



การวิเคราะห์ข้อมูลทางสถิติ  
ของอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562

จัดทำโดย  
นายพัศตร์ภูมิ ตาแพร่ 62010609

เสนอ  
ผศ. ดร.สุรินทร์ กิตติธรรมกุล

รายงานนี้เป็นส่วนหนึ่งของรายวิชา Probability and Statistics 0107625301

ภาคเรียนที่ 2 ปีการศึกษา 2563

คณะวิศวกรรมศาสตร์ ภาควิชาวิศวกรรมคอมพิวเตอร์

สถาบันเทคโนโลยีเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

## คำนำ

รายงานการวิเคราะห์ข้อมูลทางสถิติของอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562 จัดทำขึ้นเพื่อทดสอบความรู้ความเข้าใจของผู้จัดทำ ในการวิเคราะห์ข้อมูลทางสถิติของอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562 และนำเสนอข้อมูลจากการวิเคราะห์ข้อมูลอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562 ซึ่งเป็นส่วนหนึ่งในรายวิชา 0107625301 Probability and Statistics

ผู้จัดทำหวังเป็นอย่างยิ่ง ว่ารายงานฉบับนี้ จะเป็นประโยชน์แก่ผู้ที่สนใจข้อมูลทางสถิติของข้อมูลอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562

นายพัทธ์ภูมิ ตาแพร่

ผู้จัดทำ

## สารบัญ

	หน้า
คำนำ	ก
สารบัญ	ข
เนื้อหา	
รายละเอียดข้อมูลที่ทำการวิเคราะห์	1
การวิเคราะห์ค่าสถิติพื้นฐาน	3
การวิเคราะห์ฟังก์ชันการแจกแจงความน่าจะเป็น และความน่าจะเป็นสะสม	8
การวิเคราะห์ช่วงค่าเฉลี่ยประชากร	12
การวิเคราะห์ความสัมพันธ์ โดย Linear Regression	16
บรรณานุกรม	ค
ภาคผนวก	

## รายละเอียดข้อมูลที่ทำกรวิเคราะห์

### ข้อมูลอภิปันธ์ (Metadata) :

ชื่อข้อมูล : อุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562

#### คำอธิบายข้อมูล

ข้อมูลอุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม ประกอบด้วย อุบัติเหตุที่เกิดขึ้นบนถนนทางหลวงชนบท และทางด่วนปี 2562

#### วันที่สร้างชุดข้อมูล

23 ส.ค. 2563

#### วันที่ปรับปรุงข้อมูลล่าสุด

6 ม.ค. 2564

#### ชื่อผู้ติดต่อ

ศูนย์เทคโนโลยีสารสนเทศและการสื่อสาร สำนักรับผิดกระทรวงคมนาคม

### พจนานุกรมข้อมูล (Data Dictionary)

- ปีที่เกิดเหตุ : ปีที่เกิดอุบัติเหตุ
- วันที่เกิดเหตุ : วันที่เกิดอุบัติเหตุ
- เวลา : เวลาเกิดอุบัติเหตุ
- เวลาที่รายงาน : เวลาที่ทำการรายงานเข้าสู่ระบบ
- ACC\_CODE : รหัสแทนของอุบัติเหตุครั้งนั้น
- หน่วยงาน : หน่วยงานที่ทำการรายงานเข้าสู่ระบบ
- สายทาง : เส้นทางที่เกิดอุบัติเหตุ
- ก.ม. : หลักกิโลเมตรที่เกิดอุบัติเหตุ
- จังหวัด : จังหวัดที่เกิดอุบัติเหตุ
- ลักษณะรถคันที่ 1 : ลักษณะรถที่เป็นต้นเหตุของอุบัติเหตุ

- บริเวณที่เกิดอุบัติเหตุ/ลักษณะทาง : ลักษณะทางบริเวณที่เกิดเหตุ
- มูลเหตุสันนิษฐาน : ข้อสันนิษฐานถึงสาเหตุการเกิดอุบัติเหตุ
- ลักษณะอุบัติเหตุ : ลักษณะของอุบัติเหตุ
- จำนวนรถที่เกิดอุบัติเหตุ : จำนวนรถที่เกิดอุบัติเหตุ (รวมคันที่1)
- จำนวนผู้เสียชีวิต : จำนวนผู้เสียชีวิตในอุบัติเหตุ
- จำนวนผู้บาดเจ็บ : จำนวนผู้บาดเจ็บในอุบัติเหตุ
- สภาพอากาศ : ลักษณะสภาพอากาศขณะที่เกิดอุบัติเหตุ
- LATITUDE : พิกัดละติจูด
- LONGITUDE : พิกัดลองจิจูด

### คอลัมน์ที่ทำการวิเคราะห์

- จำนวนผู้บาดเจ็บ : จำนวนผู้บาดเจ็บในอุบัติเหตุ
- จำนวนผู้เสียชีวิต : จำนวนผู้บาดเจ็บในอุบัติเหตุ

ผู้จัดทำมีความสนใจข้อมูล 2 คอลัมน์นี้ เนื่องจากผู้จัดทำมีความสนใจที่จะศึกษาข้อมูลอุบัติเหตุบนท้องถนน โดยมุ่งเน้นไปที่การศึกษาความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ เพื่อตอบข้อสงสัยว่า ความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ มีความสัมพันธ์กันหรือไม่ ลักษณะใด

### แหล่งที่มาของข้อมูล

<https://data.go.th/dataset/roadaccident>

## การวิเคราะห์ค่าสถิติพื้นฐาน

คอลัมน์ที่เลือกไว้วิเคราะห์ในครั้งนี้ : จำนวนผู้บาดเจ็บในอุบัติเหตุ ,จำนวนผู้เสียชีวิตในอุบัติเหตุ

ตัวอย่าง ข้อมูลที่ใช้ในการวิเคราะห์

ปีที่เกิดเหตุ	วันที่เกิดเหตุ	เวลา	รถคันที่ 1	จำนวนผู้เสียชีวิต	รวมจำนวนผู้บาดเจ็บ
2019	01/01/2019	00:00	รถจักรยานยนต์	0	2
2019	01/01/2019	00:03	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	2
2019	01/01/2019	00:05	รถจักรยานยนต์	1	0
2019	01/01/2019	00:20	รถจักรยานยนต์	0	1
2019	01/01/2019	00:25	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	0
2019	01/01/2019	00:30	รถปิคอัพบรรทุก 4 ล้อ	0	2
2019	01/01/2019	00:30	รถจักรยานยนต์	0	2
2019	01/01/2019	00:35	รถจักรยานยนต์	1	0
2019	01/01/2019	00:40	รถจักรยานยนต์	3	0
2019	01/01/2019	00:45	รถจักรยานยนต์	0	1
2019	01/01/2019	00:45	รถจักรยานยนต์	1	0
2019	01/01/2019	00:45	รถจักรยานยนต์	0	1
2019	01/01/2019	01:00	รถจักรยานยนต์	0	1
2019	01/01/2019	01:00	รถจักรยานยนต์	0	1
2019	01/01/2019	01:00	รถจักรยานยนต์	1	0
2019	01/01/2019	01:04	รถปิคอัพบรรทุก 4 ล้อ	0	1
2019	01/01/2019	01:15	รถจักรยานยนต์	0	2
2019	01/01/2019	01:20	รถจักรยานยนต์	1	1
2019	01/01/2019	01:20	รถจักรยานยนต์	0	1
2019	01/01/2019	01:30	รถปิคอัพบรรทุก 4 ล้อ	0	5
2019	01/01/2019	01:30	รถปิคอัพบรรทุก 4 ล้อ	0	1
2019	01/01/2019	01:30	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	1
2019	01/01/2019	01:30	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	0
2019	01/01/2019	01:30	รถจักรยานยนต์	1	1
2019	01/01/2019	01:30	รถปิคอัพบรรทุก 4 ล้อ	0	1
2019	01/01/2019	01:30	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	1	0

