

文章编号: 1003-8701(2008)04-0049-04

主成分权重法在番茄果实商品性综合评价上的应用探讨

陈 贤, 吴兴恩, 杨 德, 龚元圣 *

(云南农业大学, 昆明 650201)

摘 要: 本试验以主成分权重法对 7 个番茄品系的果实商品性状进行了综合评价分析, 构建了综合评价函数。结果表明: 应用主成分法提取了 2 个主成分, 累计方差贡献率达 86.59%, 变量的权重偏重于营养品质和外观品质, 合乎评价的要求。将 7 个品系的商业性分为 3 个层次, 品系间在营养品质和耐贮品质的得分上分化较大, 外观品质上分化较小, 利于品系评选。第 1 层次是 YH02-6 和 YH02-9, 第 2 层次是 YH02-1 和 YH02-7, 第 3 层次 YH02-11、YH02-8 和 YH02-5。此方法把 AHP 法和主成分法结合起来, 考虑了变量测度的转换, 使评价客观、分数直观、层次分明, 对番茄果实商品性的综合评价有较好的适用性。

关键词: AHP; 主成分分析; 测度转换; 综合评价

中图分类号: S641.2

文献标识码: A

Discussion on Comprehensive Evaluation of Commercial Traits of Tomato Fruits Using the Method of Weighting of Principal Component

CHEN Xian, WU Xing-en, YANG De, GONG Yuan-sheng*

(Yunnan Agricultural University, Kunming 650201, China)

Abstract: Commercial traits of 7 tomato cultivars were comprehensively evaluated using the method of weighting of principal component in the experiment. The comprehensive evaluation function was built up. The results showed that with the factor analyzed, 2 principal components were abstracted and their cumulative rate of variance was 86.59%. The weights of the nourishment quality and the external appearance quality were higher, which was met up to the demand of the comprehensive evaluation. According to the scores of commercial traits, the cultivars varied considerably in the nourishment quality traits and the preservative quality traits, and less in the external appearance quality traits. This was helpful to pick out good one. The cultivars were divided into three classes, of which 'YH02-6' and 'YH02-9' were in the first class, 'YH02-1' and 'YH02-7' were in the second class, and 'YH02-11', 'YH02-8' and 'YH02-5' were in the lowest class. This approved that this analysis method was a practical and efficient way to analyze the commercial traits of the fruits of tomato comprehensively. It integrated the methods of AHP and the principal component analysis to make the evaluating job easy, objective, and obvious.

Key words: AHP; Principal component analysis; Transforming of variable; The comprehensive evaluation

收稿日期: 2008-03-25

基金项目: 云南省自然科学基金项目 2001-C0035M)

作者简介: 陈 贤 (1972-), 男, 硕士, 讲师, 主要从事园林园艺植物遗传育种、苗木生产、试验统计分析等方面的教学与研究。

通讯作者: 龚元圣, 硕士, 讲师, E-Mail: yuansheng789@vip.sina.com

番茄是我国人们喜爱的蔬菜之一, 在追求产量的同时, 对番茄品质提出了更多更高的消费要求, 涵盖果实外观、风味和营养品质的商品性状的综合评价, 引起了育种、栽培、贮藏和销售流通等环节专业人士的关注。

主成分分析法应用于综合评价是对综合评价

体系涵盖的多变量通过无量纲标准化处理,将其原相关的多个随机变量,以方差贡献率作为信息的测度标准,降维为不相关的几个新变量(主成分),构建评价函数,对参评的项目进行综合得分的评价排序,属于综合评价方法中的客观赋权法^[1-3]。在实际应用中,评价函数的构建通常是分步进行的,变量的线性组合形成主成分,主成分的线性组合形成综合评价得分,对参评项目的综合评价得分的分析也是分为变量对主成分影响和主成分对综合得分的影响两个环节进行的,有时对变量在同一个主成分线性组合的解释是牵强的,每个变量参与多次计算,变量值必须标准化转换才可代入评价函数分析,计算繁琐,问题变得复杂,且综合得分有负数,评价成绩不直观,这些都影响着主成分综合评价法的推广应用。

本研究尝试借鉴 AHP 法 (Analytic Hierarchy Process) 的分析思路,应用主成分法计算番茄的商品性状各变量的权重,按 AHP 法的思路进行综合评价,即,变量值在分析前经过 3 种测度的转换,均转化为正向变量,然后进行以权重值为系数的线性评价函数的分析,不需对变量进行标准化转换,也不需因子分析的正交变换和主成分意义的解释,可以大大简化思路,为生产者,销售者在鲜果产品选择时提供客观、方便、快捷的分析方法。

