บทที่ 2
สถิติที่ใช้ในการวิจัย

ในบทนี้จะกล่าวถึงวิธีการประมาณฟังก์ชันการอยู่รอดของกรมธรรม์ประกันชีวิต สำหรับข้อมูลที่ มีปัญหาทางด้านการจัดเก็บข้อมูล เนื่องจากข้อจำกัดทางด้านเวลาและงบประมาณ ทำให้ข้อมูลที่จัดเก็บได้เป็นข้อมูลที่ไม่สมบูรณ์ โดยข้อมูลที่เก็บรวบรวมได้ส่วนใหญ่จะเป็นข้อมูลที่มีค่าถูกตัดทิ้งทางขวา ประเภทที่ 1 โดยศึกษาลักษณะของการแจกแจงการอยู่รอด 3 แบบคือ การแจกแจงแบบไวบูลล์ การแจกแจงแบบลอกนอร์มอล และการแจกแจงแบบพาเรโต โดยกำหนดให้แต่ละการแจกแจงมีค่าเฉลี่ยเท่ากับ 6 เนื่องจากการศึกษาข้อมูลของการประกันชีวิตในอดีตของกรมธรรม์ที่มีระยะเวลาการชำระเบี้ยประกันภัย 20 ปี พบว่าผู้เอาประกันจะชำระเบี้ยประกันภัยโดยเฉลี่ยประมาณ 6 ปี

2.1 ทฤษฎีพื้นฐาน

2.1.1 ประเภทของการถูกตัด (Type of Censoring)

1. การถูกตัดประเภทที่ 1 (Type I Censoring)

ข้อมูลที่ถูกตัดประเภทที่ 1 เกิดขึ้นเนื่องจากมีการกำหนดเวลาการสิ้นสุดการเก็บข้อมูลไว้ล่วงหน้า (Fixed Censoring Time) ด้วยค่าคงที่ Tc ตัวอย่างเช่นการศึกษาระยะเวลาคงอยู่ของกรมธรรม์ที่กำหนดระยะเวลาในการศึกษาไว้เท่ากับ 5 ปี แล้วทำการเก็บรวบรวมข้อมูลโดยเริ่มทำการบันทึกข้อมูลตั้งแต่ผู้เอาประกันภัยเริ่มทำประกันจนกระทั่งขาดส่งเบี้ยประกัน หรือมีการยกเลิกกรมธรรม์ก่อนครบกำหนดสัญญาเอาประกันภัย ผู้เอาประกันภัยที่ขาดการส่งเบี้ยประกันในช่วงก่อนที่กรมธรรม์จะมีมูลค่ากรมธรรม์เป็นผลทำให้เกิดการขาดอายุกรมธรรม์ (Lapsation) แต่ถ้าผู้เอาประกันภัยขาดการส่งเบี้ยประกันหลังจากที่กรมธรรม์จะมีมูลค่ากรมธรรม์แล้วผู้เอาประกันภัยจะได้รับมูลค่าเวนคืนเงินสด หรือเปลี่ยนกรมธรรม์เป็นแบบกรมธรรม์ใช้เงินสำเร็จ หรือกรมธรรม์แบบขยายระยะเวลา จนกระทั่งขาดการชำระเบี้ยประกันภัย จะถือว่าข้อมูลของกรมธรรม์ไม่ถูกตัดทิ้ง สำหรับข้อมูลของกรมธรรม์ที่มีผลบังคับ(Inforce) คือกรมธรรม์ที่ยังส่งเบี้ยประกันภัยจนสิ้นสุดเวลาที่ทำการศึกษา นั้นข้อมูลที่เก็บรวบรวมได้จะมีลักษณะเป็นข้อมูลที่มีค่าถูกตัดประเภทที่ 1 โดยเวลาที่ถูกกำหนดไว้ล่วงหน้าเท่ากับ 5 ปี

ถ้า $T_1, T_2, T_3, \ldots, T_n$ เป็นค่าสังเกตที่ไม่ถูกตัดที่มีการแจกแจงเหมือนกันและเป็นอิสระกัน Tc เป็นเวลาที่ค่าสูงสุดกำหนดไว้ล่วงหน้า จะได้ตัวแปรสุ่มของค่าสังเกต $Y_1, Y_2, Y_3, \ldots, Y_n$ ซึ่ง

$$Y_i = \begin{cases} T_i & ; T_i \le T_c \\ T_c & ; T_i > T_c \end{cases}$$

โดยมีฟังก์ชันความควรจะเป็น (Likelihood Function) ดังนี้

$$L(y_i) = \begin{cases} f(y_i) & \text{ถ้าค่าสังเกตไม่ถูกตัดทิ้ง} \\ P(T_i > T_c) = S(T_c) & \text{ถ้าค่าสังเกตถูกตัด} \end{cases}$$

และมีฟังก์ชันความควรจะเป็นรวมดังนี้

$$L = \prod_{i \in u} f(y_i) \prod_{i \in c} S(T_c)$$

$i \in u$ หมายถึงค่าสังเกตที่ไม่ถูกตัดทิ้ง

$i \in c$ หมายถึงค่าสังเกตที่ถูกตัดทิ้ง

2. การถูกตัดประเภทที่ 2 (Type II Censoring)

ในบางกรณีไม่อาจกำหนดเวลา หรือค่าสูงสุดของการตัดข้อมูลที่เหมาะสมได้ ดังนั้นจะกำหนดจำนวนค่าสังเกตที่ไม่ถูกตัดแทน นั่นคือเมื่อค่าสังเกตที่ไม่ถูกตัดเกิดขึ้นครบจำนวนที่กำหนดแล้วก็หยุดการทดลอง เช่นการทดลอบอายุการใช้งานของหลอดไฟ จะกำหนดจำนวนหลอดไฟที่เสื่อมสภาพไว้ล่วงหน้า เริ่มทดลองโดยเปิดหลอดไฟทำงานทั้งหมดเริ่มบันทึกเวลาและนับจำนวนหลอดไฟที่เสื่อมสภาพ เมื่อได้จำนวนหลอดไฟที่เสื่อมสภาพครบแล้วก็จะหยุดทำการทดลอง

ถ้า N คือจำนวนข้อมูลทั้งหมดและกำหนด n คือจำนวนค่าสังเกตที่ไม่ถูกตัด $n \le N$ ให้ $T_1 \le T_2 \le T_3 \ldots \le T_n$ เป็นค่าสังเกตที่ไม่ถูกตัด และ $T_{n+1} \le T_{n+2} \le T_{n+3} \ldots \le T_N$ เป็นค่าสังเกตที่ถูกตัด ซึ่ง $T_i \ge T_n$; $i = n+1, n+2, \ldots, N$ ไม่ทราบค่าที่แท้จริงของค่าสังเกตดังนั้น Y_i เป็นตัวแปรสุ่มของค่าสังเกตซึ่ง

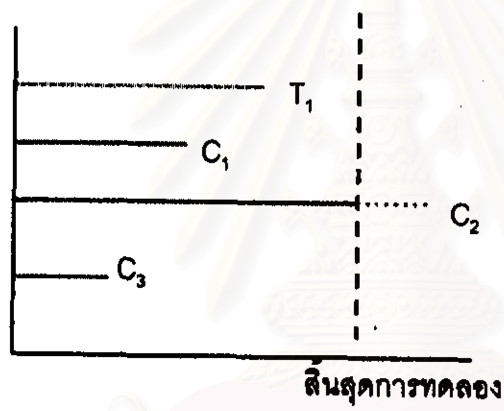
$$Y_i = \begin{cases} T_i & \text{เมื่อ } i \le n \\ T_n & \text{เมื่อ } i = n+1, n+2, \ldots, N \end{cases}$$

ฟังก์ชันความหนาแน่นร่วมของค่าสังเกตคือ

$$\frac{N!}{(N-n)!}f(y_1)f(y_2)...f(y_n)[S(y_n)]^{N-n}$$

3. การตัดแบบสุ่ม (Random Censoring)

