第一课.3 混合策略纳什均衡

王彬，暨南大学经济学系，binwang@jnu.edu.cn

阅读材料：吉本斯 1.3.A。完成作业 1.9。

1. 概率

随机试验是指结果不确定也无法预测、直到观察才能知道结果的实验。例如抛硬币就是一个随机试验，在抛硬币前你无法预知抛硬币后是正面还是反面，结果是不确定的，但抛硬币后，你可以观测到抛硬币的结果。

样本空间是指随机试验所有可能的结果组成的集合。例如抛一次硬币这个随机试验中，样本空间就是{正，反}；抛一次骰子的随机试验中，样本空间就是{1,2,3,4,5,6}。

事件是样本空间的子集，由样本空间中的一部分元素组成。例如抛骰子游戏中，样本空间是{1,2,3,4,5,6}，那偶数点这个事件是{2,4,6}，是样本空间的子集。

随机变量是从样本空间到实数的函数。例如抛一次硬币这个随机试验中，我们可以给样本空间中的“正”赋值 1，给样本空间中的“反”赋值 0，用X代表抛一次硬币的随机变量，则：

$$X = \begin{cases} 1 \ \textit{正面向上} \\ 0 \ \textit{反面向上} \end{cases}$$

概率是随机试验的样本空间中结果或者事件发生的测度。概率的范围在[0,1]之间，0 意味着事件不可能发生，而 1 表示事件确定发生。概率大数值越高，则事件发生的可能性越大。在一次抛硬币的随机事件中，正面向上与反面向上这两个结果或者说事件的概率相等，都等于 0.5。概率也可以理解为当随机试验重复多次时结果发生的频率，例如将硬币抛的次数越来越多时，正面朝上的频率会越来越接近 0.5。

概率的性质：假设A和B都是随机试验的事件，A和B的概率是$P(A)$和$P(B)$，则：

- 如果$A \subseteq B$，则$P(A) \le P(B)$。
- 若$A_1, A_2, \cdots, A_n$满足$A_i \cap A_j = \Phi$, $A_1 \cup \cdots \cup A_n = \Omega$,则$P(A_1) + \cdots + P(A_n) = 1$。

概率分布是指对随机变量**所有**可能取值以及相应的概率的描述。例如，抛硬币随机试验中，概率分布是指随机变量即抛硬币的结果X的所有取值{1,0}以及相应的概率$P(X=1) = \frac{1}{2}, P(X=0) = \frac{1}{2}$。

期望，是随机试验中可能结果的概率乘以其随机变量取值的总和。它反映随机变量平均取值的大小。即若随机变量$X = x_1, \cdots, x_n$，并且$P(X = x_i) = p_i$，则其期望为：

$$E(X)=\sum_{i=1}^{n}p_i x_i$$

例如，掷骰子游戏中，骰子的点数 1,2,3,4,5,6 为随机变量 X，X 可能的取值为 1、2、3、4、5、6，每个取值的概率为$\frac{1}{6}$，则X的期望为：

$$E(X)=1*\frac{1}{6}+2*\frac{1}{6}+3*\frac{1}{6}+4*\frac{1}{6}+5*\frac{1}{6}+6*\frac{1}{6}=\frac{21}{6}$$

如果我们将掷骰子的游戏重复 n 次，随着 n 的数值越来越大，则 n 次数字的平均值越来越接近其期望$\frac{21}{6}$。

在博弈论中，我们可以引入概率探讨不确定性的策略，不确定性的策略带来的收益由期望收益决定。例如两个参与者的博弈中，双方都是 A 和 B 两种策略，其博弈双矩阵为：

		参与人 2: A	参与人 2: B
参与人 1	A	(1,2)	(0,1)
	B	(-1,3)	(2,-1)

如果参与人 1 知道参与人 2 的策略具有不确定性，并且知道参与人 2 以$\frac{1}{3}$的概率选 A、以$\frac{2}{3}$的概率选 B，那参与人 1 选 A 的收益是什么呢？我们知道，在参与人 1 选择 A 的情况下，如果参与人 2 选 A，则参与人 1 在策略组合(A,A)下获得收益为 1，而这个概率是$\frac{1}{3}$；同样，如果参与人 2 选 B，则参与人 1 在策略组合(A,B)下获得收益为 0，而这个概率是$\frac{2}{3}$。因此参与人 1 选择 A 的期望收益为$1\cdot\frac{1}{3}+0\cdot\frac{2}{3}=\frac{1}{3}$。

