

未成年人心理健康评价体系

摘要

现代社会未成年人心理问题发生率逐步上升，已成为一个不可忽视的问题。本文通过建立数学模型，构建了未成年人心理健康评价体系，并在应用层面作出展示和分析，在此基础上也为学校、家长提出了相关的改进建议。

针对问题一，根据给出的问卷调查数据建立心理健康评价体系的评价问题。为建立评价体系，对数据的处理可以大致分为 3 步：先将 10 个指标分为三类——人际关系因素：强迫症状、偏执、敌对、人际关系紧张与敏感；生活因素：适应不良、情绪不平衡、心理不平衡；学习因素：学习压力、抑郁、焦虑。然后采用稳健主成分分析中的 MCD 法，MCD 迭代算法可以降低异常数据的干扰，从而构建出稳健均值矩阵和稳健协方差矩阵，随后对稳健协方差矩阵计算出主成分，通过主成分分析法算出每一类的得分，再通过层次分析法分析每一类的权重，人际关系因素权重 0.2605，学习因素权重 0.1062，生活因素权重 0.6333，并计算得到线上问卷成绩。接着，再次使用层次分析法算得线上问卷得分权重 0.25 和线下部分得分权重 0.75，以此算出最终成绩，再以原先各等级所占分值为依据，同样将所有人划分成五个等级 A、B、C、D、E，此处这五个等级是心理健康水平递减的指标。

针对问题二，分析该地区未成年人的心理健康状况。数据中包含 1500 份线上问卷与线下部分成绩和粗略的评价，但其中问卷成绩指标的计算不贴合实际，可舍弃之。依实际经验，线下部分得分的可信度和精确度比线上高，我们根据层次分析法为线下成绩赋予较高的权重，将线上线下成绩相结合得到最终成绩。首先计算得到总体问卷信息的 Cronbach α 可信度为 0.6279，这说明数据具有一定的可靠性。再将题目所提供的样本数据代入问题一的评价模型中，可以算得人际关系因素权重 0.2605，学习因素权重 0.1062，生活因素权重 0.6333。接着又可以算得线下部分得分权重为 0.75，线上问卷得分权重为 0.25，最后得到总得分。在分析该地区未成年人心理健康状况时，可以先分析五个等级每个人数的占比，A 等占 0.2%；B 等占 12.0%；C 等占 78.8%；D 等占 7.3%；E 等占 1.6%。然后分析男女生之间的差异，大致可以得到以下结论：(1) 绝大部分人的心理将康状况都处于中等水平；(2) 强迫症状和学习压力对男女生的影响都很强烈；(3) 女生的心理将康状况普遍比男生的良好；(4) 男生心理健康状况的自我认知水平不如女生；(5) 男生比女生更不易于产生焦虑、抑郁的问题。

针对问题三，将分析结果与建议呈现给当地学校和家长，撰写一篇调查报告。可以在前两问的基础上，在成因都合理的基础上对学校、教师、家庭、学生四个方面提出建议。

关键词：MCD 法 主成分分析 层次分析 Cronbach α 可信度

一、 问题重述

1.1 问题背景

未成年人，作为一个比较特殊的群体，是国家和民族未来的希望，然而，现今未成年人的生存发展面临诸多新问题，尤其以心理问题较为突出，引起了全社会的关注与重视。在未成年人这个阶段，出于对未来的美好追求和对自己一定程度的要求，使他们产生了比较大的心理压力，随之产生了一些心理问题在现代社会，未成年人心理问题发生率逐步上升，已成为一个不可忽视的问题。构建未成年人心理健康评价体系刻不容缓，意义重大。

1.2 问题提出

通过一定的方法获得了某地区 1500 位参与者心理健康测评的线下得分和线上问卷得分，问卷主要针对 10 个指标：强迫症状、偏执、敌对、人际关系紧张与敏感、抑郁、焦虑、学习压力、适应不良、情绪不平衡以及心理不平衡。我们需要完成以下任务：

- (1) 建立合理的未成年人心理健康评价体系数学模型；
- (2) 通过第一问所建立的评价模型，分析这一地区未成年人的心理健康状况；
- (3) 根据前两问，给该地区的学校和家长写一份报告说明我们的分析结果及相关的建议。

二、 问题分析

2.1 问题一的分析

我们需要建立合适的心理健康评价模型，现将 10 个指标，分成三类，分别是人际关系因素：强迫症状、偏执、敌对、人际关系紧张与敏感；生活因素：适应不良、情绪不平衡、心理不平衡；学习因素：学习压力、抑郁、焦虑。

为了使数据更加稳健，我们通过基于 MCD 法的稳健主成分分析，先获取相应数据的稳健协方差矩阵 S_0 ，再对 S_0 进行主成分分析，计算得到累计贡献率达到 85% 的主成分，并通过这些主成分计算相应参与者各类的得分。接着通过层次分析法，建立线上问卷三个类的判断矩阵，计算各部分对应的权重，得到相应参与者的线上问卷得分。

最后再次通过层次分析法，建立线上问卷得分和线下部分得分的判断矩阵，计算各部分对应的权重，最终求得每位参与者的总成绩。再对总成绩分段分成五个等级，获得每个人的心理健康程度。

2.2 问题二的分析

题目要求分析该地区未成年人的心理健康状况，就是要根据题中中学生心理调查结果，在代入计算前，可以先把题目所给的 1500 份数据进行 Cronbach α 信度分析，来确定其可信度。

然后通过问题一所建立的评价模型，我们可以得到每个人的线上问卷得分、线下部分得分、总成绩与等级，可以先简单分析每个等级对应的人数，从总体上说明该地区未成年人的心理健康状况。然后可以更深层次和精细化研究调查数据，我们选择从性别上体现出的异同点、造成心理压力的主要因子和心理问题产生的原因三方面整理数据、展开分析。

更深层次上，我们可以首先研究心理健康状况在不同性别间的异同点。先在问题一评价体系的基础上，对男女生分别研究心理状况等级分布。其次，男女生心理问题构成的主要因子也必然有所区别，因此我们返回线上调查问卷，分别整理男女生在各项心理指标的得分情况，各项心理指标的总分对人数作平均即可体现出该项心理问题的严重程度和普遍性。在心理指标间展开横向对比可以得出未成年人最容易产生的心理问题和各类问题在未成年人心理压力中所占比重；而对男女生的心理指标展开纵向对比则可以得出男女生在心理状况上存在的不同，有利于进一步的原因分析。最后，线下成绩和线上成绩既要相互结合，体现未成年人的综合心理状况成绩，又应分别看待，线下成绩可以客观地反映调查对象的心理健康水平，线上调查问卷结果则以自查的方式反映未成年人对自身可能存在的心理问题知觉程度。因此对比分析线上与线下成绩，即可体现出未成年人对自身心理问题的了解和调节情况，为解决未成年人心理问题提出指导性意见，为问题三打下基础。

2.3 问题三的分析

借助前两问建立的评价模型和分析结果，阐述造成这种现象的原因，并从学校、教师、家庭、学生四个层次，提出建议，帮助学校和家长了解和改善未成年人心理健康，真正做到为国家和社会贡献根红苗正的年轻力量。

