บทที่ 2
ทฤษฎีและการแจกแจงที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงทฤษฎีที่เกี่ยวข้องกับการแจกแจงไวบูลล์ ซึ่งเป็นรายละเอียดเกี่ยวกับสถิติลำดับ (order statistics) ทฤษฎีค่าที่สุด (extreme value theory) อัตราการเสี่ยง (harzard rate) ความเชื่อถือได้ (reliability) โดยทฤษฎีเหล่านี้จะเป็นที่มาของการแจกแจงไวบูลล์ นอกจากนี้ก็จะให้รายละเอียดเกี่ยวกับการศึกษาตัวแปรสุ่มเวลาได้แก่ การแจกแจงปัวซอง การแจกแจงเอ็กซ์โปเนนเชียล และการแจกแจงแกมมา ซึ่งจะทำให้เข้าใจการแจกแจงไวบูลล์มากยิ่งขึ้น

2.1 สถิติลำดับ (Order Statistics)

สถิติลำดับ มีบทบาทที่สำคัญในทฤษฎีความเชื่อถือได้ (reliability theory) โดยเฉพาะอย่างยิ่งในการทดสอบอายุการใช้งาน (life testing) ความทนทานของผลิตภัณฑ์ เนื่องจากในการทำงานของผลิตภัณฑ์ที่มีส่วนประกอบหลาย ๆ ส่วน แต่ละส่วนมีความทนทานหรืออายุการใช้งานแตกต่างกัน ถ้ามีส่วนประกอบส่วนใดส่วนหนึ่งชำรุด ก็จะทำให้ระบบการทำงานทุกส่วนของผลิตภัณฑ์ชำรุดตามไปด้วย ดังนั้นอายุการใช้งานของผลิตภัณฑ์จึงเท่ากับอายุการใช้งานต่ำสุดของแต่ละส่วนประกอบ ซึ่งหมายถึงสถิติลำดับที่น้อยที่สุด (the smallest order statistics) นั่นเอง

นิยาม 2.1 ให้ $x_1, x_2, x_3, \ldots, x_n$ เป็นตัวอย่างสุ่มขนาด n จากประชากรที่มีฟังก์ชันความหนาแน่น $f_X(x;\theta)$ และฟังก์ชันการแจกแจง $F_X(x)$ โดยที่

ให้ $x_{(1)} = \min(x_1, x_2, \ldots, x_n)$ จะสังเกตได้ว่า $x_{(1)}$ เป็นค่าใด ๆ ก็ได้ของค่าสังเกต x_i ทั้งหมด n ค่า จึงเรียก $x_{(1)}$ ว่าสถิติลำดับที่หนึ่ง (first order statistic)

$x_{(2)} = x$ ที่มีค่าต่ำสุดเป็นอันดับสองนับจากค่าน้อยของ $x_1, x_2, \ldots, x_n$ เรียกว่าสถิติลำดับที่สอง (second order statistic)

$\vdots$

$x_{(n)} = \max(x_1, x_2, \ldots, x_n)$ ซึ่งเป็น $x_{(n)}$ ที่มีค่ามากที่สุดในบรรดา $x_1, x_2, \ldots, x_n$ จึงเรียก $x_{(n)}$ ว่าสถิติลำดับที่ n (nth order statistic)

ดังนั้น $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ จึงเป็นสถิติลำดับ (order statistics) ของตัวอย่างสุ่ม $x_1, x_2, \ldots, x_n$ โดยที่ $x_{(1)} \le x_{(2)} \le \ldots \le x_{(n)}$

โดยทั่วไปแล้ว $x_{(i)}$ จะเรียกว่าสถิติลำดับที่ i (ith order statistic) และจะมีค่าสังเกต i-1 ค่า ก่อนสถิติลำดับที่ i (ก่อน $x_{(i)}$) วัตถุประสงค์ของเรื่องนี้เป็นการศึกษาคุณสมบัติสถิติลำดับที่ i

ในการทดสอบอายุการใช้งาน ที่ตัวอย่างสุ่มประกอบด้วยเวลาที่วัตถุเกิดขัดข้องในการทำงานหลาย ๆ ลักษณะ โดยที่ $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ เป็นค่าสังเกตที่เกิดขึ้นตามลำดับ นั่นคือ ค่าสังเกตแรกจะเป็นเวลาที่เกิดความขัดข้อง (failure) น้อยที่สุด โดยที่ค่าสังเกตแต่ละค่าเป็นสถิติลำดับค่าหนึ่ง

2.1.1 การแจกแจงสถิติลำดับที่ i (Distribution of ith Order Statistics)

ในการพิจารณาสถิติลำดับที่ i , $x_{(i)}$ ที่เกิดจากฟังก์ชันความหนาแน่น $f_X(x;\theta)$ และฟังก์ชันการแจกแจง $F_X(x;\theta)$ ถ้า $x_1, x_2, \ldots, x_n$ เป็นตัวอย่างสุ่มขนาด n หรือค่าสังเกต n ค่า สิ่งที่เราต้องการทราบก็คือ ฟังก์ชันความหนาแน่นของ $x_{(i)}$ หรือ $f_{X(i)}(x;\theta)$

ให้ E แทนเหตุการณ์ที่สถิติลำดับที่ i , $x_{(i)}$ เป็นค่าที่อยู่ระหว่าง x และ $x+\Delta x$ ซึ่งหมายความว่ามีค่าสังเกต i-1 ค่าจะเกิดขึ้นก่อน x และ n - i จะเกิดขึ้นหลังจาก $x+\Delta x$ จากการใช้ multinomial probability mass function จะได้ว่า

$$P\{E\} = P\left\{x \le x_{(i)} \le x+\Delta x\right\}$$

$$= \frac{n!}{(i-1)!\ \ 1!(n-i)!}\left[F_x(x)\right]^{i-1} f_X(x)\Delta x\left[1-F_x(x)\right]^{n-i}$$

โดยข้อจำกัดที่ว่าเมื่อ $\Delta x \longrightarrow 0$ จะได้

$$f_{x(i)}(x) = \frac{i(i-1)!\ \ \ n!}{(i-1)!\ \ (n-i)!\ \ \ i!}\left[F_x(x)\right]^{i-1}\left[1-F_x(x)\right]^{n-i} f_x(x)$$

$$= i\binom{n}{i}\left[F_x(x)\right]^{i-1}\left[1-F_x(x)\right]^{n-i} f_x(x)$$

โดยเฉพาะอย่างยิ่ง ถ้า $i=1$ แล้ว จะได้ฟังก์ชันความหนาแน่นของสถิติลำดับที่หนึ่งซึ่งมีค่าน้อยที่สุด เป็น

$$f_{x(1)}(x) = 1\binom{n}{1}\left[F_x(x)\right]^{i-1}\left[1-F_x(x)\right]^{n-i}f_x(x)$$

$$= \frac{n(n-1)!}{1!\ (n-i)!}\left[F_x(x)\right]^{1-1}\left[1-F_x(x)\right]^{n-1}f_x(x)$$

$$= n\left[1-F_x(x)\right]^{n-1}f_x(x)$$

และถ้า $i=n$ แล้วจะให้ฟังก์ชันความหนาแน่นสถิติลำดับที่ n หรือสถิติลำดับสุดท้ายซึ่งมีค่ามากที่สุด เป็น

$$f_{x(n)}(x) = n\binom{n}{n}\left[F_x(x)\right]^{n-1}\left[1-F_x(x)\right]^{n-n}f_x(x)$$

$$= n\left[F_x(x)\right]^{n-1}f_x(x)$$

ดังนั้นฟังก์ชันการแจกแจงของ $x_{(1)}$ และ $x_{(n)}$ หาได้ดังนี้

จาก $$F_{x(1)}(x) = P(x_{(1)} \le x)$$

$$= 1-P(x_{(1)} \ge x)$$

แต่ $$P(x_{(1)} \ge x) = P(x_{(1)} \ge x, x_{(2)} \ge x, \cdots, x_{(n)} \ge x)$$

$$= P(x_{(1)} \ge x)\,P(x_{(2)} \ge x)\cdots P(x_{(n)} \ge x)$$

$$= \left[1-F_x(x)\right]\left[1-F_x(x)\right]\cdots\left[1-F_x(x)\right]$$

$$= \left[1-F_x(x)\right]^n$$

$$\therefore\ F_{x(1)}(x) = 1-\left[1-F_x(x)\right]^n$$

ในทำนองเดียวกัน จะได้

$$F_{x(n)}(x) = P(x_{(n)} \le x)$$

$$= P(x_{(1)} \le x, x_{(2)} \le x, \cdots, x_{(n)} \le x)$$

$$= P(x_{(1)} \le x) \cdot P(x_{(2)} \le x) \cdots P(x_{(n)} \le x)$$

$$= F_x(x) \cdot F_x(x) \cdots F_x(x)$$

$$= [F_x(x)]^n$$

$$\therefore F_{x(n)}(x) = [F_x(x)]^n$$

ตัวอย่างของการประยุกต์ใช้ในปัญหาเกี่ยวกับความเชื่อถือได้ซึ่งรูปแบบที่พบเสมอก็คือ ระบบการทำงานที่เป็นระบบแบบลูกโซ่ (chain model) หรือระบบที่ทำงานต่อเนื่องกัน สำหรับระบบที่ทำงานแบบลูกโซ่ จะเหมาะสมกับระบบที่ประกอบด้วยส่วนประกอบหรือข้อต่อที่เชื่อมต่อกัน ควรจะมีจำนวนส่วนประกอบจำนวนน้อย ๆ (n เล็ก) เพราะถ้าส่วนประกอบระบบแต่ละส่วนมีข้อบกพร่อง แต่เมื่อทำงานร่วมกันเป็นระบบที่ทำงานต่อเนื่องกันเป็นลูกโซ่ ระบบก็ยังเป็นระบบที่แข็งแรงสามารถทำงานได้เป็นปกติ แต่ในความเป็นจริงแล้วเมื่อส่วนประกอบหรือข้อต่อของระบบส่วนใดส่วนหนึ่งที่บกพร่องอยู่เกิดการชำรุดขึ้น ก็จะทำให้ระบบทั้งหมดชำรุดตามไปด้วย

สมมติว่าการแจกแจงอายุการใช้งานของส่วนประกอบ มีการแจกแจงเอ็กซ์โปเนนเชียล $F_x(x)$ จะเป็นการแจกแจงของสถิติลำดับน้อยที่สุด หรือสถิติลำดับที่หนึ่ง $x_{(1)}$

นั่นคือ $F_{x(1)}(x) = 1-[1-F_x(x)]^n$

เนื่องจาก $1-F_x(x) = e^{-\lambda x}$

ดังนั้นจะได้ $F_{x(1)}(x) = 1-e^{-n\lambda x}$

เมื่อหาอนุพันธ์อันดับหนึ่งของ $F_{X(1)}(x)$ จะได้ฟังก์ชันความหนาแน่นของการแจกแจงอายุการใช้งาน เป็น

$$f_x(x) = \lambda n e^{-\lambda n x}$$

ซึ่ง $f_X(x)$ ที่ได้นี้เป็นฟังก์ชันความหนาแน่นของการแจกแจงเอ็กซ์โปเนนเชียลอีกรูปแบบหนึ่งด้วยพารามิเตอร์ λn ซึ่งคุณสมบัตินี้ก็คือสถิติลำดับที่น้อยที่สุด (the smallest order statistic) จากการแจกแจงเอ็กซ์โปเนนเชียล จะเป็นการแจกแจงเอ็กซ์โปเนนเชียลด้วย เราเรียกคุณสมบัตินี้ว่าการจำลองแบบตัวเอง (self-reproducting) ส่วนการแจกแจงอื่นที่มีคุณสมบัตินี้ก็คือการแจกแจงไวบูลล์ ซึ่งจะแสดงให้ทราบต่อไปในบทที่ 3

ในการหาค่าเฉลี่ยและความแปรปรวนของ $f_{X(1)}(x)$ เปรียบเทียบกับ ค่าเฉลี่ยและความแปรปรวนของ $f_X(x)$ คงจะจำได้ว่า เมื่อ X มีการแจกแจงแบบเอ็กซ์โปเนนเชียล ซึ่งมีฟังก์ชันความหนาแน่นเป็น $f(x) = \lambda e^{-\lambda x}$; $x \geq 0$, $\lambda > 0$ แล้วค่าเฉลี่ยและความแปรปรวนเป็น

$$E(X) = \mu = \frac{1}{\lambda} = \theta$$

และ $$V(X) = \sigma^2 = \frac{1}{\lambda^2} = \theta^2$$

ดังนั้นค่าเฉลี่ยและความแปรปรวนของ $X_{(1)}$ เป็นดังนี้

เนื่องจาก $E(X) = nE(X_{(1)})$

ดังนั้น $E(X_{(1)}) = \dfrac{1}{n\lambda} = \dfrac{\theta}{n}$

และ จาก $V(X) = n^2 V(X_{(1)})$

ดังนั้น $V(X_{(1)}) = \dfrac{1}{n^2\lambda^2} = \left(\dfrac{\theta}{n}\right)^2$

สามารถแสดงการแจกแจงอายุการใช้งานของส่วนประกอบที่แข็งแรงที่สุด หรือสถิติลำดับที่ n, $x_{(n)}$ ซึ่งกำหนดโดย

$$F_{x(n)}(x) = [F_x(x)]^n$$

$$= (1-e^{-\lambda x})^n$$

ดังนั้น $$f_{x(n)}(x) = n\lambda(1-e^{-\lambda x})^{n-1}e^{-\lambda x}$$

2.1.2 การแจกแจงร่วมของสถิติลำดับ r ลำดับแรก

(Joint Distribution of The First r Order Statistics)

จากการศึกษาการแจกแจงสถิติลำดับที่ i ในหัวข้อ 2.1.1 จะได้ว่า

$$P\{E\} = P\{x_1 \le x_{(1)} \le x_1 + \Delta x_1, x_2 \le x_{(2)} \le x_2 + \Delta x_2, \cdots,$$

$$x_n \le x_{(r)} \le x_r + \Delta x_r\}$$

โดยที่เหตุการณ์ E แทนเหตุการณ์ที่ไม่มีค่าสังเกตใดเลยที่เกิดขึ้นก่อน x_1 การเกิดขึ้นของค่าสังเกตลำดับที่หนึ่งเกิดขึ้นระหว่าง x_1 และ $x_1 + \Delta X_1$ และไม่มีค่าสังเกตใดเลยที่เกิดขึ้นระหว่าง $x_1 + \Delta x_1$ และ x_2 ส่วนค่าสังเกตลำดับที่สองเกิดขึ้นระหว่าง x_2 กับ $x_2 + \Delta x_2$ เช่นนี้ไปเรื่อย ๆ ในที่สุดการเกิดค่าสังเกต n - r ค่าสังเกตสุดท้าย จะเกิดขึ้นหลังจาก $x_r + \Delta x_r$ จากการใช้ multinomial probability mass function จะได้

$$P(E) = \frac{n!}{0!\ 1!\ 0!\ 1!\ldots(n-r)!} f_x(x_1)\Delta x_1 f_x(x_2)\Delta x_2 \cdots f_x(x_r)\Delta x_r [1-F(x_r)]^{n-r}$$

โดยข้อจำกัดที่ว่าเมื่อ $\Delta x \to 0$ จะได้

$$f_{x(1),x(2),\ldots,x(r)}(x_1, x_2, \ldots, x_r) = \frac{n!}{(n-r)!}[1-F_{(x_r)}]^{n-r} \prod_{i=1}^{r} f_x(x_i)$$

หากว่าเลือกทุกลำดับจาก n ค่าสังเกต นั่นคือ ถ้า $r = n$ แล้ว

$$f_{x(1),x(2),\ldots,x(n)}(x_1, x_2, \ldots, x_n) = n! \prod_{i=1}^{n} f_x(x_i)$$

2.2 ทฤษฎีค่าที่สุด (Extreme Value Theory)

ในเรื่องนี้เป็นการนำเสนอส่วนหนึ่งของทฤษฎีค่าที่สุด (extreme-value) ประเด็นสำคัญของทฤษฎีนี้มุ่งที่จะศึกษา การชำรุด การสึกกร่อนของเครื่องมือ วัสดุอุปกรณ์ เนื่องจากการแตก หัก หรือการทำให้วัตถุเสียด้วยสาเหตุอื่นจากการใช้งาน ฉนวนไฟฟ้าขาด การสึกกร่อนของโลหะ ซึ่งเป็นการศึกษาของ Epstein (1960) ซึ่งได้เขียนรายงานเกี่ยวกับเรื่องนี้อย่างแพร่หลาย

ผลการศึกษา การประยุกต์ใช้ทฤษฎี ได้มุ่งเน้นอธิบายด้วยวิธีการที่มีระบบกฎเกณฑ์ เราอาจคาดหวังการแจกแจงที่แน่นอน เพื่ออธิบายการเกิดการแตกหัก หรือการสึกกร่อนของวัตถุ การประยุกต์ใช้เกี่ยวกับเรื่องนี้ได้ถูกศึกษาอย่างกว้างขวางโดย Gumbel (1958) และ Gnedenk (1943)

ทฤษฎีค่าที่สุด จะแบ่งออกเป็น 2 กรณี คือ

1. การแจกแจงของค่าที่น้อยที่สุด (the distribution of smallest values)
2. การแจกแจงของค่าที่มากที่สุด (the distribution of largest values)

แต่ในที่นี้ จะศึกษาเฉพาะกรณีการแจกแจงของค่าที่น้อยที่สุดที่เกี่ยวข้อง การแจกแจงไวบูลล์เท่านั้น

การแจกแจงค่าที่น้อยที่สุด (the distribution of smallest values)

จากการศึกษาในเรื่องการแจกแจงของสถิติลำดับที่ i ในหัวข้อ 2.1 ซึ่งจะได้การแจกแจงของสถิติลำดับที่หนึ่ง (the distribution of the first order statistic) มีรูปแบบดังนี้

$$F_{x(1)}(x) = 1-\left[1-F_x(x)\right]^n$$

ผลการศึกษาได้ประยุกต์ใช้ในกรณีของการแจกแจงแบบเอ็กซ์โปเนนเชียล เมื่อ

กำหนดให้ $1-F_x(x) = e^{-\lambda x}$ ดังนั้น

$$F_{x(1)}(x) = 1-e^{-\lambda xn}$$

ถ้าฟังก์ชัน $F_X(x)$ ไม่ใช่รูปแบบที่ชัดเจนและเข้าใจง่าย ตัวอย่างเช่นในการศึกษากำลังของฟังก์ชันการแจกแจงแบบปกติ ซึ่งเป็นเรื่องที่ไม่ใช่จะศึกษาได้ง่าย ๆ ดังนั้น การสรุปผลของ $F_{X(1)}(x)$ จึงเป็นปัญหาที่ยุ่งยาก แต่มีวิธีที่แก้ปัญหานี้ได้ โดยใช้เทคนิคของ Cramer (1946) ซึ่งสามารถประยุกต์ใช้ได้อย่างเหมาะสม โดยมีข้อสมมติว่า n มีขนาดใหญ่

นิยามตัวแปรสุ่ม η_n ดังนี้

$$\eta_n = nF_x(x_{(1)})$$

ให้ $\Gamma_n(u) = P(\eta_n \le u) \quad , \quad 0 \le u \le n$

จะได้ $\Gamma_n(u) = P\left[nF_x(x_{(1)}) \le u\right]$

$$= P\left[F_x(x_{(1)}) \le \frac{u}{n}\right]$$

$$= P\left[x_{(1)} \le F_x^{-1}\left(\frac{u}{n}\right)\right]$$

จาก $F_{x(1)}(x) = P(x_{(1)} \le x)$

ดังนั้น $\Gamma_n(u) = F_{x(1)}\left[F_x^{-1}\left(\frac{u}{n}\right)\right]$

และจาก $F_{x(1)}(x) = 1-\left[1-F_x(x)\right]^n$

$$\therefore \Gamma_n(u) = F_{x(1)}\left[F_x^{-1}\left(\frac{u}{n}\right)\right]$$

$$= 1-\left[1-F_x\left(F_x^{-1}\left(\frac{u}{n}\right)\right)\right]^n$$

$$= 1-\left[1-\left(\frac{u}{n}\right)\right]^n$$

ซึ่งเป็นจริงสำหรับทุก $F_X(x)$

กำหนดให้ $\Gamma(u) = \lim_{n\to\infty} \Gamma_n(u)$

$$= \lim_{n\to\infty}\left[1-\left(1-\left(\frac{u}{n}\right)\right)^n\right]$$

$$= \lim_{n\to\infty}\left[1-\left(1+\left(-\frac{u}{n}\right)\right)^n\right]$$

เนื่องจาก $e^{u} = \lim_{n\to\infty}\left(1+\frac{u}{n}\right)^{n}$

ดังนั้น $\Gamma(u) = 1-e^{-u}$, $u \geq 0$

ผลที่ได้นี้เป็นผลเนื่องจากสาเหตุยิ่งไปกว่านั้นอีก ก็เนื่องจากลำดับของฟังก์ชันการแจกแจง (the sequence of distribution functions) ของ $\Gamma_n(u)$ ลู่เข้าสู่ $1 - e^{-u}$ โดยที่ลำดับของตัวแปรสุ่ม η_n เป็นการแจกแจงของตัวแปรสุ่มที่เรียกว่า η

