บทที่ 1

บทนำ

<u>ความเป็นมาและความสำคัญของปัญหา</u>

ความน่าจะเป็นที่คนที่ขณะนี้อายุ x จะมีชีวิตอยู่รอดต่อไปอีก 1 ปีข้างหน้าหรือ p_x เป็นองค์ประกอบที่สำคัญองค์ประกอบหนึ่งของการศึกษางานทางด้านคณิตศาสตร์ประกันชีวิต เพราะไม่ว่าจะคำนวณหาอัตราเบี้ยประกันภัย (Premium Rates) เงินสำรอง (Reserves) การวิเคราะห์สินทรัพย์ (Asset Share Analysis) การวิเคราะห์ค่าใช้จ่าย (Expense Analysis) รวมทั้งมูลค่าที่ริบไม่ได้ (Nonforfeiture Values) และอื่นๆ ก็ต้องใช้ค่า p_x มาเป็นส่วนประกอบสำคัญในการคำนวณทั้งสิ้น

ในการคำนวณค่า p_x นั้น โดยหลักการจะอาศัยการแจกแจงของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต (Future Life Time)

$p_x = P[T(X) > 1]$

เมื่อ T(X) คือ ระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต (The Future Life Time) ของคนอายุ x

ในทางปฏิบัติ การหาการแจกแจงดังกล่าวยากที่จะทำได้ เนื่องจากการหาระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคตนั้น จะต้องบันทึกตั้งแต่วันที่เริ่มศึกษาจนกระทั่งคนที่ถูกศึกษาได้เสียชีวิต ซึ่งต้องใช้เวลาและเสียค่าใช้จ่ายเป็นจำนวนมาก ดังนั้นเพื่อแก้ปัญหานี้จึงจำเป็นที่จะต้องเก็บข้อมูลเกี่ยวกับระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคตเท่าที่จะสามารถบันทึกได้ โดยกำหนดช่วงเวลาสนใจศึกษา ขึ้นมาเป็นช่วงเวลาที่จำกัด เช่น 1 ปี 5 ปี และ 10 ปี เป็นต้น ฉะนั้นข้อมูลที่ได้จะเป็นข้อมูลที่ไม่สมบูรณ์ (Incomplete Data) คือการที่ไม่ทราบค่าสังเกตของตัวแปรที่สนใจ ซึ่งแบ่งได้เป็นค่าสังเกตมีลักษณะเป็นข้อมูลที่ถูกตัดทิ้ง (Censored Data) เนื่องจากไม่สามารถบันทึกค่าจริงของข้อมูลได้ อันเนื่องมาจากได้ทำการทดลองเพียงชั่วระยะเวลาหนึ่งเมื่อหยุดทำการทดลองจึงไม่สามารถทราบค่าจริงของข้อมูลได้ และข้อมูลออกจากช่วงก่อนสิ้นสุดการศึกษา (Withdrawal) ตัวอย่างเช่น ในการประกันภัย การครบกำหนดอายุของกรมธรรม์ (Ender) การขาดอายุของ

กรมธรรม์ (Lapse) การเวนคืนกรมธรรม์ (Surrender) เป็นต้น ถือได้ว่าเป็นข้อมูลที่ออกจากกลุ่มก่อนสิ้นช่วงการศึกษา (Withdrawal) ทำให้ไม่สามารถทราบอายุยาวนานของผู้ถือกรมธรรม์นั้นได้

จากสภาพปัญหาดังกล่าว ได้มีผู้สนใจศึกษาปัญหาในการประมาณค่าความน่าจะเป็นที่จะอยู่รอดสำหรับข้อมูลที่ไม่สมบูรณ์ และได้สร้างทฤษฎีเพื่อประมาณค่าความน่าจะเป็นนี้ขึ้นมาคือ ดรอลีทเท (Drolette) และเชียง (Chiang) ผู้วิจัยจึงสนใจที่จะทำการศึกษาและเปรียบเทียบการประมาณค่า p_x จากข้อมูลที่ไม่สมบูรณ์ ซึ่งจะศึกษาในลักษณะของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต (Future Life Time) T ภายใต้การแจกแจง 2 แบบ คือ การแจกแจงแบบไวบูลล์ (Weibull Distribution) และการแจกแจงแบบกอมเพิรตซ์ (Gompertz Distribution) และภายใต้การแจกแจงของระยะเวลาการออกจากกลุ่มก่อนสิ้นช่วงศึกษา (W) 2 แบบ คือ การแจกแจงแบบสม่ำเสมอ (Uniform Distribution) และการแจกแจงแบบเบต้า (Beta Distribution) และใช้วิธีการประมาณทางสถิติ 4 วิธี ดังต่อไปนี้

