บทที่ 3

วิธีดำเนินการวิจัย

ในการศึกษาเพื่อพิจารณาฟังก์ชันการคงอยู่ ค่าเฉลี่ย และค่ามัธยฐานจำนวนปีการคงอยู่ของกรมธรรม์ประกันชีวิตเมื่อพิจารณาเฉพาะกรมธรรม์ที่สิ้นผลบังคับลงก่อนกำหนด รวมทั้งเปรียบเทียบฟังก์ชันการคงอยู่ และวิเคราะห์ปัจจัยที่มีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ วิธีดำเนินการศึกษาจะประกอบไปด้วย การเก็บรวบรวมข้อมูลเกี่ยวกับกรมธรรม์ประกันชีวิตเฉพาะที่สิ้นผลบังคับลงก่อนกำหนดในรอบปีพ.ศ. 2544 และปัจจัยที่คาดว่าจะมีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิต ศึกษาหาค่าประมาณฟังก์ชันการคงอยู่ และเปรียบเทียบฟังก์ชันการคงอยู่ในแต่ละปัจจัยที่คาดว่ามีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด ศึกษาตัวแบบเชิงพหุของการวิเคราะห์ข้อมูลการคงอยู่เพื่อหาค่าประมาณพารามิเตอร์ จนถึงการตรวจสอบความเหมาะสมของตัวแบบการคงอยู่ที่ได้

3.1 การเก็บรวบรวมข้อมูล

3.1.1 รวบรวมข้อมูลกรมธรรม์ประกันชีวิตประเภทสามัญเฉพาะที่สิ้นผลบังคับลงก่อนกำหนดในรอบปีพ.ศ. 2544 และปัจจัยที่คาดว่าจะมีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิต

ก. รวบรวมข้อมูลเกี่ยวกับกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด* ในรอบปีพ.ศ. 2544 เนื่องมาจากการไม่ชำระเบี้ยประกันภัย ซึ่งจะรวมถึงกรมธรรม์ที่สิ้นผลบังคับลงอันเนื่องมาจาก การขาดอายุ การเวนคืน การเปลี่ยนเป็นกรมธรรม์ใช้เงินสำเร็จ หรือการเปลี่ยนเป็นกรมธรรม์ขยายเวลา เว้นแต่ กรมธรรม์ที่สิ้นผลบังคับลงอันเนื่องมาจากการแปลงสัญญา การยกเลิก การตาย การครบกำหนด รวมทั้งกรมธรรม์ที่มีการชำระเบี้ยประกันภัยครั้งเดียว

ข. รวบรวมปัจจัยที่คาดว่าจะมีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิต โดยเป็นการเก็บรวบรวมข้อมูลจากหนังสือ วารสาร และเอกสารอ้างอิงต่างๆ

---

* การเก็บรวบรวมเฉพาะข้อมูลกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนดเป็นที่เรียบร้อย โดยไม่มีการนำเอากรมธรรม์ประกันชีวิตที่ยังคงมีผลบังคับอยู่มาร่วมพิจารณาด้วย ข้อมูลที่ใช้ในการวิจัยครั้งนี้ จึงมีแต่เพียงข้อมูลที่มีค่าสังเกตสมบูรณ์แล้วทั้งสิ้น

3.2 ตัวแปรที่ใช้ในการวิจัย

จากงานวิจัยที่เกี่ยวข้องในบทที่ 2 ทำให้ทราบถึงปัจจัยต่างๆ ที่มีผลกระทบต่อระยะเวลาการสิ้นผลบังคับลงก่อนกำหนดของกรมธรรม์ประกันชีวิต แต่เนื่องจากลักษณะการดำรงชีวิต ทัศนคติ รสนิยม วัฒนธรรม และปัจจัยภายนอกอื่นๆ ของแต่ละภูมิภาค จึงอาจส่งผลต่อพฤติกรรมของผู้เอาประกันภัยให้มีความแตกต่างกัน ในส่วนนี้ จึงได้เพิ่มตัวแปรที่ใช้ในการวิจัยมากขึ้นกว่าผลงานวิจัยที่ได้กล่าวถึง จากการรวบรวมตัวแปรต่างๆ ทั้งทางด้านผู้เอาประกันภัย และบริษัทประกันชีวิตในประเทศไทย เพื่อให้เหมาะสมต่อสภาพแวดล้อม และการดำเนินชีวิตที่ต่างกันออกไป โดยตัวแปรที่ใช้ในการวิจัยครั้งนี้ คือ

1. อายุที่เริ่มทำประกันชีวิต (Age)* กลุ่มตัวอย่างที่อายุต่างกัน จะมีพฤติกรรมการตัดสินใจทำประกันชีวิตที่แตกต่างกัน ตามแนวคิด และทฤษฎีเกี่ยวกับลักษณะทางประชากร ที่ว่าอายุเป็นตัวกำหนดความเหมือนและความแตกต่าง ในเรื่องความคิด พฤติกรรม และความยากง่ายในการชักจูง ทำให้ทัศนะคติและความรู้สึกนึกคิด ตลอดจนความต้องการในสิ่งต่างๆ ของคนต่างรุ่นต่างวัยไม่เหมือนกัน เช่น วัยกลางคนและคนสูงวัย มักคิดถึงเรื่องความปลอดภัยในชีวิตและทรัพย์สิน การรักษาพยาบาล การมีบ้านและที่ดินเป็นของตัวเอง (ปรมะ สตะเวทิน, 2538: 112-113)

การวิจัยครั้งนี้ จะแบ่งตามลักษณะอายุของผู้เอาประกันภัยในบริษัทประกันชีวิตที่นำมาศึกษาและโครงสร้างตามลักษณะประชากร ซึ่งแบ่งออกเป็น 4 ช่วง คือ

- อายุตั้งแต่ 0-20 ปี
- อายุตั้งแต่ 21-30 ปี
- อายุตั้งแต่ 31-40 ปี
- อายุตั้งแต่ 40 ปีขึ้นไป

2. เพศผู้เอาประกันภัย (Sex) ชายและหญิงจะมีการตัดสินใจทำประกันชีวิตที่แตกต่างกัน โดยเพศชายจะมีแนวโน้มในการตัดสินใจทำประกันชีวิตมากกว่าเพศหญิง อาจอธิบายได้ว่า เพศที่ต่างกันจะมีความแตกต่างทางด้านความคิด ค่านิยม ทัศนคติ และความต้องการ ตามที่สังคมและวัฒนธรรมได้กำหนดบทบาทและกิจกรรมของคนทั้งสองเพศไว้ต่างกัน (เพ็ญจมาส ศิริกิจวัฒนา, 2543: 126)

- ชาย
- หญิง

---

* โดยนับอายุใกล้วันครบรอบวันเกิดที่สุด (age nearest birthday)

