Argue that this fact proves $\phi_{\mathbf{X}}(\mathbf{t})$ uniquely determines the distribution of $\mathbf{X}$.

Hint: Use parts (a) and (b) to show that the distribution of $\mathbf{X}+\mathbf{Y}$ depends on $\mathbf{X}$ only through $\phi_{\mathbf{X}}$. Then note that $\mathbf{X}+\mathbf{Y} \stackrel{d}{\to} \mathbf{X}$ as $\sigma^2 \to 0$.

Exercise 4.3 Use the Continuity Theorem to prove the Cramér-Wold Theorem, Theorem 4.12.

Hint: $\mathbf{a}^\top \mathbf{X}_n \stackrel{d}{\to} \mathbf{a}^\top \mathbf{X}$ implies that $\phi_{\mathbf{a}^\top \mathbf{X}_n}(1) \to \phi_{\mathbf{a}^\top \mathbf{X}}(1)$.

Exercise 4.4 Suppose $\mathbf{X} \sim N_k(\boldsymbol{\mu}, \Sigma)$, where Σ is invertible. Prove that

$$(\mathbf{X}-\boldsymbol{\mu})^\top \Sigma^{-1} (\mathbf{X}-\boldsymbol{\mu}) \sim \chi^2_k.$$

Hint: If Q diagonalizes Σ, say $Q\Sigma Q^\top = \Lambda$, let $\Lambda^{1/2}$ be the diagonal, nonnegative matrix satisfying $\Lambda^{1/2}\Lambda^{1/2} = \Lambda$ and consider $\mathbf{Y}^\top\mathbf{Y}$, where $Y = (\Lambda^{1/2})^{-1} Q(\mathbf{X}-\boldsymbol{\mu})$.

Exercise 4.5 Let $X_1, X_2, \ldots$ be independent Poisson random variables with mean $\lambda = 1$. Define $Y_n = \sqrt{n}(\overline{X}_n - 1)$.

(a) Find $\mathrm{E}\,(Y_n^+)$, where $Y_n^+ = Y_n I\{Y_n > 0\}$.

(b) Find, with proof, the limit of $\mathrm{E}\,(Y_n^+)$ and prove Stirling's formula

$$n! \;\sim\; \sqrt{2\pi}\, n^{n+1/2} e^{-n}.$$

Hint: Use the result of Problem 3.12.

Exercise 4.6 Use the Continuity Theorem to prove Theorem 2.19.

Hint: Use a Taylor expansion (1.5) with $d = 2$ for both the real and imaginary parts of the characteristic function of $\overline{X}_n$.

Exercise 4.7 Use the Cramér-Wold Theorem along with the univariate Central Limit Theorem (from Example 2.12) to prove Theorem 4.9.

4.2 The Lindeberg-Feller Central Limit Theorem

The Lindeberg-Feller Central Limit Theorem states in part that sums of independent random variables, properly standardized, converge in distribution to standard normal as long as a certain condition, called the Lindeberg Condition, is satisfied. Since these random variables do not have to be identically distributed, this result generalizes the Central Limit Theorem for independent and identically distributed sequences.

4.2.1 The Lindeberg and Lyapunov Conditions

Suppose that $X_1, X_2, \ldots$ are independent random variables such that $\mathrm{E}\, X_n = \mu_n$ and $\mathrm{Var}\, X_n = \sigma_n^2 < \infty$. Define

$$\begin{aligned} Y_n &= X_n - \mu_n, \\ T_n &= \textstyle\sum_{i=1}^n Y_i, \\ s_n^2 &= \mathrm{Var}\; T_n = \textstyle\sum_{i=1}^n \sigma_i^2. \end{aligned}$$

Instead of defining Y_n to be the centered version of X_n, we could have simply taken μ_n to be zero without loss of generality. However, when these results are used in practice, it is easy to forget the centering step, so we prefer to make it explicit here.

Note that T_n/s_n has mean zero and variance 1. We wish to give sufficient conditions that ensure $T_n/s_n \stackrel{d}{\to} N(0,1)$. We give here two separate conditions, one called the Lindeberg condition and the other called the Lyapunov condition. The *Lindeberg Condition* for sequences states that

$$\text{for every } \epsilon > 0,\ \frac{1}{s_n^2}\sum_{i=1}^n \mathrm{E}\left(Y_i^2 I\left\{|Y_i| \ge \epsilon s_n\right\}\right) \to 0 \text{ as } n \to \infty; \tag{4.8}$$

the *Lyapunov Condition* for sequences states that

$$\text{there exists } \delta > 0 \text{ such that } \frac{1}{s_n^{2+\delta}}\sum_{i=1}^n \mathrm{E}\left(|Y_i|^{2+\delta}\right) \to 0 \text{ as } n \to \infty. \tag{4.9}$$

We shall see later (in Theorem 4.17, the Lindeberg-Feller Theorem) that Condition (4.8) implies $T_n/s_n \to N(0,1)$. For now, we show only that Condition (4.9) — the Lyapunov Condition — is stronger than Condition (4.8). Thus, the Lyapunov Condition also implies $T_n/s_n \to N(0,1)$:

Theorem 4.14 The Lyapunov Condition (4.9) implies the Lindeberg Condition (4.8).

