# Lecture 20 – Limit theorems

**Last Modified:** November 5, 2026

## Outline:


20.1 Law of Large Numbers . . . . . 1
20.2 Central Limit Theorem . . . . . 2


## 20.1 Law of Large Numbers

The LLN says that the sample mean converges to the population mean regardless of the underlying distribution of the samples. To appreciate the power of this result, let's prove it directly for normals. Suppose $X_1, \ldots, X_n \overset{\text{iid}}{\sim} \mathcal{N}(\mu, \sigma^2)$. Let's prove directly that $\bar{X}_n \xrightarrow{p} \mu$. We know that a sum of independent normal random variables is normal, with $\bar{X}_n \sim \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right)$. Fix any $\delta > 0$. If $Z \sim \mathcal{N}(0,1)$ and $\Phi$ is its cdf, then

$$\begin{aligned}
\mathbb{P}(|\bar{X}_n - \mu| > \delta) &= \mathbb{P}\left(\left|\frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}\right| > \frac{\sqrt{n}\,\delta}{\sigma}\right) \\
&= \mathbb{P}\left(|Z| > \frac{\sqrt{n}\,\delta}{\sigma}\right) \\
&= 2\left\{1 - \Phi\left(\frac{\sqrt{n}\,\delta}{\sigma}\right)\right\} \longrightarrow 0.
\end{aligned}$$

The last step follows because $\sqrt{n}\,\delta/\sigma \to \infty$ and $\Phi(t) \to 1$ as $t \to \infty$. Since this holds for every $\delta > 0$, we have proved $\bar{X}_n \xrightarrow{p} \mu$.

We heavily used properties of the normal distribution here: the sample mean is exactly normal, and standardizing gives a standard normal distribution whose tail probabilities we can evaluate. Remarkably, the law of large numbers tells us that the same convergence holds much more broadly.

**Theorem 1** (Law of Large Numbers (LLN))**.** Let $\{X_n\}$ be a sequence of iid random variables with finite expectation $\mathbb{E}[X_1] = \mu$. Then

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{p} \mu.$$

The LLN says the sample mean converges to the population mean. In the context of parameter estimation, we say that the sample mean is a *consistent estimator* of the population mean, because it converges to it in probability.

Although the LLN is formulated for means, we can use it in other contexts, too. For example, we now show the sample variance converges to the population variance:

**Proposition 1.** Let $\hat{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$ be the sample variance of $n$ i.i.d. random variables $X_1, \ldots, X_n$, with $\mathbb{E}[X_1] = \mu$ and $\mathsf{Var}(X_1) = \sigma^2$. Then

$$\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2.$$

*Proof.* By the LLN, $\bar{X}_n \xrightarrow{p} \mu$ and $n^{-1}\sum_{i=1}^n X_i^2 \xrightarrow{p} \mathbb{E}[X_1^2] = \mu^2 + \sigma^2$. Furthermore, we can rewrite the sample variance as follows:

$$\begin{aligned}\hat{\sigma}_n^2 &= \frac{1}{n}\sum_{i=1}^n (X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^n X_i^2 - \frac{2}{n}\sum_{i=1}^n X_i\bar{X}_n + \bar{X}_n^2 \\ &= \frac{1}{n}\sum_{i=1}^n X_i^2 - (\bar{X}_n)^2.\end{aligned}$$

The continuous mapping theorem says $-(\bar{X}_n)^2 \xrightarrow{p} -\mu^2$. By the theorem that says a sum of two sequences converging in probability converges in probability, we conclude

$$\hat{\sigma}_n^2 \xrightarrow{p} \mu^2 + \sigma^2 - \mu^2 = \sigma^2.$$

□

**Remark 1.** The version of the sample variance that divides by $n-1$ instead of $n$ also converges to $\sigma^2$, since we can write one as $(n-1)/n$ times the other one, and $(n-1)/n \to 1$ as $n \to \infty$.

## 20.2 Central Limit Theorem

The CLT says more: it tells us not just that $\bar{X}_n$ converges to the population mean $\mu$, but also what the *distribution of* $\bar{X}_n$ is on its way to $\mu$.

