บทที่ 4
การสร้างตัวประมาณค่าด้วยวิธีต่าง ๆ

ตัวประมาณค่าแบบวิธี Maximum Likelihood

4.1 บทนำ Maximum Likehood เป็นวิธีการทางหนึ่งที่ใช้กันมากในทางสถิติ เพราะว่าวิธีการสร้างตัวประมาณค่าด้วยวิธีนี้จะได้ตัวประมาณค่าที่มีคุณสมบัติที่ดีหลายอย่าง ยกตัวอย่างเช่น

- เป็นตัวประมาณค่าที่มีประสิทธิภาพที่สุด
- เป็นตัวประมาณค่าที่มีคุณสมบัติของความแนบนัยต่อพารามิเตอร์ เป็นต้น

แนวคิดเบื้องต้นเกี่ยวกับการสร้างตัวประมาณค่าโดยวิธีนี้ก็คือ พยายามเลือกค่าของพารามิเตอร์ของประชากรที่เป็นไปได้ โดยให้ค่าที่เลือกนี้เหมาะกับค่าสังเกตที่ได้จากกลุ่มตัวอย่าง เช่น ในการทดลองโยนเหรียญ 10 ครั้ง ปรากฏว่าขึ้นหน้าหัว 4 ครั้ง เราจะประมาณค่า π (สัดส่วนของความสำเร็จในการโยนเหรียญแล้วได้หน้าหัว) ของประชากรได้เท่าไร เป็นที่ทราบกันอยู่ว่าสัดส่วนจริงของความสำเร็จในการที่จะโยนเหรียญแล้วได้หน้าหัวคือ $0<\pi<1$ ถ้าทดลองเลือกดูว่า $\pi = 0.1$ จะเป็นค่าที่สมเหตุสมผลหรือไม่ โดยการศึกษาว่าความน่าจะเป็น $\pi = .1$ จะมีค่าน่าเชื่อถือแค่ใด ถ้าให้ S เป็นตัวแปรแทนจำนวนครั้งที่โยนเหรียญ n ครั้งแล้วได้หัว ดังนั้นฟังก์ชั่นของความน่าจะเป็น S

$$f(s;\pi) = \binom{n}{s} \pi^s(1-\pi)^{n-s} \quad ; \quad s = 0,1,2,\ldots,$$

กำหนดให้ $n = 10,\ s = 4$

จงหาความน่าจะเป็นของ $\pi; L(\pi) = \binom{10}{4} \pi^4(1-\pi)^6; 0<\pi<1$

ซึ่งถ้าหากว่าเราลองเลือกค่า $\pi = 0.1$

ในการโยนเหรียญ 10 ครั้ง แล้วผลลัพธ์จะเป็นหัว 4 ครั้งจะมีความน่าจะเป็น

$$L(\pi=0.1) = \binom{10}{4}(.1)^4(.9)^6$$

$$= .011$$

ซึ่งในกรณีนี้ความน่าจะเป็นที่เหตุการดังกล่าวจะมีค่าต่ำมากถ้าทดลองต่อไปโดยเลือก $\pi = .8$

$$L(\pi = .8) = \binom{10}{4} .8^4 .2^6$$

$$= .006$$

ซึ่งมีค่ายิ่งน้อยกว่าเมื่อเลือก $\pi = .1$

ในการเลือกว่าจะแทนค่า π ด้วยค่าใดในทางสถิติมีหลักการว่าควรจะเลือกค่าที่ทำให้ความน่าจะเป็นมีค่าสูงสุด ปัญหาก็คือในช่วง 0 ถึง 1 ค่า π เท่าไรจึงจะทำให้มีความเหมาะสมกับข้อมูลที่เราได้รับจากการทดลอง

ขอให้พิจารณาจากการทดลองแทนค่า π ต่าง ๆ กัน ดังนี้

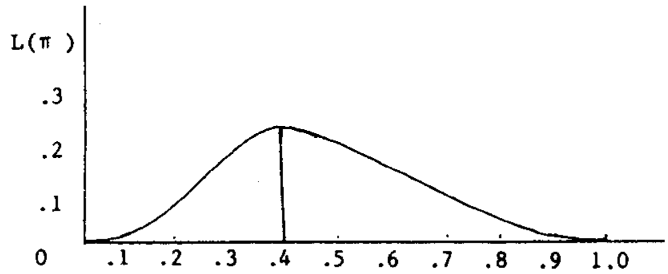
กำหนดข้อมูลที่ได้จากการทดลองโยนเหรียญ 10 ครั้งจะได้ผลลัพธ์เป็นหน้าหัว 4 ครั้ง

π	$L(\pi) = \binom{10}{4} \pi^4(1-\pi)^{10-4}$	
0	0	
.1	.011	
.2	.088	
.3	.200	
.4	.251	→ Max
.5	.205	
.6	.111	
.7	.037	
.8	.006	
1	0	

จากตารางที่แสดงไว้คงจะเป็นแนวทางให้เห็นว่าควรจะเลือกค่า $\pi = .4$ ซึ่งจะได้ค่าประมาณที่มีโอกาสถูกต้องสูงสุด ในเมื่อข้อมูลที่ได้จากตัวอย่างมีลักษณะดังกล่าว ถ้าจะพิจารณาค่า $\pi = .4$ ที่เราเลือกก็จะพบว่าค่า .4 นี้ที่แท้จริงก็คือ สัดส่วนของจำนวนครั้งที่จะโยนได้หัวเมื่อคิดมาจากการทดลอง (กลุ่มตัวอย่าง $n = 10$ นั่นเอง)

สรุปได้ว่า ค่าที่ใช้เป็นตัวแทน π ก็คือ p (จากตัวอย่าง)

จากตัวอย่างที่กล่าวสรุปออกมาเป็นกราฟจะได้ผลดังนี้



รูปแสดงฟังก์ชั่นของความน่าจะเป็น ที่จะเลือก π ณ. ที่ค่าต่าง ๆ กัน

ถ้าจะเขียนในรูปทั่ว ๆ ไป ในกรณีของการทดลองที่มีลักษณะเป็นการแจกแจงทวินาม โดยที่กำหนดให้มีการทดลองตัวอย่างขนาด n โดยที่ S คือจำนวนครั้งของความน่าจะเป็นที่จะได้รับผลสำเร็จ s จากการทดลอง n ครั้ง

$$p(s/\pi) \;=\; \binom{n}{s} \pi^{s}(1-\pi)^{n-s} \quad ; \quad s = 0,1,2,\ldots$$

ในกรณีที่เราเจาะจงกำหนดว่า n และ s เป็นค่าที่ได้รับจากการทดลองแล้ว เราก็จะสามารถเขียนฟังก์ชั่นความน่าจะเป็นของ S ($p(s/\pi)$) ออกมาได้เป็นฟังก์ชั่นความน่าจะเป็นของตัวแปร ได้คือ $L(\pi) = \binom{n}{s} \pi^{s} (1-\pi)^{n-s}$