รูปที่ 1 แสดงตัวอย่างข้อมูลที่ใช้ในการวิเคราะห์

## ผลการวิเคราะห์ค่าสถิติพื้นฐาน

```
=====
จำนวนอุบัติเหตุทั้งหมด : 19966 ครั้ง
จำนวนผู้บาดเจ็บทั้งหมด : 18191 คน
จำนวนผู้เสียชีวิตทั้งหมด : 3258 คน

-----

(median) มัธยฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : 1 คน
(mode) ฐานนิยม จำนวนผู้บาดเจ็บในอุบัติเหตุ : 0 คน
(mean) ค่าเฉลี่ย จำนวนผู้บาดเจ็บในอุบัติเหตุ : 0.91 คน
(SD) ส่วน เบี่ยงเบนมาตรฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : 1.70

(median) มัธยฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0 คน
(mode) ฐานนิยม จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0 คน
(mean) ค่าเฉลี่ย จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0.16 คน
(SD) ส่วน เบี่ยงเบนมาตรฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0.49
=====
```

รูปที่ 2 แสดงผลการผลการวิเคราะห์ค่าสถิติพื้นฐาน

จากการคำนวณค่าสถิติพื้นฐาน ได้ผลลัพธ์ดังนี้

### สถิติพื้นฐานของจำนวนผู้บาดเจ็บในอุบัติเหตุ

ค่ามัธยฐานของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 1 คน

ค่าฐานนิยมของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 0 คน

ค่าเฉลี่ยของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 0.91 คน

ส่วนเบี่ยงเบนมาตรฐานของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 1.70 คน

### สถิติพื้นฐานของจำนวนผู้เสียชีวิตในอุบัติเหตุ

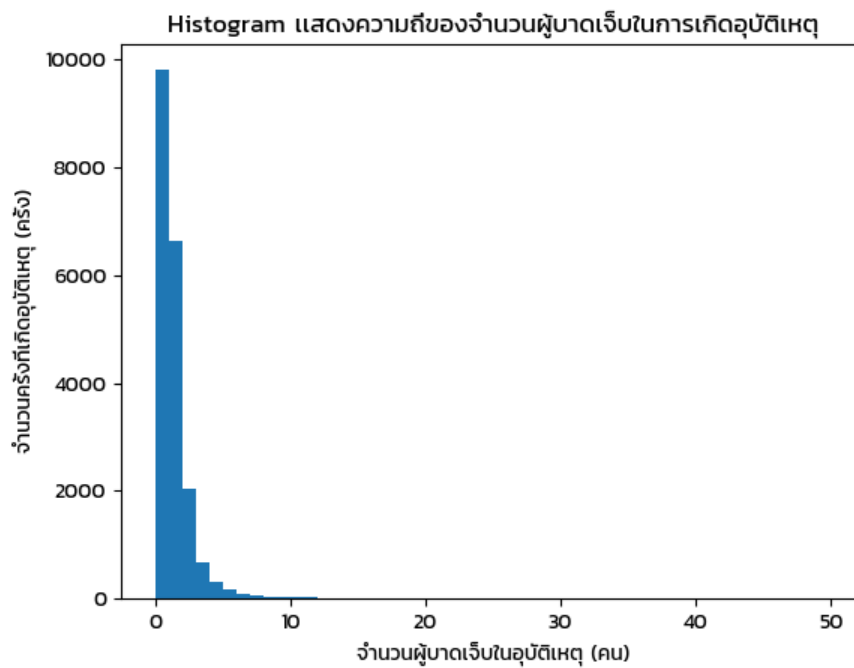
ค่ามัธยฐานของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0 คน

ค่าฐานนิยมของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0 คน

ค่าเฉลี่ยของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0.16 คน

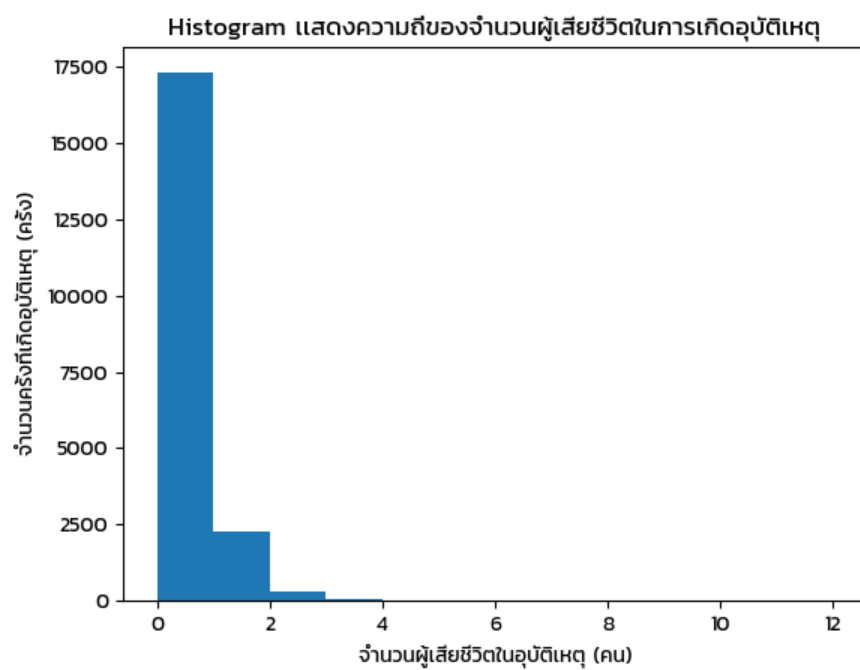
ส่วนเบี่ยงเบนมาตรฐานของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0.49 คน

## ผลการวิเคราะห์ Histogram



จาก Histogram แสดงความถี่ของจำนวนผู้บาดเจ็บในการเกิดอุบัติเหตุ พบว่า โดยส่วนใหญ่แล้ว การเกิดอุบัติเหตุบนท้องถนนในแต่ละครั้ง จะมีจำนวนผู้บาดเจ็บ 0-1 คน





จาก Histogram แสดงความถี่ของจำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุ พบว่า โดยส่วนใหญ่แล้ว การเกิดอุบัติเหตุบนท้องถนนในแต่ละครั้ง จะมีจำนวนผู้เสียชีวิต 0-1 คน

