บทที่ 7

การแจกแจงทวินาม และการแจกแจงปกติ
(Binomial and Normal Distribution)

7.1 การแจกแจงทวินาม (Binomial distribution)

การแจกแจงทวินามเป็นการแจกแจงข้อมูลที่ได้จากตัวแปรที่มีลักษณะไม่ต่อเนื่อง (Discrete variable) การแจกแจงข้อมูลจะเป็นการแจกแจงแบบทวินามก็ต่อเมื่อการทดลองนั้นมีโอกาสเกิดผลได้ 2 อย่าง คือ สำเร็จ (success) กับไม่สำเร็จ (failure) นิยมให้ p แทน ความน่าจะเป็นของความสำเร็จ และ q แทน ความน่าจะเป็นของความไม่สำเร็จ

ถ้าให้ x เป็นจำนวนครั้งที่เกิดผล ซึ่งมีความน่าจะเป็นในแต่ละครั้งเป็น p

f (x) จะเป็นการแจกแจงทวินามก็ต่อเมื่อ

$$f(x) = \binom{n}{x} p^{x} q^{(n-x)} \quad , \quad x = 0, 1, 2, \dots, n$$

ตัวอย่าง 1 หลอดไฟบริษัทหนึ่ง มีความน่าจะเป็นที่จะมีอายุใช้งานต่ำกว่า 1 ปี เท่ากับ 0.1 ถ้านายช่างของบริษัทสุ่มหลอดมาตรวจ 10 หลอด ความน่าจะเป็นที่จะได้หลอดที่มีอายุใช้งานต่ำกว่า 1 ปี 3 หลอด เป็นเท่าไร

$$f(x) = \binom{10}{3} (0.1)^3 (0.9)^{10-3}$$

$$= \frac{10\,!}{3\,!\,(10-3)\,!} (0.1)^3 (0.9)^7$$

$$= 12'0 \times 001 \times .478$$

$$= .057$$

7.2 คุณสมบัติของการแจกแจงทวินาม

ถ้าตัวแปรสุ่ม X มีการแจกแจงทวินาม จะได้ว่า

$$E(X) = np$$

$$v(X) = npq$$

พิสูจน์

$$E(X) = \sum_{x=o}^{n} x \binom{n}{x} p^{x} q^{n-x}$$

$$= 0 + \sum_{x=1}^{n} x \frac{n!}{x!(n-x)!} p^{x} q^{n-x}$$

$$= \sum_{x=1}^{n} x \frac{n(n-1)!}{x(x-1)!(n-x)!} p\ p^{X-l} \quad 4^{n-x}$$

$$= np \sum_{x=1}^{n} \frac{(n-1)!}{(x-1)!(n-x)!} p^{x-1} q^{n-x}$$

ให้ $m = n - 1,\ y = x - 1$

$$= np \sum_{y=o}^{m} \frac{m!}{y!(y-m)!} p^{y} q^{m--r}$$

$$= np \sum_{y=o}^{m} \binom{n}{y} p^{y} q^{m-y}$$

$$= np (p+q)^{m}$$

$$= np$$

$$V(X) = E(X^{2}) - [E(X)]^{2}$$

$$E(X^{2}) = \sum_{x=o}^{n} x^{2} \binom{n}{x} p^{x} q^{n-x}$$

$$= \sum_{x=1}^{n} x^{2} \frac{n!}{x!(n-x)!} p^{x} q^{n-x}$$

$$A_{x=1}\ X \frac{n!}{(X-l)!(n-x)!} p^{x} q^{n-x}$$

ให้ $m = n - 1,\ y = x - 1$

$$= np \sum_{y=o}^{m} (y+1) \frac{m!}{y!(m-y)!} p^{Y} q^{m-y}$$

$$= np \left[\sum_{y=0}^{m} y \binom{n}{y} p^{y} q^{m-y} + \sum_{y=0}^{m} \binom{m}{y} p^{y} q^{m-y} \right]$$

$$= np \left[mp + 1 \right]$$

$$= np \left[(n-1) p + 1 \right]$$

$$= n^2 p^2 - np^2 + np$$

$$V(X) = n^2 p^2 - np^2 + np - (np)^2$$

$$= np - np^2$$

$$= np (1-p)$$

$$= npq$$

7.3 การแจกแจงปกติ (Normal distribution)

