

高校课程考试成绩的非均衡多因素方差分析

刘笑涵, 乔 舰

中国矿业大学理学院, 北京

收稿日期: 2024年11月27日; 录用日期: 2024年12月23日; 发布日期: 2024年12月31日

摘 要

本文基于非均衡数据多因素方差分析理论对高校某公共基础课程考试成绩进行了统计分析。文中先对相关方差分析理论基础进行了综述, 然后基于R软件对某课程考试的真实数据集进行了详细的方差分析研究, 给出了丰富的图表展示, 结果表明本文所设定的两个自变量的交互影响不显著, 但学院类别和学生性别对平均成绩影响显著。基于此分析结果, 本文给出了一些教学指导建议。

关键词

非均衡数据多因素方差分析, 课程考试成绩, 离差平方和

Multi-Factor ANOVA of Unbalanced Data on College Course Test Scores

Xiaohan Liu, Jian Qiao

School of Science, China University of Mining and Technology, Beijing

Received: Nov. 27th, 2024; accepted: Dec. 23rd, 2024; published: Dec. 31st, 2024

Abstract

This paper makes a statistical analysis on college course test scores based on the theory of Multi-factor ANOVA of unbalanced data. Firstly, the theoretical basis of relevant ANOVA is summarized, then based on R software, the detailed ANOVA process for a real data sets of a course exam is conducted, a rich chart shows is given, the result shows that the interaction influence of the two independent variables is not significant, but the college category and student gender have a significant influence on the average score. Based on the analysis results, this paper gives some teaching guidance suggestions.

Keywords

Multi-Factor ANOVA of Unbalanced Data, Course Test Scores, Sum of Squares of Deviations

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

课程考试成绩是学生学习能力的集中体现,是课程教学效果的重要体现,是教学总结的重要手段。对课程成绩进行统计分析是教学效果评价的重要组成部分,对指导教学、提高教学质量具有重要意义。以往对于考试成绩的统计分析,大多数集中在对于考试成绩的描述性统计[1][2]、基于多元统计分析的考试成绩分析[3][4]、基于回归分析的成绩分析[5][6]。鉴于非均衡数据多因素方差分析理论基础较为复杂¹,以往使用方差分析进行考试成绩统计分析多使用单因素方差分析[7],本文对多因素方差分析理论基础进行了综述并基于 R 软件对某高校某公共基础课程考试成绩进行了统计分析。

2. 理论基础

作为一种统计分析模型,方差分析需要样本数据满足一定的假设条件。尽管相关假设条件中的一条或几条不满足时,仍可技术性的进行相关统计分析,但分析结果可能会给出错误结论。接下来给出方差分析三个假设条件、检验方法以及相关假设条件不满足时的处理方式。

假设一:样本数据的独立性:一般样本数据的独立性是通过实验设计或实验条件的控制来实现的,而非通过正式统计检验;只需考虑每个单元内每个样本点是否与其它样本点有关联或每个样本点是否对其它样本点造成影响即可。例如若观测数据是对某个个体的重复观测,则应考虑相关性。

假设二:样本数据的正态性:方差分析作为特殊的回归分析模型,要求模型残差应服从或渐近服从正态分布²。样本数据的正态性可通过对相应回归模型残差的可视化 QQ 图来实现,也可基于每个单元内样本数据的正态性 shapiro-wilk 检验进行。当单元数较多或单元内样本量较小时,前者更加适宜。如果数据本身或经过变换(对数变换、平方根变换等)后仍不渐近服从正态分布,可通过 kruskal-wallis 非参数检验进行方差分析。

假设三:样本数据的方差齐性:不同单元内样本数据方差相等称为方差齐性。方差齐性可通过可视化的 boxplot 图的离散性进行,也可以通过 levene Test 检验或 Bartlett 检验进行;若方差齐性不满足,可通过 welch test 检验进行方差分析³。

数据是否均衡对应方差分析理论有着本质区别,究其原因在于非均衡数据多因素方差分析中,作为自变量的类别变量之间存在共线性[9],不再正交,进而自变量对因变量变异的解释存在重复。将非均衡数据多因素方差分析单独看待比将其作为均衡数据方差分析的拓展要更加合理[10]。

对于非均衡数据多因素方差分析,基于各因素及交互项离差平方和计算方式的不同,借用 SAS 软件

¹按照单因素不同水平或多因素不同水平组合分组后每个单元内样本量是否相等,方差分析可划分为均衡数据和非均衡数据两种类型。数据均衡与否对于单因素方差分析没有影响。

²方差分析对于样本数据的非正态性具有一定的稳健性[8],即方差分析对样本数据正态性要求并不十分苛刻,只要每个单元内样本容量比较大,峰度绝对值小于 10 且偏度绝对值小于 3,仍可进行方差分析。

³Bartlett 检验适用于服从正态分布的样本数据; leveneTest 检验既适用于服从正态分布,也适用于服从非正态分布或分布不明的样本数据; welch test 检验不需要方差齐性假设而 kruskal-wallis 检验既不需要正态性假设亦不需要方差齐性假设。

的标识, 可分为类型 I、II、III 及 IV 四种类型, 其中类型 IV 专门用于某些单元内样本容量为零的特殊情形, 当每个单元内样本容量均不小于 1 时, 类型 III、IV 结果一样[11]。以含交互项的两因素方差分析为例, 它们对于各因素及交互项离差平方和的计算, 分别如图 1~3 所示, 各因素及交互项离差平方和可通过添加相应因素及交互项的回归模型的残差平方和的差值进行计算。可以看到无论采用哪种类型, 相同样本数据的交互项离差平方和与误差平方和相同, 区别在于各因素对应离差平方和的计算。类型 I 方差分析与各因素进入模型的顺序有关, 先进入模型的因素会将自身与其它因素共有的离差平方和占为己有, 后进入因素只能得到去除与先进入因素共有离差平方和后的剩余离差平方和⁴; 类型 II 方差分析在计算各因素对应离差平方和时, 将自身与其它因素共有的离差平方去除, 但将交互项离差平方和占为己有; 类型 III 方差分析在计算各因素对应离差平方和时, 会将自身与其它因素以及交互项共有的离差平方全部去除。

