Conditional Probability and Bivariate Distributions

รศ.ดร. วีระชาติ กิเลนทอง
มหาวิทยาลัยหอการค้าไทย

©Kilenthong 2019

Main issue

- ความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability)
  - ความเป็นอิสระต่อกัน (Independent Events)
  - ทฤษฎีบทของเบส์ (Bayes’ Theorem)
- การแจกแจงของสองตัวแปร (Bivariate Distributions)
  - การแจกแจงตามขอบ (Marginal Distributions)
  - การเป็นอิสระต่อกันระหว่างตัวแปรสุ่ม
- เมื่อเราพิจารณาตัวแปรสุ่ม เราควรต้องถามว่า ตัวแปรสุ่มนั้นมีการแจกแจงอย่างไร? เพราะ
  - การแจกแจง (distribution) คือผลของการรวบรวมคุณสมบัติเชิงความน่าจะเป็น (probabilistic properties) ทั้งหมดของตัวแปรสุ่ม

นิยามของความน่าจะเป็นแบบมีเงื่อนไข (Conditional Probability)

- เรามีคำถามมากมายในชีวิตประจำวันที่อยู่ในรูปแบบของความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) ยกตัวอย่างเช่น เราอาจจะอยากทราบว่า ผู้ชายมีโอกาสป่วยเป็นโรคหัวใจมากน้อยแค่ไหน?
- ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) ยังเป็นรากฐานที่สำคัญของการพยากรณ์ (prediction) หรือการประมาณการณ์ (estimation)

Definition

กำหนดให้ A และ B เป็นเหตุการณ์ใดๆ ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) ของเหตุการณ์ A เมื่อกำหนดให้ว่าเหตุการณ์ B เกิดขึ้น คือ

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} \tag{1}$$

โดยที่ $Pr(B) > 0$ จะต้องมีค่ามากกว่าศูนย์เท่านั้น หาก $Pr(B)$ มีค่าเท่ากับศูนย์แล้ว $Pr(A|B)$ จะไม่มีความหมาย

ตัวอย่าง: การคำนวณหาความน่าจะเป็นแบบมีเงื่อนไข (conditional probability)

- สมมุติว่าเราทราบว่าผลรวมของการโยนลูกเต๋าสองครั้ง Y มีค่าเป็นเลขคู่ แล้วความน่าจะเป็นที่ Y จะมีค่าน้อยกว่า 8 เป็นเท่าใด?
- กำหนดให้ A คือเหตุการณ์ที่ Y จะมีค่าน้อยกว่า 8 และ B คือเหตุการณ์ที่ Y มีค่าเป็นเลขคู่นั้นคือ เราต้องการหา "$Pr\left(A|B\right)$"
- เริ่มจากการหาความน่าจะเป็นของเหตุการณ์ B

$$Pr\left(B\right) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} + \frac{5}{36} + \frac{3}{36} + \frac{1}{36} = \frac{18}{36} = \frac{1}{2}$$

- ผลรวมของการโยนลูกเต๋าสองครั้ง Y มีค่าเป็นเลขคู่และมีค่าน้อยกว่า 8 ได้เป็น

$$Pr\left(A \cap B\right) = \frac{1}{36} + \frac{3}{36} + \frac{5}{36} = \frac{9}{36} = \frac{1}{4}$$

- ดังนั้น ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) ของเหตุการณ์ A เมื่อกำหนดให้ว่าเหตุการณ์ B เกิดขึ้น มีค่าเท่ากับ

$$Pr\left(A|B\right) = \frac{Pr\left(A \cap B\right)}{Pr\left(B\right)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2}$$

กฎการคูณของความน่าจะเป็นแบบมีเงื่อนไข (multiplication rule for conditional probabilities)

Theorem

กำหนดให้ A และ B เป็นเหตุการณ์ใดๆ ถ้า $Pr(B) > 0$ แล้ว

$$Pr(A \cap B) = Pr(B)\, Pr(A|B) \tag{2}$$

และในทำนองเดียวกัน ถ้า $Pr(A) > 0$ แล้ว

$$Pr(A \cap B) = Pr(A)\, Pr(B|A) \tag{3}$$

Theorem

กำหนดให้ $A_1, \ldots, A_n$ เป็นเหตุการณ์ใดๆ โดยที่ $Pr(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$ แล้ว

$$Pr(A_1 \cap \cdots \cap A_n) = Pr(A_1)\, Pr(A_2|A_1) \cdots Pr(A_n|A_1 \cap \cdots \cap A_{n-1}) \tag{4}$$

ตัวอย่าง: กฎการคูณของความน่าจะเป็นแบบมีเงื่อนไข

- พิจารณาการทดลองที่ต้องจับลูกบอลครั้งละลูกทั้งหมด 4 ครั้ง โดยไม่ใส่ลูกบอลกลับเข้าไป จากกล่องที่มีลูกบอลสีแดง ρ ลูก และ ลูกบอลสีฟ้า β ลูก
- คำถาม: ความน่าจะเป็นที่ผลของการจับลูกบอลเป็น สีแดง สีฟ้า สีแดง สีฟ้า มีค่าเท่าใด?
- กำหนดให้เหตุการณ์ R_j แทนเหตุการณ์ที่จับได้ลูกบอลสีแดงในการจับครั้งที่ j^{th} และ B_j แทนเหตุการณ์ที่จับได้ลูกบอลสีฟ้าในการจับครั้งที่ j^{th}
- ดังนั้น สิ่งที่เราต้องการหาคือ

$$Pr\left(R_1 \cap B_2 \cap R_3 \cap B_4\right) = Pr\left(R_1\right) Pr\left(B_2 | R_1\right) Pr\left(R_3 | R_1 \cap B_2\right) Pr\left(B_4 | R_1 \cap B_2 \cap R_3\right)$$

- เราเหลือเพียงแค่ต้องหาค่าของความน่าจะเป็นแบบมีเงื่อนไขแต่ละพจน์ซึ่งมีค่าเท่ากับ

$$Pr\left(R_1\right) = \frac{\rho}{\rho+\beta}, Pr\left(B_2 | R_1\right) = \frac{\beta}{\rho+\beta-1},$$
$$Pr\left(R_3 | R_1 \cap B_2\right) = \frac{\rho-1}{\rho+\beta-2}, Pr\left(B_4 | R_1 \cap B_2 \cap R_3\right) = \frac{\beta-1}{\rho+\beta-3}$$

- ดังนั้น ความน่าจะเป็นที่ผลของการจับลูกบอลเป็น สีแดง สีฟ้า สีแดง สีฟ้า มีค่าเท่ากับ

$$Pr\left(R_1 \cap B_2 \cap R_3 \cap B_4\right) = \frac{\rho}{\rho+\beta} \times \frac{\beta}{\rho+\beta-1} \times \frac{\rho-1}{\rho+\beta-2} \times \frac{\beta-1}{\rho+\beta-3}$$

ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) เป็นมาตรวัดความน่าจะเป็น (probability measure) อันหนึ่ง

ความน่าจะเป็นแบบมีเงื่อนไข (conditional probability) สอดคล้องกับสัจพจน์ของความน่าจะเป็น

1. $Pr(A|B) \geq 0$ ซึ่งเป็นผลมาจากการที่ $Pr(A \cap B) \geq 0$ และ $Pr(B) > 0$
2. $Pr(S|B) \geq 0$ ซึ่งเป็นผลมาจากการที่

$$Pr(S|B) = \frac{Pr(S \cap B)}{Pr(B)} = \frac{Pr(B)}{Pr(B)} = 1$$

ซึ่งเป็นผลมาจากการที่ $Pr(S \cap B) = Pr(B)$

3. พิจารณา

$$Pr\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) = \frac{Pr\left(\left[\bigcup_{i=1}^{\infty} A_i\right] \cap B\right)}{Pr(B)} = \frac{Pr\left(\bigcup_{i=1}^{\infty} [A_i \cap B]\right)}{Pr(B)}$$

เนื่องจาก $A_1, A_2, \ldots$ ไม่มีส่วนร่วมต่อกัน (disjoint events) ทำให้ $A_1 \cap B, A_2 \cap B, \ldots$ ไม่มีส่วนร่วมต่อกันเช่นเดียวกัน ดังนั้น

$$Pr\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right) = \frac{Pr\left(\bigcup_{i=1}^{\infty} [A_i \cap B]\right)}{Pr(B)} = \sum_{i=1}^{\infty} \frac{Pr(A \cap B)}{Pr(B)} = \sum_{i=1}^{\infty} Pr(A_i|B)$$

การแบ่งส่วน (partition) และ กฎของความน่าจะเป็นรวม (Law of total probability)

Definition

กำหนดให้ S คือปริภูมิตัวอย่าง (sample space) ของการทดลองอันหนึ่ง และ $B_1, B_2, \ldots, B_k$ คือเหตุการณ์ใดๆ ในปริภูมิตัวอย่าง S ที่ไม่มีส่วนร่วมต่อกัน (disjoint) และ $\bigcup_{i=1}^{k} B_i = S$ ในกรณีนี้ เราเรียกเหตุการณ์เหล่านี้ว่าเป็นการแบ่งส่วน (partition) ของปริภูมิตัวอย่าง S

Theorem (กฎของความน่าจะเป็นรวม (Law of total probability))

กำหนดให้ $B_1, B_2, \ldots, B_k$ เป็นการแบ่งส่วน (partition) ของปริภูมิตัวอย่าง S และ $Pr\left(B_i\right) > 0$ สำหรับ $i = 1, \ldots, k$ ดังนั้น สำหรับเหตุการณ์ $A \subset S$ ใดๆ

$$Pr\left(A\right) = \sum_{i=1}^{k} Pr\left(B_i\right) Pr\left(A|B_i\right) \tag{5}$$

ตัวอย่าง : กฎของความน่าจะเป็นรวม (Law of total probability)

- สมมุติว่ามีกล่องบรรจุลูกบอลทั้งหมด 2 กล่อง โดยกล่องแรกมีลูกบอลสีฟ้า 70 ลูก และลูกบอลสีแดง 30 ลูก ส่วนกล่องที่สองมีลูกบอลสีฟ้า 10 ลูก และลูกบอลสีแดง 30 ลูก
- กำหนดให้ B_1 และ B_2 คือเหตุการณ์ที่กล่องที่หนึ่งและกล่องที่สองถูกเลือกตามลำดับ
- **คำถาม** : หากการเลือกกล่องเป็นแบบสุ่ม ความน่าจะเป็นของการจับได้ลูกบอลสีฟ้าเป็นเท่าใด?
- กำหนดให้ A คือเหตุการณ์ที่จับได้ลูกบอลสีฟ้า ดังนั้น ความน่าจะเป็นของการจับได้ลูกบอลสีฟ้าสามารถเขียนในรูปของกฎของความน่าจะเป็นรวมได้เป็น

$$Pr(A) = Pr(B_1)\,Pr(A|B_1) + Pr(B_2)\,Pr(A|B_2)$$

- เนื่องเราสุ่มเลือกกล่อง ดังนั้น $Pr(B_1) = Pr(B_2) = \frac{1}{2}$ $Pr(A|B_1) = \frac{7}{10}, Pr(A|B_2) = \frac{1}{4}$
- ดังนั้น ความน่าจะเป็นของการจับได้ลูกบอลสีฟ้าเท่ากับ

$$Pr(A) = \frac{1}{2} \times \frac{7}{10} + \frac{1}{2} \times \frac{1}{4} = \frac{19}{40}$$

- ข้อสังเกต: หากเราคิดโดยรวมเอาทั้งสองกล่องเข้าด้วยกันจะทำให้มีลูกบอลสีฟ้า 80 ลูก และลูกบอลสีแดง 60 ลูก ซึ่งในกรณีนี้จะได้ความน่าจะเป็นของการจับได้ลูกบอลสีฟ้าเท่ากับ $\frac{4}{7}$ แทน

นิยามของความเป็นอิสระต่อกัน (Independence)

- เหตุการณ์สองเหตุการณ์จะเป็นอิสระต่อกันก็ต่อเมื่อ เราทราบว่ามีเหตุการณ์หนึ่งเกิดขึ้นแล้ว **ไม่ได้มีผลกระทบ** ต่อความน่าจะเป็นของอีกเหตุการณ์หนึ่ง กล่าวคือ $Pr(A|B) = Pr(A)$ และ $Pr(B|A) = Pr(B)$

Definition

เหตุการณ์ A และ B **เป็นอิสระต่อกัน (independent)** ถ้า

$$Pr(A \cap B) = Pr(A)\,Pr(B) \tag{6}$$

Definition

เหตุการณ์ $A_1, \ldots, A_n$ เป็นอิสระต่อกัน (independent) ถ้าทุกๆ กลุ่มเหตุการณ์ k เหตุการณ์ใดๆ $A_{i_1}, \ldots, A_{i_k}$ สำหรับ $k = 2, \ldots, n$

$$Pr(A_{i_1} \cap \cdots \cap A_{i_k}) = Pr(A_{i_1}) \cap \cdots \cap Pr(A_{i_k}) \tag{7}$$

โดยที่ดัชนี $i_j \in \{1, 2, \ldots, n\}$ คือดัชนีที่บอกว่าเป็นเหตุการณ์ใด เช่น $i_1 = 4$ หมายความว่า $A_{i_1} = A_4$

ตัวอย่างของความเป็นอิสระต่อกัน (Independence)

