

文章编号:1005-0523(2015)02-0054-05

基于生存分析方法的步行行人闯红灯行为研究

王海婵, 郑长江

(河海大学土木与交通学院, 江苏 南京 210098)

摘要:将海德归因理论应用到步行行人过街行为中,提出了步行行人闯红灯归因理论,该理论分析了步行行人闯红灯行为及其主要影响因素,利用生存分析方法,研究步行行人的闯红灯行为,具体分析步行行人闯红灯的影响因素,然后以中小城市的交叉口例,分析影响因素的影响效果,结果表明,该理论与方法可以有效地研究影响步行行人闯红灯行为的影响因素。

关键词:步行行人;闯红灯;归因理论;生存分析方法

中图分类号:U491

文献标志码:A

“中国式闯红灯”,即“凑够一批人就过街,与红绿灯无关”的现象在我国很普遍^[1]。我国新交通法规的发布,加大了对闯红灯的机动车的惩罚力度,在一定程度上明显减少了机动车闯红灯数量^[2]。然而,步行行人闯红灯问题始终是我国城市交通管理的一个重点和难点。北京等发达城市经过多年整治使这一现象有所改善,但中小城市的步行行人闯红灯问题依然很严重。行人作为交通参与者中的弱势群体,极易受到机动车的伤害。行人闯红灯不仅违反交通规则,而且严重影响交通秩序,提高了交通事故发生的可能性^[3]。

从信号交叉口步行行人闯红灯的原因着手,将海德归因理论应用于步行行人闯红灯行为中,建立步行行人闯红灯的归因理论,从而对步行行人闯红灯行为及其影响因素进行分析,通过对中小城市的交叉口进行实地调查,利用生存分析方法,具体地对步行行人等待绿灯时间长短及2个重要影响因素作分析,并拟合出一定条件下不同影响因素与生存函数曲线的一元线性回归方程。

1 步行行人闯红灯归因理论

1.1 步行行人闯红灯行为

步行行人到达路口后,看到所显示的信号灯(刺激),判断交叉口情况(态度),从而采取相应的行为(反应)^[4]。如图1所示,为步行行人过街行为的决策图。

根据信号灯的情况,将步行行人的过街行为分为以下4种:全红,全绿,红—黄—绿,绿—黄—红。以上行为的具体含义是^[5]:

全红,步行行人在过街过程中,信号灯始终显示红灯;全绿,过街过程中,信号灯始终显示绿灯;红—黄—绿,过街过程中,信号灯由红灯到黄灯过渡到绿

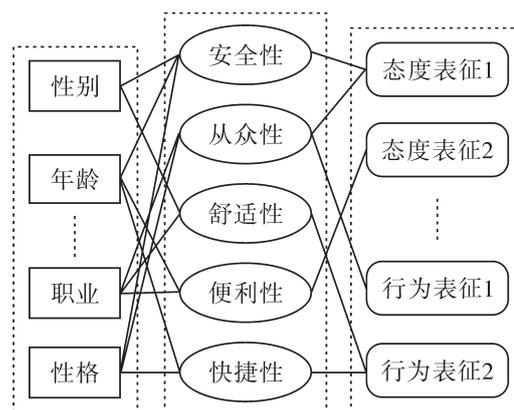


图1 行人过街行为决策图

Fig. 1 Decision-making figure of pedestrians' crossing behavior

收稿日期:2014-10-27

基金项目:江苏省城市智能交通重点实验室开放项目(JTKT2014002)

作者简介:王海婵(1989—),女,硕士研究生,研究方向为交通运输规划与管理;郑长江(1966—),男,教授,博士,研究方向为交通运输规划管理、交通运输信息与控制、城市交通安全工程。

灯;绿—黄—红,过街过程中,信号灯由绿灯到黄灯过渡到红灯。

其中,步行行人因为各种原因,闯红灯行为会出现在全红,红—黄—绿,绿—黄—红3种行为过程中。

1.2 海德归因理论

归因指通过关于自己或他人行为的分析,指出行为的性质或推断其原因的过程,即从行为的结果来论述行为的原因^[6-7]。海德指出,为了预测别人的行为,并有效地控制环境,问题的关键是分析别人行为或事件的原因。

