Conditional Expectation

รศ.ดร. วีระชาติ กิเลนทอง
มหาวิทยาลัยหอการค้าไทย

ค่าคาดหมายแบบมีเงื่อนไข (Conditional Expectation)

- การคาดการณ์ (prediction) หรือการประมาณค่า (estimation) ที่ได้รับความสนใจมักจะอยู่ในรูปของค่าคาดหมายแบบมีเงื่อนไข (conditional expectation)
  - การคาดการณ์ (prediction) หรือการประมาณค่า (estimation) มักจะต้องเริ่มจากการที่ "**รู้อะไรบางอย่าง**" แล้วจึงพยายามที่จะบอกว่าจะเกิดอะไรขึ้นกับตัวแปรหรือสิ่งที่สนใจ
  - ยกตัวอย่างเช่น
    - รายได้ของครัวเรือนที่หัวหน้าครัวเรือนเรียนจบปริญญาตรีแตกต่างจากครัวเรือนที่หัวหน้าครัวเรือนจบอาชีวศึกษานั้นมีค่าเท่าใด
    - อัตราผลตอบแทนของกองทุนจะเป็นเท่าใดหากอัตราผลตอบแทนของตลาดเท่ากับ 10 เปอร์เซ็นต์ต่อปี

นิยาม: ค่าคาดหมายแบบมีเงื่อนไข (Conditional Expectation)

Definition

กำหนดให้ X และ Y เป็นตัวแปรสุ่ม โดยที่ค่าคาดหมายของ Y มีค่าจำกัด นั่นคือ $\mu_Y < \infty$ สำหรับกรณีของตัวแปรสุ่มต่อเนื่อง ค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ Y เมื่อทราบค่าของ $X = x$ เท่ากับ

$$E\left[Y|X=x\right] = \int_{-\infty}^{\infty} y f_2\left(y|x\right) dy \tag{1}$$

และค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ X เมื่อทราบค่าของ $Y = y$ เท่ากับ

$$E\left[X|Y=y\right] = \int_{-\infty}^{\infty} x f_1\left(x|y\right) dx \tag{2}$$

สำหรับกรณีของตัวแปรสุ่มไม่ต่อเนื่อง ค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ Y เมื่อทราบค่าของ $X = x$ เท่ากับ

$$E\left[Y|X=x\right] = \sum_{y} y f_2\left(y|x\right) dy \tag{3}$$

และค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ X เมื่อทราบค่าของ $Y = y$ เท่ากับ

$$E\left[X|Y=y\right] = \sum_{x} x f_1\left(x|y\right) dx \tag{4}$$

สรุป: ค่าคาดหมายแบบมีเงื่อนไข (Conditional Expectation)

- ค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ Y เมื่อทราบค่าของ $X = x$ ซึ่งแทนด้วย $E[Y|X = x]$ นั่นจะถูกพิจารณาเป็น“**ค่าคงที่ค่าหนึ่ง**”
- แต่หากนำเอาค่าคาดหมายแบบมีเงื่อนไขสำหรับค่าต่างๆ ของ x มารวบรวมเข้าใน “**รูปแบบของการแจกแจง (distribution)**” ซึ่งเขียนในรูป $E[Y|X]$ จะได้ผลลัพธ์ออกมาเป็นตัวแปรสุ่มแบบหนึ่ง
- โดยสรุป
    - $E[Y|X = x]$ เป็นค่าคงที่
    - $E[Y|X]$ เป็นตัวแปรสุ่ม

