HW1:博弈论与多臂老虎机算法基础

姓名： 学号： 日期：

1.1. 占优策略均衡与纳什均衡的关系

证明如下关于占优策略均衡与纳什均衡的关系的结论：

1.如果每个参与人i都有一个占优于其它所有策略的策略s_i^*，那么$s^* = (s_1^*, ..., s_n^*)$是纳什均衡；

2.如果每个参与人i都有一个严格占优于其它所有策略的策略s_i^*，那么$s^* = (s_1^*, ..., s_n^*)$是博弈的唯一纳什均衡。

Solution:

1.我们要证明$s^* = (s_1^*, s_2^*, ..., s_n^*)$是一个纳什均衡。根据纳什均衡的定义，我们需要证明对于任何参与人i，他都无法通过单方面改变策略来提高收益。也就是要证明：

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \quad \forall s_i \in S_i$$

根据题设，对于每个参与人i，策略s_i^*是他所有策略中的一个（弱）占优策略。

根据占优策略的定义，对于任何参与人i，他的策略s_i^*满足：无论其他参与人选择什么策略组合s_{-i}，选择s_i^*的收益总是不会低于选择任何其他策略s_i的收益。即：

$$u_i(s_i^*, s_{-i}) \geq u_i(s_i, s_{-i}) \quad \forall s_i \in S_i, \forall s_{-i} \in S_{-i}$$

既然这个条件对所有的对手策略组合s_{-i}都成立，那么它自然也对s_{-i}^*(即其他参与人选择其各自占优策略的组合) 成立。

因此，我们有：

$$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*) \quad \forall s_i \in S_i$$

这个不等式对所有参与人$i \in N$都成立。这完全符合纳什均衡的定义。

因此，由每个参与人的占优策略构成的策略组合s^*是一个纳什均衡。

2.这个证明包含两部分：a)证明s^*是一个纳什均衡；b)证明这个纳什均衡是唯一的。

a)证明s^*是一个纳什均衡

根据严格占优策略的定义，对于任何参与人i，他的策略s_i^*满足：无论其他参与人选择什么策略组合s_{-i}，选择s_i^*的收益总是严格高于选择任何其他不等于s_i^*的策略s_i的收益。即：

$$u_i(s_i^*, s_{-i}) > u_i(s_i, s_{-i}) \quad \forall s_i \in S_{i(s_i \neq s_i^*)}, \forall s_{-i} \in S_{-i}$$

同样，这个条件对所有的对手策略组合s_{-i}都成立，自然也对s_{-i}^*成立。所以：

$$u_i(s_i^*, s_{-i}^*) > u_i(s_i, s_{-i}^*) \quad \forall s_i \in S_{i(s_i \neq s_i^*)}$$

这显然满足纳什均衡的条件$u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$。因此，$s^*$是一个纳什均衡。

b)证明这个纳什均衡是唯一的

我们使用反证法。

假设除了$s^* = (s_1^*, ..., s_n^*)$之外，还存在另一个不同的纳什均衡，我们称之为$s' = (s_1', ..., s_n')$。由于$s' = s^*$，那么必然存在至少一个参与人 j，使得他的策略s_j'与他的严格占优策略s_j^*不同，即$s_j' \neq s_j^*$。

现在我们来考察参与人j的收益。根据严格占优策略s_j^*的定义，对于任意的对手策略组合s_{-j}，选择s_j^*的收益都严格大于选择任何其他策略s_j'的收益。

$$u_j\left(s_j^*, s_{-j}\right) > u_j\left(s_j', s_{-j}\right) \quad \forall s_j' \neq s_j^*, \forall s_{-j} \in S_{-j}$$

这个条件对于对手策略组合s_{-j}'(即在s'均衡中，除j以外的其他人的策略)也成立。所以，我们有：

$$u_j\left(s_j^*, s_{-j}'\right) > u_j\left(s_j', s_{-j}'\right)$$

这个不等式意味着：在其他参与人都选择s_{-j}'的情况下，参与人j如果将自己的策略从s_j'单方面变更为s_j^*，他的收益将会严格增加。

但这与我们最初的假设"s'是一个纳什均衡"矛盾。因为根据纳什均衡的定义，任何参与人都不可能通过单方面改变策略来获得更高的收益。

矛盾，故假设是错误的。

因此，不存在其他任何纳什均衡。由每个参与人的严格占优策略构成的策略组合s'是该博弈的唯一纳什均衡。

证毕！

1.2.N人古诺竞争

假设在古诺竞争中，一共有J家企业。当市场中所有企业总产量为q时，市场价格为$p(q) = a - bq$。且每个企业生产单位产品的成本都是同一个常数c，即企业i的产量为q_i时该企业的成本为$c_i(q_i) = c \cdot q_i$。假设$a > c \geq 0, b > 0$。

1.求纳什均衡下所有企业的总产量以及市场价格；

2.讨论均衡价格随着J变化的情况，你有什么启示？

3.讨论$J \to \infty$的均衡结果，你有什么启示？

Solution:

1. 企业i的利润π_i取决于它自己的产量q_i和所有其它企业的总产量$\sum_{j \neq i} q_j$。

$$\pi_i(q_i, q_{-i}) = p \cdot q_i - c \cdot q_i$$

将价格函数$p(q) = a - b\left(q_i + \sum_{j \neq i} q_j\right)$代入:

$$\pi_i(q_i, q_{-i}) = \left[a - b\left(q_i + \sum_{j \neq i} q_j\right)\right] q_i - cq_i$$

$$\pi_i(q_i, q_{-i}) = aq_i - bq_i^2 - b\left(\sum_{j \neq i} q_j\right) q_i - cq_i$$

古诺模型的核心假设是，每个企业在决定自己的产量时，都将其他企业的产量视为给定值。因此，企业i会选择q_i来最大化其自身利润π_i。我们通过对π_i求关于q_i的一阶导数并令其等于 0 来求解：

$$\frac{\partial \pi_i}{\partial q_i} = a - 2bq_i - b\sum_{j\neq i} q_j - c = 0$$

