

การคำนวณค่าคลาดเคลื่อนมาตรฐานที่ไม่ถูกต้องของโปรแกรม SPSS ในการทดลองแบบแฟกทอเรียลกรณีตัวแบบอิทธิพลผสม

Wrong Standard Error Calculation from SPSS Program for Factorial Experiments : Mixed Model

กมล บุขบา

ภาควิชาคณิตศาสตร์และสถิติ คณะวิทยาศาสตร์และเทคโนโลยี

มหาวิทยาลัยธรรมศาสตร์ ศูนย์รังสิต ปทุมธานี 12121

การทำงานกับโปรแกรม SPSS สำหรับวินโดวส์ (SPSS for Windows) ผู้ใช้สามารถทำงานโดยใช้เมาส์เลือกเมนูต่างๆ และโต้ตอบกับกล่องข้อความ (dialog box) หรือการเขียนคำสั่งที่อยู่ในแฟ้มประเภท syntax file สำหรับชุดพื้นฐาน (base system) ของโปรแกรม SPSS นั้นผู้ใช้สามารถจัดการกับข้อมูล เช่น การดัดแปลงข้อมูล การเลือกบางรายการ รวมทั้งการจัดการเกี่ยวกับแฟ้มและการให้น้ำหนักตัวแปร สามารถวิเคราะห์ข้อมูลได้ทั้งสถิติเบื้องต้นและสถิติขั้นสูง ทั้งสถิติประเภทที่อิงและไม่อิงพารามิเตอร์ และผู้ใช้ยังสามารถจัดการเกี่ยวกับผลลัพธ์ทั้งที่เป็นตารางและกราฟ รวมทั้งการส่งผลลัพธ์ไปใช้ในโปรแกรมอื่นๆ ได้หลายรูปแบบ [1]

จุดเด่นของโปรแกรม SPSS อยู่ในการทำงานกับข้อมูลที่ได้จากการสำรวจตามความหมายของชื่อโปรแกรม (Statistical Package for the Social Sciences) ซึ่งสะท้อนให้เห็นว่าโปรแกรมได้รับการพัฒนามาจากสายงานใด ทำให้โปรแกรมสามารถจัดการกระทำกับข้อมูลที่มีจำนวนมากได้ โดยงานต่างๆ ที่ต้องการทำกับข้อมูลจากการสำรวจสามารถใช้โปรแกรม SPSS ได้เกือบทั้งหมด จุดด้อยที่สำคัญข้อหนึ่งของโปรแกรมคือ การใช้โปรแกรมเพื่อวิเคราะห์ข้อมูลจากแผนแบบการทดลอง (experimental designs) แบบต่างๆ [2] อย่างไรก็ตามนักทดลองส่วนใหญ่ยังคงใช้โปรแกรม SPSS กันอย่างแพร่หลายเนื่องจากความสะดวกและค่าใช้จ่ายที่ไม่สูงมากในการจัดหาโปรแกรม