ข้อมูลลักษณะนี้เกิดขึ้นคล้ายกับการตัดประเภทที่ 1 คือมีการกำหนดค่าสูงสุดหรือกำหนดเวลาไว้ล่วงหน้า แต่การตัดข้อมูลอาจจะเกิดขึ้นก่อนนั้นได้ เช่นการทดลองทางการแพทย์ที่คนไข้ที่ทำการศึกษาถอนตัวออกจากการทดลองก่อนสิ้นสุดการทดลอง หรือคนไข้ยังมีชีวิตอยู่รอดเมื่อสิ้นสุดการทดลอง หรือ คนไข้เสียชีวิตเนื่องจากสาเหตุอื่นที่ไม่เกี่ยวกับสิ่งที่กำลังศึกษาทดลอง เป็นต้น จึงไม่สามารถทราบค่าที่แน่นอนของค่าสังเกตนั้นได้ รูปที่ 2.1 แสดงลักษณะความเป็นไปได้ของข้อมูลถูกตัดแบบสุ่ม



รูปที่ 2.1 แผนภาพแสดงการเกิดค่าถูกตัดแบบสุ่ม

คนไข้คนที่ 1 เข้าทำการทดลองตั้งแต่เริ่มการทดลอง และเสียชีวิตที่เวลา T_1 จะถือว่าเป็นค่าสังเกตที่ไม่ถูกตัดทิ้ง

คนไข้คนที่ 2 เข้าทำการทดลองตั้งแต่เริ่มการทดลอง และถอนตัวจากการทดลองที่เวลา C_1 จะถือว่าเป็นค่าสังเกตที่ถูกตัดทิ้ง

คนไข้คนที่ 3 เข้าทำการทดลองตั้งแต่เริ่มการทดลอง และมีชีวิตอยู่รอดจนสิ้นสุดการทดลอง จะถือว่าเป็นค่าสังเกตที่ถูกตัดทิ้ง C_2

คนไข้คนที่ 4 เข้าทำการทดลองตั้งแต่เริ่มการทดลอง และเสียชีวิตเนื่องจากสาเหตุอื่นที่เวลา C_3 จะถือว่าเป็นค่าสังเกตที่ถูกตัดทิ้ง

ถ้า $T_1, T_2, T_3, \ldots, T_n$ เป็นตัวแปรสุ่มของค่าสังเกตที่ไม่ถูกตัดที่มีการแจกแจงเหมือนกันและเป็นอิสระกันมีฟังก์ชันการอยู่รอดและฟังก์ชันความหนาแน่น S และ f ตามลำดับ และ $C_1, C_2, C_3, \ldots, C_n$ เป็นตัวแปรสุ่มของค่าสังเกตที่ถูกตัดที่มีการแจกแจงเหมือนกันและเป็นอิสระกันมีฟังก์ชันการอยู่รอดและฟังก์ชันความหนาแน่น G และ g ตามลำดับ

ดังนั้น T_i และ C_i ; $i = 1,2,3,\ldots,N$ เป็นอิสระ จากนิยามการตัดทิ้งแบบสุ่ม นิยามให้ $Y_i = \min(T_i, C_i)$ จะได้ค่าสังเกตสุ่ม $Y_1, Y_2, Y_3, \ldots, Y_n$ ดังนี้

$$Y_i = \begin{cases} T_i & \text{ถ้า } T_i \le C_i \text{ (ไม่ถูกตัดทิ้ง)} \\ C_i & \text{ถ้า } T_i > C_i \text{ (ถูกตัดทิ้ง)} \end{cases}$$

$$\delta_i = \begin{cases} 1 & \text{ถ้า } T_i \le C_i \text{ (ไม่ถูกตัดทิ้ง)} \\ 0 & \text{ถ้า } T_i > C_i \text{ (ถูกตัดทิ้ง)} \end{cases}$$

จะมีฟังก์ชันความควรจะเป็นดังนี้

$$L(y_i, \delta_i) = \begin{cases} f(y_i)G(y_i) & \text{ถ้า } \delta_i = 1 \\ g(y_i)S(y_i) & \text{ถ้า } \delta_i = 0 \end{cases}$$

และมีฟังก์ชันความควรจะเป็นรวมดังนี้

$$L = \left[\prod_{i \in N} f(y_i) \prod_{i \in N} G(y_i)\right]\left[\prod_{i \in C} g(y_i) \prod_{i \in C} S(y_i)\right]$$

เนื่องจาก $G(y_i)$ และ $g(y_i)$ ไม่เกี่ยวข้องกับพารามิเตอร์ที่สนใจ จึงละไว้โดยใช้ฟังก์ชันความควรจะเป็นดังนี้

$$L = \left[\prod_{i \in N} f(y_i) \prod_{i \in C} S(y_i)\right]$$

$i \in u$ หมายถึงค่าสังเกตที่ไม่ถูกตัดทิ้ง

$i \in c$ หมายถึงค่าสังเกตที่ถูกตัดทิ้ง

4. การตัดแบบอื่น (Other Type of Censoring)

นอกจากการตัดประเภทที่ 1 การตัดประเภทที่ 2 และการตัดแบบสุ่ม ซึ่งที่กล่าวมาเป็นการตัดทางขวา ยังมีการตัดทางซ้าย (Left Censoring) และการตัดทางซ้ายและทางขวา (Left and Right Censoring) ข้อมูลที่ถูกตัดทางซ้าย เช่น การรับประกันภัยรถยนต์จะกำหนดความเสียหายส่วนแรก (Deductible) ที่ผู้เอาประกันภัยจะต้องรับผิดชอบ ดังนั้นข้อมูลการจ่ายค่าสินไหมทดแทนที่บริษัทเก็บไว้จะเป็นข้อมูลเฉพาะกรณีที่ค่าเสียหายที่เกิดขึ้นมากกว่ากว่าค่าเสียหายส่วนแรก และถ้าบริษัทรับประกันภัยนี้ได้ส่งต่อให้กับบริษัทประกันต่อ ซึ่งบริษัทรับประกันภัยจะกำหนดความรับผิดสูงสุดของบริษัทเอาไว้ ความสูญเสียส่วนที่เกินจากที่กำหนดนี้บริษัทรับประกันต่อจะเป็นผู้รับผิดชอบ ดังนั้นข้อมูลที่บริษัทรับประกันภัยมีอยู่จึงเป็นความสูญเสียที่มากกว่าความเสียหายส่วนแรกและไม่เกินจำนวนสูงสุดที่รับผิดชอบ

2.1.2 ฟังก์ชันการอยู่รอดและฟังก์ชันการสูญเสีย (Survival Function and Hazard Function)

ให้ T เป็นตัวแปรสุ่มของระยะเวลาการอยู่รอด

$f(t)$ เป็นฟังก์ชันความหนาแน่นของ T (Probability density Function)

$F(t)$ เป็นฟังก์ชันการแจกแจงของ T (Distribution Function)

$S(t)$ เป็นฟังก์ชันการแจกแจงการอยู่รอดของ T (Survival Function)

$h(t)$ เป็นฟังก์ชันความสูญเสียหรืออัตราการสูญเสียของ T (Hazard Function)

ซึ่งนิยามของฟังก์ชัน $S(t)$ คือความน่าจะเป็นที่ตัวแปรสุ่ม T จะมีค่ามากกว่า t

$$
\begin{aligned}
S(t) &= \Pr(T > t) \\
&= \int_t^{\infty} f(t)\,dt \\
&= 1 - \Pr(T \le t) \\
&= 1 - F(t)
\end{aligned}
$$

โดยที่ $S(t)$ มีคุณสมบัติดังนี้

1. $S(t)$ เป็นฟังก์ชันไม่เพิ่ม (Nonincreasing Function)
2. $S(t)$ เป็นฟังก์ชันต่อเนื่องของ t
3. $S(t) = 1$ เมื่อ $t = 0$
4. $S(t) = 0$ เมื่อ $t = \infty$

