arXiv:2304.05024v1 [cs.DM] 11 Apr 2023

简单折扣决斗的纳什均衡研究

Athanasios Kehagias

2023年4月12日

摘要

我们将一个双人静态决斗 博弈构建并研究为非零和折扣随机博弈。 玩家 P_1， P_2 原地站立，在每个回合中，一名或两名玩家可以向“另一名”玩家开枪。如果 P_n 向 P_m 开枪（$m \neq n$），则要么击中并杀死对方（概率为 p_n），要么未击中对方，且 P_m 不受影响（概率为 $1 - p_n$）。这一过程持续进行，直到至少一名玩家死亡；如果始终无人死亡，博弈将持续无限多个回合。每名玩家在自己存活的每个回合都获得单位收益；杀死对手不产生收益。我们证明，始终射击策略是一个纳什均衡，此外，该博弈还存在采用平稳策略和非平稳策略的“合作型”（即不射击）纳什均衡。文中还指出并讨论了它与重复囚徒困境之间的某种相似性。

1 引言

本文研究一种按回合进行的双人静态决斗 博弈。玩家 P_1， P_2 站在原地，在每个回合中，一名或两名玩家可能 向“另一名”玩家射击。如果 P_n 向 P_m射击（$m \neq n$），那么要么击中并杀死对方，要么未击中对方且 P_m 不受影响；相应的概率分别为 p_n和 $1-p_n$。该过程持续进行，直到至少一名玩家死亡；也可能没有人死亡，博弈持续无限多个回合。我们将上述过程表述为一种非零和折扣随机博弈。 博弈规则和玩家的收益函数将在下一节中介绍。

关于静态决斗的研究寥寥无几。据我们所知，它实际上仅被作为研究“静态三人决斗”的初步步骤；在这种决斗中，三名平稳玩家彼此射击。关于静态三人决斗的早期研究有 [11, 16, 18, 19, 20] ，其中设定的博弈规则保证恰好有一名幸存者（“赢家”）。 [17]提出了更一般的分析，探讨玩家之间“合作”的可能性。 [13, 14, 15, 25]进一步研究了这一观点。关于三人决斗的近期论文包括[1, 2, 3, 5, 6, 7, 8, 9, 22, 23, 24]。 [1]

[1]我们还要指出，关于一种截然不同的对决博弈类型，已有大量文献；这类博弈本质上属于时机博弈 [4, 10, 12]。不过，这些文献与本文研究的博弈无关。

虽然上述论文聚焦于静态三人决斗的各种形式，但我们认为，静态决斗本身就很有趣，却未得到应有的关注。特别是，我们将证明，在我们的表述下，静态决斗与重复囚徒困境有某种相似性，并且在非平稳策略下存在“合作型”纳什均衡。

本文结构如下。第 2 节将对博弈进行严格定义。第 3 节将介绍若干平稳和非平稳策略，并计算它们的期望收益。第 4 节将证明，先前定义的某些策略组合构成纳什均衡。第 5 节将讨论所得结果，以及静态决斗与重复囚徒困境之间的联系。最后，第 6 节将总结我们的结果，并提出一些未来的研究方向。

2 博弈

博弈涉及玩家 P_1、P_2，并以离散的时间步（回合） $t \in \{1,2,...\}$进行。在时间 t 的状态 为

$$s(t) = (s_1(t), s_2(t)) \in S = \{(1,1),(1,0),(0,1),(0,0),(\tau,\tau)\}.$$

对于 $n \in \{1,2\}$， s_n （t）是 P_n在 $t \in \{0,1,2,...\}$ 时的状态，可以是

> 1：当 P_n 存活时；0：当 P_n 在当前回合死亡时；τ：当一名或两名玩家在之前的回合中死亡时。

P_n在 $t \in \{1,2,...\}$ 的行动是 $f_n(t)$，可以是1（P_n 开枪）或0（P_n 不开枪）。如果 $s_n(t-1) \neq 1$，P_n 不能向 t 开枪，且 $f_n(t)$必须等于0；如果 $s_n(t-1) = 1$，那么 $f_n(t)$可以是0或1。当 $f_n(t) = 1$时， $s_{-n}(t) = 0$ （即 P_{-n} 死亡[2]）的概率为 $p_n \in (0,1)$，而$s_{-n}(t) = 1$ 的概率为 $1 - p_n$。我们设 $f(t) = (f_1(t)\ f_2(t))$，且$\mathbf{p} = (p_1, p_2)$。注意，我们假设 p_1和 p_2 都不同于0和1。

博弈从初始状态 $s(0)$ 开始；显然，我们感兴趣的主要情形是 $s(0) = (1,1)$。在 $t \in \{1,2,...\}$ 时，玩家同时选择动作 $f_1(t)$、 $f_2(t)$，博弈根据条件状态转移概率$\Pr(s(t) | s(t-1)\ f(t))$转移到状态 $s(t)$。图1给出了状态转移图，其中依赖于动作的转移概率标注在边旁；很容易验证这些概率符合博弈规则。图中显示，从 (1,1) 开始的博弈会持续无限轮，并且存在两种可能。

1.博弈始终处于 (1,1)（不会有玩家被杀死）。

2.在某个 t' 博弈移动到状态 $s \in \{(1,0),(0,1),(0,0)\}$ （一名或两名玩家被杀死）时，并且在 $t'+1$ 博弈移动到终止状态（τ,τ），此后永远停留在那里。

[2]在续篇中，我们使用标准的博弈论记号，据此 $s_{-1} = s_2, s_{-2} = s_1$。玩家、行动等也使用相同的记号。

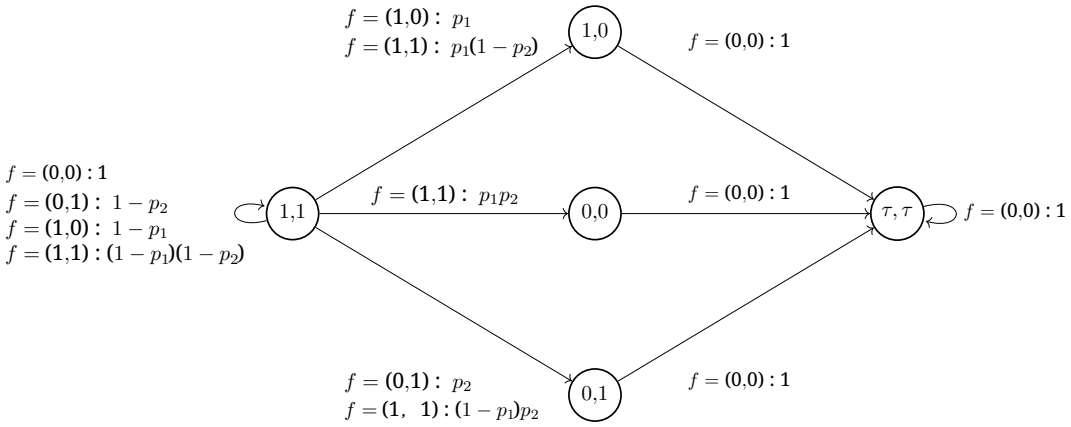

图1：静态决斗博弈的状态转移图。

A 有限历史 是一个序列 $h = s(0)\ f(1)\ s(1)\ ...f(T)\ s(T)$；无限历史 是一个$h = s(0)\, f(1)\, s(1)...$。可允许 历史是符合博弈规则的历史；所有可允许的有限（分别为无限）历史的集合记为 H （分别为 H^∞）。对于每个历史h 以及 $n \in \{1, 2\}$，我们将 P_n 的总收益函数 定义为

$$Q_n(h) = \sum_{t=0}^{\infty} \gamma^t q_n(s(t))$$

其中 $\gamma \in (0,\ 1)$ 是折扣因子，而 $q_n :\ S \to \mathbb{R}$ 是 P_n 的阶段收益函数，定义如下。

$$\begin{array}{ll} q_1(\tau,\tau) = 0, & q_2(\tau,\tau) = 0, \\ q_1(1,1) = 1, & q_2(1,1) = 1, \\ q_1(1,0) = \frac{1}{1-\gamma}, & q_2(1,0) = 0, \\ q_1(0,1) = 0, & q_2(0,1) = \frac{1}{1-\gamma}, \\ q_1(0,0) = 0, & q_2(0,0) = 0. \end{array}$$