三、 模型假设

为了更好地建立并求解模型，不妨对模型作出以下合理的假设：

- (1) 所得的 15 个表格的数据存在异常，且异常是随机的；
- (2) 数据的异常可以通过合适的方法进行处理；
- (3) 参与者秉持公平公正原则进行答题，没有任何偏向性；
- (4) 参与者具有一定的代表性，可以代表该地区整体的心理健康状况。

四、 主要符号说明

符号	符号意义	符号	符号意义
α	Cronbach α 可信度	T_j	第 j 次迭代得到的 样本均值矩阵
S_j	第 j 次迭代得到的 样本协方差矩阵	$d_j(i)$	第 i 个点在第 j 次迭代时 得到的马氏距离
$\chi^2_{p,0.975}$	置信度为 0.975 的自由度 为 p 的卡方分布的分位数	H_j	第 j 次迭代得到的马氏距离最小 的 h 个样本组成的集合
T_0	稳健样本均值矩阵	S_0	稳健样本协方差矩阵
λ_i	第 i 个特征值	e_i	第 i 个特征值对应的 正交单位化的特征向量
ψ_m	前 m 个主成分的累计贡献率	y_i	第 i 个主成分
C.I.	一致性指标	R.I.	平均随机一致性指标
R.I.	一致性比率	ω	层次分析法中和法算得的权重

五、 模型的建立和求解

5.1 问题一模型的建立和求解

5.1.1 线上问卷 10 个指标的分类

为了凸显层次感，我们不妨将线上问卷得到的 10 个指标进行分类，分成以下三类：

表 1: 10 个指标的分类

人际关系因素	学习因素	生活因素
强迫症状	焦虑	适应不良
偏执	抑郁	情绪不平衡
敌对	学习压力	心理不平衡
人际关系紧张与敏感		

由于题目所给数据众多，包括线下成绩、等级、十项问卷心理健康因子调查结果，我们可以对所有数据指标按下图分层划类，并以此为基本思路计算心理调查总成绩：

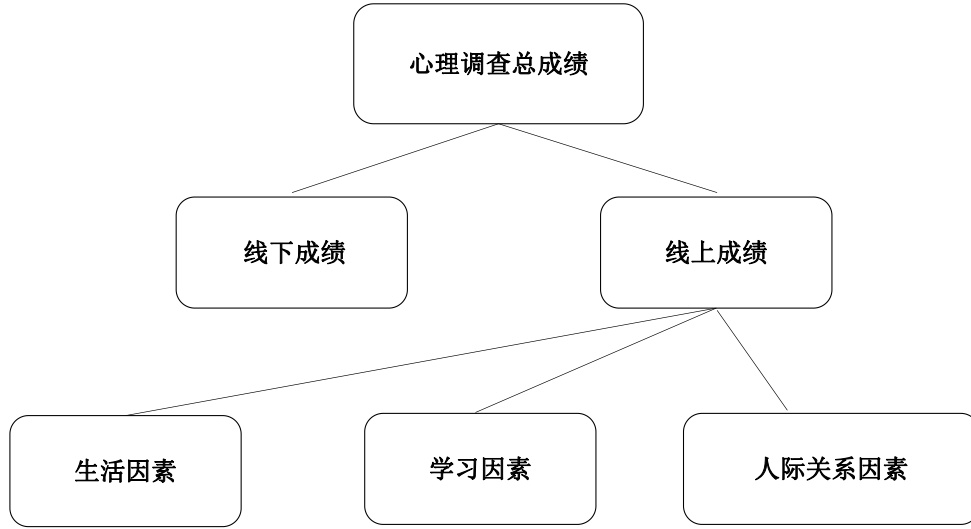


图 1: 所有数据的层次图

5.1.2 利用 MCD 法获得稳健的协方差矩阵

由于所得的数据无法保证不存在异常点，所以我们需要借助 MCD 法获得稳健的协方差矩阵^[2]，来增加我们所要建立的评价模型的准确度。MCD 法具有渐进正态性，收敛速度快，结果准确性较高，稳健性很强，是一种很好的主成分分析前的预处理。

MCD 法的基本思想是借助马氏距离，通过多次迭代选取 h 个可以使得协方差矩阵的行列式最小的数据，并以行列式最小的数据的马氏距离 $d_m(i)$ 作为权重 w_i 获得的依据，加权平均获得最终稳健的协方差矩阵 $S_0 = \frac{\sum_{i=1}^m w_i (x_i - T_1)(x_i - T_1)^T}{\sum_{i=1}^m w_i - 1}$ ，其中 $T_0 = \frac{\sum_{i=1}^m w_i x_i}{\sum_{i=1}^m w_i}$ ，以进行主成分分析。

5.1.3 利用稳健样本协方差矩阵进行主成分分析计算三类得分

主成分分析作为一种常用的方法，经常用于数据分析。大致想法就是通过计算样本的协方差矩阵或标准化后的相关系数矩阵，获得到达一定累计贡献率的主成分 y_1, y_2, \dots, y_p ，通过这部分主成分求得所有样本的主成分得分。

5.1.4 利用层次分析法综合三类因素求得线上问卷得分

层次分析法可以通过两两比较的方法确定决策方案相对重要度的总排序。既可以分解、判断、综合模型的基本特征，也能克服决策者主观判断的缺陷。^[5]

层次分析法的主要步骤就是：确定各因素间的相对重要程度，构造 $n \times n$ 的判断矩阵 A ，对判断矩阵 A 进行一致性检验，如不通过，调整判断矩阵使其能通过一致性检验，如通过，计算其每列归一化之后每行的平均值 (确定权重的和法)，即 $\omega_i = \frac{1}{n} \cdot \sum_{j=1}^n \frac{a_{ij}}{\sum_{k=1}^n a_{kj}}$ ，通过权重

算得线上问卷得分。

线上问卷得分由人际关系、学习、生活三项因素构成，不同因素包含不同的心理健康因子，但在实际生活中，三个因素给未成年人带来的心理压力大小有所不同，因此我们对三项因素进行两两比较，使用层次分析法给出权重序列，再用和法结合三项因素计算得到线上成绩综合分。通过人为划定，可确定

表 2: 判断矩阵的含义与数值

	人际关系	学习	生活
人际关系	1	3	$\frac{1}{3}$
学习	$\frac{1}{3}$	1	$\frac{1}{5}$
生活	3	5	1

一致性检验结果为 $C.R. = \frac{C.I.}{R.I.} = 0.0370 < 0.01$ ，这表明我们建立的判断矩阵是合理可行的。

得到的 $\omega = (0.2605, 0.1062, 0.6333)^T$ ，即人际关系因素权重 0.2605，学习因素权重 0.1062，生活因素权重 0.6333。以此权重计算：线上问卷得分 = $0.2605 \times$ 人际关系因素得分 + $0.1062 \times$ 学习因素得分 + $0.6333 \times$ 生活因素得分。

5.1.5 利用层次分析法综合问卷得分与线下得分获得总得分和等级

该步的方法与上一小节一致，不过由于主要两个因素，获得的判断矩阵 A 一定可以通过一致性检验，判断矩阵如下：

表 3: 判断矩阵的含义与数值

	线下	线上
线下	1	3
线上	$\frac{1}{3}$	1

可以直接通过确定权重的和法确定其权重 $\omega = (0.25, 0.75)^T$ ，即线上问卷得分的权重为 0.25，线下部分得分的权重为 0.75，并算得总得分。在得到总成绩之后，以数据说明中的 A 、