2. 网球赛

在网球赛中，击球者可以选择击打网球到对侧球场到左边和右边，而回球者也会选择移动到此侧的左边还是右边回球。如果回球者猜对了击球者的策略，那很大可能会赢得这一分，如果猜错了，则可能无法得分。

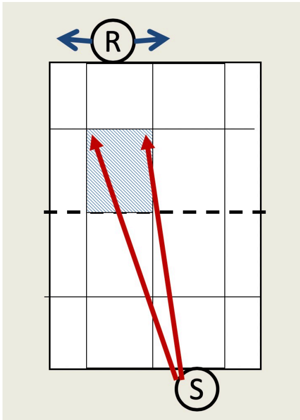
案例 1 球场博弈：球赛中有击球者和回球者，击球者和回球者选择相同（左或右），则击球者得到收益$1/4$，而回球者得到$3/4$；若击球者与回球者选择不同，则击球者得到收益$3/4$，而回球者得到$1/4$。

这也是一个典型的完全信息的静态博弈，球场博弈$G=\{S_1,S_2;u_1,u_2\}$的三个要素为：

1) 两个参与人{击球者，回球者}；
2) 击球者和回球者的策略空间相同$S_1=S_2=\{左，右\}$；

3) 对于四个策略组合（左，左）、（左，右）、（右，左）、（右，右），击球者的收益u_1为 1/4、3/4、3/4、1/4；回球者的收益u_2为 3/4、1/4、1/4、3/4。

图 1 网球场



用双矩阵来表示的话：

图 2 球场博弈

		回球者	
		左	右
击球者	左	(1/4,$\underline{3/4}$)	($\underline{3/4}$,1/4)
	右	($\underline{3/4}$,1/4)	(1/4,$\underline{3/4}$)

当回球者选择左时，击球者的最优反应为右，当回球者选择右时，击球者的最优反应为左；当击球者选择左时，回球者的最优反应为左，当击球者选择右时，回球者的最优反应为右。因此球场博弈并没有“纯策略”纳什均衡。

现实的球场中，击球回球不只一次，击球者不会总是用同样的策略，一旦被对手识破确定性的策略，对手总能针对性的出手取得更高的收益。所以现实中击球者或者回球者总是不确定性的出手，让对方摸不清己方的策略，从而达到某种均衡，那我们怎么描述这种不确定性的策略呢？在博弈论中，我们把这种不确定性的策略叫做混合策略。

3. 混合策略

混合策略定义为参与者策略空间S_i上的一个概率分布，例如，在球场博弈中，击球者的混合策略$P_1=(p_{11},p_{12})$表示击球者击打向对方左边的概率为p_{11}，击打向对方右边的概率为p_{12}，因为P_1是一个概率分布，所以我们有$0\le p_{11},p_{12}\le 1,p_{11}+p_{12}=1$。同

理，我们也可以定义回球者的混合策略$P_2=(p_{21},p_{22})$。而(P_1,P_2)是博弈的一个混合策略组合。

纯策略是混合策略的特殊情况，纯策略可以看作概率分布为$P_i=(0,0,\dots,1,0,\dots,0)$（第$i$个策略的概率为 1，其余策略的概率为 0）的混合策略。例如，球场博弈中，击球者选择左的纯策略等价于$P_1=(1,0)$的混合策略，而选择右的纯策略等价于$P_2=(0,1)$的混合策略。所以$(P_1,P_2)=\big((1,0),(0,1)\big)$这个混合策略组合等价于（左，右）这个纯策略组合。

那参与人的收益函数是什么呢？再以球场博弈为例，假设击球者的混合策略为$P_1=(p_{11},p_{12})=(\frac{1}{9},\frac{8}{9})$，而回球者的混合策略为$P_2=(p_{21},p_{22})=(\frac{1}{8},\frac{7}{8})$。这个问题就回到了博弈论的开创者之一冯诺伊曼和摩根斯坦创立的期望效用。当面临不确定性的结果时，理性人以随机事件带来收益的期望值为基础做出选择。

击球者在混合策略组合(P_1,P_2)下的收益可以分为两步：首先计算击球者选择左（右）时，因为回球者的不确定性P_2所带来的期望收益；然后计算击球者因为自身选择左右的不确定性P_1所带来的期望收益。

