บทที่ 1

บทนำ

1.1 ที่มาและความสำคัญของปัญหา

โดยทั่วไปแล้วการแจกแจงแบบเอกซ์โปเนนเชียลเป็นการแจกแจงหนึ่งซึ่งเราจะนำมาใช้เป็นตัวแบบในการศึกษาเกี่ยวกับระยะเวลาที่ต้องรอคอยจนกระทั่งเหตุการณ์ที่เราสนใจศึกษาเกิดขึ้น เช่น การศึกษาเกี่ยวกับอายุการใช้งานของหลอดไฟฟ้าตราหนึ่ง หรือ การศึกษาเกี่ยวกับระยะเวลาที่ใช้ในการรักษาโรคของยาชนิดหนึ่ง เป็นต้น

จากคุณสมบัติของการแจกแจงแบบเอกซ์โปเนนเชียลดังกล่าวนี้ได้มีนักสถิติหลายท่านทำการศึกษาเกี่ยวกับระยะเวลาที่ต้องรอคอยจนกระทั่งเหตุการณ์ที่สนใจเกิดขึ้นซึ่งอยู่ในสถานการณ์เดียวกันของประชากรที่มีมากกว่า 1 กลุ่ม เช่น การศึกษาเกี่ยวกับอายุการใช้งานของเครื่อง-จักรอุตสาหกรรมชนิดหนึ่งในหลาย ๆ ตรา(ชุด) และมีข้อตกลงเบื้องต้นว่าเครื่องจักรที่นำมาศึกษานั้นจะต้องเป็นเครื่องจักรที่ใช้งานอยู่ในลักษณะเดียวกัน และมีอายุการใช้งานโดยเฉลี่ยอยู่ในช่วงเวลาเดียวกัน หรือ การศึกษาเกี่ยวกับประสิทธิภาพของยาชนิดหนึ่งในหลาย ๆ ฉลาก (ชื่อ) ที่ใช้ในการรักษาโรคชนิดเดียวกัน โดยที่การแจกแจงซึ่งเหมาะสมที่จะนำมาใช้เป็นตัวแบบในการวิเคราะห์ปัญหาในลักษณะดังกล่าวนี้คือ **การแจกแจงแบบเอกซ์โปเนนเชียล 2 พารามิเตอร์**[1] ซึ่งมีพารามิเตอร์แสดงตำแหน่ง (location parameter) และพารามิเตอร์แสดงสเกล (scale parameter) พารามิเตอร์แสดงตำแหน่งในที่นี้หมายถึง ระยะเวลาที่ต่ำที่สุด หรือ ระยะเวลาที่สั้นที่สุด (minimum time) หรือระยะเวลาในการรับประกัน (guarantee time) ก่อนที่จะ

---

[1] Singh.N.H. "The Likelihood Ratio Test for the Equality of Location Parameter of $k \geq 2$ Exponential Distribution," <u>Technometric</u> 25 (1983): 193-195.

เกิดความเสียหายแก่สิ่งที่เรานำมาศึกษานั้น ๆ เป็นครั้งแรก และพารามิเตอร์แสดงสเกล หมายถึง ระยะเวลาเฉลี่ยของอายุการใช้งานของสิ่งที่เรานำมาศึกษาซึ่งวัดได้หลังจากเกิดความเสียหายเป็นครั้งแรกแก่สิ่งที่เรานำมาศึกษานั้น ๆ และโดยทั่วไปแล้วในการศึกษาเกี่ยวกับข้อมูลอายุ (lifetime data) จะกำหนดให้สิ่งที่นำมาศึกษานั้นมีอายุการใช้งานโดยเฉลี่ยอยู่ในช่วงเวลาเดียวกัน

จากข้อกำหนดดังกล่าวผู้วิจัยจึงสนใจที่จะทำการศึกษาเปรียบเทียบการเท่ากันของพารา-มิเตอร์แสดงตำแหน่ง ซึ่งหมายถึงการเปรียบเทียบระยะเวลาที่ต่ำที่สุดของสิ่งที่เรานำมาศึกษาจนกระทั่งเหตุการณ์ที่เราสนใจเกิดขึ้น ในที่นี้จะพิจารณาเปรียบเทียบระหว่าง k ประชากรโดยที่ k ประชากรที่ทำการศึกษานี้อาจพิจารณาได้ในกรณีตัวอย่าง เช่น ในทางอุตสาหกรรมอาจเปรียบเทียบการเท่ากันของอายุการใช้งานที่ต่ำที่สุดของเครื่องจักร k ชนิดที่ใช้งานอย่างเดียวกัน หรือในทางการแพทย์อาจจะทำการเปรียบเทียบระยะเวลาที่สั้นที่สุดซึ่งใช้ในการรักษาโรคของยา k ชนิดที่ใช้ในการรักษาโรคอย่างเดียวกัน เป็นต้น ในการศึกษากรณีดังกล่าวนี้นักสถิติหลายท่านได้ทำการศึกษาและคิดตัวสถิติขึ้นเพื่อทดสอบการเท่ากันของพารามิเตอร์แสดงตำแหน่งพอสรุปได้ดังนี้

