รายงานผลการวิจัย

เรื่อง ตัวแบบเชิงเส้นวางนัยทั่วไปสำหรับการติดตามระยะยาวของจำนวนการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

Generalized Linear Models for Longitudinal Study of Accident for Highways in Mountainous Areas

ได้รับการจัดสรรงบประมาณวิจัย ประจำปี 2558

จำนวน 50,000 บาท

หัวหน้าโครงการวิจัย **นายทวีศักดิ์ จันทร์งาม**

ผู้ร่วมวิจัย **นายชัยวัฒน์ โฆษภัทรพิมพ์**

งานวิจัยเสร็จสิ้นสมบูรณ์

28 กุมภาพันธ์ 2562

กิตติกรรมประกาศ

โครงการวิจัยเรื่อง ตัวแบบเชิงเส้นวางนัยทั่วไปสำหรับการศึกษาติดตามระยะยาวของจำนวนการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขา สำเร็จลุล่วงได้ด้วยการได้รับทุนอุดหนุนการวิจัยจากสำนักวิจัยและส่งเสริมวิชาการเกษตร มหาวิทยาลัยแม่โจ้ ประจำปีงบประมาณ 2558 ผู้วิจัยขอขอบคุณสำนักอำนวยความปลอดภัย กรมทางหลวง และสำนักทางหลวงที่ 1 จังหวัดเชียงใหม่ ที่ให้ความอนุเคราะห์ข้อมูลในการวิจัยครั้งนี้ ขอขอบคุณสาขาวิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้ ที่อนุเคราะห์เรื่องสถานที่ และอุปกรณ์บางอย่างที่ใช้ในการดำเนินการวิจัยให้เสร็จสิ้นสมบูรณ์

ผู้วิจัย

สารบัญ

สารบัญตาราง

สารบัญภาพ

ตัวแบบเชิงเส้นวางนัยทั่วไปสำหรับการศึกษาติดตามระยะยาวของจำนวนการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

Generalized Linear Models for Longitudinal Study of Accident for Highways in Mountainous Areas

ทวีศักดิ์ จันทร์งาม[1] และชัยวัฒน์ โฆษภัทรพิมพ์[1]

Taweesak Channgam[1] and Chaiwat Kosapattarapim[1]

[1]สาขาวิชาสถิติ คณะวิทยาศาสตร์ มหาวิทยาลัยแม่โจ้ จ.เชียงใหม่ 50290

บทคัดย่อ

การวิจัยนี้ มีวัตถุประสงค์ เพื่อศึกษาปัจจัยต่าง ๆ ที่ส่งผลต่อจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขาด้วยตัวแบบเชิงสถิติ โดยใช้ข้อมูลการติดตามระยะยาวระหว่างปี 2555 – 2559 ของช่วงกิโลเมตรทั้งสิ้น 105 กิโลเมตร จากถนน 26 สาย ของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ ผลการวิจัยพบว่า ตัวแบบที่มีความเหมาะสมกับข้อมูลจำนวนการเกิดอุบัติเหตุ คือ ตัวแบบผสมเชิงเส้นนัยทั่วไปสำหรับข้อมูลจำนวนนับที่มีศูนย์มาก (GLMMs with Zero inflated count data) เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) ซึ่งพบว่า จำนวนช่องจราจร ประเภทเกาะกลางถนน ชนิดผิวจราจร ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี มีผลต่อจำนวนการเกิดอุบัติเหตุ สำหรับตัวแบบที่มีความเหมาะสมกับข้อมูลจำนวนผู้บาดเจ็บ คือ ตัวแบบผสมเชิงเส้นนัยทั่วไปสำหรับข้อมูลจำนวนนับที่มีศูนย์มากเมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ Exchangeable ซึ่งพบว่า ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี ชนิดผิวจราจร และความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร ส่วนข้อมูลจำนวนผู้เสียชีวิตไม่สามารถหาตัวแบบที่เหมาะสมได้

คำสำคัญ: ตัวแบบผสมเชิงเส้นนัยทั่วไป, จำนวนการเกิดอุบัติเหตุ, จำนวนผู้บาดเจ็บ, จำนวนผู้เสียชีวิต, ทางหลวงในพื้นที่ลาดชันแบบภูเขา

Abstract

The purposes of this research were to determine the factors affecting the number of traffic accidents, the number of injury accident and number of fatal accidents for highways in mountainous areas using the statistical models. Follow-up data between 2012 and 2016 were drawn from the Chiang Mai Rural Road District Office. Results found that, the appropriable model for the number of traffic accidents data is the Generalized Linear Mixed Models (GLMMs) for Zero inflated count data with the covariance structure of AR(1). Number of Lane, Road median, Road surface and Annual average daily traffic (AADT) were significantly associated with the number of traffic accidents. The GLMMs for Zero inflated count data with the covariance structure of Exchangeable appropriated for the number of injury accident data and we found that AADT, road surface and average slope after 300 meter of one kilometer were associated with the number of injury accident. However, this study did not found the appropriable model for the number of fatal accidents.

Key words: Generalized Linear Mixed Models, traffic accidents, injury accident, fatal accidents, highways in mountainous areas

คำนำ

องค์การอนามัยโลก (World Health Organization: WHO) ได้กำหนดเรื่องความปลอดภัยบนถนน เป็นประเด็นที่ทั่วโลกไม่ว่าจะเป็นภาครัฐหรือภาคเอกชนต้องให้ความสำคัญ ร่วมมือกันผลักดันเป็นวาระของโลกและชาติ โดยได้กำหนดคำขวัญที่ใช้ในการรณรงค์เกี่ยวกับความปลอดภัยบนถนนไว้ว่า “Road Safety is No Accident” หรือในภาษาไทยใช้คำว่า “สำนึกดี ขับขี่ปลอดภัย ร่วมใจลดอุบัติเหตุ” จากรายงานขององค์การอนามัยโลกในปี 2547 พบว่ามีผู้ได้รับบาดเจ็บจากอุบัติเหตุทางถนนทั่วโลกเฉลี่ยวันละ 140,000 คน เสียชีวิตเฉลี่ยวันละ 3,000 คน และพิการเฉลี่ยวันละ 15,000 คน (WHO, 2004) จากรายงานดังกล่าวทำให้เกิดความสูญเสียทั้งทางตรงและทางอ้อมต่อผู้ที่ประสบอุบัติเหตุ เช่น การเสียชีวิต การบาดเจ็บ ความพิการ และการสูญเสียทรัพย์สินเพื่อการรักษาตัว การสูญเสียหัวหน้าครอบครัว การสูญเสียรายได้ เป็นต้น นอกจากความสูญเสียที่เกิดจากผู้ที่ประสบอุบัติเหตุแล้ว ยังพบว่าเกิดความสูญเสียที่ผลต่อสังคม เช่น รัฐต้องแบกรับภาระค่าใช้จ่ายที่เกิดจากการรักษาตัวของผู้ประสบอุบัติเหตุ เป็นต้น

สำหรับประเทศไทย องค์การอนามัยโลกได้จัดให้ประเทศไทยเป็นประเทศที่มีความเสี่ยงสูงในด้านความไม่ปลอดภัยทางถนนเมื่อเทียบกับประเทศอื่น ๆ ในภูมิภาคเดียวกัน โดยพบว่ามีจำนวนผู้เสียชีวิตจากอุบัติเหตุทางถนนเฉลี่ย 19 คน ต่อประชากร 100,000 คนต่อปี (WHO, 2009) และจากรายงานจำนวนผู้เสียชีวิตจากอุบัติเหตุจราจรทางบกประจำปีงบประมาณ พ.ศ. 2556 ของศูนย์อำนวยการความปลอดภัยทางถนน พบว่ามีจำนวนผู้เสียชีวิตจากอุบัติเหตุจราจรทางบกสูงถึง 8,631 คน (ศูนย์อำนวยการความปลอดภัยทางถนน, 2557) จากรายงานการเกิดอุบัติเหตุข้างต้นสามารถจำแนกปัจจัยที่เกี่ยวข้องกับการเกิดอุบัติได้เป็น 4 กลุ่มหลัก ๆ คือ คน สภาพยานพาหนะ สภาพถนน และสภาพสิ่งแวดล้อม เมื่อพิจารณาจะพบว่าปัจจัยที่เกี่ยวกับคนนั้นสามารถป้องกันได้ด้วยการเสริมสร้างความรู้ความเข้าใจเกี่ยวกับการขับขี่ ปัจจัยที่เกี่ยวกับสภาพยานพาหนะสามารถป้องกันได้ด้วยการออกกฎหมายเพื่อควบคุมมาตรฐานของยานพาหนะได้ ส่วนปัจจัยที่เกี่ยวกับสภาพถนนและสภาพแวดสิ่งแวดล้อมเป็นสิ่งที่ควบคุมได้ยากที่สุด โดยเฉพาะถนนที่อยู่ในช่วงภูมิประเทศที่เป็นแบบภูเขา ซึ่งจะพบว่า เส้นทางจะมีความลาดชันสูง ซึ่งอาจก่อให้เกิดอุบัติเหตุได้ง่ายถ้าผู้ขับขี่ไม่ระมัดระวัง หรือสภาพของถนนไม่ดี อย่างเช่น การเกิดอุบัติเหตุรถประจำทางสายเชียงใหม่ – อุดรธานี ตกเหวบริเวณทางหลวง หมายเลข 11 ตำบลบ่อเหล็กลอง จังหวัดแพร่ เมื่อวันที่ 17 ตุลาคม พ.ศ. 2557 ทำให้มีผู้เสียชีวิตจำนวน 1 คน และมีผู้ได้รับบาดเจ็บรวบ 18 คน การเกิด

อุบัติเหตุรถทัวร์เสียหลักตกสะพานห้วยตอง สาย 12 หล่มสัก-ขอนแก่น กิโลเมตรที่ 374+300 ในวันที่ 27 ธันวาคม พ.ศ. 2556 ทำให้มีผู้เสียชีวิตสูงถึง 29 คน (ไทยรัฐออนไลน์, 2557) การเกิดอุบัติเหตุรถยนต์ตกเขาในขณะกำลังขึ้นภูทับเบิก จังหวัดเพชรบูรณ์ ทำให้มีผู้ได้รับบาดเจ็บ 1 คน (ASTVผู้จัดการออนไลน์, 2557) เป็นต้น และเมื่อเกิดอุบัติเหตุขึ้น ณ ช่วงกิโลเมตรใดกิโลเมตรหนึ่ง เจ้าหน้าที่ตำรวจหรือเจ้าหน้าที่รัฐจะทำการบันทึกข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตลงในฐานข้อมูล ซึ่งโดยทั่วไปมักจะพบว่าการเกิดอุบัติเหตุจะเกิดขึ้นซ้ำบริเวณเดิม ทำให้ข้อมูลดังกล่าวมีลักษณะที่เรียกว่า ข้อมูลที่มีการวัดซ้ำ หรือข้อมูลที่มีการติดตามระยะยาว (Longitudinal Data)

ในทางสถิติได้มีการศึกษาตัวแบบสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวง ณ เวลาใดเวลาหนึ่ง ซึ่งเป็นข้อมูลการนับ (count) ที่มีลักษณะการแจกแจงปัวซง (Poisson Distribution) (Bäumer *et al.*, 2000) โดยได้มีการนำหลักการของตัวแบบเชิงเส้นวางนัยทั่วไป (Generalized Linear Models, GLMs) มาใช้พัฒนาตัวแบบข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวง แต่เนื่องจากวิธี GLMs มีข้อตกลงเบื้องต้นว่าค่าสังเกตของตัวแปรตามหรือตัวแปรที่สนใจจะต้องเป็นอิสระต่อกันในช่วงเวลาที่ศึกษา ในขณะที่การวิเคราะห์ข้อมูลที่มีการติดตามในระยะยาวอย่างเช่น ข้อมูลการเกิดอุบัติเหตุบนทางหลวง เป็นข้อมูลที่มีการเก็บบันทึกอย่างต่อเนื่องทุกปีในกิโลเมตรเดียวกันของถนน ทำให้ค่าสังเกตข้อมูลที่มีการเก็บซ้ำในแต่ละปีของช่วงกิโลเมตรเดียวกันของถนนมีความสัมพันธ์กันเกิดขึ้น ดังนั้นในปี ค.ศ. 1986 นักสถิติที่ชื่อว่า Liang และ Zeger (1986) ได้นำเสนอวิธี Generalized Estimating Equation (GEE) ที่พัฒนาต่อมาจากวิธี GLMs สำหรับการสร้างตัวแบบภาพรวมของค่าเฉลี่ยประชากร (Population Average: PA) โดยได้มีการนำโครงสร้างความสัมพันธ์ที่สมมติขึ้นสำหรับข้อมูลที่เก็บซ้ำของหน่วยศึกษาหนึ่ง ๆ กำหนดเป็น Working Correlation เข้ามาพิจารณาในขั้นตอนของการประมาณค่าพารามิเตอร์ นอกจากนี้ข้อมูลที่มีการเก็บซ้ำแบบติดตามระยะยาวสามารถวิเคราะห์ด้วยวิธีตัวแบบผสมเชิงเส้นวางนัยทั่วไป (Generalized Linear Mixed Models, GLMMs) ซึ่งเป็นตัวแบบที่อธิบายแต่ละหน่วยบุคคลหรือหน่วยศึกษา (Subject-Specific Models) โดยตัวแบบนี้พัฒนามาจากตัวแบบภาพรวมค่าเฉลี่ยประชากร (PA Model) และมีการเพิ่มเทอมอิทธิพลสุ่ม (Random Effect) เพื่อแสดงลักษณะเฉพาะของแต่ละบุคคล เช่น การยอมให้ค่าพื้นฐานหรือที่เรียกว่าค่าคงที่ (Intercept) ของแต่ละบุคคลมีค่าแตกต่างกันได้แบบสุ่ม เรียกเป็น Random

Intercept เพื่อให้อธิบายความแตกต่างของแต่ละบุคคลได้ชัดเจนกว่าการอธิบายผลในภาพรวมของค่าเฉลี่ยประชากร

จากปัญหาและความสำคัญที่กล่าวข้างต้น ผู้วิจัยจึงให้ความสนใจพัฒนาตัวแบบทางสถิติสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่มีการเก็บข้อมูลซ้ำแต่ละปีในระยะเวลาติดตาม 5 ปี ของช่วงถนนกิโลเมตรเดียวกัน เมื่อตัวแปรตามที่สนใจ คือ จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ที่มีการแจกแจงปัวซง (Poisson Distribution) ด้วยวิธี GEE และวิธี GLMMs พร้อมทั้งพิจารณาตัวแบบที่เหมาะสมกับข้อมูล เพื่อเป็นประโยชน์ต่อหน่วยงานที่เกี่ยวข้องสามารถนำวิธีการพัฒนาตัวแบบไปใช้กับข้อมูลในช่วงเวลาต่อไป จะได้ผลลัพธ์ข้อมูลที่ทันสมัยและใช้เป็นข้อมูลสนับสนุนในการวางแผนการป้องกันการเกิดอุบัติเหตุในพื้นที่ดังกล่าว รวมทั้งสามารถเป็นแนวทางในการนำวิธีทางสถิติมาประยุกต์ใช้กับข้อมูลการเกิดอุบัติเหตุอื่น ๆ เมื่อตัวแปรตามที่สนใจเป็นข้อมูลจำนวนนับที่มีการเก็บข้อมูลซ้ำ

วัตถุประสงค์ของการวิจัย

1. เพื่อศึกษาปัจจัยต่าง ๆ ที่ส่งผลต่อจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

2. เพื่อพัฒนาตัวแบบเชิงสถิติสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชัดแบบภูเขาที่มีการเก็บข้อมูลซ้ำในระยะเวลาติดตาม 5 ปี เมื่อสมมติโครงสร้างความสัมพันธ์ของข้อมูลในหน่วยศึกษามีรูปแบบ First-order Autoregressive: AR(1) และ Exchangeable

3. เพื่อหาตัวแบบเชิงสถิติที่เหมาะสมสำหรับทำนายข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชัดแบบภูเขา

ประโยชน์ที่คาดว่าจะได้รับ

1. ทราบถึงปัจจัยต่าง ๆ ที่ส่งผลต่อจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

2. ได้ตัวแบบทางสถิติที่สามารถทำนายข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่มีการเก็บข้อมูลซ้ำ

การตรวจเอกสาร

1) นิยามศัพท์พื้นฐานเกี่ยวกับอุบัติเหตุ

เพื่อให้เกิดความเข้าตรงกันเกี่ยวกับคำศัพท์ต่าง ๆ ที่ใช้ในรายงานการวิจัยนี้ ผู้วิจัยจึงขอกำหนดคำนิยามศัพท์ โดยอ้างอิงจากสำนักอำนวยความปลอดภัย กรมทางหลวง ดังนี้

1.1) อุบัติเหตุ (Traffic Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นบนทางหลวง ทั้งนี้อาจมีผู้เสียชีวิตหรือผู้บาดเจ็บหรือเกิดความเสียหายต่อทรัพย์สิน

1.2) อุบัติเหตุที่มีผู้เสียชีวิต (Fatal Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นแล้วทำให้เกิดผู้เสียชีวิต ทั้งนี้อาจมีผู้บาดเจ็บหรือทรัพย์สินเสียหายร่วมด้วยก็ได้

1.3) อุบัติเหตุที่มีผู้บาดเจ็บ (Injury Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นแล้วทำให้เกิดผู้บาดเจ็บหรือทรัพย์สินเสียหาย

1.4) อุบัติเหตุที่เกิดความเสียหาย (Damage Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นแล้วทำให้เกิดทรัพย์สินเสียหายเท่านั้น

1.5) อัตราการเกิดอุบัติเหตุ (Accident Rate) หมายถึง จำนวนรายการการเกิดอุบัติเหตุต่อตัวแปรต่าง ๆ ที่นิยมนำมาเปรียบเทียบตามหลักสากล เช่น จำนวนการเกิดอุบัติเหตุที่เกิดในช่วงระยะเวลาที่พิจารณา

1.6) ปริมาณจราจรเฉลี่ยต่อวันตลอดปี (Average Annual Daily Traffic หรือ AADT) หมายถึง จำนวนยานพาหนะที่วิ่งผ่านจุดหนึ่งจุดใดหรือทางตอนหนึ่งตอนใดตลอดปีหารด้วยจำนวนวันในปีนั้น

2) ปัจจัยทั่วที่ส่งผลต่อการเกิดอุบัติเหตุบนทางหลวง

จากการทบทวนวรรณกรรมที่เกี่ยวข้องกับการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขา สามารถสรุปปัจจัยที่ส่งผลต่อการเกิดอุบัติในลักษณะดังกล่าวได้ 3 ปัจจัยหลัก ดังนี้

2.1) ผู้ใช้ถนน

ผู้ใช้ถนน เป็นองค์ประกอบหลักในระบบการจราจรและการขนส่งทางหลวง จากรายงานอุบัติเหตุบนทางหลวงพบว่าเกิดจากผู้ใช้ถนนมากถึงร้อยละ 95.62 (สำนักอำนวยความปลอดภัย กรมทางหลวง, 2549) นอกจากการเข้าใจพฤติกรรมทางกายภาพและจิตของผู้ใช้ถนนแล้ว

ยังพบว่าปัจจัยที่เกี่ยวกับผู้ใช้ถนนที่ผลต่อการเกิดอุบัติเหตุยังขึ้นอยู่กับ 3 ส่วนที่สำคัญ คือ การประมวลผลข่าวสารข้อมูล ลักษณะการมองเห็น และข้อมูลข่าวสารที่จำเป็นของผู้ใช้ถนน

2.2.1) การประมวลข่าวสารข้อมูล (Information Processing) ประกอบไปด้วย

1) กระบวนการขับรถ (Driving Task) หมายถึง กิจกรรมหลัก 3 ส่วนที่สำคัญในการขับขี่รถยนต์ คือ การนำร่อง การนำทาง และการควบคุมรถ

2) การคาดการณ์ล่วงหน้า (Expectancy) หมายถึง ประสบการณ์ในการขับขี่รถยนต์ของผู้ใช้ถนน อนึ่งถ้าผู้ใช้ถนนมีประสบการณ์ในการขับขี่รถยนต์และใช้ทางก็จะสามารถลดเวลาการตอบสอง (reaction times) ลงได้ และทำให้ผู้ขับขี่สามารถลดการรับปริมาณข่าวสารข้อมูลลงได้เมื่อขับรถอีกครั้ง ทั้งนี้การคาดการณ์ล่างหน้านี้ขึ้นอยู่กับประสบการณ์และระยะเวลาของผู้ใช้ถนน

3) เวลาตอบสนอง (Reaction Time) หมายถึง ระยะเวลาในการประมวลผลข้อมูลข่าวสารนั้น ๆ ของผู้ใช้ถนนว่าสามารถตอบสนองต่อสัญญาณต่าง ๆ ได้รวดเร็วหรือไม่

4) ความจำ (Memory) หมายถึง ความสามารถใจการจดจำข้อมูลข่าวสารต่อการใช้ถนน เช่น สามารถจดจำป้ายเตือนให้ควบคุมความเร็ว เป็นต้น

5) ความล้าจากการประมวลผลข้อมูลจำนวนมาก (Hysteresis Effect) หมายถึง การที่ผู้ใช้ถนนได้รับข้อมูลข่าวสารจำนวนมากไประยะหนึ่ง อาจทำให้ความสามารถในการประมวลผลลดลง แม้ว่าจะลดปริมาณการรับข้อมูลเหล่านั้นลงก็ตาม ทั้งนี้ยังคงทำให้ความสามารถในการประมวลผลของผู้ใช้ถนนยังคงอยู่ในระดับต่ำ

2.2.2) ลักษณะการมองเห็น (Visual Characteristics)

ผู้ใช้ทางสามารถรับรู้ข้อมูลข่าวสารผ่านการมองข้อมูลจราจร เช่น ป้ายจราจร ได้ร้อยละ 90 ส่วนที่เหลือเป็นการรับรู้ผ่านทางเสียง การสั่นสะเทือน การโยนตัว อาทิ เมื่อรถวิ่งด้วยความเร็วสูงขึ้นการมองเห็นในมุมกว้างจะแคบลง เช่น ที่ความเร็ว 30 กิโลเมตรต่อชั่วโมง มุมการมองเห็นจะลดลงจาก 180 องศาเหลือเพียง 100 องศา ดังนั้นความเร็วยิ่งมากจะทำให้องศาของการมองเห็นลดลง ทั้งนี้จำเป็นต้องออกแบบป้ายจราจร สัญญาจราจร ให้อยู่ในทัศนะวิสัยของผู้ใช้ถนนมองเห็นได้ชัดเจน เป็นต้น

2.2.3) ข้อมูลข่าวสารที่ผู้ใช้ทางต้องการ (Information need for road users)

ข้อมูลข่าวสารที่จำเป็นต่อผู้ใช้ทางต้องมีลักษณะที่สำคัญ คือ สัญญาณต้องมีความเด่นชัด (conspicuity) ข่าวสารต้องอ่านได้ (legibility) ข่าวสารต้องเข้าใจได้ (comprehensibility) และข่าวสารที่ได้รับต้องเป็นจริง (credibility) ทั้งนี้ถ้ามีการให้ข้อมูลข่าวสารแก่ผู้ใช้ทางที่ถูกต้องเหมาะสมตามหลักแล้ว ก็จะสามารถลดการเกิดอุบัติเหตุลงได้

2.2) ยานพาหนะ

ยานพาหนะเป็นปัจจัยหนึ่งที่มีความสัมพันธ์กับถนน ถ้าการออกแบบถนนไม่คำนึงถึงลักษณะของยานพาหนะที่ใช้แล้ว ก็จะทำให้ความสามารถของยานพาหนะลดลง เช่น การเคลื่อนที่ การมองเห็น การเลี้ยว การหยุด รวมถึงความสามารถในการป้องกันการเกิดอุบัติเหตุ

2.2.1 การเคลื่อนที่ (Maneuverability) เป็นคุณลักษณะของยานพาหนะที่ขึ้นอยู่กับ ขนาด ความยาว ความกว้าง ความสูง และความหนักของรถ

2.2.2 การมองเห็น (Visibility) เป็นคุณลักษณะที่สำคัญของยานพาหนะที่ต้องออกแบบให้สามารถมองเห็นถนนและบริเวณโดยรอบถนนได้อย่างเหมาะสมและได้มาตรฐาน

2.2.3 ลักษณะการเข้าโค้ง (Cornering characteristics) เป็นคุณลักษณะของยานพาหนะที่ขึ้นอยู่กับ การลอยตัว ร่องล้อ ฐานล้อ และตำแหน่งของจุดศูนย์กลางความโน้มถ่วงของยานพาหนะ

2.2.4 การหยุด (Breaking characteristics) ยานพาหนะต้องมีการคำนวณระยะการมองเห็นสำหรับการหยุดรถ (stopping sight distance) ได้ตามมาตรฐานที่กำหนด

2.3) ถนนและสิ่งแวดล้อมข้างทาง

ถนนและสิ่งแวดล้อม เป็นปัจจัยที่สามที่ส่งผลต่อการเกิดอุบัติเหตุบนทางหลวงซึ่งสามารถแยกออกเป็น 4 องค์ประกอบ ได้ดังนี้