1. วิธีทางคณิตศาสตร์ประกันภัย (The Actuarial Method)
2. ตัวประมาณ เอ (Estimator A)
3. ตัวประมาณ บี (Estimator B)
4. ตัวประมาณ ซี (Estimator C)

<u>วัตถุประสงค์ของการวิจัย</u>

เพื่อศึกษาและเปรียบเทียบวิธีการประมาณความน่าจะเป็นที่คนอายุ x ปี จะมีชีวิตอยู่รอดต่อไปอีก 1 ปี ข้างหน้า สำหรับข้อมูลแบบไม่สมบูรณ์ซึ่งจะเสนอตามวิธีดังกล่าวไว้แล้วข้างต้น

<u>สมมติฐานของการวิจัย</u>

ภายใต้ข้อมูลแบบไม่สมบูรณ์นี้ ตัวประมาณ ซี (Estimator C) จะให้ค่าประมาณใกล้เคียงกับค่าจริงมากกว่า วิธีทางคณิตศาสตร์ประกันภัย , ตัวประมาณ เอ และ ตัวประมาณ บี

ข้อตกลงเบื้องต้น

1. ในการวิจัยครั้งนี้ เวลาเริ่มต้นสำหรับค่าสังเกตมีลักษณะเป็นการศึกษาแบบติดตามผล (Follow-up Study) ซึ่งเป็นแบบที่เป็นไปได้ในทางปฏิบัติแบบหนึ่ง และช่วยให้ง่ายแก่การบันทึกและการวิเคราะห์ข้อมูล โดยการศึกษาลักษณะดังกล่าวนี้จะกำหนดให้ผู้ที่เข้ามาในช่วงเวลาที่สนใจเป็นจุดเดียวกัน คือ t = 0 และศึกษาลักษณะของข้อมูลที่ถูกตัดปลายทางขวา (Right Censoring) ซึ่งแต่ละคนจะถูกสังเกตจนกระทั่งเสียชีวิต หรือจนกระทั่งสิ้นสุดช่วงการศึกษา คือ t = 1 ขึ้นอยู่กับว่าสิ่งไหนจะเกิดขึ้นก่อน ส่วนผู้ที่ออกจากกลุ่มก่อนสิ้นช่วงการศึกษา ก็จะถูกสังเกตจนกระทั่งเสียชีวิต หรือจนกระทั่งสิ้นสุดช่วงการศึกษา คือ t = 1 เช่นกัน

2. การแจกแจงของค่าที่ถูกตัดและค่าที่ไม่ถูกตัด เป็นอิสระต่อกัน

3. ศึกษากรณีของประเภทค่าที่ถูกตัดทิ้ง เป็นการตัดทิ้งทางขวา

ขอบเขตการวิจัย

ในการวิจัยครั้งนี้จะทำการวิจัยภายใต้ขอบเขตดังนี้

1. จะทำการเปรียบเทียบการประมาณความน่าจะเป็นของคนอายุ X ปี จะอยู่รอดต่อไปอีก 1 ปี ข้างหน้า (p_x) โดยประมาณด้วยวิธี

1. วิธีทางคณิตศาสตร์ประกันภัย (The Actuarial Method)
2. ตัวประมาณ เอ (Estimator A)
3. ตัวประมาณ บี (Estimator. B)
4. ตัวประมาณ ซี (Estimator C)

2. ข้อมูลที่นำมาใช้ในการวิจัย มีการแจกแจง 2 แบบ สำหรับระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต คือ

2.1 การแจกแจงแบบไวบูลล์ (Weibull Distribution) ซึ่งมี k และ n เป็นพารามิเตอร์ โดยรายละเอียดของค่าพารามิเตอร์จะแสดงอยู่ในภาคผนวก และมีฟังก์ชันความหนาแน่นของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต ดังต่อไปนี้

$$f(t) = \begin{cases} kt^n \exp\left(-\frac{kt^{n+1}}{n+1}\right) & , t \geq 0 \ , \ k > 0 \ , \ n > 0 \\ 0 & , t < 0 \end{cases}$$

