$$-2\ln\lambda \simeq \sum_i^r\sum_j^c \frac{(X_{ij}-np_{ij})^2}{np_{ij}} > k'$$ [1]

และเมื่อสมมุติฐานหลักเป็นจริงจะพบว่า

$$-2\ln\lambda \simeq \sum_i^r\sum_j^c \frac{(X_{ij}-np_{i.}p_{.j})^2}{np_{i.}p_{.j}}$$ [2]

$$\simeq \sum_i^r\sum_j^c \frac{(X_{ij}-n(\frac{X_{i.}}{n}\ \frac{X_{.j}}{n}))^2}{n(\frac{X_{i.}}{n}\ \frac{X_{.j}}{n})}$$

$$= \sum_i^r\sum_j^c \frac{(X_{ij}-\frac{X_{i.}X_{.j}}{n})^2}{\frac{X_{i.}X_{.j}}{n}} \sim \chi^2_{(r-1)(c-1)}$$

จาก $\alpha = \Pr\{\text{Reject } H_0|H_0\}$

$= \Pr\{$ปฏิเสธสมมุติฐานหลักเมื่อสมมุติฐานหลักเป็นจริง$\}$

$= \Pr\{-2\ln\lambda > k' | p_{ij} = p_{i.}p_{.j}\ ;\ i = 1, 2, \ldots, r\ ;\ j = 1, 2, \ldots, c\}$

$$=> \Pr\{\sum_i^r\sum_j^c \frac{(X_{ij}-\frac{X_{i.}X_{.j}}{n})^2}{\frac{X_{i.}X_{.j}}{n}} > k'\} = \Pr\{\chi^2_{(r-1)(c-1)} > \chi^2_{(r-1)(c-1),1-\alpha}\}$$

$$=> \quad k' = \chi^2_{(r-1)(c-a),1-\alpha}$$

นั่นคือ ปฏิเสธสมมุตฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\sum_i^r\sum_j^c \frac{(X_{ij}-\frac{X_{i.}X_{.j}}{n})^2}{\frac{X_{i.}X_{.j}}{n}} > \chi^2_{(r-1)(c-a),1-\alpha}$$

---

[1] เสนอโดย Karl Pearson

[2] $\hat{p}_{i.} = \frac{X_{i.}}{n}$, $\hat{p}_{.j} = \frac{X_{.j}}{n}$

สำหรับการคำนวณหา β และ π จะขอยกเว้นไว้ไม่กล่าวถึง ขณะเดียวกัน การพัฒนาตัวทดสอบสำหรับ 3CT, 4CT และ CT ที่จำแนกมากกว่า 4 ทางขึ้นไป ก็จะงดไว้ไม่กล่าวถึง เพราะวิธีพัฒนาตัวทดสอบจะเป็นไปในทำนองเดียวกัน แต่จะสลับซับซ้อนกว่าอยู่บ้างโดยเฉพาะในกรณีของการหาค่าความน่าจะเป็นของเทอมสุดท้าย[1] เช่น

$$H_0 : p_{ijk} = p_{i..}p_{.j.}p_{..k} \text{ vs } H_1 : p_{ijk} \neq p_{i..}p_{.j.}p_{..k} ;$$

$i = 1, 2, \ldots, r ; j = 1, 2, \ldots, c ; k = 1, 2, \ldots, s$

จะพบว่า $$p_{rcs} = 1 - \sum_{i}^{r}\sum_{j}^{c-1}\sum_{k}^{s} p_{ijk} - \sum_{i}^{r-1}\sum_{k}^{s} p_{ick} - \sum_{k}^{s-1} p_{rck}$$

ค่า p_{rcs}, $p_{r..}$, $p_{.c.}$ และ $p_{..s}$ สามารถคำนวณหาได้โดยอาศัยภาพ 3 มิติ โดยวาดรูปในลักษณะลูกเต๋าซ้อน ๆ กัน เรื่องนี้ของดไว้ไม่กล่าวถึงโดยละเอียดแต่ขอเว้นไว้เป็นแบบฝึกหัดการคำนวณค่อนข้างซับซ้อนแต่มิใช่เรื่องที่ยากเกินสติปัญญาความสามารถ

ตัวอย่าง 6.21 เป็นที่น่าสังเกตว่าคนที่มีร่างกายสูงใหญ่ (น้ำหนักตัวมาก) มักจะเป็นบุคคลที่มีพฤติกรรมผิดปกติ และมีแนวโน้มก่ออาชญากรรมได้มากกว่าคนที่มีร่างกายเล็กกว่า

จากการทดสอบโดยใช้แบบสอบถามพฤติกรรมแก่บุคคลทั่วไป 493 คน ปรากฏข้อมูลจำนวนบุคคลจำนวนตามพฤติกรรม และน้ำหนักตัวดังนี้

พฤติกรรม	น้ำหนักตัว (ปอนด์)				
	9-120	120-130	130-140	140-150	150^{+}
ปกติ	21	51	94	106	124
ก้าวร้าว	15	18	34	15	15

จงทดสอบดูว่าน้ำหนักและพฤติกรรมเชิงอาชญากรรมของบุคคลเกี่ยวข้องกันหรือไม่ให้ใช้ระดับนัยสำคัญ 5%

---

[1] อ่าน Mood, Graybill and Boe **Introduction to the Theory of Statistics**, 3rd ed. หน้า 459

วิธีทำ จากข้อมูล สามารถจัดตารางได้ดังนี้

พฤติกรรม	น้ำหนักตัว					รวม ($X_{i.}$)
	90-120	120-130	130-140	140-150	150+	
ปกติ	21	51	94	106	124	396 ($X_{1.}$)
ก้าวร้าว	15	18	34	5	15	97 ($X_{2.}$)
รวม ($X_{.j}$)	36	69	128	121	139	493
	($X_{.1}$)	($X_{.2}$)	($X_{.3}$)	($X_{.4}$)	($X_{.5}$)	

H_0 : น้ำหนักและพฤติกรรมเชิงอาชญากรรมของบุคคลมิได้เกี่ยวข้องกัน หรือพฤติกรรมเชิงอาชญากรรมมิได้ผันแปรไปด้วยสาเหตุอันสืบเนื่องหรือพัวพันมาจากน้ำหนักตัวหรือความสูงใหญ่ของร่างกาย

H_1 : น้ำหนักและพฤติกรรมเชิงอาชญากรรมของบุคคลมีส่วนเกี่ยวข้องพัวพันกัน

จากตารางพบว่า $X_{1.} = 396$, $X_{2.} = 97$, $X_{..} = 493 = n$

$X_{.1} = 36$, $X_{.2} = 69$, $X_{.3} = 128$, $X_{.4} = 121$, $X_{.5} = 129$

และเนื่องจาก $E_{ij} = \frac{X_{i.}X_{.j}}{n}$; $i = 1, 2$; $j = 1, 2, ..., 5$

ดังนั้น

$$E_{11} = \frac{396 \times 36}{493} = 28.917 \; , \; E_{12} = \frac{396 \times 69}{493} = 55.424$$

$$E_{13} = \frac{396 \times 128}{493} = 102.815 \; , \; E_{14} = \frac{396 \times 121}{493} = 97.193$$

$$E_{15} = \frac{396 \times 139}{493} = 111.651 \; , \; E_{21} = \frac{97 \times 36}{493} = 7.083$$

$$E_{22} = \frac{97 \times 69}{493} = 13.576 \; , \; E_{23} = \frac{97 \times 128}{493} = 25.185$$

$$E_{24} = \frac{97 \times 121}{493} = 23.807 \; , \; E_{25} = \frac{97 \times 139}{493} = 27.349$$

$$\chi^2_c \;=\; \sum_{i}^{2}\sum_{j}^{5} \frac{(X_{ij} - E_{ij})^2}{E_{ij}} = 21.65$$

$$\chi^2_{(r-1)(c-a),1-\alpha} \;=\; \chi^2_{4,.95} \;=\; 9.488$$

จะเห็นว่า $\chi^2_c > 9.488$ ดังนั้นเราจึงปฏิเสธสมมุติฐานหลักและเชื่อว่าน้ำหนักตัวหรือความสูงใหญ่ของร่างกาย มีส่วนเกี่ยวเนื่องกับพฤติกรรมเชิงอาชญากรรม

ตัวอย่าง 6.22 เป็นที่กล่าวขวัญกันว่ายา A มีประสิทธิภาพในการรักษาไข้หวัดสูงมาก จากการทดลองให้ยาแก่คนไข้ 164 ราย โดย 82 ราย ให้รับประทานยา A อีก 82 รายให้รับประทานน้ำตาลที่อัดเม็ด และเคลือบผิวเหมือนยา A ปรากฏผลการรักษาดังนี้

ตัวยา	อาการของไข้หวัด		
	ดีขึ้น	ทรุดลง	ทรงตัว
ยา A	52	10	20
น้ำตาล	44	12	26

จงทดสอบสมมุติฐานว่า ยา A และน้ำตาลไม่ได้ให้ผลแตกต่างกันเลยในการรักษาไข้หวัด ให้ใช้ระดับนัยสำคัญ 5%

วิธีทำ H_0 : $p_{ij} = p_{i.}p_{.j}$; i = 1, 2, ; j = 1, 2, 3,

H_1 : $p_{ij} = p_{ij} \neq p_{i.}p_{.j}$

หรือ H_0 : อาการไข้หวัดที่เปลี่ยนแปลงไปมิได้เกี่ยวข้องกับชนิดของตัวยาที่รักษา[1]

H_1 : อาการไข้หวัดเปลี่ยนแปลงไป เพราะอิทธิพลของตัวยาที่ใช้รักษา

จากข้อมูลพบว่า

[1] เราตีความหมาย H_0 ไปได้หลายแบบ แบบทั่ว ๆ ไปก็คือ A ไม่เกี่ยวข้องกับ B ตามตัวอย่างนี้คือ "อาการไข้หวัดไม่เกี่ยวข้องกับชนิดของตัวยาที่ใช้รักษา" แต่ถ้าจะให้มีความหมายมากขึ้นอาจกล่าวว่า "อาการไข้หวัดมิได้เปลี่ยนแปลงไปในทางหนึ่งทางใดตามชนิดของตัวยาที่ใช้รักษา" หรือ "ตัวยาที่ใช้รักษามิได้ให้ผลในการรักษาแตกต่างกันเลย" ก็ได้ ทั้งนี้เราอาจมองจาก A ไปหา B หรือ B ไปหา A ก็ได้

$$X_{1.} = \sum_{j}^{3} X_{1j} = 52 + 10 + 20 = 82;\ X_{2.} = \sum_{j}^{3} X_{2j} = 44 + 12 + 26 = 82$$

$$X_{.1} = \sum_{i}^{2} X_{i1} = 52 + 44 = 96\ ;\ X_{.2} = \sum_{i}^{2} X_{i2} = 10 + 12 = 22$$

$$X_{.3} = \sum_{i}^{2} X_{i3} = 20 + 26 = 46$$

$$X_{..} = \sum_{i}^{3}\sum_{j}^{2} X_{ij} = \sum_{i}^{3} X_{i.} = \sum_{j}^{2} X_{.j} = n = 164$$

ตัวยา	อาการไข้หวัด			รวม ($X_{i.}$)
	ดีขึ้น	ทรุดลง	ทรงตัว	
ยา A	52	10	20	82 ($X_{1.}$)
น้ำตาล	44	12	26	82 ($X_{2.}$)
รวม ($X_{.j}$)	96	22	46	164
	($X_{.1}$)	($X_{.2}$)	($X_{.3}$)	

$$\because\ E_{ij} = \frac{X_{i.}X_{.j}}{n}\ ;\ j = 1, 2,;\ j = 1, 2, 3$$

ดังนั้น

$$E_{11} = \frac{82 \times 96}{164} = 48\ ,\ E_{12} = \frac{82 \times 22}{164} = 11\ ,\ E_{13} = \frac{82 \times 46}{164} = 23$$

$$E_{21} = \frac{82 \times 96}{164} = 48\ ,\ E_{22} = \frac{82 \times 22}{164} = 11\ ,\ E_{23} = \frac{82 \times 46}{164} = 23$$

$$\chi_c^2 = \sum_{i}^{2}\sum_{j}^{3} \frac{(X_{ij} - E_{ij})^2}{E_{ij}} = 1.631$$

$$\chi^2_{(r-1)(c-a),1-a} = \chi^2_{2,.95} = 5.991$$

จะเห็นว่า $\chi_c^2 < 5.991$ เราจึงไม่อาจปฏิเสธสมมุติฐานหลัก และเชื่อว่าอาการเปลี่ยนแปลงของไข้หวัดไม่พัวพันเกี่ยวข้องกับชนิดของตัวยา หรือนัยหนึ่งไม่ว่าจะใช้ตัวยาใดหรือไม่ก็ตามก็ไม่มีผลต่ออาการของไข้หวัด

ตัวอย่าง 6.23 จากการสังเกตพบว่านักศึกษาที่เรียนวิชาทฤษฎีความน่าจะเป็นได้ดีมักจะเรียนวิชาทฤษฎีสถิติได้ดี

จากการบันทึกเกรดวิชาทั้งสอง ปรากฏจำนวนผู้สอบได้เกรดต่าง ๆ ดังนี้

วิชาทฤษฎีความน่าจะเป็น	วิชาทฤษฎีสถิติ			
	A	B	C	D
A	38	24	8	2
B	16	12	8	8
C	16	19	35	23
D	3	6	7	13

