arXiv:2303.01602v1 [stat.ME] 2 Mar 2023

Mission Imputable：对删失协变量进行插补时校正 Berkson 误差

Kyle F. Grosser*生物统计学系，北卡罗来纳大学教堂山分校 Gillings 全球公共卫生学院 Sarah C. Lotspeich 统计科学系，维克森林大学 Tanya P. Garcia 生物统计学系，北卡罗来纳大学教堂山分校 Gillings 全球公共卫生学院

2023年3月6日

摘要

为临床试验选择结局时，针对亨廷顿病这一致命性神经退行性疾病检验实验性疗法，分析人员会对潜在结局如何随时间变化进行建模。然而，亨廷顿病受试者在疾病进展的不同阶段接受观察的情况很常见。为考虑这些差异，分析人员在对潜在结局建模时将距临床诊断时间作为协变量，但该协变量常常存在删失。一种常用解决方案是插补，即利用基于其他数据建立的删失协变量模型所生成的预测值，对删失值进行插补，然后分析插补数据集。然而，当插补模型设定错误时，我们的结局模型估计值可能会产生偏差。为解决这一问题，我们开发了一种新方法，称为“ACE 插补”。首先，我们将插补值建模为真实协变量值的误差版本。

*作者谨此致谢美国国家环境卫生科学研究所资助项目 *T32ES007018* 和 *P30ES010126*，以及美国国家神经疾病与卒中研究所资助项目 *K01NS099343*。

然后，我们利用半参数理论对这些误差进行校正。具体而言，我们推导出一种结局模型估计量：即使删失协变量是使用设定错误的插补模型进行插补的，该估计量仍然相合。模拟结果表明，即使插补模型设定错误，ACE 插补在经验上仍然无偏；相比之下，多重插补会产生 $> 100\%$ 的偏差。将我们的方法应用于一项亨廷顿病研究，可以确定旨在减缓疾病进展的临床试验的结局。

关键词： 删失数据、亨廷顿病、插补校正、测量误差、半参数理论

1 引言

1.1 亨廷顿病临床试验的统计难题

目前，临床试验正在进行中，以检验旨在减缓亨廷顿病进展的实验性疗法。亨廷顿病是一种遗传性疾病，会导致认知能力和运动功能逐渐受损。迄今尚未开发出针对该疾病的有效治疗方法；其中一个主要难题是寻找一种结局，以帮助评估实验性疗法是否产生了效果（Langbehn and Hersch，2020年）。临床试验研究人员希望避免选择随时间变化过慢的结局，因为这样将难以确定实验性疗法是否确实减缓了疾病进展。这是因为，对于变化缓慢的结局而言，无论接受何种治疗，在试验期间其进展都不会发生太大变化。

因此，理想的结局应当是在临床试验过程中能够轻松检测到变化的结局，但要确定这样的结局在统计学上颇具挑战。通常，从疾病进展程度各不相同的个体（有些病情较其他人更为严重）中收集用于衡量认知障碍和运动障碍随时间变化的数据。这些差异可能会妨碍我们寻找适用于临床试验的最佳结局。例如，如果不调整个体所处的疾病进展阶段不同这一因素，符号数字模态测验（SDMT）评分（认知障碍的一种衡量指标）图显示，在研究过程中几乎没有变化，甚至毫无变化（图 1A）。为捕捉疾病进展，我们采用距 临床诊断时间，指受试者被观察的时间与其 临床诊断时间之间的差值（即临床医生判定受试者的运动障碍明确归因于亨廷顿病的那一天）（Kieburtz et al., 1996）。现在，我们不再单纯按研究时间绘制符号数字模式测验评分，而是将其作为距临床诊断时间的函数进行绘制（即进行调整

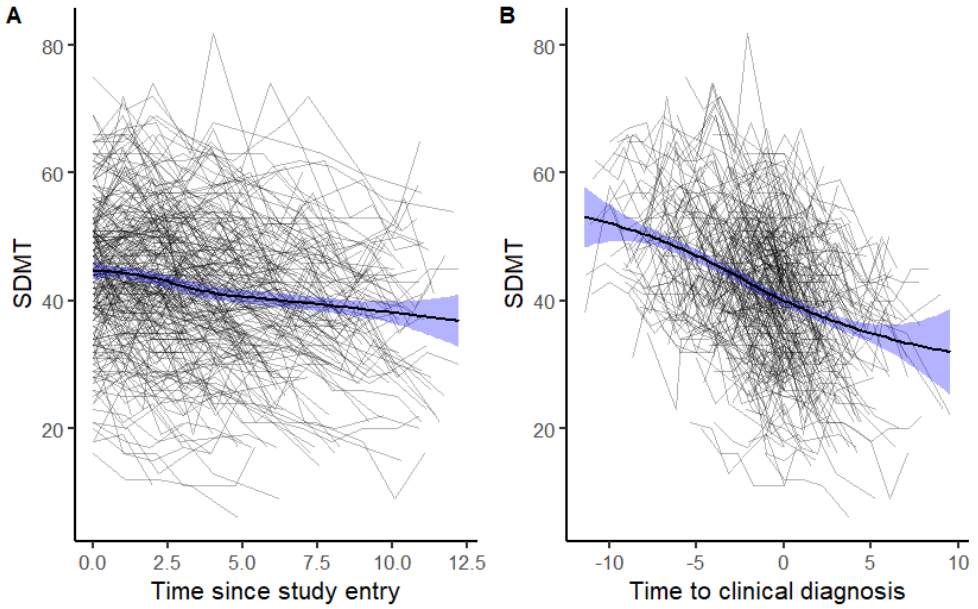

图 1：未删失（即已接受临床诊断）受试者的符号数字模式测验（SDMT）评分分别相对于（A）入组后的时间和（B）距临床诊断时间绘制。细线表示各个体随时间的变化，粗线表示概括总体轨迹的局部加权回归曲线。

疾病进展），由此可见，在诊断前后紧邻的一段时间内，认知障碍会迅速加重（图 1B）。基于调整后的认知障碍变化轨迹得出的这些发现与既往临床发现一致，而基于未调整轨迹得出的发现则不一致（Paulsen 等人，2008）。

将认知障碍建模为距临床诊断时间的函数，是校正个体疾病进展程度差异的一种常见方法（Long 等人，2014年）。对于来自已完成诊断的队列的数据，在分析中进行这种校正会很简单。然而，由于亨廷顿病的进展缓慢，

在某个队列中，至少有些受试者的临床诊断时间往往无法观察到。尽管如此，亨廷顿病具有完全外显性，即任何遗传了该基因突变的人最终都应符合诊断标准。因此，我们知道，对于尚未确诊的受试者，其临床诊断时间必定晚于最后一次观察时间。这种现象称为右删失。因此，为亨廷顿病临床试验选择结局时，我们必须准确地对给定右删失协变量——临床诊断时间时潜在结局的变化进行建模。

1.2 使用删失协变量进行统计建模

人们可能会想用删失受试者最后一次观测的时间（例如，最后一次访视）代替其临床诊断时间，然后用这些替代值对潜在结局进行建模。然而，这些右删失的替代值必然小于真实的临床诊断时间。因此，这类“朴素”分析往往会导致模型估计值存在偏差（Austin 和 Brunner，2003年）。另一种做法是从数据中剔除删失受试者，然后仅使用未删失数据对潜在结局进行建模。然而，已有充分证据表明，这类“完整案例”分析所得的参数估计值不如使用完整数据集时所得的估计值精确（即效率较低）（Fitz maurice 等，2011年）。完整案例分析还可能导致参数估计值存在偏差，并使假设检验的第一类错误率膨胀（Austin 和 Brunner，2003年）。

因此，我们不希望像完整案例分析那样“丢弃”未删失数据，但也不能像朴素分析那样“照原样”使用数据。插补——即先对删失协变量值进行插补，再分析插补数据集——提供了一种更有前景的解决方案。这些插补值可以通过多种方式，根据其他完全观测到的数据建立的删失协变量模型（即“插补模型”）生成，

包括随机抽样（Bernhardt 等人，2014年；Wei 等，2018年）、条件均值（Atem 等人，2019）和条件分位数（Wang 和 Feng，2012年；Yu 等人，2021年）。

尽管针对删失协变量的插补仍是一个不断发展的研究领域，但我们可以从有关“传统”缺失数据插补的大量文献中汲取许多经验。重要的是，众所周知，在满足两个关键假设的情况下，对传统缺失数据进行插补可以得到相合的结果模型参数估计值：i) 数据为随机缺失；ii) 插补模型设定正确（Little 和 Rubin，2019）。具体而言，随机缺失假设要求：给定其他完全观测到的变量后，变量“缺失”的概率与缺失变量本身条件独立。对于删失数据，这一假设显然不成立，因为变量 X 被变量 C 右删失的概率（即 $P(C < X)$）直接取决于 X。例如，对于右删失数据，如果 $X = 2$，则可能被删失的变量 X 比在 $X = 4$的情况下更有可能被变量 C 删失，因为 $P(C < 4) \leq P(C < 2)$（通常严格不等）。幸运的是，Bernhardt 等（2014）和 Wang and Feng（2012）通过理论分析和模拟表明，使用插补处理删失协变量仍然可以获得相合的参数估计值。这些发现表明，插补删失协变量时，随机缺失假设并非必要假设。然而，所有上述插补方法都依赖于对删失协变量建立模型；当该插补模型设定错误时，可能会给我们的结局模型估计值引入偏差（Yucel 和 Demirtas，2010；Black 等人，2011）。

1.3 概述

本文提出了一种新方法，用于在给定删失协变量的情况下，对连续型结果的线性混合效应模型进行相合估计。我们在现有删失协变量插补

方法的基础上，通过考虑插补模型设定错误来减少偏差；当插补模型设定错误时，这些方法会导致结局模型估计值产生偏差。为考虑插补模型可能存在的设定错误，我们使用 Berkson 误差模型（Carroll 等人，2006）对真实的（但被删失的）协变量值与其插补替代值之间的差异进行建模。随后，为避免可能错误设定这种“插补误差”的分布（因为这可能导致偏差和效率低下），我们调整了 Garcia 和 Ma（2016）提出的一种灵活的半参数方法，使插补误差可以服从任意分布。我们在第 2 节介绍所提出的方法——“插补误差的主动校正”（简称“ACE 插补”）。在第 3 节中，我们证明了即使插补模型存在设定错误，ACE 插补仍能产生一个可识别、相合且渐近正态的估计量。在第 4 节中，模拟结果表明，当插补模型正确时，ACE 插补优于针对删失协变量的竞争性解决方案；即使底层插补模型存在设定错误，它仍保持无偏。在第 5 节中，我们展示了将 ACE 插补应用于亨廷顿病数据，有助于确定可用于检验亨廷顿病实验性疗法的结局。第 6 节通过讨论潜在局限性和未来工作对本文作出总结。

2 方法

为了更准确地确定亨廷顿病临床试验的结局（即变化速度足够快、从而能够可靠地识别有效疗法的结局），我们必须准确地对潜在结局随时间的进展进行建模，并将其表示为临床诊断时间这一右删失协变量的函数。本节将介绍我们的 ACE 插补方法，该方法在给定右删失协

变量的情况下估计纵向模型的参数。为实现这一目标，我们首先采用一种现有的协变量删失解决方案，即根据 Cox模型插补删失值。随后，鉴于当插补模型设定错误时，该方法可能产生偏差，我们通过校正此类模型设定错误所导致的误差来减少这种偏差。

2.1 符号表示与纵向模型

我们利用为 n 名受试者（索引为 i, $i = 1$、.....、 n）收集的数据，对潜在结局进行纵向建模；每名受试者有 m_i 个观测值（索引为 j, $j = 1$、.....、 m_i）。受试者 i 在第 j 次访视时的结局为 Y_{ij}，观测时间为 s_{ij}。我们还观测临床诊断时间 X_i、 p_a维协变量 $\boldsymbol{Z}_{ij}^a$，以及 p_b维协变量 $\boldsymbol{Z}_{ij}^b$。为考虑同一受试者各结局之间的聚类，我们纳入一个 p_b维向量，表示未观测到的受试者特异性随机效应 $\boldsymbol{b}_i$。例如，在亨廷顿病研究中， Y_{ij}是临床试验所考察的结局， s_{ij} 是自研究开始以来的观测时间（年）， X_i 是相对于研究开始时间的临床诊断时间（年）， $\boldsymbol{Z}_{ij}^a$ 为基线年龄、性别、受教育年限和遗传信息， $\boldsymbol{b}_i$ 是受试者特异性随机截距（因此， $\boldsymbol{Z}_{ij}^b$ 只是一个全 **1** 向量）。

为考　虑 Y_{i1}、.....、 Y_{im_i}之间的聚类，我们采用线性混合模　　l:

$$Y_{ij} = \alpha(s_{ij} - X_i) + \boldsymbol{\beta}^T \boldsymbol{Z}_{ij}^a + \boldsymbol{b}_i^T \boldsymbol{Z}_{ij}^b + \epsilon_{ij}, \quad \boldsymbol{b}_i \sim f_{\boldsymbol{b}}, \quad \epsilon_{ij} \sim \text{Normal}(0, \sigma^2).$$

其中， α 是与（$s_{ij} - X_i$）相关的参数， $\boldsymbol{\beta}$ 是与 $\boldsymbol{Z}_{ij}^a$ 相关的 p_a维参数向量， σ^2 是随机误差 ϵ_{ij}的方差。

我们假设各受试者彼此独立，并且给定 $\boldsymbol{b}_i \sim f_{\boldsymbol{b}}$ 后，第 i 位受试者的结局 Y_{ij} 条件独立。通常，在混合模型中，分析人员会假设 $\boldsymbol{b}$ 服从某种分布，例如均值为 **0**、协方差矩阵未知的多元正态分布。为提高模型灵活性，我们改为允许 $\boldsymbol{b}$ 服从

任意分布 f_b。这样做是为了避免可能对 f_b 设定错误，因为模型设定错误可能导致偏差和效率降低（Garcia 和 Ma，2016年）。

我们将 $(s_{ij} - X_i)$ 称为“距临床诊断时间”（而 X_i 本身表示临床诊断时间）。举例来说，当受试者 i 尚未接受临床诊断、距离诊断还有两年时，$(s_{ij} - X_i) = -2$；而在受试者 i 接受临床诊断两年后，$(s_{ij} - X_i) = 2$。如第 1 节所述，我们使用这个差值，而非仅使用 s_{ij}，以反映有些受试者的疾病进展程度比其他受试者更深这一事实。

该模型面临的一项主要挑战是，X_i 常常会发生右删失；在我们分析的亨廷顿病数据中，X_i 有 $\approx 78\%$ 的受试者发生右删失。因此，我们观测到的不是 X_i，而是 $W_i = \min(X_i, C_i)$ 和 $\Delta_i = I(X_i \leq C_i)$，其中 C_i 是连续的随机右删失时间，Δ_i 是删失指示变量。我们将 C_i 称为“右删失”时间，因为当临床诊断时间 X_i 被 C_i删失时，我们知道 X_i 位于 C_i的右侧，即 $X_i > C_i$。

总之，对于 $i = 1, \ldots, n$，观测到的数据为 $(\boldsymbol{Y}_i, \boldsymbol{s}_i, \boldsymbol{Z}_i^a, \boldsymbol{Z}_i^b, W_i, \Delta_i)$，其中 $\boldsymbol{Y}_i = (\boldsymbol{Y}_{i1}, \ldots, \boldsymbol{Y}_{im_i})^T$, $\boldsymbol{s}_i = (s_{i1}, \ldots, s_{im_i})^T$, $\boldsymbol{Z}_i^a = (\boldsymbol{Z}_{i1}^a, \ldots, \boldsymbol{Z}_{im_i}^a)^T$，以及 $\boldsymbol{Z}_i^b = (\boldsymbol{Z}_{i1}^b, \ldots, \boldsymbol{Z}_{im_i}^b)^T$。我们希望在 X_i 存在右删失的情况下，估计 $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^{\mathrm{T}}, \sigma^2)^{\mathrm{T}}$。

2.2 诊断时间的删失值插补

插补是解决协变量删失的一种有力方法，因为它能让我们“补上”被删失的临床诊断时间。随后，我们便可以使用限制性最大似然估计（REML）等传统方法分析这个补全后的数据集。此外，保留所有受试者的数据，可以提高参数估计值的效率，相较于完整病例分析尤为如此。分析插补数据集的效率无法达到分析从未被删失的数据集的水平，但仍比直接丢弃删失受试者更有效率。

因此，为了在存在协变量删失的情况下估计 $\boldsymbol{\theta}$，我们首先对 X_i的删失值进行插补。具体而言，我们采用基于Cox模型的条件均值插补（Atem 等人，2019；Lotspeich等人，2022），这种方法可对删失进行调整，且不会对临床诊断时间 X_i施加严格的分布假设。首先，我们拟合一个Cox模型，以诊断时间 X_i为结局，并纳入一组时不变协变量 $\boldsymbol{V}_i$。协变量 $\boldsymbol{V}_i$ 可以是 $\boldsymbol{Z}_i^a$ 或 $\boldsymbol{Z}_i^b$ 中的部分变量，也可以是辅助变量。Cox模型假设，给定 $\boldsymbol{V}_i$ 时 X_i 的风险函数为

$$\lambda(x|\boldsymbol{V}_i) = \lambda_0(x)\exp(\boldsymbol{\gamma}^T\boldsymbol{V}_i),$$

其中，$\lambda_0(x)$ 和 $\boldsymbol{\gamma}$ 分别为基线风险函数和协变量效应。$\boldsymbol{\gamma}$ 的各元素表示与 $\boldsymbol{V}_i$中相应元素对应的对数风险比。拟合此模型后，我们将临床诊断时间 X_i 替换为:

$$\widehat{X}_i \;=\; \Delta_i X_i + (1-\Delta_i)\mathrm{E}(X_i|X_i > C_i, \boldsymbol{V}_i). \tag{1}$$

对于未删失受试者，X_i 保持不变（即，对于满足 $\Delta_i = 1$ 的 i，$X_i = X_i$）。对于删失受试者，我们将 X_i 替换为其在给定 $X_i > C_i$ 和 $\boldsymbol{V}_i$ 条件下的条件均值（即，对于满足 $\Delta_i = 0$ 的 i，$\sim X_i = \mathrm{E}(X_i|X_i > C_i,\ \boldsymbol{V}_i)$）。Lotspeich等人（2022年）表明，这一条件期望可近似为 $\mathrm{E}(X_i|X_i > C_i,\ \boldsymbol{V}_i) =$

$$C_i + \tfrac{1}{2}\left[\frac{\sum_{j=1}^{n-1}\mathrm{I}(W_{(j)}\geq C_i)\left\{S_0\left(W_{(j+1)}\right)^{\exp(\boldsymbol{\gamma}^T\boldsymbol{V}_i)} + S_0\left(W_{(j)}\right)^{\exp(\boldsymbol{\gamma}^T\boldsymbol{V}_i)}\right\}\left(W_{(j+1)} - W_{(j)}\right)}{S_0(C_i)^{\exp(\boldsymbol{\gamma}^T\boldsymbol{V}_i)}}\right],$$

其中，$W_{(1)} < W_{(2)} < \cdots < W_{(n)}$ 是 $W = \min(X, C)$的有序取值，而 $S_0(x|\boldsymbol{V}_i)$是给定 $\boldsymbol{V}_i$ 时 X_i 的基线生存函数。为了计算这一近似值，我们使用标准统计软件估计对数风险比 $\boldsymbol{\gamma}$，并使用 Breslow 估计（Breslow，1972）估计基线累积风险函数 $S_0(\cdot)$。

2.3 为什么插补会引入误差

要使插补得到相合的参数估计值，我们必须正确设定此 Cox^模型。当模型设定错误时，插补值 X_i 可能与临床诊断的真实时间 X_i相差甚远。正如我们在第 4 节中通过实证所示，这种“插补^误差”（即 X_i 与 X_i之间的差异）可能导致严重偏差。

