arXiv:2603.06072v1 [stat.ME] 6 Mar 2026

不完全信息下竞争性库存与定价的层次贝叶斯动态博弈：学习、可信风险与均衡

Debashis Chatterjee*[1]

[1]统计系，维斯瓦·巴拉蒂大学，桑蒂尼克坦，印度[2]S. N. Bose 国家基础科学中心，加尔各答，印度

2026年3月9日

摘要

我们提出了一种不完全信息下竞争性库存与定价的层次贝叶斯动态博弈。两家公司在面临两层不确定性——未知市场需求与私有对手特征——的同时，反复选择订货量与价格。该框架将关于需求与替代的贝叶斯学习与关于对手类型的策略性信念更新相结合。为使决策对后验不确定性保持稳健，我们引入一种可信风险准则，该准则奖励预期未来利润，同时惩罚后验预测离散度。由此得到一种保守均衡概念，其中企业同时进行学习、竞争与适应。

本文给出了模型构建、信息结构、后验更新机制、均衡定义，以及一种基于信念状态动态规划的计算策略。仿真研究表明，贝叶斯学习对于取得良好表现至关重要，而可信风险规则在不确定性下尤其适合作为操作性正则化项。基于高维蛋白质表达数据集的真实数据示例表明，同一不确定性感知贝叶斯原则能够揭示具有生物学可解释性的亚组和潜在状态。所提出的框架为贝叶斯博弈论与运筹学搭建了统一的桥梁，在不确定且信息受限环境中的竞争性决策方面具有实际意义。

关键词：贝叶斯博弈；动态竞争；不完全信息；贝叶斯学习；竞争性库存；不确定性下的定价；可信风险决策规则；均衡分析；运筹学；不确定性感知推断

1 引言

博弈论研究理性决策者之间的策略互动，而贝叶斯分析则提供了一种连贯的机制，用于表示和更新不确定性。当参与者必须在不了解环境或彼此的完整信息时进行策略性行动，这两者的结合便顺理成章。这一奠基性洞见归功于Harsanyi [1967, 1968], ，他通过引入由自然抽取的类型来表示不完全信息，并将原有的策略问题转化为不完全信息博弈。这促成了现代概念的形成

*通讯作者：debashis.chatterjee@visva-bharati.ac.in

贝叶斯博弈和贝叶斯纳什均衡。后续研究对信念层级和类型空间进行了形式化处理 [Mertens和Zamir, 1985, Aumann和Heifetz, 2002], 而静态博弈和重复博弈中的贝叶斯学习则由Jordan [1991, 1995]进行了研究。

在运筹学中，许多竞争性决策问题本质上具有博弈论特征。只要一家企业的行动会影响另一家企业的需求或利润，库存与定价问题就会演变为博弈。产品互为替代品、缺货引发需求溢出，或顾客在不同企业之间转换时，情况尤其如此。运筹学领域的重要贡献包括研究替代情形下的随机库存博弈 [Avsarand Baykal-G¨ursoy, 2002], 竞争性报童模型 [Serin, 2007], 以及针对未知需求分布的贝叶斯库存模型 [Azoury, 1985, Kamath 和 Pakkala, 2002, Zhang 和 Chen, 2006, Zhang 等人, 2016, Zhou 和 Luo, 2023]。然而，现有文献仍有很大空间来构建一个能够同时刻画以下因素的统一框架：

(i) 多家企业之间的战略竞争，

(ii) 关于竞争对手运营类型的私人信息，

(iii) 对潜在需求与替代关系进行序贯贝叶斯学习，

(iv) 由后验不确定性引致的显式风险惩罚。

本文提出了这样一个框架。我们关注一个重复双寡头市场，其中两家公司在有限或无限时域内销售可替代产品。每家企业都了解自身的运营类型——例如，其私有的边际采购成本、服务效率或残值处理能力——但不了解竞争对手的真实类型。同时，双方都无法完全了解共同的需求环境：市场规模、交叉价格敏感度和缺货替代强度都是部分未知的。企业会随着时间推移观察实现销售和缺货情况，依据贝叶斯定理更新信念，随后策略性地选择后续订货量和价格。

本文处于四类文献的交汇处。

首先，贝叶斯博弈的一般理论始于Harsanyi [1967, 1968], ，他通过类型分布表述了不完全信息博弈。此后，Mertens和Zamir [1985]进一步完善了认知论和形式化的信念空间处理方法，而Au- mann和Heifetz [2002]对其进行了综述。

其次，博弈中的贝叶斯学习研究信念如何在重复的策略互动中演变。重要贡献包括Jordan [1991]关于标准型博弈的研究，以及Jordan[1995]关于重复博弈的研究。

第三，运筹学领域的竞争性库存与替代博弈包括 Avsar 和 Baykal-G¨ursoy [2002], ，他们将可替代需求下的库存控制分析为随机博弈；以及 Serin [2007], ，其研究了竞争性报童行为。

第四，库存和定价领域的贝叶斯需求学习见于 Azoury [1985], Ka-math 和 Pakkala [2002], Zhang 和 Chen [2006], ，以及较新的动态更新研究，例如 Zhang 等人 [2016], Zhou和Luo [2023]。这些论文与本文高度相关，但通常聚焦于单一决策者，或仅研究竞争，而未采用完整的 Harsanyi 型私人信息结构。

我们的研究将这些研究方向融为一体，构建了一个统一的动态均衡框架，从而填补了这一空白。

1.1 研究目标与创新性

本文的主要目标是构建一个连贯的贝叶斯博弈论框架，用于分析企业在重复性竞争库存与定价决策中，如何一边了解市场，一边考虑竞争对手的策略。具体而言，本文旨在通过单一模型统一以下四个要素：关于竞争对手特征的不完全信息、对市场需求与替代的序贯学习、依状态而定的策略决策，以及在后验不确定性下选择保守行动。

本文的目标并非只是将贝叶斯更新附加到标准库存问题中，而是构建一个动态均衡框架，使信念本身成为不断演化的状态的一部分，从而让战略行为同时取决于运营变量和累积信息。从这个意义上说，本文旨在将贝叶斯学习纳入博弈之中，而非置于博弈之外。

第二个目标在于方法论。我们旨在说明，不确定性不仅应当被估计，还应当在决策中得到规范性运用。因此，本文提出的可信风险原则将后验不确定性转化为有章可循的战略行为。这使该框架既能为均衡理论提供参考，也能为运营情境中的实际决策提供支持；在这些情境下，信息不足时采取激进行动可能代价高昂。

本文的创新性 本研究的创新之处在于，将贝叶斯博弈论、序贯学习与运筹学联系起来，纳入一个统一的动态框架。

首先，本文将竞争性库存与定价建模为一种同时包含市场不确定性和战略不确定性的贝叶斯动态博弈。企业不仅学习需求状况，还会随着时间推移更新对竞争对手特征的信念。这种联合处理在结构上比将学习与竞争分开研究的方法更加一体化。

其次，本文提出了一项可信风险决策原则，明确对后验不确定性施加惩罚。这使模型具备规范性不确定性感知组件：企业不只是估计未来，还会在后验不确定性仍然较大时采取保守行动。这一特征在重复运营场景中尤为重要，因为过度自信的行动可能导致本可避免的损失。

第三，模型通过增广信念状态表示来表达，从而能够在统一的动态规划框架中研究学习与竞争。这建立了贝叶斯滤波、战略选择与均衡分析之间清晰的联系。

本文通过两方面的证据支持理论：一是在预期的竞争环境中提供基于仿真的证据，二是提供真实数据示例，展示底层可信风险贝叶斯原则在复杂高维数据分析中的更广泛用途。

2 模型构建

2.1 参与者、时域与行动

考虑两家企业，分别以 $i \in \{1, 2\}$ 编号，它们在时期 $t = 1$、.....、T 进行竞争。企业 i 在每个时期选择：

$$a_{it} = (q_{it}, p_{it}),$$

其中

- $q_{it} \in \mathcal{Q}_i \subset \mathbb{R}_+$ 是订货量，

- $p_{it} \in \mathcal{P}_i \subset \mathbb{R}_+$ 是售价。

定义 $a_t = (a_{1t},\ a_{2t})$。

令 I_{it} 表示企业 i的期初库存。则订购后的可用库存为

$$S_{it} = I_{it} + q_{it}.$$

2.2 私人类型

每家企业都有一个私人类型

$$\tau_i = (c_i, h_i, s_i) \in \mathcal{T}_i,$$

其中:

- c_i =边际采购成本,
- h_i =持有成本,
- s_i =残值或处置价值。

自然从共同先验中抽取 $(\tau_1,\ \tau_2)$

$$\mu_0(\tau_1, \tau_2).$$

企业 i 仅观察到 τ_i，而无法观察到 τ_j 对于 $j \neq i$。

这是第一层贝叶斯层：对竞争对手运营特征的不完全信息。

2.3 潜在需求状态

设共同市场状态为

$$\theta = (\alpha_0, \alpha_1, \alpha_2, \beta, \lambda, \sigma^2),$$

其中:

- α_0 表示基准市场规模,
- α_i 表示产品特定偏好参数,
- $\beta > 0$ 表示交叉价格替代,
- $\lambda \geq 0$ 表示缺货驱动溢出需求,
- σ^2 是需求噪声方差。

企业并不完全了解向量 θ 。为其指定了一个先验分布

$$\pi_0(\theta).$$

这是第二个贝叶斯层：关于市场基本要素的统计不确定性 .

2.4 需求系统

我们提出以下结构性需求方程：

$$D_{it}^{*} = \alpha_0 + \alpha_i - \eta p_{it} + \beta p_{jt} + \lambda Z_{jt} + \varepsilon_{it}, \qquad i \neq j, \tag{1}$$

其中：

- $\eta > 0$ 表示自身价格敏感度，
- $Z_{jt} = \mathbb{1}\{D_{jt}^{*} > S_{jt}\}$ 表示竞争对手缺货，
- ε_{it} 独立同分布$\sim N(0, \sigma^2)$，以 θ为条件。

观测销量受到可用库存的截尾：

$$Y_{it} = \min\{D_{it}^{*},\, S_{it}\}. \tag{2}$$

期末库存的演变如下

$$I_{i,t+1} = S_{it} - Y_{it}. \tag{3}$$

注释 2.1。 方程（1）通过 βp_{jt} 刻画了价格竞争，并通过 λZ_{jt}刻画了缺货引发的替代。这比标准的单周期报童模型更丰富，因为策略性替代既通过价格，也通过库存可用性产生。

3 信息结构与信念状态

3.1 历史记录

设时刻 t 的公共历史为

$$H_t^c = \{(p_{1\ell}, p_{2\ell}, Y_{1\ell}, Y_{2\ell}, Z_{1\ell}, Z_{2\ell}) : \ell = 1, \ldots, t-1\}.$$

企业 i 在时刻 t 的私人信息为

$$H_{it} = (H_t^c, \tau_i, I_{it}).$$

Hen况，策略状态包括公开观察和私人类型信息 n.

3.2 后验信念

在每个时期，企业都会维持：

(a) 对共同市场状态的后验信念，

$$\pi_t(\theta) = \mathbb{P}(\theta \mid H_t^c),$$

(b) 对竞争对手类型的后验，

$$\mu_{it}(\tau_j) = \mathbb{P}(\tau_j \mid H_t^c, \tau_i), \qquad j \neq i.$$

因此，与最优博弈相关的增广状态为

$$X_{it} = (I_{it}, \tau_i, \pi_t, \mu_{it}).$$

定义 3.1（信念状态马尔可夫表示）。 如果在给定 $X_t = （X_{1t}， X_{2t}）$和当前行动 a_t 的条件下，$(X_{1,\ t+1},\ X_{2,\ t+1})$ 的分布独立于更早的历史记录，则称该动态博弈具有信念状态马尔可夫表示。

这种表示至关重要，因为它使得能够在由后验超参数扩展的状态空间上使用动态规划。

4 层级贝叶斯设定

4.1 先验结构

具体而言，假设

$$\boldsymbol{\psi} := (\alpha_0, \alpha_1, \alpha_2, \eta, \beta, \lambda)^\top \sim N_6(m_0, \Sigma_0), \qquad \sigma^2 \sim \mathrm{IG}(a_0, b_0),$$

且 $\boldsymbol{\psi}$ 与 σ^2 先验独立。

对于私人成本，假设

$$c_i \in \{c_L, c_H\}, \qquad \mathbb{P}(c_i = c_H) = \rho_i,$$

对于离散化的其他类型分量，也可作类似处理。这为关于竞争对手类型的事中信念提供了计算可处理性。

4.2 似然

给定行动和缺货指示变量，实现的未删失需求方程可写为

$$D^*_{it} = x_{it}^\top \boldsymbol{\psi} + \varepsilon_{it},$$

其中协变量向量为

$$x_{it} = (1, \mathbb{1}\{i=1\}, \mathbb{1}\{i=2\}, -p_{it}, p_{jt}, Z_{jt})^\top.$$

如果能够观测到未删失需求，后验更新将遵循标准的正态-逆伽马递推。在销量被删失的情况下，可以采用数据增广方法，将缺货发生时的潜变量D^*_{it} 视为缺失。

4.3 使用数据增强进行后验更新

如果 $Y_{it} < S_{it}$，则不会发生删失，并且

$$D^*_{it} = Y_{it}.$$

如果 $Y_{it} = S_{it}$，则

$$D^*_{it} \geq S_{it},$$

因此，潜在需求服从右截断正态分布：

$$D^*_{it} \mid \boldsymbol{\psi}, \sigma^2, H_t^c \sim N(x_{it}^\top \boldsymbol{\psi}, \sigma^2) \text{ truncated to } [S_{it}, \infty).$$

