บทที่ 9

การวิเคราะห์ความแปรปรวน (Analysis of Variance: ANOVA)

40503011 สถิติสำหรับวิศวกรและนักวิทยาศาสตร์

9.1 การวิเคราะห์ความแปรปรวน

ในการทดสอบสมมุติฐานเกี่ยวกับค่าเฉลี่ยระหว่างประชากรสองชุดที่เป็นอิสระต่อกัน ตัวสถิติทดสอบที่ใช้คือตัวสถิติ Z หรือตัวสถิติ T ซึ่งขึ้นอยู่กับการแจกแจงของประชากร ขนาดของตัวอย่าง และการทราบค่าความแปรปรวนของประชากรหรือไม่

แต่ถ้ามีตัวอย่างสุ่มจากประชากรมากกว่า 2 ประชากร (k กลุ่ม เมื่อ $k \geq 2$) โดยความแปรปรวนของ k ประชากรไม่ต่างกัน และต้องการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากรทั้ง k กลุ่มพร้อมกัน เทคนิคที่จะนำมาใช้ในการทดสอบสมมติฐานนี้เรียกว่า **การวิเคราะห์ความแปรปรวน**

9.1 การวิเคราะห์ความแปรปรวน

การวิเคราะห์ความแปรปรวนมาจากภาษาอังกฤษว่า Analysis of Variance เขียนแทนด้วย ANOVA ซึ่งเป็นวิธีแยกความผันแปรรวมของข้อมูลออกเป็นส่วนๆ ตามแหล่งที่มา (Source of Variation: SOV) โดยที่มาของความแปรปรวนนั้นแบ่งได้เป็น 2 ลักษณะ คือ ทราบและไม่ทราบสาเหตุ โดยแหล่งความผันแปรที่ทราบสาเหตุอาจมี 1 แหล่งหรือมากกว่าก็ได้ตัวอย่างเช่น

9.1 การวิเคราะห์ความแปรปรวน

1. นักวิชาการเกษตรต้องการเปรียบเทียบผลผลิตของข้าวโพด 3 พันธุ์ จึงทำการทดลองปลูกข้าวโพดทั้ง 3 พันธุ์นั้น พันธุ์ละหลายๆ แปลง โดยแต่ละพันธุ์นั้นพยายามที่จะให้ปุ๋ย น้ำ และการดูแลรักษาให้สม่ำเสมอกัน เมื่อถึงเวลาเก็บผลผลิตได้ข้อมูลดังตารางที่ 9.1

พันธุ์ข้าวโพด		
ก	ข	ค
36	57	50
33	53	41
48	43	47
	54	42
	48	

9.1 การวิเคราะห์ความแปรปรวน

2. อาจารย์ท่านหนึ่งสุ่มนักเรียนที่มีผลการเรียนใกล้เคียงกันทั้งเพศหญิงและเพศชายมาจำนวนหนึ่ง แบ่งเป็น 3 กลุ่ม และทดลองสอนด้วยวิธีการสอน 3 วิธี ได้แก่

กลุ่มที่ 1 สอนเนื้อหาในห้องเรียนและให้นักเรียนทำแบบฝึกหัดด้วยตนเองนอกเวลา

กลุ่มที่ 2 ให้นักเรียนศึกษาเนื้อหาด้วยตนเอง แต่ผู้สอนควบคุมการทำแบบฝึกหัด

กลุ่มที่ 3 สอนเนื้อหาและฝึกทำแบบฝึกหัดภายในเวลา

9.2 การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว (One-way Analysis of Variance)

การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียวเป็นการวิเคราะห์ความแปรปรวน เมื่อทราบแหล่งที่ทำให้เกิดความผันแปรเพียง 1 แหล่ง

ข้อตกลงเบื้องต้น

1. ประชากรทุกกลุ่มมีการแจกแจงแบบปรกติ
2. ประชากรทุกกลุ่มเป็นอิสระต่อกัน
3. ความแปรปรวนของประชากรทุกกลุ่มเท่ากัน

9.2 การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว
(One-way Analysis of Variance)

