Uniformity and the delta-method

Maximilian Kasy José L. Montiel Olea

October 27, 2014

Introduction

- Many procedures for estimation and inference:
  - motivated by asymptotic behavior
  - for fixed parameter values.
- Often, such procedures behave poorly
  - in finite samples
  - for some parameter regions.
- Such problems can arise,
  if approximations are not uniformly valid.
- Can lead to
  1. large mean squared error for estimators,
  2. undercoverage for confidence sets,
  3. distorted rejection rates for tests,
  4. ...

Examples in econometrics

1. Instrumental variables:
   poor behavior for weak instruments
2. Inference under partial identification:
   poor behavior near point-identification
3. Estimation after model selection:
   poor behavior around the critical values for model selection
4. Time series:
   poor behavior near unit roots

- Unifying theme?
- One important tool in asymptotics:
  Delta-method
- Taylor expansions to approximate functions of random variables
- Problems $\Leftrightarrow$
  Large remainder for some parameter values
- This paper:
  - A sufficient and necessary condition
  - for uniform negligibility
  - of the remainder.

Roadmap

- Literature
- Preliminaries:
  - Definitions
  - Uniformity and inference
  - Uniform continuous mapping theorem
- Uniform delta method:
  - Necessary and sufficient condition
  - Simpler sufficient conditions
- Applications:
  - Stylized examples: $|t|$, $1/t$, $\sqrt{t}$, $\cos(t^2)$
  - Weak instruments, moment inequalities

Preliminaries

Notation:

- $\theta \in \Theta$ indexes the distribution of the observed data
- $\mu = \mu(\theta)$ is some finite dimensional function of θ
- asymptotics wrt n
- F: cumulative distribution functions
- S, T, X, Y and Z: random variables / vectors

Definition (bounded Lipschitz metric)

- $\mathbf{BL_1}$: *set of all functions h on $\mathbb{R}^k$ such that*
  1. $|h(x)| \leq 1$ *and*
  2. $|h(x) - h(x')| \leq \|x - x'\|$ *for all* x, x'
- *bounded Lipschitz metric on the set of random variables:*

$$d_{BL}^{\theta}(X_1, X_2) := \sup_{h \in \mathbf{BL_1}} \left| E^{\theta}[h(X_1)] - E^{\theta}[h(X_2)] \right| .$$

- van der Vaart and Wellner (1996, section 1.12): convergence in distribution of X_n to X $\Leftrightarrow$ convergence of $d_{BL}^{\theta}(X_n, X)$ to 0.

Definition (Uniform convergence)

1. *X_n converges uniformly in distribution to Y_n if*

$$d_{BL}^{\theta_n}(X_n, Y_n) \to 0$$

   for all sequences $\{\theta_n \in \Theta\}$.

2. *X_n converges uniformly in probability to Y_n if*

$$P^{\theta_n}(\|X_n - Y_n\| > \varepsilon) \to 0$$

   for all $\varepsilon > 0$ and all sequences $\{\theta_n \in \Theta\}$.

Lemma (Characterization of uniform convergence)

1. *X_n converges uniformly in distribution to Y_n iff*

$$\sup_\theta d^\theta_{BL}(X_n, Y_n) \to 0$$

2. *X_n converges uniformly in probability to Y_n iff*

$$\sup_\theta P^\theta(\|X_n - Y_n\| > \varepsilon) \to 0$$

for all $\varepsilon > 0$.

Remarks

- Definition of convergence:
  sequence X_n toward another sequence Y_n
- Special case $Y_n = X$
- Uniform convergence in distribution safeguards
  - for large n
  - against poor approximation
  - for some θ.
- Next slide:
  uniform convergence in distribution $\Rightarrow$ uniform validity of inference procedures

Lemma (Uniform confidence sets)

- *Suppose* $Z_n = Z_n(\mu) \to^d Z$ *uniformly, where*
- Z *is continuously distributed and*
- *the distribution of* Z *does not depend on* θ.
- *Let* z *be the* $1-\alpha$ *quantile of the distribution of* Z.

