

บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

การศึกษาในบทนี้มีวัตถุประสงค์เพื่อทำการทบทวนทฤษฎีและสรุปสาระสำคัญจากเอกสารที่เกี่ยวข้องกับปัจจัยอุบัติเหตุและแบบจำลองอุบัติเหตุ ในการทบทวนงานวิจัยที่ผ่านมาพบว่ายังไม่มีรายงานการวิจัยใดในประเทศไทยที่ทำการศึกษาในเรื่องแบบจำลองอุบัติเหตุที่มีกรณีศึกษาเป็นถนนในประเทศไทย ส่วนในต่างประเทศนั้นมีการศึกษาและวิจัยเกี่ยวกับเรื่องเหล่านี้มากมาย ซึ่งจะได้สรุปในลำดับต่อไป

2.1 คำนิยามพื้นฐาน

กองวิศวกรรมจราจร กรมทางหลวง ได้กำหนดนิยามต่างๆ เหล่านี้ไว้ในเอกสารสถิติอุบัติเหตุปี 2541 ซึ่งมีคำดังต่อไปนี้

ก. อุบัติเหตุ (Traffic Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นบนทางหลวงอาจมีคนตาย/บาดเจ็บ หรือเกิดความเสียหายต่อทรัพย์สิน

ข. อุบัติเหตุที่เกิดคนตาย (Fatal Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นแล้วทำให้เกิดคนตาย อาจมีคนบาดเจ็บ หรือทรัพย์สินเสียหายด้วยก็ได้ เช่น อุบัติเหตุเกิดขึ้นแล้วครั้งหนึ่งมีคนตาย 10 คน บาดเจ็บ 12 คน ถือว่าเป็นอุบัติเหตุที่เกิดความตายเพียง 1 ราย

ค. อุบัติเหตุที่เกิดคนบาดเจ็บ (Injury Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นแล้วทำให้คนบาดเจ็บอาจมีทรัพย์สินเสียหายด้วยก็ได้ แบ่งได้ 2 กลุ่ม

ค.1 บาดเจ็บสาหัส (Severe Injury Accident) หมายถึง ผู้บาดเจ็บจากอุบัติเหตุดังกล่าวไม่สามารถรักษาให้หายได้ ภายใน 3 สัปดาห์

ค.2 บาดเจ็บเล็กน้อย (Slightly Injury Accident) หมายถึง ผู้บาดเจ็บจากอุบัติเหตุดังกล่าวสามารถรักษาให้หายได้ภายในเวลาอันสั้น หรือไม่นาน

ง. อุบัติเหตุที่เกิดการเสียหาย (Damage Accident) หมายถึง อุบัติเหตุที่เกิดขึ้นแล้วทำให้ทรัพย์สินเสียหายเท่านั้น

จ. อัตราการเกิดอุบัติเหตุ (Accident Rate, AR) หมายถึง จำนวนรายการเกิดอุบัติเหตุต่อตัวแปรต่างๆที่นิยมมาเปรียบเทียบตามหลักสากล เช่น อัตราการเกิดอุบัติเหตุเทียบกับปริมาณการใช้รถใช้ถนนที่บริเวณทางหลวงนั้นๆ ซึ่งมีหน่วยเป็นล้านคัน-กิโลเมตร และเรียกค่าของอัตราการเกิดอุบัติเหตุนี้ว่าจำนวนอุบัติเหตุต่อล้านคัน-กิโลเมตร หรือ

$$AR = \frac{A_i \times 10^6}{365 \times AADT \times L_i}$$

เมื่อ AR คือ อัตราอุบัติเหตุบนช่วงถนนที่ i (อุบัติเหตุต่อ ล้านคัน-กิโลเมตร)
 A_i คือ จำนวนอุบัติเหตุบนช่วงถนนที่ i ต่อปี
 $AADT$ คือ ปริมาณจราจรเฉลี่ยต่อวันตลอดปี (คันต่อวัน, vpd)
 L_i คือ ความยาวช่วงถนนที่ i (กิโลเมตร, km)

ฉ. ปริมาณจราจรเฉลี่ยต่อวันตลอดปี (Average Annual Daily Traffic; AADT) หมายถึง จำนวนยานพาหนะที่วิ่งผ่านจุดหนึ่งจุดใดหรือทางตอนหนึ่งตอนใดตลอดปีหารด้วยจำนวนวันในปีนั้น

2.2 ปัจจัยทางถนน การจราจรและสภาพแวดล้อมที่มีอิทธิพลต่อการเกิดอุบัติเหตุ

จากงานวิจัยที่ผ่านมาพบว่าการเลือกปัจจัยที่มีอิทธิพลต่ออุบัติเหตุนำมาสัมพันธ์กับจำนวนอุบัติเหตุโดยการสร้างแบบจำลองทางคณิตศาสตร์ ภายในแบบจำลองเหล่านั้นมีการเลือกพิจารณาปัจจัยอยู่หลายตัวและมีความแตกต่างกันไปในแต่ละแบบจำลองขึ้นอยู่กับลักษณะสภาพแวดล้อม สภาพการจราจร สภาพภูมิประเทศ สภาพภูมิอากาศ ในบริเวณที่ทำการศึกษาวิจัยและเก็บข้อมูล นอกจากนี้ยังขึ้นกับความสมบูรณ์ของข้อมูล

ปัจจัยที่มีอิทธิพลต่อการเกิดอุบัติเหตุ ได้แก่

2.2.1 ปริมาณจราจรและปริมาณการใช้รถ (Traffic Volume and Vehicle Exposure)

Kihberg et al. (1968) ทำการศึกษาเรื่องอัตราอุบัติเหตุสัมพันธ์กับการออกแบบองค์ประกอบทางเรขาคณิตของถนนในเขตนอกเมือง ได้อ้างอิงงานวิจัยหลายการศึกษา (Vey, 1937;

Belmont, 1953; Raff, 1953; Schoppert, 1957) ที่มีผลยืนยันว่าอัตราอุบัติเหตุเป็นสัดส่วนโดยตรงกับปริมาณจราจรบนทางหลวง กล่าวคือ เมื่อปริมาณจราจรเพิ่ม อัตราอุบัติเหตุเพิ่ม

Jacobs(1976) ทำการศึกษาเรื่องอัตราอุบัติเหตุบนถนนในเขตนอกเมืองประเทศเคนยาและจาไมก้า จากการศึกษาวิจัยพบว่าที่ระดับนัยสำคัญ 5 % อัตราอุบัติเหตุของทั้ง 2 ประเทศมีความสัมพันธ์กับปริมาณจราจรกล่าวคือ อัตราอุบัติเหตุเพิ่มขึ้นตามปริมาณจราจรที่เพิ่มขึ้น

