STAT6018 Research Frontiers in Data Science

Topic II: Introduction to empirical process theory

Yu Gu, PhD
Assistant Professor

Department of Statistics & Actuarial Science
The University of Hong Kong

Table of Contents

Table of Contents

Glivenko-Cantelli (GC) class

Definition 1 (GC class)

A function class $\mathcal{F}$ is called P-GC if

$$\|\mathbb{P}_n - P\|_{\mathcal{F}} \xrightarrow{a.s.} 0$$

under the probability measure P.

- $\|Q\|_{\mathcal{F}} = \sup_{f \in \mathcal{F}} |Qf|$
- uniform almost sure convergence across $\mathcal{F}$

GC theorem with bracketing

Bracket number $N_{[]}(\epsilon, \mathcal{F}, \|\cdot\|)$:

- minimum number of brackets $[\ell, u]$ with $\|\ell - u\| < \epsilon$ needed to cover $\mathcal{F}$
- entropy with bracketing: $\log N_{[]}(\epsilon, \mathcal{F}, \|\cdot\|)$

Theorem 2 (GC with bracketing)

Let $\mathcal{F}$ be a class of P-measurable functions such that

$$N_{[]}(\epsilon, \mathcal{F}, L_1(P)) < \infty, \qquad \text{for every } \epsilon > 0.$$

Then $\mathcal{F}$ is P-GC.

GC theorem with bracketing (cont.)

Proof.

For every $f \in [\ell_i, u_i]$, we have

$$\begin{cases} (\mathbb{P}_n - P)f \le \mathbb{P}_n u_i - P\ell_i \le (\mathbb{P}_n - P)u_i + \|u_i - \ell_i\|_{L_1(P)} \\ (\mathbb{P}_n - P)f \ge \mathbb{P}_n \ell_i - P u_i \ge (\mathbb{P}_n - P)\ell_i - \|u_i - \ell_i\|_{L_1(P)} \end{cases}$$

Thus,

$$\begin{cases} \sup_{f\in\mathcal{F}}(\mathbb{P}_n - P)f \le \max_i(\mathbb{P}_n - P)u_i + \epsilon \overset{a.s.}{\to} \epsilon \\ \inf_{f\in\mathcal{F}}(\mathbb{P}_n - P)f \ge \min_i(\mathbb{P}_n - P)\ell_i - \epsilon \overset{a.s.}{\to} -\epsilon \end{cases} \quad \text{(by SLLN)}$$

$$\Rightarrow \limsup_n \|\mathbb{P}_n - P\|_{\mathcal{F}} \le \epsilon \text{ almost surely.}$$

Letting $\epsilon \downarrow 0$ yields the desired result. □

GC theorem without bracketing

Covering number $N(\epsilon, \mathcal{F}, \|\cdot\|)$:

- minimum number of balls $B(f;\epsilon) := \{g : \|g - f\| \leq \epsilon\}$ needed to cover $\mathcal{F}$
- entropy without bracketing: $\log N(\epsilon, \mathcal{F}, \|\cdot\|)$

Envelope function F: $|f(x)| \leq F(x)$ for every $x \in \mathcal{X}$ and $f \in \mathcal{F}$

Theorem 3 (GC without bracketing)

Let $\mathcal{F}$ be a class of P-measurable functions with envelope F such that $PF < \infty$. Let $\mathcal{F}_M$ be the class of functions $f\mathbb{1}\{F \leq M\}$ when f ranges over $\mathcal{F}$. Then $\mathcal{F}$ is P-GC if and only if

$$n^{-1} \log N(\epsilon, \mathcal{F}_M, L_1(\mathbb{P}_n)) \xrightarrow{p} 0, \quad \forall \epsilon, M > 0.$$

GC theorem without bracketing (cont.)