จงทดสอบดูว่าข้อสังเกตข้างต้นเป็นความจริงหรือไม่

วิธีทำ $H_0 : p_{ij} = p_{i.}p_{.j}$; $i = 1, 2, 3, 4$; $j = 1, 2, 3, 4$

ข้อมูลความถี่สังเกตและความถี่คาดหมายปรากฏดังนี้ ข้อมูลในวงเล็บคือข้อมูลความถี่คาดหมาย

วิชาทฤษฎีความน่าจะเป็น	วิชาทฤษฎีสถิติ				รวม
	A	B	C	D	
A	38(20.79)	24(18.93)	8(18)	2(14.28)	72
B	16(12.71)	12(11.57)	8(11)	8(8.72)	44
C	16(25.13)	19(22.88)	35(21.75)	23(17.25)	87
D	3(8.38)	6(7.63)	7(7.25)	13(5.75)	29
รวม	67	61	58	46	232

$$\chi^2_c = \sum_i^4 \sum_j^4 \frac{(X_{ij} - E_{ij})^2}{E_{ij}} = 66.173168$$

$$\chi^2_{(r-1)(c-a),1-\alpha} = \chi^2_{9,.95} = 16.92$$

จะเห็นได้ว่า $\chi_c^2 > 16.92$ ดังนั้นจึงปฏิเสธสมมุติฐานหลักและเชื่อว่าผลการเรียนวิชาทั้งสองเกี่ยวข้องหรือมีอิทธิพลสืบเนื่องถึงกัน หรือผู้เรียนวิชาทฤษฎีความน่าจะเป็นได้ดี จะเรียนวิชาทฤษฎีสถิติได้ดีด้วย

6.3.4 ดัชนีการกระจาย (Index of Dispersion)

ในหลายกรณีของงานทดสอบ เรามักสนใจเปรียบเทียบความแตกต่างระหว่างค่าสัดส่วน หรือนัยหนึ่งต้องการทราบว่าข้อมูลที่มีอยู่ โดยบันทึกมาจากหลายแหล่งนั้น เป็นข้อมูลจากกลุ่มประชากรทวินามเดียวกันหรือไม่ หรือในกรณีเฉพาะเมื่อ p มีค่าน้อย และ $n \to \infty$ ข้อมูลเหล่านั้นมาจากกลุ่มประชากรพัวซองเดียวกันหรือไม่ หรือเสนอเป็นสมมุติฐานทางสถิติได้ดังนี้

$$H_0 \ : \ p_1 = p_2 = \ldots = p_k = p$$

และในกรณีเฉพาะ

$$H_0 \ ; \ \lambda_1 = \lambda_2 = \ldots = \lambda_k = \lambda$$

ให้ $X_i; i = 1, 2, \ldots, k$ คือจำนวน success จากกลุ่มตัวอย่าง n หน่วยที่สุ่มมาจากกลุ่มประชากรแบบทวินามที่ i ที่มีพารามิเตอร์ $p_i; i = 1, 2, \ldots, k$ หรือนัยหนึ่ง X_i คือความถี่ของ p_i ในกลุ่มตัวอย่างขนาด n เท่ากันที่สุ่มมาจากกลุ่มประชากรทวินามที่ $i; i = 1, 2, \ldots, k$

ดังนั้น $n - X_i$ จึงเป็นความถี่ของ q_i หรือจำนวน failure ในกลุ่มตัวอย่างที่สุ่มมาจากกลุ่มประชากรที่ i เดียวกัน $i = 1, 2, \ldots, k$

ปัญหาก็คือ กลุ่มประชากรทั้งหลายเหล่านั้นเป็นกลุ่มประชากรทวินามเดียวกันหรือไม่? ดัชนีที่จะช่วยตอบคำถามนี้ก็คือค่า p_i ซึ่งเป็นพารามิเตอร์ของประชากรทวินาม ถ้าประชากรกลุ่มต่าง ๆ มีค่าพารามิเตอร์ p เท่ากัน แสดงว่ากลุ่มประชากรเหล่านั้นมีลักษณะละม้ายกัน (Homogeneous) ซึ่งเราพออนุโลมว่าเป็นกลุ่มเดียวกัน หรือตัวอย่างที่สุ่มมาจากกลุ่มประชากรต่าง ๆ สุ่มมาจากกลุ่มประชากรเดียวกัน

ด้วยเหตุนี้ เราจึงตอบคำถามข้างต้นได้ว่า ถ้าสามารถแสดงให้เห็นว่า $p_i = p; i = 1,2,\ldots,k$ เมื่อใด เมื่อนั้นก็เชื่อได้ว่าประชากรกลุ่มต่าง ๆ ก็คือกลุ่มเดียวกัน

ดังนั้นสิ่งที่ต้องการตัดสินใจในขั้นต้นก็คือสมมุติฐานต่อไปนี้คือ

$$H_0 \ : \ p_i = p; i = 1, 2, \ldots, k$$

หรือ $H_0 \ : \ p_1 = p_2 = \ldots = p_k = p$

$H_1 \ : \ p_i \neq p$ สำหรับบางค่าของ $i; i = 1, 2, \ldots, k$

โดยอาศัยวิธีการของ 2CT โดยอนุโลมเราสามารถสร้างตารางจำแนกได้ดังนี้

						รวม
	X_1	X_2	X_3	...	X_k	$\sum_i^k X_i$
	$n - X_1$	$n - X_2$	$n - X_3$	...	$n - X_k$	$kn - \sum_i^k X_i$
รวม	n	n	n	...	n	kn

จาก $$E_{ij} = n\hat{p}_{i.}\hat{p}_{.j} = \frac{X_{i.}X_{.j}}{n}$$

ดังนั้น $$E_{1j} = \frac{n\Sigma X_i}{kn} = \frac{\Sigma X_i}{k} = \overline{X}\,;\ j = 1, 2, \ldots, k$$

$$E_{2j} = \frac{n(kn - \Sigma X_i)}{kn} = n - \overline{X}\,;\ j = 1, 2, \ldots, k$$

ดังนั้น $$\chi_c^2 = \sum_i^2 \sum_j^k \frac{(X_{ij} - E_{ij})^2}{E_{ij}}$$

$$= \frac{(X_1 - \overline{X})^2}{\overline{X}} + \frac{(X_2 + \overline{X})^2}{\overline{X}} + \ldots + \frac{(X_k - \overline{X})^2}{\overline{X}}$$

$$+ \frac{\{(n - X_1) - (n - \overline{X})\}^2}{n - \overline{X}} + \frac{\{(n - X_2) - (n - \overline{X})\}^2}{n - \overline{X}}$$

$$+ \ldots + \frac{\{(n - X_k) - (n - \overline{X})\}^2}{n - \overline{X}}$$

$$= \sum_i^k \frac{(X_i - \overline{X})^2}{\overline{X}} + \sum_i^k \frac{(X_i - \overline{X})^2}{n - \overline{X}}$$

$$= \left(\frac{1}{\overline{X}} + \frac{1}{n - \overline{X}}\right) \sum_i^k (X_i - \overline{X})^2$$

$$= \frac{n - \overline{X} + \overline{X}}{\overline{X}(n - \overline{X})} \sum_i^k (X_i - \overline{X})^2$$

$$= \{\sum_i^k (X_i - \overline{X})^2\} / \{\overline{X}(n - \overline{X})/n\}$$

$$= \frac{\Sigma(X_i - \overline{X})^2}{\overline{X}(1 - \frac{\overline{X}}{n})}$$

พิจารณาเศษจะพบว่า $\sum_{i}^{k}(X_i - \overline{X})^2 = (k-1)S^2$

พิจารณาส่วนจะพบว่า $\overline{X}(1 - \frac{\overline{X}}{n}) = \hat{\mu}(1 - \frac{\hat{\mu}}{n}) = n\hat{p}(1 - \frac{n\hat{p}}{n}) = n\hat{p}\hat{q}$

แสดงว่า $\overline{X}(1 - \frac{\overline{X}}{n})$ เป็นตัวประมาณค่าของความแปรปรวนของตัวแปรสุ่มทวินาม

นั่นคือ
$$\chi_c^2 = \frac{\sum^{k}(X_i - \overline{X})^2}{\overline{X}(1 - \frac{\overline{X}}{n})} \sim \chi^2_{(r-1)(k-1)} = \chi^2_{(k-1)}$$

และเมื่อพิจารณา χ_c^2 จะเห็นว่า $\chi_c^2 \simeq \frac{(k-1)s^2}{n\hat{p}\hat{q}}$ ซึ่งอยู่ในรูปอัตราส่วนของความแปรปรวน (Variance Ratio)[1] เมื่อใดก็ตามที่ X_i แตกต่างไปจาก $\overline{X}$ (ซึ่ง $\overline{X} = \hat{\mu} = n\hat{p}$) มากเพียงใด ย่อมหมายความว่า X_i เบี่ยงเบนมากไปจากค่าเฉลี่ยของกลุ่มประชากรทวินามร่วม (b(n, p)) มากเพียงนั้น ซึ่งในกรณีเช่นนี้จะสังเกตได้ว่าเศษของ χ_c^2 จะมีค่าสูงขึ้นอันจะมีผลให้ χ_c^2 มีค่าสูงขึ้น และมีแนวโน้มให้ $\chi_c^2 > \chi^2_{(r-1)(c-1),1-\alpha}$ หรือปฏิเสธสมมุติฐานหลักได้ง่ายขึ้นเราเรียก χ_c^2 ในรูปอัตราส่วนข้างต้นว่า ดัชนีการกระจายแบบทวินาม (Binomial Index of Dispersion)

ดังนั้นเราจึงปฏิเสธสมมุติฐานหลัก $H_0 : p_i = p$ vs $H_1 : p_i \neq p \,;\, i = 1, 2, \ldots, k$ ณ ระดับนัยสำคัญ α เมื่อ ดัชนีการกระจายทวินามมีค่ามากกว่า $\chi^2_{(k-1),1-\alpha}$ นั่นคือปฏิเสธสมมุติฐานหลักเมื่อ

$$\frac{\sum(X_i - \overline{X})^2}{\overline{X}(1 - \frac{\overline{X}}{n})} > \chi^2_{k-1,1-\alpha} \qquad \text{..........(1)}$$

หรือ
$$\frac{(k-1)s^2}{n\hat{p}\hat{q}} > \chi^2_{k-1,1-\alpha}{}^{2} \qquad \text{..........(2)}$$

---

[1] ดูการทดสอบ $H_0 : \sigma_X^2 = \sigma_Y^2$ ในบทที่ 7

[2] $\hat{p} = \frac{\overline{X}}{n}$, $\hat{q} = 1 - \hat{p}$, $r = 2$, $c = k$, $(r-1)(c-1) = (2-1)(k-1) = k-1$

ในกรณีเฉพาะเมื่อ $n \to \infty$ จะพบว่าเมื่อ $n \to \infty$ แล้ว $\frac{X}{n} = \hat{p} \to 0$ ซึ่งในกรณีเช่นนี้เป็นกรณีของการแจกแจงแบบพัวซอง[1]

ดังนั้นกรณีเฉพาะนี้สมมุติฐานที่ใช้ทดสอบคือ

$H_0 : \lambda_1 = \lambda_2 = \ldots = \lambda_k$ หรือ $\lambda_i = \lambda; i = 1, 2, \ldots, k$

$H_1 : \lambda_1 \neq \lambda; i = 1, 2, \ldots, k$

ทั้งนี้ λ_i คือค่าคาดหมายและความแปรปรวนของตัวแปรสุ่มพัวซอง และปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\sum_{i}^{k} \frac{(X_i - \overline{X})^2}{\overline{X}} > \chi^2_{k-1,1-\alpha} \qquad \text{..........(1)}$$

หรือ

$$\frac{(k-1)s^2}{\overline{X}} > \chi^2_{k-1,1-\alpha} \qquad \text{..........(2)}$$

และเนื่องจาก $\hat{\lambda} = n\hat{p}$ แสดงว่า $\hat{\lambda} = \overline{X}$ และด้วยเหตุที่ $E(U) = V(U) = \lambda$ ถ้า $U \sim$ Poisson (λ) ดังนั้น จึงพบว่า $\chi^2_c = \frac{(k-1)s^2}{\overline{X}} = \frac{(k-1)s^2}{\hat{\lambda}}$ ซึ่งก็คืออัตราส่วนระหว่างความแปรปรวนเช่นเดียวกัน

ดังนั้นตัวทดสอบที่ใช้อีกรูปหนึ่งก็คือ $\chi^2_c = \frac{(k-1)s^2}{\hat{\lambda}}$ นั่นคือปฏิเสธสมมุติฐานหลัก ณ ระดับนัยสำคัญ α เมื่อ

$$\frac{(k-1)s^2}{\hat{\lambda}} > \chi^2_{k-1,1-\alpha} \qquad \text{..........(3)}$$

อนึ่งเมื่อพิจารณาสูตรที่ (1) จะพบว่า $\chi^2_c = \frac{\sum^{k}(X_i - \overline{X})^2}{\overline{X}}$ คำนวณมาจากตาราง 2CT โดยใช้เฉพาะแถวที่ 1 เท่านั้น ดังนั้นในทางปฏิบัติถ้าต้องการทดสอบดูว่าข้อมูลจากกลุ่มตัวอย่างที่สุ่มมาจากหลายแหล่ง (กลุ่มประชากร) น่าจะเป็นกลุ่มตัวอย่างที่สุ่มมาจากกลุ่มประชากรพัวซองที่มีพารามิเตอร์ λ เดียวกันหรือไม่ จึงไม่มีความจำเป็นต้องจำแนกข้อมูลเป็น 2CT ก็ได้

