Convergence of Random Variables

Types of Convergence

Convergence in Probability

$X_n \xrightarrow{\mathbb{P}} X$ as $n \to \infty$ if for every $\varepsilon > 0$, we have:

$$P(|X_n - X| > \varepsilon|) \to 0 \text{ as } n \to \infty$$

Suppose $X_n \sim \text{Ber}(1/2)$. We will check if $X_n \xrightarrow{\mathbb{P}} \text{Ber}(1/2)$.

For $\varepsilon \in (0, 1)$, the event $\{|X_n - X| > \varepsilon\}$ is equivalent to $\{|X_n - X| = 1\}$. This occurs with probability $1/2$, which does not go to 0, so this does not converge in probability!

Convergence in Distribution

$X_n \rightsquigarrow X$ as $n \to \infty$ if:

$$P(X_n \leq x) \to P(X \leq x) \text{ as } n \to \infty$$

for all x at which the cdf for X is continuous

For the above example, we do have $X_n \rightsquigarrow \text{Ber}(1/2)$

Intuition:

- Converge in probability: focus on as n increases, probability that X_n deviates significantly from X is low
- Converge in distribution: focus on as n increases, the distributions are the same

Relationship Between Convergence

If $X_n \xrightarrow{\mathbb{P}} X$ then $X_n \rightsquigarrow X$

Convergence to a Constant

If $X_n \rightsquigarrow c$ for some deterministic constant c, then $X_n \xrightarrow{\mathbb{P}} c$

Convergence in Probability of Sums and Products

If $X_n \xrightarrow{\mathbb{P}} X$ and $Y_n \xrightarrow{\mathbb{P}} Y$ then:

$$X_n + Y_n \xrightarrow{\mathbb{P}} X + Y$$
$$X_n Y_n \xrightarrow{\mathbb{P}} XY$$

Slutsky's Theorem

If $X_n \rightsquigarrow X$ and $Y_n \rightsquigarrow c$ for a deterministic constant c then:

$$X_n + Y_n \rightsquigarrow X + c$$
$$X_n Y_n \rightsquigarrow Xc$$

For general Y_n, this does not hold. In fact, consider the example where $X_n \sim \mathcal{N}(0, 1)$ and $Y_n = -X_n$

- Because normal variables are symmetric, $Y_n \sim \mathcal{N}(0, 1)$
- $X_n + Y_n = 0$
- But we could choose $Y = X$ as our limits (instead of $Y = -X$)
- Then $X_n + Y_n$ does not converge to $2X$

Intuition:

- From just the descriptions of X_n and Y_n individually, we can describe the distribution of the limits individually
- But we don't have information abotu their correlation
- We need this information to determine the distribution of sum of limits

Properties

Continuous Mapping Theorem

If $X_n \xrightarrow{\mathbb{P}} X$ then for continuous functions g, we have:

$$g(X_n) \xrightarrow{\mathbb{P}} g(X)$$

Similar statement can be made for convergence in distribution.

Delta Method

Let Y_n be a sequence of random variables such that:

$$\frac{Y_n - \mu}{\sigma/\sqrt{n}} \rightsquigarrow Y \sim \mathcal{N}(0,1)$$

Then, for any differentiable g such that $g'(\mu) \neq 0$, we have:

$$\frac{g(Y_n) - g(\mu)}{\sigma/\sqrt{n}} \rightsquigarrow \mathcal{N}(0, g'(\mu)^2)$$

This is typically applied to Y_n being a sample average of n samples.

Proof Sketch:

- First order Taylor expand g around μ:
  - $g(Y_n) - g(\mu) = g'(\mu)(Y_n - \mu)$
- Divide both sides by $\sigma/\sqrt{n}$ and note the RHS converges to $g'(\mu)\mathcal{N}(0,1)$, which is equivalent to $\mathcal{N}(0, g'(\mu)^2)$

Intuition:

- If we have some estimator Y_n for μ such that Y_n is normally distributed, then we can say that $g(Y_n)$ is also normally distributed with a specific variance

Slutsky's Theorem Example

When we only know the sample std dev and not the actual std dev and want to make a statistical test with CLT, we may be tempted to just replace the population std dev with the sample one

I.e. we have $\frac{\overline{X_n}-\mu}{\sigma/\sqrt{n}} \approx \mathcal{N}(0,1)$ but we don't know σ.

We do know $\hat{\sigma}$, and the LLN tells us $\sigma/\hat{\sigma} \xrightarrow{\mathbb{P}} 1$ since the sample variance is unbiased estimator.

Then Slutsky's Theorem tells us that multiplying this with the above gives us:

$$\frac{\overline{X_n} - \mu}{\hat{\sigma}/\sqrt{n}} \approx \mathcal{N}(0,1)$$