นิยาม: ค่าคาดหมายแบบมีเงื่อนไข (Conditional Expectation) ของ Y เมื่อทราบ X

Definition

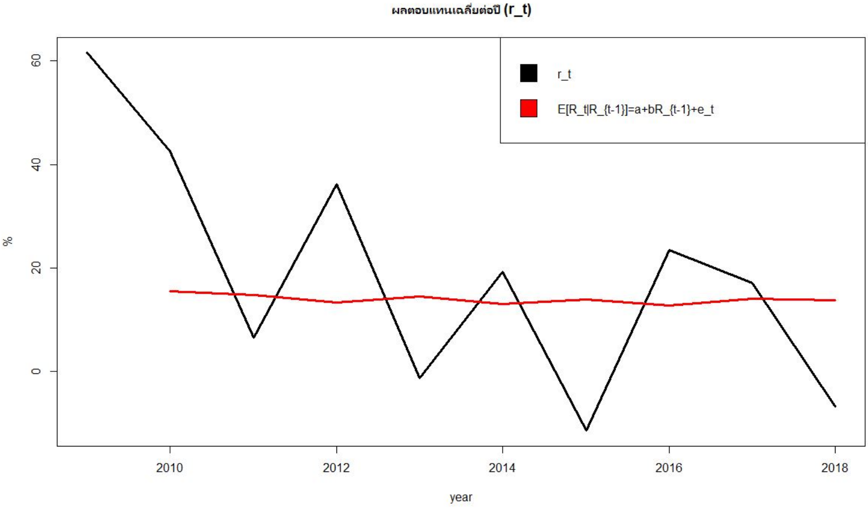
กำหนดให้ $h(x) = E[Y|X = x]$ เป็นฟังก์ชันของ x นิยามให้ $E[Y|X] \equiv h(X)$ เป็น ค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ Y เมื่อทราบ X

- $E[Y|X]$ เป็นฟังก์ชันของตัวแปรสุ่ม X ดังนั้นจึงเป็นตัวแปรสุ่มเช่นกัน

ตัวอย่าง: ค่าคาดหมายแบบมีเงื่อนไขของอัตราผลตอบแทนในตลาดทุน

- ตัวอย่างต่อไปนี้แสดงค่าคาดหมายแบบมีเงื่อนไขของผลตอบแทนของตลาดหลักทรัพย์ในปีปัจจุบันภายใต้เงื่อนไขที่ว่าเราทราบผลตอบแทนการลงทุนในปีที่ผ่านมา $E[R_t|R_{t-1} = r_{t-1}]$
- เพื่อความสะดวกในการนำเสนอ จึงแบ่งอัตราผลตอบแทนเป็น 10 ปี (2009-2018)
- ประเด็นที่ต้องการให้สังเกตในตัวอย่างนี้คือ การที่ค่าคาดหมายของอัตราผลตอบแทนในปัจจุบันสำหรับแต่ละช่วงอัตราผลตอบแทนในปีที่ผ่านมานั้นเป็นค่าคงที่ค่าหนึ่ง: $E[R_t|R_{t-1} = r_{t-1}]$

กราฟแสดงค่าคาดหมายแบบมีเงื่อนไขของอัตราผลตอบแทนในตลาดทุน



ตัวอย่าง: การประยุกต์ใช้กับการทดลองทางคลีนิค

Example

พิจารณาการทดลองทางคลินิค (clinical trial) โดยกำหนดให้ X_i แทนตัวแปรสุ่มที่บ่งบอกถึงผลลัพธ์ของการรักษาสำหรับผู้ป่วย i นั่นคือ $X_i = 1$ ถ้ารักษาได้สำเร็จ แต่ $X_i = 0$ ถ้าไม่สำเร็จ ในขณะเดียวกัน นักวิจัยยังไม่ทราบว่า ความน่าจะเป็นที่จะรักษาสำเร็จมีค่าเท่าใด จึงต้องพิจารณาให้ความน่าจะเป็นดังกล่าวเป็นตัวแปรสุ่ม ซึ่งแทนด้วย P ในขณะเดียวกัน สมมุติให้ $X_1, \ldots, X_n$ มีการแจกแจงแบบมีเงื่อนไขเมื่อทราบว่า $P = p$ เหมือนกัน นั่นคือ $Pr\left(X_i = 1 | P = p\right) = p$ สำหรับทุก $i = 1, \ldots, n$

- สิ่งที่เราสนใจในตัวอย่างนี้คือ ค่าคาดหมายของจำนวนผู้ป่วยที่จะหายจากอาการป่วยหรือรักษาสำเร็จ ซึ่งแทนด้วย $X = X_1 + \cdots + X_n$
- คุณสมบัติเชิงเส้นของค่าคาดหมายแบบมีเงื่อนไขทำให้สามารถสรุปได้ว่า

$$E\left[X | P = p\right] = \sum_{i=1}^{n} E\left[X_i | P = p\right] = np$$

- เนื่องจาก P เป็นตัวแปรสุ่ม ดังนั้น $E\left[X | P\right] = nP$ เป็นตัวแปรสุ่มด้วย

กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation)

