Chapter 2.11: Central Limit Theorem for Processes

Lasse Vuursteen

May 25, 2020

So far, the Donsker Theorem's were based on identical distributions.

$$\mathbb{G}_n := n^{-1/2}(\mathbb{P}_n - P) \rightsquigarrow \mathbb{G}\ell^\infty(\mathscr{F}).$$

Under the hood here are iid observations:

$$n^{-1/2}\sum_{i=1}^{n}(f(Z_i) - \mathbb{E}f(Z_i))$$

$Z_1, \dots, Z_n$ are iid, $f \in \mathscr{F}$.

What if $Z_1, \dots, Z_n$ are not iid?

In the case $\mathscr{F} = \{x \mapsto x\}$, we have the Lindeberg CLT

Theorem

Let $Z_{n1}, \dots, Z_{nm_n}$ independent, $\mathbb{E}Z_{ij} = 0$,

$$\sum_{i=1}^{m_n} Var(X_{ni}) = 1$$

and

$$L_n(\varepsilon) := \sum_{i=1}^{m_n} \mathbb{E}X_{ni}^2\{|X_{ni}| > \varepsilon\} \to 0$$

for all $\varepsilon > 0$. Then,

$$\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni} \rightsquigarrow N(0,1).$$

Chapter 2.11 “generalizes” the previous theorem to empirical processes indexed by more demanding $\mathscr{F}$’s.

That is, it generalizes Donsker type theorem to accomodate for sequence of SP’s $Z_{n1}, \ldots, Z_{nm_n}$.

The results in Chapter 2.11 allow for different probability spaces for each of the Z_{ni}'s.

$Z_{n1}, \ldots, Z_{nm_n}$ takes values in

$$\prod_{i=1}^{m_n} (\mathscr{X}_{ni}, \mathscr{A}_{ni}, P_{ni})$$

$$\left\{ \sum_{i=1}^{m_n} Z_{ni}(f) - P_{ni} Z_{ni}(f) : f \in \mathscr{F} \right\}$$

The results in Chapter 2.11 allow for different probability spaces for each of the Z_{ni}'s.

$Z_{n1}, \ldots, Z_{nm_n}$ takes values in

$$\prod_{i=1}^{m_n} (\mathscr{X}_{ni}, \mathscr{A}_{ni}, P_{ni})$$

The stochastic processes for which we want weak convergence in $\ell^\infty(\mathscr{F})$ are

$$\left\{ \sum_{i=1}^{m_n} Z_{ni}(f) - \mathbb{E} Z_{ni}(f) : f \in \mathscr{F} \right\} \text{ for } n \in \mathbb{N}.$$

So for example $Z_{ni}(f)$ could be

$$Z_{ni}(f) := \frac{1}{\sqrt{n}} f(X_i)$$

for some independent random variables X_i which are not identically distributed, and then we recover something looking a lot like the empirical process.

A function class $\mathscr{F}$ is called *P-Donsker* if

$$\mathbb{G}_n := n^{-1/2}(\mathbb{P}_n - P) \rightsquigarrow \mathbb{G} \text{ in } \ell^\infty(\mathscr{F}).$$

In Chapter 2.5 we have seen Donsker theorems in different flavours (different types of conditions on $\mathscr{F}$):

- Based on the uniform entropy type conditions

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon||F||_{Q,2}, \mathscr{F}, L_2(Q))} d\varepsilon < \infty.$$

- Based on bracketing entropy

$$\int_0^\infty \sqrt{\log N_{[]}(\varepsilon, \mathscr{F}, L_2(P))} d\varepsilon < \infty.$$

These conditions are generally *not* comparable.

Donsker Theorem's based on uniform entropy

Theorem

Let class $\mathscr{F}$ of measurable functions satisfying

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon||F||_{Q,2}, \mathscr{F}, L_2(Q))} d\varepsilon < \infty$$

If furthermore $\mathscr{F}_\delta$ and $\mathscr{F}_\infty^2$ are P-measurable for every $\delta > 0$ and F is a square integrable envelope, then $\mathscr{F}$ is P-Donsker.

Donsker Theorem's based on bracketing

Theorem

Any class $\mathscr{F}$ of measurable functions satisfying

$$\int_0^\infty \sqrt{\log N_{[]}(\varepsilon, \mathscr{F}, L_2(P))}\, d\varepsilon < \infty$$

is P-Donsker.

Chapter 2.11 gives versions of these two theorems for the non-iid setting; sufficient conditions for weak convergence based on

- a uniform entropy condition,
- a bracketing entropy condition,
- a “Random entropy” condition.

