10. การแจกแจงแบบไคสแควร์ และการวิเคราะห์ความแปรปรวน

1. การใช้ไคสแควร์ทดสอบความเป็นอิสระ
2. การใช้ไคสแควร์ทดสอบสารรูปสนิทดี
3. การเปรียบเทียบระหว่าง k สัดส่วน
4. การวิเคราะห์ความแปรปรวนแบบจำแนกทางเดียว
5. การวิเคราะห์ความแปรปรวนแบบจำแนก 2 ทาง
6. การอนุมานความแปรปรวนของ 1 ประชากร
7. การอนุมานความแปรปรวนของ 2 ประชากร

10.1 เหตุใดจึงต้องใช้การทดสอบแบบไคสแควร์?

เราใช้การทดสอบแบบ χ^2 สำหรับเปรียบเทียบ k สัดส่วน ซึ่งถ้า $k > 2$ จะใช้การทดสอบแบบ Z ไม่ได้ และยังใช้การทดสอบ χ^2 สำหรับทดสอบความเป็นอิสระกันของ 2 คุณลักษณะ ใช้ทดสอบความแปรปรวนของ 1 ประชากร และใช้สร้างช่วงเชื่อมั่นของความแปรปรวนของประชากร

10.2 เหตุใดจึงต้องทำการวิเคราะห์ความแปรปรวน?

เราใช้การวิเคราะห์ความแปรปรวนเพื่อคำนวณค่า F สำหรับทดสอบความแตกต่างระหว่างค่าเฉลี่ยของประชากร ซึ่งแต่เดิมเราใช้ Z-test และ t-test ทดสอบความแตกต่างของ 2 ประชากร แต่เมื่อเกิน 2 ประชากรต้องใช้ F-test และการสร้างตารางวิเคราะห์ความแปรปรวนเป็นหนทางอันหนึ่ง เพื่อให้ได้ค่าสถิติ F

10.3 เราควรใช้การทดสอบแบบใดสำหรับสภาวการณ์ต่อไปนี้

ก) **ต้องการทราบว่าวิธีการส่งเสริมการขาย 3 วิธี จะทำให้จำนวนขายโดยเฉลี่ยแตกต่างอย่างมีนัยสำคัญหรือไม่?**

ใช้ F - test

ข) **ต้องการทราบว่ายาสีฟัน X ได้รับความนิยมระหว่างหญิงและชายเท่ากันหรือไม่?**

ใช้ Z - test หรือ χ^2 - test

ค) **ต้องการเปรียบเทียบสัดส่วนลูกค้าที่นิยมผงซักฟอก 4 ชนิด**

ใช้ χ^2 - test

10.4 จะใช้การแจกแจงแบบใด หรือการทดสอบแบบใดที่เหมาะสมกับการเปรียบเทียบระหว่างกลุ่มต่าง ๆ ต่อไปนี้

ก) **เปอร์เซนต์แรงงานในกลุ่มอายุ : 16-23, 24-31, 32-39, 40-47, 48-55 และ 50 ขึ้นไป**

ใช้ χ^2 - test และจะมีการแจกแจงแบบ χ^2

ข) รายได้เฉลี่ยของกลุ่มอายุ : 16-23, 24-31, 32-39, 40-47, 48-55, และ 50 ขึ้นไป

ใช้ F - test และมีการแจกแจงแบบ F_{γ_1,γ_2}

ค) รายได้เฉลี่ยของหญิงและชายที่มีอายุ 16 - 56 ปี

ใช้ t - test จะมีการแจกแจงแบบ $t(n_1 + n_2 - 2)$

ใช้ F - test จะมีการแจกแจงแบบ F_{γ_1,γ_2}

ง) ความแปรปรวนหรือการกระจายของรายได้หญิงและชายที่มีอายุ 16 - 56 ปี

ใช้ .F – test จะมีการแจกแจงแบบ $F_{(n_1-1),(n_2-1)}$

10.5 เราจะทราบ df สำหรับเปิดตาราง χ^2 เพื่อทดสอบความเป็นอิสระได้อย่างไร?

ตัวสถิติ X^2 จะมีการแจกแจงแบบ χ^2 ด้วย $df = (r-1)(C-1)$ ในเมื่อ r = จำนวนแถว, C = จำนวนคอลัมน์ ของตารางคอนทินเจ็นซี

10.6 ในการทดสอบความเป็นอิสระ ถ้าค่า X^2 ที่คำนวณได้ เป็นค่าเล็กมาก เราจะปฏิเสธสมมติฐานว่างเปล่าได้ไหม? เพราะเหตุใด?

ถ้าค่า X^2 เป็นค่าเล็กมาก แสดงว่า ค่าสังเกต และค่าคาดหมาย มีความขัดแย้งกันน้อย นั่นคือ ค่าสังเกตมีค่าใกล้เคียงกับค่าทฤษฎีซึ่งหามาจากสมมติฐานว่าลักษณะคู่นั้นเป็นอิสระกัน ดังนั้น จึงจะยอมรับ Ho ว่าคุณลักษณะคู่นั้นเป็นอิสระกัน

10.7 ในการทดสอบความเป็นอิสระ เราจะปฏิเสธ Ho ได้ไหม ถ้า X^2 ที่คำนวณได้เป็นค่าที่โตมาก? เพราะเหตุใด?

ถ้า X^2 เป็นค่าโต แสดงว่า ค่าสังเกตมีความขัดแย้งกับค่าทฤษฎีมาก ค่าทฤษฎีหามาจากสมมติฐานว่างเปล่าว่า คุณลักษณะทั้งคู่เป็นอิสระกัน ดังนั้น ถ้ามีความแตกต่างกับค่าทฤษฎีมาก ข้อมูลจะไปสนับสนุน สมมติฐานรองว่าคุณลักษณะคู่นั้นไม่เป็นอิสระกัน นั่นคือ เราจะปฏิเสธ-Ho เมื่อ X^2 ที่คำนวณได้เป็นค่าโตเกินไป

10.8 ตารางคอนทินเจนซีคืออะไร? สมมติฐานของการทดสอบว่าอย่างไร?

ตารางคอนทินเจนซี คือตารางแยกคุณลักษณะ 2 อย่างไว้ทางด้านแถว (แถวนอน) และคอลัมน์ (แนวตั้ง) เพื่อใช้ทดสอบว่าคุณลักษณะคู่นั้น มีความสัมพันธ์กันหรือไม่ โดยมีสมมติฐานดังนี้

Ho : คุณลักษณะคู่นั้นเป็นอิสระกัน

Ha : คุณลักษณะคู่นั้นไม่เป็นอิสระกัน

10.9 ในการลงคะแนนเสียงกฎหมายฉบับหนึ่ง มีผู้แทนที่ออกเสียงสนับสนุนและคัดค้าน โดยจำแนกตามพรรคการเมือง ดังนี้

การออกเสียง	พรรคร่วมรัฐบาล	พรรคฝ่ายค้าน
คัดค้าน	250	200
สนับสนุน	400	150

จงทดสอบด้วยระดับนัยสำคัญ 5% ว่า ผลการออกเสียงไม่มีความสัมพันธ์กับพรรคการเมือง

1) Ho : ผลการออกเสียง และพรรคการเมืองเป็นอิสระกัน

2) Ha : ผลการออกเสียง และพรรคการเมืองไม่เป็นอิสระกัน

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(2-1)(2-1)=1\ df,.05} = 3.84$

5)

Oij

	พรรคร่วมรัฐบาล	พรรคฝ่ายค้าน	
คัดค้าน	250	200	450
สนับสนุน	400	150	550
	650	350	**1000**

E_{ij}

	พรรคร่วมรัฐบาล	พรรคฝ่ายค้าน	
คัดค้าน	292.5	157.5	450
สนับสนุน	357.5	192.5	550
	650	350	**1000**

$$E_{ij} = \frac{R_i \times C_j}{n} \text{ เช่น } E_{11} = \frac{450 \times 650}{1000} = 292.5$$

$$E_{21} = \frac{550 \times 650}{1000} = 351.5$$

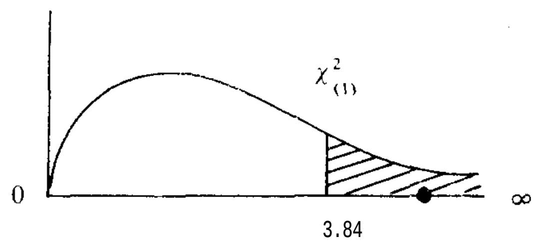
$$E_{12} = \frac{450 \times 350}{1000} = 157.5$$

$$E_{22} = \frac{550 \times 350}{1000} = 192.5$$

$$X^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} (O_{ij} - E_{ij})^2 / E_{ij}$$

$$= \frac{(250 - 292.5)^2}{292.5} + \frac{(200 - 157.5)^2}{157.5} \quad \frac{(400 - 357.5)^2}{357.5} \quad \frac{(150 - 192.5)^2}{192.5}$$

$$= 32.078$$



6) $X^2 = 32.078$ ตกอยู่ในเขตวิกฤต จึงปฏิเสธ Ho และยอมรับ Ha สรุปว่า การออกเสียงมีความสัมพันธ์กับพรรคการเมือง

10.10 ต้องการทดสอบว่าความสัมฤทธิ์ผลในการงานเป็นอิสระกับความสัมฤทธิ์ผลทางการศึกษาหรือไม่ ได้สุ่มพนักงานมา 100 คน และจำแนกใส่ตารางคอนทินเจ็นซีขนาด 3 × 3 ดังนี้

ผลการทำงาน	ความสัมฤทธิผลในการศึกษา						
	A		**B**		**C**		
ดีมาก	10	**(10)**	5	(6)	5	(4)	20
ปานกลาง	20	(20)	12	(12)	8	(8)	40
เลว	20	(20)	13	(12)	7	(8)	40
	50		**30**		**20**		**100**

ถ้าใช้ $\alpha = .05$ จะสรุปว่าความสัมฤทธิ์ผลในการทำงานเป็นอิสระกับความสัมฤทธิ์ผลด้านการศึกษาได้หรือไม่?

1) Ho : ความสัมฤทธิ์ผลในการงาน และการศึกษาเป็นอิสระกัน

2) Ha : ความสัมฤทธิ์ผลในการงานและการศึกษาไม่เป็นอิสระกัน

3) $\alpha = .05$

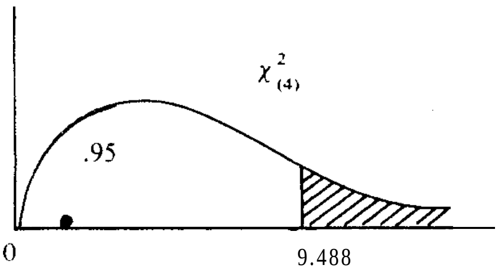
4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(3-1)(3-1) = 4\ df,\ .05} = 9.488$

5) $X^2 = \sum_{i=1}^{3} \sum_{j=1}^{3} (O_{ij} - E_{ij})^2/E_{ij}$

$E_{ij} = \frac{(R_i \times C_j)}{n}$ คือค่าในวงเล็บในตารางคอนทินเจ็นซี ดังนั้น

$$X^2 = \frac{(10 - 10)^2}{10} \quad \frac{(5 - 6)^2}{6} \qquad \frac{(7 - 8)^2}{8}$$

$= 0.63$



6) $X^2 = 0.63$, ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho ว่า ความสัมฤทธิ์ผลในการงานและการศึกษาเป็นอิสระกัน

10.11 ผู้ผลิตสินค้าต้องการทราบว่า ความนิยมสินค้าแบบต่าง ๆ มีความสัมพันธ์กับเพศของลูกค้าหรือไม่ จากการสุ่มผู้ซื้อสินค้าของบริษัทมา 1,000 ราย ได้ข้อมูลดังนี้

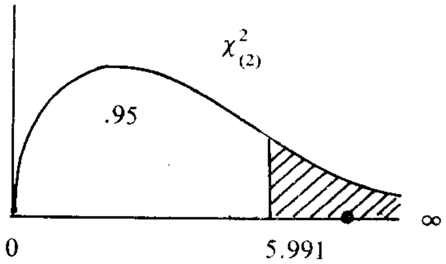
เพศ	แบบที่นิยมซื้อ						รวม
	I		II		III		
ชาย	100	(160)	100	(100)	200	(140)	400
หญิง	300	(240)	150	(150)	150	(210)	600
รวม	400		250		350		1,000

