

葡萄酒综合评价

王文涛¹ 江兵兵¹ 李阳¹ 陈六新²

(1.重庆邮电大学计算机学院 重庆 400065 ; 2.重庆邮电大学数理学院 重庆 400065)

摘要:本文旨在围绕葡萄理化指标、葡萄酒理化指标、以及葡萄酒质量三者进行研究。运用秩检验、Cronbach 系数、系统聚类分析、多元线性回归和灰色关联度等理论,解决判断评酒员评分可信度、葡萄酒评价与葡萄等级分类、研究葡萄与葡萄酒理化指标间联系等问题。实践证明,本文提出的模型具有较高的实用性,可为解决相关问题奠定理论基础。

关键词:葡萄酒 秩检验

中图分类号:S663.1

文献标识码:A

文章编号:1674-2060(2012)07-0026-04

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,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。本文以2012年高教社杯全国大学生数学建模竞赛A题中提到的葡萄酒评价为例,选取题目所给附录中红葡萄做为研

表1 评分差异性显著性

项目	总分	外观得分	香味得分	口感得分	整体得分
显著性概率 P	0.0151	0.0261	0.0165	0.1366	0.4926

表2 主成分分析结果

主成分	特征值	贡献量	累计贡献量	主成分	特征值	贡献量	累计贡献量
1	4.693	0.2761	0.2761	2	3.5033	0.2061	0.4821
3	1.729	0.1017	0.5838	4	1.5127	0.089	0.6728
5	1.1096	0.0653	0.7381	6	0.8684	0.0511	0.7892

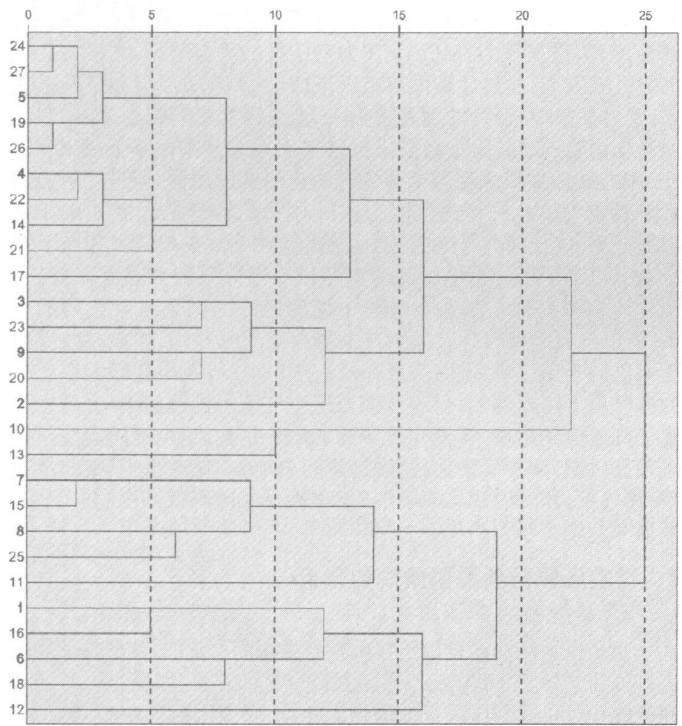


图1 红葡萄分级聚类图

究对象,先后研究分析了评酒员评分的可信度、葡萄的等级分类,葡萄与葡萄酒理化指标间的联系,酿酒葡萄理化指标和葡萄酒理化指标对葡萄酒质量的影响。

1、模型假设

- (1)所得葡萄理化数据详实可靠;
- (2)评委打分相互独立且公正;
- (3)可信度高的一组评酒人员对葡萄酒的评价结果可视为葡萄酒质量;
- (4)不考虑酿造工艺、酿造环境等外部因素的影响;

2、符号约定

I_{red}^j :第*i*组对第*j*种红葡萄酒的评分($i=1,2;j=1,2,L,27$);
 P :品酒人员对葡萄酒评分的一致显著性概率;