การทดลองแบบแฟกทอเรียล

การทดลองแบบแฟกทอเรียลหมายถึงการทดลองที่มีการออกแบบทรีทเมนต์ในลักษณะแฟกทอเรียล (factorial treatment designs) [3] การออกแบบทรีทเมนต์ในลักษณะนี้หมายถึงทรีทเมนต์ที่ใช้ในการทดลองประกอบด้วยทุกทางเลือกที่เป็นไปได้ของระดับต่างๆ ของปัจจัยหลายๆ ปัจจัยที่กำลังศึกษาตัวอย่างเช่น ในการศึกษาการใช้ยาฆ่าแมลงเพื่อปรับปรุงคุณภาพเนื้อโค ปัจจัยที่สนใจศึกษาได้แก่ปริมาณยาฆ่าแมลงซึ่งแบ่งเป็น 4 ระดับ (เช่น ร้อยละ 0.0002 หรือ 0.0004 โดยน้ำหนักเนื้อโค เป็นต้น) และระยะเวลาที่ใช้ในการนวดผสมซึ่งแบ่งเป็น 3 ระดับ (เช่น 10 หรือ 15 นาที เป็นต้น) ดังนั้นการออกแบบทรีทเมนต์แบบแฟกทอเรียลจะประกอบด้วยจำนวนทรีทเมนต์ทั้งสิ้น 12 ทรีทเมนต์ หรือ ตัวอย่างของการทดลองเพื่อศึกษาสภาวะที่เหมาะสมในการลดปริมาตรของสารละลายแอนโธไซยานินจากการระเหยด้วยวิธีสุญญากาศโดยใช้เครื่อง Vacuum Rotary Evaporator การทดลองนี้มีปัจจัยที่เกี่ยวข้องคือความดันและอุณหภูมิของเครื่องมือดังกล่าว สมมติว่านักทดลองต้องการศึกษาที่ความดัน 3 ระดับและอุณหภูมิ 3 ระดับ ดังนั้นจะมีจำนวนทรีทเมนต์ที่เป็นไปได้ทั้งหมด 9 ทรีทเมนต์ เช่น การใช้เครื่องมือที่ความดัน 800 mbar และอุณหภูมิ 45 องศาเซลเซียส เป็นต้น ข้อดีของการทดลองแบบแฟกทอเรียลคือมีประสิทธิภาพมากกว่าการทดลองที่ละปัจจัย และจำเป็นที่จะต้องใช้ในการทดลองในลักษณะนี้กรณีที่มีหรือคาดว่าจะมีปฏิสัมพันธ์ (interactions) ระหว่างปัจจัยที่ต้องการศึกษาเพื่อไม่ทำให้การสรุปผลผิดพลาด นอกจากนี้ยังช่วยให้นักทดลองสามารถประมาณอิทธิพลของปัจจัยหนึ่งที่มีระดับต่างๆ ของปัจจัยอีกปัจจัยหนึ่งได้ [4]

ตัวแบบอิทธิพลผสม

สมมติว่าในการทดลองเพื่อศึกษาอิทธิพลของปริมาณยางมะลกอที่มีต่อการปรับปรุงคุณภาพเนื้อนั้น นักทดลองสนใจศึกษาปริมาณยางมะลกอเฉพาะ 4 ระดับดังต่อไปนี้เท่านั้น กล่าวคือ ร้อยละ 0.0, 0.0002, 0.0004, และ 0.0006 ของน้ำหนักเนื้อโค ดังนั้นนักทดลองจึงได้ระบุระดับทั้งสี่ของปริมาณยางมะลกอในการทดลอง นั่นคือประชากรที่สนใจของระดับประกอบด้วยระดับทั้งสี่ที่ใช้ในการทดลอง และถ้าหากมีการทำการทดลองในลักษณะนี้อีก นักทดลองก็ยังคงที่จะใช้ปริมาณยางมะลกร้อยละ 0.0, 0.0002, 0.0004, และ 0.0006 อีกเช่นกัน อิทธิพลของยางมะลกอในการทดลองนี้จะเรียกว่าเป็น*อิทธิพลคงที่ (fixed effects)* ซึ่งการสรุปผลจะจำกัดเฉพาะกลุ่มของระดับที่ใช้ศึกษาในการทดลอง [3] และความสนใจของนักทดลองจะอยู่ที่การประมาณค่าหรือการทดสอบสมมติฐานเกี่ยวกับค่าเฉลี่ยของตัวแปรที่สนใจในแต่ละระดับหรือในแต่ละทรีทเมนต์เท่านั้น อย่างไรก็ตามยังมีการทดลองในอีกลักษณะหนึ่งที่มีการสุ่มระดับของปัจจัยจากกลุ่มประชากรของระดับที่มีจำนวนระดับมากกว่าทำให้ระดับที่ใช้ในการทดลองครั้งหนึ่งๆ เป็นเพียงแค่ตัวอย่างสุ่มเท่านั้น โดยจะใช้ผลที่ศึกษาจากตัวอย่างของระดับที่ใช้ในการทดลองขยายไปยังระดับทุกระดับที่มีในประชากร ทำให้ความสนใจของการศึกษาสำหรับการทดลองในลักษณะนี้ไม่จำกัดเฉพาะเพียงแค่ตัวอย่างของระดับที่ใช้ในการทดลอง แต่ต้องการประมาณค่าหรือทดสอบสมมติฐานเกี่ยวกับความผันแปรของอิทธิพลนี้ซึ่งเรียกว่า*อิทธิพลสุ่ม (random effects)* [4]

