第六讲：不完全信息动态博弈

兴军亮

Junliang.Xing@ia.ac.cn

2020年10月29日

中国科学院大学
University of Chinese of Academy of Sciences

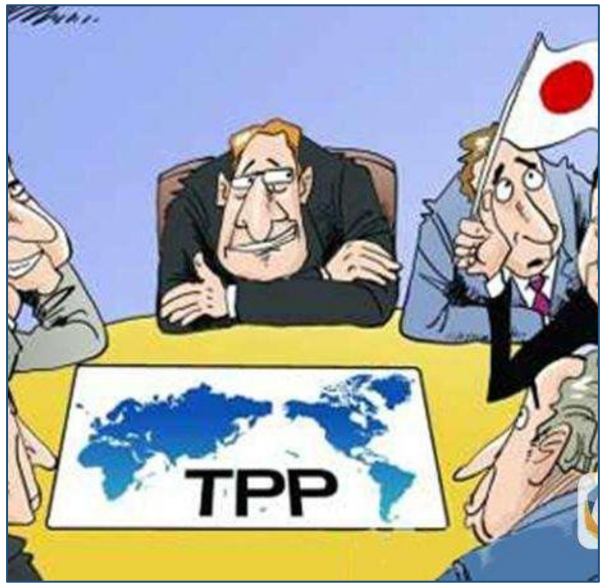
CASIA
自动化研究所
Institute of Automation

本讲提纲

- 精炼贝叶斯纳什均衡
- 信号传递博弈及示例
- 其他相关的均衡概念
- 不完全信息重复博弈

不完全信息动态博弈

- 实际中的博弈问题往往不但存在不完全信息，而且会持续多个阶段。博弈参与者会在观察到先行动者所选择的动作来推断其类型或者修正对其类型的先验信息（概率分布），然后选择自己的最优行动。

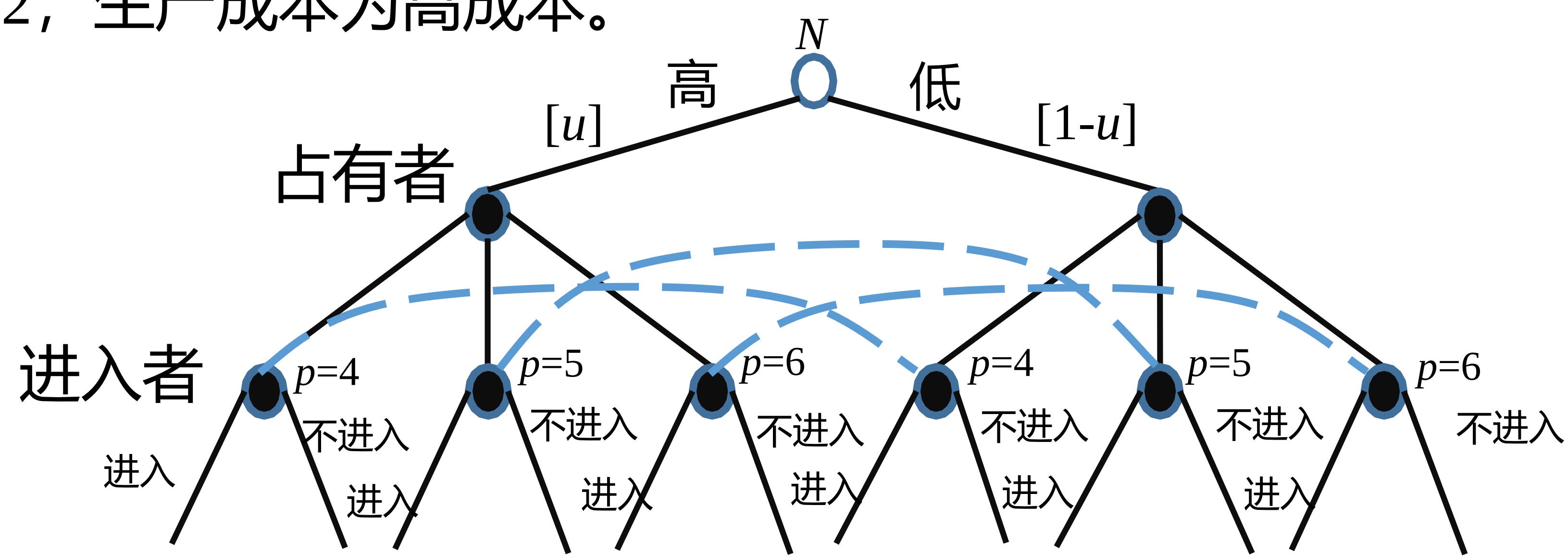


定义和表示

- 不完全信息动态博弈：博弈过程中参与人预先不知道关于博弈的所有信息（最终体现在支付/偏好等），参与人的行动有先后顺序，且后行动者能够观察到先行动者行动的博弈。
- 不完全信息动态博弈的展开式表示：
  - **参与人集合：**$i \in \mathbf{N}, \mathbf{N} = \{1,2, \ldots, n\}$，**用**$N$**表示自然**
  - **参与人的行动顺序：谁在什么时候行动**
  - **参与人的行动空间：每次行动时，有什么选择**
  - **信息集合：每次行动结束，参与人知道些什么**
  - **效益函数：每次行动时，参与人知道什么**
  - **外生事件（即自然选择）的概率分布**

举例：不完全信息下的市场进入阻挠博弈

- 博弈描述：
  - 两阶段博弈：第一阶段占有者垄断市场，进入者考虑是否进入；第二阶段如果进入者进入，两者进行库诺特博弈，否则占有者继续垄断。占有者有高和低两种成本，有三种价格选择p=4, 5, 6，高低成本对应利润分别为2, 6, 7 和6, 9, 8。进入者进入成本为2，生产成本为高成本。



第一阶段	(2, 0)	(2, 0)	(6, 0)	(6, 0)	(7, 0)	(7, 0)	(6, 0)	(6, 0)	(9, 0)	(9, 0)	(8, 0)	(8, 0)
第二阶段	(3, 1)	(7, 0)	(3, 1)	(7, 0)	(3, 1)	(7, 0)	(5, -1)	(9, 0)	(5, -1)	(9, 0)	(5, -1)	(9, 0)

举例：不完全信息下的市场进入阻挠博弈

- 沿用静态博弈的思路：
  - 对于占有者而言，第二阶段的收益都是相同的。从第一阶段的结果来看，如果占有者是高成本的，它的最优选择是$p = 6$；如果占有者是低成本的，它的最优选择是$p = 5$。
  - 对于进入者而言，如果占有者是高成本的，它的最优选择是进入，如果占有者是低成本的，它的最优选择是不进入。而只有进入者认为占有者是高成本的概率u大于1/2时，才会选择进入。
  - 与静态博弈不同的是，在观察到占有者第一阶段的价格选择后，进入者可以修正对占有者的先验概率u，因为占有者的价格选择可能包含着有关其成本函数的信息。
  - 因此，尽管$p= 6$对高成本占有者是有利可图的，但这将导致进入者推断占有者是高成本的，进入者将选择进入。因此高成本的占有者在第一阶段就不会选择$p = 6$。
- 因此，采用静态分析的方法会得到不合理的均衡。

不完全信息动态博弈

- 基本思路
  - 参与人的类型，只有参与人自己知道。
  - 后行动者可以通过观察先行动者所选择的行动来推断其类型或修正对其类型的先验信念（概率分布），选择自己的最优行动。
  - 先行动者预测到自己的行为将被后行动者所利用，就会设法选择传递对自己最有利的信息。
  - 参与人双方都只知道对方行动的类型以及采取这样行动的概率然后双方不断试探，根据对方的反应调整自己的策略，最终取得最优决策。总之，**不完全信息动态博弈不仅是参与人选择行动的过程，而且是参与人不断修正信念的过程。**
- 直观例子
  - 黔驴技穷
  - 信号传递

不完全信息动态博弈

- 不完全信息动态博弈的执行顺序

1. 自然赋予博弈各方类型向量$\Theta = \{\theta_1, \dots, \theta_n\}$
2. 参与人得知自己的类型，但不知其他人的类型
3. 第一个行动者根据类型行动，传递出某种信息
4. 下一个行动者根据上个行动者的行动更新对其类型的先验推断，得到后验推断，然后做出行动
5. 下来的行动者不断重复上述过程直到博弈终止

贝叶斯法则（Bayesian's Rule）

- 贝叶斯法则的基本思想
  - 当面临不确定性时，在任何一个时间点上，人们对某件事情发生的可能性有一个初步判断。然后人们会根据新到的**信息**来修正这一判断。修正之前的初步判断成为“**先验概率**”，修正之后的判断称为“**后验概率**”。
  - 贝叶斯法则是根据新信息从先验概率得到后验概率的基本方法。
  - 贝叶斯法则并不是一个技术性法则，而是人们修正信念的唯一合理方法（**最优估计**）。

Likelihood：似然函数

How probable is the evidence given that our hypothesis is true?

Prior：先验概率

How probable was our hypothesis before observing the evidence?

Posterior：后验概率

$$P(H|e) = \frac{P(e|H)P(H)}{P(e)}$$

How probable is our hypothesis given the observed evidence? (Not directly computable)

Marginal：边界分布

How probable is the new evidence under all possible hypotheses?

贝叶斯法则（Bayesian's Rule）

- 贝叶斯法则举例分析：好人坏人的例子
  - 我们把所有的人分为好人(GP)和坏人(BP)，所有的事分为好事(GT)和坏事(BT)，那么一个人干好事的概率p(GT)等于他是好人的概率p(GP)乘以好人做好事的概率p(GT|GP)，加上他是坏人的概率p(BP)乘以坏人做好事的概率p(GT|BP)

$$p(GT) = p(GT|GP)p(GP) + p(GT|BP)p(BP)$$

  - 假设我们观测到了一个人干了一件好事，那么这个人是好人的后验概率是：

$$p(\mathrm{GP}|\mathrm{GT}) = \frac{p(\mathrm{GT}|\mathrm{GP})p(\mathrm{GP})}{p(\mathrm{GT})}$$

贝叶斯法则（Bayesian's Rule）

- 贝叶斯法则举例分析：好人坏人的例子
  - 假设我们认为一个人是好人的先验概率是1/2，那么观察到他干了一件好事之后，我们如何修正他是好人的先验概率依赖于我们认为这件好事好到了什么程度。
    - 第一种情况：这是一件很好的好事，好人一定干，坏人不会干，那么：

$$p(GP|GT) = \frac{1\times 1/2}{1\times 1/2+0\times 1/2}=1$$

    - 第二种情况：这是一件一般的好事，好人会干，坏人也会干，那么：

$$p(GP|GT) = \frac{1\times 1/2}{1\times 1/2+1\times 1/2}=\frac{1}{2}$$

    - 第三种情况：这是一件不错的好事，好人肯定会干，坏人有可能干，也可能不干，那么：

$$p(GP|GT) = \frac{1\times 1/2}{1\times 1/2+1/2\times 1/2}=\frac{3}{4}$$

贝叶斯法则（Bayesian's Rule）

- 贝叶斯法则的推导过程（以不完全信息博弈为例）
  - $\{\theta^k\}_{k=1}^{K}$为参与人可能的类型集合，$\{a^h\}_{h=1}^{H}$为可能的行动集合
  - 则参与人 i 属于θ^k的先验概率$p(\theta^k) \geq 0, \sum_{k=1}^{K} p(\theta^k) = 1$
  - 假设i 属于θ^k， 选择a^h的条件概率为$p\left(\theta^k \middle| a^h\right)$，则边际概率为$p\left(a^h\right) = \sum_{k=1}^{K} p\left(a^h \middle| \theta^k\right) p(\theta^k)$
  - 贝叶斯法则推导过程：假如观测到了参与人i选择了行动a^h，那么参与人 i属于类型θ^k的后验概率为：

$$p\left(\theta^k \middle| a^h\right) = \frac{p(\theta^k, a^h)}{P(a^h)} = \frac{p\left(a^h \middle| \theta^k\right) p(\theta^k)}{\sum_{k=1}^{K} p(a^h | \theta^k) p(\theta^k)}$$