为减少这种偏差，我们对插补误差进行建模。Haber等人（2020）指出，插补误差属于 Berkson 误差，当同一组中的受试者被赋予相同的缺失变量值时，就会产生这种误差。当特征相似的受试者都被赋予同一个值（例如估计的组均值）时，由于个体真实值与该组均值之间存在未观测到的差异，便会产生误差。采用条件均值^插补时，我们用 $X_i = \mathrm{E}(X_i|X_i > C_i,\ \boldsymbol{V}_i)$替代删失的 X_i。因此，如果两个受试者 i和 j具有相同的特征，即（$C_i,\ \boldsymbol{V}_i$）$=$（$C_j,\ \boldsymbol{V}_j$），我们就会为他们分别赋予相同的插补^ ^值，即 $X_i = X_j$。因此，我们知道，X_i 与$\sim X_i$ 之间的关系符合 Berkson误差模型。我们将其表示为

$$X_i = \widehat{X}_i + U_i$$

^存在测量误差 $U_i \sim f_U$ （Carroll 等人，2006）。根据公式（1）中 X_i 的定义，当受试者的临床诊断时间未删失时（即对于满足 $\Delta_i = 1$ 的 i，$U_i = 0$），不存在插补误差，因为我们无需对该受试者进行插补。

2.4 校正插补误差

由于我们知道，X_i 的真实（但经删失）值与其插补^替代值 X_i 之间的误差可能导致偏差（第 4 节），因此我们对这一插补误差（U_i）进行调整，以便更准确地估计 $\boldsymbol{\theta}$。为在最灵活的建模假设下估计 $\boldsymbol{\theta}$，我们不会对随机效应 $\boldsymbol{b}_i$或插补

误差 U_i作任何分布假设。相反，我们将随机效应和插补误差视为所谓的潜变量（即未观测到的变量），因此我们的纵向混合效应模型属于广义线性潜变量模型这一类。对于这类模型，存在一种半参数方法（Garcia 和 Ma，2016），该方法利用一种中间量来估计模型参数，而这种中间量发挥着类似于经典充分完备统计量的作用。该方法得到的参数估计量具有相合性、有效性，并且对模型设定错误具有稳健性。

将这种半参数方法扩展到我们的问题后，我们在附录 A 中证明，$\widehat{}$感兴趣的估计量 $\boldsymbol{\theta}_{\mathrm{ACE}}$是以下估计方程的解：

$$\sum_{i=1}^{n}\{\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\}\equiv\sum_{i=1}^{n}\left\{\Delta_i\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})+(1-\Delta_i)\boldsymbol{S}_{\mathrm{eff}}^{*}(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\right\}=\mathbf{0},\quad(2)$$

其中，$\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})$ 是未删失数据的得分向量 $\boldsymbol{D}_i=(\boldsymbol{s}_i,X_i,\boldsymbol{Z}_i^a,\boldsymbol{Z}_i^b)$，而$\boldsymbol{S}^*_{\mathrm{eff}}(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})$ 是删失数据的得分向量 $\widehat{\boldsymbol{D}}_i=(\boldsymbol{s}_i,\widehat{X}_i,\boldsymbol{Z}_i^a,\boldsymbol{Z}_i^b)$。这些得分向量的推导表明，最初构造这些得分向量需要求解一个计算速度慢且数值不稳定的问题。然而，借助多元正态分布和线性投影理论的性质，我们推导出了简明的闭式得分向量（见附录 B）。

在附录 C.1 中，我们证明，$\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})=$

$$\sigma^{-4}\begin{bmatrix}\sigma^2\boldsymbol{Z}_i^{aT}(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\}\\ \sigma^2(\boldsymbol{s}_i-X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\}\\ \frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i-\mathrm{E}(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)\}-\boldsymbol{\zeta}^T(\boldsymbol{D}_i;\boldsymbol{\theta})\{\boldsymbol{Y}_i-\mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)\}\end{bmatrix},$$

其中

$$\begin{aligned}\mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)&=\boldsymbol{P}_{\boldsymbol{Z}_i^b}\boldsymbol{Y}_i+(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\\ \mathrm{E}(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)&=\sigma^2(m_i-p_b)+||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})||_2^2,\end{aligned}$$

其中 $\boldsymbol{T}_i = \sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$、 $\boldsymbol{P}_{\boldsymbol{Z}_i^b} = \boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\boldsymbol{Z}_i^{bT}$，以及 $\boldsymbol{\zeta}(\boldsymbol{D}_i;\ \boldsymbol{\theta}) = \boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha($ $\boldsymbol{S}_i - X_i\mathbf{1}_{m_i})$。^在附录 C.2 中，我们证明了 $\boldsymbol{S}^*{}_{\text{eff}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta}) =$

$$\sigma^{-4}\begin{bmatrix} \sigma^2\boldsymbol{Z}_i^{aT}(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\} \\ \sigma^2(\boldsymbol{s}_i - \widehat{X}_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\} \\ \frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i^{\text{aug}},\widehat{\boldsymbol{D}}_i)\} - \boldsymbol{\zeta}^T(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\{\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}_i^{\text{aug}},\widehat{\boldsymbol{D}}_i)\} \end{bmatrix},$$

其中

$$\begin{aligned} \mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}_i^{\text{aug}},\widehat{\boldsymbol{D}}_i) &= \boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)}\boldsymbol{Y}_i + (\boldsymbol{I}_{m_i} - \boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)})\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta}) \\ \mathrm{E}(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i^{\text{aug}},\widehat{\boldsymbol{D}}_i) &= \sigma^2(m_i - p_b) + ||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\text{aug}},\widehat{\boldsymbol{D}}_{\mathrm{i}})||_2^2, \end{aligned}$$

其中 $\boldsymbol{T}_{\text{aug}_i} = \sigma^{-2}(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)^T\boldsymbol{Y}_i$, $\boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)} = (\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)((\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)^T(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b))^{-1}(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)^T$，且 $\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta}) = \boldsymbol{Z}_i^a\boldsymbol{\beta} + \alpha(\boldsymbol{S}_i - X_i\mathbf{1}_{m_i})$。

^构造 $S_{\text{eff}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$ 和 $S^*_{\text{eff}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$ 看似复杂，但实际涉及的矩阵代数计算很简单；也许更重要的是，它完全避免了所有同时涉及随机效应 $\boldsymbol{b}$ 和插补误差 U的项。我们利用多元正态分布和线性投影算子的性质，证明包含 $\boldsymbol{b}$ 或 U 的所有项都会从有效得分向量中消失（附录C），从而得出了这一结果。因此，ACE 插补既不要求我们为 $(f_{\boldsymbol{b}},\ f_U)$ 指定形式，也不要求我们估计这些未知分布。因此，所提出的方法不同于传统的最大似然估计；后者要求我们假设 $(f_{\boldsymbol{b}},\ f_U)$ 具有特定形式，例如多元正态分布。当这些分布假设正确时，最大似然估计量的效率会高于半参数或非参数方法；但当这些假设不正确时，最大似然估计量会产生偏差。所提出的方法也不同于传统的半参数估计量，后者要求指定干扰分布；尽管这些方法无论所指定的

分布是否正确，都能产生相合估计量（这是相较于传统最大似然估计的一项优势），但当所指定的分布不正确时，效率会有所损失（Garcia 和 Ma，2016年）。由于我们的方法无需指定干扰分布 (f_b, f_U)，$\hat{\boldsymbol{\theta}}_{\mathrm{ACE}}$因此无需任何关于干扰分布的先验知识，就能同时实现一致性和最优效率。

3 新型估计量的性质

3.1 可识别性

现在我们讨论求解公式（2）的估计量 $\hat{\boldsymbol{\theta}}_{\mathrm{ACE}}$ 的性质。首先讨论 $\hat{\boldsymbol{\theta}}_{\mathrm{ACE}}$ 的可识别性：

定理 *1* 令 $\mathbf{1}_n$ 表示一个全 *1* 向量，$\boldsymbol{I}_n$ 表示一个单位矩阵，并定义 $\boldsymbol{P}_{\boldsymbol{M}} = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T$，其中 $\boldsymbol{M}$ 是一个矩阵。如果矩阵 $\boldsymbol{N} = \sum_{i=1}^{n}\left\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^{b*}})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})\right\}$ 是非奇异的（即，如果 $\boldsymbol{N}$ 可逆），那么 $\hat{\boldsymbol{\theta}}_{\mathrm{ACE}}$ 是可识别的。

我们在附录 D 中证明定理 1。定理 1 中 $\boldsymbol{N}$ 的闭式形式使我们能够轻松检验给定数据集的可识别性：我们可以计算 $\boldsymbol{N}$，如果它是可逆的，我们就知道，在存在干扰分布（f_b、f_U）的情况下，可以唯一估计 $\hat{\boldsymbol{\theta}}_{\mathrm{ACE}}$。

3.2 一致性

接下来，为了证明 $\hat{\boldsymbol{\theta}}_{\mathrm{ACE}}$ 是 $\boldsymbol{\theta}_0$的一致估计量，我们必须首先证明公式（2）中的估计函数是无偏的（即估计函数的均值为 0）。为此，我们首先强调能够得到无偏估计函数的简单假设。

命题 1 考虑以下条件：

(C1) $\boldsymbol{Y}_i$ 在给定 $\boldsymbol{D}_i$的条件下与 C_i 条件独立。

(C2) $\boldsymbol{Y}_i$ 在给定 $\boldsymbol{D}_i$的条件下与 C_i 条件独立。

若 *(C1)* 成立，则 $\mathrm{E}\{\Delta_{\mathrm{i}}\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})\}=\mathbf{0}$。若 *(C2)* 成立，则 $\mathrm{E}\{(1-\Delta_{\mathrm{i}})\boldsymbol{S}^{*}{}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}})\}=\mathbf{0}$。

我们在附录E.2中证明命题 1。回顾一下，$\boldsymbol{D}_i=(\boldsymbol{Z}_i^a,\boldsymbol{s}_i,X_i,\boldsymbol{Z}_i^b)$，且$\widehat{\boldsymbol{D}}_i=(\boldsymbol{Z}_i^a,\boldsymbol{s}_i,\widehat{X}_i,\boldsymbol{Z}_i^b)$。根据这些定义，条件 (C1) 等价于以下说法：一旦我们观察到所有感兴趣的协变量，删失变量本身就不会为我们提供有关结局的新信息。条件 (C2) 与条件 (C1) 类似，只是将 $\boldsymbol{D}_i=(\boldsymbol{Z}_i^a,\boldsymbol{s}_i,X_i,\boldsymbol{Z}_i^b)$ 替换为 $\widehat{\boldsymbol{D}}_i=(\boldsymbol{Z}_i^a,\boldsymbol{s}_i,\widehat{X}_i,\boldsymbol{Z}_i^b)$（即，将 X_i 替换为插补值 $\widehat{X}_i$）。这些条件独立假设是否成立，取决于当前的模型和数据。

鉴于这一无偏性，我们通过将似然型估计量的反函数定理（Foutz，1977）应用于 $\boldsymbol{\theta}_{\mathrm{ACE}}$，来确立其一致性。该定理要求：1）我们的估计函数无偏；2）满足定理 2 中列出的技术条件：

定理 2 考虑以下正则条件：

(C3) 参数 $\boldsymbol{\theta}$ 的定义域 $\boldsymbol{\Theta}$ 是一个紧集。

(C4) 估计方程 $\sum_i \boldsymbol{S}^{\mathrm{full}}{}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})$ 以及 $\mathrm{E}\{\sum_{\mathrm{i}}\boldsymbol{S}^{\mathrm{full}}{}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\ \boldsymbol{\theta})\}$ 是 $\boldsymbol{\theta}$ 在 $\boldsymbol{\theta}_0$ 的某个邻域内的充分光滑函数。

（*C5*）我们的估计方程的期望值$\mathrm{E}\{\sum_{\mathrm{i}}\boldsymbol{S}^{\mathrm{full}}{}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta})\}$有唯一解，并且$\mathrm{E}\{\sup_{\boldsymbol{\theta}\in\Theta}|\sum_{\mathrm{i}}\boldsymbol{S}^{\mathrm{full}}{}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta})|\}$ 的每个分量都有限。

$\widehat{\ }$ $\widehat{\ }$如果条件（*C1*）–（*C5*）成立，那么 $\boldsymbol{\theta}_{\mathrm{ACE}}$ 是 $\boldsymbol{\theta}_0$ 的相合估计（即 $\boldsymbol{\theta}_{\mathrm{ACE}}$ 依概率收敛到 $\boldsymbol{\theta}_0$）。

$\widehat{\ }$ $\boldsymbol{\theta}_{\mathrm{ACE}}$ 的一致性$f_{\boldsymbol{b}}$ f_U尤其令人满意，因为它不依赖于对干扰分布（$f_{\boldsymbol{b}}$、 f_U）的正确设定。事实上，甚至无需为这两个分布中的任何一个指定$\widehat{\ }$工作模型， $\boldsymbol{\theta}_{\mathrm{ACE}}$ 仍能达到一致性。这与现有针对删失协变量的基于插补的解决方案（Wang 和 Feng，2012年；Bernhardt 等人，2014年；Wei 等人，2018年；Atem 等人，2019年；Yu 等人，2021年）不同，后者依赖于对底层插补模型的正确设定。

3.3 渐近正态性

$\widehat{\ }$$\boldsymbol{\theta}_{\mathrm{ACE}}$ 不仅相合，而且渐近正态，如下所示。

定理 3 考虑以下正则条件：

(C6) 矩阵 $n^{-1}\sum_{i=1}^{n}\frac{\partial}{\partial\boldsymbol{\theta}}\{\boldsymbol{S}^{\mathrm{full}}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})\}$ 在概率意义下，在 $\boldsymbol{\theta}_0$ 的某个邻域内一致收敛于矩阵 $E[\frac{\partial}{\partial\boldsymbol{\theta}}\{\boldsymbol{S}^{\mathrm{full}}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})\}]$。

（*C7*）矩阵 $E[\frac{\partial}{\partial\boldsymbol{\theta}}\{\boldsymbol{S}^{\mathrm{full}}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})\}]$ 是 $\boldsymbol{\theta}$ 在邻域 $\boldsymbol{\theta}_0$ 内的有界、光滑函数，并且是非奇异的。

在条件（*C1*）–（*C7*）下，我们有 $n^{1/2}(\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}}-\boldsymbol{\theta}_0)$ 依分布收敛于 *Normal*（0，$\mathcal{V}$），其中矩阵 $\mathcal{V}=\mathcal{A}^{-1}\mathcal{B}(\mathcal{A}^{-1})^{T}$ 满足

$$\mathcal{A} \;=\; \mathrm{E}\left\{\frac{\partial\left\{\boldsymbol{S}^{\mathrm{full}}_{\mathrm{eff}},\boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta}_0)\right\}}{\partial\boldsymbol{\theta}^{\mathrm{T}}}\right\},\quad \mathcal{B}=\mathrm{Var}\{\boldsymbol{S}^{\mathrm{full}}_{\mathrm{eff}},\boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta}_0)\}.$$

定理 3 的证明见附录 F。这一结果非常有力，因为如果条件 (C1)–(C2)（命题 1）、(C3)–(C5)（定理 2）和 (C6)–(C7)（定理 3）均

成立，那么我们就能对感兴趣的参数进行渐近有效推断，$\boldsymbol{\theta}$。例如，我们可以为 $\boldsymbol{\theta}$ 构造Wald型置信区间，并对 $\boldsymbol{\theta}$ 进行渐近有效的假设检验。此外，无须指定或估计干扰分布($f_{\boldsymbol{b}}$, f_U)，即可进行这种渐近推断。再者，这种方法不同于现有方法，后者要求正确指定插补模型。

4 模拟研究

接下来，我们将在模拟研究中比较 ACE 插补与竞争方法的表现。竞争方法是多重条件均值插补：我们执行条件均值插补 M 次，通过对多个插补数据集应用限制性最大似然来估计 $\boldsymbol{\theta}$，并汇总 M 组参数估计值。ACE 插补与竞争方法的核心区别在于，我们的方法会校正插补模型不正确时产生的误差，而竞争方法不会对这些误差进行校正。

4.1 数据生成

我们为 $n = 1000$ 名受试者模拟数据，每名受试者有 $m = 3$个观测值。我们通过两种方式模拟临床诊断时间 X_i 。在第一种模拟情境中，我们根据包含两个不同协变量的Cox模型生成 X_i；随后，在此情境中，我们使用包含这两个协变量的Cox模型对 X_i 的删失值进行插补，因此插补模型设定正确。在第二种模拟情境中，我们根据包含同一协变量线性项和二次项的Cox模型生成 X_i；在此情境中对 X_i 的删失值进行插补时，我们省略二次项，因此

设定错误插补模型。用于生成这些数据的风险函数为 :

$$\lambda(t|\boldsymbol{V}_i) = \lambda_0(t)\exp(\eta_{11}V_{i1} - \eta_{12}V_{i2}), \quad (3)$$

$$\lambda(t|\boldsymbol{V}_i) = \lambda_0(t)\exp(\eta_{21}V_{i1} - \eta_{22}V_{i1}^2), \quad (4)$$

分别对应正确设定的插补模型和设定错误的插补模型。我们将对数风险比设为 $(\eta_{11},\ \eta_{12})^{\mathrm{T}} = (1,\ 0.5)^{\mathrm{T}}$ 和 $(\eta_{21},\ \eta_{22})^{\mathrm{T}} = (1,\ 0.25)^{\mathrm{T}}$。协变量 $V_{i1},\ V_{i2}$ 相互独立，且服从均值为 0、方差为 1 的正态分布。为了生成符合方程 (3) 和 (4) 中 Cox模型的 X_i ，我们采用 Bender等（2005）提出的策略。

我们根据以下线性混合模型生成 Y_{ij}：

$$Y_{ij} = \beta Z_{ij}^a + \alpha(s_{ij} - X_i) + b_i Z_{ij} + \epsilon_{ij}, \quad b_i \sim f_b, \quad \epsilon_{ij} \sim \mathrm{Normal}(0, \sigma^2),$$

其中 $\boldsymbol{\theta} = (\beta,\ \alpha,\ \sigma^2)^{\mathrm{T}} = (1,\ 1,\ 1)^{\mathrm{T}}$。我们生成 $Z_{ij}^a \sim$服从正态分布 Normal(0，1)，$Z_{ij} \sim$服从正态分布 Normal(5，1)，$b_i \sim$ 服从正态分布 Normal(0，1)，以及 $s_{ij} = j - 1$。设置 $s_{ij} = j - 1$ 后，生成的数据中，同一受试者的后续观测值之间以 1 个时间单位均匀间隔，且观测从时间 $= 0$开始；这一选择使 s_{ij} 与 PREDICT-HD数据相符，该数据中后续观测值之间的平均间隔为 ≈ 1.18 年。最后， X_i 受到随机右删失变量 C_i 的删失，该变量由具有不同速率参数 λ_C的指数分布生成。取 $\lambda_C = 0.125$、0.5 和 2 时，删失程度分别为轻度（$\approx 25\%$）、中等（$\approx 50\%$）和重度（$\approx 75\%$）。

4.2 模型评估方法

对于任何被删失的 X_i ，我们使用条件均值插补来插补删失值（第 2.2 节）。首先，我们使用设定正确的模型进行插补，即根据方程（3），同时纳入 V_{i1} 和 V_{i2} 的线性项。在第二种

情形中，我们省略方程（4）中的二次项 V_i^2，从而错误设定插补模型，并拟合以下错误的插补模型： $\lambda^*(t|V_{i1}) = \lambda_0(t)\exp(\eta_{21}^* V_{i1})$。相对于方程（4）中的真实数据生成机制，该模型设定有误。