- พิจารณาการโยนลูกเต๋าหนึ่งครั้ง กำหนดให้ A เป็นเหตุการณ์ที่ผลการโยนเป็นเลขคี่ และ
- B เป็นเหตุการณ์ที่ผลการโยนเป็นตัวเลขใดตัวเลขหนึ่งในเซ็ต $\{1, 2, 3, 4\}$
- ความน่าจะเป็นของทั้งสองเหตุการณ์เท่ากับ $Pr(A) = \frac{1}{2}$ และ $Pr(B) = \frac{2}{3}$
- เราต้องการตรวจสอบว่าสองเหตุการณ์นี้เป็นอิสระต่อกันหรือไม่?
- เนื่องจาก $A \cap B = \{1, 3\}$ ดังนั้น $Pr(A \cap B) = \frac{1}{3}$ ซึ่งมีค่าเท่ากับ $Pr(A) \times Pr(B) = \frac{1}{2} \times \frac{2}{3}$
- ดังนั้น เราจึงสามารถสรุปได้ว่า A และ B เป็น “**อิสระต่อกัน (independent)**”

ตัวอย่าง: ความเป็นอิสระต่อกันของกลุ่มย่อยไม่การันตีความเป็นอิสระต่อกันของกลุ่มใหญ่

- พิจารณาการโยนเหรียญสองครั้งติดต่อกัน โดยมีปริภูมิตัวอย่าง $S = \{HH, HT, TH, TT\}$ โดยที่ H แทนผลการโยนที่เป็นหัว และ T แทนผลการโยนที่เป็นก้อย
- กำหนดให้ $A = \{HH, HT\}, B = \{HH, TH\}, C = \{HH, TT\}$,
- ดังนั้น $Pr(A) = Pr(B) = Pr(C) = \frac{1}{2}$
- ในขณะเดียวกัน เราสามารถแสดงได้ว่า $A \cap B = A \cap C = B \cap C = \{HH\}$ ซึ่งทำให้เราสามารถคำนวณได้ว่า

$$Pr(A \cap B) = Pr(A \cap C) = Pr(B \cap C) = \frac{1}{4}$$

- โดยสรุปเหตุการณ์ A, B และ C เป็นอิสระต่อกันเป็นคู่ (pairwise independent) แต่เหตุการณ์เหล่านี้ “**ไม่ได้เป็นอิสระต่อกัน**” เพราะ

$$Pr(A \cap B \cap C) = \frac{1}{4}$$

มีค่าแตกต่างจาก $Pr(A)\, Pr(B)\, Pr(C)$

ความเป็นอิสระต่อกัน (Independence) และความเป็นอิสระต่อกันแบบมีเงื่อนไข (Conditionally Independence)

Theorem

กำหนดให้ $A_1, \ldots, A_n$ *เป็นเหตุการณ์ใดๆ โดยที่* $Pr\left(A_1 \cap \cdots \cap A_{n-1} \cap A_n\right) > 0$ *ดังนั้น* $A_1, \ldots, A_n$ *เป็นอิสระต่อกัน (independent) ก็ต่อเมื่อ สำหรับเซ็ตที่ไม่มีส่วนร่วมกันสองเซ็ต* $\{i_1, \ldots, i_m\}$ *และ* $\{j_1, \ldots, j_l\}$ *ซึ่งเป็นสับเซ็ตของ* $\{1, \ldots, n\}$

$$Pr\left(A_{i_1} \cap \cdots \cap A_{i_m} | A_{j_1} \cap \cdots \cap A_{j_l}\right) = Pr\left(A_{i_1} \cap \cdots \cap A_{i_m}\right) \tag{8}$$

Definition

กลุ่มเหตุการณ์ $A_1, \ldots, A_n$ เป็นอิสระต่อกันแบบมีเงื่อนไข (conditionally independent) ภายใต้เงื่อนไขของเหตุการณ์ B ถ้า

$$Pr\left(A_{i_1} \cap \cdots \cap A_{i_m} | B\right) = Pr\left(A_{i_1} | B\right) \cdots Pr\left(A_{i_m} | B\right) \tag{9}$$

สำหรับทุกๆ กลุ่มย่อย $A_{i_1}, \ldots, A_{i_m}$ ของ $2 \leq m \leq n$ เหตุการณ์

ความสัมพันธ์ระหว่างการเป็นอิสระต่อกัน (independent) และการไม่มีส่วนร่วมต่อกัน (mutually exclusive)

- ถึงแม้ว่าชื่อของทั้งสองหลักการจะดูมีส่วนคล้ายกันไม่น้อย แต่ในความเป็นจริงแล้ว กลุ่มของเหตุการณ์โดยทั่วไปจะไม่สามารถสอดคล้องกับทั้งสองหลักการพร้อมกันได้
- เนื่องจากการเป็น **อิสระต่อกัน (independent)** เป็นเงื่อนไขที่กำหนดว่า การเรียนรู้เกี่ยวกับเหตุการณ์หนึ่งจะต้อง **ไม่มีผล** ต่อความน่าจะเป็นของเหตุการณ์หนึ่ง
- ในขณะที่ **การไม่มีส่วนร่วม (mutually exclusive or disjoint)** กลับบอกว่าเมื่อทราบว่าเหตุการณ์หนึ่งเกิดแล้วย่อม **ทราบทันทีว่าอีกเหตุการณ์หนึ่งเกิดขึ้นไม่ได้** กล่าวคือ ภายใต้หลักการไม่มีส่วนร่วมต่อกัน การเรียนรู้เกี่ยวกับเหตุการณ์หนึ่งมีผลต่อความน่าจะเป็นของเหตุการณ์หนึ่งเสมอ ยกเว้นในกรณีที่ เหตุการณ์อันหลังนั้นมีความน่าจะเป็นเท่ากับศูนย์อยู่แล้ว
- เราสรุปได้ว่าหลักการสองอันนี้จะเป็น **จริงพร้อมกันไม่ได้** ยกเว้นในกรณีที่ เหตุการณ์ทุกเหตุการณ์ยกเว้นเพียงหนึ่งอันมีความน่าจะเป็นเท่ากับศูนย์

นิยามของทฤษฎีบทของเบส์

Theorem (ทฤษฎีบทของเบส์ (Bayes' theorem))

กำหนดให้ $B_1, B_2, \ldots, B_k$ *เป็นการแบ่งส่วน (partition) ของปริภูมิตัวอย่าง* S *โดยที่* $Pr(B_i) > 0$ *สำหรับ* $i = 1, \ldots, k$ *และ* A *เป็นเหตุการณ์ใดๆ ที่* $Pr(A) > 0$ *ดังนั้น สำหรับ* $i = 1, \ldots, k$

$$Pr(B_i|A) = \frac{Pr(B_i)\, Pr(A|B_i)}{\sum_{j=1}^{k} Pr(B_j)\, Pr(A|B_j)} \tag{10}$$

