

# 名词解释

## Lecture 1: 数据要素市场概述

- **生产要素 (Factors of Production)**
  - **解释:** 指进行社会生产经营活动时所需要的各种投入品。传统经济学认为主要有四种：**土地 (Land)**，指一切自然资源，其报酬为租金；**劳动力 (Labour)**，指人的体力和脑力劳动，其报酬为工资；**资本 (Capital)**，指用于生产其他物品的设备、建筑物等，其报酬为利息或股利；**企业家才能 (Entrepreneurship)**，指组织和管理其他要素进行生产的能力。在数字经济时代，**数据 (Data)** 已被公认为一种新的、关键的生产要素。
- **循环流量图 (Circular Flow Diagram)**
  - **解释:** 一个简化的、直观的市场经济模型，用于展示经济中货币、物品和服务如何在家庭与企业这两个决策者之间，通过物品与服务市场和生产要素市场进行循环流动的。它揭示了经济的总体组织结构和参与者之间的相互交易关系。
  - **核心组成:**
    - **决策者:** 家庭和企业。
    - **市场:** 物品与服务市场、生产要素市场。
    - **流向:**
      - **实物流 (投入与产出):** 生产要素从家庭流向企业，物品与服务从企业流向家庭。
      - **货币流:** 货币以支出形式从家庭流向企业，以收入（工资、租金、利润）形式从企业流向家庭。
- **绝对优势 (Absolute Advantage)**
  - **解释:** 比较不同生产者生产效率的概念。如果一个生产者使用比另一个生产者更少的投入（如时间、劳动力）就能生产出同样数量的某种物品，那么这个生产者在生产该物品上就具有绝对优势。
  - **定义:** 生产者生产一种物品所需投入较少，就可以说该生产者在生产这种物品上有绝对优势。
- **机会成本 (Opportunity Cost)**
  - **解释:** 指为了得到某种东西而必须放弃的价值最高的替代品。它是决策制定中的一个核心概念，衡量了选择一个选项所放弃的次优选项的价值。例如，农民Frank生产1单位牛肉的机会成本是其因此放弃生产的4单位土豆。
- **比较优势 (Comparative Advantage)**
  - **解释:** 比较不同生产者机会成本的概念。如果一个生产者以低于另一个生产者的机会成本生产一种物品，那么他在生产该物品上就具有比较优势。贸易的基础是比较优势而非绝对优势，即使一方在所有产品上都具有绝对优势，双方仍然可以通过专门生产自己具有比较优势的产品并进行贸易而获益。
- **产权 (Property Rights)**
  - **解释:** 指个人或群体拥有、使用、处置和从资产中获取收益的合法权利。在数据市场中，明确的数据产权是数据流通和交易的基础，产权不明可能导致**公地悲剧** (Tragedy of the Commons)，即公共资源因缺乏明确所有者而被过度使用和耗尽。

## Lecture 2: 数据合规与安全

- **安全多方计算 (Secure Multi-Party Computation, MPC)**
  - **解释:** 一种密码学协议，允许多个互不信任的参与方在不泄露各自私有输入数据的前提下，共同完成对某个函数的计算。其目标是在保护各方数据隐私的同时，得到正确的计算结果。经典例子是姚氏百万富翁问题。

- **联邦学习 (Federated Learning)**

- **解释:** 一种分布式的机器学习技术, 允许在多个持有本地数据的设备 (如手机、医院) 上联合训练一个共享模型, 而无需将原始数据集中起来。参与方只交换模型参数 (如梯度), 原始数据不出本地, 从而在保护数据隐私和安全的同时实现协同建模。

- **数据脱敏 (Data Masking)**

- **解释:** 根据特定的规则和技术, 对原始数据中的敏感信息 (如姓名、电话、身份证号) 进行变形、替换或遮蔽处理, 使其不再能直接识别到具体个人, 同时又保留数据的可用性和格式一致性。这是在非生产环境 (如测试、开发) 中使用数据的常用隐私保护方法。

