ECON 481-3
LECTURE 13: INFERENCE IN MOMENT INEQUALITY MODELS

Ivan A. Canay
Northwestern University

1. Canay, I.A. and A.M. Shaikh (2017): "Practical and Theoretical Advances for Inference in Partially Identified Models", In B. Honore, A. Pakes, M. Piazzesi, & L. Samuelson (Eds.), Advances in Economics and Econometrics: Volumen 2: Eleventh World Congress.
2. Ho, K. and A. M. Rosen (2017): "Partial Identification in Applied Research: Benefits and Challenges", In B. Honore, A. Pakes, M. Piazzesi, & L. Samuelson (Eds.), Advances in Economics and Econometrics: Volumen 2: Eleventh World Congress.

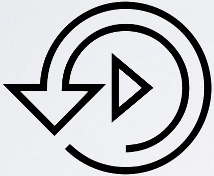
Past & Future

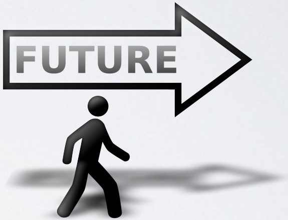
LAST CLASS

- Review of Subsampling
- Uniformity issues with Subsampling
- Parameter at the Boundary
- Asymptotic Size of Subsampling

TODAY

- Inference in MI Models
- Examples
- Confidence Regions
- LF and SS critical values





Motivation

Partially Identified Models:

- Param. of interest is not uniquely determined by distr. of obs. data.
- Instead, limited to a set as a function of distr. of obs. data.

  (i.e., the identified set)
- Due largely to pioneering work by C. Manski, now ubiquitous.

  (many applications!)

Inference in Partially Identified Models:

- Focused mainly on the construction of confidence regions.
- Most well-developed for moment inequalities.
- Important practical issues remain subject of current research.

Simplest Example: Missing Data

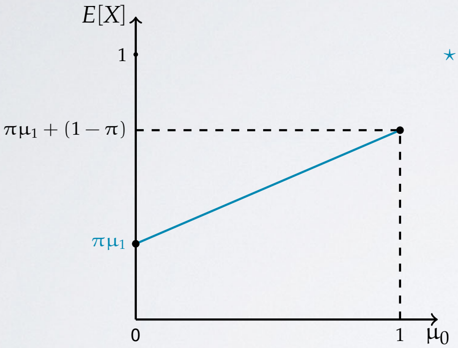
Example (Missing Data)

Data:	$\{X_i, Z_i\}$ i.i.d. with support $[0,1] \times \{0,1\}$.
Missing:	X_i observed if $Z_i = 1$.
Parameter of interest:	$\theta = E[X] = \pi \cdot \mu_1 + (1-\pi) \cdot \mu_0$.
Identified parameters:	$\mu_1 = E[X \mid Z = 1]$ and $\pi = P\{Z_i = 1\} \in (0,1)$.

Simplest Example: Missing Data

Example (Missing Data)

Data:	$\{X_i, Z_i\}$ i.i.d. with support $[0,1] \times \{0,1\}$.
Missing:	X_i observed if $Z_i = 1$.
Parameter of interest:	$\theta = E[X] = \pi \cdot \mu_1 + (1-\pi) \cdot \mu_0$.
Identified parameters:	$\mu_1 = E[X \mid Z=1]$ and $\pi = P\{Z_i = 1\} \in (0,1)$.

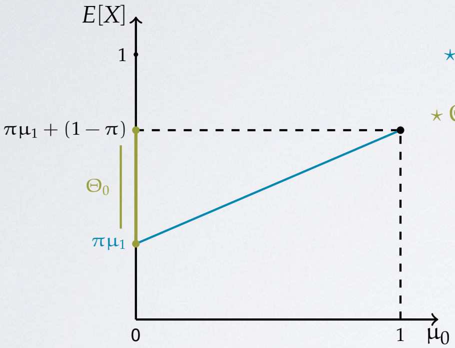


$\star\ E[X] = \pi\mu_1 + (1-\pi)\mu_0$

Simplest Example: Missing Data

Example (Missing Data)

Data:	$\{X_i, Z_i\}$ i.i.d. with support $[0,1] \times \{0,1\}$.
Missing:	X_i observed if $Z_i = 1$.
Parameter of interest:	$\theta = E[X] = \pi \cdot \mu_1 + (1-\pi) \cdot \mu_0$.
Identified parameters:	$\mu_1 = E[X \mid Z = 1]$ and $\pi = P\{Z_i = 1\} \in (0,1)$.



