บทที่ 4
การแจกแจงแบบทวินามและไฮเปอร์จีออเมตริก
(Binomial and Hpergeomatric Distribution)

4.1 (Bernnoulli Trails)

Definition1

การทดลองแบบเบอนูลลี่เป็นการทดลองที่ในแต่ละครั้งของการทดลอง จะประกอบด้วยผลที่เป็นไปได้เพียง 2 อย่างเท่านั้น ได้แก่ความสำเร็จ (success) กับความไม่สำเร็จ (failure) โดยที่ความน่าจะเป็นของการเกิดความสำเร็จในการทดลองแต่ละครั้งเป็น p ความน่าจะเป็นของการเกิดความไม่สำเร็จในการทดลองแต่ละครั้งเป็น (1-p)

ในแต่ละครั้งของการทดลองแบบเบอนูลลี่ ถ้ากำหนดให้ตัวแปรสุ่ม(Random variable) X มีค่าได้เพียง 2 ค่า คือ

X = 1 เมื่อเกิดความสำเร็จ (success)

X = 0 เมื่อเกิดไม่ความสำเร็จ (failure)

เรียก x ส่ามีการแจกแจงแบบแบบเบอนูลลี่ ที่มี p.d.f. เป็น

$$f(x) = p^x(1-p)^{1-x},\ \ \mathrm{x} = 0{,}1$$

อันดับ(sequence) ของการทดลองแบบเบอนูลลี่ n ครั้ง ถ้าให้ x_i เป็นผลที่ได้จากการทดลองครั้งที่ i ผลที่ได้จะออกมาเป็นอันดับที่ของ 1 และ 0

Example 1 ทอดลูกเต๋าเที่ยงตรง 1 ลูก 4 ครั้ง ถ้าลูกเต๋าขึ้นแต้ม 6 จะนับว่าเป็นความสำเร็จ ถ่ขึ้นแต้มอื่นเป็นความไม่สำเร็จ จดผลของการทดลองไว้ สมมติว่าครั้งแรก ขึ้นแต้ม 3 ครั้งที่สองขึ้นแต้ม 4 ครั้งที่สามขึ้นแต้ม 6 และครั้งที่สี่ขึ้นแต้ม 1 นั่นคือ

เราจะได้อันดับของการทดลองนี้ประกอบด้วย 1 และ 0 คือ (0,0,1,0) โดยที่แต่ละครั้งการทดลองเป็นอิสระต่อกัน ความน่าจะเป็นที่จะได้ผลของการทดลองครั้งนี้เท่ากับ (5/6) (5/6) (1/6) (5/6)

ในอันดับของการทดลองแบบเบอนูลลี่ เราสนใจแต่จำนวนครั้งของการเกิดความสำเร็จในการทดลองทั้งหมด n ครั้ง โดยไม่สนใจว่าจะเกิดในครั้งใดของการทดลอง เราจะเรียกชนิดของการทดลองนี้ว่า การทดลองแบบทวินาม(Binomial Experiment)

4.2 การทดลองแบบทวินาม (Binomial Experiment)

Definition2

การทดลองแบบทวินามเป็นการทดลองที่กระทำซ้ำกัน n ครั้งโดยแต่ละครั้งของการทดลอง จะประกอบด้วยผลที่เป็นไปได้เพียง 2 อย่างเท่านั้น ได้แก่ความสำเร็จ (success) กับความไม่สำเร็จ (failure)

ถ้า x เป็นจำนวนครั้งของการเกิดความสำเร็จในการทดลอง ซึ่งกระทำซ้ำๆกัน n ครั้ง โยที่การทดลองแต่ละครั้งเป็นอิสระต่อกัน และความน่าจะเป็นของการเกิดความสำเร็จในแต่ละครั้งมีค่าคงที่ เราจะเรียกตัวแปร x ว่าเป็นตัวแปรสุ่มแบบทวินาม (Binomial Random variable)

เช่น ในการโยนเหรียญเที่ยงตรง 1 อัน 5 ครั้ง การโยนเหรียญแต่ละครั้งเป็นอิสระต่อกัน เหรียญอาจจะขึ้นหัวหรือก้อยก็ได้ ถ้าเราให้การขึ้นหัวเป็นความสำเร็จ หากขึ้นหัว 3 ครั้ง ความสำเร็จก็มีค่าเท่ากับ 3

