บทที่ 3
ระเบียบวิธีวิจัย

3.1 การรวบรวมข้อมูล

การศึกษาครั้งนี้เป็นการศึกษาเชิงปริมาณ(Quantitative Research)คือกระบวนการค้นหาความจริงโดยอาศัยข้อมูลที่เป็นเชิงปริมาณ และนำวิธีการทางสถิติมาใช้ในการวิเคราะห์

3.1.1 แหล่งที่มาของข้อมูลศึกษาเป็นข้อมูลทุติยภูมิ(Secondary data) เก็บรวบรวมจากงบดุลและงบกำไรขาดทุนของธนาคารพาณิชย์ไทย ระดับการวัด(Level of Measurement)ของข้อมูลเป็นมาตราอัตราส่วน(Ratio Scale) ประเภทของข้อมูลเป็นข้อมูลไม่ต่อเนื่อง(Discrete data)

3.1.2 ประชากรได้แก่ธนาคารพาณิชย์ไทยทั้ง 15 แห่งได้แก่ ธนาคารกรุงเทพ จำกัด (มหาชน) ธนาคารกรุงไทย จำกัด(มหาชน) ธนาคารกสิกรไทย จำกัด(มหาชน) ธนาคารไทยพาณิชย์ จำกัด(มหาชน) ธนาคารกรุงศรีอยุธยา จำกัด(มหาชน) ธนาคารทหารไทย จำกัด(มหาชน) ธนาคารมหานคร จำกัด(มหาชน) ธนาคารนครหลวงไทย จำกัด(มหาชน) ธนาคารกรุงเทพพาณิชย์การ จำกัด (มหาชน) ธนาคารศรีนคร จำกัด(มหาชน) ธนาคารเอเซีย จำกัด(มหาชน) ธนาคารไทยทนุ จำกัด (มหาชน) ธนาคารนครธน จำกัด(มหาชน) ธนาคารสหธนาคาร จำกัด(มหาชน) ธนาคารแหลมทอง จำกัด(มหาชน) และกลุ่มตัวอย่างรวบรวมจากประชากรทั้งหมดโดยเก็บข้อมูลงบการเงินในช่วงปี 2531-2543 รวม 195 ตัวอย่าง

กลุ่มตัวอย่างแบ่งแยกเป็นธนาคารพาณิชย์ที่ไม่ล้มละลาย 175 ตัวอย่างหรือเท่ากับร้อยละ 89.7 และเป็นธนาคารที่อยู่ในภาวะการล้มละลาย 20 ตัวอย่างหรือเท่ากับร้อยละ 10.3 โดยแบ่งเป็นธนาคารพาณิชย์ที่ล้มละลายในปี 2540 จำนวน 4 ธนาคารได้แก่ ธนาคารมหานคร ธนาคารนครหลวงไทย ธนาคารกรุงเทพพาณิชย์การ และธนาคารศรีนคร ส่วนปี 2541 มีจำนวนธนาคารที่ล้มละลาย 7 แห่งได้แก่ธนาคาร 4 แห่งในปี 2540 ซึ่งธนาคารมหานครและกรุงเทพพาณิชย์การได้เลิกกิจการไปรวมกับธนาคารสหธนาคาร ธนาคารนครธน(สแตนดาร์ดชาร์เตอร์ดนครธน) และธนาคารแหลมทอง (ยูโอบีรัตนสิน) สำหรับปี 2542 จำนวนธนาคารที่ยังคงอยู่ในเกณฑ์ล้มละลายมี 5 แห่งได้แก่ธนาคาร 4 แห่งในปี 2540 และธนาคารนครธน(สแตนดาร์ดชาร์เตอร์ดนครธน) ปี 2543 ยังมีมีธนาคารที่ล้มละลาย 4 แห่งคือธนาคารมหานครและกรุงเทพพาณิชย์การที่ได้เลิกกิจการไป รวมถึงธนาคารนครหลวงไทยและศรีนครที่อยู่ระหว่างการควบคุมของทางการ

3.1.3 ตัวแปรที่ใช้ในการศึกษาพิจารณาจากปัจจัยภายในซึ่งได้มาจากการศึกษาผลงานวิจัยที่เกี่ยวข้องและกรอบแนวความคิดทางทฤษฎีแบ่งเป็น 4 หมวดโดยหมวดแรกพิจารณาด้านอัตราส่วนแสดงสภาพคล่องจากสัดส่วนสินเชื่อต่อเงินฝาก หมวดที่สองอัตราส่วนแสดงคุณภาพสินทรัพย์พิจารณาสัดส่วนค่าเผื่อหนี้สงสัยจะสูญต่อเงินให้สินเชื่อ และดอกเบี้ยค้างรับต่อเงินให้สินเชื่อ หมวดที่สามอัตราส่วนแสดงความเพียงพอของทุนจากสัดส่วนเงินฝากต่อสินทรัพย์รวม ส่วนของผู้ถือหุ้นต่อสินทรัพย์รวม และหนี้สินต่อส่วนของผู้ถือหุ้น หมวดที่สี่แสดงประสิทธิภาพในการทำกำไรพิจารณาจากสัดส่วนรายได้ดอกเบี้ยและเงินปันผลต่อรายได้รวม รายได้มิใช่ดอกเบี้ยต่อรายได้รวม ค่าเผื่อหนี้สงสัยจะสูญต่อรายได้รวม รายได้รวมต่อจำนวนสาขา และรายได้รวมต่อจำนวนพนักงาน

3.2 แบบจำลองและสถิติที่ใช้ในการวิจัย

3.2.1 แบบจำลองที่ใช้ในการวิจัย

$$Y_i = \alpha + \beta_1 LD + \beta_2 LP + \beta_3 IL + \beta_4 DA + \beta_5 EA + \beta_6 DE + \beta_7 GP + \beta_8 RO + \beta_9 RP + \beta_{10} PB + \beta_{11} PM + e$$

โดย Y_i : การล้มละลายของธนาคารพาณิชย์

$Y_i = 1$ ธนาคารพาณิชย์ไม่ล้มละลาย

$Y_i = 0$ ธนาคารพาณิชย์ล้มละลาย

การล้มละลายของธนาคารพาณิชย์หมายถึงการที่ธนาคารพาณิชย์ใดๆถูกทางการสั่งให้ลดทุนจดทะเบียนก่อนที่จะให้มีการเพิ่มทุนจดทะเบียนใหม่ หรือถูกทางการสั่งให้เลิกกิจการธนาคารพาณิชย์