- **差分隐私 (Differential Privacy)**

- **解释:** 一种提供可量化、严格隐私保护保证的数学模型。其核心思想是在查询结果中加入经过校准的随机噪声, 使得从数据集中添加或删除任意一条记录对查询结果的影响微乎其微。这样, 任何攻击者都无法通过分析查询输出来推断出关于数据集中任何特定个体的信息。
- **数学定义:** 对于任意两个仅相差一条记录的相邻数据集  $D_1$  和  $D_2$ , 一个随机化算法  $A$  满足  $\epsilon$ -差分隐私, 如果对于其输出域中的任何子集  $O$ , 都满足:

$$\Pr[A(D_1) \in O] \leq e^\epsilon \cdot \Pr[A(D_2) \in O]$$

其中, 隐私预算  $\epsilon$  是一个小的正数, 控制着隐私保护的强度。

- **零知识证明 (Zero-Knowledge Proof, ZKP)**

- **解释:** 一种密码学协议, 其中一方 (证明者 Prover) 可以向另一方 (验证者 Verifier) 证明自己知道某个秘密或某个论断为真, 而无需透露任何关于该秘密或论断本身的任何信息。
- **三大特性:** **完整性** (如果论断为真, 总能证明成功)、**可靠性** (如果论断为假, 无法欺骗验证者) 和 **零知识性** (验证者除了知道论断为真外, 学不到任何额外信息)。

## Lecture 3: 非合作博弈论基础 (一)

- **效用 (Utility)**

- **解释:** 在经济学中, 效用是衡量消费者从消费一组商品或服务中获得的满意度或满足感的度量。**效用函数**  $u(x)$  是一个将消费组合  $x$  映射到一个实数的函数, 用以表示消费者对该消费组合的偏好程度。理性人假设认为, 经济参与者会选择使自己效用最大化的行动。

- **边际效用递减规律 (Law of Diminishing Marginal Utility)**

- **解释:** 指在一定时间内, 随着消费者对某种商品消费量的不断增加, 他从每增加一单位消费中所获得的额外满足感 (即边际效用) 是递减的。

- **市场失灵 (Market Failure)**

- **解释:** 指自由市场在自发调节下, 无法实现资源最优配置 (帕累托最优) 的情况。导致市场失灵的主要原因包括**外部性**、**垄断**、**公共物品**和**信息不对称**。

- **外部性 (Externality)**

- **解释:** 指一个经济主体 (个人或企业) 的活动对其他人的福利产生了影响, 但这种影响没有通过市场价格得到反映。**负外部性**是产生有害影响 (如污染), **正外部性**是产生有益影响 (如植树)。

- **信息不对称 (Asymmetric Information)**

- **解释:** 指在市场交易中, 一方参与者拥有比另一方更多或更好的信息。这会导致**逆向选择** (劣质产品驱逐优质产品, 如二手车“柠檬市场”) 和**道德风险** (一方在交易后采取对自己有利但损害对方利益的行动) 等问题, 从而造成市场失灵。

- **帕累托最优 (Pareto Optimality)**

- **解释:** 也称帕累托效率，是资源分配的一种理想状态。在这种状态下，任何形式的资源重新分配，都不可能在不使任何人的境况变坏的前提下，使至少一个人的境况变得更好。福利经济学第一定理指出，在完全竞争、无外部性等理想条件下，市场均衡是帕累托最优的。

## Lecture 4: 非合作博弈论基础（二）

- **纳什均衡 (Nash Equilibrium)**

- **解释:** 非合作博弈论的核心解概念。它描述的是这样一种策略组合状态：在此状态下，假设其他参与者的策略保持不变，没有任何一个参与者能通过单方面改变自己的策略来获得更高的收益。因此，纳什均衡是一个稳定的、没有参与者愿意单方面偏离的策略组合。
- **数学定义:** 一个策略组合  $s^* = (s_1^*, \dots, s_n^*)$  是一个纳什均衡，如果对于每个参与者  $i$  和其任意策略  $s_i$ ，都满足：

