arXiv:2301.09124v1 [math.ST] 22 Jan 2023

两阶段自适应设计下所选治疗均值的估计

Masihuddin[1*] 和 Neeraj Misra[2†]

[1,2] 印度理工学院坎普尔分校数学与统计学系，印度北方邦坎普尔 208016

摘要：自适应设计常用于临床和药物开发研究，以优化可用资源的利用。本文考虑采用两阶段自适应设计估计所选（较优）治疗效果的问题。设有两种治疗，其疗效分别服从均值不同且方差共同未知的两个正态分布。均值效应较大的治疗被称为较优治疗。在设计的第一阶段，将两种治疗分别独立施用于不同组的 n_1 名受试者，并选择样本均值较大的治疗作为较优治疗。在第二阶段，将所选治疗进一步施用于 n_2 名受试者。本文利用上述自适应设计，研究所选治疗均值的估计问题。我们扩展了 Cohen 和 Sackrowitz（1989）的结果，求得在第二阶段有多个观测值时，所选治疗均值效应的一致最小方差条件无偏估计量（UMVCUE）。我们证明，最大似然估计量（基于第一阶段和第二阶段数据计算的加权样本均值）在估计所选治疗的均值效应时是极小极大且可容许的。对于 σ^2 已知的情况，我们还提出了一些代入估计量：在 Masihuddin和Misra（2022）提出的部分估计量中，以合并样本方差代替共同方差 σ^2。我们通过模拟研究比较了所选治疗均值效应的各种估计量的表现。此外，我们还提供了一个真实数据应用示例。

AMS 2010学科分类：62F07 · 62F10 · 62C20

关键词与短语：两阶段自适应设计；所选治疗；缩放均方误差；缩放偏差；极小极大；UMVCUE；最大似然估计量；不可容许估计量。

*电子邮件：masih.iitk@gmail.com, masihst@iitk.ac.in

†电子邮件：neeraj@iitk.ac.in

1 引言

自适应设计在临床药物发现与开发中的重要性日益提升。在临床研究中，通常需要评估多种新治疗方案，但由于资源有限（时间、可用患者、预算等），只能选择观察到的响应最佳的一种或两种治疗方案，在大规模临床试验中进一步评估。两阶段自适应设计主要是在第一阶段从多个候选治疗中找出安全有效的治疗，然后在第二阶段使用独立样本验证其特性。

关于临床试验中自适应设计的全面深入综述，可参阅 Pallmann 等（2018）。关于两阶段自适应设计中推断程序的详细讨论，可参阅 Bauer 和 Kieser（1999）、Sampson 和 Sill（2005）、Stallard 和 Friede（2008）、Bowden 和 Glimm（2008）、Carreras 和 Brannath（2013）、Kimani 等（2013）、Chiu 等（2018）、Kimani 等（2020）以及 Robertson 等（2022）。

本文考虑一种两阶段自适应设计：第一阶段从候选治疗中选出较优者，第二阶段估计所选治疗的平均治疗效应。众所周知，在许多情况下，仅凭单阶段数据无法得到所选均值的无偏估计量。例如，正态总体和二项总体的所选均值均无法进行无偏估计（见 Putter 和 Rubinstein（1968）以及 Tappin（1992））。然而，纳入两个阶段数据的朴素估计可能会导致选择偏倚。为解决这一问题，可以采用 Rao-Blackwell 化技术。该技术将无偏的第二阶段样本均值对一个完全充分统计量进行条件化，从而得到一致最小方差条件无偏估计量（UMVCUE）。UMVCUE 的一个突出性质是，在所有所选均值的条件无偏估计量中，它的方差（换言之，均方误差（MSE））最小。

最初，Cohen 和 Sackrowitz（1989）在高斯情形下，基于排序与选择框架研究了所选治疗均值的两阶段估计。他们分别考虑了方差已知和未知的情形。作者得出了所选正态均值的 UMVCUE，该估计量使用两个阶段的数据。Cohen 和 Sackrowitz（1989）的研究有一个局限：假设自适应设计的第二阶段只有单次观测。此后，包括 Tappin（1992）以及 Sampson 和 Sill（2005）在内的许多研究者扩展了他们的工作。对于方差已知且为公共方差的情形，Bowden 和 Glimm（2008）将 UMVCUE 扩展到第一阶段和第二阶段样本量不相等的情况，其中感兴趣的参数是 j^{th} 个候选对象中最优的 k。Kimani 等（2013）研究了两阶段自适应无缝试验后对所选最有效治疗进行点估计，并将其与对照进行比较；在该试验中，第一阶段可以进行治疗选择，也可以因无效而提前停止。Bowden 和 Glimm（2014）利用两阶段淘汰劣者设计的多阶段对应设计，给出了所选均值的无偏和近似无偏估计。

在许多实际情形中，假定治疗效应的方差已知可能并不恰当。对于共同未知方差的情形，基于

Cohen 和 Sackrowitz（1989）的研究，我们推导了自适应设计第二阶段存在多个观测值时所选治疗均值的 UMVCUE。尽管 Robertson 和 Glimm（2019年）声称，他们的 UMVCUE 适用于第二阶段有多个观测值的情形，但由于他们对第二阶段数据进行了条件化，而所依据的统计量并非完全充分统计量，因此其方法并非最优。我们扩展后的 UMVCUE 弥补了这一不足。我们还在缩放均方误差准则下，得到了所选治疗均值的极小极大估计量，并证明该极小极大估计量也是可容许的。

本文其余部分安排如下：第2节介绍全文将用到的一些符号和预备知识。第2.1小节推导所选治疗均值的 UMVCUE。第3节证明，在共同未知方差的情况下，用于估计所选治疗均值的朴素估计量（即第一阶段和第二阶段样本均值的加权平均）是极小极大且可容许的。第3.1小节介绍一些用于估计所选治疗均值的其他估计量。为了依据缩放均方误差和缩放偏差准则，对各种竞争估计量的表现进行数值评估，我们在第4节开展了一项模拟研究。此外，本文第5节还考虑了一个真实数据集作为示例。

2 所选治疗均值的估计

下文将使用以下符号：

- $\mathbb{R}$：实数轴（$-\infty, \infty$）；
- $\mathbb{R}^k$： k 维欧几里得空间， $k \in \{2, 3, \ldots\}$；
- N（θ, σ^2）：均值为 $\theta \in \mathbb{R}$、标准差为 $\sigma \in$ 的正态分布（$0, \infty$）；
- ϕ（$\cdot$）： N（0, 1）的概率密度函数（p.d.f.）；
- Φ（$\cdot$）： N（0, 1）的累积分布函数（c.d.f.）；
- $Beta$（a, b）：形状参数为 $a > 0$ 和 $b > 0$的贝塔分布；
- $B(\alpha, \beta) = \frac{\Gamma}{\Gamma(\alpha)\Gamma(\beta)}\Gamma(\alpha+\beta)$， $\alpha > 0$、 $\beta > 0$表示通常的贝塔函数， $\Gamma(\cdot)$表示通常的伽马函数；
- 对于实数 x 和 y

$$I(x \geq y) = \begin{cases} 1, & \text{if } x \geq y \\ 0, & \text{if } x < y. \end{cases}$$

