LECTURE NOTES IN PROBABILITY THEORY

GUAN-YU CHEN

1. Introduction

1.1. **Probabilities on finite sets.** Let's consider an experiment with finite number of possible outcomes, say $\omega_1, ..., \omega_n$. The set $\Omega = \{\omega_1, ..., \omega_n\}$ is called the sample space and ω_i is named as a sample point. An event is a subset of Ω. A probability on Ω is a function $\mathbb{P}$ defined on Ω satisfying

(1) $0 \le \mathbb{P}(\omega_i) \le 1$ for all $1 \le i \le n$.
(2) $\sum_{i=1}^n \mathbb{P}(\omega_i) = 1$.

The probability of event $E \subset \Omega$ is defined by

$$\mathbb{P}(E) = \sum_{\omega_i \in E} \mathbb{P}(\omega_i).$$

Let $\mathcal{F}$ be the collection of all subsets of Ω. The triple $(\Omega, \mathcal{F}, \mathbb{P})$ is named a probability space.

Remark 1.1. It is clear that if E and F are mutually disjoint events, then $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F)$. Moreover, it is clear that $\mathbb{P}(\Omega) = 1$, $\mathbb{P}(\emptyset) = 0$, and $\mathbb{P}(E \cup F) = \mathbb{P}(E) + \mathbb{P}(F) - \mathbb{P}(E \cap F)$, where $\emptyset$ is the empty set.

Let E, F be events and assume that $\mathbb{P}(E) > 0$. The conditional probability of F given E is defined by

$$\mathbb{P}(F|E) = \frac{\mathbb{P}(E \cap F)}{\mathbb{P}(E)}.$$

Two events are said to be independent if $\mathbb{P}(E \cap F) = \mathbb{P}(E)\mathbb{P}(F)$.

Remark 1.2. Any event E is independent of Ω and $\emptyset$. Furthermore, if E and F are independent, then a set in $\{E, E^c, \Omega, \emptyset\}$ and a set in $\{F, F^c, \Omega, \emptyset\}$ are independent, where E^c denotes the complement of E in Ω.

Remark 1.3. Let $E_1, ..., E_m$ be mutually disjoint events such that $\bigcup_{i=1}^m E_i = \Omega$ and $\mathbb{P}(E_i) > 0$ for $1 \le i \le m$. Bayes's formula says that, for any event F satisfying $\mathbb{P}(F) > 0$,

$$\mathbb{P}(E_i|F) = \frac{\mathbb{P}(E_i)\mathbb{P}(F|E_i)}{\sum_{j=1}^m \mathbb{P}(E_j)\mathbb{P}(F|E_j)}, \quad \forall 1 \le i \le m.$$

A random variable is a real-valued function defined on Ω. For any random variable X and any subset $B \subset \mathbb{R}$, we write $\{X \in B\}$ for the set $\{\omega \in \Omega | X(\omega) \in B\}$. If $B = [a, b]$, we also write $\{a \le X \le b\}$ for $\{X \in [a, b]\}$. For random variables $X_1, X_2, ...$ and subsets $B_1, B_2, ...$, both $\{X_1 \in B_1, X_2 \in B_2, ...\}$ and $\{X_i \in B_i, \forall i\}$ denote the set $\bigcap_i \{X_i \in B_i\}$.

Two random variables X, Y are said to be independent if $\{X \le x\}$ and $\{Y \le y\}$ are independent for all $x, y \in \mathbb{R}$. For any random variable X, the distribution of X is defined by $F_X(x) = \mathbb{P}(X \le x)$ and the expectation and variance are defined by

$$\mathbb{E}(X) := \sum_{i=1}^n X(\omega_i)\mathbb{P}(\omega_i), \quad \mathrm{Var}(X) = \mathbb{E}(X - \mathbb{E}X)^2.$$

Remark 1.4. Let X be a random variable and F_X be the distribution of X.

(1) F_X is a right-continuous non-decreasing function.
(2) $F_X(x) \to 0$ as $x \to -\infty$ and $F_X(x) \to 1$ as $x \to \infty$.