B、C、D、E 五个等级的人数比例对所有人的总成绩进行赋级，赋级公式如下：

$$\text{等级} = \begin{cases} \text{A, } 0.872 < \text{总成绩} \leq 1 \\ \text{B, } 0.744 < \text{总成绩} \leq 0.872 \\ \text{C, } 0.487 < \text{总成绩} \leq 0.744 \\ \text{D, } 0.231 < \text{总成绩} \leq 0.487 \\ \text{E, } 0 \leq \text{总成绩} \leq 0.231 \end{cases} \quad (5.1)$$

等级意思如下：

- (1) A 级：该生目前的心理状态、生活和学习状态非常良好，处于平稳状态，即使遇上问题与挑战也能较为积极的处理。
- (2) B 级：该生目前的生活和学习状态基本良好，处于较为平稳状态，但有一些压力和烦恼，不过还能较为合理的处理这些困难。
- (3) C 级：该生目前心理状态一般，有一定的心理压力，不能在生活和学习中全力发挥实力，情绪起伏较大，容易受外界因素影响。
- (4) D 级：该生目前可能面临着较多的困扰，容易感到不被理解、孤独甚至恐惧。
- (5) E 级：该生目前存在较为强烈的心理压力或强烈的心理问题的困扰。

如此，一个完整的心理健康评价模型便已建立。以下通过问题二代入样本，呈现出应得的所有数据。

5.2 问题二模型的建立和求解

5.2.1 数据预处理

这 1500 份数据中包含有 10 个指标、线下成绩，我们数据预处理大致有以下几步：

- (1) 1500 份数据中有 40 份缺少线上问卷的 10 个指标的得分，由于数据缺失过多以及原数据的庞大性，我们直接将其舍去；
- (2) 对线上问卷得到的 10 个指标进行信度分析，确定其可信度；
- (3) 计算总成绩时，表 12 中只含有线下部分成绩对应的总得分，不妨使用随机函数生成一个属于相应区间的线下部分得分。同时所有表中有部分参与者没有线下成绩，我们将其剔除，若碰到得分为 0 的，依然保留。表 5 的第 99 位参与者线下成绩为 4，总得分为 32，显然录入反了，我们不妨将其调回原位。

5.2.2 线上问卷部分的信度分析

由于本题有 1500 位同学进行了心理健康的评估，所以针对一份问卷的设置与答题质量的好坏，我们有很大的必要去判断我们所获数据的可信度，这是一种问卷内部可信度的检测，可以使用 Cronbach α 方法^[1] 进行检验。

Cronbach α 方法的公式如下：

$$\alpha = \frac{n}{n-1} \left(1 - \frac{\sum s_i^2}{s^2} \right) \quad (5.2)$$

式 (5.2) 中， n 是题量， s_i^2 是第 i 个题所有人得分的方差， s^2 是所有人所有题所得总分的方差。可信度表如下：

表 4: Cronbach α 与内部一致性的关系

α	内部一致性
$\alpha \geq 0.9$	优秀
$0.7 \leq \alpha < 0.9$	良好
$0.6 \leq \alpha < 0.7$	可接受
$0.5 \leq \alpha < 0.6$	弱
$\alpha < 0.5$	不可接受

由于在所给数据中，有 40 位同学仅有线下部分成绩，而没有完成线上问卷，我们不妨在检验时，先删除这 40 位同学，而表 12 中缺少了线下部分成绩，因此在最后检验的时候仅检验 1460 位同学线上问卷的 10 个指标的可信度，最后得到的可信度为

$$\alpha = 0.6279. \quad (5.3)$$

由表 (4) 可知，所获得的 1460 份数据在可接受范围内，但内部一致性不算良好，我们需要对其进行一定的数据处理。

在利用 Matlab 求 Cronbach α 时，为方便数据的输入，我们先将提供的 15 份 Excel 表格的名字删除只留下数字，并将表 6 中的班级列删除，将剩下列向前移，保持与其他表格的一致性。

5.2.3 利用 MCD 法获得线上问卷三类层次的稳健的协方差矩阵

MCD 法的大致步骤如下：

- (1) 已知原数据矩阵为 $X_{m \times n} = (x_1, x_2, \dots, x_m)^T$ ，其中的每个分量 x_i 是个 $n \times 1$ 的向量，从中随机选取 $h (h < n)$ 个样本数据，构成一个集合 $\{H_1\}$ ，计算这 h 个数据的样本均值矩

阵 T_1 和样本协方差矩阵 S_1 ，计算公式如式 (5.4)：

$$T_1 = \frac{1}{h} \sum_{i \in H_1} x_i, \quad S_1 = \frac{1}{h} \sum_{i \in H_1} (x_i - T_1)(x_i - T_1)^T \quad (5.4)$$

这就是初始迭代值。其中， $h \in [0.5m, m]$ ，一般来说， h 越小，处理结果抵抗异常值的能力就越强，当样本数较多时，选取 $h = 0.75m$ ；样本数较少时，选取 $h = 0.9m^{[3]}$ 。

(2) 以 T_1 和 S_1 为标准，计算全部样本的马氏距离

$$d_1(i) = \sqrt{(x_i - T_1)^T S_1^{-1} (x_i - T_1)}, (i = 1, 2, \dots, m). \quad (5.5)$$

(3) 选取马氏距离中最小的 h 个数，构成新的集合 $\{H_2\}$ ，计算这 h 个数的样本均值矩阵 T_2 和样本协方差矩阵 S_2 ，如式 (5.6)：

$$T_2 = \frac{1}{h} \sum_{i \in H_2} x_i, \quad S_2 = \frac{1}{h} \sum_{i \in H_2} (x_i - T_2)(x_i - T_2)^T \quad (5.6)$$

(4) 以 T_2 和 S_2 为标准，计算全部样本的马氏距离

$$d_2(i) = \sqrt{(x_i - T_2)^T S_2^{-1} (x_i - T_2)}, (i = 1, 2, \dots, m). \quad (5.7)$$

(5) 重复至 $\det(S_{m-1}) = \det(S_m)$ 或 $\det(S_m) = 0$ ，并得到此时马氏距离最小的 h 个数所构成的集合 $\{H_m\}$ ，以及对应的样本均值矩阵 T_m 和样本协方差矩阵 S_m 。

(6) 以 T_m 和 S_m 为标准，计算全部样本的马氏距离

$$d_m(i) = \sqrt{(x_i - T_m)^T S_m^{-1} (x_i - T_m)}, (i = 1, 2, \dots, m). \quad (5.8)$$

由于 $d_m(i)$ 近似服从一个自由度为 n 的卡方分布，取置信度为 97.5%，构造权函数如式 (5.9)

$$w_i = \begin{cases} 1, & d_m(i) \leq \sqrt{\chi_{p,0.975}^2} \\ 0, & d_m(i) > \sqrt{\chi_{p,0.975}^2} \end{cases} (i = 1, 2, \dots, m). \quad (5.9)$$

(7) 以 w_i 为权重，构造全部样本的样本均值矩阵 T_0 和样本协方差矩阵 S_0 ，公式如式 (5.10)

$$T_0 = \frac{\sum_{i=1}^m w_i x_i}{\sum_{i=1}^m w_i}, \quad S_0 = \frac{\sum_{i=1}^m w_i (x_i - T_0)(x_i - T_0)^T}{\sum_{i=1}^m w_i - 1}. \quad (5.10)$$

