Covariance, Correlation, and Moments

รศ.ดร. วีระชาติ กิเลนทอง
มหาวิทยาลัยหอการค้าไทย

นิยาม: ความแปรปรวนร่วมและสหสัมพันธ์ (Covariance and Correlation)

- ความแปรปรวนร่วมและสหสัมพันธ์ (Covariance and Correlation) คือ “**ความสัมพันธ์ระหว่างตัวแปรสุ่ม**” เป็นสิ่งที่นักวิเคราะห์ต้องการทราบ

Definition

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่ค่าคาดหมายมีค่าจำกัด และ $E[X] = \mu_X$ และ $E[Y] = \mu_Y$ ความแปรปรวนร่วม (covariance) ของ X และ Y นิยามได้เป็น

$$Cov[X, Y] = E[(X - \mu_X)(Y - \mu_Y)] \tag{1}$$

ถ้าค่าคาดหมายมีค่าจำกัด

- ความแปรปรวนร่วม (covariance) ก็คือค่าคาดหมายของฟังก์ชันของ “**ตัวแปรสุ่ม**” อย่างหนึ่งนั่นเอง

ทฤษฎี: ความแปรปรวนร่วม(Covariance)

Theorem

สำหรับตัวแปรสุ่ม X และ Y ที่ความแปรปรวนร่วม (covariance) มีค่าจำกัดใดๆ

$$Cov\left[X, Y\right] = E\left[XY\right] - E\left[X\right]E\left[Y\right] \tag{2}$$

Proof.

กำหนดให้ $E\left[X\right] = \mu_X$ และ $E\left[Y\right] = \mu_Y$

$$\begin{aligned} Cov\left[X, Y\right] = E\left[(X - \mu_X)(Y - \mu_Y)\right] &= E\left[XY - \mu_X Y - \mu_Y X + \mu_X \mu_Y\right] \\ &= E\left[XY\right] - \mu_X E\left[Y\right] - \mu_Y E\left[X\right] + \mu_X \mu_Y \\ &= E\left[XY\right] - \mu_X \mu_Y = E\left[XY\right] - E\left[X\right]E\left[Y\right] \end{aligned}$$

$\square$

ข้อสังเกตของความแปรปรวนร่วม (Covariance)

- ข้อจำกัดอย่างหนึ่งคือการที่ขนาดของความแปรปรวนร่วม (covariance) **ขึ้นอยู่กับขนาดของแต่ละตัวแปร**
- การเปรียบเทียบค่าความแปรปรวนร่วม (covariance) บอกได้**ยาก** ว่าตัวแปรคู่ใดมีความสัมพันธ์ระหว่างกัน (association) มากกว่ากัน
- เราต้องหานิยามค่าสถิติที่บอกถึง **ความสัมพันธ์แต่ไม่ขึ้นอยู่กับขนาดของแต่ละตัวแปรโดยตรง** หรือมีการปรับขนาด (re-scaling) นั่นเอง

นิยาม: สหสัมพันธ์ (Correlation)

Definition

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่ความแปรปรวน (variance) มีค่าจำกัด นั่นคือ $\sigma_X^2 < \infty$ และ $\sigma_Y^2 < \infty$ สหสัมพันธ์ (correlation) ของ X และ Y นิยามได้เป็น

$$\rho[X, Y] = \frac{Cov[X, Y]}{\sigma_X \sigma_Y} \tag{3}$$

- ค่าสมบูรณ์ของสหสัมพันธ์มีค่าน้อยกว่าหรือเท่ากับหนึ่งเสมอ

ทฤษฎี: สหสัมพันธ์ (Correlation)

Theorem

สำหรับตัวแปรสุ่ม X และ Y ใดๆ ที่ $E[XY]$ หาค่าได้

$$(E[XY])^2 \leq E\left[X^2\right] E\left[Y^2\right] \tag{4}$$

ความสัมพันธ์นี้เป็นจริงในรูปแบบสมการก็ต่อเมื่อมีจำนวนจริงที่ไม่เท่ากับศูนย์ $a \neq 0$ และ $b \neq 0$ ที่ทำให้ $aX + bY = 0$ ด้วยความน่าจะเป็นเท่ากับหนึ่ง

Proof.