การจะเลือกค่าของ π เพื่อให้ความน่าจะเป็นสูงที่สุดนี่ก็คือการหาค่า π ที่ทำให้ฟังก์ชั่น L เป็นค่าสูงที่สุด เราทำได้โดยการทดลองแทนค่า π ต่าง ๆ กันลงไปในฟังก์ชั่นซึ่งจะเป็นวิธีที่การที่ค่อนข้างเสียเวลาและยุ่งยาก ยิ่งในกรณีที่เป็นฟังก์ชั่นที่ซับซ้อน จะยิ่งทำให้การคำนวณหาลำบากมากยิ่งขึ้นอีก

แนวทางหนึ่งที่จะพิจารณาว่าฟังก์ชั่น (L) จะมีค่าสูงสุดเมื่อ π มีค่าเท่าใด ก็คือการใช้หลักการทางคณิตศาสตร์เข้าช่วย โดยใช้หลักการที่ว่าฟังก์ชั่นใดเมื่อเรานำไปสร้างกราฟ ค่าที่จุดสูงสุดหรือจุดต่ำสุด (ดูรูปที่ 1 ประกอบ) ในกรณีของเราจะสนใจจุดสูงสุดของฟังก์ชั่น ซึ่ง ณ. ที่จุดนี้ความลาดชันของกราฟจะมีค่าเป็นศูนย์ การจะหาความชันที่จุดนี้ได้ก็โดยการหาอนุพันธ์ของฟังก์ชั่นเมื่อเทียบกับพารามิเตอร์ (เช่น π เป็นต้น)

สรุปได้ดังนี้ $\frac{\partial L(\pi)}{\partial \pi} = 0; \hat{\pi} = f(n,s)$

ดังนั้นค่า $\hat{\pi}$ ที่ได้คือค่าที่ทำให้ฟังก์ชั่น $\hat{\pi}$ มีค่าสูงสุดตามที่ต้องการ

ตัวอย่าง

$$L(\pi) = \pi^s(1-\pi)^{n-s} \quad 1/$$

$$\frac{\partial L(\pi)}{\partial \pi} = (\hat{\pi}^s(n-s)(1-\hat{\pi})^{n-s-1}(-1) + s\hat{\pi}^{s-1}(1-\hat{\pi})^{n-s})$$

$$= 0$$

$$-\hat{\pi}(n-s) + s(1-\hat{\pi}) = 0$$

$$-n\hat{\pi} + s = 0$$

$$\hat{\pi} = s/n$$

หรือที่ใช้กันทั่วไปคือ $\hat{\pi} = p$

4.2 Principle of Maximum Likelihood Estimator

หลักการทั่วไปในการหาตัวประมาณค่าที่ทำให้ฟังก์ชั่นร่วมมีค่าสูงสุด (Maximum Likelihood Estimato)

กำหนดให้ $X_1,\ldots,X_n$ เป็นกลุ่มตัวอย่างจากตัวแปรเชิงสุ่ม X ที่ให้ค่าดังนี้ $x_1,\ldots,x_n$ เราจะสร้างฟังก์ชั่นร่วม (Likelihood Function) ได้ดังนี้

$$L(x_1,\ldots,x_n;\theta) = f(x_1;\theta)\ldots f(x_n;\theta)$$

1/ จะเห็นว่า $L(\pi)$ เป็นฟังก์ชั่นที่ไม่ต่อเนื่องแต่เราพอจะใช้วิธีการหาอนุพันธ์เข้าช่วยได้

ในกรณีที่ X เป็นตัวแปรชนิดที่ไม่ต่อเนื่องจะได้ว่า

$$L(x_1,\ldots,x_n;\theta) \quad = \quad P(X_1 = x_1, X_2 = x_2,\ldots,X_n = x_n)$$

หลังจากสุ่มตัวอย่าง $X_1,\ldots,X_n$ แล้ว ถ้าค่าที่ได้รับเป็น $x_1,\ldots,x_n$ ปัญหาต่อไปจะพบก็คือ $L(x_1,x_2,\ldots,x_n;\theta)$ จะเป็นฟังก์ชั่นของพารามิเตอร์ θ และหน้าที่ที่ต้องทำก็คือจะต้องเลือก θ อย่างไร จึงจะทำให้ฟังก์ชั่น L มีค่าสูงที่สุด

นั่นคือ

$$L(x_1,x_2,\ldots,x_n;\theta_k) > L(x_1,x_2,\ldots,x_n;\theta_j) \ ; \ j = 1,2,\ldots \quad ;\ldots\ldots\ldots(1)$$
$$j \neq k$$

ถ้า θ_k มีคุณสมบัติตาม (1) θ_k คือ Maximum Likelihood Estimator ของ θ (โดยอาศัยประโยชน์ข้อมูลจาก $X_1,\ldots,X_n$)

คำจำกัดความ การประมาณค่าโดยใช้วิธีการทำให้ฟังก์ชั่นร่วมมีค่าสูงที่สุดของพารามิเตอร์ ถ้าหากได้ว่า $\hat{\theta}$ เป็นฟังก์ชั่นของกลุ่มตัวอย่าง $X_1,\ldots,X_n$ เป็นค่าหนึ่งของ θ ที่ทำให้ฟังก์ชั่นร่วม $L(x_1,\ldots,x_n;\theta)$ มีค่าสูงที่สุดด้วย

เราเรียก $\hat{\theta}$ นี้ว่าตัว Maximum Likelihood Estimator ของ θ

ขั้นตอนในการหา

1. สร้าง $L(x_1,\ldots,x_n\theta) = f(x_1;\theta)\ldots f(x_n;\theta)$
2. $\dfrac{\partial L(x_1,\ldots,x_n;\theta)}{\partial\theta} = 0$

   หาค่า $\hat{\theta}_{mle}$ ออกมาในฟังก์ชั่นของ $(X_1,\ldots,X_n)$

3. เป็นกรณีพิเศษของขั้นที่ 2 ก็คือ ในกรณีที่ $L(x_1,\ldots,x_n;\theta)$ เป็นฟังก์ชั่นที่ยุ่งยากในการดิฟเฟอร์เรนชิเอททันที เช่นอาจจะอยู่ในรูป ฟังก์ชั่นของการแจกแจงแบบปกติ ฟังก์ชั่นของการแจกแจงแบบเอ็กโปเนนเชียล เป็นต้น

1/ ฟังก์ชั่นก็คือฟังก์ชั่นร่วมของตัวแปรเชิงสุ่มที่ได้จากการทดลอง (Joint pdf.) นั่นเอง

เราอาจหาค่าของ mle โดยวิธี ดิฟเฟอร์เรนชิเอท $\ln L(x_1,\ldots,x_n;\theta)$ เทียบกับ θ แทน ซึ่งจะได้ผลลัพธ์เช่นเดียวกันกับ $\frac{\partial L}{\partial \theta}(x_1,\ldots,x_n;\theta)$

พิสูจน์ได้ว่า ถ้าดิฟเฟอร์เรนชิเอท $\ln L(x_1,\ldots,x_n;\theta)$ เทียบกับ θ จะได้ว่าค่าของ $\frac{\partial L}{\partial \theta}(x_1,\ldots,x_n;\theta)$ จะมีค่าสูงสุดด้วย

$$\frac{\partial \ln}{\partial \theta} L(x_1, x_2,\ldots, x_n, \theta) = 0$$

$$(1/L(x_1,\ldots,x_n,\theta))\frac{\partial \ln}{\partial \theta_k} L(x_1,\ldots,x_n;\theta_k) = 0$$

