Inference with Rational Agents

- Recall that in our setting, agents are rational (i.e.: have complete and transitive preferences) and are utility maximizing agents:

$$u_i^* = \arg \max_{u \in U(x)} \{U(x_i, u; \theta)\}$$

- Where:
  - x_i is the state of the "environment"
  - u_i^* is the action/control taken by the agent
  - θ is the type (private information) that we wish to infer

- And we can use Inverse Optimization to **infer** the private information θ.

Inference with Rational Agents

- The inference is based on observation of the agent interacting with the environment. We collect observations of state-action pairs:

$$((x_1, u_1^*), (x_2, u_2^*), ..., (x_n, u_n^*))$$

- And we posed our inference as the following optimization problem:

$$\hat{\theta} = \underset{\theta \in \Theta}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$u_i^* = \arg \max_{u \in U(x_i)} \{U(x_i, u, \theta)\}, \forall i \in \{1, ..., n\}$$

- Where Θ is some bounded set of possible type values.

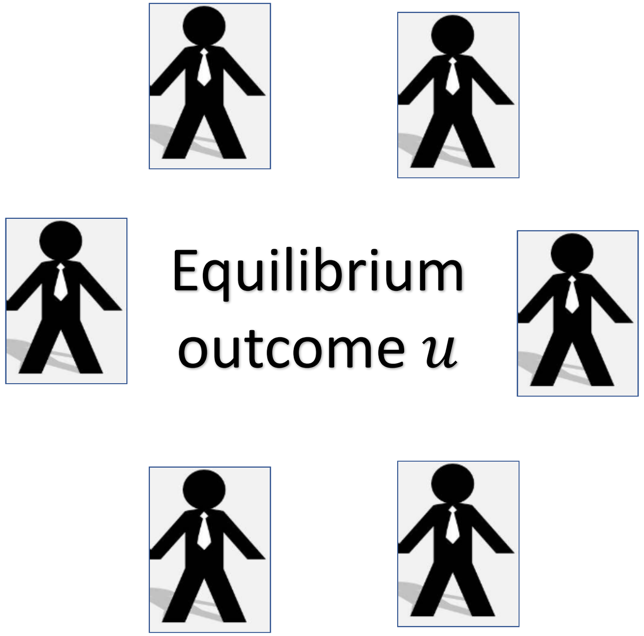
Inference with Rational Agents

- This problem is called an **Inverse Optimization Problem.**
- It is a feasibility problem, where the feasible region is given by solutions of optimization problems.
- These in general are very hard to solve, because a computer cannot handle constraints that themselves are optimization problems.
- We studied a simple case, where the solution is in fact tractable.
    - Linear constraints and quadratic utilities
    - "Weighted sum" of criterions, where the private information are the weights
    - Suboptimal Agents
    - Kernel-based Utilities

Inference with Multiple Rational Agents

- On this lecture, we will explore how can we extend the Inverse Optimization formulation to the case where there are multiple agents interacting with each other.

- First we need a concept to characterize such interaction. We will assume the agents interact until some outcome is reached, which we call an **Equilibrium**:



The Nash Equilibrium

- On such concept is the **Nash Equilibrium(NE).** It is a very natural notion that appears in many different contexts (even in nature and in animal behavior).

- Let’s start with an example (perhaps the most famous) of the *Prisoner’s Dilemma:*

- Two Prisoners are in jail and are brought to the prosecutors room. The prosecutor offers a bargain and asks which of them committed the crime. Each prisoner can respond:
  - (1) Blame the other
  - (2) Stay silent

- If both prisoners stay silent they will serve 1 year in prison. If both blame each other they will serve 2 years in prison.

- If one stays silent and the other blames them, the one who blames goes free, and the other serves 3 years in prison.

The Prisoner's Dilemma

- In this environment, we assume that both agents (the prisoners) desire to minimize the time they spent in jail.

- They only have two action choices: $u \in \{\text{blame the other}, \text{stay silent}\}$.

- The prisoner's dilemma is a special kind of game:
    - It has a finite action space
    - It has a finite number of agents
    - It is a static game (there are no "state space" nor dynamics)

- In this game, both agent's act "at the same time", without prior knowledge of what the other will do.