- สำหรับกรณีที่ $E\left[X^2\right] = \infty$ หรือ $E\left[Y^2\right] = \infty$ ไม่จำเป็นต้องพิสูจน์อะไรมากเพราะพจน์ด้านขวาเท่ากับอนันต์ดังนั้นอสมการที่ 4 เป็นจริง
- ในกรณีที่ $E\left[X^2\right] = 0$ นั้นสามารถพิสูจน์ได้โดยเริ่มจาก $Pr(X = 0) = 1$ ดังนั้น $E[XY] = 0$ ซึ่งมีผลทำให้อสมการที่นี้ป็นจริงเช่นกัน
- $E\left[Y^2\right] = 0$ ก็สามารถพิสูจน์ได้ในทำนองเดียวกัน

□

พิสูจน์: สหสัมพันธ์ (Correlation)

Proof.

- กรณีที่ $0 < E\left[X^2\right] < \infty$ และ $0 < E\left[Y^2\right] < \infty$ ซึ่งสามารถพิสูจน์ได้โดยเริ่มจากการพิจารณาความสัมพันธ์ต่อไปนี้ สำหรับค่าจำนวนจริงที่ไม่เท่ากับศูนย์ a, b ใดๆ
  - $$E\left[(aX+bY)^2\right] = a^2E\left[X^2\right] + b^2E\left[Y^2\right] + 2abE[XY] \geq 0 \tag{5}$$
    $$E\left[(aX-bY)^2\right] = a^2E\left[X^2\right] + b^2E\left[Y^2\right] - 2abE[XY] \geq 0 \tag{6}$$
  - หากเลือก $a = \sqrt{E\left[Y^2\right]}$ และ $b = \sqrt{E\left[X^2\right]}$ จะสามารถเขียนสมการที่ 5 และ 6 ตามลำดับ ได้ใหม่เป็น
    $$E[XY] \geq -\sqrt{E\left[X^2\right]E\left[Y^2\right]} \tag{7}$$
    $$E[XY] \leq \sqrt{E\left[X^2\right]E\left[Y^2\right]} \tag{8}$$
    $$(E[XY])^2 \leq E\left[X^2\right]E\left[Y^2\right] \tag{9}$$
  - สรุปได้ว่า
    $$E\left[(aX+bY)^2\right] = 0 \Rightarrow Pr(aX+bY=0) = 1$$
    $$E\left[(aX-bY)^2\right] = 0 \Rightarrow Pr(aX-bY=0) = 1$$

ทฤษฎี: อสมการโคชีและชวาร์ซ (Cauchy-Schwarz Inequality)

Theorem (อสมการโคชีและชวาร์ซ (Cauchy-Schwarz Inequality))

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่ความแปรปรวน (variance) มีค่าจำกัด นั่นคือ $\sigma_X^2 < \infty$ *และ* $\sigma_Y^2 < \infty$ *สหสัมพันธ์ (correlation) ของ X และ Y มีค่าอยู่ระหว่าง* $[-1, 1]$ *นั่นคือ*

$$-1 \leq \rho\left[X, Y\right] \leq 1 \tag{10}$$

และความแปรปรวนร่วม (covariance) ของ X และ Y มีความสัมพันธ์กับความแปรปรวน (variance) ของแต่ละตัวแปรสุ่มดังนี้

$$\left(Cov\left[X, Y\right]\right)^2 \leq \sigma_X^2 \sigma_Y^2 \tag{11}$$

โดยที่ อสมการที่ จะเป็นจริงในรูปแบบสมการก็ต่อเมื่อ มีจำนวนจริงที่ไม่เท่ากับศูนย์ $a, b \neq 0$ *และจำนวนจริง* c *ที่ทำให้* $aX + bY = c$ *ด้วยความน่าจะเป็นเท่ากับ 1*

พิสูจน์: ทฤษฎีของอสมการโคชีและชวาร์ซ (Cauchy-Schwarz Inequality)

Proof.