2.11.1: Random entropy

We start by defining the "process" equivalent of the empirical L_2 metric.

$$d_n(f,g) := \sum_{i=1}^{m_n} (Z_{ni}(f) - Z_{ni}(g))^2 .$$

Compare to:

$$||f-g||^2_{L_2(\mathbb{P}_n)} := \int (f-g)^2 d\mathbb{P}_n.$$

2.11.1: Random entropy

The random entropy condition is formulated in terms of the random semimetric d_n

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)}\, d\varepsilon \xrightarrow{P^*} 0 \quad \text{for every } \delta_n \searrow 0.$$

2.11.1: Random entropy

We need to following measurability assumption:

$$(x_1, \dots, x_{m_n}) \mapsto \sup_{\rho(f,g)<\delta} |\sum_{i=1}^{m_n} e_i(Z_{ni}(f) - \mathbb{E}Z_{ni}(f))|$$

$$(x_1, \dots, x_{m_n}) \mapsto \sup_{\rho(f,g)<\delta} |\sum_{i=1}^{m_n} e_i(Z_{ni}(f) - \mathbb{E}Z_{ni}(f))^2|$$

need to be measurable for the completion of $\prod_{i=1}^{m_n}(\mathscr{X}_{ni}, \mathscr{A}_{ni}, P_{ni})$.

Theorem (2.11.1)

Let $(\mathscr{F}, \rho)$ totally bounded semimetric space. Assume the Lindeberg condition

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||_{\mathscr{F}}^2 \{||Z_{ni}||_{\mathscr{F}} > \varepsilon\} \to 0 \text{ for all } \varepsilon > 0.$$

Assume furthermore that

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \to 0$$

for all $\delta_n \to 0$ and entropy condition

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)} d\varepsilon \overset{P^*}{\to} 0.$$

Then, $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ is asymptotically ρ-equicontinuous. If the covariance of the marginals converges pointwise in $\mathscr{F} \times \mathscr{F}$, we have $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ converges weakly in $\ell^\infty(\mathscr{F})$.

Before going to the proof, let's briefly compare with the regular Lindeberg CLT.

Theorem (Lindeberg CLT)

Let $Z_{n1}, \ldots, Z_{nm_n}$ independent, $\mathbb{E}Z_{ij} = 0$,

$$\sum_{i=1}^{m_n} Var(Z_{ni}) = 1$$

and

$$L_n(\varepsilon) := \sum_{i=1}^{m_n} \mathbb{E}Z_{ni}^2\{|Z_{ni}| > \varepsilon\} \to 0$$

for all $\varepsilon > 0$. Then,

$$\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni} \rightsquigarrow N(0, 1).$$

Theorem (2.11.1)

Let $(\mathscr{F}, \rho)$ totally bounded semimetric space. Assume the Lindeberg condition

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||^2_{\mathscr{F}} \{||Z_{ni}||_{\mathscr{F}} > \varepsilon\} \to 0 \textit{ for all } \varepsilon > 0.$$

Assume furthermore that

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \to 0$$

for all $\delta_n \to 0$ and entropy condition

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)} d\varepsilon \xrightarrow{P^*} 0.$$

Then, $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ is asymptotically ρ-equicontinuous. If the covariance of the marginals converges pointwise in $\mathscr{F} \times \mathscr{F}$, we have $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ converges weakly to a Gaussian process in $\ell^\infty(\mathscr{F})$.

Proof of 2.11.1:

By Theorem 1.5.4, for convergence in distribution in $\ell^\infty(\mathscr{F})$ it is enough to show

- convergence of maringals
- asymptotic tightness.

Theorem 1.5.7: X_α is asymptotically tight if and only if $X_\alpha(t)$ is asympotically tight in $\mathbb{R}$ for every t and there exists a semimetric ρ on T such that (T, ρ) is totally bounded and X_α is asymptotically uniformly ρ-equicontinuous in probability.

So asymptotic ρ-equicontinuity is (as usual) enough here.

Proof of 2.11.1: Convergence of marginals

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||^2_{\mathscr{F}} \{||Z_{ni}||_{\mathscr{F}} > \varepsilon\} \to 0 \text{ for all } \varepsilon > 0.$$

is much stronger than the regular Lindeberg condition needed for marginal in $f \in \mathscr{F}$:

$$\sum_{i=1}^{m_n} E^* Z_{ni}(f)^2 \{|Z_{ni}(f)| > \varepsilon\} \to 0 \text{ for all } \varepsilon > 0.$$

Proof of 2.11.1: Convergence of marginals

Theorem (Lindeberg CLT)

Let $Z_{n1}, \dots, Z_{nm_n}$ independent, $\mathbb{E}Z_{ij} = 0$,

$$\sum_{i=1}^{m_n} Var(Z_{ni}) = 1$$

and

$$L_n(\varepsilon) := \sum_{i=1}^{m_n} \mathbb{E}Z_{ni}^2\{|Z_{ni}| > \varepsilon\} \to 0$$

for all $\varepsilon > 0$. Then,

$$\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni} \rightsquigarrow N(0,1).$$

Apply (multivariate version) to $Z_{n1}(f_i), \dots, Z_{nm_n}(f_i)$.

Proof of 2.11.1: Convergence of marginals

To get the convergence of marginals that we are after, we need

$$\mathrm{Cov}_n(f,g) := \sum_{i=1}^{m_n} \left[\mathbb{E}Z_{ni}(f)Z_{ni}(g) - \mathbb{E}Z_{ni}(f)\mathbb{E}Z_{ni}(g)\right]$$

to convergence pointwise for $(f,g) \in \mathscr{F} \times \mathscr{F}$.