## ผลการวิเคราะห์ XY Scatter Plot

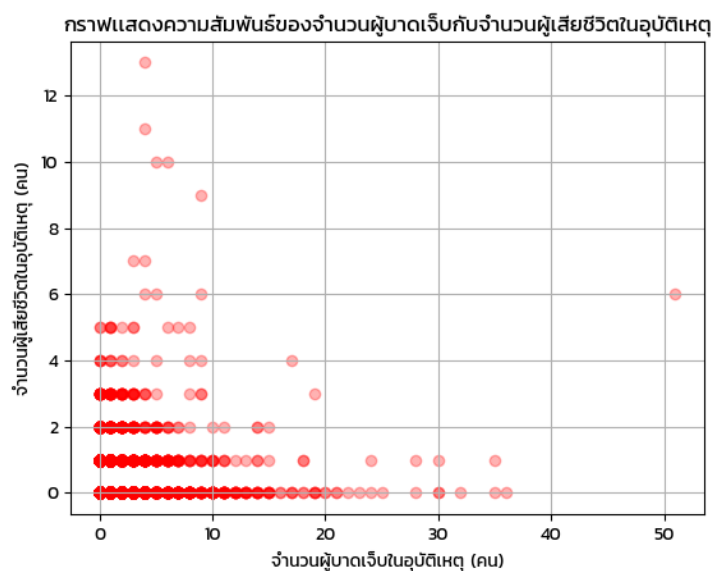
### ตัวแปรที่เลือกในการทำ XY Scatter Plot

ตัวแปรต้น : จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)

ตัวแปรตาม : จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)

### เหตุผลในการเลือกตัวแปร

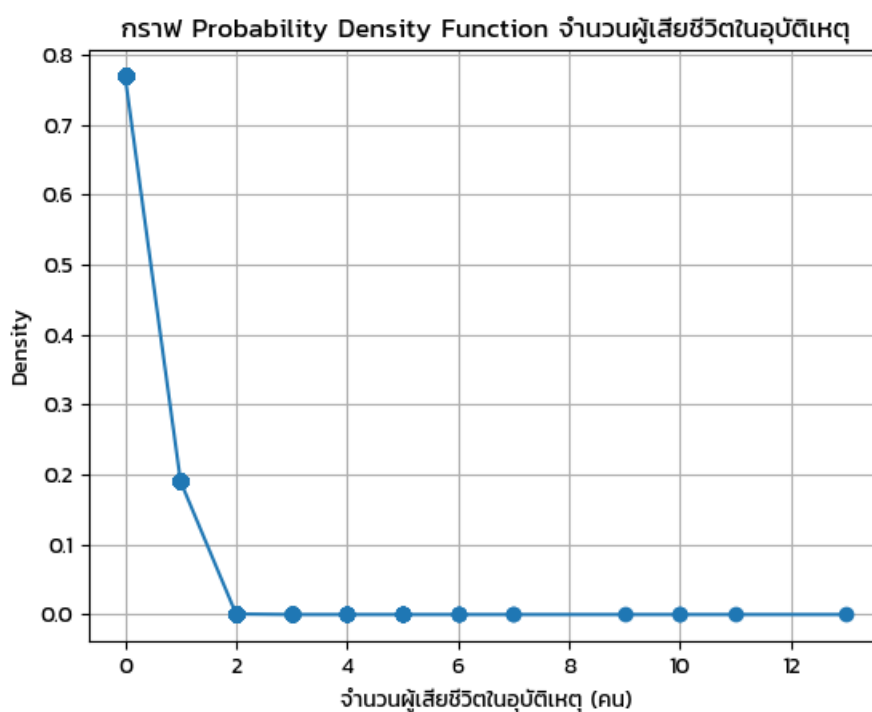
เนื่องจาก XY Scatter Plot เป็นกราฟสำหรับแสดงแนวโน้มและความสัมพันธ์ของข้อมูลชนิด Numeric ต่อข้อมูลชนิด Numeric ซึ่งจำนวนผู้บาดเจ็บในอุบัติเหตุ และจำนวนผู้เสียชีวิตในอุบัติเหตุ นั้นเป็นข้อมูลชนิด Numeric ประกอบกับเพื่อที่จะสังเกตแนวโน้มและความสัมพันธ์ของจำนวนผู้บาดเจ็บในการเกิดอุบัติเหตุ และจำนวนผู้เสียชีวิตในอุบัติเหตุ แต่ละครั้ง



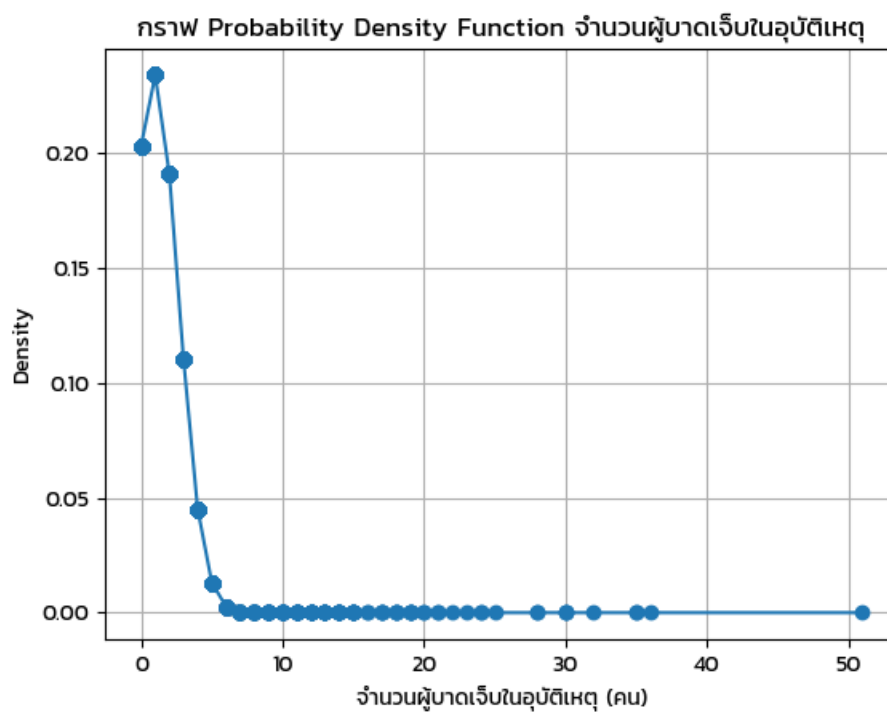
## การวิเคราะห์ฟังก์ชันการแจกแจงความน่าจะเป็น และความน่าจะเป็นสะสม

คอลัมน์ที่เลือกใช้วิเคราะห์ในครั้งนี้ : จำนวนผู้บาดเจ็บในอุบัติเหตุ , จำนวนผู้เสียชีวิตในอุบัติเหตุ

