Journal of Research Methodology, Volume 17, Number 1 (January - April 2004)

วารสารวิธีวิทยาการวิจัย ปีที่ 17 ฉบับที่ 1 (มกราคม - เมษายน 2547)

Survival Analysis

Yothin Sawangdee
Warangkana Polprasert
Nisaporn Wattanasupt

ABSTRACT

Survival Analysis is a widely used statistical methods which arises in several different disciplines and in particular terminologies such as in sociology called Event History Analysis (EHA), in engineering called Reliability Analysis, and in economy called Duration Analysis. Actually, this method concerns about the occurrence and timing of event which offers the probability of risk in survival or failure times of event. The key thing for survival analysis is "time" which has 2 types; the origin time and the censoring time. In order to get this information, time interval on the occurrence of events are needed such as working status during 5 years after graduated.

There are 2 main methods of survival analysis. First is bio-medical traditional including Life table and Kaplan-Meier methods. And second is Multivariate Analysis using Cox regression which classified to be Discrete Time Regression and Continuous Time Regression. Nowadays, the work of survival analysis can calculate on Statistical Program for Social Science (SPSS), which is a popular statistical package for researchers and graduate students and the analysis in this article used SPSS Version 9.

Journal of Research Methodology, Volume 17, Number 1 (January - April 2004)

วารสารวิธีวิทยาการวิจัย ปีที่ 17 ฉบับที่ 1 (มกราคม - เมษายน 2547)

การวิเคราะห์การรอดชีพ

โยธิน แสวงดี
วรางคณา ผลประเสริฐ
นิสาพร วัฒนศัพท์

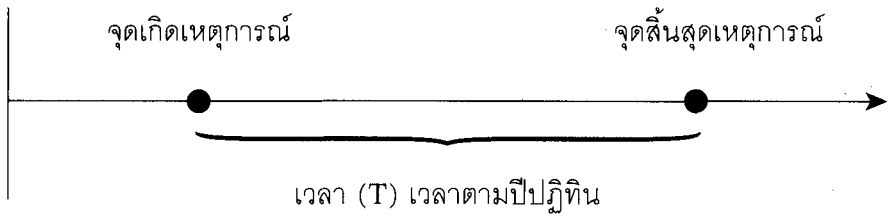
บทคัดย่อ

การวิเคราะห์การรอดชีพเป็นวิธีการวิเคราะห์ที่นิยมใช้กับหลาย ๆ สาขาวิชา โดยแต่ละสาขาวิชาจะมีชื่อเรียกแตกต่างกันออกไป เช่น สาขาสังคมวิทยาจะเรียกว่า Event History Analysis (EHA) สาขาวิศวกรรมศาสตร์จะเรียกว่า Reliability Analysis สาขาเศรษฐศาสตร์จะเรียกว่า Duration Analysis การวิเคราะห์การรอดชีพเป็นการวิเคราะห์ถึงโอกาสที่จะเกิดเหตุการณ์ที่นักวิจัยสนใจโดยลักษณะของเหตุการณ์จะมี 2 เหตุการณ์ คือ เกิดกับไม่เกิดเหตุการณ์ ในช่วงเวลาที่เราสนใจศึกษา ดังนั้นในการวิเคราะห์การรอดชีพ ข้อมูลที่จำเป็นคือ ช่วงเวลาที่เราศึกษา และเหตุการณ์ที่เราสนใจในช่วงเวลานั้น ๆ เช่น การมีงานทำในช่วงเวลา 5 ปีหลังจากจบการศึกษา เป็นต้น

การวิเคราะห์โอกาสรอดชีพมีวิธีการศึกษา 2 ลักษณะใหญ่ ๆ คือ 1) วิธีตารางชีพ (Life Table) ซึ่งแบ่งเป็น 2 วิธี ได้แก่ วิธีการตารางชีพ (Life Tables) และวิธีการของ Kaplan-Meier Method 2) วิธีแบบ Multivariate Analysis ที่ใช้ Cox-Regression ซึ่งแบ่งเป็น 2 วิธี ได้แก่ Discrete Time Regression Model และ Continuous Time Hazard Model การวิเคราะห์การรอดชีพทั้งสองแบบนี้ในปัจจุบันได้มีโปรแกรมสำเร็จรูปช่วยในการวิเคราะห์หลายโปรแกรม แต่ในเอกสารฉบับนี้จะยกตัวอย่างการวิเคราะห์โดยใช้โปรแกรมสำเร็จรูป SPSS Version 9 เนื่องจากโปรแกรมดังกล่าวเป็นโปรแกรมที่ใช้กันอย่างแพร่หลายในกลุ่มนักวิจัย และนักศึกษาระดับบัณฑิตศึกษาในประเทศไทย

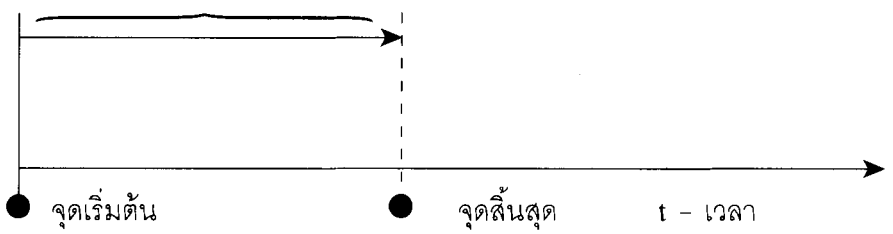
การวิเคราะห์การรอดชีพ (Survival Analysis)

ในปัจจุบันซึ่งเป็นยุคโลกาภิวัฒน์ที่มีการค้นคว้าทดลองอย่างไร้พรมแดนและมีนักวิจัยจำนวนมากที่ต้องการรู้ว่านับตั้งแต่เริ่มเกิดเหตุการณ์ต่าง ๆ ที่ต้องการศึกษาไปจนถึงการสิ้นสุดเหตุการณ์นั้น ๆ โอกาสในการอยู่รอดหรือประสบความสำเร็จและโอกาสที่จะล้มเหลวจะเป็นเท่าใด ยกตัวอย่างเช่น นับแต่เกิดภาวะวิกฤตเศรษฐกิจในประเทศไทยจนถึงสิ้นสุดเหตุการณ์ โอกาสที่บริษัทจะอยู่รอดมีเท่าใด สถิติที่นิยมใช้ในการวิเคราะห์โอกาสอยู่รอดดังกล่าว คือ "การวิเคราะห์การรอดชีพ" ซึ่งในปัจจุบันเป็นวิธีที่นิยมใช้กับหลายสาขา แต่จะเรียกแตกต่างกัน สาขาสังคมวิทยาจะเรียกว่า Event History Analysis (EHA) สาขาวิศวกรรมศาสตร์เรียกว่า Reliability Analysis สาขาเศรษฐศาสตร์เรียกว่า Duration Analysis สิ่งสำคัญสำหรับการวิเคราะห์ คือ การมีเวลา (Time) เข้ามาเกี่ยวข้อง นั่นคือตัวแปรเวลา (The Random Variable T-length of Time) จะต้องนับตั้งแต่เกิดเหตุการณ์ไปจนถึงเหตุการณ์สิ้นสุด ระยะเวลาของเหตุการณ์โดยทั่วไปพิจารณาจากเวลาตามปีปฏิทิน เช่น จำนวนวัน จำนวนเดือน หรือจำนวนปี ที่นับจากจุดเริ่มต้นไปทางด้านขวา (การวิเคราะห์ที่เริ่มต้นนับย้อนไปทางซ้าย ระยะเวลาติดลบ จะไม่นิยมเพราะไม่มีเหตุการณ์เกิดขึ้น) ดังแผนภูมิต่อไปนี้



