บทที่ 2

ฟังก์ชั่นของเวลาการอยู่รอด

(Function of Survival Time)

ในการอธิบายการแจกแจงการอยู่รอด (survival distribution) นั้น ต้องใช้ฟังก์ชั่นของเวลาการอยู่รอด คือ ฟังก์ชั่นความหนาแน่น, ฟังก์ชั่นการอยู่รอด และฟังก์ชั่นการสูญเสีย ทั้งนี้เพราะ ฟังก์ชั่นเหล่านั้นสามารถบอกถึงรูปแบบของการแจกแจงการอยู่รอดได้ ดังกล่าวแล้วในตอนที่ 1.1 และถ้าทราบฟังก์ชั่นใดฟังก์ชั่นหนึ่งใน 3 ฟังก์ชั่นนี้แล้ว ก็จะสามารถหาอีก 2 ฟังก์ชั่นได้ จากความสัมพันธ์ของฟังก์ชั่น (ดูตอนที่ 2.2)

ในทางปฏิบัติแล้ว ฟังก์ชั่นของเวลาการอยู่รอดทั้ง 3 ฟังก์ชั่นข้างต้น สามารถประมาณได้จากตัวอย่าง ซึ่งสามารถทำได้ทั้งวิธีพาราเมตริกและวิธีนอนพาราเมตริก ดังกล่าวแล้วข้างต้น

2.1 <u>นิยาม</u>

ให้ตัวแปรสุ่ม T แทนเวลาการอยู่รอด และให้ f(t) และ F(t) แทนฟังก์ชั่นความหนาแน่น (density function) และฟังก์ชั่นการแจกแจง (distribution function) ของเวลาการอยู่รอด ตามลำดับ

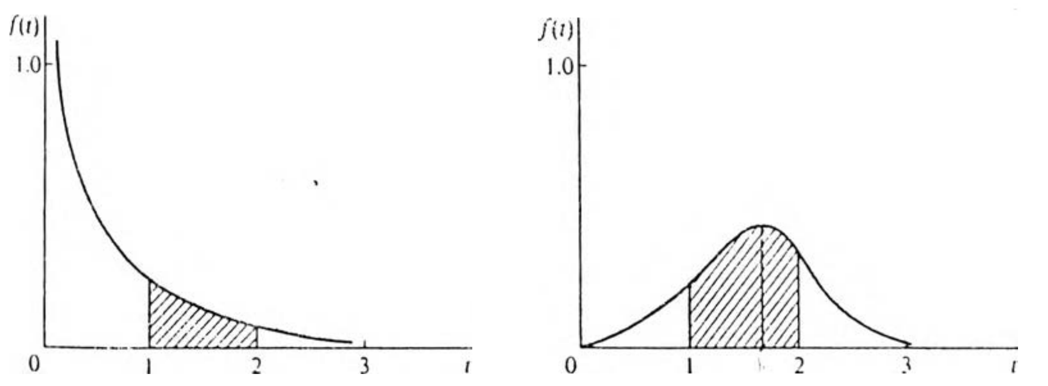
(1) ฟังก์ชั่นความหนาแน่น (density function) ตัวแปรสุ่ม T เป็นตัวแปรสุ่มแบบต่อเนื่อง ดังนั้น ฟังก์ชั่นความหนาแน่นของ T มีค่าเท่ากับลิมิตของความน่าจะเป็นที่แต่ละหน่วยตัวอย่างสูญเสียในช่วงเวลาสั้น ๆ จาก t ถึง $t + \Delta t$ ต่อหน่วยเวลา Δt หรือ เป็นความน่าจะเป็นของการสูญเสียในช่วงเวลา t ถึง $t + \Delta t$ ต่อหน่วยเวลา Δt นั่นคือ

$$f(t) = \lim_{\Delta t \to 0} P(\text{แต่ละหน่วยตัวอย่างสูญเสียในช่วง } (t, t + \Delta t)) / \Delta t \qquad (2.1)$$