上述数值表明，每位玩家在其存活的每个回合中都会获得一个收益单位；收益 $q_1(1,0) = \frac{1}{1-\gamma}$ 包含了无限收益序列 $\sum_{t=0}^{\infty} \gamma^t 1 = \frac{1}{1-\gamma}$（当 P_1 杀死 P_2 并在之后无限多个回合中存活时，就会出现这种情况）。请注意，玩家杀死对手不会直接获得收益，但可以间接获益，因为这样就消除了自己被杀的可能性。

一个策略对于 P_n 是一个函数 $\sigma_n :\ H \to [0, 1]$，它为每个有限历史 h对应概率

$$x_n = \sigma_n(h) = \Pr(\text{“}P_n \text{ shoots at } P_{-n}\text{”}).$$

一项平稳策略 是一个仅取决于当前状态 s的 σ_n，因此我们简单记作$x_n = \sigma_n(s)$。一个策略组合是一个向量 $\sigma = (\sigma_1, \sigma_2)$。我们将所有可容许

策略（即符合博弈规则的策略）记为 Σ，所有可容许平稳策略的集合记为 Σ。

给定初始状态 $s(0)$ 以及分别由 P_1 和 P_2 采用的策略 σ_1 和 σ_2，在所有无限历史的集合上定义了一个概率测度。由于 $\gamma \in (0,1)$，总期望收益

$$\forall n \in \{1,2\} : \overline{Q}_n\left(s\left(0\right), \sigma_1, \sigma_2\right) = \mathbb{E}\left(Q_n\left(h\right) | s\left(0\right), \sigma_1, \sigma_2\right)$$

定义良好。

因此，我们将同时静态决斗表述为折扣随机博弈，记作 $\Gamma(s(0)\ \gamma, \mathbf{p})$，或者简写为 $\Gamma(s(0)\ \gamma)$，其中(p_1, p_2) 是固定的。我们主要关注<丈量非零和 博弈 $\Gamma((1,1)\ \gamma)$。我们假设 P_1 和 P_2 试图达到<丈量纳什均衡 （NE），即满足以下条件的策略组合$(\widehat{\sigma}_1, \widehat{\sigma}_2)$:

$$\forall n \in \{1,2\} : \forall \sigma_n \in \Sigma : \overline{Q}_n\left((1,1), \widehat{\sigma}_n, \widehat{\sigma}_{-n}\right) \geq \overline{Q}_n\left((1,1), \sigma_n, \widehat{\sigma}_{-n}\right).$$

3 一些基本策略及其收益 S

在本节中，我们将介绍几种策略，以供后续探讨纳什均衡时使用。

3.1 平稳策略 σ^S

当 $P_n(n \in \{1,2\})$ 使用平稳可容许策略 σ_n 时，对于 $s \in \{(1,\ 0),\ (0,\ 1),\ (0,\ 0),\ (\tau,\ \tau)\}$，有 $\sigma_n(s) = 0$，而 σ_n 完全由值 $\sigma_n(1,\ 1) = x_n$确定。因此，我们有时会用 $Q_n((1,\ 1)\ x_1,\ x_2)$ 代替 $Q_n((1,\ 1)\ \sigma_1, \sigma_2)$。

令 $V_1^S(x_1,\ x_2) = \overline{Q_1}((1,\ 1)\ x_1,\ x_2)$；为简便起见，我们也将其简称为 V_1^S。那么 V_1^S满足方程

$$\begin{aligned} V_1^S = 1 + \gamma p_1 x_1 \left(x_2\left(1-p_2\right) + \left(1-x_2\right)\right) \frac{1}{1-\gamma} \\ + \gamma\left(x_1\left(1-p_1\right) + \left(1-x_1\right)\right)\left(x_2\left(1-p_2\right) + \left(1-x_2\right)\right) V_1^S, \end{aligned} \tag{1}$$

可通过以下推理得到。当博弈处于状态 $s = (1,\ 1)$ 时，P_1 的期望收益为当前状态的一单位收益，加上后续状态 s' 的折扣期望收益；后续状态有以下几种可能。

1. $s' = (1,0)$，当 P_1 开枪并命中 P_2，且 P_2 要么开枪但未命中，要么不开枪；相应的概率为 $p_1 x_1\left(x_2(1-p_2) + (1-x_2)\right)$。该情况下的总期望收益为 $Q_1((1,0)\ x_1, x_2) = \frac{1}{1-\gamma}$。

2. $s' = (1,1)$，当双方 P_1 和 P_2 要么开枪但未命中，要么不开枪；相应的概率为 $(x_1(1-p_1) + (1-x_1))(x_2(1-p_2) + (1-x_2))$。在这种情况下，我们回到了初始状态 $(1,1)$，额外的总期望收益又是$Q_1((1,1)\ x_1,\ x_2)$。

3. 我们还可能转移到 (0, 1) 和 (0,0)，但这会使 P_1 获得零收益，因此在 (1) 中未予考虑。

求解 (1)，经过一些代数计算[3], 可得 that

$$V_1^S(x_1, x_2) = \frac{1-\gamma(1-p_1x_1(1-p_2x_2))}{(1-\gamma)(1-\gamma(1-p_1x_1)(1-p_2x_2))}. \tag{2}$$

在下文中，我们将经常使用归一化的总期望收益。在这种情况下，它是

$$v_1^S(x_1, x_2) = (1-\gamma)\overline{Q}_1((1,1), x_1, x_2) = \frac{1-\gamma(1-p_1x_1(1-p_2x_2))}{1-\gamma(1-p_1x_1)(1-p_2x_2)} \tag{3}$$

$V_2^S(x_1, x_2) = Q_2((1, 1) \overline{x_1}, x_2)$ 和 $v_2^S(x_1, x_2) = (1-\gamma)Q_2((1, 1) \overline{x}_1, x_2)$ 的公式可通过在 (2)-(3) 中交换索引 1 和 2 得到。

3.2 合作策略 σ^C

静态的"合作型"（该名称将在第5节中得到解释）策略 σ^C 被定义为

$\sigma^C(1,1) = 0$，这意味着玩家从不射击。

显然， σ^C是 σ^S，其中 $x_n = 0$和$Q_n\left((1, 1) \sigma^C, \sigma^C\right) = Q_n((1, 1), 0, 0)$。因此，通过在 (2) 中设定 $x_1 = x_2 = 0$,一我们得到$V_1^C = Q_1\left((1, 1) \sigma^C, \sigma^C\right) = Q_1((1, 1), 0, 0)$。由此可得期望收益和归一化总期望收益为

$$V_1^C = \frac{1}{1-\gamma}, \qquad v_1^C = (1-\gamma)V_1^C = 1$$

$V_2^C = Q_2\left(\sigma^C, \sigma^C\right), \overline{v}_2^c = (1-\gamma)$ V_2^C的公式可通过交换上述公式中的索引 1 和 2 得到。