我们模拟了 1000 个数据集，并针对每个数据集，使用以下方法估计 $\boldsymbol{\theta}$：

1. *Oracle*估计量：我们对完整的未删失数据集应用限制性最大似然。该方法得到的结果，就是在数据从未被删失的情况下所能得到的结果。限制性最大似然是利用未删失数据估计线性混合模型的标准方法（Fitzmaurice 等，2011）。当然，当临床诊断时间已被删失时，这种方法在实践中不可行；因此，在模拟中，我们将这些估计值视为“金标准”或“先知”。

2. 多重条件均值插补（*MCMI*）：我们还采用一种结合条件均值插补的多重插补方法。我们将上文第 2.2 节所述的插补过程重复 $M = 15$ 次，但插补模型中的 $\hat{\boldsymbol{\eta}}$ ——即对数风险比的估计向量——改为从正态分布$\{\hat{\boldsymbol{\eta}}, \text{Cov}(\hat{\boldsymbol{\eta}})\}$ 中抽取（Cole 等，2006）。随后，我们使用多重插补的 Rubin 法则合并这 M 组参数估计值。我们预计，当插补模型设定正确时，该方法是无偏的；而当插补模型设定错误时，则会产生偏差。

3. *ACE* 插补（*ACE*）：我们应用所提出的方法，并预计无论插补模型是否设定正确，该方法都具有无偏性。

为评估这些方法的性能，我们计算经验偏差、标准误估计的经验均值、参数估计值的经验标准差，以及偏差平方的经验均值。此外，我们还报告名义覆盖率为95%的Wald型置信区间的观测覆盖率。我们

不报告 σ^2 的标准误估计和覆盖概率，因为使用限制性最大似然方法时， σ^2 的变异性通常并非关注重点。中等删失和重度删失下的结果分别见表1和表2。轻度删失下的结果见附录H中的表4和表5。

我们使用现有和新开发的软件，在 R 中执行这些插补和估计程序（R 核心团队，2022年）。我们使用*survival*软件包中的 `coxph()` 函数估计所有 Cox模型（Therneau，2022年）；使用*imputeCensoRd* 软件包中的 `condl mean impute()` 函数执行条件均值插补（Lotspeich等人，2022年）；并使用*lme4* 软件包中的 `lmer()` 函数执行限制性最大似然（Bates等人，2015年）。为实施 ACE 插补，我们开发了一个 R 软件包 *ACEimpute*；该软件包包含 `eff score vector()` 函数，用于构建公式（2）中所示的估计方程。最后，为求解该估计方程的根，我们使用*geex* 软件包中的`m estimate()` 函数（Saul和Hudgens，2020年）。

4.3 插补模型设定正确

当插补模型设定正确时，ACE 插补能得到高度准确的参数估计值（表1）。在轻度、中等和重度删失情况下，ACE 插补对每个参数的平均经验偏差均为 < 0.003。由于所有真实参数均为 1，这相当于平均百分比偏差为 $< 1\%$。此外，平均标准误估计能够反映参数估计值的真实变异性。综合来看，这些结果表明，对 α 和 β可以进行有效推断；具体而言，在各种删失率下， α 和 β 的 95% 置信区间的观测覆盖概率均介于 94% 和 96% 之间。

此外，ACE 插补优于竞争方法 MCMI。我们原以为竞争方法会得出无偏的参数估计值，因为条件均值插补——

（随后采用普通最小二乘回归，而非限制性最大似然）在横断面研究中经验证是无偏的（Atem 等人，2019）。然而，在我们的纵向模拟中，MCMI 对 α 的估计平均偏差为 $\geq 20\%$。在三种删失率下，这种方法对 σ^2 的估计偏差也均为 $\geq 150\%$。我们发现，MCMI 对 β 的估计确实准确（平均偏差为 ≈ 0），但这些 β 估计值的变异性高于 ACE 插补所得的估计值。根据这些发现，MCMI 会导致对 α 和 σ^2 的估计严重偏倚。值得庆幸的是，ACE 插补有效消除了这种估计偏差，同时能更有效地估计 β。

4.4 插补模型设定错误

表2展示了从插补模型中省略二次项（因此导致模型设定错误）时的结果。可以看到，在这种情况下，竞争方法 MCMI 的表现远差于插补模型设定正确时的表现。尽管这种方法对β的估计仍相当准确，但在三种删失率下，它对 α 的估计均有 $> 100\%$ 的偏差，并且严重高估了 σ^2，平均偏差为 $>> 100\%$。这种偏差是意料之中的；当插补模型设定错误时，我们应当预期，用来替代删失 X_i的条件均值 E($X_i|X_i > C_i,\ V_{i1}$)会离真实 X_i更远。相比之下，我们发现，插补模型设定错误并未妨碍 ACE 插补的表现；在所考察的两种情况下，其表现都非常相似。即使从插补模型中省略了二次项，ACE 插补仍能得到准确的参数估计值并进行可靠的推断。因此，即使我们事先不知道正确的插补模型（先验地），ACE 插补仍使我们能够相合地估计含有删失协变量的线性混合模型的参数；在生物学场景中，这种情况很容易发生（Haber 等，2020）。

删失	参数	方法	Bias	SEE	ESE	MSE	CPr
None	α	先知	-0.001	0.009	0.010	0.000	0.928
	β	先知	0.000	0.022	0.022	0.000	0.940
	σ^2	先知	-0.002		0.032	0.001	
中等	α	ACE	0.000	0.026	0.024	0.001	0.960
(≈ 50%)		MCMI	0.312	0.033	0.107	0.109	0.000
	β	ACE	0.001	0.026	0.025	0.001	0.958
		MCMI	0.001	0.040	0.040	0.002	0.966
	σ^2	ACE	-0.000	0.036	0.037	0.001	0.940
		MCMI	2.498		2.947	14.914	
重度	α	ACE	0.001	0.028	0.029	0.001	0.947
(≈ 75%)		MCMI	0.206	0.044	0.072	0.048	0.014
	β	ACE	0.001	0.028	0.028	0.001	0.952
		MCMI	-0.001	0.045	0.046	0.002	0.950
	σ^2	ACE	-0.003	0.040	0.040	0.002	0.940
		MCMI	3.356		3.426	22.989	

表1：插补模型正确设定时的模拟结果。结果展示了三种方法：ACE 插补、MCMI（多重条件均值插补）和 Oracle估计量。我们报告了经验偏差、平均标准误估计（SEE）、经验标准误（ESE）、平均均方误差（MSE），以及名义覆盖率为95%的 Wald型置信区间的观测覆盖概率（CPr）。

5 建模亨廷顿病的进展

亨廷顿病是一种致命的神经退行性疾病，会导致运动、认知和精神病学领域的功能受损。该病由重复的胞嘧啶-

删失	参数	方法	Bias	SEE	ESE	MSE	覆盖率
None	α	先知	-0.000	0.003	0.003	0.000	0.954
	β	先知	0.001	0.022	0.022	0.000	0.953
	σ^2	先知	-0.001		0.032	0.001	
中等	α	ACE	0.000	0.026	0.026	0.001	0.948
($\approx 50\%$)		MCMI	1.591	0.264	6.569	45.639	0.005
	β	ACE	0.001	0.026	0.026	0.001	0.949
		MCMI	-0.069	0.327	1.137	1.296	0.955
	σ^2	ACE	-0.002	0.037	0.038	0.001	0.943
		MCMI	≥ 100		≥ 100	≥ 100	
重度	α	ACE	0.000	0.029	0.029	0.001	0.946
($\approx 75\%$)		MCMI	0.716	0.326	3.085	10.018	0.188
	β	ACE	0.000	0.029	0.029	0.001	0.948
		MCMI	-0.073	0.332	1.263	1.598	0.949
	σ^2	ACE	-0.001	0.041	0.040	0.002	0.951
		MCMI	≥ 100		≥ 100	≥ 100	

表2：插补模型设定错误时的模拟结果。结果展示了三种方法：ACE 插补、MCMI（多重条件均值插补）和 Oracle估计量。我们报告了经验偏差、平均标准误估计（SEE）、经验标准误（ESE）、平均均方误差（MSE），以及名义覆盖率95%的 Wald型置信区间的观测覆盖概率（CPr）。

亨廷顿蛋白基因中的腺嘌呤-鸟嘌呤（CAG）突变，以及携带超过36个CAG重复序列的人，注定会患上亨廷顿病（McColgan 和 Tabrizi，2018年）。这为研究该疾病的临床医生创造了独特的机会；通过基因检测，可以招募注定会患上亨廷顿病的受试者（“基因

突变携带者”），并检验旨在延缓病情进展的疗法。

规划亨廷顿病临床试验时，一个关键考量是选择随时间变化足够快的结局，以便分析人员更容易检测治疗组与安慰剂组之间的差异。为帮助进行这项选择，分析人员一直在比较各种可能结局在基因突变携带者中的变化速度（Langbehn and Hersch，2020年；Paulsen 等人，2014年）。为量化这些结局的变化速度，分析人员利用未接受治疗的受试者数据，为这些结局估计纵向模型；这些“疾病进展模型”随后会给出一个指标，反映可能的结局在未接受治疗的受试者中随时间进展的速度。因此，通过比较估计得到的疾病进展模型，我们可以轻松比较这些结局的进展速度。

为推动这项持续进行的研究，我们分析了 PREDICT-HD 的数据。PREDICT-HD 是一项针对基因突变携带者的纵向研究，这些受试者在确诊前招募（Paulsen 等人，2008年）。具体而言，我们评估了用于量化亨廷顿病症状的七种潜在结局：(1) 总运动评分（TMS）；(2) 符号数字模式测验评分；(3–5) Stroop词语、颜色和干扰测验的得分；(6) 总功能能力（TFC）；以及 (7) 亨廷顿病综合统一评定量表（cUHDRS）。TMS 用于评估受试者的运动障碍。符号数字模式测验和 Stroop 测验用于评估认知障碍程度。TFC 用于评估受试者执行日常任务的功能能力。cUHDRS 是由 TFC、TMS、符号数字模式测验评分和 Stroop词语测试得分构成的线性组合（Schobel 等人，2017年）。

对于每个结局 Y_{ij}，我们拟合以下线性混合效应模型：

$$Y_{ij} = \beta_0 + \boldsymbol{\beta}^{\mathrm{T}} \boldsymbol{Z}_{ij} + \alpha(s_{ij} - X_i) + b_i + \epsilon_{ij}, \tag{5}$$

具有固定截距 β_0、固定斜率（$\boldsymbol{\beta}^{\mathrm{T}}$、 α）和随机截距 b_i。假设随机截距 b_i 服从分布 f_b，但不对该分布作具体规定，且

假设随机误差 ϵ_{ij} 服从均值为 0、方差为 σ^2的正态分布。对于每位受试者 i，模型纳入其他协变量 $\boldsymbol{Z}_{ij} =$ （年龄$_i$，性别$_i$，受教育年限$_i$，CAG$_i$），这些变量均在基线测量。年龄$_i$ 指基线年龄；性别$_i$ 表示受试者 i 为男性时取 1、为女性时取 0；受教育年限$_i$ 指受教育年数；CAG$_i$ 指 CAG 重复次数（Paulsen 等人，2014年）。变量 s_{ij} 和 X_i 分别表示研究时间（年）和临床诊断时间（年），因此 $s_{ij} - X_i$ 表示距临床诊断时间。如第 1 节所述，我们选择使用距临床诊断时间（$s_{ij} - X_i$），而不只是研究时间（s_{ij}），以考虑到受试者进入研究时所处的疾病进展阶段不同这一事实。

为了比较基因突变携带者的潜在结局如何进展，我们比较了针对每种结局（复合统一亨廷顿病评定量表、总运动评分、符号数字模式测验评分、总功能能力和 Stroop测验评分）拟合方程（5）中的模型时，距临床诊断时间的估计斜率（即 $\hat{\alpha}$）。根据“标准化斜率”对结局进行排名，该斜率的计算方法是将斜率的绝对值除以其估计标准误（即 $|\hat{\alpha}|/\text{SE}(\hat{\alpha})$）。

分析前，我们采用与 Long 等人（2017年）相似的排除标准。纳入的受试者必须是亨廷顿病携带者（即具有 ≥ 36 个 CAG 重复次数），且在研究入组时尚未获得临床诊断（即首次访视时的诊断置信度等级为 ≤ 3 ）。我们还筛选出仅包含所有感兴趣的结局和协变量 Z_{ij} 均有完整数据的访视，这仅排除了 25 名受试者（2%）和 260 次访视（4%）。按照这些标准，我们得到一个分析数据集，其中包含 1,102 名不同的受试者，共有 5,612 个观测值。附录H中的表6提供了描述性统计。

5.1 插补删失的临床诊断时间

在 1,102 名受试者中，244 人（22.1%）在研究期间获得了临床诊断，858 人（77.9%）未获得诊断（即其诊断时间为右删失）。我们使用基于 Cox 模型的条件均值插补（第 2.2 节）对所有删失的诊断时间进行插补。

我们将 CAG-年龄乘积、TMS 和符号数字模式测验评分作为 Cox模型协变量。CAG-年龄乘积定义为 $X_{i,\ \text{age}} \times (X_{i,\ \text{CAG}} - 34)$（Zhang 等人，2011年），表示受试者一生中累积的“疾病负担”。Long 等人（2017年）表明，与其他备选协变量组合相比，这三个变量对临床诊断时间的预测效果更好；我们认为，这些变量也是预测临床诊断时间的直观选择。已有充分证据表明，CAG 重复次数与临床诊断时间呈负相关（Long 等人，2017年）。此外，TMS 和符号数字模式测验评分分别衡量受试者的运动能力和认知能力，而众所周知，这两者都会随着受试者接近临床诊断时间而下降。使用这些协变量，我们拟合了一个风险函数为 $\lambda(t|\boldsymbol{V}_i) = \lambda_0(t)\exp(\boldsymbol{\gamma}\boldsymbol{V}_i)$ 的 Cox模型，其中协变量 $\boldsymbol{V}_i$包括受试者 i 在$\boldsymbol{\gamma} = (\gamma_1,\ \gamma_2,\ \gamma_3)^{\mathrm{T}}$ 时第1次访视的 TMS、符号数字模式测验评分和 CAG-年龄乘积值，相应的对数风险比为这些值。拟合该模型之前，我们使用协变量的样本均值和标准差（见附录H中的表6）对其进行中心化和标准化，以减少可能的共线性。利用该 Cox模型的估计值，$\widehat{\ }$我们实施条件均值插补，将 X_i 的值替换为 $X_i = \Delta_i X_i + (1 - \Delta_i)\,\mathrm{E}(X_i | X_i > C_i,\ \boldsymbol{V}_i)$。

5.2 亨廷顿病损伤的排名指标

使用该插补数据集，我们针对七项结局指标分别估计方程（5）中的线性混合效应模型。我们计算每项潜在结局随时间变化的速度

，使用标准化斜率 $|\hat{\alpha}|/\mathrm{SE}(\hat{\alpha})$ 并对其进行排名。我们针对三种方法分别进行一次排名：i）完整病例分析（CCA），仅对未删失受试者应用限制性最大似然；ii）多重条件均值插补（见第 4.2 节）；以及 iii）ACE 插补。表3展示了这些排名结果。三种方法均认为，复合统一亨廷顿病评定量表的进展速度最快。这一发现与 Langbehn 和 Hersch（2020）的结论相反；他们指出，在所分析的结局中，总运动评分变化最快，因此是首选结局。

TMS仅衡量运动障碍，而复合统一亨廷顿病评定量表是总功能能力、总运动评分、符号数字模式测验评分和Stroop词语测验得分的线性组合；因此，复合统一亨廷顿病评定量表能够更全面地衡量亨廷顿病所致的障碍。因此，确定一种能够改善受试者复合统一亨廷顿病评定量表评分的治疗，可能会更全面地改善其生活质量。这种全面改善是复合统一亨廷顿病评定量表被倡导作为针对已确诊受试者的亨廷顿病临床试验主要结局的主要原因。由于我们的分析仅使用了基线时尚未确诊患者的数据，结果表明，对于在这一“症状前”阶段招募的患者，复合统一亨廷顿病评定量表也可作为一种易于检测的结局指标。

需要注意的是，由于 cUHDRS 是另外四项测量指标的线性组合，因此在给定数据集中，它可能更容易出现缺失。考虑到这一潜在问题，研究人员或许会考虑使用仅测量一种症状的结局来衡量亨廷顿病所致的障碍（例如，符号数字模式测验评分、TMS）。根据我们的结果，仅测量一种症状的结局中，最佳选择是符号数字模式测验评分。事实上，已有研究表明，认知障碍比运动障碍更早出现（Paulsen 等人，2008）。我们的分析与这一既往发现一致，而其他方法则指向运动障碍。

方法	结局	Rank	估计值	SE	标准化斜率	相对于最大值的比例
CCA	复合统一亨廷顿病评定量表	1	-0.506	0.011	44.179	1.000
	TMS	2	2.304	0.062	37.292	0.844
	Stroop词语	3	-2.106	0.070	30.118	0.682
	SDMT	4	-1.258	0.043	29.205	0.661
	Stroop颜色	5	-1.608	0.060	26.629	0.603
	TFC	6	-0.253	0.010	24.104	0.546
	Stroop干扰	7	-0.762	0.044	17.419	0.394
ACE	复合统一亨廷顿病评定量表	1	-0.253	0.013	20.064	1.000
	SDMT	2	-0.757	0.041	18.303	0.912
	Stroop词语	3	-1.177	0.070	16.883	0.841
	TMS	4	0.968	0.064	15.026	0.749
	Stroop颜色	5	-0.802	0.059	13.532	0.674
	TFC	6	-0.102	0.009	11.466	0.571
	Stroop干扰	7	-0.314	0.037	8.489	0.423
MCMI	复合统一亨廷顿病评定量表	1	-0.250	0.005	45.714	1.000
	TMS	2	1.068	0.024	44.250	0.968
	Stroop词语	3	-1.170	0.039	29.876	0.654
	SDMT	4	-0.739	0.026	28.226	0.617
	TFC	5	-0.109	0.004	27.069	0.592
	Stroop颜色	6	-0.807	0.033	24.189	0.529
	Stroop干扰	7	-0.338	0.025	13.613	0.298

表3：CCA：完整病例分析。ACE：ACE 插补。MCMI：多重条件均值插补。cUHDRS：亨廷顿病综合统一评定量表。TMS：总运动评分。SDMT：符号数字模式测试。TFC：总功能能力。标准化斜率 $= |\hat{\alpha}/\text{SE}(\hat{\alpha})|$，用于衡量直至临床诊断期间结局随时间变化的速度。排名：按同一种方法估计的标准化斜率排序。相对于最大值的比例：相对于同一种方法估计的最大标准化斜率的比例。

5.3 计算临床试验的样本量

接下来，我们计算一项假设性亨廷顿病临床试验所需的样本量。然后，我们比较基于 ACE 插补、条件均值插补和完整病例分析得出的样本量估计值。我们给出以符号数字模式测验评分为主要结局的临床试验中，检测治疗效应所需的样本量。为简便起见，我们假设采用平衡设计，即治疗组和安慰剂组的样本量相同。为考察治疗效应，我们分别使用以下线性模型比较安慰剂组（下标为 p）与治疗组（下标为 t）的符号数字模式测验评分变化轨迹：

$$\begin{aligned} Y_p &= \beta_p + \alpha_p(s - X) \\ Y_t &= \beta_t + \alpha_t(s - X), \end{aligned}$$

其中，β表示截距，α表示距临床诊断时间的斜率，$s - X$。利用这些模型，我们旨在对量 $\delta = \alpha_t - \alpha_p$进行推断。

具体而言，为了考察所研究的治疗是否减缓了认知障碍（以符号数字模式测验评分衡量），我们分别定义原假设H_0： $\delta = 0$ 和备择假设 H_A： $\delta > 0$。为了计算在目标检验效能下检验这些假设所需的样本量，我们从 δ的 Z统计量开始：$\hat{\delta}\sqrt{n_g}\{\text{Var}(\delta)\}^{-1}$。根据该统计量，在第一类错误率为 κ、检验效能为 $1 - \Gamma$ 时，每组所需的样本量为

$$\left[\frac{\sqrt{\text{Var}(\delta)}\{\Phi^{-1}(1-\Gamma) - \Phi^{-1}(\kappa)\}}{-d}\right]^2,$$

