บทที่ 2

ทฤษฎีและตัวสถิติที่เกี่ยวข้อง

ในการวิจัยครั้งนี้ ได้ทำการศึกษาเปรียบเทียบแผนภูมิควบคุมคุณภาพที่ใช้สำหรับตรวจวัดกระบวนการผลิตที่มีการเปลี่ยนแปลงในค่าเฉลี่ย ซึ่งทำการเปรียบเทียบแผนภูมิควบคุมทั้ง 3 แบบ คือ แผนภูมิควบคุมค่าเฉลี่ยชนิดตัวแปรเดียว แผนภูมิควบคุม Hotelling และแผนภูมิควบคุมเชิงพหุ Shewhart $\overline{X}$

ให้ $(X_{11},X_{21}),(X_{12},X_{22}),...(X_{1n},X_{2n})$ แทนค่าวัดลักษณะเฉพาะของผลิตผลที่ผลิตออกมา แบ่งเป็นกลุ่มย่อย ๆ กลุ่มละจำนวน n การแบ่งกลุ่มย่อย ควรให้มีสมาชิกในกลุ่มมีความเหมือนกันมากที่สุด เช่น ผลิตจากคาบเวลาเดียวกันหรือผลิตจากเครื่องจักรตัวเดียวกัน

$$\overline{X}_{it} = \frac{X_{i1}+X_{i2}+...+X_{in}}{n} \quad , t=1,2,3,... \ ; i=1,2$$

โดยที่ข้อมูล $(\overline{X}_{1t},\overline{X}_{2t})$ จะนำไปใช้ในแต่ละแผนภูมิควบคุมต่อไปในการตรวจสอบวิธีว่าอยู่ในการควบคุมหรือไม่ โดยที่ค่าผลิตภัณฑ์มีการแจกแจงแบบทวิคูณปกติ ที่มี ค่าเฉลี่ยของตัวแปรสุ่ม X_1 และ X_2 เป็น μ_1 และ μ_2 ค่าความแปรปรวนตัวแปรสุ่ม X_1 และ X_2 เป็น σ_1^2 และ σ_2^2 ตามลำดับและค่าสัมประสิทธิ์สหสัมพันธ์เป็น ρ ซึ่งในการวิจัยครั้งนี้จะสมมติว่าทราบค่า $\mu_1,\mu_2,\sigma_1^2,\sigma_2^2$ และ ρ กล่าวคือ สามารถกำหนดค่าให้เป็นไปตามสถานการณ์ที่ต้องการศึกษา

2.1 ตัวแบบที่ใช้ในการสร้างตัวแปรสุ่ม X_1 และ X_2 สำหรับกรณีวิจัยโดยหลักการจำลองข้อมูล

สำหรับกรณีใช้หลักการจำลองข้อมูล ข้อมูลที่นำมาศึกษาได้จากตัวแบบอนุกรมเวลา โดยมีตัวแบบดังนี้

$$X_{it} = \mu_{0i} + (\delta_i\sigma_i) + \varepsilon_{it} \quad ; i=1,2 \ , t=1,2,3,... \qquad (2.1)$$

โดยที่ X_{it} คือ ค่าวัดตัวแปรสุ่ม X_i ; $i=1,2$ ณ เวลาที่ t

μ_{0i} คือ ค่าเฉลี่ยประชากรของตัวแปรสุ่ม X_i ; $i=1,2$ เมื่อกระบวนการผลิตอยู่ภายในสถานการณ์ปกติ

δ_i คือ ค่าคงที่ซึ่งแสดงถึงขนาดการเปลี่ยนแปลงไปของค่าเฉลี่ยต่อส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X_i ; $i=1,2$

σ_i คือ ส่วนเบี่ยงเบนมาตรฐานของตัวแปรสุ่ม X_i ; $i = 1,2$

ε_{it} คือ ค่าความคลาดเคลื่อนที่สุ่มของตัวแปรสุ่ม X_i ; $i = 1,2$ ณ เวลา t ซึ่ง ε_{1t} และ ε_{2t} มีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเท่ากับ 0 และค่าความแปรปรวนเท่ากับ 1 โดยที่ ε_{1t} และ ε_{2t} มีความสัมพันธ์กันกล่าวคือ $-1 \le \rho(\varepsilon_{1t},\varepsilon_{2t}) \le 1$

n คือ ขนาดตัวอย่างที่ใช้ในการศึกษาเพื่อหาค่าเฉลี่ยตัวอย่าง $(\bar{X}_i\ ; i = 1,2)$

