

基于岭回归分析的烤烟焦油含量预测模型构建

牛慧伟¹, 许自成^{1*}, 李青常², 王龙宪³, 邵惠芳¹, 焦敬华³

(1. 河南农业大学 烟草学院, 河南 郑州 450002; 2. 中国烟草总公司 郑州烟草研究院, 河南 郑州 450001; 3. 中国烟草总公司辽宁省公司, 辽宁 沈阳 110012)

摘 要: 为建立烤烟焦油含量的预测模型, 对烤烟总糖含量(X_1)、还原糖含量(X_2)、总氮含量(X_3)、烟碱含量(X_4)、钾含量(X_5)、氯含量(X_6)、糖碱比(X_7)、氮碱比(X_8)和钾氯比(X_9) 9 项化学指标进行基于相关分析和统计检验的多重共线性诊断, 并对其与焦油含量(Y)进行了岭回归分析。结果表明, 烤烟总糖、还原糖、总氮、烟碱、糖碱比和氮碱比 6 项化学指标间存在共线性, 且 9 项化学指标与焦油含量间均存在极显著相关关系, 表明采用岭回归方法建立的以该 9 项化学指标为自变量的烤烟焦油含量多元线性回归模型是合理的。当岭回归参数为 0.08 时的烤烟焦油含量预测模型为 $Y = 18.8009 - 0.0007X_1 - 0.0342X_2 + 1.6252X_3 + 0.6911X_4 - 0.9686X_5 + 0.2927X_6 - 0.0299X_7 - 3.5193X_8 - 0.0568X_9$ ($R^2=0.8175$)。回归方程通过显著性检验($P<0.01$), 且回归系数的符号均与相关分析结果相一致, 有效地使运用最小二乘法估计时符号不合理的回归系数变得合理。

关 键 词: 烤烟; 焦油; 化学成分; 多重共线性诊断; 岭回归分析

中图分类号: TB302.2 文献标志码: A 文章编号: 1007-1032(2012)03-0245-06

Model for predicting tar content of flue-cured tobacco based on ridge regression analysis

NIU Hui-wei¹, XU Zi-cheng^{1*}, LI Qing-chang², WANG Long-xian³, SHAO Hui-fang¹, JIAO Jing-hua³

(1.College of Tobacco Science, Henan Agricultural University, Zhengzhou 450002, China; 2.Zhengzhou Tobacco Research Institute of China Tobacco Corporation, Zhengzhou 450001, China; 3.Liaoning Provincial Tobacco Corporation of China Tobacco Corporation, Shenyang 110012, China)

Abstract: In order to build the model for predicting the tar content of flue-cured tobacco, multicollinearity of 9 chemical indexes including total sugar (X_1), reduced sugar (X_2), total nitrogen (X_3), nicotine (X_4), potassium (X_5), chlorine (X_6), total sugar to nicotine ratio (X_7), total nitrogen to nicotine ratio (X_8) and potassium to chlorine ratio (X_9) was diagnosed through correlation and statistic test, while the ridge regression between tar content and chemical components was also conducted. The results indicated that collinearity existed among total sugar, reduced sugar, total nitrogen, nicotine, total sugar to nicotine ratio, and total nitrogen to nicotine ratio. And there was extremely significant correlation between the tar content and the 9 chemical indexes mentioned above. Therefore, it was reasonable to establish the multiple linear regression model using 9 chemical components as independent variables based on ridge regression. The prediction model for tar content using ridge regression at parameter $K=0.08$ was $Y = 18.8009 - 0.0007X_1 - 0.0342X_2 + 1.6252X_3 + 0.6911X_4 - 0.9686X_5 + 0.2927X_6 - 0.0299X_7 - 3.5193X_8 - 0.0568X_9$ ($R^2=0.8175$), and the regression equation passed the significance test ($P<0.01$). The regression coefficients properties of ridge regression equation were consistent with the results of correlation analysis, which made the unreasonable symbols of regression coefficients in the least square estimation reasonable.

