

第九讲：信息不完善的静态博弈

——贝叶斯博弈 Bayesian Games

期中考试(共182位同学)

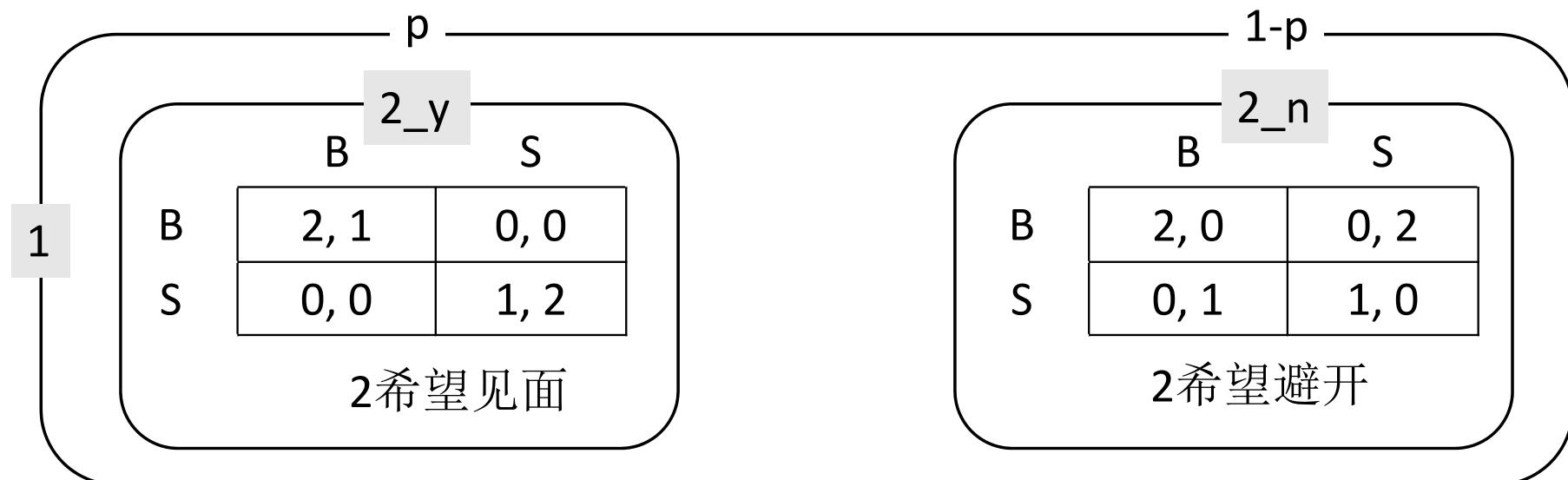
- 100分: 39位同学, 占21%
- 中位数: 96分, (50%的同学得分在此之上)
- 90分以上: 135位同学, 占74%
- 80分以上: 163位同学, 占90%

不完善信息(Imperfect Information)

- 博弈者所拥有的信息可能是不充分的，有时文献中区别两种情况
 - 不完美信息(imperfect information): 关于他人在以前的阶段已经采取的行动的信息不充分
 - 只出现在动态博弈里
 - 不完全信息(incomplete information): 关于博弈和博弈者的某些特征（例如博弈者偏好、行动成本等等）的信息不充分
 - 可出现在静态或动态博弈中
- 但我们按所用主要参考书（Osborne）的术语，都统称“不完善信息”（imperfect information）