สำหรับการทดลองที่มีปัจจัยหลายปัจจัยที่สนใจศึกษา เช่น การทดลองแบบแฟกทอเรียล นักทดลองอาจมีความสนใจศึกษาค่าเฉลี่ยของตัวแปรที่สนใจจำแนกตามระดับของปัจจัยหนึ่ง และสนใจศึกษาความผันแปรของอิทธิพลของอีกปัจจัยหนึ่ง นั่นคือการทดลองจะประกอบด้วยปัจจัยที่มีทั้งอิทธิพลคงที่และอิทธิพลสุ่ม ตัวแบบทางสถิติที่ใช้กับการทดลองที่มีทั้งอิทธิพลคงที่และอิทธิพลสุ่มจะเรียกว่า “**ตัวแบบอิทธิพลผสม (mixed model)**” [3]

ตัวอย่างเช่น ในการศึกษาความแตกต่างของวิธีการทางเคมีสองวิธีที่ใช้ในการตรวจวิเคราะห์ไตรกลีเซอไรด์ (triglycerides) ในเซรัมมนุษย์ [3] ถ้าปัจจัยที่สนใจศึกษาได้แก่วิธีการทางเคมี (สองวิธี) และวันที่ทำการทดลอง (สี่วัน) โดยมี

การออกแบบทรีทเมนต์แบบแฟกทอเรียล จะเห็นว่าอิทธิพลของวิธีการทางเคมีเป็นอิทธิพลคงที่เนื่องจากผู้วิจัยสนใจศึกษาและสรุปผลเพียงแค่วิธีการสองวิธีนี้เท่านั้นและถ้ามีการทำการทดลองนี้อีกครั้งก็สามารถใช้วิธีการสองวิธีนี้ในการทดลองได้อีก ในทางตรงกันข้ามอิทธิพลของวันเป็นอิทธิพลสุ่มเนื่องจากวันสี่วันที่ทำการทดลองจะพิจารณาว่าเป็นตัวอย่างสุ่มของวันต่างๆ ที่นักวิจัยสามารถที่จะทำการทดลองได้ และถ้าทำการทดลองซ้ำอีกครั้งก็ไม่สามารถที่จะทดลองในวันสี่วันเดิมที่ได้เคยทำการทดลองไปแล้ว

การใช้โปรแกรม SPSS และผลลัพธ์

จากตัวอย่างการศึกษาวิธีการทางเคมีสองวิธีในการตรวจวิเคราะห์ไตรกลีเซอไรด์ ถ้าในแต่ละวันทำการเตรียมตัวอย่างเซรัมสองตัวอย่างเพื่อใช้กับวิธีการทั้งสองวิธี และลำดับของการตรวจวิเคราะห์ปริมาณไตรกลีเซอไรด์ด้วยเครื่องมือชนิดเดียวกันโดยนายช่างเทคนิคคนเดียวกันเป็นไปอย่างสุ่มตามแผนแบบการทดลองสุ่มสมบูรณ์ (completely randomized design) ที่มีสองซ้ำ สมมติว่าปริมาณไตรกลีเซอไรด์ที่วัดได้ (mg/dl) ในตัวอย่างเซรัมแสดงในตารางที่ 1 (ข้อมูลจาก [3])

การวิเคราะห์ข้อมูลในตารางที่ 1 ด้วยโปรแกรม SPSS สำหรับวินโดวส์ ให้ดำเนินการตามขั้นตอนต่อไปนี้

Analyze

General Linear Model

Univariate...