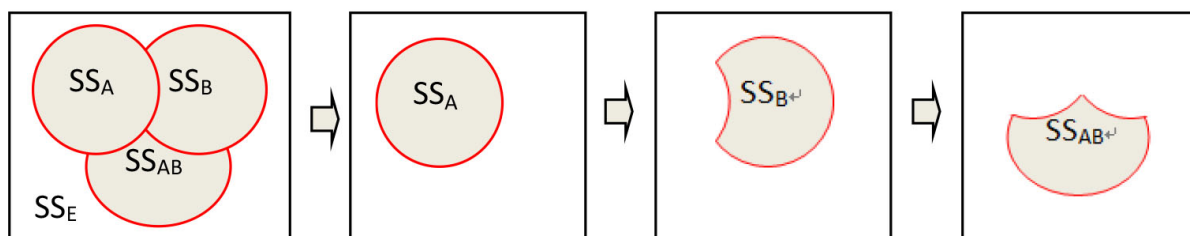


Figure 1. Type I analysis of variance calculation of factorial variation

图 1. 类型 I 方差分析各因素离差平方和计算

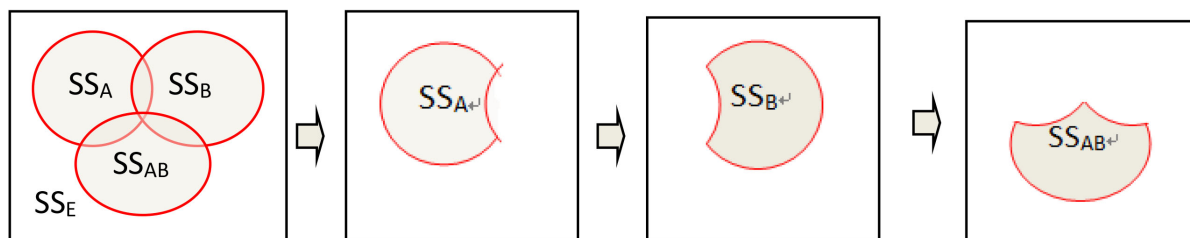


Figure 2. Type II analysis of variance calculation of factorial variation

图 2. 类型 II 方差分析各因素离差平方和计算

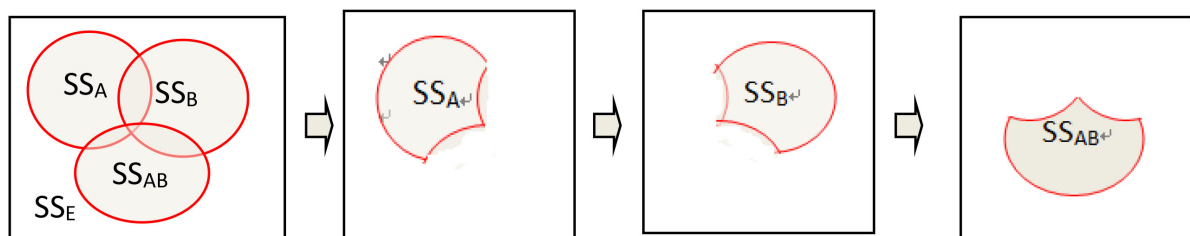


Figure 3. Type III analysis of variance calculation of factorial variation

图 3. 类型 III 方差分析各因素离差平方和计算

对于非均衡数据多因素方差分析类型的选择, 目前比较一致的结论是: 若实际分析中研究因素之间确实存在主次之分, 主次因素共有离差平方和应分配给主因素, 应选择类型 I; 若理论或实际分析中确认研究因素之间不存在交互效应, 应选择类型 II, 反之若交互效应存在, 则应选择类型 III⁵。由于类型 III 方差分析在计算各因素对应离差平方和时, 只计算各因素对应的“纯”效应, 而类型 I、II 在各因素对应离

⁴ 只有类型 I 方差分析满足总的离差平方和等于各因素及交互项离差平方和与误差平方和之和。

⁵ 类型 II、III 的选择涉及交互效应存在与否的检验, 该检验存在争议, 进而导致类型 II、III 的选择存在争议[12]。

差平方和中混杂了其它因素的影响, 故而受到包括 SAS、SPSS、STATISTICA 等众多统计软件的青睐, 将其作为缺省选项[9]。开源的 R 软件对于非均衡数据多因素方差分析比较包容, 类型 I 是 R 软件主程序包命令“aov”的缺省选项, 类型 II、III 可通过程序包“car”的程序“Anova”实现。

多因素方差分析运行中需注意: (1) 先进行交互项的显著性检验。若交互项显著, 无需再进行主效应分析, 应进行固定其中一个因素在某一个水平条件下, 另一个因素不同水平的效应之差是否显著的简单主效应分析或交互作用对照分析; 若交互项不显著, 则应基于类型 II 方差分析进行主效应的显著性检验。(2) 鉴于多因素方差分析模型是一个过参数模型, 相应模型的参数估计必须限制未知参数个数, 作为处理方式之一的各因素水平的对照编码对于非均衡数据多因素方差分析至关重要, 尤其对于类型, 若不进行各因素水平的对照编码处理, 结果将与理论结果大相径庭, 但对于其它两种类型影响不大。

3. 考试成绩分析

本文原始数据集(名称: gs)来源于某大学 2020 级选修概率论与数理统计课程的 1600 名本科生(不含降级、参军、休(退)学、转专业等情况)考试成绩, 其中学生卷面考试成绩对应变量 cj 是一个定量变量, 百分制; 学生所属学院对应变量 xy 是一个定性变量, 取值 A、B、……、G, 表示 7 个学院; 学生性别对应变量 xb 是一个定性变量, 取值 M、F, 分别表示男性、女性。

原始数据的描述性统计如表 1 所示, 对应因变量为 cj。相应散点图、协同图、箱线图、豆图分别如图 4~7 所示。对应图像代码见附录。

Table 1. Descriptive statistics of raw data
表 1. 原始数据描述性统计

xb	xy	均值	标准差	样本量
F	A	79.25	12.29	28
	B	80.31	11.70	39
	C	74.45	13.07	67
	D	72.41	13.89	98
	E	78.37	13.28	86
	F	79.02	13.42	41
	G	74.91	14.98	117
M	A	73.09	14.16	112
	B	73.14	12.42	100
	C	73.08	14.13	182
	D	68.92	13.84	138
	E	75.53	15.76	314
	F	73.13	14.19	226
	G	73.35	14.42	52

