Statistics 212: Lecture 19 (April 16, 2025)

Lévy's Characterization of Brownian Motion, Dubins-Schwarz Theorem

Instructor: Mark Sellke

Scribe: Baiyun Jing, Knut Vanderbush

Summary. In this lecture, we're going to use Lindeberg's exchange method to prove two CLTs. The first will be a usual CLT for a sequence of independence random variables, and the second will be a martingale CLT. A consequence of the martingale CLT will be Lévy's characterization of Brownian Motion, which states that if $X_t, X_t^2 - t$ are martingales then X_t is BM. A corollary of Lévy's characterization of Brownian Motion will be the Dubins-Schwartz Theorem, which states that any continuous martingale is a time-changed BM.

1 Central Limit Theorem with Lindeberg's Exchange Method

We first state the usual central limit theorem for independent random variables.

Theorem 1.1. *Let $X_1, ..., X_n$ be a sequence of independent, not necessarily identically distributed, random variables with $E[X_i] = 0, E[X_i^2] = 1$ and $E[|X_i|^3] \le C$ for some finite C. Then*

$$\frac{1}{\sqrt{n}}\big(X_1 + \cdots + X_n\big) \xrightarrow{d} N(0,1). \tag{1}$$

Proof. In this proof, we use Lindeberg's exchange method. The idea is to replace each X_k with a Gaussian that has the same first and second moments, and argue that each of the replacements doesn't change the distribution of the sum by much. Hence the original sum behaves approximately like the sum of Gaussians which is $N(0,1)$.

By the Portmanteau Theorem, convergence in distribution is equivalent to

$$\mathbb{E}\Big[f\Big(\frac{1}{\sqrt{n}}\big(X_1 + \cdots + X_n\big)\Big)\Big] \to \mathbb{E}[f(Z)], \quad Z \sim N(0,1) \tag{2}$$

for all uniformly continuous and bounded $f : R \to R$. In addition, recall that smooth and bounded functions are *dense* in the space of uniformly continuous and bounded functions, so in fact we only need to show Equation (2) for all smooth and bounded $f : R \to R$ with f', f'', f''' uniformly bounded.

Let $g_1, ..., g_n$ be iid $N(0,1)$ random variables, and consider

$$\sum_{k=0}^{n-1}\left|\mathbb{E}\left[f\left(\frac{1}{\sqrt{n}}(X_1 + \cdots + X_k + g_{k+1} + \cdots + g_n)\right) - f\left(\frac{1}{\sqrt{n}}(X_1 + \cdots + X_k + X_{k+1} + g_{k+2} + \cdots + g_n)\right)\right]\right|.$$

It suffices to show that the sum goes to 0, and we will proceed by bounding each summand. To do so, recall the Taylor expansion formula

$$f(a+b) = f(a) + bf'(a) + \frac{b^2}{2} f''(a) + \mathcal{O}(|b|^3),$$

which suggests that for each k, we can let

$$b_k = g_{k+1}, \quad \tilde{b}_k = X_{k+1}, \quad a_k = \text{sum of all other terms}$$

and Taylor expand around a_k. Notice that $\mathbb{E}[b_k] = \mathbb{E}[\tilde{b}_k] = 0$ and $\mathbb{E}[b_k^2] = \mathbb{E}[\tilde{b}_k^2] = 1$ and $a_k \perp\!\!\!\perp (b_k, \tilde{b}_k)$. Thus, the original kth summand is

$$\left|\mathbb{E}\left[f\Big(\frac{a_k+b_k}{\sqrt{n}}\Big) - f\Big(\frac{a_k+\tilde{b}_k}{\sqrt{n}}\Big)\right]\right| \le \mathcal{O}\left(\mathbb{E}[|b_k|^3 + |\tilde{b}_k|^3] \times n^{-3/2}\right),$$

since b_k and $\tilde{b}_k$ have the same first and second moments, meaning the first and second order terms in the Taylor expansion cancel out and only the third order terms are left. The final expression is at most $\mathcal{O}(n^{-3/2})$ since $\mathbb{E}|\tilde{b}_k|^3 = \mathbb{E}|X_{k+1}|^3 \le C$ and $\mathbb{E}|b_k| = \mathbb{E}|g_{k+1}|^3$ which is also bounded.