ตารางที่ 1 ปริมาณไตรกลีเซอไรด์ (mg/dl) ในตัวอย่างเซรัมจากการทดลองแบบแฟกทอเรียลที่มีวิธีการทางเคมีเป็นอิทธิพลคงที่และวันเป็นอิทธิพลสุ่ม

วิธีการทางเคมี	วันที่			
	1	2	3	4
1	142.3	134.9	148.6	152.0
	144.0	146.3	156.5	151.4
2	142.9	125.9	135.5	142.9
	147.4	127.6	138.9	142.3

จากนั้นเลือกตัวแปรตามที่แทนปริมาณไตรกลีเซอไรด์ (ตัวแปร level) ใส่ในช่อง Dependent Variable ทำการระบุปัจจัยทั้งสองโดยการเลือกตัวแปรที่แทนวิธีการทางเคมี (ตัวแปร method) เข้าในช่อง Fixed Factor(s) และเลือกตัวแปรที่แทนวันที่ทำการทดลอง (ตัวแปร day) เข้าในช่อง Random Factor(s) ทำการเลือกปุ่ม Options เพื่อให้กล่องข้อความอีกอันหนึ่งปรากฏ จากนั้นกำหนดให้โปรแกรมแสดงค่าเฉลี่ยจำแนกตามวิธีการทางเคมีและจำแนกตามวิธีการทางเคมีในแต่ละวันโดยการเลือกตัวแปร method และตัวแปร method*day เข้าในช่อง Display Means for ตามลำดับ แล้วกดปุ่ม OK เพื่อให้โปรแกรมแสดงผลลัพธ์ การทำงานตามเมนูดังกล่าวมีคำสั่งของโปรแกรม SPSS ดังภาพที่ 1

unianova level by method day /random = day

/method = sstype(3) /intercept = include

/emmeans = tables(method)

/emmeans = tables(method*day)

/criteria = alpha(.05)

/design = method day method*day .

ภาพที่ 1 คำสั่งของโปรแกรม SPSS สำหรับการทดลองแบบแฟกทอเรียลกรณีตัวแบบอิทธิพลผสมสำหรับข้อมูลในตารางที่ 1

ผลลัพธ์ที่ได้เฉพาะการแสดงผลค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์จำแนกตามตัวแปร method หรือวิธีการทางเคมี พร้อมค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยและขอบเขตของช่วงเชื่อมั่นร้อยละ 95 ของค่าเฉลี่ยประชากรของปริมาณไตรกลีเซอไรด์แสดงในภาพที่ 2 ส่วนผลลัพธ์ในทำนองเดียวกันแต่จำแนกตามวิธีการทางเคมีในแต่ละวันแสดงในภาพที่ 3

Chemistry method				
Dependent Variable: LEVEL				
Chemistry method	Mean	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound
1	147.000	1.342	143.904	150.096
2	137.925	1.342	134.829	141.021

ภาพที่ 2 ค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์จำแนกตามวิธีการทางเคมีและค่าคลาดเคลื่อนมาตรฐานที่คำนวณจากโปรแกรม SPSS

จากผลลัพธ์ในภาพที่ 2 การตรวจวิเคราะห์ปริมาณไตรกลีเซอไรด์ด้วยวิธีการทางเคมีวิธีที่ 1 และ 2 ให้ค่าเฉลี่ยเท่า

กับ 147.000 และ 137.925 mg/dl ตามลำดับโดยมีค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์เท่ากับ 1.342 mg/dl

2. Chemistry method * Day					
Dependent Variable: LEVEL					
Chemistry method	Day	Mean	Std. Error	95% Confidence Interval	
				Lower Bound	Upper Bound
1	1	143.150	2.685	136.959	149.341
	2	140.600	2.685	134.409	146.791
	3	152.550	2.685	146.359	158.741
	4	151.700	2.685	145.509	157.891
2	1	145.150	2.685	138.959	151.341
	2	126.750	2.685	120.559	132.941
	3	137.200	2.685	131.009	143.391
	4	142.600	2.685	136.409	148.791

