

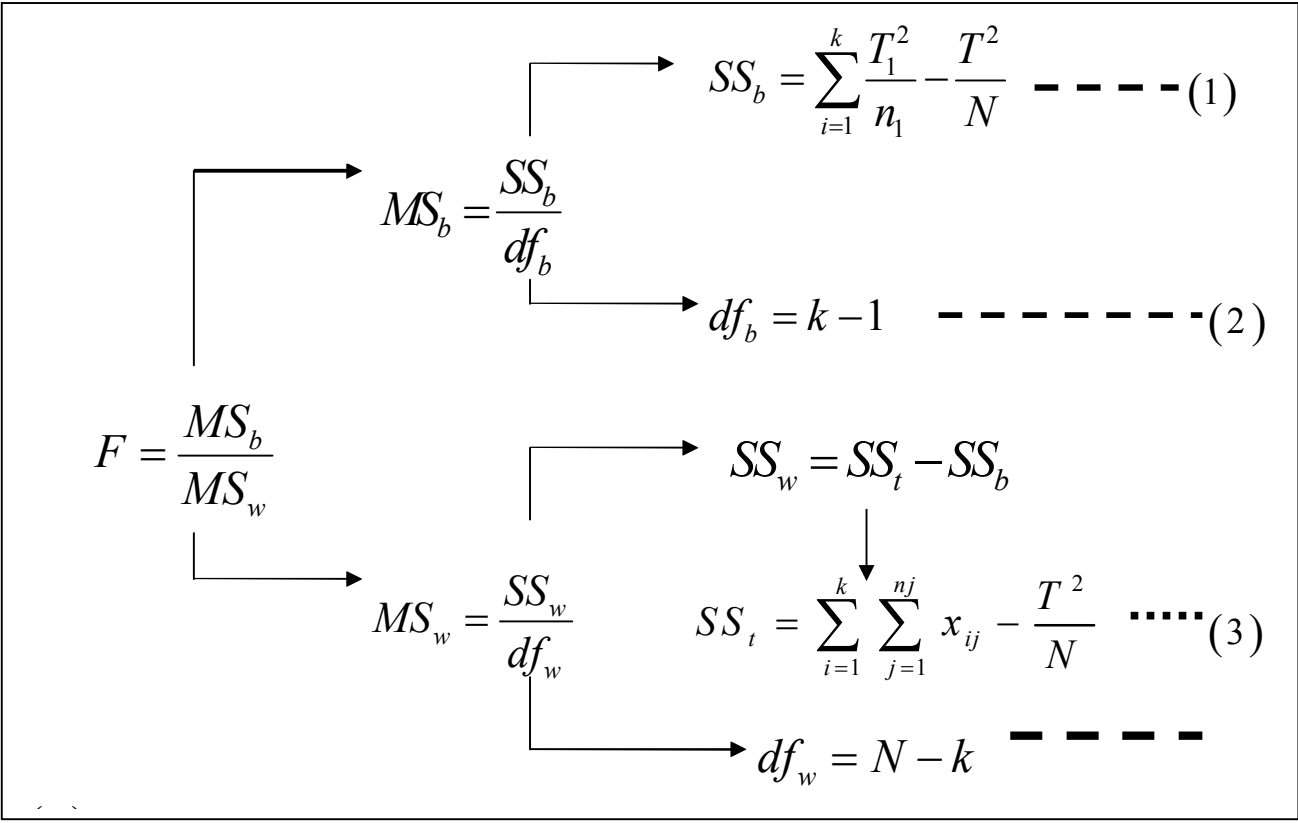
การวิเคราะห์หาค่าความแปรปรวน (Analysis of variance: Anova)