ฟังก์ชัน η ได้จากการหาอนุพันธ์ของ $1 - e^{-u}$

ดังนั้น $f_\eta(u) = \Gamma'(u)$

$$= e^{-u} \quad , \quad u \geq 0$$

เมื่อ $\eta_n = nF_x(x_{(1)})$

จะเห็นได้ชัดเจนว่าลำดับของตัวแปรสุ่ม $x_{(1)}$ มีการแจกแจงร่วมกันด้วยตัวแปรสุ่ม Y ซึ่ง

$Y = F_x^{-1}\left(\frac{\eta}{n}\right)$ ด้วยตัวแปรสุ่ม η ดังนิยาม

สำหรับสัญญลักษณ์ Γ ที่ใช้ในนิยาม $\Gamma_n(u)$ นี้ไม่ใช่สัญญลักษณ์ที่ใช้ในฟังก์ชันแกมมา ซึ่งจะเขียนเหมือนกัน แต่ความหมายแตกต่างกัน

ดังนั้น การแจกแจงจำกัดของสถิติลำดับที่หนึ่ง $x_{(1)}$ จึงกำหนดโดยการแจกแจงของ Y

ถ้ากำหนดให้ $f_x(x) = \frac{\beta(x-\gamma)^{\beta-1}}{\alpha^{\beta}}$, $\gamma \leq x \leq \alpha+\gamma ; \gamma \geq 0$ และ $\alpha, \beta > 0$

จาก $f_x(x) = \dfrac{\beta(x-\gamma)^{\beta-1}}{\alpha^\beta}$

จะได้ $f_x(x) = \int_\gamma^x f_x(x)dx$

$$= \int_\gamma^x \frac{\beta(x-\gamma)^{\beta-1}}{\alpha^\beta}dx$$

ให้ $u = x - \gamma$

$du = dx$

ดังนั้น $F_x(x) = \int_\gamma^x \dfrac{\beta u^{\beta-1}}{\alpha^\beta}du$

$$= \frac{\beta}{\alpha^\beta}\int_\gamma^X u^{\beta-1}du$$

$$= \frac{\beta}{\alpha^\beta}\left[\frac{u^\beta}{\beta}\right]$$

$$= \frac{u^\beta}{\alpha^\beta}$$

$$= \left(\frac{x-\gamma}{\alpha}\right)^\beta \quad , \quad \gamma \le x \le \alpha+\gamma$$

จากนิยาม $\eta_n = nF_x(x_{(1)})$

จะได้ $\eta_n = n\left(\dfrac{x_{(1)}-\gamma}{\alpha}\right)^\beta$

$$\frac{\eta_n}{n} = \left(\frac{x_{(1)}-\gamma}{\alpha}\right)^\beta$$

$$\left(\frac{x_{(1)}-\gamma}{\alpha}\right) = \left(\frac{\eta_n}{n}\right)^{\frac{1}{\beta}}$$

$$\therefore \; x_{(1)} = \alpha\left(\frac{\eta_n}{n}\right)^{\frac{1}{\beta}}+\gamma$$

เมื่อ $n \to \infty$ และ $\eta_n \xrightarrow{L} \eta$

$$x_{(1)} \xrightarrow{L} \alpha\left(\frac{\eta}{n}\right)^{\frac{1}{\beta}}+\gamma$$

ดังนั้น $F_{x(1)}(x) = 1-\exp\left[-n\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] \quad , \quad x \geq \gamma$

จะได้ $f_{x(1)}(x) = \eta\beta\left(\frac{x-\gamma}{\alpha}\right)^{\beta-1}\exp\left[-n\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right]$

ซึ่งฟังก์ชันความหนาแน่นของการแจกแจงของ $X_{(1)}$ ที่ได้นี้เป็นฟังก์ชันความหนาแน่นของการแจกแจงแบบไวบูลล์ และเป็นทฤษฎีสถิติที่เกิดขึ้นอย่างมีประสิทธิภาพ (Weibull, 1939)

กล่าวอีกนัยหนึ่งการแจกแจงไวบูลล์สามารถหาได้ จากการพิจารณารูปแบบอื่นอีก ตัวอย่างเช่น การแจกแจงที่ถูกต้องแน่ชัด (the exact distribution) ของสถิติลำดับที่หนึ่งที่มาจากการแจกแจงไวบูลล์ถือว่าเป็นไวบูลล์กลุ่มหนึ่ง

ให้ $F_{x(1)}(x) = 1-\exp\left[-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] \quad , \quad x \geq \gamma \;\; ; \;\; \alpha, \beta > 0 \;\; ; \;\; \gamma \geq 0$

$$= 0 \qquad\qquad , \quad x < \gamma$$

และ $F_{x(1)}(x) = 1-\exp\left[-n\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] \quad , \quad x \geq \gamma$

ซึ่งจะพบว่า $F_X(x)$ เป็นการแจกแจงไวบูลล์

$$\text{เนื่องจาก} \quad F_{x(1)}(x) = 1-[1-F_x(x)]^n$$

$$\text{ดังนั้น} \quad F_{x(1)}(x) = 1-\left[1-\left(1-\exp\left[-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right]\right)\right]^n$$

$$F_{x(1)}(x) = 1-\exp\left[-n\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right]$$

ซึ่งเป็นการแจกแจงไวบูลล์อีกรูปแบบหนึ่ง

Gumbel (1958) เรียกการแจกแจงจำกัดรูปแบบที่ได้ดังกล่าวว่า type III asymptotic distribution of the smallest extreme ซึ่งการแจกแจงเหล่านี้จะเกิดขึ้นได้เมื่อพบ 2 เงื่อนไขต่อไปนี้

1. เกี่ยวข้องกับฟังก์ชันความหนาแน่นพื้นฐานที่นิยามถูกจำกัดขอบเขตโดยที่ $F_X(x) = 0$ เมื่อ $x<\gamma$ สำหรับค่าจำกัดบางค่าของ γ

2. $F_x(x)$ มีลักษณะเช่นเดียวกับ $\alpha(x-\gamma)^{\beta}$ สำหรับ α ,β ที่มากกว่าศูนย์ โดยที่ $x \to \gamma$

ถ้าขอบเขตของฟังก์ชันความหนาแน่น (range of the density) ที่ไม่ถูกจำกัดมาจาก $\lim\limits_{x\to\infty}(-x)^n F_x(x) = \beta$ และ α,β บางค่ามากกว่าศูนย์ แล้ว

การแจกแจงจำกัดของสถิติลำดับที่น้อยที่สุดที่เสนอโดย Gumbel (1958) เรียกว่า type II asymptotic distribution of smallest extreme การแจกแจงแอสซิมโทติกที่มีลักษณะเช่นนี้จึงไม่มีประโยชน์ในการศึกษาในเรื่องของความขัดข้องเสียหาย

แต่ถ้าฟังก์ชันความหนาแน่นพื้นฐาน $f_X(x)$ เป็นรูปแบบที่มีแนวโน้มเข้าสู่ 0 โดยที่ $x \to -\infty$ การแจกแจงจำกัดของสถิติลำดับที่หนึ่งซึ่งถูกเสนอโดย Gumbel เรียกว่า type I asymptotic distribution of the smallest extreme ซึ่งการแจกแจงจำกัดของสถิติลำดับที่หนึ่ง มีรูปแบบเป็น

$$F_{x(1)}(x) = 1-\exp\left[-\exp\left(\frac{x-\gamma}{\alpha}\right)\right]$$

ซึ่ง α และ γ เป็นค่าคงที่ที่มากกว่าศูนย์ และ $-\infty < x < \infty$

Gumbel ได้ใช้ type I asymptotic distribuiton อย่างกว้างขวาง ในการศึกษาเกี่ยวกับปรากฏการณ์ธรรมชาติที่มีค่าสุด (extremal phenomena) ด้วยเหตุผลนี้จึงเรียกกลุ่มการแจกแจงนี้ว่า การแจกแจงกัมเบล (gumbel distribution)

สรุปแล้วเราสามารถกำหนดรูปแบบการแจกแจงที่เป็นไปได้ไว้ตายตัวได้ 3 รูปแบบ สำหรับการแจกแจงของสถิติลำดับที่น้อยที่สุด ดังนี้

รูปแบบที่ 1 (type I)

$$F_{x(1)}(x) = 1-\exp\left[-\exp\left(\frac{x-\gamma}{\alpha}\right)\right] \ , \ -\infty < x < \infty \ ; \ \alpha > 0$$

รูปแบบที่ 2 (type II)

$$F_{x(1)}(x) = 1-\exp\left[-\left(\frac{-x-\gamma}{\alpha}\right)^{-\beta}\right] \ , \ -\infty < x \le \gamma \ ; \ \alpha,\beta > 0$$

รูปแบบที่ 3 (type III)

$$F_{x(1)}(x) = 1-\exp\left[-\left(\frac{x-\gamma}{\alpha}\right)^{\beta}\right] \ , \ \gamma \le x < \infty \ ; \ \alpha,\beta > 0 \, , \gamma \ge 0$$

การแจกแจงทั้ง 3 รูปแบบนี้มีคุณสมบัติที่เรียกว่า "คุณสมบัติปิดตัวเอง" (self-locking property) ซึ่งหมายความว่า ถ้าฟังก์ชันการแจกแจงพื้นฐาน $F_X(x)$ เป็นรูปแบบใดแบบหนึ่งของทั้ง 3 รูปแบบ แล้วการแจกแจงสถิติลำดับน้อยที่สุด $x_{(1)}$ สำหรับทุก ๆ n มีรูปแบบเหมือนกัน เพียงแต่เปลี่ยนแปลงพารามิเตอร์ที่เหมาะสมเท่านั้น

2.3 แนวความคิดเกี่ยวกับอัตราการเสี่ยง (The Hazard-rate Concept)

ในการวัดความเชื่อถือได้ของเครื่องมือ เครื่องใช้ วัสดุอุปกรณ์ ที่เกิดผิดปกติ จากการใช้งาน ซึ่งจะเกิดขึ้นนาน ๆ ครั้ง เป็นการเกิดที่ผิดพลาด ขัดข้อง ล้มเหลวหรือเสียหายในการใช้งาน ในช่วงเวลาใดเวลาหนึ่ง การแจกแจงความขัดข้อง (failure distribution) เป็นการแสดงถึงความพยายามที่จะอธิบายด้วยวิธีการทางคณิตศาสตร์ของอายุการใช้งานของเครื่องมือ เครื่องใช้อุปกรณ์ สิ่งประดิษฐ์ต่าง ๆ ที่มีสาเหตุเกี่ยวกับวัตถุ หรือเป็นลักษณะโดยรวม อาจจะเป็นผล

เนื่องมาจากการชำรุดการขัดข้องในการทำงาน ของวัสดุอุปกรณ์ที่เกิดขึ้นในทันทีทันใดโดยเฉพาะ ในการนำเสนอเรื่องนี้เป็นการยากที่จะแยกต้นเหตุด้านสภาพของวัตถุ และการอธิบายทางคณิตศาสตร์ของทุกเหตุการณ์ได้ ดังนั้นทางเลือกของการแจกแจงของความเสียหาย จึงถือว่าเป็นศิลปะอย่างหนึ่ง ถ้าการทดลองอันหนึ่งที่เชื่อมั่นอยู่บนการบันทึกการสังเกตที่แท้จริงของเวลาที่อุปกรณ์จะเกิดการขัดข้องในขณะทำงาน กระทั่งรู้ข้อแตกต่างระหว่างฟังก์ชันความน่าจะเป็นที่ไม่สมมาตร (nonsymmetrical probability function) ต่าง ๆ การแจกแจงความขัดข้อง ก็ยังคงเผชิญกับปัญหาอีก เพราะว่าการแจกแจงไม่สมมาตร (nonsymmetric distribution) เป็นความแตกต่างที่สำคัญที่ปลายหางของกราฟ และค่าสังเกตที่แท้จริงจะมีน้อย โดยเฉพาะอย่างยิ่งที่ปลายหางทางด้านขวามือ (right-hand tail) เนื่องจากข้อจำกัดของขนาดตัวอย่าง

ในการตรวจสอบปัญหาเหล่านี้ เป็นสิ่งที่น่าสนใจถึงแนวความคิดในอันที่เป็นไปได้ที่จะทราบข้อแตกต่างระหว่างฟังก์ชันการแจกแจงความแตกต่าง (the different distribution functions) บนพื้นฐานของการพิจารณารูปร่างลักษณะวัตถุ ดังนั้นแนวความคิดนี้จึงอยู่บนพื้นฐานของฟังก์ชันอัตราการขัดข้อง (the failure-rate function) ซึ่งเป็นที่รู้จักกันในบทความเกี่ยวกับความเชื่อถือได้ เช่นอัตราการเสี่ยง ในทางสถิติประกันภัยอัตราการเสี่ยง จะถูกเรียกว่าพลังของความตาย (foree of mortality) ในทฤษฎีค่าที่สุด เรียกว่าฟังก์ชันความหนาแน่น และในทางเศรษฐศาสตร์มีบทบาทที่มีผลต่อกันและกันเรียกว่า Mill's Ratio

กำหนดให้ F(x) เป็นฟังก์ชันการแจกแจงของเวลาที่วัตถุจะเสียด้วยตัวแปรสุ่ม X และให้ f(x) เป็นฟังก์ชันความหนาแน่นน่าจะเป็น (probability density function) ดังนั้นอัตราการขัดข้อง เขียนแทนด้วย h(x) จะนิยามได้ดังนี้

$$h(x) = \frac{f(x)}{1-F(x)}$$

ในที่นี้ 1- F(x) เรียกว่าความเชื่อถือได้ ที่เวลา x ซึ่งจะเขียนแทนด้วย R(x) หรือ $\bar{F}(x)$

$$\text{นั่นคือ} \quad h(x) = \frac{f(x)}{R(x)}$$

อัตราการเสี่ยงเป็นฟังก์ชันของเวลาฟังก์ชันหนึ่งมีการอธิบายด้วยความน่าจะเป็น กล่าวคือ $h(x)\Delta x$ หมายถึงความน่าจะเป็นที่อุปกรณ์จะมีอายุ x จะชำรุดในช่วง $(x, x+\Delta x)$ คือ

$$h(x) = \lim_{\Delta x \to 0}\left[\frac{P\left\{\text{วัสดุอุปกรณ์ที่อายุ } x \text{ จะเสียในช่วง } (x, x+\Delta x) \middle| \text{โดยที่มีอายุการใช้งานเริ่มต้นที่ } x\right\}}{\Delta x}\right]$$

บนพื้นฐานของการพิจารณาลักษณะรูปร่างของวัตถุ สิ่งหนึ่งก็คือการมีสิทธิ์ที่จะเลือกรูปแบบฟังก์ชัน h(x) สำหรับวัสดุอุปกรณ์โดยเฉพาะ วิธีหนึ่งก็คือการหาอนุพันธ์ของสมการ h(x) ที่ได้มาจาก f(x) และ F(x) ที่สามารถจะค้นพบได้ใหม่อีก

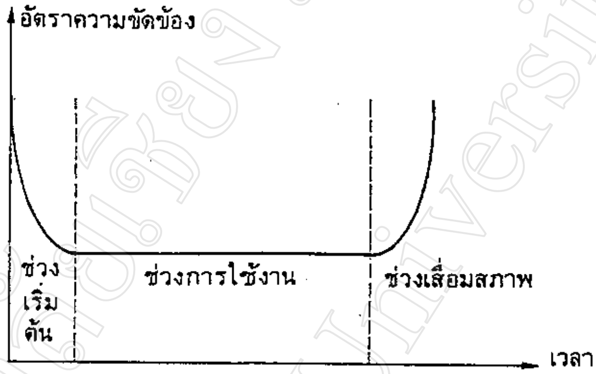
ทางเลือกของ h(x) โดยทั่วไปมี 3 ช่วงของการผิดพลาด ซึ่งเป็นที่รู้จักกันดี เนื่องจากมีคุณลักษณะพิเศษที่เกี่ยวข้องกับเวลา

ช่วงแรก เรียกว่า ช่วงเริ่มต้นที่จะขัดข้อง (initial failure) จะแสดงให้เห็นด้วยตัววัตถุเองในช่วงสั้น ๆ หลังจากเวลาเริ่มต้นที่ $x = 0$ และเริ่มต้นอย่างค่อยเป็นค่อยไปทีละเล็กทีละน้อยจนกระทั่งลดลงระหว่างช่วงแรกของการปฏิบัติงาน ตัวอย่างที่จะทำให้เข้าใจดีขึ้นของช่วงแรกนี้จะเห็นได้จากตารางชีพมาตรฐาน (the standard human mortality table) จากตารางสมมติว่าเด็กที่เกิดจนกระทั่งอายุ 10 ขวบ เด็กคนหนึ่งมีโอกาสที่เสียชีวิตด้วยความไม่สมบูรณ์ของร่างกายอันเกิดจากกรรมพันธุ์ แต่จะมีชีวิตผ่านช่วงนี้ไปได้เกือบทั้งหมดอย่างแน่นอน โดยไม่ขึ้นอยู่กับความบกพร่องหรือความไม่สมบูรณ์ของร่างกาย

ช่วงที่สอง เรียกว่าช่วงที่มีโอกาสเสียหรือช่วงเสี่ยงที่จะเสีย (chance failure) จะเกิดขึ้นระหว่างช่วงที่วัสดุอุปกรณ์แสดงอัตราการขัดข้อง ซึ่งคงที่ค่าหนึ่ง โดยทั่วไปน้อยกว่าผลที่ได้ระหว่างช่วงเริ่มต้น สาเหตุการขัดข้องของวัตถุนี้เป็นผลเนื่องจากความผิดปกติอย่างรุนแรงขณะใช้งาน และไม่สามารถจะทราบสาเหตุหรือสภาพแวดล้อมที่เหมาะสมที่จะเกิดเหตุการณ์ความผิดพลาดขึ้นระหว่างการควบคุมเวลาของอุปกรณ์ที่กำลังใช้งานอยู่ สำหรับตัวอย่างตารางชีพซึ่งเป็นสิ่งที่สมมติว่าคนที่ตายในช่วงอายุ 10 ปีถึง 30 ปี โดยทั่วไปแล้วจะเสียชีวิตด้วยอุบัติเหตุหรือปัจจัยสุ่มภายนอก

ช่วงที่สาม เรียกว่า ช่วงเสื่อมสภาพหรือสึกหรอ (wear-out failure) ซึ่งเกี่ยวข้องกับการใช้อุปกรณ์เป็นเวลานาน ๆ ทำให้ถูกทำลายทีละเล็กทีละน้อย หรือวัตถุได้รับการกระแทกเป็นเวลานาน ๆ จะทำให้วัตถุเสียความยืดหยุ่น จึงทำให้วัตถุหรืออุปกรณ์เสื่อมสภาพสำหรับตัวอย่างในตารางชีพ จะพบว่าหลังจากอายุ 30 ปีแล้ว สัดส่วนที่เสียชีวิตก็จะเพิ่มขึ้น เนื่องมาจากการมีอายุมากขึ้นนั่นเอง

ทั้ง 3 ช่วงของความการขัดข้องถูกเสนอในลักษณะพิเศษด้วยโค้งอ่างอาบน้ำ (the bath-tub curve) ดังรูปที่ 2.1



รูปที่ 2.1 แสดงอัตราความขัดข้องของผลิตภัณฑ์

2.4 ความเชื่อถือได้ (The Reliability)

ในการทดสอบอายุการใช้งานของผลิตภัณฑ์ เครื่องมืออุปกรณ์ต่าง ๆ ในงานทางด้านวิทยาศาสตร์และวิศวกรรมศาสตร์ จะเรียกว่า ความเชื่อถือได้ของผลิตภัณฑ์หรือเครื่องมืออุปกรณ์ ซึ่งหมายถึงความสามารถในการที่ผลิตภัณฑ์หรืออุปกรณ์นั้นยังทำงานได้เป็นปกติภายใต้เงื่อนไขและภายในช่วงเวลาที่กำหนด ความเชื่อถือได้ สามารถวัดค่าได้โดยกำหนดเป็นความน่าจะเป็นที่ผลิตภัณฑ์หรืออุปกรณ์จะทำงานได้ภายใต้เงื่อนไขและช่วงเวลาที่กำหนด สังเกตว่าการควบคุมคุณภาพของผลิตภัณฑ์ จะกระทำ ณ เวลาที่ผลิตสินค้า และเป็นการควบคุมให้สินค้ามีลักษณะคุณภาพตรงตามข้อกำหนดหรือมาตรฐาน อย่างไรก็ตามสินค้าที่มีคุณภาพดี ไม่จำเป็นว่าจะต้องมีความเชื่อถือได้สูง ชิ้นส่วนหรือส่วนประกอบต่าง ๆ อาจผ่านการตรวจสอบตามข้อกำหนดต่าง ๆ แต่เมื่อประกอบแต่ละชิ้นส่วนเข้าด้วยกัน ความจำเป็นในการที่แต่ละชิ้นส่วนหรืออะไหล่ต้องทำงานสัมพันธ์กัน อาจทำให้ความเชื่อถือได้ของสินค้าลดลง นอกจากนี้สินค้าที่มีความเชื่อถือได้สูงภายใต้สภาพแวดล้อมหนึ่ง อาจจะมีความเชื่อถือได้ลดลงภายใต้อีกสภาพแวดล้อมหนึ่ง เช่น แบตเตอรี่รถยนต์ อาจทำงานได้ดีที่อุณหภูมิประเทศไทย แต่อาจใช้งานไม่ได้ถ้านำไปใช้ในประเทศที่อากาศหนาวจนอุณหภูมิต่ำกว่าศูนย์องศาเซนติเกรด เป็นต้น