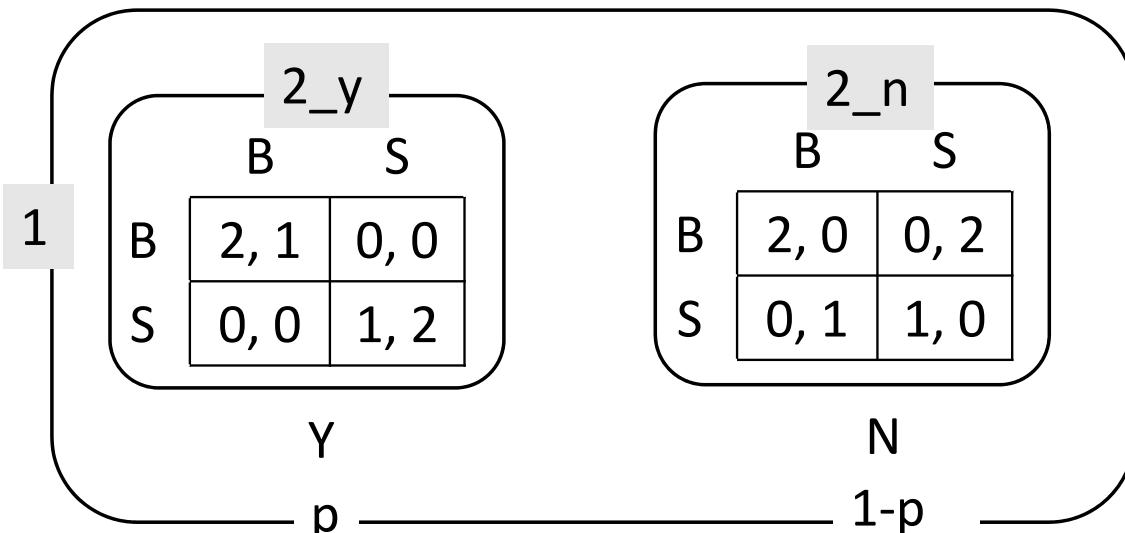
例：信息不完善性别战争(BOS)博弈

- 博弈者2有两种可能的偏好：希望和1见面，或者希望避开1
- 博弈者2知道自己的偏好：所以她有两个类型——y型和n型
- 博弈者1不知道自己面对的2内心的偏好；他只知道第一种情况的概率是 p ，第二种情况的概率是 $1-p$



一些基本要素

- 世界的状态(state)
 - 有两种: Y: 1与喜欢他的2博弈
N: 1与不喜欢他的2博弈
- 博弈者的类型(type)
 - 类型从根本上反映的是博弈者掌握的信息
 - 这里2有两种类型; 1只有一种类型
- 博弈者的信念(belief): 世界处于某种状态中的概率
 - 每一种类型的博弈者, 根据自己掌握的信息, 形成关于自己在每个世界状态的信念 (概率判断)
 - 1的信念: 处在世界状态Y的概率是 p , 处于世界状态N的概率 $1-p$
 - 2的每种类型的信念: 2_y 知道自己在世界状态Y, 2_n 知道自己在世界状态N



博弈者的策略与收益

- 每一种类型的博弈者都各自选择自己的行动；而每一方博弈者的策略包括自己所有类型的行动选择
 - 就好像在动态博弈中，每一方博弈者的策略需覆盖自己所有的决策节点
- 刚才的博弈中
 - 博弈者1只有一个类型，他的可能策略是B和S
 - 博弈者2有两个类型，每一种类型都可以选B或S，所以2的可能策略是BB、BS、SB、SS 四种
 - 博弈的策略型如下

		博弈者2			
		BB	BS	SB	SS
博弈者1	B				
	S				

博弈者的策略与收益

- 每一种类型的博弈者，根据自己的信念，最大化自己的期望收益
- 先看博弈者1：只有一种类型；他相信自己在Y世界的概率是 p ，在N世界的概率是 $1-p$
- 博弈者1的期望收益

		博弈者2			
		BB	BS	SB	SS
博弈者1	B	2	2p	2-2p	0
	S	0	1-p	p	1

- 如果 $p=1/2$ ，可以圈出1的最优反应
- 1的最优反应函数（图像）： $(B, BB), (B, BS), (B, SB), (S, SS)$

博弈者的策略与收益

- 博弈者2有两种类型，每一种类型分别按信念选择自己的最优反应

		2_y	
		B	S
B	B	2, 1	0, 0
	S	0, 0	1, 2
		2_n	
		B	S
B	B	2, 0	0, 2
	S	0, 1	1, 0

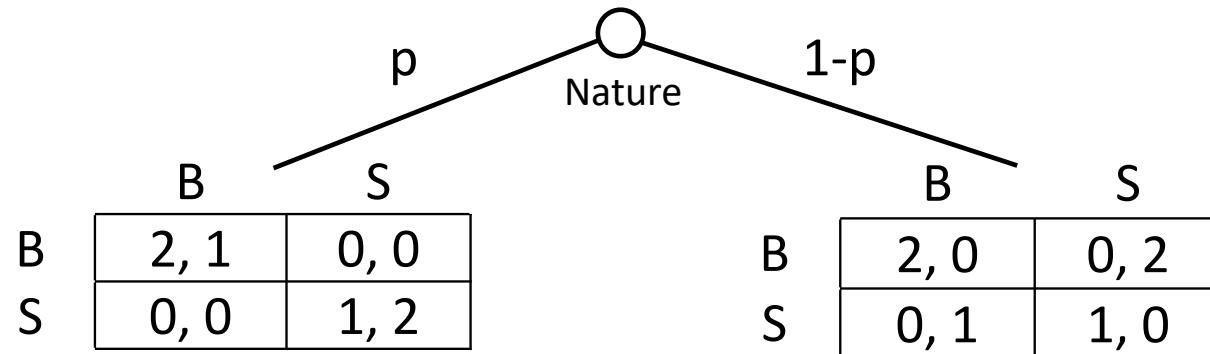
- “拼装”起来，博弈者2的最优反应函数：(B, BS), (S, SB)
- 之前已经得到，1的最优反应函数是：(B, BB), (B, BS), (B, SB), (S, SS)
- 双方最优反应函数的交集——均衡：(B, BS)
- 贝叶斯纳什均衡：给定其他博弈者的策略，每一方博弈者的每种类型所选的策略是最优的