- ทฤษฎีบทต่อไปนี้แสดงวิธีการคำนวณหา ”**ค่าคาดหมายของตัวแปรสุ่มจากค่าคาดหมายแบบมีเงื่อนไข**” โดยอาศัยหลักการที่ว่าค่าคาดหมายแบบมีเงื่อนไข $E[Y|X]$ เป็นตัวแปรสุ่มที่ขึ้นอยู่กับตัวแปรสุ่มที่เป็นเงื่อนไข ซึ่งในที่นี้หมายถึงตัวแปรสุ่ม X
- หากเราดำเนินการหาค่าคาดหมายโดยใช้การแจกแจงของตัวแปรสุ่ม X อีกครั้งหนึ่งก็จะได้ค่าคาดหมายของ X (ไม่ใช่ตัวแปรสุ่มอีกต่อไป)
- หลักการนี้มีประโยชน์อย่างมากในการวิเคราะห์เชิงเศรษฐศาสตร์และการเงิน ซึ่งมักเรียกคุณสมบัตินี้ว่า “**กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation)**”
- ในขณะที่ นักสถิติเรียกทฤษฎีบทนี้ว่ากฎของความน่าจะเป็นรวมสำหรับค่าคาดหมาย (Law of Total Probability for Expectations)

ทฤษฎี: กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation)

Theorem

กำหนดให้ X และ Y เป็นตัวแปรสุ่ม โดยที่ค่าคาดหมายของ Y มีค่าจำกัด นั่นคือ $\mu_Y < \infty$ ดังนั้น

$$E\left[E\left[Y|X\right]\right] = E\left[Y\right] \tag{5}$$

Proof.

เพื่อประหยัดพื้นที่ขอนำเสนอเฉพาะการพิสูจน์สำหรับตัวแปรสุ่มต่อเนื่องดังต่อไปนี้

$$E\left[E\left[Y|X\right]\right] = \int_{-\infty}^{\infty} E\left[Y|x\right] f_1\left(x\right) dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y g_2\left(y|x\right) f_1\left(x\right) dx$$

เนื่องจาก $g_2\left(y|x\right) f_1\left(x\right) = f\left(x, y\right)$ ทำให้สรุปได้ว่า

$$E\left[E\left[Y|X\right]\right] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y f\left(x, y\right) dx = E\left[Y\right]$$

$\square$

ตัวอย่าง: การหาค่าคาดหมายโดยประยุกต์ใช้กฎการหาค่าคาดหมายซ้ำ

Example

สมมุติว่า X และ Y เป็นตัวแปรสุ่มที่มีการแจกแจงเอกรูป (uniform distribution) พิจารณาการสุ่มที่เริ่มจากการสุ่มเลือกค่า x จากการแจกแจงเอกรูปที่มีค่าอยู่ในช่วง $[0,1]$ หลังจากนั้นจึงสุ่มเลือกค่า y จากการแจกแจงเอกรูปที่มีค่าอยู่ในช่วง $[x,1]$ นั่นคือการแจกแจงแบบมีเงื่อนไขของ Y เมื่อทราบค่า $X=x$ เป็นการแจกแจงเอกรูปที่มีค่าอยู่ในช่วง $[x,1]$ คำถามก็คือ ค่าคาดหมาย (expectation) ของ Y มีค่าเท่าใด?

- ค่าคาดหมายแบบมีเงื่อนไขของ Y เมื่อทราบค่า $X=x$ มีค่าเท่ากับ

$$E[Y|X=x]=\frac{1+x}{2}\Rightarrow E[Y|X]=\frac{1+X}{2}$$

- ดังนั้น เมื่อประยุกต์ใช้กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation) เพื่อหาค่าคาดหมาย (expectation) ของ Y จะได้ว่า

$$E[Y]=E[E[Y|X]]=E\left[\frac{1+X}{2}\right]=\frac{1+\frac{1}{2}}{2}=\frac{3}{4}$$

- ทั้งนี้สมการที่สามประยุกต์ใช้ผลลัพธ์ที่ว่าค่าคาดหมายของตัวแปรสุ่มที่มีการแจกแจงเอกรูปในช่วง $[0,1]$ เท่ากับ $E[X]=\frac{1}{2}$

ตัวอย่าง: การหาค่าคาดหมายโดยประยุกต์ใช้กฎการหาค่าคาดหมายซ้ำเมื่อทราบว่า $E[Y|X] = aX + b$

Example

สมมุติว่า $E[Y|X] = aX + b$ สำหรับค่าคงที่ a และ b ใดๆ คำถามคือ ค่าคาดหมาย $E[XY]$ ในรูปของ $E[X]$ และ $E[X^2]$ มีค่าเท่าใด?