ในการศึกษาความเชื่อถือได้หรือการทดสอบอายุการใช้งานของชิ้นส่วน เครื่องมือ อุปกรณ์ต่าง ๆ ในระหว่างใช้งาน โดยทั่วไปแล้วจะใช้ตัวแบบเอ็กซ์โปเนนเชียลภายใต้ข้อสมมุติว่าในการทำงานของเครื่องมืออุปกรณ์นั้นจะแสดงอัตราการขัดข้องหรืออัตราเสี่ยง มีค่าคงที่ตลอดเวลา ถ้าอัตราการขัดข้องคงที่ตลอดเวลา ขณะที่ใช้งาน ก็จะทำให้ความเชื่อถือได้ของเครื่องมืออุปกรณ์นั้นคงที่ด้วย แต่ในความเป็นจริงในขณะที่อุปกรณ์เครื่องมือเครื่องใช้ต่าง ๆ ถูกใช้งาน จะแสดงอัตราขัดข้อง ไม่คงที่ตลอดเวลา อาจจะแสดงอัตราขัดข้องเพิ่มขึ้น ลดลง หรือคงที่ก็ได้ ในกรณีที่อัตราการขัดข้องไม่คงที่ จะใช้ตัวแบบเอ็กซ์โปเนนเชียลในการทดสอบความเชื่อถือได้ของเครื่องมือ อุปกรณ์ จึงไม่เหมาะสม ดังนั้นถ้าหากว่า เครื่องมือ อุปกรณ์ แสดงอัตราการขัดข้องเพิ่มขึ้น หรือลดลงในขณะที่ใช้งาน ตัวแบบไวบูลล์ จึงเป็นฟังก์ชันที่เหมาะสมในการศึกษาความเชื่อถือได้ ในการทดสอบอายุการใช้งานของเครื่องมือ อุปกรณ์ต่าง ๆ

2.4.1 ความเชื่อถือได้กับเวลาใช้งาน

ดังที่ได้กล่าวไปแล้ว ความเชื่อถือได้ของสินค้าหรือผลิตภัณฑ์ มิได้คงที่ตลอดเวลา โดยขึ้นกับเวลาใช้งานที่ผ่านไป โดยปกติในช่วงแรกของการใช้งานอาจมีโอกาสที่จะเสียได้บ่อย เนื่องจากสินค้าหรือผลิตภัณฑ์อาจเก็บไว้นานก่อนนำออกจำหน่ายชิ้นส่วนบางอย่างอาจชำรุดเสียหาย หรือหลังจากผลิตภัณฑ์ประกอบกันแล้วชิ้นส่วนแต่ละส่วนประกอบอาจทำงานขัดข้อง ไม่ราบรื่นสอดคล้องสัมพันธ์กันเนื่องจากการใช้งานใหม่จึงแสดงอัตราการขัดข้องเพิ่มขึ้นในช่วงแรกของการใช้งาน เมื่อได้ทำการแก้ไขทดแทนชิ้นส่วนหรือส่วนประกอบที่ชำรุด การทำงานของเครื่องมืออุปกรณ์ก็จะดีขึ้น อัตราการขัดข้องก็จะลดลง และอยู่ในอัตราคงที่จนถึงช่วงเมื่อหมดอายุใช้งาน ชิ้นส่วนต่าง ๆ เริ่มเสื่อมสภาพ อัตราการเสียหรือการที่ต้องหยุดซ่อมเครื่องมืออุปกรณ์ ก็จะเริ่มเพิ่มมากขึ้นเรื่อย ๆ เนื่องจากการกระแทกหรือใช้งานเป็นเวลานาน

จากรูปที่ 2.1 จะเห็นว่าอัตราการขัดข้องเทียบกับเวลาใช้งาน สามารถกำหนดเป็นความสัมพันธ์ภายใต้สภาพการทำงานระบบ ๆ หนึ่ง ณ เวลา t ใด ๆ โดยที่เราสนใจอัตราการขัดข้องที่จะเกิดขึ้น

ถ้ากำหนดให้ $f(t)$ เป็นฟังก์ชันการแจกแจงความน่าจะเป็นที่ชิ้นส่วนของอุปกรณ์จะขัดข้อง ความน่าจะเป็นที่ชิ้นส่วนของอุปกรณ์จะขัดข้องระหว่าง เวลา t ถึง $t + \Delta t$ มีค่าเท่ากับ $f(t) . \Delta t$

และความน่าจะเป็นที่ชิ้นส่วนของอุปกรณ์หรือเครื่องมือจะขัดข้อง ระหว่างช่วงเวลา 0 ถึงเวลา t เขียนเป็นฟังก์ชันการแจกแจงสะสมได้เป็น

$$F(t) = \int_0^t f(x)\,dx$$

ดังนั้นฟังก์ชันความเชื่อถือได้ (reliability function) คือ

$$\begin{aligned} R(t) &= 1 - P[\text{ชิ้นส่วนของอุปกรณ์จะเสียก่อนเวลา } t] \\ &= 1 - \int_0^t f(x)\,dx \\ &= 1 - F(t) \qquad \text{------} \quad (1) \end{aligned}$$

เมื่อฟังก์ชันความเชื่อถือได้เป็นค่าความน่าจะเป็นที่ชิ้นส่วนของอุปกรณ์หรือเครื่องมือจะใช้งานได้จนถึงเวลา t

ดังนั้นความน่าจะเป็นที่ชิ้นส่วนของอุปกรณ์จะขัดข้องในช่วงเวลา t ถึง $t + \Delta t$ กำหนดได้จาก

$$F(t + \Delta t) - F(t)$$

เมื่อกำหนดให้ A เป็นเซตของเหตุการณ์ที่ชิ้นส่วนของอุปกรณ์เกิดการขัดข้องในช่วงเวลา t ถึง $t + \Delta t$ และให้ B เป็นเซตของเหตุการณ์ที่ชิ้นส่วนของอุปกรณ์มีอายุการใช้งานจนถึงเวลา t

จะได้ความน่าจะเป็นแบบมีเงื่อนไขของการขัดข้องในระหว่างช่วงเวลา t ถึง $t + \Delta t$ เมื่อกำหนดให้ชิ้นส่วนของอุปกรณ์มีอายุการใช้งานจนถึงเวลา t เป็น

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$= \frac{F(t + \Delta t) - F(t)}{R(t)}$$

จะได้อัตราเฉลี่ยของการขัดข้องในช่วงเวลา t ถึง t + Δt โดยกำหนดให้ชิ้นส่วนของอุปกรณ์สามารถทำงานได้จนถึงเวลา t คือ

$$\left[\frac{F(t+\Delta t)-F(t)}{R(t)}\right]\times\left[\frac{1}{\Delta t}\right] = \left[\frac{F(t+\Delta t)-F(t)}{\Delta t}\right]\times\left[\frac{1}{R(t)}\right]$$

เมื่อ $\Delta t \rightarrow 0$ จะได้อัตราขัดข้อง

$$h(t) = \lim_{\Delta t\to 0}\left[\frac{F(t+\Delta t)-F(t)}{\Delta t}\right]\times\left[\frac{1}{R(t)}\right]$$

$$h(t) = \frac{F'(t)}{R(t)}$$

เมื่อ $$F'(t) = \frac{dF(t)}{dt}$$

เนื่องจากอนุพันธ์ของฟังก์ชันการแจกแจงสะสมเมื่อเทียบกับตัวแปรสุ่ม คือฟังก์ชันความหนาแน่น (density function) ของตัวแปรนั้น ๆ ดังนั้น

$$\frac{dF(t)}{dt} = f(t)$$

หรือ $$F'(t) = f(t)$$

ดังนั้น $$h(t) = \frac{f(t)}{R(t)} \quad \ldots\ldots \quad (2)$$

หรือ $$h(t) = \frac{F'(t)}{1-F(t)} \quad \ldots\ldots \quad (3)$$

กล่าวอีกนัยหนึ่ง ก็คือ ในการพิจารณาระบบการปฏิบัติงานของเครื่องมือ อุปกรณ์ ที่เวลา $t = 0$ แล้วเริ่มสังเกตการปฏิบัติงานของเครื่องมือ อุปกรณ์ จนกระทั่งเกิดการขัดข้อง ถ้าให้ X เป็นสัญลักษณ์แทนเวลาของความขัดข้อง ตัวแปรสุ่ม X นี้เป็นตัวแปรสุ่มต่อเนื่อง และมีค่าอยู่ในช่วง $[0, \infty)$ ซึ่งเป็นตัวแปรสุ่มที่มีฟังก์ชันความหนาแน่นน่าจะเป็น $f(x)$, มีฟังก์ชันความเชื่อถือได้ $R(x)$ และฟังก์ชันอัตราการขัดข้อง $h(x)$ แล้วจะได้ความสัมพันธ์ของทั้ง 3 ฟังก์ชันเป็น

$$h(x) = \frac{f(x)}{R(x)}$$

จากรูปที่ 2.1 จะได้

$$h(t) = \lim_{\Delta t \to 0} P(t \le x \le t + \Delta t \,|\, x \le t)\frac{1}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{\text{ความน่าจะเป็นที่อุปกรณ์จะขัดข้องในช่วงเวลา } [t, t+\Delta t]}{\text{ความน่าจะเป็นที่อุปกรณ์จะขัดข้องในช่วงเวลา } [t, \infty)} \cdot \frac{1}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{\int_t^{t+\Delta t} f(x)dx}{\int_t^{\infty} f(x)dx} \cdot \frac{1}{\Delta t}$$

$$= \lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{R(t)} \cdot \frac{1}{\Delta t}$$

$$= \left[\lim_{\Delta t \to 0} \frac{F(t+\Delta t) - F(t)}{\Delta t}\right] \cdot \frac{1}{R(t)}$$

$$= \frac{F'(t)}{R(t)}$$

$$= \frac{f(t)}{R(t)}$$

หรือ $f(t) = h(t)R(t)$

ดังนั้นจะเห็นว่าอัตราการขัดข้อง $h(x)$ แสดงในรูปสมการของการแจกแจงความขัดข้องและเวลาการแจกแจงความขัดข้องและเวลาสามารถแสดงในรูปสมการของฟังก์ชันความขัดข้องได้ จากสมการ (1), $R(t) = 1 - F(t)$

ถ้าหาอนุพันธ์ทั้ง 2 ข้าง ของสมการจะได้

$$\frac{d[R(t)]}{dt} = \frac{d[1-F(t)]}{dt}$$

จะได้ $R'(t) = -F'(t)$

หรือ $F'(t) = -R'(t)$

เนื่องจาก $F'(t) = f(t)$

จากสมการที่ (2) และ (3) จะได้

$$h(t) = \frac{-R'(t)}{R(t)}$$

ถ้าเราอินทิเกรตทั้ง 2 ข้าง จะได้

$$\int_0^t h(x)dx = -\int_0^t \frac{R'(x)}{R(x)}dx$$

$$= -[\ln R(x)]\Big|_0^t$$

$$= -[\ln R(t) - \ln R(0)]$$

แต่ $R(0) = 1$ เนื่องจากชิ้นส่วนของอุปกรณ์จะไม่ชำรุดก่อนเวลา $t = 0$ เนื่องจากยังไม่ได้ใช้งานเลย ความเชื่อถือได้หรือความน่าจะเป็นที่อุปกรณ์จะไม่เสียก่อนเวลา $t = 0$ จึงเท่ากับ 1

และ $\ln R(0) = \ln 1$

$$= 0$$

ดังนั้น $\int_0^t h(x)dx = -\ln R(t)$

หรือ $\ln R(t) = -\int_0^t h(x)dx$

จะได้ว่า

$$\exp[\ln R(t)] = \exp\left[-\int_0^t h(x)dx\right]$$

$$R(t) = \exp\left[-\int_0^t h(x)dx\right] \quad \ldots\ldots\ldots (4)$$

จากสมการที่ 2, $h(t) = \dfrac{f(t)}{R(t)}$

$$f(t) = h(t)R(t)$$

ดังนั้นแทน R(t) จากสมการที่ (4) จะได้

$$f(t) = h(t)\left\{\exp\left[-\int_0^t h(x)dx\right]\right\} \quad \ldots\ldots\ldots (5)$$

จากรูปที่ 2.1 จะเห็นได้ว่าอัตราการขัดข้องในช่วงการใช้งานของอุปกรณ์เครื่องใช้มีค่าค่อนข้างคงที่ (λ) ไม่ขึ้นกับเวลา t เมื่อ $\lambda > 0$

นั่นคือ $h(x) = \lambda$

เมื่อแทน $h(x) = \lambda$ ในสมการที่ (5) จะได้

$$f(t) = \lambda\left\{\exp\left[-\int_0^t \lambda dx\right]\right\}$$

$$= \lambda e^{-\lambda t} \quad \text{เมื่อ} \quad t \geq 0 \quad \ldots\ldots (6)$$

จะสังเกตว่า f(t) มีการแจกแจงแบบเอ็กซ์โปเนนเชียล เมื่อกำหนดให้อัตราขัดข้องมีค่าคงที่

ในทำนองเดียวกัน เมื่อแทน $h(x) = \lambda$ ในสมการที่ (4)

จะได้ $$R(t) = \exp\left[-\int_0^t \lambda dx\right]$$

$$= e^{-\lambda t}$$

ในกรณีนี้การแจกแจงระหว่างความขัดข้องกับเวลา จะเป็นแบบเอ็กซ์โปเนนเชียล ดังนั้นสำหรับการแจกแจงแบบเอ็กซ์โปเนนเชียล เวลาเฉลี่ยระหว่างการขัดข้อง (mean time between failure) หรือ MTBF เท่ากับ $\frac{1}{\lambda}$ และความแปรปรวนเท่ากับ $\frac{1}{\lambda^2}$

ในความเป็นจริงอัตราการขัดข้องอาจไม่คงที่ตลอดเวลาการใช้งาน แต่ขึ้นอยู่กับอายุการใช้งานที่ผ่านไป อย่างไรก็ตามการเปลี่ยนแปลงมักจะค่อยเป็นค่อยไป ฟังก์ชันที่ใช้ประมาณอัตราการขัดข้องคือ

$$h(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \quad ; \quad t \geq 0 \quad \text{........ (7)}$$

เมื่อ α และ β เป็นค่าคงที่ โดย $\alpha > 0$ และ $\beta > 0$ เนื่องจาก t ยกกำลัง β -1

1. ถ้า $\beta > 1$ จะทำให้อัตราการขัดข้องเพิ่มสูงขึ้น
2. ถ้า $\beta < 1$ จะทำให้อัตราการขัดข้องลดลง
3. ถ้า $\beta = 1$ จะทำให้อัตราการขัดข้องคงที่

   นั่นคือ $h(t) = \alpha$

ดังนั้น $f(t) = \frac{1}{\alpha}e^{-\frac{t}{\alpha}}$, $t \geq 0$ ซึ่งเป็นการแจกแจงเอ็กซ์โปเนนเชียลเช่นเดียวกันกับสมการที่ (6) เมื่อ $\lambda = \frac{1}{\alpha}$ นั่นเอง

เมื่อแทน $h(x) = \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1}$ ลงในสมการที่ (4) จะได้

$$R(t) = \exp\left[-\int_0^t h(x)dx\right]$$

$$= \exp\left[-\int_0^t \frac{\beta}{\alpha}\left(\frac{x}{\alpha}\right)^{\beta-1} dx\right]$$

$$= \exp\left[-\left(\frac{x}{\alpha}\right)^{\beta}\right]\Big|_0^t$$

$$= \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right]$$

$$\therefore\ R(t) = \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \ ; \quad t > 0 \qquad \text{..........} (8)$$

เมื่อแทนค่าสมการที่ (7) ลงในสมการที่ (5) จะได้

$$f(t) = \frac{\beta}{\alpha}\left(\frac{t}{\alpha}\right)^{\beta-1} \exp\left[-\left(\frac{t}{\alpha}\right)^{\beta}\right] \ ; \ t \geq 0 \ , \alpha > 0, \ \beta > 0$$

ซึ่งเป็นฟังก์ชันความหนาแน่นของการแจกแจงไวบูลล์ (Weibull distribution) ด้วยพารามิเตอร์ α , β

2.4.2 ความเชื่อถือได้ของระบบ

ความเชื่อถือได้ของชิ้นส่วนใด ๆ หมายถึง ความน่าจะเป็นที่ชิ้นส่วนนั้นจะทำงานได้ขนถึงช่วงเวลาที่กำหนด ภายใต้สภาพเงื่อนไขการทำงานที่กำหนด ตัวอย่างเช่น ความเชื่อถือได้ของหลอดไฟที่จะใช้งานได้ขนถึง 7,500 ชั่วโมง มีค่าเท่ากับ 0.95 หมายถึงว่าหลอดไฟนั้นถ้าใช้งานตามสภาพปกติจะมีโอกาสใช้งานได้มากกว่าหรือเท่ากับ 7,500 ชั่วโมง ด้วยความน่าจะเป็นเท่ากับ 0.95

ชิ้นส่วนหรือสินค้าเมื่อนำมาประกอบกันเข้าเป็นระบบ ความเชื่อถือได้ของระบบ จะขึ้นอยู่กับวิธีการประกอบและความเชื่อถือได้ของแต่ละชิ้นส่วนการประกอบชิ้นส่วนเข้าเป็นระบบอาจทำโดยวิธีอนุกรม (series) วิธีขนาน (parallel) และผสมกันระหว่างอนุกรมและขนาน

ระบบอนุกรม หมายถึงระบบที่ประกอบด้วยชิ้นส่วนที่ทำให้การทำงานของระบบจะทำได้ก็ต่อเมื่อทุก ๆ ชิ้นส่วนต้องทำงานได้ ตัวอย่างเช่น เครื่องรับโทรทัศน์จะใช้งานได้ก็ต่อเมื่อส่วนรับเสียง ส่วนขยายเสียง และส่วนกำลังจะต้องทำงานได้ทั้งหมด ถ้าส่วนใดส่วนหนึ่งเสียจนใช้งานไม่ได้ เครื่องรับโทรทัศน์ก็ใช้งานไม่ได้ ถ้ากำหนดให้แต่ละชิ้นส่วนมีค่าความเชื่อถือได้เป็น R_i เมื่อมี n ชิ้นส่วนในระบบ ดังนั้นสำหรับระบบอนุกรม ค่าความเชื่อถือได้ของระบบ R_i คำนวณได้จากสมการคือ

$$R_s = \prod_{i=1}^{n} R_i \qquad \text{.............(9)}$$

ตัวอย่างที่ 2.4.1 พิจารณาตัวอย่างการคำนวณค่าความเชื่อถือได้ของระบบที่มี 4 ส่วนประกอบดังรูป

$$\boxed{R_1 = 0.98} - \boxed{R_2 = 0.97} - \boxed{R_3 = 0.95} - \boxed{R_4 = 0.99}$$

ค่าความเชื่อถือได้ของระบบคือ

$$\begin{aligned} R_s &= R_1 \times R_2 \times R_3 \times R_4 \\ &= (0.98)(0.97)(0.95)(0.99) \\ &= 0.894 \end{aligned}$$

ในระบบที่มีความซับซ้อนและประกอบด้วยชิ้นส่วนจำนวนมาก เช่น ยานอวกาศ อาจประกอบด้วยชิ้นส่วนต่าง ๆ ถึง 4,000 ชิ้น ถ้ากำหนดว่าแต่ละชิ้นส่วนมีความเชื่อถือได้เท่ากับ 0.9999 ดังนั้นความเชื่อถือได้ของยานอวกาศจะมีค่าเท่ากับ

$$R_s = (0.9999)^{4000} = 0.67$$

ในการออกแบบระบบที่ประกอบด้วยชิ้นส่วนจำนวนมาก ถ้าความเชื่อถือได้ของระบบถูกกำหนดให้มีค่าตามที่ต้องการ ก็สามารถคำนวณได้ว่าค่าความเชื่อถือได้ของแต่ละชิ้นส่วนจะต้องมีค่าเท่าใด