The Prisoner’s Dilemma

- This allow us to write the interaction problem (i.e.: the game) in what is called the *Normal Form Representation*:

	Stay Silent	Blame
Stay Silent	$-1, -1$	$-3, 0$
Blame	$0, -3$	$-2, -2$

- The Nash Equilibrium can be described informally in words:

- “It is the outcome, where any agent has no incentive to **unilaterally** deviate from their own chosen strategy(action choice).”

The Prisoner's Dilemma

- Game in the Normal Form can be solved by looking for unilateral deviations:

	Stay Silent	Blame
Stay Silent	$\cancel{-1,-1}$	$-3, 0$
Blame	$0, -3$	$-2, -2$

- Hence the NE is the outcome $(-2, -2)$, where both agents blame each other.

- Note that this is **not** the best outcome possible! If the agents cooperated (or had mutual trust), they could have both remained silent to reach a strictly better outcome.

The Nash Equilibrium

- As this example shows, the NE does not capture the notion of “optimality” of outcome.

- It captures the notion of agent’s behaving strategically, since they **know** that the other agents behave strategically as well.

- Let’s formalize the NE definition.

- Let $U_i\left(x, u_1, \dots, u_i, \dots u_p\right)$ be the Utility Function of agent $i \in \{1, \dots, p\}$.

- Note that their Utility depend on their own actions u_i but also depends on the actions of the other $p-1$ players.

- To simplify notation we will define $u_{-i} = \left(u_1, u_2, \dots, u_{i-1}, u_{i+1}, \dots, u_p\right)$

The Nash Equilibrium

- Let $U_i(x)$ be the set of possible actions, the agent i can chose from.

- Here x denote the state of the environment.

- For example, in the Prisoners Dilemma, the agents are *symmetric* (they have the same Utility Function), defined as:

$$U_i(\text{blame}, \text{blame}) = -2$$

$$U_i(\text{blame}, \text{stay silent}) = 0$$

$$U_i(\text{stay silent}, \text{blame}) = -3$$

$$U_i(\text{blame}, \text{blame}) = -1$$

- and in this case there are no state, so we omit x.

The Nash Equilibrium

- We say that a action vector (also called **Strategy Profile** in Game Theory) $(u_1^*, \ldots, u_p^*)$ is a Nash Equilibrium if:

$$u_i^* = \arg\max_{u \in U_i(x)} \{U_i(x, u, u_{-i}^*)\}, \forall i \in \{1, ..., p\}$$

- The equation both, precisely stats that if each agent plays u_i^*, then each of them will have no incentive to deviate from their chosen action.

- Another way of looking at it, is to consider the **Best-Response Set:**

$$\mathcal{B}_i(u_{-i}) = \{u \in U_i(x) : u \in \arg\max\{U_i(x, u, u_{-i})\}\}, \forall i \in \{1, ..., p\}$$

- This set captures the best-response and agent can make given that the other agents play u_{-i}.

The Nash Equilibrium

- Then an strategy profile $\left(u_1^*, \ldots, u_p^*\right)$ is a Nash Equilibrium if and only if:

$$u_i^* \in \mathcal{B}_i(u_{-i}^*),\ \forall i \in \{1, \ldots, p\}$$

- So a NE occurs when each agent is “best-responding” to every other agent.

- This formulation highlights the fact that the Nash Equilibrium is not (in general) a solution to some optimization problem.

- It is actually a solution to a **fixed-point** equation (given by the best-response sets).

- So one might asks: Does a Nash Equilibrium always exists?

The Matching Pennies

- Consider the following game where either match two pennies or not:

~~1, −1~~	~~−1, 1~~
~~−1, 1~~	~~1, −1~~

- The Matching Pennies is a zero-sum game. There are no NE in this game.

- So a NE may not exist. Then we need to adapt our Equilibrium concept in order to ensure that we can always find it.

The Mixed-Strategy Nash Equilibrium

- We will adapt the equilibrium concept to allow the agent's to **randomize** their actions

- Hence, the action space is now continuous:
  - Given a set of M actions, the agents will choose $u_i^m \in [0,1]$ to be the probability of selecting action m. And we will require that:

$$\sum_{m=1}^{M} u_i^m = 1, \forall i \in \{1, ..., p\}$$

  - So note now that the action $u_i \in [0,1]^M$ is a vector, for each agent.

  - And we will assume that the Agent's maximize the **Expected Utility** over their actions.