สมมติฐานของการทดสอบ $H_0 : \mu_1 = \mu_2 = \cdots = \mu_k$

$H_1 :$ มีค่าเฉลี่ยอย่างน้อย 1 คู่ที่ต่างกัน

	ประชากร					
	1	2	3	$\cdots$	k	
	x_{11}	x_{21}	x_{31}	$\cdots$	x_{k1}	
	x_{12}	x_{22}	x_{32}	$\cdots$	x_{k2}	
	$\vdots$	$\vdots$	$\vdots$	$\ddots$	$\vdots$	
	x_{1n_1}	x_{2n_2}	x_{3n_3}	$\cdots$	x_{kn_k}	
ขนาดตัวอย่าง	n_1	n_2	n_3	$\cdots$	n_k	N
ผลรวม	$x_{1.}$	$x_{2.}$	$x_{3.}$	$\cdots$	$x_{k.}$	$x_{..}$
ค่าเฉลี่ย	$\overline{x}_{1.}$	$\overline{x}_{2.}$	$\overline{x}_{3.}$	$\cdots$	$\overline{x}_{k.}$	$\overline{x}_{..}$

9.2 การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว

(One-way Analysis of Variance)

เมื่อ $x_{i.} = \sum_{j=1}^{n_i} x_{ij}$ แทนผลรวมของค่าสังเกตที่เลือกมาจากประชากรที่ i

$\bar{x}_{i.}$ แทนค่าเฉลี่ยของตัวอย่างจากประชากรที่ i

$x_{..}$ แทนผลรวมของค่าสังเกตทั้งหมด

$\bar{x}_{..}$ แทนค่าเฉลี่ยรวม

n_i แทนจำนวนหน่วยทดลองที่ได้รับ treatment i

N แทนจำนวนค่าสังเกตทั้งหมดในตัวอย่าง

ตัวสถิติที่ในการทดสอบ

การหาตัวสถิติทดสอบจะพิจารณาจากการแยกผลรวมกำลังสองของความแตกต่างของค่าเฉลี่ยและค่าสังเกตุออกเป็นส่วน ๆ ได้ดังนี้

$$\sum_{i=1}^{k}\sum_{j=1}^{n_j}(x_{ij}-\bar{x}_{..})^2=\sum_{i=1}^{k}(\bar{x}_{i.}-\bar{x}_{..})^2+\sum_{i=1}^{k}\sum_{j=1}^{n_j}(x_{ij}-\bar{x}_{i.})^2$$

$$SST = SSA + SSW$$

โดยที่

$$SST=\sum_{i=1}^{k}\sum_{j=1}^{n_j}(x_{ij}-\bar{x}_{..})^2=\sum_{i=1}^{k}\sum_{j=1}^{n_j}x_{ij}{}^2-\frac{x_{..}^2}{N} \qquad ;CT=\frac{x_{..}^2}{N}$$

$$SSA=\sum_{i=1}^{k}(\bar{x}_{i.}-\bar{x}_{..})^2=\sum_{i=1}^{k}\frac{x_{i.}^2}{n_i}-\frac{x_{..}^2}{N}$$

$$SSW = SST - SSA$$

ตัวสถิติที่ในการทดสอบ

จากนั้นหาผลรวมกำลังสองเฉลี่ย (Mean square) ได้ดังนี้

$$MSA = \frac{SSA}{k-1}$$

โดย k-1 คือองศาความเป็นอิสระของประชากร

$$MSW = \frac{SSW}{n-k}$$

โดย N-k คือองศาความเป็นอิสระของความคลาดเคลื่อน

และตัวสถิติทดสอบ คือ $F = \dfrac{MSA}{MSW}$

โดยจะปฏิเสธสมมติฐาน H_0 เมื่อ

1. $F > f_{\alpha,(k-1,n-k)}$
2. p – value < α

ตัวสถิติที่ในการทดสอบ

เพื่อความสะดวกในการหาค่าสถิติ F จะเขียนเป็นตารางวิเคราะห์ความแปรปรวน (ANOVA) ได้ดังนี้

