

# 葡萄酒的评价

## 摘要

本文运用概率模型，综合评价和逐步回归分析方法解决了葡萄酒评价问题。

对于问题一，首先进行数据处理，通过 P-P 图和单样本 K-S 检验等方法确定品酒员对葡萄酒的评价近似服从正态分布，从而建立了概率模型，提出了两组评酒员的评价结果不存在显著性差异的原假设，采用了配对样本 T 检验的检验方法，发现在显著性水平  $\alpha=0.05$  的条件下，两组评酒员的评价结果存在显著性差异；通过比较两组品酒员的组内方差，发现第二组品酒员关于红白葡萄酒的评价结果更可信。

对于问题二，首先我们对数据进行了初步预处理，将部分明显异于其他样本的数据进行修正，补齐，然后进行标准化处理，运用主成分分析法，以主成分的贡献率为权重，我们建立对葡萄酒的主成分综合评价模型，从而根据酿酒葡萄的理化指标和葡萄酒的质量对酿酒葡萄进行分级。我们将葡萄从优到劣分为了四个等级，结果发现样本葡萄大多集中在二、三级，红葡萄样本中编号为 2、3、9、23 的葡萄质量最优，为一级葡萄，编号为 11、18、25 质量相对不好，为四级葡萄。

对于问题三，我们采用了逐步回归的方法，建立酿酒葡萄一级理化指标与葡萄酒一级理化指标间的回归模型，得出葡萄酒各个理化指标关于酿酒葡萄一级理化指标的函数关系，并具体分析了酿酒葡萄理化指标对葡萄酒各个理化指标的影响。

对于问题四，借用问题三建立的逐步回归模型，从外观、香气、口感三个方面分析了酿酒葡萄和葡萄酒的理化指标对葡萄酒质量的影响，可求解出了酿酒葡萄和葡萄酒的理化指标对葡萄酒的外观、香气、口感的回归函数；通过样本组中的预测值和实际值比较，判断出通过葡萄和葡萄酒的理化指标来评价葡萄酒的质量的方案是可行的。

**关键词：**T 检验   主成分分析   逐步回归

## 一、 问题重述

确定葡萄酒质量时一般是通过聘请一批有资质的评酒员进行品评。每个评酒员在对葡萄酒进行品尝后对其分类指标打分，然后求和得到其总分，从而确定葡萄酒的质量。酿酒葡萄的好坏与所酿葡萄酒的质量有直接的关系，葡萄酒和酿酒葡萄检测的理化指标会在一定程度上反映葡萄酒和葡萄的质量。附件 1 给出了某一年份一些葡萄酒的评价结果，附件 2 和附件 3 分别给出了该年份这些葡萄酒的和酿酒葡萄的成分数据。请尝试建立数学模型讨论下列问题：

1. 分析附件 1 中两组评酒员的评价结果有无显著性差异，哪一组结果更可信？
2. 根据酿酒葡萄的理化指标和葡萄酒的质量对这些酿酒葡萄进行分级。
3. 分析酿酒葡萄与葡萄酒的理化指标之间的联系。
4. 分析酿酒葡萄和葡萄酒的理化指标对葡萄酒质量的影响，并论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量？

## 二、 模型假设

1. 假设题目所给的数据真实可靠；
2. 假设评酒员给出的葡萄酒评价能够准确反映葡萄酒质量；
3. 不考虑理性性质的二级指标；
4. 假设葡萄酒的质量只与酿酒葡萄的好坏有关，忽略酿造过程中的其他因素；
5. 葡萄酒质量中的香气评定主要受葡萄酒中芳香物质的影响。

## 三、 定义与符号说明

$x_{ij}$ : 第  $i$  个品酒员对  $j$  种酒样品的打分；

$\bar{d}$ : 样本  $D_1, D_2, \dots, D_n$  的样本均值的观察值；

$s_d^2$ : 样本  $D_1, D_2, \dots, D_n$  的样本方差的观察值；

$d_i$ : 样本  $D_1, D_2, \dots, D_n$  的中的第  $i$  ( $i=1, 2, \dots, n$ ) 个样本值；

R1: 第一组品酒员对红葡萄酒的评价结果；

R2: 第二组品酒员对红葡萄酒的评价结果；

W1: 第一组品酒员对白葡萄酒的评价结果；

W2: 第二组品酒员对白葡萄酒的评价结果；

R: 相关系数矩阵

#### 四、 问题分析

问题一要对两组评酒员的评价结果进行比较,判断是否存在显著性差异、哪一组更可靠。建立显著性检验的模型判断结果是否存在显著性差异,通过比较两组数据的方差可以判断哪一组评价结果更为可靠。

问题二要根据酿酒葡萄的理化指标和葡萄酒的质量对这些酿酒葡萄进行分级。葡萄酒的质量由评酒员的打分来确定,运用主成分分析法,以主成分的贡献率为权重,建立对葡萄酒的主成分综合评价模型。

问题三要分析酿酒葡萄与葡萄酒的理化指标之间的联系,我们采用了逐步回归的方法,建立酿酒葡萄一级理化指标与葡萄酒一级理化指标间的回归模型,得出酿酒葡萄一级理化指标对葡萄酒各个理化指标的关系,分析出它们间的联系。

问题四要求分析酿酒葡萄和葡萄酒的理化指标对葡萄酒质量的影响,并论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量。为了分析葡萄与葡萄酒的理化指标对葡萄酒质量的影响,分别对葡萄酒质量的评价因素:外观、香气、口感,根据理化指标作逐步回归。这样就可以找出主要影响葡萄酒质量评价的理化指标。根据假设:葡萄酒质量中的香气评定主要受葡萄酒中芳香物质的影响,做逐步回归方程,找到芳香物质与香气评价的关系。通过葡萄酒质量三个评价因素的逐步回归方程,就可以分析得到理化指标对葡萄酒质量的影响。

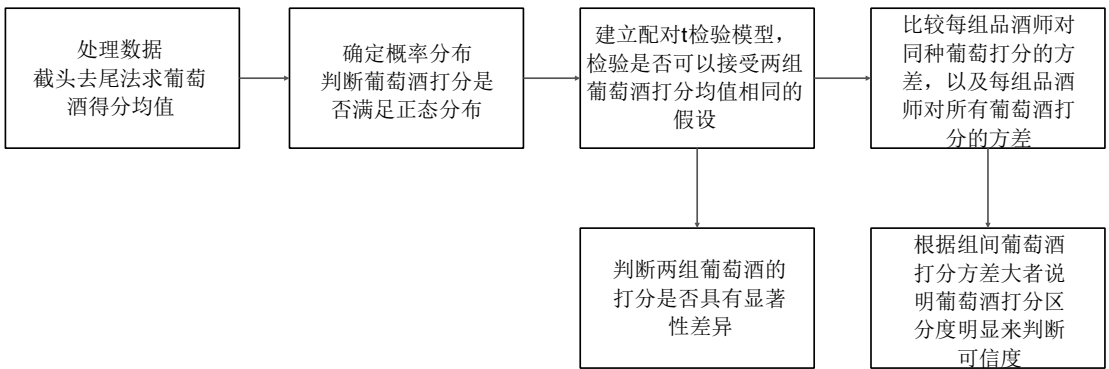
利用上面做出的逐步回归方程,分别对外观、香气、口感3个评价因素做预测评分,将这个评分与原来品酒员的评分做比较,从而判断能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量。

#### 五、 模型的建立与求解

##### 5.1 问题一模型的建立与求解

###### 5.1.1 问题一模型分析

问题一要求分析两组评酒员的评价结果有无显著性差异,并判断哪一组结果更可信。本文建立显著性检验的模型判断结果是否存在显著性差异,而比较两组数据的方差可以判断哪一组评价结果更为可靠。



问题一流程图

##### 数据处理

对问题分析,附件1中品酒员的评分带有主观性,为进行量化分析,我们采取截头去尾的方法对数据进行处理,剔除某些不合理的数据,可得第j种酒样品

最终得分为：

$$x_j = \frac{1}{n-2} \left( \sum_{i=1}^n x_{ij} - \max x_{ij} - \min x_{ij} \right)$$

其中， $x_{ij}$  表示第  $i$  个品酒员对  $j$  种酒样品的打分， $i=1,2,\cdots,m$ ， $j=1,2,\cdots,n$  ( $m,n$  分别表示品酒员总的人数和某类葡萄酒总样品数)。

