WIKIPEDIA

Continuous mapping theorem

In probability theory, the **continuous mapping theorem** states that continuous functions preserve limits even if their arguments are sequences of random variables. A continuous function, in Heine's definition, is such a function that maps convergent sequences into convergent sequences: if $x_n \to x$ then $g(x_n) \to g(x)$. The *continuous mapping theorem* states that this will also be true if we replace the deterministic sequence $\{x_n\}$ with a sequence of random variables $\{X_n\}$, and replace the standard notion of convergence of real numbers "→" with one of the types of convergence of random variables.

This theorem was first proved by Henry Mann and Abraham Wald in 1943,[1] and it is therefore sometimes called the **Mann–Wald theorem**.[2] Meanwhile, Denis Sargan refers to it as the **general transformation theorem**.[3]

Contents

Statement

Proof

- Convergence in distribution
- Convergence in probability
- Almost sure convergence

See also

References

Statement

Let $\{X_n\}$, X be random elements defined on a metric space S. Suppose a function g: $S \to S'$ (where S' is another metric space) has the set of discontinuity points D_g such that $\Pr[X \in D_g] = 0$. Then[4][5]

$$
\begin{aligned}
X_n &\xrightarrow{\text{d}} X \quad &\Rightarrow \quad g(X_n) &\xrightarrow{\text{d}} g(X);\\
X_n &\xrightarrow{\text{p}} X \quad &\Rightarrow \quad g(X_n) &\xrightarrow{\text{p}} g(X);\\
X_n &\xrightarrow{\text{a.s.}} X \quad &\Rightarrow \quad g(X_n) &\xrightarrow{\text{a.s.}} g(X).
\end{aligned}
$$

where the superscripts, "d", "p", and "a.s." denote convergence in distribution, convergence in probability, and almost sure convergence respectively.

Proof

This proof has been adopted from (van der Vaart 1998, Theorem 2.3)

Spaces S and S' are equipped with certain metrics. For simplicity we will denote both of these metrics using the $|x - y|$ notation, even though the metrics may be arbitrary and not necessarily Euclidean.

Convergence in distribution

We will need a particular statement from the portmanteau theorem: that convergence in distribution $X_n \xrightarrow{d} X$ is equivalent to

$$\mathbb{E}f(X_n) \to \mathbb{E}f(X) \text{ for every bounded continuous functional } f.$$

So it suffices to prove that $\mathbb{E}f(g(X_n)) \to \mathbb{E}f(g(X))$ for every bounded continuous functional f. Note that $F = f \circ g$ is itself a bounded continuous functional. And so the claim follows from the statement above.

Convergence in probability

Fix an arbitrary $\varepsilon > 0$. Then for any $\delta > 0$ consider the set B_δ defined as

$$B_\delta = \big\{x \in S \mid x \notin D_g : \ \exists y \in S : \ |x - y| < \delta, \ |g(x) - g(y)| > \varepsilon\big\}.$$

This is the set of continuity points x of the function $g(\cdot)$ for which it is possible to find, within the δ-neighborhood of x, a point which maps outside the ε-neighborhood of $g(x)$. By definition of continuity, this set shrinks as δ goes to zero, so that $\lim_{\delta \to 0} B_\delta = \emptyset$.

Now suppose that $|g(X) - g(X_n)| > \varepsilon$. This implies that at least one of the following is true: either $|X - X_n| \geq \delta$, or $X \in D_g$, or $X \in B_\delta$. In terms of probabilities this can be written as

$$\Pr\big(\big|g(X_n) - g(X)\big| > \varepsilon\big) \leq \Pr\big(|X_n - X| \geq \delta\big) + \Pr(X \in B_\delta) + \Pr(X \in D_g).$$

On the right-hand side, the first term converges to zero as $n \to \infty$ for any fixed δ, by the definition of convergence in probability of the sequence $\{X_n\}$. The second term converges to zero as $\delta \to 0$, since the set B_δ shrinks to an empty set. And the last term is identically equal to zero by assumption of the theorem. Therefore, the conclusion is that

$$\lim_{n\to\infty} \Pr\big(\big|g(X_n) - g(X)\big| > \varepsilon\big) = 0,$$

which means that $g(X_n)$ converges to $g(X)$ in probability.

Almost sure convergence

By definition of the continuity of the function $g(\cdot)$,

$$\lim_{n\to\infty} X_n(\omega) = X(\omega) \quad \Rightarrow \quad \lim_{n\to\infty} g(X_n(\omega)) = g(X(\omega))$$

at each point $X(\omega)$ where $g(\cdot)$ is continuous. Therefore,

$$\begin{aligned} \Pr\Big(\lim_{n\to\infty} g(X_n) = g(X)\Big) &\geq \Pr\Big(\lim_{n\to\infty} g(X_n) = g(X), \ X \notin D_g\Big) \\ &\geq \Pr\Big(\lim_{n\to\infty} X_n = X, \ X \notin D_g\Big) = 1, \end{aligned}$$

because the intersection of two almost sure events is almost sure.

By definition, we conclude that $g(X_n)$ converges to $g(X)$ almost surely.

See also

- Slutsky's theorem
- Portmanteau theorem

References

1. Mann, H. B.; Wald, A. (1943). "On Stochastic Limit and Order Relationships". *Annals of Mathematical Statistics*. **14** (3): 217–226. doi:10.1214/aoms/1177731415 (https://doi.org/10.1214%2Faoms%2F1177731415). JSTOR 2235800 (https://www.jstor.org/stable/2235800).
2. Amemiya, Takeshi (1985). *Advanced Econometrics* (https://books.google.com/books?id=0bzGQE14CwEC&pg=pA88). Cambridge, MA: Harvard University Press. p. 88. ISBN 0-674-00560-0.
3. Sargan, Denis (1988). *Lectures on Advanced Econometric Theory*. Oxford: Basil Blackwell. pp. 4–8. ISBN 0-631-14956-2.
4. Billingsley, Patrick (1969). *Convergence of Probability Measures*. John Wiley & Sons. p. 31 (Corollary 1). ISBN 0-471-07242-7.
5. Van der Vaart, A. W. (1998). *Asymptotic Statistics* (https://books.google.com/books?id=UEuQEM5RjWgC&pg=PA7). New York: Cambridge University Press. p. 7 (Theorem 2.3). ISBN 0-521-49603-9.