

走行環境が自動運転自動車に与える ヒヤリハット要因の分析

田島 友祐¹・笛田 薫²・三浦 雄作³・三樹 孝博³

(受付 2025 年 7 月 1 日; 改訂 9 月 18 日; 採択 9 月 25 日)

要 旨

本論文では、自動運転車両の走行実験データを用いて、走行環境がヒヤリハット発生に与える影響を分析した。具体的には、ヒヤリハット発生時の走行環境として、走行モード、道路形状、行動類型、道路情報、周辺状況をカテゴリ変数化し、ゼロ切断ポアソン回帰モデルを構築することで、ヒヤリハット数の予測ならびにその要因を分析した。さらに、それらの2変数間の交互作用項を導入し、モデルの変数選択を行い、解釈性を検討した。その結果、ヒヤリハット発生に有意な影響を与える走行環境の組合せが明らかになった。また、走行モードによらず行動類型と周辺状況の交互作用がヒヤリハットの発生数の増加に影響することが判明した。本研究の成果は、今後、巡回型バスなどの無人自動運転サービスの安全性向上に資する知見を提供するものと思われる。

キーワード：ゼロ切断ポアソン回帰, 交互作用, リスク分析。

1. はじめに

近年、自動運転車両(Autonomous Vehicles, AV)の技術は急速に進展しており、レベル4に相当する運転手不要のモビリティサービスの実現に向けた実証実験が国内外で進められている。自動運転の実現に合わせ、法制度の整備も着実に進行している(国土交通省 2023, 2024)。自動運転の実現は、交通事故の削減、巡回バスやタクシーの増加による高齢者や交通弱者への移動支援、さらには、近年需要が高まっている物流や配達などで生じている慢性的なドライバー不足の解消といった、現代社会が抱える多くの課題への解決策として期待されている(Fagnant and Kockelman, 2015; Litman, 2020)。

滋賀大学では、あいおいニッセイ同和損害保険と連携し、群馬大学および同大学発スタートアップである日本モビリティが実施した自動運転車両の走行実験データを用いて、自動運転車両の走行ルートに関するリスク評価モデルの構築に取り組んでいる。自動運転車両においては、人間による運転ミスや判断の遅れといった従来の事故原因は回避可能と考えられる。一方で、自動運転車両に搭載されているセンサーや走行アルゴリズムに基づく判断が環境要因と複

¹ 大妻女子大学 データサイエンス学部: 〒102-8357 東京都千代田区三番町 12 番地; yusuke.tajima@otsuma.ac.jp

² 滋賀大学 データサイエンス学部: 〒522-8522 滋賀県彦根市馬場 1 丁目 1-1; kfueda@biwako.shiga-u.ac.jp

³ あいおいニッセイ同和損害保険株式会社 商品・CSV × DX 企画部: 〒150-8488 東京都渋谷区恵比寿 1-28-1; t-miki@aioinissaydowa.co.jp, yusakumiura1218@aioinissaydowa.co.jp

雑に相互作用することが考えられるが、その相互作用については未だ不明な点が多く、安全に運用するための課題が多く存在するとされている。

実運用に向けて重要なことは、事故には至らなかったものの潜在的な危険が存在する「ヒヤリハット」事象に注目することである。ヒヤリハットは、重大事故に繋がる手前の予兆を示しており、それを分析することによって、重大事故の要因を把握することができ、安全な走行ルートや運用管理が可能になると考えられる。このようなヒヤリハットの要因は、走行環境(道路形状、交通状況、周辺の構造物や人の動きなど)によって大きく影響されるため、データに基づく定量的な分析が求められる(Kim et al., 2017; 伊藤 他, 2020)。

本論文では、今後日本各地で展開される、中型のバスなどの無人自動運転モビリティサービスに向けて、走行環境がヒヤリハット発生に与える影響を明らかにすることを目的とする。具体的には、走行環境に関するカテゴリ変数(走行モード、道路形状、交通情報、周辺状況など)をもとに、ゼロ切断ポアソン回帰モデル(zero-truncated Poisson regression)を用いてヒヤリハットの件数を予測し、各要因およびその交互作用がリスクに与える影響を分析する。なお、モデルの妥当性の確認のためにゼロ切断負の二項回帰モデルによる解析結果も併用する。

2. 自動運転車両におけるヒヤリハット数予測に向けた実験およびデータ概要

2.1 自動運転車両による走行実験の概要

本研究では、群馬大学および日本モビリティ株式会社によって実施された走行実験データを用いる。このデータは、自動運転車両の走行中に発生するヒヤリハット(本論文では、車両に搭載されている加速度センサが予め設定した閾値を超過したイベントと定義する)の予測を目的として、複数地域で実施された走行実験データである。走行実験は都市部や郊外をはじめとする多様な環境で行われている。実験に用いられている自動運転車両には、加速度センサだけでなく、カメラも搭載されており、走行実験中の走行環境についても連続的に記録されている。

2.2 計測データの構成

本研究で用いた走行実験データは、2017年から2021年にかけて実施された、計55箇所での走行実験の記録である。このうち、走行実験の記録の一部に欠損があるものや、記録管理の関係で連携が困難なデータを除外し、最終的に40箇所の走行実験データを対象として分析を行った。

これらの40箇所の走行実験データ全体の総走行距離は約18,600kmであり、そのうち8,828kmが自動運転による走行である。手動運転を行う必要がある場合として、走行実験ルート内での工事区間や、やむを得ない事情によるルート変更など自動運転の実施が困難と判断される場が考えられる。なお、40箇所の走行実験で各ルートの走行距離や自動運転・手動運転の走行距離にばらつきが見られるが、40箇所すべてで見ると自動運転と手動運転がほぼ同程度の割合で実施されている。