นิยามฟังก์ชัน $h(t)$ แทนฟังก์ชันการสูญเสียที่มีค่าเท่ากับลิมิตของความน่าจะเป็นที่ตัวแปรสุ่ม T จะมีค่าอยู่ในช่วงสั้นๆ $(t, t+\Delta t)$ ต่อหน่วยเวลา Δt เมื่อกำหนดค่า $T > t$ และ $h(t)$ กำหนดดังนี้

$$
\begin{aligned}
h(t) &= \lim_{\Delta t \to 0} \frac{\Pr(t \le T < t + \Delta t / T \ge t)}{\Delta t} \\
&= \lim_{\Delta t \to 0} \frac{\left[F(t+\Delta t) - F(t)\right]}{(1 - F(t))\Delta t} \\
&= \frac{d}{dt} F(t) \frac{1}{1 - F(t)} \\
&= \frac{f(t)}{1 - F(t)} \\
&= \frac{f(t)}{S(t)}, t > 0
\end{aligned}
$$

กรณีเมื่อ T เป็นตัวแปรสุ่มที่ไม่ต่อเนื่อง และมีค่าเป็น $t_1, t_2, t_3, \ldots, t_n$ โดยที่ $0 \le t_1 < t_2 < t_3 < \ldots < t_n$ ดังนั้นฟังก์ชันความน่าจะเป็น (Probability Function)กำหนดได้ดังนี้

$$
p(t_j) = \Pr(T = t_j) \qquad j = 1,2,3,\ldots
$$

และฟังก์ชันการอยู่รอด $S(t)$ คือ

$$
\begin{aligned}
S(t) &= \Pr(T \ge t_j) \\
&= \sum_{j = j \ge t} p(t_j)
\end{aligned}
$$

ดังนั้นฟังก์ชันความสูญเสีย h(t) แสดงได้ดังนี้

$$
\begin{aligned}
h(t) &= \Pr(T = t_j / T \ge t_j) \\
&= \frac{p(t_j)}{S(t_j)}
\end{aligned}
$$

2.1.3 ฟังก์ชันความหนาแน่น ฟังก์ชันการแจกแจง ฟังก์ชันการอยู่รอด และค่าคาดหวัง ของการแจกแจงภายใต้การศึกษานี้

สำหรับในการวิจัยครั้งนี้พิจารณารูปแบบการแจกแจงการอยู่รอด 3 รูปแบบ คือ

1. การแจกแจงไวบูลล์ (Weibull Distribution) มี τ และ c เป็นพารามิเตอร์

$$f(t) = \begin{cases} c\tau t^{\tau-1}\exp\left[-ct^{\tau}\right] & ,t \ge 0, c>0, \tau>0 \\ 0 & ,t<0 \end{cases}$$

$$F(t) = 1-\exp\left[-ct^{\tau}\right]$$

$$S(t) = \exp\left[-ct^{\tau}\right]$$

$$E(t) = \frac{\Gamma\left(1+\frac{1}{\tau}\right)}{c^{\frac{1}{\tau}}}$$

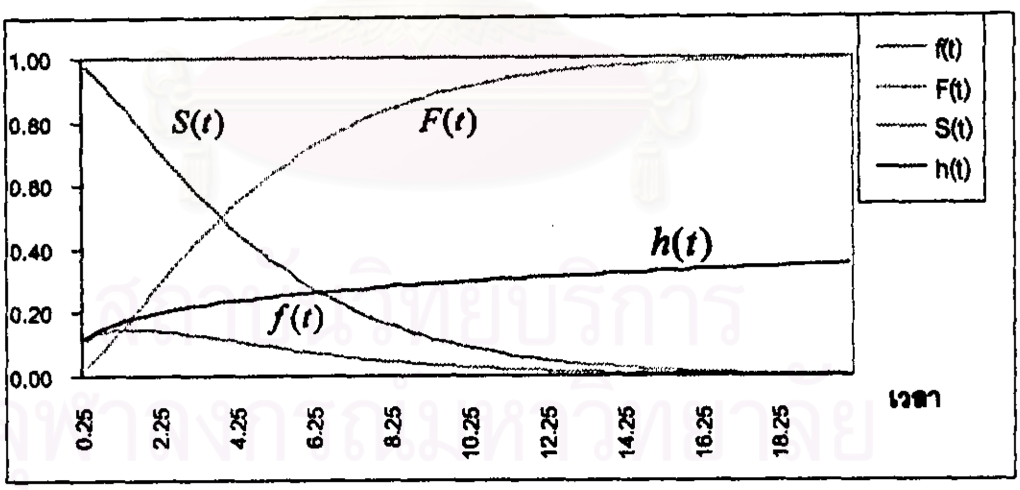
โดยกำหนดให้

$$\tau = 1.5$$

$$c = 0.05$$

$$E(t) = 6$$

สำหรับรูปแบบของฟังก์ชันความหนาแน่น ฟังก์ชันการแจกแจงสะสม ฟังก์ชันการอยู่รอด และฟังก์ชันภาวะภัย ของการแจกแจงแบบไวบูลล์ ได้แสดงไว้ดังรูปที่ 2.2



รูปที่ 2.2 แสดงความสัมพันธ์ของฟังก์ชัน $f(t), F(t), S(t)$ และ $h(t)$ เมื่อมีการแจกแจงแบบไวบูลล์

2. การแจกแจงลอกนอร์มอล (Lognormal Distribution) มี μ และ δ เป็น พารามิเตอร์

$$f(t) = \begin{cases} \dfrac{\exp\left[-\dfrac{(\ln t-\mu)^2}{2\delta^2}\right]}{\sqrt{2\pi}\delta t} & ,t > 0, -\infty < \mu < \infty, \delta > 0 \\ 0 & ,t \le 0 \end{cases}$$

$$F(t) = \phi\left[\frac{\ln(t)-\mu}{\delta}\right]$$

เมื่อ $\phi(x)$ คือฟังก์ชันการแจกแจงปกติมาตรฐาน

$$\phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x \exp\left[-\frac{u^2}{2}\right]du$$

$$S(t) = 1-\phi\left[\frac{\ln(t)-\mu}{\delta}\right]$$

$$E(t) = \exp(\mu + \frac{\delta^2}{2})$$

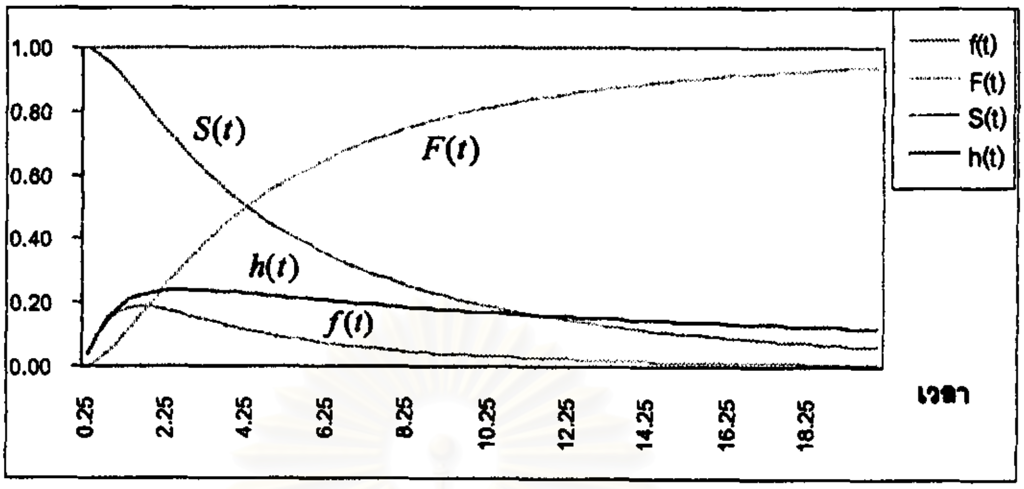
โดยกำหนดให้

μ = 1.67

δ = 0.5

E(t) = 6

สำหรับรูปแบบของฟังก์ชันความหนาแน่น ฟังก์ชันการแจกแจงสะสม ฟังก์ชันการอยู่รอด และฟังก์ชันภาวะภัย ของการแจกแจงแบบลอกนอร์มอล ได้แสดงไว้ดังรูปที่ 2.3