Then

$$C_n := \{m : Z_n(m) \leq z\}$$

is such that

$$P^{\theta_n}(\mu(\theta_n) \in C_n) \to 1-\alpha$$

for any sequence θ_n.

Theorem (Uniform continuous mapping theorem)

Let $\psi(x)$ be a Lipschitz-continuous function of x.

1. *Suppose X_n converges uniformly in distribution to Y_n. Then $\psi(X_n)$ converges uniformly in distribution to $\psi(Y_n)$.*
2. *Suppose X_n converges uniformly in probability to Y_n. Then $\psi(X_n)$ converges uniformly in probability to $\psi(Y_n)$.*

The uniform delta-method

Setting

- sequence of numbers r_n (eg. $r_n = \sqrt{n}$)
- sequence of random variables T_n
- such that

$$S_n := r_n(T_n - \mu) \to^d S$$

  uniformly
- all distributions and μ indexed by θ
- corresponding sequence

$$X_n := r_n(\phi(T_n) - \phi(\mu))$$

- goal: approximate the distribution of X_n by the distribution of

$$X := \frac{\partial \phi}{\partial x}(\mu) \cdot S.$$

- first order Taylor expansion of ϕ:

$$\phi(t) = \phi(m) + \frac{\partial \phi}{\partial m}(m)(t-m) + o(t-m)$$

- normalized remainder

$$\Delta(t,m) := \frac{1}{\|t-m\|} \left\| \phi(t) - \phi(m) - \frac{\partial \phi}{\partial m}(m) \cdot (t-m) \right\|.$$

- $$p(\varepsilon, \varepsilon', m) := \int_{\|s\| \leq 1} \mathbf{1}\left(\Delta\left(m + \varepsilon \cdot s, m\right) > \varepsilon'\right) ds.$$

- necessary and sufficient condition for uniform delta-method: bound on $p(\varepsilon, \varepsilon', m)$
- sufficient condition:
  bound on Δ

Assumption (Uniform convergence of S_n)

Let $S_n := r_n(T_n - \mu)$.

1. *$S_n \to^d S$ uniformly.*
2. *S is continuously distributed for all θ.*
3. *The collection $\{S(\theta)\}_{\theta \in \Theta}$ is tight.*
4. *The density of S satisfies*

$$\begin{aligned} \underline{f} \leq f_\theta(s) & \qquad \forall\, s:\ \|s\| < \overline{s},\ \forall\, \theta \\ f_\theta(s) \leq \overline{f} & \qquad \forall\, s,\ \forall\, \theta \end{aligned}$$

Leading example:

- $S \sim N(0, \Sigma(\theta))$, with
- uniform lower and upper bounds on the eigenvalues of $\Sigma(\theta)$.

- Define

$$
\begin{aligned}
X_n &= r_n(\phi(T_n) - \phi(\mu)), \\
\widetilde{T}_n &= \mu + \frac{1}{r_n} S \\
\widetilde{X}_n &= r_n(\phi(\widetilde{T}_n) - \phi(\mu)) \\
X &= \frac{\partial \phi}{\partial \mu}(\mu) \cdot S.
\end{aligned}
$$

- Approximate X_n by $\widetilde{X}_n$ (uniformly):
  straightforward under assumption on uniform convergence of S_n
- Approximate $\widetilde{X}_n$ by X (uniformly):
  requires uniform bound on remainder of Taylor approximation

Theorem (Uniform delta method – part 1)

Suppose

- *assumption on uniform convergence of S_n holds, and*
- *ϕ is continuously differentiable everywhere in $\mu(\Theta)$.*

Then:

1. $$X_n \to^d \tilde{X}_n$$
   uniformly if $\partial\phi/\partial\mu$ is bounded.
2. $$\tilde{X}_n \to^p X$$
   uniformly if and only if
   $$p(\varepsilon, \varepsilon', m) \leq \delta(\varepsilon, \varepsilon') \quad (1)$$
   for all $\varepsilon, \varepsilon'$, all $m \in \mu(\Theta)$ and some function δ where
   $$\lim_{\varepsilon \to 0} \delta(\varepsilon, \varepsilon') = 0.$$

Theorem (Uniform delta method – part 2)

3. *A sufficient condition for condition* (1)*:*

$$\Delta(t,m) \leq \widetilde{\delta}(\|t-m\|). \qquad (2)$$

for some function $\widetilde{\delta}$ *where* $\lim_{\varepsilon\to 0}\widetilde{\delta}(\varepsilon) = 0$.

