

# CLT Proof

MATH/STAT 355: Statistical Theory

Spring 2026

## Central Limit Theorem

Let  $Y_1, \dots, Y_n$  be a random sample from a distribution with mean  $\mu$  and variance  $\sigma^2 < \infty$ , and define  $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$ . Then,

$$\sqrt{n} \left( \frac{\bar{Y}_n - \mu}{\sigma} \right) \xrightarrow{d} Z,$$

where  $Z \sim N(0, 1)$ .

## Central Limit Theorem (stronger assumptions)

Let  $Y_1, \dots, Y_n$  be a random sample from a distribution whose moment generating function exists in a neighborhood of zero (i.e.,  $M_Y(t)$  exists  $\forall |t| < h$  for some  $h > 0$ ). Let  $\mu$  and  $\sigma^2$  represent the mean and variance of that distribution. (Note that  $\mu$  and  $\sigma^2$  exist and are finite since the MGF exists.) Then,

$$\sqrt{n} \left( \frac{\bar{Y}_n - \mu}{\sigma} \right) \xrightarrow{d} Z,$$

where  $Z \sim N(0, 1)$ .

**NOTE:** The assumption about the existence of the MGF is a stronger assumption than we actually need to prove the CLT, but it makes our proof easier! To prove the CLT without this assumption, we use something called a *characteristic function* instead of moment generating functions. Look up characteristic functions if you're interested!

# Proof Toolkit

To prove the CLT, we'll rely on the following results:

- (a) If  $V = aW + b$ , then

$$M_V(t) = e^{bt} M_W(at).$$

- (b) If  $W_1, \dots, W_n$  are independent and  $W = W_1 + \dots + W_n$ , then

$$M_W(t) = \prod_{i=1}^n M_{W_i}(t).$$

- (c) Let  $W_1, \dots, W_n$  be a sequence of random variables with MGFs  $M_{W_1}(t), \dots, M_{W_n}(t)$ , and let  $V$  be another random variable with MGF  $M_V(t)$ . If

$$\lim_{n \rightarrow \infty} M_{W_n}(t) = M_V(t) \quad \forall t \text{ in some neighborhood of } 0,$$

then  $W_n \xrightarrow{d} V$ .

- (d) The Taylor expansion of  $e^x$  about zero is

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

- (e) Let  $a_1, a_2, \dots, a_n$  be a sequence of real numbers such that  $\lim_{n \rightarrow \infty} a_n = a$ . Then

$$\lim_{n \rightarrow \infty} \left(1 + \frac{a_n}{n}\right)^n = e^a.$$

We won't prove any of these results (you probably have in previous math classes!), but we'll use them throughout our proof today.

# Proof

We'll break our proof of the “stronger assumptions” version of the CLT into several steps.

To begin, you can assume that all of the assumptions stated in the CLT Theorem are true:  $Y_1, \dots, Y_n$  are a random sample (iid) from a distribution with mean  $\mu$ , variance  $\sigma^2 < \infty$ , and MGF,  $M_Y(t)$ , that exists in a neighborhood of zero.

**Step 1:** Define  $X_i = \frac{Y_i - \mu}{\sigma}$ . Show that  $E[X_i] = 0$  and  $V[X_i] = 1$ .

**Step 2:** Define  $Z_n = \sqrt{n} \left( \frac{\bar{Y}_n - \mu}{\sigma} \right)$ . Show that  $Z_n = \frac{1}{\sqrt{n}} \sum_{i=1}^n X_i$ .

We must show  $Z_n \xrightarrow{d} N(0, 1)$ . Since the MGFs exist, it suffices to show

$$\lim_{n \rightarrow \infty} M_{Z_n}(t) = e^{t^2/2},$$

since  $e^{t^2/2}$  is the MGF for a standard normal distribution.

We'll do this in a few more steps...

**Step 3:** Find the MGF of  $\sum_{i=1}^n X_i$ .



**Step 4:** Find the MGF of  $Z_n$ .



**Step 5:** Using the Taylor expansion of  $e^x$  about zero, show that  $M_{Z_n}(t) = \left(1 + \frac{t^2}{2!n} + \frac{t^3 E[Y^3]}{3!n^{3/2}} + \dots\right)^n$ .

**Step 6:** Finally, show that  $M_{Z_n}(t) \xrightarrow{n \rightarrow \infty} e^{t^2/2}$ .

