บทที่ 2

งานวิจัยและทฤษฎีที่เกี่ยวข้อง

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

2.1 ข้อสมมุติฐานในการวิเคราะห์

วิธีการหนึ่งที่ใช้ประเมินความน่าจะเป็น (Probability) ของค่าใช้จ่ายโครงการก่อสร้างอาคารก็คือวิธีการ "Monte Carlo Simulation" วิธีการนี้จะมีการตั้งสมมุติฐานว่าค่าใช้จ่ายทุก ๆ ตัวมีศักยภาพสูงสำหรับความผันแปร (Variability) ค่าใช้จ่ายของรายการต่าง ๆ ถูกประมาณขึ้นต่อหน่วยพื้นที่แล้วหาค่าการกระจายทางสถิติและนำมาคำนวณหาค่าใช้จ่ายทั้งหมดของโครงการ วิธีนี้ถูกทำซ้ำ ๆ และวัดค่าการกระจายสะสมของค่าใช้จ่ายทั้งหมดของโครงการ และนำมาคำนวณหาค่าความน่าจะเป็น (Probability) ของโครงการ หรือความ-ผันแปรของค่าใช้จ่ายของโครงการ

วิธี Monte Carlo Technique สามารถที่จะหาค่าความเสี่ยงของกำไร (Profit Risk) สำหรับค่าใช้จ่ายทั้งหมดของโครงการ ความน่าจะเป็นในการประมาณซึ่งเป็นธรรมชาติของความไม่แน่นอนของค่าใช้จ่าย โดยค่าใช้จ่ายรวมของโครงการเป็น Model ของความไม่แน่นอนมากกว่าจำนวนเงินคงที่จำนวนหนึ่ง (Fix Sum) องค์ประกอบของค่าใช้จ่ายต่าง ๆ (Cost Component) ของแต่ละโครงการสามารถแสดงรายการโดยใช้หลักการกระจายทางสถิติ ซึ่งประกอบด้วยค่าเฉลี่ย (Mean) และค่าความผันแปร (Variance) และค่าใช้จ่ายรวมของโครงการจะเป็น "Random Variable" ซึ่งเกิดจากการรวมกันของ"Random Number" โดย

$$C_{tot} = \sum_{i=1}^{n} C_i \qquad \text{-------------(1)}$$

เมื่อ C_{tot} = ค่าใช้จ่ายรวมของโครงการต่อหน่วยพื้นที่

C_i = องค์ประกอบของค่าใช้จ่ายของโครงการต่อหน่วยพื้นที่

ถ้าต้องการพิจารณาความผันแปรของค่าใช้จ่าย (Cost Variation) ในทุกส่วนขององค์ประกอบของค่าใช้จ่ายก็คือ รายละเอียดในการประมาณราคาที่ถูกประมาณและสรุปขึ้นมา เมื่อพิจารณาแล้วส่วนใหญ่จะไม่มีความแน่นอนเนื่องจากความผันแปรขององค์ประกอบของค่าใช้จ่ายต่าง ๆ ดังนั้นรายการต่าง ๆ ขององค์ประกอบของค่าใช้จ่ายซึ่งมีศักยภาพสูงของความผันแปรและถูกพิจารณาเป็น "Random Variable" ฉะนั้นองค์ประกอบของค่าใช้จ่ายต่างๆ จะไม่มีความแน่นอนตายตัว จะมีความผันแปร ซึ่งค่า C_{tot} ในสมการที่ (1) จะประกอบไปด้วยส่วนหนึ่งคือ "Fix Sum" และส่วนของ "Random Component" โดย C_i สามารถหาค่าการกระจายผันแปรและรวบรวมเป็น C_{tot} โดย Monte Carlo Simulation สามารถที่จะหาค่าความน่าจะเป็นผันแปรขององค์ประกอบของค่าใช้จ่ายต่าง ๆ ของโครงการ ซึ่งประกอบด้วยการหาค่าความผันแปรของ C_i ที่กระจายและรวบรวมรายการเหล่านี้กับส่วนของค่าใช้จ่ายที่ไม่ผันแปร (Fix Sum) ก็จะได้ค่าใช้จ่ายรวมของโครงการ กระบวนการนี้จะถูกทำซ้ำหลายครั้งโดยแต่ละครั้งค่า C_{tot} จะถูกคำนวณและนำมาสร้าง Histrogram และ Cumulative Distribution Function (CDF) และนำค่าเหล่านี้มาใช้ในการประมาณความน่าจะเป็นของความสำเร็จของโครงการ หรือความน่าจะเป็นของงบประมาณที่จะใช้จ่ายในโครงการ การใช้วิธีการมอนติคาร์โล ในการหาความน่าจะเป็นของการประมาณราคาค่าก่อสร้างของโครงการนั้นสิ่งแรกที่จำเป็นต้องสร้างคือ การกระจายทางสถิติของส่วนความผันแปรขององค์ประกอบของค่าใช้จ่ายต่าง ๆ ประการที่สอง ตัวความผันแปรนั้นไม่อิสระ ความผันแปรเหล่านั้นต้องหาความสัมพันธ์ซึ่งเป็นปัญหาในการใช้วิธีการมอนติคาร์โล

2.2 ทฤษฎีที่เกี่ยวข้อง

2.2.1 การกระจายทางสถิติของค่าใช้จ่ายต่างๆ

ในการหาการกระจายทางสถิตินั้น ตัวพารามิเตอร์ของการกระจายที่เกี่ยวข้องก็คือ ค่าเฉลี่ย (Mean) ค่าความผันแปร (Variance) และขอบเขตของเหตุการณ์ (Range of Event) ซึ่งไม่เพียงพอในการอธิบาย จำนวนสุ่ม (Random Number) ภายใต้การกระจายของความผันแปร องค์ประกอบของค่าใช้จ่ายต่าง ๆ จึงจำเป็นต้องใช้วิธีการมอนติคาร์โล

ในการหาความน่าจะเป็นของค่าใช้จ่ายของโครงการ ข้อมูลอาจจะมีการกระจายเป็นแบบต่าง ๆ เช่น Normal , Lognormal , Triangular , Beta หรือ Uniform Distribution ซึ่งจากการทดสอบตัวอย่างของ AbouRizk และ Halpin (1990) ได้พิสูจน์ว่า การกระจายแบบ Lognormal จะมีค่าใกล้เคียงความจริงมากที่สุด รองลงมาก็คือ การกระจายแบบ Beta สำหรับโครงการก่อสร้างอาคาร

2.2.2 ความสัมพันธ์ระหว่างองค์ประกอบของค่าใช้จ่ายต่าง ๆ ของโครงการ

ในการหาความน่าจะเป็นของค่าใช้จ่ายของโครงการนั้น สิ่งที่จะทำให้เกิดความผิดพลาด (Error) ก็คือ องค์ประกอบของค่าใช้จ่ายที่สมมุติว่าอิสระ การเปลี่ยนแปลงของค่าใช้จ่ายเหล่านี้จะไม่มีผลต่อ ค่าใช้จ่ายตัวอื่น ๆ ซึ่งข้อสมมุติฐานเหล่านี้ทำให้เกิดความไม่ถูกต้อง

แต่อย่างไรก็ตาม ถ้าความสัมพันธ์ระหว่าง องค์ประกอบของค่าใช้จ่ายต่าง ๆ เหล่านี้มีเล็กน้อยก็ให้สมมุติว่า องค์ประกอบของค่าใช้จ่ายต่าง ๆ เป็นอิสระต่อกัน การประมาณค่าใช้จ่ายต้องใช้วิธีการประมาณในรูปของความสัมพันธ์

2.2.3 วิธีการประมาณค่าใช้จ่ายต่าง ๆ ที่มีความสัมพันธ์กัน

ในการรวมกันขององค์ประกอบของค่าใช้จ่ายต่าง ๆ ซึ่งมีความสัมพันธ์กัน โดยพิจารณาดังตัวอย่าง ซึ่งค่าใช้จ่ายของโครงการประกอบด้วย 10 รายการ C_1 ถึง C_{10} จะได้ว่า

$$C_{tot} = \sum_{n=1}^{10} C_i \qquad \text{------------(2)}$$

สมมุติว่า C_4, C_5, C_6 มีความสัมพันธ์กันมาก และ C_9 มีความสัมพันธ์กับ C_{10} ซึ่งกำหนดให้เป็น C' เป็น C'' จะได้ว่า

$$C' = C_4 + C_5 + C_6 \qquad \text{------------(3)}$$

$$C'' = C_9 + C_{10} \qquad \text{------------(4)}$$

ผู้ประมาณราคาสามารถหาการกระจายและพารามิเตอร์ของ C' และ C'' และส่วนที่เหลืออยู่ขององค์ประกอบของค่าใช้จ่ายต่าง ๆ โดยสมมุติว่ามีความอิสระต่อกัน ($C_1, C_2, C_3, C', C_7, C_8$ และ C'') ดังนั้นสมการ (2) สามารถเขียนใหม่ตามวิธีการของ Monte Carlo Simulation ได้ดังนี้

$$C_{tot} = C_1 + C_2 + C_3 + C' + C_7 + C_8 + C'' \qquad \text{------------(5)}$$

2.2.4 วิธีการวิเคราะห์หาความสัมพันธ์

ในการวิเคราะห์หาความสัมพันธ์ สำหรับการกระจายของค่าใช้จ่ายต่าง ๆ เมื่อองค์ประกอบของค่าใช้จ่ายต่าง ๆ มีความสัมพันธ์กัน จำเป็นต้องรวมฟังก์ชั่นต่าง ๆ เข้าด้วยกัน (Joint Density Function) ถ้าปราศจากการรวมฟังก์ชั่นต่าง ๆ เข้าด้วยกัน จะไม่สามารถนำ Monte Carlo Simulation มาใช้ได้ ซึ่งการรวมฟังก์ชั่นต่าง ๆ เข้าด้วยกันก็คือการประมาณค่าใช้จ่ายของโครงการ หรือการวิเคราะห์ความเสี่ยงขององค์ประกอบของค่าใช้จ่าย ซึ่งถ้าองค์ประกอบของค่าใช้จ่ายต่าง ๆ ถูกสมมุติว่ามีการ กระจายแบบ Lognormal Distribution ในการหาความสัมพันธ์ขององค์ประกอบของ ค่าใช้จ่ายเหล่านั้นสามารถใช้ "Multivariate Normal Distribution" โดย Multivariate Lognormal Distribution สามารถแปลงเป็น Multivariate Normal Distribution แล้วนำมาวิเคราะห์หาความสัมพันธ์ของความผันแปรของค่าใช้จ่ายของโครงการ

2.2.5 สัมประสิทธิ์ความสัมพันธ์ (Correlation Coefficient)

การวิเคราะห์เพื่อหาสัมประสิทธิ์ความสัมพันธ์ (Correlation Coefficient) ระหว่างตัวแปร 2 ตัว หาได้จากความสัมพันธ์ดังนี้

$$\rho = \frac{Cov\ (X,Y)}{\sigma_x . \sigma_y} \quad \text{-----------(6)}$$

ซึ่งเราสามารถประมาณค่า จากข้อมูลที่สังเกตุได้ดังนี้

$$\hat{\rho} = \frac{\frac{1}{n-1} \sum_{i=1}^{n} X_i Y_i - n \bar{X} \bar{Y}}{S_x S_y} \quad \text{-----------(7)}$$

ค่า $\hat{\rho}$ อยู่ระหว่าง -1 ถึง +1

ค่า $\hat{\rho}$ เข้าใกล้ +1 หรือ -1 แสดงว่า X, Y มีความสัมพันธ์กันเป็นเส้นตรงมาก

