Journal of Research Methodology Volume 11, Number 1 (January - June 1998)
วารสารวิธีวิทยาการวิจัย ปีที่ 11 ฉบับที่ 1 (มกราคม - มิถุนายน 2541)

Variance Component Analysis in Education

Sirichai Kanjanawasee

ABSTRACT

Educational data are usually multi-level hierarchical nested data, such as students are nested in class, classes are nested in school, and schools are nested in sector. According to multi-level structure, "variance component analysis" should be performed. Variance component analysis is a statistical procedure for estimating the variance components in a general linear model. Variance component analysis depicts the decomposition of total variance for each variable that can be attributed to multi-level structure. The information from the analysis will be useful for making decision of the further advanced analysis.

Journal of Research Methodology Volume 11, Number 1 (January - June 1998)
วารสารวิธีวิทยาการวิจัย ปีที่ 11 ฉบับที่ 1 (มกราคม - มิถุนายน 2541)

การวิเคราะห์ส่วนประกอบความแปรปรวนทางการศึกษา

ศิริชัย กาญจนวาสี

บทคัดย่อ

ข้อมูลของตัวแปรทางการศึกษามีลักษณะเป็นข้อมูลพหุระดับที่สอดแทรกลดหลั่น เช่น ข้อมูลตัวแปรเกี่ยวกับสังกัด โรงเรียน ห้องเรียน/ครู และนักเรียน ในการทำความเข้าใจตัวแปรทางการศึกษาจึงควรมีการศึกษาถึงความแปรปรวนของตัวแปรแต่ละระดับของข้อมูล การวิเคราะห์ส่วนประกอบความแปรปรวนเป็นเทคนิคการวิเคราะห์ทางสถิติเชิงเส้นตรงที่จะช่วยให้ทราบว่า ตัวแปรที่สนใจมีส่วนประกอบของความผันแปรเนื่องมาจากแหล่งหรือระดับใด มากน้อยเพียงไร การรู้ถึงความแปรปรวนในแต่ละระดับ นอกจากจะเป็นสารสนเทศสำคัญในการบรรยายตัวแปร ยังเป็นประโยชน์ต่อการเลือกโมเดลการวิเคราะห์ขั้นสูงได้อย่างเหมาะสม

บทนำ

ข้อมูลทางการศึกษาที่ได้จากการสำรวจหรือใช้ในงานวิจัยมักเป็นข้อมูลหลายระดับ (Multi-Level Data) เช่น ข้อมูลระดับนักเรียน ชั้นเรียน โรงเรียน ชุมชน สังกัด เป็นต้น ข้อมูลหลายระดับนี้มีลักษณะของข้อมูลที่มีสมาชิกต่าง ๆของหน่วยย่อยอยู่รวมกันเป็นหน่วยใดหน่วยหนึ่งที่ใหญ่ขึ้นตามลำดับ หรือเรียกว่าการเป็นระดับที่สอดแทรกลดหลั่นกัน (hierarchical nested data) โดยนักเรียนสอดแทรกอยู่ในชั้นเรียน (students nested in class) ชั้นเรียนสอดแทรกอยู่ในโรงเรียน (classes nested in school) โรงเรียนสอดแทรกอยู่ในสังกัด (schools nested in sector)

จากการที่ข้อมูลทางการศึกษามีลักษณะธรรมชาติเป็นข้อมูลหลายระดับที่สอดแทรกลดหลั่นกัน ดังนั้นการวิเคราะห์ข้อมูลของตัวแปรทางการศึกษาจึงควรทราบถึงความแปรปรวนของค่าตัวแปรในระดับต่าง ๆ กันว่ามีส่วนประกอบย่อย ๆ แต่ละส่วนแตกต่างกันอย่างไรตามระดับข้อมูล หรือตัวแปรนั้นมีความแปรปรวนอันเนื่องมาจากความแตกต่างระหว่างนักเรียน ชั้นเรียน โรงเรียน ชุมชน หรือสังกัดมากน้อยเพียงไร