LD : อัตราส่วนเงินให้สินเชื่อต่อเงินฝาก

LP : อัตราส่วนค่าเผื่อหนี้สงสัยจะสูญต่อเงินให้สินเชื่อ

IL : อัตราส่วนดอกเบี้ยค้างรับต่อเงินให้สินเชื่อ

DA : อัตราส่วนเงินฝากต่อสินทรัพย์รวม

EA : อัตราส่วนส่วนของผู้ถือหุ้นต่อสินทรัพย์รวม

DE : อัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น

GP : อัตราส่วนรายได้ดอกเบี้ยและเงินปันผลต่อรายได้รวม

RO : อัตราส่วนรายได้มิใช่ดอกเบี้ยต่อรายได้รวม

RP : อัตราส่วนค่าเผื่อหนี้สงสัยจะสูญต่อรายได้รวม

PB : อัตราส่วนรายได้รวมต่อจำนวนสาขา

PM : อัตราส่วนรายได้รวมต่อจำนวนพนักงาน

α : ค่าคงที่(constant)

β_i : ค่าสัมประสิทธิ์หน้าตัวแปรที่นำมาอธิบาย(coefficiences)

e : ค่าความคลาดเคลื่อน(error term)

3.2.2 สมมุติฐานของการวิจัย

อัตราส่วนเงินให้สินเชื่อต่อเงินฝาก(LD) หมายถึง สัดส่วนเงินให้สินเชื่อของธนาคารหารด้วยสัดส่วนเงินฝากทั้งหมดของธนาคาร แสดงถึงสภาพคล่องของธนาคารจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มีค่ามากเท่ากับธนาคารมีการให้กู้ยืมมากจะทำให้สภาพคล่องของธนาคารน้อยลงและอาจทำให้ธนาคารเกิดปัญหาได้

สมมุติฐานหลัก ; $H_0 = 0$

(เงินให้สินเชื่อต่อเงินฝากไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; $H_a \neq 0$

(เงินให้สินเชื่อต่อเงินฝากมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนค่าเผื่อหนี้สงสัยจะสูญต่อเงินให้สินเชื่อ(LP) หมายถึง สัดส่วนค่าเผื่อหนี้สงสัยจะสูญของธนาคารหารด้วยเงินให้สินเชื่อของธนาคาร แสดงถึงคุณภาพสินทรัพย์ของธนาคารจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มีค่ามากขึ้นหมายถึงธนาคารไม่ได้รับชำระเงินต้นคืนจำนวนมากและอาจทำให้ธนาคารเกิดปัญหาได้

สมมุติฐานหลัก ; $H_0 = 0$

(ค่าเผื่อหนี้สงสัยจะสูญต่อเงินให้สินเชื่อไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; $H_a \neq 0$

(ค่าเผื่อหนี้สงสัยจะสูญต่อเงินให้สินเชื่อมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนดอกเบี้ยค้างรับต่อเงินให้สินเชื่อ(IL) หมายถึง อัตราส่วนดอกเบี้ยค้างรับหารด้วยเงินให้สินเชื่อของธนาคาร แสดงถึงคุณภาพสินทรัพย์ของธนาคารจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มีค่ามากขึ้นหมายถึงธนาคารไม่ได้รับชำระดอกเบี้ยคืนจำนวนมากและอาจส่งผลให้ธนาคารขาดทุนได้

สมมุติฐานหลัก ; $H_0 = 0$

(ดอกเบี้ยค้างรับต่อเงินให้สินเชื่อไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; $H_a \neq 0$

(ดอกเบี้ยค้างรับต่อเงินให้สินเชื่อมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนเงินฝากต่อสินทรัพย์รวม(DA) หมายถึงสัดส่วนเงินฝากหารด้วยสินทรัพย์รวมของธนาคาร แสดงถึงระดับความเพียงพอของทุนของธนาคารจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มีค่ามากจะช่วยให้ธนาคารสามารถปล่อยสินเชื่อได้มาก และทำให้มีรายได้จากดอกเบี้ยมาก

สมมุติฐานหลัก ; $H_0 = 0$

(เงินฝากต่อสินทรัพย์รวมไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; $H_a \neq 0$

(เงินฝากต่อสินทรัพย์รวมมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนส่วนของผู้ถือหุ้นต่อสินทรัพย์รวม(EA) หมายถึง สัดส่วนของผู้ถือหุ้นหารด้วยสินทรัพย์รวมของธนาคารซึ่งหากมีค่ามากขึ้น แสดงถึงระดับความเพียงพอของทุนของธนาคารจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากหมายถึงธนาคารมีส่วนของทุนมากทำให้ดำเนินธุรกรรมต่างๆได้มากขึ้น

สมมุติฐานหลัก ; $H_0 = 0$

(ส่วนของผู้ถือหุ้นต่อสินทรัพย์รวมไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; $H_a \neq 0$

(ส่วนของผู้ถือหุ้นต่อสินทรัพย์รวมมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนหนี้สินต่อส่วนของผู้ถือหุ้น(DE) หมายถึง สัดส่วนหนี้สินหารด้วยส่วนของผู้ถือหุ้นของธนาคาร แสดงถึงระดับความเพียงพอของทุนของธนาคารจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากหมายถึงธนาคารมีหนี้สินมากเมื่อเปรียบเทียบกับส่วนของทุนมากทำให้มีความเสี่ยงต่อการล้มละลายมากขึ้น

สมมุติฐานหลัก ; $H_0 = 0$

(หนี้สินต่อส่วนของผู้ถือหุ้นไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; $H_a \neq 0$

(หนี้สินต่อส่วนของผู้ถือหุ้นมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนรายได้ดอกเบี้ยและเงินปันผลต่อรายได้รวม(GP) หมายถึง สัดส่วนรายได้ดอกเบี้ยและปันผลหารด้วยรายได้รวมของธนาคาร แสดงถึงแสดงถึงประสิทธิภาพในการทำกำไรจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากธนาคารจะมีรายได้มากและส่งผลให้มีกำไรมากตามไปด้วย ดังนั้นธนาคารจึงมีความมั่นคงมากขึ้น

สมมุติฐานหลัก ; H_0 = 0

(รายได้ดอกเบี้ยและปันผลต่อรายได้รวมไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; H_a ≠ 0

(รายได้ดอกเบี้ยและปันผลต่อรายได้รวมมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนรายได้มิใช่ดอกเบี้ยต่อรายได้รวม(RO) หมายถึง สัดส่วนรายได้มิใช่ดอกเบี้ยหารด้วยรายได้รวมของธนาคาร แสดงถึงแสดงถึงประสิทธิภาพในการทำกำไรจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากธนาคารจะมีรายได้มากและส่งผลให้มีกำไรมากตามไปด้วย ดังนั้นธนาคารจึงมีความมั่นคงมากขึ้น

สมมุติฐานหลัก ; H_0 = 0

(รายได้มิใช่ดอกเบี้ยต่อรายได้รวมไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; H_a ≠ 0