(8) 矩阵 S_0 就是最后处理所求得的稳健样本协方差矩阵，可以用来做主成分分析。

最后，利用 Matlab 求得的人际关系因素的稳健样本协方差矩阵 S_0 为

$$\begin{pmatrix} 0.0077 & 0.0077 & 0.0077 & 0.0077 \\ 0.0077 & 0.0077 & 0.0077 & 0.0077 \\ 0.0077 & 0.0077 & 0.0077 & 0.0077 \\ 0.0077 & 0.0077 & 0.0077 & 0.0077 \end{pmatrix} \quad (5.11)$$

学习因素的稳健样本协方差矩阵 S_0 为

$$\begin{pmatrix} 0.0025 & 0.0025 & 0.0025 \\ 0.0025 & 0.0025 & 0.0025 \\ 0.0025 & 0.0025 & 0.0025 \end{pmatrix} \quad (5.12)$$

生活因素的稳健样本协方差矩阵 S_0 为

$$\begin{pmatrix} 0.0012 & 0.0012 & 0.0012 \\ 0.0012 & 0.0012 & 0.0012 \\ 0.0012 & 0.0012 & 0.0012 \end{pmatrix} \quad (5.13)$$

5.2.4 利用稳健协方差矩阵进行主成分分析以计算三类得分

主成分分析法的大致步骤如下^[4]：

- (1) 求得稳健样本协方差矩阵 S_0 的非零特征值与对应的正交单位化的特征向量，分别设为 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k > 0$ 和 e_1, e_2, \cdots, e_k 。
- (2) 计算前 $m(m \leq k)$ 个样本主成分的累计贡献率 ψ_m ，计算方式如式 (5.14)

$$\psi_m = \frac{\sum_{i=1}^m \lambda_i}{\sum_{i=1}^k \lambda_i} \times 100\% \quad (5.14)$$

- (3) 按累计贡献率达到 85% 写出前 p 个主成分，即

$$\begin{cases} y_1 = e_{11}x_1 + e_{21}x_2 + \cdots + e_{n1}x_n \\ y_2 = e_{12}x_1 + e_{22}x_2 + \cdots + e_{n2}x_n \\ \cdots \\ y_p = e_{1p}x_1 + e_{2p}x_2 + \cdots + e_{np}x_n \end{cases} \quad (5.15)$$

- (4) 把样本 x_i 代入第 j 个主成分，算得

$$\hat{y}_{ij} = e_{1j}x_{i1} + e_{2j}x_{i2} + \cdots + e_{nj}x_{in}, (j = 1, 2, \cdots, p) \quad (5.16)$$

(5) 每个样本的得分为

$$\hat{y}_i = \frac{\sum_{i=1}^p \lambda_i \hat{y}_{ip}}{\sum_{i=1}^p \lambda_i} \quad (5.17)$$

利用 Matlab，我们可以得到三个因素累计贡献率达到 85% 的主成分都是第一个主成分，且第一个主成分的贡献率均已达到 100%，第一个主成分的公式如下：

$$\begin{aligned} \text{人际关系: } y_1 &= 0.5x_1 + 0.5x_2 + 0.5x_3 + 0.5x_4 \\ \text{学习: } y_1 &= 0.5774x_5 + 0.5774x_6 + 0.5774x_7 \\ \text{生活: } y_1 &= 0.5774x_8 + 0.5774x_9 + 0.5774x_{10} \end{aligned} \quad (5.18)$$

算得的每一类的得分分别存于 score_renji.xls、score_xuexi.xls 和 score_shenghuo.xls 中。

5.2.5 利用层次分析法计算线上问卷得分

层次分析法以及后续等级的基本步骤如下：^[6]

(1) 建立判断矩阵，判断矩阵标度及其含义如下表：

表 5: 判断矩阵标度及其含义

序号	重要性程度	c_{ij}	序号	重要性程度	c_{ij}
1	i, j 元素同等重要	1	6	i 元素比 j 元素稍不重要	$\frac{1}{3}$
2	i 元素比 j 元素稍显重要	3	7	i 元素比 j 元素明显不重要	$\frac{1}{5}$
3	i 元素比 j 元素明显重要	5	8	i 元素比 j 元素强烈不重要	$\frac{1}{7}$
4	i 元素比 j 元素强烈重要	7	9	i 元素比 j 元素极端不重要	$\frac{1}{9}$
5	i 元素比 j 元素极端重要	9			

通过人为划定，可确定

表 6: 判断矩阵的含义与数值

	人际关系	学习	生活
人际关系	1	3	$\frac{1}{3}$
学习	$\frac{1}{3}$	1	$\frac{1}{5}$
生活	3	5	1

(2) 进行一致性检验，计算一致性比率 $C.R. = \frac{C.I.}{R.I.} = 0.0370 < 0.01$ ，通过一致性检验。

(3) 通过一致性检验后，采取常用的权重计算方法中的和法，即权重为

$$\omega_i = \frac{1}{n} \cdot \frac{\sum_{j=1}^n a_{ij}}{\sum_{k=1}^n a_{kj}}, (i = 1, 2, \dots, n). \quad (5.19)$$

(4) 以所得到的 $\omega = (0.2605, 0.1062, 0.6333)^T$ ，即人际关系因素权重 0.2605，学习因素权重 0.1062，生活因素权重 0.6333。

(5) 以上述权重计算线上问卷得分。

同时我们可以得到 1460 位参与者的问卷得分，限于篇幅只展示前 10 位 (即表 1 中 21 ~ 30) 的问卷得分，其余的问卷得分在 online_score.xls 中。

表 7: 前 10 位参与者的问卷得分

参与者编号	得分	参与者编号	得分
1	1.801851572	6	1.801851572
2	1.801851572	7	1.801851572
3	1.801851572	8	1.90786308
4	1.801851572	9	1.964799019
5	1.801851572	10	1.950702678

5.2.6 利用层次分析法计算总得分与等级

(1) 该环节的判断矩阵如下：

表 8: 判断矩阵的含义与数值

	线下	线上
线下	1	3
线上	$\frac{1}{3}$	1

由于该判断矩阵是 2 阶的，一致性检验一定可以通过，可以直接计算权重。

(2) 权重 $\omega = (0.25, 0.75)^T$ ，即线上问卷得分的权重为 0.25，线下部分得分的权重为 0.75。

(3) 由于线上问卷得分为越小越好，而线下部分得分为越大越好，我们不妨将线上问卷得分反向，取其中最大的元素 c_{\max} ，再让每个元素 $\tilde{x}_i = c_{\max} - x_i$ 。再对两部分的得分进行归一化，使它们均落在 $[0, 1]$ 之间，公式如式 (5.20)：

$$\hat{x}_i = \frac{x_i - \min\{x_i\}}{\max\{x_i\} - \min\{x_i\}}. \quad (5.20)$$

- (4) 使用归一化后的两部分得分，以线上问卷得分 0.25 权重和线下部分得分 0.75 权重，计算总成绩。(为方便编程，将线上问卷得分与线下部分得分合并，存于 on_off_score.xls 中)
- (5) 在得到总成绩之后，以数据说明中的 A、B、C、D、E 五个等级的人数比例对所有人的总成绩进行赋级。

