บทที่ 6
การประมาณค่าพารามิเตอร์โดยวิธีของเบย์
(Bayes Estimator)

ทฤษฎีการตัดสินใจ (Decision Theory)

6.1 บทนำ

การดำเนินกิจกรรมทุกอย่างไม่ว่าจะเป็นงานระดับซับซ้อนมาก เช่น ในโครงการอวกาศ หรือแม้แต่ในการดำเนินชีวิตแบบเรียบ ๆ ง่าย ๆ ของชาวชนบท ผู้ดำเนินการทุกคนล้วนแต่ต้องมีปัญหาการตัดสินใจว่าจะดำเนินการอย่างไรจึงจะถูกต้อง คำว่าถูกต้องในที่นี้ถ้าจะพูดในเชิงธุรกิจก็คือ การเลือกทางที่จะได้กำไรสูงสุดนั่นเอง

โดยปกติแล้วการที่จะดำเนินการอย่างไรอย่างหนึ่งมักจะมีแนวทางให้เลือกอยู่หลายทาง วิธีการเลือกว่าจะดำเนินการทางใดก็ยังขึ้นอยู่กับผู้เลือกว่าจะคาดการว่าควรจะเลือกทางใดที่จะรับกับเหตุการณ์ที่จะเกิดขึ้นในอนาคต เช่น เกษตรกรก่อนที่จะลงมือเพาะปลูกก็จะต้องตัดสินใจว่าเขาควรจะเพาะปลูกอะไรดี ปัจจัยหลักที่จะต้องนำมาเป็นพื้นฐานในการตัดสินใจคือ

1. สภาพดินฟ้าอากาศ เช่นปริมาณน้ำฝน
2. ราคาของผลผลิตในท้องตลาด เป็นต้น

ในการศึกษาในชั้นต้นนี้เพื่อให้ง่ายแก่การอ้างอิงจะขอพิจารณาเพียงปัจจัยเดียวคือปัจจัยในเรื่องของปริมาณน้ำฝนในฤดูกาลเพาะปลูกนั้น

กำหนดว่า ปริมาณน้ำฝน (θ) อาจจะเกิดขึ้นอยู่ 2 ระดับคือ

θ_1 ฝนตกสม่ำเสมอ ตลอดฤดูกาลนั้น

θ_2 ปริมาณฝนตกน้อยมาก

(เรียก θ ว่า State of Nature)

ทางที่จะเลือกดำเนินการ (Action) ของเกษตรกร มีอยู่ 2 ทางคือ

a_1 ปลูกอ้อย (ซึ่งเหมาะกับปริมาณน้ำฝนพอสมควร)

a_2 ปลูกสัปปะรด (ซึ่งเหมาะกับปริมาณน้ำฝนน้อย)

พิจารณาเกี่ยวกับ State of Nature และ Action จะประมวลได้ว่า

ถ้า $\theta = \theta_1$ เกษตรกรควรจะเลือก a_1 ....(1)

ถ้า $\theta = \theta_2$ เกษตรกรควรจะเลือก a_2 ....(2)

การที่จะเลือกว่าเป็นแบบที่ 1 หรือที่ 2 มีอุปสรรคอยู่ที่ว่า เราไม่สามารถทราบได้ว่า ในอนาคต θ_1 หรือ θ_2 จะเกิดขึ้น ดังนั้น ถ้าเกษตรกรเลือก Action ที่ไม่ตรงกับ State of Nature ที่เกิดขึ้น ผลเสียหายก็จะตามมา ผลเสียหายที่เกิดขึ้นจะใช้คำว่า loss

แนวทางที่จะพิจารณาต่อไปนี้แทนที่เราจะมุ่งที่กำไรสูงสุด (Maximize profit) เราจะมุ่งในด้านที่จะลดค่าเสียหายให้ต่ำที่สุดแทน (Minnimize loss)

ตัวอย่าง loss table ที่สร้างขึ้นจากตัวอย่างในการที่เกษตรกรจะเลือกแนวทางที่จะเพาะปลูก ค่าของ loss จะใช้เป็นมาตราการวัดก็ขึ้นอยู่กับลักษณะของปัญหาเป็นสำคัญ เช่น อาจวัดด้วยค่าของเงิน ความพอใจ ฯลฯ เป็นต้น (โดยปกติค่าที่จะใช้วัดเราเรียกรวมเป็นคำกลาง ๆ ว่า Ultility

การที่ผู้เลือกไม่สามารถที่จะรู้ State of Nature ว่าจะเป็นเช่นไร จึงทำให้ผู้เลือกต้องวิธีการเดาเป็นเบื้องต้น

กรณีของตัวอย่าง ถ้าเลือก a_1 และ θ_1 เกิดขึ้น loss จะมีค่า 0
ถ้าเลือก a_2 และ θ_2 เกิดขึ้น loss จะมีค่า 2
ถ้าเลือก a_1 และ θ_1 เกิดขึ้น loss จะมีค่า 3
ถ้าเลือก a_2 และ θ_2 เกิดขึ้น loss จะมีค่า 1

ถ้าเราจะยึดถือเอาหลักการของการป้องกันการสูญเสียมากที่สุด โดยการยึดถือเอาแนวทางที่ปฏิบัติดังกล่าวต่อไปนี้ "ให้เลือก Action ที่ค่า loss ต่ำที่สุดจากบรรดาค่าที่สูงสุดทั้งหลายของ loss ไม่ว่า State of Nature จะเปลี่ยนแปลงไปเช่นใด (Minimax Criterior)"

เขียนเป็นรูปแบบได้ดังนี้คือ Min of Max $1(a;\theta)$ อ่านว่า Minimize of Maximum losses

พิจารณาจากตารางประกอบ

a \ θ	θ_1	θ_2	Max $l(a;\theta)$	Min of Max $l((a;\theta)$
a_1	0	2	2	2
a_2	3	1	3	

ดังนั้นเราควรจะเลือก a_1

การเลือก Action โดยแนวทางของ Minimax Criterion วิธีนี้เรียกว่า Minimax Loss

ตัวอย่างที่ 2 กำหนดให้มี action ที่จะเลือกอยู่ 4 ทาง ในขณะที่มี state of nature อยู่ 3 ทาง ปรากฏตาราง loss ดังนี้ จงหา action โดยวิธีการของ Minimax Decision

Actions	States of nature θ_1	θ_2	θ_3	$\max\limits_{\theta} [l(a_i,\theta]$
a_1	4	7	3	7
a_2	5	2	4	5
a_3	8	6	10	10
a_4	3	1	9	9

$$\min_{A} \left\{ \max_{\theta} [(A,\theta)] \right\} = 5$$

ดังนั้น คำตอบของ action ที่ได้คือ a_2

รูปแบบทั่ว ๆ ไปของการเลือก action โดยวิธี

Minimax Decision ก็คือ

กำหนดให้ $A = \{ a_1, a_2, \ldots, a_m \}$ be the class of all posible actions available to the decision maker

$\theta = \{\theta_1, \theta_2, \ldots, \theta_n\}$ be the class of all states of nature

$l(a_i, \theta_j)$ = the loss from using action a_i when the state of nature is θ_j

GENERAL DECISION TABLE

States of nature

Actions	θ_1	θ_2	.....	θ_j	.....	θ_n
a_1	$l(a_1,\theta_1)$	$l(a_1,\theta_2)$	.....	$l(a_1,\theta_j)$	.....	$l(a_1,\theta_n)$
a_2	$l(a_2,\theta_1)$	$l(a_2,\theta_2)$	.....	$l(a_2,\theta_j)$	.....	$l(a_2,\theta_n)$
:	.....	.....	.....	.....	.....	.....
a_i	$l(a_i,\theta_1)$	$l(a_i,\theta_2)$	.....	$l(a_i,\theta_j)$	.....	$l(a_i,\theta_n)$
:	.....	.....	.....	.....	.....	.....
a_m	$l(a_m,\theta_1)$	$l(a_m,\theta_2)$	.....	$l(a_m,\theta_j)$	.....	$l(a_m,\theta_n)$

จากหลักเกณฑ์ดังกล่าวข้างต้น เราจะนำไปใช้กับแนวทาง ในการประมาณค่าพารามิเตอร์ดังนี้

A = Action Space ประกอบด้วยเซทของ $(a_1,a_2,\ldots,a_k)$ ซึ่งเทียบได้กับ $d(x_1,x_2,\ldots,x_n)$ โดยที่ $d(x_1,\ldots,x_n)$ คือ Decision ซึ่งบางทีเรียกว่า Strategy

ค่า $x_1,\ldots,x_n$ ได้จากการสุ่มตัวอย่าง (สุ่มการทดลอง)

ดังนั้น $A = (d_1(x_1,\ldots,x_n),\ldots,d_k(x_1,\ldots,x_n))$

$d(x_1,\ldots,x_n)$ เป็นฟังชั่น d ของตัวแปรเชิงสุ่มที่จะนำไปประมาณค่าพารามิเตอร์

θ = Parameter Space

(θ อาจจะเป็นชนิดต่อเนื่องหรือไม่ก็ได้)

ตัวอย่างเช่น ตัวแปรเชิงสุ่ม X มีการแจกแจงแบบปกติ $N(\mu,1)$ State of Nature ในตัวอย่างนี้ก็คือ $\theta = \mu$

$$\theta = \{\mu : -\infty < \mu < \infty\}$$

$$A = \{\hat{\mu} : -\infty < \hat{\mu} < \infty\}$$

ถ้าสุ่มตัวอย่าง $X_1,X_2,\ldots,X_n$

สมมุติว่าเราสร้าง Decision function ขึ้นมาได้ดังนี้คือ

$$d_1(x_1,\ldots,x_n) = \overline{X}$$

$$d_2(x_1,\ldots,x_n) = \Sigma X_i^2$$

$$d_3(x_1,\ldots,x_n) = \Sigma X_i^5$$

หน้าที่ของนักสถิติก็คือ จะเลือก Decision function ใดมาประมาณค่า θ ค่าของ Decision function ที่ได้มาเรียกว่าค่าคงที่ที่ใช้ประมาณค่าของพารามิเตอร์ (Point estimate)

6.2 การตัดสินใจโดยวิธีการของเบย์เมื่อปราศจากข้อสนเทศที่สนับสนุน (Bays Decision Procedure Without Data)

loss function ใช้สัญลักษณ์แทนว่า $l(a;\theta)$ หรือ $l(\hat{\theta};\theta)$ ทั้งนี้ขึ้นอยู่กับจุดมุ่งหมายของการนำไปใช้

