PROBABILITY THEORY: LECTURE NOTES 4

CHANDRA NAIR

Disclaimer: These notes are created from a variety of great references in the field, most notably, Varadhan's lecture notes and Dembo's lecture notes.

5. Independence

Definition 5.1. Events A and B are independent if $P(A \cap B) = P(A)P(B)$.

Definition 5.2. Random variables X and Y are independent if $\forall A, B \in \mathcal{B}_R$,

$$P\{\omega : X(\omega) \in A, Y(\omega) \in B\} = P\{\omega : X(\omega) \in A\} P\{\omega : Y(\omega) \in B\}.$$

The definition of independence extends to a finite number of random variables. $X_1, ...X_n$ are independent if $\forall A_i \in \mathcal{B}_R$,

$$P(X_1 \in A_1, ..., X_n \in A_n) = \prod_{i=1}^{n} P(X_i \in A_i).$$

Exercise 5.1. Two random variables X and Y defined on the same probability space are independent iff the measure induced by the joint mapping $(\omega_1, \omega_2) \to (X(\omega_1), Y(\omega_2))$, is the *product measure*.

Exercise 5.2. Show that for independent random variables X and Y and measurable functions f and g where $E[|f(X)|]$ and $E[|g(Y)|]$ are finite

$$E[f(X)g(Y)] = E[f(X)]E[g(Y)].$$

Remark 5.1. This implies that $\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$. Note that the reverse does not hold, meaning that

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t) \nRightarrow X \text{ and } Y \text{ are independent}$$

On the other hand note that

$$E[e^{i(t_1 X + t_2 Y)}] = E[e^{it_1 X}]E[e^{it_2 Y}] \quad \forall t_1, t_2 \iff X \text{ and } Y \text{ are independent}$$

5.1. **Weak Law of Large Numbers.**

Theorem 5.1. *$X_1, ..., X_n$ are independent random variables, satisfying $E(X_i) = 0$ and $E(X_i^2) = \sigma_i^2 < B$. Then $\frac{S_n}{n} \to 0$ in measure, where $S_n = X_1 + ... + X_n$.*

Proof. Note that

$$E\left(\frac{S_n}{n}\right)^2 = \frac{1}{n^2}E(X_i^2) \leq \frac{B}{n}.$$

Hence

$$\epsilon^2 P\left(\left|\frac{S_n}{n}\right| > \epsilon\right) \leq E\left(\frac{S_n}{n}\right)^2 \leq \frac{B}{n}.$$

□

Date: October 21, 2019.

Definition 5.3. A collection of random variables $\{X_\alpha\}$ is called *uniformly integrable* (U-I) if

$$\lim_{M\to\infty} \sup_\alpha E(|X_\alpha| 1_{|X_\alpha|>M}) = 0.$$

Theorem 5.2. *$X_1, ..., X_n, ..$ are independent and uniformly integrable random variables with mean zero, then $\frac{S_n}{n} \to 0$ in measure.*

Proof. Given $\delta > 0$, take M (it exists by uniform integrability) such that

$$E\left(|X_i| 1_{|X_i|>M}\right) < \delta, \quad \forall i.$$

Truncate X_i to two parts X_i^T and Y_i^T as follows

$$\begin{array}{ll} X_i^T = X_i \mathbf{1}_{\{|X_i| \le M\}} & E[X_i^T] = a_i \\ Y_i^T = X_i \mathbf{1}_{\{|X_i| > M\}} & E[Y_i^T] = -a_i \\ G_n^T = X_1^T + ... + X_n^T - \sum_{i=1}^n a_i & B_n^T = Y_1^T + ... + Y_n^T + \sum_{i=1}^n a_i. \end{array}$$

Note that $|a_i| < \delta$ for all i. Clearly $X_i = X_i^T + Y_i^T$. Now consider the quantity

$$\begin{aligned} E[|\frac{S_n}{n}|] &= E[|\frac{G_n^T + B_n^T}{n}|] \le E[|\frac{G_n^T}{n}|] + E[|\frac{B_n^T}{n}|] \\ &\le E[|\frac{G_n^T}{n}|] + 2\delta. \end{aligned} \tag{5.1}$$

Where the last inequality follows from

$$E[|\frac{B_n^T}{n}|] = E[|\frac{\sum Y_i^T + a_i}{n}|] \le \sum \frac{E[|Y_i^T|] + |a_i|}{n} \overset{U-I}{\le} 2\delta.$$