Hence we conclude that

$$\sum_{k=0}^{n-1}\left|\mathbb{E}\left[f\left(\frac{1}{\sqrt{n}}(X_1+\cdots+X_k+g_{k+1}+\cdots+g_n)\right) - f\left(\frac{1}{\sqrt{n}}(X_1+\cdots+X_k+X_{k+1}+g_{k+2}+\cdots+g_n)\right)\right]\right| \le \mathcal{O}(n^{-1/2}),$$

which converges to 0 as $n \to \infty$. In other proofs, it is not uncommon to use this method of replacing each summand by a Gaussian and bounding the differences. □

Remark. It is also true (but more complicated to show) that properly scaled averages of independent heavy tailed random variables often converge to a "stable distribution". In fact for any $\alpha \in (0,2]$, there exist α-stable laws such that:

$$\frac{X_1+\cdots+X_n}{n^{1/\alpha}} \stackrel{d}{=} X_1.$$

The CLT impies that the only 2-stable laws are centered Gaussian, but for other α there are more. The main tool to study them is characteristic functions. See also: Feller's Probability textbook, the Wikipedia page on "Stable distribution".

Remark. If additional moments match, then Taylor expanding further in the above proof gives a faster rate of convergence. A common situation is that 3rd moment match whenever $X_i \stackrel{d}{=} -X_i$ has symmetric distribution.

2 Martingale Central Limit Theorem

Let $\{X_{n,i}, 1 \le i \le m(n)\}$ be a triangular array for $n \in \mathbb{Z}_+$. In our application $m(n) = n$. Define:

$$S_{n,k} = \sum_{i=1}^{k} X_{n,i}, \qquad V_{n,k}^2 = \sum_{i=1}^{k} \mathbb{E}[X_{n,i}^2 \mid \mathcal{F}_{n,i-1}]$$

Theorem 2.1. *Assume:*

- *(Martingale difference)* $\mathbb{E}[X_{n,i} \mid \mathcal{F}_{n,i-1}] = 0$.
- $V_{n,m(n)}^2 = 1$.

- *(Lindeberg condition)*

$$\forall \delta > 0, \quad \lim_{n\to\infty} \sum_{i=1}^{m(n)} \mathbb{E}\left[X_{n,i}^2 \cdot \mathbf{1}_{\{|X_{n,i}|>\delta\}}\right] = 0.$$

Then, $S_{n,m(n)} \xrightarrow{d} \mathcal{N}(0,1)$.

Note that the Lindeberg condition is intuitively the "weakest" thing that one might require if one hopes for a central limit theorem. It is saying that the variance contribution from each individual martingale difference term getting large becomes negligible when higher moments are bounded.

Proof. We often drop the n subscript for simplicity. Let $g_1, \dots, g_{m(n)} \sim \mathcal{N}(0,1)$ i.i.d., and define:

$$Y_i = \sigma_{n,i} g_i, \qquad \sigma_{n,i}^2 = \mathbb{E}[X_{n,i}^2 \mid \mathscr{F}_{n,i-1}]$$

We want to bound:

$$\sum_{k=1}^{m(n)} \left|\mathbb{E}\left[f\left(X_1 + \cdots + +X_{k-1} + Y_k + Y_{k+1} + \cdots + Y_{m(n)}\right) - f\left(X_1 + \cdots + X_{k-1} + X_k + Y_{k+1} + \cdots + Y_{m(n)}\right)\right]\right| \tag{3}$$

Since the Lindeberg condition is weaker than 3rd moments, we need to be more careful with Taylor expansion error. We will use the fact that given a smooth function $f \in C_b^3(\mathbb{R})$, for any $\varepsilon > 0$, there exist δ, C such that $\forall a, b \in \mathbb{R}$:

(a)

$$|b| \le \delta \Rightarrow |f(a+b) - f(a) - bf'(a) - \tfrac{1}{2}b^2 f''(a)| \le \varepsilon b^2$$

(b)

$$\forall b, \quad |f(a+b) - f(a) - bf'(a) - \tfrac{1}{2}b^2 f''(a)| \le C|b|^2.$$

We apply this with

$$a = X_1 + \cdots + X_k,$$
$$b_k = X_{k+1},$$
$$\tilde{b}_k = Y_{k+1}.$$

(The rest of the proof will be outlined in class next lecture, due to some confusion on Mark's end.)