ประยุกต์ใช้ทฤษฎีบทก่อนหน้านี้ เพื่อแสดงว่า

$$(Cov\left[X,Y\right])^2 = \left(E\left[(X-\mu_X)(Y-\mu_Y)\right]\right)^2 \le E\left[(X-\mu_X)^2\right]E\left[(Y-\mu_Y)^2\right] = \sigma_X^2\sigma_Y^2$$

ซึ่งจะเป็นจริงในรูปแบบสมการก็ต่อเมื่อมีจำนวนจริงที่ไม่เท่ากับศูนย์ $a \neq 0$ และ $b \neq 0$ ที่ทำให้ $aX + bY = 0$ ด้วยความน่าจะเป็นเท่ากับหนึ่ง เช่นเดียวกับกรณีของทฤษฎีก่อนหน้านี้ ยิ่งไปกว่านั้นความสัมพันธ์ที่ได้นำไปสู่ข้อสรุปที่ว่า

$$-\sigma_X\sigma_Y \le Cov\left[X,Y\right] \le \sigma_X\sigma_Y \Rightarrow -1 \le \frac{Cov\left[X,Y\right]}{\sigma_X\sigma_Y} \le 1 \Rightarrow -1 \le \rho\left[X,Y\right] \le 1$$

□

- “**ค่าสมบูรณ์ของสหสัมพันธ์มีค่าน้อยกว่าหรือเท่ากับหนึ่ง**” เสมอ

นิยาม: สหสัมพันธ์เชิงบวกและเชิงลบ (positively and negatively correlated)

Definition

ตัวแปรสุ่ม X และ Y มีสหสัมพันธ์เชิงบวก (positively correlated) ถ้า $\rho[X, Y] > 0$ และมีสหสัมพันธ์เชิงลบ (negatively correlated) ถ้า $\rho[X, Y] < 0$ ส่วนกรณีที่ $\rho[X, Y] = 0$ หมายความว่า X และ Y ไม่มีสหสัมพันธ์ (uncorrelated)

- ความสัมพันธ์ระหว่างความเป็นอิสระต่อกัน (independent) และการไม่มีสหสัมพันธ์ (uncorrelated) ซึ่งกล่าวว่า "**ตัวแปรสุ่มที่เป็นอิสระต่อกันต้องไม่มีสหสัมพันธ์**"

ทฤษฎี: การไม่มีสหสัมพันธ์ (uncorrelated)

Theorem

ถ้าตัวแปรสุ่ม X และ Y เป็นอิสระต่อกัน (independent) และ $\sigma_X^2 < \infty$ และ $\sigma_Y^2 < \infty$ แล้ว

$$Cov\,[X, Y] = \rho\,[X, Y] = 0 \tag{12}$$

Proof.

เนื่องจาก X และ Y เป็นอิสระต่อกัน (independent)

- $E\,[XY] = E\,[X]\,E\,[Y]$ ซึ่งหมายความว่า $E\,[XY] - E\,[X]\,E\,[Y] = Cov\,[X, Y] = 0$
- ผลที่ตามมาอีกอย่างหนึ่งก็คือ $\rho\,[X, Y] = 0$

$\square$

- ตัวแปรสุ่มที่ไม่มีสหสัมพันธ์อาจจะไม่เป็นอิสระต่อกันก็ได้
- สะท้อนให้เห็นว่า การเป็นอิสระต่อกัน (independent) นั้นเป็น "**ข้อจำกัดที่เข้มข้นมากกว่า**" การไม่มีสหสัมพันธ์ (uncorrelated)

ตัวอย่าง: ค่าสหสัมพันธ์ และ การเป็นอิสระต่อกัน

Example

กำหนดให้ X เป็นตัวแปรสุ่มที่มีค่าที่เป็นไปได้สามค่าคือ $\{-1, 0, 1\}$ และมีความน่าจะเป็นเท่ากับ

$$Pr\left(X = -1\right) = Pr\left(X = 0\right) = Pr\left(X = 1\right) = \frac{1}{3}$$

สมมุติให้ $Y = X^2$ ดังนั้น X และ Y ไม่เป็นอิสระต่อกันอย่างชัดเจน ส่วนที่เหลืออยู่คือการตรวจสอบว่า X และ Y มีสหสัมพันธ์ต่อกันหรือไม่?

