

Intersection classification and analysis of traffic violation behavior in Naturalistic Driving Study

Xuehan MA, Xichan ZHU, Zhixiong MA

School of Automotive Studies, Tongji University, Shanghai, China, 201804

Email: 18818260059@163.com

Abstract: This paper classified the intersection by the geometric types, traffic control and lane types. About 400 km mileage of video data was processed and a total of 277 safe intersection through scenarios was obtained, in which crossroads and pedestrian crossing have highest proportions. Based on the 780 dangerous conditions, 421 cases of intersections through scenarios were investigated, in which crossroad and pedestrian crossing have higher proportions, and the T-intersection is far more dangerous. In dangerous cases, there are 68 involving driving violation, mainly running on solid lines, speeding and not giving way according to traffic rules; 39 cases are about other road users driving violation, mainly vehicle changing lanes in intersections, non-motor vehicles and pedestrian running on motorways, vehicle running on solid lines. Finally, binomial logistic regression model was used to analyze the factors affecting driving violation. Lane type, road type, traffic density and light are the most significant affecting factors.

Keywords: intersection type, driving violation, binomial logistic regression model

自然驾驶研究中路口通行分类及违规行为分析

马雪寒, 朱西产, 马志雄

同济大学汽车学院, 上海, 中国, 201804

Email: 18818260059@163.com

摘 要: 本文从路口几何类型、交通控制类型和车道类型三个方面, 对路口进行了分类。通过处理约 400 公里的自然驾驶视频数据, 统计出共 277 例路口通行场景, 其中十字路口与人行横道占比最高。在自然驾驶数据库 780 例危险工况的基础上, 筛选出了 421 例路口通行场景, 其中十字路口与人行横道比例较高, T 形路口干路远支路侧相对发生危险的概率较高。危险工况中, 本车违规有 68 例, 主要是压实线、超速和不按规定让行; 其他道路使用者违规 39 例, 主要是机动车路口变道、非机动车及行人走机动车道和机动车压实线。最后利用二项 logit 回归模型分析了影响违规驾驶的因素, 车道类型、道路类型、交通密度和光照是最为显著的影响因素。

关键词: 路口类型; 安全工况; 危险工况; 违规驾驶; 二项 logit 回归模型

1 引言

路口通行是交通中最为复杂的场景之一, 行人、骑车人与机动车之间的冲突频繁, 易发生交通违规和交通事故。而路口种类繁多, 不同类型路口的交通状况各不相同, 难以枚举, 给自然驾驶研究带来了很大困难。

目前市场上的 ADAS 产品, 大多在纵向或侧向上, 即车辆的速度和转向上, 提醒驾驶员或直接控制车辆, 适用于相对简单的工况。针对路口通行场景下的 ADAS, 福特提出了基于 V2I 信号灯的车速引导和基于 V2V 的左转辅助技术; 沃尔沃 City Safety 能够探测碰撞轨迹中的迎面来车, 存在碰撞危险时, 发出警告或制动干预, 但使用条件有诸多限制。总体而言, 针对路口通行的产品开发还处于初级阶段, 无法应对复杂的交通状况。

随着 ADAS 系统的发展, 可以预见未来将出现完全由人类驾驶员控制的车辆与 ADAS 辅助驾驶的车辆(或自动驾驶车辆)通过同一路口的情况, ADAS 系统能否准确识别人类驾驶员意图、恰当处理其他道路参与者发生违规行为的情况, 极大的影响着车辆的安全性、舒适性和驾乘人员对系统的接受度。国内外学者对路口危险情况进行了大量研究, 但对于驾驶员的违规行为研究较少。然而路口通行的违规行为层出不穷, 极大影响了交通效率和驾驶体验。

本文将从自然驾驶研究的角度，提出路口分类方法并结合自然驾驶数据进行相关研究；利用二项 logit 模型分析自然驾驶中路口通行违规行为的影响因素。

2 路口分类方法

路口种类众多、交通环境复杂，常涉及交通规则及路权的问题，本文基于自然驾驶研究提出一种路口分类方法。人类驾驶员在路口通行中，首先根据路口几何类型决定可选的驾驶路径；其次，在路口几何类型基础上，根据有无红绿灯选择基本的通行准则；最后，在即将进入路口时选择合适的车道。因此路口通行的要素归纳为：路口几何类型，交通控制类型与车道类型。

根据路口的几何类型可以进行如表 1 的分类。部分主路与支路交叉或者次干路与支路交叉，是干路中心隔离封闭、支路只准右转通行的交叉路，本文将其列为右进右出交叉口。城市中无信号灯控制的人行横道较为常见，也将其列为一种几何类型^[8]。T 形路口三个方向通行方式分别为：直行/右转/掉头（主路上近支路一侧，TInear），直行/左转/掉头（主路上远支路一侧，TIfar），左转/右转/掉头（支路 Tlbp, bypass）。

Table 1. Intersection geometry type

表 1. 路口几何类型及其代码

交叉口类型	代号
十字形交叉	I(Intersection)
T 形交叉	TI (T Intersection)
Y 形交叉	YI (Y Intersection)
X 形交叉	XI (X Intersection)
多叉形交叉	MI (Multi-Intersection)
错位交叉	OI (Offset Intersection)
环形交叉	RI (Rotary Intersection)
右进右出交叉	RIROI (Right-in& Right-out Intersection)
人行横道	PC (Pedestrian Crossing)
道路与铁路平面交叉	GC (Grade Crossing)

