Adversarial Sampling

NBER Law & Economics, July 29, 2026

Alma Cohen[1,2] Alon Klement[3] Zvika Neeman[1,4] Eilon Solan[4]

[1] Berglas School of of Economics, Tel Aviv University
[2] Harvard Law School
[3] Buchmann Faculty of Law, Tel Aviv University
[4] European University Institute (EUI), Florence
[5] School of Mathematical Sciences, Tel Aviv University

Motivation

- Many institutional settings require selecting a small representative subset from a large population under conditions of conflict.
    - jury selection,
    - Multi-District Litigation (MDL) bellwether trials,
    - citizen assemblies

Motivation

- This is a problem because:
  - random sampling can generate a very unrepresentative sample,
  - informed parties have opposing interests over the composition of the chosen sample.
- We provide a first-best solution to this problem through a new selection procedure.

Formal Description of Problem

- **Given:** A population $\{x_1, \dots, x_n\}$ and a complete and transitive ranking $\succeq$.
- The objective of a planner who does not know $\succeq$ is to obtain a representative sample of the population with size $k \ll n$.
- **Important.** Two players agree on the ranking $\succeq$, but may disagree on the specific values of $\{x_1, \dots, x_n\}$.
- Player 1 and 2 prefer sampling high and low, ranking items, respectively.
- How is it possible to harness the players' knowledge to obtain a representative sample?

Suggested Solutions

- Jury selection
  - Random selection
  - Struck jury and Strike & Replace
  - Median of medians
- Multi-district litigation (MDL): selection by parties
- Citizen assemblies: theoretical computer science methods that maximize selection probabilities subject to demographic fairness constraints.

Related Literature

- Papers related to the three examples: Jury selection, Multi-district litigation (MDL), citizen assemblies, other?!
- Nash implementation
- Papers in mechanism design study Pareto-efficient selection. This is uninformative when preferences are diametrically opposed, since in that case every outcome is Pareto efficient.
- Cut & Choose mechanisms ("Cake-cutting")

MODEL

Player I prefers items that are ranked higher, while Player II prefers items that are ranked lower.

Need to define players' preferences over samples of size k.

Definition (Utility functions)

A utility function for Player I, II, is a function u_I, u_{II}, that assigns a real number to every sample $y = (y_i)_{i=1}^k$ of k items, such that $u_I(y) \geq u_I(y')$, $u_I(y) \leq u_I(y')$ if y if y is "shifted to the right" relative to y'.

Note:

1. Players' utilities satisfy vNM axioms (expected utility)
2. Allow for $u_I \neq u_{II}$.

A selection procedure, or mechanism, is a game-form that selects a (possibly random) set of k indices.

Definition (Mechanism)

A mechanism is a tuple $M = \langle A_I, A_{II}, f \rangle$ where

- A_I and A_{II} are sets of messages for Players I and II, resp.;
- f is a function that for every $a_I \in A_I, a_{II} \in A_{II}$, returns a probability distribution over the collection of all subsets of indices $\{1, \dots, n\}$ that have size k.

1. Equivalence between extensive and strategic forms implies sequential mechanisms also covered ("anything goes").
2. Since we implement the first-best, description of class of selection procedures is anyway less important.

- All the mechanisms described earlier … random sampling, party choice, strike & replace, etc.

- **Cut & Choose.** One player partitions the set of indices $\{1, \ldots, n\}$ into k disjoint subsets, and the other player selects one index from each subset.

  The choosing player prefers:

  - "no overlap": $\{1, 2, 3, 4, 5\}$ and $\{6, 7, 8\}$ is better than $\{1, 2, 4, 5, 7\}$, $\{3, 6, 8\}$

  - "shifting small subsets to the right": $\{1, 2, 3, 4, 5\}$ and $\{6, 7, 8\}$ is better than $\{1, 2, 3\}$ and $\{4, 5, 6, 7, 8\}$

Taken together, a selection procedure (mechanism) and players' preferences induce a Mechanism Game.

By Nash's Theorem, every mechanism game admits an equilibrium (possibly in mixed strategies).

We focus on this equilibrium as our prediction for how the mechanism game would be played.

Measuring Representativeness

- How can we measure the representativeness of a sample?
  - sample should include median, 5% quantile, 95% quantile?
  - Sample average should try to match the population average (what does that even mean if players disagree about values of items?)
- We measure representativeness by the distance between the population and the sample CDFs.