- หาว่า $E\left[X\right] = 0$
- ค่าสหสัมพันธ์ของ X และ Y

$$\begin{aligned} Cov\left[X, Y\right] &= E\left[XY\right] - E\left[X\right]E\left[Y\right] = E\left[XY\right] \\ &= Pr\left(X = -1, Y = 1\right)\left(-1\right) + Pr\left(X = 0, Y = 0\right)\left(0\right) \\ &\quad + Pr\left(X = 1, Y = 1\right)\left(1\right) \\ &= -\frac{1}{3} + 0 + \frac{1}{3} = 0 \end{aligned}$$

- X และ Y ไม่มีสหสัมพันธ์ (uncorrelated) ทั้งที่ไม่เป็นอิสระต่อกัน

ข้อสังเกต: ค่าสหสัมพันธ์ และ การเป็นอิสระต่อกัน

- ตัวอย่างนี้คือความสัมพันธ์ระหว่าง X และ Y เป็นแบบไม่เป็นเชิงเส้น (nonlinear)
  - มีส่วนสำคัญที่ทำให้เกิดปรากฏการณ์ที่ตัวแปรสุ่มสองตัวที่ไม่เป็นอิสระต่อกัน แต่กลับไม่มีสหสัมพันธ์ (uncorrelated)
- แต่ถ้าความสัมพันธ์ของตัวแปรเป็นแบบเชิงเส้น (linear) จะสามารถสรุปได้ว่า **ตัวแปรสุ่มสองตัวที่ไม่เป็นอิสระต่อกันจะมีสหสัมพันธ์ต่อกัน**

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มที่ $0 < \sigma_X^2 < \infty$ และ $Y = aX + b$ สำหรับค่าคงที่ $a \neq 0$ และ b แล้วสามารถสรุปได้ว่า

1. *ถ้า $a > 0$ แล้ว $\rho[X, Y] = 1$*
2. *ถ้า $a < 0$ แล้ว $\rho[X, Y] = -1$*

สหสัมพันธ์ใช้วัดความสัมพันธ์เชิงเส้นเป็นหลัก

Theorem

กำหนดให้ X และ Y เป็นตัวแปรสุ่มที่ความแปรปรวน (variance) มีค่าจำกัด นั่นคือ $\sigma_X^2 < \infty$ และ $\sigma_Y^2 < \infty$ ถ้า $|\rho[X,Y]| = 1$ แล้ว จะต้องมีค่าคงที่ $a \neq 0$, $b \neq 0$, และ c ที่ทำให้ $aX + bY = c$ ด้วยความน่าจะเป็นเท่ากับหนึ่ง

Proof.

เริ่มจากอสมการโคชีและชวาร์ซ (Cauchy-Schwarz Inequality)

$$(Cov[X,Y])^2 \leq \sigma_X^2 \sigma_Y^2$$

ซึ่งจะเป็นจริงในรูปแบบสมการก็ต่อเมื่อมีจำนวนจริงที่ไม่เท่ากับศูนย์ $a, b \neq 0$ และจำนวนจริง c ที่ทำให้ $aX + bY = c$ ด้วยความน่าจะเป็นเท่ากับ 1 ซึ่งหมายความว่า ถ้า

$$(Cov[X,Y])^2 = \sigma_X^2 \sigma_Y^2$$

ก็ต่อเมื่อมีจำนวนจริงที่ไม่เท่ากับศูนย์ $a, b \neq 0$ และจำนวนจริง c ที่ทำให้ $aX + bY = c$ ด้วยความน่าจะเป็นเท่ากับ 1 $\square$

ทฤษฎี: ผลรวมเชิงเส้นของความแปรปรวนร่วมสำหรับตัวแปรสุ่ม 2 ตัว

Theorem

ถ้า X และ Y เป็นตัวแปรสุ่มที่ความแปรปรวน (variance) มีค่าจำกัด นั่นคือ $\sigma_X^2 < \infty$ และ $\sigma_Y^2 < \infty$ แล้ว

$$Var\left[aX + bY + c\right] = a^2 Var\left[X\right] + b^2 Var\left[Y\right] + 2abCov\left[X, Y\right] \tag{13}$$

- ทฤษฎีบทต่อไปนี้แสดงวิธีการคำนวณค่าความแปรปรวนของผลรวมของตัวแปรสุ่มสองตัว ซึ่งรวมทั้งกรณีที่ตัวแปรสุ่มทั้งสองมีสหสัมพันธ์ต่อกันด้วย