- ในมุมมองของสถิติแบบเบส์ (Bayesian Statistics) $Pr(B_i)$ มักจะถูกเรียกว่าความน่าจะเป็นก่อน (prior probability) ซึ่งสะท้อนถึงความรู้หรือความเชื่อที่ผู้วิเคราะห์มีต่อเหตุการณ์ B_i ซึ่งจะถูกปรับปรุงเมื่อมีข้อมูลใหม่เพิ่มเข้ามาซึ่งในที่นี้หมายถึงเหตุการณ์ A
- ผลลัพธ์ที่ได้จากการปรับปรุง (updating) $Pr(B_i|A)$ มักจะถูกเรียกว่าความน่าจะเป็นหลัง (posterior probability)

ตัวอย่างของการประยุกต์ใช้ทฤษฎีบทของเบส์

- พิจารณาเทคโนโลยีการตรวจการติดเชื้ออันหนึ่ง โดยหากคนใดติดเชื้อจริง ความน่าจะเป็นที่ผลการตรวจจะออกมาเป็นบวกหรือบอกว่าติดเชื้อมีค่าเท่ากับ 0.90 ในขณะที่ ความน่าจะเป็นที่ผลการตรวจของคนที่ไม่ติดเชื้อจะออกมาเป็นบวกมีค่าเท่ากับ 0.10
- ข้อมูลในภาพรวมยังบอกอีกว่า โอกาสในการติดเชื้อในประชากรอยู่ที่ประมาณ 1 ใน 10,000
- สมมุติว่า ผู้ชายคนหนึ่งไปตรวจและพบว่าได้ผลการตรวจเป็นบวก คำถามคือ ความน่าจะเป็นที่จริงๆ แล้วเขาติดเชื้อมีค่าเท่าใด?
- กำหนดให้ A แทนเหตุการณ์ที่ผลของการตรวจเชื้อออกมาเป็นบวก ส่วน B_1 แทนเหตุการณ์ที่ผู้ชายคนนั้นติดเชื้อ และ B_2 แทนเหตุการณ์ที่เขาไม่ได้ติดเชื้อ
- สิ่งที่เราต้องการทราบคือ

$$Pr\left(B_1|A\right) = \frac{Pr\left(B_1\right) Pr\left(A|B_1\right)}{Pr\left(B_1\right) Pr\left(A|B_1\right) + Pr\left(B_2\right) Pr\left(A|B_2\right)} = \frac{0.0001 * 0.90}{0.0001 * 0.90 + 0.9999 * 0.10}$$

ซึ่งมีค่าเท่ากับ 0.0009

- เพื่อให้เห็นผลของข้อมูลภาพรวม เราลองสมมุติว่าโอกาสในการติดเชื้อในประชากรอยู่ที่ประมาณ 1 ใน 100 ในกรณีนี้ ความน่าจะเป็นที่จริงๆ แล้วเขาติดเชื้อมีค่าเท่ากับ

$$Pr\left(B_1|A\right) = \frac{0.01 \times 0.90}{0.01 \times 0.90 + 0.99 \times 0.10} = 0.0833$$

- ซึ่งชี้ให้เห็นว่า หากโอกาสในการติดเชื้อในประชากรไม่สูงมากนัก การตรวจการติดเชื้อที่มีระดับความเชื่อมั่นไม่ถึงร้อยเปอร์เซนต์อาจจะไม่คุ้มค่าก็เป็นได้

การแจกแจงร่วมของสองตัวแปร (Bivariate Distributions)

Definition

การแจกแจงร่วม (joint distribution) ของตัวแปรสุ่ม X และ Y คือเซ็ตของความน่าจะเป็นทุกอันที่อยู่ในรูปแบบ $Pr\left((X,Y) \in C\right)$ สำหรับเซ็ตของจำนวนจริงสองจำนวน C ทุกเซ็ต โดยที่ $\{(X,Y) \in C\}$ เป็นเหตุการณ์

Definition

ตัวแปรสุ่ม X และ Y มีการแจกแจงร่วมไม่ต่อเนื่อง (discrete joint distribution) ถ้าค่าจำนวนจริงทั้งสองจำนวน (x,y) ที่เป็นไปได้ของ (X,Y) มีจำนวนจำกัด (finite) หรืออย่างมากต้องเป็นอนุกรมอนันต์ที่นับได้ (countable)

ทฤษฎีการแจกแจงร่วมไม่ต่อเนื่องของสองตัวแปร

Theorem

ถ้า X และ Y ต่างเป็นตัวแปรสุ่มที่มีการแจกแจงไม่ต่อเนื่อง แล้ว (X, Y) มีการแจกแจงร่วมไม่ต่อเนื่อง (discrete joint distribution)

Proof.

- ค่าที่เป็นไปได้ของตัวแปรสุ่มแต่ละตัวมีจำนวนจำกัด (finite) แล้วค่าจำนวนจริงทั้งสองจำนวน (x, y) ที่เป็นไปได้ของ (X, Y) จะต้องมีจำนวนจำกัดตามไปด้วย
- ถ้าแต่ละตัวมีค่าที่เป็นไปได้เป็นอนันต์แต่นับได้ (countably infinite) แล้วค่าจำนวนจริงทั้งสองจำนวน (x, y) ที่เป็นไปได้ของ (X, Y) จะต้องเป็นอนันต์แต่นับได้ (countably infinite)

$\square$

นิยามและคุณสมบัติ: การแจกแจงร่วมไม่ต่อเนื่องของสองตัวแปร

Definition

ฟังก์ชันความน่าจะเป็นร่วม (joint probability function) ของตัวแปรสุ่ม X และ Y คือฟังก์ชัน f ซึ่งนิยามสำหรับทุกค่า (x, y) ในระนาบ xy โดยที่

$$f(x, y) = Pr(X = x \text{ และ } Y = y) \tag{11}$$

- คุณสมบัติ
  1. ถ้า (x, y) เป็นค่าที่เป็นไปไม่ได้สำหรับ (X, Y) แล้ว $f(x, y) = 0$
  2. สำหรับเซ็ตของจำนวนจริงสองจำนวน C ใดๆ

$$Pr((X, Y) \in C) = \sum_{(x,y) \in C} f(x, y) \tag{12}$$

  3. ค่าผลรวมของฟังก์ชันความน่าจะเป็นร่วมจากทุกค่า (x, y) ที่เป็นไปได้จะต้องเท่ากับหนึ่ง นั่นคือ

$$\sum_{(x,y) \in \mathbb{R}^2} f(x, y) = 1 \tag{13}$$

ตัวอย่าง: ฟังก์ชันความน่าจะเป็นร่วมไม่ต่อเนื่อง

- ตัวอย่างการประยุกต์ใช้ฟังก์ชันความน่าจะเป็นร่วม (joint probability function) ในชีวิตประจำวันคือการนำเสนอข้อมูลในรูปแบบตารางไขว้ (cross tabulation) ซึ่งช่วยให้สามารถสังเกตความสัมพันธ์ระหว่างตัวแปรได้โดยสะดวก