ได้อธิบายมาแล้วถึงความหมายคร่าว ๆ ของ loss function แล้วจะเห็นได้ว่า loss function ทำหน้าที่คล้าย ๆ กับมาตรการวัดความสูญเสียที่จะเกิดขึ้นในการดำเนินการใด ๆ ซึ่งขึ้นอยู่กับภาวะการที่ไม่แน่นอน

ลักษณะที่สำคัญของ loss function

1. $l(a;\theta)$ เป็น real valued non-negative function[1/]
2. ค่าของ $l(a;\theta)$ เป็น 0 แสดงว่า a เป็น Action ที่ถูกต้อง[2/]

รายละเอียดในเรื่องของ loss function จะแยกพิจารณาการศึกษาออกมาเป็น 2 พวกคือ

1. Avoidable loss
2. Unavoidable loss

พิจารณาจากตัวอย่างเพื่อให้อธิบายได้ง่ายขึ้น

ตารางที่ 1

State of Nature	Action		
	Increase rate	Usual rate	Reduced rate
θ_1 Prosperity	1	3	5
θ_2 S'tability	3	1	3
θ_3 Recession	7	4	1

---

1 หนังสือบางเล่มยอมให้ค่า $l(a;\theta)$ เป็นค่าลบได้ ซึ่งนั่นก็คือ กำไร (profit)

2 ในบางปัญหาที่มีเรื่องของ Avoidable loss เข้ามาเกี่ยวข้องข้อที่ 2 อาจเปลี่ยนแปลงได้

จากตารางจะเห็นได้ว่า

Avoidable loss สำหรับ Action ใด ๆ ก็จะมีค่าเท่ากับ 1 (ถึงแม้ว่าจะสามารถทราบได้ว่า State of Nature จะเป็นเช่นใด)

ดังนั้น Avoidable loss สำหรับในตัวอย่างนี้ก็คือ loss ที่เกิดจาก Ignorance of the State of Nature

เพื่อเป็นการเสี่ยงที่จะเลือก Action โดยเข้าไปเกี่ยวข้องกับ Avoidable loss เราจะสร้าง regret (Opportunity loss) เพื่อใช้เป็นแนวทางในการเลือก Action แทน loss function

สร้าง regret table ได้ว่า

State of Nature	Action		
	a_1	a_2	a_3
θ_1	0	2	4
θ_2	2	0	2
θ_3	6	3	0

เราสามารถที่จะประยุกต์การหา Action โดยวิธี Minimax loss มาเป็น Minimax regret แทนได้

โดยที่ Minimax regret = Min (Max($a_1;\theta$),...,Max($a_k;\theta$))

ดังนั้น Minimax regret สำหรับปัญหานี้ก็คือ Action a_2

เราอาจจะเขียนเป็นรูปทั่วไปได้ดังนี้

regret function = $r(a_i, \theta_j)$

ดังนั้น

$$r(a_i, \theta_j) = l(a_i, \theta_j) - \min_{A} l(a_k, \theta_j) \quad i = 1,2,\ldots, m$$
$$j = 1, 2, \ldots, n$$

สำหรับแต่ละ state of nature θ_j

ในบางครั้งนั้นถ้าหากเราได้ว่า

$$l(a_i, \theta_j) = \min_{A} [l(a_k, \theta_j)]$$

สำหรับ ค่า i = t ดังนั้นจะได้ว่า

$$r(a_t,\theta_j) = l(a_t,\theta_j) - l(a_t,\theta_j) = 0$$

การจะใช้ตาราง loss หรือตาราง Regret นั้น จะให้คำตอบการเลือก action ในทิศทางเดียวกัน ซึ่งเราจะสามารถแสดงให้ดูได้โดยใช้ตัวอย่างต่อไปนี้

LOSS TABLE

Actions	States of nature θ_1	θ_2	θ_3	$\max_{\theta} [l(a_i,\theta)]$
a_1	3	8	1	8
a_2	9	6	5	9
a_3	7	4	8	8
a_4	5	5	6	6

$$\min_{A} \left\{\max_{\theta} [l(A,\theta)]\right\} = 6$$

เลือกaction ที่ a_4

Regret Table

Actions	States of nature θ_1	θ_2	θ_3	$\max_{\theta} [l(a_i,\theta)]$
a_1	0	4	0	4
a_2	6	2	4	6
a_3	4	0	7	7
a_4	2	1	5	5

$$\min_{A} \left\{\max_{\theta} [l(A,\theta)]\right\} = 4$$

เลือก action ที่ a_1

การตัดสินใจโดยวิธีการของเบย์

(Bays Decision Procedure Without Data)

วิธีการทั้งหมดที่กล่าวมานี้เป็นวิธีการเลือก Action โดยตั้งอยู่บนพื้นฐานที่ว่าผู้ตัดสินใจเลือก Action โดยที่ไม่ทราบลักษณะการแจกแจงของ State of Nature เลย

ถ้าหากว่าเรามีความรู้ในเรื่องลักษณะของ State of Nature ก็จะเป็นส่วนที่ช่วยให้เลือก Action ได้ถูกต้องยิ่งขึ้น

Expected loss

สมมุติว่าผู้ดำเนินการตัดสินใจเชื่อว่าความน่าจะเป็นของ prosperity, stability ละ recession มีค่าคือ $p_1 = .7, p_2 = .2$ และ $p_3 = .1$ ดังนั้นเราจะหาค่าของ Expected loss $= E(l(a;\theta))$ ได้ดังนี้

ตารางที่ 2

probability p	State of Nature	Action		
		a_1	a_2	a_3
$p_1 = 0.70$	θ_1	1	3	5
$p_2 = 0.20$	θ_2	3	1	3
$p_3 = 0.10$	θ_3	7	4	1
Expected loss $l(a;\theta)$		2.0	2.7	4.2

ตัวอย่างการคำนวณหาค่า Expected loss

$$E(l(a_1;\theta)) = E(l(a_1)) = 1(0.70) + 3(0.20) + 7(0.10)$$
$$= 2.0$$
$$E(l(a_2)) = 2.7$$
$$E(l(a_3)) = 4.2$$

a_1 เป็น Action ที่ให้ค่า Expected loss ต่ำที่สุดคือเพียง 2.0 ดังนั้นในปัญหานี้เราควรจะเลือก Action a_1

หลักเกณฑ์ทั่วไปในการเลือก Action โดยอาศัย ค่าของ Expected loss ก็คือ

Minimum expected loss $= \text{Min}(l(a_1),\ldots,l(a_k))$

Expected regret

การใช้ Expected regret ก็เนื่องจากเหตุผลที่กล่าวมาแล้วข้างต้นในเรื่องของ Avoidable loss ดังนั้นเราจะเลือก Action โดยใช้หลักการของ Minimum $E(r(a;\theta))$ แทน

probability p	State of Nature	Action		
		a_1	a_2	a_3
$p_1 = 0.7$	θ_1	0	2	4
$p_2 = 0.2$	θ_2	2	0	2
$p_3 = 0.1$	θ_3	6	3	0
		1	1.7	3.2

ตารางที่ 3

พิจารณาจากตารางที่ได้ จะพบค่า Expected regret มีความสัมพันธ์กับ Expected loss คือ ถ้าเราทราบค่า Expected loss อยู่แล้ว ก็ให้นำค่าของ Avoidable loss หักออกจากค่าของ Expected loss แต่ละค่าก็จะได้ค่า Expected regret

ตัวอย่างที่ 1 กำหนดให้ตาราง loss และค่าของความน่าจะเป็นของ θ (prior distribution of θ) เป็นดังนี้ จงหา action ที่ดีที่สุด

	State of Nature		
Actions	θ_1	θ_2	θ_3
a_1	4	7	3
a_2	5	2	4
a_3	8	6	10
a_4	3	1	9

$$P_\theta(\theta_j) = P(\theta = \theta_j) = \begin{cases} \frac{1}{2} & \theta = \theta_1 \\ \frac{1}{3} & \theta = \theta_2 \\ \frac{1}{6} & \theta = \theta_3 \end{cases}$$

then the expected loss for action a_1 is

$$\begin{aligned} E_\theta[l(a_1,\theta)] &= \frac{1}{2}\, l(a_1,\theta_1) + \frac{1}{3}\, l(a_1,\theta_2) + \frac{1}{6}\, l(a_1,\theta_3) \\ &= \frac{1}{2}(4) + \frac{1}{3}(7) + \frac{1}{6}(3) \\ &= \frac{29}{6} \end{aligned}$$

For actions a_2, a_3, and a_4, the expected losses are

$$\begin{aligned} E_\theta[l(a_2,\theta)] &= \frac{1}{2}(5) + \frac{1}{3}(2) + \frac{1}{6}(4) \\ &= \frac{23}{6} \end{aligned}$$

$$\begin{aligned} E_\theta[l(a_3,\theta)] &= \frac{1}{2}(8) + \frac{1}{3}(6) + \frac{1}{6}(10) \\ &= \frac{46}{6} \end{aligned}$$

$$E_\theta[l(a_4,\theta)] = \frac{1}{2}(3) + \frac{1}{3}(1) + \frac{1}{6}(9)$$
$$= \frac{20}{6}$$

ดังนั้น action ที่จะเลือกใช้ก็คือ action a_4 เพราะว่าให้ค่า $E_\theta[l(a_4,\theta)]$ ต่ำกว่าบรรดา action อื่น ๆ

ตัวอย่างที่ 2 ในร้านขายของแห่งหนึ่งต้องการจะสั่งสต๊อกสินค้าประเภทเครื่องกันหนาว ไว้เพื่อขายในฤดูหนาวที่จะมาถึง ซึ่งการสต๊อกสินค้าจะมีอยู่ 4 ระดับ แต่ละระดับขึ้นอยู่กับสภาพดินฟ้าอากาศ ซึ่งจะเป็นไปได้อยู่ 3 ระดับ การสต๊อกของนี้เราสามารถจะตาราง loss อันเนื่องมาจากสภาพดินฟ้าอากาศ ได้ดังนี้คือ