จงใช้ $\alpha = .05$ ทดสอบว่า เพศ และความนิยมไม่มีความสัมพันธ์กัน

1) Ho : เพศและความนิยมไม่มีความสัมพันธ์กัน

2) Ha : เพศและความนิยมมีความสัมพันธ์กัน

3) $\alpha = .05$



4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(2-1)(3-1),.05} = \chi^2_{2,.05} = 5.991$

5) $X^2 = \sum_{i=1}^{2} \sum_{j=1}^{3} (O_{ij} - E_{ij})^2/E_{ij}$ $\boxed{E_{ij} = \frac{R_i \times C_j}{n}}$

$$= \frac{(100-160)^2}{160} + \frac{(100-100)^2}{100} + \ldots + \frac{(150-210)^2}{210}$$

$$= 80.35$$

6) $X^2 = 80.35$ ตกอยู่ในเขตวิกฤต จึงปฏิเสธ Ho และยอมรับ Ha สรุปว่าเพศและความนิยมมีความสัมพันธ์กัน

10.12 **ผู้ผลิตวัคซีนป้องกันโรคหวัด ต้องการทดสอบวัคซีนชนิดใหม่ โดยแบ่งผู้ทดลองเป็น 2 กลุ่ม และฉีดยาให้กลุ่มหนึ่งซึ่งมี 30 คน ส่วนอีก 20 คนไม่ฉีดยา โดยถือเป็นกลุ่ม "ควบคุม"** (control) เพื่อใช้เปรียบเทียบได้ข้อมูลดังนี้

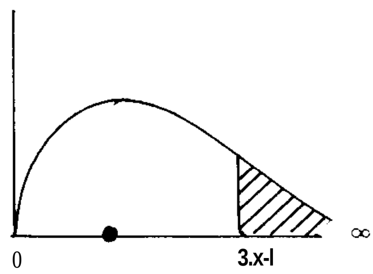
ผล	ฉีดวัคซีน		ไม่ฉีดวัคซีน		รวม
เป็นหวัด	10	(12)	10	(8)	20
ไม่เป็นหวัด	20	(18)	10	(12)	30
รวม	30		20		50

วัคซีนมีผลในการรักษาโรคหวัดหรือไม่? $\alpha = .05$

1) ผลการรักษา และการฉีดยาไม่เกี่ยวข้องกัน

2) ผลการรักษา และการฉีดยามีความสัมพันธ์กัน

3) $\alpha = .05$



4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(1)(1) = 1\ df,\ .05} = 3.84$

5) $X^2 = \sum_{i=1}^{2} \sum_{j=1}^{2} (O_{ij} - E_{ij})^2/E_{ij}$

$E_{ij} = R_i \times C_j/n$ และคือค่าในวงเล็บในตารางคอนทินเจ็นซี

$$X^2 = \frac{(IO - 12)^2}{12} + \frac{(10 - 8)^2}{8} + \frac{(20 - 18)^2}{18} + \frac{(10 - 12)^2}{12}$$

$$= 1.39$$

6) $X^2 = 1.39$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า ผลการรักษาและการฉีดยาไม่เกี่ยวข้องกัน นั่นคือ วัคซีนไม่มีผลในการรักษา

10.13 ในการทดสอบการแจกแจงของข้อมูล (test of goodness of fit) **เรามีวิธีการหา** df **อย่างไร**

ในการทดสอบทั่ว ๆ ไป df = k − 1 แต่การทดสอบการแจกแจงของข้อมูล อาจต้องประมาณค่าพารามิเตอร์บางตัว จึงต้องหัก df เท่ากับจำนวนพารามิเตอร์ที่ต้องประมาณค่าด้วย ดังนั้น

df = k − p − 1

k = จำนวนชั้นหรือลักษณะที่จำแนกข้อมูล

p = จำนวนพารามิเตอร์ที่ต้องประมาณค่า

1 = total

10.14 ในการทดสอบการแจกแจงของข้อมูล ถ้า X^2 ที่คำนวณได้เป็นค่าที่เล็กเกินไป เราจะปฏิเสธ Ho **ได้ไหม? เพราะเหตุใด?**

ในการทดสอบการแจกแจงของข้อมูล (test of goodness of fit) มีสมมติฐาน ดังนี้

1) Ho : ข้อมูลมีการแจกแจงแบบ ....

2) Ha : ข้อมูลไม่มีการแจกแจงแบบ ....

และตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{k} (O_i - E_i)^2/E_i$

ค่า E_i ได้จากสมมติฐานว่างเปล่า นั่นคือ เป็นจำนวนคาดหมายภายใต้การแจกแจงที่อ้างถึง

ดังนั้น ถ้า O_i และ E_i มีค่าใกล้เคียงกัน จะทำให้ X^2 เป็นค่าเล็ก จะสนับสนุน Ho ว่าข้อมูลมีการแจกแจงตามที่อ้าง แต่ถ้า O_i และ E_i มีความขัดแย้งกันสูง จะทำให้ค่าสถิติ X^2 มีค่าสูงด้วย แสดงว่า ข้อมูลไม่มีการแจกแจงตามที่อ้างไว้ใน Ho จึงต้องปฏิเสธ Ho

10.15 ในการทดสอบการแจกแจงของข้อมูล ถ้าเราคำนวณได้ค่า X^2 เป็นค่าที่ใหญ่เกินไป เราจะปฏิเสธ Ho ไหม? เพราะเหตุใด

ข้อนี้อธิบายในข้อ 10.14 แล้ว จึงสรุปรวมกันได้ ดังนี้

ถ้า X^2 ที่คำนวณได้เป็นค่าเล็กเกินไป (เล็กกว่าค่าตาราง) จะยอมรับ Ho

ถ้า X^2 ที่คำนวณได้เป็นค่าโตเกินไป (โตกว่าค่าตาราง) จะปฏิเสธ Ho

10.16 ในการทดสอบการแจกแจงของข้อมูล เราจะใช้การทดสอบแบบ 2 ด้าน คือมีเขตวิกฤตอยู่ 2 ด้านได้ไหม? จงอธิบาย

การทดสอบการแจกแจง เป็นการทดสอบแบบด้านเดียว เพราะในสมมติฐาน Ha ก็มิได้กำหนดทิศทางอะไร เพียงแต่แย้งว่า ข้อมูลไม่มีการแจกแจงตามอ้าง ซึ่งถ้าจะพิจารณาตัวสถิติที่ใช้ทดสอบแล้ว คือ $X^2 = \sum_i (O_i - E_i)^2/E_i$ จะเห็นว่า เราควรจะปฏิเสธ Ho เฉพาะกรณีที่มีความขัดแย้งระหว่าง $O_i - E_i$ มากเกินไปเท่านั้น นั่นคือ เราจะปฏิเสธเฉพาะค่าที่โตเกินไปของตัวสถิติ X^2 เท่านั้น เขตวิกฤตจึงอยู่ด้านเดียวทางปลายหางด้านขวามือ

10.17 ตัวเลขในตารางเลขสุ่ม (random digit table) **ควรมีลักษณะไม่เอียงเฉ เนื่องจากผู้สร้างตารางได้พยายามให้ตัวเลขทุกตัวมีโอกาสปรากฏตัวเท่า ๆ กัน เพื่อที่จะทดสอบคุณสมบัติข้อนี้ จึงได้ทำการสุ่มมา 100 จำนวน จากตารางเลขสุ่ม และนับจำนวนครั้งที่แต่ละตัว**

ปรากฏ มีดังนี้

เลข	0	1	2	3	4	5	6	7	8	9	รวม
จำนวนครั้ง	8	11	10	14	7	12	6	9	13	10	100

เราจะปฏิเสธว่าตัวเลขในตารางไม่เป็นแบบสุ่มที่ระดับนัยสำคัญ 5% ได้ไหม?

1) Ho : ตัวเลขเป็นแบบสุ่ม นั่นคือ ต้องมีการแจกแจงแบบยูนิฟอร์ม ($\pi_i = \frac{1}{10}$)

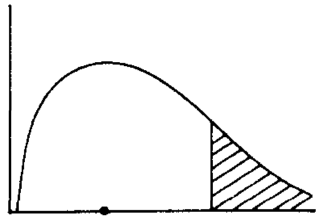
2) Ha : ตัวเลขไม่เป็นแบบสุ่ม นั่นคือ ไม่มีการแจกแจงแบบยูนิฟอร์ม ($\pi_i \neq \frac{1}{10}$)

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(10-1)=9,.05} = 16.919$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{10} (O_i - E_i)^2/E_i$

$E_i \quad n\pi_i = 100\left(\frac{1}{10}\right) = 10$



ดังนั้น $X^2 = \frac{(8-10)^2}{10} + \frac{(11-10)^2}{10} + \ldots + \frac{(10\ 10)^2}{10}$

$= 60/10 = 6.0$

เลข	0	1	2	.	9	10	รวม
O_i	8	11	10	..	13	10	100
E_i	10	10	10	....	10	10	100

6) $X^2 = 6.0$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่าตัวเลขในตารางเป็นแบบสุ่ม

10.18 จำนวนอุบัติเหตุในกรุงเทพฯ ใน 1 สัปดาห์ มีดังนี้

วัน	อาทิตย์	จันทร์	อังคาร	พุธ	พฤหัส	ศุกร์	เสาร์	รวม
จำนวน	28	12	10	7	8	11	24	100
E_i	25	10	10	10	10	10	25	100

จงทดสอบว่าอุบัติเหตุในวันสุดสัปดาห์เป็นวันละ 25% และวันธรรมดาเป็นวันละ 10%
ด้วย $\alpha = .025$

1) $Ho : \pi_1 : \pi_2 : \ldots : \pi_6 : \pi_7 = 25 : 10 : 10 : 10 : 25$

2) $Ha : \pi_1 : \pi_2 : : \pi_6 : \pi_7 \neq 25 : 10 : 10 : 10 : 25$

3) a = .025

4) จะปฏิเสธ เมื่อ $X^2 > \chi^2_{(7-1),.025} = 14.449$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{7} (O_i - E_i)^2/E_i$

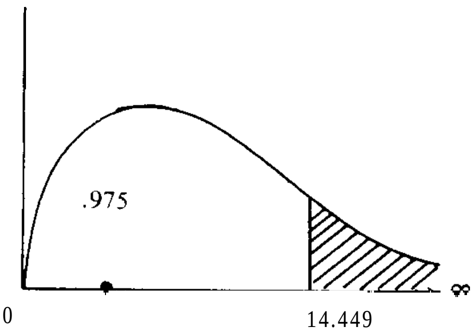
O_i คือจำนวนอุบัติเหตุที่ได้จากตัวอย่าง

E_i ต้องหาโดยสมมติว่า Ho เป็นจริง นั่นคือ $E_i = n\pi_i$

$E_1 = 100(.25) = 25$ (วันอาทิตย์)

$E_2 = 100(.10) = 10$

} วันธรรมดา



$E_6 = 100(.10) = 10$

$E_7 = 100(.25) = 25$ (วันเสาร์)

รวม 100

ดังนั้น $X^2 = \frac{(28-25)^2}{25} + \frac{(12 \quad 10)^2}{10} + \ldots + \frac{(11-10)^2}{10} + \frac{(24-25)^2}{25}$

$= 2.2$

6) $X^2 = 2.2$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า อุบัติเหตุในวันสุดสัปดาห์เกิดขึ้นวันละ 25% และวันธรรมดาวันละ 10%

10.19 โยนเหรียญสมดุลย์ 4 อัน พร้อมๆ กัน 160 ครั้ง และนับจำนวนเหรียญที่หงายด้านหัว ได้ผลดังนี้

จำนวนหัว	0	1	2	3	4	รวม
ความถี่	16	35	55	48	6	160

จงใช้ $\alpha = .05$ ทดสอบว่า เป็นเหรียญสมดุลย์ทั้ง 4 อัน

1) Ho : เหรียญทั้ง 4 อัน เป็นเหรียญสมดุลย์

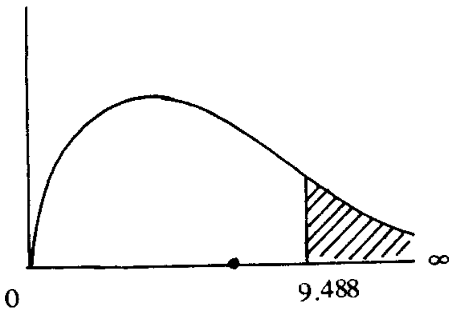
(นั่นคือ ข้อมูลต้องมีการแจกแจงแบบทวินามด้วย $\pi = .5$)