Population and Sample CDFs

For any population $x = \{x_i\}_{i=1}^n$ and ranking $\succsim$ on x, the CDF of x is given by:

$$F_x(x_i) := \frac{1}{n} \cdot \#\{j \in \{1, \dots, n\} \colon x_j \precsim x_i\}, \quad \forall i \in \{1, \dots, n\}.$$

For any sample $y = (y_i)_{i=1}^k$ of k items from x, the CDF of y is given by:

$$F_y(x_i) := \frac{1}{k} \cdot \#\{j \in \{1, \dots, k\} \colon y_j \precsim x_i\}, \quad \forall i \in \{1, \dots, n\}.$$

- In F_x steps are $\frac{1}{n}$; in F_y steps are $\frac{1}{k}$.

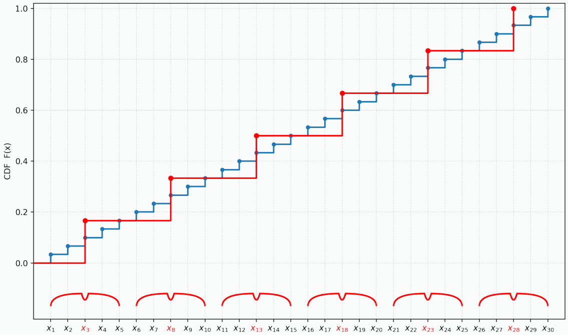

Figure 1: F_x and F_y when $n = 30, k = 6, y = \{x_3, x_8, x_{13}, x_{18}, x_{23}, x_{28}\}$.

Given a population $x = \{x_i\}_{i=1}^n$, a ranking $\succsim$ on x, and a sample $y = (y_i)_{i=1}^k$ of k, we measure how well the sample y represents the population x by the:

- **Kolmogorov-Smirnov (KS)** statistic:

$$\mathcal{KS}(x, y) := \max_{1 \le i \le n} |F_x(x_i) - F_y(x_i)|.$$

- L_1 statistic:

$$L_1(x, y) := \frac{1}{n} \sum_{i=1}^{n} |F_x(x_i) - F_y(x_i)|.$$

- **Cramér-von Mises (CvM)** statistic:

$$\mathcal{C}v\mathcal{M}(x, y) := \frac{1}{n} \sum_{i=1}^{n} (F_x(x_i) - F_y(x_i))^2.$$

Definition (KS-Optimality)

A mechanism $M^* = (A_I^*, A_{II}^*, f^*)$ is KS-optimal if for every population x of size n and every ranking $\succsim$ on x, for every equilibrium (a_I^*, a_{II}^*) of M^* and for every equilibrium (a_I, a_{II}) of every other mechanism $M = (A_I, A_{II}, f)$, every sample y^* that may be obtained under (a_I^*, a_{II}^*), and every sample y that may be obtained under (a_I, a_{II}),

$$\mathcal{KS}(x, y) \geq \mathcal{KS}(x, y^*).$$

Definition (L_1-Optimality)

A mechanism $M^* = (A_I^*, A_{II}^*, f^*)$ is L_1-optimal if for every population x of size n and every ranking $\succsim$ on x, for every equilibrium (a_I^*, a_{II}^*) of M^* and for every equilibrium (a_I, a_{II}) of every other mechanism $M = (A_I, A_{II}, f)$, every sample y^* that may be obtained under (a_I^*, a_{II}^*), and every sample y that may be obtained under (a_I, a_{II}),

$$L_1(x, y) \geq L_1(x, y^*).$$

Definition (CvM-Optimality)

A mechanism $M^* = (A^*_I, A^*_{II}, f^*)$ is CvM-optimal if for every population x of size n and every ranking $\succsim$ on x, for every equilibrium (a^*_I, a^*_{II}) of M^* and for every equilibrium (a_I, a_{II}) of every other mechanism $M = (A_I, A_{II}, f)$, every sample y^* that may be obtained under (a^*_I, a^*_{II}), and every sample y that may be obtained under (a_I, a_{II}),

$$\mathcal{C}v\mathcal{M}(x, y) \geq \mathcal{C}v\mathcal{M}(x, y^*).$$

The Cut, Strike & Choose Mechanism

Suppose that $n = (2m+1)k$.

- The partitioning player partitions the population into k equally sized subsets (with $2m+1$ items in each).
- The partitioning player strikes out m items from one of the chosen subsets.
- The choosing player chooses one item from each subset.