因此，可以进行 Gibbs 类型的后验递推：

(i) 为删失观测值抽取潜变量 D_{it}^{*}；

(ii) 根据多元正态分布更新 $\psi \mid \sigma^{2},\ D^{*}$；

(iii) 根据逆伽马分布更新 $\sigma^{2} \mid \psi,\ D^{*}$。

注 4.1。这一贝叶斯滤波步骤可将需求学习直接整合到博弈状态中。后验超参数，而非完整的原始历史数据，成为充分的状态描述量。

5 利润结构

企业 i 单期利润为

$$\Pi_i(a_t, \theta, \tau_i) = p_{it} Y_{it} - c_i q_{it} - h_i (I_{i,t+1})^{+} + s_i (I_{i,t+1})^{-}_{\text{disp}}, \tag{4}$$

其中，最后一项表示残值收益或处置成本，具体取决于符号约定。对于标准的库存解释，我们可以简单地写作

$$\Pi_i = p_{it} Y_{it} - c_i q_{it} - h_i (I_{i,t+1}) + s_i I_{i,t+1},$$

若 $I_{i,\ t+1} \geq 0$ 始终成立。具体的记账约定可以灵活处理。

6 可信风险目标

本文的一项核心贡献是定义了一个可信风险目标，该目标明确考虑了后验预测不确定性。给定当前信念状态 X_{it} 和策略组合 $\sigma = (\sigma_1,\ \sigma_2)$，定义延续价值

$$V_i^{\sigma}(X_{it}) = \mathbb{E}^{\sigma}\left[\sum_{u=t}^{T} \delta^{u-t} \Pi_i(a_u, \theta, \tau_i) \,\middle|\, X_{it} \right].$$

经典的贝叶斯期望效用最大化采用 $\mathbb{E}[V_i^{\sigma} \mid X_{it}]$。我们则定义

$$\mathcal{J}_i^{\sigma}(X_{it}) = \mathbb{E}[V_i^{\sigma}(X_{it}) \mid X_{it}] - \kappa_i \sqrt{\mathrm{Var}(V_i^{\sigma}(X_{it}) \mid X_{it})}, \tag{5}$$

其中，$\kappa_i \geq 0$ 是企业特定的不确定性厌恶参数。

注 6.1. 式（5）中的惩罚项将后验不确定性转化为保守行动。它并非极大极小意义上的模糊厌恶，而是一种贝叶斯后验预测风险调整。当 $\kappa_i = 0$ 时，我们便得到通常的贝叶斯期望收益。

7 策略空间与均衡

7.1 行为策略

企业 i 的（马尔可夫）行为策略是一个可测映射

$$\sigma_i : \mathcal{X}_i \to \Delta(\mathcal{A}_i),$$

其中 $\mathcal{X}_i$ 是信念状态空间，而 $\Delta(\mathcal{A}_i)$ 是行动空间 $\mathcal{A}_i = \mathcal{Q}_i \times \mathcal{P}_i$ 上的概率测度集合。

在确定性情形下，可以写作

$$a_{it} = \sigma_i(X_{it}).$$

7.2 中期最佳响应

固定对手策略 σ_j。参与者 i 的中期最佳响应满足

$$\sigma_i^* \in \arg\max_{\sigma_i} \mathcal{J}_i^{(\sigma_i,\sigma_j)}(X_{it}) \qquad \text{for all admissible } X_{it}. \tag{6}$$

定义 7.1（可信风险马尔可夫完美贝叶斯纳什均衡）。策略组合$\sigma^* = (\sigma_1^*, \sigma_2^*)$ 若满足以下条件，则称为可信风险马尔可夫完美贝叶斯纳什均衡（CR-MPBNE）：

（a）在可能的情况下，按照贝叶斯法则更新信念；

（b）每个σ_i^* 都关于当前信念状态 X_{it}可测；

(c) 对每家企业 i及每个可容许状态 X_{it}，

$$\mathcal{J}_i^{(\sigma_i^*,\sigma_j^*)}(X_{it}) \geq \mathcal{J}_i^{(\sigma_i,\sigma_j^*)}(X_{it}) \quad \forall \sigma_i.$$

8 动态规划刻画

令 $W_i(X_{it})$ 表示企业 i 的均衡价值函数。那么，贝尔曼方程组 is

$$W_i(X_{it}) = \max_{a_{it}\in\mathcal{A}_i} \Big\{ \mathbb{E}[\Pi_i(a_{it}, a_{jt}, \theta, \tau_i) \mid X_{it}] - \kappa_i \sqrt{\mathrm{Var}[\Pi_i(a_{it}, a_{jt}, \theta, \tau_i) \mid X_{it}]} + \delta\, \mathbb{E}[W_i(X_{i,t+1}) \mid X_{it}, a_{it}, a_{jt}] \Big\}, \tag{7}$$

其中，a_{jt} 由 $\sigma_j^*(X_{jt})$ 诱导。

由于$X_{i,\ t+1}$ 包含后验更新，延续项同时纳入实物库存演变和信息积累。正是在这里，贝叶斯学习与博弈论密不可分。在加强紧性与连续性条件后，相应的贝尔曼算子在有界连续价值函数上定义良好，并具有附录 A 中所用的不动点性质。

9 结构假设

为确保存在性和可处理性，我们作出以下假设。

假设 9.1（紧致行动空间）。对每个 i,

$$\mathcal{Q}_i = [0, \bar{q}_i], \qquad \mathcal{P}_i = [\underline{p}_i, \bar{p}_i]$$

其中 $0 < p_{_i} < \bar{p}_i < \infty$ 和 $\bar{q}_i \underline{<} \infty$ 是有限常数。

假设 9.2（需求正则性）。对于每个可容许状态与行动对，条件需求矩有限，且关于行动连续。

假设 9.3（凹性）。 给定竞争对手的行动和当前信念，一步奖励加上折扣后的延续价值的后验均值关于 (q_{it}, p_{it}) 为凹函数。

假设 9.4（有界可信惩罚）。式（5）中的后验预测方差项有限，且关于行动连续。

为给出附录 A 中报告的严格存在性证明，我们施加一组强化的正则性条件，包括紧致度量信念状态空间、单期可信风险收益有界且连续、受控转移核弱连续，以及采用混合马尔可夫策略以确保最优反应对应为凸值。此处陈述的假设刻画了主要的经济结构，而附录则列出了完整形式化证明所需的更强技术条件。

10 均衡的存在性

我们首先概述主要存在性结果。在强化的技术正则性条件下，这些结果的严格版本见附录A。

命题 10.1（中期最佳回应的存在性）。假设行动空间 $\mathcal{A}_i$ 是紧致的，单期可信风险收益关于自身行动连续，且延续价值有界且连续。那么，对于任意固定的可测对手策略σ_j，参与者 i在每个可容许状态下的中期优化问题都至少存在一个最大化解。

证明。 中期目标关于自身行动连续，因为单期可信风险收益连续，且延续价值有界且连续。由于行动空间是紧致的，魏尔斯特拉斯极值定理表明该目标可取得最大值。在强化正则性假设下的完整严谨证明见附录 A；参见命题 A.9。

□

定理 10.2（强化正则性条件下 CR-MPBNE 的存在性）。 假设信念状态空间是紧致度量空间，行动空间是紧致的，单期可信风险收益有界且连续，受控转移核弱连续，并且参与者可以采用混合马尔可夫策略，使相应的最佳回应对应非空、凸值且上半连续。那么至少存在一个可信风险马尔可夫完美贝叶斯纳什均衡。

证明。 在所述正则性条件下，每位参与者的贝尔曼算子在有界连续价值函数上定义良好，中期最佳响应存在，且最佳回应对应非空。采用混合策略后，对应具有凸值性，并且由连续性假设可得其上半连续性。因此，角谷不动点定理可推出一个均衡不动点。强化假设下的严谨证明见附录 A；参见定理A.13。

□

11 特殊情形

我们的框架嵌套了若干重要模型。

11.1 经典竞争性报童博弈

If:

- 时域为单周期,
- 价格由外生因素固定,
- $\kappa_i = 0$,
- 所有需求参数均已知,
- 类型是共同知识,

那么该模型就简化为一个标准的竞争性报童问题，类似于 Serin [2007]。

11.2 带替代的随机库存博弈

If:

- 需求参数已知,
- 只有库存会动态变化,
- 不存在后验学习,

那么，该模型就变成了 Avsar 和 Baykal- G¨ursoy [2002]所研究的那类随机库存博弈。

11.3 贝叶斯单智能体库存学习

若移除竞争且只剩下一家企业，该模型便退化为贝叶斯库存学习模型，例如 Azoury [1985], Kamath andPakkala [2002], Zhang 和 Chen [2006]。

12 计算

由于信　念状态空间可能很大，我们提出了一套实用的计算流程　　ne.

12.1 后验压缩

用超参数表示后验 $\pi_t(\theta)$

$$\zeta_t = (m_t, \Sigma_t, a_t, b_t),$$

或者在存在非共轭特征时使用粒子近似。

12.2 类型信念更新

对于离散的对手类型空间 $\mathcal{T}_j = \{\tau_j^{(1)}, \ldots, \tau^{(K)}{}_j\}$,

$$\mu_{it}^{(k)} \propto \mu_{i,t-1}^{(k)} L\left(H_t^c \mid \tau_j^{(k)}, \tau_i, \zeta_{t-1}\right), \qquad k = 1, \ldots, K,$$

随后进行归一化。

12.3 近似动态规划

价值函数可近似表示为:

- 基于网格的值迭代,
- 拟合值迭代,
- 策略迭代,
- 基于仿真的贝叶斯动态规划。

算法 1 后验学习均衡迭代

1：初始化先验超参数 ζ_0、类型信念 μ_{i0}以及初始策略 $\sigma_i^{(0)}$ 2：对 $m=0$，1，2，.....直至收敛执行 3：对每个代表性信念状态网格点 X 执行 4：对 $i=1$，2执行 5：在 $\sigma_j^{(m)}$ 下计算后验预测销售额／利润矩6：求解可信风险最优响应：$\sigma_i^{(m+1)}(X) = \arg\max \{$后验平均收益 $-\kappa_i \times$ 后验标准差 $+\delta \times$ 延续价值$\}_{a_i}$7：结束循环 8：结束循环 9：模拟轨迹，按贝叶斯法则更新 ζ_t 和 μ_{it} 10：结束循环 11：输出近似均衡策略 $\widehat{\sigma}_1$、$\widehat{\sigma}_2$

13 解释与管理启示

该模型得出若干定性启示。

(M1) **后验不确定性会抑制激进竞争**。当 κ_i 较大或后验方差较高时，企业的行为会更加保守，避免过度备货和极端定价。(M2) **学习可能加剧或缓和竞争**。对高需求市场有更深入的学习可能会促使企业激进备货。相反，了解到替代强度较弱可能会减轻竞争压力。(M3) **私人成本不对称会形成策略性筛选**。低成本企业可以采取更激进的行动，随着时间推移，其行动可能会揭示有关其类型的信息，从而间接改变竞争对手的信念。(M4) **库存和定价会成为信息信号**。在重复市场中，订货量和价格不仅是运营决策，也会揭示有关信念、类型和预期市场状况的信息。

14 可能的理论扩展

还可以进行若干在数学上颇具意义的扩展。

（E1）**平均场推广：**将模型从双寡头扩展至 N 竞争企业，并研究贝叶斯平均场均衡。（E2）**网络替代：**使替代遵循产品或地点之间的图结构。（E3）**分布鲁棒性：**用围绕后验预测律的瓦瑟斯坦或KL球鲁棒化替代后验方差惩罚。（E4）**机制设计层：**引入设定信息披露规则的平台或监管者，从而形成贝叶斯斯塔克尔伯格博弈。

15 仿真研究

在本节中，我们研究所提出的均衡策略Proposed Bayesian CredibleRisk能否在不完全信息下的重复性竞争库存-定价博弈中带来更优表现。仿真研究的目的有三：

1. 评估后验学习相对于非学习经典基准是否提升策略表现；2. 评估我们的可信风险修正是否比纯粹的风险中性贝叶斯学习者产生更理想的运营行为；3. 联合考察利润、学习精度和动态市场行为，而不是仅通过单一标量指标。

15.1 模拟设计

我们模拟了一个持续 $T = 30$ 个时期的重复双寡头市场，并针对每种竞争策略进行了150次独立蒙特卡洛重复实验。在每次重复实验中，各企业的私人成本均在期初抽取一次，随后在整个时域内保持不变。需求根据结构模型生成，

$$D_{it} = \beta_0 + \beta_1 p_{it} + \beta_2 p_{jt} + \beta_3 \mathbf{1}\{\text{rival stockout}\} + \varepsilon_{it}, \qquad i \neq j,$$

其真实参数值为

$$\beta_0 = 45, \qquad \beta_1 = -3.6, \qquad \beta_2 = 1.2, \qquad \beta_3 = 7.5,$$