2.3.1 วิศวกรรมงานทาง ประกอบไปด้วย ส่วนต่าง ๆ ของการออกแบบถนน เช่น ความกว้างถนน ระดับแนวนอน ระดับแนวตั้ง ความชัน ระยะการมองเห็น พื้นถนน ความฝืดของพื้นถนน ความกว้างของไหล่ทาง และเกาะกลาง

2.3.2 วิศวกรรมจราจร เป็น เครื่องมือที่ใช้ในการจัดการการจราจร เช่น เครื่องหมายจราจรต่าง ๆ หมุดแบ่งช่องจราจร เขตจำกัดความเร็วในระดับต่าง ๆ และการควบคุมจุดเข้าออกของทางเชื่อม เป็นต้น

2.3.3 วัตถุหรือสิ่งกีดขวางข้างทาง เช่น เสาไฟฟ้า ต้นไม้ ป้ายและเสาสัญญาณไฟจราจร ราวกันอันตราย ขอบสะพาน ทางระบายน้ำ ร้านค้า ชุมชม และขอบข้างถนน เป็นต้น

2.3.4 สภาพแวดล้อมรอบข้างทาง เช่น สิ่งแวดล้อมต่าง ๆ ที่อยู่นอเหนือการควบคุมของมนุษย์ เช่น สภาพอากาศ เป็นต้น

3) ปัจจัยจำเพาะที่ส่งผลต่อการเกิดอุบัติเหตุบนทางหลวงที่มีความลาดชันแบบภูเขา

ผู้วิจัยได้ทำการทบทวนวรรณกรรมเพื่อศึกษาปัจจัยจำเพาะบางอย่างที่อาจจะส่งผลต่อการเกิดอุบัติเหตุบนทางหลวงที่มีความลาดชันแบบภูเขา ดังนี้

3.1) ปริมาณจราจร (AADT) จากงานวิจัยของ Kihberg and Tharp (1968, อ้างใน เมษา, 2555) วนิดา และ ลิลี่ (2553) เอกนรินทร์ (2547, อ้างใน เมษา, 2555) สำรวล (2543, อ้างใน เมษา, 2555) และ Caliendo *et al.* (2007, อ้างใน เมษา, 2555) พบว่าปริมาณจราจรมีผลต่อการเกิดอุบัติเหตุบนทางหลวง ทั้งทางหลวงแนวราบและทางลาดชัน โดยพบว่าเมื่อปริมาณการจราจรเพิ่มขึ้นจะส่งผลให้อัตราการเกิดอุบัติเหตุเพิ่มขึ้นด้วย

3.2) ความลาดชัน (Slope) จากงานวิจัยของ Kihberg and Tharp (1968, อ้างใน เมษา, 2555) Shankar *et al.* (1995, อ้างใน เมษา, 2555) สำรวล (2543, อ้างใน เมษา, 2555) ปฏิวัติ (2550, อ้างใน เมษา, 2555) เสริมศักดิ์ (2545, อ้างใน เมษา, 2555) และ เมษา (2555) พบว่า ความลาดชันของถนน เป็นปัจจัยหนึ่งที่ส่งผลต่อการเกิดอุบัติเหตุโดยเฉพาะทางลาดชันแบบภูเขา ทั้งนี้ความลาดชันที่พิจารณาอาจแบ่งได้หลายประเภท เช่น ความลาดชันสูงสุด ความลาดชันเฉลี่ย ความลาดชันก่อนหน้าช่วงถนน และความลาดชันหลังช่วงถนน โดยถ้าค่าความลาดชันเหล่านี้มีค่ามากจะส่งผลทำให้อัตราการเกิดอุบัติเหตุมากขึ้นด้วย

3.2) ความกว้างถนน จากงานวิจัยของ Caliendo *et al.* (2007, อ้างใน เมษา, 2555) Wang *et al.* (2009, อ้างใน เมษา, 2555) ปฏิวัติ (2550, อ้างใน เมษา, 2555) และเอกนรินทร์ (2547, อ้างใน เมษา, 2555) พบว่า ความกว้างถนน เป็นอีกหนึ่งปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุ โดยถ้าค่าความกว้างถนนมีค่ามากจะส่งผลทำให้อัตราการเกิดอุบัติเหตุมากขึ้น

4) ข้อมูลหรือค่าสังเกตการติดตามระยะยาว (Longitudinal data)

ข้อมูลการติดตามระยะยาว เป็นข้อมูลที่มีการเก็บซ้ำ (Repeated measures) จากหน่วยศึกษาเดิม โดยมีการเก็บข้อมูล ณ ช่วงเวลาต่าง ๆ การเก็บข้อมูลซ้ำลักษณะนี้จะพบในการศึกษาทางการแพทย์ การเกษตร และการประกันภัย เป็นต้น โดยค่าสังเกตของตัวแปรตาม (Y) จากแต่ละหน่วยศึกษา (Subject/cluster) จะมีการเก็บซ้ำตามช่วงเวลาที่กำหนด ส่วนตัวแปรอิสระ (X) อาจจะเป็นตัวแปรที่มีค่าคงที่ในทุกช่วงเวลาเช่น จำนวนทางแยก ความลาดชันของถนน เป็นต้น หรือเป็นตัวแปรที่มีค่าเปลี่ยนแปลงตามเวลาเช่นเดียวกับตัวแปรตามก็ได้ เช่น อายุของถนน เป็นต้น ลักษณะโครงสร้างของข้อมูลการศึกษาระยะยาวมีรูปแบบ ดังตารางที่ 1 (Diggle *et al.*, 2002)

เมื่อ y_{it} คือ ค่าสังเกตของตัวแปรตาม Y จากหน่วยศึกษาที่ i ณ เวลาที่ t และตัวแปรอิสระมีจำนวน k ตัว คือ $X_1, X_2, \ldots, X_k$ โดยที่ x_{itk} คือ ค่าสังเกตของตัวแปร X_k จากหน่วยศึกษา ตัวแบบทางสถิติสำหรับข้อมูลศึกษาระยะยาวอาจพิจารณาจาก 2 ตัวแบบ (Diggle *et al.*, 2002) ดังนี้

4.1) ตัวแบบตามขอบ หรือ ตัวแบบค่าเฉลี่ยประชากร (Marginal models or Population average models) ตัวแบบนี้มีเป้าหมาย คือ การอนุมานทางสถิติไปที่ค่าเฉลี่ยของประชากรเป็นหลัก และทำการประมาณค่าพารามิเตอร์ด้วยวิธีการ GEE โดยการสร้างตัวแบบความสัมพันธ์ระหว่างค่าเฉลี่ยประชากรกับตัวแปรอิสระผ่านฟังก์ชันเชื่อมโยง ส่วนความสัมพันธ์ของค่าสังเกตภายในหน่วยศึกษาจะมีการสร้างเป็นตัวแบบต่างหาก แล้วนำผลที่ได้มาใช้ในการปรับค่าประมาณพารามิเตอร์และค่าความแปรปรวนของพารามิเตอร์การถดถอยในตัวแบบความสัมพันธ์ เพื่อหาค่าประมาณของพารามิเตอร์ที่เหมาะสมของตัวแบบ

4.2) ตัวแบบอิทธิพลผสม (Mixed effects Models or Subject-specific models) ตัวแบบนี้จะมุ่งเน้นการอนุมานทางสถิติไปที่หน่วยศึกษาเป็นหลัก โดยการเพิ่มเทอมอิทธิพลสุ่ม (Random effects) เข้ามาในตัวแบบ โดยยอมให้เซตย่อยของสัมประสิทธิ์การถดถอยสามารถมีค่าเปลี่ยนแปลงได้อย่างสุ่มจากหน่วยหนึ่งไปยังอีกหน่วยหนึ่ง โดยมีการนำความสัมพันธ์ของค่าสังเกตภายในหน่วยศึกษาเข้ามาพิจารณาด้วยผ่านทางเทอมอิทธิพลสุ่ม

ตารางที่ 1 โครงสร้างของข้อมูลการศึกษาติดตามระยะยาว

หน่วยศึกษา (Subject)	ตัวแปรตาม (Y)	ตัวแปรอิสระ (X)
		$X_1, X_2, \ldots, X_k$
1	y_{11}	$x_{111} \quad x_{112} \quad \cdots \quad x_{11k}$
	y_{12}	$x_{121} \quad x_{122} \quad \cdots \quad x_{12k}$
	$\vdots$	$\vdots$
	y_{1t}	$x_{1t1} \quad x_{1t2} \quad \cdots \quad x_{1tk}$
2	y_{21}	$x_{211} \quad x_{212} \quad \cdots \quad x_{21k}$
	y_{22}	$x_{221} \quad x_{222} \quad \cdots \quad x_{22k}$
	$\vdots$	$\vdots$
	y_{2t}	$x_{2t1} \quad x_{2t2} \quad \cdots \quad x_{2tk}$
$\vdots$	$\vdots$	$\vdots$
S	y_{s1}	$x_{s11} \quad x_{s12} \quad \cdots \quad x_{s1k}$
	y_{s2}	$x_{s21} \quad x_{s22} \quad \cdots \quad x_{s2k}$
	$\vdots$	$\vdots$
	y_{st}	$x_{st1} \quad x_{st2} \quad \cdots \quad x_{stk}$

การสร้างตัวแบบทางสถิติสำหรับข้อมูลการศึกษาติดตามระยะยาวจะมุ่งศึกษาอธิบายหรือแปลผลในภาพรวมของค่าเฉลี่ยประชากร หรืออธิบายผลในลักษณะเฉพาะของแต่ละหน่วยศึกษา อย่างไรก็ตามตัวแบบทั้งสองได้พัฒนาจากตัวแบบเชิงเส้นนัยทั่วไป (GLMs)

5) ตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Models, GLMs)

ตัวแบบเชิงเส้นทั่วไป (General linear models) มีข้อกำหนดเกี่ยวกับตัวแปรตามต้องมีการแจกแจงแบบปรกติ แต่ในบางสถานการณ์ตัวแปรตามอาจมีรูปแบบการแจกแจงอย่างอื่น เช่น มีลักษณะข้อมูลเป็นเชิงกลุ่ม (Categorical data) หรือข้อมูลจำนวนนับ (Count) ซึ่งเป็นลักษณะของตัวแปรชนิดไม่ต่อเนื่อง ดังนั้นการวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระจะไม่สามารถเขียนให้อยู่ในรูปความสัมพันธ์เชิงเส้นทั่วไปได้ ในปี ค.ศ.1972 นักสถิติที่ชื่อว่า Nelder and Wedderburn (2002) ได้นำเสนอตัวแบบเชิงเส้นนัยทั่วไป (GLMs) เพื่อให้สามารถใช้ได้กับข้อมูลที่

มีลักษณะไม่ต่อเนื่อง (Discrete data) โดยตัวแบบนี้ได้พัฒนาต่อจากตัวแบบเชิงเส้นทั่วไป โดยเป็นตัวแบบสำหรับข้อมูลที่มีการแจกแจงในวงศ์เลขชี้กำลัง (Exponential family) ซึ่งจะครอบคลุมถึงการแจกแจงทวินาม (Binomial distribution) การแจกแจงปัวซง (Poisson distribution) การแจกแจงเรขาคณิต (Geometric distribution) การแจกแจงทวินามเชิงลบ (Negative binomial distribution) การแจกแจงแกมมา (Gamma distribution) และการแจกแจงแบบเลขชี้กำลัง (Exponential distribution) เป็นต้น (Myers *et al.*, 2002)

ตัวแบบ GLMs ประกอบด้วยองค์ประกอบ 3 ส่วน (Fitzmaurice *et al.*, 2004) คือ

1) การแจกแจงของตัวแปรสุ่ม (Y) ที่เป็นตัวแปรตาม อาจเป็นข้อมูลต่อเนื่องหรือไม่ต่อเนื่อง สามารถเขียนอยู่ในวงศ์เลขชี้กำลัง ได้ดังนี้

$$f_Y(y;\theta,\phi) = \exp\left\{\frac{\left[y\theta - b(\theta)\right]}{a(\phi)} + c(y,\phi)\right\} \tag{1}$$

เมื่อ $a(\phi), b(\theta)$ และ $c(y,\phi)$ เป็นฟังก์ชันใด ๆ ที่กำหนด

โดย $b(\theta)$ เป็นฟังก์ชันในเทอมของ θ เท่านั้น

และ $c(y,\phi)$ เป็นฟังก์ชันในเทอมของ y และ ϕ

พารามิเตอร์ θ เรียกว่า พารามิเตอร์ตำแหน่ง (Location parameter) ส่วน ϕ เรียกว่า พารามิเตอร์การกระจาย (Dispersion parameter) สำหรับฟังก์ชัน $a(\phi)$ มักจะมีรูปแบบเป็น $a(\phi) = \phi/w$ และฟังก์ชัน $c = c(y,\phi/w)$ โดยที่ w แทนค่าน้ำหนักที่ทราบค่าของแต่ละค่าสังเกต โดยทั่วไปจะกำหนดให้ $w = 1$ สำหรับทุก ๆ ค่าสังเกต

ค่าเฉลี่ยและค่าความแปรปรวนของตัวแปรตาม (Y) ที่เขียนในวงศ์เลขชี้กำลัง สามารถหาได้จาก

$$E(Y) = \frac{db(\theta)}{d\theta} = b'(\theta) = \mu \text{ ซึ่งจะได้ว่า } \theta = b'^{-1}(\mu) \text{ และ}$$

$$\operatorname{var}(Y) = \frac{d^2 b(\theta)}{d\theta^2} a(\phi)$$
$$= \mathrm{b}''(\theta) a(\phi)$$
$$= \frac{d\mu}{d\theta} a(\phi)$$

โดยที่ $b'(\theta)$ คือ การหาอนุพันธ์อันดับที่ 1 ของฟังก์ชัน $b(\theta)$ เมื่อเทียบกับพารามิเตอร์ θ

$b''(\theta)$ คือ การหาอนุพันธ์อันดับที่ 2 ของฟังก์ชัน $b(\theta)$ เมื่อเทียบกับพารามิเตอร์ θ

ความแปรปรวนของตัวแปรตาม (Y) ที่เขียนในวงศ์เลขชี้กำลัง สามารถเขียนเป็นผลคูณระหว่างฟังก์ชันสองฟังก์ชัน คือ $b''(\theta)$ ซึ่งเป็นฟังก์ชันที่ขึ้นกับ θ หรือค่าเฉลี่ยกับ $a(\phi)$ ซึ่งเป็นฟังก์ชันที่เป็นอิสระจาก θ ดังนั้นฟังก์ชันความแปรปรวนของ Y จะถูกพิจารณาเหมือนเป็นฟังก์ชันของค่าเฉลี่ย (μ) (Myers, 2002) คือ

จาก $\operatorname{var}(Y) = b''(\theta) a(\phi)$

$$= \frac{d\mu}{d\theta} a(\phi) \text{ หรือ } b''\left[b'^{-1}(\mu)\right] a(\phi)$$

โดยที่ $b''(\theta) = \frac{d\mu}{d\theta} = \operatorname{var}(\mu)$

ซึ่ง $\operatorname{var}(\mu)$ จะถูกเรียกเป็น Variance Function ที่อธิบายความสัมพันธ์ระหว่างค่าเฉลี่ยและค่าความแปรปรวน ดังนั้น $\operatorname{var}(Y) = a(\phi)\operatorname{var}(\mu)$ หรือ $\operatorname{var}(Y) = \phi \operatorname{var}(\mu)$ ถ้า $a(\phi) = \phi$

ในการศึกษาครั้งนี้ ตัวแปรตามหรือ Y เป็น จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต โดยที่ค่าสังเกต y จะมีค่าเป็นศูนย์หรือจำนวนเต็มบวกเท่านั้น ดังนั้นฟังก์ชันการแจกแจงความน่าจะเป็นของค่าสังเกต y จะมีการแจกแจงปัวซง ดังสมการข้างล่างนี้

$$f(y;\mu) = \frac{e^{-\mu}\mu^{y}}{y!} \quad \text{เมื่อ } y = 0,1,2,\ldots$$

สามารถเขียนฟังก์ชันของ y ให้อยู่ในรูปแบบของฟังก์ชันแบบเลขชี้กำลังได้ดังนี้

$$f(y;\mu)=\exp\left[y\log(\mu)-\mu-\log(y!)\right]$$

โดยที่ $\theta=\log(\mu)$, $b(\theta)=\mu=\exp(\theta)$, $a(\phi)=1$ หรือ $\phi=1$ และ $c(y,\phi)=-\log(y!)$

จากการจัดรูปทำให้พารามิเตอร์การกระจาย ของการแจกแจงปัวซง มีค่าเท่ากับ 1 นั้นคือ $\phi=1$ ส่วนค่าเฉลี่ยและความแปรปรวนของตัวแปรตาม Y คำนวณได้จาก

$$E(Y)=b'(\theta)=\exp(\theta)=\mu \text{ และ } \operatorname{var}(Y)=b''(\theta)=\exp(\theta)=\mu$$

2) องค์ประกอบระบบ (Systematic component) หรือโครงสร้างหลักของ GLMs สามารถเขียนได้ในรูปแบบของตัวทำนายเชิงเส้น (Linear Predictor) ของตัวแปรอิสระ $X_1,\ldots,X_k$ ที่มีความสัมพันธ์กับเวกเตอร์พารามิเตอร์ β ดังสมการ

$$\eta=X\beta;\eta_i=x_i'\beta=\beta_0+\sum_{i=1}^{k}\beta_i x_i \quad ;i=1,2,\ldots,k$$

โดยที่ η แทน ตัวทำนายเชิงเส้นเป็นเวกเตอร์ขนาด *n x 1*

X แทน เมตริกซ์ของตัวแปรอิสระที่มีขนาด *n x p* ประกอบด้วย n ค่าสังเกต

x_i' แทน เวกเตอร์ของค่าสังเกตของตัวแปรอิสระ $X_1,\ldots,X_k$ ในแถวข้อมูลที่ i โดย $x_i'=\begin{pmatrix}1 & x_1 & x_2 & \cdots & x_k\end{pmatrix}$

β แทน เวกเตอร์ของพารามิเตอร์ $(\beta_1,\ldots,\beta_p)'$ ที่มีขนาด *p x 1; p = k+1*

3) ฟังก์ชันเชื่อมโยง (Link function) แสดงความสัมพันธ์ระหว่างค่าเฉลี่ยของตัวแปรตาม (μ) กับตัวทำนายเชิงเส้น (η) โดยเขียนแทนฟังก์ชันเชื่อมโยง ด้วยสัญลักษณ์ $g(\cdot)$ ดังสมการ

$$g(\mu_i)=\eta_i=\beta_0+\beta_1x_1+\beta_2x_2+\cdots+\beta_kx_k$$

สำหรับการแจกแจงปัวซงจะเชื่อมความสัมพันธ์ระหว่าง μ กับ η ด้วยฟังก์ชันของล็อก เรียกว่า ฟังก์ชันเชื่อมโยง ของล็อก (Log Link Function) นั่นคือ $g(\mu)=\log\mu$ สามารถเขียนแทน $g(\mu)$ ด้วย η ดังนี้

$$\log(\mu_i) = \eta_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

หรือเรียกเป็นตัวแบบล็อกลิเนียร์ (Log-linear model) ซึ่งเขียนได้ดังสมการ

$$\log(\mu_i) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

การประมาณค่าพารามิเตอร์ β ในตัวแบบ GLMs สามารถหาได้จากวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method; ML) กรณีตัวแปรตามมีการแจกแจงปัวซง สามารถจัดให้อยู่ในวงศ์เลขชี้กำลัง จะได้ค่าความแปรปรวนของตัวแปรตาม Y เป็น $\text{var}(Y) = \phi\,\text{var}(\mu)$ โดยที่ $\text{var}(\mu) = \mu$ ส่วนพารามิเตอร์การกระจาย ϕ เป็นค่าที่ได้จากการประมาณค่าจากข้อมูลที่มีการแจกแจงปัวซง โดยจะได้ $\phi = 1$

การประมาณพารามิเตอร์ β ด้วยวิธีภาวะน่าจะเป็นสูงสุด (ML) เริ่มด้วยการกำหนดฟังก์ชันภาวะน่าจะเป็น (Likelihood function) ดังนี้

$$l = \log L = \sum_{i=1}^{n} \left[\frac{y_i\theta_i - b(\theta_i)}{a(\phi)} + c(y_i, \phi) \right]$$

และทำการหาอนุพันธ์อันดับที่หนึ่งเทียบกับ β แล้วกำหนดให้เท่ากับศูนย์ ได้สมการประมาณค่า (Estimation equation) เขียนในรูปเวกเตอร์ได้ดังนี้

$$\frac{\partial l}{\partial \beta} = \sum_{i=1}^{n} \left(\frac{\partial \theta_i}{\partial \beta} \right) \left(\frac{y_i - \mu_i}{\phi} \right) = 0$$

เนื่องจากสมการดังกล่าวอยู่ในรูปแบบไม่เป็นเชิงเส้น (Non-linear) ดังนั้นจึงไม่สามารถหาค่าประมาณ β ของสมการนี้ได้โดยตรง จึงต้องใช้วิธีการคำนวณเชิงตัวเลขแบบวนซ้ำ (Numerical Iteration) เช่น วิธีนิวตัน-ราฟสัน (Newton-Raphson Algorithm) และวิธีคะแนนฟิชเชอร์ (Fisher-Scoring Algorithm) เป็นต้น สามารถใช้โปรแกรมทางสถิติมาช่วยวิเคราะห์ เช่น โปรแกรม R เป็นต้น

6) ตัวแบบสมการประมาณค่านัยทั่วไป (Generalized Estimating Equations, GEE)

ตัวแบบเชิงสถิติสำหรับข้อมูลที่มีการติดตามระยะยาวที่มีการวัดซ้ำค่าสังเกตของตัวแปรตามที่สนใจภายใต้หน่วยศึกษาเดียวกันมักจะพบว่าค่าสังเกตดังกล่าวมีความสัมพันธ์กันเอง

(Correlated data) เพื่อให้วิธีการทางสถิติมีความสอดคล้องกับข้อมูลมากยิ่งขึ้น Liang and Zeger (1986) ได้พัฒนาและนำเสนอวิธีการทางสถิติที่ชื่อว่า สมการประมาณค่านัยทั่วไป (GEEs) ที่ขยายมาจาก ตัวแบบเชิงเส้นนัยทั่วไป สำหรับใช้วิเคราะห์ข้อมูลที่มีการติดตามระยะยาว โดยตัวแบบที่ได้จะเป็นตัวแบบตามขอบ ซึ่งมีการอธิบายผลของค่าประมาณพารามิเตอร์ในภาพรวมค่าเฉลี่ยประชากร พร้อมทั้งพิจารณาโครงสร้างความสัมพันธ์ของค่าสังเกตของตัวแปรตามที่มีการวัดซ้ำบนหน่วยศึกษาหนึ่ง ๆ เรียกความสัมพันธ์ดังกล่าวว่า ความสัมพันธ์ภายในหน่วยศึกษา (Within-Subject) โดยอยู่บนพื้นฐานของการกำหนดรูปแบบของโครงสร้างความสัมพันธ์ในลักษณะของเมทริกซ์ความสัมพันธ์ที่เรียกว่า Working Correlation Matrix ($R_i(\rho)$) สำหรับสมการประมาณค่านัยทั่วไป มีลักษณะดังนี้

กำหนดให้ T เป็นจำนวนเวลาการวัดซ้ำของจำนวนหน่วยศึกษา S หน่วย และกำหนดให้ y_{it} เป็นค่าสังเกตของหน่วยศึกษาที่ $i = 1,2,\ldots,S$ จากการวัดซ้ำครั้งที่ $t = 1,2,\ldots,T$ และมีความสัมพันธ์กับเวกเตอร์ค่าสังเกตของตัวแปรอิสระ $X_1, X_2, \ldots, X_k$ ที่ได้จากการวัดครั้งที่ t คือ $\mathbf{x}'_{it} = (1, x_{1t}, \ldots, x_{kt})$ เมื่อ $k = 1,2,\ldots,k$ เนื่องจากสมการประมาณค่านัยทั่วไปขยายมาจากตัวแบบเชิงเส้นนัยทั่วไป ดังนั้นจึงมีข้อตกลงเบื้องต้นที่เหมือนกัน ยกเว้นสมการประมาณค่านัยทั่วไปจะไม่คำนึงถึงข้อสมมติเกี่ยวกับการแจกแจงของตัวแปรตาม กล่าวคือ ตัวแปรตามไม่จำเป็นต้องอยู่ในวงศ์เลขชี้กำลังก็ได้ แต่อย่างไรก็ตามค่าเฉลี่ย ฟังก์ชันความแปรปรวน และฟังก์ชันเชื่อมโยง ของตัวแบบเชิงเส้นจะต้องมีคุณสมบัติเป็นไปตามของตกลงของตัวแบบเชิงเส้นนัยทั่วไป ดังนั้นจึงสรุปได้ว่า สมการประมาณค่านัยทั่วไป มีคุณลักษณะที่สำคัญ 3 ประการ ดังนี้

