การทดสอบสมมติฐานทางสถิติสองด้านเมื่อประชากรมีการแจกแจงแบบปัวซอง

รุจิเรข ดีเสียง*

1. คำนำ

การอนุมานทางสถิติ (Statistical Inference) ที่สำคัญลักษณะหนึ่งคือ การทดสอบสมมติฐาน (Testing of Hypothesis) ซึ่งหมายถึงการที่ "One wishes to decide whether or not hypothesis that has been formulated is correct" (Lehmann, E.L., 1986, p.68) กล่าวอีกนัยหนึ่งคือการทดสอบสมมติฐานทางสถิตินี้เป็นการตัดสินใจปฏิเสธ (reject) หรือไม่ปฏิเสธ (not reject หรือ accept) สมมติฐานที่ตั้งขึ้นซึ่งในที่นี้จะใช้สัญญลักษณ์แทนด้วย H และ K โดยที่ H หมายถึง null hypothesis และ K หมายถึง alternative hypothesis

กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงเป็นแบบ P_θ ซึ่งเป็นสมาชิกของ Class $\mathscr{P} = \{P_\theta : \theta \in \Omega\}$ ให้ H และ K แทน subclass ของ $\mathscr{P}$ โดยที่ $H \cap K = \phi$ และ $H \cup K = \mathscr{P}$ และให้ $\Omega_H = \{\theta : \theta \in H\}, \Omega_K = \{\theta : \theta \in K\}$ โดยที่ $\Omega_H \cup \Omega_K = \Omega$

ในการทดสอบสมมติฐาน เราจะพยายามหาแบบทดสอบ (test) ซึ่งเป็นกฎเกณฑ์ (rule) หรือวิธีการ (procedure) ที่ใช้ในการตัดสินใจปฏิเสธหรือไม่ปฏิเสธสมมติฐานที่ตั้งขึ้น โดยพิจารณาจากค่าสถิติเพื่อการทดสอบ (test statistic)

2. ขอบเขตของการศึกษา

สำหรับการทดสอบสมมติฐานทางสถิติในบางครั้ง เราไม่สามารถหาแบบทดสอบ (test) ที่เราต้องการได้ ทั้งนี้เนื่องจากการทดสอบสมมติฐานทางสถิติขึ้นอยู่กับสิ่งต่าง ๆ ที่สำคัญ ดังนี้

1. สมมติฐาน H และ K ที่ต้องการทดสอบ ซึ่งหมายถึงลักษณะของสมมติฐานและพารามิเตอร์ที่ต้องการทดสอบ
2. การแจกแจงของตัวแปรสุ่มที่เกี่ยวข้อง
3. ขนาดตัวอย่างและขนาดของการทดสอบ (sample size and size of the test)

* อาจารย์ประจำภาควิชาสถิติประยุกต์ คณะวิทยาศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้า ลาดกระบัง, นักศึกษาปริญญาเอก สาขาสถิติ คณะสถิติประยุกต์ สถาบันบัณฑิตพัฒนบริหารศาสตร์

2.1 สมมติฐานที่สนใจ

สมมติฐานที่สนใจศึกษาในครั้งนี้เป็นสมมติฐานสองด้าน (two-sided hypotheses) 3 แบบ ดังนี้

แบบที่ 1 $H_1 : \theta \le \theta_1$ หรือ $\theta \ge \theta_2$ เทียบกับ $K_1 : \theta_1 < \theta < \theta_2$

แบบที่ 2 $H_2 : \theta_1 \le \theta \le \theta_2$ เทียบกับ $K_2 : \theta < \theta_1$ หรือ $\theta > \theta_2$

แบบที่ 3 $H_3 : \theta \le \theta_0$ เทียบกับ $K_3 : \theta \ne \theta_0$

แบบทดสอบที่ได้จะเป็นแบบทดสอบสองด้าน (two-sided tests) ซึ่งจะมีรูปแบบแตกต่างกันขึ้นกับลักษณะของสมมติฐานที่ต้องการทดสอบ