Wang (1997) อ้างงานวิจัยของ Chapman (1973) อธิบายแนวคิดที่สำคัญในการวิเคราะห์อุบัติเหตุคือ แนวคิดเกี่ยวกับปริมาณการใช้รถ (Concept of exposure) แนวคิดนี้เพื่อเลี่ยงการตีความหมายผิดเกี่ยวกับสถานะของอุบัติเหตุ (Accident situations) เพราะจำนวนอุบัติเหตุที่เกิดขึ้นสูงไม่ได้หมายความว่าความน่าจะเป็น (Probability) สำหรับการเกิดอุบัติเหตุสูง เขาใช้คำจำกัดความของปริมาณการใช้รถเป็นจำนวนโอกาส (Opportunities) สำหรับการเกิดอุบัติเหตุ ที่ต้องเกิดขึ้นแน่นอน ในช่วงเวลาและพื้นที่ศึกษาที่กำหนด

2.2.2 แนวทางราบและแนวทางตั้ง (Horizontal Alignment and Vertical Alignment)

Zegeer et al. (1991) กล่าวว่าโค้งแคบ (sharper curve) มีผลให้เกิดอัตราอุบัติเหตุที่สูงกว่าโค้งที่กว้าง

Mcgee et al. (1995) แนะนำว่าความลาดชันมีอิทธิพลต่ออัตราอุบัติเหตุบนโค้งราบ กล่าวคือ โค้งราบในช่วงลาดลง (downgrade) มีโอกาสเกิดอุบัติเหตุสูงกว่าโค้งราบที่ตั้งอยู่บนที่ราบ

Vijayalakshmi (1997) อ้างงานวิจัยของ Vallete et al. (1981) พบว่าโดยส่วนใหญ่รถบรรทุกจะเกิดอุบัติเหตุบนทางลาดชันมากกว่าทางราบ และอุบัติเหตุที่เกี่ยวข้องกับรถบรรทุกจะเกิดบนถนนในเขตเมืองมากกว่าถนนในเขตนอกเมือง

Vijayalakshmi (1997) ทำการศึกษาเรื่องประสิทธิผลทางด้านความปลอดภัยของการขยายถนนสองช่องจราจรในเขตนอกเมือง กรณีศึกษาในเมือง Madras ประเทศอินเดีย โดยสร้างแบบจำลองทำนายจำนวนอุบัติเหตุ จำนวนอุบัติเหตุที่มีผู้บาดเจ็บ และจำนวนอุบัติเหตุที่มีผู้เสียชีวิต ที่สัมพันธ์กับปัจจัยทางเรขาคณิตที่เกี่ยวข้อง ได้แก่ ปริมาณการใช้รถ ความเร็ว ความกว้างผิวจราจร องศาโค้งราบ เปอร์เซ็นต์ความลาดชัน จากการศึกษาวิจัยพบว่า องศาโค้งราบเป็นตัวทำนายที่ดีที่สุดสำหรับแบบจำลองทั้ง 3 ชนิด

Wang (1997) ใช้วิธี Multiple Linear Regression และ Poisson Regression มาทำการศึกษาเปรียบเทียบ และพัฒนาแบบจำลองทำนายจำนวนอุบัติเหตุ จำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิต ที่สัมพันธ์กับปัจจัยทางเรขาคณิตที่เกี่ยวข้องบนถนน Sun Yat-Sen Freeway ประเทศไต้หวัน จากการศึกษาวิจัยพบว่าเปอร์เซ็นต์ความลาดชันเป็นตัวแปรตามที่มีนัยสำคัญกับทุกแบบ

จำลอง โดยจำนวนอุบัติเหตุ จำนวนผู้บาดเจ็บและจำนวนผู้เสียชีวิตจะเพิ่มขึ้นตามความลาดชันช่วงลาดลง (downgrade slope) และจะลดลงตามความลาดชันช่วงลาดขึ้น (upgrade slope)

2.2.3 ความกว้างของช่องจราจรและไหล่ทาง (Lane width and Shoulder width)

Belmont (1954) ทำการศึกษาผลของความกว้างไหล่ทางต่อการเกิดอุบัติเหตุบนถนนทางตรงสองช่องจราจรในเขตนอกเมืองที่มีปริมาณจราจรน้อย ถนนที่ใช้เป็นตัวอย่างในการศึกษายาว 533 ไมล์ ซึ่งมีอุบัติเหตุเกิดขึ้น 1,333 อุบัติเหตุ ความกว้างไหล่ทางแบ่งออกเป็น 3 กลุ่ม คือ น้อยกว่า 6 ฟุต เท่ากับ 6 ฟุต และมากกว่า 6 ฟุต พบว่าบนช่วงถนนที่มีความกว้างไหล่ทาง 6 ฟุต มีอัตราอุบัติเหตุต่ำกว่าช่วงถนนที่มีความกว้างไหล่ทางน้อยกว่า 6 ฟุต และบนช่วงถนนที่มีความกว้างไหล่ทาง 6 ฟุต ในช่วงปริมาณจราจรมากกว่า 5000 คันต่อวันมีอัตราอุบัติเหตุต่ำกว่าช่วงถนนที่มีความกว้างไหล่ทางเกิน 6 ฟุต ผลที่ได้จากการศึกษาดังกล่าวแตกต่างจากผลในทางทฤษฎี (ไหล่ทางที่กว้างขึ้นจะทำให้มีอัตราอุบัติเหตุลดลง) เขาสนับสนุนตามความจริงที่ว่า ในกรณีที่ไหล่ทางน้อยกว่า 3 ฟุตจะทำให้พื้นที่บนไหล่ทางไม่เพียงพอที่จะมีการแซง (Inadequate shoulder maneuver room) และไหล่ทางแคบทำให้การตัดกันของรถในกระแสดูจราจรมีมากขึ้น และกรณีที่ความกว้างไหล่ทางมากกว่า 6 ฟุตจะทำให้ผู้ขับขี่บนไหล่ทางเกิดความประมาทและมีโอกาสที่จะชนรถที่จอดอยู่ได้

Haefner (1970) อ้างงานวิจัยของ Stohner (1956) และ Billion (1957) กล่าวว่ามีการศึกษาหลายงานวิจัยในรัฐนิวยอร์กให้การปกป้องว่าไหล่ทางที่กว้างกว่าปลอดภัยกว่า

Zegeer and Deacon (1988) ทำการศึกษาผลทางด้านความปลอดภัยของทางหลวงในการออกแบบความกว้างช่องจราจร ความกว้างไหล่ทาง และประเภทไหล่ทาง พบว่าเมื่อเพิ่มความกว้างช่องจราจรและความกว้างไหล่ทางจะทำให้อัตราอุบัติเหตุลดลง การขยายความกว้างช่องจราจรมีผลทางด้านความปลอดภัยมากกว่าการเพิ่มความกว้างไหล่ทาง

2.2.4 ระยะมองเห็น (Sight Distance)

Hedman (1990) อ้างถึง Guide to Road Safety Engineering ประเทศสวีเดนว่า อัตราอุบัติเหตุจะลดลงเมื่อเพิ่มความยาวเฉลี่ยของระยะมองเห็น

