

Investigation and Prospect Forecast of Bicycle Sharing Based on Statistical Modeling

Weiwei Yan¹, Feng Hu²

¹Jining Haida Xingzhi Middle School, Jining Shandong

²School of Statistics, Qufu University, Jining Shandong

Email: 1490164087@qq.com, hufengqf@163.com

Received: Dec. 27th, 2018; accepted: Jan. 9th, 2019; published: Jan. 16th, 2019

Abstract

As we enter the new era, bicycle sharing has become synonymous with green travel and easy health. However, any new type of business activity always comes with chaos in the beginning such as barbarous growth and disordered management. Bicycle sharing is no exception. Situations such as disordered placement and severe damage of bicycles are troubling the users. There are many factors affecting the satisfaction of users. This paper studies the basic characteristics of users and the objective factors that affect user satisfaction. Firstly, the independent test and the lining analysis are applied to the basic characteristics of the user, and the exploratory factor analysis and the confirmatory factor analysis are applied to the influencing factors. Secondly, the Logistic model is used to predict from the two aspects of the user's basic characteristics and influencing factors. The results show that there is a significant relationship between the basic characteristics of users and user satisfaction. Safety hazards and green environmental protection have significant effects on user satisfaction.

Keywords

Bicycle Sharing, Contingency Analysis, Path Analysis, Logistic Regression Model

基于统计建模下共享单车的调查分析与前景预测

严围围¹, 胡 锋²

¹济宁海达行知中学, 山东 济宁

²曲阜师范大学统计学院, 山东 济宁

Email: 1490164087@qq.com, hufengqf@163.com

收稿日期: 2018年12月27日; 录用日期: 2019年1月9日; 发布日期: 2019年1月16日

摘要

随着我们进入新时代, 共享单车成了绿色出行、易于健康的代名词, 但随着时间的推移, 车辆过度投放、管理无序等问题引起人们担忧; 影响使用者满意度的因素众多, 本文从使用者基本特征和影响使用者满意度的客观因素两个角度展开研究。首先针对使用者基本特征运用独立性检验与列联分析、针对影响因素运用探索性因子分析和验证性因子分析; 其次从使用者基本特征和影响因素两个角度利用Logistic模型进行预测。结果表明: 使用者基本特征与使用者满意度存在显著性关系, 安全隐患、绿色环保对使用者满意度存在显著性影响。

关键词

共享单车, 列联分析, 路径分析, Logistic模型

Copyright © 2019 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

共享单车、网购、支付宝、高铁成为中国“新四大发明”的代名词, 共享单车搭着“互联网+”的顺风车飞速发展。目前我国共有互联网租赁自行车运营企业 30 多家, 累计投放车辆超过 1000 万辆, 累计服务超过十亿人次(可见参考文献[1])。共享单车的生存现状与未来的发展趋势为研究热点。有桩的公共自行车的研究方面, 研究者们从存在的影响因素方面分析了使用者使用公共自行车的满意度, 但对于共享单车使用者基本特征是否影响满意度、共享单车未来的一个发展趋势? 这方面研究较少。本文针对使用者基本特征, 利用对应分析探究了性别、年龄、职业、使用目的与使用者认同度的列联分析; 针对影响共享单车的 23 个因素采用李克特量表将其表示, 运用探索性因子分析和验证性因子分析对其进行分析, 因子分析进行降维提取主要影响因素, 路径模型分析不同影响因素间的联系; 最后, 从提取的五大影响因素、使用者基本特征两个不同角度建立 Logistic 模型对共享单车前景预测。结果表明: 使用者基本特征与使用者满意度存在显著性关系, 安全隐患、绿色环保对使用者满意度存在显著性影响。

2. 测量工具与模型

2.1. 李克特量表

利用李克特量表调研使用者的满意度。该量表采用 5 级李克特量表(以“完全不认同、不认同、不确定、基本认同、非常认同”分别由低到高赋 1、2、3、4、5 分)。通过灵活便捷、安全隐患、绿色环保、制度风险、性能舒适五个维度来测量使用者满意度。具体如表 1 所示。

