

《博弈论基础》

蒋文华

浙江大学公共管理学院

Mb : 13989895432 (679432)

Email : jwh0422@163.com

第七讲 混合策略与监督博弈

第十三章 混合策略

第十四章 监督博弈

第十三章 混合策略

纯策略和混合策略

纯策略：如果一个策略要求参与者在每一个给定信息情况下只选择一种特定的行动。

混合策略：如果一个策略要求参与者在给定信息情况下以某种概率分布随机地选择不同的行动。

混合策略

在 n 个参与人博弈的策略式表述 $G = \{S_1, \dots, S_n; u_1, \dots, u_n\}$ 中, 假定参与人 i 有 K 个纯策略: $S_i = \{s_{i1}, \dots, s_{iK}\}$, 那么, 概率分布 $\sigma_i = (\sigma_{i1}, \dots, \sigma_{iK})$ 称为 i 的一个混合策略, 这里 $\sigma_{iK} = \sigma(s_{ik})$ 是 i 选择 s_{ik} 的概率, 对于所有的 $k=1, \dots, K$, $0 \leq \sigma_{iK} \leq 1, \sum \sigma_{iK} = 1$.

随机药物检测

		运动员	
		服药	未服药
IOC	抽检	1 , -1	-1 , 1
	未抽检	-1 , 1	1 , -1

点球射门

攻方的策略是：攻左边，或攻右边

守方的策略是：守左边，或守右边
(以守方的方向为准)

		守	
		左边	右边
攻	左边	-1 , 1	1 , -1
	右边	1 , -1	-1 , 1

石头、剪子、布

浙江大学发研究报告：石头剪刀布百战百胜策略

浙江大学、浙江工商大学和中科院理论物理研究所的研究人员通过实验发现了石头剪刀布的一个制胜策略。研究人员招募了 360 名学生，将他们分成六组，随机配对玩 300 轮石头剪刀布游戏，在每一轮中获胜的学生将会获得少量人民币奖励。通过观察学生使用的策略，他们发现了获胜者或失利者习惯使用的游戏策略。

如果你的剪刀输给了对手的石头，下一轮你更有可能出能战胜石头的布；如果你是获胜者，那么下一轮你更有可能沿用相同的出手。赢家保持现状输家做出改变的策略（胜留败走）。

石头剪刀布的制胜策略：如果你是输家，下一轮换用能打败对手的出手；如果你是赢家，下一轮不要再使用原来的出手。

比如：你用石头打败了对手的剪刀，那么下一轮你不能再出石头，而应该出剪刀，因为对方很有可能会出布。

假设我和你一起玩“石头、剪子、布”的游戏，如果我告诉你说，我准备出“石头”，

请问：你会出什么？（赌注为 10 元）

请问：事先的告知会影响你的选择吗？

特别提示：

小得失相信人，大得失怀疑人！

石头剪子布（高级版）

	石头	剪子	布
石头	0 , 0	1 , -1	-5 , 5
剪子	-1 , 1	0 , 0	2 , -2
布	5 , -5	-2 , 2	0 , 0

试问：出石头和出布的学生，哪个多？

第十四章 监督博弈

雇主与雇员的监督博弈

V 是雇员的贡献，即雇员为雇主创造的价值； W 是雇员的工资， H 是雇员的付出。 C 是雇主检查的成本。同时假定 $H < W < V$ ， $W > C$ 。

V W H

2 个问题:

1、求解混合策略纳什均衡。

2、求解雇主支付 W 为多少时，其期望收益最大？

设雇主检查的概率为 α ，不检查的概率为 $1-\alpha$ ，雇员偷懒的概率为 β ，不偷懒的概率为 $1-\beta$ 。那么，在这种情况下，雇主的期望收益（用 $\Pi_{\text{雇主}}$ 来表示）是：

$$\Pi_{\text{雇主}} = (-C) * \alpha * \beta + (V - W - C) * \alpha * (1 - \beta) + (-W) * (1 - \alpha) * \beta + (V - W) * (1 - \alpha) * (1 - \beta)$$

雇员的期望收益（用 $\Pi_{\text{雇员}}$ 来表示）是：

$$\Pi_{\text{雇员}} = 0 * \alpha * \beta + (W - H) * \alpha * (1 - \beta) + W * (1 - \alpha) * \beta + (W - H) * (1 - \alpha) * (1 - \beta)$$

双方期望收益最大化的

必要条件是一阶导数为零。

$$d\Pi_{\text{雇主}}/d\alpha = 0$$

$$d\Pi_{\text{雇员}}/d\beta = 0$$

通过计算得出

$$\alpha = H/W$$

$$\beta = C/W$$

混合策略均衡解：雇主以 H/W 的概率检查，当雇员以 C/W 的概率偷懒，达到均衡状态。

雇主检查的概率 = H/W , 雇员偷懒的概率 = C/W

H 越大, 雇主越容易检查。

C 越大, 雇员越容易偷懒。

W 越大, 雇员越少偷懒, 雇主越少检查。

2、雇主支付 W 为多少时，其期望收益最大？

雇主检查的概率 $= H/W$ ，雇员偷懒的概率 $= C/$

W

计算出雇主和雇员各自的期望收益是：

$$\pi_{\text{雇主}} = (-W) \beta + (V-W)(1-\beta) = V-W-V\beta = V-W-VC/W$$

$$\pi_{\text{雇员}} = W-H$$

雇主检查的概率 = H/W ，雇员偷懒的概率 = C/W

雇主的期望收益可分为两部分：一部分是雇员努力工作给雇主带来的净收益为 $V-W$ ，另一部分是雇员偷懒给雇主带来的净损失 $V\beta$ ，这里的 V 是雇员的贡献， β 是雇员的偷懒概率，在均衡状态下 $\beta = C/W$ 。

当我们求出了雇主的期望收益后，就可以进一步计算出当雇主给雇员支付多少工资 (W) 时，其期望收益达到最大值。

雇主的期望收益对工资一阶导数为 0，是雇主实现收益最大化的必要条件。

$$d\pi_{\text{雇主}}/dW=0$$

由此求出 $W=$

雇员的 $w =$

$$\frac{V - H}{C}$$

V 越大，工资越高。

C 越大，工资越高。

C 越大，雇员越容易分享到剩余价值。

越是难以被监督的职业越需要职业道德。

$$\pi_{\text{雇主}} = (-W) \beta + (V - W)(1 - \beta) = V - W - V\beta = V - W - VC/W$$

$$\pi_{\text{雇员}} = W - H$$

特别提示：

选一个容易偷懒的职业，并努力不偷懒，你离成功就不远了！

特别提示：

成功的路上并没有你想象的那么拥挤！

谢 谢

THANK YOU