

百科探索

基于线性回归模型的中国新能源汽车发展研究

胡艺缤 余青昊 黄俊炜 韦露 朱治衡*

(桂林理工大学南宁分校 广西崇左 532500)

摘要: 新能源汽车因其较低的环境污染特性备受关注,成为未来汽车产业发展的重要方向。本研究旨在通过线性回归分析,构建一个反映新能源汽车销量的模型。结果表明,产量对新能源汽车销量呈正向影响,即增加产量可能促使销量的提升。然而,公共类充电桩数量和动力电池产量对销量呈负向影响,可能由于充电基础设施的不足以及动力电池的高成本影响了新能源汽车的市场竞争力。配套动力电池企业数量对销量呈正向影响,表明优质的动力电池供应链可能有助于提升新能源汽车的市场地位。

关键词: 线性回归模型; 新能源汽车; 公共类充电桩数量; 动力电池产量; 充电基础设施

Study on the Development of New Energy Vehicles in China Based on Linear Regression Models

Yibin Hu、Qinghao Yu、Junwei Huang、Lu Wei、Zhiheng Zhu*

(Guilin University of Technology At Nanning, Chongzuo, Guangxi, 532500)

Abstracts: New energy vehicles have attracted much attention because of their lower environmental pollution characteristics, and have become an important direction for the future development of the automotive industry. This study aims to construct a model reflecting the sales volume of new energy vehicles through linear regression analysis. The results show that production has a positive effect on the sales of new energy vehicles, i.e., increasing production may lead to an increase in sales. However, the number of public charging columns and the production of power batteries have a negative impact on sales, possibly due to the lack of charging infrastructure and the high cost of power batteries, which affects the competitiveness of new energy vehicles in the market. The number of supporting power battery companies has a positive impact on sales, suggesting that a high-quality power battery supply chain may help to improve the market position of new energy vehicles.

Keywords: linear regression model; new energy vehicles; number of public-type charging columns; power battery production; charging infrastructure

1 引言

随着环境污染和能源消耗问题日益凸显,新能源汽车作为一种环保、高效的交通工具,受到了广泛关注和推崇。新能源汽车采用非常规车用燃料为动力源,具有低污染、低能耗等特点,被认为是未来汽车发展的重要方向。在全球范围内,新能源电动汽车得到了消费者和政府的青睐。电动公交车和家用电动汽车成为新能源电动汽车的代表,其在城市交通和个人出行领域发挥了重要作用。尤其是在中国,政府积极推动新能源电动汽车的发展,并出台了一系列优惠政策,以促进新能源汽车的普及和推广。在这样的背景下,新能源电动汽车的发展面临着一系列问题和挑战。例如,电池技术的进一步提升和成本的降低、充电设施的建设 and 普及、续航里程的增加、安全性和可靠性的提高等。解决这些问题将推动新能源电动汽车行业的可持续发展,为环境保护和能源可持续利用做出贡献。

2 模型建立

在建立线性回归模型时,我们假设因变量 Y 是自变量 x_1, x_2, \dots, x_n 的线性组合,并引入回归系数 $\beta_0, \beta_1, \beta_2, \dots, \beta_n$,得到一个基本的线性回归模型。计算公式如下:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

其中, Y 代表因变量,是我们要预测或分析的目标变量。在建立线性回归模型时,我们希望能够通过自变量的信息来预测或解释因变量的变化。 β_0 称为截距项

(Intercept),表示当所有自变量 (x_1, x_2, \dots, x_n) 的值都为0时,因变量(Y)的预测值^[1]。截距项是线性回归模型在自变量为0时的起始点。 $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ 分别表示自变量 x_1, x_2, \dots, x_n 对因变量 y 的影响程度,即自变量每变动一个单位时,因变量的变动量。这些回归系数是模型的斜率。 x_1, x_2, \dots, x_n 代表自变量,也称为特征变量。这些自变量是我们用来预测或解释因变量的变化^[2]。

线性回归模型假设因变量与自变量之间的关系是线性的,即因变量是自变量的线性组合。我们通过找到最优的回归系数 $(\beta_0, \beta_1, \beta_2, \dots, \beta_n)$,使得模型预测值(Y 的估计值)尽可能接近实际观测值,从而建立了一个线性关系的模型。然后,我们可以使用这个模型来预测新数据的因变量值,或者分析回归系数的符号和大小,了解自变量对因变量的影响程度。

2.3 计算预测值

(1) 使用模型中的回归系数,计算每个样本点的模型预测值 \hat{y}_{hat} 。对于每个样本点,将自变量 x_1, x_2, \dots, x_n 的值代入模型中,得到对应的因变量的预测值 \hat{y}_{hat} 。计

算线性回归模型的预测值 \hat{y}_{hat} 是为了得到每个样本点在当前建立的模型下的因变量的估计值^[3]。这些预测值可以用来评估模型的拟合效果或进行新数据的预测。对于每个样本点，我们都得到了模型对于自变量的预测值 \hat{y}_{hat} 。这些预测值表示了在当前建立的线性回归模型下，自变量 x_1, x_2, \dots, x_n 所对应的因变量 y 的估计值。