Proof of 2.11.1: Asymptotic equicontinuity

A net $X_\alpha : \Omega \to \ell^\infty(T)$ is *asymptotically uniformly ρ-equicontinuous in probability* if for every $\varepsilon, \eta > 0$ there exists a $\delta > 0$ such that

$$\limsup_{\alpha} P^* \left(\sup_{\rho(s,t)<\delta} |X_\alpha(s) - X_\alpha(t)| > \varepsilon \right) < \eta.$$

Proof of 2.11.1: Asymptotic equicontinuity

So it is enough to show that for arbitrary $t > 0$, there is some $\delta_n \searrow 0$ such that

$$P^* \left(\sup_{\rho(f,g)<\delta_n} |\sum_{i=1}^{m_n} (Z_{ni}^\circ(f) - Z_{ni}^\circ(g))| > t \right) \to 0$$

where $Z_{ni}^\circ = Z_{ni} - \mathbb{E}Z_{ni}$.

Proof of 2.11.1: Asymptotic equicontinuity

Lemma (2.3.7)

For arbitrary stochastic processes $Z_1, \dots, Z_n$

$$P^*\left(\|\sum_{i=1}^{n} Z_i\|_{\mathscr{F}} > x\right) \leq \frac{2}{\beta_n(x)} P^*\left(\|\sum_{i=1}^{n} \varepsilon_i Z_i\|_{\mathscr{F}} > x\right)$$

for every $x > 0$ and $\beta_n(x) \leq \inf_f P(|\sum_{i=1}^{n} Z_i(f)| < x/2)$.

Proof of 2.11.1: Asymptotic equicontinuity

We have

$$\sup_{\rho(f,g)<\delta_n} P(|\sum_{i=1}^{m_n} Z_{ni}^\circ(f) - Z_{ni}^\circ(g)| > t/2) \leq \frac{1}{2}$$

by the assumption

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}\left(Z_{ni}(f) - Z_{ni}(g)\right)^2 \to 0.$$

So by Lemma 2.3.7:

$$P^*\left(\sup_{\rho(f,g)<\delta_n} |\sum_{i=1}^{m_n} (Z_{ni}^\circ(f) - Z_{ni}^\circ(g))| > t\right) \leq$$

$$4P\left(\sup_{\rho(f,g)<\delta_n} |\sum_{i=1}^{m_n} \varepsilon_i (Z_{ni}(f) - Z_{ni}(g))| > t/4\right)$$

Proof of 2.11.1: Asymptotic equicontinuity

By the measurability assumptions, we can condition

$$4P\left(\sup_{\rho(f,g)<\delta_n}|\sum_{i=1}^{m_n}\varepsilon_i(Z_{ni}(f)-Z_{ni}(g))|>t/4\right)=$$
$$4P_Z P_\varepsilon\left(\sup_{\rho(f,g)<\delta_n}|\sum_{i=1}^{m_n}\varepsilon_i(Z_{ni}(f)-Z_{ni}(g))|>t/4\,\middle|\,Z_{n1},\dots,Z_{nn}\right)$$

We focus first on bounding the inner probability statement.

Proof of 2.11.1: Asymptotic equicontinuity

For fixed $Z_{n1}, \dots, Z_{nn}$, first define the random subset $A_n \subset \mathbb{R}^{m_n}$ consisting of

$$(Z_{n1}(f) - Z_{n1}(g), \dots, Z_{nm_n}(f) - Z_{nm_n}(g))$$

where (f,g) range over $\{(f,g) \in \mathscr{F} \times \mathscr{F} : \rho(f,g) < \delta_n\}$.

We can write

$$P_\varepsilon\left(\sup_{\rho(f,g)<\delta_n} |\sum_{i=1}^{m_n} \varepsilon_i (Z_{ni}(f) - Z_{ni}(g))| > t/4\right) = P_\varepsilon\left(\sup_{a \in A_n} |\sum_{i=1}^{m_n} \varepsilon_i a_i| > t/4\right)$$

and bound by

$$\frac{4\mathbb{E}_\varepsilon \sup_{a \in A_n} |\sum_{i=1}^{m_n} \varepsilon_i a_i|}{t}$$

Proof of 2.11.1: Asymptotic equicontinuity

Corollary (2.2.8)

Let $\{X_t : t \in T\}$ separable sub-Gaussian process for metric d. Then we have for all $t_0 \in T$

$$\mathbb{E} \sup_t |X_t| \lesssim \mathbb{E}|X_{t_0}| + \int_0^\infty \sqrt{\log N(\varepsilon, T, d)} d\varepsilon$$

We aim to apply this to $\{\sum_{i=1}^n \varepsilon_i a_i : a \in A_n\}$, with $d = ||\cdot||_2$ and taking $t_0 = a_0$ corresponding to (f, f), so the first term on the right hand side vanishes.

Proof of 2.11.1: Asymptotic equicontinuity

Letting $a_i = Z_{ni}(f) - Z_{ni}(g)$, the sub-Gaussianity required by Corollary 2.2.8 is

$$P_\varepsilon(|\sum_{i=1}^{m_n} \varepsilon_i a_i| > x) \lesssim e^{-\frac{x^2}{||a||_2}}$$

This is given by Hoeffding's lemma.

