

บทที่ 2

สถิติที่ใช้ในการวิจัย

ในบทนี้ กล่าวถึงวิธีการปรับแก้ค่าประมาณความน่าจะเป็นที่จะเสียชีวิต ซึ่งมี 3 วิธี คือ วิธีการปรับแก้แบบวิทเทคเกอร์ (Whittaker Graduation) วิธีการปรับแก้แบบเบส์เซียน (Bayesian Graduation) และวิธีการปรับแก้แบบอินครีซิงเบส์เซียน (Increasing Bayesian Graduation) ซึ่งทั้ง 3 วิธีดังกล่าวนี้ จะศึกษาเพื่อการปรับแก้ค่าประมาณความน่าจะเป็นที่จะเสียชีวิต (q_x') ที่ประมาณด้วยวิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด (Maximum Likelihood Estimation) โดยจะศึกษาภายใต้การแจกแจงของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต (Future Life Time, T) แจกแจงแบบเอกซ์โพเนนเชียล (Exponential Distribution) และการแจกแจงของระยะเวลาการถอนตัว (Withdrawal Time, W) 2 รูปแบบ คือการแจกแจงแบบสม่ำเสมอ (Uniform Distribution) และการแจกแจงแบบแกมมา (Gamma Distribution) โดยมีรายละเอียดดังนี้

1. วิธีการประมาณ q_x แบบภาวะน่าจะเป็นสูงสุด

วิธีการประมาณแบบภาวะน่าจะเป็นสูงสุด เป็นวิธีการประมาณ q_x วิธีหนึ่งที่ใช้สำหรับข้อมูลการประกันชีวิตแบบไม่สมบูรณ์ ซึ่งวิธีการนี้จะใช้ได้เมื่อตัวอย่างสุ่มมีการแจกแจงแบบมีพารามิเตอร์

กำหนดให้

$$\delta_i = \begin{cases} 1 & \text{เมื่อคนที่ } i \text{ เสียชีวิตในช่วง } (x, x+1] \\ 0 & \text{อื่นๆ} \end{cases}$$

$$\gamma_i = \begin{cases} 1 & \text{เมื่อคนที่ } i \text{ ถอนตัวออกจากช่วง } (x, x+1] \\ 0 & \text{อื่นๆ} \end{cases}$$

m เป็นจำนวนคน (ขนาดตัวอย่าง) ที่เริ่มต้นอายุ x ปี

d เป็นจำนวนคนที่เสียชีวิตในช่วง $(x, x+1)$

t_i เป็นระยะเวลาที่จะมีชีวิตอยู่จนกระทั่งออกจากกลุ่มของคนที่ i ซึ่งมีอายุเริ่มต้นเป็น x ปี สาเหตุการออกจากกลุ่ม คือ การเสียชีวิต การถอนตัวออกจากช่วง หรือการมีชีวิตอยู่ ณ สิ้นช่วงเวลาศึกษา

${}_i p_x^{(d)}$ เป็นความน่าจะเป็นที่คนอายุ x จะมีชีวิตอยู่ต่อไปอีกอย่างน้อย t_i ปี

${}_i p_x^{(w)}$ เป็นความน่าจะเป็นที่คนอายุ x จะยังคงอยู่ในช่วงที่ศึกษาอีกอย่างน้อย

t_i ปี

$\mu_{x+t_i}^{(d)}$ เป็นพลังของมรณะ

$\mu_{x-t_i}^{(w)}$ เป็นพลังของการถอนตัว

ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function, L) คือ

$$L = \prod_{i=1}^m {}_i p_x^{(d)} \cdot {}_i p_x^{(w)} \cdot (\mu_{x+t_i}^{(d)})^{\delta_i} \cdot (\mu_{x+t_i}^{(w)})^{\gamma_i} \quad (2.1)$$