SOV	df	SS	MS	F
ระหว่างกลุ่ม (Anomg)	$df_A = k-1$	SSA	$MSA = \frac{SSA}{k-1}$	$F_{cal} = \frac{MSA}{MSW}$
ภายในกลุ่ม (Within)	$df_W = N-k$	SSW	$MSW = \frac{SSW}{N-k}$	
รวม	$df_T = N-1$	SST		

ถ้า $F_{cal} > f_{\alpha,(df_A,df_W)}$ จะปฏิเสธสมมติฐานหลัก (H_0)

ตัวอย่างที่ 9.1 สุ่มผ้าทอด้วยเส้นใยสังเคราะห์ที่มีฝ้ายผสมอยู่ 15, 20, 25, 30 และ 35 เปอร์เซ็นต์ มาอย่างละ 5 ผืน โดยทดสอบสภาพความยืดหยุ่นของผ้าแต่ละผืนได้ผลดังตารางที่ 9.4

ตารางที่ 9.4 ความยืดหยุ่นของผ้าที่มีฝ้ายผสมอยู่แตกต่างกัน

	% ฝ้ายที่ผสมในเนื้อผ้า					
	15	**20**	**25**	**30**	**35**	
	7	12	14	19	7	
	7	17	18	25	10	
	15	12	18	22	11	
	11	18	19	19	15	
	9	18	19	23	11	
ผลรวม	49	77	88	108	54	$X_{..} = 376$
ค่าเฉลี่ย	9.8	15.4	17.6	21.6	10.8	$\overline{X}_{..} = 15.04$

จงทดสอบว่าผ้าที่มีปริมาณฝ้ายผสมอยู่แตกต่างกันมีความยืดหยุ่นแตกต่างกันหรือไม่ ที่ระดับนัยสำคัญ 0.05

วิธีทำ

สมมติฐานการทดสอบ คือ

H_0 : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$ (ผ้าที่มีฝ้ายผสมอยู่แตกต่างกันให้ความยืดหยุ่นไม่ต่างกัน)

H_1 : มีค่าเฉลี่ยอย่างน้อย 1 คู่ที่ต่างกัน

การคำนวณมีดังนี้

$$k = 5, \quad n_1 = n_2 = \cdots = n_5 = 5, \quad df_T = 24, \; df_A = 4, \; df_W = 20$$

$$CT = \frac{376^2}{25} = 5,655.04$$

$$SST = 7^2 + 7^2 + 15^2 + \cdots + 15^2 + 11^2 - CT = 636.96$$

$$SSA = \frac{49^2}{5} + \frac{77^2}{5} + \frac{88^2}{5} + \frac{108^2}{5} + \frac{54^2}{5} - CT = 475.76$$

$$SSW = 636.96 - 475.76 = 161.20$$

$$MSA = \frac{475.76}{4} = 118.94, \quad MSW = \frac{161.20}{20} = 8.06$$

$$F_{cal} = \frac{118.94}{8.06} = 14.76$$

ANOVA

SOV	df	SS	MS	F
ระหว่างกลุ่มที่มีฝ้ายผสมต่างกัน	4	475.76	118.94	14.76
ภายในกลุ่มที่มีฝ้ายผสมอยู่เท่ากัน	20	161.20	8.06	
รวม	24	636.96		

ค่าวิกฤต คือ $f_{0.05,(4,20)}$ = 2.87 เนื่องจาก $F_{cal} > f_{0.05,(4,20)}$ จึงปฏิเสธ H_0 ดังนั้นสรุปว่าผ้าที่มีปริมาณฝ้ายผสมอยู่แตกต่างกัน มีค่าเฉลี่ยของความยืดหยุ่นแตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.05

ตัวอย่าง 1 นักวิชาการผู้วิจัยสนใจศึกษาว่าคะแนนสอบเฉลี่ยวิชาความถนัดเชิงกลของนักศึกษาระดับ ป.วส. ชั้นปีที่ 2 ของวิทยาลัยช่างกล 3 แห่ง แตกต่างกันหรือไม่ โดยได้สุ่มตัวอย่างนักศึกษาวิทยาลัยละ 7 คน การทดสอบปรากฏผลดังนี้

