บทที่ ๓

ทฤษฎีเบื้องต้นเกี่ยวกับ Order Statistics

๓.๑ คำนำ

ทฤษฎีเกี่ยวกับ order statistics นั้นถ้าดูจากหนังสือด้านสถิติหรือคณิตศาสตร์และวารสารที่เกี่ยวข้องในสาขาวิชาทั้งสองนี้จะเห็นว่ามีอยู่หลายทฤษฎีด้วยกัน แต่อันที่จริงแล้วอาจประมวลไว้เป็นทฤษฎีหลักสำคัญ ๆ เพียงสองสามทฤษฎีเท่าที่จะได้เสนอไว้ในบทนี้เท่านั้น แต่ที่เห็นกันว่ามีหลายทฤษฎีก็เพราะการที่พยายามนำเอาหลักสำคัญ ๆ เหล่านี้ไปใช้กับฟังก์ชันของการกระจายเฉพาะแบบซึ่งมีอยู่อย่างมากมายนั่นเอง อาทิเช่น นำไปใช้กับการกระจายแบบปกติ หรือการกระจายแบบแกมม่า เป็นต้น

๓.๒ การกระจายของ Order Statistics

ให้ $x_1, x_2, \ldots, x_n$ เป็นตัวอย่างขนาด n หน่วยที่ถูกเลือกจากประชากรที่มีการกระจายของความน่าจะเป็นเป็นฟังก์ชันต่อเนื่อง $f(x)$ โดยที่มี $x_{(1)}$ เป็นค่าต่ำสุดในตัวอย่างชุดนี้ $x_{(2)}$ เป็นค่าต่ำที่ต่อจากค่าต่ำสุด และเป็นเช่นนั้นเรื่อย ๆ ไปจนถึง $x_{(n)}$ เป็นค่าสูงสุดในตัวอย่างชุดดังกล่าว ดังที่เราเรียก $x_{(i)}$ ว่าเป็น order statistic ตัวที่ i ของกลุ่มตัวอย่าง $x_1, x_2, \ldots, x_n$ แล้ว เราสามารถหาฟังก์ชันของการกระจายร่วมของ $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ หรือ $g(x_{(1)}, x_{(2)}, \ldots, x_{(n)})$ ได้ดังนี้

$$g(X_{(1)},X_{(2)},\ldots,X_{(n)}) = n!f(X_{(1)})f(X_{(2)})\ldots f(X_{(n)}), \quad (3.1.1)[1]$$

$$-\infty < X_{(1)} < X_{(2)} < \ldots < X_{(n)} < \infty$$

พิสูจน์ สมมุติให้มีการเลือกตัวอย่างออกมาขนาด ๓ หน่วย (n=3) ให้ตัวอย่างดังกล่าวเป็น X_1, X_2, X_3 โดยที่ $-\infty < X_i < \infty$ เราจะได้

$X_{(1)} = X_1$, $X_{(2)} = X_2$, $X_{(3)} = X_3$, ถ้า $X_1 < X_2 < X_3$,

$X_{(1)} = X_1$, $X_{(2)} = X_3$, $X_{(3)} = X_2$, " $X_1 < X_3 < X_2$,

$X_{(1)} = X_2$, $X_{(2)} = X_1$, $X_{(3)} = X_3$, " $X_2 < X_1 < X_3$,

$X_{(1)} = X_2$, $X_{(2)} = X_3$, $X_{(3)} = X_1$, " $X_2 < X_3 < X_1$,

$X_{(1)} = X_3$, $X_{(2)} = X_1$, $X_{(3)} = X_2$, " $X_3 < X_1 < X_2$,

$X_{(1)} = X_3$, $X_{(2)} = X_2$, $X_{(3)} = X_1$, " $X_3 < X_2 < X_1$,

โดยที่ region ทั้งหกอันนี้อันได้แก่ $X_1 < X_2 < X_3$, $X_1 < X_3 < X_2$, ..., $X_3 < X_2 < X_1$. ต่างก็เป็น disjoint และ $-\infty < X_1 < \infty$, $-\infty < X_2 < \infty$, $-\infty < X_3 < \infty$.

