บทที่ ๑

บทนำ

ความเป็นมาและความสำคัญของปัญหา

จากการศึกษาวิชาการออกแบบการวิจัย (Statistical Design in Education Research) ได้นำเอาวิธีกำลังสองน้อยที่สุดมาแก้ปัญหาทางสถิติหลายตอน ผู้วิจัยจึงใคร่จะศึกษาทฤษฎี, หลักการข้อตกลงเบื้องต้นหรือเงื่อนไขและข้อจำกัดของการนำวิธีกำลังสองน้อยที่สุดไปใช้ จึงได้ปรึกษาและขอคำแนะนำจากอาจารย์ ดร. นิยม ปุราคำ ซึ่งก็ได้สนับสนุนให้ทำการวิจัยเรื่องนี้

วิธีกำลังสองน้อยที่สุดเป็นวิธีประมาณค่าที่เที่ยงตรงและมีประสิทธิภาพสูง สามารถนำไปใช้แก้ปัญหาต่าง ๆ ทางสถิติได้อย่างกว้างขวางเช่น โคแครน (William G. Cochran) ได้กล่าวถึงปัญหาที่เกิดจากการรวบรวมข้อมูลและการนำวิธีกำลังสองน้อยที่สุดไปช่วยแก้ไขเพื่อให้การวิเคราะห์ข้อมูลถูกต้องยิ่งขึ้น[1] เอ็ดเวิร์ด (Allen L. Edwards) ใช้วิธีกำลังสองน้อยที่สุดหาสูตรที่ใช้ในการวิเคราะห์ข้อมูลที่ได้จากการวิจัย[2] เมนเดนฮอลล์ (William Mendenhall) ได้ใช้หลักการของวิธีกำลังน้อยที่สุดวิเคราะห์ข้อมูลที่ได้จากการทดลองทางวิทยาศาสตร์[3] ยัง (Pauline V. Young) กล่าวถึงการวิเคราะห์อนุกรมเวลาเพื่อหาแนวโน้มวิธีต่าง ๆ และสรุป

---

[1] William G. Cochran and Gertrnde M. Cox, Experimental Designs (2 ded., New York: John Wiley & Sons Inc., c 1957), p.80

[2] Allen L. Edwards, Experimental Design in Psychological Research (3 ded.; New York: Holt, Rinehart and Winston, Inc., c 1968), p. 350.

[3] William, Mcdnenhall, Introduction to Linear Models and The Design and Analysis of Experiments (California: Wadsworth Publishing Company, Inc., c 1968), p. 129.

ว่าวิธีกำลังสองน้อยที่สุดเป็นวิธีที่ให้ผลใหล้เคียงที่สุด[๔] เป็นต้น นอกจากนี้วิธีกำลังสองน้อยที่สุดยังสามารถนำไปใช้ได้หลายสาขาวิชาเช่น ทางการศึกษา, จิตวิทยา, สังคมศาสตร์, เศรษฐศาสตร์, ประชากรศาสตร์, บริหารธุรกิจ, และวิทยาศาสตร์ จึงสมควรจะได้รวบรวมทฤษฎีและหลักการนำวิธีกำลังสองน้อยที่สุดไปประยุกต์ใช้เพื่อเป็นเอกสารอ้างอิงแก่ผู้ที่มีความจำเป็นต้องใช้วิธีกำลังสองน้อยที่สุดในงานวิจัยและวิเคราะห์ข้อมูลต่าง ๆ

ความมุ่งหมายของการวิจัย

เพื่อศึกษาทฤษฎีและหลักการนำวิธีกำลังสองน้อยที่สุดไปใช้แก้ปัญหาทางสถิติทางด้านต่าง ๆ โดยแบ่งศึกษาเป็นเรื่อง ๆ ดังนี้

๑. ประวัติและพัฒนาการของวิธีกำลังสองน้อยที่สุด

๒. คุณสมบัติและหลักการของวิธีกำลังสองน้อยที่สุด

๓. ปัญหาการประมาณค่าตัวพารามิเตอร์ในสมการแบบต่าง ๆ และการหาสมการถดถอย

๔. ปัญหาการนำวิธีกำลังสองน้อยที่สุดไปใช้ในการวางแผนและวิเคราะห์ข้อมูลในการวิจัย รวมทั้งการแก้ปัญหาที่อาจจะเกิดขึ้นในระหว่างการวิจัย

๕. ปัญหาการนำวิธีกำลังน้อยที่สุดไปใช้ในการวิเคราะห์อนุกรมเวลาเพื่อหาแนวโน้มระยะยาว

๖. การทดสอบสมมุติฐานและหาขอบเขตของความเชื่อมั่นของค่าที่ได้จากวิธีกำลัง สองน้อยที่สุด

ขอบเขตของการวิจัย

๑. การวิจัยนี้ไม่มีการทดลองปฏิบัติเพื่อยืนยันทฤษฎีต่าง ๆ

---

[4] Pauline V. Young, Scientific Social Surveys and Research (4thed.; New Jersey: Prentice-Hall, Inc., c 1939), p. 342.

