Empirical Processes with Applications in Statistics

4. Tools to work with empirical processes

Johan Segers

UCLouvain (Belgium)

Conférence Universitaire de Suisse Occidentale

Programme Doctoral en Statistique et Probabilités Appliquées

Les Diablerets, February 4–5, 2020

So far so good, but…

So far: general framework for proving weak convergence

$$\mathbb{G}_n = \sqrt{n}\,(\mathbb{P}_n - P) \rightsquigarrow \mathbb{G}, \qquad n \to \infty$$

in $\ell^\infty(\mathcal{F})$ for suitable families of functions $\mathcal{F} \subset L_2(P)$

How to find the asymptotic distribution of estimators of "parameters" arising as statistical functionals ϕ?

$$\hat{\vartheta}_n = \phi(\mathbb{P}_n),$$
$$\vartheta = \phi(P)$$

$\implies$ *Functional delta method*

Often, the limit distribution is unknown and complicated: in practice?
$\implies$ *Bootstrap*

Main sources for this lecture: van der Vaart and Wellner (1996), Kosorok (2008)

Tools to work with empirical processes

Functional delta method

Bootstrap

Ordinary delta method

Random vectors X_n in $\mathbb{R}^d$. Suppose $\theta \in \mathbb{R}^d$ and $0 < r_n \to \infty$ and

$$Z_n =: r_n(X_n - \theta) \xrightarrow{d} Z, \qquad n \to \infty$$

Delta method. Let $\phi : \mathbb{R}^d \to \mathbb{R}^p$ be differentiable at θ. Then

$$r_n \{\phi(X_n) - \phi(\theta)\} \xrightarrow{d} \phi'(\theta) Z$$

where $\phi'(\theta) \in \mathbb{R}^{p \times d}$ is the Jacobian of ϕ at θ.

Proof: Write

$$r_n \{\phi(X_n) - \phi(\theta)\} = r_n \left\{\phi(\theta + r_n^{-1} Z_n) - \phi(\theta)\right\} = g_n(Z_n)$$

By differentiability of ϕ, for all sequences $z_n \to z \in \mathbb{R}^d$,

$$\lim_{n\to\infty} g_n(z_n) = g(z) = \phi'(\theta) z$$

Apply the extended continuous mapping theorem. □

Delta method in normed spaces?

Empirical process $\mathbb{G}_n f = \sqrt{n}(\mathbb{P}_n f - Pf)$ indexed by P-Donsker $\mathcal{F} \in L_2(P)$:

$$\mathbb{G}_n \rightsquigarrow \mathbb{G}, \qquad \text{in } \ell^\infty(\mathcal{F})$$

"Smooth" transformation

$$\phi : \ell^\infty(\mathcal{F}) \to \mathbb{E}$$

where $\mathbb{E}$ is some other normed real vector space
Often $\mathbb{E}$ is $\mathbb{R}^p$ or $\ell^\infty(T)$ for some set T

Functional delta method? Asymptotic distribution of *statistical functionals*

$$\sqrt{n}\{\phi(\mathbb{P}_n) - \phi(P)\} \rightsquigarrow \phi'_P(\mathbb{G}), \qquad \text{in } \mathbb{E}$$

with "derivative" $\phi'_P : \ell^\infty(\mathcal{F}) \to \mathbb{E}$, linear and bounded

Differentiable mappings between normed spaces

Map $\phi : \mathbb{D} \to \mathbb{E}$ between normed real vector spaces

Definition: Hadamard differentiability.
$\phi : \mathbb{D} \to \mathbb{E}$ is *Hadamard differentiable* at θ if there exists bounded linear $\phi'_\theta : \mathbb{D} \to \mathbb{E}$ such that, in $\mathbb{E}$,

$$\lim_{t \downarrow 0} \frac{1}{t} \{\phi(\theta + th_t) - \phi(\theta)\} = \phi'_\theta(h)$$

whenever $\lim_{t \downarrow 0} h_t = h$ in $\mathbb{D}$.