2.2 การแจกแจงของตัวแปรสุ่มที่เกี่ยวข้อง

กำหนดให้ $X_1,\ldots, X_n$ เป็นตัวอย่างสุ่ม (random sample) จากประชากรที่มีการแจกแจงแบบปัวซองที่มีพารามิเตอร์ θ ซึ่งมีฟังก์ชันความน่าจะเป็นดังนี้

$$P_\theta(x) = \frac{e^{-n\theta}\, \theta^{\sum_{i=1}^{n} x_i}}{\Pi x_i\,!} \qquad ; x_i = 0, 1, 2,\ldots$$

$$= e^{-n\theta}\, e^{(\ell n\theta)\sum_{i=1}^{n} x_i} \frac{1}{\prod_{i=1}^{n} x_i\,!}$$

$$= c(\theta)\, e^{\eta(\theta)T(x)}\, h(x)$$

จะได้ว่า $P_\theta(x)$ เป็นสมาชิกของแฟมมิลี่แบบเอกซ์โปเนนเชียล (exponential family) ที่มี $c(\theta) = e^{-n\theta}$, $\eta(\theta) = \ell n\theta$ ซึ่งเป็นฟังก์ชันเพิ่มของ θ, $T(x) = \sum_{i=1}^{n} x_i$ และ $h(x) = \dfrac{1}{\prod_{i=1}^{n} x_i\,!}$

2.3 ขนาดตัวอย่างและขนาดของการทดสอบ

สำหรับการศึกษาในครั้งนี้ จะพิจารณาการทดสอบสมมติฐานเมื่อขนาดตัวอย่างมีขนาด เท่ากับ 20, 30, 40 และ 50 และขนาดของการทดสอบเป็น 0.01 และ 0.05

3. ผลการศึกษา

เนื่องจาก $P_\theta(x)$ เป็นสมาชิกของแฟมมิลี่แบบเอกซ์โปเนนเชียล และ $\eta(\theta) = \ell n\theta$ ซึ่งเป็นฟังก์ชั่นเพิ่มของ θ และสถิติเพื่อการทดสอบคือ $\sum_{i=1}^{n} X_i$ จะได้แบบทดสอบดังนี้

3.1 การทดสอบสมมติฐานแบบที่ 1

แบบทดสอบที่มีกำลังสูงที่สุดเสมอขนาด α (UMP test of sized α) ที่ใช้ในการทดสอบ $H_1 : \theta \le \theta_1$ หรือ $\theta \ge \theta_1$ เทียบกับ $K_1 : \theta_1 < \theta < \theta_2$ (ดู Lehmann, E.L., 1986, Theorem 6, p.101-102) ได้แก่

$$\phi(x_1,\ldots,x_n) = \begin{cases} 1 & \text{ถ้า} \quad c_1 < \sum_{i=1}^{n} x_i < c_2 \\ \gamma_1 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = c_1 \\ \gamma_2 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = c_2 \\ 0 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i < c_1 \text{ หรือ } \sum_{i=1}^{n} x_i > c_2 \end{cases}$$

โดยที่ c_1, c_2 เป็นค่าคงที่ และ γ_1, γ_2 เป็นความน่าจะเป็นที่ทำให้

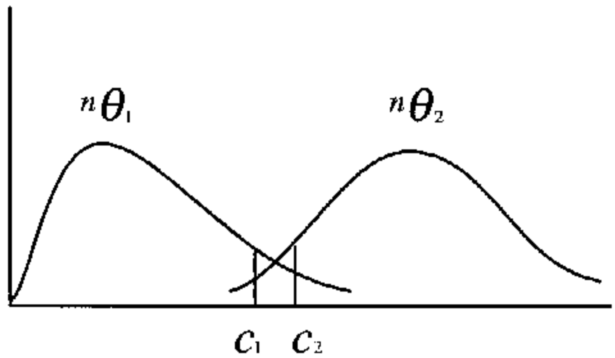
$$E_{\theta_1}\phi(x_1,\ldots,x_n) = E_{\theta_2}\phi(x_1,\ldots,x_n) = \alpha$$

เพื่อให้สามารถหาแบบทดสอบที่ต้องการ เมื่อกำหนดขนาดตัวอย่างและขนาดการทดสอบได้ จำเป็นจะต้องกำหนดค่าพารามิเตอร์ที่ต้องการทดสอบให้เหมาะสมด้วย โดยพิจารณาจากลักษณะแบบทดสอบ (test) ซึ่งขึ้นอยู่กับสมมติฐานที่ต้องการทดสอบ จากการพิจารณาแบบทดสอบที่ได้พบว่าจะปฏิเสธสมมติฐาน ถ้าค่าสถิติทดสอบอยู่ระหว่างค่าคงที่สองค่า (c_1, c_2) เมื่อขนาดของการทดสอบเป็น α กล่าวคือ

$$P\left[c_1 < \sum_{i=1}^{n} X_i < c_2 \mid \theta = \theta_1\right] \le \alpha$$