2) Ha : เหรียญ 4 อัน ไม่สมดุลย์ทั้งหมด

(นั่นคือ ข้อมูลมีการแจกแจงแบบทวินามด้วย $\pi \neq .5$)

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(5-1),.05} = 9.488$



5) ตัวสถิติที่ใช้ทดสอบคือ

$$X^2 = \sum_{i=1}^{5} \frac{(O_i - E_i)^2}{E_i}$$

O_i คือ ข้อมูลที่เก็บจากการทดลอง

E_i ค่าคาดหมายแบบสมมติฐานว่างเปล่า นั่นคือ ถ้าข้อมูลมีการแจกแจงแบบทวินาม ด้วย $n = 4$ และ $\pi = .5$ (เพราะเป็นเหรียญสมดุลย์)

$$P(X = 0) = \binom{4}{0}\left(\frac{1}{2}\right)^0\left(\frac{1}{2}\right)^4 = \frac{1}{16}\ ;\ E_1 = \frac{1}{16}(160) = 10$$

$$P(X = 1) = \binom{4}{1}\left(\frac{1}{2}\right)^1\left(\frac{1}{2}\right)^3 = \frac{4}{16}\ ;\ E_2 = \frac{4}{16}(160) = 40$$

$$P(X = 2) = \binom{4}{2}\left(\frac{1}{2}\right)^2\left(\frac{1}{2}\right)^2 = \frac{6}{16}\ ;\ E_3 = \frac{6}{16}\ 160) = 60$$

$$P(X = 3) = \binom{4}{3}\left(\frac{1}{2}\right)^3\left(\frac{1}{2}\right)^1 = \frac{4}{16}\ ;\ E_4 = \frac{4}{16}(160) = 40$$

$$P(X = 4) = \binom{4}{4}\left(\frac{1}{2}\right)^4\left(\frac{1}{2}\right)^0 = \frac{1}{16}\ ;\ E_5 = \frac{1}{16}(160) = 10$$

จำนวนหัว	0	1	2	3	4	รวม
O_i	16	35	55	48	6	100
E_i	10	40	60	40	10	100

ดังนั้น $$X^2 = \frac{(16 - 10)^2}{10} + \frac{(35 - 40)^2}{40} + \frac{(35 - 60)^2}{60} + \frac{(48 - 40)^2}{40} + \frac{(6 - 10)^2}{10}$$

$$= 7.842$$

6) X^2 ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่าเหรียญทั้ง 4 อันเป็นเหรียญสมดุลย์

10.20 นักชีววิทยาได้ทำการผสมพันธ์ถั่วลันเตา ได้ลูกผสมที่มีลักษณะต่าง ๆ ดังนี้

ลำต้นสูงและมีสีสวยงาม	**186 ต้น**
ลำต้นสูงและไม่มีสี	**66 ต้น**
ลำต้นเตี้ยแคระและมีสีสวยงาม	**54 ต้น**
ลำต้นเตี้ยแคระและไม่มีสี	**14 ต้น**
	320

ถ้าตามหลักพันธุกรรมซึ่งท่านเมนเดลได้สร้างทฤษฎีไว้ จะได้ลักษณะดังกล่าวในอัตราส่วน 9 : 3 : 3 : 1 ตามลำดับ จงทดสอบโดยใช้ $\alpha = .01$ ว่าผลการทดลองเป็นไปตามทฤษฎีของเมนเดลหรือไม่?

1) Ho : ผลการทดลองเป็นไปตามทฤษฎีของเมนเดล นั่นคือ

$$\pi_1 : \pi_2 : \pi_3 : \pi_4 = 9 : 3 : 3 : 1$$

2) Ha : ผลการทดลองไม่เป็นไปตามทฤษฎีของเมนเดล นั่นคือ

$$\pi_1 : \pi_2 : \pi_3 : \pi_4 \neq 9 : 3 : 3 : 1$$

3) $\alpha = .01$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(4-1),.01} = 11.345$

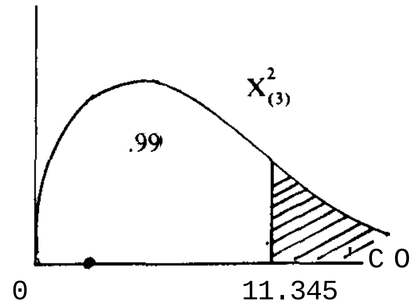
5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{4} \frac{(O_i - E_i)^2}{E_i}$

$$
\left.\begin{aligned}
E_1 &= n\pi_1 = 320\left(\frac{9}{16}\right) = 180 \\
E_2 &= n\pi_2 = 320\left(\frac{3}{16}\right) = 60 \\
E_3 &= n\pi_3 = 320\left(\frac{3}{16}\right) = 60 \\
E_4 &= n\pi_4 = 320\left(\frac{1}{16}\right) = 20
\end{aligned}\right\} E_i
\qquad
\left.\begin{aligned}
186 \\ 66 \\ 54 \\ 14
\end{aligned}\right\} O_i
$$

$$\underline{\underline{180}}$$

$$X^2 = \frac{(186-180)^2}{180} + \frac{(66-60)^2}{60} + \frac{(54-60)^2}{60} + \frac{(14-20)^2}{20}$$

$= 3.2$



6) $X^2 = 3.2$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho

และสรุปว่า ข้อมูลมีการแจกแจงแบบทฤษฎีของเมนเดล

10.21 **ในการนำสินค้าตัวใหม่ออกสู่ตลาด ผู้ผลิตได้แบ่งตลาดผู้ซื้อเป็น 10 ส่วน โดยคาดว่าแต่ละส่วนจะมีจำนวนประชากร และผู้มีอำนาจซื้อไม่ต่างกัน และได้ทำการโฆษณาสินค้าโดยสั่งโฆษณาหลายชนิด โดยขอให้ผู้สนใจผลิตภัณฑ์ส่งแบบสอบถามไปยังบริษัทว่าจะติดต่อกับตัวแทนจำหน่ายในท้องที่ของตนได้ที่ใด ในจำนวนแบบสอบถามที่ส่งกลับคืนมายังผู้ผลิต 200 ฉบับ โดยจำแนกตามท้องที่ มีดังนี้**

ท้องที่	1	2	3	4	5	6	7	8	9	10	รวม
จำนวน	22	23	18	16	21	17	19	23	20	21	200

จงทดสอบว่า ทั้ง 10 ท้องที่มีผลตอบสนองต่อผลิตภัณฑ์เท่ากันหรือไม่? $\alpha = .025$

1) Ho : $\pi_i = \frac{1}{10}$ $i = 1, 2, \ldots, 10$ (ทุกท้องที่มีผลตอบสนองเท่ากัน)

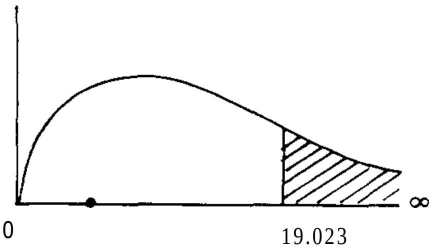
2) Ho : $\pi_i \neq \frac{1}{10}$ (ทุกท้องที่มีผลตอบสนองไม่เท่ากัน)

3) $\alpha = .025$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(10-1),.025} = 19.023$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{10} \frac{(O_i - E_i)^2}{E_i}$

$E_i = n\pi_i = 200\left(\frac{1}{10}\right) = 20$



ท้องที่	1	2	3	4	5	6	7	8	9	10	รวม
O_i	22	23	18	16	21	17	19	23	20	21	200
E_i	20	20	20	20	20	20	20	20	20	20	200

$$X^2 = \frac{(22-20)^2}{20} + \frac{(23-20)^2}{20} + \dots + \frac{(21-20)^2}{20}$$

$$= 54/20 = 2.7$$

6) ค่าสถิติ $X^2 = 2.7$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า แต่ละท้องที่มีผลตอบสนองต่อผลิตภัณฑ์ไม่ต่างกัน

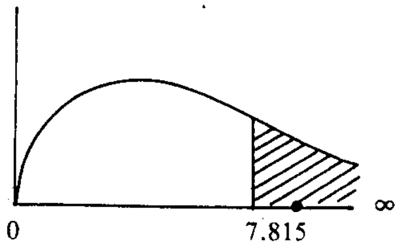
10.22 จำนวนอุบัติเหตุต่อสัปดาห์บนทางหลวงสายหนึ่งซึ่งเก็บสถิติรวม 80 สัปดาห์ มีดังนี้

จำนวนอุบัติเหตุร้ายแรง	0	1	2	3	4 หรือมากกว่า	รวม
ความถี่	3	22	33	16	6	80

จงทดสอบว่าข้อมูลมีการแจกแจงแบบปัวซอง ด้วย $\alpha = .05$
(ต้องประมาณค่า μ จากตัวอย่าง ดังนั้น df จึงหายไปอีก 1 df)

1) Ho : ข้อมูลมีการแจกแจงแบบปัวซอง

2) Ha : ข้อมูลไม่มีการแจกแจงแบบปัวซอง



3) $a = .05$

4) จะปฏิเสธ เมื่อ $X^2 > \chi^2_{(5-1-1) = 3\ df, .05} = 7.815$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{5} (O_i - E_i)^2/E_i$

E_i ต้องหาจาก Ho : ข้อมูลมีการแจกแจงแบบปัวซอง

การแจกแจงแบบปัวซองมีพารามิเตอร์เพียงตัวเดียว คือ μ

$$\hat{\mu} = \overline{X} = \frac{0(3) + 1(22) + 2(33) + 3(16) + 4(6)}{80}$$

$$= 160/80 = 2.0$$

จากตารางการแจกแจงแบบปัวซอง เมื่อ $\mu = 2$ จะได้ความน่าจะเป็น ดังนี้

ความน่าจะเป็น $= \pi_i$	$E_i = n\pi_i$	O_i	$(O_i - E_i)$	$(O_i - E_i)^2/E_i$
$P(X = 0) = .1353$	$E_1 = 80(.1353) = 10.824$	3	− 7.824	5.655
$P(X = 1) = .2707$	$E_2 = 80(.2707) = 21.656$	22	0.344	.005
$P(X = 2) = .2707$	$E_3 = 80(.2707) = 21.656$	33	11.344	5.942
$P(X = 3) = .1804$	$E_4 = 80(.1804) = 14.432$	16	1.568	0.170
$P(X \geq 4) = .1429$	$E_5 = 80(.1429) = 11.432$	6	− 5.432	2.581
1.000	80.00	80	0	14.353

6) $X^2 = 14.353$ อยู่ในเขตวิกฤต จึงปฏิเสธ Ho และยอมรับ Ha สรุปว่า ข้อมูลไม่มีการแจกแจงแบบปัวซอง

หมายเหตุ df = 5 − 1 − 1 = 3 = 5 − parameter ที่ต้องประมาณค่าคือ μ − total

10.23 โยนเหรียญ 6 เหรียญพร้อมกัน 1280 ครั้ง ได้ผลดังนี้

จำนวนหัว	0	1	2	3	4	5	6	รวม
ความถี่	26	140	274	420	290	112	18	1280

จงทดสอบว่าเหรียญ 6 อันนี้ เป็นเหรียญสมดุลย์หรือไม่, $\alpha = .01$

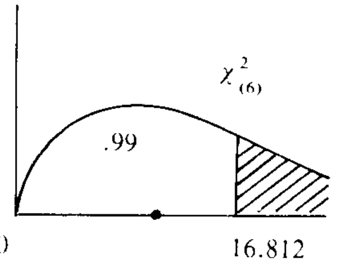
1) Ho : เหรียญทั้ง 6 เป็นเหรียญสมดุลย์ นั่นคือ จำนวนหัวต้องมีการแจกแจงแบบทวินาม ที่มี $\pi = \frac{1}{2}$

2) Ha : เหรียญไม่สมดุลย์ทั้งหมด นั่นคือ จำนวนหัวมีการแจกแจงแบบทวินามที่มี $\pi \neq \frac{1}{2}$

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(7-1),.01} = 16.812$

5) ตัวสถิติที่ใช้ทดสอบ คือ $\sum_{i=1}^{7} \frac{(O_i - E_i)^2}{E_i}$



$E_i = n\pi_i$; π_i ได้จากตารางการแจกแจงแบบทวินาม เมื่อ $n = 6, \pi = \frac{1}{2} = 0.5$ ดังนี้

จำนวนหัว = X	0	1	2	3	4	5	6	รวม
ความน่าจะเป็น = π	$\frac{1}{64}$	$\frac{6}{64}$	$\frac{15}{64}$	$\frac{20}{64}$	$\frac{15}{64}$	$\frac{6}{64}$	$\frac{1}{64}$	1.0
$E_i = n\pi_i$	20	120	300	400	300	120	20	12x0
O_i = ความถี่	26	140	274	420	290	112	18	1280
$(O_i - E_i)$	-6	-20	26	20	10	8	2	0
$(O_i - E_i)^2/E_i$	1.80	3.33	2.25	1.00	.33	.53	.20	9.44

6) $X^2 = 9.44$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า เหรียญ 6 อันเป็นเหรียญที่สมดุลย์

10.24 ถ้าใช้ $\alpha = .10$ จะสรุปว่าข้อมูลต่อไปนี้มีการแจกแจงแบบปัวซองที่มี $\mu = 2$ ได้ไหม?

