Asymptotic theory

Stochastic convergence

Stochastic convergence

Let $x_1, x_2, \ldots$ be a sequence of a scalar random variable that converge to a random variable x. Both x_N and x are defined on probability space $(\Omega, \mathcal{F}, Pr)$. Let us denote the distribution functions of x_N and x by F_N and F respectively.

Remark: x may be a fixed, real number. It is then a degenerated random variable that takes on a given value with a probability one.

Convergence in probability

The sequence x_N converges in probability to x if for every $\epsilon > 0$

$$\lim_{N\to\infty} Pr(|x_N - x| < \epsilon) = 1$$

It is abbreviated as

$$p\lim x_N = x$$

or

$$x_N \to^p x$$

Convergence almost surely

The sequence x_N converges almost surely or with a probability one to x if for every $\epsilon > 0$

$$Pr(\lim_{N\to\infty} |x_N - x| < \epsilon) = 1$$

It is abbreviated as

$$x_N \to^{a.s.} x$$

Sometimes it is called a strong convergence.

Convergence in quadratic mean

The sequence x_N converges in quadratic mean to x if

$$\lim_{N\to\infty} E(x_N - x)^2 = 0$$

It is abbreviated as

$$x_N \to^{q.m.} x$$

Convergence in distribution

The sequence x_N converges in distribution to x if for every real number c

$$\lim_{N\to\infty} F_N(c) = F(c)$$

It is abbreviated as

$$x_N \to^d x$$

Sometimes it is called a weak convergence.

Remark: It does not require convergence of p.d.f.s

Convergence properties

Convergence properties

1. $x_N \to^{a.s.} x \Rightarrow x_N \to^{p} x \Rightarrow x_N \to^{d} x$
2. $x_N \to^{q.m.} x \Rightarrow x_N \to^{p} x \Rightarrow x_N \to^{d} x$
3. Suppose x is fixed, nonstochastic then
   - $x_N \to^{q.m.} x \Rightarrow \lim E(x_N) = x$
   - $x_N \to^{q.m.} x \Rightarrow \lim Var(x_N) = \lim E(x_N - Ex_N)^2 = 0$
   - $x_N \to^{p} x \Leftrightarrow x_N \to^{d} x$

Convergence properties

Convergence properties

1. $x_N \rightarrow^{a.s.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
2. $x_N \rightarrow^{q.m.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
3. Suppose x is fixed, nonstochastic then
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim E(x_N) = x$
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim Var(x_N) = \lim E(x_N - Ex_N)^2 = 0$
   - $x_N \rightarrow^{p} x \Leftrightarrow x_N \rightarrow^{d} x$

Convergence properties

Convergence properties

1. $x_N \rightarrow^{a.s.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
2. $x_N \rightarrow^{q.m.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
3. Suppose x is fixed, nonstochastic then
    - $x_N \rightarrow^{q.m.} x \Rightarrow \lim E(x_N) = x$
    - $x_N \rightarrow^{q.m.} x \Rightarrow \lim Var(x_N) = \lim E(x_N - Ex_N)^2 = 0$
    - $x_N \rightarrow^{p} x \Leftrightarrow x_N \rightarrow^{d} x$

Convergence properties

Convergence properties

1. $x_N \rightarrow^{a.s.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
2. $x_N \rightarrow^{q.m.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
3. Suppose x is fixed, nonstochastic then
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim E(x_N) = x$
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim Var(x_N) = \lim E(x_N - Ex_N)^2 = 0$
   - $x_N \rightarrow^{p} x \Leftrightarrow x_N \rightarrow^{d} x$

Convergence properties

Convergence properties

1. $x_N \rightarrow^{a.s.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
2. $x_N \rightarrow^{q.m.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
3. Suppose x is fixed, nonstochastic then
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim E(x_N) = x$
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim Var(x_N) = \lim E(x_N - Ex_N)^2 = 0$
   - $x_N \rightarrow^{p} x \Leftrightarrow x_N \rightarrow^{d} x$