และ $$P\left[c_1 < \sum_{i=1}^{n} X_i < c_2 \mid \theta = \theta_2\right] \le \alpha$$

ถ้าพิจารณาจากกราฟความน่าจะเป็นควรมีลักษณะดังนี้



นั่นคือค่าพารามิเตอร์ θ_1 และ θ_2 ควรมีค่าต่างกัน และจากการ trial จะกำหนดให้ พารามิเตอร์ที่ต้องการทดสอบ θ_1 และ θ_2 มีค่าเป็น 0.2 และ 0.75 ตามลำดับ ซึ่งแบบทดสอบที่ได้เมื่อกำหนดขนาดตัวอย่างและขนาดของการทดสอบ แสดงไว้ดังตารางที่ 1(ดูตัวอย่างการคำนวณในภาคผนวก)

ตารางที่ 1 ความน่าจะเป็น (γ_1, γ_2) และค่าคงที่ (c_1, c_2) สำหรับการทดสอบสมมติฐาน $H_1 : \theta \le 0.2$ หรือ $\theta \ge 0.75$ เทียบกับ $K_1 : 0.2 < \theta < 0.75$

ขนาดของการทดสอบ (α)	ขนาดตัวอย่าง (n)	พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$	$\gamma_2 = P(\sum_{i=1}^{n} X_i = c_1)$	$\gamma_2 = P(\sum_{i=1}^{n} X_i = c_2)$	ค่าคงที่ c_1	c_2
0.01	20	4, 15	-	-	-	-
	30	6, 22.5	0.595029	0.630031	12	13
	40	8, 30	0.300595	0.771825	15	18
	50	10, 37.5	0.400789	0.543968	18	24
0.05	20	4, 15	0.139937	0.899526	7	9
	30	6, 22.5	0.200967	0.555582	10	15
	40	8, 30	0.537884	0.786063	13	21
	50	10, 37.5	-	-	-	-

ข้อสังเกต ไม่มีแบบทดสอบที่ต้องการสำหรับการทดสอบสมมติฐาน $H_1 : \theta \le 0.2$ หรือ $\theta \ge 0.75$ เทียบกับ $K_1 : 0.2 < \theta < 0.75$ เมื่อขนาดตัวอย่างเป็น 20 และขนาดของการทดสอบเป็น 0.01 กับขนาดตัวอย่างเป็น 50 และขนาดของการทดสอบเป็น 0.05

3.2 การทดสอบสมมติฐานแบบที่ 2

แบบทดสอบที่ไม่เอนเอียงที่มีกำลังสูงที่สุดเสมอขนาด α (UMP Unbiased test of sized α) ที่ใช้ในการทดสอบ $H_2 : \theta_1 \le \theta \le \theta_2$ เทียบกับ $K_2 : \theta < \theta_1$ หรือ $\theta > \theta_2$ (ดู Lehmann, E.L., 1986, p.135) ได้แก่

$$\phi(x_1,\ldots,x_n) = \begin{cases} 1 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i < c_1 \text{ หรือ } \sum_{i=1}^{n} x_i > c_2 \\ \gamma_1 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = c_1 \\ \gamma_2 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = c_2 \\ 0 & \text{ถ้า} \quad c_1 < \sum_{i=1}^{n} x_i < c_2 \end{cases}$$

โดยที่ c_1, c_2 เป็นค่าคงที่ และ γ_1, γ_2 เป็นความน่าจะเป็น

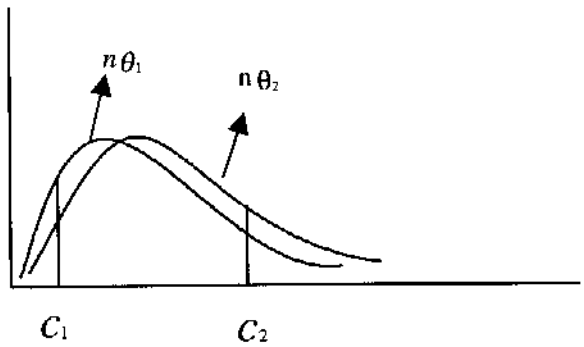
ที่ทำให้ $E_{\theta_1}\phi(X_1,\ldots, X_n) = E_{\theta_2}\phi(X_1,\ldots, X_n) = \alpha$