วิทยาลัย		
1	**2**	**3**
20	30	22
31	28	35
25	25	27
29	30	24
30	24	32
22	26	24
27	20	25
184	183	189

$$\sum_{i=1}^{k}\sum_{j=1}^{n_j} x_{ij}^{2} = 15040$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_j} x_{ij} = x.. = 556$$

จงทดสอบสมมุติฐานที่ระดับนัยสำคัญ 0.05

1. ตั้งสมมติฐานในการทดสอบ $H_0 : \mu_1 = \mu_2 = \mu_3$ $H_1 :$ มีค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่าง

2. คำนวณตัวสถิติทดสอบ 2.1 CT=

2.2 SST=

2.3 SSA=

2.4 SSW= SST - SSA =

3. ANOVA

SOV	df	SS	MS	F

4 สรุปผล

ตัวอย่าง 2 โรงงานอุตสาหกรรม 3 โรง ผลิตชิ้นส่วนอิเลคทรอนิคส์สำหรับใช้ในคอมพิวเตอร์ชนิดหนึ่งเหมือนๆ กัน ในการเปรียบเทียบเส้นผ่านศูนย์กลางของชิ้นส่วนดังกล่าวที่ผลิต ได้สุ่มเลือกตัวอย่างชิ้นส่วนที่ผลิตจากโรงงานทั้ง 3 มาเพื่อวัดเส้นผ่านศูนย์กลางผลปรากฏดังนี้

	โรงงานอุตสาหกรรม		
	1	2	3
	1.9	1.2	1.7
	1.8	1.5	1.8
	2.1	1.3	1.6
	2.0	1.7	1.4
	1.7		1.8
	1.7		
รวม	11.2	5.7	8.3

$$\sum_{i=1}^{k}\sum_{j=1}^{n_j} x_{ij}^{\ 2} = 43.2$$

$X_{..} = 25.2$

จงตรวจสอบว่ามีความแตกต่างระหว่างค่าเฉลี่ยเส้นผ่านศูนย์กลางของชิ้นส่วนอีเลคทรอนิคส์ที่ผลิตมาจากแต่ละโรงงานหรือไม่ที่ระดับนัยสำคัญ 0.05

1. ตั้งสมมติฐานในการทดสอบ $H_0 : \mu_1 = \mu_2 = \mu_3$ $H_1 :$ มีค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่าง

2. คำนวณตัวสถิติทดสอบ 2.1 CT=

2.2 SST=

2.3 SSA=

2.4 SSW= SST -SSA =

3. ANOVA

SOV	df	SS	MS	F

4 สรุปผล

9.3 การเปรียบเทียบเชิงซ้อน

จากการวิเคราะห์ความแปรปรวน ถ้าปฏิเสธ H_0 แสดงว่ามีค่าเฉลี่ยของประชากรอย่างน้อย 1 คู่ที่แตกต่างกัน ดังนั้นจึงควรทำการวิเคราะห์ต่อเพื่อพิจารณาว่ามีค่าเฉลี่ยคู่ใดบ้างที่ต่างกัน การทดสอบหลังจากการวิเคราะห์ความแปรปรวน เรียกว่า การเปรียบเทียบเชิงซ้อน (Multiple Comparisons)

การเปรียบเทียบเชิงซ้อนมีหลายวิธี ในที่นี้จะกล่าวถึงเพียงวิธี Least Significant Difference โดยเรียกย่อว่า LSD

วิธี LSD เป็นการทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากรแบบจับคู่ โดยมีจำนวนคู่ทดสอบทั้งหมด ${}^{k}C_{2}$ คู่ แต่ละคู่จะมีสมมติฐานการทดสอบ สูตรการคำนวณรวมทั้งเกณฑ์การตัดสินใจดังนี้

$$\boxed{\begin{array}{ll} H_0 : \mu_i = \mu_j & ; i, j = 1, 2, \ldots, k; i \neq j \\ H_1 : \mu_i \neq \mu_j & \end{array}}$$