(รายได้มิใช่ดอกเบี้ยต่อรายได้รวมมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนค่าเผื่อหนี้สงสัยจะสูญต่อรายได้รวม(RP) หมายถึง สัดส่วนค่าเผื่อหนี้สงสัยจะสูญหารด้วยรายได้รวมของธนาคาร แสดงถึงแสดงถึงประสิทธิภาพในการทำกำไรจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากธนาคารจะมีค่าใช้จ่ายมากและส่งผลให้ขาดทุนมากตามไปด้วย ดังนั้นความมั่นคงของธนาคารจึงลดลง

สมมุติฐานหลัก ; H_0 = 0

(ค่าเผื่อหนี้สงสัยจะสูญต่อรายได้รวมไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; H_a ≠ 0

(ค่าเผื่อหนี้สงสัยจะสูญต่อรายได้รวมมีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนรายได้รวมต่อจำนวนสาขา(PB) หมายถึง สัดส่วนรายได้รวมหารด้วยจำนวนสาขาของธนาคาร แสดงถึงแสดงถึงประสิทธิภาพในการทำกำไรจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากธนาคารจะมีกำไรมากและธนาคารจะมีความมั่นคงมาก

สมมุติฐานหลัก ; H_0 = 0

(รายได้รวมต่อจำนวนสาขาไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; H_a ≠ 0

(รายได้รวมต่อจำนวนสาขามีผลต่อการล้มละลายของธนาคารพาณิชย์)

อัตราส่วนรายได้รวมต่อจำนวนพนักงาน(PM) หมายถึง สัดส่วนรายได้รวมหารด้วยจำนวนพนักงานของธนาคาร แสดงถึงแสดงถึงประสิทธิภาพในการทำกำไรจึงตั้งสมมุติฐานว่าหากสัดส่วนนี้มากธนาคารจะมีกำไรมากและธนาคารจะมีความมั่นคงมาก

สมมุติฐานหลัก ; H_0 = 0

(รายได้รวมต่อจำนวนพนักงานไม่มีผลต่อการล้มละลายของธนาคารพาณิชย์)

สมมุติฐานรอง ; H_a ≠ 0

(รายได้รวมต่อจำนวนพนักงานมีผลต่อการล้มละลายของธนาคารพาณิชย์)

3.2.3 สถิติที่ใช้ในการวิจัย

การอ้างอิงถึงประชากรโดยใช้ข้อมูลจากตัวอย่างโดยการประมาณค่าและการทดสอบสมมุติฐานจำเป็นต้องทราบถึงลักษณะการแจกแจงของประชากรที่สนใจศึกษาหากประชากรมีการแจกแจงแบบปกติหรือใกล้เคียงแบบปกติจึงจะสามารถใช้ Z หรือ t หรือ F เป็นสถิติทดสอบซึ่งรวมเรียกว่าการทดสอบสมมุติฐานที่ใช้พารามิเตอร์(Parametric Tests)

ในกรณีที่ประชากรไม่ได้มีการแจกแจงปกติหรือตัวอย่างมีขนาดเล็กต้องใช้วิธีการทดสอบที่ไม่ใช้พารามิเตอร์(Nonparametric Tests) ประสิทธิภาพของการทดสอบโดยไม่ใช้พารามิเตอร์นั้นด้อยกว่าวิธีที่ใช้พารามิเตอร์ เนื่องจากการทดสอบที่ไม่ใช้พารามิเตอร์ใช้ลำดับที่ของข้อมูลเท่านั้นในการวิเคราะห์ในขณะที่การทดสอบที่ใช้พารามิเตอร์นำค่าของข้อมูลจริงมาวิเคราะห์[1]

เครื่องมือทางสถิติเชิงคุณภาพสำหรับการทดสอบแบบไม่ใช้พารามิเตอร์สำหรับ ใช้ในการทำนายภาวะการนำไปสู่การล้มละลายของระบบธนาคารพาณิชย์มีหลายวิธี โดยวิธีแรกเป็นการวิเคราะห์อัตราส่วนทางการเงิน(Financial ratios analysis)เป็นการพิจารณาทีละอัตราส่วนเปรียบเทียบกันเองระหว่างธนาคาร ธนาคารใดมีอัตราส่วนที่ดีกว่าเมื่อเปรียบเทียบกับกลุ่มหมายถึงมีประสิทธิภาพมากกว่า แต่มีข้อจำกัดคือหาข้อสรุปได้ยากว่าอัตราส่วนที่เหมาะเพื่อใช้เป็นเกณฑ์มาตรฐานควรมีค่าเท่าใด เช่น อาจมีบางธนาคารมีกำไรน้อยและมีส่วนของผู้ถือหุ้นต่ำน่าจะหมายถึงเกิดปัญหาในการทำธุรกิจ แต่หากธนาคารนั้นมีสภาพคล่องสูงก็อาจทำให้ภาพรวมธนาคารนั้นไม่เกิดปัญหายังสามารถดำเนินธุรกิจได้ ดังนั้นการใช้การวิเคราะห์อัตราส่วนทางการเงินแต่ละตัวมาวิเคราะห์ความสามารถในการประกอบธุรกิจอาจจะไม่นำไปสู่ข้อสรุปที่ตรงกันได้

---

[1] กัลยา วานิชย์บัญชา ,*การทดสอบสมมุติฐานที่ไม่ใช้พารามิเตอร์* ,การวิเคราะห์สถิติ:สถิติเพื่อการตัดสินใจ , หน้า 215-216 .

เครื่องมือต่อมาได้แก่แบบจำลอง The Z-score ของ Edward I. Altman ซึ่งได้ใช้วิธี Discriminant Analysis มาวิเคราะห์ความมั่นคงของธุรกิจในประเทศสหรัฐอเมริกาครั้งแรกในปี 1968 และพัฒนาจนกระทั่งปี 1983 เป็นแบบจำลองทางเศรษฐมิติที่ใช้วิเคราะห์ความมั่นคงของธุรกิจในประเทศสหรัฐอเมริกา แบบจำลองนี้เป็นแนวความคิดทางด้านการวิเคราะห์โดยรวมการวิเคราะห์อัตราส่วนทางการเงิน(ratios analysis)เข้ากับการใช้เครื่องมือทางสถิติ(statistical techniques)ซึ่งจะได้สมการ

$$Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 1.0X_5$$

โดยที่ Z = Z - Score Index

X_1 = Current Assets / Total Assets อัตราส่วนนี้เป็นการวัดสินทรัพย์สภาพคล่องของธนาคารพาณิชย์ นอกเหนือจากอัตราส่วนอื่นในประเภทเดียวกันนี้ อาทิเช่น Current Ratio และ Quick Ratio