ตัวอย่างที่ 2.4.2 ถ้าต้องการออกแบบระบบที่ประกอบด้วยชิ้นส่วน 1,000 ชิ้น ต่อกันแบบอนุกรมโดยกำหนดว่าความเชื่อถือได้ของระบบจะต้องมีค่า 0.99 ดังนั้น

$$0.99 = R_i^{1000}$$

ถ้าใส่ลอการิทึมทั้งสองข้างของสมการจะได้

$$\ln 0.99 = 1000 \ln R_i$$

ดังนั้น $R_i = 0.99998995$

หมายความว่าชิ้นส่วนแต่ละชิ้นจะต้องได้รับการออกแบบให้มีค่าความเชื่อถือได้สูงถึง 0.99998995 ทีเดียว

ระบบขนาน ค่าความเชื่อถือได้ของระบบสามารถเพิ่มได้ด้วยการประกอบชิ้นส่วนในลักษณะขนาน ระบบที่มีชิ้นส่วนขนานหรือสามารถทดแทนกันได้ จะไม่สามารถทำงานได้ก็ต่อเมื่อทุกชิ้นส่วนไม่สามารถทำงานได้ ถ้าให้ค่าความน่าจะเป็นที่ชิ้นส่วนจะขัดข้องเป็น F_i ดังนั้น

$$F_i = 1 - R_i$$

ค่าความเชื่อถือได้ของระบบขนานคือ

$$R_P = 1 - \pi_{i=1}^{n} F_i$$

หรือ $$R_P = 1 - \pi_{i=1}^{n} (1 - R_i) \qquad \text{.......... (10)}$$

ตัวอย่างที่ 2.4.3 **พิจารณาระบบที่มีส่วนประกอบตัวอย่างขนาน 3 ชิ้น แต่ละส่วนประกอบมีความเชื่อถือได้ 0.65 ดังนั้นความเชื่อถือได้ของระบบ คือ**

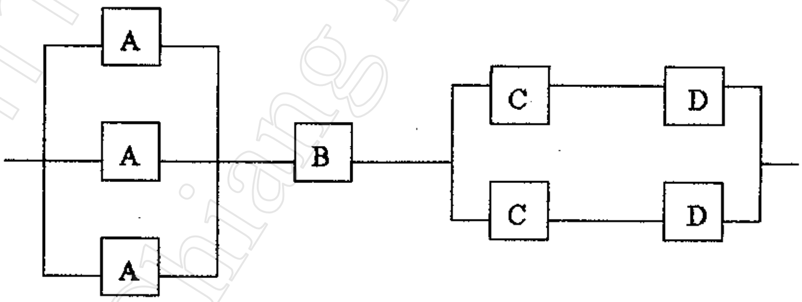
$$R_P = 1 - (1 - 0.65)^3$$
$$= 1 - 0.35^3$$
$$= 0.957125$$

สำหรับระบบที่ประกอบด้วยชิ้นส่วนที่ต่อกันแบบผสม การคำนวณความเชื่อถือได้สามารถทำได้โดยคำนวณค่าความเชื่อถือได้โดยแบ่งเป็นระบบย่อย ๆ แล้วจึงรวมเป็นระบบใหญ่

ตัวอย่างที่ 2.4.4 **พิจารณาระบบที่มีองค์ประกอบต่าง ๆ ดังแสดงในรูป องค์ประกอบต่าง ๆ แบ่งเป็นชิ้นส่วนย่อยซึ่งแต่ละชิ้นส่วนมีค่าความเชื่อถือได้คือ**

$R_A = 0.72 \qquad R_B = 0.96$

$R_C = 0.87 \qquad R_D = 0.91$



ขั้นแรกคำนวณค่าความเชื่อถือได้ของส่วน A จะได้

$$R_{PA} = 1 - (1 - 0.72)^3$$
$$= 1 - (0.28)^3$$
$$= 0.978$$

เนื่องจากชิ้นส่วน C และ D ต่อกันแบบอนุกรม ดังนั้น

$$R_{SCD} = (0.87)(0.91)$$
$$= 0.7917$$

แต่เนื่องจากชิ้นส่วน C และ D ถูกนำมาต่อแบบขนานกันอีกดังนั้น

$$R_{PCD} = 1 - (1 - 0.7917)^2$$

$$= 0.9566$$

ความเชื่อถือของระบบเกิดจากอนุกรม R_{PA} , R_B และ R_{PCD} เข้าด้วยกันจะได้

$$R_S = (0.978)(0.96)(0.9566)$$

$$= 0.898$$

2.5 การแจกแจงแบบปัวซอง (Poisson Distribution)

การแจกแจงแบบปัวซองเป็นเรื่องของการศึกษาถึงสถานการณ์ของการเกิดปรากฏการณ์ที่อุบัติขึ้นโดยสุ่ม (random occurrence) ส่วนใหญ่จะเป็นอุบัติการณ์ในช่วงของเวลา (time interval) มากกว่าอย่างอื่น เช่น ปรากฏการณ์ที่เกิดขึ้นภายในช่วงเวลาใดเวลาหนึ่ง (fixed time interval) แต่ช่วงหรือขอบเขตของการเฝ้าติดตามนับ (count) จำนวนการปรากฏการณ์ที่อุบัติขึ้นนี้มิได้หมายถึงเฉพาะช่วงหรือขอบเขตของเวลาเท่านั้น หากรวมไปถึงการเฝ้าติดตามนับจำนวนปรากฏการณ์ที่อุบัติขึ้นในขอบเขตของพื้นที่ ขอบเขตของปริมาตร ระยะทาง จุดนัดพบ หรืออื่น ๆ อีกด้วย เช่น จำนวนครั้งของโทรศัพท์ที่หมุนเข้ามายังโอเปอเรเตอร์ในช่วงเวลาหนึ่ง (โดยเฉพาะช่วงเวลาที่มีผู้ใช้โทรศัพท์มาก)

จำนวนอิเล็กตรอนที่พุ่งออกมาจากขั้วลบในหลอดสุญญากาศ , จำนวนดาวในบริเวณหนึ่ง (ปริมาตร) ของทางช้างเผือก, จำนวนเซลล์เม็ดโลหิตที่สามารถเห็นได้จากกล้องจุลทัศน์ (จำนวนที่มองเห็นได้หมายถึงจำนวนที่ปรากฏในพื้นที่หน่วยหนึ่งที่กว้างขวางเท่ากับความกว้างยาวของเลนส์) จำนวนอนุภาค α ที่พุ่งออกมาจากแหล่งสารกัมมันตรังสีในช่วงเวลาหนึ่ง จำนวนตำหนิ “ตามด” บนแผ่นไม้อัดขนาดพื้นที่หนึ่ง จำนวนปุ่มขรุขระบนเส้นลวดขนาดความยาวหนึ่ง จำนวนอุบัติเหตุบนท้องถนนสายหนึ่งในสัปดาห์หนึ่ง จำนวนอุกกาบาต และสะเก็ดดาวที่หล่นลงมาเกาะบนผิวของดาวเทียมที่โคจรครบ 1 รอบ จำนวนจุลินทรีย์ในน้ำปริมาตรหนึ่ง จำนวนรอยร้าวบนผิวของแผ่นกระเบื้องขนาดพื้นที่หนึ่ง จำนวนยวดยานที่วิ่งผ่านจุดสังเกตการณ์หนึ่ง จำนวนยวดยานที่จอดคอยสัญญาณไฟเขียว ณ สี่แยกแห่งหนึ่ง จำนวนพายุดีเปรสชันที่พัดผ่านตำบลหนึ่งในแต่ละปี จำนวนบุคคลที่เข้ามาติดต่อสมัครงาน ณ สำนักจัดหางานในแต่ละวัน จำนวนเรือที่เข้าเทียบท่าในแต่ละวัน ฯลฯ

จากตัวอย่างสถานการณ์ข้างต้น คงพอเข้าใจได้ว่า สถานการณ์ที่พึงสังเคราะห์เข้ากับการแจกแจงแบบปัวซองนั้น

1) ต้องเป็นสถานการณ์ที่สามารถนับจำนวนปรากฏการณ์ได้ (counting event)

2) อัตราเฉลี่ยของการเกิดปรากฏการณ์ (average rate of occurrence) จะต้องคงที่ และ

3) จะต้องกำหนดช่วงหรือเขตของการศึกษาให้แน่นอน อาจเป็นช่วงเวลา ระยะทาง จุด พื้นที่ ปริมาตรที่แน่นอน หรืออื่นใดก็ได้

การศึกษาเพื่อพัฒนาฟังก์ชันของความน่าจะเป็นของตัวแปรสุ่มปัวซองจำเป็นจะต้องมีข้อตกลงดังต่อไปนี้ และเพื่อความสะดวกจะขอเริ่มศึกษาด้วยขอบเขตของ "ช่วงเวลา" ส่วนช่วงหรือเขตอื่นก็สามารถสรุปผลได้โดยนัยเดียวกัน

1. จำนวนปรากฏการณ์ที่อุบัติขึ้นในช่วงเวลาคนละช่วง (nonoverlapping time interval) เป็นตัวแปรสุ่มที่เป็นอิสระต่อกัน

2. ความน่าจะเป็นที่จะเกิดปรากฏการณ์เพียง 1 ครั้ง ในช่วงเวลา Δt สั้น ๆ มีค่าเท่ากับ $m(\Delta t)$ โดยที่ m เป็นจำนวนเต็มบวกใด ๆ อีกนัยหนึ่ง ความน่าจะเป็นที่จะเกิดปรากฏการณ์เพียง 1 ครั้ง ในช่วงเวลา Δt สั้น ๆ มีค่าเป็นสัดส่วนกับความยาวของช่วงเวลา นั่นคือ $P_1(\Delta t) = m(\Delta t)$

เมื่อ m คืออัตราเฉลี่ยของการเกิดปรากฏการณ์ (mean rate of occurrence) หมายถึงอัตราเฉลี่ยที่แสดงถึงจำนวนปรากฏการณ์ที่อุบัติขึ้นใน 1 หน่วยของเวลา (ช่วงเวลา) เช่นเรือเข้าเทียบท่าเฉลี่ยวันละ $m = 5$ ลำ อุบัติเหตุรถยนต์ชนกันในชั่วโมงคับขัน ชั่วโมงละ $m = 7$ คัน

$P_1(\Delta t)$ คือ ความน่าจะเป็นที่จะเกิดปรากฏการณ์ 1 ครั้งในช่วงระยะเวลา Δt มีค่าเท่ากับ $m(\Delta t)$

3. ความน่าจะเป็นที่จะเกิดปรากฏการณ์เกินกว่า 1 ครั้งในช่วงเวลา Δt สั้น ๆ มีค่าน้อยมาก (เมื่อเทียบกับ $m(\Delta t)$) จนเราสามารถตัดออกจากการพิจารณาได้ นั่นคือ $\sum_{k=2}^{n} P_k(\Delta t) \cong 0$

4. ความน่าจะเป็นที่จะไม่มีปรากฏการณ์ใดเกิดขึ้นเมื่อเริ่มต้นศึกษามีค่าเท่ากับ 1 นั่นคือ $P_0(0) = 1$

เราจะนำข้อตกลงเหล่านี้ไปใช้พัฒนาทฤษฎีของการแจกแจงแบบปัวซองดังต่อไปนี้

ทฤษฎี 2.5.1 ถ้า "ข้อตกลง" ทั้ง 4 ประการข้างต้นเป็นจริงจำนวนปรากฏการณ์ที่อุบัติขึ้นในช่วงเวลา t จะมีการแจกแจงแบบปัวซอง มีพารามิเตอร์เป็น $\lambda = mt$

$\lambda = mt$ แสดงว่า λ คืออัตราถัวเฉลี่ยของการเกิดปรากฏการณ์ในช่วงเวลายาว t หน่วยทั้งนี้เพราะ m คืออัตราถัวเฉลี่ยของการเกิดปรากฏการณ์ในช่วงเวลาสั้น ๆ (Δt) เช่น จำนวนอนุภาค α ที่พุ่งจากที่เก็บวินาทีละ 3 อนุภาค แสดงว่า $m = 3$ ถ้าเราทำการบันทึกข้อมูลครั้งละ 10 วินาที แสดงว่า $t = 10$ ดังนั้น $\lambda = mt = 30 = v$ อัตราถัวเฉลี่ยของจำนวนอนุภาค α ที่พุ่งออกจากที่เก็บในทุก ๆ 10 วินาที

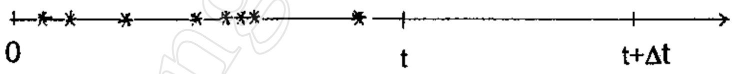
นั่นคือ ถ้าให้ตัวแปรสุ่ม X แทนจำนวนปรากฏการณ์ที่อุบัติขึ้นในช่วงเวลา t แล้ว

$$P(X = x) = f_x(x) = \frac{e^{-mt}(mt)^x}{x!} \quad ; \quad x = 0, 1, 2, \ldots$$

พิสูจน์ วิธีที่ 1

ให้ t เป็นกำหนดเวลาหลังเวลาเริ่มต้นที่ทำให้ช่วงเวลา (0, t] มีความยาว t หน่วย และช่วงเวลา $(t, t + \Delta t)$ มีความยาว Δt หน่วย ดังภาพ

ให้ $P_n(t)$ เป็นฟังก์ชันแทนความน่าจะเป็นที่จะมีปรากฏการณ์ที่สนใจอุบัติขึ้น n ครั้ง ในช่วงเวลาที่มีความยาว t หน่วย



เครื่องหมาย * แสดงจำนวนปรากฏการณ์ที่อุบัติในช่วงแห่งเวลาที่มีความยาว t หน่วย

การพิสูจน์จะดำเนินการหาความน่าจะเป็นที่จะเกิดปรากฏการณ์ในช่วงเวลา $(0, t + \Delta t)$ ตั้งแต่ 0, 1, 2, ... ครั้งเรื่อยไปตามลำดับดังนี้

$P_0(t+\Delta t)$ = P{ไม่มีปรากฏการณ์ใดอุบัติขึ้นในช่วงเวลา $(0, t+\Delta t]$}

= P{ไม่มีปรากฏการณ์ใดอุบัติอุบัติขึ้นในช่วงเวลา (0, t] หรือไม่มีปรากฏการณ์อุบัติขึ้นในช่วงเวลา $(t, t + \Delta t]$}

= P{ไม่มีปรากฏการณ์ใดอุบัติขึ้นในช่วงเวลา (0, t]} + P{ไม่มีปรากฏการณ์ใดอุบัติขึ้นในช่วงเวลา $(t, t + \Delta t]$}

นั่นคือ $P_0(t+\Delta t) = P_0(t)P_0(\Delta t)$ เนื่องจากข้อตกลงที่ 1

แต่ $P_0(t+\Delta t) = 1 - P($มีปรากฏการณ์อุบัติขึ้นในช่วงเวลา $(t, t+\Delta t])$ อย่างน้อย 2 ครั้ง)

$= 1 - \{P($มีปรากฏการณ์อุบัติขึ้น 1 ครั้งในช่วงเวลา $(t, t+\Delta t]) +$ P (มีปรากฏการณ์อุบัติขึ้นตั้งแต่ 2 ครั้งขึ้นไปในช่วงเวลา $(t, t+\Delta t])\}$

$$\cong 1 - P_1(\Delta t) - \sum_{k=2}^{\infty} P_k(\Delta t)$$

นั่นคือ $P_0(\Delta t) \cong 1 - m(\Delta t)$ เนื่องจากข้อตกลงที่ 2 และที่ 3

ดังนั้น $P_0(t+\Delta t) \cong P_0(t)(1 - m\Delta t)$

$$\cong P_0(t) - m\, P_0(t)\, \Delta t$$

$$\frac{P_0(t+\Delta t) - P_0(t)}{\Delta t} \cong -mP_0(t)$$

แต่ $$\lim_{\Delta x \to 0} \frac{f(x+\Delta x) + f(x)}{\Delta x} = f'(x)$$

ดังนั้น $$\lim_{\Delta t \to 0} \frac{P_0(t+\Delta t) + P_0(t)}{\Delta t} = P_0'(t)$$

$$P_0'(t) = -mP_0(t)$$

$$\frac{P_0'(t)}{P_0(t)} = -m$$

$$\frac{d}{dt}\ln P_0(t) = -m$$ เนื่องจาก $$\frac{d}{dx}\ln f(x) = \frac{1}{f(x)}\frac{d}{dx}f(x) = \frac{f'(x)}{f(x)}$$

อินทิเกรตทั้งสองด้านเทียบ t จะได้

$$\int \frac{d}{dt}\ln P_0(t) = \int -m dt$$

$$\ln P_0(t) = -mt + c$$

ถ้าให้ $t = 0$ จะได้ $c = 0$ เนื่องจากข้อตกลงที่ 4 ; $P_0(0) = 1$, $\ln 1 = 0, 0 = 0$

นั่นคือ $\ln P_0(t) = -mt$

ดังนั้น $P_0(t) = e^{-mt}$ หรืออีกนัยหนึ่งความน่าจะเป็นที่จะไม่เกิดปรากฏการณ์

ใดในช่วงเวลายาว t หน่วย เมื่ออัตราเฉลี่ยของการเกิดปรากฏการณ์ใด ๆ ในช่วงเวลาความยาวนี้ (t หน่วย) มีค่าเท่ากับ mt

พิจารณา

$P_1(t+\Delta t)$ = P(เกิดปรากฏการณ์ 1 ครั้งในช่วงเวลา $(0, t + \Delta t]$)

= P{เกิดปรากฏการณ์ 1 ครั้งในช่วงเวลา $(0, t]$ แต่ไม่เกิดขึ้นในช่วงเวลา $(t, t+\Delta t]$ หรือไม่เกิดปรากฏการณ์ใดขึ้นในช่วงเวลา $(0, t]$ แต่ไปเกิดขึ้น 1 ครั้งในช่วงเวลา $(t, t+\Delta t]$}

$= P_1(t)P_0(\Delta t) + P_0(t)P_1(\Delta t)$

$= P_1(t)(1 - m(\Delta t)) + P_0(t)\, m\Delta t$ เนื่องจากข้อตกลงที่ (2)

$P_1(\Delta t) = m(\Delta t)$ และสมการ (1)

$= P_1(t) + m\Delta t(-P_1(t) + P_0(t))$

$$P_1(t + \Delta t) = P_1(t) + m\Delta t(-P_1(t) + P_0(t))$$

$$\frac{P_1(t+\Delta t) - P_1(t)}{\Delta t} = m(-P_1(t) + P_0(t))$$

ดังนั้น $$\lim_{\Delta t \to 0} \frac{P_1(t+\Delta t) - P_1(t)}{\Delta t} = -mP_1(t) + mP_0(t))$$

$$P_1'(t) = -mP_1(t) + mP_0(t)$$

$$P_1'(t) = -mP_1(t) + me^{-mt}$$: เนื่องจาก $P_0(t) = e^{-mt}$

จาก differential equation of order 1 degree 1

$$\frac{du}{dx} + R(x)u = S(x)$$; เมื่อ $y = u(x)$ $R(x)$ และ $S(x)$ เป็นฟังก์ชันของ x

($R(x)$ และ $S(x)$ เป็นเทอมคงที่ได้) จะได้คำตอบทั่วไปดังนี้คือ

$$ue^{\int R(x)dx} = \int s(x)e^{\int R(x)dx}dx + c$$

ดังนั้น

จากสมการ $$P_1'(t) = -mP_1(t) + me^{-mt}$$

$$\frac{d}{dt}P_1(t) + mP_1(t) = me^{-mt}$$

$$P_1(t)e^{\int -mtdt} = \int me^{\int -mt}, e^{\int mdt}de + c$$ ในที่นี้ $u = P_1(t)$, $R = m$, $S = me^{-mt}$

$$P_1(t)e^{mt} = mt + c$$

นั่นคือ $$P_1(t) = (mt)e^{-mt}$$: $c = 0$ ตามข้อตกลงที่ (4)

ในทำนองเดียวกัน (โดยการพิสูจน์ติดต่อกันมาเป็นลำดับ) จะได้

$$P_2(t) = \frac{(mt)^2 e^{-mt}}{2!}, P_3(t) = \frac{(mt)^3 e^{-mt}}{3!} \ldots, P_n(t) = \frac{(mt)^n e^{-mt}}{3n!} \ldots$$

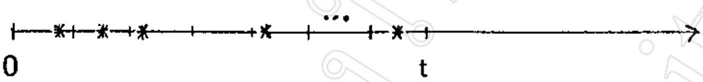
นั่นคือ $$P_x(t) = \frac{(mt)^x e^{-mt}}{x!}; X = 0, 1, 2, \ldots$$

ให้ $$\lambda = mt$$

$$P_x(t) = \frac{\lambda^x e^{-\lambda}}{X!}; X = 0, 1, 2, \ldots$$

นั่นคือ $$f_x(x) = \frac{e^{-\lambda}\lambda^x}{X!}; X = 0, 1, 2, \ldots$$

พิสูจน์ วิธีที่ 2 แบ่งช่วงเวลา $(0, t)$ ออกเป็นช่วงเวลาย่อย ๆ n ช่วงขนาดความยาวช่วงละ $\Delta t = t/n$ ดังภาพ



