Convergence in Distribution

Dr. Syring

March 2, 2020

Definition

Let F_{X_n} denote the sequence of CDFs of a sequence of random variables X_n and let F_X denote the CDF of a random variable X. We say that X_n converges in distribution to X if

$$\lim_{n \to \infty} F_{X_n}(x) = F_X(x)$$

at every point x where F is continuous.

The issue of continuity is mostly a technical one to avoid problems with situations like the following. Suppose $P(X_n = 1/n) = 1$ and $P(X = 0) = 1$. Then, it certainly feels like X_n converges in distribution to X but $F_{X_n}(0) = 0$ for every n , so the limit is not equal to $F_X(0)$. But, 0 is a point of discontinuity of F_X , and F_X agrees with F_{X_n} , in the limit, at every other point (the points of continuity). This is why we make this technical point part of our definition.

Central Limit Theorem: very short proof

We will give a short proof of the CLT using moment generating funtions. Details are also in Theorem 5.3.1 in the textbook.

Consider iid random variables $X_1, X_2, \ldots, X_n$ $n \to \infty$ with moment generating functions $M_X(t)$ that exist in a neighborhood of 0. Then

$$\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} \to Z$$

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

Proof:
First note that

$$Z_n := \sqrt{n}\frac{\bar{X}_n - \mu}{\sigma} = \frac{1}{\sqrt{n}}\sum_{i=1}^{\infty}\left(\frac{X_i - \mu}{\sigma}\right) =: \frac{1}{\sqrt{n}}\sum_{i=1}^{\infty} Z_i$$

Second, note the MGF of $\sqrt{n}\frac{\bar{X}_n - \mu}{\sigma}$ can be written (using independence)

$$M_{Z_n}(s) = Ee^{sZ_n} = Ee^{s\frac{1}{\sqrt{n}}\sum_{i=1}^{\infty}\left(\frac{X_i-\mu}{\sigma}\right)} = \prod_{i=1}^{n} Ee^{s\frac{1}{\sqrt{n}}\left(\frac{X_i-\mu}{\sigma}\right)} = \prod_{i=1}^{n} M_{Z_i}(t/\sqrt{n})$$

where $t = s$. Finally, take a 2 term Taylor expansion of $M_{Z_i}(t/\sqrt{n})$ around zero making note that $M_{Z_i}(0) = 1$, $M_{Z_i}^{'}(0) = 0$, $M_{Z_i}^{''}(0) = 1$, and $\exists \eta \in (0, t/\sqrt{n})$ s.t.

$$M_{Z_i}(t/\sqrt{n}) = M_{Z_i}(0) + M_{Z_i}^{'}(0)t/\sqrt{n} + M_{Z_i}^{''}(\eta)\frac{t^2}{2n}$$

$$= 1 + M_{Z_i}^{''}(\eta)\frac{t^2}{2n}$$

Then,

$$M_{Z_n}(s) = \left(1 + M_{Z_i}^{''}(\eta)\frac{t^2}{2n}\right)^n \to e^{t^2/2}$$

as $n \to \infty$ because $M_{Z_i}^{''}(\eta) \to 1$.

More advanced versions of the proof do not require moment generating functions. Recall that the existence of the MGF implies the tails of the distribution are sufficiently light. A weaker but similar condition would restrict the variances of the random variables; see Lindeberg-Feller CLT.

Relationship to convergence in probability

Look at $F_{X_n}(x)$ at a point of continuity x . By definition $F_{X_n}(x) = P(X_n \leq x)$. By law of total probability

$$P(X_n \leq x) = P(X_n \leq x \cap \ |X_n - X| \ \leq \epsilon) + P(X_n \leq x \cap \ |X_n - X| \ > \epsilon).$$

If $X < x + \epsilon$ then the first event on the RHS definitely happens, so its probability is bounded by $P(X < x + \epsilon)$. If $|X_n - X| \ > \epsilon$ happens then the second event on the RHS happens so its probability is bounded by $P(|X_n - X > \ |\epsilon)$, which limits to zero if $X_n \xrightarrow{i.p.} X$. So, we see that

$$\limsup_{n \to \infty} F_{X_n}(x) \leq F(x + \epsilon)$$

The same sort of argument says

$$\liminf_{n \to \infty} F_{X_n}(x) \geq F(x - \epsilon)$$

And, since ϵ is arbitrary the limit exists and X_n converges to X in distribution.