Symmetrization (Theorem 1.26):

$$E\left\|\mathbb{P}_n - P\right\|_{\mathcal{F}} \leq 2E\left\|\mathbb{P}_n^o\right\|_{\mathcal{F}}$$

Proof of sufficiency.

$$\begin{aligned} E\left\|\mathbb{P}_n - P\right\|_{\mathcal{F}} &\leq 2E_X E_\varepsilon \left\|\frac{1}{n}\sum_{i=1}^n \varepsilon_i f(X_i)\right\|_{\mathcal{F}} && \text{(symmetrization)} \\ &\leq 2E_X E_\varepsilon \left\|\frac{1}{n}\sum_{i=1}^n \varepsilon_i f(X_i)\right\|_{\mathcal{F}_M} + 2P[F\mathbb{1}\{F > M\}] && \text{(triangle inequality)} \end{aligned}$$

For sufficiently large M, $P[F\mathbb{1}\{F > M\}]$ is arbitrarily small.

GC theorem without bracketing (cont.)

Maximal inequality for Rademacher linear combinations (Corollary 1.25):

$$E \max_{1\le i\le N} |\xi_i| \le C\sqrt{\log N} \max_{1\le i\le N} \|a^{(i)}\|$$

Proof of sufficiency (cont.)

Let $\mathcal{G}$ denote the ϵ-cover associated with $N(\epsilon, \mathcal{F}_M, L_1(\mathbb{P}_n))$. For any $f \in \mathcal{F}_M$,

$$\begin{aligned}
\left|\frac{1}{n}\sum_{i=1}^n \varepsilon_i f(X_i)\right| &\le \left|\frac{1}{n}\sum_{i=1}^n \varepsilon_i g(X_i)\right| + \left|\frac{1}{n}\sum_{i=1}^n \varepsilon_i [f(X_i) - g(X_i)]\right| \\
&\le \left\|\frac{1}{n}\sum_{i=1}^n \varepsilon_i g(X_i)\right\|_{\mathcal{G}} + \epsilon \\
&\le C\sqrt{\frac{\log N(\epsilon, \mathcal{F}_M, L_1(\mathbb{P}_n))}{n}} \max_{g\in\mathcal{G}} \sqrt{\mathbb{P}_n g^2} + \epsilon \quad \text{(maximal inequality)} \\
&\xrightarrow{p} \epsilon
\end{aligned}$$

GC theorem without bracketing (cont.)

Proof of sufficiency (cont.)

Letting $\epsilon \downarrow 0$ yields $\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i f(X_i)\|_{\mathcal{F}_M} \xrightarrow{p} 0$. Since $\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i f(X_i)\|_{\mathcal{F}_M} \leq M$, it follows by the dominated convergence theorem that $E_X E_\varepsilon \left\|\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i f(X_i)\right\|_{\mathcal{F}_M} \to 0$.
Thus, we conclude that $E\|\mathbb{P}_n - P\|_{\mathcal{F}} \to 0$.
By Lemma 2.4.5 of VW, $\|\mathbb{P}_n - P\|_{\mathcal{F}}$ is a reverse sub-martingale, thus converges almost surely to a constant, which must be 0 by the convergence in mean. □

GC theorem with uniform covering

Corollary 4

Let $\mathcal{F}$ be a class of P-measurable functions with envelope F such that $PF < \infty$. Then $\mathcal{F}$ is P-GC if

$$\sup_Q N(\epsilon \|F\|_{L_1(Q)}, \mathcal{F}, L_1(Q)) < \infty, \quad \forall \epsilon > 0,$$

where the supremum is over all probability measures Q with $0 < QF < \infty$.

Proof.

Assume that $PF > 0$ (otherwise the result is trivial). There exists an $\eta \in (0, \infty)$ such that $1/\eta < \mathbb{P}_n F < \eta$ for all n large enough. For any $\epsilon > 0$, there exists a K_ϵ such that with probability 1,

$$\log N(\epsilon\eta, \mathcal{F}, L_1(\mathbb{P}_n)) \leq \log N(\epsilon \mathbb{P}_n F, \mathcal{F}, L_1(\mathbb{P}_n)) \leq K_\epsilon$$

for all n large enough. Thus, for any $\epsilon, M > 0$,

$$\log N(\epsilon, \mathcal{F}_M, L_1(\mathbb{P}_n)) \leq \log N(\epsilon, \mathcal{F}, L_1(\mathbb{P}_n)) = O_p(1).$$

The desired result follows by Theorem 3. □

Table of Contents

Donsker class

Definition 5

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

$$\mathbb{G}_n = \sqrt{n}(\mathbb{P}_n - P) \xrightarrow{d} \mathbb{G},$$

where $\mathbb{G}$ is a tight[a] random element in $\ell^\infty(\mathcal{F})$.