กำหนดให้การเสียชีวิต และการถอนตัวเป็นเหตุการณ์ที่เป็นอิสระกัน ดังนั้นการประมาณค่าความน่าจะเป็นที่คนอายุ x จะเสียชีวิตภายใน 1 ปีข้างหน้า คือ ในช่วง $(x, x+1)$ เทอม ${}_i p_x^{(w)} \cdot (\mu_{x+t_i}^{(w)})^{\gamma_i}$ จะเป็นค่าคงที่ ดังนั้น สมการ (2.1) เขียนใหม่ได้ดังนี้

$$L = \prod_{i=1}^m {}_i p_x^{(d)} \cdot (\mu_{x+t_i}^{(d)})^{\delta_i} \quad (2.2)$$

ภายใต้ข้อสมมติ เมื่อพลังของมรณะมีค่าคงที่ (Constant Force of Mortality Assumption) กำหนดให้ $\mu_{x+t_i}^{(d)} = \mu_x^{(d)}$ จะได้ว่า

$$L = (\mu_x^{(d)})^d \prod_{i=1}^m e^{-\mu_x^{(d)} t_i} \quad (2.3)$$

จากสมการ (2.3) หาอนุพันธ์ $\frac{d \ln L}{d \mu_x^{(d)}} = 0$ จะได้ว่า

$$\mu_x^{(d)} = \frac{d}{\sum_{i=1}^m t_i} \quad (2.4)$$

และ เราสามารถหาค่าประมาณ q_x ได้ดังนี้

$$q_x' = 1 - \exp(-\mu_x^{(d)}) \quad (2.5)$$

2. วิธีการปรับแก้แบบวิทแทคเกอร์¹

การปรับแก้โดยวิธีนี้เป็น การนำหลักพื้นฐานของการปรับ คือ ความราบเรียบ (Smooth) และความพอดี (Fit) มารวมเข้าด้วยกัน โดยอาศัยหลักการถ่วงน้ำหนัก และการหาค่าผลต่างของค่าปรับที่ระดับต่าง ๆ และจากนั้นใช้วิธีการแก้สมการเพื่อหาค่าต่ำสุด ซึ่งวิธีการปรับนี้คิดค้นโดย อี ที วิทแทคเกอร์ (E T Whittaker) โดยมีรายละเอียดดังนี้

กำหนดให้ค่าอายุ x ในช่วงที่ศึกษา แทนด้วย $1, 2, \dots, n$ ทั้งนี้เพื่อให้ง่ายต่อการคำนวณ ดังนั้น สามารถเขียนสมการพื้นฐานได้ดังนี้

$$M = F + hS = \sum_{x=1}^n w_x (q_x'' - q_x')^2 + h \sum_{x=1}^{n-z} (\Delta^z q_x'')^2 \quad (2.6)$$

โดยที่ z เป็นค่าแสดงดีกรีของโพลีโนเมียล ซึ่งในการวิจัยครั้งนี้จะศึกษาที่ $z = 1, 2, 3$ และ 4

w_x เป็นค่าถ่วงน้ำหนัก

$$w_x = \frac{n_x}{q_x' (1 - q_x')}$$

h เป็นค่าเฉลี่ยของ w_x

$$h = \left(\sum_{x=1}^n w_x \right) / n$$

จากสมการ (2.6) หาค่า q_x'' ได้โดยการแก้สมการหาค่าต่ำสุด ดังนี้

กำหนด

$q' = [q'_1 \ q'_2 \ \dots \ q'_n]'$ เป็นเวกเตอร์ของค่าประมาณเบื้องต้น

$q'' = [q''_1 \ q''_2 \ \dots \ q''_n]'$ เป็นเวกเตอร์ของค่าประมาณที่ปรับแก้แล้ว

w เป็นเมตริกซ์ทแยงมุมขนาด $n \times n$ ซึ่งสมาชิกในแนวทแยงมุมคือค่า w_1, w_2, \dots, w_n

k_z เป็นเมตริกซ์ของสัมประสิทธิ์ทวินามลำดับ z ขนาด $(n-z) \times n$

จากสมการ (2.6) เขียนใหม่ให้อยู่ในรูปสมการของเมตริกซ์และเวกเตอร์ ได้ดังนี้

$$(q'' - q')' w (q'' - q') + h (k_z q'')' k_z q'' = (q'' - q')' w (q'' - q') + h (k_z q'')' k_z q'' \quad (2.7)$$