The idea is that the future terms

$$R_{k+1:m} \equiv Y_{k+1} + \cdots + Y_{m(n)}$$

are "just Gaussian noise", and in fact **the total variance is known at time** k. This is because we assumed $V_{n,m}^2 = 1$, and our replacements preserve conditional variance. Hence the total conditional variance of these future terms is:

$$W_k \equiv \sum_{i=k+1}^{m(n)} \mathbb{E}[Y_{n,i}^2 | \mathscr{F}_{n,i-1}] = \sum_{i=k+1}^{m(n)} \mathbb{E}[X_{n,i}^2 | \mathscr{F}_{n,i-1}] = V_{n,m}^2 - \sum_{i=1}^{k} \mathbb{E}[Y_{n,i}^2 | \mathscr{F}_{n,i-1}] = 1 - \sum_{i=1}^{k} \mathbb{E}[Y_{n,i}^2 | \mathscr{F}_{n,i-1}].$$

Crucially the sum in the last expression is $\mathscr{F}_{n,k-1}$-measurable, hence so is the "total future variance" W_k. This implies that conditionally on $\mathscr{F}_k$, the law of $R_{k+1:m}$ is always $N(0, W_k)$. Hence we bound each sum in (3) by averaging over the future, via the function

$$f_k \equiv f * N(0, W_k)$$

which is given explicitly by

$$f_k(x) = \mathbb{E}^{z \sim N(0, W_k)}[f(x+z)].$$

For $j \in \{1,2,3\}$ the j-th derivative is similarly given by $f_k^{(j)}(x) = \mathbb{E}^{z\sim N(0,W_k)}[f^{(j)}(x+z)]$, and so each f_k inherits the smoothness bounds of f. Replacing $R_{k+1:m}$ with exogenous Gaussian randomness (see Lemma 2.2 below for a careful justification), the k-th term in (3) is equal to

$$\left|\mathbb{E}\left[f_k(X_1+\cdots+Y_k) - f_k(X_1+\cdots+X_k)\right]\right| \le \left|\mathbb{E}\left[\mathbb{E}\left[f_k(X_1+\cdots+Y_k) - f_k(X_1+\cdots+X_k)\,|\mathscr{F}_{k-1}\right]\right]\right|$$

We now match Taylor expansions around $X_1+\cdots+X_{k-1}$, using the 3rd moment bound for Y_k, to find that this is bounded in expectation by

$$C(f)\cdot\mathbb{E}[\epsilon X_k^2 + C(X_k^2\cdot 1_{|X_k|\ge\delta}) + |Y_k|^3].$$

For the first part, we have

$$\mathbb{E}[\epsilon\sum_k X_k^2 + Y_k^2] = 2\epsilon$$

which tends to zero. For the second part, the Lindeberg condition implies

$$\mathbb{E}\sum_k[X_k^2\cdot 1_{|X_k|\ge\delta}] \to 0.$$

The Y_k terms are also not difficult to handle. With $\sigma_k = \sigma_{n,k}$ as above, recalling that $\sum_k \sigma_k^2 = 1$ almost surely, we have

$$\sum_k |Y_k|^3 = \sum_k \sigma_i^3 \le \max_k(\sigma_k)\left(\mathbb{E}\sum_k \sigma_k^2\right) = \max_k \sigma_k.$$

Hence to control the $|Y_k|^3$ terms, it suffices to show that $\max_k \sigma_k \to 0$ in probability (we want this in expectation, but it is at most 1 almost surely). Indeed fixing say $c \in (0,0.1)$, let τ be the first time with $\sigma_k \ge c > 0$. Then for $\delta \ll c$, we have

$$\sum_i \mathbb{E}\left[X_{n,i}^2\cdot\mathbf{1}_{\{|X_{n,i}|>\delta\}}\right] \ge \mathbb{E}\left[1_{\tau<\infty}\cdot\mathbb{E}\left[X_{n,\tau}^2\cdot\mathbf{1}_{\{|X_{n,\tau}|>\delta\}}\right]\right] \ge \mathbb{E}\left[1_{\tau<\infty}\cdot\mathbb{E}\left[X_{n,\tau}^2\right]\right] - \delta^2 \ge c\cdot\mathbb{P}[\tau<\infty] - \delta^2.$$

