ECON 60532: *Lecture 3*

Alastair R. Hall

The University of Manchester

Spring semester, 2012

Outline of today's lecture

Outline of today's lecture

- Notions of Convergence
  - Some background
  - Convergence in probability: definition and comparison to other concepts

Outline of today's lecture

- Notions of Convergence
  - Some background
  - Convergence in probability: definition and comparison to other concepts
- Further properties of convergence in probability

Outline of today's lecture

- Notions of Convergence
  - Some background
  - Convergence in probability: definition and comparison to other concepts
- Further properties of convergence in probability
- Convergence in distribution

Random variables

A random variable is a mathematical device used to model the outcome of a statistical experiment.

Random variables

A random variable is a mathematical device used to model the outcome of a statistical experiment.

To present a more formal definition, we require the concept of a *probability space*.

Random variables

A random variable is a mathematical device used to model the outcome of a statistical experiment.

To present a more formal definition, we require the concept of a *probability space*.

Definition: A probability space consists of the triple $(\Omega, \mathcal{F}, P(.))$ where:

- Ω is the sample space;
- $\mathcal{F}$ is the σ-algebra associated with Ω;
- $P(.)$ is a probability measure defined over $\mathcal{F}$.

Random variables

Definition: A random variable V on $(\Omega, \mathcal{F}, P(.))$ is a real-valued function over the sample space Ω denoted $V(\omega)$ for $\omega \in \Omega$ such that for any real number v, $\{\omega \,|\, V(\omega) < v\} \in \mathcal{F}$.

Random variables

Definition: A random variable V on $(\Omega, \mathcal{F}, P(.))$ is a real-valued function over the sample space Ω denoted $V(\omega)$ for $\omega \in \Omega$ such that for any real number v, $\{\omega \,|\, V(\omega) < v\} \in \mathcal{F}$.

Example: Statistical experiment is to toss a coin and record whether a head or tail lands upwards.

- $\Omega = \{H, T\}$
- $\mathcal{F} = \{\emptyset, H, T, \{H, T\}\}$
- $P(\omega) = \theta,$ for $\omega = H,$ $P(\omega) = 1 - \theta,$ for $\omega = T$ (for some $0 \leq \theta \leq 1$)
- $V(\omega) = 0,$ for $\omega = H;$ $V(\omega) = 1,$ for $\omega = T$

Random variables

Associated with a rv is its distribution function.

Random variables

Associated with a rv is its distribution function.

Definition: The distribution function $F(v)$ of a random variable $V(\omega)$ is given by

$$F(v) = P(\omega \,|\, V(\omega) < v)$$

Random variables

Associated with a rv is its distribution function.

Definition: The distribution function $F(v)$ of a random variable $V(\omega)$ is given by

$$F(v) = P(\omega \,|\, V(\omega) < v)$$

Associated with the distribution is the expectation operator.

Random variables

Associated with a rv is its distribution function.

Definition: The distribution function $F(v)$ of a random variable $V(\omega)$ is given by

$$F(v) = P(\omega \,|\, V(\omega) < v)$$

Associated with the distribution is the expectation operator.

Definition: The expectation of $m(V)$ is given by:

$$E[m(V)] \; = \int_{-\infty}^{\infty} m(v)dF(v) \; = \lim_{\substack{v_\ell \to -\infty \\ v_u \to \infty}} \int_{-v_\ell}^{v_u} m(v)dF(v)$$

Existence of moments

It should be noted that a particular moment may not exist.

Example: Consider $E[V^s]$ when V has a Student's t-distribution with k degrees of freedom and s is a positive even number:

$$E[V^s] \;=\; c' \int_{-\infty}^{\infty} \frac{v^s}{(k \,+\, v^2)^{(k+1)/2}} dv$$

where c' is a constant.

Existence of moments

It should be noted that a particular moment may not exist.

Example: Consider $E[V^s]$ when V has a Student's t-distribution with k degrees of freedom and s is a positive even number:

$$E[V^s] \;=\; c' \int_{-\infty}^{\infty} \frac{v^s}{(k\,+\,v^2)^{(k+1)/2}} dv$$

where c' is a constant.