确定概率分布

正态分布检验：  
在进行假设检验之前，我们判断两组葡萄酒的打分是否满足正态分布，首先绘制两组葡萄酒得分条形图如 1.1 和图 1.2 所示。

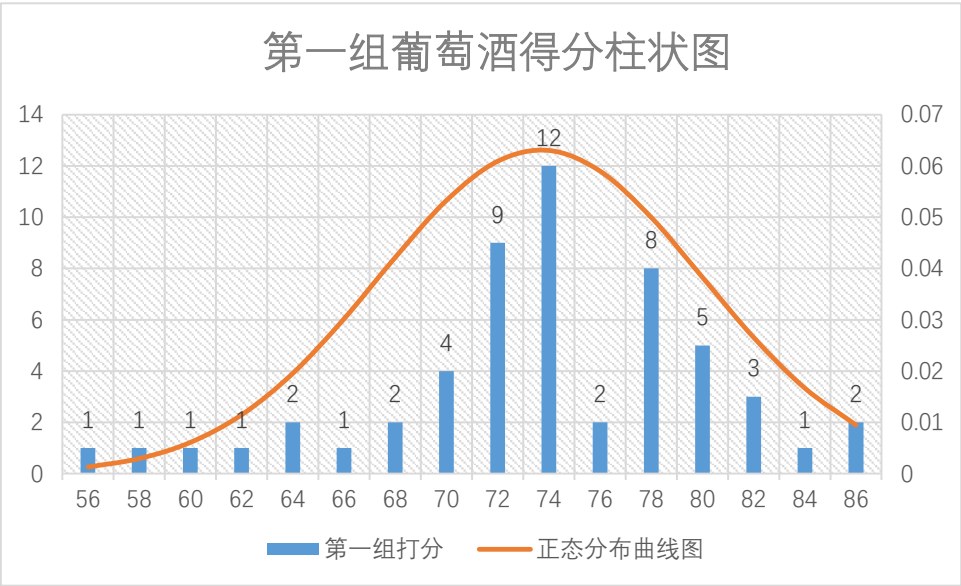


图 1.1 第一组葡萄酒得分柱状图

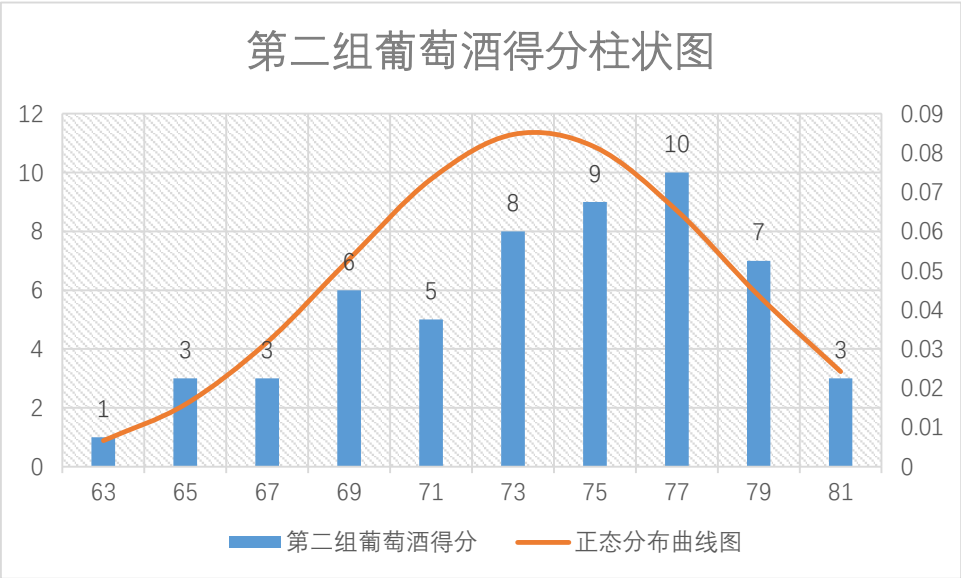


图 1.2 第二组葡萄酒得分柱状图

从图 1.1 和图 1.2 可以看出两组葡萄酒的得分都近似满足正态分布。

然后将处理过的数据导入 spss，进行 P-P 图和单样本 K-S 检验，以第一组红葡萄酒评价结果为例（详见图 1），从图 1 可以看出第一组数据的散点近似为一条直线，且与对角线大致重叠，其余三组评价结果图见附录；而从表 1 中可看出四组葡萄酒的单样本 K-S 检验结果的双边检验结果均大于 0.05。因此可确定品酒员对葡萄酒的评价近似服从正态分布。

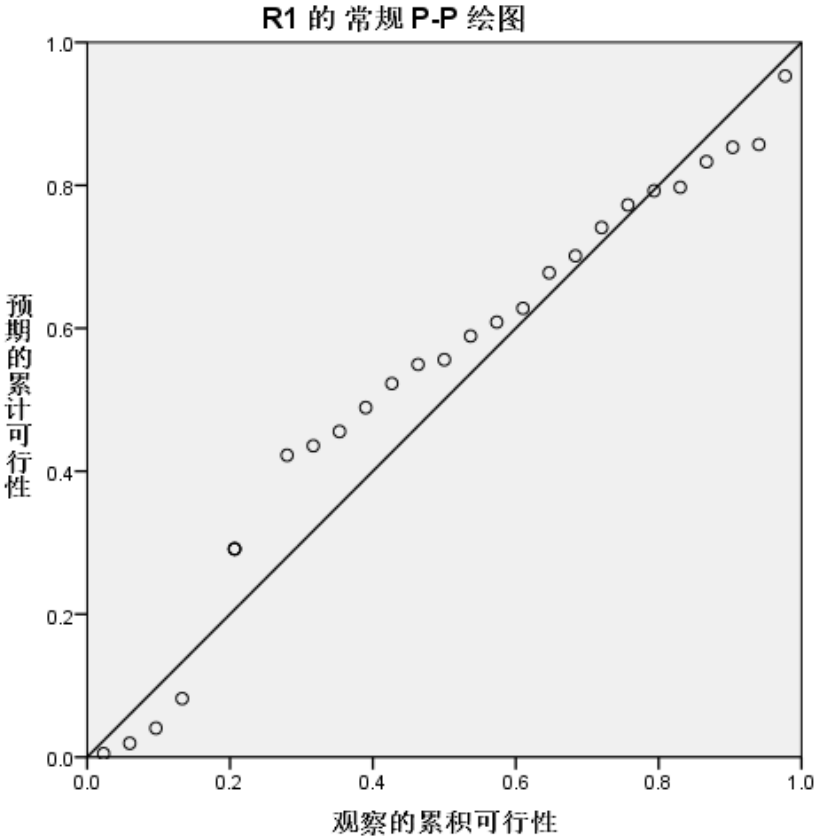


图 1 第一组红葡萄酒评价结果的 P-P 图

表 1 四组葡萄酒评价结果的单样本 K-S 检验结果

		W1	W2	R1	R2
数字		28	28	27	27
正态参数 a, b	平均值	75.0714286	77.3035714	72.9537037	70.6435185
	标准偏差	4.83314518	3.40893919	7.42171872	3.98445050
最极端差分	绝对	0.087	0.135	0.163	0.116
	正	0.065	0.069	0.106	0.090
	负	-0.087	-0.135	-0.163	-0.116
检验统计		0.087	0.135	0.163	0.116
渐近显著性（双尾）		0.200 <sup>c, d</sup>	0.200 <sup>c, d</sup>	0.063 <sup>c</sup>	0.200 <sup>c, d</sup>

注：a:检验分布是正态分布；  
b:根据数据计算；  
c:Lilliefors 显著性校正；

d:这是真正显著性的下限;  
R1:第一组品酒员对红葡萄酒的评价结果;  
R2:第二组品酒员对红葡萄酒的评价结果;  
W1:第一组品酒员对白葡萄酒的评价结果;  
W2:第二组品酒员对白葡萄酒的评价结果。

### 5.1.2 问题一模型建立

#### 配对样本 T 检验

对每一对数据而言, 它们是不同的组评酒员对同一种葡萄酒(红白)评价的结果, 因此, 它们不是两个独立的随机变量的观察值, 适合用配对样本 T 检验法检验评价结果的显著性差异。

设有  $n$  对相互独立的观察结果:  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 。

令  $D_1 = x_1 - y_1, D_2 = x_2 - y_2, \dots, D_n = x_n - y_n$ , 则  $D_1, D_2, \dots, D_n$  相互独立。又由于  $D_1, D_2, \dots, D_n$  是由同一因素引起, 可认为它们服从同一分布。现假设  $D_1 \sim N(\mu, \sigma^2)$ , 也就是说  $D_1$  是正态总体  $N(\mu, \sigma^2)$  的一个样本, 其中  $\mu, \sigma^2$  未知。

我们可基于这一样本检验假设:

$H_0: \mu = 0$  (两组品酒员对酒的评价结果不存在显著性差异)

$H_1: \mu \neq 0$  (两组品酒员对酒的评价结果存在显著性差异)

分别记样本  $D_1, D_2, \dots, D_n$  的样本均值和样本方差的观察值为  $\bar{d}, s_d^2$ , 在显著性水平为  $\alpha$  的条件下检验问题的拒绝域为:

$$t = \left| \frac{\bar{d}}{s_d / \sqrt{n}} \right| \geq t_{\alpha/2}(n-1)$$

#### 5.1.2.2 两组评价结果显著性差异评价

根据样本值计算检验统计量的观测值, 判断是否落入拒绝域, 若落入拒绝域中, 即

$$P \left\{ \left| \frac{\bar{d}}{s_d / \sqrt{n}} \right| \geq t_{\alpha/2}(n-1) \right\} \leq \alpha \quad (5.1.1)$$

则拒绝  $H_0$ , 也就是说, 在显著性水平  $\alpha=0.05$  的条件下认为两组评酒员对葡萄酒的评价结果不同, 即存在显著性差异; 反之则不存在显著性差异。

#### 5.1.2.2 配对样本 T 检验评价结果可信度

查阅资料可知, 标准偏差是反映一组测量数据离散程度的统计指标, 可用来检验评价结果是否可信, 且数值越小代表可信度越高。

样本标准差的观察值  $s_d$  :

$$s_d = \sqrt{s_d^2} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (d_i - \bar{d})^2} \tag{5.1.2}$$

其中， $d_i$ ， $\bar{d}$  分别为  $D_1, D_2, \dots, D_n$  的中的第  $i$  ( $i=1, 2, \dots, n$ ) 个样本值和样本均值观察值。

5.1.3 问题一模型求解

将预先处理好的数据导入 spss，以红葡萄酒为例，将两组中每一种红葡萄酒的样品两两配对，从而进行样本 T 检验，白葡萄酒的处理方法类似，最终配对样本统计，配对样本相关性以及检验结果分别如表 2，表 3 和表 4 所示。