$I(x > y)$ 的定义类似。

考虑两种治疗，分别记为 τ_1 和 τ_2，其中治疗 τ_i 的疗效由$N(\mu_i, \sigma^2)$ $i = 1, 2$ 描述； $\mu_1 \in \mathbb{R}$ 和 $\mu_2 \in \mathbb{R}$ 是未知的平均治疗效应， σ^2 (σ >0) 是共同的未知治疗效应方差。我们将 max $\{\mu_1, \mu_2\}$ 对应的治疗定义为较优或有希望的治疗。我们考虑一种包含两个阶段的自适应设计，并在中期进行一次数据驱动选择。在设计的第一阶段（即阶段 1），将治疗 τ_1 施用于 n_1 名受试者，并将治疗τ_2 独立施用于另一组 n_1 名受试者。令 $\overline{X}_i$, $i = 1, 2$ 分别表示对应两种治疗的样本均值（平均治疗效应估计值）。为了选择较优治疗，我们考虑一种自然的选择规则：选择样本均值较大的治疗作为较优治疗（关于这一自然选择规则的最优性，参见 Bahadur 和 Goodman（1952）、Eaton（1967）以及 Misra 和 Dhariyal（1994））。令 $Q \in \{1, 2\}$ 表示所选治疗 τ_Q 的指标（即，若 $\overline{X}_1 > \overline{X}_2$，则为 $Q = 1$；若 $\overline{X}_2 \geq X_1$，则为 $Q = 2$）。随后将治疗 τ_Q 带入第二阶段（即两阶段设计的阶段 2）进行进一步分析。在阶段 2，将所选治疗 τ_Q 独立施用于另外 n_2 名受试者。令 $\overline{Y}$ 表示所选治疗 τ_Q在阶段 2 的样本均值。我们的目标是估计 μ_Q，即所选治疗的均值效应。注意，$\mu_Q \equiv \mu_Q\left(\mu_1, \mu_2, \overline{X}_1, \overline{X}_2\right)$ 是一个随机参数，它取决于 μ_1, μ_2, $\overline{X}_1$ 和$\overline{X}_2$；当 $\overline{X}_1 > \overline{X}_2$时等于 μ_1，当 $\overline{X}_2 \geq \overline{X}_1$时等于 μ_2。显然， $\overline{X}_i \sim N\left(\mu_i, \frac{\sigma^2}{n_1}\right)$，$i = 1, 2$ 相互独立分布，且在给定 Q的条件下， $\overline{Y} \sim N\left(\mu_Q, \frac{\sigma^2}{n_2}\right)$。

本文还将通篇使用以下符号：
$\underset{\sim}{T} = (\overline{X}_1, \overline{X}_2, \overline{Y}, S^2)$; $S^2 = \sum_{i=1}^{2}\sum_{j=1}^{n_1} X_{ij}^2 + \sum_{j=1}^{n_2} Y_j^2$; $\underline{\theta} = (\mu_1, \mu_2, \sigma)$; $\Theta = \mathbb{R}^2 \times (0, \infty)$; $\overline{X}_Q = \max(\overline{X}_1, \overline{X}_2)$（ $\overline{X}_1$ 和 $\overline{X}_2$中的最大值）； $\overline{X}_{3-Q} = \min(\overline{X}_1, \overline{X}_2)$（ $\overline{X}_1$ 和 $\overline{X}_2$中的最小值）。此外，对于任意$\underline{\theta} \in \Theta$， $\mathbb{P}_{\underline{\theta}}(\cdot)$表示由 $\underset{\sim}{T} = (\overline{X}_1, \overline{X}_2, \overline{Y}, S^2)$诱导的概率测度，其中$\underline{\theta} \in \Theta$ 为真实参数值； $\mathbb{E}_{\underline{\theta}}(\cdot)$表示概率测度 $\mathbb{P}_{\theta}(\cdot)$下的期望算子， $\underline{\theta} \in \Theta$。

本文重点关注以下定义的所选治疗均值效应的点估计：

$$\begin{aligned}\mu_Q \equiv \mu_Q\left(\mu_1, \mu_2, \overline{X}_1, \overline{X}_2\right) &= \begin{cases}\mu_1, & \text{if } \overline{X}_1 \geq \overline{X}_2 \\ \mu_2, & \text{if } \overline{X}_1 < \overline{X}_2\end{cases}, \\ &= \mu_1 I\left(\overline{X}_1 \geq \overline{X}_2\right) + \mu_2 I\left(\overline{X}_2 > \overline{X}_1\right), \end{aligned} \tag{2.1}$$

在缩放平方误差损失函数下

$$L_{\underset{\sim}{T}}(\underline{\theta}, a) = \left(\frac{a - \mu_Q}{\sigma}\right)^2, \quad \underline{\theta} \in \Theta, \ a \in \mathcal{A} = \mathbb{R}. \tag{2.2}$$

值得一提的是，统计量 $\left(\frac{n_1\overline{X}_Q + n_2\overline{Y}}{n_1+n_2}, \overline{X}_{3-Q}, S^2, Q\right)$ 是最小充分的，但并不完备。然而，给定 Q 后，统计量 $\left(\frac{n_1\overline{X}_Q + n_2\overline{Y}}{n_1+n_2}, \overline{X}_{3-Q}, S^2\right)$ 是一个完全充分统计量。因此， $\underset{\sim}{T} = (\overline{X}_1, \overline{X}_2, \overline{Y}, S^2)$ 是充分统计量，而 μ_Q 仅通过观

测值依赖于 $\underset{\sim}{T}$。因此，我们可以将注意力仅限于那些只通过 $\underset{\sim}{T}$ 依赖于观测值的估计量（见 Misra 和 Singh（1993））。

在缩放平方误差损失函数 (2.2) 下，估计量 $d(\underset{\sim}{T})$ 的风险函数（亦称缩放均方误差）定义为

$$R(\underline{\theta}, d) = \mathbb{E}_{\underline{\theta}} \left(\frac{d(\underset{\sim}{T}) - \mu_Q}{\sigma} \right)^2, \ \underline{\theta} \in \Theta.$$

假设我们有定义在 Θ 上的先验分布（密度） Π 。那么，估计量d相对于先验分布 Π 的贝叶斯风险定义为

$$\begin{aligned} r(\Pi, d) &= \mathbb{E}_{\Pi} \left(R(\underline{\theta}, d) \right) \\ &= \int_{\Theta} \int_{\zeta} L_T(\underline{\theta}, d(\underset{\sim}{t})) f(\underset{\sim}{t}|\underline{\theta}) \Pi(\underline{\theta}) d\underline{t} \ d\underline{\theta}, \end{aligned}$$

其中， ζ 表示 $\underset{\sim}{T}$的支撑集。

在所有针对 μ_Q的估计量 d 中，使贝叶斯风险 $r(\Pi, d)$ 最小的估计量 d_Π ，称为相对于先验分布 Π的贝叶斯估计量。

若估计量 $d(\underset{\sim}{T})$ 满足以下条件，则称其为条件无偏:

$$\mathbb{E}_{\underline{\theta}}((d(\underset{\sim}{T})|Q = q)) = \mu_q, \ \forall \ \underline{\theta} \in \Theta.$$

用于估计所选治疗 μ_Q 的均值效应的朴素估计量，是两个阶段样本均值的加权平均，即

$$d_M(\underset{\sim}{T}) = \frac{n_1 \overline{X}_Q + n_2 \overline{Y}}{n_1 + n_2}. \tag{2.3}$$

显然， $d_M(\underset{\sim}{T})$ 是 μ_Q 的最大似然估计量（MLE）。

在下一小节中，将在共同方差未知的假设下推导所选治疗均值 μ_Q 的 UMVCUE。

2.1 扩展的 UMVCUE

定理 2.1. 给定 Q 时， μ_Q 的两阶段一致最小方差条件无偏估计量为

$$d_U(\underset{\sim}{T}) = \frac{Z}{n_1 + n_2} - \sqrt{\frac{n_1}{n_2(n_1 + n_2)}} \widetilde{S} \frac{(1 - V^{*2})^c}{2^{2c} c B(c, c) I_{c,c}\left(\frac{V^*+1}{2}\right)}, \tag{2.4}$$

where, $Z = n_1 \overline{X}_Q + n_2 \overline{Y}$, $V = \frac{\sqrt{\frac{n_1(n_1+n_2)}{n_2}}}{\widetilde{S}} \left(\frac{Z}{n_1+n_2} - \overline{X}_{3-Q} \right)$, $V^* = \min\{V, 1\}$,

$\widetilde{S}^2 = \sum_{i=1}^{2} \sum_{j=1}^{n_1} X_{ij}^2 + \sum_{j=1}^{n_2} Y_j^2 - (n_1 + n_2) \left(\frac{Z}{n_1+n_2} \right)^2 - n_1 \overline{X}_{3-Q}^2$, $c = \frac{2n_1+n_2-3}{2}$,

$I_{c\,c}\ u$ $Beta\ c\ c$ $B\ c\ c$ $\frac{}{\Gamma\ 2c}$ $c >$, ( ) 是 (,) 分布 0 的累积分布函数，而 (,) $= {}^{\Gamma}{}_{(c)\Gamma(c)}\ ()$ 是通常的贝塔函数。