As $v \to \pm\infty$:

- $v^s \to \infty$;
- $(k\,+\,v^2)^{(k+1)/2} \to \infty$

Existence of moments

It should be noted that a particular moment may not exist.

Example: Consider $E[V^s]$ when V has a Student's t-distribution with k degrees of freedom and s is a positive even number:

$$E[V^s] \;=\; c' \int_{-\infty}^{\infty} \frac{v^s}{(k\,+\,v^2)^{(k+1)/2}} dv$$

where c' is a constant.

As $v \to \pm\infty$:

- $v^s \to \infty$;
- $(k\,+\,v^2)^{(k+1)/2} \to \infty$

Does the integral exist? $E[V^s]$ exists for all $s < k$.

Random sequences

What is random sequence?

Random sequences

What is random sequence?

Example (i): Suppose coin is tossed repeatedly

Random sequences

What is random sequence?

Example (i): Suppose coin is tossed repeatedly

- the sample space, Ω, consists of all possible sequences of heads and tails, e.g. a typical element ω is then of the form $\{H, T, T, H, H, H, T, H, \ldots\}$.

Random sequences

What is random sequence?

Example (i): Suppose coin is tossed repeatedly

- the sample space, Ω, consists of all possible sequences of heads and tails, e.g. a typical element ω is then of the form $\{H, T, T, H, H, H, T, H, \ldots\}$.
- Define $V_T(\omega)$ to be proportion of heads in the first T tosses.

Random sequences

What is random sequence?

Example (i): Suppose coin is tossed repeatedly

- the sample space, Ω, consists of all possible sequences of heads and tails, e.g. a typical element ω is then of the form $\{H, T, T, H, H, H, T, H, \ldots\}$.
- Define $V_T(\omega)$ to be proportion of heads in the first T tosses.
- $\{V_T(\omega)\} = \{V_1(\omega), V_2(\omega), V_3(\omega), \ldots\}$ is a random sequence.

Random sequences

What is random sequence?

Example (i): Suppose coin is tossed repeatedly

- the sample space, Ω, consists of all possible sequences of heads and tails, e.g. a typical element ω is then of the form $\{H, T, T, H, H, H, T, H, \ldots\}$.
- Define $V_T(\omega)$ to be proportion of heads in the first T tosses.
- $\{V_T(\omega)\} = \{V_1(\omega), V_2(\omega), V_3(\omega), \ldots\}$ is a random sequence.

Example (ii): IV estimator, $\{\hat{\theta}_T\} = \{\hat{\theta}_q, \hat{\theta}_{q+1}, \ldots\}$ is a random sequence.

Notions of convergence

There are a number of ways of characterizing the convergence of a random sequence.

Our focus is on: convergence in probability and convergence in distribution.

However we briefly discuss two other concepts: almost sure convergence and convergence in the r^{th} mean

Notions of convergence

There are a number of ways of characterizing the convergence of a random sequence.

Our focus is on: convergence in probability and convergence in distribution.

However we briefly discuss two other concepts: almost sure convergence and convergence in the r^{th} mean

Discussion organized as follows:

- almost sure convergence, convergence in probability, convergence in the r^{th} mean;
- convergence in distribution.

Convergence of deterministic sequence

Definition: The sequence $\{v_T;\ T = 1, 2, \ldots\}$ of real numbers is said to converge to a limit, v, if for every $\epsilon > 0$ there is a positive, finite integer $T(\epsilon)$ such that

$$|v_T - v| < \epsilon \quad \text{for} \quad T > T(\epsilon)$$

Convergence of deterministic sequence

Definition: The sequence $\{v_T;\ T = 1, 2, \ldots\}$ of real numbers is said to converge to a limit, v, if for every $\epsilon > 0$ there is a positive, finite integer $T(\epsilon)$ such that

$$|v_T - v| < \epsilon \quad \text{for} \quad T > T(\epsilon)$$

- If this limit exists then it is denoted by $\lim_{T\to\infty} v_T = v$ or $v_T \to v$ as $T \to \infty$.