สามารถเขียนให้อยู่ในรูปเวกเตอร์ได้ดังนี้

$$\underset{\sim}{X}_t = \underset{\sim}{\mu}_0 + \underset{\sim}{\gamma} + \underset{\sim}{\varepsilon}_t \quad ; \quad t = 1,2,3,... \tag{2.2}$$

เมื่อ เวกเตอร์สุ่ม $\underset{\sim}{X}_t = [X_{1t} \quad X_{2t}]^T$

เวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\mu}_0 = [\mu_{01} \quad \mu_{02}]^T$

เวกเตอร์ระดับการเปลี่ยนแปลงของค่าเฉลี่ย $\underset{\sim}{\gamma} = [\delta_1\sigma_1 \quad \delta_2\sigma_2]^T$

เวกเตอร์ความคลาดเคลื่อนในการสุ่ม $\underset{\sim}{\varepsilon}_t = [\varepsilon_{1t} \quad \varepsilon_{2t}]^T$

ดังนั้นในการวิจัยครั้งนี้ $\underset{\sim}{X}_t \sim N_2\left(\underset{\sim}{\mu},\Sigma\right)$ มีการแจกแจงแบบทวิคูณปกติ ซึ่งเวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\mu} = \underset{\sim}{\mu}_0 + \underset{\sim}{\gamma}$ และเมทริกซ์ความแปรปรวนร่วม $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho \\ \rho & \sigma_2^2 \end{bmatrix}$ เมื่อ σ_1^2 และ σ_2^2 มีค่าเท่ากับ 1

โดยที่ $\bar{X}_1 \sim N\left(\mu_{01},\dfrac{\sigma_1^2}{n}\right)$ และ $\bar{X}_2 \sim N\left(\mu_{02},\dfrac{\sigma_2^2}{n}\right)$

ถ้ากระบวนการผลิตมีค่าเฉลี่ยมีการเปลี่ยนแปลงไปจากเดิมในการวิจัยครั้งนี้จะพิจารณา 2 กรณี คือ

กรณีที่ 1 $\delta_1 = 0$ แต่ $\delta_2 \neq 0$ และกรณีที่ 2 $\delta_1 \neq 0$ และ $\delta_2 \neq 0$

ดังนั้น เวกเตอร์ค่าเฉลี่ยใหม่ $\underset{\sim}{\mu}_1 = \begin{bmatrix} \mu_{11} \\ \mu_{12} \end{bmatrix}$

โดยที่ $\mu_{11} = \mu_{01} + \delta_1\sigma_1$ และ $\mu_{12} = \mu_{02} + \delta_2\sigma_2$

เมื่อ $X_1 \sim N\left(\mu_{11},\sigma_1^2\right)$ และ $X_2 \sim N\left(\mu_{12},\sigma_2^2\right)$

ทำการสุ่มชุดตัวอย่างขนาด n โดยทำการวัดค่าในแต่ละตัวแปร X_i ; $i = 1,2$

$$E(\bar{X}_i) = E\left(\sum_{j=1}^{n} X_{ij} \Big/ n\right) = \frac{1}{n} E(X_{i1} + X_{i2} + \ldots + X_{in}) \quad ; \quad i = 1,2$$

$$= (\mu_{0i} + \delta_i \sigma_i)$$

$$Var(\bar{X}_i) = Var\left(\sum_{j=1}^{n} X_{ij} \Big/ n\right) = \frac{1}{n^2} Var(X_{i1} + X_{i2} + \ldots + X_{in}) \quad ; \quad i = 1,2$$

$$= \frac{\sigma_i^2}{n}$$

เนื่องจาก $X_{i1}, X_{i2}, X_{i3}, \ldots, X_{in}$; $i = 1,2$ เป็นอิสระต่อกัน

ดังนั้น $\bar{X}_1 \sim N\left(\mu_{01}, \frac{\sigma_1^2}{n}\right)$ และ $\bar{X}_2 \sim N\left(\mu_{02}, \frac{\sigma_2^2}{n}\right)$

เมื่อ $\mu_{11} = \mu_{01} + \delta_1 \sigma_1$ และ $\mu_{12} = \mu_{02} + \delta_2 \sigma_2$ เขียนในรูป เวกเตอร์ค่าเฉลี่ยตัวอย่าง $\begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} = \bar{\underset{\sim}{X}} \sim N_2\left(\underset{\sim}{\mu}_1, \Sigma\right)$ โดยที่เวกเตอร์ค่าเฉลี่ยใหม่ $\underset{\sim}{\mu}_1 = \begin{bmatrix} \mu_{11} \\ \mu_{12} \end{bmatrix}$ และเมทริกซ์ความแปรปรวนร่วม $\Sigma = \frac{1}{n}\begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$