3.3 背叛策略 σ^D

平稳的"背叛"策略 σ^D （其名称将在第 5 节中说明）定义为

$\sigma^D(1,1) = 1$，这意味着玩家总是以概率为一开枪。

显然， σ^D是 σ^S，其中 $x_n = 1$且 $Q_n\left((1, 1) \sigma^D, \sigma^D\right) = Q_n((1,1), 1, 1)$。因此，在（2）中令 $x_1 = x_2 = 1$，即可得到$V_1^D = Q_1\left((1, 1) \sigma^D, \sigma^D\right) = Q_1((1, 1), 1, 1)$。于是，总期望收益及其归一化后的值为

$$V_1^D = \frac{1-\gamma(1-p_1(1-p_2))}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}$$
$$v_1^D = (1-\gamma)V_1^D(\gamma) = \frac{1-\gamma(1-p_1(1-p_2))}{1-\gamma(1-p_1)(1-p_2)}$$

通过在上述方程中交换索引 1 和 2，可得到 $V_2^D = Q_2\left(\sigma^D, \sigma^D\right), v_2^D = (1-\gamma)V_2^D$的公式。

[3]求得该解所需的计算由计算机代数系统 Maple 完成，并经人工验证。本文其余部分所需的其他计算（有时相当复杂）也是如此。

3.4 提前射击策略 $\sigma^{DC(K)}$

非平稳的"早射"策略 $\sigma^{DC(K)}$ 是只在时间 1、2、..... K 开枪（概率为一），其中 K 是该策略的参数。令 $V^{DC\ (K)\ 1} = Q_1\left((1,\ 1)\ \ \sigma^{DC(K)},\ \ \sigma^{DC(K)}\right)$；则 $V_1^{DC(K)}$ 满足方程：

$$
\begin{aligned}
V_1^{DC(K)} &= 1 \\
&+ \gamma\left(p_1\left(1-p_2\right)\frac{1}{1-\gamma} + \left(1-p_1\right)\left(1-p_2\right)\right) \\
&+ \gamma^2\left(\left(1-p_1\right)p_1\left(1-p_2\right)^2\frac{1}{1-\gamma} + \left(1-p_1\right)^2\left(1-p_2\right)^2\right) \\
&+ ... \\
&+ \gamma^{K-1}\left(\left(1-p_1\right)^{K-2}p_1\left(1-p_2\right)^{K-1}\frac{1}{1-\gamma} + \left(1-p_1\right)^{K-1}\left(1-p_2\right)^{K-1}\right) \\
&+ \gamma^{K}\left(\left(1-p_1\right)^{K-1}p_1\left(1-p_2\right)^{K}\frac{1}{1-\gamma} + \left(1-p_1\right)^{K}\left(1-p_2\right)^{K}V_1^C\right)
\end{aligned}
\tag{4}
$$

这个方程的合理性如下所述。

在时间 $t=0$，P_1 获得一单位收益。

在时间 $t=1$、..... $K-1$，期望收益如下。

(a) 以概率$(1-p_1)^{t-1}p_1(1-p_2)^t$，$P_1$ 在时间 $t'=1,...,t-1$ 未命中，并在时间 t'成-功，而 P_2 在时间 $t'=1,...,t$未命中。在这种情况下，P_1 收到收益$1^{1-\gamma}$。(b) 以概率$(1-p_1)^t(1-\overline{p_2})^t$，$P_1$ 和 P_2 都在时间 $t'=1,...,t$未命中。在这种情况下，P_1 收到收益 1。(c) 所有其他可能性都会产生零收益，因此未将其纳入方程。

3. 在时间 $t=K$，期望收益如下。

(a) 以概率$(1-p_1)^{K-1}p_1(1-p_2)^K$，$P_1$ 在时间 $t'=1$,...，$K-1$ 成功于时间 $t'=K$，而 P_2 在时间 $t'=1$, ...，K未命中。在这种情况下，P_1 获得收益 $\frac{1}{1-\gamma}$。

(b)以概率$(1-p_1)^K(1-p_2)^K$，P_1 和 P_2 均在时间 $t'=1,...,K$未命中。在这种情况下，双方玩家在后续时间内将永远不会射击，因此我们回到了初始状态$(1,1)$，且额外的总期望收益为 $V_1^C = Q_1\left((1,1)\ \ \sigma^C,\sigma^C\right)$。

(c) 其他所有可能情况的收益均为零，因此不纳入方程。

将 V_1^C 的表达式代入 (4)，我们得到总期望收益和归一化总期望收益的以下表达式：

$$V_1^{DC(K)} = \frac{1+\gamma^{K+1}(1-p_1)^K(1-p_2)^K p_2 - \gamma(1-p_1+p_1p_2)}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}$$
$$v_1^{DC(K)} = (1-\gamma)V_1^{DC(K)} = \frac{1+\gamma^{K+1}(1-p_1)^K(1-p_2)^K p_2 - \gamma(1-p_1+p_1p_2)}{1-\gamma(1-p_1)(1-p_2)}$$

通过在上述方程中交换索引 1 和 2，即可得到 $V_2^{DC(K)}$，$v_2^{DC(K)}$ 的公式。

3.5 延迟射击策略 $\sigma^{CD(K)}$

非平稳的“延迟射击”策略 $\sigma^{CD(K)}$ 是指仅在时间 K、$K+1$.....射击（概率为一），其中 K 是该策略的一个参数。令 $V^{CD}{}_{(K)1} = Q_1\left((1,\ 1)\ \sigma^{CD(K)},\ \sigma^{CD(K)}\right)$；则 $V_1^{CD(K)}$ 满足以下方程：

$$V_1^{CD(K)} = 1+\gamma+...+\gamma^{K-1}+\gamma^K\left(p_1(1-p_2)\frac{1}{1-\gamma}+(1-p_1)(1-p_2)V_1^D\right) \tag{5}$$

该方程的依据如下。

1. 在时间 $t=0$、1、.....、$K-1$，P_1 获得一单位的折现收益。

2. 在时间 $t=K$，有以下几种可能。

   (a) 以概率 $p_1(1-p_2)$，P_1 命中 P_2，而 P_2 未命中。在此情况下，P_1 获得收益 $\frac{1}{1-\gamma}$。
   (b) 以概率 $(1-p_1)(1-p_2)$，P_1 和 P_2 均未命中。在此情况下，P_1 和 P_2均恢复为策略 σ^D，P_1 获得总期望收益 $V_1^D = Q_1\left((1,1)\ \sigma^D,\sigma^D\right)$。(c) 所有其他可能性均产生零收益，因此不包含在方程中。

将先前得到的 V_1^D 表达式代入（5），我们得到总期望收益和归一化总期望收益的以下表达式：

$$V_1^{CD(K)} = \frac{1-\gamma^K p_2 - \gamma(1-p_1)(1-p_2)}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}$$
$$v_1^{CD(K)} = (1-\gamma)V_1^{CD(K)} = \frac{1-\gamma^K p_2 - \gamma(1-p_1)(1-p_2)}{1-\gamma(1-p_1)(1-p_2)}$$