Proof: Assume that the Lyapunov Condition is satisfied and fix $\epsilon > 0$. Since $|Y_i| \ge \epsilon s_n$ implies $|Y_i/\epsilon s_n|^\delta \ge 1$, we obtain

$$\begin{aligned} \frac{1}{s_n^2}\sum_{i=1}^n \mathrm{E}\left(Y_i^2 I\left\{|Y_i| \ge \epsilon s_n\right\}\right) &\le \frac{1}{\epsilon^\delta s_n^{2+\delta}}\sum_{i=1}^n \mathrm{E}\left(|Y_i|^{2+\delta} I\left\{|Y_i| \ge \epsilon s_n\right\}\right) \\ &\le \frac{1}{\epsilon^\delta s_n^{2+\delta}}\sum_{i=1}^n \mathrm{E}\left(|Y_i|^{2+\delta}\right). \end{aligned}$$

Since the right hand side tends to 0, the Lindeberg Condition is satisfied. ■

Example 4.15 Suppose that we perform a series of independent Bernoulli trials with possibly different success probabilities. Under what conditions will the proportion of successes, properly standardized, tend to a normal distribution?

Let $X_n \sim \text{Bernoulli}(p_n)$, so that $Y_n = X_n - p_n$ and $\sigma_n^2 = p_n(1-p_n)$. As we shall see later (Theorem 4.17), either the Lindeberg Condition (4.8) or the Lyapunov Condition (4.9) will imply that $\sum_{i=1}^n Y_i/s_n \xrightarrow{d} N(0,1)$.

Let us check the Lyapunov Condition for, say, $\delta = 1$. First, verify that

$$\mathrm{E}\,|Y_n|^3 = p_n(1-p_n)^3 + (1-p_n)p_n^3 = \sigma_n^2[(1-p_n)^2 - p_n^2] \le \sigma_n^2.$$

Using this upper bound on $\mathrm{E}\,|Y_n|^3$, we obtain $\sum_{i=1}^n \mathrm{E}\,|Y_i|^3 \le s_n^2$. Therefore, the Lyapunov condition is satisfied whenever $s_n^2/s_n^3 \to 0$, which implies $s_n \to \infty$. We conclude that the proportion of successes tends to a normal distribution whenever

$$s_n^2 = \sum_{i=1}^n p_n(1-p_n) \to \infty,$$

which will be true as long as $p_n(1-p_n)$ does not tend to 0 too fast.

4.2.2 Independent and Identically Distributed Variables

We now set the stage for proving a central limit theorem for independent and identically distributed random variables by showing that the Lindeberg Condition is satisfied by such a sequence as long as the common variance is finite.

Example 4.16 Suppose that $X_1, X_2, \ldots$ are independent and identically distributed with $\mathrm{E}\,(X_i) = \mu$ and $\mathrm{Var}\,(X_i) = \sigma^2 < \infty$. The case $\sigma^2 = 0$ is uninteresting, so we assume $\sigma^2 > 0$.

Let $Y_i = X_i - \mu$ and $s_n^2 = \mathrm{Var}\,\sum_{i=1}^n Y_i = n\sigma^2$. Fix $\epsilon > 0$. The Lindeberg Condition states that

$$\frac{1}{n\sigma^2}\sum_{i=1}^n \mathrm{E}\left(Y_i^2 I\{|Y_i| \ge \epsilon\sigma\sqrt{n}\}\right) \to 0 \quad \text{as } n \to \infty. \tag{4.10}$$

Since the Y_i are identically distributed, the left hand side of expression (4.10) simplifies to

$$\frac{1}{\sigma^2}\,\mathrm{E}\left(Y_1^2 I\{|Y_1| \ge \epsilon\sigma\sqrt{n}\}\right). \tag{4.11}$$

To simplify notation, let Z_n denote the random variable $Y_1^2 I\{|Y_1| \ge \epsilon\sigma\sqrt{n}\}$. Thus, we wish to prove that $\mathrm{E}\,Z_n \to 0$. Note that Z_n is nonzero if and only if $|Y_1| \ge \epsilon\sigma\sqrt{n}$. Since this event has probability tending to zero as $n \to \infty$, we conclude that $Z_n \xrightarrow{P} 0$ by the definition of convergence in probability. We can also see that $|Z_n| \le Y_1^2$, and we know that $\mathrm{E}\,Y_1^2 < \infty$. Therefore, we may apply the Dominated Convergence Theorem, Theorem 3.22, to conclude that $\mathrm{E}\,Z_n \to 0$. This demonstrates that the Lindeberg Condition is satisfied.