Convergence of deterministic sequence

Definition: The sequence $\{v_T;\ T = 1, 2, \ldots\}$ of real numbers is said to converge to a limit, v, if for every $\epsilon > 0$ there is a positive, finite integer $T(\epsilon)$ such that

$$|v_T - v| < \epsilon \quad \text{for} \quad T > T(\epsilon)$$

- If this limit exists then it is denoted by $\lim_{T\to\infty} v_T = v$ or $v_T \to v$ as $T \to \infty$.
- Not every sequence has a limit.

Almost sure convergence

Consider the random sequence $V_1(\omega), V_2(\omega), V_3(\omega), \ldots$.

Fix $\omega = \tilde{\omega}$ say, so sequence is deterministic.

Almost sure convergence

Consider the random sequence $V_1(\omega), V_2(\omega), V_3(\omega), \ldots$.

Fix $\omega = \tilde{\omega}$ say, so sequence is deterministic.

If $V(\tilde{\omega})$ is the limit of $V_T(\tilde{\omega})$ then for every $\epsilon > 0$, there exists $T(\epsilon)$ such that

$$|V_T(\tilde{\omega}) - V(\tilde{\omega})| < \epsilon$$

for all $T > T(\epsilon)$.

Almost sure convergence

Consider the random sequence $V_1(\omega), V_2(\omega), V_3(\omega), \ldots$.

Fix $\omega = \tilde{\omega}$ say, so sequence is deterministic.

If $V(\tilde{\omega})$ is the limit of $V_T(\tilde{\omega})$ then for every $\epsilon > 0$, there exists $T(\epsilon)$ such that

$$|V_T(\tilde{\omega}) \; - \; V(\tilde{\omega})| \; < \; \epsilon$$

for all $T > T(\epsilon)$.

Now define set of values for ω for which $V(.)$ is the limit of $V_T(.)$:

$$D \; = \; \{\omega : lim_{T\to\infty} V_T(\omega) \; = \; V(\omega)\}$$

Almost sure convergence

Definition: V_T converges almost surely to V if $P(D) = 1$, or equivalently that

$$P(\lim_{T\to\infty} |V_T(\omega) - V(\omega)| < \epsilon) = 1$$

for every $\epsilon > 0$.

Almost sure convergence is denoted $V_T \stackrel{a.s.}{\to} V$.

Convergence in probability

Consider $p_T(\epsilon) = P(|V_T(\omega) - V(\omega)| < \epsilon)$.

Convergence in probability

Consider $p_T(\epsilon) \ = \ P(|V_T(\omega) \ - \ V(\omega)| \ < \ \epsilon)$.

If $\{p_T(\epsilon)\} \rightarrow 1$ for all $\epsilon > 0$ then $V_T(\omega)$ has "converged" to $V(\omega)$.

Convergence in probability

Consider $p_T(\epsilon) \;=\; P(|V_T(\omega) \;-\; V(\omega)| \;<\; \epsilon)$.

If $\{p_T(\epsilon)\} \rightarrow 1$ for all $\epsilon > 0$ then $V_T(\omega)$ has "converged" to $V(\omega)$.

Definition: V_T converges in probability to V if for every $\epsilon > 0$ and every $\delta > 0$, there exists $T(\epsilon, \delta)$ such that for all $T > T(\epsilon, \delta)$,

$$P(|V_T(\omega) \;-\; V(\omega)| \;<\; \epsilon) \;>\; 1 \;-\; \delta$$

or equivalently if $lim_{T\rightarrow\infty} p_T(\epsilon) = 1$ for every $\epsilon > 0$.

This type of convergence is denoted $V_T \xrightarrow{p} V$.

Convergence in probability

This definition holds for rv V.

If V is a *degenerate* rv then we have two important items of terminology.

Definition: If $V_T \xrightarrow{p} c$ where c is a constant then c is refered to as the probability limit of V_T and this is written as $plim V_T = c$.

Definition: If $\hat{\theta}_T$ is an estimator of the unknown parameter θ_0 and $\hat{\theta}_T \xrightarrow{p} \theta_0$ then $\hat{\theta}_T$ is said to be a consistent estimator of θ_0.