2.2 การแจกแจงแบบทวิคูณปกติ (Bivariate Normal Distribution)

ในกรณีที่พิจารณาเพียงตัวแปรสุ่ม 1 ตัว คือ ตัวแปรสุ่ม X มีการแจกแจงแบบปกติ ซึ่งมีค่าเฉลี่ยเท่ากับ μ และค่าความแปรปรวนเท่ากับ σ^2 จะมีฟังก์ชันความน่าจะเป็นดังนี้

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right] \quad ; \quad -\infty < x < \infty$$

ในกรณีที่พิจารณาตัวแปรสุ่ม 2 ตัวคือ X_1 และ X_2

ให้ เวกเตอร์สุ่ม $\underset{\sim}{X} = [X_1 \quad X_2]^T$

โดยที่ เวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\mu} = [\mu_1 \quad \mu_2]^T$

และ เมทริกซ์ของความแปรปรวนร่วม $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$

ดังนั้น เวกเตอร์สุ่ม $\underset{\sim}{X}$ จะมีฟังก์ชันความน่าจะเป็นดังนี้

$$f\left(\underset{\sim}{x}\right) = \frac{1}{2\pi}\det\left(\Sigma^{-\frac{1}{2}}\right)\exp\left(-\frac{\left(\underset{\sim}{x}-\underset{\sim}{\mu}\right)^T \Sigma^{-1}\left(\underset{\sim}{x}-\underset{\sim}{\mu}\right)}{2}\right)$$

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

$; \quad -\infty < x_i < \infty \ , i = 1,2$

ซึ่งจะแทนการแจกแจงแบบทวิคูณปกติ ด้วยสัญลักษณ์ $N_2\left(\underset{\sim}{\mu},\Sigma\right)$ เมื่อ $\underset{\sim}{\mu}$ เป็นเวกเตอร์ค่าเฉลี่ยและ Σ เป็นเมทริกซ์ของความแปรปรวนร่วมของประชากรและ ρ เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่าง ตัวแปรสุ่ม X_1 และ X_2

กรณีที่ ตัวแปรสุ่ม X_1 และ X_2 เป็นอิสระต่อกัน ซึ่งค่าสัมประสิทธิ์สหสัมพันธ์ (ρ) เท่ากับ 0 ดังนั้น $\Sigma = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$

เวกเตอร์สุ่ม $\underset{\sim}{X} \sim N_2\left(\underset{\sim}{\mu},\Sigma\right)$ จะมีฟังก์ชันความน่าจะเป็นดังนี้

$$f\left(\underset{\sim}{x}\right) = \frac{1}{2\pi\sigma_1\sigma_2}\exp\left\{-\frac{1}{2}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\}; \quad -\infty < x_i < \infty \ , i = 1,2$$

2.3 การแจกแจงไคสแควร์ที่ไม่เบี่ยงเบนไปจากศูนย์กลาง(Non-central chi-square Distribution)

ถ้า เวกเตอร์สุ่ม $\underset{\sim}{X} = [X_1 \quad X_2]^T$ มีการแจกแจงแบบทวิคูณปกติ หรือเขียนแทนด้วย $\underset{\sim}{X} \sim N\left(\underset{\sim}{\mu},\Sigma\right)$

เมื่อ $\underset{\sim}{\mu} = [\mu_1 \quad \mu_2]^T$ เป็นเวกเตอร์ค่าเฉลี่ย

และ $\Sigma = \begin{bmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 \\ \rho\sigma_1\sigma_2 & \sigma_2^2 \end{bmatrix}$ เป็นเมทริกซ์ของความแปรปรวนร่วมของประชากร

ดังนั้น $Z=\left(\underset{\sim}{X}-\underset{\sim}{\mu}_0\right)^T \Sigma^{-1}\left(\underset{\sim}{X}-\underset{\sim}{\mu}_0\right)$ เมื่อ $\underset{\sim}{\mu}_0=[\mu_{01} \quad \mu_{02}]^T$ เป็นเวกเตอร์ค่าเฉลี่ยประชากรเริ่มต้นของตัวแปรสุ่ม X_1 และ X_2