ซึ่งมีคุณสมบัติดังนี้ (ก) $f(t) \geq 0$ เมื่อ $t \geq 0$ และ $f(t) = 0$ เมื่อ $t < 0$

(ข) $\int_{-\infty}^{\infty} f(t)\, dt = 1$

กราฟของ $f(t)$ เรียกว่า เส้นโค้งความหนาแน่น (density curve) ใช้ในการหาค่าอัตราหรือความถี่สูงสุดของการสูญเสีย และบอกสัดส่วนของการสูญเสียในช่วงเวลาใดเวลาหนึ่ง ดังแสดงในรูปที่ 1



(ก) (ข)

รูปที่ 1 แสดงเส้นโค้งความหนาแน่น

ในรูปที่ 1 (ก) อัตราสูงสุดของการสูญเสียจะอยู่ที่จุดเริ่มต้นของเวลา และอัตราจะลดลงเมื่อเวลาเพิ่มขึ้น ในรูปที่ 1 (ข) ความถี่สูงสุดของการสูญเสียจะอยู่ประมาณจุด 1.7 และสัดส่วนของตัวอย่างที่สูญเสียระหว่างเวลา 1 ถึง 2 จะเท่ากับพื้นที่ส่วนที่แรเงาระหว่างเส้นโค้งกับแกน t

(2) ฟังก์ชั่นการอยู่รอด (survival function) ให้แทนด้วย S(t) มีค่าเท่ากับความน่าจะเป็นที่แต่ละหน่วยตัวอย่างมีการอยู่รอดนานกว่าเวลา t นั่นคือ

$$S(t) = P\ (\text{แต่ละหน่วยตัวอย่างมีการอยู่รอดนานกว่าเวลา } t)$$
$$= P\ (T>t) \qquad (2.2)$$
$$= 1 - F(t)$$

มีคุณสมบัติดังนี้ (ก) S(t) เป็นฟังก์ชั่นไม่เพิ่ม (nonincreasing function)

นั่นคือ $S(t) > S(t+x)$ เมื่อ $x > 0$

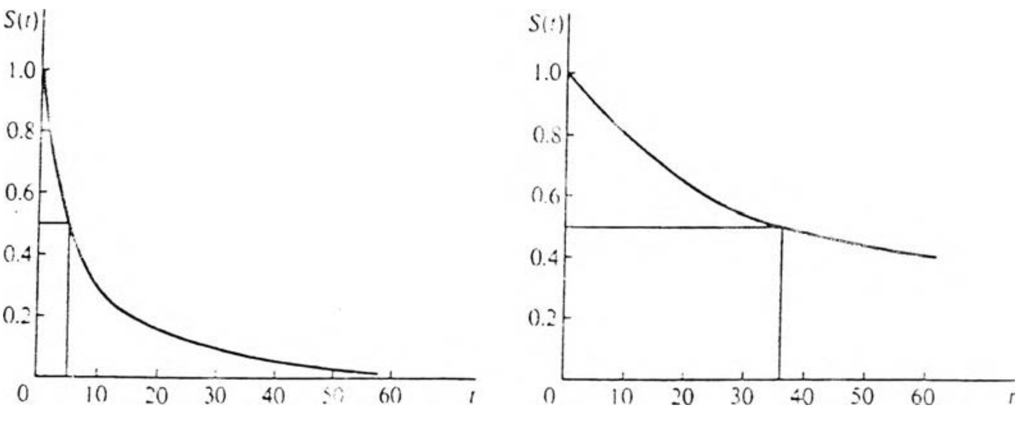
(ข) S(t) เป็นฟังก์ชั่นต่อเนื่องของ t

(ค) $S(t) = 1$ เมื่อ $t = 0$

(ง) $S(t) = 0$ เมื่อ $t = \infty$

ดังนั้น ฟังก์ชั่นการอยู่รอดที่เวลาอย่างน้อย 0 คือ 1 และที่เวลาอนันต์ (∞) คือ 0 และกราฟของ S(t) เรียกว่า เส้นโค้งการอยู่รอด (survival curve) ในรูปที่ 2 (ก) จะเป็นรูปที่มีอัตราการอยู่รอดต่ำหรือเวลาการอยู่รอดสั้น รูปที่ 2 (ข) เป็นรูปที่มีอัตราการอยู่รอดสูงหรือเวลาการอยู่รอดยาว

