

2013 第十届五一数学建模联赛

承 诺 书

我们仔细阅读了五一数学建模联赛的竞赛规则。

我们完全明白，在竞赛开始后参赛队员不能以任何方式（包括电话、电子邮件、网上咨询等）与本队以外的任何人（包括指导教师）研究、讨论与赛题有关的问题。

我们知道，抄袭别人的成果是违反竞赛规则的，如果引用别人的成果或其它公开的资料（包括网上查到的资料），必须按照规定的参考文献的表述方式在正文引用处和参考文献中明确列出。

我们郑重承诺，严格遵守竞赛规则，以保证竞赛的公正、公平性。如有违反竞赛规则的行为，我们愿意承担由此引起的一切后果。

我们授权五一数学建模联赛赛组委会，可将我们的论文以任何形式进行公开展示（包括进行网上公示，在书籍、期刊和其他媒体进行正式或非正式发表等）。

我们参赛选择的题号为（从 A/B/C 中选择一项填写）：_____ A _____

我们的参赛报名号为：_____ 2013200650 _____

参赛组别（研究生或本科或专科）：_____ 本科 _____

所属学校（请填写完整的全名）_____ 西南交通大学 _____

参赛队员（打印并签名）：1. _____ 黄博 _____

2. _____ 秦万家 _____

3. _____ 吴丝雨 _____

日期：_____ 2013 _____ 年 _____ 5 _____ 月 _____ 1 _____ 日

获奖证书邮寄地址：_____ 四川成都高新区西部园区西南交通大学 _____ 邮政编码 _____ 611756 _____

2013 第十届五一数学建模联赛

编 号 专 用 页

竞赛评阅编号（由竞赛评委会评阅前进行编号）：

评阅记录

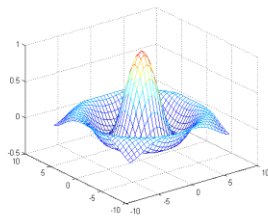
评 阅 人						
评 分						
备 注						

_____裁剪线_____裁剪线_____裁剪线_____

竞赛评阅编号（由竞赛评委会评阅前进行编号）：

参赛队伍的参赛号码：（请各参赛队提前填写好）：2013200650

2013 第十届五一数学建模联赛



题 目 A 题：大学生体质健康评价问题

摘 要

评价大学生体质健康状况一般是通过体质健康测试进行评价，并且对大学生体质健康的评价将为如何提高体质健康水平有现实指导意义。体质健康测试项目包括身高体重、肺活量、台阶试验、立定跳远、握力或坐位体前屈五项。

本文通过对某高校大一新生 36 个班级共 1000 多名学生的体质与健康测试结果数据进行分析，采用频数分析法、相关性分析、箱线图分析法、聚类分析法、单因素方差分析法以及主成分分析法，借助统计软件 SPSS 和数据软件 Excel，分析了体重与其他测试项目的相关性及其影响，运用箱线图和聚类分析法找出并剔除测试结果中的异常数据，建立了基于生源地和学生体质健康状况的单因素方差分析模型，确定了生源地与学生体质健康状况有无显著性差异，最后通过补充相关信息，结合主成分分析法建立权重模型，确定了体质健康评分公式，建立体质健康评价模型，提出对不达标学生如何在校期间提高自身身体素质的建议报告书。

针对问题一，首先，采用体质指数 BMI 作为体重对体质健康状况影响的分析指标，根据 SPSS19.0 对体质指数 BMI 进行频数分析的运行结果对学生进行分组，在不同组别中，对体重与其他各项大学生体质健康指标进行相关性分析，得出不同体重级别对学生的体质健康状况的影响程度。然后，对 1、2、3 班同学的测试数据进行描述性分析，绘制箱线图，找出各项数据中的离群值（即异常值），检验测试结果的正确性和准确性。

针对问题二，以 1、2、3 班剔除异常数据后的测试数据为样本，根据生源地对比各项体质健康测试项目结果用 SPSS 软件进行聚类分析，得到测试结果聚类树状图，而将测试结果分成 8 个等级，与八个地区无显著的相关性。然后，建立单因素方差分析模型，根据 Levene 检验法和 SPSS19.0 的运行结果得出显著性值均大于显著性水平，即无显著性影响。

针对问题三，对 1 班同学的体质健康测试数据进行主成分分析，根据 SPSS 软件运行结果得出主成分矩阵和方差解释表，并结合相关权重计算公式，建立权重模型，确定各体质测试项目所占的权重，计算体质测试各项单项得分和总分，对比《国家体质健康标准》对学生体质进行体质健康评价。

针对问题四，查阅身体机能和运动机能方面相关资料，根据人体运动中新陈代谢规律、心肺功能训练原理和力量训练设计基本原则，为不同不合格的项目设计相对应的训练方式和强度。利用权重模型确定的各项目权重和体质健康测试分数，计算分配对应的训练方式和强度。

关键词：相关性分析 箱线图 聚类分析 单因素方差分析 主成分分析 权重模型

一、问题重述

评价大学生体质健康状况一般是通过体质健康测试进行评价,体质健康测试主要包括身体形态、身体机能、身体素质等方面,主要项目包括身高体重、肺活量、台阶试验、立定跳远、握力(男生)或坐位体前屈(女生)五项,根据各项测试结果计算得分,然后加权求和得到其总分,对比《国家体质健康标准》对学生体质进行评价。由于测试过程中学生未能按照要求规范测试,导致测量结果中出现一些偏差,进而影响了体质健康的测试结果的正确性和准确性。影响大一新生的体质健康状况的因素很多,其中体重是体现体质健康状况的重要指标,生源地也是影响体质健康状况的因素之一。对大学生体质健康的评价将为如何提高体质健康水平有现实指导意义,对于未达标的大一新生来说,就如何让学生在在校期间提高自身的体质健康写一份建议报告书,其中包括提高体质健康水平的措施和手段,如何量化提高体质健康指标等问题。附表1为某高校大一新生36个班级共1000多名学生的体质与健康测试项目和结果,附表2为大一新生各项测试评分标准,请尝试建立数学模型讨论下列问题:

- 1、根据附表1中测试结果分析体重对体质健康的影响。
- 2、根据附表1中测试结果,建立数学模型检验测试结果的正确性和准确性,找出附表1中1、2、3班同学的可能偏差测试结果。
- 3、分析生源地与体能健康是否具有显著差别?
- 4、建立体质健康评价模型,评价该校学生的体质健康状况,并对1班的30名同学进行体质健康评价。
- 5、对于未达标的大一新生来说,就如何让学生在在校期间提高自身的体质健康写一份建议报告书(其中包括提高体质健康水平的措施和手段,如何量化提高体质健康指标等问题)。

二、问题分析

(1) 问题一, 体重是体现体质健康状况的重要指标, 由于肺活量、握力、立定跳远、坐位体前屈、台阶试验的测试结果都受身高体重的影响, 我们要分析体重对体质健康的影响, 首先要将学生划分为正常和非正常两组, 划分的标准采用体质指数(BodyMass Index, BMI)进行分组, 通过 Excel 计算产生 BMI 值。并采用 SPSS19.0 对 BMI 进行频数分析、对体重及各项测试项目与大学生体质健康指标进行相关分析, 得出不同体重级别对学生的体质健康状况的影响程度。测试结果偏差的分析, 先对1、2、3班同学的测试数据进行描述性分析, 根据计算所得均值、方差、置信区间等值可初步判断出测试结果的准确性, 绘制箱线图, 找出各项数据中的离群值(即可能测试偏差结果)。

(2) 本问题先应用聚类分析法, 建立聚类分析的数学模型, 以1、2、3班剔除异常数据后的测试数据为样本, 利用 SPSS19.0 对样本进行聚类分析, 得到测试结果聚类树状图, 这样就可以从图中看出样本测试结果的等级, 对比生源地区的划分, 判断生源地与体能健康是否具有显著差别。再利用单因素方差分析法, 建立单因素方差的数学模型, 采用的是 Levene 检验法, 计算出显著性值与显著性水平的关系, 判断生源地的差异对体重、身高、肺活量、台阶测试、握力(坐位体前屈)和立定跳远的指标没有显著性影响。

(3) 问题三, 首先, 根据基于主成分分析的权重确定方法, 利用 SPSS19.0 对1班同学的体质健康测试数据进行主成分分析, 得到主成分矩阵和方差解释表, 利

用主成分矩阵和方差解释表结合相关权重计算公式,建立基于主成分分析的权重模型,确定各体质测试项目所占的权重。然后,对比《国家体质健康标准》,在Excel表格中计算出1班同学的体质健康测试各项的单项得分,根据有权重模型的得出的各测试项目的权重,进行加权求和得出总分,进行体质健康评价。

(4) 问题四,查阅补充身体机能和运动机能方面相关资料,根据人体运动中新陈代谢规律、心肺功能训练原理和力量训练设计基本原则,并考虑身体机能疲劳的出现,为不同不合格的项目设计相对应的训练方式和强度,绘制相应的表格。利用权重模型确定的各项目权重和体质健康测试分数,计算分配对应的训练方式和强度。

三、模型假设

1. 假设学生参加体质测试时,其身体状态处于正常水平,即测试数据能准确反映同学的体质健康状况。
2. 假设参加测试的学生无生理疾病,即偏差的体质测试结果全部由于测试不规范造成的。
3. 假设问题二中生源地的分析样本是独立的随机样本,保证变异的可加性。
4. 假设对比《国家体质健康标准》对学生体质进行体质健康评分时,计算的单项得分若小于10时,单项得分记为0分。