ดังนั้น $\frac{\partial \ln}{\partial \theta} L(x_1, x_2,\ldots, x_n, \theta) = 0$ ด้วย

ดังนั้นในกรณีที่ X เป็นตัวแปรเชิงสุ่มชนิดที่มีการแจกแจงแบบพัวซอง ที่มีพารามิเตอร์คือ θ ดังนั้นจะได้ว่า

$$f(x;\theta) = \frac{e^{-x}x^{\theta}}{\theta!} \; ; \; x=0,1,2,\ldots$$

หรือในกรณีที่ X เป็นตัวแปรเชิงสุ่ม ที่มีการแจกแจงแบบปกติที่มีค่าเฉลี่ย และความแปรปรวนเป็น θ_1 และ θ_2 ตามลำดับ

$$f(x;\theta_1,\theta_2) = \frac{1}{\sqrt{2\pi\theta_2}} e^{-(x-\theta_1)^2/2\theta_2}; \; -\infty<\theta_1<\infty, \theta_2>0$$

หรือในกรณีที่ X เป็นตัวแปรเชิงสุ่มที่มีการแจกแจงแบบทวินาม มีพารามิเตอร์เป็น p

---

$L(x_1,\ldots,x_n;\theta)$ ถือว่าเป็นฟังก์ชั่นของ $L(\theta/x_1,\ldots,x_n)$ เพราะ $x_1,\ldots,x_n$ เป็นค่าที่ทราบแล้วจากการทดลอง เปรียบเหมือนกับการใช้ผลของการทดลองมาช่วยสร้างค่าประมาณของตัวพารามิเตอร์นั่นเอง โดยปกติข้อมูลที่ได้มามักจะมีข้อมูลที่อธิบายลักษณะของพารามิเตอร์แฝงอยู่ด้วย เราจึงนำประโยชน์ข้อนี้มาใช้

$$f(x;p) = p^x(1-\theta)^{1-x}\ ;\ x=0,1$$

ตัวอย่างเหล่านี้ล้วนแต่จะทำให้การหาค่าสูงสุดของฟังก์ชั่นเป็นไปด้วยความลำบาก ดังนั้นถ้าเจอปัญหาเช่นนี้ก็ให้เลี่ยงไปใช้วิธีการ $\frac{\partial}{\partial\theta} \ln L(x_1,x_2,\ldots,\theta)$ แทนวิธีการใช้ $\frac{\partial}{\partial\theta} L(x_1,x_2,\ldots,x_n;\theta)$ ซึ่งก็จะให้คำตอบเช่นเดียวกัน

ข้อสังเกตุ *การที่เราสามารถนำฟังก์ชั่นร่วมมาดิฟเฟอร์เรนชิเอทเทียบกับค่าพารามิเตอร์ แสดงว่าค่าพารามิเตอร์นั้นเป็นตัวแปรซึ่งขัดกับความรู้เดิมที่เราทราบว่า ค่าพารามิเตอร์ของประชากรเป็นค่าคงที่ สาเหตุที่เป็นเช่นนี้ก็เพราะว่าในกรณีที่ตัวพารามิเตอร์เป็นค่าที่ไม่สามารถทราบได้เราจึงต้องพยายามหาค่าของมันที่จะเป็นไปได้ในทางทฤษฎี ซึ่งค่าของมันมีมากมายตัวอย่างเช่นค่าคาดหมายของตัวแปร ที่มีการกระจายเป็นปกติ (μ, σ^2) ค่า $-\infty<\mu<\infty$ มีค่ามากมาย เราจึงต้องใช้วิธีการทางคณิตศาสตร์มาช่วยในการหาค่าประมาณ โดยวิธีก็จะต้องทำให้ฟังก์ชั่นร่วมมีค่าสูงสุดนั่นเอง นั่นหมายถึงว่าตัวพารามิเตอร์จะกลายเป็นตัวแปรนั่นเอง

หรือในกรณีที่มีตัวพารามิเตอร์หลายตัว เช่น $\theta_1,\ldots,\theta_k$ การที่จะหา mle ของพารามิเตอร์แต่ละตัวก็มีหลักการคล้ายกับกรณีของพารามิเตอร์ตัวเดียว กล่าวคือ

ให้หาค่า $\theta_1,\ldots,\theta_k$ ที่ทำให้ฟังก์ชั่น $L(x_1,\ldots,x_n;\ \theta_1,\ldots,\theta_k)$ มีค่าสูงสุดเมื่อเทียบกับพารามิเตอร์แต่ละตัว

นั่นคือ
$$\frac{\partial L}{\partial\theta_1}(x_1,\ldots,x_n;\ \theta_1,\ldots,\theta_k) = 0$$
$$\frac{\partial L}{\partial\theta_2}(x_1,\ldots,x_n;\ \theta_1,\ldots,\theta_k) = 0$$
$$\vdots \qquad\qquad \vdots$$
$$\frac{\partial L}{\partial\theta_k}(x_1,\ldots,x_n;\ \theta_1,\ldots,\theta_k) = 0$$

แล้วหาค่า $\hat{\theta}_{1mle},\ldots,\hat{\theta}_{kmle}$ ในฟังก์ชั่นของค่าสังเกตุ $X_1,\ldots,X_n$ หรืออาจจะหาได้จาก

$$\frac{\partial \ln}{\partial \theta_1} \quad L(x_1,..,x_n;\ \theta_1,\ldots,\theta_k) = 0$$

$$\frac{\partial \ln}{\partial \theta_2} \quad L(x_1,\ldots,x_n;\ \theta_1\ldots,\ \theta_k) = 0$$

$$\vdots \qquad\qquad\qquad\qquad \vdots$$

$$\frac{\partial \ln}{\partial \theta_k} \quad L(x_1,\ldots,x_n;\ \theta_k) = 0$$

ก็จะได้ผลเช่นเดียวกันตามที่พิสูจน์ไว้แล้ว ในกรณีของพารามิเตอร์ตัวเดียว

4.3 การหาค่าประมาณโดยให้ฟังก์ชั่นร่วมมีค่าสูงสุดในประชากรที่มีการแจกแจงแบบปกติ (MLE for Normal Population)

จากการสุ่มตัวอย่างขนาด 3 คือ X_1,X_2,X_3 โดยที่ประชากรมีการแจกแจงแบบปกติ มีค่าคาดหมายและความแปรปรวนเป็น μ และ σ^2 ให้หาค่าประมาณของ μ โดยวิธีการของ mle

$$p(x_1/\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp-(x_1-\mu)^2/(1/2\sigma^2)\ . \ ; \quad -\infty<\mu<\infty$$

ในทำนองเดียวกันจะได้ว่าความน่าจะเป็นของ X_1,X_2 ก็คือ

$$p(x_2/\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp-(x_2-\mu)^2/(1/2\sigma^2) \quad ; \quad -\infty<\mu<\infty$$

$$p(x_3/\mu) = \frac{1}{\sqrt{2\pi}\sigma} \exp-(x_3-\mu)^2/(1/2\sigma^2) \quad ; \quad -\infty<\mu<\infty$$