Note that

$$E((X_i^T - a_i)^2) \le (M + |a_i|)^2 \le (M + \delta)^2.$$

Therefore (from Cauchy-Schwartz)

$$E[|\frac{G_n^T}{n}|] \le \sqrt{E[(\frac{G_n^T}{n})^2]} \le \sqrt{\frac{1}{n}(M+\delta)^2}.$$

By Markov's inequality

$$P(|\frac{S_n}{n}| > \epsilon) \le \frac{E[|\frac{S_n}{n}|]}{\epsilon} \le \frac{\frac{1}{\sqrt{n}}(M+\delta) + 2\delta}{\epsilon}.$$

Taking limsup we get $P(|\frac{S_n}{n}| > \epsilon)$ is upper bounded by $\frac{2\delta}{\epsilon}$; but since $\delta > 0$ is arbitrary, we are done. □

5.2. **Strong Law of Large Numbers.** So far we considered convergence in measure of a sequence of random variables. We could get away without showing that one can actually define an infinite sequence of independent random variables. But when we talk about the strong law; then we do need to show that one can define such a probability measure. To do so, let us start with a Lemma.

Lemma 5.3. *Given any Borel set $A \subset \mathbb{R}^n$ and a probability measure P, for any $\epsilon > 0$ there exists a closed and bounded set $K_\epsilon \subset A$ such that $P(A \setminus K_\epsilon) < \epsilon$.*

Proof. Boundedness of K_ϵ is simple. Just truncate the space to within a closed ball, whose outside probability is at most, say $\frac{\epsilon}{3}$. Let us try to prove it for $n = 1$. let $A = (a, b]$. By right continuity the probability of $[a + \frac{1}{n}, b]$ goes to $(a, b]$. Hence the statement is true for such intervals, as well as finite disjoint union of such intervals. The set of all such A is a monotone class (why?) (Hint: truncate the countable union to a finite union with a small probability loss; for each set now approximate by closed set, and take their finite union. Quantifiers can be chosen show that total probability loss is small.)

For larger n start with finite disjoint union of rectangles and proceed similarly. $\square$

A family of probability measures P_n on $\mathbb{R}^n$ is said to be *consistent* if $P_{n+1}((A, \mathbb{R})) = P_n(A)$ for all A, a Borel set in $\mathbb{R}^n$.

Let $\Omega = \{x_1, ..., x_n, ...\} = \mathbb{R}^\infty$ be the space of all real sequences. Consider the natural sigma field, Σ, generated by the field of all finite dimensional cylinder sets, i.e sets of the for $A \times \mathbb{R} \times \mathbb{R} \cdots$, where A is Borel set in $\mathbb{R}^n$ for some n.

Theorem 5.4 (Kolmogorov's consistency theorem)**.** *Given a consistent family of finite-dimensional distributions P_n, there is a unique P on (Ω, Σ) such that for every n, under the natural projection, the induced measure $P\Pi_n^{-1}(A) = P_n(A)$ for all $A \in \mathcal{B}_n$.*

Proof. Consider the field of all finite dimensional cylinder sets. For A in this field define $P(A) = P_n(A)$ (defined uniquely because of the finite-dimensional consistency of the family). Suffices to show that this P is countably additive on the field (Caratheodory's extension theorem does the rest). Assume not. Hence there is $A_n \downarrow \varnothing$ but $P(A_n) \geq \delta > 0$. Since A_n is a finite dimensional cylinder set of $B_{k_n} \in \mathbb{R}^{k_n}$, from lemma, we can find a closed set $K_{k_n} \subset B_{k_n}$ with $P_n(B_{k_n} \setminus K_{k_n}) \leq \frac{\delta}{2^{n+2}}$. Let $C_n \subset A_n$ denote the closed cylinder generated by K_{k_n}. Let $D_n = \cap_{m \leq n} C_m$. Note that

$$A_n \setminus D_n = \cup_{m \leq n}(A_n \setminus C_m) \subset \cup_{m \leq n}(A_m \setminus C_m).$$

Hence

$$P(A_n \setminus D_n) \leq \sum_{m=1}^{n} P(A_m \setminus C_m) = \sum_{m=1}^{n} P_n(B_{k_m} \setminus K_{k_m}) \leq \sum_{m=1}^{n} \frac{\delta}{2^{m+2}} \leq \frac{\delta}{4}.$$

Therefore $P(D_n) \geq \frac{3\delta}{4}$, hence D_n is non-empty closet set of Ω. Further $D_n \downarrow \varnothing$. Take a sequence of points (in n) ω^n from D_n. Since D_n are non-empty decreasing cylinder sets, by considering the sequence of its projections and using the diagonalization argument we can find a limit point $(x_1, ..., x_m, ...)$ such that this limit point belongs to D_m for every m, contradicting $D_m \downarrow \varnothing$. $\square$

We first prove a vanilla version of the main theorem by imposing a constraint on the fourth moment of the random variables.