9.3 การเปรียบเทียบเชิงซ้อน

การคำนวณค่า LSD ที่ระดับนัยสำคัญ α

$$LSD = t_{\frac{\alpha}{2}, df_W} \sqrt{MSW\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

ถ้า $n_i = n_j = n$ สำหรับทุกค่า $i, j = 1, 2, \ldots, k; i \neq j$ แล้ว

$$LSD = t_{\frac{\alpha}{2}, df_W} \sqrt{\frac{2MSW}{n}}$$

เกณฑ์การตัดสินใจคือ จะปฏิเสธสมมติฐาน $H_0 : \mu_i = \mu_j$ เมื่อ $\left|\bar{X}_{i.} - \bar{X}_{j.}\right| \geq LSD$

จากตัวอย่าง 9.1

% ฝ้ายที่ผสมในเนื้อผ้า				
15	20	25	30	35
7	12	14	19	7
7	17	18	25	10
15	12	18	22	11
11	18	19	19	15
9	18	19	23	11

% ฝ้ายที่ผสมในเนื้อผ้า	15	20	25	30	35
$\overline{X}_{i.}$	9.8	15.4	17.6	21.6	10.8
n_i	5	5	5	5	5

SOV	df	SS	MS	F
ระหว่างกลุ่มที่มีฝ้ายผสมต่างกัน	4	475.76	118.94	14.76
ภายในกลุ่มที่มีฝ้ายผสมอยู่เท่ากัน	20	161.20	8.06	
รวม	24	636.96		

ค่าวิกฤต คือ $f_{0.05,(4,20)} = 2.87$ → $F_{cal} > f_{0.05,(4,20)}$ จึงปฏิเสธ H_0

เนื่องจาก $k = 5$ จึงทำการจับคู่ทดสอบทั้งหมด ${}^5C_2 = 10$ คู่

การคำนวณค่า *LSD*

เนื่องจาก $n_1 = n_2 = n_3 = n_4 = n_5 = 5$ ดังนั้นค่า *LSD* สำหรับใช้เปรียบเทียบในทุกคู่ของค่าเฉลี่ย คือ

$$LSD_{0.05} = t_{0.025,20}\sqrt{\frac{2MSW}{n}} = 2.086\sqrt{\frac{2(8.06)}{5}} = 2.086(1.795) = 3.7$$

เปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยแต่ละคู่ $\left|\overline{X}_{i.} - \overline{X}_{j.}\right|$ กับค่า *LSD*

$$LSD_{0.05} = t_{0.025,20}\sqrt{\frac{2MSW}{n}} = 2.086\sqrt{\frac{2(8.06)}{5}} = 2.086(1.795) = 3.7$$

$H_0: \mu_i = \mu_j,\ H_1: \mu_i \neq \mu_j$	$\lvert \overline{X}_{i.} - \overline{X}_{j.} \rvert$
$H_0: \mu_1 = \mu_2,\ H_1: \mu_1 \neq \mu_2$	$\lvert \overline{X}_{1.} - \overline{X}_{2.} \rvert = \lvert 9.8 - 15.4 \rvert = 5.6^*$
$H_0: \mu_1 = \mu_3,\ H_1: \mu_1 \neq \mu_3$	$\lvert \overline{X}_{1.} - \overline{X}_{3.} \rvert = \lvert 9.8 - 17.6 \rvert = 7.8^*$
$H_0: \mu_1 = \mu_4,\ H_1: \mu_1 \neq \mu_4$	$\lvert \overline{X}_{1.} - \overline{X}_{4.} \rvert = \lvert 9.8 - 21.6 \rvert = 11.8^*$
$H_0: \mu_1 = \mu_5,\ H_1: \mu_1 \neq \mu_5$	$\lvert \overline{X}_{1.} - \overline{X}_{5.} \rvert = \lvert 9.8 - 10.8 \rvert = 1.0$
$H_0: \mu_2 = \mu_3,\ H_1: \mu_2 \neq \mu_3$	$\lvert \overline{X}_{2.} - \overline{X}_{3.} \rvert = \lvert 15.4 - 17.6 \rvert = 2.2$
$H_0: \mu_2 = \mu_4,\ H_1: \mu_2 \neq \mu_4$	$\lvert \overline{X}_{2.} - \overline{X}_{4.} \rvert = \lvert 15.4 - 21.6 \rvert = 6.2^*$
$H_0: \mu_2 = \mu_5,\ H_1: \mu_2 \neq \mu_5$	$\lvert \overline{X}_{2.} - \overline{X}_{5.} \rvert = \lvert 15.4 - 10.8 \rvert = 4.6^*$
$H_0: \mu_3 = \mu_4,\ H_1: \mu_3 \neq \mu_4$	$\lvert \overline{X}_{3.} - \overline{X}_{4.} \rvert = \lvert 17.6 - 21.6 \rvert = 4.0^*$
$H_0: \mu_3 = \mu_5,\ H_1: \mu_3 \neq \mu_5$	$\lvert \overline{X}_{3.} - \overline{X}_{5.} \rvert = \lvert 17.6 - 10.8 \rvert = 6.8^*$
$H_0: \mu_4 = \mu_5,\ H_1: \mu_4 \neq \mu_5$	$\lvert \overline{X}_{4.} - \overline{X}_{5.} \rvert = \lvert 21.6 - 10.8 \rvert = 10.8^*$

