

# 土木工程专业本科生课程成绩多元统计分析

阎利

(安阳工学院 建筑工程系,河南 安阳 455000)

**摘要:**以114名土木工程专业本科生的13门课程成绩为样本,应用SAS软件分别进行正态性检验、线性相关分析和主成分分析,认为结构力学、混凝土结构设计和高等数学等课程成绩最能代表该专业学生的总体学习效果,并对改进土木工程专业的教学工作提出了建议。

**关键词:**土木工程;课程成绩;主成分分析;SAS

**中图分类号:**G642.47

**文献标志码:**A

**文章编号:**1005-2909(2010)04-0118-04

本科生在校期间要学习数10门课程,每门课程的考试成绩直观地反映着学生对该课程掌握的程度。对于课程成绩的统计分析,一般是在期末考试结束后分课程进行,并从中获取有价值的信息,用以改进以后的教学工作。但是,这些不同课程成绩之间的相关性如何?哪些课程最能体现学生的总体成绩?如何更加客观公正地评价学生4年来的综合学习效果?笔者以某大学2005级土木工程专业114名本科生在校期间13门主干课程的考试成绩为样本,应用SAS统计软件的UNIVARIATE、PRINCOMP等过程进行分析和讨论<sup>[1]</sup>,以期揭示这些枯燥分数背后所隐藏着的深刻教学规律,为进一步提高土木工程专业的教育教学质量提供决策参考。

## 一、样本描述与基本统计分析

某大学2005级土木工程专业入学时共有124名本科生,入学后先后有10名学生退学或留级,因此本文选取114名顺利完成学业学生的课程成绩作为研究对象。该专业4年来先后开设了30余门课程,经过土木工程专业视角的初步遴选,确定以13门主干课程成绩 $X_i(i=1, 2, \dots, 13)$ 作为变量。其中,公共课方面选取大学英语( $X_1$ )、高等数学( $X_2$ )和大学物理( $X_3$ ),专业基础课方面选取画法几何与工程制图( $X_4$ )、理论力学( $X_5$ )、材料力学( $X_6$ )、结构力学( $X_7$ )、土力学( $X_8$ )、房屋建筑学( $X_9$ )、土木工程材料( $X_{10}$ ),专业课方面选取混凝土结构设计( $X_{11}$ )、钢结构基本原理( $X_{12}$ )和土木工程施工( $X_{13}$ )。上述114名学生13门主干课程成绩的基本统计分析见表1。

由表1可以看出,绝大多数课程成绩的均值在70分左右(仅结构力学均值

收稿日期:2010-05-25

基金项目:安阳工学院教学改革研究项目——工科大学生实践能力培养途径与方法研究(2009)

作者简介:阎利(1975-),男,安阳工学院建筑工程系副教授,博士,主要从事土木工程和工程管理研究,

(E-mail) yanli@tju.edu.cn.  
欢迎访问重庆大学期刊社 <http://qks.cqu.edu.cn>

略低于 60 分),标准差一般在 10 分左右(仅结构力学的标准差略高于 20 分),偏斜度值接近于 0(表明分布近似于左右对称)且一般为负值(表明在均值左侧的数据更分散),峰度值多数较小(越接近于 0 正态性越好,大于 0 表示数据中含有较多远离均值的数据,小于 0 表示均值两侧极端数据较少),W 统计量均接近于 1(越接近于 1 正态性越好),所选样本的正态性总体较好。

### 二、线性相关分析

线性相关分析可通过计算各门课程成绩的相关系数,分析学生某门课程掌握情况对其他课程掌握情况的影响程度<sup>[2]</sup>。由表 2 可以看出,在统计的 13 门课程中,有 12 门课程成绩间的相关系数为正值(正相关),仅画法几何、工程制图(X4)与土力学(X8)的相关系数为负值。这表明,多数课程成绩间

具有较强的可比性,即某门课程成绩的高低,可由另外一门课程成绩的高低分布规律大致推断,学生某门课程的成绩高,其他课程的成绩一般也较高。

以结构力学(X7)和混凝土结构设计(X11)为例,其相关系数最高(0.8053),这表明结构力学掌握较好的学生,混凝土结构设计掌握得也较好。另外,土木工程专业开设的多数专业基础课和专业课对数理逻辑思维能力依赖性较重,而高等数学(X2)与绝大多数课程的相关系数在 0.4 以上,这有力地验证了在大学一年级学好高等数学对土木工程专业本科生有深远的影响。画法几何与工程制图(X4)与多数课程成绩的相关系数较小,说明该课程的性质与其他课程存在较显著差异,教师在合理引导学生转变思维方式、改进学习方法等方面还应进一步加强。