The Mixed-Strategy Nash Equilibrium

- A strategy profile $(u_1^*, \dots, u_p^*)$ is a Mixed-Strategy Nash Equilibrium if:

$$u_i^* = \arg \max_{u \in U_i(x), u \in \Delta} \{\mathbb{E}_{(u, u_{-i}^*)}[U_i(x, u, u_{-i}^*)]\}, \forall i \in \{1, \dots, p\}$$

- Where:

$$\Delta = \{u \in [0,1]^M : \sum_{m=1}^{M} u^m = 1\}$$

- Is the probability simplex. And $U_i(x)$ is some constraint set over the probabilities.

- Nevertheless, as before we can characterize the mixed-strategy, as the Best-Response:

$$\mathcal{B}_i(u_{-i}) = \{u \in \Delta : u \in \arg \max_{u \in U_i(x)} \{\mathbb{E}_{(u, u_{-i})}[U_i(x, u, u_{-i})]\}\}, \forall i \in \{1, \dots, p\}$$

The Mixed-Strategy Nash Equilibrium

- A strategy profile $(u_1^*, \dots, u_p^*)$ is a Mixed-Strategy Nash Equilibrium if:

$$u_i^* \in \mathcal{B}_i(u_{-i}^*),\ \forall i \in \{1, \dots, p\}$$

- Now it turns out that every game with finite action space and a finite number players **always** have a Mixed-Strategy Nash Equilibrium.

- The proof lies beyond our scope.
  - It is based on the convergence of fixed-point equations
  - The proof is actually related to Dynamic Programming

The Matching Pennies

- Let's find the Mixed-strategy NE of the Matching Pennies problem:

		
	$1, -1$	$-1, 1$
	$-1, 1$	$1, -1$

- The game is a zero-sum symmetric game.

- Upon inspection we can verify if both players simply "flip" the coins
  - That is they "play" each action with probability 0.5

- That is the Mixed-strategy NE of this game.

The Mixed-Strategy Nash Equilibrium

- For a general game it is actually hard to find the mixed-strategy NE:

$$u_i^* = \arg \max_{u \in U_i(x), u \in \Delta} \{\mathbb{E}_{(u, u^*_{-i})}[U_i(x, u, u^*_{-i})]\}, \, \forall i \in \{1, ..., p\}$$

- However in some cases it is actually a tractable problem.
- One such case is a zero-sum game with two players
- In such a game we write the Normal Form with a single matrix A (called the payoff matrix).
- One agent seeks to maximize, the other seeks to minimize.
    - One example is the previous Matching Pennies game.

Example: The zero-sum game

- For this kind of game we can the Mixed-Strategy NE as:

$$(u_1^*, u_2^*) = \min_{u_1 \in \Delta} \max_{u_2 \in \Delta} \{u_1^\top A u_2\}$$

- Where agent 1 is the “row player”. And agent 2 is the “column player”.
- This can be converted into a linear program, via Duality, which can be solved very efficiently by the computer.
- We can even add polyhedral constraints on top of the probability simplex Δ.
- But what is the meaning of the mixed-strategy NE? What “randomization” means?

The Mixed-Strategy Nash Equilibrium

- This is actually a fairly deep question.

- It seems odd to consider an Equilibrium concept where agent’s are doing something “random”.

- One interpretation of it has to do with **beliefs.**

- We can look at the probabilities we assign to other agents actions as our beliefs that the other agent will select those actions.

- What about our own probabilities? They reflect the fact that we also have beliefs on which action is the best action to take, even though we are not certain which one.

Inverse-Decision Making

- Let's return to our Inverse-Decision Making problem.

- The equilibrium concept we will utilize is the concept of Mixed-Strategy Nash Equilibrium.

- Without loss of generality, we can write that the action $u_i \in \mathbb{R}^{m_i}$ belongs to some polyhedral set:

$$u_i \in \{u \in \mathbb{R}^{m_i} : A_i x + B_i u = b_i, u_i \geq 0\}, \quad i \in \{1, ..., p\}$$

- Where x is the environment state. Note that we allow agents to have a different set of actions. So:
  - The probability simplex Δ can be written as the above.
  - Any additional (linear) constraint can be written as the above as well.
  - We will drop the Expectation of the computations to make notation easier.

Inverse-Decision Making

- In addition, we assume each agent has a type (private information) vector $\theta_i \in \Theta$.