其中，$\Phi^{-1}(\cdot)$ 表示正态分布 Normal(0,1) 的逆累积分布函数，而 d 是假定的 δ 真值。为简化起见，我们假设 $\sqrt{\text{Var}(\delta)} = 1$。

假设我们特别关注的是：要以 80% 的检验效能检出 10% 的治疗效应，所需的样本量为 $\alpha_t = (1 - 0.1)\alpha_p$, $\kappa = 0.05$, 和 $1 - \Gamma = 0.8$。

借用第5.2节中 α 的模型估计值作为安慰剂组在 $s-X$ 上的斜率（即将这些估计值用作 α_p），我们可以根据每种方法计算所需的样本量。根据 ACE 插补得出的 $\hat{\alpha}_p = -0.757$，我们发现，为达到预期的检验效果，每组需要 1079 人。相比之下，根据完整病例分析进行计算，所需样本量更小（$\hat{\alpha}_p = -1.258$，需要 391 人）；而 MCMI 得出的所需样本量更大（$\hat{\alpha}_p = -0.739$，需要 1134 人）。

鉴于支持 ACE 插补的理论和实证证据，我们认为，对于所提出的检验，每组 1079 人是最可靠的样本量。因此，以此样本量为参照，完整病例分析得出的样本量会大幅降低我们检测 10% 治疗效应的检验效能。尽管 MCMI 得出的样本量与 ACE 插补得出的样本量更为接近，但考虑到亨廷顿病的患病率很低（每 100,000 人中有 12 人；Wexler等人，2016年），每组再招募 55 名受试者既会增加成本，也可能难以实现。附录H的图2给出了所有方法在治疗效应分别为 10%、15% 和 20% 时对应的功效曲线。

6 讨论

由于亨廷顿病会导致多方面的功能损害，临床医生在设计临床试验时必须从许多可能的结局中进行选择。为指导这一选择，分析人员一直试图根据基因突变携带者的结局随时间下降的速度，对可能的结局进行排名。只有建立准确的疾病进展模型，才能进行这种排名；而这依赖于了解受试者被临床诊断为亨廷顿病的时间。当临床诊断时间被删失时——这在亨廷顿病观察性研究中十分常见——我们可以对其进行插补，但此时结局模型估计值会受到插补模型设定是否正确的影响；当

我们错误设定插补模型时，可能会产生偏差并降低效率。

本文介绍了 ACE 插补：即使使用设定错误的插补模型，在给定插补协变量的情况下，该方法仍能准确估计纵向模型。我们证明，无需为随机效应 $\boldsymbol{b}$ 或插补误差 U设定分布，也可以计算该估计量。此外，即使不了解这些干扰分布的先验知识，该估计量仍能实现一致性和渐近正态性。无需估计或假定 $\boldsymbol{b}$ 或 U 的分布，这一点不同于典型的半参数方法；后者需要假定干扰分布，虽然能够保持一致性，但在所假定的分布不正确时会遭受效率损失。

尽管 ACE 插补产生的估计量具有可识别性、相合性和渐近正态性，但这种新方法并非没有局限。如附录 G 所示，当对于所有受试者而言，固定效应设计矩阵的一列属于随机效应设计矩阵的列空间时，所提出的估计方程可能无法收敛。例如，当我们在模型中同时纳入固定截距和随机截距时，就可能出现这种“列空间问题”。附录 G 展示了更为严谨的细节，但我们可以通过以下方式直观地理解这一问题：例如，当我们同时纳入固定截距和随机截距时，固定截距就是随机截距分布的均值。因此，当我们试图同时估计固定截距，并将随机截距的分布（固定截距即为该分布的均值）视为干扰参数时，就会产生冲突。基于这一理解，也许可以要求随机效应的干扰分布具有均值$\mathbf{0}$ ，从而将我们感兴趣的固定效应与干扰分布区分开来，以克服这一列空间问题。未来仍需进一步研究如何解决这一问题。

补充材料

补充材料包含我们的定理和命题的所有技术细节。开源 R软件包 ACEimpute 以及用于复现模拟研究的代码可在 https://github.com/Tanya-Garcia-Lab/ACEimpute 获取。

参考文献

Atem, F. D., Sampene, E., and Greene, T. J. (2019年). 改进用于随机删失预测变量的线性回归条件插补方法. 《医学研究统计方法》, 28(2):432–444.

Austin, P. C. and Brunner, L. J. (2003年). 存在天花板效应时的 I 类错误膨胀. 《美国统计学家》, 57(2):97–104.

Bates, D., M¨achler, M., Bolker, B., and Walker, S. (2015年). 使用 lme4 拟合线性混合效应模型. 《统计软件杂志》, 67(1):1–48.

Bender, R., Augustin, T., and Blettner, M. (2005年). 生成生存时间以模拟 Cox比例风险模型. 《医学统计学》, 24(11):1713–1723.

Bernhardt, P. W., Wang, H. J., and Zhang, D.（2014年）。通过多重插补，对受检测限约束的协变量进行灵活建模，以分析生存数据。《计算统计与数据分析》，69：81–91。

Black, A. C., Harel, O. 和 Betsy McCoach, D.（2011年）。多层数据的缺失数据处理方法：模型设定错误的影响。《应用统计学杂志》，38(9):1845–1865。

Breslow, N. E.（1972）。对 Cox 博士论文的讨论意见。皇家统计学会期刊，*B* 辑，34:216–217。

Carroll, R., Ruppert, D., Stefanski, L. 和 Crainiceanu, C.（2006）。非线性模型中的测量误差：现代观点。伦敦：CRC 出版社，第 2 版。

Cole, S. R., Chu, H. 和 Greenland, S.（2006）。测量误差校正的多重插补。国际流行病学杂志，35(4):1074–1081。

Fitzmaurice, G. M., Laird, N. M. 和 Ware, J. H.（2011）。应用纵向分析。波士顿：威利出版社。

Foutz, R. V.（1977）。关于似然方程唯一相合解的研究。《美国统计学会杂志》，72(357):147–148。

Garcia, T.P. 和 Ma, Y.（2016）。无分布假设的随机截距逻辑回归模型的最优估计量。斯堪的纳维亚统计学杂志，43(1):156–171。

Haber, G., Sampson, J. 和 Graubard, B.（2020年）。Berkson 误差导致的偏差：用预测值代替观测协变量时的问题。生物统计学。

Kieburtz, K., Feigin, A., McDermott, M., Como, P., Abwender, D., Zimmerman, C., Hickey, C., Orme, C., Claude, K., Sotack, J. 等（1996）。盐酸瑞马西胺治疗亨廷顿病的对照试验。《运动障碍》：运动障碍学会官方期刊，11(3)：273–277。

Langbehn, D. R. 和 Hersch, S.（2020年）。亨廷顿病前驱期试验的临床结局和筛选标准。《运动障碍》，35(12):2193–2200。

Little, R. J. 和 Rubin, D. B.（2019年）。《缺失数据统计分析》，第793卷。John Wiley &Sons。

Long, J. D., Langbehn, D. R., Tabrizi, S. J., Landwehrmeyer, B. G., Paulsen, J. S., Warner, J., 和 Sampaio, C.（2017年）。亨廷顿病预后指数的验证。《运动障碍》，32(2):256–263。

Long, J. D., Paulsen, J. S., Marder, K., Zhang, Y., Kim, J.-I., Mills, J. A., 以及 PREDICT-HD 亨廷顿病研究组的 R.（2014年）。追踪亨廷顿病进展中的运动障碍。《运动障碍》，29(3):311–319。

Lotspeich, S. C., Grosser, K. F., 和 Garcia, T. P.（2022年）。校正删失协变量的条件均值插补并提高可用性。《生物计量学杂志》。

McColgan, P. 和 Tabrizi, S. J.（2018年）。亨廷顿病：临床综述。《欧洲神经病学杂志》，25(1):24–34。

Paulsen, J. S., Langbehn, D. R., Stout, J. C., Aylward, E., Ross, C. A., Nance, M., Guttman, M., Johnson, S., MacDonald, M., Beglinger, L. J., 等（2008）。亨廷顿病在确诊数十年前的检出：PREDICT-HD 研究。《神经病学、神经外科学与精神病学杂志》，79(8):874–880。

Paulsen, J. S., Long, J. D., Johnson, H. J., Aylward, E. H., Ross, C. A., Williams, J. K., Nance, M. A., Erwin, C. J., Westervelt, H. J., Harrington, D. L., 等（2014年）。前驱期亨廷顿病的临床和生物标志物变化显示试验可行性：PREDICT-HD 研究十年。《衰老神经科学前沿》，6:78。

R 核心团队（2022年）。*R*：统计计算语言与环境。R 计算基金会，奥地利维也纳。

Saul, B. C. 和 Hudgens, M. G.（2020年）。使用 geex 在 R 中进行 m估计的微积分。《统计软件杂志》，92(2)：1–15。

Schobel, S. A., Palermo, G., Auinger, P., Long, J. D., Ma, S., Khwaja, O. S., Trundell, D., Cudkowicz, M., Hersch, S., Sampaio, C., 等（2017年）。早期亨廷顿病中的运动、认知和功能衰退共同构成一个单一的进展因子。《神经病学》，89(24):2495– 2502。

Therneau, T. M.（2022年）。用于*R*生存分析的软件包。R软件包版本 3.3-1。

Wang, H. J. 和 Feng, X.（2012年）。针对含删失协变量的m回归进行多重插补。《美国统计学会杂志》，107(497):194–204。

Wei, R., Wang, J., Jia, E., Chen, T., Ni, Y. 和 Jia, W.（2018年）。GSimp：一种基于吉布斯采样器的左删失缺失值插补方法，适用于代谢组学研究。《公共科学图书馆计算生物学》，14(1):1–14。

Wexler, A., Wild, E. J. 和 Tabrizi, S. J.（2016年）。George Huntington：探究、同理心与希望的传承。《大脑》，139(8)：2326–2333。

Yu, T., Xiang, L. 和 Wang, H. J.（2021年）。针对受检测限约束的协变量生存数据的分位数回归。《生物统计学》，77(2)：610–621。

Yucel, R. M. 和 Demirtas, H.（2010年）。非正态随机效应对多重插补推断的影响：模拟评估。《计算统计学与数据分析》，54(3)：790–801。

Zhang, Y., Long, J. D., Mills, J. A., Warner, J. H., Lu, W., Paulsen, J. S., P.-H. 研究者，以及亨廷顿研究组的 C.（2011年）。对疾病进展进行指标化，适用于

研究入组时有亨廷顿病患病风险的个体。《美国医学遗传学杂志*B*辑：神经精神遗传学》，156(7)：751–763。

附录

A 含干扰参数时估计 θ 的有效得分向量

为应用 Garcia 和 Ma（2016年）的半参数框架，注意到模型中第 i个受试者的密度为：

$$f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\text{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})=\int f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{b}_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{b}_i;\boldsymbol{\theta})f_{\boldsymbol{D}_i,\boldsymbol{b}_i,\text{mis}}(\boldsymbol{D}_i,\boldsymbol{b}_i)d\mu(\boldsymbol{b}_i)$$

若 $\Delta_i=1$ 且

$$f_{\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i,\text{mis}}(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})=\int f_{\boldsymbol{Y}_i|\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i}(\boldsymbol{Y}_i|\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i;\boldsymbol{\theta})f_{\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i,\text{mis}}(\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i)d\mu(\boldsymbol{b}_i,U_i)$$

若 $\Delta_i=0$。这里，下标 $_{\text{mis}}$ 用于表示由于 $(\boldsymbol{D}_i,\boldsymbol{b}_i)$（或 $(\boldsymbol{D}_i,\boldsymbol{b}_i,U_i)$）的联合密度存在模型设定错误而导致模型设定错误的项。然后，采用 Garcia 和 Ma（2016年）中的类似方法，可得未删失受试者的有效得分向量为：

$$\boldsymbol{S}_{\text{eff},\text{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})=\boldsymbol{S}_{\boldsymbol{\theta},\text{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})-\text{E}_{\text{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\text{i}},\boldsymbol{b}_{\text{i}})|\boldsymbol{Y}_{\text{i}},\boldsymbol{D}_{\text{i}}\}\tag{6}$$

其中 $\boldsymbol{S}_{\boldsymbol{\theta},\text{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})=\frac{\partial}{\partial\boldsymbol{\theta}}\log\{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\text{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})\}$。这里，$\boldsymbol{h}(\boldsymbol{D}_i,\boldsymbol{b}_i)$ 是一个均值为0 且满足以下条件的函数

$$\text{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\text{mis}}(\boldsymbol{Y}_{\text{i}},\boldsymbol{D}_{\text{i}};\boldsymbol{\theta})|\boldsymbol{D}_{\text{i}},\boldsymbol{b}_{\text{i}}\}=\text{E}[\text{E}_{\text{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\text{i}},\boldsymbol{b}_{\text{i}})|\boldsymbol{Y}_{\text{i}},\boldsymbol{D}_{\text{i}}\}|\boldsymbol{D}_{\text{i}},\boldsymbol{b}_{\text{i}}].$$

类似地，删失受试者的有效得分向量为

$$\boldsymbol{S}^{*}_{\mathrm{eff,mis}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) = \boldsymbol{S}^{*}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) - \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\widehat{\boldsymbol{D}}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}, \mathrm{U_i})|\boldsymbol{Y}_{\mathrm{i}}, \widehat{\boldsymbol{D}}_{\mathrm{i}}\} \tag{7}$$

其中 $\boldsymbol{S}^{*}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta}) = \frac{\partial}{\partial \boldsymbol{\theta}} \log\{f_{\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i, \mathrm{mis}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})\}$，而 $\boldsymbol{h}(\widehat{\boldsymbol{D}}_i, \boldsymbol{b}_i, U_i)$ 是一个均值为0的函数，并满足

$$\mathrm{E}\{\boldsymbol{S}^{*}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}}, \widehat{\boldsymbol{D}}_{\mathrm{i}}; \boldsymbol{\theta})|\widehat{\boldsymbol{D}}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}, \mathrm{U_i}\} = \mathrm{E}[\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\widehat{\boldsymbol{D}}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}, \mathrm{U_i})|\boldsymbol{Y}_{\mathrm{i}}, \widehat{\boldsymbol{D}}_{\mathrm{i}}\}|\widehat{\boldsymbol{D}}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}, \mathrm{U_i}].$$

然而，正如 Garcia 和 Ma（2016年）所指出的，要找到满足这些条件的函数 $\boldsymbol{h}(\boldsymbol{D}_i, \boldsymbol{b}_i)$ 和 $\boldsymbol{h}(\boldsymbol{D}_i, \boldsymbol{b}_i, U_i)$ 并非易事。在 B节中，我们简化了公式（6）和公式（7）给出的有效得分向量。我们利用多元正态分布和线性投影理论的性质实现了这一简化。如此一来，我们完全无需使用函数 $\boldsymbol{h}(\boldsymbol{D}_i, \boldsymbol{b}_i)$ 和 $\boldsymbol{h}(\boldsymbol{D}_i, \boldsymbol{b}_i, U_i)$，并给出了两个有效得分向量的简单闭式。此外，我们在附录C中表明，无需指定 $(\boldsymbol{D}, \boldsymbol{b}_i)$ 或 $(\boldsymbol{D}, \boldsymbol{b}_i, U_i)$ 的联合分布，也能计算这些有效得分向量；这一点影响重大，因为如果必须指定这些联合分布，就有可能将其设定错误，进而降低 $\boldsymbol{\theta}_{\mathrm{ACE}}$ 的整体效率； $\boldsymbol{\theta}_{\mathrm{ACE}}$ 是我们针对 $\boldsymbol{\theta}$ 的估计量。

B 简化有效得分向量

B.1 变换 未删失受试者响应向量的变换

为简化未删失受试者的有效得分向量，我们将：1）对删失受试者的 $\boldsymbol{Y}_i$ 进行变换；2）确立关于给定该变换及 $(\boldsymbol{D}_i, \boldsymbol{b}_i)$ 时 $\boldsymbol{Y}_i$ 的条件分布的五个性质；然后 3）利用这些性质简化上文公式（6）中给出的复杂有效得分向量形式。具体而言，我们

通过 $\boldsymbol{T}_i = \sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$ 变换 $\boldsymbol{Y}_i$。在命题 2 中，我们概述了关于给定 $(\boldsymbol{T}_i, \boldsymbol{D}_i, \boldsymbol{b}_i)$ 时 $\boldsymbol{Y}_i$ 的条件分布的五个关键性质:

命题 2 考虑记为 $\boldsymbol{T}_i = \sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$ 的变换后的响应变量。如果随机效应设计矩阵 $\boldsymbol{Z}_i^b$ 满列秩（即，如果 $\mathrm{rank}(\boldsymbol{Z}_i^b) = p_b$），那么，

(A) $\mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}_i, \boldsymbol{D}_i, \boldsymbol{b}_i) = \boldsymbol{\zeta}(\boldsymbol{D}_i; \boldsymbol{\theta}) + \boldsymbol{P}_{\boldsymbol{Z}_i^b}\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\boldsymbol{D}_i; \boldsymbol{\theta})\}$，其中 $\boldsymbol{P}_{\boldsymbol{M}} = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T$是矩阵 $\boldsymbol{M}$ 的正交投影算子。

(B) 在给定 $(\boldsymbol{T}_i, \boldsymbol{D}_i)$ 的条件下， $\boldsymbol{Y}_i$ 和 $\boldsymbol{b}_i$ 相互独立。

(C) $\mathrm{E}(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i, \boldsymbol{D}_i, \boldsymbol{b}_i) = \sigma^2(m_i - p_b) + ||\mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}_i, \boldsymbol{D}_i, \boldsymbol{b}_i)||_2^2$，其中 $||\cdot||_2^2$ 表示 *L2*范数。

(D) 对于任意向量值函数 $\boldsymbol{g}(\boldsymbol{T}_i, \boldsymbol{D}_i)$，如果 $\mathrm{E}\{\boldsymbol{g}(\boldsymbol{T}_i, \boldsymbol{D}_i)|\boldsymbol{D}_i, \boldsymbol{b}_i\} = \boldsymbol{0}$，则可得 $\boldsymbol{g}(\boldsymbol{T}_i, \boldsymbol{D}_i) = \boldsymbol{0}$。

(E) 给定随机变量 $\boldsymbol{R}$， $\boldsymbol{R}$ 与 $\boldsymbol{T}_i = \sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$ 在给定$(\boldsymbol{Y}_i, \boldsymbol{Z}_i^b)$的条件下相互独立。

为证明命题 2，我们首先回到所关注的结果模型 :

$$Y_{ij} = \boldsymbol{\beta}^T\boldsymbol{Z}_{ij}^a + \alpha(s_{ij} - X_i) + \boldsymbol{b}_i^T\boldsymbol{Z}_{ij}^b + \epsilon_{ij}$$
$$\boldsymbol{b}_i \sim f_{\boldsymbol{b}}, \quad \epsilon_{ij} \sim \mathrm{Normal}(0, \sigma^2).$$

由该模型可知，给定（$\boldsymbol{Z}_{ij}^a$， s_{ij}， X_i， $\boldsymbol{Z}_{ij}^b$， $\boldsymbol{b}_i$）时，响应 $\boldsymbol{Y}_{ij}$服从一元正态分布 $\mathrm{Normal}\{\boldsymbol{\beta}^T\boldsymbol{Z}_{ij}^a + \alpha(S_{ij} - X_i) + \boldsymbol{Z}_{ij}^b\boldsymbol{b}_i, \sigma^2\}$。回想一下，我们假设给定第 i个受试者的随机效应 $\boldsymbol{b}_i$后，该受试者的响应相互独立。根据这一事实以及 Y_{ij}的条件分布，我们知道，给定（$\boldsymbol{D}_i$， $\boldsymbol{b}_i$）时， m维响应向量 $\boldsymbol{Y}_i$ 服从多元

