \leq n]}] \text{ and by iid}$$

Since $\left|\frac{n E(X_1 1_{[|X_1|\leq n]})}{n} - E(X_1)\right| = E(X_1 1_{[|X_1|>n]}) \to 0$ as $n\to\infty$,
the result follows.

Pf 2: For $t \in \mathbb{R}$,

$$\varphi_{\bar{X}_n}(t) = E(e^{it\bar{X}_n}) = E(e^{\frac{it\sum X_i}{n}}) = [E(e^{it\frac{X_1}{n}})]^n \quad \because X_n\text{'s iid.}$$
$$= [\varphi_X(\tfrac{t}{n})]^n = [1 + \tfrac{it}{n}\mu + o(\tfrac{t}{n})]^n \quad \because E|X_1| < \infty \text{ and } E(X_1) = \mu$$
$$\to e^{it\mu} \quad \text{as } n\to\infty$$
$$= \varphi_\mu(t)$$

Hence the continuity thm implies that $\bar{X}_n \Rightarrow \mu$ and hence $\bar{X}_n \xrightarrow{P} \mu$.

#

• <u>Theorem</u> (Kolmogorov-Feller WLLN without First Moment Assumption)

Suppose that $\{X_n, n\ge 1\}$ are "iid" r.v.'s. Then

$$n P(|X_1| > n) \to 0$$

iff

$$\frac{S_n - n E(X_1 \mathbb{1}_{[|X_1|\le n]})}{n} \xrightarrow{P} 0. \quad \text{truncated by } \{n\}$$

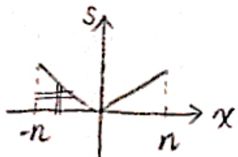
Pf: ($\Rightarrow$) We want to show that the 2 conditions of the general WLLN hold.

If $n P(|X_1| > n) \to 0$, (with $b_n = n$)

(1) $\sum_{k=1}^{n} P(|X_k| > n) = n P(|X_1| > n) \to 0$, by $\{X_n\}$ are iid

(2) $\frac{1}{n^2} \sum_{k=1}^{n} E(X_k^2 \mathbb{1}_{[|X_k|\le n]})$

$= \frac{1}{n} \int_{\Omega} |X_1|^2 \mathbb{1}_{[|X_1|\le n]} \, dP = \frac{1}{n} \int_{\{x: |x|\le n\}} x^2 \, dF(x)$



$= \frac{1}{n} \int_{|x|\le n} \left(\int_{s=0}^{|x|} 2s \, ds \right) dF(x)$

$= \frac{1}{n} \int_{s=0}^{n} 2s \left(\int_{s<|x|\le n} dF(x) \right) ds$, by Fubini's thm by $s>0$.

$= \frac{1}{n} \int_0^n 2s \left(P[|X_1| > s] - P[|X_1| > n] \right) ds$

$= \frac{1}{n} \int_0^n 2s \, P[|X_1| > s] \, ds - \frac{1}{n} P[|X_1| > n] \int_0^n 2s \, ds$

$= \frac{2}{n} \int_0^n s \, P[|X_1| > s] \, ds - n P[|X_1| > n] \to 0$,

Since $n P[|X_1| > n] \to 0$ and hence so does its average, i.e.

$$\frac{1}{n} \int_0^n s \, P[|X_1| > s] \, ds \to 0.$$

#

- <u>Thm</u> (Empirical Distribution Function).

Let $\{X_n, n\geq 1\}$ be iid r.v.'s with unknown distribution F.
Define the empirical distribution function $F_n(x)$ as

$$F_n(x) = \frac{1}{n}\sum_{k=1}^{n} 1_{[X_k\leq x]}.$$

- Then for $x\in\mathbb{R}$,

$$F_n(x) \xrightarrow{P} F(x) \text{ as } n\to\infty.$$

Pf: $n F_n(x) = \sum_{k=1}^{n} 1_{[X_k\leq x]}$.

$\because$ $\{X_k, k\geq 1\}$ are iid $\Rightarrow$ $\{1_{[X_k\leq x]}, k\geq 1\}$ are iid, too.

and $1_{[X_k\leq x]} = \begin{cases} 1 & \text{w.p. } F(x) = P[X_k\leq x], \\ 0 & \text{w.p. } 1-F(x). \end{cases}$

$\therefore$ $n F_n(x) \sim \text{binomial}(n, F(x))$.

$$E(F_n(x)) = \frac{1}{n}\cdot n F(x) = F(x),$$

$$Var(F_n(x)) = \frac{1}{n^2} n F(x)(1-F(x)) = \frac{1}{n} F(x)(1-F(x)) \leq \frac{1}{4n}$$

Then

$$P(|F_n(x)-F(x)|>\varepsilon) \leq \frac{Var(F_n(x))}{\varepsilon^2} \leq \frac{1}{4n\varepsilon^2} \to 0 \text{ as } n\to\infty.$$

Pf 2: $\{X_k, k\geq 1\}$ are iid $\Rightarrow$ $\{1_{[X_k\leq x]}, k\geq 1\}$ are iid with mean $F(x)$,
the weak law of large numbers implies

$$F_n(x) \xrightarrow{P} F(x) \text{ as } n\to\infty.$$

#

- <u>Example</u> (A Weak Law without Finite Mean)

The typical example is Pareto-like densities, i.e. densities where the tails drop off like a negative power of x, possibly multiplied by (a power of) a logarithm.

Suppose that $\{X_n, n\geq 1\}$ are iid r.v.'s with pdf

$$f(x)=\frac{c}{x^2\log|x|},\quad |x|>2.$$

(1) The mean does not exist, i.e. $E(|X|)=\infty$.

$$\int_{|x|>2}\frac{c|x|}{x^2\log|x|}=2c\int_2^\infty\frac{1}{x\log x}dx=+\infty$$

(2) With $b_n=n$ in the general WLLN,

(2.1) Check the first condition:

$$n\,P(|X|>n)=2n\int_n^\infty\frac{c}{x^2\log x}dx\sim n\frac{c}{n\log n}=\frac{c}{\log n}\to 0 \text{ as } n\to\infty.$$

(2.2) Check the second condition: (Not necessarily need to check it: Kolmogorov-Feller WLLN).

$$\frac{1}{n^2}n\,E\left(X^2 I_{[|X|\leq n]}\right)=\frac{2}{n}\int_2^n x^2\frac{c}{x^2\log x}dx=\frac{2c}{n}\int_2^n\frac{c}{\log x}dx$$

$$\sim\frac{c}{n}\frac{n}{\log n}=\frac{c}{\log n}\to 0 \text{ as } n\to\infty.$$

Hence the general weak law of large numbers implies the weak law holds, that is,

$$\frac{S_n-\mu_n}{b_n}\xrightarrow{P}0 \text{ as } n\to\infty,$$

where $S_n=\sum_{k=1}^n X_k$, $b_n=n$,

$$\mu_n=\sum_{k=1}^n E\left(X_k\cdot 1_{[|X_k|\leq n]}\right)=n\,E\left(X_k 1_{[|X|\leq n]}\right)=0$$

$\because$ the truncated symmetric r.v. is symmetric, too. #

• Example (General WLLN holds, but NOT Kolmogorov-Feller WLLN)

Let $\{X_n, n\geq 1\}$ be iid Cauchy $(0,1)$ r.v.'s.