1 材料与方法

1.1 试验材料

供试材料 7 个樱桃番茄育种品系: YH02-1、YH02-5、YH02-6、YH02-7、YH02-8、YH02-9 和 YH02-11,果实形态特征分别为红色近圆形果、黄色梨形果、红色圆形果、红色圆形果、红色梨形果、红色李形果和黄色圆形果。

1.2 试验方法

2005 年在云南农大园林园艺农场,按单因素随机区组试验设计进行试验^[4],处理数(品系数) $t=7$,区组数 $b=3$,每个小区 24 株,依据土壤差异的垂直方向划分区组,保护地栽培。

在果实完熟期(红熟期)采收,以品系的小区为品质测定分析单元,测定能反映番茄果实商品性的性状,包括外观品质(单果重和果色)、营养品质^[5](维生素 C、可溶性固形物、糖酸比和水分含量)和常温下贮藏 21 d 的变质率,其中以单果重指标表示果实的大小,这是因为品系中存在不同的果形,用果径指标不能客观衡量果实的大小。

1.3 分析方法

1.3.1 评价指标变量的测度转换

采用比值法的 3 种测度^[6]进行转换,使所有变量均转换为正向变量(越高越好),定量数据直接转换,定性数据先进行分级评分,再进行转换。

对于正向性求高指标的定量数据(维生素 C、单果重、果色得分),采用上限测度:

$$y_{ij}(\%) = x_{ij} / x_{\max} \times 100, (x_{ij} \leq x_{\max})$$

对于负向性求低指标的定量数据(果实水分含量、变质率),采用下限测度:

$$y_{ij}(\%) = x_{\min} / x_{ij} \times 100, (x_{ij} \geq x_{\min})$$

对于要求适中的定量数据(可溶性固形物、糖酸比),采用适中性测度:

$$y_{ij}(\%) = x_{ij} / (|x_{ij} - x_0| + x_0) \times 100$$

式中: i 为第 i 个品系, j 为第 j 个性状, y_{ij} 为第 i 个品系第 j 个性状指标的测度值, x_{ij} 为性状观察值, x_{\max} 为性状指标最大值, x_{\min} 为性状指标最小值, x_0 为性状指标最优的适中值。

番茄果实颜色根据大众消费者的喜爱(偏好红色番茄)分级评分进入分析,确定打分标准为:深红, 90; 浅红, 80; 黄色, 70; 杂色, 60。

1.3.2 主成分法求权重

对 7 个品系组成的 21 个样本(1 个小区当作 1 个样本,每个品系 3 个小区)在 7 个评价指标变量(X_1, X_2, \dots, X_7)的测度变量(Y_1, Y_2, \dots, Y_7)上的计算值构建 21×7 维矩阵($y_{1,1}, y_{1,2}, \dots, y_{21,7}$),应用 SPSS 软件进行主成分分析,根据方差贡献率 80% 的标准提取 n 个主成分,用各主成分的特征向量的绝对值计算各变量的权重系数。

$$j = \sum_{k=1}^n (|U_{kj}| \cdot C_k)$$

$$C_k = \frac{k}{\sum_{k=1}^n k}$$

$$U_{kj} = \frac{A_{kj}}{\sqrt{k}}$$

式中: j 为第 j 个变量的权系数; U_{kj} 为第 k 个主成分的特征向量在第 j 个变量的分量; C_k 为第 k 个主成分的方差相对贡献率; k 第 k 个主成分的特征根; n 为提取的主成分数; A_{kj} 为第 k 个主成分在第 j 个变量的载荷值。

1.3.3 权系数的归一(权重)

