STAT 430-1, Fall 2024
HW4

Tuorui Peng[1]

Contents

Notation: I use $\fallingdotseq$ for Fourier transform and $\risingdotseq$ for inverse Fourier transform. i.e. $f(x) \fallingdotseq \phi(t)$ and $\phi(t) \risingdotseq f(x)$.

Exercise 1 Weak LLN for weakly correlated random variables

Note that $\mathbb{E}\left[S_n\right] = 0$, $\forall n$, using Chebyshev's inequality, we have

$$\mathbb{P}\left(|S_n| \geq n\varepsilon\right) \leq \frac{var(S_n)}{n^2\varepsilon^2}$$

Now we study the variance of S_n: denote $var(X) := \sigma^2$, upper bound of $r(\cdot)$ being $\sup_{k\in\mathbb{N}^+} r(k) := R$.

1. Note that $r(k) \xrightarrow{k\to\infty} 0$, which implies that $\forall\delta > 0$, $\exists N$ s.t. $\forall n \geq N$ we have $0 < r(n) \leq \delta$, then we have

$$\begin{aligned} var(S_n) =& n\sigma^2 + \sum_{i=1}^{n-1}(n-i)r(i) \\ \leq& n\sigma^2 + \sum_{i=1}^{N}(n-i)R + \sum_{i=N+1}^{n-1}(n-i)\delta \\ =& n\sigma^2 + \frac{N(2n-N-1)R}{2} + \frac{(n-N-2)(n+N)\delta}{2} \\ =& O(n) + O(n^2\delta) \end{aligned}$$

2. Here we consider first sending $n\to\infty$, then $\delta\to 0$ to get: $\forall\varepsilon > 0$

$$\mathbb{P}\left(\left|\frac{S_n}{n}\right| \geq \varepsilon\right) \leq \frac{var(S_n)}{n^2\varepsilon^2} \lesssim \frac{1}{n\varepsilon^2} + \frac{\delta}{\varepsilon^2} \xrightarrow{n\to\infty} \frac{\delta}{\varepsilon^2} \xrightarrow[\delta\in\mathbb{Q}]{\delta\to 0} 0$$

Thus we have $\dfrac{S_n}{n} \xrightarrow{\mathrm{p}} 0$, as desired.

[1] TuoruiPeng2028@u.northwestern.edu

Exercise 2 Coupon collector's problem

1. We have the following bounds for $\mathbb{P}(T_n > k)$:

$$(1-1/n)^k \overset{(a)}{\leq} \mathbb{P}(T_n > k) \overset{(b)}{\leq} n(1-1/n)^k.$$

Proof. By writing the event $T_n > k$ as the union of n events, we have

(a) We have

$$\begin{aligned}\mathbb{P}(T_n > k) =& \mathbb{P}\left(\bigcup_{i=1}^{n}\left\{\text{no draw of coupon } i \text{ in the first } k \text{ draws}\right\}\right)\\ \geq& \mathbb{P}\left(\left\{\text{no draw of coupon 1 in the first } k \text{ draws}\right\}\right)\\ =& (1-1/n)^k\end{aligned}$$

(b) We have

$$\begin{aligned}\mathbb{P}(T_n > k) =& \mathbb{P}\left(\bigcup_{i=1}^{n}\left\{\text{no draw of coupon } i \text{ in the first } k \text{ draws}\right\}\right)\\ \leq& \sum_{i=1}^{n}\mathbb{P}\left(\left\{\text{no draw of coupon } i \text{ in the first } k \text{ draws}\right\}\right)\\ =& n(1-1/n)^k\end{aligned}$$

□

2. Then we prove $\lim\limits_{n\to\infty}\mathbb{P}\left(\dfrac{T_n}{n\log n} \geq 1+\varepsilon\right) = 0$:

Proof. First notice the following limits for any given $\varepsilon > 0$:

$$\lim_{n\to\infty}\frac{\log\left[n(1-\frac{1}{n})^{(1+\varepsilon)n\log n}\right]}{\log n} = \lim_{n\to\infty} 1 + (1+\varepsilon)\log\left(1-\frac{1}{n}\right)^n = -\varepsilon$$

which means there always exists some N s.t. $\forall n > N$ we have

$$\frac{\log\left[n(1-\frac{1}{n})^{(1+\varepsilon)n\log n}\right]}{\log n} \leq -\varepsilon + \frac{\varepsilon}{2} = -\frac{\varepsilon}{2} \Rightarrow n(1-\frac{1}{n})^{(1+\varepsilon)n\log n} \leq n^{-\varepsilon/2} \xrightarrow{n\to\infty} 0$$