其次，将交通控制类型分为信号灯控制交叉口和无信号灯控制交叉口。

Table 2. Intersection traffic control type

表 2. 路口交通控制类型及代码

交通管理类型	代号
信号灯控制交叉口	S (Signalized Intersection)
无信号灯控制交叉口	US (Unsignalized Intersection)

第三个影响要素为车道类型。本文将人行横道、非机动车道、行人非机动车合用车道和非结构化道路也纳入车道类型体系中。

Table 3. Intersection lane type

表 3. 路口车道类型及其代码

车道类型	代号
直行车道	SL (Straight Lane)
左转专用车道	LL (Left-turn Lane)
右转专用车道	RL (Right-turn Lane)
掉头专用车道	UL (U-turn Lane)
直行左转合用车道	SLL (Straight& Left-turn Lane)
直行右转合用车道	SRL (Straight& Right-turn Lane)
直行掉头合用车道	SUL (Straight& U-turn Lane)
左转掉头合用车道	LUL (Left-turn& U-turn Lane)
左转弯右转弯合用车道	LRL (Left-turn& Right-turn Lane)

左转直行右转合用车道	LSRL (Left-turn& Straight& Right-turn Lane)
无导向车道	NGL (No Guidance Lane)
可变车道	RvL (Reversible Lane)
人行横道	PC (Pedestrian Crossing)
非机动车道	NML (Non-motor Lane)
行人非机动车合用车道	PNML (Pedestrian& Non-motor Lane)
非结构化道路 (无车道线)	UR (Unstructured Road)

据此，即可用字母组合确定某一路口类型，如“I-S-SL”表示信号灯控制的十字路口，直行车道：I表示十字路口，S表示信号控制，SL表示直行车道。同一类型的路口，通行准则相同，控制策略相似。路口 ADAS 系统的开发中，可重点研究出现频率较高、危险频繁发生的路口类型。

3 路口类型分布及违规行为统计

3.1 China-FOT 数据简介

China-FOT 项目从 2014 年 7 月持续至 2015 年 12 月，分 4 个阶段进行，每个阶段 8 至 10 辆车参加测试，共 32 名驾驶员参与测试。试验车辆为沃尔沃集团提供的 S60L。32 名驾驶员的家庭住址主要分布在上海西部，其中 7 名为女性，25 名为男性，均具有一定的驾驶经验。数据采集了驾驶员行为数据，车辆动力学参数及道路交通环境数据。

Table 4. Data and data source

表 4. 数据及数据来源

数据	数据来源
驾驶员行为	CAN 总线，车内摄像头，踏板摄像头
车辆动力学参数	CAN 总线，外接加速度传感器
道路交通环境数据	前置摄像头，前置广角摄像头。

China-FOT 项目可用数据约 3559 小时，总行程 129109 公里。本文利用的自然驾驶数据均来自中国大型实车路试项目（China-FOT）。

3.2 路口类型分布及违规行为统计

通过观看 China-FOT 视频，统计里程约 400 公里，包含城市和郊区道路。共有 277 例路口通行场景，涉及 30 种路口类型。接下来，将逐层进行统计。

“路口几何-交通控制-车道类型”分布情况如表 5：信号灯控制的十字路口、直行车道最多，共 67 例，占比 24.19%；其次，信号灯控制的十字路口、直行右转合用车道共 42 例，占比 15.16%；无信号灯控制的人行横道、无导向车道共 39 例，占比 14.08%。

Table 5. Geometry- traffic control- lane classified intersection distribution

表 5. 路口几何-交通控制-车道类型分布表

路口类型代码	路口类型	频数	占比
I-S-SL	十字路口-信号灯控制-直行车道	67	24.19%
I-S-SRL	十字路口-信号灯控制-直行右转合用车道	42	15.16%
PC-US-NGL	人行横道-无信号灯控制-无导向车道	39	14.08%
I-US-NGL	十字路口-无信号灯控制-无导向车道	22	7.94%
I-S-SLL	十字路口-信号灯控制-直行左转合用车道	17	6.14%
I-S-LL	十字路口-信号灯控制-左转专用车道	16	5.78%
I-S-NGL	十字路口-信号灯控制-无导向车道	13	4.69%

TInear-US-NGL	T 形路口干路近支路侧-无信号灯控制-无导向车道	11	3.97%
其他		22	7.94%
总计		277	100.00%

进一步统计“路口几何-交通控制类型”分布。信号灯控制的十字路口比例较多，共 163 例，占比 58.84%。无信号灯控制的人行横道 39 例，占比 14.08%。在郊区，主干道上常常设置此类人行横道。无信号灯控制的十字路口 28 例，占比 10.11%。

Table 6. Geometry- traffic control classified intersection distribution
表 6. 路口几何-交通控制类型分布表

路口类型代码	路口类型	频数	占比
I-S	十字路口-信号灯控制	163	58.84%
PC-US	人行横道-无信号灯控制	39	14.08%
I-US	十字路口-无信号灯控制	28	10.11%
TInear-US	T 形路口干路近支路侧-无信号灯控制	14	5.05%
TInear-S	T 形路口干路近支路侧-信号灯控制	9	3.25%
TIfar-S	T 形路口干路远支路侧-信号灯控制	8	2.89%
其他		16	5.78%
总计		277	100.0%

最后统计了“路口几何类型”分布。十字路口 191 例，占比 68.95%；人行横道 42 例，占比 15.16%；T 形路口近支路端 23 例，占比 8.30%；所有 T 形路口共 38 例，占比 13.72%。这三类路口总占比 97.83%。