Proof of 2.11.1: Asymptotic equicontinuity

Using 2.2.8, we obtain

$$\mathbb{E}_\varepsilon \sup_{a \in A_n} |\sum_{i=1}^{m_n} \varepsilon_i a_i| \lesssim \int_0^\infty \sqrt{\log N(\varepsilon, A_n, ||\cdot||_2)} d\varepsilon$$

Define

$$\theta_n = \sup_{a \in A_n} \sqrt{\sum_{i=1}^{m_n} a_i^2}$$

and note that if $\varepsilon > \theta_n$,

$$\sqrt{\log N(\varepsilon, A_n, ||\cdot||_2)} = 0$$

So,

$$\mathbb{E}_\varepsilon \sup_{a \in A_n} |\sum_{i=1}^{m_n} \varepsilon_i a_i| \lesssim \int_0^{\theta_n} \sqrt{\log N(\varepsilon, A_n, ||\cdot||_2)} d\varepsilon$$

Proof of 2.11.1: Asymptotic equicontinuity

For the entropy numbers we have

$$N(\varepsilon, A_n, ||\cdot||) \leq N^2(\varepsilon/2, \mathscr{F}, d_n)$$

by construction; recalling that $A_n \subset \mathbb{R}^{m_n}$ consists of

$$(Z_{n1}(f) - Z_{n1}(g), \ldots, Z_{nm_n}(f) - Z_{nm_n}(g))$$

where (f, g) range over $\{(f,g) \in \mathscr{F} \times \mathscr{F} : \rho(f,g) < \delta_n\}$.

Proof of 2.11.1: Asymptotic equicontinuity

Consequently, we have

$$\mathbb{E}_\varepsilon \sup_{a \in A_n} \Big|\sum_{i=1}^{m_n} \varepsilon_i a_i\Big| \lesssim \int_0^{\theta_n} \sqrt{2 \log N(\varepsilon/2, \mathscr{F}, d_n)}\, d\varepsilon$$

Suppose now that there exists a sequence $s_n \searrow 0$ such that $P(\theta_n > s_n) \to 0$. For such a sequence,

$$\int_0^{s_n} \sqrt{2 \log N(\varepsilon/2, \mathscr{F}, d_n)}\, d\varepsilon \to 0$$

in outer probability by assumption.

Proof of 2.11.1: Asymptotic equicontinuity

Letting $E_\eta = \{\theta_n < s_n, \int_0^{s_n} \sqrt{2\log N(\varepsilon/2, \mathscr{F}, d_n)} d\varepsilon < \eta t\}$ and putting together the previous steps

$$P_Z P_\varepsilon\big(\sup_{a\in A_n} |\sum_{i=1}^{m_n} \varepsilon_i a_i| > t/4\big)(1_E + 1_{E^c}) \lesssim P_Z \frac{4\int_0^{s_n} \sqrt{2\log N(\varepsilon/2, \mathscr{F}, d_n)} d\varepsilon}{t} 1_E + 1_{E^c}$$
$$\leq \eta P_Z(E) + P(E^c)$$

where $P(E^c) \to 0$ provided $\theta_n \xrightarrow{P} 0$ and η can be picked arbitrarily small.

Proof of 2.11.1: Asymptotic equicontinuity

It is left to prove that indeed $\theta_n \xrightarrow{P} 0$. Proof of this fact is similar as in 2.5.2. It uses:

- (again) Lindeberg condition to control $||Z_{ni}||_{\mathscr{F}}$
- Symmetrization lemma + ε-net B_n approximation of A_n.
- Hoffmann-Jorgensen A.1.5
- The cardinality of B_n (ie $(N^2(\varepsilon, \mathscr{F}, d_n))$ being bounded in probability (Exercise 2.11.1)

Proof of 2.11.1: Asymptotic equicontinuity

1. The random entropy condition (2.11.2) implies that the sequence $N(\varepsilon, \mathcal{F}, d_n)$ is bounded in probability for every $\varepsilon > 0$.
[**Hint:** For every $\delta_n \leq \varepsilon$,

$$\mathrm{P}\Big(N(\varepsilon, \mathcal{F}, d_n) \geq M_n\Big) \leq \mathrm{P}\Big(\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathcal{F}, d_n)}\, d\varepsilon \geq \delta_n \sqrt{\log M_n}\Big).$$

Given $M_n \to \infty$, choose $\delta_n \downarrow 0$ such that $\delta_n \sqrt{\log M_n}$ is bounded away from zero.]

Theorem (2.11.1)

Let $(\mathscr{F}, \rho)$ *totally bounded semimetric space. Assume the condition*

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||^2_{\mathscr{F}} \{||Z_{ni}||_{\mathscr{F}} > \varepsilon\} \to 0 \text{ for all } \varepsilon > 0.$$

Assume furthermore that

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \to 0$$

for all $\delta_n \to 0$ *and entropy condition*

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)}\, d\varepsilon \xrightarrow{P^*} 0.$$

Then, $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ *is asymptotically* ρ*-equicontinuous. If the covariance of the marginals converges pointwise in* $\mathscr{F} \times \mathscr{F}$*, we have* $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ *converges weakly in* $\ell^\infty(\mathscr{F})$.