States of nature

Actions	θ_1 $P(\theta=\theta_1) = \frac{1}{2}$	θ_2 $P(\theta=\theta_2) = \frac{1}{3}$	θ_3 $P(\theta=\theta_3) = \frac{1}{6}$	$E[l(a_i,\theta)]$
a_1	4	7	3	29/6
a_2	5	2	4	23/6
a_3	8	6	10	46/6
a_4	3	1	9	20/6

$$\min_A \{E_\theta[l(A,\theta)]\} = \frac{20}{6}$$

ตัวอย่างที่ 3 กำหนดตาราง loss ดังนี้ จงหาaction ที่ดีที่สุดโดยวิธีของ Minimax ของ $E_\theta(l(A,\theta))$

Action: volume on hand	θ_1 Excellent $P_{\theta_1} = 0.5$	θ_2 Good $P_{\theta_2} = 0.4$	θ_3 Poor $P_{\theta_3} = 0.1$	$\max_\theta [l(a_i,\theta)]$	$E_\theta[l(a_i,\theta)]$
a_1:high	−20	−12	−6	−6	−15.4
a_2:med	−16	−16	−8	−8	−15.2
a_3:low	−12	−12	−12	−12	−12.0

$$\min_A \{\max_\theta [l(A,\theta)]\} = -12$$

$$\min_A \{E_\theta[l(A,\theta)\} = -15.4$$

ถ้าเราใช้วิธีของ Minimax จะได้รับคำตอบว่า action ที่จะเลือกคือ a_3 แต่ถ้าเราใช้ Bays decision จะได้ว่า action ที่จะเลือกคือ a_1

สรุปกฎเกณฑ์ในการใช้ Bays decision
(Bays Decision Without Data)

Step 1

Use the prior distribution of θ to calculate the expected loss for each action a_i with i = 1, 2, ..., m. That is, let

$$E_\theta[l(a_i, \theta)] = \sum_{j=1}^{n} l(a_i,\theta_j) \cdot P(\theta=\theta_j) \text{ for } i = 1, 2, \ldots, m$$

Step 2

Select the action that has the smallest expected loss. Stop.

การเลือก Action มาแล้วทั้งหมดนี้เป็นเรื่องของการตัดสินใจที่เราไม่ได้ใช้ข้อมูลอื่นมาประกอบเว้นแต่ loss หรือ regret table กับฟังก์ชั่นความน่าจะเป็นของ State of Nature หรือ θ นั่นเอง เราเรียกวิธีการเช่นนี้ว่า "Decision Without Information" หรือ No Data Problem แต่ในตอนที่กล่าวถึงวิธีการตัดสินใจในการที่จะเลือกวิธีการประมาณค่าพารามิเตอร์ จะเป็นวิธีการที่อยู่ในข่ายของ "Decision with Information"

ถ้าเรากำหนดให้มี J State of Nature $\theta_1,\theta_2,\ldots,\theta_j$

มี k Action คือ $a_1,a_2,\ldots,a_k$

โดยที่ θ_i มีโอกาสที่จะเกิดขึ้นด้วยความน่าจะเป็น p_i เราจะเรียก เซท $(p_1,\ldots,p_j)$ ว่าเป็น prior distribution of θ โดยที่ $\Sigma p_i = 1$

ดังนั้นเราสามารถที่จะสร้าง Risk ได้ดังนี้

$$R(a_1) = \text{Expected regret for } a_1 = \sum^{J} p_i\, r(a_1;\theta_i)$$

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

$$R(a_k) = \text{Expected regret for } a_k = \sum^{J} p_i\, r(a_k;\theta_i)$$

Decision rule คือให้เลือก Action a ที่ตรงกับเงื่อนไขนี้ คือ

Minimum risk = $\text{Min}(R(a_1),\ldots,R(a_k))$

Table Generalized regret for the no-data

ตารางที่ 4

A priori probability	State of Nature	Action	
		$a_1 \ldots a_m \ldots, \ldots, a_k$	
p_1	θ_1	$r(\theta_1;a_1)$	
p_2	θ_2	$r(\theta_2;a_1)$	
$\vdots$	$\vdots$	$\vdots$	
p_J	θ_J	$r(\theta_J;a_1)$	$r(\theta_J;a_k)$
Expected regret R(a)		$= R(a_1)$	$R(a_k)$
		$= \Sigma p_i r(\theta_i;a_1)$	$= \Sigma p_i r(\theta_i;a_k)$

Minimum risk ที่เราหาได้เรียกว่า Bayes' risk Action คล้อยตาม Bayes' risk จะเรียกว่า Bayes' action จากตัวอย่างในตารางที่ 3[1/] จะได้ว่า Bayes' risk มีค่าเป็น 1 ส่วน Bayes' action ก็คือ a_1 นั่นเอง การเลือก Action โดยใช้หลักของ risk (หรือ Expected regret) ก็คล้อยตามกับการหา Expected disutility (Expected negative utility) นั่นคือ Bayes' risk ก็คือวิธีการเดียวกับการ Minimum expected disutility หรือ Maximum expected disutility จะกล่าวอีกนัยหนึ่งก็เหมือนกับการเลือกวิธีดำเนินการโดยที่ให้ขาดทุน(จากกำไร)น้อยที่สุด นั่นก็คือการหวังกำไรมากที่สุดนั่นเอง (เป็นการมองสิ่งเดียวกันแต่มองกันคนละด้าน)

1/ $R(a_1) = p_1 r(a_1;\theta_1) + p_2 r(a_1;\theta_2) + p_3 r(a_1;\theta_3)$

$= 0.7(0) + 0.2(2) + 0.1(6)$

$= 0.4 + 0.6 = 1.0$

$R(a_2) = 2.17$

$R(a_3) = 3.52$

Loss Function

รูปลักษณะของฟังชั่นของ loss โดยทั่วไปแยกออกเป็น 3 แบบคือ

1. $l(\theta;\hat{\theta}) = c(\theta)(\theta - \hat{\theta})^2$

   โดยที่ $c(\theta) > 0$

   เมื่อ $c(\theta) = 1$ เรียกว่า Squared-error loss function

2. $l(\theta;\hat{\theta}) = c(\theta)\,|\,\theta - \hat{\theta}\,| \quad ; c(\theta) > 0$

   เรียกว่า absolute error function

3. $l(\theta;\hat{\theta}) = A\ ;\ \hat{\theta} > \theta + k_1$

   $= B\ ;\ \theta - k_2 < \hat{\theta} < \theta + k_1$

   $= C\ ;\ \hat{\theta} < \theta - k_2$

   Step error loss function

Risk Function

ปัญหาในทางสถิติที่เกิดขึ้นในการใช้ loss function หรือ regret function ที่ตามมาก็คือ ทั้งคู่ต่างก็เป็นตัวแปรเชิงสุ่ม (ขึ้นอยู่กับ $X_1, X_2, \ldots, X_n$ ดังที่ได้อธิบายมาแล้วบ้าง)

ดังนั้นเราจึงต้องหันมาสนใจในค่าเฉลี่ยของ loss หรือ regret แทน ใช้สัญลักษณ์ $R(\theta;d)$ or $R(\theta;a)$ แทนฟังชั่นของ risk

$$R(\theta;d) = E(l(\theta;a))$$
$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} l(\theta;d(x_1,\ldots,x_n))\, f(x_1;\theta) \ldots f(x_n;\theta)\, dx_1 \ldots dx_n$$

เกณฑ์การเลือกตัวประมาณค่าของ θ (Action or Decision) จึงยึดเอา $R(\theta;d)$ เป็นแนวทางต่อไป

แบบฝึกหัด

1. โรงงานผลิตเครื่องซักผ้าแห่งหนึ่งได้สร้างแบบจำลองของเครื่องซักผ้าชนิดใหม่ขึ้นมา 3 แบบด้วยกัน ถ้าหากว่าในระบบเศรษฐกิจของประเทศอยู่ในภาวะเงินฝืดเครื่องชนิดราคาต่ำจะขายได้ดี ถ้าหากว่าเศรษฐกิจอยู่ในภาวะปกติเครื่องชนิดธรรมดาจะขายได้ดี และถ้าหากว่าภาวะเศรษฐกิจดีเครื่องชนิดพิเศษจะขายได้ดี

กำหนดให้ a_1, a_2, a_3 เป็นทางที่จะเลือก 3 ทางในการที่จะสร้างเครื่องซักผ้าขึ้นมาจำหน่าย

$\theta_1, \theta_2, \theta_3$ คือภาวะเศรษฐกิจของประเทศ คือภาวะเศรษฐกิจตกต่ำ ภาวะเศรษฐกิจปกติ และภาวะเศรษฐกิจดี ตามลำดับ

ให้ผลที่จะได้รับ (Pay off or Utility) ของบริษัทที่จะได้รับในการดำเนินกิจการเป็นดังนี้คือ

Pay off table ($u(\theta;a)$)

State of Nature		Action		
		Low price model	Regular model	Deluxe model
		a_1	a_2	a_3
Recession	θ_1	5	4	1
Stability	θ_2	4	8	5
Boom	θ_3	3	6	10

a. ถ้าใช้หลักเกณฑ์ของ Minimax ผู้จัดการโรงงานจะเลือกดำเนินทางใดในการผลิตสินค้า

b. ถ้าใช้หลักเกณฑ์ของ Maximin เพื่อให้ได้กำไรสูงสุด ผู้จัดการจะเลือกดำเนินการทางใด

ในทางปฏิบัติค่าของผลที่พึงจะได้รับสามารถหาได้จากการเฝ้าติดตามสังเกต

2. ใช้ปัญหาในข้อ ข. ถ้ากำหนดให้ความน่าจะเป็นที่จะเกิดภาวะเศรษฐกิจดังกล่าวคือ

$p(a_1) = 0.2$, $p(a_2) = 0.5$, $p(a_3) = 0.3$

a. จงหา Expected pay off สำหรับแต่ละแบบจำลองของสินค้า

b. ให้หา Bayes' action และ Bayes' pay off

3. จากปัญหาในข้อ 1 กำหนดให้ผลที่จะได้รับ เป็นไปตามตาราง

Pay off $1(\theta;a)$

State of Nature	Action		
	a_1	a_2	a_3
θ_1	1	4	7
θ_2	3	2	5
θ_3	5	4	3

a. ถ้าใช้หลักเกณฑ์ของ Minimax loss ผู้จัดการจะเลือกดำเนินการทางใด

b. จงสร้าง Regret table แล้วเราจะเลือกดำเนินการทางใด

4. ในจังหวัดเชอราปุน แบ่งลักษณะของดินฟ้าอากาศได้เป็นอย่างเดียวคือ ฝนตกหรือฝนไม่ตก คนที่อาศัยอยู่ในเมืองนี้มีวิธีการเลือกอยู่ 2 อย่างก่อนไปทำงานคือ นำเสื้อกันฝนติดตัวไปด้วย หรือไม่นำไป กำหนดให้การที่จะนำเสื้อกันฝนไปด้วยหรือไม่ก่อให้เกิดการสูญเสีย (loss) ดังตารางกำหนดไว้คือ