As this expression tends to 0 for any δ as $n\to\infty$, it is eventually smaller than say δ^2, so we conclude that $\mathbb{P}[\tau<\infty] \le 2\delta^2/c$, which is arbitrarily small for $\delta \ll c$. Since $\lim_{n\to\infty}\mathbb{P}[\tau<\infty] = 0$ for any arbitrary c, we conclude that $\max_i \sigma_i \to 0$ in probability, completing the proof. □

Lemma 2.2. *Let $(\sigma_1,\dots,\sigma_m)$ be a predictable process of non-negative reals such that $\sum_{i=1}^m \sigma_i^2 = 1$ almost surely. Let $g_1,\dots,g_m$ be IID Gaussians also adapted to the same filtration, and suppose that g_i is distributed as $N(0,1)$ conditionally on $\mathscr{F}_{i-1}$. Then*

$$S \equiv \sum_{i=1}^m \sigma_i g_i \sim N(0,1).$$

Proof. We induct on m in the backward direction. Namely, the result is clear for $m=1$, and so by inductive hypothesis we know that conditionally on $\mathscr{F}_1$, the conditional law of S is $\sigma_1 g_1 + \sqrt{1-\sigma_1^2}g'$ for an IID standard Gaussian g'. Since σ_1 is $\mathscr{F}_0$-measurable, conditionally on $\mathscr{F}_0$, both coefficients $(\sigma_1, \sqrt{1-\sigma_1^2})$ are determined by (g_1, g') are conditionally IID Gaussian, completing the proof. □

3 Lévy's Characterization of BM, Dubins–Schwarz Theorem

For now, take the Martingale central limit theorem to be true. A consequence is Lévy's Characterization of Brownian Motion.

I.e. σ_i is $\mathscr{F}_{i-1}$-measurable.

Theorem 3.1. *(Lévy's Characterization of Brownian Motion.) If $(X_t)_{t\geq 0}$ is a continuous martingale and $(X_t^2 - t)_{t\geq 0}$ is also a continuous martingale, then X_t is Brownian motion.*

The proof strategy of this result is to apply the Martingale CLT to a discretization of the martingale $X_{\frac{t}{n}}, X_{\frac{2t}{n}}, \ldots, X_{\frac{nt}{n}}$ to show that $X_t \sim N(0,t)$ and then repeat, conditioning on the past. The Lindeberg condition in this setting follows from continuity. Note that

$$\mathbb{E}[X_t^2] = t$$

and so by dominated convergence,

$$\mathbb{E}\left[X_t^2 \cdot \mathbf{1}_{\left\{\max_{1\leq i\leq n}\left|X_{\frac{it}{n}} - X_{\frac{(i-1)t}{n}}\right|>\delta\right\}}\right] \to 0 \quad \text{a.s.}, \quad n\to\infty, \quad \forall \delta > 0.$$

The latter estimate is slightly stronger than what the Lindeberg condition requires: it shows that the contribution when any one increment exceeds δ tends to zero as $n \to \infty$, while in the Lindeberg condition we only keep the large increments "à la carte". Remember here $E[X_t^2] = \sum_k E[(\Delta B_{t_k})^2]$ because increments are independent.

A corollary of Lévy's Characterization of Brownian Motion is the following Dubins–Schwarz Theorem.

Theorem 3.2. *Any continuous martingale X_t is a time-changed Brownian motion.*

We next sketch a proof for a continuous martingale that has the form of $X_t = \int_0^t \sigma_s \, dB_s$. In this case, its quadratic variation is given by $A_t = \int_0^t \sigma_s^2 \, ds$, and $X_t^2 - A_t$ is a martingale. We let

$$Y_s := X_{\tau_s}, \quad \text{where } \tau_s = \inf\{t : A_t = s\}.$$

Since τ_s is a stopping time, by the optional stopping theorem, we know that Y_s and $Y_s^2 - s$ are martingales with respect to the stopped filtration $(\mathscr{F}_{\tau_s})$. To see this, the optional stopping theorem tells us that for $t > s$, we have $\mathbb{E}[X_{\tau_t} \mid \mathscr{F}_{\tau_s}] = X_{\tau_s}$ and thus $\mathbb{E}[Y_t \mid \mathscr{F}_{\tau_s}] = Y_s$.

By Lévy's characterization, we conclude that Y_s is standard Brownian motion.

Remark. To fully prove this result for any continuous martingale X_t, one has to argue that X_t, A_t satisfies some technical properties such that there cannot be times where A_t varies but X_t doesn't. This is straightforward to see for the form of martingales we have described above.