รูปที่ 2.3 แสดงความสัมพันธ์ของฟังก์ชัน $f(t)$,$F(t)$,$S(t)$ และ $h(t)$ เมื่อมีการแจกแจงแบบลอกนอร์มอล

3. การแจกพาเรโต (Pareto Distribution) มี λ และ α เป็นพารามิเตอร์

$$f(t)=\begin{cases}\alpha\lambda^{\alpha}(\lambda+t)^{-\alpha-1} & ,t>0,\alpha>0,\lambda>1\\ 0 & ,t<0\end{cases}$$

$$F(t)=1-\left(\frac{\lambda}{\lambda+t}\right)^{\alpha}$$

$$S(t)=\left(\frac{\lambda}{\lambda+t}\right)^{\alpha}$$

$$E(t)=\frac{\lambda}{(\alpha-1)}$$

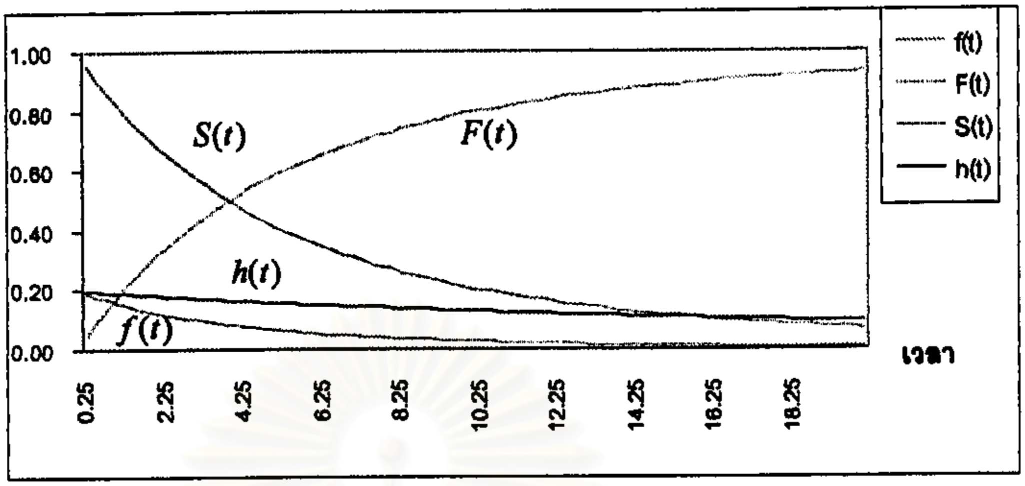
โดยกำหนดให้

λ = 24

α = 5

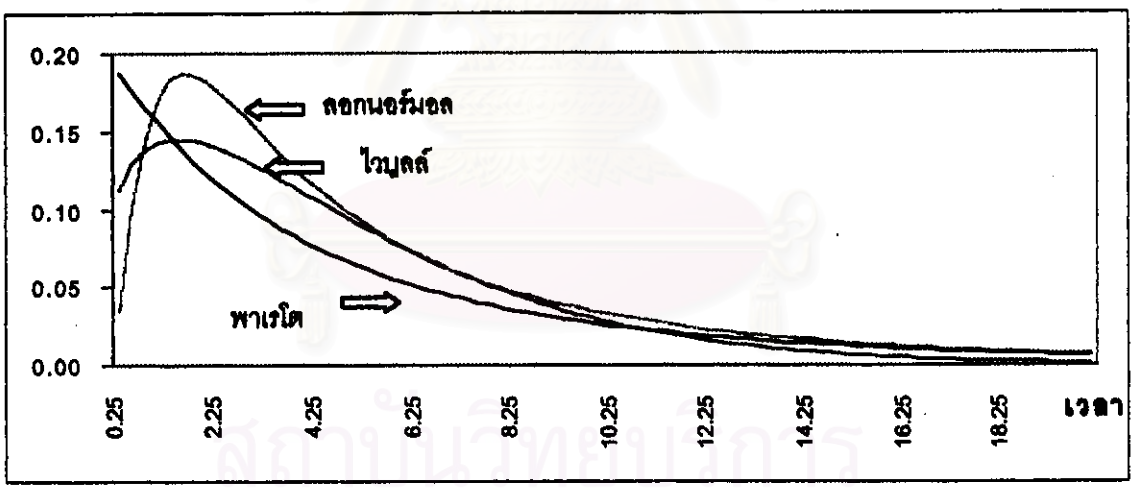
E(t) = 6

สำหรับรูปแบบของฟังก์ชันความหนาแน่น ฟังก์ชันการแจกแจงสะสม ฟังก์ชันการอยู่รอด และฟังก์ชันภาวะภัย ของการแจกแจงแบบพาเรโต ได้แสดงไว้ดังรูปที่ 2.4



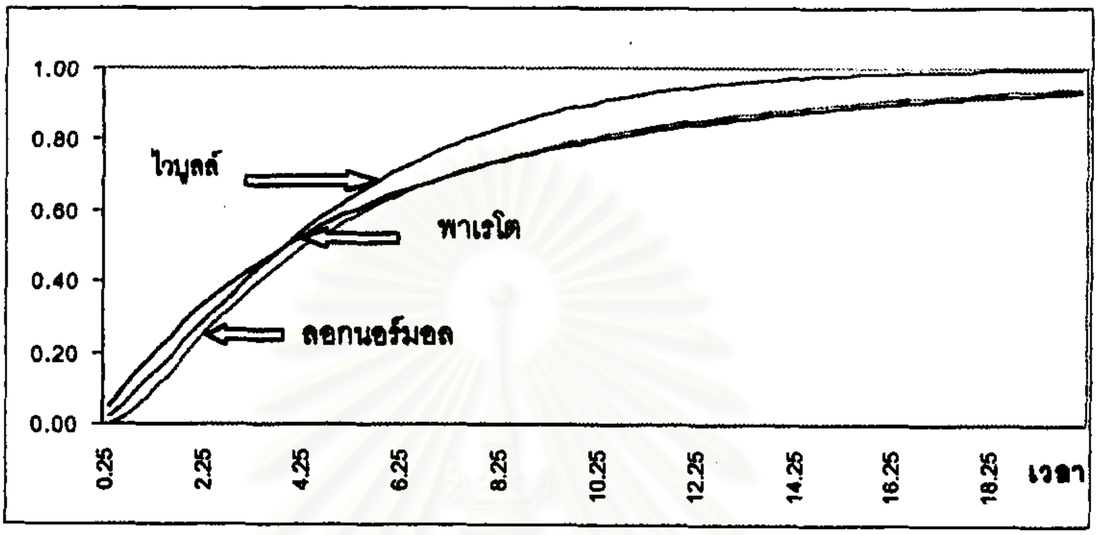
รูปที่ 2.4 แสดงความสัมพันธ์ของฟังก์ชัน $f(t), F(t), S(t)$ และ $h(t)$ เมื่อมีการแจกแจงแบบพาเรโต

สำหรับรูปแบบของฟังก์ชันความหนาแน่น ของการแจกแจงแบบไวบูลล์ ลอกนอร์มอล และพาเรโต ได้แสดงไว้ดังรูปที่ 2.5



