หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

บทที่ 5

สรุปผลการวิจัยและข้อเสนอแนะ

ในการวิจัยครั้งนี้ต้องการศึกษาเปรียบเทียบค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าพยากรณ์ที่ได้จากการประมาณค่าที่สูญหายของตัวแปรตามทั้ง 5 วิธี ในการวิเคราะห์การถดถอยเชิงเส้นพหุ เมื่อตัวแปรตามมีค่าสูญหายแบบสุ่ม โดยการประมาณค่าสูญหายของตัวแปรตามด้วยวิธีการ 5 วิธี ได้แก่ วิธีสูญหาย , วิธีค่าเฉลี่ย , วิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์

วิธีการดำเนินการวิจัยครั้งนี้ ใช้วิธีการจำลองแบบการทดลองด้วยเทคนิคมอนติคาร์โล ทำงานด้วยเครื่องคอมพิวเตอร์ AMDAHL 5860 โดยใช้ภาษาฟอร์แทรน (FORTRAN) เพื่อสร้างข้อมูลให้มีลักษณะตามแผนการทดลองที่กำหนด และกำหนดให้เครื่องคอมพิวเตอร์ ทำการทดลองซ้ำ ๆ กัน 200 รอบในแต่ละสถานการณ์

<u>สรุปผลการวิจัย</u>

ผลสรุปจากการวิเคราะห์ค่ารากที่สองของค่าเฉลี่ยของความคลาดเคลื่อนกำลังสองของค่าพยากรณ์สามารถจำแนกได้เป็น 2 ลักษณะดังนี้

1. <u>เมื่อลักษณะของตัวแปรอิสระเป็นรูปแบบที่ 1 และ 2 ซึ่งเป็นข้อมูลอนุกรมเวลา</u>

เมื่อขนาดตัวอย่างมีขนาดเล็กมาก (NM) = 10 สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 5 และสัดส่วนการสูญหายของตัวแปรตาม (PM) = 10% - 50% วิธีสูญหายเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ไม่แตกต่างจากวิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ แต่เนื่องจากวิธีสูญหายเป็นวิธีที่ง่ายในการคำนวณ จึงน่าจะใช้วิธีสูญหายในทางปฏิบัติ สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 5 และสัดส่วนการสูญหายของตัวแปรตาม (PM) = 60% - 70% วิธีการของฮันท์เป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์

ต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ และสำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน(σ) = 10 - 25 และสัดส่วนการสูญหายของตัวแปรตามทุกระดับ วิธีค่าเฉลี่ยเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ

เมื่อขนาดตัวอย่างมีขนาดเล็ก (NM) = 20 สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนทุกค่า และสัดส่วนการสูญหายของตัวแปรตาม (PM) = 10% - 60% วิธีสูญหายเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ไม่แตกต่างจากวิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ แต่เนื่องจากวิธีสูญหายเป็นวิธีที่ง่ายในการคำนวณ จึงน่าจะใช้วิธีสูญหายในทางปฏิบัติ แต่ถ้าสัดส่วนการสูญหายของตัวแปรตาม (PM) = 70% วิธีการของฮันท์เป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ

เมื่อขนาดตัวอย่างมีขนาดปานกลางและใหญ่ (NM) = 30, 50, 70 สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนทุกค่า และสัดส่วนการสูญหายของตัวแปรตามทุกระดับ วิธีสูญหายเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ไม่แตกต่างจากวิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ แต่เนื่องจากวิธีสูญหายเป็นวิธีที่ง่ายในการคำนวณ จึงน่าจะใช้วิธีสูญหายในทางปฏิบัติ

2. <u>เมื่อลักษณะของตัวแปรอิสระเป็นรูปแบบที่ 3 ซึ่งเป็นข้อมูลภาคตัดขวาง</u>

เมื่อขนาดตัวอย่างมีขนาดเล็กมาก (NM) = 10 สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 5 -15 และสัดส่วนการสูญหายของตัวแปรตาม (PM) = 10% - 60% วิธีสูญหายเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ไม่แตกต่างจากวิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ แต่เนื่องจากวิธีสูญหายเป็นวิธีที่ง่ายในการคำนวณ จึงน่าจะใช้วิธีสูญหายในทางปฏิบัติ สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 5 และสัดส่วนการสูญหายของตัวแปรตาม (PM) = 70% วิธีการของฮันท์เป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ และสำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 20 - 25 และสัดส่วนการสูญหายของตัวแปรตามทุกระดับ วิธีค่าเฉลี่ยเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ

เมื่อขนาดตัวอย่างมีขนาดเล็กและปานกลาง (NM) = 20, 30 สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 5 - 20 และสัดส่วนการสูญหายของตัวแปรตามทุกระดับ วิธีสูญหายเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ไม่แตกต่างจากวิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ แต่เนื่องจากวิธีสูญหายเป็นวิธีที่ง่ายในการคำนวณ จึงน่าจะใช้วิธีสูญหายในทางปฏิบัติ และสำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อน (σ) = 25 และสัดส่วนการสูญหายของตัวแปรตามทุกระดับ วิธีค่าเฉลี่ยเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ต่ำที่สุดเมื่อเปรียบเทียบกับวิธีการอื่น ๆ