---

[1] อ่านตอน 4.2.5 วิธีพิสูจน์ที่ 2 หน้า 134-135

สำหรับกรณีทั่วไปเมื่อขนาดตัวอย่างที่ใช้มีขนาดไม่เท่ากัน คือกลุ่มตัวอย่างที่ 1 ใช้ตัวอย่างขนาด n_1 กลุ่มที่ 2 ใช้ตัวอย่างขนาด $n_2, \ldots$, กลุ่มตัวอย่างที่ k ใช้ตัวอย่างขนาด n_k ให้ใช้ตาราง 2CT ข้างต้นได้เช่นกัน และวิเคราะห์ข้อมูลโดยนัยของ 2CT ส่วนตัวทดสอบจะลดรูปมาเป็นดัชนีของการแจกแจงแบบพัวซอง หรือไม่ยังเป็นปัญหา เรื่องนี้จึงขอเว้นไว้เป็นแบบฝึกหัด

ตัวอย่าง 6.24 เป็นที่สงสัยว่าอัตราการระบาดของโรคพืชในแปลงสาธิตที่มีอยู่ทั้งสิ้น 12 แปลง จะเป็นไปในอัตราเดียวกันหรือไม่ จากการสุ่มพืชตัวอย่างจากแปลงสาธิตทั้ง 12 แปลง ๆ ละ 90 ต้น พบว่ามีจำนวนพืชที่ติดโรคในแต่ละแปลงนี้

19, 6, 9, 18, 15, 13, 14, 15, 16, 20, 22, 14

จงทดสอบ ณ ระดับนัยสำคัญ 1% ว่า จะเชื่อได้หรือไม่ที่อัตราการระบาดของโรคพืชมีอัตราเดียวกัน

วิธีทำ $H_0 : p_i = p; i = 1, 2, \ldots, 12$ vs $H_1 : p_i \neq p; i = 1, 2, \ldots, 12$

$H_0 : p_1 = p_2 = \ldots = p_{12} = p$ vs $H_1 : p$ อย่างน้อย 1 คู่ไม่เท่ากัน

จากข้อมูลความถี่สามารถจัดเป็นตาราง 2CT ได้ดังนี้

จำนวนต้น	แปลงที่												รวม
	1	2	3	4	5	6	7	8	9	10	11	12	
เป็นโรค	19	6	9	18	15	13	14	15	16	20	22	14	181
ไม่เป็นโรค	71	84	81	72	75	77	76	75	74	70	68	76	899
รวม	90	90	90	90	90	90	90	90	90	90	90	90	1080

$$E_{ij} = \frac{X_{i.}X_{.j}}{n} ; i = 1, 2, ; j = 1, 2, \ldots, 12$$

$$E_{11} = E_{12} = \ldots = E_{1,12} = 15.083 = \overline{X}$$

$$E_{21} = E_{22} = \ldots = E_{2,12} = 74.917 = n - \overline{X} ; n = 90, kn = 1,080$$

$$\chi_c^2 = \sum_i^2 \sum_j^{12} \frac{(X_{ij} - E_{ij})^2}{E_{ij}} = \frac{(19 - 15.083)^2}{15.083} + \frac{(6 - 15.083)^2}{15.083}$$

$$+ \ldots + \frac{(76 - 74.917)^2}{74.917}$$

$$= 17.755$$

$$\chi^2_{11,.95} = 19.675$$

จะเห็นว่า $\chi_c^2 < 19.675$ เราจึงไม่อาจปฏิเสธสมมุติฐานหลักได้ และเชื่อว่าอัตราโรคระบาดของโรคพืชในแปลงสาธิตทั้ง 12 แปลงมีอัตราเดียวกัน หรือในแง่ของความน่าจะเป็นเราจะถือว่าข้อมูลต่าง ๆ ในตารางข้างต้นต่างก็สุ่มมาจากประชากรทวินามกลุ่มเดียวกัน

หมายเหตุ ตัวอย่างนี้นักศึกษาวิเคราะห์โดยใช้สูตรที่ (2) และ (3) ต่อไปนี้ก็ได้เพราะให้ผลลัพธ์ตรงกัน

$$\chi_c^2 = \{\sum_i^k (X_i - \overline{X})^2\}/\{\overline{X}(1 - \frac{\overline{X}}{n})\}$$

$$\chi_c^2 = \frac{(k-1)s^2}{n\hat{p}\hat{q}}$$

โดยที่ $n = 90, k = 12$

$$\overline{X} = \frac{\sum^k X_i}{k} = 181/12 = 15.083; \frac{\overline{X}}{n} = \frac{15.083}{90} = 0.168 = \hat{p}; \hat{q} = .832$$

$$s^2 = \frac{1}{k-1} \{ \sum_i^k X_i^2 - (\sum_i^k X_i)^2\} = \frac{1}{11} (2,953 - \frac{32,761}{12})$$

$$= \frac{222.917}{11} = 20.265$$

อย่างไรก็ตาม สูตรทั้งสองนี้พัฒนามาจากกรณีเมื่อขนาดตัวอย่างจากประชากรต่าง ๆ มีขนาดเท่ากัน ถ้าขนาดตัวอย่างต่างกันโครงสร้างของสูตรจะเปลี่ยนไปซึ่งเป็นภาระที่นักศึกษาจะต้องพัฒนาขึ้นเอง ดังนั้นเพื่อป้องกันปัญหาต่าง ๆ ที่อาจปรากฏขึ้นเราควรใช้สูตรที่ (1) ในทั้งสองกรณี

ตัวอย่าง 6.25 ในการพิสูจน์อักษร ผู้พิสูจน์อักษรต้องการทราบว่าช่างเรียงพิมพ์ 11 คน เรียงพิมพ์ผิดพลาดในอัตราเดียวกันหรือไม่

สุ่มต้นฉบับเรียงพิมพ์จากช่างเรียงพิมพ์ทั้ง 11 คน มารายละ 1,000 คำ ปรากฏคำผิดดังนี้ 15, 13, 8, 6, 11, 9, 14, 10, 16, 9, 12 คำตามลำดับ

จงทดสอบ ณ ระดับนัยสำคัญ 5% ว่าช่างเรียงพิมพ์ 11 คนเรียงคำผิดในอัตราเดียวกันหรือไม่

วิธีทำ จากข้อมูลความถี่ นำมาจำแนกเป็นตาราง 2CT ได้ดังนี้

จำนวนคำ	ช่างเรียงที่											รวม
	1	2	3	4	5	6	7	8	9	10	11	
คำผิด	15	13	8	6	11	9	14	10	16	9	12	123
คำถูก	985	987	992	994	989	991	986	990	984	991	988	10877
รวม	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	1000	11000

พิจารณาข้อมูลจะพบว่า

$$\bar{X} = 123/11 = 11.182 = \hat{\lambda}$$

$$\hat{p} = \frac{\bar{X}}{n} = 11.182/1000 = 0.11182$$ ซึ่งมีค่าต่ำมาก ขณะที่ n มีค่าสูงมาก

$(n = 1,000)$

ด้วยเหตุนี้จึงถือว่าสมมุติฐานที่มุ่งทดสอบคือ

$H_0 : \lambda_1 = \lambda_2 = \ldots = \lambda_k = \lambda$ vs $H_1 : \lambda$ อย่างน้อย 1 คู่ไม่เท่ากัน

$$\chi^2_c \simeq \frac{\sum^{11}(X_i - \overline{X})^2}{\overline{X}} = \frac{1}{11.182}(1,473 - \frac{15,129}{11}) = \frac{97.6364}{11.182}$$

$$= 8.732$$

$$\chi^2_{10,.95} = 18.307$$

จะเห็นว่า $\chi^2_c < 18.307$ ดังนั้นจึงไม่อาจจะปฏิเสธสมมุติฐานหลักและเชื่อว่าช่างเรียงทั้ง 11 คนจะเรียงพิมพ์ผิดในอัตราเดียวกัน หรือสรุปในเชิงความน่าจะเป็นได้ว่า ข้อมูลต่าง ๆ ทั้ง 11 แหล่งสุ่มมาจากกลุ่มประชากรพัวซองเดียวกัน

ขณะเดียวกัน ถ้าใช้สูตรของ 2CT โดยตรงคือ $\chi^2_c = \sum_i^2 \sum_j^{11} \frac{(X_{ij} - E_{ij})^2}{E_{ij}}$

จะพบว่า $E_{11} = E_{12} = \ldots = E_{1,11} = 11.282$

$E_{21} = E_{22} = \ldots = E_{2,11} = 988.818$

$\chi^2_c = 8.231$

จะเห็นว่าวิธีทั้งสองให้ค่า χ^2_c ใกล้เคียงกัน แต่สูตร $\chi^2 = \sum_i^{11} \frac{(X_i - \overline{X})^2}{\overline{X}}$ วิเคราะห์ง่ายกว่า

6.4 Sequential Probability Ratio Test (SPRT)

งานทดสอบสมมุติฐานบางลักษณะ โดยเฉพาะอย่างยิ่งคืองานตรวจสอบคุณภาพของวัตถุ (Sampling Inspection, Acceptance Sampling) นั้น ขนาดตัวอย่างที่ใช้จะไม่คงที่แน่นอนแต่เป็นไปในลักษณะที่สุ่มตัวอย่างมาครั้งละหน่วย ๆ เรื่อยไปจนกว่าจะทำให้ข้อมูลที่ได้ซึ่งรวมตัวกันเป็นตัวสถิติหนึ่ง ตกอยู่ในเขตวิกฤติหรือเขตยอมรับเขตใดเขตหนึ่ง หรือนัยหนึ่งขนาดตัวอย่างทำหน้าที่เป็นตัวแปรมิใช่ค่าคงที่ดังที่ผ่านมา

ดังนั้นโดยนัยของ SPRT ระนาบจะถูกจำแนกออกเป็น 3 ส่วน คือ Acceptance Zone (AZ) Rejection Zone (RZ); Indifferent Zone (IZ) ซึ่งการตัดสินใจก็เป็นไปตามความหมายของคำเหล่านี้ กล่าวคือผลจากการสุ่มตัวอย่างครั้งหนึ่ง ๆ ถ้ามีผลให้ตัวสถิติตกอยู่ใน IZ แสดงว่าเรายังไม่อาจชี้ชัดว่าควรยอมรับหรือปฏิเสธสมมุติฐานหลัก หรือยังคลุมเคลือ ไม่อาจตัดสินใจได้ วิธีที่ดีที่สุดคือเพิ่มตัวอย่างให้ใหญ่ขึ้นหรือสุ่มตัวอย่างหน่วยต่อไป เพื่อให้ได้ข้อสนเทศมากขึ้น ถ้าค่าสถิติตกอยู่ใน AZ เราจะยอมรับสมมุติฐานหลักหรือนัยหนึ่งเราไม่อาจปฏิเสธสมมุติฐานหลัก โดยนัยกลับกันถ้าค่าสถิติตกอยู่ใน RZ เราจะปฏิเสธสมมุติฐานหลัก สิ่งที่ใช้แบ่งระนาบออกเป็น 3 โซน

คือเส้นตรง 2 เส้นที่เกิดจาก Likelihood Ratio เรียกว่า Acceptance Line (AL) และ Rejection Line (RL) การคำนวณหาเส้นตรงทั้งสองมิใช่เรื่องยากเพียงอาศัยหลักและวิธีการของ Likelihood Ratio ที่เคยใช้ในตอนที่ผ่านมาเท่านั้น

วิธีการของ SPRT โดยสรุปก็คือค่อย ๆ สุ่มตัวอย่างมาคราวละ 1 หน่วย และเมื่อสุ่มตัวอย่างครั้งหนึ่ง จะคำนวณหา Likelihood Ratio ตามขนาดตัวอย่างที่มีอยู่ขณะนั้นคราวหนึ่ง ให้กระทำดังนี้เรื่อย ๆ ไปจนกว่าจะตัดสินใจได้ หรือสรุปได้ด้วยสัญลักษณ์และหลักการดังนี้

จาก NPL เราจะปฏิเสธสมมุติฐานหลักเมื่อ $\lambda = \frac{L_0}{L_1} \leqslant k\ ;\ k \geqslant 0$

ดังนั้น

กำหนดให้ $$\lambda_m = \frac{L_0(x_1, x_2, \ldots, x_m)}{L_1(x_1, x_2, \ldots, x_m)} = \frac{\prod^{m} f_X(x_i|H_0)}{\prod_{i}^{m} f_X(x_i|H_1)}$$

คือ Likelihood Ratio ของตัวแปรสุ่มจากกลุ่มตัวอย่างขนาด m เมื่อ $m = 1, 2, \ldots$ ให้ k_0 และ k_1 เป็นตัวคงที่ 2 ตัวใด ๆ ที่สามารถปรับค่าไปได้จนกระทั่งทำให้พื้นที่ในเขตวิกฤติเท่ากับ α และ β ตามลำดับ เมื่อ $0 < k_0 < k_1$

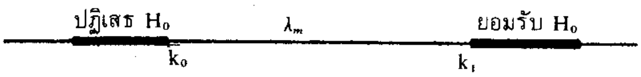
ดังนั้น

1. เมื่อสุ่มตัวอย่างมา 1 หน่วย เราจะปฏิเสธสมมุติฐานหลักเมื่อ[1]

