การแจกแจงที่เป็นที่นิยม (Popular Distributions)

รศ.ดร. วีระชาติ กิเลนทอง
มหาวิทยาลัยหอการค้าไทย

ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของการแจกแจงปกติมาตรฐานหลายตัวแปร (Multivariate Normal Distribution)

- สิ่งที่ขาดหายไปจากฟังก์ชันความหนาแน่นความน่าจะเป็น (p.d.f.) ของตัวแปรสุ่มตัวเดียว คือ **ความแปรปรวน (covariance) หรือสหสัมพันธ์ (correlation) ซึ่งจะปรากฏในกรณีที่มีตัวแปรหลายตัว**
- เริ่มจากตัวแปรสุ่มปกติมาตรฐานที่เป็นอิสระต่อกัน ซึ่งเขียนแทนด้วยเวคเตอร์สุ่ม (random vector) $Z = [Z_1, \ldots, Z_n]'$ โดยที่ X' หมายถึง เมทริกซ์สลับเปลี่ยน (transpose matrix) ของ X และ $Z_i \sim N(0,1)$ มีการแจกแจงปกติมาตรฐาน (standard normal) และเป็นอิสระต่อกัน (independent)
- นั่นคือตัวแปรสุ่มเหล่านี้มีการแจกแจงเหมือนกันและเป็นอิสระต่อกัน (identically and independently distributed หรือ i.i.d.) ซึ่งในกรณีที่การแจกแจงเป็นแบบปกติมาตรฐาน สามารถเขียนแทนด้วย $Z \sim N(\mathbf{0}, I_n)$
- ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ Z คือ

$$f(z) = \prod_{i=1}^{n} \phi(z_i) = \prod_{i=1}^{n} \left(\frac{1}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2}z_i^2\right\} = \left(\frac{1}{2\pi}\right)^{\frac{n}{2}} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} z_i^2\right\}$$

$$= \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2}z'z\right\} \tag{1}$$

ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของการแจกแจงปกติมาตรฐานหลายตัวแปร (Multivariate Normal Distribution)

- สมการข้างต้นเป็นผลมาจากการที่ตัวแปรทั้งหมดมีการแจกแจงปกติมาตรฐานเหมือนกันและเป็นอิสระต่อกัน (i.i.d.)
- ชัดเจนว่า ค่าคาดหมาย (mean) และเมทริกซ์ของค่าความแปรปรวนร่วม (covariance matrix) ของ Z มีค่าเท่ากับ $\mathbf{0}$ และ I_n ตามลำดับ
- ส่วนฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Z สามารถเขียนได้เป็น

$$\begin{aligned}\psi_{\mathbf{z}}(t) &= E\left[\exp\left\{\mathbf{Z}'\mathbf{t}\right\}\right] = E\left[\exp\left\{\sum_{i=1}^{n} Z_i t_i\right\}\right] = E\left[\prod_{i=1}^{n}\exp\{Z_i t_i\}\right] \\ &= \prod_{i=1}^{n} E\left[\exp\{Z_i t_i\}\right] = \prod_{i=1}^{n}\psi_z(t_i) = \prod_{i=1}^{n} e^{\frac{1}{2}t_i^2} \\ &= \exp\left\{\frac{1}{2}\sum_{i=1}^{n} t_i^2\right\} = \exp\left\{\frac{1}{2}\mathbf{t}'\mathbf{t}\right\} \end{aligned} \qquad (2)$$

- เราสามารถสร้างเวคเตอร์ของตัวแปรสุ่มปกติ (normal random vector) X ที่มีค่าคาดหมาย (mean) เท่ากับ $\boldsymbol{\mu}$ และเมทริกซ์ของค่าความแปรปรวนร่วม (covariance matrix) เท่ากับ $\boldsymbol{\Sigma}$

การสร้างเวคเตอร์ของตัวแปรสุ่มปกติ (normal random vector) ของการแจกแจงปกติหลายตัวแปร (Multivariate Normal Distribution)

- ในการสร้างเวคเตอร์ของตัวแปรสุ่มปกติ (normal random vector) จะคล้ายกับตัวแปรสุ่มตัวเดียวที่มีการแจกแจงปกติแบบมาตรฐาน แต่มี **ความยุ่งยากทางเทคนิค** ในการนิยามค่าความเบี่ยงเบนในรูปของเมทริกซ์ ซึ่งต้องใช้เทคนิคที่เรียกว่า **การแยกส่วนสเปกตรัล (spectral decomposition)**
- เริ่มจากคุณสมบัติที่สำคัญอันหนึ่งของเมทริกซ์ของค่าความแปรปรวนร่วม (variance-covariance matrix) $\boldsymbol{\Sigma}$ คือ มันเป็น **เมทริกซ์จัตุรัสและสมมาตร (symmetric square matrix)** และ **คุณสมบัติเป็นบวกเกือบแน่นอน (positive semidefinite)** ซึ่งทำให้สามารถใช้หลักการทางพีชคณิตเชิงเส้น (linear algebra) แยกส่วน $\boldsymbol{\Sigma}$ ออกได้เป็น

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}' \boldsymbol{\Lambda} \boldsymbol{\Gamma} \tag{3}$$

โดยที่ $\boldsymbol{\Lambda}$ เป็นเมทริกซ์แนวทแยง (diagonal matrix)

$$\boldsymbol{\Lambda} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 & 0 \\ 0 & \lambda_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_{n-1} & 0 \\ 0 & 0 & \cdots & 0 & \lambda_n \end{bmatrix} \tag{4}$$

และ $\boldsymbol{\Gamma}$ เป็นเมทริกซ์ตั้งฉาก (orthogonal matrix) นั่นคือ $\boldsymbol{\Gamma}^{-1} = \boldsymbol{\Gamma}'$

- สรุปได้ว่า $\boldsymbol{\Gamma}'\boldsymbol{\Gamma} = I$ โดยทั่วไปเพื่อความสะดวก เรามักจะเรียงลำดับ λ_i จากมากไปหาน้อย นั่นคือ $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ และเรียก λ_i แต่ละตัวว่า ค่าลักษณะเฉพาะ (eigenvalue) ส่วนคอลัมน์ที่ i แต่ละอัน (สอดคล้องกับ λ_i) ซึ่งแทนด้วย $\boldsymbol{\gamma}_i$ มักเรียกว่าเวคเตอร์ลักษณะเฉพาะ (eigenvector) ของ $\boldsymbol{\Sigma}$