การวิเคราะห์ความแปรปรวน (Analysis of Variance) เป็นวิธีการทดสอบความแตกต่างๆ ระหว่างค่าเฉลี่ยของตัวแปรอิสระที่มีองค์ประกอบย่อยมากกว่า 2 องค์ประกอบขึ้นไป เช่น ตัวแปรสถานภาพสมรส จะมีค่าย่อย 3 ค่า ได้แก่ โสด สมรส และหย่า/หม้าย โดยนำไปวิเคราะห์กับตัวแปรตาม ที่มีระดับการวัดตัวแปรเป็นระดับมาตราอันดับ (Interval Scale) หรือมาตราส่วน (Ratio Scale) การวิเคราะห์ความแปรปรวน ในที่นี้ เราเรียกว่า Anova นั้นจะเป็นการทดสอบค่าความแปรปรวนของค่าเฉลี่ยด้วย F - test ซึ่งมีสูตรดังนี้

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

เนื่องจากการคำนวณ F มีหลายขั้นตอนที่ค่อนข้างยุ่งยาก ดังนั้นอาจารย์จึงขอแสดงเป็นโครงสร้างการคำนวณดังนี้

โครงสร้างการคำนวณค่าความแปรปรวนของค่าเฉลี่ย (Anova) ด้วย F- test



สรุปสูตร

$$F = \frac{\frac{(1)}{(2)}}{\frac{(3)-(1)}{(4)}}$$

เมื่อ

MS_b = ผลรวมกำลัง 2 เฉลี่ยระหว่างกลุ่ม

T_i = ผลรวมในกลุ่มที่ i

MS_w = ผลรวมกำลัง 2 เฉลี่ยภายในกลุ่ม

T = ผลรวมทั้งหมด

SS_b = ผลรวมกำลัง 2 ระหว่างกลุ่ม

n_i = จำนวนตัวอย่างในแต่ละตัวแปร

SS_w = ผลรวมกำลัง 2 ภายในกลุ่ม

N = จำนวนตัวอย่างทั้งหมด

SS_t = ผลรวมกำลัง 2 ทั้งหมด

k = จำนวนตัวแปรทั้งหมด

df_t = ชั้นความเป็นอิสระระหว่างกลุ่ม

X_{ij} = ค่าของตัวอย่างที่ i ในตัวแปรที่ j

df_w = ชั้นความเป็นอิสระภายในกลุ่ม

การแสดงค่าต่างๆ ที่จะต้องใช้ในการคำนวณจัดอยู่ในรูปของตารางได้ดังนี้

	ตัวแปรที่ 1	ตัวแปรที่ 2	...	ตัวแปรที่ k	
	X_{11}	X_{21}	...	X_{k1}	
	X_{12}	X_{22}	...	X_{k2}	
	
	
	X_{1n_1}	X_{1n_2}	...	X_{kn_k}	
ขนาดตัวอย่าง (N_i)	n_1	n_2	...	n_k	N
ผลรวม (T_i)	T_1	T_2	...	T_k	T

ตัวอย่างที่ 1 แสดงข้อมูลกลุ่มตัวอย่าง 30 คนที่เก็บได้จากงานวิจัย เรื่องความพึงพอใจในการรับฟังวิทยุคลื่น จส.1500 ที่จะนำไปใช้คำนวณค่าสถิติอ้างอิงด้วยมือ โดยได้นำเอาข้อมูลอายุของผู้ตอบแบบสอบถามมาจัดเป็น 3 กลุ่มอายุ โดยกลุ่มที่ 1 อายุต่ำกว่า 41 ปี กลุ่มที่ 2 อายุ 41 – 50 ปี และกลุ่มที่ 3 อายุ 50 ปีขึ้นไป โดยต้องการทดสอบว่ากลุ่มอายุที่ต่างกันมีความพึงพอใจในตัวพิธีกรผู้จัดรายการแตกต่างกันหรือไม่ ที่ระดับนัยสำคัญทางสถิติที่ 0.05