1. 首先计算击球者选择左右时的期望收益：
   - 当击球者选择左时。回球者以$p_{21}=\frac{1}{8}$的概率选择左，因此击球者以$\frac{1}{8}$的概率获得纯策略(左，右)的收益$\frac{1}{4}$；回球者以$p_{22}=\frac{7}{8}$的概率选择右，因此击球者以$\frac{7}{8}$的概率获得纯策略(左，右)的收益$\frac{3}{4}$。因此击球者选择左时的期望收益为$\frac{1}{8}*\frac{1}{4}+\frac{7}{8}*\frac{3}{4}=\frac{22}{32}$。
   - 当击球者选择右时。回球者以$p_{21}=\frac{1}{8}$的概率选择左，因此击球者以$\frac{1}{8}$的概率获得纯策略(右，右)的收益$\frac{3}{4}$；回球者以$p_{22}=\frac{7}{8}$的概率选择右，因此击球者以$\frac{7}{8}$的概率获得纯策略(右，右)的收益$\frac{1}{4}$。因此击球者选择左时的期望收益为$\frac{1}{8}*\frac{3}{4}+\frac{7}{8}*\frac{1}{4}=\frac{10}{32}$。
2. 其次计算因为击球者自身选择左右的不确定性P_1带来的期望收益。击球者以$p_{11}=\frac{1}{9}$的概率选择左，收益为$\frac{22}{32}$（第一步中击球者选择左的期望收益）；击球者以$p_{12}=\frac{8}{9}$的概率选择右，收益为$\frac{10}{32}$（第一步中击球者选择右的期望收益）。因此击球者在(P_1,P_2)下的期望收益为$\frac{1}{9}*\frac{22}{32}+\frac{8}{9}*\frac{10}{32}=\frac{102}{288}$。

同理，我们也可以计算回球者在混合策略(P_1,P_2)下的期望收益。

下面我们来定义参与人的期望收益。在两个参与人的策略式博弈$G=\{S_1,S_2;u_1,u_2\}$中，令J,K分别为两个参与人策略空间S_1和S_2中纯策略的个数，即$S_1=\{s_{11},\cdots,s_{1j},\cdots,s_{1J}\},S_2=\{s_{21},\cdots,s_{2k},\cdots,s_{2K}\}$。参与人 1，2 的混合策略分别为定义在$S_1$和$S_2$上的概率分布$P_1=\{p_{11},\cdots,p_{1j},\cdots,p_{1J}\}$和$P_2=\{p_{21},\cdots,p_{2k},\cdots,p_{2K}\}$。参与人 1 知道参与人 2 将以$P_2=\{p_{21},\cdots,p_{2k},\cdots,p_{2K}\}$的概率选择$\{s_{21},\cdots,s_{2k},\cdots,s_{2K}\}$，则参与人 1 选择纯策略$S_{1j}$的期望收益为：

$$\sum_{k=1}^{K} p_{2k}u_1(s_{1j},s_{2k})$$

并且参与人 1 以$\{p_{11},\cdots,p_{1j},\cdots,p_{1J}\}$的概率选择$\{s_{11},\cdots,s_{1j},\cdots,s_{1J}\}$的期望收益为：

$$v_1(P_1,P_2)=\sum_{j=1}^{J} p_{1j}\left(\sum_{k=1}^{K} p_{2k}u_1(s_{1j},s_{2k})\right)=\sum_{j=1}^{J}\sum_{k=1}^{K} p_{1j}p_{2k}u_1(s_{1j},s_{2k})$$

同理，我们也可以定义参与人 2 在混合策略组合(P_1,P_2)下的期望收益为：

$$v_2(P_1,P_2)=\sum_{k=1}^{K} p_{2k}\left(\sum_{j=1}^{J} p_{1j}u_2(s_{1j},s_{2k})\right)=\sum_{k=1}^{K}\sum_{j=1}^{J} p_{2k}p_{1j}u_2(s_{1j},s_{2k})$$

因此，参与人在混合策略(P_1,P_2)的收益为定义在纯策略组合(S_{1j},S_{2k})上的联合概率分布下的期望收益。这是因为参与人的概率分布是独立的。

因此，在球场博弈中，我们也可以换一种方式来计算击球者在混合策略组合(P_1,P_2)下的期望收益。在混合策略组合中，击球者以$\frac{1}{9}$的概率选择左，击球者以$\frac{1}{8}$的概率选择左，而击球者在(左，左)的纯策略组合中获得收益$\frac{1}{4}$，相当于击球者在混合策略组合中以$\frac{1}{9}*\frac{1}{8}$的概率获得纯策略组合（左，左）的收益$\frac{1}{4}$。同样的道理，击球者在混合策略下会以$\frac{1}{9}*\frac{7}{8}$、$\frac{8}{9}*\frac{1}{8}$、$\frac{8}{9}*\frac{7}{8}$的概率获得纯策略组合（左，右）、（右，左）、（右，右）的收益$\frac{3}{4}$、$\frac{3}{4}$、$\frac{1}{4}$。因此击球者在混合策略中的期望收益为：