⋆ $E[X] = \pi\mu_1 + (1-\pi)\mu_0$

⋆ $\Theta_0(P) = \{\theta : \pi\mu_1 \leqslant \theta \leqslant \pi\mu_1 + (1-\pi)\}$

Simplest Example: Missing Data

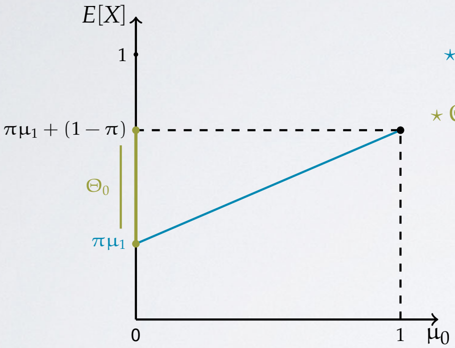
Example (Missing Data)

Data: $\{X_i, Z_i\}$ i.i.d. with support $[0,1] \times \{0,1\}$.

Missing: X_i observed if $Z_i = 1$.

Parameter of interest: $\theta = E[X] = \pi \cdot \mu_1 + (1-\pi) \cdot \mu_0$.

Identified parameters: $\mu_1 = E[X|Z=1]$ and $\pi = P\{Z_i = 1\} \in (0,1)$.



⋆ $E[X] = \pi\mu_1 + (1-\pi)\mu_0$

⋆ $\Theta_0(P) = \{\theta : \pi\mu_1 \leqslant \theta \leqslant \pi\mu_1 + (1-\pi)\}$

Moment Inequalities:

⋆ $E[m_1(W,\theta)] = E[\theta - XZ] \geqslant 0$

⋆ $E[m_2(W,\theta)] = E[1 - Z + XZ - \theta] \geqslant 0$

Partially Identified Models

- Obs. data $X \sim P \in \mathbf{P} = \{P_\gamma : \gamma \in \Gamma\}$: ($\gamma$ is possibly infinite-dim.)
- Identified set for γ:
$$\Gamma_0(P) = \{\gamma \in \Gamma : P_\gamma = P\} .$$
- Typically, only interested in $\theta = \theta(\gamma)$.
- Identified set for θ:
$$\Theta_0(P) = \{\theta(\gamma) \in \Theta : \gamma \in \Gamma_0(P)\} .$$

Example (Linear Model)

The model $\mathbf{P}$ consists of
$$Y = \theta' X + \epsilon ,$$
and a dist. P_γ specified by
$$\gamma = (\theta, P_{X,\epsilon}) \in \Gamma ,$$
where $(X, \epsilon) \sim P_{X,\epsilon}$.
Γ restricted s.t. $E_{P_\gamma}[\epsilon X] = 0$ and $E_{P_\gamma}[XX']$ invertible. Here $\theta = \theta(\gamma)$ is identified.

Partially Identified Models

- θ is identified relative to $\mathbf{P}$ if
$$\Theta_0(P) \text{ is a singleton for all } P \in \mathbf{P} .$$
- θ is unidentified relative to $\mathbf{P}$ if
$$\Theta_0(P) = \Theta \text{ for all } P \in \mathbf{P} .$$
- Otherwise, θ is partially identified relative to $\mathbf{P}$.
- $\Theta_0(P)$ has been characterized in many examples ...

... can often be characterized using moment inequalities.

Outline of Lecture

- Examples leading to moment inequalities
  - – Missing data
  - – Entry Games
  - – Revealed Preferences in Discrete Choice
- Confidence regions for partially identified models
  - – Importance of uniform asymptotic validity
- Moment inequalities: five distinct approaches
  1. Least Favorable Test
  2. subsampling
  3. Moment Selection
  4. Refined Moment Selection
  5. Two-step methods
- Subvector inference for moment inequalities
- Extensions

Example II: Entry Games

- Cross-sectional data on active firms in each market.
- Objective: estimate impact of competitors on firm profits.
- Issue: **multiple equilibria**
- The model is incomplete. Cannot use MLE.
- The model is actually **partially identified**.
- One solution is to incorporate additional restrictions:
  - Equilibrium selection assumptions (Bjorn & Vuong 1984, Berry 1992).
  - Ensure number of entrants unique (Bresnahan and Reiss 1990).
- These restrictions may not always be appropriate.
- Other approach is to use **moment inequalities**.

Example II (cont): Entry Games

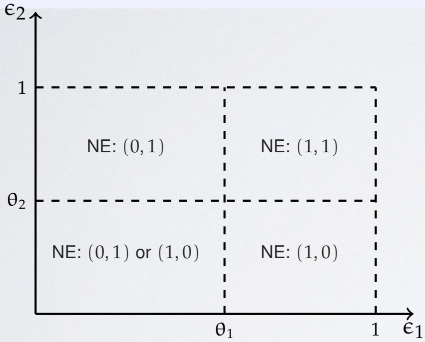
Example (2x2 entry game)

Players: $j \in \{1,2\}$ in $n \in \{1,...,N\}$ markets.