四、符号说明

- | | |
|--|--------------------------------|
| 1. r :为相关样本的相关系数。 | 2. s_{xy} : x 和 y 的协方差 |
| 3. θ :是总体 X 的一个未知参数 | 4. σ 称均方差或标准差 |
| 5. H_0 :单因素方差分析的原假设 | 6. H_1 :单因素方差分析的择备假设 |
| 7. n_i :为第 i 个总体的样本观察值个数 | |
| 8. x_{ij} :为第 i 个总体的第 j 个观察值 | |
| 9. \bar{x} :全部观察值的总均值 | 10. n : 为全部观察值的个数 |
| 11. k : 为因素水平(总体)的个数 | 12. F :所需要的检验统计量 |
| 13. F_α :给定的显著性水平 α 的临界值 | 14. F_m 为分析后得到的 m 个主成分 |
| 15. α_i :每个项目所对应的系数 | 16. β_i : 体质健康测试项目合格判定系数 |
| 17. w_i :各测试项目指标权重 | |

五、模型的建立与求解

5.1 问题一模型的建立和求解

5.1.1. 体重与其他各测试项目的相关性分析

(1) 应用原理:

频数分析,是描述性统计中最常用的方法之一,它能够了解变量取值的状况,对把握数据分布特征非常有用,适用于对分类变量以及不服从正态分布的连续性变量进行描述。通过频数分析,可以得出不同的数值出现的频数者数据落入指定区域内的频数了解数据的分布状况。

相关性分析,是研究两个变量之间关联程度的统计方法,它主要是通过计算简单相关系数来反映变量之间关系的强弱。皮尔逊相关系数常称为积差相关系数,适用于研究连续变量之间的相关程度。其计算公式如下:

$$r = \frac{s_{xy}}{s_x s_y}$$

$$s_{xy} = \frac{\sum (x - \bar{x})(y - \bar{y})}{n} = \frac{1}{n} \sum (x - \bar{x})(y - \bar{y})$$

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

其中：r——相关系数， s_{xy} ——x 和 y 的协方差

相关系数的范围在-1 到 1 之间，即 $-1 \leq r \leq 1$ ，当 $r=1$ 为完全正相关， $r=-1$ ，为完全负相关， $r=0$ 为不相关。r 的范围在 0.3-0.5 是低度相关；r 的范围在 0.5-0.8 是显著相关；r 的范围在 0.8 以上是高度相关。

(2) 体重是体现体质健康状况的重要指标，由于肺活量、握力、立定跳远、坐位体前屈、台阶试验的测试结果都受身高体重的影响，我们要分析体重对体质健康的影响，首先要将学生划分为正常和非正常两组，划分的标准采用体质指数 (BodyMass Index, BMI) 进行分组，体质指数 (BMI) 的正常范围为 18~24，其余为非正常范围。对所收集大一新生的身高、体重数据，运用 $BMI = \text{体重(kg)} / [\text{身高(m)}]^2$ 公式，通过 Excel 计算产生 BMI 值。并采用 SPSS19.0 对 BMI 进行频数分析、对体重及各项测试项目与大学生体质健康指标进行相关分析，得出不同体重级别对学生的体质健康状况的影响程度。

在 SPSS 中用进行描述性分析的模块 Descriptive Statistic 下的 Frequencies, Analyze-->Descriptive Statistic-->Frequencies。对 BMI 值进行描述性分析，可以得到如下结果：

表 5.1.1.1 BMI 分布频数一览表

	N	平均值	最大值	最小值	偏低		正常		超重		肥胖	
					N	%	N	%	N	%	N	%
男生	763		40.10	15.92	32	4.2	515	67.50	149	19.53	67	8.78
女生	306		36.87	0	19	6.21	257	83.99	25	8.17	5	1.63
总体	1069		40.10	0	51	4.77	772	72.22	174	16.28	72	6.74

表 5.1.1.1 显示，BMI 值在 18~24 之间为正常，小于 18.5 为偏低，大于或等于 24，且小于 28 为超重；大于或等于 28 为肥胖，根据 BMI 值将学生分为两组：正常和非正常（包括偏低、超重和肥胖），方便下面分类对体重及各项测试项目与大学生体质健康指标进行相关分析。

(3) 在 SPSS 中用进行相关性分析的模块 Corelate 下的 Bivariate, Analyze-->Corelate-->Bivariate。对体重与体质健康指标进行相关性分析，可以得到如下结果：

描述性统计量			
	均值	标准差	N
体重	53.84	5.029	257
台阶测试	50.20	7.027	257

相关性			
		体重	台阶测试
体重	Pearson 相关性	1	.045
	显著性 (双侧)		.472
	N	257	257
台阶测试	Pearson 相关性	.045	1
	显著性 (双侧)	.472	
	N	257	257

描述性统计量			
	均值	标准差	N
体重	53.84	5.029	257
立定跳远	1.56	.168	257

相关性			
		体重	立定跳远
体重	Pearson 相关性	1	-.077
	显著性 (双侧)		.216
	N	257	257
立定跳远	Pearson 相关性	-.077	1
	显著性 (双侧)	.216	
	N	257	257

描述性统计量			
	均值	标准差	N
体重	53.84	5.029	257
握力坐位体	16.72	8.042	257

相关性			
		体重	握力坐位体
体重	Pearson 相关性	1	.069
	显著性 (双侧)		.267
	N	257	257
握力坐位体	Pearson 相关性	.069	1
	显著性 (双侧)	.267	
	N	257	257

图 5.1.1.1 女生正常体重级别与体质健康指标的相关关系

根据图 5.1.1.1 及附表图 5.1.1.3 相关性指标可知，男生肺活量、台阶测试、握力与体重显著相关，且成正比例关系，而立定跳远与体重无明显相关。女生肺活量与体重显著相关，且成正比例关系，而立定跳远、坐位体前屈、台阶测试与体重无明显相关。

描述性统计量			
	均值	标准差	N
体重	58.51	15.060	49
立定跳远	1.52	.222	49

相关性			
		体重	立定跳远
体重	Pearson 相关性	1	-.152
	显著性 (双侧)		.298
	N	49	49
立定跳远	Pearson 相关性	-.152	1
	显著性 (双侧)	.298	
	N	49	49

描述性统计量			
	均值	标准差	N
体重	58.51	15.060	49
握力坐位体	15.88	9.072	49

相关性			
		体重	握力坐位体
体重	Pearson 相关性	1	-.294 [*]
	显著性 (双侧)		.040
	N	49	49
握力坐位体	Pearson 相关性	-.294 [*]	1
	显著性 (双侧)	.040	
	N	49	49

描述性统计量			
	均值	标准差	N
体重	58.51	15.060	49
肺活量	2742.80	530.742	49

相关性			
		体重	肺活量
体重	Pearson 相关性	1	.105
	显著性 (双侧)		.472
	N	49	49
肺活量	Pearson 相关性	.105	1
	显著性 (双侧)	.472	
	N	49	49

图 5.1.1.2 女生非正常体重级别与体质健康指标的相关关系

根据图 5.1.1.2 及附表图 5.1.1.4 相关性指标可知，男生肺活量、握力与体重显著相关，且成正比例关系，而立定跳远、台阶测试与与体重无明显相关。女生坐位体前屈、台阶测试与体重显著相关，且成反比例关系，说明非正常体重对学生的心肺功能和身体韧性造成负面影响，立定跳远与体重成反比例关系，说明非正常体重对学生的下肢爆发力与身体协调能力造成负面影响，而肺活量与体重无

明显相关

总体来说，正常体重的男女学生的体质健康指标基本呈正相关，即在正常体重范围内，体重越重，体质越健康。非正常体重的男女学生的体质健康指标受体重影响显著，一般情况下体重越重，对大学生心血管功能、身体韧性、下肢爆发力和身体协调能力造成的负面影响越大。

5.1.2. 测试数据中的异常数据筛选

(1) 应用原理：

箱线图 (Boxplot) 也称箱须图 (Box-whisker Plot)，是利用数据中的五个统计量：最小值、第一四分位数、中位数、第三四分位数与最大值来描述数据的一种方法，它也可以粗略地看出数据是否具有有对称性，分布的分散程度等信息。其中箱体的两端分别对应下四分位数 Q_1 和上四分位数 Q_3 ， Q_1 和 Q_3 之间称作四分位距 (Inter Quartile Range, IQR)。上四分位点右边 1.5 倍 IQR 和下四分位点左边 1.5 倍 IQR 位置对应的点是异常值截断点，异常值截断点之间是内限。上四分位点右边 3 倍 IQR 和下四分位点左边 3 倍 IQR 位置对应的点是极端值截断点，极端值截断点之间是外限。异常值截断点以外的数据称作异常值，其中在内限与外限之间的异常值为温和异常值或离群值 (Outlier, mild outliers)，在极限以外的为极端异常值或极端值。

设 θ 是总体 X 的一个未知参数，若存在随机区间 $[\theta_1, \theta_2]$ ，对于给定的 $0 < \alpha < 1$ ，若满 $P\{\theta_1 < \theta < \theta_2\} = 1 - \alpha$ ，则称区间 $[\theta_1, \theta_2]$ 是 θ 的置信水平 (置信度) 为 $1 - \alpha$ 的置信区间。 θ_1, θ_2 分别称为置信下限和置信上限 (双侧置信区间)， $1 - \alpha$ 为置信度， α 为显著水平。