$$v_1(P_1,P_2)=\frac{1}{9}*\frac{1}{8}*\frac{1}{4}+\frac{1}{9}*\frac{7}{8}*\frac{3}{4}+\frac{8}{9}*\frac{1}{8}*\frac{3}{4}+\frac{8}{9}*\frac{7}{8}*\frac{1}{4}=\frac{102}{288}$$

也可以通过同样的方法算出回球者在混合策略组合(P_1,P_2)下的期望收益。

在介绍了混合策略后，我们来看看球场博弈中的稳定的混合策略应该是什么？当击球者已知回球者以p_{21},p_{22}的概率选择左和右，那击球者应该如何选择呢，选择什么样的混合策略才让己方的期望收益最大呢？或者说击球者的**最优反应**是什么呢？

- 当击球者选择左时：回球者以p_{21}的概率选择左，击球者获取纯策略组合（左，左）下的收益$\frac{1}{4}$；回球者以p_{22}的概率选择右，击球者获取纯策略组合（左，右）下的收益$\frac{3}{4}$。因此击球者选择左的期望收益为$p_{21}*\frac{1}{4}+p_{22}*\frac{3}{4}$；
- 当击球者选择右时：回球者以p_{21}的概率选择左，击球者获取纯策略组合（右，左）下的收益$\frac{3}{4}$；回球者以p_{22}的概率选择右，击球者获取纯策略组合（右，右）下的收益$\frac{1}{4}$。因此击球者选择左的期望收益为$p_{21}*\frac{3}{4}+p_{22}*\frac{1}{4}$；

因此，

- 当击球者选择左的期望收益大于选择右的期望收益时，击球者会选择$P_1 = (1,0)$的混合策略，亦即选择左这个纯策略；
- 当击球者选择右的期望收益大于选择右的期望收益时，击球者会选择$P_1 = (0,1)$的混合策略，亦即选择右这个纯策略；
- 当击球者选择左的期望收益等于选择右的期望收益时，击球者会选择任意的混合策略$P_1 = (p_{11}, p_{12})$。因为两者相等时，任意的混合策略都让击球者的期望收益相等。

我们可以计算$P_2 = (p_{21}, p_{22})$的具体取值使得击球者处于上述三种情况：

- $p_{21} * \frac{1}{4} + p_{22} * \frac{3}{4} > p_{21} * \frac{3}{4} + p_{22} * \frac{1}{4}$，我们知道$p_{21} + p_{22} = 1$，所以，当$p_{21} < \frac{1}{2}$时，击球者会选择左的纯策略或者说$P_1 = (1,0)$的混合策略，即$p_{11} = 1$。
- $p_{21} * \frac{1}{4} + p_{22} * \frac{3}{4} < p_{21} * \frac{3}{4} + p_{22} * \frac{1}{4}$，我们知道$p_{21} + p_{22} = 1$，所以，当$p_{21} > \frac{1}{2}$时，击球者会选择左的纯策略或者说$P_1 = (0,1)$的混合策略，即$p_{11} = 0$。
- $p_{21} * \frac{1}{4} + p_{22} * \frac{3}{4} = p_{21} * \frac{3}{4} + p_{22} * \frac{1}{4}$，我们知道$p_{21} + p_{22} = 1$，所以，当$p_{21} = \frac{1}{2}$时，击球者会选择任意的混合策略$P_1 = (p_{11}, p_{12})$，即$p_{11} \in [0,1]$。

同理，我们也可以给定击球者的混合策略$P_1 = (p_{11}, p_{12})$从而探讨回球者的最优反应。

- 当$p_{11} > \frac{1}{2}$时，回球者会选择左的纯策略或者说$P_2 = (1,0)$的混合策略，即$p_{21} = 1$。
- 当$p_{11} < \frac{1}{2}$时，回球者会选择右的纯策略或者说$P_2 = (0,1)$的混合策略，即$p_{21} = 0$。
- 当$p_{11} = \frac{1}{2}$时，回球者会选择任何混合策略$P_2 = (p_{21}, p_{22})$，即$p_{21} \in [0,1]$。