เมื่อขนาดตัวอย่างมีขนาดใหญ่ (NM) = 50, 70 สำหรับค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนทุกค่า และสัดส่วนการสูญหายของตัวแปรตามทุกระดับวิธีสูญหายเป็นวิธีที่ให้ความคลาดเคลื่อนของค่าพยากรณ์ไม่แตกต่างจากวิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ แต่เนื่องจากวิธีสูญหายเป็นวิธีที่ง่ายในการคำนวณ จึงน่าจะใช้วิธีสูญหายในทางปฏิบัติ

จากผลสรุปดังกล่าวข้างต้น สามารถสรุปโดยสถานการณ์ดังนี้

ในสถานการณ์ที่ขนาดตัวอย่างมีขนาดเล็ก (10-20) เมื่อค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีขนาดไม่ใหญ่นัก และสัดส่วนการสูญหายของค่าของตัวแปรตามมีจำนวนมาก (60%-70%) วิธีการของฮันท์จะให้ค่าความคลาดเคลื่อน RMSE ของค่าพยากรณ์ต่ำกว่าวิธีการอื่น ๆ แต่ในขนาดตัวอย่างเดียวกัน ถ้าค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีขนาดเพิ่มขึ้น วิธีค่าเฉลี่ยจะให้ค่าความคลาดเคลื่อน RMSE ของค่าพยากรณ์ต่ำกว่าวิธีการอื่น ๆ ในทุกสัดส่วนการสูญหายของค่าของตัวแปรตาม ส่วนในสถานการณ์ที่ขนาดตัวอย่างมีขนาดปานกลางถึงใหญ่ (30-70) วิธีสูญหายจะเหมาะสมเกือบทุกกรณี นั่นคือถ้าขนาดตัวอย่างใหญ่พอ การตัดชุดข้อมูลสูญหายทิ้งจะมีผลกระทบน้อยมากกับผลการวิเคราะห์การถดถอยเชิงเส้นพหุ โดยใช้วิธีกำลังสองน้อยที่สุด

<u>ข้อเสนอแนะ</u>

เพื่อเป็นแนวทางในการเลือกวิธีการประมาณค่าสูญหายของตัวแปรตาม ในการวิเคราะห์การถดถอยเชิงเส้นพหุเพื่อการพยากรณ์ และเพื่อเป็นแนวทางสำหรับการศึกษาวิจัยนอกเหนือจากการวิจัยครั้งนี้ ดังนั้นผู้วิจัยขอเสนอแนะเป็น 2 ด้านคือ

1. ด้านการนำไปใช้ประโยชน์

กรณีที่ต้องการประมาณค่าที่สูญหายของตัวแปรตาม ในการวิเคราะห์การถดถอยเชิงเส้นพหุ และใช้วิธีกำลังสองน้อยที่สุดหาสัมประสิทธิ์การถดถอย เพื่อหาสมการถดถอยเชิงเส้นพหุในการพยากรณ์ ซึ่งวิธีการประมาณค่าสูญหายได้แก่ วิธีสูญหาย , วิธีค่าเฉลี่ย , วิธีสมการถดถอย , วิธีอีเอ็ม และวิธีการของฮันท์ มีแนวทางดังนี้

สำหรับลักษณะของตัวแปรอิสระทุกรูปแบบ โดยที่ขนาดตัวอย่างมีขนาดเล็ก (10-20) เมื่อค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีขนาดไม่ใหญ่นัก และสัดส่วนการสูญหายของค่าของตัวแปรตามมีจำนวนมาก (60%-70%) วิธีการของฮันท์จะให้ค่าความคลาดเคลื่อน RMSE ของค่าพยากรณ์ต่ำกว่าวิธีการอื่น ๆ แต่ในขนาดตัวอย่างเดียวกัน ถ้าค่าเบี่ยงเบนมาตรฐานของความคลาดเคลื่อนมีขนาดเพิ่มขึ้น วิธีค่าเฉลี่ยจะให้ค่าความคลาดเคลื่อน RMSE ของค่าพยากรณ์ต่ำกว่าวิธีการอื่น ๆ ในทุกสัดส่วนการสูญหายของค่าของตัวแปรตาม ส่วนในสถานการณ์ที่ขนาดตัวอย่างมีขนาดปานกลางถึงใหญ่ (30-70) วิธีสูญหายจะเหมาะสมเกือบทุกกรณี นั่นคือถ้าขนาดตัวอย่างใหญ่พอ การตัดชุดข้อมูลสูญหายทิ้งจะมีผลกระทบน้อยมากกับผลการวิเคราะห์การถดถอยเชิงเส้นพหุโดยใช้วิธีกำลังสองน้อยที่สุด

สำหรับกรณีที่ขนาดตัวอย่างมีขนาด 10 และสัดส่วนการสูญหายของตัวแปรตามมีขนาด 70 ไม่ควรนำไปใช้ เพราะว่าในสภาพความเป็นจริงเมื่อเกิดลักษณะเช่นนี้จะไม่นำมาพิจารณา

2. ด้านการศึกษาวิจัย

2.1 สำหรับกรณีที่ความคลาดเคลื่อนของข้อมูล (ε_t) มีการแจกแจงรูปแบบอื่น วิธีการเหล่านี้อาจไม่มีประสิทธิภาพ จึงควรที่จะศึกษาวิจัยในปัญหาดังกล่าว

2.2 ในการวิจัยครั้งนี้ได้ทำการศึกษา เฉพาะกรณีของการวิเคราะห์การถดถอยเชิงเส้นพหุ (Multiple Linear Regression) เท่านั้น ดังนั้นในการวิจัยครั้งต่อไปอาจจะทำการศึกษาในกรณีของการวิเคราะห์การถดถอยที่ไม่ใช่เชิงเส้น (Non-linear Regression)