บทที่ 1
บทนำ

ความสำคัญและที่มาของปัญหาที่ทำการวิจัย

โดยทั่วไปการคำนวณค่าความน่าจะเป็นของเหตุการณ์ที่สนใจในปริภูมิตัวอย่างมักใช้อัตราส่วนของจำนวนสมาชิกในเหตุการณ์ที่สนใจและจำนวนสมาชิกในปริภูมิตัวอย่างนั้นโดยแต่ละสมาชิกภายใต้ปริภูมิตัวอย่างอันเป็นผลลัพธ์ที่เป็นไปได้ทั้งหมดของการทดลองสุ่มเกิดขึ้นด้วยความน่าจะเป็นที่เท่ากัน หากแต่ต้องดำเนินการด้วยวิธีซ้ำๆสำหรับค่าความน่าจะเป็นของแต่ละเหตุการณ์ซึ่งทำให้เกิดความยุ่งยากล่าช้าเพื่อทำให้เกิดประสิทธิภาพในการอธิบายค่าความน่าจะเป็นของทุกเหตุการณ์ที่เป็นไปได้ทั้งหมดอย่างเป็นระบบนักคณิตศาสตร์จึงมีการกำหนดฟังก์ชันที่แสดงความสัมพันธ์ระหว่างสมาชิกในปริภูมิตัวอย่างกับค่าของจำนวนจริง (Real number) ที่เรียกว่าตัวแปรสุ่ม และฟังก์ชันที่แสดงความสัมพันธ์ระหว่างแต่ละค่าของตัวแปรสุ่มกับความน่าจะเป็นของตัวแปรสุ่มนั้นที่เรียกว่า ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม (Probability Function) หรือเรียกอีกอย่างว่าการแจกแจงความน่าจะเป็น(distribution) การจัดทำระบบดังกล่าวนี้เป็นวิวัฒนาการก้าวสำคัญของทฤษฎีความน่าจะเป็นที่ทำให้การศึกษาความน่าจะเป็นของเหตุการณ์ที่สนใจภายใต้ขอบเขตเดียวกันหรือเกี่ยวเนื่องกันเป็นไปได้อย่างครบถ้วนสมบูรณ์และสามารถนำไปประยุกต์ได้อย่างกว้างขวางสำหรับกรณีของการทดลองสุ่มอื่นๆที่มีคุณลักษณะหรือคุณสมบัติที่เหมาะสมสอดคล้องกัน

การแจกแจงความน่าจะเป็นของแต่ละตัวแปรสุ่มมีหลากหลายรูปแบบขึ้นอยู่กับลักษณะของการทดลองสุ่มที่ทำให้เกิดตัวแปรสุ่มนั้น ๆ เช่น การทดลองที่ตัวแปรสุ่มมีโอกาสเกิดขึ้นได้เพียง 2 ค่า คือค่าที่ต้องการกับค่าที่ไม่ต้องการด้วยความน่าจะเป็นเท่ากับ p และ 1-p ตามลำดับ เรียกว่าการทดลองสุ่มลักษณะนี้ว่าการทดลองเบอร์นูลี ตัวแปรสุ่มที่เกิดจากการทดลองนี้เรียกว่าตัวแปรสุ่มเบอร์นูลี และเรียกการแจกแจงความน่าจะเป็นของตัวแปรสุ่มนี้ว่าการแจกแจงเบอร์นูลีซึ่งเป็นการแจกแจงไม่ต่อเนื่อง หรือ การทดลองที่ตัวแปรสุ่มมีได้ n ค่าที่แตกต่างกันแต่ละค่ามีความน่าจะเป็นเท่ากันและเท่ากับ 1/n แล้วเรียกว่าการทดลองสุ่มลักษณะนี้ว่าการทดลองสม่ำเสมอ (uniform distribution) ตัวแปรสุ่มที่เกิดจากการทดลองนี้เรียกว่าตัวแปรสุ่มสม่ำเสมอ และเรียกการแจกแจงความน่าจะเป็นของตัวแปรสุ่มนี้ว่าการแจกแจงสม่ำเสมอ (uniform distribution) ซึ่งเป็นการแจกแจงไม่ต่อเนื่อง เป็นต้น การได้มาซึ่งการแจกแจงความน่าจะเป็นของตัวแปรสุ่มลักษณะนี้ถือเป็นการได้มาโดยวิธีสังเกตจากการทดลองสุ่ม นอกจากการได้มาด้วยวิธีการสังเกตจากการทดลองสุ่มโดยตรงแล้วอาจจะได้จากการพัฒนาจากการแจกแจงความน่าจะเป็นที่มีอยู่แล้วก็ได้ เช่น การแจกแจงเรขาคณิตได้จากการพัฒนาแจกแจงทวินามลบ (Negative Binomial Random Variable)การแจกแจงทวินามลบเกิดจากการผสมระหว่างการแจกแจงปัวซงกับการแจกแจงแกมมา การแจกแจงลินเลย์วางนัยทั่วไป (Generalized Lindley Distribution)ได้จากการพัฒนาการแจกแจงลินเลย์(Lindley Distribution) การแจกแจงลินเลย์เกิดจากการผสมระหว่างการแจกแจงเอกโปเนนเชียลกับการแจกแจงแกมมา เป็นต้น