หรือการหยิบไพ่ทีละใบ 5 ครั้ง โดยวิธีหยิบแล้วคืนเข้าที่จากไพ่ทั้งสำรับ ถ้าได้ไพ่สีแดง นับว่าเป็นความสำเร็จ ได้ไพ่ชนิดอื่นไม่สำเร็จ

จะเห็นว่าการทดลองทั้งสองชนิดมีลักษณะเหมือนกัน คือ การทดลองแต่ละครั้งเป็นอิสระต่อกัน ความน่าจะเป็นของการทดลองแต่ละครั้งมีค่าคงที่คือ $\frac{1}{2}$ เราเรียกการทดลองแบบนี้ว่าการทดลองแบบทวินาม (Binomial Experiment)

อย่างไรก็ตาม หากในการหยิบไพ่ทีละใบ 5 ครั้ง โดยวิธีหยิบแล้วไม่คืนที่ (without replacement) ความน่าจะเป็นที่จะได้ไพ่สีแดงในการหยิบครั้งแรกเป็น $\frac{1}{2}$ แต่ในการหยิบครั้งที่สองจะเป็น.....หรือ..... ขึ้นอยู่กับว่าการหยิบครั้งแรกได้ไพ่สีแดงหรือไม่ การทดลองแบบนี้จะไม่เรียกว่าเป็นการทดลองแบบทวินาม

4.3 ทฤษฎีบททวินาม (The Binomial Theorem)

ทฤษฎีบททวินามสำหรับเลขชี้กำลังเป็นเลขจำนวนเต็มบวก หากพิจารณาการกระจาย $(x+y)^3$ โดยใช้ Distributive Law ซ้ำ จนกระทั่งกระจายหมดจะได้ว่า

$$
\begin{aligned}
(x+y)^3 &= (\mathrm{x}+\mathrm{y})(\mathrm{x}+\mathrm{y})(\mathrm{x}+\mathrm{y}) \\
&= \mathrm{x}(\mathrm{x}+\mathrm{y})(\mathrm{x}+\mathrm{y}) + y(\mathrm{x}+\mathrm{y})(\mathrm{x}+\mathrm{y})
\end{aligned}
$$

$$
\begin{aligned}
(x+y)^3 &= x\{x(x+y)+y(x+y)\} + y\{x(x+y)+y(x+y)\} \\
&= xxx + xxy + xyx + xyy + yxx + yxy + yyx + yyy \\
&= xxx + (xxy + xyx + yxx) + (xyy + yxy + yyx) + yyy
\end{aligned}
$$

จะเห็นว่าแต่ละเทอมประกอบด้วย x และ y (หรือ x หรือ y อย่างเดียว) ตามกันสามตัวซึ่งตัวหนึ่งๆจะมาจากวงเล็บหนึ่งๆ

ดังนั้นเมื่อกระจาย $(x+y)^{\mathrm{n}}$ โดยใช้ Distributive Law ซ้ำๆ จนกระทั่งกระจายหมด จะพบว่าแต่ละเทอมในรูปกระจายจะประกอบด้วย x และ y (หรือ x หรือ y อย่างเดียว รวมกัน n ตัว ซึ่งตัวหนึ่งๆจะมาจากวงเล็บหนึ่ง

การรวมเทอมที่เหมือนกัน จะดูว่าเทอม$y^{\mathrm{r}}x^{n-r}$ เกิดขึ้นกี่ครั้ง ซึ่งจะเห็นได้ว่าเทอม $y^{\mathrm{r}}x^{n-r}$ ได้มาจากการเลือก *r* วงเล็บ จากทั้งหมด *n* วงเล็บ และนำ *y* จาก *r* วงเล็บที่เลือกมาคูณกับ *x* ที่มาจากอีก *(n-r)* วงเล็บที่เหลือ ซึ่งจะทำได้ $\binom{\mathrm{n}}{\mathrm{r}}$วิธี แสดงว่าสัมประสิทธิ์ของเทอม $y^{\mathrm{r}}x^{n-r}$ คือ $\binom{\mathrm{n}}{\mathrm{r}}$