表 2 配对样本统计

		平均值 (E)	数字	标准偏差	标准误差平均值
配对 1	R1	72.9537037	27	7.42171872	1.42831043
	R2	70.6435185	27	3.98445050	0.76680786
配对 2	W1	75.0714286	28	4.83314518	0.91337859
	W2	77.3035714	28	3.40893919	0.64422895

表 3 配对样本相关性

		数字	相关系数	显著性
配对 1	R1 & R2	27	0.705	0.000
配对 2	W1 & W2	28	0.234	0.232

表 4 配对样本检验

		配对差值					t	自由 度	显著 性 (双 尾)
		平 均 值 (E)	标准偏 差	标 准 误 差 平 均 值	差值的 95% 置信区 间				
					下限	上限			
配 对 1	R1 － R2	2. 31019	5. 40703	1. 04058	0. 17123	4. 44914	2. 220	26	0. 035
配 对 2	W1 － W2	-2. 23214	5. 22356	0. 98716	-4. 25763	-0. 20666	-2. 261	27	0. 032

注：R1:第一组品酒员对红葡萄酒的评价结果;  
R2:第二组品酒员对红葡萄酒的评价结果;  
W1:第一组品酒员对白葡萄酒的评价结果;  
W2:第二组品酒员对白葡萄酒的评价结果。

然后，分别求得每组品酒员对同种葡萄酒的组内打分方差，如表 5，表 6 所示。

表 5 红葡萄酒打分的组内方差

序号	第一组打分组内方差	第二组打分组内方差
1	9.63846	9.04863
2	6.30784	4.02768
3	6.76921	5.54176
4	10.3944	6.42564
5	7.87471	3.69534
6	7.72873	4.59589
7	10.179	7.91693
8	6.63409	8.06915
9	5.73973	5.0728
10	5.51362	6.0148
11	8.41229	6.16802
12	8.925	5.01221
13	6.70323	3.9101
14	6	4.81202
15	9.25023	6.42996
16	4.25441	4.4833
17	9.38142	3.02765
18	6.87103	7.0899
19	6.88315	7.42668
20	5.10338	6.25033
21	10.775	5.95912
22	7.11493	4.92612
23	5.6999	4.97661
24	8.65384	3.27448
25	8.03879	6.61312
26	5.59365	6.44636
27	7.05534	4.52769
累计求和	201.495	151.742

表 6 白葡萄酒打分的组内方差

序号	第一组打分组内方差	第二组打分组内方差
1	9.60324	5.08702
2	14.1798	7.00476
3	8.30449	11.9369
4	6.68664	6.48845
5	11.2448	5.12619



6	12.7558	4.76678
7	6.25833	6.49444
8	12.769	5.57873
9	9.63155	10.3086
10	14.5835	8.39047
11	13.3087	9.37135
12	10.7605	11.834
13	13.0678	6.83862
14	10.6875	3.98469
15	11.4717	7.35149
16	13.3417	9.06826
17	12.0074	6.20125
18	12.5118	5.49848
19	6.81175	5.10338
20	8.02496	7.07421
21	13.142	8.02496
22	11.7757	7.3212
23	6.60724	3.40588
24	10.5415	6.20842
25	5.82046	10.3199
26	8.53815	10.144
27	12.0167	5.96285
28	8.9697	5.03764
累计求和	295.4222	199.9329

#### 5.1.4 问题一结果分析

##### 两组评价结果显著性差异评价分析

由表 4 中的统计数据可知，无论是红葡萄酒还是白葡萄酒经配对数据 T 检验，双尾显著性水平均小于 0.05，可以认为这两组品酒员对红白葡萄酒的评价结果具有显著性差异。

##### 两组评价结果可信度分析

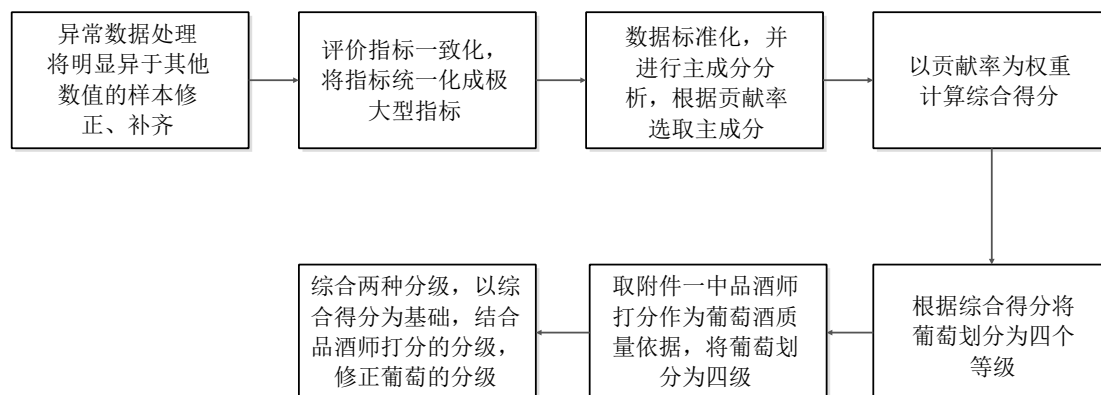
综合表 4，表 5 以及表 6 的数据可以看出，无论是对于红葡萄酒亦或是对白葡萄酒而言，第二组品酒员的组内成员间对同种酒的打分的标准差都要更小，即第二组品酒员的十个成员间的评判差别更小，由此我们可以判断出第二组品酒员关于红白葡萄酒的评价结果更可信。

## 5.2 问题二模型的建立与求解

### 5.2.1 问题二模型分析

问题二要求根据酿酒葡萄的理化指标和葡萄酒的质量，对酿酒葡萄进行分级。

对于酿酒葡萄的理化指标与酿酒葡萄之间的关系，运用主成分分析法，以主成分的贡献率为权重，建立对葡萄酒的主成分综合评价模型从而对酿酒葡萄进行分级；而葡萄酒的质量与酿酒葡萄的之间的关系，可参考问题一中第二组品酒员的打分情况对葡萄进行排序，从而修正酿酒葡萄的分级情况。



问题二流程图

## 数据处理

### 1) 异常数据处理

观察附件 2 的表格发现有数据存在明显错误，例如酿酒葡萄理化指标中白葡萄百粒质量的第三次监测值为 2226.1g，而前两次的检测值分别为 225.8g 和 224.6g，为避免异常数据对数据分析的影响，我们用前两次的检测值均值代替异常数据。

### 2) 评价指标类型的一致化处理

在已建立的指标体系中，指标集可能同时含有“极大型”和“极小型”指标，我们分别称之为优质因子和劣质因子，也存在“中间型”指标。因此在评价之前必须将评价指标的类型一致化，即要统一为极大型。

极小型：对于某个极小型指标  $x_i$ ，通过平移变换

$$x_i' = M_i - x_i \quad (x_i > 0, i = 1, 2, \dots, 27)$$

其中  $M_i$  为指标  $x_i$  可能取到的最大值。

中间型：对于某个中间型指标  $x_i$ ，要将其化为极大值指标，令

$$x_i' = \begin{cases} \frac{2(x_i - m_i)}{M_i - m_i}, & m_i \leq x_i \leq \frac{1}{2}(M_i + m_i) \\ \frac{2(M_i - x_i)}{M_i - m_i}, & \frac{1}{2}(M_i + m_i) \leq x_i \leq M_i \end{cases}$$

其中  $M_i$ ， $m_i$  分别为指标  $x_i$  可能取到的最大值和最小值。

## 5.2.2 问题二模型建立

对酿酒葡萄而言，虽然各种指标在成因上互不相同，但是不同的指标之间往往存在相关性，为找出这些相关性，本文选用主成分分析法解决这些问题。

基于主成分分析法的评价步骤如下：

### 1) 对原始数据进行标准化处理

已知涉及的指标共  $m$  个，样本对象  $n$  个，第  $i$  个样本的第  $j$  个指标值为  $a_{ij}$ ，将各项指标值  $a_{ij}$  转化为标准化指标  $a_{ij}$ ，有

$$a_{ij} = \frac{a_{ij} - \mu_j}{s_j}$$

其中第  $j$  个指标的样本均值  $\mu_j = \frac{1}{n} \sum_{i=1}^n a_{ij}$ ，第  $j$  个指标的样本标准差

$$s_j = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (a_{ij} - \mu_j)^2}, \quad j=1, 2, \dots, m, \quad \text{相应地，称}$$

$$x_j = \frac{x_j - \mu_j}{s_j}, \quad j=1, 2, \dots, m$$

为标准化指标变量。

### 2) 计算相关系数矩阵 $R$

记第  $i$  个指标与第  $j$  个指标的相关系数为  $r_{ij}$ ，则

$$r_{ij} = \frac{\sum_{k=1}^n a_{ki} * a_{kj}}{n-1}, \quad i, j=1, 2, \dots, m$$

那么相关系数矩阵为  $R = (r_{ij})_{m \times m}$ ，其中  $r_{ii} = 1$ ， $r_{ij} = r_{ji}$ 。

### 3) 计算特征值和特征向量

计算相关系数矩阵  $R$  的特征值  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m \geq 0$ ，及其对应的特征向量  $u_1, u_2, \dots, u_m$ ，其中  $u_j = [u_{1j}, u_{2j}, \dots, u_{mj}]^T$ ，由特征向量组成  $m$  个新的指标变量：

$$\begin{cases} y_1 = u_{11}x_1 + u_{21}x_2 + \dots + u_{m1}x_m \\ y_2 = u_{12}x_1 + u_{22}x_2 + \dots + u_{m2}x_m \\ \vdots \\ y_m = u_{1m}x_1 + u_{2m}x_2 + \dots + u_{mm}x_m \end{cases}$$