เมื่อจำลองภาพเพื่อคำนวณช่วงเวลา (t) ว่านับตั้งแต่เหตุการณ์เกิดขึ้นไปจนถึงเหตุการณ์สิ้นสุด หรือจุดที่กำหนดเป็นจุดตรวจสอบ จะได้ภาพดังนี้

ระยะเวลาที่จะนำไปวิเคราะห์ Survival Analysis



จากแผนภูมิทั้ง 2 จะเห็นว่า T มี 2 สัญลักษณ์ โดย T หมายถึงตัวแปรเวลา (Time) ส่วน t หมายถึงช่วงเวลาที่เป็นค่าที่นับได้จากการคำนวณเวลาจริง เป็นค่าของเวลาว่ายาวนานเท่าใด และวางหลักไว้ว่า T1, T2,...Tn ซึ่งตัวแปรเวลาตัวที่ 1 ตัวที่ 2...ไปจนถึงตัวที่ n จะต้องเป็นอิสระแก่กันและมีการกระจายในลักษณะของข้อมูลแบบต่อเนื่อง นอกจากนั้น T จะต้องมากกว่า 0 ซึ่งหมายความว่ามีเหตุการณ์เกิดขึ้น T=0 จะไม่นับ เพราะถือว่ายังไม่มีการเริ่มต้นนับช่วงเวลา ดังนั้น ข้อมูลที่จะใช้จึงเป็นข้อมูลต่อเนื่องระยะยาว (Longitudinal Data) หากเป็นข้อมูลการสำรวจแบบตัดขวาง (Cross-sectional Data) ข้อมูลที่จะนำมาใช้ต้องมีการบันทึกตามช่วงเวลาในลักษณะถามย้อนอดีต เพราะการวิเคราะห์เริ่มพิจารณาเหตุการณ์จากจุดเริ่มต้น ไปจนถึงจุดสิ้นสุดของเหตุการณ์ เช่น ในการวิเคราะห์การตาย จุดเริ่มต้นจะนับจากวันเกิด และจุดสิ้นสุดจะนับที่วันตาย แต่เนื่องจากว่าหลายสาเหตุของการตายเกิดจากการติดเชื้อ ดังนั้น จะนิยมนับเวลาเริ่มต้น ณ วันที่พิสูจน์ว่ามีการติดเชื้อและกำหนดช่วงเวลาว่า ภายใน 5 ปี 10 ปี อาจจะกำหนดเป็นสัปดาห์หรือเป็นเดือน ฯลฯ ผู้ติดเชื้อจะมีผู้ใดรอดชีพบ้างและมีผู้ใดเสียชีวิตบ้าง ในช่วงที่ทำการตรวจสอบอาจมีการนำปัจจัยทางด้านเศรษฐกิจ สังคม และประชากรศาสตร์ร่วมพิจารณาด้วย หากมีปัจจัยที่สามารถเปลี่ยนแปลงตามเวลา เช่น สถานะภาพสมรส การมีงานทำ ฯลฯ เพื่อให้ได้ผลการทำนายที่แม่นตรงจะนำตัวแปรเหล่านี้เข้าร่วมพิจารณาโดยเรียกว่า "ตัวผันแปรร่วมตามเวลา" (Time-Dependent Covariate) สำหรับหลักเบื้องต้นที่ต้องทำความเข้าใจก่อนเริ่มการวิเคราะห์คือแนวคิดเกี่ยวกับฟังก์ชั่นการเสี่ยงต่อการไม่รอดชีพ (Hazard Function) ซึ่งเนื้อหาส่วนใหญ่ในการบรรยายนำมาจากหนังสือ Survival Analysis Using The SAS System : A Practical Guide ที่เขียนโดย Paul D. Allison เมื่อปี ค.ศ. 2000 ฟังก์ชั่นของการเสี่ยงต่อการไม่รอดชีพ ณ ช่วงเวลา t มีสูตรดังนี้

$$h(t) = \lim_{\Delta t \to 0} \frac{\Pr\{t \leq T < t + \Delta t | T \geq t\}}{\Delta t}$$

โดย h = Hazard Rate

t = ช่วงเวลาที่เหตุการณ์จะเกิดขึ้น

T = เวลาที่พิจารณาโอกาสรอดชีพ

Δt = จำนวนเวลาที่เปลี่ยนแปลงไป

เนื่องจากตัวแปรระยะเวลาที่คำนวณจากช่วงเวลาหนึ่งถึงช่วงเวลาหนึ่งที่ใช้สัญลักษณ์ t เป็นตัวแปรที่มีค่าแบบต่อเนื่องและมีค่าเป็นจำนวนเต็ม โดยเหตุการณ์ที่คาดว่าจะเกิดขึ้น ณ ช่วงเวลา t แต่ละช่วงเวลานั้น โอกาสจะเกิดขึ้นอาจไม่เท่ากันและเวลาไม่สามารถย้อนกลับได้ ดังนั้น

เมื่อเริ่มนับเวลา ณ จุดเริ่มต้นเวลาช่วงที่ 1 จะน้อยกว่าช่วงที่ 2 เวลาช่วงที่ 2 จะน้อยกว่าช่วงที่ 3 และทุก ๆ ช่วงเวลาดังสัญลักษณ์ $t_1<t_2<t_3$...นั่นคือโอกาสที่เหตุการณ์จะเกิดขึ้นในแต่ละช่วงของเวลาจะมีลักษณะการกระจายแบบ Population Density Function [f(t)] ซึ่งเป็นการคำนวณโอกาสของความน่าจะเป็นจากประชากรตัวอย่างทั้งหมด และมีรูปฟังก์ชั่น เป็นดังนี้

$f(t_j)$ = $Pr\ (T=t_j)$

f(t) หมายถึง สัญลักษณ์ของ Population Density ของช่วงเวลาที่คำนวณ (t)

T คือ ตัวแปรเวลาที่มีค่าช่วงเวลาเท่ากับ t

j คือ ลำดับช่วงเวลา = 1, 2, 3 .....n

และ $\Sigma f(t_j)$ = 1 นั่นคือ ผลรวมโอกาสของความน่าจะเป็นที่เกิดเหตุการณ์ ณ แต่ละช่วงเวลาจะมีค่าสูงสุดเท่ากับ 1