จากการพิจารณาแบบทดสอบที่ได้พบว่าจะปฏิเสธสมมติฐาน ถ้าค่าสถิติมีค่าน้อยกว่าค่าคงที่ c_1 หรือมีค่ามากกว่าค่าคงที่ c_2 เมื่อขนาดของการทดสอบเป็น α กล่าวคือ

$$P[\sum_{i=1}^{n} X_i < c_1 \mid \theta = \theta_1] + P[\sum_{i=1}^{n} X_i > c_2 \mid \theta = \theta_1] \le \alpha$$

และ $$P[\sum_{i=1}^{n} X_i < c_1 \mid \theta = \theta_2] + P[\sum_{i=1}^{n} X_i > c_2 \mid \theta = \theta_2] \le \alpha$$

ถ้าพิจารณาจากกราฟความน่าจะเป็นควรมีลักษณะดังนี้



นั่นคือค่าพารามิเตอร์ θ_1 และ θ_2 ควรมีค่าต่างกันน้อย และจากการ trial จะกำหนดให้พารามิเตอร์ที่ต้องการทดสอบ θ_1 และ θ_2 มีค่าเป็น 0.25 และ 0.3 ตามลำดับ ซึ่งแบบทดสอบที่ได้เมื่อกำหนดขนาดตัวอย่างและขนาดของการทดสอบแสดงไว้ดังตารางที่ 2 (ดูตัวอย่างการคำนวณในภาคผนวก)

ตารางที่ 2 ความน่าจะเป็น (γ_1, γ_2) และค่าคงที่ (c_1, c_2) สำหรับการทดสอบสมมติฐาน $H_2 : 0.25 \le \theta \le 0.3$ เทียบกับ $K_2 : \theta < 0.25$ หรือ $\theta > 0.3$

ขนาดของการทดสอบ (α)	ขนาดตัวอย่าง (n)	พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$	$\gamma_2 = P(\sum_{i=1}^{n} X_i = c_1)$	$\gamma_2 = P(\sum_{i=1}^{n} X_i = c_2)$	ค่าคงที่	
					c_1	c_2
0.01	20	5, 6	0.054221	0.594662	1	13
	30	7.5, 9	0.259232	0.37997	2	17
	40	10, 12	0.813588	0.331916	3	21
	50	12.5, 15	0.404659	0.426229	5	25
0.05	20	5, 6	-	-	-	-
	30	7.5, 9	0.541992	0.704644	3	15
	40	10, 12	0.35337	0.020086	5	18
	50	12.5, 15	-	-	-	-

ข้อสังเกต ไม่มีแบบทดสอบที่ต้องการสำหรับการทดสอบสมมติฐาน $H_2 : 0.25 \le \theta \le 0.3$ เทียบกับ $K_2 : \theta < 0.25$ หรือ $\theta > 0.3$ เมื่อขนาดของการทดสอบเป็น 0.05 และขนาดตัวอย่างเป็น 20 และ 50

3.3 การทดสอบสมมติฐานแบบที่ 3

แบบทดสอบที่ไม่เอนเอียงที่มีกำลังสูงที่สุดเสมอขนาด α (UMP Unbiased test of sized α) ที่ใช้ในการทดสอบ $H_3 : \theta = \theta_0$ เทียบกับ $K_3 : \theta \ne \theta_0$ (ดู Lehmann, E.L., 1998, p.135-136) ได้แก่

$$\phi(x_1,\ldots,x_n) = \begin{cases} 1 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i < c_1 \text{ หรือ } \sum_{i=1}^{n} x_i > c_2 \\ \gamma_1 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = c_1 \\ \gamma_2 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = c_2 \\ 0 & \text{ถ้า} \quad c_1 < \sum_{i=1}^{n} x_i < c_2 \end{cases}$$

โดยที่ c_1, c_2 เป็นค่าคงที่ และ γ_1, γ_2 เป็นความน่าจะเป็นที่ทำให้

$$E_{\theta_0}\phi(X_1,\ldots, X_n) = \alpha$$

และ $$E_{\theta_0}[(\sum_{i=1}^{n} X_i)\phi(X_1,\ldots, X_n) = \alpha E_{\theta_0}(\sum_{i=1}^{n} X_i)$$