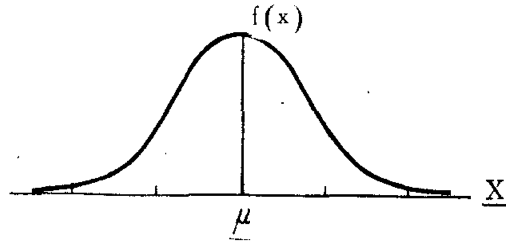
การแจกแจงปกติเป็นการแจกแจงเชิงทฤษฎี (Theorethical distribution) ของตัวแปรที่มีลักษณะต่อเนื่อง (Continuous variable) การแจกแจงปกติเป็นการแจกแจงที่ปรากฏบ่อยครั้งในการวิจัยทางการศึกษาและจิตวิทยา โดยลักษณะของการแจกแจงแบบนี้สามารถพบได้จากปรากฏการณ์ต่าง ๆ ตามธรรมชาติ เช่น ความสูง น้ำหนัก คะแนนผลการสอบ ฯลฯ ผู้ที่ค้นพบการแจกแจงปกติ และเผยแพร่การแจกแจงปกติจนเป็นที่รู้จักกันแพร่หลายคือ คาร์ล เก๊าส์ (Carl Gauss) นักคณิตศาสตร์ และฟิสิกส์ชาวเยอรมัน คาร์ล เก๊าส์ ได้กำหนดสมการของการแจกแจงปกติ ดังนี้

$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} \qquad , -\infty < x < \infty$$

เมื่อ e เป็นค่าของ natural logarithm ซึ่งมีค่าเท่ากับ 2.71828...

π มีค่าเท่ากับ 3.14159...

การแจกแจงปกติสามารถเขียนกราฟแสดงได้ดังนี้



คือรูปเป็นโค้งระฆังคว่ำ มีลักษณะสมมาตรเมื่อเทียบกับแกน f (x)

f (x) ไม่มีโอกาสเป็น 0 พื้นที่ภายใต้โค้ง f (x) เป็น $1 = \int_{-\infty}^{\infty} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} dx$

และพื้นที่ภายใต้โค้ง f (x) ระหว่าง x = a ถึง x = b = $\int_{a}^{b} \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} dx$

ในกรณีที่ $\mu = 0, \sigma = 1$ คือ

$$f(x) = \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}}$$

เราเรียก f (x) ว่าเป็นการแจกแจงปกติมาตรฐาน (Standard normal distribution)
ถ้านิยามคะแนนมาตรฐาน Z ว่า

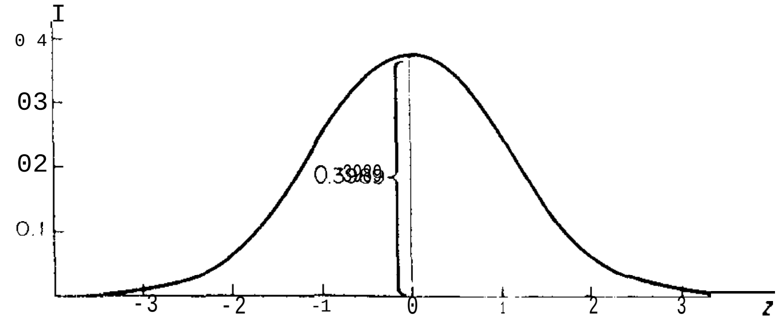
$$Z = \frac{x-\mu}{\sigma}$$

ดังนั้น Z มีการแจกแจงเป็น Standard normal distribution เพราะ

$$f(Z) = \frac{e^{-\frac{1}{2}Z^2}}{\sqrt{2\pi}}, \mu = 0, \sigma = 1$$

สำหรับ Standard normal distribution เราสามารถหาพื้นที่ในช่วงต่าง ๆ ได้โดยดูจากตารางซึ่งคำนวณไว้แล้ว สำหรับพื้นที่ในช่วง $Z = \pm 1$, $Z = \pm 2$, $Z = \pm 3$ จะประมาณ .6827, .9545, และ .9973 ตามลำดับ