โดยที่ X_1,X_2 และ X_3 เป็นอิสระต่อกัน ดังนั้นฟังก์ชั่นร่วมของ X_1,X_2 และ X_3 ก็คือ

$$p(x_1,x_2,x_3;\mu) = \prod_{i=1}^{3} \left(\frac{1}{\sqrt{2\pi}\sigma} \exp-(x_1,\mu)^2/(1/2\sigma^2)\right)$$

หรือจะเขียนอีกแบบหนึ่งได้ว่า

$L(x_1,x_2,x_3;\mu)$ หรือ $L(\mu)$

$$= \prod_{i=1}^{3} \frac{1}{\sqrt{2\pi\sigma}} \exp-(x_i-\mu)^2/(1/2\sigma^2))$$

$$L(\mu) = [\frac{1}{\sqrt{2\pi\sigma}}]^3 \exp-(\sum_{i=1}^{3}(x_i-\mu)^2/(1/2\sigma^2))$$

$$\ln L(\mu) = 3\ln(1\sqrt{2\pi\sigma}) - (1/2\sigma^2)\sum_{i=1}^{3}(x_i-\mu)^2)$$

$$\frac{\partial \ln}{\partial\mu} L(\mu) = 0 - \frac{1}{2\sigma^2} 2 \sum_{i=1}^{3}(x_i-\mu)(-1) = 0$$

$$\Rightarrow \sum_{i=1}^{3}(x_i-\hat{\mu}) = 0$$

$$\hat{\mu} = \overline{X}$$

$$\Rightarrow \hat{\mu}_{mle} = \overline{X}$$

การประมาณค่าของพารามิเตอร์โดยวิธีการทำให้ฟังก์ชั่นร่วมมีค่าสูงสุดในประชากรที่มีการแจกแจงแบบปกติในสมการถดถอย

ในที่นี้จะขอยกตัวอย่างสมการถดถอยแบบธรรมดา (มีตัวแปรอิสระตัวเดียว)

$$Y_i = \alpha+\beta X_i+e_i$$

โดยที่เรามีสมมุติฐานว่าตัว e_i (ค่าความผิดพลาด) มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเป็น 0 และมีค่าความแปรปรวนเป็น σ^2

หมายเหตุ $E(Y_i) = E(\alpha+\beta X_i+e_i) = \alpha+\beta X_i+E(e_i)$

$= \alpha+\beta X_i$; $E(e_i) = 0$ และ X_i เป็นค่าคงที่

ถ้าสุ่มค่าสังเกตมา n ค่า ฟังก์ชั่นร่วมของ $Y_1,\ldots,Y_n$ (ความน่าจะเป็นที่จะเกิด $y_1,\ldots,y_n$) คือ $p(y_1,\ldots,y_n;\alpha,\beta)$

$$\text{โดยที่} \Rightarrow p(y_1;\alpha,\beta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - (y_1 - (\alpha + \beta x_1))^2/(1/2\sigma^2))$$

$$\vdots \qquad \vdots$$

$$p(y_n;\alpha,\beta) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - (y_n - (\alpha + \beta x_n))^2/(1/2\sigma^2))$$

เนื่องจากว่า $Y_1, Y_2, \ldots, Y_n$ มีความเป็นอิสระต่อกัน

$$\text{ดังนั้น } p(y_1,\ldots,y_n;\alpha,\beta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp - (y_i - (\alpha + \beta x_i)^2/(1/2\sigma^2)$$

$$\text{หรือ } L(\alpha, \beta) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp - (\sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2/(1/2\sigma^2)$$

การที่จะหาค่าประมาณของ α,β โดยวิธี mle ก็คือการทำให้ฟังก์ชั่นร่วมมีค่าสูงที่สุดนั่นเอง พิจารณาดูว่าฟังก์ชั่นร่วม $L(\alpha,\beta)$ จะมีค่าสูงที่สุดเมื่อไร (เพื่อที่จะเลือก α,β ที่ต้องการได้)

จะได้ว่า $L(\alpha,\beta)$ จะมีค่าสูงที่สุด เมื่อ $-\sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2/(1/2\sigma^2)$ มีค่าที่ต่ำที่สุดนั่นเอง

เป็นที่ทราบกันอยู่ว่าในการวิเคราะห์สมการถดถอยนั้นค่าประมาณของ α,β ได้จากการหาค่าที่ทำให้ฟังก์ชั่น $\sum_{i=1}^{n} (y_i - \alpha - \beta x_i)^2$ มีค่าน้อยที่สุด

สรุป วิธีการหาค่า α,β โดยวิธีของ mle กับการหาค่าโดยวิธีใช้สมการถดถอยก็คือวิธีเดียวกันนั่นเอง

การหาค่าประมาณโดยวิธีการทำให้ฟังก์ชั่นร่วมมีค่าสูงสุด (mle) สำหรับประชากรที่มีการแจกแจงแบบใด ๆ

เรื่องที่กล่าวมาแล้วข้างต้นแสดงให้เห็นวิธีการหาค่าประมาณพารามิเตอร์ของประชากรที่มีการแจกแจงแบบปกติ แต่ถ้าประชากรมีการแจกแจงแบบอื่น ๆ เราก็ยังคงจะ

ใช้หลักการเช่นเดียวกันนั่นเอง คือ $\frac{\partial}{\partial\theta}$ $L(\theta)$ หรือ $\frac{\partial \ln}{\partial\theta}$ $L(\theta)$ ในที่นี้จะยกตัวอย่างของประชากรที่มีการแจกแจงแบบต่าง ๆ ให้เป็นแนวทาง

ตัวอย่างที่ 1 กำหนดให้ตัวแปร X มีการแจกแจงแบบ Exponential โดยที่มี pdf เป็น

$$f(x;\lambda) = \lambda e^{-\lambda x} \quad ; \quad x \geqslant 0,\ \lambda > 0$$

กำหนดให้ใช้ขนาดตัวอย่าง n (แต่ละค่าสังเกตเป็นอิสระต่อกัน) ให้หา mle ของ λ

$$L(\lambda) = \prod_{i=1}^{n} (\lambda e^{-\lambda x_i})$$

$$= \lambda^n e^{-\lambda \Sigma x_i}$$

$$\ln L(\lambda) = n\ln\lambda - \lambda\Sigma x_i$$

$$\frac{\partial \ln}{\partial\lambda} L(\lambda) = n/\hat{\lambda} - \Sigma x_i = 0$$

$$= n/\Sigma x_i$$

$$\Rightarrow \quad \hat{\lambda}_{mle} = 1/\bar{X}$$

ตัวอย่างที่ 2 กำหนดให้ X เป็นตัวแปรสุ่มที่มาจากประชากรที่มีการแจกแจงแบบทวินาม มีพารามิเตอร์ คือ n,p ให้หาค่าประมาณของ p โดยวิธี mle

$$f(x;p) = \binom{n}{x} p^x(1-p)^{n-x} \quad ; \quad x = 0,1,\ldots$$
$$0<p<1$$

$$L(p) = \binom{n}{x} p^x(1-p)^{n-x} \quad ;$$

$$\ln L(p) = \ln\binom{n}{x} + x\ln p + (n-x)\ln(1-p)$$

$$\frac{\partial \ln}{\partial p} L(p) = 0 + x/\hat{p} + (n-x)(-1)/(1-\hat{p})$$
$$= 0$$