ค่า $\hat{\rho} \simeq 0$ แสดงว่าความสัมพันธ์ของตัวแปรไม่เป็นเส้นตรง

โดยที่ $\bar{X}$, $\bar{Y}$ และ S_x , S_y เป็นค่าเฉลี่ยและค่าความแปรปรวนของกลุ่มตัวอย่าง n ข้อมูล ซึ่งเราสามารถประมาณค่า $r'_{ij} = \hat{\rho}$ เมื่อ r'_{ij} คือค่าความสัมพันธ์ระหว่างตัวแปรที่ i กับตัวแปรที่ j ของ Lognormal Distribution

2.2.6 การกระจายแบบ Lognormal

ในที่นี้เราตั้งสมมุติฐานว่า ข้อมูลของค่าใช้จ่ายของโครงการก่อสร้างอาคารมีการกระจายแบบ Lognormal Distribution ข้อมูลจะถูกแปลงเป็น Normal Distributions โดยสมการดังต่อไปนี้ (Johnson และ Ramberg 1978)

$$\mu_i = \ln \left(\frac{\mu_i'^2}{\sqrt{\sigma_i'^2 + \mu_i'^2}} \right) \quad \text{----------(8)}$$

$$\sigma_i^2 = \ln\left(1 + \frac{\sigma_i'^2}{\mu_i'^2}\right) \quad \text{---------(9)}$$

$$r_{ij} = \frac{1}{\sigma_i \cdot \sigma_j}\left(\ln\ 1 + r'_{ij}\left|\frac{\sigma'_i \cdot \sigma'_j}{\mu'_i \cdot \mu'_j}\right|\right) \quad \text{--------(10)}$$

เมื่อ μ_i , σ_i^2 และ r_{ij} , μ'_i , $\sigma_i'^2$, r'_{ij} เป็น Mean, Variances และ Correlation Coefficients ของ Normal Distribution และ Mean, Variances และ Correlation Coefficients ของ Lognormal Distribution

2.2.7 การกระจายแบบ Normal

N-Dimension Multivariate Distribution กับค่าเฉลี่ย คือ $(\mu_1, \mu_2, \ldots, \mu_n)^t$ (t แสดง Transpose ของเวคเตอร์) และ Covariance Matrix Z. เมื่อ σ_{ij} (Covariance ระหว่างความผันแปร i และ j) ใน Matrix ที่มี Joint Density Function ดังนี้ (Law และ Kelton 1991)

$$f_c(c) = (2\pi)^{-n/2}|Z|^{-0.5} \exp \frac{[-(c-\mu)^t \ Z^{-1}(c-\mu)]}{2} \quad \text{-----(11)}$$

เมื่อ $C = (C_1 , C_2 , \ldots.)^t$ คือ Random Variates (Cost Component) และ $|Z|$ คือ Determinate ของ Z เมื่อ $C = (C_1 , C_2 , \ldots.)^t$ เป็น Normal (μ, Z), $\mathrm{cov}\ (C_i, C_j) = \sigma_{ij} = r_{ij} \cdot \sigma_i \cdot \sigma_j$ และ Z คือ Symmestric ($\sigma_{ij} = \sigma_{ji}$) และเป็นค่าบวก Fishman (1978) แนะนำวิธีการที่หาค่า Multivariate Normal

Distribution โดยมีการรวมกันระหว่าง Finishman's Algorithm กับ Algorithm ของ John และ Ramborg (1978) ที่นำมาใช้ในการแสดง Multivariate Lognormal Distribution (Law และ Kelton 1991) โดย Lognormal Distribution เป็นทางเลือกที่ดีในการแสดงค่าใช้จ่ายของโครงการก่อสร้างทางการเงินของอาคาร และ Z เป็น Covariance Matrix และเป็น Symmetrix มีค่าเป็นบวก สามารถเขียนได้ว่า $Z = XX^t$ เมื่อ X คือ n x n Lower Triangular Matrix (Low and Kelton 1991 , Fishman 1978) โดย C สามารถเขียนได้ดังนี้

