บทที่ 2
เอกสารและงานวิจัยที่เกี่ยวข้อง

2.1 แนวคิดและทฤษฎี

2.1.1 หลักสถิติที่ใช้ในการวิจัย

จุดมุ่งหมายของงานวิจัยนี้คือต้องการแสดงให้เห็นถึงปัจจัยที่มีผลต่อการจำแนกกลุ่มบริษัทประกันวินาศภัยที่มีปัญหาทางความมั่นคงทางการเงินและบริษัทที่ไม่มีปัญหาทางการเงินออกจากกัน เพื่อใช้ในการทำนายความมั่นคงของบริษัทประกันวินาศภัยต่อไป นับแต่ในอดีตจนถึงปัจจุบันนั้นได้มีนักวิชาการหลายฝ่ายที่มุ่งศึกษาโดยการใช้วิธีการวิเคราะห์อัตราส่วน (Ratio Analysis) ซึ่งทำให้ได้ข้อสรุปที่ว่า บริษัทที่มีปัญหาทางการเงินนั้นจะมีความแตกต่างในอัตราส่วนทางการเงินไปจากบริษัทที่ไม่มีปัญหาทางการเงิน ซึ่งตามที่ได้กล่าวมาแล้วในข้างต้นที่ว่าการวิเคราะห์ด้วยระบบอัตราส่วนนั้นมีจุดอ่อนที่สำคัญ คือ การตีความหมายในผลของการวิเคราะห์นั้นอาจเกิดความสับสนและผิดพลาดได้ง่าย เนื่องจากเป็นการแยกพิจารณาอัตราส่วนทีละอัตราส่วน เช่น การมีอัตราส่วนจำนวนหนึ่งที่ให้ผลสรุปที่ขัดแย้งกับอัตราส่วนที่เหลือ ทำให้เกิดปัญหาในด้านการตีความและถ้าหากการวิเคราะห์ในรูปแบบนี้กระทำการโดยผู้วิเคราะห์หลายคนอาจจะทำให้ผลลัพธ์ยิ่งมีความแตกต่างกันออกไปอีก ดังนั้นเพื่อเป็นการแก้ไขปัญหาดังกล่าว จึงได้มีผู้พยายามนำเสนอวิธีการทางสถิติ (Statistic Approach) มาใช้แทนการวิเคราะห์ด้วยระบบอัตราส่วน

ผู้วิจัยจึงได้เลือกใช้วิธีการวิเคราะห์จำแนกกลุ่ม (Discriminant Analysis) เพื่อศึกษาและสร้างเป็นตัวแบบจำลองทางสถิติในการแสดงว่ามีอัตราส่วนทางการเงินใดบ้างที่สามารถใช้บ่งชี้ถึงสถานะความมั่นคงของบริษัทประกันวินาศภัยของประเทศไทย เพื่อการพัฒนาเป็นดัชนีชี้นำถึงสถานะความมั่นคงในลำดับต่อไป ซึ่งวิธีการวิเคราะห์จำแนกกลุ่มเป็นวิธีการหนึ่งในการวิเคราะห์ตัวแปรหลายตัวแปร (Multivariate Analysis) อันเป็นวิธีการจำแนกกลุ่มตัวอย่างที่ต้องการศึกษาออกเป็นกลุ่มย่อยๆ ตั้งแต่ 2 กลุ่มขึ้นไป เช่น การศึกษาระดับความยากจนของชุมชน โดยการแบ่งกลุ่มย่อยๆออกเป็นหมู่บ้านที่ยากจนมาก หมู่บ้านที่ยากจน หมู่บ้านฐานะปานกลาง และหมู่บ้านที่พัฒนาแล้ว หรือกลุ่มคนที่ต้องการมีบุตร 2 คนและกลุ่มคนที่ต้องการมีบุตรมากกว่า 2 คน สำหรับในงานวิจัยนี้จะมีการศึกษากลุ่มย่อย 2 กลุ่มคือ กลุ่มที่มีความมั่นคงและกลุ่มที่มีความไม่มั่นคงในสถานภาพทางการเงินของบริษัทประกันวินาศภัยในประเทศไทยที่ประกอบธุรกิจ ในช่วงระยะเวลาตั้งแต่ ปี พ.ศ.2537 – พ.ศ.2541

2.1.2 วัตถุประสงค์ของการวิเคราะห์จำแนกกลุ่ม

วัตถุประสงค์ของการวิเคราะห์จำแนกกลุ่มนี้ คือ การวิเคราะห์ (Analysis) และการจำแนก (Classification) โดยอาศัยข้อมูลที่เก็บรวบรวมมาอันจะทำให้ทราบว่าความเป็นสมาชิกของหน่วยวิเคราะห์ขึ้นอยู่กับปัจจัยใดบ้าง เพื่อการนำไปใช้ประโยชน์ในเชิงปฏิบัติในการที่จะนำความรู้ที่ได้ไปจำแนกหน่วยวิเคราะห์ใหม่ออกเป็นกลุ่มหรือประเภทได้ว่าหน่วยวิเคราะห์นั้นควรจะเป็นสมาชิกในกลุ่มใดได้อย่างถูกต้องและมีประสิทธิภาพมากกว่าการที่จะทำการจำแนกโดยไม่มีข้อมูลที่ได้จากการวิเคราะห์จำแนกกลุ่ม โดยจะทำการคำนวณหาเป็นรูปแบบของส่วนประสมเชิงเส้น (Linear combination) ของตัวแปรอิสระ

2.1.3 คุณสมบัติของวิธีการวิเคราะห์จำแนกกลุ่ม

การวิเคราะห์จำแนกกลุ่มมีคุณสมบัติหลายประการที่มีความคล้ายกับการวิเคราะห์ถดถอยเชิงพหุ (Multiple Regression) โดยเฉพาะอย่างยิ่งในส่วนรูปแบบของสมการ ทำให้ในบางครั้งจึงมีการนำเอาผลการวิเคราะห์การจำแนกกลุ่มไปเปรียบเทียบกับวิธีการวิเคราะห์ถดถอยเชิงพหุ แต่ว่าปัญหาที่ทำการศึกษาในครั้งนี้ ก็คือ การหาปัจจัย (อัตราส่วนทางการเงิน) ที่มีความสามารถแบ่งแยกกลุ่มบริษัทประกันวินาศภัยออกเป็น 2 กลุ่ม คือ กลุ่มที่มีความมั่นคงทางการเงิน และอีกกลุ่มหนึ่งคือกลุ่มที่ไม่มีความมั่นคงทางการเงิน ถ้าหากว่าเราทำการศึกษาด้วยการวิเคราะห์ถดถอยเชิงพหุ ซึ่งเป็นวิธีที่เราจักต้องทราบถึงปัจจัยที่จะมีผลต่อสถานะความมั่นคงของบริษัทประกันวินาศภัยแล้วให้น้ำหนักแก่ปัจจัยทีละตัวตามค่าสัมประสิทธิ์ถดถอยของปัจจัยนั้นๆ ซึ่งนอกจากปัญหาอันเกี่ยวกับความสัมพันธ์ของตัวแปรอิสระแล้วนั้น ยังอาจไม่มีความเหมาะสมในด้านที่ว่าปัจจัยต่างๆที่เกี่ยวข้องนั้นในแต่ละกลุ่มตัวอย่างอาจจะไม่เหมือนกัน หรือหากว่าเหมือนกันน้ำหนักของปัจจัยแต่ละตัวในแต่ละกลุ่มตัวอย่างอาจจะไม่เหมือนกัน ดังนั้นการที่เราทำการบังคับให้ปัจจัยทุกตัวมีผลต่อสถานะความมั่นคงของบริษัทประกันวินาศภัยที่ต่างกลุ่มกันนั้น จึงอาจจะไม่เหมาะสมผู้วิจัยจึงได้แสดงการเปรียบเทียบความแตกต่างและเหมือนกันระหว่างวิธีการวิเคราะห์ถดถอยเชิงพหุกับการวิเคราะห์จำแนกกลุ่มไว้ในภาคผนวก ค. และผู้วิจัยจึงได้ใช้วิธีการวิเคราะห์จำแนกกลุ่ม อันเป็นวิธีการที่จะทำให้ทราว่ามีปัจจัยใดบ้างที่สามารถใช้จำแนกถึงความแตกต่างในสถานะความมั่นคงทางการเงินระหว่างกลุ่มบริษัทประกันวินาศภัยได้ โดยที่ปัจจัยเหล่านี้มีความสามารถจำแนกกลุ่มได้ดีเพียงใด รวมทั้งการมีหลักเกณฑ์อย่างไรในการพิจารณาการจำแนกกลุ่มตัวอย่าง ซึ่งจะกล่าวในลำดับต่อไป

2.1.4 ตัวแบบทางคณิตศาสตร์

จากวัตถุประสงค์ของการวิเคราะห์จำแนกกลุ่มในการที่จะจำแนกกลุ่มตัวอย่างออกเป็นตั้งแต่ 2 กลุ่ม คือ $\P_1$ และ $\P_2$ โดยที่หลักการพิจารณาจะใช้คุณลักษณะของเวกเตอร์ (Vector of Characteristics) $X = (X_1, X_2, \ldots, X_p)$ เป็นเครื่องช่วยในการพิจารณาโดยให้ $f_1(X)$ และ $f_2(X)$ เป็น ฟังก์ชั่นความหนาแน่นความน่าจะเป็น (probability density function) ของ X ในกลุ่มตัวอย่างทั้ง 2 กลุ่มนี้ หรืออาจจะกล่าวได้ว่าวิธีการวิเคราะห์จำแนกกลุ่ม คือการพยายามหาสมการเชิงเส้น (Linear function) ในการที่จะเป็นตัวจำแนกที่ดีที่สุด (Best discriminator) ระหว่างกลุ่มทั้งสอง โดยจะมีการกำหนดค่าตัวแปรตาม (Y) ในการวิจัยนี้จะกำหนดให้ Y มีชื่อว่า Status โดยที่ Status = 0 คือ ลักษณะของกลุ่มบริษัทประกันวินาศภัยที่มีฐานะทางการเงินไม่มั่นคงหรือว่ามีปัญหาทางด้านสภาพคล่องจนต้องให้ความสนใจเป็นพิเศษในการเข้าไปตรวจสอบการดำเนินงานของบริษัท และกำหนดให้ค่า Status =1 คือ ลักษณะของกลุ่มบริษัทประกันวินาศภัยที่มีฐานะทางการเงินที่มั่นคง ซึ่งจะทำการเลือกสัมประสิทธิ์ของสมการที่ทำให้อัตราส่วนระหว่างความแตกต่างในค่าเฉลี่ย 2 กลุ่ม คือ Y_1 และ Y_2 ยกกำลังสองและความแปรปรวน (Variance) ของ Y มีค่าสูงสุด

โดยจะกำหนดให้ค่าเฉลี่ยของ X ในตัวอย่างทั้ง 2 กลุ่ม คือ μ_1 และ μ_2 และ Covariance Matrix ของทั้งสองกลุ่ม คือ $\sum_1$ และ $\sum_2$ (ในที่นี้เรามิได้สมมติว่า X มีการกระจายแบบปกติ (Normal)) แล้ว ค่าเฉลี่ยของสมการเส้นตรง Y ใน 2 กลุ่มก็จะเป็น $b'\mu_1$ และ $b'\mu_2$ และถ้าสมมติว่า $\sum_1$ เท่ากับ $\sum_2$ และเท่ากับ $\sum$ แล้ว Variance ของ Y จะเป็น $b'\sum b$ ดังนั้นเราจะต้องทำการให้มีค่าสูงสุด (Maximize)

$$\Delta = [b'(\mu_1 - \mu_2)]^2 / b'\sum b$$

ด้วยการอนุพันธ์ (Differentiate) Δ เทียบกับ b และให้ค่า Derivative =0 เราจะได้ว่า

$$\Delta' = \frac{2[b'(\mu_1 - \mu_2)](\mu_1 - \mu_2)b'\sum b - 2\sum b[b'(\mu_1 - \mu_2)]^2}{(b'\sum b)^2}$$

หรือ $(\mu_1 - \mu_2)b'\sum b = \sum b[b'(\mu_1 - \mu_2)]$

เราจะได้คำตอบว่า $(\mu_1 - \mu_2) = \sum b$ หรือ $b = \sum^{-1}(\mu_1 - \mu_2)$

หากว่าเราไม่ทราบค่าพารามิเตอร์ก็จะสามารถประมาณค่าพารามิเตอร์ต่างๆได้ โดยสมมติว่ามีตัวอย่างจำนวน n_1 ที่มีค่า Y=1 และตัวอย่างจำนวน n_2 ที่มีค่า Y=0 จะได้ว่า

$$\bar{X}_1 = \sum x_{1i}/n_1 \quad , \quad \bar{X}_2 = \sum x_{2i}/n_2$$

$\overline{X} = (n_1X_1+n_2X_2) / n_1+n_2-2$ และ

$$S = \frac{[\sum_i (X_{1i} - \bar{X}_1)(X_{1I} - \bar{X})' + \sum_i (X_{2i} - \bar{X}_1)'(X_{2i} - \bar{X}_2)']}{n_1 + n_2 - 2}$$

จะได้ว่า $b = S^{-1}(\bar{X}_1 - \bar{X}_2)$

และค่าเฉลี่ยของสมการการจำแนกกลุ่ม 2 กลุ่ม จะเป็น

$$\bar{y}_1 = \hat{b}'\bar{X}_1 = (\bar{X}_1 - \bar{X}_2)'S^{-1}\bar{X}_1$$
$$\bar{y}_2 = \hat{b}'\bar{X}_2 = (\bar{X}_1 - \bar{X}_2)'S^{-1}\bar{X}_2$$

หากมีตัวอย่างใหม่คือ X_0 เราจะได้ว่า

$$y_0 = \hat{b}'\bar{X}_0 = (\bar{X}_1 - \bar{X}_2)'S^{-1}\bar{X}_0$$

เราจะจำแนก y_0 ลงใน ¶$_1$ ถ้าหากว่า y_0 มีค่าใกล้เคียงกับ $\bar{y}_1$ มากกว่า $\bar{y}_2$ นั่นคือ ถ้า $\bar{y}_1 > \bar{y}_2$ แล้ว Y_0 จะมีค่าใกล้เคียง $\bar{y}_1$มากกว่าที่จะใกล้เคียงกับ $\bar{y}_2$ ถ้า $|Y_0 - \bar{Y}_1| < |Y_0 - \bar{Y}_2|$ หรือ $Y_0 > \frac{1}{2}(\bar{Y}_1 + \bar{Y}_2)$ ฉะนั้นจุดที่ใช้แบ่งแยก 2 กลุ่ม ก็คือ จุดกึ่งกลางระหว่างค่าเฉลี่ย 2 กลุ่ม และค่ายกกำลังสองของระยะห่างระหว่างค่าเฉลี่ยทั้งสอง เราเรียกว่า Mahalanobis's Distance และแทนโดยสัญลักษณ์ D^2

$$D^2 = (\bar{Y}_1 - \bar{Y}_2)^2 = (\bar{X}_1 - \bar{X}_2)'S^{-1}(\bar{X}_1 - \bar{X}_2)$$

ในการทดสอบความมีนัยสำคัญทางสถิติ เราจำเป็นจะต้องสมมติว่าตัวแปรที่สามารถอธิบายได้ (Explanatory variable) ของทั้ง 2 กลุ่ม มาจากประชากรที่มีการแจกแจงปกติ (Normal population) ที่มีค่าเฉลี่ยเท่ากับ μ_1 และ μ_2 และมีเมตริกซ์ความสัมพันธ์ร่วม (Covariance matrix) เท่ากันด้วย คือ เท่ากับ $\sum$ การทดสอบนัยสำคัญทางสถิติ คือการทดสอบว่าประชากร 2 กลุ่ม มีความแตกต่างในค่าเฉลี่ย อย่างมีนัยสำคัญหรือไม่ เราจะใช้ตัวสถิติ F

$$F = \frac{n_1 n_2 (n_1 + n_2 - p - 1)}{(n_1 + n_2)(n_1 + n_2 - 2)p} D^2$$

ตัวสถิติ F จะมีองศาอิสระ (degree of freedom) เท่ากับ p และ n_1+n_2-p-1

จะเห็นได้ว่าในการได้ค่า $\hat{b}$ นั้น ไม่มีข้อสมมติใดๆเกี่ยวกับตัวแปร แต่ในการทดสอบความมีนัยสำคัญนั้นจำเป็นจะต้องมี

2.1.5 **ความน่าจะเป็นโดยหลักเกณฑ์ หรือความน่าจะเป็นก่อน** (Prior Probability) **และต้นทุนการจำแนกกลุ่มผิดพลาด** (Misclassification Costs)

กำหนดให้กลุ่ม π_1 และกลุ่ม π_2 มีจำนวนตัวอย่างเป็นสัดส่วน p_1 และ ะ p_2 ตามลำดับ และเลือกพื้นที่ R_1 และ R_2 ซึ่งหากว่าตัวอย่างตกในพื้นที่ R_1 แล้ว เราจะทำการจำแนกตัวอย่างนั้นลงกลุ่ม π_1 และถ้าตัวอย่างนั้นตกลงในพื้นที่ R_2 แล้ว เราจำแนกตัวอย่างนั้นลงในกลุ่ม π_2

สมมติให้ c_1 เป็นต้นทุนในการจำแนกตัวอย่างซึ่งแท้จริงแล้วเป็นสมาชิกของกลุ่ม π_1 แต่ไปจำแนกลงในกลุ่ม π_2 และ c_2 เป็นต้นทุนที่เกิดจากการจำแนกสมาชิกของกลุ่ม π_2 แต่ไปลงในกลุ่ม π_1 ซึ่งต้นทุนในการจำแนกผิดพลาดทั้งหมดจะเท่ากับ

$$C = c_1 p_1 \int_{R2} f_1(x)dx + c_2 p_2 \int_{R1} f_2(x)dx$$

เพราะว่า $$\int_{R2} f_1(x)dx + \int_{R1} f_2(x)dx = 1$$

จะได้ว่า $$C = c_1 p_1[1 - \int_{R1} f_1(x)dx] + c_2 p_2 \int_{R1} f_2(x)dx$$

$$C = c_1 p_1 + \int_{R1}[c_2 p_2 f_2(x)dx - c_1 p_1 f_1(x)dx$$

C นี้จะมีค่าต่ำสุด (Minimized) หากเราเลือก R_1 ที่ทำให้

$$c_2 p_2 f_2(x)dx < c_1 p_1 f_1(x)dx$$

หรือ

$$\frac{f_1(x)}{f_2(x)} > \frac{c_2 p_2}{c_1 p_1}$$

โดยสมมติให้ X มีการแจกแจงแบบปกติ โดยมีค่าเฉลี่ยของประชากร 2 กลุ่ม เป็น μ_1 และ μ_2 โดยมีเมตริกซ์ความสัมพันธ์ร่วม (Covariance matrix) เท่ากับ $\sum$

เพราะว่า

$$f_1(x) = (2\pi)^{-p/2} |\Sigma|^{-1/2} \exp\left[-\tfrac{1}{2}(x - \mu_i)' \Sigma^{-1} (x - \mu_i)\right]$$

จะได้ว่า

$$\frac{f_1(x)}{f_2(x)} = \exp\left[-\tfrac{1}{2}(x - \mu_1)' \Sigma^{-1} (x - \mu_1) + \tfrac{1}{2}(x - \mu_2)' \Sigma^{-1} (x - \mu_2)\right]$$

$$\frac{f_1(x)}{f_2(x)} = \exp\left[(\mu_1 - \mu_2)' \Sigma^{-1} x - \tfrac{1}{2}(\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 + \mu_2)\right]$$

แล้ว Take Log จะได้ว่า

$$(\mu_1 - \mu_2)' \Sigma^{-1} x > \ln \frac{c_2 p_2}{c_1 p_1} + \frac{1}{2}(\mu_1 - \mu_2)' \Sigma^{-1} (\mu_1 + \mu_2)$$

$$b'(x) > \ln \frac{c_2 p_2}{c_1 p_1} + \frac{1}{2} b'(\mu_1 + \mu_2)$$

หากแทนค่า μ_1 ด้วย $\bar{X}_1$ และ μ_2 ด้วย$\bar{X}_2$ และ $\sum$ ด้วย S แล้ว และถ้า $C_2P_2 = C_1P_1$ ผลลัพธ์ที่ได้ก็เป็นเช่นเดียวกับค่าของสัมประสิทธิ์การจำแนกกลุ่มที่เรา Derive ได้มาตั้งแต่แรก

ถ้า $\Sigma_1 \neq \Sigma_2$ จะทำให้สมการการจำแนกกลุ่ม ไม่ได้เป็นสมการในรูปแบบสมการเชิงเส้น แต่จะเป็นในรูปของสมการกำลังสอง เพราะว่า

$$\ln \frac{f_1(x)}{f_2(x)} = -\frac{1}{2} x'(\Sigma_1^{-1} - \Sigma_2^{-1})X + X'(\Sigma_1^{-1} \mu_1 - \Sigma_2^{-1} \mu_2) + cons\tan t$$

ดังนั้นจากที่ผ่านมาในข้างต้น เราจะสามารถสรุปได้ว่าสมการจำแนกกลุ่มที่ได้ จะมีความสามารถสูงสุด (Optimal) ก็ต่อเมื่อความผันแปรของตัวแปรอิสระ ซึ่งก็คือเวกเตอร์ X ในกลุ่มที่ 1 และกลุ่มที่ 2 เท่ากัน หรือใกล้เคียงกัน โดยที่เราต้องสมมติว่าตัวแปรอิสระทุกตัวในแต่ละกลุ่มจะต้องมีการแจกแจงแบบปกติด้วย (Normal distribution)

2.1.6 ลักษณะของการจำแนกกลุ่มที่ได้ถูกต้อง

ลักษณะของการวิเคราะห์ที่ได้จะแสดงออกมาในรูปของเมตริกซ์ที่มีลักษณะดังนี้

		ลักษณะจากการจำแนก (Classification characteristic) กลุ่มที่1	กลุ่มที่2	
ลักษณะที่แท้จริง (Actual Characteristic)	กลุ่มที่1	n_{11}	n_{12}	n_1
	กลุ่มที่ 2	n_{21}	n_{22}	n_2
		n_1	n_2	n

กำหนดให้จำนวน n_{ij} คือ ตัวอย่างที่เป็นสมาชิกของกลุ่มที่ i แต่จำแนกลงกลุ่มที่ j ฉะนั้นสัดส่วนของการจำแนกกลุ่มได้ถูกต้องจึงเท่ากับ $(n_{11} + n_{22}) / n$ ซึ่งในการพิจารณาถึงประสิทธิภาพของการวิเคราะห์จำแนกกลุ่มนี้จะเน้นพิจารณาถึงสัดส่วนของการจำแนกกลุ่มได้ถูกต้องเป็นหลัก กล่าวคือ ถ้าสมการ

จำแนกกลุ่มที่ได้สามารถจำแนกกลุ่มตัวอย่างได้ถูกต้องในระดับสัดส่วนที่สูง ก็ถือว่าสมการนี้มีประสิทธิภาพที่ดี

ในการพิจารณาสัดส่วนของการจำแนกกลุ่มได้ถูกต้องโดยรวมนั้น เราจำเป็นต้องคำนึงถึงค่าสัดส่วนของกลุ่มตัวอย่างในแต่ละกลุ่มสมาชิกด้วย โดยในที่นี้สมมติให้ประชากรประกอบด้วยกลุ่มตัวอย่าง 2 กลุ่ม คือ กลุ่มที่1 และกลุ่มที่ 2 ในกลุ่มที่ 1 มีจำนวนหน่วยตัวอย่างเป็นสัดส่วน p และในกลุ่มที่ 2 มีสัดส่วนเป็น 1-p ถ้าหากตัวแปรอิสระไม่มีอำนาจในการจำแนกกลุ่มเลย ก็คาดได้ว่าจะสามารถจำแนกหน่วยตัวอย่างลงกลุ่มได้ถูกต้องในสัดส่วน p ถ้าเราจำแนกตัวอย่างลงในกลุ่มที่ 1 แต่ถ้า $p < \frac{1}{2}$ เราจะจำแนกหน่วยตัวอย่างลงในกลุ่มที่ 2

ถ้าหากว่าเราสามารถทราบได้ล่วงหน้าว่าสัดส่วนของตัวอย่างเป็นสมาชิกในกลุ่มที่ 1 เท่ากับ p แล้ว ตัวแบบจำลองของการจำแนกกลุ่มได้ถูกต้องที่เหมาะสมควรจะเป็นดังนี้

กำหนดให้ p เป็นสัดส่วนของสมาชิกในกลุ่มที่ 1 ของประชากร

α เป็นสัดส่วนในการจำแนกตัวอย่างลงในกลุ่มที่ 1

ดังนั้น โอกาสในการจำแนกตัวอย่างได้อย่างถูกต้องทั้งหมดนั้น ก็คือ ความน่าจะเป็นในการจำแนกกลุ่มได้ถูกต้อง (P(Correct)) เท่ากับ ความน่าจะเป็นที่มีเงื่อนไขของการจำแนกกลุ่มได้ถูกเมื่อกำหนดว่าถูกจัดให้อยู่ในกลุ่ม I (P(Correct/Classified type I)) x ความน่าจะเป็นของการถูกจัดให้อยู่ในกลุ่ม I (P(Classifies type I)) + ความน่าจะเป็นที่มีเงื่อนไขของการจำแนกกลุ่มได้ถูกเมื่อกำหนดว่าถูกจัดให้อยู่ในกลุ่ม II (P(Correct/Classified type II)) x ความน่าจะเป็นของการถูกจัดให้อยู่ในกลุ่ม II (P(Classifies type II))

ความน่าจะเป็นในการจำแนกกลุ่มได้ถูกต้อง (P(Correct)) = $p\alpha + (1-p)(1-\alpha)$

ฉะนั้นในกรณีที่มีการกำหนดความน่าจะเป็นโดยหลักเกณฑ์ หรือความน่าจะเป็นก่อน (Prior Probability) ในการจำแนกกลุ่มแล้วนั้น เกณฑ์ที่ใช้ในการจำแนกกลุ่มที่เหมาะสม ก็คือ

$$C_{pro} = \alpha^2 + (1-\alpha)^2$$

ในกรณีที่ P = 0.50

P(Correct) = 0.50 ไม่ว่า α จะมีค่าเป็นเท่าไรก็ตาม

ดังนั้นหากวัตถุประสงค์ในการจำแนกกลุ่ม คือ ต้องการให้การจำแนกกลุ่มจำแนกกลุ่มให้มีสัดส่วนของการจำแนกกลุ่มได้ถูกต้องสูงสุด เกณฑ์ที่ใช้ในการจำแนกกลุ่ม ก็คือ

$$C_{max} = \text{Max}(\alpha, 1-\alpha)$$

ซึ่งในการศึกษาครั้งนี้จะได้กำหนดค่าของ Prior Probability เท่ากับ 0.50

2.1.7 การพิจารณาผลการวิเคราะห์จำแนกกลุ่ม

การพิจารณาผลของการจำแนกกลุ่มดังที่ได้กล่าวไว้ในข้างต้นว่า เราควรต้องคำนึงว่ามีตัวแปรอิสระใดบ้างที่สามารถจำแนกความแตกต่างระหว่างกลุ่มได้ดี โดยที่ตัวแปรอิสระเหล่านี้มีความสามารถในการจำแนกกลุ่มได้ดีเพียงใด ซึ่งเราพิจารณาได้จากที่มาของสมการจำแนกกลุ่ม ซึ่งการพิจารณาสมการจำแนกกลุ่ม กระทำได้โดยที่

กำหนดให้ $P(I)$ = ความน่าจะเป็นโดยหลักเกณฑ์หรือความน่าจะเป็นก่อน (Prior - Probability) ที่ตัวอย่างหนึ่งจะเป็นสมาชิกของกลุ่มที่ 1

$P(I/X_i)$ = ความน่าจะเป็นที่มีเงื่อนไข (Condition (Posterior) Probability) ที่ตัวอย่างหนึ่งจะเป็นสมาชิกของกลุ่มที่ 1 โดยกำหนดว่า เราทราบค่าของตัวแปรอิสระนั้นๆแล้ว (X_i)

(X_i/I) = Likelihood ที่ตัวอย่างจะมีค่าของตัวแปรอิสระเป็น X_i โดยกำหนดว่าตัวอย่างนั้นเป็นสมาชิกของกลุ่มที่ 1

$P(II)$ = ความน่าจะเป็นโดยหลักเกณฑ์หรือความน่าจะเป็นก่อน Prior Probability ที่ตัวอย่างหนึ่งจะเป็นสมาชิกของกลุ่มที่ 2

$P(II/X_i)$ = ความน่าจะเป็นที่มีเงื่อนไข (Condition (Posterior) Probability) ที่ตัวอย่างหนึ่งจะเป็นสมาชิกของกลุ่มที่ 2 โดยกำหนดว่า เราทราบค่าของตัวแปรอิสระนั้นๆแล้ว (X_i)

$l(X_i/II)$ = Likelihood ที่ตัวอย่างจะมีค่าของตัวแปรอิสระเป็น X_i โดยกำหนดว่าตัวอย่างนั้นเป็นสมาชิกของกลุ่มที่ 2

จากทฤษฎีบทของเบส์ (Bayes' Theorem) จะได้ว่า

$$\frac{P(I/X_1)}{P(II/X_1)} = \frac{\ell(X_1/I)}{\ell(X_1/II)} * \frac{P(I)}{P(II)}$$

หรือ Posterior odds = Likelihood Ratio (L.R.) * Prior odds

เราอาจจะใช้ในรูปของ Log ได้ คือ

Log(Posterior odds) = Log(Likelihood Ratio (L.R.)) * Log(Prior odds)

หากว่าเราสมมติให้ X_i ในแต่ละกลุ่มมีการแจกแจงแบบปกติแล้ว และมี Covariance เท่ากัน เราจะได้ว่า

$$\text{Log(Likelihood Ratio (L.R.))} = b_0+b_1x_1+b_2x_2+ \ldots +b_px_p$$

ซึ่งก็คือ สมการจำแนกกลุ่มนั่นเอง โดยที่ถ้าขนาดจำนวนตัวอย่างของทั้ง 2 กลุ่มเท่ากันแล้ว Prior odds ก็จะมีค่าเป็น 1 จะทำให้ Posterior odds มีค่าเท่ากับ Likelihood Ratio แต่ถ้า Prior odds มีค่าไม่เท่ากับ 1 แล้ว

พลตำรวจ สถาบันวิทยบริการ
จุฬาลงกรณ์มหาวิทยาลัย

Log(Posterior odds) = $b_0+b_1x_1+b_2x_2+ \ldots +b_px_p$+ Log(Prior odds)

แต่ว่า พจน์หลังมีค่าเป็นค่าคงที่ ฉะนั้น

Log(Posterior odds) = $b_0+b_1x_1+b_2x_2+ \ldots +b_px_p$

2.1.8 การพิจารณาตัวแปรอิสระ

จากการได้มาซึ่งสมการจำแนกกลุ่มข้างต้น ทำให้เห็นได้ว่า ขนาดและเครื่องหมายของสัมประสิทธิ์ของสมการจำแนกกลุ่มนี้ จะเป็นตัวกำหนดผลของตัวแปรอิสระแต่ละตัว และขนาดของค่าสัมประสิทธิ์แต่ละตัวนั้นจะถูกกระทบโดยค่าตัวเลข (scale) ของตัวแปรอิสระที่ใช้ ดังนั้นหากมีการหาร (Normalized) ตัวแปรอิสระแต่ละตัวด้วยค่าส่วนเบี่ยงเบนมาตรฐานของตัวแปรอิสระแต่ละตัวแล้ว ก็จะสามารถขจัดอิทธิพลของการใช้จำนวนขนาดที่แตกต่างกันของตัวแปรอิสระต่างๆได้ ซึ่งการคำนวณหาค่าสัมประสิทธิ์ที่ปรับให้อยู่ในรูปมาตรฐาน (Standardize; b_Δ^*)ทำได้โดย

$$b_\Delta X_\Delta = b_\Delta^* X_\Delta^* = b_\Delta^* X_\Delta/\sigma_\Delta$$

เพราะฉะนั้น $b_\Delta^* = b_\Delta\sigma_\Delta$

ถ้า $|b_\Delta^*| > |X_k^*|$ เราก็สรุปได้ว่า X_Δ มีความสามารถในการจำแนกความแตกต่างระหว่าง 2 กลุ่มได้ดีกว่า X_k เพราะว่า หาก X_Δ^* เปลี่ยนแปลงไป 1 หน่วย จะทำให้ตัวแปรตามเปลี่ยนแปลงไปได้มากกว่าที่ X_k^* เปลี่ยนแปลงไป 1 หน่วย เนื่องจากว่าการวิเคราะห์จำแนกกลุ่มนี้อยู่บนพื้นฐานของ Statistical Distance ที่มีหน่วยวัดเป็นส่วนเบี่ยงเบนมาตรฐาน

ส่วนเครื่องหมายของค่าสัมประสิทธิ์ b_Δ^* ก็เป็นตัวกำหนดทิศทางของผลที่เกิดขึ้นของ X_Δ กล่าวคือ ถ้า b_Δ^* เป็นเครื่องหมายบวกแล้ว หากว่า X_Δ มีค่าเพิ่มขึ้นก็ทำให้ตัวแปรตามมีค่าเพิ่มขึ้นและถ้ายิ่งตัวแปรตามมีค่าเพิ่มขึ้นมากเพียงไร ก็ยิ่งทำให้โอกาสที่ตัวอย่างนั้นๆจะเป็นสมาชิกของกลุ่มที่ 1 ได้มากขึ้น ซึ่งในการประมาณค่าสัมประสิทธิ์นั้น เราต้องการที่จะประมาณค่าสัมประสิทธิ์ให้ดีที่สุดเพื่อการจำแนกกลุ่มที่ไประสิทธิภาพ ซึ่งตามหลักการทางสถิติแล้วนั้นถ้าจำนวนตัวอย่างมีมากเพียงใดผลของการประมาณก็จะยิ่งดีมากขึ้นเท่านั้น โดยสมมติว่า กลุ่มที่ 1 มีจำนวนหน่วยตัวอย่างเท่ากับ 900 ตัวอย่าง กลุ่มที่ 2 มีจำนวนตัวอย่างเท่ากับ 100 ตัวอย่าง หากว่าในการสร้างสมการจำแนกกลุ่มเราได้ใช้ตัวอย่างจากกลุ่มที่ 2 ทั้ง 100 ตัวอย่าง แล้วเลือกจำนวนตัวอย่างจากกลุ่มที่ 1 ให้มีจำนวนเท่ากับในกลุ่มที่ 2 คือ 100 ตัวอย่าง ค่าความน่าจะเป็นโดยหลักเกณฑ์ (Prior Probability) ที่หน่วยตัวอย่างจะเป็นสมาชิกของกลุ่มที่ 1 เท่ากับ 0.5 แต่หากว่าเราใช้จำนวนตัวอย่างทั้งหมดในการสร้างสมการจำแนกกลุ่ม (กลุ่มที่ 1 จำนวน 900 ตัวอย่าง และกลุ่มที่ 2 จำนวน 100 ตัวอย่าง) ค่าค่าความน่าจะเป็นโดยหลักเกณฑ์ (Prior Probability) ก็จะเท่ากับ 0.9 ซึ่งการเปลี่ยนแปลงเช่นนี้ไม่ส่งผลกระทบต่อ b_Δ แต่จะ

กระทบเฉพาะ b_0 เท่านั้น ดังนั้นหากว่าต้องการทราบว่าตัวแปรอิสระตัวใดที่สามารถจำแนกความแตกต่างระหว่างกลุ่มทั้งสองออกจากกันได้ดีเราก็ควรจะใช้ข้อมูลทุกหน่วยตัวอย่าง ซึ่งหมายความว่าเราจะใช้ตัวอย่างทุกตัวอย่างไม่ใช่ใช้ตัวแปรทุกตัว เพราะว่าในบางครั้งค่าสัมประสิทธิ์ที่ประมาณได้จะไม่มีความคงที่ (Stable) เนื่องจากว่าความสัมพันธ์ระหว่างตัวแปรบางคู่มีความสัมพันธ์กันที่สูง

2.1.9 การประมาณค่าสัมประสิทธิ์ของสมการจำแนกกลุ่ม

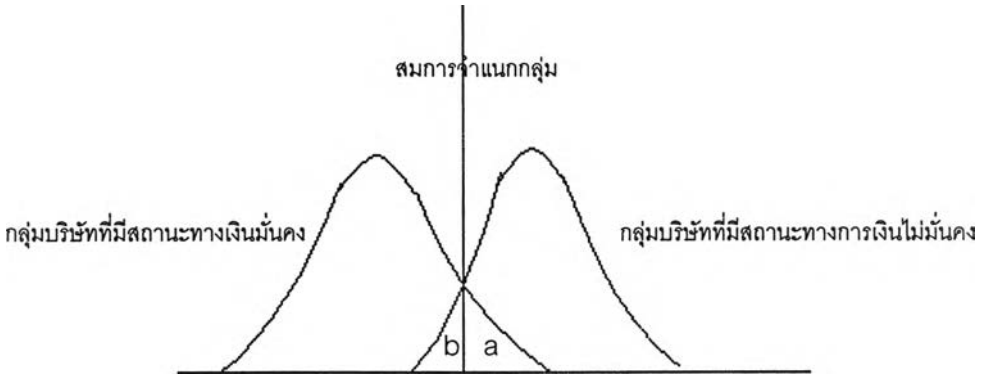
ภายหลังจากการศึกษาความสัมพันธ์ระหว่างตัวแปรอิสระแล้ว จะมีการคัดเลือกตัวแปรอิสระเหล่านี้เพื่อมาสร้างเป็นสมการที่เรียกว่า สมการจำแนกกลุ่ม ซึ่งวิธีการนี้จะหาสมการจำแนกกลุ่มที่ดีที่สุด โดยลดความผิดพลาดในการจำแนกกลุ่มให้น้อยที่สุดจากข้อมูลที่มีอยู่ไปยังกลุ่มย่อยต่างๆ รวมทั้งแยกตัวแปรอิสระที่สามารถใช้จำแนกข้อมูลได้ดีออกจากตัวแปรอิสระที่มีความสามารถที่น้อยกว่าได้ ซึ่งรูปแบบความสัมพันธ์ระหว่างตัวแปรตามกับตัวแปรอิสระจะอยู่ในรูปแบบเชิงเส้น ดังนี้

$$Z = W_1X_1 + W_2X_2 + ... + W_nX_n + e$$

โดยที่

Z	=	ตัวแปรตามคะแนนจำแนกกลุ่ม (Discriminant Score) เป็นคะแนนที่ได้จากสมการจำแนกกลุ่ม เพื่อใช้จัดแบ่งกลุ่มระหว่างบริษัทที่มีสถานะทางการเงินมั่นคง กับบริษัทที่มีสถานะการเงินไม่มั่นคง
X_i	=	ตัวแปรอิสระ หรือตัวแปรจำแนกกลุ่ม คือ อัตราส่วนทางการเงินต่างๆ ที่ใช้แยกกลุ่มบริษัทที่มีฐานะการเงินมั่นคงและไม่มั่นคงโดยที่ i = 1,2, 3,....,n
W_i	=	ค่าสัมประสิทธิ์แสดงความสำคัญของตัวแปรอิสระ โดยที่ i = 1, 2, 3,...,n หรือน้ำหนักของอัตราส่วนทางการเงินที่มีผลต่อการจำแนกกลุ่ม
e	=	ค่าความคลาดเคลื่อนที่เกิดขึ้น

รูปที่ 2 การจำแนกกลุ่มโดยสมการจำแนกกลุ่มตามความมั่นคงในสถานะทางการเงิน



สมการแสดงความสัมพันธ์ หมายถึง สมการจำแนกกลุ่มหรือเส้นแบ่งกลุ่มประชากรออกเป็น 2 กลุ่มคือ กลุ่มที่มีความมั่นคงและกลุ่มที่มีความไม่มั่นคง ซึ่งในการวิจัยนี้ จะทำการแบ่งกลุ่มบริษัทประกันวินาศภัยออกเป็น 2 กลุ่ม คือ กลุ่มที่มีสถานะการเงินมั่นคงและไม่มั่นคง โดยลักษณะของการจำแนกกลุ่มที่เกิดขึ้นมีลักษณะการกระจายตัวที่สามารถแยกความแตกต่างออกเป็น 2 กลุ่มได้อย่างชัดเจน ดังแสดงในรูปที่ 2

จากรูปที่ 2 จะสามารถแบ่งกลุ่มบริษัทประกันวินาศภัยออกเป็น 2 กลุ่มตามสถานะความมั่นคงทางการเงินได้อย่างชัดเจน โดยที่กลุ่มที่1 เป็นกลุ่มบริษัทที่มีสถานะการเงินมั่นคง โดยจะอยู่ด้านขวาของเส้นจะถือเป็นกลุ่มบริษัทที่มีสถานะการเงินมั่นคง และกลุ่มที่2 เป็นกลุ่มที่มีสถานะการเงินไม่มั่นคงและจะอยู่ทางด้านซ้ายของเส้นแบ่งกลุ่ม ซึ่งพบว่ามีบริษัทประกันวินาศภัยที่มีสถานะการเงินมั่นคงถูกจัดให้เป็นกลุ่มบริษัทที่มีสถานะการเงินไม่มั่นคงแสดงด้วยพื้นที่ส่วน a โดยพื้นที่ส่วนนี้เป็นสัดส่วนการจัดกลุ่มของบริษัทประกันวินาศภัยที่มีสถานะการเงินมั่นคงที่ถูกจัดว่ามีสถานะการเงินไม่มั่นคงหรือสัดส่วนของการจัดกลุ่มผิด ในขณะเดียวกันส่วนพื้นที่ b หมายถึง สัดส่วนของบริษัทที่มีสถานะการเงินไม่มั่นคงแต่ถูกจัดเป็นกลุ่มบริษัทที่มีสถานะการเงินมั่นคงด้วยสมการจำแนกกลุ่ม จึงเป็นสัดส่วนการจัดกลุ่มผิดเช่นเดียวกัน ซึ่งหลักเกณฑ์ของการวิเคราะห์จำแนกกลุ่มต้องการลดความผิดพลาดของเปอร์เซ็นต์การจัดกลุ่มผิดให้น้อยที่สุด จึงสรุปได้ว่าวิธีการนี้จะหาความสัมพันธ์ระหว่างตัวแปรอิสระที่ใช้จำแนกกลุ่มตัวอย่างออกจากกันแล้วนำเอามาทำการประมาณค่าของหน่วยวิเคราะห์นั้นว่าควรจะเป็นสมาชิกของกลุ่มใดในลักษณะที่คล้ายคลึงกับการวิเคราะห์ถดถอย ซึ่งค่าที่ได้จากการประมาณที่ใกล้เคียงกับค่าเฉลี่ยกลางของกลุ่ม (Group Centroid) ใดมากที่สุด หน่วยวิเคราะห์นั้นก็จะเป็นสมาชิกของกลุ่มนั้น ซึ่งจากรูปแบบของสมการจำแนกกลุ่มในข้างต้น ค่าของ W_i ที่ประมาณได้จะ

ต้องเป็นไปได้ในลักษณะที่ทำให้ค่าความเป็นสมาชิกของแต่ละกลุ่มแตกต่างกันมากที่สุด กล่าวคือ คะแนนจำแนกกลุ่มที่ได้จากสมการควรจะเป็นไปในลักษณะที่ทำให้อัตราส่วนของผลรวมกำลังสองระหว่างกลุ่มต่อผลรวมกำลังสองภายในกลุ่มที่ค่าสูงที่สุดด้วย ดังนั้นการหาน้ำหนักของตัวแปร (W_i) จะสามารถได้จากสมการเมตริกซ์ $(A^{-1}B - \lambda I)W = 0$

โดยที่ A^{-1} คือ Inverse ของ within-groups sums of squares and cross product matrix

B คือ between – groups sums of squares and cross product matrix

λ คือ eigen value

I คือ identity matrix

การหาค่า within – groups SSCP matrix (A) สมาชิกแต่ละตัวของเมตริกซ์ A จะหาได้จาก

$$A_{\alpha\alpha} = \sum_{g=1}^{k}\sum_{i=1}^{n_g} X_{\alpha gi}^2 - \sum_{g=1}^{k}\left[(\sum_{i=1}^{n_g} X_{\alpha gi})^2 / n_g\right]$$

$$A_{\alpha\beta} = \sum_{g=1}^{k}\sum_{i=1}^{n_g} X_{\alpha gi}X_{\beta gi} - \sum_{g=1}^{k}\left[(\sum_{i=1}^{n_g} X_{\alpha gi})(\sum_{i=1}^{n_g} X_{\beta gi}) / n_g\right]$$

โดยที่ $A_{\alpha\alpha}$ คือ สมาชิกของ within – groups SSCP matrix แถวที่ α สดมภ์ที่ α

$A_{\alpha\beta}$ คือ สมาชิกของ within – groups SSCP matrix แถวที่ α สดมภ์ที่ β

การหาค่า Between – groups SSCP matrix (B) สมาชิกแต่ละตัวของเมตริกซ์ B จะหาได้จาก

$$b_{\alpha\alpha} = \sum_{g=1}^{k}\left[(\sum_{i=1}^{n_g} X_{\alpha gi})^2 / n_g\right] - \left[(\sum_{g=1}^{k}\sum_{i=1}^{n_g} X_{\alpha gi})^2 / N\right]$$

$$b_{\alpha\beta} = \sum_{g=1}^{k}\left[(\sum_{i=1}^{n_g} X_{\alpha gi})(\sum_{i=1}^{n_g} X_{\beta gi}) / n_g\right] - \left[(\sum_{g=1}^{k}\sum_{i=1}^{n_g} X_{\alpha gi})(\sum_{g=1}^{k}\sum_{i=1}^{n_g} X_{\beta gi}) / N\right]$$

โดยที่ $b_{\alpha\alpha}$ คือ สมาชิกของ between – groups SSCP matrix แถวที่ α สดมภ์ที่ α

$b_{\alpha\beta}$ คือ สมาชิกของ between – groups SSCP matrix แถวที่ α สดมภ์ที่ β

$X_{\alpha gi}, X_{\beta gi}$ คือ คะแนนในตำแหน่งที่ i ของกลุ่มที่ g ของตัวแปรในกลุ่ม

n_g คือ ขนาดของกลุ่ม

k คือ จำนวนกลุ่ม

N คือ ผลรวมทั้งหมดของกลุ่มตัวอย่าง (Total Sample Size) $= \sum_{g=1}^{k} n_g$

การหาค่าไอเค่น (Eigen Value)

การหาค่า Eigen Value สามารถกระทำได้โดย

1. นำ λ ซึ่งเป็น Unknown scalar ไปลบออกจาก diagonal element matrix $A^{-1} B$
2. นำเมตริกซ์ที่ได้จากข้อที่ 1 มาหา determinant แล้วให้เท่ากับศูนย์ (0)
3. แก้สมการหาค่า λ

การทดสอบนัยสำคัญของสมการการจำแนกกลุ่ม จะใช้ Barlett's V ซึ่งมีการกระจายโดยประมาณแบบไคกำลังสอง (chi-square) ที่ df = p+k-2r โดยคำนวณได้จากสูตร

$$V_i = [N-1-(p+k)/2] \ln(1+\lambda_i)$$

โดยที่ V_i คือ ค่าที่ใช้เปรียบเทียบในการทดสอบความมีนัยสำคัญ

N คือ ขนาดของกลุ่มตัวอย่าง

p คือ จำนวนตัวแปร

k คือ จำนวนกลุ่ม

df คือ องศาความเป็นอิสระ (degree of freedom)

r คือ จำนวนค่าของ Eigen value

2.2 วิธีการสร้างสมการการวิเคราะห์จำแนกกลุ่ม

ขั้นตอนการสร้างสมการจำแนกกลุ่มนั้นมี 2 วิธีการ ก็คือ

2.2.1 **วิธีทางตรง** (Direct Method) เป็นวิธีการวิเคราะห์โดยใช้ตัวแปรอิสระทุกตัวเข้ามาวิเคราะห์เพื่อสร้างเป็นสมการจำแนกกลุ่ม โดยไม่คำนึงว่าตัวแปรอิสระตัวใดที่มีความแตกต่างระหว่างกลุ่มอย่างมีนัยสำคัญ ซึ่งวิธีการนี้ผู้วิจัยเห็นว่าจะทำให้ไม่สามารถทราบได้ว่าระหว่างกลุ่มบริษัทประกันวินาศภัยที่มีความมั่งคงทางการเงินและกลุ่มที่ไม่มีความมั่นคงนั้น มีความแตกต่างกันในอัตราส่วนทางการเงินใดบ้าง

2.2.2 **วิธีแบบขั้นตอน** (Stepwise Method) เป็นวิธีการคัดเลือกตัวแปรเพื่อสร้างสมการจำแนกกลุ่มโดยการสร้างสมการจำแนกกลุ่มแบบขั้นตอนนี้จะเป็นวิธีการที่คัดเลือกตัวแปรทีละตัวมาเข้าสมการโดยจะหาตัวแปรที่ดีที่สุดในการจำแนกกลุ่มมาเข้าในสมการเป็นตัวแรก จากนั้นก็จะหาตัวแปรที่ดีที่สุดตัวต่อไปมาเข้าในสมการเพื่อการปรับปรุงแก้ไขให้สมการมีความสามารถในการจำแนกกลุ่มได้ดีขึ้นต่อๆไป ในขั้นตอนต่อๆมาก็จะเป็นการนำตัวแปรที่ดีที่สุดแต่ละตัวที่เหลือเข้ามาในสมการ ซึ่งในแต่ละขั้นตอนตัวแปรที่ได้รับการเลือกมาก่อนอาจจะถูกตัดทิ้งออกไปก็ได้ หากพบว่าเมื่อนำมารวม

กับตัวแปรอื่นๆแล้วไม่สามารถช่วยให้สมการสามารถจำแนกกลุ่มได้ดีขึ้น ซึ่งการคัดเลือกตัวแปรเข้ามาทีละขั้นตอนนั้นจะมีความคล้ายคลึงกับการวิเคราะห์ถดถอยพหุแบบขั้นตอน

วิธีการสร้างสมการจำแนกกลุ่มแบบขั้นตอนนี้ยังมีหลักเกณฑ์ในการคัดเลือกตัวแปรเข้ามาในสมการจำแนกกลุ่ม อีก 5 หลักเกณฑ์การคัดเลือก คือ

1. วิธี Wilks' lambda
2. วิธี Mahalanobis distance
3. วิธี Unexplained variance
4. วิธี Smallest F ratio
5. วิธี Rao's V

ผู้วิจัยจะเลือกใช้วิธีการของ Wilk ซึ่ง วิธีการของ Wilk เป็นวิธีการพิจารณาอัตราส่วน F ของตัวแปรอิสระหลายตัวทั้งหมด (Overall multivariate F ratio) ในการทดสอบความแตกต่างระหว่างกลุ่ม ตัวแปรอิสระที่ให้ค่า F สูงสุดจะลดค่า Wilks' lambda ให้ต่ำลง เนื่องจากค่า Wilks' lambda เป็นมาตรวัดความแตกต่างระหว่างกลุ่ม โดยวิธีการนี้จะคำนึงถึงความแตกต่างระหว่างกลุ่มและภายในกลุ่มด้วย ค่าของ Wilks' lambda ที่ต่ำ แสดงได้ว่า สมการที่ได้สามารถจำแนกกลุ่มได้ดี ซึ่งไม่ว่าจะใช้วิธีการคัดเลือกตัวแปรอิสระเข้าสมการด้วยวิธีใดก็ตาม ตัวแปรอิสระที่ได้รับการคัดเลือกก็ต่อเมื่อค่าของอัตราส่วน F เชิงส่วน (Partial F ratio) มีค่ามากกว่าที่กำหนดไว้ ซึ่งค่า F เชิงส่วนนี้ใช้วัดความสามารถในการแบ่งแยกที่เกิดจากการนำเอาตัวแปรอิสระตัวใดตัวหนึ่งเข้ามาเมื่อได้พิจารณาถึงความสามารถในการแบ่งแยกที่เกิดจากการนำเอาตัวแปรอิสระตัวอื่นๆที่ได้รับการคัดเลือกแล้ว สำหรับการทดสอบด้วยค่า F เชิงส่วนนี้จะได้กระทำก่อนที่จะนำตัวแปรอิสระเข้าสู่ขั้นตอนของการวิเคราะห์สมการจำแนกกลุ่ม และนอกจากนี้มีงานวิจัยที่ได้มีการศึกษาปัญหาด้วยวิธีวิเคราะห์จำแนกกลุ่มด้วยวิธีการคัดเลือกตัวแปรทั้ง 5 วิธีที่ได้กล่าวถึงในข้างต้น อันได้ผลสรุปของการศึกษาว่า จำนวนตัวแปรอิสระที่เข้ามาในสมการจำแนกกลุ่มเป็นตัวแปรเดียวกันทั้งสิ้นและค่าสัมประสิทธิ์ที่ได้ตลอดจนความสามารถในการจำแนกกลุ่มได้ถูกต้องล้วนมีความใกล้เคียงกันมากโดยอาจกล่าวได้ว่าได้ผลสรุปของการวิเคราะห์ที่เหมือนกันก็ว่าได้[8]

ในการคัดเลือกตัวแปรอิสระเข้ามาในสมการมีหลักเกณฑ์อีกว่าตัวแปรอิสระนั้นจะต้องไม่มีความสัมพันธ์กันมากในเชิงเส้น มิฉะนั้นจะไม่สามารถแก้สมการได้ เช่นเดียวกับการวิเคราะห์ถด

---

[8] พีรเทพ รุ่งชีวิน และคณะ. ปัจจัยที่มีอิทธิพลต่อการตัดสินใจไปทำงานในภาคตะวันออกเฉียงเหนือ. การประชุมทางวิชาการสถิติประยุกต์ ครั้งที่3, 12-13 พฤษภาคม 2526. คณะสถิติประยุกต์ สถาบันพัฒนบริหารศาสตร์.

ถอย ด้วยเหตุนี้จึงจะต้องมีการกำหนดว่าตัวแปรอิสระควรจะมีความสัมพันธ์กันเพียงใดจึงจะสามารถยอมรับได้ ซึ่งเท่ากับว่าเป็นการกำหนดระดับที่ยอมรับได้นั่นเอง (Level of tolerance)

ระดับที่ยอมรับได้ คือ มาตรวัดอัตราของความสัมพันธ์เชิงเส้นระหว่างตัวแปรอิสระโดยกำหนดให้ตัวแปรอิสระแต่ละตัวผลัดกันเข้ามาเป็นตัวแปรตามแล้วพิจารณาความสัมพันธ์ดัวตัวแปรอิสระตัวอื่นๆกับตัวแปรอิสระตัวนั้นๆ ซึ่งระดับที่ยอมรับได้ที่ต่ำ หมายความได้ว่า ตัวแปรอิสระตัวนั้นเกือบไม่มีความสัมพันธ์เชิงเส้นกับตัวแปรอิสระตัวอื่นๆ โดยที่ตัวแปรอิสระใดที่มีค่าระดับที่ยอมรับได้ที่ต่ำกว่าที่ระบุไว้ ตัวแปรอิสระตัวนั้นก็จะไม่ถูกนำเข้ามา นอกจากนั้นถ้าหากว่าตัวแปรอิสระตัวใดจะทำให้ระดับที่ยอมรับได้ลดต่ำลงไปก็จะไม่ถูกนำเข้ามาในการพิจารณา ซึ่งในที่นี้ผู้วิจัยกำหนดค่าระดับที่ยอมรับได้เท่ากับ 0.01

2.2.3 การคำนวณค่าของคะแนนจำแนกกลุ่ม

เมื่อได้สมการจำแนกกลุ่มแล้วต่อมา คือการหาคะแนนของหน่วยวิเคราะห์แต่ละหน่วย โดยการนำเอาค่าสัมประสิทธิ์ของตัวแปรมาคูณกับค่าของตัวแปรอิสระ (อัตราส่วนทางการเงิน) ในแต่ละหน่วยตัวอย่างแล้ว ทำการรวมผลคูณทั้งหมดก็จะได้เป็นคะแนนของหน่วยวิเคราะห์นั้นๆ ที่เรียกว่า Z-score ซึ่งคะแนนดังกล่าวนี้เป็นคะแนนที่จะใช้ชี้และเป็นหลักเกณฑ์ที่ใช้ตัดสินว่าหน่วยวิเคราะห์นั้นๆ ควรจะเป็นสมาชิกในกลุ่มใด โดยที่การคำนวณหาค่า Z-score ได้ดังนี้

$$Z_{score} = \frac{N_A Z_A + N_B Z_B}{N_A + N_B}$$

โดยที่

Z_{score} = คะแนนในการจำแนกกลุ่ม (Z-score)

N_A = จำนวนบริษัทในกลุ่มที่มีฐานะการเงินไม่มั่นคง

N_B = จำนวนบริษัทในกลุ่มที่มีฐานะการเงินมั่นคง

Z_A = ค่าคะแนนเฉลี่ยของกลุ่ม (Group Centroid) บริษัทที่มีฐานะการเงินไม่มั่นคง

Z_B = ค่าคะแนนเฉลี่ยของกลุ่ม (Group Centroid) บริษัทที่มีฐานะการเงินมั่นคง

2.2.4 ค่าสถิติที่ใช้ในการประเมินสมการจำแนกกลุ่ม

ค่าทางสถิติที่สำคัญที่ใช้ในการพิจารณาสมการจำแนกกลุ่ม มีอยู่ 4 ค่า คือ

2.2.4.1 สัดส่วนของการจำแนกกลุ่มได้ถูกต้องโดยรวม เป็นสิ่งที่สำคัญอย่างยิ่งในการใช้พิจารณาถึงประสิทธิภาพและความสามารถของสมการตัวแบบในการจำแนกกลุ่มตัวอย่าง โดย

ที่เราจะพิจารณาในส่วนเฉพาะที่ตัวแบบสมการจำแนกกลุ่มสามารถจำแนกหน่วยตัวอย่างได้ถูกต้องตามลักษณะที่แท้จริงของตัวอย่างนั้น ซึ่งวิธีการพิจารณาได้แสดงไว้ในข้างต้น หัวข้อ 2.1.6 ลักษณะของการจำแนกกลุ่มที่ได้ถูกต้อง

2.2.4.2 ค่าEigenvalue คือ อัตราส่วนการผันแปรระหว่างกลุ่มต่อการผันแปรภายในกลุ่ม โดยที่ถ้าค่า Eigenvalue มีค่าสูง ก็หมายความว่า สมการนั้นยิ่งมีความเชื่อถือมากขึ้น สำหรับวิธีการหาค่า Eigenvalue ได้แสดงไว้ในข้างต้นแล้ว หัวข้อ 2.1.9 การประมาณค่าสัมประสิทธิ์ของสมการจำแนกกลุ่ม

2.2.4.3 ค่าความสัมพันธ์แคนนอนนิคัล (Canonical correlation) เป็นมาตรวัดสัมพันธ์ของสมการกับกลุ่มของตัวแปร ซึ่งระบุถึงการเป็นสมาชิกของกลุ่มนั้นๆ ของตัวแปรตามที่ชี้ให้เห็นว่าการเป็นสมาชิกของกลุ่ม มีความสัมพันธ์กับสมการที่ได้มามากน้อยเพียงใด หรืออาจกล่าวได้ว่า Canonical correlation กำลังสอง หมายถึง สัดส่วนของการผันแปรของสมการที่อธิบายโดยกลุ่ม และหากเปรียบเทียบกับการวิเคราะห์ความแปรปรวนทางเดียว (One way analysis of variance) ค่าของ Canonical correlation ก็คือ ค่า Eta หรือ Correlation ratio และในกรณีที่มี 2 กลุ่ม ค่าความสัมพันธ์แคนนอนนิคัล คือ ค่าสัมประสิทธิ์เพียร์สัน (Pearson correlation coefficient) ระหว่างคะแนนจำแนกกลุ่มกับตัวแปรกลุ่ม ค่าของ Cannaonial correlation ที่ต่ำ หมายความว่า สมการนั้นไม่สามารถใช้ในการคาดคะเนความเป็นสมาชิกของกลุ่มนั้นๆได้ดี

2.2.4.4 ค่าวิลค์แลมด้า (Wilks' lambda) เป็นอัตราส่วนของค่าการผันแปรผลรวมยกกำลังสองภายในกลุ่มต่อผลรวมยกกำลังสองของทั้งหมด กล่าวคือ เป็นสัดส่วนของการผันแปรทั้งหมดที่ไม่ได้อธิบายด้วยความแตกต่างระหว่างกลุ่ม โดยที่ค่า Wilks' lambda ที่น้อย หมายความว่า สมการที่ได้มีความผันแปรระหว่างกลุ่มมากกว่าความผันแปรภายในกลุ่ม ซึ่งค่า Wilks' lambda เท่ากับ 1 เมื่อความแตกต่างระหว่างกลุ่มไม่มีความแตกต่างจากค่าความแตกต่างภายในกลุ่มนั่นเอง

2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง

ในอดีตได้มีความพยายามจากนักวิชาการหลายฝ่ายในต่างประเทศต้องการที่จะศึกษาการใช้ข้อมูลทางการเงิน ซึ่งจัดให้อยู่ในรูปของชุดอัตราส่วนทางการเงินมาเป็นเครื่องมือในการพยากรณ์ภาวะล้มเหลวทางธุรกิจ Beaver (ค.ศ.1967) ได้ศึกษาการใช้อัตราส่วนทางการเงินมาเป็นเครื่องมือในการพยากรณ์ภาวะความล้มเหลวทางธุรกิจโดยการสุ่มตัวอย่างบริษัทที่ประกอบธุรกิจประสบความล้มเหลวในช่วงระหว่างปี ค.ศ.1954 - ค.ศ.1964 จำนวน 79 บริษัทใน 38 ต่าง ประเภทธุรกิจ โดยศึกษาด้วยวิธี Paired Sample พบว่ามี 6 อัตราส่วนทางการเงินจากจำนวน 30 อัตราส่วน

ที่สามารถแสดงความแตกต่างอย่างมีนัยสำคัญในการวิเคราะห์ความแตกต่างของค่าเฉลี่ย (Mean) ระหว่างบริษัทสองกลุ่มคือบริษัทที่ล้มเหลวและไม่ล้มเหลว อัตราส่วนเหล่านั้นคือ 1.อัตราส่วนกระแสเงินสดต่อหนี้สินรวม (cash flow to total debt ratio) 2.อัตราส่วนกำไรสุทธิต่อสินทรัพย์รวม (net income to total assets ratio) 3.อัตราส่วนหนี้สินหมุนเวียนและหนี้สินระยะยาวต่อสินทรัพย์รวม (current plus long-term liabilities to total assets ratio) 4.อัตราส่วนเงินทุนหมุนเวียนต่อสินทรัพย์รวม (working capital to total assets ratio) 5.อัตราส่วนหมุนเวียนสภาพคล่อง (current ratio) และ6. อัตราส่วนระยะเวลาการจ่ายหนี้ (no-credit interval ratio) ซึ่งนอกจากนี้พบต่อไปว่าความแตกต่างของบริษัทที่ล้มเหลวและบริษัทที่มีความมั่นคงในอัตราส่วนทางการเงินเหล่านี้ที่สามารถแสดงถึงความแตกต่างของบริษัทที่มีความมั่นคงและไม่มั่นคงมีมากขึ้น เมื่อใกล้ปีที่เหตุการณ์จริงที่บริษัทประสบความล้มเหลวในการดำเนินงาน นอกจากนั้นโดยเฉลี่ยแล้วบริษัทที่มั่นคงจะมีความเบี่ยงเบนของอัตราส่วนทางการเงินในแต่ละปีน้อยมาก

Altman (ค.ศ.1968) ได้ทำการศึกษาและพัฒนาตัวแบบจำลองสำหรับการพยากรณ์ภาวะล้มละลายโดยการวิเคราะห์ตัวแปรหลายตัว (Multivariate Analysis) ประกอบเข้ากับการวิเคราะห์ข้อมูลด้วยเทคนิคการวิเคราะห์แบบจำแนกกลุ่มเปรียบเทียบกลุ่มอัตราส่วนทางการเงิน 5 กลุ่ม คือ กลุ่มสภาพคล่อง กลุ่มความสามารถในการทำกำไร กลุ่มการบริหารโครงสร้างหนี้สิน กลุ่มการบริหารการดำเนินงาน และกลุ่มความแข็งแกร่งของเงินทุน รวมทั้งสิ้น 22 อัตราส่วน โดยได้นำเอาตัวอย่างบริษัทที่ล้มเหลวจำนวน 33 บริษัทในธุรกิจว่าด้วยเงื่อนไขที่ระบุใน chapterX ภายใต้กฎหมายการล้มละลายแห่งชาติของสหรัฐอเมริกา ในระหว่างปี ค.ศ.1946 - ค.ศ.1965 หลังจากนั้นสร้างสมการจำแนกกลุ่มเพื่อใช้ในการจำแนกสถานะของกิจการ แบบจำลองดังกล่าวสรุปได้ว่าบริษัทที่มีค่า Z-score จากการคำนวณตามแบบจำลองที่มากกว่า 2.99 เป็นบริษัทที่ไม่ล้มละลาย ส่วนบริษัทที่มีค่า Z-score ต่ำกว่า 1.81 จะเป็นบริษัทที่มีโอกาสล้มละลาย ส่วนบริษัทที่มีค่า Z-score อยู่ระหว่าง 1.81 ถึง 2.99 เรียกว่าเป็น zone of ignorance คือ ไม่สามารถระบุได้ว่าเป็นบริษัทที่ล้มละลายได้อย่างแน่นอน การศึกษาของ Altman แสดงให้เห็นว่าแบบจำลองที่ได้มีความถูกต้องของการพยากรณ์เฉลี่ยถึง 95 เปอร์เซ็นต์ โดยเฉพาะอย่างยิ่งภายในระยะเวลา 2 ปีก่อนเหตุการณ์ล้มละลายจริงจะเกิดขึ้น

ภายหลังจากที่ Altman เสนอตัวแบบการพยากรณ์ภาวะล้มละลายของธุรกิจด้วยวิธีการจำแนกกลุ่ม ต่อมาในช่วงปี ค.ศ.1970 - ค.ศ.1980 ได้มีนักวิชาการอีกหลายคนได้ ทำการศึกษาต่อเนื่องจากการศึกษาของ Beaver และ Altman สำหรับการทบทวนผลการศึกษาในต่างประเทศซึ่งล้วนแต่เป็นผลงานการศึกษาในแนวทางเดียวกันกล่าวคือเป็นการใช้วิธีการวิเคราะห์จำแนกกลุ่มมาศึกษาแต่ได้ขยายขอบเขตการศึกษาออกไปในลักษณะต่างๆกัน ไม่ว่าจะเป็นการศึกษาของ Edmister

(ค.ศ.1972) โดยมีวัตถุประสงค์เพื่อทดสอบวิธีการวิเคราะห์อัตราส่วนทางการเงินในการพยากรณ์ความล้มเหลวของธุรกิจขนาดเล็กกลุ่มตัวอย่างที่ใช้ในการศึกษา คือ บริษัทที่ทำการกู้ยืมกับ Small Business Administration ในช่วงปี ค.ศ.1954 - ค.ศ.1969 ได้ผลสรุปว่าประสิทธิภาพในการวิเคราะห์อัตราส่วนทางการเงินนั้นจะขึ้นอยู่กับการคัดเลือกอัตราส่วนทางการเงินและวิธีการที่ใช้ในการวิเคราะห์ ทั้งนี้อัตราส่วนทางการเงินที่พบว่ามีความสำคัญในการพยากรณ์ คือ 1.อัตราส่วนกระแสเงินสดต่อหนี้สินหมุนเวียน (fund flow to current liabilities ratio) 2.อัตราส่วนกระแสหมุนเวียนในการดำเนินงานต่อยอดขาย (net worth to sales ratio) 3.อัตราส่วนสินทรัพย์หมุนเวียนเร็วต่อหนี้สินหมุนเวียน (quick assets to current liabilities ratio) 4.อัตราส่วนหนี้สินหมุนเวียนต่อกระแสหมุนเวียนในการดำเนินงาน (current liabilities to net worth) 5.อัตราส่วนสินค้าคงคลังต่อยอดขาย (inventory to sales) และ 6. อัตราส่วนยอดขายต่อเงินทุนในการดำเนินงาน (sales to working capital)

การศึกษาการจำแนกระดับกำไรในกลุ่มธุรกิจธนาคารในสหรัฐอเมริกาของ Haslen และ Longbrake (ค.ศ.1971) โดยการนำอัตราส่วนทางการเงินจำนวน 46 อัตราส่วน นำมาศึกษาถึงระดับความสามารถในการทำกำไรของธนาคารพาณิชย์ในสหรัฐอเมริกาที่ประกอบธุรกิจในปี ค.ศ.1963 มาศึกษาด้วยวิธีการวิเคราะห์ถดถอยแบบขั้นตอน (Stepwise Regression) เพื่อให้ได้สมการจำแนกกลุ่ม จนกระทั่งได้อัตราส่วนการเงิน 8 อัตราส่วนที่มีนัยสำคัญใช้ในการจำแนกกลุ่ม ซึ่ง สมการที่ได้นี้ให้ผลการจำแนกกลุ่มของธนาคาร ถูกต้องถึง 86.5 เปอร์เซ็นต์

Sinkey (ค.ศ.1975) ได้ทำการพยากรณ์ธนาคารที่ประสบปัญหาทางการเงินของประเทศสหรัฐอเมริกาในระหว่างปี ค.ศ.1969 – ค.ศ.1972 กลุ่มตัวอย่างที่ใช้ในการศึกษา คือ ธนาคารที่มีรายชื่อปรากฏในรายการของสถาบันประกันเงินฝากมาเป็นกลุ่มของธนาคารที่มีปัญหาทางการเงินจำนวน 110 ธนาคารแล้วทำการเลือกธนาคารที่ไม่มีปัญหาทางการเงินอีก 110 ธนาคาร โดยวิธีจับคู่ตามขนาดของธนาคารและจำนวนสาขา เขาเลือกตัวแปรมาศึกษาเพื่อจำแนกความแตกต่างทั้งหมด 10 ตัวแปรโดยที่ตัวแปรที่มีความสำคัญที่สุด คือ ตัวแปรในด้าน ประสิทธิภาพการบริหาร จากนั้นนำสมการที่ประมาณได้มาจำแนกกลุ่มตัวอย่าง พบว่ามีความน่าเชื่อถือของตัวแบบ จึงสรุปว่าสามารถนำวิธีการนี้ไปใช้ในการสร้างระบบสัญญาณเตือนภัยล่วงหน้าได้ในอนาคต

สำหรับการศึกษาในธุรกิจประกันภัยได้มีนักวิชาการที่ให้ความสนใจศึกษาด้วยวิธีการวิเคราะห์จำแนกกลุ่มเป็นจำนวนมากในการพยากรณ์ภาวะการล้มละลายของธุรกิจ โดยใช้ข้อมูลทางการเงิน (Trieschmann and Pinches, ค.ศ.1973, ค.ศ.1974; Harmelink, ค.ศ.1974; Ohlson, ค.ศ.1980; Eck, ค.ศ.1982; Richardson and Davidson, ค.ศ.1984; Taffler, ค.ศ.1982; Izan, ค.ศ.1984; Takashasital, ค.ศ.1984; Zmijewski, ค.ศ.1984; Zavgren, ค.ศ.1985; Barniv and Raveh,

ค.ศ.1986; Harrington and Nelson, ค.ศ.1986; Hershbarger and Miller, ค.ศ.1986; Barniv and Smith, ค.ศ.1987; Ambrose, J. M. and J. A. Seward, ค.ศ.1988; Barniv, R. and J. B. Mcdonald, ค.ศ.1992) Trieschmann และ Pinches (ค.ศ.1973) ศึกษาธุรกิจประกันวินาศภัยของประเทศออสเตรเลียและได้พยากรณ์การล้มละลายในระหว่าง ปี ค.ศ.1966 – ค.ศ.1971 โดยเลือก 26 บริษัทเป็นตัวอย่างของบริษัทที่ล้มละลายแล้วทำการสุ่มตัวอย่างบริษัทที่ไม่ล้มละลายมาจับคู่อีก 26 บริษัท เพื่อสร้างตัวแบบในการพยากรณ์ อัตราส่วนในการศึกษานี้มีถึง 70 อัตราส่วน พบว่ามี 6 อัตราส่วนที่สามารถจำแนกความแตกต่างระหว่างกลุ่มบริษัทที่ล้มละลายและไม่ล้มละลายออกจากกัน คือ 1. อัตราส่วนค่าตัวแทนนายหน้าต่อสินทรัพย์รวม (agent balance to total assets) 2.อัตราส่วนเบี้ยประกันภัยรับตรงต่อเงินสำรองส่วนเกิน (direct premium written to surplus) 3.อัตราส่วนรวม (combined ratio) 4.อัตราส่วนต้นทุนพันธบัตรต่อมูลค่าตลาดของพันธบัตร (bond-cost to bond market) 5.อัตราส่วนต้นทุนหุ้นทุนบวกอัตราส่วนค่าใช้จ่ายในการปรับค่าเสียหาย (stock cost plus loss adjustment expenses) และ 6.อัตราส่วนค่าใช้จ่ายในการรับประกันภัยต่อเบี้ยประกันภัยรับสุทธิ (underwriting expenses to net premium written) ซึ่งให้ผลการพยากรณ์ถูกต้องถึง 94 เปอร์เซ็นต์

วิธีการวิเคราะห์จำแนกกลุ่มนี้ ยังคงได้รับการพัฒนาให้มีความสามารถในการจัดจำแนกกลุ่มได้ดียิ่งขึ้น โดยต้องให้มีความแปรปรวนระหว่างกลุ่มต่อความแปรปรวนภายในกลุ่มสูงที่สุด ซึ่งจะต้องมีสมมมติฐานมารองรับ (Thornton, ค.ศ.1977; Barniv and Reveh, ค.ศ.1989) รวมทั้งมีการนำเสนอเทคนิคการจำแนกกลุ่ม 2 ขั้นขึ้นเพื่อประสิทธิภาพในการพยากรณ์ที่มีความถูกต้องมากยิ่งขึ้น โดยการใช้ตัวแปรหุ่นทั้งหมด (Dummy variable) (Ambrose and Seward, ค.ศ.1988)

สำหรับในประเทศไทยนั้นได้มีนักวิชาการได้ทำการศึกษาที่เกี่ยวข้องกับปัญหาความมั่นคงทางการเงิน ตัวอย่างเช่น การศึกษาความมั่นคงและตัวชี้ความมั่นคงของสถาบันการเงิน (ประศาสน์ ตั้งมติธรรม, พ.ศ.2530) การศึกษาพฤติกรรมการถือสินทรัพย์และความมั่นคงของธนาคารพาณิชย์ไทย (ภาณุพงศ์ นิธิประภา และอัจนา วัฒนานุกิจ, พ.ศ.2530) การศึกษาความมั่นคงในกลุ่มบริษัทเงินทุนและบริษัทเครดิตฟองซิเอร์ (สันติ ถิรพัฒน์ และคณะ, พ.ศ.2530) พบว่าผลการศึกษาในข้างต้นนี้มุ่งศึกษาในกลุ่มของธนาคารพาณิชย์ บริษัทเงินทุนหลักทรัพย์ และบริษัทเครดิตฟองซิเอร์ด้วยวิธีการวิเคราะห์จำแนกกลุ่มมาศึกษา โดยใช้ข้อมูลทางการเงินจากงบดุล งบกำไรขาดทุน เพื่อทำการจำแนกกลุ่มของสถาบันการเงินที่มีปัญหาและไม่มีปัญหาออกจากกัน แต่สำหรับธุรกิจประกันวินาศภัยในประเทศไทยนั้น การศึกษาเรื่องความมั่นคงในสถานะทางการเงินยังไม่เป็นที่แพร่หลายนัก ดังนั้นงานวิจัยนี้จึงมุ่งศึกษาที่จะวิเคราะห์การจำแนกกลุ่มเพื่อการทำนายความไม่มั่นคงในสถานะทางการเงินของบริษัทประกันวินาศภัยไทยว่ามีบริษัทประกันวินาศภัยใดในธุรกิจนี้ที่มีแนวโน้มที่จะประสบ

ปัญหาทางการเงินจนต้องมีการเข้าควบคุมอย่างใกล้ชิด เพื่อมิให้เกิดปัญหาการล้มละลายซึ่งจะส่งผลกระทบอย่างยิ่งต่อระบบเศรษฐกิจและภาพพจน์ของธุรกิจประกันวินาศภัยของประเทศเฉกเช่นเดียวกับการที่บริษัทเงินทุนและบริษัทเงินทุนหลักทรัพย์ต้องประสบอยู่ในระยะเวลาไม่นานมานี้