通过 Matlab 编程，我们可以获得所有数据，限于篇幅，我们展示每个等级的两个人，其余数据在 final_score.xls 中。

表 9: 每个等级各两个人的相关数据

序号	总成绩	等级	序号	总成绩	等级
551	0.875411207	A	661	0.872340895	A
42	0.794565264	B	130	0.748248522	B
1	0.593273942	C	2	0.612504711	C
7	0.458658557	D	19	0.407184063	D
44	0.256410256	E	84	0.214861775	E

5.2.7 各等级人数的比例及其分析

(以下数据均在 analysis_data.xlsx 中。) 通过问题一的模型和前几节的求解，我们可以得到每个等级的人数及其比例 (剔除空白数据后，共有 1456 位参与者)，如表 (10) 与图 (2):

表 10: 每个等级的人数及其比例

等级	人数	比例
A	3	0.00206044
B	174	0.119505495
C	1148	0.788461538
D	107	0.073489011
E	24	0.016483516

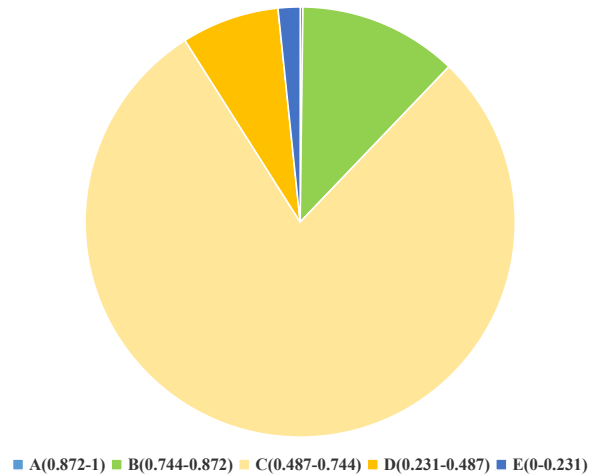


图 2: 每个等级的人数比例

总体来看, 该地区的中学生心理健康水平呈现中间宽, 头尾窄的特点, 即绝大部分未成年人心理状态处于一般水平, 而心理状态非常良好或者有强烈心理压力和心理问题的人极少。

5.2.8 男女之间的数据横向比较

将不同性别的人群分开 (仅在结果中分开, 不单独计算男生或女生作为样本总体的具体成绩), 统计其线上问卷平均得分以及线下部分平均得分, 如表 (11), 进行数据的横向比较。

表 11: 男女之间的线上问卷平均得分以及线下部分平均得分

性别	线上问卷平均得分	线下部分平均得分
男	0.37	0.67
女	0.34	0.77

针对中学生男女生群体来说, 男生的线上线下成绩分别为 0.37, 0.67(成绩越高表示心理状态越佳); 女生的线上线下成绩为 0.34, 0.77。两个群体均是线下部分得分高于线上问卷得分, 这也证实了线下部分得分比线上问卷的得分要重要。

还可以统计在等级 C、D、E, 即心理健康状态并非良好的群体中, 线上问卷部分 10 个指标的平均分, 如表 (12) 和图 (3):

表 12: C、D、E 等级中 10 个指标的平均得分

指标	男	女	指标	男	女
偏执	1.607241379	1.530764431	敌对	1.325626959	1.30675507
人际关系紧张与敏感	1.510454545	1.50648986	抑郁	1.315407524	1.440234009
焦虑	1.456974922	1.576131045	学习压力	1.839890282	1.861700468
适应不良	1.362899687	1.322106084	情绪不平衡	1.610187207	1.61862069
心理不平衡	1.378730408	1.2825117	强迫症状	2.395736677	2.346474259

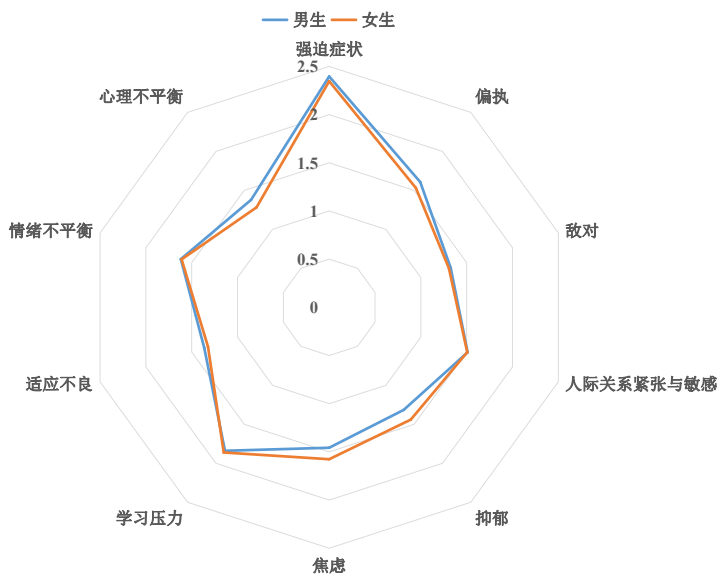


图 3: C、D、E 等级中 10 个指标男女生的平均得分

从 10 个指标入手 (此处得分表示得分越高问题越大)，我们很显然能够发现无论是男生还是女生，在强迫症状和学习压力上得分都偏高，说明强迫症状和学习压力对两个群体造成的心理问题都偏重，并且相较于学习压力，强迫症状的影响会更强。

5.2.9 男女之间的数据纵向比较

我们可以将男女生五个等级的百分比计算出来，如表 (13) 和图 (4)：

表 13: 男女生五个等级人数的占比

性别	A 人数	A 占比	B 人数	B 占比	C 人数	C 占比
男	1	0.0014388489	56	0.0805755396	539	0.775539568
女	2	0.0026281209	118	0.155059133	609	0.800262812
性别	D 人数	D 占比	E 人数	E 占比		
男	84	0.120863309	15	0.0215827338		
女	23	0.0302233903	9	0.0118265440		

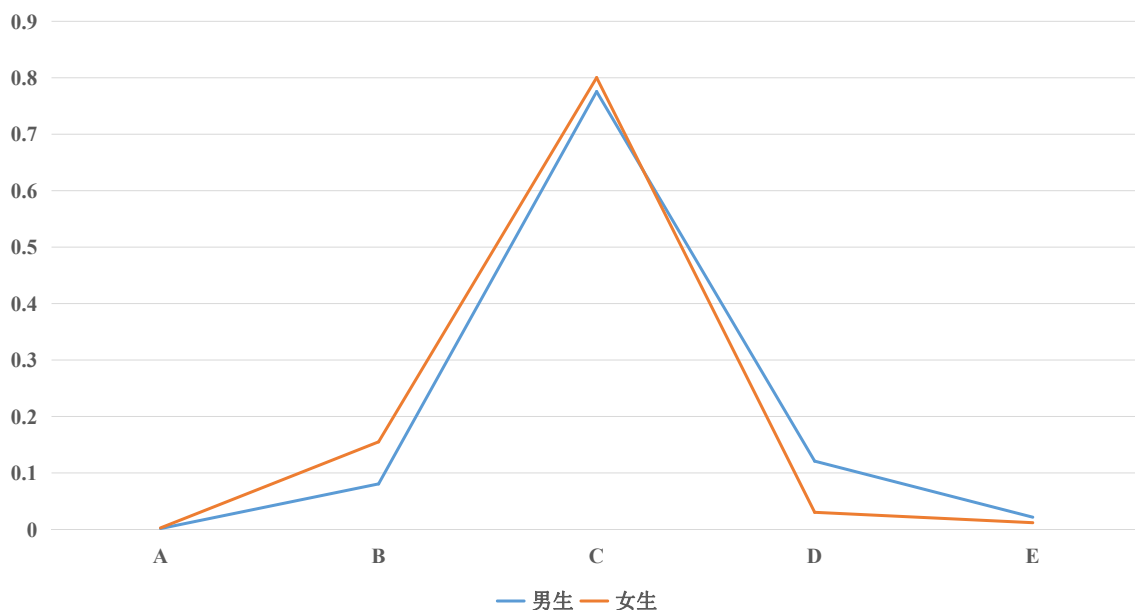


图 4: 男女生五个等级人数比例的对比

根据我们的测评标准来看，定义 C 等以下 (不含 C 等) 为心理状况不良者，男生心理状况不良率为 7.1%，女生心理状况不良率为 2.1%，男生心理状况不良率显著高于女生，该地区中学生女生的心理健康状况显著优于男生。