$V_2^{CD(K)}$、$v_2^{CD(K)}$ 的公式可通过在上述公式中交换下标1和2得到。

3.6 周期性射击策略 $\sigma^{P}(M)$

非平稳"周期射击"策略 $\sigma^{P(M)}$是只在时间 $M+1, 2M+2, ...$开火，其中 M是策略参数。令 $V^{P}{}_{(M)1} = Q_1\left((1,1)\ \sigma^M, \sigma^M\right)$; 通过与前几种情况类似的推理，我们可知 $V^{P}{}_{(M)\,1}$ 满足方程

$$\begin{aligned} V_1^{P(M)} &= 1 + \gamma + ... + \gamma^M \\ &\quad + \gamma^{M+1}\left(p_1\left(1-p_2\right)\frac{1}{1-\gamma} + \left(1-p_1\right)\left(1-p_2\right)V_1^{P(M)}\right) \end{aligned} \tag{6}$$

解方程 (6)，可得总期望收益和归一化总期望收益的表达式如下:

$$\begin{aligned} V_1^{P(M)} &= \frac{1-\gamma^{M+1}\left(1-p_1\left(1-p_2\right)\right)}{\left(1-\gamma\right)\left(1-\gamma^{M+1}\left(1-p_1\right)\left(1-p_2\right)\right)} \\ v_1^{P(M)} &= \frac{1-\gamma^{M+1}\left(1-p_1\left(1-p_2\right)\right)}{1-\gamma^{M+1}\left(1-p_1\right)\left(1-p_2\right)} \end{aligned}$$

$V^{P}{}_{(M)\,2}$， $v_2^{P(M)}$ 的公式可通过在上述公式中交换索引 1 和 2 得到。

3.7 冷酷策略

任何 P_n 的策略 σ_n 都可用于定义相应的冷酷策略 $\widetilde{\sigma}_n$，如下所示:

> $\widetilde{\sigma}_n$: 只要 P_{-n} 使用 σ_n， P_n 也使用 σ_n; 如果在第 t 个回合， P_{-n} 偏离 σ_n，则 P_n 在此后所有回合中都使用 σ^D 。

例如，冷酷合作策略 $\widetilde{\sigma}^C$ 规定: 只要 P_{-n}也不射击， P_n 就永不射击; 如果在第 t 回合 P_{-n} 开枪，那么从第 $t+1$ 回合开始， P_n将始终以概率为一开枪。类似地，我们可以得到:

1. 冷酷背叛策略 $\widetilde{\sigma}^D$ （它与 σ^D 相同）， 2. 冷酷早射策略 $\widetilde{\sigma}^{DC(K)}$， 3. 冷酷晚射策略 $\widetilde{\sigma}^{CD(K)}$， 4. 冷酷周期性射击策略 $\widetilde{\sigma}^{P(M)}$。

4 纳什均衡

现在我们将给出一系列命题; 每个命题都表明，某一策略对是（平稳或非平稳）纳什均衡; 有时，这一结论仅在 γ 以及可能的 p_1， p_2 取值范围内成立。

命题 4.1 对于每个 $\gamma \in (0,1)$， $\Gamma\left((1,1)\ \gamma\right)$ 的唯一平稳纳什均衡是 $\left(\sigma^C, \sigma^C\right)$ 和 $\left(\sigma^D, \sigma^D\right)$。

证明。 假设 P_1 （相应地， P_2）采用平稳策略 σ^S，其中 $\sigma^S(1,1) = x_1$ （相应地，σ^S 其中 $\sigma^S(1,1) = x_2$ ）。那么 P_1 的收益为

$$V_1^S(x_1, x_2) = \frac{1 - \gamma(1 - p_1 x_1(1 - p_2 x_2))}{(1-\gamma)(1 - \gamma(1 - p_1 x_1)(1 - p_2 x_2))}.$$

现在假设 P_1 转而选择 $\sigma_S(1, 1) = y_1$；他的收益变为

$$V_1^S(y_1, x_2) = \frac{1 - \gamma(1 - p_1 y_1(1 - p_2 x_2))}{(1-\gamma)(1 - \gamma(1 - p_1 y_1)(1 - p_2 x_2))}.$$

我们来看归一化收益之差

$$\begin{aligned}
\delta v_1 &= (1-\gamma)\left(Q_1((1,1), x_1, x_2) - Q_1((1,1), y_1, x_2)\right) \\
&= \frac{1 - \gamma(1 - p_1 x_1(1 - p_2 x_2))}{1 - \gamma(1 - p_1 x_1)(1 - p_2 x_2)} - \frac{1 - \gamma(1 - p_1 y_1(1 - p_2 x_2))}{1 - \gamma(1 - p_1 y_1)(1 - p_2 x_2)} \\
&= \frac{\gamma^2 p_1 p_2 x_2 (x_1 - y_1)(1 - p_2 x_2)}{((1 - \gamma(1 - p_2 x_2)(1 - p_1 y_1))(1 - \gamma(1 - p_2 x_2)(1 - p_1 x_1)))}
\end{aligned}$$

现在， P_1 没有动机从 x_1 转向 y_1，当且仅当 $\delta v_1 \geq 0$，这等价于

$$\gamma^2 p_1 p_2 x_2 (x_1 - y_1)(1 - p_2 x_2) \geq 0$$

类似地， P_2 没有动机从 x_2 转向 y_2，当且仅当

$$\gamma^2 p_1 p_2 x_1 (x_2 - y_2)(1 - p_1 x_1) \geq 0$$

因此，对于 $n \in \{1, 2\}$，以下结论成立。

1. 如果 $(x_1, x_2) = (0, 0)$， P_n 没有改变 x_n的动机；$(x_1, x_2) = (0, 0)$；因此 $\left(\sigma^C, \sigma^C\right)$ 是纳什均衡。

2. 如果 $(x_1, x_2) = (1,1)$， P_n 没有动力改变 x_n；$(x_1, x_2) = (1, 1)$；因此 $\left(\sigma^D, \sigma^D\right)$ 是一个纳什均衡。

3. 如果 $(x_1, x_2) \in (0,1) \times (0,1)$，那么它不可能是纳什均衡，因为 P_n 有动机（单方面）从 x_n 改为 1。

证明至此完成。 ■

现在我们将开始考察由冷酷策略组合得到的纳什均衡。

命题 4.2 对于每个 $\gamma \in (0, 1)$， $\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right)$ 都是 $\Gamma((1, 1) \gamma)$ 的纳什均衡。

证明。假设 P_1 和 P_2 都使用 $\widetilde{\sigma}^C$。那么 P_1的收益是

$$V_1^C = Q_1\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right) = Q_1\left(\sigma^C, \sigma^C\right) = \frac{1}{1-\gamma}.$$

现在假设 P_1 偏离了 $\widetilde{\sigma}^C$。只需考察 P_1 在 $t = 1$时偏离的情况；此外，在 P_1 偏离后（即从 $t = 2$开始）， P_2 将改用 σ^D，而 P_1 没有理由不在任何 $t \geq 2$开枪。因此， P_1 实质上采用的是策略 $\sigma_1 = \sigma^D$，其总期望收益将为

$$\begin{aligned} V_1 = Q_1\left(\sigma^D, \widetilde{\sigma}^C\right) &= 1 + \gamma\left(p_1 \frac{1}{1-\gamma} + (1-p_1)\left(1 + \gamma Q_1\left(\sigma^D, \sigma^D\right)\right)\right) \\ &= 1 + \gamma\left(p_1 \frac{1}{1-\gamma} + (1-p_1)\left(1 + \gamma \frac{1-\gamma\left(1-p_1\left(1-p_2\right)\right)}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}\right)\right) \\ &= \frac{1 - p_2(1-p_1)\gamma^3 - (1-p_2)(1-p_1)\gamma}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)} \end{aligned}$$