Table 7. Geometry classified intersection distribution
表 7. 路口几何类型分布表

路口类型代码	路口类型	频数	占比
I	十字路口	191	68.95%
PC	人行横道	42	15.16%
TInear	T 形路口干路近支路侧	23	8.30%
TIfar	T 形路口干路远支路侧	11	3.97%
RIROI	右进右出路口	4	1.44%
TIbp	T 形路口支路端	4	1.44%
RI	环形路口	2	0.72%
总计		277	100.00%

路口 ADAS 系统开发中，尤其需要关注十字路口和人行横道的通行。其中信号灯控制的十字路口和无信号灯控制的人行横道，占比尤其高。

此次数据统计中，驾驶员大多遵守交通规则，本车明显违章行为共 15 例，其中压实线 10、压导向车道线变道 5 例。考虑到数据样本量有限，个别驾驶员的不良习惯会大大增加某类违章行为的数量，因此在后续的违规行为分析中，不使用以上违规行为数据。

3.3 危险工况下路口类型分布及违规行为统计

3.3.1 危险工况筛选

事实上, China-FOT 数据量庞大, 人工处理效率较低, 难以获取整体数据。而危险是行车中应当极力避免的情况, 也是国内外学者的研究重点。孙晓宇^[5]采用三层危险场景筛选标准提取了出整个 China-FOT 数据中的危险工况。由于此次数据中不含事故, 以 SeMiFOT^[6]中的分级为基础, 主要关注其中的等级 2-等级 4, 即 2.接近型干涉、3.碰撞相关干涉和 4.近碰撞。第一层利用纵向加速度、横向加速度、横摆角速度与速度之间的关系初次制定危险标准及危险等级(共三级); 第二层通过平均制动压力变化率及 TTC (Time to Collision) 值进行等级修正; 最后利用模糊综合评价法, 对上述五个变量进行综合评价, 并结合不同速度制定不同的危险等级阈值^[5]。最终采集到 780 例危险工况, 筛选精度为 85.10%, 同时截取了纵向加速度峰值出现时刻的前 10s 以及后 5s 作为危险场景数据片段。

本文基于 780 例危险工况, 统计危险工况下各路口类型分布及路口通行中的违规行为。

3.3.2 危险工况下路口类型分布

780 例危险工况中, 共有 421 例发生在路口区域, 占比 53.97%; 能准确辨识路口几何类型、交通控制类型和车道类型的有 399 例, 共涉及 49 种路口类型。

“路口几何-交通控制-车道类型”分布情况如表 8: 信号灯控制的十字路口、直行车道最多, 共 81 例, 占比 20.30%; 其次是信号灯控制的十字路口、直行右转合用车道, 共 56 例, 占比 14.04%; 无信号控制的人行横道、无导向车道, 共 50 例, 占比 12.53%。

Table 8. Geometry- traffic control- lane classified intersection distribution under dangerous condition
表 8. 危险工况下路口几何-交通控制-车道类型分布表

路口类型代码	路口类型	频数	占比
I-S-SL	十字路口-信号灯控制-直行车道	81	20.30%
I-S-SRL	十字路口-信号灯控制-直行右转合用车道	56	14.04%
PC-US-NGL	人行横道-无信号灯控制-无导向车道	50	12.53%
I-S-SLL	十字路口-信号灯控制-直行左转合用车道	35	8.77%
I-S-LL	十字路口-信号灯控制-左转专用车道	32	8.02%
I-S-NGL	十字路口-信号灯控制-无导向车道	17	4.26%
TIfar-S-SL	T 形路口干路远支路侧-信号灯控制-直行车道	13	3.26%
I-S-RL	十字路口-信号灯控制-右转专用车道	10	2.51%
其他		41	10.28%
总计		399	100.00%

前文统计的 400 公里是整体自然驾驶数据的随机抽样, 具有一定代表性, 后文均简称整体分布。

路口类型整体分布与危险工况下的分布相比较可知, 整体分布与危险工况分布排名前 3 的路口类型相同。信号灯控制的十字路口、直行车道(I-S-SL), 信号灯控制的十字路口、直行右转合用车道(I-S-SRL), 无信号灯控制的人行横道、无导向车道(PC-US-NGL), 这 3 类路口在车辆行驶中出现频率较高, 发生危险的概率较高。但是比例有所不同, 危险工况占比偏低, 说明此类路口频繁出现, 驾驶员比较熟悉, 相对发生危险的概率有所降低。相对发生危险的概率为某一路口类型在危险工况下占比与整体分布下占比的比值。

Table 9. Contrast between overall intersection type distribution and distribution under dangerous condition I
表 9. 路口类型整体分布与危险工况下的分布对比 I

安全工况下路口类型		危险工况下路口类型	
I-S-SL 十字路口-信号灯控制-直行车道	24.19%	I-S-SL 十字路口-信号灯控制-直行车道	20.30%

I-S-SRL 十字路口-信号灯控制-直行右转合用车道	15.16%	I-S-SRL 十字路口-信号灯控制-直行右转合用车道	14.04%
PC-US-NGL 人行横道-无信号灯控制-无导向车道	14.08%	PC-US-NGL 人行横道-无信号灯控制-无导向车道	12.53%
总计	53.43%	总计	47.68%