高斯噪声的标准差为 $\sigma = 4.5$。因此，自身价格会抑制需求，竞争对手价格会通过替代效应提升需求，而竞争对手缺货则会产生额外的正向溢出效应。

模拟对所有方法使用了相同的经济环境：折扣因子 δ =0.98，低边际成本和高边际成本分别为6和10，持有成本0.8，残值1.5，价格网格$\{8, 9, \ldots, 16\}$，以及数量网格 $\{20, 25, \ldots, 65\}$。贝叶斯需求学习者从相同的先验均值初始化

$$(35.0,\ -2.0,\ 0.5,\ 3.0)$$

具有先验标准差

$$(10.0,\ 2.0,\ 2.0,\ 4.0).$$

15.2 对比方法

我们比较了以下三种策略。

1. **所提方法：**
   Proposed Bayesian CredibleRisk。这是本文提出的方法。它对需求系统和竞争对手类型信念进行贝叶斯更新，然后通过最大化如下形式的后验可信风险 目标来选择行动

$$\text{后验平均利润} \ - \ \kappa \times \text{后验预测标准差},$$

   其中 $\kappa = 0.60$。

2. **贝叶斯风险中性学习器：**贝叶斯风险中性。此方法采用与我们的方法相同的贝叶斯后验更新机制，但将可信风险惩罚设为零，即 $\kappa = 0$。

3. **经典非学习基准：**
   经典静态先验。此方法采用固定的基于先验的启发式方法，不会随时间动态学习。

因此，该比较经过精心设计，旨在分离学习和可信风险正则化各自的影响。我们的方法与经典静态先验之间的比较衡量了贝叶斯学习和自适应战略响应的价值；我们的方法与贝叶斯风险中性之间的比较则衡量了额外的保守不确定性校准所带来的价值。

15.3 评估指标

对于每种方法，我们记录了：

- 总折现市场利润；
- 企业层面的折现利润；
- 需求参数的最终后验均方误差（均方误差）；
- 随时间变化的缺货率；
- 随时间变化的平均价格和订货量；
- 关于竞争对手是否为高成本企业的后验信念。

为量化绩效比较中的不确定性，我们还计算了最终折现市场利润均值之差和最终后验均方误差均值之差的自助法置信区间。

15.4 主要数值结果

表1 汇总了主要统计数据。几个观察结果显而易见 .

首先，两种贝叶斯学习程序均大幅优于非学习的经典基准。经典静态先验策略的平均总折现市场利润仅为67.01，中位数为0.00，而两种基于学习的方法的平均总折现市场利润均约为1593–1597。这是一次显著的改善，有力地支持了后验学习在所提博弈中的重要性。

其次，在三种方法中，我们所提方法取得了最大平均总折现市场利润（1597.30），以及最大总折现市场利润中位数（1678.98）。因此，就市场层面折现收益这一首要经济准则而言，我们的方法是本实验中表现最好的。

第三，与贝叶斯风险中性的比较要比与静态基准的比较紧密得多。我们所提方法与风险中性贝叶斯学习者之间的平均折现市场利润差异仅为约4.01个单位，相对于模拟规模而言在经济上很小，并且在后文表3所报告的自助法分析中在统计上可忽略不计。

第四，最终后验均方误差结果显示，两种贝叶斯学习者的表现再次十分接近，但在这里，所提方法并非均方误差最低的方法。具体而言，贝叶斯风险中性的平均最终均方误差为 17.3283，而所提方法为 17.6573。因此，在当前配置下，可信风险机制似乎略微提高了盈利能力，却没有提升终端参数估计精度。

表 1：三种竞争策略的主要模拟结果。所提方法为 Proposed Bayesian CredibleRisk。

策略	平均总利润	总利润标准差	总利润中位数	平均最终均方
贝叶斯风险中性	1593.29	1211.32	1608.33	17.
经典静态先验	67.01	1347.33	0.00	30.
Proposed Bayesian CredibleRisk	**1597.30**	1351.41	**1678.98**	17.

表2 进一步阐明了相较于两种基准方法，所提方法的相对增益。相较于静态先验策略，所提方法使平均总折现市场利润提高约2283。66%，并使平均最终均方误差降低约42。72%。然而，相较于贝叶斯风险中性学习器，平均利润增益仅为0。25%，且采用所提方法时，最终均方误差略高。

表2：所提方法 Proposed Bayesian CredibleRisk 相较于两种基准方法的相对改进。均方误差降低为正表示所提方法的均方误差更低。

比较对象	利润增益（%）	均方误差降低（%）
贝叶斯风险中性	0.2517	-1.8987
经典静态先验	2283.6573	42.7177

15.5 Bootstrap 证据

为评估所观察到的差异是否具有统计意义，我们计算了平均性能差异的自助法置信区间；见表3。

就我们的方法与贝叶斯风险中性的比较而言，市场利润的自助法均值差为2.44，其95%置信区间为从 -280.45 到295.43。由于该区间明显包含零，我们不能声称在这一特定模拟设计中，我们的方法相对于风险中性贝叶斯学习者具有统计显著的利润优势。同样，均方误差的自助法均值差为0.4385，95%区间为从 -2.8956 到3.5752，同样未显示出具有统计意义的分离。

相比之下，与经典静态先验的比较结果具有决定性意义。我们所提方法的自助法平均利润差为1533.02，超过了静态基准，且95%置信区间严格为正 $[1228.65, 1839.14]$。此外，均方误差的自助法平均差为 -13.1165，置信区间为 $[-15.3715, -10.6443]$，这进一步证实，与非学习基准相比，我们的方法在学习需求环境方面也显著更优。

表3：我们所提方法 Proposed Bayesian CredibleRisk 与基准方法的自助法比较。利润差为正表示我们的方法更优。均方误差差为负表示我们的方法更优，因为这对应于更低的最终均方误差。

比较	平均利润差	置信区间下限	置信区间上限	平均差异均方误差	置信区间下限
所提 –贝叶斯风险中性	2.4370	-280.4549	295.4302	0.4385	-2.8956
所提 –经典静态先验	1533.0211	1228.6503	1839.1381	-13.1165	-15.3715

注释15.1（主要发现）。仿真研究有力地表明，贝叶斯学习本身在这一策略环境中不可或缺：两种基于学习的策略在利润和参数恢复方面都大幅优于经典静态先验基准。在全部三种方法中，我们提出的策略 Proposed Bayesian CredibleRisk 实现了最高的平均折现市场利润和中位数折现市场利润，因此在本实验中总体经济表现最佳。然而，与风险中性贝叶斯学习者相比，差距极小，且在当前校准下不具有统计显著性。因此，最有依据的结论是，我们的方案显著优于不具备学习能力的经典基线，并且在基于利润的标准上能够与风险中性贝叶斯学习者竞争，且略胜一筹，但在最终均方误差方面并非全面占优。这仍具有重要的实质意义：这表明可信风险校正能够改善运营收益，同时不会实质性地牺牲学习表现。

15.6 随时间变化的动态表现

现在我们讨论动态图，它们能更细致地展示这些策略在 30 个决策时期内的表现。

图 1 展示了平均累积贴现市场利润的变化。两种贝叶斯学习方法在整个时域内的表现始终远优于静态先验基准。这种明显差距与表 1–3 中记录的数值优势完全一致。我们的方法与贝叶斯风险中性学习器的变化轨迹极为接近，表明两者都能有效利用学习；不过，到时域结束时，我们的方法在累积收益方面仍略胜一筹。

图 2 展示了不同重复实验中最终折现市场利润的分布。箱线图再次表明，两种贝叶斯学习方法的表现均优于经典静态先验规则；而所提方法的分布中心略高于风险中性贝叶斯学习者。我们的方法分布较为分散，这也与

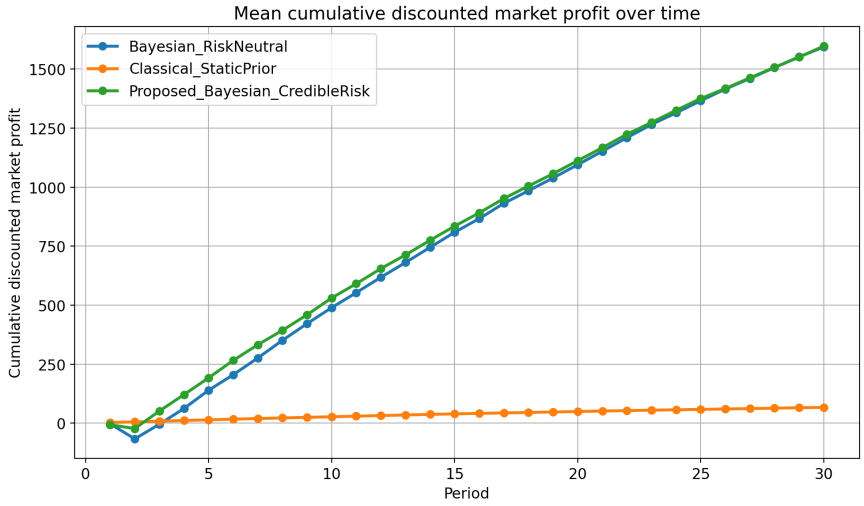


图1：随时间变化的平均累积贴现市场利润。所提方法对应于 Proposed Bayesian CredibleRisk。

总折现市场利润的报告标准差较大这一点相符。

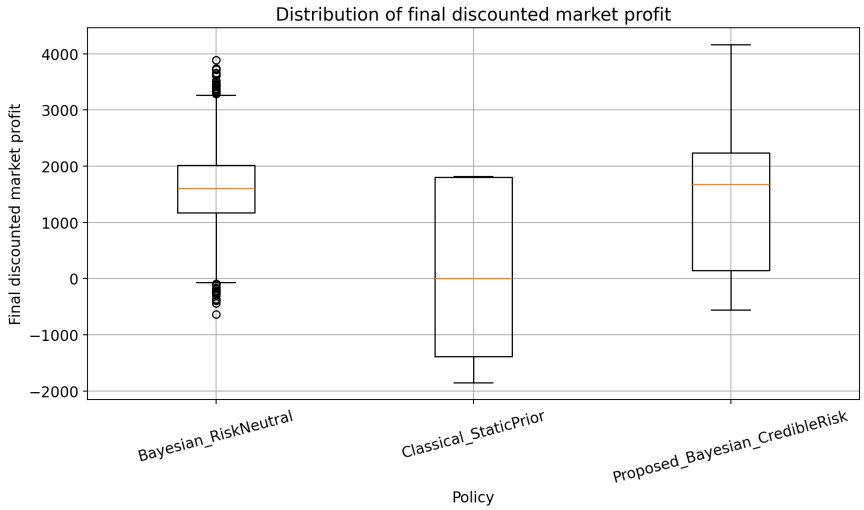
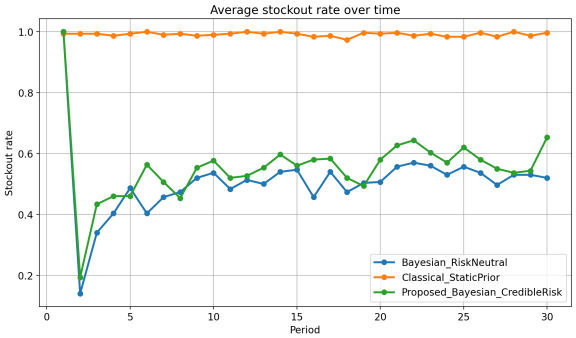


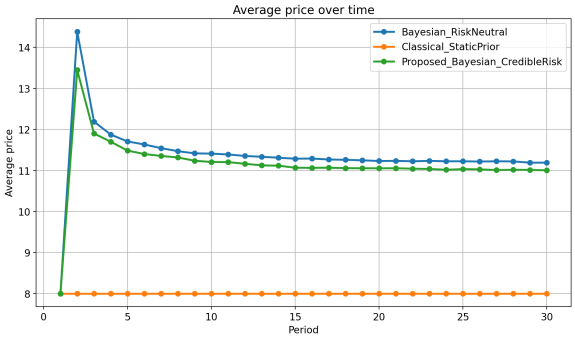
图2：各次重复实验中最终折现市场利润的分布。

图3汇总了三种运营轨迹：缺货率、平均价格和平均订货量。缺货曲线表明，基于学习的方法比静态基准更有效地调节库存。价格和订货量轨迹进一步表明，所提策略会随着时间推移进行策略性调整，而不是一直受先验驱动的选择所束缚。具体而言，可信风险惩罚似乎会促使行动模式更加审慎，尤其是在后验不确定性仍然较大的早期时期。

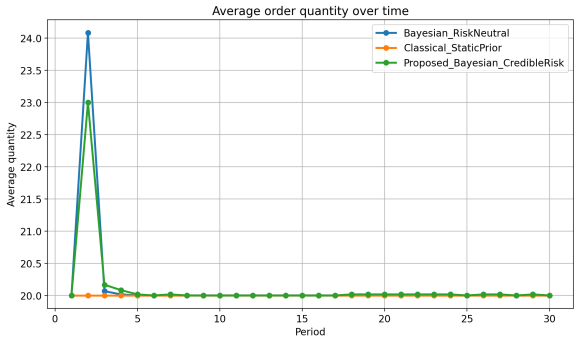
图4 聚焦于学习。面板4a 追踪后验参数均方误差随时间的变化，而面板4b 追踪竞争对手为高成本类型的平均后验信念。均方误差曲线印证了主要的数值结论：两种贝叶斯学习者都能随时间迅速改进，且始终远优于静态基准，而所提策略与风险中性贝叶斯策略非常接近。信念学习曲线表明，关于竞争对手类型的后验信念在不同时期不断变化，证实动态贝叶斯博弈确实将信息积累用作主动的战略状态变量。