3. แบบประกันชีวิต (Plan of Life Insurance) ปัจจัยสำคัญที่จะเป็นเหตุเอื้ออำนวยให้เกิดการตัดสินใจซื้อกรมธรรม์ประกันชีวิตของแต่ละบุคคลนั้น ย่อมแตกต่างไปตามสถานภาพของบุคคลนั้นๆ คือ การต้องการได้รับความคุ้มครอง หรือค่าทดแทนในรูปตัวเงินที่จะเกิดขึ้นในอนาคต แบบประกันชีวิตในแต่ละแบบ จึงมีลักษณะเงื่อนไขความคุ้มครอง ผลประโยชน์จากการออมทรัพย์ รวมถึงเบี้ยประกันภัยที่จะต้องชำระในแต่ละงวดแตกต่างกัน ลักษณะและเงื่อนไขที่แตกต่างกันนี้ จึงส่งผลถึงพฤติกรรมการเลือกซื้อกรมธรรม์ประกันชีวิต และแรงจูงใจในการชำระเบี้ยประกันภัยของผู้เอาประกันภัย (สากล ธนสัตยาวิบูล, 2538) แบบกรมธรรม์ประกันชีวิตโดยทั่วไป แบ่งออกเป็น 4 แบบ คือ

- แบบตลอดชีพ (Whole-Life Insurance)
- แบบชั่วระยะเวลา (Term Insurance)
- แบบสะสมทรัพย์ (Endowment Insurance)
- แบบบำนาญ (Annuity Insurance)

4. งวดการชำระเบี้ยประกันภัย (Mode of Payment) มีทั้งกรณีชำระเบี้ยประกันภัยเพียงครั้งเดียว (งวดเดียวสำหรับการคุ้มครองตามสัญญาประกันชีวิต) หรือผ่อนชำระเป็นรายงวด งวดละเท่าๆ กัน โดยอาจชำระเป็นงวดรายปี ราย 6 เดือน ราย 3 เดือน หรือรายเดือน ซึ่งโดยเปรียบเทียบแล้ว การชำระเบี้ยประกันภัยที่ผ่อนชำระเป็นรายงวดบ่อยครั้งมากขึ้น จำนวนเบี้ยประกันภัยที่จะต้องชำระต่องวดการจ่าย ก็จะเป็นจำนวนเงินที่ลดน้อยลงตามความถี่ของการชำระเบี้ยประกันภัยเช่นเดียวกัน ทั้งนี้ก็เพื่อให้สนองตอบกับความต้องการของผู้เอาประกันภัยบางคน ที่ไม่สามารถชำระเบี้ยประกันภัยเป็นจำนวนเงินมากๆ ได้ (สุธรรม พงศ์สำราญ และคณะ 2542: 50) โดยทั่วไปมีการแบ่งงวดการชำระเบี้ยประกันภัยออกเป็น 4 งวด คือ

- รายเดือน (Monthly)
- ราย 3 เดือน (Quarterly)
- ราย 6 เดือน (Semi Annually)
- รายปี (Annually)

5. จำนวนเงินเอาประกันภัย (Sum Insured) ในการคำนวณเบี้ยประกันภัย ปัจจัยในเรื่องของจำนวนเงินเอาประกันภัย จะเป็นส่วนสำคัญต่อเบี้ยประกันภัยว่าจะสูงหรือต่ำอย่างไร มีกรมธรรม์ประกันชีวิตจำนวนไม่น้อยที่สิ้นผลบังคับลง อันเนื่องมาจากผู้เอาประกันภัยประสบปัญหาทางด้านการเงิน ทำให้ไม่สามารถชำระเบี้ยประกันภัยต่อไปได้ ดังนั้น การลดจำนวนเงินเอาประกันภัยลง ก็จะช่วยทำให้ภาระในการชำระเบี้ยประกันภัยลดจำนวนลงไปด้วย (พรรณิภา วรกัลป์, 2542: 102)

การวิจัยครั้งนี้ จะแบ่งจำนวนเงินเอาประกันภัยออกเป็น 3 ช่วง โดยพิจารณาตามฐานนิยมของจำนวนเงินเอาประกันภัยที่มีความถี่สูงสุด คือ

- จำนวนเงินเอาประกันภัยต่ำกว่า 100,000 บาท
- จำนวนเงินเอาประกันภัยเท่ากับ 100,000 บาท
- จำนวนเงินเอาประกันภัยมากกว่า 100,000 บาท

6. ระยะเวลาเอาประกันภัย (Period Insured) ลักษณะของกรมธรรม์ประกันชีวิตที่แตกต่างกัน ส่งผลให้ระยะเวลาคุ้มครองของสัญญาแตกต่างกัน การเลือกความคุ้มครองของผู้เอาประกันภัยโดยส่วนใหญ่ จึงมักเลือกจากระยะเวลาการคุ้มครองที่เหมาะสมกับตัวเอง สามารถให้ความคุ้มครองต่อครอบครัว และให้ประโยชน์ในด้านการออมทรัพย์ การลงทุน หรือมีรายได้ยามเจ็บป่วยจากการทุพพลภาพ หรือชราภาพ ระยะเวลาเอาประกันภัยในแผนความคุ้มครองต่างๆ จึงมีผลต่อการตัดสินใจเลือกซื้อกรมธรรม์ประกันชีวิต และการต่ออายุสัญญากรมธรรม์ประกันชีวิต เนื่องจากบางครั้งความต้องการในเรื่องของความคุ้มครองหรือการออมทรัพย์ อาจมีการเปลี่ยนแปลงไม่เป็นไปตามที่เลือกไว้ในครั้งแรก (สากล ธนสัตยาวิบูล, 2538)

การวิจัยครั้งนี้ จึงแบ่งระยะเวลาเอาประกันภัยออกตามแผนความคุ้มครอง ที่ผู้เอาประกันภัยของบริษัทประกันชีวิตที่ทำการศึกษานิยมเลือกความคุ้มครอง คือ

- ระยะเวลาเอาประกันภัยไม่เกิน 20 ปี
- ระยะเวลาเอาประกันภัยตั้งแต่ 21-35 ปี
- ระยะเวลาเอาประกันภัยตั้งแต่ 36-55 ปี
- ระยะเวลาเอาประกันภัยตั้งแต่ 55 ปีขึ้นไป

7. ระยะเวลาชำระเบี้ยประกันภัย (Period of Payment) ถ้าระยะเวลาของการส่งเบี้ยประกันภัยมีระยะเวลานาน ผู้ที่ถือกรมธรรม์ประกันชีวิตอยู่ จึงอาจต้องเผชิญกับการตัดสินใจว่าจะต่ออายุสัญญากรมธรรม์ดีหรือไม่ โดยพิจารณาจากระยะเวลาที่ต้องส่งเบี้ยประกันภัยต่อไปในอนาคต ถ้าระยะเวลาชำระเบี้ยประกันภัยในอนาคตยังอีกยาวนาน ก็อาจส่งผลให้ผู้ถือกรมธรรม์ตัดสินใจที่จะหยุดส่งเบี้ยประกันภัย (โสภา คงธนาคมธัญกิจ, 2538) การวิจัยในครั้งนี้ จะแบ่งระยะเวลาชำระเบี้ยประกันภัยออกเป็นช่วงตามความเหมาะสมของข้อมูล คือ

- ระยะเวลาชำระเบี้ยประกันภัยไม่เกิน 20 ปี
- ระยะเวลาชำระเบี้ยประกันภัยมากกว่า 20 ปี