ซึ่งเราสามารถหา joint density ของ Xi's ได้

$$q(X_1, X_2, X_3) = f(X_1)\, f(X_2)\, f(X_3), \quad (3.1.2)$$

$$-\infty < X_1 < \infty,$$
$$-\infty < X_2 < \infty,$$
$$-\infty < X_3 < \infty.$$

[1] Robert V. Hogg and Allen T. Craig, Introduction to Mathematical Statistics (2nd ed.; New York: The Macmillan Company, 1965), p. 168.

แต่สำหรับการหา joint density ของ $X_{(i)}$'s เราจะต้องดูไปที่ละ region ดังนี้

ใน region ที่ $x_1 < x_2 < x_3$ เราจะได้

$h(X_{(1)}, X_{(2)}, X_{(3)}) = f(X_{(1)})\, f(X_{(2)})\, f(X_{(3)}),\ X_{(1)} < X_{(2)} < X_{(3)}$

โดยที่ Jacobian = 1

ใน region ที่ $x_1 < x_3 < x_2$ เราจะได้

$h(X_{(1)}, X_{(2)}, X_{(3)}) = f(X_{(1)}) f(X_{(3)}) f(X_{(2)}),\ X_{(1)} < X_{(2)} < X_{(3)}$

โดยที่ Jacobian = 1

ในทำนองเดียวกันเราจะได้ในอีก ๔ regions ดังนี้

$h(X_{(1)}, X_{(2)}, X_{(3)}) = f(X_{(2)}) f(X_{(1)}) f(X_{(3)}),\ X_{(1)} < X_{(2)} < X_{(3)}$

$h(X_{(1)}, X_{(2)}, X_{(3)}) = f(X_{(2)}) f(X_{(3)}) f(X_{(1)}),\ X_{(1)} < X_{(2)} < X_{(3)}$

$h(X_{(1)}, X_{(2)}, X_{(3)}) = f(X_{(3)}) f(X_{(1)}) f(X_{(2)}),\ X_{(1)} < X_{(2)} < X_{(3)}$

$h(X_{(1)}, X_{(2)}, X_{(3)}) = f(X_{(3)}) f(X_{(2)}) f(X_{(1)}),\ X_{(1)} < X_{(2)} < X_{(3)}$

จะเห็นได้ว่า density ของ $X_{(i)}$'s ในทั้งหก regions ต่างก็อยู่บน region $X_{(1)} < X_{(2)} < X_{(3)}$ เดียวกัน ดังนั้น joint density ของ $X_{(i)}$'s จะได้จากผลรวมของ ๖ regions นั่นเอง

นั่นคือ $h(X_{(1)}, X_{(2)}, X_{(3)}) = 6 f(X_{(1)}) f(X_{(2)}) f(X_{(3)}),$

$$-\infty < X_{(1)} < X_{(2)} < X_{(3)} < \infty$$

หรือ $h(X_{(1)}, X_{(2)}, X_{(3)}) = 3!\, f(X_{(1)}) f(X_{(2)}) f(X_{(3)}),$

$$-\infty < X_{(1)} < X_{(2)} < X_{(3)} < \infty$$

ด้วยวิธีการข้างต้นนี้เราสามารถพิสูจน์ได้ว่า

$h(X_{(1)}, X_{(2)}, \ldots, X_{(n)}) = n!\, f(X_{(1)}) f(X_{(2)}) \ldots f(X_{(n)}),$

$$-\infty < X_{(1)} < X_{(2)} < \cdots < X_{(n)} < \infty$$

๓.๓ การกระจายร่วมกันของ $X_{(r)}$ และ $X_{(s)}$

ถ้า $X_{(r)}$ และ $X_{(s)}$ เป็นตำแหน่งที่ r และ s, $r < s$, ของ order statistic จากสุ่มตัวอย่างขนาด n หน่วยจากประชากรที่มีการกระจายของโอกาสแห่งความน่าจะเป็นเป็น $f_\theta(x)$ แล้ว การกระจายร่วมกันของ $X_{(r)}$ และ $X_{(s)}$,