2.2. 结构方程模型

结构方程模型通常采用路径图来表达, □表示显变量, ○表示潜变量, →表示变量间的影响关

系; 箭头开始处为潜变量, 末端为显变量; δ 、 ε 表示变量 x 、 y 测量误差, λ 表示潜变量与显变量间的因子载荷。关于结构方程模型的更多细节可见参考文献[2]。

Table 1. Five latent factors
表 1. 五大潜在因子

潜在因子	含义	影响因素数量
灵活便捷	扫码取车还车方便、意愿选择路途与途径	4
安全隐患	维修不及时, 存在安全隐患、受天气影响大	5
绿色环保	减少噪声、废气污染、绿色环保、低碳出行	4
制度风险	收费制度不合理、押金制度不合理	4
性能舒适	车身性能好、设计合理、舒适度高	2

3. 数据分析与处理

3.1. 问卷组成

调查问卷由三部分:

- 1) 使用者的基本特征: 性别、年龄、职业等;
- 2) 使用者的用车特征: 使用频率、用车用途、用车期望等;
- 3) 客观影响因素: 由 23 个题项组成的研究变量。

3.2. 数据清洗与检验

运用 SPSS22.0 分析调查数据, 利用箱线图进行剔除离群值。剔除数据为: 所用时间不合理: 658 s、640 s、46 s、724 s、337 s、305 s、450 s、62 s; 得分不合理: 49 分、78 分、69 分、78 分、122 分、77 分、78 分、35 分、34 分。

正式调查共发放有效问卷 300 份, 收回有效问卷 212 份, 有效率达 70.6%; 有效样本量大于 10 倍量表的题项数, 可进行正式实证分析。利用 SPSS22.0 软件对量表题项进行效度和信度检验; 运行结果: 克隆巴哈系数为 0.901, KMO 的度量为 0.857, Bartlett 的球形度检验值为 2435.171, P 值等于 0.000, 即量表效度、信度较好。

4. 模型建立与实证分析

4.1. 使用者认同度的列联分析

采用问卷第三部分量表处的调查数据作为模型的研究数据; 以 19 个影响因素作为模型的解释变量, 用 $x_i (i = 1, 2, 3, \dots, 19)$ 表示, 探究对使用者认同度的影响程度。采用问卷第一部分使用者的基本特征及所提取的五个主成分为研究数据, 从两个不同角度利用 Logistic 回归模型来预测共享单车的未来发展前景。具体使用者的基本特征见表 2。

针对使用者基本特征进行对应分析及独立性检验, 利用 SPSS22.0 运算结果如下:

利用 SPSS22.0 软件分析得表 3。结果表明: 性别、年龄、职业、使用目的均与使用者认同度呈显著性关系。

4.2. 探索性因子分析

影响使用者对共享单车认同度的 19 个显变量以特征值大于 1 为提取标准进行主成分分析, 运行结果如表 4 所示, 前五个主成分的累计贡献率达到了 70.80%, 选取五个主成分。

Table 2. Basic feature type of users
表 2. 使用者基本特征类型

变量	数据类型	分类情况
满意度(P)	定序三水平(P_1, P_2, P_3)	不满意、较满意、满意
性别(X_1)	定类两水平(X_{11}, X_{12})	男、女
年龄(X_2)	定序四水平($X_{21}, X_{22}, X_{23}, X_{24}$)	≤ 18 、 $19\sim 27$ 、 $28\sim 45$ 、 ≥ 45
职业(X_3)	定类五水平($X_{31}, X_{32}, X_{33}, X_{34}, X_{35}$)	在校学生、政府事业人员、企业人员、个体经营者、其他

Table 3. Basic feature test value
表 3. 基本特征检验值

类别	χ^2	df	P
性别	19.962	4	0.001
年龄	35.313	12	0.000
职业	36.656	16	0.002
使用目的	51.482	20	0.000

Table 4. Principal component cumulative contribution rate
表 4. 主成分累计贡献率

Importance of components	Comp.1	Comp.2	Comp.3	Comp.4	Comp.5
Proportion of variance	17.48	16.88	14.94	11.57	9.93
Cumulative proportion	17.48	34.36	49.30	60.87	70.80