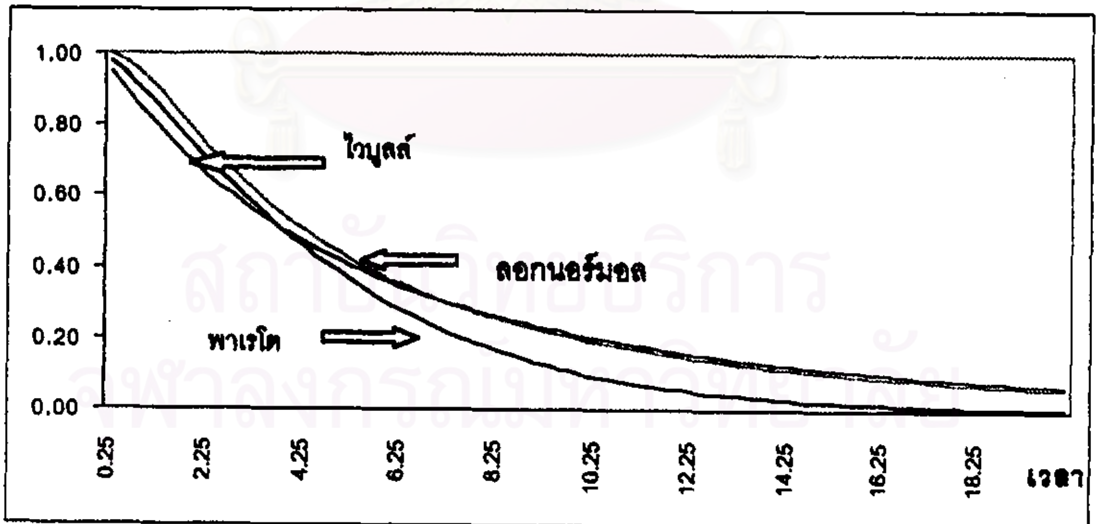
รูปที่ 2.5 เปรียบเทียบฟังก์ชันความหนาแน่นของการแจกแจงแบบไวบูลล์ ลอกนอร์มอล และพาเรโต

สำหรับรูปแบบของฟังก์ชันสะสม ของการแจกแจงแบบไวบูลล์ ลอกนอร์มอล และพาเรโต ได้แสดงไว้ดังรูปที่ 2.6



รูปที่ 2.6 เปรียบเทียบฟังก์ชันสะสมของการแจกแจงแบบไวบูลล์ ลอกนอร์มอล และพาเรโต

สำหรับรูปแบบของฟังก์ชันการอยู่รอด ของการแจกแจงแบบไวบูลล์ ลอกนอร์มอล และพาเรโต ได้แสดงไว้ดังรูปที่ 2.7



รูปที่ 2.7 เปรียบเทียบฟังก์ชันการอยู่รอดของการแจกแจงแบบไวบูลล์ ลอกนอร์มอล และพาเรโต

2.1.3 วิธีการประมาณค่า

1. วิธีลิมิตผลคูณ (Product Limit (PL) Method)[1]

วิธีลิมิตผลคูณ เป็นวิธีประมาณที่ถูกพัฒนาขึ้นโดย Kaplan และ Meier ในปี ค.ศ.1958 โดยใช้ข้อมูลจากค่าสังเกตทุกค่าซึ่งจะกำหนดเวลาที่แน่นอนในการศึกษาค่าสังเกต โดยจะแบ่งการบันทึกข้อมูลออกเป็นช่วง ในการสังเกตแต่ละครั้งจะกำหนดขอบเขตของเวลาการสังเกตแต่ละช่วง โดยที่

X_i เป็นค่าสังเกตที่ทำการศึกษา

T_i เป็นตัวแปรสุ่มของเวลาการอยู่รอด ที่มีการแจกแจงเหมือนกัน และเป็นอิสระกัน

δ_i เป็นจำนวนค่าสังเกตที่หยุดการชำระเบี้ยประกันภัยในแต่ละช่วง

λ_i เป็นจำนวนค่าสังเกตที่ชำระเบี้ยประกันภัยในแต่ละช่วง

n_i เป็นจำนวนค่าสังเกตทั้งหมดที่เหลืออยู่ในช่วงที่ทำการศึกษา $\left(n_i = \delta_i + \lambda_i\right)$

P_i เป็นความน่าจะเป็นที่จะอยู่รอดที่เวลา T_i โดยที่อยู่รอดมาแล้วที่เวลา T_{i-1}

จะได้ว่า

$$p_i = \Pr(T > t_i / T > t_{i-1})$$

$$\hat{p}_i = \frac{n_i - \delta_i}{n_i}$$

ดังนั้นจะได้ว่า

$$S(t_i) = \Pr(T > t_i)$$

$$= \Pr(T > t_1) \times \Pr(t > t_2 / T > t_1) \times \ldots \times \Pr(T > t_i / T > t_{i-1})$$

$$= p_1 \cdot p_2 \cdots p_i$$

สำหรับการประมาณจะได้ว่า

$$\hat{S}(t) = \prod_{i=1}^{t} \hat{p}_i \qquad \text{———} \qquad (2.1)$$

$$\hat{S}(t) = \prod_{i=1}^{t} \left(\frac{n_i - \delta_i}{n_i} \right) \qquad (2.2)$$

---

[1] Kaplan EL,and Meier P. "Nonparametric Estimation from incomplete Observation," Journal of the American Statistical 53(1958) : 457-481

2. วิธีทางคณิตศาสตร์ประกันชีวิต (Actuarial Method or Life-table Estimation)[2]

วิธีทางคณิตศาสตร์ประกันชีวิต สำหรับวิธีการประมาณนี้จะทำการประมาณอัตราค่ากลาง(The Central Rate : m_i) ก่อน โดยจะใช้ข้อมูลจากค่าสังเกตทุกค่าซึ่งกำหนดเวลาที่แน่นอนในการศึกษาค่าสังเกต โดยจะแบ่งการบันทึกข้อมูลออกเป็นช่วง ในการสังเกตแต่ละครั้งจะกำหนดขอบเขตของเวลาการสังเกตแต่ละช่วง โดยที่

m_i เป็นอัตราค่ากลางในแต่ละช่วงที่ทำการศึกษา

จะได้ว่า

$$\hat{m}_i = \frac{\delta_i}{n_i - \frac{\delta_i}{2}}$$

ภายใต้ความสัมพันธ์ของอัตราค่ากลางและกำลังของมรณะ (Force of Mortality) เป็นค่าคงที่ จะได้ว่า

$$\mu_i = m_i$$
$$p_i = \Pr(T > t_i / T > t_{i-1})$$
$$\hat{p}_i = e^{-\hat{m}_i}$$

จากการประมาณ $\hat{S}(t)$ ด้วยวิธีลิมิตผลคูณ ในสมการที่ (2.1) จะได้ว่า

$$\hat{S}(t) = \prod_{i=1}^{t} e^{-m_i} \qquad (2.3)$$

3. วิธีประมาณแบบคลาสสิค (Classical Estimation Method)

วิธีประมาณแบบคลาสสิค จะกำหนดสมมติฐานว่า ผลการทดลองแต่ละผลการทดลองมีโอกาสเกิดขึ้นเท่าๆ กัน ภายใต้ผลการทดลองจำนวนจำกัด ดังนั้นจึงทำให้สามารถกำหนดนิยามของความน่าจะเป็นของเหตุการณ์ A ใดๆ ได้ดังนี้

---

[2]Dick London , FSA , Survival Models and Their Estimation : ACTEX Publications Winsted and Avon, Connecticut ,1988 : 110-111.

$$P(A) = \frac{\text{จำนวนเหตุการณ์ A จะเกิดขึ้นได้}}{\text{จำนวนวิธีทั้งหมดซึ่งการทดลองจะให้ผลการทดลองได้}}$$

ซึ่งจากแนวความคิดดังกล่าว เราสามารถประยุกต์ใช้กับการประมาณหาค่า p_{t_i} จะได้ว่า

$$\hat{p}_{t_i} = \frac{D_{t_i}}{E_{t_i}}$$

โดยที่ D_{t_i} = จำนวนกรมธรรม์ที่ยังชำระเบี้ยประกันภัยระหว่างช่วงที่ t_i และ t_{i+1}