ดังนั้นความน่าจะเป็นที่จะปรากฏการณ์ใด ๆ อุบัติขึ้น x ครั้งในช่วงเวลา $(0, t)$ จึงมีเท่ากับความน่าจะเป็นที่จะเกิดปรากฏการณ์ขึ้นในแต่ละช่วงเวลาย่อย x ช่วง ๆ ละ 1 ครั้ง

แต่ในแต่ละช่วงเวลาย่อย ๆ หนึ่งอาจมีปรากฏการณ์อุบัติขึ้นหรือไม่ก็ได้ (อุบัติ-ไม่อุบัติ) แสดงว่าการเกิดปรากฏการณ์ขึ้นในช่วงเวลาย่อยหนึ่ง ๆ มีการแจกแจงแบบเบอร์นูลลิ และเนื่องจากการเกิดปรากฏการณ์ในแต่ละช่วงเวลาย่อยนั้นเป็นอิสระต่อกัน (ข้อตกลงที่ 1) ดังนั้น จำนวนปรากฏการณ์ที่จะอุบัติขึ้น x ครั้ง $(x = 0, 1, 2, \ldots)$ ในช่วงเวลาย่อย n ช่วง จึงมีการแจกแจงแบบทวินาม

เนื่องจากค่าความน่าจะเป็นที่จะมีปรากฏการณ์ใดอุบัติขึ้นในแต่ละช่วงเวลาย่อยมีค่าเท่ากับ $m\Delta t$ (ข้อตกลงที่ 2)

ดังนั้น P (มีปรากฏการณ์อุบัติขึ้น X ครั้งใน n ช่วงเวลาย่อย)

$$= \binom{n}{x}\left(\frac{mt}{n}\right)^x\left(1-\frac{mt}{n}\right)^{n-x} \text{ เมื่อ } \Delta t = \frac{t}{n}$$

$$= \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}$$

$$= \frac{n!}{x!(n-x)!}\,\frac{\lambda^x}{n^x}\cdot\frac{(1-\lambda/n)^n}{(1-\lambda/n)^x}$$

$$= \frac{n(n-1)(n-2)\ldots(n-x+1)(n-x)!}{n.n.n\ldots n(n-x)!}\,\frac{\lambda^x}{x!}\,\frac{(1-\lambda/n)^n}{(1-\lambda/n)^x}$$

$$= 1(1-1/n)(1-2/n)\cdots\left(1-\frac{x-1}{n}\right)\frac{\lambda^x}{x!}\,\frac{(1-\lambda/n)^n}{(1-\lambda/n)^x}$$

ถ้าแบ่งช่วงเวลา (0, t) ให้ถี่ยิ่งขึ้น กล่าวคือ $n \to \infty$

ดังนั้น $\lim_{n\to\infty} 1-\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{x-1}{n}\right)\frac{\lambda^n}{x!}\frac{(1-\lambda/n)^n}{(1-\lambda/n)^x} \;=\; \frac{\lambda}{x!}e^{-\lambda}$

กล่าวคือ เมื่อ $n \to \infty$ จะได้ $\left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}$ และ $\left(1-\frac{\lambda}{n}\right) \to 1$

ดังนั้น

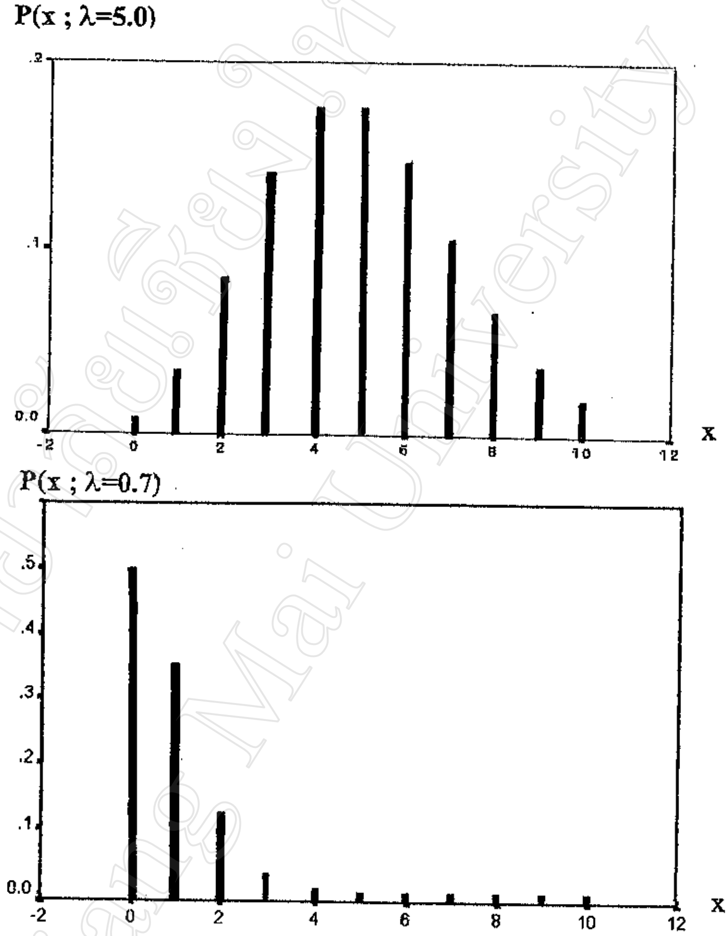
P (เกิดปรากฏการณ์ x ครั้งในช่วงเวลา (0, t) ที่แบ่งเป็นช่วงเวลาย่อย ๆ n ช่วง) = $\frac{e^{-\lambda}\lambda^x}{x!}$

$$P_x(t) \;=\; f_x(x) \;=\; \frac{e^{-\lambda}\lambda^x}{x!} \quad ; \; x = 0, 1, 2, \ldots$$

หมายเหตุ กระบวนการศึกษาถึงการแจกแจงของตัวแปรสุ่มใด ๆ ที่แสดงจำนวนครั้งของปรากฏการณ์ที่อุบัติขึ้นในช่วงเวลา (ระยะทาง พื้นที่ ปริมาตร จุด ฯลฯ) ที่แน่นอน โดยยึดถือข้อตกลงทั้ง 4 ประการดังกล่าวข้างต้น และผลของการศึกษาพบว่าการ แจกแจงของตัวแปรสุ่มนั้น มีการแจกแจงแบบปัวซองเราเรียกกระบวนการศึกษานี้ว่า "กระบวนการปัวซอง" (poisson process)

อนึ่ง การจะให้ข้อยุติลงไปว่าสถานการณ์หนึ่งใดจะมีการแจกแจงแบบปัวซอง (ทั้ง ๆ ที่เป็นเรื่องของการนับเหมือนกัน) นั้นเป็นเรื่องที่ควรระมัดระวังเพราะอาจผิดพลาดได้ง่าย ขอให้คำนึงถึงข้อตกลงทั้ง 4 ประการนั้นไว้เสมอทุกครั้งที่จะมีการลงยุติ อย่างเช่น การศึกษาการแจกแจงของจำนวนไข่ของแมลงที่ปรากฏอยู่ตามพื้นที่เพาะปลูก เรื่องนี้แม้จะเป็นการศึกษาถึงการนับจำนวนไข่แมลงที่นับในฤดูกาลหนึ่ง ๆ (ก่อนฤดูเพาะปลูก) ซึ่งน่าจะเป็นการแจกแจงแบบปัวซองได้เพราะกำหนดเวลาของการศึกษาจำกัดแน่นอนและเป็นเรื่องของการนับ (counting event) แต่มิอาจสรุปว่าเป็นการแจกแจงแบบปัวซองได้ เพราะโดยธรรมชาติแมลงจะลงวางไข่เป็นกระจุกไข่และตัวอ่อนจะอยู่รวมกันเป็นกระจุก (cluster) ซึ่งขัดกับข้อตกลงที่ 1 ข้างต้น

ภาพการแจกแจงแบบปัวซองปรากฏดังนี้



รูปที่ 2.2 แสดงแผนภาพการแจกแจงปัวซอง เมื่อ $\lambda = 5.0, 0.7$

ทฤษฎี 2.5.2 ถ้าตัวแปรสุ่ม X มีการแจกแจงแบบปัวซองมีพารามิเตอร์เป็น $\lambda = mt$ แล้วตัวแปร X จะมี mgf ค่าคาดหมายและความแปรปรวนดังนี้

$$M_x(t) = e^{\lambda(e^t-1)}, E(X) = V(X) = \lambda$$

พิสูจน์

$$f_x(x) = \frac{e^{-\lambda}\lambda^x}{X!} \; ; \; X = 0, 1, 2, \ldots$$

$$M_x(t) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!}$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

โดยอาศัยการกระจายเทเลอร์

$$e^{u} = 1+u+\frac{u^2}{2!}+\frac{u^3}{3!}+\cdots = \sum_{x=0}^{\infty}\frac{u^k}{k!}$$

จะได้ $M_x(t) = e^{-\lambda}e^{\lambda e^t} = e^{\lambda(e^t - t)}$

$$E(X) = M_x'(t)\big|_{t=0}$$

$$= \lambda e^t e^{\lambda(e^t-1)}\big|_{t=0}$$

$$= \lambda$$

$$V(X) = M_x''(t)\big|_{t=0} - \lambda^2$$

$$M_x''(t) = \lambda^2 e^{2t} e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}$$

$$M_r''(t)\big|_{t=0} = \lambda^2 + \lambda$$

ดังนั้น $V(X) = \lambda$

นั่นคือ เมื่อตัวแปรสุ่ม X มีการแจกแจงแบบปัวซองแล้ว $M_x(t) = e^{\lambda(e^t-1)}$ และ $E(X) = V(X) = \lambda$

ทฤษฎี 2.5.3 เมื่อ $(X_1, X_2, \ldots, X_n)$ เป็นกลุ่มตัวอย่างที่สุ่มมาจากกลุ่มประชากรที่มีการแจกแจงแบบปัวซองที่มีพารามิเตอร์ λ แล้ว ตัวสถิติ $Y = \sum_{i}^{n} X_i$ จะมีการแจกแจงแบบปัวซองที่มีพารามิเตอร์เป็น $n\lambda$

พิสูจน์ จาก $f_{x_i}(x) = \dfrac{e^{-\lambda}\lambda^x}{x!}$; $x = 0, 1, 2, \ldots$, $i = 1, 2, \ldots, n$

$$M_{x_i}(t) = e^{\lambda(e^t-1)} \; ; \; i = 1, 2, \ldots, n$$

$$M_y(t) = \prod_{i}^{n} M_{x_i}(t) = e^{n\lambda(e^t-1)}$$

$$f_y(y) = \frac{e^{n\lambda}(n\lambda)^y}{y!} \quad ; \ y = 0, 1, 2, \ldots$$

$$E(Y) = V(Y) = n\lambda$$

ทฤษฎี 2.5.4 เมื่อ $X_1, X_2, \ldots, X_r$ เป็นตัวแปรสุ่มที่มีการแจกแจงแบบปัวซองที่มีพารามิเตอร์เป็น $\lambda_1, \lambda_2, \ldots, \lambda_r$ ตามลำดับแล้ว ตัวแปรสุ่ม $Y = \sum_i^r X_i$ จะมีการแจกแจงแบบปัวซองที่มีพารามิเตอร์ $\sum_i^r \lambda_i$

ข้อสังเกต

1. กรณีที่ตัวแปรสุ่ม X มีการแจกแจงแบบปัวซองมีพารามิเตอร์เป็น λ แล้ว จะมี mgf ค่าคาดหมายและความแปรปรวนเป็น $M_x(t) = e^{\lambda(e^t-1)}$ และ $E(X) = V(X) = \lambda$ นั้น ข้อนี้ก็มีลักษณะของการสังเกตเพื่อหาการแจกแจงค่าคาดหมาย และความแปรปรวนได้เช่นเดียวกับตัวแปรสุ่มแบบอื่น ๆ คือ เราสามารถทราบการแจกแจงค่าคาดหมายและค่าความแปรปรวนได้จาก mgf กล่าวคือ ตัวแปรสุ่มที่มี mgf เป็น $e^{\lambda(e^t-1)}$ เราสามารถทราบได้ทันทีว่าตัวแปรสุ่มนั้นมีการแจกแจงแบบปัวซองมีค่าคาดหมายเท่ากับ λ และความแปรปรวนเท่ากับ λ เช่น

$$M_x(t) = e^{0.1(e^t-1)} \text{ ก็แสดงว่า } f_x(x) = \frac{e^{-0.1}(0.1)^x}{x!} \quad ; \ x = 0, 1, 2, \ldots$$

2. ทฤษฎี 2.5.2 เป็นเรื่องของการแจกแจงของตัวสถิติ หลักการของทฤษฎีนี้จะถูกนำไปใช้ในทางสร้างตัวทดสอบสำหรับตรวจสอบนัยสำคัญของพารามิเตอร์ λ และการประมาณค่าพารามิเตอร์ λ

3. ทฤษฎี 2.5.3 เป็นกรณีทั่ว ๆ ไปของทฤษฎี 2.5.2 ทฤษฎีนี้เรียกว่า reproductive property ของตัวแปรสุ่มปัวซองเป็นการยืนยันให้เห็นว่าตัวแปรสุ่มแบบปัวซองที่มีพารามิเตอร์ต่างกัน (หรือเหมือนกัน ถ้ามีพารามิเตอร์เหมือนกันเรามักหมายถึงกรณีสุ่มตัวอย่างมาจากกลุ่มประชากรที่มีพารามิเตอร์นั้น) นั้น สามารถรวมตัวกันได้เป็นตัวแปรสุ่มแบบปัวซองเช่นเดิม

ตัวอย่าง 2.5.1 โดยเฉลี่ยตัวทรานซิสเตอร์ในเครื่องคอมพิวเตอร์จะเสื่อมคุณภาพ 1 ตัว ในทุก 100 ชั่วโมงทำงาน และเมื่อใดก็ตามที่มีตัวทรานซิสเตอร์เสื่อมคุณภาพตั้งแต่ 3 ตัวขึ้นไปเครื่องจะหยุดทำงาน

ก. โปรแกรมหนึ่งใช้เวลาเครื่อง (computer time) 20 ชั่วโมง จงหาความน่าจะเป็นเครื่องยังทำงานได้ดีในช่วงเวลานี้

ข. ถ้าเครื่องหยุดทำงานทันทีที่ตัวทรานซิสเตอร์เสื่อมคุณภาพ 2 ตัว จงคำนวณหาค่าความน่าจะเป็นในข้อ ก.

วิธีทำ ให้ X = จำนวนตัวทรานซิสเตอร์ที่เสื่อมคุณภาพภายในเครื่องคอมพิวเตอร์ เนื่องจากในทุก 100 ชั่วโมงทำงานโดยเฉลี่ยจะมีตัวทรานซิสเตอร์เสื่อมคุณภาพ 1 ตัว

ดังนั้น จะมีตัวทรานซิสเตอร์เสื่อมคุณภาพ $= \frac{1}{100} = 0.01$ ตัวในทุก 1 ชั่วโมง จะได้ $m = .01$

ก. โปรแกรมหนึ่งใช้เวลาเครื่อง 20 ชั่วโมง

ดังนั้นจะมีตัวทรานซิสเตอร์เสื่อมคุณภาพในช่วงเวลา 20 ชั่วโมง $= mt = (0.01)(20) = 0.02$ ตัว

P(เครื่องจักรยังคงทำงานได้ดีในช่วงเวลา 20 ชั่วโมงนี้)

$$= \sum_{x=0}^{2} \frac{e^{-0.2}(0.2)^x}{x!}$$

$$= 1 - \sum_{x=3}^{\infty} \frac{e^{-0.2}(0.2)^x}{x!}$$

$$= 1 - 0.0011485$$

$$= 0.9988515$$

ข. ถ้าเครื่องหยุดทำงานทันทีที่มีตัวทรานซิสเตอร์เสื่อมคุณภาพ 2 ตัว

P(เครื่องยังคงทำงานได้ดีในช่วงเวลา 20 ชั่วโมงนี้)

$$= \sum_{x=0}^{1} \frac{e^{-0.2}(0.2)^x}{x!}$$

$$= 1 - \sum_{x=2}^{\infty} \frac{e^{-0.2}(0.2)^x}{x!}$$

$$= 1 - 0.0175231$$

$$= 0.9824769$$

ตัวอย่าง 2.5.2 บริษัทจำหน่ายอุปกรณ์และเครื่องมือทำสถิติเกี่ยวกับการจำหน่ายเครื่องมือผ่าตัดเอาไว้และพบว่า โดยเฉลี่ยจะจำหน่ายเครื่องมือผ่าตัดได้เพียงเดือนละ 2 เครื่อง โดยปกติการสั่งซื้อเครื่องมือประเภทนี้จะต้องสั่งล่วงหน้า 1 ปี เพราะผู้ผลิตไม่ปรารถนาจะผลิตมากนัก เพราะเกรงว่าจะล้าสมัย ขณะเดียวกันบริษัทผู้จำหน่ายก็ไม่ปรารถนาจะสั่งสินค้ามาคงคลังไว้มาก เพราะต้องเสียค่าใช้จ่ายเรื่องการดูแลรักษาและทำให้เสียค่าใช้จ่ายโดยไม่จำเป็น อยากทราบว่าบริษัทผู้จำหน่ายควรสั่งสินค้าประเภทนี้เข้าคงคลังไว้กี่เครื่องจึงจะมั่นใจได้ถึง 99% เป็นอย่างน้อยว่าสินค้าดังกล่าวจะมีไว้เพียงพอและเหมาะสมกับความต้องการของลูกค้าในช่วงเวลา 1 ปี

วิธีทำ ให้ X = ปริมาณความต้องการเครื่องมือผ่าตัดในช่วงเวลา 1 ปี

k = จำนวนสินค้า (เครื่องมือผ่าตัด) ที่บริษัทผู้จำหน่ายควรสั่งเข้าคงคลัง

บริษัทจำหน่ายเครื่องมือผ่าตัดได้โดยเฉลี่ยเดือนละ 2 เครื่อง

ดังนั้น $m = 2$ และ $\lambda = mt = 2(12) = 24$ เครื่องต่อปี

จาก $P(X \leq k) \quad \geq \quad 0.99$

$$\sum_{x=0}^{k} \frac{e^{24} 24^x}{x!} \quad \geq \quad 0.99$$

โดยอาศัยทฤษฎีการเข้าสู่เกณฑ์กลาง

$$k \cong 35$$

นั่นคือบริษัทตัวแทนจำหน่ายควรคงคลังเครื่องมือผ่าตัดไว้ปีละประมาณ 35 เครื่อง จึงจะเชื่อถือได้ว่าเป็นจำนวนที่พอเหมาะกับความต้องการของผู้ใช้ในแต่ละปี

ตัวอย่าง 2.5.3 การตรวจเม็ดโลหิตแดงในเลือดของมนุษย์สามารถกระทำได้ด้วยการตรวจนับจำนวนได้จากตัวอย่างเลือด (specimen) โดยอาศัยกล้องจุลทัศน์ ถ้าในเลือดปริมาณหนึ่งของคนปกติจะมีเม็ดโลหิตแดงอยู่ 20 เม็ด จงหาค่าความน่าจะเป็นที่ตัวอย่างเลือดของคนปกติผู้หนึ่งจะมีเม็ดโลหิตแดงต่ำกว่า 15 เม็ด

วิธีทำ ให้ X = จำนวนเม็ดโลหิตแดงในตัวอย่างเลือดของคนปกติ

$\lambda = 20 =$ จำนวนเม็ดโลหิตแดง (ถัวเฉลี่ย) ในตัวอย่างเลือดปริมาณหนึ่ง

$$P(X<15) = \sum_{x=0}^{14} \frac{e^{-20}20^x}{x!} = 0.10486$$

นั่นคือ ความน่าจะเป็นที่จะปรากฏเม็ดโลหิตแดงในตัวอย่างเลือดของคนปกติต่ำกว่า 15 เม็ด มีค่าเท่ากับ 0.10486 หรือโอกาสที่คนปกติจะมีเม็ดโลหิตแดงต่ำกว่า 15 เม็ด ในปริมาณเลือด ตัวอย่างปริมาตรหนึ่ง ซึ่งจะต้องมีเม็ดเลือดประมาณ 20 เม็ด นั้น เป็นไปได้เพียง 10.48% เท่านั้น

2.6 การแจกแจงแบบเอ็กซ์โปเนนเชียล (Exponential Distribution)

การแจกแจงแบบเอ็กซ์โปเนนเชียลเป็นการแจกแจงของตัวแปรสุ่มที่แสดงระยะเวลาระหว่างเหตุการณ์ (time between event) เมื่อสิ่งที่มุ่งศึกษานั้นหมายถึง "เหตุการณ์" หรือแสดงอายุการใช้งานของวัตถุสิ่งของเมื่อสิ่งที่มุ่งศึกษาหมายถึง "อายุ"