正态分布随机变量 X 的密度函数曲线呈中间高两边低、对称的钟形，期望 (均值) $EX = \mu$ ，方差 $DX = \sigma^2$ ，记作 $X \sim N(\mu, \sigma^2)$ ， σ 称均方差或标准差，当 $\mu = 0$ ， $\sigma = 1$ 时，称为标准正态分布，记作 $X \sim N(0, 1)$ 。正态分布完全由均值 μ 和方差 σ^2 决定，它的偏度为 0，峰度为 3。95% 的数值落在距均值左右 2 个标准差的范围内，即 $P\{\mu - 3\sigma \leq X \leq \mu + 3\sigma\} = 0.95$ 。

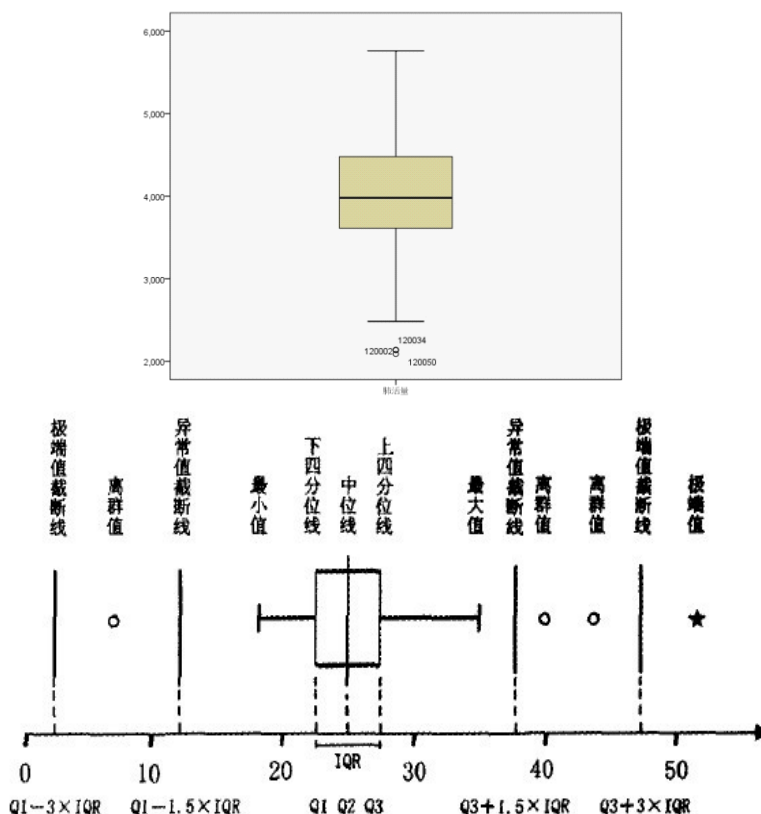
(2) 在 SPSS 中用进行描述性分析的模块 Descriptive Statistic 下的 Explore, Analyze-->Descriptive Statistic-->Explore。对肺活量、握力、立定跳远、坐位体前屈、台阶试验、身高和体重进行描述性分析，以肺活量这一指标为例可以得到如下结果：

图 5.1.2.1 肺活量描述性指标

肺活量	均值		3965.2639	85.86236
	均值的 95% 置信区间	下限	3794.0592	
		上限	4136.4686	
	5% 调整均值		3988.4691	
	中值		3980.5000	
	方差		530808.817	
	标准差		728.58627	
	极小值		2089.00	
	极大值		5762.00	
	范围		3673.00	
	四分位距		874.25	
	偏度		-.512	.283
	峰度		.734	.559

根据计算所得均值、方差、置信区间等值可初步判断出测试结果存在一定性的不准确性。以肺活量这一指标为例可得箱线图如下：

图 5.1.2.2 肺活量数据箱线图



如图所示，进行比较，可以很容易地得出，中间箱型区域为正常值，而学号为 120002、120034、120050 为离群值，也就是异常值。下面我们以 1、2、3 班同学的测试结果为样本进行分析，结合测试项目评分标准，得出可能偏差测试结果如下：

男生：①肺活量：学号为 120002、120034、120050 同学的肺活量指数过低，明显偏离评分标准指数范围，存在偏差；②台阶测试：学号为 120017、120045、120053 同学的台阶测试结果过低，明显偏离评分标准结果范围，存在偏差；③握力、身高测试结果未出现偏差；④立定跳远：学号为 120082 同学的立定跳远成绩过低，明显偏离评分标准结果范围，存在偏差；⑤体重：学号为 120043、120040、120068、120044、120033 同学的体重偏离正常范围，但是实际中可能会出现这种情况，因此我们再结合该同学其他测试指标的结果综合判断，学号为 120043、120044 同学的体重过重，但其他测试指标均达到了优秀水平，与实际情况不符，因此其结果存在偏差。

女生：①身高、肺活量、坐位体前屈和立定跳远的测试结果未出现偏差；②学号为 120055、同学的体重及台阶测试高于正常值，但仍在评分标准范围之内，认定其结果没有偏差。

综上所述，在 1、2、3 班同学中，其测试结果出现偏差的同学的学号为 120002、120034、120050、120017、120045、120053、120082、120043、120044。

5.2 问题二模型的建立和求解

5.2.1 建立聚类分析模型判断生源地与体能健康是否具有显著差别

(1) 应用原理：

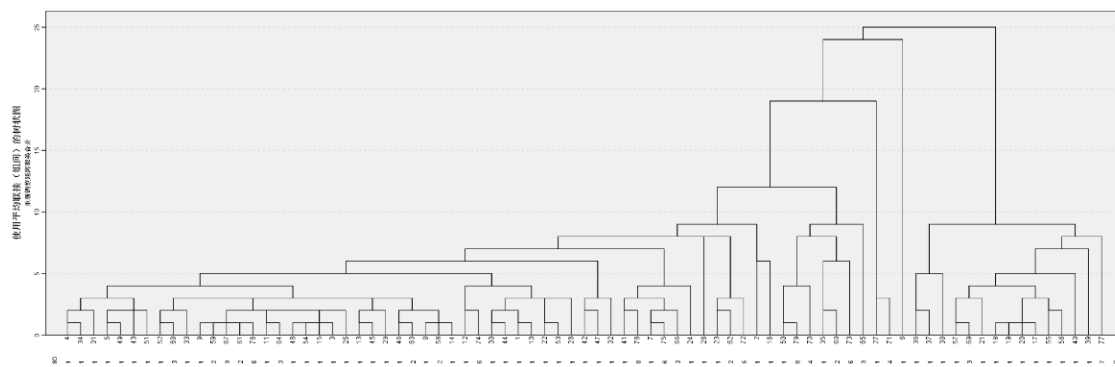
聚类分析法，基本思想是：距离相近的样品（或变量）先聚成类，距离相远

的后聚成类，过程一直进行下去，每个样品（或变量）总能聚到合适的类中。系统聚类过程是：假设总共有 n 个样品（或变量），第一步将每个样品（或变量）独自聚成一类，共有 n 类；第二步根据所确定的样品（或变量）“距离”公式，把距离较近的两个样品（或变量）聚合成一类，其他的样品（或变量）仍各自聚为一类，共聚成 $n-1$ 类；第三步将“距离”最近的两个类进一步聚成一类，共聚成 $n-2$ 类；……以上步骤一直进行下去，最后将所有的样品（或变量）聚成一类。为了直观地反映以上的系统聚类过程，可以把整个分类系统地画成一张谱系图。

（2）我们以 1、2、3 班体质测试结果的准确值为分析样本，即把第二问中有偏差的测试结果去除后的值。

在 SPSS 中用多元统计模块下的聚类分析，以肺活量、台阶测试、握力（坐位体前屈）、立定跳远为因变量，学生的生源地为标识量进行系统聚类分析，把系统聚类法分出的类别与原生源地的类别进行比较，得出结论，分类结果的树状图如下图所示：

图 5.2.1.1 系统聚类树状图



由图中我们可以看出，用系统聚类法把相似性高的学生分为了八类，而这八类与原生源地的八个地区无显著的相关性，因此判断学生的生源地对学生的体质状况无显著相关。

5.2.2 建立单因素方差分析模型

（1）应用原理：

单因素方差分析法，假设一般提法 $H_0: m_1 = m_2 = \dots = m_k$ ，自变量对因变量没有显著影响； $H_1: m_1, m_2, \dots, m_k$ 不全相等，自变量对因变量有显著影响。假定从第 i 个总体中抽取一个容量为 n_i 的简单随机样本，第 i 个总体的样本均值为该样本的全部观察值总和除以观察值的个数，计算公式为

$$\bar{x}_i = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i} \quad (i = 1, 2, \dots, k)$$

式中： n_i 为第 i 个总体的样本观察值个数

x_{ij} 为第 i 个总体的第 j 个观察值

全部观察值的总均值：

$$\bar{\bar{x}} = \frac{\sum_{i=1}^k \sum_{j=1}^{n_i} x_{ij}}{n} = \frac{\sum_{i=1}^k n_i \bar{x}_i}{n}, \quad \text{式中：} n = n_1 + n_2 + \dots + n_k$$

总误差平方和 SST，反映全部观察值的离散状况，其计算公式为：

$$SST = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

水平项平方和 SSA，反映各总体的样本均值之间的差异程度，该平方和既包括随机误差，也包括系统误差，计算公式为：

$$SSA = \sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_i - \bar{x})^2 = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2$$

误差项平方和 SSE，反映每个样本各观察值的离散状况，该平方和反映的是随机误差的大小，计算公式为：

$$SSE = \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

总离差平方和(SST)、误差项离差平方和(SSE)、水平项离差平方和 (SSA) 之间的关系 $SST = SSA + SSE$ ，即

$$\sum_{i=1}^k \sum_{j=1}^{n_i} (\bar{x}_{ij} - \bar{x})^2 = \sum_{i=1}^k n_i (\bar{x}_i - \bar{x})^2 + \sum_{i=1}^k \sum_{j=1}^{n_i} (x_{ij} - \bar{x})^2$$

