Probability

Donghyun Park

December 11, 2025

This post illustrates fundamental results on Durrett's Probability book [Dur19]

Chapter 2. Law of Large Numbers

In this chapter the most important technique is **to convert in probability convergence to almost surely convergence**. I shall build up some major theorems.

Theorem 1 (Theorem 2.1.21 (Kolmogorov's Extension Theorem)). *Given probability measures μ_n on $(\mathbb{R}^n, \mathcal{R}^n)$ that are consistent*

$$\mu_{n+1}((a_1, b_1] \times \cdots \times (a_n, b_n] \times \mathbb{R}) = \mu_n((a_1, b_1] \times \cdots \times (a_n, b_n])$$

Then there is a unique probability measure on $(\mathbb{R}^{\mathbb{N}}, \mathcal{R}^{\mathbb{N}})$ such that

$$P(w : w_i \in (a_i, b_i], 1 \leq i \leq n) = \mu_n((a_1, b_1] \times \cdots \times (a_n, b_n])$$

This theorem is important to construct i.i.d variables. Here, $(\mathbb{R}^n, \mathcal{R}^n)$ plays important role, for infinite product of different spaces, such as $\Omega_1 \times \Omega_2 \times \cdots$ Kolmogorov's Extension theorem version does not hold.
- Convergence **in probability** : Y_n converges to Y in probability if for all $\epsilon > 0$, $P(|Y_n - Y| > \epsilon) \to 0$

Basic Weak Law of Large Number theorem is:

Theorem 2 (Theorem 2.2.3). *$X_1, X_2, \cdots$ uncorrelated random variables with $\mathbb{E}X_i = \mu$ and $var(X_i) \leq C < \infty$. If $S_n = X_1 + \cdots + X_n$ then $S_n/n \to \mu$ in L^2 and in probability.*

Now, one of the most important technique follows.

Truncation

By truncating large values, we can prove **'finite version' of theorem and then send them to infinity** and complete the argument.

Theorem 3 (Theorem 2.2.11). *For each n, $X_{n,k}$ are independent $1 \leq k \leq n$. Let $b_n > 0$ with $b_n \to \infty$ and $\bar{X}_{n,k} = X_{n,k}1_{(|X_{n,k}| \leq b_n)}$. If*
(1) $\sum_{k=1}^{n} P(|X_{n,k}| > b_n) \to 0$
(2) $b_n^{-2} \sum_{k=1}^{n} \mathbb{E}\bar{X}_{n,k}^2 \to 0$
If $S_n = X_{n,1} + \cdots + X_{n,n}$ and $a_n = \sum_{k=1}^{n} \mathbb{E}\bar{X}_{n,k}$ then

$$(S_n - a_n)/b_n \to 0$$

in probability.

This means, if **truncated parts are not large enough (1)** and **Variance compared to some b_n is small enough (2)** then untruncated version also holds.

Theorem 4 (Theorem 2.2.12 (Weak Law of Large Numbers)). *Let $X_1, X_2, \cdots$ be i.i.d. with*

$$xP(|X_i| > x) \to 0$$

as $x \to \infty$. Let $S_n = X_1 + \cdots + X_n$ and $\mu_n = \mathbb{E}(X_1 1_{(|X_1| \leq n)})$. Then $S_n/n - \mu_n \to 0$ in probability.

If $\mathbb{E}|X_i| < \infty$ then for $\mu = \mathbb{E}X_1$ by above theorem $S_n/n \to \mu$ in probability holds.

Borel-Cantelli Lemma

Borel-Cantelli Lemma is the art of probability measure.

Theorem 5 (Theorem 2.3.1). *If $\sum_{n=1}^{\infty} P(A_n) < \infty$ then*

$$P(A_n \ i.o.) = 0$$

Borel Cantelli Lemma is the key converting **in probability argument to almost surely convergence**.