$$f(x_{(r)}, x_{(s)}) = \frac{n!}{(r-1)!(1)!(s-r-1)!(1)!(n-s)!}\left[F(x_{(r)})\right]^{(r-1)}$$
$$\cdot f_\theta(x)\left[F(x_{(s)}) - F(x_{(r)})\right]^{(s-r-1)}$$
$$\cdot f_\theta(x)\left[1 - F(x_{(s)})\right]^{(n-s)} \qquad (3.2.1)$$
[2]

หมายความว่าจะมีค่าจากตัวอย่างจำนวน $(r-1)$ ตัวที่มีค่าต่ำกว่า $X_{(r)}$, อีก ๑ ตัวมีค่าเท่ากับ $X_{(r)}$, อีก $(s-r-1)$ ตัวมีค่าอยู่ระหว่าง $X_{(r)}$ กับ $X_{(s)}$, อีก ๑ ตัวมีค่าเท่ากับ $X_{(s)}$, และอีก $(n-s)$ ตัวมีค่าสูงกว่า $X_{(s)}$.

โดยอาศัยทฤษฎีของ Multinomial distribution[3] เราสามารถอธิบายการกระจายร่วมกันของ $X_{(r)}$ และ $X_{(s)}$ ได้ดังนี้

โดยที่ $(r-1) + 1 + (s-r-1) + 1 + (n-s) = n =$ ขนาดของตัวอย่าง เพราะขนาดของตัวอย่างเป็น n หน่วย ย่อมทำให้ทางที่จะเกิดค่าของ x_i จากตัวอย่างนั้นเป็นได้ $n!$ ทาง

---

[2] Ibid., p.173.

[3] สำหรับทฤษฎี Multinomial อาจหาดูได้จากหนังสือ Advanced Statistical Methods in Biometric Research ของ C.Radhakrishna Rao ฉบับพิมพ์ปี 1952 ของสำนักพิมพ์ John Wiley and Sons, Inc. พิมพ์ที่ New York. หน้า ๓๔ – ๓๕

∵ $[F(X_{(r)})] = P[X < X_{(r)}]$ = โอกาสที่ x จะมีค่าต่ำกว่าค่าของ $x_{(r)}$

∴ โอกาสที่ x จำนวน $(r-1)$ ตัวจะมีค่าต่ำกว่า $x_{(r)} = \dfrac{[F(X_{(r)})]^{r-1}}{(r-1)!}$

เหตุที่ต้องหารด้วย $(r-1)!$ นั้นเพราะในจำนวน $(r-1)$ ตัวของ x ที่มีค่าต่ำกว่าค่าของ $x_{(r)}$ นั้น เราไม่คำนึงถึงว่า x ตัวใดจะเกิดก่อนหรือเกิดทีหลัง..

ด้วยวิธีการเดียวกันนี้เราจะได้

โอกาสที่ x จำนวน ๑ ตัวมีค่าเท่ากับ $X_{(r)} = f_{\theta}\dfrac{(X)}{(1)!}$

โอกาสที่ x จำนวน $(s-r-1)$ ตัวจะมีค่าอยู่ระหว่าง $x_{(r)}$ กับ $x_{(s)}$

$$= \frac{[F(X_{(s)}) - F(X_{(r)})]^{(s-r-1)}}{(s-r-1)!}$$

โอกาสที่ x จำนวน ๑ ตัวจะมีค่าเท่ากับ $X_{(s)} = f_{\theta}\dfrac{(X)}{(1)!}$

โอกาสที่ x จำนวน $(n-s)$ ตัวจะมีค่าสูงกว่า $x_{(s)} = \dfrac{[1-F(X_{(s)})]^{(n-s)}}{(n-s)!}$