(1) We first compute chf of it.

Let $Y_1, Y_2 \overset{iid}{\sim} exp(1)$. So $Y_1 - Y_2 \sim$ Laplace with $f_{Y_1-Y_2}(x) = \frac{1}{2}e^{-|x|}$.

$$\varphi_{Y_1-Y_2}(t) = E(e^{it(Y_1-Y_2)}) = E(e^{itY_1}\cdot e^{-itY_2}) = \varphi_{Y_1}(t)\,\varphi_{Y_2}(-t)$$
$$= \frac{1}{1-it}\,\frac{1}{1+it} = \frac{1}{1+t^2}$$

By the inversion formula of chf's:

$$f_{Y_1-Y_2}(x) = \frac{1}{2}e^{-|x|} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\varphi(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\frac{1}{1+t^2}\,dt$$

By a change of variable: $x\to t$, $t\to x$

$$\frac{1}{2}e^{-|t|} = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-itx}\frac{1}{1+x^2}\,dx \Rightarrow \varphi_{X_j}(t) = e^{-|t|}.$$

(2) $\varphi_{\bar{X}_n}(t) = [\varphi(\frac{t}{n})]^n = (e^{-|\frac{t}{n}|})^n = e^{-|t|} = \varphi_{X_j}(t)$.

The uniqueness thm of chf's implies that

$$\bar{X}_n \overset{d}{=} X_1 \text{ for all } n \Rightarrow \bar{X}_n \overset{P}{\nrightarrow} 0 \text{ as } n\to\infty.$$

⇒ Counter-intuitive: In general, $\bar{X}_n$ would have better precision than X_1.

⇒ The WLLN does NOT hold, which, in turn is no contradiction since the mean of the Cauchy distribution does not exist.

(3) We now check $nP(|X_1|>n)\to 0$ or not.

$$nP(|X_1|>n) = n\frac{2}{\pi}\int_n^{\infty}\frac{1}{1+x^2}\,dx \quad \text{let } x=\tan\theta,\ 1+x^2 = 1+\tan^2\theta = \sec^2\theta,\ dx = \sec^2\theta\,d\theta.$$
$$= \frac{2n}{\pi}\int_{\tan^{-1}n}^{\tan^{-1}\infty}\frac{1}{\cancel{\sec^2\theta}}\cancel{\sec^2\theta}\,d\theta$$
$$= \frac{2n}{\pi}\left(\frac{\pi}{2} - \tan^{-1}n\right) = \frac{2n}{\pi}\tan^{-1}\frac{1}{n} \quad \because \tan^{-1}n + \tan^{-1}\frac{1}{n} = \frac{\pi}{2}.$$
$$\to \frac{2}{\pi} \neq 0 \text{ as } n\to\infty \quad \because n\tan^{-1}\frac{1}{n}\to 1.$$

So Kolmogorov-Feller WLLN does NOT hold.

#

- Example (General WLLN holds, but NOT Kolmogorov-Feller WLLN).

Suppose that $\{X_n, n\geq 1\}$ are iid r.v.'s with pdf

$$f(x) = \frac{1}{2x^2}, \quad \text{for } |x|>1.$$

(1) The mean does NOT exist.

$$E(|X_1|) = \int_{|x|>1} \frac{|x|}{2x^2}\,dx = 2\int_1^\infty \frac{x}{2x^2}\,dx = \log x\Big|_1^\infty = \infty.$$

(2) The Kolmogorov-Feller WLLN does NOT hold.

$$nP[|X_1|>n] = 2n\int_n^\infty \frac{1}{2x^2}\,dx = n\left[-\frac{1}{x}\right]_n^\infty = 1 \not\to 0 \text{ as } n\to\infty.$$

(3) With $b_n = n\log n$ instead, the General WLLN holds.

- $$nP[|X_1|>n\log n] = 2n\int_{n\log n}^\infty \frac{1}{2x^2}\,dx = n\left[-\frac{1}{x}\right]_{n\log n}^\infty = \frac{1}{\log n} \to 0 \text{ as } n\to\infty;$$
- $$\begin{aligned}\frac{1}{(n\log n)^2}\, n\, E\left(X_1^2 1_{[|X_1|\leq n\log n]}\right) &= \frac{1}{n(\log n)^2}\int_{1<|x|\leq n\log n} \frac{x^2}{2x^2}\,dx \\ &= \frac{1}{n(\log n)^2}\, 2\int_1^{n\log n} \frac{1}{2}\,dx \\ &= \frac{1}{n(\log n)^2}(n\log n - 1) \to 0 \text{ as } n\to\infty.\end{aligned}$$

The General WLLN implies

$$\frac{S_n - a_n}{b_n} \xrightarrow{P} 0 \text{ as } n\to\infty,$$

where $b_n = n\log n$, $a_n = E\left(X_1 1_{[|X_1|\leq n\log n]}\right) = 0$ $\because X_1 1_{[|X_1|\leq n\log n]}$ is symmetric.

#.

- Example (The St. Petersburg Game)

It, the St. Petersburg paradox, was invented the formalism to handle r.v.'s with infinite expectation seemed paradoxical.

Toss a coin repeatedly until heads appears. If head happens at trial number n you receive 2^n dollars. What is a fair price for you to participate in this game.

(1) The r.v. X behind the game is w.p.

$$P(X=2^n)=\frac{1}{2^n},\ n=1,2,\cdots$$

(2) The mean does not exist: $E(X)=\sum_{n=1}^{\infty}2^n\frac{1}{2^n}=\infty$.