Theorem 5.5. *$X_1, ..., X_n$ are independent random variables such that $E(X_i) = 0$ and $E[X_i^4] = c < \infty$. Then $\frac{S_n}{n} \to 0$ almost surely.*

Proof. Consider

$$E\left(\frac{S_n}{n}^4\right) = \frac{1}{n^4}\left(\sum_i E(X_i^4) + 6\sum_{i<j} E(X_i^2)(X_j^2)\right). \tag{5.2}$$

Note that all the other cross terms are zero. (why?). Further $E(X_i^2) \leq \sqrt{c}$. Therefore

$$E\left(\frac{S_n}{n}^4\right) \leq \frac{1}{n^4}(n + 3n(n-1))c \leq \frac{3c}{n^2}.$$

Therefore, by Markov's inequality,

$$P(|\frac{S_n}{n}| > \epsilon) \leq \frac{1}{\epsilon^4} E\left(\frac{S_n}{n}^4\right) \leq \frac{3c}{\epsilon^4 n^2}.$$

Let $A_n = \{\omega : |\frac{S_n}{n}| > \epsilon\}$. Since $\sum_n P(A_n) < \infty$, from Borel-Cantelli Lemma we know that $P(A_n \ i.o.) = 0$.

Note that the set $B_\epsilon := A_n \ i.o$ is same as the set of omega's for which $\limsup_n |\frac{S_n}{n}| > \epsilon$. We have $P(B_\epsilon) = 0$ for any $\epsilon > 0$; hence $P(\cup_m B_{1/m}) = 0$. Therefore the set of omega's for which $\limsup_n |\frac{S_n}{n}| > 0$ has probability 0, establishing almost sure convergence.

□

Theorem 5.6. *Let $X_1, \ldots, X_n, \ldots$ be pairwise independent and identically distributed non-negative random variables (say same as X) with $E(X) = \mu < \infty$ and $S_n = \sum_{i=1}^n X_i$. Then $\frac{S_n}{n} \to \mu \ \ a.s.$*

Proof. Step 1: (Truncation) Let $Y_k = X_k 1_{(X_k \leq k)}$. Let $T_n = \sum_{k=1}^n Y_k$. Define $Z_k = X_k - Y_k$. Note that

$$\begin{aligned}\sum_{k=1}^\infty P(Z_k > 0) &= \sum_{k=1}^\infty P(X > k) \\ &\leq \int_0^\infty P(X > t)dt = \int_0^\infty \int 1_{X>t} dP dt \\ &= \int \int_0^X 1_{X>t} dt dP = \int X dP = \mu < \infty.\end{aligned}$$

Hence (by Borel-Cantelli) $P(Z_k > 0 \ i.o.) = 0$. Let $A = \{\omega : Z_k(w) = 0 \text{ eventually}\}$. Thus $P(A) = 1$.

Let $A_m = \{\omega : Z_k(\omega) = 0 \ \forall k \geq m\}$. We know that $A_m \uparrow A$. For all $\omega \in A_m$ we have $\frac{\sum_{k=1}^n Z_k}{n} = \frac{\sum_{k=1}^{\min\{n,m\}} Z_k}{n}$. Thus

$$0 \leq \liminf_n \frac{\sum_{k=1}^n Z_k}{n} \leq \limsup_n \frac{\sum_{k=1}^n Z_k}{n} \leq \limsup_n \frac{\sum_{k=1}^m Z_k}{n} = 0.$$

Taking union over m we obtain that

$$0 = \lim_n \frac{\sum_{k=1}^n Z_k}{n} = \lim_n \frac{S_n - T_n}{n}, \forall \omega \in A.$$

As $T_n \leq S_n$ implies that $\limsup_n \frac{E(T_n)}{n} \leq \mu$.