  - 贝叶斯法则要求$P(a^h) > 0$，否则后验概率没有定义。如果$P(a^h) = 0$，则我们允许$p\left(a^h \middle| \theta^k\right)$取[0,1]之间任何值
  - 在动态博弈中，$P(a^h) = 0$对应非均衡路径上的信息集

精炼贝叶斯均衡

- 精炼贝叶斯的相关概念
  - 假定博弈有n个参与人，参与人i的类型为θ_i，$p(\theta_{-i}|\theta_i)$是参与人i认为其他$n-1$个参与人的类型$\theta_{-i}=(\theta_1,\dots,\theta_{i-1},\theta_{i+1},\dots,\theta_n)$的先验概率。
  - 令$S_i=\{s_i(\theta_i)\}$是参与人的策略空间，s_i是一个特定的策略。
  - 令$a_{-i}^h=\left(a_1^h,\dots,a_{i-1}^h,a_{i+1}^h,\dots,a_n^h\right)$是在第$h$个信息集上参与人$i$观测到的其他个$n-1$参与人的行动组合，这个行动组合是策略组合$s_{-i}=(s_1,\dots,s_{i-1},s_{i-2},\dots,s_n)$的一部分。
  - 令$\tilde{p}_i\left(\theta_{-i}\middle|a_{-i}^h\right)$是在观测到$a_{-i}^h$的情况下参与人 i认为其他$n-1$个参与人属于类型 θ_{-i}的后验概率，$\tilde{p}_i$是所有后验概率的集合即 $\tilde{p}_i$ 包含了参与人 i在每一个信息集上的后验概率。
  - 令$u_i(s_i,s_{-i},\theta_i)$是参与人 i的效用函数。

精炼贝叶斯均衡

- 精炼贝叶斯均衡定义：

精炼贝叶斯均衡

精炼贝叶斯均衡是一个策略组合$s^*(\theta)=\left(s_1^*(\theta_1),\dots,s_n^*(\theta_n)\right)$和一个后验概率组合 $\tilde{p}=(\tilde{p}_1,\dots,\tilde{p}_n)$，满足：

（P）对于所有的参与人i，在每一个信息集合h，

$$s_i^*(s_{-i},\theta_i)\in\max_{s_i}\sum_{\theta_{-i}}\tilde{p}\left(\theta_{-i}\middle|a^h_{-i}\right)u_i(s_i,s_{-i},\theta_i)$$

（B） $\tilde{P}\left(\theta_{-\mathrm{i}}\middle|a^h_{-i}\right)$是使用贝叶斯法则从先验概率$p_i\left(\theta_{-i}\middle|\theta_i\right)$、观测到的$a^h_{-i}$和最优策略$s^*_{-i}(*)$得到的（在可能的情况下）。

精炼贝叶斯均衡

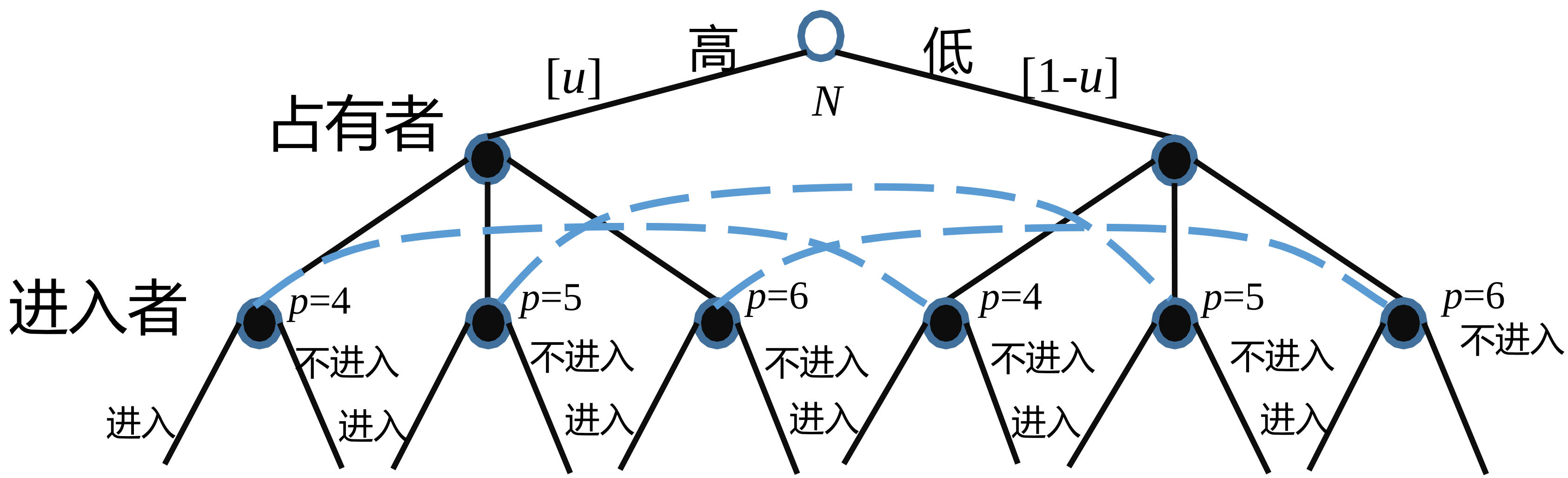
- 关于贝叶斯纳什均衡的一些解释
  - （P）是精炼条件（perfectness condition），给定其他参与人的策略s_{-i}和参与人的后验概率$\tilde{p}_i\left(\theta_{-i}\middle|a^h_{-i}\right)$，每个参与人$i$的策略在所有从信息集 h开始的后续博弈上都是最优的，即所有参与人都是序贯理性的（sequentially rational）。
  - （B）对应的是贝叶斯法则的运用。如果参与人是多次行动的，修正概率涉及贝叶斯法则的重复运用。其中策略本身是不可观测的，因此参与人i 只能根据观测到的行动组合a_{-i}修正概率，但它假定所观测到的行动是最优策略s^*_{-i}规定的行动。
  - “在可能的情况下”：如果a_{-i}不是均衡策略下的行动，即a_{-i}是一个零概率事件。根据贝叶斯法则，任何$\tilde{p}_i\left(\theta_{-i}\middle|a^h_{-i}\right)$都是允许的，只要它与均衡策略相容。

精炼贝叶斯均衡

- 贝叶斯纳什均衡定义的关键
  - 精炼贝叶斯均衡是均衡策略和均衡信念的结合：给定信念$\tilde{p} = (\tilde{p}_1, \dots, \tilde{p}_n)$，策略$s^* = (s_1^*, \dots, s_n^*)$是最优的；给定策略$s^* = (s_1^*, \dots, s_n^*)$，信念$\tilde{p} = (\tilde{p}_1, \dots, \tilde{p}_n)$是使用贝叶斯法则从均衡策略和所观测到的行动得到的。
  - 因此，精炼贝叶斯均衡是一个不动点，后验概率依赖于策略，策略依赖于后验概率，因此完全信息博弈中使用的逆向归纳法（backward induction）在不完全信息中不再适用（如果我们不知道先行者如何选择，我们就不可能知道后行动者应该如何选择），必须使用前向法（forward manner）进行贝叶斯修正。

示例回顾：不完全信息市场进入阻挠博弈

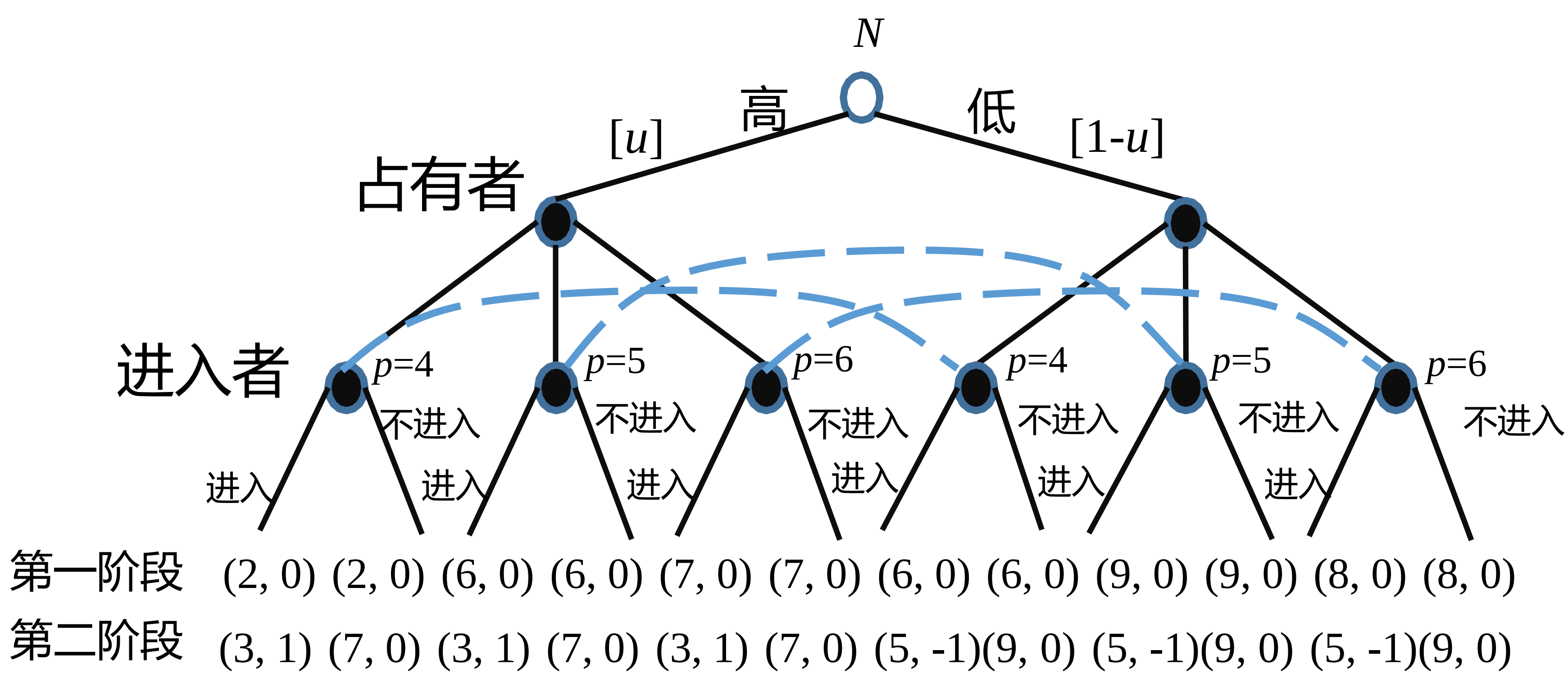
- 占有者有两个潜在类型：高成本和低成本；进入者只有一个类型：进入成本；因此只有进入者需要修正信念。
- 令 $\tilde{u}(p)$是进入者在观测到占有者的价格选择后认为在位者是高成本的后验概率（注意：这里p代表占有者的价格而非概率）。