X_2 = Retained Earnings / Total Assets กำไรสะสมเป็นบัญชีที่แสดงถึงจำนวนของรายรับที่เกิดจากการลงทุน(Reinvested Earnings) ตลอดจนการประสบภาวะขาดทุนของธุรกิจในช่วงระยะเวลาการดำเนินงาน นอกจากนี้ยังหมายถึงส่วนเกินจากการทำมาหาได้ของธุรกิจ อัตราส่วนนี้เป็นอัตราส่วนที่มุ่งเน้นการพิจารณาถึงเงินทุนสะสมจากแหล่งภายใน โดยการวัดความสามารถในการทำกำไรตลอดอายุของกิจการด้วยข้อสมมุติฐานที่ว่าธุรกิจจะดำเนินธุรกิจต่อไปในอนาคต(Going Concern)

X_3 = Profit Before Taxes / Total Assets อัตราส่วนนี้เป็นการวัดประสิทธิภาพในการใช้สินทรัพย์ในการดำเนินธุรกิจตามปกติของธนาคารว่าเป็นอย่างไร บรรลุเป้าหมายในการจัดการหรือไม่ หากอัตราส่วนนี้สูงแสดงว่าธนาคารนั้นๆ ใช้สินทรัพย์อย่างมีประสิทธิภาพ

X_4 = Net Worth / Total Liabilities เป็นการพิจารณาถึงสัดส่วนแห่งภาระผูกพันทางการเงินหรือโครงสร้างเงินทุนของกิจการ ตัวแปร X_4 ในที่นี้จะใช้ส่วนของเจ้าของสุทธิตามมูลค่าทางบัญชี (the book value of net worth) แทนส่วนของเจ้าของตามมูลค่าราคาตลาดของหุ้น(market value of net worth) แทนส่วนของเจ้าของตามมูลค่าราคาตลาดของหุ้น(market value of equity) ตามสูตรเดิม ทั้งนี้เพื่อเป็นการขจัดความผิดปกติของข้อมูลราคาหุ้นตามราคาตลาดซึ่งไม่สะท้อนภาพที่เป็นจริงของมูลค่าส่วนของเจ้าของ อีกทั้งจากการศึกษาของ Altman ได้ระบุถึงการแทนค่า Market Value of Equity ด้วย Net Worth ตามมูลค่าทางบัญชี ว่าจะสามารถเพิ่มประสิทธิภาพในการทำนายความมั่นคงของหน่วยธุรกิจได้ดี ซึ่งในการศึกษาเกี่ยวกับเรื่อง Bankruptcy ในระยะต่อมาอัตราส่วนนี้ก็ได้ถูกนำมาใช้ในฐานะที่เป็นอัตราส่วนที่มีประสิทธิภาพในการแบ่งแยกกลุ่มบริษัทที่มีปัญหาทางการเงินออกจากบริษัทที่ไม่มีปัญหา

X_5 = Total Interest and Dividend Income / Total Assets อัตราส่วนนี้เป็นการวัดผลการดำเนินงานและถือเป็นเครื่องวัดความสามารถของผู้บริหารและสินทรัพย์ธนาคารพาณิชย์ว่ามีประสิทธิภาพในอันที่จะก่อให้เกิดรายได้ในการบรรลุเป้าหมายของการจัดการหรือไม่หากอัตราส่วนนี้มีค่าสูงย่อมแสดงว่าธนาคารนั้นๆ ใช้สินทรัพย์อย่างมีประสิทธิภาพ

การวิเคราะห์หาค่าดัชนีรวม หรือ ค่า Z-Score ของธนาคารพาณิชย์ทั้งระบบได้ใช้ตัวเลขที่เป็นค่าเฉลี่ยในช่วงระยะเวลาที่ศึกษาเป็นฐานในการคำนวณหาค่าดัชนีรวม สำหรับค่าดัชนีรวมที่ได้จะใช้เป็นเกณฑ์สำหรับพิจารณาว่า ธนาคารพาณิชย์มีความมั่นคงทางการเงินหรือไม่ อย่างไร เมื่อเปรียบเทียบกับบรรทัดฐานของระบบ โดยที่ค่าดัชนี Z-Score ที่ได้จะถือเป็นบรรทัดฐานของระบบในการชี้วัดความมั่นคงทางการเงินของแต่ละธนาคาร หากค่า Z-Scores ของธนาคารใดมีค่าต่ำกว่าค่า Z-Score ของทั้งระบบ ธนาคารนั้นจะถูกจัดเป็นธนาคารที่มีความมั่นคงต่ำ โดยเฉพาะหากค่าดัชนีของธนาคารใดที่คำนวณได้ยิ่งมีค่าต่ำกว่าค่าดัชนีของทั้งระบบที่เป็นบรรทัดฐานความมั่นคงของธนาคารนั้นก็จะยิ่งลดลงหรือไม่มีความมั่นคง ในทางตรงกันข้ามหากค่าดัชนี Z-Score ของธนาคารใดมีค่าสูงกว่าค่าดัชนีของทั้งระบบ ธนาคารนั้นจะถูกจัดเป็นธนาคารที่มีความมั่นคง หากค่าดัชนีของธนาคารใดยิ่งมีค่าสูงกว่าค่าดัชนีของทั้งระบบ ธนาคารนั้นก็จะมีความมั่นคงสูงตามไปด้วย อย่างไรก็ตามเนื่องจากแบบจำลองนี้ถูกพัฒนาขึ้นตามระบบเศรษฐกิจของประเทศสหรัฐอเมริกาซึ่งมีการกำหนดค่าสัมประสิทธิ์(coefficients)ของตัวแปรอิสระไว้แน่นอน หากนำแบบจำลองมาใช้ในประเทศอื่นๆที่มีระบบเศรษฐกิจต่างกันความสามารถในการพยากรณ์ก็อาจแตกต่างกันด้วย

ซึ่งต่อมาภาณุพงศ์และอัจฉราได้ใช้เครื่องมือ Multiple discriminant analysis ในการสร้างแบบจำลองทำนายความมั่นคงของธนาคารพาณิชย์ในประเทศไทย โดยมีการปรับตัวแปรอิสระให้เหมาะสมมากขึ้น Discriminant Analysis เป็นเครื่องมือทางสถิติที่ใช้ทำนายความน่าจะเป็นของการเกิดเหตุการณ์ตั้งแต่ 2 เหตุการณ์ขึ้นไปซึ่งเป็นการทดสอบโดยไม่ใช้พารามิเตอร์[2] ที่มีตัวแปรตามเป็นตัวแปรเชิงคุณภาพ(qualitative)และมีสมการเป็น

$$Z_i = b_1 X_{1i} + b_2 X_{2i} \ldots\ldots + b_n X_{ni}$$

โดยที่ X_{ji} = ตัวแปรอิสระแต่ละตัว

b_j = ค่าสัมประสิทธิ์ของตัวแปรแต่ละตัว

Z_i = คะแนนของค่า discriminant(Z-score)

---

[2] William G. Zikmund , *Multivariate Analysis* , Business Research Methods , P. 535

การทำงานของ discriminant Analysis เกิดจากการหาความสัมพันธ์ของ linear combination และเพื่อป้องกันการเกิดปัญหาตัวแปรอิสระมีความสัมพันธ์กัน(multicollinearity) แบบจำลองมีการนำตัวแปรเข้าโดยวิธี stepwise[3] คือนำตัวแปรอิสระเข้าสมการครั้งละ 1 ตัว(forward)โดยเลือกตัวแปรอิสระที่มีความสัมพันธ์กับตัวแปรตามมากที่สุดเข้าเป็นตัวแรก จากนั้นจึงเลือกตัวแปรอิสระที่เหลือเข้าสมการโดยคำนวณค่าสัมประสิทธิ์สหสัมพันธ์เชิงส่วนระหว่างตัวแปรตามและตัวแปรอิสระที่เหลือทีละตัว จากนั้นจึงตัดตัวแปรอิสระออก(backward)โดยกำหนด entry และ remove ป้องกันไม่ให้นำตัวแปรเดิมเข้าหรือออกจากสมการจนกระทั่งไม่มีตัวแปรใดออกจากสมการอีก

การกำหนด entry และ remove อาศัยสถิติ wilks' lambda likelihood criterion และจากการประมวลผลจะได้แบบจำลอง discriminant Analysis ซึ่งค่าสัมประสิทธิ์บอกถึงความสำคัญของตัวแปรอิสระแต่ละตัวในสมการที่เกิดจากความสัมพันธ์ของน้ำหนักตัวแปรระหว่างกลุ่มที่แบ่งแยก 2 กลุ่ม โดยมีค่า Z-score เป็นตัวชี้วัดการแบ่งแยกออกเป็น 2 กลุ่มอธิบายได้ดังนี้

Z	=	0	หมายถึงมีโอกาสที่จะเกิดการล้มละลายร้อยละ 50
Z	<	0	หมายถึงมีโอกาสที่จะล้มละลายน้อยกว่าร้อยละ 50
Z	>	0	หมายถึงมีโอกาสที่จะล้มละลายมากกว่าร้อยละ 50

ดังนั้นจุดตัดจึงอยู่ที่ค่า Z-score เท่ากับ 0 หากค่ามากกว่า 0 จะทำนายว่าล้มละลายแต่ถ้ามีค่าน้อยกว่า 0 จะทำนายว่ามีความมั่นคง ผลของการทำนายทั้งหมดจะอยู่ในรูปแบบของตาราง hit ratio ซึ่งมีการเปรียบเทียบค่าทำนายกับค่าจริงและคำนวณความถูกต้องหรือผิดพลาดในอัตราร้อยละเพื่อพิจารณาถึงความสามารถในการพยากรณ์

ส่วนเครื่องมือทางสถิติที่ใช้ในการทดสอบแบบใช้พารามิเตอร์(Parametric Tests) สำหรับแบบจำลองเชิงคุณภาพซึ่งตัวแปรตามมีสองทางเลือก(Binary-choice models)มีด้วยกัน 4 เครื่องมือ[4] โดยเครื่องมือแรกได้แก่ Linear Probability Model เป็นแบบจำลองเชิงคุณภาพเพื่อใช้งานกรณีที่ตัวแปรตามมี 2 ทางเลือก(binary-choice models) การสร้างแบบจำลอง Linear Probability Model เริ่มจากสมการถดถอย(regression)คือ

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

---

[3] Amir D. Aczel , *Discriminant Analysis* , Complete Business Statistics ,P. 790

[4] Robert S. Pindyck and Daniel L. Rubinfeld, *Econometric Models and Economic Forecasts, Models of Qualitative choice*, p.298-333.

โดยที่ $Y_i = 1$ สำหรับทางเลือกที่ 1

$Y_i = 0$ สำหรับทางเลือกที่ 2

ε_i มีการแจกแจงแบบปกติและมีค่าเฉลี่ย(mean)เท่ากับ 0

ทำการใส่ค่าเฉลี่ย(expected value) ให้กับตัวแปรอิสระดังสมการ

$$E(Y)_i = \alpha + \beta X_i$$

สามารถอธิบายการกระจายของ Y จาก $P_i = \text{Prob}(Y_i = 1)$และ $1\text{-}P_i = \text{Prob}(Y_i = 0)$ได้ดังนี้

$$E(Y)_i = 1(P_i) + 0(1 - P_i) = P_i$$

ดังนั้นสมการของ linear probability model จึงอธิบายได้ดังนี้

$$P_i = \begin{cases} \alpha + \beta X_i \rightarrow 0 < \alpha + \beta X_i < 1 \\ 1 \quad\quad\quad \rightarrow \alpha + \beta X_i \geq 1 \\ 0 \quad\quad\quad \rightarrow \alpha + \beta X_i \leq 0 \end{cases}$$

ค่าความน่าจะเป็นในการกระจายของ error term ในแบบจำลองจะถูกแทนที่โดยค่าของ Y_i (1,0) และความสัมพันธ์ระหว่างค่าความน่าจะเป็นของ P_i และ X_i เมื่อมีข้อสมมุติว่า error term มีค่าเฉลี่ยเท่ากับ 0 เป็นดังนี้

$$E(\varepsilon)_i = (1 - \alpha - \beta X_i)P_i + (-\alpha - \beta X_i)(1 - P_i) = 0$$

และหาค่า P_i ได้เป็น

$$P_i = \alpha + \beta X_i$$

และ

$$1 - P_i = 1 - \alpha - \beta X_i$$

คำนวณหาค่าความแปรปรวน(variance)ของ error term

$$E(\varepsilon_i^2) = \begin{cases} (1 - \alpha - \beta X_i)^2 P_i + (-\alpha - \beta X_i)^2 (1 - P_i) \\ (1 - \alpha - \beta X_i)^2 (\alpha + \beta X_i) + (\alpha + \beta X_i)^2 (1 - \alpha - \beta X_i) \\ (1 - \alpha - \beta X_i)(\alpha + \beta X_i) = P_i(1 - P_i) \end{cases}$$

หรือ

$$\sigma_i^2 = E(\varepsilon_i^2) = E(Y_i)[1 - E(Y_i)]$$

ค่าสังเกตุของ P_i เข้าสู่ 0 หรือ 1 จะมีค่าความแปรปรวนต่ำ ในขณะที่ P_i เข้าสู่ 0.5 จะมีค่าความแปรปรวนสูง ดังนั้นความแปรปรวนของ error term ไม่คงที่เกิดปัญหา heteroscedasticity ซึ่ง

ทำให้แบบจำลองขาดประสิทธิภาพ(efficiency) ถึงแม้ว่าค่ากำลังสองน้อยที่สุด(least squares)จะยังคงมีความสอดคล้อง(consistent) และ unbiased การแก้ปัญหา heteroscedasticity โดยใช้วิธี weighted least-squares ยังไม่สามารถบอกได้ว่าค่าทำนาย(Y_i)ที่คำนวณได้จะอยู่ในช่วง(0,1) หากค่าทำนาย(Y_i) อยู่นอกช่วง(0,1)จะต้องตัดค่าสังเกตุที่เกินออกหรือปรับค่าสังเกตุให้เท่ากับ 0.01 หรือ 0.99 เพื่อให้ค่า (Y_i) ทำนายอยู่ในช่วงที่กำหนด

เครื่องมือต่อมาได้แก่ Probit Model เป็นแบบจำลองเชิงคุณภาพเพื่อใช้งานกรณีที่ตัวแปรตามมี 2 ทางเลือก(binary-choice models) แบบจำลองนี้สามารถแก้ปัญหาของ linear probability model ได้ดีขึ้นเนื่องจากรูปแบบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระเป็นเส้นโค้งซึ่งอธิบายผลได้ดีกว่าแบบเส้นตรง รวมถึง การที่ค่าทำนาย(Y_i)จะอยู่ระหว่าง(0,1)ได้ดีขึ้นอันเนื่องมาจากการใช้ cumulative probability function ซึ่งทำให้ความน่าจะเป็นของการกระจายเป็นดังนี้

$$P_i = F(\alpha + \beta X_i) = F(Z_i)$$

แบบจำลอง probit เริ่มต้นจากแบบจำลอง linear probability แต่แตกต่างที่มีฟังก์ชั่นอยู่ในรูปแบบของ cumulative normal probability function ซึ่งสามารถเขียนเป็นแบบจำลองได้เป็น

$$Z_i = \alpha + \beta X_i$$

ข้อมูลจากค่าสังเกตุทั้งหมดจะเป็นตัวแบ่งแยกค่า Z_i ออกเป็น 2 ทางเลือกโดยทางเลือกแรกเป็นกลุ่มที่มีค่า Z_i สูงและทางเลือกที่สองเป็นกลุ่มที่มีค่า Z_i ต่ำ หากสมมุติให้ค่า Y เป็นตัวแทนของตัวแปรหุ่น(dummy)ที่เท่ากับ 1 เมื่อเลือกทางเลือกแรกและเท่ากับ 0 เมื่อเลือกทางเลือกที่สอง และกำหนดให้ค่า Z^*_i แทนค่าวิกฤตที่เป็นจุดตัดซึ่งจะสามารถอธิบายความสัมพันธ์ระหว่างค่า Z_i และค่า Z^*_i ได้ดังนี้

ผลการทำนายเป็นทางเลือกที่หนึ่งถ้า $Z_i > Z^*_i$

ผลการทำนายเป็นทางเลือกที่สองถ้า $Z_i \leq Z^*_i$

แบบจำลอง probit สมมุติให้ค่า Z^*_i เป็นตัวแปรเชิงสุ่มที่มีการกระจายแบบปกติ ดังนั้นความน่าจะเป็นที่ค่า Z^*_i น้อยกว่าหรือเท่ากับค่า Z_i สามารถคำนวณได้จาก cumulative normal probability function ดังต่อไปนี้

$$P_i = F(Z_i) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{Z_i} e^{-s^2/2}\, ds$$

โดยที่ S เป็นตัวแปรเชิงสุ่มที่มีการกระจายแบบปกติด้วยค่าเฉลี่ยเท่ากับ 0 และมีความแปรปรวนเท่ากับ 1 ส่วน P_i เป็นความน่าจะเป็นของการเกิดเหตุการณ์ที่มีค่าอยู่ระหว่าง(0,1) หรือความน่าจะเป็นจะเกิดขึ้นภายใต้พื้นที่โค้งปกติตั้งแต่ $-\infty$ ถึง Z_i

sample observation ได้แก่ $Y_1, Y_2, Y_3, \ldots, Y_N$ โดยที่ Y_i ต้องมีการแจกแจงแบบปกติ หากทางเลือกที่ 1 ถูกเลือก n_1 ครั้งและทางเลือกที่ 2 ถูกเลือก n_2 ครั้งและ($n_1 + n_2 = N$) จะสามารถกำหนด Likelihood function ได้เป็น

$$L = \text{Prob}(Y_1)\,\text{Prob}(Y_2) \ldots\ldots \text{Prob}(Y_N)$$

ค่าความน่าจะเป็นของการเลือกทางเลือกที่สองเท่ากับ 1 หักด้วยค่าความน่าจะเป็นของการเลือกทางเลือกที่หนึ่งดังสมการหรือหากกำหนด L คือ Likelihood Function

$$\log L = \sum_{i=1}^{n_1} \log P_i + \sum_{i=n_1+1}^{N} \log(1 - P_i)$$

ตัวแปรตามจะขึ้นอยู่กับ Logarithm ของอัตราส่วนระหว่างความน่าจะเป็นที่จะเกิดผลสำเร็จต่อความน่าจะเป็นที่จะไม่ก่อให้เกิดผลสำเร็จ(odds) เนื่องจากการที่ประชากรมีความแตกต่างกันทำให้กลุ่มตัวอย่างบางกลุ่มมีความแตกต่างกัน ในทางสถิติจะต้องมีกลุ่มตัวอย่างที่มีคุณสมบัติเหมือนกันดังนั้นแบบจำลอง probit จึงต้องมีการประมาณค่ากลุ่มตัวอย่าง(Maximum-Likelihood Estimation) ซึ่งทำให้ parameter estimations ของ α และ β เหมือนการเลือกตัวอย่างที่ต้องการ ทำการ differentiate ที่ log-likehood function และทำการ maximize เพื่อให้ได้ค่าเพื่อให้ได้ α^* และ β^*

การประมาณค่าโดยกระบวนการ maximum-likelihood จะมีผลให้ค่าพารามิเตอร์ consistent และ efficient asymptotically คือค่าพารามิเตอร์ทุกตัวมีการกระจายแบบปกติและทำให้สามารถทดสอบระดับนัยสำคัญทางสถิติ เช่น สถิติ t หรือ F ได้ และเมื่อได้ค่าพารามิเตอร์ α^* และ β^* จากการ maximize และได้สมการ probit แล้วจึงทำการทดสอบความมีนัยสำคัญของค่าสัมประสิทธิ์แต่ละตัวในแบบจำลองโดยทดสอบสมมุติฐานว่าค่า β มีค่าเท่ากับ 0 หรือไม่โดยใช้เครื่องมือThe Likelihood Ratio Test (λ) ซึ่งเป็นการเปรียบเทียบระหว่างฟังก์ชั่นความน่าจะเป็นสูงสุดที่ไม่มีข้อจำกัด : $L(\beta_{UR})$ และฟังก์ชั่นความน่าจะเป็นสูงสุดที่มีข้อจำกัด: $L(\beta_R)$

$$\lambda = \frac{L(\beta_R)}{L(\beta_{UR})}$$

ค่า λ อยู่ในช่วง 0 - 1 หากค่าที่ได้เข้าใกล้ 1 หมายถึงสมมุติฐานหลักเป็นจริง และหากค่าเข้าใกล้ 0 ก็จะปฏิเสธสมมุติฐานหลัก สำหรับการทดลองที่มีกลุ่มตัวอย่างขนาดใหญ่(large sample size) นั้น The Likelihood Ratio Test จะถูกนำมาเปรียบเทียบกับค่าไคว์สแคว : X^2m

$$-2(L(\beta_R) - L(\beta_{UR})) \sim X^2m \quad , m = \text{number of restrictions}$$

หากค่า X^2m มากกว่าค่าวิกฤตจะปฏิเสธสมมุติฐานหลักหรือหมายถึง β ไม่เท่ากับ 0 จากนั้นจึงทำการทดสอบความเหมาะสมของแบบจำลอง(goodness of fit)จากค่า The log-likelihood index : ρ

$$\rho = 1 - \left[\frac{L_0}{L_{\max}}\right]$$

โดยที่ L_0 หมายถึงค่าเริ่มต้นของ Likelihood function

L_{max} หมายถึงค่าสูงสุดของ Likelihood function

ค่าดัชนีหรือค่า R^2 ในสมการถดถอย (regression) ทั่วไปค่า R^2 อยู่ระหว่าง 0 ถึง 1 ถ้าค่าที่ได้อยู่ใกล้ 1 หมายถึงมีความเหมาะสมมาก อย่างไรก็ตามในแบบจำลองที่เรียกว่า binary dependent variable Model นั้นค่า R^2 ไม่จำเป็นต้องเข้าใกล้ 1 และขอบเขตบนที่เป็นไปได้ของค่า R^2 เท่ากับ 0.33

เครื่องมือ Logit Model เป็นแบบจำลองที่ช่วยงานวิเคราะห์ทางด้านคุณภาพเพื่อใช้งานกรณีที่ตัวแปรตามมี 2 ทางเลือก(binary-choice models) เริ่มใช้ครั้งแรกโดย Berkson(1944) จากการวิเคราะห์แบบจำลองด้วยวิธีกำลังสองน้อยที่สุดแบบถ่วงน้ำหนักที่ทำซ้ำ(iteratively weighted least squares) หลักของการทำงานของแบบจำลอง logit คำนวณจากการสะสมความน่าจะเป็นของ logit (comulative logistic probability function) แบบจำลอง logit มีคุณสมบัติคล้ายกับ probit มากคือรูปแบบความสัมพันธ์ระหว่างตัวแปรตามและตัวแปรอิสระเป็นเส้นโค้งซึ่งอธิบายผลได้ดีกว่าแบบเส้นตรงและทำให้ค่าทำนาย(Y_i)จะอยู่ระหว่าง(0,1)ได้ดี แต่มีข้อแตกต่างคือฟังก์ชั่นความน่าจะเป็นของ cumulative logistic มีการกระจายส่วนหางอ้วนกว่าเล็กน้อยเมื่อเทียบกับ cumulative normal และมีความสะดวกในการใช้คำนวณมากกว่าทำให้นิยมใช้แบบจำลอง logit แทน probit

$$P_i = F(Z_i) = F(\alpha + \beta X_i)$$

หรือ

$$P_i = \frac{1}{1 + e^{-Z_i}}$$

หรือ

$$P_i = \frac{1}{1 + e^{-(\alpha + \beta\chi_i)}}$$

คูณด้วย $1 + e^{-Zi}$

$$P_i = (1 + e^{-Z_i}) = 1$$

ดังนั้น

$$e^{-Z_i} = \frac{P_i}{(1 - P_i)}$$

take log เข้าไปทั้งสองข้าง

$$Z_i = (\alpha + \beta X_i) = \log \left[\frac{P_i}{1 - P_i} \right]$$

sample observation ได้แก่ $Y_1, Y_2, Y_3, \ldots, Y_N$ โดยที่ Y_i ต้องมีการแจกแจงแบบปกติ หากทางเลือกที่ 1 ถูกเลือก n_1 ครั้งและทางเลือกที่ 2 ถูกเลือก n_2 ครั้งและ($n_1 + n_2 = N$) จะสามารถกำหนด Likelihood function ได้เป็น

$$L = \text{Prob}(Y_1)\,\text{Prob}(Y_2) \ldots\ldots\ldots \text{Prob}(Y_N)$$

ค่าความน่าจะเป็นของการเลือกทางเลือกที่สองเท่ากับ 1 หักด้วยค่าความน่าจะเป็นของการเลือกทางเลือกที่หนึ่งดังสมการหรือหากกำหนด L คือ Likelihood Function

$$\log L = \sum_{i=1}^{n_1} \log P_i + \sum_{i=n_1+1}^{N} \log(1 - P_i)$$

ตัวแปรตามจะขึ้นอยู่กับ Logarithm ของอัตราส่วนระหว่างความน่าจะเป็นที่จะเกิดผลสำเร็จต่อความน่าจะเป็นที่จะไม่ก่อให้เกิดผลสำเร็จ(odds)

เนื่องจากการที่ประชากรมีความแตกต่างกันทำให้กลุ่มตัวอย่างบางกลุ่มมีความแตกต่างกันในทางสถิติจะต้องมีกลุ่มตัวอย่างที่มีคุณสมบัติเหมือนกันดังนั้นแบบจำลอง logit จึงต้องมีการประมาณค่ากลุ่มตัวอย่าง(Maximum-Likelihood Estimation) ซึ่งทำให้ parameter estimations ของ α

และ β เหมือนการเลือกตัวอย่างที่ต้องการ ทำการ differentiate ที่ log-likehood function และทำการ maximize เพื่อให้ได้ค่าเพื่อให้ได้ α^* และ β^*

$$\frac{\partial(\log L)}{\partial\alpha} = \sum_{i=1}^{n} \frac{\partial P_i / \partial\alpha}{P_i} - \sum_{i=n_1+1}^{N} \frac{\partial P_i / \partial\alpha}{(1-P_i)} = 0$$

$$\frac{\partial(\log L)}{\partial\beta} = \sum_{i=1}^{n} \frac{\partial P_i / \partial\beta}{P_i} - \sum_{i=n_1+1}^{N} \frac{\partial P_i / \partial\beta}{(1-P_i)} = 0$$