$\lambda_1 = \frac{L_0(x_1)}{L_1(x_1)} \leqslant k_0$ ยอมรับสมมุติฐานหลักเมื่อ $\lambda_1 = \frac{L_0(x_1)}{L_1(x_1)} \geqslant k_1$ และไม่ตัดสินใจและสุ่มตัวอย่างหน่วยที่ 2 เมื่อ $k_0 < \lambda_1 < k_1$

---

[1] แนวคิด SPRT พัฒนาดังนี้:- จาก NPL เราจะปฏิเสธสมมุติฐานหลักเมื่อ $\lambda = \frac{L_0}{L_1} \leqslant k$ และยอมรับสมมุติฐานหลักเมื่อ $\frac{L_0}{L_1} > k$ ดังนั้นถ้ากำหนดค่า k เสียใหม่เป็น k_0 และ k_1 เราจึงปฏิเสธสมมุติฐานหลักเมื่อ $\frac{L_0}{L_1} \leqslant k_0$ และยอมรับสมมุติฐานหลักเมื่อ $\frac{L_0}{L_1} \geqslant k_1$ ดังไดอะแกรม



เขตที่ว่างอยู่ก็คือ $k_0 < \frac{L_0}{L_1} < k_1$ จึงเป็นเขตที่ไม่ใช่ทั้งยอมรับสมมุติฐานหลักและปฏิเสธสมมุติฐานหลัก และให้ชื่อเสียใหม่ว่าเขตไม่ตัดสินใจ

2. สุ่มตัวอย่างมาอีก 1 หน่วย เราจะปฏิเสธสมมุติฐานหลักเมื่อ

$\lambda_2 = \dfrac{L_0(x_1, x_2)}{L_1(x_1, x_2)} \leqslant k_0$ ยอมรับสมมุติฐานหลักเมื่อ $\lambda_2 = \dfrac{L_0(x_1, x_2)}{L_1(x_1, x_2)} \geqslant k_1$ ไม่ตัดสินใจ

และสุ่มตัวอย่างเพิ่มอีก 1 หน่วย ถ้า $k_0 < \lambda_2 < k_1$

3. สุ่มตัวอย่างมาอีก 1 หน่วย เราจะปฏิเสธสมมุติฐานหลักเมื่อ

$\lambda_3 = \dfrac{L_0(x_1, x_2, x_3)}{L_1(x_1, x_2, x_3)} \leqslant k_0$ ยอมรับสมมุติฐานหลักเมื่อ $\lambda_3 \geqslant k$

และสุ่มตัวอย่างเพิ่มอีก 1 หน่วย ถ้า $k_0 < \lambda_3 < k_1$ ให้ดำเนินการเช่นนี้เรื่อย ๆ ไปจนกว่า $\lambda_m \leqslant k_0$ หรือ $\lambda_m \geqslant k_1$; $m = 1, 2, \ldots$

เขตวิกฤติสำหรับ SPRT จึงนิยามด้วย $R = \bigcup_{m=1}^{\infty} R_m$ โดยที่

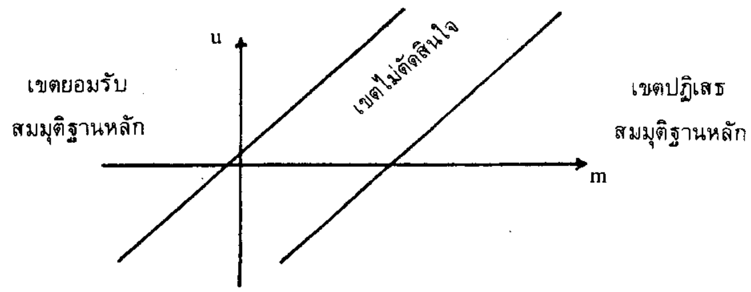
$$R_m = \{(x_1, x_2, \ldots, x_m) : k_0 < \lambda_j < k_1 \; ; \; j = 1, 2, \ldots, m-1 \; ; \; \lambda_m < k_0\}$$

และเขตยอมรับสำหรับ SPRT นิยามด้วย $A = \bigcup_{m=1}^{\infty} A_m$ โดยที่

$$A_m = \{(x_1, x_2, \ldots, x_m) : k_0 < A_j < k_1 \; ; \; j = 1, 2, \ldots, m-1; \; \lambda_m \geqslant k_1\}$$

หรือนัยหนึ่งการตัดสินใจจะเลื่อนไปเรื่อย ๆ ตราบใดที่ $k_0 < \lambda_j < k_1$ จนกระทั่งเมื่อใช้ขนาดตัวอย่างเท่ากับ m แล้วมีผลทำให้ $\lambda_m \leqslant k_0$ เราจะตัดสินใจปฏิเสธสมมุติฐานหลักทันที หรือยอมรับสมมุติฐานหลักทันทีที่ $\lambda_m \geqslant k_1$

ถ้าให้ $u(x_1, x_2, \ldots, x_m; \theta_0, \theta_1)$ เป็นตัวสถิติที่ได้มาจาก Likelihood Ratio ตามนัยข้างต้น ระนาบจะถูกแบ่งออกเป็น 3 โซน โดยอาศัย $u(x_1, x_2, \ldots, x_m; \theta_0, \theta_1)$ และ m ดังนี้



สำหรับค่า k_0 และ k_1 นั้นสามารถคำนวณหาได้จากสมการ α และ β เมื่อ

α = Pr{ปฏิเสธสมมุติฐานหลักทั้ง ๆ ที่สมมุติฐานหลักถูกต้อง} และ

β = Pr{ยอมรับสมมุติฐานหลักทั้ง ๆ ที่สมมุติฐานรองถูกต้อง}

ดังนั้นเราจึงคำนวณหาค่า k_0, k_1 ได้จากสมการต่อไปนี้

$$\alpha = \sum_{m=1}^{\infty} \int\int_{R_m} \cdots \int L_0(x_1, x_2, \ldots, x_m)dx_1dx_2\ldots dx_m$$

$$\beta = \sum_{m=1}^{\infty} \int\int_{A_m} \cdots \int L_1(x_1, x_2, \ldots, x_m)dx_1dx_2\ldots dx_m$$

ปัญหาก็คือ สมการทั้งสองเป็นสมการที่ซับซ้อนการคำนวณหาค่า k_0 และ k_1 จะต้องค่อย ๆ รวมพื้นที่ไปเรื่อย ๆ จนกว่าพื้นที่จะเท่ากับ α และ β จึงค่อยแก้สมการในภายหลัง จึงมิใช่เรื่องง่าย ดังนั้นในทางปฏิบัติเรานิยมประมาณค่า k_0 และ k_1 ขึ้นใช้ดังนี้คือ

$$k_0 \simeq \frac{\alpha}{1-\beta} \quad \text{หรือ } k_0 \text{ มีค่าอย่างน้อยเท่ากับ } \frac{\alpha}{1-\beta}$$

และ

$$k_1 \simeq \frac{1-\alpha}{\beta} \quad \text{หรือ } k_1 \text{ มีค่าไม่เกิน } \frac{1-\alpha}{\beta}$$

ค่าประมาณดังกล่าวข้างต้นพัฒนาขึ้นมาดังนี้

n . จาก $a = \Pr\{\text{Reject } H_0|H_0\}$

$$= \sum_{m=1}^{\infty} \int\int_{R_m} \ldots \int L_0(x_1, x_2, \ldots x_m)dx_1dx_2\ldots dx_m$$ [1]

---

[1] สมมุติ $\alpha = .05$ และปฏิเสธสมมุติฐานหลักเพื่อใช้ตัวอย่าง 4 หน่วย

เมื่อ $H_0 : \theta \neq \theta_0$ vs $H_1 : \theta = \theta_1$

$$\begin{aligned} .05 = {} & \int_{R_1} f_X(x_1, \theta_0)dx_1 + \int_{R_2}\int f_X(x_1, \theta_0)f_X(x_2, \theta_0)dx_1,dx_2 \\ & + \int_{R_3}\int\int f_X(x_1, \theta_0)f_X(x_2, \theta_0)f_X(x_3, \theta_0)dx_1dx_2dx_3 \\ & + \int_{R_4}\int\int\int f_X(x_1, \theta_0)f_X(x_2, \theta_0)f_X(x_3, \theta_0)f_X(x_4, \theta_0)dx_1dx_2dx_3dx_4 \end{aligned}$$

แต่เนื่องจากเราจะปฏิเสธสมมุติฐานหลักเมื่อ $\frac{L_0}{L_1} \leqslant k_0$ หรือ $L_0 \leqslant k_0 L_1$

ดังนั้น $$\alpha \leqslant \sum_{m=1}^{\infty} \int\int_{R_m} \ldots \int k_0 L_1(x_1, x_2, \ldots, x_m) dx_1 dx_2 \ldots dx_m$$

$$\leqslant k_0 \sum_{m=1}^{\infty} \int\int_{R_m} \ldots \int L_1(x_1, x_2, \ldots, x_m) dx_1 dx_2 \ldots dx_m$$

แต่เนื่องจาก R_m คือเขตวิกฤติที่ปฏิเสธสมมุติฐานหลัก

$$=> \quad \alpha \leqslant k_0 . Pr\{Reject\ H_0 | H_1\}$$

$$\leqslant k_0 . (1 - \beta)$$

$$=> \quad k_0 \geqslant \alpha / 1 - \beta$$

ข. จาก $1 - \alpha = Pr\{Accept\ H_0 | H_0\}$ [1]

$$= \sum_{m=1}^{\infty} \int\int_{A_m} \ldots \int L_0(x_1, x_2, \ldots, x_m) dx_1 dx_2 \ldots dx_m$$

แต่เนื่องจากเราจะยอมรับสมมุติฐานหลักเมื่อ $\frac{L_0}{L_1} \geqslant k_1$ หรือ $L_0 \geqslant k_1 L_1$

$$=> \quad 1 - \alpha \geqslant \sum_{m=1}^{\infty} \int\int_{A_m} \ldots \int k_1 L_1(x_1, x_2, \ldots, x_m) dx_1 dx_2 \ldots dx_m$$

$$=> \quad 1 - \alpha \geqslant k_1 \sum_{m=1}^{\infty} \int\int_{A_m} \ldots \int L_1(x_1, x_2, \ldots, x_m) dx_1 dx_2 \ldots dx_m$$

แต่เนื่องจาก A_m คือเขตที่ยอมรับสมมุติฐานหลัก

$$=> \quad 1 - \alpha \geqslant k_1 Pr\{Accept\ H_0 | H_1\}$$

$$\geqslant k_1 \beta$$

$$=> \quad k_1 \leqslant \frac{1 - \alpha}{\beta}$$

นั่นคือ $k_0 \geqslant \frac{\alpha}{1 - \beta}$ และ $k_1 \leqslant \frac{1 - \alpha}{\beta}$

ดังนั้นโดยนัยของ SPRT

---

[1] การพิสูจน์จะเริ่มจากสมการ $\alpha, \beta, 1 - \alpha$ หรือ $1 - \beta$ ก็ได้

ก. ปฏิเสธสมมุติฐานหลักเมื่อ $\lambda_m \leqslant k_0 \simeq \dfrac{\alpha}{1-\beta}$

ข. ยอมรับสมมุติฐานหลักเมื่อ $\lambda_m \geqslant k_1 \simeq \dfrac{1-\alpha}{\beta}$

ค. ไม่ตัดสินใจหรือดำเนินการสุ่มตัวอย่างต่อไปเมื่อ $\dfrac{\alpha}{1-\beta} < \lambda_m < \dfrac{1-\alpha}{\beta}$

ตัวอย่าง 6.26 สุ่มตัวแปรมา m หน่วย เมื่อ m = 1, 2, ... มาจากกลุ่มประชากรปกติ $N(\mu, 1)$ จงทดสอบสมมุติฐานต่อไปนี้ โดยวิธี SPRT

$$H_0 : \mu = \mu_0 \text{ vs } H_1 : \mu = \mu_1 > \mu_0$$

วิธีทำ กำหนดให้ความเสี่ยงประเภทที่ 1 และความเสี่ยงประเภทที่ 2 เท่ากับ α และ β ตามลำดับ และให้ m คือขนาดตัวอย่างที่มีค่าผันแปรไปได้คือ m = 1, 2, ...