1) ค่าเฉลี่ยตามขอบ (Marginal mean) ของตัวแปรตาม Y_{it} คือ $E(Y_{it}) = \mu_{it}$ ขึ้นอยู่กับชุดตัวแปรอิสระ โดยอาศัยความสัมพันธ์ผ่านฟังก์ชันเชื่อมโยง ในรูปแบบเชิงเส้น ดังนี้

$$g(\mu_{it}) = \eta_{it} = \mathbf{x}'_{it}\beta \qquad (2)$$

เมื่อ $\beta = (\beta_1, \beta_2, \ldots, \beta_p)'$ เป็นสัมประสิทธิ์การถดถอยขนาดเวคเตอร์ $p \times 1$

2) ค่าความแปรปรวนโดยรวม (Marginal variance) ของตัวแปรตาม Y_{it} จะเป็นฟังก์ชันที่ขึ้นอยู่กับค่าเฉลี่ย

$$\mathrm{Var}(Y_{it}) = \varphi\,\mathrm{Var}(\mu_{it}) \qquad (3)$$

เมื่อ $\mathrm{Var}(\mu_{it})$ เป็นฟังก์ชันที่ทราบค่าของ μ_{it} และ φ เป็นพารามิเตอร์การกระจายในตัวแบบวางนัยทั่วไปที่ทราบค่าหรือประมาณจากข้อมูลได้

3) ความสัมพันธ์ของค่าสังเกตของตัวแปรตามที่มีการวัดซ้ำของหน่วยศึกษาเดียวกัน คือ $\mathrm{Corr}(y_{it}, y_{ik}) = \rho_{tk}$ ในสมการประมาณค่านัยทั่วไปจะพิจารณาถึงโครงสร้างความสัมพันธ์ในรูปแบบของ Working Correlation Matrix ต่าง ๆ เช่น Independent, Exchangeable, First-order Autoregressive: AR(1), Toeplitz และ Unstructured เป็นต้น

สำหรับการศึกษาครั้งนี้เป็นการสร้างตัวแบบเชิงสถิติสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่มีการเก็บข้อมูลซ้ำในระยะเวลาติดตาม 5 ปี โดยตัวแปรตาม คือ จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ซึ่งเป็นตัวแปรจำนวนนับ (Count) มีการแจกแจงปัวซง สามารถวิเคราะห์ด้วยวิธี GEE ได้ดังนี้

ตัวแปรตามสามารถกำหนดความสัมพันธ์กับชุดตัวแปรอิสระหรือตัวทำนาย ผ่านฟังก์ชันเชื่อมโยง ที่เรียกว่า Log Link function ดังนี้

$$g(\mu_{it}) = \log(\mu_{it}) = \mathbf{x}'_{it}\beta \tag{4}$$

ฟังก์ชันความแปรปรวนของตัวแปรตาม เมื่อถูกกำหนดด้วยค่าของชุดตัวแปรอิสระจะขึ้นอยู่กับค่าเฉลี่ยของตัวแปรตาม ดังนั้น จากสมการ (3) เมื่อกำหนดให้ $\mathrm{Var}(\mu_{it}) = \mu_{it}$ และ $\varphi = 1$ จะได้ว่า

$$\mathrm{Var}(Y_{it}) = \mu_{it} \tag{5}$$

ความสัมพันธ์ระหว่างค่าสังเกตของหน่วยศึกษาเดียวกัน ที่ปรากฏใน Working Correlation Matrix จะสมมติให้อยู่ในรูปแบบของ Independent, Exchangeable, First-order Autoregressive: AR(1), Toeplitz และ Unstructured ดังนี้

1) รูปแบบ Independent มีโครงสร้างความสัมพันธ์ดังต่อไปนี้

$$\text{Corr}\left(y_{it}, y_{ip}\right) = \begin{cases} 1 & ,t = k \\ 0 & ,t \neq k \end{cases} \text{ เขียนในรูปเมทริกซ์ คือ } \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$$

2) รูปแบบ Exchangeable ถูกนำเสนอ Shoukri และ Chaudhary (2007) มีโครงสร้างความสัมพันธ์ดังต่อไปนี้

$$\text{Corr}\left(y_{it}, y_{ik}\right) = \begin{cases} 1 & ,t = k \\ \rho & ,t \neq k \end{cases} \text{ เขียนในรูปเมทริกซ์ คือ } \begin{bmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{bmatrix}$$

ความสัมพันธ์ของค่าสังเกตภายในหน่วยศึกษาเดียวกันจะมีค่าเท่ากัน เท่ากับ ρ ณ ช่วงเวลาที่ $t \neq p$

3) รูปแบบ First-order Autoregressive: AR(1) มีโครงสร้างความสัมพันธ์ดังต่อไปนี้

$$\text{Corr}\left(y_{it}, y_{i,t+j}\right) = \rho^{j} \text{ สำหรับ } j = 0,1,\ldots,n_i - t$$

$$\text{เขียนในรูปเมทริกซ์ คือ } \begin{bmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{bmatrix}$$

ความสัมพันธ์ของค่าสังเกตภายในหน่วยศึกษาเดียวกันจะมีค่าลดลงตามระยะห่างของช่วงเวลาที่วัดข้อมูลซ้ำ ซึ่งค่าความสัมพันธ์จะอยู่ในรูปของ ρ^{j} เมื่อ j คือ จำนวนช่วงเวลาก่อนหน้า (lag time) เมื่อ $j = 0,1,\ldots,n_i - t$

4) รูปแบบ Toeplitz มีโครงสร้างความสัมพันธ์ดังต่อไปนี้

$$\mathrm{Corr}\left(y_{it}, y_{i,t+j}\right) = \begin{cases} 1 & ,t = k \\ \rho_j & ,t = 1,2,\ldots,n_i - j \end{cases}$$

เขียนในรูปเมทริกซ์ คือ

$$\begin{bmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{bmatrix}$$

5) รูปแบบ Unstructured มีโครงสร้างความสัมพันธ์ดังต่อไปนี้

$$\mathrm{Corr}\left(y_{it}, y_{ik}\right) = \begin{cases} 1 & ,t = k \\ \rho_{tp} & ,t \neq k \end{cases}$$ เขียนในรูปเมทริกซ์ คือ

$$\begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots & \rho_{1p} \\ \rho_{12} & 1 & \rho_{23} & \cdots & \rho_{2p} \\ \rho_{13} & \rho_{23} & 1 & \cdots & \rho_{3p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{1p} & \rho_{2p} & \rho_{3p} & \cdots & 1 \end{bmatrix}$$

สำหรับการประมาณค่า β หรือค่าสัมประสิทธิ์การถดถอยในสมการประมาณค่านัยทั่วไปจะขยายจากสมการ เมื่อความสัมพันธ์ของข้อมูลเป็นอิสระต่อกัน สามารถเขียนได้ดังนี้

$$\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} \left[V_i\left(\hat{\rho}\right) \right]^{-1} \left[y_i - \mu_i \right] = 0 \quad ; \mu_i' = \left[\mu_{i1}, \mu_{i2}, \ldots, \mu_{it} \right] \qquad (6)$$

เมื่อ y_i เป็นค่าสังเกตที่ i ของตัวแปรตาม Y

เนื่องจากสมการ (6) เป็นสมการที่อยู่ในรูปแบบไม่เป็นเชิงเส้น ในการประมาณค่า β ไม่สามารถแก้สมการนี้ได้โดยตรง จึงต้องใช้วิธีการคำนวณเชิงตัวเลขแบบวนซ้ำของวิธี Quasi-Likelihood เพื่อประมาณค่าพารามิเตอร์ และวิธี Robust สำหรับประมาณค่า ρ โดยใช้คำสั่งในโปรแกรมสำเร็จรูป R และสามารถสรุปเป็นขั้นตอนได้ดังนี้

1) ประมาณค่าเริ่มต้น (Initial value) ของพารามิเตอร์ β ด้วยตัวแบบเชิงเส้นนัยทั่วไป ในกรณีนี้จะกำหนดให้ความสัมพันธ์ของค่าสังเกตภายในหน่วยศึกษาเดียวกันเป็นอิสระต่อกัน

2) คำนวณ โครงสร้างความสัมพันธ์ของ Working Correlation Matrix (R) โดยประมาณค่าโครงสร้างความสัมพันธ์ภายในแต่ละหน่วยศึกษา คือ $R_i(\rho)$ จากค่าที่คำนวณได้ของ β ณ ปัจจุบัน แล้วคำนวณค่าเศษเหลือแบบ Pearson $\left(e_{it}\right)$ ได้ดังนี้ $e_{it} = \dfrac{y_{it} - \mu_{it}}{\sqrt{V(\mu_{it})}}$ จากนั้นนำค่า e_{it} ที่ได้มาคำนวณค่าประมาณของ ϕ โดยที่ $\phi = \dfrac{1}{M-p}\sum_{i=1}^{S}\sum_{j=1}^{t} e_{ij}^2$ โดยที่ M คือจำนวนค่าสังเกตทั้งหมดของข้อมูล $(M = S \times T)$ และประมาณค่าของ ρ จากฟังก์ชันของ ϕ และ e_{it} ตามรูปแบบความสัมพันธ์ของ $R_i(\rho)$ ที่กำหนด

3) คำนวณเมทริกซ์ความแปรปรวนร่วม (Working Covariance Matrix) ของหน่วยศึกษาที่ i ได้จาก $V_i(\rho) = \phi A_i^{1/2} R_i(\rho) A_i^{1/2}$

4) ทำการปรับค่าพารามิเตอร์ β ไปเรื่อย ๆ โดยการเพิ่มจำนวนการวนซ้ำดังนี้

$$\beta_{r+1} = \beta_r - \left[\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1} \frac{\partial \mu_i}{\partial \beta}\right]^{-1} \left[\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1} (y_i - \mu_i)\right]$$

5) คำนวณซ้ำในข้อ 2 ถึง 4 จนกระทั่งค่าประมาณของพารามิเตอร์ β นั้นคือ $\hat{\beta}$ ลู่เข้าสู่ค่าคงที่

สำหรับค่าประมาณความแปรปรวนของ $\hat{\beta}$ นั้นคือ $\mathrm{Cov}(\hat{\beta})$ ที่ได้จากการกำหนดโครงสร้างความสัมพันธ์ของค่าสังเกตภายในหน่วยศึกษาเดียวกันจะเรียกว่า ตัวประมาณค่าความแปรปรวนด้วยวิธีแซนวิช (Sandwich variance estimator) โดย Liang และ Zeger (1986) ได้เสนอไว้ดังนี้

$$\hat{\mathrm{C}}\mathrm{ov}(\hat{\beta}) = \left[\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1} \frac{\partial \mu_i}{\partial \beta}\right]^{-1} \left[\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1} (y_i - \mu_i)(y_i - \mu_i)' V_i^{-1} \frac{\partial \mu_i}{\partial \beta}\right] \left[\sum_{i=1}^{N} \frac{\partial \mu_i'}{\partial \beta} V_i^{-1} \frac{\partial \mu_i}{\partial \beta}\right]^{-1}$$

7. ตัวแบบการประมาณค่านัยทั่วไปสำหรับจำนวนนับมีศูนย์มาก (Generalized Estimating Equations based Zero-inflated Models)

ในบางกรณีพบว่า ค่าสังเกตของตัวแปรตามที่มีลักษณะเป็นจำนวนนับ (count data) มีศูนย์มาก เช่น การศึกษาติดตามจำนวนอุบัติเหตุที่เกิดบนทางหลวง 5 ปี อาจพบว่ามีจำนวนการไม่

เกิดอุบัติในช่วงเวลาที่ติดตามเป็นจำนวนมาก ซึ่งจะกำหนดเหตุการณ์ให้เป็นศูนย์ อย่างไรก็ตาม ลักษณะการแจกแจงของตัวแปรตามดังกล่าวก็จะยังมีการแจกแจงปัวซง หรือ การแจกแจงทวินามเชิงลบ แต่เป็นกรณีพิเศษที่เรียกว่า การแจกแจงปัวซงที่มีจำนวนศูนย์มาก (Zero-inflated Poisson Distribution: ZIP) และ การแจกแจงทวินามเชิงลบที่มีจำนวนศูนย์มาก (Zero-inflated Negative Binomial Distribution: ZINB)

7.1) ตัวแบบการประมาณค่านัยทั่วไปสำหรับการแจกแจงปัวซงที่มีจำนวนศูนย์มาก (Generalized Estimating Equations based Zero-inflated Poisson Model)

ตัวแบบนี้พัฒนาขึ้นภายใต้ข้อมูลการศึกษาระยะยาวที่มีค่าสังเกตเป็นจำนวนนับ (count data) มีความสัมพันธ์กันเองภายในหน่วยศึกษา (subjects/cluster) เดียวกัน และพบว่าค่าสังเกตดังกล่าวมีค่าศูนย์มากภายใต้การแจกแจงปัวซง กำหนดให้ Y_{it} เป็นเป็นจำนวนนับของหน่วยศึกษาที่ $i=1,2,\ldots,S$ จากการวัดซ้ำครั้งที่ $t=1,2,\ldots,T$ โดย Y_{it} จะมีรูปแบบการแจกแจงผสม (Mixture distribution) ระหว่างตัวแปรสุ่มที่มีค่าเป็น 0 ด้วยความน่าจะเป็น p_{it} และตัวแปรสุ่มที่มีค่ามากกว่า 0 มีการแจกแจงปัวซงที่มีค่าเฉลี่ย λ_{it} ด้วยความน่าจะเป็น $1-p_{it}$ ดังนั้นกำหนดให้ W_{it} เมื่อ $w_{it}=0,1,2,\ldots,$ เป็นตัวแปรสุ่มที่มีการแจกแจงปัวซงและมีฟังก์ชันมวลความน่าจะเป็นดังนี้

$$\Pr\left(W_{it}=w_{ij}\right)=f\left(w_{it}|\lambda_{it}\right)=\frac{e^{-\lambda_{it}}\lambda_{it}^{w_{it}}}{w_{it}!},$$

ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่ม W_{it} กำหนดได้โดย $E\left(W_{it}\mid\lambda_{it}\right)=\lambda_{it}$ และ $V\left(W_{it}\mid\lambda_{it}\right)=\lambda_{it}$ ตามลำดับ เมื่อพิจารณากำหนดให้ Y_{it} มีการแจกแจงผสมระหว่างตัวแปรสุ่มที่มีค่าเป็น 0 ด้วยความน่าจะเป็น p_{it} และตัวแปรสุ่ม W_{it} มีการแจกแจงทวินามเชิงลบที่มีค่าเฉลี่ย λ_{it} ด้วยความน่าจะเป็น $1-p_{it}$ แล้ว ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม Y_{it} สามารถกำหนดได้ดังนี้

$$\begin{aligned}\Pr\left(Y_{it}=y\right) &= \begin{cases} p_{it}+\left(1-p_{it}\right)\Pr\left(W_{ij}=0\right) & \text{ถ้า } y=0 \\ \left(1-p_{it}\right)\Pr\left(W_{ij}=y\right) & \text{ถ้า } y\geq 0 \end{cases} \\ &= \begin{cases} p_{it}+\left(1-p_{it}\right)e^{-\lambda_{it}} & \text{ถ้า } y=0 \\ \left(1-p_{it}\right)\dfrac{e^{-\lambda_{it}}\lambda_{it}^{y}}{y!} & \text{ถ้า } y\geq 0 \end{cases}\end{aligned}$$

โดยตัวแปรสุ่ม Y_{it} จะมีค่าเฉลี่ยและความแปรปรวนดังนี้

$$E(Y_{it}) = (1 - p_{it})\lambda_{it}$$

และ

$$V(Y_{it}) = (1 - p_{it})\lambda_{it}(1 + p_{it}\lambda_{it})$$

เมื่อค่าสังเกตเป็นอิสระต่อกันแล้ว ตัวแบบสำหรับข้อมูลค่าสังเกตที่เป็นจำนวนนับและมีศูนย์มากจะถูกพัฒนาโดยการใช้ตัวแบบปัวซงที่มีศูนย์มาก (ZIP) และสามารถกำหนดความสัมพันธ์กับชุดตัวแปรอิสระหรือตัวทำนาย ผ่านฟังก์ชันเชื่อมโยง 2 ตัวให้กับ p_{it} และ λ_{it} ด้วย Log Link function และ Logit function ได้ดังนี้

$$g(\lambda_{it}) = \log(\lambda_{it}) = \mathbf{x}'_{it}\beta \quad (7)$$

โดยที่ $\mathbf{x}'_{it}$ เป็นชุดตัวแปรอิสระหรือตัวทำนายของ λ_{it}

และ
$$g(p_{it}) = \text{logit}(p_{it}) = \mathbf{z}'_{it}\gamma \quad (8)$$

โดยที่ $\mathbf{z}'_{it}$ $\mathbf{x}'_{it}$ เป็นชุดตัวแปรอิสระหรือตัวทำนายของ p_{it}

ทั้งนี้ชุดตัวแปรอิสระหรือตัวทำนายของ λ_{it} และ p_{it} จะเป็นตัวเดียวกันหรือแตกต่างกันก็ได้ ฉะนั้นแล้วค่าเฉลี่ยของ Y_{it} ที่กำหนดโดย $E(Y_{it}) = \mu_{it} = (1 - p_{it})\lambda_{it}$ ก็จะขึ้นอยู่กับพารามิเตอร์ β และ γ

สำหรับการประมาณค่า β และ γ ในสมการประมาณค่านัยทั่วไปจะพิจารณาภายใต้ความสัมพันธ์ระหว่างค่าสังเกตของหน่วยศึกษาเดียวกัน $R_i(\rho)$ (Hall and Zhang, 2004) ดังนี้

$$\sum_{i=1}^{S} \begin{pmatrix} \dfrac{\partial \mu_i'}{\partial \beta} \\ \dfrac{\partial \mu_i'}{\partial \gamma} \end{pmatrix} \left[V_i(\rho, \beta, \gamma)\right]^{-1} \left[y_i - \mu_i\right] = 0 \quad (9)$$

โดยที่ y_i เป็นค่าสังเกตที่ i ของตัวแปรตาม Y_i, $\mu_i' = [\mu_{i1}, \mu_{i2}, \ldots, \mu_{it}]$, $V_i(\rho, \beta, \gamma) = A_i^{1/2} R_i(\rho) A_i^{1/2} R_i$ เมื่อ $A_i = Diag[V(Y_{it})], t = 1, 2, \ldots, T$ อย่างไรก็ตาม Hall and Zhang (2004) ได้ให้ข้อเสนอแนะว่า สมการ (9) ไม่สามารถใช้ในการประมาณค่า β และ γ ได้ เนื่องจากพารามิเตอร์ทั้งสองตัวต้องใช้สารสนเทศร่วมกัน ดังนั้นจึงเสนอการประมาณค่าด้วยการแยกสมการ

ด้วยการสร้างตัวแปรแฝง u_{it} (latent variable) ขึ้นมาเพื่อใช้กำหนดการแจกแจงศูนย์มากหรือการแจกแจงปัวซงของ Y_{it}

กำหนดให้ $u_{it}=0$ เมื่อ $Y_{it}\sim Poi(\lambda_{it})$, $u_{it}=1$ เมื่อ Y_{it} มีการแจกแจงศูนย์มาก และ $\Pr(u_{it}=1)=p_{it}$ ดังนั้นสมการประมาณค่านัยทั่วไปสำหรับ γ สามารถกำหนดได้ดังนี้

$$\sum_{i=1}^{S}\frac{\partial p_i'}{\partial\gamma}\left[V_{\gamma i}\right]^{-1}\left[u_i-p_i\right]=0 \tag{10}$$

โดยที่ $u_i=\left[u_{i1},\ldots,u_{iT_i}\right]'$, $p_i'=\left[p_{i1},\ldots,p_{iT_i}\right]$ และ $p_{it}(\gamma)=\dfrac{\exp(\mathbf{z}_{it}'\gamma)}{1+\exp(\mathbf{z}_{it}'\gamma)}$ แสดงนัยด้วย

$$\frac{\partial p_{it}}{\partial\gamma}=z_{it}\frac{\exp(\mathbf{z}_{it}'\gamma)}{\left[1+\exp(\mathbf{z}_{it}'\gamma)\right]^2}$$

และ $V_{\gamma i}=A_i^{1/2}R_{1i}(\rho_1)A_i^{1/2}R_i$ เป็นเมทริกซ์ความแปรปรวนของตัวแปรสุ่ม u_i เมื่อ $A_i=$ $Diag\left[p_{i1}(1-p_{i1}),p_{i2}(1-p_{i2}),\ldots,p_{iT_i}(1-p_{iT_i})\right]$ ของความแปรปรวนของตัวแปรสุ่ม u_{it} ที่ลำดับ t^{th} และ $R_{1i}(\rho_1)$ เป็นเมทริกซ์สหสัมพันธ์ของ u_i ดังนั้นสามารถประมาณค่า β ในสมการประมาณค่านัยทั่วไป ได้ดังนี้

$$\sum_{i=1}^{S}\frac{\partial\lambda_i'}{\partial\beta}\left[V_{\beta i}\right]^{-1}Diag\left[(1-\mu_i)(y_i-\lambda_i)\right]=0 \tag{11}$$

ให้ $\lambda_i'=\left(\lambda_{i1},\lambda_{i2},\ldots,\lambda_{iT_i}\right)$ เมื่อ $\log(\lambda_{it})=\mathbf{x}_{it}'\beta$ ฉะนั้น $\lambda_{it}=\exp(\mathbf{x}_{it}'\beta)$ และ $\dfrac{\partial\lambda_{it}}{\partial\beta}=x_{it}\exp(\mathbf{x}_{it}'\beta)$ จากเดิม $V_{\beta i}=\left(D_i^{1/2}R_{2i}(\rho_2)D_i^{1/2}\right)$ เป็นเมทริกซ์ความแปรปรวนของ Y_i เมื่อ $D_i=Diag\left[\lambda_{i1}(1+\tau\lambda_{i1}),\lambda_{i2}(1+\tau\lambda_{i2}),\ldots,\lambda_{iT_i}(1+\tau\lambda_{iT_i})\right]$ และ $R_{2i}(\rho_2)$ เป็นเมทริกซ์สหสัมพันธ์ของ Y_i เมทริกซ์ทแยงมุม $Diag[1-u_i]=Diag\left[1-u_{i1},\ldots,1-u_{iT_i}\right]$ ในสมการ (11) ที่เป็นตัวชี้วัดของ Y_{it} ที่มาจากการแจกแจงปัวซงเท่านั้น (เช่น $u_{it}=0$) ซึ่งเป็นตัวที่ใช้ในการประมาณค่าสมการที่ (X)

เมื่อกำหนดให้ $u_{it}, i=1,2,\ldots,S, t=1,2,\ldots,T_i$ การประมาณค่า β และ γ ไม่สามารถแก้สมการ (10) และ (11) ได้โดยตรง จึงต้องใช้วิธีการคำนวณเชิงตัวเลขแบบวนซ้ำของวิธีคะแนนฟิชเชอร์ (Fisher-Scoring Algorithm) เพื่อประมาณค่าพารามิเตอร์ดังกล่าว อย่างไรก็ตาม u_{it} เป็นเพียงตัวแปรสุ่มแฝง (latent variable, unobserved variables) เท่านั้นการแก้สมการดังกล่าว

จะไม่สามารถใช้งานได้โดยตรง ฉะนั้นจึงใช้วิธีการ expectation-solution (Hall and Zhang, 2004; Rosen *et al.*, 2000) ในทุก ๆ รอบวนซ้ำ ด้วยการพิจารณาแทนค่าเฉลี่ยแบบมีเงื่อนไขของ u_{it} เมื่อกำหนด Y_{it} ในสมการ (10) และ (11) ในการประมาณค่าพารามิเตอร์ทั้งหมด (λ และ β) โดยใช้คำสั่งในโปรแกรมสำเร็จรูป R

7.2) ตัวแบบการประมาณค่านัยทั่วไปสำหรับการแจกแจงทวินามเชิงลบที่มีจำนวนศูนย์มาก (Generalized Estimating Equations based Zero-inflated Negative Binomial Model)

ตัวแบบนี้พัฒนาขึ้นภายใต้ข้อมูลการศึกษาระยะยาวที่มีค่าสังเกตเป็นจำนวนนับ (count data) มีความสัมพันธ์กันเองภายในหน่วยศึกษา (subjects/cluster) เดียวกัน และพบว่าค่าสังเกตดังกล่าวมีค่าศูนย์มากภายใต้การแจกแจงทวินามเชิงลบ เพื่อแก้ปัญหาความแปรปรวนมากกว่าค่าเฉลี่ยในตัวแบบปัวซง (Over-dispersion problem) กำหนดให้ Y_{it} เป็นจำนวนนับของหน่วยศึกษาที่ $i=1,2,\ldots,S$ จากการวัดซ้ำครั้งที่ $t=1,2,\ldots,T$ โดย Y_{it} จะมีรูปแบบการแจกแจงผสม (Mixture distribution) ระหว่างตัวแปรสุ่มที่มีค่าเป็น 0 ด้วยความน่าจะเป็น p_{it} และตัวแปรสุ่มที่มีค่ามากกว่า 0 มีการแจกแจงทวินามเชิงลบที่มีค่าเฉลี่ย λ_{it} ด้วยความน่าจะเป็น $1-p_{it}$ ดังนั้นกำหนดให้ W_{it} เมื่อ $w_{it}=0,1,2,\ldots,$ เป็นตัวแปรสุ่มที่มีการแจกแจงทวินามเชิงลบและมีฟังก์ชันมวลความน่าจะเป็นดังนี้

$$\Pr\left(W_{it}=w_{ij}\right)=f\left(w_{it}\,|\,\lambda_{it},\tau\right)=\frac{\Gamma\left(w_{it}+\frac{1}{\tau}\right)}{w_{it}!\,\Gamma\left(\frac{1}{\tau}\right)}\left(\frac{1}{1+\tau\lambda_{ij}}\right)^{\frac{1}{\tau}}\left(\frac{\tau\lambda_{ij}}{1+\tau\lambda_{ij}}\right)^{w_{it}},$$

