

doi: 10.3969/j.issn.1674-8425(z).2016.08.025

# 基于多元回归模型的公路客运量预测分析

鲁 亚

(天津大学 管理与经济学部, 天津 300072)

**摘 要:** 基于多元线性回归理论选取我国1993—2012年间的公路客运量等数据, 分析国内生产总值和人口总数对公路客运量的影响程度, 并对我国公路客运量进行中短期预测。结果表明: 国内生产总值和人口总数对公路客运量有显著的正相关关系, 且多元回归模型的预测精度很高, 适合进行公路客运量的中短期预测。

**关 键 词:** 公路; 客运量; 多元回归模型; 预测

**中图分类号:** F062

**文献标识码:** A

**文章编号:** 1674-8425(2016)08-0152-04

## Prediction of Traffic Volume based on Multiple Regression Model

LU Ya

(College of Management and Economics, Tianjin University, Tianjin 300072, China)

**Abstract:** Based on the theory of multivariate linear regression, this paper selected highway passenger traffic volume data of 1993-2012 and analysed the influence of GDP and population to highway passenger traffic volume, and forecasted the highway passenger transportation volume of china in the short-term. The results show that the GDP and population have a positive relationship to the highway passenger volume, and the forecast accuracy of the multiple regression models is very high and is suitable for short-term prediction for highway passenger traffic volume.

**Key words:** highway; traffic volume; multiple regression model; prediction

交通运输是国民经济的重要组成部分, 对于国民经济的发展有着至关重要的作用, 是国民经济发展的基础和命脉。公路运输作为交通运输体系的一员, 在交通运输体系中占相当大的比重, 很大程度上直接影响着国民经济的持续健康高效发展。对一个地区而言, 公路运输的发展水平直接体现了当地的经济发展水平, 也直接影响政府的

投资方向和资源配置。改革开放以来, 随着我国经济的突飞猛进, 交通运输也有了翻天覆地的变化, 公路运输更是实现了跨越式发展。不论是从公路客运量还是公路周转量来看, 公路运输已然成为我国交通运输行业的主要客运方式, 在国民经济和社会发展进程中扮演越来越重要的角色。

在这样的大背景下, 科学准确地预测我国客

**收稿日期:** 2015-09-28

**基金项目:** 国家自然科学基金资助项目(18990)

**作者简介:** 鲁亚(1982—), 女, 海南海口人, 硕士研究生, 主要从事管理科学与工程、企业管理研究, E-mail: myluya@163.com。

**引用格式:** 鲁亚. 基于多元回归模型的公路客运量预测分析[J]. 重庆理工大学学报(自然科学), 2016(8): 152-155.

**Citation format:** LU Ya. Prediction of Traffic Volume based on Multiple Regression Model[J]. Journal of Chongqing University of Technology (Natural Science), 2016(8): 152-155.

运量对于我国公路运输业的发展有着十分重要的意义。不仅为政府部门制定交通规划提供参考,也便于分析影响公路客运量的各种因素及其重要性程度,促进我国资源进一步合理配置,加快实现交通运输现代化的进程。在公路客运量的预测方面,国内外学者进行了不少研究,公路客运量的预测模型主要有以下几种:线性回归分析模型、灰色预测模型、自回归积分滑动平均模型等<sup>[1-9]</sup>。其中线性回归分析模型不仅可以在公路客运量的预测方面有较高的精度,而且便于分析影响公路客运量的相关因素及其影响程度,深受国内外学者的青睐。本文基于多元线性回归模型,选取影响公路客运量较显著的2个因素:国内生产总值和人口总数,建立了公路客运量的多元线性回归模型,分析国内生产总值和人口总数对公路客运量的影响程度,并在此基础上预测我国公路客运量。

1 多元线性回归模型

多元线性回归模型是一元线性回归模型的扩展。实际中影响因变量的因素很多,比如公路客运量。影响公路客运量的因素包括国内生产总值、人口总数、公路长度、国内消费水平以及铁路等运输客运量等。此时,需要有2个或2个以上的自变量去解释因变量,与一元线性回归模型相似,多元线性回归模型如下:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n + \varepsilon$$

其中: $\beta_0$ 为常数项; $\beta_1, \beta_2, \cdots, \beta_n$ 为回归系数,意为当其他自变量不变时,某一自变量变化一个单位时因变量的变化值。

建立多元线性回归模型时有如下假设:零均值假定;同方差和无自相关假定;随机扰动项和解释变量不相关;无多重共线性假定;正态性假定。在以上假定的基础上,以残差平方和最小为目标,根据最小二乘法求解多元线性回归模型,这个过程可以通过Eviews、SPSS和SAS等统计软件实现。