การวิเคราะห์ส่วนประกอบความแปรปรวนเป็นเทคนิคการวิเคราะห์ทางสถิติอย่างหนึ่งที่จะตอบคำถามว่า "ความแปรปรวนของค่าในตัวแปรทางการศึกษามีส่วนประกอบย่อย ๆ แต่ละส่วนแตกต่างกันอย่างไรตามระดับข้อมูล" เช่น นักเรียน ชั้นเรียน โรงเรียน เป็นต้น ซึ่งจะทำให้ทราบการกระจายตัวของตัวแปรตามและตัวแปรอิสระทั้งหลายที่สนใจระหว่างระดับชั้นต่าง ๆ ของข้อมูล

โมเดลการวิเคราะห์

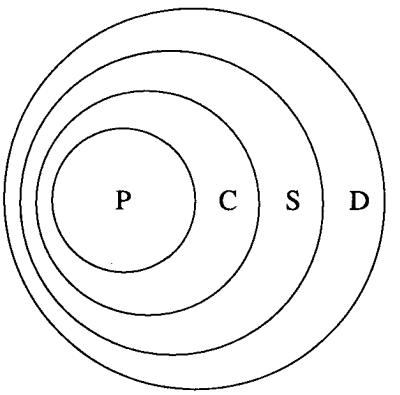
ข้อมูลทางการศึกษาเป็นข้อมูลที่มีลักษณะเป็นพหุระดับที่สอดแทรกลดหลั่นกันดังกล่าวข้างต้น การวิเคราะห์ส่วนประกอบความแปรปรวนจะช่วยให้ผู้วิจัยทราบการกระจายตัวของตัวแปรระหว่างระดับชั้นของข้อมูล สมมุติในกรณีที่เราสนใจข้อมูลของตัวแปร X ในระดับนักเรียน (P) ห้องเรียน (C) โรงเรียน (S) และสังกัด (D)

เราจะพบว่านักเรียนเป็นหน่วยย่อยในห้องเรียน มีนักเรียนหลายคน (P_i) รวมกันอยู่ในห้องเรียนเดียวกัน (C) หรือ P_i สอดแทรกอยู่ใน C ซึ่งสามารถเขียนเป็นสัญลักษณ์ได้ว่า P:C (P nested in C)

ในขณะเดียวกันห้องเรียนเป็นหน่วยย่อยในโรงเรียน มีห้องเรียนหลายห้อง (C_i) รวมกันอยู่ในโรงเรียนเดียวกัน (S) หรือ C_i สอดแทรกอยู่ใน S ซึ่งสามารถเขียนเป็นสัญลักษณ์ได้ว่า C:S (C nested in S)

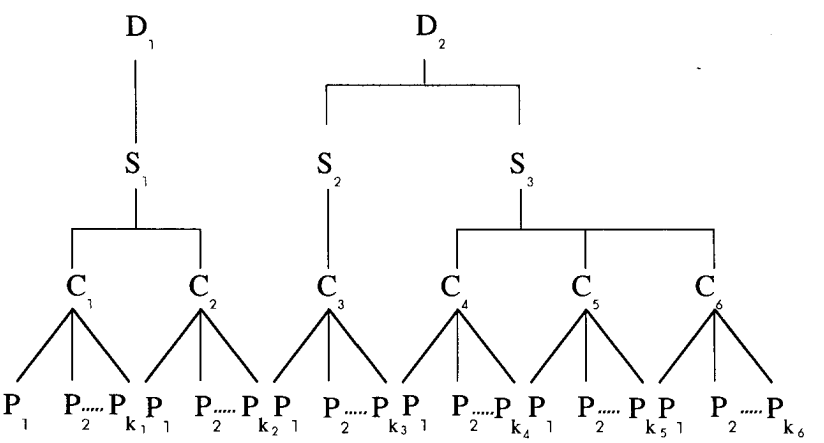
ในทำนองเดียวกัน โรงเรียนเป็นหน่วยย่อยในสังกัด มีโรงเรียนหลายโรง (S_i) รวมกันอยู่ในสังกัดเดียวกัน (D) หรือ S_i สอดแทรกอยู่ใน D ซึ่งสามารถเขียนเป็นสัญลักษณ์ได้ว่า S:D (S nested in D)