ทฤษฎี: ผลรวมเชิงเส้นของความแปรปรวนร่วมสำหรับตัวแปรสุ่มหลายตัว

Theorem

ถ้า $X_1, \ldots, X_n$ *เป็นตัวแปรสุ่มที่ความแปรปรวน (variance) มีค่าจำกัด นั่นคือ* $Var[X_i] < \infty$ *แล้ว*

$$Var\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 Var[X_i] + 2a_i a_j \sum_{i<j} Cov[X_i, X_j] \tag{14}$$

Proof.

$$Var\left[\sum_{i=1}^{n} a_i X_i\right] = Cov\left[\sum_{i=1}^{n} a_i X_i, \sum_{j=1}^{n} a_j X_j\right] = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j Cov[X_i, X_j]$$

$$= \sum_{i=1}^{n} a_i^2 Cov[X_i, X_i] + \sum_{i \neq j} a_i a_j Cov[X_i, X_j]$$

$$= \sum_{i=1}^{n} a_i^2 Var[X_i] + 2a_i a_j \sum_{i<j} Cov[X_i, X_j]$$

ตัวอย่าง: ขอบเขตของค่าคาดหมายและความแปรปรวนของกลุ่มหลักทรัพย์ (mean-variance frontier)

Example (ขอบเขตของค่าคาดหมายและความแปรปรวนของกลุ่มหลักทรัพย์ (mean-variance frontier))

กำหนดให้ R_i แทนอัตราผลตอบแทนรวม (gross return) ต่อปีของกองทุนรวม (mutual fund) i สำหรับ $i = 1, \ldots, n$ และกำหนดให้ $\alpha_i \in [0, 1]$ โดยที่ $\sum_i \alpha_i = 1$ แทนสัดส่วนการลงทุนในกองทุน i ของกลุ่มหลักทรัพย์ (portfolio) ที่สนใจ

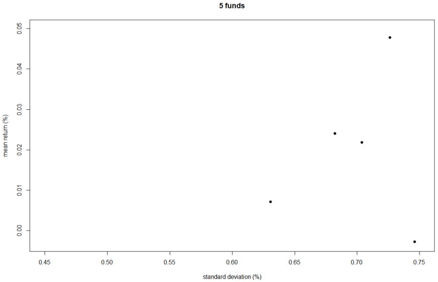
สิ่งที่ต้องการทราบคือกลุ่มหลักทรัพย์ (portfolio) คืออัตราผลตอบแทนรวมของกลุ่มหลักทรัพย์ (portfolio) หลังจากลงทุนเป็นเวลาหนึ่งปี

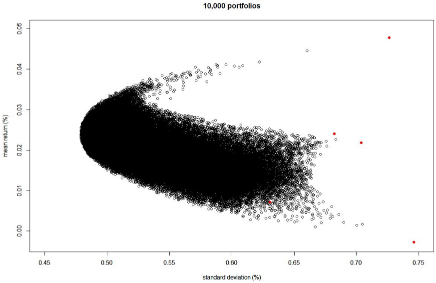
$$R = \sum_{i=1}^{n} \alpha_i R_i$$

ซึ่งเป็นตัวแปรสุ่มเพราะเป็นฟังก์ชันของตัวแปรสุ่ม

รูปแสดงขอบเขตของค่าคาดหมายและความแปรปรวนของกลุ่มหลักทรัพย์ (mean-variance frontier)

- ยกตัวอย่างกองทุนรวม 5 กองทุน ซึ่งประกอบไปด้วย ABSM, 1AMSET50, BTP, CGLTF, ABAG และใช้ข้อมูลอัตราผลตอบแทนรวมสุทธิตั้งแต่11/9/2017 จนถึง 11/9/2018 และแสดงเป็นกราฟความสัมพันธ์และจัดกลุ่มหลักทรัพย์ได้คือ





โมเมนต์และฟังก์ชันก่อกำเนิดโมเมนต์ (Moments and Moment Generating Function)