Fink and Krammes (1995) ศึกษาผลของความยาว Tangent และระยะมองเห็นต่ออัตราอุบัติเหตุบริเวณโค้งราบบนถนนสองช่องจราจรในเขตนอกเมือง โดยสร้างแบบจำลองถดถอยเชิงเส้นอย่างง่าย สัมพันธ์จำนวนอุบัติเหตุที่เกิดขึ้นกับองศาโค้งราบที่จำแนกระยะมองเห็นสั้น ปานกลาง และยาว ทำการเก็บรวบรวมข้อมูลอุบัติเหตุบนถนนสองช่องจราจรในเขตนอกเมืองของรัฐ

ต่างๆ 3 รัฐในประเทศสหรัฐอเมริกา ระยะเวลา 5 ปี นำข้อมูลมาจำแนกโค้งที่มีคุณสมบัติใกล้เคียงกันได้ 563 โค้ง แล้วนำมาเรียงลำดับระยะมองเห็น โดยข้อมูล 25 % มีระยะมองเห็นต่ำกว่า 61 เมตร จะจำแนกว่ามีระยะมองเห็นสั้น ข้อมูล 50 % มีระยะมองเห็นในช่วงระหว่าง 61-213.5 เมตร จะจำแนกว่ามีระยะมองเห็นปานกลาง และข้อมูลอีก 25 % ระยะมองเห็นมากกว่า 213.5 เมตร จะจำแนกว่ามีระยะมองเห็นยาว จากผลการศึกษาวิจัยพบว่าองศาโค้งราบเป็นตัวทำนายที่ดีในการหาอัตราการเกิดอุบัติเหตุบริเวณโค้งราบ ถึงแม้ว่าผลการจำแนกระยะมองเห็นเพื่อแสดงความสัมพันธ์ระหว่างองศาโค้งกับจำนวนครั้งของอุบัติเหตุจะไม่สามารถนำแบบจำลองไปใช้ในการทำนายจำนวนครั้งของอุบัติเหตุได้ แต่ผลการจำแนกนั้นแนะนำว่าทางตรงที่ยาวก่อนเข้าโค้ง (long approach tangent Length) มีผลต่อความปลอดภัย กล่าวคือในทางตรงยาวๆ รถจะวิ่งด้วยความเร็วสูงเมื่อมาถึงโค้งที่มีรัศมีน้อย (sharp curve) คนขับจะปรับตัวไม่ทัน จึงทำให้เกิดอุบัติเหตุง่าย

Vijayalakshmi (1997) อ้างงานวิจัยของ Olson (1984) กล่าวว่าระยะมองเห็นถูกจำกัดเนื่องจากมีอุปสรรค (obstruction) บริเวณด้านใน (inside curve) โค้งราบ บริเวณทางแยก หรือ บริเวณในหุบเขา ถึงแม้บางครั้งจะสามารถกำจัดอุปสรรคในบริเวณโค้งราบและบริเวณทางแยกโดยไม่มี การเปลี่ยนแปลงลักษณะเรขาคณิตของถนน เช่น การตัดต้นไม้ แต่การกำจัดอุปสรรคในหุบเขาสามารถทำได้โดยการเปลี่ยนแนวทางโค้งเท่านั้น ดังนั้น Olson จึงรวบรวมข้อมูลอุบัติเหตุในบริเวณที่มีและไม่มี การจำกัดการมองเห็นเนื่องจากโค้งโค้ง เขาพบว่าบริเวณที่มีการจำกัดระยะมองเห็นจำกัดมีจำนวนอุบัติเหตุเป็น 52 % ของทั้งหมด ซึ่งมากกว่าบริเวณที่ระยะมองเห็นไม่ถูกจำกัด

2.2.5 ความเร็ว (Speed)

Kihberg (1967) อ้างอิงงานของ Bitzl (1954) ได้รายงานว่าการใช้ความเร็วที่สูงเกินกำหนดก่อให้เกิดอุบัติเหตุถึง 24 % บนทางด่วนในประเทศเยอรมัน

Mcknight and Klein (1990) จากการศึกษาเปรียบเทียบอุบัติเหตุที่เกิดบนถนนสายนอกเมืองระหว่างรัฐเนื่องจากการเพิ่มพิกัดความเร็วจาก 55 ไมล์ต่อชั่วโมง เป็น 65 ไมล์ต่อชั่วโมง พบว่าพิกัดความเร็วที่เพิ่มขึ้นเป็น 65 ไมล์ต่อชั่วโมงส่งผลให้มีจำนวนผู้ขับขี่ขับรถเร็วเพิ่มขึ้น 48 % ซึ่งมีผลให้จำนวนอุบัติเหตุที่เกิดบนถนนสายเพิ่มขึ้น

2.2.6 จำนวนทางแยก และจำนวนจุดต่อเชื่อม (Number of Intersections and Junctions)

Jacobs (1976) จากการศึกษาเรื่องอัตราอุบัติเหตุบนถนนในเขตนอกเมืองประเทศเคนยาและจาไมก้า และใช้วิธี Multiple Linear Regression อธิบายความสัมพันธ์ระหว่างอัตราอุบัติเหตุกับลักษณะทางเรขาคณิตของทาง พบว่าในประเทศจาไมก้า ที่ระดับนัยสำคัญ 5 % ความกว้างถนนและจำนวนจุดต่อเชื่อมต่อกิโลเมตรเป็นตัวแปรที่มีความสัมพันธ์ต่ออัตราอุบัติเหตุมากที่สุด

Saito et al. (1998) ทำการศึกษาความสัมพันธ์ระหว่างอัตราอุบัติเหตุกับลักษณะของถนนและสภาพจราจรบนถนนสองช่องจราจรนอกเมือง ในเกาะฮอกไกโด สรุปว่าในเรื่องความสัมพันธ์ระหว่างอัตราอุบัติเหตุกับความหนาแน่นของทางแยกโดยแบ่งความหนาแน่นออกเป็น 3 กลุ่ม ได้แก่ ความหนาแน่นทางแยกน้อยกว่า 1 อยู่ในช่วง 1-2 และมากกว่า 2 ทางแยกต่อกิโลเมตร พบว่าอัตราอุบัติเหตุมีค่าต่ำสุดเมื่อความหนาแน่นของทางแยกอยู่ในช่วง 1-2 ทางแยกต่อกิโลเมตร