$$\Rightarrow \quad \hat{p}_{mle} = x/n$$

ตัวอย่างที่ 3 ให้ $X_1,...,X_n$ เป็นตัวแปรเชิงสุ่มชุดหนึ่งที่สุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ มีพารามิเตอร์ μ และ σ^2 ให้หาค่าประมาณของ μ และ σ^2 โดยวิธีการของ mle (ตัวอย่างนี้เป็นกรณีของหลายตัวแปร)

$$f(x;\mu,\sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp - (x_i - \mu)^2/(1/2\sigma^2)$$

$$L(\mu,\sigma^2) = \sum_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right) \exp - (x_i - \mu)^2/(1/2\sigma^2)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp - \sum_{i=1}^{n} (x_i - \mu)^2/(1/2\sigma^2)$$

$$\ln L(\mu,\sigma^2) = \ln\left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n - \left(\frac{1}{2\sigma^2}\right) \sum_{i=1}^{n} (x_i - \mu)^2$$

$$\frac{\partial \ln L(\mu, \sigma)^2}{\partial \mu} = \frac{1}{\sigma^2} \Sigma(x_i - \hat{\mu}) = 0$$

$$\hat{\mu}_{mle} = \overline{X}$$

$$\frac{\partial}{\partial \sigma^2} \ln L(\mu,\sigma^2) = \frac{1}{2\hat{\sigma}^4} \Sigma(x_i - \mu)^2 - \frac{n}{2\hat{\sigma}^2} = 0$$

$$\hat{\sigma}^2_{mle} = \sum_{i=1}^{n} (x_i - \mu)^2/n$$

ให้พิจารณาค่า $\hat{\sigma}^2_{mle}$ ซึ่งอยู่ในรูปฟังก์ชั่นของ x_i กับ μ เราสามารถที่จะแทนค่า $\hat{\mu}_{mle} = \overline{X}$ ลงไปใน $\hat{\sigma}^2_{mle} = \sum_{i=1}^{n} \frac{(x_i - \bar{x})^2}{n}$ ได้ทั้งนี้เนื่องจาก mle มีคุณสมบัติของ Invariance property กล่าวคือถ้าเราได้ว่า $\hat{\theta}_{mle}$ เป็น mle ของ θ เราจะได้ตามมาว่า ฟังก์ชั่นของ $\hat{\theta}_{mle}$ จะเป็น mle ของฟังก์ชั่นของ θ ด้วยนั่นคือ

เมื่อ $\overline{X}$ เป็น mle ของ $\hat{\mu}$

$\Sigma \frac{(x_i - \bar{x})^2}{n}$ เป็น mle ของ $\Sigma \frac{(x_i - \mu)^2}{n}$ ด้วย

ในการหาค่า mle ของพารามิเตอร์ของประชากรโดยวิธีการดิฟเฟอร์เรนชิเอท ไม่สำเร็จ อันเกิดขึ้นเนื่องจากลักษณะของฟังก์ชั่นความน่าจะเป็นอยู่ในรูปที่ไม่สามารถ แก้ปัญหาได้ด้วยวิธีนี้ จะต้องใช้วิธีการอื่นเข้าช่วย ดังตัวอย่างที่ยกมาแสดง

1. $f(x;\theta) = 1/\theta$; $0<\theta<\infty$, $0<x\leqslant\theta$

กำหนดตัวแปร $X_1,\ldots,X_n$

$$\Rightarrow L(\theta) = 1/\theta^n$$

$$\ln L(\theta) = \ln 1 - n\ln\theta$$

$$\frac{\partial \ln L(\theta)}{\partial\theta} = -n/\theta = 0$$

แก้สมการหาค่า θ ในฟังก์ชั่นของ $X_1,\ldots,X_n$ ไม่ได้โดยวิธีนี้

2. $f(x;\theta) = 1$; $\theta-\frac{1}{2}<x<\theta+\frac{1}{2}$
; $-\infty<\theta<\infty$

กำหนดตัวแปร $X_1,\ldots,X_n$

$$\Rightarrow L(\theta) = 1$$

$$\frac{\partial L}{\partial\theta}(\theta) = 0$$

หา mle ของ θ ไม่สำเร็จโดยวิธีนี้

ในปัญหาที่หา mle โดยวิธีการทางแคลคูลัส (ดิฟเฟอร์เรนชิเอท) ไม่สำเร็จ เราจะใช้ตัวสถิติที่เกี่ยวกับอันดับมาเป็นแนวทางในการหา mle ของพารามิเตอร์

ตัวสถิติที่เกี่ยวข้องกับอันดับก็คือ Order Statistic

ตัวอย่างที่ 4 กำหนดให้ $f(x;\theta) = 1/\theta$; $0<x\leqslant\theta$, $0<\theta<$

ให้ใช้ขนาดตัวอย่าง n $X_1,X_2,\ldots,X_n$

$$L(\theta) = 1/\theta^n \; ; \; 0<x_i\leqslant\theta \, , \, 0<\theta<\infty$$

ถ้าเรียงค่า $X_1,\ldots,X_n$ ที่ได้จากอันดับน้อยที่สุดไปมากที่สุดจะได้ $Y_1\leqslant Y_2\leqslant\ldots\leqslant Y_n$

แต่เนื่องจาก $0<x_i<\theta \rightarrow 0<Y_1\leqslant Y_2\leqslant\ldots\leqslant Y_n\leqslant\theta$

การที่จะเลือกตัวประมาณค่าของ θ ในตัวอย่างนี้ ถ้าจะยึดหลักที่ทำให้ค่า $L(\theta)$ สูงที่สุด จะเห็นว่าเราควรเลือก θ ที่จะทำให้ค่าของ $L(\theta)$ สูงสุดนั่นคือเลือก θ ที่จะทำให้ค่าของ $1/\theta^n$ มีค่าสูงสุด นั่นคือ ต้องเลือก θ ที่มีค่าน้อยที่สุด นั่นคือเลือก $\theta\geqslant Y_n \Rightarrow \theta = Y_n$ $\Rightarrow L(\theta) = 1/Y_n^n$ [1/]

การที่เราไม่เลือก $\theta = Y_1$ ก็เพราะว่าขัดกับ (1)

ดังนั้น $\hat{\theta}_{mle}$ ก็คือ $\mathrm{Max}(X_i)$; $i=1,\ldots,n$

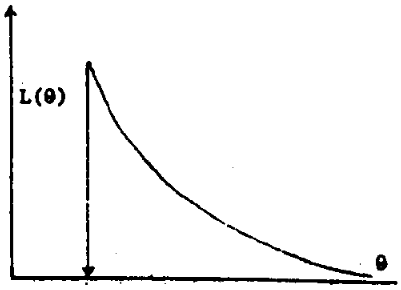
และเมื่อพิจารณาค่าของ Y_n จะพบว่า $E(Y_n) = n\theta/(n+1) \neq \theta$