State of Nature	Actions	
	ใส่เสื้อกันฝน	ไม่ใส่เสื้อกันฝน
	a_1	a_2
ฝนตก θ_1	2	6
ฝนไม่ตก θ_2	5	0

a. คนในเมืองนี้จะเลือกวิธีการใด ถ้าใช้หลักเกณฑ์ของ Minimax loss ค่า loss ที่เกิดขึ้นมีค่าเท่าไร

b. ถ้าใช้ Minimax regret จะได้วิธีใด regret มีค่าเท่าใด

c. ถ้าหากมีเครื่องมือที่สามารถทำนายได้อย่างแม่นยำว่าวันไหนฝนจะตกวันไหนจะไม่ตกเราจะเลือกวิธีใด ในสถานะการณ์แต่ละอย่างที่เกิดขึ้น

5. จากปัญหาในข้อ 4 ถ้าหากว่ากรมอุตุนิยมวิทยาของเมืองนั้นประกาศว่าโอกาสที่ฝนจะตกในวันหนึ่งมีค่า 70% ถามว่า Bayes' action คืออะไร ให้ค่า Bayes' risk เป็นเท่าไร

6.3 การตัดสินใจเมื่อมีข้อสนเทศ

(Decision with Information)

การประมาณโดยวิธีการของเบย์

การหาค่าประมาณของพารามิเตอร์เพื่อศึกษาถึงคุณลักษณะของประชากรนั้นโดยปกติแล้ว สมมุติฐานเบื้องต้นที่ใช้กันก็คือไม่ทราบคุณลักษณะด้านใด ๆ ของพารามิเตอร์ทั้งสิ้น เราจึงสร้างค่าประมาณโดยวิธีการต่าง ๆ ที่เหมาะสมกับแต่ละปัญหา เช่นใช้ Maximum-likelihood Estimator, Moment Estimator เป็นต้น

เป็นที่น่าสังเกตว่าในบางครั้งเราสามารถจะศึกษาถึงคุณลักษณะของพารามิเตอร์ได้ (ถึงแม้ว่าจะไม่ทราบค่า) คุณลักษณะที่ว่านี้ก็คือการแจกแจงของตัวพารามิเตอร์

ตัวอย่างเช่น ในโรงงานผลิตถ่านไฟฉายแห่งหนึ่ง ได้สังเกตว่าสัดส่วนของถ่ายไฟฉายที่ไม่ได้มาตรฐานคือ p จะมีลักษณะการแจกแจงเป็น $h(p) = 6p(1-p) ; 0 \leqslant p \leqslant 1$ ซึ่งในกรณีเช่นนี้ก็หมายความว่าตัวพารามิเตอร์ p ทำหน้าที่คล้ายกับตัวแปรเชิงสุ่ม คือเราไม่ทราบค่าแต่พอจะรู้จักฟังชั่นการแจกแจง ปัญหาต่อไปก็คือเมื่อทราบฟังชั่นการแจกแจงของ p แล้วเราจะนำฟังก์ชั่นนี้ไปใช้ประโยชน์ได้อย่างไรในการที่จะประมาณค่า p

ก่อนที่จะศึกษาถึงวิธีการหาค่าประมาณของ p ต่อไปโดยอาศัยฟังก์ชั่นการแจกแจง p เราจะต้องทำความเข้าใจกับพื้นฐานที่เกี่ยวข้องนั่นคือ ในเรื่องของการประมาณค่าพารามิเตอร์ด้วยวิธีการต่าง ๆ ที่กล่าวมาแล้วแต่แรก ล้วนแต่อาศัยข้อมูลของพารามิเตอร์ที่แอบแฝงมาในรูปของตัวแปรเชิงสุ่มที่ได้ทั้งสิ้น โดยไม่มีการอ้างอิงคุณลักษณะใด ๆ ของพารามิเตอร์

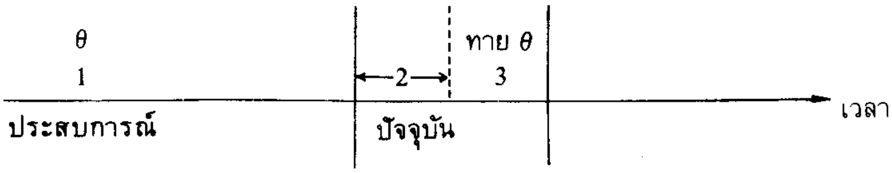
กำหนดให้ $X_1, X_2, \ldots, X_n$ เป็นตัวแปรเชิงสุ่มชุดหนึ่งที่สุ่มจากประชากรที่มี pdf. $f(x;\theta)$

จะได้ว่า $f(x_1, x_2, \ldots, x_n; \theta) = \prod_{i=1}^{n} f(x_i; \theta)$

แล้วจึงนำ $f(x_1, x_2, \ldots, x_n; \theta)$ ไปสร้างค่าประมาณของพารามิเตอร์ ที่กล่าวมาแล้วเป็นการสรุปวิธีการของการสร้างค่าประมาณด้วยวิธีการอื่นที่ไม่ใช่เบย์

วิธีการประมาณค่าของเบย์

พิจารณาจากรูปประกอบ



1. จากอดีต เราใช้ประสบการณ์ในการสร้างฟังก์ชั่นการแจกแจงของพารามิเตอร์ (ช่วงที่ 1) $p(\theta)$
2. ปัจจุบัน (ในช่วงที่ 2) เราสุ่มการทดลองมาชุดหนึ่งแต่ละตัวแปรมี pdf. $f(\theta/x_1,x_2,\dots,x_n)$ โดยที่ ได้ค่า $x_1,x_2,\dots,x_n$
3. ในช่วงที่ 3 เราจะนำผลที่ได้จาก 1, 2 มาสร้าง pdf. คือ $f(x/\theta)$

จุดประสงค์ในการสร้าง $f(\theta/x)$ ก็เพื่อที่จะศึกษาหาค่าประมาณของ θ หลังจากที่เราได้ค่าสังเกต $x_1,x_2,\dots,x_n$ แล้ว

นอกเหนือไปจากขั้นตอนทั้งสามที่กล่าวมาแล้ว ยังมีส่วนที่เข้ามาเกี่ยวข้องอีกคือ loss function วิธีการของเบย์ก็คือสร้างค่าประมาณของพารามิเตอร์ โดยที่มีข้อแม้ว่าตัวประมาณ ที่สร้างขึ้นจะต้อง Minnimize Bayes' risk (B(d))

โดยทั่วไป $R(\theta;d) = E(l(\theta;\hat{\theta}))$

$B(d) = E(R(\theta;d))$

บางครั้งเราเรียก Bayes' risk นี้ว่า Expected risk

การศึกษาเรื่องของ B(d) โดยใช้กรณีย์ของฟังชั่นต่อเนื่องเป็นตัวอย่างจะได้ว่า

$$
\begin{aligned}
B(d) &= E(R(\theta;d)) \\
&= \int_{\infty}^{\infty} R(\theta;d)\, p(\theta)\, d\theta \\
&= \int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} l(\theta;d(x_1,\dots,x_n))g(x_1,\dots,x_n/\theta)\, dx_1\dots dx_n\right\} p(\theta)d\theta
\end{aligned}
$$

เปลี่ยนลำดับของอินทรีเกรชั่นจะได้ว่า

$$B(d) = \int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty} l(\theta;d(x_1,\ldots,x_n))g(x_1,\ldots,x_n/\theta)p(\theta)d\theta\right\} dx_1\ldots dx_n \qquad \ldots(1)$$

พิจารณา

(a) $g(x_1,\ldots,x_n/\theta)p(\theta) = h(x_1,\ldots,x_n,\theta)$

(b) $h(\theta/x_1,\ldots,x_n) = \dfrac{g(x_1,\ldots,x_n/\theta)p(\theta)}{k(x_1,\ldots,x_n)}$

(c) $k(x_1,\ldots,x_n) = \int_{-\infty}^{\infty} h(x_1,\ldots,x_n,\theta)\, d\theta$

เราเรียก $h(\theta/x_1,\ldots,x_n)$ ว่า Posterior density

$p(\theta)$ ว่า Prior density

สมการที่ 1 จุดมุ่งหมายคือหา d ที่ Minimize ฟังชั่นดังกล่าว (B(d))
การที่จะ Minimize ฟังชั่นที่ 1 นั้นเรามุ่งที่จะ Minimize เพราะ

$$\int_{-\infty}^{\infty} l(\theta;d(x_1,\ldots,x_n))\ g(x_1,\ldots,x_n/\theta)p(\theta)\ d\theta \qquad \ldots(2)$$

ก็พอเพียง

นำผลที่ได้จาก (a),(b),(c) มาแทนใน 2 จะได้ว่า

$$\int_{-\infty}^{\infty} l(\theta;d(x_1,\ldots,x_n))h(\theta/x_1,\ldots,x_n)\ k(x_1,\ldots,x_n)\ d\theta$$

ผลสุดท้าย ฟังก์ชั่นที่จะ Minimize จะเหลือเพียง

$$\int_{-\infty}^{\infty} l(\theta;d(x_1,\ldots,x_n))\ h(\theta/x_1,\ldots,x_n)\ d\theta$$

$$\rightarrow \quad v(\theta;x_1,\ldots,x_n) = \int_{-\infty}^{\infty} l(\theta;d(x_1,\ldots,x_n))\ h(\theta/x_1,\ldots,x_n)\ d\theta$$

เราเรียก $v(\theta;x_1,\ldots,x_n)$ ว่า Posteriori risk

ทฤษฎี $\hat{\theta}$ ซึ่งอยู่ในรูปฟังชั่นของ ($x_1,\ldots,x_n$ ที่เราสังเกตมา) โดยที่ $\hat{\theta}$ เป็นค่าที่ Minimize a Posteriori risk $v(\theta;x_1,\ldots,x_n)$ เราจะเรียก ตัวประมาณค่าที่ได้โดยวิธีนี้ว่า ตัวประมาณที่ได้จากวิธีการของเบย์

หมายเหตุ

ในกรณีที่เป็นฟังชั่นชนิดไม่ต่อเนื่อง (discrete) วิธีการก็ยังคงเดิมทุกอย่างเพียงแต่เปลี่ยนจากเครื่องหมายอินทรีเกร์ชั่นมาเป็นเครื่องหมายผลรวมแทน

ตัวอย่างที่ 1 กำหนดให้ $X_1,\ldots,X_n$ เป็นตัวแปรเชิงสุ่มที่มาจากประชากรที่มี pdf.