ดังนั้นฟังก์ชั่นการรอดชีพของสมการในลักษณะ Continuous Time คือ

$$S(t) = \int_t^{\infty} f(u)\, du$$

โดย S เป็นสัญลักษณ์ของฟังก์ชั่นที่ย่อมาจากคำว่า Survival ขณะที่รูปฟังก์ชั่นของ Discrete Time ที่จำนวนเวลามีหน่วยเป็นจำนวนเต็มเช่น 0, 1 คือ

$$S(t_j) = \sum_{k=j}^{\infty} f(t_k)$$

ยกตัวอย่าง เช่น ถ้าเริ่มนับเหตุการณ์ตั้งแต่ช่วงเวลาที่ 3 (j=3) และสิ้นสุดที่ช่วงเวลาที่ 5 จะได้ Survival Function ดังนี้

ดังนั้น $S(t_3) = f(t_3) + f(t_4) + f(t_5)$

เหตุผลที่ Survival Function นับตั้งแต่ 3 ขึ้นไป เนื่องจากว่า จากเวลาเริ่มต้น t_1 ถึง t_2 เหตุการณ์ยังไม่เกิดขึ้น เพื่อเป็นการติดตามอย่างต่อเนื่องในระยะเวลาที่มีหน่วยเป็นจำนวนเต็มตลอดไป จึงเริ่มติดตามต่อไปจากเวลา ณ 3, 4, 5 ที่เป็นช่วงเวลาสุดท้าย ดังนั้น ค่าโอกาสการไม่รอดชีพในแต่ละเวลา (t) จะต้องมีค่าเท่ากับ 0 นั่นคือต้องกำหนดให้เป็นค่าโอกาสความน่าจะเป็นของการตายที่มีค่าต่ำสุดคือ ศูนย์ (0) แล้วพิจารณาว่าเมื่อเวลาเปลี่ยนแปลงไป คือ $t + \Delta t$ โอกาสรอดจะเป็นอย่างไร สำหรับค่า Hazard Rate บางตำราใช้สัญลักษณ์ $\lambda(t)$ ดังสูตรต่อไปนี้

$$\lambda(t) = \lim_{\Delta t \to 0} \frac{Pr\{t \le T < t + \Delta t | T \ge t\}}{\Delta t}$$

และหากวางหลักไว้ว่าให้โอกาสเสี่ยงต่อการตายคงที่ และกำหนดให้ h(t) เป็น Hazard Rate ณ เวลา t และ λ เป็นโอกาสเสี่ยงต่อการตาย จะได้รูปสมการดังนี้

$$h(t) = \lambda$$

อนึ่ง หากทำการกำหนดค่า logarithm ที่ค่า Hazard Rate เช่น log h(t) และใช้ λ เป็นสัญลักษณ์แทนค่า log h(t) จะได้รูปสมการดังนี้

$$\log h(t) = \lambda$$

ดังนั้น Hazard Function ซึ่งเป็นค่าโอกาสเสี่ยงต่อการไม่รอดชีพ ณ เวลา t_j อาจกำหนดเป็นรูปสมการ ให้ λ เป็นสัญลักษณ์แทน Hazard Rate ดังนี้

$$\lambda_j = \Pr (T = t_j \mid T \geq t_j)$$

$$= \frac{f(t_j)}{S(t_j)}$$

โดยที่ f คือ ฟังก์ชั่นไม่รอดชีพ (Failure Function)

S คือ ฟังก์ชั่นของการรอดชีพ (Survival Function)

หรืออาจจะเขียนได้ดังนี้

$$\lambda_j = \frac{S(t_j) - S(t_j + 1)}{S(t_j)}$$

ซึ่งเป็นรูปของสัดส่วนและเป็นค่าโอกาสของความน่าจะเป็นที่มีค่าอยู่ระหว่าง 0 - 1 หากเขียนในรูปของ Likelihood Function จะได้รูปสมการ

$$\log \frac{\lambda_{ij}}{1 - \lambda_{ij}} = B_j X_i$$

โดยที่ B คือ สัมประสิทธิ์สหสัมพันธ์ถดถอย

X คือ ตัวแปรอิสระ

แต่เนื่องจากว่าใน Hazard Function ของโอกาสเสี่ยงที่เหตุการณ์จะเกิดขึ้นมีลักษณะคล้ายแนวคิดของ Benjamin Compertz (1825) ที่วางหลักไว้ในกฎการตายว่าโอกาสการตายจะเพิ่มขึ้นแบบอนุกรมเลขาคณิต (Geometrical Progression) รูปสมการจึงเป็นดังนี้

$$\log h(t) = \lambda + \alpha \log t$$

หากถอดค่า logarithm ออก จะได้รูปสมการดังนี้

$$h(t) = \lambda t^{\alpha}$$

ในการวิเคราะห์โอกาสรอดชีพมีวิธีการศึกษา 2 วิธีคือ 1)โดยวิธีตารางชีพ (Life Tables) ซึ่งมี 2 วิธี เช่น วิธีการตารางชีพใน Function Life Tables ของ SPSS version 9 และวิธีการของ Kaplan-Meier Method ซึ่งมีใน SPSS versions 9 เช่นกัน 2) วิธีแบบ Multivariate Analysis ที่ใช้ Cox-Regression ใน SPSS version 9 ซึ่งแบ่งเป็น Discrete Time Regression Model และ Continuous Time Hazard Model รายละเอียดทั้งวิธีตารางชีพและวิธี Multivariate Analysis จะได้กล่าวต่อไป

โครงสร้างและการจัดข้อมูลเพื่อการวิเคราะห์โอกาสรอดชีพ

1. การกำหนดเวลาเริ่มต้น เนื่องจาก Survival Analysis สามารถนำไปใช้ในการวิเคราะห์เหตุการณ์ชีพหลายเหตุการณ์ หลักในการกำหนดช่วงเวลาเริ่มต้นดังนี้

(1) เลือกช่วงเวลาที่คิดว่าเหมาะสมที่สุดที่จะนำไปสู่การเกิดพฤติกรรมเสี่ยงนั้นๆ เช่น ในการวิเคราะห์การอยู่รอดในการครองเรือน ระยะเวลาเริ่มต้นจะนับจากวันที่ทำการสมรส เช่นเดียวกับการวิเคราะห์โอกาสรอดชีพเนื่องจากการติดเชื้อโรค ช่วงเวลาเริ่มต้นควรนับจากวันที่พิสูจน์ว่ามีการติดเชื้อ

(2) ในการทำวิจัยแบบทดลองที่มีการควบคุมกลุ่มตัวอย่าง (Case Control) ช่วงเวลาเริ่มต้นจะนับตั้งแต่ เริ่มการให้การทดลอง เช่น ฉีดยา ให้สารเคมี ฯลฯ ว่าหลังจากจุดเริ่มต้น โอกาสการรอดชีพเป็นอย่างไร โดยอาจกำหนดจุดของการสิ้นสุดไว้

2. โครงสร้างของข้อมูล ข้อมูลที่ใช้ จะต้องประกอบด้วยตัวแปรที่สำคัญ 2 ตัวแปร คือ

(1) ตัวแปรที่บอกถึงช่วงเวลา (Time) หรือบาง Software จะใช้คำว่า (DUR) (ในโปรแกรม SAS ใช้ DUR) ซึ่งหมายถึง Duration โดยจะนับตั้งแต่ช่วงเวลาของการเริ่มต้น ไปจนถึงการสิ้นสุดของเหตุการณ์ที่มีการตรวจสอบ