ในปี ค.ศ.1953 Epstien และ Tsao เสนอตัวสถิติทดสอบ LR (Likelihood Ratio Test Statistics) เพื่อทดสอบการเท่ากันของพารามิเตอร์แสดงตำแหน่งระหว่างประชากร 2 กลุ่ม แต่วิธีการดังกล่าวมีข้อจำกัดในการคำนวณที่ซับซ้อนและในการนำไปใช้ กล่าวคือใช้ได้กับประชากร 2 กลุ่มเท่านั้น ในปี ค.ศ.1963 Hogg และ Tanis เสนอตัวสถิติทดสอบ IP (Iterated Procedure Test Statistics) ซึ่งมีแนวความคิดมาจากวิธี LR ที่ Epstien และ Tsao เสนอไว้ Hogg และ Tanis ปรับปรุงค่าของตัวสถิติทดสอบ IP ให้มีการแจกแจงลู่เข้าสู่การแจกแจงเอฟ ซึ่งสามารถหาค่าวิกฤตได้จากตารางการแจกแจงเอฟโดยทั่วไป นอกจากนี้ยังได้ปรับปรุงตัวสถิติทดสอบเพื่อใช้ทดสอบในกรณีที่มีประชากรมากกว่า 2 กลุ่ม ($k \geq 2$) ในปี ค.ศ.1981 M.L.Tiku เสนอตัวสถิติทดสอบ TIKU (Tiku's Test Statistics) เพื่อใช้ทดสอบการเท่ากันของพารามิเตอร์แสดงตำแหน่งระหว่างประชากร 2 กลุ่ม ในปี ค.ศ.1983 Singh เสนอตัวสถิติทดสอบ Modified LR ทดสอบการเท่ากันของพารามิเตอร์แสดงตำแหน่งในกรณีที่มีประชากรมากกว่า 2 กลุ่ม ($k \geq 2$) ในปี ค.ศ.1985

Kambo และ Awad นำตัวสถิติทดสอบ TIKU มาทดสอบในกรณีที่มีประชากรมากกว่า 2 กลุ่ม ทั้งในกรณีที่มีค่าสังเกตสมบูรณ์และในกรณีที่ค่าสังเกตมีค่าถูกตัดทิ้ง ในส่วนของการวิจัยนั้น ปี ค.ศ. 1983 M.L.Tiku นำตัวสถิติทดสอบ LR,IP และ TIKU มาเปรียบเทียบอำนาจการทดสอบในกรณีที่มีประชากร 2 กลุ่ม พบว่าตัวสถิติทดสอบ TIKU มีอำนาจการทดสอบสูงกว่าตัวสถิติทดสอบตัวอื่นแทบทุกกรณีทั้งในกรณีที่มีค่าสังเกตสมบูรณ์และค่าสังเกตไม่สมบูรณ์

จากผลการวิจัยดังกล่าวนี้ผู้วิจัยจึงสนใจทำการศึกษาเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบ 3 ตัวคือ

1) LR

2) IP

3) TIKU

โดยจะทำการทดสอบการเท่ากันของพารามิเตอร์แสดงตำแหน่งระหว่างประชากร 2 กลุ่มและมากกว่า 2 กลุ่มภายใต้การวิเคราะห์ข้อมูลทั้งหมด (complete data) และการวิเคราะห์ข้อมูลซึ่งมีข้อมูลถูกตัดทิ้ง (censored data)

1.2 วัตถุประสงค์ของการวิจัย

เพื่อศึกษาและเปรียบเทียบอำนาจการทดสอบของตัวสถิติทดสอบ 3 ตัวข้างบนที่ใช้ทดสอบการเท่ากันของพารามิเตอร์แสดงตำแหน่งของการแจกแจงแบบเอกซ์โปเนนเชียล 2 พารามิเตอร์

1.3 สมมุติฐานของการวิจัย

ตัวสถิติทดสอบ TIKU จะให้อำนาจการทดสอบสูงสุดภายใต้ข้อมูลในสถานการณ์ต่าง ๆ ตามหลักการของตัวสถิติทดสอบนี้ซึ่งอยู่ในหัวข้อที่ 2.3 ของบทที่ 2