走行実験中に用いられる車両には加速度センサおよび周囲の状況を記録するカメラが搭載されている。本研究では、そのうちヒヤリハットが発生した時点の車両情報ならびに走行環境について記録がされているデータを用いる。具体的には、自動運転か手動運転かを制御している走行モード、道路形状、行動類型、道路情報、および周辺状況に着目する。以下に、本研究で着目した各項目について説明する。

- 道路形状：単路、交差点の区別に加え、見通しの良好／不良、走行環境に関する情報を示す。
- 行動類型：車両の挙動として、直進、右折、左折の進行方向を示す。

- 道路情報：歩車区分に関する情報であり、ガードレールや段差の有無などを示す。
- 道路情報2：中央分離帯の有無やその構造に関する情報を含む。歩者区分を表す道路情報と同様にガードレールや段差の有無などを示す。
- 周辺状況：市街地か否かなど、道路周辺の土地利用に関する情報を示す。

ここで、ヒヤリハットとは、車載の加速度センサが事前に設定された閾値を超過したイベントを指し、車両のみだしや歩行者や自転車等の飛び出しに対して、事故を避けるための急ブレーキなどの急減速時に発生する。ハインリッヒの法則(Heinrich's Law) (Heinrich, 1931)やバードの法則(Bird's Accident Ratio) (Bird and Germain, 1985)にあるように1つの大きな事故の背後には軽微な事故があり、その背後にはヒヤリハットがあるということが知られている。本分析で扱う40箇所走行実験データには、合計406件のヒヤリハットが発生しており、このヒヤリハットについて走行環境をもとに発生要因についての分析を実施する。

2.3 データのカテゴリ化

本分析の実施のために、ヒヤリハット発生時の走行モード、道路形状、行動類型、道路情報、道路情報2、および周辺状況についてのカテゴリ定義を表1に示す。

ヒヤリハットとして記録された合計406件の走行環境について表1の数値変換を施し、同一の組合せパターンごとに分類した。表2はその一部であり、ヒヤリハットの数は1以上であり、そのパターン(水準組合せ)は48とおり、ヒヤリハットの数は最小値で1、最大値で74となっている。

表 1. 走行環境に関する変数のカテゴリ化。

項目	カテゴリ	対応する水準
走行モード	ヒヤリハット発生前が手動運転	0
	ヒヤリハット発生前が自動運転	1
道路形状	単路	0
	交差点	1
行動類型	直進	0
	右折	1
	左折	2
道路情報(歩車区分)	何もしないもしくは線のみ	0
	段差もしくはガードレールがある	1
道路情報2(中央分離帯)	何もしないもしくは線のみ	0
	段差もしくはガードレールがある	1
周辺状況	市街地(住宅街)	0
	市街地(住宅街でない)	1
	非市街地	2

表 2. ダミー変数の組合せごとの出現回数。

ID	走行モード	道路形状	行動類型	道路情報	道路情報2	周辺状況	出現回数
1	0	0	2	1	0	1	74
2	1	1	0	1	0	0	41
3	0	0	0	0	0	2	25
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
48	0	1	0	1	1	2	1

3. ゼロ切断ポアソン回帰

自動車の走行運転時のヒヤリハットの予測のような問題を扱う際、予め定めた単位時間あたりにある事象が生じる回数を予測する問題としてポアソン回帰 (Gourieroux et al., 1984; Long, 1997; Long and Freese, 2006; Winkelmann, 2008; Cameron and Trivedi, 2013) が用いられることが多い。ポアソン回帰は、従属変数 Y が非負の整数値をとる場合に適した回帰モデルとして知られている。ポアソン回帰モデルでは、従属変数 Y_i の期待値 $\lambda_i = E[Y_i]$ が説明変数に対して次のように定義される。

$$\lambda_i = \exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip})$$

これはリンク関数として対数リンク関数(log-link function)を用いていることを意味する。

$$\log \lambda_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}$$

このとき、 Y_i は平均 λ_i のポアソン分布に従うと仮定される。

$$Y_i \sim \text{Poisson}(\lambda_i)$$

ポアソン分布の確率分布は以下の通りである。

$$P(Y_i = y) = \frac{\lambda_i^y e^{-\lambda_i}}{y!}, \quad y = 0, 1, 2, \dots$$

しかし、今回の分析に用いるデータはヒヤリハットが発生したという条件のもとでの走行環境に関するデータであるため、ヒヤリハットが0となるデータは存在しない。そのため、ポアソン回帰を実施することは難しい。そこで、従属変数が常に正の整数をとる場合の回帰モデルであるゼロ切断ポアソン回帰(zero-truncated Poisson regression)を用いる。通常のポアソン回帰では、従属変数 Y_i は0を含む非負の整数であることを仮定しているが、観測対象が「発生した件数」のみに限定されている場合、すなわち $Y_i = 0$ がデータに含まれないとき用いるのがゼロ切断ポアソン回帰 (Gurmu, 1991; Grogger and Carson, 1991) である。

ここでは、 $Y_i > 0$ の条件のもとでの発生件数を X_i とする。 X_i がゼロ切断ポアソン分布に従うとは、次のような Y_i の条件付き分布を意味する。

$$P(X_i = x) = P(Y_i = x | Y_i > 0) = \frac{\lambda_i^x e^{-\lambda_i}}{x!(1 - e^{-\lambda_i})}, \quad x = 1, 2, 3, \dots$$

これは、 $Y_i = 0$ の確率を除いて再正規化したポアソン分布である。このとき、

$$E[X_i] = \frac{\lambda_i}{1 - e^{-\lambda_i}} \quad V[X_i] = \frac{\lambda_i}{1 - e^{-\lambda_i}} \left(1 - \frac{\lambda_i e^{-\lambda_i}}{1 - e^{-\lambda_i}} \right)$$