证明。令 $\underline{S}=\left(\underline{X_1},\underline{X_2},\underline{Y}\right)$ 表示合并后的第一阶段和第二阶段样本观测值向量。于是，基于第一阶段和第二阶段数据，给定 $Q=q$ 时 $\underline{S}$ 的联合概率密度函数（概率密度函数）可写为

$$f_q(\underline{x_1},\underline{x_2},\underline{y}|\underline{\theta})\propto\exp\left(\frac{1}{\sigma^2}\left[\sum_{j=1}^{n_1}x_{qj}+\sum_{j=1}^{n_2}y_j\right]\mu_q+\frac{\sum_{j=1}^{n_1}x_{(3-q)j}}{\sigma^2}\mu_{3-q}-\frac{1}{2\sigma^2}\left[\sum_{j=1}^{n_1}x_{qj}^2+\sum_{j=1}^{n_1}x_{(3-q)j}^2+\sum_{j=1}^{n_2}y_j^2\right]\right).$$

由联合密度 $f_q(\underline{x_1},\underline{x_2},\underline{y}|\underline{\theta})$ 的表达式可知，给定 Q 后，统计量$\left(Z,\overline{X}_{3-Q},\widetilde{S}^2\right)$ 是完全充分统计量。由于给定 Q 时， $\overline{Y}$是 μ_Q的无偏估计量，因此， μ_Q 的 UMVCUE 是在完全充分统计量 $\left(Z,\overline{X}_{3-Q},\widetilde{S}^2\right)$ 条件下对 $\overline{Y}$ 进行 Rao-Blackwell 化的结果。因此，所求的所选均值 μ_Q 的 UMVCUE 为 $\mathbb{E}_{\underline{\theta}}\left(\overline{Y}\middle|\left(Z,\overline{X}_{3-Q},\widetilde{S}^2,Q\right)\right)$。定义

，

$$U=\frac{\sqrt{\frac{n_2(n_1+n_2)}{n_1}}\left(\overline{Y}-\frac{Z}{n_1+n_2}\right)}{\widetilde{S}}.$$

为了证明该 UMVCUE 与（2.4）相同，需要证明

$$\mathbb{E}_{\underline{\theta}}\left(U\middle|\left(Z,\overline{X}_{3-Q},\widetilde{S}^2\right)\right)=-\frac{(1-V^{*2})^c}{2^{2c}cB(c,c)I_{c,c}\left(\frac{V^*+1}{2}\right)}.\tag{2.5}$$

2.5），我们需要求出 U 在给定 $\left(Z,\overline{X}_{3-Q},\widetilde{S}^2,Q\right)$条件下的条件概率密度函为证明（ 数 .

~注意，对 S^2的表达式中的各项重新排列后，可将 U^2 的表达式写为

$$U^2=\frac{\frac{n_2(n_1+n_2)}{n_1}\left(\overline{Y}-\frac{Z}{n_1+n_2}\right)^2}{\sum_{i=1}^{2}\sum_{j=1}^{n_1}(X_{ij}-\overline{X}_i)^2+\sum_{j=1}^{n_2}(Y_j-\overline{Y})^2+\frac{n_2(n_1+n_2)}{n_1}\left(\overline{Y}-\frac{Z}{n_1+n_2}\right)^2}.\tag{2.6}$$

与 Cohen 和 Sackrowitz（1989）以及 Robertson和Glimm（2019）中的做法相同，可以验证，在给定 $\left(Z,\overline{X}_{3-Q},\widetilde{S}^2,Q\right)=$（$z$、 x_{3-q}、 $\widetilde{s}$、 q）的条件下， U 的条件密度为

$$f_{U|\left(Z,\overline{X}_{3-Q},\widetilde{S}^2,Q\right)}(u|(z,x_{3-q},\widetilde{s},q))=\frac{(1-u^2)^{c-1}}{\int_{-1}^{v^*}(1-u^2)^{c-1}du},\quad -1<u<v^*,$$

其中，对于 $v=\frac{\sqrt{n_1}}{\sqrt{(n_1+n_2)}\widetilde{s}^{n_2}}\left(\frac{z}{n_1+n_2}-x_{3-Q}\right)$, $v^*=\min\{v,1\}$。

因此，我们有

$$\mathbb{E}_{\underline{\theta}}\left(U\middle|\left(Z,\overline{X}_{3-Q},\widetilde{S}^2,Q\right)\right)=\frac{\int_{-1}^{V^*}u(1-u^2)^{c-1}\,du}{\int_{-1}^{V^*}(1-u^2)^{c-1}du}$$
$$=-\frac{(1-V^{*2})^c}{2^{2c}cB(c,c)I_{c,c}\left(\frac{V^*+1}{2}\right)}.$$

因此，结论得证。 ■

现在，我们将证明，由（2.3）定义的朴素估计量 $d_M(\underline{T})$，在缩放均方误差准则下，对于估计 μ_Q 是极小极大且可容许的。

3 极小极大估计量与可容许估计量

$\sum_{i=1}^{2}\sum_{j=1}^{n_1}(X_{ij}-\overline{X}_i)^2+\sum_{j=1}^{n_2}(Y_j-\overline{Y})^2$令 $S^{*2}=\frac{}{2n_1+n_2-3}$ 为第一阶段和第二阶段数据的合并样本方差，~~并令 $W=(2n_1+n_2-3)S^{*2}$~~，使得 $\frac{W}{\sigma^2}\sim\chi^2_{2n_1+n_2-3}$。假设未知参数向量 $\underline{\theta}=(\mu_1,\mu_2,\sigma^2)\in\Theta$ 是随机向量 $\underline{P}=(P_1,P_2,P_3)$ 的一个实现，该随机向量具有指定的概率分布函数。考虑一列先验分布（密度）$\{\xi_m\}_{m\geq1}$，其对应于 $\underline{P}=(P_1,P_2,P_3)$，满足：

(i) 对于任意固定的 $p_3>0$，给定 $P_3=p_3$ 后，条件地，P_1 和 P_2 相互独立且同分布为 $N(0,mp_3)$；

(ii) 随机变量 P_3 服从逆指数分布，其概率密度函数为

$$\Pi_{1,P_3}(p_3)=\frac{1}{p_3^2}e^{-\frac{1}{p_3}},\ p_3>0.$$

回顾一下，$\overline{X}_i\ (i=1,2)$ 是来自 i^{th} 总体的第一阶段样本的样本均值，而 $\overline{Y}$ 是从第一阶段选定的总体中抽取的样本的第二阶段样本均值。于是，在先验分布 ξ_m 下，给定 $(\overline{X}_1,\overline{X}_2,\overline{Y},W)$ 时，(P_1,P_2,P_3) 的联合后验分布 $=(x_1,x_2,y,w)$ 满足：

(i) 对任意 $p_3>0$，在给定条件下，随机变量 P_1 和 P_2 分别相互独立分布为 $N\left(\theta_{1,\ m},\tau^2_{1,\ m}\right)$ 和 $N\left(\theta_{2,\ m},\tau^2_{2,\ m}\right)$，其中

$$\left(\theta_{1,m},\theta_{2,m},\tau^2_{1,m},\tau^2_{2,m}\right)=\begin{cases}\left(\frac{n_1x_1+n_2y}{n_1+n_2+\frac{1}{m}},\frac{n_1x_2}{n_1+\frac{1}{m}},\frac{1}{\frac{n_1+n_2}{p_3}+\frac{1}{mp_3}},\frac{1}{\frac{n_1}{p_3}+\frac{1}{mp_3}}\right), & \text{if } x_1\geq x_2\\ \left(\frac{n_1x_1}{n_1+\frac{1}{m}},\frac{n_1x_2+n_2y}{n_1+n_2+\frac{1}{m}},\frac{1}{\frac{n_1}{p_3}+\frac{1}{mp_3}},\frac{1}{\frac{n_1+n_2}{p_3}+\frac{1}{mp_3}}\right), & \text{if } x_1<x_2\end{cases}.$$

(ii) 随机变量 P_3 服从逆伽马分布，其概率密度函数为

$$\Pi_{2,m}(p_3)=\frac{(v_m)^{\frac{n+7}{2}}}{\Gamma(n+7)}\frac{e^{-\frac{v_m}{p_3}}}{p_3^{\frac{n+9}{2}}}$$

其中， $n=2n_1+n_2-3$ 和

$$v_m=\begin{cases}1+\frac{w}{2}+\frac{n_1x_1^2}{2}+\frac{n_1x_2^2}{2}+\frac{n_2y^2}{2}-\frac{(n_1x_1+n_2y)^2}{2(n_1+n_2+\frac{1}{m})}-\frac{(n_1x_2)^2}{2(n_1+\frac{1}{m})}, & \text{if } x_1\geq x_2\\ 1+\frac{w}{2}+\frac{n_1x_1^2}{2}+\frac{n_1x_2^2}{2}+\frac{n_2y^2}{2}-\frac{(n_1x_2+n_2y)^2}{2(n_1+n_2+\frac{1}{m})}-\frac{(n_1x_1)^2}{2(n_1+\frac{1}{m})}, & \text{if } x_1<x_2\end{cases},$$

$m=1$，2，.....