从散点图可以看到不同学院学生人数差距较大, 数据具有右拖尾现象; 由协同图可知女同学平均成绩比男同学高, 男同学考试成绩更加集中一些; 由箱线图可知学生平均成绩差距较大但四分位间距差距不大, 且 A 学院的女同学平均成绩最高, D 学院男同学平均成绩最低, 男同学平均成绩较女同学要低一些, 存在个别离群值; 由豆图可知各单元考试成绩所服从的分布都是单峰的, 部分单元考试成绩极差较大, 部分单元存在个别离群值。

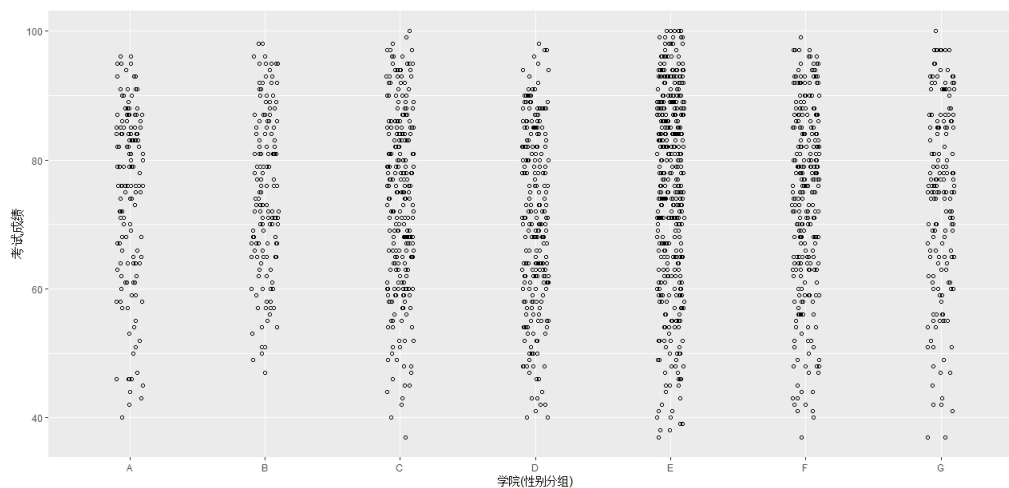


Figure 4. Scatter plot
图 4. 散点图

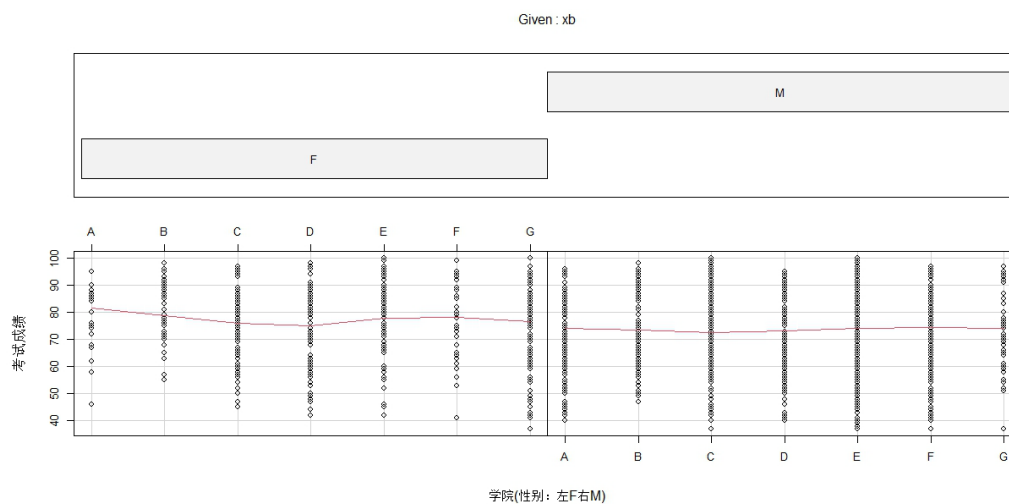


Figure 5. Collaborative diagram
图 5. 协同图

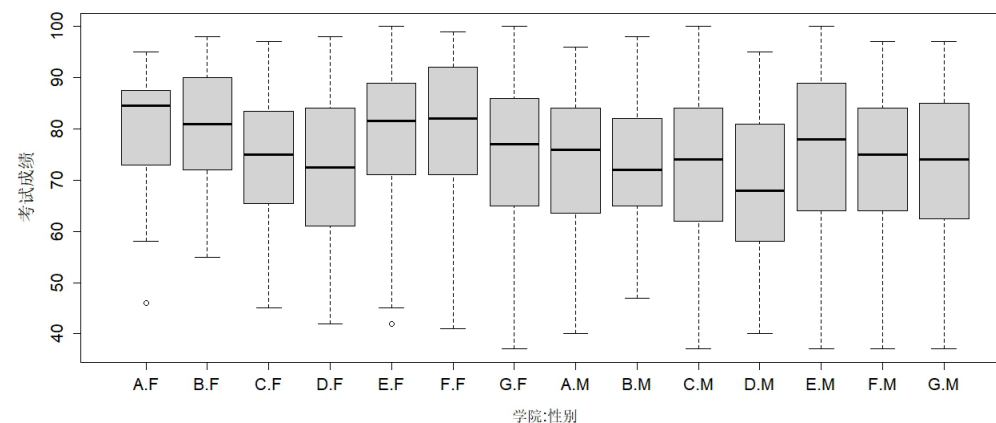


Figure 6. Box-plot
图 6. 箱线图

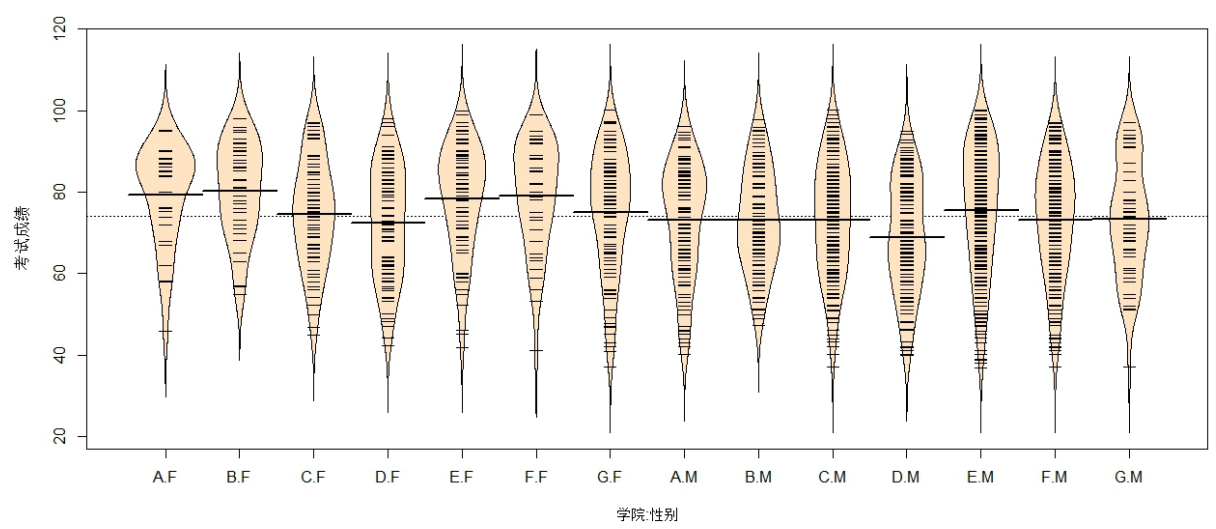


Figure 7. Bean-plot
图 7. 豆图

接下来进行正态性检验，图 8 为相应回归模型残差的 QQ 正态性检验图示，可以看到模型残差并未完全处于拟合直线上，小部分甚至超出了相应置信区间，可以确定模型残差并不完全服从正态分布。按照自变量交互项(xy: xb)分组后每个单元样本的 Shapiro-Wilk 正态性检验结果如表 2，原假设是对应单元样本服从正态分布。在 5%的显著性水平下，只有两个单元内的样本不拒绝原假设，进一步证实了以上 QQ 正态性检验的结论。尝试进行因变量的各种函数变换，都未能满足正态分布的要求。考虑到方差分析对于样本数据非正态性的稳健性，继续进行接下来的检验过程。