การสร้างเวคเตอร์ของตัวแปรสุ่มปกติ (normal random vector) ของการแจกแจงปกติหลายตัวแปร (Multivariate Normal Distribution) Con’ t

- จากการที่ เราสามารถเขียนเมทริกซ์ของค่าความแปรปรวนร่วม (variance-covariance matrix) $\boldsymbol{\Sigma}$ ในรูปของค่าลักษณะเฉพาะ (eigenvalue) และเวคเตอร์ลักษณะเฉพาะ (eigenvector) ได้ดังนี้

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}' \boldsymbol{\Lambda} \boldsymbol{\Gamma} = \sum_{i=1}^{n} \lambda_i \gamma_i \gamma_i' \tag{5}$$

- ประโยชน์ที่สำคัญของการแยกส่วนนี้ก็คือ การได้มาซึ่งเมทริกซ์แนวทแยง (diagonal matrix) $\boldsymbol{\Lambda}$ ซึ่งค่าลักษณะเฉพาะ (eigenvalue) λ_i แต่ละค่ามีค่าไม่ติดลบ (nonnegative) ซึ่งช่วยให้สามารถนิยามรากที่สอง (square root) ของ $\boldsymbol{\Lambda}$ ได้เป็น

$$\boldsymbol{\Lambda}^{\frac{1}{2}} = \begin{bmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_{n-1}} & 0 \\ 0 & 0 & \cdots & 0 & \sqrt{\lambda_n} \end{bmatrix} \tag{6}$$

- ซึ่งมีคุณสมบัติคล้ายกับรากกำลังสองของจำนวนจริง นั่นคือ

$$\boldsymbol{\Lambda}^{\frac{1}{2}} \boldsymbol{\Lambda}^{\frac{1}{2}} = \boldsymbol{\Lambda} \tag{7}$$

การสร้างเวคเตอร์ของตัวแปรสุ่มปกติ (normal random vector) ของการแจกแจงปกติหลายตัวแปร (Multivariate Normal Distribution) Con' t

- เราสามารถเขียนสมการการแยกส่วนสเปกตรัล (spectral decomposition) ใหม่ได้เป็น

$$\boldsymbol{\Sigma} = \boldsymbol{\Gamma}'\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Gamma} = \left(\boldsymbol{\Gamma}'\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Gamma}\right)\left(\boldsymbol{\Gamma}'\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Gamma}\right) \tag{8}$$

- ใช้คุณสมบัติการตั้งฉาก (orthogonal) ของเมทริกซ์ $\boldsymbol{\Gamma}$ นั่นคือ $\boldsymbol{\Gamma}\boldsymbol{\Gamma}' = I$ ผลที่ตามมาจากสมการที่ 8 คือ ทำให้สามารถนิยามรากที่สอง (square root) ของ $\boldsymbol{\Sigma}$ ได้เป็น

$$\boldsymbol{\Sigma}^{\frac{1}{2}} = \boldsymbol{\Gamma}'\boldsymbol{\Lambda}^{\frac{1}{2}}\boldsymbol{\Gamma} \tag{9}$$

- หากเมทริกซ์ของค่าความแปรปรวนร่วม (variance-covariance matrix) $\boldsymbol{\Sigma}$ เป็นบวกแน่นอน (positive definite หรือ p.d.) แล้วค่าลักษณะเฉพาะ (eigenvalue) มีค่าเป็นบวกทุกค่า นั่นคือ $\lambda_i > 0$ สำหรับทุกๆ $i = 1, \ldots, n$ ซึ่งมีผลทำให้ค่าดีเทอร์มิแนนต์ (determinant) ของ $\boldsymbol{\Lambda}^{\frac{1}{2}}$ มีค่ามากกว่าศูนย์อย่างแน่นอน ทำให้สามารถหาเมทริกซ์ส่วนกลับ (inverse matrix) ของ $\boldsymbol{\Lambda}^{\frac{1}{2}}$ ได้ ดังนั้น จึงสามารถนิยามเมทริกซ์ส่วนกลับ (inverse matrix) ของ $\boldsymbol{\Sigma}^{\frac{1}{2}}$ ได้เป็น

$$\boldsymbol{\Sigma}^{-\frac{1}{2}} = \boldsymbol{\Gamma}'\boldsymbol{\Lambda}^{-\frac{1}{2}}\boldsymbol{\Gamma} \tag{10}$$

โดยที่ $\boldsymbol{\Lambda}^{-\frac{1}{2}}$ แทนเมทริกซ์ส่วนกลับ (inverse matrix) ของ $\boldsymbol{\Lambda}^{\frac{1}{2}}$

ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของการแจกแจงปกติหลายตัวแปร (Multivariate Normal Distribution)

- สร้างเวคเตอร์ของตัวแปรสุ่มปกติ (normal random vector) $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ ที่มีค่าคาดหมาย (mean) เท่ากับ $\boldsymbol{\mu}$ และเมทริกซ์ของค่าความแปรปรวนร่วม (covariance matrix) เท่ากับ $\boldsymbol{\Sigma}$ โดยกำหนดให้

$$X = \boldsymbol{\Sigma}^{\frac{1}{2}} Z + \boldsymbol{\mu} \tag{11}$$

- ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X เท่ากับ

$$\begin{aligned} \psi_{\mathbf{x}}(t) &= E\left[\exp\left\{X' t\right\}\right] = E\left[\exp\left\{\left(\boldsymbol{\Sigma}^{\frac{1}{2}} Z + \boldsymbol{\mu}\right)' t\right\}\right] = \exp\left\{\boldsymbol{\mu}' t\right\} E\left[\exp\left\{Z' \left(\boldsymbol{\Sigma}^{\frac{1}{2}} t\right)\right\}\right] \\ &= \exp\left\{\boldsymbol{\mu}' t\right\} \exp\left\{\frac{1}{2}\left(\boldsymbol{\Sigma}^{\frac{1}{2}} t\right)' \left(\boldsymbol{\Sigma}^{\frac{1}{2}} t\right)\right\} = \exp\left\{\boldsymbol{\mu}' t + \frac{1}{2} t' \boldsymbol{\Sigma} t\right\} \end{aligned} \tag{12}$$