ดังนั้นการออกแบบโมเดลการวิเคราะห์จึงสามารถเขียนเป็นสัญลักษณ์ได้ว่า P:C:S:D ส่วนประกอบความแปรปรวนของตัวแปร X แต่ละส่วนตามระดับข้อมูลสามารถแสดงด้วยไดอะแกรมดังภาพ



ภาพที่ 1 ไดอะแกรมของโมเดล P:C:S:D

ในกรณีที่เราสนใจศึกษา 2 สังกัด แต่ละสังกัด (D) มีโรงเรียน 1 และ 2 โรงตามลำดับ แต่ละโรง (S) มี 1-3 ห้องเรียน (C) แต่ละห้องเรียนมีนักเรียน (P) หลายคน สามารถแสดงไดอะแกรมดังภาพที่ 2



ภาพที่ 2 ไดอะแกรมของโมเดล 2 สังกัด

ตัวแปร X มีส่วนประกอบย่อย ๆ แต่ละส่วนตามระดับข้อมูลซึ่งสามารถเขียนเป็นโมเดลทางคณิตศาสตร์เชิงเส้นได้ดังนี้

$$
\begin{aligned}
X_{PCSD} &= \mu && \text{[grand mean]} \\
&\quad + (\mu_D - \mu) && \text{[sector effect]} \\
&\quad + (\mu_{SD} - \mu_D) && \text{[school effect]} \\
&\quad + (\mu_{CSD} - \mu_{SD}) && \text{[class effect]} \\
&\quad + (X_{PCSD} - \mu_{CSD}) && \text{[residual effect]}
\end{aligned}
$$

ดังนั้น $\sigma^2 (X_{PCSD}) = \sigma^2_D + \sigma^2_{S:D} + \sigma^2_{C:D:S} + \sigma^2_{P:C:S:D}$

เมื่อ $\sigma^2 (X_{PCSD})$ = ความแปรปรวนของตัวแปร X ในระดับนักเรียน

σ^2_D = ความแปรปรวนระหว่างสังกัด

$\sigma^2_{S:D}$ = ความแปรปรวนระหว่างโรงเรียนภายในสังกัด

$\sigma^2_{C:D:S}$ = ความแปรปรวนระหว่างห้องเรียนภายในโรงเรียน

$\sigma^2_{P:C:S:D}$ = ความแปรปรวนระหว่างนักเรียนภายในห้องเรียน

วิธีการประมาณค่าส่วนประกอบของความแปรปรวน

การประมาณค่าส่วนประกอบของความแปรปรวนสามารถกระทำได้ด้วยวิธีหลัก 4 วิธี (Corbeil and Searle, 1976 ; Liu and Senturia, 1977 ; นงลักษณ์ วิรัชชัย, 2535)

1) การใช้วิธีวิเคราะห์ความแปรปรวน (AOV) โดยการเลือกโมเดลวิเคราะห์ที่เหมาะสมกับโครงสร้างแหล่งความแปรปรวนของข้อมูลเพื่อคำนวณค่าคาดหมายของกำลังสองเฉลี่ย (expected mean square) ของแต่ละแหล่งความแปรปรวน จากนั้นจึงหาค่าความแปรปรวนของแต่ละส่วนที่ต้องการ

2) การใช้วิธีความเป็นไปได้สูงสุด (Maximum Likelihood estimation : ML) เพื่อประมาณค่าส่วนประกอบความแปรปรวนแต่ละส่วนที่มีความเป็นไปได้สูงสุด หรืออาจใช้ REML (Restricted ML)

3) การใช้วิธีกำลังสองที่ไม่ลำเอียงต่ำสุด (Minimum Norm Quadratic Unbiased Estimation : MINQUE) เพื่อประมาณค่าส่วนประกอบความแปรปรวนแต่ละส่วนที่ให้ค่าความคลาดเคลื่อนกำลังสองต่ำสุด

4) การใช้วิธีของเบส์ (Bayesian estimation)