Equivalently,

- The first player chooses k disjoint subsets: one with $m+1$ items, and all the others with $2m+1$ items (m items are not assigned and cannot be chosen).
- the second player chooses one item from each set.

Equivalently,

- The first player strikes m items;
- The second player chooses one item;
- The first player strikes $2m$ items;
- The choosing player chooses one item;
- And so on until k items are chosen.

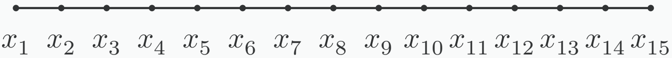

Figure 2: $n = 15, k = 1, m = 7$, or $k = 3, m = 2$, or $k = 5, m = 1$

Observe that the CSC mechanism is symmetric. The same sample is chosen regardless of which player performs what role.

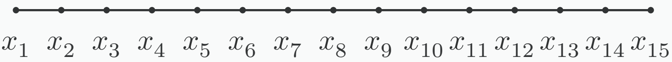

Figure 3: $n = 15, k = 1, m = 7$, or $k = 3, m = 2$, or $k = 5, m = 1$

Theorem

Suppose that $n = k(2m+1)$. Then, the Cut, Strike & Choose mechanism chooses the first-best sample (what the planner would have chosen if she knew the ranking herself). It is KS-optimal, L_1-optimal, and CvM-optimal. Moreover, for every population x and ranking $\succeq$, for every sample y' that is not equivalent to the sample that is produced by the CSC mechanism:

$$\mathcal{KS}(x, y') > \mathcal{KS}(x, y) = \frac{m}{n},$$

$$L_1(x, y') > L_1(x, y) = \frac{m}{4n},$$

$$\mathcal{C}v\mathcal{M}(x, y') > \mathcal{C}v\mathcal{M}(x, y).$$

Population and CSC Sample CDFs

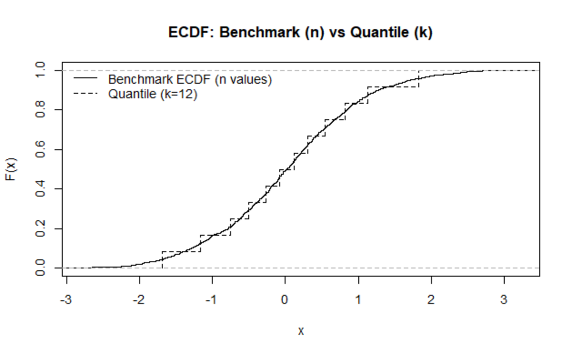

Figure 4: F_x and F_y under the CSC mechanism for $n = 972, k = 12, m = 40$.

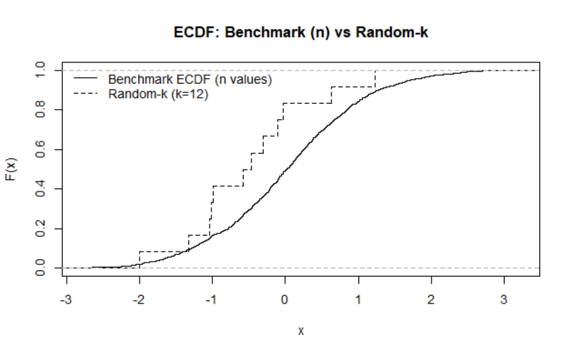

Figure 5: F_y under random selection $(n = 972, k = 12, m = 40)$.

CSC vs Random Sample

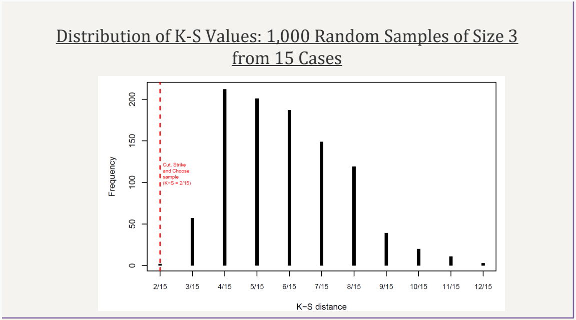

Figure 6: CSC vs Random Samples with 1000 draws from a population with $n = 15$ and $k = 3$.