Actions: $Y_{j,n} \in \{1,0\}$ firm j's action (entry or not) in market n.

Payoff: $\pi_{j,n} = (\varepsilon_{j,n} - \theta_j Y_{-j,n})1\{Y_{j,n} = 1\}$.

- $\varepsilon_{j,n} \in [0,1]$ firm j's benefit of entry.
- $\theta_j \in (0,1)$ firm j's sensitivity to competition.



Example II (cont): Entry Games

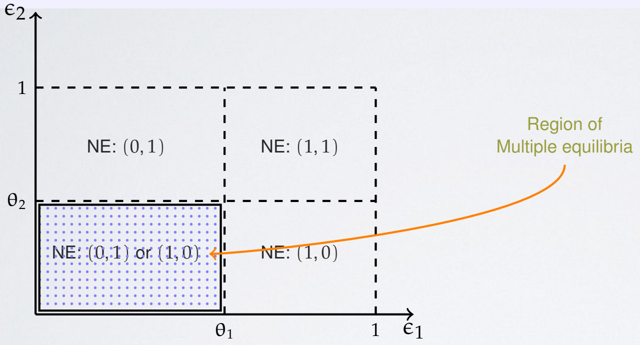
Example (2x2 entry game)

Players: $j \in \{1,2\}$ in $n \in \{1,...,N\}$ markets.

Actions: $Y_{j,n} \in \{1,0\}$ firm j's action (entry or not) in market n.

Payoff: $\pi_{j,n} = (\varepsilon_{j,n} - \theta_j Y_{-j,n})1\{Y_{j,n} = 1\}$. Econometrician observes Y_n

- $\varepsilon_{j,n} \in [0,1]$ firm j's benefit of entry. Econ: $\varepsilon_{j,n} \sim U[0,1]$.
- $\theta_j \in (0,1)$ firm j's sensitivity to competition. Econ: $\theta_0 = (\theta_1, \theta_2)$.



Example II (cont): Entry Games

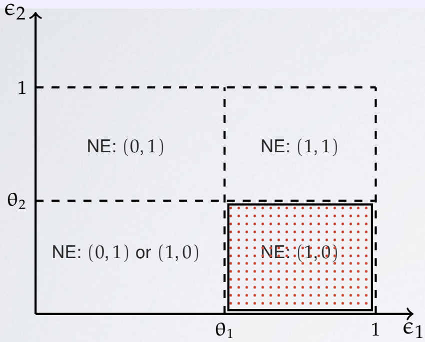
Example (2x2 entry game)

Players: $j \in \{1,2\}$ in $n \in \{1,...,N\}$ markets.

Actions: $Y_{j,n} \in \{1,0\}$ firm j's action (entry or not) in market n.

Payoff: $\pi_{j,n} = (\varepsilon_{j,n} - \theta_j Y_{-j,n})1\{Y_{j,n} = 1\}$. Econometrician observes Y_n

- $\varepsilon_{j,n} \in [0,1]$ firm j's benefit of entry. Econ: $\varepsilon_{j,n} \sim U[0,1]$.
- $\theta_j \in (0,1)$ firm j's sensitivity to competition. Econ: $\theta_0 = (\theta_1, \theta_2)$.



Without further assumptions:

$\star$ $P(1,1) = (1-\theta_1)(1-\theta_2)$

$\star$ $\theta_2(1-\theta_1) \leqslant \quad P(1,0)$

Example II (cont): Entry Games

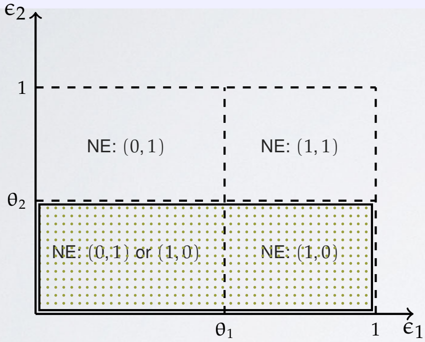
Example (2x2 entry game)

Players: $j \in \{1,2\}$ in $n \in \{1,...,N\}$ markets.

Actions: $Y_{j,n} \in \{1,0\}$ firm j's action (entry or not) in market n.

Payoff: $\pi_{j,n} = (\varepsilon_{j,n} - \theta_j Y_{-j,n}) 1\{Y_{j,n} = 1\}$. Econometrician observes Y_n

- $\varepsilon_{j,n} \in [0,1]$ firm j's benefit of entry. Econ: $\varepsilon_{j,n} \sim U[0,1]$.
- $\theta_j \in (0,1)$ firm j's sensitivity to competition. Econ: $\theta_0 = (\theta_1, \theta_2)$.



