บทที่ 2

สถิติที่ใช้ในการวิจัย

การปรับค่าประมาณความน่าจะเป็นที่คนอายุ x ปี จะเสียชีวิตภายใน 1 ปีข้างหน้าหรือค่า q'_x สำหรับข้อมูลที่ถูกตัดปลาย เพื่อให้ได้ค่าประมาณที่ปรับแล้วหรือค่า q''_x โดยนำเสนอวิธีการ 3 วิธีดังนี้ การปรับค่าโดยใช้ค่าเฉลี่ยถ่วงน้ำหนักเคลื่อนที่ การปรับค่าโดยใช้รูปแบบฟังก์ชันและการปรับค่าโดยใช้ส่วนโค้งพหุนามองศาสาม สำหรับการหาค่า q'_x จะใช้วิธีการประมาณแบบคณิตศาสตร์ประกันภัย ซึ่งทำการศึกษาในลักษณะของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต ภายใต้ 1. การแจกแจงแบบไวบูลล์ 2. การแจกแจงแบบกอมเพิรตซ์ และระยะเวลาที่จะเกิดการถอนตัวภายใต้ 1. การแจกแจงแบบสม่ำเสมอ 2. การแจกแจงแบบแกมมา โดยจะแสดงรายละเอียดวิธีการประมาณแบบคณิตศาสตร์ประกันภัย และวิธีการปรับค่าแต่ละวิธีตามลำดับต่อไปนี้

<u>การประมาณแบบคณิตศาสตร์ประกันภัย</u>

การประมาณแบบคณิตศาสตร์ประกันภัย เป็นวิธีการหนึ่งในการหาค่า q'_x โดยอาศัยหลักการทางสถิติในการหาค่าความน่าจะเป็น ซึ่งคำนวณได้จากจำนวนคนที่เสียชีวิตหารด้วยระยะเวลาที่เสี่ยงภัยทั้งหมด ภายใต้สมมติฐานและขั้นตอนในการคำนวณดังนี้

จากการศึกษาค่าความน่าจะเป็นที่คนอายุ x ปีจะเสียชีวิตภายใน 1 ปีข้างหน้าหรือในช่วง $(x,x+1]$ จะได้สมการค่าคาดหวังของจำนวนตัวอย่างที่เสียชีวิต ดังนี้

$$E(D) \quad = \quad \sum_{i=1}^{m} {}_{t_i}q_x \; - \; \sum_{k=1}^{w} {}_{1-t_k}q_{x+t_k} \quad = \quad d \qquad (2.1)$$

เมื่อ ${}_{t_i}q_x$ คือความน่าจะเป็นที่คนอายุ x ปี จะเสียชีวิตภายในช่วง $(x,x+t_i]$

${}_{1-t_k}q_{x+t_k}$ คือความน่าจะเป็นที่คนอายุ x ปี จะเสียชีวิตภายในช่วง $(x+t_k,x+1]$

m คือจำนวนตัวอย่างที่ศึกษาในช่วง $(x,x+1]$

w คือจำนวนตัวอย่างที่ถอนตัวจากขนาดตัวอย่าง m ที่เวลา $x+t_k$ ในช่วง $(x,x+1]$

d คือจำนวนตัวอย่างที่เสียชีวิตจากขนาดตัวอย่าง m ในช่วง $(x,x+1]$

D คือตัวแปรสุ่มของจำนวนที่เสียชีวิต

t_i คือระยะเวลาที่จะมีชีวิตอยู่จนกระทั่งเสียชีวิตของตัวอย่างที่ i จากขนาดตัวอย่าง m

t_k คือระยะเวลาที่จะมีชีวิตอยู่จนกระทั่งถอนตัวของตัวอย่างที่ k จากขนาดตัวอย่าง m

จากวิธีการประมาณทางคณิตศาสตร์ประกันภัย และภายใต้สมมติฐานของบาลดุกซ์ซี่ (Balducci 's Assumption)[2] จะได้