นั่นคือ

$$f(X_{(r)}, X_{(s)}) = \frac{n!}{(r-1)!(1)!(s-r-1)!(1)!(n-s)!}[F(X_{(r)})]^{(r-1)} f_{\theta}(X)$$

$$\cdot [F(X_{(s)}) - F(X_{(r)})]^{(s-r-1)} f_{\theta}(X)[1-F(X_{(s)})]^{(n-s)}$$

จาก joint density, $f(X_{(r)}, X_{(s)})$, เราสามารถหาค่าของ joint density ใด ๆ ได้โดยการแทนค่าของ r และ s เช่น ถ้า $r=1$, $s = n$ เราจะได้ joint density ของ $X(\max)$ และ $X(\min)$ ดังนี้

$$f(x_{\min}, x_{\max}) = n(n-1)[F(X_{(n)}) - F(X_{(1)})]^{n-2} f(X_{(1)}) f(X_{(n)}).$$

<u>พิสูจน์</u>

$$\because f(X_{(1)}, X_{(n)}) = \int_{x_{(1)}}^{x_{(n)}} \cdots \int_{x_{(1)}}^{x_{(4)}} \int_{x_{(1)}}^{x_{(3)}} h(X_{(1)}, X_{(2)}, \ldots, X_{(n)})\, dx_{(2)}\, dx_{(3)} \cdots dx_{(n-1)}$$

$$-\infty < X_{(1)} < X_{(2)} < \cdots < X_{(n)} < \infty$$

$$= n!\, f(X_{(1)}) f(X_{(n)}) \int_{x_{(1)}}^{x_{(n)}} \cdots \int_{X_{(1)}}^{x_{(4)}} \int_{X_{(1)}}^{x_{(3)}} f(x_{(2)}) f(X_{(3)}) \cdots$$

$$f(X_{(n-1)})\; dx_{(1)}\, dx_{(2)} \cdots dx_{(n-1)}$$

เราจะแยกออกมาทำการ integrate ทีละส่วนได้ดังนี้

$$\int_{x_{(1)}}^{x_{(3)}} f(x_{(2)})\, dx_{(2)} = F(x_{(3)}) - F(x_{(1)})$$

$$\int_{x_{(1)}}^{x_{(4)}} \int_{x_{(1)}}^{x_{(3)}} f(x_{(3)}) f(x_{(2)})\, dx_{(2)}\, dx_{(3)} = \int_{x_{(1)}}^{x_{(4)}} [F(x_{(3)}) - F(x_{(1)})] f(x_{(3)})\, dx_{(3)}$$

$$= \frac{[F(x_{(4)}) - F(x_{(1)})]^2}{2}$$

$$\text{และ } \frac{1}{2}\int_{x_{(1)}}^{x_{(5)}} [F(x_{(4)}) - F(x_{(1)})]^2 f(x_{(4)})\, dx_{(4)} = \frac{1}{3!}\{F(x_{(5)}) - F(x_{(1)})\}^3$$

ในที่สุดเราจะได้

$$f(x_{(1)}, x_{(n)}) = n(n-1)[F(x_{(n)}) - F(x_{(1)})]^{n-2} f(x_{(1)}) f(x_{(n)})$$

เป็นที่น่าสังเกตว่า เมื่อเราทราบค่าของขนาดของตัวอย่าง ค่าของคะแนนสูงสุดและต่ำสุดของตัวอย่าง และทราบฟังก์ชันของการกระจายของความน่าจะเป็น เราก็สามารถที่จะทราบถึงการกระจายร่วมกันของค่าคะแนนสูงสุดและคะแนนต่ำสุดของข้อมูลตัวอย่างชุดนั้นได้

จากการกระจายร่วมกันของคะแนนสูงสุดและคะแนนต่ำสุด, $f(X_{(1)}, X_{(n)})$, ถ้าเราอาศัยคุณสมบัติบางประการในทางคณิตศาสตร์ เช่น transformation property เราสามารถที่จะหาการกระจายที่น่าสนใจอื่น ๆ ได้อาทิเช่น การกระจายของพิสัย, $R = X_{(n)} - X_{(1)}$ หรือการกระจายของกึ่งพิสัย, $\frac{X_{(n)} + X_{(1)}}{2}$ ได้