- เริ่มจาก เราสามารถพิจารณาให้ตัวแปรสุ่ม X เป็นค่าคงที่ได้ในกรณีของค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ของ Y เมื่อทราบ X ทั้งนี้เพราะเราทราบค่า X แล้ว ดังนั้น

$$E[XY|X] = XE[Y|X] = aX^2 + bX$$

- ประยุกต์ใช้กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation) เพื่อคำนวณหาค่า $E[XY]$

$$E[XY] = E[E[XY|X]] = E[aX^2 + bX] = aE[X^2] + bE[X]$$

ตัวอย่าง: การหาค่าคาดหมายโดยใช้ค่าคาดหมายแบบมีเงื่อนไข

Example

สมมุติให้การทดลองแต่ละครั้งเป็นอิสระต่อกันและมีความน่าจะเป็นที่การทดลองนี้จะสำเร็จในแต่ละครั้งมีค่าเท่ากับ p ถ้าอยากจะทดลองให้สำเร็จติดต่อกัน k ครั้ง คำถามคือ เราน่าจะทดลองกี่ครั้ง

- กำหนดให้ N_k คือจำนวนการทดลองที่จะพบความสำเร็จติดต่อกัน k ครั้ง
- กำหนดให้ M_k คือ ค่าคาดหวังของจำนวนการทดลองที่มีความสำเร็จติดต่อกัน k ครั้ง

$$M_k = E[N_k] = E[E[N_k|N_{k-1}]] \tag{6}$$

- เมื่อกำหนดให้ค่าคาดหวังของจำนวนการทดลองที่จะพบความสำเร็จติดต่อกัน k ครั้ง, (N_k) เมื่อการทดลองก่อนหน้าสำเร็จ คือ

$$\begin{aligned} E[N_k|N_{k-1}] &= p(N_{k-1}+1) + (1-p)(N_{k-1}+1+M_k) \\ E[N_k|N_{k-1}] &= N_{k-1}+1+(1-p)M_k \end{aligned} \tag{7}$$

ตัวอย่าง: การหาค่าคาดหมายโดยใช้ค่าคาดหมายแบบมีเงื่อนไข (ต่อเนื่อง)

Example

- ประยุกต์ใช้กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation) เพื่อคำนวณหาค่าจาก (7)

$$\begin{aligned} E[E[N_k|N_{k-1}]] &= E[N_{k-1} + 1 + (1-p)M_k] \\ M_k &= M_{k-1} + 1 + (1-p)M_k \\ M_k &= \frac{1}{p} + \frac{M_{k-1}}{p} \end{aligned} \tag{8}$$

- เราทราบด้วยว่าค่าคาดหวังของจำนวนการทดลองที่จะพบความสำเร็จ 1 ครั้ง คือ

$$E[N_1] = M_1 = \frac{1}{p} \tag{9}$$

ตัวอย่าง: การหาค่าคาดหมายโดยใช้ค่าคาดหมายแบบมีเงื่อนไข (ต่อเนื่อง)

Example

- เราสามารถหาค่าคาดหวังของจำนวนการทดลอง (8) ที่ทดลองสำเร็จ k ครั้ง

$$\begin{aligned} M_2 &= \frac{1}{p} + \frac{1}{p^2} \\ M_3 &= \frac{1}{p} + \frac{1}{p^2} + \frac{1}{p^3} \\ &\vdots \\ M_k &= \sum_{i=1}^{k} \frac{1}{p^k} = \frac{1-p^k}{p^k(1-p)} \end{aligned} \tag{10}$$

ข้อสังเกตค่าคาดหมายแบบมีเงื่อนไข (Conditional Expectation)

- ค่าคาดหมายแบบมีเงื่อนไขมีความสัมพันธ์แบบเชิงเส้น (linear relationship) และทราบค่าคาดหมาย (mean) μ_X, μ_Y ค่าความแปรปรวน (variance) σ_X^2, σ_Y^2 และค่าสหสัมพันธ์ (correlation) ρ ของตัวแปรสุ่มทั้งสองตัว X และ Y ก็จะสามารถระบุค่าคงที่ a และ b ในรูปค่าสถิติดังกล่าวและเขียนค่าคาดหมายแบบมีเงื่อนไขได้เป็น

$$E\left[Y|X\right] = \mu_Y + \frac{\rho\sigma_Y}{\sigma_X}\left(X - \mu_X\right) \tag{11}$$