因此，在缩放平方误差损失函数（2.2）下，关于先验分布 ξ_m，所选治疗均值 μ_Q的贝叶斯估计量为

$$\begin{aligned}d_{\xi_m}\left(\underset{\sim}{T}\right)&=\begin{cases}\frac{\mathbb{E}^{\underset{\sim}{P}|T}\left(\frac{P_1}{P_3}\right)}{\mathbb{E}^{\underset{\sim}{P}|T}\left(\frac{1}{P_3}\right)}, & \text{if } \overline{X}_1\geq\overline{X}_2\\ \frac{\mathbb{E}^{\underset{\sim}{P}|T}\left(\frac{P_2}{P_3}\right)}{\mathbb{E}^{\underset{\sim}{P}|T}\left(\frac{1}{P_3}\right)}, & \text{if } \overline{X}_1<\overline{X}_2\end{cases}\\&=\begin{cases}\frac{n_1\overline{X}_1+n_2\overline{Y}}{n_1+n_2+\frac{1}{m}}, & \text{if } \overline{X}_1\geq\overline{X}_2\\ \frac{n_1\overline{X}_2+n_2\overline{Y}}{n_1+n_2+\frac{1}{m}}, & \text{if } \overline{X}_1<\overline{X}_2\end{cases}\\&=\frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2+\frac{1}{m}},\ m=1,2,\ldots.\end{aligned}\tag{3.1}$$

贝叶斯规则 $d_{\xi_m}\left(\underset{\sim}{T}\right)$ 的后验风险为

$$\begin{aligned}r_{d_{\xi_m}}\left(\underset{\sim}{T}\right)&=\mathbb{E}^{\underset{\sim}{P}|T}\left[\frac{\left(\frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2+\frac{1}{m}}-\mu_Q\right)^2}{P_3}\right]\\&=\frac{1}{n_1+n_2+\frac{1}{m}},\end{aligned}$$

它与 $\underset{\sim}{T}=(\overline{X}_1,\overline{X}_2,\overline{Y},W)$ 相互独立。因此，估计量 d_{ξ_m} 的贝叶斯风险为

$$r^*_{d_{\xi_m}}(\xi_m)=\frac{1}{n_1+n_2+\frac{1}{m}},\ m=1,2,\ldots\tag{3.2}$$

应用 Masihuddin和Misra（2022年）的引理 2.2（ii），我们得到估计量$d_M(\underset{\sim}{T}) = \frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2}$的风险为

$$R(\underline{\theta}, d_M) = \mathbb{E}_{\underline{\theta}} \left(\frac{\frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2} - \mu_Q}{\sigma} \right)^2$$
$$= \frac{1}{n_1+n_2}. \tag{3.3}$$

The因此，在先验分布 ξ_m 下，朴素估计量 $d_M(\underset{\sim}{T})$ 的贝叶斯风险为

$$r^*_{d_M}(\xi_m) = \frac{1}{n_1+n_2},\ m = 1, 2, \ldots \tag{3.4}$$

现在，我们给出以下定理，以证明自然估计量d_M的极小极大性。

定理 3.1。在缩放平方误差损失函数 (2.2) 下，自然估计量 $d_M(\underset{\sim}{T}) = \frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2}$是估计所选治疗均值 μ_Q 的极小极大估计量。

证明。设 d 是任意其他估计量。由于由 (3.1) 给出的 d_{ξ_m} 是在先验分布 $\xi_m, m = 1, 2, \ldots$ 下对 μ_Q 的贝叶斯估计量，我们有：

$$\begin{aligned}
\sup_{\underline{\theta}\in\Theta} R(\underline{\theta}, d) &\geq \int_\Theta R(\underline{\theta}, d)\xi_m(\underline{\theta})d\underline{\theta} \\
&\geq \int_\Theta R(\underline{\theta}, d_{\xi_m})\xi_m(\underline{\theta})d\underline{\theta} \\
&= r^*_{d_{\xi_m}}(\xi_m) = \frac{1}{n_1+n_2+\frac{1}{m}}, \quad m = 1, 2, \ldots \quad (\text{using } (3.2)) \\
\Rightarrow \sup_{\underline{\mu}\in\Theta} R(\underline{\mu}, d) &\geq \lim_{m\to\infty} r^*_{d_{\xi_m}}(\xi_m) = \frac{1}{n_1+n_2} = \sup_{\underline{\theta}\in\Theta} R(\underline{\theta}, d_M), \quad (\text{using } (3.4))
\end{aligned}$$

这意味着 $d_M(\underset{\sim}{T}) = \frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2}$ 是估计 μ_Q 的极小极大估计量。 ■

现在我们将运用不变性原理。在缩放平方误差损失函数（2.2）下，对所选治疗均值 μ_Q 进行估计的问题，在仿射变换群以及置换群下都是不变的。

容易验证，任何仿射且置换等变的 μ_Q 估计量都具有以下形式

$$d_\psi(\underset{\sim}{T}) = \frac{n_1\overline{X}_Q + n_2\overline{Y}}{n_1+n_2} - \widetilde{S}\psi\left(\frac{D}{\widetilde{S}}\right), \tag{3.5}$$

对于某个函数 $\psi:\ \mathbb{R} \to \mathbb{R}$，其中 $D = \frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2} - \overline{X}_{3-Q}$。将所有形式为 (3.5) 的仿射且置换等变估计量构成的类记为 $\mathscr{E}_1$。显然，最大似然估计量 $d_M(\underset{\sim}{T})$ 和 UMVCUE d_U 分别由 (2.3) 和 (2.4) 定义，且均属于类 $\mathscr{E}_1$.

接下来，我们将证明最大似然估计量 d_M 在仿射且置换等变估计量类 $\mathscr{E}_1$ 中是可容许的。注意，任意估计量 $d \in \mathscr{E}_1$ 的风险函数都通过 $\mu = \frac{\theta_2-\theta_1}{\sigma}$依赖于 $\underline{\theta} \in \Theta$，其中 $\theta_2 = \max\{\mu_1, \mu_2\}$，且 $\theta_1 = \min\{\mu_1, \mu_2\}$。因此，任意估计量 $d \in \mathscr{E}_1$ 的贝叶斯风险都通过 $\frac{P_{[2]}-P_{[1]}}{P_3}$的分布，依赖于 $\underline{P} = (P_1, P_2, P_3)$ 的先验分布，其中 $P_{[1]} = \min\{P_1, P_2\}$， $P_{[2]} = \max\{P_1, P_2\}$.