Without further assumptions:

- ⋆ $P(1,1) = (1-\theta_1)(1-\theta_2)$
- ⋆ $P(1,0) \leqslant \theta_2$

Example II (cont): Entry Games

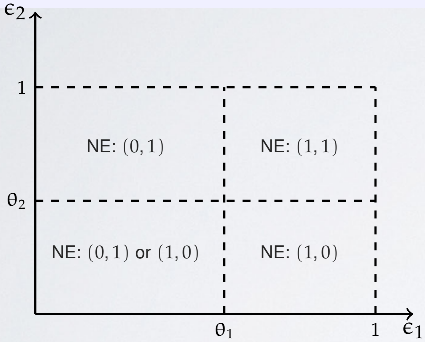
Example (2x2 entry game)

Players: $j \in \{1,2\}$ in $n \in \{1,...,N\}$ markets.

Actions: $Y_{j,n} \in \{1,0\}$ firm j's action (entry or not) in market n.

Payoff: $\pi_{j,n} = (\varepsilon_{j,n} - \theta_j Y_{-j,n}) 1\{Y_{j,n} = 1\}$. Econometrician observes Y_n

- $\varepsilon_{j,n} \in [0,1]$ firm j's benefit of entry. Econ: $\varepsilon_{j,n} \sim U[0,1]$.
- $\theta_j \in (0,1)$ firm j's sensitivity to competition. Econ: $\theta_0 = (\theta_1, \theta_2)$.



Without further assumptions:

$\star\ P(1,1) = (1-\theta_1)(1-\theta_2)$

$\star\ \theta_2(1-\theta_1) \leqslant P(1,0) \leqslant \theta_2$

Moment Inequalities:

$\star\ \mathbb{E}[Y_1 Y_2 - (1-\theta_1)(1-\theta_2)] = 0$

$\star\ \mathbb{E}[Y_1(1-Y_2) - \theta_2(1-\theta_1)] \geqslant 0$

$\star\ \mathbb{E}[\theta_2 - Y_1(1-Y_2)] \geqslant 0$

Example II (cont): Entry Games

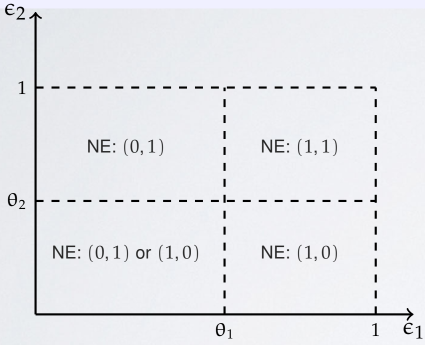
Example (2x2 entry game)

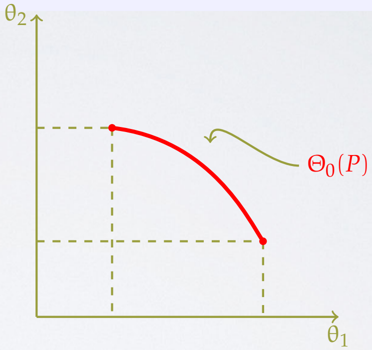
Players: $j \in \{1, 2\}$ in $n \in \{1, ..., N\}$ markets.

Actions: $Y_{j,n} \in \{1, 0\}$ firm j's action (entry or not) in market n.

Payoff: $\pi_{j,n} = (\varepsilon_{j,n} - \theta_j Y_{-j,n})1\{Y_{j,n} = 1\}$. Econometrician observes Y_n

- $\varepsilon_{j,n} \in [0, 1]$ firm j's benefit of entry. Econ: $\varepsilon_{j,n} \sim U[0, 1]$.
- $\theta_j \in (0, 1)$ firm j's sensitivity to competition. Econ: $\theta_0 = (\theta_1, \theta_2)$.





Example I (cont): Entry Games

- "Market Structure and Multiple Equilibria in Airline Markets" (Ciliberto and Tamer, 09)
- Complete information, static entry game (airlines, market = city pair)
- **Simplified version** with 2 firms deliver

$$Y_{1,m} = I\{X'\lambda_1 + \delta_1 Y_{2,m} + \epsilon_{1,m} \geqslant 0\}$$
$$Y_{2,m} = I\{X'\lambda_2 + \delta_2 Y_{1,m} + \epsilon_{2,m} \geqslant 0\}$$

- **Multiple equilibria** exist when $\epsilon_{j,m}$ in a range where both (1,0) and (0,1) satisfy these conditions.
- Model implies UB and LB on outcome probabilities for $Y = (Y_1, Y_2)$:

$$LB_{(1,0)}(\gamma, x) \leqslant P\{Y = (1,0)|X = x\} \leqslant UB_{(1,0)}(\gamma, x)$$

  - LB is probability (1,0) is unique outcome of game
  - UB is probability (1,0) is one outcome of game
- Both can be simulated as functions of $\gamma = (\lambda, \delta, F_\epsilon(e))$ and X.