我们可以用图来表示，给定回球者的混合策略$P_2 = (p_{21}, p_{22})$，击球者的最优反应是什么。

图 3 击球博弈最优反应图

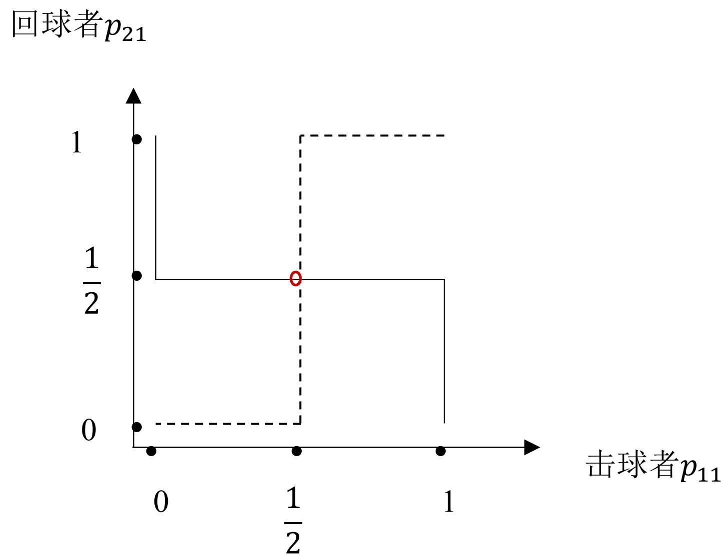


图 3 表示球场博弈中两个参与者的最优反应图。其中，实线是击球者面对回球者选择左的概率为p_{21}时己方选择左的概率p_{11}的最优反应曲线，而虚线是回球者面对击球者选择左的概率为p_{11}时己方选择左的概率p_{21}的最优反应曲线。

图中可以看出，在实线与曲线交叉时，即最优混合策略组合等于(P_1^*, P_2^*)，$P_1^* = P_2^* = \left(\frac{1}{2}, \frac{1}{2}\right)$时，双方的混合策略都是对方策略的最优反应，也就是说只要对方选择此混合策略组合中的混合策略，己方就不会偏离此混合策略组合。而这恰恰是**纳什均衡**的概念。我们也可以由此定义混合策略纳什均衡。

博弈$G = (S_1, \cdots, S_n; u_1, \cdots, u_n)$中，令$P_1, \cdots, P_n$为$n$个参与人定在各自策略空间上的概率分布，我们就称$P_i$为参与人$i$的混合策略，并且$(P_1, \cdots, P_n)$为博弈$G$的一个混合策略组合。给定其他参与人的混合策略组合$P_{-i}^* = P_1^*, \cdots, P_{i-1}^*, P_{i+1}^*, \cdots, P_n^*$，$P_i^*$为参与人$i$的最优选择，即

$$v_i(P_i^*, P_{-i}^*) \geq v_i(P_i, P_{-i}^*), \text{对于任意的 } P_i$$

我们就称$P^* = (P_1^*, \cdots, P_n^*)$为博弈$G$的**混合策略纳什均衡**。其中$P_i^*$最大化$v_i(P_i, P_{-i}^*)$的值，我们也称$P_i^*$为其他参与人策略组合$P_{-i}^*$的最优反应。

纳什均衡是不是一定存在呢？在一定条件下，答案是肯定的，这也是约翰纳什最大的贡献。

纳什定理（纳什均衡的存在性）：在n个参与者的策略式博弈$G = \{S_1, \cdots, S_n; u_1, \cdots, u_n\}$中，如果参与者人数$n$以及每个参与者的策略空间$S_i$都是有限的，则博弈存在至少一个纳什均衡，均衡可能包含混合策略。

纳什定理保证了相当广泛的博弈中均衡的存在性，但前面所应用的古诺模型以及贝特兰德模型都无法应用纳什定理，因为其策略空间是无限的，这说明纳什定理是一个充分条件，而不是必要条件。

4. 案例

4.1. 猜谜博弈

两人各自拿一枚硬币，如果都正面向上或者都是反面向上，则参与人 2 给参与人 1 一元钱，如果两人的正反面不同，则参与人 1 给参与人 2 一元钱。

图 4 猜谜博弈

		参与人 2	
		正	反
参与人1	正	($\underline{1}$，-1)	(-1，$\underline{1}$)
	反	(-1，$\underline{1}$)	($\underline{1}$，-1)