(3) The suitable solution is to set the fee as a funtion of n, as b_n, s.t.

$$\frac{S_n}{b_n}\xrightarrow{P}1,\ \text{i.e. } \frac{S_n-b_n}{b_n}\xrightarrow{P}0,\ \text{as } n\to\infty.$$

$\because P(|X|>2^n)=P(X>2^n)=\sum_{k=n+1}^{\infty}\frac{1}{2^k}=\frac{\frac{1}{2^{n+1}}}{1-\frac{1}{2}}=\frac{1}{2^n}$

$\therefore$ Set $2^n\sim x\log_2 x$, $n=[\log_2(x\log_2 x)]$. Take $b_x=x\log_2 x$,

$$x P(|X|>x\log_2 x)=xP(X>x\log_2 x)$$

$$=x\left(\frac{1}{2}\right)^{[\log_2(x\log_2 x)]}\le x\left(\frac{1}{2}\right)^{\log_2(x\log_2 x)}$$

$$=x\frac{1}{x\log_2 x}\to 0 \text{ as } x\to\infty.$$

$\because b_n=n\log_2 n\in RV(1)$, the WLLN implies that

$$\frac{S_n-nE(X_1 1_{[|X_1|\le n\log_2 n]})}{n\log_2 n}\xrightarrow{P}0. \text{ as } n\to\infty.$$

$$nE(X_1 1_{[|X_1|\le n\log_2 n]})=n\sum_{k=1}^{[\log_2(n\log_2 n)]}2^k\frac{1}{2^k}\qquad \text{by } 2^k=n\log_2 n,\ k=[\log_2(n\log_2 n)]$$

$$=n[\log_2(n\log_2 n)]$$

$$\sim n\log_2 n\qquad \text{by } \lim_{n\to\infty}\frac{[\log_2(n\log_2 n)]}{\log_2 n}=\lim_{n\to\infty}\left(1+\frac{1}{\log n}\right)=1$$

Hence,

$$\frac{S_n-n\log_2 n}{n\log_2 n}\xrightarrow{P}0\Rightarrow\frac{S_n}{n\log_2 n}\xrightarrow{P}1 \text{ as } n\to\infty$$

#.

• <u>Theorem</u> (Cantelli's Thm: SLLN for iid $\{X_n\}$ with $E[X_n^4]<\infty$)

If X_n are iid and $E[X_n]=m$ with $E[X_n^4]<\infty$, then

$$P[\lim_n \frac{S_n}{n}=m]=1. \text{ i.e. } \frac{S_n}{n}\to_{a.s.} m$$

Pf: • WLOG, assume $m=0$. Hence it suffices to show that

$$P[|\tfrac{1}{n}S_n|>\epsilon \text{ i.o.}]=0 \text{ for each } \epsilon.$$

Since $E[X_i^4]<\infty$, let $E[X_i^2]=\sigma^2$, $E[X_i^4]=\xi^4$.

• The Markov inequality implies that

$$P[|\tfrac{1}{n}S_n|>\epsilon] < \frac{1}{n^4\epsilon^4}E[S_n^4],$$

• and $E[S_n^4]$ will be shown to have order n^2.

$\because$ $E[S_n^4]=\Sigma E[X_\alpha X_\beta X_\gamma X_\delta]$, the four indices ranging independently from 1 to n.

① $E[X_i^4]=\xi^4$, with n terms,

② $E[X_i^2X_j^2]=E[X_i^2]E[X_j^2]=\sigma^4$, for $i\neq j$,

because there are n choices for the α, three ways to match it with β,γ,δ, and $n-1$ choices for the value common to the remaining two indices, with $3n(n-1)$ terms.

③ $E[X_i^2X_jX_k]=E[X_i^2]E[X_j]E[X_k]=0$,

④ $E[X_i^3X_j]=E[X_i^3]E[X_j]=0$

⑤ $E[X_iX_jX_kX_l]=0$.

$\therefore$ $E[S_n^4]=n\xi^4+3n(n-1)\sigma^4\le Kn^2$, where K does not depend on n.

So $P[|\tfrac{1}{n}S_n|>\epsilon]\le \frac{1}{n^4\epsilon^4}Kn^2=\frac{K}{n^2\epsilon^4}$

and by the first Borel-Cantelli lemma,

$$\sum_n P[|\tfrac{1}{n}S_n|>\epsilon]\le \sum_n \frac{K}{n^2\epsilon^4}<\infty$$

implies that

$$P[|\tfrac{1}{n}S_n|>\epsilon \text{ i.o.}]=0, \text{ as required.}$$

#

- <u>Theorem</u> (Kolmogorov's Inequality)

  Suppose that $X_1, \cdots, X_n$ are "independent" with $E[X_n]=0$ and $\sigma_n^2 = E[X_n^2] < \infty$.

  For $\alpha > 0$,

  $$P[\max_{1\le k\le n} |S_k| \ge \alpha] \le \frac{1}{\alpha^2} Var[S_n]$$

  Pf: • Let A_k be the set where $|S_k| \ge \alpha$ but $|S_j| < \alpha$ for $j<k$.

  Since A_k are disjoint,

  • $Var[S_n] = E[S_n^2]$ by $E[X_n]=0$

  $$= \sum_{k=1}^{n} \int_{A_k} S_n^2 \, dP = \sum_{k=1}^{n} \int_{A_k} (S_n - S_k + S_k)^2 \, dP$$

  $$= \sum_{k=1}^{n} \int_{A_k} [S_k^2 + 2S_k(S_n - S_k) + (S_n - S_k)^2] \, dP$$

  $$\ge \sum_{k=1}^{n} \int_{A_k} [S_k^2 + 2S_k(S_n - S_k)] \, dP.$$

  Since A_k and S_k are measurable $\sigma(X_1, \cdots, X_k)$ and

  $S_n - S_k$ is measurable $\sigma(X_{k+1}, \cdots, X_n)$,

  and since the means are all 0, it follows that

  $$\int_{A_k} S_k(S_n - S_k) \, dP = \int_{A_k} S_k \, dP \cdot \overset{0}{E(S_n - S_k)} = 0.$$

  $$Var[S_n] \ge \sum_{k=1}^{n} \int_{A_k} S_k^2 \, dP \ge \sum_{k=1}^{n} \alpha^2 P(A_k) = \alpha^2 P[\max_{1\le k\le n} |S_k| \ge \alpha].$$

  #

- Remark:

  By Chebychev's inequality,

  $$P[|S_n| \ge \alpha] \le \frac{1}{\alpha^2} Var[S_n].$$

  Furthermore, for sums of indep. r.v.'s,

  if $\max_{1\le k\le n} |S_k|$ is large, then $|S_n|$ is probably large, too.

• <u>Theorem</u> (Etemadi's Inequality).

Suppose $X_1, \cdots, X_n$ are "independent". For $\alpha > 0$,

$$P[\max_{1 \le k \le n} |S_k| \ge 3\alpha] \le 3 \max_{1 \le k \le n} P[|S_k| \ge \alpha].$$

pf: • Let B_k be the set where $|S_k| \ge 3\alpha$ but $|S_j| < 3\alpha$ for $j < k$.