- รูปแบบความสัมพันธ์นี้ที่ได้มีความลักษณะคล้ายคลึงกับความสัมพันธ์ที่ได้จากการประมาณค่าแบบกำลังสองน้อยที่สุด (Ordinary Least Square หรือที่เรียกสั้นๆ ว่า OLS) ถัดๆ ไป โดยจะเป็นได้ว่าค่าสัมประสิทธิ์ของ X นั้นขึ้นอยู่กับค่าสหสัมพันธ์เป็นหลัก โดยเฉพาะอย่างยิ่งในการกำหนดว่าค่าจะเป็นบวกหรือลบ
- ส่วนต่อไปนำเสนอความแปรปรวนแบบมีเงื่อนไข (conditional variance) ซึ่งใช้หลักการเดียวกับค่าคาดหมายแบบมีเงื่อนไข หรือในมุมหนึ่งก็คือค่าคาดหมายแบบมีเงื่อนไขของฟังก์ชันของตัวแปรสุ่มนั่นเอง

ค่าความแปรปรวนแบบมีเงื่อนไข (Conditional Variance)

Definition

กำหนดให้ X และ Y เป็นตัวแปรสุ่ม ค่าความแปรปรวนแบบมีเงื่อนไข (conditional variance) ของ Y เมื่อค่าของ $X = x$ เท่ากับ

$$Var[Y|x] = E\left[(Y - E[Y|x])^2 |x\right] \tag{12}$$

- ในทำนองเดียวกันกับค่าคาดหมายแบบมีเงื่อนไข $Var[Y|x]$ เป็นค่าคงที่ แต่ $Var[Y|X]$ เป็นตัวแปรสุ่ม
- ทฤษฎีบทต่อไปนี้แสดงถึงวีธีการแยกส่วนค่าความแปรปรวน (variance decomposition) และ "**ทฤษฎีบทนี้ว่า กฎของความน่าจะเป็นรวมสำหรับค่าความแปรปรวน (Law of Total Probability for Variances)**"

กฎของความน่าจะเป็นรวมสำหรับค่าความแปรปรวน (Law of Total Probability for Variances)

Theorem

สมมุติให้ X และ Y เป็นตัวแปรสุ่มที่มีค่าคาดหมายและค่าความแปรปรวนที่จำกัด แล้ว

$$Var[Y] = E[Var[Y|X]] + Var[E[Y|X]] \tag{13}$$

Proof.

เริ่มจากความแปรปรวนแบบมีเงื่อนไข (conditional variance)

$$Var[Y|X] = E\left[Y^2|X\right] - E[Y|X]^2$$

ซึ่งเป็นตัวแปรสุ่ม ดังนั้น

$$E[Var[Y|X]] = E\left[E\left[Y^2|X\right]\right] - E\left[E[Y|X]^2\right] = E\left[Y^2\right] - E\left[E[Y|X]^2\right]$$

$$Var[E[Y|X]] = E\left[E[Y|X]^2\right] - (E[E[Y|X]])^2 = E\left[E[Y|X]^2\right] - E\left[Y^2\right] \tag{14}$$

$$E[Var[Y|X]] + Var[E[Y|X]] = E\left[Y^2\right] - E\left[Y^2\right] = Var[Y] \tag{15}$$

□

ค่าคาดหมายและค่ามัธยฐานในฐานะตัวทำนาย (Mean and Median as Predictors)

- หัวข้อนี้นำเสนอแนวคิดพื้นฐานของการทำนายในทางสถิติ โดยจะชี้ให้เห็นว่า ค่าสถิติที่แตกต่างกัน เช่น ค่าคาดหมาย (mean) ค่ามัธยฐาน (median) เป็นต้น ต่างสามารถเป็นตัวทำนาย (predictor) ที่เหมาะสมได้ขึ้นอยู่กับเงื่อนไขที่ใช้ในการกำหนดความเหมาะสมหรือจุดประสงค์ที่ต้องการตอบ
    - ค่าคาดหมาย (mean) เป็นตัวทำนายที่เหมาะสมหากจุดประสงค์คือ การทำให้ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (mean square error: m.s.e.) มีค่าต่ำที่สุด
    - ค่ามัธยฐาน (median) เป็นตัวทำนายที่เหมาะสมหากจุดประสงค์คือ การทำให้ค่าคาดหมายของค่าสัมบูรณ์ของค่าคลาดเคลื่อน (mean absolute error: m.a.e.) มีค่าต่ำที่สุด
- บทเรียน: นักศึกษาควรจะเห็นภาพที่ชัดเจนมากขึ้นว่า ทำไมนักสถิติจึงให้ความสนใจกับค่าสถิติบางค่าเป็นพิเศษ เช่น ค่าคาดหมาย (mean) ค่ามัธยฐาน (median) เป็นต้น