การสร้างหรือพัฒนาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มเป็นสาระสำคัญในทฤษฎีความน่าจะเป็นที่นักคณิตศาสตร์ได้ดำเนินการมาเป็นเวลาช้านาน ซึ่งจะเป็นประโยชน์ให้การคำนวณ

ความน่าจะเป็นของการเกิดเหตุการณ์ในหลากหลายรูปแบบทำได้ง่ายและรวดเร็วขึ้น และการได้มาซึ่งการแจกแจงความน่าจะเป็นของตัวแปรสุ่มแบบนี้ยังแสดงถึงความก้าวหน้าในการพัฒนาเนื้อหาสาระทางคณิตศาสตร์ที่เกี่ยวกับความน่าจะเป็นอีกด้วย และการสร้างหรือพัฒนาการแจกแจงความน่าจะเป็นของตัวแปรสุ่มมักจะดำเนินการไปพร้อมๆ กับการพัฒนาคุณสมบัติเชิงความน่าจะเป็นของตัวแปรสุ่มอีกด้วย ได้แก่ ฟังก์ชันความน่าจะเป็นของตัวแปรสุ่ม (probability function) ฟังก์ชันความน่าจะเป็นสะสม(cumulative probability function) ค่าเฉลี่ย (mean) ค่าความแปรปรวน (variance) ค่าสัมประสิทธิ์ความเบ้ (skewness coefficient) ค่าสัมประสิทธิ์ความโด่ง (kurtosis coefficient) ฟังก์ชันโมเมนต์เวียนบังเกิด (moment generation function) และฟังก์ชันลักษณะเฉพาะ (characteristic function)

สำหรับการศึกษาในครั้งนี้เป็นการสร้างการแจกแจงความน่าจะเป็นของตัวแปรสุ่มขึ้นมาใหม่ โดยการผสมกันระหว่างการแจกแจงทวินามลบ (Negative Binomial Distribution) กับการแจกแจงลินเลย์วางนัยทั่วไป (Generalized Lindley Distribution) โดยจะเรียกการแจกแจงแบบใหม่นี้ว่า การแจกแจงทวินามลบ - ลินเลย์วางนัยทั่วไป (Negative Binomial – Generalized Lindley Distribution) นับว่าเป็นการสร้างการแจกแจงแบบไม่ต่อเนื่องที่สำคัญให้เกิดขึ้นในวงวิชาการคณิตศาสตร์ได้อีกการแจกแจงหนึ่ง

วัตถุประสงค์

1. เพื่อสร้างฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มทวินามลบ - ลินเลย์วางนัยทั่วไป
2. เพื่อศึกษาคุณสมบัติเชิงความน่าจะเป็นของตัวแปรสุ่มที่มีการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไป
3. เพื่อศึกษาและเปรียบเทียบการประมาณค่าพารามิเตอร์ของการแจกแจงแบบทวินามลบ-ลินเลย์วางนัยทั่วไปโดยใช้วิธีภาวะน่าจะเป็นสูงสุด (maximum likelihoodmethod) และวิธีโมเมนต์ (moment method)
4. เพื่อศึกษาการแจกแจงแบบทวินามลบ-ลินเลย์วางนัยทั่วไปในเชิงประยุกต์กับข้อมูลไม่ต่อเนื่อง เช่นจำนวนผู้เสียชีวิตจากการเกิดอุบัติเหตุ จำนวนผู้ป่วย เป็นต้น