4. *If*
   - *the domain of* ϕ *is compact and convex*
   - ϕ *is everywhere continuosly differentiable on its domain*

   then
   - $\partial\phi/\partial\mu$ *is bounded and*
   - *condition* (2) *holds.*

- compact and convex domain of continuously differentiable ϕ: sufficient for uniformity
- too restrictive for most applications
- but suggests where problems might occur:
  1. neighborhood of boundary points not included in the domain: near 0 for $|t|$, $1/t$, $\sqrt{t}$
  2. infinity: $\cos(t^2)$

- One of our assumptions: uniform convergence of S_n
- Special case: uniform CLT
- Follows from CLTs for triangular arrays, eg.

Lemma (Uniform central limit theorem)

- *Let Y_i be i.i.d.*
- *with mean $\mu(\theta)$ and variance $\Sigma(\theta)$.*
- *Assume that $E\left[\|Y_i^{2+\varepsilon}\|\right] < M$.*

Then

$$S_n := \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Y_i - \mu(\theta))$$

converges uniformly in distribution to the tight family $S \sim N(0, \Sigma(\theta))$.

Applications

- Our necessary condition is violated in several applications
- Stylized examples we will discuss next: $|t|$, $1/t$
- In the paper:
  - $\sqrt{t}$, $\cos(t^2)$
  - weak instruments
  - moment inequalities

Recall

- Sufficient condition:

$$\Delta(t,m) \leq \widetilde{\delta}(\|t-m\|),$$

  $\lim_{\varepsilon \to 0} \widetilde{\delta}(\varepsilon) = 0.$
- Sufficient and necessary condition:

$$p(\varepsilon,\varepsilon',m) \leq \delta(\varepsilon,\varepsilon'),$$

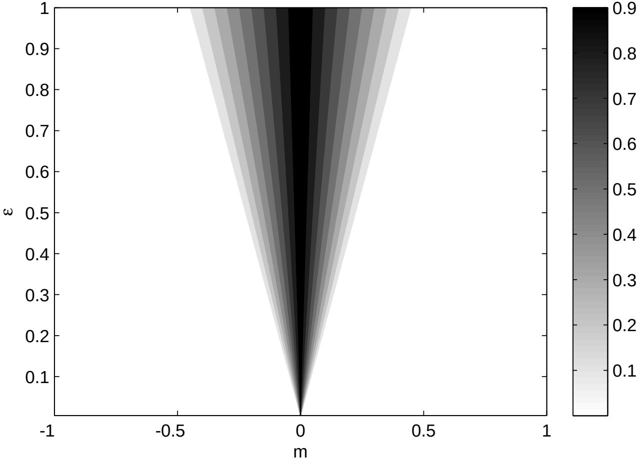
  $\lim_{\varepsilon \to 0} \delta(\varepsilon,\varepsilon') = 0.$
- Graphically:
  Level sets of $p(\varepsilon,\varepsilon',m)$, given ε',
  have to be bounded away from $\varepsilon = 0$ (m-axis).

Example 1: $\phi(t) = |t|$

- Stylized version of moment inequality-type problems.
- Domain: $\mathbb{R}\backslash\{0\}$
-
$$\phi(t) = |t|$$
$$\partial_m \phi(m) = \text{sign}(m)$$
-
$$\begin{aligned}\Delta(t,m) &= \frac{1}{|t-m|}\left|\phi(t) - \phi(m) - \partial_m \phi(m)\cdot(t-m)\right| \\ &= \frac{1}{|t-m|}\left||t| - |m| - \text{sign}(m)\cdot(t-m)\right| \\ &= \mathbf{1}(t\cdot m \leq 0)\frac{2\cdot|t|}{|t-m|}.\end{aligned}$$

- $$p(\varepsilon, \varepsilon', m) = \max\left(1 - \frac{2|m|/\varepsilon}{2-\varepsilon'}, 0\right).$$
- Consider
$$\varepsilon' = 1$$
$$\varepsilon_n = 1/n$$
$$m_n = 1/(4n)$$
- Then
$$p(1/n, 1, 1/(4n)) = 1/2 \not\to 0.$$
- Our necessary condition is violated.