ภาพที่ 3 ค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์จำแนกตามวิธีการทางเคมีในแต่ละวันและค่าคลาดเคลื่อนมาตรฐานที่คำนวณจากโปรแกรม SPSS

ผลลัพธ์ในภาพที่ 3 แสดงค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์ที่ตรวจวิเคราะห์ด้วยวิธีการทางเคมีวิธีที่ 1 และ 2 จำแนกตามวันที่ทำการทดลอง เช่น วิธีการที่ 1 ในวันที่ 1 ให้ปริมาณไตรกลีเซอไรด์โดยเฉลี่ย 143.150 mg/dl โดยมีค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยของวิธีหนึ่งๆ ในวันที่หนึ่งๆ เป็น 2.685 mg/dl เป็นต้น

การคำนวณค่าคลาดเคลื่อนมาตรฐานที่ถูกต้อง

ถ้าให้ A แทนปัจจัยที่เป็นอิทธิพลคงที่และ B แทนปัจจัยที่เป็นอิทธิพลสุ่ม ตัวแบบทางสถิติชนิดที่ไม่มีเงื่อนไขบังคับเกี่ยวกับเทอมปฏิสัมพันธ์ (unrestricted model) สำหรับการทดลองนี้ได้แก่

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

โดยที่ $i = 1, 2, \dots, a$ ระดับ, $j = 1, 2, \dots, b$ ระดับ, และ $k = 1, 2, \dots, r$ ซ้ำ ในที่นี้ μ แทนค่าเฉลี่ยประชากร, α_i แทนอิทธิพลคงที่ของปัจจัย A, β_j แทนอิทธิพลสุ่มที่เป็นอิสระกันของปัจจัย B ซึ่งมีแจกแจงแบบปกติโดยมีค่าเฉลี่ยเป็น 0 และความแปรปรวน σ_B^2 , $(\alpha\beta)_{ij}$ แทนอิทธิพลปฏิสัมพันธ์กลุ่มที่เป็นอิสระกัน มีการแจกแจงแบบปกติโดยมีค่าเฉลี่ย 0 และความแปรปรวน σ_{AB}^2 , และ ε_{ij} แทนความคลาดเคลื่อนสุ่มจากการทดลองที่เป็นอิสระกัน มีการแจกแจงแบบปกติโดยมีค่า

เฉลี่ย 0 และความแปรปรวน σ^2 นอกจากนี้ยังตกลงว่าตัวแปรสุ่ม β_j , $(\alpha\beta)_{ij}$, และ ε_{ijk} เป็นอิสระต่อกันแบบทุกคู่ (pairwise independent) อีกด้วย

การคำนวณค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยจะต้องทำการหาความแปรปรวน (Var) ของค่าเฉลี่ยก่อน ความแปรปรวนของค่าเฉลี่ยจำแนกตามปัจจัย A ($\bar{Y}_{i..}$) และค่าเฉลี่ยของแต่ละระดับของปัจจัย A ที่ระดับหนึ่งๆ ของปัจจัย B ($\bar{Y}_{ij.}$) คำนวณได้ดังนี้

$$\begin{aligned} Var(\bar{Y}_{i..}) &= Var\left(\frac{1}{br} \sum_j \sum_k (\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk})\right) \\ &= Var\left(\mu + \alpha_i + \frac{1}{br} r \sum_j \beta_j + \frac{1}{br} r \sum_j (\alpha\beta)_{ij} + \frac{1}{br} \sum_j \sum_k \varepsilon_{ijk}\right) \\ &= \frac{1}{b^2} \sum_j Var(\beta_j) + \frac{1}{b^2} \sum_j Var((\alpha\beta)_{ij}) + \frac{1}{b^2 r^2} \sum_j \sum_k Var(\varepsilon_{ijk}) \\ &= \frac{1}{b^2} b\sigma_B^2 + \frac{1}{b^2} b\sigma_{AB}^2 + \frac{1}{b^2 r^2} br\sigma^2 \\ &= \frac{\sigma_B^2}{b} + \frac{\sigma_{AB}^2}{b} + \frac{\sigma^2}{br} = \frac{r\sigma_B^2 + r\sigma_{AB}^2 + \sigma^2}{br} \quad (1) \end{aligned}$$

$$\begin{aligned} Var(\bar{Y}_{ij.}) &= Var\left(\frac{1}{r} \sum_k (\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk})\right) \\ &= Var\left(\mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \frac{1}{r} \sum_k \varepsilon_{ijk}\right) \\ &= Var(\beta_j) + Var((\alpha\beta)_{ij}) + \frac{1}{r^2} \sum_k Var(\varepsilon_{ijk}) \\ &= \sigma_B^2 + \sigma_{AB}^2 + \frac{\sigma^2}{r} = \frac{r\sigma_B^2 + r\sigma_{AB}^2 + \sigma^2}{r} \quad (2) \end{aligned}$$

จากนั้นจึงทำการประมาณค่าความแปรปรวนของค่าเฉลี่ยข้างต้นโดยอาศัยค่าคาดหวัง ($E[.]$) ของกำลังสองเฉลี่ยซึ่งสามารถแสดงได้ว่า ([3] และ [4])

$$E[MSE] = \sigma^2 \quad (3)$$

$$E[MSAB] = \sigma^2 + r\sigma_{AB}^2 \quad (4)$$

$$E[MSB] = \sigma^2 + r\sigma_{AB}^2 + ra\sigma_B^2 \quad (5)$$

เมื่อ MSE, MSAB และ MSB คือผลบวกกำลังสองเฉลี่ยของความคลาดเคลื่อน, ของปฏิสัมพันธ์ระหว่างปัจจัย A กับ B, และของปัจจัย B ตามลำดับซึ่งสามารถคำนวณได้จากข้อ

มูลจริง นั่นคือสามารถประมาณความแปรปรวนของค่าเฉลี่ยด้วยวิธีการโมเมนต์ได้ดังนี้

$$\text{จาก (3)} \quad \sigma^2 \approx MSE \quad (6)$$

$$\text{จาก (4)} \quad \sigma^2 + r\sigma_{AB}^2 \approx MSAB \quad (7)$$

$$\text{จาก (4), (5)} \quad r\sigma_B^2 \approx (MSB - MSAB)/a \quad (8)$$

ดังนั้นสามารถประมาณความแปรปรวนของค่าเฉลี่ยได้โดยอาศัยการแทนค่า (6), (7), และ (8) ใน (1) และ (2) ทำให้ได้ว่า

$$\begin{aligned} Var(\bar{Y}_{i..}) &\approx \frac{(MSB - MSAB)/a + MSAB}{br} \\ &= \frac{(a-1)MSAB + MSB}{abr} \quad (9) \end{aligned}$$

$$\begin{aligned} Var(\bar{Y}_{ij.}) &\approx \frac{(MSB - MSAB)/a + MSAB}{r} \\ &= \frac{(a-1)MSAB + MSB}{ar} \quad (10) \end{aligned}$$

อีกหนึ่งตัวแบบที่ไม่มีเงื่อนไขบังคับเกี่ยวกับเทอมปฏิสัมพันธ์นี้เป็นตัวแบบที่โปรแกรม SPSS ใช้ในการคำนวณอัตราส่วนของสถิติเอฟในตารางวิเคราะห์ความแปรปรวนกรณีตัวแบบอิทธิพลผสม สำหรับตัวแบบที่มีเงื่อนไขบังคับเกี่ยวกับเทอมปฏิสัมพันธ์การหาค่าคาดหวังของกำลังสองเฉลี่ยจะต่างไปจากนี้ผู้ที่สนใจอาจศึกษาได้จาก [3] และ [4] อย่างไรก็ตามความแปรปรวนของค่าเฉลี่ยยังคงมีค่าเท่ากับ (9) และ (10)