เช่นในกรณีที่จะหาการกระจายของพิสัย

$$\because f(x_{(1)}, x_{(n)}) = n(n-1)\{F(x_{(n)}) - F(x_{(1)})\}^{n-2} f(x_{(1)}) f(x_{(n)})$$

เราให้ $R = X_{(n)} - X_{(1)}$

$S = X_{(1)}$

(ค่าสัมบูรณ์ของ Jacobian = 1)

เราจะได้

$$g(R,S) = n(n-1)[F(R+S) - F(S)]^{n-2} f(S) f(R+S)$$

จากนั้นถ้าเรา integrate g(R, S) with respect to S
เราก็จะได้ marginal density เป็น density ของพิสัยตามต้องการ

ตัวอย่างที่ ๓.๓.๑

ถ้า $f(X) = 1 \quad , \quad 0 < X < 1$

โดยมี $x_1, x_2, \ldots, x_6$ เป็นตัวอย่างขนาด ๖ หน่วย

$$\therefore F(x) = \int_0^x f(x)\,dx = \int_0^x (1)\,dx$$
$$= X$$

$$\Rightarrow \quad F(R+S) = R+S \quad , \quad F(S) = (S)$$

$$\therefore g(R,S) = 6(6-1)[(R+S)-S]^{6-2}$$
$$= 30R^4 \qquad \begin{matrix} 0<S<1-R \\ 0<R<1 \end{matrix}$$

$$\therefore g(R) = \int_0^{1-R} 30R^4\,dS = 30R^4\int_0^{1-R}(1)\,dS$$
$$= 30R^4(1-R) \quad , \quad 0<R<1 .$$

๓.๔ การกระจายของ $x_{(r)}$

ถ้า $X_{(r)}$ เป็นค่าตำแหน่งที่ r ของ order statistic จากสุ่มตัวอย่างขนาด n หน่วยจากประชากรที่มีการกระจายของโอกาสแห่งความน่าจะเป็นเป็น $f_\theta(X)$ แล้ว ฟังก์ชันของการกระจายของ $x_{(r)}$, $f(X_{(r)})$, จะเป็นดังนี้

$$f(x_{(r)}) = \frac{n!}{(r-1)!\,1!\,(n-r)!}[F(x_{(r)})]^{(r-1)} f_\theta(x)\,[1-F(x_{(r)})]^{n-r} ,$$
$$-\infty < x_{(1)} < x_{(2)} < \cdots < x_{(n)} < \infty$$

หมายความว่าจะมีค่าจากตัวอย่างจำนวน (r-1) ตัวที่มีค่าต่ำกว่า $x_{(r)}$ และอีก (n - r) ตัวจะมีค่าสูงกว่า $x_{(r)}$ การอธิบายสามารถทำได้เช่นเดียวกับที่ได้แสดงไว้ใน ๓.๒ ทุกประการ

จากฟังก์ชั่นของการกระจายของ $X_{(r)}$ เราสามารถที่จะหาฟังก์ชั่นของการกระจายของค่าคะแนนต่ำสุด, $X_{(1)}$, ของคะแนนสูงสุด, $X_{(n)}$, ของมัธยฐาน, $X_{(\frac{n}{2})}$, ของเปอร์เซ็นต์ไทล์, $X_{(\frac{Pn}{100})}$, ของควอไทล์, $X_{(\frac{Q}{4}n)}$ ของเดไซล์, $X_{(\frac{D}{10}n)}$ ฯลฯ ดังตัวอย่างต่อไปนี้

$$f(X_{(1)}) = n\left[1-F(X_{(1)})\right]^{n-1} f_{\theta}(X)$$

$$f(X_{(n)}) = n\left[F(X_{(n)})\right]^{n-1} f_{\theta}(X)$$