Example III: Revealed Pref. in Disc. Choice

- Discrete choice demand models have **revealed preference** foundation (McFadden (1974), Berry (1994), BLP (1995))
- This approach builds on Pakes(2010) and Pakes,Porter,Ho and Ishii (2015)
- Main idea is as follows:

  Firms have **profits** $\pi_j(Y_j, Y_{-j}; X)$. The behavioral assumption is that

  $$\sup_{y \in \mathcal{Y}} E[\pi_j(Y_j = y, Y_{-j}; X)|I_j] \leqslant E[\pi_j(Y_j = S_j, Y_{-j}; X)|I_j] \quad a.s. \quad I_j$$

  - $\mathcal{Y}$: set of actions
  - S_j: strategy actually played by player j
  - I_j: Information set at the time of making the decision
- Leads to **moment inequalities**
- PPHI: analyze the number of ATMs chosen by banks.

QUESTIONS?

Confidence Regions

- If θ is identified relative to $\mathbf{P}$ (so, $\theta = \theta(P)$), then we require that

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}} P\{\theta(P) \in C_n\} \geqslant 1-\alpha \; .$$

- **Now** we require that

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}} \inf_{\theta\in\Theta_0(P)} P\{\theta \in C_n\} \geqslant 1-\alpha \; .$$

- Refer to as conf. region for points in id. set that are uniformly consistent in level.

- **Remark**: May also be interested in conf. regions for identified set itself:

$$\liminf_{n\to\infty} \inf_{P\in\mathbf{P}} P\{\Theta_0(P) \subseteq C_n\} \geqslant 1-\alpha \; .$$

- See Chernozkukov et al. (2007) and Romano & Shaikh (2010).

Test Inversion

- **Duality**: C_n can be constructed by inverting tests of each of the individual null hypotheses

$$H_\theta : \theta \in \Theta_0(P) .$$

- More specifically, suppose that for each θ a test of H_θ, $\phi_n(\theta)$, is available that satisfies

$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}} \sup_{\theta\in\Theta_0(P)} E_P[\phi_n(\theta)] \leqslant \alpha .$$

- It follows that C_n equal to the set of $\theta \in \Theta$ for which H_θ is accepted is uniformly consistent in levels,

$$C_n = \{\theta \in \Theta : \phi_n(\theta) = 0\} .$$

- **Computational note**: this requires to explore the parameter space Θ.

Confidence Regions (cont.)

Uniform consistency in level

$$\liminf_{n \to \infty} \inf_{P \in \mathbf{P}} \inf_{\theta \in \Theta_0(P)} P\{\theta \in C_n\} \geqslant 1 - \alpha \, .$$

Pointwise consistency in level

$$\liminf_{n \to \infty} P\{\theta \in C_n\} \geqslant 1 - \alpha \text{ for all } P \in \mathbf{P} \text{ and } \theta \in \Theta_0(P) \, .$$

- **Pointwise**: possible that $\forall n \; \exists P \in \mathbf{P}$ and $\theta \in \Theta_0(P)$ with cov. prob. $\ll 1 - \alpha$.
- In well-behaved prob., distinction is entirely technical issue.
- In less well-behaved prob., distinction is **more important**.
- Some "natural" conf. reg. may need to restrict $\mathbf{P}$ in non-innocuous ways.

  (e.g., may need to assume model is "far" from identified.)

Example

Example

Let $W_i = (L_i, U_i), i = 1, \ldots, n$ be i.i.d. $P \in \mathbf{P}$ with

$$\mathbf{P} = \{N(\mu, \Sigma) : \mu = (\mu_L, \mu_U) \in \mathbf{R}^2 \text{ with } \mu_L < \mu_U\} ,$$

where Σ is a known covariance matrix with unit variances.

- Suppose there is a parameter of interest θ.
- θ is known to belong to the identified set

$$\Theta_0(P) = [\mu_L(P), \mu_U(P)] .$$

- Consider the confidence region

$$C_n = \left[\bar{L}_n - \frac{z_{1-\alpha}}{\sqrt{n}}, \bar{U}_n + \frac{z_{1-\alpha}}{\sqrt{n}}\right] \quad \text{where} \quad \bar{L}_n = \frac{1}{n}\sum_{i=1}^{n} L_i \quad \text{and} \quad \bar{U}_n = \frac{1}{n}\sum_{i=1}^{n} U_i .$$

Example

Claim: C_n is pointwise consistent in level.