ในขณะเดียวกันกับพื้นที่ในช่วงหนึ่ง ๆ มีค่าคงที่ ความสูงหรือ Ordinate หรือ f (Z) ก็มีค่าคงที่ ซึ่งสามารถดูได้จากตารางซึ่งเขาคำนวณไว้เช่นกัน ความสูงที่จุด $\mu = 0$ จะประมาณ .3989



คือถ้า $X = \mu = 0, \quad \sigma = 1$ จะได้ว่า

$$f(x) = \frac{e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}}{\sigma\sqrt{2\pi}} = \frac{1}{\sqrt{2\pi}} \simeq .3989$$

$E(X)$ เมื่อ $f(x)$ เป็น normal $= \mu$

และ $V(X) = \sigma^2$

การแจกแจงปกตินี้ อาจจะใช้เป็นการแจกแจงประมาณ (approximate) ของการแจกแจงไบโนเมียลได้ ในกรณีที่ค่า p หรือ q ในการแจกแจงไบโนเมียลมีค่าน้อยใกล้ศูนย์ (0) และ n มีค่ามาก ๆ

$$Z = \frac{x - np}{\sqrt{npq}}$$ จะมีการแจกแจงใกล้การแจกแจงปกติ

7.4 คุณสมบัติของการแจกแจงปกติ

1. เป็นโค้งรูประฆังคว่ำ (bell shaped) ส่วนสูงของโค้งจะมากน้อยเพียงใด ขึ้นอยู่กับค่าของความแปรปรวน กล่าวคือถ้าความแปรปรวนมีค่าน้อย ส่วนสูงของโค้งจะมีมาก แต่ถ้าความแปรปรวนมีค่ามาก ส่วนสูงของโค้งจะน้อย

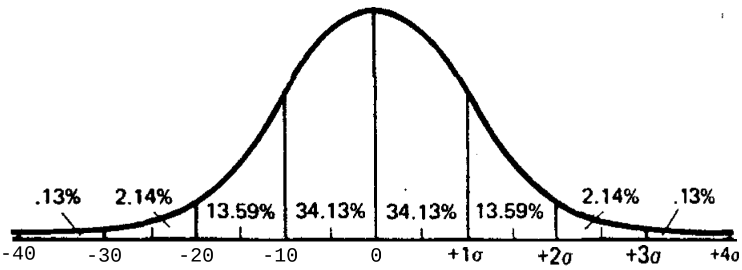
2. เป็นโค้งที่มีลักษณะสมมาตร (symmetry) คือลักษณะทางซ้ายและทางขวาของโค้งจะเหมือนกัน

3. เป็นโค้งที่มีจุดสูงสุดเพียงจุดเดียว (unimodal) และจุดสูงสุดนี้จะอยู่ตรงกึ่งกลางโค้ง ส่วนปลายของโค้งจะไม่จรดกับแกนนอน (horizontal axis)

4. เป็นโค้งที่มีค่าเฉลี่ย มัธยฐาน และฐานนิยมอยู่ที่จุดเดียวกัน นั่นคือ ค่าเฉลี่ย มัธยฐาน และฐานนิยม จะมีค่าเท่ากัน