The preceding argument, involving the Dominated Convergence Theorem, is quite common in proofs that the Lindeberg Condition is satisfied. Any beginning student is well-advised to study this argument carefully.

Note that the assumptions of Example 4.16 are not strong enough to ensure that the Lyapunov Condition (4.9) is satisfied. This is because there are some random variables that have finite variances but no finite $2+\delta$ moment for any $\delta > 0$. Construction of such an example is the subject of Exercise 4.10. However, such examples are admittedly somewhat pathological, and if one is willing to assume that $X_1, X_2, \ldots$ are independent and identically distributed with $\mathrm{E}\,|X_1|^{2+\delta} = \gamma < \infty$ for some $\delta > 0$, then the Lyapunov Condition is much easier to check than the Lindeberg Condition. Indeed, because $s_n = \sigma\sqrt{n}$, the Lyapunov Condition reduces to

$$\frac{n\gamma}{(n\sigma^2)^{1+\delta/2}} = \frac{\gamma}{n^{\delta/2}\sigma^{2+\delta}} \to 0,$$

which follows immediately.

4.2.3 Triangular Arrays

It is sometimes the case that $X_1, \ldots, X_n$ are independent random variables — possibly even identically distributed — but their distributions depend on n. Take the simple case of the binomial(n, p_n) distribution as an example, where the probability p_n of success on any trial changes as n increases. What can we say about the asymptotic distribution in such a case? It seems that what we need is some way of dealing with a sequence of sequences, say, $X_{n1}, \ldots, X_{nn}$ for $n \ge 1$. This is exactly the idea of a triangular array of random variables.

Generalizing the concept of "sequence of independent random variables," a triangular array or random variables may be visualized as follows:

$$\begin{array}{llll} X_{11} & & & \leftarrow \text{independent} \\ X_{21} & X_{22} & & \leftarrow \text{independent} \\ X_{31} & X_{32} & X_{33} & \leftarrow \text{independent} \\ & \vdots & & \end{array}$$

Thus, we assume that for each n, $X_{n1}, \ldots, X_{nn}$ are independent. Carrying over the notation from before, we assume E $X_{ni} = \mu_{ni}$ and Var $X_{ni} = \sigma^2_{ni} < \infty$. Let

$$\begin{aligned} Y_{ni} &= X_{ni} - \mu_{ni}, \\ T_n &= \textstyle\sum_{i=1}^n Y_{ni}, \\ s_n^2 &= \operatorname{Var} T_n = \textstyle\sum_{i=1}^n \sigma^2_{ni}. \end{aligned}$$

As before, T_n/s_n has mean 0 and variance 1; our goal is to give conditions under which

$$\frac{T_n}{s_n} \xrightarrow{d} N(0,1). \tag{4.12}$$

Such conditions are given in the Lindeberg-Feller Central Limit Theorem. The key to this theorem is the *Lindeberg condition* for triangular arrays:

$$\text{For every } \epsilon > 0,\ \frac{1}{s_n^2} \sum_{i=1}^n \mathrm{E}\left(Y_{ni}^2 I\left\{|Y_{ni}| \geq \epsilon s_n\right\}\right) \to 0 \text{ as } n \to \infty. \tag{4.13}$$

Before stating the Lindeberg-Feller theorem, we need a technical condition that says essentially that the contribution of each X_{ni} to s_n^2 should be negligible:

$$\frac{1}{s_n^2} \max_{i \leq n} \sigma^2_{ni} \to 0 \text{ as } n \to \infty. \tag{4.14}$$

Now that Conditions (4.12), (4.13), and (4.14) have been written, the main result may be stated in a single line:

Theorem 4.17 *Lindeberg-Feller Central Limit Theorem:* Condition (4.13) holds if and only if Conditions (4.12) and (4.14) hold.

A proof of the Lindeberg-Feller Theorem is the subject of Exercises 4.8 and 4.9. In most practical applications of this theorem, the Lindeberg Condition (4.13) is used to establish asymptotic normality (4.12); the remainder of the theorem's content is less useful.