$\{B_k\}$ are disjoint.

• $P[\max_{1 \le k \le n} |S_k| \ge 3\alpha]$

$\le P(\bigcup_{k=1}^{n} [|S_k| \ge 3\alpha]) = P(\bigcup_{k=1}^{n} B_k) \le \sum_{k=1}^{n} P(B_k)$

$\le \sum_{k=1}^{n} \left(P(B_k \cap [|S_n| \ge \alpha]) + P(B_k \cap [|S_n| < \alpha]) \right)$

$\le P[|S_n| \ge \alpha] + \sum_{k=1}^{n-1} P(B_k \cap [|S_n| < \alpha])$ $\quad \because B_n \cap [|S_n| < \alpha] = \phi$

$\le P[|S_n| \ge \alpha] + \sum_{k=1}^{n-1} P(B_k \cap [|S_n - S_k| > 2\alpha])$ $\quad \because 3\alpha \le |S_k| \le |S_n - S_k| + |S_n|$
$\le |S_n - S_k| + \alpha$

$= P[|S_n| \ge \alpha] + \sum_{k=1}^{n-1} P(B_k) P[|S_n - S_k| > 2\alpha]$ $\quad \because B_k \perp\!\!\!\perp [|S_n - S_k| > 2\alpha]$.

$\le P[|S_n| \ge \alpha] + \max_{1 \le k \le n} P[|S_n - S_k| \ge 2\alpha]$

$\le P[|S_n| \ge \alpha] + \max_{1 \le k \le n} \left(P[|S_n| \ge \alpha] + P[|S_k| \ge \alpha] \right)$. $\quad \because |S_n| < \alpha$ and $|S_k| < \alpha$
implies $|S_n - S_k| < |S_n| + |S_k|$
$< 2\alpha$

$\le 3 \max_{1 \le k \le n} P[|S_k| \ge \alpha]$.

#

- <u>Theorem</u> (Levy's Thm)

For an "independent" sequence $\{X_n\}$, let $S_n = \sum_{j=1}^{n} X_j$. Then S_n converge a.s. iff S_n converge in pr.

Pf: Since $S_n \xrightarrow{a.s.} S$ implies $S_n \xrightarrow{P} S$,

- it suffices to show that if $S_n \xrightarrow{P} S$, then $S_n \xrightarrow{a.s.} S$. Since

$$P[|S_{n+j} - S_n| \geq \epsilon] \leq P[|S_{n+j} - S| \geq \tfrac{\epsilon}{2}] + P[|S_n - S| \geq \tfrac{\epsilon}{2}]$$

- $S_n \xrightarrow{P} S$ implies $\lim_n \sup_{j \geq 1} P[|S_{n+j} - S_n| \geq \epsilon] = 0$. (1)
- By Etemadi's inequality,

$$P\left[\max_{1\leq j\leq k} |S_{n+j} - S_n| \geq \epsilon\right] \leq 3 \max_{1\leq j\leq k} P\left[|S_{n+j} - S_n| \geq \tfrac{\epsilon}{3}\right],$$

Since the set on the left are nondecreasing in k,

$$\begin{aligned}\lim_{k\to\infty} P\left[\max_{1\leq j\leq k} |S_{n+j} - S_n| \geq \epsilon\right] &= P\left[\lim_{k\to\infty} \max_{1\leq j\leq k} |S_{n+j} - S_n| \geq \epsilon\right] \\ &= P\left[\sup_{k\geq 1} |S_{n+k} - S_n| \geq \epsilon\right] \\ &\leq \lim_{k\to\infty} 3 \max_{1\leq j\leq k} P\left[|S_{n+j} - S_n| \geq \tfrac{\epsilon}{3}\right] \\ &= 3 \sup_{k\geq 1} P\left[|S_{n+k} - S_n| \geq \tfrac{\epsilon}{3}\right]\end{aligned}$$

$$\lim_n P\left[\sup_{k\geq 1} |S_{n+k} - S_n| \geq \epsilon\right] \leq 3 \lim_n \sup_{k\geq 1} P\left[|S_{n+k} - S_n| \geq \tfrac{\epsilon}{3}\right] = 0 \quad (2), \text{ by (1).}$$

- Let $E(n,\epsilon)$ be the set where $\sup_{j,k\geq n} |S_j - S_k| > 2\epsilon$, and put $E(\epsilon) = \bigcap_n E(n,\epsilon)$. Then $E(n,\epsilon) \downarrow E(\epsilon)$, and (2) implies $P(E(\epsilon)) = 0$.

  Now $\bigcup_\epsilon E(\epsilon)$, where the union extends over positive rational ϵ, contains the set where the sequence $\{S_n\}$ is not fundamental (does not have the Cauchy property), and this set therefore has probability 0.

#

• Theorem (Kolmogorov's Convergence Criterion)

Suppose $\{X_n, n \geq 1\}$ are "indep." r.v.'s. If $E[X_j^2] < \infty$ and

$$\sum_{j=1}^{\infty} Var(X_j) < \infty,$$

then

$$\sum_{j=1}^{\infty} (X_j - EX_j) \text{ converges almost surely.}$$

Sketch: show $\{S_n\}$ is L_2 Cauchy

$\Downarrow$

Cauchy in probability $\Leftrightarrow$ i.p. convergence

$\Downarrow$ by Lévy's thm

a.s. convergence.

pf: • WLOG, suppose $E(X_j) = 0$.

The sum of variance becomes $\sum_{j=1}^{\infty} E X_j^2 < \infty$

• This implies that $\{S_n\}$ is L_2 Cauchy since, for $m < n$,

$$\|S_n - S_m\|_2^2 = Var(S_n - S_m) = \sum_{j=m+1}^{n} E(X_j^2) \to 0 \text{ as } n, m \to \infty \because \sum_j E X_j^2 < \infty.$$

• So $\{S_n\}$, being L_2 Cauchy, is also Cauchy in probability since

$$P[|S_n - S_m| > \epsilon] \leq \epsilon^{-2} Var(S_n - S_m) \to 0 \text{ as } n, m \to \infty \text{ from above.}$$

• By Lévy's theorem, $\{S_n\}$ is almost surely convergent.

