The proof we did not give

Lecture 3 proved the central limit theorem by swapping summands for Gaussians one at a time. There is an older proof, and we have had the tools for it since Lecture 1: the characteristic function turns sums into products, and a product of n nearly-equal factors is an exponential. This assignment gives that proof, then runs the identical computation on a different array and gets Poisson instead — so the Gaussian/Poisson boundary becomes one calculation done twice. It ends by locating what the argument cannot do, which is why Lecture 4 exists.

RECALLED FROM LECTURE For a random variable X on $\mathbb{R}$, $\varphi_X(t) = \mathbb{E}[e^{itX}]$. It satisfies $\varphi_X(0) = 1$, $|\varphi_X| \le 1$, and $\varphi_{aX+b}(t) = e^{itb}\varphi_X(at)$. If $X \perp Y$ then $\varphi_{X+Y} = \varphi_X\varphi_Y$. A standard Gaussian has $\varphi(t) = e^{-t^2/2}$ (Lemma 1.2). We write $S_n = n^{-1/2}\sum_{i\le n} X_i$ for i.i.d. X_i with mean 0 and variance 1, and $g \sim \mathcal{N}(0,1)$.

IMPORTED, WITHOUT PROOF **Lévy's continuity theorem.** If $\varphi_{Y_n}(t) \to \psi(t)$ for every $t \in \mathbb{R}$ and ψ is continuous at 0, then $\psi = \varphi_Y$ for some random variable Y and $Y_n \Rightarrow Y$. This is the one thing below that we take on faith; everything else is calculus.

1. What the transform knows about moments

The transform is a generating function: differentiating at $t = 0$ pulls down moments.

Problem 1. CORE Let $\mathbb{E}|X|^k < \infty$. Differentiating under the integral k times, show $\varphi_X^{(k)}(0) = i^k\,\mathbb{E}[X^k]$. Check it against $\varphi(t) = e^{-t^2/2}$ by computing $\varphi''(0)$ and $\varphi^{(4)}(0)$, and confirm you recover $\mathbb{E}g^2 = 1$ and $\mathbb{E}g^4 = 3$.

For the proof we need slightly more than the value of the second derivative: we need a Taylor expansion whose error we can control uniformly in n.

Problem 2. CORE Suppose $\mathbb{E}X = 0$ and $\mathbb{E}X^2 = 1$. Show that

$$\varphi_X(t) \;=\; 1 - \tfrac{t^2}{2} + r(t), \qquad \frac{r(t)}{t^2} \longrightarrow 0 \;\text{ as } t \to 0.$$

Hint: $|e^{i\theta} - 1 - i\theta + \theta^2/2| \le \min(|\theta|^3/6,\ \theta^2)$. *Apply it with* $\theta = tX$, *take expectations, and use dominated convergence — the second bound in the minimum is what lets you avoid assuming a third moment.*

2. Products of nearly-equal factors

Problem 3. CORE Let $z_n \to z$ in $\mathbb{C}$. Show $\left(1 + \frac{z_n}{n}\right)^n \to e^z$. *Hint: for complex* a, b *with* $|a|, |b| \le M$ *one has* $|a^n - b^n| \le nM^{n-1}|a - b|$; *take* $b = e^{z_n/n}$.

Problem 4. CORE Combine the last two problems. Show that for each fixed t,

$$\varphi_{S_n}(t) \;=\; \Big[\varphi_{X_1}\big(t/\sqrt{n}\big)\Big]^n \;\longrightarrow\; e^{-t^2/2},$$

and conclude by Lévy's theorem that $S_n \Rightarrow \mathcal{N}(0,1)$. This is the central limit theorem. Note where independence was used, and where identical distribution was used.

3. The same computation, a different limit

Nothing in §2 was about Gaussians. It was about a product of n factors each of the form $1 + z/n$. Change what sits in the numerator and the limit changes with it.

Problem 5. CORE Let $B \sim \mathrm{Bernoulli}(p)$. Show $\varphi_B(t) = 1 + p\,(e^{it} - 1)$. Now let $B_{n,1}, \dots, B_{n,n}$ be independent $\mathrm{Bernoulli}(\lambda/n)$ and $N_n = \sum_i B_{n,i}$. Show $\varphi_{N_n}(t) \to \exp(\lambda(e^{it} - 1))$. Finally, sum the series $\sum_k e^{-\lambda}\frac{\lambda^k}{k!}e^{itk}$

to identify that limit as the characteristic function of Poisson(λ), so $N_n \Rightarrow \text{Poisson}(\lambda)$.