คำว่า "เหตุการณ์" ที่กล่าวถึงในตอนต้นอาจหมายถึงอะไรก็ได้ เช่น ระยะเวลาการพูดโทรศัพท์ ระยะเวลาที่รถคอยผ่านสัญญาณไฟแดง ระยะเวลาที่รถบรรทุกเสียไปในด่านชั่งน้ำหนัก ระยะเวลาทิ้งห่างระหว่างรถแต่ละคันที่วิ่งผ่านจุดใดจุดหนึ่ง (ระยะเวลาทิ้งห่างระหว่างรถแต่ละคันที่ผ่านจุด ๆ หนึ่ง อาจเป็นสี่แยกไฟแดง ด่านตรวจความเร็ว หลักกิโลเมตร ฯลฯ ถ้า

ยิ่งสั้น ความหนาแน่นของการจราจรก็ยิ่งสูง ถ้าช่วงทิ้งห่างยาวขึ้น ความหนาแน่นของการจราจรก็เบาบาง) ระยะเวลาที่คนไข้แต่ละคนรอคอยเข้ารับบริการตรวจรักษาจากแพทย์ ระยะเวลาที่โมเลกุลใด ๆ ของน้ำที่จะอยู่ในถังเก็บน้ำเมื่อมีการสูบน้ำเข้าถังและจ่ายน้ำไปตามท่ออยู่ตลอดเวลา ฯลฯ เหล่านี้ล้วนเป็นตัวอย่างที่แสดงสถานการณ์ของการศึกษาเมื่อประเด็นที่มุ่งศึกษาหมายถึงเหตุการณ์ จะเห็นได้ว่า "เวลาระหว่างเหตุการณ์" หมายถึง ระยะเวลาระหว่างเหตุการณ์ทำนองเดียวกันที่เกิดขึ้นแล้วเหตุการณ์ดังกล่าวก็กำลังจะเกิดขึ้นอีก ส่วน "อายุใช้งาน" ที่กล่าวถึงหมายถึง ความทนทานของวัตถุสิ่งของโดยวัดด้วยเวลาตั้งแต่เริ่มใช้ (t = 0) จนกระทั่งวัตถุนั้นเสียหรือเสื่อมสภาพ เช่น อายุใช้งานของทรานซิสเตอร์ อายุใช้งานของเครื่องใช้ไฟฟ้าและอุปกรณ์ไฟฟ้าต่าง ๆ อายุใช้งานของวัสดุสิ้นเปลืองและวัสดุยาวนานต่าง ๆ เช่น ชิ้นส่วนอะไหล่รถยนต์ เครื่องบินและอื่น ๆ เป็นต้น

จากที่กล่าวมาทั้งหมดจึงเห็นได้ว่า ตัวแปรสุ่มที่มีการแจกแจงแบบเอ็กซ์โปเนนเชียลโดยส่วนใหญ่เป็นเรื่องของตัวแปรเวลา (time vairable) มากกว่าอย่างอื่น ๆ ที่เริ่มต้นนับเมื่อเริ่มใช้วัตถุ (กรณีอายุใช้งาน) หรือเริ่มนับเมื่อปรากฏการณ์ทำนองเดียวกันได้เกิดขึ้นเสร็จสิ้นไปแล้ว เวลาดังกล่าวจะนับติดต่อกันไปเรื่อย ๆ จนกระทั่งปรากฏการณ์ที่รอคอยนั้นอุบัติขึ้นอีก

ณ จุดนี้จึงสามารถใช้ข้อสรุปเกี่ยวกับลักษณะของการแจกแจงเอ็กซ์โปเนนเชียลได้เป็น 4 ประการ

1. การแจกแจงแบบเอ็กซ์โปเนนเชียลเป็นการศึกษาถึงระยะเวลาที่รอคอย (contineous waiting time) จนกระทั่งเกิดปรากฏการณ์ที่มุ่งศึกษาเป็นครั้งแรก

2. การแจกแจงแบบเอ็กซ์โปเนนเชียลมีความเกี่ยวพันกับการแจกแจงแบบปัวซองอย่างลึกซึ้ง กล่าวคือ การแจกแจงแบบปัวซองศึกษาถึงจำนวนครั้ง (เฝ้านับ) ของการเกิดอุบัติการณ์ที่มุ่งศึกษาในช่วงเวลา (ระยะทาง พื้นที่ ปริมาตร จุด) ที่แน่นอนช่วงหนึ่ง ส่วนการแจกแจงแบบเอ็กซ์โปเนนเชียลถือว่าเวลาเป็นตัวแปร (ไม่คงที่เช่นกรณีปัวซอง) และนับเวลาหรือบันทึกเวลาติดต่อกันไปเรื่อย ๆ จนกระทั่งเกิดอุบัติการณ์ปัวซองเป็นครั้งแรก (occurrence of poisson event) (เฝ้าคอย) อย่างไรก็ตาม เวลาอาจไม่จำเป็นต้องเริ่มที่ 0 เสมอไป ณ จุดนี้เราจึงเห็นได้ว่า การแจกแจงทั้งสองแบบมีลักษณะร่วมกันคือเฝ้านับปรากฏการณ์ที่จะอุบัติขึ้นโดยสุ่ม (random occurrence)

3. การแจกแจงแบบเอ็กซ์โปเนนเชียลเป็นอีกโฉมหน้าหนึ่งของการแจกแจงแบบอนุกรมเรขาคณิต เพียงแต่การแจกแจงแบบอนุกรมเรขาคณิตศึกษาถึงสถานการณ์ของตัวแปรสุ่มแบบไม่ต่อเนื่อง ส่วนการแจกแจงแบบเอ็กซ์โปเนนเชียลศึกษาถึงสถานการณ์ของตัวแปรสุ่มแบบต่อเนื่อง

4. ฟังก์ชันการแจกแจงของตัวแปรสุ่มเอ็กซ์โปเนนเชียลพัฒนาขึ้นมาจากความน่าจะเป็นที่จะเกิดอุบัติการณ์ปัวซองเป็นครั้งแรกดังนี้

ให้ตัวแปรสุ่ม T = ช่วงเวลาที่ต้องรอคอยจนกระทั่งเกิดอุบัติการณ์ปัวซองเป็นครั้งแรก

X = จำนวนครั้งของการเกิดปรากฏการณ์ที่อุบัติขึ้นในช่วงเวลา (0, t) ของกระบวนการปัวซอง

ดังนั้น $P(T > t)$ จึงหมายถึงความน่าจะเป็น ที่ยังไม่เกิดอุบัติการณ์ปัวซองใดในช่วงเวลายาว t หน่วย โดยที่ จะเรียก $P(T > t)$ ว่า reliability หมายถึงโอกาสที่วัตถุยังคงใช้งานต่อไปได้เมื่อใช้มาแล้วถึง t หน่วยเวลา (หรือยังไม่เสื่อมสภาพเมื่อใช้ถึง t หน่วยเวลา) เช่น หลอดไฟมีอายุใช้งานโดยเฉลี่ย 300 ชั่วโมง $P(T > 300)$ จะหมายถึงความน่าจะเป็นที่หลอดไฟจะมีอายุยืนยาวกว่าอายุเฉลี่ย (เพราะใช้มา 300 ชั่วโมงก็ยังไม่ขาด) หรือความน่าจะเป็นที่หลอดไฟจะยังคงใช้ได้เป็นปกติ เมื่อผ่านการใช้มาแล้ว 300 ชั่วโมง หรือความน่าจะเป็นที่หลอดไฟจะยังไม่ขาดในช่วงเวลา (0, 300)

เนื่องจาก $P(T > t)$ = P(ไม่เกิดอุบัติการณ์ปัวซองในช่วงเวลายาว t หน่วย)

$$= P(X = 0)$$

จะได้ $$P(T > t) = \frac{e^{-\lambda t}(\lambda t)^0}{0!} = e^{-\lambda t} \; : \; t \geq 0$$

เมื่อ λ = อัตราถัวเฉลี่ยของการเกิดอุบัติการณ์ต่อ 1 หน่วยเวลา

$$1 - P(T \leq t) = e^{-\lambda t} \qquad : t \geq 0$$

$$1 - F_T(t) = e^{-\lambda t}$$

$$F_T(t) = P(T \leq t) = 1 - e^{-\lambda t}$$

$$\frac{d}{dt}F_T(t) \quad = \quad \frac{d}{dt}(1-e^{-\lambda t})$$

$$F_T(t) \quad = \quad \lambda e^{-\lambda t} \qquad ; \ t > 0$$

นั่นคือ $F_T(t) = \lambda e^{-\lambda t}$; $t \geq 0$ คือฟังก์ชันความน่าจะเป็นที่จะเกิดอุบัติการณ์ปัวซองเป็นครั้งแรกในช่วงเวลาขนาดความยาว t หน่วยเมื่อ λ คืออัตราถัวเฉลี่ยของจำนวนอุบัติการณ์ที่จะเกิดขึ้นใน 1 หน่วยเวลา λ ในที่นี้ความหมายเช่นเดียวกันกับ λ ในการแจกแจงแบบปัวซอง เมื่อนำมาศึกษาในแง่ของเอ็กซ์โปเนนเชียล จึงหมายถึงระยะเวลาถัวเฉลี่ยระหว่างเหตุการณ์หรืออัตราการเสื่อมสภาพของวัตถุต่อหน่วยเวลา (time between event หรือ failure rate) ทั้งนี้แล้วแต่ว่าสถานการณ์ที่กำลังศึกษาอยู่นั้นเป็น "เหตุการณ์" หรือ "อายุการใช้งาน"

นิยาม 2.6.1 ตัวแปรสุ่ม X จะมีการแจกแจงแบบเอ็กซ์โปเนนเชียล ถ้า X มี pdf ดังนี้คือ

$$f_x(x) = \lambda e^{-\lambda x} \quad ; \text{ เมื่อ } x \geq 0$$
$$= 0 \qquad ; \text{ เมื่อ } x < 0$$

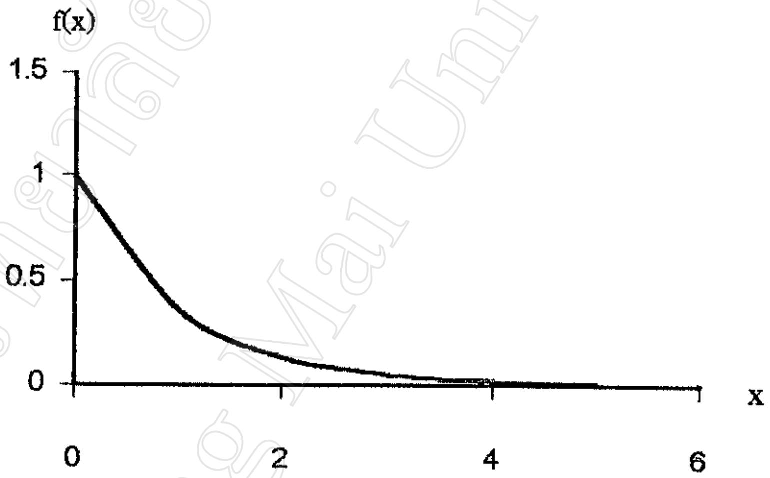
ข้อสังเกต โดยปกตินิยมใช้ T เป็นตัวแปรสุ่มเอ็กซ์โปเนนเชียลหรือตัวแปรสุ่มใดก็ตามที่เกี่ยวข้องกับเวลา แต่ในที่นี้ใช้ X เพราะเคยใช้ X ในความหมายของตัวแปรสุ่มมาโดยตลอด

นิยาม 2.6.2 ถ้าตัวแปรสุ่ม X มีการแจกแจงแบบเอ็กซ์โปเนนเชียลแล้ว X จะมี mgf ค่าคาดหมายและความแปรปรวนดังนี้

$$M_x(t) = \frac{\lambda}{\lambda - t} \quad ; \quad t < \lambda \quad , \quad E(X) = \frac{1}{\lambda} \text{ และ } V(X) = \frac{1}{\lambda^2}$$

เราจะพบว่า ค่าคาดหมายของตัวแปรสุ่มเอ็กซ์โปเนนเชียลมีค่าเท่ากับ $\frac{1}{\lambda}$ แสดงว่า E(X) มีค่าเท่ากับส่วนกลับ (inverse) ของอัตราการเสื่อมสภาพของวัตถุหรือระยะเวลาถัวเฉลี่ย เช่นอายุของหลอดสุญญากาศมีค่าเท่ากับ 100 ชั่วโมง ($\frac{1}{\lambda}$) แสดงว่าอัตราการเสื่อมสภาพ (λ) ของหลอดสุญญากาศมีค่าเท่ากับ 0.01 หน่วยต่อชั่วโมง (หรือหลอดสุญญากาศจะค่อย ๆ เสื่อมสภาพไป 0.01 หน่วยในทุกชั่วโมงที่ใช้งาน) หรือนายแพทย์ให้บริการรักษาคนไข้โดยเฉลี่ย ($\frac{1}{\lambda}$) ชั่วโมงละ 5 คน แสดงว่าคนไข้คนหนึ่งได้รับบริการจากแพทย์โดยเฉลี่ย (λ) คนละ 20 นาที

เนื่องจากในกระบวนการปัวซอง มีข้อตกลงว่าอุบัติการณ์ที่ปรากฏในแต่ละช่วงเวลาเป็นอิสระต่อกัน นั่นหมายความว่าปรากฏการณ์ที่จะอุบัติในอนาคตเป็นอิสระกับการเกิดปรากฏการณ์ในอดีตและปัจจุบัน ลักษณะนี้เรียกว่า memoryless property และโดยเหตุที่การแจกแจงแบบเอ็กซ์โปเนนเชียลมีความสัมพันธ์กับการแจกแจงแบบปัวซองอย่างใกล้ชิด เหตุที่การแจกแจงแบบเอ็กซ์โปเนนเชียลมีความสัมพันธ์กับการแจกแจงแบบปัวซองอย่างใกล้ชิด (ดูข้อสรุปที่ 4) ดังนั้นการแจกแจงแบบเอ็กซ์โปเนนเชียลจึงมีคุณสมบัตินี้ด้วย ภาพของ pdf ของการแจกแจงแบบเอ็กซ์โปเนนเชียลปรากฏดังนี้



รูปที่ 2.3 กราฟฟังก์ชันความหนาแน่นของการแจกแจงเอ็กซ์โปเนนเชียล เมื่อ $\lambda = 1$

ทฤษฎี 2.6.1 ตัวแปรสุ่มเอ็กซ์โปเนนเชียลมีคุณสมบัติ memoryless นั่นคือ
ถ้า X = ระยะเวลาระหว่างเหตุการณ์หรืออายุใช้งานของวัตถุแล้ว

$$P(X \le t | X > t_0) = P(X \le t - t_0)$$

พิสูจน์

$$F_{x|x>t_0}(t) = P(X \le t | X > t_0)$$

$$= \frac{P(X \le t) \cap (X > t_0)}{P(X > t_0)}$$

$$= \frac{P(t_0 < X \le t)}{P(X > t_0)} \quad \text{เนื่องจาก} \quad t \ge t_0$$

$$= \frac{F_x(t) - F_x(t_0)}{1 - F_x(t_0)}$$

$$= \frac{(1 - e^{-\lambda t}) - (1 - e^{-\lambda t_0})}{1 - (1 - e^{-\lambda t_0})} \quad ; \text{ ข้อสรุปที่ 4}$$

$$= \frac{e^{-\lambda t_0} - e^{-\lambda t}}{e^{-\lambda t_0}}$$

$$= 1 - e^{-\lambda(t - t_0)} \quad ; \quad t \geq t_0$$

จะได้ $$\frac{d}{dt} F_{x|x>t_0}(t) = \frac{d}{dt}(1 - e^{-\lambda(t-t_0)}) = \lambda e^{-\lambda(t-t_0)} \quad ; \quad t \geq t_0$$

ดังนั้น $$f_{x|x>t_0}(t) = \lambda e^{-\lambda(t-t_0)} \quad ; \quad t \geq t_0$$

หมายความว่า ถ้า X หมายถึงอายุใช้งานของวัตถุแล้วจะทำให้มองเห็นภาพได้ชัดเจน ว่า $P(X < t | X = t_0) = P(X < t - t_0)$ ซึ่งหมายความว่าเมื่อวัตถุใช้งานมาแล้ว t_0 หน่วยเวลาแต่ยังไม่เสื่อมสภาพ แสดงว่าวัตถุคงจะเสื่อมสภาพในช่วงเวลาที่เหลืออยู่คือ $(t - t_0)$ ทั้งนี้การพิจารณาอายุใช้งานของวัตถุเราจะไม่สนใจเวลาใช้งานที่ผ่านมาแล้ว หรือถือว่าตราบใดที่วัตถุนั้นยังคงใช้งานได้ก็ถือเสมือนว่าวัตถุนั้นเป็นของใหม่อยู่เสมอ เช่นตราบใดที่ฟิวส์ยังคงใช้ได้ (ยังไม่ละลาย) แม้จะได้ใช้มานานแล้ว t_0 ชั่วโมง เราก็ถือว่าฟิวส์นั้นยังใช้ได้ดีเหมือนของใหม่ พฤติกรรมการเสื่อมสภาพของวัตถุจึงมิได้ขึ้นอยู่กับเวลาที่ได้ใช้งานมาแล้วนานเพียงใดแต่ขึ้นอยู่กับระยะเวลาการใช้งานแต่ละครั้งในภายหน้าว่าจะใช้ติดต่อกันไปนาน (หักโหม) เพียงใด

นั่นคือ ความน่าจะเป็นที่วัตถุจะเสื่อมสภาพเมื่อใช้งานมาแล้ว t_0 หน่วยเวลา คือความน่าจะเป็นที่วัตถุจะเสื่อมสภาพในช่วงเวลาภายหน้า

ตัวอย่าง 2.6.1 คอนเดนเซอร์ขนาดหนึ่งมีอายุใช้งานโดยเฉลี่ย 100 ชั่วโมง ถ้าพบว่าเมื่อใช้งานผ่านไปแล้ว t ชั่วโมง 90% ของคอนเดนเซอร์เหล่านั้นยังคงใช้งานได้ดี อยากทราบว่าคอนเดนเซอร์ถูกใช้งานมาแล้วกี่ชั่วโมง

วิธีทำ เมื่อใช้งานมาแล้ว t ชั่วโมงร้อยละ 90 ของคอนเดนเซอร์ยังคงทำงานได้ดี

จาก $$P(X > t) = 0.90$$

$$\because \; P(X > t) = 1 - P(X \leq t) = e^{-\lambda t}$$

$$E(X) = \frac{1}{\lambda} = 100$$

$$\lambda = 0.01$$

จะได้ $\qquad e^{-0.01t} = 0.90$

ดังนั้น $\qquad \ln(0.90) = -0.01t$

$$\therefore\ t = -100\ \ln(0.90)$$

$$= 10.54$$

แสดงว่าคอนเดนเซอร์เหล่านั้นเพิ่งใช้งานมาเพียง 10.54 ชั่วโมง

ตัวอย่าง 2.6.2 นำหลอดอิเล็กทรอนิกส์ n หลอดมาใช้งาน (เป็นอิสระต่อกัน) และทราบว่า หลอดชนิดดังกล่าวมีอายุใช้งานเฉลี่ย 100 ชั่วโมง ถ้าใช้งานไปแล้ว t ชั่วโมง จงหาความน่าจะเป็นที่อย่างน้อยครึ่งหนึ่งของหลอดดังกล่าวยังคงทำงานได้เป็นปกติ

วิธีทำ ให้ X = อายุใช้งานของหลอดอิเล็กทรอนิกส์

ดังนั้น

q = P(หลอดเสื่อมสภาพในระหว่าง (0, t) ชั่วโมง) = $P(X \le t)$

= $F_x(t)$

= $1 - (0.005e^{-0.005t})$

p = P(หลอดยังคงใช้งานได้ดีเมื่อใช้ไปแล้ว t ชั่วโมง) = $P(X > t)$

= $1 - F_x(t)$

= $0.005e^{-0.005t}$

การที่หลอดจะเสื่อมสภาพนั้น การเสื่อมสภาพอาจปรากฏกับหลอดใด ๆ ก็ได้ ดังนั้น จากหลอดทั้งสิ้น n หลอด การที่หลอดยังคงทำงานได้เป็นปกติถึง k หลอด จึงปรากฏขึ้นได้ถึง $\binom{n}{k}$ วิธี ดังนั้น

ความน่าจะเป็นที่อย่างน้อยครึ่งหนึ่งของหลอดทั้งสิ้น n หลอดยังคงทำงานได้จึงมีค่าดังนี้

$$\sum_{k=n/2}^{n}\binom{n}{k}(0.005e^{-0.005t})^{k}(1-0.005e^{-0.005t})^{n-k} \quad ; \text{ เมื่อ } n \text{ เป็นเลขคู่}$$