Now

$$\begin{aligned} \left(1-\gamma\right)\left(Q_1\left(\widetilde{\sigma}^C, \widetilde{\sigma}^C\right) - Q_1\left(\sigma^D, \widetilde{\sigma}^C\right)\right) &= \left(1-\gamma\right)\left(V_1^C - V_1\right) \\ &= 1 - \frac{1 - p_2(1-p_1)\gamma^3 - (1-p_2)(1-p_1)\gamma}{1-\gamma\left(1-p_1\right)\left(1-p_2\right)} \\ &= \frac{\gamma^3 p_2\left(1-p_1\right)}{1-\gamma\left(1-p_1\right)\left(1-p_2\right)} > 0 \end{aligned}$$

因此， P_1 没有理由偏离 $\widetilde{\sigma}^C$。对 P_2也可以证明同样的结论。因此， $\widetilde{\sigma}^C\big)$ 是一个纳什均衡。

■

In t在下一个命题中，只有当"足够大"时，该策略组合才是纳什均衡 γ.

命题 4.3 存在某个 $\gamma_0 \in$ （**0**，**1**），使得：对于所有 $\gamma \in$ （γ_0，**1**）以及所有 $K \in \mathbb{N}$、$\left(\widetilde{\sigma}^{DC(K)},\ \widetilde{\sigma}^{DC(K)}\right)$ 都是 Γ （（**1**，**1**） γ）的纳什均衡。

证明。 回想一下，当双方玩家都使用 $\widetilde{\sigma}^{DC(K)}$ 时， P_1 获得收益

$$V_1^{DC(K)} = Q_1\left((1,1), \widetilde{\sigma}^{DC(K)}, \widetilde{\sigma}^{DC(K)}\right) = \frac{1 + \gamma^{K+1}(1-p_1)^K(1-p_2)^K p_2 - \gamma(1-p_1+p_1p_2)}{\left(1-\gamma\right)\left(1-\gamma\left(1-p_1\right)\left(1-p_2\right)\right)}$$

让我们证明， P_1 没有动机采用偏离策略 σ_1。

[4]这是以下事实的推论，我们将在本文余下部分经常用到这一事实。如果 P_n 从某个时间 t 开始采用平稳策略 σ_n，那么 P_{-n} 的最佳响应也会是平稳策略。这是因为，对于固定的平稳 σ_n， P_{-n} 必须求解一个 马尔可夫决策过程，而该问题的最优策略是平稳的。更多详情请参见 [21]。

1. **情况 I：** 我们先考虑在时间 $t \in \{K+1, K+2, ...\}$ 进行偏离的策略；也就是说，在博弈进入禁射阶段后才开枪。实际上，我们只需考虑会在 $t = K+1$ 开枪且以概率为一开枪的 σ_1 。在这种情况下

$$\begin{aligned} Q_1\left((1,1), \widetilde{\sigma}^{DC(K)}, \widetilde{\sigma}^{DC(K)}\right) &= A + \gamma^{K+1} Q_1\left((1,1), \widetilde{\sigma}^{DC(K)}, \widetilde{\sigma}^{DC(K)}\right) \\ &= A + \gamma^{K+1} Q_1\left((1,1), \sigma^C, \sigma^C\right) = A + \gamma^{K+1} V_1^C \\ Q_1\left((1,1), \sigma^1, \widetilde{\sigma}^{DC(K)}\right) &= A + \gamma^{K+1} Q_1\left((1,1), \sigma^1, \widetilde{\sigma}^{DC(K)}\right) \\ &= A + \gamma^{K+1} Q_1\left((1,1), \sigma^1, \widetilde{\sigma}^{DC(K)}\right) \end{aligned}$$

其中， A 是对时间 $t \in \{0, ..., K\}$ 求和得到的期望收益，对于 P_1采用的两种策略而言，该收益相同。现在，出于通常的原因， P_1 会继续向 $t \in \{K+2, K+3, ...\}$ 射击，因此我们有

$$\begin{aligned} V_1 &= Q_1\left((1,1), \sigma^1, \widetilde{\sigma}^{DC(K)}\right) = p_1 \frac{1}{1-\gamma} + (1-p_1)\left(1 + \gamma Q_1\left(\sigma^D, \sigma^D\right)\right) \\ &= p_1 \frac{1}{1-\gamma} + (1-p_1)\left(1 + \gamma \frac{1-\gamma(1-p_1(1-p_2))}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}\right) \\ &= \frac{1 - p_2(1-p_1)\gamma^2 - (1-p_2)(1-p_1)\gamma}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))} \end{aligned}$$

于是我们有

$$\begin{aligned} &(1-\gamma)\left(Q_1\left((1,1), \widetilde{\sigma}^C, \widetilde{\sigma}^C\right) - Q_1\left((1,1), \sigma^D, \widetilde{\sigma}^C\right)\right) \\ &= (1-\gamma)\left(V_1^C - V_1\right) \\ &= 1 - \frac{1 - p_2(1-p_1)\gamma^2 - (1-p_2)(1-p_1)\gamma}{1-\gamma(1-p_1)(1-p_2)} \\ &= \frac{p_2(1-p_1)\gamma^2}{1-\gamma(1-p_1)(1-p_2)} > 0 \end{aligned}$$

因此， P_1 没有理由向 $t > K$射击。

2. **情况 II：** 接下来我们考虑在 $t \in \{1, 2, ..., K\}$ 时偏离的策略，即在射击阶段不开火。同样，在第一次偏离之后， P_1 没有理由不开火。因此，我们只需考虑满足以下条件的策略： σ_1 （a）在某个 $t = L \in \{1, 2, ..., K\}$ 时不开火；（b）在所有 $t \in \{1, 2, ..., L-1, L+1, ...\}$ 时都开火。接着，按照通常的

论证,

$$
\begin{aligned}
V_1 &= Q_1\left((1,1),\sigma_1,\widetilde{\sigma}^{DC(K)}\right)\\
&= 1\\
&+\gamma\left(p_1(1-p_2)\frac{1}{1-\gamma}+(1-p_1)(1-p_2)\right)\\
&+\gamma^2\left((1-p_1)p_1(1-p_2)^2\frac{1}{1-\gamma}+(1-p_1)^2(1-p_2)^2\right)\\
&+...\\
&+\gamma^{L-1}\left((1-p_1)^{L-2}p_1(1-p_2)^{L-1}\frac{1}{1-\gamma}+(1-p_1)^{L-1}(1-p_2)^{L-1}\right)\\
&+\gamma^{L}\left((1-p_1)^{L-1}p_1(1-p_2)^{L}\frac{1}{1-\gamma}+(1-p_1)^{L}(1-p_2)^{L}\left(1+\gamma V_1^D\right)\right)\\
&= 1+\sum_{k=1}^{L-1}\gamma^k\left((1-p_1)^{k-1}p_1(1-p_2)^k\frac{1}{1-\gamma}+(1-p_1)^k(1-p_2)^k\right)\\
&+\gamma^L(1-p_1)^{L-1}(1-p_2)^L p_1\frac{1}{1-\gamma}\\
&+\gamma^L(1-p_1)^L(1-p_2)^L\left(1+\gamma\frac{1-\gamma(1-p_1(1-p_2))}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}\right)
\end{aligned}
$$