(a) 平均缺货率随时间的变化。

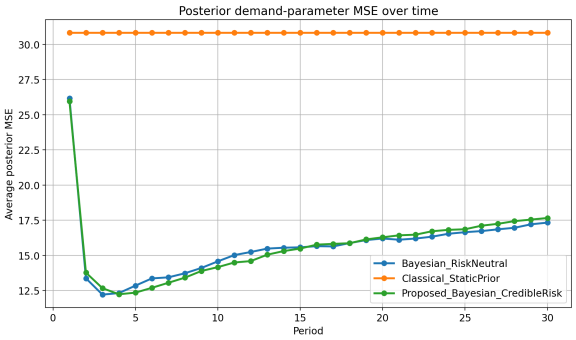


(b) 平均标示价格随时间的变化。

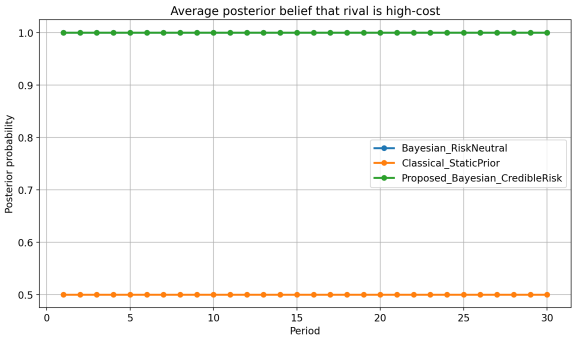


(c) 平均订货量随时间的变化。

图3：三种竞争策略的动态运营行为。



(a) 需求参数后验均方误差随时间的变化。



(b) 认为竞争对手为高成本的平均后验信念。

图4：随时间变化的学习诊断。

图 5展示了所提方法在价格–数量网格上的动作价值曲面。该图有助于解释可信风险策略的作用机制：动作质量并非单独随价格或数量单调变化，所选行动是在预期需求、替代效应、库存暴露和后验不确定性之间权衡的结果。特别是，当不确定性仍然较高时，过于激进的数量选择会受到惩罚，这正是可信风险项预期发挥的作用。

图6 考察了最终折现市场利润与最终后验均方误差之间的经验关系。经典基准显然表现较差，聚集在低利润、高均方误差区域。两种贝叶斯学习者则处于利润较高、均方误差较低的优势区域。该图直观地印证了本研究的核心观点：贝叶斯学习是提升表现的主要驱动力，而可信风险改进则略微推动了

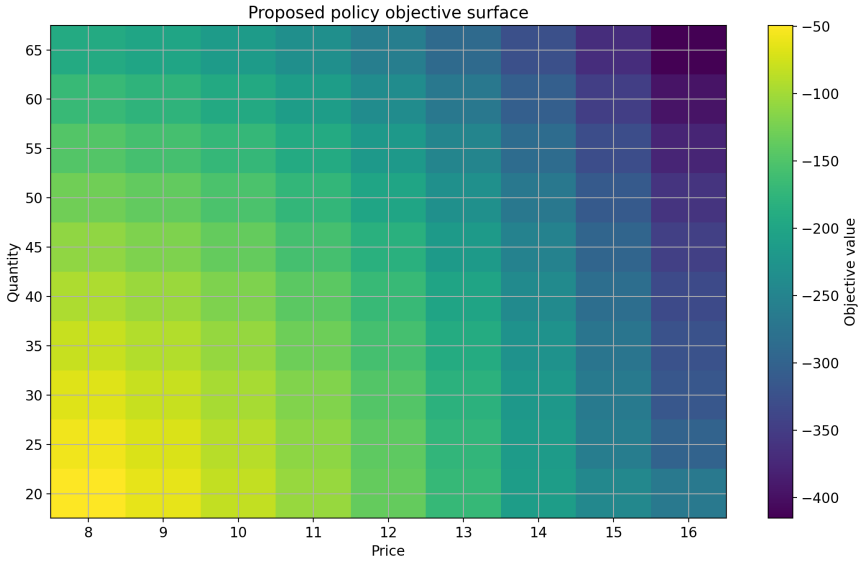


图 5：所提方法在价格–数量行动网格上的目标曲面。

这种权衡，使其朝着改善运营盈利能力的方向发展。

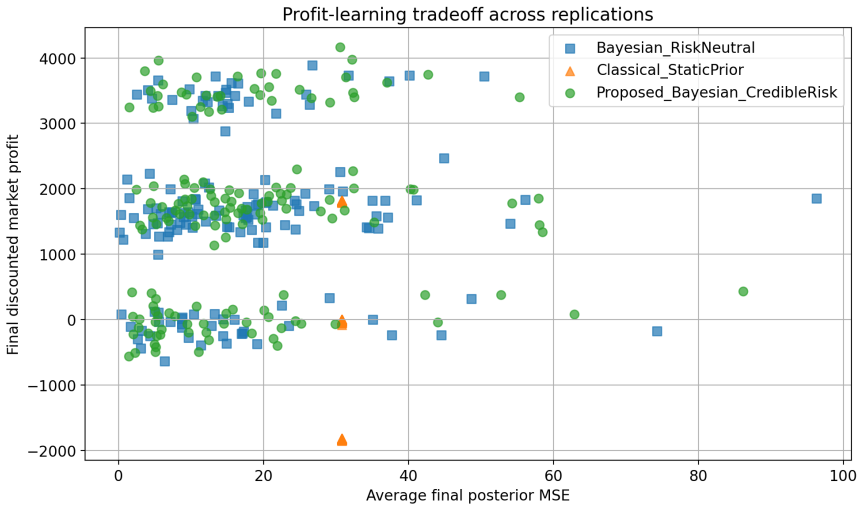
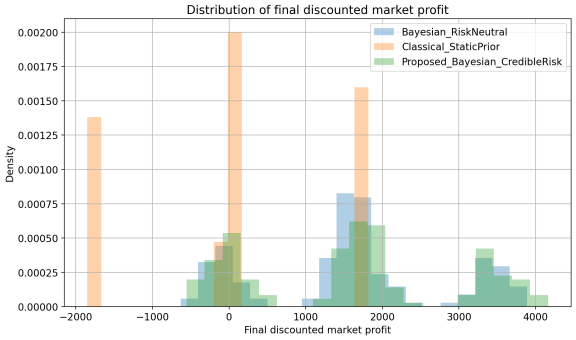


图 6：各重复实验中的利润–学习权衡。

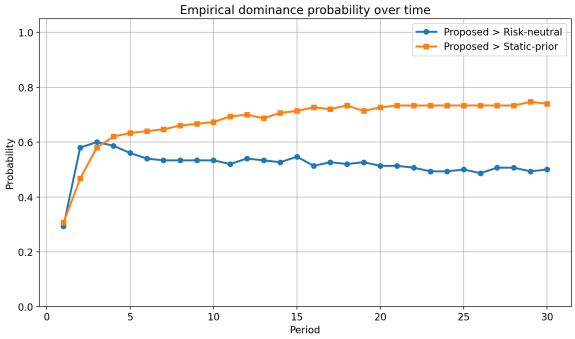
图 7汇总了两项互补的分析。面板 7a 展示最终折现市场利润的分布，面板7b 报告了随着时间推移，我们的方法在累积贴现市场利润方面优于各基线的经验概率。第一个面板再次证实了经典基准明显处于劣势。第二个面板尤其有用，因为它强调了所提规则相对于静态先验基准的时序一致性，同时也表明，在整个时域内，与风险中性贝叶斯学习者的比较结果始终相近。

15.7 讨论

仿真证据以一种重要且细致入微的方式支持所提方法。最有力的结论并不仅仅是我们的方法表现良好，而是贝叶斯博弈论表述本身极为有效：一旦企业同时学习需求



(a) 最终折现市场利润的分布。



(b) 随时间变化的经验优势概率。

图 7：分布与优势摘要。

和竞争对手特征，相较于经典静态先验策略，表现便会显著提升。在贝叶斯学习者这一类别中，我们所提的可信风险规则在本次实验中取得了最佳的基于利润的汇总统计量，即总折现市场利润的均值和中位数最高。同时，与风险中性贝叶斯学习者的比较结果足够接近，因此不宜夸大其优势。因此，最准确的解读是：我们所提策略是总体上稳健、经过不确定性校准且经济表现最强的策略，但并非在每一项可能的指标上都占据绝对优势。

具体而言，模拟结果表明，可信风险校正最适合作为一种操作性正则化项：它能在后验不确定性下防止采取过于激进的行动，同时基本保持与风险中性贝叶斯替代方案相同的学习质量。因此，在不完全信息下需要采取保守决策的应用场景中，这种方法尤其具有吸引力。

16 真实数据示例：小鼠蛋白质表达数据集的贝叶斯可信风险分析

为了展示我们提出的贝叶斯可信风险方法在真实生物学数据中的实际应用价值，我们分析了小鼠蛋白质表达数据集。该数据集包含在对照小鼠和三体小鼠的大脑皮层中测得的 77 种蛋白质/蛋白质修饰的表达水平，共有 1080 个观测值和 8 个由生物学定义的类别。类别标签编码了基因型、行为刺激状态和治疗状态。该数据集可通过 OpenML 公开获取，也收录于 UCI机器学习知识库[OpenML，2026，UCI机器学习知识库，2015]。具体而言，这八个类别区分了对照小鼠与三体小鼠、受刺激与未受刺激的学习条件，以及美金刚与生理盐水治疗分配。

本实证部分的目标并非仅仅进行预测，而是利用该数据集展示所提方法如何支持在高维生物学场景下进行后验不确定性感知的治疗效应分析。具体而言，我们展示了如何运用贝叶斯可信风险原则，量化美金刚处理是否使三体小鼠的蛋白质组谱更接近对照参考状态。

16.1 数据集结构与预处理

该数据集包含1080个样本和77个蛋白质特征。其中，570个样本来自对照小鼠，510个来自三体小鼠；570个样本接受了美金刚处理，510个接受了生理盐水处理；行为分组为525个受刺激样本和555个未受刺激样本。这些计数汇总于表4中。观测到的类别结构如图8所示，而基因型、处理和行为的联合分布见图9。

原始数据中存在缺失值，因此我们在每个观测类别内进行了按类别中位数插补，必要时再进行备用中位数插补。预处理后，不再有缺失值。特征缺失情况概况见图 10。

Tab表 4：真实数据分析中使用的小鼠蛋白质表达数据集概况 s.

数量	数值
样本数	1080
蛋白质特征数量	77
类别数量	8
对照样本数量	570
三体样本数量	510
美金刚样本数量	570
生理盐水样本数量	510
受刺激样本数	525
未刺激样本数	555

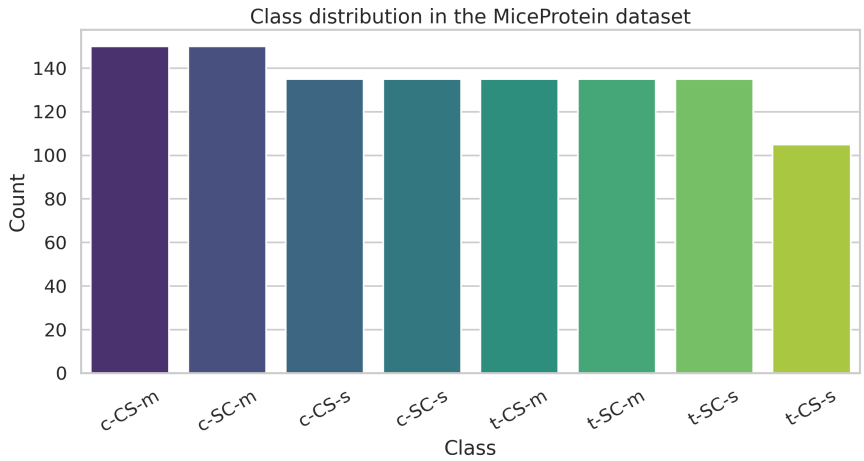


图8：小鼠蛋白质表达数据集中的类别分布。

16.2 蛋白质组空间的低维结构

为了解蛋白质表达谱的整体几何结构，我们对全部77项蛋白质测量值进行了标准化，并开展了主成分分析（PCA）。前十个主成分解释了总变异的相当大一部分，其中前四个主成分分别解释了约