解得：

$$q_i = \frac{a-c}{2b} - \frac{1}{2}\sum_{j\neq i} q_j$$

由于所有企业都有相同的成本函数，我们可以预期一个对称的纳什均衡，即所有企业都生产相同的产量。令此均衡产量为q^*，则$q_i = q^*$对所有i成立。

在这种情况下，其他企业的总产量为$\sum_{j\neq i} q_j = (J-1)q^*$。将其代入最优反应函数：

$$q^* = \frac{a-c}{2b} - \frac{1}{2}(J-1)q^*$$

解得：

$$q^* = \frac{a-c}{b(J+1)}$$

均衡总产量：

$$Q^* = J \cdot q^* = J \cdot \frac{a-c}{b(J+1)} = \frac{J}{J+1}\frac{a-c}{b}$$

均衡市场价格：

$$P^* = a - bQ^* = \frac{a+cJ}{J+1}$$

2. 对P^*求关于J的导数：

$$\frac{dP^*}{dJ} = \frac{c(J+1)-(a+cJ)}{(J+1)^2} = \frac{c-a}{(J+1)^2}$$

由于$c-a<0, (J+1)^2>0$，因此

$$\frac{dP^*}{dJ} < 0$$

这表明，均衡价格P^*是企业数量J的一个递减函数。

启示：市场中的竞争者越多，市场竞争就越激烈，从而导致市场价格越低。

- 当$J=1$(垄断)时，价格为$P^* = \frac{a+c}{2}$，最高。
- 随着J的增加，每个企业所占的市场份额和市场势力都在减小。为了争夺消费者，企业间的价格竞争压力增大，最终拉低了均衡价格。
- 这揭示了市场结构（企业数量）对市场结果（价格）有决定性的影响。增加市场准入和鼓励竞争是降低价格、提高消费者福利的有效途径。

3.考察当J趋向于无穷大时的极限情况。

$$\lim_{J\to\infty} Q^* = \lim_{J\to\infty}\left(\frac{J}{J+1}\frac{a-c}{b}\right) = \left(\lim_{J\to\infty}\frac{J}{J+1}\right)\left(\frac{a-c}{b}\right) = \frac{a-c}{b}$$

$$\lim_{J\to\infty} P^* = \lim_{J\to\infty}\frac{a+cJ}{J+1} = c$$

启示：当市场中的企业数量趋于无穷多时，市场价格趋近于边际成本c。

- 这正是完全竞争市场的均衡结果。在一个完全竞争的市场中，没有企业拥有市场势力，它们都是价格的接受者，长期的均衡条件就是价格等于边际成本，企业的经济利润为零。
- 它表明古诺模型可以看作是连接垄断和完全竞争的桥梁
  - 当 $J = 1$ 时，模型描述的是垄断。
  - 当 J 为较小的数时，模型描述的是寡头垄断。
  - 当 J 趋于无穷大时，模型的结果趋同于完全竞争。
- 竞争的程度是决定市场效率的关键。

1.3.公地悲剧

假设有I个农场主，每个农场主均有权利在公共草地上放牧奶牛。一头奶牛产奶的数量取决于在草地上放牧的奶牛总量N：当$N < \overline{N}$时，n_i头奶牛产生的收入为$n_i \cdot v(N)$；而当$N \geq \overline{N}$时，$v(N) \equiv 0$。假设每头奶牛的成本为c，且$v(0) > c, v' < 0, v'' < 0$，所有农场主同时决定购买多少奶牛，所有奶牛均会在公共草地上放牧(注：假设奶牛的数量可以是小数，也就是无需考虑取整的问题)。

1.将上述情形表达为策略式博弈；

2.求博弈的纳什均衡下所有农场主购买的总奶牛数(可以保留表达式的形式，不用求出具体解)；

3.求所有农场主效用之和最大(社会最优)情况下的总奶牛数(可以保留表达式的形式，不用求出具体解)，与上一问的结果比较，你能从中得到什么启示？

Solution:

1.将上述情形表达为策略式博弈

- **Players**:I个农场主，集合$P = \{1, 2, ..., I\}$
- **Strategies**:每个农场主i的策略是选择一个非负的奶牛数量n_i行放牧。因此，农场主i的策略空间为$S_i = [0, \infty)$。一个策略组合就是所有农场主选择的数量向量$(n_1, n_2, ..., n_I)$。
- **Payoffs**:每个农场主i的收益π_i是其所有奶牛带来的总收入减去总成本。
  - 总放牧数量 $N = \sum_{j=1}^{I} n_j$
  - 每头牛的收入（或价值）为$v(N)$（分段）
  - 每头牛的成本为c

因此，农场主i的收益函数$\pi_i(n_i, n_{-i})$为：

$$\pi_i(n_1, ..., n_I) = \begin{cases} n_i \cdot v\left(\sum_{j=1}^{I} n_j\right) - c \cdot n_j & \text{, if} \sum_{j=1}^{I} n_j < \overline{N} \\ -c \cdot n_i & \text{, if} \sum_{j=1}^{I} n_j \geq \overline{N} \end{cases}$$

其中n_{-i}代表除农场主i之外其他所有农场主选择的奶牛数量。

2.在纳什均衡中，每个农场主都选择自己的最优奶牛数，以在给定其他农场主选择的情况下最大化自己的收益。我们假设存在一个$N < \overline{N}$的内部解。

农场主i的优化问题是：

$$\max_{n_i \geq 0} \pi_i = n_i \cdot v(n_i + N_{-i}) - c \cdot n_i$$

其中$N_{-i} = \sum_{j \neq i} n_j$是其他农场主的奶牛总数，被农场主$i$视为常数。

为了找到最优解，我们对π_i求关于n_i的一阶导数，并令其等于零：

$$\frac{\partial \pi_i}{\partial n_i} = \frac{\partial}{\partial n_i}[n_i \cdot v(n_i + N_{-i}) - cn_i] = 0$$