ฟังก์ชั่นการอยู่รอดหรือเส้นโค้งการอยู่รอด ใช้ในการหาค่ามัธยฐานการอยู่รอด (median survival time) หรือเปอร์เซ็นไทล์ที่ 50 และเปอร์เซ็นไทล์อื่น ๆ และยังใช้ในการเปรียบเทียบการแจกแจงการอยู่รอดของ 2 กลุ่ม หรือมากกว่า เช่น ในรูปที่ 2 (ก) เป็นรูปแสดงเส้นโค้งของฟังก์ชั่นการอยู่รอดของคนไข้โรคมะเร็งในเม็ดเลือด ที่ไม่ได้รับการรักษา และรูปที่ 2 (ข) เป็นเส้นโค้งฟังก์ชั่นการอยู่รอดของคนไข้ที่ได้รับการรักษา ได้ค่ามัธยฐานการอยู่รอดประมาณ 5 และ 36 ตามลำดับ



รูปที่ 2 แสดงเส้นโค้งการอยู่รอด

(3) ฟังก์ชั่นการสูญเสีย (failure, hazard function) ให้แทนด้วย h(t) มีค่าเท่ากับลิมิตของความน่าจะเป็นที่แต่ละหน่วยตัวอย่างสูญเสียในช่วงเวลาสั้น ๆ จาก t ถึง t + Δt ต่อหน่วยเวลา Δt เมื่อแต่ละหน่วยตัวอย่างมีการอยู่รอด (อายุ) ถึง t นั่นคือ

$$h(t) = \lim_{\Delta t \to 0} P(\text{แต่ละหน่วยตัวอย่างที่มีอายุ } t \text{ สูญเสียในช่วง } (t, t + \Delta t))/\Delta t \quad (2.3)$$

ฟังก์ชั่นการสูญเสียเป็นอัตราการสูญเสียที่มีเงื่อนไข (conditional failure rate) ซึ่งเงื่อนไขก็คือ หน่วยตัวอย่างแต่ละหน่วยจะต้องมีการอยู่รอดจนถึงเวลา t (มีอายุเท่ากับ t)

ฟังก์ชันการสูญเสียมีคุณสมบัติ ดังนี้

(ก) $h(t) \geq 0$ เมื่อ $-\infty < t < \infty$

(ข) $\lim_{t \to -\infty} \int_{-\infty}^{t} h(t)\,dt = 0$ และ

$$\lim_{t \to \infty} \int_{-\infty}^{t} h(t)\,dt = \infty$$

ฟังก์ชันการสูญเสีย สามารถเขียนในเทอมของฟังก์ชันการแจกแจง $F(t)$ และฟังก์ชันความหนาแน่น $f(t)$ ได้ดังนี้