Then we have

$$\mathbb{P}\left(\frac{T_n}{n\log n} > 1+\varepsilon\right) \leq n(1-\frac{1}{n})^{(1+\varepsilon)n\log n} \xrightarrow{n\to\infty} 0$$

□

3. Similarly we prove $\lim\limits_{n\to\infty}\mathbb{P}\left(\dfrac{T_n}{n\log n} \geq 1-\varepsilon\right) = 1$:

Proof. Notice the following limits for any given $\varepsilon > 0$:

$$\lim_{n\to\infty} \frac{\log\left[(1-\frac{1}{n})^{(1-\varepsilon)n\log n}\right]}{\log n} = \lim_{n\to\infty}(1-\varepsilon)\log\left(1-\frac{1}{n}\right)^n = -1+\varepsilon$$

which means there always exists some N s.t. $\forall n > N$ we have

$$\frac{\log\left[(1-\frac{1}{n})^{(1-\varepsilon)n\log n}\right]}{\log n} \geq -1+\varepsilon-\frac{\varepsilon}{2} = -1+\frac{\varepsilon}{2} \Rightarrow (1-\frac{1}{n})^{(1-\varepsilon)n\log n} \geq n^{-1+\varepsilon/2}, \quad n > N$$

Then we have

$$\begin{aligned}\sum_{n=1}^{\infty}\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon\right) \geq& \sum_{n=1}^{N}\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon\right) + \sum_{n=N+1}^{\infty}\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon\right)\\ \geq& \sum_{n=1}^{N}\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon\right) + \sum_{n=N+1}^{\infty} n^{-1+\varepsilon/2} = \infty\end{aligned}$$

and notice that we have independence of $\{T_n\}$, thus by Borel-Cantelli lemma (2nd kind), we have

$$\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon \text{ i.o.}\right) = 1 \Rightarrow \lim_{n\to\infty}\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon\right) = 1$$

$\square$

4. Combining the two sides, we have

$$\begin{cases}\lim_{n\to\infty}\mathbb{P}\left(\frac{T_n}{n\log n} > 1+\varepsilon\right) = 0\\ \lim_{n\to\infty}\mathbb{P}\left(\frac{T_n}{n\log n} > 1-\varepsilon\right) = 1\end{cases} \quad \Rightarrow \frac{T_n}{n\log n} \xrightarrow{\text{p}} 1$$

Exercise 3 "Almost" Law of the Iterated Logarithm

Denote the density function of standard normal distribution as $\phi(x)$, and its derivative as $\phi'(x) = -x\phi(x)$.

3.(a)

For $X \sim N(0,1)$:

- Using integration by parts, we have

$$\mathbb{P}(X \geq x) = \int_x^\infty \phi(t)\,\mathrm{d}t = \int_x^\infty (-\frac{1}{t})\phi'(t)\,\mathrm{d}t = \frac{1}{x}\phi(x) - \int_x^\infty \frac{1}{t^2}\phi(t)\,\mathrm{d}t \leq \frac{1}{x}\phi(x) \tag{1}$$

- Using integration by parts again, we have

$$\begin{aligned}\mathbb{P}(X \geq x) =& \frac{1}{x}\phi(x) - \int_x^\infty \frac{1}{t^2}\phi(t)\,\mathrm{d}t = \frac{1}{x}\phi(x) - \int_x^\infty (-\frac{1}{t^3})\phi'(t)\,\mathrm{d}t\\ =& \frac{1}{x}\phi(x) - \frac{1}{x^3}\phi(x) + \int_x^\infty \frac{3}{t^4}\phi(t)\,\mathrm{d}t \geq \frac{1}{x}\phi(x) - \frac{1}{x^3}\phi(x)\end{aligned} \tag{2}$$