- ที่ผ่านมา คุณสมบัติทั้งหมดของตัวแปรสุ่มในรูปของ **ฟังก์ชันการแจกแจง (distribution function)** ไม่ว่าจะเป็นในรูปของฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) หรือฟังก์ชันความน่าจะเป็นสะสม (C.D.F.)
- "**ค่าคาดหมาย (expectation)**" เป็นเครื่องมือที่ช่วย "**สรุปคุณสมบัติของการแจกแจง**" ให้เข้าใจได้ง่ายและสะดวกมากยิ่งขึ้น **แต่** ค่าคาดหมาย (expectation) อันเดียว "**ไม่สามารถบ่งบอก**" ถึงคุณสมบัติของการแจกแจงได้ครบถ้วน
- หากรวบรวมเอาค่าคาดหมายของฟังก์ชันยกกำลังของตัวแปรสุ่มสำหรับทุกๆ ค่ายกกำลังที่เป็นจำนวนเต็มบวกแล้วก็จะได้คุณสมบัติที่ครบถ้วนตามที่ต้องการ นักสถิติมักเรียกค่าคาดหมายของฟังก์ชันยกกำลังของตัวแปรสุ่มว่า **ค่าโมเมนต์ (moment)**

โมเมนต์และฟังก์ชันก่อกำเนิดโมเมนต์ (Moments and Moment Generating Function)

- ที่สำคัญเราสามารถรวบรวมเอาค่าโมเมนต์ทั้งหมดนั้นมาไว้ด้วยกันในรูปของฟังก์ชัน โดยเรียกฟังก์ชันนี้ว่า **ฟังก์ชันก่อกำเนิดโมเมนต์ (moment generating function หรือ m.g.f.)**
- การแจกแจงอันใดอันหนึ่ง **มีฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ได้เพียงอันเดียว (unique)**
- โมเมนต์ (moment) ที่ k ของ X คือค่าคาดหมายของ X^k ซึ่งแทนด้วย $E\left[X^k\right]$
    - $\mu = E\left[X\right]$ แทนค่าคาดหมายของ X
- ส่วน $E\left[(X-\mu)^k\right]$ นั้นจะถูกเรียกว่า **โมเมนต์ศูนย์กลาง (central moment)** ที่ k ของ X
    - โมเมนต์ศูนย์กลางที่สองก็คือค่าความแปรปรวน $E\left[(X-\mu)^2\right] = Var(X)$
    - ส่วนโมเมนต์ศูนย์กลางที่สามนั้นใช้เพื่อบ่งบอกถึงความสมมาตร (symmetry) ของการแจกแจงซึ่งวัดด้วยค่าสถิติที่เรียกว่าความเบ้ (skewness) ซึ่งนิยามได้เป็น $\frac{E\left[(X-\mu)^3\right]}{\sigma^3}$

ทฤษฎี: โมเมนต์ที่มีค่าจำกัด

- ทฤษฎีบทนี้แสดงว่าหากโมเมนต์ที่ k มีค่าจำกัด ค่าโมเมนต์ที่ต่ำกว่านั้น (สำหรับ $j < k$) จะมีค่าจำกัดเสมอ หมายความว่า หากเราสามารถหาค่าโมเมนต์ในอันดับที่สูงของการแจกแจงได้ **เราย่อมสามารถหาค่าโมเมนต์ในอันดับที่ต่ำลงมาได้เสมอ**

Theorem

ถ้า $E\left[|X|^k\right] < \infty$ สำหรับค่าจำนวนเต็มบวก k แล้ว $E\left[|X|^j\right] < \infty$ สำหรับค่าจำนวนเต็มบวก j ใดๆ ที่ $j < k$

Proof.

กำหนดให้ j และ k โดยที่ $j < k$

$$\begin{aligned}
E\left[|X|^j\right] &= \int_{-\infty}^{\infty} |x|^j f(x)\,dx = \int_{|x|\le 1} |x|^j f(x)\,dx + \int_{|x|>1} |x|^j f(x)\,dx \\
&\le \int_{|x|\le 1} f(x)\,dx + \int_{|x|>1} |x|^j f(x)\,dx \le \int_{|x|\le 1} f(x)\,dx + \int_{|x|>1} |x|^k f(x)\,dx \\
&\le \int_{|x|\le 1} f(x)\,dx + \int_{|x|>1} |x|^k f(x)\,dx + \int_{|x|\le 1} |x|^k f(x)\,dx \\
&= Pr\left(|X| \le 1\right) + E\left[|X|^k\right] < \infty
\end{aligned}$$

$\square$

นิยาม: ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.)