得到：

$$1 \cdot v(n_i + N_{-i}) + n_i \cdot v'^{(n_i + N_{-i})} \cdot 1 - c = 0$$

将$N = n_i + N_{-i}$代回，得到：

$$v(N) + n_i v'(N) - c = 0$$

这是一个农场主i的最优反应必须满足的条件。由于所有农场主都是同质的（成本相同），我们可以寻找一个对称的纳什均衡，即所有农场主都选择相同的数量$n_i = n^{\text{NE}}$

在这种情况下，总数量$N^{\text{NE}} = I \cdot n^{\text{NE}}$，或者说$n^{\text{NE}} = \frac{N^{\text{NE}}}{I}$。将此代入一阶条件中：

$$v(N^{\text{NE}}) + \frac{N^{\text{NE}}}{I} v'(N^{\text{NE}}) - c = 0$$

该等式隐式地定义了纳什均衡下的总奶牛数N^{NE}

3.社会最优指的是所有农场主的总效用（或称社会总福利）最大化。一个社会计划者的目标是选择一个总数量N来最大化所有人的收益之和。

社会总福利W为：

$$W(N) = \sum_{i=1}^{I} \pi_i = \sum_{i=1}^{I} (n_i v(N) - cn_i) = \left(\sum_{i=1}^{I} n_i\right) v(N) - c\left(\sum_{i=1}^{I} n_i\right)$$

$$W(N) = N \cdot v(N) - cN$$

为了找到社会最优的总奶牛数N^{SO},我们对$W(N)$求关于N的一阶导数，并令其等于零：

$$\frac{dW}{dN} = \frac{d}{dN}[N \cdot v(N) - cN] = 0$$

得到:

$$1 \cdot v(N) + N \cdot v'(N) - c = 0$$

该等式隐式地定义了社会最优的总奶牛数N^{SO}

$$v\left(N^{\text{SO}}\right) + N^{\text{SO}} v'\left(N^{\text{SO}}\right) - c = 0$$

Insight:

这里的比较结果取决于$v'(N)$的符号，这代表了外部性的方向。

- $v'(N) < 0$:这是最典型的情况，即过度放牧导致草地退化，每头牛的产出下降（负外部性）。
  - NE 条件: $v(N^{\text{NE}}) - \frac{N^{\text{NE}}}{I} |v'(N^{\text{NE}}| = c$
  - SO 条件: $v\left(N^{\text{SO}}\right) - N^{\text{SO}} \left|v'\left(N^{\text{SO}}\right)\right| = c$
  - 在社会最优水平N^{SO}，个体农场主发现他的边际收益$v\left(N^{\text{SO}}\right) - \frac{N^{\text{SO}}}{I} \left|v'\left(N^{\text{SO}}\right)\right| > c$，因为他只承担了$\frac{1}{I}$的负外部性成本。因此他有动机继续增加奶牛，导致最终的均衡数量$N^{\text{NE}} > N^{\text{SO}}$
  - 当存在负外部性且资源产权不清晰时，个体的理性决策会导致对公共资源的过度使用，造成社会总福利的损失。这就是“公地悲剧”的核心。
- $v'(N) > 0$:这是一种不寻常但有趣的设定，意味着牛群有集聚效应，牛越多，每头牛的产出越高（正外部性），直到某个临界点$\overline{N}$。
  - NE 条件：$v(N^{\text{NE}}) + \frac{N^{\text{NE}}}{I} v'(N^{\text{NE}}) = c$

- SO 条件：$v(N^{\mathrm{SO}}) + N^{\mathrm{SO}} v'(N^{\mathrm{SO}}) = c$
- 在社会最优水平N^{SO}，个体农场主发现他的边际收益$v(N^{\mathrm{SO}}) + \frac{N^{\mathrm{SO}}}{I} V'(N^{\mathrm{SO}}) < c$，因为他只获得了$\frac{1}{I}$的正外部性收益。由于边际收益小于边际成本，他没有动机把牛增加到社会最优水平。这导致最终的均衡数量$N^{\mathrm{NE}} < N^{\mathrm{SO}}$
- 当存在正外部性时，个体决策者因为无法获得其行为产生的全部社会收益，会导致对该行为的供给不足。这也是一种市场失灵，可以称之为“公地喜剧”或“反公地悲剧”。

Appendix:我们运用实分析方法和泛函分析中的思想，对解的存在性、唯一性进行分析，并给出一个形式解的表示。

首先，把两个核心方程写成$F(N) = 0$的形式。我们求解的目标就是找到这两个函数的根N^*。

- NE:$F_{\mathrm{NE}}(N) \equiv v(N) + \frac{N}{I} v'(N) - c = 0$
- SO:$F_{\mathrm{SO}} \equiv v(N) + N v'(N) - c = 0$

注意，社会最优的函数$F_{\mathrm{SO}}(N)$其实是社会总福利函数$W(N) = N v(N) - cN$的一阶导数，即$F_{\mathrm{SO}} = W'(N)$。同样，$F_{\mathrm{NE}}(N)$也与个体利润函数的一阶导数密切相关。

我们可以使用介值定理来证明解的存在性。

考察$N \to 0$的情况：

$$F_{\mathrm{SO}}(0) = v(0) + 0 \cdot v'(0) - c = v(0) - c$$

$$F_{\mathrm{NE}}(0) = v(0) + 0 \cdot v'(0) - c = v(0) - c$$

根据题目假设$v(0) > c$，所以在这两种情况下，当$N = 0$时函数值都大于 0.