ดังนั้นการตัดสินใจจะเลื่อนไปและสุ่มตัวอย่างเพิ่มเข้ามาคราวละ 1 หน่วยทุกครั้งที่

$$k_0 < \frac{L_0(x_1, x_2, \ldots, x_m)}{L_1(x_1, x_2, \ldots, x_m)} < k_1$$

$$\frac{\alpha}{1-\beta} < \frac{L_0(x_1, x_2, \ldots, x_m)}{L_1(x_1, x_2, \ldots, x_m)} < \frac{1-\alpha}{\beta}$$

$$\frac{\alpha}{1-\beta} < \frac{(1/\sqrt{2\pi})^m \exp\{-\frac{1}{2}\Sigma(x_i-\mu_0)^2\}}{(1/\sqrt{2\pi})^m \exp\{-\frac{1}{2}\Sigma(x_i-\mu_1)^2\}} < \frac{1-\alpha}{\beta}$$

$$\frac{\alpha}{1-\beta} < \exp\{+\frac{1}{2}\sum_i^m (x_i-\mu_1)^2 - \frac{1}{2}\sum_i^m (x_i-\mu_0)^2\} < \frac{1-\alpha}{\beta}$$

$$\frac{\alpha}{1-\beta} < \exp\{(\mu_0-\mu_1)\sum_i^m x_i + \frac{m}{2}\mu_1^2 - \frac{m}{2}\mu_0^2\} < \frac{1-\alpha}{\beta}$$

$$\ln(\frac{\alpha}{1-\beta} < (\mu_0-\mu_1)\sum_i^m x_i + \frac{m}{2}\mu_1^2 - \frac{m}{2}\mu_0^2 < \ln(\frac{1-\alpha}{\beta})$$

$$\ln\left(\frac{\alpha}{1-\beta}\right) - \frac{m}{2}\mu_1^2 + \frac{m}{2}\mu_0^2 < (\mu_0 - \mu_1)\sum_i^m x_i < \ln\left(\frac{1-\alpha}{\beta}\right) - \frac{m}{2}\mu_1^2 + \frac{m}{2}\mu_0^2$$

$$\frac{1}{\mu_0 - \mu_1}\ln\left(\frac{\alpha}{1-\beta}\right) + \frac{m}{2}(\mu_0 + \mu_1) \leqslant \sum_i^m x_i > \frac{1}{\mu_0 - \mu_1}\ln\left(\frac{1-\alpha}{\beta}\right) + \frac{m}{2}(\mu_0 + \mu_1)$$

..........(1)

และยอมรับสมมุติฐานหลักเมื่อ

$$\sum_i^m x_i \leqslant \frac{1}{\mu_0 - \mu_1}\ln\left(\frac{1-\alpha}{\beta}\right) + \frac{m}{2}(\mu_0 + \mu_1) \qquad \text{..........(2)}$$

ปฏิเสธสมมุติฐานหลักเมื่อ

$$\sum_i^m x_i \geqslant \frac{1}{\mu_0 - \mu_1}\ln\left(\frac{\alpha}{1-\beta}\right) + \frac{m}{2}(\mu_0 + \mu_1) \qquad \text{..........(3)}$$

ข้อสังเกต

ขอให้สังเกตว่าในสมการทั้งสาม $\sum_i^m x_i$ เป็นตัวสถิติ U ส่วนเขตวิกฤติและเขตยอมรับเป็นฟังก์ชั่นของตัวแปรสุ่ม m พิจารณาสมการที่ (2) และ (3) จะพบว่าเราสามารถจัดสมการให้เป็นรูปสมการเส้นตรงชนิด Intercept Form ได้ดังนี้

ให้ $\qquad y = \sum^m x_i$

ดังนั้นจากสมการที่ (2) จะพบว่ายอมรับสมมุติฐานหลักเมื่อ

$$\sum^m x_i \leqslant \frac{1}{\mu_0 - \mu_1}\ln\left(\frac{1-\alpha}{\beta}\right) + m \cdot \frac{\mu_0 + \mu_1}{2}$$

$$=> \qquad y \leqslant \frac{\mu_0 + \mu_1}{2} \cdot m + \frac{1}{\mu_0 - \mu_1}\ln\left(\frac{1-\alpha}{\beta}\right)$$

เมื่อ $\frac{\mu_0 + \mu_1}{2}$ คือความชัน (Slope) และ $\frac{1}{\mu_0 - \mu_1}\ln\left(\frac{1-\alpha}{\beta}\right)$ คือจุดตัดบนแกน [1] และจากสมการที่ (3) จะพบว่าปฏิเสธสมมุติฐานหลักเมื่อ

$$Y \quad a \quad \frac{\mu_0 + \mu_1}{2} \cdot m + \frac{1}{\mu_0 - \mu_1} h \left(\frac{L}{1-\beta}\right)$$

[1] y = mx + c เมื่อ m คือความชันและ c คือจุดตัดบนแกน y

เมื่อ $\frac{\mu_0 + \mu_1}{2}$ คือความชันและ $\frac{1}{\mu_0 - \mu_1} \ln \left(\frac{\alpha}{1-\beta} \right)$ คือจุดตัดบนแกน y

เพื่อให้สามารถมองเห็นภาพได้ชัดเจนขอกำหนดตัวอย่างให้เป็นกรณีเฉพาะดังนี้

$H_0 : \mu = 2$ vs $H_1 : \mu = 3$ กำหนดให้ $\alpha = \beta = .05$

จะพบว่า

$$\frac{\mu_0 + \mu_1}{2} = \frac{2+3}{2} = 2.5$$

$$\frac{1}{\mu_0 - \mu_1} \ln \left(\frac{1-\alpha}{\beta} \right) = \frac{1}{2-3} \ln \left(\frac{.95}{.05} \right) = -2.94444$$

$$\frac{1}{\mu_0 - \mu_1} \ln \left(\frac{\alpha}{1-\beta} \right) = \frac{1}{2-3} \ln \left(\frac{.05}{.95} \right) = +2.9444$$

นั่นคือยอมรับสมมุติฐานหลักเมื่อ

$$y \leqslant 2.5m - 2.94444$$

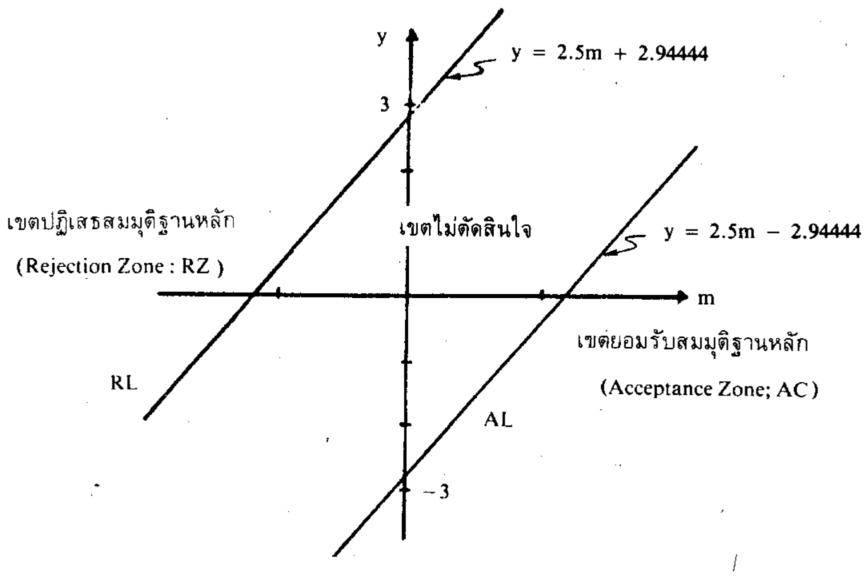
ปฏิเสธสมมุติฐานหลักเมื่อ

$$y \geqslant 2.5m + 2.94444$$

และสุ่มตัวอย่างเพิ่มเติมเมื่อ

$$2.5m - 2.94444 < y < 2.5m + 2.94444$$

จะเห็นว่าระนาบถูกแบ่งออกเป็น 3 โซนดังภาพ



การตัดสินใจให้พล็อตคู่ลำดับ (m, y) ลงในระนาบ ถ้าคู่ลำดับยังคงตกอยู่ใน IZ ให้สุ่มตัวอย่างต่อไป ถ้าคู่ลำดับตกอยู่ใน AZ หรือ RZ เขตใดเขตหนึ่ง ให้หยุดสุ่มตัวอย่าง และตัดสินยอมรับหรือปฏิเสธสมมุติฐานทันที

ตัวอย่าง 6.27 ข้อมูลต่อไปนี้ เป็นผลจากการตรวจสอบสภาพของวัตถุ ซึ่งดำเนินการตรวจคราวละ 1 หน่วยคือ

ggggg ggggg gbggb ggbbg เมื่อ b = bad g = good

ซึ่งการตรวจสภาพนี้มุ่งตรวจดูว่าวัตถุมีอัตราการเสื่อมสภาพเกินกว่า 5% หรือไม่ โดยการกำหนดเป็นข้อสมมุติฐานไว้ดังนี้

$H_0 : p = 0.05$ ns $H_1 : p = 0.15$

เมื่อ p แสดงสัดส่วนของวัตถุที่เสื่อมสภาพ

การทดสอบครั้งนี้ต้องการให้เกิดความเสี่ยง $\alpha = 10\%$ และ $\beta = 20\%$ จงทดสอบสมมุติฐานโดยนัยของ SPRT และตรวจดูว่าการตัดสินใจจะเกิดขึ้นเมื่อตรวจสอบคุณภาพของวัตถุถึงหน่วยที่เท่าไร

วิธีทำ ตัวแปรสุ่ม X มีการแจกแจงแบบเบอร์นุลลิมี pdf ดังนี้คือ

$f_X(x) = p^x q^{1-x}$; x = 0, 1 (1 = good, 0 = bad)

จากกรณีทั่วไปของสมมุติฐานที่มุ่งทดสอบเกี่ยวกับพารามิเตอร์ p จะพบว่า

$H_0 : p = p_0$ vs $H_1 : p = p_1 > p_0$

ดังนั้น โดยนัยของ SPRT เราจะสุ่มตัวอย่างต่อไปเมื่อ

$$\frac{\alpha}{1-\beta} < \frac{p_0^{\Sigma x_i} q_0^{m-\Sigma x_i}}{p_1^{\Sigma x_i} q_1^{(m-\Sigma x_i)}} < \frac{1-\alpha}{\beta}$$

$$\frac{\alpha}{1-\beta} < \left(\frac{p_0}{p_1}\right)^{\Sigma x_i}\left(\frac{q_0}{q_1}\right)^{m}\left(\frac{q_1}{q_0}\right)^{\Sigma x_i} < \frac{1-\alpha}{\beta}$$

$$\ln\left(\frac{\alpha}{1-\beta}\right) - m(\ln q_0 - \ln q_1) < \Sigma x_i\{(\ln p_0 - \ln p_1) + (\ln q_1 - \ln q_0)\}$$

$$< \ln\left(\frac{1-\alpha}{\beta}\right) - m(\ln q_0 - \ln q_1)$$

$$=> \frac{\ln\left(\frac{\alpha}{1-\beta}\right) - m(\ln q_0 - \ln q_1)}{(\ln p_0 - \ln p_1) + (\ln q_1 - \ln q_0)} \geqslant \sum_{i}^{m} x_i \geqslant \frac{\ln\left(\frac{1-\alpha}{\beta}\right) - m(\ln q_0 - \ln q_1)}{(\ln p_0 - \ln p_1) + (\ln q_1 - \ln q_0)}$$

ให้ $p_0 = 0.05,\ p_1 = 0.15,\ q_0 = 0.95,\ q_1 = 0.85$

$\alpha = 0.10$ และ $\beta = 0.20$

ดังนั้น

$$\frac{\ln\left(\frac{\alpha}{1-\beta}\right) - m(\ln q_0 - \ln q_1)}{(\ln p_0 - \ln p_1) + (\ln q_1 - \ln q_0)} = \frac{\ln\left(\frac{.10}{.80}\right) - m(\ln .95 - \ln .85)}{(\ln .05 - \ln .15) + (\ln .85 - \ln .95)}$$

$$= \frac{-2.079 - .111226m}{-1.2098}$$

$$= 0.09m + 1.719$$

$$\frac{\ln\left(\frac{1-\alpha}{\beta}\right) - m(\ln q_0 - \ln q_1)}{(\ln p_0 - \ln p_1) + (\ln q_1 - \ln q_0)} = \frac{\ln\left(\frac{.90}{.20}\right) - m(\ln .95 - \ln .85)}{(\ln .05 - \ln .15) + (\ln .85 - \ln .95)}$$

$$= \frac{1.504 - .111226m}{-1.2098} = 0.09m - 1.243$$

นั่นคือให้ดำเนินการสุ่มตัวอย่างต่อไปนี้ถ้า

$$\frac{-2.079 - 0.111226m}{-1.2098} \geqslant y \geqslant \frac{1.504 - 0.111226m}{-1.2098}$$

$$=> \quad 0.09m + 1.719 \geqslant y \geqslant 0.09m - 1.243$$

นั่นคือสุ่มตัวอย่างหน่วยต่อไปเมื่อ

$$0.09m - 1.243 \leqslant y \leqslant 0.09m + 1.719$$

ปฏิเสธสมมุติฐานหลักเมื่อ

$$y > 0.09m + 1.719$$

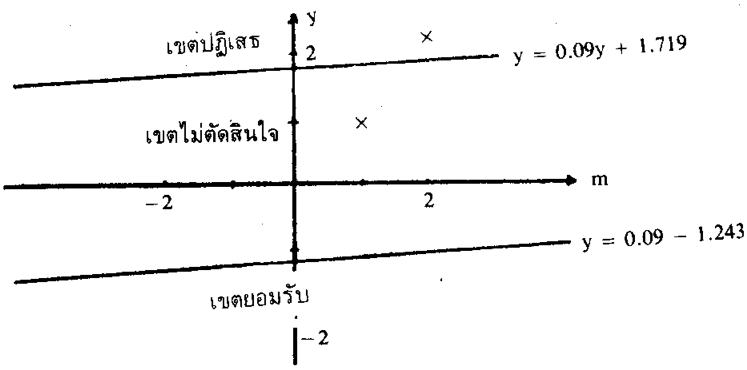
$\ln .95 = -0.051$, $\ln .85 = -0.163$, $\ln .05 = -2.996$, $\ln .15 = -1.897$

$\ln .10 = -2.303$, $\ln .80 = -0.223$, $\ln .90 = -0.105$

ยอมรับสมมุติฐานหลักเมื่อ

$$y < 0.09m - 1.243$$

และ IZ, AZ, RZ ปรากฏดังภาพ



จากข้อมูลสามารถจัดเป็นคู่ลำดับ (m, y) เมื่อ $g = 1, b = 0$ ได้ดังนี้

ขนาดตัวอย่าง	ผล	$y = \sum_{i}^{m} x_i$	คู่ลำดับ (m, y)
1	g	$1 = 1$	(1, 1)
2	gg	$1+1 = 2$	(2, 2)
3	ggg	$1+1+1 = 3$	(3, 3)
4	gggg	$1+1+1+1 = 4$	(4, 4)
5	ggggg	$1+1+1+1+1 = 5$	(5, 5)
.			
.			
.			

