

2015 年(第四届)全国大学生统计建模 大赛参赛论文

论文名称：打车软件影响下的西安市出租车
运营市场研究和统计分析¹

参赛学校：西安理工大学

参赛队员：刘二嫚 余 菲 王赵汉

指导教师：王金霞 肖燕婷

提交日期：2015 年 6 月 29 日

¹注:该论文获得由中国统计教育学会举办的“2015 年(第四届)全国大学生统计建模大赛”市场调查分析类本科生组三等奖。

摘 要

近几年滴滴和快的等打车软件在一些一线城市比较流行,导致正常出租车运营受到影响,且随着“烧钱”大战,打车软件引起了人们的广泛关注。为研究西安市出租车市场受打车软件的影响度,调查小组查阅相关文献资料并设计调查问卷,于2015年5月随机抽取西安市8个区展开“西安市居民乘坐出租车情况调查”的问卷调查。对调查数据采用相关性分析、主成分分析,且建立多元线性回归模型,可得:西安市出租车数量近几年呈上升趋势但幅度不大,并预测2016年出租车保有量为12510辆;居民倾向于打车软件的便利性、额外经济补贴、打车方便等特性,满意度较高;西安市居民的性别、年龄、职业对出行方式的选择具有一定影响:男性、青年、职员更倾向于乘坐出租车;通过相关性的分析以及多元回归模型的建立,得到居民满意度与出租车乘坐价格、诚信度、可靠性、安全性等因素相关性较高,且西安市居民大多在有突发事件、乘车高峰期、地理位置偏僻等情况下使用打车软件,弥补了传统出租车的不足。最后通过分析打车软件的利弊,对出租车市场管理提出了相关建议。

关键词:打车软件,出租车市场,满意度,相关性分析,多元线性回归模型

一、选题背景

打车软件是一种智能手机应用,乘客可以便捷地通过手机发布打车信息,并立即和抢单司机直接沟通,大大提高了打车效率。

2013年,上海出租行业兴起一款“手机打车软件”,用户在网上下载软件后,输入起点和目的地,自愿选择“是否支付小费”,出租车司机则可根据线路、是否有小费等选择接受订单。调查发现,这种拼小费竞价打车的模式引来了不少质疑声,认为这是变相涨价,并使行业监管出现“灰色地带”。手机打车软件的产生,例如:滴滴、快的、Ube等,且在短时间内数量疯长。打车软件的这种新兴打车模式给传统的出租车市场带来了一场变革,打破了传统的扬招和电话预订车,也打破了传统出租车市场的垄断地位。

截止到2013年5月7日,安卓平台上11家主流应用商店的打车类软件客户端总体下载量已超过百万,用户主要集中在北上广等一线城市。由于出租车司机与打车者之间信息不对称,导致非高峰时段出租车空载、高峰期和恶劣天气下司机拒载等现象频发,而手机打车软件通过加价等手段,提高了打车成功几率,实现了司机和打车者双赢,因而在大城市日益走俏。

打车软件虽然火,但是问题也存在很多。滴滴打车和快的打车近日“疯狂烧钱”^[1]推广,引发全民打车热潮。打车软件在提高资源利用效率的同时,也带来了老人打车难、司机挑活增多、蓝牌车运营更加隐蔽等系列问题。市民对于打车软件既爱又恨。打车软件在为市民提供方便的同时,也带来了一些问题。

其一,影响交通驾驶安全。营运过程中司机边开车,边盯手机屏幕,准备随时抢单,注意力分散,存在严重安全隐患。有出租车企业反映,近日出租车剐蹭事故有上升趋势;

其二,影响行业营运秩序。由于软件对驾驶员缺乏严格的认证机制,让黑车和假的有机可乘,扰乱出租车市场秩序,同时也威胁乘客的人身财产安全;

其三,影响行业服务质量。由于一些软件自动向司机端推送订单,其推送范围取决于在一定召车区域内装载客户端的司机数目,而并非该区域所有出租车,

实际上难以实现车辆资源的最有效调配。同时挑客拒载等现象增多,近日,交委信息中心统计的出租车行业投诉总量同比增长 90%,其中拒载类投诉同比增加超过 50%。此外,支付方式等服务质量纠纷增多,信息中心接到支付类投诉,已占到其投诉总量的五成以上。

同时,虽然多款打车软件已经接入官方统一电召平台,但在实际运营中,存在“两张皮”的情况:只有部分订单向统一平台发送,部分订单则自行向司机手机客户端推送,绕开监管。

面对打车软件存在的益处与弊端,它对出租车市场是如何影响的呢?基于此,团队选取了所在城市—西安市进行问卷调查,来研究打车软件影响下的西安市出租车市场如何变化,并提出相关建议与意见。

二、文献综述

由于近几年滴滴和快的等打车软件在一些一线城市比较流行,使得正常出租车运营受到影响,并且引起人们的广泛关注。从最初的对居民出行方式调查,发展到出租车数量对城市规划的影响,均提出了针对性的政策建议体系。但是已有的研究数据对目前的出租车发展规模形式仍然十分局限,众多学者只分析调查了问卷中的相关指标,缺少宏观把握。

为了解市民对坐出租车的出行方式,吉波^[2]等学者先对出租车行业现状需求影响因素进行分析;再对大城市出租车需求提出对策;最后对经营模式的种类、各类经营模式的利弊以及经营效率的比较,并且以广州市内的出租车运营方式为例,对以上三方面进行检验。

其后王克^[3]所进行的论文,先是对城市的出租车保有量的影响因素并分析,明确了出租车与城市其他客运方式的竞争关系;之后根据出租车的保有量,利用系统动力学法,建立了城市出租车系统保有量的预测模型;最后以大连市核心区为实例,通过对大连市的系统分析并且收集相关数据,对大连市主城区保有量进行求解,并对未来出租车保有量进行了预测。

三、研究假设

- 1、假设抽样的人群、街道具有代表性。

- 2、假设统计数据的微小误差对结果影响不大可以忽略。
- 3、假设司机的驾驶技术、经验相当，都能达到中等以上水平。
- 4、假设在研究时间内，西安市出租车市场没有巨大变化。
- 5、假设小部分缺失的数据对统计分析结果没有影响。

四、建模过程

（一）西安市出租车保有量分析及预测

在进行打车软件影响分析之前，通过查阅西安市出租车管理处官网，得到 5 个年份西安市出租车保有量，数据如下表 1：

表 1 近几年西安市出租车保有量					
年份	2006 年	2009 年	2011 年	2014 年	2015 年
保有量	11177	11879	11720	12115	12180

对西安市出租车保有量数据进行相关性分析，得到近几年西安市出租车数量变化趋势图，如下图。由图可以看出近几年西安市出租车数量整体上随年份增加，使用 Excel 软件对现有出租车保有量进行拟合，得出指数模型匹配度最高，相对误差最小。根据当前拟合趋势，预计 2016 年西安市出租车保有量将达到 12510 辆。

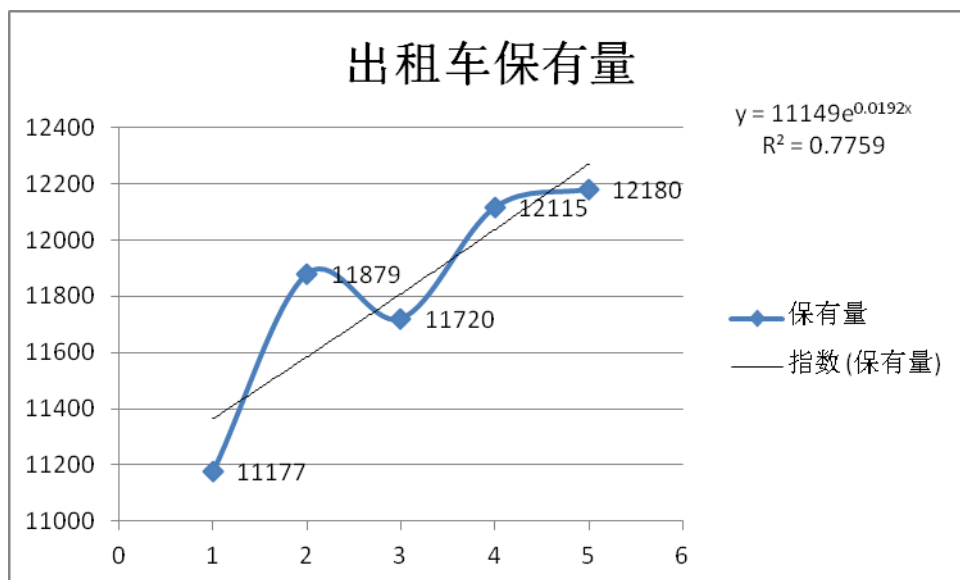


图 1 2016 年西安市汽车保有量预测

（二）西安市居民对打车软件满意度及其相关性分析

通过现状分析得出在打车软件影响下，主要影响顾客使用打车软件的指标，

总结出影响出租车行业顾客满意度的因素，提出策略并改进。

1. 打车软件满意度分析的研究

(1) 前期: 问卷调查准备

研究对象：西安市居民及出租车司机

样本量：主要由调查目的、调查性质、精度要求等因素决定，同时考虑实际情况下的可实施性。本调查数据以百分比形式呈现，在确定样本量时不考虑总体，且允许抽样误差为 $E = 6\%$ ，在 95% 的置信水平下 (置信度 $Z_{\frac{\alpha}{2}} = 1.96$)，样本变异程度 $p = 0.5$ ，由公式：