- Hence a strategy profile $\left(u_1^*, \dots, u_p^*\right)$ is a Nash Equilibrium if:

$$
\begin{aligned}
u_i^* = \arg\max_{u \geq 0} & \{U_i(x, u, u_{-i}^*; \theta_i)\},\ \forall i \in \{1, \dots, p\} \\
\text{s.t.:} \quad & A_i x + B_i u = b_i, \quad i \in \{1, \dots, p\}
\end{aligned}
$$

- Where we note that all agents share the same state vector x.
  - After all they all interact in the same environment.

- As before, our goal is to **infer** they types of each agent.

Inverse-Decision Making

- Suppose we collect data points, from the agent's actions:

$$((x^{(1)}, u^{(1)*}), (x^{(2)}, u^{(2)*})), \ldots, (x^{(n)}, u^{(n)*})))$$

- Where $u^{(j)*} = \left(u_1^{(j)*} \ldots, u_p^{(j)*}\right)$ is the Nash Equilibrium of the game when the state is $x^{(j)}$.

- Then we write the following Inverse Optimization Problem:

$$(\hat{\theta}_1, \ldots, \hat{\theta}_p) = \underset{\theta_i \in \Theta, \forall i \in \{1,\ldots,p\}}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$u_i^{(j)*} = \operatorname{argmax}_{\text{s.t.: } A_i x^{(j)} + B_i u = b_i, u \geq 0} \{U_i(x^{(j)}, u, u_{-i}^{(j)*}; \theta_i)\}, \, \forall i \in \{1, \ldots, p\}, \forall j \in \{1, \ldots, n\}$$

Inverse-Decision Making

- This problem is hard to solve in general.

- However, as we did for the single agent case, we can reformulate the problem and, under some special cases, the problem becomes tractable.

- Let's first restate the NE definition:

$$u_i^* = \arg\max_{u \in \geq 0} \{U_i(x, u, u_{-i}^*; \theta_i)\}, \ \forall i \in \{1, ..., p\}$$
$$\text{s.t.:} \quad A_i x + B_i u = b_i, \qquad i \in \{1, ..., p\}$$

- We will the notion of **Variational Inequalities** to characterize the NE profile.

- Let $P_i = \{u \geq 0 \colon A_i x + B_i u = b_i\}$

Inverse-Decision Making

- Then we can write the problem as:

$$u_i^* = \arg\max_{u \in P_i} \{U_i(x, u, u_{-i}^*; \theta_i)\}, \; \forall i \in \{1, ..., p\}$$

- If u_i^* is the optimal solution to the above problem, it must be that if we move from u_i^* in any direction that is feasible, we must reach a lower objective value.

- Formally this means that u_i^* is optimal if and only if:

$$\nabla_u U_i(x, u_i^*, u_{-i}^*; \theta_i)^\top (u - u_i^*) \leq 0, \; \forall u \in P_i$$

- Or equivalently (by multiplying by -1):

$$-\nabla_u U_i(x, u_i^*, u_{-i}^*; \theta_i)(u - u_i^*) \geq 0, \; \forall u \in P_i$$

Inverse-Decision Making

- The inequality:

$$-\nabla_u U_i(x, u_i^*, u_{-i}^*; \theta_i)(u - u_i^*) \geq 0, \; \forall u \in P_i$$

- Has to hold for every agent $i \in \{1, \dots, p\}$. Then if we define:

$$F(u_1, \dots, u_p) = \begin{bmatrix} -\nabla_{u_1} U_1(x, u; \theta_1) \\ \vdots \\ -\nabla_{u_p} U_p(x, u; \theta_p) \end{bmatrix}$$

- Then a strategy profile $(u_1^*, \dots, u_p^*)$ is a Nash Equilibrium if and only if:

$$F(u^*)^\top (u - u^*) \geq 0, \; \forall u \in P$$

- Where $P = P_1 \times \cdots \times P_p$ is the joint feasible region of all agents.

Inverse-Decision Making

- The inequality:

$$F(u^*)^\top (u - u^*) \geq 0, \, \forall u \in P$$

- Is called a **Variational Inequality:** $\mathrm{VI}(F, P)$

- Hence a (mixed-strategy) NE can be seen as the solution of this particular VI.

- This is nice, but not enough: this inequality needs to hold for every feasible joint action vector u. This cannot be implemented by a computer.

- Now we need to use the fact the Utility Functions are concave.