* หมายถึงความแตกต่างระหว่างค่าเฉลี่ย $\lvert \overline{X}_{i.} - \overline{X}_{j.} \rvert$ มีค่ามากกว่าค่า LSD จึงปฏิเสธ H_0

เพื่อให้การสรุปผลทำได้ง่ายขึ้น อาจใช้วิธีเรียงลำดับค่าเฉลี่ยจากน้อยไปมากแล้วลากเส้นเชื่อมโยงคู่ของค่าเฉลี่ยที่ไม่แตกต่างกันไว้ด้วยกัน ดังนั้นค่าเฉลี่ยที่อยู่บนเส้นตรงเดียวกันจะไม่ต่างกันทางสถิติ

% ฝ้ายที่ผสมในเนื้อผ้า	15	35	20	25	30
สถิติ	$\overline{X}_{1.}$	$\overline{X}_{5.}$	$\overline{X}_{2.}$	$\overline{X}_{3.}$	$\overline{X}_{4.}$
ค่าเฉลี่ย	9.8	10.8	15.4	17.6	21.6

สรุปผลได้ว่า ที่ระดับนัยสำคัญ 0.05

(1) ผ้าที่มีฝ้ายผสมอยู่ 30% ให้ความยืดหยุ่นดีที่สุดและแตกต่างจากผ้าที่มีฝ้ายผสมอยู่ 15%, 20%, 25%, และ 35 %

(2) ผ้าที่มีฝ้ายผสมอยู่ 20% และ 25% มีความยืดหยุ่นไม่แตกต่างกัน แต่มีความยืดหยุ่นมากกว่าผ้าที่มีฝ้ายผสมอยู่ 15% หรือ 35%

(3) ผ้าที่มีฝ้ายผสมอยู่ 15% มีความยืดหยุ่นไม่แตกต่างจากผ้าที่มีฝ้ายผสมอยู่ 35%

ตัวอย่าง 3 บริษัทผลิตกระดาษแห่งหนึ่งพบว่า ความเหนียวของกระดาษขึ้นอยู่กับความเข้มข้นของเยื่อไม้ที่ใช้ทำเยื่อกระดาษ จึงทำการทดลองผลิตกระดาษโดยใช้ความเข้มข้นของเยื่อไม้ต่างกัน ได้แก่ 5% 10% 15% และ 20% แล้วทำการวัดค่าความเหนียวของกระดาษที่เลือกจากแต่ละกลุ่ม ได้ข้อมูลตามตาราง

	5%	10%	15%	20%
	9	12	13	19
	10	11	15	23
	8	13	15	19
	11	14	17	20
	8	15	17	21
	8			21
total	54	65	77	123
n	6	5	5	6

$$\sum_{i=1}^{k}\sum_{j=1}^{n_j} x_{ij}^{2} = 5079$$

$$\sum_{i=1}^{k}\sum_{j=1}^{n_j} x_{ij} = x_{..} = 319$$

จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.05 ว่าค่าความเหนียวของกระดาษขึ้นกับความเข้มข้นของเยื่อไม้ที่แตกต่างกันหรือไม่