正态分布 $\text{Normal}_{m_i}\{\boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha(\boldsymbol{S}_i-X_i\mathbf{1}_{m_i})+\boldsymbol{Z}_i^b\boldsymbol{b}_i,\ \sigma^2\boldsymbol{I}_{m_i}\}$，其中 $\mathbf{1}_{m_i}$表示一个m_i维全 1 向量，$\boldsymbol{I}_{m_i}$ 表示 $m_i\times m_i$ 单位矩阵。为简化记号，令 $\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})=\boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha(\boldsymbol{S}_i-X_i\mathbf{1}_{m_i})$。我们将其视为 $\boldsymbol{Y}_i$的线性预测值中的固定效应分量。

接下来，我们推导给定 $(\boldsymbol{D}_i,\boldsymbol{b}_i)$ 时 $(\boldsymbol{Y}_i,\boldsymbol{T}_i)$ 的联合条件分布。根据我们对 $\boldsymbol{T}_i=\sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$ 的定义，可得 $(\boldsymbol{Y}_i^T,\boldsymbol{T}_i^T)^T=(\boldsymbol{I}_{m_i},\sigma^{-2}\boldsymbol{Z}_i^b)^T\boldsymbol{Y}_i$。由此可知，在给定 $(\boldsymbol{D}_i,\boldsymbol{b}_i)$ 的条件下，$(\boldsymbol{Y}_i,\boldsymbol{T}_i)\sim$正态分布$_{m_i}[(\boldsymbol{I}_{m_i},\sigma^{-2}\boldsymbol{Z}_i^b)^T\{\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i\},\sigma^2(\boldsymbol{I}_{m_i},\sigma^{-2}\boldsymbol{Z}_i^b)^T(\boldsymbol{I}_{m_i},\sigma^{-2}\boldsymbol{Z}_i^b)]$。具体而言，

$$\begin{aligned}\begin{pmatrix}\boldsymbol{I}_{m_i}\\ \sigma^{-2}\boldsymbol{Z}_i^{bT}\end{pmatrix}\{\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i\} &= \begin{pmatrix}\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i\\ \sigma^{-2}\boldsymbol{Z}_i^{bT}\{\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i\}\end{pmatrix}\\ \sigma^2\begin{pmatrix}\boldsymbol{I}_{m_i}\\ \sigma^{-2}\boldsymbol{Z}_i^{bT}\end{pmatrix}\begin{pmatrix}\boldsymbol{I}_{m_i} & \sigma^{-2}\boldsymbol{Z}_i^b\end{pmatrix} &= \sigma^2\begin{pmatrix}\boldsymbol{I}_{m_i} & \sigma^{-2}\boldsymbol{Z}_i^b\\ \sigma^{-2}\boldsymbol{Z}_i^{bT} & \sigma^{-4}\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b\end{pmatrix}.\end{aligned}$$

接下来，我们建立给定 $(\boldsymbol{T}_i,\boldsymbol{D}_i,\boldsymbol{b}_i)$ 时 $\boldsymbol{Y}_i$ 的条件分布。根据前面的结果，我们知道，给定 $(\boldsymbol{T}_i,\boldsymbol{D}_i,\boldsymbol{b}_i)$ 时，$\boldsymbol{Y}_i$ 也服从正态分布，其

$$\begin{aligned}\mathrm{E}(\boldsymbol{Y}_\mathrm{i}|\boldsymbol{T}_\mathrm{i},\boldsymbol{D}_\mathrm{i},\boldsymbol{b}_\mathrm{i}) &= \boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i+\sigma^2\boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}[\boldsymbol{T}_i-\sigma^{-2}\boldsymbol{Z}_i^{bT}\{\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i\}]\\ &= \sigma^2\boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\boldsymbol{T}_i+(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\{\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i\},\end{aligned}$$

其中，$\boldsymbol{P}_{\boldsymbol{Z}_i^b}=\boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\boldsymbol{Z}_i^{bT}$是投影到 $\boldsymbol{Z}_i^b$列空间上的正交投影算子。利用这一点，我们注意到$(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{Z}_i^b\boldsymbol{b}_i=0$，因为 $\boldsymbol{P}_{\boldsymbol{Z}_i^b}\boldsymbol{Z}_i^b=\boldsymbol{Z}_i^b$，且根据我们的假设，$\boldsymbol{Z}_i^b$ 满列秩。因此，我们得出

$$\begin{aligned}\mathrm{E}(\boldsymbol{Y}_\mathrm{i}|\boldsymbol{T}_\mathrm{i},\boldsymbol{D}_\mathrm{i},\boldsymbol{b}_\mathrm{i}) &= \sigma^2\boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\boldsymbol{T}_i+(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\\ &= \boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i+(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\\ &= \boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{P}_{\boldsymbol{Z}_i^b}\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\},\end{aligned}$$

由此，命题 2 (A) 得证。接下来，注意到

$$\mathrm{Cov}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}}) \;=\; \sigma^2\boldsymbol{I}_{m_i}-\sigma^2\boldsymbol{Z}_i^b(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\boldsymbol{Z}_i^{bT}=\sigma^2(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b}).$$

总之，

$$\boldsymbol{Y}_i|(\boldsymbol{T}_i,\boldsymbol{D}_i)\sim\mathrm{Normal}_{m_i}[\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{P}_{\boldsymbol{Z}_i^b}\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\},\sigma^2(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{\boldsymbol{Z}_i^b})].$$

注意，该分布不涉及随机效应 $\boldsymbol{b}_i$。因此我们得出结论——由于多元正态分布完全由其均值和协方差确定——给定 $(\boldsymbol{T}_i,\ \boldsymbol{D}_i)$ 时，$\boldsymbol{Y}_i$ 和 $\boldsymbol{b}_i$ 条件独立。这证实了命题 2 (B)。

接下来，我们注意到，对于一个均值为 $\boldsymbol{\mu}$、方差-协方差矩阵为 $\boldsymbol{\Sigma}$ 的随机向量 $\boldsymbol{Y}$，$E(\boldsymbol{Y}^T\boldsymbol{Y})=\boldsymbol{\mu}^T\boldsymbol{\mu}+tr(\boldsymbol{\Sigma})$。据此，根据 $\boldsymbol{Y}_i$ 在给定 $(\boldsymbol{T}_i,\ \boldsymbol{D}_i,\ \boldsymbol{b}_i)$ 条件下的条件分布，可得

$$\begin{aligned}\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}}) &= \mathrm{tr}\{\sigma^2(\boldsymbol{I}_{\mathrm{m_i}}-\boldsymbol{P}_{\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{b}}})\}+||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})||_2^2\\ &= \sigma^2\mathrm{tr}(\boldsymbol{I}_{\mathrm{m_i}}-\boldsymbol{P}_{\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{b}}})+||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})||_2^2\\ &= \sigma^2(m_i-p_b)+||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})||_2^2,\end{aligned}$$

因为 $\boldsymbol{I}_m-\boldsymbol{P}$ 本身是一个秩为 $m-r$ 的正交投影算子，而正交投影算子的迹等于其秩。这就证明了命题 2 (C)。

命题 2 (D) 可由以下计算得出，其中给定任意函数$\boldsymbol{g}(\boldsymbol{T}_i,\ \boldsymbol{D}_i)$，使得 $\mathrm{E}\{\boldsymbol{g}(\boldsymbol{T}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}})|\boldsymbol{b}_{\mathrm{i}},\ \boldsymbol{D}_{\mathrm{i}}\}=\mathbf{0}$:

$$\begin{aligned}&\mathbf{0}=\mathrm{E}\{\boldsymbol{g}(\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})|\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}}\}\\ =\;&\int\boldsymbol{g}(\boldsymbol{T}_i,\boldsymbol{D}_i)f(\boldsymbol{T}_i|\boldsymbol{D}_i,\boldsymbol{b}_i)d\mu(\boldsymbol{T}_i)\\ =\;&\int\boldsymbol{g}(\boldsymbol{T}_i,\boldsymbol{D}_i)\exp\Big[-\frac{\sigma^4}{2}\{\boldsymbol{T}_i-\sigma^{-2}\boldsymbol{Z}_i^{bT}(\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i)\}^T\times\\ &(\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\{\boldsymbol{T}_i-\sigma^{-2}\boldsymbol{Z}_i^{bT}(\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i)\}\Big]d\mu(\boldsymbol{T}_i),\end{aligned}$$

由此可得　　根据公式（8）中 $(\boldsymbol{Y}_i,\ \boldsymbol{T}_i)$ 的条件分布。接下来，定义

$$k(\boldsymbol{T}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) \ = \ \exp\Big[-\frac{\sigma^4}{2}\{\boldsymbol{T}_i-\sigma^{-2}\boldsymbol{Z}_i^{bT}(\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i)\}^T\times (\boldsymbol{Z}_i^{bT}\boldsymbol{Z}_i^b)^{-1}\{\boldsymbol{T}_i-\sigma^{-2}\boldsymbol{Z}_i^{bT}(\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{Z}_i^b\boldsymbol{b}_i)\}\Big]d\mu(\boldsymbol{T}_i),$$

它对所有（$\boldsymbol{T}_i,\boldsymbol{D}_i,\boldsymbol{\theta}$）均为正。因此，事实是

$$\mathbf{0}=\int k(\boldsymbol{T}_i,\boldsymbol{D}_i;\boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{T}_i,\boldsymbol{D}_i)d\mu(\boldsymbol{T}_i)$$

对所有 $\boldsymbol{\theta}$，都有 $k(\boldsymbol{T}_i,\boldsymbol{D}_i;\ \boldsymbol{\theta})\boldsymbol{g}(\boldsymbol{T}_i,\boldsymbol{D}_i)\ =\mathbf{0}$。然而，由于 $k(\boldsymbol{T}_i,\boldsymbol{D}_i;\ \boldsymbol{\theta})$ 对所有（$\boldsymbol{T}_i,\boldsymbol{D}_i,\boldsymbol{\theta}$）均为正，只有当 $\boldsymbol{g}(\boldsymbol{T}_i,\boldsymbol{D}_i)\ =\mathbf{0}$ 时，这才成立。由此得证命题 2 (D)。命题 2 (E) 由引理 1 得出：

引理 1 给定随机变量 $\boldsymbol{R}_1$、 $\boldsymbol{R}_2$ 和 $\boldsymbol{R}_3$，以及向量值函数 $\boldsymbol{g}(\cdot)$，在给定$(\boldsymbol{R}_2,\boldsymbol{R}_3)$的条件下，$\boldsymbol{R}_1$与 $\boldsymbol{g}(\boldsymbol{R}_2,\boldsymbol{R}_3)$ 相互独立。

为证明引理 1，注意：

$$\begin{aligned}&P\{\boldsymbol{R}_1\le\boldsymbol{r}_1,\boldsymbol{g}(\boldsymbol{R}_2,\boldsymbol{R}_3)\le\boldsymbol{k}|\boldsymbol{R}_2=\boldsymbol{r}_2,\boldsymbol{R}_3=\boldsymbol{r}_3\}\\ =\ &P(\boldsymbol{R}_1\le\boldsymbol{r}_1|\boldsymbol{R}_2=\boldsymbol{r}_2,\boldsymbol{R}_3=\boldsymbol{r}_3)I\{\boldsymbol{g}(\boldsymbol{r}_2,\boldsymbol{r}_3)\le\boldsymbol{k}\}\\ =\ &P(\boldsymbol{R}_1\le\boldsymbol{r}_1|\boldsymbol{R}_2=\boldsymbol{r}_2,\boldsymbol{R}_3=\boldsymbol{r}_3)P\{\boldsymbol{g}(\boldsymbol{R}_2,\boldsymbol{R}_3)\le\boldsymbol{k}|\boldsymbol{R}_2=\boldsymbol{r}_2,\boldsymbol{R}_3=\boldsymbol{r}_3\}.\end{aligned}$$

命题 2 (E) 可直接由引理 1 得出，因为 $\boldsymbol{T}_i=\sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$ 是（$\boldsymbol{Y}_i,\boldsymbol{Z}_i^b$）的函数。因此，命题 2 得证。

B.2　Tran删失受试者的响应向量变换

同样地，我们将对删失受试者的响应向量进行变换，建立关于给定该变换后的响应条件分布的五个性质，然后利用这些性质简化公式（7）中给出的复杂有效得分向量。这五个性质概述于命题3中：

命题3 考虑变换后的响应变量 $\boldsymbol{T}_{\mathrm{aug}\,i} = \sigma^{-2}(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)^T \boldsymbol{Y}_i$。如果矩阵 $(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)$ 具有满列秩（即，如果 排名$\{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)\} = p_b + 1$），则，

P1* $E(\boldsymbol{Y}i|\boldsymbol{T}aug\ i\ \boldsymbol{D}_i, \boldsymbol{b}_i, U_i) = \boldsymbol{\zeta}(\boldsymbol{D}_i; \boldsymbol{\theta}) + \boldsymbol{P}(\mathbf{1}_{m_i}, \boldsymbol{z}_i^b)\{\boldsymbol{Y}i - \boldsymbol{\zeta}(\boldsymbol{D}_i; \boldsymbol{\theta})\}$，其中矩阵 $\boldsymbol{P}_M = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T$是矩阵 $\boldsymbol{M}$的正交投影算子。

P2* 在给定$(\boldsymbol{T}_i^{\mathrm{aug}}, \boldsymbol{D}_i)$的条件下，$\boldsymbol{Y}_i$与$(\boldsymbol{b}_i, U_i)$相互独立。

P3* $\mathrm{E}(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_i, \boldsymbol{b}_i, U_i) = \sigma^2(m_i - p_b) + ||\mathrm{E}(\boldsymbol{Y}_i|\boldsymbol{T}^{\mathrm{aug}}{}_i\ \widehat{\boldsymbol{D}}_i, \boldsymbol{b}_i, \mathrm{U_i})||_2^2$，其中 $||\cdot||_2^2$表示 *L2*范数。

P4* 对于任意函数 $\boldsymbol{g}(\boldsymbol{T}_i^{\mathrm{aug}}, \boldsymbol{D}_i)$，若$\mathrm{E}\{\boldsymbol{g}(\boldsymbol{T}_{\mathrm{aug}\,i}\ \boldsymbol{D}_i)|\boldsymbol{D}_i, \boldsymbol{b}_i, \mathrm{U}_i\} = \mathbf{0}$，则可得$\boldsymbol{g}(\boldsymbol{T}_i^{\mathrm{aug}}, \boldsymbol{D}_i) = \mathbf{0}$。

P5* 给定任意随机变量 $\boldsymbol{R}$，$\boldsymbol{R}$与 $\boldsymbol{T}_{\mathrm{aug}\,i}$在给定$(\boldsymbol{Y}_i$、$\boldsymbol{Z}_i^b)$的条件下独立。

我们开始证明命题3，但省略其余部分，因为后续内容与命题 2 的证明类似。回顾一下，对于删失受试者，结果模型等价于

$$
\begin{aligned}
&Y_{ij} = \boldsymbol{\beta}^T \boldsymbol{Z}_{ij}^a + \alpha(s_{ij} - \widehat{X}_i - U_i) + \boldsymbol{b}_i^T \boldsymbol{Z}_{ij}^b + \epsilon_{ij} \\
&\boldsymbol{b}_i \sim f_{\boldsymbol{b}}, \quad U_i \sim f_U, \quad \epsilon_{ij} \sim \mathrm{Normal}(0, \sigma^2)
\end{aligned}
$$

由于插补条件均值 X_i 通过 Berkson 误差模型与删失 X_i 相关，$X_i = X_i + U_i$。因此，在给定变量 $(\boldsymbol{D}_i, \boldsymbol{b}_i, U_i)$ 的条件下，m_i 维响应向量 $\boldsymbol{Y}_i$ 服从多元正态分布 $\mathrm{Normal}_{m_i}\{\boldsymbol{Z}_i^a\boldsymbol{\beta} + \alpha(\boldsymbol{S}_i - \widehat{X}_i\mathbf{1}_{m_i} - U_i\mathbf{1}_{m_i}) + \boldsymbol{Z}_i^b\boldsymbol{b}_i, \sigma^2\boldsymbol{I}_{m_i}\}$。为简化该多元正态分布的均值-，我们沿用记号 $\boldsymbol{\zeta}(\boldsymbol{D}_i; \boldsymbol{\theta}) = \boldsymbol{Z}_i^a\boldsymbol{\beta} + \alpha(\boldsymbol{s}_i - X_i\mathbf{1}_{m_i})$。采用此记号后，给定 $(\widehat{\boldsymbol{D}}_i, \boldsymbol{b}_i, U_i)$ 时 $\boldsymbol{Y}_i$的条件均值为 $\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) + (\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)(-\alpha U_i, \boldsymbol{b}_i^T)^T$。类似

于 B.1节，我们提出对删失受试者的响应变量进行如下变换：$\boldsymbol{T}_{\text{增广}\,i} = \sigma^{-2}(\mathbf{1}_{m_i},\ \boldsymbol{Z}_i^b)^T \boldsymbol{Y}_i$。这与对未删失受试者的响应向量所使用的变换相同，只是 $\boldsymbol{Z}_i^b$ 增加了一个全 1 向量（因此上标为“增广”）。此后，命题3的证明与命题 2 的证明类似，其中将 $\boldsymbol{D}_i$ 替换为 $\widehat{\boldsymbol{D}}_i$，将 $\boldsymbol{Z}_i^b$ 替换为 $(\mathbf{1}_{m_i},\ \boldsymbol{Z}_i^b)$，将 $\boldsymbol{b}_i$ 替换为 $(\boldsymbol{b}_i,\ U_i)$，并将 $\boldsymbol{T}_i$ 替换为 $\boldsymbol{T}_{\text{增广}\,i}$。

B.3 求有效得分向量的闭式 s

为了简化未删失受试者的有效得分向量（公式（6）），首先回顾一下，该有效得分向量的原始形式要求我们找到一个函数 $\boldsymbol{h}(\boldsymbol{D}_i,\ \boldsymbol{b}_i)$，使其满足

$$\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}; \boldsymbol{\theta}) | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}\} = \mathrm{E}[\mathrm{E}_\mathrm{mis}\{\boldsymbol{h}(\boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}) | \boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\} | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}],$$

其中 $\boldsymbol{S}_{\boldsymbol{\theta},\ \mathrm{mis}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i) = \frac{\partial}{\partial \boldsymbol{\theta}} \log f_{\boldsymbol{Y}_i,\ \boldsymbol{D}_i,\ \text{误分类}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$。注意到

$$\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}; \boldsymbol{\theta}) | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}\} = \mathrm{E}[\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}; \boldsymbol{\theta}) | \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}\} | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}]$$

根据全期望定律。此外,

$$\mathrm{E}[\mathrm{E}_\mathrm{mis}\{\boldsymbol{h}(\boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}) | \boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\} | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}] = \mathrm{E}[\mathrm{E}_\mathrm{mis}\{\boldsymbol{h}(\boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}) | \boldsymbol{Y}_\mathrm{i}, \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\} | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}]$$