表3 27种红葡萄理化指标综合评价结果

种类	1	2	3	4	5	6	7	8	9
得分	-0.25	0.51	0.34	-0.45	0.31	-0.39	-0.31	0.33	0.28
种类	10	11	12	13	14	15	16	17	18
得分	0.15	-0.63	-0.29	0.12	-0.43	0.33	-0.58	0.2	-0.46
种类	19	20	21	22	23	24	25	26	27
得分	0.18	0.4	-0.17	-0.23	0.57	0.3	0.19	0.07	-0.07

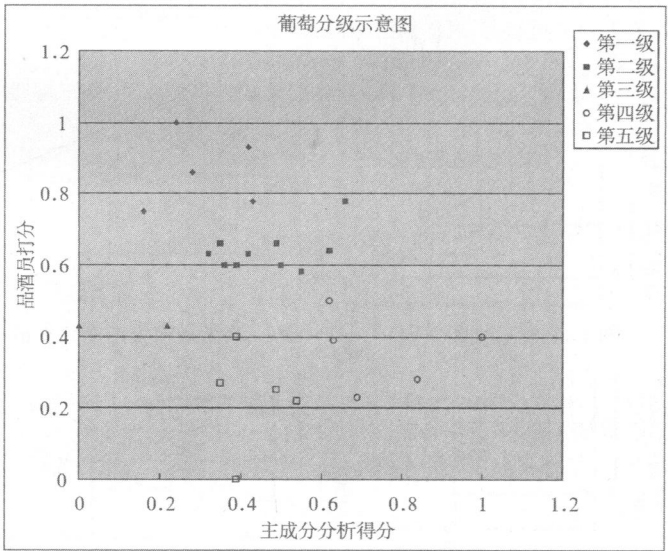


图2 红葡萄分级图

作者简介:王文涛(1992-),男,汉,重庆邮电大学,本科学士,研究方向:信息安全;
江兵兵(1991-),男,汉,重庆邮电大学,本科学士,研究方向:信息安全;
李阳(1990-),男,汉,重庆邮电大学,本科学士,研究方向:网络工程。

α :品酒人员对葡萄酒评分的内部一致性系数;

y_{gi} :对酿酒葡萄理化指标的主成分分析中的第*i*个主成分($i=1,2,L,6$);

G_{gi} :酿酒葡萄理化指标的主成分分析中的评价指标;

M :酿酒葡萄理化指标的主成分分析综合评价得分;

a_i :酿酒葡萄的理化指标($i=1,2,L,20$);

b_j :葡萄酒中的理化指标($j=1,2,L,9$);

$x_i(i)$:葡萄酒或酿酒葡萄主要理化指标标准化数据($i=1,2,L,27; j=1,2,L, k$);

3、模型的建立及求解

3.1 对不同评价结果的差异与可信度分析

3.1.1 两配对样本秩检验验证差异显著性

要检验两个样本是否存在差异,采用Wilcoxon提出的秩检验方法^[1]。检验的基本步骤为:

(1)建立假设: H_0 :两组样本不存在显著差异; H_1 :两组样本存在显著差异。

(2)编排秩次,求秩和设两个样本的数据量分别为 n_1 和 n_2 ,再令 R_1 代表样本1的秩, R_2 代表样本2的秩,则:

$$Z = \frac{\frac{R_1 - R_2}{n_1 - n_2}}{\sqrt{\frac{(n_1 + n_2)^2 (n_1 + n_2 + 1)}{12n_1n_2}}} \quad (1)$$

(3)做出推断:求出Z值后可判断两个样本的均值差的显著性。

首先计算出第一组对红葡萄酒的打分($I_{\text{g1}}^{(1)}, I_{\text{g1}}^{(2)} L I_{\text{g1}}^{(27)}$)、第二组对红葡萄酒的打分($I_{\text{g2}}^{(1)}, I_{\text{g2}}^{(2)} L I_{\text{g2}}^{(27)}$),为全面分析两组评酒员的评价结果有无显著性差异,对评酒员所给关于外观、香气、口感、整体评价四个指标的评分结果做差异显著性检验,在显著性概率 <0.05 范围内可认为两样本存在显著差异。题目中另给出四个指标的相加之后的总分,由数据计算得总分显著性与四个评分项目显著性见表1。