#

• Theorem (Converse of Kolmogorov Convergence Criterion)

Suppose $\{X_n, n \geq 1\}$ are indep., uniformly bounded r.v.'s. Then

$$\sum_{j=1}^{\infty} Var(X_j) < \infty$$

iff

$$\sum_{j=1}^{\infty} (X_j - EX_j) \text{ converges almost surely.}$$

- <u>Lemma</u> (Kronecker's Lemma)

Let $\{x_k\}$ be a sequence of real numbers, $\{a_k\}$ a sequence of numbers with $a_k > 0$ and $a_k \uparrow \infty$. Then

$$\sum_k \frac{x_k}{a_k} \text{ converges} \Rightarrow \frac{1}{a_n}\sum_{k=1}^{n} x_k \to 0.$$

pf: For $1 \le n \le \infty$, let

$$b_n = \sum_{j=1}^{n} \frac{x_j}{a_j}.$$

If we also write $a_0 = 0$, $b_0 = 0$, we have

$$b_n - b_{n-1} = \frac{x_n}{a_n},\quad x_n = a_n(b_n - b_{n-1}), \text{ and}$$

$$\frac{1}{a_n}\sum_{j=1}^{n} x_j = \frac{1}{a_n}\sum_{j=1}^{n} a_j(b_j - b_{j-1}) = \frac{1}{a_n}\left(a_n b_n + \sum_{j=1}^{n-1} a_j b_j - \sum_{j=1}^{n} a_j b_{j-1}\right)$$

$$= \frac{1}{a_n}\left(a_n b_n + \sum_{j=0}^{n-1} a_j b_j - \sum_{j=0}^{n-1} a_{j+1} b_j\right) \quad \text{by } a_0 = 0,\ b_0 = 0$$

$$= b_n - \frac{1}{a_n}\sum_{j=0}^{n-1} (a_{j+1} - a_j) b_j$$

Since $a_{j+1} - a_j \ge 0$ and $\frac{1}{a_n}\sum_{j=0}^{n-1}(a_{j+1} - a_j) = 1$, and $b_n \to b_\infty < \infty$, we have $\frac{1}{a_n}\sum_{j=1}^{n} x_j \to b_\infty - b_\infty = 0$.

#

• <u>Corollary</u> (Kolmogorov Sufficient Condition)

Let $\{X_n, n \geq 1\}$ be ① indep. r.v.'s satisfying ② $E(X_n^2) < \infty$, and hence $E(X_n) < \infty$, too.

Suppose we have a monotone sequence $b_n \uparrow \infty$.

If ③ $\sum_k Var(\frac{X_k}{b_k}) < \infty$,

then $\frac{S_n - E(S_n)}{b_n} \xrightarrow{a.s.} 0$

Pf: If $\sum_k Var(\frac{X_k}{b_k}) < \infty$, the Kolmogorov convergence criterion implies

$\sum_k \left(\frac{X_k - EX_k}{b_k}\right)$ converges a.s.

and the random Kronecker's lemma implies that

$$\frac{1}{b_n} \sum_{k=1}^{n} (X_k - EX_k) = \frac{S_n - E(S_n)}{b_n} \xrightarrow{a.s.} 0.$$

This is the Kolmogorov sufficient condition.

#

• <u>Theorem</u> (Necessary and Sufficient Conditions for Convergence of $\sum_n X_n$: The Kolmogorov Three-Series Thm).

Suppose that $\{X_n\}$ is "independent". Let $X_n^{(c)} = X_n 1_{[|X_n| \le c]}$.

Consider the three series

$$\sum_n P[|X_n| > c], \quad \sum_n E[X_n^{(c)}], \quad \sum_n Var[X_n^{(c)}].$$

In order that $\sum_n X_n$ converges a.s., it is necessary that the three series converge for all positive c and sufficient that they converge for some positive c.

• Pf of Sufficiency. Suppose that the three series converge, by Kolmogorov's convergence criterion,

$$\sum_n Var[X_n^{(c)}] < \infty \Rightarrow \sum_n (X_n^{(c)} - E[X_n^{(c)}]) \text{ converges a.s.,}$$

and since $\sum_n E[X_n^{(c)}]$ converges, so does $\sum_n X_n^{(c)}$.

Since $\sum_n P[X_n \ne X_n^{(c)}] = \sum_n P[|X_n| > c] < \infty$,

$\{X_n\}$ and $\{X_n^{(c)}\}$ are equivalent sequences and hence $\sum_n X_n$ converges a.s.

• Pf of Necessity. Suppose that $\sum_n X_n$ converges a.s, and fix $c > 0$.

Since $X_n \to 0$ a.s., it follows that $\sum_n X_n^{(c)}$ converges a.s., and by the second Borel-Cantelli lemma,

$$P[|X_n| > c \text{ i.o.}] = 0 \Rightarrow \sum_n P[|X_n| > c] < \infty. \Rightarrow \text{1st series converges}$$

Let $S_n^{(c)} = \sum_{k=1}^n X_k^{(c)}$, $M_n^{(c)} = E[S_n^{(c)}] = \sum_n E[X_n^{(c)}]$, $(s_n^{(c)})^2 = Var[S_n^{(c)}] = \sum_n Var[X_n^{(c)}]$.

If $s_n^{(c)} \to \infty$, then since $X_n^{(c)} - m_n^{(c)}$ are uniformly bounded, $m_n^{(c)} = E[X_n^{(c)}]$, it follows by the central limit thm that

$$\lim_n P\left[x < \frac{S_n^{(c)} - M_n^{(c)}}{s_n^{(c)}} \le y\right] = \frac{1}{\sqrt{2\pi}} \int_x^y e^{-\frac{t^2}{2}} dt. \quad (1)$$

And since $\sum_n X_n^{(c)}$ converges a.s., $s_n^{(c)} \to \infty$ also implies $\frac{S_n^{(c)}}{s_n^{(c)}} \to 0$ a.s.

So that $\frac{S_n^{(c)}}{s_n^{(c)}} \to_p 0$ and hence $\lim_n P\left[\left|\frac{S_n^{(c)}}{s_n^{(c)}}\right| > \epsilon\right] = 0. \quad (2)$

But (1) and (2) stand in contradiction.

(next pg. cont.)

