

书面作业按时提交即有 8 分。临近期末公布答案供其他同学复习备考使用，逾期提交扣 1 分，答案公布后、登成绩前提交再扣 1 分（此举意在鼓励同学们按时完成规定的作业题，即便是对答案抄，也有一个强化记忆的过程）。

第 1 次书面作业

2023 年采用的评分细则：第 1 题 (2) (3) 每小问 0.5 分，第 2 题 1 (0.4 + 0.3 + 0.3) 分。

第 1 题

对于回归效应的认识，题干表述不够清晰，很多人理解为全体样本意义下的平均值，得出了“一样高”的答案（在 2023 版本中，这一答案也被视为有效）。

另也有很多人在错误的逻辑框架下“推导”出了符合题目本意的答案。“回归效应的悖论”实质原因是计算均值所使用的数据集不同，因此学生的解题过程中一旦出现 $s > t$ 、 $s < t$ 这样的分类讨论，即可视为无效推理。

第 2 次书面作业

2023 年采用的评分细则：第 1 题每小问 0.4 分，第 2 题每小问 0.4 分，第 3 题 0.4 分。

第 2 题

很多同学 (1) ② 的答案错写为 $\frac{1}{1+e^{-1}}$ ，这是因为遗漏了 Sigmoid 函数在指数项上的负号。看似容易记混，但实际上只需要牢记单调性（激活信号需与原始输入呈递增关系）即可区分。

第 3 题

关于二阶导（Hessian 矩阵），很多人采用了不严谨的符号表述，以

$$\mathbf{x}_n \left(\frac{dy_n}{d\mathbf{w}} \right)^T$$

为例，以下给出两个错误示例，主要错因在于混淆了内积与外积：

$$\mathbf{x}_n \cdot \frac{dy_n}{d\mathbf{w}}; \mathbf{x}_n^T \frac{dy_n}{d\mathbf{w}}; \dots$$

这些符号给出标量 \mathbb{R} ，而不是矩阵 $\mathbb{R}^{M \times M}$ 。还有些同学扭扭捏捏、含糊其辞地省略了乘号或是干脆摆烂跳步，明显看出没有理解矩阵乘法的形状匹配要求。这里建议对矩阵微积分不熟悉的同学按分量形式书写，纯粹做标量微积分，反而更加清晰明了，实际这样操作的同学也都的确拿到了本题的满分。2023 版本中，对这一表述问题所做的处理是扣除 0.1 分。