Theorem 6 (Theorem 2.3.2). *$X_n \to X$ in probability if and only if for every subsequence $X_{n(m)}$ there is a further subsequence $X_{n(m_k)}$ that converges almost surely to X.*

proof constructs sequence of positive numbers ϵ_k so that

$$\sum_{k=1}^{\infty} P(|X_{n(m_k)} - X| > \epsilon_k) < \infty$$

By upgrading convergence in probability to convergence almost surely, now we can prove something unnatural to prove when we have only convergence in probability.

Theorem 7 (Theorem 2.3.4). *If f is continuous and $X_n \to X$ in probability then $f(X_n) \to f(X)$ in probability. If f is bounded then $\mathbb{E}(f(X_n)) \to \mathbb{E}(f(X))$*

The second Borel-Cantelli lemma is opposite argument.

Theorem 8 (Theorem 2.3.7 (Second Borel-Cantelli lemma)). *If A_n are independent and $\sum P(A_n) = \infty$ then $P(A_n \ i.o.) = 1$*

I want to finally note **this technique** can be applied to other problem.

Theorem 9 (Theorem 2.3.9). *If $A_1, A_2, \cdots$ are pairwise independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$ then*

$$\frac{\sum_{m=1}^{n} 1_{A_m}}{\sum_{m=1}^{n} P(A_m)} \to 1 \quad a.s.$$

Proof. $X_m = 1_{A_m}$ and $S_n = X_1 + \cdots + X_n$. The Chebyshev inequality gives

$$P(|S_n - \mathbb{E}S_n| > \delta \mathbb{E}S_n) \leq \frac{1}{\delta^2 \mathbb{E}S_n} \to 0$$

so $S_n/\mathbb{E}S_n \to 1$ in probability. Now we would change into almost sure convergence.

$$n_k = \inf\{n : \mathbb{E}S_n \geq k^2\}$$

and $T_k = S_{n_k}$. Then using Borel-Cantelli,

$$P(|T_k/\mathbb{E}T_k - 1| > \delta \ i.o.) = 0$$

so $T_k/\mathbb{E}T_k \to 1$ almost surely, and then for such events, using monotonicity

$$\frac{T_k(\omega)}{\mathbb{E}T_{k+1}} \leq \frac{S_n(\omega)}{\mathbb{E}S_n} \leq \frac{T_{k+1}(\omega)}{\mathbb{E}T_k}$$

so we can interpolate intermediate terms. $\qquad\square$

Strong Law of Large Numbers

Strong Law of Large Number uses **Truncation** and **Borel Cantelli lemma**, as a result induces powerful consequence.

Theorem 2.4.1 (Strong Law of Large Numbers) $X_1, X_2, \cdots$ pairwise independent identically distributed r.v. with $\mathbb{E}|X_i| < \infty$. Let $\mathbb{E}X_i = \mu$ and $S_n = X_1 + \cdots + X_n$ then $S_n/n \to \mu$ a.s. as $n \to \infty$

Theorem 10 (Theorem 2.4.5). *Let $X_1, X_2, \cdots$ i.i.d with $\mathbb{E}X_i^+ = \infty$ and $\mathbb{E}X_i^- < \infty$. If $S_n = X_1 + \cdots + X_n$ then $S_n/n \to \infty$ a.s.*

Theorem 11 (Theorem 2.4.9 (Glivenko-Cantelli Theorem)). *Let F be the distribution of X_i and $X_1, X_2, \cdots$ i.i.d. If we let $F_n(x) = \frac{1}{n} \sum_{m=1}^{n} 1_{(X_m \leq x)}$ then*

$$\sup_x |F_n(x) - F(x)| \to 0 \ a.s.$$

Convergence of Random Series

Here is another approach. Previous section used Borel Cantelli and Truncation but here, **Tail Algebra** and **Kolmogorov's maximal inequality** plays the role.