再看回表 (11)，就这两个群体进行比较，男生的线上得分偏高，而线下得分偏低，说明男生群体对自我的认知相较于女生弱，更不易把握自己所遇到的心理压力和问题，即更不易发现自己存在心理压力或者问题。

然后看表 (13) 和图 (4)，根据线上总成绩来看，在各项心理问题中，男生相较于女生更不易焦虑和抑郁，更容易心里不平衡。我们推测这是由于性格导致的，说明该地区男生性格相较于女生总体上更加阳光开朗，而更容易争强好胜，产生落差感，造成心里不平衡。

5.3 问题三模型的建立和求解

5.3.1 该地区心理健康的分析结果

当下，未成年人是实现伟大复兴中国梦的希望，而未成年人心理健康问题是否存在已成为影响他们是否健康成长的重要因素。为调查本地区未成年人心理状况，我们基于本地区 1500 份中学生心理调查问卷数据进行研究，对本地区中学生心理健康状况做出以下分析：

- (1) 绝大部分未成年人心理健康状况处于中等水平，而无论是心理健康状况很好或者很不好的人极少；
- (2) 对男生和女生来说，强迫症状和学习压力 (主要是强迫症状) 都是导致其心理健康状况变化的重要因素；
- (3) 女生的心理健康状况普遍比男生的良好；
- (4) 男生群体对自我心理健康状况的认知程度要弱于女生；
- (5) 男生比女生更不易于产生焦虑、抑郁的问题。

5.3.2 针对该地区心理健康的建议

我们结合之前的分析，可以提出以下几点建议^[7]：

- (1) 重视中学生的心理健康教育。当前的中学教育更关心的是学生的学习成绩，而忽略了学生的心理发展和心理需求。中学生正处在身心全面发展的时期，他们在自己学习、生活 and 自我意识等方面会遇到各种心理问题。正如我们前面的调查分析表明，大多数学生或多或少的都出现了一些心理压力或问题。研究表明，人的心理活动和人体的生理功能之间存在着内在的相关联系，如果处在良好的情绪状态，可以使生理功能处在最佳状态，那么学生专注性就得到提高，从而提高了学习效率，一举两得，何乐而不为？
- (2) 拓宽心理教育知识宣传方式。中学生的三观尚在形成之中，有很多的心理知识、见闻是他们没有的。除了以往常用的心理班会等宣传手段，学校还可在每天人流量大的地方比如食堂、宿舍附近设立心灵长廊，学生饭后散步是能驻足观望；抑或是举办心理知识竞赛，给优胜者喜欢的奖品，让学生在快乐中学习，健康中成长。
- (3) 学校与心理医疗机构合作开展教育。与学校等教育部门相比，专业的机构在技术条件、医疗经验上有无法比拟的优势。我们希望学校能加强和专业心理机构的交流与合作，通过座谈会、讲座、面对面教学等方式对学校的师资队伍进行心理教育指导，为学生提供更优质的心理保健和治疗。
- (4) 将心理健康教育与平时教学相结合。心理健康教育与其他教育一样需要通过日常教育教学活动来进行，如同春风化雨，润物细无声。作为教师，既要关注学生的心理健康，同时也要时常自检自己的心理状态，通过不断提升自己的心理素质，产生人格魅力，以此

对学生在心理上产生潜移默化的熏陶和教育作用。建议学校在教师培养上，将心理学纳入教师培养计划，使教师不仅为本专业教学能手，同时对心理学有一定程度上的了解。

- (5) 学会倾听，增强有效沟通。很多家长对着孩子训话，让孩子在一旁听话。但是家里应该是孩子吐露真心的地方。在家中，家长应意识到说话的机会更多是给孩子的，特别是根据调查问卷测出有焦虑、抑郁、心理不平衡的孩子，家长更要重视，鼓励孩子积极发言，不能因为工作忙或是自己心情不好而忽略与孩子的交流。当孩子说话时，家长可以用适当的方式比如表情管理，让孩子感受到自己讲话是有在被倾听的，是被尊重的，孩子的吐槽也是他们释放自己心理压力的一种方式。只有懂得孩子的苦恼，才能一针见血地缓解与处理孩子的心理压力与问题。
- (6) 关注孩子的情绪波动。当孩子出于明显的情绪波动时，往往面临着心理压力，这时家长们可以选择转移孩子的注意力，安排一些活动，来赢得心理自愈的恢复时间。或者是引导宣泄：吃东西、大声喊出来、写日记等，使得情绪更好的流淌。同时家长在家中还应主动调整好自己的情绪，言传身教这四个字相信各位家长都不陌生。在日常生活中，向孩子展示健康的情绪调节方法，比如倾诉和分享等，让孩子耳濡目染，学会在情绪波动时进行自我调节。

总之，家庭和学校相互配合教育是中学生心理健康教育的有效手段。只要家庭和学校配合得当，必然能大大缓解学生的心理压力，还以学生纯真健康的心灵。

六、 模型的评价与推广

6.1 模型的评价

6.1.1 模型的优点

本文问题一先将 10 个指标化为三类，选用基于 MCD 法的主成分分析、层次分析法进行分析。MCD 法具有渐进正态性，收敛速度快，且不必剔除异常数据，结果准确性较高，稳健性较强，获取的稳健协方差矩阵对后续主成分分析的帮助很大。利用主成分分析法计算每类的得分，通过降维，减轻了我们的处理数据量，减少了问题的复杂度。最后通过两次层次分析，选择主观却又合理的方法计算参与者的线上问卷得分与总成绩。

问题二在通过问题一所建立的模型基础上先借助信度分析，评价出一份问卷以及其答题质量的好坏程度，这也是能告诉数据处理者数据可行度以及完成后续数据处理的必备步骤。然后代入样本进问题一的模型，从表面到深层次，通过人数占比差异、男女间的差异等方面说明该地区未成年人的心理状况，能够较好地说明问题并加以分析。

问题三结合前两问，向学校和家长分析结果和提出建议，并从多个层次入手，比较完备地给出了我们的建议，对学校和家长有较大的参考价值。

6.1.2 模型的缺点

问题一建立的模型中使用了层次分析法，加入了人们的主观思想，这既是其优点，也是其缺点，主观的同时无法完全保证其客观性，这一点有待改进。

问题二考虑问题的方面尚不够丰富，无法很全面、多层次地说明该地区未成年人的心理状况，应该尝试再丰富分析的角度。

问题三向学校和家长提出建议的角度也尚不够丰富，值得拓宽层次。

6.2 模型的推广

问题一所使用的模型是分析和处理数据常用的模型，可以对其进行推广，尤其是其中处理异常数据的 MCD 法，通过 MCD 法的思想，还可以延伸出 Fast-MCD 法，值得人们借鉴思考。在评价时，无法避免人们的主观思维，于是正面主观性，使用的层次分析法也可以推广，引用到其他数据分析的场所，发挥其强大的作用。