จำนวนลูกค้าต่อชั่วโมง	0	1	2	3	4	5	รวม
จำนวนชั่วโมง (f)	10	19	31	26	11	3	100

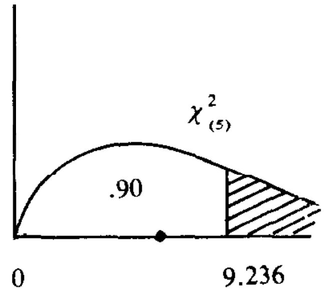
1) Ho : ข้อมูลมีการแจกแจงแบบปัวซอง ซึ่งมี $\mu = 2$

2) Ha : ข้อมูลมีการแจกแจงแบบปัวซองที่มี $\mu \neq 2$

3) $\alpha = .10$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(6-1) = 5, .10} = 9.236$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{6} \frac{(O_i - E_i)^2}{E_i}$



$E_i = n\pi_i$, π_i ได้จากการแจกแจงแบบปัวซองที่มี $\mu = 2$

จำนวนลูกค้า/ชั่วโมง = x	0	1	2	3	4	5 หรือมากกว่า	รวม
ความน่าจะเป็น = π_i	.1353	.2707	.2707	.1804	.0902	.0527	1.0
$E_i = n\pi_i$	13.53	27.07	27.08	18.04	9.02	5.27	100
Oi	10.00	19.00	31.00	26.00	11.00	3.00	100
$(O_i - E_i)^2/E_i$	0 92	2.41	0.57	3.51	0.43	0.98	8.82

6) $X^2 = 8.82$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่าข้อมูลมีการแจกแจงแบบปัวซอง ที่มี $\mu = 2$

10.25 ให้ทดสอบว่า ข้อมูลในตารางข้างล่างมีการแจกแจงแบบปกติหรือไม่ โดยใช้ $\alpha = .05$ และกำหนดความถี่คาดหมายภายใต้การแจกแจงแบบปกติให้ ดังนี้

คะแนน	51 - 60	61 - 70	71 - 80	81 - 90	91 - 100	รวม
O_i	3	10	44	50	13	120
E_i	2	17	50	41	10	120

1) Ho : ข้อมูลมีการแจกแจงแบบปกติ

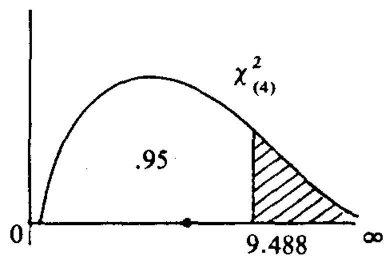
2) Ha : ข้อมูลไม่มีการแจกแจงแบบปกติ

3) $\alpha = 0.05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(5-1), .05} = 9.488$

5) ตัวสถิติที่ใช้ทดสอบ คือ

$$X^2 = \Sigma \frac{(O_i - E_i)^2}{E_i}$$



$$= \frac{(3-2)^2}{2} + \frac{(10-17)^2}{17} + \frac{(44-50)^2}{50} + \frac{(50-41)^2}{41} + \frac{(13-10)^2}{10}$$

$$= 6.98$$

6) $X^2 = 6.98$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่าข้อมูลมีการแจกแจงแบบปกติ

หมายเหตุ กรณีนี้โจทย์ไม่บอกว่า หาค่า E_i มาได้อย่างไร หากต้องหาค่า E_i เอง และต้องหาค่าประมาณของ μ และ σ df จะต้องหายไปอีก 2 df ดังนั้น ต้องเทียบค่า X^2 กับ $\chi^2_{2,.05} = 5.991$ แต่ $X^2 = 6.98 > 5.991$ จึงปฏิเสธ Ho และสรุปว่า ข้อมูลไม่มีการแจกแจงแบบปกติ

10.26 บริษัทจัดทำโฆษณาสินค้าได้ประเมินผลการโฆษณาโดยการใช้โทรศัพท์สอบถามว่าได้รับชมโฆษณาชิ้นนั้นหรือไม่ โดยได้ออกอากาศใน ใช้ช่วงเวลาต่าง ๆ กัน 3 ครั้ง เมื่อสัปดาห์ก่อน บริษัทเชื่อว่า ผู้รับชมโทรทัศน์ จะมีโอกาสได้รับชมโฆษณาแต่ละชิ้นด้วยโอกาส .4 และเชื่อว่า การได้รับชมแต่ละครั้งเป็นอิสระกัน เพราะใช้เวลาและสถานีต่างกัน ผลการสำรวจผู้ชม 200 คน ได้ผลดังนี้

จำนวนโฆษณาที่ได้รับชม	0	1	2	3	รวม
จำนวนผู้รับชม	46	73	58	23	200

จงใช้ $\alpha = .10$ **ทดสอบว่า จำนวนผู้รับชม มีการแจกแจงแบบทวินามด้วย** $\pi = .4$

1) Ho : ข้อมูลมีการแจกแจงแบบทวินามด้วย $\pi = .4$

2) Ha : ข้อมูลมีการแจกแจงแบบทวินามด้วย $\pi \neq .4$

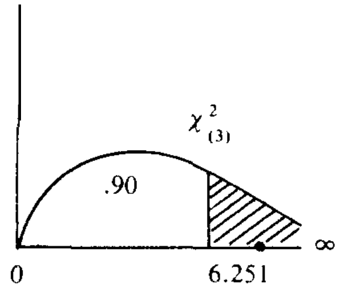
3) $\alpha = .10$

4) **จะปฏิเสธเมื่อ** $X^2 > \chi^2_{(4-1),.10} = 6.251$

5) **ตัวสถิติที่ใช้ทดสอบคือ** $X^2 = \frac{\Sigma(O_i - E_i)^2}{E_i}$

$E_i = n\pi_i$, π_i ได้จากการแจกแจงแบบทวินามที่มี $n = 3$, $\pi = .4$

จำนวนโฆษณา = X	0	1	2	3	รวม
จำนวนผู้รับชม = O_i	46	73	58	23	200
ความน่าจะเป็น = π_i	.216	.432	.288	.064	1.000
$E_i = n\pi_i$	43.2	86.4	57.6	12.8	200
$(O_i - E_i)^2/E_i$	0.18	2.08	.003	8.13	10.393



6) $X^2 = 10.393$ ตกอยู่ในเขตวิกฤต จึงปฏิเสธ Ho และยอมรับ Ha, สรุปว่าจำนวนโฆษณาที่ได้รับชมมีการแจกแจงแบบทวินาม ซึ่งมี $\pi \neq .4$

10.27 จำนวนผู้ประสบอุบัติเหตุกระดูกหักต่อวันในโรงพยาบาลหนึ่ง โดยเก็บจากสถิติที่สุ่มมา 300 วัน มีดังนี้

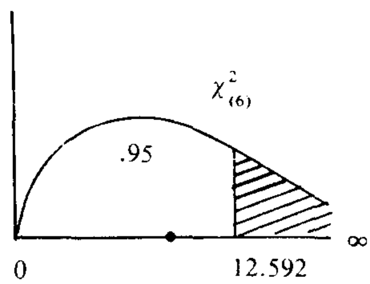
จำนวนคนไข้ต่อวัน	0	1	2	3	4	5	6 ขึ้นไป	รวม
จำนวนวัน	25	45	63	71	48	26	22	300

ถ้าใช้ $\alpha = 0.05$ จะสรุปว่า จำนวนคนไข้กระดูกหักของโรงพยาบาลนี้ มีการแจกแจงแบบปัวซอง ที่มีค่าเฉลี่ย $\mu = 3$ ได้หรือไม่?

1) Ho : จำนวนคนไข้ต่อวัน มีการแจกแจงแบบปัวซองที่มี $\mu = 3$

2) Ha : จำนวนคนไข้ต่อวัน มีการแจกแจงแบบปัวซองที่มี $\mu \neq 3$

3) $\alpha = .05$



4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(7-1)\ .05} = 12.592$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{7} (O_i - E_i)^2/E_i$

$E_i = n\pi_i$, π_i ได้จากตารางการแจกแจงแบบปัวซองที่มี $\mu = 3$ ดังนี้

จำนวนคนไข้ต่อวัน = X	0	1	2	3	4	5	6 ขึ้นไป	รวม
จำนวนวัน = O_i	25	45	63	71	48	26	22	300
ความน่าจะเป็น = π_i	.0498	1494	.2240	.2240	.1680	.1008	.0840	I .0
$E_i = n\pi_i$	14.94	44.82	67.2	67.2	50.4	30.24	25.2	300
$(O_i - E_i)^2/E_i$	6.77	0.0007	0.2625	0.21	0.11	0.59	0.41	8.35

6) $X^2 = 8.35$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า ข้อมูลมีการแจกแจงแบบปัวซอง ด้วย $\mu = 3$ คือ โดยเฉลี่ยมีคนไข้กระดูกหัก 3 คนต่อวัน

10.28 ถ้าแบ่งการทำงานของพนักงานดับเพลิงของสถานีดับเพลิงแห่งหนึ่งเป็น 3 ผลัด ๆ ละ 8 ชั่วโมง เชื่อว่ามีโอกาส 30% ที่สัญญาณไฟไหม้จะดัง และมีข้อมูลที่เก็บมาใน 60 วัน ดังนี้

จำนวนผลัดที่มีสัญญาณเพลิงไหม้	0	1	2	3	รวม
จำนวนวัน	16	27	11	6	60

ให้ใช้ $\alpha = .05$ ทดสอบว่า ข้อมูลมีการแจกแจงแบบทวินาม ที่มี $\pi = .3$

1) Ho : ข้อมูลมีการแจกแจงแบบทวินามด้วย $\pi = .3$

2) Ha : ข้อมูลมีการแจกแจงแบบทวินามด้วย $\pi \neq .3$

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(3-1)=2,.05} = 5.991$

5) ตัวสถิติที่ใช้ทดสอบคือ $X^2 = \sum_{i=1}^{3} (O_i - E_i)^2/E_i$

$E_i = n\pi_i$, π_i ได้จากตารางการแจกแจงแบบทวินาม ที่มี $n = 3, \pi = .3$

จำนวนผลัดที่มีการแจ้งสัญญาณเพลิงไหม้ = X	0	1	2	3	รวม
จำนวนวัน = O_i	16	27	11	6	60
ความน่าจะเป็น	.3430	.4410	.1890	.0270	1.0
$E_i = n\pi_i$	20.58	26.46	11.34	1.62	60
	20.58	26.46	12.96		60
O_i	16	27	17		60
$(O_i - E_i)^2/E_i$	1.02	0.01	1.26		2.29

6) $X^2 = 2.29$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า ข้อมูลมีการแจกแจงแบบทวินามด้วย $\pi = .3$