SST 反映全部数据总的误差程度；SSE 反映随机误差的大小；SSA 反映随机误差和系统误差的大小；如果原假设成立，则表明没有系统误差，组间平方和 SSA 除以自由度后的均方与组内平方和 SSE 和除以自由度后的均方差异就不会太大；如果组间均方显著地大于组内均方，说明各水平(总体)之间的差异不仅有随机误差，还有系统误差；判断因素的水平是否对其观察值有影响，实际上就是比较组间方差与组内方差之间差异的大小

SSA 的均方，记为 MSA，计算公式为

$$MSA = \frac{SSA}{k - 1}$$

SSE 的均方，记为 MSE，计算公式为

$$MSE = \frac{SSE}{n - k}$$

将 MSA 和 MSE 进行对比，即得到所需要的检验统计量 F。当 H_0 为真时，二者的比值服从分子自由度为 $k-1$ 、分母自由度为 $n-k$ 的 F 分布，即

$$F = \frac{MSA}{MSE} \sim F(k - 1, n - k)$$

将统计量的值 F 与给定的显著性水平 α 的临界值 F 进行比较，作出对原假设 H_0 的决策，根据给定的显著性水平，在 F 分布表中查找与第一自由度 $df_1 = k-1$ 、第二自由度 $df_2 = n-k$ 相应的临界值 F_α 。若 $F > F_\alpha$ ，则拒绝原假设 H_0 ，表明均值之间的差异是显著的，所检验的因素对观察值有显著影响；若 $F < F_\alpha$ ，则不拒绝原假设 H_0 ，不能认为所检验的因素对观察值有显著影响。

(2) 我们以 1、2、3 班体质测试结果的准确值为分析样本，即把第二问中有偏差的测试结果去除后的值。

在 SPSS 中用方差分析的模块 Compare Mean 下的 One-Way ANOVA，Analyze-->Compare Mean-->One-Way ANOVA。对肺活量、握力、立定跳远、坐

位体前屈、台阶试验、身高和体重进行单因素方差分析，以肺活量这一指标为例可以得到如下结果：

图 5.2.2.1 单因素方差分析数据表

描述								
肺活量								
	N	均值	标准差	标准误	均值的 95% 置信区间		极小值	极大值
					下限	上限		
1	57	3615.0877	936.41330	124.03097	3366.6238	3863.5516	1229.00	5762.00
2	6	3918.8333	600.85187	245.29675	3288.2780	4549.3887	3129.00	4542.00
3	6	3664.6667	892.44824	364.34047	2728.0997	4601.2337	1919.00	4258.00
4	2	4650.0000	125.86501	89.00000	3519.1478	5780.8522	4561.00	4739.00
5	1	4328.0000	4328.00	4328.00
6	4	3881.7500	675.33664	337.66832	2807.1387	4956.3613	3031.00	4485.00
7	1	2570.0000	2570.00	2570.00
8	2	4429.0000	309.71277	219.00000	1646.3412	7211.6588	4210.00	4648.00
总数	79	3698.0253	889.74884	100.10457	3498.7324	3897.3182	1229.00	5762.00

方差齐性检验				
肺活量				
Levene 统计量	df1	df2	显著性	
.806 ^a	5	71	.549	

a. 在计算肺活量的方差齐性检验时，将忽略仅有一个案例的组。

ANOVA					
肺活量					
	平方和	df	均方	F	显著性
组间	5376784.471	7	768112.067	.967	.462
组内	56372149.47	71	793973.936		
总数	61748933.94	78			

由方差齐性检验表，我们这里采用的是 Levene 检验法，Levene 统计量为 0.806，由于显著性 p 值为值 0.549，明显大于显著性水平（0.05），满足方差分析的前提条件，因此可以进行方差分析。

由单因素方差分析表，我们可以看到方差来源包括组间、组内和总的离差平方和，最后一列中由于显著性值为 0.462，大大于显著性水平，因此接受零假设，认为不同的生源地学生对肺活量的值没有显著性影响。

之后，我们又分别作了体重、身高、台阶测试、握力（坐位体前屈）和立定跳远的单因素方差分析，其显著性值分别为：0.086、0.275、0.055、0.140、0.687，均大于显著性水平（0.05），因此我们可以判定生源地的差异对体重、身高、肺活量、台阶测试、握力（坐位体前屈）和立定跳远的指标没有显著性影响。

综上，我们可以知道，不同地区的学生的体能健康没有显著性差别。

5.3 问题三模型的建立和求解

5.3.1. 确定各测试项目的权重

（1）应用原理：基于主成分分析的权重确定方法，

1)构造样本阵

$$X = \begin{bmatrix} x_1^T \\ x_2^T \\ \vdots \\ x_n^T \end{bmatrix} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

其中， x_{ip} 表示第 i 组样本数据中的第 p 个变量的值。

2)对样本阵 X 进行变换得 $Y=[y_{ij}]_{n \times p}$ ，其中

$$y_{ij} = \begin{cases} x_{ij}, & \text{对正指标} \\ -x_{ij}, & \text{对逆指标} \end{cases}$$

3)对 Y 做标准化变换得标准化阵

$$Z = \begin{bmatrix} z_1^T \\ z_2^T \\ \vdots \\ z_n^T \end{bmatrix} = \begin{pmatrix} z_{11} & z_{12} & \cdots & z_{1p} \\ z_{21} & z_{22} & \cdots & z_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{np} \end{pmatrix}$$

其中， $z_{ij} = \frac{y_{ij} - \overline{y_j}}{s_j}$ ， $\overline{y_j}$ ， s_j ，分别为 Y 阵中第 j 列的均值和标准差。

4)计算标准化阵 z 的样本相关系数阵

$$R=[r_{ij}]_{p \times p} = \frac{Z^T Z}{n-1}$$

5)求特征值

$|R - \lambda I_p| = 0$ ，解得 p 个特征值 $\lambda_1 \geq \lambda_2 \geq \cdots \geq 0$ 。

6)确定 m 值，使信息的利用率达到 80% 以上。确定方法为

$$\frac{\sum_{j=1}^m \lambda_j}{\sum_{j=1}^p \lambda_j} \geq 0.8$$

对每个 λ_j ， $j=1,2,\cdots,m$ 。解方程组 $Rb = \lambda_j b$ ，得单位向量 $b_j^0 = \frac{b_j}{\|b_j\|}$ 。

7)求出 $z_i = (z_{i1}, z_{i2}, \cdots, z_{ip})^T$ 的 m 个主成分分量 $u_{ij} = z_i^T b_j^0$ ， $j=1,2,\cdots,m$ 。得决策矩阵

$$U = \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_p^T \end{bmatrix} = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1p} \\ u_{21} & u_{22} & \cdots & u_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ u_{p1} & u_{p2} & \cdots & u_{pn} \end{pmatrix}$$

(2)建立权重模型

首先确定的初级权重模型即是主成分模型

$$\begin{cases} F_1 = u_{11}w_1 + u_{12}w_2 + \cdots + u_{1L}w_L \\ F_2 = u_{21}w_1 + u_{22}w_2 + \cdots + u_{2L}w_L \\ \quad \quad \quad \cdots \\ F_m = u_{m1}w_1 + u_{m2}w_2 + \cdots + u_{mL}w_L \end{cases} \quad (5.3.1.1)$$

式中, F_1, F_2, \dots, F_m 为分析后得到的 m 个主成分; a_{ij} 为决策矩阵中系数。需要指出的是, 在用 SPSS 软件进行主成分分析时, 得到不是决策矩阵系数 u_{ij} 而是初始因子载荷 f_{ij} 。二者 满足如下关系

$$u_{ij} = \frac{f_{ij}}{\sqrt{\lambda_j}}, j=1, 2, \dots, m \quad (5.3.1.2)$$

在此基础上构建综合评价函数:

$$F_Z = \sum_{j=1}^m (\lambda_j / \kappa) F_j = a_1 w_1 + a_2 w_2 + \dots a_L w_L, \quad (5.3.1.3)$$

$$\kappa = \lambda_1 + \lambda_2 + \dots + \lambda_m$$

式中, a_1, a_2, \dots, a_L 即指标 w_1, w_2, \dots, w_L 在主成分中的综合重要度。

$$V_{Zi} = \sum_{j=1}^L a_j p_{ij}, i=1, 2, \dots, h \quad (5.3.1.4)$$

可得各指标权重为

$$w_i = V_{Zi} / \sum_{i=1}^h V_{Zi} \quad (5.3.1.5)$$

由式(5.3.1.3)、式(5.3.1.4)、式(5.3.1.5)可得二级权重模型

$$\begin{cases} F_Z = \sum_{j=1}^m (\lambda_j / \kappa) F_j = a_1 w_1 + a_2 w_2 + \dots a_L w_L, \\ V_{Zi} = \sum_{j=1}^L a_j p_{ij} \\ w_i = V_{Zi} / \sum_{i=1}^h V_{Zi} \end{cases} \quad (5.3.1.6)$$

(2) 为了建立体质健康评价模型, 根据《国家体质健康标准》对学生体质进行评价, 附表 2 中(男生: sheet1; 女生: sheet2)为各项目评价标准, 需要先确定各项体质健康指标的权重, 再进行综合评分, 根据评分评价学生的体质健康状况。

在 SPSS 中用进行多元统计分析的模块 Dimension Reduction 下的 Factor, Analyze-->Dimension Reduction-->Factor。对肺活量、握力、立定跳远、坐位体前屈、台阶试验进行多元统计分析, 并建立权重模型, 确定各项体质健康指标的权重。