- ประยุกต์ใช้ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Z ในสมการที่ 2 โดยใช้ $\boldsymbol{\Sigma}^{\frac{1}{2}} t$ แทน t สังเกตด้วยว่า ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ในที่นี้สามารถหาค่าได้ตราบเท่าที่เมทริกซ์ของค่าความแปรปรวนร่วม (variance-covariance matrix) $\boldsymbol{\Sigma}$ เป็นบวกเกือบแน่นอน (p.s.d.) ในขณะที่ ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ X จะหาค่าได้ก็ต่อเมื่อ $\boldsymbol{\Sigma}$ เป็นบวกแน่นอน (p.d.) ทั้งนี้เพราะการแปลงฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) จำเป็นต้องใช้ค่าดีเทอร์มิแนนต์ (determinant) ของ $\boldsymbol{\Sigma}^{-\frac{1}{2}}$

ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของการแจกแจงปกติหลายตัวแปร (Multivariate Normal Distribution)

หาก $\boldsymbol{\Sigma}$ เป็นบวกแน่นอน (p.d.) แล้ว จะสามารถเขียนสมการที่ 11 ใหม่ได้เป็น

$$Z = \boldsymbol{\Sigma}^{-\frac{1}{2}} \left(X - \boldsymbol{\mu} \right) \tag{13}$$

ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็น (p.d.f.) ของ X คือ

$$f(x) = \frac{1}{(2\pi)^{\frac{n}{2}} \left| \det \boldsymbol{\Sigma} \right|^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2} \left(x - \boldsymbol{\mu} \right)' \boldsymbol{\Sigma}^{-1} \left(x - \boldsymbol{\mu} \right) \right\} \tag{14}$$

ทฤษฎี: ความสัมพันธ์เชิงเส้นของการแจกแจงปกติหลายตัวแปร (Multivariate Normal Distribution)

Theorem

กำหนดให้ X เป็นเวคเตอร์ของตัวแปรสุ่ม n ตัว ซึ่งมีการแจกแจงร่วมแบบปกติ นั่นคือ $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ และ $Y = AX + b$ โดยที่ A เป็นเมทริกซ์ของค่าคงที่ขนาด $m \times n$ และ b เป็นเวคเตอร์ของค่าคงที่ m ตัว แล้ว $Y \sim N\left(A\boldsymbol{\mu} + b, A\boldsymbol{\Sigma}A'\right)$

Proof.

ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Y เท่ากับ

$$\begin{aligned}
\psi_Y(t) &= E\left[\exp\left\{Y't\right\}\right] = E\left[\exp\left\{(AX+b)'t\right\}\right] \\
&= \exp\left\{b't\right\} E\left[\exp\left\{X'\left(A't\right)\right\}\right] = \exp\left\{b't\right\}\exp\left\{\boldsymbol{\mu}'\left(A't\right) + \frac{1}{2}\left(A't\right)'\boldsymbol{\Sigma}\left(A't\right)\right\} \\
&= \exp\left\{(A\boldsymbol{\mu}+b)'t + \frac{1}{2}t'\left(A\boldsymbol{\Sigma}A'\right)t\right\}
\end{aligned}$$

นั่นคือ $Y \sim N\left(A\boldsymbol{\mu} + b, A\boldsymbol{\Sigma}A'\right)$ □

ประโยชน์ในการแปลงตัวแปรสุ่มในรูปแบบฟังก์ชันเชิงเส้น (linear transformation)

- ทฤษฎีบทด้านบนยังช่วยให้สามารถหาการแจกแจงของตัวแปรสุ่มเพียงบางส่วนในกรณีที่การแจกแจงร่วมเป็นแบบปกติ โดยเริ่มจากการที่สามารถแบ่งเวคเตอร์ X ออกได้เป็นสองส่วน คือ

$$X = \begin{bmatrix} X_m \\ X_k \end{bmatrix} \tag{15}$$

  โดยที่ X_m คือเวคเตอร์ของตัวแปรสุ่มที่สนใจ โดยในที่นี้กำหนดให้มีขนาด m ในขณะที่ตัวแปรสุ่ม X_k มีขนาด $k = n - m$

- กำหนดให้ $b = \mathbf{0}$ และ

$$A = \begin{bmatrix} I_{mm} & \mathbf{0}_{mk} \\ \mathbf{0}_{km} & \mathbf{0}_{kk} \end{bmatrix} \tag{16}$$

  จะทำให้ฟังก์ชันที่เกิดจากการแปลงค่าโดยใช้ A และ b เท่ากับ

$$Y = \begin{bmatrix} I_{mm} & \mathbf{0}_{mk} \end{bmatrix} \begin{bmatrix} X_m \\ X_k \end{bmatrix} = X_m$$

สรุป: การแปลงตัวแปรสุ่มในรูปแบบฟังก์ชันเชิงเส้น (linear transformation)

- เราสามารถแบ่ง ค่าคาดหมาย (mean) และเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix) ของ X ได้เป็น

$$\boldsymbol{\mu} = \begin{bmatrix} \boldsymbol{\mu}_m \\ \boldsymbol{\mu}_k \end{bmatrix} \tag{17}$$

และ

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{mm} & \boldsymbol{\Sigma}_{mk} \\ \boldsymbol{\Sigma}_{km} & \boldsymbol{\Sigma}_{kk} \end{bmatrix} \tag{18}$$

โดยที่ $\boldsymbol{\Sigma}_{mk}$ คือเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix) ระหว่าง X_m และ X_k

- ค่าคาดหมาย (mean) และเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix) ของ Y สามารถแบ่งส่วนได้เป็น

$$\begin{bmatrix} I_{mm} & \mathbf{0}_{mk} \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu}_m \\ \boldsymbol{\mu}_k \end{bmatrix} = \boldsymbol{\mu}_m \tag{19}$$

และ

$$\begin{bmatrix} I_{mm} & \mathbf{0}_{mk} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{mm} & \boldsymbol{\Sigma}_{mk} \\ \boldsymbol{\Sigma}_{km} & \boldsymbol{\Sigma}_{kk} \end{bmatrix} \begin{bmatrix} I_{mm} \\ \mathbf{0}_{mk} \end{bmatrix} = \boldsymbol{\Sigma}_{mm} \tag{20}$$