Furthermore, $\mathbb{E}(X) = \int_{-\infty}^{\infty} x dF_X(x)$, where the right side is known as the Riemann-Stieljes integral.

A sequence of random variables $X_1, X_2, ...$ is said to be identically distributed if they have the same distribution. A sequence of random variables $X_1, X_2, ...$ is said to be independent if, for any n and $x_1, ..., x_n \in \mathbb{R}$, $\mathbb{P}(X_i \le x_i, \forall 1 \le i \le n) = \prod_{i=1}^n \mathbb{P}(X_i \le x_i)$.

Law of large numbers: Consider a sequence of i.i.d. (independent and identically distributed) random variables, $X_1, X_2, ...$ Set $S_n = X_1 + \cdots + X_n$. Suppose $-\infty < \mu = \mathbb{E}(X_1) < \infty$. Then,

(1) (weak version) $\mathbb{P}(|S_n/n - \mu| > \epsilon) \to 0$ as $n \to \infty$ for all $\epsilon > 0$.
(2) (strong version) $\mathbb{P}(S_n/n \to \mu) = 1$.

Central limit theorem: Let $X_1, X_2, ...$ be i.i.d. random variables and set $S_n = \sum_{i=1}^n X_i$. Suppose $\mu = \mathbb{E}X_1$ exists and $0 < \sigma^2 = \mathrm{Var}(X_1) < \infty$. Then,

$$\lim_{n\to\infty} \mathbb{P}\left(a \le \frac{S_n - n\mu}{\sqrt{n\sigma^2}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx, \quad \forall a < b.$$

Remark 1.5. In the law of large numbers, the weak version says that S_n/n converges to μ in probability and the strong version says that S_n/n converges to μ almost surely or with probability 1. The central limit theorem says that $(S_n - n\mu)/(\sqrt{n}\sigma)$ converges in distribution to the standard normal random variable.

1.2. **Independent tosses of a fair coin.** Let n be a positive integer. Consider an experiment of tossing a fair coin independently for n times, that is,

(a) There are 2^n outcomes, which are all $\{H, T\}$-valued n-vectors, where "H" and "T" represent for "Head" and "Tail".
(b) Each of the 2^n outcomes are equally likely to occur, namely, every n-vector has probability 2^{-n}.

Let $(\Omega_n, \mathcal{F}_n, \mathbb{P}_n)$ be the corresponding probability space. It is easy to see that Ω_n consists of all $\{H, T\}$-valued n-vectors, $\mathcal{F}_n$ is the collection of all subsets of Ω_n and $\mathbb{P}_n$ is the probability on $\mathcal{F}_n$, which is uniform over all sample points in Ω_n.

Concerning the experiment of flipping a fair coin for infinitely many times, we set $\Omega = \{\omega = (\omega_1, \omega_2, ...) | \omega_i \in \{H, T\}, \forall i\}$ and define

$$\mathcal{E}(x) = \{\omega \in \Omega | (\omega_1, ..., \omega_n) = x\}, \quad \forall x \in \Omega_n,\ n \ge 1,$$

and

$$\mathcal{E}(A) = \bigcup_{x \in A} \mathcal{E}(x), \quad \forall A \in \mathcal{F}_n,\ n \ge 1.$$

Clearly, $\mathcal{E}(A) = A \times \Omega$ for all $A \in \bigcup_{n\ge1} \mathcal{F}_n$. By defining $\mathcal{E}(\mathcal{F}_n) := \{\mathcal{E}(A) | A \in \mathcal{F}_n\}$, one has $\mathcal{E}(\mathcal{F}_n) \subset \mathcal{E}(\mathcal{F}_{n+1})$. To discuss the law of large numbers and the central limit theorem, we address the following assumption.

Suppose there is a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, where $\mathcal{F}$ is some structured collection of subsets of Ω (say, a σ-field) containing $\bigcup_{n\ge1} \mathcal{E}(\mathcal{F}_n)$ and $\mathbb{P}(\mathcal{E}(A)) = \mathbb{P}_n(A)$ for all $A \in \mathcal{F}_n$ and $n \ge 1$.