Combining (1) and (2), we have

$$\left(\frac{1}{x}-\frac{1}{x^3}\right)\phi(x) \overset{(2)}{\leq} \mathbb{P}(X \geq x) \overset{(1)}{\leq} \frac{1}{x}\phi(x)$$

3.(b)

We notice the following limits for $n \to \infty$:[2]

$$\frac{\log \left[\frac{1}{x}\phi(x)\Big|_{x=\sqrt{1+\varepsilon}\sqrt{2\log n}} \right]}{\log n} \xrightarrow{n\to\infty} -(1+\varepsilon) \tag{1}$$

$$\frac{\log \left[\left(\frac{1}{x} - \frac{1}{x^3}\right)\phi(x)\Big|_{x=\sqrt{1-\varepsilon}\sqrt{2\log n}} \right]}{\log n} \xrightarrow{n\to\infty} -(1-\varepsilon) \tag{2}$$

Now using the two bounds from part (a), we have

1. For $\varepsilon > 0$, we have

$$\mathbb{P}\left(X_n > \sqrt{1+\varepsilon}\sqrt{2\log n} \right) \leq \frac{1}{x}\phi(x)\Big|_{x=\sqrt{1+\varepsilon}\sqrt{2\log n}}$$

using the limit (1), we have for such $\varepsilon/2 > 0$, there exists N s.t. $\forall n > N$ we have

$$\frac{\log \left[\frac{1}{x}\phi(x)\Big|_{x=\sqrt{1+\varepsilon}\sqrt{2\log n}} \right]}{\log n} \leq -(1+\varepsilon) + \frac{\varepsilon}{2} = -1 - \frac{\varepsilon}{2} \Rightarrow \frac{1}{x}\phi(x)\Big|_{x=\sqrt{1+\varepsilon}\sqrt{2\log n}} \leq n^{-1-\varepsilon/2}, \quad \forall n > N$$

Then we have

$$\sum_{n=1}^{\infty} \mathbb{P}\left(X_n > \sqrt{1+\varepsilon}\sqrt{2\log n} \right) \leq \sum_{n=1}^{N} \mathbb{P}\left(X_n > \sqrt{1+\varepsilon}\sqrt{2\log n} \right) + \sum_{n=N+1}^{\infty} n^{-1-\varepsilon/2} < \infty$$

Then by Borel-Cantelli lemma (1st kind), we have

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{X_n}{\sqrt{2\log n}} > \sqrt{1+\varepsilon} \right) = 0, \quad \forall \varepsilon > 0$$

2. For $\varepsilon > 0$, we have

$$\mathbb{P}\left(X_n > \sqrt{1-\varepsilon}\sqrt{2\log n} \right) \geq \left(\frac{1}{x} - \frac{1}{x^3}\right)\phi(x)\Big|_{x=\sqrt{1-\varepsilon}\sqrt{2\log n}}$$

using the limit (2), we have for such $\varepsilon/2 > 0$, there exists N s.t. $\forall n > N$ we have

$$\frac{\log \left[\left(\frac{1}{x} - \frac{1}{x^3}\right)\phi(x)\Big|_{x=\sqrt{1-\varepsilon}\sqrt{2\log n}} \right]}{\log n} \geq -(1-\varepsilon) - \frac{\varepsilon}{2} = -1 + \frac{\varepsilon}{2}$$
$$\Rightarrow \left(\frac{1}{x} - \frac{1}{x^3}\right)\phi(x)\Big|_{x=\sqrt{1-\varepsilon}\sqrt{2\log n}} \geq n^{-1+\varepsilon/2}, \quad \forall n > N$$

Then we have

$$\sum_{n=1}^{\infty} \mathbb{P}\left(X_n > \sqrt{1-\varepsilon}\sqrt{2\log n} \right) \geq \sum_{n=1}^{N} \mathbb{P}\left(X_n > \sqrt{1-\varepsilon}\sqrt{2\log n} \right) + \sum_{n=N+1}^{\infty} n^{-1+\varepsilon/2} = \infty$$

And we notice that $\{X_n\}$ are independent, thus by Borel-Cantelli lemma (2nd kind), we have

$$\mathbb{P}\left(\limsup_{n\to\infty} \frac{X_n}{\sqrt{2\log n}} > \sqrt{1-\varepsilon} \right) = 1, \quad \forall \varepsilon > 0$$