Measurelike processes and uniform entropy

We now will turn to the case that $\mathscr{F}$ consists of measurable functions $f : \mathscr{X} \to \mathbb{R}$ satisfying

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon||F||_{Q,2}, \mathscr{F}, L_2(Q))} d\varepsilon < \infty.$$

This is the uniform entropy condition we have used in Chapter 2.5 to verify the Donsker property for $\mathscr{F}$ in for example Theorem 2.5.2.

We wish to relate the above to the random entropy condition of the previous theorem.

We call Z_{ni} *measure like* with respect to random measures μ_{ni} if

$$(Z_{ni}(f) - Z_{ni}(f))^2 \leq \int (f-g)^2 d\mu_{ni}$$

for every $f, g \in \mathscr{F}$

Lemma 2.11.6

Lemma

Let $\mathscr{F}$ be a class of measurable functions with envelope function F. Let $Z_{n1}, \ldots, Z_{nm_n}$ be measurelike with respect to μ_{ni}'s. Suppose $\mathscr{F}$ satisfies

$$\int_0^\infty \sup_Q \sqrt{\log N(\varepsilon \|F\|_{Q,2}, \mathscr{F}, L_2(Q))}\, d\varepsilon < \infty.$$

and the Q contains $\sum_{i=1}^{m_n} \mu_{ni}$ and $\sum_{i=1}^{m_n} \mu_{ni} F^2 = O_P^*(1)$, then

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)}\, d\varepsilon \xrightarrow{P^*} 0.$$

Lemma 2.11.6: Proof

Since Z_{ni} is measure like, we have random measures μ_{ni} satisfying

$$(Z_{ni}(f) - Z_{ni}(f))^2 \leq \int (f-g)^2 d\mu_{ni}$$

for every $f, g \in \mathscr{F}$.

Write $\mu_n = \sum_{i=1}^{m_n} \mu_{ni}$. Note that

$$d_n^2(f,g) \leq \int (f-g)^2 d\mu_n.$$

So,

$$N(\varepsilon, \mathscr{F}, d_n) \leq N(\varepsilon, \mathscr{F}, L_2(\mu_n))$$

We have

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)} d\varepsilon = \int_0^{\frac{\delta_n}{||F||}} \sqrt{\log N(\varepsilon ||F||, \mathscr{F}, d_n)} d\varepsilon ||F||$$

$$\leq \int_0^{\frac{\delta_n}{||F||}} \sqrt{\log N(\varepsilon ||F||, \mathscr{F}, L_2(\mu_n))} d\varepsilon ||F|| := J'$$

where $||F|| \equiv ||F||_{\mu_n} = \mu_n F^2$.

Write

$$J(\delta) = \int_0^{\delta} \sup_Q \sqrt{\log N(\varepsilon ||F||_{Q,2}, \mathscr{F}, L_2(Q))} d\varepsilon$$

The proof is finished by noting that

- $J' \leq J(\infty)\eta$ on the set $\{||F||_{\mu_n} \leq \eta\}$, can make arb. small by choice η.
- On $\{||F||_{\mu_n} \leq \eta\}$, $J' \leq J(\delta_n/\eta)$ which is will converge to 0 in probability for every $\delta_n \to 0$ by assumption.

Example: weighted empirical processes

Let

- $X_{n1}, \ldots, X_{nm_n}$ be independent random elements in measurable space $(\mathscr{X}, \mathscr{X})$ and corresponding laws P_{ni}
- $P_{ni} f$ exist for all $n, i = 1, \ldots, m_n$ and f in measurable class of real valued functions $\mathscr{F}$
- Let c_{ni} constants where $i = 1, \ldots, m_n$ and define

$$\mathbb{G}_n := \sum_{i=1}^{m_n} c_{ni}(f(X_{ni}) - P_{ni} f).$$

Example: weighted empirical processes

This process $\mathbb{G}_n := \sum_{i=1}^{m_n} c_{ni}(f(X_{ni}) - P_{ni}f)$ can be related to a measurelike process Z_{ni} by setting

$$Z_{ni} = c_{ni}\delta_{X_{ni}}$$

which are measurelike for the measures

$$\mu_{ni} = c_{ni}^2 \delta_{X_{ni}} :$$

Recall: Z_{ni} is measurelike if

$$(Z_{ni}(f) - Z_{ni}(f))^2 \leq \int (f-g)^2 d\mu_{ni}$$

for possibly random measures μ_{ni}.

Example: weighted empirical processes

Suppose that $\mathscr{F}$ satisfies the uniform entropy condition and that

$$\max_{1 \leq i \leq m_n} |c_{ni}| \to 0$$

and

$$\sum_{i=1}^{m_n} c_{n,i}^2 P_{ni} \leq P$$

for P a probability measure with $P^* F^2 < \infty$.

Then (under measurability conditions) Lemma 2.11.6 + Theorem 2.11.1 now yield that $\mathbb{G}_n$ converges weakly in $\ell^\infty(\mathscr{F})$ to a Gaussian process provided we have the required marginal convergence.

Now we turn to CLT's for processes based on the bracketing entropy condition on $\mathscr{F}$.

In particular we will be proving two generalization of Theorem 2.5.6.