であるから、

$$\mu_i = \frac{\lambda_i}{1 - e^{-\lambda_i}}$$

とおき、期待値の対数を説明変数で表現する点は通常のポアソン回帰と同様である：

$$\log \mu_i = \beta^T \mathbf{x}_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip}.$$

ただし、この場合、 $\mu_i > 1$ であるので、 $\beta^T \mathbf{x}_i > 0$ となるような β を推定する必要がある。ゼロ切断ポアソン回帰における β の推定法については、様々な研究がなされている (Gurmu, 1991; Grogger and Carson, 1991; Long, 1997; Van der Heijden et al., 2003; Simonoff, 2003; Tian et al., 2018; Li et al., 2023; Godwin, 2024)。本論文では、RのパッケージであるVGAM (Yee and Ma, 2024)を用いてゼロ切断ポアソン回帰を実施する。

ヒヤリハット数の検討を進めるにあたり、ゼロ切断ポアソン回帰では過分散の影響などにより十分な当てはまりが得られない可能性がある。そこで、本研究では負の二項回帰も実施し、ゼロ切断ポアソン回帰との比較による検討も進める。負の二項回帰は、ポアソン回帰を一般化したものであり、次の式で示されるものである。ここで、 n はサイズパラメータ、 $\mu_i^{(0)}$ は平均パラメータであり、サイズパラメータによって過分散に対する調整がなされる。

$$\Pr(Y_i = x | \mu, \phi) = \frac{\Gamma(n+x)}{\Gamma(n)x!} \left(\frac{n}{\mu_i^{(0)}+n} \right)^n \left(\frac{\mu_i^{(0)}}{\mu_i^{(0)}+n} \right)^x, \quad x = 0, 1, 2, \dots$$

この場合にも、ゼロ切断ポアソン回帰と同様に観測値に 0 が現れないゼロ切断負の二項回帰モデルを用いる。このとき、

$$\begin{aligned} P(X_i = x) &= P(Y_i = x | Y_i > 0) \\ &= \frac{1}{1 - \left(\frac{n}{\mu_i^{(0)}+n} \right)^n} \frac{\Gamma(n+x)}{\Gamma(n)x!} \left(\frac{n}{\mu_i^{(0)}+n} \right)^n \left(\frac{\mu_i^{(0)}}{\mu_i^{(0)}+n} \right)^x, \quad x = 1, 2, 3, \dots \end{aligned}$$

となる。

$$\mu_i = E[X_i] = \frac{\mu_i^{(0)}}{1 - \left(\frac{n}{\mu_i^{(0)}+n} \right)^n}$$

したがって、リンク関数は

$$\log \mu_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_p x_{ip}$$

である。使用する R のパッケージは VGAM である。

4. ゼロ切断ポアソン回帰を用いたヒヤリハット数の予測

4.1 主効果のみの予測

図 1 は、走行モード、道路形状、行動類型、道路情報、道路情報 2、周辺状況のカテゴリ変数を用いてゼロ切断ポアソン回帰により推定したヒヤリハットの予測値と実測値の散布図であ

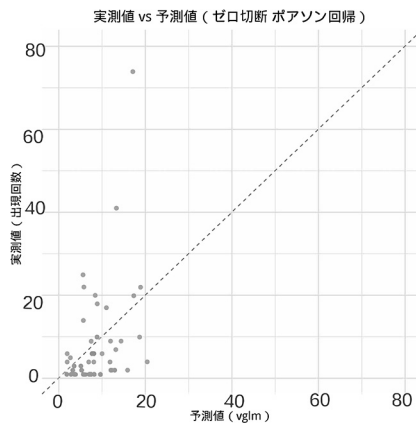


図 1. 主効果のみでのゼロ切断ポアソン回帰。実測値と予測値の比較。

る。横軸は予測値、縦軸は実測値を示しており、斜めの点線は予測値と実測値が一致する線を示している。図より、推定精度は高くなく、すべての実測値に対して予測値が0~20の範囲に集中しており、十分な推定精度をもつ推定ではないことがわかる。

4.2 交互作用項を考慮した予測

よりあてはまりのよい推定とするために、走行モード、道路形状、行動類型、道路情報、道路情報2、周辺状況のカテゴリ変数の2変数間の交互作用項を導入したモデルを用いて、改めてヒヤリハットの推定を実施する。表3は用いた交互作用項をまとめたものである。このうち、ヒヤリハットの数か0となっている交互作用項については考慮から除外している。

本論文において、3変数以上の交互作用については、分析に用いるデータ数が十分でないため、観測値が存在しない水準組合せがあり、推定不能となるため考慮しない。今後、実証実験数の増加や実環境による走行が増えた際には、3因子交互作用が及ぼすヒヤリハットへの影響を考慮することも考えられ、走行環境との影響についてより詳細な分析ができると考えられる。

ポアソン回帰では、過分散問題(over-dispersion)が指摘されることがある。そこで、ゼロ切断負の二項回帰でも推定を行い、比較検討を実施している。図2は、不要な交互作用項を除外

表3. 交互作用項の一覧。

交互作用項
走行モード_ダミー:道路形状_ダミー
走行モード_ダミー:行動類型_ダミー
走行モード_ダミー:道路情報_ダミー
走行モード_ダミー:道路情報2_ダミー
走行モード_ダミー:周辺状況_ダミー
道路形状_ダミー:行動類型_ダミー
道路形状_ダミー:道路情報_ダミー
道路形状_ダミー:道路情報2_ダミー
道路形状_ダミー:周辺状況_ダミー
行動類型_ダミー:道路情報_ダミー
行動類型_ダミー:道路情報2_ダミー
行動類型_ダミー:周辺状況_ダミー
道路情報_ダミー:周辺状況_ダミー
道路情報2_ダミー:周辺状況_ダミー