Note that $E(Z_n) = E(X 1_{X>n})) \to 0$, hence (by Cesaro-sum)

$$\lim_n \frac{1}{n} E(S_n - T_n) = 0.$$

Therefore the theorem is proved if we establish that

$$\lim_n \frac{T_n - E(T_n)}{n} \to 0 \ a.s. \tag{5.3}$$

Step 2: (Subsequence convergence). Let $\alpha > 1$ be arbitrary and let $k(n) = \lfloor \alpha^n \rfloor$. Then note that

$$\begin{aligned}\sum_{n=1}^{\infty} P(|T_{k(n)} - E(T_{k(n)})| \geq \epsilon k(n)) &\leq \sum_{n=1}^{\infty} \frac{\mathrm{var}(T_{k(n)})}{\epsilon^2 k(n)^2} \\ &= \sum_{n=1}^{\infty} \frac{1}{\epsilon^2 k(n)^2} \sum_{m=1}^{k(n)} \mathrm{var}(Y_m) \\ &= \sum_{m=1}^{\infty} \mathrm{var}(Y_m) \left(\sum_{n:k(n)\geq m} \frac{1}{\epsilon^2 \lfloor \alpha^n \rfloor^2} \right) \\ &\leq \sum_{m=1}^{\infty} \frac{\mathrm{var}(Y_m)}{\epsilon^2 m^2} \left(\sum_{n=0}^{\infty} \frac{4}{\alpha^{2n}} \right) \\ &= 4 \frac{\alpha^2}{\epsilon^2(\alpha^2 - 1)} \sum_{m=1}^{\infty} \frac{\mathrm{var}(Y_m)}{m^2} \\ &\leq 16 \frac{\alpha^2}{\epsilon^2(\alpha^2-1)} E(X) < \infty.\end{aligned}$$

The first inequality follows from Chebychev's inequality. To obtain the second inequality let n_0 be the smallest n such that $k(n) \geq m$. Then note that $\forall n \geq n_0$ we have

$$\lfloor \alpha^n \rfloor = \lfloor \alpha^{n_0} \alpha^{n-n_0} \rfloor \geq \lfloor \alpha^{n_0} \rfloor \lfloor \alpha^{n-n_0} \rfloor \geq m \frac{\alpha^{n-n_0}}{2}.$$

The third inequality follows from Lemma 5.7 below.

Hence (by Borel-Cantelli)

$$\lim_n \frac{T_{k(n)} - E(T_{k(n)})}{k(n)} = 0 \ a.s. \implies \lim_n \frac{T_{k(n)}}{k(n)} = \mu \ a.s.$$

since $\frac{E(T_{k(n)})}{k(n)} \to \mu$.

Step 3: (Sandwich) Note that

$$\frac{\alpha^{n+1}-1}{\alpha^n} \leq \frac{k(n+1)}{k(n)} \leq \frac{\alpha^{n+1}}{\alpha^n - 1}, \implies \lim_n \frac{k(n+1)}{k(n)} = \alpha.$$

For every m let n_m be such that $k(n_m) \leq m < k(n_m+1)$. We know that

$$\frac{k(n_m)}{k(n_m+1)} \frac{T_{k(n_m)}}{k(n_m)} \leq \frac{T_m}{m} \leq \frac{k(n_m+1)}{k(n_m)} \frac{T_{k(n_m+1)}}{k(n_m+1)}.$$

Thus on the set where $\lim_n \frac{T_{k(n)}}{k(n)} = \mu$ we have

$$\frac{\mu}{\alpha} \leq \liminf_m \frac{T_m}{m} \leq \limsup_m \frac{T_m}{m} \leq \alpha\mu.$$

Since $\alpha > 1$ is arbitrary, we have $\lim_m \frac{T_m}{m} = \mu \ a.s.$ $\square$

Lemma 5.7. *Let $X_1, \ldots, X_n$ be pairwise independent and identically distributed non-negative random variables (say same as X) with $E(X) = \mu < \infty$. Let $Y_k =$*