- Stronger than *Gateaux differentiability* (fixed $h_t \equiv h$)
- Weaker than *Fréchet differentiability* (bounded h_t)
- In $\mathbb{R}^d$, all three definitions are the same as the usual one
- Similar definition if ϕ has domain $\mathbb{D}_\phi \subset \mathbb{D}$

Slight generalization in view of applications

Definition: H-dif tangentially to a set.
$\phi : \mathbb{D}_\phi \to \mathbb{E}$ is *Hadamard differentiable at* $\theta \in \mathbb{D}_\phi$ *tangentially to* $\mathbb{D}_0 \subset \mathbb{D}$ if there exists bounded linear $\phi'_\theta : \mathbb{D}_0 \to \mathbb{E}$ such that, whenever $\{h_t\}_{t>0} \subset \mathbb{D}$ satisfies

- $\theta + th_t \in \mathbb{D}_\phi$ for all sufficiently small $t > 0$
- $\lim_{t\downarrow 0} h_t = h \in \mathbb{D}_0$

then

$$\lim_{t\downarrow 0} \frac{1}{t} \{\phi(\theta + th_t) - \phi(\theta)\} = \phi'_\theta(h)$$

Chain rule for differentiation applies as usual: facilitates proving H-dif

Functional delta method

- Normed linear spaces $\mathbb{D}$ and $\mathbb{E}$
- $\mathbb{D}_\phi \subset \mathbb{D}$
- $\phi : \mathbb{D}_\phi \to \mathbb{E}$
- $T_n : \Omega \to \mathbb{D}_\phi$
- $0 < r_n \to \infty$ and $\theta \in \mathbb{D}_\phi$

Theorem: Functional delta method. If

1. If $r_n(T_n - \theta) \rightsquigarrow Z$ in $\mathbb{D}$ with Z taking values only in $\mathbb{D}_0$
2. ϕ is Hadamard differentiable tangentially to $\mathbb{D}_0$

then, in $\mathbb{E}$,

$$r_n \{\phi(T_n) - \phi(\theta)\} \rightsquigarrow \phi'_\theta(Z), \qquad n \to \infty$$

Asymptotic normality of statistical functionals

Suppose $\mathcal{F} \subset L_2(P)$ is P-Donsker.

Setting $\mathbb{D} = \ell^\infty(\mathcal{F})$ and $T_n = \mathbb{P}_n$ and $\theta = P$, we find, in $\mathbb{E}$,

$$\sqrt{n}\{\phi(\mathbb{P}_n) - \phi(P)\} \rightsquigarrow \phi'_P(\mathbb{G}), \qquad n \to \infty$$

provided

- $\mathbb{P}_n$ takes values in $\mathbb{D}_\phi \subset \ell^\infty(\mathcal{F})$
- $\mathbb{G}$ takes values in $\mathbb{D}_0 \subset \ell^\infty(\mathcal{F})$
- $\phi : \mathbb{D}_\phi \to \mathbb{E}$ is Hadamard-differentiable at P tangentially to $\mathbb{D}_0$

The $\mathbb{E}$-valued weak limit $\phi'_P(\mathbb{G})$ is Gaussian as well

For every continuous linear functional $\rho : \mathbb{E} \to \mathbb{R}$, the random variable $\rho \circ \phi'_P(\mathbb{G})$ is Gaussian

Example: sample quantiles

Iid random variables $X_1, X_2, \ldots \sim F$ on $\mathbb{R}$, empirical distribution function $\mathbb{F}_n$.