考察$N > 0$的情况： 为了保证有解，我们需要函数在定义域$\left(0, \overline{N}\right)$内的某一点变为负值。例如，我们假设当$N$趋近于草地承载上限$\overline{N}$时，放牧的总收益接近于零，即$v\left(\overline{N}\right) \approx 0$。如果此时$v'\left(\overline{N}\right)$不是一个巨大的正数，那么$F\left(\overline{N}\right) \approx -c < 0$。

解的唯一性取决于函数$F(N)$的严格单调性。如果一个函数是严格单调的（始终递增或始终递减），它最多只会与x轴相交一次，因此根是唯一的。我们通过分析$F(N)$的导数来判断其单调性。

社会最优情况：

$$F'_{\mathrm{SO}} = \frac{d}{dN}[v(N) + N v'(N) - c] = v'(N) + [1 \cdot v'(N) + N \cdot v''(N)] = 2v'(N) + N v''(N)$$

纳什均衡情况：

$$F'_{\mathrm{NE}} = \frac{d}{dN}\left[v(N) + \frac{N}{I} v'(N) - c\right] = v'(N) + \left[\frac{1}{I} v'(N) + \frac{N}{I} v''(N)\right] = \left(1 + \frac{1}{I}\right) v'(N) + \frac{N}{I} v''(N)$$

题设条件为$v'(N) > 0$和$v'' < 0$。在两种情况下，导数的符号都是不确定的（一个正项和一个负项之和）。这意味着仅凭题目给出的条件，我们无法保证解的唯一性。$F(N)$可能不是单调函数，这意味着可能存在多个N足一阶条件，即可能存在多个均衡点或最优点。

在标准的微观经济学模型中，为了保证解的唯一性，通常会施加一个更强的凹性条件。

对于社会最优，要求社会福利函数$W(N)$是严格凹函数，即$W''(N) < 0$。由上可知，$W''(N) = 2v'(N) + N v''(N)$。因此，唯一性需要满足$2v'(N) + N v''(N) < 0$。

对于纳什均衡，要求个体利润函数π_i对n_i是严格凹函数，即$\frac{\partial^2 \pi_i}{\partial n_i^2} < 0$。计算可得$\frac{\partial^2 \pi_i}{\partial n_i^2} = 2v'(N) + n_i v''(N)$。唯一性需要满足$2v'(N) + n_i v''(N) < 0$。

如果补充上这些标准凹性条件，那么$F(N)$就是严格单调递减函数，其根N^*必然是唯一的。

我们可以将求解方程 $F(N) = 0$的问题，转化为寻找一个算子的不动点的问题。

社会最优(SO):

$$v(N) + Nv'(N) - c = 0 \Rightarrow Nv'(N) = c - v(N)$$

如果$v'(N) \neq 0$，则$N = \frac{c-v(N)}{v'(N)}$ 我们可以定义一个算子T_{SO}:

$$T_{\mathrm{SO}} = \frac{c - v(N)}{v'(N)}$$

社会最优解N^{SO}是算子T_{SO}的不动点，即满足$N^{\mathrm{SO}} = T_{\mathrm{SO}}(N^{\mathrm{SO}})$

纳什均衡(NE):

$$v(N) + \frac{N}{I}v'(N) - c = 0 \Rightarrow \frac{N}{I}v'(N) = c - v(N)$$

如果$v'(N) \neq 0$，则$N = \frac{I(c-v(N))}{v'(N)}$ 我们可以定义一个算子T_{NE}:

$$T_{\mathrm{NE}}(N) = \frac{I(c - v(N))}{v'(N)}$$

纳什均衡N^{NE}是算子T_{NE}的不动点，即满足$N^{\mathrm{NE}} = T_{\mathrm{NE}}(N^{\mathrm{NE}})$

这种不动点的表达方式就是一种形式解。它没有给出具体的数值，但它精确地刻画了"解"这个对象的数学结构。

根据巴拿赫不动点定理，如果一个算子T是在一个完备度量空间上的一个压缩映射，那么它有且仅有一个不动点。验证一个算子是否为压缩映射，需要证明$|' (N)| < 1$。这同样需要知道$v(N)$的具体形式或者对其导数有更强的约束。

1.4.教育作为一种信号

本练习题表明，大学教育除了扩展学生的知识之外，还向未来的雇主传递了一种形式的信号。一个去求 职的年轻人可能是高能力的，也可能是低能力的。假设 1/4 的高中毕业生是高能力的，剩下的是低能 力的。一个新近的高中毕业生，他知道自己是否是高能力的，在申请一份工作之前，可以选择去国外旅 行一年或者进入大学（假定他无法同时做这两件事）。雇主希望雇用一个人填充空缺，但他不知道工作 申请者是否是高能力的；他所知道的仅仅是这个申请者是进入了大学还是去国外旅行了。雇主从雇用 一个工人中所得到的收益，只取决于被雇用工人的能力（而不取决于他的教育水平），而这个年轻人的 收益取决于高中毕业后做了什么、他的才能（因为高能力的学生更享受学习的过程）以及他是否得到了 这份工作。下面的表格描述了这些收益:

		年轻人	
		国外旅行	进入大学
雇主	雇用	0，6	0，2
	不雇用	3，3	3，−3

年轻人不是高能力的收益矩阵

		年轻人	
		国外旅行	进入大学
雇主	雇用	8，6	8，4
	不雇用	3，3	3，1

年轻人是高能力的收益矩阵

1.将上述博弈表达为一个不完全信息博弈;

2.求出博弈所有的贝叶斯纳什均衡。

Solution:

1.我们可以使用信号博弈的框架来描述这个情形，该框架包含参与人、类型、行动、信念和收益。

- Players:
  - 发送方 (Sender, S): 年轻人（求职者）。
  - 接收方 (Receiver, R): 雇主。
- Timing:
  1. “自然”（Nature）决定年轻人的类型 ϑ（即能力），并只有年轻人自己知道。
  2. 年轻人（发送方）根据自己的类型，选择一个行动（发送一个信号）
  3. 雇主（接收方）观察到年轻人的行动，但不知道其真实类型。雇主根据观察到的信号更新自己的信念，并决定是否雇佣
  4. 双方根据年轻人的类型、年轻人的行动和雇主的行动获得各自的收益。
- Types:
  - 年轻人的类型空间为$\theta \in \{高能力, 低能力\}$
  - 根据题设，自然选择类型的先验概率为：$P(\theta = 高能力) = \frac{1}{4}, P(\theta = 低能力) = \frac{3}{4}$
- Actions:
  - 年轻人的行动空间为$M = \{进入大学, 国外旅行\}$
  - 雇主的行动空间为$A = \{雇用, 不雇用\}$
- Beliefs:
  - 雇主的信念是指在观察到特定信号后，对年轻人是高能力的概率判断。我们用$\mu(高能力|行动)$
- Payoffs:

如果年轻人是低能力：

雇主	国外旅行	进入大学
雇用	(0,6)	(0,2)
不雇用	(3,3)	(3,-3)

如果年轻人是高能力：

雇主	国外旅行	进入大学
雇用	(8,6)	(8,4)
不雇用	(3,3)	(3,1)

2.首先，我们来确定雇主的最优策略。雇主的决策取决于他对求职者是高能力的信念μ。

雇主选择“雇用”的期望收益为：

$$\mathbb{E}[U(雇用)] = \mu \cdot 8 + (1 - \mu) \cdot 0 = 8\mu$$

雇主选择“不雇用”的期望收益为：

$$\mathbb{E}[U(不雇用)] = \mu \cdot 3 + (1 - \mu) \cdot 3 = 3$$

雇主会选择雇用，当且仅当$\mathbb{E}[U(雇用)] \geq \mathbb{E}[U(不雇用)]$，即

$$8\mu \geq 3 \Rightarrow \mu \geq \frac{3}{8}$$

所以，雇主的决策规则是：

- 如果信念$\mu \geq \frac{3}{8}$，则雇用
- 如果信念$\mu < \frac{3}{8}$，则不雇用

现在我们寻找两种类型的均衡：分离均衡和混同均衡。

类型一：分离均衡

在分离均衡中，不同类型的年轻人会选择不同的行动。

假设：高能力者选择进入大学，低能力者选择国外旅行

1. 信念：在这个均衡路径上，雇主观察到“进入大学”便知其为高能力者，观察到“国外旅行”便知其为低能力者。

$$\mu(\text{高能力}|\text{ 进入大学}) = 1$$

$$\mu(\text{高能力}|\text{ 国外旅行}) = 0$$

2. 雇主行动:
   - 观察到“进入大学”：$\mu = 1 > \frac{3}{8}$，雇主选择雇用。
   - 观察到“国外旅行”：$\mu = 0 < \frac{3}{8}$，雇主选择不雇用。
3. 年轻人动机检验:
   - 高能力者：选择进入大学(均衡路径)，被雇用，收益为 4.若偏离至国外旅行，不被雇用，收益为 3.高能力者不会偏离。
   - 低能力者：选择国外旅行(均衡路径)，不被雇用，收益为 3。若偏离至进入大学，收益为 2.低能力者不会偏离。

这是一个有效的精炼贝叶斯纳什均衡。

策略：高能力年轻人选择进入大学，低能力年轻人选择国外旅行。雇主雇用上大学的人，不雇用去旅行的人。

信念：雇主相信上大学的人是高能力的($\mu = 1$)，去旅行的人是低能力的($\mu = 0$)。

类型二：混同均衡

在混同均衡中，所有类型的年轻人选择相同的行动。

假设：所有年轻人都选择国外旅行

1. 信念：均衡路径上: 雇主观察到“国外旅行”，但他无法分辨求职者类型，因此他的信念等于先验概率：$\mu(\text{高能力}|\text{国外旅行}) = P(\text{高能力}) = \frac{1}{4}$.非均衡路径上: 雇主观察到“进入大学”。由于这在均衡中不会发生，贝叶斯法则不适用。我们可以设定一个非均衡路径信念$\mu_U = \mu(\text{高能力}|\text{进入大学})$
2. 雇主行动:观察到“国外旅行”：$\mu = \frac{1}{4} < \frac{3}{8}$，雇主选择不雇用。观察到“进入大学”：行动取决于$\mu_U$。为了维持均衡，我们需要设定一个让年轻人不想偏离的信念。最能阻止偏离的信念是让雇主不雇用，即要求$\mu_U < \frac{3}{8}$。
3. 年轻人动机检验: 假设非均衡路径信念为$\mu_U = 0$，则雇主在任何情况下都不会雇用。
   - 高能力者:选择国外旅行（均衡路径），不被雇用，收益为 3。若偏离至进入大学，$\mu_U = 0$，同样不被雇用，收益为 1。高能力者不会偏离。
   - 低能力者:选择国外旅行（均衡路径），不被雇用，收益为 3。若偏离至进入大学，$\mu_U = 0$，同样不被雇用，收益为 -3。低能力者不会偏离。

这是一个有效的精炼贝叶斯纳什均衡。

策略：所有年轻人（无论高能力或低能力）都选择国外旅行。雇主不雇用任何求职者（无论其是旅行还是上大学）。

信念：雇主相信去旅行的人是高能力的概率为$\frac{1}{4}$。雇主相信上大学的人是高能力的概率$\mu_U < \frac{3}{8}$（非均衡路径信念）。

综上所述，该博弈存在两个精炼贝叶斯纳什均衡:
1. 一个分离均衡: 高能力者通过上大学来彰显自己，并成功获得工作；低能力者则选择旅行。教育在这里成功地起到了信号作用。

2. 一个混同均衡：所有人都选择旅行，雇主因为无法分辨好坏而选择不雇用任何人。教育在此未能起到信号作用，导致了一个对高能力者和雇主都不利的结果。

1.5.混合策略的不完全信息解释

考虑以下抓钱博弈（**grab the dollar**）：桌子上放 1 块钱，桌子的两边坐着两个参与人，如果两人同时去抓钱，每人罚款 1 块；如果只有一人去抓，抓的人得到那块钱；如果没有人去抓，谁也得不到什么。 因此，每个参与人的策略是决定抓还是不抓。

抓钱博弈描述的是下述现实情况：一个市场上只能有一个企业生存，有两个企业在同时决定是否进入。 如果两个企业都选择进入，各亏损 100 万；如果只有一个企业进入，进入者盈利 100 万；如果没有企 业进入，每个企业既不亏也不盈。

