

研究生挂科情况数据分析报告

雪云风

北京理工大学

2024 年 11 月 4 日





① 引言

② 数据分析

③ 总结

④ 附录

1 引言

2 数据分析

3 总结

4 附录



研究生阶段是学生从学术学习向独立研究转型的重要时期，然而许多学生在适应新的学习和研究环境时，常常遇到困难。学业压力、时间管理能力不足、对课程内容的理解不够深入，以及缺乏有效的学习策略等，都会导致挂科。

在我校（北京理工大学）数学与统计学院中，研究生挂科的现象引人担忧，且研究生挂科可能导致一系列严重后果，包括但不限于延毕、学术压力增加、职业前景受限、心理健康问题。

因此，针对该情况进行数据分析，以探究研究生挂科的原因，并为缓解挂科问题提供指导方案。

对北京理工大学数学与统计学院 2023-2024 学年中出现过挂科情况的研究生成绩单数据（共 12 份，具体见附录）进行数据分析，得到以下三个结论。

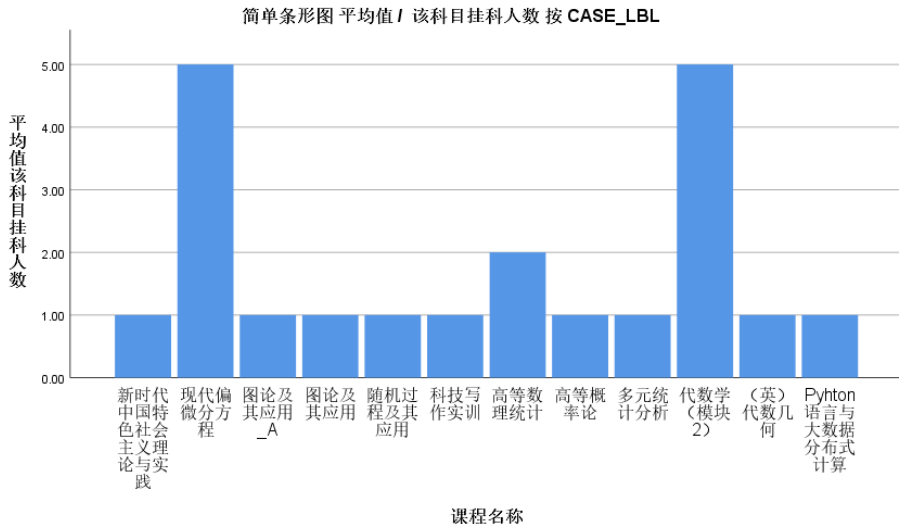
- 结论一：挂科的情况主要集中在专业（必修/选修）课
- 结论二：专业课挂科的情况主要集中于课程现代偏微分方程与代数学（模块二）
- 结论三：专业课挂科人数与平均分存在显著的负相关