由表1可知每组评价结果中总分、外观、香味存在显著性差异,口感、整体没有显著性差异,可以认为两组评酒评分存在显著性差异。

3.1.2 评委评分可信分析模型

为对样本的可信度进行分析,采用Cronbach提出的 α 系数^[2], α 系数评价各项得分间的一致性,属于内在一致性系数。相关系数愈高,相关性愈高,即内部一致性愈高。

$$\alpha = \frac{K}{K-1} \times \frac{1 - \sum D(x_i)}{D(x)} \quad (2)$$

上式中 K 表示被测样本总数, $D(x)$ 表示评分总分的方差, $\sum D(x_i)$ 表示第*i*个评委对酒打分的方差,计算出数值判断结果是否可信。一般认为 α 系数大于0.7属于很可信,大于0.9属于十分可信。

红葡萄酒被测样本总数 $K=27$,计算出评分总分的方差,被测试酒第*i*个评委给分的方差,然后综合计算每组10个评委的可信度,利用Matlab计算得到第一组评委对红葡萄酒的评分结果可靠性 $\alpha_{\text{g1}}=0.862$,属于很可信, $\alpha_{\text{g2}}=0.935$,属于十分可信,因此,有理由认为在对红葡萄酒的评分中,第二组评委评分可信度明显优于第一组。

3.2 酿酒葡萄分级模型

3.2.1 酿酒葡萄分级指标选取

葡萄酒的质量与葡萄的质量密切相关,我们利用第二组评酒员的评价结果,分析葡萄酒的质量与葡萄理化指标的联系,从而将葡萄进行分级。

使用Matlab7.0将数据进行标准化,由此构成一组无量纲序列。

主成分分析试图在力保数据信息丢失最少的原则下,通过研究指标体系的内在结构联系,把多指标转化成少数几个相互独立且包含原有指标大部分信息(70%—80%以上)综合指标的多元统计方法,即对高维变量空间进行降维处理。

根据查阅的相关研究论文^{[4][5][6][7]},主要指标包含资料中的所有17项葡萄酒理化指标,对17个评价指标进行主成分分析,相关系数矩阵的前6个特征值及其贡献率如表2。

可以看出前6个特征值的累计贡献率为78.92%,主成分分析效果良好。下面选取前6个主成分进行综合评价。参考6个特征根对应的特征向量表,可得六个主成分值计算公式:

$$\begin{aligned} y_{\text{g1}} &= -0.0333\%a + 0.0661\%L + 0.0566\% \\ y_{\text{g2}} &= 0.2724\%a - 0.5515\%L + 0.1283\% \\ y_{\text{g3}} &= -0.2723\%a + 0.5966\%L + 0.2366\% \\ y_{\text{g4}} &= 0.1088\%a - 0.1928\%L + 0.3033\% \\ y_{\text{g5}} &= 0.0624\%a - 0.0270\%L - 0.3146\% \\ y_{\text{g6}} &= 0.0402\%a + 0.2031\%L + 0.1533\% \end{aligned}$$

根据6个主成分的表达式,将27种红葡萄原始的17个指标的标准化数据带入其中。就可以得到红葡萄理化指标的六个主成分值。分别以六个主成分的贡献率为权重,构建红葡萄理化指标综合评价模型:

$$G_{\text{gi}} = 0.2761y_{\text{g1}} + 0.2061y_{\text{g2}} + 0.1017y_{\text{g3}} + 0.0890y_{\text{g4}} + 0.0653y_{\text{g5}} + 0.0511y_{\text{g6}} \quad (3)$$

把27种红葡萄的6个主成分值带入上式,可以得到它们理化指标的综合评价(见表3)。

通过主成分综合分析计算得出的红葡萄理化指标的综合得分 $G_{\text{gi}}^{(i)}(i=1,2L,27)$ 。在解决问题一时,已证明第二组的评酒员的评分更合理,故以第二组评酒员的评分作为葡萄酒的质量,得红葡萄酒质量数据 $I_{\text{g2}}^{(i)}(i=1,2L,27)$ 。