利用方差最大化对因子进行旋转, 得到表 5, 将五个公共因子命名为: 灵活便捷、安全隐患、绿色环保、制度风险、性能舒适; 分析五大因子发现, 前三个因子占比依次为: 17.48%, 16.88%, 14.94%, 即使用者在使用共享单车的过程中看重共享单车的灵活便捷、安全隐患、绿色环保。

Table 5. Factor load map
表 5. 旋转后因子载荷图

显变量	潜变量	
车锁设计合理	0.730	
扫码取车还车方便	0.754	
意愿选择路途与途径	0.744	灵活便捷
取放方便, 随停随取	0.709	
受天气影响大	0.668	
无年龄限制, 儿童骑车危险	0.711	
维修不及时, 存在安全隐患	0.707	安全隐患
停放随意, 不易找车	0.647	
散落城市, 影响市容	0.642	
使用对象制度有待提高		0.594
骑车易于健康		0.762
减少噪声、废气污染		0.758
绿色环保、低碳出行		0.735

Continued

押金制度不合理	0.791	
收费制度不合理	0.795	
实名注册, 存在安全隐患	0.696	制度风险
扫码可能泄露个人信息	0.574	
车身性能好	0.809	性能舒适
设计合理、舒适度高	0.877	

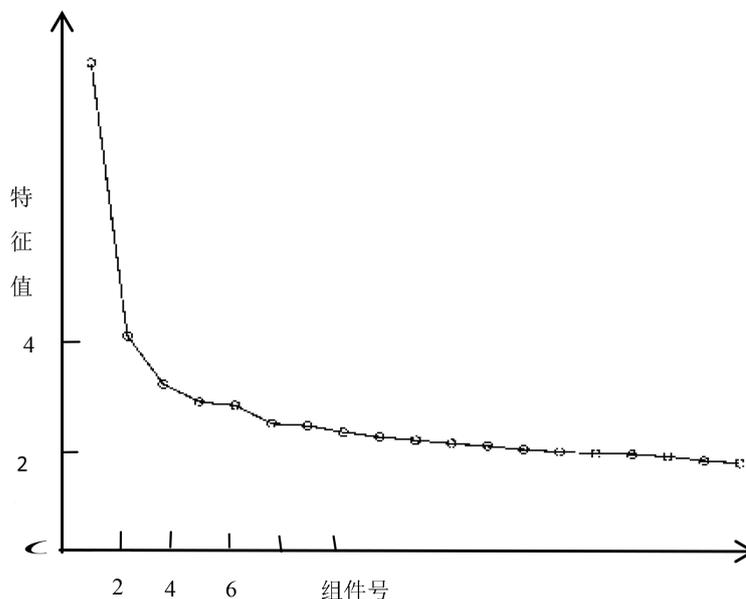


Figure 1. Gravel map
图 1. 各因子对应碎石图

通过图 1 知: 碎石图显示: 第一个因子的特征值很高, 对解释变量原有的贡献最大; 第五个因子以后的特征值对解释原有变量的贡献率很小, 基本可以忽略, 因此提取五个因子是较为合适的。因子得分模型为:

$$F_1 = -0.049x_1 - 0.086x_2 + 0.297x_3 + 0.296x_4 + 0.275x_5 + 0.272x_6 + 0.020x_7 \\ - 0.121x_8 - 0.077x_9 - 0.013x_{10} + 0.063x_{11} + 0.095x_{12} + 0.130x_{13} - 0.018x_{14} \\ + 0.045x_{15} + 0.034x_{16} + 0.030x_{17} - 0.206x_{18} - 0.233x_{19}$$

$$F_2 = -0.065x_1 - 0.050x_2 + 0.017x_3 - 0.108x_4 - 0.006x_5 + 0.046x_6 - 0.199x_7 \\ - 0.001x_8 - 0.137x_9 + 0.125x_{10} + 0.063x_{11} + 0.322x_{12} + 0.289x_{13} + 0.266x_{14} \\ - 0.123x_{15} - 0.096x_{16} - 0.095x_{17} + 0.209x_{18} + 0.230x_{19}$$