下述定理证明，在仿射且置换等变估计量类 $\mathscr{E}_1$中，估计量 d_M是可容许的。由于该定理的证明与Masihuddin和Misra（2021）定理5.2的证明思路相同，此处省略。

定理3.2. 自然估计量 $d_M(\underset{\sim}{T}) = \frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2}$ 在缩放均方误差准则下，在类 $\mathscr{E}_1$中可用于估计所选治疗均值 μ_Q，且是可容许的。

4 一些附加 的朴素估计量和模拟研究

为了估计所选治疗均值，可以将未知方差 σ^2 的估计值代入Masihuddin和Misra（2022）针对 σ^2 已知情形推导出的估计量（δ_0^{RB}和 δ_1）的表达式中，从而得到一些额外的朴素估计量。在方差未知的情形下，可得到两个这样的朴素估计量:

$$d_{U1}(\underset{\sim}{T}) = d_0^{RB,\widehat{\sigma}}(\underset{\sim}{T}) = Z_1 + \frac{\sqrt{n_2}S^*}{\sqrt{n_1(n_1+n_2)}} \frac{\phi\left(\frac{\sqrt{n_1(n_1+n_2)}}{\sqrt{n_2}S^*}(Z_1-Z_2)\right)}{\Phi\left(\frac{\sqrt{n_1(n_1+n_2)}}{\sqrt{n_2}S^*}(Z_1-Z_2)\right)} \tag{4.1}$$

and

$$d_{U2}(\underset{\sim}{T}) = d_1^{\widehat{\sigma}}(\underset{\sim}{T}) = \begin{cases} \frac{(n_1+n_2)Z_1+n_1Z_2}{2n_1+n_2}\left\{\frac{\Phi\left(\frac{Z_1-Z_2}{\widehat{\sigma_1}}\right)-\Phi\left(\frac{n_1(Z_1-Z_2)}{(2n_1+n_2)\widehat{\sigma_1}}\right)}{\Phi\left(\frac{Z_1-Z_2}{\widehat{\sigma_1}}\right)}\right\} \\ +\frac{\widehat{\sigma_1}\phi\left(\frac{n_1(Z_1-Z_2)}{(2n_1+n_2)\widehat{\sigma_1}}\right)+Z_1\Phi\left(\frac{n_1(Z_1-Z_2)}{(2n_1+n_2)\widehat{\sigma_1}}\right)}{\Phi\left(\frac{Z_1-Z_2}{\widehat{\sigma_1}}\right)}, & \text{if } Z_1 > Z_2\,, \\ \frac{(n_1+n_2)Z_1+n_1Z_2}{2n_1+n_2}, & \text{if } Z_1 \le Z_2 \end{cases} \tag{4.2}$$

其中， $\widehat{\sigma_1} = S^*\sqrt{\frac{n_2}{n_1(n_1+n_2)}}$， $Z_1 = \frac{n_1\overline{X}_Q+n_2\overline{Y}}{n_1+n_2}$、 $Z_2 = \overline{X}_{3-Q}$ 和 S^{*2} 是基于第一阶段和第二阶段数据计算的合并样本方差。

现在，我们报告一项模拟研究，按照缩放均方误差和偏差准则，对估计量d_{U_1}、 d_{U_2}、最大似然估计量 d_M 和UMVCUE d_U 进行性能比较。

我们通过蒙特卡洛模拟比较了所选治疗均值 (μ_Q) 的各种估计量在风险（缩放均方误差）和偏差方面的表现。数值研究中考察了以下估计量： d_M、 d_U、 d_{U_1}和 d_{U_2}（见 (2.3)、(2.4)、(4.1) 和 (4.2)）。由于这些 μ_Q 估计量的缩放均方误差和偏差取决于参数 $\underline{\theta} =$ （μ_1、 μ_2、 σ） $\in \Theta$

通过 μ （≥ 0），我们绘制了不同样本量配置 n_1 和 n_2下，各种估计量的模拟风险和缩放偏差随 $\mu \geq 0$ 变化的曲线。基于 100,000 次模拟得到的缩放均方误差和偏差模拟值绘于图 4.1–4.7。在图 4.1–4.2 中，我们绘制了所提出估计量的模拟缩放均方误差随 μ变化的曲线。在图 4.3–4.5 中，对于固定的总样本量 $n = n_1 + n_2$ 和 μ，我们绘制了缩放均方误差模拟值随信息分数（$= \frac{n_1}{n_1+n_2}$）变化的曲线。各种估计量的缩放偏差绘于图 4.6–4.7。

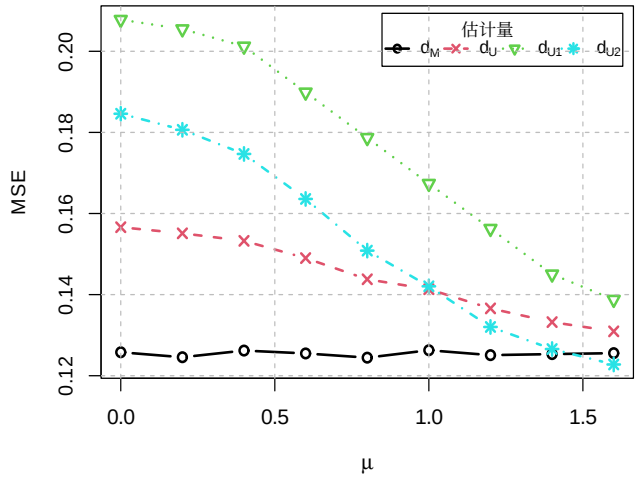
根据模拟研究得出以下结论：

(i) 与 (3.3) 一致，自然估计量 d_M 的缩放均方误差对于标准化治疗效应差 μ 的所有取值均为常数（$=\frac{1}{n_1+n_2}$）。(ii) 除 d_M 外，所有其他估计量（d_U、 d_{U1} 和 d_{U2}）的缩放均方误差均随着 μ 的增大而减小。当 μ 较大时，与其他估计量相比，估计量 d_{U2} 的缩放均方误差表现更好。(iii) 对于 $n_1 = n_2$，UMVCUE d_U 的缩放均方误差表现与估计量 d_{U1} 相似。对于 $n_1 \geq n_2$，估计量 d_{U2} 优于估计量 d_U 和 d_{U1}。还观察到，就缩放均方误差而言，估计量 d_{U2} 始终优于 d_{U1}。(iv) 当 n_2 较小且 n_1 相对较大时，估计量 d_{U1} 和 d_{U2} 的缩放均方误差与 d_M 的非常接近。当 n_1 较小而 n_2 较大时，UMVCUE d_U 的缩放均方误差更接近 d_M 的缩放均方误差。(v) 当 μ 和总样本量 $n = n_1 + n_2$ 固定且取较小值时，UMVCUE d_U 的缩放均方误差随着信息分数 $\frac{n_1}{n}$ 的增大而增大，而 d_{U1} 和 d_{U2} 的缩放均方误差则随着信息分数的增大而减小。当 μ 较大时，估计量 d_U 和 d_M 的缩放均方误差表现相似。(vi) 除 UMVCUE d_U 外，所有竞争估计量（d_M、 d_{U1} 和 d_{U2}）的缩放偏差均随着 μ 的增大而减小。当 μ 较大时，这些估计量的缩放偏差表现与 d_U 相当。另一个有趣的现象是，μ_Q 的所有竞争估计量在 μ 为零时达到最大偏差，并且最大偏差随着 μ 的增大而减小。(vii) 当样本量 n_1 和 n_2 较大时，所有竞争估计量的缩放均方误差和缩放偏差均趋近于零。

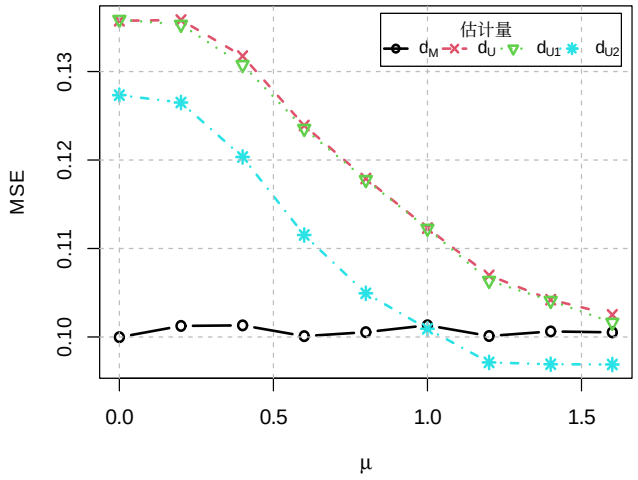
当缩放均方误差是选择合适估计量的关键标准时，我们推荐使用估计量d_M 和 d_{U2}。对于 n_1 和 n_2 的某些特定配置（ n_1 较小且 n_2较大），UMVCUEd_U 也是一个不错的竞争者。