สมมุติฐาน

การแจกแจงทวินามลบ -ลินเลย์วางนัยทั่วไปสามารถประยุกต์กับ ข้อมูลไม่ต่อเนื่องได้

ขอบเขตการวิจัย

1. ทำการศึกษาคุณสมบัติเชิงความน่าจะเป็นที่เกี่ยวข้องกับตัวแปรสุ่มที่มีการแจกแจงแบบทวินามลบ-ลินเลย์วางนัยทั่วไป ดังต่อไปนี้

1.1 ฟังก์ชันความน่าจะเป็น

1.2 ฟังก์ชันความน่าจะเป็นสะสม

1.3 คุณสมบัติสถิติเชิงคณิตศาสตร์ซึ่งประกอบไปด้วย

1.3.1 ค่าเฉลี่ย (mean)

1.3.2 ค่าความแปรปรวน (variance)

1.3.3 ค่าสัมประสิทธิ์ความเบ้ (skewness coefficient)

1.3.4 ค่าสัมประสิทธิ์ความโด่ง (kurtosis coefficient)

1.4 ฟังก์ชันโมเมนต์เวียนบังเกิด (moment generation function) และฟังก์ชันลักษณะเฉพาะ (characteristic function)

1.5 การสร้างค่าตัวแปรสุ่มทวินามลบ-ลินเลย์วางนัยทั่วไป

1.6 การทดสอบภาวะรูปสนิทดี (maximum likelihood criterion) โดยใช้ไคสแควร์ (Chi-Square)

2. การประมาณค่าพารามิเตอร์ (parameter estimation) โดยใช้วิธีการประมาณค่าพารามิเตอร์ 2 วิธีคือวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์

2.1 จำลองข้อมูลตามสถานการณ์ที่กำหนดโดยใช้โปรแกรม R ของแต่ละขนาดตัวอย่างที่ใช้ในการศึกษาเท่ากับ 50,100,200 และ 500 โดยทำการจำลองซ้ำ 500 ครั้งในแต่ละสถานการณ์ที่กำหนดและประมาณค่าพารามิเตอร์โดยวิธีภาวะน่าจะเป็นสูงสุด และวิธีโมเมนต์ของแต่ละสถานการณ์ โดยใช้โปรแกรม R

2.2 สร้างค่าตัวแปรสุ่มตามสถานการณ์ดังกล่าวข้างต้นโดยทำการจำลองซ้ำ 500 ครั้งในแต่ละสถานการณ์ที่กำหนด

2.3 เปรียบเทียบประสิทธิภาพของวิธีการประมาณค่าพารามิเตอร์โดยพิจารณาจากค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองระหว่างค่าประมาณกับค่าจริง (mean square error:MSE)

3. นำการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไปมาศึกษากับข้อมูลจริงเพื่อแสดงการประยุกต์กับข้อมูลทางด้านอุบัติเหตุ ซึ่งจะทำการศึกษาดังนี้

3.1 การทดสอบภาวะสารูปสนิทดี

3.2 การประมาณค่าพารามิเตอร์ด้วยวิธีภาวะน่าจะเป็นสูงสุดและวิธีโมเมนต์

3.3 เปรียบเทียบประสิทธิภาพระหว่างวิธีภาวะน่าจะเป็นสูงสุดกับวิธีโมเมนต์

ประโยชน์ที่คาดว่าจะได้รับ

1. ได้การแจกแจงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ - ลินเลย์วางนัยทั่วไป
2. ได้ทราบถึงคุณสมบัติเชิงความน่าจะเป็นของตัวแปรสุ่มทวินามลบ-ลินเลย์วางนัยทั่วไป
3. ได้ทราบถึงผลการเปรียบเทียบวิธีการประมาณค่าพารามิเตอร์ของการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไปโดยใช้วิธีภาวะน่าจะเป็นสูงสุดและวิธีโมเมนต์
4. เป็นพื้นฐานในการประยุกต์ใช้กับข้อมูลที่ลักษณะการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไป
5. ใช้เป็นเอกสารอ้างอิงสำหรับผู้ที่สนใจตัวแปรสุ่มที่มีการแจกแจงทวินามลบ-ลินเลย์วางนัยทั่วไป