表 1 课程名称对应的变量及各门课程成绩的基本统计分析

变量	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>
课程名称	大学英语	高等数学	大学物理	画法几何与工程制图	理论力学	材料力学	结构力学	土力学	房屋建筑学	土木工程材料	混凝土结构设计	钢结构基本原理	土木工程施
均值	72.6	76.3	76.6	78.6	74.1	75.6	59.5	79.8	72.8	68.7	74.7	72.6	67.1
标准差	6.63	12.2	8.62	9.20	11.0	11.5	20.6	9.13	11.5	8.21	10.3	12.3	10.7
变异系数	9.134312	15.94524	11.25098	11.71291	14.83576	15.1425	34.66451	11.43496	15.76695	11.95312	13.74711	16.83073	15.88063
偏斜度	-0.62735	-0.59487	-0.11043	-0.03363	-0.69248	-0.23297	-0.55066	-1.15979	-0.76112	0.645274	-0.4072	-1.48696	-0.84853
峰度	0.517718	0.812047	-0.46825	-0.58405	1.509601	-0.36417	-0.18465	3.478727	1.330407	-0.00138	0.051789	4.2397	1.494989
W 统计量	0.965572	0.958015	0.97273	0.965577	0.962701	0.975522	0.95036	0.934282	0.967224	0.917548	0.968752	0.904547	0.958132
p 值	0.0442	0.0079	0.1735	0.0442	0.0237	0.2711	0.0011	0.0001	0.0622	0.0001	0.0841	0.0001	0.0082

表 2 各变量间的线性相关系数矩阵

	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>	X <sub>5</sub>	X <sub>6</sub>	X <sub>7</sub>	X <sub>8</sub>	X <sub>9</sub>	X <sub>10</sub>	X <sub>11</sub>	X <sub>12</sub>	X <sub>13</sub>
X <sub>1</sub>	1.0000	0.5641	0.5946	0.1953	0.6115	0.3294	0.5243	0.2194	0.6118	0.4908	0.5064	0.2802	0.3352
X <sub>2</sub>	0.5641	1.0000	0.5559	0.4504	0.5426	0.4267	0.5976	0.3167	0.4712	0.3562	0.5526	0.4264	0.4265
X <sub>3</sub>	0.5946	0.5559	1.0000	0.0931	0.6066	0.2906	0.5434	0.3052	0.4889	0.3536	0.4193	0.3559	0.3518
X <sub>4</sub>	0.1953	0.4504	0.0931	1.0000	0.1751	0.3417	0.2149	-0.1024	0.2177	0.1783	0.1564	0.1577	0.1596
X <sub>5</sub>	0.6115	0.5426	0.6066	0.1751	1.0000	0.2645	0.4526	0.1925	0.3969	0.3561	0.3909	0.2289	0.3199
X <sub>6</sub>	0.3294	0.4267	0.2906	0.3417	0.2645	1.0000	0.6196	0.2750	0.5205	0.4356	0.6298	0.4361	0.4509
X <sub>7</sub>	0.5243	0.5976	0.5434	0.2149	0.4526	0.6196	1.0000	0.5013	0.5518	0.4504	0.8053	0.5932	0.5272
X <sub>8</sub>	0.2194	0.3167	0.3052	-0.1024	0.1925	0.2750	0.5013	1.0000	0.3173	0.2804	0.5549	0.4848	0.3283
X <sub>9</sub>	0.6118	0.4712	0.4889	0.2177	0.3969	0.5205	0.5518	0.3173	1.0000	0.4974	0.5594	0.3996	0.3990
X <sub>10</sub>	0.4908	0.3562	0.3536	0.1783	0.3561	0.4356	0.4504	0.2804	0.4974	1.0000	0.5254	0.3438	0.3703
X <sub>11</sub>	0.5064	0.5526	0.4193	0.1564	0.3909	0.6298	0.8053	0.5549	0.5594	0.5254	1.0000	0.6427	0.5270
X <sub>12</sub>	0.2802	0.4264	0.3559	0.1577	0.2289	0.4361	0.5932	0.4848	0.3996	0.3438	0.6427	1.0000	0.4812
X <sub>13</sub>	0.3352	0.4265	0.3518	0.1596	0.3199	0.4509	0.5272	0.3283	0.3990	0.3703	0.5270	0.4812	1.0000