1. 求抓钱博弈的纯策略纳什均衡；

	抓	不抓
抓	−1,-1	1,0
不抓	0,1	0,0

2. 求抓钱博弈的混合策略纳什均衡；

现在考虑同样的博弈但具有如下不完全信息：如果参与人i赢了，他的利润是$1+\theta_i$(而不是 1). 这里θ_i是参与人的类型，参与人i自己知道θ_i，但另一个参与人不知道。假定θ_i在$[-\varepsilon,\varepsilon]$区间上均匀分布。

	抓	不抓
抓	−1,-1	$1+\theta_1$,0
不抓	0,$1+\theta_2$	0,0

由于两个参与人的情况完全对称，故考虑如下对称贝叶斯纳什均衡（两个人的策略相同）形式：参与人$i(i=1,2)$的策略均为

$$s_{i(\theta_i)}=\begin{cases}\text{抓},\text{如果}\theta_i\geq\theta^*\\\text{不抓},\text{如果}\theta_i<\theta^*\end{cases}$$

即θ^*是两个参与人抓或不抓的类型分界阈值，其中θ^*是一个待计算确定的参数。

3.求θ^*;

4.当$\varepsilon\to0$时，上述贝叶斯纳什均衡会收敛于什么？从中你能得到怎样的启示。

Solution:

1.我们分析给定的收益矩阵：

参与者 1/参与者 2	抓	不抓
抓	−1,-1	1,0
不抓	0,1	0,0

纳什均衡是指在给定对方策略的情况下，没有任何一方有动机单方面改变自己的策略。我们逐一检查四个策略组合：

- (抓，抓)：收益为(-1, −1)。此时，如果参与人 1 单方面改为“不抓”，他的收益将从-1 变为 0。因为有动机改变策略，所以这不是一个纳什均衡。

- (抓，不抓)：收益为$(1,0)$。对参与人 1：坚持“抓”得到 1，改为“不抓”得到 0。他没有动机改变。对参与人 2：坚持“不抓”得到 0，改为“抓”得到-1。他没有动机改变。双方都没有动机改变，因此 (抓, 不抓) 是一个纯策略纳什均衡。
- (不抓，抓)：收益为$(0,1)$。与上一种情况对称。对参与人 1：坚持“不抓”得到 0，改为“抓”得到-1。他没有动机改变。对参与人 2：坚持“抓”得到 1，改为“不抓”得到 0。他没有动机改变。双方都没有动机改变，因此 (不抓, 抓) 是一个纯策略纳什均衡。
- (不抓，不抓)：收益为$(0,0)$。此时，如果参与人 1 单方面改为“抓”，他的收益将从 0 变为 1。因为有动机改变策略，所以这不是一个纳什均衡。

因此，该博弈有两个纯策略纳什均衡，分别是 (抓, 不抓) 和 (不抓, 抓)。

2.在混合策略均衡中，每个参与人选择不同策略的概率使得对手对于其纯策略的选择是无差异的。

设参与人 1 选择“抓”的概率为p，选择"不抓"的概率为$1-p$；设参与人 2 选择"抓"的概率为q，选择"不抓"的概率为$1-q$。

为了让参与人 1 对“抓”和“不抓”无差异，这两个策略的期望收益必须相等：

$$\mathbb{E}_1(\text{抓}) = \mathbb{E}_1(\text{不抓})$$

即：

$$q \cdot (-1) + (1-q) \cdot (1) = q \cdot (0) + (1-q) \cdot (0)$$

解得：

$$q = \frac{1}{2}$$

因此，该博弈存在一个混合策略纳什均衡，即双方都以 $\frac{1}{2}$ 的概率选择“抓”。

3. 现在我们进入不完全信息博弈。我们寻找一个对称的贝叶斯纳什均衡，其形式为：当类型$\theta_i \geq \theta^*$时选择“抓”，当$\theta_i < \theta^*$时选择"不抓"。在临界点$\theta_i = \theta^*$时，参与人i对于“抓”和“不抓”这两个选择的期望收益是无差异的。我们以参与人 1 为例，假设他的类型$\theta_1 = \theta^*$。

他选择"抓"的期望收益为：

$$\mathbb{E}_1(\text{抓}|\ \theta_1 = \theta^*) = P(\text{参与人 2 抓}) \cdot (-1) + P(\text{参与人 2 不抓}) \cdot (1 + \theta_1)$$

他选择"不抓"的期望收益为：

$$\mathbb{E}_1(\text{抓}|\ \theta_1 = \theta^*) = P(\text{参与人 2 抓}) \cdot (0) + P(\text{参与人 2 不抓}) \cdot (0) = 0$$

根据参与人 2 的策略，他选择“抓”的条件是$\theta_2 \geq \theta^*$。 由于θ_2在$[-\varepsilon, \varepsilon]$上均匀分布，我们可以计算这个概率：

$$P(\text{参与人 2 抓}) = P(\theta_2 \geq \theta^*) = \frac{\varepsilon - \theta^*}{\varepsilon - (-\varepsilon)} = \frac{\varepsilon - \theta^*}{2\varepsilon}$$

$$P(\text{参与人 2 不抓}) = P(\theta_2 < \theta^*) = 1 - \frac{\varepsilon - \theta^*}{2\varepsilon} = \frac{\varepsilon + \theta^*}{2\varepsilon}$$

在$\theta_1 = \theta^*$时，令两种选择的期望收益相等：

$$\mathbb{E}_1(\text{抓}|\ \theta_1 = \theta^*) = \mathbb{E}_1(\text{不抓}|\ \theta_1 = \theta^*) = 0$$

$$\left(\frac{\varepsilon - \theta^*}{2\varepsilon}\right) \cdot (-1) + \left(\frac{\varepsilon + \theta^*}{2\varepsilon}\right) \cdot (1 + \theta^*) = 0$$