Example 4.18 As an extension of Example 4.15, suppose $X_n \sim \text{binomial}(n, p_n)$. The calculations here are not substantially different from those in Example 4.15, so we use the Lindeberg Condition here for the purpose of illustration. We claim that

$$\frac{X_n - np_n}{\sqrt{np_n(1-p_n)}} \xrightarrow{d} N(0,1) \tag{4.15}$$

whenever $np_n(1-p_n) \to \infty$ as $n \to \infty$. In order to use Theorem 4.17 to prove this result, let $Y_{n1}, \ldots, Y_{nn}$ be independent and identically distributed with

$$P(Y_{ni} = 1 - p_n) = 1 - P(Y_{ni} = -p_n) = p_n.$$

Then with $X_n = np_n + \sum_{i=1}^n Y_{ni}$, we obtain $X_n \sim \text{binomial}(n, p_n)$ as specified. Furthermore, E $Y_{ni} = 0$ and Var $Y_{ni} = p_n(1-p_n)$, so the Lindeberg condition says that for any $\epsilon > 0$,

$$\frac{1}{np_n(1-p_n)} \sum_{i=1}^n \mathrm{E}\left(Y_{ni}^2 I\left\{|Y_{ni}| \geq \epsilon\sqrt{np_n(1-p_n)}\right\}\right) \to 0. \tag{4.16}$$

Since $|Y_{ni}| \leq 1$, the left hand side of expression (4.16) will be identically zero whenever $\epsilon\sqrt{np_n(1-p_n)} > 1$. Thus, a sufficient condition for (4.15) to hold is that $np_n(1-p_n) \to \infty$. One may show that this is also a necessary condition (this is Exercise 4.11).

Note that any independent sequence $X_1, X_2, \ldots$ may be considered a triangular array by simply taking $X_{n1} = X_1$ for all $n \geq 1$, $X_{n2} = X_2$ for all $n \geq 2$, and so on. Therefore, Theorem 4.17 applies equally to the Lindeberg Condition (4.8) for sequences. Furthermore, the proof of Theorem 4.14 is unchanged if the sequence Y_i is replaced by the array Y_{ni}. Therefore, we obtain an alternative means for checking asymptotic normality:

Corollary 4.19 Asymptotic normality (4.12) follows if the triangular array above satisfies the *Lyapunov Condition* for triangular arrays:

$$\text{there exists } \delta > 0 \text{ such that } \frac{1}{s_n^{2+\delta}} \sum_{i=1}^n \mathrm{E}\left(|Y_{ni}|^{2+\delta}\right) \to 0 \text{ as } n \to \infty. \tag{4.17}$$

Combining Theorem 4.17 with Example 4.16, in which the Lindeberg condition is verified for a sequence of independent and identically distributed variables with finite positive variance, gives the result commonly referred to simply as "The Central Limit Theorem":

Theorem 4.20 *Univariate Central Limit Theorem for iid sequences:* Suppose that $X_1, X_2, \ldots$ are independent and identically distributed with $\mathrm{E}\,(X_i) = \mu$ and $\operatorname{Var}(X_i) = \sigma^2 < \infty$. Then

$$\sqrt{n}\left(\frac{\overline{X}_n - \mu}{\sigma}\right) \xrightarrow{d} N(0,1). \tag{4.18}$$

The case $\sigma^2 = 0$ is not covered by Example 4.16, but in this case limit (4.18) holds automatically.

We conclude this section by generalizing Theorem 4.20 to the multivariate case, Theorem 4.9. The proof is straightforward using theorem 4.20 along with the Cramér-Wold theorem, theorem 4.12. Recall that the Cramér-Wold theorem allows us to establish multivariate convergence in distribution by proving univariate convergence in distribution for arbitrary linear combinations of the vector components.

Proof of Theorem 4.9: Let $\mathbf{X} \sim N_k(\mathbf{0}, \Sigma)$ and take any vector $\mathbf{a} \in \mathbb{R}^k$. We wish to show that

$$\mathbf{a}^\top \left[\sqrt{n}\left(\overline{\mathbf{X}}_n - \boldsymbol{\mu}\right)\right] \stackrel{d}{\to} \mathbf{a}^\top \mathbf{X}.$$

But this follows immediately from the univariate Central Limit Theorem, since $\mathbf{a}^\top(\mathbf{X}_1 - \boldsymbol{\mu}), \mathbf{a}^\top(\mathbf{X}_2 - \boldsymbol{\mu}), \ldots$ are independent and identically distributed with mean 0 and variance $\mathbf{a}^\top \Sigma \mathbf{a}$. ∎

We will see many, many applications of the univariate and multivariate Central Limit Theorems in the chapters that follow.