Table 2. Shapiro-Wilk normality test for each unit sample
表 2. 各单元样本的 Shapiro-Wilk 正态性检验

xy	xb	w	p
A	F	0.8951	0.0088
	M	0.9523	0.0005
B	F	0.9539	0.1109
	M	0.9823	0.2014
C	F	0.9803	0.3645
	M	0.9823	0.0208
D	F	0.9734	0.0437
	M	0.9734	0.0085
E	F	0.9497	0.0021
	M	0.9579	0.0000
F	F	0.9398	0.0310
	M	0.9714	0.0002
G	F	0.9568	0.0008
	M	0.9690	0.1913

可以看到无论是原始数据图 4~7，还是模型残差正态性检验图 8，都显示原始数据集中存在个别离群

值, 方差分析前进行一下离群值检验, 结果显示本数据集对应样本点不存在离群值。进一步可给出离群值、高杠杆值、强影响点的综合图示, 结果如图 9 所示: 该图横坐标的两条垂线分别对应平均帽子值的 2 倍和 3 倍, 纵坐标的三条水平线分别对应模型标准化残差取值-2, 0 和 2。一般来说纵坐标绝对值大于 2 的样本点都是离群值; 纵坐标大于 3 倍平均帽子值的样本点都是高杠杆值点; 圆圈面积正比于影响程度, 面积越大的点对模型拟合造成的影响越强。