### 三、主成分分析

主成分分析法是把多个指标化为少数几个综合指标并基本保留所含信息,进而通过对新指标的比

较达到分析目的多元统计方法<sup>[3,4]</sup>。借助 SAS 软件由相关阵出发进行主成分分析,输出结果见表 3、表 4。

表3 各主成分的特征值、方差、贡献率和累积贡献率

主成分 (PRIN)	特征值	方差	贡献率	累积贡献率
PRIN1	6.09494	4.69294	0.468842	0.46884
PRIN2	1.40200	0.22778	0.107846	0.57669
PRIN3	1.17422	0.35953	0.090325	0.66701
PRIN4	0.81469	0.19265	0.062668	0.72968
PRIN5	0.62204	0.07890	0.047849	0.77753
PRIN6	0.54314	0.05806	0.041780	0.81931
PRIN7	0.48508	0.03824	0.037314	0.85662
PRIN8	0.44684	0.05703	0.034372	0.89100
PRIN9	0.38981	0.04293	0.029985	0.92098
PRIN10	0.34688	0.07081	0.026683	0.94766
PRIN11	0.27607	0.03046	0.021236	0.96890
PRIN12	0.24561	0.08692	0.018893	0.98779
PRIN13	0.15869	-	0.012207	1.00000

表3给出了各主成分的特征值、方差、贡献率和累积贡献率。其中,第一主成分(PRIN1)的贡献率高达46.9%,前7个主成分的累积贡献率已达85.7%。若按照85%的判断标准,只需前7个主成分就能很好地概括这组数据的信息。另外,这13个主成分特征值的算术平均值 $\bar{\lambda}$ 为1.00000,若按照大于 $\bar{\lambda}$ 的特征值个数确定主成分个数的标准,则仅需前3个主成分就能较好地概括这组数据的信息,此时前3个主成分的累积贡献率达到66.7%,基本满足评价要求。

表4给出了前7个主成分的特征矢量。第一主成分对应的第一特征矢量的各个分量值均为正值,数值上相差不大且均在0.13以上,可以认为第一主成分PRIN1代表学生的总体学习水平,即综合学习成绩。其中,分量值最大的3个变量对应的课程分

别是:结构力学( $X_7$ )、混凝土结构设计( $X_{11}$ )和高等数学( $X_2$ ),这三门课程的成绩最能代表学生综合学习成绩的好坏。这三门课程又分别属于专业课、专业基础课和公共课,因此可进一步推断,高等数学( $X_2$ )成绩最能代表学生公共课学习的效果,结构力学( $X_7$ )成绩最能代表学生专业基础课学习的效果,而混凝土结构设计( $X_{11}$ )成绩最能代表学生专业课掌握的情况。第一主成分的得分可由如下公式计算得出(其他主成分的分析 and 计算方法以此类推):

$$\begin{aligned} \text{PRIN1} = & 0.292212x_1 + 0.306458x_2 + \\ & 0.278143x_3 + 0.130036x_4 + 0.255147x_5 + \\ & 0.277839x_6 + 0.347155x_7 + 0.217362x_8 + \\ & 0.299754x_9 + 0.259549x_{10} + 0.342526x_{11} + \\ & 0.270499x_{12} + 0.262016x_{13} \end{aligned}$$

式中, $x_i$ 代表某名学生的 $X_i$ 课程的考试成绩。

表4 前7个主成分对应的特征矢量

	PRIN1	PRIN2	PRIN3	PRIN4	PRIN5	PRIN6	PRIN7
$X_1$	0.292212	0.352733	-0.214955	-0.187285	0.079595	-0.018073	0.168544
$X_2$	0.306458	0.228866	0.114187	0.390544	0.128339	0.127446	0.000545
$X_3$	0.278143	0.257315	-0.353692	0.196959	-0.044429	-0.153762	0.145021
$X_4$	0.130036	0.325723	0.705073	0.218162	0.193426	0.242500	0.104091
$X_5$	0.255147	0.405387	-0.275996	0.161371	-0.147894	0.067467	-0.425829
$X_6$	0.277839	-0.139414	0.384109	-0.231439	-0.011799	-0.466756	-0.382289
$X_7$	0.347155	-0.147453	0.014134	0.098134	0.070450	-0.225405	-0.232269
$X_8$	0.217362	-0.455753	-0.282346	0.197980	0.388192	0.255201	-0.053622
$X_9$	0.299754	0.084793	-0.006211	-0.377024	0.153561	-0.336011	0.560150
$X_{10}$	0.259549	0.023150	0.013074	-0.621162	-0.009000	0.642312	-0.137447
$X_{11}$	0.342526	-0.258458	0.026063	-0.047796	0.110553	-0.060362	-0.211884
$X_{12}$	0.270499	-0.363442	0.073955	0.245221	-0.006620	0.155474	0.398024
$X_{13}$	0.262016	-0.176207	0.092404	0.084703	-0.851005	0.096697	0.139459