### Probability Density Function

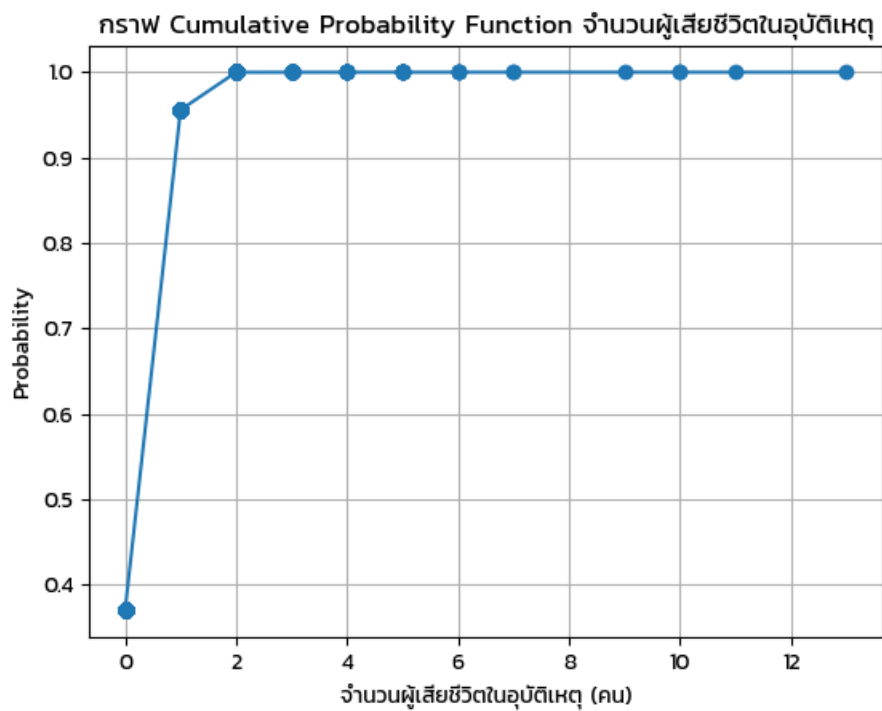


จากกราฟ Probability Density Function จำนวนผู้เสียชีวิตในอุบัติเหตุ พบว่า ในการเกิดอุบัติเหตุแต่ละครั้ง โอกาสที่จะมีผู้เสียชีวิต 0 คน (ไม่มีผู้เสียชีวิต) มีโอกาสเกิดขึ้นสูงที่สุด และในทางกลับกัน โอกาสที่จำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุตั้งแต่ 2 คนขึ้นไป มีโอกาสเกิดขึ้นน้อยมาก

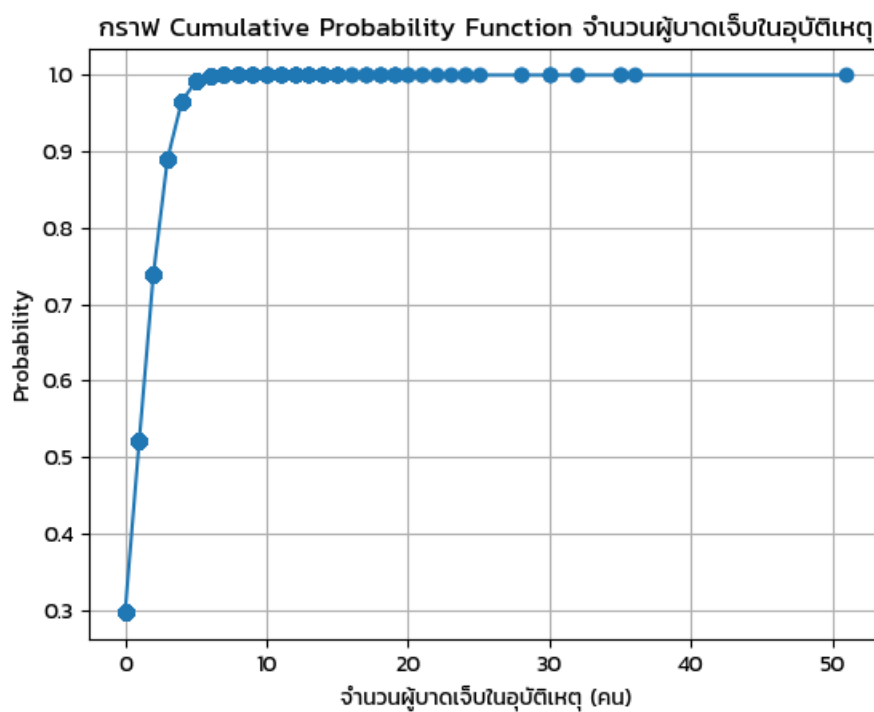


จากกราฟ Probability Density Function จำนวนผู้บาดเจ็บในอุบัติเหตุ พบว่า ในการเกิดอุบัติเหตุแต่ละครั้ง โอกาสที่จะมีผู้บาดเจ็บประมาณ 1-2 คน มีโอกาสเกิดขึ้นสูงที่สุด และในทางกลับกัน โอกาสที่จำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุตั้งแต่ประมาณ 5 คนขึ้นไป มีโอกาสเกิดขึ้นน้อยมาก

## Cumulative Probability Function



จากกราฟ Cumulative Probability Function จำนวนผู้เสียชีวิตในอุบัติเหตุ พบว่า ในการเกิดอุบัติเหตุแต่ละครั้ง มักมีผู้เสียชีวิตไม่เกิน 2 คน หรือกล่าวคือ โอกาสที่จำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุแต่ละครั้งจะน้อยกว่า 2 มีค่าสูง



จากกราฟ Cumulative Probability Function จำนวนผู้บาดเจ็บในอุบัติเหตุ พบว่า ในการเกิดอุบัติเหตุแต่ละครั้ง มักมีผู้บาดเจ็บไม่เกิน 5 คน หรือกล่าวคือ โอกาสที่จำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุแต่ละครั้งจะน้อยกว่า 5 มีค่าสูง