海德认为,一个人的行为必有原因,其原因或者决定于外界环境,或者决定于主观条件。归因可以分为:①情境归因,即外部原因,如个体周围的环境因素;②个体倾向归因,即内部原因,如情绪,需要,态度,兴趣等。个体的内部原因与外部原因共同作用造就个体的行为^[8]。

1.3 步行行人闯红灯归因理论

步行行人闯红灯行为的原因是多样的,既有内在因素,又有外在因素,内外因素共同作用,共同影响行人闯红灯的行为。

为了预见步行行人闯红灯的行为,并有效地控制步行行人过街环境,关键在于对步行行人闯红灯的行为做出原因分析。其原因可归为两类,一是内因,步行行人交通系统内部因素的影响,二是外因,步行行人交通系统以外的其他因素影响^[9]。步行行人闯红灯行为是外因与内因的共同作用结果。

1) 外因:信号设置不合理,缺乏有效的处罚措施,天气环境差,交通环境复杂等。

2) 内因:自律意识、安全意识低,从众、侥幸心理,赶时间等。

研究其中的几个主要影响因素^[10]:①步行行人在红灯期间到达路口后,等待绿灯亮起所需时间的长短^[11];②步行行人到达信号交叉口时,正在等待绿灯的行人数量;③步行行人到达信号交叉口时,正在闯红灯过街的行人数量。

2 生存分析方法

2.1 生存分析的含义

生存分析是将终点事件(事件的结果)和生存时间(出现这一结果所经历的时间)结合起来分析的一种统计分析方法。生存分析的两个重要概念:生存函数与删失数据。生存函数是指以时间为自变量,以相应时刻的生存率为函数值的函数,它能反映每一时刻事件的生存率。删失数据指在研究时间结束时,某些个体上还没有发生所观测的含有这些事件的数据。因此,生存分析方法适合用来研究动态事件的发生过程规律以及影响事件发生的影响因素^[12]。

2.2 生存函数

生存函数表示生存时间大于给定时间的概率。以生存时间函数作为分析。生存时间指从某个起点事件开始到某个终点事件发生所经历的时间。生存时间数据可分为删失数据和完全数据。

针对在红灯期间到达信号交叉口的步行行人,事件起点为步行行人在红灯期间到达斑马线的时刻,事件终点为步行行人开始过街的时刻,生存时间为步行行人过街前的等待忍耐时间。事件的结果有两种,遵守交规(等到红灯结束绿灯开始再过街)和违反交规(闯红灯)。红灯期间到达交叉口并等待至绿灯才过街的步行行人,无法知道其最大等待忍耐时间,为删失数据;闯红灯的步行行人,可以知道其最大等待忍耐时间,为完全数据^[13-14]。

这里,生存时间表示步行行人过街前的等待忍耐时间,用符号 t_s 表示,其分布函数为

$$F(t) = P_r(t_s \leq t) = \int_0^t f(u) du \quad (1)$$

生存函数表示步行行人在信号交叉口前的等待时间大于 t 的概率,用符号 $S(t)$ 表示,其数学表达式为

$$S(t) = 1 - F(t) = P_r(t_s > t) = \int_t^{\infty} f(u) du \quad (2)$$

2.3 非参数估计方法

步行行人闯红灯分析中包含删失数据,利用生存分析中的非参数方法来估算步行行人在信号交叉口等待时间的生存函数。设生存时间数据(包括删失数据的和非删失数据)有 n 个,将其从小到大排列 $t_1 \leq t_2 \leq t_3 \leq \dots \leq t_n$,则:

$$S(t_i) = \prod_{j^*} \{1 - d(j)/r(j)\} \quad (3)$$

式中: $S(t_i)$ 为生存时间大于 t_i 的概率; j^* 为生存时间 $t_j < t_i$ 的所有 j 的集合,这里 i 和 j 是序号, $j < i$; $d(j)$ 为在 t_j 时间段内发生闯红灯事件的行人数量, $r(j)$ 为在 t_j 时间段内未发生闯红灯事件的行人数量;其中,把 $1 - d(j)/r(j)$ 写成 $P_j = [r(j) - d(j)]/r(j)$, P_j 为 t_j 时间段内的生存概率^[13]。

$$S(t_i) = P_1 \times \dots \times P_{t_i} \quad (4)$$

3 实例分析

3.1 调查分析

中小城市交叉口较小,步行行人数量较多,而且行人的交通意识淡薄,闯红灯现象很严重。选取中小城市盐城市的黄海路与小海路交叉口进行调查。该交叉口是主干路与次干路的十字交叉口,但行人量较大,闯红灯现象严重,故作为调查对象。

采用人工观测法,将信号周期的整个红灯期间作为数据收集的一个单元,步行行人在红灯期间到达的作为有效样本,主要调查数据为 $t_i - t_{i-1}$ 时间段内发生闯红灯行为的行人数量与未发生闯红灯行为的行人数量。

等待忍耐时间指从步行行人到达停止线开始等待的时刻到开始过街的时段的这一时间段。完全数据指闯红灯的步行行人样本,删失数据指红灯期间到达并等到绿灯后才开始过街步行行人样本。选取其中2个重要影响因素进行研究,如表1所示,为其变量含义。

表1 影响因素变量

Tab. 1 Variable of influencing factors

变量	变量类型	变量含义
N_w	离散变量	步行行人到达信号交叉口时,正在等待绿灯的行人数量
N_c	离散变量	步行行人到达信号交叉口时,正在闯红灯过街的行人数量

3.2 基本数据分析

调查时,步行行人在红灯期间到达交叉口,记录0~80 s不同等待时间下选择闯红灯的步行行人数量与等待红灯的步行行人数量。调查的有效样本数为669,其中265个行人发生了闯红灯行为,其中,步行行人最大的等待忍耐时间是96 s,即步行行人遵守交通规则一直等到绿灯才过街;最小的等待忍耐时间是0 s,即步行行人在红灯期间到达信号交叉口时直接闯红灯过街了。

该实例中, t_i 表示生存时间,生存时间数据个数 $n=40$ 个, $i=1,2,3,4,\dots,40$;2 s为一间隔, $t_1=2$, $t_2=4$, $t_3=6$, $t_4=8$, $t_5=10$, \dots , $t_{40}=80$ s; j 是序号,若 $i=10$,则 $j=1,2,3,4,\dots,9$ 。

$P_1 = [r(1) - d(1)]/r(1)$ 表示在 t_1 时间段内的生存概率,其中 $d(1)$ 表示在 t_1 时间段内发生闯红灯行为的行人数量, $r(1)$ 表示在 t_1 时间段内未发生闯红灯行为的行人数量, $d(1)$ 与 $r(1)$ 为观测数据, $S(t_1) = P_1$ 表示生存时间大于 t_1 的概率;

以此类推, $P_2 = [r(2) - d(2)]/r(2)$, $S(t_2) = P_1 \times P_2$; \dots ;

$P_{40} = [r(40) - d(40)]/r(40)$, $S(t_{40}) = P_1 \times P_2 \times P_3 \times \dots \times P_{40}$ 。

将调查数据按上述方法进行处理,得出等待时间与生存函数的关系如图2所示。从图上可以看出信号

交叉口处步行行人遵守交通规则的几率,闯红灯行为随等待绿灯时间变化的趋势。可知,随着等待时间的增加,交叉口处步行行人遵守交通规则的概率是逐渐降低的,即闯红灯概率是增加的:①在0~5 s内生存函数下降较快,表示大部分步行行人直接闯红灯过街;②6~75 s内生存函数下降较平缓,表示步行行人等待绿灯的忍耐时间有限,因人而异,等待时间越长,闯红灯的概率越高;③75~80 s内可见部分步行行人不愿闯红灯,能够等待红灯结束,到信号灯变为绿灯才过街。