(2) ตัวแปร Status เป็นตัวแปรที่กำหนดว่าตัวอย่างที่ตรวจสอบในช่วงเวลาที่กำหนด (Time หรือ Duration) มีสถานะอย่างไร รอดชีพหรือตาย จะกำหนดให้ 1 คือ รอดชีพ 0 คือ ตาย หรือหย่าร้าง ฯลฯ

นอกจาก 2 ตัวแปรหลักที่สำคัญข้างต้น นักวิจัยจะนำตัวแปรผันร่วม เข้าร่วมพิจารณาด้วยก็ได้ ถ้าตัวแปรผันร่วมเป็นตัวแปรที่คงที่ ไม่มีการเปลี่ยนแปลงเมื่อเวลาเปลี่ยนไป เช่น สีผิว เพศ สามารถนำตัวแปรเหล่านี้เพิ่มเข้าไปในสมการได้ โดยไม่ต้องคำนวณค่าที่อาจจะเปลี่ยนแปลงไปตามเวลา อนึ่ง หากมีตัวแปรผันร่วมที่เปลี่ยนแปลงไปตามเวลา นักวิจัยต้องมีการคำนวณค่าที่

เปลี่ยนแปลงไปก่อนแล้วจึงจัดเป็นตัวแปรและเพิ่มเข้าไปในสมการทำนาย ดังนั้น ในการจัดโครงสร้างข้อมูล แฟ้มข้อมูลต้องประกอบด้วยตัวแปรหลักๆ อาทิ Time และ Status ตัวอย่างต่อไปนี้ มาจากแฟ้มข้อมูล Software ของ SPSS version 9 เป็นข้อมูลของผู้ป่วยโรคมะเร็งเต้านม วิธีใช้ข้อมูลชุดนี้ไปที่ C:/Program Files/SPSS.9/Breast Cancer Survival.Sav ตัวแปรในแฟ้มข้อมูลชุดนี้มีจำนวนตัวอย่างผู้ที่เป็นมะเร็งเต้านม 1,207 ราย และมีตัวแปร ดังนี้

id	age	pathsize	inpos	histgrad	er	pr	status	pathscat	in-yesno	time
1	60	99.00	0	3	0	0	0	99	0	9.47
2	79	99.00	0	4	2	2	0	99	0	8.60
3	82	99.00	0	2	2	2	2	99	0	19.33
4	66	99.00	0	2	1	1	0	99	0	16.33
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
1206	48	2.50	4	3	0	0	0	2	1	45.20
1207	73	2.40	0	3	1	1	0	2	0	6.10

โดย

id	คือ Identification Number ซึ่งหมายถึงลำดับที่ของตัวอย่าง
age	คือ อายุ ซึ่งมีระดับการวัดเป็นอัตราส่วน
pathsize	คือ Pathologic Tumor Size ซึ่งหมายถึง ขนาดของก้อนเนื้อมะเร็ง มีหน่วยวัดเป็นเซนติเมตร ซึ่งเป็นมาตราวัดแบบอัตราส่วน
inpos	คือ Positive Axillary Lymph Nodes ซึ่งหมายถึง จำนวนต่อมน้ำเหลืองโตใต้รักแร้ มีหน่วยวัดเป็นระดับอันดับ
histgrad	คือ Histologic Grade ซึ่งหมายถึง ระดับความรุนแรงของโรค มีหน่วยวัดเป็นระดับอันดับ
er	คือ Estrogen Receptor Status ซึ่งหมายถึง การทำงานของ Receptor ฮอร์โมน Estrogen มีหน่วยวัดเป็นอันดับ
pr	คือ Progesterone Receptor Status ซึ่งหมายถึง การทำงานของ Receptor ฮอร์โมน Progesterone มีหน่วยวัดเป็นอันดับ
status	คือ Status โดย 0 หมายถึง เสียชีวิต ขณะที่ 1 คือรอดชีพ
pathscat	คือ Pathological Tumor Size ซึ่งหมายถึง ขนาดก้อนเนื้อที่มีพยาธิสภาพ ได้จากการ Recode ตัวแปร Pathsize มีหน่วยวัดเป็นอันดับ

in-yesno คือ Lymph Nodes ซึ่งหมายถึง ต่อมน้ำเหลืองโต โดยที่ 1 หมายถึง มีต่อมน้ำเหลืองโต และ 0 หมายถึง ไม่มีต่อมน้ำเหลืองโต

time คือ ระยะเวลาที่นับเป็นเดือนตั้งแต่พิสูจน์ว่าเป็นมะเร็งเต้านม มีหน่วยวัดเป็นระดับช่วง

จากข้อมูลชุดนี้ นำมาวิเคราะห์โอกาสรอดชีพ โดยใช้วิธีการต่างๆ ข้างต้นดังนี้

1) การวิเคราะห์โอกาสรอดชีพโดยใช้ตารางชีพ

ตารางชีพเป็นเครื่องมือที่สำคัญอันหนึ่งที่ใช้ในการวิเคราะห์โอกาสรอดชีพ (ปราโมทย์ ประสาทกุล และปัทมา ว่าพัฒนวงศ์, 2545) ดังนั้นในการวิเคราะห์โอกาสรอดชีพสามารถใช้คำสั่งในโปรแกรมสำเร็จรูปต่างๆ ที่สร้างขึ้นได้ ต่อไปนี้เป็นการสร้างตารางชีพจาก SPSS version 9 จากแฟ้มข้อมูล Breast Cancer กำหนดให้ มี 2 ตัวแปร ที่สำคัญคือ

1) ตัวแปรตาม ต้องมีระดับการวัดแบบต่อเนื่อง อาจจะเป็นแบบช่วง หรือ แบบอัตราส่วนก็ได้ เพราะตัวแปรนี้ ต้องแบ่งออกเป็นอันตรภาคชั้น เพื่อสะดวกในการวิเคราะห์เป็นกลุ่มๆ เช่น ถ้าเป็นตัวแปรอายุ เมื่อถูกแบ่งเป็นอันตรภาคชั้นจะเป็นบุคคลในแต่ละกลุ่มอายุ

2) ตัวแปร Status ต้องมีค่าเป็น 1 หากรอดชีพและมีค่าเป็น 0 หากตาย ต่อไปนี้คือคำสั่งใน SPSS Version 9 ที่ใช้สร้างตารางชีพ ขั้นตอนในการสั่งงานเป็นดังนี้

1) ไปที่ Analysis

2) ไปที่ Survival

3) จากคำสั่ง Survival ไปที่ Life Tables

4) จากนั้นตรงช่อง Time ให้ใส่ตัวแปรตามลงไปในที่นี้ตัวแปรตามคือ อายุ ที่นับตั้งแต่การพิสูจน์ได้ว่าเป็นมะเร็งเต้านม ถึงปัจจุบัน(วันที่ทำการสำรวจ) ก็ใส่ตัวแปร age