1. ตั้งสมมติฐานในการทดสอบ $H_0 : \mu_1 = \mu_2 = \mu_3$ $H_1 :$ มีค่าเฉลี่ยอย่างน้อย 1 คู่แตกต่าง

2. คำนวณตัวสถิติทดสอบ 2.1 $CT = \frac{319^2}{22} = 4625.5$

2.2 SST= $5079 - 4625.5 = 453.5$

2.3 SSA= $\left(\frac{54^2}{6} + \frac{65^2}{5} + \frac{77^2}{5} + \frac{123^2}{6} \right) - 4625.5 = 412.8$

2.4 SSW= SST - SSA =453.5-412.8=40.7

3. ANOVA

SOV	df	SS	MS	F
ระหว่างกลุ่ม (Among of group)	3	412.8	137.6	60.85
ภายในกลุ่ม (within group)	18	40.7	2.261	
รวม (Total)	21	453.5		

4 สรุปผล

ตัวอย่าง 4 จากตัวอย่างที่ 3 จงทดสอบสมมติฐานว่าระดับความเข้มข้นเยื่อกระดาษระดับใดที่ส่งผลให้ค่าเฉลี่ยความเหนียวของกระดาษแตกต่างกันโดยใช้วิธี LSD ที่ระดับนัยสำคัญ 0.05

Difference of Levels	Difference of Means	$LSD = t_{0.025,18}\sqrt{MSW\left(\frac{1}{n_i}+\frac{1}{n_j}\right)}$
10% - 5%	4.000	$2.101\sqrt{2.261(\frac{1}{6}+\frac{1}{5})} = 1.913$
15% - 5%	6.400	$2.101\sqrt{2.261(\frac{1}{6}+\frac{1}{5})} = 1.913$
20% - 5%	11.500	$2.101\sqrt{2.261(\frac{1}{6}+\frac{1}{6})} = 1.824$
15% - 10%	2.400	$2.101\sqrt{2.261(\frac{1}{5}+\frac{1}{5})} = 1.99$
20% - 10%	7.500	$2.101\sqrt{2.261(\frac{1}{5}+\frac{1}{6})} = 1.913$
20% - 15%	5.100	$2.101\sqrt{2.261(\frac{1}{5}+\frac{1}{6})} = 1.913$

ตัวอย่าง 9.2 ในการทดลองเผาอิฐที่มีความหนาแน่นเท่ากัน 17 ก้อน ด้วยไฟที่มีอุณหภูมิต่างกัน 4 ระดับ คือ 100, 125, 150 และ 175° F แล้ววัดความหนาแน่นของอิฐหลังจากที่ทำการเผาตามกรรมวิธีทำอิฐเรียบร้อยแล้ว จงทดสอบสมมติฐานที่ระดับนัยสำคัญ 0.01 ว่าอุณหภูมิที่ใช้เผาอิฐมีผลกระทบต่อความหนาแน่นของอิฐหลังจากการเผาไฟหรือไม่

อุณหภูมิ (° F)			
100	**125**	**150**	**175**
21.8	22.2	21.9	21.6
21.9	21.9	21.8	21.7
21.7	22.3	21.8	21.8
21.6	22.0	21.6	
21.7		21.5	

การอ่านผลวิเคราะห์จาก EXCEL

ตัวอย่าง 9.2 ในการทดลองเผาอิฐที่มีความหนาแน่นเท่ากัน 17 ก้อน ด้วยไฟที่มีอุณหภูมิต่างกัน 4 ระดับ คือ 100, 125, 150 และ 175° F แล้ววัดความหนาแน่นของอิฐหลังจากที่ทำการเผาตามกรรมวิธีทำอิฐเรียบร้อยแล้ว จงทดสอบสมมติฐานที่**ระดับนัยสำคัญ 0.01** ว่าอุณหภูมิที่ใช้เผาอิฐมีผลกระทบต่อความหนาแน่นของอิฐหลังจากการเผาไฟหรือไม่