问题二的分析与问题三的建议都比较实际，在实际生活中的可行性较高，值得其他问题的分析者进行借鉴。

七、 参考文献

- [1] 徐鹰. 统计分析在语言研究中的应用 [M]. 华南理工大学出版社: 广州,2018:202-203.
- [2] 成可. 稳健稀疏主成分回归的研究及应用 [D]. 山东: 山东大学,2019.
- [3] 王斌会, 陈一非. 基于 MCD 的稳健主成分算法及其实证分析 [J]. 数理统计与管理,2006,25(4):462-468.
- [4] 李素兰. 数据分析与 R 软件 [M]. 科学出版社: 北京,2019:137-139.
- [5] 周凯, 邬学军, 宋军全. 数学模型 [M]. 浙江大学出版社: 浙江,2017:85-89.
- [6] 赵华伟. 大学生心理健康状况 APH 模型评价体系的构建与应用——以鲁东大学为例 [J]. 吉首大学学报 (社会科学版),2015(z1):202-205.
- [7] 关于加强中小学生心理健康教育的建议 [EB/OL]. <http://www.gaszx.gov.cn/Item/6781.aspx>,2020-07-15.

八、 附录

附录一：信度分析的 Matlab 代码

文件名：Cronbachs_alpha_all.m

```

1 data = [];
2
3 for i = 1:15
4
5     if i == 1 | i == 11
6         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
7         data = [data; x(21:100, 7:16)];
8     else
9         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
10        data = [data; x(:, 7:16)];
11    end
12
13 end
14
15 [m, n] = size(data);
16 sumvariance = 0;
17 variance = 0;
18
19 for i = 1:n
20     sumvariance = sumvariance + var(data(:, i));
21 end
22
23 variance = var(sum(data, 2));
24
25 alpha = 10/9 * (1 - sumvariance / variance);

```

附录二：计算人际关系因素稳健协方差矩阵与问卷得分的 Matlab 代码与结果 Excel
 文件名：mcd_PCA_renji.m 和 score_renji.xls

```

1 data = [];
2
3 for i = 1:15
4
5     if i == 1 | i == 11
6         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
7         data = [data; x(21:100, 7:10)];

```

```

8     else
9         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
10        data = [data; x(:, 7:10)];
11    end
12
13 end
14
15 [m, n] = size(data);
16 h = floor(0.75 * m);
17 % 随机选取h个样本数据(固定随机数种子)
18 rng(h);
19 P = randi(m, h, 1);
20 chosen_data = data(P, :);
21 % 计算T和S1
22 T = mean(chosen_data)';
23 S1 = cov(chosen_data, 1);
24 % 先判断S1是否符合标准, 不符合重新生成S2, 直到符合标准
25 if abs(det(S1)) < 1e-14
26     d = zeros(1, m);
27
28     for i = 1:m
29         d(i) = sqrt((data(i, :) - T)' * inv(S1) * (data(i, :) - T))
30         ;
31     end
32 else
33     d = zeros(1, m);
34
35     for i = 1:m
36         d(i) = sqrt((data(i, :) - T)' * inv(S1) * (data(i, :) - T))
37         ;
38     end
39
40 [B, id] = sort(d);

```

```

40     minid = id(1:h)';
41
42     while true
43         clear chosen_data T
44         chosen_data = data(minid, :);
45         T = mean(chosen_data)';
46         S2 = cov(chosen_data, 1);
47
48         if (det(S1) == det(S2)) | (abs(det(S2)) < 1e-14)
49             break;
50         else
51             clear d
52
53             for i = 1:m
54                 d(i) = sqrt((data(i, :) - T)' * inv(S2) * (data(i, :)
55                     ' - T));
56             end
57
58             clear B id minid
59             [B, id] = sort(d);
60             minid = id(1:h)';
61         end
62
63         S1 = S2;
64         clear S2
65     end
66 end
67
68 w = zeros(1, m);
69
70 for i = 1:m
71
72     if d(i) <= sqrt(chi2inv(0.975, 10))

```

```

73         w(i) = 1;
74     else
75         w(i) = 0;
76     end
77
78 end
79
80 % 计算稳健均值矩阵和稳健协方差矩阵
81 T0 = sum(w' .* data) / sum(w);
82 S0 = zeros(4);
83
84 for i = 1:4
85     S0 = S0 + w(i) * (data(i, :) - T0) * (data(i, :) - T0)';
86 end
87
88 S0 = S0 / (sum(w) - 1);
89 % 根据稳健协方差矩阵进行主成分分析
90 [coeff, latent, explained] = pcacov(S0);
91 % 计算每个人的问卷得分
92 score = zeros(1, m);
93
94 for i = 1:m
95     score(i) = (coeff(:, 1)' * data(i, :))';
96 end
97
98 score = zeros(1, m);
99 latent_cum_sum = cumsum(latent);
100 latent_sum = sum(latent);
101
102 for i = 1:m
103
104     for j = 1:length(latent_cum_sum)
105         score(i) = score(i) + latent_cum_sum(j) / latent_sum * (coeff
            (:, j)' * data(i, :))';

```

```

106     end
107
108 end
109
110 % 写入score_renji.xls中
111 xlswrite('score_renji.xls', score');

```

附录三：计算学习因素稳健协方差矩阵与问卷得分的 Matlab 代码与结果 Excel
 文件名：mcd_PCA_xuexi.m 和 score_xuexi.xls

```

1 data = [];
2
3 for i = 1:15
4
5     if i == 1 | i == 11
6         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
7         data = [data; x(21:100, 11:13)];
8     else
9         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
10        data = [data; x(:, 11:13)];
11    end
12
13 end
14
15 [m, n] = size(data);
16 h = floor(0.75 * m);
17 % 随机选取h个样本数据(固定随机数种子)
18 rng(h);
19 P = randi(m, h, 1);
20 chosen_data = data(P, :);
21 % 计算T和S1
22 T = mean(chosen_data)';
23 S1 = cov(chosen_data, 1);
24 % 先判断S1是否符合标准, 不符合重新生成S2, 直到符合标准
25 if abs(det(S1)) < 1e-14
26     d = zeros(1, m);

```



```

27
28     for i = 1:m
29         d(i) = sqrt((data(i, :) - T)' * inv(S1) * (data(i, :) - T))
30         ;
31     end
32 else
33     d = zeros(1, m);
34
35     for i = 1:m
36         d(i) = sqrt((data(i, :) - T)' * inv(S1) * (data(i, :) - T))
37         ;
38     end
39
40     [B, id] = sort(d);
41     minid = id(1:h)';
42
43     while true
44         clear chosen_data T
45         chosen_data = data(minid, :);
46         T = mean(chosen_data)';
47         S2 = cov(chosen_data, 1);
48
49         if (det(S1) == det(S2)) | (abs(det(S2)) < 1e-14)
50             break;
51         else
52             clear d
53
54             for i = 1:m
55                 d(i) = sqrt((data(i, :) - T)' * inv(S2) * (data(i, :)
56                     ' - T));
57             end
58
59             clear B id minid