统计“路口几何-交通控制类型”分布，信号灯控制的十字路口，共 237 例，占比 59.40%；无信号灯控制的人行横道，共 51 例，占比 12.78%。这两类路口共占比 72.18%，最容易发生危险。有信号控制的 T 形路口干路远支路侧，共 27 例，占比 6.77%。

Table 10. Geometry- traffic control classified intersection distribution under dangerous condition
表 10. 危险工况下路口几何-交通控制类型分布表

路口类型代码	路口类型	频数	占比
I-S	十字路口-信号灯控制	237	59.40%
PC-US	人行横道-无信号灯控制	51	12.78%
Tlfar-S	T 形路口干路远支路侧-信号灯控制	27	6.77%
I-US	十字路口-无信号灯控制	18	4.51%
Tlnear-S	T 形路口干路近支路侧-信号灯控制	18	4.51%
Tlfar-US	T 形路口干路远支路侧-无信号灯控制	12	3.01%
Tlnear-US	T 形路口干路近支路侧-无信号灯控制	12	3.01%
Tlbp-S	T 形路口支路端-信号灯控制	11	2.76%
其他		13	3.26%
总计		399	100.00%

对比表 11 可知，前 3 类路口类型有所不同。可见，信号灯控制的 T 形路口干路远支路侧(Tlfar-S)，虽然在整体分布中占比较低 2.89%，但在危险工况下占比 6.77%，相对发生危险的概率较高。信号灯控制的十字路口(I-S)与无信号灯控制的人行横道(PC-US)，在行车过程中出现频率较高，发生危险的概率较高，相对发生危险的概率较低。

Table 11. Contrast between overall intersection type distribution and distribution under dangerous condition II
表 11. 路口类型整体分布与危险工况下分布对比 II

安全工况下路口类型		危险工况下路口类型	
I-S 十字路口-信号灯控制	58.84%	I-S 十字路口-信号灯控制	59.40%
PC-US 人行横道-无信号灯控制	14.08%	PC-US 人行横道-无信号灯控制	12.78%
I-US 十字路口-无信号灯控制	10.11%	Tlfar-S T 形路口干路远支路侧-信号灯控制	6.77%
总计	83.03%	总计	78.95

统计“路口几何类型”分布，其中十字路口 255 例，占比最高 63.91%，其次是人行横道 52 例，占比 13.03%，T 形路口干路远支路侧 9.77%。所有 T 形路口占比 20.55%。发生危险的路口几何类型，以十字路口、人行横道、T 形路口为主，共占比 97.49%。

Table 12. Geometry classified intersection distribution under dangerous condition
表 12. 危险工况下路口几何类型分布表

路口类型代码	路口类型	频数	占比
I	十字路口	255	63.91%
PC	人行横道	52	13.03%
Tlfar	T 形路口干路远支路侧	39	9.77%
Tlnear	T 形路口干路近支路侧	30	7.52%

Tlbp	T 形路口支路端	13	3.26%
YI	Y 型路口	4	1.00%
RI	环形路口	3	0.75%
RIROI	右进右出口	3	0.75%
总计		399	100.00%

对比表 13 可知，十字路口和人行横道在行车过程中出现频率较高，发生危险的概率也较高，而相对发生危险的概率较低。T 形路口干路远支路侧(Tlfar)，整体分布中占比 3.97%，危险工况下占比 9.77%，相对发生危险的概率较高。T 形路口干路近支路侧(Tlnear)，整体分布中占比 8.3%，危险工况下占比 7.52%，相对发生危险的概率较低。

Table 13. Contrast between overall intersection type distribution and distribution under dangerous condition III

表 13. 路口类型整体分布与危险工况下分布对比 III

安全工况下路口类型		危险工况下路口类型	
I 十字路口	68.95%	I 十字路口	63.91%
PC 人行横道	15.16%	PC 人行横道	13.03%
Tlnear T 形路口干路近支路侧	8.3%	Tlfar T 形路口干路远支路侧	9.77%
总计	92.41%	总计	86.71%

3.3.3 危险工况下违规行为

根据《中华人民共和国道路交通安全法实施条例》和《上海市道路交通管理条例》，结合自然驾驶数据，对各类违规行为做出如下分类。其中本车违规 10 类，其他交通参与者（包括行人、非机动车和机动车）违规共 15 种。

Table 14. Subject vehicle traffic violation type list

表 14. 本车违规类型表

违规类型	代码
压实线	1
超速	2
不按规定让行	3
连续变道	4
进入堵塞路口	5
逆行	6
路口超车	7
路口变道	8
不按导向行驶	9
走非机动车道	10

Table 15. other road user traffic violation type list

表 15. 其他交通参与者违规类型

违规类型	代码
机动车违规停车	0
机动车压实线	1
机动车超速	2
机动车不按规定让行	3
机动车连续变道	4
机动车进入堵塞路口	5

机动车逆行	6
机动车路口超车	7
机动车路口变道	8
机动车不按导向行驶	9
机动车走非机动车道	10
非机动车及行人走机动车道	11
非机动车及行人不走人行横道，横穿车道	12
非机动车及行人逆行	13
非机动车及行人路口内通行	14

对超速进行说明：信号灯控制的路口通行，转弯时，车速不超过 30km/h；直行时，不得超过标志标线限速及规定限速。无信号灯控制的路口通行，转弯时，车速不超过 30km/h；直行时，不得超过标志标线限速及规定限速。