E_{t_i} = จำนวนกรมธรรม์ ณ ช่วงที่ t_i

$$= \sum_{i=1}^{n} (r_i)$$

เมื่อ r_i เป็นระยะเวลาของกรมธรรม์ที่ i หยุดชำระเบี้ยประกันภัยซึ่ง $0 < r_i \leq 1$

n เป็นจำนวนกรมธรรม์ต้นช่วงเวลาที่ t_i

จากการประมาณ $\hat{S}(t)$ ด้วยวิธี ลิมิตผลคูณ ในสมการที่ (2.1) จะได้ว่า

$$\hat{S}(t) = \prod_{i=1}^{t} \frac{D_{t_i}}{E_{t_i}} \qquad (2.4)$$

4. วิธีนอนพาราเมตริกแบบเบส์ (Bayesian Nonparametric Method)[3]

วิธีนอนพาราเมตริกแบบเบส์นี้ เป็นวิธีทางสถิติที่ต้องอาศัยข้อมูลจากกลุ่มตัวอย่างที่นำมาศึกษา และข้อมูลหรือประสบการณ์ในอดีตเข้าช่วยในการประมาณค่า โดยจากข้อมูลหรือประสบการณ์ในอดีตนี้ทำให้สามารถประมาณการแจกแจงก่อน (Prior Distribution) ได้ และเมื่อทราบการแจกแจงก่อน และข้อมูลจากกลุ่มตัวอย่างที่นำมาศึกษาแล้ว จะสามารถทราบการแจกแจงภายหลัง (Posterior Distribution) ได้ โดยจะนำการแจกแจงดังกล่าวนี้ไปใช้ในการประมาณค่าที่ต้องการ

---

[3] Feruson Thomas S. and Eswar G. Phadia, "Bayesian Nonparametric Estimation Based on Censored Data," The Annals of Statistics 7(1979) : 163-186

ได้ ซึ่ง Susarla และ Van Ryzin เป็นผู้เสนอรูปแบบของการแจกแจงก่อนโดยกระบวนการดีริชเลต์ (Dirichlet process prior) ซึ่งรูปแบบของการแจกแจงที่นำเสนอนี้ไม่ต้องทราบการแจกแจงของฟังก์ชัน

สำหรับการวิจัยครั้งนี้ กำหนดให้ $T_1, T_2, T_3, \ldots, T_n$ เป็นตัวอย่างสุ่มจากฟังก์ชันการแจกแจงสะสม (Cumulative Distribution Function) F และ T_c เป็นค่าคงที่ซึ่งทราบค่า และให้ $Z_i = \min(T_i, T_c) \quad ; i = 1,2,..,n$

ในที่นี้จะหาตัวประมาณเบส์ภายใต้การกำหนดให้ฟังก์ชันความสูญเสียเป็นความคลาดเคลื่อนกำลังสอง (squared-error loss function)

$$l(F(t) - \hat{F}(t)) = (F(t) - \hat{F}(t))^2$$

ดังนั้น ตัวประมาณเบส์สามารถหาได้จากการหาค่าคาดหวังของการแจกแจงภายหลังของ F เมื่อกำหนด (Z_i, δ_i)

สำหรับการประมาณค่าฟังก์ชันการอยู่รอด ด้วยวิธีตัวประมาณเบส์โดยวิธีการประมาณที่ไม่ใช้พารามิเตอร์ในที่นี้พิจารณาการแจกแจงก่อนของ F เป็นกระบวนการนิวทรัลทางขวา (Process neutral to the right)

เมื่อ $F(t)$ เป็นฟังก์ชันการแจกแจงสุ่มที่เป็นนิวทรัลทางขวา (neutral to the right) แล้ว $F(t)$ สามารถเขียนได้ในรูปของ

$$F(t) = 1 - e^{-Y_t} \qquad (2.5)$$

โดยที่ Y_t เป็นกระบวนการซึ่งมีการเพิ่มขึ้นอย่างอิสระ (Independent increment) ซึ่งมีคุณสมบัติดังนี้

1) Y_t เป็นฟังก์ชันไม่ลดลง (Nondecreasing)
2) Y_t ต่อเนื่องทางขวา (Right-continuous)
3) $\lim_{t \to -\infty} Y_t = 0$
4) $\lim_{t \to +\infty} Y_t = \infty$

กำหนดให้ $S(t)$ เป็นฟังก์ชันการอยู่รอดซึ่งนิยามได้ดังนี้

$$S(t) = 1 - F(t) \qquad (2.6)$$

ดังนั้น จากสมการ (2.5) และ (2.6) จะได้ว่าการหาค่าคาดหวังของการแจกแจงภายหลังของ F เมื่อกำหนด (Z_i, δ_i) คือ การหาโมเมนต์ที่ 1 ของ Y_t เพื่อลดปัญหาการหาค่าของฟังก์ชันที่ให้โมเมนต์

$$E\left[S(t)\middle|\left(Z_i,\delta_i\right)\right] = E\left[\left(1-F(t)\right)\middle|\left(Z_i,\delta_i\right)\right]$$
$$= E\left[\left(1-\left(1-e^{-Y_t}\right)\right)\middle|\left(Z_i,\delta_i\right)\right]$$
$$= E\left[\left(e^{-Y_t}\right)\middle|\left(Z_i,\delta_i\right)\right]$$

นั่นคือ $E\left[S(t)\middle|\left(Z_i,\delta_i\right)\right] = M_t\left(1\middle|\left(Z_i,\delta_i\right)\right)$ ———— (2.7)

โดยที่ $M_t\left(\theta\middle|\left(Z_i,\delta_i\right)\right)$ เป็นฟังก์ชันที่ให้โมเมนต์ภายหลัง(Posterior Moment Generating function) ของ Y_t ซึ่งในที่นี้นิยามฟังก์ชันที่ให้โมเมนต์เป็น

$$M_t(\theta) = E\left(e^{-\theta Y_t}\right) \text{ ———— } (2.8)$$

กำหนดให้ $u_1,u_2,\ldots,u_k$ คือ ค่าที่แตกต่างกันของ $t_1,t_2,\ldots,t_k$ โดยที่ $u_1 > u_2 > \ldots > u_k$

$\rho_1,\rho_2,\ldots,\rho_k$ คือ จำนวนของค่าสังเกตที่ไม่ถูกตัดทิ้งที่ $u_1,u_2,\ldots,u_k$ ตามลำดับ

$\lambda_1,\lambda_2,\ldots,\lambda_k$ คือ จำนวนของค่าสังเกตที่มีค่าถูกตัดทิ้งที่ $u_1,u_2,\ldots,u_k$ ตามลำดับ

h_j เป็นจำนวนของ t_i ที่มีค่ามากกว่า u_j

$j(t)$ เป็นจำนวนของ u_j ที่มีค่าน้อยกว่าหรือเท่ากับ t

ดังนั้น จะได้ว่า
$$h_i = \sum_{i=j+1}^{k}\left(\rho_i + \lambda_i\right) \text{ ———— } (2.9)$$

เมื่อ F เป็นฟังก์ชันการแจกแจงสุ่มที่เป็นนิวทรัลทางขวา และให้ $T_1,T_2,\ldots,T_n$ เป็นตัวอย่างสุ่มขนาด n จาก F แล้วจะได้ว่า การแจกแจงภายหลังของ F เมื่อกำหนดข้อมูล (Z_i,δ_i) จะเป็นนิวทรัลทางขวา[6] และมีฟังก์ชันที่ให้โมเมนต์ภายหลังของ Y_t เป็น

$$M_t\left(\theta\middle|\left(Z_i,\delta_i\right)\right) = \frac{M_t\left(\theta+h_{j(t)}\right)}{M_t\left(h_{j(t)}\right)}\cdot\prod_{i=1}^{j(t)}\left[\frac{M_{u_i}^{-}\left(\theta+h_{i-1}\right)}{M_{u_i}^{-}\left(h_{i-1}\right)}\cdot\frac{C_{u_i}\left(\theta+h_i+\lambda_i,\rho_i\right)}{C_{u_i}\left(h_i+\lambda_i,\rho_i\right)}\cdot\frac{M_{u_i}\left(h_i\right)}{M_{u_i}\left(\theta+h_i\right)}\right] \text{ — } (2.10)$$

โดยที่ ถ้า u เป็นจุดที่กำหนดไว้ล่วงหน้าของ Y_t

$$C_u(\alpha,\beta) = \int_0^\infty e^{-\alpha s}\left(1-e^{-s}\right)^{\beta} dG_u(s) \text{ ———— } (2.11)$$

และ ถ้า u ไม่ได้เป็นจุดที่กำหนดไว้ล่วงหน้าของ Y_t

$$C_u(\alpha,\beta) = \int_0^\infty e^{-\alpha s}\left(1-e^{-s}\right)^{\beta-1} dH_u(s) \quad \text{ถ้า } \beta \ge 1$$
$$= 1 \qquad \beta = 0 \text{ ———— } (2.12)$$

[6] Doksum , K., "Tailfree and neutral random probabilities and their posterior distribution," The Annals of Probability 2(1974) : 183-201.