Population and sample *quantiles*: for $0 < p < 1$,

$$Q(p) = \inf\{x \in \mathbb{R} \mid F(x) \geqslant p\} = \phi(F)$$
$$Q_n(p) = \inf\{x \in \mathbb{R} \mid \mathbb{F}_n(x) \geqslant p\} = \phi(\mathbb{F}_n)$$

Asymptotic normality of

$$\sqrt{n}\{Q_n(p) - Q(p)\}$$

by functional delta method:

- $\mathbb{D} = \ell^\infty(\mathbb{R})$ and $\mathbb{E} = \mathbb{R}$
- $\mathbb{D}_\phi = \{$all distribution functions on $\mathbb{R}\}$
- $T_n = \mathbb{F}_n$ and $\theta = F$ and $r_n = \sqrt{n}$
- Limit process $\mathbb{G}(x) = B(F(x))$, $x \in \mathbb{R}$, with B a Brownian bridge

“Vervaat’s lemma”

If F has derivative f at $Q(p)$, then ϕ is H-dif at F tangentially to

$$\mathbb{D}_0 = \{h \in \ell^\infty(\mathbb{R}) \mid h \text{ is continuous at } Q(p)\}$$

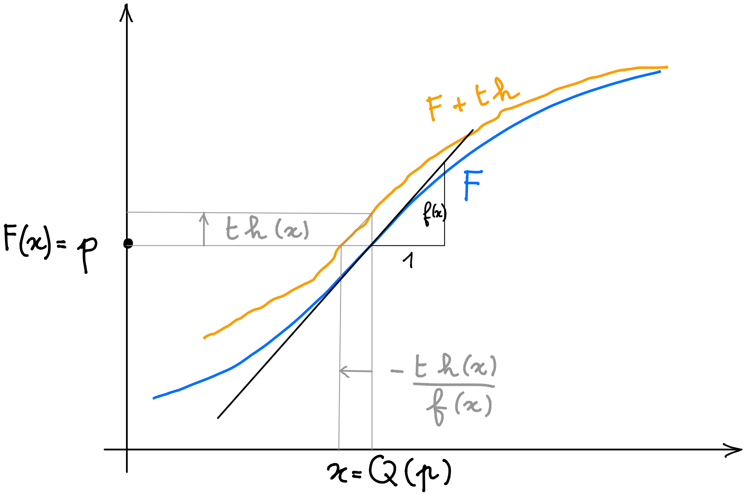
with derivative map

$$\phi'_F(h) = -\frac{h(Q(p))}{f(Q(p))}, \qquad h \in \mathbb{D}_0$$

By the functional delta method,

$$\sqrt{n}\{Q_n(p) - Q(p)\} \xrightarrow{d} -\frac{B(p)}{f(Q(p))} \sim \mathcal{N}\left(0, \frac{p(1-p)}{(f(Q(p)))^2}\right)$$

Extension: *empirical quantile process* $\{Q_n(p)\}_{p \in [a,b]}$


F + t h
F
f(x)
1
t h (x)
F(x) = p
- t h(x) / f(x)
x = Q(p)

Example: copulas

The *copula* C of a d-variate random vector $X = (X_1, \ldots, X_d)$ with joint cdf F and continuous marginals $F_1, \ldots, F_d$ is the joint cdf of

$$(U_1, \ldots, U_d) = (F_1(X_1), \ldots, F_d(X_d))$$

A copula is a cdf on $[0,1]^d$ with uniform $[0,1]$ marginals. Explicit formula:

$$C(u_1, \ldots, u_d) = F\Big(F_1^{-}(u_1), \ldots, F_d^{-}(u_d)\Big)$$

with F_j^{-} the quantile function, defined carefully as

$$F_j^{-}(u_j) = \begin{cases} \inf\left\{x \in \mathbb{R} \mid F_j(x) \geqslant u_j\right\}, & 0 < u_j \leqslant 1 \\ \sup\left\{x \in \mathbb{R} \mid F_j(x) = 0\right\}, & u_j = 0 \end{cases}$$