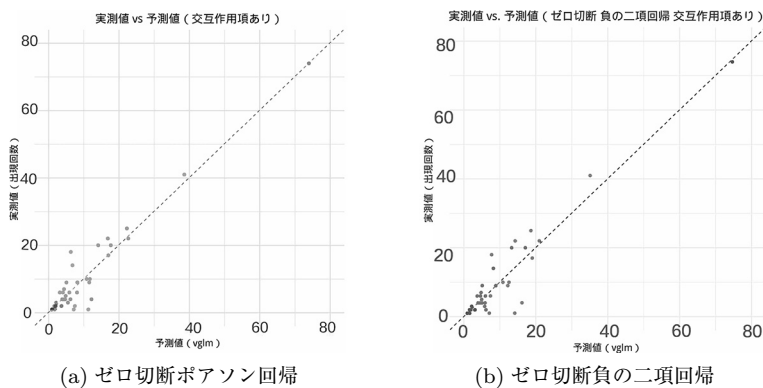


図2. 交互作用項をすべて用いた回帰。実測値と予測値の比較。

表 4. 推定結果と有意性(ゼロ切斷ポアソン回帰).

効果名	推定値	有意性
(Intercept)	2.8374	***
走行モード_ダミー 1	-0.3854	
道路形状_ダミー 1	-0.7338	*
行動類型_ダミー 1	-2.3714	
行動類型_ダミー 2	-1.6812	***
道路情報_ダミー 1	-0.1837	
道路情報 2_ダミー 1	-1.0257	**
周辺状況_ダミー 1	-1.3750	**
周辺状況_ダミー 2	0.2696	
走行モード_ダミー 1:道路形状_ダミー 1	1.1589	***
走行モード_ダミー 1:行動類型_ダミー 1	-1.6288	
走行モード_ダミー 1:行動類型_ダミー 2	-18.1724	
走行モード_ダミー 1:道路情報_ダミー 1	0.8553	*
走行モード_ダミー 1:道路情報 2_ダミー 1	-0.1119	
走行モード_ダミー 1:周辺状況_ダミー 1	-1.0141	**
走行モード_ダミー 1:周辺状況_ダミー 2	-19.1795	
道路形状_ダミー 1:行動類型_ダミー 1	-16.3506	
道路形状_ダミー 1:行動類型_ダミー 2	-2.9793	
道路形状_ダミー 1:道路情報_ダミー 1	0.1066	
道路形状_ダミー 1:道路情報 2_ダミー 1	0.3108	
道路形状_ダミー 1:周辺状況_ダミー 1	-0.2658	
道路形状_ダミー 1:周辺状況_ダミー 2	-15.1947	
行動類型_ダミー 1:道路情報_ダミー 1	0.4531	
行動類型_ダミー 2:道路情報_ダミー 1	0.3345	
行動類型_ダミー 1:道路情報 2_ダミー 1	0.7148	
行動類型_ダミー 2:道路情報 2_ダミー 1	1.2907	.
行動類型_ダミー 1:周辺状況_ダミー 1	16.9911	
行動類型_ダミー 2:周辺状況_ダミー 1	3.2628	***
行動類型_ダミー 1:周辺状況_ダミー 2	1.0536	
行動類型_ダミー 2:周辺状況_ダミー 2	0.3013	
道路情報_ダミー 1:周辺状況_ダミー 1	1.1098	.
道路情報_ダミー 1:周辺状況_ダミー 2	-0.4562	
道路情報 2_ダミー 1:周辺状況_ダミー 1	1.1485	***
道路情報 2_ダミー 1:周辺状況_ダミー 2	-1.1076	

***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, .: $p < 0.1$

したモデルでのヒヤリハットの推定結果を示しており、図 1 と同様の形式の散布図となっている。左はゼロ切斷ポアソン回帰、右はゼロ切斷負の二項回帰による推定結果である。

図 2, 表 4 のモデルでは、どちらの回帰においても、交互作用を考慮しないモデル(図 1)と比較して予測値の当てはまりが改善していることが確認できる。また、ゼロ切斷ポアソン回帰ならびにゼロ切斷負の二項回帰の予測値と観測値の関係を比較すると、ほぼ同等の当てはまりであるが、本研究ではヒヤリハットの出現回数が多い観測値の特徴を抽出することが目的であり、その意味ではわずかにゼロ切斷ポアソン回帰の方が当てはまりが良いと判断できる。

4.3 変数選択によるモデルの簡素化

表 4 は、ゼロ切斷ポアソン回帰ですべての 2 因子交互作用までを考慮したモデルの各効果

表5. 交互作用項を含むゼロ切断ポアソン回帰モデルの比較(対数尤度・パラメータ数・AIC).

モデル		対数尤度	パラメータ数	AIC
1	主効果のみ	-286.1136	9	590.2272
2	交互作用項をすべて用いたモデル	-104.8937	34	277.7874
3	モデル2から係数は大きい が有意差がない交互作用項を削除	-135.1193	24	318.2386
4	モデル3から行動類型:道路情報 を削除	-136.3659	22	316.7318
5	モデル4から走行モード:道路情報 を削除	-137.6000	21	317.2000
6	モデル5から走行モード:道路情報2 を削除	-138.2256	20	316.4512
7	モデル6から道路形状:道路情報と 道路形状:道路情報2を削除	-139.3560	18	314.7120

注: モデル3では, 以下の交互作用項を削除した:

走行モード:行動類型, 走行モード:周辺状況, 道路形状:行動類型, 道路形状:周辺状況, 行動類型:道路情報2

(変数)の推定値ならびに対応する p 値をまとめたものである. モデルのあてはまりは大きく改善されたが, ヒヤリハット回数が最も多い(74回)組合せなど, 予測値が実測値と一致しており, 特定のデータに対する過適合(オーバーフィッティング)が生じている可能性が考えられるとともに, 推定モデルの各効果の p 値が大きく, ヒヤリハットの数の予測に寄与しないと考えられる変数も多く含まれていることがわかる. 表4の中で, 推定値は大きい有意でない交互作用項として次のものが挙げられる.