เมื่อประชากรมีการแจกแจงแบบปัวซองที่มีพารามิเตอร์ θ จะได้ $E_{\theta_0}(\sum_{i=1}^{n} X_i) = n\theta_0$

นั่นคือ $E_{\theta_0}[(\sum_{i=1}^{n} X_i)\phi(X_1,\ldots, X_n)] = \alpha n\theta_0$

สำหรับสมมติฐานในลักษณะนี้ส่วนใหญ่สามารถหาแบบทดสอบที่ต้องการได้ โดยในที่นี้จะกำหนดพารามิเตอร์ที่ต้องการทดสอบมีค่าเป็น $\frac{1}{2}$ ซึ่งแบบทดสอบที่ได้เมื่อกำหนดขนาดตัวอย่างและขนาดของการทดสอบ แสดงไว้ดังตารางที่ 3 (ดูตัวอย่างการคำนวณในภาคผนวก)

ตารางที่ 3 ความน่าจะเป็น (γ_1, γ_2) และค่าคงที่ (c_1, c_2) สำหรับการทดสอบสมมติฐาน $H_3 : \theta = \frac{1}{2}$ เทียบกับ $K_3 : \theta \neq \frac{1}{2}$

ขนาดของการทดสอบ (α)	ขนาดตัวอย่าง (n)	พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$	$\gamma_1 = P(\sum_{i=1}^{n} X_i = c_1)$	$\gamma_2 = P(\sum_{i=1}^{n} X_i = c_2)$	ค่าคงที่	
					c_1	c_2
0.01	20	10	0.13569	0.721283	3	9
	30	15	0.552083	0.431034	6	26
	40	20	0.092953	0.880434	10	33
	50	25	0.770396	0.416387	13	39
0.05	20	10	0.809523	0.7890622	4	17
	30	15	0.493127	0.220551	8	23
	40	20	0.332553	0.075764	12	29
	50	25	0.301528	0.046168	16	35

4. สรุปผลการศึกษา

ในการทดสอบสมมติฐานแบบสองด้าน ถึงแม้ว่าโดยทั่วไปจะสามารถหาแบบทดสอบที่มีกำลังสูงที่สุดเสมอหรือแบบทดสอบที่ไม่เอนเอียงที่มีกำลังสูงที่สุดเสมอ แต่ในทางปฏิบัติอาจหาแบบทดสอบไม่ได้สำหรับประชากรที่มีการแจกแจงบางอย่างหรือขนาดตัวอย่างและ/หรือขนาดของการทดสอบบางค่าก็ได้ ดังผลของการศึกษาข้างต้น

ภาคผนวก

ตัวอย่างการคำนวณค่าในตารางที่ 1

กำหนดให้ ขนาดตัวอย่าง (n) = 20 ขนาดของการทดสอบ (α) = 0.05 และพารามิเตอร์ที่ต้องการทดสอบคือ $\theta_1 = 0.2$ และ $\theta_2 = 0.75$ จะได้ว่าสถิติเพื่อการทดสอบ ($\sum_{i=1}^{n} x_i$) มีการแจกแจงแบบปัวซองที่มีพารามิเตอร์เป็น $n\theta$ กล่าวคือ

เมื่อ $n = 20, \theta_1 = 0.2$ พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$ คือ $20(0.2) = 4$ และเมื่อ $n = 20, \theta_2 = 0.75$ พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$ คือ $20(0.75) = 15$ เราสามารถหาค่า $c_1, c_2, \gamma_1, \gamma_2$ จากเงื่อนไขดังนี้

$$E_{\theta_1}\phi(X_1,\ldots,X_n) = \alpha$$

นั่นคือ $$\sum_{\sum_{i=1}^{n} x_i = c_1+1}^{c_2-1} P_4(x) + \gamma_1 P_4(c_1) + \gamma_2 P_4(c_2) = 0.05 \qquad \text{..........①}$$

และ $$E_{\theta_2}\phi(X_1,\ldots,X_n) = \alpha$$

นั่นคือ $$\sum_{\sum_{i=1}^{n} x_i = c_1+1}^{c_2-1} P_{15}(x) + \gamma_1 P_{15}(c_1) + \gamma_2 P_{15}(c_2) = 0.05 \qquad \text{..........②}$$