Let

$$
\begin{aligned}
\delta v_1(\gamma) &= (1-\gamma)\left(Q_1\left((1,1),\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right)-Q_1\left((1,1),\sigma_1,\widetilde{\sigma}^{DC(K)}\right)\right)\\
&= V_1^{DC(K)}-V_1.
\end{aligned}
$$

注意， $\delta v_1(\gamma)$ 对所有 $\gamma\in[0,1]$ 都有明确定义且连续，因为因子 $(1-\gamma)$ 抵消了分母中的 $(1-\gamma)$ 因子。

$$
Q_1\left((1,1),\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right)-Q_1\left((1,1),\sigma_1,\widetilde{\sigma}^{DC(K)}\right).
$$

经过大量代数运算[5] 我们发现,

$$
\delta v_1(1)=\frac{(1-p_1)^K(1-p_2)^K p_2}{p_1+(1-p_1)p_2}>0.
$$

由于 $\delta v_1(\gamma)$ 是连续的，因此存在某个 $\gamma_0\in(0,1)$，使得对于每个 $\gamma\in(\gamma_0,1)$ 以及每个 $K\in\mathbb{N}$， $\delta v_1(\gamma)$都为正。因此，对于这些取值， P_1 在射击阶段没有偏离的动机。

综合情况 I和II，我们看到 P_1 没有动机偏离 $\widetilde{\sigma}^{DC(K)}$。同理可证明， P_2也是如此。因此 $\left(\widetilde{\sigma}^{DC(K)},\widetilde{\sigma}^{DC(K)}\right)$ 是一个纳什均衡。 ■

接下来我们给出一个否定结果：双方延迟射击不是纳什均衡。

[5]再次使用 Maple。

命题4.4 对于每个 $\gamma \in (0,1)$ 以及每个 $K \in \mathbb{N}$, $\left(\widetilde{\sigma}^{CD(K)}, \widetilde{\sigma}^{CD(K)}\right)$ 都不是$\Gamma((1,1) \ \gamma)$的纳什均衡。

证明。 回想一下

$$V_1^{CD(K)} = Q_1\left(\widetilde{\sigma}^{CD(K)}, \widetilde{\sigma}^{CD(K)}\right) = \frac{1-\gamma^K p_2 - \gamma(1-p_1)(1-p_2)}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}$$

我们只需证明， P_1 有一个有利可图的偏离策略 σ_1。令 σ_1 为：不向 $t \in \{1,2,...,K-2\}$射击，而向 $t \in \{K-1, K, ...\}$射击；换言之，在射击阶段开始前一个回合就开始射击。接着，根据通常的论证，P_1的收益为

$$\begin{aligned} V_1 = Q_1\left((1,1), \sigma_1, \widetilde{\sigma}^{CD(K)}\right) &= \sum_{k=0}^{K-2} \gamma^k + \gamma^{K-1}\left(p_1 \frac{1}{1-\gamma} + (1-p_1)(1+\gamma V_D)\right) \\ &= \frac{\left(-\gamma^{K+1} p_2(1-p_1) + 1 - \gamma(1-p_2)(1-p_1)\right)}{(1-\gamma)(1-\gamma(1-p_2)(1-p_1))} \end{aligned}$$

经过适当代换和代数计算，我们得到

$$\begin{aligned} \delta v_1 &= (1-\gamma)\left(Q_1\left(\widetilde{\sigma}^{CD(K)}, \widetilde{\sigma}^{CD(K)}\right) - Q_1\left(\sigma_1, \widetilde{\sigma}^{CD(K)}\right)\right) \\ &= -\frac{p_2 \gamma^K (1-\gamma(1-p_1))}{(1-\gamma(1-p_2)(1-p_1))} < 0 \end{aligned}$$

由于 P_1 有动机转而采用 σ_1 ， 而且 $\left(\widetilde{\sigma}^{CD(K)}, \ \widetilde{\sigma}^{CD(K)}\right)$ 不是纳什均衡。 ■

He

命题 4.5 对于每个 $M \in \mathbb{N}$，都存在一个 $\delta_M > 0$ ，使得：如果

$$\begin{aligned} \gamma_M &= \frac{9}{10}, \\ p_M &= \frac{1-e^{-M}}{10}, \\ I_M &= (\gamma_M - \delta_M, \gamma_M + \delta_M) \times (p_M - \delta_M, p_M + \delta_M) \times (p_M - \delta_M, p_M + \delta_M), \end{aligned}$$

那么，对于每个$(\gamma, p_1, p_2) \in I_M$， $\left(\widetilde{\sigma}^{P_{(M)}} \widetilde{\sigma}^{P(M)}\right)$都是 $\Gamma((1,1) \ \gamma, \mathbf{p})$的纳什均衡。

证明。 回顾一下，

$$V^{P(M)} = Q_1\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right) \ = \frac{1-\gamma^{M+1}(1-p_1(1-p_2)))}{(1-\gamma)(1-\gamma^{M+1}(1-p_1)(1-p_2))}$$

我们将证明，对于每个 $(\gamma, \ p_1, \ p_2) \in I_M$， P_1都没有动机偏离 $\widetilde{\sigma}^{P(M)}$（关于$P_2$ 的证明相同）。

假设 P_1 使用某种策略 σ_1，在某个 $t \neq i \cdot (M+1)$ 向 P_2 开枪。基于通常的理由，只需考虑 P_1 在第一期以概率为一开枪的策略。因此，假设 P_1 在所有 $t \in (1, ..., K)$ 中都不射击，然后在某个 $t' = K+1 \leq M$ 向 P_2 开枪。那么存在以下两种可能。

1. 以概率 p_1： P_2 被杀， P_1 获得收益 $\frac{1}{1-\gamma}$。 —

2. 概率为 $1-p_1$： P_2 未命中， P_1 获得收益 1；在之后的所有回合中，P_2 总会以概率为一的概率向 P_1 射击。在这种情况下， P_1在时间 $t'' > t'$ 的最佳回应是始终以概率为一的概率向 P_2 射击；因此，从第（$t'+1$）回合开始，双方都采用 σ^D 策略。在这种情况下，P_1 获得的总期望收益为 $Q_1\left(\sigma^D, \sigma^D\right)$。

因此，假设 P_1 会先向 $t = K+1 \in \{1, ..., M\}$射击，根据上述推理， P_1的总期望收益为

$$V_1 = Q_1\left(\sigma_1, \widetilde{\sigma}^{P(M)}\right) = \left(\sum_{k=0}^{K} \gamma^k\right) + \gamma^{K+1}\left(p_1 \frac{1}{1-\gamma} + (1-p_1) Q_1\left(\sigma^D, \sigma^D\right)\right)$$

代入 $Q_1\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right)$ 和 $Q_1\left(\sigma^D, \sigma^D\right)$ 的值，并进行大量代数运算后，我们得到

$$\begin{aligned}\delta v_1(\gamma, p_1, p_2) &= (1-\gamma)\left(Q_1\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right) - Q_1\left(\sigma_1, \widetilde{\sigma}^{P(M)}\right)\right) \\ &= \frac{p_2\gamma^{K+2}\left(-(1-p_1)^2(1-p_2)\gamma^{M+1} + (1-p_1)(1-p_2)\gamma^{M-K} - \gamma^{M-K-1} + 1 - p_1\right)}{\left(1-\gamma^{M+1}(1-p_1)(1-p_2)\right)\left(1-\gamma(1-p_1)(1-p_2)\right)}\end{aligned}$$