$$E(D) \quad = \quad mq_x - q_x\sum_{k=1}^{w}(1-t_k) \quad = \quad d \tag{2.2}$$

ดังนั้นสามารถหาค่าประมาณของค่า q_x ได้คือ

$$q'_x \quad = \quad \frac{d}{m-\sum_{k=1}^{w}(1-t_k)} \tag{2.3}$$

จากนั้นจึงทำการปรับค่าประมาณของค่า q_x ที่ได้ ด้วยวิธีการปรับค่าทั้ง 3 วิธี

<u>การปรับค่าโดยใช้ค่าเฉลี่ยถ่วงน้ำหนักเคลื่อนที่</u>

การปรับค่าโดยใช้ค่าเฉลี่ยถ่วงน้ำหนักเคลื่อนที่ เป็นวิธีการปรับค่าอย่างง่าย ซึ่งมีชื่อเรียกเดิมว่า รูปแบบเส้นตรงเชิงประกอบ (Linear - Compound Formular) ต่อมาได้รับการพัฒนาโดย แอล ดีฟอเรสท์ (E.L. DeForest) วิธีการปรับค่านี้เป็นการปรับค่าโดยนำค่า q'_x มาทำการเฉลี่ยถ่วงน้ำหนักด้วยค่า ๆ หนึ่ง เพื่อให้ได้ค่าประมาณที่ปรับแล้วมีความราบเรียบ ซึ่ง มีสมการคือ

---

[2] Dick London "Survival Models and Their Estimation" Winsted, Connecticut ACTEX Publications, 1988, p 65

$$q''_x = \sum_{r=-n}^{n} a_r q'_{x+r} \qquad (2.4)$$

โดยที่ $a_r = a_{-r}$ สำหรับ $r = 1, 2, \ldots, n$

เมื่อ a_r คือสัมประสิทธิ์แสดงระยะห่างและทิศทางจาก x ซึ่งเป็นดัชนีของค่า q'_x ที่สัมประสิทธิ์นี้คูณอยู่

n คือจำนวนเต็มบวกที่กำหนดพิสัยในการปรับค่า

ภายใต้สมมติฐานที่ว่าในช่วง $[x-n, x+n]$ จะแทนค่า q_x ด้วยพหุนามองศาสาม (Third Degree Polynomial) โดยมีสมการเงื่อนไขคือ

$$\sum_{r=-n}^{n} a_r = 1 \quad \text{และ} \quad \sum_{r=-n}^{n} r^2 a_r = 0$$

$$a_0 + 2\sum_{r=1}^{n} a_r = 1 \quad \text{และ} \quad \sum_{r=1}^{n} r^2 a_r = 0 \qquad (2.5)$$

การหาค่า a_r จะใช้รูปแบบค่า R_z น้อยที่สุด (Minimum - R_z Formula) ซึ่งมีสมการคือ

$$R_z^2 = \frac{Var\left[\Delta^z Q''_x\right]}{Var\left[\Delta^z Q'_x\right]}$$

$$= \frac{1}{C(2z,z)} \sum_{r=-n-z}^{n} \left(\Delta^z a_r\right)^2 \qquad (2.6)$$

เมื่อ Q'_x คือตัวแปรสุ่มค่าประมาณของค่า q_x

Q''_x คือตัวแปรสุ่มค่ากประมาณของค่า q_x ที่ปรับค่าแล้ว

Δ^z คือตัวดำเนินการผลต่างสืบเนื่องข้างหน้าลำดับที่ z (The z Order Forward Difference Operator)

แทนค่า $R = C(2z,z).R_z^2$ ลงในสมการ (2.6)

จะได้ $$R = \sum_{r=-n-z}^{n}\left(\Delta^{z} a_r\right)^2 \qquad (2.7)$$

โดยการหาอนุพันธ์ (Differentiation) ของฟังก์ชัน R เทียบกับค่า a_r จากสมการ (2.7) และให้สมการอนุพันธ์มีค่าเท่ากับศูนย์ จะได้ค่า a_r อยู่ในรูปพหุนามกำลังคู่ (Even - Power Polymial) ซึ่งมีสมการคือ

$$a_r = \left[(n+1)^2 - r^2\right]\left[(n+2)^2 - r^2\right]..\left[(n+z)^2 - r^2\right]\left[h + k\left(n^2 - r^2\right)\right] \qquad (2.8)$$