เมื่อ S เป็นค่าที่เพิ่มขึ้นของ Y_t ที่ u_t นั่นคือ $S = Y_t - Y_{t-1}$

$G_u(s)$ เป็นการแจกแจงก่อนของ S ที่ u

$H_u(s)$ เป็นการแจกแจงภายหลังของ S ที่ u เมื่อกำหนด $T = u$

ดังนั้น จากสมการ (2.7) และ (2.10) จะได้ว่า

$$M_t\left(1\middle|(Z_i,\delta_i)\right)=\frac{M_t\left(1+h_{J(t)}\right)}{M_t\left(h_{J(t)}\right)}\cdot\prod_{i=1}^{J(t)}\left[\frac{M^-_{u_i}\left(1+h_{i-1}\right)}{M^-_{u_i}\left(h_{i-1}\right)}\cdot\frac{C_{u_i}\left(1+h_i+\lambda_i,\rho_i\right)}{C_{u_i}\left(h_i+\lambda_i,\rho_i\right)}\cdot\frac{M_{u_i}\left(h_i\right)}{M_{u_i}\left(1+h_i\right)}\right] \quad \text{——} \quad (2.13)$$

และเมื่อกำหนดให้

$$R_t(h)=\frac{M_t(h+1)}{M_t(h)} \quad \text{——} \quad (2.14)$$

$$r_u(\alpha,\beta)=C_u(\alpha+1,\beta)/C_u(\alpha,\beta) \quad \text{——} \quad (2.15)$$

ดังนั้น $E\left[S(t)\middle|(Z_i,\delta_i)\right]=M_t\left(1\middle|(Z_i,\delta_i)\right)$

$$\hat{S}(t)=R_t\left(h_{J(t)}\right)\prod_{i=1}^{J(t)}\frac{R^-_{u_i}\left(h_{i-1}\right)}{R_{u_i}\left(h_i\right)}\cdot r_{u_i}\left(h_i+\lambda_i,\rho_i\right) \quad \text{——} \quad (2.16)$$

กำหนดให้

$Y_t = \lim_{s\to t} Y_s$

$M^-{}_t(\theta)$ เป็นฟังก์ชันที่ให้โมเมนต์ (Moment Generating Function) ของ Y_t^-

$M_t(\theta)$ เป็นฟังก์ชันที่ให้โมเมนต์ (Moment Generating Function) ของ Y_t

โดยที่ ฟังก์ชันที่ให้โมเมนต์ของ Y_t สามารถเขียนให้อยู่ในรูป

$$M_t(\theta)=e^{\gamma(t)\int_0^{\infty}\left(e^{-z\theta}-1\right)dN(z)} \quad \text{——} \quad (2.17)$$

เมื่อ γ(t) เป็นฟังก์ชันต่อเนื่องที่ไม่ลดลง ที่มี $\lim_{t\to-\infty}\gamma(t)=0$ และ $\lim_{t\to+\infty}\gamma(t)=+\infty$

$N(z)$ เป็นฟังก์ชันต่อเนื่องที่ไม่ลดลงซึ่งเป็นตัววัด (measure) ใดๆ บน (0,∞)

ในการวิจัยครั้งนี้ได้ศึกษาการแจกแจงก่อนเป็น 2 แบบ คือ

4.1<u>แบบกระบวนการดีริชเลต์</u> (Dirichlet Process)

ในที่นี้พิจารณากระบวนการดีริชเลต์ $\mathcal{D}(\alpha)$ ซึ่ง Susarla และVan Ryzin เป็นผู้นำเสนอ โดยกำหนดให้ α เป็นฟังก์ชันต่อเนื่องทางขวาที่ไม่ลดลง (nondecreasing right-continuous) บนจำนวนจริง(R) โดยที่มี $\lim_{t\to-\infty}\alpha(t)=0$ และ $\lim_{t\to+\infty}\alpha(t)=\alpha(R)<\infty$

เมื่อ $\mathcal{D}(\alpha)$ เป็นฟังก์ชันการแจกแจงสุ่มที่เป็นนิวทรัลทางขวาแล้ว จะสามารถเขียนฟังก์ชันที่ให้โมเมนต์ของ Y_t ได้ในรูป

$$
\begin{aligned}
M_t(\theta) &= \frac{\Gamma\left(\alpha(R)\right)}{\Gamma\left(\alpha(t)\right)\Gamma\left(\alpha(R)-\alpha(t)\right)}\int_0^\infty e^{-\left(\alpha(R)-\alpha(t)+\theta\right)y}\left(1-e^{-y}\right)^{\alpha(t)-1}dy \\
&= \frac{\Gamma\left(\alpha(R)\right)\Gamma\left(\alpha(R)-\alpha(t)+\theta\right)}{\Gamma\left(\alpha(R)-\alpha(t)\right)\Gamma\left(\alpha(R)+\theta\right)} \\
&= e^{\int_0^\infty \left(e^{-z\theta}-1\right)dN_t(z)} \qquad \text{——— (2.18)}
\end{aligned}
$$

โดยที่ $\quad dN_t(z)=\dfrac{e^{-\alpha(R)z}\left(e^{-\alpha(t)z}-1\right)}{z\left(1-e^{-z}\right)}dz$

เมื่อ $0<\alpha(t)<\alpha(R)$ และจากฟังก์ชันที่ให้โมเมนต์ของ Y_t ในสมการที่ (2.18) กำหนดให้มีการแจกแจงเป็น $\mathrm{H}(\alpha(t),\alpha(R)-\alpha(t))$ ดังนั้น สำหรับ $s<t$ จะได้ว่า

$$Y_t-Y_s\in\mathrm{H}(\alpha(t)-\alpha(s),\alpha(R)-\alpha(t))$$

โดยที่สามารถหาค่าของฟังก์ชัน R และ r ได้นั่นคือ จากสมการ (2.14) และ (2.18) จะได้

$$R_t(h)=\frac{\alpha(R)-\alpha(t)+h}{\alpha(R)+h} \qquad \text{——— (2.19)}$$

และสำหรับการหาค่า $r_u(a,b)$ จะแยกพิจารณา 2 กรณีคือ

1. เมื่อ u เป็นจุดที่ถูกกำหนดไว้ล่วงหน้า (Prior fixed point) ของ Y_t

ดังนั้น u จะเป็นจุดที่ไม่ต่อเนื่องของ α และถ้าให้ $\Delta(u)=\alpha(u)-\alpha^-(u)$ แล้วฟังก์ชันการแจกแจงก่อนของ $Y_u-Y_u^-$ คือ $\mathrm{H}(\Delta(u),\alpha(R)-\alpha(u))$ และจากสมการ (2.11) และ (2.18) จะได้

$$C_u(a,b)=\frac{\Gamma\left(\alpha(R)-\alpha(u)+\Delta(u)\right)\Gamma\left(\alpha(R)-\alpha(u)+a\right)}{\Gamma\left(\alpha(u)\right)\Gamma\left(\alpha(R)-\alpha(u)\right)\Gamma\left(\alpha(R)-\alpha(u)+\Delta(u)+a+b\right)} \qquad \text{——— (2.20)}$$

