Chapter 3.9: Independence Empirical Processes

Weak convergence and Empirical Processes by Aad van der Vaart & Jon Wellner

Lasse Vuursteen

November 16, 2020

Let $(\mathscr{X} \times \mathscr{Y}, \mathscr{A} \times \mathscr{B}, H)$ be a probability space and $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ an iid sample from H. We wish to test whether X_i is independent from Y_i.

Formally, we wish to test

$$H_0 : H = P \times Q,$$

versus

$$H_1 : H \neq P \times Q.$$

So we have a composite null- and alternative hypothesis, where under the null hypothesis X_i has some distribution marginal P, Y_i has some marginal distribution Q and X_i is independent from Y_i.

Let $\mathscr{F}$ and $\mathscr{G}$ classes of real valued measurable functions defined on the sample spaces of X_i and Y_i respectively.

Define the class

$$\mathscr{F} \times \mathscr{G}$$

of real valued measurable functions $f \times g$ with $f \in \mathscr{F}$, $g \in \mathscr{G}$ defined by

$$f \times g(x,y) = f(x)g(y).$$

This is itself a class of real valued measurable functions for the corresponding product sigma algebra.

We will be testing

$$H_0 : ||H - P \times Q||_{\mathscr{F} \times \mathscr{G}} = 0$$

versus

$$H_1 : ||H - P \times Q||_{\mathscr{F} \times \mathscr{G}} > 0.$$

$$\mathbb{H}_n := \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i} \times \delta_{Y_i}$$

where $\delta_{X_i} \times \delta_{Y_i}(f \times g) = f(X_i)g(Y_i)$. Letting $\mathbb{P}_n$ and $\mathbb{Q}_n$ denote the empirical measures for X_i and Y_i respectively, we have

$$\mathbb{P}_n \times \mathbb{Q}_n(f \times g) = \mathbb{P}_n f \cdot \mathbb{Q}_n g$$

We consider the test statistic

$$K_n = \sqrt{n}||\mathbb{H}_n - \mathbb{P}_n \times \mathbb{Q}_n||_{\mathscr{F}\times\mathscr{G}}$$

and the *independence empirical process* $\{\mathbb{Z}_n h : h \in \mathscr{F} \times \mathscr{G}\}$ given by

$$\mathbb{Z}_n := \sqrt{n}[(\mathbb{H}_n - \mathbb{P}_n \times \mathbb{Q}_n)(f \times g) - (H - P \times Q)(f \times g)]$$

Under H_0,

$$K_n = ||\mathbb{Z}_n||_{\mathscr{F}\times\mathscr{G}}.$$

Under H_1,

$$K_n = ||\mathbb{Z}_n + \sqrt{n}(H - P \times Q)||_{\mathscr{F}\times\mathscr{G}} \geq \sqrt{n}||H - P \times Q||_{\mathscr{F}\times\mathscr{G}} - ||\mathbb{Z}_n||_{\mathscr{F}\times\mathscr{G}},$$

where the first term $\to \infty$.

Consistent testing based on K_n is possible if $||\mathbb{Z}_n||_{\mathscr{F}\times\mathscr{G}} = O_H(1)$ under the null- *and* alternative hypothesis.

- Under the null, $K_n = O_H(1)$.
- Under H_1, K_n is unbounded in probability.

An (H free) critical value can be obtained through bootstrapping, similarly to what we saw in Amine's presentation last week.

Consistent testing based on K_n is possible if $||\mathbb{Z}_n||_{\mathscr{F}\times\mathscr{G}} = O_H(1)$ under the null- *and* alternative hypothesis. In this case:

- under the null, $K_n = O_H(1)$,
- under H_1, K_n is unbounded in probability.

An (H free) critical value can be obtained through bootstrapping, similarly to what we saw in Amine's presentation last week.

Theorem (Theorem 3.9.1)

Let $\mathscr{F}$ and $\mathscr{G}$ be classes of measurable functions on measurable spaces $(\mathscr{X}, \mathscr{A})$ and $(\mathscr{Y}, \mathscr{B})$, respectively. If $\mathscr{F} \times \mathscr{G}$, $\mathscr{F}$ and $\mathscr{G}$ are H-Donsker and $||P||_{\mathscr{F}} < \infty$ and $||Q||_{\mathscr{G}} < \infty$, then $\mathbb{Z}_n$ converges in distribution in $\ell^{\infty}(\mathscr{F} \times \mathscr{G})$ to a Gaussian process $\mathbb{Z}_H$.