โดยที่ Z ว่ามีการแจกแจงไคสแควร์เบี่ยงเบนไปจากศูนย์กลาง ซึ่งมีระดับองศาความเป็นอิสระ(d.f.) เท่ากับ 2 และ พารามิเตอร์เบี่ยงเบนไปจากศูนย์กลางคือ $\lambda=\left(\underset{\sim}{\mu}-\underset{\sim}{\mu}_0\right)^T \Sigma^{-1}\left(\underset{\sim}{\mu}-\underset{\sim}{\mu}_0\right)$ เขียนสัญลักษณ์แทนด้วย $Z\sim \chi^2_{(2,\lambda)}$ ดังนั้นฟังก์ชันความน่าจะเป็นของ Z จะอยู่ในรูปของ

$$f(z)=\sum_{j=0}^{\infty}\frac{\exp\left(-\frac{\lambda}{2}\right)\left(\frac{\lambda}{2}\right)^j}{j!}\frac{z^j\exp(-z/2)}{\Gamma(1+j)2^{1+j}}$$

กรณี ถ้า $\underset{\sim}{\overline{X}}=[\overline{X}_1 \quad \overline{X}_2]^T$ เป็นเวกเตอร์ค่าเฉลี่ยของตัวอย่างสุ่มขนาด n จากประชากรที่มาการแจกแจงแบบทวิคูณปกติ กล่าวคือ $\underset{\sim}{\overline{X}}\sim N\left(\underset{\sim}{\mu},\frac{1}{n}\Sigma\right)$ ซึ่งมีระดับองศาความเป็นอิสระ (d.f.) เท่ากับ 2 และพารามิเตอร์เบี่ยงเบนไปจากศูนย์กลางคือ $\lambda=n\left(\underset{\sim}{\mu}-\underset{\sim}{\mu}_0\right)^T \Sigma^{-1}\left(\underset{\sim}{\mu}-\underset{\sim}{\mu}_0\right)$ ซึ่งรายละเอียดเพิ่มเติมสามารถศึกษาได้จาก Anderson (1984:75-77)

2.4 แผนภูมิควบคุมที่ใช้ในการทดสอบ

จุดมุ่งหมายของระเบียบวิธีทางสถิติในการควบคุมคุณภาพคือการแยกความแปรผันที่เกิดขึ้นโดยบังเอิญ (Chance variation) ออกจากความแปรผันที่มีสาเหตุระบุได้ (Assignable variation) เพื่อที่จะทราบได้ เมื่อมีความแปรผันที่มีสาเหตุระบุได้ เกิดขึ้นขณะดำเนินการผลิต

ในการศึกษากลุ่มของข้อมูลใดๆจากกระบวนการผลิต ถ้าความแปรผันที่เกิดขึ้นมีแบบแผนที่เป็นไปตามกฎเกณฑ์ทางสถิติและมีเหตุผลที่เชื่อได้ว่าความแปรผันดังกล่าวเกิดขึ้นเพราะความแปรผันที่เกิดขึ้นโดยบังเอิญ ถือว่าไม่มีความแปรผันที่มีสาเหตุระบุได้เกิดขึ้น เมื่อทำการทดสอบซ้ำมากๆครั้งอาจประมาณปริมาณและลักษณะของความแปรผันได้โดยอาศัยทฤษฎีสถิติ จะกล่าวได้ว่ากระบวนการผลิตอยู่ภายใต้การควบคุม (in control)

ในทางตรงกันข้าม ถ้าผลที่ได้รับมีลักษณะไม่เป็นตามธรรมชาติ ซึ่งมีเหตุผลที่เชื่อได้ว่าจะต้องมีสิ่งผิดปกติที่ระบุได้ 1สาเหตุหรือมากกว่า 1สาเหตุเกิดขึ้น ตามเงื่อนไขนี้จะกล่าวได้ว่ากระบวนการผลิตอยู่นอกการควบคุม (out of control)

การใช้แผนภูมิควบคุมในการวิเคราะห์ข้อมูลหรือควบคุมกระบวนการผลิต จะต้องมีการพิจารณาในเรื่องการเกิดขึ้นของความคลาดเคลื่อน 2 ประเภทคือ

1. ความคลาดเคลื่อนประเภทที่ 1 (Type I error) เกิดจากการตัดสินใจว่า กระบวนการผลิตอยู่นอกการควบคุม ทั้งที่ความจริงแล้วกระบวนการผลิตอยู่ภายใต้การควบคุม ความน่าจะเป็นที่จะเกิดความคลาดเคลื่อนประเภทนี้ ก็คือ α
2. ความคลาดเคลื่อนประเภทที่ 2 (Type II error) เกิดจากการตัดสินใจว่า กระบวนการผลิตอยู่ภายใต้การควบคุม ทั้งที่ความจริงแล้วกระบวนการผลิตอยู่นอกการควบคุม ความน่าจะเป็นประเภทนี้ ก็คือ β