5) ในช่วง Display Time Intervals ที่เขียนไว้ว่า 0 Thru.....By..... ให้มีค่าสูงสุดของอายุจากการสั่งให้เครื่องหาความถี่ (Frequencies) พบว่าค่าสูงสุดซึ่งในที่นี้เท่ากับ 88 ก็ใส่ 88 ลงไป ตรงช่อง Thru สำหรับ ช่อง By นั้นการใส่ค่าขึ้นอยู่กับเราจะแบ่งอันตรภาคชั้นออกเป็นกี่ช่วง เช่น ในตัวอย่างนี้แบ่งอันตรภาคชั้นออกเป็น 10 ช่วง ก็ใส่เลข 10 ลงไป ซึ่งตัวอย่างตรงนี้คือ 0 thru 88 by 10

6) ในช่อง Status กำหนดตัวแปร Status ลงไปในช่อง แล้วกำหนดว่าค่าในช่อง Define Event โดยระบุ Single Value ให้เป็น 0 หมายถึงการไม่รอดชีพหรือตาย

7) แล้วดำเนินการต่อไปตามคำสั่ง Continue

8) กำหนดคำสั่ง OK หลังจากนั้นจะได้ ตารางชีพดังนี้

This subfile contains : 1207 observations

Life Table

Survival Variable Time (months)

Intrvl Start Time	Number Entrng this Intrvl	Number Wdrawn During Intrvl	Number Exposd to Risk	Number of Termnl Events	Propn Termi-nating	Propn Sur-viving	Cumul Propn Surv at End	Proba-bility Densty	Hazard Rate	Intrvl Start Time	SE of Cumul Sur-viiving	SE of Proba-bility Densty	SE of Hazard Rate
.0	1207.0	1.0	1206.5	102.0	.0845	.9155	.9155	.0085	.0088	.0	.0080	.0008	.0009
10.0	1104.0	12.0	1098.0	152.0	.1384	.8616	.7887	.0127	.0149	10.0	.0118	.0010	.0012
20.0	940.0	15.0	932.5	136.0	.1458	.8542	.6737	.0115	.0157	20.0	.0136	.0009	.0013
30.0	789.0	7.0	785.5	126.0	.1604	.8396	.5656	.0108	.0174	30.0	.0144	.0009	.0015
40.0	656.0	16.0	648.0	137.0	.2114	.7886	.4460	.0120	.0236	40.0	.0145	.0010	.0020
50.0	503.0	8.0	499.0	125.0	.2505	.7495	.3343	.0112	.0286	50.0	.0139	.0009	.0025
60.0	370.0	5.0	367.5	103.0	.2803	.7197	.2406	.0094	.0326	60.0	.0127	.0009	.0032
70.0	262.0	4.0	260.0	70.0	.2692	.7308	.1758	.0065	.0311	70.0	.0114	.0007	.0037
80.0	188.0	3.0	186.5	71.0	.3807	.6193	.1089	.0067	.0470	80.0	.0094	.0008	.0054
90.0+	114.0	1.0	113.5	113.0	.9956	.0044	.0005	**	**	90.0+	.0007	**	**

** These calculations for the last interval are meaningless.

The median survival time for these data is 45.49

การแปลผลตารางชีพ

จากตารางชีพข้างต้น ตัวแปรระยะเวลามีระดับการวัดเป็นเดือนที่ป่วยเป็นมะเร็งเต้านมนับแต่ตรวจพบจนถึงปัจจุบัน โดยมีทั้งผู้ที่เสียชีวิตและผู้ที่รอดชีพ และค่าอันตรภาคชั้นถูกแบ่งออกเป็น 10 ช่วง ช่วงเริ่มต้นนับตั้งแต่ช่วง 0 ถึง 9 เดือน พบว่ามีจำนวนตัวอย่างเข้าสู่อันตรภาคชั้นนี้ทั้งหมด 1207 ตัวอย่าง ซึ่งเป็นจำนวนตัวอย่างทั้งหมด และในจำนวนนี้ มี 1 ราย เสียชีวิตระหว่าง 0 ถึง 9 เดือน (Number of Withdrawn During Interval) ดังนั้น ประชากรเสี่ยง (Number Exposed to Risk) ในช่วง 0 ถึง 9 เดือน คือ 1206.5 คน ซึ่งคำนวณโดยหลักการคาดประมาณจำนวนเสี่ยงเป็นดังนี้

$$\text{Number Exposed to Risk} = \text{Number Entering this Interval} - \frac{\text{Number Withdrawn During Interval}}{2}$$

$$\text{Number Exposed to Risk at time}_0 = 1207 - \frac{1}{2}$$

$$= 1207 - 0.5$$

$$= 1206.5$$

ในหัวข้อ Number of Terminal Events หมายถึงจำนวนตัวอย่างที่มีระยะเวลาการเป็นมะเร็งเต้านมในช่วงอันตรภาคชั้นนี้ (0 ถึง 9 เดือน) ที่ยังมีชีวิตอยู่ สำหรับหัวข้อ Proportion Terminating อธิบายถึงค่าโอกาสของความน่าจะเป็นของการตายที่เกิดขึ้นในอันตรภาคชั้นนั้น คำนวณโดยหลักดังนี้

$$\text{Proportion Terminating} = \frac{\text{Number of Terminal Events}}{\text{Number of Exposed to Risk}}$$

เช่น $$\text{Proportion Terminating} = \frac{102.0}{1206.5}$$

$$= 0.0845$$

หรือ $$\text{Proportion Terminating} = \frac{152.0}{1098.0}$$

$$= 0.1484$$ เป็นต้น

ซึ่งในทางประชากรศาสตร์จะเรียก Proportion Terminating นี้ว่าเป็นค่า ${}_nq_x$ หรืออัตรามรณะ ส่วน Proportion Surviving คือสัดส่วนของผู้รอดชีวิตในอันตรภาคชั้นนั้นๆ ในทางประชากรศาสตร์คือค่า P_x โดยมีวิธีคำนวณดังนี้

Proportion Surviving = 1 - Proportion Terminating

เช่น P_0 = 1 - 0.0845 = 0.9155

หรือ P_{10} = 1 - 0.1384 = 0.8616 เป็นต้น

สำหรับช่วง Cummulative Proportion Surviving at End เรียกว่าเป็นค่าจำนวนคนเริ่มต้นในแต่ละอายุ I_x ในตารางชีพมีหลักในการคำนวณดังนี้

1) สำหรับค่าในอันตรภาคชั้นเริ่มต้น ค่า Cumulative Proportion Surviving จะเท่ากับค่า Proportion Surviving

2) ในอันตรภาคชั้นถัดไปใช้วิธีการคำนวณดังนี้

Cumulative Proportion Surviving at End ของชั้นถัดไป = Proportion Surviving at End ของชั้นที่ผ่านมา * Proportion Surviving ของอันตรภาคชั้นนั้นๆ

