บทที่ 2

สถิติที่ใช้ในการวิจัย

ในการศึกษาเพื่อหาวิธีการประมาณค่าความน่าจะเป็นที่จะอยู่รอด ซึ่งมีด้วยกัน 4 วิธีคือ วิธีทางคณิตศาสตร์ประกันภัย ตัวประมาณ เอ ตัวประมาณ บี และตัวประมาณ ซี จะศึกษาภายใต้การแจกแจงของระยะเวลาที่จะมีชีวิตอยู่ต่อไปในอนาคต (Future Lifetime) T สองรูปแบบ คือ การแจกแจงแบบไวบูลล์ และการแจกแจงแบบกอมเพิรตซ์ และระยะเวลาการออกจากช่วงก่อนสิ้นสุดการศึกษา (Withdrawal Time) W อีกสองรูปแบบคือ การแจกแจงแบบสม่ำเสมอ และการแจกแจงแบบเบต้า โดยกำหนดให้

N_x เป็น จำนวนคน (ขนาดตัวอย่าง) ที่เริ่มต้นอายุ X ปี

m_x เป็น จำนวนคนที่ถูกสังเกตจนถึงสิ้นช่วงการศึกษา ที่เริ่มต้นอายุ X ปี

n_x เป็น จำนวนคนที่ออกจากกลุ่ม (Withdraw) ไปก่อนที่จะสิ้นช่วงที่ศึกษาที่เริ่มต้นอายุ X ปี

s_x เป็น จำนวนคนจากกลุ่มคน m_x ที่อยู่รอดจนถึงอายุ X+1 ปี

w_x เป็น จำนวนคนจากกลุ่มคน n_x ที่อยู่รอดจนถึงอายุ X+1 ปี

d_x เป็น จำนวนคนจากกลุ่มคน m_x ที่เสียชีวิตก่อนถึงอายุ X+1 ปี

d'_x เป็น จำนวนคนจากกลุ่มคน n_x ที่เสียชีวิตก่อนถึงอายุ X+1 ปี

D_x เป็น จำนวนคนทั้งหมดที่เริ่มต้นอายุ X ปีและเสียชีวิตก่อนถึงอายุ X+1 ปี

โดย N_x สามารถแบ่งออกได้เป็น 2 กลุ่ม คือ m_x และ n_x โดยในแต่ละกลุ่ม จะแบ่งเป็นกลุ่มคนที่อยู่รอดจนถึงสิ้นช่วงที่ศึกษา และ กลุ่มคนที่เสียชีวิตในระหว่างช่วงที่ศึกษา สามารถสรุปได้ดังนี้

กลุ่มคน	จำนวนทั้งหมด	กลุ่มคนที่ถูกสังเกตจนถึงสิ้นช่วงการศึกษา	กลุ่มคนที่ออกจากกลุ่ม ไปก่อนที่จะสิ้นช่วงที่ศึกษา
ทั้งหมด	N_x	m_x	n_x
อยู่รอด	$s_x + w_x$	s_x	w_x
เสียชีวิต	D_x	d_x	d'_x

นั่นคือ

$$
\begin{aligned}
N_x &= s_x + w_x + D_x \\
m_x &= s_x + d_x \\
n_x &= w_x + d'_x \\
D_x &= d_x + d'_x
\end{aligned}
$$

พิจารณากลุ่มคน m_x ซึ่งจะมี p_x เป็นความน่าจะเป็นที่จะอยู่รอด และ $q_x = 1 - p_x$ ในช่วง (X,X+1) และ s_x เป็นจำนวนคนที่อยู่รอดจนถึงสิ้นช่วง จะมีการแจกแจงแบบทวินาม (Binomial Distribution) คือ

$$C_1 p_x^{s_x}(1-p_x)^{d_x}$$

พิจารณา กลุ่มคน w_x สมมติให้

$p_x\left(\frac{1}{2}\right)$ เป็นฟังก์ชันความน่าจะเป็นที่คนที่ออกจากกลุ่มจะเสียชีวิตภายในระยะเวลา 1 ปี

ดังนั้น การแจกแจงความน่าจะเป็นของ w_x คือ

$$C_2\left[p_x\left(\tfrac{1}{2}\right)\right]^{w_x}\left[1-p_x\left(\tfrac{1}{2}\right)\right]^{d'_x}$$