表 5.3.1.1 原始数据表

班号	学号	肺活量	台阶测试	男生：握力 体重指数；女 生：坐位体前屈	立定跳远
1	120003	3604	46	40.2	2.31
1	120004	5762	40	49.4	2.17
1	120007	4005	50	39.4	2.15
1	120008	4893	46	44.4	2.1
1	120010	3983	45	43.3	1.93
1	120011	3670	74	49	2.3
1	120012	3958	47	46.4	2.23
1	120013	4032	51	53	1.93
1	120014	3978	53	45.9	2.17
1	120015	4115	45	43.9	2.31
1	120016	3571	52	38.2	2.39
1	120018	2704	47	38.4	2
1	120020	3711	55	52.5	2.16
1	120021	4043	48	43.7	1.9
1	120022	3741	48	46.2	2.1
1	120023	5024	46	32.4	2.37
1	120002	2134	65	47.1	1.88
1	120009	3809	50	50.4	1.98
1	120005	3870	54	40.8	2.23
1	120006	3590	58	39.2	2.14
1	120001	4561	65	59.4	2.22
1	120017	4387	32	33.7	2.14
1	120019	3652	61	31.7	2.22
1	120024	2450	51	22.6	1.78
1	120025	3011	50	20.1	1.67
1	120026	3245	49	16.8	1.64
1	120027	2522	47	19.3	1.65
1	120029	1708	51	11.2	1.71
1	120028	2570	63	13.5	1.65

以 1 班同学的体质健康测试结果为样本，对肺活量、握力、立定跳远、坐位体前屈、台阶试验各项体质健康指标进行主成分分析。经 SPSS 软件分析得表 5.3.1.2 方差解释表、表 5.3.1.3 成分矩阵表和图 5.3.1.1 因子碎石图。

表 5.3.1.2 方差解释表

解释的总方差						
成份	初始特征值			提取平方和载入		
	合计	方差的 %	累积 %	合计	方差的 %	累积 %
1	2.341	58.523	58.523	2.341	58.523	58.523
2	.963	24.070	82.593			
3	.445	11.129	93.723			
4	.251	6.277	100.000			

提取方法：主成份分析。

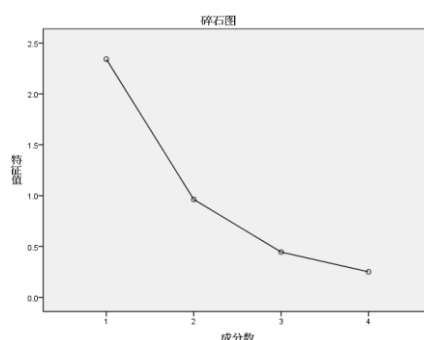
表 5.3.1.3 成分矩阵表

成份矩阵 ^a	
	成份
	1
男生：握力 体重指数；女 生：坐位体前屈	.900
立定跳远	.887
肺活量	.823
台阶测试	.258

提取方法：主成份。

a. 已提取了 1 个成份。

图 5.3.1.1 因子碎石图



如图所示,表 5.3.1.2 所示为因子分析的方差解释表,是相关系数矩阵特征值、方差贡献率及累计方差贡献率的计算结果。表 5.3.1.3 所示为成分矩阵表。图 5.3.1.1 所示是因子碎石图,横坐标为因子数目,纵坐标为特征值。从图中可以看出,因子的特征值越高,对解释原有变量的贡献就越大。

将表 5.3.1.3 成分矩阵表结果经式(5.3.1.2)转换后代入初始权重模型,再结合上述结果以及表 5.3.1.1 和表 5.3.1.2 共同代入二级权重模型即式(5.3.1.6)可得

$w = (w_1, w_2, w_3, w_4) = (0.3, 0.4, 0.12, 0.12)$, 即指标集(肺活量, 台阶试验, 立定跳远, 握力或坐位体前屈)对应的权重集(0.3, 0.4, 0.18, 0.12)。

5.3.2 计算体质测试得分

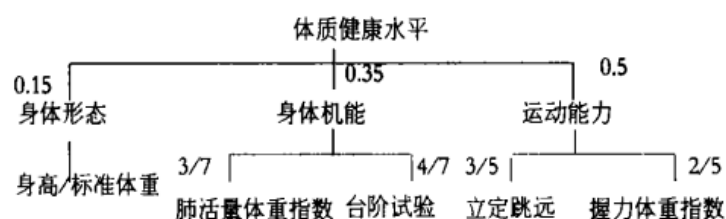
(1) 根据我国高校现采用的大学生体质健康评价体系,我们考虑将身高、体重也作为体质健康评价模型的测试评价指标,并决定用身体形态、身体机能、运动能力三个方面综合评价大学生的体质健康水平,依次建立大学生体质健康水平综合评价的层次结构。其中,身体形态包括身高和体重两项指标,身体机能包括肺活量体重指数和台阶试验两项指标,运动能力包括立定跳远和握力体重指数或坐位体前屈两项指标。

身高体重的评分用克托莱指数 K (体重 W /身高 $H \times 1000$) 计算,克托莱指数是反映人体形态发育水平和匀称程度的有效指标,们采用了个体克托莱指数与同年龄同性别的平均指数的比值作为新的指数,用这个新指数来评价大学的身体形态,所以标准模式指标数设为 1,即个体 W/H 与平均 W/H (19 岁以上克托莱指数趋于稳事实上其值为 347) 的比值,若比值 > 1 ,则分数 $= 100 \times 347 / \text{克托莱指数 } K$,反之,若比值 < 1 ,则分数 $= 100 \times \text{克托莱指数 } K / 347$ 。

以 1 班同学的体质健康测试结果为样本,对身高体重、肺活量、握力、立定跳远、坐位体前屈、台阶试验各项体质健康指标进行主成分分析。重复 (1) 中步骤,经 SPSS 软件分析及结合相关公式建立权重模型,可得

$w = (w_1, w_2, w_3, w_4, w_5) = (0.15, 0.15, 0.2, 0.3, 0.2)$, 即指标集(身高体重, 肺活量, 台阶试验, 立定跳远, 握力或坐位体前屈)对应的权重集(0.15, 0.15, 0.2, 0.3, 0.2)。综合分析,可得大学生体质健康水平综合评价层次结构及各层次权重如下图所示。

图 5.3.2.2 大学生体质健康水平综合评价层次结构及各层次权重



(2) 根据附表 2 中(男生: sheet1; 女生: sheet2)项目评价标准计算肺活量、台阶试验、立定跳远、握力或坐位体前屈测试项目的单项得分, 再根据(1)中所述的克托莱指数法计算身高体重的单项得分, 录入 Excel 表格中。通过(1)中权重模型的建立与计算, 得出各项体质健康测试指标的权重为 $w = (w_1, w_2, w_3, w_4, w_5) = (0.15, 0.15, 0.2, 0.3, 0.2)$, 即指标集(身高体重, 肺活量, 台阶试验, 立定跳远, 握力或坐位体前屈)对应的权重集为(0.15, 0.15, 0.2, 0.3, 0.2)。

在根据公式 $S = w_1s_1 + w_2s_2 + w_3s_3 + w_4s_4 + w_5s_5$, 式中, S 为体质健康测试的总分, s_1, s_2, s_3, s_4, s_5 为各项体质健康测试项目(身高体重, 肺活量, 台阶试验, 立定跳远, 握力或坐位体前屈)的单项得分, 计算给学生的体质健康测试的总分, 根据个单项得分和总分评价学生的体质健康状况, 建立体质健康评价模型。

以 1 班的 29 名同学为样本, 计算各项体质健康测试项目的单项得分和体质健康测试的总分, 生成表 5.3.2.4 —1 班学生体质测试成绩汇总表和表 5.3.2.5 —1 班学生体质测试结果统计表, 并进行体质健康评价。