$\because P\left[x < \frac{S_n^{(c)} - M_n^{(c)}}{s_n^{(c)}} < y\right]$

$\leq P\left[x < \frac{S_n^{(c)} - M_n^{(c)}}{s_n^{(c)}} < y, \left|\frac{S_n^{(c)}}{s_n^{(c)}}\right| < \epsilon\right] + P\left[x < \frac{S_n^{(c)} - M_n^{(c)}}{s_n^{(c)}} < y, \left|\frac{S_n^{(c)}}{s_n^{(c)}}\right| \geq \epsilon\right]$

$\leq P\left[x < \frac{S_n^{(c)} - M_n^{(c)}}{s_n^{(c)}} < y, \left|\frac{S_n^{(c)}}{s_n^{(c)}}\right| < \epsilon\right] + P\left[\left|\frac{S_n^{(c)}}{s_n^{(c)}}\right| \geq \epsilon\right]$

$\therefore$ For all sufficiently large n,

$$P\left[x < \frac{S_n^{(c)} - M_n^{(c)}}{s_n^{(c)}} < y, \left|\frac{S_n^{(c)}}{s_n^{(c)}}\right| < \epsilon\right] \geq \frac{1}{\sqrt{2\pi}} \int_x^y e^{-\frac{t^2}{2}} dt > 0 \text{ (if } x < y)$$

But then

$$-\epsilon < \frac{S_n^{(c)}}{s_n^{(c)}} < \epsilon, \quad x - \frac{S_n^{(c)}}{s_n^{(c)}} < -\frac{M_n^{(c)}}{s_n^{(c)}} < y - \frac{S_n^{(c)}}{s_n^{(c)}} \Rightarrow x - \epsilon < -\frac{M_n^{(c)}}{s_n^{(c)}} < y + \epsilon,$$

and this cannot hold simultaneously for, say, $(x-\epsilon, y+\epsilon) = (-1, 0)$ and $(0, 1)$.
Thus $s_n^{(c)}$ cannot go to ∞, i.e. $(s_n^{(c)})^2 = \sum_n Var[X_n^{(c)}] < \infty \Rightarrow$ 2nd series converges.
And now it follows by Kolmogorov's converges criterion,

$$\sum_n Var[X_n^{(c)}] < \infty \Rightarrow \sum_n (X_n^{(c)} - E[X_n^{(c)}]) \text{ converges a.s.},$$

$\because \sum_n X_n^{(c)}$ converges a.s. $\Rightarrow \sum_n E[X_n^{(c)}]$ converges.

#

- **Proposition** (Moments and Tail Probability)

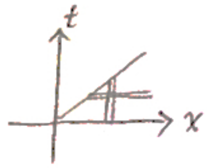
(a) For "nonnegative" X,

$$E[X] = \int_0^\infty P[X>t]\,dt = \int_0^\infty P[X\ge t]\,dt$$

(b) Even if X is not nonnegative,

$$\int_{[X>\alpha]} X\,dP = \alpha P[X>\alpha] + \int_\alpha^\infty P[X>t]\,dt, \quad \alpha \ge 0.$$

Pf: (a) Suppose X is nonnegative,



$$E[X] = \int_0^\infty x\,dF(x) = \int_0^\infty \int_0^x dt\,dF(x)$$

$$= \int_0^\infty \int_t^\infty dF(x)\,dt = \int_0^\infty P[X>t]\,dt \quad \because 1 \ge 0, \text{ by Fubini's thm.}$$

Since $P[X=t] > 0$ for at most countably many values of t,

$P[X>t]$ and $P[X\ge t]$ differ only on a set of Lebesgue measure 0,

hence $E[X] = \int_0^\infty P[X>t]\,dt = \int_0^\infty P[X\ge t]\,dt$

(b) Replacing X by $X I_{[X>\alpha]}$, for $\alpha \ge 0$, part (a) implies that, for $X>0$,

$$\int_{[X>\alpha]} X\,dP = \int_0^\infty P[X I_{[X>\alpha]} > t]\,dt$$

$$= \int_0^\alpha P[X I_{[X>\alpha]} > t]\,dt + \int_\alpha^\infty P[X I_{[X>\alpha]} > t]\,dt$$

$$= \alpha P[X>\alpha] + \int_\alpha^\infty P[X>t]\,dt$$

As long as $\alpha \ge 0$, this holds even if X is not nonnegative:

$$\int_{[X>\alpha]} X\,dP = \int_{[X^+>\alpha]} X^+\,dP - \int_{[X^+>\alpha]} X^-\,dP \quad \text{by } \alpha\ge 0, [X>\alpha]=[X^+>\alpha].$$

$$= \int_{[X^+>\alpha]} X^+\,dP - 0 \quad \because X^-=0 \text{ on } [X^+>\alpha] \Rightarrow \int_{[X^+>\alpha]} X^-\,dP = 0$$

$$= \alpha P[X^+>\alpha] + \int_\alpha^\infty P[X^+>t]\,dt$$

$$= \alpha P[X>\alpha] + \int_\alpha^\infty P[X>t]\,dt.$$

#

- **Theorem** (Moments, Tail Probabilities, and Convergence a.s.)

We have

$$\sum_{n=1}^{\infty} P[|X|^r > n] \le E(|X|^r) \le 1 + \sum_{n=1}^{\infty} P[|X|^r > n]$$

So that the following are equivalent for $\{X_n\}$ "i.i.d" r.v.'s:

(i) $E(|X_n|^r) < \infty$ i.e. $X_n \in L_r$

(ii) $\sum_{n=1}^{\infty} P[|X_n| > n^{\frac{1}{r}}\epsilon] < \infty$ for $\epsilon > 0$

(iii) $P([|X_n| > n^{\frac{1}{r}}\epsilon]\ i.o.) = 0$ for $\epsilon > 0$

(iv) $\frac{X_n}{n^{1/r}} \xrightarrow{a.s.} 0$ as $n \to \infty$ (or $\frac{|X_n|}{n^{1/r}} \xrightarrow{a.s.} 0$ as $n \to \infty$)

Pf: • If $\Lambda_n = \{n \le |X|^r < n+1\}$,

$$E(|X|^r) = \sum_{n=0}^{\infty} \int_{\Lambda_n} |X|^r dP, \quad \text{by } |X|^r \ge 0$$

so $\sum_{n=0}^{\infty} nP(\Lambda_n) \le E(|X|^r) \le \sum_{n=0}^{\infty}(n+1)P(\Lambda_n) = 1 + \sum_{n=0}^{\infty} nP(\Lambda_n)$

since $\sum_{n=0}^{\infty} nP(\Lambda_n) = \sum_{n=0}^{\infty} n\, P(|X|^r \in [n, n+1)) = \sum_{n=1}^{\infty} P[|X|^r > n]$,

the inequalities follow.

• For (i) $\Leftrightarrow$ (ii):

$\because$ X_n's are iid, so that the X_n are identically distributed,

$$E(|X_n|^r) < \infty \text{ iff } \sum_{n=1}^{\infty} P[|X_n|^r > n] < \infty$$

$$\text{iff } \sum_{n=1}^{\infty} P[|\tfrac{X_n}{\epsilon}|^r > n] < \infty$$

$$\text{iff } \sum_{n=1}^{\infty} P[|X_n| > n^{1/r}\epsilon] < \infty$$

• For (ii) $\Leftrightarrow$ (iii):

$\because$ X_n's are iid, of course indep., the Borel zero-one law implies

$$\sum_{n=1}^{\infty} P[|X_n| > n^{1/r}\epsilon] < \infty \text{ iff } P[|X_n| > n^{1/r}\epsilon\ i.o.] = 0$$

• For (iii) $\Leftrightarrow$ (iv):

$$0 = P[|X_n| > n^{1/r}\epsilon\ i.o.] = P[|\tfrac{X_n}{n^{1/r}}| > \epsilon\ i.o.]$$

iff $\frac{X_n}{n^{1/r}} \to 0$ a.s.