Proof: We can write

$$\begin{aligned}\mathbb{Z}_n(f,g) &:= \sqrt{n}[(\mathbb{H}_n - \mathbb{P}_n \times \mathbb{Q}_n)(f \times g) - (H - P \times Q)(f \times g)] \\ &= \sqrt{n}[(\mathbb{H}_n - H)(f \times g) - (\mathbb{P}_n - P)f - Pf(\mathbb{Q}_n - Q)g] \\ &= \sqrt{n}(\mathbb{H}_n - H)((f - Pf) \times (g - Qg)) - \sqrt{n}(\mathbb{P}_n - P)f(\mathbb{Q}_n - Q)g.\end{aligned}$$

- First equality: definition.
- Second equality: add-subtract $\mathbb{Q}_n g P_f$.
- Third equality: note that $H(f \times Qg) = PfQg$.

We end up with

$$\mathbb{Z}_n(f,g) = \sqrt{n}(\mathbb{H}_n - H)((f - Pf) \times (g - Qg)) - \sqrt{n}(\mathbb{P}_n - P)f(\mathbb{Q}_n - Q)g.$$

Since $\mathscr{F}$ and $\mathscr{G}$ are H-Donsker, the last term is $o_H(1)$ by Slutsky's Lemma.

Since $\mathscr{F} \times \mathscr{G}$ is H-Donsker, the first term converges weakly to a Gaussian process $\mathbb{Z}_H$, which finishes the proof.

Furthermore, we see that the covariance function of $\mathbb{Z}_H$ under the null hypothesis is

$$\operatorname{cov}(\mathbb{Z}_H(f_1 \times g_1), \mathbb{Z}_H(f_2 \times g_2)) = (Pf_1f_2 - Pf_1Pf_2)(Qg_1g_2 - Qg_1Qg_2).$$

Theorem 3.9.1 gives sufficiency conditions for the test statistic K_n to behave in way that we can use to reliably test.

We do not know the distribution of K_n under the null generally, so we are going to bootstrap to obtain a critical point that 'approximates' to asymptotic critical value

$$c := \inf\{t > 0 : P \times Q(||\mathbb{Z}_{P\times Q}||_{\mathscr{F}\times\mathscr{G}} > t) \leq \alpha\}.$$

- Chapter 3.9 provides a proof of a regular bootstrap procedure that succesfully approximates c.
- It is stated in the end of the chapter that permutation bootstrap should also work, but there is no proof as of yet for this.

Note that since we do not know whether we live under the null- or alternative hypothesis we need a bootstrap procedure such that our bootstrap test statistic $||\mathbb{Z}_n||_{\mathscr{F}\times\mathscr{G}}$ behaves the same under the null- *and* the alternative hypothesis.

In order to assure this, we sample $(\hat{X}_1,\hat{Y}_1),(\hat{X}_2,\hat{Y}_2),\ldots,(\hat{X}_n,\hat{Y}_n)$ from $\mathbb{P}_n\times\mathbb{Q}_n$ and consider

$$\hat{\mathbb{H}}_n := \frac{1}{n}\sum_{i=1}^{n}\delta_{\hat{X}_i}\times\delta_{\hat{Y}_i},$$

and

$$\hat{\mathbb{P}}_n\times\hat{\mathbb{Q}}_n = \left(\frac{1}{n}\sum_{i=1}^{n}\delta_{\hat{X}_i}\right)\times\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{\hat{Y}_i}\right).$$

We obtain the *bootstrap independence process*

$$\hat{\mathbb{Z}}_n := \sqrt{n}(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}\times\hat{\mathbb{Q}}_n).$$

We will define the critical region for $\hat{K}_n := ||\hat{\mathbb{Z}}_n||_{\mathscr{F}\times\mathscr{G}}$ as

$$\hat{c}_n = \inf\{t > 0 : P_{\hat{X}\times\hat{Y}}(\hat{K}_n > t) \leq \alpha\}.$$

Assuming some things, we shall establish that both under the null- and alternative hypothesis,

$$\hat{c}_n \to c \text{ a.s.}$$

so we can define an asymptotically consistent test of level α through

$$\begin{cases} \text{REJECT} & \text{if } K_n > \hat{c}_n \\ \text{FAIL TO REJECT} & \text{if } K_n \leq \hat{c}_n. \end{cases}$$

Theorem (Theorem 3.9.3)

Let $\mathscr{F}$ and $\mathscr{G}$ be separable classes of measurable functions on measurable spaces $(\mathscr{X},\mathscr{A})$ and $(\mathscr{Y},\mathscr{B})$, respectively, such that $\mathscr{F}\times\mathscr{G}$ satisfies the uniform entropy condition for envelope functions F, G and $F\times G$ that are H-square integrable.