ตัวอย่าง เช่น

0.7887 = 0.9155 * 0.8616

หรือ 0.6737 = 0.7887 * 0.8542 เป็นต้น

สำหรับช่อง Probability Density หากเปรียบเทียบกับตารางชีพในทางประชากรศาสตร์ก็คือ ค่าจำนวนคนตายในช่วงอายุ x ($_nd_x$) และถ้าคูณด้วย 100 จะกลายเป็นร้อยละของการไม่รอดชีพสามารถคำนวณได้จากผลต่างของค่า Cumulative Proportion Surviving at End ต่อค่าอันตรภาคชั้น เช่น ในอันตรภาคชั้นเริ่มต้นจะคำนวณได้จากโอกาสรอดชีพทั้งหมด คือ 1 ลบด้วย Cumulative Proportion Surviving at End แล้วหารด้วยค่าอันตรภาคชั้นซึ่งเท่ากับ 10 ดังนี้

$$\text{Probability Density time}_0 = \frac{1-0.9155}{10} = 0.0085 \text{ หรือ}$$

$$\text{Probability Density time}_0 = \frac{0.9155 - 0.7887}{10} = 0.0127 \text{ นั่นเอง}$$

ช่องถัดไปคือ Hazard Rate หรืออาจเรียกว่าอัตราเสี่ยงต่อการตายหลักการคำนวณ คือ

$$\text{Hazard Rate ของชั้นนั้น} = \frac{2 \text{ (Probability Density ของชั้นนั้น)}}{\text{(Comulative Proportion Survival at End ของชั้นที่ผ่านมา) + Proporting Surviving ของชั้นนั้นๆ}}$$

ตัวอย่าง เช่น Hazard Rate ของอันตรภาคชั้นระหว่าง 0 ถึง 9 เท่ากับ 0.0088 คำนวณจากหลักดังนี้

$$0.0088 = \frac{2\ (0.0085)}{1 + 0.9155}$$

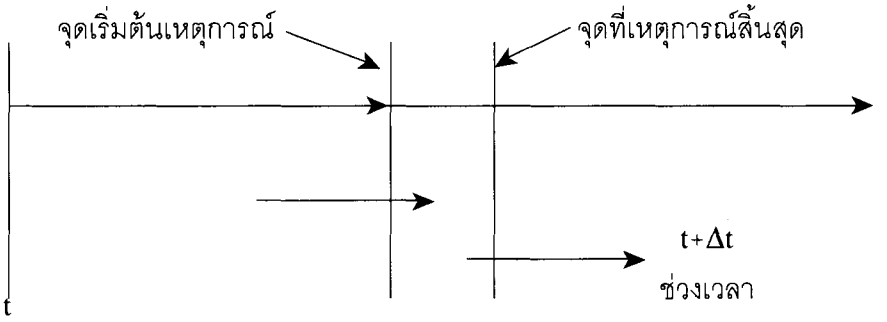
เมื่อพิจารณาค่า Hazard Rate จะเห็นว่าเมื่อระยะเวลาผ่านไปเพิ่มขึ้นตามลำดับของอันตรภาคชั้น โอกาสเสี่ยงต่อการตายจะเพิ่มขึ้นด้วย จะสังเกตเห็นได้ชัดเมื่อ Plot เป็น Line Graph ซึ่งกำหนดที่ Options ของ SPSS ในทำนองเดียวกัน Survival Rate หาก Plot ด้วย Line Graph จะเห็นว่ามีทิศทางลดลงตามเวลาที่เพิ่มขึ้น อนึ่งเนื่องจากว่า Hazard Rate ที่ได้จากการคำนวณเป็นค่าของแต่ละอันตรภาคชั้นที่กำหนดให้เท่ากับ 10 ในการเสนอตารางการวิเคราะห์จำเป็นต้องคำนวณหาค่าที่ถูกต้องจึงต้องปรับ Hazard Rate ให้เป็นค่าเฉลี่ย วิธีการปรับคือ นำค่า Hazard Rate ที่ได้จากตารางหารด้วยค่าอันตรภาคชั้นจะได้ดังนี้

อันตรภาคชั้น	ระยะช่วงของอันตรภาคชั้น	Hazard Rate จากตาราง	ตัวหาร	ค่า Hazard Rate ที่ปรับแล้ว
0-9	10	0.0088	10	0.0009
10-19	10	0.0149	10	0.0015
20-29	10	0.0157	10	0.0016
30-39	10	0.0174	10	0.0017
40-49	10	0.0236	10	0.0024
50-59	10	0.0286	10	0.0029
60-69	10	0.0326	10	0.0033
70-79	10	0.0311	10	0.0031
80-89	10	0.0470	10	0.0047
90 ขึ้นไป	10	-		

ในการวิเคราะห์โดยใช้ตารางชีพใน SPSS Versions 9 สามารถจำแนกโอกาสการตายเปรียบเทียบได้ในแต่ละกลุ่มของตัวแปร โดยตัวแปรนั้นจะถูกกำหนดให้เป็นตัวแปรอิสระและต้องมีระดับของการวัดเป็นกลุ่มหรืออันดับ ในการเปรียบเทียบว่าในแต่ละกลุ่มจะแตกต่างกันอย่างมีนัยสำคัญทางสถิติหรือไม่ จะพิจารณาจากค่า Model Chi-square พร้อมด้วยค่า Degree of Freedom ตลอดจนค่านัยสำคัญทางสถิติของ Model Chi-square นั้น การกำหนดให้ SPSS Versions 9 คำนวณค่า Overall Comparison Statistics ให้ไปที่ Options Compare

2) การวิเคราะห์โอกาสรอดชีพด้วยวิธีการของ Kaplan-Meier (KM)

วิธีนี้มีลักษณะคล้ายกับวิธีตารางชีพ แต่จำนวนตัวอย่างไม่มากกว่า 30 และอยู่บนพื้นฐานของความคิดที่ว่าโอกาสความน่าจะเป็น ณ ช่วงเวลาใดเวลาหนึ่งต้องสูงกว่าหรือมากกว่า ณ ช่วงเวลาเริ่มต้น (t) และช่วงเวลาต้องไม่มีค่าติดลบ โอกาสความน่าจะเป็นที่เหตุการณ์จะเกิดขึ้นมุ่งไปในทิศทางของจำนวนเวลาที่เพิ่มขึ้นไปทางขวา ดังรูปที่แสดงนี้



ดังนั้น $t_1 < t_2 < t_3 ... < t_k$ ถ้าหากกำหนดให้ t_j เป็นเวลาที่จะเกิดเหตุการณ์หนึ่ง และ n_j เป็นบุคคลหนึ่งๆ d_j เป็นเหตุการณ์ตายที่จะเกิดขึ้นของบุคคลนั้นๆ ฟังก์ชั่นการรอดชีพของ Kaplan Meier คือ

$$(t) = \prod_{j:\, tj \le t} (1 - \frac{d_j}{n_j})$$

การวิเคราะห์ตามวิธี Kaplan -Meier ใน SPSS Version 9 กำหนดคำสั่งดังนี้

1) ไปที่ Analysis

2) ไปที่ Survival แล้วเลือก Kaplan-Meier

3) เลือกตัวแปร Time แล้วกำหนดลงในช่อง Time

4) กำหนดตัวแปร Status ให้เป็น Single Value กำหนดให้ 1 เป็นโอกาสไม่รอดชีพ 0 เป็นรอดชีพ ซึ่งการกำหนดค่า Status ใน Kaplan - Meier Method จะสลับกับการกำหนด status ในการคำนวณตามวิธีของตารางชีพ