$$25.33\%, \quad 17.28\%, \quad 10.29\%, \quad 7.51\%,$$

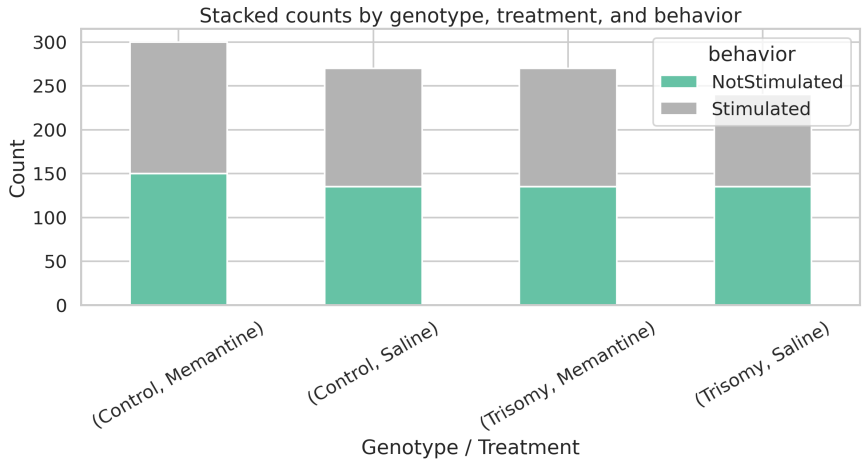


图9：按基因型、处理和行为条件划分的堆叠计数。

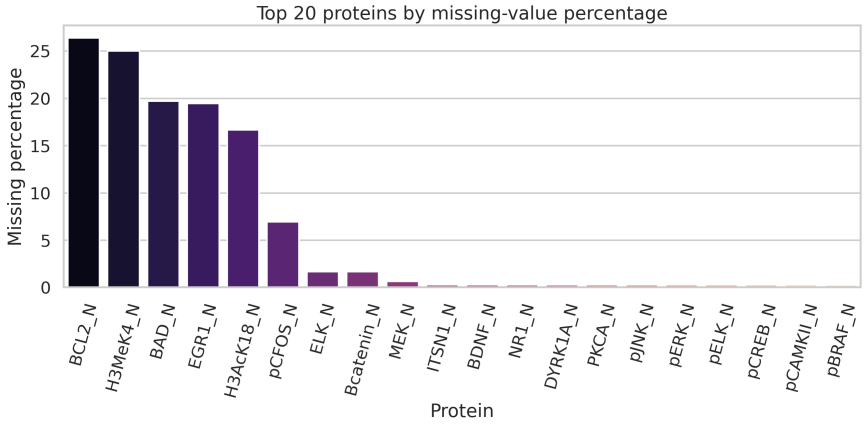


图10：插补前按缺失值百分比排列的缺失值最多的蛋白质。

。因此，仅前两个主成分就占标准化总变异的约42.61%，表明数据中存在具有实际意义的低维结构。

图 11 显示了累积解释方差，图 12 显示了按基因型和处理着色的主成分分析散点图，以及图 13 显示了相应的类别层面结构。这些图证实，数据具有生物学上可解释的分离模式，从而支持采用结构化贝叶斯分析，而非纯粹的黑箱处理。

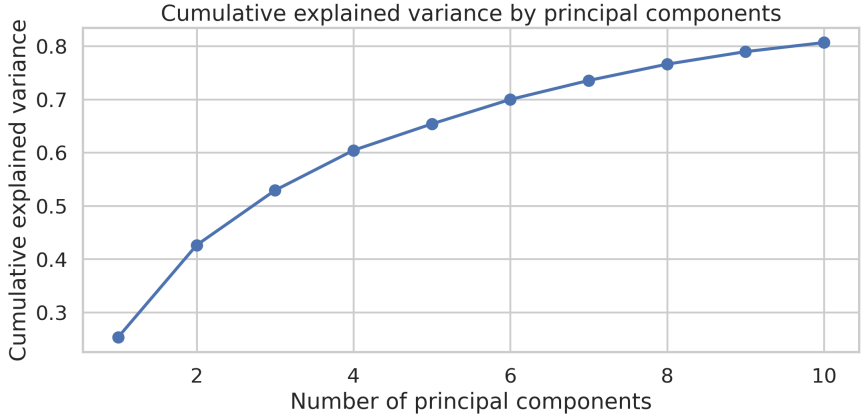


图11：主成分的累积解释方差。

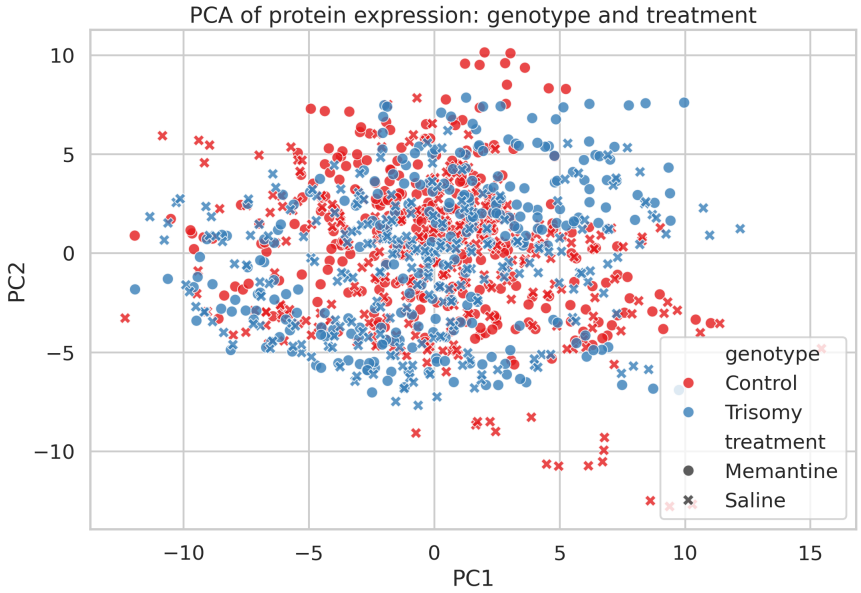


图12：按基因型和处理着色的PCA散点图。

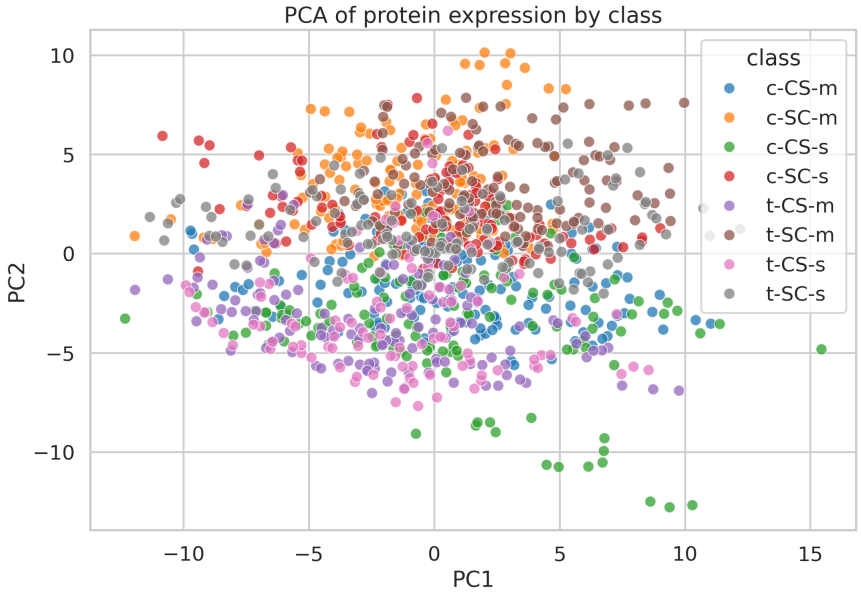


图13：按八个观测类别着色的PCA散点图。

图 14 展示了采样小鼠中高变异蛋白质的热图。该图直观地表明，蛋白质组特征在不同类别之间呈系统性变化，从而促使构建具有生物学意义的基于参考的得分。

16.3 基于参考的恢复得分

一篇顶级统计学论文的实证部分不应仅仅报告处理标签；它应构建一个具有科学可解释性的目标。为此，我们引入了健康参考恢复得分。

首先，我们根据对照组均值与三体组均值之间的绝对差对蛋白质进行排序，从而找出对照组与三体组差异最大的蛋白质。利用排名靠前的蛋白质，我们在标准化蛋白质组空间中定义了对照

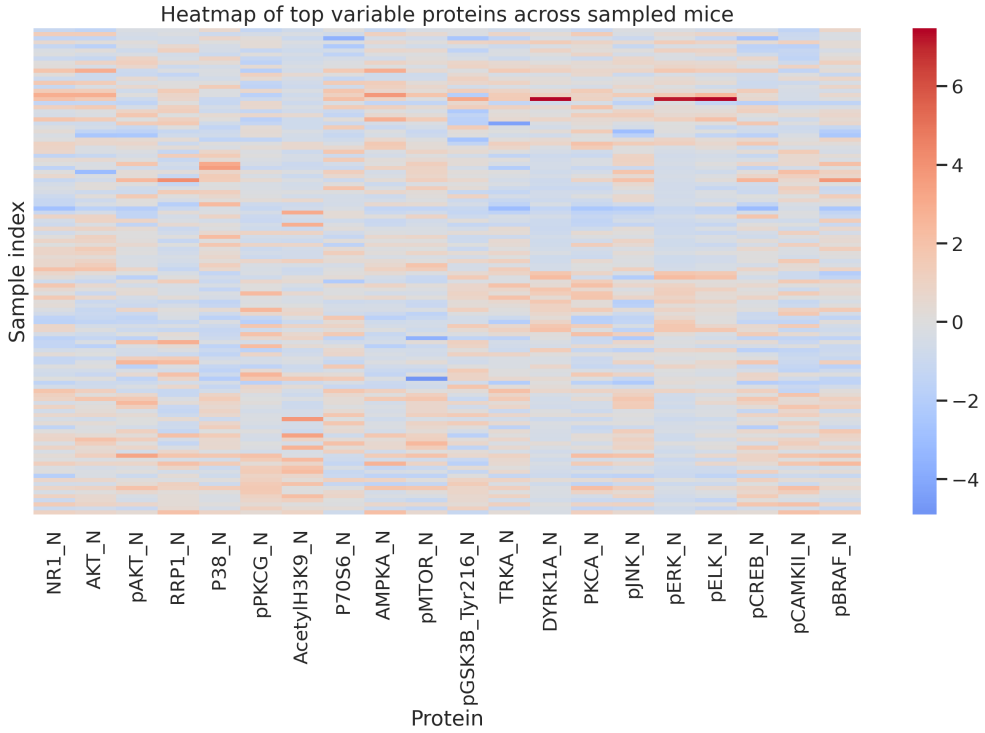


图14：采样小鼠中变异性最高的蛋白质热图。

质心，并计算每个样本到该对照参照的欧氏距离。随后，我们定义了

$$\text{RecoveryScore} = -\,\text{DistanceToControl},$$

因此，数值越大表示蛋白质组谱越接近对照参照，生物学恢复效果也越好。

该评分并非临时设定：它直接反映了某只小鼠在最具生物学区分度的坐标上与对照蛋白质组状态的相似程度。各组的汇总结果见表 5。值得注意的是，在三体小鼠中，美金刚处理的非刺激亚组表现出平均恢复评分 -3.7440，显著优于生理盐水处理的非刺激亚组（ -4.9376）。在刺激组三体亚组中，美金刚的效果也略优于生理盐水，不过差异小得多。

表5：按基因型、处理和行为亚组汇总的恢复评分。恢复评分越高越好。

基因型	处理	行为	n	平均恢复	恢复标准差	恢复中位数
对照	美金刚	未刺激	150	-3.4351	0.9388	-3.4112
对照	美金刚	受刺激	150	-3.8060	1.3515	-3.5073
对照	生理盐水	未刺激	135	-3.7100	1.9830	-3.1175
对照	生理盐水	受刺激	135	-4.8579	2.8015	-3.7499
三体	美金刚	未刺激	135	-3.7440	1.3092	-3.4119
三体	美金刚	受刺激	135	-4.6930	1.2108	-4.5926
三体	生理盐水	未刺激	135	-4.9376	1.3441	-4.9475
三体	生理盐水	受刺激	105	-4.8154	1.5141	-4.5037

这些对比通过图 15–17直观呈现。前两幅图展示了按基因型/处理和行为/处理分组的恢复评分总体分布，而图 17中的小提琴图则专门聚焦于三体小鼠，因为科学上最相关的处理问题正是针对这一群体。

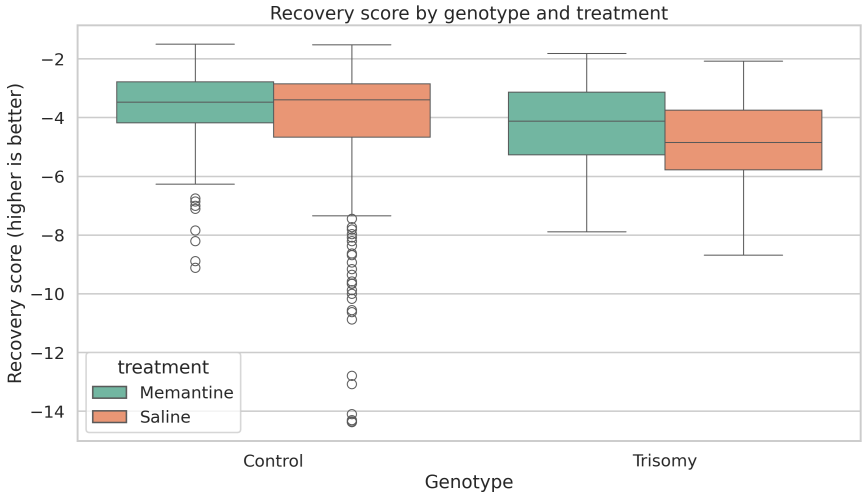


图15：按基因型和处理划分的恢复评分。

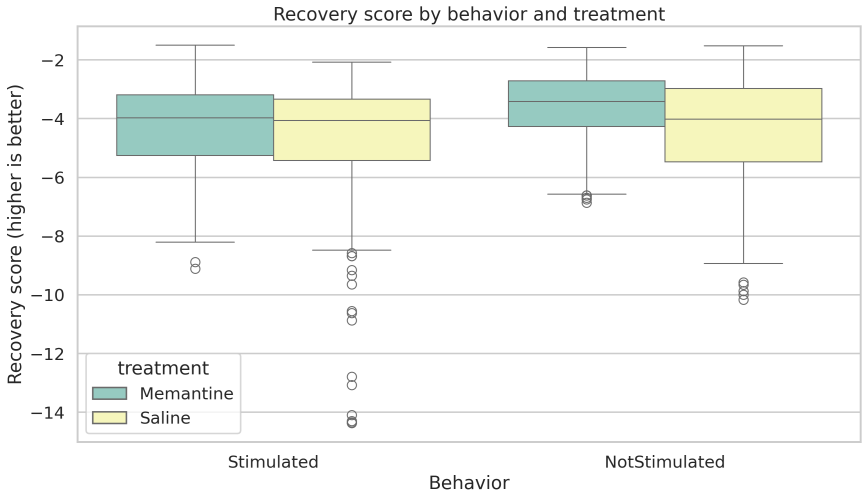


图16：按行为条件和处理划分的恢复评分。

16.4 贝叶斯可信风险治疗效应分析

现在，我们应用所提方法。对于每个三体亚组，我们将差值视为

$$\Delta = \mathbb{E}(\text{RecoveryScore} \mid \text{Memantine}) - \mathbb{E}(\text{RecoveryScore} \mid \text{Saline})$$