2.4.1 แผนภูมิควบคุมเชิงเดี่ยว $\bar{X}$ (Univariate $\bar{X}$ - Chart)

แผนภูมิควบคุมค่าเฉลี่ยเชิงเดี่ยว เป็นแผนภูมิที่แสดงความผันแปรของค่าเฉลี่ยต่างๆ ของตัวอย่างผลิตภัณฑ์ที่สุ่มมาตรวจสอบว่ากระบวนการผลิตอยู่ภายใต้ขอบเขตควบคุมของคุณภาพที่กำหนดไว้หรือไม่ นั่นคือใช้ควบคุมคุณภาพของผลิตภัณฑ์โดยมีค่าเฉลี่ย (Mean) เป็นตัวกำหนดคุณภาพของผลิตภัณฑ์นั้น

ลักษณะของแผนภูมิประกอบด้วยเส้นแกนกลาง ขีดจำกัดการควบคุมเส้นบนและขีดจำกัดการควบคุมเส้นล่าง โดยมีแกนนอนแทนกลุ่มที่ของตัวอย่างแต่ละชุด (Subgroup number) และแกนตั้งแทนค่าเฉลี่ยของตัวอย่าง $(\bar{X})$

ที่มาของแผนภูมิควบคุมเชิงเดี่ยว $\bar{X}$

ในกรณีที่ต้องการทดสอบสมมติฐานค่าเฉลี่ยแยกทีละตัวแปร

H_{0i} : $\mu_i = \mu_{0i}$ กระบวนการผลิตอยู่ภายใต้การควบคุมทางสถิติ

แย้งกับ H_{1i} : $\mu_i \neq \mu_{0i}$ กระบวนการผลิตไม่อยู่ภายใต้การควบคุมทางสถิติ

เมื่อ μ_{0i} คือค่าเฉลี่ยที่แท้จริงของตัวแปร X_i

ในกรณีที่พิจารณาขอบเขต 3σ จะมีค่า $P(\text{Type I error}) = \alpha = 0.00135$

ข้อสมมติ (Assumption): เนื่องจากตัวแปร X_i แต่ละหน่วยต่างมีการแจกแจงแบบปกติ นั่นคือ $X_{i1}, X_{i2}, \ldots, X_{in} \sim N\left(\mu_i, \sigma_i^2\right)$; $i = 1,2$ โดยทราบค่า σ^2 จากหลักการทดสอบสมมติฐานสามารถทราบได้ว่า

จะปฏิเสธ H_{0i} เมื่อ $\dfrac{\bar{X}_i - \mu_i}{\sigma_i/\sqrt{n}} > 3$ หรือ $\dfrac{\bar{X}_i - \mu_i}{\sigma_i/\sqrt{n}} < -3$

ดังนั้น จะได้ $\bar{X}_i > \mu_{0i} + 3\left(\sigma_i/\sqrt{n}\right)$ และ $\bar{X}_i < \mu_{0i} - 3\left(\sigma_i/\sqrt{n}\right)$

ในที่นี้ถือว่าทราบค่าทั้ง μ_{0i} และ σ_i^2

ขอบเขตควบคุม

การควบคุมขอบเขตการควบคุมค่าเฉลี่ย $\bar{X}$ นิยมใช้พื้นที่ภายใต้เส้นโค้งปกติในช่วงความเชื่อมั่น 99.73% นั่นคือที่ระดับนัยสำคัญ (α) 0.27% หรืออีกนัยหนึ่ง ก็คือ การใช้ขอบเขตควบคุม 3σ มีความหมายว่าจำนวนร้อยละ99.73 ที่ข้อมูลตกอยู่ในช่วง $\mu_0 + 3\sigma/\sqrt{n}$ และ $\mu_0 - 3\sigma/\sqrt{n}$ จำนวนร้อยละ 0.27 ตกอยู่นอกขอบเขตของการยอมรับหรือโดยเฉลี่ยแล้วจะมี 27 ตัวอย่างในจำนวน 10,000 ตัวอย่างที่มีค่าเฉลี่ยอยู่นอกการควบคุม

เพราะฉะนั้น ถ้ากระบวนการผลิตอยู่ภายใต้การควบคุม ตลอดทุกกลุ่มตัวอย่างย่อยในช่วงเวลา t ดังนั้นจะได้ขอบเขตควบคุมของแผนภูมิควบคุมเชิงเดี่ยว $\bar{X}$ ได้ดังนี้

$$\boxed{\begin{aligned} UCL_i &= \mu_{0i} + 3\left(\sigma_i/\sqrt{n}\right) \\ LCL_i &= \mu_{0i} - 3\left(\sigma_i/\sqrt{n}\right) \quad ; i = 1,2 \end{aligned}}$$