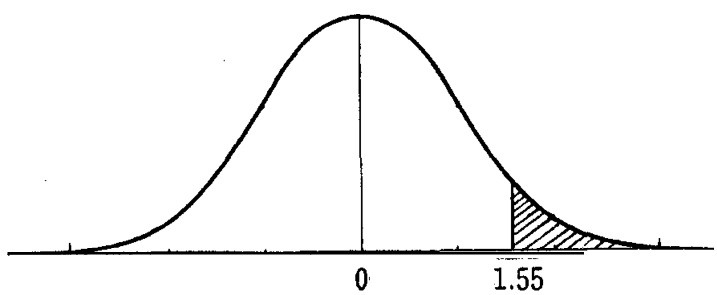
7.5 การหาพื้นที่ใต้โค้งปกติ

โดยวิธีการทางแคลคูลัส ทำให้นักคณิตศาสตร์สามารถที่จะหาพื้นที่ในส่วนต่าง ๆ ที่อยู่ใต้โค้งได้ โดยสัดส่วนของพื้นที่ใต้โค้งปกติจะเป็นดังรูป



ในทางปฏิบัติ สามารถคำนวณหาความน่าจะเป็นระหว่างจุดหนึ่งถึงอีกจุดหนึ่งได้ โดยความน่าจะเป็นจะมีค่าเท่ากับพื้นที่ภายใต้โค้งระหว่างจุดนั้น ๆ และถ้ากำหนดค่า Z มาให้ก็จะสามารถหาพื้นที่ใต้โค้งได้จากตารางในภาคผนวก วิธีการเปิดตารางหาค่าพื้นที่ใต้โค้งมีดังนี้

1. หาพื้นที่ใต้โค้งจากจุด $Z = 1.55$ ถึง ∞



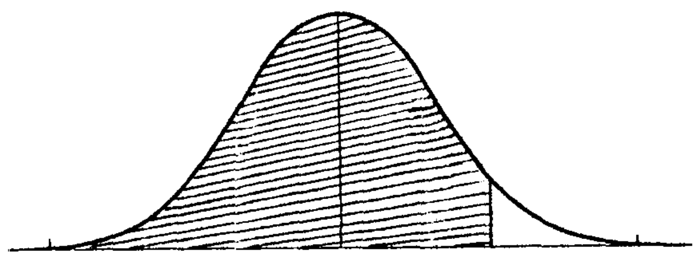
ถ้าต้องการหาค่า P $(Z > 1.55)$ จากความรู้เดิมจะได้ว่าพื้นที่ใต้โค้งตั้งแต่ $Z=0$ ถึง $Z = \infty$ มีค่าเท่ากับครึ่งหนึ่งของพื้นที่ทั้งหมดหรือมีค่าเท่ากับ .50

ดังนั้นพื้นที่ใต้โค้งตั้งแต่ $Z = 1.55$ ถึง ∞ = พื้นที่ใต้โค้งตั้งแต่ $Z = 0$ ถึง $Z=\infty$

— พื้นที่ใต้โค้งตั้งแต่ $Z = 0$ ถึง $Z=1.55$

$$
\begin{aligned}
\therefore P(Z > 1.55) &= .50 - P(0 < Z < 1.55) \\
&= .50 - .4394 \\
&= .0216
\end{aligned}
$$

2. หาพื้นที่ใต้โค้งจากจุด $Z = -\infty$ ถึง 1.55

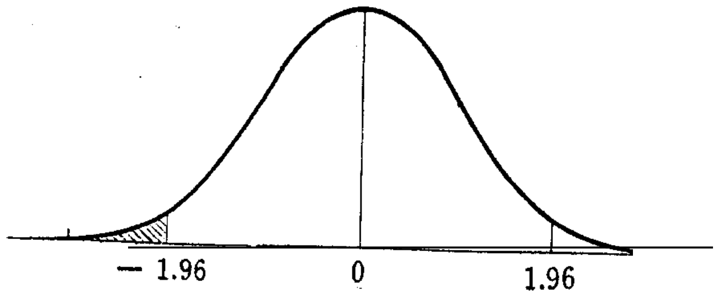


จากรูปต้องการหาพื้นที่ใต้โค้งตั้งแต่ $Z = -\infty$ ถึง 1.55 ซึ่งสามารถคำนวณได้ดังนี้

$$
\begin{aligned}
P(Z < 1.55) &= .50 + P(0 < Z < 1.55) \\
&= .50 + .4394 \\
&= .9394
\end{aligned}
$$