ดังนั้น $\hat{\theta}_{mle}$ จึงไม่จำเป็นต้องเป็นค่าประมาณ ที่ปราศจากความเอียงเฉ

[1/] พิจารณา $L(\theta) = 1/\theta^n$ จะเห็นว่า โดยวิธีของ ตัวกะประมาณที่ทำให้ $L(\theta)$ มีค่าสูงสุดคือ Y_1 แต่เมื่อพิจารณาสมการที่ (1) จะพบว่า Y_1 มีค่าห่างจากพารามิเตอร์ θ มากที่สุด การใช้ Y_1 เป็นตัวกะประมาณจึงไม่เหมาะสมในประการทั้งปวง ส่วน Y_n แม้ว่าจะไม่ทำให้ $L(\theta)$ มีค่าสูงที่สุด แต่ Y_n กลับมีค่าใกล้ θ มากที่สุด จึงเลือกใช้ Y_n เป็นตัวกะประมาณ

เรื่องนี้เป็นเรื่องที่ต้องใช้วิจารณญาณหลายด้านประกอบกัน ตัวประมาณค่าที่ได้บางครั้งจึงเป็นตัวประมาณค่าที่มีความเอียงเฉได้ (Biased Estimator)

พิจารณาจากรูปประกอบ



ถ้าต้องการปรับปรุง $\hat{\theta}_{mle}$ ให้เป็นค่าประมาณที่ปราศจากความเอียงเฉของ θ จะได้ $\hat{\theta} = \frac{n+1}{n} \mathrm{Max}(X_1,\ldots,X_n)$

ตัวอย่างที่ 5 กำหนดให้ใช้ $X_1,\ldots,X_n$ จากตัวแปรสุ่ม X ที่มีการแจกแจงอย่างสม่ำเสมอในช่วง $(\theta-\frac{1}{2}\leqslant x\leqslant\theta+\frac{1}{2})$; $0<\theta<\infty$ จงหา mle ของ θ

จาก $(f(x;\theta) \;=\; 1 \;;\; \theta-\frac{1}{2}\leqslant x\leqslant\theta+\frac{1}{2})$

$$L(x_1,\ldots,x_n;\theta) \;=\; 1$$

จะเห็นว่า ไม่ว่าจะเลือกค่า θ อย่างไร ค่าสูงสุดของ $L(x_1,\ldots,x_n;\theta)$ ก็ยังคงเป็น 1 พิจารณาจากปัญหานี้จะได้ว่า ถ้านำค่า $X_1,\ldots,X_n$ ที่สังเกตได้มาเรียงกันอยู่ในรูปของ Order Statistics คือ $Y_1\leq Y_2\leq\ldots\leq Y_n$ โดยที่เซทของ $Y_1,\ldots,Y_n$ ก็ยังคงมีข้อจำกัดดังนี้

$$\theta-\frac{1}{2}\leqslant Y_1<Y_2 \ldots < Y_n\leqslant\theta+\frac{1}{2}$$

จะเห็นได้ว่าตัวสถิติที่น่าสนใจของเราในที่นี้ก็คือ

$$\theta-\frac{1}{2}\leqslant Y_1 \quad \text{และ} \quad \theta+\frac{1}{2}\geqslant Y_n$$

$$\theta-\frac{1}{2}\leqslant Y_1 \;=>0\leqslant Y_1+\frac{1}{2}$$

$$\theta+\frac{1}{2}\geqslant Y_n \;=>\theta\geqslant Y_n-\frac{1}{2}$$

$$=> \; Y_n-\frac{1}{2}\leqslant\theta\leqslant Y_1+\frac{1}{2}$$

ดังนั้นเซทของ mle สำหรับ θ ก็คือ ค่าสังเกตใด ๆ ที่อยู่ในช่วง $Y_n - \frac{1}{2} \rightarrow Y_1 + \frac{1}{2}$ ดังภาพ

$$\overset{\theta}{\overline{\hspace{6cm}}}$$

$$Y_n - \frac{1}{2} \qquad\qquad\qquad\qquad\qquad Y_1 + \frac{1}{2}$$

ดังนั้น เซทของ $\hat{\theta}_{mle}$ ของ θ ก็คือ

$$Y_n - \frac{1}{2} + a(Y_1 + \frac{1}{2} - (Y_n - \frac{1}{2})) = Y_n - \frac{1}{2} + a(Y_1 - Y_n + 1) \quad ^{1/} \; ; \; 0<a<1$$

ตัวอย่างที่ 6 กำหนดให้ $X_1,\ldots,X_n$ เป็นตัวอย่างขนาด n ที่สุ่มมาจากประชากรที่มีการแจกแจงแบบยูนิฟรอมในช่วง $(\mu - \sqrt{3}\sigma, \mu + \sqrt{3}\sigma)$ โดยที่ $-\infty<\mu<\infty$; $\sigma>0$ จงหา mle ของ μ และ σ

เนื่องจาก $\quad f(x;\mu,\sigma) = \dfrac{1}{2\sqrt{3}\sigma} \quad ; \quad \mu - \sqrt{3}\sigma < x < \mu + \sqrt{3}\sigma$

$$\Rightarrow L(\mu,\sigma) = \left(\frac{1}{2\sqrt{3}\sigma}\right)^n ; \quad \mu - \sqrt{3}\sigma < x_i < \mu + \sqrt{3}\sigma$$

ต้องเลือก μ และ σ เพื่อให้ (μ,σ) มีค่าสูงสุดภายใต้เงื่อนไขที่กำหนดให้ แต่เนื่องจากเราไม่สามารถดิฟเฟอร์เรนชิเอท L หรือ lnL ได้ จึงต้องหันมาพิจารณาตัวสถิติที่เกี่ยวกับอันดับแทน นั่นคือ

จาก $X_1, X_2, \ldots, X_n \Rightarrow Y_1 \leqslant Y_2 \leqslant \ldots \leqslant Y_n \; ; \; i = 1,2,\ldots,n$

1/ ตัวอย่างเช่น $a = 0 \Rightarrow \hat{\theta}_{mle} = Y_n + \frac{1}{2}$

$a = 1 \Rightarrow \hat{\theta}_{mle} = Y_1 + \frac{1}{2}$

$a = \frac{1}{2} \Rightarrow \hat{\theta}_{mle} = \frac{Y_n + Y_1}{2}$ mid range

เป็นต้น

จากตัวอย่างนี้จะได้ว่า $\hat{\theta}_{mle}$ ไม่จำเป็นต้องมีตัวเดียวอาจจะมีหลายตัวก็ได้

เนื่องจาก $\mu-\sqrt{3}\sigma \leqslant x_i \leqslant \mu+\sqrt{3}\sigma$

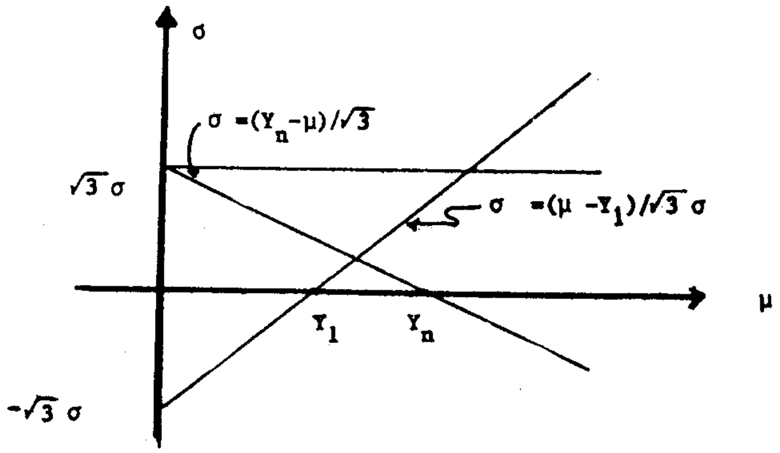
$=>$ $\mu-\sqrt{3}\sigma \leqslant Y_1 \leqslant Y_2 \leqslant \ldots \leqslant Y_n \leqslant \mu+\sqrt{3}\sigma$