เกณฑ์การตัดสินใจ

กระบวนการผลิตจะไม่อยู่ภายใต้การควบคุม เมื่อค่าสถิติ $\bar{X}_t$ มีค่ามากกว่าขอบเขตควบคุมบน $\left(\bar{X}_t > UCL\right)$ หรือ มีค่าน้อยกว่าขอบเขตควบคุมล่าง $\left(\bar{X}_t < LCL\right)$

ในการทดสอบสมมติฐานค่าเฉลี่ยพร้อมกันทั้ง 2 ตัวแปร

$H_0 : (\mu_1 = \mu_{01}$ และ $\mu_2 = \mu_{02})$ จะถือว่า H_0 จริง ถ้าสมมติฐานย่อยคือ H_{01} และ H_{02} เป็นจริงทั้งคู่

ดังนั้น ค่าความน่าจะเป็นของความคลาดเคลื่อนชนิดที่ 1 (Type I – error) จะเกิดขึ้น ได้เมื่อปฏิเสธ H_0 โดยทั้งที่ H_{01} และ H_{02} เป็นจริง การปฏิเสธ H_0 นั้นหมายความว่า $\mu_1 \neq \mu_{01}$ หรือ $\mu_2 \neq \mu_{02}$

ถ้า $\bar{X}_1$ และ $\bar{X}_2$ เป็นไม่เป็นอิสระต่อกัน สามารถคำนวณหาระดับนัยสำคัญสำหรับ H_0 ได้ดังนี้

$$P(\text{ปฏิเสธ } H_0 \mid H_0 \text{ จริง}) = 1 - P(\text{ไม่ปฏิเสธ } H_0 \mid H_0 \text{ จริง})$$
$$= 1 - P(\text{ไม่ปฏิเสธ } H_{01} \text{ และ } H_{02} \mid H_0 \text{ จริง})$$

ถ้า $\bar{X}_1$ และ $\bar{X}_2$ เป็นอิสระต่อกัน กล่าวคือสามารถคำนวณหาระดับนัยสำคัญสำหรับ H_0 ได้ดังนี้

$$P(\text{ปฏิเสธ } H_0 \mid H_0 \text{ จริง}) = 1 - (P(\text{ไม่ปฏิเสธ } H_{01} \mid H_{01} \text{ จริง}) \times P(\text{ไม่ปฏิเสธ } H_{02} \mid H_{02} \text{ จริง}))$$

(ทั้งนี้เพราะการทดสอบ H_{01} และ H_{02} เป็นอิสระต่อกัน)

$$= 1 - (1-\alpha)(1-\alpha)$$
$$= 1 - (1-\alpha)^2$$

2.4.2 แผนภูมิควบคุมHotelling[1]

ที่มาของแผนภูมิควบคุม Hotelling

ข้อสมมติเบื้องต้นกำหนดให้ เวกเตอร์สุ่ม $\underset{\sim}{X} = [X_1 \quad X_2]^T$ มีการแจกแจงแบบทวิคูณปกติ (Bivariate normal distribution) สามารถเขียนแทนด้วย $\underset{\sim}{X} \sim N_2\left(\underset{\sim}{\mu}, \Sigma\right)$

โดยที่ เวกเตอร์ค่าเฉลี่ยเป็น $\underset{\sim}{\mu} = [\mu_1 \quad \mu_2]^T$

เมทริกซ์ความแปรปรวนร่วมเป็น $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$ เมื่อ $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$

และค่าสัมประสิทธิ์สหสัมพันธ์ $\rho = \dfrac{\sigma_{12}}{\sqrt{\sigma_1^2\sigma_2^2}}$

---

[1] Montgomery,D.C. Introduction to Statistical Quality Control.3rd.Edition(New York:John Wiley&Sons),1997c.),pp.510-515.

