

# 利用主成分分析法评价 6 个燕麦品种

赵晓军

青海省海北州草原工作站,青海海北 810299

**摘要** 为了研究 6 个燕麦品种,利用 IBM SPSS Statistics 22 软件对 6 个燕麦试验品种的经济性状数据进行主成分分析,并对主成分赋值进行排序。试验结果表明:青引 2 号燕麦表现最好,加燕 2 号燕麦次之,林纳燕麦最差。这说明利用主成分分析法评价燕麦品种,可以简洁明了地对试验品种的优劣进行排序,并且可以摒弃不同度量指标因量纲不同带来的差异,可以有效减少系统误差,提高试验的准确性。

**关键词** 主成分分析;燕麦;品种

燕麦(*Avena sativa* L.)属禾本科早熟禾亚科燕麦属植物,具有鲜草产量高、营养价值高和适应性强的生物学特点<sup>[1]</sup>,在青海全省范围内被广泛种植和利用。近年来,饲草种植技术推广部门不断推广以燕麦为主的各种饲草种植,各级农牧业主管部门推进以“粮改饲”、饲草料基地建设等饲草种植补贴项目,特别是在 2019 年初青海省青南地区发生特大雪灾之后,以燕麦为主的青干草、青贮饲草在抗灾保畜工作中作用巨大,极大地保护了广大农牧民的牲畜财产和生态畜牧业生产少受损失,广大农牧民逐渐意识到饲草种植的重要性和迫切性:平时补饲、灾时抗灾。在这种情况下,如何因地制宜地选择燕麦等饲草品种,显得尤为重要。本文根据笔者 2019 年田间试验数据,利用主成分分析(principal component analysis)对 6 个燕麦品种的叶长、叶宽、茎粗度、株高 4 个单株经济性状进行研究,以期对 6 个燕麦品种的优劣进行评价。

本研究种植燕麦试验地位于青海省海北州牧科所(原二分厂),地理位置为 100°59'E、北纬 36°54'E,属高原大陆性气候<sup>[2]</sup>。2019 年全年降水量 405.4 mm,年平均气温 8.6 °C,最热月均温 17.6 °C,最冷月均温-9.7 °C,极端最高温度 24.7 °C,极端最低温度-25.4 °C,无霜期 95 d,初霜期 9 月 11 日,终

霜期 5 月 29 日,年积温( $\geq 0$  °C)1 882 °C,年有效积温( $\geq 10$  °C)1 285 °C(2019 年气象资料由海北州气象站提供)。试验地土壤类型为黑钙土,土壤有机质、碱解氮、速效磷、速效钾含量分别为:3835 g/kg、2.58 mg/kg、1.36 mg/kg、21.69 mg/kg,土壤 pH 为 8.21。

## 1 材料与方法

### 1.1 研究材料

试验选用青引 2 号、青引 3 号、林纳、青海 444、加燕 2 号和白燕 11 号 6 个燕麦品种,燕麦种子由青海省畜牧兽医科学院提供。

### 1.2 研究方法

1)研究数据采集。2019 年 4 月下旬在试验地播种,每个试验品种种植面积 25 m<sup>2</sup>(长 5 m×宽 5 m),播种方式为人工条播,行距 30 cm,播种深度 2~3 cm,播种量为 10 kg/666.67 m<sup>2</sup>,播种前施底肥 64%磷酸二铵 10 kg/666.67 m<sup>2</sup>,拔节期追施 46%尿素 5 kg/666.67 m<sup>2</sup>。初花期对 6 个试验品种随机选取 10 个植株测定叶长、叶宽、茎粗度、株高 4 个单株经济性状数据,取平均数备用。

2)研究数据分析。主成分分析是指利用降维的思想将多个原始变量通过线性组合提炼出少数彼

此独立的新变量(主成分)的多元统计分析方法<sup>[3]</sup>。在研究中,主成分分析将复杂的具有关联性的多维数据通过降维简化成几个彼此之间互不关联而又能充分反映原有信息量的少数综合因子,从而使研究过程和结果变得简单化。本研究采用 IBM SPSS Statistics 22 进行主成分分析的相关运算,得出计算结果并进行分析。

## 2 结果与分析

2019 年 7 月下旬,6 个燕麦试验品种生育期进入初花期,对叶长、叶宽、茎粗度、株高 4 个单株经济性状数据进行测定,并计算平均值,如表 1 所示。利用 IBM SPSS Statistics 22 软件对 6 个燕麦试验品种的 4 个经济性状数据进行标准化值(Z)计算,并进行主成分分析,主要结果见表 2~表 4。表 3 中根据主成分特征值大于 0.5 的原则进行主成分的抽取,共抽取了 2 个主成分,第 1 个主成分对总方差