คนที่	อายุ	พอใจ พิธีกร	กลุ่มที่	คนที่	อายุ	พอใจ พิธีกร	กลุ่มที่
1	33	5	1	16	36	5	1
2	43	2	2	17	40	4	1
3	47	4	2	18	37	3	1
4	54	3	3	19	50	4	2
5	60	3	3	20	32	3	1
6	45	4	2	21	44	4	2
7	59	3	3	22	49	3	2
8	39	5	1	23	46	1	2
9	49	4	2	24	49	2	2
10	47	3	2	25	48	4	2
11	46	4	2	26	41	5	2
12	52	2	3	27	42	4	2
13	53	3	3	28	52	3	3
14	59	3	3	29	31	4	1
15	37	5	1	30	42	2	2

จากตาราง ผู้ตอบแบบสอบถามคนที่ 1 มีอายุ 33 ปี และมีความพึงพอใจในตัวพิธีกรผู้จัดรายการระดับ 5 โดยนำ 5 ที่ได้มาใส่ในตารางกลุ่มอายุที่ 1 ต่ำกว่า 41 ปี ทำแบบนี้ไปเรื่อยจนถึงคนที่ 30 ผลที่ได้ออกมาดังตารางทางด้านล่าง

	กลุ่มอายุที่ 1 ต่ำกว่า 41 ปี	กลุ่มอายุที่ 2 41 - 50 ปี	กลุ่มอายุที่ 3 50 ปีขึ้นไป	รวม
	5	2	3	
	5	4	3	
	5	4	3	
	5	4	3	
	4	3	3	
	3	4	3	
	3	4	3	
	4	4		
		3		
		1		
		2		
		4		
		5		
		4		
		2		
ขนาดตัวอย่าง (N_i)	8 คน	15 คน	7 คน	30 คน
ผลรวม (T_i)	34	50	21	105
ค่าเฉลี่ย (\bar{x})	4.25	3.33	3.00	

กำหนดสมมติฐานทางสถิติ $H_o : \mu_1 = \mu_2 = \mu_3$

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$

การหาค่า **F** จากการคำนวณ

$$\begin{aligned}
 1. \text{ หาค่า } SS_b \text{ จากสูตร } SS_b &= \sum_{i=1}^k \frac{T_i^2}{n_i} - \frac{T^2}{N} \\
 &= \left(\frac{34^2}{8} + \frac{50^2}{15} + \frac{21^2}{7} \right) - \frac{105^2}{30} \\
 &= (14435 + 166.6667 + 63) - 367.5 \\
 SS_b &= 6.6667
 \end{aligned}$$

2. หาค่า SS_t จากสูตร $SS_t = \sum_{i=1}^k \sum_{j=1}^{nj} x_{ij} - \frac{T^2}{N}$

$$= \left(\begin{array}{l} 5^2 + 5^2 + 5^2 + 5^2 + 4^2 + 3^2 + 3^2 + 4^2 + 2^2 + 4^2 + 4^2 + 4^2 + 3^2 + 4^2 + 4^2 \\ + 4^2 + 3^2 + 1^2 + 2^2 + 4^2 + 5^2 + 4^2 + 2^2 + 3^2 + 3^2 + 3^2 + 3^2 + 3^2 + 3^2 + 3^2 \end{array} \right) - \frac{105^2}{30}$$