8. ประเภทกลุ่มความเสี่ยง (Classification of Risk) ในการกำหนดอัตราเบี้ยประกันภัย บริษัทจะพิจารณาถึงขนาดความเสี่ยงภัย ในส่วนที่เกี่ยวกับลักษณะโดยทั่วไปของผู้เอาประกันภัย และสภาพสิ่งแวดล้อมในสังคม ดังนั้น การที่บริษัทจะรับทำการเสี่ยงภัยรายใด จึงต้องมีการคัดเลือกภัยว่าควรจัดอยู่ในภัยประเภทกลุ่มความเสี่ยงใด ซึ่งบริษัทอาจรับประกันชีวิตรายนั้นไว้ในอัตราเบี้ยประกันภัยธรรมดาสำหรับภัยประเภทกลุ่มความเสี่ยงนั้น หรือรับประกันภัยโดยเพิ่มอัตราเบี้ยประกันภัยขึ้นเป็นพิเศษ กรณีที่เป็นภัยประเภทต่ำกว่ามาตรฐาน (สุธรรม พงศ์สำราญ และคณะ, 2542)

ผู้วิจัยคาดว่า ผู้เอาประกันภัยที่จัดอยู่ในภัยประเภทต่ำกว่ามาตรฐาน ต้องมีการชำระเบี้ยประกันภัยที่สูงขึ้น จึงน่าจะมีอิทธิพลต่อการสิ้นผลบังคับลงของกรมธรรม์ประกันชีวิตสูงกว่าผู้เอาประกันภัยที่จัดอยู่ในภัยประเภทปกติ หรือในทางตรงกันข้าม ผู้เอาประกันภัยที่จัดอยู่ในภัยประเภทต่ำกว่ามาตรฐาน ซึ่งโดยส่วนใหญ่จะกำหนดมาจากสุขภาพของผู้เอาประกันภัย ที่อาจมีผลต่อการมีชีวิตอยู่ของบุคคลเหล่านั้น ทำให้ผู้เอาประกันภัยรู้ถึงความไม่แน่นอนที่อาจเกิดขึ้นกับชีวิตของตนเอง มีความสนใจที่จะต่ออายุสัญญากรมธรรม์ประกันชีวิต มากกว่าผู้เอาประกันภัยที่ไม่มีปัญหาในเรื่องของสุขภาพ การวิจัยครั้งนี้ จึงแบ่งประเภทกลุ่มความเสี่ยงของผู้เอาประกันภัยออกเป็น

- กลุ่มความเสี่ยงปกติ (Standard Risk) เกณฑ์ที่บริษัทประกันชีวิตกำหนดขึ้นให้เป็นการเสี่ยงปกติ โดยไม่ต้องชำระเบี้ยประกันภัยเพิ่มเติม หรือมีเงื่อนไขจำกัดอื่นใด
- กลุ่มความเสี่ยงสูง (Substandard Risk) ผู้เอาประกันภัยรายใดมีร่างกายไม่สมบูรณ์ มีอาชีพเสี่ยงอันตรายมาก หรืออยู่ใกล้กับสิ่งแวดล้อมไม่ดี ก็จะถือว่าภัยดังกล่าวมีความเสี่ยงสูงกว่าปกติ ซึ่งจะต้องเพิ่มเบี้ยประกันภัย

9. สัญญาเพิ่มเติม (Rider) ในการทำประกันชีวิต ผู้เอาประกันภัยสามารถซื้อความคุ้มครองเพิ่มเติมแนบท้ายกรมธรรม์ประกันชีวิตได้ เช่น สัญญาเพิ่มเติมการประกันอุบัติเหตุ การทุพพลภาพ ค่ารักษาพยาบาล ค่าชดเชยรายวันในการเข้ารับการรักษาพยาบาล การได้รับยกเว้นการชำระเบี้ยประกันภัย เป็นต้น สัญญาเพิ่มเติมกรมธรรม์ประกันชีวิตดังกล่าว โดยส่วนมากจะมีระยะเวลาในการให้ความคุ้มครองเพียงหนึ่งปี และบริษัทจะต่ออายุสัญญาให้ปีต่อปี ดังนั้น เมื่อผู้เอาประกันภัยได้ทำประกันชีวิตไว้โดยได้ซื้อสัญญาเพิ่มเติมต่างๆ ไว้ด้วย เบี้ยประกันภัยที่จะต้องชำระในแต่ละงวดก็จะเพิ่มสูงขึ้น (พรรณิภา วรกัลป์, 2542: 102) ในที่นี้พิจารณาออกเป็น

- มีสัญญาเพิ่มเติม
- ไม่มีสัญญาเพิ่มเติม

10. สาขาของบริษัทประกันชีวิต (Distribution Channels) กรมธรรม์ประกันชีวิตมีหลายประเภท หลายแบบ การมีสาขาของบริษัทประกันชีวิตเพิ่มขึ้น จึงเป็นการอำนวยความสะดวกต่อผู้เอาประกันภัยในกรณีที่ต้องการติดต่อกับบริษัท หรือมีปัญหาที่จะสอบถาม โดยไม่ต้องติดต่อกับสำนักงานใหญ่โดยตรง ทำให้ผู้เอาประกันภัยได้รับความสะดวกสบาย นอกจากนี้ สาขาของบริษัทประกันชีวิตจะมีความแตกต่างในเรื่องของการให้บริการ ลักษณะการขาย การควบคุมดูแลของตัวแทนประกันชีวิตตามลักษณะพื้นที่ต่างๆ ที่แตกต่างกันไป (ศรีสุกิจ อำนาจวรประเสริฐ, 2529)

ผู้วิจัยคาดว่า สาขาของบริษัทประกันชีวิตในพื้นที่ต่างๆ จึงน่าจะส่งผลต่อการสิ้นผลบังคับลงของกรมธรรม์ประกันชีวิต จากการให้บริการและความเอาใจใส่ต่อผู้เอาประกันภัย การวิจัยในครั้งนี้ จะแยกพิจารณากรุงเทพมหานครออกจากพื้นที่ภาคกลาง คือ

- กรุงเทพมหานคร
- ภาคกลาง (ไม่รวมกรุงเทพมหานคร)
- ภาคตะวันออกเฉียงเหนือ
- ภาคเหนือ
- ภาคใต้

3.2.1 การกำหนดตัวแปรสำหรับการพิจารณาตัวแบบเชิงพหุ

age หมายถึง อายุที่เริ่มทำประกันชีวิต (ปี)*

sex หมายถึง เพศผู้เอาประกันภัย กำหนดให้

ชาย	สัญลักษณ์เป็น 1
หญิง	สัญลักษณ์เป็น 0

plan_2 หมายถึง กรมธรรม์ประกันชีวิตแบบชั่วระยะเวลา กำหนดให้

แบบชั่วระยะเวลา	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

plan_3 หมายถึง กรมธรรม์ประกันชีวิตแบบสะสมทรัพย์ กำหนดให้

แบบสะสมทรัพย์	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

plan_4 หมายถึง กรมธรรม์ประกันชีวิตแบบบำนาญ กำหนดให้

แบบบำนาญ	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

mode_2 หมายถึง งวดการชำระเบี้ยประกันภัย ราย 3 เดือน กำหนดให้

ราย 3 เดือน	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

mode_3 หมายถึง งวดการชำระเบี้ยประกันภัย ราย 6 เดือน กำหนดให้

ราย 6 เดือน	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

mode_4 หมายถึง งวดการชำระเบี้ยประกันภัย รายปี กำหนดให้

รายปี	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

---

* การกำหนดตัวแปรในที่นี้ จะใช้สำหรับการพิจารณาเฉพาะตัวแบบเชิงพหุ ตัวแปรอิสระที่เป็นตัวแปรแบบต่อเนื่อง (Continueous vaiable) เช่น อายุที่เริ่มทำประกันชีวิต และจำนวนเงินเอาประกันภัย การวิจัยในส่วนนี้ จะพิจารณาตามลักษณะของตัวแปรแบบต่อเนื่อง