- สรุปได้ว่า $Y = X_m \sim N\left(\boldsymbol{\mu}_m, \boldsymbol{\Sigma}_{mm}\right)$ กล่าวคือ การแจกแจงร่วมเป็นแบบปกติ แล้วการแจกแจงของตัวแปรสุ่มบางส่วนก็จะมีการแจกแจงปกติเช่นกัน โดยที่ค่าคาดหมาย (mean) และค่าความแปรปรวน (variance) ของชุดตัวแปรบางส่วนสามารถหาได้จากค่าสถิติของตัวแปรสุ่มทั้งหมดได้ไม่ยากนัก

ตัวอย่าง: การหาฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ X_1 และ X_2

Example

พิจารณาตัวแปรสุ่มสองตัว X_1 และ X_2 ที่มีการแจกแจงปกติ (normally distributed) ด้วยค่าคาดหมาย (μ_1, μ_2) และเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix)

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

จะเห็นได้ว่า det $\boldsymbol{\Sigma} = \sigma_1^2\sigma_2^2 - \sigma_{12}^2 = \sigma_1^2\sigma_2^2\left(1-\rho^2\right)$ โดยที่ ρ แทนค่าสหสัมพันธ์ระหว่าง X_1 และ X_2 ดังนั้น ส่วนกลับ (inverse) ของ $\boldsymbol{\Sigma}$ เท่ากับ

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)} \begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}$$

- สามารถคำนวณได้ว่า

$$\begin{aligned}(\mathbf{x}-\boldsymbol{\mu})'\,\boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu}) &= [x_1-\mu_1, x_2-\mu_2]\frac{1}{\sigma_1^2\sigma_2^2\left(1-\rho^2\right)}\begin{bmatrix} \sigma_2^2 & -\sigma_{12} \\ -\sigma_{12} & \sigma_1^2 \end{bmatrix}\begin{bmatrix} x_1-\mu_1 \\ x_2-\mu_2 \end{bmatrix} \\ &= \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\end{aligned}$$

วิธีทำ: การหาฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ X_1 และ X_2

Example

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ X_1 และ X_2 คือ

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} e^{-\frac{Q}{2}} \tag{21}$$

โดยที่

$$Q = \frac{1}{1-\rho^2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\left(\frac{x_1-\mu_1}{\sigma_1}\right)\left(\frac{x_2-\mu_2}{\sigma_2}\right) + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right] \tag{22}$$

- ถ้า $\rho = 0$ แล้ว ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ X_1 และ X_2 คือ

$$\begin{aligned} f(x_1, x_2) &= \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\} \\ &= \left[\frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2}\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2}\right]\left[\frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2}\left(\frac{x_2-\mu_2}{\sigma_2}\right)^2}\right] = f_1(x_1) f_2(x_2) \end{aligned}$$

- หมายความว่า X_1 และ X_2 เป็นอิสระต่อกัน กล่าวคือ ในกรณีของการแจกแจงปกติ (normal distribution) การไม่มีสหสัมพันธ์ (uncorrelated) หมายถึงการเป็นอิสระต่อกัน (independent) ด้วย

ทฤษฎี: ความเป็นอิสระต่อกัน (independent) ของตัวแปรสุ่ม

Theorem

กำหนดให้ $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ และสามารถแยกส่วนได้เป็นสองส่วนคือ $X = (X_m, X_k)$ ซึ่งสอดคล้องกับสมการที่ 15, 17, และ 18 ดังนั้น X_m และ X_k เป็นอิสระต่อกัน (independent) ก็ต่อเมื่อ (if and only if) $\boldsymbol{\Sigma}_{mk} = \mathbf{0}$

Proof.

พิจารณาฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X ในรูปของ t_m และ t_k

$$\begin{aligned}\psi(t_m, t_k) &= \exp\left\{\boldsymbol{\mu_m}' t_m + \boldsymbol{\mu_k}' t_k + \frac{1}{2}\left[t_m' \boldsymbol{\Sigma}_{mm} t_m + t_k' \boldsymbol{\Sigma}_{mm} t_k + t_m' \boldsymbol{\Sigma}_{mk} t_k + t_k' \boldsymbol{\Sigma}_{km} t_m\right]\right\} \\ &= \exp\left\{\boldsymbol{\mu_m}' t_m + \boldsymbol{\mu_k}' t_k + \frac{1}{2}\left[t_m' \boldsymbol{\Sigma}_{mm} t_m + t_k' \boldsymbol{\Sigma}_{mm} t_k\right]\right\} \exp\left\{\frac{1}{2}\left[t_m' \boldsymbol{\Sigma}_{mk} t_k + t_k' \boldsymbol{\Sigma}_{km} t_m\right]\right\} \end{aligned} \tag{23}$$

$$\begin{aligned}\psi(t_m, t_k) &= \exp\left\{\boldsymbol{\mu_m}' t_m + \frac{1}{2}\left[t_m' \boldsymbol{\Sigma}_{mm} t_m\right]\right\} \exp\left\{\boldsymbol{\mu_k}' t_k + \frac{1}{2}\left[t_k' \boldsymbol{\Sigma}_{mm} t_k\right]\right\} \\ &= \exp\left\{\boldsymbol{\mu_m}' t_m + \boldsymbol{\mu_k}' t_k + \frac{1}{2}\left[t_m' \boldsymbol{\Sigma}_{mm} t_m + t_k' \boldsymbol{\Sigma}_{mm} t_k\right]\right\} \end{aligned} \tag{24}$$

$$\exp\left\{\frac{1}{2}\left[t_m' \boldsymbol{\Sigma}_{mk} t_k + t_k' \boldsymbol{\Sigma}_{km} t_m\right]\right\} = 1 \Rightarrow t_m' \boldsymbol{\Sigma}_{mk} t_k + t_k' \boldsymbol{\Sigma}_{km} t_m = 0$$

ซึ่งจะเป็นจริงก็ต่อเมื่อ $\boldsymbol{\Sigma}_{mk} = \boldsymbol{\Sigma}_{km}' = \mathbf{0}$ □

ทฤษฎี: การแจกแจงแบบมีเงื่อนไข (conditional distribution) เป็นการแจกแจงแบบปกติ (normal)

Theorem

กำหนดให้ $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *และสามารถแยกส่วนได้เป็นสองส่วนคือ* $X = (X_m, X_k)$ *ซึ่งสอดคล้องกับสมการที่ 15, 17, และ 18 และสมมุติว่า* $\boldsymbol{\Sigma}$ *เป็นบวกแน่นอน (positive definite) ดังนั้น การแจกแจงแบบมีเงื่อนไข (conditional distribution) ของ* X_m *เมื่อทราบ* X_k *เป็นแบบปกติ (normal) นั่นคือ*

$$X_m \middle| X_k \sim N\left(\boldsymbol{\mu}_m + \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\left(X_k - \boldsymbol{\mu}_k\right), \boldsymbol{\Sigma}_{mm} - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\Sigma}_{km}\right) \tag{25}$$

การพิสูจน์: การแจกแจงแบบมีเงื่อนไข (conditional distribution) เป็นการแจกแจงแบบปกติ (normal)

Proof.