Tail Algebra $\mathcal{T} = \bigcap_n \sigma(X_n, X_{n+1}, \cdots)$ is always probability 0 or 1 (**Kolmogorov's 0-1 law**)

Theorem 12 (Theorem 2.5.5 (Kolmogorov's maximal inequality)). *Suppose $X_1, \cdots, X_n$ are independent with $\mathbb{E}X_i = 0$ and $var(X_i) < \infty$. If $S_n = X_1 + \cdots + X_n$ then*

$$P(\max_{1 \leq k \leq n} |S_k| \geq x) \leq x^{-2} var(S_n)$$

As a corollary, the random series converges with probability one for following situation:

Theorem 13 (Theorem 2.5.6). *$X_1, X_2, \cdots$ independent and have $\mathbb{E}X_n = 0$. If*

$$\sum_{n=1}^{\infty} var(X_n) < \infty$$

Then with probability one, $\sum_{n=1}^{\infty} X_n(\omega)$ converges.

For truncated version, here is Kolmogorov's three-series theorem,

Theorem 14 (Theorem 2.5.8). *$X_1, X_2, \cdots$ be independent. Let $A > 0$ and let*

$$Y_i = X_i 1_{(|X_i| \leq A)}$$

. *Then $\sum_{n=1}^{\infty} X_n$ converges a.s. if and only if three holds.*
(1) $\sum_{n=1}^{\infty} P(|X_n| > A) < \infty$
(2) $\sum_{n=1}^{\infty} \mathbb{E}Y_n$ converges
(3) $\sum_{n=1}^{\infty} var(Y_n) < \infty$

I shall prove Exercise 2.5.10 of Durrett Probability. Here is the statement.

Theorem 15 (Theorem (P.Levy)). *Let $X_1, X_2, \cdots$ be independent and $S_n = X_1 + \cdots + X_n$. If $\lim_{n \to \infty} S_n$ exists in probability then it also exists a.s.*

This lemma is essential.

Lemma 1. *$X_1, X_2, \cdots$ independent and $S_{m,n} = X_{m+1} + \cdots + X_n$ then*

$$P(\max_{m < j \leq n} |S_{m,j}| > 2a) \min_{m < k \leq n} P(|S_{k,n}| \leq a) \leq P(|S_{m,n}| > a)$$

Proof.

$$A_k = \{\omega : \max_{m < j \leq n} |S_{m,j}| = |S_{m,k}| > 2a, \ |S_{m,m+1}|, \cdots, |S_{m,k-1}| \leq 2a\}$$

then

$$\cup_{k=m+1}^{n} A_k \cap \{\omega : |S_{k,n}| \leq a\} \subset \{|S_{m,n}| > a\}$$

Each event A_k and $\{\omega : |S_{k,n}| \leq a\}$ are independent so Lemma holds.

Now proving the theorem, since $\lim_{n \to \infty} S_n$ exists in probability,

$$\lim_{n \to \infty} P(|S_n - S| > \epsilon) = 0$$

so $\lim_{n \to} \sup_{j \geq 1} P(|S_{n+j} - S_n| \geq \epsilon) = 0$

$$P(\max_{1 \leq j \leq k} |S_{n+j} - S_n| \geq \epsilon) \leq P(|S_{n+k} - S_n| \geq \epsilon/3)$$

$$+ \sum_{j=1}^{k-1} P((|S_{n+k} - S_n| < \epsilon/3) \cap (|S_{n+j} - S_n| \geq \epsilon, \ |S_{n+i} - S_n| < \epsilon \ \forall i < j))$$

$$\leq P(|S_{n+k} - S_n| \geq \epsilon/3)$$

$$+ \sum_{j=1}^{k-1} P((|S_{n+j} - S_n| \geq \epsilon, \ |S_{n+i} - S_n| < \epsilon \forall i < j)) P(|S_{n+j} - S_{n+k}| > \frac{2}{3}\epsilon)$$

$$\leq P(|S_{n+k} - S_n| \geq \epsilon/3) + \max_{1 \leq j \leq k-1} P(|S_{n+k} - S_{n+j}| \geq \frac{2}{3}\epsilon)$$

$$\leq 3 \max_{0 \leq j \leq k-1} P(|S_{n+j} - S_{n+k}| \geq \epsilon/3) \to 0$$

$\square$

Chapter 3. Central Limit Theorems

We move on our interest into **converge in distributions (or Weak convergence)**. In this perspective, the **characteristic function** is important object.