ดังนั้น

$$r_u(a,b)=\frac{\alpha(R)-\alpha(u)+a}{\alpha(R)-\alpha(u)+\Delta(u)+a+b} \qquad (2.21)$$

1. เมื่อ u ไม่ใช่จุดที่ถูกกำหนดไว้ล่วงหน้า ของ Y_i สามารถคำนวณได้จากสมการที่ (2.20) โดยพิจารณาที่ $\Delta(u)=0$ และ b = 0

ดังนั้นค่าคาดหวังภายหลังของ S(t) เมื่อการแจกแจงก่อนของ F เป็นกระบวนการดีริชเลต์จะคำนวณได้จาก

$$E(S(t)/(Z_i,\delta_i))=\frac{\alpha(R)-\alpha(t)+h_{j(t)}}{\alpha(R)+n}\cdot\prod_{i=1}^{j(t)}\frac{\left(\alpha(R)-\alpha^-(u_i)+h_{i-1}\right)\left(\alpha(R)-\alpha(u_i)+h_i+\lambda_i\right)}{\left(\alpha(R)-\alpha(u_i)+h_i\right)\left(\alpha(R)-\alpha^-(u_i)+h_i+\lambda_i+\rho_i\right)} \qquad (2.22)$$

ให้

$$K=\prod_{i=1}^{j(t)}\frac{\left(\alpha(R)-\alpha^-(u_i)+h_{i-1}\right)\left(\alpha(R)-\alpha(u_i)+h_i+\lambda_i\right)}{\left(\alpha(R)-\alpha(u_i)+h_i\right)\left(\alpha(R)-\alpha^-(u_i)+h_i+\lambda_i+\rho_i\right)}$$

จากสมการ (2.22) พบว่า ถ้า u_i เป็นจุดที่เกิดการตัดทิ้งของข้อมูล (Censoring Point) แล้วจะได้ว่า $\rho_i=0$ และทำให้ $h_{i-1}=h_i+\lambda_i$ ดังนั้น

$$K=\prod_{i=1}^{j(t)}\frac{\left(\alpha(R)-\alpha(u_i)+h_i+\lambda_i\right)}{\left(\alpha(R)-\alpha(u_i)+h_i\right)}$$

และเมื่อ u_i ไม่ใช่จุดการตัดทิ้งของข้อมูล จะได้ว่า $\lambda_i=0$ และทำให้ $h_{i-1}=h_i+\rho_i$ ดังนั้น

$$K=1$$

นั่นคือสามารถคำนวณฟังก์ชันการอยู่รอดจากค่าคาดหวังภายหลังของ S(t) เมื่อการแจกแจงก่อนของ F เป็นกระบวนการดีริชเลต์ ดังนี้

$$E(S(t)/(Z_i,\delta_i)) = \frac{\alpha(R)-\alpha(t)+h_{j(t)}}{\alpha(R)+n} \cdot \prod_{i=1}^{j(t)} \frac{\left(\alpha(R)-\alpha(u_i)+h_i+\lambda_i\right)}{\left(\alpha(R)-\alpha(u_i)+h_i\right)} \qquad (2.23)$$

ในทางปฏิบัติแล้ว จะต้องมีการกำหนดการแจกแจงก่อนการทดลอง (Prior guess) ของ S(t) โดยนำความรู้จากข้อมูลในอดีตมาช่วยในการตัดสินใจ และเพื่อเป็นประโยชน์ในการประมาณค่าพารามิเตอร์ โดยที่การแจกแจงก่อนการทดลองของ S(t) สำหรับกระบวนการดีริชเลต์ สามารถคำนวณได้ดังนี้คือ

$$E(S(t)) = \frac{\alpha(t)}{\alpha(R)}$$

ถ้าให้การแจกแจงก่อนการทดลองของ S(t) คือ $S_o(t)$ และ α (R) เป็นขนาดตัวอย่างก่อนการทดลอง (Prior sample size) แล้วจะได้ว่า

$$\alpha(t) = S_o(t).\alpha(R) \qquad (2.24)$$

4.2 <u>แบบกระบวนการโฮโมจีเนียสอย่างง่าย</u> (Simple Homogeneous Process)

กำหนดให้ Y_t เป็นกระบวนการโฮโมจีเนียสอย่างง่ายสามารถเขียนฟังก์ชันที่ให้โมเมนต์ของ Y_t ได้ดังนี้

$$M_t(\theta) = \exp\left[\gamma(t)\int_0^{\infty}\left(e^{-z\theta}-1\right)e^{-\alpha z}\left(1-e^{-z}\right)^{-1}dz\right] \qquad (2.25)$$

$$dN(z) = e^{-\alpha z}\left(1-e^{-z}\right)^{-1}dz$$

ถ้าให้

$$\varphi(\alpha,\beta,N) = \int_0^{\infty} e^{-(\alpha+\tau)z}(1-e^{-z})^{\beta-1}dz$$

$$=\Gamma(\alpha+\tau)\Gamma(\beta)/\Gamma(\alpha+\beta+\tau) \qquad (2.26)$$

กรณีที่ $\beta = 1$ จะได้ว่า

$$\varphi(\alpha, 1, N) = \frac{\Gamma(\alpha + \tau)}{\Gamma(\alpha + 1 + \tau)} = \frac{1}{\alpha + \tau}$$

จากสมการที่ (2.12) และ (2.14) เราสามารถเขียนให้อยู่ในรูปของ $\varphi(\alpha, \beta, N)$ ได้ดังนี้

$$C_u(\alpha, \beta) = \varphi(\alpha, \beta, N) / \varphi(0, 1, N) \qquad \text{ถ้า } \beta \geq 1$$

$$= 1 \qquad \beta = 0 \quad \text{———} \quad (2.27)$$

ดังนั้นจะได้ว่า

$$R_t(h) = e^{-\gamma(t)\varphi(h,1,N)}$$

$$R_t(h) = e^{-\gamma(t)/(h+\tau)} \quad \text{———} \quad (2.28)$$

นั่นคือสามารถคำนวณฟังก์ชันการอยู่รอดจากค่าคาดหวังภายหลังของ S(t) เมื่อการแจกแจงก่อนของ F เป็นกระบวนการโฮโมจีเนียสอย่างง่าย ดังนี้

$$E\left(S(t) / (Z_i, \delta_i)\right) = e^{-\frac{\gamma(t)}{h_{j(t)}+\tau}} \prod_{i=1}^{j(t)} \left[e^{\frac{\gamma(u_i)(h_{i-1}-h_i)}{(h_{i-1}+\tau)(h_i+\tau)}} \left(\frac{h_i + \lambda_i + \tau}{h_i + \lambda_i + \delta_i + \tau} \right) \right] \quad \text{———} \quad (2.29)$$

ในทางปฏิบัติแล้ว จะต้องมีการกำหนดการแจกแจงก่อนการทดลอง (Prior guess) ของ S(t) โดยนำความรู้จากข้อมูลในอดีตมาช่วยในการตัดสินใจ และเพื่อเป็นประโยชน์ในการประมาณค่าพารามิเตอร์ โดยที่การแจกแจงก่อนการทดลองของ S(t) สำหรับกระบวนการโฮโมจีเนียสอย่างง่าย สามารถคำนวณได้ดังนี้คือ

$$E(S(t)) = M_t(1) = e^{-\gamma(t)/\tau} \quad \text{———} \quad (2.30)$$

ดังนั้นถ้าให้การแจกแจงก่อนการทดลองของ S(t) คือ $S_o(t)$ แล้วจะได้ว่า

$$S_o(t) = e^{-\gamma(t)/\tau} \quad \text{———} \quad (2.31)$$

และเมื่อกำหนดค่าให้กับพารามิเตอร์ τ จะสามารถคำนวณค่าพารามิเตอร์ $\gamma(t)$ สำหรับทุกค่าของ t ได้ดังนี้

$$\gamma(t) = -\tau \log S_o(t) \quad \text{———} \quad (2.32)$$