สมมติฐานเพื่อการทดสอบ

$$H_0 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} = \begin{bmatrix} \mu_{01} \\ \mu_{02} \end{bmatrix}$$ กระบวนการผลิตอยู่ภายใต้การควบคุมทางสถิติ

$$H_1 : \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \neq \begin{bmatrix} \mu_{01} \\ \mu_{02} \end{bmatrix}$$ กระบวนการผลิตไม่อยู่ภายใต้การควบคุมทางสถิติ

ตัวสถิติทดสอบ $$Z^2 = n\Delta^2 = n\left(\underset{\sim}{\bar{x}} - \underset{\sim}{\mu_0}\right)^T \Sigma^{-1}\left(\underset{\sim}{\bar{x}} - \underset{\sim}{\mu_0}\right)$$

เนื่องจากเวกเตอร์สุ่ม $\underset{\sim}{X} = [X_1 \quad X_2]^T$ และ $\underset{\sim}{X} \sim N_2\left(\underset{\sim}{\mu}, \Sigma\right)$

ดังนั้น $$\underset{\sim}{Z} = \left(T^T\right)^{-1}\left(\underset{\sim}{X} - \underset{\sim}{\mu_0}\right)$$

โดยที่ $\Sigma = T^T T$ แยกส่วนประกอบโดยใช้วิธีการ Cholesky

จะได้ว่า $\Sigma^{\frac{1}{2}}$ เป็นรากที่สองเมตริกซ์จตุรัสของ Σ $(\Sigma = \Sigma^{\frac{1}{2}}\Sigma^{\frac{1}{2}})$

เวกเตอร์ $\underset{\sim}{Z} \sim N_2\left(\underset{\sim}{0}, I\right)$ เนื่องจากตัวแปรสุ่ม Z_1 และ Z_2 เป็นอิสระต่อกัน

เพราะฉะนั้น $\sum_{i=1}^{2} Z_i^2 = \underset{\sim}{Z}^T \underset{\sim}{Z}$ มีการแจกแจงแบบไคสแควร์ χ_2^2

$$\underset{\sim}{Z}^T \underset{\sim}{Z} = \left(\underset{\sim}{X} - \underset{\sim}{\mu}\right)^T \Sigma^{-1}\left(\underset{\sim}{X} - \underset{\sim}{\mu}\right)$$

เมื่อ $$\Sigma^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} \frac{1}{\sigma_1^2} & -\frac{\rho}{\sigma_1\sigma_2} \\ -\frac{\rho}{\sigma_1\sigma_2} & \frac{1}{\sigma_2^2} \end{bmatrix}$$

ถ้า $\underset{\sim}{X} \sim N_2\left(\underset{\sim}{\mu}, \Sigma\right)$ ดังนั้น $\left(\underset{\sim}{X} - \underset{\sim}{\mu}\right)^T \Sigma^{-1}\left(\underset{\sim}{X} - \underset{\sim}{\mu}\right)$ มีการแจกแจงแบบ χ^2

กำหนดระดับนัยสำคัญ $\alpha_{overall}$ ดังนั้นจะได้ค่าวิกฤตคือ $\chi^2_{(2,\alpha_{overall})}$

เขตปฏิเสธ

จะปฏิเสธ H_0 ถ้า $$\chi^2 > \chi^2_{(2,\alpha_{overall})}$$

เพราะฉะนั้น ถ้ากระบวนการผลิตอยู่ภายใต้การควบคุม ตลอดทุกกลุ่มตัวอย่างย่อยในช่วงเวลา t ดังนั้นจะได้ขอบเขตควบคุมของแผนภูมิควบคุมได้ดังนี้

$$UCL = \chi^2_{(2,\alpha_{overall})}$$
$$LCL \approx 0$$

เกณฑ์การตัดสินใจ

กระบวนการผลิตจะไม่อยู่ภายใต้การควบคุม เมื่อค่าสถิติ χ^2 มีค่ามากกว่าขอบเขตควบคุมบน $\chi^2 > \chi^2_{(2,\alpha_{overall})}$

2.4.3 แผนภูมิควบคุมเชิงพหุ Shewhart $\bar{X}$ (Multivarilate Shewhart $\bar{X}$ - chart)[1]

เนื่องจากเวกเตอร์สุ่ม $\underset{\sim}{X} = [X_1 \quad X_2]^T$ โดยที่ เวกเตอร์สุ่ม $\underset{\sim}{X} \sim N_2\left(\underset{\sim}{\mu},\Sigma\right)$

เมื่อ เวกเตอร์ค่าเฉลี่ย $\underset{\sim}{\mu} = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$

เมทริกซ์ความแปรปรวนร่วม $\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{21} & \sigma_2^2 \end{bmatrix}$ ซึ่งค่าความแปรปรวนร่วม $\sigma_{12} = \sigma_{21} = \rho\sigma_1\sigma_2$ และ ค่าสัมประสิทธิ์สหสัมพันธ์ $\rho = \frac{\sigma_{12}}{\sqrt{\sigma_1^2\sigma_2^2}}$

ทำการแปลงเวกเตอร์สุ่ม $\underset{\sim}{X}$ ให้การแจกแจงแบบปกติมาตรฐาน

$$\begin{bmatrix} \dfrac{X_1 - \mu_1}{\sigma_1} \\ \dfrac{X_2 - \mu_2}{\sigma_2} \end{bmatrix} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \underset{\sim}{Z}$$

[1] Ali A.Houshmand and Saeed Golnabi.Multivariate Shewhart Xbar - Chart (http://interstat.stat.vt.edu/InterStat/Articles/S99004.pdf. 1999),pp.1-10.