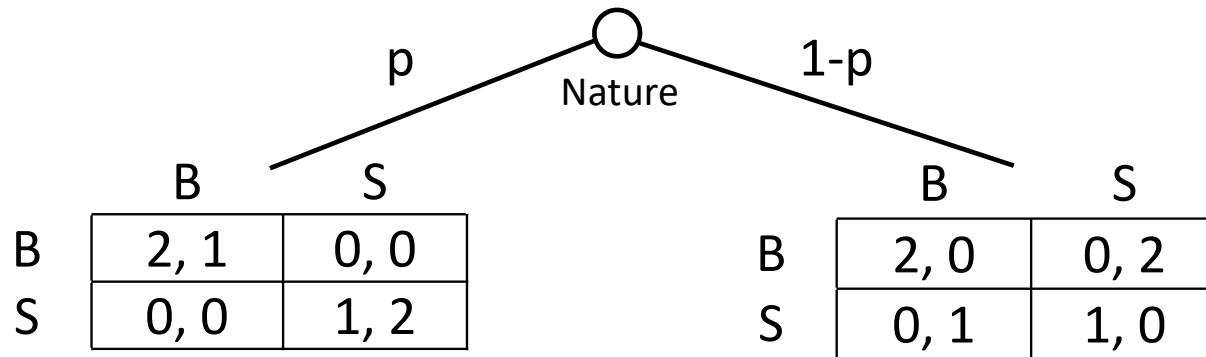
- 能不能在一个收益矩阵中把双方的期望收益都写出来？

		博奕者2			
		BB	BS	SB	SS
博奕者1	B	2,	2p,	2-2p,	0,
	S	0,	1-p,	P,	1,

海萨尼转换

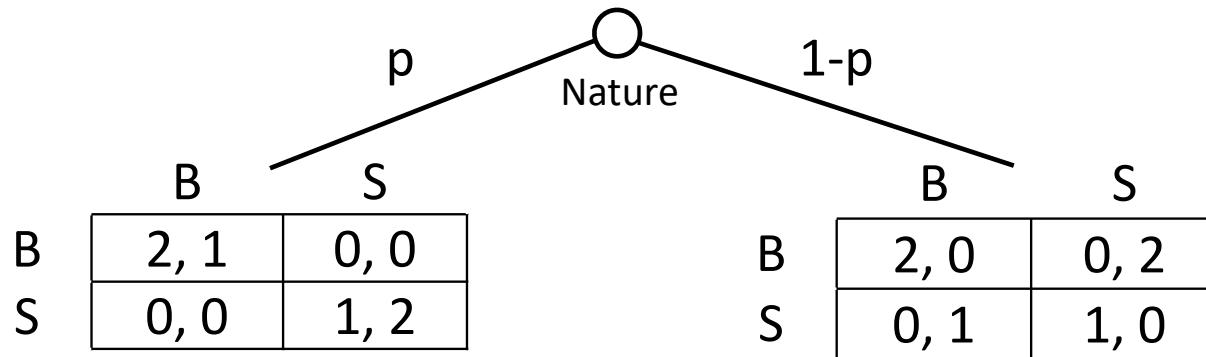
- 另一种视角：“自然”先行动，按概率决定世界状态
- 将博弈者2看作一个统一的博弈者，拥有两个决策节点，要选择策略以最大化自己总的期望收益





- 把博弈者2看作是一个统一的决策者， 拥有两个决策节点，要选择策略以最大化自己总的期望收益
- 收益矩阵

		博弈者2			
		BB	BS	SB	SS
博弈者1	B	2, p	2p, 2-p	2-2p, 0	0, 2-2p
	S	0, 1-p	1-p, 0	p, 1+p	1, 2p