Weak Convergence

Weak convergence is defined: for sequence of distribution functions F_n converges weakly to a limit F if $F_n(y) \to F(y)$ for all y a continuity points of F.

There are some equivalent statements of weak convergence.
- $X_n \Rightarrow X_\infty$ iff every bounded continuous function g, $\mathbb{E}g(X_n) \to \mathbb{E}g(X_\infty)$
- $X_n \Rightarrow X_\infty$ iff $\forall G$ open,

$$\liminf_{n \to \infty} P(X_n \in G) \geq P(X_\infty \in G)$$

- $X_n \Rightarrow X_\infty$ iff $\forall K$ closed,

$$\limsup_{n \to \infty} P(X_n \in K) \leq P(X_\infty \in K)$$

- $X_n \Rightarrow X_\infty$ iff $\forall A$ Borel set with $P(X_\infty \in \partial A) = 0$,

$$\lim_{n \to \infty} P(X_n \in A) = P(X_\infty \in A)$$

All these proofs use that we can construct Y_n with distribution F_n such that $Y_n \to Y_\infty$ almost surely, with Y_∞ has distribution F_∞.

Tightness

In the series of distribution functions, **tightness** condition is important to guarantee that the series converge into distribution again. If not, the general theorem "Helly's selection theorem" holds.

Theorem 16 (Theorem 3.2.12 (Helly's selection theorem)). *For every sequence F_n of distribution functions, there exists a subsequence $F_{n(k)}$ and right continuous nondecreasing F that*

$$\lim_{k \to \infty} F_{n(k)}(y) = F(y)$$

at continuity points of F.

To guarantee that this limit is again a distribution function,

Theorem 17 (Theorem 3.2.13). *Every subsequential limit is the distribution function if and only if F_n is* **tight**. *There exists M for each $\epsilon > 0$ such that*

$$\limsup_{n \to \infty} 1 - F_n(M_\epsilon) + F_n(-M_\epsilon) \leq \epsilon$$

Characteristic Functions

Characteristic Function is like a **key** of distribution functions. This function **Encodes** the distribution function and formulates problem into much easier object to handle.
First of all, the definition of the characteristic function is

$$\varphi(t) = \mathbb{E}e^{itX}$$

We can recover the distribution function from teh characteristic function by the inversion formula.

Theorem 18 (Theorem 3.3.11 (The Inversion Formula)). *If $a < b$ then*

$$\lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \frac{e^{-ita} - e^{-itb}}{it} \varphi(t)dt = \mu(a,b) + \frac{1}{2}\mu(\{a,b\})$$

Or, if φ is good :

Theorem 19 (Theorem 3.3.14). *If $\int |\varphi(t)|dt < \infty$, then measure μ has bounded continuous density*

$$f(y) = \frac{1}{2\pi} \int e^{-ity} \varphi(t)dt$$

Weak Convergence in Characteristic function

This theorem is the heart of Chapter 3. Weak convergence can be encoded as function limit

Theorem 20 (Theorem 3.3.17 (Continuity Theorem)). *Let μ_n probability measures with characteristic functions φ_n*
(1) If $\mu_n \Rightarrow \mu_\infty$, then $\varphi_n(t) \to \varphi_\infty(t)$ for all t.
*(2) If Characteristif functions $\varphi_n(t)$ converges pointwise to a limit $\varphi(t)$ (Do not know it is characteristic funciton) and $\varphi(t)$ is continuous at 0, then μ_n **is tight and converges weakly to a measure μ which has a characteristic funtion φ***

Proof. Using the identity

$$\int_{-u}^{u} 1 - e^{itx}dt = 2u - \frac{2sin(ux)}{x}$$

$$\frac{1}{u}\int_{-u}^{u}(1 - \varphi_n(t))dt = 2\int\left(1 - \frac{sin(ux)}{ux}\right)\mu_n(dx)$$