$X_k 1_{(X_k \le k)}$. *Then*

$$\sum_{k=1}^{\infty} \frac{\text{var}(Y_k)}{k^2} \le 4E(X).$$

Proof. Recall that we have $Y_k \ge 0$. Note that

$$\begin{aligned}
\sum_{k=1}^{\infty} \frac{\text{var}(Y_k)}{k^2} &\le \sum_{k=1}^{\infty} \frac{E(Y_k^2)}{k^2} \\
&= \sum_{k=1}^{\infty} \frac{\int \int_0^\infty 2t 1_{Y_k > t} dt dP}{k^2} \\
&\overset{(a)}{=} \sum_{k=1}^{\infty} \frac{\int_0^\infty 2t P(Y_k > t) dt}{k^2} \\
&\le \sum_{k=1}^{\infty} \frac{\int_0^k 2t P(X > t) dt}{k^2} \\
&\overset{(b)}{=} \int_0^\infty 2t P(X > t) \left(\sum_{k: k \ge \max\{t,1\}} \frac{1}{k^2} \right) dt \\
&\le \int_0^\infty P(X > t) \left(2t \frac{\pi^2}{6} 1_{t \le 1} + 2t \left(\frac{\pi^2}{6} - 1 \right) 1_{1 < t \le 2} + \frac{2t}{t-1} 1_{t > 2} \right) dt \\
&\le \int_0^\infty 4 P(X > t) dt \\
&= 4E(X).
\end{aligned}$$

Here $(a), (b)$ use Fubini, the second inequality uses that $X_k \ge Y_k$ implying $P(Y_k > t) \le P(X \ge t)$, and the third inequality is just a case decomposition of t into three intervals. □

Theorem 5.8. *(Kolmogorov's zero-one law) Suppose $X_1, ..., X_n, ...$ are independent random variables. Let $\mathcal{F}_n = \sigma(X_n, ...)$ and define $\mathcal{F}_\infty = \underset{n}{\cap} \mathcal{F}_n$. Then*

$$\forall A \in \mathcal{F}_\infty, P(A) \in \{0, 1\}.$$

Proof. We will see that $A \in \mathcal{F}_\infty$ is independent of itself. W.l.o.g. let $P(A) > 0$.

Given any $n \ge 1$, we know that $A \in \mathcal{F}_{n+1}$. Since $A \in \mathcal{F}^\infty \subset \mathcal{F}^{n+1}$ it is independent of $\mathcal{F}'_n := \sigma(X_1, ..., X_n)$. Therefore A is also independent of sets in the field $\mathcal{F}' = \cup_n \mathcal{F}'_n$.

Consider the class of sets, $\mathcal{A}$, that are independent of A. Define two countably probability additive probability measures on $\sigma(\mathcal{F}')$, one according to $(i) \frac{P(A \cap B)}{P(A)}$ and the other according to $(ii) P(B)$. Note that these two probability measures agree on the field $\mathcal{F}'$; hence must agree on $\sigma(\mathcal{F}')$.

Note that $\mathcal{F}_n$ is generated by finite dimensional cylinder sets; but each of these finite dimensional cylinder sets will be in some $\mathcal{F}'_n$. Hence $\mathcal{F}_n \subset \sigma(\mathcal{F}')$. Thus

We have

$$\underset{n}{\cap} \mathcal{F}_n = \mathcal{F}_\infty \subset \sigma(\mathcal{F}').$$

and since $A \in \mathcal{F}^\infty$, A is independent of itself hence $P(A) = P^2(A)$. □

6. Central limit theorem

6.1. Some useful lemmas regarding real and complex numbers.

Lemma 6.1.
$$\left|e^{-z}-1+z\right| \leq |z|^2,\ 0 \leq |z| \leq 1.$$

Proof. From Taylor expansion, suffices to show that
$$\sum_{n=2}^{\infty} \frac{|z|^n}{n!} \leq |z|^2.$$
This follows immediately since $\frac{|z|^n}{n!} \leq \frac{|z|^2}{2^{n-1}}$ for $|z| \leq 1, n \geq 2$. □

Lemma 6.2.
$$\left|e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!}\right| \leq \min\left\{\frac{|x|^{n+1}}{(n+1)!}, \frac{2|x|^n}{n!}\right\}.$$

Proof. Using integration by parts, verify that
$$e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!} = \frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} ds.$$
Therefore, we will bound the right-hand-side in two ways. Observe that
$$\left|e^{ix} - \sum_{m=0}^{n} \frac{(ix)^m}{m!}\right| = \left|\frac{i^{n+1}}{n!}\int_0^x (x-s)^n e^{is} ds\right| \leq \left|\frac{1}{n!}\int_0^{|x|} s^n ds\right| = \frac{|x|^{n+1}}{(n+1)!}.$$
On the other hand applying the same bound to $n-1$ yields
$$\left|e^{ix} - \sum_{m=0}^{n-1} \frac{(ix)^m}{m!}\right| \leq \frac{|x|^n}{n!}.$$
From triangle inequality, the second upper bound follows. □