โดยที่ $a_r = a_{-r}$ สำหรับ $r = 1, 2, \ldots, n$

และ $a_{n+1} = \ldots = a_{n+z} = 0$

จากสมการ (2.5)และ (2.8) สามารถสร้างสมการรูปทั่วไปเพื่อหาค่า h และ k ได้ดังนี้

$$h.S_{n,z} + k.S_{n-1,z+1} = 1$$

และ $$h.S_{n,Z+1} + k.S_{n-1,z+2} = (n+z+1)^2 \qquad (2.9)$$

โดยที่ $$S_{m,n} = \sum_{r=-n}^{n}\left[(n+1)^2 - r^2\right]\ldots\left[(n+m)^2 - r^2\right]$$

$$= C(2m+2n+1, 2m+1).(m!)^2 \qquad (2.10)$$

การหาค่า h และk จะกระทำโดยการหาผลเฉลยของระบบสมการพีชคณิตเชิงเส้น (Linear Algebraic Equation System) ด้วยกฎของคราเมอร์ (Cramer 's Rule) ซึ่งได้สมการดังนี้

$$h = \frac{\begin{vmatrix} 1 & S_{n-1,z+1} \\ (n+z+1)^2 & S_{n-1,z+2} \end{vmatrix}}{\begin{vmatrix} S_{n,z} & S_{n-1,z+1} \\ S_{n,z+1} & S_{n-1,z+2} \end{vmatrix}}$$

$$= \frac{S_{n-1,z+2} - (n+z+2)^2 . S_{n-1,z+1}}{S_{n,z} . S_{n-1,z+2} - S_{n,z+1} . S_{n-1,z+1}}$$

และ

$$k = \frac{\begin{vmatrix} S_{n,z} & 1 \\ S_{n,z+1} & (n+z+1)^2 \end{vmatrix}}{\begin{vmatrix} S_{n,z} & S_{n-1,z+1} \\ S_{n,z+1} & S_{n-1,z+2} \end{vmatrix}}$$

$$= \frac{(n+z+1)^2 . S_{n,z} - S_{n,z+1}}{S_{n,z} . S_{n-1,z+2} - S_{n,z+1} . S_{n-1,z+1}} \qquad (2.11)$$

หลังจากได้ค่า h และ k แล้ว สามารถหาค่า a_r และค่าประมาณของค่า q_x ที่ปรับแล้วได้จากสมการ (2.8) และ (2.4) ตามลำดับ

<u>การปรับค่าโดยใช้รูปแบบฟังก์ชัน</u>

การปรับค่าโดยใช้รูปแบบฟังก์ชัน เป็นวิธีการปรับค่าโดยแทนค่าประมาณของค่า q_x ด้วยรูปแบบฟังก์ชันที่เหมาะสมทางด้านคณิตศาสตร์ประกันชีวิต ในที่นี้จะนำเสนอรูปแบบฟังก์ชันที่สอดคล้องกับระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต ที่ได้ศึกษาในการทำวิจัยนี้ ซึ่งมี 2 รูปแบบคือ 1. รูปแบบไวบูลล์ (Weibull Form) 2. รูปแบบกอมเพริตซ์ (Gompertz Form) โดยจะแสดงรายละเอียดวิธีการของแต่ละรูปแบบดังต่อไปนี้

1. รูปแบบไวบูลล์

ภายใต้กำลังของมรณะ (Force of Mortality) ของรูปแบบไวบูลล์จะได้ว่า

$$\mu_x = kx^n \qquad , k > 0 , n > 0 \qquad (2.12)$$

โดยการแปลงรูป (Transformation) จากสมการ (2.12) จะได้

$$\log \mu_x = \log k + n \log x \qquad (2.13)$$

กำหนดให้ $$\mu_x^* = a + bx^* \qquad (2.14)$$

โดยที่ $\mu_x^* = \log \mu_x'$, $x^* = \log x$,

$a = \log k$ และ $b = n$

เมื่อ $\mu'_{x+0.5} = \dfrac{q'_x}{1-0.5q'_x}$

การหาค่า a และ b จะกระทำได้โดยวิธีกำลังสองน้อยที่สุด (Least Square Method) ซึ่งมีสมการคือ