3.2.2 采用系统聚类法将酿酒葡萄分级

系统聚类法^[3]是聚类分析方法中最常用的一种方法。它的优点在于可以指出由粗到细的多种分类情况,典型的系统聚类结果可由一个聚类图展示出来。系统归类过程与计算类和类之间的距离有关,采用不同的距离定义,有可能得出不同的聚类结果。利用最短距离法对酿酒葡萄做聚类分析,从而将葡萄分类。

步骤如下:设 $\Omega = \{w_1, w_2, L, w_n\}$

(1)计算*n*个本点两两之间的距离 $\{l_{ij}\}$,记为矩阵 $L = (l_{ij})_{n \times n}$;

(2)首先构造*n*个类,每一个类中只包含一个样本点,每一类的平台高度均为零;

(3)合并距离最近的两类为新类,并且以这两类间的距离值作为聚类图中的平台高度;

(4)计算新类与当前各类的距离,若类个数已等于1,转步骤5,否则,转步骤3;

(5)画聚类图;

(6)决定类的个数和类。

使用最短距离法来测量类与类之间的欧式距离,即称其为系统聚类法中的最短距离法(又称最近邻法),最先由Florek等人1951年和Sneath1957年引入。

首先,对红葡萄进行聚类由附录数据2可得,27种红葡萄对应 $\Omega = \{w_1, w_2, L, w_{27}\}$,葡萄品质为($G_{\text{gi}}^{(i)}, I_{\text{g2}}^{(i)}$),使用绝对值距离来测量点与点之间的距离,使用最短距离法来测量类与类之间的距离,即:

$$l(w_i, w_j) = \sum_{k=1}^2 |G_{\text{gi}}^{(ik)} - I_{\text{g2}}^{(jk)}| \quad (4)$$

$$L(G_p, G_q) = \min_{\substack{w_i \in G_p \\ w_j \in G_q}} \{l(w_i, w_j)\} \quad (5)$$

由距离公式利用SPSS软件计算出距离矩阵并绘制出聚类图如下:

由图1分析,计算取高度 $h=17$ 分级,将红葡萄分类所得结果做

定性分析。以主成分分析得分做 x 轴,葡萄酒质量,即第二组评委打分做 y 轴,将数据标准化处理后,用EXCEL绘制红葡萄分类(见图2)。

葡萄的分类主要依据酿酒葡萄主要理化指标得分与评酒员对该种葡萄所酿的葡萄酒的打分,由上图可直观看出,分类结果还是较为满意的。结合图1与图2作出红葡萄的分级表,第一级为最好,以此类推,见表4。

3.3 分析酿酒葡萄与葡萄酒理化指标间的联系

某些理化指标例如:单宁,总酚,花色苷等在葡萄中存在,在葡萄酒中也存在,其中必然存在一定的联系。数据有酿造前葡萄指标的数据和酿造后葡萄酒指标的数据,基于此种问题,适合于灰色关联度分析。灰色关联度分析是基于系统内参考因素和比较因素之间的关联度大小对系统行为特征进行量化分析,指在系统发展过程中,如果两个因素变化的态势是一致的,即同步变化程度较高,则可以认为两者关联较大,灰色关联度分析对于一个系统发展变化态势提供了量化的度量,适合动态的历程分析。步骤如下:

(1)选取参考数列和比较数列。

将酿酒葡萄理化指标做自变量,假设各指标对应参数 $a_i (i=1,2,L,20)$,葡萄酒理化指标做因变量,设各指标对应参数 $b_j (j=1,2,L,9)$ 。分别计算每个葡萄酒理化指标和20个酿酒葡萄理化指标间的灰色关联度。

$$\xi_i(k) = \frac{\min_s \min_t |b_i(t) - a_s(t)| + \rho \max_s \max_t |b_i(t) - a_s(t)|}{|b_i(k) - a_i(k)| + \rho \max_s \max_t |b_i(t) - a_s(t)|} \quad (6)$$

上式(6)为比较数列 x_i 对参考数列 b_i 在 k 时刻的关联系数,其中 $\rho \in [0,1]$ 为分辨系数。称上式中 $\min_s \min_t |b_i(t) - a_s(t)|$ 项, $\max_s \max_t |b_i(t) - a_s(t)|$ 项分别为两级最小差及两级最大差。

(2)参考数列和比较数列初始值的无量纲处理。为保证结果的正确性,原始数据必须消除量纲,此处采用均值化变化处理。

(3)关联度计算 $r_i = \frac{1}{n} \sum_{k=1}^n \xi_i(k)$ 为数列 a_i 对参考数列 b_i 的关联度。

关联度是把各个时刻的关联系数集中为一个平均值,亦即把过于分散的信息集中处理。

按上述步骤,计算出第1个红葡萄酒理化指标中的花色苷与20个酿酒葡萄理化指标间的灰色关联度如表5。

由表5可以看出,红葡萄酒中的花色苷与酿酒葡萄中的花色苷关联度是最大的,此外,花色苷与总黄酮和黄酮醇的关联度也很大。

表4 红葡萄分级表

组别	红葡萄的种类
一级	3, 23, 9, 20, 2
二级	24, 27, 5, 19, 26, 4, 22, 14, 21, 17
三级	10, 13
四级	1, 16, 6, 18, 12
五级	7, 15, 8, 25, 11

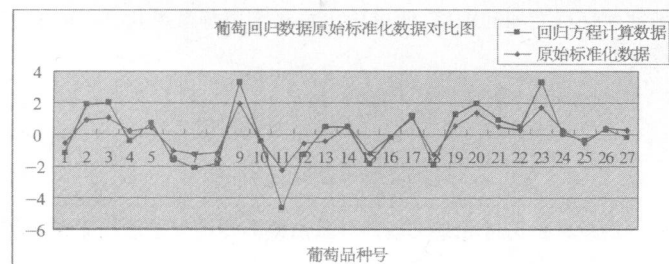


图3 红葡萄酒理化指标对酒质量对比图

同理可计算出9个红葡萄酒理化指标与20个红葡萄酒理化指标的关联度。

查阅的相关研究论文^{[4][5][6]},花色苷属于类黄酮,可由总黄酮与黄酮醇转化而来,所得结果比较符合实际,由此可知酿酒葡萄理化指标与葡萄酒理化指标之间存在联系:两者共有的指标间灰色关联度最大,例如酒中存在的花色苷与葡萄中存在的花色苷;与其相近的指标,即可以转化的指标间关联也很明显。

3.4 酿酒葡萄和葡萄酒理化指标与葡萄酒质量间的关系模型

由3.3可知,间接因素指标与直接因素指标有些是重叠的,例如:花色苷,葡萄总黄酮等。分开讨论葡萄酒理化指标对葡萄酒质量的影响,葡萄酒理化指标对葡萄酒质量的影响,由此分析,建立多元线性回归模型。

首先将原始数据消除量纲,选红葡萄酒中包括芳香物质总量在内的22个指标做非参数相关性分析,通过SPSS求解得到到红葡萄酒主要指标与红葡萄酒质量相关性指数较大的前9种物质:花色苷、单宁、总酚、葡萄总黄酮、PH值、果皮颜色a、DPPH、黄酮醇、固酸比。

由资料^{[4][5][7]}可知,葡萄酒芳香类物质影响葡萄酒质量中的香味,葡萄中有73种芳香物,香味主要来源于丁酸乙酯、己醇、苯乙醇。结合非参数相关分析确定影响葡萄酒质量的葡萄酒理化指标有10种:花色苷、单宁、总酚、酒总黄酮、白藜芦醇、DHHP、色泽L、丁酸乙酯、己醇、苯乙醇。

建立的红葡萄酒质量与葡萄酒理化指标间的线性函数:

$I_{\text{红2}} = \sum_{i=1}^n \beta_i x_i + \varphi$, 式中 $I_{\text{红2}}$ 为红葡萄酒质量, $x_i (i=1,2,L,n)$ 是通过SPSS计算相关性分析选出的影响葡萄酒质量的 n 个主要指标, $\beta_i (i=1,2,L,n)$ 分别为 n 个指标前的系数, φ 为常数。代入数据得:

(1)红葡萄酒理化指标与红葡萄酒质量之间的关系模型:

$$I_{\text{红2}} = -0.357x_1^{(1)} + L + 0.187x_9^{(1)} - 6.308 \times 10^{-15} \quad (7)$$

检验: $R^2 = 0.810$ 表明回归方程拟合程度较好; $F = 8.030 > F_{1-\alpha}(9,17) = 4.30$, 表明回归方程显著; $P = 0.00014 = 0.05$, 回归模型成立。

(2)红葡萄酒理化指标与红葡萄酒质量关系模型:

$$I_{\text{红2}} = -0.787x_1^{(2)} + L - 0.207x_{10}^{(2)} - 8.11 \times 10^{-15} \quad (8)$$

检验: $R^2 = 0.715$ 表明回归方程拟合程度较好; $F = 4.023 > F_{1-\alpha}(10,16) = 4.0$, 表明回归方程显著性较好; $P = 0.007 = 0.05$, 回归模型成立。

表5 酿酒葡萄理化指标间的灰色关联度

理化指标	花色苷	总黄酮	黄酮醇	褐变度	DPPH 自由基	白藜芦醇	总酚
关联度	0.9112	0.8426	0.8052	0.7959	0.7828	0.7753	0.7736
理化指标	果皮颜色-b	氨基酸	果梗比	多酚氧化酶	果皮颜色-a	单宁	固酸比
关联度	0.7714	0.748	0.7396	0.7385	0.7341	0.7236	0.7047
理化指标	可溶性固体	VC 含量	总糖	可滴定酸	果皮颜色-L	PH 值	
关联度	0.7015	0.6938	0.6478	0.6301	0.6282	0.585	

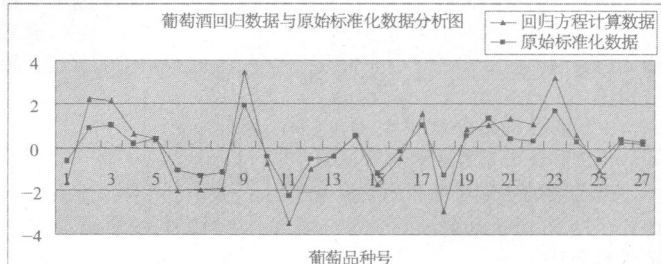


图4 红葡萄酒理化指标对酒质量对比图

从上述可以看出,多元回归模型通过 F 检验与 P 检验,有理由相信葡萄与葡萄酒的理化指标与葡萄酒质量之间存在联系,但可能不是线性关系。建立了多元线性回归模型解释葡萄理化指标与葡萄酒质量间的关系,可将红葡萄每个指标数据代入回归方程,计算出回归方程反应的葡萄酒质量,结合原始葡萄酒质量数据,即品酒员的打分,画出对比图。

从上图3可以看出,回归模型可大体解释红葡萄理化指标与红葡萄酒质量间的关系。从上图4可以得出与图3相同的结论,即回归模型可大体解释红葡萄酒理化指标与红葡萄酒质量间的关系,取多元线性回归方程相关系数作为影响程度,得红葡萄与红葡萄酒质量相关程度为81%,红葡萄酒理化指标和红葡萄酒质量相关程度为71%。

接下来,对其联系做定性分析。利用灰色关联度理论,计算得出葡萄的理化指标与葡萄酒质量间的灰色关联度在0.95以上的指标有:PH值、果皮颜色L、可溶性固体、总糖、还原性糖、固酸比、可滴定酸、DPPH、果梗比、总酚、葡萄总黄酮、多酚氧化酶、氨基酸总量、单宁、果皮颜色a。

葡萄酒理化指标与葡萄酒质量间的灰色关联度在0.75以上指标有:色泽a、丁酸乙酯、色泽b、总酚、单宁、己酸、苯乙酸、酒总黄酮、色泽L、DPPH、花色苷。

从定性角度来看,灰色关联度计算得到的葡萄酒质量影响较大的指标和利用SPSS软件做相关性分析选取的指标大致相同,酿酒葡萄理化指标和葡萄酒对葡萄酒质量都存在影响。

4、结语

.....上接第25页

压由 P_2 降为 P_1 后,空压机的耗电量随排压降低而降低,而排气量却是增加的,如图2。此时既达到了节能的目的,增加的空气量还有利于提高氧气产量。所以工作点A的经济性比工作点B好,所以在压缩机运行中要尽量使工作点趋近A点。