To quantify the model of flipping coins, we set $X_i(\omega) = \mathbf{1}_H(\omega_i)$. It is easy to check that $X_1, X_2, ...$ are i.i.d. with $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = 0) = 1/2$. The law of large number says

$$\lim_{n\to\infty} \mathbb{P}(|S_n/n - 1/2| > \epsilon) = 0,\ \forall \epsilon > 0, \quad \mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}\right) = 1,$$

where $S_n = \sum_{i=1}^n X_i$, while the central limit theorem refers to the limit of

$$\lim_{n\to\infty} \mathbb{P}\left(a \le \frac{S_n - n/2}{\sqrt{n/4}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx, \quad \forall a < b.$$

In the next three subsections, we will give rigorous proofs of the above theorems.

1.3. **The weak law of large numbers for i.i.d. Bernoulli sequences.**

Theorem 1.1 (Weak law of large numbers)**.** *Let S_n be the number of heads in the first n independent tosses of a fair coin. Then, for $\epsilon > 0$,*

$$\lim_{n\to\infty} \mathbb{P}(|S_n/n - 1/2| \ge \epsilon) = 0.$$

To prove the above theorem, we need the Chebyshev inequality.

Proposition 1.2 (Chebyshev inequality)**.** *Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space with $|\Omega| < \infty$ and X be a random variable on Ω. Then, for any $\epsilon > 0$,*

$$\mathbb{P}(|X| \ge \epsilon) \le \frac{\mathbb{E}X^2}{\epsilon^2}.$$

Exercise 1.1. Prove the above proposition.

Proof of Theorem 1.1. For $n \ge 1$ and $1 \le i \le n$, let $X_{n,i}$ be the random variable defined by

$$X_{n,i}(\omega) = \mathbf{1}_H(\omega_i) = \begin{cases} 1 & \text{if the } i\text{-th entry of } \omega \text{ is } H, \\ 0 & \text{if the } i\text{-th entry of } \omega \text{ is } T, \end{cases} \quad \forall \omega = (\omega_1, ..., \omega_n) \in \Omega_n.$$

Set $S_n' = X_{n,1} + X_{n,2} + \cdots + X_{n,n}$. Obviously, one has

$$\{|S_n/n - 1/2| \ge \epsilon\} = \mathcal{E}(\{|S_n'/n - 1/2| \ge \epsilon\}).$$

This implies

$$\mathbb{P}(|S_n/n - 1/2| \ge \epsilon) = \mathbb{P}_n(|S_n'/n - 1/2| \ge \epsilon).$$

Observe that

$$\mathbb{E}X_{n,i}^2 = \mathbb{E}X_{n,i} = \mathbb{P}_n(\{(\omega_1, ..., \omega_n) | \omega_i = H\}) = 1/2$$

and, for $i \ne j$,

$$\mathbb{E}(X_{n,i}X_{n,j}) = \mathbb{P}_n(\{(\omega_1, ..., \omega_n) | \omega_i = \omega_j = H\}) = 1/4.$$

Using the linearity of the expectation, we have

$$\forall 1 \le i \le n, \quad \mathbb{E}(X_{n,i} - 1/2)^2 = \mathbb{E}X_{n,i}^2 - \mathbb{E}X_{n,i} + 1/4 = 1/4$$

and

$$\forall i \ne j, \quad \mathbb{E}[(X_{n,i} - 1/2)(X_{n,j} - 1/2)] = 0.$$

Write

$$\frac{S_n'}{n} - \frac{1}{2} = \frac{1}{n} \sum_{i=1}^n (X_{n,i} - 1/2).$$

As a result of the above computations, we obtain

$$\begin{aligned} \mathbb{E}\left(\frac{S_n'}{n} - \frac{1}{2}\right)^2 &= \frac{1}{n^2} \sum_{i,j=1}^n \mathbb{E}[(X_{n,i} - 1/2)(X_{n,j} - 1/2)] \\ &= \frac{1}{n^2} \sum_{i=1}^n \mathbb{E}(X_{n,i} - 1/2)^2 = \frac{1}{4n}. \end{aligned}$$

Consequently, the Chebyshev inequality implies

$$\mathbb{P}_n(|S'_n/n - 1/2| \geq \epsilon) \leq 1/(4n\epsilon^2) \to 0,$$

as $n \to \infty$. □

Remark 1.6. In the above proof, the last inequality says that the probability $\mathbb{P}(|S_n/n - 1/2| \geq \epsilon)$ converges to 0 at least polynomially. We refer the readers to the topic of large deviation for a precise estimation of this convergence.