$$SS = \sum_{x=x}^{x+v-1} w_x\left(\mu_x^* - a - bx^*\right)^2 \tag{2.15}$$

โดยที่ $w_x = \dfrac{m_x}{\mu_x'(1-\mu_x')}$

เมื่อ w_x คือค่าถ่วงน้ำหนัก (Weights) ที่อายุ x ปี

m_x คือขนาดตัวอย่างของข้อมูลที่นำมาศึกษาที่อายุ x ปี

v คือจำนวนค่าของอายุในช่วงที่สนใจศึกษา

โดยการหาอนุพันธ์ของฟังก์ชัน SS จากสมการ (2.15) เทียบกับค่า a และ b ตามลำดับ และให้สมการอนุพันธ์มีค่าเท่ากับศูนย์ จะได้สมการดังนี้

$$a\sum w_x + b\sum x^* w_x = \sum w_x \mu_x^*$$

และ
$$a\sum w_x + b\sum (x^*)^2 w_x = \sum x^* w_x \mu_x^* \tag{2.16}$$

โดยที่ $\sum = \sum_{x=x}^{x+v-1}$

ซึ่งสามารถแสดงในรูปเมตริกซ์เพื่อหาค่า a และ b ได้คือ

$$x^{*'} w x^* y = x^{*'} w \mu_x^* \tag{2.17}$$

โดยที่ $x^{*} = \begin{bmatrix} 1 & \log x \\ 1 & \log x+1 \\ \vdots & \vdots \\ 1 & \log x+v-1 \end{bmatrix}$ เป็นเมตริกซ์ขนาด $(v \times 2)$

$w = \begin{bmatrix} w_x & & O \\ & \ddots & \\ O & & w_{x+v-1} \end{bmatrix}$ เป็นเมตริกซ์ขนาด $(v \times v)$

$\mu_x^{*} = \begin{bmatrix} \log \mu'_x \\ \vdots \\ \log \mu'_{x+v-1} \end{bmatrix}$ เป็นเมตริกซ์ขนาด $(v \times 1)$

$y = \begin{bmatrix} a \\ b \end{bmatrix}$ เป็นเมตริกซ์ขนาด (2×1)

เมื่อ $x^{*'}$ คือการสับเปลี่ยน (Transpose) ของเมตริกซ์ x^{*}

หลังจากได้ค่า a และ b แล้ว จะทำให้ทราบค่า k และ n ได้จากสมการ (2.14)และสามารถหาค่าประมาณของค่า q_x ที่ปรับแล้วได้ดังนี้

$$q_x^{*} = 1-\exp\left[1\int_0^1 \mu_{x+s}\, ds\right]$$

$$= 1-\exp\left[\frac{-k}{n+1}\left\{(x+1)^{n+1} - x^{n+1}\right\}\right] \qquad (2.18)$$

2. รูปแบบกอมเพริตซ์

ภายใต้กำลังของมรณะ (Force of Mortality) ของรูปแบบกอมเพริตซ์จะได้ว่า

$$\mu_x = Bc^{x} \quad , B > 0 , c > 1 \qquad (2.19)$$

โดยการแปลงรูป (Transformation) จากสมการ (2.19) จะได้

$$\log \mu_x \quad = \quad \log B + x \log c \tag{2.20}$$

กำหนดให้ $\mu_x^* \quad = \quad a + bx^*$ (2.21)

โดยที่ $\mu_x^* = \log \mu_x'$, $x^* = x$,

$a = \log B$ และ $b = \log c$

เมื่อ $\mu'_{x+0.5} = \dfrac{q_x'}{1-0.5q_x'}$

การหาค่า a และ b จะกระทำได้โดยวิธีกำลังสองน้อยที่สุด (Least Square Method) ซึ่งมีสมการคือ

$$SS \quad = \quad \sum_{x=x}^{x+v-1} w_x\left(\mu_x^* - a - bx^*\right)^2 \tag{2.22}$$