2.4 计算残差

计算残差是线性回归模型中的一个重要步骤，用于衡量模型预测值与实际观测值之间的差异。残差是实际观测值与模型预测值之间的误差，通过计算残差可以评估模型的拟合效果。对于每个样本点，将实际观测值 y 减去对应的模型预测值 \hat{y}_{hat} ，得到残差值 $residual$ ^[4]。残差表示了模型预测值与实际观测值之间的误差，即模型的预测偏差。计算公式为：

$$residual = y - \hat{y}_{hat} \tag{3}$$

其中， $residual$ 是残差， y 是实际观测值， \hat{y}_{hat} 是模型预测值。

将得到的残差值保存起来，以便后续分析和使用。通常，我们会对所有样本点的残差进行统计分析，例如计算残差的均值、方差等，来评估模型的拟合效果。

通过计算残差，我们可以了解线性回归模型对数据的拟合程度。较小的残差表示模型能够较好地拟合数据，预测值与实际观测值之间的误差较小。较大的残差可能意味着模型还需要改进或者数据存在一些异常值或非线性关系。通过 SPSS，我们得到残差统计如下表所示。

表1：残差统计

	最小值	最大值	平均值	标准偏差	个案数
预测值	-6252.30	727394.8	256609.7	197034.3	43
残差	-28874.93	23097.00	.000	12118.80	43
标准预测值	-1.334	2.389	.000	1.000	43
标准残差	-2.266	1.813	.000	.951	43

从表 1 中我们可以得出，平均残差接近于 0，说明模型整体对客运量的预测效果较好，并且标准预测值和标准残差接近于 0，表示预测值和残差在平均值附近波动，模型的预测相对稳定。

3 最小二乘法优化

最小二乘法是一种常用的优化方法，用于计算线性回归模型的最优回归系数，使得残差平方和 (RSS) 最小化。最小二乘法的目标是通过调整回归系数的值，使得模型的预测值与实际观测值之间的误差平方的总和最小^[5]。

计算每个样本点的残差，即实际观测值 y 与模型预测值 \hat{y}_{hat} 之间的误差 ($residual = y - \hat{y}_{hat}$)。残差平方和 (RSS) 是将所有样本点的残差的平方求和，用于衡量整个模型的预测误差。

$$RSS = \sum(residual^2) \tag{5}$$

通过调整回归系数 ($\beta_0, \beta_1, \beta_2, \dots, \beta_n$) 的值，使得 RSS 最小化。最小化 RSS 的目标是找到最优的回归系数，使得模型的预测值尽可能接近实际观测值，即模型的拟合效果最好。将 RSS 作为优化的目标函数。回归系数 $\beta_0, \beta_1, \beta_2, \dots, \beta_n$ 是目标函数的变量。我们需要找到一组回归系数的值，使得目标函数最小化。使用数学和统计的方法，对目标函数进行优化，求解最优的回归系数。这个过程通常涉及对目标函数求导并令导数为零，或者使用数值优化算法来寻找最小值。一旦求解得到最优的回归系数 ($\beta_0, \beta_1, \beta_2, \dots, \beta_n$)，我们就得到了最优的线性回归模型。这个模型能够在当前数据集下最好地拟合自变量与因变量之间的线性关系。

通过最小二乘法的优化过程，我们能够找到最优的回归系数，使得线性回归模型能够最小化预测误差，从而得到一个较好的线性关系模型。

使用最优回归系数：将最优的回归系数 ($\beta_0, \beta_1, \beta_2, \dots, \beta_n$) 代入线性回归模型的数学表达式中。最终我们得到线性回归模型为：

$$y = -36067.029 + 1.023 * \text{产量} - 0.002 * \text{公共类充电桩数量} - 0.116 * \text{动力电池产量} + 909.920 * \text{配套动力电池数量} \tag{6}$$

其中， y 表示新能源汽车销量。从中我们可以得到，新能源汽车销量与产量呈现正向线性关系，即新能源汽车产量每增加 1 辆，预计销量将增加 1.023 辆，并且与公共类充电桩数量、动力电池企业数量呈稍微负向线性关系，与配套动力电池企业数量呈正向线性关系。

4 结论

研究发现，新能源汽车的产量与销量呈正向关系，即随着产量的增加，新能源汽车的销量也相应增加。这可能表明生产规模的扩大有助于提高市场供给，从而推动新能源汽车销售。结果显示，公共类充电桩数量与销量呈负相关关系，暗示着充电基础设施的不足可能是制约新能源汽车销售的一个因素。加强充电基础设施建设可能有助于消除充电难题，提高新能源汽车的使用便利性。研究表明，动力电池产量与销量呈负相关关系，这可能反映了动力电池成本较高、技术水平相对滞后等因素影响了新能源汽车市场的竞争力。降低动力电池成本、提高技术水平可能是促进新能源汽车发展的关键。

参考文献：

[1]方杰;温忠麟;梁东梅;李霓霓.基于多元回归的调节效应分析[J].心理科学,2015,38(03).
 [2]张宇山.多元线性回归分析的实例研究[J].科技信息,2009,(09):54-56.
 [3]杨华龙;刘金霞;郑斌.灰色预测 GM(1,1)模型的改进及应用[J].数学的实践与认识,2011,41(23):39-46.
 [4]陈陌;郭亚军;于振明.改进型序关系分析法及其应用[J].系统管理学报,2011,20(03):352-355.
 [5]郭金玉;张忠彬;孙庆云.层次分析法的研究与应用[J].中国安全科学学报,2008,(05).