**Theorem 2** (Central Limit Theorem)**.** Let $\{X_n\}$ be a sequence of iid random variables with $\mathbb{E}[X_1] = \mu$ and $\mathsf{Var}(X_1) = \sigma^2$. Then

$$\sqrt{\frac{n}{\sigma^2}}\big(\bar{X}_n - \mu\big) \xrightarrow{d} \mathcal{N}(0,1).$$

**Remark 2.** Informally, this result tells us that for large $n$,

$$\bar{X}_n \approx \mathcal{N}\left(\mu, \frac{\sigma^2}{n}\right).$$

Thus the CLT tells us the approximate distribution of $\bar{X}_n$, which is very useful e.g. in hypothesis testing.

Note that it does *not* make sense to say $\bar{X}_n \xrightarrow{d} \mathcal{N}\big(\mu, \sigma^2/n\big)$ as $n \to \infty$! Whenever you take a limit, the thing you are converging to is not allowed to depend on $n$.

In hypothesis testing based on the sample mean, when we also don't know the variance, we reject based on the value of $(\bar{X}_n - \mu)/\hat{\sigma}_n/\sqrt{n}$. We now study the limiting distribution of this test statistic.

**Proposition 2.** Let $X_1, X_2, \ldots$ be i.i.d. with $\mathbb{E}[X_1] = \mu$ and $\mathsf{Var}(X_1) = \sigma^2$. Then

$$\frac{\bar{X}_n - \mu}{(\hat{\sigma}_n/\sqrt{n})} \xrightarrow{d} \mathcal{N}(0,1).$$

*Proof.* The continuous mapping theorem gives $\hat{\sigma}_n \xrightarrow{p} \sigma$ and $\sigma/\hat{\sigma}_n \xrightarrow{p} 1$, using that $\hat{\sigma}_n^2 \xrightarrow{p} \sigma^2$. By the CLT,

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \xrightarrow{d} \mathcal{N}(0,1).$$

Applying Slutsky's theorem to the product gives

$$\frac{\sqrt{n}(\bar{X}_n - \mu)}{\hat{\sigma}_n} = \frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \frac{\sigma}{\hat{\sigma}_n} \xrightarrow{d} \mathcal{N}(0,1).$$

□

**Example 1** (Blood-pressure changes with unknown variance)**.** Return to the motivating example from Lecture 19. Let $X_i$ be the change in blood pressure for patient $i$ after taking a medication, measured in mmHg. Suppose $X_1, \ldots, X_n$ are i.i.d. with unknown distribution, mean $\mu$, and variance $0 < \sigma^2 < \infty$. We test

$$H_0 : \mu = 0 \qquad \text{versus} \qquad H_1 : \mu \neq 0.$$

Here $H_0$ does not specify $\sigma^2$, so we cannot standardize using a known null variance. Instead, use

$$T_n = \frac{\sqrt{n}\,\bar{X}_n}{\hat{\sigma}_n}.$$

By the proposition, under $H_0$ we have $T_n \xrightarrow{d} \mathcal{N}(0,1)$. Thus we reject when

$$|T_n| > z_{\alpha/2}, \qquad \text{equivalently,} \qquad |\bar{X}_n| > \frac{\hat{\sigma}_n}{\sqrt{n}} z_{\alpha/2}.$$

For each fixed population distribution satisfying $H_0$ and the stated assumptions,

$$\mathbb{P}_{H_0}(|T_n| > z_{\alpha/2}) \longrightarrow \mathbb{P}(|Z| > z_{\alpha/2}) = \alpha, \qquad Z \sim \mathcal{N}(0,1).$$

We say that this test has *asymptotic level* $\alpha$.

The CLT is often used in conjunction with the Delta method.

**Theorem 3** (Delta method)**.** Let $\{Y_n\}$ be a sequence of random variables such that

$$\sqrt{n}\big(Y_n - \theta\big) \xrightarrow{d} \mathcal{N}\big(0,\, v\big).$$

If $g$ is a differentiable function at $\theta$ with derivative $g'(\theta) \neq 0$, then

$$\sqrt{n}\Big(g(Y_n) - g(\theta)\Big) \xrightarrow{d} \mathcal{N}\big(0,\ [g'(\theta)]^2 v\big).$$

Thus for example, if $Y_n = \bar{X}_n$ and $g(y) = y^2$, then

$$\sqrt{n}\Big((\bar{X}_n)^2 - \mu^2\Big) \xrightarrow{d} \mathcal{N}\big(0,\ 4\mu^2\sigma^2\big).$$