$$n = \frac{Z_{\frac{\alpha}{2}}^2 p (1 - p)}{E^2} \tag{1}$$

得出 $n = 267$ ，考虑实际情况的可实施性，取 $n = 300$ 份。

问卷准备情况：团队根据调查目的自制问卷调查表 (即《打车软件影响下的西安市居民满意度调查》)，对西安市居民展开问卷调查，用比例分配法、规模分配法进行抽样。

第一阶段：采用整群抽样在西安市 9 个区中选取 6 个区 (新城区、碑林区、莲湖区、灞桥区、未央区、雁塔区)，运用比例分配法将样本量 (300 份调查问卷) 分给各个区，并开始对西安市居民的问卷调查。

由表 2 可以看出，新城区、碑林区、灞桥区各抽样 40 人，莲湖区抽样 45 人，未央区抽样 55 人，雁塔区抽样 80 人。

表 2 西安市抽样点情况			
区名	人数 ^[2] (万人)	比例 (%)	抽样人数
新城区	58.9739	13.15	40
碑林区	61.471	13.71	40
莲湖区	69.8513	15.58	45
灞桥区	59.5124	13.27	40
未央区	80.6811	18.00	55

雁塔区	117.8529	26.29	80
合计	448.3426	100.00	300

第二阶段：在西安市所抽区内以街道为样本单元进行再次抽取，抽取方式按所在城区与街道规模的比例，运用规模分配法对问卷进行合理分配。另外，为方便收集数据，在所抽街道采取偶遇抽样方式，结合现场填答式对西安市居民进行问卷调查。

本次问卷调查，共发放调查表 300 份，回收调查表 300 份，剔除不符合要求（回答不完整、不按要求回答）的问卷，回收有效问卷 278 份，问卷有效回收率为 97.54%。

表 3 西安市抽样街道情况						
区名	街道数	抽中街道数	应收样本数	实收样本数	有效问卷数	问卷有效率
新城区	9	3	40	40	35	94.59
碑林区	8	3	40	40	39	100.00
莲湖区	9	3	45	45	41	95.35
灞桥区	9	3	40	40	35	97.22
未央区	10	4	55	55	53	98.15
雁塔区	8	5	80	80	75	98.68
合计	53	21	300	300	278	97.54

（2）中期：问卷调查表设计及分析

在大规模问卷调查前，组员于 2015 年 5 月 20 日至 5 月 22 日期间进行了预调查。发放样本 50 份，回收有效问卷 45 份，有效回收率为 90%。通过预调查，对结果运用统计分析法进行信度与效度分析，根据分析结果筛选问卷题项，调整问卷结构，最终得到一份正式问卷调查表。

问卷调查表形式：问卷分两部分（个人基本信息、调查问题），基本信息主要包括调查对象的年龄、性别、收入等；调查问题则包括方便程度、准时性、

耗时性、价格、快捷程度、诚信度、可靠性和安全性等。本问卷据调查目的展开设计，问题内容的系统性为后期数据分析处理提供了便利，同时问题数量也易于被调查对象接受，保证了问卷调查结果的真实性、适用性。

问卷信度与效度：

.信度分析即可靠性，指根据测验工具所得到的结果的一致性或稳定性，反映被测特征真实程度的指标，其大致分为三类：稳定系数、等值系数、内在一致系数。信度分析方法主要有四种：重测信度法、复本信度法、折半信度法、信度系数法。本问卷采用目前最常用的 信度系数法，这种方法适用于态度、意见式问卷的信度分析，其公式为：

$$a = \frac{k}{k-1} \left(1 - \frac{\sum S_i^2}{S_T^2} \right) \tag{2}$$

其中， k 为量表中题项的总数， S_i^2 为第 i 题得分的题内方差， S_T^2 为全部题项总得分的方差。

通常认为：信度系数在 0~1 之间（即 0.7 以上表信度良好）。按照年龄层次划分调查对象为老年人、中年人、青年人，并结合 7 个项目的数据求出信度系数，结果见表 4：调查对象的信度系数均超过 0.7，反映调查数据的可靠性、真实性。

表 4 信度分析		
项目	项目数	信度系数
老年人分值	7.0000	.702
中年人分值	7.0000	.724
青年人分值	7.0000	.715

.效度分析即有效性，指测量工具或手段能够准确测出所需测量事物的程度。效度分为三种类型：内容效度、准则效度、结构效度。其分析方法有三种：单项与总和相关效度分析、准则效度分析、结构效度分析。本文采用单项与总和相关效度分析，即计算每个题项得分与题项总和的相关系数，根据相关是否显著判断是否有效。若量表中有反意题项，应将其逆向处理后再计算总分。

(3) 后期：问卷研究方法

定性定量分析：浏览学者相关文献，结合本文问题及问卷调查表理解文献定

量分析的方法与原理；通过与组员及老师沟通交流，确定问题突破口，对已学习到的分析方法进行验证、修改，形成适合本文的定量分析法。最终，针对问卷调查数据采用 R 软件进行描述统计分析、主成分分析、多元回归分析、相关性分析。

多元回归分析，不仅能够分析解释变量对因变量影响的大小，还可以通过多元回归方程对因变量进行与猜测和控制，主要用来研究一个因变量与一个或多个解释变量之间的相互依存关系，并估计、预测解释变量对变量的研究，是研究变量之间非确定关系的多元统计分析方法。

2.打车软件满意度与各项满意度的相关系数分析

本次调查数据主要来自问卷调查表，运用相关分析法分析现阶段西安市打车软件的方便程度、准时性、耗时性、价格等 8 个方面对总体满意度的影响。对不同类型的变量采用不同的相关系数来度量，常用的相关系数主要有 Pearson 简单相关系数、Spearman 等级相关系数和 Kendall 相关系数等。本文采用 Pearson 相关系数，其公式为：

$$r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}} \tag{3}$$

相关系数的绝对值越大，相关性越强；相关系数越接近于 1 或-1，相关度越强；相关系数越接近于 0，相关度越弱。利用相关分析的思想，可以根据西安市打车软件各分项满意度与总体满意度的相关系数，来确定对总体满意度有最大影响的城市人居环境满意度指标，从而得到决定总体满意度的各分项指标，哪个分项指标具有更大的影响作用。因此，计算打车软件总体满意度指标与其他各指标满意度之间的相关关系。

针对打车软件总体满意度指标与其他各指标满意度，采用主成分分析法，得出影响打车软件满意度的因素主要有软件方便程度、耗时性、诚信度。该结论较好的反映了居民在使用软件过程中所看重的因素。

表 5 打车软件总体满意度与其影响因素的相关性						
评价项目	很不满意	不满意	一般	满意	很满意	总体满意度
方便程度	0.1538	0	0.6924	1.6924	0.9615	3.5001
准时性	0.1538	0.2308	1.0386	1.0768	0.577	3.077