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

กำหนดให้ใช้ Squared error loss $l(\theta;\hat{\theta}) = (\theta - \hat{\theta})^2$

และ θ มีการแจกแจงแบบสม่ำเสมอในช่วง $(0 \leqslant \theta \leqslant 1)$

จงหาค่าประมาณของ θ ด้วยวิธีการของเบย์

$$g(x_1,\ldots,x_n/\theta) = \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}$$

$$p(\theta) = 1$$

$$k(x_1,\ldots,x_n) = \int_0^1 \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}.1\ d\theta$$

$$= \frac{(\Sigma x_i)!\ (n-\Sigma x_i)!}{(n+1)!}$$

$$h(\theta/x_1,\ldots,x_n) = \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i} / \left[\frac{(\Sigma x_i)!\ (n-\Sigma x_i)!}{(n+1)!}\right]$$

$$= (n+1)!\ \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i} / [\ (\Sigma x_i)!\ (n-\Sigma x_n)!\ ]$$

$\hat{\theta} = d(x_1,\ldots,x_n)$ a posteriori risk

$$v(\theta;x_1,\ldots,x_n) = \int_0^1 (\theta-\hat{\theta})^2 \frac{(n+1)!\ \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}}{(\Sigma x_i)!\ (n-\Sigma x_n)!} d\theta$$

$$\frac{\partial v}{\partial \hat{\theta}}(\theta;x_1,\ldots,x_n) = \int_0^1 2(\theta-\hat{\theta}) \frac{(n+1)!\ \theta^{\Sigma x_i}(1-\theta)^{n-\Sigma x_i}}{(\Sigma x_i)!\ (n-\Sigma x_i)!} d\theta = 0$$

$$\hat{\theta} = (\Sigma x_i + 1)/(n+2)$$

Bayes' estimator $\hat{\theta}$ คือ $\hat{\theta} = (\Sigma x_i + 1)/(n+2)$

ตัวอย่างที่ 2 กำหนดให้ตัวแปรเชิงสุ่ม $X_1,\ldots,X_n$ ที่สุ่มมาจากประชากรที่มีการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเป็น μ และความแปรปรวนเป็น 1

ให้หาค่าประมาณของค่าเฉลี่ยของประชากรโดยวิธีการของเบย์

$$f(x/\mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-\mu)^2}$$

$$g(x_1,x_1,\ldots,x_n/\mu) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}(\Sigma x_i - 2\mu\Sigma x_i + n\mu^2)}$$

สมมุติว่า μ เป็นตัวแปรเชิงสุ่ม ที่มีฟังก์ชั่นการแจกแจงเป็น

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

$$g(x_1,x_2,\ldots,x_n,\mu) = \frac{1}{(2\pi)^{(n+1)/2}} \exp\{-\frac{1}{2}(\Sigma x_i^2 + (n+1)\mu^2 - 2n\bar{x})\}$$

$$k(x_1,x_2,\ldots,x_n) = \frac{1}{(2\pi)^{(n+1)/2}} \exp(-\frac{1}{2}\Sigma x_i^2) \int_{-\infty}^{\infty} \exp\{-\frac{1}{2}(n+1)\mu^2 - 2\mu n\bar{x}\} d\mu$$

โดยวิธีการสร้างให้เป็นสมการกำลังสอง จำนวนที่อยู่ภายใต้เครื่องหมายอินทริกัล จะได้ว่า ผลที่ได้ของ $k(x_1,x_2,\ldots,x_n)$ จะเป็น

$$\frac{1}{(2\pi)^{n/2}} \exp\left[-\frac{1}{2}(\Sigma x_i^2 - \frac{n^2\bar{x}^2}{(n+1)})\right] \left\{ \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \exp\left[-\frac{1}{2}(n+1)(\mu - \frac{n\bar{x}}{n+1})^2\right] d\mu \right.$$

$$k(x_1,x_2,\ldots,x_n) = \frac{1}{(n+1)^{1/2}(2\pi)^{n/2}} \exp\left[-\frac{1}{2}\Sigma x_i^2 - \frac{n^2\bar{x}^2}{n+1}\right]$$

$$h(\mu/x_1,x_2,\ldots,x_n) = \frac{(2\pi)^{-(n+1)/2} \exp\left\{-\frac{1}{2}\left[\Sigma x_i^2 + (n+1)\mu^2 - 2n\bar{x}\mu\right]\right\}}{(2\pi)^{-(n/2)}(n+1)^{-1/2} \exp\left[-\frac{1}{2}(\Sigma x_i^2 - \frac{n^2\bar{x}^2}{(n+1)})\right]}$$

$$= \frac{(n+1)^{1/2}}{(2\pi)^{1/2}} \exp\left\{-\frac{1}{2}(n+1)\left[\mu - \frac{n\bar{x}\mu}{n+1}\right]^2\right\}$$

จะเห็นได้ว่า Conditional distribution of μ, given $x_1,\ldots,x_n$

มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น $\bar{X}n/(n+1)$ และความแปรปรวนเป็น $(n+1)^{-1}$ กำหนดให้ใช้ Sqared-error loss function จะได้ว่า ค่าประมาณโดยวิธีการของเบย์ ก็คือค่าของ $\hat{\theta}$ ซึ่งเป็นฟังก์ชั่นของ $x_1,\ldots,x_n$ ซึ่ง Minimizes $v(\hat{\mu};x_1,\ldots,x_n)$

$$v(\hat{\mu};x_1,\ldots,x_n) = \int_{-\infty}^{\infty}(\hat{\mu}-\mu)^2 . h(\mu/x_1,\ldots,x_n)\,d\mu$$

$$= \frac{(n+1)^{1/2}}{(2\pi)^{1/2}}\int_{-\infty}^{\infty}(\hat{\mu}-\mu)^2 \exp\left\{-\frac{1}{2}(n+1)\left(\mu-\frac{n\bar{x}}{n+1}\right)^2\right\}d\mu$$

$$= \hat{\mu}^2 - \frac{2\hat{\mu}\bar{x}n}{n+1} + \frac{1}{n+1} + \frac{\bar{x}^2n^2}{(n+1)^2}$$

$$\partial v/\partial\hat{\mu} \rightarrow 2\hat{\mu} - \frac{2\bar{x}n}{(n+1)} = 0$$

$$\hat{\mu}_{Bayes} = \Sigma X_i/(n+1)$$

ตัวอย่างที่ 3 สมมุติว่าร้านขายวัสดุก่อสร้างแห่งหนึ่งต้องการจะสั่งสต๊อกของไว้เพื่อขายในช่วงน้ำท่วม โดยที่เขาทราบจากข้อมูลในอดีตว่า ความต้องการของลูกค้าในช่วงเวลาดังกล่าวมีการแจกแจงแบบปกติ ที่มีค่าความแปรปรวนเป็น 10 และค่าเฉลี่ยเป็นค่าใดค่าหนึ่งต่อไปนี้คือ 5, 10 หรือ 15 ดังนั้นการสต๊อกสินค้าของเขาจึงมี 3 ระดับ คือ ระดับต่ำ ระดับกลาง และระดับสูง

ให้ X = ความต้องการวัสดุของลูกค้า
θ = ค่าเฉลี่ยของตัวแปรเชิงสุ่ม X

ให้ prior distribution ของ θ มีค่าดังนี้

$$P(\theta=\theta_i) = \begin{cases} 0.4 & \theta=\theta_1=5 \\ 0.5 & \theta=\theta_2=10 \\ 0.1 & \theta=\theta_3=15 \end{cases}$$

$$Q_X(x|\theta = \theta_j) = \frac{1}{\sqrt{2\pi}\sqrt{10}} e^{-1/2[(x-\theta_j)/\sqrt{10}]^2} \quad \text{for } \theta_j = 5, 10, 15$$

ดังนั้น codition distribution of x given θ ก็คือ prior distribution of θ ก็คือ

BAYES DECISION WITHOUT DATA
STATES OF NATURE (MEAN DEMAND)

Actions: volume to stock	θ_1 $P_\theta(5) = 0.2$	θ_2 $P_\theta(10) = 0.5$	θ_3 $P_\theta(15) = 0.3$	$\underset{\theta}{E}[l(a_i,\theta)]$
a_1 : high	−8	−12	−20	−13.6
a_2 : medium	−10	−14	−14	−13.2
a_3 : low	−12	−12	−12	−12.0
			$\underset{A}{\min}\left\{\underset{\theta}{E}[l(A,\theta)]\right\} =$	−13.6

เนื่องจากว่า

$$h_\theta(\theta_j|X = x) = \frac{Q_X(x|\theta = \theta_j)P_\theta(\theta_j)}{\sum_{k=1}^{3} Q_X(x|\theta = \theta_k)P_\theta(\theta_k)}$$

ดังนั้นจะได้ว่า

$$h_\theta(\theta_j|X = 10) = \frac{\frac{1}{\sqrt{2\pi}\sqrt{10}} e^{-1/2[(10-\theta_j)/\sqrt{10}]^2} P_\theta(\theta_j)}{\sum_{k=1}^{3}\left\{\frac{1}{\sqrt{2\pi}\sqrt{10}} e^{-1/2[(10-\theta_k)/\sqrt{10}]^2} P_\theta(\theta_k)\right\}}$$