Note that the law of large numbers does not mean $S_n/n = 1/2$ (in any suitable sense). In fact, one has

$$\max_{0 \leq k \leq n} \mathbb{P}(S_n = k) = \mathbb{P}(S_n = \lfloor n/2 \rfloor), \quad \lim_{n\to\infty} \mathbb{P}(S_n/n = 1/2) = 0. \tag{1.1}$$

where $\lfloor t \rfloor := \max\{n \in \mathbb{Z} | n \leq t\}$. To see a proof, we need the following facts.

Lemma 1.3. *For the n independent tosses of a fair coin, there are $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ n-vectors with exactly k heads appear.*

Stirling's formula

$$n! = \sqrt{2\pi} e^{-n} n^{n+1/2} (1 + \epsilon_n), \tag{1.2}$$

where ϵ_n converges to 0 as n tends to infinity. More precisely, it holds true that

$$\sqrt{2\pi} n^{n+1/2} e^{-n+1/(12n+1)} < n! < \sqrt{2\pi} n^{n+1/2} e^{-n+1/(12n)} \tag{1.3}$$

and this implies that $\epsilon_n = \frac{1}{12n} + O\left(\frac{1}{n^2}\right)$.

The first equality in (1.1) is obvious from Lemma 1.3. To see the limit, we let S'_n be the random variable in the proof of Theorem 1.1. Note that

$$\mathbb{P}(S_n/n = 1/2) = \mathbb{P}_n(S'_n/n = 1/2) = \begin{cases} 0 & \text{if } n \text{ is odd} \\ 2^{-n}\binom{n}{n/2} & \text{if } n \text{ is even} \end{cases}.$$

In the case that $n = 2m$, one has

$$2^{-2m}\binom{2m}{m} = 2^{-2m}\frac{(2m)!}{m!m!} = 2^{-2m}\frac{\sqrt{2\pi}e^{-2m}(2m)^{2m+1/2}(1+\epsilon_{2m})}{2\pi e^{-2m}m^{2m+1}(1+\epsilon_m)^2} = \frac{1}{\sqrt{\pi m}}(1+\delta_m).$$

Clearly, $\delta_m \to 0$ as $m \to \infty$.

In fact, there is a more general estimation. For $0 \leq k \leq n$, we write

$$\mathbb{P}(S_n = k) = \sqrt{\frac{2}{\pi n}}(1 + \delta_{n,k}). \tag{1.4}$$

Exercise 1.2. Let $K_n > 0$ be a sequence satisfying $K_n = o(\sqrt{n})$ and $\delta_{n,k}$ be the constant in (1.4). Show that

$$\max_{k:|k-n/2|<K_n} |\delta_{n,k}| \to 0, \quad \text{as } n \to \infty.$$

Hint: Use Stirling's formula to derive

$$\delta_{n,k} = \frac{n^{n+1}(1+\epsilon_n)}{(2k)^{k+1/2}(2n-2k)^{n-k+1/2}(1+\epsilon_k)(1+\epsilon_{n-k})} - 1.$$

One may conclude from the above exercise that, for $K_n \to \infty$ and $K_n = o(\sqrt{n})$,

$$\mathbb{P}(|S_n - n/2| \leq K_n) = \frac{2\sqrt{2}K_n}{\sqrt{\pi n}}(1 + o(1)).$$

1.4. **The central limit theorem for i.i.d. Bernoulli sequences.**

Theorem 1.4 (The central limit theorem). *Let S_n denote the number of heads in the first n independent tosses of a fair coin and Φ be a function defined by*