[2] which is simple to verify by noticing that

$$\frac{\log\log n}{\log n} \xrightarrow{n\to\infty} 0$$

Combining the two sides, we have

$$\begin{cases} \mathbb{P}\left(\limsup_{n\to\infty} \dfrac{X_n}{\sqrt{2\log n}} > \sqrt{1+\varepsilon}\right) = 0, \quad \forall\varepsilon > 0 \\ \mathbb{P}\left(\limsup_{n\to\infty} \dfrac{X_n}{\sqrt{2\log n}} > \sqrt{1-\varepsilon}\right) = 1, \quad \forall\varepsilon > 0 \end{cases} \Rightarrow \limsup_{n\to\infty} \frac{X_n}{\sqrt{2\log n}} = 1, \quad \text{a.s.}$$

3.(c)

Notice that for Normal distribution, we have

$$S_n = \sum_{i=1}^{n} X_i \sim N(0,n)$$

Define a constant ε accroding to $C = \sqrt{2}\sqrt{1+\varepsilon}$, then we have

$$\mathbb{P}\left(\frac{S_n}{\sqrt{n\log n}} > C\right) = \mathbb{P}\left(\frac{X}{\sqrt{2\log n}} > \sqrt{1+\varepsilon}\right) \leq n^{-1-\varepsilon/2} \text{ eventually}$$

Then following the same steps as in part (b), we have

$$\sum_{n=1}^{\infty} \mathbb{P}\left(\frac{S_n}{\sqrt{n\log n}} > C\right) < \infty$$

thus by Borel-Cantelli lemma (1st kind), we have

$$\forall C < \infty,\ \mathbb{P}\left(\limsup_{n\to\infty} \frac{S_n}{\sqrt{n\log n}} > C\right) = 0 \Rightarrow \limsup_{n\to\infty} \frac{S_n}{\sqrt{n\log n}} < C, \quad \text{a.s.}$$

Exercise 4 Poisson approximation to the Binomial distribution

Note that in HW3 we computed the following Fourier transform (characteristic function) for Poisson distribution and Binomial distribution:

$$\begin{aligned} \text{Binom: } & \binom{n}{x} p^x (1-p)^{n-x} \risingdotseq (pe^{it} + (1-p))^n \\ \text{Poisson: } & \frac{\lambda^x}{x!} e^{-\lambda} \risingdotseq e^{\lambda(e^{it}-1)} \end{aligned}$$

To show the convergence in distribution, we need to show the convergence of characteristic functions pointwisely. To do so, we notice that

$$\lim_{n\to\infty} \text{Binom}(n,p_n) \risingdotseq \lim_{n\to\infty} (p_n e^{it} + (1-p_n))^n = \lim_{n\to\infty} (1 + \frac{1}{n}(e^{it}-1)np_n)^n = e^{(e^{it}-1)\lambda} \fallingdotseq \frac{\lambda^x}{x!} e^{-\lambda} \sim \text{Poi}(\lambda)$$

thus finishes the proof of convergence in distribution.

Exercise 5 Exponential approximation to the geometric distribution

Note that in HW3 we computed the following Fourier transform (characteristic function) for Exponential distribution and Geometric distribution:

$$\begin{aligned} \text{Expo: } & \lambda e^{-\lambda x} \risingdotseq \frac{\lambda}{\lambda - it} \\ \text{Geometric: } & p(1-p)^x \risingdotseq \frac{p}{1-(1-p)e^{it}} \end{aligned}$$

To show the convergence in distribution, we need to show the convergence of characteristic functions pointwisely. To do so, we notice that

$$
\begin{aligned}
\lim_{p\to 0} p\mathrm{Geom}(p) \eqqcolon& \lim_{p\to 0} \frac{p}{1-(1-p)e^{itp}}\\
=& \lim_{p\to 0} \frac{1}{(1+(p-1)it)e^{itp}}\\
=& \frac{1}{1-it} \eqqcolon e^{-x} \sim \mathrm{Exp}(1)
\end{aligned}
$$

thus finishes the proof of convergence in distribution.