Left hand side converges as $n \to \infty$ to $\frac{1}{u}\int_{-u}^{u} 1 - \varphi(t)dt$. Right hand side, when integrated on $|x| \geq 2/u$ is bigger than

$$\mu_n(\{x : |x| > 2/u\})$$

$$\frac{1}{u}\int_{-u}^{u}(1 - \varphi_n(t))dt \geq \mu_n\{x : |x| > 2/u\}$$

Now, continuity of φ at 0 is used to bound

$$\frac{1}{u}\int_{-u}^{u}(1 - \varphi(t))dt \to 0$$

So tightness can be shown. $\qquad\square$

Here in the proof, **smoothness of the characteristic function at 0 gives impact on the decay of distribution function**. This idea can be also found on Fourier analysis. For instance, f is C^k then $\hat{f}(n) = o(|n|^{-k})$.

Theorem 21 (Theorem 3.3.18). *If $\int |x|^n \mu(dx) < \infty$ then φ is C^n with*

$$\varphi^{(n)}(t) = \int (ix)^n e^{itx}\mu(dx)$$

Polya's Criteria

Special case of Ch.f. is when the Density function is

$$f(x) = \frac{1 - cos(x)}{\pi x^2}$$

$$\varphi(t) = (1 - |t|)^+$$

Using this distribution we can verify **when does the given function $\varphi(t)$ becomes Characteristic function of some distribution**.

Theorem 22 (Theorem 3.3.22. (Polya's criterion)). *If $\varphi(t)$ real nonnegative, $\varphi(0) = 1$, $\varphi(-t) = \varphi(t)$ and φ decreasing convex on $(0, \infty)$.*

$$\lim_{t\downarrow 0}\varphi(t) = 1$$

$$\lim_{t\uparrow\infty}\varphi(t) = 0$$

Then there is a probability measure on $(0, \infty)$ so that

$$\varphi(t) = \int_{0}^{\infty}\left(1 - \left|\frac{t}{s}\right|\right)^+ \mu(ds)$$

Thus, as a continuous sum (or integral) of Ch.f.s in above formula φ is characteristic function.
A big application of the Polya's criterion is, for the function

$$\varphi(t) = \exp(-|t|^\alpha)$$

It is a characteristic function if $0 < \alpha \le 2$ but not a characteristic function if $\alpha > 2$
Now, we shall seek on the application of characteristic functions.

Momentum Problem

If we have all the values of $\int x^k dF(x)$ $k = 1, 2, \cdots$. Then can we find F?
In general, the answer is no.

Counterexapmle 1. suggested by Heyde (1963) was lognormal density

$$f_0(x) = (2\pi)^{-1/2} x^{-1} \exp(-(\log x)^2/2)$$

Then $f_a(x) = f_0(x)(1 + a\sin(2\pi \log x))$ have the same moments for $-1 \le a \le 1$
In this case moments are $\mathbb{E}X^n = \exp(n^2/2)$.

Counterexample 2. For $\lambda \in (0, 1)$ and $-1 \le a \le 1$

$$f_{a,\lambda}(x) = c_\lambda \exp(-|x|^\lambda)\{1 + a\sin(\beta|x|^\lambda \text{sgn}(x))\}$$

$\beta = \tan(\lambda\pi/2)$ and c_λ normalizing constant, have the same moments. In this case, moments are approximately
size of $\mathbb{E}X^n \asymp \Gamma(\frac{n-\lambda-1}{\lambda})$
Here, $ln(\Gamma(z)) \sim z ln(z) - z$ so

$$\Gamma\left(\frac{n}{\lambda}\right)^{1/n} = \exp\left(\frac{1}{n}ln(\Gamma(\frac{n}{\lambda}))\right) \sim \exp\left(\frac{1}{\lambda}ln(n)\right) = n^{1/\lambda}$$

$$\sum_{k=1}^{\infty} \frac{1}{\mathbb{E}(X^{2k})^{1/2k}} \sim \sum_{k=1}^{\infty} \frac{1}{(2k)^{1/\lambda}} < \infty$$

The criteria is determined by this formula. Which is **Carleman's condition**.