- 把博弈者2看作是一个统一的决策者，拥有两个决策节点，要选择策略以最大化自己总的期望收益
- 收益矩阵（设定 $p=1/2$ ）

		博奕者2			
		BB	BS	SB	SS
博奕者1	B	2, p	2p, 2-p	2-2p, 0	0, 2-2p
	S	0, 1-p	1-p, 0	p, 1+p	1, 2p

- 贝叶斯纳什均衡是(B, BS); 和前面的结果完全一样
- 两者等价的理由：每一种类型的2都选择了最优应对 \Leftrightarrow 2的总期望收益最优

贝叶斯博弈的构成

- 世界状态(states): $\omega \in \Omega$
 - 博弈者
 - 博弈者*i*可采取的行动: $a_i \in A_i$
- 关于世界状态的前提概率(prior): 是公共信息, 又称公共前提 (common prior)
- “自然” 按前提概率行动, 决定世界状态

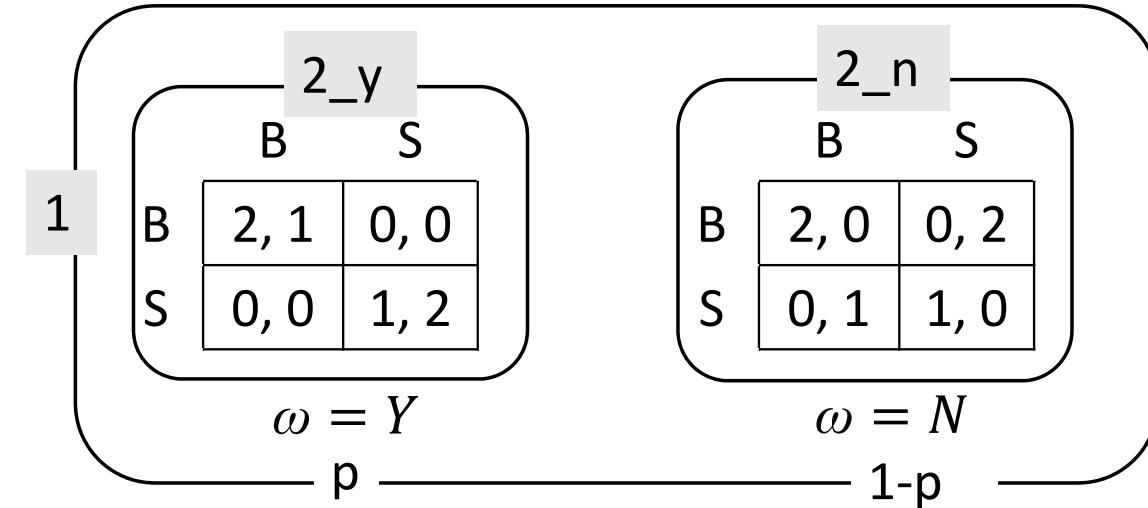
	B	S
B	2, 1	0, 0
S	0, 0	1, 2

$$\begin{matrix} \omega = Y \\ p \end{matrix}$$

	B	S
B	2, 0	0, 2
S	0, 1	1, 0

$$\begin{matrix} \omega = N \\ 1-p \end{matrix}$$

贝叶斯博弈的构成



- 每一位博弈者各自收到关于世界状态的某些信号(signal): t_i
 - 信号是私人信息: 博弈者只知道自己收到的信号, 不知道别人收到的信号内容
 - 每个博弈者收到的信号是世界状态的函数: $t_i = \tau_i(\omega)$, 根据信号, 博弈者可以进一步推断自己在哪个世界状态 (的概率)
 - 不同博弈者所收到信号的信息含量可以不同: 例如上面例子中, 博弈者2在两个不同的世界里会收到不同的信号, 而博弈者1在两个世界里收到同一个信号
- 博弈者收到的信号决定了其掌握的信息, 从而划分为不同的类型
 - 上面例子里, 博弈者2有两个类型, 博弈者1只有一个类型