ดังนั้น

$$(x+y)^n = \sum_{r=0}^{n} \binom{n}{r} y^r x^{n-r}$$

ซึ่งเทอมที่ *(r+1)* คือ $\binom{\mathrm{n}}{\mathrm{r}} y^{\mathrm{r}} x^{n-r}$

เราอาจจะแสดงการกระจาย$(x+y)^{\mathrm{n}}$ ในรูปอื่นๆต่อไปได้เช่น
เราอาจแสดงการกระจาย $(x+y)^n$ ในรูปอื่นๆได้เช่น

$$(x+y)^n = \binom{n}{0}x^n + \binom{n}{1}x^{n-1}y + \binom{n}{2}x^{n-2}y^2 + \ldots + \binom{n}{r}x^{n-r}y^r + \ldots + \binom{n}{n}y^n$$

$$= \binom{n}{0} + \binom{n}{1}x + \binom{n}{2}x^2 + \ldots + \binom{n}{r}x^r + \ldots + \binom{n}{n}x^n$$

$$= 1 + \binom{n}{1}x + \binom{n}{2}x^2 + \ldots + \binom{n}{r}x^r + \ldots + x^n$$

ทั้งนี้เพราะ $\binom{n}{0} = \binom{n}{n} = 1$

<u>Example 2</u> จงกระจาย $(x+y)^6$

<u>วิธีทำ</u> $(x+y)^6 = \binom{6}{0}x^6 + \binom{6}{1}x^5y + \binom{6}{2}x^4y^2 + \binom{6}{3}x^3y^3 + \binom{6}{4}x^2y^4 + \binom{6}{5}xy^5 + y^6$

$$= x^6 + 6x^5y + 15x^4y^2 + 20x^3y^3 + 15x^2y^4 + 6xy^5 + y^6$$

4.4 ความน่าจะเป็นของการแจกแจงแบบทวินาม *(The Binomial Probability Distribution)*

พิจารณาการทดลองแบบทวินาม ที่ได้จากการโยนเหรียญ *3* อัน *1* ครั้ง

S = {TTT, TTH, THT, HTT, THH, HTH, HHT, HHH}

สมมติให้ x *เป็นจำนวนเหรียญที่ขึ้น หัว*

Outcome	x
TTT	0
TTH	1
THT	1
HTT	1
THH	2
HTH	2
HHT	2
HHH	3

สมมติให้ความน่าจะเป็นที่เหรียญต่ละอันจะขึ้นหัวเป็น p

และให้ความน่าจะเป็นที่เหรียญต่ละอันจะขึ้นก้อยเป็น $1-p=q$, $p+q = 1$

เพราะการโยนเหรียญแต่ละอันเป็นเหตุการณ์ที่เป็นอิสระต่อกัน

ดังนั้น $P(TTT) = P(T) \times P(T) \times P(T) = q \times q \times q = q^3$

ดังนั้น $P(TTH) = P(T) \times P(T) \times P(H) = q \times q \times p = q^2p$

ในทำนองเดียวกันเราสามารถหาความน่าจะเป็นของผลที่อาจเกิดขึ้นได้ในการทดลองแต่ละครั้ง ดังตาราง

Outcome	x	P (X =x)
TTT	0	q^3
TTH	1	pq^2
THT	1	pq^2
HTT	1	pq^2
THH	2	p^2q
HTH	2	p^2q
HHT	2	p^2q
HHH	3	p^3

ซึ่งอาจนำมาเขียนใหม่ได้ดังนี้

x	*0*	*1*	*2*	*3*
P (X =x)	q^3	$3pq^2$	$3p^2q$	p^3

จะเห็นได้ว่า ความน่าจะเป็นของ *x* จะเรียงลำดับ ออกมาตรงกับเทอมต่างๆในการกระจาย $(q+p)^3$

ถ้าการทดลองโยนเหรียญ *1* อัน *n* ครั้ง
จะเห็นได้ว่า ความน่าจะเป็นของ x จะเรียงลำดับออกมาตรงกับเทอมต่างๆในการกระจาย $(p+q)^3$

$$(p+q)^3 = q^3+3pq^2+p^2q+p^3$$

$$\mathrm{P}(\mathrm{X}=x_i)=q^3+3pq^2+p^2q+p^3$$

$$=(q+p)^3=1$$

ดังนั้น x จะมีค่าเป็น 0,1,2,…, n

$\therefore$ ความน่าจะเป็นของ x จะเรียงลำดับออกมา ตรงกับเทอมต่างๆในการกระจาย $(q+p)^n$ ซึ่งมีเทอมที่ x เป็น $\binom{n}{x}p^xq^{n-x}$

ดังนั้น $\mathrm{P}(\mathrm{X}=x)=\binom{n}{x}p^xq^{n-x}$

$$\sum_{x=0}^{n}\mathrm{P}(\mathrm{X}=x)=\sum_{x=0}^{n}\binom{n}{x}p^xq^{n-x}$$

$$=q^n+\binom{n}{1}pq^{n-1}+\binom{n}{2}p^2q^{n-2}+\ldots+\binom{n}{x}pq^{n-x}+\ldots+p^n$$

$$=(p+q)^n=1$$

Definition 3 ถ้า *x* เป็นจำนวนของความสำเร็จในการทดลองสุ่ม *n* ครั้ง ของการทดลองแบบทวินาม ความน่าจะเป็นของความสำเร็จในการทดลองแต่ละครั้งเป็น *p* และความน่าจะเป็นของความไม่สำเร็จในการทดลองแต่ละครั้งเป็น *q* แล้วเราจะเรียกการแจกแจงความน่าจะเป็นของ *x* ว่าเป็นการแจกแจงแบบทวินาม ซึ่งมี

$$(p+q)^n=\sum_{x=0}^{n}\binom{n}{x}p^xq^{n-x}$$

Example 3 ความน่าจะเป็นที่คนไข้จะหายจากการเป็นโรคชนิดหนึ่งเป็น 0.4 ถ้ามีคนไข้เป็นโรคนี้ 15 คน จงหาความน่าจะเป็นที่

ก) มีผู้ที่หายจากโรคนี้อย่างน้อยที่สุด 10 คน

ข) มีผู้หายจากโรคนี้ตั้งแต่ 3 ถึง 8 คน

วิธีทำ ให้ x เป็นจำนวนที่หายจากโรคนี้

$p = \mathrm{P}$ (แต่ละคนจะหายจากโรคนี้) $= 0.4$

$q = \mathrm{P}$ (แต่ละคนจะไม่หายจากโรคนี้) $= 0.6$

ก) $$\mathrm{P}(x \geq 10) = \sum_{x=10}^{15}\binom{15}{x}(0.4)^x(0.6)^{15-x}$$
$$= \binom{15}{10}(0.4)^{10}(0.6)^5 + \binom{15}{11}(0.4)^{11}(0.6)^4 + \ldots + \binom{15}{15}(0.4)^{15}$$
$$= 0.0245 + 0.0074 + 0.0003 + \ldots + 0.000$$
$$= 0.0338$$

ข) $$\mathrm{P}(3 \leq x \leq 8) = \sum_{x=3}^{8}\binom{15}{3}(0.4)^x(0.6)^{15-x}$$

4.5 Mean and Variable of the Binomial Distribution

ในการทดลองแบบทวินาม ถ้า x เป็นจำนวนความสำเร็จที่เกิดจากการทดลองชนิดสุ่ม n ครั้ง ตัวเลขที่นับได้อาจแทนด้วยสัญญาลักษณ์ $\sum \mathrm{I}_j$

โดยให้ I_j เป็นผลที่ได้จากการทดลองครั้งที่ j ซึ่งในกรณีนี้ I_j จะมีเพียง 2 ค่า คือ สำเร็จ กับ ไม่สำเร็จ โดยที่

$\mathrm{I}_j = 1$ ถ้าการทดลองครั้งที่ j สำเร็จด้วยความน่าจะเป็น p

$\mathrm{I}_j = 0$ ถ้าการทดลองครั้งที่ j ไม่สำเร็จด้วยความน่าจะเป็น q

$$\mathrm{E}(\mathrm{I}) = \sum_{j=1}^{n}\mathrm{I}_j \cdot \mathrm{P}(\mathrm{I}_j) = (1 \times p) + (0 \times q) = p$$
$$\because \quad x = \sum_{j=1}^{n}\mathrm{I}_j$$
$$\therefore \quad \mu = \mathrm{E}\left(\sum_{j=1}^{n}\mathrm{I}_j\right) = \sum_{j=1}^{n}\left(\mathrm{E}(\mathrm{I}_j)\right)$$
$$= p + p + \ldots + p \ \ (n \text{ เทอม}) = np$$

$$\text{Var}(I_j) = E[I_j - E(I_j)]^2$$
$$= E(I_j^2) - (E(I_j))^2$$
$$= (1^2 \cdot p + 0^2 \cdot q) - p^2 \;\; = p(1-p)$$

$$\therefore \text{ Var } (x) = \text{Var } \left[\sum_{j=1}^{n} I_j\right]$$
$$= \sum_{j=1}^{n} \text{Var } (I_j)$$
$$= pq + pq + \ldots pq \;\; (n \text{ เทอม}) \;\; = npq$$