Example

Claim: C_n is not uniformly consistent in levels:

$$\inf_{P\in\mathbf{P}} \inf_{\theta\in\Theta_0(P)} P\{\theta \in C_n\} = 1 - 2\alpha < 1 - \alpha \,,$$

QUESTIONS?

Moment Inequalities

- **Data:** $W_i, i = 1, \ldots, n$ are i.i.d. with common distr. $P \in \mathbf{P}$.
- Numerous examples of partially identified models give rise to **moment inequalities**:

$$\Theta_0(P) = \{\theta \in \Theta : E_P[m(W_i, \theta)] \leqslant 0\} ,$$

where m takes values in $\mathbf{R}^k$.

- **Goal**: Confidence regions for points in the id. set that are uniformly consistent in level.

UNIFORM INTEGRABILITY CONDITION

$$\sup_{P \in \mathbf{P}} \sup_{\theta \in \Theta_0(P)} E_P \left[\left(\frac{m_j(W_i, \theta) - \mu(\theta, P)}{\sigma_j(\theta, P)} \right)^2 I \left\{ \frac{m_j(W_i, \theta) - \mu(\theta, P)}{\sigma_j(\theta, P)} > t \right\} \right] \to 0 ,$$

as $t \to \infty$.

- Mild condition that ensures CLT and LLN hold unif. over $P \in \mathbf{P}$ and $\theta \in \Theta_0(P)$.

Moment Inequalities: Test

- **How**: Construct tests $\phi_n(\theta)$ of
$$H_\theta : E_P[m(W_i, \theta)] \leqslant 0$$
that provide unif. asym. control of Type I error, i.e.,
$$\limsup_{n\to\infty} \sup_{P\in\mathbf{P}} \sup_{\theta\in\Theta_0(P)} E_P[\phi_n(\theta)] \leqslant \alpha \; .$$

- Given such $\phi_n(\theta)$,
$$C_n = \{\theta \in \Theta : \phi_n(\theta) = 0\}$$
satisfies desired coverage property.

- Below describe **five different tests**, all of form
$$\phi_n(\theta) = I\{T_n(\theta) > \hat{c}_n(\theta, 1-\alpha)\} \; .$$

Notation

Some basic notation:

$\hat{P}_n$ = empirical distribution of $W_i, i = 1, \ldots, n$.

$\mu(\theta, P) = E_P[m(W_i, \theta)]$.

$\bar{m}_n(\theta)$ = sample mean of $m(W_i, \theta)$.

$\hat{\Omega}_n(\theta)$ = sample correlation of $m(W_i, \theta)$.

$\sigma_j^2(\theta, P) = \text{Var}_P[m_j(W_i, \theta)]$.

$\hat{\sigma}_{n,j}^2(\theta)$ = sample variance of $m_j(W_i, \theta)$.

$\hat{D}_n(\theta) = \text{diag}(\hat{\sigma}_{n,1}(\theta), \ldots, \hat{\sigma}_{n,k}(\theta))$.

Test Statistic

Test statistic

For an appropriate choice of $T(x, V)$, we use

$$T_n(\theta) = T\left(\hat{D}_n^{-1}(\theta)\sqrt{n}\bar{m}_n(\theta), \hat{\Omega}_n(\theta)\right) .$$

$$T_n^{\text{mmm}}(\theta) = \sum_{1 \leqslant j \leqslant k} \max\left\{\frac{\sqrt{n}\bar{m}_{n,j}(\theta)}{\hat{\sigma}_{n,j}(\theta)}, 0\right\}^2$$

$$T_n^{\max}(\theta) = \max\left\{\max_{1 \leqslant j \leqslant k} \frac{\sqrt{n}\bar{m}_{n,j}(\theta)}{\hat{\sigma}_{n,j}(\theta)}, 0\right\}$$

$$T_n^{\text{ad,qlr}}(\theta) = \inf_{t \in \mathbf{R}^k : t \leqslant 0} \left(\hat{D}_n^{-1}(\theta)\sqrt{n}\bar{m}_n(\theta) - t\right)' \tilde{\Omega}_n(\theta)^{-1} \left(\hat{D}_n^{-1}(\theta)\sqrt{n}\bar{m}_n(\theta) - t\right) ,$$

where

$$\tilde{\Omega}_n(\theta) = \max\{\epsilon - \det(\hat{\Omega}_n(\theta)), 0\} I_k + \hat{\Omega}_n(\theta)$$

for some fixed $\epsilon > 0$, with I_k denoting the k-dimensional identity matrix.