Figure: The Integrated remainder $p(\varepsilon, 1, m)$ for $\phi(t) = |t|$.



Example 2: $\phi(t) = 1/t$

- Stylized version of weak instrument-type problems.
- Domain: $\mathbb{R}^{++}$
- $$\phi(t) = 1/t$$
  $$\partial_m \phi(m) = -1/m^2$$
- $$\begin{aligned}
\Delta(t,m) &= \frac{1}{|t-m|} \left|\phi(t) - \phi(m) - \partial_m\phi(m)\cdot(t-m)\right| \\
&= \frac{1}{|t-m|} \left|\frac{1}{t} - \frac{1}{m} + \frac{t-m}{m^2}\right| \\
&= \frac{1}{|t-m|} \left|\frac{m\cdot(m-t) + t\cdot(t-m)}{m^2\cdot t}\right| \\
&= \left|\frac{t-m}{m^2\cdot t}\right|.
\end{aligned}$$

- for $m \geq \varepsilon \geq \frac{m^2\varepsilon'}{1-m^2\varepsilon'}$, $m^2\varepsilon' < 1$

$$p(\varepsilon, \varepsilon', m) = 2\left(1 - \frac{m}{\varepsilon} \cdot \frac{1}{\frac{1}{m^2\varepsilon'} - m^2\varepsilon'}\right).$$

- Consider

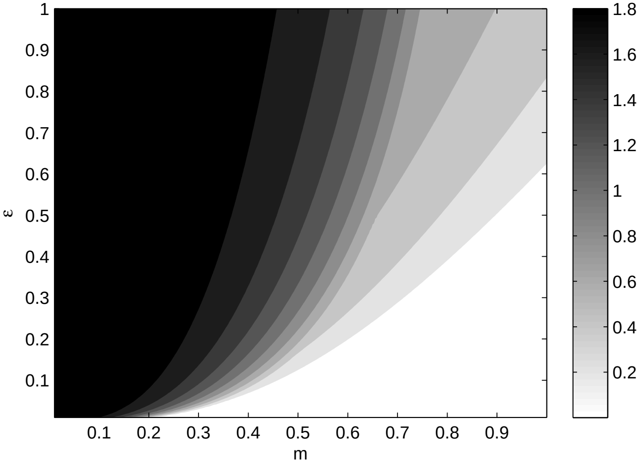
$$\varepsilon' = 1$$
$$\varepsilon_n = 1/n$$
$$m_n = 1/n$$

- Then

$$p(1/n, 1, 1/n) = -\frac{2}{n^2 - 1/n^2} \to 2.$$

- Our necessary condition is violated.

Figure: The Integrated remainder $p(\varepsilon, 1, m)$ for $\phi(t) = 1/t$.



Conclusion

- Problems with asymptotic approximations
- if approximations not uniformly valid.
- Can lead to
  - large mean squared error,
  - undercoverage,
  - distorted rejection rates.
- One important cause: large remainder of the delta-method
- We provide an easy-to-check condition which is necessary and sufficient for uniform negligibility of this remainder.

Thanks for your time!

Literature

very incomplete list:

- **Uniformity in statistics**:
  Rao (1963), Loh (1984), Hall et al. (1995), Bahadur and Savage (1956)
- **Weak instruments**:
  Staiger and Stock (1997), Moreira (2003), Andrews et al. (2006), Andrews and Mikusheva (2014)
- **Inference under partial identification, moment inequalities**:
  Imbens and Manski (2004)
- **Pretesting and model selection**:
  Leeb and Pötscher (2005), Guggenberger (2010a), Guggenberger (2010b)
- **Unit roots**:
  Stock and Watson (1996), Mikusheva (2007)
- **Sufficient condition** for uniform delta-method:
  Belloni et al. (2013)

Back