的贡献率为 71.641%,第 2 个主成分对总方差的贡献率为 21.319%,2 个主成分累计贡献率达 92.960%,这说明抽取的 2 个主成分基本上可以反映出原始试验数据中 92.960%的数据信息,符合主成分分析用于属于数据降维分析的基本思路。

从 2 个主成分的载荷矩阵可以看出(表 4),叶宽、茎粗度、株高在第 1 个主成分上有较高的载荷,这说明第 1 个主成分主要反映了叶宽、茎粗度、株高 3 个生产性状指标的信息;同理,第 2 个主成分主要反映了叶长 1 个生产性状指标的信息。

利用表 3、表 4 中 2 个主成分的基础特征值和载荷数值,计算得出 2 个主成分的特征向量,如表 4 所示。根据 2 个主成分的特征向量和 4 个经济性状数据值,分别计算 6 个燕麦试验品种在每个主成分上的赋值 F1、F2,进而根据表 3 中 2 个主成分的特征值所占主成分特征值之和的比例为权重 F,计算主成分综合赋值并排序,如表 5 所示。

表 1 6 个燕麦试验品种主要生产性状

序号	燕麦试验品种	叶长		叶宽		茎粗度		株高	
		平均值/cm	Z	平均值/cm	Z	平均值/cm	Z	平均值/cm	Z
1	青引 2 号	18.799	0.469	1.074	-1.306	0.392	-1.081	77.882	-1.194
2	青引 3 号	17.072	-0.847	1.348	-0.202	0.388	-1.154	80.865	-0.992
3	林纳	16.297	-1.438	1.358	-0.161	0.485	0.614	99.594	0.281
4	青海 444	18.068	-0.088	1.536	0.556	0.448	-0.061	103.891	0.573
5	加燕 2 号	19.047	0.658	1.270	-0.516	0.465	0.249	93.421	-0.139
6	白燕 11 号	19.818	1.246	1.802	1.629	0.530	1.434	117.104	1.471

表 2 6 个燕麦试验品种主要生产性状数据相关系数矩阵

序号	生产性状	叶长	叶宽	茎粗度	株高
1	叶长	1.000	0.286	0.309	0.313
2	叶宽	0.286	1.000	0.744	0.900
3	茎粗度	0.309	0.744	1.000	0.930
4	株高	0.313	0.900	0.930	1.000

表 3 6 个燕麦试验品种主要生产性状数据解释的总方差

成分序号	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	2.866	71.641	71.641	2.866	71.641	71.641
2	0.853	21.319	92.960	0.853	21.319	92.960
3	0.257	6.428	99.388			
4	0.024	0.612	100.000			

表 4 6 个燕麦试验品种主要生产性状数据主成分载荷矩阵及特征向量

序号	生产性状	主成分载荷矩阵		特征向量	
		1	2	A1	A2
1	叶长	0.458	0.889	0.398	1.021
2	叶宽	0.913	-0.153	0.562	-0.424
3	茎粗度	0.928	-0.127	0.567	-0.386
4	株高	0.980	-0.152	0.583	-0.422

表 5 6 个燕麦试验品种主要生产性状数据主成分赋值及排序

序号	燕麦试验品种	主成分赋值			品种排序
		F <sub>1</sub>	F <sub>2</sub>	F	
1	青引 2 号	-1.330	41.664	40.334	1
2	青引 3 号	-1.206	1.800	0.594	3
3	林纳	-0.108	-37.423	-37.531	6
4	青海 444	0.413	-11.600	-11.186	5
5	加燕 2 号	0.023	18.189	18.212	2
6	白燕 11 号	2.208	-12.630	-10.422	4

### 3 讨 论

本研究利用 IBM SPSS Statistics 22 软件对 6 个燕麦试验品种的经济性状数据进行标准化值和主成分分析,并对主成分赋值进行排序,青引 2 号燕麦表现最好,加燕 2 号燕麦次之,林纳燕麦最差,这与试验实际情况基本符合。利用主成分分析法评价燕麦品种,对不同经济性状的数据进行无纲化处理,利用标准化值进行计算,摒弃了不同经济性状之间量纲不同带来的差异,可以有效减少系统误差,提高试验的准确性。

### 参 考 文 献

- [1] 赵德.青海省共和县 6 个燕麦品种引种比较试验[J].畜牧与饲料科学,2017,38(5):36-37.
- [2] 陈秉龙.环青海湖地区 4 个老芒麦品种比较试验[J].青海草业,2017,26(3):19-21.
- [3] 张水清,黄绍敏,郭斗斗.主成分分析在潮土土壤肥力评价中的应用[J].河南农业科学,2011,40(4):82-86.

【责任编辑:胡 敏】