Critical Value I

Useful to define

$$J_n(x, s(\theta), \theta, P) = P\left\{T\left(\hat{D}_n^{-1}(\theta)Z_n(\theta) + \hat{D}_n^{-1}(\theta)s(\theta), \hat{\Omega}_n(\theta)\right) \leqslant x\right\} ,$$

where

$$Z_n(\theta) = \sqrt{n}(\bar{m}_n(\theta) - \mu(\theta, P)) .$$

Easy to estimate for a given function $s(\theta)$, e.g.,

1. Nonparametric bootstrap estimator: $J_n(x, s(\theta), \theta, \hat{P}_n)$
2. Asymptotic Approximation estimator: $J_n(x, s(\theta), \theta, \tilde{P}_n(\theta))$, where

$$Z_n(\theta) \sim N(0, \hat{\Sigma}_n(\theta)) \quad \text{under} \quad \tilde{P}_n(\theta) .$$

Difficult to estimate

$$J_n(x, \sqrt{n}\mu(\theta, P), \theta, P) = P\{T_n(\theta) \leqslant x\}$$

See, e.g., Andrews (2000).

Critical Value II

- **Goal:** to estimate the distribution of $T_n(\theta)$,

$$P\{T_n(\theta) \leqslant x\} = J_n(x, \sqrt{n}\mu(\theta, P), \theta, P) \ .$$

- **Problem:** it is not possible to estimate $\sqrt{n}\mu(\theta, P)$ consistently.

- Its natural estimator $\sqrt{n}\bar{m}_n(\theta)$ satisfies

$$|\sqrt{n}\bar{m}_n(\theta) - \sqrt{n}\mu(\theta, P)| \xrightarrow{d} |N(0, \Sigma(\theta, P))|$$

  under any fixed $\theta \in \Theta_0(P)$ and $P \in \mathbf{P}$, where $\Sigma(\theta, P) = \text{Var}_P[m(W_i, \theta)]$.

- **Five different tests** distinguished by how they circumvent this problem.

- **Trick:** exploit that $T(x, V)$ is weakly increasing in each component of its first argument.

QUESTIONS?

Least Favorable Test

$$H_\theta : E_P[m(W_i,\theta)] = \mu(\theta,P) \leqslant 0 \ .$$

- **Main Idea**: exploit **monotonicity** of $T(\cdot, V)$.
- $\sqrt{n}\mu(\theta,P) \leqslant 0$ for any $P \in \mathbf{P}$ and $\theta \in \Theta_0(P)$ thus imply

$$J_n^{-1}(1-\alpha, \sqrt{n}\mu(\theta,P), \theta, P) \leqslant J_n^{-1}(1-\alpha, 0, \theta, P) \ .$$

- Choosing

$$\hat{c}_n(1-\alpha,\theta) = \text{ estimate of } J_n^{-1}(1-\alpha,0,\theta,P)$$

  therefore leads to valid tests.
- 0_k is the **least favorable value** of the nuisance parameter $\sqrt{n}\mu(\theta,P)$

  "All moments are binding": $\mu(P,\theta) = 0$.
- See Rosen (2008) and Andrews & Guggenberger (2009).

  Closely related work by Kudo (1963) and Wolak (1987, 1991).

Least Favorable Test

Least Favorable Test

The least favorable test takes the form

$$\phi_n^{\text{lf}}(\theta) \equiv I\{T_n(\theta) > \widehat{J}_n^{-1}(1-\alpha, 0_k, \theta)\} ,$$

where $\widehat{J}_n(x, 0_k, \theta)$ equals either $J_n(x, 0_k, \theta, \hat{P}_n)$ or $J_n(x, 0_k, \theta, \tilde{P}_n(\theta))$.

- These tests are uniformly consistent in levels.
- In our simple example, this test uses

$$C_n = \left[\bar{L}_n - \frac{z_{1-\alpha/2}}{\sqrt{n}}, \bar{U}_n + \frac{z_{1-\alpha/2}}{\sqrt{n}} \right] ,$$

  instead of

$$C_n = \left[\bar{L}_n - \frac{z_{1-\alpha}}{\sqrt{n}}, \bar{U}_n + \frac{z_{1-\alpha}}{\sqrt{n}} \right] .$$

- In other words, the least favorable confidence region assumes $\mu_U(P) - \mu_L(P) = 0$.

Least Favorable Test

- **Remark**: Deemed "conservative," but criticism not entirely fair:
  - In Gaussian setting, these tests are (α- and d-) admissible.
  - Some are even maximin optimal among restricted class of tests.
  - See Lehmann (1952) and Canay & Shaikh (2016).
- Nevertheless, **unattractive**:
  - Tend to have best power against alternatives with **all** moments > 0.
  - As θ varies, many alternatives with only **some** moments > 0.
  - May therefore not lead to smallest confidence regions.
- Following tests incorporate info. about $\sqrt{n}\mu(\theta, P)$ in some way.

  $\Longrightarrow$ better power against such alternatives.

Subsampling

- **Main Idea**: Fix $b = b_n < n$ with $b \to \infty$ and $b/n \to 0$.