Key words: flue-cured tobacco; tar; chemical components; multicollinearity diagnostics; ridge regression analysis

收稿日期: 2011-12-14

基金项目: 中国烟草总公司科技重大专项(110201101002(TS-02)); 中国烟草总公司辽宁省公司重点科技项目(LY201001)

作者简介: 牛慧伟(1986—), 女, 河南辉县人, 硕士研究生, 主要从事烟草品质生态及烟草营养与烟叶质量评价研究, niuhuiwei1986@163.com; *通信作者, zcxu@sohu.com

国际上普遍把焦油含量作为评判卷烟安全性的重要指标^[1]。烟叶焦油的检测耗时、繁琐^[2],而烟叶相关理化指标的检测相对简捷迅速。大量研究^[3-11]结果表明,烤烟化学成分与焦油含量间有着密切的关系,建立烤烟化学成分对焦油含量的预测模型,对低危害烟草制品的设计生产、提高烟草的吸食安全性等均具有重要意义。

基于烤烟或卷烟化学成分的焦油含量预测模型构建常采用多元回归分析和基于神经网络等数据挖掘技术^[6,12-15],其中多数回归方程是采用普通最小二乘法(least squares, LS)建立的^[16-17]。模型中的自变量(烤烟化学成分含量)间存在密切关系^[18-19],而普通最小二乘方程忽视了“自变量不存在近似线性关系”这一应用条件^[20-22],因而基于存在缺陷的模型所作的预测也就失效^[23]。目前,处理多重共线性常用的方法有岭回归、逐步回归、主成分回归、偏最小二乘法等^[24-25]。关于逐步回归方法已有报道^[10,26],但当共线性较为严重时,变量自动筛选的方法并不能完全解决问题^[24],会损失某些信息^[20];岭回归虽为有偏估计,但能有效地控制回归系数的标准误差大小^[24],精度得到极大提高^[27],同时通过减小LS估计中因共线性而产生的大方差,从而达到克服共线性影响的目的^[20]。岭回归主要在医学^[28]、经济^[29]等领域应用较多,而在烟草研究中的应用^[30]较少。笔者运用岭回归方法,在对烤烟常规化学成分及其协调性指标进行多重共线性诊断的基础上,建立烤烟化学成分与焦油含量的岭回归方程,旨在为烤烟焦油含量预测提供参考依据。

1 材料与方法

1.1 材料

供试材料来自云南、贵州、四川、河南、湖南、湖北、广东、广西、福建、陕西、安徽和辽宁等28个烤烟主产区2009年的主栽品种,包括B2F、C3F和X2F共3个等级的初烤烟叶样品58个。

1.2 样品制备及化学成分含量的测定

所有样品的卷制采用相同的工艺和相同的烟用材料。在样品卷制前,对各烟叶样品进行留样,以备化学分析测试用;总糖、还原糖、总氮、烟碱、

钾、氯的含量均采用文献[31]方法进行测定。

1.3 烤烟焦油含量预测模型构建

以烤烟总糖含量(X_1)、还原糖含量(X_2)、总氮含量(X_3)、烟碱含量(X_4)、钾含量(X_5)、氯含量(X_6)、糖碱比(X_7)、氮碱比(X_8)和钾氯比(X_9)9个化学成分指标作为自变量,建立焦油含量(Y)的岭回归预测模型。建模的基本步骤为:先运用相关分析和统计检验方法对9个自变量进行多重共线性诊断;在此基础上再运用岭回归方法建立烤烟焦油含量多元线性回归模型。

1.3.1 多重共线性诊断

自变量的简单相关系数矩阵可作为多重共线性的经验式诊断依据,当某些自变量间的相关系数较大时可能存在多重共线性^[21]。此方法只限于在考察2个变量之间的共线性关系时能够给出较正确的判断^[20],对多个自变量间的共线性诊断,还需进一步采用更精确的统计检验方法进行判定。

共线性诊断常用的统计量有容忍度(tolerance)、方差膨胀因子(variance inflation, VIF)、特征根(eigenvalue)、条件指数(condition index)和方差比例(variance proportion)等^[21,24]。当自变量容忍度小于0.1,或 VIF 大于5(也可放宽至大于10)时,表明自变量间存在严重的多重共线性。对最小二乘法建立的模型中的常数项及所有自变量计算主成分,若前面几个主成分特征根数值较大,后面几个较小甚至接近于0,或其中几个条件指数较大(大于30),或某个主成分对2个或多个自变量的贡献(即方差比例)均较大(大于0.5)时,则提示存在多重共线性。