Table: สัดส่วนของกองทุนในประเทศไทยช่วงปี 2017-2018

	ประเภทของกองทุนในประเทศไทย (Y)			
ผลตอบแทนของกองทุน (X)	**Equity** ($Y=0$)	**Fixed** ($Y=1$)	**Mixed** ($Y=2$)	**ผลรวมทั้งหมด**
ขาดทุน ($X=0$)	0.03	0.03	0.01	0.07
กำไร ($X=1$)	0.46	0.27	0.20	0.93
ผลรวมทั้งหมด	**0.49**	**0.30**	**0.21**	**1.00**

- ตารางนี้แสดงสัดส่วนของกองทุนโดยแบ่งตามประเภทการลงทุนและผลกำไรที่ได้จากการลงทุนในช่วง 1 ปี ซึ่งแบ่งเป็นสองกลุ่มคือ
  - กลุ่มที่ผลตอบแทนมากกว่าเงินลงทุน (กำไร)
  - กลุ่มที่ผลตอบแทนน้อยกว่าเงินลงทุน (ขาดทุน)
- ตัวเลขในแต่ละช่องนั้นแทนฟังก์ชันความน่าจะเป็นร่วมของตัวแปรสุ่มสองตัว
  - ผลตอบแทนของกองทุน
  - ประเภทการลงทุนของกองทุน

นิยาม: ฟังก์ชันความน่าจะเป็นต่อเนื่องร่วม

- เช่นเดียวกันกับการแจกแจงต่อเนื่องของตัวแปรสุ่มตัวเดียว เราสามารถนิยาม **การแจกแจงต่อเนื่องของตัวแปรสุ่มสองตัวได้โดยใช้ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.)**

Definition

ตัวแปรสุ่ม X และ Y มีการแจกแจงต่อเนื่องร่วม (continuous joint distribution) ถ้ามีฟังก์ชันที่ไม่เป็นลบ (nonnegative function) f ที่

$$Pr\left((X,Y) \in C\right) = \iint_C f(x,y)\,dxdy \tag{14}$$

- สำหรับทุกสับเซ็ตของจำนวนจริงสองจำนวน C ใดๆ ที่ผลของการอินทิเกรตสามารถหาค่าได้ โดยเรียกฟังก์ชัน f ว่า ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint probability density function)
- เราเรียกส่วนปิดคลุม (closure) ของเซ็ต $\{(x,y) : f(x,y) > 0\}$ ว่า ส่วนค้ำจุน (support) ของ (X,Y)

ทฤษฎีการแจกแจงต่อเนื่องร่วม (continuous joint distribution)

Theorem

ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint probability density function) สอดคล้องกับเงื่อนไขต่อไปนี้

$$f(x,y) \geq 0, \text{ สำหรับ } (x,y) \in \mathbb{R}^2 \tag{15}$$

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y)\, dxdy = 1 \tag{16}$$

Theorem

การแจกแจงต่อเนื่องร่วม (continuous joint distribution) นิยามบนระนาบ xy ใดๆ สอดคล้องกับเงื่อนไขต่อไปนี้

1. *ความน่าจะเป็นของจุดใดๆ หรือของอนุกรมอนันต์ของจุดใดๆ ในระนาบ xy มีค่าเท่ากับศูนย์เสมอ*
2. *ถ้า g เป็นฟังก์ชันต่อเนื่อง (continuous function) ของตัวแปรหนึ่งตัวที่นิยามบนช่วง $[a,b]$ แล้ว เซ็ตหรือเหตุการณ์ $\{(x,y) : y = g(x), a < x < b\}$ และ $\{(x,y) : x = g(y), a < y < b\}$ ต่างมีความน่าจะเป็นเท่ากับศูนย์*

ตัวอย่าง: คำนวณหาค่าคงที่มาตรฐานของสำหรับฟังก์ชันความหนาแน่นของความน่าจะเป็นต่อเนื่องร่วม

- เช่นเดียวกับกรณีของตัวแปรสุ่มตัวแปรเดียว เราต้องคำนวณหาค่าคงที่มาตรฐาน (normalizing constant) สำหรับฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.)

Example

กำหนดให้ตัวแปรสุ่ม X และ Y มีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ดังต่อไปนี้

$$f(x,y) = \begin{cases} cx^2y, & \text{สำหรับ } x^2 \leq y \leq 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

คำถามก็คือ ค่าคงที่มาตรฐาน (normalizing constant) c ต้องมีค่าเท่าใดเพื่อให้ฟังก์ชัน $f(x,y)$ เป็นฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.)?

- คำตอบหาได้โดยใช้สมการที่ 16
- กำหนดให้ S คือส่วนค้ำจุน (support) ของ (X,Y)

วิธีทำ: คำนวณหาค่าคงที่มาตรฐานสำหรับฟังก์ชันความหนาแน่นของความน่าจะเป็นต่อเนื่องร่วม

Example

- การอินทิเกรตสองชั้นของฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ได้เป็น

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy = \iint_S cx^2 ydxdy = \int_{-1}^{1} cx^2\left[\int_{x^2}^{1} ydy\right]dx$$

$$= \int_{-1}^{1} cx^2\left[\left.\frac{y^2}{2}\right|_{x^2}^{1}\right]dx = c\int_{-1}^{1} x^2\left[\frac{1}{2}-\frac{x^4}{2}\right]dx$$

$$= \frac{4}{21}c$$

- เนื่องจากผลการอินทิเกรตสองชั้นของฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) จะต้องมีค่าเท่ากับหนึ่ง (สมการที่ 16) ดังนั้น ค่าคงที่มาตรฐาน (normalizing constant) $c = \frac{21}{4}$

$$f(x,y) = \begin{cases} \frac{21}{4}x^2y, & \text{สำหรับ } x^2 \le y \le 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- ข้อสังเกตในทางคณิตศาสตร์อันหนึ่งคือ ลำดับในการอินทิเกรตไม่มีผลต่อการหาค่าคงที่มาตรฐานในข้อนี้

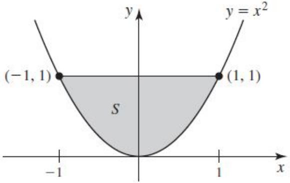
ตัวอย่าง: การคำนวณหา $Pr(X \geq Y)$

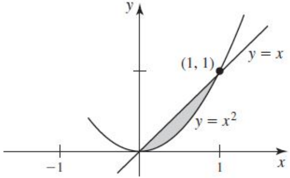
Example

พิจารณาตัวแปรสุ่ม X และ Y ซึ่งมีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) เช่นเดียวกับในตัวอย่างเดียวกับตัวอย่างที่แล้ว คำถามก็คือ ความน่าจะเป็นที่ $X \geq Y$ มีค่าเท่ากับเท่าใด เราสามารถหาคำตอบนี้ได้โดยใช้สมการที่ 14 ดังต่อไปนี้

$$Pr(X \geq Y) = \int_0^1 \int_{x^2}^{x} \frac{21}{4} x^2 y dy dx = \frac{3}{20}$$