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Temp100	5	108.7	21.74	0.013
Temp125	4	88.4	22.1	0.033333
Temp150	5	108.6	21.72	0.027
Temp175	3	65.1	21.7	0.01

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.437647	3	0.145882	**6.773109**	0.005445	5.73938
Within Groups	0.28	13	0.021538			
Total	0.717647	16				

การอ่านผลวิเคราะห์จาก EXCEL

สมมติฐานในการทดสอบ

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

H_1 : มีค่าเฉลี่ยอย่างน้อย 1 คู่ที่แตกต่าง

ค่าสถิติมีค่าเท่ากับ 6.773

ค่าวิกฤต มีค่าเท่ากับ 5.73938

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Temp100	5	108.7	21.74	0.013
Temp125	4	88.4	22.1	0.033333
Temp150	5	108.6	21.72	0.027
Temp175	3	65.1	21.7	0.01

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.437647	3	0.145882	**6.773109**	0.005445	5.73938
Within Groups	0.28	13	0.021538			
Total	0.717647	16				

สามารถสรุปผลได้ว่า การเผาอิฐที่อุณหภูมิต่างกันทำให้ค่าเฉลี่ยความหนาแน่นหลังเผาอิฐแตกต่างกันอย่างมีนัยสำคัญที่ระดับ 0.01

7. สุ่มหัวเทียนที่ใช้กับรถจักรยานยนต์ขนาด 125 ซีซี 4 ยี่ห้อ มายี่ห้อละ 5 อัน เพื่อตรวจสอบคุณภาพของหัวเทียนแต่ละยี่ห้อ โดยวัดระยะทางที่รถวิ่งได้จากการใช้หัวเทียนนั้นๆ และนำข้อมูลระยะทางมาวิเคราะห์ความแปรปรวน โดยผลการวิเคราะห์ข้อมูลบางส่วนแสดงในตาราง ANOVA ข้างล่างนี้

ตารางวิเคราะห์ความแปรปรวนสำหรับระยะทาง (กิโลเมตร)

SOV	df	SS	MS	F
ระหว่างกลุ่มหัวเทียนต่างยี่ห้อ	ก)	ง)	ฉ)	ช)
ภายในกลุ่มหัวเทียนยี่ห้อเดียวกัน	ข)	จ)	14,713.69	
รวม	ค)	310,500.76		

จงหาค่าต่างๆ ในตารางวิเคราะห์ความแปรปรวนข้างต้น และ ซ) สรุปผลการวิเคราะห์ความแปรปรวน ที่ระดับนัยสำคัญ $\alpha = 0.05$

จาก ตัวอย่างเอกสารเพิ่มเติม

แบบฝึกหัดข้อ 4 เพื่อการปรับปรุงคุณภาพของเทปบันทึกเสียง ผู้ผลิตจึงทดลองเคลือบเทปบันทึกเสียงด้วยสารเคลือบที่แตกต่างกัน 4 ชนิด คือ A, B, C และ D หลังจากนั้นทดลองบันทึกเสียงและวัดระดับเสียงที่เพี้ยนเฉลี่ย ได้ข้อมูลดังนี้ **จงทดสอบที่ระดับนัยสำคัญ 0.05**

Anova: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
A	5	60	12	9.5
B	4	68	17	10
C	7	112	16	2
D	6	84	14	2.8

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	72.36364	3	24.12121	4.618956	0.014483	3.159908
Within Groups	94	18	5.222222			
Total	166.3636	21				

จงตอบคำถามต่อไปนี้

1. สมมติฐานในการทดสอบ

----

2. ค่าสถิติทดสอบมีค่าเท่ากับ---

3. ค่าวิกฤตมีค่าเท่ากับ---

4. สามารถสรุปผลการทดสอบสมมติฐานได้ว่า--

ถ้าปฏิเสธสมมติฐาน จงแสดงวิธีการตรวจสอบว่ามีชนิดใดบ้างที่ให้ผลแตกต่างกัน