分析贝叶斯博弈

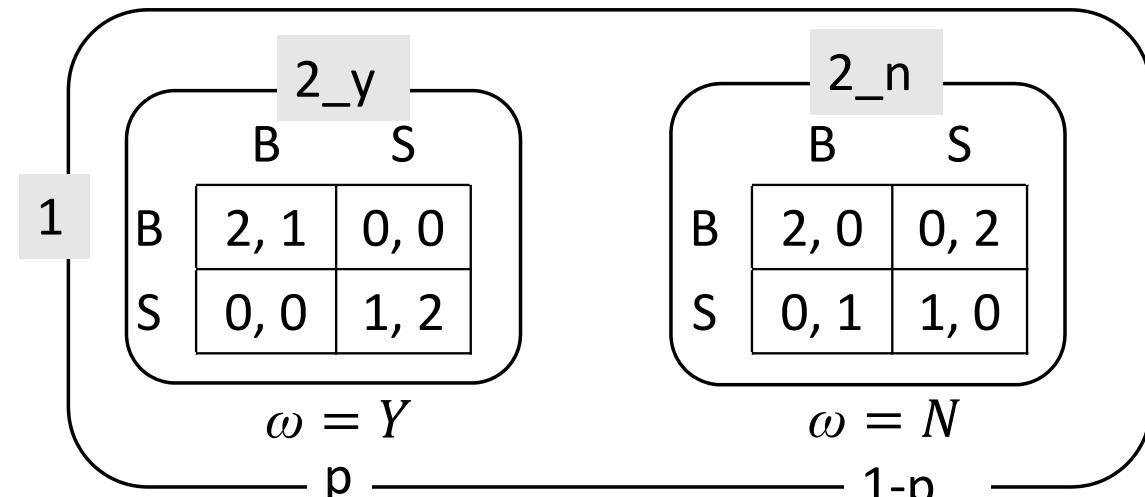
- 博弈者的信念 (belief)

- 每一类型的博弈者 t_i 都形成关于自己所处世界状态的概率判断: $\Pr(\omega | t_i)$
- 原则上, 信念是博弈者基于前提概率和自己获得的信号, 应用贝叶斯法则推论得到的, “贝叶斯博弈”由此得名
 - 本课中, 通常我们可以直观地确定博弈者的信念 (无需使用贝叶斯公式)

- 博弈者的每一类型 (基于信念) 选择策略最大化自己的期望收益

- 对于类型 t_i 的博弈者, 策略组合 (a_i, a_{-i}) 带给她的期望收益

$$\sum_{\omega \in \Omega} \Pr(\omega | t_i) \cdot u_i(a_i, a_{-i}, \omega)$$



贝叶斯博弈的纳什均衡

贝叶斯纳什均衡（简称贝叶斯均衡）

- 将每一方博弈者 i 的每一种类型 t_i 看作一个行动者
- 在一个策略组合中，如果对于任意类型的行动者 t_i 而言，给定其他行动者的策略，以及自己关于世界状态的信念，自己现在所选的策略是最优的（即最大化自己的期望收益），那么该策略组合称为贝叶斯纳什均衡

例：双方各有两种类型的BOS

- 博弈者1和2都各有两种可能的偏好态度：y, n
- 世界状态
 - yy, yn, ny, nn
- 前提概率(Common Prior)
 - 1是y和n的概率分别为q和1-q
 - 2是y和n的概率分别为p和1-p
 - 1的偏好和2的偏好在概率上是独立的

	p	$1-p$												
	B S	B S												
q	<table border="1" style="margin-left: auto; margin-right: 20px;"> <tr><td>B</td><td>2, 1</td><td>0, 0</td></tr> <tr><td>S</td><td>0, 0</td><td>1, 2</td></tr> </table>	B	2, 1	0, 0	S	0, 0	1, 2	<table border="1" style="margin-left: auto; margin-right: 20px;"> <tr><td>B</td><td>2, 0</td><td>0, 2</td></tr> <tr><td>S</td><td>0, 1</td><td>1, 0</td></tr> </table>	B	2, 0	0, 2	S	0, 1	1, 0
B	2, 1	0, 0												
S	0, 0	1, 2												
B	2, 0	0, 2												
S	0, 1	1, 0												
	yy	yn												
$1-q$	<table border="1" style="margin-left: auto; margin-right: 20px;"> <tr><td>B</td><td>0, 1</td><td>2, 0</td></tr> <tr><td>S</td><td>1, 0</td><td>0, 2</td></tr> </table>	B	0, 1	2, 0	S	1, 0	0, 2	<table border="1" style="margin-left: auto; margin-right: 20px;"> <tr><td>B</td><td>0, 0</td><td>2, 2</td></tr> <tr><td>S</td><td>1, 1</td><td>0, 0</td></tr> </table>	B	0, 0	2, 2	S	1, 1	0, 0
B	0, 1	2, 0												
S	1, 0	0, 2												
B	0, 0	2, 2												
S	1, 1	0, 0												
	ny	nn												

- 信号和信息
 - 1和2各自知道自己的偏好，但不知道对方的偏好
- 由此确定了博弈者的类型
- 每一类型博弈者的信念(Belief)