1.4 ข้อตกลงเบื้องต้น

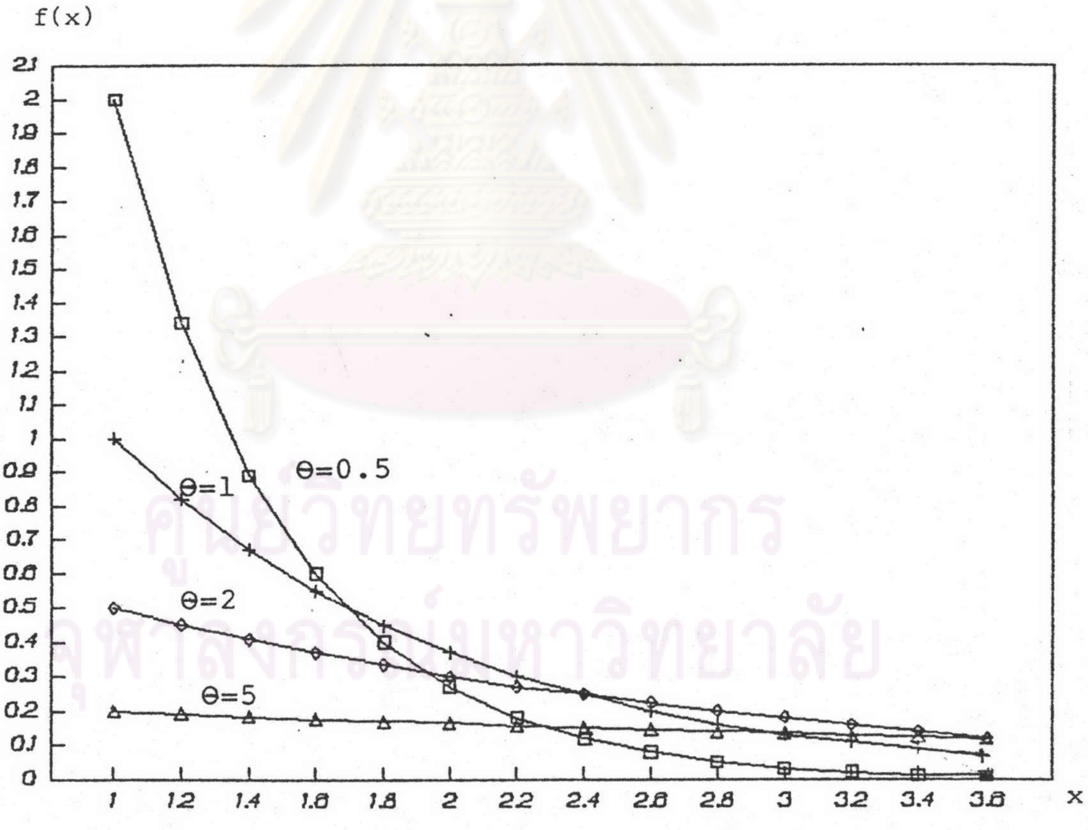
ในการวิจัยในครั้งนี้จะใช้การแจกแจงแบบเอกซ์โปเนนเชียล 2 พารามิเตอร์ โดยที่ฟังก์ชันความหนาแน่นอยู่ในรูปของ

$$f(x_j) = \begin{cases} \dfrac{1}{\theta}\exp\left(-\dfrac{(x-\beta_j)}{\theta}\right) & ,x_j \geq \beta_j,\ \theta > 0 \\ & ,j = 1,2,\ldots,k \\ 0 & \text{อื่น ๆ} \end{cases}$$