$$C = \mu + X.N \qquad \text{----------(12)}$$

$$C_i = \mu_i + \sum_{j=1}^{i} X_{ij} N_j \qquad \text{----------(13)}$$

เมื่อ N_j เป็น Independently Normal Distribution Random Variable กับ Mean = 0 และค่า Variance = 1 [IID Normal (0,1)] และ X_{ij} เป็นสมาชิกของ Matrix เมื่อ i คือ Row และ j คือ Column ของ Matrix ซึ่งสามารถหาได้ดังนี้ (Fishman 1978)

$$X_{i1} = \frac{\sigma_{i1}}{\sigma_{11}} \qquad \text{สำหรับ } i = 1,\dots,n \qquad \text{----------(14)}$$

ถ้า i = 2

$$X_{ii} = \sqrt{\sigma ii - \sum_{j=1}^{i-1} X_{ij}} \qquad \text{----------(15)}$$

ถ้า $i = n$ แล้ว ไม่สามารถหาค่า X ได้ ให้เพิ่มค่า i อีก 1

$$i \longrightarrow i+1 \text{ , } X_{ij} = \frac{\sigma ij - \sum_{m-1}^{j-1} X_{im} \cdot X_{jm}}{X_{jj}} \text{ สำหรับ } j = 2,\ldots,i-1 \qquad \text{--------(16)}$$

แล้วกลับไปยังสมการที่ (15)

การคำนวณค่า C_i ตามสมการที่ (13) นั้น ต้องแปลงข้อมูลเป็น Normal ก่อน เมื่อได้ค่า C_i เรียบร้อยแล้วจึงแปลงค่า C_i กลับเป็น Lognormal ดังเดิมโดย

$$LnC_i = \mu_i + \sum_{j=1}^{i} X_{ij} \cdot N_j \qquad \text{----------(17)}$$

$$C_i = e^{LnC i} \qquad \text{----------(18)}$$

หลังจากค่าใช้จ่ายของรายการต่าง ๆ ที่มีความสัมพันธ์กัน ถูกประมาณขึ้นตามสมการที่ (18) แล้ว ค่าเหล่านี้จะนำมารวมกับค่าใช้จ่ายที่เป็นอิสระต่อกัน ก็จะเป็นค่าใช้จ่ายรวมของโครงการ กระบวนการนี้จะถูกกระทำซ้ำหลายครั้ง โดยแต่ละครั้งจะได้ค่าใช้จ่ายรวมของโครงการ และนำค่าใช้จ่ายรวมของโครงการเหล่านี้มาหาค่าการกระจายทางสถิติ โดยค่าใช้จ่ายรวมของโครงการที่ประมาณขึ้นมานี้ จะถูกตรวจสอบโดยวิธี Chi-Square (x^2) โดยใช้ระดับความเชื่อถือ 1% Significance Level

2.2.8 การทดสอบการกระจายแบบ Chi - Square

พิจารณาตัวแปรสุ่มที่มีข้อมูล n ตัว การทดสอบการกระจายด้วยวิธี Chi-Square จะเป็นการเปรียบเทียบถึงความถี่ของ $n_1, n_2, n_3 \ldots n_k$ (k ค่า)ของตัวแปรกับความถี่ $e_1, e_2, e_3 \ldots e_k$ ซึ่งการกระจายแบบ Chi-Square (x^2) โดยมี Degree of Freedom (f=k-1) และ $C_{\infty, f}$ เป็นค่าที่เหมาะสมของการกระจายแบบ x^2 ที่ Cumulative Probability (∞) แล้วการกระจายทางทฤษฎีที่กำหนดขึ้นจะเป็นแบบจำลองที่ยอมรับได้จะต้องมีระดับนัยสำคัญ (Significance Level) ∞

$$\sum_{i=1}^{k} \frac{(n_i - e_i)^2}{e_i} < C_{\infty, f} \qquad \text{----------(19)}$$

การใช้ x^2 Test สำหรับ Goodness of Fit โดยทั่ว ๆ ไปจะต้องมี $k \geq 5$ และ $e_i \geq 5$ ถึงจะได้ผลเป็นที่พอใจ