## การวิเคราะห์ช่วงค่าเฉลี่ยประชากร

คอลัมน์ที่เลือกไว้วิเคราะห์ในครั้งนี้ : จำนวนผู้บาดเจ็บในอุบัติเหตุ

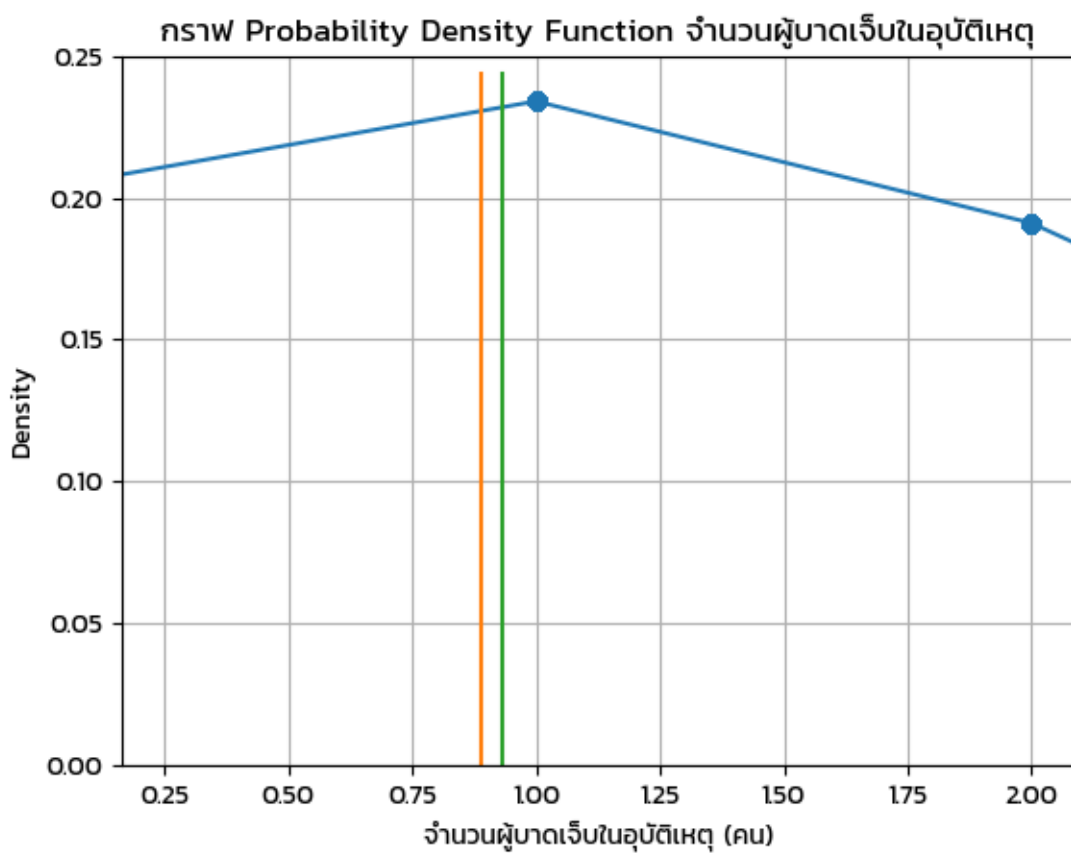
การวิเคราะห์ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ

(Confidence Interval of Mean)

```
[Confidence Level 90%] Confidence Interval of Mean : (0.891325989994717, 0.9309630157653632)
[Confidence Level 95%] Confidence Interval of Mean : (0.8875292882214763, 0.934759717538604)
[Confidence Level 99%] Confidence Interval of Mean : (0.8801088499767542, 0.942180155783326)
```

รูปที่ 3 แสดงผลการคำนวณช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level ต่างๆ

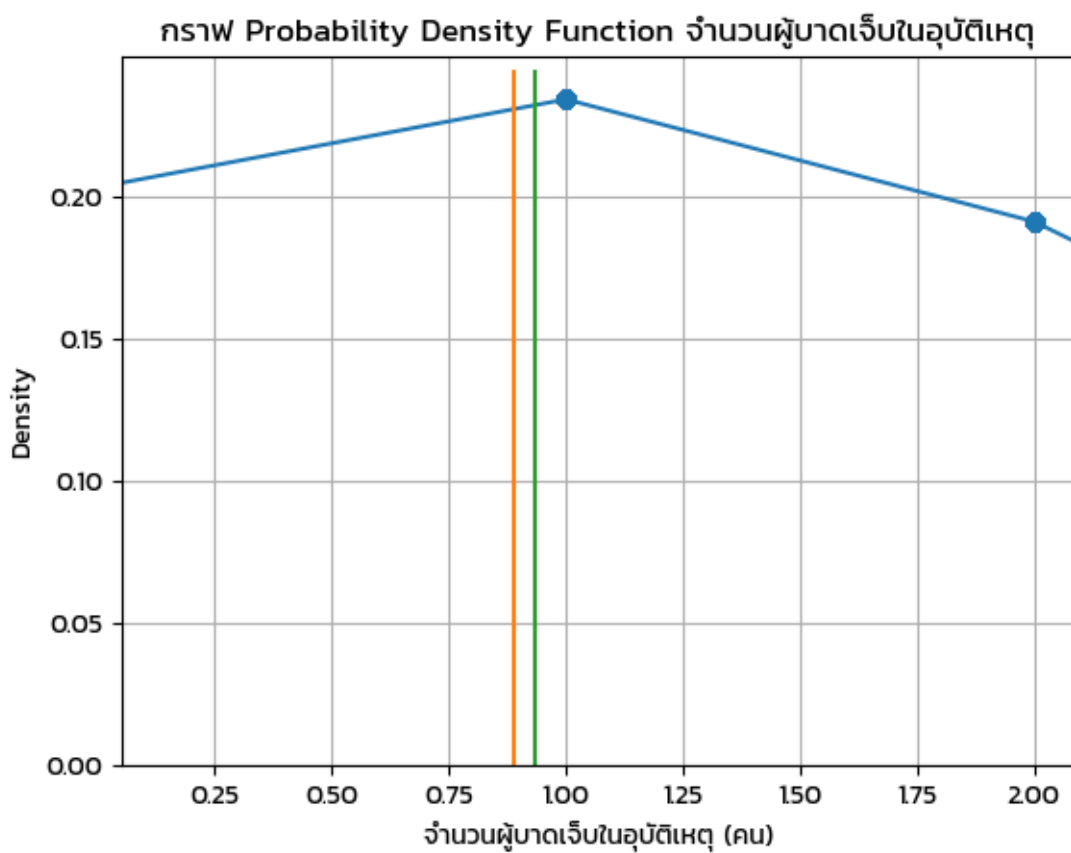
ช่วงค่าเฉลี่ยประชากร ที่ Confidence Level : 90%



จากการวิเคราะห์ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level 90% ได้ว่า ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ คือ [ 0.8913 , 0.9310 ] คน หรือสามารถกล่าวได้ว่า ค่าเฉลี่ยประชากร ( $\mu$ ) ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level 90% อยู่ในช่วง [ 0.8913 , 0.9310 ] คน