Lemma 6.3. *Let $z_1, .., z_n$ and $\omega_1, .., \omega_n$ be complex numbers whose absolute value is bounded by θ. Then*
$$\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| \leq \theta^{n-1}\sum_{m=1}^{n}|z_m - \omega_m|.$$

Proof. Note that
$$\begin{aligned}\left|\prod_{m=1}^{n} z_m - \prod_{m=1}^{n} w_m\right| &\leq \left|z_1\left(\prod_{m=2}^{n} z_m - \prod_{m=2}^{n} w_m\right)\right| + \left|(z_1-\omega_1)\prod_{m=2}^{n} w_m\right| \\ &\leq \theta\left|\prod_{m=2}^{n} z_m - \prod_{m=2}^{n} w_m\right| + \theta^{n-1}|z_1-\omega_1|.\end{aligned}$$
Bounding the first term (induction) finishes the proof.

□

Lemma 6.4. *Let X be a random variable such that $E[X] = 0$, $E[X^2] = \sigma^2$. Then*
$$\lim_{n\to\infty} n\left(\phi_X\left(\frac{t}{\sqrt{n}}\right) - 1 + \frac{E[X^2]}{2}\frac{t^2}{n}\right) = 0.$$

Proof.

$$n\left(\phi_X\left(\frac{t}{\sqrt{n}}\right)-1+\frac{E[X^2]}{2}\frac{t^2}{n}\right)$$
$$= nE\left(e^{iX\frac{t}{\sqrt{n}}}-1-iX\frac{t}{\sqrt{n}}+\frac{X^2}{2}\frac{t^2}{n}\right).$$

Let us split the integral into two parts $|X|>\frac{\epsilon\sqrt{n}}{2|t|}$ and $|X|\leq\frac{\epsilon\sqrt{n}}{2|t|}$. Consider

$$E\left(\left|n(e^{iX\frac{t}{\sqrt{n}}}-1)-iX\sqrt{n}t+\frac{X^2}{2}t^2\right|1_{|X|>\frac{\epsilon\sqrt{n}}{2|t|}}\right)$$
$$\leq 2nP\left(|X|>\frac{\epsilon\sqrt{n}}{2|t|}\right)+|t|\sqrt{n}E\left(|X|1_{|X|>\frac{\epsilon\sqrt{n}}{2|t|}}\right)+\frac{t^2}{2}E\left(|X|^2 1_{|X|>\frac{\epsilon\sqrt{n}}{2|t|}}\right).$$

Each of the term goes to 0 as n goes to infinity. The last term clearly goes to zero (by dominated convergence theorem); and the first two terms are upper bounded by a constant times last term. For instance,

$$n1_{|X|>\frac{\epsilon\sqrt{n}}{2|t|}}\leq 4\frac{t^2}{\epsilon^2}|X|^2 1_{|X|>\frac{\sqrt{n}}{2|t|}}.$$

Now consider

$$E\left(\left|n(e^{iX\frac{t}{\sqrt{n}}}-1)-iX\sqrt{n}t+\frac{X^2}{2}t^2\right|1_{|X|\leq\frac{\epsilon\sqrt{n}}{2|t|}}\right).$$

From Lemma 6.2 we have $|e^{iz}-1-iz+\frac{z^2}{2}|\leq\frac{|z|^3}{6},|z|^2$. Hence (for $\epsilon<1$)

$$\left|n(e^{iX\frac{t}{\sqrt{n}}}-1)-iX\sqrt{n}t+\frac{X^2}{2}t^2\right|1_{|X|\leq\frac{\epsilon\sqrt{n}}{2|t|}}\leq\frac{n}{6}\left(\frac{|Xt|}{\sqrt{n}}\right)^3 1_{|X|\leq\frac{\epsilon\sqrt{n}}{2|t|}}\leq|Xt|^2.$$

Hence by dominated convergence theorem, the expectation goes to zero. □

Theorem 6.5. *(Central Limit Thoerem, Vanilla Version) if $X_1,...,X_n$ are i.i.d random variables with $E[X_i]=0$, $E[X_i^2]=\sigma^2$. Then*

$$\frac{X_1+...+X_n}{\sqrt{n}}\overset{w}{\Rightarrow}\mathcal{N}(0,\sigma^2).$$

Proof.