表 5.3.2.4 1 班学生体质测试成绩汇总表

学生来源	身高	体重	克托莱指数	分数	肺活量	台阶测试	单项分数	男生: 握力 女生: 坐位体前屈	男生: 握力 女生: 坐位体前屈	单项分数	立定跳远	单项分数	肺活量体重指数	单项分数	总分
1	167.20	66.70	398.92	86.99	3604.00	46.00	60.00	40.20	60.27	63.98	2.31	69.00	54.03	50.00	66.04
1	176.20	67.00	380.25	91.26	5762.00	40.00	10.00	49.40	73.73	76.70	2.17	61.50	86.00	100.00	64.48
1	172.80	67.10	388.31	89.36	4005.00	50.00	66.00	39.4	58.72	62.82	2.15	60.50	59.69	64.70	67.02
1	174.50	67.50	386.82	89.70	4893.00	46.00	60.00	44.4	65.78	68.80	2.10	43.33	72.49	80.25	64.25
1	173.00	57.70	333.53	96.12	3983.00	45.00	50.00	43.30	75.04	78.00	1.93	0.00	69.03	76.00	51.42
1	183.50	65.20	355.31	97.66	3670.00	74.00	94.00	49.00	75.15	78.15	2.30	68.40	56.29	61.30	78.79
1	178.80	71.60	400.45	86.65	3958.00	47.00	61.50	46.40	64.80	67.80	2.23	64.50	55.28	60.30	67.25
1	171.70	70.00	407.69	85.11	4032.00	51.00	69.00	53.00	75.71	78.53	1.93	0.00	57.60	62.60	51.66
1	172.30	67.60	392.34	88.44	3978.00	53.00	75.00	45.90	69.90	70.43	2.17	61.50	58.85	63.80	70.37
1	171.60	63.10	367.72	94.37	4115.00	45.00	50.00	43.90	69.57	71.70	2.31	69.00	65.21	70.20	69.73
1	170.50	60.10	352.49	98.44	3571.00	52.00	72.00	38.20	63.56	66.60	2.39	75.60	59.42	64.40	74.83
1	167.50	56.10	334.93	96.52	2704.00	47.00	61.50	38.40	68.45	70.80	2.00	12.50	48.20	16.00	47.09
1	167.80	71.80	427.89	81.10	3711.00	55.00	76.50	52.50	73.12	76.10	2.16	61.00	51.69	37.00	66.54
1	169.50	66.90	394.69	87.92	4043.00	48.00	63.00	43.70	63.32	68.30	1.90	0.00	60.43	65.40	49.26
1	168.60	60.70	360.02	96.38	3741.00	48.00	63.00	46.20	76.11	78.83	2.10	43.33	61.63	66.60	65.81
1	170.40	60.10	352.70	98.38	5024.00	46.00	60.00	32.40	53.91	59.00	2.37	74.00	83.59	99.20	75.64
2	171.70	57.00	331.97	95.67	2134.00	65.00	87.00	47.10	82.63	85.60	1.88	0.00	37.44	0.00	48.87
2	167.00	71.00	425.15	81.62	3809.00	50.00	66.00	50.40	70.99	73.50	1.98	0.00	53.65	48.00	47.34
3	173.70	61.60	354.63	97.85	3870.00	54.00	75.75	40.80	66.23	69.15	2.23	64.50	62.82	67.80	73.18
3	162.10	55.60	343.00	98.85	3590.00	58.00	79.00	39.20	70.50	72.75	2.14	60.00	64.57	69.90	73.66
4	167.30	61.10	365.21	95.00	4561.00	65.00	87.00	59.40	97.22	100.00	2.22	64.00	74.65	83.40	83.36
5	162.50	51.60	317.54	91.51	4387.00	32.00	0.00	33.70	65.31	68.30	2.14	60.00	85.02	100.00	60.39
6	176.40	66.30	375.85	92.32	3652.00	61.00	82.00	31.70	47.81	24.00	2.22	64.00	55.08	60.10	63.26
1	162.30	50.40	310.54	89.49	2450.00	51.00	77.00	22.60	22.60	100.00	1.78	74.00	48.61	65.60	80.86
1	162.80	48.60	298.53	86.03	3011.00	50.00	76.00	20.10	20.10	95.20	1.67	65.00	61.95	85.50	79.47
1	162.30	54.90	338.26	97.48	3245.00	49.00	75.00	16.80	16.80	85.10	1.64	63.50	59.11	81.10	77.86
1	164.00	54.80	334.15	96.30	2522.00	47.00	69.00	19.30	19.30	92.33	1.65	64.00	46.02	63.00	75.36
1	154.40	53.00	343.26	98.92	1708.00	51.00	77.00	11.20	11.20	74.75	1.71	68.00	32.23	0.00	65.59
7	158.70	50.60	318.84	91.88	2570.00	63.00	91.50	13.50	13.50	78.63	1.65	64.00	50.79	68.70	77.31

表 5.3.2.5 1 班学生体质测试结果统计表

	优秀		良好		及格		不及格	
	人数	百分比	人数	百分比	人数	百分比	人数	百分比
身高体重	18	62.07%	11	37.93%	0	0.00%	0	0.00%
肺活量	3	10.34%	5	17.24%	15	51.72%	6	20.69%
台阶试验	2	6.90%	11	37.93%	12	41.38%	4	13.80%
立定跳远	0	0.00%	1	3.44%	20	68.97%	8	27.59%
男生：握力； 女生：坐位体前屈	4	13.79%	9	31.03%	14	48.28%	2	6.90%
总分	0	0.00%	8	27.59%	15	51.72%	6	20.69%
总人数：29人								

从表 5.3.2.4 —1 班学生体质测试成绩汇总表和表 5.3.2.5 —1 班学生体质测试结果统计表中可知，身高体重成绩普遍偏高，全部处于优秀和良好的水平；肺活量成绩偏低，近 50%的同学都处于及格范围内，还有近 20%的同学处于不及格范围内，需要加强提高；台阶测试成绩，近 80%的同学都处于及格和良好范围内，不及格率将近 14%，而优秀范围内的学生只占到了总数的 7%，说明学生的心肺机能水平需要提高；立定跳远成绩普遍偏低，近 70%的同学都处于及格内，不及格率将近 28%，而优秀范围内的学生人数为 0，说明学生的下肢爆发力和身体协调能力亟待提高；握力和坐位体前屈，近 50%的同学都处于及格范围内，还有近 30%的同学处于良好范围内，而不及格范围内的学生只占到了总数的 7%。从总分来看，1 班同学的测试成绩普遍偏低，近 50%的同学都处于及格范围内，还有近 20%的同学处于不及格范围内和近 30%的同学处于不及格范围内，而优秀率为 0，说明该班同学需要加强体质锻炼，尤其是加强关于提高心肺功能和下肢爆发力及身体协调能力的体育锻炼。

5.4 问题四建议报告书

《关于体质健康未达标大一新生在校期间如何提

高自身体质健康的建议报告书》

背景：我国大多数高校学生体质健康合格率未达到国家要求，对于未达标的大一新生，需要一个对自己的体质健康做出合理的测试和评价，并根据测评结果选择有效的锻炼方式。本报告书为学生对自己的体质健康的测评提供了方法，并且量化地提出了相应的提高健康水平的部分锻炼方式，并对其他在大学中常见的锻炼方式作出建议。

内容：

一、体育锻炼与新陈代谢和体质提升的原理

1、体育锻炼时的供能方式

(1) 磷酸原

①反应式： $ATP + H_2O = ADP + Pi + \text{能量}$ (ATP 酶)； $CP + ADP = ATP + C$ (肌酸激酶)； $2ADP = ATP + AMP$ (肌激酶)。②维持时间：肌肉中 ATP 含量很少，只能维持最大强度运动 1 秒。肌酸总量 120g，95%在肌肉，能维持最大强度运动 6-8 秒。③运动反应：强度越大，供能比例越大。④运动适应：速度、力量训练可显著提高 ATP 酶、肌酸激酶、肌激酶活性，增大 CP 储量，使磷酸原供能水平提高。

(2) 糖酵解

①总反应式： $G = 2HL + \text{能量}$ 。②维持时间：可维持最大强度运动 2-3 分钟。③运动反应：肌肉利用 CP 的同时，糖酵解过程被激活。在最大强度运动 30-60 秒时，

糖酵解达到最大速率，此后供能速率逐渐下降。乳酸积累是酵解不能持续下去的主要原因。④运动适应：速度耐力、力量耐力训练可显著提高糖酵解限速酶（己糖激酶、PF 激酶、丙酮酸激酶）活性，增大肌糖原贮备，增强机体抗酸能力，使糖酵解供能水平提高。

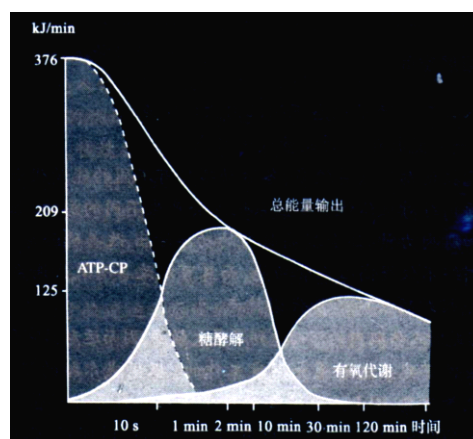
（3）有氧代谢

①总反应式： $G+6O_2=6H_2O+6CO_2+能量$ 。②持续时间：糖氧化可维持小强度运动 1-2 小时。脂肪氧化理论上不受运动时间限制，但实际上其氧化对糖有依赖性。

③运动反应：糖氧化最先启动，脂肪氧化在运动 5 分钟后逐渐增强，蛋白质氧化在运动 30 分钟后才参与供能。运动强度越小，糖供能比例越小，脂肪供能比例越大；糖贮备越少，蛋白质供能比例越大，但最大不会超过 18%。④运动适应：有氧耐力训练可显著提高有氧代谢限速酶（丙酮酸脱氢酶、柠檬酸合成酶、异柠檬酸脱氢酶、 α -酮戊二酸脱氢酶）活性，增大肝糖原贮备，增强心肺功能，使有氧代谢供能水平提高。

（4）不同代谢方式的能量输出与全力运动时的时间的关系图。

图 5.4.1 全力运动时骨骼肌三大功能系统的关系图



2、体育锻炼与体形

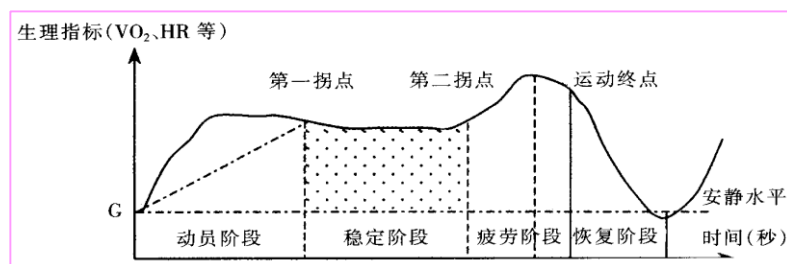
从现有的研究结果来看,有氧运动减肥目前仍是比较安全有效的减肥方式,不同的人其运动的强度与时间长短应有不同。研究表明人在运动过程中随着锻炼时间延长,脂肪供能比例增大,如在 40min、90min、180min 的连续运动时,脂肪酸供能分别占总耗能的 27%、37%、50%。而脂肪水解酶在运动进行 20min 后才能被激活,故对于体形偏胖的学生,宜进行较长时间（大于 40min）的有氧运动,如步行、跑步、骑行、游泳、体操及球类运动。而对于体形偏瘦的同学则需要通过锻炼促进机体对营养物质的吸收,进行以中等运动量(每分钟心率在 130~160 次之间)的有氧锻炼,以增长肌肉的方式增加体重。