Convergence properties

Convergence properties

1. $x_N \rightarrow^{a.s.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
2. $x_N \rightarrow^{q.m.} x \Rightarrow x_N \rightarrow^{p} x \Rightarrow x_N \rightarrow^{d} x$
3. Suppose x is fixed, nonstochastic then
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim E(x_N) = x$
   - $x_N \rightarrow^{q.m.} x \Rightarrow \lim Var(x_N) = \lim E(x_N - Ex_N)^2 = 0$
   - $x_N \rightarrow^{p} x \Leftrightarrow x_N \rightarrow^{d} x$

Convergence properties

Slutsk'y theorem

If $g : R \rightarrow R$ is a continuous functions then

- $x_N \rightarrow^p x \Rightarrow g(x_N) \rightarrow^p g(x)$
- $x_N \rightarrow^d x \Rightarrow g(x_N) \rightarrow^d g(x)$
- $x_N \rightarrow^{a.s.} x \Rightarrow g(x_N) \rightarrow^{a.s.} g(x)$

where

$$\lim g(x_N) = g(\lim x_N)$$

Asymptotic theory

Laws of Large Number and Central Limit Theorem

Laws of large numbers

Khinchine's Theorem

Let x_n be a sequence of i.i.d. random variables with $E(x_n) = \mu < \infty$. Then

$$\bar{x}_N := \frac{1}{N}\sum_{n=1}^{N} x_n \to^p \mu$$

Laws of large numbers

Chebyshev's Theorem

Let x_n be a sequence of independent random variables with $E(x_n) = \mu < \infty$ and $Var(x_n) \leq c < \infty$ for some finite constant c. Then

$$\bar{x}_N \to^p \mu$$

Central Limit Theorem

Lindberg-Levy CLT

Let x_n be a sequence of K-dimensional i.i.d. random variables with a mean μ and a covariance matrix Σ. Then

$$\sqrt{N}(\bar{x_N} - \mu) \rightarrow^d N(0, \Sigma)$$

Asymptotic theory

Asymptotic properties of estimators

Properties of estimators

Properties of an estimator $\hat{\beta}$:

- $\hat{\beta}$ is unbias estimator of β if $E(\hat{\beta}) = \beta$
- $\hat{\beta}$ is (weakly) consistent if $\hat{\beta}_N \to^p \beta$
- $\hat{\beta}$ is strongly consistent if $\hat{\beta}_N \to^{a.s.} \beta$
- $\hat{\beta}$ is asymptotically normal if

$$\sqrt{N}(\hat{\beta} - \beta) \to^d N(0, \Sigma)$$

Properties of estimators

Properties of an estimator $\hat{\beta}$:

- $\hat{\beta}$ is unbias estimator of β if $E(\hat{\beta}) = \beta$
- $\hat{\beta}$ is (weakly) consistent if $\hat{\beta}_N \rightarrow^p \beta$
- $\hat{\beta}$ is strongly consistent if $\hat{\beta}_N \rightarrow^{a.s.} \beta$
- $\hat{\beta}$ is asymptotically normal if

$$\sqrt{N}(\hat{\beta} - \beta) \rightarrow^d N(0, \Sigma)$$

Properties of estimators

Properties of an estimator $\hat{\beta}$:

- $\hat{\beta}$ is unbias estimator of β if $E(\hat{\beta}) = \beta$
- $\hat{\beta}$ is (weakly) consistent if $\hat{\beta}_N \rightarrow^p \beta$
- $\hat{\beta}$ is strongly consistent if $\hat{\beta}_N \rightarrow^{a.s.} \beta$
- $\hat{\beta}$ is asymptotically normal if

$$\sqrt{N}(\hat{\beta} - \beta) \rightarrow^d N(0, \Sigma)$$

Properties of estimators

Properties of an estimator $\hat{\beta}$:

- $\hat{\beta}$ is unbias estimator of β if $E(\hat{\beta}) = \beta$
- $\hat{\beta}$ is (weakly) consistent if $\hat{\beta}_N \to^p \beta$
- $\hat{\beta}$ is strongly consistent if $\hat{\beta}_N \to^{a.s.} \beta$
- $\hat{\beta}$ is asymptotically normal if

$$\sqrt{N}(\hat{\beta} - \beta) \to^d N(0, \Sigma)$$

Properties of estimators

Properties of an estimator $\hat{\beta}$:

- $\hat{\beta}$ is unbias estimator of β if $E(\hat{\beta}) = \beta$
- $\hat{\beta}$ is (weakly) consistent if $\hat{\beta}_N \to^p \beta$
- $\hat{\beta}$ is strongly consistent if $\hat{\beta}_N \to^{a.s.} \beta$
- $\hat{\beta}$ is asymptotically normal if

$$\sqrt{N}(\hat{\beta} - \beta) \to^d N(0, \Sigma)$$

Linear combination of estimators

Suppose $\hat{\beta}$ is an estimator of a $(K \times 1)$ vector β with

$$\sqrt{N}(\hat{\beta} - \beta) \rightarrow^{d} N(0, \Sigma)$$

Then for $A \neq 0$ the asymptotic distribution of $\sqrt{N}A(\hat{\beta} - \beta)$ is

$$\sqrt{N}A(\hat{\beta} - \beta) \rightarrow^{d} N(0, A\Sigma A')$$

Remark: Also if $p\lim \hat{A} = A$ then

$$\sqrt{N}\hat{A}(\hat{\beta} - \beta) \rightarrow^{d} N(0, A\Sigma A')$$

Linear combination of estimators

Suppose $\hat{\beta}$ is an estimator of a $(K \times 1)$ vector β with

$$\sqrt{N}(\hat{\beta} - \beta) \to^{d} N(0, \Sigma)$$

Then for $A \neq 0$ the asymptotic distribution of $\sqrt{N}A(\hat{\beta} - \beta)$ is

$$\sqrt{N}A(\hat{\beta} - \beta) \to^{d} N(0, A\Sigma A')$$

Remark: Also if $p\lim \hat{A} = A$ then

$$\sqrt{N}\hat{A}(\hat{\beta} - \beta) \to^{d} N(0, A\Sigma A')$$

Example 1

Suppose $K = 2$ and $\beta = (2, 2)'$ and

$$\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix}$$

What is the asymptotic distribution of $\hat{\beta}_1 - \hat{\beta}_2$?

We know that

$$\hat{\beta}_1 - \hat{\beta}_2 = [1, -1] \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = [1, -1]\hat{\beta}$$

Example 1

Suppose $K = 2$ and $\beta = (2, 2)'$ and

$$\Sigma = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix}$$

What is the asymptotic distribution of $\hat{\beta}_1 - \hat{\beta}_2$?

We know that

$$\hat{\beta}_1 - \hat{\beta}_2 = [1, -1] \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \end{bmatrix} = [1, -1]\hat{\beta}$$

Example 1

Therefore

$$\sqrt{N}[1,-1](\hat{\beta}-\beta) \to^d N\left(0, [1,-1]\begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$$

Since $[1,-1]\beta = [1,-1](2,2)' = 0$ and

$$[1,-1]\begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 3$$

Then $[1,-1](\hat{\beta}-\beta) = [1,-1]\hat{\beta}$ and

$$\sqrt{N}(\hat{\beta}_1 - \hat{\beta}_2) \to^d N(0,3)$$

Example 1

Therefore

$$\sqrt{N}[1,-1](\hat{\beta}-\beta) \rightarrow^d N\left(0,[1,-1]\begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right)$$

Since $[1,-1]\beta = [1,-1](2,2)' = 0$ and

$$[1,-1]\begin{bmatrix} 1 & 0.5 \\ 0.5 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 3$$

Then $[1,-1](\hat{\beta}-\beta) = [1,-1]\hat{\beta}$ and

$$\sqrt{N}(\hat{\beta}_1-\hat{\beta}_2) \rightarrow^d N(0,3)$$

Example 1

Therefore

$$\sqrt{N}[1,-1](\hat{\beta}-\beta) \rightarrow^{d} N\left(0,[1,-1]\left[\begin{array}{cc} 1 & 0.5 \\ 0.5 & 3 \end{array}\right]\left[\begin{array}{c} 1 \\ -1 \end{array}\right]\right)$$