这个博弈中不存在纯策略纳什均衡。根据纳什定理，有限博弈必然存在纳什均衡，既然没有纯策略纳什均衡，此博弈必然存在混合策略纳什均衡(P_1, P_2)。设$P_1 = (p, 1-p), P_2 = (q, 1-q)$。

(1) 当参与人 2 以q的概率选择正，以$1-q$的概率选择反时：
- 若参与人 1 选择正，则其期望收益为$1 \cdot q + (-1) \cdot (1-q) = 2q-1$;
- 若参与人 1 选择反，则其期望收益为$(-1) \cdot q + 1 \cdot (1-q) = 1-2q$;

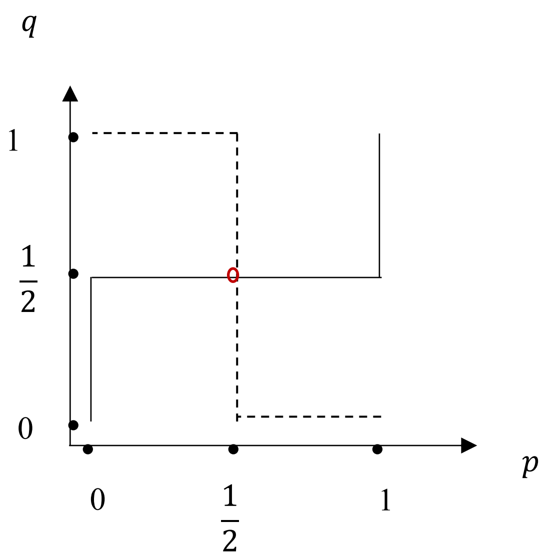
当$q > \frac{1}{2}$时，$2q-1 > 1-2q$，参与人 1 的最优反应为正，即$p=1$；当$q < \frac{1}{2}$时，$2q-1 < 1-2q$，参与人 1 的最优反应为反，即$p=0$; 当$q = \frac{1}{2}$时，$2q-1 = 1-2q$，参与人 1 的最优反应为任何概率选择正，即$p \in [0,1]$。

(2) 当参与人 1 以p的概率选择正，以$1-p$的概率选择反时：
- 若参与人 2 选择正，则其期望收益为$(-1) \cdot p + 1 \cdot (1-p) = 1-2p$;
- 若参与人 2 选择反，则其期望收益为$1 \cdot p + (-1) \cdot (1-p) = 2p-1$;

当$p < \frac{1}{2}$时，$1-2p > 2p-1$，参与人 2 的最优反应为正，即$q=1$；当$p > \frac{1}{2}$时，$1-2p < 2p-1$，参与人 2 的最优反应为反，即$q=0$；当$p = \frac{1}{2}$时，$1-2p = 2p-1$，参与人 2 的最优反应为任意概率选择正，即$q \in [0,1]$。

因此此博弈的纳什均衡为$(P_1^*，P_2^*), P_1^* = \left(\frac{1}{2}, \frac{1}{2}\right), P_2^* = \left(\frac{1}{2}, \frac{1}{2}\right)$。

图 5 猜谜博弈最优反应图



实线为参与人 1 选择正的概率p的最优反应曲线；虚线为参与人 2 选择正的概率q的最优反应曲线。

4.2. 鹰鸽博弈

案例 2 鹰鸽博弈：两只动物为争夺猎物而战，他们可以表现得很勇猛（鹰派），也可以表现得退让（鸽派），如果双方都退让，得收益 3，如一方勇猛一方退让，则勇猛一方得 4，退让一方得 1，如果双方都表现勇猛，则得 0.

图 6 鹰鸽博弈

		动物 2 鸽	动物 2 鹰
动物 1	鸽	(3,3)	(1,4)
动物 1	鹰	(4,1)	(0,0)

鹰鸽博弈中有两个纯纳什均衡（鸽，鹰）、（鹰，鸽）。接下来我们来探讨下鹰鸽博弈的混合策略纳什均衡。设两个动物的混合策略组合为(P_1, P_2)，其中$P_1 = (p, 1-p), P_2 = (q, 1-q)$。

(1) 当动物 2 以q的概率选择鸽，以$1-q$的概率选择鹰时：
- 若动物 1 选择鸽，则其期望收益为$3 \cdot q + 1 \cdot (1-q) = 2q + 1$;
- 若动物 1 选择鹰，则其期望收益为$4 \cdot q + 0 \cdot (1-q) = 4q$;