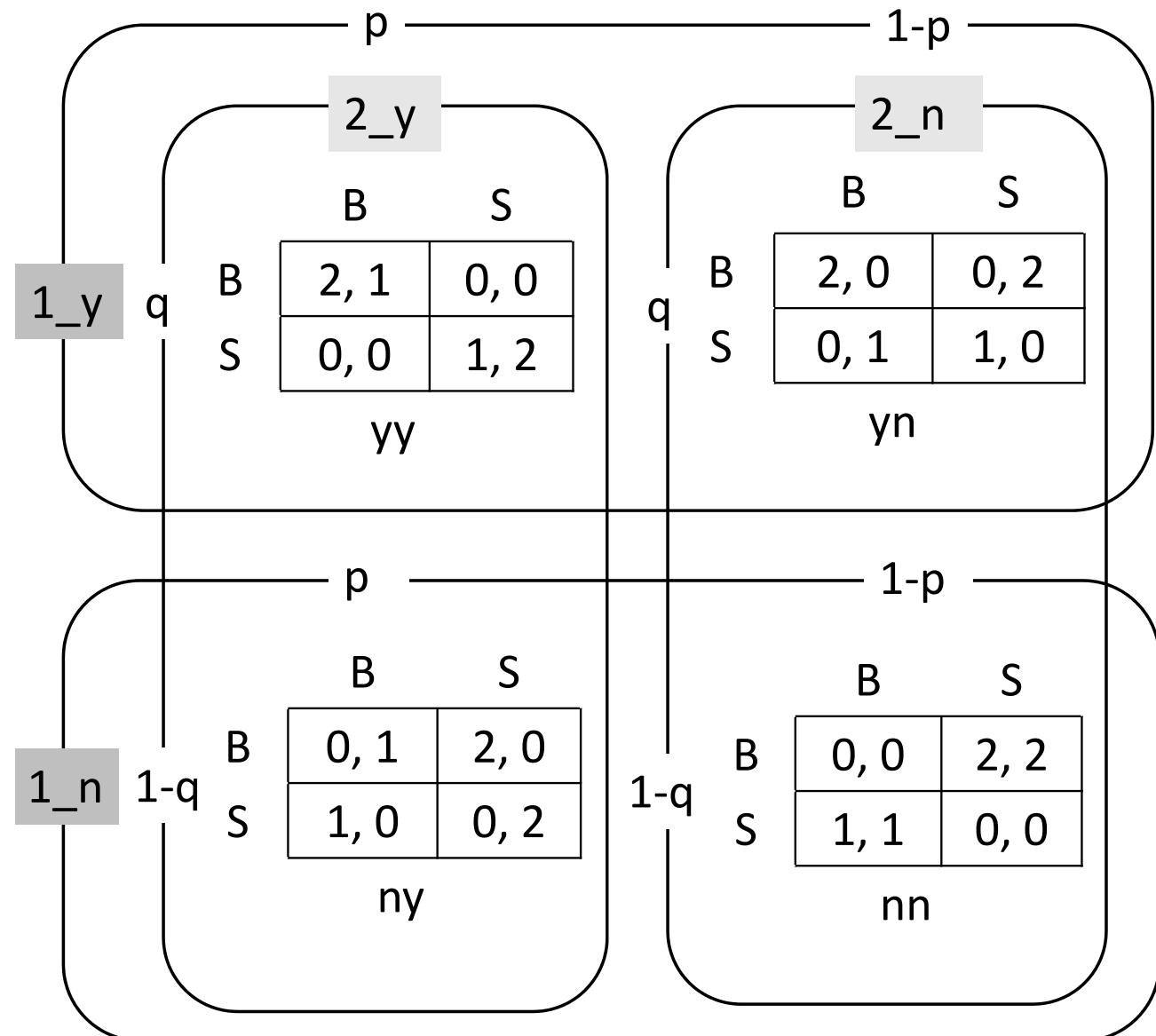
$$t_1=y: \Pr(yy|t_1=y)=p, \Pr(yn|t_1=y)=1-p$$

