

Some Final Comments on Empirical Processes

1 Sufficient conditions for Pollard's Entropy Condition and the Euclidean Property

In practice, instead of verifying directly if the family of functions \mathcal{M} that conforms a stochastic process $v_T(\cdot)$ satisfies the entropy condition, we rely on sufficient conditions that are relatively more easy to verify.

1.1 Theorem 2 in Andrews (1994) –Handbook of Econometrics Chapter–

First, let us define what a function of bounded variation is.

Total variation of a function and functions of bounded variation

Take a metric space (X, d) . A function $\gamma[a, b] \rightarrow X$ is **of bounded variation** (or it has bounded variation in the set $[a, b]$) if $\exists M$ such that for each partition $P = \{a = t_0 < t_1 < \dots < t_n = b\}$ of $[a, b]$,

$$V(\gamma, P) = \sum_{k=1}^n d(\gamma(t_k), \gamma(t_{k-1})) \leq M$$

The **total variation** V_γ of γ is defined by:

$$V_\gamma = \sup \{V(\gamma, P) : P \text{ is a partition of } [a, b]\}$$

A known result says that a function is of bounded variation if and only if it is the difference between two monotonic functions, but bounded variation is a property that is relatively easy to verify directly.

We next define three different classes of functions...

Type I Class of functions

A class of functions \mathcal{F} on \mathcal{W} is type I class if it is of the form

$$\mathcal{F} = \{f : f(w) = h(w'\theta) \forall w \in \mathcal{W}, \text{ some } \theta \in \Theta \subset \mathbb{R}^k \text{ and } h \in V_k\}$$

where V_k is some set of functions from \mathbb{R} to \mathbb{R} with bounded variation less than or equal to $K < \infty$.

Type II Class of functions

A class of functions \mathcal{F} on \mathcal{W} is type II class if it is of the form

$$\mathcal{F} = \{f : f(w) = f(w, \tau) \text{ for some } \tau \in \mathcal{T} \text{ a bounded subset of Euclidean space,}\}$$

and $f(\cdot, \tau)$ is Lipschitz in τ , that is: $|f(\cdot, \tau_1) - f(\cdot, \tau_2)| \leq B(\cdot) \|\tau_1 - \tau_2\| \forall \tau_1, \tau_2 \in \mathcal{T}$ and some function $B(\cdot) : \mathcal{W} \rightarrow \mathbb{R}$.

Type III Class of functions

Type III classes of functions are smooth inside some subset of \mathcal{W} . A class of functions \mathcal{F} on \mathcal{W} is type III class if

- (i) each $f \in \mathcal{F}$ depends on w only through a subvector w_a of dimension $k_a \leq k$,
- (ii) for some $q > k_a/2$, $C < \infty$ and some connected, compact set $\mathcal{W}_a^* \subset \mathcal{W}_a$, each $f \in \mathcal{F}$ satisfies the smoothness condition that $\forall w \in \mathcal{W}$ and $w + h \in \mathcal{W}$,

$$f(w + h) = \sum_{v=0}^{[q]} \frac{1}{v!} B_v(h_a, w_a) + R(h_a, w_a) \quad \text{with} \quad R(h_a, w_a) \leq C \|h_a\|^q$$

where each $B_v(h_a, w_a)$ is homogenous of degree v in h_a and (q, C, \mathcal{W}_a^*) do not depend on f , w or h .

- (iii) for some $K < \infty$ and all $f \in \mathcal{F}$, $f(w) = K \forall w \in \mathcal{W}$ such that $w_a \in \mathcal{W}_a - \mathcal{W}_a^*$.

If every $f \in \mathcal{F}$ has partial derivatives of order $[q]$ on $\mathcal{W}^* = \{w \in \mathcal{W} : w_a \in \mathcal{W}_a^*\}$, then the expression in (ii) would be a Taylor approximation of order $[q]$.

Theorem 2 in Andrews (1994) –Handbook of Econometrics– If \mathcal{F} is a class of functions of type I, II or III, then Pollard’s entropy condition holds with envelope given by:

$$F(\cdot) = \text{Max} \{1, \sup_{f \in \mathcal{F}} |f(\cdot)|\} \text{ for Type I functions,}$$

$$F(\cdot) = \text{Max} \{1, \sup_{f \in \mathcal{F}} |f(\cdot)|, B(\cdot)\} \text{ for Type II functions,}$$

$F(\cdot)$ is some constant function for Type III functions, because they must be uniformly bounded in absolute value.