3、体育锻炼与身体机能疲劳的出现

应用动态数学建模分析法研究表明,人体在运动过程中,心血管和呼吸系统的机能变化表现出两个明显的拐点:

第一拐点:标志进入工作状态(动员阶段)结束、稳定工作状态开始。

第二拐点:标志稳定工作状态结束、人体整体工作效率明显下降、疲劳开始。



4、体育锻炼与心肺功能

心肺功能指的是人的摄氧和转化氧气成为能量的能力。在整个过程，牵涉心脏制血及泵血功能、肺部摄氧及交换气体能力、血液循环系统携带氧气至全身各部位的效率，以及肌肉使用这些氧气的功能。心肺功能良好，也反应身体主要机能都可健康运作，从而可推断出患慢性疾病如心血管病、内分泌系统疾病、呼吸系统疾病的机会较低。

健身锻炼使人心跳加速，心搏有力，每分钟心输出量增加，心肌的微循环全面扩张；锻炼时肌肉活动产生的二氧化碳刺激了人体的呼吸中枢，使呼吸频率加快，肺容量加大，与此同时呼吸肌（膈肌、肋间外肌和肋间内肌）和呼吸辅助肌得到了锻炼，特别是膈肌的上下运动幅度增长。长期锻炼能使心脏的重量、容积增大，安静时的心率变缓，心肌的室壁增厚，使其每次收缩变得强韧有力。经常锻炼的人安静时心率可能只有每分钟 40—50 次，剧烈运动时的最高心率每分钟可达 220 次。缺乏锻炼的人在应激状态下最高心率超过每分钟 190 次时，就会出现恶心、呕吐、头晕、面色苍白、出冷汗等虚脱症状。

5、体育锻炼与肌肉数量和力量的增加。

肌肉增长的原理是超量恢复，在剧烈运动使肌肉纤维受伤之后，会有一段时间低于原先水平，随后超出原先水平，最后恢复正常，如果练的太勤，肌肉水平只会降低。小肌肉群至少需要 48 小时恢复，大肌肉群则需要 72 小时。如果超量锻炼，那只好等到肌肉完全不痛了再进行下一次。

肌肉力量的训练其强度应为最大力量的三分之二以上，较长时间的保持用力的状态有利于肌肉体积的增长，而较短时间、较快频率的练习可以增加肌肉的协调能力。

力量训练方法设计的基本原则：①大力量设计：高负荷、低重复次数、完全休息；②肌肉体积：中等负荷、大训练量、完全或不完全休息；③力量耐力：低强度、大运动量、不完全休息。

二、不达标的新生分类

对不达标的新生来说，无论男女，至少有 1 项按照《国家体质健康标准》被判定为不合格，而最多不合格项数位 4 项（男生不包括坐位体前屈，女生不包括握力），故可按不合格的项目不同将不达标的新生分为 $C_4^1 + C_4^2 + C_4^3 + C_4^4 = 15$ 类。

三、分类量化锻炼方式

每周锻炼的频率：根据肌肉超量恢复的原理，肌肉在损伤后需要 48~72 小时的时间恢复，故在考虑对不达标新生确定量化训练方式时，每周以训练 3 次为宜。

锻炼时间段：根据自身条件及附件表 5.4.2 人体机能 24 小时规律时间表，合理安排自己的锻炼时间。建议安排下午五点到六点和晚上八点到十点之间。

不同类型的达标新生的建议训练方式和强度：根据人体运动中新陈代谢规律、心肺功能训练原理和力量训练设计基本原则，为不同不合格的项目的设计相对应的训练方式和强度。为了保证不同类型的达标学生锻炼的强度不至于过大，对于每个不达标的学生，根据不合格的项目和对应项目的体质健康指标权重，计算每个项目所对应的系数

$$\alpha_i = \frac{w_i}{\sum_{i=1}^5 w_i \beta_i} \quad (1)$$

式中， w_i -----各项体质健康权重， $w_1=0.15$ ， $w_2=0.15$ ， $w_3=0.3$ ， $w_4=0.2$ ， $w_5=0.2$ ，分别依次对应身高体重、肺活量、立定跳远、握力或坐位体前屈、台阶测试。

β_i -----体质健康测试项目合格判定系数，该项合格则为 0，不合格为 1

建议运动方案和强度：

表 5.4.1 建议运动方案和强度量化表

不合格的项目	锻炼方式及强度（不论男女，训练的频率为每周进行 3 次，除慢速长跑不用休息外，其余训练方式组间休息 2~3min）			
	男生		女生	
	推荐锻炼项目（根据喜好和条件每项不合格的项目选取其中一种锻炼方式即可）	训练组数（组）	推荐锻炼项目（根据喜好和条件每项不合格的项目选取其中一种锻炼方式即可）	训练组数（组）
身高、体重	1.偏胖：慢跑或骑行，20min 一组	$\alpha_1 \times 2 \sim 3$	1.偏胖：慢跑或骑行，20min 一组	$\alpha_1 \times 2 \sim 3$
	2.偏瘦：.哑铃练习：有针对性地进行哑铃锻炼。选择能完成 8~12 次的重量为一组	$\alpha_1 \times 3 \sim 5$	2.偏瘦：.哑铃练习：有针对性地进行哑铃锻炼。选择能完成 8~12 次的重量为一组	$\alpha_1 \times 3 \sim 5$
肺活量	1.中长跑训练：稍做准备活动后进行，1000 米一组	$\alpha_2 \times 2 \sim 3$	1.中长跑训练：稍做准备活动后进行，800 米一组	$\alpha_2 \times 2 \sim 3$
	2.慢速长跑：15min 一组，	$\alpha_2 \times 2 \sim 3$	2.慢速长跑：12min 一组，	$\alpha_2 \times 2 \sim 3$
立定跳远	1.脚尖行走：前脚掌着地行走，两臂在体侧前后自然摆动。100 米为一组	$\alpha_3 \times 6 \sim 8$	1.高抬腿跑：大腿高抬与上体成直角，用前脚掌着地。30 米一组。	$\alpha_3 \times 2$
	2.变速跑：疾跑 50 米，慢跑 50 米交替进行。100 米为一组，	$\alpha_3 \times 4 \sim 6$	2.后蹬腿跑：摆动腿屈关节领先向前摆出，然后大腿积极下压。30 米一组。	$\alpha_3 \times 2$
	3.跳绳跑：边跳绳边跑。100 米 1 组	$\alpha_3 \times 4 \sim 6$	3.单足跳：一腿屈膝提起，另一腿前脚掌用力蹬地向前跳跃，前脚掌着地。30 米一组。	$\alpha_3 \times 2$
	4.往返跑：距 6~8 米两端放置或选定两个标志，练习者触摸一端标志后快速移动并触摸另一标志。往返 15 次为一组	$\alpha_3 \times 3 \sim 5$	4.五级蛙跳：连续五级蛙跳，充分展腹，膝踝关节充分伸展。	$\alpha_3 \times 5$

握力	1.抛接沙袋:将 1~2 公斤的沙袋用单手向不同的方向进行抛接,也可以两人对抛对接。 8~10 次一组。	$\alpha_4 \times 4 \sim 6$	—	—
	2.静力掰腕:肘关节置于桌面,同侧手相握,向掌心方向用力。静止保持 30 秒为 1 组。	$\alpha_4 \times 3 \sim 5$		
	3.哑铃练习:有针对性地进行哑铃锻炼。选择能完成 8~12 次的重量为一组。	$\alpha_4 \times 3 \sim 5$		
	4.屈臂静力悬垂:手握单杠,身体吊挂,保持静止坚持到疲劳。一次为 1 组。	$\alpha_4 \times 2 \sim 3$		
坐位体前屈	—	—	1.提踵练习:踝关节处于伸的状态然后做屈的动作。30 次为 1 组。	$\alpha_4' \times 3 \sim 5$
			2.屈腿仰卧起坐:仰卧,大、小腿屈成大约 90° ,上体前屈,肘关节触及膝盖。30 次为 1 组。	$\alpha_4' \times 3 \sim 5$
			3.仰卧交叉剪腿:仰卧,双手置于体侧,收腹提腿与支撑面成 30° ,双腿做外展内收的交叉剪腿动作。30 次为 1 组。	$\alpha_4' \times 3 \sim 5$
台阶测试	1.中长跑训练:稍做准备活动后进行,1000 米一组。	$\alpha_5 \times 2 \sim 3$	1.中长跑训练:稍做准备活动后进行,800 米一组。	$\alpha_5 \times 2 \sim 3$
	2.慢速长跑:15min 一组。	$\alpha_5 \times 2 \sim 3$	2.慢速长跑:12min 一组。	$\alpha_5 \times 2 \sim 3$

方案使用方式举例:

赛题中附表 1 中一位同学健康测试和评分如下:

班号	学号	性别	身高	体重	克托莱指数	分数	台阶测试	单项分数
1	120018	1	167.50	56.10	334.93	96.52	47.00	61.50

握力	握力体重指数	单项分数	立定跳远	单项分数	肺活量	肺活量体重指数	单项分数	总分
38.40	68.45	70.80	2.00	12.50	2704.00	48.20	16.00	47.09