จะเห็นว่าเมื่อนำคู่ลำดับ (m, y) ไปพล็อตลงในระนาบ จุดโคออร์ดิเนตปรากฏอยู่ใน AZ เมื่อ $m = 2$ ดังนั้นจึงปฏิเสธสมมุติฐานหลัก และการทดลองจะกระทำโดยสุ่มตัวอย่างอยู่เพียง 2 ครั้ง ก็ลงข้อยุติได้แล้ว

แบบฝึกหัด

1. จงคำนวณหาความเสี่ยงประเภทที่ 2 สำหรับสมมุติฐาน

$$H_0 : \mu = \mu_0 = 10 \text{ vs } H_1 : \mu = \mu_1 = 11$$

ให้ใช้ระดับนัยสำคัญ 5% โดยถือว่าตัวแปรสุ่มมีการแจกแจงแบบปกติ มีส่วนเบี่ยงเบนมาตรฐานเท่ากับ 2.5 และขนาดตัวอย่าง $n = 16$

(0.518)

2. ในข้อ 1. ท่านควรสุ่มตัวอย่างมากี่ชุด จึงจะประกันได้ว่า เมื่อ H_1 จริง H_1 จะต้องได้รับการยอมรับถึง 90%

3. ตัวแปรสุ่มแบบตัดตอน X มีการแจกแจงดังนี้คือ

k	1	2	3	4	5	6	7	8	9	10
$Pr(X=k\|H_0)$	.02	.18	.07	.03	.17	.24	.02	.08	.01	.18
$Pr(X=k\|H_1)$	.10	.13	.04	.27	.01	.05	.14	.03	.11	.12

โดยมีกติกาการตัดสินใจว่า "เมื่อสุ่มตัวอย่างมา 1 หน่วย เราจะปฏิเสธสมมุติฐานหลักถ้า $k = 1$ หรือ 7 หรือ 9 มิเช่นนั้นจะยอมรับ

ก. จงหาความเสี่ยงประเภทที่ 1 และ Power ของกติกาดังกล่าว

ข. จงหาเขตวิกฤติ ณ. ระดับนัยสำคัญ 5%

4. เมื่อตัวแปรสุ่ม X มี pdf เป็น

$$f_X(x) = (1 + \lambda)x \quad ; \quad 0 \leqslant x \leqslant 1$$
$$= 0 \quad \text{ถ้า } x \text{ มีค่าเป็นอย่างอื่น}$$

จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$$H_0 : \lambda = \lambda_0 \text{ vs } H_1 : \lambda = \lambda_1 > \lambda_0$$

5. เมื่อตัวแปรสุ่ม X มี pdf เป็น $f_X(x) = \lambda^2 x e^{-\lambda x}$; $x \geqslant 0$; $\lambda \geqslant 0$

จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$$H_0 : \lambda = \lambda_0 \text{ vs } H_1 \lambda = \lambda > \lambda_0$$

6. ถ้าตัวแปรสุ่ม X มี pdf. เป็น $f_X(x) = pq^{x-1}$; $x = 1, 2, \ldots$
สร้างตัวสำหรับทดสอบสมมุติฐาน

$$H_0 : p = p_0 = .30 \text{ vs } H_1 : p = p < p_0$$

7. เป็นที่ทราบกันดีว่า อัตราเฉลี่ยของกระบวนการผลิตหนึ่งคงที่เท่ากับ 10 หน่วยเสมอ จงทดสอบสมมุติฐาน

$$H_0 : \sigma^2 = \sigma_0^2 = 2 \text{ vs } H_1 : \sigma^2 = \sigma_1^2 = 1.5$$

ถ้า $\sum_{i}^{n} (x_i - 10)^2 = 20$ และ $n = 15$

8. บริษัทผู้ผลิตอุปกรณ์อิเลคโทรนิคโฆษณาว่า อายุเฉลี่ยของอุปกรณ์ของตนมีอายุใช้งานเฉลี่ยเกินกว่า 1,000 ชั่วโมง และบริษัทผู้ใช้อุปกรณ์เหล่านี้จะตกลงซื้อไว้เป็นจำนวนมาก ถ้าคำโฆษณาของผู้ผลิตเป็นความจริง

    จากการสุ่มตัวอย่างอุปกรณ์เป็นตัวอย่าง 25 ชิ้น พบว่า อายุเฉลี่ยมีค่าเท่ากับ 1,050 ชั่วโมง และทราบว่า $\sigma = 100$ ชั่วโมง

    ก. อยากทราบว่าบริษัทผู้ใช้อุปกรณ์จะตกลงใจซื้อหรือไม่

    ข. ถ้าอายุเฉลี่ยจริงเท่ากับ 1,050 ชั่วโมง จงหา Power

    ค. ถ้า Power เท่ากับ .90 อยากทราบว่าควรสุ่มตัวอย่างมาทดลองกี่หน่วย

9. บริษัทผู้จำหน่ายกาแฟผงสำเร็จรูป กล่าวว่าน้ำหนักสุทธิของกาแฟคือ 1 ปอนด์ สุ่มตัวอย่างกาแฟผงมา 25 กระป๋อง พบว่าน้ำหนักสุทธิเฉลี่ยเท่ากับ 0.985 ปอนด์ และจากประสบการณ์ทราบว่า $\sigma = 0.03$ ปอนด์
จงทดสอบคำโฆษณาของบริษัทว่าเป็นความจริงหรือไม่ และหา Power ที่น้ำหนักสุทธิจริงมีค่าเท่ากับ 0.98 ปอนด์ (ไม่จริง $\pi = 0.954$)

10. ทราบว่าความเหนียวของโลหะชนิดหนึ่ง มีความแปรปรวนเท่ากับ 36 psi และถ้ามีการปรับปรุงการผลิตคาดว่าจะทำให้ความแปรปรวนลดลง จากการบันทึกข้อมูลปรากฏข้อมูลดังนี้

4	9	0	6	8
1	13	2	13	9
14	7	15	16	2

    ก. ท่านจะให้คำแนะนำแก่บริษัทอย่างไร ($\alpha = 0.05$)

    ข. จงคำนวณหาความเสี่ยงที่จะปฏิเสธความจริงที่ว่า $\sigma = 4$

    ค. ควรสุ่มตัวอย่างมากี่หน่วยจึงจะทำให้ความเสี่ยงในข้อ ข. ลดลง 50%

11. ต้องการทดสอบสมมุติฐานว่า ความต้านทานเฉลี่ยของตัวต้านทานชนิดหนึ่งมีค่าเท่ากับ 120 โอห์มหรือไม่ ถ้าหากว่าค่าความต้านทานเฉลี่ยเปลี่ยนแปลงไปเพียง 2 โอห์มเมื่อไรก็ต้องมีการปรับปรุงแบบของวงจร สมมุติว่า $\alpha = 0.01, \beta = 0.05, \sigma = 1.5$ อยากทราบว่าควรสุ่มตัวอย่างมาทดลองกี่หน่วย ควรใช้กติกาการใดในการตัดสินใจ? ถ้าความต้านทานเฉลี่ยจริงมีค่าเท่ากับ 121.5 โอห์ม

12. จากกลุ่มตัวอย่างขนาด n = 10 ของกลุ่มประชากรปกติ พบว่า $s^2 = 100$ จงทดสอบสมมุติฐาน

$$H_0 : \sigma^2 = \sigma_0^2 = 70 \text{ vs } H_1 : \sigma^2 = \sigma_1^2 \neq 70$$

ให้ใช้ $\alpha = 0.01$

ถ้า $\sigma_2 = 112$ จงคำนวณหาความเสี่ยงที่จะปฏิเสธสมมุติฐานหลัก

13. สุ่มตัวอย่างลวดที่ใช้ทำสายของเครื่องดนตรีประเภทเครื่องสายมา 10 เส้น ปรากฏความเหนียวดังนี้

ความเหนียว (psi)				
149,278	150,851	151,578	149,871	153,023
152,405	149,971	150,250	151,461	151,515

โดยมุ่งหมายที่จะทดสอบสมมุติฐานต่อไปนี้

$$H_0 : \mu = \mu_0 = 150{,}000 \text{ vs } H_1 : \mu = \mu_1 > 150{,}000$$

ก. ท่านจะสรุปผลว่าเส้นลวดมีความเหนียวเกินกว่า 150,000 psi ได้หรือไม่

ข. จงหาความน่าจะเป็นที่จะปฏิเสธสมมุติฐานว่า $\mu = 150{,}000$ psi เมื่อค่าจริงของความเหนียวเท่ากับ 151,000 psi

หมายเหตุ psi = pound per square inch

14. ตัวแปรสุ่ม X มี pdf เป็น $f_X(x) = 1/\lambda ; 0 < x < \lambda$

$= 0$ เมื่อ x มีค่าอื่น

โดย λ เป็นตัวพารามิเตอร์ที่ยังไม่ทราบค่า

ก. จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$$H_0 : \lambda = \lambda_0 = 1 \text{ vs } H_1 : \lambda = \lambda_1 < 1$$

$$H_0 : \lambda = \lambda_0 = 1 \text{ vs } H_1 : \lambda = \lambda_1 \neq 1$$

ข้อแนะนำ ให้ใช้ Order Statistics ช่วย

ข. ถ้าสุ่มตัวอย่างมา 3 หน่วยและตัดสินใจปฏิเสธสมมุติฐานหลัก เมื่อ $\max(x_1, x_2, x_3) \leqslant 6$ จงคำนวณหา α และ ถ้า $\lambda = 0.5$

จงคำนวณหา $\beta(0.5)$ และ $\Pi(0.5)$

15. ตัวแปรสุ่ม X มีการแจกแจงแบบเบอร์นุลลิ มี pdf. เป็น

$$f_X(x) = p^x q^{1-x} \; ; \; x = 0, 1$$

ก. จงสร้างตัวทดสอบสำหรับการทดสอบสมมุติฐานต่อไปนี้ กำหนดให้ใช้ $n = 10$

(1) $H_0 : p \leqslant \frac{1}{2}$ vs $H_1 : p > \frac{1}{2}$ และ

(2) $H_0 : p = \frac{1}{2}$ vs $H_1 : p = \frac{1}{4}$

ข. สำหรับข้อ (1) ถาใช้กติกาการตัดสินใจว่า "ปฏิเสธสมมุติฐานหลัก เมื่อ $\Sigma x \geqslant 6$ จงคำนวณหา β, Π พร้อมทั้งวาดโค้ง OC และโค้ง Power Function

16. เมื่อตัวแปรสุ่ม X มี pdf. เป็น $f_X(x) = \lambda x^{\lambda-1} \; ; 0 < x < 1$

ก. เมื่อใช้ $n = 2$ จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$H_0 : \lambda = 1$ vs $H_1 : \lambda = 2$

ทั้งนี้ให้ใช้ระดับนัยสำคัญ α และ $\frac{1}{2}(1 - \ln 2)$ ตามลำดับ

ข. สำหรับสมมุติฐาน $H_0 : \lambda < 1$ vs $H_1 : \lambda > 1$ เมื่อใช้กลุ่มตัวอย่างขนาด $n = 2$ และใช้กติกาการตัดสินใจเป็น "ปฏิเสธสมมุติฐานหลัก เมื่อ $\frac{3}{4} x_1 \leqslant x_2$" จงคำนวณหา α, β และ Π

17. เมื่อตัวแปรสุ่ม X มี pdf. เป็น $f_X(x) = \lambda x^{\lambda-1}; 0 < x < 1$

ก. ในการทดสอบสมมุติฐาน

$H_0 : \lambda < 1$ vs $H_1 : \lambda = \lambda_1 > 1$

โดยสุ่มตัวอย่างมาเพียง 1 หน่วย และตัดสินใจปฏิเสธสมมุติฐานหลักเมื่อ $x \geqslant \frac{1}{2}$ จงคำนวณหา α, β และ Π ของกติกานี้ ให้กำหนดค่า λ_1 เองตามต้องการ

ข. จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$H_0 : \lambda = 2$ vs $H_1 : \lambda = 1$

$H_0 : \lambda \geqslant 2$ vs $H_1 : \lambda < 2$

$H_0 : \lambda = 1$ vs $H_1 : \lambda \neq 1$

ทั้งนี้ให้ใช้ขนาดตัวอย่าง $n = 2, 3, 4$ และขยายความสู่กรณีทั่วไป เมื่อขนาดตัวอย่างเท่ากับ n

18. เมื่อตัวแปรสุ่ม X มี pdf. เป็น $f_X(x) = 2\theta x + 1 - \theta \; ; 0 < x < 1, \; -1 < \theta < 1$

ก. จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$H_0 : \theta = 0$ vs $H_1 : \theta \neq 0$

ข. จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$H_0 : \theta_1 = 0$ vs $H_1 : \theta_1 = 0$

โดยตัวทดสอบดังกล่าวจะต้อง minimize $(\alpha + \beta)$

19. เมื่อตัวแปรสุ่ม X มีการแจกแจงแบบยูนิฟอร์ม $(0, \theta)$ โดยสุ่มตัวอย่างมา 2 หน่วย เพื่อทดสอบสมมุติฐาน

$H_0 : \theta \leqslant 1$ vs $H_1 : \theta > 1$

และปฏิเสธสมมุติฐานหลักเมื่อ $x_1 + x_2 \geqslant 1$

ก. จงคำนวณหา α และ Π

ข. จงสร้างตัวทดสอบที่มีขนาดเขตวิกฤติเท่ากับข้อ ก. แล้วเปรียบเทียบกับกติกาในข้อ ก.