Comments: Type I classes are useful for dealing with M-estimation problems with non-smooth objective functions. Type III classes are useful for dealing with non-parametric estimation problems (as we will see with Kernel-smoothed nonparametric estimators). We can combine classes of functions that satisfy the entropy condition to construct more complicated classes that also satisfy it. Theorem 3 in Andrews states this:

Theorem 3 in Andrews (1994) –Handbook of Econometrics–

This Theorem illustrates the **stability of the entropy condition that results from mixing and matching classes with this property**. Let \mathcal{G} and \mathcal{G}^* be classes of $r \times s$ matrix-valued functions, and let \mathcal{H} be a class of $s \times u$ -valued functions. Define

$$\mathcal{G} \oplus \mathcal{G}^* = \{g + g^* : g \in \mathcal{G}, g^* \in \mathcal{G}^*\},$$

$$\mathcal{G}\mathcal{H} = \{gh : g \in \mathcal{G}, h \in \mathcal{H}\},$$

$$\mathcal{G} \vee \mathcal{G}^* = \{g \vee g^* : g \in \mathcal{G}, g^* \in \mathcal{G}^*\} \text{ (element-by-element maximum),}$$

$$\mathcal{G} \wedge \mathcal{G}^* = \{g \wedge g^* : g \in \mathcal{G}, g^* \in \mathcal{G}^*\} \text{ (element-by-element minimum),}$$

$$|\mathcal{G}| = \{|g| : g \in \mathcal{G}\},$$

$$\mathcal{G}^{-1} = \{g^{-1} : g \in \mathcal{G}\} \text{ (if } r = s \text{ and } g(\cdot) \text{ is invertible for all } w \in \mathcal{W} \text{).}$$

If \mathcal{G} , \mathcal{G}^* and \mathcal{H} satisfy Pollard’s entropy condition with scalar envelopes G , G^* and H respectively, then so do each of the following classes:

$\mathcal{G} \cup \mathcal{G}^*$ with envelope $G \vee G^*$.

$\mathcal{G} \oplus \mathcal{G}^*$ with envelope $G + G^*$.

\mathcal{GH} with envelope $(G \vee 1)(H \vee 1)$.

$\mathcal{G} \vee \mathcal{G}^*$ with envelope $G \vee G^*$.

$\mathcal{G} \wedge \mathcal{G}^*$ with envelope $G \vee G^*$.

$|\mathcal{G}|$ with envelope G .

These mixing-and-matching results are the consequence of the corresponding stability properties of packing and covering numbers.

2 Final comments

1. Make sure to read and understand the examples in pages 2273-2275 in Andrews (1994) –Handbook of Econometrics–.
2. Remember that what we are after is the property of **stochastic equicontinuity**. We may use this property to “go all the way” and establish a Functional Central Limit Theorem for our empirical process $v_T(\cdot)$ (remember weak-convergence result in the Proposition at the bottom of page 2251 in Andrews), or we may use it to establish asymptotic normality of an M -estimator (see Section 3.2 in Andrews), or a semiparametric estimation problem (see Section 3.4. We will discuss this in more detail when we study semiparametric estimation). See also Section 3.3.
3. The methods we studied (the ones that lead to Pollard’s entropy condition) are called “Combinatorial results” because the machinery used to establish the sufficient conditions (e.g, Types I,II and III classes of functions, Euclidean classes of functions) is based on combinatorial results. See for example Theorem 16 in Pollard (1984), page 19).

4. We did not have time to review the **bracketing methods** for establishing stochastic equicontinuity. The logic behind them is very similar to the combinatorial results we did review: the goal is to see how well we can approximate any function in a class \mathcal{F} with a finite collection of functions from that class. The behavior of the cardinality of such finite collection as we improve the degree of approximation is the key. See Section 5 in Andrews (1994).
5. Please **read Pakes and Pollard (1989)**, you should find it easier to read now. It focuses on Euclidean classes of functions and processes (which are also called VC-classes) via properties of their subgraphs. It is a classic paper, it includes more sufficient conditions, similar to those in Theorem 2 in Andrews (see above). It also illustrates Simulation estimation methodology.
6. The argument we used in the proof of Lemma 3.4 in Pollard (1990) is called “**chaining**” (the specific construction of a sequence that linked any point $t \in T_m$ to the point t_0). In fact, the inequality in the statement of the lemma is called a **chaining inequality**.