โดยในที่นี้ เราจะเริ่มอินทิเกรตจากตัวแปร y ก่อนและมีขอบเขตของการอินทิเกรตดังแสดงในรูปต่อไปนี้





นิยามและคุณสมบัติ: Mixed Distribution

Definition

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่มีการแจกแจงไม่ต่อเนื่อง (discrete) และต่อเนื่อง (continuous) ตามลำดับ และถ้ามีฟังก์ชัน $f(x, y)$ ที่นิยามบนระนาบ xy โดยที่เซ็ตจำนวนจริง A และ B ใดๆ

$$Pr\left(X \in A, Y \in B\right) = \int_B \sum_{x \in A} f(x, y)\, dy \tag{17}$$

ในกรณีที่ผลการอินทิเกรตมีค่าจริง (integral exists) แล้ว เราจะเรียกฟังก์ชัน f ว่า ฟังก์ชันความน่าจะเป็นร่วม (joint probability function: joint p.f.) ของ X และ Y

คุณสมบัติที่

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่มีการแจกแจงไม่ต่อเนื่อง (discrete) และต่อเนื่อง (continuous) ตามลำดับ ฟังก์ชันความน่าจะเป็นร่วม (joint p.f.) ของ X และ Y สอดคล้องกับเงื่อนไขต่อไปนี้

$$f(x, y) \geq 0, \text{ สำหรับ } (x, y) \text{ ใดๆ}, \tag{18}$$

$$\int_{-\infty}^{\infty} \sum_{i=1}^{\infty} f(x_i, y)\, dy = 1 \tag{19}$$

ฟังก์ชันการแจกแจงสะสมร่วม (joint cumulative distribution function หรือ joint C.D.F.)

Definition

ฟังก์ชันการแจกแจงสะสมร่วม (joint cumulative distribution function หรือ joint C.D.F.) ของตัวแปรสุ่ม X และ Y คือ

$$F(x,y) = Pr(X \leq x, Y \leq y) \tag{20}$$

ซึ่งเป็นฟังก์ชันที่นิยามสำหรับค่า $(x,y) \in \mathbb{R}^2$

- หากตัวแปรสุ่ม X และ Y เป็นแบบต่อเนื่อง เราสามารถเขียนฟังก์ชันการแจกแจงสะสมร่วม (joint C.D.F.) ของตัวแปรสุ่ม X และ Y ได้เป็น

$$F(x,y) = Pr(X \leq x, Y \leq y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f(r,s)\, drds \tag{21}$$

คุณสมบัติของฟังก์ชันการแจกแจงสะสมร่วม (joint cumulative distribution function หรือ joint C.D.F.)

คุณสมบัติที่

กำหนดให้ X และ Y เป็นตัวแปรสุ่มต่อเนื่อง (continuous random variable) ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) สามารถคำนวณจากฟังก์ชันการแจกแจงสะสมร่วม (joint C.D.F.) ได้ดังต่อไปนี้

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2 F(x,y)}{\partial y \partial x} \tag{22}$$

สำหรับค่าจำนวนจริง x และ y ใดๆ ที่ค่าอนุพันธ์กำลังสอง (second-order derivative) หาค่าได้ (exist)

ตัวอย่าง

Example

พิจารณาตัวแปรสุ่ม X และ Y ที่มีฟังก์ชันการแจกแจงร่วม (joint C.D.F.) เป็น

$$F(x,y) = \frac{1}{156} xy \left(x^2 + y\right)$$

สำหรับค่า (x,y) ที่ $0 \leq x \leq 3$ และ $0 \leq y \leq 4$
เราสามารถหาฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ได้จากอนุพันธ์กำลังสอง (second-order derivative) ของฟังก์ชันการแจกแจงร่วม ดังต่อไปนี้

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{1}{156} \frac{\partial}{\partial x} \left[x^3 + 2xy\right] = \frac{1}{156} \left[3x^2 + 2y\right]$$

นิยาม: การแจกแจงตามขอบ (Marginal Distributions)

Definition

กำหนดให้ X และ Y เป็นตัวแปรสุ่ม ฟังก์ชันการแจกแจงสะสมตามขอบ (marginal C.D.F.) ของตัวแปรสุ่ม X และ Y สามารถนิยามได้เป็น

$$F_1(x) = Pr(X \leq x), \tag{23}$$

$$F_2(y) = Pr(Y \leq y) \tag{24}$$

เราสามารถประยุกต์ใช้หลักการของทฤษฎีเซ็ตที่ว่า

$$\{X \leq x\} = \{X \leq x\} \cap \{Y < \infty\}, \tag{25}$$

$$\{Y \leq y\} = \{Y \leq y\} \cap \{X < \infty\} \tag{26}$$

เพื่อเขียนฟังก์ชันการแจกแจงสะสมตามขอบ (marginal C.D.F.) ของตัวแปรสุ่ม X และ Y ใหม่ได้เป็น

$$F_1(x) = Pr(X \leq x, Y < \infty) = \lim_{y \to \infty} F(x, y), \tag{27}$$

$$F_2(y) = Pr(X < \infty, Y \leq y) = \lim_{x \to \infty} F(x, y) \tag{28}$$

ทฤษฎี: การแจกแจงตามขอบ (Marginal Distributions)

Theorem

ถ้า X และ Y เป็นตัวแปรสุ่มต่อเนื่องที่มี $f(x,y)$ แทนฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) แล้ว ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม X และ Y คือ

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \text{ สำหรับ } -\infty < x < \infty \tag{29}$$

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)\,dx, \text{ สำหรับ } -\infty < y < \infty \tag{30}$$

การพิสูจน์: การแจกแจงตามขอบ (Marginal Distributions)

Proof.