```

```

58         [B, id] = sort(d);
59         minid = id(1:h)';
60     end
61
62     S1 = S2;
63     clear S2
64 end
65
66 end
67
68 w = zeros(1, m);
69
70 for i = 1:m
71
72     if d(i) <= sqrt(chi2inv(0.975, 10))
73         w(i) = 1;
74     else
75         w(i) = 0;
76     end
77
78 end
79
80 % 计算稳健均值矩阵和稳健协方差矩阵
81 T0 = sum(w' .* data) / sum(w);
82 S0 = zeros(3);
83
84 for i = 1:3
85     S0 = S0 + w(i) * (data(i, :) - T0) * (data(i, :) - T0)';
86 end
87
88 S0 = S0 / (sum(w) - 1);
89 % 根据稳健协方差矩阵进行主成分分析
90 [coeff, latent, explained] = pcacov(S0);
91 % 计算每个人的问卷得分

```

```

92 score = zeros(1, m);
93
94 for i = 1:m
95     score(i) = (coeff(:, 1)' * data(i, :))';
96 end
97
98 score = zeros(1, m);
99 latent_cum_sum = cumsum(latent);
100 latent_sum = sum(latent);
101
102 for i = 1:m
103
104     for j = 1:length(latent_cum_sum)
105         score(i) = score(i) + latent_cum_sum(j) / latent_sum * (coeff
            (:, j)' * data(i, :))';
106     end
107
108 end
109
110 % 写入score_xuexi.xls中
111 xlswrite('score_xuexi.xls', score);

```

附录四：计算生活因素稳健协方差矩阵与问卷得分的 Matlab 代码与结果 Excel
 文件名：mcd_PCA_shenghuo.m 和 score_shenghuo.xls

```

1 data = [];
2
3 for i = 1:15
4
5     if i == 1 | i == 11
6         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
7         data = [data; x(21:100, 14:16)];
8     else
9         x = xlsread(sprintf('%s%s', num2str(i), '.xlsx'));
10        data = [data; x(:, 14:16)];
11    end

```

```

12
13 end
14
15 [m, n] = size(data);
16 h = floor(0.75 * m);
17 % 随机选取h个样本数据(固定随机数种子)
18 rng(h);
19 P = randi(m, h, 1);
20 chosen_data = data(P, :);
21 % 计算T和S1
22 T = mean(chosen_data)';
23 S1 = cov(chosen_data, 1);
24 % 先判断S1符不符合标准, 不符合重新生成S2, 直到符合标准
25 if abs(det(S1)) < 1e-14
26     d = zeros(1, m);
27
28     for i = 1:m
29         d(i) = sqrt((data(i, :) - T)' * inv(S1) * (data(i, :) - T))
30         ;
31     end
32 else
33     d = zeros(1, m);
34
35     for i = 1:m
36         d(i) = sqrt((data(i, :) - T)' * inv(S1) * (data(i, :) - T))
37         ;
38     end
39
40 [B, id] = sort(d);
41 minid = id(1:h)';
42
43 while true
44     clear chosen_data T

```

```

44     chosen_data = data(minid, :);
45     T = mean(chosen_data)';
46     S2 = cov(chosen_data, 1);
47
48     if (det(S1) == det(S2)) | (abs(det(S2)) < 1e-14)
49         break;
50     else
51         clear d
52
53         for i = 1:m
54             d(i) = sqrt((data(i, :) - T)' * inv(S2) * (data(i, :)
55                 ' - T));
56         end
57
58         clear B id minid
59         [B, id] = sort(d);
60         minid = id(1:h)';
61     end
62
63     S1 = S2;
64     clear S2
65 end
66
67
68 w = zeros(1, m);
69
70 for i = 1:m
71
72     if d(i) <= sqrt(chi2inv(0.975, 10))
73         w(i) = 1;
74     else
75         w(i) = 0;
76     end

```

```

77
78 end
79
80 % 计算稳健均值矩阵和稳健协方差矩阵
81 T0 = sum(w' .* data) / sum(w);
82 S0 = zeros(3);
83
84 for i = 1:3
85     S0 = S0 + w(i) * (data(i, :) - T0) * (data(i, :) - T0)';
86 end
87
88 S0 = S0 / (sum(w) - 1);
89 % 根据稳健协方差矩阵进行主成分分析
90 [coeff, latent, explained] = pcacov(S0);
91 % 计算每个人的问卷得分
92 score = zeros(1, m);
93
94 for i = 1:m
95     score(i) = (coeff(:, 1)' * data(i, :))';
96 end
97
98 score = zeros(1, m);
99 latent_cum_sum = cumsum(latent);
100 latent_sum = sum(latent);
101
102 for i = 1:m
103
104     for j = 1:length(latent_cum_sum)
105         score(i) = score(i) + latent_cum_sum(j) / latent_sum * (coeff
            (:, j)' * data(i, :))';
106     end
107
108 end
109

```

```

110 % 写入score_shenghuo.xls中
111 xlswrite('score_shenghuo.xls', score');

```

附录五：计算线上问卷得分的 Matlab 代码

文件名：online_score.m

```

1 renji = xlsread('score_renji.xls');
2 xuexi = xlsread('score_xuexi.xls');
3 shenghuo = xlsread('score_shenghuo.xls');
4 [m, n] = size(renji);
5 data = zeros(m, n);
6
7 for i = 1:m
8     data(i) = 0.2605 * renji(i) + 0.1062 * xuexi(i) + 0.6333 *
        shenghuo(i);
9 end
10
11 xlswrite('online_score', data);

```

附录六：1460 位参与者的线上问卷得分的 Excel 表格

文件名：online_score.xls

附录七：1456 位参与者 (剔除线下成绩为空的参与者) 的线上问卷得分与线下部分得分合并的 Excel 表格

文件名：on_off_score.xls

附录八：计算总成绩的 Matlab 代码

文件名：final_score.m

```

1 score = xlsread('on_off_score.xls');
2 [m, n] = size(score);
3 score_zcf = score(:, 1);
4 score_zcf_max = max(score_zcf);
5
6 for i = 1:m
7     score(i, 1) = score_zcf_max - score(i, 1);
8 end
9
10 score_zcf_jc = score(:, 1);

```

```

11 score_xxcj = score(:, 2);
12 score_zcf_jc_max = max(score_zcf_jc);
13 score_zcf_jc_min = min(score_zcf_jc);
14 score_xxcj_jc_max = max(score_xxcj);
15 score_xxcj_jc_min = min(score_xxcj);
16
17 for i = 1:m
18     score(i, 1) = (score(i, 1) - score_zcf_jc_min) / (
19         score_zcf_jc_max - score_zcf_jc_min);
20     score(i, 2) = (score(i, 2) - score_xxcj_jc_min) / (
21         score_xxcj_jc_max - score_xxcj_jc_min);
22     score(i, 3) = score(i, 1) * 0.25 + score(i, 2) * 0.75;
23 end
24 score = [(1:m)', score];
25 xlswrite('final_score.xls', score)

```

附录九：含有线上问卷得分、线下部分得分、总成绩和等级的 Excel 表格

文件名：final_score.xls

附录十：问题二相应数据的 Excel 表格

文件名：analysis_data.xlsx