2.3 แบบจำลองอุบัติเหตุ - ทฤษฎี

สมมติว่าแบ่งความยาวถนนสองช่องจราจรในเขตนอกเมือง ออกเป็น n ช่วง พิจารณาในช่วงเวลา 1 ปี หมายถึง ถ้าช่วงถนนเป็นช่วงเดียวกัน (หลักกิโลเมตรเดียวกัน) แต่เวลาเป็นคนละปี จะถือว่า ช่วงถนนช่วงนั้นเป็นคนละช่วงกัน กำหนดให้

i คือ ช่วงถนน (section) ซึ่งมีค่าเป็นจำนวนเต็มบวก

Y_i คือ ตัวแปรตามซึ่งแทนด้วยจำนวนอุบัติเหตุ จำนวนอุบัติเหตุที่เกิดบาดเจ็บ หรือจำนวนอุบัติเหตุที่เกิดคนตายบนช่วงถนนช่วงที่ i ในช่วง 1 ปี

y_i คือ ค่าของตัวแปรตาม Y_i บนถนนช่วงที่ i ในช่วงเวลา 1 ปี ซึ่งมีค่าเป็น 0 หรือจำนวนเต็มบวกอื่นๆ

x_i คือ ตัวแปรอิสระซึ่งแทนด้วยปัจจัยทางเรขาคณิตและปริมาณจราจรของช่วงถนนที่ i ที่มีผลต่อการเกิดอุบัติเหตุบนถนนสองช่องจราจร ในเขตนอกเมือง

x_{ij} คือ ตัวแปรอิสระตัวที่ j ในกลุ่มตัวแปรอิสระ x_i

j คือ ลำดับที่ของตัวแปรอิสระแต่ละตัวในแบบจำลอง

β_j คือ สัมประสิทธิ์การถดถอยตัวที่ j

V_i คือ ปริมาณการใช้รถ (Vehicle Exposure) ในช่วงเวลา 1 ปีบนช่วงถนน i (veh -km)

$$V_i = 365 \times AADT_i \times L_i \quad (2.1)$$

$AADT_i$ คือ ปริมาณจราจรรายวันเฉลี่ยตลอดปี บนช่วงถนน i (คันต่อวัน, vpd)

L_i คือ ความยาวของถนน ช่วงที่ i (กิโลเมตร)

λ_i คือ จำนวนอุบัติเหตุที่คาดว่าจะเกิดต่อคัน-กิโลเมตร เรียกว่า Rate function

$$\lambda_i = \frac{E(Y_i)}{V_i} \quad (2.2)$$

$E(Y_i)$ คือ ค่าคาดหวังของตัวแปรตาม Y_i (Expected value of Y_i)

α คือ Overdispersion Effect แสดงรายละเอียดในหัวข้อ 2.3.4

Γ คือ ค่าของฟังก์ชันแกมมา แสดงรายละเอียดในหัวข้อ 2.3.4

2.3.1 แบบจำลองการถดถอยพหุคูณเชิงเส้น (Multiple Linear Regression Model, MLR Model)

การใช้รูปแบบ MLR Model จะสมมติว่า ตัวแปรตาม Y_i มีการแจกแจงแบบปกติ มีค่าเฉลี่ย (Mean) เท่ากับ μ_i และความแปรปรวน (Variance) เท่ากับ $V_i^2 \sigma^2$ ค่าคลาดเคลื่อน (Error or Residual, ε_i) มีการแจกแจงปกติ มีค่าเฉลี่ย เท่ากับ 0 และค่าความแปรปรวน เท่ากับ σ^2 สมมติฐานนี้สามารถแทนได้ด้วยสัญลักษณ์ :

$$Y_i \sim Normal(\mu_i, V_i^2 \sigma^2)$$

$$Y_i = V_i \left[\sum_{j=1}^k x_{ij} \beta_j \right] + V_i \varepsilon_i \quad (2.3)$$

$$\varepsilon_i \sim Normal(0, \sigma^2)$$

$$\text{เมื่อ} \quad \mu_i = E(Y_i) = V_i \left[\sum_{j=1}^k x_{ij} \beta_j \right] \quad (2.4)$$

$$\text{ค่าความแปรปรวน (Variance),} \quad Var(Y_i) = V_i^2 \sigma^2$$

เมื่อ MLR Model อยู่ภายใต้สมมติฐานที่ข้อมูลอุบัติเหตุนั้น Y_i มีการแจกแจงปกติ การประมาณค่าสัมประสิทธิ์การถดถอยสามารถหาได้โดยวิธีกำลังสองน้อยที่สุดหรือวิธี Maximum Likelihood

2.3.2 แบบจำลองการถดถอยล็อกนอร์มอล (Log-Normal Regression Model, LNR Model)

LNR Model เป็นแบบจำลองผลคูณของตัวแปรอิสระ การเลือกรูปแบบ LNR Model จะสมมติให้ข้อมูลอุบัติเหตุนั้น $\log(Y_i)$ มีการแจกแจงปกติ ค่าคลาดเคลื่อน(Error or Residual, ε_i) มีการแจกแจงปกติ มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ σ^2

$$Y_i = V_i [\beta_1 \prod_{j=2}^k (1 + x_{ij})^{\beta_j}] e^{\varepsilon_i} \quad (2.5)$$

$$\log(Y_i) = \log V_i + \log \beta_1 + \sum_{j=2}^k \beta_j \log(1 + x_{ij}) + \varepsilon_i$$

$$\varepsilon_i \sim Normal(0, \sigma^2)$$

เพื่อหลีกเลี่ยง $\log(0)$ ในกรณีที่ไม่มีอุบัติเหตุนั้นในช่วงถนนนั้น จึงบวกค่าคงที่ δ ซึ่งมีค่าระหว่าง 0.001-1.000 เข้าไปในค่า Y_i ดังนั้นจะได้

$$\log(Y_i + \delta) = \log V_i + \log \beta_1 + \sum_{j=2}^k \beta_j \log(1 + x_{ij}) + \varepsilon_i$$

$\log(Y_i + \delta)$ มีการแจกแจงปกติค่าเฉลี่ยเท่ากับ v_i และความแปรปรวนเท่ากับ σ^2

$$\log(Y_i + \delta) \sim Normal(v_i, \sigma^2)$$

$$\text{เมื่อ } v_i = E(\log(Y_i + \delta)) = \log V_i + \log \beta_1 + \sum_{j=2}^k \beta_j \log(1 + x_{ij}) \quad (2.6)$$

$$\mu_i = E(Y_i) = -\delta + V_i[\beta_1 \prod_{j=2}^k (1 + x_{ij})^{\beta_j}] \quad (2.7)$$

ค่าความแปรปรวน (Variance), $Var(Y_i) = (\mu_i + \delta)^2 (e^{\sigma^2} - 1)$

เมื่อ LNR Model มีการแจกแจงปกติ การประมาณค่าสัมประสิทธิ์การถดถอยสามารถหาได้โดยวิธีกำลังสองน้อยที่สุด หรือวิธี Maximum Likelihood

2.3.3 แบบจำลองการถดถอยของพัชอง (Poisson Regression Model, PR Model)

การเลือกรูปแบบ PR Model จะต้องมีการสมมติให้ข้อมูลอุบัติเหตุ Y_i มีการแจกแจงแบบพัชอง มีค่าเฉลี่ยและความแปรปรวนเท่ากับ μ_i สมมติฐานนี้สามารถแทนได้ด้วยสัญลักษณ์ :

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

ความน่าจะเป็นที่จำนวนอุบัติเหตุ (Y_i) เท่ากับ y_i แทนได้ด้วยสมการ (2.8)

$$P(Y_i = y_i) = P(y_i) = \frac{e^{-\mu_i} \mu_i^{y_i}}{y_i!} \quad (2.8)$$

เมื่อ จำนวนอุบัติเหตุเฉลี่ยต่อคาบเวลา (Mean), $\mu_i = E(Y_i) = V_i[e^{\sum_{j=1}^k x_{ij}\beta_j}] \quad (2.9)$