และ

$$\sum_{k=(n+1)/2}^{n}\binom{n}{k}(0.005e^{-0.005t})^{k}(1-0.005e^{-0.005t})^{n-k} \quad ; \text{ เมื่อ } n \text{ เป็นเลขคี่}$$

ตัวอย่าง 2.6.3 นำอุปกรณ์อิเล็กทรอนิกส์ต่อไปนี้ต่อเข้าวงจรไฟฟ้า คือ ซิลิคอนไดโอดชนิดใช้งานเฉลี่ย 500,000 ชั่วโมง (อัตราเสื่อมสภาพคงที่เท่ากับ 0.000002) 10 ตัว ซิลิคอนทรานซิสเตอร์ชนิดอายุใช้งานเฉลี่ย 100,000 ชั่วโมง (อัตราเสื่อมสภาพคงที่เท่ากับ 0.000001) 4 ตัว คอนเดนเซอร์ชนิดอายุใช้งานเฉลี่ย 500,000 ชั่วโมง (อัตราเสื่อมสภาพคงที่เท่ากับ 0.000002) 10 ตัว และตัวต้านทานชนิดอายุใช้งานเฉลี่ย 1,000,000 ชั่วโมง (อัตราเสื่อมสภาพคงที่เท่ากับ 0.000001) 20 ตัว

การต่ออุปกรณ์ดังกล่าวเป็นการต่อแบบอนุกรม และต่างก็ทำงานอย่างเป็นอิสระต่อกัน จงหาความน่าจะเป็นที่วงจรนี้ยังคงทำงานได้เป็นปกติเมื่อผ่าน การใช้งานไปแล้ว t ชั่วโมง พร้อมทั้งคาดหมายอายุใช้งานของวงจรดังกล่าว

วิธีทำ ให้ X = อายุใช้งานของวงจรไฟฟ้า

ให้ t_i = อายุใช้งานของอุปกรณ์อิเล็กทรอนิกส์ตัวที่ i; $i = 1, 3, \ldots, 44$

เนื่องจากวงจรนี้เป็นวงจรอนุกรม ดังนั้นวงจรจะยังคงทำงานได้ก็ต่อเมื่ออุปกรณ์อิเล็กทรอนิกส์ทุกตัวยังคงทำงานได้

เนื่องจาก $P(X > t) = P(X_1 > t \text{ และ } X_2 > t \text{ และ } \ldots \text{ และ } X_{44} > t)$

$= P(X_1 > t)\,P(X_2 > t)\,\ldots\,P(X_{44} > t)$

ดังนั้น $P(X_i > t) = e^{-\lambda t}$

$P(X > t) = e^{-10(0.000002)t}\,e^{-4(0.000001)t}\,.\,e^{-10(0.000002)t}\,.\,e^{-20(0.000001)t}$

$\therefore \quad P(X > t) = e^{-0.0001t}$

นั่นคือ ความน่าจะเป็นที่วงจรนี้จะยังคงทำงานได้เป็นปกติเมื่อใช้งานมาแล้ว t ชั่วโมง มีค่าเท่ากับ $e^{-0.0001t}$ เช่น ความน่าจะเป็นที่วงจรยังคงทำงานได้เป็นปกติเมื่อใช้งานมาแล้ว 10 ชั่วโมง คือ $e^{-0.0001(10)}$ หรือ

$$P(X > 10) = e^{-00001(10)} = 0.999$$

หมายความว่า เมื่อวงจรถูกใช้งานไปแล้ว 10 ชั่วโมง 99.9% ของอุปกรณ์ไฟฟ้าจะยังคงทำงานได้ดีต่อไป

พิจารณา

$$P(X > t) = e^{-0.0001t}$$

$$1 - P(X \le t) = e^{-0.0001t}$$

$$P(X \le t) = 1 - e^{-0.0001t}$$

หรือ

$$f_X(t) = 0.0001e^{-0.0001t} \quad ; \quad t \ge 0$$

จะเห็นได้ว่า $E(X) = \dfrac{1}{\lambda} = \dfrac{1}{0.0001} = 10{,}000$ ชั่วโมง

นั่นคือวงจรไฟฟ้านี้มีอายุใช้งานโดยเฉลี่ยได้นานถึง 10,000 ชั่วโมง

ตัวอย่าง 2.6.4 อายุใช้งานของหลอดภาพโทรทัศน์มีอายุใช้งานโดยเฉลี่ยได้นานถึง 1,000 ชั่วโมง จงคำนวณหา

ก. ความน่าจะเป็นที่หลอดภาพจะมีอายุใช้งานนานเกินกว่า 1,000 ชั่วโมง

ข. ความน่าจะเป็นที่หลอดภาพจะมีอายุใช้งานได้ในระหว่าง 700 ถึง 1,000 ชั่วโมง

ค. ความน่าจะเป็นที่หลอดภาพยังคงใช้งานได้เกินกว่าชั่วโมงที่ 1,200 เมื่อผ่านการใช้งานมาแล้ว 800 ชั่วโมง

วิธีทำ ให้ X = อายุใช้งานของหลอดภาพโทรทัศน์

$X \sim Ex(\lambda = 0.001)$

ก. ความน่าจะเป็นที่หลอดภาพจะมีอายุใช้งานนานเกินกว่า 1,000 ชั่วโมง

$$P(X > 1{,}000) = \int_{1000}^{\infty}(0.001)e^{-0.001x}dx$$

หรือ

$$= e^{-(0.001)(1.000)}$$

$$= e^{-1}$$

$$= 0.36788$$

ข. ความน่าจะเป็นที่หลอดภาพจะมีอายุใช้งานระหว่าง 700 ถึง 1,000 ชั่วโมง

$$P(700 \le X \le 1{,}000) = P(X \le 1{,}000) - P(X \le 700)$$

$$= (1 - e^{-(0.001)(1.000)}) - (1 - e^{-(0.001)(700)})$$

$$= (1 - 0.36788) - (1 - 0.49659)$$

$$= 0.12871$$

ค. ความน่าจะเป็นที่หลอดภาพจะมีอายุใช้งานได้นานเกินกว่าชั่วโมงที่ 1,200 เมื่อผ่านมาแล้ว 800 ชั่วโมง

$$P(X > 1{,}200 \mid X > 800) = P(X > 1{,}200 - 800)$$

$$= P(X > 400)$$

$$= e^{-(0.001)(400)}$$

$$= 0.67032$$

2.7 การแจกแจงแบบแกมมา (Gamma Distribution)

การแจกแจงแบบแกมมาเป็นรูปแบบที่ใช้ในการคำนวณหาความน่าจะเป็นของระยะเวลาระหว่างเหตุการณ์ตั้งแต่ r , $(r \geq 2)$ เหตุการณ์ขึ้นไป (time to the r^{th} event)

ขอให้ย้อนไปพิจารณาการแจกแจงเอ็กซ์โปเนนเชียลในหัวข้อ 2.6 จะเห็นว่าเป็นเรื่องของการศึกษาความผันผวนของเวลาระหว่างเหตุการณ์ 2 เหตุการณ์ในกระบวนการปัวซอง เช่น ระยะเวลาระหว่างอุบัติเหตุ 2 ครั้ง ณ จุดสี่แยกไฟแดง ระยะเวลาระหว่างการเกิดพายุใต้ฝุ่นขนาดความเร็วลม a ไมล์ต่อชั่วโมง (a อาจเป็นค่าใดก็ได้ แล้วแต่ความสนใจของผู้ศึกษา) หรือระยะเวลารอคอยจนกระทั่งเครื่องคำนวณไฟฟ้าที่ติดตั้งใหม่เริ่มรวนเป็นต้น ถ้าพิจารณาเฉพาะในแง่อายุการใช้งานจะพบว่า อุบัติการณ์จะมีได้เพียงครั้งหนึ่งเดียว (ครั้งแรก) เท่านั้น เช่นคอยจนกระทั่งหลอดไฟขาด ซึ่งโดยทั่วไปก็มีได้เพียงครั้งเดียว ที่ต้องให้ย้อนมาพิจารณาเรื่องนี้ด้วย เพราะว่าการแจกแจงแบบอนุกรมเรขาคณิตและการแจกแจงนิเสธทวินามว่ามีส่วนคล้ายคลึง (เป็นอีกโฉมหน้าหนึ่งของกันและกัน) กับการแจกแจงแบบเอ็กซ์โปเนนเชียลและการแจกแจงแบบแกมมาตามลำดับ โดยหมายถึงเวลารอคอยด้วยกันทั้ง 4 ตัวแปร เพียงแต่การแจกแจงแบบอนุกรมเรขาคณิตเป็นเรื่องของการรอคอย (discrete waiting time) จนกระทั่งปรากฏเหตุการณ์ที่สนใจเป็นครั้งแรก (ขอให้สังเกตไว้ด้วยว่าการแจกแจงแบบอนุกรมเรขาคณิตก็มี memoryless property เช่นเดียวกับการแจกแจงแบบเอ็กซ์โปเนนเชียล) โดยที่การนับอาจจะดำเนินการนับจำนวนเหตุการณ์ที่เป็นส่วนเติมเต็มของเหตุการณ์ที่สนใจ เช่น การรอคอยจนกระทั่งเจาะพบน้ำมันในอ่าวไทยเป็นครั้งแรก การรอคอยเราจะนับที่จำนวนครั้งของการขุดเจาะที่ไม่พบน้ำมัน ส่วนการแจกแจงแบบเอ็กซ์โปเนนเชียลเป็นเรื่องของเวลาที่ใช้รอคอย (contineous waiting time) จนกระทั่งเกิดอุบัติการณ์ปัวซองที่สนใจ ขณะที่การแจกแจงแบบนิเสธทวินามเป็นเรื่องของการรอคอยจนพบเหตุการณ์ที่มุ่งสนใจครบ r ครั้ง และการแจกแจงแบบแกมมาก็ศึกษาในประเด็นเดียวกัน แต่มุ่งศึกษาถึงจำนวนเวลาที่ใช้ไปในการรอคอยนั้น ดังนี้จึงเห็นได้ว่า โดยปกติเมื่อกล่าวถึงการแจกแจงแบบแกมมาแล้วเราจะไม่พูดถึงอายุการใช้งานของวัตถุ เว้นแต่จะมีข้อกำหนดเป็นอย่างอื่น เพราะโดยทั่วไปวัตถุจะไม่เสื่อมคุณภาพเกินกว่า 1 ครั้ง ถ้าเสียแล้วก็เสียไปเลยทิ้งไป

การแจกแจงใดก็ตามที่ศึกษาถึงระยะเวลาที่รอคอยจนเกิดอุบัติการณ์ปัวซองจนครบ r ครั้ง จะมีการแจกแจงแบบแกมมาเสมอ ตัวอย่างเช่น ระยะเวลารอคอยจนกระทั่งฝนตกครบ 20 ครั้ง ระยะเวลาที่รอคอยจนกระทั่งเลนเลี้ยวขวา ณ บริเวณสี่แยกไฟแดงติดขัดเมื่อเลนเลี้ยวขวาสามารถรับจำนวนรถที่คอยเลี้ยวขวาได้เพียง 10 คัน (กรณีนี้ r = 10) ความทนทานของเสาคอนกรีตซึ่งทนแรงกระแทกขนาด 700 ปอนด์ ได้อย่างมาก 10 ครั้ง (ระยะเวลารอคอยจนกระทั่งเสาคอนกรีตร้าว) ระดับน้ำขึ้นสูงสุดของแม่น้ำเจ้าพระยาในปีหนึ่ง ๆ เป็นต้น

จากคำอธิบายและตัวอย่างข้างต้น เราจะพบว่า

1. การแจกแจงแบบแกมมาจะต้องเป็นเรื่องของการรอคอยจนกระทั่งเกิดอุบัติการณ์ปัวซองครบตามจำนวนครั้งที่ต้องการ

ความสัมพันธ์ระหว่างการแจกแจงแบบแกมมากับการแจกแจงแบบปัวซองปรากฏดังนี้

ให้ X = จำนวนอุบัติการณ์ปัวซอง (ในเรื่องที่สนใจ) ที่อุบัติขึ้นในช่วงเวลา (0, t)

ให้ T = ระยะเวลาที่ใช้ในการรอคอยจนเกิดอุบัติการณ์ปัวซองครบ r ครั้ง

ดังนั้น $P(T > t)$ = ความน่าจะเป็นที่เกิดอุบัติการณ์ปัวซองไม่ถึง r ครั้งในช่วงเวลา (0, t]

$$
\begin{aligned}
F_T(t) = P(T \le t) &= 1 - P(T > t) \\
&= 1 - \sum_{x=0}^{r-1} P(X = x) \\
&= 1 - \sum_{x=0}^{r-1} \frac{e^{-\lambda t}(\lambda t)^x}{x!}
\end{aligned}
$$

ซึ่งยังแสดงว่าการคำนวณหาค่าความน่าจะเป็นของการแจกแจงแบบแกมมากระทำได้โดยตารางปัวซองด้วย เฉพาะกรณี r เป็นเลขจำนวนเต็มบวก หรือ positive integer

2. พารามิเตอร์ของการแจกแจงแบบแกมมาจะมี 2 ตัว คือ r และ λ โดยปกติแล้วเราจะไม่ทราบค่าของ r เว้นแต่จะเป็นเรื่องของการประมาณค่าหรือ convolution ของตัวแปรสุ่มเอ็กซ์โปเนนเชียล ตัวอย่างที่เห็นได้ชัดเจนคือระดับน้ำสูงสุด (ตามเกณฑ์ของนักวิจัย) โดยปกติเป็นเรื่องที่เราไม่อาจทราบได้ว่าปีหนึ่งระดับน้ำจะขึ้นสูงสุดถึงเกณฑ์ที่กำหนดไว้กี่ครั้ง หรือในเรื่องจำนวนครั้งที่ฝนตกในปีหนึ่ง ๆ เราก็ไม่อาจทราบได้ว่ามีกี่ครั้ง ดังนี้ r จึงเป็นพารามิเตอร์ตัวหนึ่ง ส่วน λ คืออัตราถัวเฉลี่ยของการเกิดอุบัติการณ์ปัวซองต่อหนึ่งหน่วยเวลา λx คือจำนวนอุบัติการณ์ปัวซองที่เกิดขึ้นโดยถัวเฉลี่ยในช่วงเวลายาว x หน่วย

3. ในทางทฤษฎีค่า r จะแสดงรูปร่างของการแจกแจง (shape parameter) ส่วน λx เป็นค่าแสดงสเกลเป็น λx และ λ

ก่อนที่จะอธิบายการสร้างฟังก์ชันการแจกแจงแบบแกมมาจะนิยามฟังก์ชันแกมมา ซึ่งจะนำไปใช้ในการสร้างฟังก์ชันการแจกแจงความหนาแน่นของการแจกแจงแกมมาดังนี้

นิยาม 2.7.1 ในวิชา culculus ชั้นสูงจะมีนิยามการอินทิเกรต $\int_0^\infty u^{r-1}e^{-u}du$ ซึ่งหาค่าได้ มีค่าจำกัดสำหรับค่า $r>0$ แทนด้วย $\Gamma(r)$ อ่านว่าแกมมาอาร์ เรียก $\int_0^\infty u^{r-1}e^{-u}du$ ว่าเป็นฟังก์ชันแกมมา จะได้ว่า

$$\Gamma(r) = \int_0^\infty u^{r-1}e^{-u}du$$

ทฤษฎี 2.7.1 ถ้า $\Gamma(r) = \int_0^\infty u^{r-1}e^{-u}du$ สำหรับ $r>0$ แล้ว $\Gamma(r) = (r-1)\Gamma(r-1)$

พิสูจน์ ให้ $u = u^{r-1}$

$$du = (r-1)u^{r-2}$$

และ $dv = \int_0^\infty e^{-u}du$

$$v = -e^{-u}$$

ดังนั้น $\Gamma(r) = \int_0^\infty u^{r-1}e^{-u}du$

$$= -u^{r-1}e^{-u}\Big|_0^\infty + \int_0^\infty (r-1)u^{r-2}e^{-u}du$$

$$= 0+\int_0^\infty (r-1)u^{r-2}e^{-u}du$$

$$= (r-1)\cdot\int_0^\infty u^{(r-1)-1}e^{-u}du$$

$$= (r-1)\Gamma(r-1) \quad \text{เมื่อ } r>0$$

และถ้า r เป็น positive integer จะได้

$$\Gamma(r) = (r-1)\cdot(r-2)\cdot(r-3) \quad \ldots \; 3.2.1$$

$$= (r-1)!$$

นั่นคือ $\Gamma(r) = (r-1)!$

สำหรับ $(\Gamma(r))^2$ จะเขียนแทนด้วย $\Gamma^2(r)$

การสร้างฟังก์ชันการแจกแจงแบบแกมมาเราดำเนินการได้ 2 วิธี คือ

ก. โดยอาศัยการรวมตัวกัน (convolution) ของตัวแปรสุ่มเอ็กซ์โปเนนเชียลที่เป็นอิสระต่อกัน กล่าวคือ

ถ้า X_i ; i = 1,2, ...,r เป็นอิสระต่อกันและต่างก็มีการแจกแจงแบบเอ็กซ์โปเนนเชียล มีพารามิเตอร์ λ แล้ว ตัวแปรสุ่ม (ตัวสถิติ) $Y = \sum_{i}^{r} X_i$ จะมีการแจกแจงแบบแกมมา มีพารามิเตอร์ r และ λ

กรณีนี้ค่าของ r จะเป็นเลขจำนวนเต็ม (integer-valued) ดังกรณีตัวอย่างที่ยกให้เห็นในตอนต้น

ข. โดยอาศัยแนวคิด (concept) เกี่ยวกับสถานการณ์ของการแจกแจงแบบแกมมา และการปรับปรุงฟังก์ชันแกมมาให้เป็นฟังก์ชันความน่าจะเป็นที่มีพารามิเตอร์ 2 ตัว กล่าวคือ

จากฟังก์ชันแกมมา $\Gamma(r) = \int_0^\infty u^{r-1}e^{-u}du$

แปลงรูปตัวแปร u เป็น λx คือให้ $u = \lambda x$ เพื่อให้ฟังก์ชันกลายเป็นฟังก์ชันที่มีพารามิเตอร์ 2 ตัว

จะได้ $\Gamma(r) = \int_0^\infty (\lambda x)^{r-1} e^{(\lambda x)} dx \;\; ; \lambda > 0$

หารตลอดด้วย $\Gamma(r)$

จะได้ $f_x(x) = \dfrac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda r} \, ; 0 \le x \le \infty$ เรียกว่า การแจกแจงแกมมา

กรณีนี้เป็นกรณีทั่วไปของการแจกแจงแบบแกมมา และเพราะเหตุที่พัฒนามาจากฟังก์ชันแกมมา r จึงไม่จำเป็นต้องเป็นเลขจำนวนเต็ม ส่วนค่าของ λ อาจเป็นจำนวนเต็มหรือไม่ก็ได้ อนึ่งเหตุที่การแจกแจงประเภทนี้มีชื่อเรียกว่า “แกมมา” ก็ด้วยเหตุผลข้อ ข. นี้

นิยาม 2.7.2 ถ้าตัวแปรสุ่ม X มี pdf ดังต่อไปนี้คือ

เมื่อ X = ระยะเวลาทั้งหมดที่ใช้ในการรอคอยจนเกิดอุบัติการณ์ป่วยของครบ r ครั้ง

$$G(r, \lambda) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x} \quad ; \quad x \geq 0, \lambda > 0, r > 0$$

$$= 0 \; ; \text{ เมื่อ } x \text{ มีค่าเป็นอย่างอื่น}$$

แล้วตัวแปรสุ่ม X จะเป็นตัวแปรสุ่มแกมมาและเรียก pdf นี้ว่า "การแจกแจงแบบแกมมา"

ถ้า r เป็นตัวเลขจำนวนเต็ม คือ r = 1, 2, ... ซึ่งเป็นกรณีเฉพาะของการแจกแจงของแกมมา

$$f_x(x) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x} \quad ; \quad x \geq 0, \lambda > 0, r = 1, 2, \ldots$$

นิยมเรียกชื่อฟังก์ชันความน่าจะเป็นนี้ว่า erlang distribution หรือ pearson type-III distribution เพราะการแจกแจงแบบแกมม่าสอดคล้องกับระบบเพียร์สัน (pearsonian system of density function)

ถ้า pdf ใด ๆ สอดคล้องกับสมการดิฟเฟอเรนเชียล $\frac{1}{f_x(x)}\frac{d}{dx}f_x(x) = \frac{x+a}{b_0 + b_1x + b_2x^2}$ โดยที่ a , b_0 , b_1 , b_2 เป็นจำนวนคงที่แล้ว ถือว่า pdf นั้นสอดคล้องกับ pearson system ในกรณีของการแจกแจงแบบแกมมา พบว่า $\frac{1}{f_x(x)}\frac{d}{dx}f_x(x) = \frac{x-(r-1)/\lambda}{-x/\lambda}$

จะเห็นว่า $a = -\frac{r-1}{\lambda}, b_0 = 0, b_1 = \frac{-1}{\lambda}, b_2 = 0$ แสดงว่าการแจกแจงแบบแกมมาสอดคล้องกับ pearsonian system