จากคุณสมบัติของเมตริกซ์ที่ว่า $y'y$ จะได้ผลลัพธ์เป็นผลบวกกำลังสองของสมาชิกของเวกเตอร์ y ดังนั้น สมการ (2.7) เขียนใหม่ได้ดังนี้

$$(w + h k_z' k_z) q'' = w q' \quad (2.8)$$

จากสมการ (2.8) เราสามารถหาค่า q''_x ได้โดยการแก้สมการเมตริกซ์ด้วยวิธีแฟกเตอร์ไรเซชัน (Factorization Method)

3. วิธีการปรับแก้แบบเบส์เซียน²

วิธีการปรับแก้แบบเบส์เซียนเป็นวิธีที่ค่อนข้างใหม่เมื่อเปรียบเทียบกับวิธีการปรับแก้แบบค่าเฉลี่ยเคลื่อนที่ถ่วงน้ำหนัก หรือวิธีการปรับแก้แบบวิทเทคเกอร์ วิธีการปรับแก้แบบเบส์เซียนที่จะกล่าวถึงนี้ พัฒนาขึ้นโดยไคเมดอร์ฟและโจนส์ (Kimeidorf and Jones) จึงอาจเรียกวิธีการปรับแก้แบบนี้ว่าวิธีของไคเมดอร์ฟ-โจนส์ (Kimeldorf - Jones Method) ซึ่งการปรับแก้โดยวิธีนี้นอกจากอาศัยข้อมูลที่ได้จากตัวอย่างที่นำมาศึกษาแล้ว ยังอาศัยข้อมูลหรือประสบการณ์ในอดีตมาใช้ด้วย ซึ่งจากข้อมูลหรือประสบการณ์ในอดีตจะทำให้ได้การแจกแจงก่อนการทดลอง (Prior Distribution) และจากการแจกแจงก่อนการทดลองและข้อมูลจาก

² Dick London "Graduation : The Revision of Estimate" Winsted, Connecticut : ACTEX Publications, 1985, p 71

ตัวอย่างที่นำมาศึกษา สามารถหาการแจกแจงหลังการทดลอง (Posterior distribution) ได้ และสามารถหาค่าที่ปรับแก้แล้วได้จากการแจกแจงดังกล่าวนี้

กำหนด

1. ค่าอายุ x ในช่วงที่ศึกษา เพื่อให้ง่ายต่อการคำนวณให้ $x = 1, 2, \dots, n$
2. Q_i เป็นตัวแปรสุ่มแทนความน่าจะเป็นที่คนอายุ x จะเสียชีวิตภายใน 1 ปี
3. $Q = [Q_1 \ Q_2 \ \dots \ Q_n]'$ เป็นเวกเตอร์สุ่ม (Random Vector)
4. $q = [q_1 \ q_2 \ \dots \ q_n]'$ เป็นเวกเตอร์ค่าจริงของ Q
5. Q'_i เป็นตัวแปรสุ่มแทนค่าประมาณความน่าจะเป็นที่คนอายุ x จะเสียชีวิตภายใน 1 ปี และกำหนดให้ Q'_i เป็นอิสระกัน
6. $Q' = [Q'_1 \ Q'_2 \ \dots \ Q'_n]'$
7. $q' = [q'_1 \ q'_2 \ \dots \ q'_n]'$

โดยวิธีนี้กำหนดให้ Q มีการแจกแจงแบบมัลติโนร์มัล (Multinormal Distribution) มีฟังก์ชันความหนาแน่นก่อนการทดลอง (Prior Density Function) ดังนี้

$$f(q) = k_1 \exp\left[-1/2(q - q^p)' A^{-1}(q - q^p)\right] \quad (2.9)$$

โดยที่ $k_1 = \left[(2\pi)^n |A|\right]^{-1/2}$

$q^p = [q_1^p \ q_2^p \ \dots \ q_n^p]$ เป็นเวกเตอร์ค่าเฉลี่ย (Mean Vector) แสดงค่าประมาณที่ดีที่สุดของ q กรณีที่ไม่มีข้อมูลจากการทดลอง ซึ่งในการวิจัยครั้งนี้ใช้ค่า q^p จากตารางมรณะไทย 2529