Convergence in r^{th} mean

Another way to assess the convergence of a random sequence is through the expectation of the difference.

Definition: The sequence of random variables $\{V_T\}$ converges in r^{th} moment $(r \geq 1)$ to the random variable V if $E[V_T^r]$ and $E[V^r]$ exists for all T if $\lim_{T\to\infty} E[|V_T - V|^r] = 0$.

Convergence in r^{th} mean

Another way to assess the convergence of a random sequence is through the expectation of the difference.

Definition: The sequence of random variables $\{V_T\}$ converges in r^{th} moment $(r \geq 1)$ to the random variable V if $E[V_T^r]$ and $E[V^r]$ exists for all T if $\lim_{T \to \infty} E[|V_T - V|^r] = 0$.

Leading case $r = 2$ for which we have:

Definition: V_T converges in quadratic mean to V if V_T converges in the second moment to V.

Convergence in quadratic mean is denoted $V_T \stackrel{q.m.}{\rightarrow} V$.

Convergence in quadratic mean

If the random variable is an estimator, $\hat{\theta}_T$, of the unknown parameter, θ_0, then it can be shown that

$$\begin{aligned} E\left[(\hat{\theta}_T - \theta_0)^2 \right] &= E\left[(\hat{\theta}_T - E[\hat{\theta}_T])^2 \right] + (E[\hat{\theta}_T] - \theta_0)^2 \\ &= Var[\hat{\theta}_T] + \left(bias[\hat{\theta}_T] \right)^2 \end{aligned}$$

Convergence in quadratic mean

If the random variable is an estimator, $\hat{\theta}_T$, of the unknown parameter, θ_0, then it can be shown that

$$
\begin{aligned}
E\left[(\hat{\theta}_T - \theta_0)^2\right] &= E\left[(\hat{\theta}_T - E[\hat{\theta}_T])^2\right] + (E[\hat{\theta}_T] - \theta_0)^2 \\
&= Var[\hat{\theta}_T] + \left(bias[\hat{\theta}_T]\right)^2
\end{aligned}
$$

$\hat{\theta}_T \stackrel{q.m.}{\rightarrow} \theta_0$ if and only if $Var[\hat{\theta}_T] \rightarrow 0$ and $bias[\hat{\theta}_T] \rightarrow 0$ as $T \rightarrow \infty$.

Summary of relationships

It can be shown that the following holds:

- $V_T \stackrel{a.s.}{\to} V \;\Rightarrow\; V_T \stackrel{p}{\to} V$;
- $V_T \stackrel{p}{\to} V \;\not\Rightarrow\; V_T \stackrel{a.s.}{\to} V$;
- $V_T \stackrel{q.m.}{\to} V \;\Rightarrow\; V_T \stackrel{p}{\to} V$;
- $V_T \stackrel{p}{\to} V \;\not\Rightarrow\; V_T \stackrel{q.m.}{\to} V$;
- no fixed relationship between almost sure convergence and convergence in quadratic mean.

Further properties of convergence in probability

Previous discussion defined the concept convergence in probability.

We now elaborate on its properties and consider:

- Extension to vectors and matrices
- Functions of random vectors
- Orders in probability

Convergence in probability for vectors

Recall definition of convergence in probability for scalar rv's revolves around

$$P(|V_T(\omega) - V(\omega)| < \epsilon) > 1 - \delta$$

Convergence in probability for vectors

Recall definition of convergence in probability for scalar rv's revolves around

$$P(|V_T(\omega) - V(\omega)| < \epsilon) > 1 - \delta$$

For $r \times 1$ vectors $V_T(\omega)$ and $V(\omega)$ with i^{th} $V_{T,i}$ and V_i

respectively replace distance measure by

$$\|V_T - V\| = \sqrt{\sum_{i=1}^{r}(V_{T,i} - V_i)^2}$$

Convergence in probability for matrices

If V_T and V are $r \times s$ matrices with $(i,j)^{th}$ elements $V_{T,i,j}$ and $V_{i,j}$ respectively then the natural measure of distance is

$$\begin{aligned} \| V_T - V \| &= \sqrt{tr\{(V_T - V)'(V_T - V)\}} \\ &= \sqrt{\sum_{i=1}^{r} \sum_{j=s}^{r} (V_{T,i,j} - V_{i,j})^2} \end{aligned}$$

where $tr(\cdot)$ is the trace operator.