5) หากมีตัวแปรอิสระที่จะเป็นตัวแปรผันร่วม ให้กำหนดลงไปในช่อง Factor

6) ไปที่ Options เพื่อเลือกค่าทางสถิติและสั่งให้ Plot กราฟต่างๆ ตามที่ต้องการ

ตัวอย่าง ผลลัพธ์จาก Kaplan-Meier ที่ใช้ข้อมูล Breast Cancer ที่ตัดให้เหลือ 26 ตัวอย่าง ตามข้อจำกัดที่จำนวนตัวอย่างต้องไม่เกิน 30 จำแนกเป็นผู้เสียชีวิต 12 ราย ผู้รอดชีวิต 14 ราย ดังนี้

Kaplan-Meier

Survival Analysis for TIME Time (months)

Time	Status	Cumulative Survival	Standard Error	Cumulative Events	Number Remaining
13.03	1	.9615	.0377	1	25
13.10	1	.9231	.0523	2	24
24.53	0			2	23
26.00	0			2	22
29.27	1	.8811	.0646	3	21
29.53	1	.8392	.0739	4	20
41.40	1	.7972	.0812	5	19
41.53	1	.7552	.0871	6	18
42.43	1	.7133	.0918	7	17
46.60	0			7	16
47.97	1	.6687	.0963	8	15
52.70	1	.6241	.0997	9	14
53.67	0			9	13
54.33	0			9	12
54.60	1	.5721	.1041	10	11
66.00	1	.5201	.1068	11	10
71.00	0			11	9
71.63	0			11	8
73.03	1	.4551	.1115	12	7
73.53	0			12	6
87.07	0			12	5
91.93	0			12	4
100.73	0			12	3
103.30	0			12	2
107.40	0			12	1
114.07					

Number of Cases: 26 Censored: 14 (53.85%) Events: 12

	Survival Time	Standard Error	95% Confidence Interval
Mean:	76.15	7.85	(60.76, 91.54)
(Limited to	114.07)		
Median:	73.03	.	(. , .)

Percentiles

	25.00	50.00	75.00
Value	.	73.03	42.43
Standard Error	.	.	9.93

การแปลผลตาราง Kaplan-Meier

จากผลของ Kaplan-Meier ข้างต้น อธิบายได้ว่า เมื่อระยะเวลาอยู่ในช่วง 13.03 เดือน ผู้ป่วยมะเร็งเต้านมมีโอกาสรอดชีพ เท่ากับ 0.9615 และจะมีโอกาสเสียชีวิตเท่ากับ (1-0.9615) หรือเท่ากับ 0.0385 ในทำนองเดียวกันหากป่วยมาแล้วประมาณ 66 เดือน จะมีโอกาสรอดชีพ

เท่ากับ 0.5201 หรือ โอกาสตายเท่ากับ 0.4797 ซึ่งคำนวณได้จากการนำค่าโอกาสรอดชีพสูงสุดลบด้วยค่าโอกาสรอดชีพนั่นเอง (1-0.5201) ขณะที่มัธยฐาน (Median) ของระยะเวลาที่จะรอดชีพเท่ากับ 73.03 เดือน โดยค่าเฉลี่ยของระยะเวลาที่รอดชีพเท่ากับ 76.15 เดือน หากพิจารณาตาม Percentiles จะพบว่า ณ Percentiles ที่ 50 หรือครึ่งหนึ่งของผู้ป่วยจะตายเมื่อระยะเวลาประมาณ 73.03 เดือน ขณะที่ ณ Percentiles ที่ 75 พบว่าผู้ป่วยจะตายเมื่อระยะเวลาประมาณ 42.43 เดือน

3) การวิเคราะห์โอกาสรอดชีพโดยมีตัวผันแปรร่วม

เป็นการวิเคราะห์ใช้หลัก Multivariate Analysis รูปสมการ Hazard function คือ

$$\lambda(t_i x) = \lambda_o \cdot e^{\beta' x}$$

หรืออาจเขียนได้ดังนี้

$$h = h_o * \exp\{b_1(x_1)+b_2(x_2)+...+b_n(x_n)\} + e_i$$

และถ้าต้องการปลดค่า exponential สามารถกำหนดค่า logarithm ได้ดังนี้

$$\ln h = (\frac{h}{h_o}) = b_1(x_1)+b_2(x_2)+...+b_n(x_n)\} + e_i$$

ดังนั้น รูปสมการโอกาสรอดชีพใน Multivariate Analysis Model คือ

$$S_{(t,x)} = e^{-\lambda t \cdot e^{\beta x}}$$

การวิเคราะห์ด้วยวิธี Multivariate Analysis ใน SPSS version 9 จะใช้รูปสมการ Cox Regression (Cox, 1972) วิธีการ Run Cox Regression มีหลักดังนี้

1) ไปที่ Analysis

2) ไปที่ Survival

3) ไปที่ Cox Regression

4) กำหนดตัวแปร Times

5) กำหนดตัวแปร Status โดยระบุ event ให้เป็น 0 หมายถึง ตาย (หรือไม่รอด) และ 1 หมายถึง รอดชีวิต ซึ่งจะกำหนดไว้เหมือนกับวิธีของตารางชีพ

6) กำหนดตัวแปรผันร่วม ซึ่งหมายถึง ตัวแปรอิสระที่นำเข้าวิเคราะห์

7) หากตัวแปรผันร่วมมีระดับการวัดเป็นกลุ่ม ให้ระบุค่าสูงต่ำด้วยและต้องระบุ Reference Category โดยทั่วไปจะใช้ First Category

8) ใช้วิธี Enter เพราะสามารถนำค่า Log Likelihood Estimate ที่ได้จากการคำนวณไปวิเคราะห์การเพิ่มพลังทำนายของสมการในรูป Restricted Model และ Unrestricted Model

ในการแปลผลใช้หลักการเดียวกันกับการแปลผลของ Multivariate Analysis นั่นคือแปลผลค่า B (Beta) ตามทิศทางบวกหรือลบ และพิจารณาแปลผลค่า Exponential Beta สำหรับค่า Wald Chi-square ซึ่งเป็นค่าที่บอกระดับความสัมพันธ์ระหว่างตัวแปรอิสระกับตัวแปรตาม ไม่จำเป็นต้องแปรผลก็ได้ ส่วนค่า R ซึ่งหมายถึง ค่าสหสัมพันธ์จะเห็นได้ว่าหากค่า R มีค่าสูง โอกาสที่ตัวแปรนั้น ๆ จะมีนัยสำคัญทางสถิติกับตัวแปรตามจะมีสูง นอกจากนั้นเป็นการอภิปรายผลตามรูปกราฟที่แสดงถึงค่า Survival Function ที่จะลดลงตามระยะเวลาที่เพิ่มขึ้น ขณะที่ค่า Hazard Function จะเพิ่มขึ้นตามเวลา