等待时间与生存函数的曲线关系类似于递减的线性比关系。近似地用一元线性回归方程 $S(t) = at_i + b$ 来表示,其中, $S(t)$ 代表生存函数, t_i 代表等待时间, a 代表一定比例, b 代表常数。利用此曲线可预测一定红灯等待时间内步行行人闯红灯的概率。

3.3 影响因素的分析

利用上述的生存分析方法,可以得到其2个重要影响因素下的生存函数。

1) 在交叉口调查,步行行人在红灯期间到达交叉口时,分别记录正在等红灯行人数量 N_w 为0-2,2-5,>5三种情况时,0~80 s不同等待时间下闯红灯步行行人数量与等待红灯行人数量,利用上述生存分析方法,得出图3所示的3条曲线。此处数据调查建立在图1数据调查方法的基础上。

图3所示,为信号交叉口处不同等待行人数量下的生存函数。该图表明了步行行人在红灯期间到达信号交叉口时,等待绿灯的行人数量对步行行人闯红灯行为选择的影响,等待绿灯的行人数量越多,则步行行人的生存率越高,即到达信号交叉口的步行行人选择闯红灯行为的概率就越低。

计算得到图3中各曲线的一元线性回归方程分别为 $S=0.601-0.006 t(N_w=0-2$ 时); $S=0.940-0.011 t(N_w=3-5$ 时); $S=0.866-0.10 t(N_w>5$ 时)。

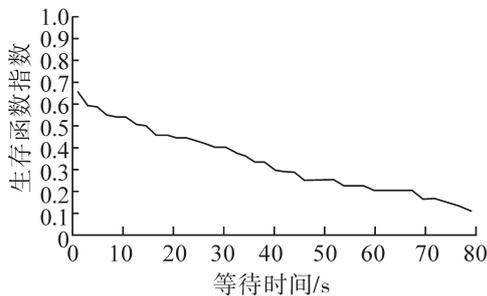


图2 信号交叉口处步行行人等待时间的生存函数

Fig. 2 Survival function of pedestrians' waiting time at signal intersection

2) 在交叉口调查,步行行人在红灯期间到达交叉口时,分布记录正在闯红灯行人数量 $N_c=0$ 和 $N_c \geq 0$ 时,0~80 s不同等待时间下到达交叉口的步行行人选择闯红灯的行人数量与等待红灯行人数量,利用上述生存分析方法,结果图4所示。此处的数据调查与图3的数据调查同时进行。

图4所示,为信号交叉口处不同闯红灯行人数量下的生存函数。从图上可以看到,当步行行人到达信号交叉口处时,不同的闯红灯行人数量对行人选择闯红灯行为的影响,正在闯红灯的行人数量越多,则行人的生存率越低,即到达信号交叉口的步行行人选择闯红灯的概率越高。

计算得图4中曲线的一元线性回归方程分别为 $S=0.813-0.008 t(N_c=0$ 时); $S=0.448-0.013 t(N_c \geq 1$ 时)。

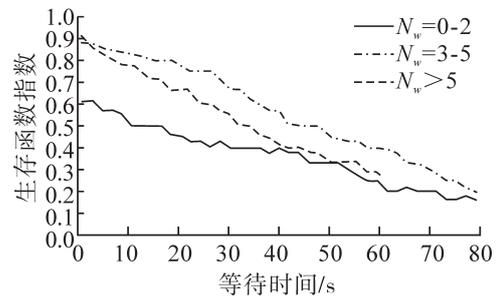


图3 信号交叉口不同等待行人数量下的生存函数

Fig. 3 Survival function under different number of pedestrians waiting at signal intersection

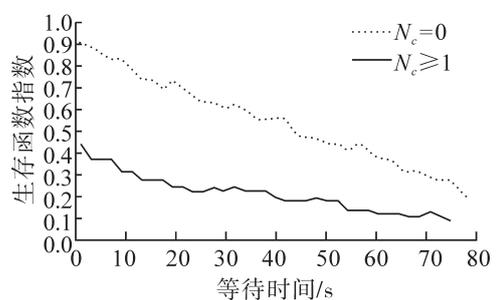


图4 信号交叉口不同闯红灯行人数量下的生存函数

Fig. 4 Survival function under different number of pedestrians who ran red light at signal intersection