2.2 การแจกแจงแบบกอมเพริตซ์ (Gompertz Distribution) ซึ่งมี B และ c เป็นพารามิเตอร์ โดยรายละเอียดของค่าพารามิเตอร์จะแสดงอยู่ในภาคผนวก และมีฟังก์ชันความหนาแน่นของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต ดังต่อไปนี้

$$f(t) = \begin{cases} Bc^t \exp\left(-\frac{B(c^t-1)}{\ln c}\right) & , t \geq 0 \ , \ B>0 \ , \ c>1 \\ 0 & , t<0 \end{cases}$$

3. ข้อมูลที่นำมาใช้ในการวิจัยมีการแจกแจง 2 แบบ สำหรับระยะเวลาออกจากกลุ่มก่อนสิ้นช่วงศึกษา คือ

3.1 การแจกแจงแบบสม่ำเสมอ (Uniform Distribution) ในช่วง (0,1) และมีฟังก์ชันความหนาแน่นดังต่อไปนี้

$$f(t) = \begin{cases} 1 & , 0 \leq w \leq 1 \\ 0 & , \text{อื่นๆ} \end{cases}$$

3.2 การแจกแจงแบบเบต้า (Beta Distribution) ซึ่งมี α_1 และ α_2 เป็นพารามิเตอร์ โดยกำหนดให้ $\alpha_1 = 5$ และ $\alpha_2 = 1.5$ และมีฟังก์ชันความหนาแน่นดังต่อไปนี้

$$f(w) = \begin{cases} \frac{\Gamma(\alpha_1+\alpha_2)}{\Gamma(\alpha_1)\Gamma(\alpha_2)} w^{\alpha_1-1}(1-w)^{\alpha_2-1} & , 0<w<1 \ , \ \alpha_1 \text{ และ } \alpha_2 > 0 \\ 0 & , \text{อื่นๆ} \end{cases}$$

4. ช่วงอายุที่สนใจศึกษา คือ 25-65 ปี

5. ขนาดตัวอย่างที่นำมาศึกษามี 4 ระดับ คือ 300 , 500 , 700 และ 1,000 ตามลำดับ

6. ศึกษาเมื่อข้อมูลมีค่าสังเกตที่ออกจากช่วงที่สนใจ (Withdraw) เป็น 10% , 20% , 30% และ 40% ตามลำดับ

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

หอสมุดกลาง สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

7. การวิจัยครั้งนี้จำลองข้อมูลให้มีสถานการณ์ตามที่ต้องการศึกษาโดยใช้เทคนิคการจำลองแบบมอนติคาร์โล (Monte Carlo Simulation Technique) ทำการจำลองข้อมูลซ้ำๆ กัน จำนวน 2,000 ครั้ง ในแต่ละสถานการณ์

<u>เกณฑ์การตัดสินใจ</u>

ในการพิจารณาว่าการประมาณค่า p_x ด้วยวิธีใดใช้ได้ดีกว่า จะพิจารณาโดยการเปรียบเทียบค่าเฉลี่ยความคลาดเคลื่อนระหว่างค่าประมาณ p_x กับค่าจริง ในรูปของค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ (Mean Absolute Percentage Error , MAPE) วิธีการใดให้ค่าเฉลี่ยเปอร์เซ็นต์ความคลาดเคลื่อนสัมบูรณ์ที่ต่ำกว่า จะเป็นวิธีการประมาณที่ดีกว่า

<u>ประโยชน์ที่คาดว่าจะได้รับ</u>

1. เพื่อทราบถึงวิธีการที่เหมาะสมในการประมาณค่า p_x ให้มีความถูกต้องและมีความน่าเชื่อถือมากขึ้น และเป็นแนวทางให้นักคณิตศาสตร์และผู้ที่สนใจศึกษา เลือกใช้วิธีการประมาณค่า p_x ที่ได้ในงานวิเคราะห์ทางด้านคณิตศาสตร์ประกันชีวิต หรือสาขาประชากรศาสตร์ได้อย่างเหมาะสมในแต่ละสถานการณ์

2. เพื่อเป็นแนวทางให้ผู้ที่สนใจศึกษาได้ทำการวิจัยต่อไป

สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย