13 การวิเคราะห์ความแปรปรวน Analysis of Variance

■ บทนำ

ความแปรปรวน (Variance) เป็นมาตรการวัดการกระจายของข้อมูลซึ่งมีความสัมพันธ์กับส่วนเบี่ยงเบนมาตรฐาน เนื่องจากความแปรปรวนสามารถคำนวณได้จากส่วนเบี่ยงเบนมาตรฐานยกกำลังสอง ความแปรปรวนจึงเป็นการวัดการกระจายของข้อมูลในรูปของพื้นที่ สำหรับการวิเคราะห์ความแปรปรวน (Analysis of Variance) หรือเรียกอย่างย่อว่า ANOVA เป็นวิธีหนึ่งในการทดสอบสมมติฐานเพื่อเปรียบเทียบค่าเฉลี่ยของประชากรมากกว่า 2 กลุ่มขึ้นไปพร้อม ๆ กัน ซึ่งยังคงใช้หลักการเปรียบเทียบระหว่างกลุ่มประชากรกลุ่มเดียวหรือ 2 กลุ่มในการใช้ z หรือ t ทดสอบ กล่าวคือ ถ้าเป็นการเปรียบเทียบกับประชากรเพียงกลุ่มเดียว ส่วนใหญ่จะเป็นการเปรียบเทียบค่าระหว่างค่าเฉลี่ยกับค่าที่ผู้วิจัยสนใจหรือค่าที่กำหนดขึ้น ส่วนการเปรียบเทียบระหว่างประชากร 2 กลุ่ม จะเป็นการเปรียบเทียบความแตกต่างของค่าเฉลี่ยระหว่างประชากรทั้ง 2 กลุ่มตามสมมติฐานที่กำหนดไว้ ในกรณีที่มีประชากรมากกว่า 2 กลุ่มหรือมีสิ่งที่ต้องการศึกษามากกว่า 2 สิ่ง ความแตกต่างระหว่างกลุ่มจะถูกวัดในรูปของส่วนเบี่ยงเบนมาตรฐานหรือความแปรปรวน ซึ่งในที่นี้ก็คือค่าเฉลี่ยของความแปรผัน (Mean Squares) โดยที่ความแปรผันเกิดจากผลรวมยกกำลังสองของความแตกต่างระหว่างค่าแต่ละค่าของค่าเฉลี่ย การวิเคราะห์ความแปรปรวน จึงเป็นการเปรียบเทียบความแตกต่างระหว่างกลุ่มกับความแตกต่างภายในกลุ่ม ในกรณีที่ประชากรมีมากกว่า 2 กลุ่มในลักษณะของความแปรผัน โดยมีเงื่อนไขว่าข้อมูลที่ได้จากกลุ่มประชากรที่นำมาทดสอบความแตกต่างของค่าเฉลี่ยข้อมูลของแต่ละกลุ่ม จะต้องมีการแจกแจงแบบปกติเท่านั้น

ในการทดสอบสมมติฐานโดยใช้วิธีการวิเคราะห์ความแปรปรวน ตัวแปรที่ศึกษาอาจมีเพียงตัวแปรเดียวหรือหลายตัวแปรที่ศึกษาพร้อม ๆ กันก็ได้ แต่ละตัวแปรอาจแยกออกได้หลายระดับหรือหลายชนิด ซึ่งเรียกส่วนย่อย ๆ นี้ว่า สิ่งทดลอง หรือ Treatment ยกตัวอย่างเช่น ต้องการเปรียบเทียบสมรรถนะของไมโครโพรเซสเซอร์ที่มีจำหน่ายในตลาดปัจจุบัน จำนวน 4 ตัวด้วยกัน เพื่อศึกษาหาไมโครโพรเซสเซอร์ที่มีสมรรถนะดีที่สุด โดยพิจารณาจากตัวแปรต่าง ๆ ที่เกี่ยวข้อง เช่น ประสิทธิภาพ ความร้อน ความเร็ว และราคา เป็นต้น หรือถ้าต้องการศึกษาวิธีการเรียนการสอนที่ดีที่สุด จำนวน 3 วิธี ที่ส่งผลให้ผู้เรียนมีผลสัมฤทธิ์ทางการเรียนสูงที่สุด โดยพิจารณาจากตัวแปรที่เป็น Treatment ได้แก่ ผลคะแนน และความพึงพอใจของผู้เรียน เป็นต้น การวิเคราะห์ความแปรปรวน จึงเหมาะสำหรับการทดสอบสมมติฐานในการวิจัยเชิงทดลองหรือการวิจัยเชิงประยุกต์มากกว่าการวิจัยทางสังคมศาสตร์ โดยตัวแปรที่ต้องการศึกษาจะมีลักษณะดังต่อไปนี้

1. ตัวแปรอิสระหรือตัวแปรต้นอาจมีเพียงตัวเดียวหรือมากกว่าหนึ่งตัว แต่ตัวแปรตามต้องมีเพียงตัวเดียวเท่านั้น

2. ลักษณะของตัวแปรอิสระหรือตัวแปรต้นอาจจำแนกออกเป็นระดับต่าง ๆ หรือจำแนกออกเป็นประเภทต่าง ๆ เช่น ระดับความสามารถของผู้เรียน จำแนกออกเป็นเก่ง ปานกลาง และอ่อน เป็นต้น

3. ลักษณะของตัวแปรตามต้องมีค่าต่อเนื่อง เช่น คะแนน น้ำหนัก ส่วนสูง ความเร็ว หรือเจตคติ เป็นต้น

การวิเคราะห์ความแปรปรวน จำแนกออกเป็น 2 ประเภทดังนี้

1. การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA)
2. การวิเคราะห์ความแปรปรวนสองทาง (Two-way ANOVA)

■ การวิเคราะห์ความแปรปรวนทางเดียว (One-way ANOVA)

การวิเคราะห์ความแปรปรวนทางเดียว หรือ One-way ANOVA เป็นวิธีการทดสอบเพื่อวิเคราะห์ความสัมพันธ์ ระหว่างตัวแปรอิสระหรือตัวแปรต้นตัวเดียวกับตัวแปรตามเพียงตัวเดียว โดยที่ตัวแปรอิสระหรือตัวแปรต้นอาจมีลักษณะเป็นตัวแปรเชิงคุณภาพ (Qualitative Variable) ที่จำแนกออกเป็นระดับหรือประเภทต่าง ๆ เช่น เก่ง-ปานกลาง-อ่อน ดีมาก-ดี-พอใช้-แย่ เป็นต้น ส่วนตัวแปรตามอาจมีลักษณะเป็นตัวแปรเชิงปริมาณ (Quantitative Variable) เพื่อศึกษาความสัมพันธ์ของตัวแปรอิสระหรือตัวแปรต้นว่าจะส่งผลอย่างไรกับตัวแปรตาม ตามสมมติฐานการวิจัยที่กำหนดไว้

ตัวอย่างหัวข้อเรื่องการวิจัยสำหรับการวิเคราะห์ความแปรปรวนทางเดียว

- การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียนที่เรียนด้วยบทเรียนคอมพิวเตอร์แบบปกติ แบบมัลติมีเดีย และแบบไฮเปอร์มีเดีย ในวิชาโครงสร้างข้อมูล ระดับ ปวส.
  ตัวแปรอิสระหรือตัวแปรต้น : บทเรียนคอมพิวเตอร์ (จำแนกออกเป็น 3 ชนิด ได้แก่ แบบปกติ แบบมัลติมีเดีย และแบบไฮเปอร์มีเดีย)
  ตัวแปรตาม : ผลสัมฤทธิ์ทางการเรียนของผู้เรียน
- การเปรียบเทียบสมรรถนะของไมโครโพรเซสเซอร์ของ Intel, Cyrix, AMD และ Motorola เมื่อนำไปใช้กับไมโครคอมพิวเตอร์ทั่วไป
  ตัวแปรอิสระหรือตัวแปรต้น : ไมโครโพรเซสเซอร์ (จำแนกออกเป็น 4 ชนิด ได้แก่ Intel, Cyrix, AMD และ Motorola)
  ตัวแปรตาม : สมรรถนะ
- การเปรียบเทียบฐานะทางสังคมของประชาชนระดับกลางที่อาศัยอยู่ในจังหวัดชลบุรี ระยอง จันทบุรี ตราด และสระแก้ว

ตัวแปรอิสระหรือตัวแปรต้น : ประชาชนระดับกลาง (จำแนกออกเป็น 5 กลุ่ม ได้แก่ จังหวัดชลบุรี ระยอง จันทบุรี ตราด และสระแก้ว)
ตัวแปรตาม : ฐานะทางสังคม

สำหรับการวิเคราะห์ความแปรปรวนทางเดียวมีหลายวิธี ได้แก่ วิธี CR-k, RB-k และ RM-k ในที่นี้จะขอกล่าวถึงวิธี CR-k และวิธี RB-k ซึ่งเป็นวิธีการวิเคราะห์ความแปรปรวนทางเดียวที่นิยมใช้กันโดยทั่วไป โดยเฉพาะวิธี CR-k ที่เป็นวิธีการวิเคราะห์ความแปรปรวนอย่างง่าย

กระบวนการวิเคราะห์ความแปรปรวนก็คือ การทดสอบสมมติฐานวิธีหนึ่งสำหรับข้อมูลที่ประกอบด้วยตัวแปรอิสระหรือตัวแปรต้นหลายตัวแปรกับตัวแปรตามเพียงตัวเดียว ขั้นตอนการวิเคราะห์ความแปรปรวน มีดังนี้

1. กำหนดสมมติฐานเป็นกลาง (H_0) และสมมติฐานตรงข้าม (H_1) ของการวิจัย ซึ่งผู้วิจัยจะต้องคาดการณ์เองว่าจะกำหนดไว้อย่างไร
2. กำหนดระดับนัยสำคัญ (α) เช่น .05, .01 หรือ .001 ตามความต้องการของผู้วิจัยและแนวโน้มของความเป็นไปได้
3. คำนวณหาค่า F จากสูตรที่กำหนดให้ โดยพิจารณาเลือกใช้สูตรตามเงื่อนไขของสูตรที่กำหนดไว้
4. เปิดตาราง F เพื่อหาค่าวิกฤติของ F ตามระดับนัยสำคัญ
5. เปรียบเทียบค่า F ที่คำนวณได้กับค่าที่ได้จากการเปิดตาราง ถ้าค่า F ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าวิกฤติที่ได้จากการเปิดตาราง แสดงว่าค่า F มีนัยสำคัญทางสถิติ ก็จะปฏิเสธ H_0 และยอมรับ H_1
6. สรุปผลที่ได้

การวิเคราะห์ความแปรปรวนทางเดียววิธี CR-k

การวิเคราะห์ความแปรปรวนทางเดียววิธี CR-k เป็นวิธีการวิเคราะห์ความแปรปรวนอย่างง่าย CR ย่อมาจาก Completed Randomized หมายถึง การสุ่มตลอด หรือการสุ่มอย่างสมบูรณ์ ส่วน k หมายถึง จำนวนกลุ่มตัวอย่าง

การวิเคราะห์ความแปรปรวนทางเดียววิธีนี้ กระทำกับกลุ่มตัวอย่างจำนวน k กลุ่ม ตามระดับ ชนิด หรือประเภทของตัวแปรอิสระหรือตัวแปรต้น เช่น ผู้เรียนเก่ง ปานกลาง และอ่อน (k = 3) ซึ่งกลุ่มตัวอย่างทั้งหมดจะต้องเกิดจากการสุ่มเท่านั้น โดยจะมีจำนวนเท่ากันหรือมีจำนวนแตกต่างกันก็ได้ กรณีนี้หมายถึงว่าตัวแปรอิสระหรือตัวแปรต้น จะถูกจำแนกออกเป็นระดับหรือประเภทต่าง ๆ ดังนี้

สมมติให้ตัวแปรอิสระหรือตัวแปรต้นแทนจำนวนกลุ่มตัวอย่างเป็น A จะได้ว่า

A_1 คือ ผู้เรียนเก่ง สมมติให้มีจำนวน 6 คน

A_2 คือ ผู้เรียนปานกลาง สมมติให้มีจำนวน 4 คน

A_3 คือ ผู้เรียนอ่อน สมมติให้มีจำนวน 5 คน

กำหนดให้ . (จุด) แทนค่าที่วัดในตัวแปรในกลุ่มตัวอย่าง การวิเคราะห์ความแปรปรวนทางเดียววิธี CR-k จะเขียนรูปแบบของการวิเคราะห์ได้ว่า

A_1	A_2	A_3
.	.	.
.	.	.
.	.	.
.	.	.
.		.
.		

ภาพที่ 13-1 รูปแบบการวิเคราะห์ความแปรปรวนทางเดียววิธี CR-k

สำหรับการวิเคราะห์ความแปรปรวนทางเดียวของตัวแปรที่ต้องการศึกษา จะต้องทำการวิเคราะห์ผลรวมของกำลังสองของความเบี่ยงเบนของค่าต่าง ๆ ที่เกี่ยวข้อง หรือที่เรียกว่า ผลรวมกำลังสอง (Sum of Squares) ของการแปรผันทั้งหมด โดยกำหนดให้

$$SS_T = SS_b + SS_w$$

เมื่อ

SS_T = ผลรวมของการแปรผันทั้งหมด (Total Sum of Squares)

SS_b = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ระหว่างกลุ่ม (Between Group Sum of Squares)

SS_w = ผลรวมของการแปรผันภายในกลุ่ม (Within Group Sum of Squares)

สามารถเขียนเป็นสูตรได้ดังนี้

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X})^2 = \sum_{j=1}^{k}n_j(\overline{X}_j-\overline{X})^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X})^2$$

เมื่อ

X_{ij} = ค่าที่ศึกษาในตัวแปรตาม

$\overline{X}$ = ค่าเฉลี่ยของค่าที่ศึกษาในตัวแปรตามทั้งหมด

$\overline{X_j}$ = ค่าเฉลี่ยของค่าที่ศึกษาในตัวแปรตามในกลุ่มที่ j

k = จำนวนกลุ่มตัวอย่างทั้งหมด

n_j = จำนวนสมาชิกของกลุ่มตัวอย่างในกลุ่มที่ j

จากสูตร $SS_T = SS_b + SS_w$ จะเห็นว่าค่าของความแปรผันรวมทั้งหมด (SS_T) เกิดจากผลรวมของความแปรผันระหว่างกลุ่ม (SS_b) บวกกับผลรวมของการแปรผันภายใน (SS_w) ซึ่งค่าของ SS_T, SS_b และ SS_w สามารถหาได้หลายวิธี โดยที่การทดสอบสมมติฐานความแปรปรวนจะใช้วิธีการทดสอบค่า F ซึ่งค่า F เกิดจากผลต่างระหว่างค่าเฉลี่ยของความแปรปรวนระหว่างกลุ่มกับค่าเฉลี่ยของความแปรปรวนภายในกลุ่ม นั่นคือ

สูตรการหาค่า F

$$F = \frac{MS_b}{MS_w}$$

เมื่อ

MS_b = ค่าเฉลี่ยของความแปรปรวนระหว่างกลุ่ม (Mean Squares Between Groups)

MS_w = ค่าเฉลี่ยของความแปรปรวนภายในกลุ่ม (Mean Squares Within Groups)

โดยที่ $$MS_b = \frac{SS_b}{k\text{-}1}$$

$$MS_w = \frac{SS_w}{N\text{-}k}$$

เมื่อ

SS_b = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ระหว่างกลุ่ม (Between Group Sum of Squares)

SS_w = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ภายในกลุ่ม (Within Group Sum of Squares)

k-1 = Degree of Freedom ของการแปรผันระหว่างกลุ่ม (df_b)

N-k = Degree of Freedom ของการแปรผันภายในกลุ่ม (df_w)

จากสูตร $SS_T = SS_b + SS_w$

$$\sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X})^2 = \sum_{j=1}^{k}n_j(\overline{X}_j-\overline{X})^2 + \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X})^2$$

$$SS_b = \sum_{j=1}^{k}n_j(\overline{X}_j-\overline{X})^2$$

$$SS_w = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(X_{ij}-\overline{X})^2$$

ตัวอย่างที่ 13-1

บริษัทพัฒนาซอฟท์แวร์คอมพิวเตอร์แห่งหนึ่ง เลือกใช้คอมพิวเตอร์ 4 ยี่ห้อ ๆ ละ 5 เครื่อง โดยเริ่มใช้งานพร้อม ๆ กัน หลังจากใช้งานภายในระยะเวลา 3 ปี พบว่าคอมพิวเตอร์ทั้ง 20 เครื่อง เกิดอาการข้อบกพร่องทั้งอาการย่อยและอาการใหญ่เป็นจำนวนครั้งตามตารางดังต่อไปนี้

	ยี่ห้อคอมพิวเตอร์			
	A	B	C	D
เครื่องที่ 1	14	15	10	13
เครื่องที่ 2	11	13	12	11
เครื่องที่ 3	12	12	11	12
เครื่องที่ 4	11	16	10	15
เครื่องที่ 5	11	12	12	13

จากอาการข้อบกพร่องของคอมพิวเตอร์ทั้ง 4 ยี่ห้อ จงทดสอบที่ระดับความเชื่อมั่น 95% ว่าเครื่องคอมพิวเตอร์มีอาการข้อบกพร่องแตกต่างกันหรือไม่

สมมติฐานการวิจัย

$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$ (คอมพิวเตอร์มีอาการข้อบกพร่องไม่แตกต่างกัน)

$H_1 : \mu_i \neq \mu_j$ (คอมพิวเตอร์มีอาการข้อบกพร่องแตกต่างกัน)
(เมื่อ $i \neq j$ และ $i, j = 1, 2, 3$ และ 4)

เมื่อ

μ_1 = จำนวนครั้งโดยเฉลี่ยที่คอมพิวเตอร์ยี่ห้อ A เกิดอาการข้อบกพร่อง

μ_2 = จำนวนครั้งโดยเฉลี่ยที่คอมพิวเตอร์ยี่ห้อ B เกิดอาการข้อบกพร่อง

μ_3 = จำนวนครั้งโดยเฉลี่ยที่คอมพิวเตอร์ยี่ห้อ C เกิดอาการข้อบกพร่อง

μ_4 = จำนวนครั้งโดยเฉลี่ยที่คอมพิวเตอร์ยี่ห้อ D เกิดอาการข้อบกพร่อง

กำหนดความเชื่อมั่นที่ 95% หรือระดับนัยสำคัญที่ .05

จากโจทย์จะได้ว่า N = 20, k = 4 ในขั้นแรกหาค่า SS_b และ SS_w ก่อน

X_{1i}	X_{2i}	X_{3i}	X_{4i}	$(X_{1i} - X_1)^2$	$(X_{2i} - X_2)^2$	$(X_{3i} - X_3)^2$	$(X_{4i} - X_4)^2$
14	15	10	13	4.84	1.96	1	.04
11	13	12	11	.64	.36	1	3.24
12	12	11	12	.04	2.56	0	.64
11	16	10	15	.64	5.76	1	4.84
11	12	12	13	.64	2.56	1	.04
59	68	55	64	6.8	13.2	4	8.8

จากตาราง $\overline{X}_1 = 11.8, \overline{X}_2 = 13.6, \overline{X}_3 = 11, \overline{X}_4 = 12.8,$

$\overline{X} = 12.3$

แทนค่าในสูตรเพื่อหา SS_b และ SS_w

$$SS_b = \sum_{j-1}^{k} n_j (\overline{X}_j - \overline{X})^2$$

$$SS_b = n_1(\overline{X}_1 - \overline{X})^2 + n_2(\overline{X}_2 - \overline{X})^2 + n_3(\overline{X}_3 - \overline{X})^2 + n_4(\overline{X}_4 - \overline{X})^2$$

$$SS_b = 5(11.8 - 12.3)^2 + 5(13.6 - 12.3)^2 + 5(11 - 12.3)^2 + 5(12.8 - 12.3)^2$$

$$SS_b = 19.4$$

หาค่า $$SS_w = \sum_{j=1i=1}^{k} \sum^{n_j} (X_{ij} - \overline{X})^2$$

$$SS_w = \sum_{i=1}^{5} (\overline{X}_{1i} - \overline{X}_1)^2 + \sum_{i=1}^{5} (\overline{X}_{2i} - \overline{X}_2)^2 + \sum_{i=1}^{5} (\overline{X}_{3i} - \overline{X}_3)^2 + \sum_{i=1}^{5} (\overline{X}_{4i} - \overline{X}_4)^2$$

$$SS_w = 6.8 + 13.2 + 4 + 8.8$$

$$SS_w = 32.8$$

ขั้นต่อไป หาค่า df

df สำหรับการแปรผันระหว่างกลุ่ม = k -1 = 4 - 1 = 3

df สำหรับการแปรผันภายในกลุ่ม = N - k = 20 - 4 = 16

df สำหรับการแปรผันทั้งหมด = N - 1 = 20 -1 = 19

ขั้นต่อไป หาค่า MS_b และ MS_w

$$MS_b = \frac{SS_b}{k\text{-}1} = \frac{19.4}{4-1} = \frac{19.4}{3} = 6.46$$

$$MS_w = \frac{SS_w}{N\text{-}k} = \frac{32.8}{20-4} = \frac{32.8}{16} = 2.05$$

ขั้นต่อไป แทนหาในสูตรเพื่อคำนวณหาค่า F

$$F = \frac{MS_b}{MS_w} = \frac{6.46}{2.05} = 3.15$$

ขั้นต่อไป เปิดตารางหาค่า F ที่ระดับนัยสำคัญ .05 ที่ df = 3, 16

ปรากฏว่า ค่า F (.05, 3, 16) จากตาราง (หน้า 394) มีค่าเท่ากับ 3.24 ซึ่งมีค่าสูงกว่าค่าที่คำนวณได้ แสดงว่าค่าที่คำนวณได้อยู่ในเขตยอมรับ (หรืออยู่นอกเขตวิกฤติ) จึงไม่มีเหตุผลใด ๆ ที่จะปฏิเสธ H_0 จึงยอมรับ H_0 สรุปได้ว่า คอมพิวเตอร์ทั้ง 4 ยี่ห้อ มีอาการข้อบกพร่องเหมือนกัน (หรือไม่แตกต่างกัน) ที่ระดับความเชื่อมั่น 95% (หรือที่ระดับนัยสำคัญ .05)

สาเหตุของความแปรปรวน	Sum of Squares	df	Mean Squares	F	ค่าวิกฤติของ F
Between Groups	19.4	3	6.46	3.15	3.24
Within Groups	32.8	16	2.05		
Total		19			

การวิเคราะห์ความแปรปรวนทางเดียววิธี RB-k

การวิเคราะห์ความแปรปรวนทางเดียววิธี CR-k ที่ผ่านมา เป็นการวิเคราะห์ความแปรปรวนโดยการสุ่มอย่างสมบูรณ์เพื่อคัดเลือกกลุ่มตัวอย่างออกเป็นประเภทหรือระดับต่าง ๆ แต่ปัญหาที่อาจจะเกิดขึ้นได้ก็คือ ถ้าหากประชากรมีความแตกต่างกันค่อนข้างมาก เช่น ผู้เรียนเก่งกับผู้เรียนอ่อน ถ้าคัดเลือกเข้าอยู่กลุ่มเดียวกัน ผลการวิเคราะห์ข้อมูลก็อาจจะเกิดความคลาดเคลื่อนได้ ซึ่งความแตกต่างระหว่างบุคคลนับว่าเป็นปัจจัยสำคัญในการจัดกลุ่มหรือจัดประเภท แต่ถ้าทราบคุณลักษณะของกลุ่มตัวอย่างแล้ว การคัดเลือกกลุ่มตัวอย่างออกเป็นประเภทหรือระดับต่าง ๆ ให้มีคุณลักษณะใกล้เคียงกัน ความแปรปรวนในส่วนนี้ที่เกิดขึ้นก็จะไม่เกี่ยวข้องกับความแปรปรวนที่เกิดจากความคลาดเคลื่อน ทำให้ผลการวิเคราะห์ข้อมูลได้ผลตรงตามเป้าหมายมากขึ้น

การวิเคราะห์ความแปรปรวนทางเดียววิธี RB-k จึงเป็นการวิเคราะห์ความแปรปรวนที่พิจารณาสมาชิกภายในกลุ่ม ซึ่ง RB ย่อมาจาก Randomized Block หมายถึง การสุ่มสมาชิกภายในกลุ่มให้มีคุณลักษณะใกล้เคียงกัน ส่วน k หมายถึง จำนวนกลุ่มตัวอย่าง และ Block หรือ bl หมายถึง ชั้นหรือพวกของสมาชิกในแต่ละกลุ่ม วิธี RB-k จะเป็นการพิจารณาจัดกลุ่มสมาชิกที่มีคุณลักษณะใกล้เคียงกันให้อยู่พวกหรือชั้นเดียวกัน เช่น มีคะแนนเท่ากัน ดังนั้น แต่ละ Block จึงมีความใกล้เคียงกัน และต่าง Block ก็จะแตกต่างกัน

สมมติให้ตัวแปรอิสระแทนวิธีการเรียนการสอนเป็น A โดยแบ่งออกเป็น 3 กลุ่ม จะได้ว่า

A_1 คือ วิธีการเรียนการสอนแบบบรรยาย และสมมติให้มีจำนวน 5 คน

A_2 คือ วิธีการเรียนการสอนแบบถามตอบ และสมมติให้มีจำนวน 5 คน

A_3 คือ วิธีการเรียนการสอนแบบศึกษาด้วยตนเอง และสมมติให้มีจำนวน 5 คน

กำหนดให้ . (จุด) แทนค่าที่วัดในตัวแปรในกลุ่มตัวอย่าง การวิเคราะห์ความแปรปรวนทางเดียววิธี RB-k จะเขียนรูปแบบการวิเคราะห์ได้ดังภาพที่ 10-2 ซึ่งสมมติให้มีสมาชิกแต่ละกลุ่มมีจำนวน 5 คน ซึ่งอยู่ภายใต้หลักพื้นฐานที่ว่าสมาชิกแต่ละ Block จะมีคุณลักษณะใกล้เคียงกัน (ความแตกต่างภายนอก Block จะมีมากกว่าความแตกต่างภายใน Block) นั่นคือ สมาชิกแต่ละ Block (bl_1, bl_2, bl_3, bl_4 และ bl_5) จะสามารถเทียบเคียงกันได้ เช่น ถ้าต้องการจัดการเรียนการสอนแบบ A_1, A_2 และ A_3 กับผู้เรียนที่จัดพวกเป็น Block โดยทำการคัดเลือกสมาชิกแต่ละ Block ให้มีความใกล้เคียงกัน ข้อมูลที่นำมาพิจารณาอาจเป็นความสามารถทางการเรียนของสมาชิกแต่ละคนก็ได้ เช่น กลุ่มเก่ง ปานกลาง และอ่อน เป็นต้น

	A_1	A_2	A_3
bl_1	.	.	.
bl_2	.	.	.
bl_3	.	.	.
bl_4	.	.	.
bl_5	.	.	.

ภาพที่ 13-2 รูปแบบการวิเคราะห์ความแปรปรวนทางเดียววิธี RB-k

ตัวอย่างที่ 13-2

ในการวิจัยเชิงทดลอง เพื่อพัฒนาบทเรียนคอมพิวเตอร์ที่มีลักษณะวิธีการนำเสนอแตกต่างกัน 4 รูปแบบ ในวิชาการโปรแกรมภาษาเชิงวัตถุ เป้าหมายเพื่อนำไปใช้กับกลุ่มตัวอย่างที่เป็นผู้เรียนจำนวน 16 คน โดยผู้เรียนทั้งหมดมีคะแนนการเรียนวิชาโครงสร้างข้อมูลและอัลกอริทึม ซึ่งเป็นรายวิชาบังคับก่อนการเรียนวิชาการโปรแกรมภาษาเชิงวัตถุ ดังนี้

ผู้เรียน :	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
คะแนน :	14	15	19	21	19	22	25	17	25	27	17	28	23	19	30	31

เมื่อพิจารณาจากคะแนนรายวิชาบังคับก่อน โดยแบ่งออกคะแนนออกเป็น 4 กลุ่ม (4 Block) จะได้ดังนี้

Block ที่ 1 : คะแนน 14 - 17 ได้แก่ ผู้เรียน B, C, I และ L
Block ที่ 2 : คะแนน 18 - 21 ได้แก่ ผู้เรียน D, E, F และ O
Block ที่ 3 : คะแนน 22 - 25 ได้แก่ ผู้เรียน G, H, J และ N
Block ที่ 4 : คะแนน 26 - 31 ได้แก่ ผู้เรียน K, M, P และ Q
จากการแบ่งผู้เรียนออกเป็น 4 Block จะได้ดังนี้

	A_1	A_2	A_3	A_4
bl_1	B	C	I	L
bl_2	D	E	F	O
bl_3	G	H	J	N
bl_4	K	M	P	Q

การวิเคราะห์ความแปรปรวนทางเดียววิธี RB-k จะต้องทำการวิเคราะห์ผลรวมของกำลังสองของความเบี่ยงเบนของค่าต่าง ๆ ที่เกี่ยวข้อง เรียกว่า ผลรวมกำลังสอง (Sum of Squares) ของการแปรผันทั้งหมด เช่นเดียวกันกับวิธี CR-k โดยกำหนดให้

$$SS_T = SS_b + SS_{bl} + SS_{a.bl}$$

เมื่อ

SS_T = ผลรวมของการแปรผันทั้งหมด (Total Sum of Squares)
SS_{bl} = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ระหว่างกลุ่ม (Between Group Sum of Squares)
SS_{bl} = ผลรวมของการแปรผันระหว่าง Block (Within Block Sum of Squares)
$SS_{a.bl}$ = ผลรวมของการแปรผันระหว่าง Block กับระดับหรือประเภทของตัวแปรอิสระ (Within Block Sum of Squares Interactions)

ค่าของ SS_T และค่าที่เกี่ยวข้อง สามารถหาได้จากสูตรต่อไปนี้

$$SS_b = \sum_{j=1}^{k} (T_j^2/bl) - T^2/N$$

$$SS_{bl} = \sum_{i=1}^{bl} (T_i^2/k) - T^2/N$$

$$SS_{a.bl} = \sum_{j=1}^{k}\sum_{i=1}^{bl} X_{ij}^2 - \sum_{j=1}^{k}(T_i^2/bl) \; - \sum_{i=1}^{bl}(T_i^2/k) \; + T^2/N$$

$$SS_T = \sum_{j=1}^{k}\sum_{i=1}^{bl} X_{ij}^2 - T^2/N$$

เมื่อ T_i = ผลรวมของคะแนนทุกตัวที่อยู่ในแถว i

N = จำนวนสมาชิกทั้งหมด มีค่าเท่ากับผลคูณของ bl กับ k

ขั้นตอนการวิเคราะห์ความแปรปรวนวิธี RB-k มีวิธีการเช่นเดียวกันกับวิธี CR-k โดยทำการสุ่มกลุ่มตัวอย่างมาจากประชากรที่มีการแจกแจงแบบปกติในขั้นแรก โดยที่ประชากรต้องมีความแปรปรวนเท่ากัน และไม่มีความสัมพันธ์ระหว่าง Block กับตัวแปรอิสระ หลังจากนั้นจึงกำหนดสมมติฐาน และทำการทดสอบสมมติฐานโดยคำนวณหาค่า F โดยการเปรียบเทียบกับค่าที่ได้จากการเปิดตารางเช่นเดียวกับวิธี CR-k ที่ผ่านมาแล้ว

สูตรการคำนวณหาค่า F

$$F = \frac{MS_b}{MS_{a.bl}}$$

เมื่อ

MS_b = ค่าเฉลี่ยของความแปรปรวนระหว่างกลุ่ม (Mean Squares Between Groups)

$MS_{ab.l}$ = ค่าเฉลี่ยของความแปรปรวนระหว่าง Block (Mean Squares Interactions)

เมื่อ

$$MS_b = SS_b/(k-1)$$

$$MS_{a.bl} = SS_{a.bl}/(k-1)(bl-1)$$

$SS_{a.bl}$ = ผลรวมของการแปรผันระหว่าง Block กับระดับหรือประเภทของตัวแปรอิสระ (Within Block Sum of Squares Interactions)

k-1 = Degree of Freedom ของการแปรผันระหว่างกลุ่ม (df_b)

bl-1 = Degree of Freedom ของการแปรผันระหว่าง Block (df_{bl})

bl = จำนวน Block

ตัวอย่างที่ 13-3

ในการพัฒนาบทเรียนคอมพิวเตอร์วิชาความรู้เบื้องต้นเกี่ยวกับคอมพิวเตอร์ ผู้วิจัยได้พัฒนาบทเรียนแตกต่างกัน 3 แบบ ได้แก่ แบบมัลติมีเดีย (A) แบบไฮเปอร์มีเดีย (B) และแบบเรียนรู้ร่วมกัน (C) เพื่อนำไปทดลองใช้กับผู้เรียน 5 กลุ่ม ๆ ละ 3 คน เพื่อศึกษาผลคะแนนที่ได้จากการใช้บทเรียน เมื่อแบ่งกลุ่มผู้เรียนตามคะแนนที่ได้รับ ปรากฏว่าได้คะแนนดังนี้

	แบบของบทเรียน		
	A	B	C
กลุ่มที่ 1	12	8	8
กลุ่มที่ 2	10	7	8
กลุ่มที่ 3	6	5	5
กลุ่มที่ 4	8	5	7
กลุ่มที่ 5	5	3	4

จากคะแนนที่ปรากฏทั้งหมด จงทดสอบที่ระดับนัยสำคัญ .05 ว่า ผลคะแนนของผู้เรียนโดยการใช้บทเรียนคอมพิวเตอร์ทั้ง 3 แบบแตกต่างกันหรือไม่

สมมติฐานการวิจัย

$H_0 : \mu_A = \mu_B = \mu_C$

$H_1 : \mu_A \neq \mu_B \neq \mu_C$

เมื่อ

μ_A = คะแนนเฉลี่ยของผู้เรียนที่เรียนด้วยบทเรียนแบบ A

μ_B = คะแนนเฉลี่ยของผู้เรียนที่เรียนด้วยบทเรียนแบบ B

μ_C = คะแนนเฉลี่ยของผู้เรียนที่เรียนด้วยบทเรียนแบบ C

กำหนดระดับนัยสำคัญที่ .05

จากโจทย์จะได้ว่า $N = 15, k = 3, bl = 5$ ขั้นแรกหาค่า SS_b, $SS_{a.bl}$, SS_{bl} และ SS_T ก่อน

Block	X_1	X_2	X_3	T_i	T_i^2	T_i^2/k
1	12	8	8	28	784	261.33
2	10	7	8	25	625	208.33
3	6	5	5	16	256	85.33
4	8	5	7	20	400	133.33
5	5	3	4	12	144	48
	41	28	32	101	2209	736.33
T_j^2/bl	336.2	156.8	204.8	$\sum_{j=1}^{k}(T_j^2/bl) = 697.8$		
T_j	8.2	5.6	6.4			
$\sum_{i=1}^{bl} X_{ij}^2$	369	172	218	$\sum_{j=1}^{k}\sum_{i=1}^{bl} X_{ij}^2 = 759$		

จากตาราง $T^2/N = 101^2/15 = 680.06$

แทนค่าต่าง ๆ ลงในสูตร

$$SS_b = \sum_{j=1}^{k}(T_j^2/bl) - T^2/N$$

$$SS_b = 697.8 - 680.06 = 17.74$$

$$SS_{bl} = \sum_{i=1}^{bl}(T_i^2/k) - T^2/N$$

$$SS_{bl} = 736.33 - 680.06 = 56.27$$

$$SS_{a.bl} = \sum_{j=1}^{k}\sum_{i=1}^{bl} X_{ij}^2 - \sum_{j=1}^{k}(T_i^2/bl) - \sum_{i=1}^{bl}(T_i^2/k) + T^2/N$$

$$SS_{a.bl} = 759 - 697.8 - 736.33 + 680.06 = 4.93$$

$$SS_T = \sum_{j=1}^{k}\sum_{i=1}^{bl} X_{ij}^2 - T^2/N$$

$$SS_T = 759 - 680.06 = 78.94$$

ขั้นต่อไป หาค่า df

การแปรผันระหว่างกลุ่ม $df_b = k - 1 = 3 - 1 = 2$

การแปรผันระหว่าง Block $df_{bl} = bl - 1 = 5 - 1 = 4$

การแปรผันของปฏิสัมพันธ์ $df_{a.bl} = (k - 1)(bl - 1) = 2 \times 4 = 8$

การแปรผันทั้งหมด $df_T = N - 1 = 15 - 1 = 14$

ขั้นต่อไป หาค่า MS_b และ $MS_{a.bl}$

$$MS_b = SS_b/(k-1)$$

$$MS_b = 17.74/(3-1) = 8.87$$

$$MS_{a.bl} = SS_{a.bl}/(k-1)(bl-1)$$

หาค่า F $$MS_{a.bl} = 4.93/8 = .61625$$

$$F = \frac{MS_b}{MS_{a.bl}}$$

$$F = \frac{8.87}{.61625} = 14.39$$

ขั้นต่อไป เปิดตารางหาค่า F (หน้า 394) ที่ระดับนัยสำคัญ .05 ที่ df = 2, 8

ปรากฏว่า ค่า F (.05, 2, 8) จากตารางมีค่าเท่ากับ 4.46 ซึ่งมีค่าน้อยกว่าค่าที่คำนวณได้ (14.39) แสดงว่าค่าที่คำนวณได้อยู่นอกเขตยอมรับ (หรืออยู่ในเขตวิกฤติ) จึงปฏิเสธ H_0 และยอมรับ H_1 สรุปตาม H_1 ได้ว่า บทเรียนคอมพิวเตอร์ทั้ง 3 แบบ ได้แก่ แบบมัลติมีเดีย (A) แบบไฮเปอร์มีเดีย (B) และแบบเรียนรู้ร่วมกัน (C) ให้ผลการเรียนแตกต่างกัน (หรือให้ผลการเรียนไม่เท่ากัน) ที่ระดับนัยสำคัญ .05

ในกรณีนี้ สามารถสรุปเป็นตารางที่ระดับนัยสำคัญ .05 ได้ดังนี้

สาเหตุของความแปรปรวน	Sum of Squares	df	Mean Squares	F	ค่าวิกฤติของ F
Between Groups	17.74	2	8.87	14.39*	4.46
Within Block	56.27	4			
Interactions	4.93	8	.61625		
Total		14			

* มีนัยสำคัญที่ระดับ .05

หากทำการทดสอบที่ระดับนัยสำคัญ .01 (หรือที่ระดับความเชื่อมั่น 99%) ปรากฏว่า ค่า F (.01, 2, 8) จากตาราง (หน้า 396) มีค่าเท่ากับ 8.65 ซึ่งมีค่าน้อยกว่าที่คำนวณได้ แสดงว่าค่าที่คำนวณได้ยังคงอยู่นอกเขตยอมรับ (หรืออยู่ในเขตวิกฤติ) จึงปฏิเสธ H_0 และยอมรับ H_1 ที่ระดับนัยสำคัญ .01 ด้วยเช่นกัน

ในกรณีนี้ สามารถสรุปเป็นตารางที่ระดับนัยสำคัญ .01 และ .05 ได้ดังนี้

สาเหตุของความแปรปรวน	Sum of Squares	df	Mean Squares	F	ค่าวิกฤติของ F
Between Groups	17.74	2	8.87		
Within Block	56.27	4		14.39*	8.65
Interactions	4.93	8	.61625	14.39**	4.46
Total		14			

** มีนัยสำคัญที่ระดับ .01 และ * มีนัยสำคัญที่ระดับ .05

■ การวิเคราะห์ความแปรปรวนสองทาง (Two-way ANOVA)

การวิเคราะห์ความแปรปรวนสองทาง หรือ Two-way ANOVA เป็นวิธีการทดสอบเพื่อวิเคราะห์ความสัมพันธ์ระหว่างตัวแปรอิสระหรือตัวแปรต้นที่เป็นสิ่งทดลองจำนวน 2 ตัวกับตัวแปรตามเพียงตัวเดียว โดยที่ตัวแปรอิสระหรือตัวแปรต้นอาจมีลักษณะเชิงคุณภาพที่จำแนกออกเป็นระดับหรือประเภทต่าง ๆ ส่วนตัวแปรตามมีลักษณะเชิงปริมาณ เพื่อศึกษาความสัมพันธ์ของตัวแปรอิสระหรือตัวแปรต้นว่าจะส่งผลอย่างไรกับตัวแปรตาม ตามสมมติฐานการวิจัยที่กำหนดไว้ โดยที่การวิเคราะห์ความแปรปรวนสองทาง นอกจากจะสามารถศึกษาผลของตัวแปรทั้งสองตัวไปพร้อม ๆ กันแล้ว ยังสามารถศึกษาผลร่วม (Interaction) ระหว่างตัวแปรทั้งสองตัวด้วยว่าตัวแปรอิสระหรือตัวแปรต้นตัวหนึ่งนอกจากจะส่งผลต่อตัวแปรตามแล้วยังส่งผลใด ๆ ต่อตัวแปรอิสระหรือตัวแปรต้นอีกตัวหนึ่งหรือไม่

ตัวอย่างหัวข้อการวิจัยที่เหมาะสมสำหรับการวิเคราะห์ความแปรปรวนสองทาง

- การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนของผู้เรียนที่มีความสามารถทางการเรียนแตกต่างกัน และเรียนด้วยบทเรียนคอมพิวเตอร์ที่มีรูปแบบแตกต่างกัน
  ตัวแปรอิสระหรือตัวแปรต้น :
  - ตัวที่ 1 : ความสามารถทางการเรียน (จำแนกออกเป็น 3 กลุ่ม ได้แก่ เก่ง ปานกลาง และอ่อน)
  - ตัวที่ 2 : บทเรียนคอมพิวเตอร์ (จำแนกออกเป็น 4 แบบ ได้แก่ แบบปกติ แบบมัลติมีเดีย แบบไฮเปอร์มีเดีย และแบบเรียนรู้ร่วมกัน)

  ตัวแปรตาม : ผลสัมฤทธิ์ทางการเรียนของผู้เรียน
- การเปรียบเทียบสมรรถนะของไมโครโพรเซสเซอร์ของ Intel, Cyrix, AMD และ Motorola ภายใต้สภาพการใช้งานที่อุณหภูมิแตกต่างกันสองสภาวะ
  ตัวแปรอิสระหรือตัวแปรต้น :

ตัวที่ 1 : ไมโครโพรเซสเซอร์ (จำแนกออกเป็น 4 ชนิด ได้แก่ Intel, Cyrix, AMD และ Motorola)

ตัวที่ 2 : อุณหภูมิ (จำแนกออกเป็น 2 ระดับ ได้แก่ เย็น และร้อน)

ตัวแปรตาม : สมรรถนะ

- การเปรียบเทียบฐานะทางสังคมของประชาชนที่มีระดับการศึกษาแตกต่างกันที่อาศัยอยู่ในจังหวัดชลบุรี ระยอง จันทบุรี ตราด และสระแก้ว

ตัวแปรอิสระหรือตัวแปรต้น :

ตัวที่ 1 : ระดับการศึกษา (จำแนกออกเป็น 3 กลุ่ม ได้แก่ สูง ปานกลาง และต่ำ)

ตัวที่ 2 : จังหวัด (จำแนกออกเป็น 5 กลุ่ม ได้แก่ ชลบุรี ระยอง จันทบุรี ตราด และสระแก้ว)

ตัวแปรตาม : ฐานะทางสังคม

สำหรับการวิเคราะห์ความแปรปรวนสองทางมีหลายวิธี เช่น วิธี CRF และวิธี RMF เป็นต้น ในที่นี้จะขอกล่าวถึงวิธี CRF ซึ่งเป็นวิธีการวิเคราะห์ความแปรปรวนสองทางที่นิยมใช้กันโดยทั่วไป

การวิเคราะห์ความแปรปรวนสองทางวิธี CRF

การวิเคราะห์ความแปรปรวนสองทางวิธี CRF หรือวิธี $CRF_{-a,b}$ เป็นวิธีการวิเคราะห์ความแปรปรวนสองทางทั่ว ๆ ไป โดยที่ CRF ย่อมาจาก Completed Randomized Factorials โดยที่ a เป็นจำนวนของตัวแปรอิสระหรือตัวแปรต้นตัวที่ 1 ส่วน b เป็นจำนวนของตัวแปรอิสระหรือตัวแปรต้นตัวที่ 2 จากตัวอย่างของหัวข้อการวิจัยที่ผ่านมา พบว่าตัวอย่างแรก ตัวแปรอิสระหรือตัวแปรต้นตัวที่ 1 แบ่งออกเป็น 3 กลุ่ม ได้แก่ เก่ง ปานกลาง และอ่อน ส่วนตัวแปรอิสระหรือตัวแปรต้นตัวที่ 2 แบ่งออกเป็น 4 แบบ ได้แก่ แบบปกติ แบบมัลติมีเดีย แบบไฮเปอร์มีเดีย และแบบเรียนรู้ร่วมกัน ดังนั้น เมื่อเขียนอยู่ในรูปแบบของการวิเคราะห์ความแปรปรวนสองทางวิธี CRF จะเขียนได้ว่า $CRF_{3,4}$

สมมติให้ตัวแปรอิสระหรือตัวแปรต้นตัวที่ 1 แทนจำนวนกลุ่มตัวอย่างเป็น A จะได้ว่า

A_1 คือ ผู้เรียนเก่ง

A_2 คือ ผู้เรียนปานกลาง

A_3 คือ ผู้เรียนอ่อน

สมมติให้ตัวแปรอิสระหรือตัวแปรต้นตัวที่ 2 แทนจำนวนรูปแบบเป็น B จะได้ว่า

B_1 คือ บทเรียนคอมพิวเตอร์แบบปกติ

B_2 คือ บทเรียนคอมพิวเตอร์แบบมัลติมีเดีย

B_3 คือ บทเรียนคอมพิวเตอร์แบบไฮเปอร์มีเดีย

B_4 คือ บทเรียนคอมพิวเตอร์แบบเรียนรู้ร่วมกัน

กำหนดให้ . (จุด) แทนค่าที่วัดในตัวแปรในกลุ่มตัวอย่าง การวิเคราะห์ความแปรปรวนสองทางวิธี CRF จะเขียนรูปแบบการวิเคราะห์ได้ว่า

	B_1	B_2	B_3	B_4
A_1	. . . . .	. . . . .	. . . . .	. . . . .
A_2	. . . . .	. . . . .	. . . . .	. . . . .
A_3	. . . . .	. . . . .	. . . . .	. . . . .

ภาพที่ 13-3 รูปแบบการวิเคราะห์ความแปรปรวนสองทางวิธี CRF

สำหรับการวิเคราะห์ความแปรปรวนสองทาง จะต้องทำการวิเคราะห์ผลรวมของกำลังสองของความเบี่ยงเบนของค่าต่าง ๆ ที่เกี่ยวข้อง หรือที่เรียกว่า ผลรวมกำลังสอง (Sum of Squares) ของการแปรผันทั้งหมด โดยกำหนดให้

$$SS_T = SS_A + SS_B + SS_{AB} + SS_W$$

เมื่อ

SS_T = ผลรวมของการแปรผันทั้งหมด (Total Sum of Squares)

SS_A = ผลรวมของการแปรผันจากตัวแปรอิสระตัวที่ 1

SS_B = ผลรวมของการแปรผันจากตัวแปรอิสระตัวที่ 2

SS_{AB} = ผลรวมของการแปรผันร่วม (หรือปฏิสัมพันธ์) ระหว่างตัวแปรอิสระตัวที่ 1 กับ ตัวที่ 2

SS_W = ผลรวมของการแปรผันภายในกลุ่ม (Within Group Sum of Squares)

ในการวิเคราะห์ความแปรปรวนสองทาง มีเงื่อนไขที่เป็นข้อตกลงเบื้องต้นว่า กลุ่มตัวอย่างที่จะต้องคัดเลือกมาจากประชากรที่มีการแจกแจงแบบปกติ โดยมีความแปรปรวนเท่ากัน และกลุ่มตัวอย่างที่ศึกษาจะต้องเป็นอิสระต่อกัน ซึ่งในการคำนวณหาผลรวมของความแปรผันทั้งหมด มีสิ่งที่จะต้องพิจารณาประกอบ ได้แก่ รูปแบบของตัวแปรอิสระหรือตัวแปรต้นที่จะวิเคราะห์ ซึ่งจำแนกเป็น 3 รูปแบบ ได้แก่ แบบสุ่ม (Random Mode) แบบกำหนดเอง (Fixed Mode) และ แบบผสม

(Mixed Mode) ในที่นี้จะกล่าวถึงเฉพาะแบบกำหนดเอง ซึ่งเป็นรูปแบบที่พบโดยทั่วไปในการวิเคราะห์ความแปรปรวนสองทาง

รูปแบบของตัวแปรอิสระหรือตัวแปรต้นที่เรียกว่า แบบกำหนดเอง หรือ Fixed Mode เป็นรูปแบบที่ผู้วิจัยเลือกกำหนดประเภทหรือระดับของตัวแปรอิสระหรือตัวแปรต้นเอง ไม่ได้เกิดจากการสุ่มตามธรรมชาติ เช่น ตัวแปรอิสระหรือตัวแปรต้นตัวที่ 1 แบ่งออกเป็น 3 กลุ่ม ได้แก่ เก่ง ปานกลาง และอ่อน และตัวแปรอิสระหรือตัวแปรต้นตัวที่ 2 แบ่งออกเป็น 4 รูปแบบ ได้แก่ แบบปกติ แบบมัลติมีเดีย แบบไฮเปอร์มีเดีย และแบบเรียนรู้ร่วมกัน ซึ่งการกำหนดประเภทของตัวแปรทั้งหมดนี้ ผู้วิจัยเป็นผู้กำหนดหรือคัดเลือกขึ้นมาเอง ไม่ได้เกิดจากการสุ่มแต่อย่างใด

สูตรในการหาค่า F สำหรับตัวแปรอิสระหรือตัวแปรต้นที่กำหนดเอง มีดังนี้

$$F_A = \frac{MS_A}{MS_W}$$

$$F_B = \frac{MS_B}{MS_W}$$

$$F_{AB} = \frac{MS_{AB}}{MS_W}$$

เมื่อ

$$MS_A = SS_A/(a-1)$$

$$MS_B = SS_B/(b-1)$$

$$MS_{AB} = SS_{AB}/(a-1)(b-1)$$

$$MS_W = SS_W/(N-ab)$$

เมื่อ

MS_A = ค่าเฉลี่ยของความแปรปรวนของ A

MS_B = ค่าเฉลี่ยของความแปรปรวนของ B

MS_{AB} = ค่าเฉลี่ยของความแปรปรวนของปฏิสัมพันธ์ระหว่าง A กับ B

MS_W = ค่าเฉลี่ยของความแปรปรวนภายในกลุ่ม

SS_A = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ของ A

SS_B = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ของ B

SS_{AB} = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ของปฏิสัมพันธ์ระหว่าง A กับ B

SS_W = ผลรวมของการแปรผัน (ผลรวมกำลังสอง) ภายในกลุ่ม

a-1 = Degree of Freedom ของการแปรผันของ A (df_A)

b-1 = Degree of Freedom ของการแปรผันของ B (df_B)

(a-1)(b-1)= Degree of Freedom ของการแปรผันของปฏิสัมพันธ์ระหว่าง A กับ B (df_{AB})

N-ab = Degree of Freedom ของการแปรผันภายในกลุ่ม (df_W)

N-1 = Degree of Freedom ของการแปรผันทั้งหมด (df_T)

ปัจจัยอีกตัวหนึ่งในการคำนวณผลรวมของการแปรผันหรือค่าของ SS ในการวิเคราะห์ความแปรปรวนสองทางวิธี CRF ก็คือ จำนวนกลุ่มตัวอย่างในแต่ละเซล ซึ่งอาจจะมีจำนวนเท่ากันหรือไม่เท่ากันก็ได้ ซึ่งในที่นี้จะขอนำเสนอเฉพาะกรณีที่มีจำนวนกลุ่มตัวอย่างในแต่ละห้องเท่ากัน ซึ่งจะสามารถหาค่า SS ได้จากสูตรดังต่อไปนี้

$$SS_A = \sum^{a} T_i^2 / n_i - T^2/N$$

$$SS_B = \sum^{b} T_j^2 / n_j - T^2/N$$

$$SS_{AB} = \sum^{a}\sum^{b} T_{ij}^2 - \sum^{a} T_i^2 / n_i - \sum^{b} T_j^2 / n_j + T^2/N$$

$$SS_W = \sum^{N} X_{ij}^2 - \sum^{a}\sum^{b} T_{ij}^2 / n_{ij}$$

$$SS_T = \sum^{N} X_{ij}^2 - T^2 / n_{ij}$$

เมื่อ

X_{ij} = คะแนนแต่ละตัว

T_i = ผลรวมของคะแนนในแถว i

T_j = ผลรวมของคะแนนในคอลัมน์ j

T_{ij} = ผลรวมของคะแนนในห้อง I j

T = ผลรวมของคะแนนทั้งหมด

n_i = จำนวนสมาชิกในแถว i

n_j = จำนวนสมาชิกในแถว j

N = จำนวนสมาชิกทั้งหมด

ตัวอย่างที่ 13-4

ในการพัฒนาบทเรียนคอมพิวเตอร์ 2 แบบ ได้แก่ แบบปกติ และแบบมัลติมีเดีย เพื่อนำไปทดลองใช้กับผู้เรียนที่มีความสามารถทางการเรียนแตกต่างกัน 3 กลุ่ม ได้แก่ กลุ่มเก่ง ปานกลาง และอ่อน โดยที่ผู้วิจัยต้องการศึกษาผลสัมฤทธิ์ที่เกิดขึ้นว่าแตกต่างกันหรือไม่อย่างไร ซึ่งผลจากการทดลองใช้บทเรียนคอมพิวเตอร์ทั้ง 2 แบบ กับกลุ่มตัวอย่างที่มีจำนวนกลุ่มละ 4 คน รวมทั้งหมด 24 คน ได้ผลคะแนนปรากฏดังตารางต่อไปนี้

	B_1		B_2		B_3	
A_1	2	3	0	3	9	10
	4	5	6	4	12	15
A_2	2	1	3	6	8	9
	3	2	7	1	11	6

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ .01

ตัวแปรอิสระหรือตัวแปรต้น ตัวที่ 1 จำนวน 2 ตัว ได้แก่

A_1 เป็นบทเรียนคอมพิวเตอร์แบบปกติ

A_2 เป็นบทเรียนคอมพิวเตอร์แบบมัลติมีเดีย

ตัวแปรอิสระหรือตัวแปรต้น ตัวที่ 2 จำนวน 3 ตัว ได้แก่

B_1 เป็นความสามารถทางการเรียนระดับเก่ง

B_2 เป็นความสามารถทางการเรียนระดับปานกลาง

B_3 เป็นความสามารถทางการเรียนระดับอ่อน

ตัวแปรตาม : ผลสัมฤทธิ์ทางการเรียน

สมมติฐานการวิจัย

1. $H_0 : \mu_{A1} = \mu_{A2}$

   $H_1 : \mu_{A1} \neq \mu_{A2}$

2. $H_0 : \mu_{B1} = \mu_{B2} = \mu_{B3}$

   $H_1 : \mu_{B1} \neq \mu_{B2} \neq \mu_{B3}$

3. $H_0 : \mu_{A1B1} = \mu_{A1B2} = \mu_{A1B3} = \mu_{A2B1} = \mu_{A2B2} = \mu_{A2B3}$

   $H_0 : \mu_{A1B1} \neq \mu_{A1B2} \neq \mu_{A1B3} \neq \mu_{A2B1} \neq \mu_{A2B2} \neq \mu_{A2B3}$

กำหนดระดับนัยสำคัญที่ .01 ($\alpha = .01$)

ขั้นแรก หาผลรวมของสมาชิกในแต่ละเซล

	B_1	B_2	B_3	T_i
A_1	14	13	46	73
A_2	8	17	34	59
T_j	22	30	80	132

จากตาราง $\Sigma T_i = \Sigma T_j = 132$

จำนวนสมาชิกทั้งหมด $N = 24$ ดังนั้น $T^2/N = 132^2/24 = 726$

จำนวน $n_i = 12$, $n_j = 8$ และ $n_{ij} = 4$

ขั้นต่อไป หาค่าผลรวมของการแปรผันจากสูตร

$$SS_A = \sum^{a} T_i^2 / n_i - T^2/N$$

$$SS_A = [\frac{73^2}{12} + \frac{59^2}{12}] - 726 = 734.16 - 726 = 8.16$$

$$SS_B = \sum^{b} T_j^2 / n_j - T^2/N$$

$$SS_B = [\frac{22^2}{8} + \frac{30^2}{8} + \frac{80^2}{8}] - 726 = 973 - 726 = 247$$

$$SS_{AB} = \sum^{a}\sum^{b} T_{ij}^2 - \sum^{a} T_i^2 / n_i - \sum^{b} T_j^2 / n_j + T^2/N$$

$$SS_{AB} = [\frac{14^2}{4} + \frac{13^2}{4} + \frac{46^2}{4} + \frac{8^2}{4} + \frac{17^2}{4} + \frac{34^2}{4}] - 734.16 - 973 + 726$$

$$SS_{AB} = 997.50 - 734.16 - 973 + 726 = 16.34$$

$$SS_W = \sum^{N} X_{ij}^2 - \sum^{a}\sum^{b} T_{ij}^2 / n_{ij}$$

$$SS_W = [2^2 + 3^2 + 4^2 + 5^2 + 0^2 + 3^2 + 6^2 + 4^2 + ...6^2] - 997.50$$

$$SS_W = 1080 - 997.50 = 82.50$$

$$SS_T = \sum^{N} X_{ij}^2 - T^2 / n_{ij}$$

$$SS_T = 1080 - 726 = 354$$

ขั้นต่อไป หาค่า df

การแปรผันของ A	df_A	=	$a - 1 = 2 - 1 = 1$
การแปรผันของ B	df_B	=	$b - 1 = 3 - 1 = 2$
การแปรผันของปฏิสัมพันธ์ A, B	df_{AB}	=	$(a - 1)(b - 1) = 1 \times 2 = 2$
การแปรผันภายในกลุ่ม	df_W	=	$N - ab = 24 - 6 = 18$
การแปรผันทั้งหมด	df_T	=	$N - 1 = 24 - 1 = 23$

ขั้นต่อไป หาค่า MS_A, MS_B, MS_{AB} และ MS_W

$$MS_A = SS_A/(a-1)$$

$$MS_A = 8.16/(2-1) = 8.16$$

$$MS_B = SS_B/(b-1)$$

$$MS_B = 247/(3-1) = 247/2 = 123.50$$

$$MS_{AB} = SS_{AB}/(a-1)(b-1)$$

$$MS_{AB} = 16.34/(2-1)(3-1) = 16.34/2 = 8.17$$

$$MS_W = SS_W/(N-ab)$$

$$MS_W = 82.50/(24-6) = 82.50/18 = 4.583$$

ขั้นต่อไป คำนวณหาค่า F

$$F_A = \frac{MS_A}{MS_W} = \frac{8.16}{4.583} = 1.78$$

$$F_B = \frac{MS_B}{MS_W} = \frac{123.50}{4.583} = 26.94$$

$$F_{AB} = \frac{MS_{AB}}{MS_W} = \frac{8.17}{4.583} = 1.78$$

ขั้นต่อไป นำค่า F ที่คำนวณได้ ไปเปรียบเทียบกับค่าที่ได้จากการเปิดตาราง (หน้า 396) แต่ละกรณี ที่ระดับนัยสำคัญ .01 จะมีค่าดังนี้

1. กรณี A เมื่อ df_A (df_1) = 1 และ df_W (df_2) = 18
   ค่าวิกฤติของ F จากการเปิดตารางที่ F (.01, 1,18) มีค่าเท่ากับ 8.29
2. กรณี B เมื่อ df_B (df_1) = 2 และ df_W (df_2) = 18
   ค่าวิกฤติของ F จากการเปิดตารางที่ F (.01, 2,18) มีค่าเท่ากับ 6.01
3. กรณี AB เมื่อ df_{AB} (df_1) = 2 และ df_W (df_2) = 18
   ค่าวิกฤติของ F จากการเปิดตารางที่ F (.01, 2,18) มีค่าเท่ากับ 6.01

ผลการทดสอบสมมติฐานทั้ง 3 ข้อ พบว่าค่า F_A และ F_{AB} ไม่มีนัยสำคัญทางสถิติที่ระดับ .01 เนื่องจากค่าที่คำนวณได้ (1.78) มีค่าต่ำกว่าค่าวิกฤติ (8.28 และ 6.01) จากตาราง จึงไม่มีเหตุผลใด ๆ ที่จะปฏิเสธ H_0 นั่นคือ การปฏิเสธ H_1 และยอมรับ H_0 ส่วนค่าของ F_B (26.94) มีค่ามากกว่าค่าวิกฤติ (6.01) จากการเปิดตารางที่ระดับนัยสำคัญ .01 จึงเป็นการปฏิเสธ H_0 และยอมรับ H_1 ซึ่งสามารถสรุปผลการทดสอบสมมติฐานได้ทั้ง 3 กรณีดังต่อไปนี้

1. บทเรียนคอมพิวเตอร์แบบปกติและบทเรียนคอมพิวเตอร์แบบมัลติมีเดีย ให้ผลสัมฤทธิ์ทางการเรียนที่ไม่แตกต่างกัน (หรือเท่ากัน)

2. กลุ่มผู้เรียนที่มีความสามารถทางการเรียนแตกต่างกัน ได้แก่ กลุ่มเก่ง ปานกลาง และอ่อน จะมีผลสัมฤทธิ์ทางการเรียนแตกต่างกัน (หรือไม่เท่ากัน)

3. ไม่มีผลร่วมหรือปฏิสัมพันธ์ใด ๆ เกิดขึ้น ระหว่างบทเรียนคอมพิวเตอร์แบบปกติกับบทเรียนคอมพิวเตอร์แบบมัลติมีเดีย

ในกรณีนี้ สามารถสรุปเป็นตารางที่ระดับนัยสำคัญ .01 ได้ดังนี้

สาเหตุของความแปรปรวน	Sum of Squares	df	Mean Squares	F	ค่าวิกฤติของ F
A	8.16	1	8.16	1.78	8.29
B	247	2	123.50	26.94**	6.01
AB	16.34	2	8.17	1.78	6.01
W	82.50	18	4.583		
Total	354	23			

** มีนัยสำคัญที่ระดับ .01

■ สหสัมพันธ์ (Correlation)

สหสัมพันธ์ (Correlation) เป็นตัวบ่งชี้ระดับและทิศทางของความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไปว่ามีความเกี่ยวข้องหรือมีความสัมพันธ์กันอย่างไร โดยจะต้องอาศัยการวิเคราะห์สหสัมพันธ์ (Correlation Analysis) ซึ่งเป็นเทคนิคทางสถิติที่ช่วยให้ทราบถึงความสัมพันธ์ของตัวแปรที่ศึกษาดังกล่าว สหสัมพันธ์มีอยู่หลายชนิด ขึ้นอยู่กับจำนวนตัวแปรและลักษณะของการวัดของตัวแปร ได้แก่

1. สหสัมพันธ์อย่างง่าย (Simple Correlation)
2. สหสัมพันธ์คาโนนิคอล (Canonical Correlation)
3. สหสัมพันธ์พหุคูณ (Multiple Correlation)

■ สหสัมพันธ์อย่างง่าย (Simple Correlation)

สหสัมพันธ์อย่างง่าย (Simple Correlation) เรียกอีกอย่างหนึ่งว่า สหสัมพันธ์แบบเพียร์สัน (Pearson Correlation) ใช้ตัวย่อว่า ρ (อ่านว่า Rho) ในกรณีที่เป็นประชากร และใช้ r หรือ r_{XY} ในกรณีที่เป็นกลุ่มตัวอย่าง เป็นดรรชนีที่บ่งชี้ความสัมพันธ์ระหว่างตัวแปร 2 ตัวว่ามีความแปรผันร่วมกันหรือไม่ เพียงใด โดยที่ r มีค่าอยู่ระหว่าง +1.00 ถึง –1.00 ถ้า r มีค่าสูงแสดงว่าตัวแปรทั้ง 2 ตัวนั้นมีการแปรผันตามกันอย่างสมบูรณ์ แต่ถ้า r มีค่าต่ำหรือติดลบ แสดงว่าตัวแปรทั้ง 2 ตัวมีค่าสูงต่ำกลับกัน และถ้า r มีค่าเท่ากับ 0 แสดงว่าตัวแปรทั้ง 2 ตัวไม่มีความสัมพันธ์กัน

ยกตัวอย่างเช่น ต้องการหาสหสัมพันธ์ระหว่างคะแนนผลการทดสอบก่อนบทเรียน (X) กับคะแนนผลสัมฤทธิ์ทางการเรียน (Y) หลังจากเรียนด้วยบทเรียนคอมพิวเตอร์ หากพบว่า r_{XY} มีค่าใกล้ 1.00 ก็ย่อมแสดงว่า หากผู้เรียนคนใดที่ได้คะแนนการทดสอบก่อนบทเรียนสูง ก็น่าจะมีผลสัมฤทธิ์ทางการเรียนสูงขึ้นตามไปด้วย แต่ถ้าหากพบในกรณีตรงกันข้ามว่า r_{XY} มีค่า –1.00 ก็จะหมายความว่า ผู้ที่ได้คะแนนการทดสอบก่อนบทเรียนสูง น่าจะได้ผลสัมฤทธิ์ทางการเรียนต่ำและต่ำกว่าผู้ที่ได้คะแนนการทดสอบก่อนบทเรียนต่ำ

สำหรับการแปลความหมายของค่าสหสัมพันธ์ สามารถกล่าวโดยสรุปได้ว่า ยิ่ง r_{XY} มีค่าสูงขึ้นหรือต่ำลงมากเท่าใดก็จะหมายถึงว่าตัวแปรทั้ง 2 มีความสัมพันธ์กันหรือไม่เกี่ยวข้องกันมากขึ้นเท่านั้น ถ้า r_{XY} ที่คำนวณได้มีค่าบวก ก็แสดงว่าตัวแปรทั้ง 2 ตัวมีความสัมพันธ์กันทางบวกหรือมีความสัมพันธ์ไปในทิศทางเดียวกัน นั่นคือ ถ้าตัวแปรตัวหนึ่งมีค่าเพิ่มขึ้น ตัวแปรอีกตัวหนึ่งก็มีค่าเพิ่มขึ้นด้วย ถ้าตัวแปรตัวหนึ่งมีค่าลดลง ตัวแปรอีกตัวหนึ่งก็จะมีค่าลดลงด้วย แต่ถ้า r_{XY} ที่คำนวณได้มีค่าเป็นลบ ก็จะหมายถึงตัวแปรทั้ง 2 ตัวมีความสัมพันธ์ไปในทิศทางตรงกันข้าม กล่าวคือ ถ้าตัวแปรตัวหนึ่งมีค่าเพิ่มขึ้น ตัวแปรอีกตัวหนึ่งจะมีค่าลดลง และถ้าตัวแปรตัวหนึ่งมีค่าลดลง ตัวแปรอีกตัวหนึ่งก็จะมีค่าเพิ่มขึ้น

ปัญหาประการสำคัญของการแปลความหมายของสหสัมพันธ์ว่ามีค่าเท่าใด จึงจะยอมรับว่ามีความสัมพันธ์กัน ซึ่ง Cohen (Runyon and Other. 1996 : 238 อ้างอิงมาจาก Cohen. 1988) ได้แนะนำว่าสหสัมพันธ์ที่มีขนาดเล็กหรือตัวแปรที่มีความความสัมพันธ์กันน้อย ค่าสหสัมพันธ์จะอยู่ระหว่าง -0.29 ถึง -0.10 หรือ 0.10 ถึง 0.29 ส่วนสหสัมพันธ์ที่มีขนาดปานกลางหรือตัวแปรที่มีความสัมพันธ์กันปานกลาง ค่าสหสัมพันธ์จะอยู่ระหว่าง -0.49 ถึง -0.30 หรือ 0.30 ถึง 0.49 และสหสัมพันธ์ที่มีขนาดใหญ่หรือตัวแปรมีความสัมพันธ์กันสูง ค่าสหสัมพันธ์จะอยู่ระหว่าง -1.00 ถึง -0.50 หรือ 0.50 ถึง 1.00 แต่ Devore and Peck (1993 : 129) ได้แนะนำเกี่ยวกับขนาดของสหสัมพันธ์ไว้ว่า ถ้าตัวแปรมีความสัมพันธ์กันสูง ค่าสหสัมพันธ์จะมีค่าน้อยกว่า -0.80 หรือมีค่ามากกว่า 0.80 ถ้าตัวแปรมีความสัมพันธ์กันปานกลาง ค่าสหสัมพันธ์จะมีค่าอยู่ระหว่าง -0.50 ถึง -0.80 หรือ 0.80 ถึง 0.50 และถ้าตัวแปรมีความสัมพันธ์กันต่ำ ค่าสหสัมพันธ์ควรมีค่าอยู่ระหว่าง -0.50 ถึง 0.50 อย่างไรก็ตามเกณฑ์การพิจารณาอาจขึ้นอยู่ปัจจัยอื่น ๆ ด้วย ซึ่งผู้วิจัยจะต้องศึกษาประเด็นอื่น ๆ ประกอบด้วย

■ การคำนวณหาสัมประสิทธิ์สหสัมพันธ์อย่างง่าย

การคำนวณหาสัมประสิทธิ์สหสัมพันธ์อย่างง่าย มีวิธีการคำนวณอยู่หลายวิธี ได้แก่ การคำนวณจากคะแนนมาตรฐาน การคำนวณจากคะแนนเบี่ยงเบน และการคำนวณจากคะแนนดิบ ซึ่งมีสูตรที่ใช้ในการคำนวณดังนี้

สูตรการคำนวณสัมประสิทธิ์สหสัมพันธ์จากคะแนนมาตรฐาน

$$r_{XY} = \frac{\sum Z_X Z_Y}{N} \qquad \text{(กรณีที่ N มีจำนวนมาก)}$$

หรือ

$$r_{XY} = \frac{\sum Z_X Z_Y}{N-1} \qquad \text{(กรณีที่ N มีจำนวนน้อย)}$$

เมื่อ

r_{XY} = สัมประสิทธิ์สหสัมพันธ์อย่างง่าย

Z_X = คะแนนมาตรฐานของตัวแปร X

Z_Y = คะแนนมาตรฐานของตัวแปร Y

N = จำนวนข้อมูลทั้งหมด

สูตรการคำนวณสัมประสิทธิ์สหสัมพันธ์จากคะแนนเบี่ยงเบน

$$r_{XY} = \frac{Z_{XY}}{\sqrt{\sum X^2 \sum Y^2}}$$

เมื่อ

r_{XY} = สัมประสิทธิ์สหสัมพันธ์อย่างง่าย

x = คะแนนเบี่ยงเบนจากค่าเฉลี่ยของตัวแปร X มีค่าเท่ากับ $x = X - \mu_X$

y = คะแนนเบี่ยงเบนจากค่าเฉลี่ยของตัวแปร Y มีค่าเท่ากับ $y = Y - \mu_Y$

สูตรการคำนวณสัมประสิทธิ์สหสัมพันธ์จากคะแนนดิบ

$$r_{XY} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\left\{N\sum X^2 - (\sum X)^2\right\}\left\{N\sum Y^2 - (\sum Y)^2\right\}}}$$

เมื่อ

r_{XY} = สัมประสิทธิ์สหสัมพันธ์อย่างง่าย

Σ_X = ผลรวมทั้งหมดของคะแนนของตัวแปร X

Σ_Y = ผลรวมทั้งหมดของคะแนนของตัวแปร Y

N = จำนวนข้อมูลทั้งหมด

ตัวอย่างที่ 13-5

จากการทดลองใช้บทเรียนคอมพิวเตอร์เพื่อการวิจัยของผู้เรียน จำนวน 10 คน ปรากฏผลคะแนนการทดสอบก่อนบทเรียน (X) และคะแนนผลสัมฤทธิ์ทางการเรียน (Y) ดังตาราง ให้หาสัมประสิทธิ์ความสัมพันธ์สหสัมพันธ์ของคะแนนทั้ง 2 ชุด

X	4	6	10	12	8	5	10	5	1	9
Y	11	10	14	16	9	14	14	11	5	16
คนที่	1	2	3	4	5	6	7	8	9	10

จากสูตร $$r_{XY} = \frac{N\sum XY - (\sum X)(\sum Y)}{\sqrt{\left\{N\sum X^2 - (\sum X)^2\right\}\left\{N\sum Y^2 - (\sum Y)^2\right\}}}$$

คนที่	X	Y	XY	X^2	Y^2
1	4	11	44	16	121
2	6	10	60	36	100
3	10	14	140	100	196
4	12	16	192	144	256
5	8	9	72	64	81
6	5	14	70	25	196
7	10	14	140	100	196
8	5	11	55	25	121
9	1	5	5	1	25
10	9	16	144	81	256
	70	120	922	592	1548

แทนค่าในสูตร

$$r_{XY} = \frac{9220 - 8400}{\sqrt{(5920 - 4900)(15480 - 14400)}} = \frac{800}{\sqrt{(1020)(1080)}}$$

$$r_{XY} = \frac{800}{1049.57} = .7622$$

สัมประสิทธิ์สหสัมพันธ์ ระหว่างคะแนนการทดสอบก่อนบทเรียนและคะแนนผลสัมฤทธิ์ทางการเรียน มีค่าเท่ากับ .7622 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์กันในเชิงบวกระดับมาก

ซึ่งแสดงว่า คะแนนการทดสอบก่อนบทเรียน มีความสัมพันธ์คล้อยตามกันกับคะแนนผลสัมฤทธิ์ทางการเรียน

■ การทดสอบนัยสำคัญของสัมประสิทธิ์สหสัมพันธ์อย่างง่าย

การคำนวณค่าสัมประสิทธิ์สหสัมพันธ์อย่างง่ายโดยใช้สูตรต่าง ๆ ที่ผ่านมา จะเป็นการคำนวณในเชิงของการประมาณ สำหรับในการวิจัยนั้นจำเป็นต้องมีการทดสอบนัยสำคัญทางสถิติของสหสัมพันธ์ที่ได้จากการคำนวณว่า สัมประสิทธิ์สหสัมพันธ์ได้มาจากประชากรที่ตัวแปรทั้งสองมีการแปรผันร่วมกันจริงหรือไม่ เพียงใด หากทดสอบนัยสำคัญแล้วพบว่ามีความแปรผันร่วมกันจริงก็จะเป็นสิ่งที่ช่วยยืนยันการค้นพบดังกล่าว แต่ถ้าหากผลการทดสอบไม่พบว่ามีการแปรผันร่วมกัน ก็แสดงว่าเกิดความคลาดเคลื่อนในการคำนวณ นอกจากนี้การทดสอบนัยสำคัญจะเป็นตัวบ่งชี้ระดับของการแปรผันร่วมกันระหว่างตัวแปรทั้งสอง ว่ามีสหสัมพันธ์กันในระดับนัยสำคัญที่เท่าใด

df	ระดับนัยสำคัญ (สองหาง)		Df	ระดับนัยสำคัญ (สองหาง)		df	ระดับนัยสำคัญ (สองหาง)	
N-2	.05	.01	N-2	.05	.01	N-2	.05	.01
1	.997	1.000	17	.456	.575	45	.288	.373
2	.950	.990	18	.444	.561	50	.273	.354
3	.878	.959	19	.432	.549	60	.250	.325
4	.811	.917	20	.423	.537	70	.232	.302
5	.754	.874	21	.413	.526	80	.217	.283
6	.707	.834	22	.404	.515	90	.205	.267
7	.666	.798	23	.396	.505	100	.195	.254
8	.632	.765	24	.388	.496	125	.174	.228
9	.606	.735	25	.381	.487	150	.159	.208
10	.576	.708	26	.374	.478	200	.138	.181
11	.553	.684	27	.367	.470	300	.113	.148
12	.532	.661	28	.361	.463	400	.098	.128
13	.514	.641	29	.355	.456	500	.088	.115
14	.497	.623	30	.349	.449	1000	.062	.081
15	.482	.606	35	.325	.418			
16	.468	.590	40	.304	.393			

ตารางที่ 13-1 ตารางค่าวิกฤติของสัมประสิทธิ์สหสัมพันธ์อย่างง่าย (บุญชม. 2547 : 407)

ขั้นตอนการทดสอบนัยสำคัญของสัมประสิทธิ์สหสัมพันธ์อย่างง่าย มีดังนี้ (บุญชม. 2547 : 108 - 109)

1. กำหนดระดับนัยสำคัญ (α) ของการทดสอบว่าจะใช้ระดับใด เช่น .05 หรือ .01 ทั้งนี้ขึ้นอยู่กับความต้องการความมั่นใจ 99% หรือ 95%

2. หาค่า df ซึ่งมีค่าเท่ากับ N-2 (เมื่อ N คือจำนวนสมาชิกในกลุ่มตัวอย่าง)

3. นำค่าระดับนัยสำคัญ (α) และ df ไปเปิดตารางหาค่าวิกฤติของ r_{XY} จากตารางวิกฤติของสัมประสิทธิ์สหสัมพันธ์อย่างง่าย

4. นำค่า r_{XY} ที่คำนวณได้ ไปเปรียบเทียบกับค่าวิกฤติที่ได้จากการเปิดตารางในขั้นที่ 3 ถ้าค่า r_{XY} ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่าวิกฤติ (กรณีที่เป็นค่าบวก) หรือมีค่าน้อยกว่าหรือเท่ากับลบค่าวิกฤติ (กรณีที่เป็นค่าบวก) แสดงว่า r_{XY} มีนัยสำคัญทางสถิติ สามารถสรุปได้ว่าตัวแปรทั้งสองมีความสัมพันธ์กันจริงที่ระดับนัยสำคัญที่กำหนดไว้ แต่ถ้าค่าที่ได้เป็นไม่เป็นไปตามที่ระบุไว้ แสดงว่าตัวแปรทั้งสองไม่มีความสัมพันธ์กันที่นัยสำคัญระดับนั้นแต่อย่างใด

จากตัวอย่างที่ผ่านมา ค่าสัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้มีค่าเท่ากับ .7622 เมื่อ N = 10 เมื่อเปิดตารางที่ df = 8 (N-2) พบว่าค่าวิกฤติมีค่าอยู่ที่ .632 ที่ α = .05 และ .765 ที่ α = .01 จึงสรุปได้ว่า .632 < .7622 < .765 แสดงว่าค่าสัมประสิทธิ์สหสัมพันธ์ดังกล่าวมีนัยสำคัญทางสถิติที่ระดับ .05 แต่ไม่มีนัยสำคัญทางสถิติที่ระดับ .01 สามารถเขียนได้ว่า r_{XY} = .7622*

ข้อสังเกตจากการทดสอบนัยสำคัญของสหสัมพันธ์ มีดังนี้

1. การทดสอบนัยสำคัญ ถ้าหากพบว่าสัมประสิทธิ์สหสัมพันธ์มีนัยสำคัญที่ระดับ .01 ก็ย่อมจะแสดงว่ามีนัยสำคัญที่ระดับ .05 ด้วยเสมอ แต่ถ้าหากผลการทดสอบมีนัยสำคัญที่ระดับ .05 จะไม่ได้หมายความว่ามีนัยสำคัญที่ระดับ .01

2. การที่สัมประสิทธิ์สหสัมพันธ์จะมีนัยสำคัญ ขึ้นอยู่กับปัจจัย 3 ประการด้วยกัน ได้แก่

   2.1 ค่าสัมประสิทธิ์สหสัมพันธ์มีค่าสูงกว่าค่าวิกฤติ (ยกเว้นกรณีเป็นค่าลบ)

   2.2 ระดับนัยสำคัญที่กำหนด เช่น ที่ระดับ .05 หรือ .01 ซึ่งมีค่าวิกฤติไม่เท่ากัน

   2.3 จำนวน N ในกลุ่ม หากพิจารณาในตารางที่ 11-1 จะพบว่า แม้ค่าสัมประสิทธิ์สหสัมพันธ์ที่คำนวณได้จะมีค่าต่ำ แต่ถ้าจำนวน N มีค่ามาก ก็อาจจะมีนัยสำคัญได้ เช่น หาก N มีจำนวนประมาณ 500 จะพบว่า ค่าวิกฤติมีค่าเพียง .088 ที่ระดับ .01 และมีค่าเพียง .115 ที่ระดับ .05 เท่านั้น

3. การเปรียบเทียบกับระหว่างค่าที่คำนวณได้กับค่าวิกฤติในตาราง หากพบว่าไม่มี df ที่ต้องการ เช่น สมมติว่ามีจำนวน N เท่ากับ 80 เมื่อหาค่า df = N - 2 จะเท่ากับ 80 – 2 = 78 ซึ่งไม่มี df ค่านี้ปรากฏในตาราง การหาค่าวิกฤติในกรณีนี้จึงสามารถใช้วิธีเทียบบัญญัติไตรยางค์ได้ โดยพิจารณาที่ df = 70 ซึ่งมีค่าวิกฤติเท่ากับ .232 และที่ df = 80 มีค่าวิกฤติเท่ากับ .217 (ที่ α

= .05) แสดงว่า df เพิ่ม 80 – 70 = 10 ค่าวิกฤติลดลง .232 - .217 = .015 แต่ถ้า df เพิ่มขึ้น 78 – 70 = 8 ค่าวิกฤติจะลดลง .015 x 8 / 10 = .012 ดังนั้น ค่าวิกฤติที่ df = 78 จึงมีค่าเท่ากับ .232 - .012 = 0.220 เป็นต้น

สัมประสิทธิ์สหสัมพันธ์นับว่ามีบทบาทและมีประโยชน์ต่อการวิจัยเป็นอย่างมาก เนื่องจากเป็นดัชนีบ่งชี้ความสัมพันธ์ระหว่างตัวแปร 2 ตัวที่ศึกษาว่ามีความสัมพันธ์กันหรือไม่ เพียงใด แต่การตีความหมายของค่าที่ได้จากสหสัมพันธ์จะต้องพิจารณาให้รอบคอบและระมัดระวัง เช่น ถ้าได้ค่าสัมประสิทธิ์สหสัมพันธ์ +.50 จะไม่ได้หมายความว่า ตัวแปรทั้งสองมีความสัมพันธ์กันในเชิงสัดส่วน 50% นอกจากนี้การวิจัยบางครั้งยังพบด้วยว่า สาเหตุที่แท้จริงของความสัมพันธ์ระหว่างตัวแปร X และตัวแปร Y เกิดจากอิทธิพลของตัวแปรนั้น ๆ เอง เช่น ตัวแปร X เป็นต้นเหตุของการเปลี่ยนแปลงของตัวแปร Y เมื่อ X เปลี่ยนจึงทำให้ Y เปลี่ยนแปลงตามไปด้วย หรือในกรณีตรงกันข้าม หรือทั้งตัวแปร X และตัวแปร Y มีความสัมพันธ์กันเนื่องจากอิทธิพลของสิ่งเดียวกันหรือสาเหตุเดียวกัน ซึ่งผู้วิจัยอาจจะไม่ทราบว่าเกิดจากสิ่งใด ทำให้มีค่าสัมประสิทธิ์สหสัมพันธ์สูง ดังนั้น การที่ตัวแปร X และตัวแปร Y มีค่าสัมประสิทธิ์สหสัมพันธ์สูง จะไม่ได้หมายความว่าตัวแปรหนึ่งมีผลทำให้ตัวแปรอีกตัวหนึ่งเปลี่ยนแปลงตาม ค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าสูงจะแสดงเพียงแต่ว่าตัวแปรทั้งสองมีการแปรผันคล้อยตามกันสูงเท่านั้น

จากตัวอย่างแรกที่ผ่านมา พบว่าสัมประสิทธิ์สหสัมพันธ์ระหว่างคะแนนการทดสอบก่อนบทเรียนและคะแนนผลสัมฤทธิ์ทางการเรียนมีค่าเท่ากับ .7622 แสดงว่าตัวแปรทั้งสองมีความสัมพันธ์กันในเชิงบวกระดับมาก หมายความว่า คะแนนการทดสอบก่อนบทเรียนมีความสัมพันธ์คล้อยตามกันกับคะแนนผลสัมฤทธิ์ทางการเรียน แม้ว่าตัวแปรทั้งสองจะมีความสัมพันธ์กันในเชิงบวกดังกล่าว แต่ก็ไม่สามารถสรุปได้อย่างชัดเจนว่าผู้เรียนที่ได้คะแนนการทดสอบก่อนบทเรียนสูงก็จะได้คะแนนผลสัมฤทธิ์ทางการเรียนสูงตามไปด้วย

■ สัมประสิทธิ์เชิงอันดับ

การหาสัมประสิทธิ์สหสัมพันธ์อย่างง่ายที่ผ่านมา เป็นการกระทำกับข้อมูลที่เป็นแบบต่อเนื่องหรือเป็นมาตราการวัดแบบอันตรภาคชั้น แต่ในการวิจัยบางครั้งเป็นการกระทำกับข้อมูลที่อยู่ในลักษณะเชิงอันดับ (Ordinal Scale) เช่น ลำดับความสำคัญ ตำแหน่งที่ ซึ่งไม่สามารถใช้วิธีการหาสัมประสิทธิ์สหสัมพันธ์อย่างง่ายที่ผ่านมาแล้วได้ จึงมีการพัฒนาสูตรสำหรับการหาสัมประสิทธิ์สหสัมพันธ์ของข้อมูลเชิงอันดับขึ้นมาใหม่ โดย Charle Spearman และใช้ตัวย่อว่า r_S หรือ r_{rho} หรือ ρ แทน โดยมีสูตรในการคำนวณดังนี้

$$r_S = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

เมื่อ

r_S = สัมประสิทธิ์สหสัมพันธ์เชิงอันดับของสเปียร์แมน
D = ผลต่างของอันดับแต่ละคู่
N = จำนวนสมาชิกหรือจำนวนคู่

ค่า r_S จะมีค่าเท่ากับ +1.00 เมื่อข้อมูลมีอันดับของแต่ละคู่เท่ากัน และจะมีค่าเป็น –1.00 เมื่อข้อมูลดังกล่าวมีอันดับกลับกัน

ตัวอย่างที่ 13-6

ผลการจัดอันดับการแข่งขันการประกวดซอฟท์แวร์ของโปรแกรมเมอร์จำนวน 10 คน ในด้านเทคนิคและวิธีการ (X) และด้านความคิดสร้างสรรค์ (Y) ปรากฏผลดังตารางต่อไปนี้ จงหาสัมประสิทธิ์สหสัมพันธ์เชิงอันดับของตัวแปรทั้งสอง

ด้านเทคนิคและวิธีการ (X)	2	1	3	5	6	10	7	9	8	4
ด้านความคิดสร้างสรรค์ (Y)	5	3	1	4	10	9	8	7	6	2

ขั้นแรก หาค่า D และ D^2 เพื่อหาค่า ΣD^2

คนที่	X	Y	D	D^2
1	2	5	3	9
2	1	3	-2	4
3	3	1	2	4
4	5	4	1	1
5	6	10	-4	16
6	10	9	1	1
7	7	8	-1	1
8	9	7	2	4
9	8	6	2	4
10	4	2	2	4
				48

ΣD^2 มีค่าเท่ากับ 48 นำไปแทนค่าในสูตร

$$r_S = 1 - \frac{6\sum D^2}{N(N^2-1)}$$

$$r_S = 1 - \frac{6(48)}{10(10^2 - 1)} = 1 - \frac{288}{10(99)} = 1 - .290 = .71$$

ดังนั้น สัมประสิทธิ์สหสัมพันธ์เชิงอันดับมีค่าเท่ากับ .71 แสดงว่าของตัวแปรทั้งสองมีความแปรผันคล้อยตามกันค่อนข้างสูง ซึ่งอาจจะสรุปต่อไปได้ว่าการแข่งขันการประกวดซอฟท์แวร์ด้านเทคนิคและวิธีการมีความสัมพันธ์กับด้านความคิดสร้างสรรค์ในเชิงบวก

■ สหสัมพันธ์คาโนนิคอล (Canonical Correlation)

สหสัมพันธ์คาโนนิคอล (Canonical Correlation) เป็นสหสัมพันธ์ระหว่างกลุ่มตัวแปรอิสระหรือตัวแปรต้นที่มีมากกว่าหนึ่งตัวขึ้นไปกับกลุ่มตัวแปรตามที่มีมากกว่าหนึ่งตัวเช่นกัน โดยทำการศึกษาและเก็บข้อมูลดังกล่าวนั้นเพียงครั้งเดียว หลังจากนั้นจึงทำการวิเคราะห์หาสัมพันธ์ระหว่างตัวแปรที่เกี่ยวข้องพร้อม ๆ กัน เพื่ออธิบายปรากฏการณ์ของตัวแปรต้นกับตัวแปรตามให้มีความชัดเจนเกี่ยวกับความสัมพันธ์ที่เกิดขึ้นระหว่างกลุ่มตัวแปรทั้งหมด

ตัวอย่างเช่น ต้องการศึกษาความสัมพันธ์ระหว่างตัวแปรของความสำเร็จในการประกอบอาชีพคอมพิวเตอร์กับตัวแปรของผลการเรียนวิชาคอมพิวเตอร์ ซึ่งตัวแปรของความสำเร็จในการประกอบอาชีพคอมพิวเตอร์ เป็นตัวแปรต้นหรือตัวแปรอิสระ (X) อาจจะประกอบด้วยหลายตัวแปร เช่น ตำแหน่งหน้าที่ (X_1) อัตราเงินเดือน (X_2) ผลงานการพัฒนาโปรแกรม (X_3) และ การยอมรับของสังคม (X_4) เป็นต้น ส่วนผลการเรียนเป็นตัวแปรตาม (Y) อาจจะประกอบด้วยหลายตัวแปรเช่นกัน ได้แก่ ผลการเรียนวิชาการโปรแกรม (Y_1) ผลการเรียนวิชาฮาร์ดแวร์ (Y_2) และผลการเรียนวิชาซอฟท์แวร์ (Y_3) เป็นต้น

การวิเคราะห์สหสัมพันธ์คาโนนิคอล จะได้ค่าสัมประสิทธิ์สหสัมพันธ์คาโนนิคอลซึ่งเป็นดรรชนีบ่งชี้ความสัมพันธ์ระหว่างกลุ่มตัวแปรต้นหรือกลุ่มตัวแปรอิสระกับกลุ่มตัวแปรตาม อีกทั้งได้ทราบอิทธิพลของตัวแปรต่าง ๆ ที่มีผลต่อค่าสัมประสิทธิ์สหสัมพันธ์คาโนนิคอล อิทธิพลต่าง ๆ ดังกล่าวพิจารณาจากค่าของ Eigen Vectors ซึ่งเป็นน้ำหนักแสดงถึงความสำคัญของตัวแปรที่ทำให้เกิดค่าสัมประสิทธิ์สหสัมพันธ์สูงที่สุด การวิเคราะห์สหสัมพันธ์คาโนนิคอลจึงเป็นการวิเคราะห์หาค่าความสัมพันธ์เชิงเส้น (Linear Correlation) ของกลุ่มตัวแปรที่เกิดขึ้นจากการรวมกันเชิงเส้น (Linear Combination) ตามค่าน้ำหนักของ Eigen Value ที่กำหนดขึ้น

ในการวิจัย การวิเคราะห์หาสหสัมพันธ์คาโนนิคอล จำเป็นต้องใช้คอมพิวเตอร์เพื่อคำนวณหาค่า เนื่องจากเป็นการคำนวณที่ซับซ้อน โดยอาศัยพื้นฐานทางด้านเมตริก (Matrix) เพื่อให้ง่ายต่อความเข้าใจต่อการหาที่มาของสูตร จึงกำหนดเงื่อนไขดังต่อไปนี้

สมมติต้องการศึกษาสหสัมพันธ์คาโนนิคอลของกลุ่มตัวแปรอิสระหรือตัวแปรต้น ซึ่งในที่นี้จะเรียกว่า ตัวพยากรณ์ ที่มีจำนวน p ตัว กับ กลุ่มตัวแปรตาม ซึ่งจะเรียกว่า ตัวเกณฑ์ ที่มีจำนวน q ตัว สามารถเขียนสัญลักษณ์แทนตัวแปรทั้งสองกลุ่มได้ดังนี้ (บุญชม. 2538 : 46 – 58)

$X_1, X_2, X_3, \ldots X_p$ แทนตัวแปรอิสระหรือตัวแปรต้น ที่เป็นตัวพยากรณ์

$Y_1, Y_2, Y_3, \ldots Y_q$ แทนตัวแปรตาม ที่เป็นตัวเกณฑ์

ขั้นต่อไป ใช้กระบวนการของ Least Square เพื่อหาน้ำหนักถดถอย (Regression Weight) ของตัวแปรแต่ละตัวทั้งสองกลุ่ม โดยทำการแปลงคะแนนแต่ละกลุ่มให้อยู่ในรูปเชิงเส้น (Linear Combination) เขียนแทนด้วย Z และ W จะได้ว่า

$Z = u_1X_1 + u_2X_2 + u_3X_3 + \ldots u_pX_p$

$W = v_1Y_1 + v_2Y_2 + v_3Y_3 + \ldots v_qX_q$

โดยที่ Z เป็นคะแนนที่แปลงจากคะแนนชุด X

และ W เป็นคะแนนที่แปลงจากคะแนนชุด Y

เมื่อนำค่า Z และ W มาหาสหสัมพันธ์ จะเรียกว่า สัมประสิทธิ์สหสัมพันธ์คาโนนิคอล โดยใช้ตัวย่อว่า r_{ZW} หรือ r_c หรือ R_c สำหรับสูตรที่ใช้หาค่าสหสัมพันธ์คาโนนิคอล มีดังนี้

$$R_c = \frac{\sum ZW}{\sqrt{(\sum Z^2)(\sum W^2)}}$$

หรือ $$R = R_{yy}^{-1} R_{yx} R_{xx}^{-1} R_{xy}$$

เมื่อคำนวณได้ค่า R แล้ว สามารถหาค่า Eigen Value ได้จากสูตรต่อไปนี้

$$|R - lI| = 0$$

เมื่อ

R	=	เมตริกของสัมประสิทธิ์สหสัมพันธ์ทั้งหมด
R_{yy}^{-1}	=	Inverse ของเมตริกของสัมประสิทธิ์สหสัมพันธ์ระหว่างกันของกลุ่มตัวแปรตาม หรือตัวเกณฑ์
R_{yx}	=	เมตริกของสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรตามแต่ละตัวกับตัวแปรต้นแต่ละตัว
R_{xx}^{-1}	=	Inverse ของเมตริกของสัมประสิทธิ์สหสัมพันธ์ระหว่างกันกลุ่มตัวแปรต้น หรือตัวพยากรณ์
R_{xy}	=	เมตริกของสัมประสิทธิ์สหสัมพันธ์ระหว่างตัวแปรต้นแต่ละตัวกับตัวแปรตามแต่ละตัว
λ	=	Eigen Value
I	=	Identity Matrix

สหสัมพันธ์คาโนนิคอลเป็นการหาความสัมพันธ์ของกลุ่มตัวแปร 2 ตัวที่ประกอบด้วยตัวแปรสมาชิกหลาย ๆ ตัว โดยมีการเก็บข้อมูลเพียงครั้งเดียวและกระทำการวิเคราะห์ข้อมูลพร้อม ๆ กันกับตัวแปรทั้งหมด เพื่อศึกษาความสัมพันธ์ของตัวแปรอิสระหรือตัวแปรต้นกับตัวแปรตาม โดยใช้เมตริก (Matrix) หรือคำนวณด้วยโปรแกรมสำเร็จรูป

■ สหสัมพันธ์พหุคูณ (Multiple Correlation)

สหสัมพันธ์พหุคูณ (Multiple Correlation) ใช้ตัวย่อว่า R เป็นดรรชนีบ่งชี้ความสัมพันธ์ระหว่างตัวแปรตามตัวหนึ่งกับตัวแปรต้นหรือตัวแปรอิสระที่มีมากกว่าหนึ่งตัว ซึ่งแตกต่างจากสหสัมพันธ์คาโนนิคอลที่ทั้งตัวแปรต้นหรือตัวแปรอิสระ กับตัวแปรตาม ต่างก็ประกอบด้วยตัวแปรหลายตัว ตัวอย่างของสหสัมพันธ์พหุคูณ ได้แก่ ต้องการศึกษาความสัมพันธ์ระหว่างผลสัมฤทธิ์ทางการเรียนวิชาคอมพิวเตอร์ ซึ่งเป็นตัวแปรตาม กับ ตัวแปรต้นหรือตัวแปรอิสระ 3 ตัว ได้แก่ ความสนใจทางด้านคอมพิวเตอร์ ความสามารถในการเขียนโปรแกรม และ ประสิทธิภาพของเครื่องคอมพิวเตอร์ที่ใช้งานอยู่ เป็นต้น จะเห็นว่ามีตัวแปรต้นหรือตัวแปรอิสระ 3 ตัว ที่ต้องการศึกษาว่าตัวแปรต้นตัวใดมีผลคล้อยตามกับผลสัมฤทธิ์ทางการเรียนวิชาคอมพิวเตอร์ ซึ่งต้องใช้สหสัมพันธ์พหุคูณในการทดสอบความแปรผันของตัวแปรต่าง ๆ ดังกล่าว

■ บทสรุป

การวิเคราะห์ความแปรปรวน (Analysis of Variance) หรือ ANOVA เป็นสถิติที่มีประโยชน์ต่อการวิจัยมากที่สุดตัวหนึ่ง เพื่อใช้ทดสอบความแตกต่างของค่าเฉลี่ยจากกลุ่มตัวอย่างมากกว่า 2 กลุ่มขึ้นไป โดยใช้กับการวิจัยที่มีตัวแปรอิสระหรือตัวแปรต้นมากกว่าหนึ่งตัวและสามารถใช้ทดสอบตัวแปรอิสระหรือตัวแปรต้นสองตัวว่ามีปฏิสัมพันธ์ร่วมกันหรือไม่ โดยการคำนวณหาค่า F ซึ่งบางครั้งมักจะเรียกการวิเคราะห์ความแปรปรวนว่า การหา F-Test

การวิเคราะห์ความแปรปรวน แบ่งออกเป็น 2 แบบ ได้แก่ การวิเคราะห์ความแปรปรวนทางเดียว หรือ One-way ANOVA กับ การวิเคราะห์ความแปรปรวนสองทาง หรือ Two-way ANOVA การวิเคราะห์ความแปรปรวนทางเดียว ใช้กับการทดสอบความแตกต่างของกลุ่มตัวอย่างที่มีการแจกแจงแบบปกติที่ประกอบด้วยตัวแปรอิสระหรือตัวแปรต้น 1 ตัว กับตัวแปรตาม 1 ตัว ส่วนการวิเคราะห์ความแปรปรวนสองทาง ใช้กับการทดสอบความแตกต่างของกลุ่มตัวอย่างที่มีการแจกแจงแบบปกติที่ประกอบด้วยตัวแปรอิสระหรือตัวแปรต้น 2 ตัว กับตัวแปรตาม 1 ตัว โดยที่ตัวแปรอิสระหรือตัวแปรต้นจะมีการจำแนกออกเป็นประเภทหรือมีระดับต่าง ๆ และส่วนใหญ่เป็นตัวแปรที่ศึกษาเกี่ยวกับคุณภาพ ส่วนตัวแปรตามมักจะเป็นการศึกษาเกี่ยวกับปริมาณ ซึ่งการทดสอบความแตกต่างของค่าเฉลี่ยที่เกิดขึ้นไม่ว่าจะเป็นการวิเคราะห์ความแปรปรวนทางเดียวหรือสองทาง จะเป็นการทดสอบตามสมมติฐานที่กำหนด โดยเปรียบเทียบค่า F ที่คำนวณได้กับค่า F ที่ได้จากตารางตามระดับนัยสำคัญที่กำหนด เพื่อสรุปผลยอมรับหรือปฏิเสธสมมติฐานการวิจัยที่กำหนดไว้

สหสัมพันธ์ (Correlation) เป็นตัวบ่งชี้ระดับและทิศทางของความสัมพันธ์ของตัวแปรตั้งแต่ 2 ตัวขึ้นไปว่ามีความเกี่ยวข้องหรือมีความสัมพันธ์กันหรือไม่ อย่างไร โดยจะต้องอาศัยการวิเคราะห์สหสัมพันธ์ (Correlation Analysis) ซึ่งเป็นวิธีทางสถิติที่ช่วยให้ทราบถึงความสัมพันธ์ของตัวแปรที่ต้องการศึกษา สหสัมพันธ์มีอยู่หลายชนิด ขึ้นอยู่กับจำนวนตัวแปรและลักษณะของการวัดของตัวแปร ได้แก่ สหสัมพันธ์อย่างง่าย สหสัมพันธ์คาโนนิคอล และ สหสัมพันธ์พหุคูณ ค่าที่ได้จากการวิเคราะห์สหสัมพันธ์เรียกว่า สัมประสิทธิ์สหสัมพันธ์ (Correlation Coefficient) การตีความหมายของค่าสัมประสิทธิ์สหสัมพันธ์ จะต้องพิจารณาอย่างรอบคอบและแปลความให้ถูกต้อง ถ้าพบว่าตัวแปรทั้งสองที่ศึกษามีค่าสัมประสิทธิ์สหสัมพันธ์สูง จะไม่ได้หมายความว่าตัวแปรหนึ่งมีผลทำให้ตัวแปรอีกตัวหนึ่งเปลี่ยนแปลงตาม ค่าสัมประสิทธิ์สหสัมพันธ์ที่มีค่าสูงจะแสดงเพียงแต่ว่าตัวแปรทั้งสองที่ศึกษามีการแปรผันคล้อยตามกันสูงเท่านั้น ไม่ได้มีความหมายในเชิงสัดส่วน รวมทั้งไม่ได้หมายความว่าการเปลี่ยนแปลงของตัวแปรหนึ่งทำให้ตัวแปรอีกตัวหนึ่งเปลี่ยนแปลงไป

■ แบบฝึกหัดท้ายบท

จงตอบคำถามต่อไปนี้

1. ความแปรปรวน คืออะไร
2. ลักษณะของข้อมูลเป็นแบบใดที่เหมาะสำหรับการวิเคราะห์ความแปรปรวน
3. การวิเคราะห์ความแปรปรวนทางเดียว เหมาะสำหรับข้อมูลแบบใด
4. การวิเคราะห์ความแปรปรวนสองทาง เหมาะสำหรับข้อมูลแบบใด
5. Mean Squares คืออะไร
6. Sum of Squares คืออะไร
7. Randomized Block หมายความว่าอย่างไร
8. Interactions ในการวิเคราะห์ความแปรปรวนสองทาง หมายถึงอะไร
9. สหสัมพันธ์คาโนนิคอล ใช้สำหรับทดสอบตัวแปรที่มีลักษณะอย่างไร
10. สหสัมพันธ์พหุคูณ ใช้สำหรับทดสอบตัวแปรที่มีลักษณะอย่างไร
11. ให้วิเคราะห์หัวข้อเรื่องการวิจัยต่อไปนี้ว่า จะต้องใช้การทดสอบแบบใด

11.1 การเปรียบเทียบผลสัมฤทธิ์ทางการเรียนและความพึงพอใจของผู้เรียน ที่เรียนด้วยบทเรียน e-Learning แบบ Adaptive และแบบ Non-Adaptive ในวิชาระบบปฏิบัติการ ระดับปริญญาตรี

11.2 การเปรียบเทียบประสิทธิภาพด้านความเร็วของไมโครคอมพิวเตอร์ จำนวน 4 ยี่ห้อ ที่มีจำหน่ายในตลาดคอมพิวเตอร์

11.3 การศึกษาความสัมพันธ์ระหว่างระบบฐานข้อมูลสินค้าคงคลังภายในองค์กร ที่ใช้วิธีการเรียงลำดับข้อมูลแบบ Bubble Sort กับวิธี Shell-Metzner Sort

11.4 การศึกษาตัวแปรที่มีผลต่อการบีบอัดข้อมูลวิธี LZW Method กับตัวแปรที่มีผลต่อการบีบอัดข้อมูลวิธี Cognitive Visual Data Compression Method

11.5 การเปรียบเทียบความคิดเห็นของผู้ใช้ PDA เกี่ยวกับการใช้ PDA เพื่อประกอบการด้านธุรกิจและการใช้งานเพื่อการบันเทิง

12. ในการวิเคราะห์ประสิทธิภาพของฮาร์ดดิสค์ที่มีจำหน่ายในตลาดคอมพิวเตอร์ จำนวน 4 ยี่ห้อ เมื่อนำไปใช้กับโปรแกรมแตกต่างกันเพื่อทดสอบการทำงานจำนวน 3 โปรแกรม พบว่าความเร็วเฉลี่ยในการค้นหาข้อมูล (Average Seek Time) ปรากฏดังตารางข้างล่าง จงทดสอบว่าความเร็วเฉลี่ยในการค้นหาข้อมูล (milli-second) ของฮาร์ดดิสค์ทั้ง 4 ยี่ห้อ แตกต่างกันหรือไม่ ที่ระดับนัยสำคัญ .05

	ยี่ห้อฮาร์ดดิสค์			
	A	B	C	D
โปรแกรม 1	11	8	12.5	9
โปรแกรม 2	10.5	7	9	8
โปรแกรม 3	15	17	12	8.5

13. ระบบสนับสนุนการตัดสินใจระบบหนึ่ง ทำงานได้ทั้งระบบปฏิบัติการ Windows, Linux และ Unix เมื่อนำไปทดลองใช้กับผู้ใช้ในองค์กรแห่งหนึ่งซึ่งผู้ใช้มีคุณลักษณะคล้ายคลึงกัน แต่มีพื้นฐานการใช้คอมพิวเตอร์แตกต่างกัน 3 กลุ่ม ได้แก่ กลุ่ม A, กลุ่ม B และ กลุ่ม C โดยมีจำนวนเท่ากันทุกกลุ่ม รวมทั้งหมด 36 คน หลังจากนั้นได้สอบถามความพึงพอใจในการใช้งาน พบว่าได้คะแนนเฉลี่ยปรากฏดังตาราง จงทดสอบที่ระดับความเชื่อมั่น 95% ว่าความพึงพอใจของผู้ใช้ทั้ง 3 กลุ่มแตกต่างกันหรือไม่

	Windows		Linux		Unix	
A	12	13	9	13	19	14
	14	12	16	14	16	15
B	18	16	12	16	18	17
	13	11	15	11	11	16
C	17	16	14	14	12	19
	12	20	15	18	13	13

14. ในการวิจัยเชิงทดลองทางด้านเทคโนโลยีสารสนเทศ เพื่อหาผลสัมฤทธิ์ของผู้เรียนจากการใช้บทเรียนคอมพิวเตอร์ที่พัฒนาขึ้น โดยใช้แบบแผนการทดลองแบบ One-Group Pretest

Posttest Design กระทำกับผู้เรียนจำนวน 14 คน ผู้วิจัยต้องการศึกษาความสัมพันธ์ระหว่างคะแนนการทดสอบก่อนบทเรียน กับคะแนนการทดสอบหลังบทเรียนของผู้เรียนว่ามีความสัมพันธ์กันหรือไม่ ผลการทดสอบปรากฏว่าได้คะแนนดังตาราง จงหาสัมประสิทธิ์สหสัมพันธ์ของคะแนนทั้ง 2 ชุด พร้อมทั้งทดสอบสมมติฐานที่ระดับนัยสำคัญ .05

O_1	8	9	11	12	9	8	7	7	10	10	12	11	12	9
O_2	14	15	17	15	11	12	12	10	12	16	17	10	15	12

15. ในการสาธิตการใช้ระบบ LMS เพื่อการจัดการเรียนรู้ในองค์กรด้วย e-Training ให้กับผู้บริหารองค์กรจำนวนกลุ่มหนึ่งได้ตัดสินใจเลือกซื้อ ผู้สาธิตได้นำเสนอ LMS จำนวน 2 ระบบ ได้แก่ ระบบ A และระบบ B เพื่อประกอบการตัดสินใจ ผลการจัดอันดับ LMS ทั้ง 2 ระบบจากผู้บริหารกลุ่มดังกล่าว ปรากฏดังตารางต่อไปนี้

LMS ระบบ A	2	1	3	5	9	8	7	6	4	10	3	8
LMS ระบบ B	3	2	7	6	4	8	10	5	1	5	2	6

จงตอบคำถามต่อไปนี้

15.1 จงกำหนดสมมติฐานการวิจัย

15.2 จงหาค่าสัมประสิทธิ์สหสัมพันธ์เชิงอันดับของสเปียร์แมน

15.3 จงทดสอบสมมติฐาน เพื่อหาข้อสรุปให้ผู้บริหารตัดสินใจเลือกซื้อ LMS ระบบ A หรือระบบ B ที่ระดับความเชื่อมั่น 99%