Then,

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n \times \mathbb{Q}_n) \rightsquigarrow \mathbb{G}_{P\times Q} \text{ in } \ell^\infty(\mathscr{F}\times\mathscr{G})$$

and

$$\hat{\mathbb{Z}}_n := \sqrt{n}(\hat{\mathbb{H}}_n - \hat{\mathbb{P}}_n \times \hat{\mathbb{Q}}_n) \rightsquigarrow \mathbb{Z}_{P\times Q} \text{ in } \ell^\infty(\mathscr{F}\times\mathscr{G})$$

given H^∞ almost every sequence $(X_1,Y_1),(X_2,Y_2),\ldots$. Here, $\mathbb{G}_{P\times Q}$ is a Brownian bridge.

The proof of the theorem comes down to an application of Theorem 2.8.9 to show

$$\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n \times \mathbb{Q}_n) \rightsquigarrow \mathbb{G}_{P\times Q} \text{ in } \ell^\infty(\mathscr{F} \times \mathscr{G}).$$

By rewriting $\hat{\mathbb{Z}}_n$ as

$$\begin{aligned}\sqrt{n}(\hat{\mathbb{H}}_n - \mathbb{P}_n \times \mathbb{Q}_n)(f \times g) - \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n) f \mathbb{Q}_n g \\ -\mathbb{P}_n f \sqrt{n}(\hat{\mathbb{Q}}_n - \mathbb{Q}_n) - \sqrt{n}(\hat{\mathbb{P}}_n - \mathbb{P}_n) f (\hat{\mathbb{Q}}_n - \mathbb{Q}_n) g\end{aligned}$$

and Slutsky's lemma, we see that the above converges weakly to

$$\mathbb{G}_{P\times Q}(f \times g) - \mathbb{G}_{P\times Q}(f \times 1) Q g - P f \mathbb{G}_{P\times Q}(1 \times g).$$

The covariance function of the above Gaussian process is

$$(f_1 \times g_1), (f_2 \times g_2) \mapsto (P f_1 f_2 - P f_1 P f_2)(Q g_1 g_2 - Q g_1 Q g_2),$$

which is that same as that of $\mathbb{Z}_H$.

Mini recall of Theorem 2.8.9: it concerned weak convergence of *empirical processes of triangular arrays*

$$P_n = \frac{1}{n}\sum_{i=1}^{n}\delta_{X_{ni}}$$

and some underlying true measure P_n which converges to some P_0 in the sense that

$$\sup_{h_1,h_2\in\mathscr{F}} |\rho_{P_n}(f,g) - \rho_{P_0}(f,g)| \to 0$$

where ρ is the variance metric.

Theorem

Assume appropriate measurability of functions in class $\mathscr{F}$, that is has an envelope F, satisfies the uniform entropy condition

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

that $P_nF^2 = O(1)$ and that P_n converges to P_0 in the above sense, as well as

$$\limsup_{n\to\infty} P_nF^2\{F \geq \varepsilon\sqrt{n}\} = 0 \text{ for every } \varepsilon.$$

Then, $\mathbb{G}_{P_n} \rightsquigarrow \mathbb{G}_{P_0}$.

As a consequence of Theorem 3.9.3, we have that $\hat{\mathbb{Z}}_n$ converges weakly to a centered Gaussian random element in $\ell^\infty(\mathscr{F} \times \mathscr{G})$ for H^∞ almost surely every sequence $(X_1, Y_1), (X_2, Y_2), \ldots$.

What's more: this weak limit has the same covariance function as $\mathbb{Z}_n$'s Gaussian weak limit; so actually their weak limits are the same.

By the continuous mapping theorem, we have that almost surely conditionally

$$\hat{K}_n := ||\hat{\mathbb{Z}}_n||_{\mathscr{F} \times \mathscr{G}} \rightsquigarrow ||\mathbb{Z}_{P \times Q}||_{\mathscr{F} \times \mathscr{G}}.$$

Consequently,

$$\hat{c}_n := \inf\{t > 0 : P_{\hat{X}\times\hat{Y}}(\hat{K}_n > t) \leq \alpha\}$$

converges almost surely to

$$\inf\{t > 0 : P \times Q(||\mathbb{Z}_{P\times Q}||_{\mathscr{F}\times\mathscr{G}} > t) \leq \alpha\}$$

so we can indeed define an asymptotically consistent test of level α through

$$\begin{cases} \text{REJECT} & \text{if } K_n > \hat{c}_n \\ \text{FAIL TO REJECT} & \text{if } K_n \leq \hat{c}_n. \end{cases}$$

Thank you for listening!