การเปรียบเทียบกับผลลัพธ์จากโปรแกรม SPSS

จากตัวอย่างการศึกษาวิธีการทางเคมีสองวิธีในการตรวจวิเคราะห์ที่ไตรกลีเซอไรด์ซึ่งมีปัจจัย A (ตัวแปร method) คือวิธีการทางเคมีและปัจจัย B (ตัวแปร day) คือวันที่ทำการทดลอง ตารางวิเคราะห์ความแปรปรวนที่ได้จากโปรแกรม SPSS แสดงในภาพที่ 4

Tests of Between-Subjects Effects					
Dependent Variable: LEVEL					
Source	Type III Sum of Squares	df	Mean Square	F	Sig.
Intercept	324729.023	1	324729.023	2257.977	.000
METHOD	431.442	1	431.442	5.347	.04
DAY	184.833	3	61.611 ^a	2.334	.252
METHOD * DAY	184.833	3	61.611 ^a	4.273	.045
Error	115.340	8	14.418 ^b		
Total					

a. MS(DAY)
b. MS(METHOD * DAY)
c. MS(Error)

ภาพที่ 4 ตารางวิเคราะห์ความแปรปรวนของปริมาณไตรกลีเซอไรด์ที่คำนวณจากโปรแกรม SPSS

จากผลลัพธ์ในภาพที่ 4 จะได้ว่า $MSE=14.418$, $MSAB=61.611$, และ $MSB = 143.814$ ดังนั้นสามารถประมาณค่าแปรปรวนของค่าเฉลี่ยได้ดังนี้

จาก (9)

$$Var(\bar{Y}_{i..}) \approx \frac{(1)(61.611) + 143.814}{(2)(4)(2)} = 12.8391 \text{ จาก}$$

(10)

$$Var(\bar{Y}_{ij.}) \approx \frac{(1)(61.611) + 143.814}{(2)(2)} = 51.3563$$

ทำให้ได้ค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์จำแนกตามวิธีการทางเคมีเท่ากับ $(12.8391)^{1/2} = 3.5832 \text{ mg/dl}$ และค่าคลาดเคลื่อนมาตรฐานของการประมาณปริมาณไตรกลีเซอไรด์โดยเฉลี่ยจำแนกตามวิธีการทางเคมีที่วันหนึ่งๆ เท่ากับ $(51.3563)^{1/2} = 7.1663 \text{ mg/dl}$ ซึ่งไม่เท่ากับผลลัพธ์จากโปรแกรม SPSS

ผลลัพธ์ที่ได้จากโปรแกรม SPSS นั้นคำนวณโดยใช้สูตรการคำนวณความแปรปรวนของค่าเฉลี่ยในกรณีที่มีตัวแบบเป็นชนิดอิทธิพลคงที่ ซึ่งมีสูตรคือ [3]

$$\begin{aligned} Var(\bar{Y}_{i..}) &= \frac{\sigma^2}{br} \\ &\approx \frac{MSE}{br} = \frac{14.418}{(4)(2)} = 1.8023 \end{aligned}$$

$$\begin{aligned} Var(\bar{Y}_{ij.}) &= \frac{\sigma^2}{r} \\ &\approx \frac{MSE}{r} = \frac{14.418}{2} = 7.209 \end{aligned}$$

ทำให้ได้ค่าคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย ของปริมาณไตรกลีเซอไรด์จำแนกตามวิธีการทางเคมี และจำแนกตามวิธีการทางเคมีที่วันหนึ่งๆ เป็น $(1.8023)^{1/2} = 1.3425 \text{ mg/dl}$ และ $(7.209)^{1/2} = 2.6850 \text{ mg/dl}$ ตามลำดับ

การวิเคราะห์ข้อมูลด้วยโปรแกรม SAS

เมื่อนำข้อมูลดังกล่าวไปวิเคราะห์ด้วยโปรแกรม SAS โดยใช้เมนู SAS Analyst เลือก Statistics ตามด้วย ANOVA และเลือก Mixed Model... หรือใช้การเขียนคำสั่งด้วย Proc Mixed ดังนี้ [5]

```
Proc mixed data=faxmix method=reml ;
```

```
class method day;
```

```
model y = method / htype=3
```

```
ddfm = satterth;
```

```
random day method*day / ;
```

```
lsmeans method ;
```

```
run;
```