[a] $\Leftrightarrow \forall \epsilon > 0, \exists$ *a compact set* $V_\epsilon \in \ell^\infty(\mathcal{F})$ *s.t.* $P(\mathbb{G}f \in V_\epsilon) > 1 - \epsilon$, *for all* $f \in \mathcal{F}$.

- The multivariate CLT ensures marginal convergence of $\mathbb{G}_n$:

$$(\mathbb{G}_n f_1, \dots, \mathbb{G}_n f_k) \xrightarrow{d} N(0, \Sigma), \quad \forall (f_1, \dots, f_k) \in \mathcal{F}$$

- It follows that $\{\mathbb{G}f : f \in \mathcal{F}\}$ must be a mean-zero Gaussian process with covariance function $E\{\mathbb{G}f_1 \mathbb{G}f_2\} = \Sigma(f_1, f_2)$.
- This and tightness determine $\mathbb{G}$ to be a *P-Brownian bridge* in $\ell^\infty(\mathcal{F})$[1].

[1] *By Lemma 1.5.3 of VW.*

Donsker with asymptotic equi-continuity

To prove the Donsker property by definition, we usually need to check:

- Marginal convergence (guaranteed by multivariate CLT)
- Tightness of the limiting process $\mathbb{G}$, which is equivalent to both of the following:
  - *Total boundedness* of $(\mathcal{F}, d)$, i.e., $N(\epsilon, \mathcal{F}, d) < \infty$ for every $\epsilon > 0$
  - *Asymptotic equicontinuity* of $(\mathcal{F}, d)$, i.e., for every $\epsilon > 0$,

$$\lim_{\delta \to 0} \limsup_{n \to \infty} \mathbb{P}^* \left(\sup_{d(f,g) \leq \delta; f,g \in \mathcal{F}} |\mathbb{G}_n(f - g)| > \epsilon \right) = 0$$

The semi-metric d is usually chosen as the $L_2(P)$ distance, and $\mathbb{P}^*$ is outer probability[2], which behaves like usual probabilities in most cases.

[2] *the infimum of the probabilities of all measurable sets that contain the event.*

Donsker with asymptotic equi-continuity (cont.)

This is formally stated in the following theorem, which follows immediately from the result of weak convergence of stochastic processes.

Theorem 6 (Donsker with asymptotic equi-continuity)

Let $\mathcal{F}$ be a class of measurable, square-integrable functions from $\mathcal{X}$ to $\mathbb{R}$ such that $\sup_{f\in\mathcal{F}} |f(x) - Pf| < \infty$, $\forall x \in \mathcal{X}$. Then $\{\mathbb{G}_n f : f \in \mathcal{F}\}$ converges weakly to a tight random element if and only if there exists a semi-metric $d(\cdot,\cdot)$ on $\mathcal{F}$ such that $(\mathcal{F}, d)$ is totally bounded and for every $\epsilon > 0$,

$$\lim_{\delta\to 0} \limsup_{n\to\infty} \mathbb{P}^* \left(\sup_{d(f,g)\le\delta; f,g\in\mathcal{F}} |\mathbb{G}_n(f-g)| > \epsilon \right) = 0.$$

Bracketing entropy integral

- In many cases, bracketing numbers grow to infinity as $\epsilon \downarrow 0$.
- Sufficient condition for Donsker class: bracketing numbers do not grow too fast with $1/\epsilon$
- **Bracketing entropy integral** measures the speed of growth:

$$J_{[]}(\delta, \mathcal{F}, L_r(P)) := \int_0^\delta \sqrt{\log N_{[]}(\epsilon, \mathcal{F}, L_r(P))} d\epsilon$$

- The above integral coverges when the bracketing entropy grows with order slower than $1/\epsilon^2$.