Since $[1,-1]\beta = [1,-1](2,2)' = 0$ and

$$[1,-1]\left[\begin{array}{cc} 1 & 0.5 \\ 0.5 & 3 \end{array}\right]\left[\begin{array}{c} 1 \\ -1 \end{array}\right] = 3$$

Then $[1,-1](\hat{\beta}-\beta) = [1,-1]\hat{\beta}$ and

$$\sqrt{N}(\hat{\beta}_1-\hat{\beta}_2) \rightarrow^{d} N(0,3)$$

Delta method

Suppose $g(\beta) = (g_1(\beta), ..., g_m(\beta))'$ is a vector-valued continuously differentiable function with $\partial g(\beta)/\partial\beta' \neq 0$ at β, then

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) \to^d N(0, \frac{\partial g(\beta)}{\partial\beta'}\Sigma\frac{\partial g(\beta)'}{\partial\beta})$$

Remark: if $\partial g(\beta)/\partial\beta' = 0$ then

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) \to^p 0$$

Delta method

Suppose $g(\beta) = (g_1(\beta), ..., g_m(\beta))'$ is a vector-valued continuously differentiable function with $\partial g(\beta)/\partial \beta' \neq 0$ at β, then

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) \rightarrow^d N(0, \frac{\partial g(\beta)}{\partial \beta'} \Sigma \frac{\partial g(\beta)'}{\partial \beta})$$

Remark: if $\partial g(\beta)/\partial \beta' = 0$ then

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) \rightarrow^p 0$$

Example 2

Suppose the true parameter $\beta = 1$ is a single valued scalar and What is an asymptotic distribution of $g(\hat{\beta}) = \hat{\beta}^3 + \hat{\beta}^2 - 2$?

Let us notice that

$$g(\beta) = 1^3 + 1^2 - 2 = 0$$

and

$$\partial g(\beta)/\partial\beta = 3\beta^2 + 2\beta = 5$$

Therefore,

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) = \sqrt{N}g(\hat{\beta}) \to^d N(0, 5 \cdot 2 \cdot 5) = N(0, 50)$$

Example 2

Suppose the true parameter $\beta = 1$ is a single valued scalar and What is an asymptotic distribution of $g(\hat{\beta}) = \hat{\beta}^3 + \hat{\beta}^2 - 2$?

Let us notice that

$$g(\beta) = 1^3 + 1^2 - 2 = 0$$

and

$$\partial g(\beta)/\partial\beta = 3\beta^2 + 2\beta = 5$$

Therefore,

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) = \sqrt{N}g(\hat{\beta}) \to^d N(0, 5 \cdot 2 \cdot 5) = N(0, 50)$$

Example 2

Suppose the true parameter $\beta = 1$ is a single valued scalar and What is an asymptotic distribution of $g(\hat{\beta}) = \hat{\beta}^3 + \hat{\beta}^2 - 2$?

Let us notice that

$$g(\beta) = 1^3 + 1^2 - 2 = 0$$

and

$$\partial g(\beta)/\partial\beta = 3\beta^2 + 2\beta = 5$$

Therefore,

$$\sqrt{N}(g(\hat{\beta}) - g(\beta)) = \sqrt{N}g(\hat{\beta}) \to^d N(0, 5 \cdot 2 \cdot 5) = N(0, 50)$$

Quadratic form

If Σ is nonsingular ($\det(\Sigma) \neq 0$), then

$$N(\hat{\beta} - \beta)'\Sigma^{-1}(\hat{\beta} - \beta) \to^{d} \chi^2(K)$$

Remark: Similarly, if $p\lim \hat{\Sigma} = \Sigma$ then

$$N(\hat{\beta} - \beta)'\hat{\Sigma}^{-1}(\hat{\beta} - \beta) \to^{d} \chi^2(K)$$