จะได้ ฟังก์ชันภาวะน่าจะเป็นสูงสุด (Likelihood Function) ของกลุ่มคน N_x เป็น

$$L(x;p_x) = p_x^{s_x}(1-p_x)^{d_x}\left[p_x\left(\tfrac{1}{2}\right)\right]^{w_x}\left[1-p_x\left(\tfrac{1}{2}\right)\right]^{d'_x}$$

ซึ่งในบทนี้จะกล่าวถึงวิธีการประมาณต่าง ๆ เป็นลำดับ ดังต่อไปนี้

1. <u>วิธีทางคณิตศาสตร์ (The Actuarial Method)</u>

วิธีนี้จะไม่แบ่งแยกคนออกเป็น 2 กลุ่ม และเป็นวิธีการประมาณค่า p_x อย่างง่ายที่สุด เพราะขั้นตอนการคำนวณไม่ยุ่งยากซับซ้อน เพียงแต่อาศัยหลักการทางสถิติในการคำนวณค่าความน่าจะเป็นคือ จำนวนคนที่อยู่รอดหารด้วยจำนวนผู้เสี่ยงภัยทั้งหมด นั่นคือ

$$p_x^o \quad = \quad 1 - \frac{\text{จำนวนผู้ที่เสียชีวิตก่อนสิ้นช่วงศึกษา}}{\text{จำนวนผู้เสี่ยงภัย}}$$

โดยสมมติว่า แต่ละคนที่จะออกจากช่วงก่อนสิ้นสุดการศึกษาจะออกจากช่วง ณ จุดกลางช่วง รูปแบบของการประมาณคือ

$$p_x^o = 1 - \frac{D_x}{N_x - \frac{1}{2}w_x} \tag{2.1}$$

2. <u>ตัวประมาณ เอ (Estimator A)</u>[1]

มีการกำหนดรูปแบบความน่าจะเป็น $p_x\left(\frac{1}{2}\right)$ โดยให้ขึ้นกับเวลาที่ออกจากกลุ่ม ภายใต้สมมติฐาน ที่ว่าการออกจากช่วงเกิดขึ้นอย่างสุ่มในช่วง (X,X+1) จะได้ว่า

$$p_x\left(\tfrac{1}{2}\right) = \int_x^{x+1} \exp\left(-\int_x^t \mu(\tau)d(\tau)\right)dt \tag{2.2}$$

ถ้ากำลังของการมรณะ (The Force of Mortality) เป็นค่าคงที่ในช่วง (X,X+1) $\mu(\tau) = \mu_x$ จะได้

$$p_x\left(\tfrac{1}{2}\right) = \int_x^{x+1} \exp\left(-(t-x)\mu_x\right)dt$$

$$= \frac{1-e^{-\mu_x}}{\mu_x} \tag{2.3}$$

$$= -\frac{(1-p_x)}{\ln p_x}$$

---

[1] Chiang, C.L., A stochastic Study of the Life Table and its Applications : III . The Follow-up Study with the Consideration of Competing Risks. Biometrics, 17, 1961

Drolette, M.E., The Effect of Incomplete Follow-up, Biometrics, 31, 1975

จะได้

$$L_a(x;p_x) = p_x^{s_x}(1-p_x)^{d_x+w_x}(\ln p_x)^{-n_x}\left[(1-p_x)+\ln p_x\right]^{d'_x} \tag{2.4}$$

$$\begin{aligned} \ln L_a(x;p_x) &= s_x \ln p_x + (d_x+w_x)\ln(1-p_x) \\ &\quad - n_x \ln(\ln p_x) + d'_x \ln\left[(1-p_x)+\ln p_x\right] \end{aligned} \tag{2.5}$$

$$\begin{aligned} \frac{\partial \ln L_a}{\partial p_x} &= \frac{s_x}{p_x} - \frac{(d_x+w_x)}{(1-p_x)} - \frac{n_x}{p_x(\ln p_x)} + \frac{d'_x(1-p_x)}{\left[(1-p_x)+\ln p_x\right]p_x} \\ &= 0 \end{aligned} \tag{2.6}$$