$$F_3 = -0.047x_1 - 0.077x_2 - 0.085x_3 - 0.021x_4 - 0.063x_5 + 0.272x_6 - 0.112x_7 \\ - 0.290x_8 + 0.012x_9 - 0.178x_{10} - 0.155x_{11} - 0.163x_{12} - 0.140x_{13} - 0.040x_{14} \\ + 0.340x_{15} + 0.325x_{16} + 0.323x_{17} + 0.132x_{18} + 0.167x_{19}$$

$$F_4 = 0.007x_1 - 0.003x_2 - 0.077x_3 + 0.013x_4 + 0.016x_5 + 0.026x_6 + 0.436x_7 \\ + 0.077x_8 + 0.430x_9 + 0.306x_{10} + 0.193x_{11} - 0.062x_{12} - 0.070x_{13} - 0.093x_{14} \\ - 0.009x_{15} - 0.015x_{16} - 0.027x_{17} - 0.092x_{18} - 0.108x_{19}$$

$$F_5 = 0.514x_1 + 0.585x_2 - 0.076x_3 - 0.106x_4 - 0.054x_5 + 0.107x_6 - 0.096x_7 - 0.083x_8 + 0.094x_9 + 0.047x_{10} - 0.138x_{11} - 0.111x_{12} - 0.070x_{13} + 0.154x_{14} - 0.062x_{15} - 0.049x_{16} - 0.052x_{17} + 0.139x_{18} + 0.047x_{19}$$

以五个公共因子为自变量，建立多元线性回归模型。运用 SPSS22.0 软件运行结果见表 6。

Table 6. Multiple linear regression parameter
表 6. 多元线性回归参数

Model	4652.408	5	9304.82	F(5,206) = 151.260	
Residual	12,672.20	206	61.52	Prob > F = 0.000	
				R-squared = 0.887	
Total	59,196.28	211	9366.34	Adj R-squared = 0.781	
				Root MSE = 2.457	
Pre	Coef.	Std. Err.	T	P > t	Corr
Cont	90.203	0.539	167.454	0.000	
X1	7.091	0.540	13.133	0.000	0.423
X2	8.730	0.540	16.169	0.000	0.521
X3	6.228	0.540	11.535	0.000	0.372
X4	5.674	0.540	10.508	0.000	0.339
X5	4.797	0.540	8.884	0.000	0.286

由表 6，性能舒适对使用者认同度影响较小。多元线性回归表达式如下：

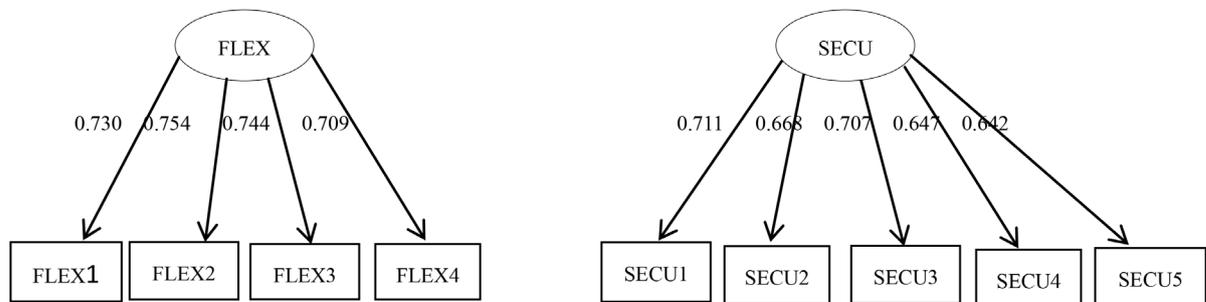
$$y = 90.203 + 7.091x_1 + 8.730x_2 + 6.228x_3 + 5.674x_4 + 4.797x_5 \tag{2}$$

该模型表示：对认同度影响最大的是安全隐患。结果表明：居民在确保安全的前提下，才会使用共享单车满足自己的使用需求；因此，共享单车若要有个较好的前景，应该从单车性能、制度保障、维修维护等方面加大治理力度。

4.3. 路径分析

以往学者针对影响因素这方面主要以结构方程为主，从建立假设到检验验证假设；而本文则仅仅利用路径模型图对上文因子分析提取的主成分进行验证。五个主成分能否较好的代表 19 个显性变量去描述潜变量使用者的认同度还有待检验。因此，进一步建立认同度之间的路径模型，来对其关系进行验证。

利用 Amos21.0 软件画出模型的初步路径图，如图所示：



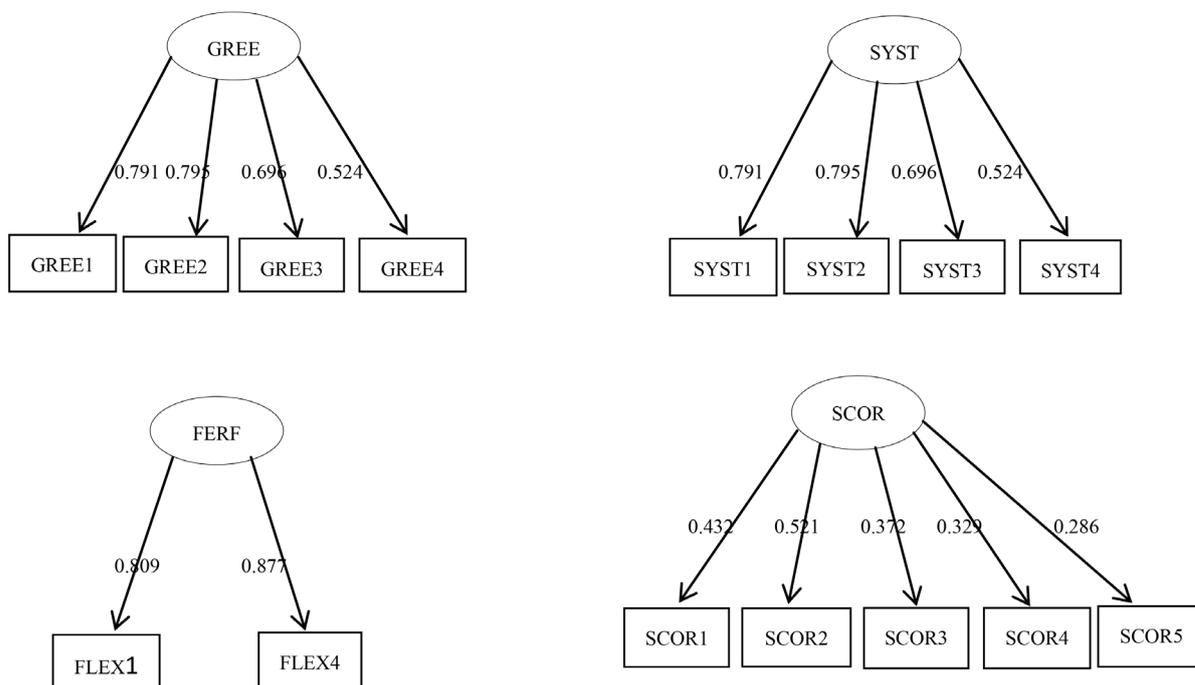
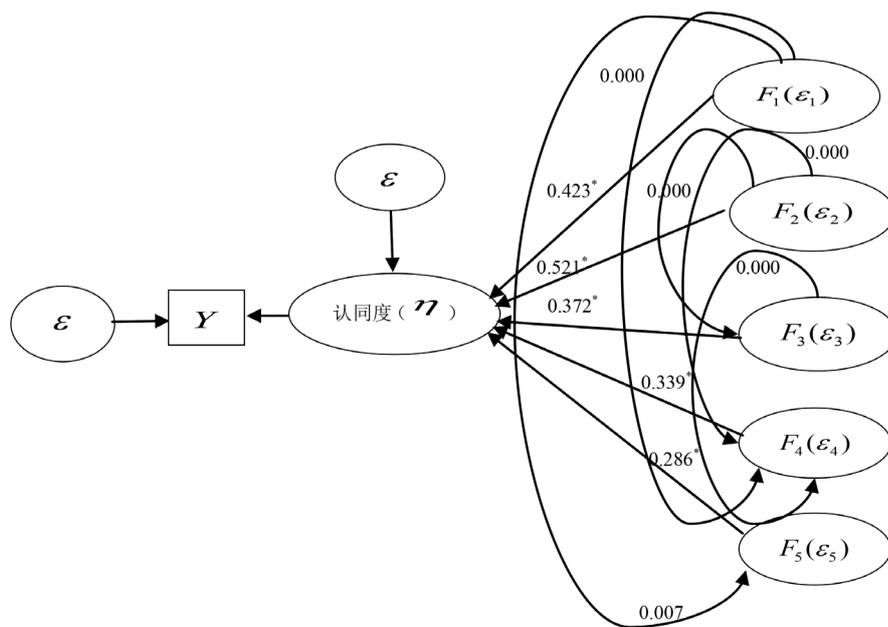