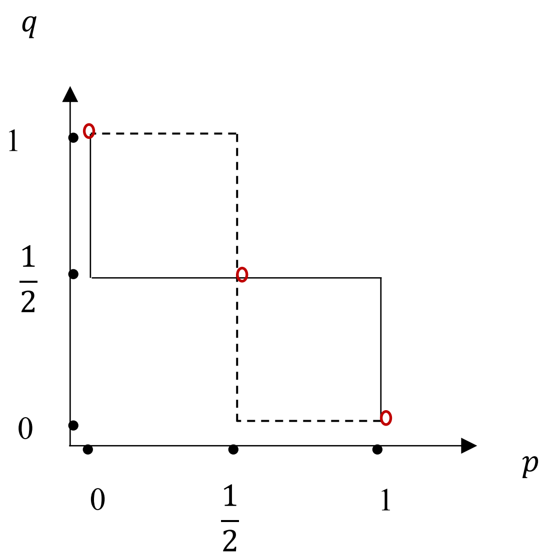
当$q < \frac{1}{2}$时，$2q + 1 > 4q$，动物 1 的最优反应为鸽，即$p = 1$；当$q > \frac{1}{2}$时，$2q + 1 < 4q$，动物 1 的最优反应为鹰，即$p = 0$; 当$q = \frac{1}{2}$时，$2q + 1 = 4q$，动物 1 的最优鹰应为任何概率选择鸽，即$p \in [0,1]$。

(2) 当动物 1 以p的概率选择鸽，以$1-p$的概率选择鹰时：
- 若动物 2 选择鸽，则其期望收益为$3 \cdot p + 1 \cdot (1-p) = 1 + 2p$;
- 若动物 2 选择鹰，则其期望收益为$4 \cdot p + 0 \cdot (1-p) = 4p$;

当$p < \frac{1}{2}$时，$1 + 2p > 4p$，动物 2 的最优反应为鸽，即$q = 1$；当$p > \frac{1}{2}$时，$1 + 2p < 4p$，参与人 2 的最优反应为鹰，即$q = 0$；当$p = \frac{1}{2}$时，$1 + 2p = 4p$，动物 2 的最优反应为任意概率选择鸽，即$q \in [0,1]$。

我们可以看到此博弈中，通过最优反应图，相交点有三个即$(P_1^*, P_2^*) = \big((1,0),(0,1)\big)$，$(P_1^*, P_2^*) = \big((0,1),(1,0)\big)$以及$(P_1^*, P_2^*) = \left(\left(\frac{1}{2}, \frac{1}{2}\right), \left(\frac{1}{2}, \frac{1}{2}\right)\right)$。前两个与纯策略纳什均衡等同。

图 7 鹰鸽博弈最优反应图



实线为动物 1 选择鸽的概率p的最优反应曲线；虚线为动物 2 选择鸽的概率q的最优反应曲线。

作业 1.9: 找到性别博弈中的混合策略。

案例 3 性别博弈：陆子豪和苏墨白出去约会吃饭，双方都有两种选择：海底捞和广州酒家。他们最主要的目的是一起吃饭，但是陆子豪更喜欢海底捞，而苏墨白更喜欢广州酒家，如果他们选择不同的地点，双方都没有效用。他们的收益函数在图 8 中表示。假设陆子豪和苏墨白的混合策略组合为$(\boldsymbol{P_1}, \boldsymbol{P_2})$，其中$\boldsymbol{P_1} = (\boldsymbol{p}, \boldsymbol{1-p}), \boldsymbol{P_2} = (\boldsymbol{q}, \boldsymbol{1-q})$.

图 8 性别博弈

		苏墨白	
		海底捞	广州酒家
陆子豪	海底捞	(2,1)	(0,0)
	广州酒家	(0,0)	(1,2)

(1)给定苏墨白的混合策略$\boldsymbol{P_2}$，找到陆子豪的最优反应；（2）给定陆子豪的混合策略$\boldsymbol{P_1}$，找到苏墨白的最优反应；（3）画出性别博弈的最优反应图，找到混合策略纳什均衡；（4）请对比纯策略纳什均衡的收益与混合策略均衡的期望收益，你能得到什么启示？

我们再来看一下鹰鸽博弈的一般情况，假设争夺的猎物的价值是v，没有抢到猎物的收益是 0，如果受伤，受伤的成本是$w > v$，若双方都出鹰派策略，则双方的收益为$\frac{v-w}{2} < 0$，若一方鹰一方鸽，鹰派得w，鸽派得 0，若两方都鸽，则双方都为$\frac{v}{2}$。