๒. วิธีกำลังสองน้อยที่สุดใช้ได้กว้างขวางในสาขาวิชาต่าง ๆ แต่การวิจัยนี้ได้แสดงตัวอย่างการประยุกต์ใช้ทางการศึกษาและการวิจัยเป็นส่วนใหญ่

ข้อตกลงเบื้องต้น

๑. ผู้วิจัยยอมรับว่าวิธีกำลังสองน้อยที่สุดเป็นวิธีที่มีประสิทธิภาพทั้งทางทฤษฎีและปฏิบัติ

๒. ความสัมพันธ์ระหว่างข้อมูลจากการสังเกตหรือการทดลองสามารถเขียนแทนได้ด้วยสมการทางคณิตศาสตร์

ความไม่สมบูรณ์ของการวิจัย

๑. การพิสูจน์ทฤษฎีและการนำวิธีกำลังสองน้อยที่สุดไปใช้โดยอาศัยพื้นฐานทางคณิตศาสตร์ชั้นสูง รวมทั้งการคำนวณโดยใช้คอมพิวเตอร์มิได้กล่าวถึงในการวิจัยนี้

๒. การวิจัยนี้มิได้แสดงหลักการหาสมการกำลังมากกว่า ๓ ขึ้นไปโดยใช้ออร์โทโกนัลโพลี่โนเมียล

๓. การวิจัยนี้มิได้กล่าวถึงค่าผิดพลาดที่เกิดจากการคำนวณโดยใช้วิธีกำลังสองน้อยที่สุด

คำจำกัดความของเทอมสถิติที่สำคัญ

ฟังก์ชั่นการกระจายของความน่าจะเป็น (Probability density function) คือ ความหนาแน่นในการกระจายของค่าต่าง ๆ ที่ได้จากฟังก์ชั่นใด ๆ ถ้าฟังก์ชั่นนั้นเป็นจริงทุกค่าของตัวแปร พื้นที่ทั้งหมดจะเท่ากับ ๑

ตัวแปรสุ่ม (Random Variables) คือตัวแปรที่อาจมีค่าต่าง ๆ เนื่องจากมีความน่าจะเป็นต่าง ๆ กัน

ค่าที่คาดหวัง (Expected Value) หมายถึงค่าที่ควรจะเป็นตามสมมุติฐานทฤษฎีหรือหลักการต่าง ๆ

ประชากร (Population) หมายถึงส่วนทั้งหมดของหน่วยที่เราต้องการจะศึกษา

กลุ่มตัวอย่าง (Sample) หมายถึงส่วนหนึ่งส่วนใดของประชากรที่ถูกเลือกมาเป็นตัวแทนในการศึกษาคุณสมบัติของประชากร

พารามิเตอร์ (Parameter) หมายถึงค่าที่คำนวณได้จากประชากร

ค่าสถิติ (Statistic) หมายถึงค่าที่คำนวณได้จากกลุ่มตัวอย่าง

การประมาณค่า (Estimation) คือวิธีหาค่าสถิติเกี่ยวกับประชากรทั้งหมดจากกลุ่มตัวอย่าง

ตัวประมาณค่า (Estimator) คือสถิติหรือฟังก์ชั่นทางคณิตศาสตร์ซึ่งใช้ประมาณค่าพารามิเตอร์โดยอาศัยข้อมูลจากกลุ่มตัวอย่างเท่านั้น

ความไม่เอนเอียง (Unbiasedness) ตัวประมาณค่า $\hat{\theta}$ จะไม่เอนเอียงเมื่อค่าที่คาดหวังของพารามิเตอร์ θ เท่ากับ θ พอดีหรือ $E(\hat{\theta}) = \theta$

ประสิทธิภาพของตัวประมาณค่า (Efficiency of the estimator) ตัวประมาณค่าจะมีประสิทธิภาพเมื่อมีความแปรปวนน้อยกว่าตัวประมาณค่าอื่น ๆ ตัวประมาณค่าที่มีความแปรปวนต่ำที่สุดจะมีประสิทธิภาพ ๑๐๐ เปอร์เซ็นต์

ฟังก์ชั่นที่น่าจะเป็นได้ (Likelihood function) คือฟังก์ชั่นที่มีการกระจายต่อเนื่องกันและขึ้นกับพารามิเตอร์ โดยค่าเฉลี่ยของฟังก์ชั่นสำหรับทุกตัวแปรเท่ากับฟังก์ชั่นนั้น