ข้อแนะนำ ดูการแจกแจงแบบเบต้า ตอน 4.2.3 หน้า 188-199 โดยเฉพาะตอน 4.2.3.2 หน้า 197

20. ตัวแปรสุ่ม X มี pdf. ดังนี้ $f_X(x) = \lambda^2 x e^{-\lambda x} ; 0 < x < \infty$

ก. ในการทดสอบสมมุติฐาน $H_0 : \lambda \leqslant 1$ vs $H_1 : \lambda > 1$

และเมื่อสุ่มตัวอย่างมาเพียง 1 หน่วย โดยจะใช้กติกาการตัดสินใจว่า "ปฏิเสธสมมุติฐานหลักเมื่อ $x \leqslant 1$" จงหา Power และ Size of Test

ข. จงสร้างตัวทดสอบสำหรับสมมุติฐาน $H_0 : \lambda = 1$ vs $H_1 : \lambda = 2$

21. สุ่มตัวอย่าง $X_1, X_2, \ldots, X_n$ มาจากกลุ่มประชากร $\lambda_1 x^{\lambda_1 - 1}; 0 < y < 1$

และสุ่มตัวอย่าง $Y_1, Y_2, \ldots, Y_n$ มาจากกลุ่มประชากร $\lambda 2 y^{\lambda_2 - 1} ; 0 < y < 1$

โดยตัวอย่างทั้ง 2 กลุ่มเป็นอิสระต่อกัน

ให้ $U_i = -\ln X_i ; i = 1, 2, \ldots, m$ และ

$V_j = -\ln Y_j ; j = 1, 2, \ldots, n$

ก. จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$H_0 : \lambda_1 = \lambda_2$ vs $H_1\ \lambda_1 \neq \lambda_2$

ข. จงแสดงให้เห็นว่าตัวทดสอบที่ต้องการคือ $T = (\Sigma U_i)/(\Sigma U_i + \Sigma V_j)$

22. ตัวแปสุ่ม X มี pdf. เป็น $f_X(x) = (1 + \theta)x^{\theta} ; 0 < x < 1, \theta > -1$

เมื่อสุ่มตัวอย่างมาเพียง 1 หน่วย

ก. จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$H_0 : \theta = 0$ vs $H_1 : \theta = 1$ และ

$H_0 : \theta \leqslant 0$ vs $H_1 : \theta > 1$

ข. ตัวทดสอบทั้งสองในข้อ ก. เป็น Uniform Most Powerful Test หรือไม่

ข้อแนะนำ สำหรับข้อ ข. ให้ดูตัวอย่างที่ 2 หน้า 281 Hogg and Craig

23. สุ่มตัวอย่างมา 4 หน่วยจากกลุ่มประชากร $N(\mu, 1)$ ได้ข้อมูลดังนี้คือ $(-.2, -.9, -.6, 1)$ จงทดสอบสมมุติฐาน

$$H_0 : \mu \leqslant 0 \text{ vs } H_1 : \mu > 0$$

24. เมื่อสุ่มตัวอย่าง $X_1, X_2, \ldots, X_n$ จากกลุ่มประชากร $N(\mu, \sigma^2)$ จงสร้างตัวทดสอบสำหรับสมมุติฐาน

$$H_0 : \mu = 6 \text{ vs } H_1 : \mu = 4$$

25. จากกลุ่มตัวอย่างขนาด $n = 100$ หน่วย $\bar{x} = 2.7$ และ $\Sigma(x_i - \bar{x})^2 = 225$ จงทดสอบสมมุติฐาน $H_0 : \mu = 3$, $H_0 : \sigma^2 = 2.5$ และ $H_0 : \mu = \sigma^2$ สมมุติว่าประชากรมีการแจกแจงแบบปกติ

26. จงสร้างตัวทดสอบสำหรับทดสอบค่าสหสัมพันธ์ (ϱ) ของกลุ่มประชากร Bivariate Normal Distribution

ข้อแนะนำ $$f_{X,Y}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1 - \varrho^2}} \cdot \exp\{-\frac{1}{2(1 - \varrho^2)}\{(\frac{x - \mu_X}{\sigma_X})^2 - 2\varrho\frac{x - \mu_X}{\sigma_X}\frac{y - \mu_Y}{\sigma_Y} + (\frac{y - \mu_Y}{\sigma_Y})^2\}\}$$

27. ทอดลูกเต๋า 300 ครั้งปรากฏผลดังนี้

หน้าที่หงาย	1	2	3	4	5	6
ความถี่	43	49	56	45	66	41

ท่านคิดว่าข้อมูลเหล่านี้ยืนยันได้หรือไม่ว่าลูกเต๋าไม่ถ่วง

28. จากการตรวจสอบลูกสุกร 64 ตัว ที่เกิดจากการผสมพันธุ์พบว่าเป็นสุกรเผือก 34 ตัว ดำ 10 ตัว และขาว 20 ตัว ขณะที่ทฤษฎีทางพันธุศาสตร์ยืนยันว่าอัตราส่วนต่าง ๆ จะต้องเป็น 9:3:4 ผลจากการบันทึกข้อมูลครั้งนี้ยืนยันหรือทดสอบความถูกต้องทางทฤษฎีได้หรือไม่

29. เราสามารถจำแนกทารกที่เกิดใหม่ได้เป็น 3 ลักษณะ ตามลักษณะทางกายภาพ และในทางพันธุศาสตร์ถือว่าอัตราส่วนระหว่างลักษณะทั้ง 3 คือ $p^2 : 2p(1 - p) : (1 - p)^2$ จากการบันทึกลักษณะทารก 109 รายพบว่าเราสามารถจำแนกทารกมีลักษณะเป็นประเภทที่หนึ่ง 10 ราย ประเภทที่สอง 53 ราย และประเภทที่สาม 46 ราย

อยากทราบว่าข้อมูลเหล่านี้สอดคล้องกับแบบจำลองทางพันธุศาสตร์ข้างต้นหรือไม่

30. จำแนกคน 1,000 คนออกเป็น กลุ่มย่อยตามเพศ และคุณภาพของสายตาได้ดังนี้

คุณภาพสายตา	เพศ	
	ชาย	หญิง
ปกติ	442	514
บอดสี	38	6

และแบบจำลองทางพันธุศาสตร์สามารถจำแนกตามเพศและคุณภาพของสายตาได้ดังนี้

$\frac{p}{2}$	$\frac{p^2}{2} + pq$
$\frac{q}{2}$	$\frac{q^2}{2}$

เมื่อ $q = 1 - p$

ก. อยากทราบว่าข้อมูลเหล่านี้สอดคล้องกับแบบจำลองทางพันธุศาสตร์หรือไม่?

ข. อยากทราบว่าคุณภาพของสายตาผันแปรไป ตามเพศหรือไม่?

31. นักวิจัยผู้หนึ่งจำแนกเด็กนักเรียน 1,725 คน ออกเป็นกลุ่มย่อยตามความฉลาดและฐานะทางเศรษฐกิจของผู้ปกครองซึ่งคาดหมายเอาจากการแต่งกาย ปรากฏข้อมูลดังนี้

การแต่งกายของผู้ปกครอง	สติปัญญา		
	โง่ทึบ	ฉลาด	ฉลาดมาก
แต่งกายดีมาก	81	322	233
แต่งกายดี	141	457	153
มอมแมม	127	163	48

จงทดสอบความเป็นอิสระ ณ. ระดับนัยสำคัญ 1%

32. เป็นที่เชื่อว่าเซรุ่มชนิดใหม่จะช่วยป้องกันไข้หวัดได้ จากการทดลองแก่คน 2 กลุ่ม ๆ ละ 500 คน ในระยะเวลา 1 ปี โดยที่กลุ่มแรกได้รับเซรุ่ม กลุ่มที่ 2 ไม่ได้รับ ผลการทดลองเป็นดังนี้

	การเป็นหวัด		
	ไม่เป็น	เป็น 1 ครั้ง	เป็นมากกว่า 1 ครั้ง
ไม่ได้รับเซรุ่ม	252	145	103
ไม่ได้รับเซรุ่ม	224	136	140

อยากทราบว่าเซรุ่มมีผลต่อการป้องกันหรือไม่

33. แบบจำลองทางพันธุศาสตร์แสดงอัตราส่วนบุคคลผู้มีกลุ่มเลือดทั้ง 4 ไว้ดังนี้คือ

$$O:A:B:AB = q^2 : p^2 + 2pq : r^2 + 2pr : 2pr \quad \text{เมื่อ } p + q + r = 1$$

จากการบันทึกข้อมูลพบบุคคลมีกลุ่มเลือดต่าง ๆ ดังนี้คือ 0, 324; A, 436; B, 132; AB, 58 ท่านเชื่อว่าแบบจำลองพันธุศาสตร์ข้างต้นถูกต้องหรือไม่?

34. จงสร้างตัวทดสอบสำหรับ $r \times c \times k$ Contingency Table

ข้อแนะนำ การสร้าง $r \times c \times k$ CT เป็นเทคนิคที่ขยายไปจาก $r \times c$ CT ความยุ่งยากมีเพียงการคำนวณหา p_{rck} และ X_{rck} เท่านั้น นอกนั้นมิได้ยากเกินวิสัย วิธีที่ทำให้ดูง่ายคือพยายามวาดภาพออกมาในรูป 3 มิติ คือรูปลูกเต๋าซ้อน ๆ กันกว้าง r แถว ยาว c แถวสูง k แถว

35. แกลตันทำการบันทึกประวัติครอบครัวต่าง ๆ ไว้ 78 ครอบครัว โดยจำแนกออกเป็นกลุ่ม ๆ คือ สีตาของเด็ก สีตาของบิดามารดา และสีตาของปู่ย่า ปรากฏข้อมูล $2 \times 2 \times 2$ CT ดังนี้

		ปู่-ย่า			
		จาง		เข้ม	
		บิดา-มารดา			
		จาง	เข้ม	จาง	เข้ม
เด็ก	จาง	1928	552	596	508
	เข้ม	303	395	225	501

อยากทราบว่าสีตาของเด็กผันแปรไปตามสีตาของบิดามารดาและปู่ย่าหรือไม่?

36. จงใช้ 2×2 CT หาการแจกแจงจริงของ MLRT สำหรับขนาดตัวอย่าง $n = 2$

37. $X_1, X_2, \ldots, X_n$ เป็นกลุ่มตัวอย่างจากกลุ่มประชากร $N(\mu, \sigma_0^2)$ ให้ λ เป็น MLRT สำหรับ $H_0 : \mu = \mu_0$ vs $H_1 : \mu = \mu_1 \neq \mu_0$ จงหาการแจกแจงของ $-2 \ln \lambda$ แล้วเปรียบเทียบดูว่า test ที่ได้จาก MLRT ตามปกติกับที่ได้จากทฤษฎี 7.1 คือ $-2 \ln \lambda$ มีคุณภาพต่างกันหรือไม่ เพียงใด

ข้อแนะนำ กำหนดให้ใช้ α เดียวกันแล้วเปรียบเทียบ $\beta(\mu)$ หรือ $\Pi(\mu)$

38. ผลการทดลองเบอร์นูลลิปรากฏดังนี้ อยากทราบว่าท่านจะสามารถยืนยันหรือสรุปได้หรือไม่ว่า $p = \frac{1}{2}$ เมื่อ s = success, f = failure

sfffs, sfsff, sffff, ffsfs, sffff,

sffsf, ffffs, ffsff, sffff, ssfff.