第一阶段	(2, 0)	(2, 0)	(6, 0)	(6, 0)	(7, 0)	(7, 0)	(6, 0)	(6, 0)	(9, 0)	(9, 0)	(8, 0)	(8, 0)
第二阶段	(3, 1)	(7, 0)	(3, 1)	(7, 0)	(3, 1)	(7, 0)	(5, -1)	(9, 0)	(5, -1)	(9, 0)	(5, -1)	(9, 0)

示例回顾：不完全信息市场进入阻挠博弈

- 首先证明占有者选择单阶段最优垄断价格（即高成本者选择$p = 6$，低成本者选择$p = 5$）并不是精炼贝叶斯均衡：因为即使在高成本时选择$p = 5$可以获得更大收益
- （$13 = 6 + 7 > 7 + 3 = 10$）。



示例回顾：不完全信息市场进入阻挠博弈

- 下来证明：当$u < 1/2$时，精炼贝叶斯均衡是：不论成本高低，占有者选择$p = 5$；当观测到$p = 6$时，进入者选择进入，否则不进入。

- 证明：首先容易得到占有者无论成本高低，选择$p = 5$都是最优的；占有者选择$p = 5$时，进入者不能从观测到的价格得到任何新的信息，即$\tilde{u}(5) = \frac{1\times u}{(1\times u)+1\times(1-u)} = u < \frac{1}{2}$，进入者选择进入的期望收益是$u \times 1 + (1-u) \times (-1) = 2u - 1 < 0$，选择不进入的期望收益是0，因此进入者的最优策略是不进入。

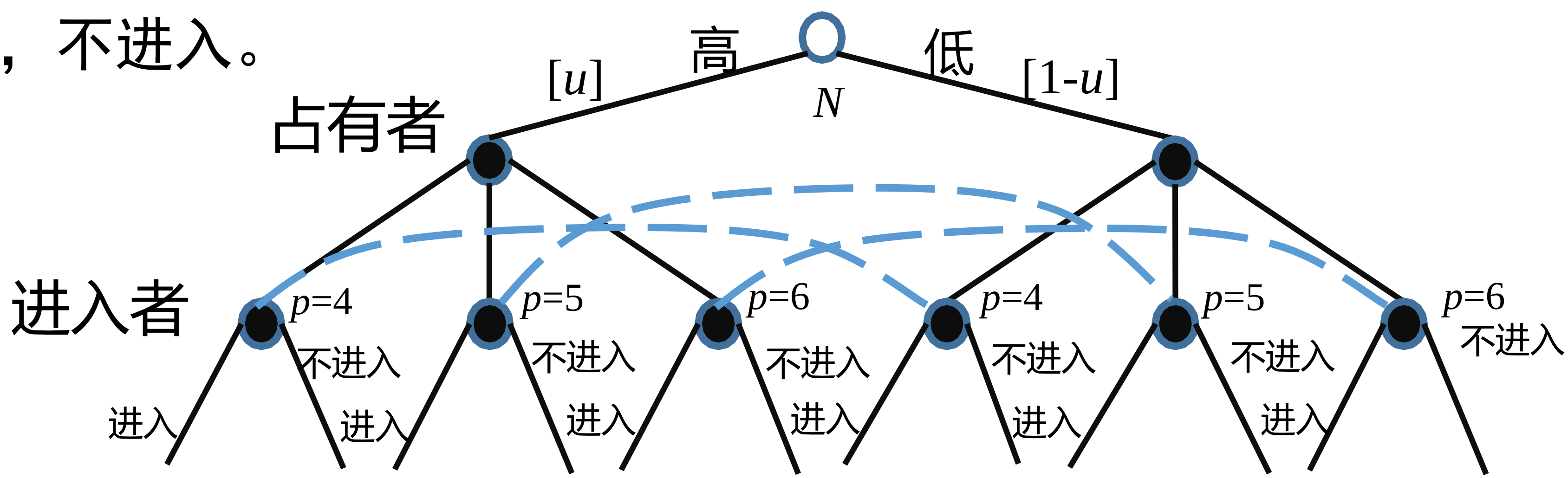
- 上述均衡称为**混同均衡**，因为两类占有者选择相同的价格，用于隐藏自己是高成本这个事实。

示例回顾：不完全信息市场进入阻挠博弈

- 最后证明：当$u \geq 1/2$时，精炼贝叶斯均衡是：低成本在位者选择$p = 4$，高成本占有者选择$p = 6$；当观测到$p = 4$时，进入者选择不进入，否则进入。
- 证明：首先不同类型占有者选择相同价格$p = 5$，因为$\tilde{u}(5) \geq 1/2$，进入者将进入。下来考虑低成本占有者：选择$p = 4$，进入者不进入，占有者收益为$6 + 9$；选择$p = 5$，进入者进入，占有者收益为$9 + 5$；因此$p = 4$最优；考虑高成本占有者：选择$p = 4$，进入者不进入，占有者收益为$2 + 7$；选择$p =\ 6$，占有者收益$7 + 3$。最后考虑进入者后验概率和策略。看到$p = 6$，进入；看到$p = 4$，不进入。
- 上述均衡称为**分离均衡**，因为两类占有者选择不同的价格，用于证明自己真实的成本。

示例回顾：不完全信息市场进入阻挠博弈

- 分离均衡证明回顾：首先不同类型在位者选择相同价格$p = 5$，进入者将进入。下来考虑低成本在位者：选择$p = 4$，进入者不进入，在位者收益为$6 + 9$；选择$p = 5$，进入者进入，在位者收益为$9 + 5$；因此$p = 4$最优；考虑高成本在位者：选择$p = 4$，进入者不进入，在位者收益为$2 + 7$；选择$p = 6$，进入者收益$7 + 3$。最后考虑进入者后验概率和策略。看到$p = 6$，选择进入；看到$p = 4$，不进入。



第一阶段 (2, 0) (2, 0) (6, 0) (6, 0) (7, 0) (7, 0) (6, 0) (6, 0) (9, 0) (9, 0) (8, 0) (8, 0)

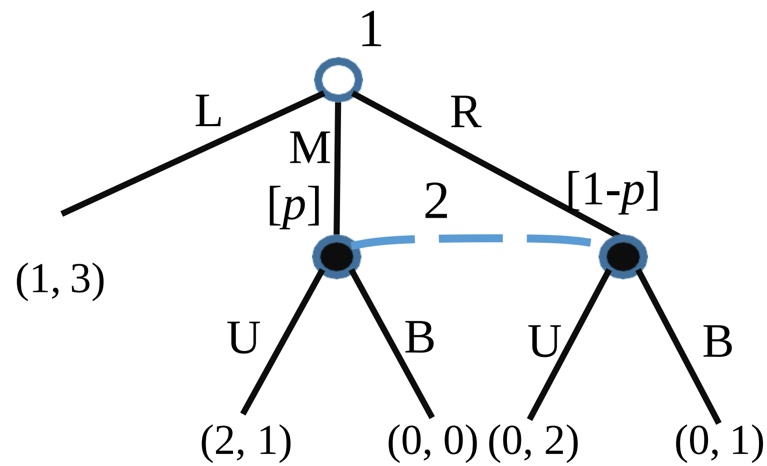
第二阶段 (3, 1) (7, 0) (3, 1) (7, 0) (3, 1) (7, 0) (5, -1)(9, 0) (5, -1)(9, 0) (5, -1)(9, 0)

不完美信息的精炼贝叶斯博弈

- 因为不完全信息博弈能够通过海萨尼转换为不完美信息博弈，因此精炼贝叶斯均衡也适用于不完美信息博弈。
- 贝叶斯法则在不完美信息博弈中的运用
  - 在不完全信息博弈中，参与人根据观测到其他参与人的行动和其他参与人的最优策略，使用贝叶斯法则修正对其他参与人类型的信念。
  - 在不完美信息博弈中，参与人观测不到其他参与人的行动，通过观测到博弈是否进入自己的信息集，修正自己处于该信息集的每一个决策结的概率。
  - 尽管贝叶斯法则在非均衡路径上没有定义，但如何规定非均衡路径上的后验概率是至关重要的。

不完美信息的精炼贝叶斯均衡

- 不完美信息博弈举例1：
  - 参与人1首先行动，选择L、M或R
  - 选择L，博弈结束；
  - 选择M或R，参与人2选择U或B

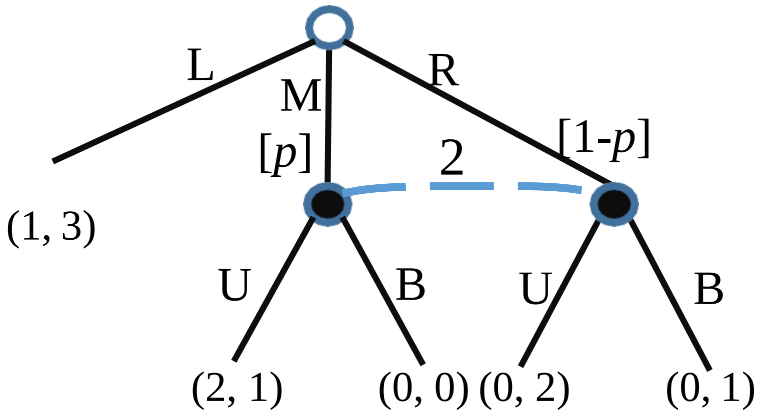


- 不完美信息
  - 参与人2在作出自己决策时（U or B）时并不知道参与人 1 选择了M 或是 R，尽管他知道 L没有被选择。
- 均衡分析
  - 两个纯策略纳什均衡(L, B)和(M, U)
  - 这两个同时也是子博弈精炼纳什均衡
  - 但(L, B)却不是合理的均衡
  - 需要通过精炼贝叶斯均衡剔除

参与人2 / 参与人1	U	B
L	1, 3	1, 3
M	2, 1	0, 0
R	0, 2	0, 1

不完美信息的精炼贝叶斯均衡

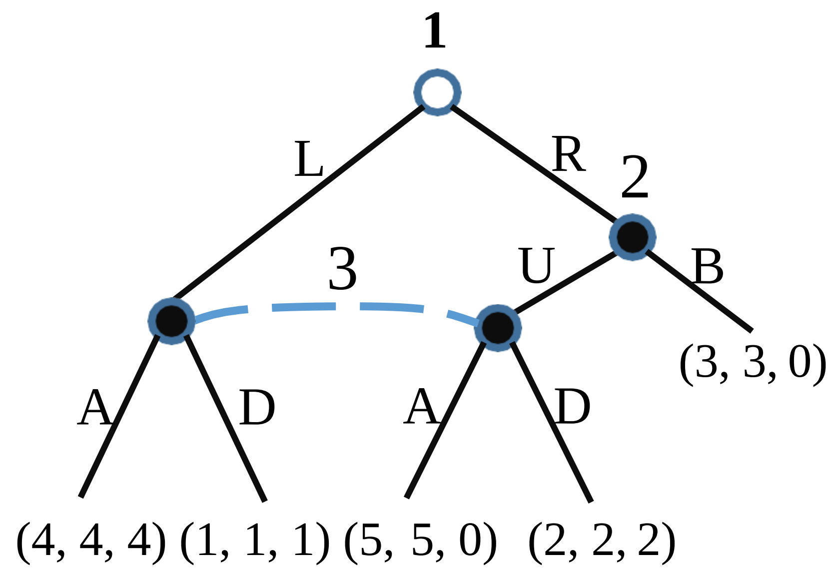
- 不完美信息博弈举例1：
  - 参与人1首先行动，选择L、M或R
  - 选择L，博弈结束；
  - 选择M或R，参与人2选择U或B