$$t_1=n: \Pr(ny|t_1=n)=p, \Pr(nn|t_1=n)=1-p$$

$$t_2=y: \Pr(yy|t_2=y)=q, \Pr(ny|t_2=y)=1-q$$

$$t_2=n: \Pr(yn|t_2=n)=q, \Pr(nn|t_2=n)=1-q$$

- 寻找最优反应和贝叶斯纳什均衡



博弈者1的收益矩阵

	BB	BS	SB	SS
$t_1=y$				
S				

博弈者2的收益矩阵

$t_2=y$ $t_2=n$

	BB	BS	SB	SS
$t_1=n$				
S				

博弈者1的收益矩阵

		BB	BS	SB	SS	
		B	2	$2p$	$2-2p$	0
		S	0	$1-p$	p	1

博弈者2的收益矩阵

$t_2=y$ $t_2=n$

		BB	BS	SB	SS	
		B	0	$2-2p$	$2p$	2
		S	1	p	$1-p$	0

参数值设定: $p=1/2$, $q=2/3$

博弈者1的收益矩阵

		BB	BS	SB	SS	
		B	2	$2p$	$2-2p$	0
$t_1=y$	B	0	$1-p$	p	1	
	S					

博弈者2的收益矩阵

		$t_2=y$	$t_2=n$
		B	2
$t_1=y$	B	0	$2-2p$
	S	1	p

		BB	BS	SB	SS	
		B	0	$2-2p$	$2p$	2
$t_1=n$	B	1	p	$1-p$	0	
	S					

- 1的最优反应: (BS, BB), (BB, BS), (BB, SB), (SB, SS)

参数值设定: $p=1/2$, $q=2/3$

博弈者1的收益矩阵

		BB	BS	SB	SS	
		B	2	$2p$	$2-2p$	0
$t_1=y$	B	0	$1-p$	p	1	
	S	1	p	$1-p$	0	

		BB	BS	SB	SS	
		B	0	$2-2p$	$2p$	2
$t_1=n$	B	1	p	$1-p$	0	
	S	1	p	$1-p$	0	

博弈者2的收益矩阵

		$t_2=y$	
		B	S
$t_1=y$	BB		
	BS		

		$t_2=n$	
		B	S
$t_1=n$	BB		
	BS		

- 1的最优反应: (BS, BB), (BB, BS), (BB, SB), (SB, SS)

参数值设定: $p=1/2$, $q=2/3$

博弈者1的收益矩阵

		BB	BS	SB	SS	
		B	2	$2p$	$2-2p$	0
$t_1=y$	B	0	$1-p$	p	1	
	S	1	p	$1-p$	0	
		BB	BS	SB	SS	
$t_1=n$	B	0	$2-2p$	$2p$	2	
	S	1	p	$1-p$	0	

博弈者2的收益矩阵

		$t_2=y$	$t_2=n$
		B	S
BB	B	1	0
	S	q	$2-2q$
BS	B	$1-q$	$2q$
	S	0	2
SB	B	0	$2-2q$
	S	q	$2q$
SS	B	1	0
	S	$1-q$	$2q$

- 1的最优反应: (BS, BB), (BB, BS), (BB, SB), (SB, SS)

参数值设定: $p=1/2$, $q=2/3$

博弈者1的收益矩阵

		BB	BS	SB	SS	
		B	2	$2p$	$2-2p$	0
$t_1=y$	B	0	$1-p$	p	1	
	S	1	$2-2p$	$2p$	2	
		BB	BS	SB	SS	
$t_1=n$	B	0	$2-2p$	$2p$	2	
	S	1	p	$1-p$	0	

博弈者2的收益矩阵

		$t_2=y$		$t_2=n$	
		B	S	B	S
		BB	1	0	2
		BS	q	$2-2q$	$2q$
		SB	$1-q$	$2q$	$2-2q$
		SS	0	2	1

- 1的最优反应: (BS, BB), (BB, BS), (BB, SB), (SB, SS)
- 2的最优反应: (BB, BS), (BS, BS), (BS, SS), (SB, SB), (SB, SS), (SS, SB)
- 贝叶斯纳什均衡 : (BB, BS), (SB, SS)

- 也可以在一个统一的收益矩阵里求解，但是通常计算量更大
- 如何得出下面这个收益矩阵中双方的收益？（有兴趣的同学可以课下自己尝试）

	BB	BS	SB	SS
BB	2q, p			
BS				
SB				
SS				

贝叶斯博弈和贝叶斯均衡：要点小结

- 贝叶斯均衡：每一方的每一种类型的博弈者都要选择自己的最优策略
- 分析贝叶斯博弈，首先要明确每一方博弈者拥有的可能的**类型**
 - 注意：类型是根据博弈者掌握的信息来划分的
- 我们一般分别从每一方的每一种类型的视角出发，找出其最优反应
 - 每种类型的博弈者根据自己的信念（对于自己处在各种世界状态中的概率判断）来计算期望收益
- 之后再“拼装”，得到该博弈方的整体策略