Theorem 23. *If $\sum_{k=1}^{\infty} \frac{1}{\mu_{2k}^{1/2k}} = \infty$ then there is at most one distribution function F with $\mu_k = \int x^k dF(x)$.*

The Central Limit Theorem

Now we are ready to focus on the central limit theorem. By our previous works, central limit theorem is an easy consequence by study of characteristic functions.

Theorem 24 (Theorem 3.3.17). *is again very powerful theorem.*

Theorem 25 (Theorem 3.4.1). *$X_1, X_2, \cdots$ i.i.d with $\mathbb{E}X_i = \mu$ and $var(X_i) = \sigma^2 \in (0, \infty)$. If $S_n = X_1 + \cdots + X_n$ then*

$$(S_n - n\mu)/\sigma n^{1/2} \Rightarrow \chi$$

the standard normal distribution.

Proof. Assume $\mu = 0$. $\varphi_{X_i}(t) = 1 - \frac{\sigma^2 t^2}{2} + o(t^2)$ so $\varphi_{S_n/\sigma n^{1/2}}(t) = \left(1 - \frac{t^2}{2n} + o(n^{-1})\right)^n \to exp(-t^2/2)$. Which
is continuous at 0 and $exp(-t^2/2)$ is characteristic function of standard normal distribution. $\square$

As we did on the law of large numbers, we can **truncate** the series of random variables to obtain general results.

Theorem 26 (Theorem 3.4.10 (The Lindeberg-Feller theorem)). *$X_{n,m}$ for $1 \le m \le n$ are independent r.v.s
with $\mathbb{E}X_{n,m} = 0$. Suppose*
(1) $\sum_{m=1}^{n} \mathbb{E}X_{n,m}^2 \to \sigma^2 > 0$
(2) $\forall \epsilon > 0$, $\lim_{n\to\infty} \sum_{m=1}^{n} \mathbb{E}(|X_{n,m}|^2; |X_{n,m}| > \epsilon) = 0$
Then $S_n = X_{n,1} + \cdots + X_{n,n} \Rightarrow \sigma \cdot \chi$

Remark. $X_{n,m} = X_m/n$ deduces into C.L.T theorem

Proof. Again we are doing on the characteristic functions. $\varphi_{n,m}(t)$ be the characteristic function of $X_{n,m}$. By the accurate examination formula

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

$$|\varphi_{n,m}(t) - (1 - t^2 \sigma_{n,m}^2/2)| \leq \mathbb{E}(|tX_{n,m}|^3 \wedge 2|tX_{n,m}|^2) \leq \epsilon t^3 \mathbb{E} X_{n,m}^2 + 2t^2 \mathbb{E}(|X_{n,m}|^2; |X_{n,m}| > \epsilon)$$

Giving us

$$\limsup_{n \to \infty} \sum_{m=1}^{n} |\varphi_{n,m}(t) - (1 - t^2 \sigma_{n,m}^2/2)| \leq \epsilon t^3 \sigma^2$$

So

$$\prod_{m=1}^{n} \varphi_{n,m}(t) - exp(-t^2 \sigma^2/2) \to 0$$

$\square$

Using this truncation technique, we can handle infinite variance cases. See Example 3.4.13 of Durrett's Probability book.

So when does the truncation technique can be used on? Gnedenko and Kolmogorov found the necessary and sufficient condition.