4) การวิเคราะห์โอกาสรอดชีพเมื่อมีตัวแปรผันร่วมที่เปลี่ยนแปลงไปตามเวลา

ดังที่ได้อธิบายในตอนต้นเกี่ยวกับการวิเคราะห์โอกาสรอดชีพที่ได้กล่าวถึงตัวแปรผันร่วมที่สามารถเปลี่ยนแปลงสถานะได้ตามเวลา อาทิ สถานภาพสมรส สถานะการมีงานทำ ฯลฯ ในการกำหนดค่าสถานะและสร้างตัวแปร จะปฏิบัติตามหลักดังนี้

1) ให้ใช้สถานะปัจจุบัน (ขณะที่เกิดเหตุการณ์) ของตัวแปรนั้น ดังเช่น มีงานทำให้ใช้ข้อมูลในปัจจุบันว่า ขณะที่เกิดเหตุการณ์มีงานทำหรือไม่ นั่นคือ ใช้หลัก Dummy Variable ในการสร้างตัวแปร เช่น กำหนดให้ 1 = มีงานทำ, 0 = ไม่มีงานทำ หรือหากเป็นสถานะภาพสมรส ให้ใช้สถานะภาพสมรสปัจจุบัน (ขณะที่เกิดเหตุการณ์) เช่น 1 = แต่งงาน, 0 = อื่น ๆ

2) ให้ใช้หลัก Time lag นั่นคือ ย้อนหลังไปสู่ข้อมูลของตัวแปรที่เป็นตัวแปรผันร่วมตามเวลานั้น ๆ โดยย้อนกลับไปจากสภาพปัจจุบันซึ่งมีเหตุการณ์เกิดขึ้น อาจย้อนกลับไป 1 เดือน 3 เดือน หรือ 12 เดือน ก็ได้ขึ้นอยู่กับวัตถุประสงค์ของการวิจัยว่าจะวิเคราะห์ถึงผลกระทบในระยะสั้นหรือระยะยาว จากนั้นสร้างตัวแปรด้วยหลัก Dummy Variable เช่น ตัวแปร สถานะภาพสมรสเมื่อ 1 เดือน จะเป็นดังนี้ (lag 1) 1.โสด, 0. อื่น ๆ หรือ สถานะภาพสมรส เมื่อ 3 เดือนที่ผ่านมา (lag 2) 1.โสด และ 0. อื่น ๆ ตัวแปรผันร่วมตามเวลาตัวอื่น ๆ ก็ใช้หลักการสร้างตัวแปรแบบเดียวกัน แล้วใส่ตัวแปรที่สร้างไปใน Multivariate Model พร้อม ๆ กับตัวแปรที่เป็นสถานภาพขณะที่มีเหตุการณ์เกิดขึ้นก็จะสามารถได้ผลการทำนายออกมา ในรูปของ Proportional Hazard Model รูปสมการตัวอย่างเป็นดังนี้

$$\log\ h_i(t) = \alpha(t) + \beta_1 X_{i1} + \beta_2 X_{i2(t)} + e_i$$

โดยที่ X_1 = ตัวแปรที่ไม่ผันแปรตามเวลา เช่น เพศ

X_2 = ตัวแปรที่ผันแปรตามเวลา เช่น การทำงานซึ่งอาจจะจำแนกได้โดยใช้หลัก Time lag โดยสร้างตัวแปร Dummy Variable ขึ้นมาอีกหลายตัว ซึ่งจะได้รูปสมการดังนี้

$$\log h_i(t) = \alpha(t) + \beta_1 X_{i1} + \beta_2 X_{i2(t)} + \beta_3 X_{i2(t-1)} + \beta_4 X_{i2(t-2)} + e_i$$

แล้วพิจารณาค่า Regression Coefficient ว่าช่วงเวลาใดมีผลมากที่สุด อาทิช่วงเวลา ณ t หรือ t-1 หรือ t-2 โดยที่ t คือเวลาปัจจุบัน นั่นคือ หากค่า Regression Coefficient ณ t มีค่าสูงที่สุด (โดยพิจารณาถึงทิศทางของค่า Regression Coefficient แล้วว่าเป็นค่าบวกหรือลบ) จะแปลผลได้ว่า สถานะของตัวผันแปรร่วม (ซึ่งตัวอย่างที่แสดงนี้ คือ X_{i2}) ณ เวลาปัจจุบันมีผลมากที่สุดต่อโอกาสเสี่ยงแห่งการรอดชีพ (Hazard Rate) เมื่อเปรียบเทียบกับสถานะ ณ เวลา t-1(Lag1)และ t-2 (Lag2) แล้ว อนึ่งหากสถานะของตัวผันแปรร่วมดังตัวอย่างกำหนดให้เป็น X_{i2} ณ เวลา t เวลา t-1(Lag1) และ t-2 (Lag 2) มีสถานะเหมือนกัน เช่น สมมุติว่าตัวแปร X_{i2} ในที่นี้ คือ การมีงานทำ ณ เวลา t ให้มีค่าเป็น 1 และเป็น 0 คือ อื่น ๆ ทำนองเดียวกัน ณ เวลา t-1 หรือ Lag1 บุคคลนั้น ๆ ก็มีงานทำนั่นคือ มีสถานะการทำงานเป็น 1 และ ณ เวลา t-2 หรือ Lag2 บุคคลคนเดียวกันก็มีงานทำนั่นคือ สถานะการมีงานทำมีค่าเป็น 1 ดังนั้น หากค่า Regression Coefficient ณ เวลา t มากกว่า ณ เวลา t-1 และมากกว่า ณ เวลา t-2 โดยมีทิศทางความสัมพันธ์กับ Hazard Rate ไปในทิศทางเดียวกัน เช่น เป็นบวกจะแปรผลได้ว่าการมีงานทำจะมีผลต่อ Hazard Rate เมื่อเวลาเพิ่มขึ้น และหากจะแปลผลว่า การมีงานทำมีผลต่อ Hazard Rate เท่าใด วิธีการคือ นำค่า Regression Coefficient ของตัวแปรผันร่วมตัวนั้น รวมกันทั้งหมด เช่น ผลของการมีงานทำต่อ Hazard rate ณ ที่นี้ คือ :

$$\Sigma\beta_{x2} = \beta_2 + \beta_3 + \beta_4 \quad \text{เป็นต้น}$$

เอกสารอ้างอิง

ปราโมทย์ ประสาทกุล และปัทมา ว่าพัฒนวงศ์. (2544) **ตารางชีพ : เครื่องมือสำคัญทางประชากรศาสตร์.** กรุงเทพฯ : บริษัทอมรินทร์พริ้นติ้งแอนด์พับลิชชิ่ง จำกัด (มหาชน).

Allison, P. (2000). **Survival Analysis Using The SAS System : A Practical Guide.** SAS Institute Inc; North Carolina. USA.

Choe, M.K. (1999). **Notes on Dynamic Event History Analysis.** Hawaii: East-West Center.

Cox, D.R. (1972). "Regression Models and Life Tables" **Journal of the Royal Statistical Society,** Series B, No. 34 : 187-220.

Gompertz,B. (1825). "On the Nature of Function Expressive of The Law of Mortality **Philosophical Transactions.** 27 (1): 513-585.