根据命题 2（E），因此，我们对 $\boldsymbol{h}(\boldsymbol{D}_i,\ \boldsymbol{b}_i)$ 的条件变为

$\mathrm{E}[\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_\mathrm{i}\ \boldsymbol{D}_\mathrm{i};\ \boldsymbol{\theta}) | \boldsymbol{T}_\mathrm{i}\ \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}\} | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}] = \mathrm{E}[\mathrm{E}_\mathrm{mis}\{\boldsymbol{h}(\boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}) | \boldsymbol{Y}_\mathrm{i}\ \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\} | \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}]$.接下来，我们利用命题 2 (B) 得出以下两个结果：

$$\begin{aligned}
\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}; \boldsymbol{\theta}) | \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}\} &= \mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}; \boldsymbol{\theta}) | \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\} \\
\mathrm{E}_\mathrm{mis}\{\boldsymbol{h}(\boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}) | \boldsymbol{Y}_\mathrm{i}, \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\} &= \mathrm{E}_\mathrm{mis}\{\boldsymbol{h}(\boldsymbol{D}_\mathrm{i}, \boldsymbol{b}_\mathrm{i}) | \boldsymbol{T}_\mathrm{i}, \boldsymbol{D}_\mathrm{i}\}.
\end{aligned}$$

结合这些结果，有效得分因此必须满足

$$\mathrm{E}[\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}|\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}}] = \mathrm{E}[\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}|\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}}].$$

结合命题 2 (D)，这意味着

$$\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} = \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\},$$

这进一步意味着

$$\begin{aligned}
&\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} = \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} \\
= \;& \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} = \mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}
\end{aligned}$$

根据命题 2 (E)，有效得分向量因此可以写为

$$\begin{aligned}
\boldsymbol{S}_{\mathrm{eff},\mathrm{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) &= \boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) - \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{h}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} \\
&= \boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) - \mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}. \qquad (8)
\end{aligned}$$

重要的是，这是一个闭式解。我们可以简化公式（8）中的第一项：

$$\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) = \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\},$$

其中“模型”得分向量 $\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i,\ \boldsymbol{b}_i;\ \boldsymbol{\theta}) = \frac{\partial}{\partial\boldsymbol{\theta}}\log f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,\ \boldsymbol{b}_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i,\ \boldsymbol{b}_i)$ 不受模型设定错误的影响。我们还可以简化公式（8）中的第二项：

$$\begin{aligned}
\mathrm{E}\{\boldsymbol{S}_{\boldsymbol{\theta},\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} &= \mathrm{E}[\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}] \\
&= \int\left\{\int \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{b}_i;\boldsymbol{\theta}) f_{\mathrm{mis}}(\boldsymbol{b}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)d\mu(\boldsymbol{b}_i)\right\}\times \\
&\quad f(\boldsymbol{Y}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)d\mu(\boldsymbol{y}_i) \\
&= \int \boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{b}_i;\boldsymbol{\theta}) f_{\mathrm{mis}}(\boldsymbol{Y}_i,\boldsymbol{b}_i|\boldsymbol{T}_i,\boldsymbol{D}_i)d\mu(\boldsymbol{b}_i)d\mu(\boldsymbol{y}_i) \\
&= \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}.
\end{aligned}$$

这将得到以下结果

$$\begin{aligned}\boldsymbol{S}_{\text{eff,mis}}(\boldsymbol{Y}_i, \boldsymbol{D}_i; \boldsymbol{\theta}) &= \mathrm{E}_{\text{mis}}\{\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\text{i}}, \boldsymbol{D}_{\text{i}}, \boldsymbol{b}_{\text{i}}; \boldsymbol{\theta})|\boldsymbol{Y}_{\text{i}}, \boldsymbol{D}_{\text{i}}\} - \\ &\quad \mathrm{E}_{\text{mis}}\{\boldsymbol{S}_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\text{i}}, \boldsymbol{D}_{\text{i}}, \boldsymbol{b}_{\text{i}}; \boldsymbol{\theta})|\boldsymbol{T}_{\text{i}}, \boldsymbol{D}_{\text{i}}\}.\end{aligned}$$

与上述工作类似，可以证明，删失受试者的这一有效得分向量等价于

$$\begin{aligned}\boldsymbol{S}^*_{\text{eff,mis}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) &= \mathrm{E}_{\text{mis}}\{\boldsymbol{S}^*_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\text{i}}, \widehat{\boldsymbol{D}}_{\text{i}}, \boldsymbol{b}_{\text{i}}, \mathrm{U}_{\text{i}}; \boldsymbol{\theta})|\boldsymbol{Y}_{\text{i}}, \widehat{\boldsymbol{D}}_{\text{i}}\} - \\ &\quad \mathrm{E}_{\text{mis}}\{\boldsymbol{S}^*_{\boldsymbol{\theta}}(\boldsymbol{Y}_{\text{i}}, \widehat{\boldsymbol{D}}_{\text{i}}, \boldsymbol{b}_{\text{i}}, \mathrm{U}_{\text{i}}; \boldsymbol{\theta})|\boldsymbol{T}^{\text{aug}}_{\text{i}}, \widehat{\boldsymbol{D}}_{\text{i}}\}.\end{aligned}$$

C 计算有效得分向量

C.1 未删失受试者

回顾章节（B.1），给定 $(\boldsymbol{D}_i, \boldsymbol{b}_i)$ 时，单个个体对应的响应向量 $\boldsymbol{Y}_i$ 服从正态分布$_{m_i}($ $\boldsymbol{Z}^a_i\boldsymbol{\beta} + \alpha(\boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \boldsymbol{Z}^b_i\boldsymbol{b}_i, \sigma^2\boldsymbol{I}_{m_i})$。因此，我们知道，单个个体的所有观测值的密度 $f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{b}_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i, \boldsymbol{b}_i; \boldsymbol{\theta}) =:$

$$(2\pi\sigma^2)^{-m_i/2}\exp\left[-\frac{||\boldsymbol{Y}_i - \{\boldsymbol{Z}^a_i\boldsymbol{\beta} + \alpha(\boldsymbol{s}_i - X\mathbf{1}_{m_i}) + \boldsymbol{Z}^b_i\boldsymbol{b}_i\}||^2_2}{2\sigma^2}\right]$$

因此，对数似然为 $\log\{f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{b}_i}(\cdot)\} =$

$$-\frac{m_i}{2}\log(2\pi\sigma^2) - \frac{||\boldsymbol{Y}_i - \{\boldsymbol{Z}^a_i\boldsymbol{\beta} + \alpha(\boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \boldsymbol{Z}^b_i\boldsymbol{b}_i\}||^2_2}{2\sigma^2} \tag{9}$$

C.1.1 β的有效得分向量

由公式（9）中的对数似然可得

$$\begin{aligned}\boldsymbol{S}_{\boldsymbol{\beta}}(\boldsymbol{Y}_i, \boldsymbol{D}_i, \boldsymbol{b}_i; \boldsymbol{\theta}) &\equiv \frac{\partial}{\partial \boldsymbol{\beta}} \log\{f_{\boldsymbol{Y}_i|\boldsymbol{D}_i, \boldsymbol{b}_i}(\cdot)\} \\ &= \frac{\boldsymbol{Z}_i^{aT}[\boldsymbol{Y}_i - \{\boldsymbol{Z}_i^a\boldsymbol{\beta} + \alpha(\boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \boldsymbol{Z}_i^b\boldsymbol{b}_i\}]}{\sigma^2} \\ &= \frac{\boldsymbol{Z}_i^{aT}\boldsymbol{Y}_i - \boldsymbol{k}_1(\boldsymbol{D}_i, \boldsymbol{b}_i)}{\sigma^2},\end{aligned}$$

其中 $\boldsymbol{k}_1(\boldsymbol{D}_i, \boldsymbol{b}_i) = \boldsymbol{Z}_i^{aT}\{\boldsymbol{Z}_i^a\boldsymbol{\beta} + \alpha(\boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \boldsymbol{Z}_i^b\boldsymbol{b}_i\}$。根据命题 2 的(B)和(E)部分，我们有：

$$\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{k}_1(\boldsymbol{D}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}})|\boldsymbol{Y}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}\} = \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{k}_1(\boldsymbol{D}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}\}.$$

这意味着 β、 $\boldsymbol{S}_{\mathrm{eff},\,\beta}(\boldsymbol{Y}_i, \boldsymbol{D}_i; \boldsymbol{\theta})$的有效得分向量等价于：

$$\begin{aligned}\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i, \boldsymbol{D}_i; \boldsymbol{\theta}) &= \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\boldsymbol{\beta}}(\boldsymbol{Y}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}; \boldsymbol{\theta})|\boldsymbol{Y}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}\} - \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\boldsymbol{\beta}}(\boldsymbol{Y}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}; \boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}\} \\ &= \sigma^{-2}\Big[\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{aT}}\boldsymbol{Y}_{\mathrm{i}} - \boldsymbol{k}_1(\boldsymbol{D}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}})|\boldsymbol{Y}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}\} - \\ &\quad \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{aT}}\boldsymbol{Y}_{\mathrm{i}} - \boldsymbol{k}_1(\boldsymbol{D}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}\}\Big] \\ &= \sigma^{-2}\{\mathrm{E}_{\mathrm{mis}}(\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{aT}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{Y}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}}) - \mathrm{E}_{\mathrm{mis}}(\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{aT}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}})\} \\ &= \frac{\boldsymbol{Z}_i^{aT}\{\boldsymbol{Y}_i - \mathrm{E}_{\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}})\}}{\sigma^2}.\end{aligned}$$

请回忆，$_{\mathrm{mis}}$ 下标表示可能受模型设定错误影响的期望。这里，由于 $\mathrm{E}_{\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}})$ 仅取决于假定的结果模型——我们假设该模型正确——因此这一期望不受模型设定错误影响；所以我们省略 $_{\mathrm{mis}}$ 下标，如下所示：

$$\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i, \boldsymbol{D}_i; \boldsymbol{\theta}) = \frac{\boldsymbol{Z}_i^{aT}\{\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}, \boldsymbol{D}_{\mathrm{i}})\}}{\sigma^2}$$

其中，根据命题 2，$\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})\ =\boldsymbol{\zeta}(\boldsymbol{D}_{\mathrm{i}};\ \boldsymbol{\theta})+\boldsymbol{P}_{\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{b}}}\{\boldsymbol{Y}_{\mathrm{i}}\ -\boldsymbol{\zeta}(\boldsymbol{D}_{\mathrm{i}};\ \boldsymbol{\theta})\}$ 其中 $\boldsymbol{P}_{\boldsymbol{M}}=\boldsymbol{M}(\boldsymbol{M}^{T}\boldsymbol{M})^{-1}\boldsymbol{M}^{T}$对于矩阵 $\boldsymbol{M},\boldsymbol{T}_{i}=\sigma^{-2}\boldsymbol{Z}_{i}^{bT}\boldsymbol{Y}_{i},$ 且 $\boldsymbol{\zeta}(\boldsymbol{D}_{i};\ \boldsymbol{\theta})\ =\boldsymbol{Z}_{i}^{a}\boldsymbol{\beta}+\alpha(\boldsymbol{S}_{i}-X_{i}\mathbf{1}_{m_{i}})$。鉴于此，我们可以进一步简化 $\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_{i},\boldsymbol{D}_{i};\boldsymbol{\theta})$。请看：

$$\begin{aligned}\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_{i},\boldsymbol{D}_{i};\boldsymbol{\theta}) &= \frac{\boldsymbol{Z}_{i}^{aT}\{\boldsymbol{Y}_{i}-\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})-\boldsymbol{P}_{\boldsymbol{Z}_{i}^{b}}\{\boldsymbol{Y}_{i}-\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})\}}{\sigma^{2}}\\ &= \frac{\boldsymbol{Z}_{i}^{aT}(\boldsymbol{I}_{m_{i}}-\boldsymbol{P}_{\boldsymbol{Z}_{i}^{b}})\{\boldsymbol{Y}_{i}-\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})\}}{\sigma^{2}}.\end{aligned}$$

C.1.2 α的有效得分

α的有效得分记为$\boldsymbol{S}_{\mathrm{eff},\ \alpha}(\boldsymbol{Y}_{i},\ \boldsymbol{D}_{i};\boldsymbol{\theta})$，其推导与$\boldsymbol{S}_{\mathrm{eff},\ \boldsymbol{\beta}}(\boldsymbol{Y}_{i},\ \boldsymbol{D}_{i};\boldsymbol{\theta})$的推导类似。

$$\begin{aligned}S_{\mathrm{eff},\alpha}(\boldsymbol{Y}_{i},\boldsymbol{D}_{i};\boldsymbol{\theta}) &= \frac{(\boldsymbol{s}_{i}-X_{i}\mathbf{1}_{m_{i}})^{T}\{\boldsymbol{Y}_{i}-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}}{\sigma^{2}}\\ &= \frac{(\boldsymbol{s}_{i}-X_{i}\mathbf{1}_{m_{i}})^{T}(\boldsymbol{I}_{m_{i}}-\boldsymbol{P}_{\boldsymbol{Z}_{i}^{b}})\{\boldsymbol{Y}_{i}-\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})\}}{\sigma^{2}}.\end{aligned}$$

C.1.3 σ^2的有效得分

根据公式（9）中的对数似然，我们得到

$$\begin{aligned}\boldsymbol{S}_{\sigma^{2}}(\boldsymbol{Y}_{i},\boldsymbol{D}_{i},\boldsymbol{b}_{i};\boldsymbol{\theta}) &\equiv \frac{\partial}{\partial\sigma^{2}}\log\{f_{\boldsymbol{Y}_{i}|\boldsymbol{D}_{i},\boldsymbol{b}_{i}}(\cdot)\}\\ &= \frac{||\boldsymbol{Y}_{i}-\{\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})+\boldsymbol{Z}_{i}^{b}\boldsymbol{b}_{i}\}||_{2}^{2}}{2\sigma^{4}}-\frac{m_{i}}{2\sigma^{2}}\\ &= \frac{\boldsymbol{Y}_{i}^{T}\boldsymbol{Y}_{i}}{2\sigma^{4}}-\frac{\boldsymbol{Y}_{i}^{T}\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})}{\sigma^{4}}-\frac{\boldsymbol{Y}_{i}^{T}\boldsymbol{Z}_{i}^{b}\boldsymbol{b}_{i}}{\sigma^{4}}+\frac{||\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})+\boldsymbol{Z}_{i}^{b}\boldsymbol{b}_{i}||_{2}^{2}}{2\sigma^{4}}-\frac{m_{i}}{2\sigma^{2}}\\ &= \frac{\boldsymbol{Y}_{i}^{T}\boldsymbol{Y}_{i}}{2\sigma^{4}}-\frac{\boldsymbol{Y}_{i}^{T}\boldsymbol{\zeta}(\boldsymbol{D}_{i};\boldsymbol{\theta})}{\sigma^{4}}-\frac{\boldsymbol{T}_{i}^{T}\boldsymbol{b}_{i}}{\sigma^{2}}+\frac{k_{2}(\boldsymbol{D}_{i},\boldsymbol{b}_{i})}{2\sigma^{2}},\end{aligned}$$

其中 $k_{2}(\boldsymbol{D}_{i},\boldsymbol{b}_{i})\ =\sigma^{-2}||\boldsymbol{\zeta}(\boldsymbol{D}_{i};\ \boldsymbol{\theta}_{i})\ +\boldsymbol{Z}_{i}^{b}\boldsymbol{b}_{i}||_{2}^{2}-m_{i}$。根据命题 2 的 (B) 部分和 (E) 部分，我们有：

$$\mathrm{E}_{\mathrm{mis}}\{\mathrm{k}_{2}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}=\mathrm{E}_{\mathrm{mis}}\{\mathrm{k}_{2}(\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\}.$$

再次根据命题 2 的 (B) 部分和 (E) 部分，$\mathrm{E}_{\mathrm{mis}}\{\boldsymbol{T}^{\mathrm{T}}{}_{\mathrm{i}}\ \boldsymbol{b}_{\mathrm{i}}|\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}}\} =$

$$\begin{aligned} & \mathrm{E}_{\mathrm{mis}}(\boldsymbol{T}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{b}_{\mathrm{i}}|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}) = \boldsymbol{T}_{\mathrm{i}}^{\mathrm{T}}\mathrm{E}_{\mathrm{mis}}(\boldsymbol{b}_{\mathrm{i}}|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}) \\ = \quad & \boldsymbol{T}_{i}^{T}\mathrm{E}_{\mathrm{mis}}(\boldsymbol{b}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}) = \mathrm{E}_{\mathrm{mis}}(\boldsymbol{T}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{b}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}). \end{aligned}$$

由此可得，σ^2 的有效得分（记作 $S_{\mathrm{eff},\sigma^2}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})$）为:

$$\begin{aligned} S_{\mathrm{eff},\sigma^2}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) \quad = \quad & \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\sigma^2}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} - \\ & \mathrm{E}_{\mathrm{mis}}\{\boldsymbol{S}_{\sigma^2}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\} \\ = \quad & \mathrm{E}_{\mathrm{mis}}\left\{\frac{\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}}{2\sigma^4} - \frac{\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{\zeta}(\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})}{\sigma^4}|\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\right\} - \\ & \mathrm{E}_{\mathrm{mis}}\left\{\frac{\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}}{2\sigma^4} - \frac{\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{\zeta}(\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})}{\sigma^4}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}\right\} \\ = \quad & \frac{1}{\sigma^4}\Big[\frac{1}{2}\left\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}_{\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\right\} - \\ & \boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})^T\{\boldsymbol{Y}_i - \mathrm{E}_{\mathrm{mis}}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}\Big], \end{aligned}$$

其中，在这些条件期望中，我们同样可以忽略下标 $_{\mathrm{mis}}$ ，因为给定 $(\boldsymbol{T}_i,\boldsymbol{D}_i)$ 时，$\boldsymbol{Y}_i$ 的条件分布不存在模型设定错误。

C.1.4 $\boldsymbol{\theta}$ 的有效得分向量

重要的是，我们已经证明，$\boldsymbol{S}_{\mathrm{eff},\ \mathrm{mis}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$ 的每个分量都不会受到模型设定错误的影响。为强调这一点，我们将 $\boldsymbol{\theta} = (\boldsymbol{\beta},\ \alpha,\ \sigma^2)^T$（对应于未删失受试者）的有效得分向量简称为 $\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$。结合这些结果，我们发现

$$\sigma^4\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) = \begin{bmatrix} \sigma^2\boldsymbol{Z}_i^{aT}(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\} \\ \sigma^2(\boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\} \\ \frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\} - \boldsymbol{\zeta}^{\mathrm{T}}(\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})\{\boldsymbol{Y}_{\mathrm{i}} - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\} \end{bmatrix}. \quad (10)$$

回顾命题 2 的 (A) 和 (C) 部分，

$$\begin{aligned}\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\boldsymbol{b}_{\mathrm{i}}) &= \boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})+\boldsymbol{P}_{\boldsymbol{Z}_i^b}\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})\}\\ \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}}) &= \sigma^2(m_i-p_b)+||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})||_2^2,\end{aligned}$$

其中 $\boldsymbol{P_M}=\boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T$对于矩阵 $\boldsymbol{M}$, $\boldsymbol{T}_i=\sigma^{-2}\boldsymbol{Z}_i^{bT}\boldsymbol{Y}_i$, 以及 $\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})=\boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha(\boldsymbol{S}_i-X_i\mathbf{1}_{m_i})$。

C.2 删失受试者

回顾第 (B.2) 节，在给定 $(\boldsymbol{D}_i,\ \boldsymbol{b}_i,\ U_i)$ 的条件下，来自一名删失受试者的响应向量 $\boldsymbol{Y}_i$ 服从多元正态分布 $\mathrm{Normal}_{m_i}\{\boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha(\boldsymbol{S}_i-\widehat{X}_i\mathbf{1}_{m_i}-U_i\mathbf{1}_{m_i})+\boldsymbol{Z}_i^b\boldsymbol{b}_i,\ \sigma^2\boldsymbol{I}_{m_i}\}$。因此，我们知道，来自单一个体的所有观测值的密度为 $f_{\boldsymbol{Y}_i|\widehat{\boldsymbol{D}}_i,\ \boldsymbol{b}_i,\ U_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i,\ \boldsymbol{b}_i,\ U_i;\boldsymbol{\theta})$ =

$$(2\pi\sigma^2)^{-m_i/2}\exp\left[-\frac{||\boldsymbol{Y}_i-\{\boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha(\boldsymbol{s}_i-\widehat{X}_i\mathbf{1}_{m_i})+(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)(-\alpha U_i,\boldsymbol{b}_i^T)^T\}||_2^2}{2\sigma^2}\right].$$