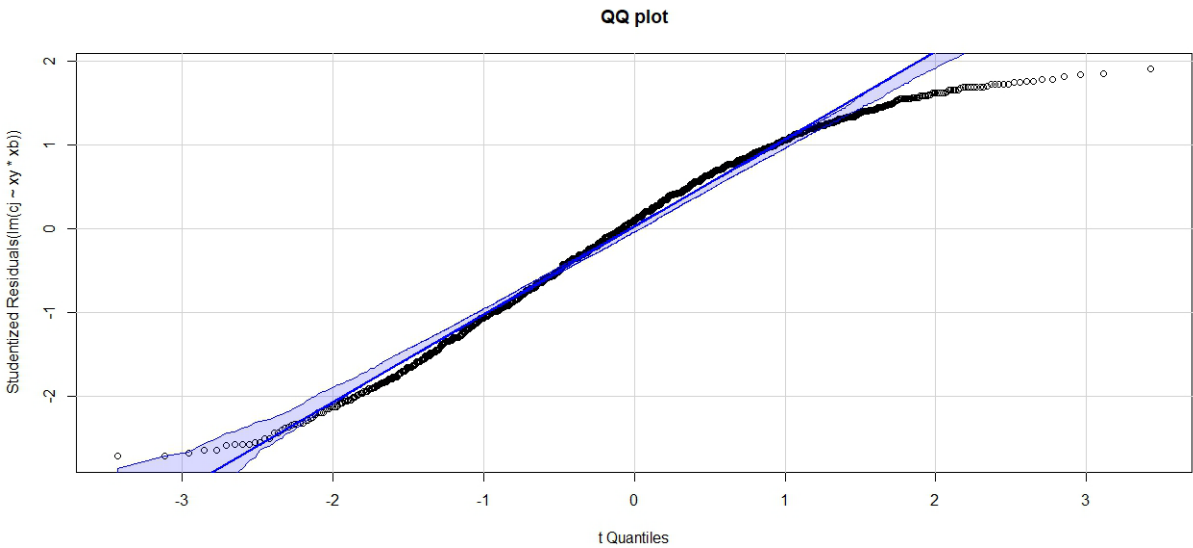


Figure 8. QQ normality test for model residuals
图 8. 模型残差的 QQ 正态性检验

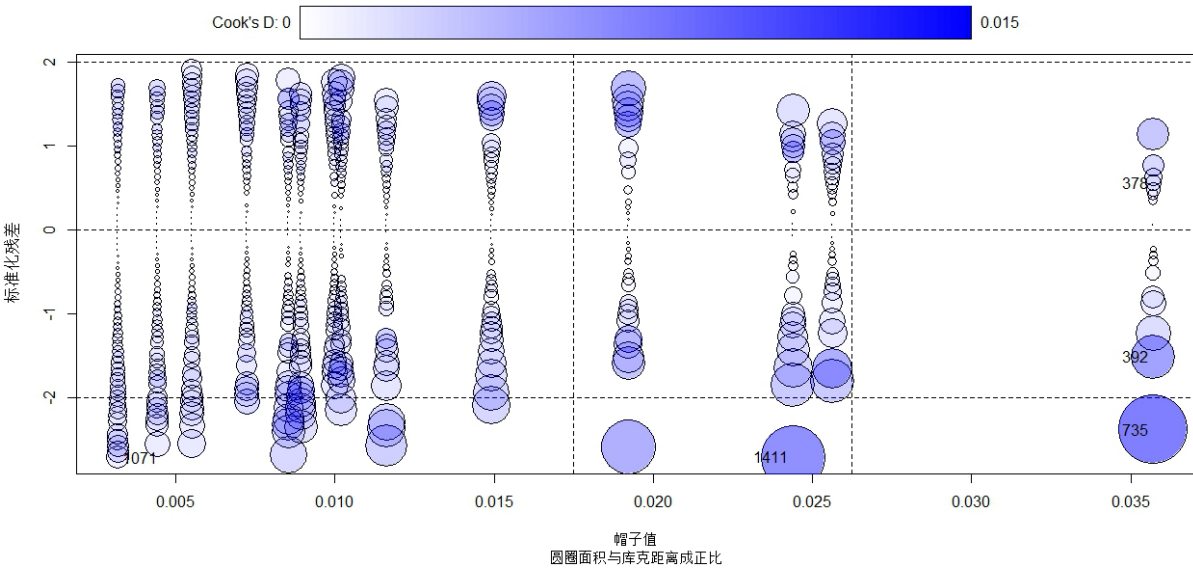


Figure 9. Influence plot for sample data
图 9. 样本数据影响点图示

接下来进行方差齐性检验, 原假设是各单元样本同方差。两种不同检验方法得到的结果如表 3 所示: 在 5% 的显著性水平下, 各单元样本数据不拒绝原假设, 即各单元样本方差齐性假设成立。图 5 箱线图中四分位间距相差不大也证实了这点。

Table 3. Test for homogeneity of variances after grouping
表 3. 分组后样本数据的方差齐性检验

分组依据	检验依据	自由度	检验统计量	p
xy:xb	leveneTest	13	1.5813	0.0835
	Bartlett	13	16.837	0.2069

综上，方差分析的三个假设条件，本文样本数据基本上都满足，接下来进行方差分析。图 10 给出了各单元平均成绩的置信度为 95%的置信区间的折线图，可以看到不同学院、不同性别学生的平均成绩确实存在着一定差异。图 11 给出了不同性别学生平均成绩的折线图，两线近乎平行没有交点，可初步认定学院(xy)与性别(xb)两个变量之间不存在交互作用，同院系女同学比男同学平均成绩高，不同院系学生平均成绩存在差异。

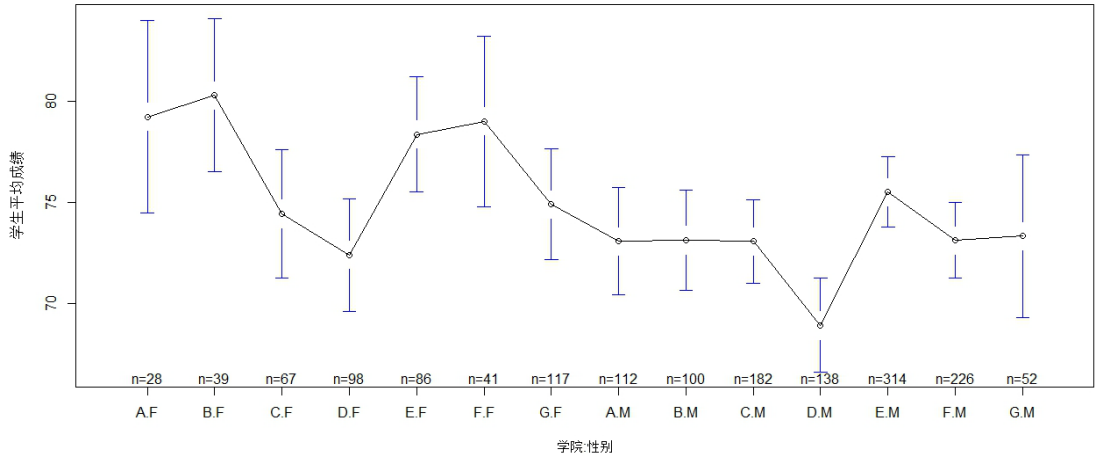


Figure 10. Line chart of 95% confidence intervals for mean values of each unit sample
图 10. 各单元样本均值的 95%置信区间折线图

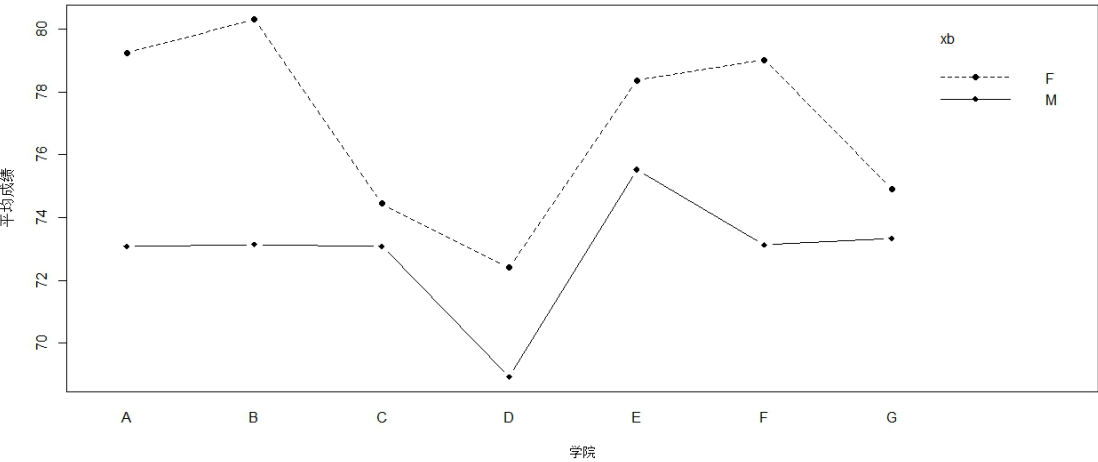


Figure 11. Line chart of average grades by gender
图 11. 不同性别学生平均成绩折线图

接下来通过三种不同平方和类型进行本文所对应的非均衡样本数据的方差分析。结果如表 4 所示。三种类型的方差分析都显示按交互项(xy: xb)分组后对应学生平均成绩差异并不明显，即两个自变量的交