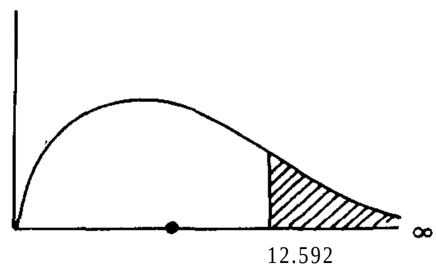
หมายเหตุ ค่า E_i เมื่อ $(X = 3) = 1.62$ ครั้ง ซึ่งน้อยกว่า 5 จึงต้องยุบรวมกับชั้นถัดไป ทำให้ df ลดลงไป 1 df และต้องรวมค่า O_i ด้วย

10.29 ร้านสรรพสินค้าเก็บสถิติลูกค้าที่ชำระเงินค่าสินค้า ณ เคาท์เตอร์ต่างๆ ในช่วง 5 นาที โดยการสุ่มเวลา 5 นาที ต่างๆ มา 800 ช่วงเวลา พบว่ามีจำนวนลูกค้ารอชำระเงินดังนี้

จำนวนลูกค้า ณ. เคาท์เตอร์	0	1	2	3	4	5	6 ขึ้นไป	รวม
จำนวนครั้ง	36	117	194	167	138	94	54	800

ก) ให้ทดสอบว่า ข้อมูลมีการแจกแจงแบบปัวซองที่มี $\mu = 3$ หรือไม่? $\alpha = .05$

ข) ร้านควรจัดตั้งเคาท์เตอร์ชำระเงินกี่ตัว



ก) 1) Ho : ข้อมูลมีการแจกแจงแบบปัวซองที่มี $\mu = 3$

2) Ha : ข้อมูลมีการแจกแจงแบบปัวซองที่มี $\mu \neq 3$

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $X^2 > \chi^2_{(7-1)\ .05} = 12.592$

5) $x^2 = \sum_{i=I}^{7} (0_i - E_i)^2/E_i$, $E_i = n\pi_i$; π_i จากตารางปัวซอง $\mu = 3$

จำนวนลูกค้า = x	0	1	2	3	4	5	6 ขึ้นไป	รวม
ความน่าจะเป็น = π_i	.0498	.1494	.2240	.2240	.1680	.1008	.0840	800
$E_i = n\pi_i$	39.84	119.52	179.2	179.2	134.4	80.64	67.2	800
จำนวนครั้ง = 0 ,	36.0	117.0	194.0	167.0	138.0	94.0	54.0	800
$(0_i - E_i)^2/E_i$	.37	.053	1.22	.83	.096	2.21	2.59	7.369

6) $X^2 = 7.369$ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho ว่า ข้อมูลมีการแจกแจงแบบปัวซองที่มี $\mu = 3$ นั่นคือ จำนวนลูกค้าเข้าชำระเงิน ณ เคาท์เตอร์ ในช่วง 5 นาทีจะมีโดยเฉลี่ย 3 ราย ดังนั้น จึงควรติดตั้งเคาท์เตอร์ชำระเงิน 3 แห่ง

10.30 จงแสดงข้อสมมุติของการวิเคราะห์ความแปรปรวน

ข้อสมมุติของการวิเคราะห์ความแปรปรวน คือ ข้อมูลต้องมาจากตัวอย่าง ซึ่งสุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ k ประชากร ตัวอย่างสุ่ม k กลุ่มนั้นต้องเป็นอิสระกัน และความแปรปรวนของประชากรปกตินั้นต้องเป็นเอกภาพ นั่นคือ $\sigma_1^2 = \sigma_2^2 \ldots = \sigma_k^2 = \sigma^2$ (แต่อาจไม่ทราบค่าแน่นอนของ σ^2 จึงต้องประมาณด้วย S_p^2 = MSE)

10.31 การวิเคราะห์ความแปรปรวนเป็นการทดสอบแบบด้านเดียวหรือ 2 ด้าน?

เป็นการทดสอบแบบด้านเดียว เพราะตัวสถิติ $F = \frac{MSA}{MSE}$ โดยทฤษฎีแล้ว จะมีค่าเป็น 1 ขึ้นไป-ทั้งนี้ เพราะตามทฤษฎี MSA วัดความแปรปรวนระหว่างกลุ่ม จะต้องสูงกว่า MSE ซึ่งวัดความแปรปรวนภายในกลุ่ม เราจึงจะปฏิเสธ เฉพาะเมื่อ MSA มีค่าโตเกินไป จนทำให้ F ใหญ่กว่าค่าเปิดตาราง ดังนั้น เขตวิกฤตจึงอยู่แต่ปลายหางด้านมาก คือ ด้านขวามือ

10.32 สมมติฐานของการวิเคราะห์ความแปรปรวนมีว่าอย่างไร?

สมมติฐานของการวิเคราะห์ความแปรปรวน คือ

Ho : $\mu_1 = \mu_2 = \ldots = \mu_t = \mu$

หมายความว่า ประชากรแบบปกติ k ประชากรนั้นมีค่าเฉลี่ยไม่ต่างกัน

Ha : μ_j ไม่เท่ากันทั้งหมด

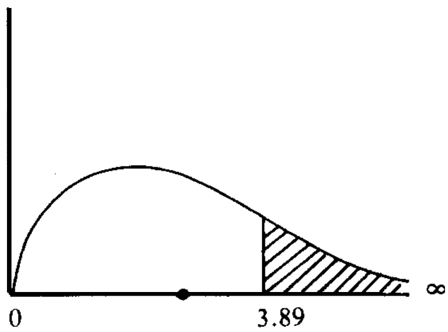
$j = 1, 2, \ldots, k$

หรือ มีอย่างน้อย กลุ่มที่ต่างกัน

10.33 จัดลูกไก่แบบสุ่มให้กินอาหารสูตรต่าง ๆ 3 สูตร แต่ละสูตรมีลูกไก่ 5 ตัว และเมื่อครบกำหนดการทดลอง ได้ชั่งน้ำหนักที่เพิ่มขึ้น ดังนี้

สูตร 1	สูตร 2	สูตร 3
4	3	6
4	4	7
7	5	7
7	6	7
8	7	8

สมมุติว่า น้ำหนักที่เพิ่มขึ้นของลูกไก่มีการแจกแจงแบบปกติ จงใช้ $\alpha = .05$ ทดสอบว่าสูตรทั้ง 3 ให้น้ำหนักเพิ่มต่างกันหรือไม่?



1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu$

(น้ำหนักเพิ่มเฉลี่ยของสูตรทั้ง 3 ไม่ต่างกัน)

2) Ha : มีอย่างน้อย 1 สูตรที่ต่างกัน

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $F > f_{(k-1),(N-k),\alpha}$

$$F > f_{2,12,.05} = 3.89$$

5) $T_1 = 30, n_1 = 5, \overline{X}_1 = 6$ $\quad k = 3$

$T_2 = 25, n_2 = 5, \overline{X}_2 = 5$ $\quad N = \Sigma n_j = 15$

$T_3 = 35,\ n_3 = 5,\ \bar{X}_3 = 7 \qquad G = \Sigma T_j = 90$

$\Sigma\Sigma X_{ij}^2 = (4^2 + 4^2 + + 7^2 + 8^2) = 576 \qquad CF = G^2/N = 540$

$SS\ (\text{total}) = \Sigma\Sigma X_{ij}^2 - CF = (576 - 540) = 36$ มี $df = N - 1$

$= 15 - 1 = 14$

$\Sigma \frac{T_j^2}{n_j} = \frac{30^2}{5} + \frac{25^2}{5} + \frac{35^2}{5} \quad \frac{2750}{5} = 550$

SS (ระหว่างกลุ่ม) $= SSA = \Sigma \frac{T_j^2}{n_j} - CF$

$= 550 - 540 = 10$ มี $df = (k \quad 1) = 2$

SS (ภายในกลุ่ม) = S S E $= SS\ (\text{total}) - SSA$

$= 36 - 10$

$= 26$ มี $df = (N \quad k) = (15 - 3) = 12$

SOV	df	SS	MS = SS/df	F - ratio
ระหว่างกลุ่ม (A)	2	10	5	2.3
ภายในกลุ่ม (E)	12	26	2.17	
Total	14	36		

ตัวสถิติที่ใช้ทดสอบ คือ $F = \frac{MSA}{MSE} = \frac{5}{2.17} = 2.3$

6) $F = 2.3 < 3.89$ จึงไม่ปฏิเสธ Ho สรุปว่า น้ำหนักเพิ่มเฉลี่ยของลูกไก่ เมื่อกินอาหารสูตรต่าง ๆ ไม่ต่างกันอย่างมีนัยสำคัญ

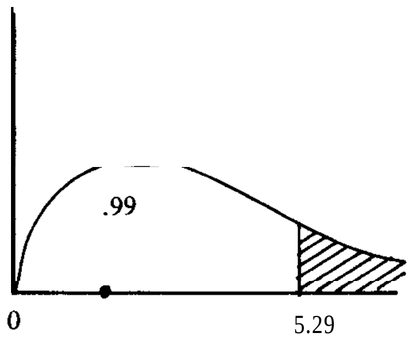
ข้อสังเกต สูตรอาหารดังกล่าวอาจมีอิทธิพลทำให้น้ำหนักเฉลี่ยต่างกันได้ ถ้าให้ขนาดตัวอย่างโตกว่านี้ สำหรับข้อมูลที่ได้ เมื่อใช้ $n = 5$ ยังไม่สามารถตรวจพบความแตกต่าง (ซึ่งอาจมี)

10.34 โรงงานแห่งหนึ่งมีเครื่องจักรชนิดเดียวกัน 4 เครื่อง แต่ละเครื่องมีผู้ควบคุม 1 คน เมื่อสุ่มตัวอย่างชั่วโมงการทำงานมา 5 ชั่วโมง และเก็บข้อมูลด้วยจำนวนสินค้าชำรุดต่อ

ชั่วโมง พบว่ามีจำนวนสินค้าชำรุดแต่ละชั่วโมง มีดังนี้

เครื่องจักร 1	7	7	8	8	5	$T_1 = 35, \bar{X}_1 = 7$
เครื่องจักร 2	2	3	3	3	4	$T_2 = 15, \bar{X}_2 = 3$
เครื่องจักร 3	10	9	9	9	8	$T_3 = 45, \bar{X}_3 = 9$
เครื่องจักร 4	3	3	6	6	7	$T_4 = 25, \bar{X}_4 = 5$

สมมุติว่าจำนวนชำรุดรายชั่วโมงมีการแจกแจงแบบปกติ จงใช้ $\alpha = .01$ **ทดสอบว่าจำนวนชำรุดจากเครื่องทั้ง 4 ไม่ต่างกัน**



1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu$

2) Ha : มีอย่างน้อย 1 เครื่องที่ต่างกัน

3) a = .01

4) จะปฏิเสธ H o เมื่อ $F > f_{.01,(4-1),(20-4)} = f_{.01,3,16} = 5.29$

5) $n_1 = n_2 = n_3 = n_4 = n = 5,\ N = \Sigma n_j = 20$

$G = \Sigma T_j = 120,\ CF = G^2/N = 720$

$\Sigma\Sigma X_{ij}^2 = (7^2 + 7^2 + \ldots + 6^2 + 7^2) = 844$

$SS(total) = SST = \Sigma\Sigma X_{ij}^2 - CF$

$= 844 - 720 = 124$

$df(total) = N - 1 = (20 - 1) = 19$ df

$\Sigma T_j^2/n_j = \frac{35^2}{5} + \frac{15^2}{5} + \frac{45^2}{5} + \frac{25^2}{5} = \frac{4100}{5} = 820$

SS (ระหว่างกลุ่ม) = $SSA = \Sigma T_j^2/n_j - CF$

$= 820 - 720 = 100$

$df_A = (k - 1) = (4 - 1) = 3$ df

S S (ภายในกลุ่ม) = $SSE = (SST - SSA)$

$= (124 - 100) = 24$

$$df_E = (N - k) = (20 - 4) = 16$$

SOV	df	SS	MS = SS/df	F - ratio
ระหว่างกลุ่ม	3	100	33.33	1.26
ภายในกลุ่ม	16	24	26.5	
รวม	19	124		

6) $F = 1.26 < 5.29$ จึงยอมรับ Ho สรุปว่า จำนวนผลผลิตชำรุด โดยถัวเฉลี่ยต่อชั่วโมงของเครื่องจักรทั้ง 4 เครื่องไม่ต่างกัน อย่างมีนัยสำคัญ