ค่าความแปรปรวน (Variance), $Var(Y_i) = \mu_i$

เนื่องจาก PR Model อยู่ภายใต้สมมติฐานที่ข้อมูลอุบัติเหตุ Y_i มีการแจกแจงแบบพัชองการประมาณค่าสัมประสิทธิ์การถดถอยของแบบจำลองนี้จึงแตกต่างจาก MLR Model และ LNR Model คือจะใช้วิธีการประมาณค่าสัมประสิทธิ์การถดถอยได้เฉพาะวิธี Maximum Likelihood เท่านั้น

2.3.4 แบบจำลองทวินามเชิงลบ (Negative Binomial Regression Model, NBR Model)

การเลือกรูปแบบ NBR Model จะต้องมีการสมมติให้ข้อมูลอุบัติเหตุ Y_i มีการแจกแจงแบบทวินามเชิงลบ มีค่าเฉลี่ย μ_i และค่าความแปรปรวนเท่ากับ $\mu_i + \alpha\mu_i^2$ โดยที่ $\alpha > 0$ สมมติฐานนี้สามารถแทนได้ด้วยสัญลักษณ์ :

$$Y_i \sim \text{Negative Binomial}(\mu_i, \mu_i + \alpha\mu_i^2)$$

ความน่าจะเป็นที่จำนวนอุบัติเหตุ (Y_i) เท่ากับ y_i แทนได้ด้วยสมการ (2.10)

$$P(Y_i = y_i) = P(y_i) = \frac{\Gamma(\alpha^{-1} + y_i)}{y! \Gamma(\alpha^{-1})} p^{\alpha^{-1}} (1-p)^{y_i} \quad (2.10)$$

เมื่อ
$$\mu_i = E(Y_i) = V_i [e^{\sum_{j=1}^k x_{ij} \beta_j}] \quad (2.11)$$

$$\text{Var}(Y_i) = \mu_i + \alpha\mu_i^2 \quad \text{โดยที่ } \alpha > 0 \quad (2.12)$$

$$p = \frac{1}{1 + \alpha\mu_i} \quad (2.13)$$

เมื่อค่า α มีค่าเข้าใกล้ศูนย์ การแจกแจงความน่าจะเป็นรูปแบบทวินามเชิงลบในสมการที่ 2.10 จะมีค่าเข้าใกล้การแจกแจงความน่าจะเป็นรูปแบบพัวซองในสมการที่ 2.8

เมื่อตัวแปร Y_i ใน NBR Model มีการแจกแจงแบบทวินามเชิงลบ การประมาณค่าสัมประสิทธิ์การถดถอยสามารถหาได้โดยวิธี Maximum Likelihood เช่นเดียวกันกับ PR Model

2.4 การประมาณค่าพารามิเตอร์ของแบบจำลองโดยวิธี Maximum Likelihood

การประมาณค่าพารามิเตอร์ของแบบจำลองโดยวิธี Maximum Likelihood เป็นวิธีประมาณค่าสัมประสิทธิ์การถดถอยที่ทำให้ ค่า Log Likelihood ($L(\beta)$) มีค่ามากที่สุดด้วยหลักการทางแคลคูลัส และโดยที่ฟังก์ชันความน่าจะเป็นนี้อยู่ในรูปของผลคูณของฟังก์ชันความน่าจะเป็น

ของตัวแปรสุ่ม ดังนั้นอาจจะสะดวกต่อการวิเคราะห์มากขึ้นถ้าทำฟังก์ชันให้อยู่ในรูปฟังก์ชัน Logarithm ดังในสมการที่ 2.14

$$L(\beta) = \log\left(\prod_{i=1}^n P(y_i)\right) \quad (2.14)$$

โดย $\prod$ คือ ผลคูณอันดับ เช่น $\prod_{i=1}^2 X_i = X_1 X_2$
 i คือ ช่วงถนน (section) ซึ่งมีค่าเป็นจำนวนเต็มบวก
 $L(\beta)$ คือ Loglikelihood Function ที่ประเมินค่า β

Miaou et al. (1992), Miaou and Lum (1993) ใช้วิธี Maximum Likelihood ในการประมาณค่าสัมประสิทธิ์การถดถอยในแบบจำลองการถดถอยพัวซอง ซึ่งเป็นวิธีที่ใช้มากที่สุด และเป็นวิธีการโดยตรงสำหรับการประมาณค่าพารามิเตอร์ เนื่องจากเป็นวิธีที่ง่ายและค่าพารามิเตอร์ที่ได้จากการประมาณโดยวิธีนี้จะทำให้ข้อมูลตัวอย่างในสมการมีค่าใกล้เคียงของประชากรมากที่สุด ขั้นตอนการประมาณด้วยวิธี Maximum likelihood ดังนี้

- จัดทำ Log Likelihood Function โดยการจัดรูปแบบฟังก์ชันของการกระจายความน่าจะเป็นให้อยู่ในรูปความน่าจะเป็นในการเกิดอุบัติเหตุ บนถนนช่วงหนึ่ง ในระยะเวลา 1 ปี

$$\begin{aligned} L(\beta) &= \log\left(\prod_{i=1}^n p(y_i)\right) \\ &= \log\left(\prod_{i=1}^n \frac{\mu^{y_i} e^{-\mu_i}}{y_i!}\right) \\ &= \sum_{i=1}^n [y_i \log(\mu_i) - \mu_i - \log(y_i!)] \\ &= \sum_{i=1}^n [y_i (\sum_{j=1}^k x_{ij} \beta_j) + y_i \log(V_i) - V_i e^{\sum_{j=1}^k x_{ij} \beta_j} - \log(y_i!)] \end{aligned}$$

- หาค่าพารามิเตอร์โดยการหาอนุพันธ์ $L(\beta)$ เทียบกับตัวแปร $\beta_1, \beta_2, \beta_3, \dots, \beta_j$ ให้เท่ากับศูนย์ ดังสมการที่ 2.15

$$\begin{aligned}\frac{\partial L(\beta)}{\partial \beta_j} &= \sum_{i=1}^n \{y_i - V_i e^{\sum_{j=1}^k x_{ij} \beta_j}\} x_{ij} \\ &= \sum_{i=1}^n \{y_i - V_i e^{\sum_{j=1}^k x_{ij} \beta_j}\} x_{ij} = 0 \quad \text{เมื่อ } j = 1, 2, 3, \dots, k\end{aligned} \quad (2.15)$$

นอกจากรูปแบบ Log Likelihood Function ในการประมาณค่าสัมประสิทธิ์การถดถอย ในแบบจำลองการถดถอยพัวของ แล้วยังมี Log Likelihood Function ในแบบจำลองการถดถอยรูปแบบอื่นๆ ซึ่งมีรูปแบบดังนี้ (A.J. Duncan, 1974)

- รูปแบบ Log Likelihood Function ในแบบจำลอง Negative Binomial Regression

$$L(\beta, \alpha) = \sum_{i=1}^n \left[\sum_{j=0}^{y_i} \log(1 + \alpha y_i) - \log(1 + \alpha y_i) - y_i \log \mu_i - (y_i + \frac{1}{\alpha}) \log(1 + \alpha \mu_i) - \log(y_i!) \right]$$