其中， $y_j$  为第  $j$  个主成分， $j=1, 2, \dots, m$ 。

### 4) 选择 $p(p \leq m)$ 个主成分，计算综合评价

a) 计算特征值  $\lambda_j (j=1, 2, \dots, m)$  的信息贡献率和累积贡献率。

称

$$b_j = \frac{\lambda_j}{\sum_{k=1}^m \lambda_k}, j = 1, 2, \dots, m$$

为主成分  $y_j$  的信息贡献率；而且称

$$\alpha_p = \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^m \lambda_k}$$

为主成分  $y_1, y_2, \dots, y_p$  的累积贡献率。当  $\alpha_p$  接近于1（ $\alpha_p=0.85, 0.90, 0.95$ ）时，则选择前  $p$  个指标变量  $y_1, y_2, \dots, y_p$  作为  $p$  个主成分，代替原来的5个指标变量，从而对  $p$  个主成分进行综合分析。

#### b) 计算综合得分

$$Z = \sum_{j=1}^p b_j y_j \quad (5.2.1)$$

其中， $b_j$  为第  $j$  个主成分的信息贡献率。

### 5.2.3 问题二模型求解

以红葡萄为例，利用 matlab 首先进行数据标准化并求得相关系数矩阵（程序见附录）从矩阵中可以发现某些指标具有很强的相关性，如果直接用这些指标进行分析不仅会使运算量过大，而且还会影响分级客观性。运用主成分分析法，可求得红葡萄相应主成分的特征值及其贡献率如表 7。

**表 7 红葡萄特征值及其贡献率**

序号	特征值	贡献率	累积贡献率
1	6.92	22.31	22.31
2	4.70	15.17	37.48
3	4.16	13.41	50.88
4	3.09	9.95	60.84
5	1.96	6.33	67.17
6	1.77	5.71	72.87
7	1.54	4.95	77.83
8	1.30	4.20	82.02
9	1.08	3.48	85.50
10	0.74	2.38	87.88

11	0.73	2.34	90.22
12	0.66	2.14	92.35
13	0.47	1.51	93.87
14	0.37	1.20	95.07
15	0.34	1.09	96.16
16	0.32	1.02	97.19
17	0.23	0.76	97.94
18	0.18	0.57	98.51
19	0.13	0.42	98.93
20	0.09	0.29	99.23
21	0.08	0.25	99.48
22	0.06	0.20	99.68
23	0.04	0.12	99.80
24	0.03	0.10	99.90
25	0.03	0.09	99.99
26	0.00	0.01	100.00
27	0.00	0.00	100.00
28	0.00	0.00	100.00
30	0.00	0.00	100.00
31	0.00	0.00	100.00
32	0.00	0.00	100.00

可以看出，前 11 个特征值的累积贡献率达到 90%以上，主成分分析效果较好。现选取前 11 个主成分进行综合评价，求出前 11 个特征向量对应的特征向量，由此可得 11 个主成分。

然后分别以 11 个主成分的贡献率为权重，构建主成分综合评价模型：

$$Z=0.2231y_1+0.1517y_2+0.1341y_3+0.0995y_4+0.0633y_5+0.0571y_6+0.0495y_7+0.042y_8+0.0348y_9+0.0238y_{10}+0.0234y_{11}$$

把各样本酿酒葡萄的 11 个主成分值代入上式，可以得到红葡萄的排名和综合评价结果，如表 8 所示。

**表 8 红葡萄排名和综合评价结果**

名次	对应得分的红葡萄序号	综合评价值
1	3	1.71889
2	9	1.36977
3	23	0.88569
4	2	0.86935
5	17	0.85283
6	21	0.58774
7	5	0.44454

8	1	0.32952
9	24	0.30408
10	14	0.23348
11	12	0.21192
12	6	0.2038
13	22	0.0131
14	8	-0.0063
15	19	-0.1455
16	15	-0.2287
17	20	-0.2692
18	16	-0.2795
19	18	-0.4082
20	7	-0.5099
21	13	-0.5107
22	4	-0.5621
23	26	-0.603
24	27	-0.656
25	25	-1.0684
26	10	-1.2169
27	11	-1.5603

为进行红葡萄的分级,我们规定综合评价值大于 0.85 即可认定是一级葡萄;综合评价值大于 0 且小于 0.85 则为二级葡萄;综合评价值大于-1 且小于 0 则为三级葡萄;剩下的则是四级葡萄。由此可求得红葡萄的分级情况,如表 9 所示。

**表 9 酿酒红葡萄的分级**

一级葡萄（大于 0.85）	2	3	9	23								
二级葡萄（大于 0）	1	5	6	12	14	17	21	22	24			
三级葡萄（大于-1）	4	7	8	13	15	16	18	19	20	26	27	
四级葡萄（小于-1）	10	11	25									

对葡萄酒的质量而言,将第二组品酒员红葡萄酒的综合得分进行排序,若表 7 中的关于葡萄的分级情况与第二组品酒员评分结果相差大于五个,则根据品酒员评级进行个例修正,具体情况见表 10。

**表 10 参考品酒员打分的酿酒红葡萄修正后的分级**

一级葡萄（大于 0.85）	2	3	9	23							
二级葡萄（大于 0）	5	10	14	17	19	20	21	22	24	26	
三级葡萄（大于-1）	1	4	6	7	8	12	13	15	16	27	
四级葡萄（小于-1）	11	18	25								

同理可求得参考品酒员打分的酿酒红葡萄修正后的分级情况,如表 11 所示。

表 11 参考品酒员打分的酿酒白葡萄修正后的分级

一级葡萄（大于 0.85）	5	9											
二级葡萄（大于 0）	10	12	15	17	18	20	21	22	23	24	25	27	28
三级葡萄（大于-1）	1	2	3	4	6	7	11	13	14	19	26		
四级葡萄（小于-1）	8	16											

5.2.4 问题二结果分析

我们将红葡萄从优到劣分为了四个等级，其中编号为 2、3、9、23 的葡萄为一级葡萄，编号为 5、10、14、17、19、20、21、22、24、26 为二级葡萄，编号为 1、4、6、7、8、12、13、15、16、27 为三级葡萄，编号为 11、18、25 为四级葡萄。

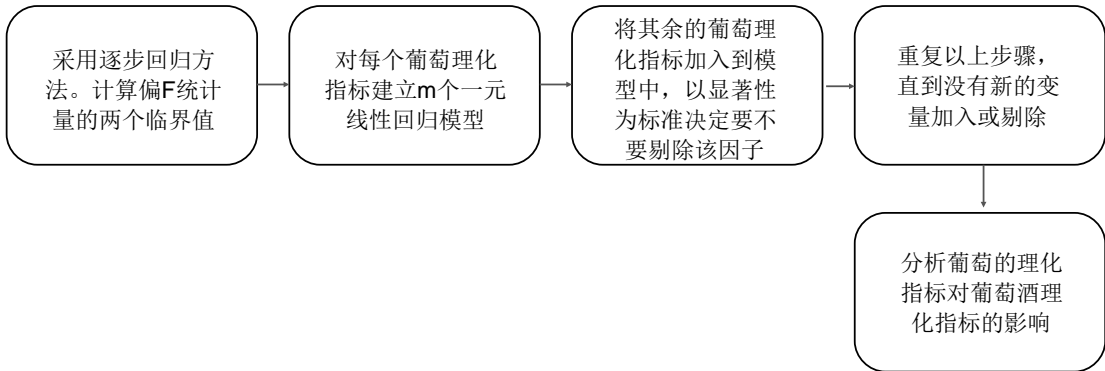
同理，我们将白葡萄从优到劣同样分为了四个等级，其中编号为 5、9 的葡萄为一级葡萄，编号为 10、12、15、17、18、20、21、22、23、24、25、27、28 为二级葡萄，编号为 1、2、3、4、6、7、11、13、14、19、26 为三级葡萄，编号为 8、16 为四级葡萄。

无论是红葡萄还是白葡萄，酿酒葡萄主要处于二、三等级。

5.3 问题三模型的建立与求解

5.3.1 问题三模型分析

问题三要求分析酿酒葡萄与葡萄酒的理化指标之间的联系，我们采用了逐步回归的方法，建立酿酒葡萄一级理化指标与葡萄酒一级理化指标间的回归模型，得出酿酒葡萄一级理化指标对葡萄酒各个理化指标的函数关系。



问题三流程图

5.3.2 问题三模型建立

逐步回归分析

若候选的自变量集合为  $S = \{x_1, x_2, \dots, x_m\}$ ，然后将变量一个一个得引入，引入变量的条件是通过偏  $F$  统计量的检验。同时，每引入一个新的变量后，对已入选方程的老变量进行检验，将经检验认为不显著的变量剔除，此过程经过若干步，直到既不能引入新变量，又不能剔除老变量为止。

设模型中已有  $l-1$  个自变量，记这些自变量的集合为  $A$ ，当不在中的一个自变量  $x_k$  加入到这个模型中时，偏  $F$  统计量的一般形式为

$$F = \frac{SSE(A) - SSE(A, x_k)}{SSE(A, x_k)/n-l-1} = \frac{SSR(x_k|A)}{MSE(A, x_k)}$$