Theorem

Any class $\mathscr{F}$ of measurable functions satisfying

$$\int_0^\infty \sqrt{\log N_{[]}(\varepsilon, \mathscr{F}, L_2(P))}d\varepsilon < \infty$$

is P-Donsker.

Generalizations will be in two directions:

- generalization in terms non-iid setting (Theorem 2.11.9)
- weakening the entropy conditions in terms of existence of certain Gaussian processes

Generalizations will be in two directions:

- generalization in terms non-iid setting (Theorem 2.11.9)
- weakening the entropy conditions in terms of existence of certain Gaussian processes (Theorem 2.11.11)

We will be looking at the bracketing numbers $N_\varepsilon \equiv N_{[]}(\varepsilon, \mathscr{F}, L_2^n)$, ie the minimum number of sets $\mathscr{F}_{\varepsilon j}^n$ such that

$$\mathscr{F} = \cup_{j=1}^{N_\varepsilon} \mathscr{F}_{\varepsilon j}^n$$

and

$$\sum_{i=1}^{m_n} E^* \sup_{f,g \in \mathscr{F}_{\varepsilon j}^n} |Z_{ni}(f) - Z_{ni}(g)|^2 \leq \varepsilon^2.$$

Theorem (2.11.9 - Bracketing entropy)

Let $Z_{n1}, \dots, Z_{nm_n}$ independent stochastic processes with finite second moments indexed by $(\mathscr{F}, \rho)$, a totally bounded semimetric space. Suppose

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||_{\mathscr{F}} \{||Z_{ni}||_{\mathscr{F}} > \eta\} \to 0 \text{ for all } \eta > 0,$$

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \to 0$$

for all $\delta_n \to 0$ and entropy condition

$$\int_0^{\delta_n} \sqrt{\log N_{[]}(\varepsilon, \mathscr{F}, L_2)} d\varepsilon \to 0.$$

Then, $\sum_{i=1}^{m_n} (Z_{ni} - \mathbb{E}Z_{ni})$ is asymptotically tight in $\ell^\infty(\mathscr{F})$ provided we have marginal convergence.

Proof of both Theorem 2.11.9 (and 2.11.11) are based on a chaining argument similar to that of 2.5.6.

The argument is quite long / intensive from a bookkeeping perspective so I will leave them to be DBY.

The entropy condition of the previous theorem can be refined further.

Two definitions: a semimetric ρ is called Gaussian if it is of the form

$$\rho(f,g) = \left(\mathbb{E}(G(f) - G(g))^2\right)^{1/2}$$

where G is some tight, centered Gaussian in $\ell^\infty(T)$.

Call a semimetric ρ *Gaussian-dominated* if it is bounded above on $\mathscr{F} \times \mathscr{F}$ by a Gaussian semimetric.

It can be shown that any semimetric on $\mathscr{F} \times \mathscr{F}$ is Gaussian dominated if

$$\int_0^\infty \sqrt{\log N(\varepsilon, \mathscr{F}, \rho)}\, d\varepsilon < \infty.$$

(this is problem 2.11.4)

Theorem (2.11.11)

Let $Z_{n1}, \ldots, Z_{nm_n}$ independent stochastic processes with finite second moments indexed by $\mathscr{F}$. Let

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||_{\mathscr{F}} \{||Z_{ni}||_{\mathscr{F}} > \eta\} \to 0 \text{ for all } \eta > 0.$$

Suppose further that there exists a Gaussian dominated semimetric ρ such that

$$\sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \leq \rho^2(f,g) \text{ for all } (f,g) \in \mathscr{F} \times \mathscr{F}$$

and

$$\sup_{t>0} \sum_{i=1}^{m_n} t^2 P^* \left(\sup_{f,g \in B_\varepsilon} |Z_{ni}(f) - Z_{ni}(g)| > t \right) \leq \varepsilon^2$$

for every ρ-ball of radius less than ε $B_\varepsilon \subset \mathscr{F}$ and for every n.

Then, $\sum_{i=1}^{m_n} (Z_{ni} - \mathbb{E} Z_{ni})$ is asymptotically tight in $\ell^\infty(\mathscr{F})$, and hence converges weakly provided we have marginal convergence.

Example (Corollary?): Jain-Marcus Theorem

Let $Z_{n1}, \dots, Z_{nm_n}$ independent stochastic processes with finite second moments indexed by $\mathscr{F}$ such that

$$|Z_{ni}(f) - Z_{ni}(g)| \leq M_{ni}\rho(f,g)$$

for independent random variables $M_{n1}, \dots, M_{n,m_n}$ and a semimetric ρ such that

$$\int_0^\infty \sqrt{\log N(\varepsilon, \mathscr{F}, \rho)} d\varepsilon < \infty,$$

and

$$\sum_{i=1}^{m_n} \mathbb{E} M_{ni}^2 = \mathscr{O}(1).$$

If the Lindeberg condition is satisfied, $\sum_{i=1}^{m_n} (Z_{ni} - \mathbb{E} Z_{ni})$ is asymptotically tight in $\ell^\infty(\mathscr{F})$, and hence converges weakly provided we have marginal convergence.