39. ถ้าการทดลองในข้อ 38. เป็นการทดลองทอดลูกเต๋า ให้ s = หน้าเอี่ยว อยากทราบว่า $p = 1/6$ หรือไม่

40. ถ้าการทดลองในข้อ 38 คือการทอดลูกเต๋าคราวละ 2 ลูก ให้ s = ลูกเต๋าหงายหน้าเดียวกัน อยากทราบว่า $p = 6/36$ หรือไม่

41. เมื่อสุ่มตัวอย่างจากกลุ่มประชากรที่มี pdf. ดังนี้คือ $f_X(x) = 1/\lambda \,;\, 0 < x < \lambda$ จงพัฒนา SPRT สำหรับ $H_0 : \lambda = \lambda_0$ vs $H_1 : \lambda = \lambda_1 > \lambda_0$

42. ถ้า $X_1, X_2, \ldots, X_n$ เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบเบอร์นุลลิคือ

$$\Pr(X = 1) = p = 1 - \Pr(X = 0)$$

เพื่อต้องการทดสอบสมมุติฐาน

$$H_0 : p = \frac{1}{4} \text{ vs } H_1 : p = 3/4$$

โดยใช้กติกาการตัดสินใจดังนี้

(1) สุ่มตัวอย่างต่อไปนี้ถ้า $\frac{n}{2} - 2 < \Sigma x_i < \frac{n}{2} + 2$

(2) ยอมรับ H_0 ถ้า $\Sigma x_i \leqslant \frac{n}{2} - 2$

(3) ยอมรับ H_1 ถ้า $\Sigma x_i \geqslant \frac{n}{2} + 2$

อยากทราบว่ากติกาเหล่านี้คือ SPRT หรือไม่

43. เมื่อ X เป็นตัวแปรสุ่มแบบพัวซองที่มีค่าเฉลี่ยเท่ากับ λ และเราต้องการทดสอบสมมุติฐาน

$$H_0 : \lambda = 1 \text{ vs } H_1 : \lambda = 2$$

ก. เมื่อกำหนด $\alpha = \beta = .05$ จงคำนวณหาขนาดตัวอย่าง n (fixed sample size)

ข. จงสร้างตัวทดสอบสำหรับ SPRT

44. สมมุติตัวแปรสุ่ม X มี pdf. ดังนี้คือ $f_X(x) = 2x/\lambda^2 \,;\, 0 < x < \lambda$
เมื่อ λ เป็นพารามิเตอร์ที่ไม่ทราบค่าและต้องการทดสอบสมมุติฐานต่อไปนี้

$$H_0 : \lambda = \lambda_0 = 1 \text{ vs } H_1 : \lambda = \lambda_1 < 1$$

ก. จงสร้างตัวทดสอบ

ข. ถ้าปฏิเสธสมมุติฐานหลักภายใต้ข้อสนเทศ 2 ชุดคือ X_1, X_2 เมื่อ $\max(X_1, X_2) < .7$

(1) จงคำนวณหา α

(2) จงคำนวณหาความน่าจะเป็นที่จะยอมรับสมมุติฐานหลักทั้ง ๆ ที่ตามความเป็นจริงแล้ว $\lambda = .84$

45. ตัวแปรสุ่ม X มีการแจกแจงแบบพัวซองมีพารามิเตอร์เท่ากับ λ, และเราต้องการทดสอบสมมุติฐาน

$$H_0 : \lambda = 2 \text{ vs } H_1 : \lambda < 2$$

ถ้าจำนวนอุบัติการณ์ที่สนใจปรากฏในช่วงเวลา [0, 5] รวม 6 ครั้ง ท่านจะปฏิเสธสมมุติฐานหลักหรือไม่?

46. ทอดเหรียญ 800 ครั้ง หงายหน้าหัว 432 ครั้งอยากทราบว่าเหรียญถ่วงหรือไม่

47. ทอดลูกเต๋า 150 ครั้ง ปรากฏผลดังนี้

หน้าที่ออก	1	2	3	4	5	6
ความถี่	29	19	19	27	26	30

อยากทราบว่าลูกเต๋าถ่วงหรือไม่

48. จากการบันทึกข้อมูลในอดีตเชื่อว่าเครื่องจักรที่ใช้อยู่สามารถผลิตสินค้าได้เป็นอัตราส่วนดังนี้ คือ เสียซ่อมไม่ได้ 5% เสียซ่อมได้ 3% และดี 92% จากการผลิตสินค้า 500 ชิ้น ปรากฏว่า 40 ชิ้นเสียซ่อมไม่ได้ และ 28 ชิ้นเสียซ่อมได้ อยากทราบว่าอัตราเหล่านี้สอดคล้องกับข้อมูลที่คาดหมายไว้หรือไม่

49. จากการบันทึกจำนวนอุบัติเหตุในโรงงานปรากฏผลดังนี้

เดือน	ความถี่	เดือน	ความถี่
มกราคม	18	กรกฎาคม	16
กุมภาพันธ์	6	สิงหาคม	24
มีนาคม	8	กันยายน	13
เมษายน	12	ตุลาคม	9
พฤษภาคม	21	พฤศจิกายน	8
มิถุนายน	7	ธันวาคม	14

จงทดสอบดูว่าจำนวนอุบัติเหตุเปลี่ยนแปลงไปตามฤดูกาลหรือไม่?

50. ในช่วงเวลาเร่งด่วนตลอด 7 วันที่ผ่านมาพบว่าจำนวนอุบัติเหตุบนท้องถนนมีจำนวนทั้งสิ้น 15, 9, 18, 12, 13, 10 และ 14 รายตามลำดับอยากทราบว่าจำนวนอุบัติเหตุเหล่านี้แจกแจงแบบพัวซองหรือไม่?

51. ข้อมูลต่อไปนี้แสดงจำนวนความต้องการสินค้าชนิดหนึ่งต่อวัน จากการบันทึก 1,000 วัน ปรากฏผลดังนี้

ความต้องการต่อวัน	ความถี่
0	626
1	274
2	80
3	15
4	4
5	1

อยากทราบว่าปริมาณความต้องการสอดคล้องกับการแจกแจงแบบพัวซองหรือไม่

52. จากการบันทึกข้อมูลอายุการใช้งานของหลอดอิเลคโทรนิคส์ชนิดเดียวกัน 49 ตัวปรากฏข้อมูลดังนี้

1.2	13.7	38.9	72.4	102.8	151.6	203.0
2.2	15.1	47.9	73.6	108.5	152.6	204.3
4.9	15.2	48.4	76.8	128.7	164.2	229.5
5.0	23.9	49.3	83.8	133.6	166.8	253.1
6.8	24.3	53.2	95.1	144.1	178.6	304.1
7.0	25.1	55.6	97.9	147.6	185.2	341.7
2.1	35.8	62.7	99.9	150.6	187.1	354.4

อยากทราบว่าข้อมูลเหล่านี้มาจากกลุ่มประชากรแบบเอกโพเนนเชียลหรือไม่

53. จำแนกคน 1,000 คนออกเป็นกลุ่มย่อยตามลักษณะนิสัยการสูบบุหรี่และการเล่นการพนัน ปรากฏผลดังนี้

	สูบบุหรี่	ไม่สูบบุหรี่
เล่นการพนัน	120	30
ไม่เล่นการพนัน	479	371

อยากทราบนิสัย 2 อย่างนี้เกี่ยวข้องกันหรือไม่

54. จากการบันทึกข้อมูล สีผม และสีตาของคน 6,800 คนปรากฏผลดังนี้

สีตา	สีผม			
	น้ำตาลอ่อน	น้ำตาลแก่	ดำ	แดง
ฟ้า	1,768	807	189	47
ฟ้าเขียว	946	1,387	746	43
น้ำตาล	115	438	288	16

อยากทราบสีผมและสีตามีส่วนเกี่ยวข้องกันหรือไม่

55. จากการทดลองล้างฟิล์ม 2 วิธี คือวิธี ก. และ ข. ผลการล้างฟิล์มปรากฏดังนี้

วิธีล้างฟิล์ม	ผลการล้างฟิล์ม		
	ปกติ	สีซีด	สีเข้มเกินไป
ก.	115	15	10
ข.	120	12	8

อยากทราบว่าผลการล้างฟิล์มขึ้นอยู่กับวิธีการล้างฟิล์มหรือไม่?

56. ใน 2 × 2 CT เมื่อจำนวนความถี่ใน cell ต่าง ๆ คือ a, b, c และ d ตามลำดับ จงแสดงให้เห็นว่า

$$\chi_c^2 = \frac{(a+b+c+d)(ad-bc)^2}{(a+b)(c+d)(b+d)(a+c)}$$

57. จากข้อมูลต่อไปนี้

x	0	1	2	3	4	5	6	7	8
ความถี่	2	4	10	15	19	12	8	7	1

อยากทราบว่าข้อมูลเหล่านี้น่าจะสอดคล้องกับกลุ่มประชากรใด

แนะนำ ให้ลองสร้างฮิสโตแกรมดูก่อนแล้ว เทียบดูว่าลักษณะของฮิสโตแกรมหรือเส้นโค้งที่ลากเชื่อมยอดกราฟแท่งสามารถสงเคราะห์เข้าสู่ pdf. ใด

58. จำนวนอุบัติเหตุทางรถยนต์ต่อสัปดาห์ปรากฏดังนี้ 12, 8, 20, 2, 14, 10, 15, 6, 9, 4 สมมุติว่าจำนวนอุบัติเหตุมีการแจกแจงแบบพัวซอง จงทดสอบดูว่าความถี่เหล่านี้สอดคล้องกับดัชนีวัดการกระจายพัวซองหรือไม่

59. นำปลากระป๋อง 5 ยี่ห้อมาตรวจสอบดูคุณภาพ โดยนำมายี่ห้อละ 2 โหล ผลการตรวจสอบปรากฏว่าแต่ละยี่ห้อมีปลากระป๋องที่คุณภาพไม่สูงพอเป็นจำนวนชนิดละ 4, 10, 6, 2, 8 กระป๋องตามลำดับ อยากทราบว่าปลากระป๋องทั้ง 5 ยี่ห้อมีคุณภาพทัดเทียมกันหรือไม่?

60. Hardy-Weinberg Formular เสนอว่าจำนวนลูกแมลงวันที่เกิดจากการผสมข้ามพันธุ์จะมีอัตราส่วนของลักษณะต่าง ๆ ดังนี้คือ $q^2 : 2pq : p^2$ เมื่อ $p + q = 1$ จากการทดลองครั้งหนึ่งพบว่า จำนวนลูกแมลงวันตามลักษณะทั้งสามมีดังนี้คือ 42, 52, 22 อยากทราบว่า ข้อมูลเหล่านี้จะสอดคล้องกับสูตรข้างบนหรือไม่ กำหนดให้

ก. $p = .5$

ข. $\hat{q} = \dfrac{n_1 + n_2/2}{n_1 + n_2 + n_3}$ เมื่อ n_1, n_2, n_3 คือจำนวน Observe

61. จากโจทย์ข้อ 60. จงพิสูจน์ว่า $\hat{q}$ คือ MLE. ของ q

62. ให้ $\alpha = .2, \beta = .2$ และ $H_0 : p = .5$ vs $H_1 : p = .4$
จงสร้างตัวทดสอบ SPRT และวาดกราฟแสดงเขตการตัดสินใจ

63. ให้ $\alpha = \beta = .1$ และ $H_0 : \sigma = 8$ vs $H_1 : \sigma = 10$ เมื่อ $X \sim N(0, \sigma^2)$
จงสร้าง SPRT

64. เมื่อตัวแปรสุ่ม X มี pdf. เป็น $f_X(x) = \theta e^{-\theta x}$; $x \geqslant 0$
กำหนดให้ $\alpha = .1, \beta = .2$ จงสร้าง SPRT

65. เมื่อตัวแปรสุ่ม X มี pdf เป็น $f(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$; $x = 0, 1, 2, \ldots$ จงสร้าง SPRT สำหรับ

$$H_0 : \lambda = \lambda_0 \text{ vs } H_1 : \lambda = \lambda_1 > \lambda_0$$

และเมื่อกำหนดให้ $\alpha = \beta = .10$ จงทดสอบสมมุติฐาน

$$H_0 : \lambda = 2 \text{ vs } H_1 : \lambda = 3$$

66. เมื่อ $X \sim b(n, p)$ และ $P \sim BT(a, b)$ จงหาการแจกแจงผสม

67. เมื่อ $X \sim$ Poisson (λ) และ $\lambda \sim G(\lambda, r)$ จงหาการแจกแจงผสม

68. เมื่อ $X \sim g(\lambda, r)$ และ $r \sim G(\alpha, \beta)$ จงหาการแจกแจงผสม

ข้อแนะนำ โจทย์ข้อ 66-68 เป็นเรื่องของการแจกแจงผสมซึ่งกล่าวไว้แล้วในบทที่ 4 ตอน 4.19 ในที่นี้จะนำมาให้ทำเป็นแบบฝึกหัดเพราะเห็นว่าเป็นเรื่องที่เกี่ยวข้องกันกับ pdf และ Bayes Solution ก่อนทำขอให้นักศึกษาย้อนไปอ่านรายละเอียดในตอนนั้นให้เข้าใจเสียก่อน

69. จากการเปรียบเทียบวิธีสอน 2 วิธีแก่กลุ่มนักเรียน 2 กลุ่มที่มีระดับสติปัญญาทัดเทียมกัน ผลสัมฤทธิ์ทางการศึกษาปรากฏดังนี้

เกรด	A	B	C	D	F	รวม
กลุ่มที่ 1	15	25	32	17	11	100
กลุ่มที่ 2	9	8	29	28	16	100

อยากทราบว่าวิธีสอนทั้งสองวิธีมีผลเสมอกันหรือไม่?

70. จากการบันทึกข้อมูลจำนวนบุตรของครอบครัวที่สมรสแล้วอย่างน้อย 10 ปี แต่ไม่เกิน 15 ปี ปรากฏข้อมูลจำแนกตามจำนวนบุตรและระดับรายได้ของครอบครัวดังนี้

ระดับรายได้	จำนวนบุตร			
	0	1	2	2^+
$6,000^-$	11	24	49	44
6,000-12,000	9	28	39	31
$12,000^+$	6	15	16	15

อยากทราบว่าการมีบุตรเกี่ยวข้องกับระดับรายได้ของครอบครัวหรือไม่