$$E[e^{it\frac{X_1+...+X_n}{\sqrt{n}}}]=\prod_{i=1}^n E[e^{it\frac{X_i}{\sqrt{n}}}]$$
$$=\phi_X(\frac{t}{\sqrt{n}})^n.$$

Using previous lemma we obtain

$$\phi_X(\frac{t}{\sqrt{n}})^n\to e^{-\sigma^2t^2/2}$$

which is the characteristic function of $\mathcal{N}(0,\sigma^2)$. □

Theorem 6.6. *(Lindberg-Feller Theorem) For each n, let $X_{n,m}, 1\leq m\leq n$, be independent random variables with $E(X_{n,m})=0$. Suppose*

(1) $\lim_n\sum_{m=1}^n E(X_{n,m}^2)=\sigma^2>0$
(2) For all $\epsilon>0$, $\lim_n\sum_{m=1}^n E\left(X_{n,m}^2 1_{|X_{n,m}|>\epsilon}\right)=0.$

Then $S_n = X_{n,1} + \cdots + X_{n,n} \overset{w}{\Rightarrow} \mathcal{N}(0, \sigma^2)$.

Proof. Let $\phi_{n,m}(t) = E\left(e^{itX_{n,m}}\right)$ and $\sigma_{n,m}^2 = E(X_{n,m}^2)$. Suffices (by Levy's continuity theorem) to show that

$$\lim_{n\to\infty} \prod_{m=1}^{n} \phi_{n,m}(t) = e^{-t^2\sigma^2/2}.$$

Let $a_n = \sup_{1\le m\le n} \sigma_{n,m}^2$. Note that

$$a_n = sup_{1\le m\le n} E(X_{n,m}^2) \le \epsilon^2 + \sum_{m=1}^{n} E\left(X_{n,m}^2 1_{|X_{n,m}|>\epsilon}\right).$$

Therefore $\limsup_n a_n \le \epsilon^2$, but since $\epsilon > 0$ is arbitrary $a_n \to 0$.

Hence for any $t > 0$, there exists n_0 large enough so that for all $n \ge n_0$, $|1 - \frac{t^2 a_n}{2}| \le 1$. Applying Lemma 6.3 (with $\theta = 1$, $n \ge n_0$) we obtain

$$\begin{aligned}
&\left|\prod_{m=1}^{n} \phi_{n,m}(t) - \prod_{m=1}^{n} (1 - \frac{t^2\sigma_{n,m}^2}{2})\right| \\
&\le \sum_{m=1}^{n} |\phi_{n,m}(t) - (1 - \frac{t^2\sigma_{n,m}^2}{2})| \\
&= \sum_{m=1}^{n} \left| E\left(e^{itX_{n,m}} - 1 - itX_{n,m} + \frac{t^2 X_{n,m}^2}{2}\right)\right| \\
&\overset{(a)}{\le} \sum_{m=1}^{n} E\left(\min\left\{\frac{|tX_{n,m}|^3}{6}, |tX_{n,m}|^2\right\}\right) \\
&\le \sum_{m=1}^{n} E\left(\frac{|t|\epsilon}{6}|tX_{n,m}|^2 1_{|X_{n,m}|\le\epsilon} + |tX_{n,m}|^2 1_{|X_{n,m}|>\epsilon}\right).
\end{aligned}$$

In the above (a) follows from Lemma 6.2. Hence

$$\limsup_n \left|\prod_{m=1}^{n} \phi_{n,m}(t) - \prod_{m=1}^{n}\left(1 - \frac{t^2\sigma_{n,m}^2}{2}\right)\right| \le \epsilon\frac{|t|^3\sigma^2}{6},$$

however since $\epsilon > 0$ is arbitrary, the limit is zero.

To complete the proof, it suffices to show that

$$\limsup_n \left|\prod_{m=1}^{n}\left(1 - \frac{t^2\sigma_{n,m}^2}{2}\right) - \prod_{m=1}^{n} e^{-t^2\sigma_{n,m}^2/2}\right| = 0.$$

Again applying Lemma 6.3 (with $\theta = 1$, $n \ge n_0$) we obtain

$$\begin{aligned}
&\left|\prod_{m=1}^{n}\left(1 - \frac{t^2\sigma_{n,m}^2}{2}\right) - \prod_{m=1}^{n} e^{-t^2\sigma_{n,m}^2/2}\right| \\
&\le \sum_{m=1}^{n} \left|e^{-t^2\sigma_{n,m}^2/2} - \left(1 - \frac{t^2\sigma_{n,m}^2}{2}\right)\right| \\
&\overset{(a)}{\le} \sum_{m=1}^{n} \frac{t^4\sigma_{n,m}^4}{2} \le \frac{t^4 a_n}{2}\sum_{m=1}^{n}\sigma_{n,m}^2.
\end{aligned}$$

In the above (a) follows from Lemma 6.1. Since $a_n \to 0$ and $\sum_{m=1}^n \sigma_{n,m}^2 \to \sigma^2$, we are done.