\#

• <u>Theorem</u> (Kolmogorov's SLLN)

Let $\{X_n, n\geq 1\}$ be a sequence of "iid" r.v.'s, let $S_n = \sum_{k=1}^{n} X_k$

(a) If $E|X_k|<\infty$ and $E[X_k]=\mu$, then

$$\frac{S_n}{n} \xrightarrow{a.s.} \mu \quad \text{as } n\to\infty$$

(b) If $\frac{S_n}{n} \xrightarrow{a.s.} c$ as $n\to\infty$, then

$$E|X_k|<\infty \text{ and } c = E[X_k]$$

(c) If $E|X_k|=\infty$, then

$$\limsup_n \frac{|S_n|}{n} = +\infty \quad \text{a.s.}$$

Pf: (a) • Set $X_n' = X_n 1_{[|X_n|\leq n]}$, $n\geq 1$, $S_n' = \sum_{j=1}^{n} X_j'$. Since

$\sum_n P[X_n \neq X_n'] = \sum_n P[|X_n|>n] = \sum_n P[|X_1|>n] < \infty$, $\because E(|X_1|)<\infty$ $(r=1)$. (Result)

$\{X_n\}$ and $\{X_n'\}$ are equivalent sequences. Therefore

$$\frac{S_n}{n} \xrightarrow{a.s.} E(X_1) \text{ iff } \frac{S_n'}{n} \xrightarrow{a.s.} E(X_1), \text{ i.e. } \frac{S_n' - E(S_n)}{n} \xrightarrow{a.s.} 0.$$

• So it suffices to consider the truncated sequence.

$$\because \left|\frac{S_n' - E(S_n')}{n} - \frac{S_n' - E(S_n)}{n}\right| = \left|\frac{n E(X_1) - \sum_{j=1}^{n} E(X_1 1_{[|X_1|\leq j]})}{n}\right|$$

$$= \left|\frac{\sum_{j=1}^{n} E(X_1 1_{[|X_1|>j]})}{n}\right| \to 0$$

The last step follows from

$|E(X_1 1_{[|X_1|>n]})| \leq E(|X_1| 1_{[|X_1|>n]}) \to 0$ and hence

the Cesaro averages converge to 0.

Thus,

$$\frac{S_n' - E(S_n)}{n} \xrightarrow{a.s.} 0 \text{ iff } \frac{S_n' - E(S_n')}{n} \xrightarrow{a.s.} 0, \text{ i.e. } \frac{1}{n}\sum_{j=1}^{n}(X_j' - E(X_j')) \xrightarrow{a.s.} 0.$$

• By Kolmogorov's convergence thm, it suffices to check

$\sum_j Var\left(\frac{X_j'}{j}\right) < \infty$, then $\sum_j \left(\left(\frac{X_j'}{j}\right) - E\left(\frac{X_j'}{j}\right)\right)$ converges a.s.,

by Kronecker's lemma, $\frac{1}{n}\sum_{j=1}^{n}(X_j' - E(X_j')) \xrightarrow{a.s.} 0$, the result follows.

(next pg. cont.)

$\cdot \sum_j \mathrm{Var}\left(\frac{X_j'}{j}\right) = \sum_j \frac{1}{j^2} \mathrm{Var}(X_j') \le \sum_j \frac{1}{j^2} E(X_j')^2$

$= \sum_j \frac{1}{j^2} E\left(X_1^2 1_{[|X_1| \le j]}\right)$ by $\{X_n\}$ iid.

$= \sum_{j=1}^{\infty} \sum_{k=1}^{j} \frac{1}{j^2} E\left(X_1^2 1_{[k-1 < |X_1| \le k]}\right)$

$= \sum_{k=1}^{\infty} \left(\sum_{j=k}^{\infty} \frac{1}{j^2}\right) E\left(X_1^2 1_{[k-1 < |X_1| \le k]}\right)$ $\because$ summands ≥ 0.

$\because \sum_{j=k}^{\infty} \frac{1}{j^2} \le \sum_{j=k}^{\infty} \int_{j-1}^{j} \frac{1}{x^2}\,dx = \int_{k-1}^{\infty} \frac{1}{x^2}\,dx = \frac{1}{k-1} \le \frac{2}{k}$, provided $k \ge 2$.

$\therefore \sum_{k=2}^{\infty} \left(\sum_{j=k}^{\infty} \frac{1}{j^2}\right) E\left(X_1^2 1_{[k-1 < |X_1| \le k]}\right)$

$\le \sum_{k=2}^{\infty} \frac{2}{k} E\left(X_1^2 1_{[k-1 < |X_1| \le k]}\right)$

$\le \sum_{k=2}^{\infty} \frac{2}{k} \cdot k\, E\left(|X_1| 1_{[k-1 < |X_1| \le k]}\right)$ $\because |X_1| \le k$

$\le 2E(|X_1|) < \infty$

(b) $\frac{X_n}{n} = \frac{S_n - S_{n-1}}{n} = \frac{S_n}{n} - \frac{n-1}{n}\frac{S_{n-1}}{n-1} \xrightarrow{a.s.} C - C = 0.$

Since $\frac{X_n}{n} \xrightarrow{a.s.} 0$ implies $E(|X_1|) < \infty$ and hence from (a). (Result)

(c) We know that

$E(|X_1|) < \infty$ iff $P(|X_n| > n \text{ i.o.}) = 0$

So $E(|X_1|) = \infty$ iff $P(|X_n| > n \text{ i.o.}) = 1$

But $|X_n| = |S_n - S_{n-1}| \le |S_n| + |S_{n-1}|$,

$|S_n| < \frac{n}{2}$ and $|S_{n-1}| < \frac{n}{2}$ implies $|X_n| < n$,

equivalently,

$|X_n| > n$ implies $|S_n| > \frac{n}{2}$ or $|S_{n-1}| > \frac{n}{2}$

For any $A > 0$,

$P(|X_n| > An \text{ i.o.}) = 1$ implies $P\left[|S_n| > \frac{An}{2} \text{ i.o.}\right] = 1$

This means that for each A there is a null set $N(A)$ s.t.

if $\omega \in (N(A))^c$, then $\limsup_{n\to\infty} \frac{|S_n|}{n} \ge \frac{A}{2}$.