$$= (397) - \frac{105^2}{30}$$

$$= 397 - 367.5$$

$$SS_t = 29.5$$

3. หาค่า SS_w จากสูตร $SS_w = SS_t - SS_b$

$$= 29.5 - 6.6667$$

$$SS_w = 22.8333$$

4. หาค่า df_b จากสูตร $df_b = k - 1$

$$= 3 - 1$$

$$df_b = 2$$

5. หาค่า df_w จากสูตร $df_w = N - k$

$$= 30 - 3$$

$$df_w = 27$$

6. หาค่า $MS_b = \frac{SS_b}{df_b}$

$$= \frac{6.6667}{2}$$

$$MS_b = 3.3334$$

7. หาค่า MS_w จากสูตร $MS_w = \frac{SS_w}{df_w}$

$$= \frac{22.8333}{27}$$

$$MS_w = 0.8457$$

8. แทนค่าหาค่า F จากการคำนวณ $F = \frac{MS_b}{MS_w}$

$$= \frac{3.3334}{0.8457}$$

ค่า F จากการคำนวณ $= 3.9416$

โดยเราสามารถเขียนสรุปค่าที่ได้ทั้งหมดมาเขียนเป็นตาราง ดังนี้

แหล่งความแปรปรวน	SS	df	MS	F
ระหว่างกลุ่ม (b)	6.6667	2	3.3334	**3.9416
ภายในกลุ่ม (w)	22.8333	27	0.8457	
รวม	29.5000	29		

การหาค่า F จากตาราง

การเปิดค่า F จากตาราง เราต้องทราบค่า 2 ค่าได้แก่

1. ระดับนัยสำคัญทางสถิติ กำหนดให้ $\alpha = 0.05$ โดยไม่ต้องหาร 2 เพราะเป็นการวิเคราะห์ความแปรปรวนทางเดียว ค่าที่ได้จึงมีค่าตั้งแต่ 0 ขึ้นไป ดังนั้นการตั้งสมมติฐานจะสามารถตั้งสมมติฐานได้แบบทางเดียวเท่านั้น คือ แตกต่างหรือไม่แตกต่าง ซึ่งสอดคล้องกับสมมติฐานด้วยค่า F ที่คำนวณได้จากโปรแกรม SPSS โดยให้ผลเป็น Sig. (ทางเดียว) ไม่ใช่ Sig.2-tailed

2. df (ชั้นความเป็นอิสระ) ทดสอบความแตกต่างระหว่างค่าเฉลี่ยของตัวแปรมากกว่า 2 ตัวขึ้นไป

$$df_1 = \text{ค่าที่คำนวณได้จาก } df_b = 2$$

$$df_2 = \text{ค่าที่คำนวณได้จาก } df_w = 27$$

ตารางแสดงค่าวิกฤตของการแจกแจง F ที่ระดับนัยสำคัญ 0.05 (บางส่วน)

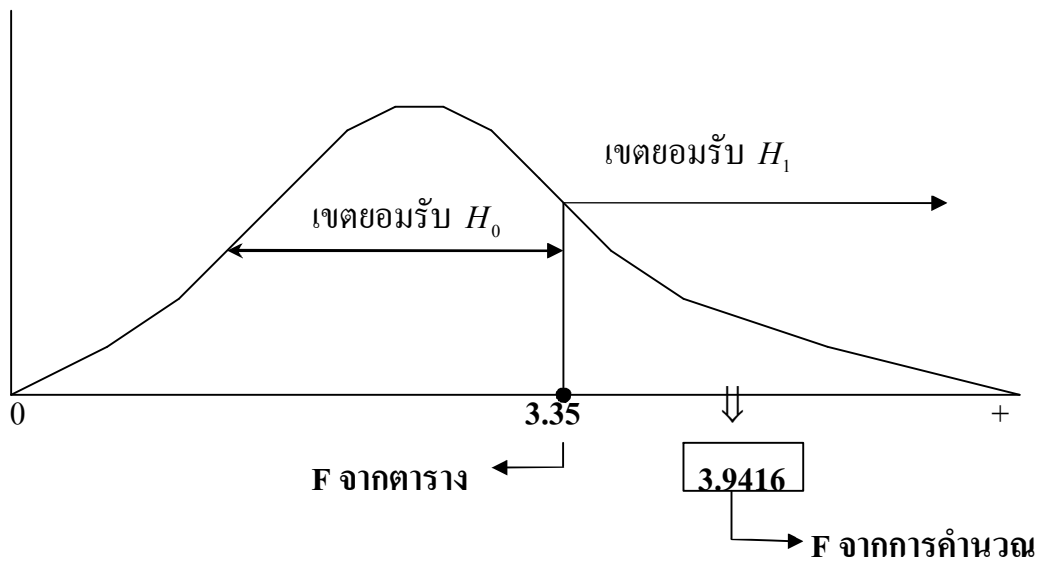
df_2	df_1				
	1	2	3	4	5
25	4.24	3.39	2.99	2.76	2.60
26	4.23	3.37	2.98	2.74	2.59
27	4.21	3.35	2.96	2.73	2.57
28	4.20	3.34	2.96	2.71	2.56