当需要同时控制缩放偏差和缩放均方误差时，我们推荐使用估计量d_U 和 d_M。

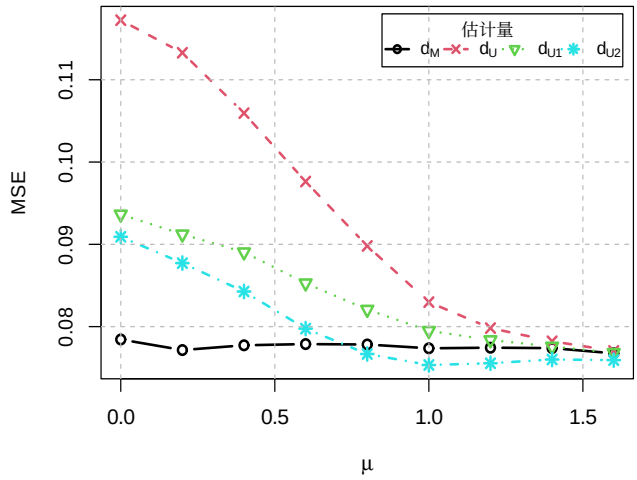
所选治疗均值的估计



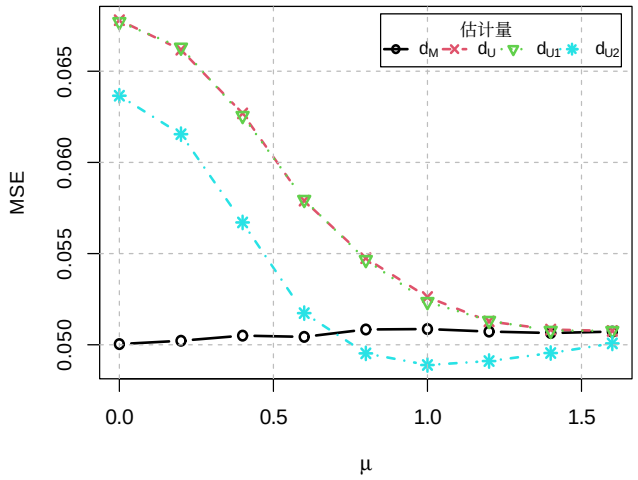
(a) $n_1 = 3, n_2 = 5$



(b) $n_1 = 5, n_2 = 5$



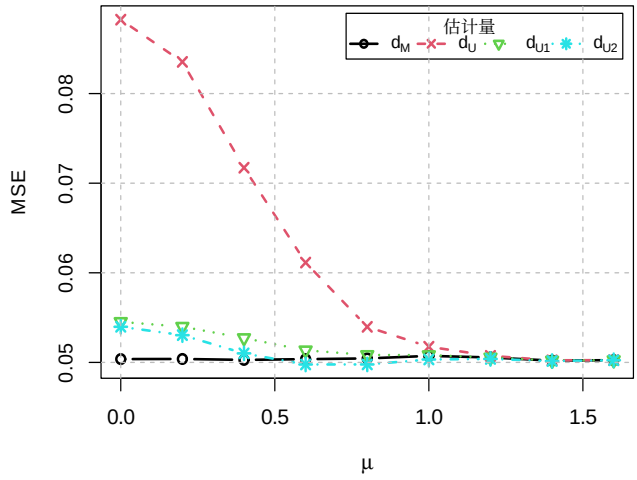
(c) $n_1 = 8, n_2 = 5$



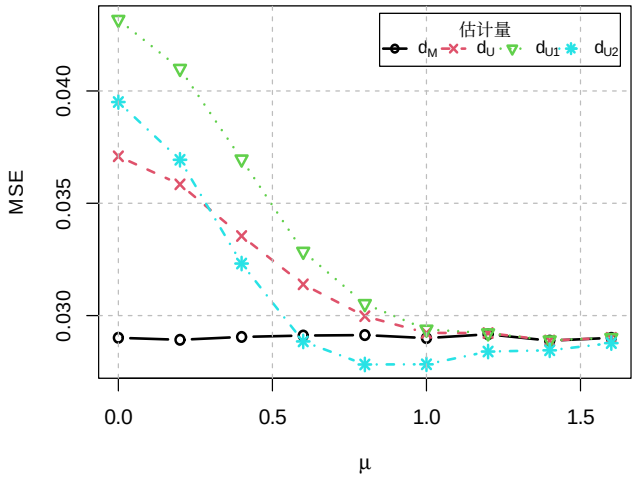
(d) $n_1 = 10, n_2 = 10$

图 4.1：在 n_1 和 n_2的不同配置下，各种竞争估计量的缩放均方误差图

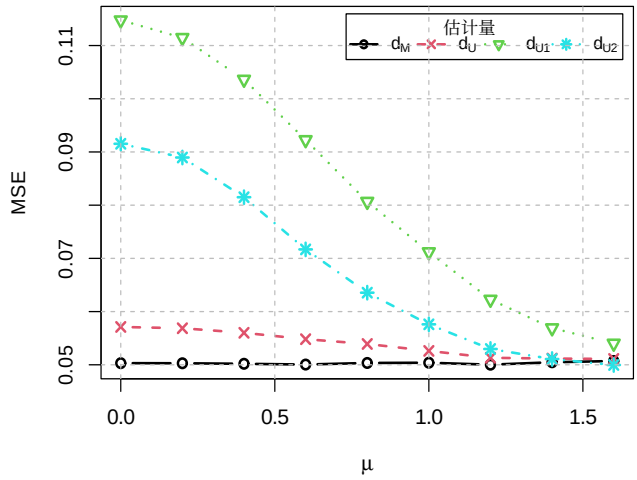
所选治疗均值的估计



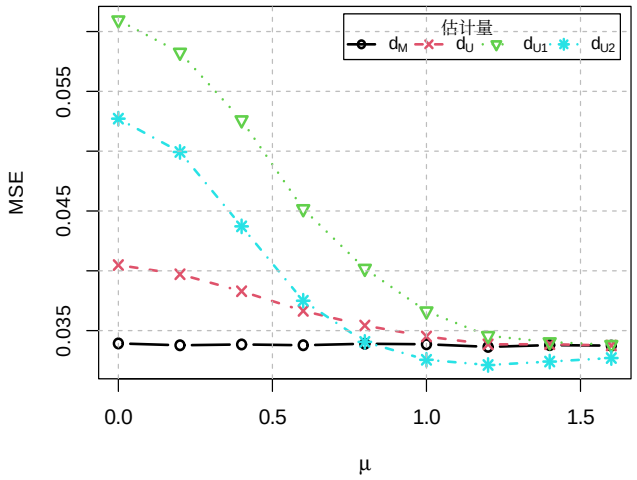
(a) $n_1 = 15, n_2 = 5$



(b) $n_1 = 15, n_2 = 20$



(c) $n_1 = 5, n_2 = 15$