เริ่มจากนิยามของฟังก์ชันการแจกแจงสะสมตามขอบ (marginal C.D.F.) ของ X

$$F_1(x) = Pr(X \leq x, Y < \infty) = \int_{-\infty}^{x} \int_{-\infty}^{\infty} f(r,y)\, dydr = \int_{-\infty}^{x} \left[\int_{-\infty}^{\infty} f(r,y)\, dy \right] dr \qquad (31)$$

ในขณะเดียวกัน เราสามารถนิยามฟังก์ชันการแจกแจงสะสมตามขอบ (marginal C.D.F.) ของ X ในรูปฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ได้เป็น

$$F_1(x) = Pr(X \leq x) = \int_{-\infty}^{x} f_1(r)\, dr$$

ดังนั้น เราสามารถสรุปได้ว่า

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)\, dy \quad \text{และ} \quad f_2(y) = \int_{-\infty}^{\infty} f(x,y)\, dx$$

□

การหาการแจกแจงตามขอบ (Marginal Distributions) ด้วย การใช้อนุพันธ์

- จากทฤษฎีข้างต้นสรุปได้ว่า

$$f_1(x) = \frac{dF_1(x)}{dx}, \tag{32}$$

$$f_2(y) = \frac{dF_2(y)}{dx} \tag{33}$$

- เราสามารถประยุกต์ใช้หลักการที่ว่า สามารถแทนการอินทิเกรตได้ด้วยการหาผลรวม (summation) หากตัวแปรสุ่มเป็นแบบไม่ต่อเนื่อง (discrete)

ทฤษฎีของการแจกแจงตามขอบเมื่อตัวแปรสุ่มเป็นแบบต่างๆ

Theorem

ถ้า X และ Y เป็นตัวแปรสุ่มไม่ต่อเนื่อง (discrete) ที่มี $f(x,y)$ แทนฟังก์ชันความน่าจะเป็นร่วม (joint p.f.) แล้ว ฟังก์ชันความน่าจะเป็นตามขอบ (marginal p.f.) ของตัวแปรสุ่ม X และ Y คือ

$$f_1(x) = \sum_y f(x,y) \quad \text{and} \quad f_2(y) = \sum_x f(x,y) \tag{34}$$

โดยที่ส่วนห้อยของการหาผลรวม (summation) ในสมการที่ 34 หมายถึงให้หาผลรวมจากทุกๆ ค่าของ y หรือ x ตามลำดับ

Theorem

ถ้า X เป็นตัวแปรสุ่มไม่ต่อเนื่อง (discrete) และ Y เป็นตัวแปรสุ่มต่อเนื่อง (continuous) ที่มี $f(x,y)$ แทนฟังก์ชันความน่าจะเป็นร่วม (joint p.f.) แล้ว ฟังก์ชันความน่าจะเป็นตามขอบ (marginal p.f.) ของตัวแปรสุ่ม X คือ

$$f_1(x) = Pr(X = x) = \int_{-\infty}^{\infty} f(x,y)\,dy, \tag{35}$$

ส่วนฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม Y คือ

$$f_2(y) = \sum_x f(x,y), \tag{36}$$

ตัวอย่าง: การแจกแจงตามขอบ (Marginal Distributions)

Example

พิจารณาตัวแปรสุ่ม (X, Y) ที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) เท่ากับ

$$f(x,y) = \begin{cases} 4xye^{-x^2-y^2}, & \text{สำหรับ } 0 < x < \infty, 0 < y < \infty, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม X เท่ากับ

$$f_1(x) = \int_0^\infty 4xye^{-x^2-y^2} dy = 2xe^{-x^2} \int_0^\infty 2ye^{-y^2} dy$$
$$= 2xe^{-x^2} \left[-e^{-y^2} \Big|_0^\infty \right] = 2xe^{-x^2}$$

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม Y เท่ากับ

$$f_2(y) = \int_0^\infty 4xye^{-x^2-y^2} dx = 2ye^{-y^2}$$

สรุปตัวอย่าง: การแจกแจงตามขอบ (Marginal Distributions)

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม X ไม่ขึ้นอยู่กับค่าของตัวแปรสุ่ม Y
- เช่นเดียวกับกรณีของฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของตัวแปรสุ่ม Y ที่ขึ้นอยู่กับค่า y เพียงอย่างเดียว
- ผลลัพธ์นี้เป็นคุณสมบัติที่สำคัญในทางสถิติที่เราเรียกว่า ความเป็นอิสระต่อกัน (independence)

นิยาม: การเป็นอิสระต่อกัน (independent)

Definition

ตัวแปรสุ่มสองตัว (X, Y) เป็นอิสระต่อกัน (independent) ถ้า

$$Pr\left(X \in A \text{ และ } Y \in B\right) = Pr\left(X \in A\right) Pr\left(Y \in B\right) \quad (37)$$

สำหรับทุกเซ็ตของจำนวนจริง A และ B ที่ $\{X \in A\}$ และ $\{Y \in B\}$ เป็นเหตุการณ์

- เราสามารถสรุปได้ว่า ถ้า (X, Y) เป็นอิสระต่อกัน (independent) แล้ว

$$Pr\left(X \leq x \text{ และ } Y \leq y\right) = Pr\left(X \leq x\right) Pr\left(Y \leq y\right) \quad (38)$$

สำหรับทุกๆ จำนวนจริง x และ y

ทฤษฎี: การเป็นอิสระต่อกัน (independent)

Theorem

กำหนดให้ $F(x,y)$ แทนฟังก์ชันการแจกแจงสะสม (C.D.F.) ของ (X,Y) ส่วน $F_1(x)$ แทนฟังก์ชันการแจกแจงสะสมตามขอบ (marginal C.D.F.) ของ X และ $F_2(y)$ แทนฟังก์ชันการแจกแจงสะสมตามขอบ (marginal C.D.F.) ของ Y ดังนั้น X และ Y เป็นอิสระต่อกัน (independent) ก็ต่อเมื่อ (if and only if)

$$F(x,y) = F_1(x)F_2(y) \tag{39}$$

สำหรับทุกๆ จำนวนจริง x และ y

- เราสามารถแปลงเงื่อนไขที่ (39) ให้อยู่ในรูปของฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) และฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) โดยใช้ความสัมพันธ์เชิงอนุพันธ์ของการแจกแจงสะสมและฟังก์ชันความหนาแน่นของความน่าจะเป็น ได้ดังต่อไปนี้

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial F_1(x)}{\partial x}\frac{\partial F_2(y)}{\partial y} = f_1(x)f_2(y) \tag{40}$$

การเป็นอิสระต่อกัน: การแยกตัวประกอบ (factorization)

Theorem

ตัวแปรสุ่มสองตัว (X, Y) เป็นอิสระต่อกัน (independent) ก็ต่อเมื่อ (if and only if) เงื่อนไขการแยกตัวประกอบ (factorization)

$$f(x, y) = f_1(x) f_2(y) \tag{41}$$

เป็นจริงสำหรับทุกๆ จำนวนจริง x และ y

- การประยุกต์ใช้ทฤษฎีนี้ โดยเฉพาะเงื่อนไขที่ว่าสมการที่ (41) จะต้องเป็นจริงสำหรับทุกๆ จำนวนจริง x และ y
- ตัวแปรสุ่มสองตัวที่สอดคล้องกับเงื่อนไขดังกล่าวภายใต้ข้อจำกัดพิเศษบางประเภทเท่านั้น อาจจะไม่เป็นอิสระต่อกัน

ตัวอย่าง: การตรวจสอบความเป็นอิสระต่อกัน (เป็นอิสระต่อกัน)

Example

พิจารณาตัวแปรสุ่ม X และ Y ที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) เป็น

$$f(x,y) = \begin{cases} 9x^2y^2, & \text{สำหรับ } 0 \le x \le 1 \text{ และ } 0 \le y \le 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