โดยที่ $w_x = \dfrac{m_x}{\mu_x'(1-\mu_x')}$

เมื่อ w_x คือค่าถ่วงน้ำหนัก (Weights) ที่อายุ x ปี

m_x คือขนาดตัวอย่างของข้อมูลที่นำมาศึกษาที่อายุ x ปี

v คือจำนวนค่าของอายุในช่วงที่สนใจศึกษา

โดยการหาอนุพันธ์ของฟังก์ชัน SS จากสมการ (2.22) เทียบกับค่า a และ b ตามลำดับ และให้สมการอนุพันธ์มีค่าเท่ากับศูนย์ จะได้สมการดังนี้

$$a\sum w_x + b\sum x^* w_x \quad = \quad \sum w_x \mu_x^*$$

และ

$$a\sum w_x + b\sum (x^*)^2 w_x \quad = \quad \sum x^* w_x \mu_x^* \tag{2.23}$$

โดยที่ $\sum = \sum_{x=x}^{x+v-1}$

ซึ่งสามารถแสดงในรูปเมตริกซ์เพื่อหาค่า a และ b ได้คือ

$$x^{*'} w x^{*} y \quad = \quad x^{*'} w \mu_x^{*} \tag{2.24}$$

โดยที่ $x^{*} = \begin{bmatrix} 1 & x \\ 1 & x+1 \\ \vdots & \vdots \\ 1 & x+v-1 \end{bmatrix}$ เป็นเมตริกซ์ขนาด $(v \times 2)$

$w = \begin{bmatrix} w_x & & O \\ & \ddots & \\ O & & w_{x+v-1} \end{bmatrix}$ เป็นเมตริกซ์ขนาด $(v \times v)$

$\mu_x^{*} = \begin{bmatrix} \log \mu_x' \\ \vdots \\ \log \mu_{x+v-1}' \end{bmatrix}$ เป็นเมตริกซ์ขนาด $(v \times 1)$

$y = \begin{bmatrix} a \\ b \end{bmatrix}$ เป็นเมตริกซ์ขนาด (2×1)

เมื่อ $x^{*'}$ คือการสับเปลี่ยน (Transpose) ของเมตริกซ์ x^{*}

หลังจากได้ค่า a และ b แล้ว จะทำให้ทราบค่า B และ c ได้จากสมการ (2.21)และสามารถหาค่าประมาณของค่า q_x ที่ปรับแล้วได้ดังนี้

$$q_x'' \quad = \quad 1-\exp\left[1\int_0^1 \mu_{x+s}\, ds\right]$$

$$= \quad 1-\exp\left[\frac{-B}{\ln c} c^x (c-1)\right] \tag{2.25}$$

<u>การปรับค่าโดยใช้ส่วนโค้งพหุนามองศาสาม</u>

การปรับค่าโดยใช้ส่วนโค้งพหุนามองศาสาม เป็นวิธีการปรับค่าโดยพิจารณาการเคลื่อนไหวของค่า q'_x แล้วทำการแบ่งช่วงอายุที่สนใจศึกษาออกเป็นช่วงย่อย ๆ และแทนแต่ละช่วงย่อยนี้ด้วยส่วนโค้งพหุนามองศาสาม (Cubic Splines) ซึ่งแต่ละส่วนโค้งจะเชื่อมติดกันอย่างต่อเนื่องและราบเรียบที่จุดแบ่งใด ๆ

กำหนดให้ช่วง $[a,b]$ เป็นช่วงอายุที่สนใจศึกษา และทำการแบ่งช่วง $[a,b]$ นี้ออกเป็น $(n+1)$ ช่วงที่อายุ $x = k_1 , k_2 , \ldots , k_n$ เมื่อ k เป็นจำนวนจริงใด ๆ และ n เป็นจำนวนเต็มบวก

การปรับค่าจะทำโดยการแทนแต่ละช่วงย่อยด้วยพหุนามองศาสาม เพื่อให้ได้ค่าประมาณของค่า q_x ที่ปรับแล้วดังนี้

$$q''_x = \begin{cases} p_0(x) & ,a \le x \le k_1 \\ p_i(x) & ,k_i \le x \le k_{i+1} \\ p_n(x) & ,k_n \le x \le b \end{cases} \qquad (2.26)$$

ภายใต้เงื่อนไขคือ

$$p_{i-1}(k_i) = p_i(k_i) \quad ,$$
$$p'_{i-1}(k_i) = p'_i(k_i) \quad ,$$

และ $$p''_{i-1}(k_i) = p''_i(k_i) \quad ,i = 1, 2, \ldots , n$$

โดยที่ $p_0(x) = c_1 + c_2 x + c_3 x^2 + c_4 x^3$

และ $p_i(x) = p_0(x) + c_5 (x-k_1)^3 + \ldots + c_{i+4} (x-k_i)^3 \quad , i = 1, 2 ,\ldots, n$