sum หมายถึง จำนวนเงินเอาประกันภัย (หน่วย : หมื่นบาท)

period_2 หมายถึง ระยะเวลาเอาประกันภัยในช่วง 21-35 ปี กำหนดให้

ช่วงตั้งแต่ 21-35 ปี	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

period_3 หมายถึง ระยะเวลาเอาประกันภัยในช่วง 36-55 ปี กำหนดให้

ช่วงตั้งแต่ 36-55 ปี	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

period_4 หมายถึง ระยะเวลาเอาประกันภัยตั้งแต่ 55 ปีขึ้นไป กำหนดให้

ช่วงตั้งแต่ 55 ปีขึ้นไป	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

pay หมายถึง ระยะเวลาชำระเบี้ยประกันภัย กำหนดให้

มากกว่า 20 ปี	สัญลักษณ์เป็น 1
ไม่เกิน 20 ปี	สัญลักษณ์เป็น 0

risk หมายถึง ประเภทกลุ่มความเสี่ยง กำหนดให้

กลุ่มความเสี่ยงสูง	สัญลักษณ์เป็น 1
กลุ่มความเสี่ยงปกติ	สัญลักษณ์เป็น 0

rider หมายถึง สัญญาเพิ่มเติม กำหนดให้

มี	สัญลักษณ์เป็น 1
ไม่มี	สัญลักษณ์เป็น 0

channel_2 หมายถึง สาขาบริษัทประกันชีวิตในภาคกลาง (ไม่รวมกรุงเทพฯ) กำหนดให้

ภาคกลาง (ไม่รวมกรุงเทพฯ)	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

channel_3 หมายถึง สาขาบริษัทประกันชีวิตในภาคตะวันออกเฉียงเหนือ กำหนดให้

ภาคตะวันออกเฉียงเหนือ	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

channel_4 หมายถึง สาขาบริษัทประกันชีวิตในภาคเหนือ กำหนดให้

ภาคเหนือ	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

channel_5 หมายถึง สาขาบริษัทประกันชีวิตในภาคใต้ กำหนดให้

ภาคใต้	สัญลักษณ์เป็น 1
อื่นๆ	สัญลักษณ์เป็น 0

survtime หมายถึง ระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิต (ปี)

status* หมายถึง สถานะของกรมธรรม์ประกันชีวิต โดยกำหนดให้มีค่าเท่ากับ 1

ยกตัวอย่างการลงรหัสตัวแปรสำหรับการวิจัย** เช่น กรมธรรม์ที่สิ้นผลบังคับลงก่อนกำหนดในปีพ.ศ. 2544 ประกอบไปด้วย ผู้ถือกรมธรรม์ประกันชีวิตเพศหญิง อายุ 24 ปี ที่ได้ทำกรมธรรม์ประกันชีวิตแบบสะสมทรัพย์ไว้กับสาขาของบริษัทประกันชีวิตในจังหวัดสระแก้ว โดยเริ่มทำประกันชีวิตเมื่อวันที่ 1 เมษายน พ.ศ. 2541 และได้สิ้นผลบังคับลงเนื่องจากการขาดส่งเบี้ยประกันภัยไปเมื่อวันที่ 18 พฤศจิกายน พ.ศ. 2544 รวมระยะเวลาที่ได้รับความคุ้มครองกับบริษัทมาแล้วเป็นเวลา 3.64 ปี รายละเอียดข้อมูลเกี่ยวกับผู้เอาประกันภัยรายนี้ คือ ได้ทำประกันชีวิตไว้เป็นจำนวนเงินเอาประกันภัย 200,000 บาท ระยะเวลาเอาประกันภัย 20 ปี ชำระเบี้ยประกันภัย 20 ปี โดยแบ่งชำระเบี้ยประกันภัยออกเป็นราย 6 เดือน ผู้เอาประกันภัยรายนี้ถูกจัดอยู่ในกลุ่มประเภทความเสี่ยงสูง พร้อมทั้งได้ซื้อสัญญาเพิ่มเติมแนบท้ายกรมธรรม์หลัก การลงรหัสในตัวแปรจะเป็นดังต่อไปนี้

ตารางที่ 1 การลงรหัสตัวแปรสำหรับการวิจัย

age	sex	plan_2	plan_3	plan_4	mode_2	mode_3
24	0	0	1	0	0	1

mode_4	sum	period_2	period_3	period_4	pay	risk
0	20	0	0	0	0	1

rider	channel_2	channel_3	channel_4	channel_5	survtime	status
1	1	0	0	0	3.64	1

* เนื่องจากข้อมูลที่ใช้ในการวิจัยครั้งนี้ เป็นข้อมูลที่มีความสมบูรณ์ทั้งสิ้น คือ พิจารณาเฉพาะแต่เพียงกรมธรรม์ที่สิ้นผลบังคับลงก่อนกำหนด จึงทราบดีอยู่แล้วว่าทุกกรมธรรม์ที่นำมาศึกษาทั้งหมดนี้ ได้ประสบกับเหตุการณ์ที่สนใจเป็นที่เรียบร้อย สถานะของกรมธรรม์ที่ใช้ในการวิจัย จึงถูกกำหนดให้มีค่าเท่ากับ 1 ทุกค่าสังเกต (สำหรับการวิจัยที่มีค่าสังเกตไม่สมบูรณ์ หน่วยสังเกตนั้นจะถูกกำหนดให้มีค่าเท่ากับ 0) เพื่อการวิเคราะห์ในโปรแกรมสำเร็จรูป SAS

** ตัวอย่างข้อมูลที่ได้จากบริษัทประกันชีวิตอยู่ในภาคผนวก ก. หน้า 121.

3.3 การวิเคราะห์ข้อมูล

จากข้อมูลจำนวนกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนดในรอบปีพ.ศ. 2544 เป็นข้อมูลที่มีการเก็บเฉพาะค่าสังเกตที่ประสบกับเหตุการณ์ที่สนใจ คือ กรมธรรม์ที่ตกอยู่ในกลุ่มการศึกษาครั้งนี้ เป็นกรมธรรม์เฉพาะที่สิ้นผลบังคับลงก่อนกำหนดแล้วทุกกรมธรรม์ และทราบระยะเวลาการคงอยู่ของกรมธรรม์เป็นที่แน่นอน โดยไม่ได้พิจารณากรมธรรม์อื่นนอกเหนือภายใต้การศึกษาด้วย ข้อมูลที่ใช้ในการวิจัยครั้งนี้ จึงเป็นข้อมูลที่ค่าสังเกตมีความสมบูรณ์ทั้งสิ้น และสามารถนำมาใช้ได้กับวิธีการทางสถิติที่ได้กล่าวถึงในบทที่ 2

3.3.1 การประมาณฟังก์ชันการคงอยู่

การประมาณฟังก์ชันการคงอยู่ (Survival function) ฟังก์ชันความหนาแน่น (Density function) และฟังก์ชันภาวะภัย (Hazard function) สำหรับค่าสังเกตสมบูรณ์และไม่สมบูรณ์นั้น เวลาการคงอยู่ส่วนใหญ่มักไม่ทราบการแจกแจง ดังนั้น การประมาณด้วยตัวสถิติไม่อิงพารามิเตอร์ จึงสามารถนำมาประยุกต์ใช้ และเข้าใจได้ง่ายในการวิเคราะห์การคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด ก่อนที่จะทำการศึกษาถึงตัวแบบที่เหมาะสมของข้อมูลในส่วนต่อไป

3.3.1.1 วิธีลิมิตผลคูณ (Product-Limit Method)

การประมาณด้วยวิธีลิมิตผลคูณ จะได้ค่าประมาณดังนี้

ก) ค่าประมาณฟังก์ชันการคงอยู่

ข) ค่าเฉลี่ยจำนวนปีการคงอยู่ของกรมธรรม์ที่สิ้นผลบังคับลงก่อนกำหนด

ค) ค่ามัธยฐานจำนวนปีการคงอยู่ของกรมธรรม์ที่สิ้นผลบังคับลงก่อนกำหนด

3.3.2 เปรียบเทียบฟังก์ชันการคงอยู่ในแต่ละปัจจัยที่คาดว่าจะมีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิต

โดยทั่วไปแล้ว ตัวแปรอิสระส่วนใหญ่จะเป็นตัวแปรนามกำหนด (nominal or categorical variable) คือ มีการแบ่งกลุ่มของข้อมูล (category) ออกเป็นกลุ่มย่อยๆ* เช่น ปัจจัยในเรื่องของแบบกรมธรรม์ประกันชีวิตที่มีการแบ่งออกเป็น 4 แบบ คือ แบบตลอดชีพ แบบชั่วระยะเวลา แบบ

---

*เนื่องจากการพิจารณาเส้นโค้งการคงอยู่ และการทดสอบสมมติฐานฟังก์ชันการคงอยู่ในแต่ละกลุ่มประชากร ตัวแปรที่ใช้จะต้องเป็นตัวแปรนามกำหนด ดังนั้น กรณีตัวแปรอิสระที่เป็นตัวแปรแบบต่อเนื่อง (continueous variable) เช่น อายุที่เริ่มทำประกันชีวิต จำนวนเงินเอาประกันภัย การวิจัยในส่วนนี้ จะทำการพิจารณาออกเป็นช่วงตามที่กำหนดไว้ในส่วนที่ 3.2 แต่เมื่อทำการประมาณค่าพารามิเตอร์ด้วยตัวแบบเชิงพหุ จึงพิจารณาเป็นตัวแปรแบบต่อเนื่องดังส่วนที่ 3.2.1

สะสมทรัพย์ และแบบบำนาญ เป็นต้น จึงสามารถนำมาพิจารณาเปรียบเทียบความแตกต่างของฟังก์ชันการคงอยู่ได้ในแต่ละปัจจัย* ในที่นี้จะพิจารณา 2 วิธี คือ

ก) การตรวจสอบโดยพิจารณาจากเส้นโค้งการคงอยู่

การพิจารณาโดยดูจากเส้นโค้งการคงอยู่ แกนตั้งจะแทนฟังก์ชันการคงอยู่ และแกนนอนแทนระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด โดยภาพแสดงความสัมพันธ์ของฟังก์ชันการคงอยู่กับระยะเวลาการคงอยู่ที่ได้ จะแสดงให้เห็นถึงฟังก์ชันการคงอยู่ที่ต่ำ (เวลาการคงอยู่สั้น) หรือฟังก์ชันการคงอยู่สูง (เวลาการคงอยู่ยาว) ในการเปรียบเทียบการแจกแจงการคงอยู่ของประชากรมากกว่า หรือเท่ากับ 2 ประชากร

ข) การตรวจสอบสมมติฐาน

$$H_0 : S_1(t) = S_2(t) = \ldots = S_k(t)$$
$$H_1 : S_j(t) \neq S_g(t) \qquad \text{เมื่อ } j \neq g$$

โดย 1) วิธีทดสอบ Gehan's Generalized Wilcoxon

2) วิธีทดสอบ Log Rank

ตัวสถิติทดสอบ คือ

$$\chi^2 = (Z_1(\tau), Z_2(\tau), \ldots, Z_{k-1}(\tau)) \sum^{-1} (Z_1(\tau), Z_2(\tau), \ldots, Z_{k-1}(\tau))'$$

ซึ่งมีการแจกแจงไคสแควร์ ที่องศาความเป็นอิสระ k-1

โดยปฏิเสธสมมติฐานว่าง เมื่อระดับนัยสำคัญของการทดสอบไคสแควร์น้อยกว่าระดับนัยสำคัญที่กำหนดไว้

3.3.3 การประมาณค่าพารามิเตอร์ด้วยตัวแบบเชิงพหุ

ขั้นตอนแรกของการศึกษาตัวแบบเชิงพหุ คือ จะใช้ Cox PH model เป็นตัวแบบในการพิจารณาตัวแปรอิสระที่มีความสัมพันธ์กับเวลาการคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด โดยคุณสมบัติที่สำคัญของ Cox PH model คือ ไม่ได้มีการกำหนดรูปแบบของ $h_0(t)$ ซึ่งส่งผลให้ได้ค่าประมาณพารามิเตอร์ที่ดีของตัวแบบเชิงพหุในอัตราส่วนภาวะภัย (hazard ratio) ที่สนใจ รวมทั้งยังเป็นวิธีที่มีประสิทธิภาพ เนื่องจากค่าประมาณพารามิเตอร์ที่ได้จาก Cox PH model

---

* การลงรหัสตัวแปรสำหรับการเปรียบเทียบฟังก์ชันการคงอยู่ในระดับที่แตกต่างกันในแต่ละปัจจัยนั้น การลงรหัสตัวแปรสำหรับการวิจัยจะกำหนดเพียง 1 ตัวแปร แต่มีค่าได้เท่ากับระดับในปัจจัยนั้นๆ เช่น แบบกรมธรรม์ประกันชีวิต จะมีทั้งหมด 4 แบบ ก็จะให้ตัวแปร plan หมายถึง แบบกรมธรรม์ประกันชีวิต ซึ่งมีค่าเท่ากับ 1 สำหรับกรมธรรม์แบบตลอดชีพ ค่าเท่ากับ 2 สำหรับกรมธรรม์แบบชั่วระยะเวลา ค่าเท่ากับ 3 สำหรับกรมธรรม์แบบสะสมทรัพย์ และมีค่าเท่ากับ 4 เมื่อเป็นกรมธรรม์แบบบำนาญ เป็นต้น

จะมีค่าใกล้เคียงกับผลลัพธ์ที่ได้จากตัวแบบการถดถอยอิงพารามิเตอร์ (Parametric regression model) ที่มีการแจกแจงถูกต้องและเหมาะสมกับข้อมูลนั้น

3.3.3.1 Cox Proportional Hazard Model

ให้ $h(t|Z)$ เป็นฟังก์ชันภาวะภัยที่เวลา t สำหรับหน่วยสังเกตที่มีปัจจัยเสี่ยงเป็น Z ความสัมพันธ์ระหว่างการแจกแจงของเวลาการคงอยู่และ Z จะแสดงในรูปของฟังก์ชันภาวะภัย คือ

$$h(t|Z) = h_0(t)\exp(\sum_{k=1}^{p}\beta_k Z_k)$$

เมื่อ $h_0(t)$ เป็นอัตราภาวะภัยพื้นฐานค่าเลือก และ $\beta = (\beta_1,..., \beta_p)'$ เป็นเวกเตอร์ของพารามิเตอร์

ก) การประมาณพารามิเตอร์ด้วยวิธีความควรจะเป็นบางส่วน

เนื่องจากค่าสังเกตที่ได้จากข้อมูลกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนดในรอบปีพ.ศ. 2544 นั้น กรมธรรม์ส่วนใหญ่จะสิ้นผลบังคับลงในช่วงเวลาเดียวกัน จึงทำให้ค่าสังเกตที่ได้มามีค่าซ้ำกันมาก การวิจัยครั้งนี้ จึงจะพิจารณาเฉพาะวิธี Efron ในการประมาณค่าพารามิเตอร์ ซึ่งเหมาะสมกับกรณีที่ค่าสังเกตของข้อมูลมีค่าซ้ำกันมาก (Hosmer et al., 1998: 106) และทำการทดสอบสมมติฐานค่าประมาณพารามิเตอร์จาก $H_0 : \beta = \beta_0$ ว่าเป็นไปตามสมมติฐานว่างที่ตั้งขึ้นหรือไม่ โดยวิธี Wald's test, Likelihood ratio test และ Score test ที่มีในโปรแกรมสำเร็จรูป SAS

$$\chi^2_W = (b - \beta_0)'I(b)(b - \beta_0)$$

$$\chi^2_{LR} = 2[LL(b) - LL(\beta_0)]$$

$$\chi^2_{SC} = U(\beta_0)' I^{-1}(\beta_0)U(\beta_0)$$

โดยมีการแจกแจงแบบไคสแควร์ ที่องศาความเป็นอิสระ p ภายใต้สมมติฐานว่าง

ข) การเลือกตัวแปรอิสระเข้าตัวแบบเชิงพหุด้วยวิธีการถดถอยแบบขั้นบันได

วิธีการถดถอยแบบขั้นบันได (stepwise regression) เป็นการคัดเลือกตัวแปรอิสระ โดยใช้อัตราส่วนความควรจะเป็นเข้ามาพิจารณาในการคัดเลือกตัวแปรเข้าตัวแบบ ซึ่งเป็นการทดสอบที่ดีที่สุดใน 3 วิธีการทดสอบ คือ Wald 's test, Likelihood ratio test และ Score test (Hosmer et al., 1998: 180) นอกจากนี้ ขั้นตอนในการดำเนินการของวิธีการถดถอยแบบขั้นบันได ยังช่วยลดปัญหาเมื่อตัวแปรอิสระที่นำมาวิเคราะห์มีจำนวนมาก ลดปัญหาการนำตัวแปรอิสระเข้าตัวแบบเพื่อทดสอบทีละตัว โดยตัวแปรที่นำเข้าหรือออกจากตัวแบบ จะสอดคล้องกับ AIC ที่มีค่าน้อยลง

สรุปการเลือกตัวแปรอิสระโดยวิธีการถดถอยแบบขั้นบันได

ขั้นที่ 1 เลือกตัวแปรอิสระตัวแรกเข้าตัวแบบ โดยใช้เกณฑ์ของวิธีการเลือกตัวแปรแบบเลือกเข้า (forward selection) คือ จะเลือกตัวแปรอิสระที่มีนัยสำคัญกับตัวแปรตามมากที่สุดเข้าตัวแบบ

- ถ้าพบว่าไม่มีตัวแปรใดมีนัยสำคัญจะหยุด และถือว่าไม่มีตัวแปรอิสระตัวใดมีความสัมพันธ์กับตัวแปรตาม
- แต่ถ้ามีตัวแปรอิสระที่ผ่านเกณฑ์เข้าตัวแบบ จะทำต่อไปในขั้นที่ 2

ขั้นที่ 2 จะใช้เกณฑ์ของวิธีการเลือกตัวแปรแบบเลือกเข้า และวิธีการกำจัดตัวแปรแบบถอยหลัง (backward elimination) นั่นคือ จะพิจารณาเลือกตัวแปรอิสระเข้าตัวแบบจากเกณฑ์ที่กำหนดใน p_E และในขณะเดียวกัน จะพิจารณาตัวแปรอิสระที่อยู่ในตัวแบบว่าสมควรถูกเลือกออกจากตัวแบบหรือไม่ โดยใช้เกณฑ์ p_R ที่กำหนด

ขั้นที่ 3 ทำซ้ำขั้นตอนที่ 2 จนกระทั่งไม่มีตัวแปรอิสระตัวใดเข้าหรือออกจากตัวแบบอีก

ค) การตรวจสอบค่าตกค้างสำหรับ Cox PH model

เมื่อได้ Cox PH model จากการพิจารณาเลือกตัวแปรอิสระที่ดีที่สุด ที่ควรอยู่ในตัวแบบโดยวิธีการถดถอยแบบขั้นบันได ขั้นตอนต่อไป จะเป็นการตรวจสอบตัวแบบว่าเป็นตัวแบบที่มีความเหมาะสมกับข้อมูลหรือไม่ โดยสิ่งที่สนใจในการตรวจสอบมี 5 ประการ คือ

1. ตรวจสอบความเหมาะสมของตัวแบบ
2. ตรวจสอบตัวแปรอิสระที่ได้มาจากวิธีการคัดเลือกตัวแปรเข้าตัวแบบว่ามีรูปแบบฟังก์ชันดีที่สุด
3. ตรวจสอบความเพียงพอสำหรับการพยากรณ์เวลาการคงอยู่ของหน่วยสังเกตที่กำหนดไว้
4. ตรวจสอบตัวแบบว่าเป็นไปตามข้อสมมุติของ Cox PH model
5. ตรวจสอบประสิทธิภาพในแต่ละหน่วยสังเกตที่อยู่ในตัวแบบที่เหมาะสม

1) การตรวจสอบความเหมาะสมของตัวแบบโดย Cox-Snell residuals

สำหรับข้อมูล (T_j, δ_j, Z_j) j =1,2...,n กำหนดให้ $Z_j = (Z_{j1}, Z_{j2}, ..., Z_{jp})'$ เป็นตัวแปรอิสระที่ถูกกำหนดเอาไว้ล่วงหน้า และให้ $h(t|Z_j) = h_0(t)\exp(\sum_{k=1}^{p}\beta_k Z_{jk})$ มีความเหมาะสมกับตัวแบบ ดังนั้น เมื่อตัวแบบมีความเหมาะ

สม ผลลัพธ์ของตัวแปรสุ่มจะมีการแจกแจงเป็นยูนิฟอร์มช่วงหน่วย (unit interval) หรือตัวแปรสุ่ม $U = H(T_j \mid Z_j)$ มีการแจกแจงเป็นแบบเลขชี้กำลัง ที่มีฟังก์ชันภาวะภัยเท่ากับหนึ่ง ในที่นี้ $H(t \mid Z_j)$ คือ ฟังก์ชันภาวะภัยสะสมที่แท้จริง สำหรับตัวอย่างที่มีเวกเตอร์ตัวแปรอิสระ Z_j

ถ้าตัวแบบมีความเหมาะสม และค่า b มีค่าใกล้เคียงกับค่าจริงของ β Cox-Snell residuals (r_j) จะพิจารณาเหมือนกับตัวอย่างที่มีค่าไม่สมบูรณ์ จากการแจกแจงแบบเลขชี้กำลังที่มีฟังก์ชันภาวะภัยเป็นค่าคงที่เท่ากับหนึ่ง

การพิจารณา r_j ว่าเป็นเหมือนกับตัวอย่างที่ได้มาจากเลขชี้กำลังหน่วย (unit exponential) จะคำนวณมาจากฟังก์ชันภาวะภัยสะสม r_j ($\hat{H}_r(r_j)$) จากตัวประมาณ Nelson-Aalen ถ้าการแจกแจงแบบเลขชี้กำลังหนึ่งหน่วย (a unit exponential distribution) มีความเหมาะสมกับข้อมูล ดังนั้น การลงจุดความสัมพันธ์ (plot) ระหว่าง $\hat{H}_r(r_j)$ กับ r_j จึงควรเป็นเส้นตรงผ่านจุดกำเนิด

2) การพิจารณาฟังก์ชันของตัวแปรอิสระโดยใช้ Martingale residuals

ในการพิจารณารูปแบบฟังก์ชันดีที่สุดสำหรับตัวแปรอิสระ ที่กำหนดภายใต้ Cox PH model กำหนดว่า เวกเตอร์ตัวแปรอิสระ Z ถูกแบ่งออกเป็นเวกเตอร์ Z^* ซึ่งเป็นรูปแบบฟังก์ชันที่เหมาะสมของ Cox PH model และตัวแปรอิสระ Z_1 ที่ยังไม่ได้กำหนดรูปแบบฟังก์ชันที่จะใช้ เมื่อกำหนดว่า Z_1 เป็นอิสระต่อ Z^* และให้ $f(Z_1)$ เป็นฟังก์ชันที่เหมาะสมของตัวแปรอิสระ Cox PH model ที่ได้จะเป็น

$$H(t \mid Z^*, Z_1) = H_0(t) \exp(\beta^* Z^*) \exp[f(Z_1)]$$

ในการหารูปแบบฟังก์ชัน จะกำหนดให้ Cox PH model มีพื้นฐานอยู่บน Z^* เพื่อคำนวณ Martingale residuals ($\hat{M}_j$) และลงจุดค่าตกค้าง (residual plots) นี้กับค่า Z_1 สำหรับหน่วยสังเกตที่ j

3) การตรวจสอบความเพียงพอสำหรับค่าประมาณโดย Deviance residuals

การตรวจสอบตัวแบบสำหรับค่า outliers เมื่อได้ Cox PH model สุดท้ายที่มีความเหมาะสมกับข้อมูล จะพิจารณาเช่นเดียวกับสูตรสมการการถดถอยทั่วไป เพื่อแสดงให้เห็นถ้าค่าสังเกตนั้นมีค่าพยากรณ์ไม่แม่นยำจากตัวแบบที่เหมาะสม

ค่าตกค้างประเภทนี้ จะมีค่าเป็นศูนย์เมื่อ $\hat{M}_j$ เป็นศูนย์ โดยค่าลอการิทึมจะมีแนวโน้มทำให้ค่าตกค้างมีค่าสูงขึ้น เมื่อ $\hat{M}_j$ มีค่าเข้าใกล้หนึ่ง และมีค่าลดลงเมื่อค่า $\hat{M}_j$ มีค่าเป็นลบมาก

4) การตรวจสอบข้อสมมุติของ Cox PH model โดย Score residuals

การลงจุดความสัมพันธ์ของ $W_k(t)$ กับเวลา ควรพิจารณาได้เหมือนกับแนวเดินเชิงสุ่ม (tied down random walk) ถ้าฟังก์ชันภาวะภัยสำหรับระดับความแตกต่างของตัวแปรอิสระไม่เป็นสัดส่วน ค่าที่แสดงจากความสัมพันธ์จะมีค่ามากเกินกว่าค่าสัมบูรณ์ (absolute value) ที่บางค่าสังเกต

5) การตรวจสอบอิทธิพลของค่าสังเกตที่ได้จากหน่วยสังเกต

ภาพการลงจุด (scatter plots) ในการตรวจสอบอิทธิพลของค่าสังเกตบนค่าประมาณของ β ค่าเฉลี่ยที่ดีที่สุด ในการตรวจสอบอิทธิพลของค่าสังเกตที่กำหนดไว้ในกระบวนการประมาณคือ การเปรียบเทียบค่าประมาณ b ที่ได้จากการประมาณ β จากข้อมูลทั้งหมด ไปยังค่าประมาณ $b_{(j)}$ ที่ได้มาจากข้อมูลของค่าสังเกตที่กำหนดไว้ ลบออกจากตัวอย่าง ถ้า $b - b_{(j)}$ มีค่าเข้าใกล้ศูนย์ ค่าสังเกตที่ j จะมีอิทธิพลน้อยต่อค่าประมาณ ในขณะที่ความเบี่ยงเบนมาก จะเป็นการบอกถึงอิทธิพลที่มาก ตามลำดับ

การประมาณจะมีพื้นฐานอยู่บน Score residuals ที่ได้กำหนดไว้ก่อนหน้านี้ ให้ $S_{jk} = S_{jk}(\infty)$ ที่ซึ่ง S_{jk} ถูกกำหนดไว้ในสมการ (2.6.4) นั่นคือ เมื่อตัวแปรอิสระทุกตัวถูกกำหนดไว้ที่เวลาเริ่มต้น และข้อมูลมีค่าสังเกตสมบูรณ์ทุกค่า (δ_j จะถูกกำหนดให้มีค่าเท่ากับ 1) จะมีค่าเป็น

$$S_{jk} = [Z_{jk} - \bar{Z}_k(T_j)] - \sum_{t_h \le T} [Z_{jk} - \bar{Z}_k(t_h)] \exp\{b' Z_j\} [\hat{H}_0(t_h) - \hat{H}_0(t_{h-1})]$$

สำหรับ j = 1,...,n และ k = 1,...,p

เทอม $[Z_{jk} - \bar{Z}_k(T_j)]$ คือ Schoenfeld's partial residuals ซึ่งหมายถึงความแตกต่างระหว่างตัวแปรอิสระ Z_{jk} ที่ช่วงเวลาล้มเหลว และค่าคาดหวังของตัวแปรอิสระ ณ เวลานี้ แสดงได้ว่า $b - b_{(j)}$ คือ การประมาณโดย $\Delta = I(b)^{-1}(S_{j1},\ldots S_{jp})'$ เมื่อ I(b) คือ ค่าสังเกตของสารสนเทศฟิชเชอร์ (Observed Fisher Information) ความสัมพันธ์สำหรับค่านี้กับแต่ละตัวแปรอิสระ Z_{jk} จะแสดงถึงอิทธิพลของค่าสังเกตที่ j บนตัวแปรอิสระลำดับที่ k

การนำตัวแบบกึ่งอิงพารามิเตอร์ (Semi-parametric model) หรือ Cox PH model มาใช้ในการพิจารณานั้น มีเงื่อนไขที่ว่า ข้อมูลการคงอยู่ที่ศึกษาจะต้องเป็นไปตามข้อสมมุติ Cox PH model คือ ฟังก์ชันภาวะภัยของกลุ่มตัวอย่างที่ศึกษาต้องเป็นสัดส่วนกัน (Proportional hazard function) (Therneau and Grambsch, 2001) ถ้าพบว่าตัวแบบนั้นไม่เป็นไปตามข้อสมมุติ วิธีการต่อจากนี้ จะเปลี่ยนตัวแบบในการประมาณค่าพารามิเตอร์จากเดิมที่เคยใช้ Cox PH model มาเป็นการประมาณค่าพารามิเตอร์จาก AFT model แทน