Donsker theorem with bracketing

Theorem 7 (Donsker with bracketing)

Suppose that $\mathcal{F}$ is a class of measurable functions satisfying

$$J_{[]}(1, \mathcal{F}, L_2(P)) < \infty.$$

Then $\mathcal{F}$ is P-Donsker.

Donsker theorem with bracketing (cont.)

The proof of Theorem 7 uses the following maximal inequality:

Lemma 8 (Maximal inequality)

For any class $\mathcal{F}$ of measurable functions $f : \mathcal{X} \to \mathbb{R}$ satisfying $Pf^2 < \delta^2$,

$$E^* \|\mathbb{G}_n\|_{\mathcal{F}} \lesssim J_{[]}(\delta, \mathcal{F}, L_2(P)) + \sqrt{n} P^*[F \mathbb{1}\{F > \sqrt{n} a(\delta)\}],$$

where $x \lesssim y$ means $x \leq cy$ for some constant $c > 0$, F is an envelope function of $\mathcal{F}$, and $a(\delta) = \delta / \sqrt{\log N_{[]}(\epsilon, \mathcal{F}, L_2(P))}$.

See Lemma 19.34 of van der Vaart (1998)[3] for the proof.

[3] *van der Vaart, A. W. (1998). Asymptotic statistics, volume 3 of Cambridge Series in Statistical and Probabilistic Mathematics. Cambridge University Press, Cambridge.*

Donsker theorem with bracketing (cont.)

Proof of Theorem 7.

$\forall \epsilon > 0$, $N_{[]}(\epsilon, \mathcal{F}, L_2(P))$ is finite, so $(\mathcal{F}, \|\cdot\|_{L_2(P)})$ is totally bounded.
Define $\mathcal{G} = \{f - g : f, g \in \mathcal{F}\}$. It is easy to see that $G = 2F$ is an envelope for $\mathcal{G}$ and $N_{[]}(2\epsilon, \mathcal{G}, L_2(P)) \leq N_{[]}^2(\epsilon, \mathcal{F}, L_2(P))$.
Let $\mathcal{G}_\delta = \{f - g : f, g \in \mathcal{F}, \|f - g\|_{L_2(P)} \leq \delta\}$. By Lemma 8, there exists a finite number $a(\delta) = \delta / \sqrt{\log N_{[]}(\epsilon, \mathcal{G}_\delta, L_2(P))}$ s.t.

$$\begin{aligned}
\mathbb{E}^* \|\mathbb{G}_n\|_{\mathcal{G}_\delta} &\lesssim J_{[]}(\delta, \mathcal{G}_\delta, L_2(P)) + \sqrt{n} P[G \mathbb{1}\{G > a(\delta)\sqrt{n}\}] \\
&\leq J_{[]}(\delta, \mathcal{G}, L_2(P)) + \sqrt{n} P[G \mathbb{1}\{G > a(\delta)\sqrt{n}\}].
\end{aligned}$$

The second term on RHS is bounded by $a(\delta)^{-1} P[G^2 \mathbb{1}\{G > a(\delta)\sqrt{n}\}]$ and hence converges to 0 as $n \to \infty$ for every δ.
By assumption, $J_{[]}(\delta, \mathcal{G}, L_2(P)) \lesssim J_{[]}(\delta, \mathcal{F}, L_2(P)) \to 0$ as $\delta \to 0$.
Thus, by Markov's inequality, the asymptotic equi-continuity condition holds.
The desired result then follows by Theorem 6. □

Uniform entropy integral

Like GC theorems, an alternative sufficient condition for Donsker property is based on the **uniform entropy integral**:

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

where F is an envelope of $\mathcal{F}$, and the supremum is taken over all finitely discrete probability measures Q with $QF^2 > 0$.