$\square$

Lemma 6.7 (Borel-Cantelli 2)**.** *If an infinite sequence of mutually independent events A_i satisfy $\sum_i P(A_i) = \infty$, then $P(A_n$ i.o.$) = 1$.*

Proof. Define

$$B_k = \cup_{m \geq k} A_m.$$

Then, by independence,

$$P(B_k) = 1 - \prod_{m \geq k} (1 - P(A_m)) \geq 1 - \prod_{m \geq k} e^{-P(A_m)} = 1.$$

where the inequality follows since $(1-x) \leq e^{-x}, 0 \leq x \leq 1$. Now $B_k \downarrow A_n$ *i.o.* and we are done. $\square$

6.2. **Convergence to a Poisson random variable.** A Poisson(λ) random variable, Z, takes values in $\mathbb{N}$ and satisfies $P(Z = n) = \frac{\lambda^n e^{-\lambda}}{n!}$.

Theorem 6.8. *Let $X_{n,m}, 1 \leq m \leq n$ be independent non-negative integer valued random variables with $P(X_{n,m} = 1) = p_{n,m}$, and $P(X_{n,m} \geq 2) = \epsilon_{n,m}$. If*

(1) $\sum_{n=1}^m p_{n,m} \to \lambda \in (0, \infty)$
(2) $\max_{1 \leq m \leq n} p_{n,m} \to 0$, and
(3) $\sum_{m=1}^n \epsilon_{n,m} \to 0$,

then $S_n := \sum_{m=1}^n X_{n,m} \overset{w}{\Rightarrow} Z$, where Z is Poisson(λ).

Proof. Note that $E\left(e^{itS_n}\right) = \prod_{m=1}^n E\left(e^{itX_{n,m}}\right)$. From Lemma 6.3 (with $\theta = 1$) we have

$$\begin{aligned}
&\left| \prod_{m=1}^n E\left(e^{itX_{n,m}}\right) - \prod_{m=1}^n \left(1 - p_{n,m} + p_{n,m} e^{it}\right) \right) \\
&\quad \leq \sum_{m=1}^n \left| E\left(e^{itX_{n,m}}\right) - \left(1 - p_{n,m} + p_{n,m} e^{it}\right) \right| \\
&\quad = \sum_{m=1}^n \left| E\left(-\epsilon_{n,m} + E\left(e^{itX_{n,m}} 1_{X_{n,m} \geq 2}\right)\right) \right| \leq 2 \sum_{m=1}^n \epsilon_{n,m} \to 0.
\end{aligned}$$

From Lemma 6.1 we know that as long as $p_{n,m} \leq \frac{1}{2}$

$$\left| e^{-p_{n,m}(1-e^{it})} - \left(1 - p_{n,m} + p_{n,m} e^{it}\right) \right| \leq 4 p_{n,m}^2.$$

Again, from Lemma 6.3 (with $\theta = 1$) we have

$$\begin{aligned}
&\left| \prod_{m=1}^n e^{-p_{n,m}(1-e^{it})} - \prod_{m=1}^n \left(1 - p_{n,m} + p_{n,m} e^{it}\right) \right) \\
&\quad \leq \sum_{m=1}^n \left| e^{-p_{n,m}(1-e^{it})} - \left(1 - p_{n,m} + p_{n,m} e^{it}\right) \right| \\
&\quad \leq 4 \sum_{m=1}^n p_{n,m}^2 \leq 4 \left(\max_{1 \leq m \leq n} p_{n,m} \right) \sum_{m=1}^n p_{n,m} \to 0.
\end{aligned}$$

Finally as

$$\left| e^{-\sum_{m=1}^{n} p_{n,m}(1-e^{it})} - e^{-\lambda(1-e^{it})} \right| \to 0,$$

we are done. $\square$