10.35 เมื่อใช้ปุ๋ย 3 ชนิดในแปลงปลูกสตรอเบอรี่ 3 แปลง โดยแปลงที่ 1 ปลูก 5 ต้น แปลงที่ 2 ปลูก 4 ต้น แปลงที่ 3 ปลูก 6 ต้น แต่ละต้นใส่ปุ๋ยในจำนวนเท่ากัน ได้ผลผลิตต่อต้น ดังนี้

ปุ๋ย I	3	4	4	5	7		$T_1 = 23, n_1 = 5, \bar{X}_1 = 4.6$
ปุ๋ย II	2	3	4	5			$T_2 = 14, n_2 = 4, \bar{X}_2 = 3.5$
ปุ๋ย III	4	5	6	6	7	8	$T_3 = 36, n_3 = 6, \bar{X}_3 = 6.0$

จงทดสอบนัยสำคัญ โดยใช้ $\alpha = .01$ และ $\alpha = .05$

1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu$

(ผลผลิตต่อต้นเมื่อใช้ปุ๋ยชนิดต่าง ๆ ไม่ต่างกัน)

2) Ha : มีอย่างน้อย 1 กลุ่มที่ต่างกัน

3) $\alpha = .05$ และ $\alpha = .01$

4) ถ้าใช้ $\alpha = .01$ จะปฏิเสธ Ho เมื่อ $F > f_{2,12;.01} = 6.93$

$$\boxed{\begin{aligned} \gamma_1 &= (k - 1) = (3 - 1) = 2 \\ \gamma_2 &= (N - k) = (15 - 3) = 12 \end{aligned}}$$

ถ้าใช้ $\alpha = .05$ จะปฏิเสธ Ho เมื่อ

$F > f_{2,12;.05} = 3.89$

5) $N = \Sigma n_j = 5 + 4 + 6 = 15$

$G = \Sigma T_j = 23 + 14 + 36 = 73$

$CF = G^2/N = (73)^2/15 = \boxed{355.27}$

$$\Sigma T_j^2/n_j = \frac{23^2}{5} + \frac{14^2}{4} + \frac{36^2}{6} = \boxed{370.8}$$

$SSA = 370.8 - 355.27 = \boxed{15.53}$, $df = (3 - 1) = 2$

$\Sigma\Sigma X_{ij}^2 = (3^2 + 4^2 + \ldots + 7^2 + 8^2) = \boxed{395}$

$SS(\text{total}) = SST = \Sigma\Sigma X_{ij}^2 - CF$

$= 395 - 355.27$

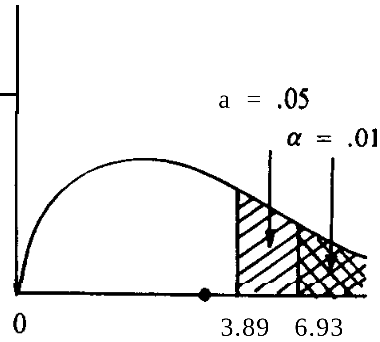
$= \boxed{39.73}$, $df = (N - 1) = (15 - 1) = 14$

SS (ภายในกลุ่ม) $= SSE = (SST - SSA)$

$= 39.73 - 15.53 = \boxed{24.2}$

$df = (N - k) = (15 - 3) = 12$

sov	df	SS	df	F • ratio
ระหว่างกลุ่ม = A	2	15.53	7.765	3.84
ภายในกลุ่ม = E	12	24.20	2.02	
	14	39.73		



6) F = 3.84 ไม่อยู่ในเขตวิกฤต ทั้ง a = .05 และ α = .01 จึงยังไม่ปฏิเสธ Ho ทั้ง 2 กรณี และสรุปว่า ปุ๋ยทั้ง 3 ชนิด มีอิทธิพล โดยให้ผลผลิตเฉลี่ยต่อต้นไม่ต่างกัน

10.36 สุ่มหลอดไฟชนิดต่าง ๆ มา 4 ชนิด ๆ ละ 3 หลอด ได้ผลรวมอายุการใช้งานของแต่ละชนิด เป็น 1,000 ชั่วโมง ดังนี้

$T_1 = 50, T_2 = 40, T_3 = 40, T_4 = 50$

และ $\Sigma\Sigma X_{ij}^2 = 2778$ ถ้าสมมุติว่าอายุการใช้งานของหลอดไฟมีการแจกแจงแบบปกติ ให้ทดสอบว่าอายุการใช้งานโดยเฉลี่ยของ 4 ชนิดนั้น แตกต่างกัน เมื่อใช้ $\alpha = .05$ หรือไม่?

1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu$

(อายุเฉลี่ยของหลอดไฟทั้ง 4 ชนิดไม่ต่างกัน)

2) Ha : มีอย่างน้อย 1 ชนิดที่ต่างกัน

3) a = .05

4) จะปฏิเสธ Ho เมื่อ F > $f_{3,8;.05} = 4.07$, $\gamma_1 = (k - 1) = (4 - 1) = 3$

$\gamma_2 = (N \quad k) = (12 - 4) = 8$

5) $n_1 = n_2 = n_3 = n_4 = 3$, $k = 4$

$N = \Sigma n_j = 4(3) = 12$

$G = \Sigma T_j = (50 + 40 + 40 + 50) = 180$

$CF = G^2/N = (180)^2/12 = \boxed{2700}$ มี I df

$\Sigma\Sigma X_{ij}^2 = 2778$

$SS(total) = (\Sigma\Sigma X_{ij}^2 \quad CF) = (2778 - 2700) = \boxed{78}$

และมี $df = (N - 1) = (12 - 1) = 11$

$$\Sigma T_j^2/n_j = \frac{50^2}{3} + \frac{40^2}{3} + \frac{40^2}{3} + \frac{50^2}{3}$$

$$= 8200/3 = 2733.33$$

SS (ระหว่างกลุ่ม) = SSA = $\Sigma T_j^2/n_j - CF$

$= 2733.33 - 2700$

$= \boxed{33.33}$ มี $df = (k - 1) = 3$

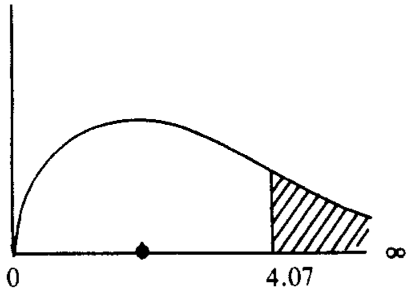
SS (ภายในกลุ่ม) = SSE = (SS (total) − SSA)

$= (78 - 33.33) = 44.67$

$df_E = N - k = 12 - 4 = 8$

sov	df	SS	MS	F ratio
ระหว่างกลุ่ม = A	3	33.33	11.11	1.99
ภายในกลุ่ม = E	8	44.67	5.58	
รวม	11	78		

6) F = 1.99 ไม่อยู่ในเขตวิกฤต
จึงยอมรับ Ho ว่าอายุเฉลี่ยของ
หลอดไฟ 4 ชนิดไม่ต่างกันอย่าง
มีนัยสำคัญ



10.37 ถ้าแบ่งพนักงานในบริษัทหนึ่งตามกลุ่มอายุ 3 กลุ่ม แต่ละกลุ่มสุ่มมา 4 คน พบว่า ผลผลิตเฉลี่ย และความแปรปรวนของแต่ละกลุ่ม มีดังนี้

$\bar{X}_1 = 7.5 \qquad \bar{X}_2 = 10 \qquad \bar{X}_3 = 8.75$

$S_1^2 = 3.0 \qquad S_2^2 = 1.0 \qquad S_3^2 = 1.25$

สมมุติว่า ผลผลิตมีการแจกแจงแบบปกติ จงทดสอบความแตกต่างของผลผลิตของกลุ่มอายุต่าง ๆ โดยใช้ $\alpha = .05$

1. Ho : $\mu_1 = \mu_2 = \mu_3 = \mu$
2. Ha : μ_j ไม่เท่ากันทั้งหมด ; j = 1, 2, 3
3. a = .05
4. $k = 3,\ n = 4,\ N = 12,\ \gamma_1 = (k - 1) = 2;\ \gamma_2 = (N - k) = 9$

   ดังนั้น จะปฏิเสธ Ho เมื่อ $F > f_{2,9;.05} = 4.26$

5) $n_1 = n_2 = n_3 = n = 4, \quad k = 3, N = \Sigma n_j = 12$

$T_j = n_j \bar{X}_j = n\bar{X}_j = 4\bar{X}_j$

$T_1 = 30,\ T_2 = 40,\ T_3 = 35,\ G = \Sigma T_j = 105$

MSE $= S_p^2$, เมื่อ n_j เท่ากันทุกกลุ่ม

$$S_p^2 = \Sigma S_j^2/k = (S_1^2 + S_2^2 + S_3^2)/3$$
$$= (3.0 + 1.0 + 1.25)/3$$
$$= 5.25/3 = 1.75$$

SSE $= (N - k)$ MSE เพราะ MSE $=$ SSE$/(N - k)$

$= 9(1.75) = \boxed{15.75}$

$CF = (105)^2/12 = 918.75$

$\Sigma T_j^2/n_j = (30^2 + 40^2 + 35^2)/4 = 3725/4 = \boxed{931.25}$

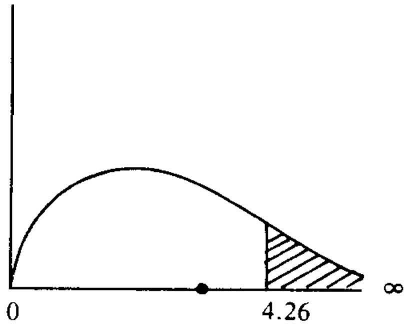
S S (ระหว่างกลุ่ม) = $\Sigma T_j^2/n_j - CF$

$= 931.25 - 918.75$

$= \boxed{12.5}$

SOV	df	SS	MS	F-ratio
ระหว่างกลุ่ม = A	2	12.50	6.25	3.57
ภายในกลุ่ม = E	9	15.75	1.75	
รวม	11	28.25		

6) F = 3.57 ไม่อยู่ในเขตวิกฤต จึงยังไม่ปฏิเสธ Ho และสรุปว่า ผลผลิตเฉลี่ยของพนักงานกลุ่มอายุต่าง ๆ ไม่ต่างกันอย่างมีนัยสำคัญ



10.38 ใช้วิธีส่งเสริมจำนวนขาย 4 วิธีๆ ละ 1 เดือน ได้ข้อมูล คือ จำนวนขายจากร้านตัวอย่าง 5 แห่ง ซึ่งใช้วิธีส่งเสริม 4 วิธี ในเดือนต่างๆ ดังนี้

แจกตัวอย่างฟรี	77	86	80	88	84	$T_1 = 415$
แถมของ 1 ชิ้น	95	92	88	91	89	$T_2 = 455$
ลดราคา	73	77	68	82	75	$T_3 = 375$
ให้ส่วนลดทางไปรษณีย์	80	84	79	70	82	$T_4 = 395$

$G = 1640$

$N = 20$

1) $Ho : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu$

(จำนวนขายจากวิธีส่งเสริม 4 วิธี ไม่ต่างกัน)

2) Ha : μ_j ไม่เท่ากันทั้งหมด ; j = 1, 2, 3, 4

ก) จงหาจำนวนขายโดยเฉลี่ยของการส่งเสริมแต่ละวิธี และหาค่าเฉลี่ยรวมยอด

$\bar{X}_1 = 415/5 = 83$

$\bar{X}_2 = 455/5 = 91$

$\bar{X}_3 = 375/5 = 75.0$

$\bar{X}_4 = 395/5 = 79.0$

ค่าเฉลี่ยรวมยอด $= \bar{\bar{X}} = G/N = 1640/20 = 82.00$

ข) จงคำนวณความแปรปรวนระหว่างค่าเฉลี่ยของตัวอย่าง

= MS (ระหว่างกลุ่ม) = SSA/(k - 1) = MSA

$SSA = \Sigma T_j^2/n_j \quad CF$

$$\Sigma T_j^2/n_j = \frac{(415^2 + 455^2 + 375^2 + 395^2)}{5} = \boxed{135,180}$$