Definition

กำหนดให้ X เป็นตัวแปรสุ่ม สำหรับค่าจำนวนจริง t ใดๆ นิยามฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X เป็น

$$\psi(t) = E\left[e^{tX}\right] \tag{15}$$

- ทฤษฎีบทนี้นำเสนอวิธีการคำนวณหาค่าโมเมนต์ต่างๆ จากฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.)

Theorem

กำหนดให้ X เป็นตัวแปรสุ่มที่ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) มีค่าจำกัดทุกๆ ค่า t ที่อยู่รอบๆ ศูนย์ (t in some open interval around zero) ดังนั้น สำหรับค่าจำนวนเต็มบวก n ใดๆ โมเมนต์ที่ n ของ X มีค่าเท่ากับอนุพันธ์ลำดับที่ n ของฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) $\psi(t)$ ที่ $t = 0$ นั่นคือ

$$E[X^n] = \psi^n(0) \tag{16}$$

โดยที่ $\psi^n(0)$ หมายถึงอนุพันธ์ลำดับที่ n ของฟังก์ชัน $\psi(t)$ ที่ $t = 0$

ทฤษฎี: ฟังก์ชันก่อกำเนิดโมเมนต์ของฟังก์ชันเชิงเส้นของตัวแปรสุ่ม

Theorem

กำหนดให้ X *เป็นตัวแปรสุ่มที่ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) เท่ากับ* $\psi_x(t)$ *และ* $Y = aX + b$ *สำหรับจำนวนจริง* a *และ* b *ใดๆ แล้ว ตัวแปรสุ่มที่ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ* Y *เท่ากับ*

$$\psi_y(t) = e^{bt}\psi_x(at) \tag{17}$$

Proof.

$$\psi_y(t) = E\left[e^{tY}\right] = E\left[e^{t(aX+b)}\right] = e^{bt}E\left[e^{(at)X}\right] = e^{bt}\psi_x(at)$$

$\square$

ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของตัวแปรสุ่มที่เป็นอิสระต่อกัน (independent)

ทฤษฎีบทต่อไปนี้แสดงถึงวิธีการหาฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของตัวแปรสุ่มที่เกิดจากการบวกกันของตัวแปรสุ่มที่เป็นอิสระต่อกัน (independent)

Theorem

สมมุติให้ $X_1, \ldots, X_n$ เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน (independent) ที่มีฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของแต่ละตัวแปรเท่ากับ $\psi_i(t)$ สำหรับ $i = 1, \ldots, n$ และกำหนดให้ $Y = \sum_{i=1}^{n} X_i$ ดังนั้น ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Y เท่ากับ

$$\psi_Y(t) = \prod_{i=1}^{n} \psi_i(t) \tag{18}$$

นิยาม: ค่ามัธยฐาน (median)

- นิยามของค่ามัธยฐาน (median) จำเป็นต้องกำหนดให้ครอบคลุมทั้งกรณีของตัวแปรสุ่มต่อเนื่องและไม่ต่อเนื่อง ดังนี้

Definition

ค่ามัธยฐาน (median) m ของการแจกแจงของตัวแปรสุ่ม X ต้องสอดคล้องกับเงื่อนไขต่อไปนี้

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

ตัวอย่าง: ในกรณีที่มีค่ามัธยฐานค่าเดียวและ มีค่าที่ชัดเจน

ตัวอย่างต่อไปนี้ช่วยให้เข้าใจนิยามของค่ามัธยฐาน (median) ได้ดียิ่งขึ้น

Example

พิจารณาตัวแปรสุ่มไม่ต่อเนื่อง X ที่มีการแจกแจงดังนี้

$$Pr\,(X=1)=0.2, Pr\,(X=3)=0.4, Pr\,(X=5)=0.1, Pr\,(X=7)=0.3$$

ในกรณีนี้ จะเห็นได้ว่าค่ามัธยฐาน $m=3$ เพราะ

$$Pr\,(X\leq 3)=0.6\geq 0.5 \text{ และ } Pr\,(X\geq 3)=0.8\geq 0.5$$

ยิ่งไปกว่านั้น ค่าที่เป็นไปได้ของตัวแปรสุ่ม X ค่าอื่นไม่สามารถเป็นค่ามัธยฐานได้ ดังนั้น $m=3$ จึงเป็นค่ามัธยฐานเพียงค่าเดียว (unique)