1.3.2 岭回归模型的建立

在自变量存在多重共线性的基础上,利用SASV8的reg过程建立岭估计的线性回归模型^[32]。建立岭回归模型的关键是岭回归参数 K 值的确定。笔者依据以下方法确定 K 值:各自变量的回归系数趋于平稳(无正负波动),且此时各自变量的 VIF 均小于5,同时残差平方和较最小二乘方程增大不太多;借助岭迹图进行判断。

1.4 数据处理

采用SPSS18.0软件进行相关分析和共线性诊断;采用SASV8和DPS软件进行岭回归分析。

2 结果与分析

2.1 烤烟各化学成分指标的多重共线性诊断

2.1.1 烤烟焦油含量预测模型各变量间的相关性

由表 1 可知,除烤烟钾与总糖、钾与还原糖、钾与氯、钾氯比与还原糖之间相关性均不显著,总糖与钾氯比、总氮与钾、钾与糖碱比、氮碱比与钾

氯比之间均相关显著外,其他化学成分含量间的相关关系均达到了极显著水平,可初步推断这些化学成分间存在较严重的多重共线性。由此可知,反应变量与自变量间有明显的线性关系,符合多重线性回归分析的条件^[24],可以将上述化学成分含量作为自变量建立关于烤烟焦油含量的多元线性回归模型。

表 1 烤烟焦油含量预测模型各变量间的相关系数

Table 1 Correlation analysis between chemical components and tar content of flue-cured tobacco									
指标	总糖含量	还原糖含量	总氮含量	烟碱含量	钾含量	氯含量	糖碱比	氮碱比	钾氯比
还原糖含量	0.977**								
总氮含量	-0.784**	-0.777**							
烟碱含量	-0.798**	-0.805**	0.864**						
钾含量	0.153	0.185	-0.313*	-0.369**					
氯含量	-0.515**	-0.514**	0.473**	0.539**	-0.17				
糖碱比	0.873**	0.857**	-0.867**	-0.927**	0.288*	-0.502**			
氮碱比	0.659**	0.668**	-0.641**	-0.909**	0.351**	-0.462**	0.884**		
钾氯比	0.279*	0.252	-0.407**	-0.396**	0.428**	-0.770**	0.356**	0.321*	
焦油含量	-0.701**	-0.711**	0.790**	0.878**	-0.470**	0.526**	-0.839**	-0.814**	-0.477**

2.1.2 烤烟各化学成分指标多重共线性诊断的统计检验

用最小二乘法建立烤烟 9 项化学成分指标对焦油含量的多元线性回归模型,其参数估计及显著性检验和自变量的共线性统计量见表 2。统计分析结果显示,该模型的决定系数 $R^2=0.825\ 7$,复相关系数 $R=0.908\ 7$ 。模型检验统计量 $F=25.266$,显著性水平 $P<0.01(P=0.000)$,模型具有统计学意义。

由表 2 可知,各自变量的回归系数均不能通过 t 测验,因而各自变量无统计学意义,且某些自变量如烟碱的回归系数符号明显与常规和相关分析结果不符,表明自变量存在多重共线性的可能性很大;从共线性统计量可知,除钾、氯和钾氯比外,其他 6 项化学成分的容忍度均小于 0.1,方差膨胀因子(VIF)均远大于 5,甚至均远超过 10,表明自变量间存在严重的多重共线性。