Donsker theorem without bracketing

Theorem 9 (Donsker without bracketing)

Suppose that $\mathcal{F}$ is a pointwise-measurable class of measurable functions satisfying $PF^2 < \infty$ and

$$J(1, \mathcal{F}, F) < \infty.$$

Then $\mathcal{F}$ is P-Donsker.

The pointwise-measurable condition suffices that there exists a countable collection $\mathcal{G}$ of functions such that each f is the pointwise limit of a sequence g_m in $\mathcal{G}$ (see Example 2.3.4 of VW for details).

Donsker theorem without bracketing (cont.)

The proof of Theorem 9 uses the following maximal inequality:

Lemma 10 (Maximal inequality)

Suppose $0 < \|F\|_{L_2(P)} < \infty$, *let* σ^2 *be any positive constant s.t.* $\sup_{f \in \mathcal{F}} Pf^2 \leq \sigma^2 \leq \|F\|^2_{L_2(P)}$. *Let* $\delta = \sigma / \|F\|_{L_2(P)}$ *and* $B = \sqrt{E \max_{1 \leq i \leq n} F^2(X_i)}$. *Then,*

$$E\|\mathbb{G}_n\|_{\mathcal{F}} \lesssim J(\delta, \mathcal{F}, F)\|F\|_{L_2(P)} + \frac{BJ^2(\delta, \mathcal{F}, F)}{\delta^2 \sqrt{n}}.$$

See Chernozhukov et al. (2014)[4] for the proof.

[4] *Chernozhukov, V., Chetverikov, D., and Kato, K. (2014). Gaussian approximation of suprema of empirical processes. Ann. Statist., 42(4):1564–1597.*

Donsker theorem without bracketing (cont.)

Proof of Theorem 9.

We first show that $(\mathcal{F}, \|\cdot\|_{L_2(P)})$ is totally bounded. For any fixed $\epsilon > 0$, there exist $f_1, \ldots, f_N \in \mathcal{F}$ s.t. $P(f_i - f_j)^2 > \epsilon^2 PF^2$, for every $i \neq j$. By LLN,

$$
\begin{aligned}
&\mathbb{P}_n(f_i - f_j)^2 \overset{a.s.}{\to} P(f_i - f_j)^2 \quad \text{and} \quad \mathbb{P}_n F^2 \overset{a.s.}{\to} PF^2 \\
\Rightarrow\ &\mathbb{P}_n(f_i - f_j)^2 > \epsilon^2 PF^2 \quad \text{and} \quad 0 < \mathbb{P}_n F^2 < 2PF^2, \quad \text{for some large } n \\
\Rightarrow\ &\mathbb{P}_n(f_i - f_j)^2 > \epsilon^2 \mathbb{P}_n F^2/2 \\
\Rightarrow\ &N \leq D(\epsilon\|F\|_{L_2(\mathbb{P}_n)}/\sqrt{2}, \mathcal{F}, L_2(P_n)) < \infty. \qquad \text{(by assumption)}
\end{aligned}
$$

Choosing $N = D(\epsilon\|F\|_{L_2(P)}, \mathcal{F}, L_2(P))$ yields the total boundedness.

Donsker theorem without bracketing (cont.)

Proof of Theorem 9 (cont.)

To verify the asymptotic equi-continuity condition, it suffices to show that

$$\lim_{\delta\to 0}\limsup_{n\to\infty} E\|\mathbb{G}_n\|_{\mathcal{G}_\delta} = 0, \tag{1}$$

where $\mathcal{G}_\delta = \{f - g : f, g \in \mathcal{F}, \|f - g\|_{L_2(P)} \le \delta\}$.
We observe that $\mathcal{G}_\delta$ has envelope $2F$ and

$$\begin{aligned}\sup_Q N(\epsilon\|2F\|_{L_2(Q)}, \mathcal{G}_\delta, L_2(Q)) &\le \sup_Q N(\epsilon\|2F\|_{L_2(Q)}, \mathcal{G}_\infty, L_2(Q))\\ &\le \sup_Q N^2(\epsilon\|F\|_{L_2(Q)}, \mathcal{F}, L_2(Q)),\end{aligned}$$

which leads to $J(\epsilon, \mathcal{G}_\delta, 2F) \lesssim J(\epsilon, \mathcal{F}, F)$ for all $\epsilon > 0$.