- (i) 走行モード1(自動運転):行動類型2(左折)
- (ii) 走行モード1(自動運転):周辺状況2(非市街地)
- (iii) 道路形状1(交差点):行動類型1(右折)
- (iv) 道路形状1(交差点):周辺状況2(非市街地)
- (v) 行動類型1(右折):周辺状況1(市街地(住宅地でない))

これらの交互作用の水準組合せをもつデータは1個あるいは2, 3個と少ないため, 係数は大きい有意ではないものとなっている. このように有意でない交互作用項を除外し, 推定精度を維持しつつモデルの簡素化と汎化性能の向上を目的として, p 値が小さい交互作用項を残して, AICに基づく変数選択(変数減少法)を実施することとした. すべての交互作用項を考慮したモデルから, AICによるモデルの当てはまりに着目しつつ, 変数の解釈性が高くなるように逐次変数選択を実施する. 表5は, ゼロ切断ポアソン回帰およびゼロ切断負の二項回帰において変数選択の過程をまとめたものである. この結果, 最終的に交互作用項として, 走行モード:道路形状, 行動類型:周辺状況, 道路情報:周辺状況, 道路情報2:周辺状況を含むモデル7を選択した.

図3は, モデル7による推定結果である. 図1ならびに図2と同様の形式の散布図となっている. 左はゼロ切断ポアソン回帰の推定結果, 右はゼロ切断負の二項回帰の推定結果である. ゼロ切断ポアソン回帰での推定のAICは278から315であり, 負の二項回帰での推定のAICは269から263となっており, 過適合はなくなりつつ, 交互作用項を考慮しないモデルよりも予測精度は高くなっている.

表6は, 交互作用項を考慮したゼロ切断ポアソン回帰モデル, 表7はゼロ切断負の二項回帰モデルの各効果の係数と p 値を示しているが, 各効果がヒヤリハットの数の予測に関係しているモデルとなっており, すべての交互作用項を考慮したモデルに対して, 推定精度を維持しつつモデルの簡素化と汎化性能の向上を図ることができた. この場合も, ゼロ切断負の二項回帰

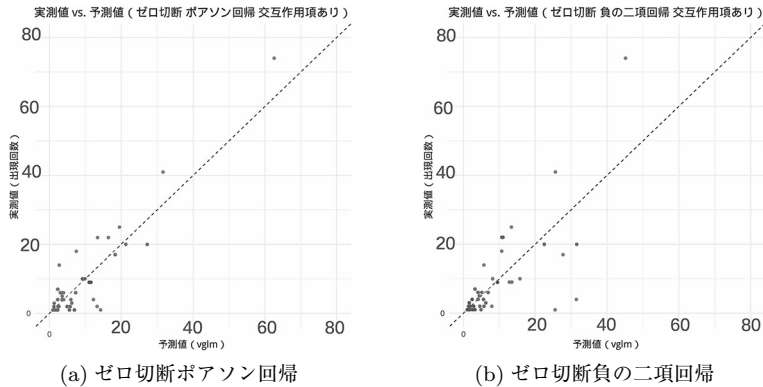


図 3. 逐次変数選択をしたモデルでの予測値と実測値の比較.

表 6. 推定係数と有意性(ポアソン).

変数	推定値	有意性
(Intercept)	2.91142	***
走行モード_ダミー 1	-0.50219	**
道路形状_ダミー 1	-1.54837	***
行動類型_ダミー 1	-3.48629	***
行動類型_ダミー 2	-1.70777	***
道路情報_ダミー 1	0.3957	*
道路情報 2_ダミー 1	-0.85589	***
周辺状況_ダミー 1	-1.66205	***
周辺状況_ダミー 2	0.05855	
走行モード_ダミー 1:道路形状_ダミー 1	2.19994	***
行動類型_ダミー 1:周辺状況_ダミー 1	3.88723	***
行動類型_ダミー 2:周辺状況_ダミー 1	3.5463	***
行動類型_ダミー 1:周辺状況_ダミー 2	2.5152	*
行動類型_ダミー 2:周辺状況_ダミー 2	0.57793	
道路情報_ダミー 1:周辺状況_ダミー 1	0.65437	
道路情報_ダミー 1:周辺状況_ダミー 2	-1.14006	***
道路情報 2_ダミー 1:周辺状況_ダミー 1	1.07014	***
道路情報 2_ダミー 1:周辺状況_ダミー 2	-0.63207	

***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, ∴ $p < 0.1$

による予測値と実測値とのプロットを比較するとゼロ切断ポアソンモデルと大差がないことがわかる。そのため、今回はモデルの簡易さならびに解釈のしやすさからゼロ切断ポアソン回帰モデルを採用した。

4.4 主効果と交互作用効果の分析

ヒヤリハットの予測モデル7の各効果についてみる。なお本研究で用いた48パターン(水準組合せ)に対するヒヤリハットの数は1以上であり、ヒヤリハットが起こったもとの解析結果であることに注意されたい。

まず、主効果については、以下のようになっている。ただし、交互作用項についても合わせて考える必要があるため、ここでは主効果の大小関係を述べるに留める。

表7. 推定係数と有意性(負の二項回帰).