Convergence in probability for matrices

If V_T and V are $r \times s$ matrices with $(i,j)^{th}$ elements $V_{T,i,j}$ and $V_{i,j}$ respectively then the natural measure of distance is

$$\begin{aligned} \|V_T - V\| &= \sqrt{tr\{(V_T - V)'(V_T - V)\}} \\ &= \sqrt{\sum_{i=1}^{r}\sum_{j=s}^{r}(V_{T,i,j} - V_{i,j})^2} \end{aligned}$$

where $tr(\cdot)$ is the trace operator.

- known as Frobenius norm - example of *matrix norm*.
- Note

$$\|V_T - V\| = \sqrt{vec(V_T - V)'vec(V_T - V)}$$

where $vec(\cdot)$ is the vec operator.

Slutsky's Theorem

Definition: $f(v)$ is a continuous function if for every $\epsilon > 0$ and every v_0 in the domain of $f(v)$ there exists a δ such that $|f(v) - f(v_0)| < \epsilon$ whenever $|v - v_0| < \delta$.

Slutsky's Theorem

Definition: $f(v)$ is a continuous function if for every $\epsilon > 0$ and every v_0 in the domain of $f(v)$ there exists a δ such that $|f(v) - f(v_0)| < \epsilon$ whenever $|v - v_0| < \delta$.

Slutsky's theorem: Let $\{V_T\}$ be a sequence of $r \times 1$ random vectors (or matrices) which converge in probability to the random vector (or matrix) V and let $f(.)$ be a real- valued vector of continuous functions then $f(V_T) \xrightarrow{p} f(V)$.

Slutsky's Theorem

Definition: $f(v)$ is a continuous function if for every $\epsilon > 0$ and every v_0 in the domain of $f(v)$ there exists a δ such that $|f(v) - f(v_0)| < \epsilon$ whenever $|v - v_0| < \delta$.

Slutsky's theorem: Let $\{V_T\}$ be a sequence of $r \times 1$ random vectors (or matrices) which converge in probability to the random vector (or matrix) V and let $f(.)$ be a real- valued vector of continuous functions then $f(V_T) \xrightarrow{p} f(V)$.

Note:

- Very important property.
- Not shared by $E[.]$, that is $E[f(V)] \neq f(E[V])$.

Orders of magnitude

So far have focused on convergence, but other aspects of sequence are also of interest, such as behaviour as function of T

$\longrightarrow$ orders of magnitude

Orders of magnitude

So far have focused on convergence, but other aspects of sequence are also of interest, such as behaviour as function of T

$$\longrightarrow \text{orders of magnitude}$$

To introduce this concept, it is once again useful to begin with deterministic sequences.

Definition: (i) The sequence $\{v_T;\ T = 1, 2, \ldots\}$ is said to be of large order of magnitude c_T if there exists a real number m such that $|v_T|/c_T < m$ for all T. This is denoted by $v_T = O(c_T)$; (ii) The sequence $\{v_T\}$ is said to be of small order of magnitude c_T if the limit of v_T/c_T is zero as $T \to \infty$. This is denoted by $v_T = o(c_T)$.

Orders of magnitude

Example: $v_T = aT^2 + bT + c$ where a, b and c are finite constants.

Orders of magnitude

Example: $v_T = aT^2 + bT + c$ where a, b and c are finite constants.

- $v_T = O(T^2)$.

Orders of magnitude

Example: $v_T = aT^2 + bT + c$ where a, b and c are finite constants.

- $v_T = O(T^2)$.
- $v_T = o(\nu_T)$ for any sequence of constants $\{\nu_T\}$ such that $\nu_T \to \infty$ as $T \to \infty$ but $T^2/\nu_T = o(1)$.