ให้เลือก จาก $\mu-\sqrt{3}\sigma \leqslant Y_1$ ให้พิจารณาเฉพาะ $Y_1 = \mu-\sqrt{3}\sigma$ ....(1)

$\mu+\sqrt{3}\sigma \leqslant Y_n$ ให้พิจารณาเฉพาะ $Y_n = \mu+\sqrt{3}\sigma$ ....(2)

แก้สมการ (1) และ (2) $=> \hat{\mu}_{mle} = \dfrac{Y_1+Y_n}{2}$

$=> \hat{\sigma}_{mle} = \dfrac{Y_n-Y_1}{2\sqrt{3}}$



4.4 คุณสมบัติของตัวประมาณค่าที่ได้มาจากวิธีที่ทำให้ฟังก์ชั่นร่วมมีค่าสูงสุด (Properties of A Maximum, Likelihood Estimator)

1. ตัว mle ไม่จำเป็นจะต้องมีคุณสมบัติของการปราศจากอคติต่อตัวพารามิเตอร์ เช่น $X_i \sim N(\mu,\sigma^2)$ จะได้ว่า $\overline{X}$ คือ mle ของ μ เราได้ว่า $E(\overline{X}) = \mu$ ในตัวอย่างที่ 4 ได้ว่า Y_n เป็น mle ของ θ โดยที่ $E(Y_n) = n\theta/(n+1)$ ดังนั้น Y_n เป็นค่าประมาณที่มีความเอียงเฉต่อพารามิเตอร์ ในกรณีที่ค่าประมาณใดที่ได้เป็นค่าประมาณที่มีอคติเราก็สามารถที่จะปรับปรุงให้เป็นค่าประมาณที่ปราศจากอคติได้ (ดูตัวอย่างที่ 4 ประกอบ)

อาจกล่าวได้อีกนัยหนึ่งว่า mle จะมีคุณสมบัติเป็น assymtotically unbiased estimator นั่นก็คือ $\lim_{n\to\infty} \hat{\theta}_{mle} = \theta$

2. mle มีคุณสมบัติเป็นตัวประมาณที่มีความแนบนัยต่อค่าพารามิเตอร์ (Consistency) ตัวอย่างเช่น ในตัวอย่างที่ 4

$$\hat{\theta}_{mle} = Y_n$$

$$E(\hat{\theta}_{mle}) = n\theta/(n+1)$$

$$V(\hat{\theta}_{mle}) = E(Y_n^2) - (E(Y_n))^2$$

$$E(Y_n^2) = \int_0^\theta y_n^2 g_n(y_n)\, dy_n$$

$$= \int_0^\theta y_n^2\, ny_n^{n-1}/\theta^n\, dy_n$$

$$= n\theta^2/(n+2)$$

$$=> V(\hat{\theta}_{mle}) = \frac{n\theta^2}{(n+2)} - \left(\frac{n}{n+1}\theta\right)^2$$

$$=> V(\hat{\theta}_{mle}) = \theta^2(n/(n+2)(n+1)^2)$$

เมื่อ $n\to\infty$ $V(\hat{\theta}_{mle}) = \lim_{n\to\infty} \theta^2(n/(n+2)(n+1)^2$

$$= 0$$

$=> \hat{\theta}_{mle} = Y_n$ มีคุณสมบัติเป็นตัวประมาณที่แนบนัยต่อพารามิเตอร์

3. mle มีคุณสมบัติของ invariance property กล่าวคือ ถ้า $\hat{\theta}_{mle}$ เป็น mle ของ θ แล้ว $g(\hat{\theta}_{mle})$ จะเป็น mle ของ $g(\theta)$

4. ตัว mle จะเป็นตัวสถิติที่ใช้อธิบายพารามิเตอร์ได้อย่างพอเพียง (Sufficiency)

กำหนดให้สุ่มตัวอย่างมาประกอบด้วย $X_1,\ldots,X_n$

โดยที่ X_i มี pdf. $f(x;\theta)$; $-\infty<\theta<\infty$ $i = 1,2,\ldots,n$

$$=> L(x_1,x_2,\ldots,x_n;\theta) = L(\theta) = f(x_1;\theta)\ldots f(x_n;\theta)$$

ถ้า $y_1 = u(x_1,\ldots,x_n)$ เป็น Sufficient Statistic

จาก Fisher Neyman Criterian ได้ว่า

$$f(x_1;\theta)\ldots f(x_n;\theta) = h(u(x_1,\ldots,x_n);\theta)\, k(x_1,\ldots,x_n)$$

โดยที่ $h(u(x_1,\ldots,x_n);\theta)$ เป็นฟังก์ชั่นของ $u(x_1,\ldots,x_n)$ กับ θ

และ $k(x_1,\ldots,x_n)$ เป็นฟังก์ชั่นของ $x_1,\ldots,x_n$ ซึ่งไม่เกี่ยวข้องกับ θ

ดังนั้น
$$\frac{\partial L(x_1,\ldots,x_n;\theta)}{\partial\theta} = \frac{\partial}{\partial\theta}(h(u(x_1,\ldots,x_n);\theta)\, k(x_1,\ldots,x_n))$$
$$= k(x_1,\ldots,x_n)\frac{\partial}{\partial\theta}h(u(x_1,\ldots,x_n);\theta) = 0$$

$$k(x_1,\ldots,x_n)\, g(u(x_1,\ldots,x_n);\hat{\theta}) = 0$$
$$\hat{\theta} = t(u(x_1,\ldots,x_n),x_1,\ldots,x_n)$$

เนื่องจากว่า $u(x_1,\ldots,x_n)$ เป็นตัวสถิติที่มีคุณลักษณะเป็น Sufficiency ของ θ

ดังนั้น $\hat{\theta}$ ซึ่งเป็นฟังก์ชั่นของตัวสถิติที่ Sufficiency จึงมีลักษณะเป็นตัวสถิติที่ Sufficiency ด้วย

5. $\hat{\theta}_{mle}$ สามารถมีค่าได้มากกว่า 1 ค่า (ดูตัวอย่างที่ 5)

6. $\lim_{n\to\infty} Var(\hat{\theta}_{mle}) = MVB$

แบบฝึกหัด

1. กำหนดให้ตัวแปรเชิงสุ่ม X มี pdf. คือ $f(x;\theta) = \exp - \frac{1}{2}(x-\theta)^2/\sqrt{2\pi}$ จงหา mle โดยใช้ขนาดตัวอย่าง n

2. กำหนดให้ตัวแปรเชิงสุ่ม X มี pdf. คือ

$$f(x;\theta) = \frac{e^{-\theta}\theta^x}{x!} \; ; \; x=0,1,2,\ldots$$

สมมุติว่าเราทดลองสุ่มตัวอย่าง ขนาด 6 หน่วยแล้วได้ข้อมูลดังนี้คือ 6, 11, 4, 8, 7 และ 6 จงหาค่าของ mle ตามเงื่อนไขนี้