ค่าคาดหมายและค่ามัธยฐานในฐานะตัวทำนาย (Mean and Median as Predictors)

- ปัญหาการทำนายที่สนใจในที่นี้ คือต้องการทำนายว่าค่าตัวเลขที่สุ่มได้จากการแจกแจงที่มีค่าคาดหมายเท่ากับ μ และค่าความแปรปรวนเท่ากับ σ^2 จะมีค่าเท่าใด? หรือก็คือ เราต้องการหาค่าตัวทำนายที่เหมาะสมนั่นเอง?
  - คำตอบที่ได้ขึ้นอยู่กับฟังก์ชันจุดประสงค์ (objective function) ที่ใช้ในการกำหนดตัวทำนาย
  - ในที่นี้จะยกตัวอย่างฟังก์ชันจุดประสงค์ 2 แบบ
    - ★ การหาตัวทำนายที่ทำให้ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) มีค่าต่ำที่สุด
    - ★ การหาตัวทำนายที่ทำให้ค่าคาดหมายของค่าสัมบูรณ์ของค่าคลาดเคลื่อน (mean absolute error: m.a.e.) มีค่าต่ำที่สุด

นิยามของค่าคาดหมายและค่ามัธยฐานในฐานะตัวทำนาย (Mean and Median as Predictors)

Definition

กำหนดให้ Y เป็นตัวแปรสุ่ม และ d คือตัวทำนาย (predictor) ดังนั้น ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) ของตัวทำนาย d นิยามได้เป็น $E\left[(Y-d)^2\right]$

- ทฤษฎีบทต่อไปนี้แสดงให้เห็นว่า ค่าคาดหมาย (mean) คือตัวทำนายที่ทำให้ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) มีค่าต่ำที่สุด

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มที่ค่าความแปรปรวนมีค่าจำกัด และกำหนดให้ $\mu = E[Y]$ ดังนั้น สำหรับค่าจำนวนจริง d ใดๆ

$$E\left[(Y-d)^2\right] \geq E\left[(Y-\mu)^2\right] \tag{16}$$

ซึ่งจะเป็นจริงในรูปสมการก็ต่อเมื่อ $d=\mu$

การพิสูจน์ทฤษฎีบทคือตัวทำนายที่ทำให้ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) มีค่าต่ำที่สุด

Proof.

- เราสามารถเขียนค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) ในรูปฟังก์ชันของตัวทำนาย d ได้เป็น

$$E\left[(Y-d)^2\right] = E\left[Y^2 - 2dY + d^2\right] = E\left[Y^2\right] - 2d\mu + d^2$$

  ซึ่งเป็นฟังก์ชันกำลังสอง (quadratic function)

- เราสามารถประยุกต์ใช้หลักการแคลคูลัสพื้นฐานในการหาค่าต่ำสุดได้ด้วยการกำหนดให้ตัวทำนายที่ทำให้ได้ค่าฟังก์ชันนี้มีค่าต่ำสุด d^* คือค่าที่ทำให้ค่าอนุพันธ์อันดับที่หนึ่งมีค่าเท่ากับศูนย์ นั่นคือ

$$\frac{\partial}{\partial d}\left(E\left[Y^2\right] - 2d^*\mu + (d^*)^2\right) = 0 \Rightarrow -2\mu + 2d^* = 0 \Rightarrow d^* = \mu$$

- เนื่องจากฟังก์ชันจุดประสงค์ (objective function) เป็นแบบกำลังสอง ค่าทำให้ได้ค่าต่ำสุดจึงมีค่าเดียว (unique) ผลที่ตามมาก็คือ สำหรับค่าจำนวนจริง d ใดๆ

$$E\left[(Y-d)^2\right] \geq E\left[(Y-\mu)^2\right]$$

□

ค่าคาดหมายและค่ามัธยฐานในฐานะตัวทำนาย (Mean and Median as Predictors)