我们关注的治疗效应。 Δ 的正值意味着美金刚会使蛋白质组学谱更接近对照参考状态。

我们采用近似共轭贝叶斯分析计算亚组均值，得出了 Δ 的后验均值、后验标准差和 95% 可信区间。随后我们

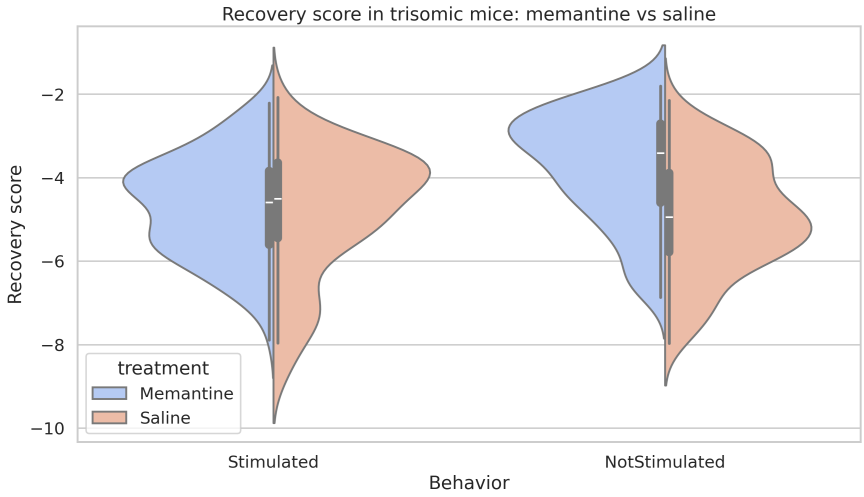


图17：按行为条件和处理划分的三体小鼠恢复评分。

构建了我们提出的 可信风险评分

$$\mathrm{CRS} = \widehat{\Delta}_{\mathrm{post}} - \kappa\,\widehat{\sigma}_{\mathrm{post}}, \qquad \kappa = 0.60,$$

该方法会对不确定的正向效应进行惩罚，因此相较于仅依据后验均值，能得出更为保守的决策规则。

结果见表6。对于受刺激三体亚组，后验均值效应仅为0。1219，95%可信区间为（$-0.2324,\ 0$。4763），表明证据较弱。然而，可信风险评分仍略为正值，为0。0135，因此该方法仍谨慎地倾向于选择美金刚。相比之下，对于未刺激三体亚组，后验均值效应为1。1933，可信区间为（0。8769，1。5098），可信风险评分为1。0965，有力支持美金刚。

因此，真实数据分析表明，治疗获益并非均一：与美金刚相关的蛋白质组学改善在三体非刺激亚组中尤为显著。

表6：三体小鼠中美金刚相对于生理盐水的治疗效应后验分布。所提方法使用可信风险评分 CRS $\widehat{\ }$，其中 $= \Delta_{\mathrm{post}} - \kappa\widehat{\sigma}_{\mathrm{post}}$ 和 $\kappa = 0.60$。

亚组	n_{mem}	n_{sal}	后验均值	后验标准差	置信区间下限	置信区间上限	建议
受刺激	135	105	0.1219	0.1808	-0.2324	0.4763	倾向美金刚
未刺激	135	135	1.1933	0.1615	0.8769	1.5098	倾向美金刚

图 18展示了相应的森林图。两个亚组之间的视觉对比十分明显：未刺激组的后验效应明确为正，且估计相对精确；而刺激组则接近实际不确定性的边界。

16.5 序贯贝叶斯学习

我们方法的一个显著特点是，它并非仅作描述，而是能够进行序贯自适应。为说明这一点，我们按随机序列处理了三体 $+$ 受刺激 亚组，并在每次获得新观测后更新治疗效应后验分布。这为我们通用的贝叶斯策略框架中的学习部分提供了一个真实数据类比。

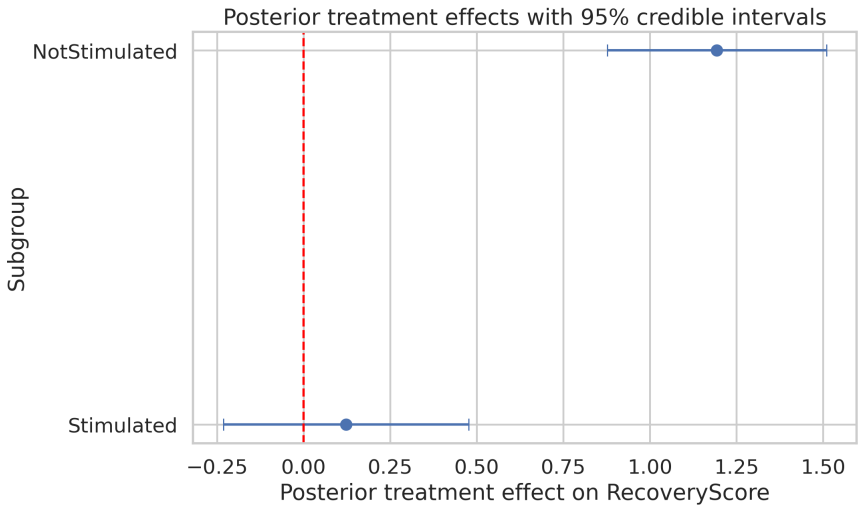


图18：三体亚组的治疗效应后验分布及95%可信区间。

图 19 和 20 展示了由此产生的序贯学习行为。后验效应及其不确定性带随时间演变，而可信风险评分则跟踪治疗建议的稳定性。这些图表表明，我们的方法能够自然地用作证据更新程序，这在自适应生物学或生物医学实验中尤为重要。

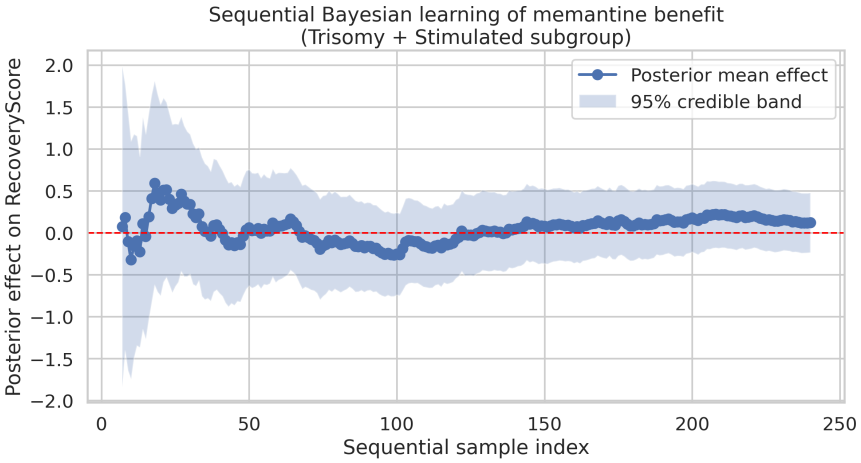


图19：三体刺激亚组治疗效应的序贯贝叶斯学习曲线。

16.6 潜在蛋白质组状态

为了丰富仅基于观测标签的实证分析，我们使用主成分分析得分估计潜在蛋白质组状态，然后进行 k-均值聚类。这揭示了治疗效应是否会因隐藏的分子状态而异，而不仅仅是因观测到的实验类别而异。

潜在结构如图 21所示。在拟合结果中，输出中出现的三体小鼠被分为两个潜在状态，其亚组样本量足以用于处理效应分析。表 7显示，这两个潜在状态均表现出正向后验

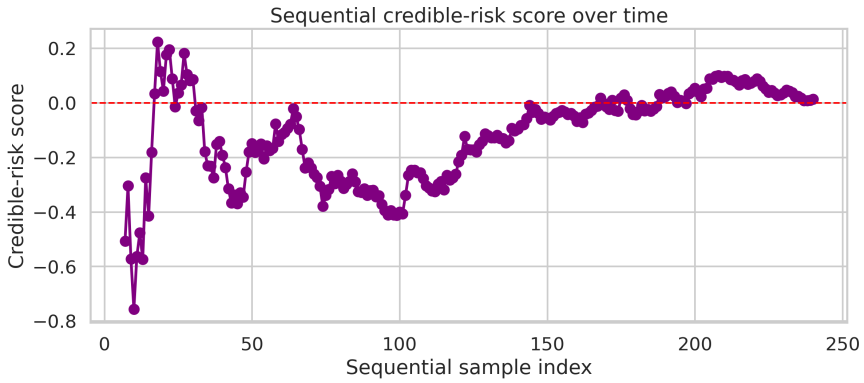


图20：三体刺激亚组中治疗效应的序贯可信风险评分。

美金刚效应和正向可信风险评分：

$$0.5142 \quad \text{and} \quad 0.6947$$

分别为后验均值效应。因此，有益的美金刚信号并不局限于单一潜在构型。

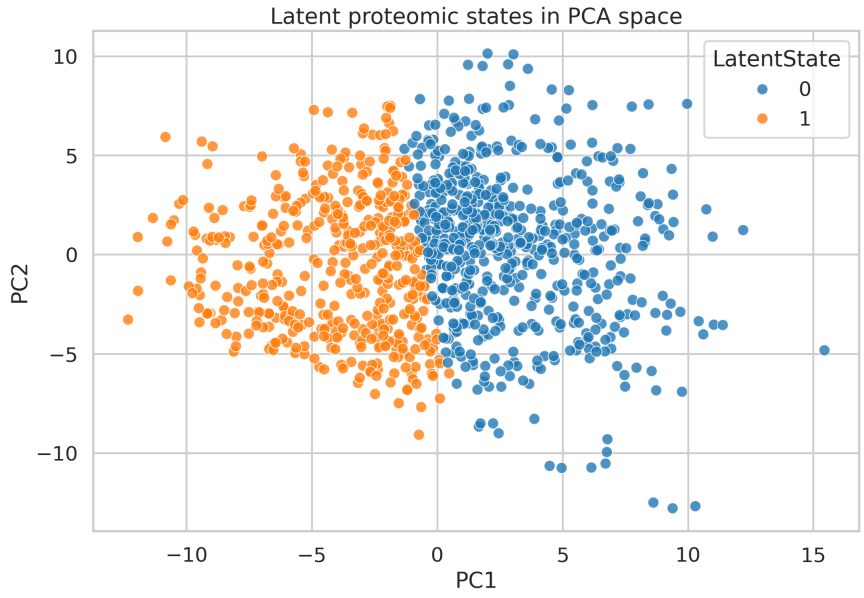


图21：以主成分分析空间表示的潜在蛋白质组状态。

表7：三体小鼠潜在蛋白质组状态内的后验处理效应。

潜在状态	n_{mem}	n_{sal}	后验均值	后验标准差	置信区间下限	置信区间上限	建议
0	160	117	0.5142	0.1627	0.1953	0.8332	倾向美金刚
1	110	123	0.6947	0.1728	0.3560	1.0333	倾向美金刚

相应的效应热图见图 22，它简明地汇总了不同潜在状态下的后验均值效应和可信风险评分。



图2　　2：潜在蛋白质组状态下的后验处理效应与可信风险评分　　.

