

ชื่อเรื่องการค้นคว้าแบบอิสระเชิงวิทยานิพนธ์ การประมาณค่าข้อมูลสูญหายจากแผนทดลอง

ชื่อผู้เขียน นางสาวรุ่งกานต์ กาใจคำ

วิทยาศาสตร์มหาบัณฑิต สาขาวิชาสถิติประยุกต์

คณะกรรมการสอบการค้นคว้าแบบอิสระเชิงวิทยานิพนธ์ :

รองศาสตราจารย์รัชณี คีตพันธ์	คดียพันธ์	ประธานกรรมการ
ผู้ช่วยศาสตราจารย์อันริกา สุปรียศิลป์	สุปรียศิลป์	กรรมการ
อาจารย์พุดพิงษ์	พุกกะมาน	กรรมการ

บทคัดย่อ

ในการวิเคราะห์ข้อมูลของแผนทดลอง หากมีปัญหาเกี่ยวกับข้อมูลสูญหาย จะทำให้การวิเคราะห์ข้อมูลขาดความแม่นยำ และข้อสรุปของการทดลองชุดนั้นผิดพลาดได้ ดังนั้นเมื่อเกิดปัญหาข้อมูลสูญหายจึงจำเป็นต้องแก้ปัญหาเกี่ยวกับข้อมูลสูญหายนั้น ซึ่งวิธีการแก้ปัญหาแบ่งได้เป็น 2 วิธี คือ การตัดค่าสังเกตชุดนั้นทิ้งไป หรือการประมาณค่าสูญหายนั้น แล้วจึงนำไปวิเคราะห์ วิธีแรกจะทำให้จำนวนค่าสังเกตน้อยลง และสูญเสียรายละเอียดของข้อมูลบางตัวไป ในขณะที่วิธีการแก้ปัญหาก็วิธีหนึ่งจะทำให้การทดลองที่ได้มีความถูกต้องแม่นยำยิ่งขึ้น

การประมาณค่าสูญหายทำได้หลายวิธี เช่นวิธีที่ใช้ค่าเฉลี่ย วิธีวิเคราะห์ความถดถอย พหุเชิงเส้น วิธีวิเคราะห์ความถดถอยพหุเชิงเส้นคิดแปลง หรือวิธีวิเคราะห์ส่วนประกอบหลัก แต่ในที่นี้จะเสนอวิธีการประมาณค่าโดยใช้กระบวนการ R ซึ่งเป็นกระบวนการที่จะได้มาซึ่งเมตริกซ์สุดท้าย โดยมีข้อสมมติของการทดลองว่าการทดลองต้องไม่มีผลกระทบร่วมเกิดขึ้นในแต่ละแถว และในแต่ละหลักของการทดลองจะต้องมีค่าสังเกตอย่างน้อย 1 ค่าสังเกต

ก

เมื่อได้เมตริกซ์สุดท้าย M สำหรับแผนทดลองแบบบล็อกสองทาง และแผนทดลองแบบบล็อกสามทางแล้ว ก็สามารถหาค่าของผลบวกกำลังสองของค่าต่างๆ ในตารางวิเคราะห์ความแปรปรวนได้ ซึ่งค่าเหล่านี้จะถูกนำไปทดสอบสมมติฐานต่าง ๆ ต่อไป

จากการศึกษาพบว่าเมื่อเปรียบเทียบวิธีประมาณค่าสูญเสียโดยวิธีใช้ค่าเฉลี่ย และวิธีประมาณค่าสูญเสียโดยใช้กระบวนการ R ในแผนทดลองแบบบล็อกสุ่ม และแผนทดลองแบบบล็อกไม่สมบูรณ์ ซึ่งจัดว่าเป็นแผนทดลองแบบบล็อกสองทางแล้ว ปรากฏว่าค่าประมาณที่ได้จากวิธีการทั้ง 2 วิธีมีค่าใกล้เคียงกัน ดังนั้นหากมีข้อมูลสูญเสียเกิดขึ้นในแผนทดลองแบบบล็อกสองทาง ผู้วิจัยสามารถเลือกใช้วิธีการประมาณค่าสูญเสียวิธีใดก็ได้ ใน 2 วิธีนี้ โดยที่วิธีใช้ค่าเฉลี่ยจะเหมาะสมเมื่อมีจำนวนข้อมูลที่สูญเสียเพียงเล็กน้อย ส่วนในแผนทดลองแบบบล็อกไม่สมบูรณ์ จะใช้วิธีนี้ได้ในกรณีที่มีค่าสังเกตสูญเสียในบล็อกเดียวกันเท่านั้น ส่วนวิธีการประมาณค่าข้อมูลสูญเสียโดยใช้กระบวนการ R จะเหมาะสมในกรณีแผนการทดลองแบบบล็อกไม่สมบูรณ์ มีข้อมูลสูญเสียเป็นจำนวนมาก และค่าสังเกต สูญหายในหลายบล็อก

Research Title Estimation of Missing Data in Experimental Design

Author Miss Rungkan Kajaikum

M.S. Applied Statistics

Examining Committee:

Assoc. Prof. Rajanee Tiyaun **Chairman**

Assist. Prof. Anthika Supriyasilp **Member**

Lecturer Putipong Bookkamana **Member**

Abstract

Analyzing data in an designed experiment with the problem of missing observations will lead to imprecision and false conclusion about the results of the experiment. Therefore, correcting the apparently missing data is necessary. Two methods used for solving the problem are exclusion of the missing observations and estimation of missing observations with calculated values. First method will result in decreasing observation numbers and losing details of some observations whereas another causes the experiment more correctly and precisely.

There are various techniques used for estimating missing data such as ; average mean, multiple linear regression, modified multiple linear regression, and principle component analyses. Hear R process, estimation technique for obtaining a final matrix subject to the assumptions of without interaction in the design and having at least one observation in each row as well as in each column of the incidence matrix, will be introduced.

Final matrix M produced from additive two-way and three-way classification designs will lead to obtaining sums of squares of various values in the analysis of variance table which are used for testing hypothesis further.

Studying estimation of missing observations by comparison between average mean and R process techniques through, as part of additive two-way classification, Randomized Block and Balanced Incomplete Block Designs finds that estimated values obtained from the two techniques are quite similar. Thus, in case of missing data in additive two-way classification, a researcher can select either of them for estimating missing observations. The average mean technique is appropriate for small numbers of missing data, however, in Balanced Incomplete Block Design, it will be applicable to missing observations existing only in the same block. Estimation of missing data through R process is suitable for Balanced Incomplete Block Designs that a lot of data are missing and that many blocks lose observations values.