- รูปแบบ Log Likelihood Function ในแบบจำลอง Multiple Linear Regression

$$L(\mu, \sigma^2) = -\frac{n}{2} \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{i=1}^n (y_i - \mu_i)^2$$

- รูปแบบ Loglikelihood Function ในแบบจำลอง Log Normal Regression

$$L(\lambda, \varsigma) = -\frac{n}{2} \log 2\pi\varsigma - \frac{1}{2\varsigma} \sum_{i=1}^n (y_i - \lambda_i)^2$$

โดย

$$\lambda_i = E(\log(y_i))$$

$$\varsigma = \sqrt{\text{Var}(\log y_i)}$$

2.5 แบบจำลองอุบัติเหตุ - การประยุกต์ใช้

Maher (1996) กล่าวว่า “ แบบจำลองอุบัติเหตุเป็นรูปแบบทางสถิติแบบหนึ่งซึ่งแสดงความสัมพันธ์ระหว่างจำนวนอุบัติเหตุเฉลี่ยต่อปีกับบริเวณหนึ่งบริเวณใด ๆ กับการไหลจราจร และลักษณะที่สามารถวัดได้ เช่น ประเภทของถนน ลักษณะทางเรขาคณิตหรือการออกแบบทางหลวง เป็นต้น ” แบบจำลองอุบัติเหตุมีการสร้างและพัฒนามาแล้วหลายรูปแบบ ความเหมาะสมในการเลือกใช้รูปแบบจำลองขึ้นอยู่กับลักษณะทางสถิติของข้อมูลอุบัติเหตุ ซึ่งมีรายละเอียดดังต่อไปนี้

2.5.1 แบบจำลองการถดถอยพหุคูณเชิงเส้น (Multiple Linear Regression Model, MLR Model)

ในการวิเคราะห์อุบัติเหตุ แบบจำลองการถดถอยพหุคูณเชิงเส้นเป็นรูปแบบที่นำมาใช้มากที่สุดและเป็นแบบดั้งเดิม แบบจำลองนี้เป็นรูปแบบที่กำหนดให้ตัวแปรตามคืออัตราอุบัติเหตุ และตัวแปรอิสระคือตัวแปรที่นำมาจากความสัมพันธ์ที่ซับซ้อนขององค์ประกอบของระบบทางหลวงและสภาพแวดล้อมที่มีผลต่ออุบัติเหตุ ได้แก่ คนขับ ยานพาหนะ ถนน สภาพอากาศ และปัจจัยที่เกี่ยวข้องกับสิ่งแวดล้อม ในการจะใช้แบบจำลอง MLR ลักษณะทางสถิติข้อมูลอุบัติเหตุต้องเป็นไปตามสมมติฐานเบื้องต้น ได้แก่ ตัวแปรตาม y ต้องเป็นตัวแปรสุ่มที่มีความต่อเนื่อง มีการแจกแจงปกติ (Normal distribution) และค่าความแปรปรวนที่คงที่

Jacobs (1976) ใช้การวิเคราะห์การถดถอยอย่างง่าย (Simple Regression Analysis) ศึกษาความสัมพันธ์ระหว่างอัตราอุบัติเหตุที่มีผู้บาดเจ็บ (Personal Injury Accident Rates) กับปริมาณจราจร และใช้การวิเคราะห์การถดถอยพหุคูณแบบขั้นตอน (Stepwise Multiple Regression Analysis) ศึกษาความสัมพันธ์ระหว่างอัตราอุบัติเหตุที่มีผู้บาดเจ็บกับลักษณะทางเรขาคณิตของถนนในประเทศเคนยา (Kenya) และประเทศจาไมก้า (Jamaica) จากผลการศึกษาวินิจฉัยสรุปว่าทั้งสองประเทศมีอัตราอุบัติเหตุที่มีผู้บาดเจ็บสัมพันธ์กับปริมาณจราจร ลักษณะทางเรขาคณิตที่มีนัยสำคัญต่ออัตราอุบัติเหตุต่อล้านคัน-กิโลเมตรในประเทศเคนยา ได้แก่ จำนวนทางแยกต่อกิโลเมตร (Number of junctions per kilometer) โค้งราบ (Horizontal curvature) โค้งตั้ง (Vertical curvature) และความขรุขระของผิวถนน (Surface irregularity) ส่วนในประเทศจาไมก้า ได้แก่ ความกว้างถนน (Road width) และจำนวนทางแยกต่อกิโลเมตร การนำแบบจำลองนี้ไปใช้ Jacobs แนะนำว่าถนนนั้นจะต้องมีสภาพจราจรและลักษณะถนนที่คล้ายคลึงกับถนนที่ใช้มาสร้างแบบจำลอง ในตอนท้ายได้สรุปผลการเปรียบเทียบที่ได้รับจากประเทศเคนยาและจาไมก้า กับประเทศที่พัฒนาแล้ว โดยกำหนดให้ระดับปริมาณการจราจรและลักษณะทางการออกแบบทางเรขาคณิตของถนนเหมือนกัน

พบว่าอัตราอุบัติเหตุของทั้ง 2 ประเทศสูงกว่าประเทศที่พัฒนาแล้ว ดังนั้น มันบ่งบอกว่ามีปัจจัยอื่นนอกเหนือจากนี้ที่เกี่ยวข้อง เช่น พฤติกรรมผู้ใช้งานและสภาพยานพาหนะ

Saito et al. (1998) ใช้การวิเคราะห์ถดถอยพหุคูณเชิงเส้นแทนความสัมพันธ์ระหว่างอัตราอุบัติเหตุกับลักษณะของถนนและสภาพจราจรในเกาะฮอกไกโด (Hokkaido) ประเทศญี่ปุ่น เพื่อการจัดการถนนสองช่องจราจรนอกเมืองให้มีประสิทธิภาพและความปลอดภัยมากขึ้น จากการศึกษาวิจัยพบว่าอัตราอุบัติเหตุมีแนวโน้มเพิ่มตามดัชนีการติดขัด (Congestion index) ที่เพิ่มขึ้นบนถนนที่มีความกว้างมาตรฐาน 8-9 เมตร ถนนที่มีความหนาแน่นของทางแยกต่อกิโลเมตร (Intersection density) 1-2 ทางแยกต่อกิโลเมตร มีอัตราการเกิดอุบัติเหตุต่ำสุดเท่ากับ 27.1 อุบัติเหตุต่อร้อยล้านคันกิโลเมตร นอกจากนี้ยังพบอีกว่าอัตราอุบัติเหตุมีแนวโน้มลดลงเมื่อพิักัดความเร็ว (Speed limit) เป็น 40 และ 50 กิโลเมตรต่อชั่วโมง ส่วนพิักัดความเร็วช่วง 60 กิโลเมตรต่อชั่วโมงมีแนวโน้มไม่ชัดเจนเพราะขนาดตัวอย่างน้อย ผลจากการวิเคราะห์จัดลำดับความสัมพันธ์ของตัวแปรพบว่าตัวแปรที่มีส่วนเสริมให้เกิดอุบัติเหตุจากมากไปน้อย ได้แก่ ความเร็วเฉลี่ยในการเดินทาง (Average travel speed) ความหนาแน่นของทางแยก เปอร์เซ็นต์รถหนัก (Percentage of heavy vehicles) และดัชนีการติดขัด ส่วนพิักัดความเร็วไม่มีนัยสำคัญ