耗时性	0.1538	0.077	1.1538	1.3848	0.3845	3.1539
价格	0.1923	0.2308	1.1538	0.7692	0.577	2.9231
快捷程度	0.1538	0.077	1.3845	0.9232	0.577	3.1155
诚信度	0.1538	0.077	1.2693	1.0768	0.577	3.1539
可靠性	0.1538	0.077	1.3845	1.0768	0.3845	3.0766
安全性	0.1538	0.2308	1.5	0.6152	0.3845	2.8843
总体满意度	0.158613	0.12505	1.197113	1.0769	0.552875	

由图 2 可以看出，研究指标与满意度有着良好的相关关系，结果反映了调查数据的可靠性，及分析过程的真实性。

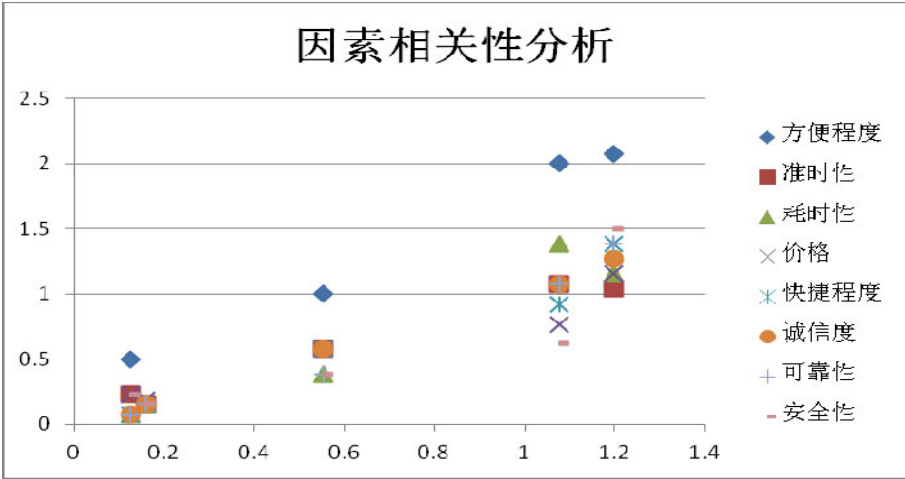


图 2 居民对打车软件的满意度与各指标相关关系

通过调查问卷数据处理，得到居民对打车软件的满意度（注：满意度分为不满意，较不满意，一般，较满意，满意 5 个等级）评价如下：

由表 6 可以看出，方便、耗时、快捷程度排前三，居民更重视出租车对个人需求的满足，同时也符合出租车的特点；在这个经济快速发展的时代，居民不免更在意打出租的目的性：快捷、省时。该结论较符合现代大多数人心理。同时，打车软件不仅给乘客带来了便利，而且在一定程度上缓解了城市交通拥堵现象，具有一定社会价值。

表 6 居民对打车软件的满意度评价				
评价项目	最小值	最大值	均值	最终排名

方便程度	1	5	3.5001	1
准时性	1	5	3.077	4
耗时性	1	5	3.1539	2
价格	1	5	2.9231	5
快捷程度	1	5	3.1155	3
诚信度	1	5	3.1539	2
可靠性	1	5	3.0766	6
安全性	1	5	2.8843	7

（三）个人属性对出行方式影响的相关性分析

相关性分析是指对两个或多个具备相关性的变量元素进行分析，从而衡量两个变量因素的相关密切程度。相关性的元素之间需要存在一定的联系或者概率才可以进行相关性分析。（注：有效问卷调查数：300 份）

表 7 居民个人属性			
因素	分类指标	参与调查人数	百分比
性别	男	158	52.7
	女	142	47.3
年龄	18 以下	20	6.67
	18~25	123	41
	25~40	102	34
	40 以上	55	18.33
职业	学生	102	34
	自由职业者	48	16
	公司职员	89	29.67
	家庭主妇	14	4.67
	其他	47	15.66
受教育程度	初中及以下	39	13
	高中及中专	56	18.67
	大专及本科	157	52.33
	研究生及以上	48	16

表 8 居民出行方式		
方式	参与调查人数	百分比
私家车	46	15.33%
出租车	93	31%
公共交通工具	115	38.33%
自行车或步行	29	9.67%
其他	17	5.67%

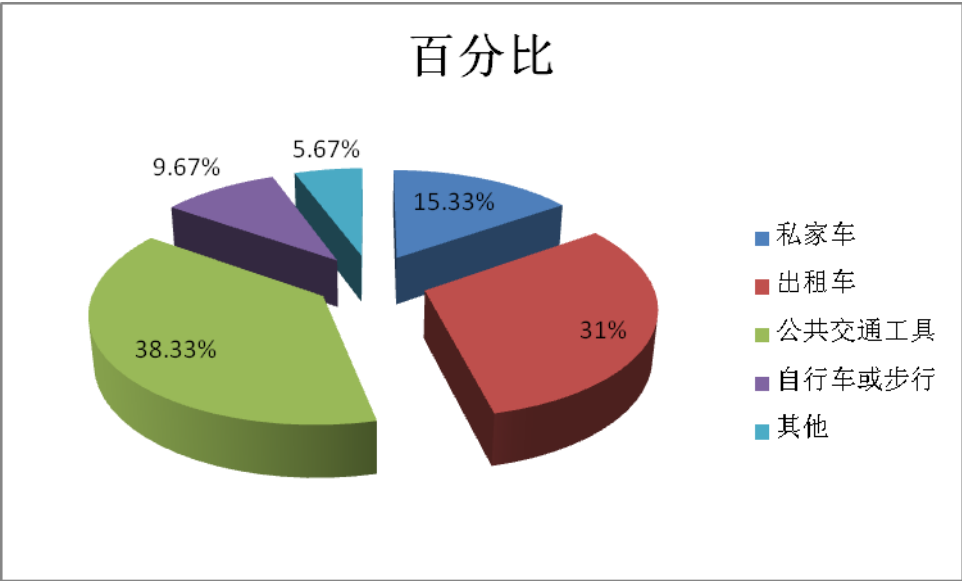


图 3 西安市居民出行方式

由居民个人属性可以看出，青年、大学、学生这三个指标更倾向于乘坐出租车，也符合现实状况；另一方面，居民出行方式所占比例可以看出，在公共交通工具满足不了居民需求时，出租车是大多数居民的第二选择。 以上结果可以反映出问卷调查数据的真实性、有效性。

根据调查问卷所获取的数据 ,对居民个人属性与居民出行方式的交叉关系进行相关性分析，画出其各分类指标的柱状图并进行如下分析。

1.性别分析：

通过表 9 与图 4 可以看出，性别各自的出行方式具有明显的不同。女性比男性更倾向选择自行车或步行方式，而男性比女性更倾向选择其他私家车和出租车方式。男性和女性在公共交通选择上无显著差别。以上结论整体上也反映出在经济快速发展的今天，尤其是交通问题日趋凸显时，出租车已然成为公交车的辅助工具，甚至“搭档”。

表 9 性别分析

X\Y	私家车	出租车	公共交通工具	自行车或步行	其他
男	30	53	54	10	11
女	16	40	61	19	6
小计	46	93	115	29	17

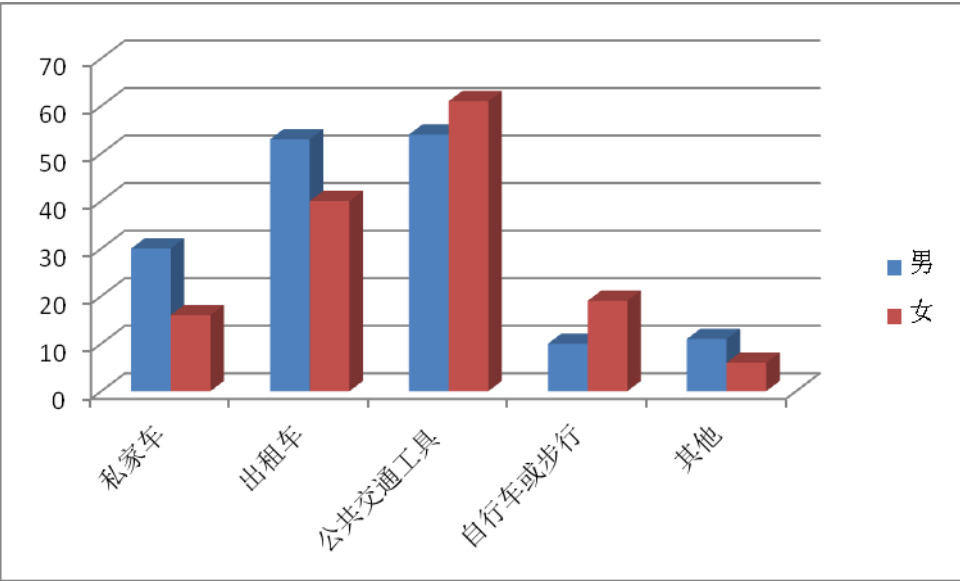


图 4 性别分析柱状图

2.年龄分析：

由表 10 与图 5 可以看出：

18 岁以下:基本都会选择公共交通工具，主体为学生；

18~25:学生群体居多，公共交通占绝大比重，对出租车和私家车的选择次之；

25~40：以出租车为主要交通工具，私家车的比例明显升高；

40 以上：以公共交通工具为主。

即：18~25 岁人群及 40 岁以上人群偏向乘坐公共交通工具，而 25~40 岁人群偏向乘坐出租车。结合问卷调查结果，以上结论与居民的日常收入、从事职业等因素有关，能较好的反映现实情况。