จากการคำนวณค่าความน่าจะเป็นและการ trial พบว่า $c_1 = 7$ และ $c_2 = 9$ โดยที่

$P[7 < \sum_{i=1}^{n} X_i < 9 \mid \theta_1 = 0.2] = 0.0298$, $P[\sum_{i=1}^{n} X_i = 7 \mid \theta_1 = 0.2] = 0.0595$,

$P[\sum_{i=1}^{n} X_i = 9 \mid \theta_1 = 0.2] = 0.0132$,

$P[7 < \sum_{i=1}^{n} X_i < 9 \mid \theta_2 = 0.75] = 0.0194$, $P[\sum_{i=1}^{n} X_i = 7 \mid \theta_2 = 0.75] = 0.0104$,

$P[\sum_{i=1}^{n} X_i = 9 \mid \theta_2 = 0.75] = 0.0324$

แทนค่าลงในสมการ ① และ ②

$$0.0298 + 0.0595\gamma_1 + 0.0132\gamma_2 = 0.05$$

$$0.0194 + 0.0104\gamma_1 + 0.0324\gamma_2 = 0.05$$

จะได้ $\gamma_1 = 0.139937$ และ $\gamma_2 = 0.899529$

นั่นคือ

$$\phi(x_1,\dots,x_{20}) = \begin{cases} 1 & \text{ถ้า} \quad 7 < \sum_{i=1}^{n} x_i < 9 \\ 0.139937 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = 7 \\ 0.899529 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i = 9 \\ 0 & \text{ถ้า} \quad \sum_{i=1}^{n} x_i < 7 \text{ หรือ } \sum_{i=1}^{n} x_i = 9 \end{cases}$$

ในที่นี้

$$E[\phi(X_1,\dots,X_{20})] = P[7 < \sum_{i=1}^{n} X_i < 9] + 0.139937P[\sum_{i=1}^{n} X_i = 7] + 0.899529P[\sum_{i=1}^{n} X_i = 9]$$

$$E[\phi(X_1,\dots,X_{20})|\theta_1 = 0.2] = 0.0298 + 0.139937(0.0595) + 0.899529(0.0132) = 0.05$$

$$E[\phi(X_1,\dots,X_{20})|\theta_2 = 0.75] = 0.0194 + 0.139937(0.0104) + 0.899529(0.0324) = 0.05$$

ตัวอย่างการคำนวณค่าในตารางที่ 2

กำหนดให้ขนาดตัวอย่าง $(n) = 20$ ขนาดของการทดสอบ $(\alpha) = 0.01$ และพารามิเตอร์ที่ต้องการทดสอบคือ $\theta_1 = 0.25, \theta_2 = 0.3$ จะได้ว่าสถิติเพื่อการทดสอบ $(\sum_{i=1}^{n} X_i)$ มีการแจกแจงแบบปัวซองที่มีพารามิเตอร์เป็น $n\theta$ กล่าวคือ

เมื่อ $n = 20, \theta_1 = 0.25$ พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$ คือ $20(0.25) = 5$

และเมื่อ $n = 20, \theta_2 = 0.3$ พารามิเตอร์ของ $\sum_{i=1}^{n} X_i$ คือ $20(0.3) = 6$

เราสามารถหาค่า $c_1, c_2, \gamma_1, \gamma_2$ จากเงื่อนไขดังนี้

$$E_{\theta_1}\phi(X_1,\dots,X_n) = \alpha$$

นั่นคือ

$$\sum_{\sum_{i=1}^{n} x_i = 0}^{c_1 - 1} P_5(x) + \sum_{\sum_{i=1}^{n} x_i = c_2 + 1}^{\infty} P_5(x) + \gamma_1 P_5(c_1) + \gamma_2 P_5(c_2) = 0.01 \quad \text{..........③}$$

และ $$E_{\theta_2}\phi(X_1,\dots,X_n) = \alpha$$

นั่นคือ

$$\sum_{\sum_{i=1}^{n} x_i=0}^{c_1-1} P_6(x) + \sum_{\sum_{i=1}^{n} x_i=c_2+1}^{\infty} P_6(x) + \gamma_1 P_6(c_1) + \gamma_2 P_6(c_2) = 0.01 \quad ....④$$