โปรแกรมคอมพิวเตอร์ที่สามารถนำมาใช้ในการประมาณค่าส่วนประกอบความแปรปรวน ได้แก่ โปรแกรม SPSS โดยใช้คำสั่ง ANOVA และคำนวณค่าความแปรปรวนแต่ละส่วนตามวิธีที่หนึ่ง แต่ถ้าจะให้สะดวกควรใช้โปรแกรม BMDP โดยใช้คำสั่ง 8 V หรือโปรแกรม SAS โดยใช้คำสั่ง VARCOMP ซึ่งจะประมาณค่าส่วนประกอบความแปรปรวนโดยตรง

ตัวอย่างการประมาณค่าส่วนประกอบของความแปรปรวน

ในที่นี้ขอยกตัวอย่างการประมาณค่าความแปรปรวนของตัวแปรตาม 4 ตัว และตัวแปรต้น 15 ตัวโดยใช้ข้อมูลและผลการวิเคราะห์จาก Kanjanawasee (1989) ตัวแปรตามทั้ง 4 ตัว เป็นตัวแปรที่วัดในระดับนักเรียนทั้งหมด ได้แก่ ผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (MATHACH) พัฒนาการของผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (ACHGROWTH) เจตคติต่อคณิตศาสตร์ (MATHATT) และพัฒนาการของเจตคติต่อคณิตศาสตร์ (ATTGROWTH) ส่วนตัวแปรอิสระเป็นตัวแปรที่วัดในระดับนักเรียน ห้องเรียน และโรงเรียน ตัวแปรอิสระที่วัดในระดับนักเรียน ประกอบด้วย ผลสัมฤทธิ์เดิม (PRIORACH) เจตคติเดิม (PRIORATT) ความมุ่งมั่นในการศึกษาต่อ (FURTHED) เศรษฐานะ (SES) และการมีส่วนช่วยทำการบ้านของพ่อแม่ (PARSCON) ตัวแปรอิสระที่วัดในระดับห้องเรียน ได้แก่ ขนาดของห้องเรียน (CLASSIZE) ประสบการณ์ของครู (TEAEXP) คุณภาพการสอน (TEAQUAL) ผลสัมฤทธิ์เฉลี่ยของห้อง (PEERACH) เจตคติต่อคณิตศาสตร์เฉลี่ยของห้อง (PEERATT) และโอกาสในการเรียนเนื้อหาที่ออกข้อสอบ (OTL) ตัวแปรอิสระที่วัดในระดับโรงเรียน ประกอบด้วย ขนาดของโรงเรียน (SCHSIZE) นโยบายการใช้เครื่องคิดเลขของโรงเรียน(CALPOL) ความเป็นผู้นำทางวิชาการของผู้บริหารโรงเรียน(INSTLEAD) และสัดส่วนของนักเรียนต่อครู (STRATIO)

ผู้เขียนใช้โปรแกรม SAS ประมาณค่าส่วนประกอบความแปรปรวนของตัวแปรตาม ซึ่งวัดในระดับนักเรียน 4 ตัว และตัวแปรอิสระระดับนักเรียน 5 ตัว ตามระดับการวัด 3 ระดับ ได้แก่ ระดับนักเรียน (STUDENT) โรงเรียน (SCHOOL) และชุมชน (SCC) โดยแยกความแปรปรวนของตัวแปรเป็น 3 ส่วน คือ ความแปรปรวนระหว่างนักเรียน (ภายในโรงเรียน) ความแปรปรวนระหว่างโรงเรียน(ภายในชุมชน) และความแปรปรวนระหว่างชุมชน นอกจากนี้ยังได้ประมาณค่าส่วนประกอบความแปรปรวนของตัวแปรอิสระระดับห้องเรียนและโรงเรียน โดยแยกความแปรปรวนของตัวแปรเป็น 2 ส่วน ได้แก่ ความแปรปรวนระหว่างโรงเรียน (ภายในชุมชน) และความแปรปรวนระหว่างชุมชน สำหรับคำสั่งที่ใช้ในการวิเคราะห์ และผลการวิเคราะห์ปรากฏในตารางที่ 1

คำสั่ง (SAS)