- กำหนดให้ $W = X_m - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}X_k$ และ

$$\begin{bmatrix} W \\ X_2 \end{bmatrix} = \begin{bmatrix} I_m & -\boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1} \\ \mathbf{0} & I_k \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

- คำนวณหาเมทริกซ์ของค่าความแปรปรวน (variance-covariance matrix) ได้เป็น

$$\begin{bmatrix} I_m & -\boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1} \\ \mathbf{0} & I_k \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{mm} & \boldsymbol{\Sigma}_{mk} \\ \boldsymbol{\Sigma}_{km} & \boldsymbol{\Sigma}_{kk} \end{bmatrix} \begin{bmatrix} I_m & \mathbf{0} \\ -\boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1} & I_k \end{bmatrix} = \begin{bmatrix} \boldsymbol{\Sigma}_{mm} - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\Sigma}_{km} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{kk} \end{bmatrix}$$

- เห็นได้ว่า W และ X_k เป็นอิสระต่อกัน (independent) ดังนั้น **การแจกแจงแบบมีเงื่อนไข $W|X_k$ จะเหมือนกับการแจกแจงของ W** ซึ่งเป็นการแจกแจงปกติ (normal distribution) ด้วยค่าคาดหมาย (mean) เท่ากับ $\boldsymbol{\mu}_m - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\mu}_k$ และสามารถสรุปได้ว่า

$$W|X_k \sim N\left(\boldsymbol{\mu}_m - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\mu}_k, \boldsymbol{\Sigma}_{mm} - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\Sigma}_{km}\right)$$

- ในขณะเดียวกัน เนื่องจากสิ่งที่เราสนใจคือการแจกแจงแบบมีเงื่อนไขเมื่อทราบ X_k ดังนั้น การบวกเพิ่มด้วยฟังก์ชันของ X_k เข้าไปย่อม**มีผลต่อเช่นเดียวกับการบวกเพิ่มด้วยค่าคงที่** ซึ่งมีผลต่อค่าคาดหมาย แต่**ไม่มีผลต่อค่าความแปรปรวน**
- เราสามารถคำนวณหา X_m ได้จาก $W + \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}X_k = X_m$ ดังนั้น การแจกแจงแบบมีเงื่อนไข (conditional distribution) ของ X_m เมื่อทราบ X_k เท่ากับ

$$X_m|X_k \sim N\left(\boldsymbol{\mu}_m - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\mu}_k + \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}X_k, \boldsymbol{\Sigma}_{mm} - \boldsymbol{\Sigma}_{mk}\boldsymbol{\Sigma}_{kk}^{-1}\boldsymbol{\Sigma}_{km}\right)$$

ตัวอย่าง: การคำนวณหาค่าคาดหมายแบบมีเงื่อนไข (conditional mean)

Example

เช่นเดียวกับตัวอย่างที่แล้ว สมมุติให้ตัวแปรสุ่ม X_1 และ X_2 มีการแจกแจงปกติ (normally distributed) ด้วยค่าคาดหมาย (μ_1, μ_2) และเมทริกซ์ความแปรปรวนร่วม (variance-covariance matrix)

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

- ใช้สมการที่ 25 เพื่อคำนวณหาค่าคาดหมายแบบมีเงื่อนไข (conditional mean) ของ X_1 เมื่อทราบ $X_2 = x_2$ ได้เป็น

$$\begin{aligned} E\left[X_1 | X_2 = x_2\right] &= \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_{12}\boldsymbol{\Sigma}_{22}^{-1}\left(X_2 - \boldsymbol{\mu}_2\right) = \mu_1 + \left(\rho\sigma_1\sigma_2\right)\left(\sigma_2^2\right)^{-1}\left(x_2 - \mu_2\right) \\ &= \mu_1 + \rho\frac{\sigma_1}{\sigma_2}\left(x_2 - \mu_2\right) \end{aligned}$$

- ฟังก์ชันเชิงเส้น (linear function) ของ x_2 สังเกตว่าเราไม่ได้กำหนดหรือบังคับให้ค่าคาดหมายแบบมีเงื่อนไข (conditional mean) นี้มีความสัมพันธ์เชิงเส้นแต่อย่างใด แต่คุณสมบัติของการแจกแจงปกติ นอกจากนี้ ยังสามารถเขียนในรูปของตัวแปรสุ่มได้เป็น

$$E\left[X_1 | X_2\right] = \mu_1 + \rho\frac{\sigma_1}{\sigma_2}\left(X_2 - \mu_2\right) \tag{26}$$

การแจกแจงไคกำลังสอง (Chi-square distribution)

- การแจกแจงไคกำลังสอง (chi-square distribution) เป็นการแจกแจงอีกแบบหนึ่งที่ได้รับความนิยม และเป็นหนึ่งในการแจกแจงแบบแกมมา (gamma distribution)
- คุณสมบัติหนึ่งที่มีความสำคัญในทางสถิติและจะมีประโยชน์สำหรับการหาการแจกแจงของค่าสถิติที่ใช้ในการทดสอบสมมุติฐานในอนาคต
- เราสามารถสร้างตัวแปรสุ่มที่มีการแจกแจงไคกำลังสอง (chi-square distribution) ได้ด้วยการนำเอาตัว **แปรสุ่มที่มีการแจกแจงปกติมาตรฐาน (standard normal) มายกกำลังสอง** กล่าวคือ ตัวแปรไคกำลังสอง (chi-square) คือกำลังสองของตัวแปรปกติ (normal)
- กำหนดให้ X เป็นตัวแปรสุ่มที่มีการแจกแจงไคกำลังสอง (chi-square distribution) ด้วยระดับความอิสระ(degree of freedom) เท่ากับ r นั่นคือ $X \sim \chi^2(r)$ ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ X คือ

$$f(x) = \begin{cases} \frac{1}{\Gamma\left(\frac{r}{2}\right)2^{\frac{r}{2}}} x^{\frac{r}{2}-1} e^{-\frac{x}{2}}, & \text{สำหรับ } 0 < x < \infty, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{27}$$