変数	推定値	有意性
(Intercept)	3.3191	***
(Intercept2)	0.726	***
走行モード_ダミー1	-1.1006	**
道路形状_ダミー1	-1.7812	***
行動類型_ダミー1	-3.8772	***
行動類型_ダミー2	-2.0443	***
道路情報_ダミー1	0.1293	*
道路情報2_ダミー1	-0.862	***
周辺状況_ダミー1	-2.0463	***
周辺状況_ダミー2	-0.7438	
走行モード_ダミー1:道路形状_ダミー1	2.6716	***
行動類型_ダミー1:周辺状況_ダミー1	3.46	***
行動類型_ダミー2:周辺状況_ダミー1	3.1075	***
行動類型_ダミー1:周辺状況_ダミー2	3.1762	*
行動類型_ダミー2:周辺状況_ダミー2	1.2342	
道路情報_ダミー1:周辺状況_ダミー1	1.3427	
道路情報_ダミー1:周辺状況_ダミー2	-0.6431	***
道路情報2_ダミー1:周辺状況_ダミー1	1.5627	***
道路情報2_ダミー1:周辺状況_ダミー2	-0.6139	

***: $p < 0.001$, **: $p < 0.01$, *: $p < 0.05$, .: $p < 0.1$

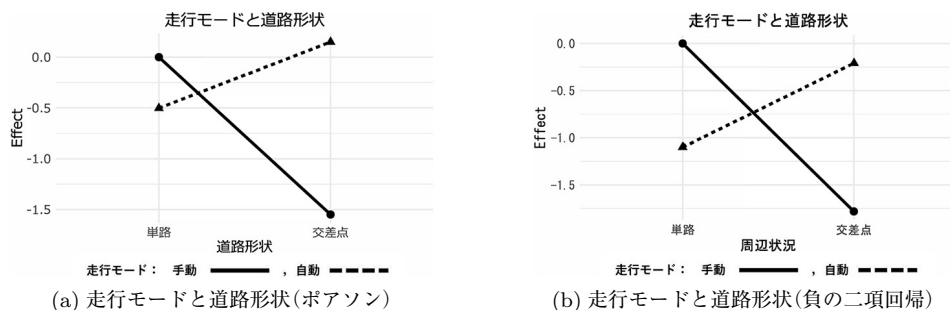


図4. 走行モードと道路形状の比較.

- 走行モード：手動>自動
- 道路形状：単路>交差点
- 行動類型：直進>左折>右折
- 道路情報(歩者区分)：段差などの障害物がある>何も無い
- 道路情報2(中央分離帯)：何も無い>段差などの障害物がある
- 周辺状況：市街地であり住宅街>非市街地>市街地だが住宅街ではない

次に、交互作用項を含めた要因効果についてみていく。図4は走行モードと道路形状、図5は行動類型と周辺状況、図6は道路情報(歩者区分)と周辺状況、図7は道路情報(中央分離帯)と周辺状況について各カテゴリでの要因効果量を示したものである。横軸ならびに各種の線はそれぞれの説明変数に対応しており、縦軸は要因効果量を表している。それぞれ、左にはゼロ切断ポアソン回帰、右にはゼロ切断負の二項回帰による結果を示している。

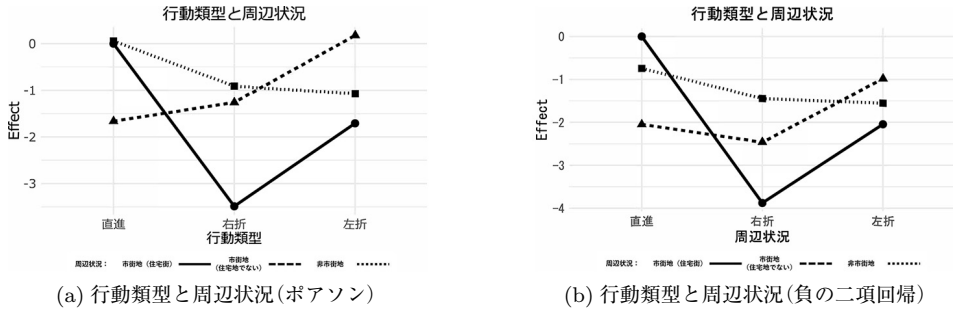


図 5. 走行モードと道路形状の比較.

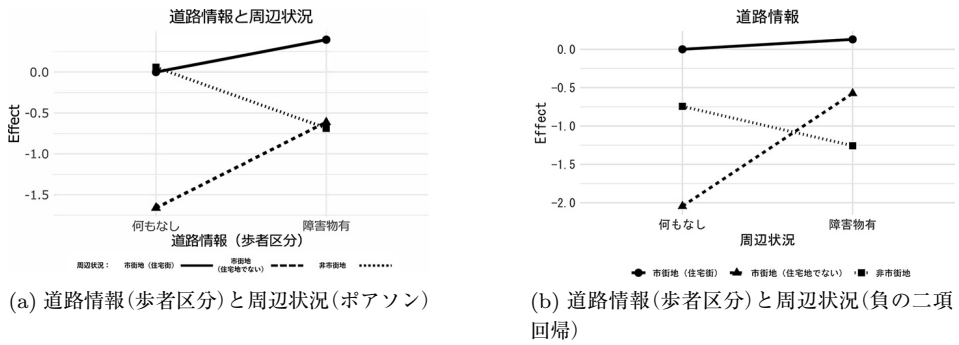


図 6. 道路情報(歩者区分)と周辺状況.