A เป็นเมตริกซ์ความแปรปรวนร่วม (Covarian Matrix) ขนาด $n \times n$

$$a_{ij} = p^2 r^{|i-j|} \quad , \quad p > 0, \quad 0 \leq r < 1$$

เมื่อ p^2 เป็นความแปรปรวนของ Q_i สำหรับทุกค่าของ i

$$p^2 = \sum_{i=1}^n \frac{q_i^p(1 - q_i^p)}{m_i} / n$$

r เป็นสัมประสิทธิ์สหสัมพันธ์ระหว่าง Q_i และ Q_j เมื่อ $|i - j| = 1$ ซึ่งในการวิจัยครั้งนี้จะศึกษาที่ $r = 0.6, 0.7, 0.8$ และ 0.9

และ กำหนดให้ $a_{ij} = 0$ สำหรับ $j > i + 10$ และ $j < i - 10$

กำหนดให้การแจกแจงที่มีเงื่อนไขของ Q' เมื่อกำหนด q มีการแจกแจงแบบ
มัลติโนร์มัล (Multinormal Distribution)

$$f(q' / q) = k_2 \exp \left[-\frac{1}{2} (q' - q)' B^{-1} (q' - q) \right] \quad (2.10)$$

$$\text{โดยที่ } k_2 = [(2\pi)^n |B|]^{-1/2}$$

B เป็นเมตริกซ์ความแปรปรวนร่วม (Covariance Matrix) เป็น
เมตริกซ์ทแยงมุมขนาด $n \times n$ ซึ่งสมาชิกในแนวทแยงมุมเป็นความแปรปรวนของ Q'_i

$$b_{ii} = \frac{q_i^p (1 - q_i^p)}{n_i}$$

จากสมการที่ (2.9) และ (2.10) สามารถเขียนฟังก์ชันความหนาแน่นหลัง
การทดลอง (Posterior Density Function) ได้ดังนี้

$$f(q / q') = k_4 \exp \left\{ -\frac{1}{2} \left[(q - q^p)' A^{-1} (q - q^p) + (q' - q)' B^{-1} (q' - q) \right] \right\} \quad (2.11)$$

$$\text{โดยที่ } k_4 = \frac{k_1 k_2}{k_3} \quad \text{เมื่อ } k_3 = f(q')$$

จากสมการ (2.11) สามารถเขียนใหม่ได้ดังนี้

$$\begin{aligned} f(q / q') &= k_4 \exp \left\{ -\frac{1}{2} \left[(q')' A^{-1} q + q'^p A^{-1} q^p - q'^p A^{-1} q - (q')' A^{-1} q^p \right. \right. \\ &\quad \left. \left. + (q')' B^{-1} (q') + (q)' B^{-1} (q) - (q)' B^{-1} (q') - (q')' B^{-1} q \right] \right\} \\ &= k_5 \exp \left\{ -\frac{1}{2} \left[(q)' (A^{-1} + B^{-1}) q - (q)' (A^{-1} q^p + B^{-1} (q')) \right. \right. \\ &\quad \left. \left. - q'^p A^{-1} + (q')' B^{-1} \right] q \right\} \end{aligned}$$

$$\text{เมื่อ } k_5 = k_4 \exp \left[-\frac{1}{2} \left(q'^p A^{-1} q^p + (q')' B^{-1} (q') \right) \right]$$

$$\text{กำหนดให้ } V = (A^{-1} + B^{-1})^{-1} (B^{-1} (q') + A^{-1} q^p)$$

$$\text{และ } C = (A^{-1} + B^{-1})^{-1}$$

ดังนั้น

$$\begin{aligned} f(q / q') &= k_5 \exp \left\{ -\frac{1}{2} \left[(q)' C^{-1} q - (q)' C^{-1} V - (q'^p A^{-1} + (q')' B^{-1}) q \right] \right\} \\ &= k_5 \exp \left\{ -\frac{1}{2} \left[(q)' C^{-1} q - (q)' C^{-1} V - V C^{-1} q + V C^{-1} V - V C^{-1} V \right] \right\} \end{aligned}$$