$$W_j = \frac{j}{\sum_{j=1}^p j}$$

式中: W_j 为第 j 个变量的权重, p 为评价指标体系

的变量数。

1.3.4 构建综合评价函数

$$F_i = \sum_{j=1}^p (W_j \cdot y_{ij})$$

式中: F_i 为第 i 个品系的综合得分, y_{ij} 为第 i 个品系第 j 个性状变量值的测度值。

2 结果与分析

2.1 原始变量的测度转换

对 21 个样本的 7 个指标的原始值进行测度转换, 其中 7 个品系的原始值的平均值见表 1, 其转换后的测度值见表 2。

表 1 7 个番茄品系果实各指标的平均值

品系	营养品质				外观品质		耐贮品质
	水分含量(%) X_1	维生素 C(mg/10 g) X_2	糖酸比(%) X_3	TSS(%) X_4	单果重(g) X_5	果色得分(%) X_6	变质率(%)Rate of X_7
1	92.04	18.92	8.58	4.47	19.32	90.00	90.00
5	94.03	16.28	10.33	4.50	10.22	70.00	82.50
6	89.08	31.83	14.82	7.40	13.48	80.00	60.00
7	92.45	21.12	8.15	6.23	14.94	90.00	97.50
8	93.64	13.05	7.62	4.47	12.62	80.00	77.50
9	91.00	24.20	14.16	6.83	8.75	81.00	55.00
11	92.56	16.57	11.06	4.62	11.22	70.00	80.00
最佳性状值	88.18	33.44	15.00	10.00	20.50	90.00	50.00
测度方法	下限测度	上限测度	适中性测度	适中性测度	上限测度	上限测度	下限测度

注: 表中品系代码 1, 5, ..., 11 表示品系 YH02-1, YH02-5, ..., YH02-11; TSS 表可溶性固形物; 最佳性状值来源于 21 个样本。

表 2 7 个番茄品系果实各指标变量的平均值的测度值(得分值)

品系	营养品质				外观品质		耐贮品质
	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
1	96	57	70	64	94	100	56
5	94	49	76	65	50	78	61
6	99	95	99	79	66	89	83
7	95	63	69	73	73	100	51
8	94	39	67	64	62	89	65
9	97	72	95	76	43	90	91
11	95	50	79	65	55	78	63
最佳性状值	100	100	100	100	100	100	100

注: 表中品系代码 1, 5, ..., 11 表示品系 YH02-1, YH02-5, ..., YH02-11; Y_1, Y_2, \dots, Y_7 表示水分含量、维生素 C、糖酸比、可溶性固形物、单果重、果色和变质率的测度变量。

2.2 主成分法求权重

2.2.1 主成分的提取

提取 2 个主成分, 累计方差贡献率为 86.591%, 已达到分析要求, 每个主成分在各变量上的特征向量的绝对值如表 3 所述, 两个主成分的方差相对贡献率计算如下:

主成分 1 的方差相对贡献率:

$$C_1 = \frac{4.052}{4.052+2.01} = 66.8\%$$

主成分 2 的方差相对贡献率:

$$C_2 = \frac{2.01}{4.052+2.01} = 33.2\%$$

表 3 主成分分析结果

主成分	特征根	累计方差 贡献率(%)	主成分特征向量的绝对值						
			Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
1	4.052	57.884	0.443	0.447	0.459	0.452	0.142	0.014	0.412
2	2.01	86.591	0.217	0.209	0.182	0.142	0.621	0.649	0.222

注: 变量 Y_1, Y_2, \dots, Y_7 分别表示水分含量、维生素 C、糖酸比、可溶性固形物、单果重、果色和变质率的测度变量。

2.3 变量权重的计算

对各变量来源于每个主成分的权系数计算,

表 4 变量的权系数及权重

参数	商品性状						
	营养品质			外观品质		耐贮品质	
	水分	维生素 C	糖酸比	TSS	单果重	果色	变质率
源于主成分 1 的权系数	0.296	0.299	0.306	0.302	0.095	0.010	0.275
源于主成分 2 的权系数	0.072	0.069	0.060	0.047	0.206	0.215	0.074
变量总权系数	0.368	0.368	0.367	0.349	0.300	0.225	0.349
变量总权重(总权系数归一)	0.158	0.158	0.158	0.150	0.129	0.097	0.150
同类性状累加权重			0.624			0.226	0.150

注: TSS 表可溶性固形物。

累加可得各变量的总权系数, 归一后可得各变量的权重, 如表 4 所述, 营养品质、外观品质和耐贮品质的权重为 0.624, 0.226, 0.150, 说明此评价体系偏重于营养品质, 兼顾外观品质和耐贮品质, 合乎客观实际。

2.4 计算各品系综合得分值

表 5 7 个番茄品系 7 个商品性状综合得分

品系	营养品质			外观品质			耐贮品质	总分
	水分	维生素 C	糖酸比	TSS	单果重	果色	变质率	
6	15.67	15.03	15.61	11.85	8.52	8.60	12.45	87.74
9	15.35	11.39	14.98	11.40	5.55	8.70	13.65	81.03
1	15.20	9.02	11.04	9.60	12.14	9.66	8.40	75.06
7	15.04	9.97	10.88	10.95	9.43	9.66	7.65	73.58
11	15.04	7.91	12.46	9.75	7.10	7.54	9.45	69.25
8	14.88	6.17	10.57	9.60	8.01	8.60	9.75	67.57
5	14.88	7.75	11.98	9.75	6.46	7.54	9.15	67.51
满分	15.80	15.80	15.80	15.00	12.90	9.70	15.00	100.00