**基于逐步回归分析方程建立一般步骤:**

**Step1:** 根据一定显著水平, 给出偏  $F$  统计量的两个临界值, 一个用作选取自变量, 记为  $F_E$ ; 另一个用作剔除自变量, 记为  $F_D$ 。一般地, 取  $F_E > F_D$ 。

**Step2:** 对每个自变量  $x_i (1 \leq i \leq m)$  拟合  $m$  个一元线性回归模型

$$y = \beta_0 + \beta_i x_i + \varepsilon$$

这时,  $A$  相当于为空集, 因此  $SSR(x_i|A) = SSR(x_i)$ ,  $MSE(A, x_i) = MSE(x_i)$ , 故对每一个  $i$ , 计算

$$F_i^{(1)} = \frac{MSR(x_i)}{MSE(x_i)}, i = 1, 2, \dots, m$$

设

$$F_{i_1}^{(1)} = \max_{1 \leq i \leq m} F_i^{(1)}$$

若  $F_{i_1}^{(1)} > F_E$ , 则选择含自变量  $x_{i_1}$  的回归模型为当前模型, 否则, 没有自变量能进入模型, 选择过程结束, 即认为所有自变量影响均不显著。

**Step3:** 在 Step2 的选出模型的基础上, 再将其余  $m-1$  的个自变量分别加入到此模型中, 得到  $m-1$  个二元回归方程, 计算

$$F_i^{(2)} = \frac{SSR(x_i|x_{i_1})}{MSE(x_i, x_{i_1})}, i \neq i_1$$

设

$$F_{i_2}^{(2)} = \max_{i \neq i_1} F_i^{(2)}$$

若  $F_{i_2}^{(2)} > F_E$ , 则选取过程结束。Step2 选择的模型为最优模型; 若  $F_D < F_{i_2}^{(2)} < F_E$  则将自变量  $x_{i_2}$  选入模型中, 即得 Step3 的模型:

$$y = \beta_0 + \beta_{i_1} x_{i_1} + \beta_{i_2} x_{i_2} + \varepsilon$$

进一步考察, 当  $x_{i_2}$  进入模型后,  $x_{i_1}$  对  $y$  的影响是否仍然显著。为此计算



$$F_{i_2}^{(2)} = \frac{SSR(x_{i_1} | x_{i_2})}{MSE(x_{i_2}, x_{i_1})}$$

若  $F_{i_2}^{(2)} \leq F_D$ ，则剔除  $x_{i_1}$ ，这时仅含有  $x_{i_2}$  的回归模型为当前模型。

**Step4:** 在 Step3 所选模型的基础上，在将余下的自变量逐个加入，拟合各个模型并计算统计量值，与比较决定是否有新变量引入，如果有新变量进入模型，还需要检验原模型中的老变量是否因这个新变量的进入而不再显著，那样就应该被剔除。

**Step5:** 重复以上步骤，直到没有新的自变量能进入模型，同时在模型之中的老变量都不能剔除，则结束选择过程，最后一个模型即为所求的最优回归模型。故所求模型为

$$\begin{cases} y = \beta_0 + \sum_{k=1}^p \beta_{i_k} x_{i_k} + \varepsilon \\ \varepsilon \sim N(0,1) \end{cases} \quad (5.3.1)$$

其中， $x_{i_k}$  为独立性强的显著变量， $p$  为显著变量的个数。

### 5.3.3 问题三模型求解

使用 spss 软件逐一对葡萄酒理化指标进行逐步回归分析，得到红葡萄酒理化指标的回归方程如表 10 所示。

**表 12 红葡萄酒理化指标及回归方程**

理化指标	回归方程
花色苷	$y_1 = 2.65x_4 - 6.74x_{26} + 437.51$
单宁	$y_2 = -0.059x_8 + 0.003x_9 + 17.34x_{10} + 0.06x_{18} - 12.52$
总酚	$y_3 = 0.008x_4 + 0.253x_{11} + 1.68$
酒总黄酮	$y_4 = -0.95x_{11} + 0.398$
DPPH 半抑制体积	$y_5 = -0.017x_{11} - 0.024$
色泽(L)	$y_6 = -0.144x_4 - 1.223x_5 - 0.124x_{15} + 3.561x_{27} + 61.908$
色泽(a)	$y_7 = -0.094x_4 - 4.818x_{27} + 69.169$
色泽(b)	$y_8 = 0.831x_5 - 0.737x_6 - 0.129x_{17} - 8.859$

同理可得，白葡萄酒理化指标及回归方程如表 11 所示。

**表 13 白葡萄酒理化指标就回归方程**

理化指标	回归方程
单宁	$y_1 = 0.242x_{12} + 4.716x_{27} - 0.014$
总酚	$y_2 = 0.174x_{13} + 0.016x_{18} - 2.503$
酒总黄酮	$y_3 = -0.359x_{11} + 0.062x_{21} - 2.948x_{19} + 2.924$
DPPH 半抑制体积	$y_4 = 0.048x_3 + 0.008x_{13} - 0.105$
色泽 (L)	$y_5 = -0.086x_{22} + 0.026x_{26} + 101.96$
色泽 (a)	$y_6 = -0.039x_6 - 1.104$
色泽 (b)	$y_7 = -0.15x_{26} + 15.509$

表格 10 与表 11 中符号含义表 12 所示。

表 14 表格 10 与表 11 中符号含义

氨基酸总量	蛋白质	VC 含量	花色苷	酒石酸
$x_1$	$x_2$	$x_3$	$x_4$	$x_5$
苹果酸	柠檬酸	多酚氧化酶活力	褐变度	DPPH 自由基
$x_6$	$x_7$	$x_8$	$x_9$	$x_{10}$
总酚	单宁	葡萄总黄酮	白藜芦醇	黄酮醇
$x_{11}$	$x_{12}$	$x_{13}$	$x_{14}$	$x_{15}$
总糖	还原糖	可溶性固形物	PH 值	可滴定酸
$x_{16}$	$x_{17}$	$x_{18}$	$x_{19}$	$x_{20}$
固酸比	干物质含量	果穗质量	百粒质量	果梗比
$x_{21}$	$x_{22}$	$x_{23}$	$x_{24}$	$x_{25}$
出汁率	果皮质量	果皮颜色 L	果皮颜色 A	果皮颜色 B
$x_{26}$	$x_{27}$	$x_{28}$	$x_{29}$	$x_{30}$