Donsker theorem without bracketing (cont.)

Proof of Theorem 9 (cont.)

Hence by Lemma 10 with $\sigma = \delta$ and envelope $2F$, we have

$$E\|\mathbb{G}_n\|_{\mathcal{G}_\delta} \le C\left\{J(\delta', \mathcal{F}, F)\|F\|_{L_2(P)} + \frac{B_n J^2(\delta', \mathcal{F}, F)}{\delta'^2\sqrt{n}}\right\},$$

where $\delta' = \sigma/(2\|F\|_{L_2(P)})$ and $B_n = 2\sqrt{E max_{1\le i\le n} F^2(X_i)}$.
Since $PF^2 < \infty$, $B_n = o(\sqrt{n})$. Thus, $\forall \eta > 0$, we can choose δ small s.t.

$$\limsup_{n\to\infty} E\|\mathbb{G}_n\|_{\mathcal{G}_\delta} \le C(\|F\|_{L_2(P)} + 1)\eta.$$

Hence, the asymptotic equi-continuity condition in (1) is satisfied and we complete the proof. □

Discussion

- Theorems 7 and 9 are based on finite bracketing entropy integral and uniform entropy integral, respectively.
- Although bracketing entropy integral involves only the true probability measure P, this gain is offset by the fact that bracketing numbers are usually larger than covering numbers.
- Thus, these two sufficient conditions for Donsker classes are not comparable.

A general Donsker theorem

Define $L_{2,\infty}(P)$-norm as $\|f\|_{L_{2,\infty}(P)} = \sup_{t>0}\{t^2 P(|f| > t)\}^{1/2}$. Note that $\|f\|_{L_{2,\infty}(P)} \leq \|f\|_{L_2(P)}$. The following general Donsker theorem combines the two entropy integrals:

Theorem 11 (General Donsker theorem)

Let $\mathcal{F}$ be a class of measurable functions such that

$$\int_0^1 \sqrt{\log N_{[]}\left(\epsilon, \mathcal{F}, L_{2,\infty}(P)\right)} d\epsilon + \int_0^1 \sqrt{\log N\left(\epsilon, \mathcal{F}, L_2(P)\right)} d\epsilon < \infty.$$

Moreover, assume that the envelope F of $\mathcal{F}$ satisfies a weak second moment, i.e., $t^2 P^*\{F(X) > t\} \to 0$ as $t \to \infty$. Then $\mathcal{F}$ is P-Donsker.

See Theorem 2.5.6 of VW for the proof.

Table of Contents

GC preservation

Theorem 12 (GC preservation)

Suppose that $\mathcal{F}_1, \ldots, \mathcal{F}_k$ are P-GC with $\max_{1 \le j \le k} \|P\|_{\mathcal{F}_j} < \infty$. Then for any continuous transformation $\phi : \mathbb{R}^k \mapsto \mathbb{R}$, the class $\mathcal{H} = \phi(\mathcal{F}_1, \ldots, \mathcal{F}_k)$ is also P-GC provided it has an integrable envelope.

See Theorem 3 of van der Vaart and Wellner (2000)[5] for the proof.

[5] *van der Vaart, A., & Wellner, J. A. (2000). Preservation theorems for Glivenko-Cantelli and uniform Glivenko-Cantelli classes. In High dimensional probability II (pp. 115-133). Boston, MA: Birkhäuser Boston.*

GC preservation (cont.)

Corollary 13

Let $\mathcal{F}$ and $\mathcal{G}$ be P-GC with respective integrable envelopes F and G. Then,

(i) $\mathcal{F} + \mathcal{G}$ *is P-GC.*

(ii) $\mathcal{F} \cdot \mathcal{G}$ *is P-GC if* $P(FG) < \infty$.