  Compute $T_n(\theta)$ on each of $N_n = \binom{n}{b}$ subsamples of data.

- Denote by $L_n(x, \theta)$ the empirical distr. of these quantities,

$$L_n(x, \theta) = \frac{1}{N_n} \sum_{\ell=1}^{N_n} I\left\{T_{b,\ell}(\theta) \leqslant x\right\} ,$$

- Use $L_n(x, \theta)$ as estimate of distr. of $T_n(\theta)$, i.e.,

$$J_n(x, \sqrt{n}\mu(\theta, P), \theta, P) .$$

- **Critical value**: choosing

$$\hat{c}_n(1 - \alpha, \theta) = L_n^{-1}(1 - \alpha, \theta)$$

  leads to valid tests.

  See Romano & Shaikh (2008) and Andrews & Guggenberger (2009).

Subsampling (cont.)

Subsampling

The subsampling test takes the form

$$\phi_n^{\text{sub}}(\theta) = I\left\{T_n(\theta) > L_n^{-1}(1-\alpha,\theta)\right\} .$$

- Note that $L_n(x)$ is a “good” estimator of

$$P\{T(\hat{D}_b(\theta)^{-1}\sqrt{b}(\bar{m}_b(\theta)-\mu(\theta,P))+\hat{D}_b(\theta)^{-1}\sqrt{b}\mu(\theta,P),\hat{\Omega}_b(\theta))\leqslant x\} ,$$

  which we denote by

$$J_b(x,\sqrt{b}\mu(\theta,P),\theta,P) .$$

- **Size b distribution:** for any $\epsilon > 0$, $L_n(x,\theta)$ satisfies

$$\sup_{x\in\mathbf{R}}\sup_{P\in\mathbf{P}}\sup_{\theta\in\Theta_0(P)} P\left\{\sup_{x\in\mathbf{R}}\left|L_n(x,\theta)-J_b(x,\sqrt{b}\mu(\theta,P),\theta,P)\right|>\epsilon\right\}\to 0 .$$

- However, we want $J_n(x,\sqrt{n}\mu(\theta,P),\theta,P)$.

Subsampling (cont.)

- **Trick**: Link J_b to J_n and then account for $\sqrt{b}\mu(\theta, P)$ vs $\sqrt{n}\mu(\theta, P)$.
- Link $J_b(x, \sqrt{b}\mu(\theta, P), \theta, P)$ with $J_n(x, \sqrt{b}\mu(\theta, P), \theta, P)$ by exploiting

$$\sup_{P \in \mathbf{P}} \sup_{\theta \in \Theta_0(P)} \sup_{s \leqslant 0} \left| J_b(x, s, \theta, P) - J_n(x, s, \theta, P) \right| \to 0 \ .$$

- Next note that

$$\sqrt{n}\mu(\theta, P) \leqslant \sqrt{b}\mu(\theta, P)$$

  for any $P \in \mathbf{P}$ and $\theta \in \Theta_0(P)$

$$\Longrightarrow J_n^{-1}(1 - \alpha, \sqrt{n}\mu(\theta, P), \theta, P) \leqslant J_n^{-1}(1 - \alpha, \sqrt{b}\mu(\theta, P), \theta, P) \ .$$

- The SS critical value is a valid **upper bound**.
- See general results in Romano & Shaikh (2012).
- **Remark**: Incorporates information about $\sqrt{n}\mu(\theta, P)$ ...

  ... but remains unattractive because **choice of b** problematic.

References

[1] Donald W. K. Andrews and Patrik Guggenberger. Validity of subsampling and “plug-in asymptotic" inference for parameters defined by moment inequalities. *Econometric Theory*, 25(3):669–709, June 2009.

[2] Donald W. K. Andrews and Gustavo Soares. Inference for parameters defined by moment inequalities using generalized moment selection. *Econometrica*, 78(1):119–158, January 2010.

[3] F. A. Bugni. Bootstrap inference in partially identified models defined by moment inequalities: Coverage of the identified set. *Econometrica*, 78(2):735–753, April 2010.

[4] Ivan Alexis Canay. El inference for partially identified models: Large deviations optimality and bootstrap validity. *Journal of Econometrics*, 156(2):408–425, June 2010.

[5] Victor Chernozhukov, H. Hong, and Elie Tamer. Estimation and confidence regions for parameter sets in econometric models. *Econometrica*, 75(5):1243–1284, 2007.

[6] G. Imbens and C. F. Manski. Confidence intervals for partially identified parameters. *Econometrica*, 72(6):1845–1857, November 2004.

[7] Joseph P. Romano and Azeem M. Shaikh. Inference for the identified set in partially identified econometric models. *Econometrica*, 78(1):169–212, September 2010.

THE END!