จัดรูปสมการใหม่ จะได้

$$f(q/q') = k_0 \exp\left[-\frac{1}{2}(q-V)'C^{-1}(q-V)\right] \quad (2.12)$$

เมื่อ $k_0 = k_3 \exp\left[-\frac{1}{2}V'C^{-1}V\right]$

จากสมการ (2.12) จะเห็นได้ว่า $f(q/q')$ แจกแจงแบบมัลติโนร์มัล (Multinormal Distribution) ที่มี V เป็นเวกเตอร์ค่าเฉลี่ย และ C เป็นเมตริกซ์ความแปรปรวนร่วม

ดังนั้น เราสามารถหาค่า q^* ได้ดังนี้

$$\begin{aligned} q^* &= E[Q/q'] \\ &= (A^{-1} + B^{-1})^{-1}(B^{-1}(q') + A^{-1}q^p) \end{aligned} \quad (2.13)$$

3. วิธีการปรับแก้แบบอินครีซีทีงเบสส์เซียน³

เป็นวิธีการปรับแก้ที่นำความรู้ที่ว่าเมื่ออายุประมาณ 30 ปีขึ้นไป อัตราภาระของคนเราจะมีความเพิ่มขึ้นเมื่ออายุเพิ่มขึ้นมาใช้พิจารณาในการปรับแก้ด้วย โดยการกำหนดข้อจำกัดขึ้นมา และการปรับแก้โดยวิธีนี้จะทำภายใต้สมมติฐานพลังของมรณะมีค่าคงที่ในช่วงอายุหนึ่งๆ และจะหาค่าที่ปรับของพลังมรณะ จากนั้นสามารถหาค่า q_x^* ได้จากค่าพลังของมรณะที่ปรับแล้ว

กำหนดช่วงอายุที่ศึกษา k ช่วง และจะศึกษาที่ค่าอายุ $x, x+1, \dots, x+k-1$

ภายใต้สมมติฐานพลังของมรณะ (Force of Mortality, $\mu(y)$) เป็นค่าคงที่ในหนึ่งช่วงอายุใดๆ ดังนั้นสำหรับ $j = 1, 2, \dots, k$ จะได้ว่า

$$\mu(y) = \theta_j, \quad x+j-1 \leq y \leq x+j$$

กำหนด

$$\theta = (\theta_1, \theta_2, \dots, \theta_k) \in R_j$$

$$\text{เมื่อ } R_j = \{\theta, \theta_1 < \theta_2 < \dots < \theta_k\}$$

$$d_j = \text{จำนวนคนที่เสียชีวิตระหว่างอายุ } x+j-1 \text{ และ } x+j$$

$$e_j = \text{ระยะเวลาที่อยู่ภายใต้การสังเกต (Exposure) ระหว่างอายุ } x+j-1 \text{ และ}$$

$x+j$

ฟังก์ชันภาวะน่าจะเป็น (Likelihood Function, L) ของ θ กำหนดโดย

$$L(\theta) \propto \prod_{j=1}^k (\theta_j^{d_j} e^{-e_j \theta_j}) \quad (2.14)$$

กำหนด

$$\begin{aligned}\theta_1 &= \phi_1 \\ \theta_2 &= \phi_1 + \phi_2 \\ &\vdots \\ \theta_k &= \phi_1 + \phi_2 + \dots + \phi_k\end{aligned}\quad (2.15)$$

$$a = (a_1, \dots, a_k)$$

$$r = (r_1, \dots, r_k)$$

$$\phi = (\phi_1, \dots, \phi_k)$$

$g(\phi / a, r)$ เป็นฟังก์ชันความหนาแน่นของการแจกแจงแบบแกมมา ซึ่งมี a, r เป็นพารามิเตอร์