จากการคำนวณค่าความน่าจะเป็นและการ trial พบว่า $c_1 = 1$ และ $c_2 = 13$ โดยที่

$$P[\sum_{i=1}^{n} X_i < 1 \mid \theta_1 = 0.25] = 0.0067,\ P[\sum_{i=1}^{n} X_i = 1 \mid \theta_1 = 0.25] = 0.0337$$

$$P[\sum_{i=1}^{n} X_i > 13 \mid \theta_1 = 0.25] = 0.0007,\ P[\sum_{i=1}^{n} X_i = 13 \mid \theta_1 = 0.25] = 0.0013$$

$$P[\sum_{i=1}^{n} X_i < 1 \mid \theta_2 = 0.3] = 0.0025,\ P[\sum_{i=1}^{n} X_i = 1 \mid \theta_2 = 0.3] = 0.0149$$

$$P[\sum_{i=1}^{n} X_i > 13 \mid \theta_2 = 0.3] = 0.0036,\ P[\sum_{i=1}^{n} X_i = 13 \mid \theta_2 = 0.3] = 0.0052$$

แทนค่าในสมการ ③ และ ④

$$0.0067 + 0.0007 + 0.0337\gamma_1 + 0.0013\gamma_2 = 0.01$$

$$0.0025 + 0.0036 + 0.0149\gamma_1 + 0.0052\gamma_2 = 0.01$$

จะได้ $\gamma_1 = 0.054211$ และ $\gamma_2 = 0.594662$

นั่นคือ

$$\phi(x_1,\ldots,x_{20}) = \begin{cases} 1 & \text{ถ้า } \sum_{i=1}^{n} x_i < 1 \text{ หรือ } \sum_{i=1}^{n} x_i > 13 \\ 0.054211 & \text{ถ้า } \sum_{i=1}^{n} x_i = 1 \\ 0.594662 & \text{ถ้า } \sum_{i=1}^{n} x_i = 13 \\ 0 & \text{ถ้า } 1 < \sum_{i=1}^{n} x_i < 13 \end{cases}$$

ในที่นี้

$$E[\phi(X_1,\ldots,X_{20}) = P[\sum_{i=1}^{n} X_i < 1] + 0.054211 P[\sum_{i=1}^{n} X_i = 1] + 0.594662 P[\sum_{i=1}^{n} X_i = 13] + P[\sum_{i=1}^{n} X_i > 13]$$

$$E[\phi(X_1,\ldots,X_{20}) \mid \theta_1 = 0.25] = 0.0067 + 0.054211(0.0337) + 0.594662(0.0013) + 0.0007$$

$$= 0.01$$

$$E[\phi(X_1,\ldots, X_{20}) \mid \theta_2 = 0.3] = 0.0025 + 0.054211(0.0149) + 0.594662(0.0052) + 0.0036$$

$$= 0.01$$

ตัวอย่างการคำนวณค่าในตารางที่ 3

กำหนดให้ขนาดตัวอย่าง $(n) = 20$ ขนาดของการทดสอบ $(\alpha) = 0.01$ และพารามิเตอร์ที่ต้องการทดสอบคือ $\theta_0 = \frac{1}{2}$ จะได้ว่าสถิติเพื่อการทดสอบ $(\sum_{i=1}^{n} X_i)$ มีการแจกแจงแบบปัวซองที่มีพารามิเตอร์เป็น $n\theta = 20(\frac{1}{2}) = 10$ เราสามารถหาค่า $c_1, c_2, \gamma_1, \gamma_2$ จากเงื่อนไขดังนี้

$$E_{\theta_0}\phi(X_1,\ldots, X_n) = \alpha$$

นั่นคือ

$$\sum_{\sum_{i=1}^{n} X_i=0}^{c_1-1} P_{10}(x) + \sum_{\sum_{i=1}^{n} X_i=c_2+1}^{\infty} P_{10}(x) + \gamma_1 P_{10}(c_1) + \gamma_2 P_{10}(c_2) = 0.01 \ldots\ldots\ldots ⑤$$

และ

$$E_{\theta_0}[(\sum_{i=1}^{n} X_i)\phi(X_1,\ldots, X_n) = E_{\theta_0}[\sum_{i=1}^{n} X_i] \cdot \alpha$$