表 2 最小二乘多元线性回归参数估计及显著性检验

项目	非标准化系数		标准系数	t 值	相伴概率	共线性统计量	
	系数	标准误差				容忍度	VIF
常数项	22.915 1	5.412	—	4.234**	0.000	—	—
总糖含量	-0.079 9	0.169	-0.191 4	-0.473	0.638	0.022	45.036
还原糖含量	-0.059 9	0.155	-0.123 8	-0.386	0.701	0.035	28.292
总氮含量	5.961 7	3.345	0.726 5	1.782	0.081	0.022	45.743
烟碱含量	-1.142 5	1.650	-0.363 5	-0.692	0.492	0.013	75.942
钾含量	-0.794 7	0.539	-0.118 1	-1.474	0.147	0.566	1.768
氯含量	-0.490 0	2.462	-0.025 4	-0.199	0.843	0.224	4.472
糖碱比	0.234 1	0.200	0.673 3	1.171	0.247	0.011	91.097
氮碱比	-13.774 3	7.739	-0.994 4	-1.780	0.081	0.012	85.972
钾氯比	-0.075 5	0.073	-0.129 5	-1.031	0.308	0.230	4.343

对最小二乘线性回归模型进行共线性诊断。由表 3 可知, 仅第 1 个主成分的特征值较大, 而后面 9 个主成分均较小, 甚至逐渐接近于 0; 第 7、8、9、10 个主成分的条件指数均大于 30, 且第 9、10 个均大于 100(第 10 个主成分的条件指数达到 233.892), 其相对应的自变量中总氮、烟碱、糖碱

比和氮碱比的方差比例均超过 50%, 总糖的方差比例也接近于 50%, 说明这 5 个自变量间存在一定程度的共线性。综合分析可知, 回归模型的自变量间存在多重共线性, 可采用岭回归方法对烤烟焦油含量进行多元线性回归模型分析。

表 3 最小二乘多元线性回归模型的共线性诊断

Table 3 Collinearity diagnostics of least square multiple linear regression

主成分	特征值	条件指数	方差比例									
			常数项	总糖	还原糖	总氮	烟碱	钾	氯	糖碱比	氮碱比	钾氯比
1	9.187	1.000	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 2	0.000 3	0.000 0	0.000 0	0.000 0
2	0.514	4.228	0.000 0	0.000 0	0.000 0	0.000 1	0.000 5	0.000 0	0.014 3	0.000 8	0.000 0	0.027 1
3	0.211	6.602	0.000 0	0.000 1	0.000 2	0.000 1	0.000 7	0.000 5	0.007 6	0.002 1	0.000 1	0.145 1
4	0.043	14.538	0.000 2	0.002 1	0.002 6	0.000 6	0.003 6	0.124 3	0.202 9	0.000 0	0.000 0	0.057 5
5	0.020	21.542	0.000 1	0.001 7	0.000 4	0.003 6	0.001 2	0.234 6	0.345 8	0.004 4	0.004 7	0.310 3
6	0.017	23.297	0.000 1	0.005 8	0.009 6	0.002 5	0.000 0	0.382 9	0.029 0	0.006 0	0.007 4	0.066 2
7	0.006	38.036	0.000 6	0.008 2	0.018 6	0.002 5	0.067 6	0.170 9	0.224 6	0.107 8	0.002 5	0.175 8
8	0.001	81.358	0.343 0	0.051 1	0.011 8	0.097 1	0.000 6	0.038 5	0.092 8	0.021 6	0.003 8	0.087 2
9	0.001	119.329	0.150 3	0.465 0	0.897 3	0.001 4	0.036 7	0.045 5	0.081 3	0.001 5	0.009 0	0.130 4
10	0.000	233.892	0.505 7	0.465 8	0.059 3	0.892 1	0.888 9	0.002 6	0.001 5	0.855 8	0.972 5	0.000 0

2.2 岭回归模型的建立

根据不同岭参数时各自变量的参数估计值作岭迹图(图 2)。由图 2 和不同岭参数 K 值下自变量回归系数及 VIF 的变化可以看出, 当 $K=0.08$ 时, 自变量的回归系数趋于平稳, 且此时各自变量的 VIF 均小于 5($X_1 \sim X_9$ 的 VIF 依次为 1.515 5、1.618 7、1.628 1、1.779 7、1.084 0、1.714 1、1.747 1、1.113 9、1.620 9), 故选岭参数为 0.08, 此时的标准化回归方程为: $Y = -8.334 0 \times 10^{-16} - 0.001 7X_1 - 0.070 6X_2 + 0.198 0X_3 + 0.219 9X_4 - 0.143 9X_5 + 0.015 2X_6 - 0.086 1X_7 - 0.254 1X_8 - 0.097 5X_9$ 。该岭回归模型的 $R^2=0.817 5$, 复相关系数 $R=0.904 1$, 比最小二乘模型略小, $F=23.897 5$, 显著性水平 $P=0.000$, 模型具有统计学意义。岭回归方程的标准误差为 0.465 47, 虽然比最小二乘法回归方程的标准误差(0.454 95)有所增