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt.$$

For $x \ge 0$, one has

$$\lim_{n\to\infty} \mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| < \frac{x}{2\sqrt{n}} \right) = \Phi(x) - \Phi(-x). \tag{1.5}$$

In particular, for $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{S_n}{n} - \frac{1}{2} < \frac{x}{2\sqrt{n}} \right) = \Phi(x).$$

Proof. Let S'_n be the random variable defined in the proof of the weak law of large numbers. As before, we have

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \frac{1}{2} \right| < \frac{x}{2\sqrt{n}} \right) = \mathbb{P}_n\left(\left| \frac{S'_n}{n} - \frac{1}{2} \right| < \frac{x}{2\sqrt{n}} \right), \quad \forall x > 0.$$

First, consider the case $n = 2m$. For $m \ge 1$ and $x \ge 0$, set

$$\begin{aligned} q_{m,x} &= \mathbb{P}_{2m}(|S'_{2m} - m| < x\sqrt{m/2}) \\ &= 2^{-2m} \sum_{k:|k-m|<x\sqrt{m/2}} \binom{2m}{k} = 2^{-2m} \sum_{j\in R_m} \binom{2m}{m+j}, \end{aligned}$$

where $R_m = \{j : |j| < x\sqrt{m/2}\}$. Write $2^{-2m}\binom{2m}{m+j} = p_m D_{m,j}$ with $p_m = \mathbb{P}_{2m}(S'_{2m} = m)$ and

$$D_{m,j}^{-1} = \left(1 + \frac{|j|}{m}\right)\left(1 + \frac{|j|}{m-1}\right) \cdots \left(1 + \frac{|j|}{m - |j| + 1}\right).$$

Using the expression $\log(1 + t) = t(1 + f(t))$, we have

$$\log D_{m,j} = -\sum_{k=0}^{|j|-1} \frac{|j|}{m-k} \left(1 + f\left(\frac{|j|}{m-k}\right)\right).$$

This leads to $D_{m,j} = e^{-(1+\epsilon_{m,j})j^2/m}$ with $\epsilon_{m,0} = 0$ and, for $j \ne 0$,

$$\epsilon_{m,j} = \frac{1}{|j|} \sum_{k=0}^{|j|-1} \frac{k}{m-k} + \frac{1}{|j|} \sum_{k=0}^{|j|-1} \frac{m}{m-k} f\left(\frac{|j|}{m-k}\right).$$

Using the fact that $f(t)/t \to -1/2$ as $t \to 0$, one can show that, as $m \to \infty$,

$$f\left(\frac{|j|}{m-k}\right) = \frac{|j|}{m-k}\left(-\frac{1}{2} + o(1)\right),$$

where $o(1)$ is uniformly for $0 \le k < |j|$ and $j \in R_m$. Consequently, the above computation yields $\epsilon_{m,j} = O(1/\sqrt{m})$ uniformly for $j \in R_m$.

Next, we write $D_{m,j} = (1 + \Delta_{m,j})e^{-j^2/m}$ with $\Delta_{m,j} = e^{-\epsilon_{m,j}j^2/m} - 1$. As $t \to 0$, one has $e^t = 1 + t(1 + o(1))$. This implies $\Delta_{m,j} = O(1/\sqrt{m})$ uniformly for $j \in R_m$ as $m \to \infty$. Recall the notation $p_m = \frac{1}{\sqrt{\pi m}}(1 + \delta_m)$ and set

$$q_{m,x} = \sum_{j\in R_m} p_m D_{m,j} = A_m + B_m,$$

where

$$A_m = \sum_{j\in R_m} \frac{e^{-j^2/m}}{\sqrt{\pi m}}, \quad B_m = \frac{1}{\sqrt{\pi m}} \sum_{j\in R_m} \left(\Delta_{m,j} e^{-j^2/m} + \delta_m + \delta_m \Delta_{m,j} e^{-j^2/m}\right).$$

Note that

$$|B_m| \le \frac{1}{\sqrt{\pi m}} \sum_{j\in R_m} (|\Delta_{m,j}| + |\delta_m| + |\Delta_{m,j}\delta_m|).$$

When $m \to \infty$, A_m converges to $\frac{1}{\sqrt{2\pi}} \int_{-x}^{x} e^{-t^2/2} dt$ and $|B_m| = O(|\delta_m| + 1/\sqrt{m})$. This proves the limit in (1.5) with even n.