โดยที่ $\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} dx$ แทนฟังก์ชันแกมมา (gamma function)

ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) และฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของการแจกแจงไคกำลังสอง (Chi-square distribution)

- ค่าแกมมาฟังก์ชันอันหนึ่งที่มีความสัมพันธ์กับการแจกแจงปกติมาตราคือค่าที่มาจากกรณีที่ระดับความอิสระ (degree of freedom) $r = 1$ ซึ่งมีค่าเท่ากับ $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$ ดังนั้น ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ $\chi^2(1)$ คือ

$$f(x) = \begin{cases} \frac{1}{\sqrt{2\pi x}} e^{-\frac{x}{2}}, & \text{สำหรับ } 0 < x < \infty, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{28}$$

- ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X เท่ากับ

$$\psi(t) = (1 - 2t)^{-\frac{r}{2}}, \text{ สำหรับ } t < \frac{1}{2} \tag{29}$$

- ค่าคาดหมาย (mean) ของ $\chi^2(r)$ เท่ากับ $\psi^1(0) = r$ และค่าความแปรปรวน (variance) ของ $\chi^2(r)$ เท่ากับ $\psi^2(0) - r^2 = r(r+2) - r^2 = 2r$

กำลังสองของการแจกแจงปกติเป็นการแจกแจงไคกำลังสอง (Chi-square distribution) ด้วยระดับความอิสระ (degree of freedom) $r = 1$

Theorem

สมมุติให้ $X \sim N(0, 1)$ *แล้ว ตัวแปรสุ่ม* $Y = X^2 \sim \chi^2(r)$

Proof.

- เริ่มจากฟังก์ชันความน่าจะเป็นสะสม (C.D.F.) ของ Y เท่ากับ

$$G(y) = Pr(Y \leq y) = Pr\left(X^2 \leq y\right) = Pr(-\sqrt{y} \leq X \leq \sqrt{y})\text{ , สำหรับ } y > 0$$

- เนื่องจาก $X \sim N(0, 1)$ ดังนั้น สำหรับ $y > 0$

$$G(y) = Pr(-\sqrt{y} \leq X \leq \sqrt{y}) = \int -\sqrt{y}^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz = 2 \int 0^{\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

- สมการสุดท้ายเป็นผลมาจากความสมมาตรของการแจกแจงปกติ หลังจากนั้นคือการเปลี่ยนตัวแปรการอินทิเกรต โดยกำหนดให้ $w = z^2$ ซึ่งส่งผลให้ $dz = \frac{dw}{2\sqrt{w}}$ ดังนั้น

$$G(y) = \int 0^y \frac{1}{\sqrt{2\pi w}} e^{-\frac{w}{2}} dw$$

ซึ่งมีค่าเท่ากับฟังก์ชันความน่าจะเป็นสะสม (C.D.F.) ของ $\chi^2(1)$ ตามที่ต้องการ

ผลบวกของตัวแปรสุ่มที่เป็นอิสระต่อกัน (independent) เป็นการแจกแจงไคกำลังสอง (Chi-square distribution)

- คุณสมบัติอีกอันหนึ่งที่การแจกแจงไคกำลังสอง (chi-square) ที่คล้ายกับคุณสมบัติของการแจกแจงปกติ (normal) คือ คุณสมบัติการรักษารูปแบบการกระจาย กล่าวคือ ผลบวกของตัวแปรสุ่มที่เป็นอิสระต่อกัน (independent) และแต่ละตัวมีการแจกแจงไคกำลังสอง (chi-square) จะมีการแจกแจงไคกำลังสอง (chi-square)

Theorem

สมมุติให้ $X_i \sim \chi^2(r_i)$ *สำหรับ* $i = 1, \ldots, n$ *และ* $X_1, \ldots, X_n$ *เป็นอิสระต่อกัน (independent) แล้ว ตัวแปรสุ่ม* $Y = \sum_{i=1}^{n} X_i \sim \chi^2(R)$ *โดยที่* $R = \sum_{i=1}^{n} r_i$

Proof.

เนื่องจาก $X_1, \ldots, X_n$ เป็นอิสระต่อกัน (independent) ดังนั้น ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Y เท่ากับ

$$\psi_y(t) = \prod_{i=1}^{n} (1-2t)^{-\frac{r_i}{2}} = (1-2t)^{-\frac{\sum_{i=1}^{n} r_i}{2}}$$

ซึ่งมีค่าเท่ากับฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของการแจกแจงไคกำลังสอง (chi-square) ที่มีระดับความอิสระ (degree of freedom) เท่ากับ $\sum_{i=1}^{n} r_i$ □

ทฤษฎีการแจกแจงไคกำลังสอง (Chi-square distribution)

Theorem

สมมุติให้ $X \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ *โดยที่* $\boldsymbol{\Sigma}$ *เป็นบวกแน่นอน (positive definite) แล้ว ตัวแปรสุ่ม* $Y = (X - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (X - \boldsymbol{\mu}) \sim \chi^2(n)$ *โดยที่* n *คือจำนวนตัวแปรสุ่มทั้งหมดใน* X

Proof.

กำหนดให้ $Z = \boldsymbol{\Sigma}^{-\frac{1}{2}} (X - \boldsymbol{\mu})$ และจากที่อภิปรายมาก่อนหน้านี้ ทำให้ทราบดีว่า $Z \sim N(\mathbf{0}, I_n)$ และกำหนดให้ $W = Z'Z = \sum_{i=1}^{n} Z_i^2$ เนื่องจาก $Z_i \sim N(0,1)$ ส่งผลให้ $Z_i^2 \sim \chi^2(1)$ ดังนั้น จึงสามารถประยุกต์ใช้ทฤษฎีบทที่แล้ว สรุปได้ว่า $\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$ นั่นคือ $W \sim \chi^2(n)$ ซึ่งคือผลลัพธ์ที่ต้องการ $\square$

ตัวอย่างการแจกแจงไคกำลังสอง (Chi-square distribution)