ซึ่งมีรูปแบบคล้ายกับตัวอย่างที่แล้ว “**ต่างกันเพียงส่วนค้ำจุน (support)เป็นสี่เหลี่ยมมุมฉาก**” แต่เนื่องจากมีการเปลี่ยนขอบเขต

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของแต่ละตัวแปรสุ่มใหม่ดังนี้

$$f_1(x) = \int_0^1 9x^2y^2dy = 3x^2, \text{ สำหรับ } 0 \le x \le 1,$$

$$f_2(y) = \int_0^1 9x^2y^2dx = 3y^2, \text{ สำหรับ } 0 \le y \le 1$$

สรุป: การตรวจสอบความเป็นอิสระต่อกัน (เป็นอิสระต่อกัน)

Example

- เงื่อนไขการแยกตัวประกอบตามสมการที่ (41) เป็นจริงสำหรับค่า $0 \leq x \leq 1$ และ $0 \leq y \leq 1$
- ส่วนกรณีอื่นๆ นั้น $f_1(x) = f_2(y) = f(x, y) = 0$ ซึ่งสอดคล้องสมการที่ (41)
- สรุปได้ว่า ตัวแปรสุ่ม X และ Y เป็นอิสระต่อกัน

- เพราะส่วนค้ำจุนของเป็นรูปสี่เหลี่ยมมุมฉาก (rectangular) ซึ่งมีขอบเขตขนานกับแกนตั้งและแกนนอน ตัวแปรสุ่มทั้งสองจะเป็น “**อิสระต่อกัน**”
- แต่ว่า ถ้าส่วนค้ำจุน (support) **เปลี่ยนไป** ตัวแปรสุ่มเหล่านี้อาจจะไม่เป็นอิสระต่อกัน

ตัวอย่าง: การตรวจสอบความเป็นอิสระต่อกัน (ไม่เป็นอิสระต่อกัน)

Example

พิจารณาตัวแปรสุ่ม X และ Y ที่มีฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) เป็น

$$f(x,y) = \begin{cases} 9x^2y^2, & \text{สำหรับ } x^2+y^2 \leq 1, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases}$$

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ให้อยู่ในรูปของการแยกตัวประกอบตามสมการที่ (41) สำหรับ (x,y) ที่สอดคล้องกับเงื่อนไข $x^2+y^2 \leq 1$
- ต้องตรวจสอบดูว่า เงื่อนไขดังกล่าวเป็นจริงในส่วนอื่นๆ ด้วยหรือไม่ โดยเริ่มจากการหาฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ของแต่ละตัวแปรสุ่ม ดังนี้

$$f_1(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} 9x^2y^2 dy = 6x^2\left(1-x^2\right)^{\frac{3}{2}},$$

$$f_2(y) = \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} 9x^2y^2 dx = 6y^2\left(1-y^2\right)^{\frac{3}{2}}$$

สรุป: การตรวจสอบความเป็นอิสระต่อกัน (ไม่เป็นอิสระต่อกัน)

Example

- จากฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (joint p.d.f.) ที่กำหนดให้
  - $f\left(\sqrt{0.75}, \sqrt{0.75}\right) = 0$ เนื่องจาก $\sqrt{0.75}^2 + \sqrt{0.75}^2 = 1.5 < 1$
- ค่าฟังก์ชันความหนาแน่นของความน่าจะเป็นตามขอบ (marginal p.d.f.) ที่จุดดังกล่าว คือ
  - $f_1\left(\sqrt{0.75}\right) = f_2\left(\sqrt{0.75}\right) = 0.5625 \neq 0$
- เงื่อนไขการแยกตัวประกอบ (factorization) ไม่เป็นจริงสำหรับค่า (x, y) ที่อยู่นอกส่วนค้ำจุน
- สรุปได้ว่า ตัวแปรสุ่ม X และ Y "**ไม่เป็นอิสระต่อกัน**"

ทฤษฎีของการแจกแจงตามขอบ (Marginal Distributions)

Theorem

*ตัวแปรสุ่มสองตัว (X, Y) มีการแจกแจงร่วมต่อเนื่อง (continuous joint distribution) สมมุติว่าส่วนค้ำจุน (support) ของตัวแปรสุ่ม (X, Y) ซึ่งนิยามโดย $\{(x, y) : f(x, y) > 0\}$ "**เป็นสี่เหลี่ยมมุมฉาก (อาจจะไม่มีขอบเขตก็เป็นได้) โดยแต่ละด้านขนานกับแกนของปริภูมิของยุคลิค (Euclidean space)**" แล้ว (X, Y) เป็น**อิสระต่อกัน (independent)** ก็ต่อเมื่อ (if and only if) เงื่อนไขการแยกตัวประกอบ (factorization)*

$$f(x, y) = f_1(x) f_2(y) \tag{42}$$

เป็นจริงสำหรับทุกๆ จำนวนจริง x และ y

ทฤษฎี: การเป็นอิสระต่อกันของตัวแปรสุ่ม

Theorem

ถ้าตัวแปรสุ่ม X และ Y เป็นอิสระต่อกัน (independent) แล้ว ตัวแปรสุ่ม $h(X)$ และ $g(Y)$ เป็นอิสระต่อกัน (independent) ไม่ว่าฟังก์ชัน h และ g จะมีรูปแบบอย่างไร

Proof.

พิจารณาเหตุการณ์ $H = \{h(X) \leq r\}$ ซึ่งสามารถแปลงให้อยู่ในรูปของเซ็ตของ x ได้เป็น $A \equiv \{x : h(x) \leq r\}$ ในทำนองเดียวกัน สามารถแปลงเหตุการณ์ $G = \{g(Y) \leq s\}$ ได้เป็น $B \equiv \{y : g(y) \leq s\}$ เนื่องจาก X และ Y เป็นอิสระต่อกัน (independent) ดังนั้น

$$Pr(X \in A \text{ และ } Y \in B) = Pr(X \in A)\, Pr(Y \in B)$$

ซึ่งสามารถเขียนในรูปของเหตุการณ์ H และ G ได้เป็น

$$\begin{aligned} Pr(h(X) \in H \text{ และ } g(Y) \in G) &= Pr(X \in A \text{ และ } Y \in B) \\ &= Pr(X \in A)\, Pr(Y \in B) \\ &= Pr(h(X) \in H)\, Pr(g(Y) \in G) \end{aligned}$$

นั่นคือ ตัวแปรสุ่ม $h(X)$ และ $g(Y)$ สอดคล้องกับเงื่อนไขการแยกตัวประกอบของความน่าจะเป็นตามสมการที่ (37) ดังนั้น เราจึงสามารถสรุปได้ว่า $h(X)$ และ $g(Y)$ เป็นอิสระต่อกัน (independent) $\square$