连续型的例子：信息不完善的库诺特竞争

- 两家企業生产同一种产品。企業*i*生产的数量为 q_i , 总产量为 $Q=q_1+q_2$, 市场对产品的需求函数为 (其中P是价格, $\alpha > c$) :

$$P(Q) = \begin{cases} \alpha - Q & \text{if } Q \leq \alpha \\ 0 & \text{if } Q > \alpha \end{cases}$$

企業*i*的收益为: $\pi_i = Pq_i - c_i q_i$

企業1只有一种成本: $c_1 = c$; 但企業2的成本有两种可能的状态: L: $c_2 = c_L$ 或 H: $c_2 = c_H$

公共前提(Common prior): 2的成本为L的概率是 θ , 为H的概率是 $1-\theta$

信号(信息): 企業2知道自己的成本, 企業1不知道对方的成本

- 由信息规定了类型: 1有一个类型, 2有两个类型
- 所以, 各个类型博弈者的信念为

- 博弈者1: 相信 $\Pr(L) = \theta$, $\Pr(H) = 1-\theta$
- 博弈者2: H型相信 $\Pr(H|t_2=H) = 1$; L型相信 $\Pr(L|t_2=L) = 1$

- 策略
 - 1只有一种类型：策略是 q_1
 - 2有两种类型：2_H策略 q_H , 2_L的策略 q_L ; 2的策略是 (q_H, q_L)
- 最优反应

企业1：选择 q_1 , 最大化期望收益 (注意世界状态有两种)

$$\begin{aligned}
 E(\pi_1) &= \theta((\alpha - Q_L)q_1 - cq_1) + (1 - \theta)((\alpha - Q_H)q_1 - cq_1) \\
 &= \theta((\alpha - q_1 - q_L)q_1 - cq_1) + (1 - \theta)((\alpha - q_1 - q_H)q_1 - cq_1) \\
 &= (\alpha - c)q_1 - \theta q_L q_1 - (1 - \theta)q_H q_1 - q_1^2
 \end{aligned}$$

一阶条件： $\alpha - c - \theta q_L - (1 - \theta)q_H - 2q_1 = 0$

最优反应： $q_1 = \frac{1}{2}(\alpha - c - (\theta q_L + (1 - \theta)q_H))$

企业2有两种类型

- 企业2_L: 选择 q_L , 最大化: $(\alpha - Q_L)q_L - c_L q_L$
 $= (\alpha - q_1 - q_L)q_L - c_L q_L$
 $= (\alpha - c_L - q_1)q_L - q_L^2$

使用一阶条件: $(\alpha - c_L - q_1) - 2q_L = 0$

最优反应: $q_L = \frac{1}{2}(\alpha - c_L - q_1)$

- 企业2_H: 类似地求得: $q_H = \frac{1}{2}(\alpha - c_H - q_1)$

贝叶斯纳什均衡, 是所有类型博弈者的最优反应函数的交点, 即求解

$$\left\{ \begin{array}{l} q_1 = \frac{1}{2}(\alpha - c - (\theta q_L + (1 - \theta)q_H)) \\ q_L = \frac{1}{2}(\alpha - c_L - q_1) \\ q_H = \frac{1}{2}(\alpha - c_H - q_1) \end{array} \right.$$

将 q_H, q_L 代入 q_1 , 解出

$$q_1 = \frac{1}{3}(\alpha - 2c + \theta c_L + (1 - \theta)c_H)$$

$$q_L = \frac{1}{3}(\alpha + c - 2c_L) - \frac{1-\theta}{6}(c_H - c_L)$$

$$q_H = \frac{1}{3}(\alpha + c - 2c_H) + \frac{\theta}{6}(c_H - c_L)$$

如果只有L ($\theta=1$)

$$q_1 = \frac{1}{3}(\alpha - 2c + c_L)$$

$$q_L = \frac{1}{3}(\alpha + c - 2c_L)$$

如果只有H ($\theta=0$)

$$q_1 = \frac{1}{3}(\alpha - 2c + c_H)$$

$$q_H = \frac{1}{3}(\alpha + c - 2c_H)$$

和完善信息的情境（1和2都知道所处世界状态）比较，现在

- 企业1的行为要考虑到2的两种可能类型及其概率
 - 1的产量，比起知道对方是L类型时多，比知道对方是H类型时少
- 企业2的每一种类型的行为也受到不完善信息的影响
 - 比起完善信息时，现在企业2的L类型的产量减少，H类型产量增加