基于以上界定，通过视频观察和 CAN 总线数据分析，得到本车各违规类型分布。此处二次违规表示车辆路口通行时有两种及以上的违规行为。

Table 16. Subject vehicle traffic violation type distribution
表 16. 本车违规类型分布表

违规类型	频数	占比
压实线	21	30.88%
路口变道	11	16.18%
不按导向行驶	6	8.82%
超速	5	7.35%
逆行	5	7.35%
路口超车	5	7.35%
二次违规	5	7.35%
未按规定让行	3	4.41%
连续变道	3	4.41%
走非机动车道	3	4.41%
进入堵塞路口	1	1.47%
总计	68	100.00%

压实线有 21 例，占比最高 30.88%，此类违规危险性相对较小、随意性高，受驾驶员习惯影响较大。路口变道有 11 例，占比 16.18%，主要因为驾驶员常在路口变到车辆较少的车道上，受交通情况影响较大。不按导向行驶有 6 例，占比 8.82%，驾驶员为了快速通过路口，在左转专用车道直行，个别驾驶员临时改变驾驶意图。

其他交通参与者各违规类型分布如表。其中二次违规表示包含两种及以上违规行为，可能由多个对象导致。

Table 17. other road user traffic violation type distribution
表 17. 其他交通参与者各类违规行为分布表

违规类型	频数	占比
机动车路口变道	8	20.51%

非机动车及行人走机动车道	8	20.51%
机动车压实线	5	12.82%
非机动车及行人不走人行横道，横穿车道	4	10.26%
机动车违规停车	3	7.69%
机动车逆行	3	7.69%
非机动车及行人路口内通行（左转，直行等）	3	7.69%
机动车不按导向行驶	2	5.13%
二次违规	2	5.13%
非机动车及行人逆行	1	2.56%
总计	39	100.00%

机动车路口变道有 8 例，占比 20.51%；路口变道危险性较大，但驾驶员为追求快速路口通行而常进行此类的操作。行人及非机动车走机动车道有 8 例，占比 20.51%。行人通行行为随意，受道路设计、信号灯周期设计等因素影响较大。机动车压实线有 5 例，占比 12.82%，随意性较大，受驾驶员驾驶习惯影响较大。

4 危险工况下违规行为影响因素相关性分析

780 例危险工况是从 32 名驾驶员、约 13 万公里的数据中筛选得到，具有相当的统计学意义。400 公里自然驾驶数据中，驾驶员数量偏低，受个别驾驶员驾驶风格影响较大。因此仅采用危险工况下的违规行为数据进行分析。780 例中有 421 例发生在路口，其中有 68 例危险工况涉及本车违规。

路口通行危险工况中的违规驾驶与驾驶员习惯、交通环境、自然环境等影响因素有关。本文利用二项 logit 模型分析驾驶员违规驾驶影响因素及其影响程度。

4.1 二项 logit 简介

4.1.1 回归模型

应变量 Y 服从二项分布，取值为 0 或 1；Y=1 的概率为 $\pi(Y=1)$ ，则 n 个自变量 X_0 、 X_1 、 X_2 ... X_n 所对应的 logit 回归模型为^[18]：

$$\pi(Y=1) = \frac{\exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n)}{1 + \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n)} \quad (1)$$

$$\log it[\pi(Y=1)] = \ln\left[\frac{\pi(Y=1)}{1 - \pi(Y=1)}\right] = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n \quad (2)$$

β_0 为常数项， β_0 、 β_1 、 β_2 ... β_n 是 X_0 、 X_1 、 X_2 ... X_n 对应的偏回归系数，由极大似然估计得到。

优势（Odds）指事件出现的概率与非事件期望出现的概率之比。二项 logit 回归模型中优势即为：

$$O = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \dots + \beta_n X_n) \quad (3)$$

p 为该模型下，某个体概率 π 的估计值，有

$$p = \frac{O}{1 + O} \quad (4)$$

4.1.2 单个自变量的检验假设

采用 Wald 统计量检验 X_i 的偏回归系数 β_i 是否为 0^[18]，

$$W_i = [\beta_i / SE(\beta_i)]^2 \quad (5)$$

自由度为 1， $SE(\beta_i)$ 为 X_i 的标准误。

如果拒绝假设 $H_0: \beta_i = 0$ ，表明该自变量对于模型的作用有统计学意义，也即有相关性。

4.1.3 模型拟合优度的评价

决定系数 R^2 反应了模型中所有自变量解释应变量 Y 变异的百分比^[18]，其值越接近 1，自变量预测能力越强。Cox and Snell 决定系数公式为

$$R_{CS}^2 = 1 - \left[\frac{-2LL_0}{-2LL_1} \right]^{2/n} \quad (6)$$

其中， n 为观察个体数， $-2LL_0$ 是只有常数项的 -2 倍对数似然值， $-2LL_1$ 为包含所有自变量的 -2 倍对数似然值。

Nagelkerke 决定系数公式为

$$R_n^2 = \frac{R_{CS}^2}{R_{CS}^2 \text{ 的最大可能值}} = \frac{1 - \left[\frac{-2LL_0}{-2LL_1} \right]^{2/n}}{1 - (-2LL_0)^{2/n}} \quad (7)$$

但此类决定系数并非真正的决定系数，而是伪决定系数，只能作为模型拟合优度的参考。

Hosmer-Lemshow 拟合优度检验将观察对象分成 k 组，比较各组应变量分类的实际观察频数与预测期望频数，通过 Pearson 卡方检验获得 Hosmer-Lemshow 统计量，检验统计量服从自由度为 $k-2$ 的卡方分布。检验结果无统计学意义，则表示模型预测值与观察值之间的差异没有统计学意义，模型拟合优度较好^[18]。