大, 但仅增大 0.010 52。岭回归方程中烟碱、氯和糖碱比的回归系数符号均与最小二乘回归方程相反, 而与相关分析的结果相符, 表明岭回归分析能够克服共线性的影响, 从而更有效地对烤烟焦油含量进行预测。

从该岭回归方程中可知, 烤烟总氮、烟碱和氯含量与焦油含量呈正相关, 其他自变量与焦油含量呈负相关, 这与相关分析的结果相一致; 对烤烟焦油含量影响最大的变量是氮碱比(X_8), 其次是烟碱含量(X_4)、总氮含量(X_3)和钾含量(X_5)。

若将标准化后的变量表述为原始变量, 则岭参数为 0.08 时的回归方程为: $Y = 18.800 9 - 0.000 7X_1 - 0.034 2X_2 + 1.625 2X_3 + 0.691 1X_4 - 0.968 68X_5 + 0.292 70X_6 - 0.029 9X_7 - 3.519 3X_8 - 0.056 87X_9$ 。

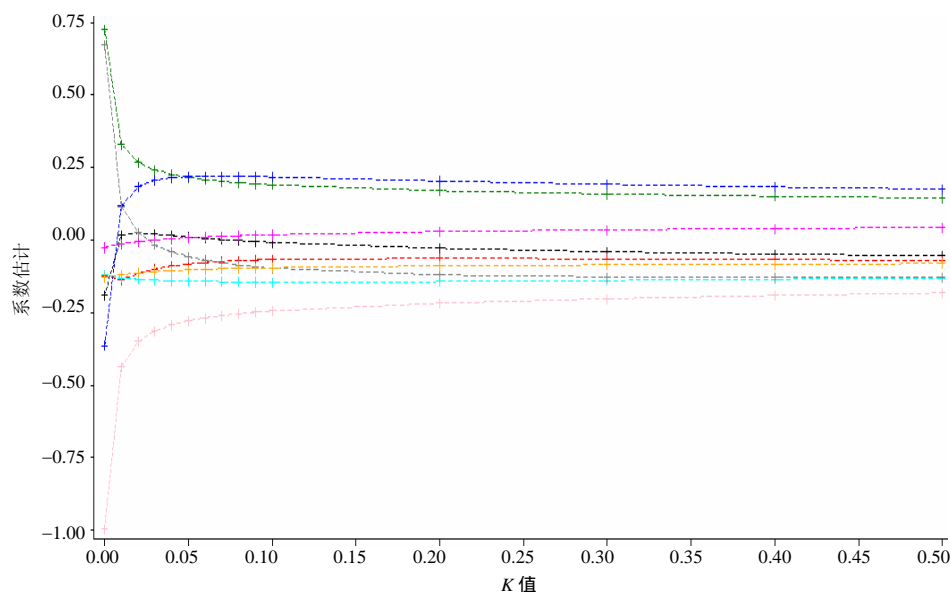


图 2 岭回归分析的岭迹图

Fig. 2 Ridge trace diagram of ridge regression analysis

3 结论与讨论

本研究结果表明, 烤烟总糖、还原糖、总氮、烟碱、糖碱比和氮碱比 6 项化学成分间存在多重共线性, 且烤烟总糖、还原糖、钾、糖碱比、氮碱比和钾氯比均与焦油含量呈极显著负相关, 总氮、烟碱和氯均与焦油含量呈极显著正相关关系, 运用岭回归方法建立以 9 项化学成分为自变量的烤烟焦油含量多元线性回归预测模型是合理的。