The two proofs are the same three lines. What differs is the normalization: in §2 we divided by $\sqrt{n}$ to hold the variance at 1; here we divided by nothing.

Problem 6. Explain, using Problem 5, what goes wrong if you insist on the Gaussian normalization: compute $\varphi(t)$ for $s_n^{-1}\sum_i (B_{n,i} - \lambda/n)$ with $s_n^2 = \lambda(1-\lambda/n)$, expand it, and identify the term that fails to vanish. Which hypothesis of Problem 2 is violated?

4. Lindeberg's condition, computed

Lecture 3 stated Lindeberg's condition for a triangular array $\{X_{n,i}\}_{i\le k_n}$, independent within rows, with mean 0 and $s_n^2 = \sum_i \operatorname{Var}(X_{n,i})$: for every $\varepsilon > 0$,

$$L_n(\varepsilon) \;=\; \frac{1}{s_n^2}\sum_i \mathbb{E}\big[X_{n,i}^2\,\mathbf{1}\{|X_{n,i}| > \varepsilon s_n\}\big] \;\longrightarrow\; 0.$$

Compute it in three cases.

Problem 7. CORE *(It holds.)* For i.i.d. X_i with mean 0 and variance 1, set $X_{n,i} = X_i$ for $i \le n$. Show $s_n^2 = n$ and $L_n(\varepsilon) = \mathbb{E}[X_1^2\mathbf{1}\{|X_1| > \varepsilon\sqrt{n}\}] \to 0$ for every $\varepsilon > 0$.

Problem 8. *(It fails: rare events.)* For the centred Bernoulli array $X_{n,i} = B_{n,i} - \lambda/n$ of Problem 5, show that $L_n(\varepsilon) \to 1$ for every $\varepsilon < 1/\sqrt{\lambda}$. Then compute $L_n(\frac{1}{2})$ exactly at $\lambda = 3$ for $n = 10$, 100, 10^4; you should get 0, 0.970, 1.000. Explain why the first entry is exactly 0.

Problem 9. *(It fails: one large summand.)* Take $X_{n,i} = \pm 1$ for $i \le n$ and one extra summand $X_{n,n+1} = \pm\sqrt{n}$, all independent and symmetric. Show $s_n^2 = 2n$ and that $L_n(\varepsilon) \ge \frac{1}{2}$ for $\varepsilon < 1/\sqrt{2}$. Then compute the limiting characteristic function of $s_n^{-1}\sum_i X_{n,i}$ — you should get $e^{-t^2/4}\cos(t/\sqrt{2})$ — and show the limit is not Gaussian by exhibiting a t at which it vanishes. *Why does that settle it?*

5. What the argument cannot do

Problem 10. *(No variance, no theorem.)* The standard Cauchy law has density $\frac{1}{\pi(1+x^2)}$ and characteristic function $e^{-|t|}$. For i.i.d. Cauchy X_i, compute $\varphi(t)$ for the average $\bar{X}_n = \frac{1}{n}\sum_i X_i$ and show that $\bar{X}_n$ has exactly the law of X_1, for every n. Averaging accomplishes nothing. Which step of Problem 2 breaks, and what is the expansion of $e^{-|t|}$ near 0?

Problem 11. *(No dependence.)* Let $X_1, \dots, X_n$ be identically distributed with mean 0 and variance 1 but *not* independent. Point to the exact equation in Problem 4 that fails, and explain why no amount of care with the Taylor expansion repairs it. Then reread the proof of Theorem 1.4 in the notes and say which of its steps used independence, and how much of it would survive if the X_i were only pairwise uncorrelated with matching second moments.

The last problem is the reason for Lecture 4. Lévy's theorem is also not quantitative: it converts pointwise convergence of φ_{S_n} into weak convergence and stops. Stein's method replaces the factorization $\varphi_{X+Y} = \varphi_X\varphi_Y$ — which is where independence is spent — with an identity checkable one summand at a time, and it returns a rate.

CORE problems are short and self-contained. The quiz on Wednesday Sep 16 draws on the core problems of this set and of Homework 4. AI assistance is permitted and encouraged throughout.