4.2 危险工况下的违规行为二项 logit 回归模型建立及分析

4.2.1 违规行为回归模型建立

本数据库驾驶员年龄大多在 30 岁左右，教育水平相当，车辆品牌均为沃尔沃，因此不单独设置相关变量。考虑到有些驾驶员是在其他道路使用者的影响下，为躲避危险被动违规，因此设置“违规时驾驶员状态”，“是否被动违规”两变量。此外，交通标志标线清晰度取 0，与强结构化道路对应；取 1，与弱结构化道路对应；取 2 与非结构化道路对应。

应变量为本车是否违规，本车没有违规取值 1、违规取值 0。

自变量涉及本车驾驶员相关自变量、自然环境自变量、交通环境自变量和自然驾驶相关自变量。

Table 18. Drive related variable
表 18. 本车驾驶员相关自变量取值表

自变量	自变量取值
性别	0: 女; 1: 男
是否有第二行为	0: 无; 1: 有
违规时驾驶员状态	0: 匀速或加速, 无调整意识; 1: 有减速或转向调整意识

Table 19. Natural environment related variable
表 19. 自然环境自变量取值表

自变量	自变量取值
天气	0: 雨雪雾; 1: 晴朗
光照	0: 白天; 1: 夜晚有灯; 2: 夜晚无灯

Table 20. Traffic environment related variable
表 20. 交通环境自变量取值表

自变量	自变量取值
路口几何类型	0: 十字; 1: T 字; 2: PC; 3: 其他
有无交通信号控制	0: 无; 1: 有
车道类型	0: 专用车道; 1: 合用车道; 2 无导向车道
道路类型	0: 城郊或乡村道路; 1: 城区道路
交通标志标线清晰度	0: 标线清晰、标志意义明确; 1 标线不清晰、标志意义模糊或无标志; 2: 无
交通密度	0: 稀疏; 1: 正常; 2: 拥挤

Table21. Naturalistic driving related variable
表 21. 自然驾驶自变量取值表

自变量	自变量取值
驾驶意图	0: 直行; 1: 右转; 2: 左转或掉头; 3: 不确定
被动违规	0: 否; 1: 是

根据上述变量标准统计 68 例违规数据。同时在 421 例场景中，随机选取 49 例未违规场景进行变量统计，总计 117 例。

利用 SPSS 进行二项 logit 建模。其中路口几何类型、车道类型、交通标志标线清晰度、交通密度和驾驶意图是分类变量，在分析时将这类变量哑变量化，均以第一项（编码为 0 的项）为参考，如图 1 所示。由于样本中没有出现夜晚无灯的情况，因此光照不再是分类变量。

“驾驶员性别”中女性共 12 例；“是否有第二行为”中有第二行为共 5 例；“违规时驾驶员状态”中进行调整的共 5 例；“天气”中雨雪共 4 项；“被动违规”中属于被动违规的共 5 例，且有 3 例进行了调整，这两类变量可能存在多重共线性，剔除“被动违规”。

初步建立模型计算得，以上 5 类变量与是否违规驾驶之间的联系无统计学意义，显著性值较高。为提高模型拟合优度，删除变量“驾驶员性别”、“是否有第二行为”、“违规时驾驶员状态”、“天气”，再次建立模型。

類別變數編碼

		次數	參數編碼		
			(1)	(2)	(3)
駕駛意图	0	63	.000	.000	.000
	1	7	1.000	.000	.000
	2	10	.000	1.000	.000
	3	37	.000	.000	1.000
路口几何类型	0	66	.000	.000	.000
	1	32	1.000	.000	.000
	2	14	.000	1.000	.000
	3	5	.000	.000	1.000
交通密度	0	50	.000	.000	
	1	49	1.000	.000	
	2	18	.000	1.000	
车道类型	0	50	.000	.000	
	1	32	1.000	.000	
	2	35	.000	1.000	
标线清晰度	0	66	.000	.000	
	1	48	1.000	.000	
	2	3	.000	1.000	

Figure 1. SPSS dummy variable code

图 1. SPSS 哑变量编码

4.2.2 建模结果及分析

模型决定系数参数 Cox and Snell 决定系数为 0.389, Nagelkerke 决定系数为 0.523。

Hosmer-Lemshow 拟合优度检验, 显著项为 0.580>0.05, 说明模型预测值与观察值之间的差异无统计学意义, 模型拟合优度较好。未剔除变量时, 显著项为 0.155, 拟合优度大大提高。