3. จงหา mle ของ θ โดยใช้ขนาดตัวอย่าง n ที่สุ่มมาจากประชากรที่มีการแจกแจงโดยมี pdf.

$$f(x;\theta) = (1+\theta)x^{\theta} \; ; \; \theta > -1, 0<x<1$$

4. กำหนดให้ pdf. ของตัวแปรเชิงสุ่ม X คือ

$$f(x;\theta) = e^{-x^2/2\theta^2}/\theta\sqrt{2\pi}$$ จงหา mle ของ θ

5. ในปัญหาที่ 4 นั้น กำหนดให้ตัวประมาณค่าของ θ^2 คือ λ จงหา mle ของ λ จาก $f(x;\lambda) = e^{-x^2/2\lambda}/\sqrt{2\pi\lambda}$ และเปรียบเทียบผลที่ได้กับข้อ 4

6. กล่องใบหนึ่งบรรจุลูกบอลล์ 10 ลูก โดยลูกบอลล์ประกอบด้วยสีขาว และสีอื่น ๆ สัดส่วนของลูกบอลล์สีขาวต่อสีอื่น ๆ คือ p กำหนดให้ X คือ จำนวนลูกบอลล์สีขาวในกล่องนั้น จงหาฟังก์ชั่นจากการแจกแจงของ X คือ $f(x;p)$ แล้วจงหา mle ของ p

7. ตัวแปรเชิงสุ่ม X มี pdf. คือ $f(x) = (\beta+1)x^{\beta}$; $0<x<1$

   ก. จงหา mle ของ β โดยใช้ขนาดตัวอย่าง n

   ข. จงหาค่าที่ได้ในข้อ ก. โดยสมมุติว่าข้อมูลที่สุ่มได้มีค่าเท่ากับ 0.3,0.8,0.27, 0.35, 0.62 และ 0.55

8. ตารางต่อไปนี้คือ ข้อมูลที่ได้จากการทดลองในงานวิจัยงานหนึ่ง สมมุติว่าผลจากการทดลองมีการแจกแจงแบบปกติ $N(\mu,\sigma^2)$ จงหา mle ของ μ และ σ^2

Table

Value of X	Frequency	Value of X	Freqency
1.0	2	3.7	123
1.3	29	4.0	82
1.6	62	4.3	48
1.9	106	4.6	27
2.2	153	4.9	14
2.5	186	5.2	5
2.8	193	5.5	1
3.1	188	Total frequencies 1370	
3.4	151		

9. สมมุติว่าตัวแปรเชิงสุ่ม T คือ เวลาที่ชิ้นส่วนอิเลคโทรนิคที่ใช้อยู่จะเสื่อมคุณภาพ โดยที่ T มีฟังก์ชั่นการแจกแจงดังนี้คือ

$$f(t) = \beta e^{-\beta(t-t_0)} \quad ; \quad t<t_0<0$$
$$= 0 \quad \text{นอกเหนือจากนี้}$$

(หมายเหตุในปัญหานี้ T ก็คือ ตัวแปรที่มีการแจกแจงแบบ exponential โดยที่เราหยุดค่า t (truncated) ที่ t_0)

ก. กำหนดเงื่อนไขว่า รู้ค่า t_0 จงหาตัวประมาณค่าของ β โดยวิธี mle

ข. ถ้ากำหนดว่าไม่รู้ค่า t_0 แต่ทราบค่า β จงหาตัวประมาณค่าของ t_0 โดยวิธี mle

10. สมมุติว่าตัวแปรเชิงสุ่ม X มีการแจกแจงแบบสม่ำเสมอและต่อเนื่องในช่วง $(-\alpha,\alpha)$ จงหาตัวประมาณค่าโดยวิธี mle ของ α

11. (a) A procedure is performed until a particular event A occures for the first time. On each repition P(A) = p. Suppose that n_1 repetitions are required. Then this experiment is repeated, and this time n_2 repetitions are required to yield the event A., If this is done k times we obtain the sample $n_1, n_2, \ldots, n_k$ based on this sample, we obtain the mle estimate of p.

(b) Suppose that k is quite large. Find the approximate value of $E(\hat{p})$ and $V(\hat{p})$, where $\hat{p}$ is the mle obtain in (a)

12. ชิ้นส่วนอุปกรณ์อิเลคโทรนิค อายุการเสื่อมจะมีการแจกแจงแบบ exponential ถ้าเราจับเวลาอายุการใช้งานของชิ้นส่วนดังกล่าวออกมาได้ดังนี้คือ 108, 212, 174, 130, 198, 169, 252, 168, 143 จะต้องใช้ตัวอย่างที่หาได้นี้ เพื่อหาค่า mle ของความน่าเชื่อถือ (reliability) ของส่วนนี้เมื่อใช้ไป 150 ชั่วโมงแล้ว

13. กำหนดให้ X เป็นเวลาการใช้งานของหลอดไฟนีออน (นับเป็นชั่วโมง) สมมุติให้ข้อมูลที่ได้มาเป็นดังนี้

1009,1085,1123,1181,1235,1249,1263,1292,1327,1338,1348,1352,1359,1368,1379,1397,
1406,1425,1437,1438,1441,1458,1483,1488,1499,1505,1509,1519,1541,1543,1548,1549,
1610,1620,1625,1638,1639,1658,1673,1682,1720,1729,1737,1752,1757,1783,1796,1809,
1828,1834,1871,1881,1936,1949,2007

จงหา mle ของ reliability สำหรับหลอดไฟนีออนเมื่อใช้ไปแล้ว 1,600 ชั่วโมง สมมุติว่าอายุการยืนยาวการใช้งานของหลอดไฟ จะมีการแจกแจงแบบปกติ

14. Let us assume that α-particles are emitted from a redioactive source according to a Poisson distribtion then $P(X=k) = \dfrac{\exp(-\lambda t).(\lambda t)^k}{k!}$ Insteded of recording the actual number or particals emitted, suppose that we note the number times no particle was emitted. Specially, suppose that 30 radioactive sources having the same strength are observed for a period of 50 seconds and that in 25 of the cases at least one particle was emitted. Obtain the mle of λ based on this information.

15. A random variable X has distribution $N(\mu,1)$. Twenty observations are taken on X but instead of recording the actual value we only note whether or not X was negative. Supposed that the evant $\{X < 0\}$ occured precisely 14 times. Using this information obtain the mle of μ.

16. Supposed that X has a gamma distribution; that is the pdf. is given by

$$f(x) \quad = \quad \frac{(\lambda x)^{r-1}e^{-\lambda x}}{\Gamma(r)}\,;\; x>0$$

Supposed that r is known. Let $X_1, X_2, \ldots, X_n$ be a random sample of X and obtain the mle of λ based on this sample.

17. Supposed that X has a Weibull distribution with pdf.

$$f(x) \quad = \quad (\lambda\alpha)x^{\alpha-1}e^{-\lambda x^{\alpha}}\,;\; x>0$$

Supposed that α is known. Find the mle of λ based on a sample of size n.