นั่นคือ

$$\sum_{\sum_{i=1}^{n} x_i=0}^{c_1-1} (\sum_{i=1}^{n} x_i)P_{10}(x) + \sum_{\sum_{i=1}^{n} x_i=c_2+1}^{\infty} (\sum_{i=1}^{n} x_i)P_{10}(x) + \gamma_1 c_1 P_{10}(c_1) + \gamma_2 c_2 P_{10}(c_2)$$

$$= 20(\frac{1}{2})(0.01) \ldots\ldots\ldots ⑥$$

จากการคำนวณค่าความน่าจะเป็นและการ trial พบว่า $c_1 = 3$ และ $c_2 = 9$ โดยที่

$P[\sum_{i=1}^{n} X_i < 3 \mid \theta = \frac{1}{2}] = 0.0028$, $P[\sum_{i=1}^{n} X_i = 3 \mid \theta = \frac{1}{2}] = 0.0076$

$P[\sum_{i=1}^{n} X_i > 19 \mid \theta = \frac{1}{2}] = 0.0035$, $P[\sum_{i=1}^{n} X_i = 19 \mid \theta = \frac{1}{2}] = 0.0037$

$P[\sum_{i=1}^{n} X_i = 1 \mid \theta = \frac{1}{2}] = 0.0005$, $P[\sum_{i=1}^{n} X_i = 2 \mid \theta = \frac{1}{2}] = 0.0023$

$P[\sum_{i=1}^{n} X_i = 20 \mid \theta = \frac{1}{2}] = 0.0019$, $P[\sum_{i=1}^{n} X_i = 21 \mid \theta = \frac{1}{2}] = 0.0009$

$P[\sum_{i=1}^{n} X_i = 22 \mid \theta = \frac{1}{2}] = 0.0004$, $P[\sum_{i=1}^{n} X_i = 23 \mid \theta = \frac{1}{2}] = 0.0002$

$$P[\sum_{i=1}^{n} X_i = 24 \mid \theta = \frac{1}{2}] = 0.0001$$

แทนค่าในสมการ ⑤ และ ⑥

$$0.0028 + 0.0035 + 0.0076\gamma_1 + 0.0037\gamma_2 = 0.01$$

$$1(0.0005) + 2(0.0023) + 3\gamma_1(0.0076) + 19\gamma_2(0.0037) + 20(0.0019) + 21(0.0009) + 22(0.0004) + 23(0.0002) + 24(0.0001) = 0.1$$

หรือ $0.0462 + 0.0227\gamma_1 + 0.0709\gamma_2 = 0.1$

จะได้ $\gamma_1 = 0.13569$ และ $\gamma_2 = 0.721283$

นั่นคือ

$$\phi(x_1,\ldots,x_{20}) = \begin{cases} 1 & \text{ถ้า } \sum_{i=1}^{n} x_i < 3 \text{ หรือ } \sum_{i=1}^{n} x_i > 19 \\ 0.13569 & \text{ถ้า } \sum_{i=1}^{n} x_i = 3 \\ 0.721283 & \text{ถ้า } \sum_{i=1}^{n} x_i = 19 \\ 0 & \text{ถ้า } 3 < \sum_{i=1}^{n} x_i < 19 \end{cases}$$

ในที่นี้

$$E[\phi(X_1,\ldots,X_n) \mid \theta = \frac{1}{2}] = 0.0028 + 0.0035 + 0.0076(0.13569) + 0.0037(0.721283) = 0.01$$

$$E[(\sum_{i=1}^{n} X_i)\phi(X_1,\ldots,X_{20}) \mid \theta = \frac{1}{2}] = 0.0462 + 0.0227(0.13569) + 0.7092(0.721283) = 0.1$$

บรรณานุกรม

Lehmann, E.L. Testing Statistical Hypotheses. 2nd ed. NewYork : John Wiley and Sons, Inc., 1986.

Dudewicz, Edward J. Modern Mathematical Statistics. New York : John Wiley and Sons, Inc., 1988.

Robert V. Hogg, ALLEN T. Craig. Introduction to Mathematical Statistics. 5th ed. America : Prentice-Hall International, Inc., 1995.

ประชุม สุวัตถี. ทฤษฎีการอนุมานเชิงสถิติ. สถาบันบัณฑิตพัฒนบริหารศาสตร์ร่วมกับสำนักพิมพ์โอเดียนสโตร์, 2527.