### 5.3.4 问题三结果分析

根据上述红、白葡萄酒理化指标的回归方程我们可以分析出酿酒葡萄与葡萄酒的理化指标之间的联系。

我们绘出各项葡萄酒理化指标与葡萄理化指标之间相关图，部分如图 3.1 和 3.2 所示。（余下见附件）。

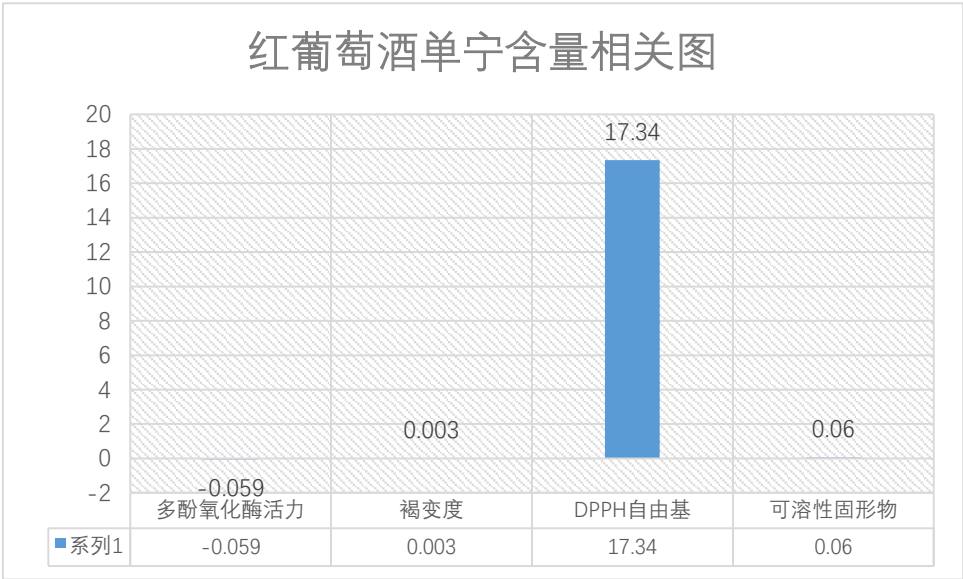


图 3.1 红葡萄酒单宁含量相关图

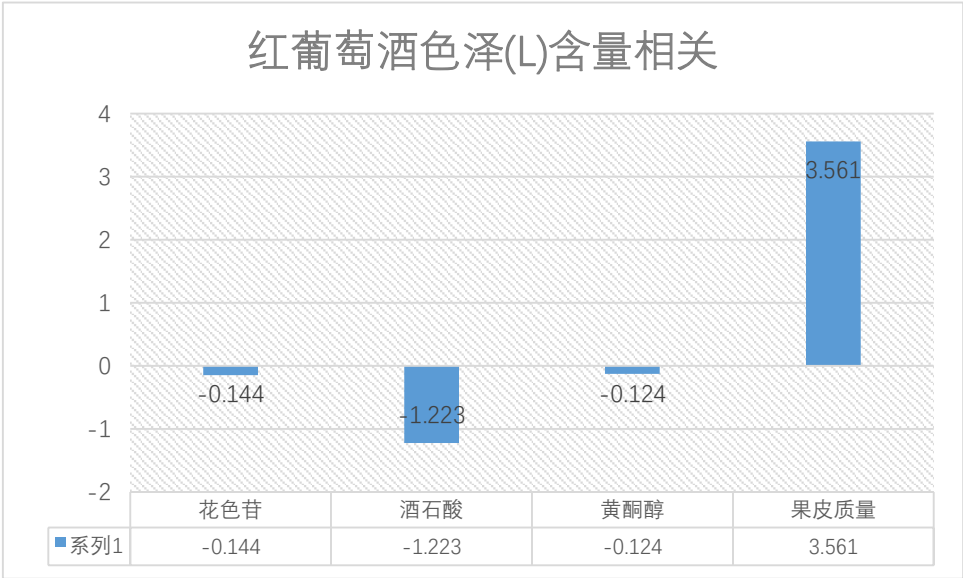


图 3.2 红葡萄酒色泽(L)含量相关图

分析如下：

**A. 红葡萄酒：**

花色苷与酿酒葡萄中花色苷和出汁率有关，酿酒葡萄中的花色苷含量越多、出汁率越低，葡萄酒中的花色苷含量越多；

单宁指标与酿酒葡萄可溶性固体物、DPPH 自由基成正相关，与多酚氧化酶活力和褐变度成负相关，主要决定于酿酒葡萄中 DPPH 自由基和褐变度；

红葡萄酒中的总酚含量主要决定于酿酒葡萄中总酚的含量；

酒总黄酮的含量与总酚成负相关，DPPH 半抑制体积与总酚含量成负相关；

色泽与酿酒葡萄中的果皮质量、酒石酸相关。

**B. 白葡萄酒：**

单宁的含量与酿酒葡萄果皮质量、单宁成正相关；

总酚含量与酿酒葡萄中葡萄总黄酮；

酒总黄酮含量与酿酒葡萄固酸比成正相关，与酿酒葡萄的 PH 值和总酚成负相关；

DPPH 半抑制体积与酿酒葡萄中葡萄总黄酮和 VC 含量成负相关。  
最后，白葡萄酒的色泽与干物含量和出汁率有关。  
部分相关性分析如图所示。

5.4 问题四模型的建立与求解

5.4.1 问题四模型分析

问题四要求分析酿酒葡萄和葡萄酒的理化指标对葡萄酒质量的影响，并论证能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量。为了分析葡萄与葡萄酒的理化指标对葡萄酒质量的影响，分别对葡萄酒质量的评价因素：外观、香气、口感，根据理化指标作逐步回归。这样就可以找出主要影响葡萄酒质量评价的理化指标。根据假设：葡萄酒质量中的香气评定主要受葡萄酒中芳香物质的影响，对葡萄酒的香气评分与葡萄和葡萄酒的芳香物质做逐步回归方程，找到芳香物质与香气评价的关系。同理，可以通过寻找葡萄酒质量另外两个评价因素的逐步回归方程，综合三个评价因素就可以分析得到葡萄酒的质量。进而理化指标对葡萄酒质量的影响。

利用上面做出的逐步回归方程，分别对外观、香气、口感3个评价因素做预测评分，将这个评分与原来品酒员的评分做比较，从而判断能否用葡萄和葡萄酒的理化指标来评价葡萄酒的质量。

5.4.2 问题四模型建立

借用问题三中建立的基于逐步回归的综合评价模型，可用来预测香气、外观、口感根据相关理化指标计算出的相应的评分。

5.4.3 问题四模型求解

对红葡萄酒外观、口感、香气的逐步回归结果如下表 13 所示：

表 5.4.1 红葡萄酒的外观、口感、香气的逐步回归结果

外观	$y_1 = -0.308x_1 - 0.216x_2 + 4.544x_3 + 0.094x_4 + 15.428$
口感	$y_2 = 0.393x_5 + 3.007x_6 + 0.078x_7 + 18.822$
香气	$y_3 = 1.188x_8 + 0.469x_9 - 0.948x_{10} + 3.613x_{11} - 4.017x_{12} - 0.506x_{13} + 20.822$

其中  $y_1, y_2, y_3$  分别代表红葡萄酒的外观、口感、香气的预测评分。

而  $x_1, x_2, x_3 \dots x_{13}$  分别代表葡萄酒色泽  $L$ 、干物质含量、果皮质量、葡萄酒白藜芦醇、果皮颜色  $a$ 、 $ph$  值、葡萄酒色泽  $b$ 、(Z)-3,7-二甲基-2,6-辛二烯醛、辛酸3-甲基丁酯、丁二酸二乙酯、2,5-二(1,1-二甲基乙基)-1,4-苯二醇、7-甲氧基-2,2,4,8-四甲基-三环[5.3.1.0(4,11)]十一碳烷、正十三烷。

对白葡萄酒外观、口感、香气的逐步回归结果如下表 14 所示：

表 5. 4. 2 白葡萄酒的外观、口感、香气的逐步回归结果

外观	$y_1 = 0.077x_1 + 0.262x_2 + 8.258$
口感	$y_2 = 0.536x_3 - 0.008x_4 - 1.6078x_5 + 32.828$
香气	$y_3 = 0.45x_6 - 0.053x_7 - 0.743x_8 + 0.202x_9 - 0.477x_{10} + 23.61$

其中  $y_1, y_2, y_3$  分别代表白葡萄酒的外观、口感、香气的预测评分；

$x_1, x_2, x_3 \dots x_{10}$  分别代表果皮颜色  $B$ 、白葡萄酒色泽  $B$ 、可滴定酸、果穗质量、白葡萄酒白藜芦醇、香叶基乙醚  $(E)-2-$ 己烯醛 1-己醇 1-丙醇 苯甲酸。

误差分析

以红葡萄酒为例，将数据代入问题四中的建立的模型得到红葡萄酒关于外观、香气、口感的预测得分，并分与附件 1 中的第二组品酒员对于红葡萄酒的各个项目的实际得分进行比较，具体内容见表 15，表 16 和表 17。

表 5. 4. 3 红葡萄酒外观预测分 and 实际得分比较

序号	预测得分	实际得分	相对误差
1	10.4645	10.7	2.20%
2	10.35828	10.1	-2.56%
3	9.807321	10.2	3.85%
4	9.763219	9.9	1.38%
5	10.61304	10.8	1.73%
6	8.645643	8.7	0.62%
7	8.288806	7.5	-10.52%
8	10.84413	10.2	-6.32%
9	10.71064	11	2.63%
10	10.22248	10.6	3.56%
11	7.255505	7	-3.65%
12	8.732478	8.7	-0.37%
13	9.656289	9.5	-1.65%
14	10.67299	10.7	0.25%
15	9.064712	10	9.35%
16	9.406147	10	5.94%
17	9.840992	10.2	3.52%
18	8.418786	7.8	-7.93%
19	10.13825	10.5	3.45%
20	8.635476	8.4	-2.80%
21	10.10918	10.2	0.89%
22	9.429272	9.4	-0.31%

23	12.07479	11.4	-5.92%
24	9.990501	10.1	1.08%
25	10.76318	10.3	-4.50%
26	11.18959	11.1	-0.81%
27	9.345826	9.9	5.60%

**表 5.4.4 红葡萄酒口感预测分和实际得分比较**

序号	预测得分	实际得分	相对误差
1	29.52302	29.1	1.45%
2	32.33221	32.7	-1.12%
3	32.30199	32.9	-1.82%
4	30.07841	30	0.26%
5	31.28752	31.5	-0.67%
6	29.83345	28.9	3.23%
7	30.08007	29.3	2.66%
8	28.43177	28	1.54%
9	31.38908	32.3	-2.82%
10	29.57005	28.8	2.67%
11	26.1959	25.8	1.53%
12	30.72902	31.1	-1.19%
13	31.54779	28.4	11.08%
14	31.30015	30.8	1.62%
15	29.11519	28.6	1.80%
16	29.63068	30.4	-2.53%
17	30.63163	32	-4.28%
18	29.27717	30.1	-2.73%
19	30.86618	30.9	-0.11%
20	30.74756	33.2	-7.39%
21	31.80648	31.6	0.65%
22	31.23377	31.2	0.11%
23	30.42754	31.7	-4.01%
24	30.6939	30.6	0.31%
25	29.29766	28.2	3.89%
26	31.74414	30	5.81%
27	29.28503	31.3	-6.44%

**表 5.4.5 红葡萄酒香气预测分和实际得分比较**

序号	预测得分	实际得分	相对误差
1	19.59038	19.9	-1.56%
2	21.73952	22.1	-1.63%
3	21.93904	22.6	-2.92%
4	20.89272	22.5	-7.14%
5	21.42089	20.9	2.49%
6	20.70295	20.1	3.00%
7	20.14516	20.1	0.22%

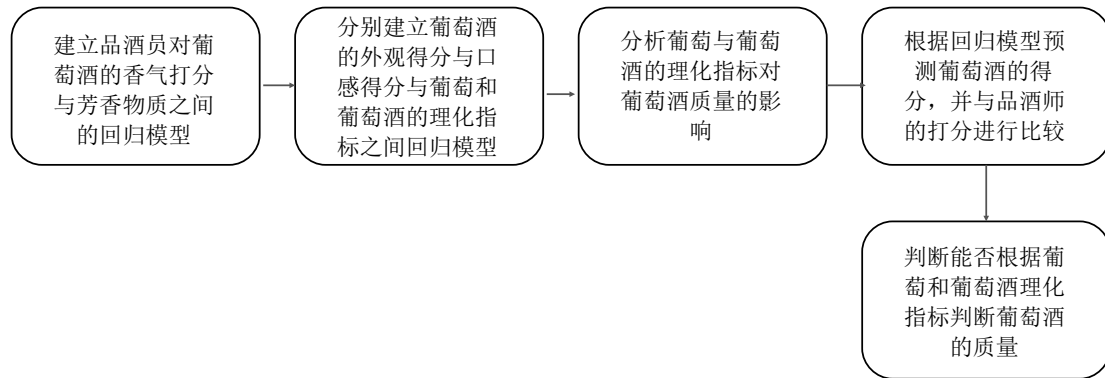
8	19.4263	19.4	0.14%
9	25.1291	25.5	-1.45%
10	22.30156	21	6.20%
11	21.48045	20.7	3.77%
12	20.22425	19.8	2.14%
13	22.1203	22.3	-0.81%
14	21.25683	21.8	-2.49%
15	20.04868	18.9	6.08%
16	20.28761	20.7	-1.99%
17	22.34643	23.3	-4.09%
18	19.68693	18.7	5.28%
19	22.63515	22.5	0.60%
20	25.05633	24.9	0.63%
21	20.71595	21.4	-3.20%
22	22.34429	22.1	1.11%
23	25.72696	25.1	2.50%
24	21.80247	22	-0.90%
25	20.95675	21.1	-0.68%
26	21.72569	22.1	-1.69%
27	21.33319	21.5	-0.78%