เมื่อ $p_i(x)$ คือพหุนามองศาสาม

การหาค่า $c_1 , c_2 , \ldots , c_{i+4}$ ทำได้โดยวิธีการกำลังสองน้อยที่สุด (Least Square Method) ซึ่งมีสมการคือ

$$SS=\sum_{x=a}^{h_1} w_x\left[q'_x-c_1-c_2x-c_3x^2-c_4x^3\right]^2+\sum_{x=h_1+1}^{h_2} w_x\left[q'_x-c_1-c_2x-c_3x^2-c_4x^3-c_5\left(x-k_1\right)^3\right]^2$$

$$+\ldots+\sum_{x=h_n+1}^{b} w_x\left[q'_x-c_1-\ldots-c_5\left(x-k_1\right)^3-\ldots-c_{n+4}\left(x-k_n\right)^3\right]^2 \qquad (2.27)$$

โดยที่ $w_x = \dfrac{m_x}{q'_x\left(1-q'_x\right)}$

เมื่อ w_x คือค่าถ่วงน้ำหนัก $\left(\text{Weights}\right)$ ที่อายุ x ปี

m_x คือขนาดตัวอย่างของข้อมูลที่นำมาศึกษาที่อายุ x ปี

$h_1,h_2,\ldots,h_n$ คือค่าจำนวนเต็มที่มากที่สุดของ x แต่น้อยกว่า (หรือเท่ากับ) ค่า $k_1,k_2,\ldots,k_n$ ตามลำดับ

โดยการหาอนุพันธ์ของฟังก์ชัน SS จากสมการ (2.27) เทียบกับค่า $k_1,k_2,\ldots,k_n$ ตามลำดับและให้สมการอนุพันธ์มีค่าเท่ากับศูนย์ ซึ่งสามารถแสดงในรูปเมตริกซ์ได้คือ

$$x'wxc = x'wq'_x \qquad (2.28)$$

โดยที่

$$x=\begin{bmatrix}
1 & a & a^2 & a^3 & 0 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
1 & h_1 & h_1^2 & h_1^3 & 0 & \cdots & 0 \\
1 & h_1+1 & \left(h_1+1\right)^2 & \left(h_1+1\right)^3 & \left(h_1+1-k_1\right)^3 & 0 \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
1 & h_n & h_n^2 & h_n^3 & \left(h_n-k_1\right)^3 & \cdots & 0 \\
1 & h_1+1 & \left(h_1+1\right)^2 & \left(h_n+1\right)^3 & \left(h_n+1-k_1\right)^3 & \cdots & \left(h_n+1-k_n\right)^3 \\
\vdots & \vdots & \vdots & \vdots & \vdots & & \vdots \\
1 & b & b^2 & b^3 & \left(b-k_1\right)^3 & \cdots & \left(b-k_n\right)^3
\end{bmatrix}$$

เป็นเมตริกซ์ขนาด $\left(v\times(n+4)\right)$

$$w = \begin{bmatrix} w_a & & & & O \\ \vdots & \ddots & & & \vdots \\ O & & & \ddots & w_b \end{bmatrix}$$ เป็นเมตริกซ์ขนาด $(v \times v)$

$$q'_x = \begin{bmatrix} q'_a \\ \vdots \\ q'_{b1} \end{bmatrix}$$ เป็นเมตริกซ์ขนาด $(v \times 1)$

$$c = \begin{bmatrix} c_1 \\ \vdots \\ c_{i+4} \end{bmatrix}$$ เป็นเมตริกซ์ขนาด $\left((i+4) \times 1\right)$

เมื่อ x' คือการสับเปลี่ยน (Transpose) ของเมตริกซ์ x

v คือจำนวนค่าของอายุในช่วง $[a,b]$

หลังจากได้ค่าแล้ว c_1 , c_2 , ... , c_{i+4} แล้วสามารถค่าประมาณของค่า q_x ที่ปรับแล้วได้จากสมการ (2.26)