วิธีกำลังสองน้อยที่สุด (The Method of Least Squares) เป็นวิธีประมาณค่าโดยทำให้ผลบวกกำลังสองของผลต่างระหว่างค่าที่ได้จากการสังเกตกับค่าตามทฤษฎีมีค่าน้อยที่สุด

สมการปกติ (Normal Equation) หมายถึงสมการที่ใช้หาผลลัพธ์ของสมการ $y = x\beta + \epsilon$ โดยการทำให้ $\frac{\partial(\epsilon'\epsilon)}{\partial\beta} = 0$ ซึ่งจะได้ $(x'x)\hat{\beta} = x'y$

เวคเตอร์ (Vector) คือปริมาณที่แทนขนาดและทิศทางซึ่งอาจใช้แทนเลขสองจำนวนหรือมากกว่านั้นรวมทั้งความน่าจะเป็นของ n มิติ

สิ่งเร้าต่อหน่วยทดลอง (Treatment) อาจหมายถึงวิธีสอน, ยาหรือปุ๋ยที่ใช้กับหน่วยทดลอง เพื่อสังเกตหรือเปรียบเทียบผลการทดลองตามต้องการ

แผนการวิจัยแบบสุ่มตัวอย่าง (Randomized Design) หมายถึงแผนการวิจัยที่เลือกกลุ่มตัวอย่างโดยให้ทุกหน่วยมีโอกาสเท่ากัน

แผนการวิจัยแบบสุ่มตามบล็อก (Randomized Blocks Design) คือแผนการวิจัยซึ่งแบ่งเป็นบล็อกแต่ละบล็อกจะได้รับสิ่งเร้าต่าง ๆ โดยวิธีสุ่มทำให้ได้ค่าคลาดเคลื่อนที่ไม่เอนเอียง

แผนการวิจัยแบบลาตินสแควร์ (Latin Square Design) คือแผนการวิจัยที่กำจัดค่าคลาดเคลื่อนโดยแบ่งเป็นแถวและสดมภ์จำนวนเท่ากันแล้วให้สิ่งเร้าต่อเซลต่าง ๆ โดยให้สิ่งเร้าชนิดหนึ่ง ๆ ปรากฏที่แถวครั้งหนึ่งและสดมภ์ครั้งหนึ่ง

แฟคทอเรียลดีไซน์ (Factorial Design) คือแผนการวิจัยที่ใช้ตรวจสอบผลของแฟคเตอร์ตั้งแต่หนึ่งขึ้นไป โดยแต่ละแฟคเตอร์ถูกแบ่งเป็น ๒ ระดับอย่างน้อย เพื่อศึกษาประสิทธิภาพของระดับต่าง ๆ ของแฟคเตอร์เหล่านี้

สมมุติฐาน (Hypothesis) คือข้อความที่เกี่ยวกับการกระจายของข้อมูล

ระดับความมีนัยสำคัญ (Level of Significance) หมายถึงความน่าจะเป็นของความคลาดเคลื่อนที่เกิดจากการไม่ยอมรับสมมุติฐานที่เป็นจริง

สัญลักษณ์ที่ใช้ในการวิจัย

$\sim$ คือค่าประมาณจากวิธีทำให้ความน่าจะเป็นมีค่าสูงสุด

$\wedge$ หรือ $*$ หมายถึงค่าประมาณที่ได้จากวิธีกำลังสองน้อยที่สุด

$\cap$ หมายถึง เซทที่เป็นไปตามข้อตกลงเบื้องต้น

$\bar{x}$ หรือ $\bar{y}$ หมายถึงมัชฌิมเลขคณิตของกลุ่มตัวอย่าง

μ หมายถึงมัชฌิมเลขคณิตของประชากร

s^2 หมายถึงความแปรปวนของกลุ่มตัวอย่าง

σ^2 หมายถึงความแปรปวนของประชากร

x_i หรือ y_i หมายถึงค่าที่ได้จากการทดลอง

e หรือ ϵ หมายถึงค่าผิดพลาดที่เกิดจากการประมาณค่าจากกลุ่มตัวอย่าง

SS หมายถึงผลบวกกำลังสอง

Q หมายถึงกำลังสองของผลต่างระหว่างค่าที่วัดได้กับค่าที่ได้จากการประมาณ

$$\sum_{i=1}^{n} x_i = (x_1 + x_2 + \ldots\ldots\ldots x_n), \quad \bar{x}_{..} = \sum_{i=j}^{n} \sum_{j=1}^{k} x_{ij}/kn$$