表 10 年龄分析

X\Y	私家车	出租车	公共交通工具	自行车或步行	其他
18 以下	1	1	13	5	0
18~25	16	33	52	13	9
25~40	23	41	24	9	5
40 以上	6	18	26	2	3
小计	46	93	115	29	17

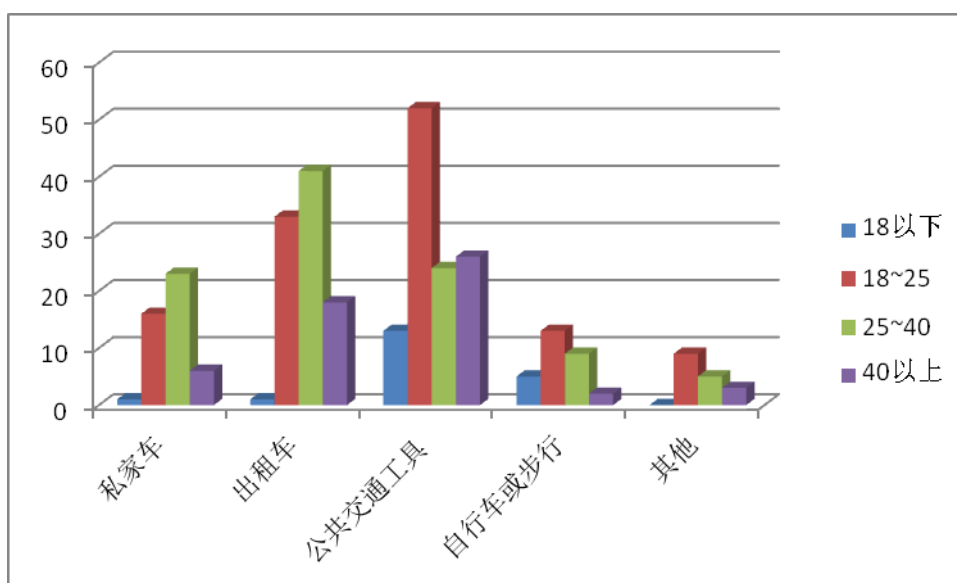


图5 年龄分析柱状图

3.职业分析：

由表 11 及图 6 可以看出：

学生：以公共交通工具为主；

自由职业者：公共交通工具和出租车占较大比重；

公司职员：以私家车和出租车为主要交通工具，这跟职员的收入密切相关；

家庭主妇：交通工具分布较为均匀，合理选择，仍是公共交通工具比重略高；

其他：交通工具分布较为均匀，合理选择，根据具体需求情况。

即学生偏向乘坐公共交通工具，而职员、自由职业者则偏向于乘坐出租车，这个结论也符合表 10 及图 5 所阐述的结果，它是与居民的日常收入、个人身份特征相关的。

表 11 职业分析

X\Y	私家车	出租车	公共交通工具	自行车或步行	其他	小计
学生	4	29	58	10	1	102
自由职业者	1	17	22	2	6	48
公司职员	32	33	16	6	2	89
家庭主妇	2	3	6	3	0	14
其他	7	11	13	8	8	47
小计	46	93	115	29	17	300

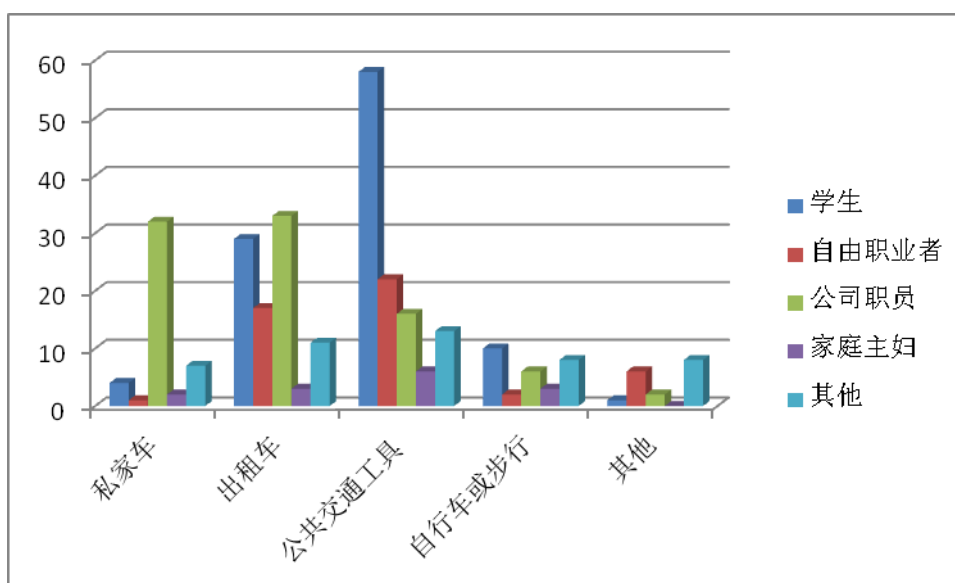


图6 职业分析柱状图

4.受教育程度分析：

由表 12 及图 7 可以看出，大专及本科生倾向于乘坐公共交通工具，研究生倾向于乘坐出租车；以上结论反映出个人属性特征对出行方式选择的影响。

通过分析，可以发现各受教育程度分类人群的出行方式的不同特征。受教育程度越高的人群越不倾向于选择步行方式链，而是越来越倾向于更为机动化的交通工具；受教育程度的高低对居民选择公共交通方式产生的影响不大。

表 12 受教育程度分析

X\Y	私家车	出租车	公共交通工具	自行车或步行	其他	小计
初中及以下	1	10	15	5	8	39
高中及中专	2	16	24	8	6	56
大专及本科	27	45	68	14	3	157
研究生及以上	16	22	8	2	0	48
小计	46	93	115	29	17	300

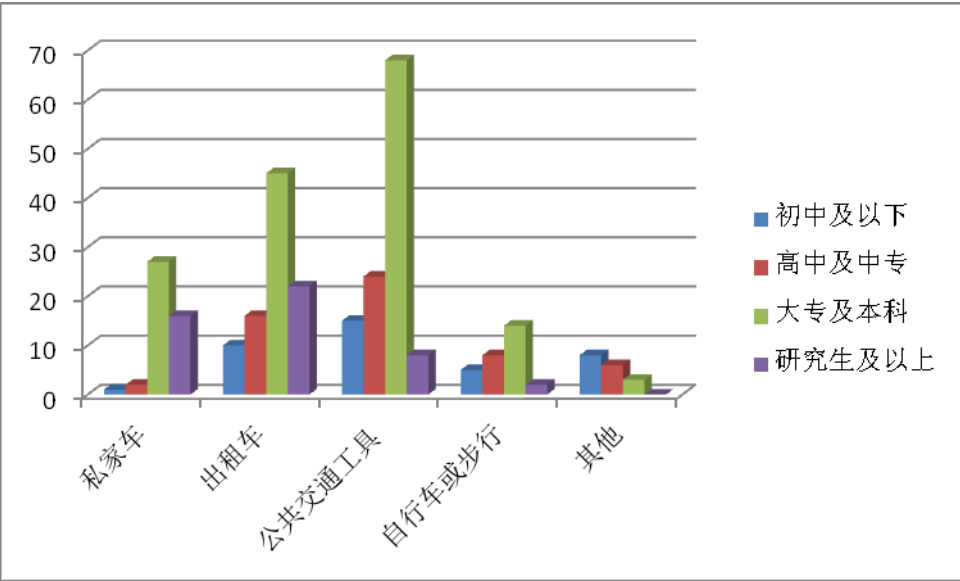


图 7 受教育程度分析

综上 ,通过对调查问卷数据进行相关分析 ,更清晰的了解了打车软件出现后 ,出租车市场的变化情况 :青年人成为软件使用主体 ,并倾向于乘坐出租车的出行方式 ;而老年人由于很少使用高科技产品 ,则受到打车软件的限制 ,出行稍不便。