For the case $n = 2m+1$, let $h > 0$. Observe that, when n is large enough, one has

$$\begin{aligned} &\left\{(w,v) \in \Omega_{2m} \times \Omega_1 : |S'_{2m}(w) - m| < \frac{x-h}{2}\sqrt{2m}\right\} \\ \subset &\left\{(w,v) \in \Omega_{2m+1} : \left|S'_{2m+1}(w,v) - \left(m + \frac{1}{2}\right)\right| < \frac{x}{2}\sqrt{2m+1}\right\} \\ \subset &\left\{(w,v) \in \Omega_{2m} \times \Omega_1 : |S'_{2m}(w) - m| < \frac{x+h}{2}\sqrt{2m}\right\} \end{aligned}$$

Letting n tend to infinity derives

$$\Phi(x-h) - \Phi(h-x) \le \liminf_{m\to\infty} \mathbb{P}_{2m+1}\left(\left|S'_{2m+1} - \frac{2m+1}{2}\right| < \frac{x}{2}\sqrt{2m+1}\right)$$

and

$$\limsup_{m\to\infty} \mathbb{P}_{2m+1}\left(\left|S'_{2m+1} - \frac{2m+1}{2}\right| < \frac{x}{2}\sqrt{2m+1}\right) \le \Phi(x+h) - \Phi(-x-h)$$

Since Φ is continuous, letting $h \to 0$ gives the desired identity in (1.5). □

1.5. **The strong law of large numbers for i.i.d. Bernoulli sequences.** The law of large numbers and the central limit theorem provide us a way of comparing the sample mean of heads S_n/n and the probability of heads 1/2. It is natural to arise the following question: Could it be possible that no matter how the first n tosses are, eventually things will settle down and smooth out in the way that

$$\lim_{n\to\infty} \frac{S_n}{n} = \frac{1}{2}.$$

Clearly, this can fail if all tosses result in heads.

Theorem 1.5 (Strong law of large numbers)**.** *Let S_n be the number of heads in the first n independent tosses of a fair coin. Then, $\mathbb{P}(S_n/n \to 1/2) = 1$.*

As before, let $X_1, X_2, ...$ be the i.i.d. sequence of random variables satisfying $\mathbb{P}(X_1 = 1) = \mathbb{P}(X_1 = 0) = 1/2$ and set $S_n = X_1 + X_2 + \cdots + X_n$. To prove the above theorem, we need the following lemma.

Lemma 1.6. *Let Ω be the sample space on which $X_1, X_2, ...$ are defined and $S_n = \sum_{i=1}^n X_i$. Then, for $\omega \in \Omega$,*

$$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = \frac{1}{2} \quad \Leftrightarrow \quad \lim_{m\to\infty} \frac{S_{m^2}(\omega)}{m^2} = \frac{1}{2}.$$

Proof. For $n \geq 1$, let m be a positive integer satisfying $m^2 \leq n < (m+1)^2$. It is obvious that $0 \leq n - m^2 \leq 2m$ and this implies

$$\begin{aligned}\left|\frac{S_n(\omega)}{n} - \frac{S_{m^2}(\omega)}{m^2}\right| &= \left|\frac{S_n(\omega)}{m^2} - \frac{S_{m^2}(\omega)}{m^2} + \left(\frac{1}{n} - \frac{1}{m^2}\right) S_n(\omega)\right| \\ &\leq \frac{|n-m^2|}{m^2} + n\left|\frac{1}{n} - \frac{1}{m^2}\right| = \frac{2|n-m^2|}{m^2} \leq \frac{4}{m}.\end{aligned}$$

Letting $m \to \infty$ gives the desired property. □

Proof of Theorem 1.5. Let $F = \{\omega : S_n(\omega)/n \nrightarrow 1/2\}$ and, for $\epsilon > 0$, set

$$F_\epsilon = \left\{\omega : \left|\frac{S_{m^2}(\omega)}{m^2} - \frac{1}{2}\right| > \epsilon \text{ for infinitely many } m\right\}.$$

Note that $F_{\epsilon_1} \subset F_{\epsilon_2}$ for $\epsilon_1 > \epsilon_2$ and $F = \bigcup_{k=1}^\infty F_{1/k}$. Consider the following two assumptions.