化简得到：

$$\theta^*(\theta^* + 2 + \varepsilon) = 0$$

解得:

$$\theta^* = 0 \quad 或 \quad \theta^* = -(2 + \varepsilon)$$

由于类型θ^*必须在分布区间$[-\varepsilon, \varepsilon]$之内，而$\varepsilon > 0$使得$-(2+\varepsilon) < -\varepsilon$，所以$\theta^* = -(2+\varepsilon)$不是一个有效的解。

因此，$\theta^* = 0$是唯一有效阈值。

4.我们已经求得阈值$\theta^* = 0$，这个值并不依赖于ε。现在我们来看当$\varepsilon \to 0$时，这个贝叶斯纳什均衡策略本身会收敛于什么。

在贝叶斯纳什均衡中，任何一个参与人i选择“抓”的（事前）概率为：

$$P(抓) = P(\theta_i \geq \theta^*) = P(\theta_i \geq 0)$$

由于θ_i在$[-\varepsilon, \varepsilon]$上均匀分布，这个概率是：

$$P(\theta_i \geq 0) = \frac{\varepsilon - 0}{\varepsilon - (-\varepsilon)} = \frac{\varepsilon}{2\varepsilon} = \frac{1}{2}$$

这个结果对所有$\varepsilon > 0$都成立。因此，当$\varepsilon \to 0$时，每个参与人选择“抓”的概率仍然是$\frac{1}{2}$。

因此，当$\varepsilon \to 0$时，这个不完全信息博弈的纯策略贝叶斯纳什均衡，收敛到了原完全信息博弈的混合策略纳什均衡。

Insight:
1. 混合策略的现实基础: 混合策略（即以一定概率随机选择行动）在现实中似乎不合常理。没人会真的通过掷硬币来做商业决策。海萨尼的理论表明，我们观察到的“随机”行为，可能并非真正的随机化，而是参与人在应对自身微小、私有的不确定性（即他们的“类型”，如θ_i）时，采取的确定性（纯策略）行为的结果。
2. 不完全信息的威力: 一个完全信息博弈中的混合策略均衡，可以被看作是一个引入了极少量不确定性（即$\varepsilon \to 0$）的不完全信息博弈的纯策略均衡的极限。
3. 从外部观察者视角: 对于一个不了解参与人私有信息（类型θ_i）的外部观察者来说，他看到参与人有时“抓”、有时“不抓”的行为，会表现为一种概率为 1/2 的随机行为。但实际上，每个参与人的决策都是基于其私有信息的理性最优选择。

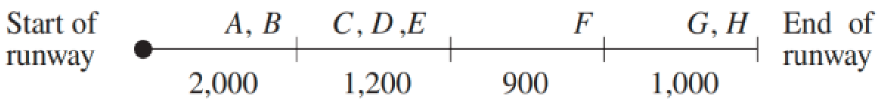
1.6.飞机跑道成本分配的沙普利置计算

机场跑道的维护费用通常是向在那个机场降落飞机的航空公司来收取的。但是轻型飞机所需的跑道长 度比重型飞机所需的跑道长度短，这就带来了一个问题，如何在拥有不同类型飞机的航空公司之间确 定公平的维护费用分摊。

定义一个成本博弈(即每个联盟的效用是成本$(N;c)$，这里N是降落在这个机场上的所有飞机的集合，$c(S)$(对于每个联盟S)是能够允许联盟中所有飞机降落的最短跑道的维护费用。)如果用沙普利值来确定 费用的分摊，证明：每段跑道的维护费用由使用那段跑道的飞机均摊。

下图描绘了一个例子，其中标号为A, B, C, D, E, F, G和H的八架飞机每天都要在这个机场降落。每 架飞机所需的跑道的整个长度由图中的区间来表示。例如，飞机F需要前三个跑道区间。每个跑道区间的每周维护费用标示在图的下面。例如，$c(A, D, E) = 3200, c(A) = 2000$和$c(C, F, G) = 5100$。在这一 例子中，$A$的沙普利值恰好等于 $2000/8 = 250$，而F的沙普利值等

于 $2000/8+1200/6+900/3=750$。你的任务是将这一性质推广到一般的情形下给出证明（提示：使用沙普利值的性质和公式的特点）。



Solution:

首先，我们将这个问题转化为一个合作博弈模型。

Players:博弈的参与人是所有使用该机场的飞机。设所有飞机的集合为$N=\{1,2,...,n\}$，其中n是飞机的总数。

Runway and Costs:假设跑道根据不同飞机的需求被逻辑上分成了m个分段。

- 设第k段跑道的建造成本或维护成本为c_k
- 一架飞机如果需要长度为L的跑道，它就需要使用所有累积长度小于等于L的跑道分段
- 我们定义U_k为所有需要使用第k段跑道（以及所有$j<k$的分段）的飞机的集合。换句话说，集合U_k中的飞机所要求的跑道长度，至少是前k段的总长度。

Cost Function:对于任何一个飞机组成的联盟（子集）$S\subset N$，其成本函数$C(S)$被定义为满足该联盟中所有飞机起降所需的最短跑道的总成本。如果联盟S为空，则$C(S)=0$。例如，如果联盟S中对跑道长度要求最高的飞机需要使用到第p段跑道，那么$C(S)=\sum_{k=1}^{p} c_k$

我们的目标是计算每架飞机i在这个合作博弈中的沙普利值$\varphi_i(C)$。普利值是一种被广泛认可的公平成本（或收益）分配方案。

根据机场跑道成本的结构，我们可以将总成本函数$C(S)$分解为m个独立的子博弈，每个子博弈对应一段跑道的成本。

我们为每段跑道$k(k=1,...,m)$定义一个子博弈成本函数$v_k(S)$:

- 如果联盟S中至少有一架飞机需要使用第k段跑道(即$S\cap U_k\neq\Phi$)，那么该联盟就需要承担第k段跑道的成本，即$v_k(S)=c_k$
- 如果联盟S中没有飞机需要使用第k段跑道(即$S\cap U_k=\Phi$)，那么成本为 0，即$v_k(S)=0$