通过观察不同岭参数 K 值下岭迹图和方差膨胀因子的变化, 得知当 $K=0.08$ 时自变量的回归系数趋于稳定, 从而得到 $K=0.08$ 时的岭回归方程。该岭回归方程的决定系数 R^2 比最小二乘法方程的略小, 标准误差比最小二乘法方程的有所增大, 但增大不多。岭回归方程中烟碱、氯和糖碱比的回归系数符号均与最小二乘方程的相反。相关分析结果表明, 焦油含量与烟碱、氯含量呈正相关关系, 与糖碱比呈负相关关系, 这与人研究结果^[3-5,7,10-11]一致。岭回归方程回归系数的符号均与此相一致, 表明岭回归有效地使运用最小二乘估计时符号不合理的回归系数变得合理。以上结论均符合岭回归的基本思想。岭回归是有偏估计, 其通过放弃最小二乘法的无偏性, 以损失部分信息、降低精度为代价来寻求效果稍差, 但回归系数更符合实际的回归方程, 所得剩余标准差比最小二乘回归模型要大, 但

其对不合理数据的耐受性远强于最小二乘法^[21,24]。运用岭回归方法的关键是 K 值的确定, 其决定了方程的回归系数和均方误差; 岭估计的残差平方和随着 K 值的增加而增加, 为使其残差平方和增大不太多, 在使岭估计的各个回归系数大体稳定的基础上, K 值应尽量小^[21,24]。本研究的岭参数 K 在实现岭回归方程各个回归系数稳定的基础上取得较小值, 保证了残差平方和较小的增幅, 因此该岭回归方程回归效果总体较好, 减弱了多重共线性的影响, 提高了回归方程的稳定性和可解释性。

从岭回归方程中可知, 烤烟总氮、烟碱和氯与焦油含量呈正相关, 其他自变量与焦油含量呈负相关, 这与相关分析的结果相一致。对烤烟焦油含量影响最大的变量是氮碱比(X_8), 其次是烟碱(X_4)、总氮(X_3)和钾(X_5), 这与邓小华等^[7]的研究结果相似; 因此, 灵活运用岭回归方法, 可以在分析各变量之间的作用和关系时带来独特而有效的帮助^[25]。

由于岭回归方法 K 值的确定有很大的人为因素, 使之存在一定的缺陷^[27]; 此外, 不同生态地区烤烟化学成分和焦油量差异均较大, 本研究烟叶样品来自 28 个不同产区, 各指标及其之间的相关关系可能受产区影响而缺乏一定的规律性, 从而影响回归结果, 故本文建立的模型尚有待进一步验证, 需根据不同产区的烟叶质量特点建立相应的模型。

参考文献:

- [1] 杜咏梅, 冯协忠, 王允白. 烟气焦油与卷烟安全性[J]. 中国烟草科学, 2002(2): 31-34.
- [2] 闫克玉, 李兴波, 李志同, 等. 烤烟(40级)烟叶焦油量和烟气烟碱的测定分析[J]. 郑州轻工业学院学报, 1994, 9(2): 52-57.
- [3] 李国栋, 于建军, 董顺德, 等. 河南烤烟化学成分与烟气成分的相关性分析[J]. 烟草科技, 2001(8): 28-30.
- [4] 于建军, 章新军, 毕庆文, 等. 烤烟烟叶理化特性对烟气烟碱、CO、焦油量的影响[J]. 中国烟草科学, 2003(3): 5-8.
- [5] 厉昌坤, 周显升, 王允白, 等. 烤烟烟叶焦油释放量与部分化学成分的关系研究[J]. 中国烟草科学, 2004(2): 25-27.
- [6] 贺英, 徐海涛, 盛志艺, 等. 综合方法对烤烟化学成分和烟气组分的相关分析[J]. 中国烟草科学, 2005(4): 1-4.
- [7] 邓小华, 周冀衡, 李晓忠, 等. 烤烟质量与焦油量的灰色关联分析[J]. 江西农业大学学报, 2006, 28(6): 850-854.
- [8] 彭玉富, 李光照, 朱智志. 灰色关联分析及其在烟叶成分与烟气成分关系研究中的应用[J]. 郑州轻工业学院学报: 自然科学版, 2007, 22(4): 20-22.
- [9] 汤朝起, 奚玉青, 张俊. 烤烟物理特性和化学成分与烟气组分的关系[J]. 四川农业大学学报, 2009, 27(4): 427-432.
- [10] 汪修奇, 邓小华, 李晓忠, 等. 湖南烤烟化学成分与焦油的相关、途径及回归分析[J]. 作物杂志, 2010(2): 32-35.
- [11] 邓小华, 周冀衡, 赵松义, 等. 单料烤烟烟气颗粒物与质量评价指标间的相关性研究[J]. 中国烟草科学, 2010, 31(1): 60-64.
- [12] 李炎强, 相秉仁, 屈凌波, 等. 卷烟主流烟气成分释放量的人工神经网络预测研究[J]. 计算机与应用化学, 2006, 23(9): 880-884.
- [13] 杨再波, 叶冲, 韩伟, 等. 基于BP神经网络在卷烟焦油预测中的应用[J]. 黔南民族师范学院学报, 2006, 6: 5-8.
- [14] 倪力军, 曾晓虹, 张立国. KNN-KSR建模方法及其在卷烟主流烟气预测中的应用[J]. 华东理工大学学报: 自然科学版, 2008, 34(4): 547-552.
- [15] 陆鸣, 李东亮, 许自成, 等. 卷烟烟气焦油量的ARMA预测模型研究[J]. 安徽农业科学, 2008, 36(14): 5696-5698.
- [16] 董平, 汪开保. 利用统计法建立卷烟焦油预测模型[J]. 烟草科技, 2001(12): 33-35.
- [17] 张志刚, 王二彬, 苏东赢. 卷烟常规化学成分与焦油的线性回归分析[J]. 烟草科技, 2003(11): 32-33.
- [18] 邓小华, 周冀衡, 李晓忠, 等. 湖南烤烟化学成分特征及其相关性[J]. 湖南农业大学学报: 自然科学版, 2007, 33(1): 24-27.
- [19] 李东亮, 许自成. 烤烟化学成分指标的灰色关联聚类分析[J]. 农业系统科学与综合研究, 2007, 23(4): 411-414, 420.
- [20] 刘国旗. 多重共线性的产生原因及其诊断处理[J]. 合肥工业大学学报: 自然科学版, 2001, 24(4): 607-610.
- [21] 马雄威. 线性回归方程中多重共线性诊断方法及其实证分析[J]. 华中农业大学学报: 社会科学版, 2008(2): 78-81, 85.
- [22] 高辉. 多重共线性的诊断方法[J]. 统计与信息论坛, 2003, 18(1): 73-76.
- [23] William H Greene. Econometric Analysis[M]. 2 ed. New York: Macmillan, 1993: 288-313.
- [24] 张文彤. SPSS 统计分析高级教程[M]. 北京: 高等教育出版社, 2004: 92-93, 114, 128, 130.
- [25] 陈玲燕. 多重共线性下的线性回归方法综述[J]. 市场研究, 2008(4): 39-41.
- [26] 王建民, 甘学文, 李晓, 等. 叶组配方卷烟烟气预测模型的建立[J]. 烟草科技, 2006(6): 5-8.
- [27] 周松青. 解决多重共线性问题的线性回归方法[J]. 江苏统计, 2000(11): 24-25.
- [28] 胡军, 姜潮, 图易宸, 等. 多元岭回归方法在医院成本分析中的应用[J]. 中国卫生统计, 2000, 17(6): 334-336.
- [29] 徐业茂, 金隼, 赵勇, 等. 基于岭回归方法的企业技术创新能力评价模型研究[J]. 机械设计与制造, 2007(10): 194-196.
- [30] 徐雅静, 汪远征, 王建民, 等. 卷烟焦油含量预测的数学模型[J]. 郑州轻工业学院学报: 自然科学版, 2005, 20(3): 35-38.
- [31] 王瑞新, 韩富根, 杨素勤. 烟草品质分析[M]. 郑州: 河南科学技术出版社, 1990.
- [32] 郭秀花, 赵连伟. SAS6.11版岭回归分析程序设计及其实例分析[J]. 数理统计与管理, 2001, 20(1): 41-44, 64.

责任编辑: 杨盛强