Inverse-Decision Making

- Then $u^* = (u_1^*, \dots, u_p^*)$ solves $\mathrm{VI}(F, P)$ (so is a NE) if and only if $\exists y = (y_1, \dots, y_p)$ such that:

$$B_i^\top y_i \leq -\nabla_{u_i} U_i(x, u_i^*, u_{-i}^*; \theta_i), \quad \forall i \in \{1, \dots, p\}$$

$$-\nabla_{u_i} U_i(x, u_i^*, u_{-i}^*; \theta_i)^\top u_i^* - (b_i - A_i x)^\top u_i^* \leq 0, \quad \forall i \in \{1, \dots, p\}$$

- This conversion is based on Duality.
  - We refer to ”Data-Driven Estimation in Equilibrium using Inverse Optimization. Bertsimas et al, 2014” for the full derivation.

- The inequalities above **can** be implemented in a computer.

Inverse-Decision Making

- Then the Inverse Optimization problem becomes:

$$(\hat{\theta}_1, ..., \hat{\theta}_p) = \underset{\theta_i \in \Theta, \forall i \in \{1,...,p\}, y^{(j)}, \forall j \in \{1,...,n\}}{\operatorname{argmin}} \quad 0$$

$$\text{s.t.:}$$

$$B_i^\top y_i^{(j)} \leq -\nabla_{u_i} U_i(x^{(j)}, u_i^{(j)*}, u_{-i}^{(j)*}; \theta_i), \quad \forall i \in \{1, ..., p\}, \forall j \in \{1, ..., n\}$$

$$-\nabla_{u_i} U_i(x^{(j)}, u_i^{(j)*}, u_{-i}^{(j)*}; \theta_i)^\top u_i^{(j)*} - (b_i - A_i x^{(j)})^\top u_i^{(j)*} \leq 0, \quad \forall i \in \{1, ..., p\}, \forall j \in \{1, ..., n\}$$

- If the Utility function is affine in θ_i then the above problem is a Linear Program and can be solved very efficiently.

- This problem is also called the *Inverse Variational Inequality Problem*.

Approximate Nash Equilibrium

- This formulation can easily incorporate notion of approximate NE.

- We say that a strategy profile $(u_1^*, \ldots, u_p^*)$ is a ϵ-approximate NE if:

$$F(u^*)^\top (u - u^*) \geq -\epsilon, \; \forall u \in P$$

- This is readily incorporated in the Inverse Optimization Problem:

$$(\hat{\theta}_1, \ldots, \hat{\theta}_p) = \underset{\theta_i \in \Theta, \forall i \in \{1,\ldots,p\}, y^{(j)}, \forall j \in \{1,\ldots,n\}}{\operatorname{argmin}} \quad \sum_{j=1}^{n} ||\epsilon_j||_2^2$$

$$\text{s.t.:}$$

$$B_i^\top y_i^{(j)} \leq -\nabla_{u_i} U_i(x^{(j)}, u_i^{(j)*}, u_{-i}^{(j)*}; \theta_i), \quad \forall i \in \{1, \ldots, p\}, \forall j \in \{1, \ldots, n\}$$

$$-\nabla_{u_i} U_i(x^{(j)}, u_i^{(j)*}, u_{-i}^{(j)*}; \theta_i)^\top u_i^{(j)*} - (b_i - A_i x^{(j)})^\top u_i^{(j)*} \leq \epsilon_j, \quad \forall i \in \{1, \ldots, p\}, \\ \forall j \in \{1, \ldots, n\}$$

Example: Airline Ticketing

- Consider two Airlines that are competing over a determined route.

- Each Airline can set prices for that route. The demand for flights depends on both prices.

- Suppose the demand for each Airline obeys the following linear equation:

$$\bar{D}_i(x, p_1, p_2; \theta_i) = \theta_{i,0} + \sum_{j=1}^{2} p_j \theta_{i,j} + \theta_{i,3} x$$

- Where $i \in \{1,2\}$. p_i is the ticket price set by company i. x represents the "state of the economy". (x can be some economic indicator, like S&P index, or something else).

- And $\theta_i = \left(\theta_{i,0}, \theta_{i,1}, \theta_{i,2}, \theta_{i,3}\right)$ are the company private information.

Example: Airline Ticketing

- The Utility Function of each company is naturally their profit:

$$U_i(x, p_1, p_2; \theta_i) = p_i \bar{D}_i(x, p_1, p_2; \theta_i)$$

- And both airlines seeks to maximize their profit. Due to regulations there is an upper bound of $\bar{p}$ on the prices. So $0 \leq p_i \leq \bar{p}$ for both airlines.