For $\theta = \theta_1 = 5$,

$$h_\theta(5|X=10) = \frac{\frac{1}{\sqrt{2\pi}\sqrt{10}} e^{-1/2[(10-5)/\sqrt{10}]^2} P_\theta(5)}{\frac{1}{\sqrt{2\pi}\sqrt{10}} \sum_{k=1}^{3} \left\{ e^{-1/2[(10-\theta_k)\sqrt{10}]^2} P_\theta(\theta_k) \right\}}$$

$$= \frac{(0.2)e^{-1/2(2.5)}}{(0.2)e^{-1/2(2.5)} + (0.5) + (0.3)e^{-1/2(2.5)}}$$

$$= 0.089$$

For $\theta = \theta_2 = 10$,

$$h_\theta(10|X=10) = \frac{0.5}{(0.2)e^{-1/2(2.5)} + 0.5 + (0.3)e^{1/2(2.5)}}$$

$$= 0.777$$

For $\theta = \theta_3 = 15$,

$$h_\theta(15|X=10) = \frac{(0.3)e^{-½(2.5)}}{(0.2)e^{-½(2.5)} + 0.5 + (0.3)e^{-½(2.5)}}$$

$$= 0.134$$

The posterior distribution of θ, given $X = 10$, is then

$$h_\theta(\theta_j|X=10) = \begin{cases} 0.089 & \theta = 5 \\ 0.777 & \theta = 10 \\ 0.134 & \theta = 15 \end{cases}$$

BAYES ANALYSIS WITH DATA

Actions: volume to stock	$\theta_1 = 5$ $h_\theta(5\|X=10) = 0.089$	$\theta_2 = 10$ $h_\theta(10\|X=10) = 0.777$	$\theta_3 = 15$ $h_\theta(15\|X=10) = 0.134$	$E_\theta[l(a_i,\theta)]$
a_1: high	−8	−12	−20	−12.716
a_2: medium	−10	−14	−14	−13.644
a_3: low	−12	−12	−12	−12.000

$\min_A \left\{ E_\theta[l(A,\theta)] \right\} = -13.644$

ดังนั้น action ที่เลือกก็คือ a_2

ข้อสังเกตุ

1. ถ้าเราใช้ Squared error loss function กับวิธีการสร้างค่าประมาณโดยอาศัยวิธีการของเบย์ จะได้ว่า $\hat{\theta}$ คือ Expected Bayes risk

$$v(\theta;x_1,\ldots,x_n) = \int_{-\infty}^{\infty} (\theta-\hat{\theta})^2\, h(\theta/x_1,\ldots,x_n)\, d\theta$$

$$\frac{\partial v}{\partial \hat{\theta}}(\theta;x_1,\ldots,x_n) = 2\int_{-\infty}^{\infty} (\theta-\hat{\theta})\, h(\theta/x_1,\ldots,x_n)\, d\theta = 0$$

$$\hat{\theta} = \int_{-\infty}^{\infty} (\theta h(\theta/x_1,\ldots,x_n)\, d\theta$$

2. ถ้าใช้ Absoluted error loss function ค่าของ Bayes estimator ก็จะเป็นมัธยฐาน (Median) ของ Posteriori function (ให้ลองทำเป็นแบบฝึกหัด โดยอาศัยแนวทางที่ว่า ฟังก์ชัน $\Sigma|x_i-a|$ จะมีค่าต่ำสุดก็ต่อเมื่อ a คือ มัธยฐาน)

นั่นก็คือ

$$\int_{-\infty}^{\hat{\theta}} h(\theta/x_1,x_2,\ldots,x_n)\, d\theta = \frac{1}{2}$$

โดยที่ $\hat{\theta}$ ที่จะหา ก็คือ มัธยฐานของ $h(\theta/x_1,x_2,\ldots,x_n)$ ในปัญหานี้ ถ้าเราสามารถจะปรับ $h(\theta/x_1,x_2,\ldots,x_n)$ ให้เป็นรูปการแจกแจงปกติได้ การหา $\hat{\theta}$ ก็คงจะไม่ลำบากนัก แต่ถ้าเราปรับไม่ได้ก็คงจะต้องใช้คอมพิวเตอร์ช่วย search หา $\hat{\theta}$ ที่ตรงตามเงื่อนไขนี้

โดยวิธีการสร้างให้เป็นสมการกำลังสอง จำนวนที่อยู่ภายใต้เครื่องหมายอินทรีกัลจะได้ว่าผลที่ได้ของ $k(x_1, x_2,\ldots,x_n)$ จะเป็น

$$\frac{1}{(2\pi)^{n/2}} \exp\left[-\tfrac{1}{2}(\Sigma x_i^2 - \frac{n^2\bar{x}^2}{n+1})\right] \{ \frac{1}{(2\pi)^{1/2}}$$

$$\int_{-\infty}^{\infty} \exp\left[-\tfrac{1}{2}(n+1)(\mu - \frac{n\bar{x}}{n+1})^2\right] d\mu$$

$$k(x_1,x_2,..,x_n) = \frac{1}{(n+1)^{1/2}(2\pi)^{n/2}} \exp\left[-\tfrac{1}{2}(\Sigma x_i^2 - \frac{n^2\,\bar{x}^2}{n+1})\right]$$

$$h(\mu/x_1,x_2,\ldots,x_n) = \frac{(2\pi)^{-(n+1)/2}\exp\{-\tfrac{1}{2}[\Sigma x_i^2 + (n+1)\mu^2 - 2n\,\bar{x}\mu]\}}{(2\pi)^{-(n/1)}(n+1)^{-1/2}\exp\left[-\tfrac{1}{2}(\Sigma x_i^2 - \frac{n^2\,\bar{x}^2}{(n+1)}\right]}$$

$$= \frac{(n+1)^{1/2}}{(2\pi)^{1/2}} \exp\{-\tfrac{1}{2}(n+1)\,[\mu - \frac{n\,\bar{x}\mu}{n+1}]^2\}$$

จะเห็นได้ว่า Conditional distribution of μ, given $x_1,\ldots, x_n$

มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็น $n\,\bar{x}\mu/(n+1)$ และความแปรปรวนเป็น $(n+1)^{-1}$ กำหนดให้ใช้ Squared-error loss function จะได้ว่า ค่าประมาณโดยวิธีการของเบย์ ก็คือค่าของ $\hat{\theta}$ ซึ่งเป็นฟังก์ชั่นของ $x_1,\ldots,x_n$ ซึ่ง Minimizes $v(\hat{\mu}\,;x_1,\ldots, x_n)$

$$v(\hat{\mu}\,;x_1,\ldots,x_n) = \int_{-\infty}^{\infty} (\hat{\mu} - \mu)^2.h\,(\mu/x_1,\ldots, x_n)\,d\mu$$

$$= \frac{(n+1)^{1/2}}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} (m - \mu)^2 \exp\{-\tfrac{1}{2}(n+1)\left(\mu - \frac{n\,\bar{x}}{n+1}\right)^2\}d\mu$$

$$= \mu^2 - \frac{2\mu\,\bar{x}\,n}{n+1} + \frac{1}{n+1} + \frac{\bar{x}^2\,n^2}{(n+1)^2}$$

$$\partial\theta/\partial\,\hat{\mu} \to 2.\hat{\mu} - \frac{2\,\bar{x}\,n}{(n+1)} = 0$$

$$\hat{\mu}_{Bayes} = \Sigma X_i/(n+1)$$

อัลกอลิทึมของวิธีการตัดสินใจโดยวิธีการของเบย์ (Algorithm - Bays Decision Procedure with Data)

Assume the distributions

$$P_\theta(\theta_j) = \text{prior distribution of } \theta \qquad j = 1, 2, \ldots, n$$
$$Q_X(x|\theta = \theta_j) = \text{conditional distribution of } X \text{ given } \theta = \theta_j$$

Step 1

Determine the posterior distribution of θ. That is, let

$$h_\theta(\theta_j|X = x) = \frac{Q_X(x|\theta = \theta_j)P_\theta(\theta_j)}{\sum_{k=1}^{n} [Q_X(x|\theta = \theta_k)P_\theta(\theta_k)]}$$

Step 2

Obtain a reliable value of the random variable X, say x^*

Step 3

Use the posterior distribution $h_\theta(\theta_j|X = x^*)$ to calculate the expected loss for each action a_i with $i = 1, 2, \ldots, m$.

$$\underset{\theta}{E}[l(a_i,\theta)] = \sum_{j=1}^{n} h_\theta(\theta_j|X = x^*)l(a_i, \theta_j) \qquad i = 1, 2, \ldots, m$$

Step 4

Select the action that has the smallest expected loss. Stop.

จากอัลกอลิทึมที่จะตัดสินใจ โดยการใช้เบย์

เรานำมาเขียนเป็นโปรแกรมได้ดังนี้คือ

หมายเหตุ จะขอยกตัวอย่างฟังก์ชั่นที่จะใช้บางอันดังนี้คือ

$$Q_X(x|\theta = \theta_j) = \frac{1}{\sqrt{2\pi}\sqrt{10}} e^{-1/2[(x-\theta_j)/\sqrt{10}]^2}$$

ซึ่งเขียนเป็นคำสั่งได้ดังนี้คือ

```
Q(X,T)=EXP(-((X-T)/SQRT(10))**2/2)/SQRT(20*3.1416)
```

โดยที่ T คือ ค่าของตัวแปรเชิงสุ่ม θ

โปรแกรมนี้จะใช้ได้ถึง 40 action และ 50 state of nature โดยที่ M จะมีความหมายถึงจำนวน action และ N ก็คือ state of nature

ดังนั้น ถ้าเราจะปรับปรุงโปรแกรมนี้เพื่อขยาย action และ state of nature ให้มีมากขึ้น เราก็เพียงแต่เปลี่ยนคำสั่งที่เกี่ยวข้องดังต่อไปนี้ .

```
DIMENSION A(M,N),P(N),EL(M),HTHETA(N),ELPOST(M)
INTEGER XSTAR,THETA(N),TJ,X,T
```