互作用不显著。类型方法显示不同学院(xy)或不同性别(xb)对应学生平均成绩差异显著，但该结果依赖于自变量进行回归模型的顺序，若回归模型自变量设置为 $xb*xy$ ，即性别变量(xb)先进入回归模型，则性别变量作用改变明显($F = 12.2464, p = 0.000479$)，学院变量作用也随着改变。类型方法结果给出的最终结论与类型方法一致，但学院变量(xy)更加显著，性别变量(xb)及交互项效应如本文前述，与类型方法结果完全一样。类型方法与类型方法结果基本一致，原因在于类型方法运行中假定交互项不显著，而本文中交互项确实不显著，两者结果差异源于不显著不代表不存在。

尽管本文中交互项(xy: xb)不显著，为使结果更加可靠可信，本文将交互项仍保留在方差分析模型中，将类型方法对应方差分析结果作为最终结论：给定 5%的显著性水平，变量学院(xy)、性别(xb)对学生平均成绩影响显著，但两者的交互项(xy: xb)对学生平均成绩并无显著影响。

Table 4. Analysis of variance results
表 4. 方差分析结果

方差来源		平方和	自由度	F	p
学院(xy)	类型 I	5239	6	4.3146	0.000252
	类型 II	6557	6	5.4	1.52 E-05
	类型 III	6238	6	5.1368	3.07 E-05
性别(xb)	类型 IIII	3797	1	18.7594	1.58 E-05
	类型 III	4277	1	21.1326	4.63 E-06
交互项(xy: xb)		1120	6	0.9226	0.4775
残差		320982	1586		

Table 5. Mean values corresponding to each level of the variable
表 5. 各变量水平对应均值

变量	加权均值	标准差	非加权均值	标准差	样本量
xb = F	76.01	13.84	76.96	0.74	476
xb = M	73.28	14.54	72.89	0.49	1124
xy = A	74.32	13.99	76.17	1.50	140
xy = B	75.15	12.60	76.72	1.34	139
xy = C	73.45	13.84	73.77	1.02	249
xy = D	70.37	13.94	70.66	0.94	236
xy = E	76.14	15.29	76.95	0.87	400
xy = F	74.04	14.21	76.08	1.21	267
xy = G	74.43	14.78	74.13	1.19	169

接下来进行变量学院(xy)、性别(xb)的主效应分析。首先给出两个变量不同水平各自的加权均值、非加权均值及相应标准差，如表 5 所示，可以看到对于本文样本数据非加权均值一般都比加权均值要大，标准差都较小；加权均值中男同学成绩的标准差大于女同学。由于性别(xb)是二分类变量，且三种类型的方差分析都已表明性别(xb)对学生平均成绩影响显著，且无论加权还是非加权均值都表明女同学比男同学平均成绩高，无需再进行其它检验。对于学院(xy)变量，加权均值的两两成对比较的显著性检验(原假设是对应两个学院学生加权平均成绩无差异)结果如表 6 左半部分所示，可以看到 D 学院与 E 学院、B 学院学生平均成绩差异显著，综合看来 D 学院学生平均成绩与其它学院差异显著，图 12 证实了这一点；图 13

给出了各学院不同性别学生平均成绩的所有可能线性组合之间差异的显著性检验结果：E 学院学生平均成绩最高，与 E 学院学生平均成绩差异显著的是 D 学院；其它学院与 D、E 学院无显著差异。对于非加权均值的两两成对比较的显著性检验(原假设是对应两个学院学生非加权平均成绩无差异)结果如表 6 右半部分所示，结果与加权均值的两两成对比较结果基本一致。

Table 6. Pairwise comparison test of means across different colleges
表 6. 不同学院均值成对比较检验

研究对象	差值	置信下限	置信上限	p 值	研究对象	差值	标准差	p 值
B-A	0.8297	-4.1990	5.8583	0.9990	A-B	-0.5542	2.02	1
C-A	-0.8716	-5.3081	3.5648	0.9974	A-C	2.4046	1.81	0.8402
D-A	-3.9528	-8.4330	0.5274	0.1252	A-D	5.5054	1.77	0.0317
E-A	1.8211	-2.3030	5.9451	0.8505	A-E	-0.7823	1.73	0.9994
F-A	-0.2840	-4.6663	4.0983	1.0000	A-F	0.0911	1.93	1
G-A	0.1046	-4.6949	4.9041	1.0000	A-G	2.0436	1.91	0.9375
C-B	-1.7013	-6.1479	2.7454	0.9191	B-C	2.9588	1.68	0.5777
D-B	-4.7824	-9.2727	-0.2921	0.0282	B-D	6.0596	1.64	0.0042
E-B	0.9914	-3.1436	5.1265	0.9922	B-E	-0.2281	1.6	1
F-B	-1.1136	-5.5062	3.2790	0.9894	B-F	0.6453	1.81	0.9998
G-B	-0.7250	-5.5340	4.0839	0.9994	B-G	2.5978	1.79	0.7742
D-C	-3.0812	-6.8965	0.7342	0.2059	C-D	3.1009	1.38	0.2746
E-C	2.6927	-0.6974	6.0828	0.2234	C-E	-3.1869	1.34	0.2047
F-C	0.5877	-3.1123	4.2876	0.9992	C-F	-2.3135	1.58	0.7652
G-C	0.9762	-3.2095	5.1619	0.9932	C-G	-0.361	1.56	1
E-D	5.7739	2.3267	9.2210	0.0000	D-E	-6.2877	1.28	<0.0001
F-D	3.6688	-0.0835	7.4211	0.0602	D-F	-5.4143	1.53	0.0075
G-D	4.0574	-0.1747	8.2894	0.0702	D-G	-3.4618	1.51	0.2501
F-E	-2.1050	-5.4240	1.2139	0.4991	E-F	0.8734	1.49	0.9972
G-E	-1.7165	-5.5695	2.1366	0.8450	E-G	2.8259	1.47	0.4639
G-F	0.3886	-3.7397	4.5168	1.0000	F-G	1.9525	1.69	0.911