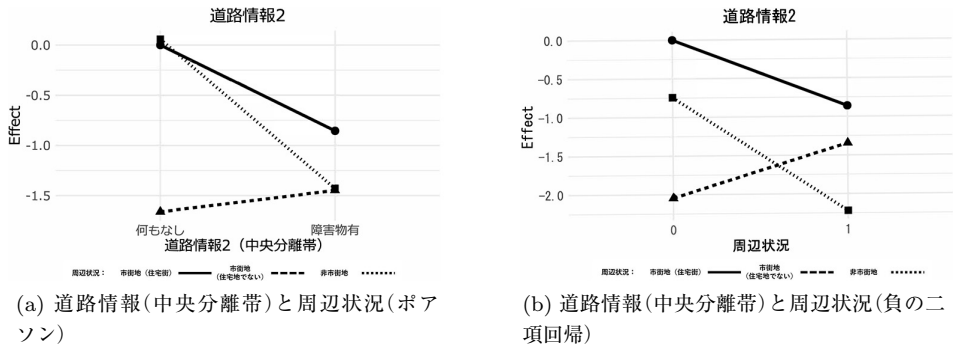


図 7. 道路情報(中央分離帯)と周辺状況.

(1) 走行モードと道路形状 図 4 より, 走行モード(手動と自動)は道路形状のみと交互作用が有意である. 手動運転と自動運転では走行する道路形状によってヒヤリハットの発生数への影響が異なることがわかる. 交差点の周辺状況の視認に加え, 交差点に入る直前の減速や進行時の徐行などを同時に処理することは, 今回の実証実験中では自動運転よりも手動運転の方がヒヤリハットの数を減少させる傾向にあることがわかる.

(2) 行動類型と周辺状況 図 5 より, 市街地(住宅地でない)における左折については, 他の行動類型と周辺状況の交互作用と異なり, ヒヤリハットを増やす要因になっていることがわかる. ここで市街地(住宅地でない)としては駅近くなどの交通量が多いところが挙げられる. そのような場所では車両だけでなく人通りや自転車も多く, 左折時の周辺状況の確認が困難であるこ

表 8. ダミー変数による走行条件と出現回数.

走行モード	道路形状	行動類型	道路情報	道路情報 2	周辺状況	出現回数
0	0	2	1	0	1	74
1	1	0	1	0	0	41

とが考えられる。また、右左折よりも直進行動の方がスピードが出、飛び出しなど不測の事態が生じることが考えられ、急ブレーキをふむなどヒヤリハットが生じていることが考えられる。

(3)道路情報(歩者区分)と周辺状況 図6より、非市街地ではヒヤリハットを減らす要因になっていることがわかる。一方、市街地では歩者区分のための段差やガードレールなどの障害物はヒヤリハットを増やす要因となっている。

(4)道路情報(中央分離帯)と周辺状況 図7より、市街地(住宅街)と非市街地ではヒヤリハットを減らす要因になっていることがわかる。一方、市街地(住宅地でない)では中央分離帯のための段差やガードレールなどの障害物の有無はヒヤリハットの増減への影響は小さい。

4.5 ヒヤリハットが多いパターンの分析

最後に、ヒヤリハットが多いパターンについてゼロ切断ポアソン回帰モデルを用いて詳細にみていく。

表8はヒヤリハットが多い上位2つのパターンをまとめたものである。まず、ヒヤリハット数が74のパターンの推定値は次のようになる。

$$\log(\mu) = 2.91142 - 1.70777 + 0.3957 - 1.66205 + 3.5463 + 0.65437 = 4.13797$$

この場合、 $\mu = e^{4.14} \approx 62.63$ となる。Interceptである2.91142だけでは平均ヒヤリハット数は $\mu = e^{2.91} \approx 19$ となるが、この場合、主効果と交互作用項を合わせて正となる項目としては

- (1) 行動類型(左折)：周辺状況(市街地(住宅街でない))
- (2) 道路情報(段差もしくはガードレールがある)：周辺状況(市街地(住宅街でない))

の2つの交互作用が影響している。よって、駅前商業地など交通量が多い所であつ、左折時あるいは段差などがある所でヒヤリハットが起りやすいことがわかる。

次に、ヒヤリハット数が41のパターンについて推定値は次のようになる。

$$\log(\mu) = 2.91142 - 0.50219 - 1.54837 + 0.3957 + 2.19994 = 3.4565$$

したがって、 $\mu = e^{3.46} \approx 31.71$ となる。この場合、同じく主効果と交互作用項を合わせ正となる項目としては

- (1) 道路情報(段差もしくはガードレールがある)
- (2) 走行モード(自動運転)：道路形状(交差点)の交互作用

の2つが影響している。すなわち段差やガードレールがあり歩者区分があるところ、あるいは自動運転の際の交差点でヒヤリハットが起りやすいことがわかる。

5. まとめ

本研究では、走行実験においてヒヤリハットが発生した時点での発生状況のデータについて、走行環境がヒヤリハット発生に与える影響を分析した。自動運転車両の走行環境要因(道路形状、行動類型、道路情報、周辺状況)と自動運転モードの交互作用がヒヤリハット発生数に

与える効果を、ゼロ切断ポアソン回帰モデルとゼロ切断負の二項回帰モデルにより検討した。

その結果、道路形状(単路, 交差点)と走行モード(手動, 自動)との相互作用がヒヤリハット発生に有意な影響を与えることが明らかとなった。行動類型や周辺状況との組合せがリスクを増大させる傾向があり、交通量が多い地域とそうでない地域でヒヤリハットとなりうる要因が異なることが見られた。特に、(i) 交通量が多い市街地(住宅街でない)で段差やガードレールによる歩者区分がある所、(ii) 自動運転の際の交差点、(iii) 交通量が多い市街地での左折時にはヒヤリハットが増加することが明らかになった。(iii) で特筆すべき点は交差点での直進や右折よりも左折時のヒヤリハットが多い点である。