Context: Modelling dependence in a margin-free way

Empirical copula process

iid sample $X_i = (X_{i1}, \ldots, X_{id})$ from F. *Empirical copula*:

$$\mathbb{C}_n(u_1, \ldots, u_d) = \mathbb{F}_n\left(\mathbb{F}_{n,1}^{-}(u_1), \ldots, \mathbb{F}_{n,d}^{-}(u_d)\right)$$

with empirical cdfs

$$\mathbb{F}_n(x) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left\{X_{i1} \leqslant x_1, \ldots, X_{id} \leqslant x_d\right\}$$

$$\mathbb{F}_{n,j}(x_j) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}\left\{X_{ij} \leqslant x_j\right\}$$

and generalized inverses as on previous slide. *Rank*-based estimator

Weak convergence in $\ell^\infty([0,1]^d)$ of the *empirical copula process*?

$$\sqrt{n}\,(\mathbb{C}_n - C) = \sqrt{n}\,\{\phi(\mathbb{F}_n) - \phi(F)\}$$

Differentiability of the copula mapping

- $\mathbb{D} = \mathbb{E} = \ell([0,1]^d)$
- $\mathbb{D}_\phi = \Big\{\text{cdf } F \text{ on } [0,1]^d \text{ with margins } F_j(0) = 0\Big\}$
- $\phi : \mathbb{D}_\phi \to \mathbb{E} : F \mapsto F(F_1^-, \ldots, F_d^-)$

Bücher and Volgushev (2013). Assume the d partial derivatives $\partial_j C$ exist and are continuous on $\{u \in [0,1]^d \mid 0 < u_j < 1\}$. Then ϕ is Hadamard differentiable at C tangentially to

$$\mathbb{D}_0 = \Big\{\alpha \in C([0,1]^d) \,\Big|\, \alpha(1,\ldots,1) = 0, \\ \alpha(x_1,\ldots,x_d) = 0 \text{ if } \min_j x_j = 0\Big\}$$

with derivative $\phi'_C : \mathbb{D}_0 \to \ell^\infty([0,1]^d)$ given by

$$\big(\phi'_C(\alpha)\big)(u) = \alpha(u) - \sum_{j=1}^{p} \partial_j C(u)\, \alpha(1,\ldots,1,u_j,1,\ldots,1)$$

Weak convergence of the empirical copula process

For $C \in \mathbb{D}_\phi$ satisfying the assumption, we obtain weak convergence

$$\sqrt{n}\,(\mathbb{C}_n - C) \rightsquigarrow \beta, \qquad n \to \infty$$

in $\ell^\infty([0,1]^d)$ with weak limit

$$\beta(u) = \mathbb{G}_C(u) - \sum_{j=1}^{d} \partial C_j(u)\, \mathbb{G}_C(1, \ldots, 1, u_j, 1, \ldots, 1), \qquad u \in [0,1]^d$$

with $\mathbb{G}_C$ a C-Brownian bridge indexed by functions $\mathbb{1}_{[0,u]}$

Weak convergence of empirical copula process under the same or similar conditions already before obtained by other authors

Tools to work with empirical processes

Functional delta method

Bootstrap

Using asymptotic distributions

Estimator $\hat{\vartheta}_n = \phi(\mathbb{P}_n)$ of $\vartheta = \phi(P)$. Asymptotic distribution

$$\sqrt{n}\left(\hat{\vartheta}_n - \vartheta\right) \rightsquigarrow \phi'_P(\mathbb{G}), \qquad n \to \infty$$

Usage:

- asymptotic confidence intervals
- critical values for hypothesis tests

Problem: limit distribution depends on unknown P via ϕ'_P and $\mathbb{G}$

Possible solutions:

- Replace P in limit distribution by $\mathbb{P}_n$
  - Consistent procedure under fairly general conditions
  - Estimated limit distribution may be difficult to handle
    Imagine a Kolmogorov–Smirnov test for the empirical copula process
- *Bootstrap*

Resampling with replacement

Empirical measure when resampling from the data with replacement:

$$\hat{\mathbb{P}}_n = \frac{1}{n}\sum_{i=1}^{n} M_{ni}\,\delta_{X_i}$$

- M_{ni}: number of times X_i was selected out of n trials
- $(M_{n1}, \ldots, M_{nn})$ multinomial $(n; 1/n, \ldots, 1/n)$ independent of $X_1, \ldots, X_n$
- δ_x is degenerate probability measure at x