（四） 打车软件出现对出租车市场的影响分析

对问卷调查结果进行汇总、相关性分析，得到以下表格与图形：

1.打车软件出现后招手打车难度变化分析

由表 13 和图 8 可以看出，53.93%居民认为打车软件的出现对打车难度无影响。结合上文研究及实际分析，打车软件的出现只是缓解了高峰期或者某些特殊情况下的传统出租车问题，并未完全取代或者完全优于传统出租车。

表 13 打车软件出现后对招手打车难度的影响	
变化	百分比
变难了	35.96%
容易了	10.11%
基本没变化	53.93%

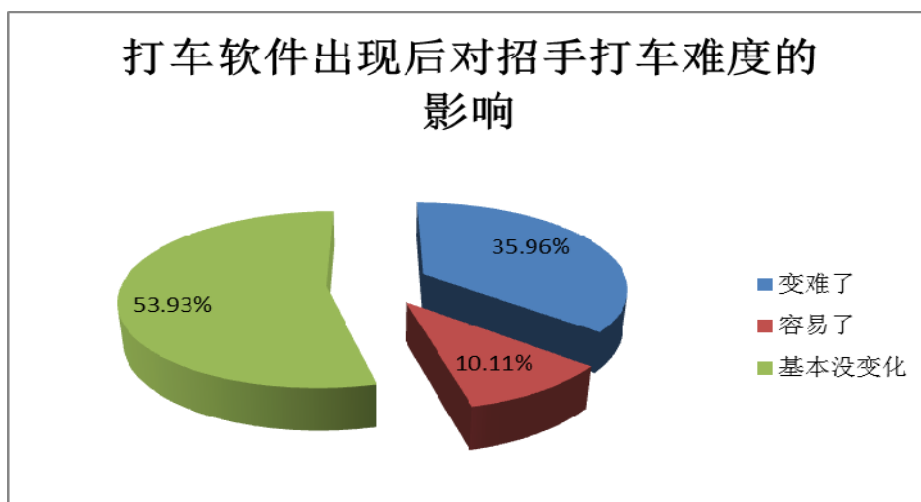


图 8 打车软件对招手打车难度的影响

2. 因他人使用打车软件而被拒载的情况分析

由表 14 和图 9 可以看出，71.91%的居民没有遇到过因打车软件而被拒载现象。考虑到调查对象的个人属性，老年人基本不用打车软件，青年及中年人间也存在一部分人不用打车软件，另外以上人群选择乘坐出租车的概率值较低，因此大部分人群不会遇到因打车软件被拒载现象。

表 14 因他人使用打车软件而被拒载的情况	
拒载情况	百分比
没遇到过	71.91%
偶尔遇到	22.47%
经常遇到	5.62%

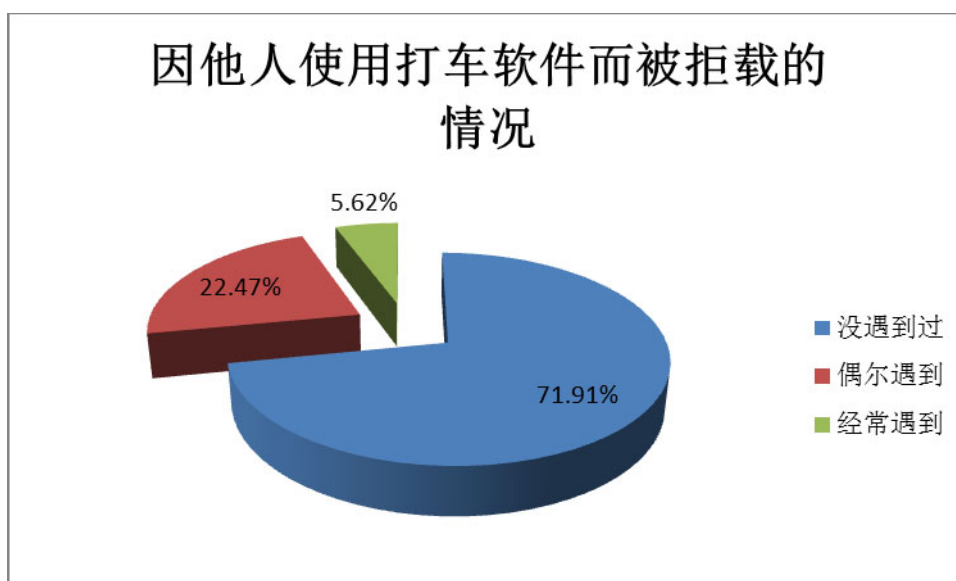


图 9 因他人使用打车软件而被拒载的情况

3.使用打车软件的原因分析

由表 15 与图 10 可以得出 ,居民选择打车软件的原因主要是它的方便性、经济上的补贴、日常打车难问题的解决。这些因素也基本体现了打车软件的优势 ,它的出现打破了原有出租车市场的垄断 ,同时给市场注入了活力 ,部分缓解了出租车市场问题。

表 15 居民使用打车软件的原因	
因素	百分比
方便	57.69%
省钱（有打车补贴）	42.31%
常规模式打车难	34.62%
不用站在风雨中打车	34.62%
利用加价抢车从而及时上车	3.85%
支付快捷无需现金支付	34.62%
有积分话费返利等一系列促销活动	7.69%
的士司机车牌等信息可自动再打车软件内记录，使得丢东西或发生纠纷时在手机端有记录可查	23.08%
其他	15.38%

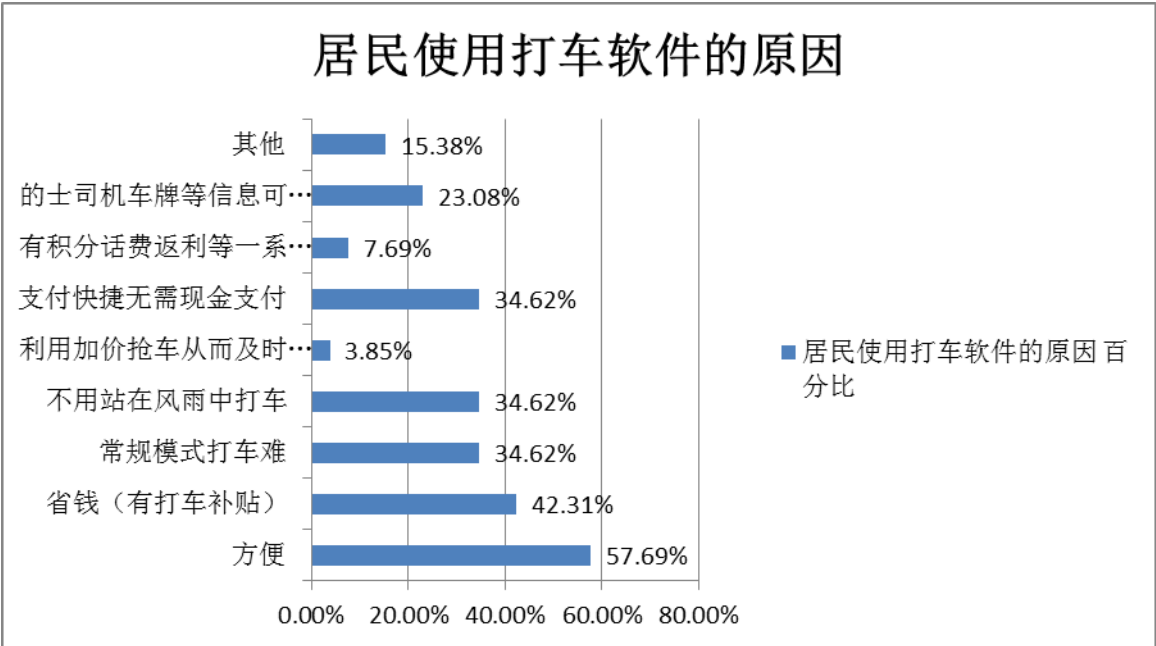


图 10 居民使用打车软件的原因

4. 关于何时使用打车软件的分析

由表 16 与图 11 可以得出 打车软件使用高峰期是在居民有突发状况需打车、自身所在地理位置不优越、上下班高峰期时。这些情况与现实良好相符 ,打车软件的出现 ,也缓解了偏远地区打不上车、突发状况及用车高峰期打不上车情况。