Definition ถ้าการแจกแจงแบบทวินาม แล้วมัชฌิมเลขคณิต (μ) ของ x จะเท่ากับ np และความแปรปรวน (σ^2) จะเท่ากับ npq ซึ่งใช้สัญญาลักษณ์ $X \sim B\ (n, p)$ แทนโดยที่

$$P(X = x) = \binom{n}{x} p^x q^{n-x} \quad ; \; x = 0,1,2,\ldots,n$$

<u>Example 4</u> จากตัวอย่างที่ผ่าน จงหา μ (ค่ากลางมัชฌิมเลขคณิต) และส่วนเบี่ยงเบนมาตรฐานของผู้ที่หายจากโรคนั้น

วิธีทำ $p = 0.4 \quad q = 0.6 \quad n = 15$

$$\mu = np \quad = (15)(0.4) \quad = 6$$

$$\sigma^2 = npq \quad = (15)(0.4)(0.6) \; = 3.6$$

$$\sigma = \sqrt{3.6} \;\; = 1.9$$

<u>4.6 การแจกแจงแบบไฮเปอร์จีออเมตริก (Hypergeometric Distribution Definition)</u>

ให้ x เป็นตัวแปรสุ่ม ซึ่งแทนจำนวนความสำเร็จที่ได้จากการสุ่มหยิบของจำนวน n สิ่งจากของ n สิ่ง ซึ่งประกอบด้วยของ 2 พวก พวกหนึ่งมี M สิ่ง อีกพวกหนึ่งมี $N-M$ สิ่ง โดยการ ไม่แทนที่ (Without replacement) จะเรียก x ว่า เป็นตัวแปรสุ่มแบบไฮเปอร์จีออเมตริก (Hypergeometric RV.)

โดยฟังก์ชั่นความน่าจะเป็นของ n กำหนดให้ได้ ดังนี้

$p(x) = \dfrac{\binom{M}{X}\binom{N-M}{n-X}}{\binom{N}{n}}$, $x<M, x<n$ และ $n-x<N-M$ การแจกแจงของความน่าจะเป็นของตัวแปรสุ่ม x นี้เราเรียกว่า Hypergeometric Distribution

4.7 The Mean and Varian of the Hypergeometric Distribution

Theorem ถ้า x มีการแจกแจงแบบไฮเปอร์จีออเมตริก ที่มีฟังก์ชันความน่าจะเป็น

เป็น $p(x) = \dfrac{\binom{m}{x}\binom{N-M}{n-x}}{\binom{N}{n}}$; $x < M$, $x < n$ และ $n-x < N-M$

แล้วมัชฌิมเลขคณิต (μ) และความแปรปรวน (σ^2) ของการแจกแจงชนิดนี้คือ

$$\mathrm{E}(x) = \mu = n\left(\frac{M}{N}\right)$$

$$\mathrm{Var}(x) = \sigma^2 = \left(\frac{N-n}{N-1}\right)\cdot n\left(\frac{M}{N}\right)\left(1-\frac{M}{N}\right)$$

Example 5 บริษัทแห่งหนึ่งมีพนักงานทั้งสิ้น 1,000 คน 800 คน จบปริญญาตรี สุ่มตัวอย่างมา 100 คน จงประมาณจำนวนพนักงานที่จบปริญญาตรี จากการสุ่มครั้งนี้พร้อมทั้งความแปรปรวน

วิธีทำ ให้ x เป็นจำนวนพนักงานที่สุ่มมา

$$\mathrm{E}(x) = n\left(\frac{M}{N}\right)$$

$$= 100\left(\frac{800}{1{,}000}\right) = 80$$

$$\mathrm{Var}(x) = \left(\frac{N-n}{N-1}\right)\cdot n\left(\frac{M}{N}\right)\left(1-\frac{M}{N}\right)$$

$$= \left(\frac{1{,}000-100}{999}\right)\cdot 100\left(\frac{800}{1000}\right)\left(1-\frac{800}{1000}\right) = 14.41$$

จำนวนผู้ที่จบปริญญาตรี 80 คน ด้วยความแปรปรวน 14.41 คน