Orders of magnitude

The extension to random sequences is as follows.

Definition: (i) The sequence of random variables $\{V_T\}$ is said to be of large order in probability c_T if for every $\epsilon > 0$ there exists positive real numbers m_ϵ and T_ϵ such that $P[|V_T|/c_T > m_\epsilon] \leq \epsilon$ for all $T \geq T_\epsilon$. This is denoted by $V_T = O_p(c_T)$; (ii) The sequence of random variables $\{V_T\}$ is said to be of small order in probability c_T if $plim(V_T/c_T) = 0$. This is denoted by $V_T = o_p(c_T)$.

Orders of magnitude

The extension to random sequences is as follows.

Definition: (i) The sequence of random variables $\{V_T\}$ is said to be of large order in probability c_T if for every $\epsilon > 0$ there exists positive real numbers m_ϵ and T_ϵ such that $P[|V_T|/c_T > m_\epsilon] \leq \epsilon$ for all $T \geq T_\epsilon$. This is denoted by $V_T = O_p(c_T)$; (ii) The sequence of random variables $\{V_T\}$ is said to be of small order in probability c_T if $plim(V_T/c_T) = 0$. This is denoted by $V_T = o_p(c_T)$.

Example: If $\hat{\theta}_T$ is an estimator θ_0 and it can be shown that $\hat{\theta}_T = \theta_0 + o_p(1)$ then this implies that $\hat{\theta}_T$ is consistent for θ_0.

Convergence in distribution

Definition of convergence in distribution is based on sequence of distribution functions.

Recall condition for convergence of deterministic functions (with scalar arguments) is:

Definition: The sequence of deterministic functions $\{f_T(\cdot)\}$ *converges to the function* $f(\cdot)$ if and only if, for all v in the domain of $f(v)$ and for every $\epsilon > 0$, there exists $T(\epsilon)$ such that

$$|f_T(v) - f(v)| < \epsilon \text{ for } T > T(\epsilon)$$

- Notice that this definition requires that for every v, there exists a T such that the difference between $f_T(v)$ and $f(v)$ becomes arbitrarily small.

Convergence in distribution

Definition: The sequence of random variables $\{V_T\}$ with corresponding distribution functions $\{F_T(\cdot)\}$ converges in distribution to the random variable V with distribution function $F(\cdot)$ if and only if there exists $T(\epsilon)$ for every ϵ such that $|F_T(c) - F(c)| < \epsilon$ for $T > T(\epsilon)$ at all points of continuity $\{c\}$.

Convergence in distribution

Definition: The sequence of random variables $\{V_T\}$ with corresponding distribution functions $\{F_T(\cdot)\}$ converges in distribution to the random variable V with distribution function $F(\cdot)$ if and only if there exists $T(\epsilon)$ for every ϵ such that $|F_T(c) - F(c)| < \epsilon$ for $T > T(\epsilon)$ at all points of continuity $\{c\}$.

- this is denoted by $V_T \xrightarrow{d} V$;

Convergence in distribution

Definition: The sequence of random variables $\{V_T\}$ with corresponding distribution functions $\{F_T(\cdot)\}$ converges in distribution to the random variable V with distribution function $F(\cdot)$ if and only if there exists $T(\epsilon)$ for every ϵ such that $|F_T(c) - F(c)| < \epsilon$ for $T > T(\epsilon)$ at all points of continuity $\{c\}$.

- this is denoted by $V_T \stackrel{d}{\rightarrow} V$;
- the distribution of V is known as the *limiting (or asymptotic) distribution* of V_T.

Convergence in distribution

Definition: The sequence of random variables $\{V_T\}$ with corresponding distribution functions $\{F_T(\cdot)\}$ converges in distribution to the random variable V with distribution function $F(\cdot)$ if and only if there exists $T(\epsilon)$ for every ϵ such that $|F_T(c) - F(c)| < \epsilon$ for $T > T(\epsilon)$ at all points of continuity $\{c\}$.