(iii) *Any continuous transformation* $\phi(\mathcal{F})$ *is P-GC provided it has an integrable envelope.*

See Corollary 9.27 of Kosorok for the proof.

Closures and convex hulls

For a class $\mathcal{F}$ of measurable functions, define the following operations.

Closure:

$$\overline{\mathcal{F}} = \left\{ f : \mathcal{X} \mapsto \mathbb{R} \,\middle|\, \exists \{f_m\} \in \mathcal{F} \text{ s.t. } f_m \to f \text{ both pointwise and in } L_2(P) \right\}$$

Symmetric convex hull:

$$\text{sconv}\mathcal{F} = \left\{ \sum_{i=1}^{\infty} \lambda_i f_i \,\middle|\, \{f_i\} \in \mathcal{F}, \sum_{i=1}^{\infty} |\lambda_i| \leq 1 \right\}$$

Donsker preservation

Theorem 14 (Donsker preservation)

Let $\mathcal{F}$ be P-Donsker. Then,

(i) *For any $\mathcal{G} \subset \mathcal{F}$, $\mathcal{G}$ is P-Donsker.*

(ii) *$\overline{\mathcal{F}}$ is P-Donsker.*

(iii) *$\text{sconv}\mathcal{F}$ is P-Donsker.*

See Theorems 2.10.1 – 2.10.3 of VW for the proofs.

Donsker preservation (cont.)

The following theorem establishes Donsker preservation under Lipschitz transformations and is one of the most useful preservation results:

Theorem 15 (Donsker preservation under Lipschitz transformations)

Suppose that $\mathcal{F}_1, \ldots, \mathcal{F}_k$ *are Donsker classes with* $\max_{1 \le j \le k} \|P\|_{\mathcal{F}_j} < \infty$. *Consider any Lipschitz transformation* $\phi : \mathbb{R}^k \mapsto \mathbb{R}$ *satisfying*

$$|\phi \circ f(x) - \phi \circ g(x)|^2 \le c^2 \sum_{j=1}^{k} \{f_j(x) - g_j(x)\}^2 ,$$

for every $f, g \in \mathcal{F}_1 \times \cdots \times \mathcal{F}_k$, *every* $x \in \mathcal{X}$, *and some constant* $c < \infty$. *Then the class* $\phi \circ (\mathcal{F}_1, \ldots, \mathcal{F}_k)$ *is Donsker provided* $\phi \circ f$ *is square integrable for at least one* $f \in \mathcal{F}_1 \times \cdots \times \mathcal{F}_k$.

See Theorem 2.10.6 and pages 196 – 198 of VW for the proof.

Donsker preservation (cont.)

Corollary 16

Let $\mathcal{F}$ and $\mathcal{G}$ be Donsker classes. Then,

(i) $\mathcal{F} \cup \mathcal{G}$ *and* $\mathcal{F} + \mathcal{G}$ *are Donsker.*

(ii) *If* $\|P\|_{\mathcal{F} \cup \mathcal{G}} < \infty$, *then the pairwise infima* $\mathcal{F} \wedge \mathcal{G}$ *and the pairwise suprema* $\mathcal{F} \vee \mathcal{G}$ *are Donsker.*

(iii) *If* $\mathcal{F}$ *and* $\mathcal{G}$ *are uniformly bounded, then* $\mathcal{F} \cdot \mathcal{G}$ *is Donsker.*

(iv) *Any Lipschitz continuous transformation* $\phi(\mathcal{F})$ *is Donsker, provided* $\|\phi(f)\|_{L_2(P)} < \infty$ *for at least one* $f \in \mathcal{F}$.

(v) *If* $\|P\|_{\mathcal{F}} < \infty$ *and* g *is a uniformly bounded, measurable function, then* $\mathcal{F} \cdot g$ *is Donsker.*

See Corollary 9.32 of Kosorok for the proof.