综合表5.4.3，表5.4.4和表5.4.5可看出实际值和理论值之间的误差较小，因此，我们认为香气、外观、口感的预测模型能较好的根据相关理化指标计算出相应的评分。这样可以利用葡萄酒和葡萄的理化指标等客观性因素计算出一个较为可靠的葡萄酒质量评分的办法，为带有个人主观因素葡萄酒的评价提供一个科学的佐证。

#### 5.4.4 问题四结果分析

由表 5.4.1 我们可以判断出，果皮质量、pH 值和四甲基-三环[5.3.1.0(4,11)]十一碳烷分别是影响红葡萄酒的外观、口感以及香气的主要因素。而从表 5.4.2 我们亦可以判断出白葡萄酒色泽 *B*、白葡萄酒白藜芦醇、乙醇分别是影响白葡萄酒外观、口感和香气的主要因素。

又综合表5.4.3，表5.4.4和表5.4.5可看出通过回归方程求得的预测值与实际值之间的相对误差基本都在5%以内，误差较小，因此，我们认为香气、外观、口感的预测模型能较好的反应品酒师们的评分，即可以反应葡萄酒的质量。以上，我们判断，可以利用葡萄酒和葡萄的理化指标等客观性因素判断出葡萄酒的质量。



问题四流程图

## 六、 结果分析与检验

对于问题一，由表 4 中的统计数据可知，无论是红葡萄酒还是白葡萄酒经配对数据 T 检验，双尾显著性水平均小于 0.05，可以认为这两组品酒员对红白葡萄酒的评价结果具有显著性差异。综合表 4，表 5 以及表 6 的数据可以看出，无论是对于红葡萄酒亦或是对白葡萄酒而言，第二组品酒员的组内成员间对同种酒的打分的标准差都要更小，即第二组品酒员的十个成员间的评判差别更小，由此我们可以判断出第二组品酒员关于红白葡萄酒的评价结果更可信。

对于问题二，我们将红葡萄从优到劣分为了四个等级，其中编号为 2、3、9、23 的葡萄为一级葡萄，编号为 5、10、14、17、19、20、21、22、24、26 为二级葡萄，编号为 1、4、6、7、8、12、13、15、16、27 为三级葡萄，编号为 11、18、25 为四级葡萄。

同理，我们将白葡萄从优到劣同样分为了四个等级，其中编号为 5、9 的葡萄为一级葡萄，编号为 10、12、15、17、18、20、21、22、23、24、25、27、28 为二级葡萄，编号为 1、2、3、4、6、7、11、13、14、19、26 为三级葡萄，编号为 8、16 为四级葡萄。

无论是红葡萄还是白葡萄，酿酒葡萄主要处于二、三等级。

对于问题三，根据红、白葡萄酒理化指标的回归方程我们可以分析出酿酒葡萄与葡萄酒的理化指标之间的联系。

**红葡萄酒：**花色苷与酿酒葡萄中花色苷和出汁率有关，酿酒葡萄中的花色苷含量越多、出汁率越低，葡萄酒中的花色苷含量越多；

单宁指标与酿酒葡萄可溶性固体物、DPPH 自由基成正相关，与多酚氧化酶活力和褐变度成负相关，主要决定于酿酒葡萄中 DPPH 自由基和褐变度；

红葡萄酒中的总酚含量主要决定于酿酒葡萄中总酚的含量；

酒总黄酮的含量与总酚成负相关，DPPH 半抑制体积与总酚含量成负相关；

色泽与酿酒葡萄中的果皮质量、酒石酸相关。

**白葡萄酒：**单宁的含量与酿酒葡萄果皮质量、单宁成正相关；

总酚含量与酿酒葡萄中葡萄总黄酮；

酒总黄酮含量与酿酒葡萄固酸比成正相关，与酿酒葡萄的 PH 值和总酚成负相关；

DPPH 半抑制体积与酿酒葡萄中葡萄总黄酮和 VC 含量成负相关。

最后，白葡萄酒的色泽与干物含量和出汁率有关。

对于问题四，由表 5.4.1 我们可以判断出，果皮质量、ph 值和四甲基-三环



[5.3.1.0(4,11)]十一碳烷分别是影响红葡萄酒的外观、口感以及香气的主要因素。

而从表 5.4.2 我们亦可以判断出白葡萄酒色泽  $B$ 、白葡萄酒白藜芦醇、乙醇分别是影响白葡萄酒外观、口感和香气的主要因素。

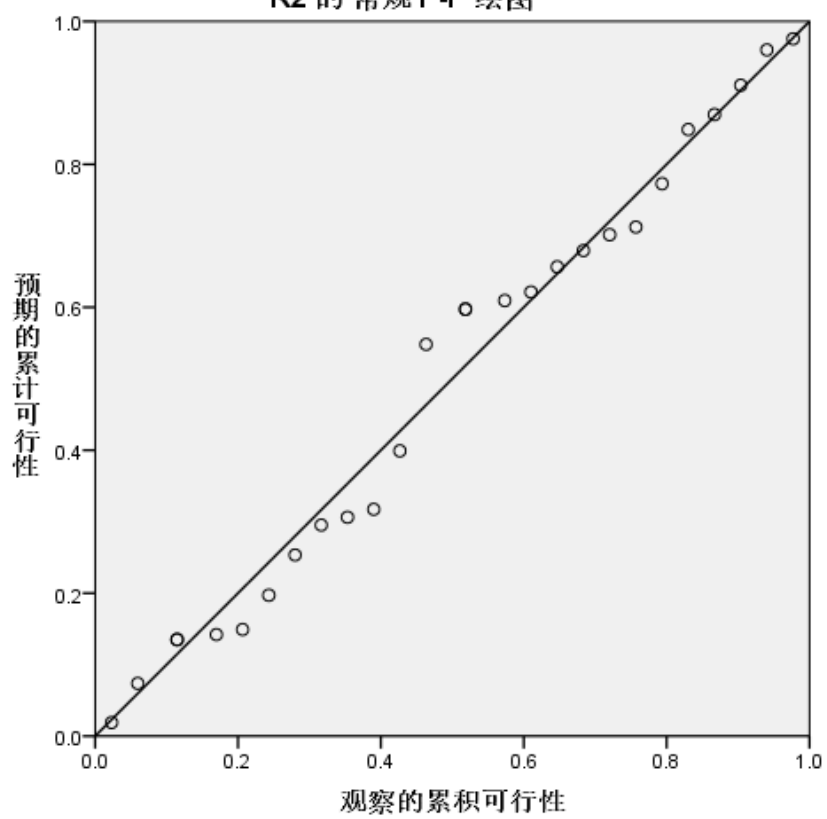
又综合表5.4.3，表5.4.4和表5.4.5可看出通过回归方程求得的预测值与实际值之间的相对误差基本都在5%以内，误差较小，因此，我们认为香气、外观、口感的预测模型能较好的反应品酒师们的评分，即可以反应葡萄酒的质量。以上，我们判断，可以利用葡萄酒和葡萄的理化指标等客观性因素判断出葡萄酒的质量。

## 七、 参考文献

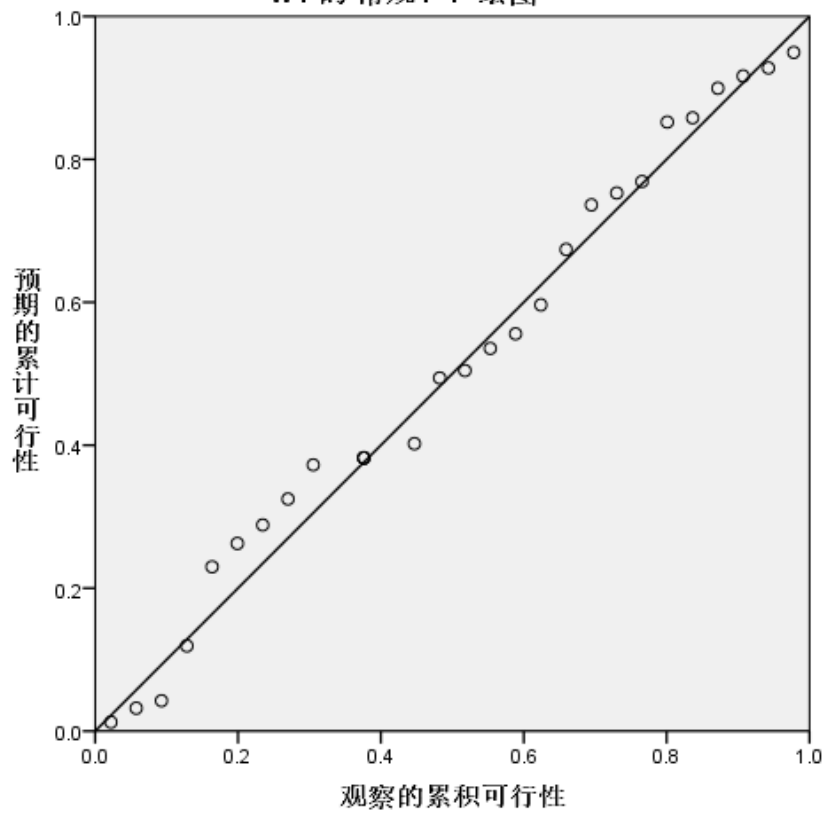
- [1] 孔告化、何铭、胡国雷，概率统计与随机过程，北京：人民邮电出版社。
- [2] 蒋辉, 我国人口增长与粮食生产关系的因子分析, 统计教育, 2005-9 第九期。
- [3] 王文静，感官评价在葡萄酒研究中的应用，科研与生产，2007-2 第 145 期
- [4] 司守奎、孙玺菁，数学建模算法与应用，北京：国防工业出版社。
- [5] 徐玖平、胡知能、李军，运筹学（II 类），北京：科学出版社，2004。