Example: Stochastic processes indexed by the unit interval

Let $Z_1, \ldots, Z_n$ independent stochastic processes indexed by $\mathscr{F} = [0,1]$, $||Z_i||_{\mathscr{F}} \leq 1$ and

$$\mathbb{E}|Z_i(f) - Z_i(g)| \leq K|f-g|$$

for independent random variables for some constant K. Then, $n^{-1/2}\sum_{i=1}^{m_n}(Z_i - \mathbb{E}Z_i)$ converges weakly in $\ell^\infty(\mathscr{F})$.

We can derive this using Theorem 2.11.11 for $Z_{ni} = n^{-1/2}Z_i$.

Example: Stochastic processes indexed by the unit interval

- Lindeberg condition is trivially satisfied.
- For every ε-sized interval

$$\sum_{i=1}^{m_n} \mathbb{E}\left(Z_{ni}(f) - Z_{ni}(g)\right)^2 \leq \rho^2(f,g)$$

  for $\rho^2(f,g) = 2K|f-g| \leq 2K\varepsilon$. This means it satisfies the conditions of 2.11.11.
- This $\rho^2(f,g) = 2K|f-g|$ has a finite entropy integral

$$\int_0^\infty \sqrt{\log N(\varepsilon, \mathscr{F}, \rho)}\, d\varepsilon < \infty$$

  so it is Gaussian dominated.

2.11.3: Classes of Functions Changing with n

Consider now

- $X_1, \ldots, X_n$ on common probability space $(\mathscr{X}, \mathscr{A})$
- $x \mapsto f_{n,t}(x)$ functions from $\mathscr{X}$ to $\mathbb{R}$ for $n \in \mathbb{N}$ and $t \in T$
- T a totally bounded semimetric space with semimetric ρ.

We wish to derive conditions for the stochastic processes

$$\left\{ n^{-1/2} \sum_{i=1}^{n} (f_{n,t}(X_i) - Pf_{n,t}) \;:\; t \in T \right\}$$

to converge weakly in $\ell^\infty(T)$.

2.11.3: Classes of Functions Changing with n

We can view these as empirical processes

$$\mathbb{G}_n f_{t,n} := n^{-1/2} \sum_{i=1}^{n} (f_{n,t}(X_i) - P f_{n,t})$$

indexed by $\mathscr{F}_n = \{f_{n,t} : t \in T\}$, so the function class changes with n.

This fits into our earlier framework upon setting

$$Z_{ni}(t) = f_{n,t}(X_i)/\sqrt{n}$$

which gives

$$\mathbb{G}_n = \sum_{i=1}^{n} (Z_{ni}(t) - \mathbb{E} Z_{ni}(t)).$$

2.11.3: Classes of Functions Changing with n

Again, the weak convergence theorems come in two flavours: uniform entropy and bracketing entropy.

For both types of entropy conditions, we assume

- There exists envelope function F_n for $\mathscr{F}_n$, with $P^* F_n^2 = \mathscr{O}(1)$.
- $P^* F_n^2 \{F_n > \eta\sqrt{n}\} \to 0$ for all $\eta > 0$.
- $\sup_{\rho(s,t)<\delta_n} P(f_{n,s} - f_{n,t})^2 \to 0$ for every $\delta_n \searrow 0$.

2.11.3: Uniform entropy

Theorem (2.11.22)

For each n, let $\mathscr{F}_n = \{f_{n,t} : t \in T\}$ *be a class of measurable functions such that* $\mathscr{F}_{n,\delta} := \{f_{n,s} - f_{n,t} : \rho(s,t) < \delta\}$ *and* $\mathscr{F}^2_{n,\delta}$ *are P-measurable for every* $\delta > 0$*. If the assumptions on the previous slide hold, as well as*

$$\sup_Q \int_0^{\delta_n} \sqrt{\log N(\varepsilon \|F_n\|_{Q,2}, \mathscr{F}_n, L_2(Q))}\, d\varepsilon \to 0 \ \text{ for every } \delta_n \searrow 0,$$

Then, $\{\mathbb{G}_n f_{n,t} : t \in T\}$ *is asymptotically tight in* ℓ^∞ *and converges in distribution to a Gaussian process provided the sequence of covariance functions* $P f_{n,s} f_{n,t} - P f_{n,s} P f_{n,t}$ *converges pointwise on* $T \times T$*.*

Proof 2.11.22

We aim at applying the earlier proved Theorem 2.11.1:

Theorem (2.11.1)

Assume

$$\sum_{i=1}^{m_n} E^*||Z_{ni}||^2_{\mathscr{F}}\{||Z_{ni}||_{\mathscr{F}} > \varepsilon\} \to 0 \textit{ for all } \varepsilon > 0,$$

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \to 0,$$

for all $\delta_n \to 0$ *and entropy condition*

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)}\, d\varepsilon \xrightarrow{P^*} 0.$$

Then, $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ *is asymptotically* ρ*-equicontinuous. If the covariance of the marginals converges pointwise in* $\mathscr{F} \times \mathscr{F}$*, we have* $\sum_{i=1}^{m_n} Z_{ni} - \mathbb{E}Z_{ni}$ *converges weakly in* $\ell^\infty(\mathscr{F})$*.*

Proof 2.11.22

We start by noting that the Lindeberg assumption is satisfied by the envelope condition $P^* F_n^2\{F_n > \eta\sqrt{n}\} \to 0$ for all $\eta > 0$.