图 9 鹰鸽博弈

		动物 2 鸽	动物 2 鹰
动物 1	鸽	$(\frac{v}{2},\frac{v}{2})$	$(0,v)$
	鹰	$(v,0)$	$(\frac{v-w}{2},\frac{v-w}{2})$

我们假设动物 1 和动物 2 的混合策略组合是(P,Q)，其中$P=(p,1-p),Q=(q,1-q)$

给定动物 2 的混合策略为Q，动物 1 的最优反应是什么呢？

(1) 若动物 1 选择鸽，其期望收益为$q\cdot\frac{v}{2}+(1-q)\cdot 0=\frac{qv}{2}$;

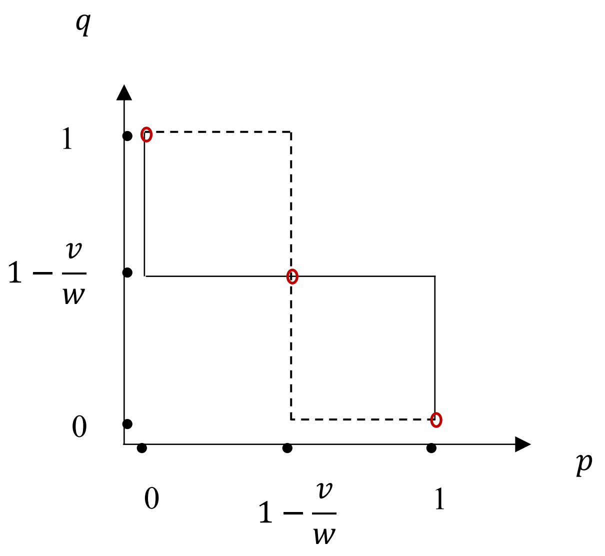
(2) 若动物 1 选择鹰，其期望收益为$q\cdot v+(1-q)\cdot\frac{v-w}{2}=\frac{(1+q)v-(1-q)w}{2}$;

当$q<1-\frac{v}{w}$时，动物 1 选择鸽，即$p^*=1$；当$q>1-\frac{v}{w}$时，动物 1 选择鹰，即$p^*=0$；当$q=1-\frac{v}{w}$时，动物 1 选择任何混合策略的收益相等，即$p^*\in[0,1]$。

给定动物 1 的混合策略为P，动物 2 的最优反应为：

当$p<1-\frac{v}{w}$时，动物 2 选择鸽，即$q^*=1$；当$p>1-\frac{v}{w}$时，动物 2 选择鹰，即$q^*=0$；当$p=1-\frac{v}{w}$时，动物 2 选择任何混合策略的收益相等，即$q^*\in[0,1]$。

图 10 鹰鸽博弈最优反应图



因此博弈中有三个纳什均衡$\big((1,0),(0,1)\big),\big((0,1),(1,0)\big),\left(\left(1-\frac{v}{w},\frac{v}{w}\right),\left(1-\frac{v}{w},\frac{v}{w}\right)\right)$。

我们可以看下这个混合策略$\left(\left(1-\frac{v}{w},\frac{v}{w}\right),\left(1-\frac{v}{w},\frac{v}{w}\right)\right)$，两个动物都以$1-\frac{v}{w}$的概率选择鸽，以$\frac{v}{w}$的概率选择鹰。并且双方的期望收益都是

$$Eu=\frac{v}{2}\left(1-\frac{v}{w}\right)$$

我们知道争斗的损失w要大于猎物的价值v，当受伤的代价越大时，混合策略中选择鹰的概率越小$\frac{v}{w}$，选择鸽的概率就越大，并且其期望收益也随着受伤的代价而增加。当代价接近猎物的价值时，双方的期望收益会更小，原因是双方都以一定概率选择鹰，而在战斗中可能受伤，而受伤的代价很大时，双方出鹰的概率小，受伤的概率小，期望收益反而增加。当受伤的代价趋近于无穷时，混合策略的收益趋近于$\left(\frac{v}{2},\frac{v}{2}\right)$，即（鸽，鸽）这个非纳什均衡的收益。

混合策略的例子告诉我们，在鹰鸽博弈中，受伤的代价大时，双方的混合策略均衡的收益会接近于(鸽，鸽)的非纳什均衡收益。当受伤的代价w接近于猎物的价值v时，双方的期望收益接近于 0。