ตัวอย่าง: ในกรณีที่มีค่ามัธยฐานค่าเดียว แต่มีค่าอยู่ระหว่างค่าของตัวแปรสุ่ม

Example

พิจารณาตัวแปรสุ่มไม่ต่อเนื่อง X ที่มีการแจกแจงต่างออกไปจากตัวอย่างก่อนหน้านี้เล็กน้อย

$$Pr\left(X=1\right)=0.1, Pr\left(X=3\right)=0.4, Pr\left(X=5\right)=0.2, Pr\left(X=7\right)=0.3$$

ซึ่งจะเห็นได้ว่า

$$Pr\left(X\leq 3\right)=0.5 \text{ และ } Pr\left(X\geq 5\right)=0.5$$

ดังนั้น ค่าจำนวนจริง m ที่อยู่ระหว่างช่วง $3\leq m\leq 5$ ล้วนเป็นค่ามัธยฐานทั้งสิ้นเพราะสอดคล้องกับสมการที่ 19 กล่าวคือ ค่ามัธยฐานในกรณีนี้ไม่มากกว่าหนึ่งค่า (not unique) ในทางปฏิบัติ จึงมักนิยมที่จะใช้ค่ากึ่งกลางของช่วงดังกล่าวแทนค่ามัธยฐาน ซึ่งในกรณีมีค่าเท่ากับ 4

ตัวอย่าง: ค่ามัธยฐานในกรณีที่มีค่ามัธยฐานมากกว่าหนึ่งค่า

กรณีที่มีค่ามัธยฐานมากกว่าหนึ่งค่าสามารถเกิดขึ้นกับการแจกแจงแบบต่อเนื่องได้เช่นกัน ดังแสดงในตัวอย่างต่อไปนี้

Example

- พิจารณาตัวแปรสุ่ม X ที่มีการแจกแจงต่อเนื่องและมีฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) เท่ากับ

$$f(x) = \begin{cases} \frac{1}{4} \text{ สำหรับ } -1 \leq x \leq 1 \\ \frac{1}{2}, \text{ สำหรับ } 3 \leq x \leq 4 \\ 0, \text{ สำหรับกรณีอื่น} \end{cases}$$

- สามารถใช้คำนวณหาฟังก์ชันความน่าจะเป็นสะสม (C.D.F.) ได้เป็น

$$F(x) = \begin{cases} 0, \text{ สำหรับ } -\infty \leq x \leq -1 \\ \frac{1}{4} + \frac{1}{4}x, \text{ สำหรับ } -1 \leq x \leq 1 \\ \frac{1}{2}, \text{ สำหรับ } 1 \leq x \leq 3 \\ -1 + \frac{1}{2}x, \text{ สำหรับ } 3 \leq x \leq 4 \\ 1, \text{ สำหรับ } 4 \leq x \leq \infty \end{cases}$$

กราฟแสดงค่ามัธยฐานในกรณีที่มีค่ามัธยฐานมากกว่าหนึ่งค่า

- ดูรูปประกอบ จะเห็นได้ว่า ค่าจำนวนจริง m ที่อยู่ระหว่าง $1 \leq m \leq 3$ ล้วนแต่มีผลทำให้ $Pr\left(X \leq m\right) \geq \frac{1}{2}$ และ $Pr\left(X \geq m\right) \geq \frac{1}{2}$ นั่นหมายความว่า ค่าจำนวนจริงเหล่านี้ล้วนเป็นค่ามัธยฐาน
- ในทำนองเดียวกับตัวอย่างก่อนหน้านี้ ในทางปฏิบัติมักนิยมที่จะใช้ค่ากึ่งกลางของช่วงดังกล่าวแทนค่ามัธยฐาน ซึ่งในกรณีมีค่าเท่ากับ 2