今後の課題として、ヒヤリハットが発生した場合の要因としてではなく、走行開始から走行終了までを一定の短時間で区切り、ヒヤリハットが発生するか否かのデータを用いたポアソン回帰などにより、どのような走行環境でヒヤリハットが発生しやすいのかをさらに詳細に分析していく。さらに、交通量などの影響がヒヤリハットにどのように関係するのかも明らかにし、自動運転車両がとっさに判断することが難しい横断についても危険度を推測することで、自動運転車両が直面するヒヤリハット数の予測と削減に繋がると考えられる。

謝 辞

本研究の実施に際し、走行データのご提供および、モデルの評価・解釈に関して実運用の観点から多大なご助言を賜りました国立大学法人群馬大学次世代モビリティ社会実装研究センターおよび当該センターの小木津准教授、ならびに日本モビリティ株式会社に、心より感謝申し上げます。

参 考 文 献

- Bird, F. E. and Germain, G. L. (1985). *Practical Loss Control Leadership*, International Loss Control Institute, Loganville, Georgia.
- Cameron, A. C. and Trivedi, P. K. (2013). *Regression Analysis of Count Data*, 2nd ed., Cambridge University Press, Cambridge, UK.
- Fagnant, D. J. and Kockelman, K. (2015). Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations, *Transportation Research Part A: Policy and Practice*, **77**, 167–181.
- Godwin, R. T. (2024). One-inflated zero-truncated count regression models, *arXiv preprint arXiv:2402.02272* (最終アクセス日 2025 年 10 月 3 日).
- Gourieroux, C., Monfort, A. and Trognon, A. (1984). Pseudo maximum likelihood methods: Applications to Poisson models, *Econometrica*, **52**(3), 701–720.
- Grogger, J. T. and Carson, R. T. (1991). Models for truncated counts, *Journal of Applied Econometrics*, **6**(3), 225–238.
- Gurmu, S. (1991). Tests for detecting over-dispersion in the positive Poisson regression model, *Journal of Business and Economic Statistics*, **9**(2), 215–222.
- Heinrich, H. W. (1931). *Industrial Accident Prevention: A Scientific Approach*, McGraw-Hill.
- 伊藤誠, 堀宗朗, 加藤真平 (2020). 自動運転車両における異常検知と事故予防のためのリスク評価, 人工知能学会論文誌, **35**(4), 562–573.
- Kim, S., Kang, J. Y., Kim, H. M. and Lee, J. (2017). Analysis of near-miss incidents in autonomous vehicles using naturalistic driving data, *IEEE Transactions on Intelligent Transportation Systems*, **18**(4), 807–818.

- 国土交通省 (2023). 道路運送車両法の一部を改正する法律(令和 4 年法律第 36 号, 2023 年 4 月施行), <https://laws.e-gov.go.jp/law/326AC0000000185/> (最終アクセス日 2025 年 10 月 3 日).
- 国土交通省 (2024). 自動運転移動サービス社会実装・事業化の手引き, https://www.road-to-the-l4.go.jp/activity/guideline/pdf/自動運転移動サービス社会実装・事業化の手引き_v2.0_.pdf (最終アクセス日 2025 年 10 月 3 日).
- Li, Xun-Jian., Sun, Yuan, Tian, Guo-Liang, Liang, Jiajuan and Shi, Jianhua (2023). Mean regression model for the zero-truncated Poisson distribution and its generalization, *Computational Statistics and Data Analysis*, **179**, 107650.
- Litman, T. (2020). Autonomous vehicle implementation predictions, Victoria Transport Policy Institute, Victoria, British Columbia.
- Long, J. S. (1997). *Regression Models for Categorical and Limited Dependent Variables*, Sage Publications, Thousand Oaks, California.
- Long, J. S. and Freese, J. (2006). *Regression Models for Categorical Dependent Variables Using Stata*, Stata Press, College Station, Texas.
- Simonoff, J. S. (2003). *Analyzing Categorical Data*, Springer-Verlag, New York.
- Tian, G. L., Liu, Y., Tang, M. L. and Jiang, X. (2018). Type I multivariate zero-truncated/adjusted Poisson distributions with applications, *International Journal of Computational and Applied Mathematics*, **344**, 132–153.
- Van der Heijden, P. G., Bustami, R., Cruyff, M. J., Engbersen, G. and Van Houwelingen, H. C. (2003). Point and interval estimation of the population size using the truncated Poisson regression model, *Statistical Modelling*, **3**(4), 305–322.
- Winkelmann, R. (2008). *Econometric Analysis of Count Data*, Springer, Berlin, Heidelberg.
- Yee, T. W. and Ma, C. (2024). Generally altered, inflated, truncated and deflated regression, *Statistical Science*, **39**(4), 568–588.

Analysis of Near-miss Factors Influenced by Driving Environment in Autonomous Vehicles

Yusuke Tajima¹, Kaoru Fueda², Yusaku Miura³ and Takahiro Miki³

¹Faculty of Data Science, Otsuma Women's University

²Faculty of Data Science, Shiga University

³Product & CSV × DX Planning Department, Aioi Nissay Dowa Insurance

This paper analyzes the influence of driving environments on the occurrence of near-miss incidents in autonomous vehicles using data from driving experiments. Specifically, we categorized the driving environment at the time of each incident in terms of driving mode, road geometry, driving behavior, road information, and surrounding conditions. Using these categorical variables, we constructed a zero-truncated Poisson regression model to predict the number of incidents and identify contributing factors. Furthermore, by introducing two-way interaction terms between variables, we conducted variable selection and examined the interpretability of the model. The results revealed specific combinations of driving environments that significantly affect incident occurrence. In particular, the interaction between driving behavior and surrounding conditions was found to increase the number of incidents regardless of the driving mode. The findings of this study are expected to contribute to the enhancement of safety in future autonomous mobility services, such as driverless shuttle buses.