การวิเคราะห์และการสร้างแบบจำลองอุบัติเหตุโดยรูปแบบจำลอง MLR มีข้อจำกัดคือการสรุปผลการวิเคราะห์จะทำให้ได้อย่างถูกต้องก็ต่อเมื่อลักษณะการเกิดอุบัติเหตุต้องเป็นไปตามสมมติฐานเบื้องต้น แต่ถ้าข้อสมมติฐานเบื้องต้นไม่เป็นจริงผลการวิเคราะห์อาจจะผิดพลาดได้ ซึ่งในความเป็นจริงแล้วข้อมูลอุบัติเหตุเป็นข้อมูลชนิดไม่ต่อเนื่อง ข้อมูลไม่มีค่าเป็นลบ ซึ่งตรงกันข้ามกับการแจกแจงปกติตามสมมติฐานที่ตั้งไว้ล่วงหน้า นอกจากนี้ยังมีกรณีที่ค่าความแปรปรวนไม่คงที่อันมีสาเหตุเนื่องมาจากอุบัติเหตุที่เกิดขึ้นมีการกระจายในรูปแบบอื่นๆ เช่น การแจกแจงแบบพัวซอง (Poisson Distribution) เป็นต้น

2.5.2 แบบจำลองการถดถอยพัวซอง (Poisson Regression Model, PR Model)

โดยธรรมชาติ อุบัติเหตุที่เกิดขึ้นมีลักษณะกระจาย ไม่ต่อเนื่อง ไม่เป็นลบ (Non-negative) และในช่วงเวลาและระยะทางที่ศึกษามีโอกาสเกิดอุบัติเหตุค่อนข้างน้อย เนื่องจากสภาพการเกิดอุบัติเหตุบังคับ ทำให้การสร้างแบบจำลองโดยใช้แบบจำลองการถดถอยพหุคูณเชิงเส้นไม่เหมาะสมจะนำไปทำนายจำนวนอุบัติเหตุ จากเหตุผลดังกล่าวจึงมีการนำเสนอแบบจำลองใหม่ ที่อาศัยการแจกแจงแบบพัวซองเข้ามาใช้

Miaou et al. (1992) นำเสนอ PR Model เพื่อสร้างความสัมพันธ์ระหว่างอุบัติเหตุของรถบรรทุก (Truck Accidents) กับตัวแปรสำคัญทางการออกแบบทางเรขาคณิตของถนน ในช่วง

ถนนใดๆ จะสมมติว่าจำนวนอุบัติเหตุของรถบรรทุกต่อปีมีการกระจายแบบพัวซอง การศึกษาวิจัยได้แนะนำว่า ปริมาณจราจรต่อช่องจราจร (AADT per lane), โค้งแนวราบ (Horizontal curvature) และ ความลาดชัน (Vertical grade) มีสหสัมพันธ์กับอัตราอุบัติเหตุที่เกี่ยวข้องกับรถบรรทุก แต่ความกว้างไหล่ทางมีความสัมพันธ์น้อยเมื่อเทียบกับตัวแปรอื่นๆ การทดสอบภาวะรูปพอดี (Goodness-of-fit) ทางสถิติบ่งบอกว่าเกิดการกระจายมากเกินไป (Overdispersion) ในแบบจำลองพัวซองซึ่งเกิดขึ้นเนื่องจากการไม่นำตัวแปรอื่นๆ ที่เกี่ยวข้องมารวมในแบบจำลอง อย่างไรก็ตามการเกิดความแปรปรวนมากก็ไม่ได้ทำให้การสรุปความสัมพันธ์ระหว่างอุบัติเหตุรถบรรทุกกับลักษณะทางเรขาคณิตที่ทำการทดสอบและตัวแปรจราจรนั้นเปลี่ยนไป

Miaou and Lum (1993) แสดงการใช้ PR Model เพื่อประเมินผลของเรขาคณิตของถนนต่ออัตราเกี่ยวข้องกับอุบัติเหตุของรถบรรทุก คาดคะเนเปอร์เซ็นต์ของจำนวนอุบัติเหตุรถบรรทุกที่คาดว่าจะลดลงหลังการปรับปรุงเรขาคณิตของทางหลวง รถบรรทุกที่นำมาพิจารณามีขนาดมากกว่า 4,545 กิโลกรัม หรือ 10,000 ปอนด์ จากการศึกษาพบว่าปัจจัยที่มีผลให้เกิดอุบัติเหตุของรถบรรทุกได้แก่ ปริมาณจราจรต่อช่องจราจร (AADT per lane) องศาของโค้งราบ (Degree of horizontal curve) ความยาวโค้งราบ (Length of horizontal curve) ความลาดชัน (Vertical grade) ความยาวความลาดชัน (Length of grade) ความกว้างผิวจราจร (Width of surface) ความกว้างไหล่ทาง (Width of shoulder) และเปอร์เซ็นต์รถบรรทุกในกระแสจราจร (% truck in the traffic stream)

จากผลการวิจัยที่ผ่านมา (Miaou et al., 1992; Miaou and Lum, 1993; Vijayalakshmi, 1997; etc.) เปิดเผยว่าแบบจำลองการถดถอยพัวซองสามารถนำเสนอผลได้ถูกต้องใกล้เคียงความเป็นจริงมากกว่าแบบจำลองถดถอยพหุคูณเชิงเส้นในการวิเคราะห์อุบัติเหตุ เพราะมันมีคุณสมบัติทางสถิติและสมมติฐานเบื้องต้นที่สอดคล้องกับลักษณะข้อมูลอุบัติเหตุมากกว่า

จากที่กล่าวมาแล้วทั้งหมด สำหรับการศึกษาค้นคว้าความปลอดภัยจราจร จึงมีการนำเสนอแบบจำลองการถดถอยพัวซองเป็นทางเลือกหนึ่งในการสร้างแบบจำลองนอกเหนือจากแบบจำลองถดถอยพหุคูณ เพราะว่าสามารถใช้ได้กับข้อมูลที่มีการแจกแจงไม่ปกติ ค่าเฉลี่ยไม่เป็นลบ และขนาดของกลุ่มข้อมูลที่มีจำนวนน้อย เช่น กรณีของข้อมูลจำนวนอุบัติเหตุ

2.5.3 แบบจำลองการถดถอยทวินามเชิงลบ (Negative Binomial Regression Model, NBR Model)

ถึงแม้ว่างานวิจัยที่ผ่านมาหลายๆ ฉบับจะนำเสนอแบบจำลองอุบัติเหตุโดยใช้แบบจำลองการถดถอยพัวซองแต่มันก็มีข้อจำกัด คือจะใช้ได้ในกรณีที่ค่าเฉลี่ยต่อค่าความแปรปรวนมีค่าเท่ากับ 1 หรือค่าเฉลี่ยต้องเท่ากับค่าความแปรปรวนนั่นเอง ดังนั้นจึงมีปัญหาเกิดขึ้นในกรณีที่ค่า