ฟังก์ชันความหนาแน่นก่อนการทดลอง (Prior Density Function) ของ ϕ กำหนดโดย

$$\text{prior}(\phi) = g(\phi / a, r) \quad (2.16)$$

จากสมการ (2.14) และ สมการ (2.16) สามารถเขียนฟังก์ชันความหนาแน่นหลังการทดลอง (Posterior Density Function) ได้ดังนี้

$$\begin{aligned}\text{post}(\phi) &\propto \prod_{j=1}^k (\theta_j^{d_j} e^{-e, \theta_j}) g(\phi / a, r) \\ &\propto \prod_{j=1}^k \theta_j^{d_j} \phi_j^{a_j-1} e^{-b_j \phi_j} \quad , \phi_j > 0, j=1, \dots, k\end{aligned}\quad (2.17)$$

$$\text{เมื่อ } b_j = r_1 + e_1 + \dots + e_k$$

$$\text{กำหนดให้ } \phi'' = (\phi_1'', \phi_2'', \dots, \phi_k'')$$

$$\theta'' = (\theta_1'', \theta_2'', \dots, \theta_k'')$$

จากสมการ (2.17) เราสามารถหาค่า ϕ'' ได้จาก

$$\frac{\partial}{\partial \phi_j} \ln \text{post}(\phi) = 0 \quad , j = 1, \dots, k$$

ซึ่งสามารถเขียนใหม่ได้ดังนี้

$$\sum_{j=1}^k \frac{d_j}{\theta_j} + \frac{a_i - 1}{\phi_j} - b_j = 0, i = 1, \dots, k \quad (2.18)$$

และใช้ขบวนการทำซ้ำ (Iteration Method) จนกระทั่งความแตกต่างของค่า ϕ'' ที่คำนวณได้ในรอบสุดท้ายกับรอบก่อนหน้านั้นมีค่าน้อยกว่า 0.01 หลังจากนั้นคำนวณหาค่า θ'' โดยใช้ความสัมพันธ์ในสมการ (2.15) และสามารถคำนวณหาค่าที่ปรับแก้ q_x'' ได้ดังนี้

$$q_{x+j-1}'' = 1 - \exp(-\theta_j'') \quad (2.19)$$

การหาค่าพารามิเตอร์ของการแจกแจงก่อนการทดลอง a และ r สามารถหาได้ดังนี้ กำหนดให้

1. $\theta^p = (\theta_1^p, \dots, \theta_k^p)$ เป็นค่าของ θ ก่อนการทดลอง
2. $\phi^p = (\phi_1^p, \dots, \phi_k^p)$ เป็นค่าของ ϕ ก่อนการทดลองซึ่งกำหนดโดย

$$\theta_i^p = \phi_1^p + \dots + \phi_j^p$$
3. $v_i^p = \text{Var}(\theta_i^p)$ และ $v^p = (v_1^p, \dots, v_k^p)$
4. $v_i' = \text{Var}(\theta_i')$ ซึ่งคำนวณจาก

$$v_i' = \exp(\theta_i' - 1) / e_i$$
5. $\sum_{i=1}^k v_i^p = m \sum_{i=1}^k v_i'$ (2.20)
6. ค่า m ในการวิจัยครั้งนี้ผู้วิจัยจะศึกษาที่ $m = 1, 5$ และ 25
7. กำหนดให้ $a_1 = a_2 = \dots = a_n = \alpha$, $\alpha > 1$

กำหนดให้ ϕ^p เป็นค่าฐานนิยมก่อนการทดลองของ ϕ นั่นคือ

$$\phi_i^p = (\alpha - 1) / r_i$$

ดังนั้น

$$\text{Var}(\phi_i) = \frac{\alpha}{r_i^2} = (\phi_i^p)^2 \alpha / (\alpha - 1)^2$$

สมการ (2.20) จะสามารถเขียนได้เป็น

$$\alpha / (\alpha - 1)^2 \sum_{i=1}^k h_i (\phi_i^p)^2 = m \sum_{i=1}^k v_i' \quad (2.21)$$

เมื่อ $h_i = k - i + 1$

จากการแก้สมการ (2.21) จะได้

$$\alpha = 1 + u + \sqrt{u(2 + u)} \quad (2.22)$$

$$\text{เมื่อ } u = \sum_{i=1}^k h_i (\phi_i^p)^2 / \left[2m \sum_{i=1}^k v_i' \right]$$

$$\text{และ } r_i = (\alpha - 1) / \phi_i^p \quad , i = 1, \dots, k$$

(2.23)