การประมาณค่าโดยกระบวนการ maximum-likelihood จะมีผลให้ค่าพารามิเตอร์ consistent และ efficient asymptotically คือค่าพารามิเตอร์ทุกตัวมีการกระจายแบบปกติและทำให้สามารถทดสอบระดับนัยสำคัญทางสถิติ เช่น สถิติ t หรือ F ได้ และเมื่อได้ค่าพารามิเตอร์ α^* และ β^* จากการ maximize และได้สมการ logit แล้ว จึงทำการทดสอบความมีนัยสำคัญของค่าสัมประสิทธิ์แต่ละตัวในแบบจำลอง โดยทดสอบสมมุติฐานว่าค่า β มีค่าเท่ากับ 0 หรือไม่โดยใช้เครื่องมือ The Likelihood Ratio Test (λ) จากนั้นจึงทำการทดสอบความเหมาะสมของแบบจำลอง(goodness of fit) จากค่า The log-likelihood index : ρ เหมือนกับการทดสอบโดยวิธี probit

เครื่องมือสุดท้ายคือ Tobit Model เป็นแบบจำลองเชิงคุณภาพเพื่อใช้งานกรณีที่ตัวแปรตามมี 2 ทางเลือก(binary-choice models) ถูกพัฒนาขึ้นมาโดย James Tobin มีรูปแบบเป็น

$$Y_i^* = \alpha + \beta X_i^* + \varepsilon_i^*$$

ข้อมูลจากค่าสังเกตุทั้งหมดจะเป็นตัวแบ่งแยกค่า Y_i ออกเป็น 2 ทางเลือกโดยทางเลือกแรกเป็นกลุ่มที่มีค่า Y_i สูงและทางเลือกที่สองเป็นกลุ่มที่มีค่า Y_i ต่ำ หากสมมุติให้ค่า Y^*_i เป็นตัวแทนของ Y_i ที่มีค่ามากกว่า 0 เมื่อเลือกทางเลือกแรกและเท่ากับ 0 เมื่อเลือกทางเลือกที่สอง และกำหนดให้ค่า 0 แทนค่าวิกฤตที่เป็นจุดตัดซึ่งจะสามารถอธิบายความสัมพันธ์ระหว่างค่า Y_i และค่า Y^*_i ได้ดังนี้

$$Y = Y_i^* \rightarrow Y_i^* > 0$$
$$Y = 0 \rightarrow Y_i^* \leq 0$$

ดังนั้นสมการที่แท้จริงสามารถเขียนได้เป็น

$$Y_i = \alpha + \beta X_i + \varepsilon_i$$

จากการที่แบบจำลอง tobit มีการประมาณค่าโดยวิธีกำลังสองน้อยที่สุด(ordinary least-squares) ดังนั้นเมื่อค่า $Y_i \geq 0$ ก็จะมีผลให้ $\varepsilon_i \geq -\alpha - \beta X_i$ ดังนั้นค่าพารามิเตอร์ α และ β ของแบบจำลองจึงไม่ consistent และ biased

James Heckman กล่าวถึงวิธีการทำให้ค่าพารามิเตอร์ของ tobit เกิดการ consistent ว่าต้องมีกระบวนการแก้ไข 2 ขั้นตอน โดยขั้นตอนแรกต้องประมาณค่า λ_i จากแบบจำลอง probit

$$P_i = F(\alpha + \beta X_i) = F(Y_i^*)$$

โดยทั่วไปแบบจำลอง probit มีการประมาณค่าโดยวิธี maximum-likelihood ซึ่งสามารถแบ่งแยก $Y_i^* > 0$ ออกจาก $Y_i^* \leq 0$ ทำให้สามารถคำนวณหาค่า $\hat{\lambda}_i$ (hazard rate) ซึ่งช่วยในการกระจายค่าเฉลี่ย(mean)ของ error term เพื่อให้ค่าพารามิเตอร์ที่ประมาณได้เกิดการ consistent

ขั้นตอนที่สองทำการเพิ่มค่า $\hat{\lambda}_i$ เข้าไปในแบบจำลองในรูปแบบตัวแปรอิสระดังสมการ

$$Y_i = \alpha + \beta X_i + \sigma \hat{\lambda}_i + u_i$$

ค่า $\hat{\lambda}_i$ ที่เพิ่มเข้าไปทำให้ค่า λ_i เป็นกลุ่มตัวอย่างขนาดใหญ่และการค่าเฉลี่ยกระจายของ error term เท่ากับ 0 ดังนั้นการประมาณค่าพารามิเตอร์ α และ β ของแบบจำลองจึง consistent อย่างไรก็ตามแบบจำลอง tobit มีข้อจำกัดเกิดขึ้นในกรณีที่ค่า $Y = 0$ ค่าความแปรปรวนของ error term จะมีความสัมพันธ์กับตัวแปรอิสระ X_i เกิดปัญหา heteroscedastics ทั้งที่ได้ผ่านกระบวนการแก้ไขทั้งสองขั้นตอนแล้วทำให้การทดสอบแบบจำลองด้วยสถิติ t เกิดการ biased

การวิจัยเพื่อทำนายภาวะการนำไปสู่การล้มละลายของระบบธนาคารพาณิชย์ครั้งนี้ตัวแปรตามเป็นตัวแปรเชิงคุณภาพที่ไม่มีความต่อเนื่อง จึงใช้เครื่องมือ Multiple discriminant analysis สำหรับการทดสอบทางสถิติแบบไม่ใช้พารามิเตอร์(nonparametric Test) เนื่องจากมีความเหมาะสม

มากที่สุดในกลุ่ม และเพื่อให้การพยากรณ์มีความสามารถในการทดสอบ(Power of Test)ขึ้นจึงใช้เครื่องมือ logit สำหรับการทดสอบทางสถิติแบบใช้พารามิเตอร์(parametric Test)เนื่องจากแบบจำลองประมาณค่าจากการสะสมความน่าจะเป็นของ logit (comulative logistic probability function) และมีการเปลี่ยนรูปโดยใช้ maximum likelihood ทำให้สถานะแบบจำลอง consistent และ unbiased มากกว่าแบบจำลอง probit และ linear probability ในขณะที่แบบจำลอง tobit มีข้อจำกัดจากการเกิดปัญหา biased ทั้งนี้การวิจัยใช้โปรแกรมสำเร็จรูป SPSSPC ในการวิเคราะห์ข้อมูล

มหาวิทยาลัยเชียงใหม่ Chiang Mai University