(1) PROC NESTED AOV; STUDENT SCHOOL SCC;
VAR MATCHACH TO PARSCON

(2) PROC NESTED AOV; SCHOOL SCC;
VAR CLASSIZE TO STRATIO

ตารางที่ 1 ส่วนประกอบความแปรปรวนและร้อยละของส่วนประกอบความแปรปรวนของตัวแปรระดับนักเรียน ชั้นเรียน /โรงเรียน และชุมชน

ตัวแปร	ความแปรปรวนทั้งหมด	ความแปรปรวนระหว่างชุมชน		ความแปรปรวนระหว่างโรงเรียน		ความแปรปรวนระหว่างนักเรียน	
		ความแปรปรวน	ร้อยละ	ความแปรปรวน	ร้อยละ	ความแปรปรวน	ร้อยละ
STUDENTS' OUTCOME							
- MATHACH	59.020	0.962	1.63	24.673	41.80	33.385	56.57
- ACHGROWTH	382.400	5.420	1.42	61.739	16.14	315.241	82.44
- MATHATT	3.214	0.008	0.25	0.187	5.90	3.019	93.93
- ATTGROWTH	1860.166	7.279	0.39	47.223	2.54	1805.664	97.07
STUDENT CHARACTERISTICS							
- PRIORACH	41.196	0.233	0.57	17.603	42.73	23.360	56.70
- PRIORATT	2.900	0	0	0.147	5.07	2.753	94.93
- FURTHED	0.889	0.021	2.36	0.086	9.67	0.782	87.96
- SES	551.212	54.203	9.83	65.960	11.97	431.049	78.20
- PARSCON	28.415	0.164	0.58	1.126	3.96	27.125	95.46
CLASS CHARACTERISTICS							
- CLASSIZE	58.551	10.065	17.19	48.486	82.81		
- TEAEXP	49.990	5.752	11.51	44.238	88.49		
- TEAQUAL	250.118	12.199	4.88	237.919	95.12		
- PEERACH	17.546	0.677	3.86	16.869	96.14		
- PEERATT	0.207	0	0	0.207	100.00		
- OTL	90.401	0	0	90.401	100.00		
SCHOOL CHARACTERISTICS							
- SCHSIZE	1157364.000	271780.000	23.48	885584.000	76.52		
- CALPOL	0.130	0.003	2.31	0.127	97.69		
- INSTLEAD	20.642	0.134	0.65	20.508	99.35		
- STRATIO	24.030	0	0	24.030	100.00		

ผลการวิเคราะห์ส่วนประกอบความแปรปรวนแสดงให้เห็นว่า ตัวแปรตามซึ่งวัดในระดับนักเรียนทั้ง 4 ตัว มีความผันแปรหรือความแตกต่างกันในระดับนักเรียน โรงเรียน และชุมชน จากการเปรียบเทียบขนาดของส่วนประกอบความแปรปรวนพบว่า มีความแตกต่างระหว่างนักเรียนภายในโรงเรียนมากที่สุด (57% ถึง 97%) รองลงมาคือความแตกต่างระหว่างโรงเรียนภายในชุมชน (3% ถึง 42%) และน้อยที่สุดเป็นความแตกต่างระหว่างชุมชน (0.3% ถึง 2%) สำหรับผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (MATHACH) มีความแตกต่างกันสูงทั้งระหว่างนักเรียนและระหว่างโรงเรียน