เมื่อ $\tau(\tau>0)$ เป็นพารามิเตอร์รูปร่างที่ใช้วัดค่าการกระจาย (over-dispersion parameter) ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่ม W_{it} กำหนดได้โดย $E(W_{it}\,|\,\lambda_{it},\tau)=\lambda_{it}$ และ $V(W_{it}\,|\,\lambda_{it},\tau)=\lambda_{it}+\tau\lambda_{it}^2$ ตามลำดับ ทั้งนี้ถ้า $\tau=0$ จะทำให้เกิดปัญหาค่าความแปรปรวนมีค่ามากกว่าค่าเฉลี่ย λ_{it} ฉะนั้นแล้วตัวแบบทวินามเชิงลบเป็นการกำหนดรูปแบบกำลังสอง $\tau\lambda_{it}^2$ ให้กับความแปรปรวนของตัวแบบปัวซง ทั้งนี้เพื่อแก้ปัญหาค่าความแปรปรวนมีค่ามากกว่าค่าเฉลี่ย (Wan *et al.*, 2012) เมื่อพิจารณากำหนดให้ Y_{it} มีการแจกแจงผสมระหว่างตัวแปรสุ่มที่มีค่าเป็น 0 ด้วยความน่าจะเป็น p_{it} และตัวแปรสุ่ม W_{it} มีการแจกแจงทวินามเชิงลบที่มีค่าเฉลี่ย λ_{it} ด้วยความน่าจะเป็น $1-p_{it}$ แล้วฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม Y_{it} สามารถกำหนดได้ดังนี้

$$\Pr(Y_{it}=y) = \begin{cases} p_{it} + (1-p_{it})\Pr(W_{ij}=0) & \text{ถ้า } y=0 \\ (1-p_{it})\Pr(W_{ij}=y) & \text{ถ้า } y \geq 0 \end{cases}$$

$$= \begin{cases} p_{it} + (1-p_{it})\left(\dfrac{1}{1+\lambda_{it}}\right)^{\frac{1}{\tau}} & \text{ถ้า } y=0 \\ (1-p_{it})\dfrac{\Gamma\left(y+\frac{1}{\tau}\right)}{y!\,\Gamma\left(\frac{1}{\tau}\right)}\left(\dfrac{1}{1+\tau\lambda_{ij}}\right)^{\frac{1}{\tau}}\left(\dfrac{\tau\lambda_{ij}}{1+\tau\lambda_{ij}}\right)^{y} & \text{ถ้า } y \geq 0 \end{cases}$$

โดยตัวแปรสุ่ม Y_{it} จะมีค่าเฉลี่ยและความแปรปรวนดังนี้

$$E(Y_{it}) = (1-p_{it})\lambda_{it}$$

และ

$$V(Y_{it}) = (1-p_{it})\lambda_{it}(1+\tau\lambda_{it}+p_{it}\lambda_{it})$$

เมื่อค่าสังเกตเป็นอิสระต่อกันแล้ว ตัวแบบสำหรับข้อมูลค่าสังเกตที่เป็นจำนวนนับและมีศูนย์มากจะถูกพัฒนาโดยการใช้ตัวแบบทวินามเชิงลบที่มีศูนย์มาก (ZINB) และสามารถกำหนดความสัมพันธ์กับชุดตัวแปรอิสระหรือตัวทำนาย ผ่านฟังก์ชันเชื่อมโยง 2 ตัวให้กับ p_{it} และ λ_{it} ด้วย Log Link function และ Logit function ได้ดังนี้

$$g(\lambda_{it}) = \log(\lambda_{it}) = \mathbf{x}'_{it}\beta \tag{12}$$

โดยที่ $\mathbf{x}'_{it}$ เป็นชุดตัวแปรอิสระหรือตัวทำนายของ λ_{it}

และ
$$g(p_{it}) = \text{logit}(p_{it}) = \mathbf{z}'_{it}\gamma \tag{13}$$

โดยที่ $\mathbf{z}'_{it}$ $\mathbf{x}'_{it}$ เป็นชุดตัวแปรอิสระหรือตัวทำนายของ p_{it}

ทั้งนี้ชุดตัวแปรอิสระหรือตัวทำนายของ λ_{it} และ p_{it} จะเป็นตัวเดียวกันหรือแตกต่างกันก็ได้ ฉะนั้นแล้วค่าเฉลี่ยของ Y_{it} ที่กำหนดโดย $E(Y_{it}) = \mu_{it} = (1-p_{it})\lambda_{it}$ ก็จะขึ้นอยู่กับพารามิเตอร์ β และ γ

สำหรับการประมาณค่า β และ γ ในสมการประมาณค่านัยทั่วไปจะพิจารณาภายใต้ความสัมพันธ์ระหว่างค่าสังเกตของหน่วยศึกษาเดียวกัน $R_i(\rho)$ (Hall and Zhang, 2004) ดังนี้

$$\sum_{i=1}^{S}\begin{pmatrix}\dfrac{\partial \mu_i'}{\partial \beta}\\ \dfrac{\partial \mu_i'}{\partial \gamma}\end{pmatrix}\left[V_i\left(\rho,\beta,\gamma,\tau\right)\right]^{-1}\left[y_i-\mu_i\right]=0 \tag{14}$$

โดยที่ y_i เป็นค่าสังเกตที่ i ของตัวแปรตาม Y_i , $\mu_i'=\left[\mu_{i1},\mu_{i2},\ldots,\mu_{it}\right]$, $V_i\left(\rho,\beta,\gamma,\tau\right)=A_i^{1/2}R_i\left(\rho\right)A_i^{1/2}R_i$ เมื่อ $A_i=Diag\left[V\left(Y_{it}\right)\right],t=1,2,\ldots,T$ อย่างไรก็ตาม Hall and Zhang (2004) ได้ให้ข้อเสนอแนะว่า สมการ (14) ไม่สามารถใช้ในการประมาณค่า β และ γ ได้ เนื่องจากพารามิเตอร์ทั้งสองตัวต้องใช้สารสนเทศร่วมกัน ดังนั้นจึงเสนอการประมาณค่าด้วยการแยกสมการด้วยการสร้างตัวแปรแฝง u_{it} (latent variable) ขึ้นมาเพื่อใช้กำหนดการแจกแจงศูนย์มากหรือการแจกแจงทวินามเชิงลบของ Y_{it}

กำหนดให้ $u_{it}=0$ เมื่อ $Y_{it}\sim NB\left(\lambda_{it},\tau\right)$, $u_{it}=1$ เมื่อ Y_{it} มีการแจกแจงศูนย์มาก และ $\Pr\left(u_{it}=1\right)=p_{it}$ ดังนั้นสมการประมาณค่านัยทั่วไปสำหรับ γ สามารถกำหนดได้ดังนี้

$$\sum_{i=1}^{S}\frac{\partial p_i'}{\partial \gamma}\left[V_{\gamma i}\right]^{-1}\left[u_i-p_i\right]=0 \text{ (XX)}$$

โดยที่ $u_i=\left[u_{i1},\ldots,u_{iT_i}\right]'$, $p_i'=\left[p_{i1},\ldots,p_{iT_i}\right]$ และ $p_{it}\left(\gamma\right)=\dfrac{\exp\left(\mathbf{z}_{it}'\gamma\right)}{1+\exp\left(\mathbf{z}_{it}'\gamma\right)}$ แสดงนัยด้วย

$$\frac{\partial p_{it}}{\partial \gamma}=z_{it}\frac{\exp\left(\mathbf{z}_{it}'\gamma\right)}{\left[1+\exp\left(\mathbf{z}_{it}'\gamma\right)\right]^2}$$

และ $V_{\gamma i}=A_i^{1/2}R_{1i}\left(\rho_1\right)A_i^{1/2}R_i$ เป็นเมทริกซ์ความแปรปรวนของตัวแปรสุ่ม u_i เมื่อ $A_i=Diag\left[p_{i1}\left(1-p_{i1}\right),p_{i2}\left(1-p_{i2}\right),\ldots,p_{iT_i}\left(1-p_{iT_i}\right)\right]$ ของความแปรปรวนของตัวแปรสุ่ม u_{it} ที่ลำดับ t^{th} และ $R_{1i}\left(\rho_1\right)$ เป็นเมทริกซ์สหสัมพันธ์ของ u_i ดังนั้นสามารถประมาณค่า β ในสมการประมาณค่านัยทั่วไป ได้ดังนี้

$$\sum_{i=1}^{S}\frac{\partial \lambda_i'}{\partial \beta}\left[V_{\beta i}\right]^{-1}Diag\left[\left(1-\mu_i\right)\left(y_i-\lambda_i\right)\right]=0 \tag{15}$$

ให้ $\lambda_i'=\left(\lambda_{i1},\lambda_{i2},\ldots,\lambda_{iT_i}\right)$ เมื่อ $\log\left(\lambda_{it}\right)=\mathbf{x}_{it}'\beta$ ฉะนั้น $\lambda_{it}=\exp\left(\mathbf{x}_{it}'\beta\right)$ และ $\dfrac{\partial \lambda_{it}}{\partial \beta}=x_{it}\exp\left(\mathbf{x}_{it}'\beta\right)$ จากเดิม $V_{\beta i}=\left(D_i^{1/2}R_{2i}\left(\rho_2\right)D_i^{1/2}\right)$ เป็นเมทริกซ์ความแปรปรวนของ Y_i เมื่อ $D_i=Diag\left[\lambda_{i1}\left(1+\tau\lambda_{i1}\right),\lambda_{i2}\left(1+\tau\lambda_{i2}\right),\ldots,\lambda_{iT_i}\left(1+\tau\lambda_{iT_i}\right)\right]$ และ $R_{2i}\left(\rho_2\right)$ เป็นเมทริกซ์

สหสัมพันธ์ของ Y_i เมทริกซ์ทแยงมุม $Diag\left[1-u_i\right]=Diag\left[1-u_{i1},\ldots,1-u_{iT_i}\right]$ ในสมการ (15) ที่เป็นตัวชี้วัดของ Y_{it} ที่มาจากการแจกแจงทวินามเชิงลบเท่านั้น (เช่น $u_{it}=0$) ซึ่งเป็นตัวที่ใช้ในการประมาณค่าสมการที่ (14)

เมื่อกำหนดให้ $u_{it}, i=1,2,\ldots,S, t=1,2,\ldots,T_i$ การประมาณค่า β และ γ ไม่สามารถแก้สมการ (14) และ (15) ได้โดยตรง จึงต้องใช้วิธีการคำนวณเชิงตัวเลขแบบวนซ้ำของวิธีคะแนนฟิชเชอร์ (Fisher-Scoring Algorithm) เพื่อประมาณค่าพารามิเตอร์ดังกล่าว อย่างไรก็ตาม u_{it} เป็นเพียงตัวแปรสุ่มแฝง (latent variable, unobserved variables) เท่านั้นการแก้สมการดังกล่าวจะไม่สามารถใช้งานได้โดยตรง ฉะนั้นจึงใช้วิธีการ expectation-solution (Hall and Zhang, 2004; Rosen *et al.*, 200) ในทุก ๆ รอบวนซ้ำ ด้วยการพิจารณาแทนค่าเฉลี่ยแบบมีเงื่อนไขของ u_{it} เมื่อกำหนด Y_{it} ในสมการ (14) และ (15) ในการประมาณค่าพารามิเตอร์ทั้งหมด (λ,β และ τ) โดยใช้คำสั่งในโปรแกรมสำเร็จรูป R

8) ตัวแบบผสมเชิงเส้นนัยทั่วไป (Generalized linear mixed models, GLMMs)

การสร้างตัวแบบเชิงสถิติสำหรับข้อมูลที่ตัวแปรตามเป็นชนิดไม่ต่อเนื่อง หรือตัวแปรตามอาจจะมีการแจกแจงแบบอื่น ๆ เช่น การแจกแจงทวินาม และการแจกแจงปัวซง เป็นต้น ซึ่งแตกต่างไปจากตัวแปรตามที่เป็นข้อมูลชนิดต่อเนื่อง ซึ่งมีการแจกแจงแบบปรกติ แต่เมื่อตัวแปรตามไม่มีการแจกแจงแบบปรกติ เรียกตัวแบบนี้ว่า ตัวแบบผสมเชิงเส้นนัยทั่วไป (GLMMs) หรือเป็นวิธีที่ใช้ในการวิเคราะห์ข้อมูลที่มีการเก็บซ้ำในระยะยาวได้ ซึ่งประกอบด้วยเทอมอิทธิพลคงที่และอิทธิพลสุ่ม เช่นเดียวกัน แต่จะมีการขยายลักษณะเฉพาะ 3 ส่วนเช่นเดียวกับตัวแบบเชิงเส้นนัยทั่วไป (Fitzmaurice *et al.*, 2004) คือ

8.1) การแจกแจงความน่าจะเป็นมีเงื่อนไขของแต่ละ y_{ij} ที่ถูกกำหนดด้วยเทอมอิทธิพลสุ่ม u_i สามารถเขียนให้อยู่ในวงศ์เลขชี้กำลัง และสามารถหาค่าความแปรปรวนมีเงื่อนไขของตัวแปรตาม Y_{ij} ที่ขึ้นอยู่กับ u_i ได้จาก

$$\operatorname{var}\left(Y_{ij}\mid u_i\right)=v\left\{E\left(Y_{ij}\mid u_i\right)\right\}\phi$$

เมื่อ $v(.)$ เป็นฟังก์ชันความแปรปรวนที่ทราบค่า ส่วน $E\left(Y_{ij} \mid u_i\right)$ เป็นฟังก์ชันของค่าเฉลี่ยแบบมีเงื่อนไข นอกจากนี้เมื่อกำหนดเทอมอิทธิพลสุ่ม u_i แล้วค่าสังเกต y_{ij} จะเป็นอิสระต่อกัน ซึ่งเรียกเป็นข้อตกลงของ Conditional Independence Assumption

8.2) ค่าเฉลี่ยของตัวแปรตามมีเงื่อนไขจะขึ้นกับอิทธิพลคงที่ และอิทธิพลสุ่ม u_i ผ่านตัวทำนายเชิงเน้นดังนี้

$$E\left(Y_{ij} \mid u_i\right) = X_{ij}\beta + Z_{ij}u_i$$

$$\eta_{ij} = X_{ij}\beta + Z_{ij}u_i \quad \text{เมื่อ } \eta_{ij} = g\left\{E\left(Y_{ij} \mid u_i\right)\right\}$$

จะได้

$$g\left\{E\left(Y_{ij} \mid u_i\right)\right\} = X_{ij}\beta + Z_{ij}u_i$$

8.3) โดยปรกติจะกำหนดให้เทอมอิทธิพลสุ่ม u_i มีการแจกแจงปรกติหลายตัวแปร (Multivariate normal distribution) คือ $u_i \sim \mathrm{N}(0,G)$ เมื่อ G เป็นเมทริกซ์ความแปรปรวนร่วม และ u_i เป็นอิสระกันกับ X_i โดยสามารถเขียนโครงสร้างความแปรปรวนร่วมของตัวแปรตามได้ดังนี้

$$\begin{aligned} \mathrm{var}\left(Y_i\right) &= \mathrm{var}\left(Z_i u_i\right) + \mathrm{var}\left(\varepsilon_i\right) \\ &= Z_i G Z_i' + R_i \end{aligned}$$

เนื่องจากการวิเคราะห์ความสัมพันธ์ภายในหน่วยศึกษาเดียวกันที่มีการวัดซ้ำจะมีรูปแบบความสัมพันธ์ของโครงสร้างความแปรปรวนร่วมของค่าความคลาดเคลื่อนคือ R_i เช่นเดียวกับตัวแบบผสมเชิงเส้น

ในการศึกษานี้ตัวแปรตาม คือ จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงที่มีความลาดชั้นแบบภูเขา ซึ่งมีการแจกแจงปัวซง สามารถศึกษาวิธี GLMMs ด้วยส่วนประกอบ 3 ส่วน ดังนี้

1) เมื่อตัวแปรตามมีการแจกแจงปัวซง สามารถหาค่า $\mathrm{var}\left(Y_{ij} \mid u_i\right)$ และ $E\left(Y_{ij} \mid u_i\right)$ คือ

$$\operatorname{var}\left(Y_{ij} \mid u_i\right) = E\left(Y_{ij} \mid u_i\right) \text{ โดยที่ค่า } \phi = 1$$

ทั้งนี้ค่าเฉลี่ยและค่าความแปรปรวนของการแจกแจงแบบปัวซงมีค่าเท่ากัน

2) ค่าเฉลี่ยของตัวแปรตามมีเงื่อนไขที่ขึ้นกับ u_i จะมีความสัมพันธ์เชิงเส้นกับเทอมอิทธิพลคงที่และเทอมอิทธิพลสุ่ม ผ่านฟังก์ชันเชื่อมโยงในรูปแบบของล็อก (Log Link) ของตัวทำนายเชิงเส้น ดังนี้

จาก $$E\left(Y_{ij} \mid u_i\right) = X_{ij}\beta + Z_{ij}u_i$$

$$\eta_{ij} = X_{ij}\beta + Z_{ij}u_i \quad \text{เมื่อ } \eta_{ij} = g\left\{E\left(Y_{ij} \mid u_i\right)\right\}$$

ดังนั้น $$\log\left\{E\left(Y_{ij} \mid u_i\right)\right\} = X_{ij}\beta + Z_{ij}u_i$$

3) เมื่อกำหนด $u_i \sim \mathrm{N}(0,G)$ เมื่อ G เป็นเมทริกซ์ความแปรปรวนร่วม และเป็นอิสระกับ X_{ij} โดยเขียนโครงสร้างความแปรปรวนร่วมของตัวแปรตามได้ดังนี้

$$\begin{aligned}\operatorname{var}\left(Y_i\right) &= \operatorname{var}\left(Z_i u_i\right) + \operatorname{var}\left(\varepsilon_i\right) \\ &= \mathrm{Z}_i G Z_i' + R_i\end{aligned}$$

ในการศึกษานี้เป็นข้อมูลการวัดซ้ำจะกำหนดให้โครงสร้างความแปรปรวนร่วมของจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวงที่มีความลาดชันแบบภูเขา มีลักษณะ AR(1) เขียนในรูปเมทริกซ์ ดังนี้

$$R_i = \operatorname{var}\left(\varepsilon_i\right) = \begin{pmatrix} \sigma^2 & \sigma^2\rho & \cdots & \sigma^2\rho^{n_i-1} \\ \sigma^2\rho & \sigma^2 & \cdots & \sigma^2\rho^{n_i-2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma^2\rho^{n_i-1} & \sigma^2\rho^{n_i-2} & \cdots & \sigma^2 \end{pmatrix}$$

โครงสร้างความแปรปรวนร่วมแบบ AR(1) จะกำหนดให้แต่ละค่าสังเกตของหน่วยศึกษาที่ i มีความแปรปรวนเท่ากัน ส่วนความแปรปรวนร่วมระหว่างค่าสังเกตจะไม่เท่ากันสำหรับหน่วยเดียวกัน แต่จะลดลงเข้าใกล้ศูนย์เมื่อช่วงห่างเวลาของการวัดซ้ำข้อมูลมากขึ้น เนื่องจากเป็นผลคูณระหว่างความแปรปรวนกับค่าความสัมพันธ์ของความล่าช้า (lag) ของเวลาในรูปยกกำลัง

นอกจากนี้ยังได้กำหนดโครงสร้างความแปรปรวนร่วมของจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวงที่มีความลาดชันแบบภูเขาเป็นลักษณะ Exchangeable เขียนในรูปเมทริกซ์ ดังนี้

$$R_i = \mathrm{var}(\varepsilon_i) = \begin{pmatrix} \sigma^2+\sigma_1 & \sigma_1 & \cdots & \sigma_1 \\ \sigma_1 & \sigma^2+\sigma_1 & \cdots & \sigma_1 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_1 & \sigma_1 & \cdots & \sigma^2+\sigma_1 \end{pmatrix}$$

ค่าสังเกตจะมีค่าความแปรปรวนร่วมในแต่ละคู่เท่ากันสำหรับหน่วยศึกษาเดียวกัน ซึ่งในการวิเคราะห์ด้วยวิธี GLMMs โครงสร้างความแปรปรวนร่วมของแต่ละหน่วยศึกษาจะมีรูปแบบ Exchangeable แต่เมื่อนำทุกหน่วยศึกษามาแสดงในภาพรวม โครงสร้างความแปรปรวนร่วมจะอยู่ในรูปแบบที่เรียกว่า ส่วนประกอบความแปรปรวน (Variance component: VC) ซึ่งบางครั้งจะเรียกรูปแบบ Exchangeable ว่าเป็นรูปแบบ VC (Littell *et al.*, 2000)

การประมาณค่าพารามิเตอร์ของวิธี GLMMs จะใช้วิธีภาวะน่าจะเป็นสูงสุด โดยอาศัยวิธีการหาอนุพันธ์ของฟังก์ชันภาวะน่าจะเป็น $L(\beta)$ อันดับที่หนึ่งเทียบกับพารามิเตอร์ที่ต้องการประมาณค่าแล้วกำหนดให้เท่ากับศูนย์ ซึ่งในทางปฏิบัติจะใช้ Log-likelihood Function หรือ $\log L(\beta)$ แทน $L(\beta)$ สามารถแสดงได้ดังนี้ (McCulloch *et al.*, 2001)

จาก $$l = \log \int f_{Y|U}(y|u) f_U(u)\,du$$

$$l = \log\left(\prod_{i=1}^{m} \int_{-\infty}^{\infty} \prod_{j=1}^{n_i} \frac{\mu_{ij}^{y_{ij}} e^{-\mu_{ij}}}{y_{ij}!} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}u_i^2}\,du_i \right)$$

$$= y'X\beta - \sum_{ij} \log y_{ij}! + \sum_i \log \int_{-\infty}^{\infty} \exp\left\{ y_{i.}u_i - \sum_j e^{X_{ij}'\beta + u_i} \right\} \times \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}u_i^2}\,du_i$$

เมื่อ y_{ij} คือ ค่าสังเกตที่ i จากการวัดครั้งที่ j โดย $i = 1,2,\ldots,m$ และ $j = 1,2,\ldots,n_i$

จากสมการดังกล่าว ไม่สามารถหาค่าพารามิเตอร์ได้โดยตรงเนื่องจากสมการอยู่ในรูปแบบไม่เป็นเชิงเส้นในเทอมของพารามิเตอร์ จะทำการคำนวณค่าด้วยวิธีภาวน่าจะเป็นสูงสุดแบบวนซ้ำ (Iterative Maximum Likelihood Estimator) โดยเริ่มต้นจากวิธีภาวะน่าจะเป็นสูงสุดหาค่าเริ่มต้นของการประมาณค่าพารามิเตอร์แล้วนำไปใช้ในกระบวนการคำนวณซ้ำต่อไป กระบวนการคำนวณซ้ำจะจบลงเมื่อค่าประมาณของพารามิเตอร์ลู่เข้า (Converge) สู่ค่าประมาณภาวะน่าจะเป็นสูงสุด โดยสามารถใช้โปรแกรมสำเร็จรูปต่าง ๆ มาช่วยในการคำนวณได้ เช่น R ฯลฯ กระบวนการคำนวณแบบวนซ้ำที่นำมาใช้ในการศึกษาครั้งนี้ คือ วิธีนิวตัว-ราฟสัน

ในการศึกษาครั้งนี้ตัวแปรตาม คือ จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวงที่มีความลาดชันแบบภูเขาที่มีการเก็บข้อมูลซ้ำในช่วงกิโลเมตรเดียวกันมีระยะติดตาม 5 ปี จะกำหนดให้ค่าคงที่ (Intercept) เป็นเทอมอิทธิพลสุ่มในวิธี GLMMs คือ ยอมให้แต่ละช่วงกิโลเมตรมีความแตกต่างกัน เรียกว่า ค่าคงที่สุ่ม (Random Intercept) และกำหนดให้โครงสร้างความแปรปรวนร่วมภายในช่วงกิโลเมตรเดียวกันของค่าสังเกตที่มีการเก็บข้อมูลซ้ำในระยะเวลาติดตาม 5 ปี หรือ R_i มีลักษณะตามรูปแบบ AR(1) และ Exchangeable