4 结论

1) 利用归因理论研究步行行人闯红灯的影响因素,提出步行行人闯红灯归因理论,定性地分析影响因素。

2) 通过生存分析方法,利用非参数估计方法,以生存函数为目标函数,定量地分析出相应影响因素对步行行人闯红灯的影响规律。

3) 根据调查可知,35%左右的步行行人偏好风险决策,在红灯期间到达交叉口时,不愿等待,违规过街;约11%的步行行人属于非风险者,遵守交通法规,一直等待到绿灯方过街。

4) 将生存函数曲线用一元线性回归方程表达出来,将生存函数与影响因素之间的关系简单明了化。

该研究可为交叉口处步行行人的红绿灯设置提供一定的参考。这里只对行人闯红灯的2个重要影响因素进行了分析,可以针对冲突方向的机动车流量对行人选择闯红灯行为的概率影响,及针对控制这些影响因素的措施进行进一步研究。

参考文献:

- [1] 张立伟."中国式闯红灯"背后的弱势行人权[N].21世纪经济报道,2012-10-26(004)
- [2] 裘晨璐,许卉莹,邵志骅,等.新交通法规对机动车闯红灯违法影响效果[J].中国公共安全:学术版,2014(1):84-87.
- [3] 许可,刘沛.徐州市行人闯红灯问题研究[J].中外企业家,2013(33):199-202.
- [4] 范春明,彭囿朗.基于"刺激-反应"理论的行人从众闯红灯过街行为分析[J].环球赛鸽科技,2013(5):90-119.
- [5] 周竹萍,王炜,任刚,等.信号控制交叉口行人个体过街行为选择模型[J].东南大学学报:自然科学版,2013(3):664-668.
- [6] 张兴红.动机归因理论在成人学习中的应用[J].社会心理科学,2006(1):67-70.
- [7] 郭念峰.心理咨询师[M].上册.北京:北京民族出版社,2002:79-82.
- [8] 时蓉华,崔丽娟,李凌.现代社会心理学[M].上海:华东师范大学出版社,2007:220-223.
- [9] 卞兆洋.因子分析在城市轨道交通发展评价中的应用[J].华东交通大学学报,2013,30(4):40-45.
- [10] 周致纳,史忠科,李迎峰.行人群体闯红灯行为决策模型[J].系统工程理论与实践,2009,29(11):177-182.
- [11] 刘光新,李克平,孙剑.信号控制交叉口行人过街等待时间研究[J].中国安全科学学报,2009,19(9):159-166,179.
- [12] 余翠玲,毕新华,齐晓云.基于生存分析方法的企业信息技术采纳研究[J].科研管理,2011,32(9):83-90,124.
- [13] 环梅,杨小宝,贾斌.基于生存分析方法的非机动车闯红灯行为研究[J].北京理工大学学报,2013,33(8):815-819.
- [14] GUO H W, GAO Z Y, YANG X B, et al. Modeling pedestrian violation behavior at signalized crosswalk in developing countries: a hazard based duration approach [J]. Traffic Injury Prevention, 2011, 12(1):96-103.

Study of Pedestrians' Behaviors of Running Red Light Based on Survival Analysis Method

Wang Haichan, Zheng Changjiang

(College of Civil and Transportation Engineering, Hohai University, Nanjing 210098, China)

Abstract: By adopting attribution theory of Hyde to illustrate the pedestrians' crossing behaviors, this paper puts forward attribution theory of pedestrians' red light running. It analyzes the main influencing factors for pedestrians' running red light. Through survival analysis on pedestrians' behaviors of running red light, it explores influencing factors in details with examples. Results show that the proposed theory and method can be used to probe into pedestrians' behaviors of running red light effectively.

Key words: pedestrian; red light running; attribution theory; survival analysis

(责任编辑 王建华 李萍)