โปรแกรมนี้จะใช้ที่ถึง 48 k bytes

```
C       ***********************************************************************
C       *                                                                     *
C       *         ***  ALGORITHMS  10.1 AND 10.2  ***                         *
C       *                                                                     *
C       *       BAYES DECISION PROCEDURE WITH AND WITHOUT DATA                *
C       *                                                                     *
C       * THIS IS A GENERAL COMPUTER PROGRAM TO CARRY CUT THE BAYES DECISION  *
C       * PROCEDURE WITH AND WITHOUT DATA.                                    *
C       *                                                                     *
C       * THIS PROGRAM IS SET UP FOR A MAXIMUN OF 40 COURSES OF ACTION AND 50 *
C       * STATES OF NATURE.  TO MODIFY THE PROGRAM TO HANDLE LARGER PROBLEMS  *
C       * WITH  M  COURSES OF ACTION AND  N  STATES OF NATURE, CHANGE THE     *
C       * DIMENSION AND INTEGER STATEMENTS TO                                 *
C       *           DIMENSION A(M,N),P(N),EL(M),ELPOST(M),HTHETA(N)           *
C       *           INTEGER XSTAR, THETA(N), TJ, X, T                         *
C       *                                                                     *
C       * THE PROGRAM,AS WRITTEN, OCCUPIES 48 K BYTES OF CORE STORAGE.        *
C       *                                                                     *
C       * IT IS DESIGNED                                                      *
C       *                                                                     *
C       *    TO READ                                                          *
C       *       CARD 1   COLS  2-80   TITLE  DESCRIPTION OF THE PROBLEM USING *
C       *                                    ANY CHARACTERS ON KEYPUNCH       *
C       *                                    ** COLUMN 1 MUST BE LEFT BLANK **  *
C       *       CARD 2   COLS  1- 5   NFLAG  FLAG TO INDICATE PRESENCE OR     *
C       *                                    ABSENCE OF DATA:  0  WITHOUT DATA *
C       *                                                      1  WITH DATA (I5) *
C       *       CARD 3   COLS  1- 5   M  # OF COURSES OF ACTION   (I5)        *
C       *                      6-10   N  # OF STATES OF NATURE   (I5)         *
C       *       CARDS 4 TO T     A(I,J)  LOSS MATRIX                          *
C       *                                READ DATA ROWWISE IN 8F10.0 FORMAT   *
C       *                                IF  N>8, CONTINUE ON A NEW CARD      *
C       *                                START EACH NEW ROW ON A NEW CARD     *
C       *       CARD T+1    P(J)  PRIOR DISTRIBUTION OF THE STATES OF NATURE  *
C       *                         READ IN 8F10.0 FORMAT.  IF  N>8, CONTINUE   *
C       *                         ON A NEW CARD.                              *
C       *       CARD T+2    THETA(J)  STATES OF NATURE  (8I10)                *
C       *                             IF  N>8, CONTINUE ON A NEW CARD         *
C       *       CARD T+3    COLS  1-10   XSTAR  SAMPLE VALUE OF X   (I10)     *
C       *            TO SOLVE MORE THAN ONE PROBLEM AT A TIME, REPEAT THE     *
C       *            READ SEQUENCE, AND STACK THE DATA ONE BEHIND THE OTHER   *
C       *                                                                     *
C       *    TO CALCULATE                                                     *
C       *       HTHETA(J)   POSTERIOR DISTRIBUTION OF THE STATE OF NATURE     *
C       *                   THETA(J)    (J=1,2,...,N)                         *
C       *       EL(I)       EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON  *
C       *                   PRIOR DISTRIBUTION  (I=1,2,...,M)                 *
C       *       SMALL       MINIMUM OF EL(I)                                  *
C       *       INDEX       ACTICN CORRESPONDING TO SMALL                     *
C       *       ELPOST(I)   EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON  *
C       *                   POSTERIOR DISTRIBUTION  (I=1,2,...,M)             *
C       *       SMPOST      MINIMUM OF ELPOST(I)                              *
C       *       INPOST      ACTION CORRESPONDING TO SMPOST                    *
C       *                                                                     *
C       *    TO PRINT                                                         *
C       *       ECHO CHECK OF DATA                                            *
C       *       EL(I), SMALL, INDEX                                           *
C       *       ELPOST(I), SMPOST, INPOST                                     *
C       *       HTHETA(J)                                                     *
C       *                                                                     *
C       ***********************************************************************
      DIMENSION A(40,50),P(50),EL(40),HTHETA(50),ELPOST(40)
      INTEGER XSTAR,THETA(50),TJ,X,T
      REAL*4 TITLE(20)
C DEFINE STATEMENT FUNCTION TO BE USED LATER
      Q(X,T)=EXP(-((X-T)/SQRT(10.))**2/2)/SQRT(20.*3.1416)
C  READ IN DATA
    5 READ(5,10,END=200)TITLE
   10 FORMAT(20A4)
      WRITE(6,6)TITLE
    6 FORMAT('1',20A4,//)
```

```
      READ(5,15)NFLAG
   15 FORMAT(2I5)
      READ(5,15)M,N
      DO 20 I=1,M
   20 READ(5,25)(A(I,J),J=1,N)
   25 FORMAT(8F10.0)
      READ(5,25)(P(J),J=1,N)
      IF(NFLAG.EQ.0) GO TO 99
      READ(5,130)(THETA(J),J=1,N)
      READ(5,130)XSTAR
  130 FORMAT(8I10)
   99 WRITE(6,100)
  100 FORMAT(9X,'STATES OF NATURE'/' ACTIONS'/)
      DO 105 I=1,M
  105 WRITE(6,110)(A(I,J),J=1,N)
  110 FORMAT( 8F10.4)
      WRITE(6,111)(P(J),J=1,N)
  111 FORMAT( //'    PRIOR DISTRIBUTION OF STATES OF NATURE',/  8F10.4)
C     ******************************************************************
C     *  STEPS 1,2                                                     *
C     *         CALCULATE THE EXPECTED LOSS FOR EACH ACTION BY USING THE *
C     *           PRIOR DISTRIBUTIONS OF THETA.                        *
C     *         SELECT THE ACTION HAVING THE SMALLEST EXPECTED LOSS    *
C     ******************************************************************
      SMALL=10.E10
      DO 30 I=1,M
      EL(I)=0.0
      DO 35 J=1,N
   35 EL(I)=EL(I) + P(J)*A(I,J)
      IF(EL(I).GE.SMALL) GO TO 30
      SMALL=EL(I)
      INDEX=I
   30 CONTINUE
C     ******************************************************************
C     *  STEP 3                                                        *
C     *         IF PROBLEM HAS DATA, GO TO STEP 4.  OTHERWISE, PROBLEM IS *
C     *           COMPLETED, SO PRINT RESULTS                          *
C     ******************************************************************
      IF(NFLAG.EQ.0) GO TO 101
C     ******************************************************************
C     *  STEP 4     (STEP 1 - ALGORITHM 10.2)                          *
C     *         CALCULATE THE POSTERIOR DISTRIBUTION OF THE STATES OF NATURE *
C     ******************************************************************
      GX=0.0
      DO 150 J=1,N
      TJ=THETA(J)
  150 GX=GX+Q(XSTAR,TJ)*P(J)
      DO 160 J=1,N
      TJ=THETA(J)
  160 HTHETA(J)=Q(XSTAR,TJ)*P(J)/GX
      WRITE(6,112)XSTAR
  112 FORMAT( //'    SAMPLE VALUE OF X IS ',I10)
      WRITE(6,113)(THETA(J),J=1,N)
  113 FORMAT( // '    STATES OF NATURE', / 8I10)
      WRITE(6,115)(HTHETA(J),J=1,N)
  115 FORMAT(//'    POSTERIOR DISTRIBUTION OF STATES OF NATURE',/8F10.4)
C     ******************************************************************
C     *  STEP 5     (STEP 3,4 - ALGORITHM 10.2)                        *
C     *         CALCULATE THE EXPECTED LOSS FOR EACH ACTION USING THE  *
C     *           POSTERIOR DISTRIBUTION                               *
C     *         SELECT ACTION HAVING SMALLEST EXPECTED LOSS            *
C     ******************************************************************
      SMPOST=10.E10
      DO 170 I=1,M
      ELPOST(I)=0.0
      DO 175 J=1,N
  175 ELPOST(I)=ELPOST(I)+HTHETA(J)*A(I,J)
      IF(ELPOST(I).GE.SMPOST) GO TO 170
      SMPOST=ELPOST(I)
      INPOST=I
  170 CONTINUE
```

```
C       ***********************************************************************
C       *  STEP 6     PRINT RESULTS                                            *
C       ***********************************************************************
  101 WRITE(6,114)
  114 FORMAT(///' RESULTS'/'+',    7('_')//)
      WRITE(6,40)(EL(I),I=1,M)
   40 FORMAT('    EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON PRIOR
     *DISTRIBUTION',/  8F10.4)
      WRITE(6,45)SMALL,INDEX
   45 FORMAT(//10X,'THE MINIMUM EXPECTED LOSS USING THE PRIOR DISTRIBUTI
     *ON OF'/10X,'THE STATE OF NATURE IS',F12.6//14X,'CHOOSE ACTION A(',
     *I2,')')
      IF(NFLAG.EQ.0) GO TO 5
      WRITE(6,180)(ELPOST(I),I=1,M)
  180 FORMAT( //'    EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON THE
     * POSTERIOR DISTRIBUTION',/  8F10.4)
      WRITE(6,185)SMPOST,INPCST
  185 FORMAT(//10X,'THE MINIMUM EXPECTED LOSS USING THE POSTERIOR DISTRI
     *BUTION'/10X,'OF THE STATE OF NATURE IS',F12.6,//14X,'CHOOSE ACTION
     * A(',I2,')')
      GO TO 5
  200 STOP
      END

/DATA

 EXAMPLE 10.3.1
    0
    4     3
  4.          7.          3.
  5.          2.          4.
  8.          6.          10.
  3.          1.          9.
  0.5       0.333330    0.16667
 EXAMPLES 10.3.2 AND 10.4.1
    1
    3     3
  -8.         -12.        -20.
  -10.        -14.        -14.
  -12.        -12.        -12.
  0.2          0.5         0.3
        5            10           15
       10
```

```
 EXAMPLE 10.3.1

           STATES OF NATURE
ACTIONS

    4.0000     7.0000     3.0000
    5.0000     2.0000     4.0000
    8.0000     6.0000    10.0000
    3.0000     1.0000     9.0000

  PRIOR DISTRIBUTION OF STATES OF NATURE
   0.5000     0.3333     0.1667
```

RESULTS

```
 EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON PRIOR DISTRIBUTION
  4.8333    3.8333    7.6667    3.3334

        THE MINIMUM EXPECTED LOSS USING THE PRIOR DISTRIBUTION OF
        THE STATE OF NATURE IS    3.333359

            CHOOSE ACTION A( 4)
```

```
 EXAMPLES 10.3.2 AND 10.4.1

        STATES OF NATURE
ACTIONS

  -8.0000  -12.0000  -20.0000
 -10.0000  -14.0000  -14.0000
 -12.0000  -12.0000  -12.0000

  PRIOR DISTRIBUTION OF STATES OF NATURE
   0.2000    0.5000    0.3000

  SAMPLE VALUE OF X IS         10

  STATES OF NATURE
       5        10        15

  POSTERIOR DISTRIBUTION OF STATES OF NATURE
   0.0891    0.7773    0.1336
```