9) การทบทวนงานวิจัยเกี่ยวกับข้อง

กฤษณ์ (2543, อ้างใน เมษา, 2555) ได้ศึกษาความสัมพันธ์การเกิดอุบัติเหตุของรถบรรทุกขนาดใหญ่กับปัจจัยด้านเรขาคณิตของทางหลวงในประเทศไทย ผลการศึกษาพบว่า การเปลี่ยนแปลงแนวดิ่งเป็นปัจจัยที่มีผลต่อการเปลี่ยนแปลงอัตราการเกิดอุบัติเหตุของรถบรรทุกขนาดใหญ่อย่างมีนัยสำคัญทางสถิติมากที่สุด รองลงมา คือ การเปลี่ยนแปลงแนวทางราบ ความกว้างผิวทาง และจำนวนทางแยก และเมื่อนำกลุ่มปัจจัยที่เป็นตัวสนับสนุนการเกิดอุบัติเหตุมาวิเคราะห์ร่วมกัน พบว่า การเปลี่ยนแปลงแนวดิ่งและความกว้างผิวทาง เป็นกลุ่มปัจจัยที่มีผลต่อการเปลี่ยนแปลงอัตราการเกิดอุบัติเหตุของรถบรรทุกขนาดใหญ่บนถนนทางหลวง 2 ช่องจราจรดังกล่าวอย่างมีนัยสำคัญทางสถิติ ซึ่งชี้ให้เห็นว่าอัตราการเกิดอุบัติเหตุของรถบรรทุกขนาดใหญ่มีแนวโน้มเพิ่มขึ้นตามลักษณะของทางหลวงที่มีค่าเฉลี่ยการเปลี่ยนแปลงแนวทางดิ่งลดลง ในขณะเดียวกันทางหลวงที่มีกว้างผิวทางแคบก็ส่งผลทำให้อัตราการเกิดอุบัติเหตุมีแนวโน้มเพิ่มขึ้นและจะมีแนวโน้มลดลงในกรณีที่ทางหลวงมีความกว้างผิวทางเพิ่มขึ้น

เสริมศักดิ์ (2545, อ้างใน เมษา, 2555) ได้พัฒนาแบบจำลองทางคณิตศาสตร์เพื่อคาดการณ์อุบัติเหตุบนถนนสองช่องจราจรในเขตนอกเมืองกับลักษณะทางเรขาคณิตของถนน จากข้อมูลแบบบันทึกรายงานอุบัติเหตุ (ส.3-02) จำนวน 3 ปี ของกรมทางหลวง โดยแบ่งระดับความรุนแรงเป็น 3 กลุ่มและ 1 รูปแบบการชน คือ จำนวนอุบัติเหตุที่เกิดขึ้นทั้งหมด จำนวนอุบัติเหตุที่เกิดคนบาดเจ็บ จำนวนอุบัติเหตุที่เกิดคนตาย และจำนวนอุบัติเหตุที่เกิดกรณีรถออกนอกถนน ผู้วิจัยได้ทำการทดลองใช้ตัวแบบความสัมพันธ์การถดถอยเชิงเส้นพหุคูณ (Multiple Linear Regression Model) ตัวแบบการถดถอยปัวซง (Poisson Regression Model) ตัวแบบการถดถอยทวินามเชิงลบ (Negative Binomial Regression Model) และตัวแบบการถดถอยแบบล็อกนอมอล (Log-normal Regression Model) ทั้งนี้พบว่าตัวแบบการถดถอยปัวซงเป็นตัวแบบที่มีความเหมาะสมกับข้อมูลมากที่สุด และพบว่า จำนวนทางเชื่อมต่อกิโลเมตรมีอิทธิพลต่อจำนวนการเกิดอุบัติเหตุ จำนวนอุบัติเหตุที่มีคนบาดเจ็บ จำนวนอุบัติเหตุที่เกิดคนตาย ส่วนตัวแบบของจำนวนอุบัติเหตุที่เกิดกรณีรถออกนอกถนน พบว่า แนวทางราบและดิ่ง เป็นปัจจัยที่มีอิทธิพลมากที่สุด

กฤษณ์ (2546, อ้างใน เมษา, 2555) ได้ศึกษาปัญหาความปลอดภัยบริเวณทางโค้งของประเทศไทย ผลการศึกษาพบว่า อุบัติเหตุร้อยละ 65 เกิดขึ้นบริเวณทางโค้ง ส่วนใหญ่เป็นอุบัติเหตุที่เกิดขึ้นกับรถคันเดียวและเกิดในลักษณะไถลออกนอกถนน โดยสันนิษฐานว่าเกิดจากการขับรถเร็วเกินกำหนด และผู้วิจัยได้ทำการตรวจสอบความปลอดภัยบนถนน โดยสำรวจกายภายของถนนพบว่ามี 6 ปัจจัยที่อาจจะเป็นสาเหตุโดยตรงของอุบัติเหตุ ได้แก่ ความเร็วที่ใช้ในการออกแบบไม่สอดคล้องกับความเร็วใช้งานของยานพาหนะ ตำแหน่งของทางเชื่อมอยู่ในระยะมองเห็นปลอดภัยไม่เพียงพอ สภาพผิวทางลื่น ทางเชื่อมบริเวณโค้งมีความลาดชัน ช่วงก่อนถนนเข้าโค้งมีลักษณะเป็นทางตรงมีความลาดลงเป็นระยะยาว และโค้งในแนบรวบที่มีรัศมีโค้งสั้นอยู่บริเวณจุดต่ำสุดของโค้งดิ่งแนวหงาย

เอกนรินทร์ (2547, อ้างใน เมษา, 2555) ได้ศึกษาปัจจัยที่มีผลต่ออุบัติเหตุจราจรบริเวณสามแยกและพัฒนาแบบจำลองทำนายจำนวนอุบัติเหตุบริเวณสามแยก สำหรับทางหลวง 2 ช่องจราจรในเขตนอกเมือง โดยมีปริมาณจราจรและลักษณะทางเรขาคณิตของทางแยกเป็นตัวร่วมในการพัฒนาแบบจำลองจำนวนอุบัติเหตุทั้งหมด จำนวนอุบัติเหตุที่เกิดผู้บาดเจ็บ และจำนวนอุบัติเหตุที่เกิดผู้เสียชีวิต ในการวิจัยครั้งนี้ได้ใช้ตัวแบบเชิงเส้นนัยทั่วไป (Generalized Linear Models) โดยเปรียบเทียบระหว่างตัวแบบถดถอยปัวซงและตัวแบบถดถอยทวินามเชิงลบ ปรากฏว่า ตัวแบบการ

ถดถอยปัวซงเป็นตัวแบบที่เหมาะสมที่สุดเมื่อนำมาใช้อธิบายการเกิดอุบัติเหตุที่ข้อมูลเป็นจำนวนนับและไม่เกิดปัญหาความแปรปรวนมีค่ามากกว่าค่าเฉลี่ย (Over-dispersion) ทั้งนี้พบว่า ปัจจัยที่มีอิทธิพลต่อจำนวนอุบัติเหตุที่เกิดผู้บาดเจ็บ ประกอบด้วย ปริมาณจราจรบนถนนสายหลัก และแนวทางราบ สำหรับปัจจัยที่มีอิทธิพลต่อการเกิดอุบัติเหตุที่เกิดผู้เสียชีวิต ประกอบด้วย ปริมาณจราจร ความเร็วออกแบบ ความกว้างไหล่ทาง ทางเชื่อม และช่องจราจรเลี้ยวซ้ายจากถนนสายหลัก

วนิดา และลี่ลี (2553) ได้ศึกษาตัวแบบทางสถิติที่เหมาะสมสำหรับข้อมูลการเกิดอุบัติเหตุจากการจราจรทางถนน โดยใช้ข้อมูลของกรมทางหลวงชนบทที่มีการเก็บข้อมูลซ้ำตั้งแต่ปี พ.ศ. 2549 – 2551 จำนวน 119 สายทาง ด้วยตัวแบบสมการวางนัยทั่วไป (Generalized Estimating Equations: GEE) เมื่อกำหนดโครงสร้างความสัมพันธ์ของการเกิดอุบัติเหตุบนสายทางเป็นแบบ Exchangeable และ First-order Autoregressive และตัวแบบผสมเชิงเส้นนัยทั่วไป (Generalized Linear Mixed Models: GLMMs) เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของการเกิดอุบัติเหตุสายทางเป็นแบบ Compound Symmetry และ First-order Autoregressive โดยตัวแปรตามเป็นการเกิดอุบัติเหตุบนสายทาง ซึ่งเป็นตัวแปรแบบทวี (Binary) และตัวแปรอิสระประกอบด้วย ความยาวถนน ค่าความเรียบของถนน ความกว้างผิวทาง ไหล่ทาง ผิวทาง และปริมาณจราจรเฉลี่ยต่อวัน ผลการศึกษาพบว่า ตัวแบบ GEE เมื่อกำหนดโครงสร้างความสัมพันธ์แบบ First-order Autoregressive และ Exchangeable มีความเหมาะสมกับข้อมูลเหมือนกัน ส่วนตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ First-order Autoregressive มีความเหมาะสมกับข้อมูลมากกว่า Compound Symmetry ในการเปรียบเทียบค่าประมาณพารามิเตอร์ในตัวแบบภาพรวมของค่าเฉลี่ยประชากร พบว่าตัวแบบ GEE และ GLMMs มีค่าใกล้เคียงกันมาก แต่ตัวแบบ GLMMs สามารถเพิ่มเทอมของอิทธิพลสุ่ม ซึ่งแสดงถึงความแตกต่างของแต่ละสายทางได้ จากตัวแบบ GEE และ GLMMs พบว่าปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุบนสายทาง ที่ระดับนัยสำคัญ 0.05 คือ ค่าความเรียบของถนน ผิวทาง และปริมาณจราจรเฉลี่ยต่อวัน

Shankar *et al.* (1995, อ้างใน เมษา, 2555) ได้ศึกษาปัจจัยด้านเรขาคณิตของถนนและสภาพอากาศที่ส่งผลกระทบต่อความถี่การเกิดอุบัติเหตุในรูปแบบการชนต่าง ๆ โดยทำการศึกษาช่วงถนน Interstate 90 ในสหรัฐอเมริกา ระยะทางรวมทั้งสิ้น 61 กิโลเมตร แบ่งช่วงถนนโดยใช้วิธีแบบจำกัด (fixed-length) ออกเป็น 10 ช่วงเท่า ๆ กัน ช่วงละ 6.1 กิโลเมตร การศึกษานี้ใช้ตัวแบบทวินามเชิงลบ เพื่อสร้างความสัมพันธ์ระหว่างการเกิดอุบัติเหตุในรูปแบบการชนต่าง ๆ กับปัจจัย

ดังกล่าวข้างต้น ผลการศึกษาพบว่า จำนวนโค้งราบที่ออกแบบความเร็วระหว่าง 80.45 ถึง 96.5 กิโลเมตรต่อชั่วโมง มีผลทำให้อุบัติเหตุลักษณะชนด้านข้างและการชนท้ายเพิ่มขึ้น แต่ลักษณะชนรถที่จอดอยู่ (Packed vehicle) ลดลง จำนวนโค้งราบที่ออกแบบความเร็วระหว่าง 96.5 ถึง 112.6 กิโลเมตรต่อชั่วโมง จำนวนโค้งในแต่ละช่วงถนน และความชันสูงสุดมีผลต่อการเกิดอุบัติเหตุทุกรูปแบบการชน รัศมีโค้งราบที่ต่ำสุดในช่วงถนน มีผลทำให้การชนด้านข้างเพิ่มขึ้นแต่ทำให้รถพลิกคว่ำลดลง ผู้วิจัยได้ทำการศึกษาความสัมพันธ์ดังกล่าวด้วยตัวแปรอิสระอื่น ๆ เพิ่มเติม พบว่า ถ้ามีหิมะตกบริเวณที่มีความลาดชัน จะส่งผลทำให้ความถี่การเกิดอุบัติเหตุลดลงทุกรูปแบบ เนื่องจากผู้ขับขี่มีความระมัดระวังมากขึ้น ถ้าหิมะตกบริเวณโค้งจะทำให้การชนด้านข้างและการชนท้ายเพิ่มขึ้น และถ้าฝนตกบริเวณทางโค้งจะทำให้เกิดความถี่การเกิดอุบัติเหตุในลักษณะชนท้าย พลิกคว่ำ และรถออกนอกถนน (Fixed object) เพิ่มขึ้นด้วย

Caliendo *et al*. (2007, อ้างใน เมษา, 2555) พัฒนาแบบจำลองทำนายจำนวนการเกิดอุบัติเหตุและจำนวนผู้บาดเจ็บสาหัส บริเวณทางตรงและทางโค้งบนทางด่วน 4 ช่องจราจรที่มีเกาะกลางในประเทศอิตาลี ระยะทาง 46.6 กิโลเมตร โดยอาศัยตัวแบบ Negative Multinomial Regression Model, Poisson Regression Model และ Negative Binomial Regression Model เพื่อหาแบบจำลองที่เหมาะสมที่สุด ด้วยการประมาณค่าจากวิธีภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Method) พบว่า ตัวแบบบริเวณทางตรง ปัจจัยที่มีอิทธิพลต่อจำนวนการเกิดอุบัติเหตุและจำนวนผู้บาดเจ็บสาหัสในเชิงบวก คือ ความยาวช่วงถนน การมีทางแยก และปริมาณจราจร ส่วนตัวแบบบริเวณทางโค้ง พบว่า จำนวนการเกิดอุบัติเหตุและจำนวนผู้บาดเจ็บสาหัสจะเพิ่มขึ้นเมื่อความยาวโค้งแคบลง จำนวนทางโค้งต่อกิโลเมตรและปริมาณจราจรเพิ่มขึ้น ผลการศึกษาพบว่าตัวแบบ Negative Multinomial Regression เป็นแบบจำลองที่เหมาะสมที่สุด เนื่องจากให้ค่าภาวะน่าจะเป็นสูงสุดดีที่สุดในการพัฒนาแบบจำลองทำนายจำนวนการเกิดอุบัติเหตุและจำนวนผู้บาดเจ็บสาหัสของทั้งบริเวณทางตรงและทางโค้ง

Fu *et al*. (2011, อ้างใน เมษา, 2555) ทำการศึกษาความสัมพันธ์ระหว่างความลาดชันและอัตราการเกิดอุบัติเหตุในพื้นที่ทางลาดลงบริเวณเขตภูมิเขาในประเทศจีน โดยแบ่งความยาวช่วงถนนจากปริมาณจราจรเฉลี่ยต่อวันตลอดปี (AADT) ออกเป็น 6 ช่วง ช่วงละ 13, 8.75, 6.88, 8.80, 14.80 และ 33.20 กิโลเมตร ตามลำดับ ตัวแปรที่ใช้ในการศึกษาความสัมพันธ์ระหว่างความลาดชันและอัตราการเกิดอุบัติเหตุ คือ ความยาวช่วงถนน ความลาดชันเฉลี่ยของแต่ละช่วงถนน ความลาดชันสูงสุดของช่วงถนน จากนั้นนำมาสร้างความสัมพันธ์ระหว่างความสูงจากระดับน้ำทะเลและ

จำนวนอุบัติเหตุเปรียบเทียบกับระยะทางในแต่ละช่วงถนนเพื่อดูการกระจายตัวของการเกิดอุบัติเหตุ พบว่าจำนวนอุบัติเหตุมีความสัมพันธ์กับความสูงจากระดับน้ำทะเล นั่นคือ เมื่อผู้ขับขี่ขับรถไปในระยะทางที่เพิ่มขึ้น จำนวนอุบัติเหตุก็จะเพิ่มขึ้นด้วย นอกจากนี้ผู้วิจัยยังทำการหาค่าความลาดชันเฉลี่ยของแต่ละช่วงถนนทั้งหมด 6 ช่วง และในแต่ละช่วงถนนได้ทำการหาความลาดชันเฉลี่ยในทุก ๆ 1, 2, 3, 4 และ 5 กิโลเมตร ก่อนช่วงถนนที่มีการเกิดอุบัติเหตุด้วย ผลการศึกษาพบว่าความลาดชันเฉลี่ยก่อนช่วงถนนที่ 2 และ 3 กิโลเมตรของทั้ง 6 ช่วงถนนมีค่าสูงสุดและจากการศึกษาความสัมพันธ์ระหว่างอัตราการเกิดอุบัติเหตุและความชันเฉลี่ยดังกล่าวโดยใช้การพยากรณ์แบบเลขชี้กำลัง (Exponential technique) และสหสัมพันธ์ของเพียร์สัน ได้ข้อค้นพบว่า ถ้าความลาดชันเฉลี่ยมีค่ามากจะส่งผลทำให้อัตราการเกิดอุบัติเหตุมากขึ้นด้วย

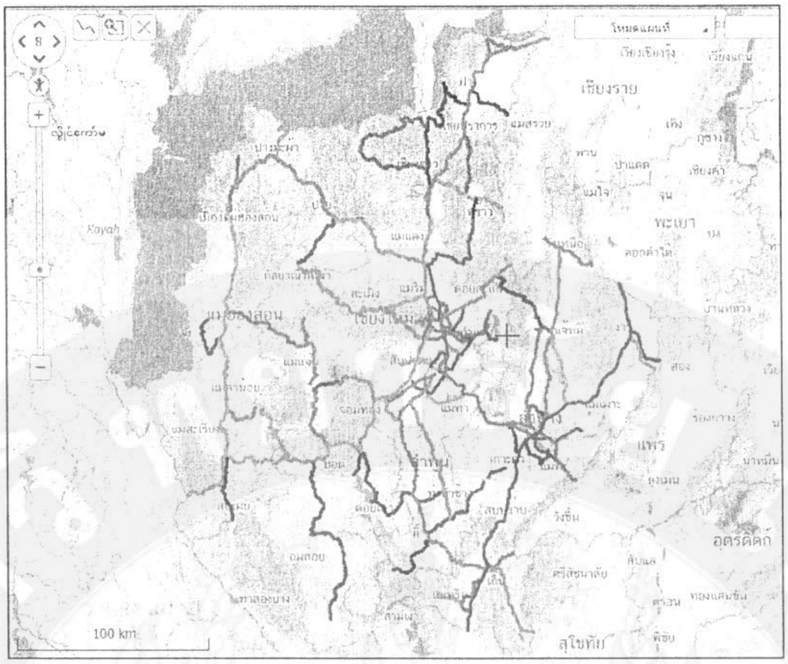
Kihberg and Tharp (1968, อ้างใน เมษา, 2555) ได้ศึกษาความสัมพันธ์ระหว่างอัตราการเกิดอุบัติเหตุกับปัจจัยทางด้านเรขาคณิตของถนนนอกเมือง โดยใช้วิธีการทางสถิติที่เรียกว่า การวิเคราะห์การถดถอย (Regression Analysis) ผลการศึกษาพบว่า เมื่อปริมาณการจราจรเพิ่มขึ้นจะส่งผลทำให้อัตราการเกิดอุบัติเหตุมีค่าเพิ่มขึ้นด้วย

อุปกรณ์และวิธีการวิจัย

1. การศึกษานี้กำหนดขอบเขตการหาตัวแบบเชิงสถิติสำหรับจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

2. การออกแบบการศึกษา (Study design) เป็นการศึกษาติดตามระยะยาวของข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวงในพื้นที่ลาดชัดแบบภูเขาในช่วงกิโลเมตรเดียวกัน เป็นระยะเวลา 5 ปี ตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559

3. ประชากรที่ศึกษา คือ ช่วงกิโลเมตรของเส้นทางบนทางหลวงแผ่นดินที่อยู่ในเขตความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ ครอบคลุมแขวงทางหลวงจังหวัดเชียงใหม่ที่ 1-3 แขวงทางหลวงแม่ฮ่องสอน แขวงทางหลวงลำปางที่ 1 และแขวงทางหลวงลำพูน ที่มีการเกิดอุบัติเหตุบนทางลาดชัน การจราจรมีทิศทางสวนทางกัน และมีภูมิประเทศเป็นแบบภูเขา ดังภาพที่ 1 และตารางที่ 2



ภาพที่ 1 แผนที่ทางหลวงในความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่

ตารางที่ 2 หมายเลขทางหลวงในความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่

ลำดับ	หมายเลขทางหลวง	ชื่อ	เริ่มต้น	สิ้นสุด	ระยะทาง (กิโลเมตร)
1	0001	กรุงเทพมหานคร-แม่สาย	18+100	994+749	970.949
2	0011	อินทร์บุรี – เชียงใหม่	0+763	563+984	545.779
3	0103	ร้อยกวาง – งาว	0+000	230+497	230.497
4	0105	แม่สอด - แม่สะเรียง	0+000	230+497	230.497
5	0106	ดอนไชย - อุโมงค์	0+000	167+204	162.553
6	0107	เชียงใหม่ - แม่จัน	4+172	240+301	236.129
7	0108	เชียงใหม่ - แม่ฮ่องสอน	4+149	353+508	349.359
8	0109	แม่สรวย - ฝาง	0+000	61+133	61.133
9	0114	ดอยติ - ลำพูน	0+000	4+854	4.854
10	0116	ป่าสัก - ท่าวังพร้าว	0+000	26+557	26.557
11	0118	เชียงใหม่ - สันป่าสัก	0+000	158+650	158.65
12	0120	พะเยา - แม่ขะจาน	0+000	60+541	60.541
13	0121	ถนนวงแหวนรอบนอกเมืองเชียงใหม่	0+000	52+957	52.957

ตารางที่ 2 (ต่อ)

ลำดับ	หมายเลขทางหลวง	ชื่อ	เริ่มต้น	สิ้นสุด	ระยะทาง (กิโลเมตร)
40	1103	ลี้ - ฮอด	0+000	67+784	67.784
41	1124	ท่าผา - วังชิ้น	0+000	52+327	52.327
42	1136	เหมืองง่า - ลำพูน	0+000	2+040	2.040
43	1141	ดอนจั่น - เชียงใหม่	0+000	7+565	7.565
44	1147	สันป่าฝ้าย - สันกำแพง	0+000	31+657	31.657
45	1150	ปิงโค้ง - เวียงป่าเป้า	0+000	82+043	82.043
46	1154	สอง - ดอนไชย	0+000	57+769	57.769
47	1156	สบทา - ท่าลี่	0+000	25+321	25.321
48	1157	ท่าล้อ - เมืองปาน	0+000	55+109	55.109
49	1178	แม่ข้อน - บ้านหลวง	0+000	74+349	74.349
50	1184	แม่อาว - ดอนมูล	0+000	73+300	73.300
51	1189	ป่าแดด-บ้านธิ	0+000	10+641	10.641
52	1192	อินทนนท์ - แม่แจ่ม	0+000	20+880	20.880
53	1194	แม่สะเรียง - แม่สามแลบ	0+000	46+652	46.652
54	1219	แม่เทย - ทุ่งหัวช้าง	0+000	16+645	16.645
55	1226	จ่าโบ่ - ปางคาม	0+000	24+793	24.793
56	1229	บ้านใหม่ - เปาสามขา	0+000	11+000	11.000
57	1230	บ้านใหม่ -ห้วยแก้ว	0+000	20+733	20.733
58	1249	แม่งอน - หนองเต่า	0+000	36+765	36.765
59	1250	ไม้แงะ - ท่าโป่งแดง	0+221	2+300	2.079
60	1252	ปางแฟน - ข่วงกอม	0+000	61+701	38.169
61	1260	ศรีบุญเรือง - โรงพยาบาลสันทราย	0+000	5+250	5.250
62	1263	ขุนยวม - แม่นาจร	0+000	66+725	20.525
63	1264	แม่พริก -ห้วยขึ้นก	0+000	20+717	20.717
64	1265	ปาย - วัดจันทร์	0+000	43+619	43.619
65	1266	แม่ลาน้อย - ละอุบ	0+000	37+800	37.800

ตารางที่ 2 (ต่อ)

ลำดับ	หมายเลขทางหลวง	ชื่อ	เริ่มต้น	สิ้นสุด	ระยะทาง (กิโลเมตร)
66	1269	สะเมิง - ต้นเกว๋น	0+000	36+319	36.319
67	1270	กองลอย - แม่แฮใต้	0+000	66+628	66.628
68	1274	ลี้ - นาแก้ว	0+000	65+780	65.780
69	1285	ทุ่งมะส้าน - ห้วยผึ้ง	0+000	15+000	15.000
70	1287	สันมะเกลือ - เมืองปาน	0+000	10+623	10.623
71	1314	ท่าตอน - แม่แหลง	0+000	26+505	26.505
72	1317	ดอนจั่น - ห้วยแก้ว	0+000	36+991	36.991
73	1322	แม่จา - รินหลวง	0+000	128+864	128.864
74	1329	นาป้อใต้ - บ้านเอื้อม	0+000	14+491	14.491
75	1335	ห้วยเดื่อ - แจ้ห่ม	0+000	41+208	41.208
76	1337	หางปอน - ประตูเมือง	0+000	42+650	42.650
77	1340	ทางเข้าแม่ละนา	0+000	1+231	1.231
78	1346	พร้าว - ไชยปราการ	0+000	37+136	37.136
79	1348	ทางเข้าโรงไฟฟ้าแม่เมาะ	0+000	17+213	17.213
80	1349	สะเมิง - วัดจันทร์	0+000	91+633	91.633
81	1352	ทางเข้าดอนไชย	0+000	2+175	2.175
82	1358	ทางเข้าดอยอ่างขาง	0+000	1+358	1.385
83	1359	ทางเข้าเชียงดาว	0+000	9+545	9.545
84	1360	ทางเข้าเวียงฝาง	0+000	8+220	0.505
85	1361	สวนดอกคำ - ม่วงงาม	0+000	5+595	4.720
86	1362	นิคมดอยเต่า - ท่าน้ำ	0+000	3+860	3.860
87	1363	ทางเข้ากองพันสัตว์ต่าง	0+000	0+091	0.971
88	1364	โรงเรียนนวมินทราชูทิศ - กองพันพัฒนาที่3	0+000	0+971	0.971
89	1365	สนามกีฬาสมโภชเชียงใหม่ 700 ปี - ศาลจังหวัดเชียงใหม่	0+000	0+876	0.876