Bootstrap: estimate the distribution of $\sqrt{n}\,\{\phi(\mathbb{P}_n) - \phi(P)\}$ by the one of

$$\sqrt{n}\left\{\phi(\hat{\mathbb{P}}_n) - \phi(\mathbb{P}_n)\right\}$$

To be shown: with large probability, the conditional distribution of the above quantity given the data $X_1, \ldots, X_n$ is close to the limit distribution on slide 18

Showing consistency of the bootstrap

Two steps:

1. Show consistency of the *bootstrapped empirical process*: with large probability, the conditional distribution of

$$\hat{\mathbb{G}}_n = \sqrt{n}\left(\hat{\mathbb{P}}_n - \mathbb{P}_n\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (M_{ni} - 1)(\delta_{X_i} - \mathbb{P}_n)$$

   given the data $X_1, \ldots, X_n$ is close to the one of $\mathbb{G}$
2. Apply (a suitable extension of) the delta method

Method extends to more general *weights* W_{ni} than $M_{ni} - 1$

- multiplier bootstrap
- wild bootstrap
- ...

Multiplier central limit theorem

Consistency in probability of the bootstrap

If $\mathcal{F} \subset L_2(P)$ has finite envelope, then $\mathcal{F}$ is P-Donsker if and only if $\hat{\mathbb{G}}_n$ is asymptotically measurable and

$$\sup_{h \in \mathrm{BL}_1} \left| \mathrm{E}_M[h(\hat{\mathbb{G}}_n)] - \mathrm{E}[h(\mathbb{G})] \right| \xrightarrow{\mathrm{P}} 0, \qquad n \to \infty$$

- E_M means conditional expectation over $(M_{n1}, \ldots, M_{nn})$ given $X_1, \ldots, X_n$
- BL_1 is the set of $h : \ell^\infty(\mathcal{F}) \to [-1, 1]$ such that

$$\left| h(z) - h(y) \right| \leqslant \|z - y\|_\infty, \qquad y, z \in \ell^\infty(\mathcal{F})$$

  "Bounded Lipschitz distance" $\sup_{h \in \mathrm{BL}_1} |\ldots|$ metrizes weak convergence

- Proof based on
  - Poissonization of $(M_{n1}, \ldots, M_{nn})$
  - (conditional) *multiplier central limit theorem*

Summary

This course:

- Weak convergence of empirical processes indexed by function classes
- Tools to leverage the power of the asymptotic theory
- Examples

What's next:

- Convergence rates different from $\sqrt{n}$ (van de Geer, 2000)
- Concentration inequalities (Giné and Guillou, 2001)
- Empirical processes indexed by estimated functions (van der Vaart and Wellner, 2007)
- Application to M and Z-estimators
- Almost sure versions, strong approximations
- Time series data
- ...

Thank you!

References

Bücher, A. and S. Volgushev (2013). Empirical and sequential empirical copula processes under serial dependence. *Journal of Multivariate Analysis 119*, 61–70.

Giné, E. and A. Guillou (2001). On consistency of kernel density estimators for randomly censored data: rates holding uniformly over adaptive intervals. *Ann. Inst. Henri Poincaré 37*, 503–522.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference*. New York: Springer Sciences+Business Media.

van de Geer, S. A. (2000). *Empirical Processes in M-Estimation*. Cambridge: Cambridge University Press.

van der Vaart, A. W. and J. A. Wellner (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics*. New York: Springer Sciences+Business Media.

van der Vaart, A. W. and J. A. Wellner (2007). Empirical processes indexed by estimated functions. In *Asymptotics: particles, processes and inverse problems*, Volume 55 of *IMS Lecture Notes Monogr. Ser.*, pp. 234–252. Beachwood, OH: Inst. Math. Statist.