令 $p_1 = p_2 = p$ ，得

$$\delta v_1(\gamma, p, p) = \frac{p\gamma^{K+2}\left(-(1-p)^3\gamma^{M+1} + (1-p)^2\gamma^{M-K} - \gamma^{M-K-1} + 1 - p\right)}{\left(1-\gamma^{M+1}(1-p)^2\right)\left(1-\gamma(1-p)^2\right)}$$

$\delta v_1(\gamma, p, p)$ 的符号与下式相同

$$\begin{aligned}f_{M,K}(\gamma, p) &= -(1-p)^3\gamma^{M+K+3} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{K+2} \\ &= f_{1,M,K}(\gamma, p) + f_{2,M,K}(\gamma, p)\end{aligned}$$

with

$$\begin{aligned}f_{M,K,1}(\gamma, p) &= (1-p)^2\gamma^{M+2} - (1-p)^3\gamma^{M+K+3} \\ f_{M,K,2}(\gamma, p) &= -\gamma^{M+1} + (1-p)\gamma^{K+2}\end{aligned}$$

现在考虑以下几种情况。

1. **情况 I:** $K \leq M-2$。则 $M-K-1 \geq 1$。对于所有 M 和 $K \in \{1, ..., M-2\}$，我们有

$$(1-p)^2 > (1-p)^3 \text{ and } \gamma^{M+2} > \gamma^{M+K+3}$$

因此我们将始终有 $f_{1,M,K}(\gamma, p) > 0$。若要使 $f_{M,K,2}(\gamma, p) > 0$ 对于特定的K，只需

$$\gamma^{M-K-1} < 1-p \Leftrightarrow \gamma < (1-p)^{\frac{1}{M-K-1}}$$

要使 $f_{M,\ K,\ 2}(\gamma,\ p) > 0$ 对于所有 $K \in \{1, ..., M-2\}$，只需使得

$$\gamma < (1-p)^{\frac{1}{M-2}} \tag{7}$$

换言之，对于所有 M，我们有：

$$\gamma \in \left(0, (1-p)^{\frac{1}{M-2}}\right) \Rightarrow (\forall K \in \{1, ..., M-2\} : f_{M,K}(\gamma, p) > 0)$$
$$\gamma \in \left(0, (1-p)^{\frac{1}{M-2}}\right) \Rightarrow (\forall K \in \{1, ..., M-2\} : \delta v_1(\gamma, p, p) > 0)$$

2. **情况 II：** $K = M-1$。在这种情况下，我们希望

$$\begin{aligned}
-(1-p)^3\gamma^{M+K+3} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{K+2} > 0 \Leftrightarrow \\
-(1-p)^3\gamma^{M+M-1+3} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{M-1+2} > 0 \Leftrightarrow \\
-(1-p)^3\gamma^{2M+2} + (1-p)^2\gamma^{M+2} - \gamma^{M+1} + (1-p)\gamma^{M+1} > 0 \Leftrightarrow \\
-(1-p)^3\gamma^{M+1} + (1-p)^2\gamma - 1 + (1-p) > 0 \Leftrightarrow \\
-(1-p)^3\gamma^{M+1} + (1-p)^2\gamma - p > 0 \Leftrightarrow \\
-(1-p)^3\gamma^{M+1} + \gamma p^2 - (2\gamma+1)p + \gamma > 0
\end{aligned}$$

定义函数

$$\overline{f}_M(\gamma, p) = -(1-p)^3\gamma^{M+1} + \gamma p^2 - (2\gamma+1)p + \gamma$$

由连续性可知，在 $(\gamma_M,\ p_M) = \left(\frac{9}{10}, \frac{1-e^{-M}}{10}\right)$ 的一个充分小邻域内，$\overline{f}_M(\gamma,\ p)$ 的符号与 $-$

$$h_1(M) = \overline{f}_M(\gamma_M, p_M) = -\frac{\left(9+e^{-M}\right)^3\left(\frac{9}{10}\right)^{M+1}}{1000} + \frac{9\left(1-e^{-M}\right)^2}{1000} + \frac{31}{50} + \frac{7e^{-M}}{25}$$

只需证明对于所有 M，$h_1(M) > 0$。为此，我们首先注意到

$$h_1(1) = -\frac{\left(9+e^{-1}\right)^3\left(\frac{9}{10}\right)^2}{1000} + \frac{9\left(1-e^{-1}\right)^2}{1000} + \frac{31}{50} + \frac{7e^{-1}}{25} = 0.60703... > 0$$

另外，令

$$h_2(M) = -\frac{\left(9+e^{-M}\right)^3\left(\frac{9}{10}\right)^{M+1}}{1000} + \frac{31}{50}$$

我们有

$$\forall M : h_1(M) > h_2(M).$$

现在，$h_2(M)$ 关于 M 严格递增，且 $h_2(2) = 0.06422\,2...$。因此

$$\forall M \geq 2 : h_1(M) > h_2(M) > h_2(2) > 0$$

因此，最终我们得到

$$\forall M \geq 1 : \overline{f}_M(\gamma_M, p_M) = h_1(M) > 0.$$

现在，要使得

$$\forall M, \forall K \in \{1, ..., M-1\} : f_{M,K}(\gamma_M, p_M) > 0$$

我们必须确保 (7)对 $(\gamma, p) = (\gamma_M, p_M)$ 成立。换句话说，我们希望 $\gamma_M < (1-p_M)^{\frac{1}{M-2}}$或者，——等价地，

$$\frac{9}{10} < \left(1 - \frac{1-e^{-M}}{10}\right)^{\frac{1}{M-2}} = \left(\frac{9}{10} + \frac{e^{-M}}{10}\right)^{\frac{1}{M-2}}.$$

这成立：由于对于所有 $M \in \mathbb{N}$ 我们都有 $\frac{9}{10} + \frac{e^{-M}}{10} < 1$，我们也有

$$\frac{9}{10} < \left(\frac{9}{10} + \frac{e^{-M}}{10}\right) < \left(\frac{9}{10} + \frac{e^{-M}}{10}\right)^{\frac{1}{M-2}}.$$

简而言之，我们已经证明了

$$\forall M, \forall K \in \{1, ..., M-1\} : f_{M,K}(\gamma_M, p_M) > 0$$
$$\forall M, \forall K \in \{1, ..., M-1\} : \delta v_1(\gamma_M, p_M, p_M) > 0$$

对于所有 M和 K， $\delta v_1(\gamma, p_1, p_2)$ 是连续函数。因此，对于所有 M，存在某个$\delta_M > 0$ 使得

$$\forall K \in \{1, ..., M-1\}, \forall (\gamma, p_1, p_2) \in I_M : \delta v_1(\gamma, p_1, p_2) > 0$$

这表明 P_1 没有动机偏离 $\widetilde{\sigma}^{P(M)}$。同样的论证也适用于 P_2。因此，对于每个 $(\gamma, p_1, p_2) \in I_M$，$\left(\widetilde{\sigma}^{P(M)}, \widetilde{\sigma}^{P(M)}\right)$ 都是 $\Gamma((1,1)\ \gamma, \mathbf{p})$的一个纳什均衡。 ■