Theorem 27 (Theorem 3.4.14). $X_1, X_2, \cdots$ be i.i.d. $S_n = X_1 + \cdots + X_n$. There exists $a_n, b_n > 0$ that $(S_n - a_n)/b_n \Rightarrow \chi$ if and only if as $y \to \infty$

$$\frac{y^2 P(|X_1| > y)}{\mathbb{E}(|X_1|^2; |X_1| \leq y)} \to 0$$

Local Limit Theorem

We can classify characteristic functions into three classes.
(1) $\varphi \equiv 1$: point mass
(2) $|\varphi(\lambda)| = 1$ for some λ but not all : Lattice distribution
(3) $|\varphi| < 1$
In the case of lattice distribution, we can argue on the all supports ; lattices that probability distribution of S_n converges to normal distribution.

Theorem 28 (Theorem 3.5.3). *If $X_1, X_2, \cdots$ i.i.d with $\mathbb{E}X_i = 0$, $\mathbb{E}X_i^2 = \sigma^2 \in (0, \infty)$ having a common lattice distribution $X_i \in b + h\mathbb{Z}$. Then for $p_n(x) = P(S_n/\sqrt{n} = x)$ the discrete probability whose support being*

$$\{(nb + hz)/\sqrt{n} : z \in \mathbb{Z}\}$$

converges uniformly to normal distribution probability density : $n(x) = \frac{1}{\sqrt{2\pi}\sigma} exp(-x^2/2\sigma^2)$

$$\sup_{x \in \{(nb+hz)/\sqrt{n}: z \in \mathbb{Z}\}} \left| \frac{\sqrt{n}}{h} p_n(x) - n(x) \right| \to 0$$

Actually, the same holds for non-lattice case.

Poisson Convergence and Poisson Process

On the other side of the limit theorem, there is a Poisson convergence.

Theorem 29 (Theorem 3.7.1). *$X_{n,m}$, $1 \leq m \leq n$ independent rvs with $P(X_{n,m} = 1) = p_{n,m}$ and $P(X_{n,m} = 0) = 1 - p_{n,m} - \epsilon_{n,m}$. If $\sum_{m=1}^{n} p_{n,m} \to \lambda$ and $\max_{1 \leq m \leq n} p_{n,m} \to 0$, $\sum_{m=1}^{n} \epsilon_{n,m} \to 0$ (Sufficient for convergence) then $S_n \Rightarrow Z$ the Poisson distribution of λ.*

Proof is again the characteristic functions.

Poisson process is the process such that $N(t)$, the $0 = t_0 < t_1 < \cdots < t_n$, $N(t_k) - N(t_{k-1})$ are independent with distribution Poisson($\lambda(t_k - t_{k-1})$)

Stable Laws

Answering the following question : when does a_n, b_n exist so that for $X_1, \cdots X_n$ i.i.d

$$\frac{S_n - b_n}{a_n} \Rightarrow Y$$

to the nondegenerate Y? CLT says if $var(X_1) < \infty$ then it does. Quite brief answer gives

Theorem 30 (Theorem 3.8.2). $X_1, \cdots$ *i.i.d that*
- $\lim_{x \to \infty} P(X_1 > x)/P(|X_1| > x) = \theta \in [0, 1]$
- $P(|X_1| > x) = x^{-\alpha} L(x), \ \alpha < 2$
- $\lim_{x \to \infty} L(tx)/L(x) = 1, \ \forall t > 0$
Then such a_n, b_n exists.

Stunning fact is that such distributions are written in

$$\exp(itc - b|t|^\alpha (1 + i\kappa \text{sgn}(t) w_{\alpha(t)}))$$

$0 < \alpha \le 2, \ -1 \le \kappa \le 1$ with $w_\alpha(t) = \tan(\pi\alpha/2)$ if $\alpha \ne 1$ and $(2/\pi) \log |t|$ if $\alpha = 1$.
and moreover, these distributions are only one who can be written as the limit $(S_n - a_n)/b_n$.

Chapter 4. Martingales

Martingales and Conditional expectation are powerful tools for analyzing process. In the Martingale's language, the **Stopping Time** becomes key to unlock some nontrivial theorems.

References

[Dur19] Rick Durrett. *Probability: Theory and Examples.* 5th. Cambridge University Press, 2019. ISBN: 978-1-108-47368-2.