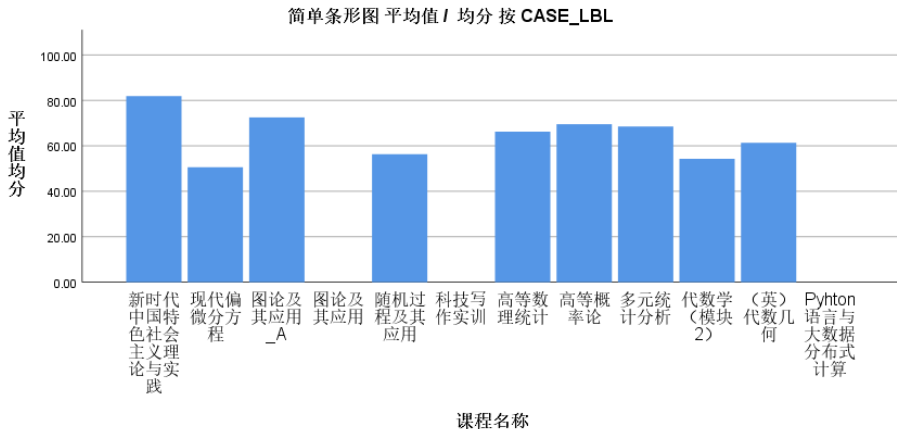
1 引言

2 数据分析

3 总结

4 附录





可以观察到，大部分的挂科情况都集中分布于专业课：将课程分为专业课与非专业课（即公选课）两类，并对课程分类与挂科人数进行独立样本 t 检验。

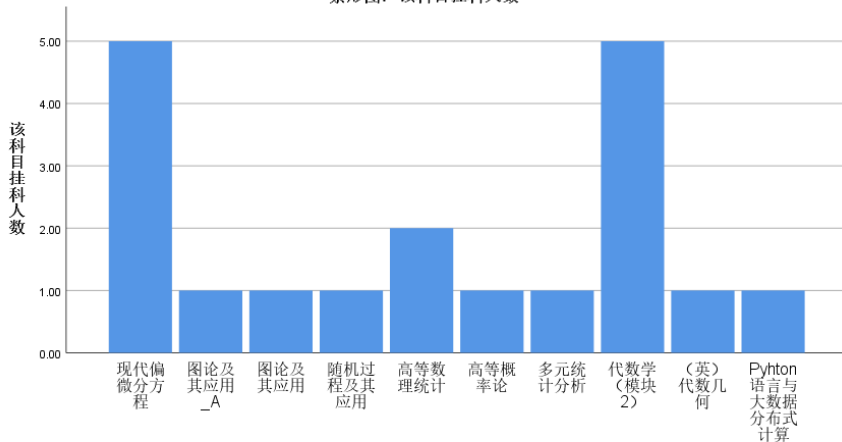
独立样本检验

		莱文方差等同性检验		平均值等同性 t 检验						
		F	显著性	t	自由度	Sig. (双尾)	平均值差值	标准误差差值	差值 95% 置信区间	
该科目挂科人数	假定等方差	4.316	.043	-1.199	47	.236	-.39929	.33294	-1.06907	.27050
	不假定等方差			-1.758	44.872	.086	-.39929	.22719	-.85690	.05833

检验显著性 $p = 0.043 < 0.05$ ，说明专业课和非专业课的平均值存在显著性差异，并得到结论：挂科主要集中在专业课。

对专业课挂科情况做可视化。

条形图：该科目挂科人数



对专业课各科目的挂科人数作 K-S 均匀分布检验

单样本柯尔莫戈洛夫-斯米诺夫检验

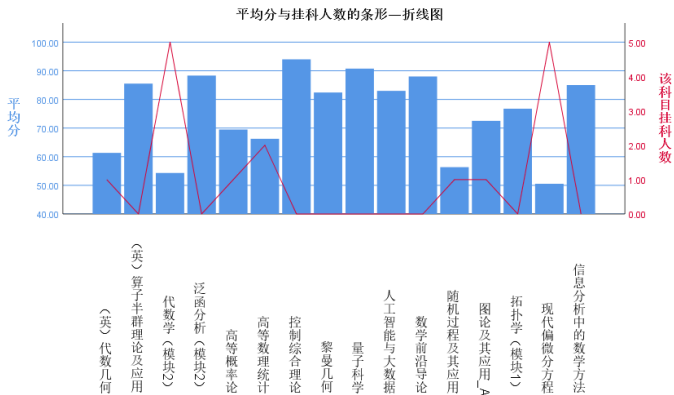
		该科目挂科人数
个案数		35
均匀参数 ^{a,b}	最小值	.00
	最大值	5.00
最极端差值	绝对	.714
	正	.714
	负	-.057
柯尔莫戈洛夫-斯米诺夫 Z		4.226
渐近显著性（双尾）		.000

a. 检验分布为均匀分布。

b. 根据数据计算。

显著性 $p = 0.00 < 0.05$ ，得到结论：
专业课各科目的挂科人数不服从均匀分布，
而更倾向于集中在少数的几个科目上：现代
偏微分方程与代数学（模块二）

由下图，直观推断出科目平均分与挂科人数具有较强的负相关性。



筛选出所有样本量大于 2 的专业课，对平均分与挂科人数做相关性分析（见下表）
得到结论：专业课挂科人数与平均分存在显著的负相关

相关性			
		该科目挂科人数	均分
该科目挂科人数	皮尔逊相关性	1	-.817**
	Sig. (双尾)		.000
	个案数	16	16
均分	皮尔逊相关性	-.817**	1
	Sig. (双尾)	.000	
	个案数	16	16

** 在 0.01 级别（双尾），相关性显著。

(a) 平均分与挂科人的 Pearson 相关性检验

相关性				
			该科目挂科人数	均分
斯皮尔曼 Rho	该科目挂科人数	相关系数	1.000	-.882**
		Sig. (双尾)	.	.000
		N	16	16
	均分	相关系数	-.882**	1.000
		Sig. (双尾)	.000	.
		N	16	16

** 在 0.01 级别（双尾），相关性显著。

(b) 平均分与挂科人的 Pearson 相关性检验

1 引言

2 数据分析

3 总结

4 附录



本文通过数据分析（可视化，卡方检验，t 检验，K-S 检验等方法），得到以下结论

- 结论一：挂科的情况主要集中在专业（必修/选修）课
- 结论二：专业课挂科的情况主要集中于课程现代偏微分方程与代数学（模块二）
- 结论三：专业课挂科人数与平均分存在显著的负相关



针对以上结论，给出缓解我院研究生挂科问题的建议：重点关注均分较低的专业课（尤其是课程现代偏微分方程与代数学（模块二）），对选修这些专业课的同学进行特别提醒，或者采取针对性措施督促/鼓励同学学习这些专业课。

本文中的数据分析存在以下不足：

- 一：仅以课程为变量进行分析，而未对学生分别展开数据分析以得到挂科原因。
- 二：未开展纵向分析，即未对各科目在 2023-2024 学年之前得数据展开分析。
- 三：样本量较小：仅有挂科同学的数据而无挂科同学的数据。

若能获得 2023-2024 学年之前的所有同学的数据以及同学的本科成绩，可以开展进一步的工作：更为细致的分析（对每位同学的学习情况分别分析），并能在时间轴上作出预测。

1 引言

2 数据分析

3 总结

4 附录



数据详情见 https://github.com/xueyunfeng123/Grade_Analysis

谢谢!