Figure 2. Measurement modeling
图 2. 测量模型

通过图 2, 各显变量与潜变量之间的路径系数大于 0.5。即; 灵活便捷、安全隐患、绿色环保、制度风险、性能舒适与使用者的认同度呈现正向关。(注: FLEX: 灵活便捷, SECU: 安全隐患, GREE: 绿色环保, SYST: 制度风险, PERF: 性能舒适, SC: 总得分)。



注: **表示模型在 0.01 水平下显著。

Figure 3. Structural Equation Modeling
图 3. 结构模型

由图 3, 灵活便捷与制度风险、性能舒适的直接关系分别为 0.000、0.000; 安全隐患与绿色环保的直接关系为 0.000。即以上两两之间正向的关系较弱; 其他两两之间不存在正向的关系, 五个主成分之间是不相关的; 其检验结果和建立结构模型前假设潜变量之间的协方差为零也是相吻合的; 而所提取的五个主成分与认同度之间的路径系数分别为: 0.423、0.521、0.372、0.339、0.286, 即潜变量间呈正相关的关系, 能很好的解释影响认同度的原因。

4.4. Logit 模型构建——前景预测

共享单车是为了解决居民出行最后一公里的问题产生的, 能够满足居民需求、迎合居民需求才是王道; 因此本部分利用提取的五个主成分和使用者基本特征对使用者的满意度构建多项 logistic 模型, 对共享单车的未来的发展前景进行简单预测。

Table 7. Model structure

表 7. Logistic 模型结果

Logistic regression				Number of obs = 212		
-2Log likelihood = 388.418				LRchi2(20) = 130.607		
				Prob > chi2 = 0.000	Pseudo = 0.460	
Pre	Coef.	Std.Err.	Wald	P > z	[95% Conf. Interval]	
Cont	3.330	0.787	17.910	0.000		
F1	1.829	0.429	18.200	0.000	6.229	2.688
F2	1.856	0.530	12.260	0.000	6.396	2.264
F3	2.313	0.505	20.949	0.000	10.102	3.752
F4	0.187	0.442	0.178	0.673	1.205	0.507
F5	1.933	0.442	19.174	0.000	6.913	2.910

以完全认同为被解释变量, 以完全不认同为哑变量, 以灵活便捷(F1)、安全隐患(F2)、绿色环保(F3)、制度风险(F4)、性能舒适(F5)为自变量建立的 Logistic 模型, 运行结果显示: 持完全认同观点的人群为 66, 占比 31.3%, 具体见表 7。模型建立如下:

$$\text{Logit}P_5 = \frac{\exp(3.330 + 1.829F_1 + 1.856F_2 + 2.313F_3 + 1.933F_4)}{1 + \exp(3.330 + 1.829F_1 + 1.856F_2 + 2.313F_3 + 1.933F_4)} \quad (3)$$

$$\ln\left(\frac{P(y = 5|x)}{P(y = 1|x)}\right) = 3.330 + 1.829F_1 + 1.856F_2 + 2.313F_3 + 1.933F_4 \quad (4)$$