- 通过精炼贝叶斯均衡剔除(L,B)
  - 当博弈进入到参与人2的信息集时，必须有一个参与人1选择M或者R的概率分布，假设其为p和$1-p$：
    - 参与人2选择U收益：$p \times 1 + (1-p) \times 2 = 2-p$
    - 参与人2选择B收益：$p \times 0 + (1-p) \times 1 = 1-p$
    - 由于$2 - p > 1 - p$，故参与人2一定会选择U，因此贝叶斯均衡(L, B)依赖于一个不可置信的危险：(L, U)
  - 给定参与人2的最优策略U，参与人1的最优策略是M，因此(M, U)是一个精炼贝叶斯均衡。

不完美信息的精炼贝叶斯均衡

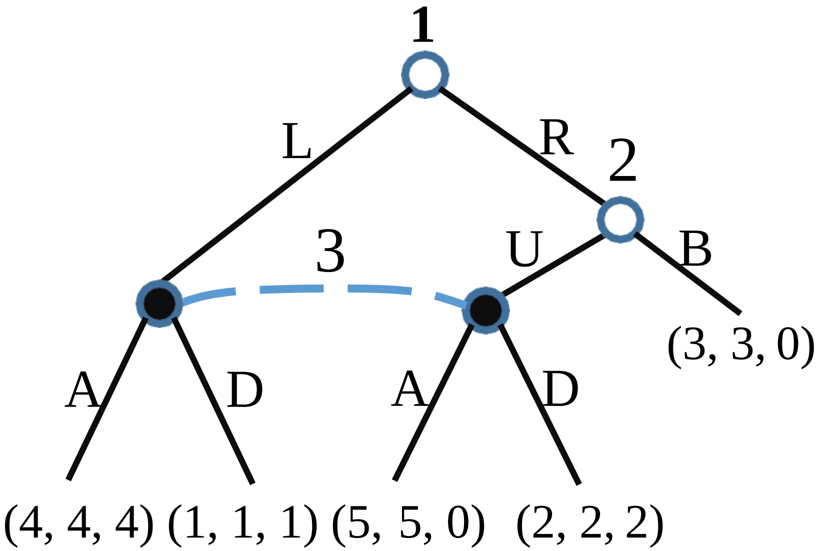
- 不完美信息博弈举例2：
  - 三个参与人，$i = 1, 2, 3$；
  - 1首先行动，选择L或R
  - 1选择R，2选择U或B（结束）
  - 1选择L，或2选择U，3选择A或D
- 不完美信息：参与人3的信息集
- 均衡分析
  - 两个纳什均衡(L, B, A)和(R, B, D)
  - 同时也是子博弈精炼纳什均衡
  - 但(L, B, A)却不是合理的均衡
  - 需要通过精炼贝叶斯均衡剔除



3	A		D	
1 \ 2	U	B	U	B
L	4,4,4	4,4,4	1,1,0	1,1,0
R	5,5,0	3,3,0	2,2,2	3,3,0

不完美信息的精炼贝叶斯均衡

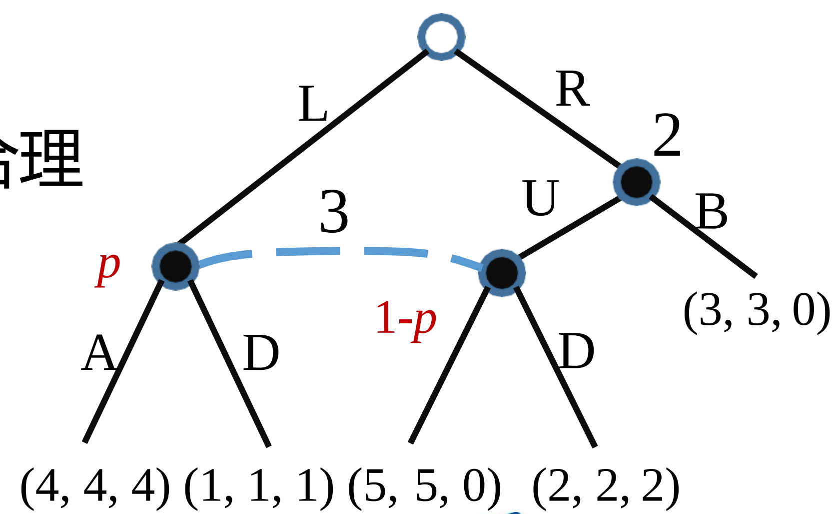
- 不完美信息博弈举例2：
  - 三个参与人，$i = 1, 2, 3$；
  - 1首先行动，选择L或R
  - 1选择R，2选择U或B（结束）
  - 1选择L，或2选择U，3选择A或D



- 利用精炼贝叶斯均衡求解均衡
  - 均衡(L, B, A)中2的选择B不在均衡路径上，需要利用精炼贝叶斯均衡检查其合理性：给定3选择A，2的最优选择是U而不是B
  - 给定2选择U而不是B，1应该选择R而不是L，进而3应该选择D
  - 假设3选择D，2应该选择B，进而1应该选择R。
  - 可以检查(R, B, D)是该博弈的另一个纳什均衡。

不完美信息的精炼贝叶斯均衡

- 不完美信息博弈举例2：精炼贝叶斯均衡要求
  - 1）每一个参与人的信息集上有一个概率分布；
  - 2）给定概率分布和其他参与人选择，每个参与人的策略最优；
  - 3）概率分布是使用贝叶斯法则从最优策略和观测到的行动得到的（在可能的情况下）；
- 参与人1和2都只有一个单结信息集，在该决策结他们决策的概率为1；参与人3的信息集有两个决策结：p和1-p;
- （L, B, A）及相关联的$p = 1$
  - 不满足（2）：B不是最优
  - 不满足（3）：进而B的$p = 1$不合理
- 检查一下均衡（R, B, D; p<2/5）

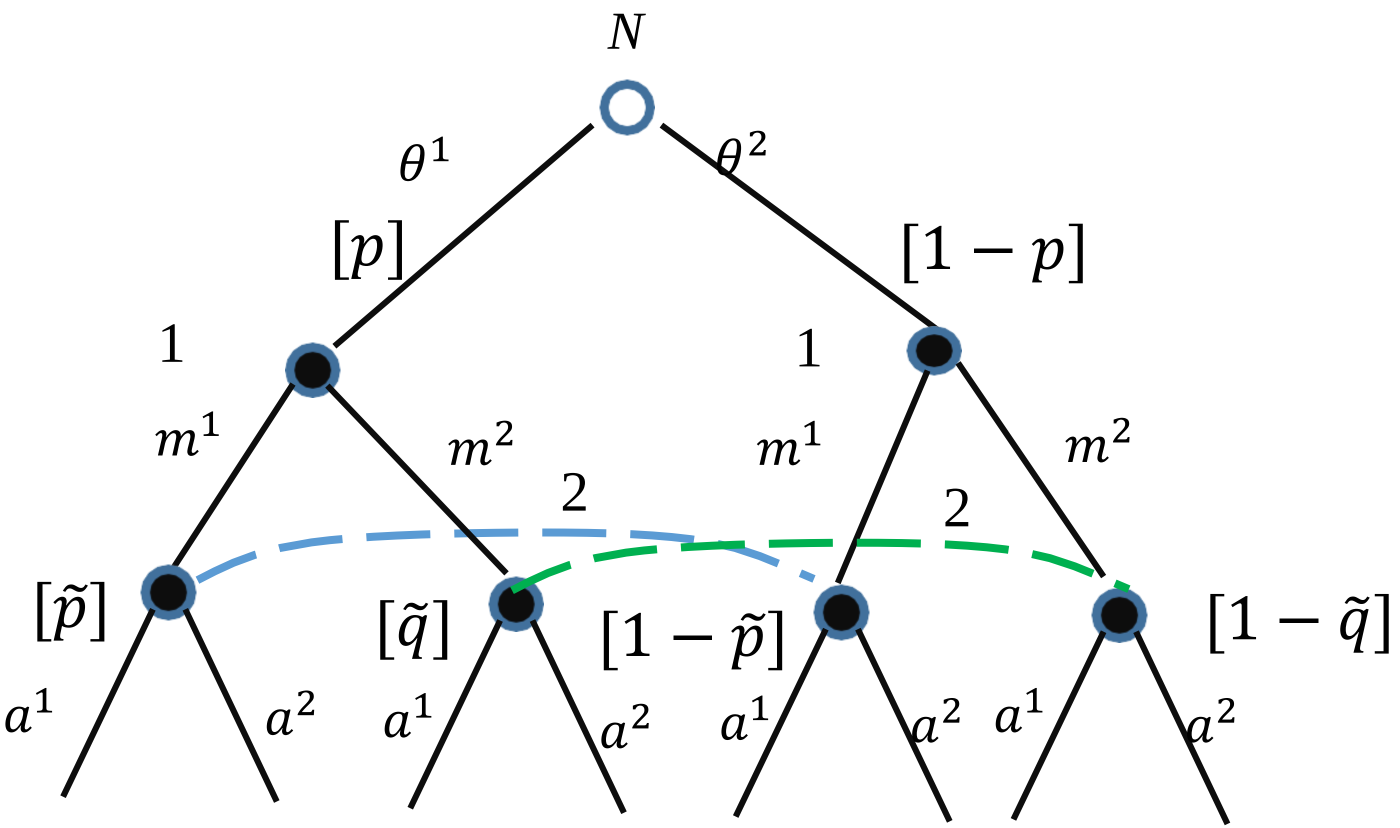


信号传递博弈介绍

- 博弈描述：两个参与人，参与人1称为信号的发送者，参与人2称为信号的接收者；参与人1的类型是私人信息，参与人2的类型是公共信息。
- 博弈顺序：
  - 首先，"自然"选择参与人1的类型$\theta \in \Theta$，$\Theta = \{\theta^1, \cdots, \theta^K\}$是参与人1的类型空间，也是参与人1的私有信息，参与人2不知道具体取值，但是参与人2知道参与1属于类型θ的先验概率是$p = p(\theta)$, 且$\sum_k p(\theta^k) = 1$。
  - 然后，参与人1在观测到类型θ之后选择发出信号$m \in M$，这里$M = \{m^1, \cdots, m^J\}$是信号空间。
  - 接下来，参与人2 观测到参与人1的信号m（但不知道具体类型θ）,使用贝叶斯法则，从先验概率$p = p(\theta)$得到后验概率$\tilde{p} = \tilde{p}(\theta|m)$，然后选择行动$a \in A$，其中$A = \{a^1, \cdots, a^H\}$是参与人2的行动空间。
  - 最后，得到支付函数$u_1(m, a, \theta)$和$u_2(m, a, \theta)$。