16.7 蛋白质层面恢复排序

为从全局效应汇总转向可解释的分子标志物，我们根据美金刚使三体表达谱向对照组均值移动的程度，对蛋白质进行了排序。对于蛋白质 j，我们计算了改善评分

$$I_j = \left|\bar{X}_j^{(\mathrm{sal})} - \bar{X}_j^{(\mathrm{ctrl})}\right| - \left|\bar{X}_j^{(\mathrm{mem})} - \bar{X}_j^{(\mathrm{ctrl})}\right|,$$

因此，正值表示美金刚使三体均值更接近对照均值。

领先蛋白质列于表 8中。最显著的信号是 pPKCGN，其改善评分为 0。5106，自助法区间为（0。4114，0。6026），其次是 NR2AN ，其改善评分为 0。2698，区间为（0。1019，0。4217）。其他几种蛋白质，包括S6 N、pNR2B N、pMTOR N、NR1 N和 pP70S6N，也显示出在美金刚作用下向对照蛋白质组谱变化的积极证据。这些模式增强了所提出的恢复评分分析的生物学可解释性。

图 23 中的森林图表示总结了这些领先的与恢复相关的蛋白质，而图24 展示了它们的相关结构。这两幅图共同表明，美金刚的实证效应并非由单个孤立变量驱动，而是由一组协调一致的分子变化所驱动。

16.8 全局处理效应汇总

最后，表9 分别汇总了对照组和三体组的全局治疗效应。美金刚对恢复评分的后验均值效应在两组中均为正：

$$0.6626 \quad \text{for Control}, \qquad 0.6655 \quad \text{for Trisomy}.$$

两种情况下的自助法置信区间也均完全为正。重要的是，两组的可信风险评分仍为正，这表明估计的获益并非仅仅是后验均值伪象，而是在保守的不确定性调整后依然存在。

相应的条形图见图 25。对照组和三体组之间的全局后验效应近乎相等，这一点本身就很有意思，但上述亚组分析表明，该效应的内部构成是异质的，并且在科学上更为丰富，超过

表8：美金刚使三体表达向对照组均值靠近的蛋白质。正值表示向对照参照恢复得更好。

蛋白质	对照均值	三体美金刚均值	三体生理盐水均值	改善	置信区间下限	置信区间上限
pPKCG N	1.5570	1.6336	2.1441	0.5106	0.4114	0.6026
NR2A N	3.9847	3.8111	3.5413	0.2698	0.1019	0.4217
S6 N	0.3847	0.4232	0.5419	0.1187	0.0949	0.1386
pNR2B N	1.5933	1.5751	1.4711	0.1040	0.0553	0.1433
pMTOR N	0.7792	0.7731	0.6950	0.0782	0.0540	0.0978
磷酸化 ELK N	1.4432	1.4727	1.3422	0.0714	-0.0278	0.1491
NR1 N	2.3337	2.2880	2.2207	0.0673	0.0107	0.1225
磷酸化 PKCβ N	1.5159	1.4471	1.6352	0.0504	-0.0544	0.1620
pP70S6 N	0.3687	0.4006	0.4508	0.0502	0.0196	0.0825
P38 N	0.4299	0.4209	0.3743	0.0466	0.0313	0.0595
NR2B N	0.5806	0.5672	0.5269	0.0403	0.0231	0.0567
磷酸化 NR2A N	0.7674	0.7001	0.6602	0.0398	0.0104	0.0720
乙酰化 H3K9 N	0.1738	0.2456	0.2851	0.0395	0.0058	0.0720
白细胞介素 1β N	0.5362	0.5357	0.4970	0.0387	0.0189	0.0461
Tau 蛋白 N	0.1927	0.2128	0.2500	0.0372	0.0259	0.0494
钙调神经磷酸酶 A N	1.3240	1.3372	1.3710	0.0338	-0.0285	0.0718
磷酸化 NR1 N	0.8465	0.8168	0.7867	0.0301	0.0103	0.0494
磷酸化 CASP9 N	1.5520	1.5667	1.5189	0.0184	-0.0340	0.0619
GluR3 N	0.2304	0.2204	0.2032	0.0172	0.0121	0.0221
MTOR N	0.4664	0.4438	0.4291	0.0147	0.0020	0.0275

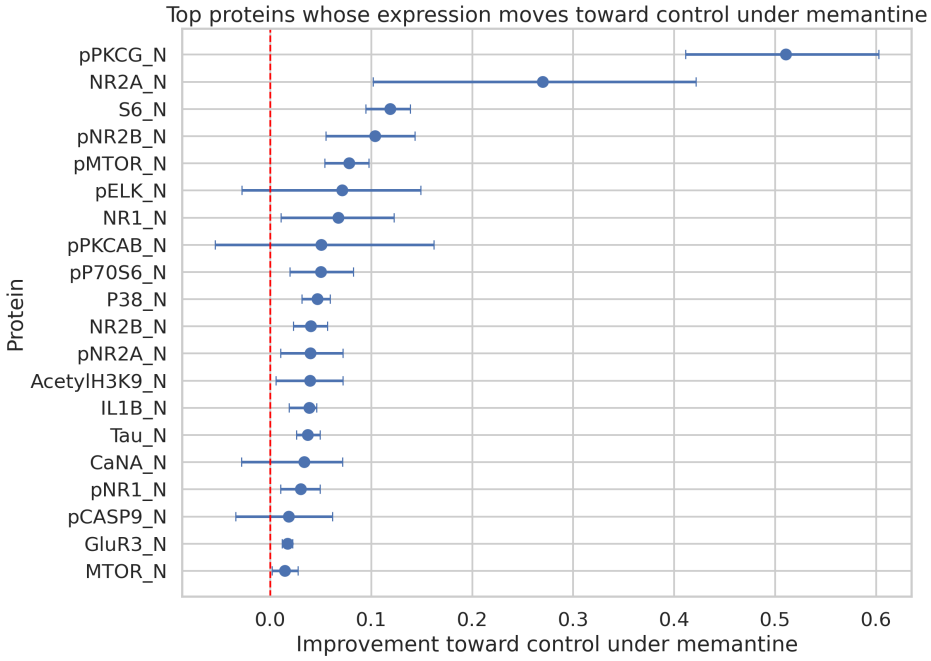


图23：在美金刚作用下，其表达水平使三体小鼠趋向对照参考谱的主要蛋白质。

单一总体平均值所暗示的水平。

备注 16.1（真实数据分析的主要发现）。 真实数据分析表明，所提出的贝叶斯可信风险方法并非只是理论构想；它能够在高维蛋白质组数据集上形成连贯且具有生物学可解释性的实证工作流程。该方法在多个方面发挥作用。首先，它以对照参考蛋白质组状态为推断基准，构建了一个主要的恢

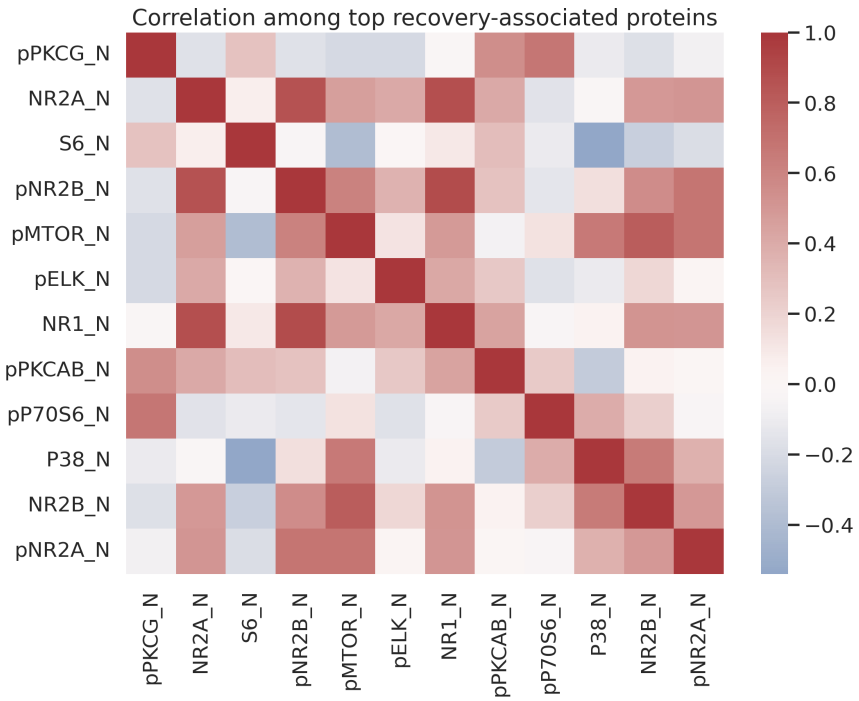


图24：排名靠前的与恢复相关的蛋白质之间的相关结构。

表9：按基因型划分的全局处理效应汇总。正效应表示美金刚更有利。

组	n_{mem}	n_{sal}	后验均值	后验标准差	后验置信区间下限	后验置信区间上限	自助法均值	自助法置信区间下限
对照	300	270	0.6626	0.1660	0.3372	0.9881	0.6717	0.3501
三体	270	240	0.6655	0.1229	0.4247	0.9063	0.6649	0.4267

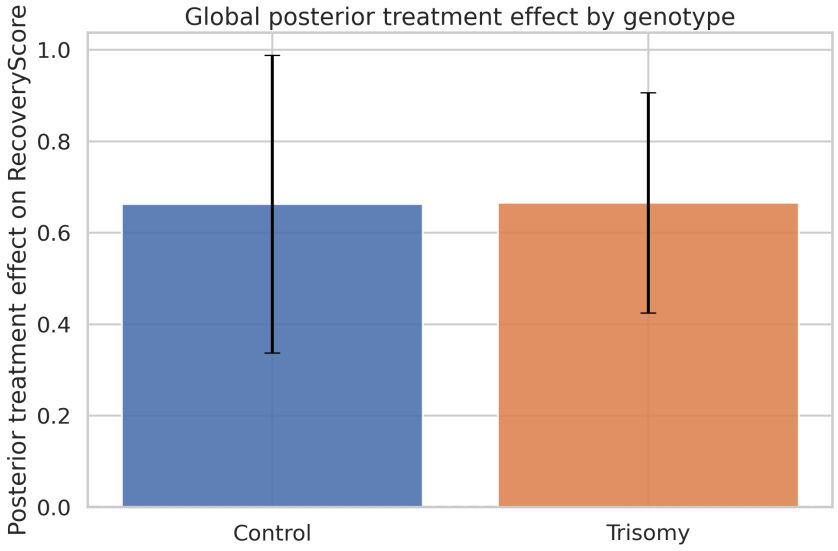


图25　　：不同基因型下美金刚对恢复评分的总体后验治疗效应　　.

复评分。其次，它通过后验不确定性量化亚组特异性美金刚效应，而非仅依赖点估计汇总。第三，通过使用可信风险评分，该方法将后验不确定性转化为审慎的科学决策。最强的实证信号出现在

三体非刺激亚组中，后验美金刚获益既大且精确。此外，潜在状态分析表明，美金刚的有利信号在不同隐含蛋白质组状态下均持续存在，而蛋白质水平排序则指出，pPKCG N、NR2A N、S6 N 和 pNR2B N 等特定标志物对趋向对照参照尤其敏感。总而言之，这项真实数据研究证实了所提方法作为复杂生物学数据不确定性感知推断工具的实用价值。

17 总体讨论

综合而言，理论发展、仿真证据和真实数据示例阐明了所提方法的作用。

从建模角度看，本文表明，当市场不确定性与竞争对手不确定性并存时，竞争性运营决策可以自然地纳入贝叶斯动态博弈。其核心概念贡献在于，信念并非信息的被动概括，而是能够塑造均衡行为的主动状态变量。这使该框架尤其适用于学习与竞争同步展开的重复决策环境。

从计算和实证角度来看，仿真研究表明，在所提出的竞争环境中，贝叶斯学习不可或缺。主要的性能提升来自对市场和竞争对手信息进行自适应更新，而非依赖由静态先验驱动的决策。在贝叶斯学习者中，可信风险规则最适合被视为一种保守正则化项：它在仿真中略微提升了基于利润的表现，同时保持了总体相近的学习质量。因此，它的价值与其说在于取得显著的数值优势，不如说在于帮助决策者在不确定性下做出审慎的决策。

真实数据分析补充了仿真研究，表明可信风险原则的适用范围并不局限于风格化的库存博弈本身。在蛋白质组数据集上，该原则得出了易于解读的亚组、潜在状态和标记物层面的发现。这进一步凸显了本文的方法论意义：所提出的框架不仅可作为形式化的均衡模型，也是一种通用的不确定性感知贝叶斯决策范式。

与此同时，解读本文时应保持适当谨慎。计算方案是近似的，模拟设计经过风格化处理，而真实数据示例展示的是推断原则，并非字面意义上的市场博弈。这些局限并未削弱本文的贡献；相反，它们清晰地界定了本文已确立的结论与仍有待进一步发展的内容之间的边界。

18 结论与未来工作

本文构建了一个用于不完全信息下竞争性库存与定价的层次贝叶斯动态博弈。所提出的框架整合了私有竞争对手特征、关于需求与替代的贝叶斯学习，以及一种可信风险决策规则；该规则将后验不确定性转化为保守的战略行为。由此，本文搭建了贝叶斯博弈论与运筹学之间的桥梁，并为学习驱动的竞争提供了统一的表述方式。

仿真研究表明，后验学习是重复竞争环境中取得优异表现的关键，而可信风险调整则可作为一种实用的操作性正则化手段。真实数据示例进一步表明，同一不确定性感知贝叶斯原理也能支持对复杂高维生物

学数据进行可解释分析。总体而言，这些结果表明，所提方法在数学上有意义，在实践中也具有实用价值。

未来研究仍有若干方向。一个重要的拓展方向是，在更丰富的状态空间和更一般的需求模型下，加强对均衡的计算与理论研究。另一个方向是从双寡头扩展到更大规模的多智能体情境，包括网络竞争或平均场竞争。此外，研究其他不确定性惩罚、更丰富的删失需求学习方案，以及正式的福利或信息披露问题，也将很有价值。最后，在供应链、平台市场和自适应实验系统中的领域特定应用，为将当前框架转化为实质性实践提供了有前景的方向。