Next, observe that $\sup_{\rho(s,t)<\delta_n} P(f_{n,s} - f_{n,t})^2 \to 0$ for every $\delta_n \searrow 0$ implies

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}\left(Z_{ni}(f) - Z_{ni}(g)\right)^2 \to 0.$$

Proof 2.11.22

Next, observe that the random semimetric of 2.11.1 satisfies

$$d_n^2(s,t) = \frac{1}{n}\sum_{i=1}(f_{n,s} - f_{n,t})^2(X_i) = \mathbb{P}_n(f_{n,s} - f_{n,t})^2,$$

so

$$N(\varepsilon, T, d_n) = N(\varepsilon, T, L_2(\mathbb{P}_n)).$$

Proof 2.11.22

Now, $N(\varepsilon, T, d_n) = N(\varepsilon, T, d_n)$ means that the entropy condition assumed in 2.11.22;

$$\sup_Q \int_0^{\delta_n} \sqrt{\log N(\varepsilon ||F_n||_{Q,2}, \mathscr{F}_n, L_2(Q))} d\varepsilon \to 0 \text{ for every } \delta_n \searrow 0,$$

implies the entropy condition of 2.11.1

$$\int_0^{\delta_n} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)} d\varepsilon \overset{P^*}{\to} 0.$$

Proof 2.11.22

We may assume wlog $F \geq 1$. Now,

$$\int_0^{\frac{\delta_n}{||F_n||_{\mathbb{P}_n,2}}} \sqrt{\log N(\varepsilon, \mathscr{F}, d_n)}d\varepsilon = \int_0^{\delta_n} \sqrt{\log N(\varepsilon||F_n||_{\mathbb{P}_n,2}, T, d_n)}d\varepsilon||F_n||_{\mathbb{P}_n,2}$$
$$= \int_0^{\delta_n} \sqrt{\log N(\varepsilon||F_n||_{\mathbb{P}_n,2}, \mathscr{F}_n, L_2(\mathbb{P}_n))}d\varepsilon||F_n||_{\mathbb{P}_n,2} \to 0$$

as

$$\sup_Q \int_0^{\delta_n} \sqrt{\log N(\varepsilon||F_n||_{Q,2}, \mathscr{F}_n, L_2(Q))}d\varepsilon \to 0 \text{ for every } \delta_n \searrow 0.$$

2.11.3: Classes of Functions Changing with n

Next, we turn to a convergence theorem based on bracketing entropy.

We still assume:

- $X_1, \ldots, X_n$ on common probability space $(\mathscr{X}, \mathscr{A})$
- $x \mapsto f_{n,t}(x)$ functions from $\mathscr{X}$ to $\mathbb{R}$ for $n \in \mathbb{N}$ and $t \in T$
- T a totally bounded semimetric space with semimetric ρ.

We wish to derive conditions for the stochastic processes

$$\left\{ n^{-1/2} \sum_{i=1}^{n} (f_{n,t}(X_i) - Pf_{n,t}) \; : \; t \in T \right\}$$

to converge weakly in $\ell^\infty(T)$.

2.11.3: Bracketing entropy

Theorem (2.11.23)

For each n, let $\mathscr{F}_n = \{f_{n,t} : t \in T\}$ be a class of measurable functions and suppose the assumptions on the previous slide hold, as well as

$$\int_0^{\delta_n} \sqrt{\log N_{[]}(\varepsilon \|F_n\|_{P,2}, \mathscr{F}_n, L_2(P))}\, d\varepsilon \to 0 \quad \textit{for every } \delta_n \searrow 0,$$

Then, $\{\mathbb{G}_n f_{n,t} : t \in T\}$ is asymptotically tight in ℓ^∞ and converges in distribution to a Gaussian process prodided the sequence of covariance functions $Pf_{n,s}f_{n,t} - Pf_{n,s}Pf_{n,t}$ converges pointwise on $T \times T$.

Theorem (2.11.9 - Bracketing CLT)

Let $Z_{n1}, \ldots, Z_{nm_n}$ independent stochastic processes with finite second moments indexed by $(\mathscr{F}, \rho)$, a totally bounded semimetric space. Suppose

$$\sum_{i=1}^{m_n} E^* ||Z_{ni}||_{\mathscr{F}} \{||Z_{ni}||_{\mathscr{F}} > \eta\} \to 0 \text{ for all } \eta > 0,$$

$$\sup_{\rho(f,g)<\delta_n} \sum_{i=1}^{m_n} \mathbb{E}(Z_{ni}(f) - Z_{ni}(g))^2 \to 0$$

for all $\delta_n \to 0$ and entropy condition

$$\int_0^{\delta_n} \sqrt{\log N_{[]}(\varepsilon, \mathscr{F}, L_2^n)} d\varepsilon \to 0 \ \text{ for all } \delta_n \searrow 0$$

Then, $\sum_{i=1}^{m_n} (Z_{ni} - \mathbb{E}Z_{ni})$ is asymptotically tight in $\ell^\infty(\mathscr{F})$ provided we have marginal convergence.

Thank you for listening!