(d) $n_1 = 10, n_2 = 30$

图 4.2：不同 n_1 和 n_2配置下各种竞争估计量的缩放均方误差图

所选治疗均值的估计

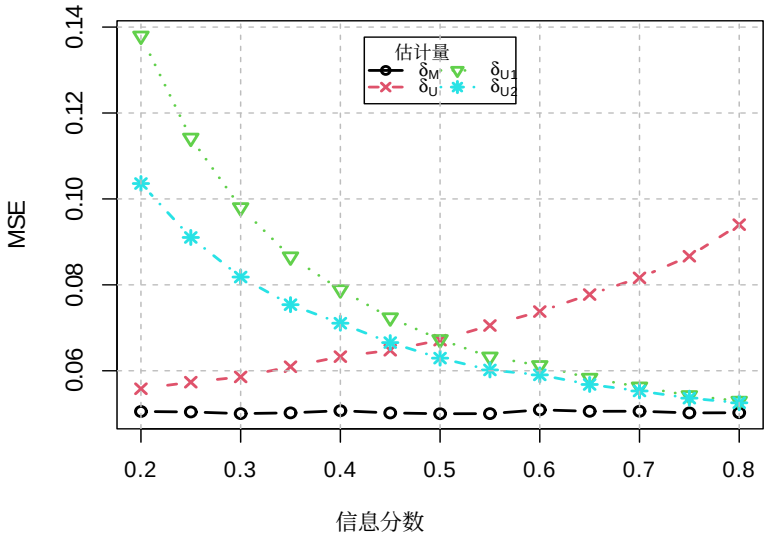


图 4.3：不同竞争估计量在固定 $n = n_1 + n_2$ =20 和 $\mu = 0.1$ 时的缩放均方误差图。

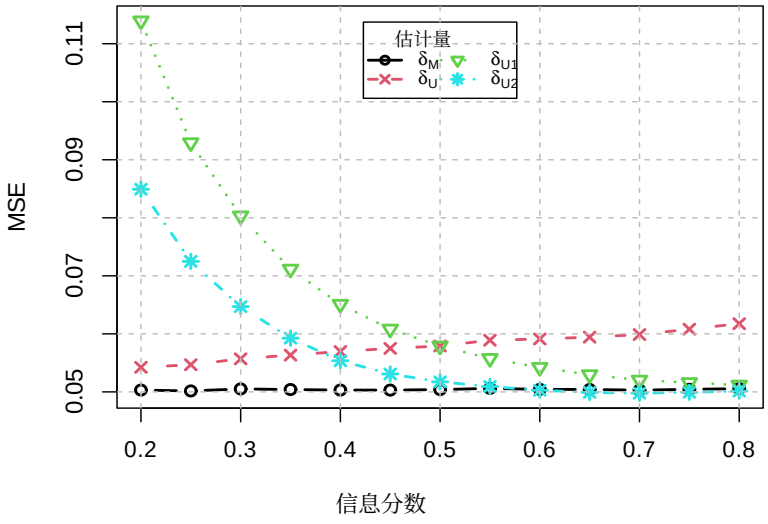


图4.4：固定 n $= n_1 + n_2 = 20$ 和 $\mu = 0.6$ 时，不同竞争估计量的缩放均方误差图。

所选治疗均值的估计

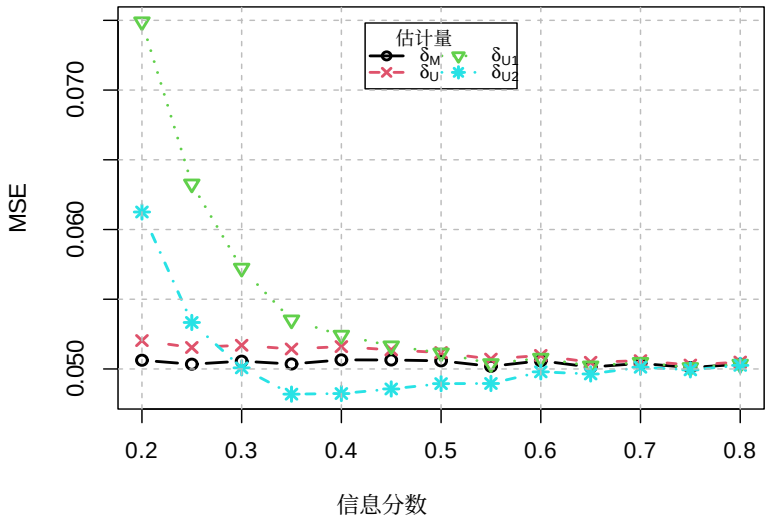


图 4.5：固定 $n = n_1 + n_2 = 20$ 和 $\mu = 1.2$ 下不同竞争估计量的缩放均方误差图。

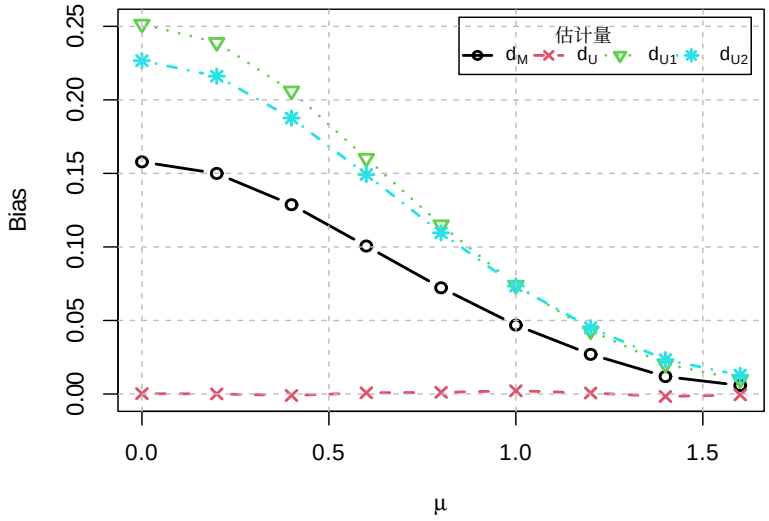
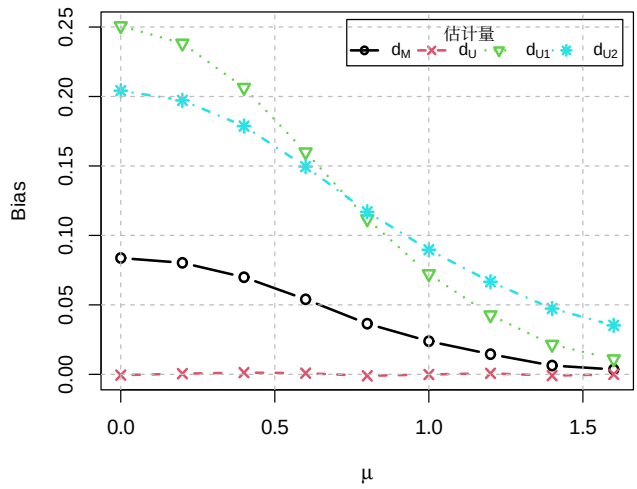
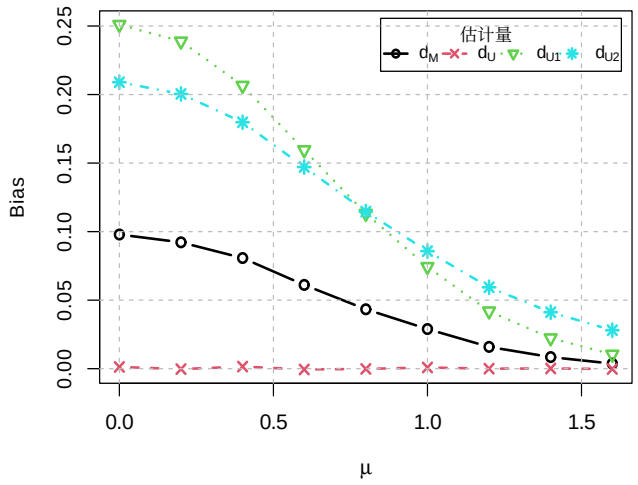