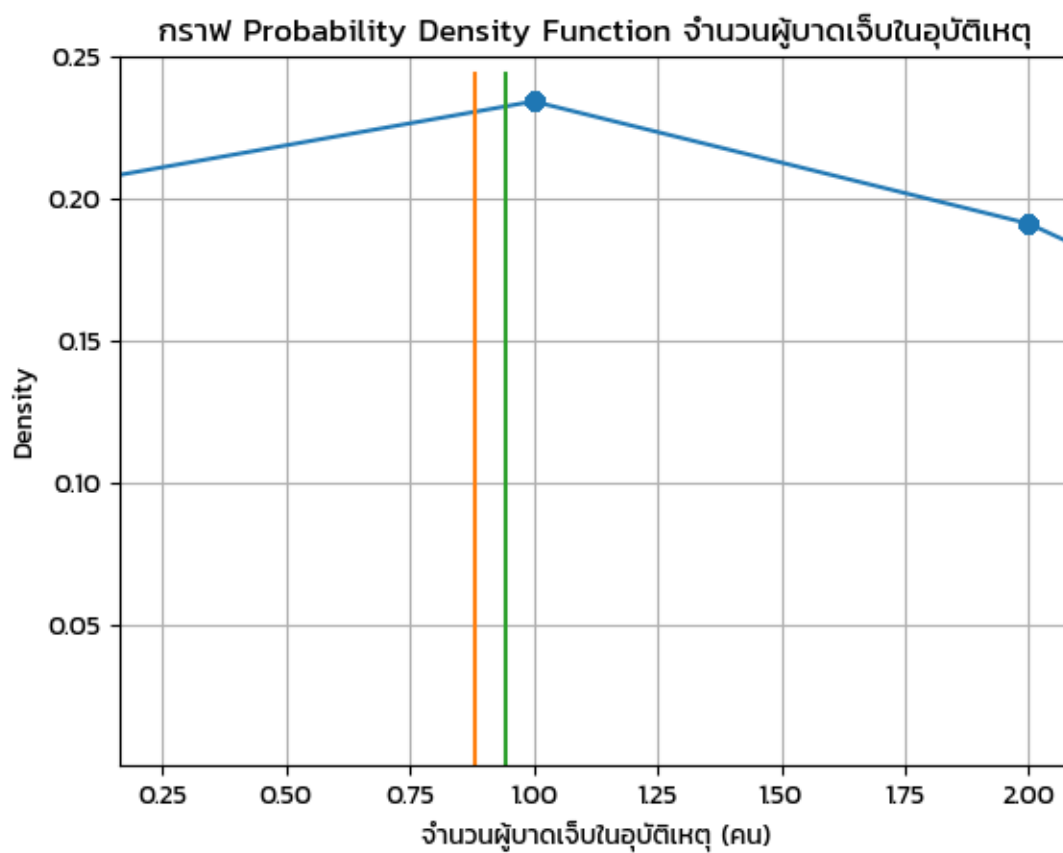


ช่วงค่าเฉลี่ยประชากร ที่ Confidence Level : 95%



จากการวิเคราะห์ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level 95% ได้ว่า ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ คือ [ 0.8875 , 0.9348 ] คน หรือสามารถกล่าวได้ว่า ค่าเฉลี่ยประชากร ( $\mu$ ) ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level 95% อยู่ในช่วง [ 0.8875 , 0.9348 ] คน

ช่วงค่าเฉลี่ยประชากร ที่ Confidence Level : 99%

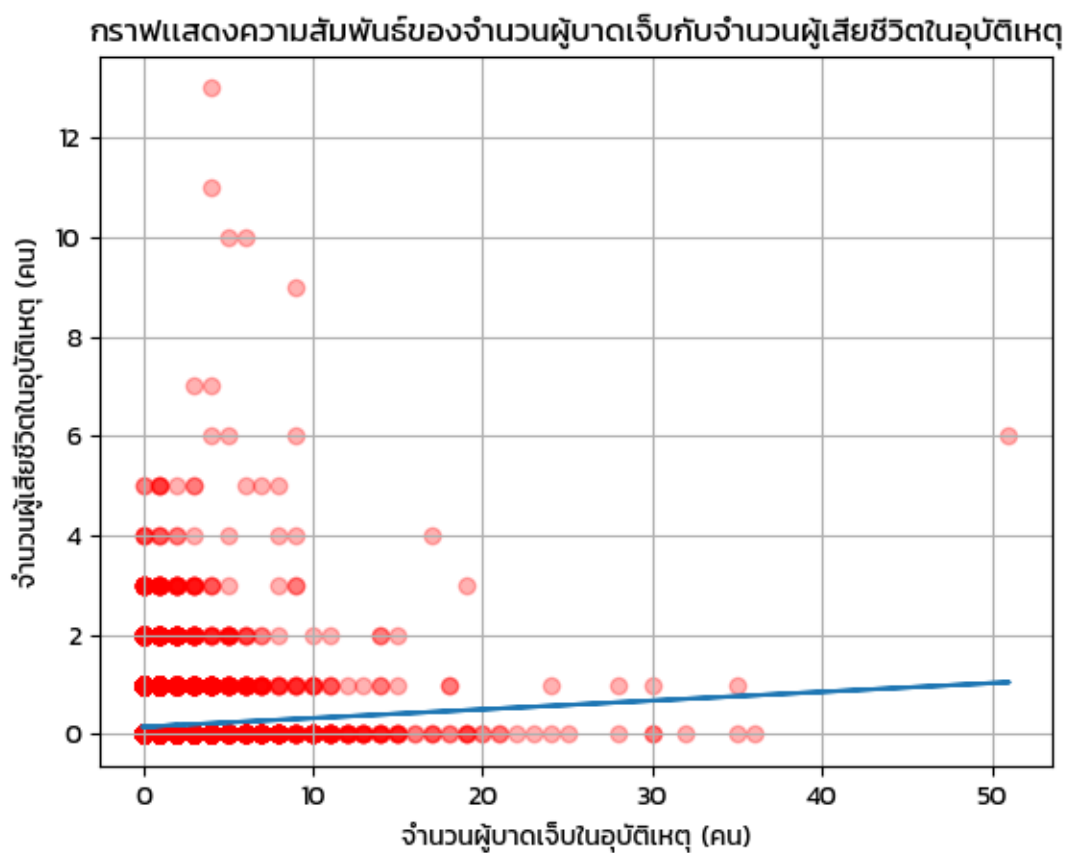


จากการวิเคราะห์ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level 99% ได้ว่า ช่วงค่าเฉลี่ยประชากร ของจำนวนผู้บาดเจ็บในอุบัติเหตุ คือ [ 0.8801 , 0.9422 ] คน หรือสามารถกล่าวได้ว่า ค่าเฉลี่ยประชากร ( $\mu$ ) ของจำนวนผู้บาดเจ็บในอุบัติเหตุ ที่ Confidence Level 99% อยู่ในช่วง [ 0.8801 , 0.9422 ] คน

## การวิเคราะห์ความสัมพันธ์ โดย Linear Regression

คอลัมน์ที่เลือกไว้วิเคราะห์ในครั้งนี้ : จำนวนผู้บาดเจ็บในอุบัติเหตุ , จำนวนผู้เสียชีวิตในอุบัติเหตุ

การวิเคราะห์ความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ โดย Linear Regression



## ผลการวิเคราะห์

```
=====
สมการ Linear Regression : Y = (0.017600744672421075*X) + 0.14714875300095107
ค่าสัมประสิทธิ์สหสัมพันธ์ (r) : 0.061045859658720183
ค่าคลาดเคลื่อนมาตรฐาน : 0.0020368131663988463
=====
```

รูปที่ 1 แสดงผลการวิเคราะห์ความสัมพันธ์ระหว่างจำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตในอุบัติเหตุ

จากผลการวิเคราะห์ความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ ได้ผลลัพธ์ ดังนี้ สมการ Linear Regression คือ  $Y = (0.02 \cdot X) + 0.15$  และค่าสัมประสิทธิ์สหสัมพันธ์ (r) คือ 0.06 โดยมีค่าคลาดเคลื่อนมาตรฐาน คือ 0.002 ซึ่งกล่าวได้ว่า ความสัมพันธ์ระหว่างจำนวนผู้บาดเจ็บและจำนวนผู้เสียชีวิตในอุบัติเหตุ นั้น มีความสัมพันธ์ในทิศทางเดียวกัน หรือกล่าวได้ว่าในการเกิดอุบัติเหตุทางท้องถนนในแต่ละครั้งนั้น “หากจำนวนผู้บาดเจ็บมีจำนวนมาก แล้วจะมีจำนวนผู้เสียชีวิตมากด้วยเช่นกัน” ทั้งนี้เนื่องจากค่าสัมประสิทธิ์สหสัมพันธ์ (r) มีค่าเพียง 0.06 ซึ่งกล่าวได้ว่าความสัมพันธ์ที่กล่าวมาข้างต้นนั้นมีความสัมพันธ์กันค่อนข้างต่ำ ซึ่งก็คือในการเกิดอุบัติเหตุแต่ละครั้งนั้น “หากจำนวนผู้บาดเจ็บมีจำนวนมาก แล้วจะมีจำนวนผู้เสียชีวิตมาก” อาจไม่สามารถกล่าวได้อย่างมั่นใจมากนัก ทั้งนี้ค่าความคลาดเคลื่อนในการวิเคราะห์นี้อยู่ที่ 0.002 ซึ่งก็คือการวิเคราะห์นี้ค่อนข้างแม่นยำ

สรุปได้ว่าความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ มีความสัมพันธ์กันค่อนข้างต่ำ โดยลักษณะความสัมพันธ์ คือ “หากจำนวนผู้บาดเจ็บมีจำนวนมาก แล้วจะมีจำนวนผู้เสียชีวิตมาก”

## บรรณานุกรม

- [1] สำนักงานปลัดกระทรวงคมนาคม. อุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562. สืบค้นเมื่อ 20 มกราคม, 2564, จาก Open Government Data of Thailand: <https://data.go.th/dataset/roadaccident>

ภาคผนวก

```

1 import csv
2 import statistics
3 import pandas as pd
4 import matplotlib.pyplot as plt
5
6 plt.rcParams['font.family'] = 'kanit'
7
8 csv_path = "../datasets/ubatiehtuthaangthnn.csv"
9
10 with open(csv_path, encoding='utf-8') as csv_file:
11     csv_reader = csv.reader(csv_file)
12
13     dead_num_list = []
14     patient_num_list = []
15
16     accident_num = 0
17     for row in csv_reader:
18
19         if accident_num != 0:
20             vechicle, dead_num, patient_num = row[10], int(row[15]), int(row[16])
21
22             dead_num_list.append(dead_num)
23             patient_num_list.append(patient_num)
24
25             accident_num += 1
26
27     ## Result
28     print("=====\n")
29
30     print("จำนวนอุบัติเหตุทั้งหมด : {} ครั้ง".format(accident_num))
31     print("จำนวนผู้บาดเจ็บทั้งหมด : {} คน".format(sum(patient_num_list)))
32     print("จำนวนผู้เสียชีวิตทั้งหมด : {} คน".format(sum(dead_num_list)))
33
34     print("\n-----\n")
35
36     print("(median) มัธยฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : {} คน".format(statistics.median(patient_num_list)))
37     print("(mode) ฐานนิยม จำนวนผู้บาดเจ็บในอุบัติเหตุ : {} คน".format(statistics.mode(patient_num_list)))
38     print("(mean) ค่าเฉลี่ย จำนวนผู้บาดเจ็บในอุบัติเหตุ : {:.2f} คน".format(statistics.mean(patient_num_list)))
39     print("(SD) ส่วนเบี่ยงเบนมาตรฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : {:.2f}".format(statistics.stdev(patient_num_list)))
40
41     print()
42
43     print("(median) มัธยฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : {} คน".format(statistics.median(dead_num_list)))
44     print("(mode) ฐานนิยม จำนวนผู้เสียชีวิตในอุบัติเหตุ : {} คน".format(statistics.mode(dead_num_list)))
45     print("(mean) ค่าเฉลี่ย จำนวนผู้เสียชีวิตในอุบัติเหตุ : {:.2f} คน".format(statistics.mean(dead_num_list)))
46     print("(SD) ส่วนเบี่ยงเบนมาตรฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : {:.2f}".format(statistics.stdev(dead_num_list)))
47
48     print("\n=====")
49
50     ## Histogram
51     plt.hist(patient_num_list, bins=range(max(patient_num_list)))
52     plt.title("Histogram แสดงความถี่ของจำนวนผู้บาดเจ็บในการเกิดอุบัติเหตุ")
53     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
54     plt.ylabel("จำนวนครั้งที่เกิดอุบัติเหตุ (ครั้ง)")
55     plt.show()
56     plt.hist(dead_num_list, bins=range(max(dead_num_list)))
57     plt.title("Histogram แสดงความถี่ของจำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุ")
58     plt.xlabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
59     plt.ylabel("จำนวนครั้งที่เกิดอุบัติเหตุ (ครั้ง)")
60     plt.show()
61
62     ## XY Scatter Plot
63     plt.scatter(patient_num_list, dead_num_list, color='red', alpha=0.3)
64     plt.title("กราฟแสดงความสัมพันธ์ของจำนวนผู้บาดเจ็บกับจำนวนผู้เสียชีวิตในอุบัติเหตุ")
65     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
66     plt.ylabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
67     plt.grid()
68     plt.show()

```

ภาคผนวกที่ 1 Source Code การวิเคราะห์ค่าสถิติพื้นฐาน

```

1 import csv
2 import numpy as np
3 import pandas as pd
4 from scipy.stats import norm
5 import matplotlib.pyplot as plt
6
7 plt.rcParams['font.family'] = 'kanit'
8
9 csv_path = "../datasets/ubatiehtuthaangthnn.csv"
10
11 with open(csv_path, encoding='utf-8') as csv_file:
12     csv_reader = csv.reader(csv_file)
13
14     dead_num_list = []
15     patient_num_list = []
16
17     accident_num = 0
18     for row in csv_reader:
19
20         if accident_num != 0:
21             vechicle, dead_num, patient_num = row[10], int(row[15]), int(row[16])
22
23             dead_num_list.append(dead_num)
24             patient_num_list.append(patient_num)
25
26         accident_num += 1
27
28     ## Data ##
29     dead_num_list = np.array(sorted(dead_num_list))
30     patient_num_list = np.array(sorted(patient_num_list))
31
32     ## Probability Density Function ##
33     pdf_dead = norm.pdf(dead_num_list, np.mean(dead_num_list), np.std(dead_num_list))
34     plt.plot(dead_num_list, pdf_dead, '-o')
35     plt.title("กราฟ Probability Density Function จำนวนผู้เสียชีวิตในอุบัติเหตุ")
36     plt.xlabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
37     plt.ylabel("Density")
38     plt.grid()
39     plt.show()
40
41     pdf_patient = norm.pdf(patient_num_list, np.mean(patient_num_list), np.std(patient_num_list))
42     plt.plot(patient_num_list, pdf_patient, '-o')
43     plt.title("กราฟ Probability Density Function จำนวนผู้บาดเจ็บในอุบัติเหตุ")
44     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
45     plt.ylabel("Density")
46     plt.grid()
47     plt.show()
48
49     ## Cumulative Probability Function ##
50     cdf_dead = norm.cdf(dead_num_list, np.mean(dead_num_list), np.std(dead_num_list))
51     plt.plot(dead_num_list, cdf_dead, '-o')
52     plt.title("กราฟ Cumulative Probability Function จำนวนผู้เสียชีวิตในอุบัติเหตุ")
53     plt.xlabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
54     plt.ylabel("Probability")
55     plt.grid()
56     plt.show()
57
58     cdf_patient = norm.cdf(patient_num_list, np.mean(patient_num_list), np.std(patient_num_list))
59     plt.plot(patient_num_list, cdf_patient, '-o')
60     plt.title("กราฟ Cumulative Probability Function จำนวนผู้บาดเจ็บในอุบัติเหตุ")
61     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
62     plt.ylabel("Probability")
63     plt.grid()
64     plt.show()