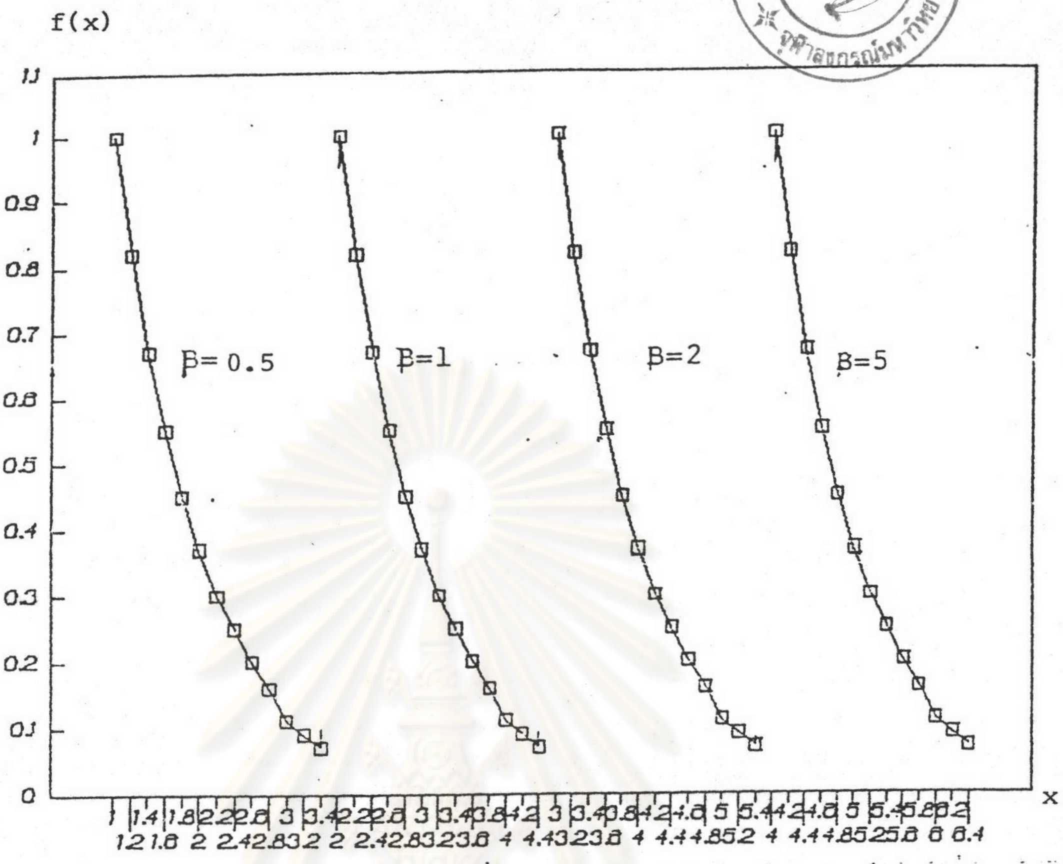
เมื่อ β_j เป็นพารามิเตอร์แสดงตำแหน่ง (location parameter)

และ θ เป็นพารามิเตอร์แสดงสเกล (scale parameter)

รูปที่ 1.1 จะแสดงภาพของฟังก์ชันความหนาแน่นข้างบนในกรณีที่ค่า β=1 และθ= 0.5,1,2 และ 5 ส่วนรูปที่ 1.2 จะแสดงภาพของฟังก์ชันความหนาแน่นในกรณีที่ค่า θ=1 และ β=0.5,1,2 และ 5 ตามลำดับ



รูปที่ 1.1 เส้นโค้งของการแจกแจงแบบเอกซ์โปเนนเชียล 2 พารามิเตอร์ เมื่อ β=1 และ θ= 0.5,1,2 และ 5



รูปที่ 1.2 เส้นโค้งของการแจกแจงแบบเอกซ์โปเนนเชียล 2 พารามิเตอร์

เมื่อ β=0.5,1,2,5 และ θ=1

1.5 ขอบเขตของการวิจัย

ในการวิจัยในครั้งนี้จะทำการวิจัยภายใต้ขอบเขตดังนี้

1) ข้อมูลที่นำมาวิจัยในครั้งนี้มีการแจกแจงแบบเอกซ์โปเนนเชียล 2 พารามิเตอร์โดยมี β เป็นพารามิเตอร์แสดงตำแหน่ง และ θ เป็นพารามิเตอร์แสดงสเกล

2) ผู้วิจัยสนใจศึกษาเมื่อค่าเฉลี่ยของประชากร (θ) เป็น 0.5,1,2,5 ตามลำดับ*

---

* ผู้วิจัยทดลองศึกษาเมื่อค่าเฉลี่ยของประชากร (θ) มีค่ามากกว่า 5 โดยพิจารณาจากเส้นโค้งของการแจกแจงดังกล่าวพบว่าลักษณะเส้นโค้งไม่แตกต่างกันกับเมื่อกำหนด θ = 5

3) ขนาดตัวอย่างที่นำมาศึกษามี 3 ระดับคือ 10,15 และ 20 ตามลำดับ*

4) ระดับความเชื่อมั่นมี 2 ระดับคือ $\alpha = 0.01$ และ 0.05

5) กำหนดค่าพารามิเตอร์แสดงตำแหน่งของประชากร (β) = 1**

6) ผู้วิจัยศึกษาเมื่อข้อมูลมีค่าสังเกตที่สมบูรณ์ และมีค่าที่ถูกตัดทิ้งทางขวา =10%, 20% และ 30% ตามลำดับ ซึ่งค่าที่ถูกตัดทิ้งทางขวาของข้อมูลจะพบมากในทางปฏิบัติ เช่น ข้อมูลที่เกี่ยวกับอายุ และข้อมูลทางการแพทย์ เป็นต้น