信号传递博弈介绍

- 信号传递博弈的执行顺序示例
  - 具体赋值：$K = J = H = 2, \tilde{p} = \tilde{p}(\theta^1|m^1), \tilde{q} = \tilde{p}(\theta^1|m^2)$,
  - 扩展式表示：



信号传递博弈的精炼贝叶斯均衡

信号传递博弈的精炼贝叶斯均衡

信号传递博弈的精炼贝叶斯均衡是一个策略组合$(m^*(\theta), a^*(m))$和后验概率$\tilde{p}(\theta|m)$，满足：

（P_1） $a^*(m) \in \arg\max_{a} \Sigma_\theta\ \tilde{p}(\theta|m) u_2(m, a, \theta)$；

（P_2） $m^*(\theta) \in \arg\max_{m} u_1(m, a^*(m), \theta)$；

（B） $\tilde{p}(\theta|m)$是参与人2使用贝叶斯法则从先验概率$p(\theta)$、观测信号m和参与人1的最优策略$m^*(\theta)$到的（在可能的情况下）。

- 说明：
  - P_1和P_2等价于精炼贝叶斯均衡中的P条件；
  - P_1指给定后验概率，参与人2对参与人1的信号做出的最优反应；
  - P_2是指预测到参与人2的最优反应，参与人1将做出的决策；
  - B限制了贝叶斯条件。

信号传递博弈的精炼贝叶斯均衡

- 信号传递博弈所有可能的精炼贝叶斯均衡的三类均衡

分离均衡	混同均衡	准分离均衡
• 不同类型的发送者以确定性的概率发送不同的信号。	• 不同类型的发送者选择相同的信号。	• 一些类型的发送者发随机选择信号，另一些类型的发送者选择特定信号。

分离均衡

- 信号传递博弈中不同类型的发送者（参与人1）以1的概率选择不同的信号，或者说没有任何类型选择与其他类型相同的信号。在分离均衡下，信号准确地揭示出类型。

信号传递博弈的分离均衡

假定$K=J=2$(即只有两个类型两个信号)，那么分离均衡意味着：如果m^{1}是类型θ^1的最优选择，那么m^1就不可能是类型θ^2的最优选择，并且，m^2一定是类型θ^2的最优选择。即：

$$u_1(m^1,a^*(m),\theta^1)>u_1\left(m^2,a^*(m),\theta^1\right);$$

$$u_1\left(m^2,a^*(m),\theta^2\right)>u_1\left(m^1,a^*(m),\theta^2\right)$$

因此，后验概率是：

$$\tilde{p}(\theta^1|m^1)=1,\tilde{p}\left(\theta^1|m^2\right)=0,$$

$$\tilde{p}\left(\theta^2|m^1\right)=0,\tilde{p}\left(\theta^2|m^1\right)=1$$

混同均衡

- 不同类型的发送者选择相同的信号，或者说，没有任何类型选择其它类型不同的信号，因此接收者不修正先验概率。

信号传递博弈的混同均衡

假定m^j是均衡策略，那么，

$$u_1\left(m^j, a^*(m), \theta^1\right) \geq u_1(m, a^*(m), \theta^1);$$
$$u_1\left(m^j, a^*(m), \theta^2\right) > u_1\left(m, a^*(m), \theta^2\right)$$
$$\tilde{p}\left(\theta^k | m^j\right) \equiv p(\theta^k)$$

- 从贝叶斯分析的角度：各个分布可能得到的观测样本是相同的，那么采样对分布的估计就没有什么帮助，只能是原先认为是什么分布，现在仍然维持原判。

准分离均衡

- 一些类型的发送者随机选择信号，另外一些类型的发送者选择特定的信号。

信号传递博弈的准分离均衡

假定类型θ^1的发送者随机地选择m^1或者m^2，而类型θ^2的发送者以1的概率选择m^2，如果这个策略组合是均衡策略组合，那么

$$u_1(m^1, a^*(m), \theta^1) = u_1\left(m^2, a^*(m), \theta^1\right);$$

$$u_1\left(m^2, a^*(m), \theta^2\right) < u_1\left(m^1, a^*(m), \theta^2\right)$$

后验概率

$$\tilde{p}(\theta^1|m^1) = \frac{\alpha \times p(\theta^1)}{\alpha \times p(\theta^1) + 0 \times p(\theta^2)} = 1;$$

$$\tilde{p}\left(\theta^1|m^2\right) = \frac{(1-\alpha) \times p(\theta^1)}{(1-\alpha) \times p(\theta^1) + 1 \times p(\theta^2)} < p(\theta^1),$$

$$\tilde{p}\left(\theta^2|m^2\right) = \frac{1 \times p(\theta^1)}{(1-\alpha) \times p(\theta^1) + 1 \times p(\theta^2)} > p\left(\theta^2\right)$$

示例：米尔格罗姆-罗布茨垄断限价模型

- 博弈描述:
  - 假定有两个企业，企业1在位者，企业2进入者。
  - 第一阶段：企业1垄断生产，价格p_1;
  - 第二阶段：企业2决定是否进入。如果企业2进入，市场变为双寡头竞争，否则，企业1仍然保持垄断地位。
  - 企业1 有两个潜在类型，为高成本H，概率为$\mu(H)$，或者低成本L，概率为$1-\mu(H)$。 类型为私有信息。对应的利润记作$M_1^\theta(p_1)$，其中$\theta = H, L$。
  - p_m^θ表示类型θ的垄断价格，$M_1^\theta = M_1^\theta(p_m^\theta)$表示最大短期垄断利润，其中$p_m^H > p_m^L$，$M_1^H < M_1^L$ （假设$M_1^\theta(p_1)$ 严格凹函数）
  - 另外，假设第一阶段企业1的类型是私有信息，企业2不知道。第二阶段，企业2如果选择并执行了进入的操作之后，他可以得知企业1的类型。这个假设是为了让第二阶段的寡头价格独立于第一阶段的寡头价格。并且把第二阶段两个企业的净利润记作D_1^θ和D_2^θ。同时假设$D_2^H > 0 > D_2^L$，共同贴现因子为δ。

示例：米尔格罗姆-罗布茨垄断限价模型

- 因为企业1的目的是想要保持垄断地位（假设$M_1^0 > D_1^0$）
- 他得想办法让企业2认为自己是低成本，从而选择不进入，自己保持垄断地位。如何做到呢？
  - 广告宣传?
  - 嘴上告知?
  - 不可置信的威胁。
  - 得传递信号：定一个低的价格p_1^L。
- 从后面的分析我们会看到，即使企业1是高成本的，他也可能会选择低价格p_1^L。因为第一阶段的垄断利润损失可能会被第二阶段继续保持垄断地位的收益所弥补。

示例：米尔格罗姆-罗布茨垄断限价模型

- 分离均衡情况的思路：
  - 先找到两个分离均衡的必要条件；
  - 然后讨论必要条件中得到的解也是均衡的充分条件；
- 两个分离均衡的必要条件：

1. 类型H的在位者不愿意选择类型L的均衡价格p_1^L；

高成本在位者选择$p_1^H = p_m^H$利润(两个阶段利润贴现之后）：$M_1^H + \delta D_1^H$
高成本在位者选择了低成本在位者的价格p_1^L时，他的利润：$M_1^H(p_1^L) + \delta M_1^H$
现在要满足第1个条件，使得p_1^H是高成本在位者的均衡：

$$(A)\ M_1^H + \delta D_1^H \geq M_1^H(p_1^L) + \delta M_1^H \quad 或\ (A')\ M_1^H - M_1^H(p_1^L) \geq \delta(M_1^H - D_1^H)$$

2. 类型L的在位者也不愿意选择类型H的均衡价格p_1^H；

低成本在位者选择p_1^L从而阻止进入时，他的利润：$M_1^L(p_1^L) + \delta M_1^L$
低成本在位者如果选择了p_1^H(即，p_m^L)，他的收益是：$M_1^L + \delta D_1^L$
同理，我们得到保证p_1^L是低成本在位者均衡的条件：

$$(B)\ M_1^L(p_1^L) + \delta M_1^L \geq M_1^L + \delta D_1^L \quad 或\ (B')\ M_1^L - M_1^L(p_1^L) \leq \delta(M_1^L - D_1^L)$$

示例：米尔格罗姆-罗布茨垄断限价模型

- 分离均衡情况的思路：
  - 先找到两个分离均衡的必要条件；
  - 然后讨论必要条件中得到的解也是均衡充分条件；
- 两个分离均衡的必要条件：
  1. 类型H的在位者不愿意选择类型L的均衡价格p_1^L；

     （A）$M_1^H+\delta D_1^H \geq M_1^H(p_1^L)+\delta M_1^H$ 或（A′）$M_1^H-M_1^H\left(p_1^L\right) \geq \delta(M_1^H-D_1^H)$

  2. 类型L的在位者也不愿意选择类型H的均衡价格p_1^H；

     （B）$M_1^L\left(p_1^L\right)+\delta M_1^L \geq M_1^L+\delta D_1^L$ 或（B′）$M_1^L-M_1^L\left(p_1^L\right) \leq \delta(M_1^L-D_1^L)$

  3. 另外，我们假设：不存在$p_1^L=p_m^H$的分离均衡。即，如果$p_1^L=p_m^H$，高成本的在位者也将选择p_1^L：

     （C）$M_1^H-M_1^H\left(p_m^L\right)<\delta(M_1^H-D_1^H)$
- 接下来，我们要找到满足上述条件的p_1^L，它将由一个区间构成，$p_1^L \in [\bar{\bar{p}}, \bar{p}]$构成一个分离均衡价格。

示例：米尔格罗姆-罗布茨垄断限价模型

- 斯宾塞-莫里斯条件（分离条件/单交叉条件）

$$(\mathrm{SM})\quad \frac{\partial}{\partial p_1}\left(M_1^H(p_1)-M_1^L(p_1)\right)>0 \text{ 或 } \frac{\partial M_1^H(p_1)}{\partial p_1}>\frac{\partial M_1^L(p_1)}{\partial p_1}$$

- 上式含义：改变价格对不同类型的企业的利润影响是不同的；特别地，高成本企业比低成本企业更愿意选择高价格（更不愿意选择低价格）。
- 很多情况满足这个条件，比如，企业成本是常数c^H和c^L，$c^H>c^L$，市场需求函数$Q(P_1)$，可以验证：

$$\frac{\partial M_1^H(p_1)}{\partial p_1}=\frac{\partial}{\partial p_1}\left((p_1-c^H)Q(P_1)\right)=Q(P_1)+(p_1-c^H)\frac{\partial Q(P_1)}{\partial p_1}$$

$$\frac{\partial M_1^L(p_1)}{\partial p_1}=\frac{\partial}{\partial p_1}\left((p_1-c^L)Q(P_1)\right)=Q(P_1)+(p_1-c^L)\frac{\partial Q(P_1)}{\partial p_1}$$

因为$c^H>c^L$，$\frac{\partial Q(P_1)}{\partial p_1}<0$，所以$\frac{\partial M_1^H(p_1)}{\partial p_1}>\frac{\partial M_1^L(p_1)}{\partial p_1}$