- Let's state the marginal utility of each company:

$$m_i(x, p_1, p_2; \theta_i) = \frac{\partial}{\partial p_i} U_i(x, p_1, p_2; \theta_i) =$$

$$m_i(x, p_1, p_2; \theta_i) = p_i \theta_{i,i} + \theta_{i,0} + \sum_{j=1}^{2} p_j \theta_{i,j} + \theta_{i,3} x$$

Example: Airline Ticketing

- Suppose we are an regulatory agent which observes the data points:

$$((x^{(1)}, p_1^{(1)}, p_2^{(1)}), (x^{(2)}, p_1^{(2)}, p_2^{(2)}), \ldots, (x^{(n)}, p_1^{(n)}, p_2^{(n)}))$$

- Our goal is to ascertain that the companies are competing fairly.

- In order words: we want to ensure that the prices are the outcome of fair competition and not the outcome of some collusion (which is illegal!).

- We can formulate a hypothesis testing based on Inverse Optimization in order to investigate that.

Example: Airline Ticketing

- First, we state the Inverse Optimization Problem:

$$\min_{\hat{\epsilon}, y, \theta_1, \theta_2} \quad \sum_{j=1}^{N} \hat{\epsilon}_j \tag{1}$$

$$\text{s.t. } y_i^{(j)} \geq m_i(x^{(j)}, p_1^{(j)}, p_2^{(j)}, \theta_1), \text{for } i \in \{1, 2\}, j \in \{1, ..., N\} \tag{2}$$

$$\bar{p} \sum_{i=1}^{2} (y_i^{(j)}) - \sum_{i=1}^{2} p_i^{(j)} m_i(x^{(j)}, p_1^{(j)}, p_2^{(j)}, \theta_i) = \hat{\epsilon}_j, j \in \{1, ..., N\}$$

$$m_i(0, 1, 1, \theta_i) = m_i(0, 1, 1, \bar{\theta}_i), \text{for } i \in \{1, 2\} \tag{3}$$

$$y_i^{(j)} = 0, \text{for } i \in \{1, 2\}, j \in \{1, ..., N\} \text{ s.t. } p_i^j < \bar{p} \tag{4}$$

$$\theta_{i,i} \leq 0, \text{for } i \in \{1, 2\} \tag{5}$$

$$\hat{\epsilon}^j \geq 0, \text{for } j \in \{1, ..., N\} \tag{6}$$

$$y^{(j)} = (y_1^{(j)}, y_2^{(j)}) \geq 0, \text{for } \forall j \in \{1, ..., N\} \tag{7}$$

Minimizing the residuals

Variational Inequalities

Normalization

Non-negativities

Example: Airline Ticketing

- There is a caveat in this model.

- We are assuming that if the Airlines are competing they do it in rational way, by maximizing their utility and try to best-respond to one another.

- This may not in fact be the case.

- But for the sake of modelling, we assume that the companies(if competing) behave according to their Nash Equilibrium strategies.

- So under the fair competition scenario, we expect the residuals $(\hat{\epsilon}_1, \dots, \hat{\epsilon}_n)$ to be statistically identical to zero.

Example: Airline Ticketing

- Alternatively, we can consider the case that the companies can deviate, albeit a little, from the equilibrium strategy.
  - This may illustrate the effect of unmodelled uncertainties in the demand function.

- For instance, we can assume that the companies deviate according to some exponential decay law, where large deviations are exponentially less likely to occur than small deviations.