การหาค่า p_x จากสมการข้างต้นนี้ จะทำได้โดยวิธีนิวตัน-ราพซ์สัน (Newton-Raphson Method)

3. <u>ตัวประมาณ บี (Estimator B)</u>[2]

มีการกำหนดรูปแบบความน่าจะเป็น $p_x\left(\frac{1}{2}\right)$ โดยให้ขึ้นกับเวลาที่ออกจากกลุ่มภายใต้สมมติฐาน ที่ว่าการออกจากช่วงเกิดขึ้นกลางช่วง

$$p_x\left(\tfrac{1}{2}\right) = \exp\left(-\frac{1}{2}\mu_x\right) = p_x^{1/2} \tag{2.7}$$

จะได้

$$L_b(x;p_x) = p_x^{s_x+(1/2)w_x}(1-p_x)^{d_x}\left(1-p_x^{1/2}\right)^{d'_x} \tag{2.8}$$

แก้สมการ $\dfrac{\partial \ln L_b(x;p_x)}{\partial p_x} = 0$ จะได้

---

[2] Chiang, C.L., Astochastic Study of the Life Table and its Applications : III . The Follow-up Study with the Consideration of Competing Risks, Biometrics, 17, 1961

Drolette, M.E., The Effect of Incomplete Follow-up, Biometrics, 31, 1975

$$\left(N_x - \frac{1}{2}n_x\right)\hat{p}_x + \frac{1}{2}d'_x\hat{p}_x^{1/2} - \left(s_x + \frac{1}{2}w_x\right) = 0 \tag{2.9}$$

$$\hat{p}_x = \left[\frac{-\frac{1}{2}d'_x + \sqrt{\frac{1}{4}d'^2_x + 4\left(N_x - \frac{1}{2}n_x\right)\left(s_x + \frac{1}{2}w_x\right)}}{2\left(N_x - \frac{1}{2}n_x\right)}\right]^2 \tag{2.10}$$

4. ตัวประมาณ ซี (Estimator C)[3]

กำหนดให้ t_j เป็นเวลาที่คนที่ j ตายในช่วง (X,X+1) ; $j = 1,...,D_x$ จะได้

$$f(t_j;x) = e^{-t_j\mu_x}\mu_x$$

$$= -p_x^{j}\ln p_x \qquad ,0 \le t_j \le 1 \quad ; \quad j = 1,\ldots,D_x$$

กำหนดให้ τ_i เป็นเวลาที่คนที่ i ออกจากช่วงในช่วง (X,X+1) ; $i = 1,...,w_x$ จะได้

$$g(\tau_i,x) = e^{-\tau_i\mu_x}$$

$$= p_x^{\tau_i} \qquad , 0 \le \tau_i \le 1 \quad ; \quad i = 1,\ldots,w_x$$

ดังนั้น ฟังก์ชันภาวะน่าจะเป็น (The Likelihood Function) เป็น

$$\begin{aligned} L_c(x;p_x) &= p_x^{s_x}\left(\prod_{i=1}^{w_x} p_x^{\tau_i}\right)\prod_{j=1}^{D_x}\left(p_x^{t_j}\ln p_x\right) \\ &= p_x^{T_x}\left(\ln p_x\right)^{D_x} \end{aligned} \tag{2.11}$$

เมื่อ $$T_x = s_x + \sum_{i=1}^{w_x}\tau_i + \sum_{j=1}^{D_x}t_j \tag{2.12}$$

[3] Chiang, C.L., A stochastic Study of the Life Table and its Applications : III . The Follow-up Study with the Consideration of Competing Risks, Biometrics, 17, 1961

$$\ln L_c(x;p_x) = T_x \ln p_x + D_x \ln(\ln p_x) \tag{2.13}$$

$$\frac{\partial \ln L_c}{\partial p_x^c} = \frac{T_x}{p_x^c} + \frac{D_x}{p_x^c \ln p_x^c} = 0 \tag{2.14}$$

$$T_x \ln p_x^c + D_x = 0 \tag{2.15}$$

$$\begin{aligned} p_x^c &= \exp\left(-\frac{D_x}{T_x}\right) \\ &= e^{-\tilde{\mu}_x^c} \end{aligned} \tag{2.16}$$

เมื่อ $\tilde{\mu}_x^c = \frac{D_x}{T_x}$