5 一些补充说明

现在让我们说明为何将这些策略称为“合作”策略和“背叛”策略。根据第 3 节的结果，对于 $n \in \{1, 2\}$，我们有

$$\overline{Q}_n\left((1,1), \sigma^C, \sigma^C\right) = \frac{1}{1-\gamma}$$
$$\overline{Q}_n\left((1,1), \sigma^D, \sigma^D\right) = \frac{1-\gamma(1-p_n(1-p_{-n}))}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))}.$$

由此可得

$$\overline{Q}_n\left((1,1), \sigma^C, \sigma^C\right) - \overline{Q}_n\left((1,1), \sigma^D, \sigma^D\right) = \frac{\gamma p_{-n}}{(1-\gamma)(1-\gamma(1-p_1)(1-p_2))} > 0.$$

简而言之，与囚徒困境中一样，对双方而言，不开枪比开枪更有利。由于在我们的表述中，击毙对手不会带来直接收益，因此 (σ^C, σ^C) 和 (σ^D, σ^D) 都是纳什均衡；然而，对双方而言， (σ^C, σ^C) 是比 (σ^D, σ^D) 更有利的纳什均衡。这就是为何将 (σ^C, σ^C) 称为合作策略，而将 (σ^D, σ^D) 称为背叛策略。

这一切可能令人惊讶，因为人们会认为，在决斗中，每个玩家的目标都是消灭对手。或许可以认为， (σ^C, σ^C)的盈利能力更高。

这源于我们选择不给杀死对手这一行为分配任何直接收益。但事实并非如此。即使存在正的"杀戮收益"， $Q_n\left((1,1)\ \sigma^C,\sigma^C\right)$ 仍然可能大于 $Q_n\left((1,1)\ \sigma^D,\sigma^D\right)$，只要 γ 足够接近一[6]。$\left(\sigma^C,\sigma^C\right)$ 之所以更优，原因在于：如果对生存分配正的收益，那么经过无限多个回合的复利累积，这一收益总是能够超过杀戮收益。因此，我们的模型可以被理解为相较于通常决斗模型[7]更为"和平主义"的版本。

现在，让我们将静态决斗与PD进行比较。在PD和决斗中，合作都比背叛更有利。虽然 $\left(\sigma^C,\ \sigma^C\right)$ 不是PD中的纳什均衡，但 $\left(\sigma^D,\ \sigma^D\right)$ 是两者中的纳什均衡。然而，决斗和重复进行的PD都拥有多个采用冷酷策略的纳什均衡；所有这类均衡的共同特征是促进合作，或者换句话说，惩罚背叛（射击）。事实上，与重复PD的情形类似，也许可以为静态决斗证明一个"民间定理"；也就是说，当 γ 足够接近1时，每个可行且个体理性的收益都是一个纳什均衡。我们计划在未来研究这个问题。

6 结论

我们将同时射击的静态决斗表述为一个折扣随机博弈。我们证明了它在平稳策略中有两个纳什均衡，即"始终射击"和"从不射击"策略；此外，还存在若干非平稳的"促进合作型"纳什均衡。未来，我们计划从多个方向扩展对静态决斗的研究。

首先，我们希望拓展研究，并针对两种变体获得类似结果：(a)杀戮收益非零的情形，以及(b)仅有终terminal收益的情形。此外，我们希望构建并研究静态决斗的一个版本，其中每个玩家都希望在尽可能短的时间内杀死对手。

其次，我们希望证明一种"民间定理"形式，即每一个可行且个体理性的收益，在 γ 足够接近 1 时，都是一个纳什均衡。

最后，我们希望将静态 $Nuel$ （即涉及 N玩家相互射击的类似决斗博弈）表述为折扣随机博弈，并将我们的结果扩展到这种情况。

参考文献

[1] Amengual，P. 和 Toral，R.《三方决斗博弈中获胜者的分布》。*AIP* 会议论文集 （2005），第 128–141 页。

[2] Amengual，P. 和 Toral，R.《三决斗的马尔可夫链分析》。《第八届格拉纳达计算物理研讨会论文集》（2005 年）。

[6]这一点，以及关于正杀戮收益情形的其他结果，将在未来的出版物中报告。[7]唐纳德·克努特也曾在三决斗的语境下提出过这一点 [17]。例如，他指出："就对手的存活而言，一个选择不射击的玩家是在确保其对手没有理由还击。"

[3] Amengual，P. 和 Toral，R. “三决斗，或弱者生存。”《科学计算》。[4] Barron，E. N.《策略博弈：导论》。约翰·威立父子出版社（2013）。[5] Bossert，W.、BramsS.J.和 Kilgour，D.M. “合作型三决斗与非合作型三决斗：几乎没有共识，但这重要吗？。”《博弈与经济行为》，第40卷（2002），第185-202页。[6] Brams，S.J. 和 Kilgour，D.M. “三决斗”。《数学杂志》，第70卷（1997），第315-326页。[7] Brams，S.J.和Kilgour，D.M. “《以轰鸣或呜咽结束的博弈》”。预印本，CV Starr应用经济学中心（2001）。[8] Brams，S.J.、Kilgour，D.M. 和 Dawson，B. “三决斗与未来。”《数学视界》，第10卷（2003），第5-8页。[9] Dorraki，M.、Allison，A. 和 Abbott，D.。“三决斗与生存策略。”《科学报告》，第9卷（2019），第1-7页。[10] Dresher，M.《策略博弈：理论与应用》。兰德公司（1961）。[11] Gardner，M.：《新的数学谜题与消遣》。（1966），第42-49页。[12] Karlin，S.，《博弈、规划与经济学中的数学方法与理论》，第2卷，Addison-Wesley（1959）。[13] Kilgour，D.M.“同时三决斗。”《国际博弈论杂志》，第1卷（1971），第229-242页。[14] Kilgour，D.M.“序贯三决斗。”《国际博弈论杂志》，第4卷（1975），第151-174页。[15] Kilgour，D. M. “无限序贯三决斗的均衡点。”《国际博弈论杂志》，第6卷（1977），第167-180页。[16] Kinnaird，C.《谜题与消遣百科全书》。Secaucus（1946）。[17] Knuth，D.E. “三决斗：一种新解法。”《娱乐数学杂志》，第6卷（1972），第1-7页。[18] Larsen，H.D.“《飞镖游戏》。”《美国数学月刊》，（1948）第640-41页。[19] Mosteller，F.《五十道具有挑战性的概率题及其解答》。Courier公司（1987）。[20] Shubik，M.“最适者一定生存吗？”，《博弈论与政治行为读本》，道布尔戴出版社。[21] Sobel，M.J.“非合作随机博弈。”《数学统计学年刊》，第42卷（1971），第1930-1935页。

[22] Toral，R.，和 Amengual，P.。“三决斗博弈中获胜者的分布。”*AIP*会议论文集，第779卷。美国物理学会（2005年）。[23] Wegener，M. 和 Mutlu，E.。“善良者、邪恶者与人脉广泛者。”《国际博弈论杂志》，第50卷（2021年），第759-771页。[24] Xu，X. “三决斗博弈。”《综合》，第185卷（2012年），第19-25页。[25] Zeephongsekul，P. “随机 N-Uels的纳什均衡点。”数学规划的新近发展。CRC出版社（1991年），第425-452页。