表 16 关于居民何时使用打车软件	
因素	百分比
上下班高峰	34.62%
节假日出去玩	19.23%
地理位置比较偏僻的地方	34.62%
天气不好的时候	15.38%
突然有用车需求时	65.38%
火车汽车站转乘	11.54%
限号	7.69%
只要打车必用打车软件	11.54%

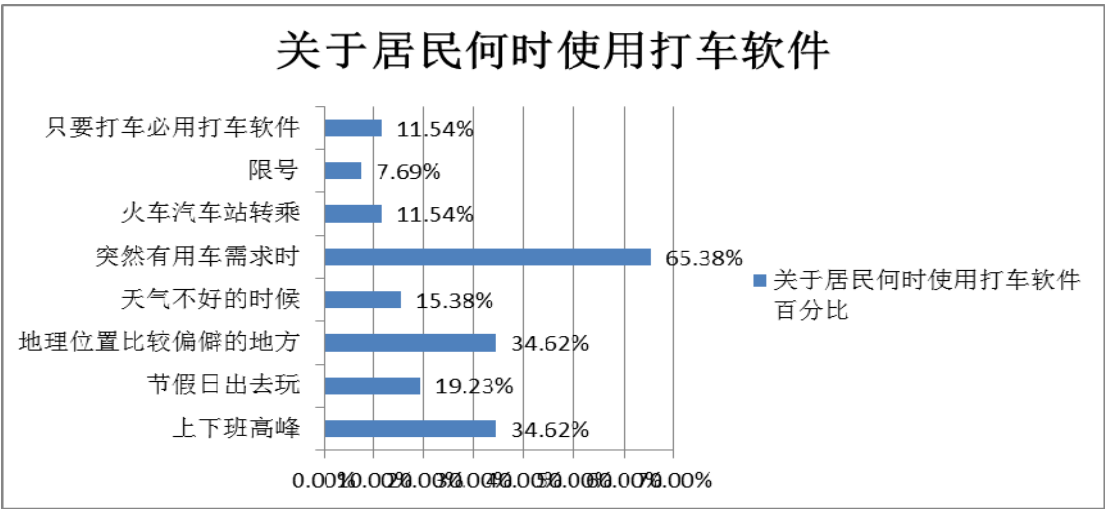


图 11 关于居民何时使用打车软件

5. 关于使用打车软件的地点分析

由表 17 与图 12 可以得出 ,居民使用打车软件的主要地点是偏僻的地理位置、学校附近。分析上文研究的打车软件作用 ,即有效缓解了传统出租车存在的问题 ,包括乘客所在地理位置偏僻这个因素。另外 ,根据校园的特征 :交通工具的需求次数多 ,可以确定学校也是打车软件的“青睐区”。

表 17 关于居民何地使用打车软件	
地点	百分比
偏僻的地方	30.77%
商场周围	11.54%
机场火车站	23.08%
学校园区	30.77%
其他	30.77%

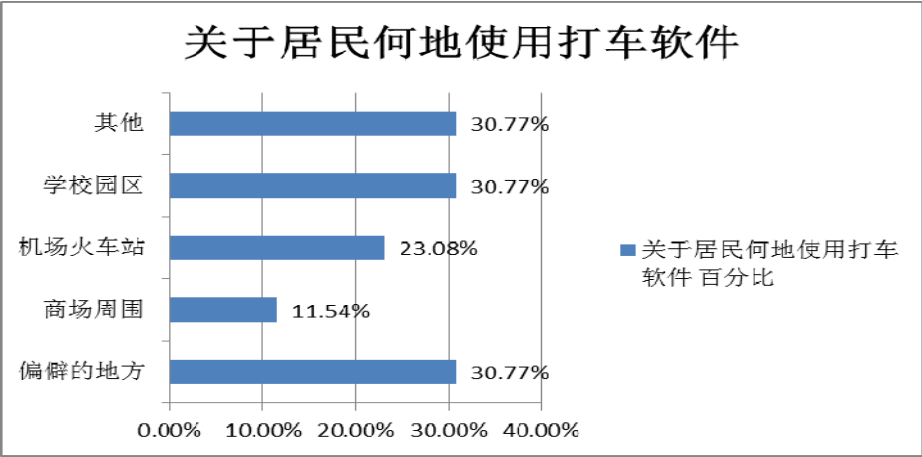


图 12 关于居民何地使用打车软件

6. 关于出租车服务选择的分析

由表 18 与图 13 可以得出，73.08%的调查对象偏向于普通叫车服务，“专车”虽然伴随着打车软件出现，但并非市场主体，仅被少部分居民接受。

表 18 关于使用打车软件叫车服务的情况	
叫车服务	百分比
普通叫车服务	73.08%
专车服务	26.92%

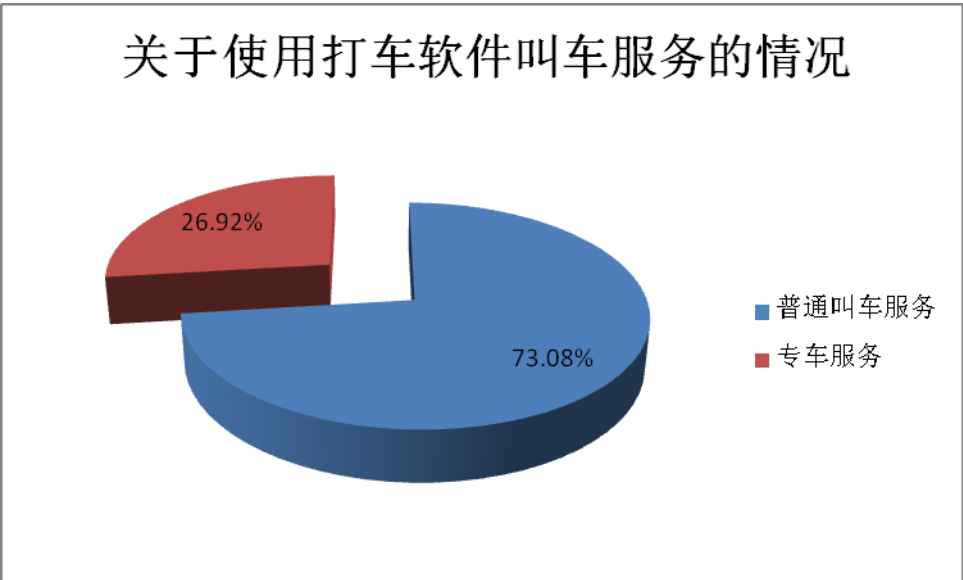


图 13 关于居民使用打车软件何种叫车服务

(五) 模型求解与检验

1. 数据标准化

数据的标准化是将数据按比例缩放,使之落入一个小的特定区间。即去除数据的单位限制,将其转化为无量纲的纯数值,便于不同单位或量级的指标能够进行比较和加权。本文采用 0-1 标准化,转化函数为:

$$x^* = \frac{x - \mu}{\sigma} \quad (4)$$

其中, μ 为所有样本数据的均值, σ 为所有样本数据的标准差。

2. 模型设定

设变量 Y 与变量 X_1, X_2, \dots, X_p 间有线性关系:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_p X_p + \varepsilon \quad (5)$$

其中 $\varepsilon \sim N(0, \sigma^2)$, $\beta_0, \beta_1, \beta_2, \dots, \beta_p$ 和 σ^2 是未知参数且 $p \geq 2$, 称模型 (5) 为多元线性回归模型

设 $(x_{i1}, x_{i2}, \dots, x_{ip}, y_i)$, $i = 1, 2, \dots, n$ 是 $(X_1, X_2, \dots, X_p, Y)$ 的 n 次独立观测模型, 则多元线性模型 (5) 可表示为:

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, i = 1, 2, \dots, n \quad (6)$$

其中 $\varepsilon_i \in N(0, \sigma^2)$, 且独立同分布。

为书写方便, 常采用矩阵形式, 令:

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix}, X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (7)$$

则多元线性模型 (6) 可表示为

$$Y = X\beta + \varepsilon \quad (8)$$

其中 Y 是由响应变量构成的 n 维向量, X 是 $n \times (p+1)$ 阶设计矩阵, β 是 $p+1$ 维向量, ε 是 n 维误差向量, 并且满足:

$$E(\varepsilon) = 0, \text{var}(\varepsilon) = \sigma^2 I_n \quad (9)$$

3. 回归系数的估计

类似于一元线性回归，求参数的估计值 $\hat{\beta}$ ，就是求最小二成函数

$$Q(\beta) = (Y - X\beta)^T (Y - X\beta) \quad (10)$$

达到最小的 β 值。

可以证明 β 的最小二乘估计

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (11)$$

从而可得经验回归方程为

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 + \cdots + \hat{\beta}_p X_p \quad (12)$$

称 $\hat{\varepsilon} = Y - X\hat{\beta}$ 为残差向量，通常取

$$\hat{\sigma}^2 = \hat{\varepsilon}^T \hat{\varepsilon} / (n - p - 1) \quad (13)$$

为 σ^2 估计，也称为 σ^2 的最小二乘估计。可以证明

$$E\hat{\sigma}^2 = \sigma^2 \quad (14)$$

可以证明 β 的协方差矩阵为

$$\text{var}(\beta) = \sigma^2 (X^T X)^{-1} \quad (15)$$

相应地， $\hat{\beta}$ 的标准差为

$$sd(\hat{\beta}_i) = \hat{\sigma} \sqrt{c_{ii}}, i = 0, 1, 2, \cdots, p \quad (16)$$

其中 c_{ii} 是 $C = (X^T X)^{-1}$ 对角线上的第 i 个元素。

4. 显著性检验

回归系数的显著性检验

$$H_{j0} : \beta_j, H_{j1} : \beta_j \neq 0, j = 0, 1, \cdots, p \quad (17)$$

当 H_{j0} 成立时，统计量

$$T_j = \frac{\hat{\beta}_j}{\hat{\sigma} \sqrt{c_{jj}}} \sim t(n - p - 1), j = 0, 1, 2, \cdots, p \quad (18)$$

其中， c_{jj} 是 $C = (X^T X)^{-1}$ 对角线上的第 j 个元素。对于给定的显著性水平 α ，检

验的拒绝域为

$$|T_j| \geq t_{\frac{\alpha}{2}}(n-p-1), j = 0, 1, 2, \dots, p \quad (19)$$

回归方程的显著性检验

$$H_0: \beta_0 = \beta_1 = \dots = \beta_p = 0, H_0: \beta_0, \beta_1, \dots, \beta_p \text{不全为} 0。$$

当 H_0 成立时，统计量

$$F = \frac{SS_R / p}{SS_E / (n - p - 1)} \sim F(n - p - 1) \quad (20)$$

其中 $SS_R = \sum_{i=1}^n (\hat{y}_i - \bar{y})^2, SS_E = \sum_{i=1}^n (y_i - \hat{y}_i)^2$ ，通常称 SS_R 为回归平方和， SS_E 为参差平方和。

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i, \hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_p x_{ip},$$

(六) 模型求解与检验

考虑西安市打车软件的方便程度、准时性、耗时性、价格、快捷程度、诚信度、可靠性和安全性等 8 个方面的评分对总体评分的影响。建立总体评分(Y)关于 X_1, X_2, \dots, X_8 的多元线性回归方程。

根据样本向量，经标准化计算后得出标准分，见表 7。

表 19 打车软件因素评分

评价项目	方便程度	准时性	耗时性	价格	快捷程度	诚信度	可靠性	安全性	总体评分
1	-2.51817	-1.69423	-2.33464	-1.14823	-2.1123	-1.65831	-1.64551	-1.51495	-1.93919
2	0.179869	0.726098	1.000562	0.618279	0.111174	0.829156	-0.61707	-0.3496	-0.09234
3	0.179869	0.726098	-0.11117	-0.26498	1.222909	0.829156	0.411377	-0.3496	-0.09234
4	-0.71948	-0.08068	-0.11117	-1.14823	-1.00056	0	-0.61707	-0.3496	-1.01577
5	1.079215	-0.88745	-1.22291	1.501536	-1.00056	-1.65831	-1.64551	-1.51495	-1.01577
6	-0.71948	-0.88745	-0.11117	-1.14823	0.111174	0	0.411377	0.81574	-0.09234
7	0.179869	-0.88745	1.000562	-1.14823	1.222909	0.829156	0.411377	0.81574	0.831082
8	0.179869	-0.08068	-0.11117	0.618279	1.222909	-0.82916	1.439818	-0.3496	0.831082
9	1.079215	1.532874	1.000562	0.618279	0.111174	1.658312	1.439818	0.81574	1.754507
10	1.079215	1.532874	1.000562	1.501536	0.111174	0	0.411377	1.981083	0.831082

使用 R 软件进行多元线性回归求解，可得结果如下表 8。其中 $R^2=0.9847$ ， $\bar{R}^2=0.8619$ ，可决系数较高，说明模型对样本的拟合较好，且回归系数与回归方程的检验都是显著的。回归方程为：

$$\hat{Y} = -1.252127 - 0.002624X_1 - 0.477543X_2 + 0.193623X_3 + 0.557010X_4 - 0.189702X_5 + 0.452672X_6 + 0.843825X_7 + 0.174967X_8$$

表 20 多元线性回归求解

Call:							
lm(formula = Y ~ X1 + X2 + X3 + X4 + X5 + X6 + X7 + X8, data = Data)							
Residuals:							
1	2	3	4	5	6	7	8
0.14293	-0.10873	-0.01222	-0.01832	-0.04398	-0.32496	0.22112	0.01466
9	10						
0.05620	0.07330						
Coefficients:							
Estimate Std. Error t value Pr(> t)							
(Intercept)	-1.252127	0.752842	-1.663	0.345			
X1	-0.002624	0.342052	-0.008	0.995			
X2	-0.477543	0.317588	-1.504	0.374			
X3	0.193623	0.499431	0.388	0.765			
X4	0.557010	0.415532	1.340	0.408			
X5	-0.189702	0.318782	-0.595	0.658			
X6	0.452627	0.381543	1.186	0.446			
X7	0.843825	0.297056	2.841	0.215			
X8	0.174967	0.325900	0.537	0.686			
Residual standard error: 0.4449 on 1 degrees of freedom							

Multiple R-squared: 0.9847, Adjusted R-squared: 0.8619
F-statistic: 8.023 on 8 and 1 DF, p-value: 0.2668 e-3

五、结论与建议

通过对调查数据的处理分析,针对西安市出租车市场得出以下结论,并提出相关建议:

(一) 结论

通过调查问卷法以及上述满意度、居民出行等相关性分析,可以得出:打车软件的出现,一定程度上缓解了传统出租车存在的问题,在有益的同时亦存在弊端,分析如下:

1. 部分人群打车难:打车软件的使用人群主要是中青年者,他们能够较快的适应并熟练运用打车软件;相对于那些不会运用和没有下载打车软件的人群来说是不公平的。再结合打车软件的补贴。司机更倾向于搭乘前者,所以会导致打车难。

2. 乘客安全受影响:出租车司机为迅速成功抢单,在行驶过程中时刻关注手机,这给乘客安全到来了巨大的隐患,尤其在高峰期。另外,司机为达到个人利益最大化,会使用多个手机安装多个打车软件,这样势必会分散开车时的注意力,威胁到乘客的安全。

3. 加价方式不合理:在打车软件的应用程序界面上,乘客自身可通过自愿加价行为吸引司机,提高打车效率。在此过程中,乘客看似自愿加价,实则若不加价,很难打到车。另外不少出租车司机利用打车软件的加价进行选择性的载客,实为一种变相的拒载。

4. 乘客的信息安全受威胁:现在的手机消费软件都是通过手机网络支付的,乘客要绑定个人银行卡号才能支付成功,而订单的提交需要提供相关个人信息,这样可能就会导致乘客信息的泄露,影响乘客的正常生活。