Comparison of Different Selection Procedures

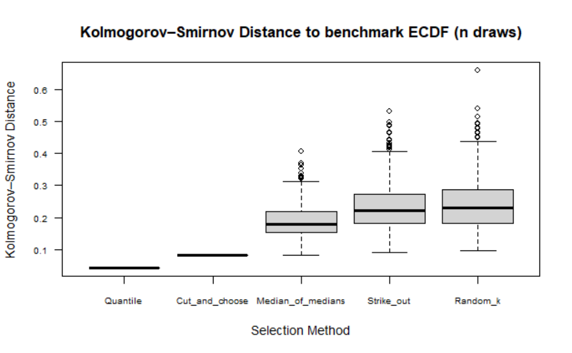

Figure 7: Comparison of the Quantile, I Cut/You Choose, Median of medians, Struck Jury and random mechanisms in terms of KS statistic for sampling 1000 times from $n = 972, k = 12, m = 40$.

Sketch of Proof of KS-Optimality (1)

Fix a population $x = (x_1, \dots, x_n)$ and a ranking $\succsim$ on x. Assume WLOG that $x_1 \precsim x_2 \precsim \cdots \precsim x_n$. In the equilibrium of the Quantile mechanism, the selected sample is equivalent to

$$\begin{aligned} y &= \left(x_{m+1}, x_{m+1+(2m+1)}, \cdots, x_{m+1+(2m+1)(k-1)}\right) \\ &= \left(x_{m+1}, x_{m+1+(2m+1)}, \cdots, x_{n-m}\right). \end{aligned}$$

We show that $\mathcal{KS}(x, y) \leq \dfrac{m}{n} < \mathcal{KS}(x, y')$, for any sample y' that is not equivalent to y.

Sketch of Proof of KS-Optimality (2)

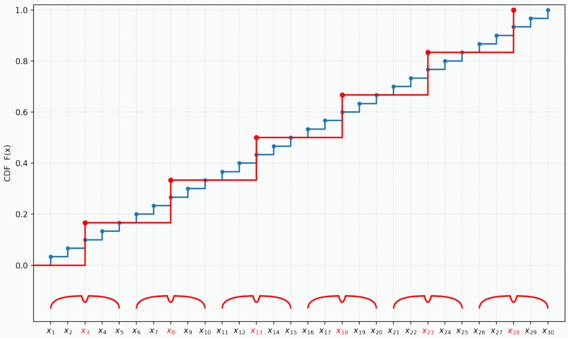

Figure 8: Sketch of proof with $n = 30, k = 6, m = 2$.

- $n = k(2m+1)$. Quantile selects $x_3, x_8, x_{13}, x_{18}, x_{23}, x_{28}$.
- Just before first jump, difference is $\frac{m}{n}$; at jump it is $\frac{1}{k} - \frac{m+1}{n} = \frac{2m+1}{n} - \frac{m+1}{n} = \frac{m}{n}$

Extensions

What if you prefer to choose a sample that consists of the median, the 5% quantile, and the 95% quantile?

Theorem

Any selection of k-quantiles is possible, through a modified version of the CSC mechanism.

Disagreement over the Ranking

Suppose that the two players have rankings $\succsim_1$ and $\succsim_2$ that are not necessarily completely opposed to each other.

Denote the sample chosen by the CSC mechanism by $y(x)$.

Denote the sample chosen by the CSC mechanism when the players rankings are given by $\succsim_i$ and $\succsim_j = -\succsim_i$ by $y_i^*(x)$.

Theorem

Given a population x and two rankings $\succsim_1$ and $\succsim_2$, we have

i. $u_1(y(x)) \geq u_1(y_1^*(x))$ *under* $\succsim_1$.
ii. $u_2(y(x)) \geq u_2(y_2^*(x))$ *under* $\succsim_2$.

Only One Informed Player

Suppose the items in the population are real numbers that are ranked by their size.

Theorem

Suppose that $n = km$. Under the Equal Cut & Random Choice mechanism, for any partition of the population into k equally sized subsets P,

$$\mu_P(y) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

Moreover, the variance $[\sigma_P^2]$ is minimized by the ordered partition, in which the items are assigned to the k subsets according to their order, so that the largest item in each subset is smaller than the smallest item in the next subset.

It follows that the player cannot affect the mean of the chosen sample. If the player is risk-averse and prefers the variance of the sample to be as small as possible, then the Equal Cut & Random Choice mechanism induces the ordered partition. In this partition, the median items of each subset in the partition coincide with the quantiles selected by the CSC mechanism, respectively, by their order.

Thank You !