จะได้ เวกเตอร์สุ่ม $\underset{\sim}{Z} \sim N_2\left(\underset{\sim}{0}, R\right)$

กำหนดให้ $\rho = \sin 2\theta$ ดังนั้น เมทริกซ์สหสัมพันธ์คือ $R = \begin{bmatrix} 1 & \sin 2\theta \\ \sin 2\theta & 1 \end{bmatrix}$

ทำการแยกตัวประกอบของเมทริกซ์สหสัมพันธ์ R โดยใช้หลักการเมทริกซ์สมมาตร

$$R^{\frac{1}{2}} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ ซึ่ง } R = R^{\frac{1}{2}}R^{\frac{1}{2}}$$

เมื่อหาอินเวอร์สของเมทริกซ์ $R^{-\frac{1}{2}}$ จะอยู่ในรูป

$$R^{-\frac{1}{2}} = \begin{bmatrix} \dfrac{\cos\theta}{\cos 2\theta} & \dfrac{-\sin\theta}{\cos 2\theta} \\ \dfrac{-\sin\theta}{\cos 2\theta} & \dfrac{\cos\theta}{\cos 2\theta} \end{bmatrix}$$

จาก $\rho = \sin 2\theta$ ทราบค่า $\theta = \dfrac{\sin^{-1}\rho}{2}$ และนำค่า θ ที่คำนวณได้ไปแทนในเมทริกซ์ $R^{-\frac{1}{2}}$

คำนวณหาตัวสถิติทดสอบสามารถเขียนอยู่ในรูป เวกเตอร์ $\underset{\sim}{M}$ ดังนี้

$$\underset{\sim}{M} = \left(R^{\frac{1}{2}}\right)^{-1} \underset{\sim}{Z} = \begin{bmatrix} \dfrac{\cos\theta}{\cos 2\theta} & \dfrac{-\sin\theta}{\cos 2\theta} \\ \dfrac{-\sin\theta}{\cos 2\theta} & \dfrac{\cos\theta}{\cos 2\theta} \end{bmatrix} \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix} = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix}$$

ดังนั้น เวกเตอร์ของตัวสถิติ $\underset{\sim}{M} \sim N_2\left(\underset{\sim}{0}, I\right)$ ซึ่ง ตัวสถิติ M_1 และ M_2 มีการแจกแจงแบบปกติ $N(0,1)$ เป็นอิสระต่อกัน

กำหนดระดับนัยสำคัญ

ขั้นแรก กำหนด $\alpha_{overall}$ ดังนั้นสามารถหา α_i

โดยที่ $\alpha_i = 1 - \sqrt{(1 - \alpha_{overall})}$

ขอบเขตปฏิเสธ

คำนวณหาค่าวิกฤต $c_i = z_{\frac{\alpha_i}{2}}$ ($z_{\frac{\alpha_i}{2}}$ มาจากตารางการแจงแจงแบบปกติมาตรฐาน)

ดังนั้น จะปฏิเสธ H_0 ถ้า $|M_1| > c_1$ หรือ $|M_2| > c_2$

เพราะฉะนั้น ถ้ากระบวนการผลิตอยู่ภายใต้การควบคุม ตลอดทุกกลุ่มตัวอย่างย่อยในช่วงเวลา t ดังนั้นจะได้ขอบเขตควบคุมของแผนภูมิควบคุมค่าเฉลี่ยได้ดังนี้

$$UCL_i = c_i = z_{\frac{\alpha_i}{2}}$$

$$LCL_i = -c_i = -z_{\frac{\alpha_i}{2}} \ ; \ i = 1,2$$

เกณฑ์การตัดสินใจ

กระบวนการผลิตจะไม่อยู่ภายใต้การควบคุม เมื่อค่าสถิติ M_1 หรือ M_2 มีค่ามากกว่าขอบเขตควบคุมบน หรือ มีค่าน้อยกว่าขอบเขตควบคุมล่าง