$CF = G^2/N = (1640)^2/20 = \boxed{134,480}$

$SSA = 135,180 - 134,480 = 700$

$MSA = 700/3 = \boxed{233.33}$

ค) จงหาค่าประมาณของ σ^2 จากความแปรปรวนระหว่างกลุ่ม

จากข้อ (ข) มีวิธีหาความแปรปรวนระหว่างกลุ่มอีกวิธีหนึ่ง คือหาตามสูตรนิยาม ดังนี้

$$V(\bar{X}_j) = S_{\bar{X}}^2 = \sum_{j=1}^{4} (\bar{X}_j - \bar{\bar{X}})^2 / (k - 1)$$

$$= \sum_{j=1}^{4} (\bar{X}_j - \bar{\bar{X}})^2 / (k - 1)$$

$$= \{(83 - 82)^2 + (91 - 82)^2 + (75 \quad 82)^2 + (79 \quad 82)^2\} / (4 - 1)$$

$$= \{140\} / 3 = 46.667 = V(\bar{X})$$

แต่ $\hat{\sigma}^2 = V(X) = n\,V(\bar{X})$ (เพราะ $S_{\bar{X}}^2 = \frac{S^2}{n}$)

$= 5(46.667)$

$= 233.33$

ง) จงหาค่าประมาณของ σ^2 จากความแปรปรวนภายในกลุ่ม

$$\hat{\sigma}^2 = S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 + (n_3 - 1)S_3^2 + (n_4 - 1)S_4^2}{n_1 + n_2 + n_3 + n_4 - 4}$$

เมื่อ $n_1 = n_2 = n_3 = n_4 = n$

$$S_p^2 = \frac{(\cancel{n-1})\,(S_1^2 + S_2^2 + S_3^2 + S_4^2)}{4(\cancel{n-1})} = \sum_{i=1}^{k} S_i^2/k$$

$$S_1^2 = \Sigma(X - \bar{X}_1)^2/n_1 - 1$$

$$= \frac{1}{4}\{(77 - 83)^2 + (86 - 83)^2 + (80 - 83)^2 + (88 - 83)^2 + (84 - 83)^2\}$$

$$= 80/4 = 20.0$$

$$S_2^2 = \frac{1}{4}\{(95 - 91)^2 + (92 - 91)^2 + (88 - 91)^2 + (91 - 91)^2 + (89 - 91)^2\}$$

$$= 30/4 = 7.5$$

$$S_3^2 = \frac{1}{4}\{(73 - 75)^2 + (77 - 75)^2 + (68 - 75)^2 + (82 - 75)^2 + (75 - 75)^2\}$$

$$= 106/4 = 26.5$$

$$S_4^2 = \frac{1}{4}\{(80 - 79)^2 + (84 - 79)^2 + (79 - 79)^2 + (70 - 79)^2 + (82 - 79)^2\}$$

$$= 116/4 = 29$$

$$S_p^2 = \Sigma S_i^2/4$$

$$= \frac{1}{4}\left(\frac{80}{4} + \frac{30}{4} + \frac{106}{4} + \frac{1161}{4}\right)$$

$$= 332/16 = 20.75$$

ดังนั้น ค่าประมาณอีกค่าหนึ่งของ σ^2 คือ MSE = 20.75

ข้อสังเกต เนื่องจากในข้อ (ข) ได้คำนวณค่า $\Sigma T_j^2/n_j$ และ CF ซึ่งเป็นสูตรเครื่องคำนวณแล้ว จึงควรหาค่าอื่นเพิ่มเติมอีกเพียงค่าเดียว คือ $\Sigma\Sigma X_{ij}^2$ เพื่อใช้ตรวจสอบความถูกต้อง

$$\Sigma\Sigma X_{ij}^2 = (77^2 + 86^2 + \ldots + 82^2) = 135{,}512$$

ดังนั้น $SS\ (total) = (\Sigma\Sigma X_{ij}^2 - CF) = (135{,}512 - 134{,}480) = 1032$

และ $SSE = (SST - SSA) = (1032 - 700) = 332$

หรือ $SSE = (\Sigma\Sigma X_{ij}^2 - \Sigma T_j^2/n_j)$

$= (135{,}512 - 135{,}180) = 332$

และ $MSE = 332/16 = 20.75$ เช่นเดียวกัน

จ) จงคำนวณค่า F และสรุปด้วยระดับนัยสำคัญ 5%

1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu$

2) Ha : μ_j ไม่เท่ากันทั้งหมด $\quad \gamma_1 = (k - 1) = (4 - 1) = 3$

3) $\alpha = .05$ $\quad \gamma_2 = (N - k) = (20 - 4) = 16$

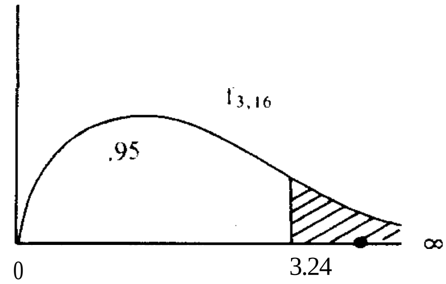
4) จะปฏิเสธ Ho เมื่อ $F > f_{3,16;.05} = 3.24$

5) $F = \hat{\sigma}_1^2/\hat{\sigma}_2^2$

$= MSA/MSE$

$= 233.33/20.75$

$= 11.24$



6) F = 11.24 > 3.24 จึงปฏิเสธ Ho และยอมรับ Ha สรุปว่า วิธีส่งเสริมการขาย 4 วิธีนั้น ให้จำนวนขายไม่เท่ากันทั้งหมด เก็บค่าต่าง ๆ ใส่ตารางวิเคราะห์ได้ดังนี้

SOV	df	SS	MS	F-ratio
A = ระหว่างกลุ่ม	3	700	233.33	11.24
E = ภายในกลุ่ม	16	332	20.75	
รวม	19	1032		

10.39 จำนวนผู้ยื่นคำร้องขอรับเงินประกันใน 1 วัน ของพนักงานบริษัทประกันภัย 5 คน มีดังนี้

							T_j	n_j
พนักงาน 1	15	17	14	11			57	4
พนักงาน 2	12	10	13	17	14		66	5
พนักงาน 3	10	14	13	15	12		64	5
พนักงาน 4	14	9	7	10	8	7	55	6
พนักงาน 5	13	12	9	14	10	9	67	6
							G = 309	26 = N

จงใช้ระดับนัยสำคัญ 1% ทดสอบว่า จำนวนผู้ร้องขอรับเงินประกันต่อวันของพนักงานทั้ง 5 คน ไม่ต่างกันอย่างมีนัยอย่างสำคัญ

1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu$

2) Ha : μ_j ไม่เท่ากันทั้งหมด, $j = 1, 2, \ldots, k$

$$k = 5$$

3) $a = 0.01$

4) จะปฏิเสธ Ho เมื่อ $F > f_{4,21;.01} = 4.37$

$\gamma_1 = (5 - 1) = 4$

$\gamma_2 = (26 \quad 5) = 21$

5) $CF = (309)^2/26 = 3672.36$

$\Sigma\Sigma X_{ij}^2 = (15^2 + 17^2 + \quad + 10^2 + 9^2) = 3873$

$\Sigma T_j^2/n_j = \left(\frac{57^2}{4} + \frac{66^2}{5} + \frac{64^2}{5} + \frac{55^2}{6} + \frac{67^2}{6} \right) = 3754.98$

ดังนั้น

$SST = \Sigma\Sigma X_{ij}^2 - CF$

$= 3873 - 3672.35 = 200.65$; $df = (20 - 1) = 19$

$SSA = \Sigma T_j^2/n_j - CF$

$= 3754.98 - 3672.35 = 82.63$; $df = (5 \quad 1) = 4$

$SSE = \Sigma\Sigma X_{ij}^2 - \Sigma T_j^2/n_j$

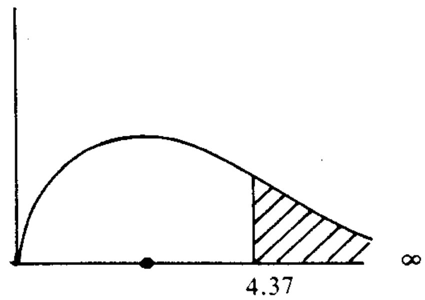
$= 3873 - 3754.98 = 118.02$; $df = (20 - 5) = 15$

หรือ $SST \quad SSA$

$= 200.65 - 82.63 = 118.02$

SOV	df	SS	MS	F-ratio
A = ระหว่างกลุ่ม	4	82.63	20.6575	2.62
E = ภายในกลุ่ม	15	118.02	7.868	
รวม	19	200.65		

6) F = 2.62 ไม่อยู่ในเขตวิกฤต จึงยังไม่ปฏิเสธ Ho สรุปว่า จำนวนผู้ยื่นคำร้องต่อวัน ต่อพนักงานประกันทั้ง 5 คน ไม่ต่างกันอย่างมีนัยสำคัญ



10.40 จากข้อมูลที่กำหนดให้ข้างล่าง 4 กลุ่มตัวอย่าง เราจะสรุปว่ามาจากประชากรที่มีค่าเฉลี่ยเท่ากัน ด้วยระดับนัยสำคัญ 5% ได้หรือไม่?

							T_j	n_j
ตัวอย่างที่ 1	17	22	25	29	30		123	5
ตัวอย่างที่ 2	29	18	20	19	30	21	137	6
ตัวอย่างที่ 3	13	14	20	18	27	16	108	6
ตัวอย่างที่ 4	21	28	20	22	18		109	5
							G = 477	22 = N

1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu$

2) Ha : μ_j ไม่เท่ากันทั้งหมด ; j = 1, 2, 3, 4, ; k = 4

3) $\alpha = .05$

4) จะปฏิเสธ Ho เมื่อ $F > f_{3,18;.05} = 3.16$

$\gamma_1 = (4 \quad = 3$

$\gamma_2 = (N - k) = (22 - 4) = 18$

5) $CF = G^2/N = (477)^2/22 = 10342.23$

$$\Sigma T_j^2/n_j = \frac{123^2}{5} + \frac{137^2}{6} + \frac{108^2}{6} + \frac{109^2}{5} = 10,474.17$$

$$\Sigma\Sigma X_{ij}^2 = 7^2 + 22^2 + \quad + 22^2 + 18^2 = 10,913$$

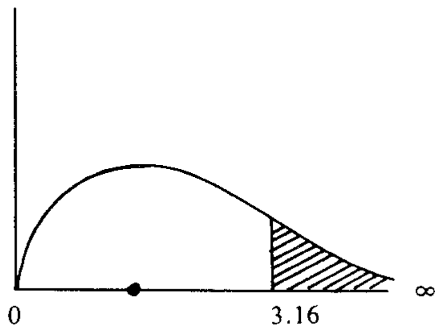
$$SST = \Sigma\Sigma X_{ij}^2 - CF$$

$$= 10,913 \quad 10,342.23$$

$$= 570.77$$

$$SSA = \Sigma T_j^2/n_j - CF$$
$$= 10,474.17 - 10,342.23$$
$$= 131.94$$
$$SSE = \Sigma\Sigma X_{ij}^2 - \Sigma T_j^2/n_j$$
$$= 10,913 - 10,474.17$$

หรือ $SSE = SST - SSA$
$$= 570.77 - 131.94$$
$$= 438.83$$



sov	df	SS	MS	F-ratio
A = ระหว่างกลุ่ม	3	131.94	43.98	1.80
E = ภายในกลุ่ม	18	438.83	24.38	
	21	570.77		

6) F = 1.80 ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho และสรุปว่า ตัวอย่างทั้ง 4 มาจากประชากรที่มีค่าเฉลี่ยเท่ากัน

10.41 ผู้จัดการโรงงานประดิษฐ์นาฬิกาต้องการทราบว่า เมื่อใช้อัตราความเร็วของสายพานลำเลียงวัสดุต่างกัน จะมีผลกระทบต่ออัตราชำรุดหรือไม่ เขาได้ทดลองใช้ความเร็ว 4 อัตรา ในช่วงการทำงาน 5 คาบเวลาๆ ละ 8 ชั่วโมง และได้บันทึกจำนวนสินค้าชำรุดในแต่ละคาบเวลาไว้ดังนี้