- ทฤษฎีบทต่อไปนี้แสดงให้เห็นว่า ข้อสรุปแบบเดียวกันนี้เป็นจริงสำหรับกรณีที่ใช้ค่าคาดหมายแบบมีเงื่อนไข (conditional mean) แทนค่าคาดหมายธรรมดา
- สมมุติว่ามีตัวแปรสุ่มอีกตัวหนึ่ง X ซึ่งมีความสัมพันธ์กับตัวแปรสุ่มที่ต้องการทราบหรือทำนายค่า Y และที่สำคัญเนื่องจากทราบค่าของ X ตัวทำนายจึงอยู่ในรูปของฟังก์ชันของ X นั่นคือ $d(X)$
- ตัวทำนายที่ทำให้ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) มีค่าต่ำที่สุดก็คือ " **ค่าคาดหมายแบบมีเงื่อนไข (conditional mean) ของ** Y **เมื่อทราบ** X"

ทฤษฎีที่ว่า ตัวทำนาย (predictor) คือค่าคาดหมายแบบมีเงื่อนไข (conditional mean)

Theorem

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่ค่าความแปรปรวนมีค่าจำกัด ตัวทำนาย (predictor) ของตัวแปรสุ่ม Y เมื่อทราบค่า X ที่ทำให้ค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) $E\left[(Y - d(X))^2\right]$ มีค่าต่ำที่สุด คือ $d^*(X) = E[Y|X]$

Proof.

กำหนดให้ $d(X)$ คือตัวทำนายใดๆ ซึ่งเป็นฟังก์ชันของ X และ $d^*(X) = E[Y|X]$ สิ่งที่ต้องการพิสูจน์คือ

$$E\left[(Y - d^*(X))^2\right] \leq E\left[(Y - d(X))^2\right]$$

ซึ่งทำได้ด้วยการประยุกต์ใช้กฎการหาค่าคาดหมายซ้ำ (Law of Iterative Expectation) ซึ่งทำให้สามารถเขียนค่าคาดหมายของค่าคลาดเคลื่อนกำลังสอง (m.s.e.) สำหรับตัวทำนาย $d(X)$ ใดๆ ในรูปของค่าคาดหมายแบบมีเงื่อนไข (conditional expectation) ได้เป็น

$$E\left[(Y - d(X))^2\right] = E\left[E\left[(Y - d(X))^2 |X\right]\right]$$

ทฤษฎีที่ว่า ตัวทำนาย (predictor) คือค่าคาดหมายแบบมีเงื่อนไข (conditional mean) Con't

Proof.

ซึ่งบอกเป็นนัยว่า สิ่งที่ต้องการพิสูจน์คือ

$$E\left[E\left[\left(Y-d^{*}\left(X\right)\right)^{2}|X\right]\right]\leq E\left[E\left[\left(Y-d\left(X\right)\right)^{2}|X\right]\right]$$

ความสัมพันธ์นี้จะเป็นจริงถ้า

$$E\left[\left(Y-d^{*}\left(X\right)\right)^{2}|X=x\right]\leq E\left[\left(Y-d\left(X\right)\right)^{2}|X=x\right] \quad (17)$$

สำหรับค่า x ใดๆ เมื่อนำเอาอสมการที่ 17 ไปเปรียบเทียบกับอสมการที่ 16 จะเห็นได้ว่า อสมการทั้งสองนั้นสมมูล (equivalent) กัน หากพิจารณาว่าตัวทำนายใน $E\left[\left(Y-d\left(X\right)\right)^{2}|X=x\right]$ นั้นเป็นค่าจำนวนจริงเช่นเดียวกับ d ในอสมการที่ 16 สรุปได้ว่า

$$E\left[\left(Y-E\left[Y|X\right]\right)^{2}|X=x\right]\leq E\left[\left(Y-d\left(X\right)\right)^{2}|X=x\right] \quad (18)$$

สำหรับฟังก์ชัน $d\left(X\right)$ และค่าจำนวนจริง x ใดๆ ดังนั้น จึงสามารถสรุปได้ว่า

$$E\left[E\left[\left(Y-d^{*}\left(X\right)\right)^{2}|X\right]\right]\leq E\left[E\left[\left(Y-d\left(X\right)\right)^{2}|X\right]\right]$$

ซึ่งนำไปสู่บทสรุปของทฤษฎีบทที่ต้องการ □

ทบทวนนิยามของค่ามัธยฐาน (median)