ตารางที่ 2 (ต่อ)

ลำดับ	หมายเลขทางหลวง	ชื่อ	เริ่มต้น	สิ้นสุด	ระยะทาง (กิโลเมตร)
90	1366	หนองฮ่อ - ศูนย์พัฒนาปิโตรเลียมฯ	0+000	1+896	1.896
91	1367	สันทรายน้อย - มหาวิทยาลัยแม่โจ้	0+000	13+123	13.123
92	1391	แม่ทะ - สถานีรถไฟแม่ทะ	0+000	2+713	2.713
93	1393	บ้านจว๊าก - แม่ทะ	0+000	1+513	1.513
94	1394	ทางเข้าตลาดแม่มาลัย	0+000	0+103	0.103
95	1395	ทางเข้าปาย	0+000	4+300	2.282
96	1396	สะเมิง - ดอยซาง	0+000	10+163	10.163
97	1398	สบมาย - เขื่อนกิ่วลม	0+000	0+672	0.672
98	1399	ทางเข้าผาบ่อง	0+000	1+507	1.507
99	1414	ดงป่าลัน - หนองมะจับ	0+000	2+455	2.455

4. ตัวอย่างที่ศึกษา คือ ช่วงกิโลเมตรของเส้นทางบนทางหลวงแผ่นดินที่อยู่ในเขตความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ ที่มีการเกิดอุบัติเหตุบนทางลาดชัน การจราจรมีทิศทางสวนทางกัน มีภูมิประเทศเป็นแบบภูเขา และมีการบันทึกข้อมูลการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ตั้งแต่ปี พ.ศ. 2555 ถึงปี พ.ศ. 2559 จำนวนทั้งสิ้น 105 กิโลเมตร จากถนน 26 สาย ดังตารางที่ 3

ตารางที่ 3 หมายเลขทางหลวงตัวอย่างที่ใช้ในการศึกษา

หมายเลขทางหลวง	ชื่อ	ตอน	ช่วงกิโลเมตรที่ศึกษา
0001	กรุงเทพมหานคร-แม่สาย	1201	728, 733, 742,747, 752, 757, 765
0011	อินทร์บุรี – เชียงใหม่	0800	500, 501, 502, 503, 509, 514, 515, 516
0103	ร้อยกวาง – งาว	0200	42

ตารางที่ 3 (ต่อ)

หมายเลขทางหลวง	ชื่อ	ตอน	ช่วงกิโลเมตรที่ศึกษา
0106	ดอนไชย - อุโมงค์	0201	28
		0202	94, 101, 113
0107	เชียงใหม่ - แม่จัน	0201	42, 62
		0203	106, 114
		0204	202
		0300	119
0108	เชียงใหม่ – แม่ฮ่องสอน	0103	72, 115
		0104	113, 135, 136, 143
		0201	174, 195, 205
		0202	222, 243
		0203	259, 272, 276, 284, 286, 295, 296
0118	เชียงใหม่ - สันป่าสัก	0100	40, 50, 52
		0200	40
0120	พะเยา - แม่ขะจาน	0200	24, 25, 28, 38, 48, 54, 55
1001	เชียงใหม่ - พร้าว	0200	59, 82
1004	ห้วยแก้ว - พระตำหนักภูพิงคราชนิเวศน์	0100	3, 14
1009	จอมทอง - ดอยอินทนนท์	0100	8, 14, 16, 41
1035	ลำปาง - วังเหนือ	0101	13, 24, 37
		0102	37
		0103	86
1039	ดงสันเงิน - ขามแดง	0100	4
1088	ออบหลวง – แม่ซา	0102	13, 15, 30, 38, 47
1095	หนองโค้ง - แม่ฮ่องสอน	0201	82
		0202	107, 116, 119, 127, 129
		0203	159, 180, 203

ตารางที่ 3 (ต่อ)

หมายเลขทางหลวง	ชื่อ	ตอน	ช่วงกิโลเมตรที่ศึกษา
1096	แม่ริม – ปางตะ	0100	6
1099	บ่อหลวง – แม่ตื่น	0100	0, 36, 44, 53, 59
		0101	13
		0102	46
1141	ดอนจั่น - เชียงใหม่	0100	6
1150	ปิงโค้ง - เวียงป่าเป้า	0102	52
1178	แม่ข้อน - บ้านหลวง	0101	16
		0200	54
1184	แม่อาว – ดอนมูล	0100	33
1192	อินทนนท์ - แม่แจ่ม	0100	15
1226	จ่าโบ่ - ปางคาม	0100	1
1249	แม่งอน - หนองเต่า	0100	9, 11, 14, 16
1263	ขุนยวม - แม่นาจร	0100	6, 8, 13
1270	กองลอย - แม่แฮใต้	0100	0
1322	แม่จา - รินหลวง	0100	22, 41
1349	สะเมิง – วัดจันทร์	0100	8

5. ข้อมูลและแหล่งข้อมูล คือ ข้อมูลอุบัติเหตุจากการรวบรวมของสำนักอำนวยความปลอดภัย กรมทางหลวง สำนักทางหลวงที่ 1 จังหวัดเชียงใหม่ ฐานข้อมูลระบบ HAIMS (Highway Accident Information Management System) และระบบสารสนเทศโครงข่ายทางหลวง (Road Net) กรมทางหลวง ระหว่างวันที่ 1 มกราคม พ.ศ. 2555 จนถึงวันที่ 31 ธันวาคม พ.ศ. 2559

5.1 ข้อมูลการเกิดอุบัติเหตุ

งานวิจัยนี้ใช้ข้อมูลทุติยภูมิรายงานการเกิดอุบัติเหตุจากฐานข้อมูลระบบ HAIMS ของสำนักอำนวยความปลอดภัย กรมทรงหลวง สำนักทางหลวงที่ 1 จังหวัดเชียงใหม่ ระหว่างวันที่ 1 มกราคม พ.ศ. 2555 จนถึงวันที่ 31 ธันวาคม พ.ศ. 2559 รวมทั้งสิ้น 5 ปี ถนนที่ศึกษา 26 สาย แบ่งความยาวช่วงถนนที่เท่ากันยาวช่วงละ 1 กิโลเมตร พบว่ามีข้อมูลที่ใช้ในการศึกษาทั้งสิ้น 109

กิโลเมตร (ชุด) โดยแบ่งข้อมูลออกเป็น 2 ส่วน คือ ส่วนแรกจำนวน 80 ชุด นำไปพัฒนาตัวแบบจำลอง (Model calibration) และส่วนที่สองจำนวน 29 ชุด เป็นข้อมูลเพื่อใช้ตรวจสอบความเที่ยงตรงของแบบจำลอง (Validation Model) โดยมีตัวแบบทั้งหมด 3 รูปแบบคือ

1) จำนวนอุบัติเหตุที่เกิดขึ้นทั้งหมด (Accidents) คือ จำนวนการเกิดอุบัติเหตุที่เกิดขึ้นทั้งหมดในช่วงถนนที่ศึกษาโดยไม่คำนึงถึงผู้บาดเจ็บและเสียชีวิต

2) จำนวนผู้บาดเจ็บ (Injury) คือ จำนวนผู้ประสบอุบัติเหตุที่เกิดอาการบาดเจ็บ ณ ที่เกิดเหตุทั้งหมดในช่วงถนนที่ศึกษา

3) จำนวนผู้เสียชีวิต (Fatality) คือ จำนวนผู้ประสบอุบัติเหตุที่เสียชีวิต ณ ที่เกิดเหตุหรือโรงพยาบาลทั้งหมดในช่วงถนนที่ศึกษา

5.2 ข้อมูลปริมาณการจราจร

ข้อมูลส่วนนี้เก็บรวบรวมจากฐานข้อมูลระบบสารสนเทศโครงข่ายทางหลวง (Road Net) กรมทางหลวง ณ วันที่ 31 ธันวาคม พ.ศ. 2559 ซึ่งประกอบไปด้วย

1) ปริมาณการจราจร (พันคันต่อปี) คือ ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี

2) ปริมาณรถหนัก (ร้อยละ) คือ ปริมาณรถโดยสารขนาดกลางและใหญ่ รถบรรทุกขนาด 2 เพลา รถบรรทุกขนาด 3 เพลา รถบรรทุกพ่วง และรถบรรทุกกึ่งพ่วง

5.3 ข้อมูลทางเรขาคณิตของถนน

ข้อมูลส่วนนี้เก็บรวบรวมจากฐานข้อมูลระบบสารสนเทศโครงข่ายทางหลวง (Road Net) กรมทางหลวง ณ วันที่ 31 ธันวาคม พ.ศ. 2559 และแบบแปลนและแบบก่อสร้างของสำนักทางหลวงที่ 1 จังหวัดเชียงใหม่ ซึ่งประกอบไปด้วย

1) จำนวนช่องจราจร คือ จำนวนช่องจราจรในช่วง 1 กิโลเมตรที่ศึกษา ประกอบไปด้วย 5 กลุ่มช่องทางจราจร คือ 2 ช่อง 4 ช่อง 6 ช่อง 8 หรือมากกว่า และอื่น ๆ

2) ประเภทเกาะกลาง คือ ลักษณะเกาะกลางถนนในช่วง 1 กิโลเมตรที่ศึกษา ประกอบไปด้วย 6 ประเภท คือ ไม่มีเกาะกลาง เกาะกลางแบบสี เกาะกลางแบบดินถมยกขึ้น เกาะกลางแบบร่อง มีอุปกรณ์กั้นกลางถนน และไม่ระบุ

3) การจราจร คือ ลักษณะการจราจรในช่วง 1 กิโลเมตรที่ศึกษา ประกอบไปด้วย 5 ลักษณะ คือ รถเดินสวนทาง รถเดินทางเดียว มีช่องเฉพาะรถโดยสาร มีช่องจราจรขึ้นเขา และอื่น ๆ

4) ชนิดผิวจราจร คือ ผิวจราจรในช่วง 1 กิโลเมตรที่ศึกษา ประกอบไปด้วย 3 ชนิด คือ คอนกรีต ลาดยาง และลูกรัง

5.4 ข้อมูลด้านความชันของช่วงถนน

ข้อมูลส่วนนี้เก็บรวบรวมจาก และแบบแปลนและแบบก่อสร้างของสำนักทางหลวงที่ 1 จังหวัดเชียงใหม่ ซึ่งประกอบไปด้วย

1) ความลาดชันเฉลี่ยในช่วงถนน (ร้อยละ) คือ ความลาดชันที่วัดจากจุดกลางของถนน คำนวณได้จาก

$$i_{500} = \frac{(A-C)+(B-C)}{500}$$

เมื่อ i_{500} แทน ความลาดชันเฉลี่ยในช่วงถนนแสดงค่าเป็นร้อยละ

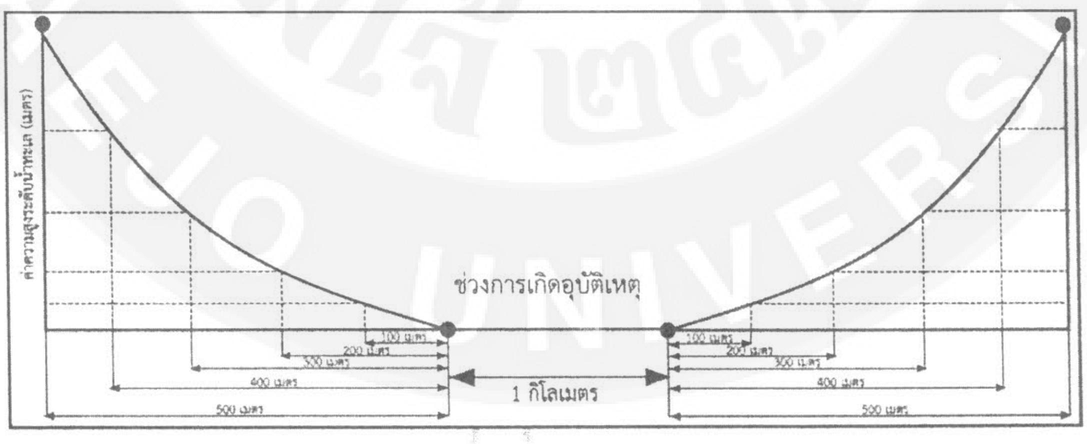
A แทน ค่าความสูงระดับน้ำทะเล ณ จุดเริ่มต้นของช่วงถนน (ซ้าย)

B แทน ค่าความสูงระดับน้ำทะเล ณ จุดสิ้นสุดของช่วงถนน (ขวา)

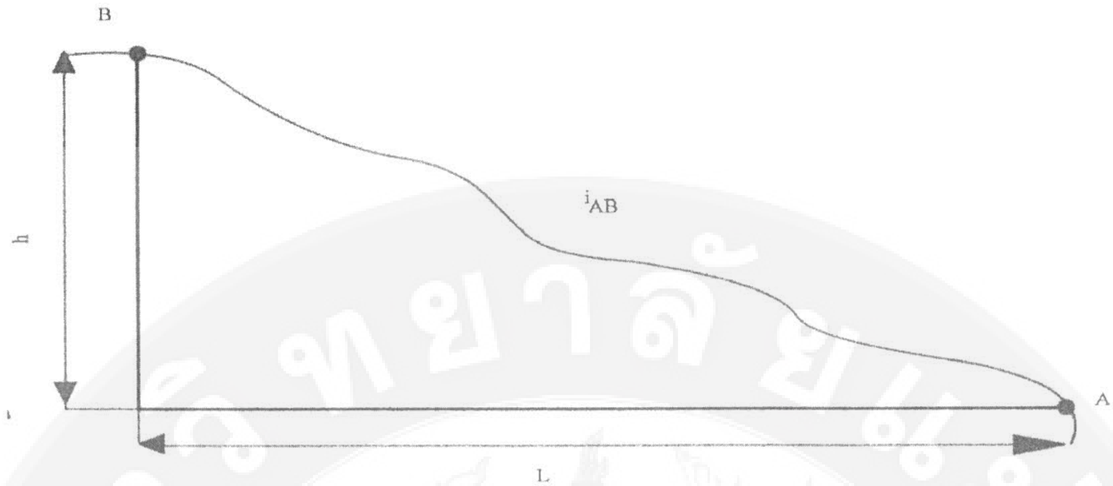
C แทน ค่าความสูงระดับน้ำทะเล ณ จุดกลางช่วงถนน

2) ความลาดชันของช่วงถนนที่อยู่ติดกัน

ผู้วิจัยพิจารณาถึงความลาดชันนอกช่วงถนนที่มีแนวโน้มต่อการเกิดอุบัติเหตุและความรุนแรงของช่วงถนนเท่านั้น นั่นคือ ความลาดชันที่มีลักษณะลาดลงสู่ช่วงถนนที่ศึกษาดังภาพที่ 2 และ 3



ภาพที่ 2 ความลาดชันเฉลี่ยก่อนหน้าและหลังช่วงถนนที่ระยะทาง 100-500 เมตร (เมษา, 2555)



ภาพที่ 3 วิธีหาค่าความลาดชันเฉลี่ยก่อนหน้าและหลังช่วงถนน (เมษา, 2555)

และคำนวณความลาดชันของช่วงถนนที่อยู่ติดกันได้ดังนี้

$$i_{AB} = \frac{h}{L_i}$$

เมื่อ i_{AB} แทน ความลาดชันเฉลี่ยของแต่ละความยาวช่วงถนน (L) แสดงค่าเป็นร้อยละ

h แทน ค่าความแตกต่างของความสูงระดับน้ำทะเลของจุด A และB เมตร

L_i แทน ความยาวของแต่ละช่วงถนน (L_i = 100, 200, 300, 400, 500) (เมตร)

6. การพัฒนาตัวแบบเชิงสถิติสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในพื้นที่ลาดชันแบบภูเขา มีขั้นตอนในการพัฒนาดังนี้

6.1 พัฒนาตัวแบบเชิงสถิติสำหรับจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในพื้นที่ลาดชันแบบภูเขาในระยะเวลาการติดตาม 5 ปี เริ่มจากการวิเคราะห์ตัวแปรเดียว (Univariate analysis) ระหว่างตัวแปรตามจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตกับตัวแปรอิสระทีละตัว (ตารางที่ 4) ด้วยตัวแบบเชิงสถิติ (GEE หรือ GLMMs) เมื่อกำหนดโครงสร้างความสัมพันธ์หรือความแปรปรวนร่วมของข้อมูลเป็นแบบ First-order Autoregressive: AR(1) และ Exchangeable (EXC) ด้วยตัวทดสอบสถิติวัลด์ (Wald) หรือ Z โดยพิจารณาตัวแปรอิสระที่ให้ค่า p-value น้อยกว่า 0.20 สำหรับนำไปวิเคราะห์ในตัวแบบหลายตัวแปร (Multivariate analysis) ต่อไป

ตารางที่ 4 ตัวแปรที่ใช้ในการศึกษาและการกำหนดรหัส

ตัวแปร	**คำอธิบาย**	**การลงรหัส**
ตัวแปรตาม		
Y1	จำนวนการเกิดอุบัติเหตุในแต่ละปี	
Y2	จำนวนผู้บาดเจ็บในแต่ละปี	
Y3	จำนวนผู้เสียชีวิตในแต่ละปี	
ตัวแปอิสระ		
ตัวแปรอิสระที่เป็นตัวแปรเชิงคุณภาพ ได้แก่		
X1	จำนวนช่องจราจร	1= 2 ช่อง
		2= 4 ช่อง
		3= อื่น ๆ
X2	ประเภทเกาะกลางถนน	1= ไม่มีเกาะกลาง
		2= มีเกาะกลาง
X3	การจราจร	1= รถเดินสวนทาง
		2= รถเดินทางเดียว
		3= มีช่องจราจรขึ้นเขา
X4	ชนิดผิวจราจร	1= คอนกรีต
		2= ลาดยาง
ตัวแปรอิสระที่เป็นตัวแปรเชิงปริมาณ ได้แก่		
X5	ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (พันคัน/ปี)	
X6	ร้อยละรถบรรทุกขนาดใหญ่	
X7	ความลาดชันเฉลี่ยในช่วงถนน (%)	
X8	ความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตร (%)	
X9	ความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (%)	
X10	ความลาดชันเฉลี่ยก่อนช่วงถนน 300 เมตร (%)	
X11	ความลาดชันเฉลี่ยก่อนช่วงถนน 400 เมตร (%)	
X12	ความลาดชันเฉลี่ยก่อนช่วงถนน 500 เมตร (%)	
X13	ความลาดชันเฉลี่ยหลังช่วงถนน 100 เมตร (%)	
X14	ความลาดชันเฉลี่ยหลังช่วงถนน 200 เมตร (%)	

ตารางที่ 4 (ต่อ)

ตัวแปร	คำอธิบาย	การลงรหัส
X15	ความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (%)	
X16	ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (%)	
X17	ความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (%)	

6.2 พัฒนาตัวแบบหลายตัวแปรสำหรับจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในพื้นที่ลาดชันแบบภูเขาในระยะเวลาการติดตาม 5 ปี ระหว่างตัวแปรตามจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตกับชุดตัวแปรอิสระที่ผ่านเกณฑ์จากตัวแบบตัวแปรเดียว ด้วยตัวแบบเชิงสถิติ (GEE หรือ GLMMs) เมื่อกำหนดโครงสร้างความสัมพันธ์หรือความแปรปรวนร่วมของข้อมูลเป็นแบบ First-order Autoregressive: AR(1) และ Exchangeable (EXC) ทดสอบตัวแปรอิสระแต่ละตัวว่าสามารถอธิบายการเปลี่ยนแปลงของตัวแปรตามได้หรือไม่ เมื่อกำหนดให้ตัวแปรอิสระอื่น ๆ มีค่าคงที่ และกำหนดระดับนัยสำคัญทางสถิติของการทดสอบ (α) เท่ากับ 0.05

สมมติฐานการทดสอบดังนี้

$H_0 : \beta_j = 0$ (ตัวแปรอิสระ X_j ไม่มีความสัมพันธ์กับตัวแปรตาม Y)

$H_1 : \beta_j = 0$ (ตัวแปรอิสระ X_j ไม่มีความสัมพันธ์กับตัวแปรตาม Y)

สำหรับวิธี GEE ใช้สถิติทดสอบวัลด์ (Wald's test) คือ

$$W=\frac{b_j-\beta_j}{SE_{b_j}}$$ สมมติฐาน $H_0 : \beta_j = 0$ จะถูกปฏิเสธเมื่อ $|W|>Z_{\alpha/2}$

แสดงว่าตัวแปรอิสระ X_j อธิบายการเปลี่ยนแปลงของตัวแปรตาม Y ได้อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

สำหรับวิธี GLMMs ใช้สถิติทดสอบ Z คือ

$$Z=\frac{b_j-\beta_j}{SE_{b_j}}$$ สมมติฐาน $H_0 : \beta_j = 0$ จะถูกปฏิเสธเมื่อ $|Z|>Z_{\alpha/2}$

แสดงว่าตัวแปรอิสระ X_j อธิบายการเปลี่ยนแปลงของตัวแปรตาม Y ได้อย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05

6.3 ตรวจสอบความเหมาะสมของตัวแบบ GEE หรือ GLMMs ที่พัฒนาขึ้น สำหรับตัวแบบ GEE สามารถพิจารณาจากค่าสถิติทดสอบไคกำลังสองส่วนเหลือเพียร์สัน (Pearson residual Chi-square test) ควรมีค่าเข้าใกล้ค่าองศาเสรี (Degree of freedom: DF) หรือค่า Pearson residual Chi-square test / DF ที่มีค่าเข้าใกล้ 1 ส่วนวิธี GLMMs สามารถพิจารณาได้จากค่าสถิติ Generalized Chi-square ที่มีค่ามากสุด และค่า Generalized Chi-square / DF ที่มีค่าเข้าใกล้ 1 และค่า AIC, BIC ที่มีค่าต่ำสุดจะเป็นตัวแบบที่ดีกว่า

ผลการวิจัย

การศึกษาตัวแบบเชิงสถิติสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในพื้นที่ลาดชันแบบภูเขาที่มีการติดตามในระยะ 5 ปี ของช่วงกิโลเมตรเดียวกันหนึ่ง ๆ โดยใช้ข้อมูลจากฐานข้อมูลระบบ HIAMS ของกรมทางหลวง ระหว่างวันที่ 1 มกราคม พ.ศ. 2555 ถึงวันที่ 31 ธันวาคม พ.ศ. 2559 พบว่ามีจำนวนช่วงกิโลเมตรของถนนที่ศึกษาทั้งหมด 105 กิโลเมตร จากถนน 26 สาย ได้ผลการวิเคราะห์ ดังนี้

1) ลักษณะทั่วไปของข้อมูลช่วงกิโลเมตรที่ศึกษาทั้งหมด

ผู้วิจัยได้ทำการวิเคราะห์ข้อมูลที่ใช้ในการศึกษา เพื่อแสดงคุณลักษณะที่สำคัญของข้อมูลดังกล่าว ด้วยค่าสถิติพรรณนา และรูปภาพ ดังต่อไปนี้

ตารางที่ 5 ลักษณะทั่วไปของข้อมูลเชิงคุณภาพ

ตัวแปร	**จำนวน (n = 105)**	**ร้อยละ (100)**
จำนวนช่องจราจร		
2	84	80.00
4	15	14.29
อื่น ๆ	6	5.71
ประเภทเกาะกลาง		
ไม่มีเกาะกลาง	87	82.86
มีเกาะกลาง	18	17.14

ตารางที่ 5 (ต่อ)

ตัวแปร	จำนวน (n = 105)	ร้อยละ (100)
การจราจร		
รถเดินสวนทาง	103	98.10
รถเดินทางเดียว	1	0.95
มีช่องจราจรขึ้นเขา	1	0.95
ชนิดผิวจราจร		
คอนกรีต	10	9.52
ลาดยาง	95	90.48

จากตารางที่ 5 พบว่า ข้อมูลทั้งหมดของช่วงกิโลเมตรของเส้นทางบนทางหลวงแผ่นดินที่อยู่ในเขตความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ ส่วนใหญ่เป็นช่วงกิโลเมตรที่มีจำนวนช่องจราจร 2 ช่องทางถึงร้อยละ 80.00 รองลงมาเป็นจำนวนช่องจราจร 4 ช่องทางร้อยละ 14.29 เป็นช่องทางแบบไม่มีเกาะกลางถึงร้อยละ 82.86 มีการจราจรแบบรถเดินสวนทางถึงร้อยละ 98.10 และส่วนใหญ่มีผิวจราจรแบบลาดยางถึงร้อยละ 90.48

ตารางที่ 6 ลักษณะทั่วไปของข้อมูลเชิงปริมาณ

ตัวแปร		มัธยฐาน	IQR	ค่าเฉลี่ย	SD
จำนวนการเกิดอุบัติเหตุ	ปี 2555	0.00	0.00	0.10	0.30
	ปี 2556	0.00	1.00	0.32	0.54
	ปี 2557	0.00	1.00	0.30	0.60
	ปี 2558	0.00	1.00	0.28	0.51
	ปี 2559	0.00	1.00	0.31	0.52
จำนวนผู้บาดเจ็บ	ปี 2555	0.00	0.00	0.09	0.50
	ปี 2556	0.00	0.00	0.22	0.80
	ปี 2557	0.00	0.00	0.77	5.11
	ปี 2558	0.00	0.00	0.21	0.58
	ปี 2559	0.00	0.00	0.41	1.70

ตารางที่ 6 (ต่อ)

ตัวแปร		มัธยฐาน	IQR	ค่าเฉลี่ย	SD
จำนวนผู้เสียชีวิต	ปี 2555	0.00	0.00	0.02	0.13
	ปี 2556	0.00	0.00	0.11	0.39
	ปี 2557	0.00	0.00	0.11	0.63
	ปี 2558	0.00	0.00	0.03	0.16
	ปี 2559	0.00	0.00	0.06	0.34
ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (พันคัน/ปี)	ปี 2555	3.15	6.00	6.95	9.07
	ปี 2556	3.21	5.85	7.28	9.59
	ปี 2557	3.19	6.64	7.87	10.09
	ปี 2558	3.71	8.98	8.22	9.63
	ปี 2559	3.60	7.76	8.90	11.52
ร้อยละรถบรรทุกขนาดใหญ่	ปี 2555	10.96	11.78	12.86	7.99
	ปี 2556	10.96	11.78	13.16	7.92
	ปี 2557	12.70	12.17	13.49	7.95
	ปี 2558	11.99	14.20	14.53	8.55
	ปี 2559	12.12	15.11	15.25	9.94
ความลาดชันเฉลี่ยในช่วงถนน (%)		3.00	9.20	7.65	12.77
ความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตร (%)		0.00	0.05	0.06	0.14
ความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (%)		0.00	0.06	0.06	0.15
ความลาดชันเฉลี่ยก่อนช่วงถนน 300 เมตร (%)		0.01	0.04	0.05	0.10
ความลาดชันเฉลี่ยก่อนช่วงถนน 400 เมตร (%)		0.01	0.05	0.04	0.09
ความลาดชันเฉลี่ยก่อนช่วงถนน 500 เมตร (%)		1.20	5.60	4.78	7.99
ความลาดชันเฉลี่ยหลังช่วงถนน 100 เมตร (%)		0.00	0.00	0.05	0.18
ความลาดชันเฉลี่ยหลังช่วงถนน 200 เมตร (%)		0.00	0.02	0.03	0.10
ความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (%)		0.00	0.02	0.03	0.09
ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (%)		0.00	0.03	0.03	0.08
ความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (%)		0.40	3.20	3.46	7.33

หมายเหตุ: IQR คือ พิสัยระหว่างควอไทล์ และ SD คือ ส่วนเบี่ยงเบนมาตรฐาน

จากสถิติพื้นฐานในตารางที่ 6 สำหรับข้อมูลทั้งหมดของช่วงกิโลเมตรของเส้นทางบนทางหลวงแผ่นดินที่อยู่ในเขตความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ พบว่า จำนวนการเกิดอุบัติเหตุในช่วงกิโลเมตรที่ศึกษาในปี 2555 และถูกติดตามจนถึงปี 2559 พบว่ามีค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของจำนวนการเกิดอุบัติเหตุเท่ากับ 0.10 ± 0.30, 0.32 ± 0.54, 0.30 ± 0.60, 0.28 ± 0.51 และ 0.31 ± 0.52 ครั้ง ตามลำดับ แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่า จำนวนการเกิดอุบัติเหตุในปี 2555 ถึง 2559 คือ 0 ครั้ง

จำนวนผู้บาดเจ็บในช่วงกิโลเมตรที่ศึกษาในปี 2555 และถูกติดตามจนถึงปี 2559 พบว่ามีค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของจำนวนผู้บาดเจ็บเท่ากับ 0.09 ± 0.50, 0.22 ± 0.80, 0.77 ± 5.11, 0.21 ± 0.58 และ 0.41 ± 1.70 ครั้ง ตามลำดับ แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่า จำนวนผู้บาดเจ็บในปี 2555 ถึง 2559 คือ 0 ครั้ง

จำนวนผู้เสียชีวิตในช่วงกิโลเมตรที่ศึกษาในปี 2555 และถูกติดตามจนถึงปี 2559 พบว่ามีค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของจำนวนจำนวนผู้เสียชีวิตเท่ากับ 0.02 ± 0.13, 0.11 ± 0.39, 0.11 ± 0.63, 0.03 ± 0.16 และ 0.06 ± 0.34 ครั้ง ตามลำดับ แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่า จำนวนผู้เสียชีวิตในปี 2555 ถึง 2559 คือ 0 ครั้ง

ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปีในช่วงกิโลเมตรที่ศึกษาในปี 2555 และถูกติดตามจนถึงปี 2559 พบว่ามีค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปีเท่ากับ 6.95 ± 9.07, 7.28 ± 9.59, 7.87 ± 10.09, 8.22 ± 9.63 และ 8.90 ± 11.52 พันคัน/ปี ตามลำดับ แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี คือ 3.15, 3.21, 3.19, 3.71 และ 3.60 ตามลำดับ

ร้อยละรถบรรทุกขนาดใหญ่ในช่วงกิโลเมตรที่ศึกษาในปี 2555 และถูกติดตามจนถึงปี 2559 พบว่ามีค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของร้อยละรถบรรทุกขนาดใหญ่เท่ากับ 12.86 ± 7.99, 13.16 ± 7.92, 13.49 ± 7.95, 14.53 ± 8.55 และ 15.25 ± 9.94 ตามลำดับ แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าร้อยละรถบรรทุกขนาดใหญ่คือ 10.96, 10.96, 12.70, 11.99 และ 12.12 ตามลำดับ

ความลาดชั้นเฉลี่ยในช่วงถนน (%) ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชั้นเฉลี่ยในช่วงถนน เท่ากับ 7.65 ± 12.77 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชั้นเฉลี่ยในช่วงถนน คือ 3.00

ความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตร (%)เท่ากับ 0.06 ± 0.14 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตรคือ 0

ความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (%)เท่ากับ 0.06 ± 0.15 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตรคือ 0

ความลาดชันเฉลี่ยก่อนช่วงถนน 300 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยก่อนช่วงถนน 300 เมตร (%)เท่ากับ 0.05 ± 0.10 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยก่อนช่วงถนน 300 เมตรคือ 0.01

ความลาดชันเฉลี่ยก่อนช่วงถนน 400 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยก่อนช่วงถนน 400 เมตร (%)เท่ากับ 0.04 ± 0.09 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยก่อนช่วงถนน 400 เมตรคือ 0.01

ความลาดชันเฉลี่ยก่อนช่วงถนน 500 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยก่อนช่วงถนน 500 เมตร (%)เท่ากับ 4.78 ± 7.99 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยก่อนช่วงถนน 500 เมตรคือ 1.20

ความลาดชันเฉลี่ยหลังช่วงถนน 100 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยหลังช่วงถนน 100 เมตร (%)เท่ากับ 0.05 ± 0.18 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยหลังช่วงถนน 100 เมตรคือ 0

ความลาดชันเฉลี่ยหลังช่วงถนน 200 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยหลังช่วงถนน 200 เมตร (%)เท่ากับ 0.03 ± 0.10 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยหลังช่วงถนน 200 เมตรคือ 0

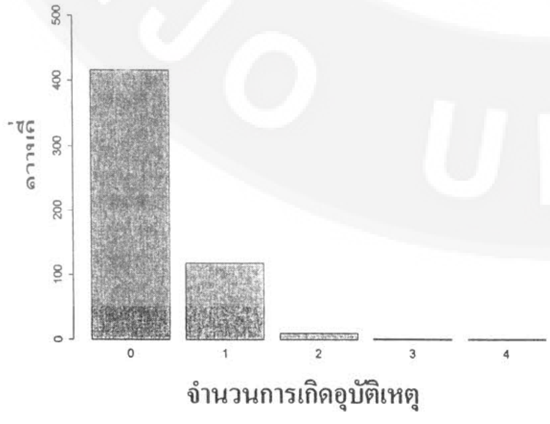
ความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (%)เท่ากับ 0.03 ± 0.09 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตรคือ 0

ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (%)เท่ากับ 0.03 ± 0.08 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตรคือ 0

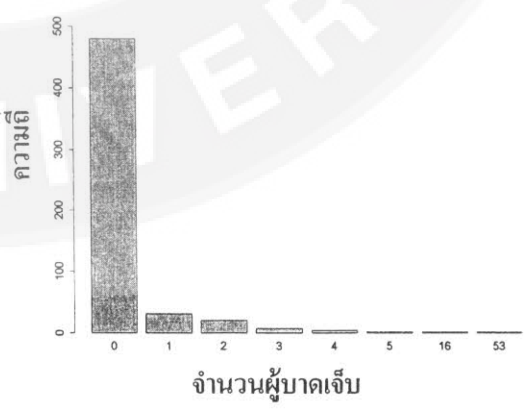
ความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (%)ของช่วงกิโลเมตรที่ศึกษา พบว่ามี ค่าเฉลี่ย ± ส่วนเบี่ยงเบนมาตรฐานของความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (%)เท่ากับ 3.46 ± 7.33 แต่เมื่อพิจารณาจากค่ามัธยฐานพบว่าความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตรคือ 0.40

ตารางที่ 7 จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต บนทางหลวง 10 ในทางลาดชันแบบภูเขาของสำนักงานทางหลวงเชียงใหม่ จำแนกตามปีที่ติดตามศึกษา

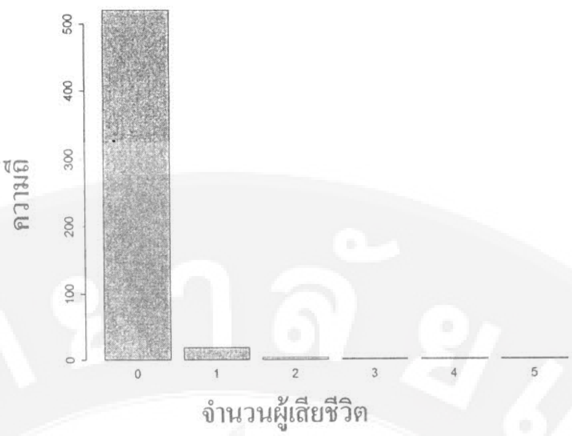
ปี	จำนวนการเกิดอุบัติเหตุ (ครั้ง)	จำนวนผู้บาดเจ็บ (ราย)	จำนวนผู้เสียชีวิต (ราย)
2555	10	9	2
2556	33	24	12
2557	32	82	12
2558	29	23	3
2559	34	45	7
รวม	**138**	**183**	**36**



จำนวนการเกิดอุบัติเหตุ



จำนวนผู้บาดเจ็บ



ภาพที่ 4 ความถี่จำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตทั้งหมด ในช่วงระยะการติดตาม 5 ปี (ระหว่างปี 2555 – 2559)

จากภาพที่ 4 พบว่า ลักษณะการกระจายตัวของจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในระยะ 5 ปีที่ติดตาม มีลักษณะการแจกแจงปัวซงที่มีศูนย์มาก เนื่องจากพบว่า ความถี่การไม่เกิดอุบัติเหตุในช่วงหลักกิโลเมตรที่ศึกษาของสายถนนเกิดขึ้นมากที่สุด รองลงมาเป็นความถี่การเกิดอุบัติเหตุ 1 ครั้ง ดังนั้น ตัวแบบเชิงสถิติที่เหมาะสมในการวิเคราะห์ข้อมูลดังกล่าว คือ ตัวแบบผสมเชิงเส้นนัยทั่วไปสำหรับข้อมูลที่มีการแจกแจงปัวซงที่มีศูนย์มาก โดยกำหนดให้หมายเลขทางหลวงเป็นอิทธิพลสุ่ม

2) ผลการวิเคราะห์ตัวแปรเดียวด้วยตัวแบบเชิงสถิติสำหรับจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

ผลจากการวิเคราะห์ตัวแปรอิสระทีละตัว (Univariate Analysis) ด้วยวิธี GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขาเป็นแบบ AR(1) และ EXC แสดงในตารางที่ 8 - 10

2.1) ตัวแบบเชิงสถิติสำหรับจำนวนการเกิดอุบัติเหตุ

ตารางที่ 8 ผลการวิเคราะห์ตัวแปรเดียวในตัวแบบ GLMMs สำหรับจำนวนการเกิดอุบัติเหตุ

ตัวแปรอิสระ	**AR(1)**				**EXC**			
	Conditional model		**Zero-inflation model**		**Conditional model**		**Zero-inflation model**	
	Est. (S.E.)	**Z value**	**Est. (S.E.)**	**Z value**	**Est. (S.E.)**	**Z value**	**Est. (S.E.)**	**Z value**
ตัวแปรเชิงคุณภาพ								
X1 [2 ช่อง]	Ref.		Ref.		Ref.		Ref.	
X1 [4 ช่อง]	0.97 (0.35)	2.76 [< 0.01]*	-	-	0.55 (0.22)	2.50 [0.01]*	-	-
X1 [อื่น ๆ]	-0.02 (0.39)	-0.05 [0.96]	-	-	-0.03 (0.41)	-0.07 [0.95]	-	-
X2 [ไม่มีเกาะกลาง]	Ref.		Ref.		Ref.		Ref.	
X2 [แบบอื่น ๆ]	0.83 (0.35)	2.35 [0.02]*	-	-	0.50 (0.21)	2.39 [0.02]*	-	-
X4 [ลาดยาง]	Ref.		Ref.		Ref.		Ref.	
X4 [คอนกรีต]	0.97 (0.40)	2.42 [0.02]*	-	-	0.65 (0.25)	2.62 [< 0.01*]	-	-
ตัวแปรเชิงปริมาณ								
X5	0.01 (0.01)	1.82 [0.07]*	-0.42 (3496.98)	0.00 [1.00]	0.02 (0.01)	3.05 [<0.01]*	-	-
X6	0.01 (0.01)	1.17 [0.24]	-0.72 (0.71)	-1.01 [0.30]	0.01 (0.01)	1.46 [0.15]	-	-
X7	-0.003 (0.01)	-0.43 [0.67]	-0.20 (2189.45)	0.00 [1.00]	-0.003 (0.01)	-0.46 [0.65]	-	-

ตารางที่ 8 (ต่อ)

ตัวแปรอิสระ	AR(1)				EXC			
	Conditional model		Zero-inflation model		Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ตัวแปรเชิงปริมาณ								
X8	0.11 (0.59)	0.19 [0.85]	-	-	0.02 (0.01)	3.05 [<0.01]*	-	-
X9	-	-	-	-	0.01 (0.01)	1.46 [0.15]	-	-
X10	-0.56 (0.93)	-0.60 [0.55]	-	-	-0.003 (0.01)	-0.46 [0.65]	-	-
X11	-0.23 (1.01)	-0.23 [0.82]	-	-	0.14 (0.62)	0.22 [0.83]	-	-
X12	-0.002 (0.01)	-0.18 [0.56]	-2.68 (53017.78)	0.00 [1.00]	-0.17 (0.66)	-0.27 [0.79]	-	-
X13	0.94 (0.66)	1.42 [0.16]	-	-	-0.58 (0.98)	-0.59 [0.56]	-	-
X14	-0.12 (0.91)	-0.14 [0.89]	-6.16 (56106.41)	0.00 [1.00]	-0.27 (1.09)	-0.24 [0.81]	-	-
X15	0.58 (0.93)	0.62 [0.53]	-	-	0.58 (0.93)	0.62 [0.53]	-	-
X16	0.95 (0.97)	0.98 [0.33]	-	-	0.95 (0.97)	0.98 [0.33]	-	-
X17	0.01 (0.01)	0.73 [0.47]	-0.29 (4795.23)	0.00 [1.00]	0.01 (0.01)	0.73 [0.47]	-	-

หมายเหตุ: * มีระดับนัยสำคัญทางสถิติที่ระดับ 0.05, ค่าใน [ ] คือ ค่า p-value และ - เกิดปัญหา Model Convergence

จากตารางที่ 8 เมื่อนำตัวแปรอิสระมาตรวจสอบโดยการวิเคราะห์ทีละตัวแปร พบว่า ในตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำนวนการเกิดอุบัติเหตุเป็นแบบ AR(1) พบว่า มีตัวแปรอิสระ 4 ตัว ที่มีความสัมพันธ์และสามารถอธิบายการเปลี่ยนแปลงของจำนวนการเกิดอุบัติเหตุอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ จำนวนช่องจราจร (X1) ประเภทเกาะกลางถนน (X2) ชนิดผิวจราจร (X4) ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) และพบว่าตัวแปรอิสระทั้ง 4 ตัวข้างต้น ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (X16) ผ่านเกณฑ์ p-value < 0.20 ในการนำไปวิเคราะห์ตัวแบบหลายตัวแปร

สำหรับตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำนวนการเกิดอุบัติเหตุเป็นแบบ EXC พบว่า มีตัวแปรอิสระ 5 ตัว ที่มีความสัมพันธ์และสามารถอธิบายการเปลี่ยนแปลงของจำนวนการเกิดอุบัติเหตุอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ จำนวนช่องจราจร (X1) ประเภทเกาะกลางถนน (X2) ชนิดผิวจราจร (X4) ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) และความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตร (X8) และพบว่าตัวแปรอิสระทั้ง 5 ตัวข้างต้น ร้อยละรถบรรทุกขนาดใหญ่ เมตร (X6) และ ความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (X9) ผ่านเกณฑ์ p-value < 0.20 ในการนำไปวิเคราะห์ตัวแบบหลายตัวแปร

2.2) ตัวแบบเชิงสถิติสำหรับจำนวนผู้บาดเจ็บ

ตารางที่ 9 ผลการวิเคราะห์ตัวแปรเดียวในตัวแบบ GLMMs สำหรับจำนวนผู้บาดเจ็บ

ตัวแปรอิสระ	**AR(1)**				**EXC**			
	Conditional model		**Zero-inflation model**		**Conditional model**		**Zero-inflation model**	
	Est. (S.E.)	**Z value**	**Est. (S.E.)**	**Z value**	**Est. (S.E.)**	**Z value**	**Est. (S.E.)**	**Z value**
ตัวแปรเชิงคุณภาพ								
X1 [2 ช่อง]	Ref.		Ref.		Ref.		Ref.	
X1 [4 ช่อง]	-0.11 (0.91)	-0.12 [0.91]	0.71 (0.62)	1.14 [0.25]	0.15 (0.64)	0.23 [0.82]	0.80 (0.55)	1.48 [0.14]
X1 [อื่น ๆ]	0.08 (0.95)	0.08 [0.93]	-0.01 (0.68)	-0.02 [0.98]	-0.58 (1.20)	-0.48 [0.63]	-0.43 (1.19)	-0.36 [0.72]
X2 [ไม่มีเกาะกลาง]	Ref.		Ref.		Ref.		Ref.	
X2 [แบบอื่น ๆ]	-1.04 (0.99)	-1.05 [0.30]	0.18 (0.93)	0.19 [0.85]	-0.19 (0.71)	-0.27 [0.79]	0.70 (0.57)	1.22 [0.22]
X4 [ลาดยาง]	Ref.		Ref.		Ref.		Ref.	
X4 [คอนกรีต]	-1.80 (1.25)	-1.45 [0.15]	-0.92 (1.78)	-0.52 [0.60]	-3.01 (0.57)	-5.32 [<0.01]*	-27.22 (8.97e+05)	0.00 [1.00]
ตัวแปรเชิงปริมาณ								
X5	-0.05 (0.03)	-1.98 [0.04]*	-0.02 (0.03)	-0.73 [0.47]	-0.04 (0.03)	-1.59 [0.11]	-0.02 (0.03)	-0.51 [0.61]
X6	-0.04 (0.04)	-1.02 [0.31]	0.01 (0.03)	0.33 [0.74]	-0.06 (0.05)	-1.25 [0.21]	-0.01 (0.04)	-0.18 [0.86]
X7	-0.004 (0.04)	-0.09 [0.93]	0.03 (0.02)	1.33 [0.18]	0.01 (0.03)	0.34 [0.74]	0.04 (0.02)	1.77 [0.08]*

ตารางที่ 9 (ต่อ)

ตัวแปรอิสระ	AR(1)				EXC			
	Conditional model		Zero-inflation model		Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ตัวแปรเชิงปริมาณ								
X8	-0.48 (1.48)	-0.32 [0.75]	-0.16 (1.26)	-0.12 [0.90]	-0.48 (1.48)	-0.32 [0.75]	-0.16 (1.26)	-0.12 [0.90]
X9	0.53 (0.95)	0.56 [0.58]	-0.63 (0.84)	-0.75 [0.45]	-0.98 (1.50)	-0.66 [0.51]	-1.81 (2.22)	-0.82 [0.41]
X10	0.73 (1.43)	0.51 [0.61]	-0.69 (1.24)	-0.56 [0.58]	-1.14 (2.50)	-0.45 [0.65]	-1.82 (2.75)	-0.66 [0.51]
X11	1.42 (1.53)	0.93 [0.35]	-0.85 (1.50)	-0.57 [0.57]	0.45 (2.87)	0.16 [0.88]	-1.12 (1.87)	-0.60 [0.55]
X12	0.02 (0.02)	0.80 [0.43]	-0.01 (0.02)	-0.43 [0.67]	0.01 (0.02)	0.78 [0.44]	-0.01 (0.02)	-0.41 [0.68]
X13	0.51 (0.57)	0.88 [0.38]	-0.31 (0.66)	-0.47 [0.64]	0.49 (0.61)	0.81 [0.42]	-0.28 (0.66)	-0.42 [0.67]
X14	1.04 (1.18)	0.89 [0.38]	-0.18 (1.36)	-0.14 [0.89]	0.91 (1.27)	0.72 [0.47]	-0.15 (1.36)	-0.11 [0.91]
X15	-	-	-	-	2.09 (1.16)	1.81 [0.07]*	-0.93 (1.41)	-0.66 [0.51]
X16	1.74 (1.83)	0.95 [0.34]	-1.12 (1.65)	-0.68 [0.50]	2.37 (1.61)	1.47 [0.14]	-0.80 (1.67)	-0.48 [0.63]
X17	0.02 (0.02)	0.91 [0.36]	- 0.003(0.02)	-0.17 [0.86]	0.03 (0.02)	1.50 [0.13]	0.001 (0.02)	0.06 [0.95]

หมายเหตุ: * มีระดับนัยสำคัญทางสถิติที่ระดับ 0.05, ค่าใน [ ] คือ ค่า p-value และ - เกิดปัญหา Model Convergence

จากตารางที่ 9 เมื่อนำตัวแปรอิสระมาตรวจสอบโดยการวิเคราะห์ทีละตัวแปร พบว่า ในตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำนวนผู้บาดเจ็บเป็นแบบ AR(1) พบว่า มีตัวแปรอิสระ 1 ตัว ที่มีความสัมพันธ์และสามารถอธิบายการเปลี่ยนแปลงของจำนวนผู้บาดเจ็บอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) และพบว่าตัวแปรอิสระทั้ง 1 ตัวข้างต้น ชนิดผิวจราจร (X4) ผ่านเกณฑ์ p-value < 0.20 ในการนำไปวิเคราะห์ตัวแบบหลายตัวแปร

สำหรับตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำนวนผู้บาดเจ็บเป็นแบบ EXC พบว่า มีตัวแปรอิสระ 2 ตัว ที่มีความสัมพันธ์และสามารถอธิบายการเปลี่ยนแปลงของจำนวนผู้บาดเจ็บอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ ชนิดผิวจราจร (X4) และ ความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (X15) และพบว่าตัวแปรอิสระทั้ง 2 ตัวข้างต้น ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (X16) และ ความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (X17) ผ่านเกณฑ์ p-value < 0.20 ในการนำไปวิเคราะห์ตัวแบบหลายตัวแปร