ส่วนประกอบความแปรปรวนของตัวแปรอิสระที่วัดในระดับนักเรียนก็เช่นกัน มีความผันแปรหรือความแตกต่างระหว่างนักเรียนภายในโรงเรียนมากที่สุด (57% ถึง 95%) รองลงมาเป็นความแตกต่างระหว่างโรงเรียนภายในชุมชน (4% ถึง 43%) และน้อยที่สุดเป็นความแตกต่างระหว่างชุมชน (0% ถึง 10%) สำหรับตัวแปรเศรษฐานะ (SES) มีความแตกต่างระหว่างชุมชน (10%) ซึ่งสูงกว่าตัวแปรอื่นๆ ที่วัดในระดับนักเรียน แสดงว่าเศรษฐานะของนักเรียนมีความแตกต่างระหว่างชุมชนมากกว่าตัวแปรอื่นๆ ในระดับเดียวกัน เป็นที่น่าสังเกตว่าทั้งตัวแปรผลสัมฤทธิ์ทางการเรียนคณิตศาสตร์ (MATHACH) และผลสัมฤทธิ์เดิม (PRIORACH) มีความแตกต่างกันมากระหว่างนักเรียนและระหว่างโรงเรียน น่าจะสะท้อนถึงการมีพื้นฐานทางคณิตศาสตร์ที่แตกต่างกันในตัวนักเรียน และความไม่เสมอภาคทางคุณภาพของการจัดการเรียนการสอนคณิตศาสตร์ระหว่างโรงเรียน

ตัวแปรอิสระในระดับชั้นเรียนและโรงเรียน ความผันแปรส่วนใหญ่เป็นความผันแปรหรือความแตกต่างระหว่างโรงเรียน (77% ถึง 100%) และมีความผันแปรค่อนข้างน้อยระหว่างชุมชน (0% ถึง 23%) สำหรับตัวแปรขนาดของโรงเรียน (SCHSIZE) ค่อนข้างมีความแตกต่างระหว่างชุมชน (23%) ส่วนตัวแปรเจตคติต่อคณิตศาสตร์เฉลี่ยของห้อง (PEERATT) โอกาสในการเรียนเนื้อหาที่ออกข้อสอบ (OTL) และสัดส่วนของนักเรียนต่อครู (STRATIO) มีความผันแปรทั้งหมดเป็นความผันแปรหรือความแตกต่างระหว่างโรงเรียน

ผลการวิเคราะห์แสดงให้เห็นว่า ความแปรปรวนที่แตกต่างกันในแต่ละระดับ น่าจะเป็นเครื่องบ่งชี้ว่าผลของตัวแปรอิสระในแต่ละระดับต่อตัวแปรตาม คงจะมีความแตกต่างกัน การวิเคราะห์ตามประเพณีนิยมที่มีการรวมวิเคราะห์ทุกตัวแปรเสมือนอยู่ในระดับเดียวกัน น่าจะไม่ให้ข้อค้นพบที่ชัดเจน การวิเคราะห์พหุระดับ (Multi-Level Analysis) ตามธรรมชาติความจริงของข้อมูลจากตัวแปรต่างระดับ น่าจะให้ข้อค้นพบที่มีความถูกต้องและชัดเจนกว่าการวิเคราะห์ตามประเพณีนิยม (Burstein, 1980a, 1980b; Kanjanawasee, 1989)

เอกสารอ้างอิง

นงลักษณ์ วิรัชชัย. (2535). การวิเคราะห์ประมาณค่าส่วนประกอบความแปรปรวน. ***วิจัยการศึกษา,*** 15(4), 9-14.

ศิริชัย กาญจนวาสี. (2535). การวิเคราะห์พหุระดับ. ***วิจัยการศึกษา,*** 15(5), 6-14.

Burstein, L. (1980 a). Analysis of Multilevel Data in Education Research and Evaluation. In Bertiner (Ed.), ***Review of Research in Education.*** Washington, D.C. : American Education Research Association.

Burstein, L. (1980 b). The Role of Levels of Analysis in the Specification of Education Effect. In R. Dreeben, and J.A. Thomas (Eds.), ***The Analysis of Educational Productivity : Microanalysis.*** Massachusetts : Ballinger Publishing Company.

Corbeil, R.R., and Searle, S.R. (1976). "A Comparison of Variance Component Estimators". ***Biometrics,*** 32 (December), 779-791.

Kanjanawasee, S. (1989). Alternative Strategies for Policy Analysis : An Assessment of School Effects on Students' Cognitive and Affective Arithmatics Outcomes in Lower Secondary Schools in Thailand. Doctoral Dissertation in Education, University of California, Los Angeles.

Liu, Lon-Mu, and Senturia, J. (1977). Computation of MINQUE Variance Component Estimates. ***Journal of the American Statistical Association,*** (December), 867-868.