Let $N = \bigcup_{A=1}^{\infty} N(A)$, then N is still a null set, and if $\omega \in N^c$,

$\limsup_{n\to\infty} \frac{|S_n|}{n} \ge \frac{A}{2}$ for every A, and therefore $+\infty$.

#.

- Corolary (The Strong Law of Large Numbers for $E[X_n]=\infty$)

Suppose that the X_n are "iid" and

$E[X_1^-]<\infty$, $E[X_1^+]=\infty$ so that $E[X_1]=\infty$.

Then $n^{-1}\sum_{k=1}^{n}X_k\to\infty$ with probability 1

Pf: Since $E[X_1^-]<\infty$, by the SLLN,

$n^{-1}\sum_{k=1}^{n}X_k^-\to E[X_1^-]$ with probability 1,

and so it suffices to prove the corollary for the case $X_1=X_1^+\geq 0$.

Suppose $X_n^{(u)}=X_n I_{[|X_n|\leq u]}$,

then $n^{-1}\sum_{k=1}^{n}X_k\geq n^{-1}\sum_{k=1}^{n}X_k^{(u)}\to E[X_1^{(u)}]$ with probability 1, by the SLLN,

$E[X_1^{(u)}]$ exists since the $X_n^{(u)}$ are uniformly bounded.

Let $u\to\infty$ the result follows.

#

• <u>Definition</u> (Empirical Distribution Function)

The empirical df for r.v.'s $X_1, \cdots, X_n$ is the distribution function $F_n(x,\omega)$ with a jump of n^{-1} at each $X_k(\omega)$:

$$F_n(x,\omega) = \frac{1}{n}\sum_{k=1}^{n} I_{(-\infty,x]}(X_k(\omega)).$$

If the X_K have a common unknown distribution function $F(x)$, then $F_n(x,\omega)$ is its natural estimate.

• <u>Theorem</u> (Glivenko-Cantelli Theorem)

Suppose that $X_1, X_2, \ldots$ are independent and have a common df F; put $D_n(\omega) = \sup_x |F_n(x,\omega) - F(x)|$. Then $D_n \to 0$ a.s..

Remark:

1. For each x, $F_n(x,\omega)$ as a function of ω is a r.v..
   By right continuity, the supremum above is unchanged if x is restricted to the rationals, and therefore D_n is a r.v.
2. The summands are iid simple r.v.'s (actually, Bernoulli r.v.'s), so by the SLLN, for each x there is a A_x of probability 0 s.t.
   $\lim_n F_n(x,\omega) = F(x)$ for $\omega \notin A_x$. $(*)$
   But the Glivenko-Cantelli thm says more, namly that $(*)$ holds for ω outside some set A of probability 0, where A does not depend on x.
   Futher, the convergence in $(*)$ is uniform in x.
3. By WLLN, $F_n(x,\omega) \xrightarrow{P} F(x)$,
   and hence $F_n(x,\omega) \Rightarrow F(x)$;
   By SLLN, $F_n(x,\omega) \xrightarrow{a.s.} F(x)$;
   By Glivenko-Cantelli Thm, $F_n(x,\omega) \xrightarrow{a.s.} F(x)$ uniformly in x;
   That is, $D_n(\omega) = \sup_x |F_n(x,\omega) - F(x)| \xrightarrow{a.s.} 0$.

Pf: As already observed, the set A_x where $(*)$ fails has probability 0.
Another application of the SLLN, with $I_{(-\infty,x)}$ in place of $I_{(-\infty,x]}$, shows that $\lim_n F_n(x-,\omega) = F(x-)$ except on B_x of probability 0.
Let the quantile function $\varphi(u) = \inf[x: u \le F(x)]$ for $0 < u < 1$, and put $x_{m,k} = \varphi(k/m)$, $m \ge 1$, $1 \le k \le m$.
So $F(\varphi(u)-) \le u \le F(\varphi(u))$; and

$$F(x_{m,k}-) \le \frac{k}{m} \le F(x_{m,k}),\quad F(x_{m,k-1}-) \le \frac{k-1}{m} \le F(x_{m,k})$$

Thus $F(x_{m,k}-) - F(x_{m,k-1}) \le \frac{k}{m} - \frac{k-1}{m} = \frac{1}{m}$,

$$F(x_{m,1}-) \le \frac{1}{m},\quad F(x_{m,k}) \ge \frac{m-1}{m} = 1 - \frac{1}{m}.$$

If $D_{m,n}(\omega) = \bigvee_{1 \le k \le m} |F_n(x_{m,k},\omega) - F(x_{m,k})| \vee |F_n(x_{m,k}-,\omega) - F(x_{m,k}-)|$.
If $x_{m,k-1} \le x < x_{m,k}$, then

$$\begin{aligned}
F_n(x,\omega) &\le F_n(x_{m,k}-,\omega) = F_n(x_{m,k}-,\omega) - F(x_{m,k}-) + F(x_{m,k}-) \\
&\le |F_n(x_{m,k}-,\omega) - F(x_{m,k}-)| + F(x_{m,k}-) \\
&\le F(x_{m,k}-) + D_{m,n}(\omega) \le F(x_{m,k-1}) + \frac{1}{m} + D_{m,n}(\omega) \\
&\le F(x) + m^{-1} + D_{m,n}(\omega);
\end{aligned}$$

$$\begin{aligned}
F_n(x,\omega) &\ge F_n(x_{m,k-1},\omega) = F_n(x_{m,k-1},\omega) - F(x_{m,k-1}) + F(x_{m,k-1}) \\
&\ge -|F_n(x_{m,k-1}) - F(x_{m,k-1})| + F(x_{m,k-1}) \\
&\ge F(x_{m,k-1}) - D_{m,n}(\omega) \ge F(x_{m,k}-) - \frac{1}{m} - D_{m,n}(\omega) \\
&\ge F(x) - m^{-1} - D_{m,n}(\omega);
\end{aligned}$$

Hence $|F_n(x,\omega) - F(x)| \le D_{m,n}(\omega) + m^{-1}$.
Together with the similar arguments for the cases $x < x_{m,1}$ and $x > x_{m,m}$, this shows that

$$D_n(\omega) \le D_{m,n}(\omega) + m^{-1}.$$

If ω lies outside the union A of all the $A_{x_{m,k}}$ and $B_{x_{m,k}}$, then $\lim_n D_{n,m}(\omega) = 0$ and hence $\lim_n D_n(\omega) = 0$.
But $P(A) \le \sum_m \sum_k [P(A_{x_{m,k}}) + P(B_{x_{m,k}})] = 0 \Rightarrow A$ has probability 0.

#