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$$

- **稳定集 (Stable Set)**

- **解释:** 在本课程的语境中，"稳定集" 指的是博弈中所有**纳什均衡**的集合。纳什均衡的本质就是一种稳定性，因为一旦达到，所有理性的参与者都没有单方面偏离的动机。

- **子博弈完美均衡 (Subgame Perfect Nash Equilibrium, SPNE)**

- **解释:** 对纳什均衡概念在动态博弈中的一种精炼。一个策略组合是子博弈完美均衡，当且仅当它在整个博弈的每一个子博弈中都构成一个纳什均衡。SPNE通过要求策略在博弈的每个阶段都是最优的，从而排除了那些包含“不可置信威胁”的纳什均衡。

- **逆向归纳法 (Backward Induction)**

- **解释:** 一种求解有限完美信息动态博弈中子博弈完美均衡的方法。它从博弈的最后一个决策点开始，向前倒推，确定每个决策点上的最优行动，直到博弈的起点。

- **完美信息博弈 (Game with Perfect Information)**

- **解释:** 一种动态博弈，其中轮到任何参与者行动时，他都完全了解博弈到目前为止的全部历史（即所有之前参与者的所有行动）。例如，象棋和围棋是完美信息博弈。

- **不完全信息博弈 (Game with Incomplete Information)**

- **解释:** 一种博弈，其中至少有一个参与者不完全了解其他参与者的某些特征，如他们的支付函数或“类型” (Type)。例如，在拍卖中，你不知道其他竞拍者对物品的真实估值。

- **贝叶斯均衡 (Bayesian Nash Equilibrium)**

- **解释:** 不完全信息博弈的均衡概念。它是一个策略组合，其中每个类型的每个参与者都选择一个最大化其期望效用的策略，这个期望是基于他对其他参与者类型的后验信念（通过贝叶斯法则从先验信念更新而来）计算的。

## Lecture 5: 合作博弈与数据估值

- **合作博弈 (Cooperative Game)**

- **解释:** 一类博弈模型，关注的是参与者之间通过形成联盟 (Coalition) 可以共同创造多少价值，以及如何在联盟成员之间公平地分配这些价值，而不关心达成协议的具体过程。

- **特征函数 (Characteristic Function)**

- **解释:** 在合作博弈中，特征函数  $v(S)$  定义了任意一个联盟  $S$ （参与者的一个子集）通过合作所能确保获得的总收益（价值）。

- **核 (Core)**

- **解释:** 合作博弈的一个解概念。它指的是所有满足**有效率**（所有参与者分配到的总收益等于大联盟的总价值）和**联盟理性**（任何子联盟获得的收益总和不小于该子联盟单独合作能创造的价值）的分配方案的集合。核中的分配方案是“稳定的”，因为没有子联盟有动机脱离大联盟自立门户。

- **Shapley 值 (Shapley Value)**

- **解释:** 合作博弈中一个独特的、公平的收益分配方案。它根据每个参与者对所有可能形成的联盟的**边际贡献**的期望值来分配总收益。Shapley值满足**有效率**、**对称性**、**零参与者**和**可加性**四个公理。
- **计算公式:** 参与者  $i$  的Shapley值  $SV_i$  为：

$$SV_i(N, v) = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!(n - |S| - 1)!}{n!} [v(S \cup \{i\}) - v(S)]$$

## Lecture 6: 多臂老虎机算法基础与应用

- **多臂老虎机问题 (Multi-Armed Bandit, MAB)**

- **解释:** 一类经典的在线决策问题。一个玩家面对  $K$  个选项（“老虎机臂”），每次选择一个臂拉动，会获得一个随机的奖励。玩家的目标是在  $T$  轮试验中，通过平衡**探索**（尝试不同的臂以了解其奖励分布）和**利用**（选择当前看起来最好的臂），来最大化累积总奖励。