示例：米尔格罗姆-罗布茨垄断限价模型

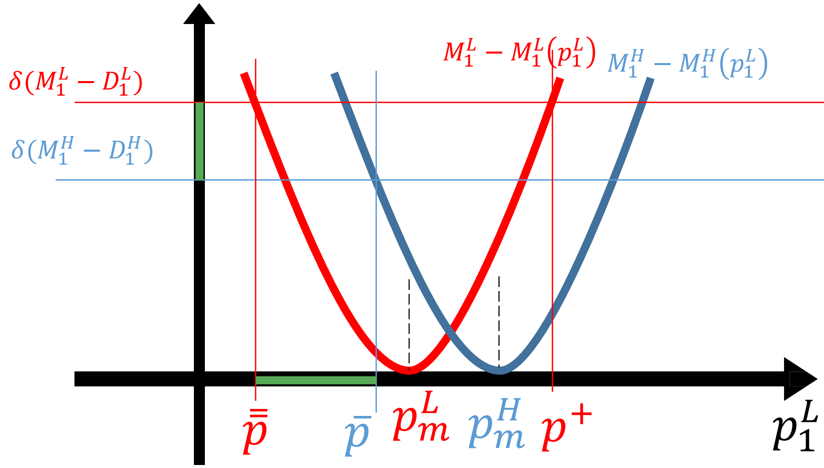
- 斯宾塞-莫里斯条件（分离条件/单交叉条件）

$$(\mathrm{SM})\quad \frac{\partial}{\partial p_1}\left(M_1^H(p_1)-M_1^L(p_1)\right)>0 \text{ 或 } \frac{\partial M_1^H(p_1)}{\partial p_1}>\frac{\partial M_1^L(p_1)}{\partial p_1}$$

- 成本类型是连续分布，上述条件变为

$$\frac{\partial^2 M(p,c)}{\partial p\partial c}=\frac{\partial^2}{\partial p\partial c}\big((p-c)Q(p)\big)=-\frac{\partial Q(p)}{\partial p}>0$$

- 这个条件保证了$y=M_1^L-M_1^L(p_1^L)$与$y=M_1^H-M_1^H(p_1^L)$在(p_1^L,y)空间只交叉一次：



$M_1^L-M_1^L(p_1^L)\le\delta(M_1^L-D_1^L)$

$M_1^H-M_1^H(p_1^L)\ge\delta(M_1^H-D_1^H)$

示例：米尔格罗姆-罗布茨垄断限价模型

- 分离均衡情况的思路：
  - 先找到两个分离均衡的必要条件；
  - 然后讨论必要条件中得到的解也是均衡充分条件；
- 斯宾塞-莫里斯条件是为了必要性提出的，接下来探讨满足这个条件的点同样是分离均衡的充分条件：

假定高成本在位者选择p_m^H，低成本在位者选择$p_1^L \in [\bar{\bar{p}}, \bar{p}]$，我们来证明它是分离均衡。

- 当观测到p_m^H时，进入者认为在位者是高成本的概率是1，选择进入；当观察到p_1^L时，进入者认为在位者高成本的概率为0，选择不进入。
- 当观测到的价格不属于这两个价格时（非均衡路径），后验概率是任意的。

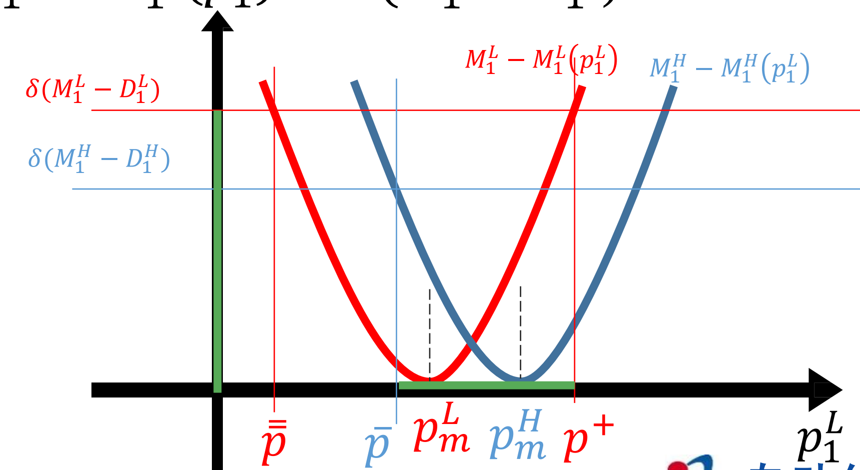
- 保证(p_m^H, p_1^L)是均衡策略最简单的办法是让$\tilde{\mu}\left(H \middle| p_1 \neq p_m^H, p_1^L\right) = 1$（观测如果不是$p_m^H$和$p_1^L$就选择进入），这使得没有任何类型的在位者有兴趣偏离所假定的均衡策略。

示例：米尔格罗姆-罗布茨垄断限价模型

- 分离均衡的结果是：
  - 高成本在位者选择p_m^H，低成本在位者选择任何$p_1^L \in [\bar{\bar{p}}, \bar{p}]$；
  - 观测者观测到p_1^L时就认为$\tilde{\mu}(H|p_1^L) = 0$，选择不进入；
  - 观测者观测到$p_1 \neq p_1^L$时就认为$\tilde{\mu}(H|p_1 \neq p_1^L) = 1$,选择进入。
- $p_1^L = \bar{p}$是最低成本均衡。
- 在所有的均衡中，低成本在位者限制自己的价格低于垄断价格以阻止进入者进入。
- 连续分离均衡对于任何$\mu(H) > 0$均存在。
- 当$\mu(H) = 0$，低成本在位者就会选择短期垄断价格p_m^L。
- 进入者对在位者拥有高成本的信念强弱影响着均衡的分析。

示例：米尔格罗姆-罗布茨垄断限价模型

- 混同均衡的情况讨论
  - 混同均衡存在的条件是：(PE) $\mu D_2^H + (1-\mu) D_2^L < 0$
    - 当该条件满足时，进入者得不到新的信息选择不进入，混同均衡存在。
    - 当该条件不满足时，在位者不能阻止进入者进入，只能采取垄断价格，混同均衡不存在。
  - 假设(PE)条件满足，p_1是一个能阻止进入者进入的混同价格。
    - p_1是混同均衡价格的必要条件是不论是高成本还是低成本的在位者都不愿意偏离p_1；如果偏离的话，在最坏的情况下诱使进入者进入。
    - p_1对低成本在位者要满足$M_1^L - M_1^L(p_1) \le \delta(M_1^L - D_1^L)$
    - p_1对高成本在位者要满足$M_1^H - M_1^H(p_1) \le \delta(M_1^H - D_1^H)$
  - 混同均衡为：
    - 高低成本都是$p_1 \in [\bar{p}, p^+]$
    - 进入者观测到p_1，就认为 $\tilde{\mu}(H|p_1) \equiv \mu(H)$，不进入。



为什么还需要其他的均衡概念？

- 不完全信息博弈存在多重精炼贝叶斯均衡：
  - **非均衡路径的后验概率本来是任意的**，那么我们通过不同的规定来对它进行处理，就会得到不同的均衡。
  - 这种现像在不完美信息动态博弈中普遍存在。
- 以信号传递博弈为例：
  - 在位者的均衡策略依赖于进入者认为什么不是在位者的均衡策略，或者说，进入者认为什么是在位者的策略均衡什么就是在位者的策略均衡，均衡是自动实现的。
  - 精炼贝叶斯均衡剔除了不可置信的策略（行动），没有剔除不可置信的信念（后验概率）。
- 如何从多重精炼贝叶斯均衡中得到更为合理的均衡结果？哪一个均衡更为合理、更可能出现呢？
  - 一些基本改进：剔除劣策略、直观标准
  - 其他更强的均衡：序贯均衡、颤抖手均衡等

剔除劣策略

- 基本思路：将“不选择劣策略”的要求扩展到非均衡路径的后验概率上。

剔除劣策略的均衡

令a_1'和a_1''是参与人1（信号发送者）的两个行动（信号），$a_1', a_1'' \in A_1$。对于参与人2（信号接收者）的所有行动$a_2', a_2'' \in A_2$，如果下列条件成立，我们说对于类型$\theta_1 \in \Theta_1$的参与人1，a_1'弱劣于a_1''：

$$u_1(a_1', a_2', \theta_1) \le u_1(a_1'', a_2'', \theta_1)$$

(至少有一个不等式对于某些a_1'', a_2''成立)

- 也就是后行动的观察者要假设前面的人是不会选择劣策略的，并以此来更新后验概率

剔除劣策略

- 例如：米尔格罗姆-罗布茨垄断限价模型
  - 在分离均衡中，不论进入者如何规定非均衡路径上的后验概率，如果高成本的在位者选择p_m^H，他的最低利润是$M_1^H + \delta D_1^H$；如果高成本的在位者选择任何$p_1 \leq \bar{p}$,他能得到的最大利润是$M_1^H(p_1) + \delta M_1^H$。根据我们之前对$\bar{p}$的定义，$M_1^H(p_1) + \delta M_1^H \leq M_1^H + \delta D_1^H$。
  - 因此，对于高成本在位者来说，$p_1 \leq \bar{p}$是否劣于p_m^H取决于进入者如何规定非均衡路径上的后验概率。
  - **如果观测到在位者选择了$\boldsymbol{p_1 \leq \bar{p}}$，进入者应该认为在位者是低成本而不可能是高成本。即$\boldsymbol{\tilde{\mu}(H|p_1 \leq \bar{p}) = 0}$，选择不进入。**
  - 给定这个后验概率，低成本的在位者不需要为了阻止进入而选择$p_1 < \bar{p}$（也就是不用再选更低的价格）。
  - 唯一合理的精炼贝叶斯分离均衡：低成本在位者选择$p_1^L = \bar{p}$，高层本的在位者选择$p_1^H = p_m^H$；如果进入者观测到$p_1 > \bar{p}$，进入，否则不进入。

直观标准

- 基本思想：将劣策略扩展到相对于均衡策略的劣策略

直观标准

假定$(a_1^*, a_2^*; \tilde{\mu})$是一个精炼贝叶斯均衡。令$u_1^*(\theta_1)$是类型$\theta_1$的参与人1的均衡效用水平。那么，$a_1' \in A_1$是参与人1相对于均衡$(a_1^*, a_2^*; \tilde{\mu})$的劣策略，如果对于参与人2所有的行动$a_2 \in A_2$，下列条件成立：

$$u_1(a_1', a_2, \theta_1) \leq u_1^*(\theta_1)$$

（至少有一个不等式对于某些$a_2 \in A_2$成立）

进一步，令$\tilde{\Theta}_1 \subset \Theta_1$是所有满足上述条件的不等式的$\theta_1$的集合。如果$\tilde{\Theta}_1 \neq \Theta_1$，那么，参与人2在非均衡路径上合理的后验概率是：

$$\sum_{\theta_1 \in \tilde{\Theta}_1} \tilde{\mu}(\theta_1 | a_1') = 0$$

直观标准

- 例如：米尔格罗姆-罗布茨垄断限价模型
  - 考虑两个策略p^*和p_m^L(低成本企业的短期垄断价格)
  - 首先，两个策略都是潜在混同均衡策略（属于$[\bar{p}, p^+]$内），根据上一讲的定义是没有劣策略的。
    - 如果在位者选择p^*，进入者认为$\tilde{\mu}(H|p^*) = 1$，选择进入，在位者选择p_m^L时，进入者认为$\tilde{\mu}\left(H|p_m^L\right) = 0$，选择不进入，则$p^*$劣于$p_m^L$。
    - 如果在位者选择p^*，进入者认为$\tilde{\mu}(H|p^*) = 0$，选择不进入，在位者选择p_m^L时，进入者认为$\tilde{\mu}\left(H|p_m^L\right) = 1$，选择进入，则$p_m^L$劣于$p^*$。
  - 假定p^*是混同均衡$\left(\tilde{\mu}(H|p^*) = \mu < \frac{1}{2}\right)$。上一节我们讨论过，高成本在位者是没有偏离$p^*$的意愿。现在来看低成本在位者。
  - 如果进入者认为$\tilde{\mu}\left(H|p_m^L\right) = 0$而选择不进入，低成本的在位者将选择$p_m^L$ 而不是p^*（这个偏离行为，将增加第一阶段利润而不改变第二阶段的利润）。
  - 因此，如果p_m^L出现，进入者的合理概率应该是$\tilde{\mu}\left(H|p_m^L\right) = 0$。
  - 结论：p^*不是一个合理的混同均衡，被剔除掉。

序贯均衡

- 基本思想是在贝叶斯均衡的基础上增加了一个新的要求：在博弈到达的每一个信息集上，参与人的行动必须由某种有关之前发生的事情的信念"合理化"。

序贯均衡

(σ,μ)是一个序贯均衡，如果它满足下列两个条件：

（S）(σ,μ)是序贯理性的：在所有的信息集h上，给定后验概率$\mu(h)$没有任何参与人i想偏离$\sigma_{i(h)}$; 即，对所有的可行策略$\sigma'_{i(h)}$,

$$u_{i(h)}(\sigma|h,\mu(h)) \geq u_{i(h)}(\sigma'_{i(h)},\sigma_{-i(h)}|h,\mu(h))$$

（C）(σ,μ)是一致的：存在一个严格混合策略组合序列$\{\sigma^m\}$和贝叶斯法则决定的概率序列μ^m, 使得(σ,μ)是(σ^m,μ^m)的极限，即

$$(\sigma,\mu) = \lim_{m\to\infty}(\sigma^m,\mu^m)$$

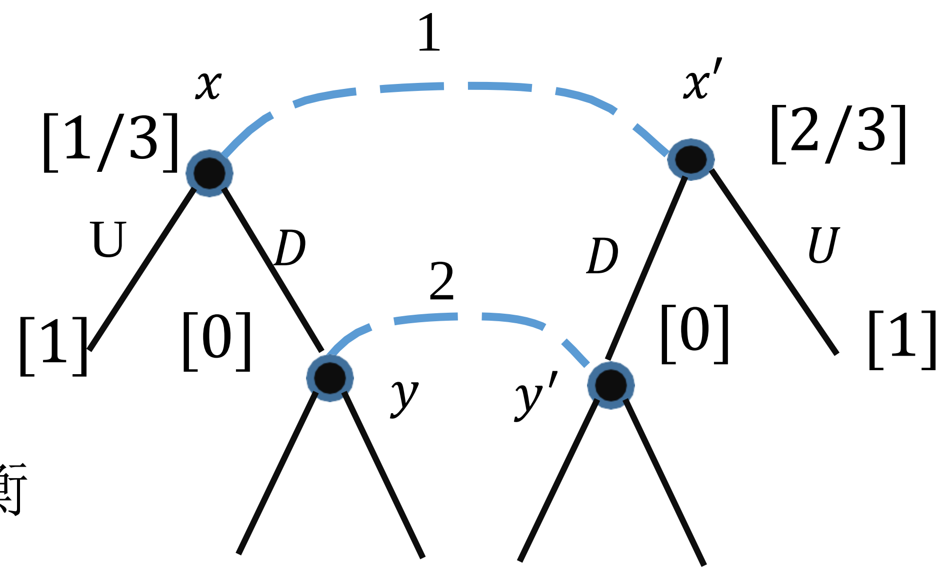
序贯均衡

- 例如：
  - 博弈到达参与人1的信息集

    参与人1认为:

    $\mu(x) = 1/3$，$\mu(x') = 2/3$.

    显然，无论出于哪个决策结，参与人1的最优决策是U，所以，参与人2的信息集是非均衡路径。



  - 当参与人1偏离了均衡路径，选择了D（在两个点偏离的可能性是相同的）：参与人2 认为$\mu(y) = 1/3$，$\mu(y') = 2/3$。
  - 因为任何$\mu(y)$都与贝叶斯法则相容，因为D是0概率事件。
  - 一致性条件(C)要求：给定一个收敛于0的序列ϵ^m，

$$\mu^m(y) = \frac{\mu^m(x)\epsilon^m}{\mu^m(x)\epsilon^m + \mu^m(x')\epsilon^m} = 1/3$$

颤抖手均衡

- 基本思想：
  - 在任何博弈中，任何参与人都有一定可能性犯错误，一个策略组合，只有当它在允许所有参与人都可能犯错是仍是每一个参与人的最优策略组合的时候，才是一个均衡。

颤抖手均衡

在n人策略式表达博弈中，纳什均衡$(\sigma_1, \cdots, \sigma_n)$ 是一个颤抖手均衡，如果对于每一个参与人i, 存在一个严格混合策略序列$\{\sigma_i^m\}$, 它满足下列两个条件：

(1) 对于每一个i， $\lim_{m\to\infty}\{\sigma_i^m\} = \sigma_i$

(2) 对于每一个i和$m = 1,2, \cdots,$

σ_i 是策略组合$\{\sigma_{-i}^m\} = \{\sigma_1^m, \cdots, \sigma_{i-1}^m, \sigma_{i-1}^m, \cdots, \sigma_n^m\}$的最优反应，

即：对任何可选择的混合策略$\sigma_i' \in \Sigma_i$,

$$u_i(\sigma_i, \sigma_{-i}^m) \geq u_i(\sigma_i', \sigma_{-i}^m)$$

颤抖手均衡举例

参与人1 \ 参与人2	L	R
U	-8, -8	0, -10
D	-10, 0	-1, -1

- 右图中(D, L)是纳什均衡：
  - 只要参与人2不选择R，D就是参与人1的最优选择；只要参与人1不选择U，L就是参与人2的最优策略。
  - 但是，如果参与人2错误地选择R，那么，参与人1的最优选择就是U而不是D。预测到这一点之后，参与人2就会选择R，使得均衡偏离。
  - 所以（D, L）不是一个颤抖手均衡。

- 右图中（U, R）是颤抖手均衡：
  - 不论参与人2犯错误的概率有多大，参与人1没有兴趣选择D；
  - 只要参与人1犯错误的概率小于2/3，参与人2就没有兴趣选择L。

完全信息有限重复博弈的悖论

- 完全信息下的重复博弈，理论上只要重复次数有限，唯一的子博弈精炼纳什均衡就是阶段博弈的均衡策略。

有限次重复博弈的均衡定理（Selten, 1978）

令Γ是阶段博弈，$\Gamma(T)$是重复T次的重复博弈（$T < \infty$）。那么，若Γ有唯一的纳什均衡，重复博弈$\Gamma(T)$的唯一子博弈精炼纳什均衡结果是阶段博弈Γ的纳什均衡重复T次（即每个阶段博弈结果都是单次博弈的均衡）。

- 有限重复囚徒困境中，囚徒每次都选择坦白不太合理。
- 现实中的试验表明：即使在有限重复博弈中，合作行为也会频繁出现。

这背后是什么原因造成的呢？

KMRW声誉模型的主要思想和结论

- 克瑞普斯（**K**reps）、米尔格罗姆（**M**ilgrom）、罗伯茨（**R**oberts）和威尔逊（**W**ilsom）四人在1982年提出的声誉模型（Reputation Model）通过将不完全信息引入到有限重复博弈中解开了这个悖论。

KMRW声誉模型的主要结论：

参与人对其他参与人支付函数或策略空间的不完全信息对均衡的结果有重要影响，导致合作行为会在有限次博弈中出现，只要博弈的次数足够长（没有必要是无限的）。特别地，“坏人”可能在相当长的一段时间时期内表现得向“好人”一样。

重复囚徒困境的KMRW声誉模型分析

- 重复囚徒困境
  - 阶段博弈的支付矩阵如右图所示：
  - 假设囚徒A有两种类型："理性的"和"非理性的"，概率分别为p和$1-p$。
  - 为方便分析，假设囚徒B是理性的。

囚徒A \ 囚徒B	坦白（C）	抵赖（D）
坦白（C）	-8, -8	0, -10
抵赖（D）	-10, 0	-1, -1

  - 理性的囚徒可以选择任何策略，非理性的囚徒只有一种策略："针锋相对"，第一次抵赖，后面根据上次对手选择做选择。
  - 博弈的顺序如下：
    1. 自然首先选择囚徒1的类型；囚徒A知道自己的类型，囚徒B只知道A属于理性的概率为p，非理性的概率为$1-p$；
    2. 两个囚徒进行第一阶段的博弈；
    3. 观测到第一阶段博弈结果后，进行第二阶段博弈；观测到第二阶段博弈结果之后，进行第三阶段博弈；如此重复直到结束。
  - 两个囚徒的效益函数都是阶段博弈的贴现值之和（为方便讨论，令贴现值$\delta = 1$）。

重复囚徒困境的KMRW声誉模型分析

• 首先考虑博弈只重复两次（$T=2$）的情况

求解：采用逆向推理法分析

囚徒 \ 次数	$t=1$	$t=2$
非理性囚徒A	D	X
理性囚徒A	C	C
囚徒B	X	C

在最后阶段（$t=2$），理性囚徒A和B都会选择C，非理性囚徒A的选择依赖于囚徒2在第一阶段的选择；

在第一阶段，根据假设非理性囚徒A选择D，理性囚徒A的最优选择是仍是C。因此只需要考虑囚徒B在第一阶段的选择X，他的选择将影响非理性囚徒A在第二阶段的选择：

如果选择X=D，囚徒B的期望收益为：

$$[p\times(-1)+(1-p)(-10)]+[p\times 0+(1-p)(-8)]=17p-18$$

如果选择X=C，囚徒B的期望收益为：

$$[p\times 0+(1-p)(-8)]+[-8]=8p-16$$

当$17p-18\geq 8p-16$，即$p\geq 2/9$，B将在第一阶段选择D，即A非理性概率不小于2/9时，A第一阶段将合作。下面分析令$p\geq 2/9$。

重复囚徒困境的KMRW声誉模型分析

- 下面考虑博弈重复三次的情况

求解： 根据重复二次的结论推理

给定$p \geq 2/9$，如果理性囚徒A和囚徒2第一阶段都选择D，那么第二、三阶段的均衡路径就和上页中的表格（左下图）相同。三阶段总的均衡路径如右下图所示。下面证明右下图是均衡的充分条件。