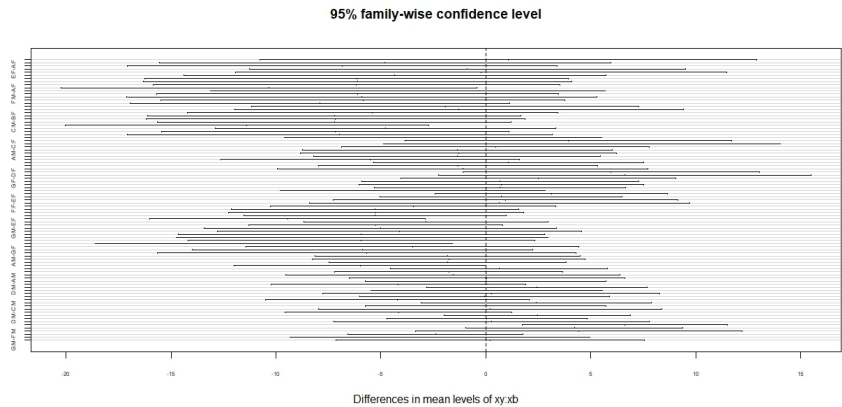


Figure 12. Pairwise comparison chart of weighted means across different colleges
图 12. 不同学院加权均值成对比较图示

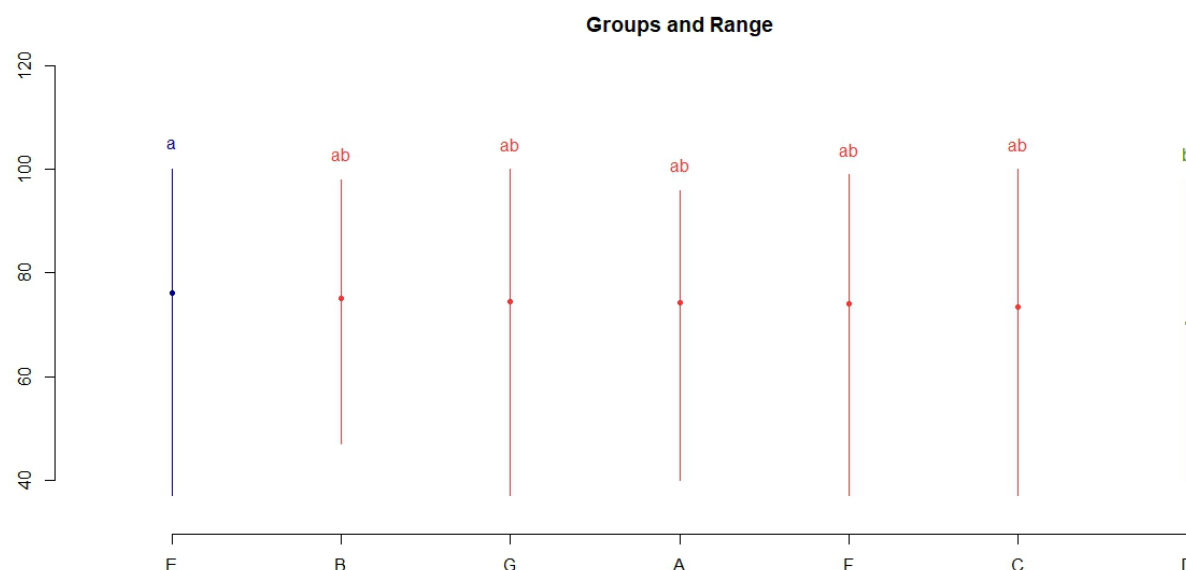


Figure 13. Composite comparison chart of weighted means across different colleges

图 13. 不同学院加权均值组合比较图示

4. 结论

本文采用非均衡数据两因素方差分析研究了高校学院类别及学生性别对平均成绩的影响。通过箱线图、豆图甄别了离群值的存在性；通过整体模型残差的 QQ 检验图和每个单元样本的 Shapiro-Wilk 检验了样本数据的正态性；通过 LeveneTest 检验和 Bartlett 检验了每个单元样本的方差齐性。结果显示，本研究每个单元样本数据具有独立性、渐近正态性和方差齐性，且不存在显著的离群值。

三种不同类型的非均衡数据方差分析都表明本文两个自变量的交互影响不显著，但学院类别和学生性别对平均成绩影响显著。对于性别变量的结论是无论加权还是非加权均值，无论哪个学院，女生平均成绩都显著高于男生；对于学院变量的成对检验结论是无论加权还是非加权均值，D 学院与 E 学院、D 学院与 B 学院平均成绩差异显著；各学院不同性别学生平均成绩的所有可能线性组合差异显著性检验表明 E 学院学生平均成绩最高，与 E 学院学生平均成绩差异显著的是 D 学院；其它学院与 D、E 学院无显著差异。

基于本文的统计分析结果，可给出以下教学指导建议：(1) 尽管不同学院的生源略有不同，学生基础有些差别，但不同学院学生的学风乃至班风也是十分重要的，平均成绩显著较低的学院应加强学生的学风管理，积极引导学生形成奋发向上、刻苦学习、勤于思考、踏实好问，并互相促进、互相鼓励的学习氛围；平均成绩显著较低的学院应积极向成绩较高的学院学习，学习人家的学生学习管理方式，借鉴人家好的经验做法；另外教务管理部门在分配任课教师时，应将那些教学经验丰富、教学能力突出、善于激发学生学习能力的教师向这些平均成绩显著较低的学院倾斜。(2) 高校中女生平均成绩显著高于男生，男生成绩两极分化更加严重，女生在课程学习方面的自我调控能力确实比男生强，这是共识，也是现实。这充分说明了分层次教学的必要性，任课教师在教学中应多给男同学一些鼓励、正面评价，积极树立良好的典型、榜样，同时应积极引导男生提高自己的自律能力，忧患意识；还有课程考试应更加多元化，增加一些需要动手能力、分析实际问题能力的题目，以使教育更加贴近生活，面向未来。