Assumption 1: If $B_1, B_2, ... \in \mathcal{F}$, then $\bigcup_{n=1}^\infty B_n \in \mathcal{F}$ and $\bigcap_{n=1}^\infty B_n \in \mathcal{F}$.

Assumption 2: If $A_n \subset \Omega$ is a sequence satisfying $A_n \subset A_{n+1}$ (resp. $A_n \supset A_{n+1}$), then

$$\mathbb{P}\left(\bigcup_{n=1}^\infty A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n). \quad \left(\text{resp. } \mathbb{P}\left(\bigcap_{n=1}^\infty A_n\right) = \lim_{n\to\infty} \mathbb{P}(A_n).\right)$$

To prove $\mathbb{P}(F) = 0$, it is equivalent to show that $\mathbb{P}(F_\epsilon) = 0$ for all $\epsilon > 0$. Now, fix $\epsilon > 0$ and set

$$E_{m_1,m_2} = \bigcup_{m=m_1}^{m_2} \left\{\omega : \left|\frac{S_{m^2}(\omega)}{m^2} - \frac{1}{2}\right| > \epsilon\right\},$$

for positive integers $m_1 \leq m_2$. For $m_1 \in \mathbb{N}$, define

$$E_{m_1} = \bigcup_{m_2=m_1}^\infty E_{m_1,m_2} = \bigcup_{m_2=m_1}^\infty \left\{\omega : \left|\frac{S_{m_2^2}(\omega)}{m_2^2} - \frac{1}{2}\right| > \epsilon\right\}.$$

Clearly, one has $F_\epsilon = \bigcap_{m_1=1}^\infty E_{m_1} = \bigcap_{m_1=1}^\infty \bigcup_{m_2=m_1}^\infty E_{m_1,m_2}$.

For convenience, if $A_n \subset A_{n+1}$ for all n, we write $\lim_n A_n$ for $\bigcup_n A_n$. If $A_n \supset A_{n+1}$ for all n, we write $\lim_n A_n$ for $\bigcap_n A_n$. Using the above notations, we may rewrite

$$E_{m_1} = \lim_{m_2\to\infty} E_{m_1,m_2}, \quad F_\epsilon = \lim_{m_1\to\infty} E_{m_1}.$$

By *Assumption 2*, one has

$$\mathbb{P}\left(\bigcap_{m_1=1}^\infty \bigcup_{m_2=m_1}^\infty E_{m_1,m_2}\right) = \lim_{m_1\to\infty}\lim_{m_2\to\infty} \mathbb{P}(E_{m_1,m_2}).$$

Note that E_{m_1,m_2} is determined by $X_{m_1}, ..., X_{m_2}$. Recall the notation $S'_n = X_{n,1} + \cdots + X_{n,n}$. By the Chebyshev inequality, one has

$$\mathbb{P}(E_{m_1,m_2}) \leq \sum_{m=m_1}^{m_2} \mathbb{P}_{m^2}\left(\left|\frac{S'_{m^2}}{m^2} - \frac{1}{2}\right| > \epsilon\right) \leq \frac{1}{4\epsilon^2}\sum_{m=m_1}^{m_2} \frac{1}{m^2}.$$

As a result, this leads to

$$\mathbb{P}(F_\epsilon) = \lim_{m_1\to\infty}\lim_{m_2\to\infty} \mathbb{P}(E_{m_1,m_2}) \leq \frac{1}{4\epsilon^2}\lim_{m_1\to\infty}\lim_{m_2\to\infty}\sum_{m=m_1}^{m_2} \frac{1}{m^2} = 0.$$

□