在多元线性回归方程中,每个自变量对因变量的解释显著性不同,也需要判断所有自变量联合起来对因变量影响的总显著性,即回归方程的

联合显著性。对回归方程的显著性检验主要包括几个方面:方差检验、 $F$ 检验、可决系数的显著性检验、 $t$ 检验等。

2 我国公路客运量的回归分析

2.1 样本选取

本文选取我国1993—2012年公路客运量数据进行统计分析,在分析影响我国公路客运量的各种因素的基础上,建立了影响我国公路客运量的多元线性回归模型。本文数据源于中华人民共和国国家统计局公布的《2013年中国统计年鉴》。

根据2013年我国统计年鉴数据,1993—2012年我国公路客运量、国内生产总值和人口总数的相关数据如表1所示。

表1 1993—2012年我国公路客运量、国内生产总值和人口总数的相关数据

年份	公路客运量/万人	国内生产总值/亿元	人口总量/万人
1993	860 719	35 333.9	118 517
1994	953 940	48 197.9	119 850
1995	1 040 810	60 793.7	121 121
1996	1 122 110	71 176.6	122 389
1997	1 204 583	78 973.0	123 626
1998	1 257 332	84 402.3	124 761
1999	1 269 004	89 677.1	125 786
2000	1 347 392	99 214.6	126 743
2001	1 402 798	109 655.2	127 627
2002	1 475 257	120 332.7	128 453
2003	1 464 335	135 822.8	129 227
2004	1 624 526	159 878.3	129 988
2005	1 697 381	184 937.4	130 756
2006	1 860 487	216 314.4	131 448
2007	2 050 680	265 810.3	132 129
2008	2 682 114	314 045.4	132 802
2009	2 779 081	340 902.8	133 450
2010	3 052 738	401 512.8	134 091
2011	3 286 220	473 104.0	134 735
2012	3 557 010	518 942.1	135 404

2.2 我国公路客运量的回归分析

王佳雨等通过对影响浙江省公路客运量的诸

多因素进行分析,发现对公路客运量有显著影响的因素主要包括国内生产总值和人口总数等。本文借助 Eviews 等数学工具,建立了公路客运量  $Y$ 、国内生产总值  $X_1$  和人口总数  $X_2$  的多元线性回归模型。

首先,考察影响公路客运量的主要因素的相关系数矩阵,如表 2 所示。

从相关系数矩阵可以看出:公路客运量和国内生产总值和人口总数的相关系数高达 0.994 447 862 0 和 0.887 213 433 6,说明公路客运量和国内生产总值及人口总数有着显著的线性

关系,可以建立如下二元线性回归模型:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \varepsilon$$

借助 Eviews 建立多元线性回归模型,作最小二乘回归分析,结果如表 3 所示。

表 2 影响公路客运量主要因素的相关系数矩阵

变量	$Y$	$X_1$	$X_2$
$Y$	1	0.994 447 862 0	0.887 213 433 6
$X_1$	0.994 447 862 0	1	0.881 306 582 3
$X_2$	0.887 213 433 6	0.881 306 582 3	1

表 3 最小二乘回归分析结果

Dependent Variable: $Y$				
Sample: 1993 2012				
Included observations: 20				
	Coefficient	Std. Error	$t$ -Statistic	Prob.
$C$	-206 894. 3	1 030 777	-0. 200 717	0. 843 3
$X_1$	5. 333 113	0. 295 397	18. 054 06	0. 000 0
$X_2$	7. 730 465	8. 426 552	0. 917 394	0. 371 8
R-squared	0. 989 449	Mean dependent var		1 799 426
Adjusted R-squared	0. 988 208	S. D. dependent var		822 560. 7
S. E. of regression	89 324. 21	Akaike info criterion		25. 775 41
Sum squared resid	1. 36E + 11	Schwarz criterion		25. 924 77
Log likelihood	-254. 754 1	Hannan-Quinn criter		25. 804 57
F-statistic	797. 103 1	Durbin-Watson stat		1. 400 613
Prob( F-statistic)	0. 000 000			

因此,公路客运量、国内生产总值和人口总数的多元回归模型为:

$$Y = -206\,894.3 + 5.333\,113X_1 + 7.730\,465X_2$$
$$(1\,030\,777) \quad (0.295\,397) \quad (8.426\,552)$$
$$t = (-0.200\,717) \quad (18.054\,06) \quad (0.917\,394)$$
$$R^2 = 0.989\,449, \bar{R}^2 = 0.988\,208$$