因此，对数似然 $\log\{f_{\boldsymbol{Y}_i|\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i,\ \boldsymbol{b}_i,\ U_i;\boldsymbol{\theta})\}=$

$$-\frac{m_i}{2}\log(2\pi\sigma^2)-\frac{||\boldsymbol{Y}_i-\{\boldsymbol{Z}_i^a\boldsymbol{\beta}+\alpha(\boldsymbol{s}_i-\widehat{X}_i\mathbf{1}_{m_i})+(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)(-\alpha U_i,\boldsymbol{b}_i^T)^T\}||_2^2}{2\sigma^2}. \quad (11)$$

C.2.1 $\boldsymbol{\beta}$ 的有效得分向量

对于 $\boldsymbol{\beta}$ 的删失受试者，其有效得分向量的推导与未删失受试者的推导类似（附录 C.1.1）。由此可得:

$$\boldsymbol{S}^*_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta}) = \frac{\boldsymbol{Z}_i^{aT}(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\}}{\sigma^2}.$$

C.2.2 α的有效得分

接下来，我们考虑与删失受试者对应的 α 的有效得分。根据公式（11）中的对数似然，我们可得：

$$\begin{aligned}\boldsymbol{S}_\alpha(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i;\boldsymbol{\theta}) &\equiv \frac{\partial}{\partial\alpha}\log\{f_{\boldsymbol{Y}_i|\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i}(\boldsymbol{Y}_i|\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i)\}\\ &= \frac{\{\boldsymbol{S}_i-(\widehat{X}_i+U_i)\mathbf{1}_{m_i}\}^T[\boldsymbol{Y}_i-\{\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})+(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)(-\alpha U_i,\boldsymbol{b}_i^T)^T\}]}{\sigma^2}\\ &= \frac{\{\boldsymbol{S}_i-(\widehat{X}_i+U_i)\mathbf{1}_{m_i}\}^T\boldsymbol{Y}_i-\boldsymbol{k}_3(\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i)}{\sigma^2},\end{aligned}$$

其中 $\boldsymbol{k}_3(\widehat{\boldsymbol{D}}_i,\boldsymbol{b}_i,U_i)=\{\boldsymbol{S}_i-(\widehat{X}_i+U_i)\mathbf{1}_{m_i}\}^T\{\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})+(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)(-\alpha U_i,\boldsymbol{b}_i^T)^T\}$。然后，再次利用命题3的(B)和(E)部分，可得：

$$\mathrm{E_{mis}}\{\boldsymbol{k}_3(\widehat{\boldsymbol{D}}_\mathrm{i},\boldsymbol{b}_\mathrm{i},\mathrm{U_i})|\boldsymbol{Y}_\mathrm{i},\widehat{\boldsymbol{D}}_\mathrm{i}\}=\mathrm{E_{mis}}\{\boldsymbol{k}_3(\widehat{\boldsymbol{D}}_\mathrm{i},\boldsymbol{b}_\mathrm{i},\mathrm{U_i})|\boldsymbol{T}_\mathrm{i}^{\mathrm{aug}},\widehat{\boldsymbol{D}}_\mathrm{i}\}.$$

接下来，我们考虑项 $\sigma^{-2}U_i\mathbf{1}_{m_i}^T\boldsymbol{Y}_i$。由于我们定义了 $\boldsymbol{T}_{\mathrm{aug}\,i}=\sigma^{-2}(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)^T\boldsymbol{Y}_i$，因此有 $\sigma^{-2}U_i\mathbf{1}_{m_i}^T\boldsymbol{Y}_i=U_iT_{i1}^{\mathrm{aug}}$，其中 $T_{\mathrm{aug}\,i1}$表示 $\boldsymbol{T}_{\mathrm{aug}\,i}$的第一个元素。然后，根据命题3的(B)和(E)部分可得 $E(U_iT_{i1}^{\mathrm{aug}}|\boldsymbol{Y}_i,\boldsymbol{D}_i)=$

$$\begin{aligned}&E(U_iT_{i1}^{\mathrm{aug}}|\boldsymbol{Y}_i,\boldsymbol{T}_i^{\mathrm{aug}},\widehat{\boldsymbol{D}}_i)=T_{i1}^{\mathrm{aug}}E(U_i|\boldsymbol{Y}_i,\boldsymbol{T}_i^{\mathrm{aug}},\widehat{\boldsymbol{D}}_i)\\ =\;& T_{i1}^{\mathrm{aug}}E(U_i|\boldsymbol{T}_i^{\mathrm{aug}},\widehat{\boldsymbol{D}}_i)=E(U_iT_{i1}^{\mathrm{aug}}|\boldsymbol{T}_i^{\mathrm{aug}},\widehat{\boldsymbol{D}}_i).\end{aligned}$$

因此，在计算 α 的有效得分时，$\boldsymbol{k}_3(\boldsymbol{D}_i$、$\boldsymbol{b}_i$、$U_i)$ 和 $U_iT_{i1}^{\mathrm{aug}}$ 的条件期望会消去。由此，我们得到

$$S_{\mathrm{eff},\alpha}^*(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})=\frac{(\boldsymbol{s}_i-\widehat{X_i}\mathbf{1}_{m_i})^T(\boldsymbol{I}_{m_i}-\boldsymbol{P}_{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i-\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\}}{\sigma^2}.$$

C.2.3 σ^2的有效得分

针对删失受试者的 σ^2 有效得分推导与附录（C.1.3）中所示的推导过程相似。因此，我们得到：

$$\begin{aligned} S^*_{\text{eff},\sigma^2}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) &= \frac{1}{\sigma^4}\Big[\frac{1}{2}\Big\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_{\mathrm{i}})\Big\} - \\ &\quad \boldsymbol{\zeta}^T(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta})\{\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_{\mathrm{i}})\}\Big]. \end{aligned}$$

C.2.4 $\boldsymbol{\theta}$的有效得分向量

结合这些结果，我们得到

$$\begin{aligned} &\sigma^4 \boldsymbol{S}^*_{\text{eff}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) \\ &= \begin{bmatrix} \sigma^2 \boldsymbol{Z}_i^{aT}(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta})\} \\ \sigma^2(\boldsymbol{s}_i - \widehat{X}_i \mathbf{1}_{m_i})^T(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta})\} \\ \frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_{\mathrm{i}})\} - \boldsymbol{\zeta}^{\mathrm{T}}(\widehat{\boldsymbol{D}}_{\mathrm{i}}; \boldsymbol{\theta})\{\boldsymbol{Y}_{\mathrm{i}} - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_{\mathrm{i}})\} \end{bmatrix}. \end{aligned} \tag{12}$$

回顾命题3的(A)和(C)部分可知

$$\begin{aligned} \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_{\mathrm{i}}, \boldsymbol{b}_{\mathrm{i}}, \mathrm{U}_{\mathrm{i}}) &= \boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta}) + \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)}\{\boldsymbol{Y}_i - \boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta})\} \\ E(\boldsymbol{Y}_i^T\boldsymbol{Y}_i|\boldsymbol{T}_i^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_i) &= \sigma^2(m_i - p_b) + ||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}}, \widehat{\boldsymbol{D}}_{\mathrm{i}})||_2^2, \end{aligned}$$

其中，对于矩阵 $\boldsymbol{M}$， $\boldsymbol{P}_{\boldsymbol{M}} = \boldsymbol{M}(\boldsymbol{M}^T\boldsymbol{M})^{-1}\boldsymbol{M}^T$增广$_i$ $= \sigma^{-2}(\mathbf{1}_{m_i}, \boldsymbol{Z}^b)^T\boldsymbol{Y}_i$，且 $\boldsymbol{\zeta}(\widehat{\boldsymbol{D}}_i; \boldsymbol{\theta})$ $=\sim \boldsymbol{Z}_i^a\boldsymbol{\beta} + \alpha(\boldsymbol{S}_i - X_i\mathbf{1}_{m_i})$。

C.3 完整估计方程

现在，我们已经推导出了分别对应未删失受试者和删失受试者的 $\boldsymbol{\theta}$ 有效得分向量。将这些向量结合起来，可得到总估计方程：

$$\sum_{i=1}^{n}\Big\{\Delta_i \boldsymbol{S}_{\text{eff}}(\boldsymbol{Y}_i, \boldsymbol{D}_i; \boldsymbol{\theta}) + (1-\Delta_i)\boldsymbol{S}^*_{\text{eff}}(\boldsymbol{Y}_i, \widehat{\boldsymbol{D}}_i; \boldsymbol{\theta})\Big\} = \mathbf{0}, \tag{13}$$

其中，删失指示变量 Δ_i 在受试者 i 未被删失时取 1，否则取 0。$\widehat{}$有效得分向量 $\boldsymbol{S}_{\text{eff}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$和 $\boldsymbol{S}^{*\text{效}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})$分别见公式（10）和公式（12）。

D 定理 1 的证明

$\widehat{}$现在我们证明 $\boldsymbol{\theta}_{\text{ACE}}$ 是可识别的，也就是说，正文中的定理 1 成立。首先考虑公式（2）中该估计方程的前 $p+1$ 个元素（即对应于（$\boldsymbol{\beta}$, α）的元素）：

$$
\begin{aligned}
\mathbf{0} = \quad & \sum_{i=1}^{n}\Big[\Delta_i\sigma^2(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\{\boldsymbol{Y}_i - (\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})(\boldsymbol{\beta}^T, \alpha)^T\} + \\
& (1-\Delta_i)\sigma^2(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - \widehat{X}_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})\{\boldsymbol{Y}_i - (\boldsymbol{Z}_i^a, \boldsymbol{s}_i - \widehat{X}_i\mathbf{1}_{m_i})(\boldsymbol{\beta}^T, \alpha)^T\}\Big].
\end{aligned}
$$

将所有不含 $(\boldsymbol{\beta}, \alpha)$ 的项移至等式左侧，可得:

$$
\begin{aligned}
& \sum_{i=1}^{n}\Big\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b}) + \\
& (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - \widehat{X}_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})\Big\}\boldsymbol{Y}_i \\
= \quad & \sum_{i=1}^{n}\Big\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \\
& (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - \widehat{X}_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - \widehat{X}_i\mathbf{1}_{m_i})\Big\}(\boldsymbol{\beta}^T, \alpha)^T.
\end{aligned}
$$

由此，我们得到以下回归系数估计值:

$$
\begin{aligned}
(\hat{\boldsymbol{\beta}}^T, \hat{\alpha})^T &= \Big[\sum_{i=1}^{n}\Big\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \\
&\quad (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})\Big\}\Big]^{-1} \times \\
&\quad \Big[\sum_{i=1}^{n}\Big\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{Y}_i + \\
&\quad (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{Y}_i\Big\}\Big] \\
&= \boldsymbol{M}_i^{-1}\Big[\sum_{i=1}^{n}\Big\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{Y}_i + \\
&\quad (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})\boldsymbol{Y}_i\Big\}\Big],
\end{aligned}
$$

假设矩阵

$$
\begin{aligned}
\boldsymbol{M}_i &= \sum_{i=1}^{n}\Big\{\Delta_i(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{\boldsymbol{Z}_i^b})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i}) + \\
&\quad (1-\Delta_i)(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})^T(\boldsymbol{I}_m - \boldsymbol{P}_{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)})(\boldsymbol{Z}_i^a, \boldsymbol{s}_i - X_i\mathbf{1}_{m_i})\Big\}
\end{aligned}
$$

可逆。接下来，我们考虑估计方程的第（$p+2$）个元素（即对应于 σ^2 的元素）：

$$
\begin{aligned}
0 &= \sum_{i=1}^{n}\Bigg(\Delta_i\left[\frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\} - \boldsymbol{\zeta}^{\mathrm{T}}(\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})\{\boldsymbol{Y}_{\mathrm{i}} - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}\right] + \\
&\quad (1-\Delta_i)\left[\frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}})\} - \boldsymbol{\zeta}^{\mathrm{T}}(\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta})\{\boldsymbol{Y}_{\mathrm{i}} - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}})\}\right]\Bigg) \\
&= \sum_{i=1}^{n}\Bigg(\Delta_i\Big[\frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \sigma^2(m_i-q) + ||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})||_2^2\} \\
&\quad -\boldsymbol{\zeta}^T(\boldsymbol{D}_i;\boldsymbol{\theta})\{\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}\Big] + \\
&\quad (1-\Delta_i)\Big[\frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i - \sigma^2(m_i-q) + ||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}},\boldsymbol{D}^*)||_2^2\} - \\
&\quad \boldsymbol{\zeta}^T(\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})\{\boldsymbol{Y}_i - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}})\}\Big]\Bigg).
\end{aligned}
$$

由此可见， σ^2 的估计值为:

$$
\begin{aligned}
\hat{\sigma^2} &= \left[\sum_{i=1}^{n}\{\Delta_i(m_i-q) + (1-\Delta_i)(m_i-q)\}\right]^{-1} \times \\
&\quad \left\{\sum_{i=1}^{n}\Bigg(\Delta_i\left[\frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i + ||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})||_2^2\} - \boldsymbol{\zeta}^{\mathrm{T}}(\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})\{\boldsymbol{Y}_{\mathrm{i}} - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}\right] + \right. \\
&\quad \left.(1-\Delta_i)\left[\frac{1}{2}\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i + ||\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}}^{\mathrm{aug}},\boldsymbol{D}^*)||_2^2\} - \boldsymbol{\zeta}^{\mathrm{T}}(\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta})\{\boldsymbol{Y}_{\mathrm{i}} - \mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}})\}\right]\Bigg)\right\}.
\end{aligned}
$$

至此，我们已确立了 $\boldsymbol{\theta}$的可识别性。

E 一致性

现在，我们转而关注新型估计量 $\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}}$ 的一致性（定理 2）。要证明该定理，我们首先需要证明完整估计函数的均值为 0

（命题 1）。不过，我们先证明引理 2，该引理阐明了命题 1 中概述的假设 (C1) 和 (C2) 与证明完整估计函数的均值为 0 所需的更严格假设之间的重要关系。

E.1 条件独立性主张之间的关系

引理 2 考虑变换 $\boldsymbol{T}_i = \sigma^2 \boldsymbol{Z}_i^{bT} \boldsymbol{Y}_i$ 和 $\boldsymbol{T}_{\text{aug}\,i} = \sigma^{-2}(\mathbf{1}_m, \boldsymbol{Z}_i^b)^T \boldsymbol{Y}_i$。

(A) 如果给定 $\boldsymbol{D}_i$ 时，$\boldsymbol{Y}_i$ 与 C_i 条件独立，那么给定 $(\boldsymbol{T}_i, \boldsymbol{D}_i)$ 时，$\boldsymbol{Y}_i$ 与 C_i 也必定条件独立。

(B) 如果给定 $\boldsymbol{D}_i$ 时，$\boldsymbol{Y}_i$ 与 C_i 条件独立，那么给定 $(\boldsymbol{T}_i^{\text{aug}}, \boldsymbol{D}_i)$ 时，$\boldsymbol{Y}_i$ 与 C_i 也必定条件独立。

我们证明引理 2 (A)，并省略引理 2 (B) 的证明，因为后者的证明与前者的证明极为相似（将 $(\boldsymbol{D}_i, \boldsymbol{T}_i)$ 替换为 $(\boldsymbol{D}_i, \boldsymbol{T}_i^{\text{aug}})$）。我们首先证明引理 2 (A) 中的前提等价于：

$$f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,C_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i, C_i) = f_{\boldsymbol{Y}_i|\boldsymbol{D}_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i).$$

由此及贝叶斯定理，我们得到：

$$\frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i|C_i}(\boldsymbol{Y}_i, \boldsymbol{D}_i|C_i)}{f_{\boldsymbol{D}_i|C_i}(\boldsymbol{D}_i|C_i)} = \frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i}(\boldsymbol{Y}_i, \boldsymbol{D}_i)}{f_{\boldsymbol{D}_i}(\boldsymbol{D}_i)}.$$

接着，由于 $\boldsymbol{T}_i$ 是 $\boldsymbol{Y}_i$ 和 $\boldsymbol{D}_i$ 的函数，因此可得

$$\frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{Y}_i, \boldsymbol{D}_i, \boldsymbol{T}_i|C_i)}{f_{\boldsymbol{D}_i|C_i}(\boldsymbol{D}_i|C_i)} = \frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{Y}_i, \boldsymbol{D}_i, \boldsymbol{T}_i)}{f_{\boldsymbol{D}_i}(\boldsymbol{D}_i)}. \tag{14}$$

现在，我们应用贝叶斯定理来修改方程（14）左侧的分母：

$$\begin{aligned} \frac{f_{\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{D}_i,\boldsymbol{T}_i|C_i)}{f_{\boldsymbol{D}_i|C_i}(\boldsymbol{D}_i|C_i)} &= f_{\boldsymbol{T}_i|\boldsymbol{D}_i,C_i}(\boldsymbol{T}_i|\boldsymbol{D}_i, C_i) \\ \Rightarrow \qquad f_{\boldsymbol{D}_i|C_i}(\boldsymbol{D}_i|C_i) &= \frac{f_{\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{D}_i,\boldsymbol{T}_i|C_i)}{f_{\boldsymbol{T}_i|\boldsymbol{D}_i,C_i}(\boldsymbol{T}_i|\boldsymbol{D}_i,C_i)}. \end{aligned}$$

然后，根据引理 2 (A) 中的假设，$\boldsymbol{Y}_i$ 在给定 $\boldsymbol{D}_i$ 的条件下与 C_i 条件独立。由于 $\boldsymbol{T}_i$是 $\boldsymbol{Y}_i$和 $\boldsymbol{D}_i$的函数，我们知道 $\boldsymbol{T}_i$在给定 $\boldsymbol{D}_i$的条件下也与 C_i条件独立。因此，$f_{\boldsymbol{T}_i|\boldsymbol{D}_i,\ C_i}(\boldsymbol{T}_i|\boldsymbol{D}_i,\ C_i) = f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i)$，并且

$$f_{\boldsymbol{D}_i|C_i}(\boldsymbol{D}_i|C_i) = \frac{f_{\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{D}_i,\boldsymbol{T}_i|C_i)}{f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i)}.$$

现在，我们应用贝叶斯定理来修改公式（14）右侧的分母：

$$\begin{aligned} \frac{f_{\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{D}_i,\boldsymbol{T}_i)}{f_{\boldsymbol{D}_i}(\boldsymbol{D}_i)} &= f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i) \\ \Rightarrow \qquad f_{\boldsymbol{D}_i}(\boldsymbol{D}_i) &= \frac{f_{\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{D}_i,\boldsymbol{T}_i)}{f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i)}. \end{aligned}$$

Wit对公式（14）的分母作这两处修改后，我们得到 ：

$$\frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i|C_i) f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i)}{f_{\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{D}_i,\boldsymbol{T}_i|C_i)} = \frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i) f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i)}{f_{\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{D}_i,\boldsymbol{T}_i)}.$$

接着，我们看到条件密度 $f_{\boldsymbol{T}_i|\boldsymbol{D}_i}(\boldsymbol{T}_i|\boldsymbol{D}_i)$ 在等式两边相消，从而得到：

$$\frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i|C_i)}{f_{\boldsymbol{D}_i,\boldsymbol{T}_i|C_i}(\boldsymbol{D}_i,\boldsymbol{T}_i|C_i)} = \frac{f_{\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i)}{f_{\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{D}_i,\boldsymbol{T}_i)},$$