方程式中的變數

	B	S.E.	Wald	df	顯著性	Exp(B)	95% EXP(B) 之信賴區間	
							下限	上限
步驟 1 ^a 路口几何类型			6.191	3	.103			
路口几何类型(1)	-.151	.697	.047	1	.828	.859	.219	3.369
路口几何类型(2)	-.727	1.312	.307	1	.580	.483	.037	6.328
路口几何类型(3)	-4.429	1.802	6.042	1	.014	.012	.000	.408
信号灯	1.721	1.136	2.295	1	.130	5.590	.603	51.816
车道类型			9.813	2	.007			
车道类型(1)	-.139	.644	.047	1	.829	.870	.246	3.076
车道类型(2)	4.386	1.492	8.645	1	.003	80.350	4.316	1495.743
城区与郊区	1.440	.710	4.116	1	.042	4.220	1.050	16.957
标线清晰度			1.940	2	.379			
标线清晰度(1)	-.438	.537	.667	1	.414	.645	.225	1.848
标线清晰度(2)	-2.526	1.983	1.622	1	.203	.080	.002	3.899
交通密度			8.859	2	.012			
交通密度(1)	1.662	.598	7.727	1	.005	5.268	1.632	17.003
交通密度(2)	-.053	.791	.004	1	.947	.949	.201	4.469
光照	-2.396	1.127	4.516	1	.034	.091	.010	.830
驾驶意图			3.768	3	.288			
驾驶意图(1)	.750	1.117	.451	1	.502	2.117	.237	18.892
驾驶意图(2)	.728	1.066	.466	1	.495	2.071	.256	16.726
驾驶意图(3)	-.837	.588	2.025	1	.155	.433	.137	1.371
常数	-2.038	1.392	2.143	1	.143	.130		

a. 步骤 1 上輸入的變數: [X1, 1]

Figure 2. SPSS model fitting parameters

图 2. SPSS 模型拟合参数

模型参数解释如下:

B: 前文所求得得的 β ;

S.E.: 参数的渐进标准误, 由 Newton-Raphson 迭代产生的信息矩阵之逆矩阵的对角元素开方获得;

Wald: Wald 卡方值, 验证选取的变量具有统计学意义, B 值是否为 0。

df: 自由度;

显著性: 检验 P 值; 如“道路类型 (城区与郊区)”变量, 在 Wald 卡方值 4.116、自由度为 1 时对应的检验 P 值为 0.042。

Exp(B): 每个自变量的优势比 OR 值; 如“道路类型 (城区与郊区)”变量, Exp(B)=4.220, 表示城市道路上违规驾驶对郊区道路上违规驾驶的优势比为 4.220。

根据计算结果所得 β 值, 建立预测概率模型:

$$p = \frac{\exp(-2.038 - 0.151X_{1a} - 0.727X_{1b} - 4.429X_{1c} + 1.721X_2 - 0.139X_{3a} + 4.386X_{3b} + 1.440X_4 - 0.438X_{5a} - 2.526X_{5b} + 1.662X_{6a} - 0.053X_{6b} - 2.396X_7 + 0.750X_{8a} + 0.728X_{8b} - 0.837X_{8c})}{1 + \exp(-2.038 - 0.151X_{1a} - 0.727X_{1b} - 4.429X_{1c} + 1.721X_2 - 0.139X_{3a} + 4.386X_{3b} + 1.440X_4 - 0.438X_{5a} - 2.526X_{5b} + 1.662X_{6a} - 0.053X_{6b} - 2.396X_7 + 0.750X_{8a} + 0.728X_{8b} - 0.837X_{8c})} \quad (8)$$

公式变量解释:

X_{1a}: 路口几何类型参数编码 (1)

X_{1b}: 路口几何类型参数编码 (2)

X_{1c}: 路口几何类型参数编码 (3)

X₂: 信号灯

X_{3a}: 车道类型参数编码 (1)

X_{3b}: 车道类型参数编码 (2)

X₄: 道路类型 (城区或郊区)

X_{5a}: 标线清晰度参数编码 (1)

X_{5b}: 标线清晰度参数编码 (2)

X_{6a}: 交通密度参数编码 (1)

X_{6b}: 交通密度参数编码 (2)

X₇: 光照

X_{8a}: 驾驶意图参数编码 (1)

X_{8b}: 驾驶意图参数编码 (2)

X_{8c}: 驾驶意图参数编码 (3)

根据统计学知识, 由每一个自变量对应的显著性值可知, 在 0.05 检验水平下, 车道类型、道路类型 (城区与郊区)、交通密度、光照与是否违规驾驶之间的联系有统计学意义; 路口几何类型、有无交通信号灯控制、交通标志标线清晰度、驾驶意图与是否违规驾驶之间的联系无统计学意义。

车道类型以专用车道为标准, 合用车道上相对标准偏向不发生违规 (-0.139), 无导向车道相对标准偏向违规 (4.386)。-0.139 数值较小, 说明差别并不太大, 合用车道由于通行方向为两个及两个以上, “不按导向行驶”可能性有所降低。无导向车道, 没有通行方向的限制, 驾驶随意性高 (共有 30 例违规行为, 涉及 7 种违规类型)。

道路类型 (城区与郊区) B 为 1.440, 说明在城市更容易发生违规情况。城市交通情况复杂, 易发生违规; 郊区交通受驾驶员驾驶习惯影响较大。

交通密度以稀疏为标准, 正常交通相对标准偏向违规 (1.662), 拥堵环境下相对标准偏向不发生违规行为 (-0.053)。交通稀疏, 车辆路口变道、不按导向行驶、路口超车以快速通过路口的可能性有所降低, 正常交通有可能导致以上行为。拥堵时, 车辆活动自由降低, 驾驶员更谨慎, 超速等违规行相对交通稀疏情况有所减少; -0.053 数值较小, 差别不大。