其中: $Y$ 为公路客运量; $X_1$ 为国内生产总值; $X_2$ 为人口总数。

2.3 模型检验

1) 经济意义检验

模型结果表明:在国内人数不变的情况下,国

内生产总值增加 1 亿元,公路客运量将增加 5.333 113 万人;在国内生产总值不变的情况下,人口总数增加 1 万人,公路客运量将增加 7.730 456 万人。换言之,公路客运量与国内生产总值、人口总数正相关。

2) 统计检验

从回归估计结果可见:模型拟合度很高,可决系数为 0.989 449,表明公路路客运量的98.944 9%可由国内生产总值和人口总数来解释。

从回归模型的  $F$  检验值来看, $F$  统计量的值为 797.103 1,Prob 统计量的值为 0.000 0,表明模

型呈现显著的线性回归趋势。

从回归模型的  $t$  检验值来看:国内生产总值和人口总数的  $t$  统计量的值分别为 18.054 06 和 0.917 394。可以看出:国内生产总值和人口总数对公路客运量均有显著的影响。

### 3) 精度检验

根据公路客运量、国内生产总值和人口总数的多元回归模型可以得出公路客运量的拟合值,如表 4 所示。

表 4 公路客运量拟合值

年份	公路客运量/万人	公路客运量预测值/万人	相对误差/%
1993	860 719	897 736.901 84	4.30
1994	953 940	976 646.777 31	2.38
1995	1 040 810	1 053 647.023 1	1.23
1996	1 122 110	1 118 822.431 6	-0.29
1997	1 204 583	1 169 964.099 0	-2.87
1998	1 257 332	1 207 693.247 2	-3.95
1999	1 269 004	1 243 748.078 3	-1.99
2000	1 347 392	1 302 010.698 5	-3.37
2001	1 402 798	1 364 525.329 2	-2.73
2002	1 475 257	1 427 855.007 3	-3.21
2003	1 464 335	1 516 448.840 9	3.56
2004	1 624 526	1 650 622.424 6	1.61
2005	1 697 381	1 790 202.433 7	5.47
2006	1 860 487	1 962 889.002	5.50
2007	2 050 680	2 232 120.676 4	8.85
2008	2 682 114	2 494 566.518 3	-6.99
2009	2 779 081	2 642 809.408 7	-4.90
2010	3 052 738	2 971 004.615 7	-2.68
2011	3 286 220	3 357 786.994 5	2.18
2012	3 557 010	3 607 418.442 6	1.42

相对误差是检验模型预测精度的重要指标,也是评价模型优劣的标准。可以看出:模型的预测精度很高,可以用来进行我国公路客运量的中短期预测。

## 3 结束语

本文建立了公路客运量、国内生产总值和人口总数的多元线性回归模型,结果表明:国内生产总值和人口总量对公路客运量显著正相关,模型精度很高;同时,模型的预测精度很高,公路客运量的预测相对误差很小,适合进行公路客运量的中短期预测。

同时,国内生产总值和人口总数对公路客运量的影响非常显著,随着国内生产总值和人口总数的增加,我国公路客运量也相应增加。我国公路客运量保持着较高的增长态势,政府部门应当做好公路交通规划,保证人民群众的出行需求,实现经济的稳步快速发展。

## 参考文献:

- [1] 王生昌,白韶波,张慧.公路客运量预测方法的比较[J].长安大学学报(自然科学版),2005(5):83-85.
- [2] 霍娅敏,李德刚.成都市公路客运量预测[J].交通标准化,2005(11):161-164.
- [3] 高红丽,霍娅敏,李德刚.高速公路客运量预测[J].交通信息与安全,2003(4):9-11.
- [4] 乔向明.2003—2005 年我国公路客运量预测分析[J].山东交通学院学报,2003(1):26-29.
- [5] 陈文华,边浩毅,王怡民等.基于 ARIMA 模型的公路客运量预测[J].重庆工商大学学报(西部论坛),2006(5):95-97.
- [6] 朱文亭,刘海林.基于灰色-马尔科夫链理论的公路客运量预测[J].交通科技与经济,2009(6):12-13.
- [7] 李晓刚,贾元华,教谷昌.基于主成分回归的公路客运量预测模型研究[J].交通标准化,2009(9):77-81.
- [8] 纪跃芝,冯延辉.应用灰色系统 GM(1,1)模型预测长春市公路客运量[J].长春工业大学学报(自然科学版),1997(2):52-56.
- [9] 彭岩,杨卓毅.基于改进 GM(1,1)模型的天津市电力消费预测研究[J].重庆理工大学学报(自然科学),2015,29(1):131-134.