ผลลัพธ์ที่ได้แสดงในภาพที่ 5 (อนึ่ง คำสั่ง Proc Mixed จะคำนวณค่าเฉลี่ยจำแนกตามปัจจัยที่เป็นอิทธิพลคงที่เท่านั้น) ซึ่งจะเห็นว่าค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยของปริมาณไตรกลีเซอไรด์จำแนกตามวิธีการทางเคมีเท่ากับ 3.5832 mg/dl และเป็นค่าที่ถูกต้อง

Least Squares Means					
method	Estimate	Error	DF	t Value	Pr > t
1	147.00	3.5832	5.17	41.03	<.0001
2	137.93	3.5832	5.17	38.49	<.0001

ภาพที่ 5 ผลลัพธ์จากโปรแกรม SAS เฉพาะค่าเฉลี่ยแบบกำลังสองน้อยที่สุดและค่าคลาดเคลื่อนมาตรฐาน

บทสรุป

ในกรณีที่มีการทดลองแบบแฟกทอเรียลที่มีปัจจัยที่มีอิทธิพลคงที่และอิทธิพลสุ่มนั้น ถ้าผู้วิจัยสนใจที่จะรายงานค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยจะต้องทำด้วย

ความระมัดระวังเป็นอย่างยิ่ง การนำค่าที่คำนวณโดยโปรแกรมสำเร็จรูปมาใช้โดยไม่มีความรู้ทางสถิติที่ดีพอจะทำให้ผลลัพธ์ที่ได้จากการทดลองหรือการวิจัยขาดความน่าเชื่อถือ โดยเฉพาะอย่างยิ่งผลลัพธ์ที่จะนำไปสู่การตัดสินใจในเรื่องที่มีความสำคัญต่อการบริหารงานและการกำหนดนโยบาย

การคำนวณค่าคลาดเคลื่อนมาตรฐานของการประมาณค่าเฉลี่ยที่ไม่ถูกต้องของโปรแกรม SPSS นี้เป็นเพียงตัวอย่างหนึ่งที่จะชี้ให้เห็นถึงความสำคัญของการเลือกใช้โปรแกรมสำเร็จรูปทางสถิติ ความเข้าใจอย่างถ่องแท้ในแนวคิดของตัวแบบสถิติและที่มาของค่าต่างๆ ในผลลัพธ์ของโปรแกรมสำเร็จรูป อย่างไรก็ตามผลลัพธ์ในส่วนอื่นที่เป็นผลลัพธ์หลักๆ เช่น ตารางวิเคราะห์ความแปรปรวน โดยทั่วไปจะมีความถูกต้อง ผู้ใช้ควรมีวิจารณญาณในการเลือกใช้คำสั่งและการเลือกใช้ผลลัพธ์เพื่อนำเสนอในรายงานวิจัย นอกจากนี้ในบางสถานการณ์ การคำนวณด้วยมือก็เป็นสิ่งที่จำเป็นอย่างหนึ่งเช่นกันเนื่องจากผลลัพธ์บางอย่างที่ได้จากโปรแกรมสำเร็จรูปบางชนิดนั้นคำนวณไม่ถูกต้อง

เอกสารอ้างอิง

- [1] SPSS Inc., SPSS Base 10.0 User's Guide., SPSS Inc., Chicago, 537 p, 1999.
- [2] Currall, J., Choosing a Statistical Analysis Package., http://www.go2hill.com/Research/Development/Statistics/r_measure_4.htm.,1991.
- [3] Kuehl, R. O., Design of Experiments: Statistical Principles of Research Design and Analysis., 2 ed., Duxbury., Pacific Grove, 666 p, 2000.
- [4] Montgomery, D.C., Design and Analysis of Experiments. 6 ed., John Wiley & Sons.,New York, 684 p, 2001.
- [5] Littell, R. C., Stroup, W. W., and Freund, R. J., SAS For Linear Models., 4 ed., SAS Publishing, Cary, 466 p, 2002.