次数 / 囚徒	$t=1$	$t=2$
非理性囚徒A	D	X
理性囚徒A	C	C
囚徒B	X	C

次数 / 囚徒	$t=1$	$t=2$	$t=3$
非理性囚徒A	D	D	D
理性囚徒A	D	C	C
囚徒B	D	D	C

重复囚徒困境的KMRW声誉模型分析

- 下面考虑博弈重复三次的情况

继续求解：首先考虑理性囚徒A在第一阶段的策略考虑。

首先注意到，当博弈重复三次时，C不一定是理性囚徒A在第一阶段的最优选择，因为尽管选择C在第一阶段得到最大收益0（如果B选择D），但暴露出来A是理性的，B在第二阶段就不会选择D，A第二阶段最大收益为-8；选择D，A第一阶段收益-1，第二阶段收益0。

给定B第一阶段选D。若A选D，B的后验概率不变，第二、三阶段就继续选择D、C，A收益为：$(-1)+0+(-8)=-9$；若A选C，暴露理性特征，B第二、三阶段选择C、C，A收益为：$0+(-8)+(-8)=-16$。因$-9>-16$，A没有兴趣偏离右下表的策略。

囚徒B / 囚徒A	坦白（C）	抵赖（D）
坦白（C）	-5, -5	0, -10
抵赖（D）	-10, 0	-1, -1

次数 / 囚徒	$t=1$	$t=2$	$t=3$
非理性囚徒A	D	D	D
理性囚徒A	D	C	C
囚徒B	D	D	C

✓

重复囚徒困境的KMRW声誉模型分析

继续求解：现在考虑囚徒B的策略考虑：(D,D,C), (C,C,C), (C,D,C)

给定理性囚徒A选择(D,C,C)，囚徒B选择(D,D,C)的期望收益为：$-1+[p(-1)+(1-p)(-10)]+[p\times 0+(1-p)(-8)]=17p-19$

如果囚徒B选择(C,C,C)（左下图），期望支付为：$0+(-8)+(-8)=-16$，因此(D,C,C)优于(D,D,C)，因$17p-19\geq -16\Rightarrow p\geq 3/17$，与假设$p\geq 2/9$相容。

如果囚徒B选择(C,D,C)（右下图），期望支付为：$0+[-10]+[p\times 0+(1-p)(-8)]=8p-18$，因此(D,C,C)优于(C,D,C)，因$17p-19\geq 8p-18\Rightarrow p\geq 1/9$，与假设$p\geq 2/9$相容。

所以，当 $p\geq 2/9$ 时，考察的策略组合是精炼纳什均衡的充分条件。

囚徒 \ 次数	$t=1$	$t=2$	$t=3$
非理性囚徒A	D	C	C
理性囚徒A	D	C	C
囚徒B	C	C	C

✗

✗

囚徒 \ 次数	$t=1$	$t=2$	$t=3$
非理性囚徒A	D	C	D
理性囚徒A	D	C	C
囚徒B	C	D	C

重复囚徒困境的KMRW声誉模型分析

- 上述过程证明三阶段囚徒困境的一个精炼纳什均衡为：理性囚徒A在第一阶段选择D，然后在第二、三阶段选择C；囚徒B在第一、二阶段选择D，在第三阶段选择C。
- 可以进一步证明，当$T > 3$时，下列策略组合构成一个精炼贝叶斯均衡：理性囚徒A在$t = 1$至$t = T - 2$阶段选择D，然后在$t = T - 1$和$t = T$阶段选择C；囚徒B在$t = 1$至$t = T - 1$阶段选择D，然后在$t = T$阶段选择C。
- 可以看出，只要 $T > 3$ ，非合作阶段数量为2，与T无关。
- 上述结论对于两个囚徒都存在理性和非理性的私有信息时仍然可以成立，只要满足$T > T^* = (3 - 2p)/7p$。
    - 当$p = 0.1$时， $T^* = 4$；当$p = 0.1$时， $T^* = 9$；。。。
- 说明：只要有一点点的不确定性，就会导致合作行为的发生，只要重复博弈的次数足够大。

KMRW声誉模型的具体内容

KMRW定理：

在T阶段重复囚徒博弈中，如果每个囚徒都有$p > 0$的概率是非理性的（即只选择“针锋相对”或“冷酷策略”），如果T足够大，那么存在$T_0 < T$，是的下列策略组合构成一个精炼贝叶斯均衡：所有理性囚徒在$t \leq T_0$阶段选择合作（抵赖），在$t > T_0$阶段选择不合作（坦白）；并且，非合作阶段的数量$(T - T_0)$只与p有关而与T无关。

说明：1）在每个合作阶段$t \leq T_0$，后验概率$\tilde{p} = p$；在每个非合作阶段，先选C的囚徒暴露自己是理性的，$\tilde{p}(C) = 0$。2）“非合作阶段的数量$(T - T_0)$只与p有关而与T无关”意味着 T只决定合作阶段的数量而不决定合作阶段的数量。

定理证明：使用数学归纳法证明。

KMRW声誉模型的进一步解释

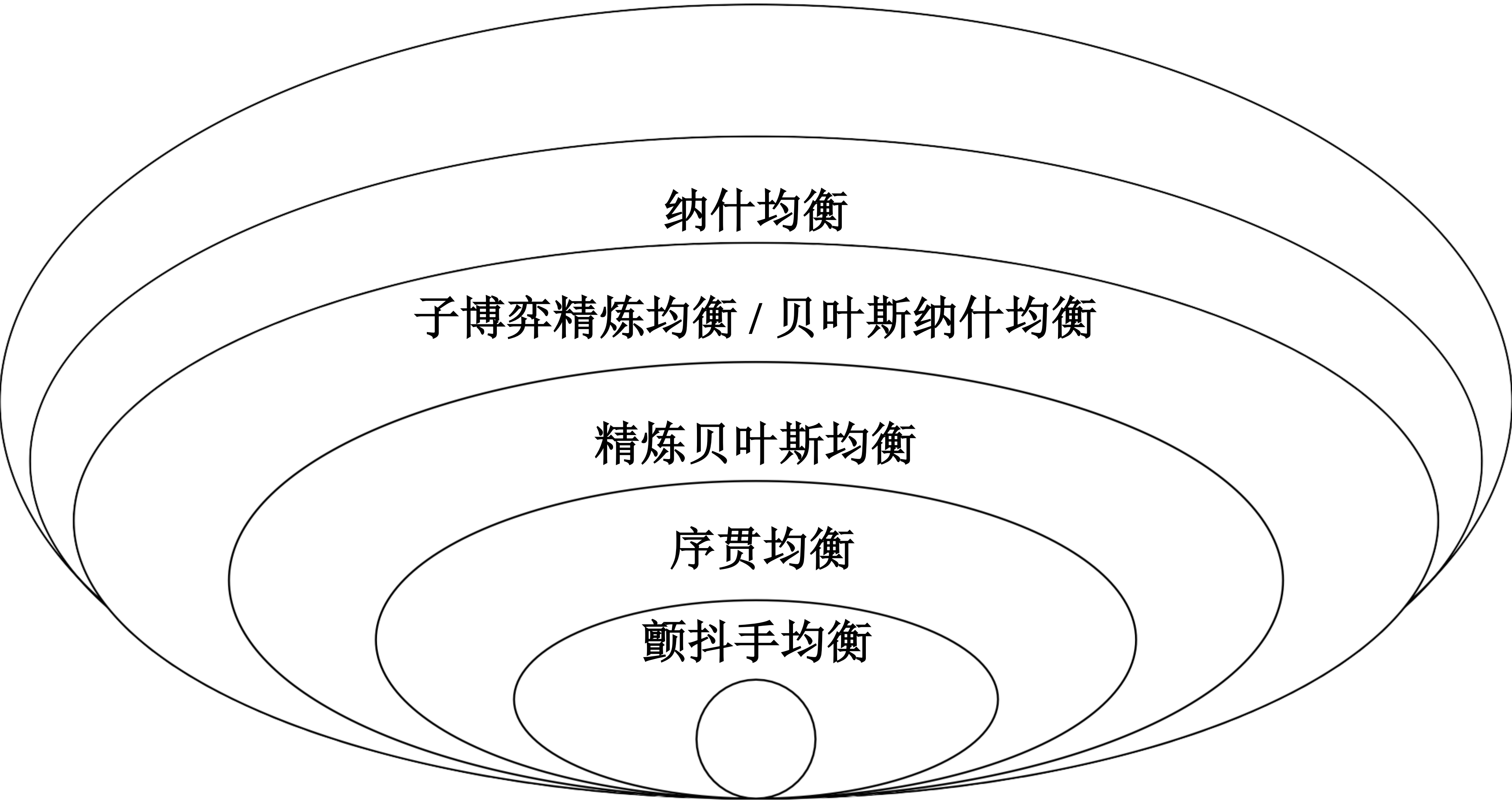
- KMRW定理的直观解释：尽管每一个囚徒在选择合作时冒着被其他囚徒出卖的风险（从而得到一个较低的现阶段支付），但如果选择不合作，就暴露了自己是非合作型的，从而失去了获得长期合作收益的可能（如果对方有可能是合作型的）。
- 如果博弈重复的次数足够多，未来收益的损失就超过了短期被出卖的损失，因此，在博弈一开始，每一个参与人都想树立一个合作形象（使对方认为自己是喜欢合作的），即使他在本性上并不是合作型的；只有在博弈快结束的时候，参与人才会一次把过去积攒的声誉利用尽，合作才会停止（因为此时的短期收益很大而未来损失很小）。

KMRW声誉模型的进一步解释

- KMRW定理为现实中观测到的大量现象提供了有力解释：
  - 连锁店悖论：在位者选择“斗争”类似于囚徒困境中的“抵赖”
  - 中国人推崇的“大智若愚”行为方式；
  - 日常生活中经常捐款的“慈善家”，做好事的“好人”；
  - 大家再举几个例子；
- KMRW定理的最重要之处在于它证明了：只要博弈重复的次数足够多，参与人有足够的耐心（之前分析假定$\delta = 1$，定理对于$\delta < 1$ 也成立，只要δ足够接近于1），即使（有关参与人的私有信息类型的）小小不确定性也可能引起均衡结构的重大改变（均衡的多重性问题）。
- 无限次重复博弈的“无名氏定理”在不完全信息有限重复博弈中也成立，只要博弈重复的次数足够多，参与人有足够的耐心，任何满足个人理性的效益均可是均衡。

博弈论中的纳什均衡小结

- 不同纳什均衡之间的关系



本次课程作业

- 作业内容：寻找存在精炼贝叶斯相关均衡概念的不完全信息动态博弈习题两道，给出习题描述、问题分析、求解过程和最终结果。
  - 评分准则：选取习题的典型性、分析求解过程的完整性以及最终结果的指导性意义。
- 提交时间：2020年11月5日17:00之前
- 提交方法：在课程网站上提交，同时提交作业Word版到助教邮箱（ kangyongxin2015@ia.ac.cn ）
- 邮件发送规范
  - 邮件主题：计算博弈第五次作业_学号_姓名
  - 附件名称：计算博弈第五次作业_学号_姓名.docx

感谢聆听！

兴军亮

Junliang.Xing@ia.ac.cn

2020年10月29日

中国科学院大学
University of Chinese of Academy of Sciences

自动化研究所
CASIA
Institute of Automation

MR垄断限价模型

$$M_1^L - M_1^L(p_1^L) \leq \delta(M_1^L - D_1^L)$$

$$M_1^H - M_1^H(p_1^L) \geq \delta(M_1^H - D_1^H)$$