图 4.6： $n_1 = 5$、 $n_2 = 3$ 的不同竞争估计量的缩放偏差图。

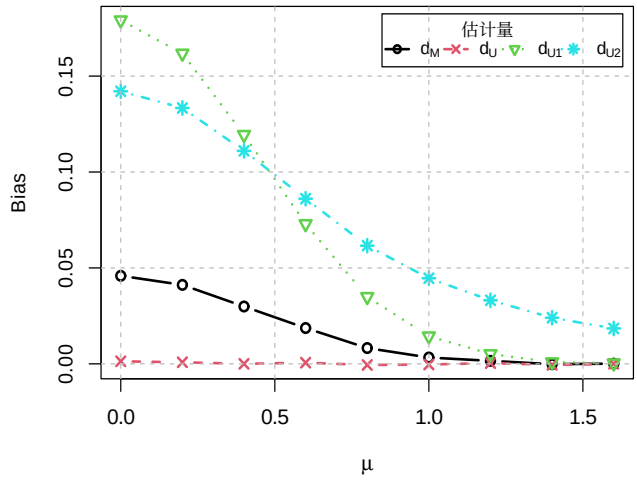
所选治疗均值的估计



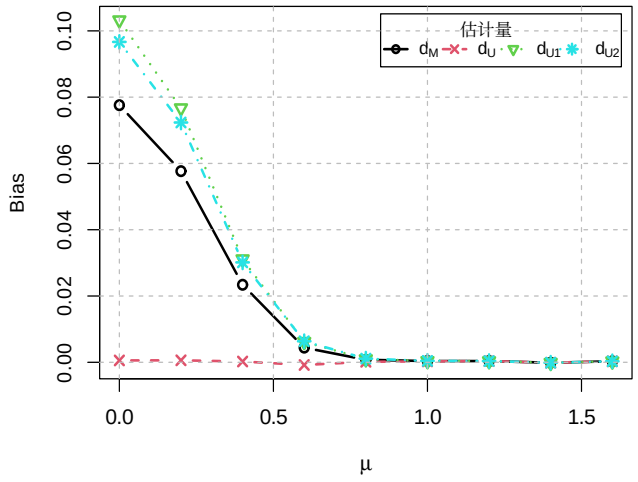
(a) $n_1 = 5, n_2 = 5$



(b) $n_1 = 5, n_2 = 8$



(c) $n_1 = 10, n_2 = 30$



(d) $n_1 = 30, n_2 = 10$

图 4.7: 不同 n_1 和 n_2 配置下各种竞争估计量的缩放偏差图

5 真实数据示例

本节将论文中的理论结果应用于一个数据集进行说明。该数据集的详细信息可通过以下链接获取：
https://vincentarelbundock.github.io/Rdatas
数据列于下方的表 1 中。该实验比较了 60 只食用不同饮食的幼鼠的体重增量。其中一半幼鼠食用低蛋白饮食，其余幼鼠食用高蛋白饮食。蛋白质来源为牛肉、谷物或猪肉。

使用 Shapiro-Wilk 正态性检验，高蛋白饮食的 p 值为 0.834，低蛋白饮食的 p 值为 0.771。据此，我们认为食用高蛋白饮食和低蛋白饮食的幼鼠体重增量所对应的总体近似服从高斯分布。通过 F 检验，两个总体的方差相等这一假设也得到接受， p 值为 0.631。因此，可以认为这些数据来自两个正态总体 $N(\mu_1, \sigma^2)$ 和 $N(\mu_2, \sigma^2)$，且 $(\widehat{\mu_1}, \widehat{\mu_2}) = (92.5, 81.6)$。

为选出有效的蛋白质饮食（即能带来最大体重增量的饮食），我们从每个总体组中抽取大小为 $n_1 = 20$ 的样本，并计算$\overline{X}_1$ 和 $\overline{X}_2$。如果 $\overline{x}_1 > \overline{x}_2$，我们就选择与高蛋白饮食对应的总体；如果$\overline{x}_2 \geq \overline{x}_1$，则选择与低蛋白饮食对应的总体。在第二阶段，我们从第一阶段选出的总体中额外抽取大小为 $n_2 = 10$ 的样本，并计算第二阶段的样本均值$\overline{Y}$。最后，根据计算得到的$\overline{X}_1$、$\overline{X}_2$、 S^2 和 $\overline{Y}$，计算估计值 d_M、 d_U、 d_{U1}和 d_{U2}。

表 1 大鼠食用高蛋白饮食后的体重增量。

7 3 102 118 104 81 10 7
第一阶段数据 100 87 117 111 98 74 56 111 95 88 82 77 86 92 大鼠食用低蛋白饮食后的体重增量。90 76 90 64 86 51 72 第一阶段数据 90 95 78 107 107 97 80 98 74 74 67 89 58 大鼠食用高蛋白饮食后的体重增量。第二阶段数据 94 79 96 98 102 102 108 91 120 105

所选治疗均值的各种两阶段估计量 μ_Q 列于下表2：

表2：所选治疗均值的各种估计值 μ_Q。

d_M	d_U	d_{U1}	d_{U2}
95.13	91.23	88.34	89.34

从表2中给出的所选治疗均值的各种估计量可以看出，两阶段自适应估计量 d_M 和 d_U 接近所选治疗均值的真实值，即 $\mu_Q = 92.5$。

6 总结

本文研究了在两阶段自适应设计下估计所选治疗均值的问题。通过考虑第二阶段有多个观测值可用的情形，我们扩展了 Cohen 和 Sackrowitz（1989）的研究。

在方差未知情形下，我们推导出了所选治疗均值的 UMVCUE，并证明最大似然估计量（第一阶段和第二阶段样本均值的加权平均，权重与相应样本量成正比）是极小极大的且可容许的。我们还提出了两个代入估计量（d_{U1} and d_{U2}）：在 Masihuddin和Misra（2022年）提出的某些估计量中，以合并样本方差代替共同方差 σ^2。我们通过模拟，尝试考察这四种估计量在不同试验设计下的优缺点。

参考文献

Bahadur, R. R. 和 Goodman, L. A.（1952）。无偏决策规则与充分统计量。《数理统计年刊》，第553–562页。Bauer, P. 和 Kieser, M.（1999）。在单项试验中结合医学治疗开发的不同阶段。《医学统计》，18(14):1833–1848。Bowden, J. 和 Glimm, E.（2008）。两阶段试验中所选治疗均值的无偏估计。《生物统计学杂志：生物科学中的数学方法杂志》，50(4):515–527。Bowden, J. 和 Glimm, E.（2014）。多阶段淘汰劣者试验中所选治疗均值的条件无偏估计与近似无偏估计。《生物统计学杂志》，56(2):332–349。Carreras, M. 和 Brannath, W.（2013）。试验中期进行治疗选择的两阶段自适应设计中的收缩估计。《医学统计》，32(10):1677–1690。Chiu, Y.-D.、Koenig, F.、Posch, M. 和 Jaki, T.（2018）。亚群选择临床试验中的设计与估计。《医学统计》，37(29):4335–4352。Cohen, A. 和 Sackrowitz, H. B.（1989）。所选均值的两阶段条件无偏估计量。《统计与概率快报》，8(3):273–278。Eaton, M. L.（1967）。排序程序的一些最优性质。《数理统计年刊》，38(1):124–137。

Kimani, P. K.、Todd, S.、Renfro, L. A.、Glimm, E.、Khan, J. N.、Kairalla, J. A. 和 Stallard, N.（2020）。基于事件发生时间数据和生物标志物驱动的亚群选择的两阶段自适应设计中的点估计与区间估计。《医学统计》，39(19):2568–2586。Kimani, P. K.、Todd, S. 和 Stallard, N.（2013）。II/III期临床试验中因无效而提前停止时的条件无偏估计。《医学统计》，32(17):2893–2910。Masihuddin 和 Misra, N.（2021）。从具有共同未知方差的两个正态总体中进行选择后的等变估计。《统计学》，55(6):1407–1438。Masihuddin 和 Misra, N.（2022）。两阶段淘汰劣者设计中所选治疗均值的估计。arXiv预印本 arXiv:2209.08567。Misra, N. 和 Dhariyal, I. D.（1994）。异方差条件下自然决策规则的非极小极大性。《统计与决策》，(12):79–98。Misra, N. 和 Singh, G.（1993）。论所选指数总体参数估计的一致最小方差无偏估计量。《印度统计学会期刊》，31(1):61–9。Pallmann, P.、Bedding, A. W.、Choodari-Oskooei, B.、Dimairo, M.、Flight, L.、Hampson, L. V.、Holmes, J.、Mander, A. P.、Odondi, L.、Sydes, M. R. 等（2018）。临床试验中的自适应设计：为何采用，以及如何实施和报告。《BMC医学》，16(1):1–15。Putter, J. 和 Rubinstein, D.（1968）。关于所选总体均值的估计，技术报告第165号。威斯康星大学统计系。Robertson, D. S.、Choodari-Oskooei, B.、Dimairo, M.、Flight, L.、Pallmann, P. 和 Jaki, T.（2022）。自适应试验设计的点估计 I：方法学综述。《医学统计》。Robertson, D. S. 和 Glimm, E.（2019）。未知方差正态情形下的条件无偏估计。《统计学通讯：理论与方法》，48(3):616–627。Sampson, A. R. 和 Sill, M. W.（2005）。淘汰劣者设计：正态情形。《生物统计学杂志：生物科学中的数学方法杂志》，47(3):257–268。Stallard, N. 和 Friede, T.（2008）。一种用于治疗选择临床试验的组序贯设计。《医学统计》，27(29):6209–6227。Tappin, L.（1992）。所选二项总体参数的无偏估计。《统计学通讯：理论与方法》，21(4):1067–1083。