结果表明, 绿色环保对共享单车前景的影响最大, 其次是性能舒适、安全隐患、灵活便捷。因此, 在未来共享单车发展的道路上, 应注重共享单车车身的性能提高, 同时应加大管理力度, 确保在安全隐患方面赢得使用者信赖。

以基本认同为被解释变量, 以完全不认同为哑变量, 以灵活便捷(F1)、安全隐患(F2)、绿色环保(F3)、制度风险(F4)、性能舒适(F5)为自变量建立的 Logistic 模型, 运行结果显示: 持基本认同观点的人群为 105, 占比 49.5%, 具体见表 8。模型建立如下:

$$\text{Logit}P_4 = \frac{\exp(4.265 + 0.758F_1 + 1.619F_2 + 2.313F_3 + 1.933F_4)}{1 + \exp(4.265 + 0.758F_1 + 1.619F_2 + 2.313F_3 + 1.933F_4)} \quad (5)$$

$$\ln\left(\frac{P(y=4|x)}{P(y=1|x)}\right) = 4.265 + 0.758F_1 + 1.619F_2 + 2.313F_3 + 1.933F_4 \quad (6)$$

结果表明, 绿色环保对共享单车前景的影响最大, 其次是性能舒适、安全隐患、灵活便捷。

Table 8. Model structure
表 8. Logistic 模型结果

Logistic regression Number of obs = 212 LRchi2(20) = 130.607						
	-2Log likelihood = 388.418		Prob > chi2 = 0.000		Pseudo = 0.460	
Pre	Coef.	Std. Err.	Wald	P > z	[95% Conf. Interval]	
Cont	4.265	0.769	30.761	0.000		
F1	0.758	0.367	4.276	0.039	1.040	4.377
F2	1.619	0.510	10.057	0.002	1.856	13.727
F3	1.326	0.448	8.770	0.003	1.566	9.060
F4	0.104	0.427	0.059	0.808	0.480	2.561
F5	1.740	0.428	16.493	0.000	2.460	13.186

4.5. 结论与建议

通过模型的建立及结果显示: 使用者基本特征与认同度之间存在显著性差异, 共享单车在划分停车布局时, 应该充分考虑周围场所的分布; 其次, 对 19 个客观影响因素进行探索性因子分析、验证性因子分析, 建立多元线性回归模型, 五个影响因素中, 安全隐患居于首位。因此企业应该重视共享单车质量、加大管理维修力度, 成立维修管理机构, 针对个别使用者不文明行为建立信用机制进行约束。最后, 建立 Logistic 回归方程对共享单车未来发展趋势进行一个预测, 进行相关系数的解释, 发现绿色环保、性能舒适是影响使用者满意度的较大影响因素, 该部分较蒋丽芹、吴宝泰[3]和崔梦蕾[4]的研究适当进行了创新。因此, 新时代促进共享事物的可持续发展应该提高性价比、环保比。探索性因子分析提取主成分, 本文结合实际情况选取 23 个影响因素, 利用因子分析提前五个主成分, 累计贡献率为 70.8%; 在蒋丽芹、吴宝泰[3]中利用因子分析法提取主成分, 累计贡献率为 64.378%。

致 谢

感谢审稿人提出的建议。

基金项目

山东省自然科学基金(ZR2017MA012), 山东省高等学校科技计划项目(J18LI07)。

参考文献

- [1] 共享单车累计投放超 1000 万辆管理新政将陆续出台——中新网[EB/OL]. <https://www.chinanews.com/cj/2017/05-24/8232308.shtml>, 2017-05-24.
- [2] 吴明隆. 结构方程模型: AMOS 的操作与应用[M]. 重庆: 重庆大学出版社, 2009.
- [3] 蒋丽芹, 吴宝泰. 基于因子分析法的城市公共自行车系统公众满意度研究——以江苏无锡为例[J]. 哈尔滨商业大学学报: 社会科学版, 2014(4): 73-80.
- [4] 崔梦蕾. 基于顾客满意度的城市居民对公共自行车出行分析——以武汉市为例[J]. 现代商贸工业, 2011(9): 125-126.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2325-2251，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：sa@hanspub.org