ความแปรปรวนต่อค่าเฉลี่ยมีค่ามากกว่า 1 คือกรณีที่ค่าความแปรปรวนมีมากเกิดการกระจายมากเกินไป (Overdispersion) หรือกรณีที่ค่าความแปรปรวนต่อค่าเฉลี่ยมีค่าน้อยกว่า 1 คือกรณีที่ค่าความแปรปรวนมีน้อยเกิดการกระจายน้อยเกินไป (Underdispersion)

Sayed and Rodriguez (1999) อ้างงานวิจัยของ Kulmala (1995) ได้แสดงให้เห็นว่าข้อมูลอุบัติเหตุโดยมากมีค่าความแปรปรวนมาก ด้วยเหตุผลนี้ทำให้มีงานวิจัยหลายฉบับ (Hadi, et al., 1995; Shankar et al., 1995; Khan, 1999; etc.) นำเสนอแบบจำลองใหม่ที่อาศัยการแจกแจงแบบทวินามเชิงลบเข้ามาใช้

Hadi et al. (1995) ใช้ NBR Model เพื่อประมาณผลของส่วนต่างๆ ของรูปตัดขวางถนนที่มีผลต่ออัตราการชนทั้งหมด อัตราการชนที่เกิดการบาดเจ็บ และอัตราการชนที่ทำให้มีคนตายสำหรับทางหลวงนอกเมืองและทางหลวงในเมืองที่ระดับปริมาณจราจรแตกต่างกัน จากการศึกษารูปได้ว่าอัตราอุบัติเหตุขึ้นอยู่กับชนิดของทางหลวง การเพิ่มความกว้างช่องจราจร ความกว้างเกาะกลาง ความกว้างไหล่ทางด้านในและ/หรือความกว้างไหล่ทางด้านนอกมีประสิทธิผลในการลดอัตราการชน

2.5.4 แบบจำลองการถดถอย Log Normal (Log Normal Regression Model, LNR Model)

Khan et al. (1999) อ้างงานวิจัยของ Miaou and Lum (1993) กล่าวว่า Miaou and Lum แสดงความสัมพันธ์ของข้อมูลอุบัติเหตุโดยใช้ NBR Model เนื่องจากข้อมูลอุบัติเหตุมีลักษณะเบ้มาก (Extremely skewed) แต่พบว่า NBR Model ก็อาจจะไม่เหมาะสม จากประสบการณ์ของ Miaou and Lum ในตอนนั้น LNR Model เป็นรูปแบบจำลองทางเลือกหนึ่งสำหรับข้อมูลอุบัติเหตุที่นำมาใช้

Kihberg et al. (1967) ใช้ LNR Model เพื่อหาความสัมพันธ์ของอุบัติเหตุยานยนต์กับประเภททางหลวง และองค์ประกอบทางการออกแบบทางหลวง (Highway design elements) การศึกษาประกอบด้วย 2 โครงการ

โครงการที่ 1 มีวัตถุประสงค์เพื่อหาอัตราอุบัติเหตุสำหรับประเภททางหลวง การสร้างแบบจำลองได้แบ่งช่วงถนนออกเป็นช่วงๆ โดยให้ช่วงถนนมีลักษณะเหมือนกันได้แก่ จำนวนช่องจราจร มี/ไม่มีเกาะกลาง และมี/ไม่มีการควบคุมทางเชื่อมเข้าออก ในขั้นตอนการวิเคราะห์ได้ใช้แบบจำลองที่มีรูปแบบ $\log A = a + b_1 \log L + b_2 \log T$ เมื่อ A คือค่าเฉลี่ยการเกิดอุบัติเหตุ L คือความยาวช่วงถนน T คือปริมาณจราจรเฉลี่ยต่อวัน (ADT) และค่า a, b₁, b₂ คือ ค่าพารามิเตอร์ และเพื่อหลีกเลี่ยง $\log(0)$ ในบางช่วงถนน จึงกำหนดให้ ใช้ $\log(A+1)$ แทนการใช้ $\log(A)$

โครงการที่ 2 ทำการศึกษาเพิ่มเติมโดยออกแบบการวิจัยให้อัตราอุบัติเหตุสัมพันธ์กับข้อมูลทางลักษณะเรขาคณิตของถนนอันได้แก่ องศาโค้ง (Degree of curvature) ปริมาณความลาดชัน (Amount of grade) จำนวนทางแยก (Number of intersections) และจำนวนโครงสร้าง (Number of structures) การสร้างแบบจำลองใช้การแบ่งช่วงถนนให้มีความยาวเท่ากัน ทำการลองแทนค่าความยาวช่วงถนนโดยเริ่มตั้งแต่ 0.1 ไมล์ และในที่สุด Kihberg et al. ได้ใช้ค่า 0.3 ไมล์ เป็นค่าที่เหมาะสมในการนำมาใช้ในการกำหนดความยาวช่วงถนนของแบบจำลองอุบัติเหตุ

จากการศึกษาวิจัยสรุปว่า (1) ประเภททางหลวงที่จำแนกโดย จำนวนช่องจราจร การมี/ไม่มีเกาะกลาง และการมี/ไม่มีการควบคุมทางเชื่อมเข้าออกมีผลต่อการเกิดอุบัติเหตุ ซึ่งทางหลวงประเภทสี่ช่องจราจรมีอัตราเกิดอุบัติเหตุสูงกว่าทางหลวงประเภทสองช่องจราจรเมื่อไม่มีเกาะกลางและไม่มีทางเชื่อมเข้าออก นอกจากนี้ทางหลวงดังกล่าวยังเป็นทางหลวงที่มีอัตราอุบัติเหตุสูงสุดในกลุ่มของประเภททางหลวง ส่วนทางหลวงที่มีอัตราอุบัติเหตุต่ำสุดในกลุ่มคือทางหลวงประเภทสี่ช่องจราจรที่มีเกาะกลางและมีทางเชื่อมเข้าออก นอกจากนี้ Kirberg et al. ยังสรุปต่อไปว่า การควบคุมทางเชื่อมเข้าออกมีผลทำให้อุบัติเหตุบนถนนน้อยลงมากที่สุด ในขณะที่การติดตั้งเกาะกลาง แม้ว่าจากการศึกษาจะยังให้ผลไม่ชัดเจน แต่การติดตั้งเกาะกลางก็มีแนวโน้มที่มีผลทำให้อุบัติเหตุลดน้อยลง (2) องค์ประกอบทางเรขาคณิตที่มีผลให้เกิดอุบัติเหตุมากที่สุดในกลุ่มคือ ทางแยก ซึ่งพบว่าอัตราอุบัติเหตุในช่วงถนนที่มีทางแยกหลายๆ จะเป็น 3 เท่าของอัตราอุบัติเหตุในช่วงถนนที่ไม่มีทางแยก