การทดสอบสมมติฐาน

เราต้องการทดสอบว่ากลุ่มอายุที่ต่างกันมีความพึงพอใจในตัวพิธีกรผู้จัดรายการแตกต่างกันหรือไม่

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

$H_1 : \mu_1 \neq \mu_2 \neq \mu_3$



ค่า **F** จากการคำนวณเท่ากับ **3.9416** ซึ่งมีค่ามากกว่าค่า **F** จากตารางซึ่งมีค่าเท่ากับ **3.35** ดังนั้นค่า **F** ที่คำนวณได้จะตกอยู่พื้นที่ยอมรับ H_1 แสดงว่า กลุ่มอายุต่างๆมีความพึงพอใจในตัวพิธีกรผู้จัดรายการแตกต่างกัน อย่างมีนัยสำคัญทางสถิติที่ระดับ **0.05**

แบบฝึกหัดหลังเรียน

การวิเคราะห์หาค่าความแปรปรวน (Analysis of variance: Anova)

ชื่อ.....นามสกุล.....รหัสประจำตัว.....กลุ่ม.....

ในการทำงานวิจัย เรื่อง ความพึงพอใจในการเข้ารับบริการสถานพยาบาล A โดยมีกลุ่มตัวอย่างเป็นประชาชนทั่วไป จำนวน 30 คน ซึ่งผลการวิเคราะห์จะนำไปใช้คำนวณค่าสถิติอ้างอิงด้วยมือ โดยใช้ระยะเวลาการเข้ารับบริการของผู้ป่วยมาจัดเป็น 3 กลุ่มอายุ โดยกลุ่มที่ 1 อายุต่ำกว่า 30 ปี กลุ่มที่ 2 อายุ 31 – 60 ปี และกลุ่มที่ 3 อายุ 60 ปีขึ้นไป โดยต้องการทดสอบว่ากลุ่มอายุที่ต่างกันมีความพึงพอใจในการเข้ารับบริการในสถานพยาบาล A แตกต่างกันหรือไม่ ที่ระดับนัยสำคัญทางสถิติที่ 0.05

คนที่	อายุ	ความพึงพอใจ สถานพยาบาล	คนที่	อายุ	ความพึงพอใจ สถานพยาบาล
1	33	3	16	21	2
2	43	2	17	40	3
3	18	4	18	37	3
4	54	1	19	28	4
5	60	3	20	32	3
6	75	4	21	44	4
7	25	3	22	81	3
8	39	3	23	46	1
9	49	4	24	19	2
10	47	2	25	48	4
11	30	5	26	41	5
12	62	5	27	77	4
13	53	3	28	17	3
14	29	1	29	31	4
15	37	2	30	64	2

จากโจทย์ที่กำหนดให้

1. จงเขียนสมมติฐานทางสถิติ
2. การหาค่า F จากการคำนวณ
3. กราฟผลการวิเคราะห์
4. สรุปผลการวิเคราะห์ข้อมูล

หมายเหตุ : 1. ให้นักศึกษาฝึกทำเอง... ทำใ้สมุด ถ้าวอกกันมาส่งอาจสอบตกได้นะครับ

2. ตารางแสดงค่าวิกฤตของการแจกแจง F ที่ระดับนัยสำคัญ 0.05 ให้ใช้ในตาราง

ตัวอย่าง