```

ภาคผนวกที่ 2 Source Code การวิเคราะห์ฟังก์ชันการแจกแจงความน่าจะเป็น และความน่าจะเป็นสะสม



```

1 import csv
2 import math
3 import numpy as np
4 import pandas as pd
5 import scipy.stats
6 import matplotlib.pyplot as plt
7
8 plt.rcParams['font.family'] = 'kanit'
9
10 csv_path = "../datasets/ubatiehtuthaangthnn.csv"
11
12 def CI_Mean(confidence_level,mean,sd,n):
13     z_value = scipy.stats.norm.ppf(1-((1-(confidence_level/100))/2))
14
15     return (mean-(z_value*(sd/math.sqrt(n))),mean+(z_value*(sd/math.sqrt(n))))
16
17 with open(csv_path, encoding='utf-8') as csv_file:
18     csv_reader = csv.reader(csv_file)
19
20     patient_num_list = []
21
22     accident_num = 0
23     for row in csv_reader:
24
25         if accident_num != 0:
26             vehicle, patient_num = row[10], int(row[16])
27
28             patient_num_list.append(patient_num)
29
30             accident_num += 1
31
32     ## Data ##
33     patient_num_list = np.array(sorted(patient_num_list))
34     patient_num_mean = np.mean(patient_num_list)
35     patient_num_sd = np.std(patient_num_list)
36
37     #Confidence Intreval of Meann: 90%
38     confidence_interval = CI_Mean(90,patient_num_mean,patient_num_sd,len(patient_num_list))
39     print("\n[Confidence Level 90%] Confidence Interval of Mean : {}".format(confidence_interval))
40
41     pdf_patient = scipy.stats.norm.pdf(patient_num_list,np.mean(patient_num_list),np.std(patient_num_list))
42     plt.plot(patient_num_list,pdf_patient,'-o')
43     plt.title("กราฟ Probability Density Function จำนวนผู้บาดเจ็บในอุบัติเหตุ")
44     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
45     plt.ylabel("Density")
46     plt.plot([confidence_interval[0],confidence_interval[0]],[0,max(pdf_patient)+0.01])
47     plt.plot([confidence_interval[1],confidence_interval[1]],[0,max(pdf_patient)+0.01])
48     plt.grid()
49     plt.show()
50
51     #Confidence Interval of Mean : 95%
52     confidence_interval = CI_Mean(95,patient_num_mean,patient_num_sd,len(patient_num_list))
53     print("\n[Confidence Level 95%] Confidence Interval of Mean : {}".format(confidence_interval))
54
55     pdf_patient = scipy.stats.norm.pdf(patient_num_list,np.mean(patient_num_list),np.std(patient_num_list))
56     plt.plot(patient_num_list,pdf_patient,'-o')
57     plt.title("กราฟ Probability Density Function จำนวนผู้บาดเจ็บในอุบัติเหตุ")
58     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
59     plt.ylabel("Density")
60     plt.plot([confidence_interval[0],confidence_interval[0]],[0,max(pdf_patient)+0.01])
61     plt.plot([confidence_interval[1],confidence_interval[1]],[0,max(pdf_patient)+0.01])
62     plt.grid()
63     plt.show()
64
65     #Confidence Interval of Mean : 99%
66     confidence_interval = CI_Mean(99,patient_num_mean,patient_num_sd,len(patient_num_list))
67     print("\n[Confidence Level 99%] Confidence Interval of Mean : {}".format(confidence_interval))
68
69     pdf_patient = scipy.stats.norm.pdf(patient_num_list,np.mean(patient_num_list),np.std(patient_num_list))
70     plt.plot(patient_num_list,pdf_patient,'-o')
71     plt.title("กราฟ Probability Density Function จำนวนผู้บาดเจ็บในอุบัติเหตุ")
72     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
73     plt.ylabel("Density")
74     plt.plot([confidence_interval[0],confidence_interval[0]],[0,max(pdf_patient)+0.01])
75     plt.plot([confidence_interval[1],confidence_interval[1]],[0,max(pdf_patient)+0.01])
76     plt.grid()
77     plt.show()
78

```

### ภาคผนวกที่ 3 Source Code การวิเคราะห์ช่วงค่าเฉลี่ยประชากร

```

1 import csv
2 import statistics
3 import pandas as pd
4 import matplotlib.pyplot as plt
5 import numpy as np
6 import scipy.stats
7
8 plt.rcParams['font.family'] = 'kanit'
9
10 csv_path = "../datasets/ubatiehtuthaangthnn.csv"
11
12 with open(csv_path, encoding='utf-8') as csv_file:
13     csv_reader = csv.reader(csv_file)
14
15     dead_num_list = []
16     patient_num_list = []
17
18     accident_num = 0
19     for row in csv_reader:
20
21         if accident_num != 0:
22             vehicle, dead_num, patient_num = row[10], int(row[15]), int(row[16])
23
24             dead_num_list.append(dead_num)
25             patient_num_list.append(patient_num)
26
27         accident_num += 1
28
29     dead_num_list = np.array(dead_num_list)
30     patient_num_list = np.array(patient_num_list)
31
32     ## Linear Regression Solve
33     (slope, intercept, r_value, p_value, std_err) = scipy.stats.linregress(patient_num_list, dead_num_list)
34
35     ## XY Scatter Plot
36     plt.scatter(patient_num_list, dead_num_list, color='red', alpha=0.3)
37     plt.plot(patient_num_list, slope*patient_num_list + intercept)
38     plt.title("กราฟแสดงความสัมพันธ์ของจำนวนผู้บาดเจ็บกับจำนวนผู้เสียชีวิตในอุบัติเหตุ")
39     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
40     plt.ylabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
41     plt.grid()
42     plt.show()
43
44     print("=====\n")
45     print("สมการ Linear Regression : Y = ({}*X) + {}".format(slope, intercept))
46     print("ค่าสัมประสิทธิ์สหสัมพันธ์ (r) : {}".format(r_value))
47     print("ค่าคลาดเคลื่อนมาตรฐาน : {}".format(std_err))
48     print("\n=====")

```

ภาคผนวกที่ 4 Source Code การวิเคราะห์ความสัมพันธ์ โดย Linear Regression