从中可知他不合格的项目为:立定跳远和肺活量

故有 $\beta_1=0$, $\beta_2=1$, $\beta_3=1$, $\beta_4=0$, $\beta_5=0$

则据式 (1) 计算得到 $\alpha_1=0$, $\alpha_2=0.33$, $\alpha_3=0.66$, $\alpha_4=0$, $\alpha_5=0$

故他应该选择训练肺活量的中长跑训练每周三次每次 1 组,或者慢速长跑每周三次每次 1 组,并结合训练立定跳远的脚尖行走每周三次每次 4—5 组或者其他训练立定跳远的项目每周三次进行相应的组数。

四、其他辅助锻炼方式及建议

由于以上介绍的提升身体素质的方法为标准化的锻炼方式,虽然能有针对性

地迅速提高不达标学生的身体健康水平，但是缺少趣味性，不包括同学们所喜爱的能综合提升各方面素质的运动方式，现将比较常见的趣味性的运动方式对身体健康的影响小做简单介绍，以供参考：

1、球类运动均能塑造体形提升心肺功能和肌肉力量，其中足球主要提升下肢力量和灵活性，乒乓球主要提升上肢手臂和手腕的力量和灵活，而篮球、排球、羽毛球、网球等既能提升四肢的力量和灵活性，还能提升腰部和肩部的力量。球类运动由于有趣味性比较容易造成运动过度，故不宜长时间和高频率地进行。

2、武术等搏击类运动最能提升身体的爆发力、反应速度、灵活性和运动协调能力，还能塑造坚韧不拔的意志，太极拳最能提升心肺功能，并且不易受伤，适合长期习练；跆拳道对下肢和腰力的训练十分有效，而拳击主要训练上肢的力量，散打和空手道对全身的训练都有较好的效果。

3、健身房的锻炼方式现已十分成熟，各种器材和锻炼方式的设计十分有效，对身体机能可以实现针对性的训练，在健身房可以实现更加全面和有效的锻炼，在经济条件允许的情况下，建议多去健身房，也建议学校开放免费健身房，建设跟多健身设施。

六、模型的推广与前景

文章第一问中创新的使用了箱线图法对数据进行了分析筛选，这种方法不仅能直观的显示出偏离平均值较大的数据，以方便筛选，提高效率，更能综合各种指标，对总体进行评价，这种方法可以应用与很多领域。随着科技进步，数据量的增大，冗余数据、偏差数据、错误数据逐渐增多，数据筛选方法的高效直观性会越来越重要。

目前，我国经济在不断的增长，而国民的体质，特别是学生群体的体质健康状况却呈现出下降趋势，针对这种情况，较好的方法就是量化的进行体重锻炼。建议书中给出的建议运动方案和强度量化表中，不仅对身体各项指标进行了量化的建议方案，而且考虑到了身体机能疲劳与体质健康指标的权重，适合于不同身体状况的人群进行分类分时进行锻炼。此运动方案及强度量化表的推广一定会使广大人群的体质健康状况得到改善，强身强国。

参考文献

- [1]. 韩启航，数学建模，北京：高等教育出版社，2003.
- [2]. 韩中庚，数学建模方法及其应用，北京：高等教育出版社，2005
- [3]. SPSS19.0 软件和实用 MATLAB 软件
- [4]. 顾兴全，武娇等，大学生体质健康综合评价中灰色公乎估模型的应用- 吉林体育学院学报- 2004(4)
- [5]. 吴亚非，李科，基于 spss 的主成分分析法在评价体系中的运用-理论探索 - 2009
- [6]. 孙向东等，箱线图法在动物卫生数据异常值检验中的运用- 统计分析 -2010

附件

图 5.1.1.3 男生正常体重级别与体质健康指标的相关关系

描述性统计量			
	均值	标准差	N
体重	62.75	6.240	515
握力坐位体	47.15	10.155	515

描述性统计量			
	均值	标准差	N
体重	62.75	6.240	515
台阶测试	51.67	7.816	515

相关性			
	体重	握力坐位体	
体重	Pearson 相关性	1	.261**
	显著性 (双侧)		.000
	N	515	515
握力坐位体	Pearson 相关性	.261**	1
	显著性 (双侧)	.000	
	N	515	515
**. 在 .01 水平 (双侧) 上显著相关。			

相关性			
	体重	台阶测试	
体重	Pearson 相关性	1	.110*
	显著性 (双侧)		.012
	N	515	515
台阶测试	Pearson 相关性	.110*	1
	显著性 (双侧)	.012	
	N	515	515
*. 在 0.05 水平 (双侧) 上显著相关。			

描述性统计量			
	均值	标准差	N
体重	62.75	6.240	515
肺活量	3988.78	783.203	515

描述性统计量			
	均值	标准差	N
体重	62.75	6.240	515
立定跳远	2.17	.195	515

相关性			
	体重	肺活量	
体重	Pearson 相关性	1	.393**
	显著性 (双侧)		.000
	N	515	515
肺活量	Pearson 相关性	.393**	1
	显著性 (双侧)	.000	
	N	515	515
**. 在 .01 水平 (双侧) 上显著相关。			

相关性			
	体重	立定跳远	
体重	Pearson 相关性	1	.069
	显著性 (双侧)		.117
	N	515	515
立定跳远	Pearson 相关性	.069	1
	显著性 (双侧)	.117	
	N	515	515

图 5.1.1.4 男生非正常体重级别与体质健康指标的相关关系

描述性统计量			
	均值	标准差	N
体重	78.20	14.508	248
肺活量	4164.56	786.533	248

描述性统计量			
	均值	标准差	N
体重	78.20	14.508	248
立定跳远	2.11	.207	248

相关性			
	体重	肺活量	
体重	Pearson 相关性	1	.344**
	显著性 (双侧)		.000
	N	248	248
肺活量	Pearson 相关性	.344**	1
	显著性 (双侧)	.000	
	N	248	248
**. 在 .01 水平 (双侧) 上显著相关。			

相关性			
	体重	立定跳远	
体重	Pearson 相关性	1	.021
	显著性 (双侧)		.741
	N	248	248
立定跳远	Pearson 相关性	.021	1
	显著性 (双侧)	.741	
	N	248	248

描述性统计量			
	均值	标准差	N
体重	78.20	14.508	248
台阶测试	50.43	7.884	248

描述性统计量			
	均值	标准差	N
体重	78.20	14.508	248
握力坐位体	48.73	10.189	248

相关性			
	体重	台阶测试	
体重	Pearson 相关性	1	.041
	显著性 (双侧)		.517
	N	248	248
台阶测试	Pearson 相关性	.041	1
	显著性 (双侧)	.517	
	N	248	248

相关性			
	体重	握力坐位体	
体重	Pearson 相关性	1	.250**
	显著性 (双侧)		.000
	N	248	248
握力坐位体	Pearson 相关性	.250**	1
	显著性 (双侧)	.000	
	N	248	248
**. 在 .01 水平 (双侧) 上显著相关。			

表 5. 4. 2 人体机能 24 小时规律时间表

月有阴晴圆缺，人体的机能状态在一天 24 小时里同样有规律可循。人体器官 24 小时工作表，可作为健康身体的参考

- 1：00 人体进入浅睡阶段，易醒。此时头脑较清楚，熬夜者想睡反而睡不着。
- 2：00 绝大多数器官处于一天中工作最慢的状态，肝脏却在紧张工作，生血

气为人体排毒。

3:00 进入深度睡眠阶段，肌肉完全放松。

4:00 “黎明前的黑暗”时刻，老年人最易发生意外。血压处于一天中最低值，糖尿病病人易出现低血糖，心脑血管患者易发生心梗等。

5:00 阳气逐渐升华，精神状态饱满。

6:00 血压开始升高，心跳逐渐加快。高血压患者得吃降压药了。

7:00 人体免疫力最强。吃完早饭，营养逐渐被人体吸收。

8:00 各项生理激素分泌旺盛，开始进入工作状态。

9:00 适合打针、手术、做体检等。此时人体气血活跃，大脑皮层兴奋，痛感降低。

10:00 工作效率最高。

10:00—11:00 属于人体的第一个黄金时段。心脏充分发挥其功能，精力充沛，不会感到疲劳。

12:00 紧张工作一上午后，需要休息。

12:00—13:00 是最佳“子午觉”时间。不宜疲劳作战，最好躺着休息半小时至一小时。

14:00 反应迟钝。易有昏昏欲睡之感，人体应激能力降低。

15:00 午饭营养吸收后逐渐被输送到全身，工作能力开始恢复。

15:00—17:00 为人体第二个黄金时段。最适宜开会、公关、接待重要客人。

16:00 血糖开始升高，有虚火者此时表现明显。阳虚、肺结核等患者的脸部最红。

17:00 工作效率达到午后时间的最高值，也适宜进行体育锻炼。

18:00 人体敏感度下降，痛觉随之再度降低。

19:00 最易发生争吵。此时是人体血压波动的晚高峰，人们的情绪最不稳定。

20:00 人体进入第三个黄金阶段。记忆力最强，大脑反应异常迅速。

20:00—21:00 适合做作业、阅读、创作、锻炼等。

21:00 大约在晚饭后1~2小时，超重者或患有高脂血症等心血管病症者，在医生指导下可进行相应的体能锻炼，以提高疗效。

22:00 适合梳洗。呼吸开始减慢，体温逐渐下降。最好在十点半泡脚后上床，能很快入睡。

23:00 阳气微弱，人体功能下降，开始逐渐进入深度睡眠，一天的疲劳开始缓解。

24:00 气血处于一天中的最低值，除了休息，不宜进行任何活动。