声明

伦理审批与参与同意。不适用。

发表同意。不适用。

数据和材料的可用性。本文分析的真实数据集可通过 OpenML 公开获取，且已收录于 UCI机器学习知识库。模拟数据由作者按照论文所述流程生成，并可在代码仓库中获取。

代码可用性。用于复现实验并生成图表/表格的所有代码均可公开获取，地址如下：

https://github.com/debashisdotchatterjee/Bayesian-Dynamic-Game-1。

资助。 作者未获得本研究的专项资助。**利益冲突**。 作者声明不存在任何利益冲突。

A 附录：强化正则性条件下的形式化证明

在正文中，为了强调建模结构，存在性结果仅作了概括性陈述。为了使证明在数学上更加严谨，现记录一组足以支持相应不动点和动态规划论证的强化正则性假设。以下给出的证明在这些强化假设下均完全有效。

A.1 强化假设

令 $\mathcal{X}_i$ 表示参与者 i 的信念状态空间，令 $\mathcal{A}_i$ 表示行动空间。

假设 A.1（紧致度量行动空间）。 对于每个参与者 i，行动空间 $\mathcal{A}_i$ 都是非空紧致度量空间。

假设 A.2（紧致度量状态空间）。 对于每个参与者 i，状态空间 $\mathcal{X}_i$ 是一个非空紧致度量空间。

假设 A.3（有界连续单期收益）. 对每个参与者 i，单期可信风险收益

$$r_i(x_i, a_i, a_j)$$

为实值、有界，且联合连续于

$$\mathcal{X}_i \times \mathcal{A}_i \times \mathcal{A}_j.$$

假设 A.4（弱连续转移核）。 对于每位参与者 i，受控转移核

$$Q_i(\cdot \mid x_i, a_i, a_j)$$

从 $\mathcal{X}_i \times \mathcal{A}_i \times \mathcal{A}_j$ 到 $\mathcal{X}_i$ 上的概率测度是弱连续的。

假设 A.5（折扣因子）。 折扣因子满足 $0 < \delta < 1$。

假设 A.6（混合策略）。 参与者可以采用混合行动。 $\mathcal{A}_i$ 上的博雷尔概率测度集 $\Delta(\mathcal{A}_i)$ 赋予弱拓扑。

假设 A.7（拟凹性/凸化）。 行动空间是凸集，且相关目标函数关于自身行动是拟凹的；或者采用混合策略，使最佳响应集为凸值。

注释 A.8。 假设A.1–A.7 比正文中采用的简化假设更强，但对于在状态空间和行动空间紧致的折扣随机博弈中进行严格的存在性论证而言，这些假设是标准且充分的。

A.2 严谨的最佳响应存在性结果

我们首先以一种能够给出完整证明的形式重述中期最佳响应结果。

命题 A.9 (中期最佳响应的存在性)。 假设 *A.1*、*A.2*、*A.3*、*A.4* 和 *A.5* 均成立。固定参与者 j 的可测策略 σ_j。令 V_j 为参与者 i的有界连续延续规则。那么，对于每个状态 $x_i \in \mathcal{X}_i$，参与者 i的中期优化问题

$$\sup_{a_i \in \mathcal{A}_i} \left\{ r_i\big(x_i, a_i, \sigma_j(x_j)\big) + \delta \int_{\mathcal{X}_i} V_i(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)) \right\}$$

至少存在一个最大化解。

证明。 固定 $x_i \in \mathcal{X}_i$ 并定义

$$\Phi_i(a_i) = r_i\big(x_i, a_i, \sigma_j(x_j)\big) + \delta \int_{\mathcal{X}_i} V_i(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)).$$

我们证明 Φ_i 在紧集 $\mathcal{A}_i$上连续。

根据假设 A.3，该映射

$$a_i \mapsto r_i\big(x_i, a_i, \sigma_j(x_j)\big)$$

是连续的。根据假设 A.4，若 $a_i^{(n)} \to a_i$ 属于 $\mathcal{A}_i$，则

$$Q_i(\cdot \mid x_i, a_i^{(n)}, \sigma_j(x_j)) \Rightarrow Q_i(\cdot \mid x_i, a_i, \sigma_j(x_j))$$

弱收敛。由于 V_i 有界且连续，Portmanteau 定理给出

$$\int_{\mathcal{X}_i} V_i(x_i')\, Q_i(dx_i' \mid x_i, a_i^{(n)}, \sigma_j(x_j)) \to \int_{\mathcal{X}_i} V_i(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)).$$

因此 Φ_i 连续。

由于 $\mathcal{A}_i$ 根据假设A.1是紧致的，魏尔斯特拉斯极值定理表明 Φ_i 在 $\mathcal{A}_i$上取得最大值。因此，至少存在一个最大化行动。 □

A.3 贝尔曼算子与值函数的存在性

现在证明，参与者的贝尔曼算子在有界连续函数空间上是一个压缩映射。

令 $C_b(\mathcal{X}_i)$ 表示定义在 $\mathcal{X}_i$上的有界连续实值函数所构成的巴拿赫空间，并赋予上确界范数

$$\|V\|_\infty = \sup_{x_i \in \mathcal{X}_i} |V(x_i)|.$$

对于固定的对手策略 σ_j，定义贝尔曼算子 $T_i : C_b(\mathcal{X}_i) \to C_b(\mathcal{X}_i)$，如下

$$(T_i V)(x_i) = \sup_{a_i \in \mathcal{A}_i} \left\{ r_i\big(x_i, a_i, \sigma_j(x_j)\big) + \delta \int_{\mathcal{X}_i} V(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)) \right\}.$$

引理 A.10 (贝尔曼算子保持有界连续性)。 在假设 *A.1–A.5*下，若 $V \in C_b(\mathcal{X}_i)$ 则 $T_i V \in C_b(\mathcal{X}_i)$。

证明。 固定 $V \in C_b(\mathcal{X}_i)$。对于每个 (x_i, a_i)，定义

$$G(x_i, a_i) = r_i\big(x_i, a_i, \sigma_j(x_j)\big) + \delta \int_{\mathcal{X}_i} V(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)).$$

根据假设 A.3 和 A.4，函数 G 关于 (x_i, a_i) 联合连续。由于 $\mathcal{A}_i$ 是紧集，根据最大值定理可知

$$x_i \mapsto \sup_{a_i \in \mathcal{A}_i} G(x_i, a_i)$$

是连续的。由 r_i 和 V的有界性可知其有界。因此 $T_i V \in C_b(\mathcal{X}_i)$。 □

引理 A.11 (贝尔曼算子是压缩映射)。 在假设 *A.1–A.5*下，算子T_i 是 $C_b(\mathcal{X}_i)$ 上的压缩映射，其压缩系数为 δ。

证明。 令 $V, W \in C_b(\mathcal{X}_i)$，并固定 $x_i \in \mathcal{X}_i$。对于每个 $a_i \in \mathcal{A}_i$,

$$\left| \int V(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)) - \int W(x_i')\, Q_i(dx_i' \mid x_i, a_i, \sigma_j(x_j)) \right| \le \|V - W\|_\infty.$$

因此,

$$\left| r_i(\cdot) + \delta \int V\, dQ_i - \Big(r_i(\cdot) + \delta \int W\, dQ_i \Big) \right| \le \delta \|V - W\|_\infty.$$

对 a_i 取上确界可保持该界，从而得到

$$|(T_i V)(x_i) - (T_i W)(x_i)| \le \delta \|V - W\|_\infty.$$

现在对 $x_i \in \mathcal{X}_i$ 取上确界，可得

$$\|T_i V - T_i W\|_\infty \le \delta \|V - W\|_\infty.$$

由于 $0 < \delta < 1$, T_i 是压缩映射。 □

推论 A.12 (价值函数的存在性与唯一性)。 在假设 *A.1–A.5*下，对于固定的对手策略 σ_j，存在唯一的价值函数 $V_i^* \in C_b(\mathcal{X}_i)$ 满足

$$T_i V_i^* = V_i^*.$$

证明。 由引理A.10， T_i 将 $C_b(\mathcal{X}_i)$映射到自身。由引理 A.11， T_i 是完备度量空间 $C_b(\mathcal{X}_i)$上的压缩映射。因此，巴拿赫不动点定理给出唯一不动点。

□

A.4 一个严格的均衡存在性定理

我们现在陈述一个强化的均衡存在性结果。该定理是正文中所使用的均衡存在性论断的数学严格版本。

定理 A.13 （强化假设下可信风险马尔可夫完美贝叶斯均衡的存在性）。 假设*A.1–A.7* 成立。进一步假设：

（*i*）对每个参与者，延续价值由上述贝尔曼算子的不动点决定；

（*ii*）对每个对手策略，最优反应对应均为非空；（*iii*）最优反应对应在混合策略空间上的乘积弱拓扑中是上半连续的；（*iv*）最优反应对应是凸值的。

因此 至少存在一个混合策略可信风险马尔可夫完美贝叶斯均衡 .

证明。令

$$\Sigma_i$$

表示参与者 i从 $\mathcal{X}_i$ 到 $\Delta(\mathcal{A}_i)$ 的可测混合马尔可夫策略集合。根据假设A.1、A.2和A.6，由普罗霍罗夫定理，混合行动空间在弱拓扑下是紧的；在通常的可测选择框架下，相应的策略空间在乘积弱拓扑下也是紧的。

定义最佳响应对应

$$\mathcal{B}_i : \Sigma_j \rightrightarrows \Sigma_i$$

将每个对手策略映射为使折现可信风险值最大化的参与者i 策略集合。根据命题A.9 和推论A.12，每个 $\mathcal{B}_i(\sigma_j)$ 均非空。根据假设，每个 $\mathcal{B}_i$ 都是上半连续且凸值的。

现在定义联合最佳回应对应

$$\mathcal{B}(\sigma_1, \sigma_2) = \mathcal{B}_1(\sigma_2) \times \mathcal{B}_2(\sigma_1)$$

从 $\Sigma_1 \times \Sigma_2$ 映射到自身。乘积策略空间是紧致且凸的，并且 $\mathcal{B}$ 具有非空、凸且紧致的值，且是上半连续的。因此，卡库塔尼不动点定理适用，从而得到一个点

$$(\sigma_1^*, \sigma_2^*) \in \Sigma_1 \times \Sigma_2$$

使得

$$(\sigma_1^*, \sigma_2^*) \in \mathcal{B}(\sigma_1^*, \sigma_2^*).$$

根据构造，每位参与者的策略都是对其他参与者策略的最佳回应，且关于当前信念状态可测。因此，该策略组合构成混合策略可信风险马尔可夫完美贝叶斯均衡。

□

参考文献

Robert J. Aumann与Aviad Heifetz。不完全信息。载于Robert J. Aumann与Sergiu Hart编，《博弈论手册：经济学应用》，第3卷，第1665–1686页。Elsevier，2002。doi: 10.1016/S1574-0005(02)03006-0。Zeynep M¨uge Avsar与Melike Baykal-G¨ursoy。可替代需求下的库存控制：一项随机博弈应用。《海军研究后勤》，49(4):359–375，2002。doi: 10.1002/ nav.10018。Katy S. Azoury。未知需求分布下动态库存模型的贝叶斯解。《管理科学》，31(9):1150–1160，1985。doi: 10.1287/mnsc.31.9.1150。John C. Harsanyi。由“贝叶斯”参与者进行的不完全信息博弈，I：基本模型。《管理科学》，14(3):159–182，1967。doi: 10.1287/mnsc.14.3.159。John C. Harsanyi。由“贝叶斯”参与者进行的不完全信息博弈，III. 博弈的基本概率分布。《管理科学》，14(7):486–502，1968。doi: 10.1287/mnsc.14.7.486。J. S. Jordan。正规形博弈中的贝叶斯学习。《博弈与经济行为》，3(1): 60–81，1991。doi: 10.1016/0899-8256(91)90005-Y。J. S. Jordan。重复博弈中的贝叶斯学习。《博弈与经济行为》，9(1):8–20，1995。doi: 10.1006/game.1995.1002。K. Rajashree Kamath与T. P. M. Pakkala。未知需求分布下动态库存模型的贝叶斯方法。《计算机与运筹研究》，29(4): 403–422，2002。doi: 10.1016/S0305-0548(00)00075-7。Jean-Fran¸cois Mertens与Shmuel Zamir。不完全信息博弈贝叶斯分析的表述。《国际博弈论杂志》，14(1):1–29，1985。doi: 10. 1007/BF01770224。OpenML。Miceprotein。https://www.openml.org/d/40966，2026。2026年访问的OpenML数据集页面。Yasemin Serin。具有相同纳什解与斯塔克尔伯格解的竞争性报童问题。《运筹学快报》，35(1):83–94，2007。doi: 10.1016/j.orl.2006.01.002。UCI机器学习知识库。小鼠蛋白质表达。https://archive.ics.uci.edu/ml/datasets/Mice+Protein+Expression，2015。数据集描述页面。Jian Zhang、Juliang Zhang与Guowei Hua。带信息更新的多周期库存博弈。《国际生产经济学杂志》，174:119–127，2016。doi: 10.1016/j. ijpe.2016.01.017。

Juliang Zhang与陈健。未知需求分布下定价与库存控制的贝叶斯解。《运筹学快报》，34(5):517–524，2006。doi: 10.1016/j.orl. 2005.08.001。

周建恒和罗尧。贝叶斯信息更新与多时期供应链筛选。《国际生产经济学杂志》，256:108750，2023。doi: 10.1016/j.ijpe. 2022.108750。