2.3) ตัวแบบเชิงสถิติสำหรับจำนวนผู้เสียชีวิต

ตารางที่ 10 ผลการวิเคราะห์ตัวแปรเดียวในตัวแบบ GLMMs สำหรับจำนวนผู้เสียชีวิต

ตัวแปรอิสระ	AR(1)				EXC			
	Conditional model		Zero-inflation model		Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ตัวแปรเชิงคุณภาพ								
X1 [2 ช่อง]	Ref.		Ref.		Ref.		Ref.	
X1 [4 ช่อง]	-	-	-	-	-0.85 (1.67)	-0.51 [0.61]	-	-
X1 [อื่น ๆ]	-	-	-	-	-0.62 (2.33)	-0.26 [0.79]	-	-
X2 [ไม่มีเกาะกลาง]	Ref.		Ref.		Ref.		Ref.	
X2 [แบบอื่น ๆ]	-1.54 (2.13)	-0.73 [0.47]	-	-	-1.76 (2.16)	-0.81 [0.42]	-	-
X4 [ลาดยาง]	Ref.		Ref.		Ref.		Ref.	
X4 [คอนกรีต]	-0.90 (2.23)	-0.41 [0.69]	-	-	-1.08 (2.26)	-0.48 [0.63]	-	-
ตัวแปรเชิงปริมาณ								
X5	-0.02 (0.06)	-0.30 [0.77]	-0.39 (4973.59)	0.00 [1.00]	-0.02 (0.05)	-0.35 [0.72]	-	-
X6	0.01 (0.01)	1.46 [0.15]	-	-	-	-	-	-
X7	-0.01 (0.04)	-0.19 [0.85]	-3.95 (64887.46)	0.00 [1.00]	-0.01 (0.04)	-0.20 [0.84]	-	-

ตารางที่ 10 (ต่อ)

ตัวแปรอิสระ	AR(1)				EXC			
	Conditional model		Zero-inflation model		Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ตัวแปรเชิงปริมาณ								
X8	-2.13 (4.93)	-0.43 [0.67]	-	-	-2.29 (4.74)	-0.48 [0.63]	-	-
X9	-0.99 (4.33)	-0.23 [0.82]	-	-	-	-	-	-
X10	-1.50 (5.98)	-0.25 [0.80]	-	-	-	-	-	-
X11	-	-	-	-	-4.83 (8.69)	-0.56 [0.58]	-	-
X12	-	-	-	-	0.01 (0.11)	0.05 [0.96]	-	-
X13	-	-	-	-	0.10 (2.22)	0.04 [0.97]	-	-
X14	0.88 (3.86)	0.23 [0.82]	-	-	-	-	-	-
X15	2.70 (3.54)	0.76 [0.45]	-	-	3.44 (3.09)	1.11 [0.27]	-	-
X16	2.18 (4.70)	0.46 [0.64]	-	-	2.94 (4.23)	0.70 [0.49]	-	-
X17	0.02 (0.05)	0.49 [0.62]	-7.52 (330713.10)	0.00 [1.00]	-	-	-	-

หมายเหตุ: * มีระดับนัยสำคัญทางสถิติที่ระดับ 0.05, ค่าใน [ ] คือ ค่า p-value และ - เกิดปัญหา Model Convergence

จากตารางที่ 10 เมื่อนำตัวแปรอิสระมาตรวจสอบโดยการวิเคราะห์ทีละตัวแปร พบว่า ในตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำนวนผู้เสียชีวิตเป็นแบบ AR(1) พบว่า มีตัวแปรอิสระ 1 ตัว ที่มีความสัมพันธ์และสามารถอธิบายการเปลี่ยนแปลงของจำนวนผู้เสียชีวิตอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) และพบว่าตัวแปรอิสระทั้ง 1 ตัวข้างต้น ชนิดผิวจราจร (X4) ผ่านเกณฑ์ p-value < 0.20 ในการนำไปวิเคราะห์ตัวแบบหลายตัวแปร

สำหรับตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมของจำจำนวนผู้เสียชีวิตเป็นแบบ EXC พบว่า มีตัวแปรอิสระ 2 ตัว ที่มีความสัมพันธ์และสามารถอธิบายการเปลี่ยนแปลงของจำนวนผู้เสียชีวิตอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ ชนิดผิวจราจร (X4) และ ความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (X15) และพบว่าตัวแปรอิสระทั้ง 2 ตัวข้างต้น ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (X16) และ ความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (X17) ผ่านเกณฑ์ p-value < 0.20 ในการนำไปวิเคราะห์ตัวแบบหลายตัวแปร

3) ผลการพัฒนาตัวแบบเชิงสถิติสำหรับจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขา

จากตัวแปรอิสระที่ผ่านการคัดกรองในส่วนที่ 2 จะถูกนำมาพิจารณาเป็นตัวแปรนำเข้าในตัวแบบที่วิเคราะห์หลายตัวแปร

3.1) ตัวแบบเชิงสถิติสำหรับจำนวนการเกิดอุบัติเหตุ

จากผลการวิเคราะห์ตัวแปรเดียวด้วยตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) พบว่าตัวแปรอิสระที่ผ่านเกณฑ์ประกอบไปด้วย

จำนวนช่องจราจร (X1) ประเภทเกาะกลางถนน (X2) ชนิดผิวจราจร (X4) ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) และ ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (X16) นำตัวแปรอิสระทั้งหมดนี้ไปวิเคราะห์ตัวแบบหลายตัวแปรได้ผลการวิเคราะห์ดังตารางที่ 11

ตารางที่ 11 ผลวิเคราะห์ตัวแบบเชิงสถิติหลายตัวแปรสำหรับจำนวนการเกิดอุบัติเหตุ เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1)

ตัวแปรอิสระ	Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ค่าคงที่ (Intercept)	-1.59 (0.23)	-6.86 [<0.01]*	-18.69	-0.004 [1.00]
X1 [2 ช่อง]	Ref.		Ref.	
X1 [4 ช่อง]	0.71 (0.40)	1.78 [0.04]*	19.37 (4496.97)	0.004 [1.00]
X1 [อื่น ๆ]	-0.12 (0.49)	-0.25 [0.80]	-1.28 (60480.25)	0.00 [1.00]
X5	0.004 (0.01)	0.35 [0.72]	-0.15 (0.20)	-0.76 [0.44]
Model fit Criteria				
Generalized Chi-Square	655.8			
Generalized Chi-Square / DF	1.28			
AIC	679.8			
BIC	731.0			

จากตารางที่ 11 พบว่า มีตัวแปรอิสระเพียง 2 ตัว คือ จำนวนช่องจราจร (X1) และปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) เท่านั้น ที่สามารถพัฒนาเป็นตัวแบบเชิงสถิติได้ (ตัวแบบไม่พบปัญหา Convergence) และพบว่าจำนวนช่องจราจร (X1) สามารถอธิบายการเปลี่ยนแปลงของจำนวนการเกิดอุบัติเหตุอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อพิจารณาความเหมาะสมของตัวแบบจากค่าสถิติ Generalized Chi-Square / DF เท่ากับ 1.28 ซึ่งมีค่าใกล้เคียง 1.00 ได้ค่า AIC เท่ากับ 679.8 และ BIC เท่ากับ 731.0

จากผลการวิเคราะห์ตัวแปรเดียวด้วยตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC พบว่าตัวแปรอิสระที่ผ่านเกณฑ์ประกอบไปด้วย

จำนวนช่องจราจร (X1) ประเภทเกาะกลางถนน (X2) ชนิดผิวจราจร (X4) ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) ร้อยละรถบรรทุกขนาดใหญ่ (X6) ความลาดชันเฉลี่ยก่อนช่วงถนน 100 เมตร (X8) และความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (X9) นำตัวแปรอิสระทั้งหมดนี้ไปวิเคราะห์ตัวแบบหลายตัวแปร ได้ผลการวิเคราะห์ดังตารางที่ 12

ตารางที่ 12 ผลวิเคราะห์ตัวแบบเชิงสถิติหลายตัวแปรสำหรับจำนวนการเกิดอุบัติเหตุ เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC

ตัวแปรอิสระ	Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ค่าคงที่ (Intercept)	-1.81 (0.25)	-7.26 [<0.01]*	0.27 (3.02)	0.09 [0.93]
	Ref.		Ref.	
X6	0.01 (0.01)	1.13 [0.26]	-0.96 (1.12)	-0.86 [0.39]
X9	-0.11 (0.67)	-0.17 [0.87]	-8.08 (22.4)	-0.36 [0.72]
Model fit Criteria				
Generalized Chi-Square	654.9			
Generalized Chi-Square / DF	1.28			
AIC	682.9			
BIC	742.6			

จากตารางที่ 12 พบว่า มีตัวแปรอิสระเพียง 2 ตัว ร้อยละรถบรรทุกขนาดใหญ่ (X6) และความลาดชันเฉลี่ยก่อนช่วงถนน 200 เมตร (X9) เท่านั้น ที่สามารถพัฒนาเป็นตัวแบบเชิงสถิติได้ (ตัวแบบไม่พบปัญหา Convergence) แต่พบว่าตัวแปรอิสระทั้ง 2 ตัว ไม่สามารถอธิบายการเปลี่ยนแปลงของจำนวนการเกิดอุบัติเหตุอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 เมื่อพิจารณาความเหมาะสมของตัวแบบจากค่าสถิติ Generalized Chi-Square / DF เท่ากับ 1.28 ซึ่งมีค่าใกล้เคียง 1.00 ได้ค่า AIC เท่ากับ 682.9 และ BIC เท่ากับ 742.6

3.2) ตัวแบบเชิงสถิติสำหรับจำนวนผู้บาดเจ็บ

จากผลการวิเคราะห์ตัวแปรเดียวด้วยตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) พบว่าตัวแปรอิสระที่ผ่านเกณฑ์ประกอบไปด้วย

ชนิดผิวจราจร (X4) และ ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) นำตัวแปรอิสระทั้งหมดนี้ไปวิเคราะห์ตัวแบบหลายตัวแปร พบว่า ผลการวิเคราะห์ตัวแบบเชิงสถิติหลายตัวแปรสำหรับจำนวนผู้บาดเจ็บ เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) ไม่พบตัวแบบที่มีความเหมาะสม เนื่องจากเกิดปัญหา Model Convergence

จากผลการวิเคราะห์ตัวแปรเดียวด้วยตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC พบว่าตัวแปรอิสระที่ผ่านเกณฑ์ประกอบไปด้วย

ชนิดผิวจราจร (X4) ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) ความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร (X15) ความลาดชันเฉลี่ยหลังช่วงถนน 400 เมตร (X16) และ ความลาดชันเฉลี่ยหลังช่วงถนน 500 เมตร (X17)นำตัวแปรอิสระทั้งหมดนี้ไปวิเคราะห์ตัวแบบหลายตัวแปรได้ผลการวิเคราะห์ดังตารางที่ 13

ตารางที่ 13 ผลวิเคราะห์ตัวแบบเชิงสถิติหลายตัวแปรสำหรับจำนวนผู้บาดเจ็บ เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC

ตัวแปรอิสระ	Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value
ค่าคงที่ (Intercept)	0.34 (0.24)	1.41 [0.16]	1.56 (0.22)	7.08 [<0.01]*
X4 [ลาดยาง]	Ref.		Ref.	

ตารางที่ 13 (ต่อ)

ตัวแปรอิสระ	Conditional model		Zero-inflation model	
	Est. (S.E.)	Z value	Est. (S.E.)	Z value
X4 [คอนกรีต]	-2.67 (0.93)	-2.88 [<0.01]*	-2.17 (52270)	0.00 [1.00]
X5	-0.01 (0.03)	-0.46 [0.65]	0.003 (0.02)	0.10 [0.92]
Model fit Criteria				
Generalized Chi-Square	578.0			
Generalized Chi-Square / DF	1.13			
AIC	606.0			
BIC	665.6			

จากตารางที่ 13 พบว่า มีตัวแปรอิสระเพียง 2 ตัว ชนิดผิวจราจร (X4) และปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี (X5) เท่านั้น ที่สามารถพัฒนาเป็นตัวแบบเชิงสถิติได้ (ตัวแบบไม่พบปัญหา Convergence) แต่พบว่าตัวแปรอิสระทั้ง 1 ตัว สามารถอธิบายการเปลี่ยนแปลงของจำนวนผู้บาดเจ็บอย่างมีนัยสำคัญทางสถิติที่ระดับ 0.05 คือ ชนิดผิวจราจร (X4) เมื่อพิจารณาความเหมาะสมของตัวแบบจากค่าสถิติ Generalized Chi-Square / DF เท่ากับ 1.13 ซึ่งมีค่าใกล้เคียง 1.00 ได้ค่า AIC เท่ากับ 606.0 และ BIC เท่ากับ 665.6

3.3) ตัวแบบเชิงสถิติสำหรับจำนวนผู้เสียชีวิต

ผลการวิเคราะห์ตัวแบบเชิงสถิติหลายตัวแปรสำหรับจำนวนผู้เสียชีวิต เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) พบว่า ไม่พบตัวแบบที่มีความเหมาะสม เนื่องจากเกิดปัญหา Model Convergence และเมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC พบว่า ไม่พบตัวแบบที่เหมาะสม เนื่องจากไม่มีตัวแปรอิสระที่เป็นไปตามเกณฑ์การคัดเลือกตัวแปร

4) ผลการวิเคราะห์ตัวแบบสำหรับจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขาด้วยวิธี GLMMs

การเปรียบเทียบผลการวิเคราะห์ตัวแบบสำหรับข้อมูลจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่มีการเก็บข้อมูลซ้ำในระยะเวลาติดตาม 5 ปี ด้วยวิธี GLMMs เมื่อโครงสร้างเป็นแบบ AR(1) และ EXC จากข้อมูลทั้งหมด แสดงผลวิเคราะห์ได้ดังนี้

4.1) ตัวแบบสำหรับจำนวนการเกิดอุบัติเหตุ

จากผลการวิเคราะห์ตัวแบบ GLMMs ในส่วนที่ 3 พบว่าตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) เป็นตัวแบบที่เหมาะสมกว่าตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC เนื่องจากมีค่า AIC และ BIC ที่ต่ำกว่า

4.2) ตัวแบบสำหรับจำนวนผู้บาดเจ็บ

จากผลการวิเคราะห์ตัวแบบ GLMMs ในส่วนที่ 3 พบว่าตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) ไม่สามารถหาตัวแบบที่เหมาะสมได้ เนื่องจากเกิดปัญหา Model Convergence ดังนั้น ตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ EXC จึงเป็นตัวแบบที่เหมาะสมสำหรับจำนวนผู้บาดเจ็บ

4.3) ตัวแบบสำหรับจำนวนผู้เสียชีวิต

จากผลการวิเคราะห์ตัวแบบ GLMMs ในส่วนที่ 3 พบว่าตัวแบบ GLMMs เมื่อกำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) และ EXC เกิดปัญหา Model Convergence ดังนั้นจึงไม่พบตัวแบบที่เหมาะสมสำหรับจำนวนผู้เสียชีวิต

วิจารณ์ผลการวิจัย

จากข้อมูลการติดตามการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่อยู่ในเขตความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ ในระยะ 5 ปี (พ.ศ. 2555 – 2559) ได้ข้อสรุปเบื้องต้นว่า ในช่วง 4 ปีหลังของการติดตาม (พ.ศ. 2556 – 2559) มีการเกิดอุบัติเหตุเป็นสองเท่าของปีแรกที่มีการติดตาม (พ.ศ. 2555) อีกทั้งมีแนวโน้มเพิ่มขึ้น ในปี พ.ศ. 2557 มีจำนวนผู้บาดเจ็บจากการเกิดอุบัติเหตุมากที่สุด (จำนวน 84 ราย) อีกทั้งพบว่า เมื่อจำนวนการเกิดอุบัติเหตุเพิ่มมากขึ้น ส่งผลทำให้เกิดจำนวนผู้บาดเจ็บมากขึ้นตามด้วย สำหรับการเสียชีวิตพบว่ามีจำนวนผู้เสียชีวิตตลอดช่วงระยะเวลาการติดตามอยู่ในช่วง 2 – 12 ราย

ลักษณะการกระจายตัวของจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในระยะ 5 ปีที่ติดตาม มีลักษณะการแจกแจงปัวซงที่มีศูนย์มาก เนื่องจากพบว่า ความถี่การไม่เกิดอุบัติเหตุในช่วงหลักกิโลเมตรที่ศึกษาของสายถนนเกิดขึ้นมากที่สุด รองลงมาเป็นความถี่การเกิดอุบัติเหตุ 1 ครั้ง ดังนั้นตัวแบบเชิงสถิติที่ใช้พัฒนาตัวแบบดังกล่าว คือ ตัวแบบเชิงเส้นผสมนัยทั่วไปสำหรับข้อมูลจำนวนนับที่มีศูนย์มาก

ตัวแปรอิสระหรือตัวแปรทำนายที่ส่งผลต่อจำนวนการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่ผ่านการวิเคราะห์จากตัวแบบตัวแปรเดียว คือ จำนวนช่องจราจร ประเภทเกาะกลางถนน ชนิดผิวจราจร ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี ส่วนตัวแปรอิสระหรือตัวแปรทำนายที่ส่งผลต่อจำนวนการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่ผ่านการวิเคราะห์จากตัวแบบหลายตัวแปรที่เหมาะสม คือ จำนวนช่องจราจร โดยพบว่าจำนวนช่องจราจร 4 ช่อง มีโอกาสเสี่ยงต่อการเกิดอุบัติเหตุมากกว่าจำนวนช่องจราจร 2 ช่อง อาจเป็นเพราะผู้ขับขี่เห็นว่าจำนวนช่องจราจรมากจึงไม่ลดความเร็วและไม่ระมัดระวังการขับขี่

ตัวแปรอิสระหรือตัวแปรทำนายที่ส่งผลต่อจำนวนผู้บาดเจ็บบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่ผ่านการวิเคราะห์จากตัวแบบตัวแปรเดียว คือ ปริมาณการจราจรเฉลี่ยต่อวันตลอดทั้งปี ชนิดผิวจราจร และความลาดชันเฉลี่ยหลังช่วงถนน 300 เมตร ส่วนตัวแปรอิสระหรือตัวแปรทำนายที่ส่งผลต่อจำนวนผู้บาดเจ็บบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่ผ่านการวิเคราะห์จากตัวแบบหลายตัวแปรที่เหมาะสม คือ ชนิดผิวจราจร โดยพบว่าชนิดผิวจราจรแบบคอนกรีต มีโอกาสเสี่ยงต่อ

การเกิดอุบัติเหตุน้อยกว่าผิวจราจรแบบลาดยาง อาจเป็นเพราะถนนหลวงส่วนใหญ่มีผิวจราจรแบบลาดยาง ซึ่งมีจำนวนการเกิดอุบัติเหตุมากกว่าและส่งผลทำให้จำนวนผู้บาดเจ็บบนทางหลวงมากกว่า

สรุปผลงานวิจัย

จากผลการวิจัย พบว่าการเกิดอุบัติเหตุบนทางหลวงในพื้นที่ลาดชันแบบภูเขาที่อยู่ในเขตความรับผิดชอบของสำนักงานทางหลวงที่ 1 จังหวัดเชียงใหม่ มีแนวโน้มเพิ่มขึ้นแบบคงที่ เมื่อจำนวนการเกิดอุบัติเหตุเพิ่มมากขึ้นจะส่งผลทำให้เกิดจำนวนผู้บาดเจ็บมากขึ้นตามด้วย ลักษณะการกระจายตัวของจำนวนการเกิดอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ในระยะ 5 ปีที่ติดตาม มีลักษณะการแจกแจงปัวซงที่มีศูนย์มาก ซึ่งสามารถพัฒนารูปแบบความสัมพันธ์ดังกล่าวกับตัวแปรอิสระหรือปัจจัยอื่น ๆ ด้วยตัวแบบผสมเชิงเส้นนัยทั่วไปสำหรับข้อมูลจำนวนนับที่มีศูนย์มากได้ พบว่า ตัวแบบเชิงสถิติที่เหมาะสมกับข้อมูลจำนวนการเกิดอุบัติเหตุ คือ ตัวแบบ GLMMs ที่กำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ AR(1) และพบปัจจัยที่ส่งผลต่อจำนวนการเกิดอุบัติเหตุที่สำคัญ คือ จำนวนช่องจราจร สำหรับตัวแบบเชิงสถิติที่เหมาะสมกับข้อมูลจำนวนผู้บาดเจ็บ คือ ตัวแบบ GLMMs ที่กำหนดโครงสร้างความแปรปรวนร่วมเป็นแบบ Exchangeable (EXC) และพบว่า ปัจจัยที่ส่งผลต่อจำนวนผู้บาดเจ็บ คือ ชนิดผิวจราจร ส่วนข้อมูลจำนวนผู้เสียชีวิตไม่สามารถหาตัวแบบที่เหมาะสมได้

เอกสารอ้างอิง

ไทยรัฐออนไลน์. 2557. **ชัชชาติ คุมเข้ม 3 มาตรการป้องกันอุบัติ เหตุ หลังทัวร์ตกเขา.** จาก http://www.thairath.co.th/content/392287. [22 พฤศจิกายน 2557].

เมษา ทิพเวช. 2555. **แบบจำลองคาดการณ์อุบัติเหตุสำหรับทางหลวงในเขตภูเขา.** วิทยานิพนธ์วิศวกรรมศาสตรมหาบัณฑิต สาขาวิชาวิศวกรรมขนส่ง. มหาวิทยาลัยเทคโนโลยีสุรนารี: นครราชศรีมา.

วนิดา ลิ่มมั่น และลี่ลี อิงศรีสว่าง. 2553. การศึกษาปัจจัยที่ส่งผลต่อการเกิดอุบัติเหตุจากการจราจรทางถนนโดยใช้ตัวแบบ Generalized Estimating Equations และ Generalized Linear Mixed Models. **ว.วิชาการพระจอมเกล้าพระนครเหนือ**. 20 (2): 311-321.

ศูนย์อำนวยการความปลอดภัยทางถนน กระทรวงมหาดไทย. 2556. **รายงานจำนวนผู้เสียชีวิตจากอุบัติเหตุจราจรทางบกประจำปีงบประมาณ พ.ศ.2556**. จาก http://www.roadsafetythailand.com/main/index.php/data-statistics-th/statnormal. [22 พฤศจิกายน 2557].

สำนักอำนวยความปลอดภัย กรมทางหลวง. 2549. **คู่มือการเฝ้าระวังและแก้ไขปัญหาการเกิดอุบัติเหตุบนทางหลวง เรื่อง วิศวกรรมจราจร (Traffic Engineering).** จาก http://bhs.doh.go.th/sites/default/files/upload/doc/book_traff.pdf. [22 พฤศจิกายน 2557].

ASTVผู้จัดการออนไลน์. 2557. **เสียว! ฟอร์จูนเนอร์แหกโค้งหวิดตกเขา ‘ภูทับเบิก’.** จาก http://www.manager.co.th/Local/ViewNews.aspx?NewsID=9570000135026]. [22 พฤศจิกายน 2557].

Bäumer, M., Hautzinger, H., and Heidemann, D. 2000. Generalized Linear Models for Analyses of Kilometrage and Accident Data. **Proceedings of Second International Conference on Transportation and Traffic Studies**. China, July 31- August 2, 2000: 58-65.

Diggle, P., P. Heagerty, K.Y. Liang and Zeger, S.L. 2002. **Analysis of Longitudinal Data**. 2nd edition. Oxford University Press: Oxford.

Fitzmaurice, G.M., Laird, N.M. and Ware, J.H. 2004. **Applied Longitudinal Analysis.**

John Wiley & Sons Inc.: New Jersey.

Hall, D.B., Zhang, Z. 2004. Marginal models for zero inflated clustered data. **Stat. Modelling.** 4: 161-180.

Hosmer, D.W. and Lemeshow, S. 2000. **Applied Logistic Regression**. John Wiley & Sons Inc.: New York.

Kutner, M.H., Nachtsheim, C.J. and Neter, J. 2004. **Applied Linear Regression Models.** 4th edition. McGraw-Hill Inc.: New York.

Liang, K.Y. and Zeger, S.L. 1986. Longitudinal data analysis using generalized linear models. **Biometrika**. 73, 13-22.

Littell, R.C., Pendergast, J. and Natarajan, R. 2000. Modelling covariance structure in the analysis of repeated measures data. **Sta. Med.** 19: 1793-1819.

McCulloch, C.E. and Searle, S.R. 2001. **Generalized, Linear, and Mixed Models.** Wiley-Interscience: New York.

Myers, R.H., Montgomery, D.C. and Vining, G.G. 2002. **Generalized Linear Models with Applications in Engineering and the Sciences.** John Wiley & Sons Inc.: New York.

Nelder, J.A. and Wedderburn, R.W.M. 1972. Generalized Linear Model. **J. Roy. Statist. Soc. A.** 135: 370-384.

Rosen, O. Jiang, W. and Tanner, M.A. 2000. Mixtures of marginal models. **Biometrika.** 87: 391-404.

Wan, T., Hua. H, Xin, T. 2012. **Applied Categorical and Count Data Analysis.** Chapman and Hall/CRC: New York.

World Health Organization. 2004. **GLOBAL STATUS REPORT ON ROAD SAFETY.** Geneva 27, Switzerland.

World Health Organization. 2009. **GLOBAL STATUS REPORT ON ROAD SAFETY.** Geneva 27, Switzerland.