我们可以证明，总成本函数$C(S)$正是所有这些子博弈成本函数之和：

$$C(S)=\sum_{k=1}^{m} v_k(S)$$

这是因为，如果联盟S中要求最长跑道的飞机需要用到第p段，那么对于所有$k\leq p$，都有$S\cap U_k\neq\Phi$，因此$v_k(S)=c_k$；而对于所有$k>p$，都有$S\cap U_k=\Phi$，因此$v_k(S)=0$。所以，$\sum v_{k(S)}=\sum_{k=1}^{p} c_k=C(S)$

沙普利值一个非常重要的性质是可加性。如果一个博弈可以分解为多个子博弈的和，那么任何一个参与人的总沙普利值，就等于其在每个子博弈中的沙普利值之和。因此，飞机i的总成本分摊为：

$$\varphi_i(C)=\varphi_i\left(\sum_{k=1}^{m} v_k\right)=\sum_{k=1}^{m}\varphi_i(v_k)$$

现在问题转化为计算飞机i在每个子博弈v_k中的沙普利值$\varphi_i(v_k)$。

对于一个特定的子博弈v_k（对应第k段跑道，成本为c_k），分两种情况讨论：

- 情况 1：飞机i不使用第k段跑道$(i\notin U_k)$

在这种情况下，飞机i的加入与否，不会影响任何联盟是否需要承担c_k的成本。也就是说，对于任何不包含i的联盟S，都有$v_k(S \cup \{i\}) = v_k(S)$。因此，飞机$i$对任何联盟的边际贡献都是 0. 所以，它的沙普利值也为 0:

$$\varphi_i(v_k) = 0 \quad , \text{ if } i \notin U_k$$

- 情况 2：飞机i使用第k段跑道($i \in U_k$)

在这种情况下，所有同样使用第k段跑道的飞机（即所有在U_k中的飞机）在这个子博弈中是对称的。因为子博弈v_k只关心联盟中是否至少有一个U_k的成员，而不关心具体是哪一个或哪几个。根据沙普利值的对称性，所有在U_k中的飞机在子博弈v_k中应该分摊相同的成本。 同时，根据沙普利值的有效性，一个博弈中所有参与人的沙普利值之和必须等于该博弈的总成本，即$\sum_{j \in N} \varphi_j(v_k) = v_k(N)$。

由于U_k非空（否则成本c_k无意义），所以$v_k(N) = c_k$。结合情况 1，我们知道只有U_k中的飞机才有非零的沙普利值。因此：

$$\sum_{j \in U_k} \varphi_j(v_k) = c_k$$

因为所有$j \in U_k$的飞机都是对称的，所以它们的沙普利值相等。设$|U_k|$为集合U_k中飞机的数量，则对于任意$i \in U_k$:

$$|U_k| \cdot \varphi_i(v_k) = c_k$$

解得：

$$\varphi_i(v_k) = \frac{c_k}{|U_k|} \quad , \text{ if } i \in U_k$$

这表明，第k段跑道的成本c_k，由所有使用它的飞机（即U_k中的飞机）平均分摊。

最后，我们将每个子博弈的结果代入第二步的公式中，得到飞机i的总成本分摊$\varphi_i(C)$:

$$\varphi_i(C) = \sum_{k=1}^{m} \varphi_i(v_k) = \sum_{k \text{ s.t. } i \in U_k} \frac{c_k}{|U_k|}$$

这个公式的文字表述是：任何一架飞机i的公平待摊费用，等于它所使用的每一段跑道的“均摊成本”之和，而每一段跑道的“均摊成本”正是该段跑道的成本除以使用该段跑道的飞机总数。

应用于题目中的例子

- 飞机 F 的成本分摊φ_F:飞机 F 需要 4,100 的跑道，因此它使用了成本为 2000、1200、900 的三个分段。
  1. 第一段 (成本 2000): 所有 8 架飞机都使用，均摊成本为 2000/8。
  2. 第二段 (成本 1200): C,D,E,F,G,H 共 6 架飞机使用，均摊成本为 1200/6。
  3. 第三段 (成本 900): F,G,H 共 3 架飞机使用，均摊成本为 900/3。
  4. 第四段 (成本 1000): F 不使用。
  5. 所以，$\varphi_F = \frac{2000}{8} + \frac{1200}{6} + \frac{900}{3} = 250 + 200 + 300 = 750$。与题目给出的结果完全一致。
- 飞机 A 的成本分摊φ_A：飞机 A 需要 2,000 的跑道，因此它只使用了成本为 2000 的第一个分段。第一段 (成本 2000): 所有 8 架飞机都使用，均摊成本为 2000/8。所以，$\varphi_A = \frac{2000}{8} = 250$。也与题目结果一致。

■

1.7.ε-贪心算法的遗憾分析

令$\varepsilon_t = t^{-\frac{1}{3}}(K \log t)^{\frac{1}{3}}$，证明：$\varepsilon$-贪心算法的遗憾界为$O\left(T^{\frac{2}{3}}(K \log T)^{\frac{1}{3}}\right)$。

提示：整体思路是先考虑求任一时刻$t+1$的期望遗憾$\mathbb{E}[R_{t+1}]$，然后对这些遗憾求和，具体步骤如下：

1. 对于时刻$t+1$，注意在前t时刻中期望出现$\sum_{i=1}^{t}\varepsilon_i$次探索，则每个臂被选中的平均次数为$\sum_{i=1}^{t}\frac{\varepsilon_i}{K}$，然后定义事件$E$为

$$|\mu_t(a)-Q_t(a)|\leq\sqrt{\frac{K\log t}{\sum_{i=1}^{t}\varepsilon_i}}$$

   则接下来的步骤与课上讲的贪心算法分析类似；
2. 证明任一时刻$t+1$的期望遗憾$\mathbb{E}[R_{t+1}]$满足

$$\mathbb{E}[R_{t+1}]\leq 3\left(\frac{1}{t}K\log t\right)^{\frac{1}{3}}+O(t^{-2})$$

3. 将上式从 1 到T求和并放锁得到遗憾界。