由此可以计算全部学生的选定数量主成分(如主成分的特征值作为权重,计算出某学生所以课程上文提及的选取前三个或前七个)的得分,并根据各成绩的综合分值,进而对全部学生4年来的学习效

果进行综合排序(见表5)。显然,主成分1-3给出的信息与主成分1-7提供的信息十分接近,即使是第一主成分给出的信息,也已满足应用需求。可见,该方法与简单地进行算术和排序方法相比,更具科学性和客观性,提供的信息量也更加丰富。

四、结语

本文将统计分析的思想与实际教学需要相结

合,通过数据分析的方法,从枯燥的课程成绩中挖掘出多条有价值的信息,如主干课程之间的相关关系,结构力学、混凝土结构设计、高等数学等课程对于土木工程专业学生的的重要意义,画法几何与工程制图与其他课程性质的差异等。以这些分析结果作参考,应用于土木工程专业今后的教育教学改革工作中,将产生良好的效果。

表5 以114名学生课程成绩的主成分1-7得分为依据进行排序的结果(仅列出前10名和后5名)

序号	学号	PRIN1	PRIN2	PRIN3	PRIN4	PRIN5	PRIN6	PRIN7	主成分 1-7得分	主成分 1-3得分	主成分 1得分
1	2005051059	5.4886	0.69828	0.28351	-0.30175	0.34804	0.36213	-0.26804	34.80190790	34.76457936	33.45268768
2	2005052048	5.3135	0.12729	-1.63607	-0.24948	-0.60727	-0.00047	-0.4046	29.86530442	30.64281815	32.38546369
3	2005051015	4.06626	1.42938	0.68687	-0.16654	0.60757	0.32452	0.37791	28.19596872	27.59413798	24.78361072
4	2005051023	4.27362	0.33488	0.26051	0.13045	1.03478	-0.76574	-0.44363	26.94170609	26.82285530	26.04745748
5	2005052047	4.09591	0.34856	-0.2243	-0.6603	0.11396	0.44536	-0.40296	24.76900213	25.18962927	24.96432570
6	2005052059	4.19228	0.019	-0.96943	-0.17802	-0.68437	1.34992	-0.51487	24.35271475	24.44000897	25.55169506
7	2005052053	3.73641	1.05292	-0.07699	0.39736	-0.71711	0.07346	0.15284	24.15067821	24.15898541	22.77319477
8	2005051010	3.65613	0.05453	1.8434	-1.26967	0.59359	0.39582	-0.08806	24.03201999	24.52490119	22.28389298
9	2005051061	3.51261	1.39587	0.44927	-0.52273	0.40982	-0.25081	0.48567	23.82212414	23.89369875	21.40914719
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
110	2005052045	-4.60129	1.03784	0.13438	-0.8187	0.03929	0.33473	0.1344	-26.82728986	-26.43174311	-28.0445865
111	2005052024	-4.31064	-2.76819	1.16998	0.49176	0.7947	0.30641	-0.16071	-27.79684716	-28.78028063	-26.2730922
112	2005051062	-4.72254	0.89777	-0.85818	-1.42524	0.01138	2.4705	-0.30046	-28.49058625	-28.53261653	-28.7835979
113	2005051047	-5.87084	0.61674	0.75972	-0.44618	-0.86107	0.70523	-0.44292	-34.75660103	-34.02566965	-35.7824175
114	2005051009	-8.54578	3.84644	0.52183	-0.13156	-0.3524	-1.74502	-1.60456	-48.13308189	-46.08056425	-52.0860164

参考文献:

[1]高惠璇.实用统计方法与SAS系统[M].北京:北京大学出版社,2002.  
 [2]吴津津,郑海鹰.用统计方法讨论大学一年级课程成绩对后继课程成绩的影响[J].温州师范学院学报(自然科学版).2003(5):66-70.  
 [3]苏斌,谢友芹.统计分析在学生成绩评估中的应用[J].系统工程理论与实践.2006(7):134-140.  
 [4]刘海生.多元统计分析法在学生成绩综合评价中的应用[J].华北科技学院学报.2002(3):77-79.

Multivariate Statistical analysis in the course scores of civil engineering majored students

YAN Li

(School of Civil Engineering, Anyang Institute of Technology, Anyang 455000, P. R. China)

**Abstract:** With a sample of 13 course scores of 114 students majored in civil engineering, statically analysis such as normality test, linear correlation analysis and principal component analysis was applied via SAS system. It was held that structural mechanics, concrete structure design and higer mathematics are the most important courses for civil engineering majored students. Improved recommendations for civil engineering education were also proposed.

**Keywords:** civil engineering; course score; principal component analysis; SAS

(编辑 梁远华)