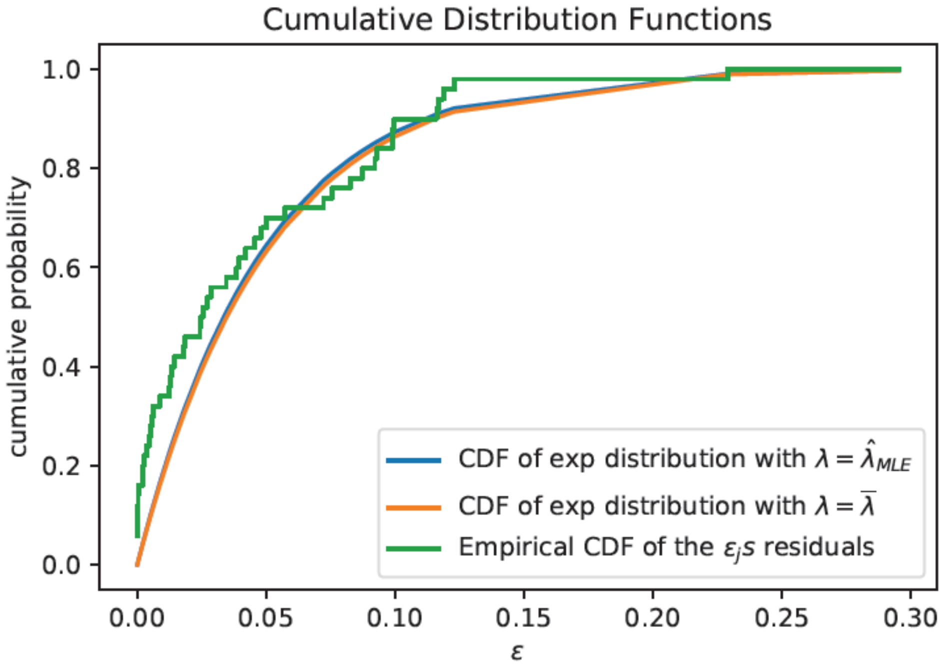
- This means that residuals should obey some exponential distribution with some rate (say rate $\bar{\lambda}$.

- Then our Hypothesis Testing boils down to ascertain whether:

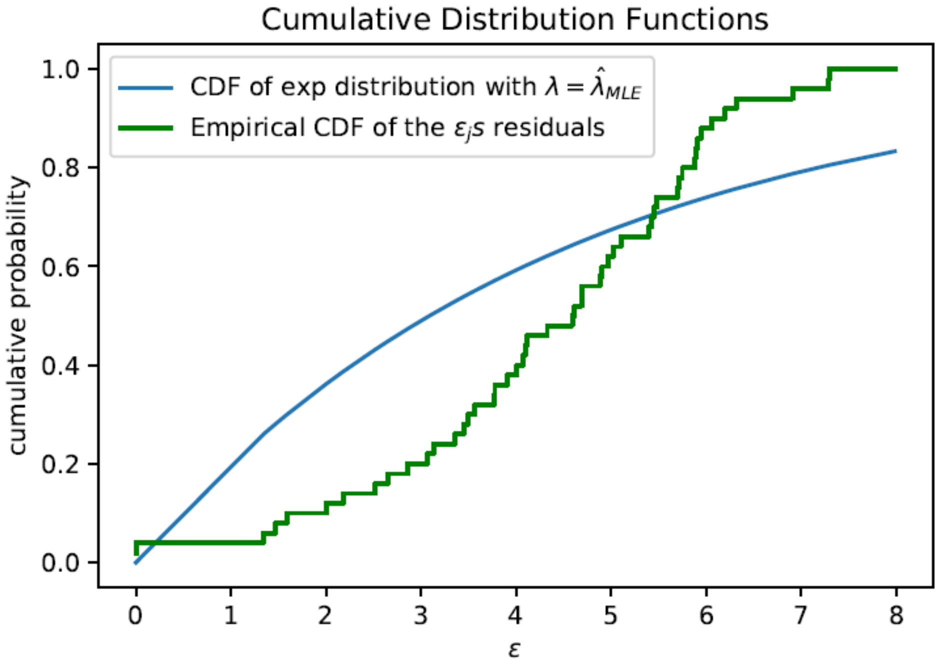
$$(\hat{\epsilon}_1, \dots, \hat{\epsilon}_n) \sim \exp(\lambda)$$

Example: Airline Ticketing

- Hence the problem boils down in analyzing the distribution of the residuals:



(fair competition)



(collusion)

- The Hypothesis Testing and residual estimation is consistent, provided we can solve the Inverse Optimization Problem.

Inverse Decision Making: Overview

- With the past two lectures we just “scrapped” the top of the iceberg for Inverse Decision Making problems.

- There are many other problems that can be cast as an Inverse Optimization problem.

- The theoretical questions and practical challenges are all there still. This is an area of active research.

- In particular, applications regarding collusion and cooperation of AI can ba analyzed in the lens of Inverse Decision Making problem.

- The underlying rational and utility assumption may also be questioned:
  - What if agents are not rational?
  - Do people really maximize “Utility” in their lives? What about robots?