RESULTS

```
 EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON PRIOR DISTRIBUTION
-13.6000  -13.2000  -12.0000

        THE MINIMUM EXPECTED LOSS USING THE PRIOR DISTRIBUTION OF
        THE STATE OF NATURE IS  -13.599999

            CHOOSE ACTION A( 1)

 EXPECTED LOSS FOR EACH COURSE OF ACTION BASED ON THE POSTERIOR DISTRIBUTION
-12.7126  -13.6437  -12.0000

        THE MINIMUM EXPECTED LOSS USING THE POSTERIOR DISTRIBUTION
        OF THE STATE OF NATURE IS  -13.643685

            CHOOSE ACTION A( 2)
```

แบบฝึกหัด

1. Consider the decision problem with the loss table below. What is the minimax action?

A: actions	θ: states of nature θ_1	θ_2	θ_3
a_1	5	3	5
a_2	8	6	3
a_3	9	4	2
a_4	8	3	4
a_5	7	1	10

2. Suppose a decision maker has obtained the following loss table:

A: actions	θ: states of nature θ_1	θ_2	θ_3
a_1	3	2	1
a_2	2	1	9
a_3	6	3	7

(a) What is the minimax action?
(b) If the prior distribution of θ is

$$P(\theta = \theta_j) = \begin{cases} 0.4 & \theta = \theta_1 \\ 0.5 & \theta = \theta_2 \\ 0.1 & \theta = \theta_3 \end{cases}$$

What is the Bayes decision

3. The first thing each morning a college student must make a decision; should he roll over and stay in bed, get up and go to class with an umbrella, or get up and go to class without an umbrella. The student has determined the loss corresponding to each decision and state of nature to be:

	States of nature θ_1: rain	θ_2: no rain
a_1: skip class	4	4
a_2: go–take umbrella	2	5
a_3: go–do not take umbrella	5	0

The local weather report has predicted rain with probability 0.7.
(a) What is the minimax action?
(b) What is the Bayes decision?
(c) Suppose the weather report revises the forecast to indicate that the probability of rain is 0.6. What is the Bayes decision?

4. In problem 1 if the priar distribution of the mean demand equipment is changed to

$$P_\theta(\theta_j) = \begin{cases} 0.2 & \theta = \theta_1 = 5 \\ 0.5 & \theta = \theta_2 = 10 \\ 0.3 & \theta = \theta_3 = 15 \end{cases}$$

What is the Bays

6.4 เปรียบเทียบวิธีการของเบย์กับวิธีธรรมดา

(Baysian Versus Classical Estimation)

เพื่อเป็นการเปรียบเทียบวิธีการประมาณค่าโดยวิธีของเบย์ และวิธี Classical Estimation ให้พิจารณาเปรียบเทียบจากตัวอย่างต่อไปนี้

ตัวอย่างที่ 1 การศึกษาถึงความยาวของแมลงหวี่ โดยการหาค่าประมาณของ θ เราใช้เครื่องมือวัดความยาว ของแมลงหวี่ที่สุ่มมา 10 ตัว ได้ $X_1,\ldots,X_{10}$ ถ้าหากเครื่องมือมีความคลาดเคลื่อน โดยที่ X ที่วัดได้ มีการแจกแจงแบบปกติ มีค่าเฉลี่ย θ โดยมีความแปรปรวน $\sigma^2 = 10$ ถ้าสุ่มแมลงหวี่มา 10 ตัว และค่าเฉลี่ย $\bar{X} = 20$

ดังนั้น Classical 95% Congidence Interval สำหรับ θ คือ

$$\theta = \bar{X} \pm 1.96\sigma/\sqrt{n}$$

$$= 20 \pm 1.96$$

และ Point estimate ของ $\theta = \bar{X} = 20$

ตัวอย่างที่ 2 นักชีววิทยาได้วิจัยพบว่าความยาวของแมลงหวี่มีการแจกแจงแบบปกติ โดยที่ค่าเฉลี่ย $\sigma_0 = 25$ และความแปรปรวน $\sigma_0^2 = 4$ นั่นแสดงว่า Prior Distribution ของจะเป็นการแจกแจงปกติ θ ° $N(\theta_0, \sigma^2)$

ดังนั้นถ้าสุ่มตัวอย่างแมลงหวี่มากลุ่มหนึ่งแล้ววัดความยาว จะได้ $\bar{X}$ โดยที่ $p(\bar{X}/\theta) = N(\theta, \sigma^2/n)$ จะได้ Posterior Distribution คือ

$$p(\theta/\bar{x}) = N\,\frac{(w_1\bar{x} + w_2\theta_0,}{w_1+w_2}\;\frac{1}{w_1+w_2}\,)$$

$$w_1 = \frac{1}{\sigma^2/n} \qquad w_2 = 1/\sigma_0^2$$

$\sigma^2 = 10, n = 10, \sigma_0^2 = 4$

$w_1 = 10/10 = 1, w_2 = 1/4$

$$\frac{w_1\bar{x}+w_2\theta_2}{w_1+w_2} = \frac{1(20) + 1/4(25)}{1+1/4} = \mathbf{21}$$

$\Rightarrow \ p(\theta/\bar{x}) = n(21, 0.8)$

ถ้ากำหนดเพิ่มเติมว่า ให้ loss function เป็น quadratic loss ค่าของ Baysian estimate จะเป็นเท่าไร (ใช้ช่วงเชื่อมั่น 95%)

เนื่องจากว่า $p(\theta/\bar{x})$ มีการแจกแจงแบบปกติ

โดยมีค่าเฉลี่ยเป็น 21 ความแปรปรวน .8 ด้วยความเชื่อมั่น 95% θ จะอยู่ในช่วง

$= \ 21 \pm 1.96\sqrt{8}$

$= \ 21 \pm 1.76$

ผลที่ได้รับนี้ คือ Baysian estimate ไปเปรียบเทียบกับ Classical estimate (20 ± 1.96)

จะเห็นว่าผลที่ได้ของ Baysian estimate ดีกว่า Classical estimate ในแง่ที่ว่าให้ความถูกต้องมากกว่า

ทั้งนี้เพราะว่า ความแปรปรวนของ Posterior dist^{n} จะน้อยกว่า dist^{n} ของ $\bar{X}$ แบบธรรมดา

พิจารณา Posterior variance ให้ $a = 1(w_1 + w_2)$

$=> \ 1/a = \{(1/(\sigma^2/n)\} + 1/\sigma_0^2$

เนื่องจาก $\sigma_0^2 \neq \infty \Rightarrow \ 1/\sigma_0^2 \neq 0$

$=> \ 1/a > 1/(\sigma^2/n)$

$a < \sigma^2/n$

6.5 ข้อได้เปรียบและเสียเปรียบในการใช้วิธีการของเบย์

(Critique of Baysian Methods)

ข้อดีของ Bayes' estimate ก็คือเป็นวิธีการทางสถิติที่ใช้ประโยชน์สูงสุดจากข้อมูลที่ได้ ในลักษณะของการพยายามให้ฟังก์ชั่นของความผิดพลาด (Loss function) มีค่าต่ำที่สุด แต่มีข้อแม้ว่าวิธีการต้องอาศัยความรู้ในเรื่องการแจกแจงของตัวพารามิเตอร์ (Prior distribution) และ Loss function เข้าช่วย

สำหรับผลที่ได้รับจากการประมาณค่าด้วยวิธีการของเบย์ให้ช่วงประมาณสั้นกว่าและน่าเชื่อถือว่าถูกต้องกว่าด้วยวิธีการของ Classical Method อีกทั้งเวลานำไปทดสอบก็สะดวกและเหมาะสมกว่า

วิธีการของเบย์เหมาะกับงานทางสังคมศาสตร์และธุรกิจ ทั้งนี้เนื่องจากงานทางสองด้านดังกล่าวมีจำนวนกลุ่มตัวอย่างค่อนข้างเล็ก

ข้อเสียของวิธีการของเบย์

โดยปกติ Loss function และ Prior distⁿ ของพารามิเตอร์ มักจะไม่ค่อยทราบซึ่งมักจะต้องเดาเอาโดยอาศัยข้อมูลในประสบการณ์ก่อน ๆ เป็นเครื่องช่วย และถ้าสร้าง Loss function หรือ Prior distⁿ ผิดพลาดก็จะยังผลให้ Baysian estiamte ผิดพลาดตามไปด้วย แต่ในวิธีการของ Classical Method ไม่ต้องอ้างอิงถึงข้อจำกัดใดที่กล่าวมาในเรื่องของเบย์เลย

ปัญหา

1. Let μ be the true I.Q. of a certain student. To measure his I.Q. he takes a test, and it is known that his test scores are normally distributed with mean μ and standard deviated 5. This student takes the I.Q. test and get a score of 130. What is the maximum-likelihood estimator of μ.

2. In prob. 2, suppose that it is known that the true I.Q.'s of students of a certain age are distributed normally with mean 100 andvariance 225; ie., assume that is distributed normally with mean 100 and variance 225. Thus $f(/\mu)$ in prob. is $n(x;\mu, 25)$ and $p(\mu)$ is $n(\ ;100,225)$. Find $q(x,\mu)$ and $k(x)$.

3. In prob. 2 show that the conditional density $h(\mu/x)$ is normally with mean $.9x+10$ and variance 45/2.

   Using the loss $1(\mu;\mu) = (\mu-\mu)^2$ in prob. 3, find the Bayes estimator of μ, the student's I.Q., if the student's test score is $x=130$. Note that it is not the same as the maximum-likelihood estimator in prob. 1.

4. The fraction defective in a day's production of a certain product is θ. Let X be an observation on one of the items of a given day's production. The distribution of X is

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

where $X = 1$ is identified with a defective item and $X = 0$ is identified with the item if it is not defective. While the proportion defective remains constant for a given day, it is noticed that θ varies from day to day and θ acts as a random variable with density function

$$p(\theta) \quad = \quad 6\theta(1-\theta);\ 0<\theta<1$$

Find the estimator of θ by using Bayes' Criterion Given the Squared loss function