- this is denoted by $V_T \xrightarrow{d} V$;
- the distribution of V is known as the *limiting (or asymptotic) distribution* of V_T.
- if $V_T \xrightarrow{d} V$, a well defined random variable, then $V_T = O_p(1)$.

Convergence of functions

As with convergence in probability, the convergence in distribution of V_T implies convergence in distribution of continuous functions of V_T.

Theorem: Let $f(.)$ be a real-valued, continuous, scalar function. If $V_T \xrightarrow{d} V$ then $f(V_T) \xrightarrow{d} f(V)$.

Convergence of functions

As with convergence in probability, the convergence in distribution of V_T implies convergence in distribution of continuous functions of V_T.

Theorem: Let $f(.)$ be a real-valued, continuous, scalar function. If $V_T \xrightarrow{d} V$ then $f(V_T) \xrightarrow{d} f(V)$.

Example: $f(\cdot) = (\cdot)^2$. Theorem $\Rightarrow$ if $V_T \xrightarrow{d} V$ then $V_T^2 \xrightarrow{d} V^2$; and so, if $V_T \xrightarrow{d} N(0,1)$ then $V_T^2 \xrightarrow{d} \chi_1^2$.

Comparison to convergence in probability

- If $V_T \xrightarrow{p} V$ then $V_T - V = o_p(1)$ and V_T and V are essentially the same in the limit.
- If $V_T \xrightarrow{d} V$ then V_T has the same distribution as V in the limit.

Comparison to convergence in probability

- If $V_T \xrightarrow{p} V$ then $V_T - V = o_p(1)$ and V_T and V are essentially the same in the limit.
- If $V_T \xrightarrow{d} V$ then V_T has the same distribution as V in the limit.

Theorem: If $V_T \xrightarrow{p} V$ then $V_T \xrightarrow{d} V$.

Comparison to convergence in probability

- If $V_T \xrightarrow{p} V$ then $V_T - V = o_p(1)$ and V_T and V are essentially the same in the limit.
- If $V_T \xrightarrow{d} V$ then V_T has the same distribution as V in the limit.

Theorem: If $V_T \xrightarrow{p} V$ then $V_T \xrightarrow{d} V$.

The reverse implication does not hold in general unless V is a degenerate random variable.

Theorem: $V_T \xrightarrow{p} c$, where c is a constant, if and only if $V_T \xrightarrow{d} c$.

Comparison to convergence in probability

As a result of the relationship between the two modes of convergence, we have the following useful result.

Theorem: If $|V_T - X_T| \xrightarrow{p} 0$ and $V_T \xrightarrow{d} V$ then $X_T \xrightarrow{d} V$.

Often establish limiting distribution as follows:

- Show $X_T = V_T + o_p(1)$
- Show $V_T \xrightarrow{d} V$
- Above theorem $\Rightarrow X_T \xrightarrow{d} V$.

Joint convergence

In general, if we are interested in the convergence in distribution of $f(V_T, X_T)$ then it is necessary to consider conditions for the joint convergence of (V_T, X_T), that is it is not sufficient to consider the conditions for their individual convergence in distribution.

One useful exception is the following.

Theorem: If $V_T \xrightarrow{d} V$ and $X_T \xrightarrow{p} c$, a constant, then the limit of the joint distribution of (V_T, X_T) exists and equals that of (V, c).

Joint convergence

In general, if we are interested in the convergence in distribution of $f(V_T, X_T)$ then it is necessary to consider conditions for the joint convergence of (V_T, X_T), that is it is not sufficient to consider the conditions for their individual convergence in distribution.

One useful exception is the following.

Theorem: If $V_T \xrightarrow{d} V$ and $X_T \xrightarrow{p} c$, a constant, then the limit of the joint distribution of (V_T, X_T) exists and equals that of (V, c).

So if $V_T \xrightarrow{d} V$ and $X_T \xrightarrow{p} c$ then we have (for instance):

- $V_T + X_T \xrightarrow{d} V + c$;
- $V_T X_T \xrightarrow{d} cV$.

Further reading

- Supplemental notes (that can be down loaded from course web page) and the references therein.