Exercises for Section 4.2

Exercise 4.8 Prove that (4.13) implies both (4.12) and (4.14) (the "forward half" of the Lindeberg-Feller Theorem). Use the following steps:

(a) Prove that for any complex numbers $a_1, \ldots, a_n$ and $b_1, \ldots, b_n$ with $|a_i| \le 1$ and $|b_i| \le 1$,

$$|a_1 \cdots a_n - b_1 \cdots b_n| \le \sum_{i=1}^n |a_i - b_i|. \tag{4.19}$$

Hint: First prove the identity when $n = 2$, which is the key step. Then use mathematical induction.

(b) Prove that

$$\left|\phi_{Y_{ni}}\left(\frac{t}{s_n}\right) - \left(1 - \frac{t^2\sigma_{ni}^2}{2s_n^2}\right)\right| \le \frac{\epsilon|t|^3\sigma_{ni}^2}{s_n^2} + \frac{t^2}{s_n^2} \operatorname{E}\left(Y_{ni}^2 I\{|Y_{ni}| \ge \epsilon s_n\}\right). \tag{4.20}$$

Hint: Use the results of Exercise 1.43, parts (c) and (d), to argue that for any Y,

$$\left|\exp\left\{\frac{\mathrm{i}tY}{s_n}\right\} - \left(1 + \frac{\mathrm{i}tY}{s_n} - \frac{t^2Y^2}{2s_n^2}\right)\right| \le \left|\frac{tY}{s_n}\right|^3 I\left\{\left|\frac{Y}{s_n}\right| < \epsilon\right\} + \left(\frac{tY}{s_n}\right)^2 I\{|Y| \ge \epsilon s_n\}.$$

(c) Prove that (4.13) implies (4.14).

Hint: For any i, show that

$$\frac{\sigma_{ni}^2}{s_n^2} < \epsilon^2 + \frac{\operatorname{E}\left(Y_{ni}^2 I\{|Y_{ni}| \ge \epsilon s_n\}\right)}{s_n^2}.$$

(d) Use parts (a) and (b) to prove that, for n large enough so that $t^2 \max_i \sigma_{ni}^2/s_n^2 \le 1$,

$$\left|\prod_{i=1}^n \phi_{Y_{ni}}\left(\frac{t}{s_n}\right) - \prod_{i=1}^n \left(1 - \frac{t^2\sigma_{ni}^2}{2s_n^2}\right)\right| \le \epsilon|t|^3 + \frac{t^2}{s_n^2}\sum_{i=1}^n \operatorname{E}\left(Y_{ni}^2 I\left\{|Y_{ni}| \ge \epsilon s_n\right\}\right).$$

(e) Use part (a) to prove that

$$\left|\prod_{i=1}^n \left(1 - \frac{t^2\sigma_{ni}^2}{2s_n^2}\right) - \prod_{i=1}^n \exp\left(-\frac{t^2\sigma_{ni}^2}{2s_n^2}\right)\right| \le \frac{t^4}{4s_n^4}\sum_{i=1}^n \sigma_{ni}^4 \le \frac{t^4}{4s_n^2} \max_{1\le i\le n} \sigma_{ni}^2.$$

Hint: Prove that for $x \le 0$, $|1 + x - \exp(x)| \le x^2$.

(f) Now put it all together. Show that

$$\left|\prod_{i=1}^n \phi_{Y_{ni}}\left(\frac{t}{s_n}\right) - \prod_{i=1}^n \exp\left(-\frac{t^2\sigma_{ni}^2}{2s_n^2}\right)\right| \to 0,$$

proving (4.12).

Exercise 4.9 In this problem, we prove the converse of Exercise 4.8, which is the part of the Lindeberg-Feller Theorem due to Feller: Under the assumptions of the Exercise 4.8, the variance condition (4.14) and the asymptotic normality (4.12) together imply the Lindeberg condition (4.13).

(a) Define

$$\alpha_{ni} = \phi_{Y_{ni}}(t/s_n) - 1.$$

Prove that

$$\max_{1\le i\le n} |\alpha_{ni}| \le 2 \max_{1\le i\le n} P\left(|Y_{ni}| \ge \epsilon s_n\right) + 2\epsilon|t|$$

and thus

$$\max_{1\le i\le n} |\alpha_{ni}| \to 0 \quad \text{as } n \to \infty.$$

Hint: Use the result of Exercise 1.43(a) to show that $|\exp\{\mathrm{i}t\} - 1| \le 2\min\{1, |t|\}$ for $t \in \mathbb{R}$. Then use Chebyshev's inequality along with condition (4.14).