Rates of Convergence

From the CLT we have

$$\sqrt{n}\frac{\bar{X}-\mu}{\sigma} \to N(0, 1)$$

and the decay of the standard deviation at rate $n^{-1/2}$ is a kind of "rate" of convergence. For example, let a_n be any divergent sequence (like $\log\log n$)

$$P(|\bar{X}_n - \mu| > a_n n^{-1/2}) \to 0$$

Since a_n is arbitrarily slow we can say the rate of convergence is $n^{-1/2}$.

The rate is not always $n^{-1/2}$. Suppose we estimate the "location" of a distribution by way of a "modal interval." Take the center of shortest interval containing half the observations as your measure of location. This estimator typically has rate of convergence $n^{-1/3}$.

In practice, rates of convergence are an intermediary result between convergence in probability to a point (consistency) and convergence in distribution (which can be used to construct tests/CIs).

Examples

1. Let X_n have CDF $F_{X_n}(x) = x - \frac{\sin(2n\pi x)}{2n\pi}$ on $x \in (0, 1)$. The limit is a uniform CDF.

2. Let X_n be the maximum of $Unif(0, \theta)$. The limit is $Exp(\theta)$.

3. Let $F_{X_n}(x) = 1 - (1 - x)^n$ for $0 \leq x \leq 1$. Limit is the point mass at zero. Beware of discontinuity.

4. Let $F_{X_n}(x) = \frac{e^{nx}}{1 + e^{nx}}$ for $x \in \mathrm{R}$. Limit is the point mass at zero.

Moment generating function technique

Theorem 5.2.10: If the MGFs converge, i.e. $\lim_{n\to\infty} M_{X_n}(t) = M_X(t)$ then X_n converges to X in distribution.

This was our approach to CLT.

Common situation: Often times we can express $M_{X_n}(t)$ in terms of a Taylor expansion and the limit is

$$\lim_{n\to\infty} [1 + b/n + \psi(n)/n]^{cn} = e^{bc}$$

if the function $\psi(n)$ has limit zero.

Application:

Let $Z_n \sim \chi^2(n)$, then it has mean n and variance $2n$. Let $Y_n = \frac{Z_n - n}{\sqrt{2n}}$ be the standardized sequence. Then, Y_n converges to standard normal.

Delta Method

Suppose $Z_n := \sqrt{n}(X_n - \theta)/\sigma$ converges to standard normal. First, this implies that Z_n is "bounded in probability", i.e.

$$P(|Z_n| > B_\epsilon) < \epsilon$$

for all large enough n . Next, consider a differentiable transformation of X_n , $g(X_n)$. What is the limiting distribution of $g(X_n)$?

A Taylor expansion provides

$$\sqrt{n}g(X_n) = \sqrt{n}g(\theta) + \sqrt{n}g'(\theta)(X_n - \theta) + o_p(\sqrt{n}|X_n - \theta|)$$

and this last expression says $\frac{o_p(\sqrt{n}|X_n - \theta|)}{\sqrt{n}|X_n - \theta|} \rightarrow 0$ Theorem 5.2.8.

Then, rearranging we have

$$\frac{\sqrt{n}[g(X_n) - g(\theta)]}{\sigma g'(\theta)^2} = \sqrt{n}|X_n - \theta| + \delta$$

where δ convergest to zero in probability. Then, $\frac{\sqrt{n}[g(X_n) - g(\theta)]}{\sigma g'(\theta)^2}$ converges to $\sqrt{n}|X_n - \theta|$ in probability, and therefore also in distribution.