5. 打车软件的副产品产生:打车软件并非只有政府出租车可以装,私家车在满足一定要求时也可以装载该软件。这就导致黑车进入出租车市场,容易引起出租车市场的混乱,对正规出租车司机亦不公平。

（二）建议

打车软件是出租车市场创新的产物,目前看来它更大的缓解了传统出租车市场的不足,但是由于国家、政府对打车软件没有进行适当的管理,持续下去,会反作用于市场,对市场产生不良影响,因此对政府建议如下:

1.打车软件行业未良好运营时,政府应当给予一定鼓励。从前文的分析以及实际调查来看,打车软件的产生是顺应市场的结果。而初期的进军市场还需政府把关,使打车软件走在正轨上,并逐渐规范化。违反或者与市场冲突的地方,要予以纠正。

2.打车软件进军出租车市场并趋于稳定时,政府要严格管理。现在看来,初期的打车软件确实存在大量不足,但是为了确保打车软件对出租车市场的有效调节,政府还不能大刀阔斧的进行管理,可能会导致打车软件中途夭折。但是稳定后的打车软件,必须要进行整顿,考虑到它已经进入居民的生活,政府亦要保障居民的权力与正规出租车司机权力,必须出台相关规范,对现存打车软件进行规范化治理。

3.打车软件在出租车行业站稳脚跟并蓬勃发展时,政府要合理调控出租车行业。技术上改进,使其与政府出租车系统融为一体,更好的组合来为居民服务,并解决传统出租车遗留的问题。制度上要合理均衡打车软件的利弊,使其正当竞争。管理上要像正规出租车一样对打车软件进行准入、安全、价格、投诉管理。

六、模型评价与改进

（一）优点

本次调查采用定量分析与定性研究相结合的方法。在定量分析中,运用问卷调查法及比例分配、规模分配法对样本量进行分配,使其更具有代表性。另外在

进行问卷调查时,首先针对调查目的进行了预调查,旨在修改问卷不足,从而制作一份真实、有效的问卷。其次,对问卷数据及题目进行了描述性分析及相关性分析,更清晰的看出了现在出租车市场的变化。

(二) 缺点

由于时间、人力等原因,本文的抽样并非完全随机,有不能科学准确地推断出打车软件影响下的居民满意度情况;其次文中的人数、车数均为往年数量,并不能完全表示现在的人口数、车辆数,会导致结果产生一定误差;最后本文针对出租车市场的分析选取满意度与居民出行两种方式,在研究整个出租车市场变化上亦会存在一定误差。

(三) 改进

1. 扩大样本容量,提高数据的全面性和可靠性。
2. 提取更科学的角度,对居民乘车满意度更深入分析。
3. 建立其他模型与本文模型进行对比,从而确定更优模型。

参考文献

- [1] 新华网:打车软件竞争观察:《为何烧 烧着谁 烧多久》.
- [2] 吉波. 大城市出租车需求对策及经营管理模式研究[D]. 广州:华南理工大学,2014.
- [3] 王克. 城市保有量研究[D]. 大连:大连交通大学,2014.
- [4] 陕西省 2010 年第六次全国人口普查主要数据公报
<http://www.shaanxi.gov.cn/0/1/9/39/100600.htm>
- [5] David Freedman. 统计学[M]. 中国:中国统计出版社,2000.
- [6] 薛毅,陈立萍. 统计建模与R软件[M]. 北京:清华大学出版社. 2007. 4.
- [7] 科克伦. 抽样技术[M]. 中国:中国统计出版社. 2010.

- [8] 张亦辉,郑有增,周朋红等. 调查问卷的信度效度分析[J]. 科研导航. 2003,(第 22 期).
- [9] 百度百科_全球最大中文百科全书 <http://baike.baidu.com/>
- [10] 吴长莉,陶秀玲.出租车行业顾客满意度影响因素分析[J].商场现代化,2007,(第 518 期).
- [11] 李雪.出租车行业的政府管制研究[J].商业文化(学术版).2010(10).
- [12] 桂虹,杨剑.出租车行业管理中的政府角色义务:基于打车软件的分析[J].企业导报,2014,(第 14 期).

附 录

附录一:

西安市居民乘坐出租车情况的调查

您好,为了更确切的了解西安市居民对打车软件的满意度及出租车乘坐情况,特制作此调查问卷,望您能抽出一点宝贵的时间来参与我们的调查,对此我们表示万分感谢!本问卷采取匿名形式且仅用于研究。

1. 请问您的性别是?

男

女

2. 您目前的职业是?

学生

自由职业者

公司职员

家庭主妇

其他

3. 请问您的年龄段处于?

18 以下

18~25

25~40

40 以上

4. 您平时出门一般选择什么方式？

私家车

出租车

公共交通工具（公交车、地铁等）

自行车或步行

其他

5. 您平常打出租车的频率是？

一周 1~3 次

一周 4~5 次

每天都坐

几乎不坐

6. 您平常采用何种方式打车？

打车软件打车

路边招手叫车

两者均可

7. 打车软件出现的近 1 年来招手叫车的难度有何变化？

变难了

容易了

基本没变化

8. 您是否遇到过因他人使用打车软件而被拒载的情况？

没遇到过

偶尔遇到

经常遇到

9. 由于打车软件的广泛使用，一些人（尤其是老人）在路面上招手打不到出租车，

您是怎样看待这一情况的？

打车软件本来就有一定的适应群体

政府应出台措施进行管理

打车软件应当设计出一款专门适合特殊人群的打车系统

所有人都应主动去学习使用打车软件

其他

10. 您使用软件打车通常选择？

滴滴打车

快的打车

Uber

其他

11. 您通常使用软件内哪种叫车服务？

普通叫车服务

专车服务

12. 您通常在什么情况下使用打车软件？

上下班高峰

节假日出去玩

地理位置比较偏僻的地方

天气不好的时候

突然有用车需求时

火车汽车站转乘

限号

只要打车必用打车软件

13. 您一般选择什么时段使用打车软件打车？

上下班高峰

周末

凌晨前后

无固定时间

14. 您用打车软件的地点？

偏僻的地方

商场周围

机场火车站

学校园区

其他

15. 您用打车软件打车被放过鸽子吗？

没有

有

16. 如果没有优惠措施您还会使用打车软件吗？

不会

会

17. 您使用打车软件的原因是？

方便

省钱（有打车补贴）

常规模式打车难

不用站在风雨中打车

利用加价抢车从而及时上车

支付快捷无需现金支付

有积分话费返利等一系列促销活动

的士司机车牌等信息可自动再打车软件内记录，使得丢东西或

发生纠纷时在手机端有记录可查

其他

18. 您认为打车软件存在的弊端是？

想打车的时候还是打不到

肆意加价影响市场秩序

比路边打车更慢

有些司机放鸽子

坐车过程中，司机一直被打车软件干扰，不安全

打车要使用 app 输入起点和终点等信息，对智能机不熟悉的人使用不方便

投诉处理困难

黑车抢单

担心个人信息泄露

其他 _____

19. 您如何看待打车软件目前推出的专车服务？

20. 您认为打车软件发展的前景如何？

很好

较好

一般

不好

难说

21. 您对打车软件有什么评价？

	很不满意	不满意	一般	满意	很满意
方便					
准时					
省时					
便宜					
快捷					
诚信度（避免诈骗，黑车）					
可靠性（能否打到车）					
安全性					

22. 您希望打车软件增加什么功能？ [填空题]

23. 您如何看待政府出台的电召平台这一政策？ [单选题] [必答题]

政府统一管理，有利于资源整合和规范，将有效缓解打车难的问题，值得肯定

政府将手伸向市场，干扰正常市场竞争秩序，有些不妥

无所谓，没太多看法

24. 您认为这一政策会对打车软件产生什么影响？ [单选题] [必答题]

统一平台，民间打车软件势必面临竞争壁垒，前景堪忧

不会有太大影响，民间打车软件已积累了客户资源，早早占领了市场

25. 您认为哪种途径更有利于解决打车难的问题？ [单选题] [必答题]

由政府按照目前出台的监管政策，有奖有罚，严格执行

政府电召平台和手机打车软件并行发展，遵循市场规律，公平竞争

让政府主要负责行业监督，规范叫车服务业，使其良性竞争发展