这等价于以下陈述：$f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{T}_i,C_i}(\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{T}_i,C_i) = f_{\boldsymbol{Y}_i|\boldsymbol{D}_i,\boldsymbol{T}_i}(\boldsymbol{Y}_i,\boldsymbol{D}_i,\boldsymbol{T}_i)$，因此 $\boldsymbol{Y}_i$ 在给定 $(\boldsymbol{D}_i,\boldsymbol{T}_i)$ 的条件下与 C_i 条件独立。这证实了引理 2（A）；引理 2（B）的证明过程类似，为简洁起见，此处省略。

E.2 命题 1 的证明

我们证明命题 1 (A)，并省略命题 1 (B) 的证明，因为后者的证明可仿照前者的证明得出（将 $(\boldsymbol{D}_i,\boldsymbol{T}_i)$ 替换为 $(\boldsymbol{D}_i,\boldsymbol{T}_i^{\text{aug}})$）。

为证明 Pr 对于命题 1 (A)，我们首先应用全期望定律，得到 :

$$\begin{aligned}\mathrm{E}\{\Delta_{\mathrm{i}}\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})\} &= \mathrm{E}[\mathrm{E}\{\Delta_{\mathrm{i}}\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}\}] \\ &= \mathrm{E}\{\mathrm{I}(\mathrm{X_i}\le\mathrm{C_i})\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}\} \\ &= I(X_i\le C_i)\mathrm{E}\{\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}\}\end{aligned}$$

因为根据定义， $\boldsymbol{D}_i=(\boldsymbol{Z}_i^a,\ \boldsymbol{s}_i,\ X_i,\ \boldsymbol{Z}_i^b)$包含 X_i。因此，要证明所需结果 $\mathrm{E}\{\Delta_{\mathrm{i}}\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})\}=0$，只需证明 $\mathrm{E}\{\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}};\boldsymbol{\theta})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}\}=0$。我们逐分量地完成这一证明。回顾 C 节 C.1.1:

$$\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})=\frac{\boldsymbol{Z}_i^{aT}\{\boldsymbol{Y}_i-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}}{\sigma^2}.$$

因此,

$$\begin{aligned}\mathrm{E}\{\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}\} &= \frac{\mathrm{E}[\boldsymbol{Z}_{\mathrm{i}}^{\mathrm{aT}}\{\boldsymbol{Y}_{\mathrm{i}}-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}]}{\sigma^2} \\ &= \frac{\boldsymbol{Z}_i^{aT}\{\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i})-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}}{\sigma^2}.\end{aligned}$$

根据引理 2，我们关于 $\boldsymbol{Y}_i$ 在给定 $\boldsymbol{D}_i$ 的条件下与 C_i 条件独立的假设，意味着 $\boldsymbol{Y}_i$与 C_i 在给定（$\boldsymbol{T}_i,\ \boldsymbol{D}_i$）的条件下条件独立。因此可得$E(\boldsymbol{Y}_i|\boldsymbol{T}_i,\ \boldsymbol{D}_i,\ C_i)=E(\boldsymbol{Y}_i|\boldsymbol{T}_i,\ \boldsymbol{D}_i)$, 进而得到 $E\{\boldsymbol{S}_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})|\boldsymbol{T}_i,\ \boldsymbol{D}_i,\ C_i\}=0$。

对应于$S_{\mathrm{eff},\ \alpha}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\boldsymbol{\theta})$的证明类似。根据 C 节 C.1.2，我们知道，对于未删失受试者，对应于 α 的有效得分向量为:

$$S_{\mathrm{eff},\alpha}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta}) = \frac{(\boldsymbol{s}_i-X_i\mathbf{1}_{m_i})^T\{\boldsymbol{Y}_i-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}}{\sigma^2}.$$

因此，我们观察到:

$$\mathrm{E}\{\boldsymbol{S}_{\mathrm{eff},\alpha}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i}\} = \frac{(\boldsymbol{s}_i-X_i\mathbf{1}_{m_i})^T\{\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C_i})-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}}{\sigma^2}.$$

同样，根据我们的假设—— $\boldsymbol{Y}_i$ 与 C_i 在给定$\boldsymbol{D}_i$的条件下条件独立——以及引理 2，我们知道 $\mathrm{E}\{\boldsymbol{S}_{\mathrm{eff},\alpha}(\boldsymbol{Y}_{\mathrm{i}}\ \boldsymbol{D}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C}_{\mathrm{i}}\}=0$。

接下来，回顾 C.1.3 节：

$$S_{\mathrm{eff},\sigma^2}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})=\frac{\frac{1}{2}\left\{\boldsymbol{Y}_i^T\boldsymbol{Y}_i-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\right\}-\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})^T\{\boldsymbol{Y}_i-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}}{\sigma^4}.$$

由此可知

$$\begin{aligned}\sigma^4\mathrm{E}\{\boldsymbol{S}_{\mathrm{eff},\sigma^2}(\boldsymbol{Y}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C}_{\mathrm{i}}\}\ &=\ \frac{1}{2}\left\{\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C}_{\mathrm{i}})-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}^{\mathrm{T}}\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\right\}-\\&\quad\boldsymbol{\zeta}(\boldsymbol{D}_i;\boldsymbol{\theta})^T\{\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}},\mathrm{C}_{\mathrm{i}})-\mathrm{E}(\boldsymbol{Y}_{\mathrm{i}}|\boldsymbol{T}_{\mathrm{i}},\boldsymbol{D}_{\mathrm{i}})\}\end{aligned}$$

同样，由于我们假设 $\boldsymbol{Y}_i$ 与 C_i 在给定 $\boldsymbol{D}_i$ 的条件下条件独立，因此该值为 0。这说明 $\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\boldsymbol{D}_i)$ 的每个分量的均值均为 0，从而证实了命题 1 (A)。如前所述，命题 1 (B) 的证明过程类似（将 $(\boldsymbol{D}_i,\boldsymbol{T}_i)$ 替换$\widehat{\ }$为 $(\boldsymbol{D}_i,\boldsymbol{T}_i^{\mathrm{aug}})$）。

F 定理 3 的证明

再次，我们用 $\boldsymbol{S}^{\mathrm{full}}{}_{\mathrm{eff}}(Y_i,\ \boldsymbol{D}_i,\ \widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})=\Delta_i\boldsymbol{S}_{\mathrm{eff}}(\boldsymbol{Y}_i,\ \boldsymbol{D}_i;\ \boldsymbol{\theta})+(1-\Delta_i)\boldsymbol{S}^*{}_{\mathrm{eff}}(\boldsymbol{Y}_i,\ \widehat{\boldsymbol{D}}_i;\ \boldsymbol{\theta})$ 表示完整的估计函数。由于在命题 1 的条件 (C1) 和 (C2) 下，$\mathrm{E}\left\{\boldsymbol{S}^{\mathrm{full}}{}_{\mathrm{eff}}(\mathrm{Y_i},\ \boldsymbol{D}_{\mathrm{i}},\ \widehat{\boldsymbol{D}}_{\mathrm{i}};\ \boldsymbol{\theta})\right\}=\mathbf{0}$，我们可以使用泰勒定理在 $\boldsymbol{\theta}_0$ 附近展开如下：

$$\begin{aligned}\mathbf{0}\ &=\ \sum_{i=1}^{n}\left\{\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(Y_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}})\right\}\\&=\ n^{-1/2}\sum_{i=1}^{n}\left\{\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(Y_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta}_0)\right\}+\\&\quad n^{-1}\sum_{i=1}^{n}\left\{\frac{\partial}{\partial\boldsymbol{\theta}^T}\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(Y_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta}^*)\right\}n^{1/2}(\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}}-\boldsymbol{\theta}_0),\end{aligned}$$

其中 $\boldsymbol{\theta}^*$ 位于连接 $\boldsymbol{\theta}_{\mathrm{ACE}}$ 和 $\boldsymbol{\theta}_0$ 的直线上。然后，根据 C6，

$$\begin{aligned}\mathbf{0} &= n^{-1/2}\sum_{i=1}^{n}\left\{\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(Y_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta}_0)\right\}+\\ &\quad \mathrm{E}\left\{\frac{\partial}{\partial\boldsymbol{\theta}^{\mathrm{T}}}\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(\mathrm{Y_i},\boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta}_0)\right\}\mathrm{n}^{1/2}(\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}}-\boldsymbol{\theta}_0)+\mathrm{o_p}(1).\end{aligned}$$

因此，

$$\begin{aligned}n^{1/2}(\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}}-\boldsymbol{\theta}_0) &= \left[\mathrm{E}\left\{\frac{\partial}{\partial\boldsymbol{\theta}^{\mathrm{T}}}\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(\mathrm{Y_i},\boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta}_0)\right\}\right]^{-1}\times\\ &\quad n^{-1/2}\sum_{i=1}^{n}\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(Y_i,\boldsymbol{D}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta}_0)+o_p(1)\end{aligned}$$

根据 C7。因此，根据中心极限定理可直接得到，$n^{1/2}(\widehat{\boldsymbol{\theta}}_{\mathrm{ACE}}-\boldsymbol{\theta}_0)$ 依分布收敛于 $\mathrm{Normal}_{p_a+2}(\mathbf{0},\mathcal{V})$，其中 $\mathcal{V}=\mathcal{A}^{-1}\mathcal{B}(\mathcal{A}^{-1})^T$ 满足

$$\begin{aligned}\mathcal{A}^{-1} &= \left[\mathrm{E}\left\{\frac{\partial}{\partial\boldsymbol{\theta}^{\mathrm{T}}}\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(\mathrm{Y_i},\boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta}_0)\right\}\right]^{-1}\\ \mathcal{B} &= \mathrm{Var}\{\boldsymbol{S}_{\mathrm{eff}}^{\mathrm{full}}(\mathrm{Y_i},\boldsymbol{D}_{\mathrm{i}},\widehat{\boldsymbol{D}}_{\mathrm{i}};\boldsymbol{\theta}_0)\}.\end{aligned}$$

G 局限性

现在，我们说明所提出的估计方程存在一个局限性：

引理 3 令 $C(\boldsymbol{M})$ 表示矩阵 $\boldsymbol{M}$ 的列空间，并令 $(\boldsymbol{Z}_i^a)_k$ 表示 $\boldsymbol{Z}_i^a$ 的第 k 列。

A 如果$(\boldsymbol{Z}_i^a)_k$属于 $C(\boldsymbol{Z}_i^b)$，则 $S_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})$ 的第 k 个元素将为0。同样地，如果$(\boldsymbol{s}_i-X_i\mathbf{1}_{m_i})$ 属于 $C(\boldsymbol{Z}_i^b)$，则 $S_{\mathrm{eff},\alpha}(\boldsymbol{Y}_i,\boldsymbol{D}_i;\boldsymbol{\theta})$ 将为0。

B 若$(\boldsymbol{Z}_i^a)_k$属于 $C\{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)\}$，则 $S^*_{\mathrm{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})$ 的第 k 个元素将为0。同理，若$(\boldsymbol{s}_i-\widehat{X}_i\mathbf{1}_{m_i})$ 属于 $C\{(\mathbf{1}_{m_i},\boldsymbol{Z}_i^b)\}$，则 $S^*_{\mathrm{eff},\alpha}(\boldsymbol{Y}_i,\widehat{\boldsymbol{D}}_i;\boldsymbol{\theta})$ 将为0。

引理 3 可直接由 C 节中给出的有效得分向量定义推出。为证明这一点，首先假设 $(\boldsymbol{Z}_i^a)_k$ 属于 $\boldsymbol{Z}_i^b$ 的列空间。由于 $\boldsymbol{P}_{\boldsymbol{Z}_i^b}$ 是投影到 $\boldsymbol{Z}_i^b$ 的列空间上的正交投影算子，因此我们知道 $\boldsymbol{I}_{m_i}(\boldsymbol{Z}_i^a)_k = (\boldsymbol{Z}_i^a)_k = \boldsymbol{P}_{\boldsymbol{Z}_i^b}(\boldsymbol{Z}_i^a)_k$。这意味着 $(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{\boldsymbol{Z}_i^b})(\boldsymbol{Z}_i^a)_k = \mathbf{0}$，等价地，意味着 $(\boldsymbol{Z}_i^a)_k^T(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{\boldsymbol{Z}_i^b}) = \mathbf{0}^T$。因此，$\boldsymbol{Z}_i^{aT}(\boldsymbol{I}_{m_i} - \boldsymbol{P}_{\boldsymbol{Z}_i^b})$ 的第 k 行等于$\mathbf{0}$，所以 $S_{\text{eff},\boldsymbol{\beta}}(\boldsymbol{Y}_i, \boldsymbol{D}_i; \boldsymbol{\theta})$ 的第 k 个元素将为 0。命题 4 是引理 3 的直接推论，它描述了这种情况所引发计算挑战的更广泛情形。

命题 4 设 k 为介于 1 和 p_a之间的固定整数。假设对于所有满足 $\Delta_i = 1$的 i，$(\boldsymbol{Z}_i^a)_k$ 属于$C(\boldsymbol{Z}_i^b)$。进一步假设对于所有满足 $\Delta_i = 0$的 i，$(\boldsymbol{Z}_i^a)_k$属于 $C\{(\mathbf{1}_{m_i}, \boldsymbol{Z}_i^b)\}$。那么，式 (2)中给出的总估计方程的第 k 个元素将为 0。

在较不一般的情形下，如果对于所有 i，$\boldsymbol{Z}_i^a$ 和 $\boldsymbol{Z}_i^b$ 有任何列“相同”，估计方程可能无法收敛。例如，在我们关注的结果模型中同时纳入固定截距和随机截距，就可能出现这种情况。

H 补充表

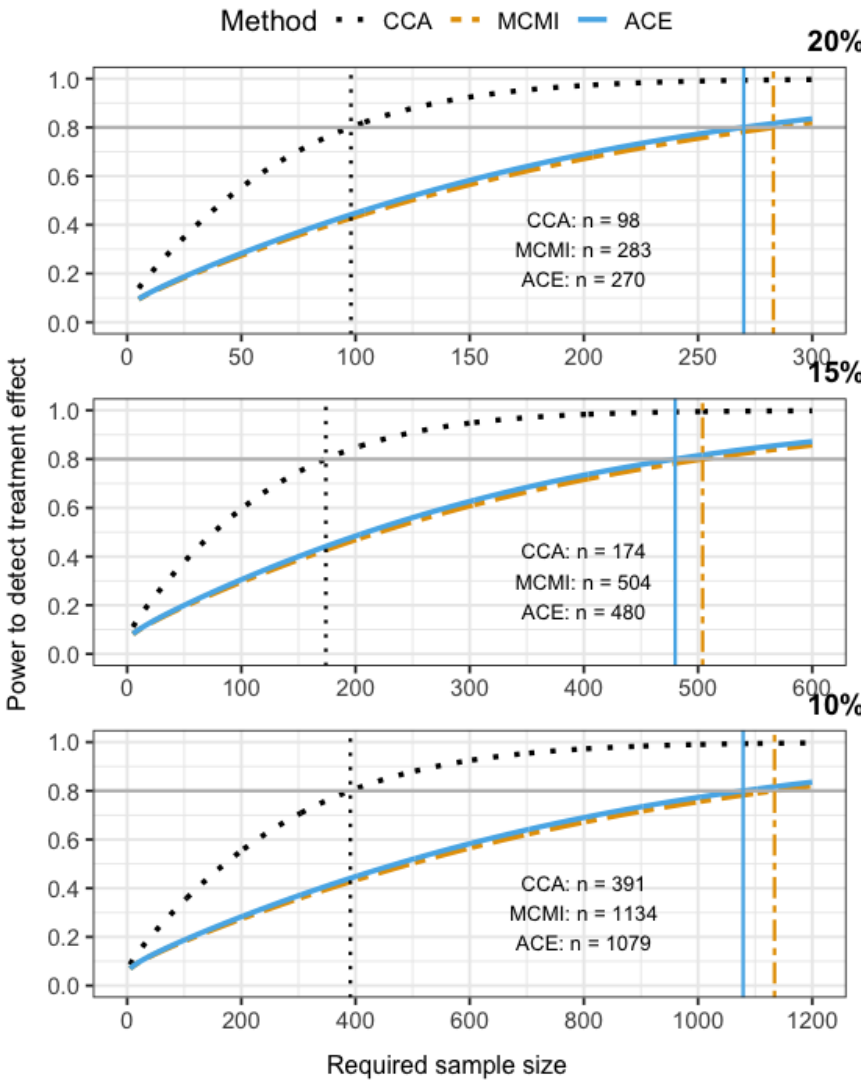

图2：在第一类错误率为 0.05 时，针对给定的每组样本量，检验效能可检测到 10%、15% 和 20% 的治疗效应。结果基于安慰剂组的估计斜率，分别采用完整病例分析（CCA）、多重条件均值插补（MCMI）和 ACE 插补（ACE）得出。

删失	参数	方法	Bias	SEE	ESE	MSE	CPr
轻度	α	ACE	-0.000	0.022	0.022	0.000	0.950
		MCMI	0.173	0.023	0.127	0.046	0.070
	β	ACE	0.000	0.024	0.024	0.001	0.942
		MCMI	0.002	0.036	0.038	0.001	0.951
	σ^2	ACE	-0.003	0.034	0.034	0.001	0.940
		MCMI	1.940		2.979	12.629	

表4：插补模型设定正确时轻度删失下的模拟结果。ACE：将所提出的估计方程应用于插补数据集。MCMI：多重条件均值插补。我们报告经验偏差、平均标准误估计（SEE）、经验标准误（ESE）、平均均方误差（MSE），以及具有名义覆盖率95%的Wald型置信区间的观测覆盖概率（CPr）。

删失	参数	方法	Bias	SEE	ESE	MSE	覆盖率
轻度	α	ACE	0.000	0.022	0.022	0.000	0.946
		MCMI	1.872	0.175	10.429	112.154	0.000
	β	ACE	0.001	0.024	0.024	0.001	0.948
		MCMI	-0.071	0.316	1.289	1.665	0.953
	σ^2	ACE	-0.002	0.034	0.035	0.001	0.944
		MCMI	≥ 100		≥ 100	≥ 100	

表5：插补模型设定错误时轻度删失下的模拟结果。ACE：将所提出的估计方程应用于插补数据集。MCMI：多重条件均值插补。我们报告经验偏差、平均标准误估计值（SEE）、经验标准误（ESE）、平均均方误差（MSE），以及具有名义覆盖率95%的Wald型置信区间的观测覆盖概率（CPr）。

	删失（$\Delta_i = 0$）	未删失（$\Delta_i = 1$）
n	858	244
Age	38.82 (10.43)	43.23 (10.26)
性别 = 男性（%）	311 (36.2)	83 (34.0)
受教育年限	14.64 (2.64)	14.11 (2.55)
CAG	42.22 (2.62)	43.48 (2.81)
CAG-年龄乘积	304.68 (78.74)	389.54 (72.88)
复合统一亨廷顿病评定量表	16.88 (1.78)	15.35 (1.94)
总运动评分	3.79 (4.27)	8.36 (6.71)
符号数字模式测验评分	52.47 (11.14)	44.57 (10.56)
总功能能力	12.83 (0.75)	12.70 (0.80)
Stroop颜色测试得分	79.20 (13.69)	70.95 (13.40)
Stroop词语测试得分	101.01 (17.28)	91.62 (16.38)
Stroop干扰测试得分	46.45 (10.25)	39.55 (9.16)
$\widehat{X}_i$	11.26 (1.35)	4.50 (2.78)

表6：第1次访视时模型协变量（$\boldsymbol{Z}_{ij}^{a}$）和潜在结局（$\boldsymbol{Y}_i$）的汇总。CAG：胞嘧啶-腺嘌呤-鸟嘌呤重复序列的数量。cUHDRS：复合统一亨廷顿病评定量表。SDMT：符号数字模式测试。 $X_i = \Delta_i X_i + (1 - \Delta_i)\mathrm{E}(X_i | X_i > C_i,\ \boldsymbol{V}_i)$。连续变量以均值（标准差）汇总。二元变量以计数（百分比）汇总。