Quadratic form

If Σ is nonsingular ($\det(\Sigma) \neq 0$), then

$$N(\hat{\beta} - \beta)'\Sigma^{-1}(\hat{\beta} - \beta) \rightarrow^{d} \chi^2(K)$$

Remark: Similarly, if $p\lim \hat{\Sigma} = \Sigma$ then

$$N(\hat{\beta} - \beta)'\hat{\Sigma}^{-1}(\hat{\beta} - \beta) \rightarrow^{d} \chi^2(K)$$

Example 3

What is the asymptotic distribution of a quadratic form of $A(\hat{\beta}-\beta)$, where A is a quadratic,nonsingular matrix?

We know that

$$\sqrt{N}A(\hat{\beta}-\beta) \to^{d} N(0, A\Sigma A')$$

So replacing $\hat{\alpha} = A\hat{\beta}$, $\alpha = A\beta$ and $\Sigma_\alpha = A\Sigma A$ we have

$$\sqrt{N}(\hat{\alpha}-\alpha) \to^{d} N(0, \Sigma_\alpha)$$

and

$$N(\hat{\alpha}-\alpha)'\hat{\Sigma_\alpha}^{-1}(\hat{\alpha}-\alpha) \to^{d} \chi^2(K)$$

Example 3

What is the asymptotic distribution of a quadratic form of $A(\hat{\beta}-\beta)$, where A is a quadratic,nonsingular matrix?

We know that

$$\sqrt{N}A(\hat{\beta}-\beta) \rightarrow^{d} N(0, A\Sigma A')$$

So replacing $\hat{\alpha} = A\hat{\beta}$, $\alpha = A\beta$ and $\Sigma_{\alpha} = A\Sigma A$ we have

$$\sqrt{N}(\hat{\alpha}-\alpha) \rightarrow^{d} N(0, \Sigma_{\alpha})$$

and

$$N(\hat{\alpha}-\alpha)'\hat{\Sigma_{\alpha}}^{-1}(\hat{\alpha}-\alpha) \rightarrow^{d} \chi^{2}(K)$$

Example 3

What is the asymptotic distribution of a quadratic form of $A(\hat{\beta}-\beta)$, where A is a quadratic,nonsingular matrix?

We know that

$$\sqrt{N}A(\hat{\beta}-\beta) \rightarrow^{d} N(0, A\Sigma A')$$

So replacing $\hat{\alpha} = A\hat{\beta}$, $\alpha = A\beta$ and $\Sigma_{\alpha} = A\Sigma A$ we have

$$\sqrt{N}(\hat{\alpha}-\alpha) \rightarrow^{d} N(0, \Sigma_{\alpha})$$

and

$$N(\hat{\alpha}-\alpha)'\hat{\Sigma}_{\alpha}^{-1}(\hat{\alpha}-\alpha) \rightarrow^{d} \chi^2(K)$$

Example 3

Since

$$\hat{\alpha} - \alpha = A(\hat{\beta} - \beta)$$

and

$$(\hat{\alpha} - \alpha)' = (A(\hat{\beta} - \beta))' = (\hat{\beta} - \beta)'A'$$

Then

$$N(\hat{\beta} - \beta)'A'(A\Sigma A')^{-1}A(\hat{\beta} - \beta) \to^{d} \chi^2(K)$$

Example 4

What happens if the matrix A is not quadratic, but has dimension $(M \times K)$, where $M < K$ and $rank(A) = M$?

Similarly to the previous example, the quadratic form will have a χ^2 distribution. It will have M degrees of freedom.

$$N(\hat{\beta} - \beta)' A' (A \Sigma A')^{-1} A(\hat{\beta} - \beta) \to^d \chi^2(M)$$

Example 4

What happens if the matrix A is not quadratic, but has dimension $(M \times K)$, where $M < K$ and $rank(A) = M$?

Similarly to the previous example, the quadratic form will have a χ^2 distribution. It will have M degrees of freedom.

$$N(\hat{\beta} - \beta)' A' (A \Sigma A')^{-1} A (\hat{\beta} - \beta) \to^{d} \chi^2(M)$$