- **遗憾 (Regret)**

- **解释:** 在MAB问题中，遗憾是衡量一个算法性能的核心指标。它被定义为“始终选择最优臂所能获得的期望总奖励”与“算法实际获得的期望总奖励”之间的差值。一个好算法的目标是最小化遗憾。

- **跟风算法 (Follow-The-Leader, FTL)**

- **解释:** 一种简单的在线学习算法。在对抗性MAB中，它在每一轮都选择到目前为止累积损失（或代价）最小的那个臂。

- **上置信界算法 (Upper Confidence Bound, UCB)**

- **解释:** MAB问题的一种经典算法。它为每个臂的平均奖励估计值构建一个置信区间的上界，然后在每一轮选择上置信界最大的那个臂。这个上界由两部分组成：当前臂的平均奖励（利用）和一个与选择次数成反比的项（探索），从而巧妙地平衡了探索与利用。

## Lecture 7-8: 拍卖与最优机制

- **机制设计 (Mechanism Design)**

- **解释:** 博弈论的一个分支，也被称为“逆向博弈论”。它从期望达成的目标（如社会福利最大化、卖家利润最大化）出发，来设计一套规则（即一个“机制”），使得理性的、自利的参与者在该规则下行动时，其均衡结果恰好能实现设计者的目标。

- **显示原理 (Revelation Principle)**

- **解释:** 机制设计中的一个基本原理。它指出，对于任何一个（间接）机制及其均衡结果，都存在一个等价的**直接显示机制**，在该机制中，参与者诚实地报告自己的私人信息（类型）本身就是一个均衡，并且能产生与原机制相同的分配结果。这大大简化了机制设计问题。

- **激励相容 (Incentive Compatible)**

- **解释:** 机制的一个性质，指在该机制下，每个参与者的最优策略都是诚实地报告自己的私人信息（类型）。如果诚实是占优策略，则称之为**占优策略激励相容 (DSIC)**。如果诚实是在贝叶斯均衡中的策略，则称之为**贝叶斯激励相容 (BIC)**。

- **VCG 机制 (Vickrey-Clarke-Groves Mechanism)**

- **解释:** 一种通用的、能实现社会福利最大化的DSIC机制。其分配规则是选择使所有参与者报告的估值之和最大的分配方案。其支付规则是，每个获胜者支付的费用等于其参与给其他所有参与者造成的“外部性”成本。

- **虚拟估值 (Virtual Valuation)**

- **解释:** 在设计最优拍卖（卖家利润最大化）时，由买家的真实估值  $t_i$  和其估值分布  $F_i(t_i)$  推导出的一个概念。在正则化条件下，最优机制等同于一个将物品分配给虚拟估值最高的买家的机制。

- **公式:**

$$c_i(t_i) = t_i - \frac{1 - F_i(t_i)}{f_i(t_i)}$$

其中  $f_i(t_i)$  是概率密度函数。

## Lecture 9-10: 数据定价与贝叶斯劝说

- **数据的版本化 (Versioning of Data)**

- **解释:** 一种定价策略，指卖家不直接出售单一、完整的“最好”版本的数据产品，而是创建并出售多个不同版本（如不同精度、不同覆盖范围、不同时效性）的产品。这可以看作是一种价格歧视，旨在吸引不同支付意愿和需求的买家，从而最大化总利润。

- **无套利原则 (Arbitrage Freeness)**

- **解释:** 市场定价的一个基本原则，指市场上不存在无风险的获利机会。在数据定价中，这意味着买家不能通过购买低版本数据产品的组合来“合成”一个高版本产品，且总成本低于直接购买高版本产品的价格。

- **贝叶斯劝说 (Bayesian Persuasion)**

- **解释:** 也称信息设计，是一种策略性信息传递模型。其中，一个有信息优势的“发送者”通过设计并承诺一个信号机制，来影响一个理性“接收者”的后验信念，从而引导接收者采取对发送者最有利的行动。发送者不能说谎，但可以选择披露信息的“颗粒度”。