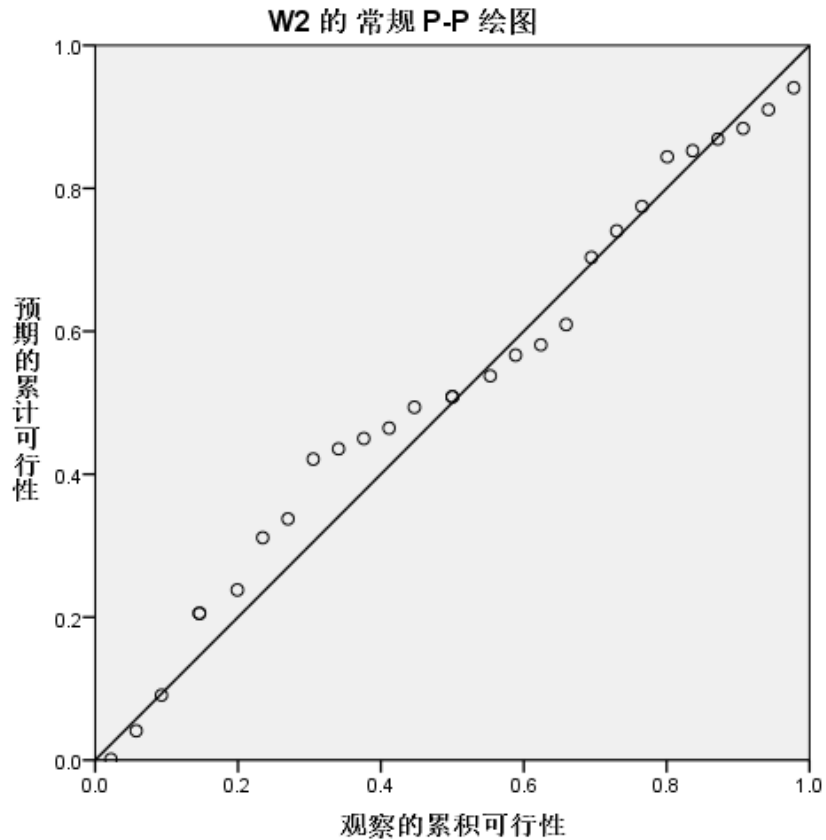
## 八、 附件

R2 的常规 P-P 绘图



W1 的常规 P-P 绘图





Matalab 代码:

问题二:

```
clear;
% load rzhibiao;
% load zhibiaogai;
load baiputaozhibiao;
%主成分分析法
baizhibiao=zscore(baizhibiao);
r=corrcoef(baizhibiao);
[x,y,z]=pcacov(r);
f=repmat(sign(sum(x)),size(x,1),1);
x=x.*f;
df=baizhibiao*x(:,1:13);
tf=df*z(1:13)/100;
[stf,ind]=sort(tf,'descend')
```

Spss 输出结果

模型摘要

模型	R	R 平方	调整后的 R 平方	标准估算的误差
----	---	------	-----------	---------

1	.757 <sup>a</sup>	.574	.557	.66535
2	.843 <sup>b</sup>	.710	.687	.55963

a. 预测变量: (常量), 出汁率

b. 预测变量: (常量), 出汁率, 百粒质量

系数<sup>a</sup>

模型	非标准化系数		标准系数	t	显著性
	B	标准错误	贝塔		
1 (常量)	-3.189E-18	.126		.000	1.000
出汁率	-.757	.128	-.757	-5.915	.000
2 (常量)	-1.340E-11	.106		.000	1.000
出汁率	-.690	.109	-.690	-6.303	.000
百粒质量	-.375	.109	-.375	-3.428	.002

a. 因变量: W 色泽 B

模型摘要

模型	R	R 平方	调整后的 R 平方	标准估算的错误
1	.459 <sup>a</sup>	.211	.180	.90534
2	.610 <sup>b</sup>	.372	.322	.82326

a. 预测变量: (常量), 百粒质量

b. 预测变量: (常量), 百粒质量, 苹果酸

系数<sup>a</sup>

模型	非标准化系数		标准系数	t	显著性
	B	标准错误	贝塔		
1 (常量)	1.639E-11	.171		.000	1.000
百粒质量	.459	.174	.459	2.635	.014
2 (常量)	1.713E-11	.156		.000	1.000
百粒质量	.480	.159	.480	3.023	.006
苹果酸	.403	.159	.403	2.538	.018

a. 因变量: W 色泽 A

模型摘要

模型	R	R 平方	调整后的 R 平方	标准估算的错误
1	.721 <sup>a</sup>	.519	.501	.70651

2	.814 <sup>b</sup>	.663	.636	.60341
---	-------------------	------	------	--------

- a. 预测变量: (常量), 干物质含量
- b. 预测变量: (常量), 干物质含量, 出汁率

#### 系数<sup>a</sup>

模型		非标准化系数		标准系数		显著性
		B	标准错误	贝塔	t	
1	(常量)	-7.143E-11	.134		.000	1.000
	干物质含量	-.721	.136	-.721	-5.300	.000
2	(常量)	-7.143E-11	.114		.000	1.000
	干物质含量	-.502	.134	-.502	-3.742	.001
	出汁率	.438	.134	.438	3.262	.003

- a. 因变量: W 色泽 L

#### 模型摘要

模型	R	R 平方	调整后的 R 平方	标准估算的错误
1	.429 <sup>a</sup>	.184	.153	.92031
2	.560 <sup>b</sup>	.314	.259	.86088
3	.687 <sup>c</sup>	.472	.406	.77078

- a. 预测变量: (常量), 葡萄总黄酮
- b. 预测变量: (常量), 葡萄总黄酮, 总糖
- c. 预测变量: (常量), 葡萄总黄酮, 总糖, VC 含量

#### 系数<sup>a</sup>

模型		非标准化系数		标准系数		显著性
		B	标准错误	贝塔	t	
1	(常量)	-4.076E-11	.174		.000	1.000
	葡萄总黄酮	.429	.177	.429	2.425	.023
2	(常量)	-3.430E-11	.163		.000	1.000
	葡萄总黄酮	.520	.171	.520	3.043	.005
	总糖	.371	.171	.371	2.171	.040
3	(常量)	-4.302E-11	.146		.000	1.000
	葡萄总黄酮	.605	.156	.605	3.873	.001

总糖	.474	.158	.474	3.006	.006
VC 含量	.414	.155	.414	2.681	.013

a. 因变量：WDPPH 半抑制体积

#### 模型摘要

模型	R	R 平方	调整后的 R 平方	标准估算的错误
1	1.000 <sup>a</sup>	1.000	.	.

a. 预测变量: (常量), 果皮颜色 B, 白藜芦醇, PH 值, 柠檬酸, 单宁, 果梗比, 褐变度, 还原糖, 苹果酸, 果皮质量, 花色苷, DPPH 自由基, 果穗质量, 多酚氧化酶活力, 总糖, 固酸比, 黄酮醇, VC 含量, 出汁率, 氨基酸总量, 果皮颜色 A, 百粒质量, 干物质含量, 酒石酸, 总酚, 蛋白质, 葡萄总黄酮

#### 系数<sup>a</sup>

模型		非标准化系数		标准系数	t	显著性
		B	标准错误	贝塔		
1	(常量)	-2.530E-10	.000	.	.	.
	氨基酸总量	2.105	.000	.2105	.	.
	蛋白质	-1.312	.000	-.1312	.	.
	VC 含量	-2.715	.000	-.2715	.	.
	花色苷	.115	.000	.115	.	.
	酒石酸	-.284	.000	-.284	.	.
	苹果酸	-2.000	.000	-.2000	.	.
	柠檬酸	-.783	.000	-.783	.	.
	多酚氧化酶活力	-1.040	.000	-.1040	.	.
	褐变度	1.188	.000	.1188	.	.
	DPPH 自由基	-.181	.000	-.181	.	.
	总酚	2.075	.000	.2075	.	.
	单宁	.421	.000	.421	.	.
	葡萄总黄酮	-3.247	.000	-.3247	.	.
	白藜芦醇	-.909	.000	-.909	.	.
	黄酮醇	.730	.000	.730	.	.
	总糖	-1.309	.000	-.1309	.	.
	还原糖	-2.533	.000	-.2533	.	.
	PH 值	-1.273	.000	-.1273	.	.
	固酸比	-2.308	.000	-.2308	.	.

干物质含量	.501	.000	.501	.	.
果穗质量	-.435	.000	-.435	.	.
百粒质量	-2.625	.000	-2.625	.	.
果梗比	-.732	.000	-.732	.	.
出汁率	-.492	.000	-.492	.	.
果皮质量	1.035	.000	1.035	.	.
果皮颜色 A	-.213	.000	-.213	.	.
果皮颜色 B	-1.490	.000	-1.490	.	.

a. 因变量: W 白藜芦醇