ขอให้สังเกตค่าของ $\Gamma(r)$ และ $(r - 1)!$ จะพบว่า $\Gamma(r)$ นั้นสามารถหาค่า $\Gamma(r)$ ได้เสมอในทุกค่า r โดยที่ r ไม่จำเป็นต้องเป็นเลขจำนวนเต็ม ส่วนกรณี $(r - 1)!$ นั้น เราจะสามารถหาค่า $(r - 1)!$ ได้เฉพาะเมื่อ r เป็นเลขจำนวนเต็มบวกเท่านั้น แสดงว่า $\Gamma(r) = (r - 1)!$ เมื่อ r เป็นเลขจำนวนเต็มบวก

กรณีเฉพาะของการแจกแจงแบบแกมมาที่สำคัญ ที่ถูกนำไปใช้มากในทางปฏิบัติคือ

ก. เมื่อ $r = 1$ และ $\lambda = c\,(c > 0)$ การแจกแจงแบบแกมมาจะลดรูปเป็นการแจกแจงแบบเอ็กซ์โปเนนเชียล มีพารามิเตอร์ λ และมี pdf ดังนี้

$$f_x(x) = \lambda e^{-\lambda x} \quad ; \quad x \geq 0, \lambda > 0$$

ข. เมื่อ $r = 1$ และ $\lambda = 1$ การแจกแจงแบบแกมมาจะลดรูปเป็นการแจกแจงแบบเอ็กซ์โปเนนเชียลมีพารามิเตอร์ $\lambda = 1$ และมี pdf ดังนี้

$$f_x(x) = e^{-x} \quad ; \quad x \geq 0$$

ค. ถ้า $r = \frac{n}{2}$ และ $\lambda = \frac{1}{2}$ การแจกแจงแบบแกมมาจะเปลี่ยนรูปเป็นการแจกแจงแบบไคสแควร์ (χ^2 – Distribution) มีพารามิเตอร์ n (degree of freedom) และมี pdf ดังนี้

$$f_x(x) = \frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right)} x^{(n/2)-1} e^{-(1/2)x} \quad ; \quad x \geq 0$$

ง. ถ้า $r = \frac{1}{2}$ และ $\lambda = \frac{1}{2}$ การแจกแจงแบบแกมมาจะเปลี่ยนรูปเป็นการแจกแจงแบบไคสแควร์มีพารามิเตอร์ 1 (degree of freedom) และมี pdf ดังนี้

$$f_x(x) = \frac{\left(\frac{1}{2}\right)^{1/2}}{\Gamma\left(\frac{1}{2}\right)} x^{-(1/2)} e^{-(1/2)x} \quad ; \quad x \geq 0$$

หรือ $$= \frac{1}{\sqrt{2\pi}} x^{-(1/2)} e^{-x/2} \quad ; \quad x \geq 0$$

$$\because \quad \Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

กรณีเฉพาะของ ค. และ ง. ถูกนำไปใช้อย่างกว้างขวางในการวิเคราะห์ความแปรปรวน การสร้างการทดสอบสำหรับทดสอบความแปรปรวนและการสร้างการทดสอบสถิติ (test statistics) สำหรับทดสอบภาวะรูปดี (goodness of fit) และทดสอบความเป็นอิสระของคุณลักษณะทางประชากร

ทฤษฎี 2.7.2 ถ้าตัวแปรสุ่ม X มีการแจกแจงแบบแกมมา $G(r, \lambda)$ มี pdf ดังนี้

$$f_x(x) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x} \quad ; \quad x \ge 0, \lambda > 0, r > 0$$

แล้ว X จะมี mgf ค่าคาดหมายและความแปรปรวนดังต่อไปนี้

$$M_x(t) = \left(\frac{\lambda}{\lambda - t}\right)^r \quad , \quad E(X) = \frac{r}{\lambda} \quad , \quad V(X) = \frac{r}{\lambda^2}$$

พิสูจน์

จาก $$f_x(x) = \frac{\lambda}{\Gamma(r)}(\lambda x)^{r-1}e^{-\lambda x} \quad ; \quad x \ge 0$$

จะได้
$$\begin{aligned} M_x(t) &= \int_0^\infty e^{tx}\frac{\lambda^r x^{r-1}}{\Gamma(r)}e^{-\lambda x}dx \\ &= \frac{\lambda^r}{\Gamma(r)}\int_0^\infty x^{r-1}e^{-(\lambda-t)x}dx \\ &= \frac{\lambda^r}{\Gamma(r)}\int_0^\infty \frac{1}{(\lambda-t)^r}u^{r-1}e^{-u}du, \quad \text{แปลงรูป } (\lambda\text{-}t)x = u \\ &= \frac{\lambda^r}{\Gamma(r)}\frac{1}{(\lambda-t)^r}\Gamma(r) \\ &= \left(\frac{\lambda}{\lambda-t}\right)^r \end{aligned}$$

$$E(X) = M'_x(t)\big|_{t=0}$$

$$M'_x(t)\Big|_{t=0} = \frac{d}{dt}\left(\frac{\lambda}{\lambda-t}\right)^r\Big|_{t=0} = \lambda^r\frac{r}{(\lambda-t)^{r+1}}\Big|_{t=0}$$

$$= \frac{r}{\lambda}$$

ดังนั้น $E(X) = \frac{r}{\lambda}$

$$M''_x(t)\Big|_{t=0} = \lambda^r \frac{r(r+1)}{(\lambda+t)^{r+2}}\Big|_{t=0} = \frac{r(r+1)}{\lambda^2}$$

และ $V(X) = M''_x(t)\Big|_{t=0} - \left(\frac{r}{\lambda}\right)^2$

$$= \frac{r(r+1)}{\lambda^2} - \frac{r^2}{\lambda^2} = \frac{r}{\lambda^2}$$

ดังนั้นถ้าตัวแปรสุ่ม X มีการแจกแจงแบบแกมมา $G(r, \lambda)$ แล้ว

$$M_x(t) = \left(\frac{\lambda}{\lambda-t}\right)^r, E(X) = \frac{r}{\lambda} \text{ และ } V(X) = \frac{r}{\lambda^2}$$

ข้อสังเกต เมื่อพิจารณากรณีเฉพาะตามหมายเหตุท้ายนิยามที่ 2.7.1 จะพบว่า

ก. ถ้า $r = 1$ และ $\lambda = c\ (c > 0)$ จะพบว่าตัวแปรสุ่ม X มี pdf, mgf, ค่าคาดหมายและความแปรปรวนดังนี้

$$f_x(x) = \lambda e^{-\lambda x}\ ; x \geq 0, \lambda > 0$$

$$M_x(t) = \frac{\lambda}{\lambda-t}, E(X) = \frac{1}{\lambda}, V(X) = \frac{1}{\lambda^2}$$

ข. ถ้า $r = 1$ และ $\lambda = 1$ ตัวแปรสุ่ม X จะมี pdf, mgf, ค่าคาดหมายและความแปรปรวนดังนี้

$$f_x(x) = e^{-x}\ ; x \geq 0$$

$$M_x(t) = \frac{1}{1-t} = (1-t)^{-1}, E(X) = 1 \text{ และ } V(X) = 1$$

ค. ถ้า $r=\frac{n}{2}$ และ $\lambda=\frac{1}{2}$ ตัวแปรสุ่ม X จะมี pdf, mgf , ค่าคาดหมายและความแปรปรวน ดังนี้

$$f_x(x) \;=\; \frac{\left(\frac{1}{2}\right)^{n/2}}{\Gamma\left(\frac{n}{2}\right)} x^{(n/2)-1} e^{-(1/2)x} \quad ; \quad x \geq 0$$

$$M_x(t) \;=\; \left(\frac{1/2}{(1/2)-t}\right)^{n/2} \left(\frac{1}{1-2t}\right)^{n/2} \;=\; (1-2t)^{-n/2}$$

$$E(X) \;=\; \frac{n/2}{1/2} = n \text{ และ } V(x) \;=\; \frac{n/2}{1/4} = 2n$$

ง. ถ้า $r=\frac{1}{2}$ และ $\lambda=\frac{1}{2}$ ตัวแปรสุ่ม X จะมี pdf, mgf , ค่าคาดหมายและความแปรปรวน ดังนี้

$$f_x(x) \;=\; \frac{\left(\frac{1}{2}\right)^{1/2}}{\Gamma\left(\frac{1}{2}\right)} x^{-1/2} e^{-x/2} \; ; x \geq 0$$

$$M_x(t) \;=\; \left(\frac{\frac{1}{2}}{\frac{1}{2}-t}\right)^{1/2} \;=\; (1-2t)^{-1/2}$$

$$E(X) \;=\; 1 \text{ และ } V(X) \;=\; 2$$

ตัวอย่าง 2.7.1

ก. จงแสดงให้เห็นว่าการแจกแจงแบบแกมมาเกิดจากการรวมตัวกันของตัวแปรสุ่มเอ็กซ์โปเนนเชียลที่เป็นอิสระต่อกัน

ข. จงแสดงให้เห็นว่าตัวแปรสุ่มแกมมามีคุณสมบัติแห่งการรวมตัวกัน เมื่อตัวแปรสุ่มทุกตัวต่างก็มีพารามิเตอร์ λ เดียวกันและเป็นอิสระต่อกัน

ค. จงแสดงให้เห็นว่าตัวแปรสุ่ม χ^2 มีคุณสมบัติแห่งการรวมตัวกัน เมื่อตัวแปรสุ่มทุกตัวเป็นอิสระต่อกัน

ก. $X \sim Ex(\lambda)$; i = 1, 2, ... r และต่างก็เป็นอิสระต่อกัน

วิธีทำ

ให้ $Y = \sum_{i}^{r} X_i$

ดังนั้น $M_Y(t) = \prod_{i}^{r} M_{x_i}(t)$

$$= \underbrace{\frac{\lambda}{\lambda - t} \cdot \frac{\lambda}{\lambda - t} \cdots \frac{\lambda}{\lambda - t}}_{r\text{ ครั้ง}}$$

$$= \left(\frac{\lambda}{\lambda - t}\right)^{r}$$

จะได้ $Y \sim G(r, \lambda)$

ข. ให้ $X_i \sim G(r_i, \lambda)$; i = 1, 2, ... , n และต่างก็เป็นอิสระต่อกัน

ให้ $Y = \sum_{i}^{n} X_i$

ดังนั้น $M_Y(t) = \prod_{i}^{n} M_{x_i}(t)$

$$= \left(\frac{\lambda}{\lambda - t}\right)^{r_1} \cdot \left(\frac{\lambda}{\lambda - t}\right)^{r_2} \cdots \left(\frac{\lambda}{\lambda - t}\right)^{r_n}$$

$$= \left(\frac{\lambda}{\lambda - t}\right)^{r_1 + r_2 + \ldots + r_n}$$

นั่นคือ Y มีการแจกแจงแบบแกมมาที่มีพารามิเตอร์เป็น $\sum_{i}^{n} r_i$ และ λ ตามลำดับ

ดังนั้น $Y \sim G\left(\sum_{i}^{n} r_i, \lambda\right)$ หรือตัวแปรสุ่มแกมมามี reproductive property

ค. ให้ $X_i \sim \chi^2_{(n_i)}; i = 1, 2, \ldots, k$ และต่างก็เป็นอิสระต่อกัน

เนื่องจาก $X_i \sim \chi^2_{(n_i)}$ ดังนั้นแสดงว่า $X_i \sim G\left(r = \frac{n_i}{2}, \lambda = \frac{1}{2}\right)$

$$\text{จะได้}\quad M_{x_i}(t) = \left(\frac{1/2}{1/2-t}\right)^{n_i/2} = \left(\frac{1}{1-2t}\right)^{n_i/2} = i = 1, 2, \dots, k$$

$$\text{ให้}\qquad Y = \sum_{i}^{k} X_i$$

$$\text{ดังนั้น}\ M_Y(t) = \prod_{i}^{n} M_{x_i}(t) = \left(\frac{1}{1-2t}\right)^{(n_1+n_2+\dots+n_k)/2}$$

$$\text{นั่นคือ}\qquad Y \sim \chi^2_{(n1+n2+\dots+nk)} = \left(\frac{1}{1-2t}\right)^{(n1+n2+\dots+nk)/2}$$

แสดงว่าตัวแปรสุ่ม χ^2 มีคุณสมบัติแห่งการรวมตัวกัน

กรณีเฉพาะที่สำคัญคือ เมื่อ $X_i \sim \chi^2_{(1)}; i = 1, 2, \dots, n$ แล้ว $Y = \sum_{i}^{n} X_i \sim \chi^2_{(n)}$

ซึ่งแสดงให้เห็นได้โดยง่ายดังนี้

เนื่องจาก $X_i \sim \chi^2_{(1)}$

$$\therefore\ M_{x_i}(t) = \left(\frac{1}{1-2t}\right)^{1/2}$$

ดังนั้นเมื่อให้ $Y = \sum_{i}^{n} X_i$

$$\text{จะได้}\ M_Y(t) = \prod_{i}^{n} M_{x_i}(t) = \left(\frac{1}{1-2t}\right)^{(1+1+\dots+1)/2}$$

$$= \left(\frac{1}{1-2t}\right)^{n/2}$$

แสดงว่าเมื่อตัวแปรสุ่ม $X_i \sim \chi^2_{(1)}$ แล้วตัวแปรสุ่ม $Y = \sum_{i}^{n} X_i$ จะมีการแจกแจงแบบ $\chi^2_{(n)}$

หมายเหตุ Convolution หมายถึงการรวมตัวของตัวแปรสุ่มประเภทเดียวกันซึ่งเมื่อรวมตัวกันแล้ว ฟังก์ชันของตัวแปรสุ่ม (Y) จะมีการแจกแจงเป็นรูปอื่น เช่นตัวแปรสุ่มเอ็กซ์โปเนนเชียลรวมตัวกันแล้วมีการแจกแจงแบบแกมมา ตัวแปรสุ่มอนุกรมเรขาคณิตรวมตัวกันแล้วมีการแจกแจงแบบทวินามนิเสธ ตัวแปรสุ่มแบบเบอร์นูลลิรวมตัวกันแล้วเป็นการแจกแจงแบบทวินาม เป็นต้น convolution มีประโยชน์มาก ในแง่ของการพัฒนาฟังก์ชันความน่าจะเป็นของตัวแปรสุ่มตัวใหม่ขึ้นมา และในแง่ของการหาฟังก์ชันความน่าจะเป็นของตัวสถิติ

Reproductive Property หมายถึงการรวมตัวของตัวแปรสุ่มประเภทเดียวกัน ซึ่งเมื่อรวมตัวกันแล้ว ฟังก์ชันของตัวแปรสุ่ม (Y) จะยังคงมีฟังก์ชันความน่าจะเป็นเช่นเดิม เช่นตัวแปรสุ่มทวินามรวมตัวกันแล้วยังคงมีการแจกแจงแบบทวินาม ตัวแปรสุ่มปัวซองรวมตัวกันแล้วยังคงมีการแจกแจงแบบปัวซอง ตัวแปรสุ่มแบบแกมมารวมตัวกันแล้วยังคงมีการแจกแจงแบบแกมมา ตัวแปรสุ่ม χ^2 รวมตัวกันแล้วยังคงมีการแจกแจงแบบ χ^2 ตัวแปรสุ่มปกติรวมตัวกันแล้วยังคงมีการแจกแจงแบบปกติ เป็นต้น อย่างไรก็ตามฟังก์ชันความน่าจะเป็นที่เกิดขึ้นใหม่นี้อาจมีโครงสร้างผิดไปจากเดิมบ้าง และอาจจำเป็นต้องมีข้อจำกัดเกี่ยวกับพารามิเตอร์

ทั้ง convolution และ reproductive property จำเป็นต้องมีเงื่อนไขว่าตัวแปรสุ่มที่จะมารวมตัวกันนั้นต้องเป็นอิสระต่อกัน ซึ่งเมื่อเป็นดังนี้จึงเป็นประโยชน์มากในทฤษฎีการสุ่มตัวอย่างและฟังก์ชันความน่าจะเป็นของตัวสถิติ ทั้งนี้เพราะหน่วยตัวอย่าง (sampling unit) ในแต่ละตัวอย่าง (sample) จะต้องเป็นอิสระต่อกันเสมอ ซึ่งจะพบประโยชน์ของทั้ง convolution และ reproductive property ในเรื่องการทดสอบสมมติฐาน

ตัวอย่าง 2.7.2 จากสถิติกระแสน้ำไหลสูงสุด (maximum stream flow) รายปีของแม่น้ำสายหนึ่ง การสังเกตกระทำติดต่อกันทุกปี ตั้งแต่ พ.ศ. 2473 ถึง พ.ศ. 2503 พบว่ากระแสน้ำสูงสุดมีการแจกแจงแบบแกมมา $G(r = 1.727,\ \lambda = 0.00672$ ลบ.ฟ.) จงคำนวณค่าคาดหมาย ความแปรปรวนและความน่าจะเป็นที่กระแสน้ำไหลสูงสุดในปีหนึ่ง ๆ จะมีค่าไม่เกิน 400 ลบ.ฟ.

วิธีทำ ให้ X = ปริมาณกระแสน้ำไหลสูงสุด

เนื่องจาก $X \sim G(1.727,\ 0.00672)$

ดังนั้น
$$E(X) = \frac{r}{\lambda} = \frac{1.727}{0.00672} = 256.7$$

$$V(X) = \frac{r}{\lambda^2} = \frac{1.727}{(0.00672)^2} = 36{,}100$$

$$P(X \le 400) = P(2\lambda x \le (2)(0.0672)(400))$$

$$= P(\chi^2_{(2r)} \le 5.4)$$

$$= P(\chi^2_{(3.452)} \le 5.4)$$

$$\cong P(\chi_4^2 \le 5.4)$$

$$\cong 1 - 0.24866$$

$$\cong 0.75134$$

หรือจะคำนวณหาโดยตรงจาก pdf ของ X ก็ได้

$$P(X \le 400) = \int_0^{400} \frac{(0.00672)^{1.727}}{\Gamma(1.727)} x^{0.727} e^{-0.00672x} dx$$

จะเห็นว่าอินทิเกรตค่อนข้างยุ่งยาก จึงสามารถคำนวณได้โดยอาศัยทฤษฎีโน้มสู่ส่วนกลาง (central limit theorem)

ได้ดังนี้

$$P(X \le 400) = P\left(\frac{X - E(X)}{\sigma_x} \le \frac{400 - 256.7}{190}\right)$$

$$\cong P(Z \le 0.76)$$

$$\cong 0.7989$$

ตัวอย่าง 2.7.3 อายุใช้งานของอุปกรณ์อิเล็กทรอนิกส์แต่ละตัวในวงจรไฟฟ้ามีการแจกแจงแบบเอ็กซ์โปเนนเชียล มีค่าโดยเฉลี่ยเท่ากับ 1,000 ชั่วโมง (หรือระยะเวลาระหว่างการเสื่อมสภาพของอุปกรณ์อิเล็กทรอนิกส์ต่าง ๆ ในวงจรมีค่าเฉลี่ย 1,000 ชั่วโมง)

ก. จงหาความน่าจะเป็นที่อุปกรณ์ชิ้นที่ 3 จะเสื่อมสภาพภายหลังจากที่ใช้ไป แล้ว 4,250 ชั่วโมง

ข. จงหาค่าคาดหมายระยะเวลาการใช้งานของวงจรจนกระทั่งอุปกรณ์อิเล็กทรอนิกส์เสื่อมสภาพครบ 3 ตัว

วิธีทำ ให้ X = อายุใช้งาน (เวลาระหว่างเหตุการณ์ที่ "อุปกรณ์แต่ละชิ้นเสื่อมคุณภาพ") ของอุปกรณ์อิเล็กทรอนิกส์

$X \sim Ex(\lambda = 0.001)$

ดังนั้น

$$Y = \sum_{i}^{3} X_i \sim G(r = 3, \lambda = 0.001)$$

จะได้ $E(Y) = \dfrac{3}{0.001}$ และ $V(Y) = \dfrac{3}{(0.001)^2}$

ก. สามารถประมาณค่าด้วย CLT ได้ดังนี้

$$\begin{aligned} P(Y > 4250) &\cong P\left(\frac{(Y - E(Y))}{\sigma_Y} > \frac{4250 - 3/(0.001)}{\sqrt{3/(0.001)^2}}\right) \\ &\cong P\left(P > \frac{1250}{1732.1}\right) \\ &\cong P(Z > 0.72) \\ &\cong 0.1921 \end{aligned}$$

ข. $E(X) = \dfrac{r}{\lambda} = \dfrac{3}{0.001} = 3000$ ชั่วโมง

นั่นคือ โดยเฉลี่ยแล้วคาดว่าวงจรไฟฟ้าจะใช้งานได้นานประมาณ 3,000 ชั่วโมง อุปกรณ์อิเล็กทรอนิกส์จึงจะเสื่อมสภาพครบ 3 ตัว