Example

พิจารณาแบบจำลองเชิงเส้น

$$Y = X\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

โดยที่ $\boldsymbol{\beta}$ คือเวคเตอร์ของพารามิเตอร์ที่ต้องการประมาณค่า (vector of estimated parameters) และสมมุติให้ค่าคลาดเคลื่อน (error terms) $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$ และสมมุติว่ามีข้อมูลแบบตัวอย่างสุ่มขนาด n ตัวอย่าง ดังนั้น ในการทดสอบสมมุติฐานที่เกี่ยวข้องกับแบบจำลองเชิงเส้นนี้ จำเป็นจะต้องหาการแจกแจงของ $\boldsymbol{\varepsilon}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon}$ ซึ่งเมื่อประยุกต์ใช้ทฤษฎีบทที่ผ่านมาจะสามารถสรุปได้ว่า $\boldsymbol{\varepsilon}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\varepsilon} \sim \chi^2(n)$

- การแจกแจงไคกำลังสอง (chi-square) ยังมีความสำคัญเพราะเป็นส่วนสำคัญในการสร้างการแจกแจงที (*t* distribution) และการแจกแจงเอฟ (*F* distribution)

การแจกแจงที (t Distribution)

- กำหนดให้ $X \sim N(0,1)$ และ $Y \sim \chi^2(r)$ และเป็นอิสระต่อกัน (independent) ดังนั้น

$$T = \frac{X}{\sqrt{\frac{Y}{r}}} \tag{30}$$

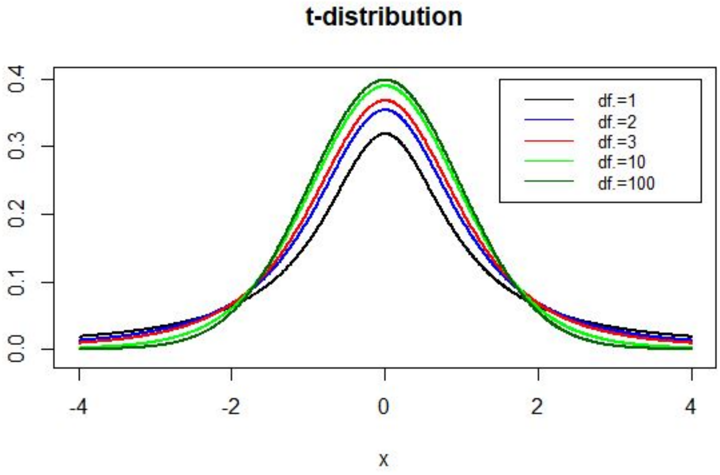
มีการแจกแจงที (t distribution) ด้วยระดับความอิสระ (degree of freedom) r ซึ่งมักแทนด้วย $t(r)$ นั่นคือ การแจกแจงที (t distribution) เกิดจากการผสมผสานกันของการแจกแจงปกติ (normal) และการแจกแจงไคกำลังสอง (chi-square)

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ T ด้วยระดับความอิสระ (degree of freedom) r คือ

$$f(t) = \frac{\Gamma\left(\frac{1+r}{2}\right)}{\Gamma\left(\frac{r}{2}\right)\sqrt{\pi r}}\left(1 + \frac{t^2}{r}\right)^{-\frac{1+r}{2}} \tag{31}$$

ในทางปฏิบัติค่าสถิติจากการแจกแจงทีมีค่าใกล้เคียงค่าสถิติจากการแจกแจงปกติอย่างมาก โดยเฉพาะอย่างยิ่งเมื่อกลุ่มตัวอย่างหรือระดับความอิสระ (degree of freedom) มีค่ามากพอ นั่นคือ $\lim_{r\to\infty} t(r) \sim N(0,1)$

กราฟแสดงการแจกแจงที่ (t Distribution)



การแจกแจงเอฟ (*F* Distribution)

- กำหนดให้ $X \sim \chi^2(r_1)$ และ $Y \sim \chi^2(r_2)$ และเป็นอิสระต่อกัน (independent) ดังนั้น

$$F = \frac{\frac{X}{r_1}}{\frac{Y}{r_2}} \tag{32}$$

มีการแจกแจงเอฟ (*F* distribution) ด้วยระดับความอิสระ (degree of freedom) r_1 และ r_2 ซึ่งมักแทนด้วย $F(r_1, r_2)$ นั่นคือ การแจกแจงที (*F* distribution) เป็น **ผลลัพธ์จากการหารกันของการแจกแจงไคกำลังสอง (chi-square) สองอัน**

- ฟังก์ชันความหนาแน่นของความน่าจะเป็นร่วม (p.d.f.) ของ *T* ด้วยระดับความอิสระ (degree of freedom) *r* คือ

$$f(x) = \begin{cases} \frac{\Gamma\left(\frac{r_1+r_2}{2}\right)\left(\frac{r_1}{r_2}\right)^{\frac{r_1}{2}}}{\Gamma\left(\frac{r_1}{2}\right)\Gamma\left(\frac{r_2}{2}\right)} x^{\frac{r_1}{2}-1}\left(1+\frac{r_1}{r_2}x\right)^{-\frac{r_1+r_2}{2}}, & \text{สำหรับ } 0 < x < \infty, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{33}$$

- นอกจากนี้ ยังสามารถแสดงได้ว่า ลิมิตของการแจกแจงเอฟ คือการแจกแจงไคกำลังสอง นั่นคือ

$$\lim_{r_2 \to \infty} F(r_1, r_2) \sim \chi^2(r_1) \tag{34}$$

$$\lim_{r_1 \to \infty} \lim_{r_2 \to \infty} F(r_1, r_2) \sim N(0, 1) \tag{35}$$

การแจกแจงของตัวแปรสุ่มไม่ต่อเนื่อง

- การแจกแจงแบบ Bernoulli (Bernoulli Distribution)
- การแจกแจงแบบ Binomial (Binomial Distribution)
- การแจกแจงแบบ Poisson (Poisson Distribution)

\- Binomial Distribution
\- แบบพอยซองน์ (Poisson Distribution)

การแจกแจงแบบ Bernoulli (Bernoulli Distribution)

- กำหนดให้ X เป็นตัวแปรสุ่มที่ไม่ต่อเนื่อง ดังนั้น ฟังก์ชันความความน่าจะเป็น (p.f.) ของ X ที่พารามิเตอร์ p คือ

$$f(x) = \begin{cases} p^x(1-p)^{1-x}, & \text{สำหรับ } x \in \{0,1\}, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{36}$$

- ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X สำหรับ $-\infty < t < \infty$ คือ

$$\psi(t) = E(e^{tX}) = pe^t + (1-p) \tag{37}$$

- ค่าคาดหมาย (mean) เท่ากับ $\psi^1(0) = p$
- ค่าความแปรปรวน (variance) เท่ากับ $\psi^2(0) - p^2 = p - p^2 = p(1-p)$

การแจกแจงแบบ Binomial (Binomial Distribution)

- กำหนดให้ X เป็นตัวแปรสุ่มที่ไม่ต่อเนื่อง ดังนั้น ฟังก์ชันความความน่าจะเป็น (p.f.) ของ X ที่พารามิเตอร์ n และ p คือ

$$f(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x}, & \text{สำหรับ } x \in \{0, 1, 2, \dots\}, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{38}$$

- ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X สำหรับ $-\infty < t < \infty$ คือ

$$\psi(t) = E(e^{tX}) = (pe^t + (1-p))^n \tag{39}$$

- ค่าคาดหมาย (mean) เท่ากับ $\psi^1(0) = np$
- ค่าความแปรปรวน (variance) เท่ากับ $\psi^2(0) - (np)^2 = (n(n-1)p^2 + np) - (np)^2 = np(1-p)$

ความสัมพันธ์ระหว่าง Bernoulli Distribution และ Binomial Distribution

Theorem

สมมุติให้ X_i *มีการแจกแจงแบบ Bernoulli Distribution ที่พารามิเตอร์* p *สำหรับ* $i = 1, \ldots, n$ *และ* $X_1, \ldots, X_n$ *เป็นอิสระต่อกัน (independent) แล้ว ตัวแปรสุ่ม* $Y = \sum_{i=1}^{n} X_i$ *จะมีการแจกแจงแบบ Binomial Distribution ที่พารามิเตอร์* n *และ* p

Proof.

เนื่องจาก $X_1, \ldots, X_n$ เป็นอิสระต่อกัน (independent) ดังนั้น ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Y เท่ากับ

$$\psi_Y(t) = E(e^{tY}) = E(e^{t\sum_{i=1}^{n} X_i}) = \prod_{i=1}^{n} E(e^{tX_i}) = (pe^t + (1-p))^n$$

ซึ่งมีค่าเท่ากับฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของการแจกแจงแบบ Binomial Distribution ที่พารามิเตอร์ n และ p $\square$

การแจกแจงแบบ Poisson (Poisson Distribution)

- กำหนดให้ X เป็นตัวแปรสุ่มที่ไม่ต่อเนื่อง ดังนั้น ฟังก์ชันความความน่าจะเป็น (p.f.) ของ X ที่พารามิเตอร์ λ คือ

$$f(x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}, & \text{สำหรับ } x \in \{0, 1, 2, ...\}, \\ 0, & \text{สำหรับกรณีอื่น} \end{cases} \tag{40}$$

- ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ X สำหรับ $-\infty < t < \infty$ คือ

$$\psi(t) = E(e^{tX}) = \sum_{x=0}^{\infty} \frac{e^{tx}e^{-\lambda}\lambda^x}{x!} = e^{\lambda(e^t-1)} \tag{41}$$

- ค่าคาดหมาย (mean) เท่ากับ $\psi^1(0) = \lambda$
- ค่าความแปรปรวน (variance) เท่ากับ $\psi^2(0) - \lambda^2 = (\lambda^2 + \lambda) - \lambda^2 = \lambda$

ผลรวมของตัวแปรสุ่มที่มีการแจกแจงแบบ Poisson จะมีการแจกแจงแบบ Poisson

Theorem

สมมุติให้ X_i มีการแจกแจงแบบ Poisson (Poisson Distribution) ที่พารามิเตอร์ λ_i สำหรับ $i = 1, \ldots, n$ และ $X_1, \ldots, X_n$ เป็นอิสระต่อกัน (independent) แล้ว ตัวแปรสุ่ม $Y = \sum_{i=1}^{n} X_i$ จะมีการแจกแจงแบบ Poisson (Poisson Distribution) ที่พารามิเตอร์ λ คือ $\lambda_y = \sum_{i=1}^{n} \lambda_i$

Proof.

เนื่องจาก $X_1, \ldots, X_n$ เป็นอิสระต่อกัน (independent) ดังนั้น ฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของ Y เท่ากับ

$$\psi_y(t) = \prod_{i=1}^{n} e^{\lambda_i(e^t - 1)} = e^{(\sum_{i=1}^{n} \lambda_i)(e^t - 1)}$$

ซึ่งมีค่าเท่ากับฟังก์ชันก่อกำเนิดโมเมนต์ (m.g.f.) ของการแจกแจงแบบ Poisson (Poisson Distribution) ที่พารามิเตอร์ $\lambda_y = \sum_{i=1}^{n} \lambda_i$ □

ความสัมพันธ์ระหว่างการแจกแจงแบบ Poisson และ การแจกแจงแบบ Binomial

Theorem

ในกรณี พารามิเตอร์ n ของการแจกแจงแบบ Binomial มีค่ามาก ($n \longrightarrow \infty$) ฟังก์ชันการแจกแจงจะเข้าสู่ การแจกแจงแบบพัวซอง (Poisson Distribution) ที่พารามิเตอร์ $\lambda = np$

Proof.

ในกรณี พารามิเตอร์ n ของการแจกแจงแบบ Binomial มีค่ามาก ($n \longrightarrow \infty$) ฟังก์ชันการแจกแจงจะเข้าสู่ การแจกแจงแบบPoisson ที่พารามิเตอร์ λ

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x} = \frac{(n)(n-1)(n-3)\ldots(n-x+1)}{x!} p^x (1-p)^{n-x}$$

ให้ $\lambda = np$

$$f(x) = \frac{\lambda^x}{x!} \left(\frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \ldots \cdot \frac{n-x+1}{n} \cdot (1 - \frac{\lambda}{n})^n \right) (1 - \frac{\lambda}{n})^{-x}$$

ถ้า ($n \longrightarrow \infty$) แล้ว

$$\lim_{n \to \infty} \frac{n}{n} \cdot \frac{n-1}{n} \cdot \frac{n-2}{n} \cdot \ldots \cdot \frac{n-x+1}{n} \cdot (1 - \frac{\lambda}{n})^n = 1 \quad \text{และ} \quad \lim_{n \to \infty} (1 - \frac{\lambda}{n})^{-x} = e^{-\lambda}$$

$$f(x) = \frac{\lambda^x}{x!} e^{-\lambda} \tag{42}$$