- ทฤษฎีบทต่อไปนี้แสดงให้เห็นว่าหากเปลี่ยนฟังก์ชันจุดประสงค์ (objective function) เป็นอย่างอื่น ตัวทำนายที่เหมาะสมก็อาจจะเปลี่ยนไปได้
- พิจารณากรณีที่ค่ามัธยฐาน (median) เป็นตัวทำนายที่เหมาะสมหากจุดประสงค์คือการทำให้ค่าคาดหมายของค่าสัมบูรณ์ของค่าคลาดเคลื่อน (m.a.e.) มีค่าต่ำสุด
- จำเป็นต้องกำหนดนิยามของค่ามัธยฐาน (median) ซึ่งครอบคลุมทั้งกรณีของตัวแปรสุ่มต่อเนื่องและไม่ต่อเนื่อง ดังนี้
- ค่ามัธยฐาน (median) มีได้มากกว่า 1 ค่า

Definition

ค่ามัธยฐาน (median) m ของการแจกแจงของตัวแปรสุ่ม X ต้องสอดคล้องกับเงื่อนไขต่อไปนี้

$$Pr\left(X \leq m\right) \geq \frac{1}{2} \text{ และ } Pr\left(X \geq m\right) \geq \frac{1}{2} \tag{19}$$

ทฤษฎีบทที่ว่า ค่ามัธยฐาน (median) เป็นตัวทำนายที่เหมาะสม

- ทฤษฎีบทต่อไปนี้แสดงว่า ค่ามัธยฐาน (median) เป็นตัวทำนายที่เหมาะสมหากจุดประสงค์คือการทำให้ค่าคาดหมายของค่าสัมบูรณ์ของค่าคลาดเคลื่อน (m.a.e.) มีค่าต่ำสุด

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มที่ค่าความคาดหมายมีค่าจำกัด และ m แทนค่ามัธยฐาน (median) ของการแจกแจงของตัวแปรสุ่ม X ดังนั้น สำหรับค่าจำนวนจริง d ใดๆ

$$E\left[|X-m|\right] \leq E\left[|X-d|\right] \tag{20}$$

การพิสูจน์ทฤษฎีบทที่ว่า ค่ามัธยฐาน (median) เป็นตัวทำนายที่เหมาะสม

Proof.

ในที่นี้จะพิจารณาเฉพาะกรณีที่ $d > m$ ส่วนกรณีตรงกันข้ามสามารถพิสูจน์ได้ด้วยขั้นตอนที่คล้ายคลึงกัน

$$\begin{aligned} E\left[|X-d|\right]-E\left[|X-m|\right] &= \int_{-\infty}^{\infty}\left(|x-d|-|x-m|\right)f(x)\,dx \\ &= \int_{-\infty}^{m}(d-m)f(x)\,dx+\int_{m}^{d}(d+m-2x)f(x)\,dx+\int_{d}^{\infty}(m-d)f(x)\,dx \end{aligned}$$

เนื่องจาก $d+m-2x \leq d+m-2d = m-d$ สำหรับจำนวนจริง x ใดๆ ที่ $m \leq x \leq d$ จึงสามารถเขียนสมการข้างบนใหม่ได้เป็น

$$\begin{aligned} E\left[|X-d|\right]-E\left[|X-m|\right] &\geq \int_{-\infty}^{m}(d-m)f(x)\,dx+\int_{m}^{d}(m-d)f(x)\,dx+\int_{d}^{\infty}(m-d)f(x)\,dx \\ &= (d-m)\left[\int_{-\infty}^{m}f(x)\,dx-\int_{m}^{\infty}f(x)\,dx\right] \\ &= (d-m)\left[2\int_{-\infty}^{m}f(x)\,dx-\left[\int_{-\infty}^{m}f(x)\,dx+\int_{m}^{\infty}f(x)\,dx\right]\right] \\ &= (d-m)\left[2\int_{-\infty}^{m}f(x)\,dx-\int_{-\infty}^{\infty}f(x)\,dx\right]=(d-m)\left[2Pr\left(X\leq m\right)-1\right] \end{aligned}$$

ในขณะเดียวกัน ค่ามัธยฐาน m จะต้องสอดคล้องกับเงื่อนไขที่ว่า $Pr\left(X \leq m\right) \geq \frac{1}{2}$ ซึ่งส่งผลให้ $2Pr\left(X \leq m\right)-1 \geq 0$ จึงสามารถสรุปได้ว่า สำหรับค่าจำนวนจริง d ใดๆ

$$E\left[|X-d|\right]-E\left[|X-m|\right] \geq 0 \Rightarrow E\left[|X-d|\right] \geq E\left[|X-m|\right]$$