7) ผู้วิจัยกำหนดจำนวนกลุ่มประชากรที่นำมาศึกษาในครั้งนี้เป็น 2,3 และ 5 กลุ่มประชากรตามลำดับโดยที่ $k \geq 5$ จัดว่าเป็นกลุ่มประชากรขนาดใหญ่[2]

สถานการณ์ที่ผู้วิจัยนำมาศึกษาในครั้งนี้มีทั้งสิ้น = 4x3x2x4x3 = 288 กรณี

1.6 เกณฑ์การตัดสินใจ

เกณฑ์การตัดสินใจว่าตัวสถิติทดสอบใดให้อำนาจการทดสอบสูงที่สุดจะพิจารณาเป็น 2 ขั้นตอน ภายใต้สมมุติฐาน

$H_o : \beta_1=\beta_2=\ldots=\beta_k=1$

เทียบกับ $H_1 : \beta_1\neq\beta_2\neq\ldots\neq\beta_k$ (อย่างน้อยที่สุด 2 ประชากรมีค่าพารามิเตอร์แสดงตำแหน่งไม่เท่ากัน)

---

* ในทางปฏิบัติการทดลองเกี่ยวกับข้อมูลอายุขนาดตัวอย่างที่ใช้โดยส่วนใหญ่มักจะกำหนดให้มีขนาดไม่ใหญ่มากนัก ทั้งนี้เพื่อเป็นการประหยัดเวลาและค่าใช้จ่ายในการทดลอง นอกจากนี้ยังเพื่อสะดวกในการคำนวณค่าต่าง ๆ อีกด้วย

** ผู้วิจัยทดลองศึกษาเมื่อค่าพารามิเตอร์แสดงตำแหน่ง (β) มีค่าเปลี่ยนไปโดยที่ค่าพารามิเตอร์แสดงสเกล (θ) คงที่ (รูปที่ 1.4.2) พบว่าลักษณะเส้นโค้งที่ได้ไม่แตกต่างกัน

[2] Kambo and Awad F.M."Testing Equality of Location Parameter of k Exponential Distribution`" Commun Statist Theory meth 14(3) (1985): 567-586.

1) พิจารณาจากความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 (Type 1 Error) ของการทดสอบในแต่ละสถานการณ์ โดยใช้เกณฑ์ของ Cochran และ Bradley โดยที่เกณฑ์ของ Cochran จะพิจารณาจากค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 กล่าวคือ ถ้าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าอยู่ในช่วง [0.007,0.015] และ [0.04,0.06] ณ ระดับนัยสำคัญ 0.01 และ 0.05 ตามลำดับ จะถือว่าตัวสถิติทดสอบนั้นสามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ ส่วนเกณฑ์ของ Bradley จะพิจารณาหลักเกณฑ์คล้ายกับเกณฑ์ของ Cochran โดยพิจารณาว่า ถ้าค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 มีค่าอยู่ในช่วง [0.005,0.015] และ [0.025,0.075] ณ ระดับนัยสำคัญ 0.01 และ 0.05 ตามลำดับ จะถือว่าตัวสถิติทดสอบนั้นควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้

2) พิจารณาค่าอำนาจการทดสอบ โดยจะเปรียบเทียบเฉพาะตัวสถิติทดสอบที่สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้เท่านั้น

1.7 คำจำกัดความ

1) ความคลาดเคลื่อนประเภทที่ 1 (type I error) คือความคลาดเคลื่อนที่เกิดจากการปฏิเสธสมมุติฐานว่าง (H_o) เมื่อสมมุติฐานว่างนั้นเป็นจริง

2) ความคลาดเคลื่อนประเภทที่ 2 (type II error) คือความคลาดเคลื่อนที่เกิดจากการยอมรับสมมุติฐานว่าง (H_o) เมื่อสมมุติฐานว่างนั้นไม่จริง

3) อำนาจการทดสอบ (power of test) คือความน่าจะเป็นที่จะปฏิเสธสมมุติฐานว่างเมื่อสมมุติฐานว่างนั้นไม่จริง

1.8 ประโยชน์ที่คาดว่าจะได้รับ

เพื่อเป็นแนวทางในการเลือกใช้ตัวสถิติทดสอบที่เหมาะสมแก่ผู้วิจัยในการวิเคราะห์ปัญหาที่เกี่ยวข้องกับข้อมูลอายุ (lifetime data)