3.2 管网特性曲线

管网特性曲线如图3所示。由于气体在管网中的流动阻力与流量的平方成正比,即: $P=AQ^2$; 式中: P 为管网的阻力损失; Q 为管网的气体流量; A 为管网的阻力系数。图3中的二次曲线代表了气体在管网内流过时,阻力损失与流量之间的关系。当气体流量为 Q 时,管网的阻力损失为 P ,当用户的工作压力(如精馏下塔压力)为 P_r 时,则压缩机的排压达到 P_c 才能满足要求。在 P_r 不变的情况下, P 越低,则空压机的排压越低。

3.3 压缩机实际工作点的确定

压缩机在实际运行中,其工作点既会受压缩机性能曲线的影响,又会受到管网特性曲线的影响。所以其工作点是压缩机性能曲线与管网特性曲线的交点(如图4)。

从压缩机工作曲线图可以看到,同一压缩机的性能曲线是恒定的。但是,当供气管网阻力发生变化时,压缩机的实际工作点将会沿着性能曲线进行移动。

4、空气系统阻力分析及改进措施

公司在用空压机的出口压力的高低取决于使用岗位工艺要求,

对葡萄酒做综合评价,从专家打分入手,确定葡萄酒质量,结合葡萄主要理化指标得分做聚类分析,对葡萄做等级分类,灰色关联度理论适用于分析葡萄与葡萄酒理化指标间内在的复杂关联,采用多元回归模型结合灰色关联度理论的方法从定量与定性角度确定酿酒葡萄和葡萄酒理化指标对葡萄酒质量影响。

值得一提的是题目中葡萄酒理化指标未给出蛋白质、氨基酸、多糖等指标,数据缺乏导致在分析红葡萄酒理化指标对酒质量影响时回归模型效果一般。

参考文献

- [1]盛骤等.概率论与数理统计[M].北京:高等教育出版社,2012.
- [2]何佳,何惧等.评分者信度的分析方法简介及比较[J].医学教育.2007,45(6):76-77.
- [3]孙鹏,张小松等.不同产地青蒿 FTIR 特征图谱的系统聚类分析和主成分分析[J].药物分析杂志.2008(10):51-54.
- [4]李华,刘曙东,王华,张予林.葡萄酒感官评价结果的统计分析方法研究[J].中国食品年报,2006,6(2):126-131
- [5]宋于洋,王炳举等.石河子地区酿酒葡萄成熟度与葡萄酒质量的研究[J].西北农业学报.2006,15(3):153-156.
- [6]于贞,姜爱莉.主成分分析法研究酚类物质对葡萄酒品质的影响[J].酿酒科技,2011,204(6):107-109.
- [7]李华,刘勇强,郭安鹏等.运用多元统计分析确定葡萄感官特性的描述符[J].中国食品学报,2007,7(4):114-118.
- [8]谢辉,樊丁宇等.统计方法在葡萄理化指标简化中的应用[J].新疆农业科学,2011,48(8):1435-1437.

而供气管网阻力在设计中虽然有一个具体值,但由于后期运行过程中的管线改造、使用工序的增加供气系统的配置调整等诸多因素影响,造成供气系统实际管偏离设计值。因此,降低空压机运行能耗还需从减小供气系统管阻着手。根据工作经验,简要总结以下几点降低空气系统阻力的措施:(1)将列管式后冷器更换为翅片式后冷器,可降低阻力约10KPa;(2)选用高效过滤器,操作上定期除灰,减小过滤器的阻力;(3)加强空压机冷却系统管理,提升员工冷却水系统操作水平,保证各级冷却系统及被冷却介质达到预期效果。

5、结语

文章论述了空气压缩机运行过程中影响工作效率的影响因素,为降低生产企业空压机运行能耗提供了参考依据。空压机低成本运行方面仍有空间,广大技术工作者仍需不断努力,不断摸索空气压缩机最佳经济运行参数,全面实现空压机的低能耗运行。

参考文献

- [1]徐忠.离心式压缩机原理[Z].西安交通大学,2000.
- [2]李云,姜培正.过程流体机械(第二版)[Z].北京,化学工业出版社,2008-09.