基金项目

本文受中国矿业大学(北京)校级课程建设与教改项目(编号：J230704)支持。

参考文献

- [1] 董倩倩. 基于 spss 的《高等数学》期末考试成绩评价与预测[J]. 教育现代化, 2019, 6(85): 242-243+252.
- [2] 何庆光. 学生考试成绩的统计分析[J]. 广西财政高等专科学校学报, 2004(2): 89-91.
- [3] 黄培丽. 多元统计在高校学生成绩统计分析中的应用研究[D]: [硕士学位论文]. 西安: 西安建筑科技大学, 2013.
- [4] 王政宇, 潘晋孝, 赵勇. 多元统计在考试成绩评定中的应用[J]. 华北工学院学报, 2000, 21(3): 277-279.
- [5] 孙毅, 刘仁云, 王松, 等. 基于多元线性回归模型的考试成绩评价与预测[J]. 吉林大学学报: 信息科学版, 2013, 31(4): 404-408.
- [6] 谭立云, 张昆龙, 李慧. 基于 R 语言的数学基础课成绩多元线性回归分析[J]. 黑龙江科技信息, 2014(32): 97.
- [7] 常广炎. 用 Matlab 实现高校课程考试成绩方差分析[J]. 电脑编程技巧与维护, 2015(22): 89-91.
- [8] Mcdonald, J.H. (2009) Handbook of Biological Statistics. 2nd Edition, Sparky House.
- [9] Smith, C.E. and Cribbie, R. (2014) Factorial Anova with Unbalanced Data: A Fresh Look at the Types of Sums of Squares. *Journal of Data Science*, **12**, 385-403. [https://doi.org/10.6339/JDS.201407_12\(3\).0001](https://doi.org/10.6339/JDS.201407_12(3).0001)
- [10] Searle, S. (1987) Linear Models for Unbalanced Data. John Wiley and Sons Inc.
- [11] 段重阳, 陈平雁. 固定效应方差分析中四种离差平方和的定义与探讨[J]. 中国卫生统计, 2013, 30(4): 470-475.
- [12] Langsrud, Ø. (2003) ANOVA for Unbalanced Data: Use Type II Instead of Type III Sums of Squares. Kluwer Academic Publishers.

附录

本文所用程序命令:

```
Gs <- read.table('clipboard', header = T): #导入数据
Softbao <- c('beanplot', 'car', 'ggplot2', 'gplots', 'multcomp', 'lsmeans', 'agricolae')
Sapply(softbao, require, character.only = TRUE): #载入相应软件包, 需提前安装
Attach(gs): #将数据集 gs 作为当前数据集
Xy <- factor(xy): #将定性变量因子化
Xb <- factor(xb)
With(gs, tapply(cj, list(xy, xb), mean)): #各单元学生平均成绩
With(gs, tapply(cj, list(xy, xb), sd)): #各单元学生成绩标准差
With(gs, tapply(cj, list(xy, xb), length)): #各单元学生样本量, 对应表 1
ps1 <- ggplot(gs, aes(x = xy, y = cj, group = xb)) + geom_point(cex = 1.5, pch = 1.0, position = position_jitter(w = 0.1, h = 0))
p1 <- ps1 + xlab("学院(性别分组)") + ylab("考试成绩"): #散点图, 对应图 4
coplot(cj~xy|xb, panel = panel.smooth, xlab = "学院(性别: 左 F 右 M)", ylab = "考试成绩"): #协同图, 对应图 5
boxplot(cj~xy: xb, cex.lab = 1.3, cex.axis = 1.3, xlab = "学院: 性别", ylab = "考试成绩"): #箱线图, 对应图 6
beanplot(cj~xy: xb, col = "bisque", method = "jitter", xlab = "学院: 性别", ylab = "考试成绩"): #豆图, 对应图 7
qqPlot(lm(cj~xy*xb), id = F, simulate = TRUE, main = 'QQ plot'): #正态性 QQ 检验, 对应图 8
by(cj, xy: xb, shapiro.test): #各单元样本的 shapiro 检验, 对应表 2
fit0 <- lm(cj~xy*xb, data = gs): #回归分析
outlierTest(fit0): #离群点检测;
influencePlot(fit0, id = TRUE, xlab = "帽子值", ylab = "标准化残差", sub = "圆圈面积与库克距离成正比"): #影响点图, 对应图 9
leveneTest(cj~xy:xb): #方差齐性检验
bartlett.test(cj~ interaction(xy, xb)): #方差齐性检验, 对应表 3
plotmeans(cj~interaction(xy, xb), p = 0.95, use.t = TRUE, xlab = "学院: 性别", ylab = "学生平均成绩"): #各单元学生成绩均值置信区间图, 对应图 10
interaction.plot(xy, xb, cj, type = 'b', legend = T, pch = c(16, 18), xlab = "学院", ylab = "平均成绩"): #不同性别学生平均成绩折线图, 对应图 11
anova1 <- aov(cj~xy*xb): #方差分析模型
anova(anova1): #类型 I 方差分析
Anova(anova1, type = 2): #类型 II 方差分析, 对应表 4
Options(contrasts = c("contr.Sum", "contr.Poly")): #对照编码
Anova(anova1, type = 3): #类型 III 方差分析
Anova2 <- aov(cj~xb*xy): #改变自变量进入模型次序
Anova(anova2): #类型 I 方差分析
```

```
TukeyHSD (anova1, 'xy', conf. level = 0.95): #不同学院加权均值成对差异 Tukey-Kramer 检验; 对非  
均衡数据, Tukey-Kramer 检验等同 TukeyHSD 检验, 对应表 6 左半部分  
Plot (TukeyHSD (anova1), cex. Axis = 0.5): #成对差异 Tukey-Kramer 检验图示, 对应图 12  
Print (lsmeans (anova1, list (pairwise~xy)), adjust = c ( "tukey" )): #不同学院非加权均值差异成对检验,  
对应表 6 右半部分  
ubtest <-scheffe. test (y = anova1, trt = "xy" ): #非均衡数据不同学院学生成绩加权均值组合比较  
plot (ubtest): # scheffe 检验图示, 对应图 13
```