光照 B 为 -2.396, 说明夜晚更容易发生违规驾驶。夜晚视觉效果差, 容易偏离车道, 压实线。

5 总结

本文提出确定路口类型的三个要素：几何类型、交通控制类型和车道类型。基于 China-FOT 数据，统计了 400 公里的路口类型分布，有 277 例路口通行场景，涉及 30 种类型，其中十字路口和人行横道比例较高；统计了 780 例危险工况下的路口类型分布，共 421 例，涉及 49 种路口类型。与整体分布进行对比，信号灯控制的十字路口与人行横道占整体比较高，容易发生危险，但相对发生危险的概率不高；T 形路口干路远支路相对发生危险的概率较高。统计了危险工况下的违章行为，本车违章压实线、路口变道和不按导向行驶占比最高；其他道路使用者违规行为中机动车路口变道、行人及非机动车走机动车道和机动车压实线占比最高。利用二项 logit 模型分析了违规行为影响因素，可知车道类型、道路类型（城区与郊区）、交通密度、光照与是否违规驾驶之间的联系有统计学意义

致谢 (Acknowledgement)

我在朱西产老师和马志雄老师的悉心指导下顺利完成了论文。老师们在论文选题、方案讨论、论文撰写以及审稿中给予了我极大的关心与帮助。他们丰富的专业知识、严谨的治学态度、谦逊的处事风格让我受益匪浅。感谢两位老师这些日子的精心教导与付出！

参考文献 (References)

- [1] Dingus TA, Klauer S, Neale V, et al. The 100-car naturalistic driving study, Phase II-results of the 100-car field experiment[J]. Behavior, 2006.
- [2] Characterizing Underground Utilities[J]. Shrp Report, 2009.
- [3] Chen Jun, Xu Liangjie, Zhu Shunying, et al. Traffic management and control [M]. Beijing: China Communications Press, 2012: 171-176.
- [4] 陈峻, 徐良杰, 朱顺应, 等. 交通管理与控制[M]. 北京: 人民交通出版社, 2012: 171-176.
- [5] Li Lin, Zhu Xichan, Liu Ying, etc. Typical Traffic Risk Scenarios Related to Pedal Cyclists [J]. Journal of Tongji University (Natural Science Edition), Volume 42 No. 7, 2014.7, pp1082-1087.
- [6] 李霖, 朱西产, 刘颖等, 涉及骑车人的典型交通危险场景, 同济大学学报(自然科学版), 第 42 卷第 7 期, 2014 年 7 月, pp1082-1087.
- [7] Li Lin. Research of Braking and Steering Control in Automotive Collision Avoidance [D]. Shanghai: Tongji University, Institute of Automotive Studies, 2015.
- [8] 李霖. 制动转向控制在事故回避中的应用研究[D]. 上海: 同济大学汽车学院, 2015.
- [9] Sun Xiaoyu. Research on the Causation and Incidents based on Naturalistic Driving Study [D]. Shanghai: Tongji University, School of Automotive Studies, 2017.
- [10] 孙晓宇. 基于自然驾驶研究的危险场景及诱导因素研究[D]. 上海: 同济大学汽车学院, 2017.
- [11] Trent Victor, Fredrik Moeschlin, Magnus H.etc. Sweden-Michigan Naturalistic Field Operational Test(SeMiFOT) Phase1, WP5 Evaluation of Methodology-Final Report ,SAFER REPORT 2010:06 PROJECT C3 SeMiFOT
- [12] Chen Mingyang. Application of Naturalistic Driving study in Development of Automated Vehicle [D]. Shanghai: Tongji University, School of Automotive Studies, 2017.
- [13] 陈名扬. 自然驾驶研究在自动驾驶汽车开发中的应用[D]. 上海: 同济大学, 2017.
- [14] Chen Mingyang, Zhu xichan, Ma zhixiong, et al. Brake Response Time under Near-crash Case with Cyclist. 2016 IEEE Intelligent Vehicle Symposium (IV) Gothenburg, Sweden, June 29-22, 2016.
- [15] Jonas Baergman, Kip Smith, Julia Werneke. Quantifying drivers' comfort-zone and dread-zone boundaries in left turn across path/opposite direction (LTAP/OD) scenarios [J]. Transportation Research part F 35, 2015: 170-184
- [16] Ding Zhao, Henry Lam, Huei Peng, et al. Accelerated Evaluation of Automated Vehicles Safety in Lane-Change Scenarios Based on Importance Sampling Techniques[J]. IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEM.
- [17] JOHN M. SCANLON, KRISTOFER D. KUSANO, HAMPTON C. GABLE. Analysis of Driver Evasive Maneuvering Prior to Intersection Crashes Using Event Data Records [J]. Traffic Injury Prevention, 2015: S182-S186.
- [18] Ulrich Sander. Opportunities and limitations for intersection collision intervention-A study of real world 'left turn across path' accidents [J]. Accidents Analysis and Prevention 99, 2017: 342-355.
- [19] Wassim G. Najm, John D. Smith, David L. Simth (NHTSA). Analysis of Crossing Path Crashes. 2001.7.
- [20] Road traffic signs and markings part2- road traffic signs: GB5768.2-2009 [S].
道路交通标志和标线第二部分-道路交通标志: GB5768.2-2009 [S].
- [21] Road traffic signs and markings part3- road traffic markings: GB5768.3-2009 [S].
道路交通标志和标线第二部分-道路交通标线: GB5768.3-2009 [S].
- [22] Yuan Zhifa, Zhou Jingyu. Multivariate statistical analysis [M]. Science Press, 2002-10.
袁志发, 周静宇. 多元统计分析[M]. 科学出版社, 2002-10.
- [23] Yu Chuanhua. SPSS and Statistical Analysis [M]. Electronic Industry Press, 2007-02.
宇传华. SPSS 与统计分析[M]. 电子工业出版社, 2007-02