3.3.3.2 Accelerated Failure-Time Model (AFT model)

เนื่องจากการแจกแจงการคงอยู่ที่มักพบบ่อยอยู่เสมอในการวิเคราะห์ข้อมูลการคงอยู่ คือ การแจกแจงแบบเลขชี้กำลัง การแจกแจงแบบไวบูลล์ การแจกแจงแบบลอก-นอร์มอล การแจกแจงแบบลอกลอจิสติค การแจกแจงแบบแกมมามาตรฐาน และการแจกแจงแบบแกมมาทั่วไป รวมทั้งเมื่อพิจารณาจากงานวิจัยที่เกี่ยวข้องกับการวิเคราะห์การคงอยู่ของกรมธรรม์ประกันชีวิตประเภทสามัญที่สิ้นผลบังคับลงก่อนกำหนด ในประเทศสิงคโปร์จำนวน 48,243 กรมธรรม์ ก็มีการแจกแจงเป็นไปตามตัวแบบไวบูลล์ (Lian et al., 1998) การวิจัยในครั้งนี้ จึงจะพิจารณาตัวแบบอิงพารามิเตอร์ที่มีการแจกแจงการคงอยู่ทั้ง 6 ตัวแบบนี้เท่านั้น โดย AFT model ถูกกำหนดจากความสัมพันธ์

$$S(t|Z) = S_0[\exp(\theta'Z)t] \qquad \text{ทุกค่า t}$$

เมื่อ Z เป็นเวกเตอร์ของตัวแปรอิสระที่มีค่าคงที่ $\theta' = (\theta_1, ..., \theta_p)$ เป็นเวกเตอร์ของพารามิเตอร์การถดถอย โดยมี $\exp(\theta'Z)$ เป็นตัวประกอบความเร่ง ที่บอกให้ทราบถึงการเปลี่ยนแปลงในค่าของตัวแปรอิสระ และตัวแบบเชิงเส้นทั่วไปสำหรับลอการิทึมของเวลา คือ

$$Y = \ln(T) = \mu + \gamma'Z + \sigma W$$

เมื่อ $\gamma' = (\gamma_1, ..., \gamma_p)$ เป็นเวกเตอร์ของพารามิเตอร์การถดถอย และ W เป็นการแจกแจงของความคลาดเคลื่อน (The error distribution)

ก) การประมาณพารามิเตอร์ด้วยวิธีความควรจะเป็นสูงสุด

ประมาณค่าพารามิเตอร์จากฟังก์ชันความควรจะเป็นตามตัวแบบอิงพารามิเตอร์จากการแจกแจงแบบต่างๆ พร้อมตรวจสอบค่าประมาณพารามิเตอร์ที่ได้จาก $H_0 : \beta = \beta_0$ ว่าเป็นไปตามสมมติฐานว่างที่ตั้งขึ้นหรือไม่ โดยวิธี Wald's test

$$\chi^2_W = (b - \beta_0)'I(b)(b - \beta_0)$$

โดยมีการแจกแจงไคสแควร์ ที่องศาความเป็นอิสระ p ภายใต้สมมติฐานว่าง

ข) การเลือกตัวแบบที่ดีที่สุดโดย Akaike 's Information Criterion

ใช้เกณฑ์การคัดเลือกตัวแบบโดยสารสนเทศของ Akaike

$$AIC_k = -2\log(ML_k) + 2p_k$$

เมื่อ p_k คือ จำนวนพารามิเตอร์ของตัวแบบที่ k โดยตัวแบบที่ให้ค่า AIC ต่ำสุด จะเป็นตัวแบบที่ให้ค่าพยากรณ์แม่นยำและถูกต้อง

ค) การตรวจสอบสำหรับ Accelerated Failure-Time Model

จากที่ได้กล่าวถึงความหลากหลายตัวแบบของข้อมูลการคงอยู่สำหรับหนึ่งตัวแปรสุ่ม (univariate survival data) และตัวแบบอิงพารามิเตอร์ต่างๆ ที่ใช้ในการศึกษาสำหรับตัวแปรอิสระต่อเวลาการคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด นอกจากนี้ จะทำการตรวจสอบความเหมาะสมของตัวแบบโดยพิจารณาจากค่าตกค้างประเภทต่างๆ

ค่าตกค้างที่จะทำการทดสอบสำหรับ AFT model คือ Cox-Snell residuals ที่ใช้ในการตรวจสอบความเหมาะสมของตัวแบบ โดย Cox-Snell residuals (r_j) ถูกกำหนดโดย

$$r_j = \hat{H}(T_j | Z_j)$$

เมื่อ $\hat{H}$ เป็นตัวแบบที่เหมาะสมสำหรับข้อมูล ดังนั้น r_j จะมีการแจกแจงแบบเลขชี้กำลังมาตรฐาน ($\lambda = 1$) เมื่อลงจุดดูความสัมพันธ์ระหว่างค่าตกค้าง r_j กับค่าประมาณภาวะภัยสะสม r_j ($\hat{H}_r(r_j)$) ที่ได้จากการประมาณ Nelson-Aalen เส้นตรงที่ได้จะมีความชันเท่ากับหนึ่ง

Martingale residuals และ Deviance residuals สำหรับ AFT model ที่ข้อมูลมีค่าสังเกตสมบูรณ์ทุกค่า (δ_j ถูกกำหนดให้มีค่าเท่ากับ 1) ดังเช่นข้อมูลที่ใช้ในการวิจัยครั้งนี้ จะถูกกำหนดเท่ากับ

$$M_j = 1 - r_j$$

$$D_j = sign(\hat{M}_j)\{-2[\hat{M}_j + \ln(1 - \hat{M}_j)]\}^{1/2}$$

การลงจุดดูความสัมพันธ์ระหว่าง Martingale residuals หรือ Deviance residuals กับเวลาการคงอยู่ของค่าสังเกต หรือตัวประกอบความเร่ง จะเป็นการตรวจสอบความเหมาะสมในรูปแบบฟังก์ชันของตัวแปรอิสระ และความเพียงพอสำหรับค่าประมาณการคงอยู่เช่นเดียวกับกรณี Cox PH model

3.4 สรุปผลที่ได้จากการวิเคราะห์ข้อมูล

สรุปผลที่ได้จากฟังก์ชันการคงอยู่ ค่าเฉลี่ย และค่ามัธยฐานจำนวนปีการคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนดทั้งหมดในรอบปีพ.ศ. 2544 และเปรียบเทียบฟังก์ชันการคงอยู่ในแต่ละปัจจัยที่คาดว่าจะมีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิต พร้อมทั้งพิจารณาตัวแบบเชิงพหุสำหรับข้อมูลการคงอยู่ เพื่อพิจารณาตัวแปรอิสระที่มีผลต่อระยะเวลาการคงอยู่ของกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนด และสรุปผลที่ได้จากการวิเคราะห์ข้อมูลเปรียบเทียบในแต่ละปัจจัย เมื่อพิจารณาเฉพาะกรมธรรม์ประกันชีวิตที่สิ้นผลบังคับลงก่อนกำหนดเท่านั้น