บทที่ 5

สรุปผลการวิเคราะห์และข้อเสนอแนะ

ในการเปรียบเทียบประสิทธิภาพของตัวประมาณค่าพารามิเตอร์ α และ β ในการวิเคราะห์ความถดถอย เมื่อตัวแปรตามบางค่ามีค่าขาดหาย โดยใช้วิธีการกำลังสองต่ำสุด วิธีการของมิลเล่อร์ วิธีการของบัคเลย์และเจมส์ สามารถสรุปผลการวิเคราะห์ได้ดังนี้

5.1 สรุปผลการวิเคราะห์

5.1.1 วิธีการกำลังสองต่ำสุด จะมีประสิทธิภาพในการประมาณค่า α และ β ได้ดีที่สุดเกือบทุกเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย ขนาดตัวอย่างและการแจกแจงของค่าขาดหายที่ศึกษา ยกเว้นกรณีในข้อ 2 และการประมาณค่าพารามิเตอร์ α, β จะดีขึ้นเมื่อเปอร์เซ็นต์เฉลี่ยของค่าขาดหายมีค่าต่ำ ๆ ในทุกขนาดตัวอย่าง และเมื่อขนาดตัวอย่างเพิ่มขึ้น การประมาณค่าพารามิเตอร์จะดีขึ้นด้วย และการประมาณค่าพารามิเตอร์ เมื่อค่าขาดหายเป็นฟังก์ชันเชิงเส้นกับค่าความคลาดเคลื่อนจะดีกว่าการแจกแจงแบบอื่น ๆ

5.1.2 วิธีการของบัคเลย์และเจมส์ จะมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ได้ดีที่สุด เมื่อค่าขาดหายมีการแจกแจงแบบแกมม่า ในกรณีดังต่อไปนี้คือ จะมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ α และ β ที่เปอร์เซ็นต์เฉลี่ยของ ค่าขาดหาย และขนาดตัวอย่างดังนี้คือ ที่เปอร์เซ็นต์ของค่าขาดหายเท่ากับ 5% เมื่อขนาดตัวอย่างเท่ากับ 20 และ 50 ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายเท่ากับ 5% ถึง 10% เมื่อขนาดตัวอย่างเท่ากับ 60 ที่เปอร์-เซ็นต์เฉลี่ยของค่าขาดหายเท่ากับ 30% ถึง 50% เมื่อขนาดตัวอย่างเท่ากับ150 และจะมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ α ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายเท่ากับ 10% เมื่อขนาดตัวอย่างเท่ากับ 20 และจะมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ β ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายและขนาดตัวอย่างดังนี้คือ ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายเท่ากับ 10% ถึง 30% เมื่อขนาดตัวอย่างเท่ากับ 50 ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายเท่ากับ 15% ถึง 50% เมื่อขนาดตัวอย่างเท่ากับ 60 ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหาย เท่ากับ 10% ถึง 50% เมื่อขนาดตัวอย่างเท่ากับ 100 ที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายเท่ากับ 5% ถึง 25% และ 60% เมื่อขนาดตัวอย่างเท่ากับ 150 และในกรณีที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายมีค่าสูง

จะให้ค่าประมาณที่ผิดปกติไปมาก ทั้งในการประมาณค่าพารามิเตอร์ α และ β

5.1.3 วิธีการของมิลเลอร์ จะมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ α และ β ต่ำกว่าวิธีการกำลังสองต่ำสุดในทุกเงื่อนไขของการศึกษา แต่จะมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ ดีกว่าวิธีการของบัคเลย์และเจมส์ที่เปอร์เซนต์เฉลี่ยของค่าขาดหายมีค่าสูง ๆ

การสรุปผลสามารถแสดงให้เห็นชัดมากขึ้น จากตารางต่อไปนี้

ตารางที่ 5.1 แสดงตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่ขนาดตัวอย่างเท่ากับ 10 จำแนกตามเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย และการแจกแจงของค่าขาดหาย

% เฉลี่ย / การแจกแจง	5 %		10 %		15 %		20 %		25 %		30 %		50 %		60 %		70 %		80 %	
	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)
ยูนิฟอร์ม	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
แกมม่า	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
ปกติ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
เชิงเส้น	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS

ตารางที่ 5.2 แสดงตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่ขนาดตัวอย่างเท่ากับ 20 จำแนกตามเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย และการแจกแจงของค่าขาดหาย

% เฉลี่ย / การแจกแจง	5 %		10 %		15 %		20 %		25 %		30 %		50 %		60 %		70 %		80 %	
	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)
ยูนิฟอร์ม	LS	LS	BJ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
แกมม่า	BJ	BJ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
ปกติ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
เชิงเส้น	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS

ตารางที่ 5.3 แสดงตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่ขนาดตัวอย่างเท่ากับ 50
จำแนกตามเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย และการแจกแจงของค่าขาดหาย

% เฉลี่ย / การแจกแจง	5 %		10 %		15 %		20 %		25 %		30 %		50 %		60 %		70 %		80 %	
	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)
ยูนิฟอร์ม	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
แกมม่า	BJ	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	LS	LS	LS	LS	LS	LS	LS
ปกติ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
เชิงเส้น	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS

ตารางที่ 5.4 แสดงตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่ขนาดตัวอย่างเท่ากับ 60
จำแนกตามเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย และการแจกแจงของค่าขาดหาย

% เฉลี่ย / การแจกแจง	5 %		10 %		15 %		20 %		25 %		30 %		50 %		60 %		70 %		80 %	
	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)
ยูนิฟอร์ม	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
แกมม่า	BJ	BJ	BJ	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	LS	LS	LS	LS	LS
ปกติ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
เชิงเส้น	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS

ตารางที่ 5.5 แสดงตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่ขนาดตัวอย่างเท่ากับ 100 จำแนกตามเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย และการแจกแจงของค่าขาดหาย

% เฉลี่ย / การแจกแจง	5 %		10 %		15 %		20 %		25 %		30 %		50 %		60 %		70 %		80 %	
	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)
ยูนิฟอร์ม	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
แกมม่า	LS	LS	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	LS	LS	LS	LS	LS
ปกติ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
เชิงเส้น	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS

ตารางที่ 5.6 แสดงตัวประมาณที่มีค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุด ที่ขนาดตัวอย่างเท่ากับ 150 จำแนกตามเปอร์เซ็นต์เฉลี่ยของค่าขาดหาย และการแจกแจงของค่าขาดหาย

% เฉลี่ย / การแจกแจง	5 %		10 %		15 %		20 %		25 %		30 %		50 %		60 %		70 %		80 %	
	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)	MSE(A)	MSE(B)
ยูนิฟอร์ม	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS
แกมม่า	LS	BJ	LS	BJ	LS	BJ	LS	BJ	LS	BJ	BJ	BJ	BJ	BJ	LS	BJ	LS	LS	LS	LS
ปกติ	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	BJ	LS	LS	LS	LS	LS	LS	LS
เชิงเส้น	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS	LS

5.1.4 จากการวิจัยสามารถสรุปผลเพื่อเป็นแนวทางในการเลือกใช้ตัวประมาณได้ดังนี้ เมื่อเกิดปัญหาตัวแปรบางค่ามีค่าขาดหาย ในการวิเคราะห์ความถดถอย วิธีการกำลังสองต่ำสุด เมื่อทำการวิเคราะห์ข้อมูลรวมทั้งข้อมูลที่ขาดหายและข้อมูลที่ไม่ขาดหาย ก็ยังคงมีประสิทธิภาพในการประมาณค่าพารามิเตอร์ α และ β ได้ดีกว่าวิธีการของมิลเลอร์ และวิธีการของบัคเลย์และเจมส์ โดยที่ให้ค่าความคลาดเคลื่อนกำลังสองเฉลี่ยต่ำสุดกว่าทั้งสองวิธี แต่วิธีการกำลังสองต่ำสุด ก็ยังคงเป็นตัวประมาณที่เอนเอียงอยู่

ในกรณีที่เปอร์เซ็นต์เฉลี่ยของค่าขาดหายมีค่าต่ำ ๆ ประมาณ 5% ถึง 15% วิธีการของบัคเลย์และเจมส์ จะให้ค่าประมาณที่เป็นค่าไม่เอนเอียงของพารามิเตอร์ แต่จะให้ค่าความคลาดเคลื่อนกำลังเฉลี่ยสูงกว่าวิธีการกำลังสองต่ำสุด ซึ่งในกรณีนี้ถ้าต้องการตัวประมาณที่ไม่เอนเอียง ก็สามารถจะเลือกใช้วิธีการของบัคเลย์และเจมส์ในการประมาณค่าได้

5.2 ข้อเสนอแนะ

5.2.1 ในการศึกษาครั้งนี้ ได้ทำการศึกษาเมื่อค่าขาดหายมีการแจกแจงเป็นแบบยูนิฟอร์ม แบบแกมม่า แบบปกติ และค่าขาดหายเป็นฟังก์ชั่นเชิงเส้นกับค่าความคลาดเคลื่อน เมื่อค่าความคลาดเคลื่อนมีการแจกแจงแบบปกติ เท่านั้น ดังนั้นในการศึกษาครั้งต่อไป อาจจะศึกษาเมื่อกรณีที่ค่าขาดหายมีการแจกแจงในรูปแบบอื่น

5.2.2 ในการศึกษาครั้งนี้ได้ทำการศึกษาเฉพาะกรณีของความถดถอยเชิงเส้นอย่างง่าย (Simple Linear Regression) เท่านั้น ดังนั้นในการศึกษาครั้งต่อไปอาจจะทำการศึกษาในกรณีของ Multiple Linear Regression

5.2.3 อาจจะทำการศึกษาวิธีการของ

5.2.3.1 Koul, Susarla, Van Ryzin (1981a)

5.2.3.2 Koul, Susarla, Van Ryzin (1981b)

5.2.3.3 Koul, Susarla, Van Ryzin (1981c)

5.2.3.4 Friedman, Stuetzle (1981)

5.2.3.5 Schmee, Hahn (1978)

5.2.3.6 Chatterjee, Mcleish (1981)

ในการศึกษาครั้งต่อไป

5.2.4 ในการเกิดปัญหาตัวแปรมีค่าขาดหายนี้ นอกจากจะเกิดปัญหานี้ในการวิเคราะห์ความถดถอยแล้วยังอาจเกิดในการวิเคราะห์ด้านอื่นๆ เช่น ในการทดสอบสมมุติฐานเกี่ยวกับ 2 ประชากร ซึ่งมีวิธีการที่จะใช้ในการแก้ปัญหาได้คือ วิธีการของ Gehan Test, วิธีการของ Mantel-Hacnszel Test, วิธีการของ Tarone-Ware Class of Tests วิธีการของ Efron Test และในการทดสอบสมมุติฐานเกี่ยวกับ k ประชากร ซึ่งมีวิธีการที่จะใช้ในการแก้ปัญหาคือ วิธีการของ Generalized Gehan Test (Breslow) วิธีการของ Generalize Mantel-Haenszel Test (Trone and Ware)[1] ซึ่งเป็นวิธีการที่น่าสนใจสำหรับการศึกษาครั้งต่อไป

---

[1]ดูรายละเอียดจาก

Rupert G. Miller. Survival Analysis. (New York : John Wiley and Sons Inc., 1981), pp. 81-118.