39. Limit Theorems

The Central Limit Theorem (CLT) [Ross 8.3]

Proposition 39.1 *The Central Limit Theorem*

Let $X_1, X_2, \ldots$ be a sequence of iid random variables having mean μ and variance σ^2. Then, the distribution of

$$\begin{aligned} Z_n &= \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma} \\ &= \frac{X_1 + X_2 + \cdots + X_n - n\mu}{\sigma\sqrt{n}} \\ &= \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \qquad \textit{where } \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i \end{aligned}$$

tends to the standard normal as $n \to \infty$. Specifically,

$$P[Z_n \leq a] \to \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{a} e^{-u^2/2} du}_{\Phi(a)} \qquad \textit{as } n \to \infty$$

[See textbook Section 8.3 for proof]

The CLT can be used to approximate probabilities:

Example 39.1: An astronomer takes iid measurements $X_1, X_2, \ldots$ of the distance of a star.

Each X_i has mean μ (the true distance) and variance $\sigma^2 = 4$ light-years^2.

How many measurements are needed to be 95% certain that the sample average $\bar{X}$ of the measurements is within ± 0.5 light-years of the true value μ?

Solution: Let

$$Z_n = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \tag{39.1}$$

By the CLT, when n is large, this is approximately $\mathcal{N}(0, 1)$.

$$\begin{aligned}
P\left[-0.5 \leq \bar{X} - \mu \leq 0.5\right] &= P\left[\frac{-0.5}{\sigma/\sqrt{n}} \leq \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq \frac{0.5}{\sigma/\sqrt{n}}\right] \\
&= P\left[-\frac{\sqrt{n}}{2\sigma} \leq Z_n \leq \frac{\sqrt{n}}{2\sigma}\right] \\
&\approx \Phi\left(\frac{\sqrt{n}}{4}\right) - \Phi\left(-\frac{\sqrt{n}}{4}\right) \\
&= 2\Phi\left(\frac{\sqrt{n}}{4}\right) - 1
\end{aligned}$$

For this to be at least 0.95, we need

$$\Phi\left(\frac{\sqrt{n}}{4}\right) \geq 0.975$$

From the $\Phi(.)$ Table [Notes #18], $\sqrt{n}/4 \geq 1.96$.

The smallest integer than makes this true is $n = 62$.

Note: This analysis assumes that with 62 observations, Z_n is well approximated by a Gaussian. In statistics, $n \geq 30$ is often assumed to be enough.

The Chebyshev inequality is not an approximation.

$$E\left[\bar{X}\right] = \mu \qquad Var\left[\bar{X}\right] = \frac{\sigma^2}{n}$$

So by Chebyshev:

$$P\left[\left|\bar{X}-\mu\right| \geq 0.5\right] \leq \frac{\sigma^2/n}{(0.5)^2} = \frac{16}{n}$$

95% confident $\Rightarrow 16/n \leq 0.05 \Rightarrow n \geq 320$ measurements are enough.

What if we don't know σ^2?

Use sample variance $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ instead of σ^2 in (39.1) to get

$$T_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n}\frac{X_i-\mu}{S} = \frac{\bar{X}-\mu}{S/\sqrt{n}} \tag{39.2}$$

<u>If</u> $X \sim \mathcal{N}(\mu, \sigma^2)$, then T_n has **Student-t** distribution with $n-1$ degrees of freedom:

- pdf of T_n depends on n but not μ or σ
- pdf of T_n similar to $\mathcal{N}(0,1)$ but has heavier tails
- as $n \to \infty$, pdf of T_n converges to that of $\mathcal{N}(0,1)$

For $\alpha \in (0,1)$, let t_α be such that $P[T_n \geq t_\alpha] = \alpha$. Then:

$$1-\alpha = P[-t_{\alpha/2} < T_n < t_\alpha/2] \tag{39.3}$$

$$= P[-t_{\alpha/2} < \frac{\bar{X}-\mu}{S/\sqrt{n}} < t_\alpha/2] \tag{39.4}$$

$$= P[-t_{\alpha/2}\frac{S}{\sqrt{n}} < \bar{X}-\mu < t_{\alpha/2}\frac{S}{\sqrt{n}}] \tag{39.5}$$

So $\bar{X}$ is within $\pm t_{\alpha/2}\frac{S}{\sqrt{n}}$ of μ with prob. $1-\alpha$.

Note: the width of the interval is now random and depends on S.

t_α is well-tabulated. With $\alpha = 0.05$:

n	$t_{\alpha/2}$
16	2.131
30	2.045
40	2.021
62	2.000
120	1.980
1000	1.962

If $n \to \infty$ then $t_{\alpha/2} \to z_{\alpha/2} = 1.960$

Large Sample Confidence Interval: If X_i are not Gaussian or σ not known, when n large enough (e.g., 30, 40 or 120), replace σ with S and use $z_{\alpha/2}$.

Strong Law of Large Numbers [Ross S8.4]

We saw earlier the *weak* law of large numbers. This suggests that there is a strong law of large numbers as well (and there is).

Proposition 39.2 *Strong Law of Large Numbers*

Let $X_1, X_2, \ldots$ be iid with commom mean $E[X_i] = \mu$. Then

$$P\left[\lim_{n\to\infty} \frac{X_1 + X_2 + \cdots + X_n}{n} = \mu\right] = 1$$