$$h(t) = f(t)/(1-F(t))$$

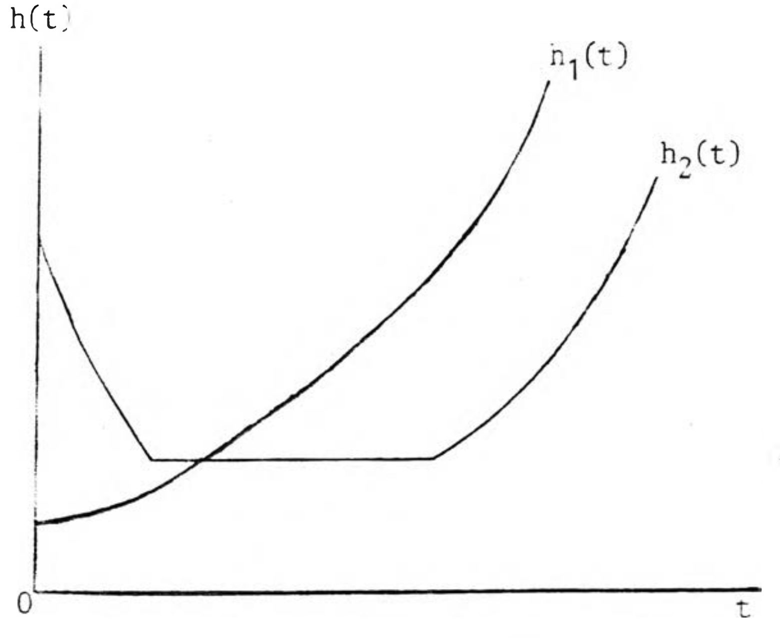
เนื่องจาก $S(t) = P(T>t)$

$= 1-P(T<t)$

$= 1-F(t)$

ดังนั้น $h(t) = f(t)/S(t)$

ฟังก์ชันการสูญเสีย เป็นอัตราการสูญเสีย ใช้หาความเสี่ยงในการสูญเสียของหน่วยตัวอย่าง และฟังก์ชันการสูญเสียอาจจะเป็นฟังก์ชันเพิ่ม (increase) หรือเป็นฟังก์ชันลด (decrease) หรือเป็นค่าคงที่ หรือมีหลายแบบรวมกัน ดังแสดงในรูปที่ 3 เส้นโค้ง $h_1(t)$ แสดงเส้นโค้งของฟังก์ชันการสูญเสีย (อัตราการสูญเสีย) ของคนไข้โรคมะเร็งในเม็ดเลือดที่ไม่ได้รับการรักษา จะเห็นว่า เมื่อเวลาเพิ่มขึ้นคนไข้จะมีความเสี่ยงต่อการสูญเสียชีวิตเพิ่มมากขึ้น และเส้นโค้ง $h_2(t)$ เป็นเส้นโค้งแสดงความเสี่ยงต่อการเสียชีวิตของคนทั่ว ๆ ไป เมื่อแรกเกิดและเมื่ออายุมากจะมีความเสี่ยงต่อการเสียชีวิตมาก และจะมีค่าคงที่ในระหว่างอายุประมาณ 18-40 ปี



รูปที่ 3 แสดงเส้นโค้งของฟังก์ชั่นการสูญเสีย

เนื่องจากว่าค่า $\Delta t.h(t)$ คือ สัดส่วนของหน่วยตัวอย่างที่มีการอยู่รอดถึง t ซึ่งจะสูญเสียในช่วงเวลาจาก t ถึง t + Δt ดังนั้น ฟังก์ชั่นการสูญเสีย คือ การวัดการสูญเสียในลักษณะของฟังก์ชั่นของอายุของแต่ละหน่วยตัวอย่าง ฟังก์ชั่นการสูญเสียจึงเป็นฟังก์ชั่นที่ให้ค่าความเสี่ยงของการสูญเสียต่อหน่วยเวลา

2.2 <u>ความสัมพันธ์ของฟังก์ชั่นการอยู่รอด</u>

ฟังก์ชั่นการอยู่รอดทั้ง 3 ชนิดที่กล่าวในตอนที่ 2.1 มีความสัมพันธ์ต่อกัน ดังนี้[1]

(ก) $f(t) = F'(t) = -S'(t)$

(ข) $h(t) = f(t)/S(t) = -S'(t)/S(t) = -d/dt \log S(t)$

(ค) $\int_0^t h(u)\,du = -\int_0^t S'(u)/S(u)\,du = -\log S(t)$

$\therefore S(t) = \exp(-\int_0^t h(u)\,du)$

และ $f(t) = h(t) \exp(-\int_0^t h(u)\,du)$

---

[1] Lee, <u>Statistical Methods for Survival Data Analysis.</u> pp. 16-17.