注: 表中品系代码 1, 5, ..., 11 表示品系 YH02-1, YH02-5, ..., YH02-11; TSS 表可溶性固形物。

3 结论与讨论

本研究应用主成分求权重的方法构建了 3 种测度结合的综合评价番茄果实商品性状的模型, 指标变量的权重分布, 偏重于营养品质和风味, 合乎评价的客观实际, 将 7 个品系的商品性分为 3 个层次, 品系间在营养品质和耐贮品质的得分上分化较大, 外观品质上分化较小, 利于品系评选, 第 1 层次是 YH02-6 和 YH02-9, 营养品质和耐贮品质的分值较高; 第 2 层次是 YH02-1 和 YH02-7, 营养品质和外观品质得分较高, 但耐贮品质得分偏低, 第 3 层次是 YH02-11、YH02-8 和 YH02-5, 营养品质和耐贮品质得分和第 2 层次差距不大, 但外观品质得分较低, 这和主观的感官评价相吻合。

主成分分析的样本容量不能局限于待评价的样本数, 构建模型的样本数应大于待评价的样本数, 各样本类型的样本容量应尽量均匀, 模型的主成分载荷值才趋于稳定, 求出的权重系数才有很好的适用性^[7-11]。本研究中, 用 7 个品系的 21 个样本来构建模型, 评价时仅评价 7 个样本。

本研究验证了此方法简单易行, 集中了 AHP 法和主成分法的优点, 使评价得分客观公正, 分数间比较直观, 评比层次分明, 避开了主成分法模型构建的复杂计算和参数选择过程, 以及主成分含义的解释, 把分析重点集中到专业上的单变量(性状)、同类变量组合(性状类别), 以及综合得分上的多维比较, 利于科研、生产、流通环节的实际操作,

将 7 个番茄品系在各变量的平均值的测度值代入评价函数, 可得各品系的得分, 如表 5 所述, 此评价体系是百分制, 量程为 [0, 100], 将 7 个品系分为 3 个层次 (表 5): 第 1 层次是 YH02-6 和 YH02-9, 第 2 层次是 YH02-1 和 YH02-7, 第 3 层次是 YH02-11, YH02-8 和 YH02-5。

有较好的推广前景。

本研究仅从番茄的品质角度进行分析, 尚未考虑消费者的偏好等市场因素对果实商品性的影响, 可以进一步研究, 建立更加完善的体系。

此方法是从评价工作实际需要提出的, 其理论考证和模型的适用性需要进一步的研究, 对其他学科的综合评价领域有借鉴意义。

参考文献:

- [1] 余锦华, 杨维权. 多元统计分析与应用[M]. 广州: 中山大学出版社, 2005: 189-229.
- [2] 裴鑫德. 多元统计分析及其应用[M]. 北京: 北京农业大学出版社, 1991: 196-276.
- [3] 蔡 艺. 主成分方法在综合评价中的应用[J]. 中国统计, 2005(2): 25-26.
- [4] 杨 德. 试验设计与分析[M]. 北京: 中国农业出版社, 2002: 135-128.
- [5] 宁正祥. 食品成分分析手册[M]. 北京: 中国轻工业出版社, 1998.
- [6] 秦海英, 程献云, 梁中喜, 等. 杂交组合灰色评判在小麦育种中的作用[J]. 种子, 2002(2): 54-55.
- [7] 刘 坚, 李新萍, 温凤桐. 综合素质评价中权系数的确定方法[J]. 统计与决策, 2005, 25(8): 48-50.
- [8] 孙晓军, 周宗奎. 探索性因子分析及其在应用中存在的主要问题[J]. 心理科学, 2005, 28(6): 1440-1442.
- [9] 殷瑞飞, 朱建平. 关于利用因子分析方法对变量分类的探讨[J]. 统计与决策, 2005(2): 20-21.
- [10] 莫惠东. 农业试验统计[M]. 上海: 上海科学技术出版社, 1984: 151-174.
- [11] 汤梦玲, 王占龙, 李志建. 因子分析法求权重评价水质的实例[J]. 邢台职业技术学院学报, 2005, 22(5): 15-16.