อัตราความเร็ว

1	2	3	4
36	29	31	36
34	34	35	28
37	34	32	34
35	36	33	32
33	32	39	30

ก) จงหาจำนวนชำรุดโดยเฉลี่ยของแต่ละอัตราความเร็ว และหาค่าเฉลี่ยรวบยอด

ข) จงหาความแปรปรวนระหว่างค่าเฉลี่ยของตัวอย่าง

ค) จงหาความแปรปรวนภายในกลุ่มตัวอย่าง

ง) จงคำนวณอัตราส่วน F และสรุปผลที่ $\alpha = .05$

$T_1 = 175,\ T_2 = 165,\ T_3 = 170,\ T_4 = 160$

$n_1 = n_2 = n_3 = n_4 = n = 5;\ k = 4$

$N = \Sigma n_j = 20,\ \ G = \Sigma T_j = 670$

ก) จำนวนชำรุดโดยเฉลี่ยของอัตราความเร็วต่าง ๆ มีดังนี้

$\overline{X}_1 = 175/5 = 35;\ \overline{X}_2 = 165/5 = 33$

$\overline{X}_3 = 170/5 = 34;\ \overline{X}_4 = 160/5 = 32$

$\overline{\overline{X}} = 670/20 = 33.5$

ข) ความแปรปรวนระหว่างค่าเฉลี่ยของตัวอย่าง คือ $V(\overline{X}_j)$

$$V(\overline{X}_j) = \sum_{j=1}^{k=4} (\overline{X}_j - \overline{\overline{X}})^2 / (k-1)$$

$$= ((35 - 33.5)^2 + (34 - 33.5)' + (33 - 33.5)^2 + (32 - 33.5)^2 \} / 3$$

$$= \{ (1.5)^2 + (0.5)^2 + (-0.5)^2 + (-1.5)' \} / 3$$

$$= 5/3 = 1.67$$

ค) จงหาความแปรปรวนภายในกลุ่มตัวอย่าง

$$S_1^2 = \sum_{i=1}^{5} (X - \overline{X}_1)^2 / (n_1 - 1). \qquad \overline{X}_1 = 35$$

$$= \{ (36 - 35)^2 \text{ t } (34 - 35)' + (37 \quad 35)^2 + (35 - 35)^2 + (33 - 35)^2 \} / 4$$

$$= \{ (1)^2 + (-1)^2 + (2)^2 + (0)^2 + (-2)^2 \} / 4$$

$$= 10/4$$

$$S_2^2 = \sum_{i=1}^{5} (X \quad \overline{X}_2)^2 / (n_2 \quad 1), \qquad \overline{X}_2 = 33$$

$$= \{ (-4)^2 + (+1)^2 + (+1)^2 + (3)^2 + (-1)^2 \} / 4$$

$$= 28/4$$

$$S_3^2 = \Sigma(X - \overline{X}_3)^2 / (n_3 - 1), \quad \overline{X}_3 = 34$$
$$= \{(-3)^2 + (1)^2 + (-2)^2 + (-1)^2 + (5)^2\} / 4$$
$$= 40/4$$
$$S_4^2 = \Sigma(X - \overline{X}_4)^2 / (n_4 - 1), \ \overline{X}_4 = 32$$
$$= \{(4)^2 + (-4)^2 + (2)^2 + (0)^2 + (-2)^2\} / 4$$
$$= 40/4$$

ความแปรปรวนภายใน กลุ่มตัวอย่าง = MSE = $S_p^2 = \Sigma S_j^2 / k$ เมื่อ n_j เท่ากัน

$$S_p^2 = \left(\frac{10}{4} + \frac{28}{4} + \frac{40}{4} + \frac{40}{4} \right) / 4$$

$$\frac{118}{16} \ 7.375$$

ง) จงคำนวณอัตราส่วนของ F และสรุปผลโดยใช้ $\alpha = .05$

$$F = \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$$

จากข้อ (ข) ได้ $V(\overline{X}_j) = S_{\overline{X}}^2 = 1.67$

แต่ $V(X) = nS_{\overline{X}}^2 = 5(1.67) = 8.35 = \hat{\sigma}_1^2$

และ $\hat{\sigma}_2^2 = S_p^2$

$$F = \frac{8.35}{7.375} = 1.13$$

$f_{3,16;.05} = 3.24$

ค่า F ที่คำนวณได้ ไม่อยู่ในเขตวิกฤต จึงยอมรับ Ho ว่า จำนวนชำรุดเฉลี่ยเมื่อใช้อัตราความเร็วต่าง ๆ กันไม่มีความแตกต่างกันอย่างมีนัยสำคัญ

10.42 ร้านอัดขยายรูป ต้องการทราบความแตกต่างของวิธีส่งเสริมการขาย 3 วิธี คือ

1) แถมฟิล์ม 1 ม้วน ฟรีสำหรับทุกม้วนที่ล้าง-อัดเป็นเวลา 6 สัปดาห์

2) ลดราคาทุก ๆ ม้วนที่ล้างอัด ม้วนละ 20 บาท เป็นเวลา 6 สัปดาห์ ต่อมา

3) คิดราคาปกติ 5 สัปดาห์ ก่อนใช้วิธีแจกของหรือลดราคา

จำนวนฟิล์มที่ลูกค้านำมาล้าง-อัดต่อสัปดาห์ เมื่อใช้วิธีต่าง ๆ มีดังนี้

วิธี	จำนวนล้าง-อัด ต่อสัปดาห์						T_j	n_j
แจกฟรี 1 ม้วน	65	79	73	55	68	74	414	6
ลดราคา 20 บาท	60	64	57	75	62	56	374	6
คิดราคาปกติ	61	54	74	59	46		294	5
k = 3							G = 1,082	17 = N

ก) จงหาจำนวนขายโดยเฉลี่ยต่อสัปดาห์ของแต่ละวิธี และค่าเฉลี่ยรวมยอด

ข) จงหาความแปรปรวนระหว่างค่าเฉลี่ยของตัวอย่าง

ค) จงหาค่าประมาณของความแปรปรวนของประชากร โดยใช้ความแปรปรวนระหว่างกลุ่ม

ง) จงหาความแปรปรวนภายในกลุ่มตัวอย่าง และหาค่าประมาณของความแปรปรวนของประชากรจากความแปรปรวนภายในกลุ่ม

จ) จงคำนวณค่า F และสรุปผล โดยใช้ $\alpha = .05$

ก) จำนวนขายเฉลี่ยของแต่ละวิธี คือ

$\bar{X}_1 = 414/6 = 69$, $\bar{X}_2 = 374/6 = 62,33$, $\bar{X}_3 = 294/5 = 58.8$

และ ค่าเฉลี่ยรวมยอดคือ $\bar{\bar{X}} = G/N = 1082/17 = 63.65$

ข) ความแปรปรวนระหว่างค่าเฉลี่ยของตัวอย่าง $= V(\bar{X}_j) = S^2_{\bar{X}} = \Sigma(\bar{X}_j - \bar{\bar{X}})^2/(k-1)$-

$$= \frac{1}{(3-1)}[(69 - 63.65)^2 + (62.33 - 63.65)^2 + (58.8 - 63.65)^2]$$

$$= \frac{1}{2}(53.8874) = 26.94$$

ค) $\hat{\sigma}^2 = V(X) \quad = \quad n\,V(\bar{X})$ เพราะ $S^2_{\bar{X}} = S^2/n$

$= \quad n\,S^2_{\bar{X}}$ แต่เนื่องจาก กลุ่มที่ 3 มีขนาดตัวอย่างต่างจากกลุ่มอื่น ๆ

จึงต้องใช้ n_j ถ่วงน้ำหนักก่อน

จากสูตรเดิม $V(\bar{X}_j) = \sum_j^k (\bar{X}_j - \bar{\bar{X}})^2/(k-1)$

ดังนั้น $V(X) \quad = \quad n_j\,V(\bar{X}_j) = \sum_j n_j(\bar{X}_j - \bar{\bar{X}})^2/(k-1)$

$$= \frac{1}{(3-1)}\{6(69-63.65)' + 6(62.33-63.65)' + 5(58.8-63.65)'\}$$

$$= 299.80/2 = 149.90$$

ง) หาความแปรปรวนภายในกลุ่มตัวอย่าง คือ S_j^2

$$S_1^2 = \Sigma(X-\bar{X}_1)^2/5 = \frac{1}{5}\{\Sigma X_1^2 - (\Sigma X_1)^2/6\}$$

$$= \frac{1}{5}(28{,}920 \quad (414)^2/6) = \frac{354}{5} = 70.8$$

$$S_2^2 = \Sigma(X-\bar{X}_2)^2/5 = \frac{1}{5}\{\Sigma X_2^2 - (\Sigma X_2)^2/n\}$$

$$= \frac{1}{5}\{23550 - (374)^2/6\} = 237.33/5 = 47.47$$

$$S_3^2 = \Sigma(X-\bar{X}_3)^2/4 = \frac{1}{4}\{\Sigma X_3^2 \quad (\Sigma X_3)^2/n\}$$

$$= \frac{1}{4}\{17710 - (294)^2/5\} = 422.8/4 = 105.7$$

ค่าประมาณของความแปรปรวนคือ

$$\hat{\sigma}^2 = S_p^2 = \frac{(n_1-1)S_1^2 + (n_2 \quad 1)S_2^2 + (n_3-1)S_3^2}{n_1+n_2+n_3-3}$$

$$= \frac{\Sigma(X-\bar{X}_1)^2 + \Sigma(X-\bar{X}_2)^2 + \Sigma(X-\bar{X}_3)^2}{N-3}$$

$$= \frac{354 + 237.33 + 422.8}{17-3}$$

$$= 1014.13/14 = 72.44$$

จ) $F = \hat{\sigma}_1^2/\hat{\sigma}_2^2$ = ความแปรปรวนระหว่างกลุ่ม/ความแปรปรวนภายในกลุ่ม

$$= 149.90/72.44$$

$$= 2.07$$

$f_{.05,2,14} = 3.74$ แต่ค่าสถิติ $F = 2.07 < 3.74$ จึงยอมรับ Ho และสรุปว่าวิธีส่งเสริมการขายทั้ง 3 วิธีให้ผลไม่แตกต่างกันอย่างมีนัยสำคัญ

10.43 ผู้จัดการฝ่ายรักษาความปลอดภัยของร้านสรรพสินค้าขนาดใหญ่ตั้งข้อสงสัยว่า จำนวนผู้ลักขโมยสิ่งของจากร้านในช่วงเทศกาลปีใหม่น่าจะสูงกว่าช่วงก่อนหรือหลังปีใหม่ เขาได้เก็บสถิติจำนวนผู้ลักขโมยไว้ 6 ปี มีดังนี้

							T_j
ก่อนเทศกาลปีใหม่	42	36	58	54	37	47	274
ระหว่างเทศกาลปีใหม่	51	38	45	32	47	46	259
หลังเทศกาลปีใหม่	37	29	35	42	31	33	207
						G =	740

จำนวนผู้ลักขโมยใน 3 ช่วงเวลาแตกต่างกันไหม เมื่อใช้ $\alpha = .05$

1) Ho : $\mu_1 = \mu_2 = \mu_3 = \mu$
2) มีอย่างน้อย 1 คู่ที่ต่างกัน
3) $\alpha = .05$
4) จะปฏิเสธ Ho เมื่อ $F > f_{2,15;.05} = 3.68$
5) $n_1 = n_2 = n_3 = 6,\ k = 3,\ N = 18,\ G = 740$

$CF = G^2/N = 30{,}422.22$

$\Sigma T_j^2/n_j = (274^2 + 259^2 + 207^2)/6 = 30{,}834.33$

$\Sigma\Sigma X_{ij}^2 = 31{,}586$

$SST = \Sigma\Sigma X_{ij}^2 - CF = 31{,}586 - 30{,}422.22 = 1163.78$

$SSA = \Sigma T_j^2/n_j - CF = 30{,}834.33 - 30{,}422.22 = 412.11$

$SSE = SST - SSA = 1163.78 - 412.11 = 751.61$

SOV	df	SS	MS	F-ratio
A = ระหว่างกลุ่ม	2	412.11	206.055	4.11
E = ภายในกลุ่ม	15	751.67	50.11	
รวม	17	1,163.78		

6) F = 4.1 I > 3.68 อยู่ในเขตวิกฤต จึงปฏิเสธ Ho และยอมรับ Ha นั่นคือ จำนวนผู้ลักขโมยใน 3 ช่วงเวลามีความแตกต่างกัน