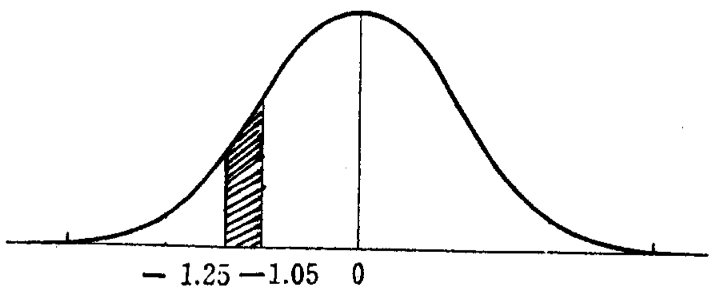
3. หาพื้นที่จากจุด $Z = -1.96$ ถึง $-\infty$

ในกรณีที่ Z มีค่าเป็นลบ กล่าวคือ Z จะอยู่ทางซ้ายมือของ 0 ก็สามารถหาค่าพื้นที่ใต้โค้งได้จากตารางเดิม ทั้งนี้เพราะพื้นที่ใต้โค้งทางซ้ายจะเท่ากับพื้น ที่ใต้โค้งทางขวา ดังตัวอย่าง



$$P(Z < -1.96) = P(Z > 1.96)$$
$$= .50 - P(0 < Z < 1.96)$$
$$= .50 - .4744$$
$$= .0256$$

4. หาพื้นที่ใต้โค้งจากจุด $Z = -1.05$ ถึง 1.25



เนื่องจาก $P(-1.25 < Z < -1.05) = P(1.05 < Z < 1.25)$

$$= P(0 < Z < 1.25) - P(0 < Z < 1.05)$$
$$= .3944 - .3749$$
$$= .0195$$

7.6 ตัวอย่างการคำนวณที่เกี่ยวกับการแจกแจงปกติ

ตัวอย่าง 2 นักศึกษารามคำแหง 1000 คนมีความสูงเฉลี่ย 150 ซม. ความเบี่ยงเบนมาตรฐานของความสูง 20 ซม. ถ้าการแจกแจงความสูงเป็นการแจกแจงปกติ อยากทราบว่ามีนักศึกษากี่คนที่มีความสูงระหว่าง 145 กับ 155 ซม.

145 ทำเป็นคะแนนมาตรฐานได้เท่ากับ $\frac{145-150}{25} = -.2$

155 ทำเป็นคะแนนมาตรฐานได้เท่ากับ $\frac{155-150}{25} = .2$

พื้นที่ภายใต้โค้งปรกติระหว่าง $z = -.2$ กับ $z = .2$ ประมาณ .1586 (ดูจากตาราง)

ดังนั้นนักศึกษาที่สูงระหว่าง 145 กับ 155 ซม. มี $1000 \times .1586$

$\approx$ 159 คน

7.7 ตัวอย่างการใช้การแจกแจงปกติ ประมาณการแจกแจงไบโนเมียล

ตัวอย่าง 3 โยนเหรียญ 10 ครั้ง จงหาความน่าจะเป็นที่จะออกหัวระหว่าง 2 ถึง 6 ครั้ง

เรารู้ว่าถ้านิยามตัวแปรสุ่ม X ว่าเป็นจำนวนครั้งที่ออกหัวจากการโยนเหรียญ n ครั้ง x จะมี f (x) เป็น Binomial function ซึ่งมี $E(X) = \mu = np$ และ $V(X) = \sigma^2 = npq$ ในที่นี้ $\mu = (10)\left(\frac{1}{2}\right) = 5$

$$\sigma = \sqrt{(10)\left(\frac{1}{2}\right)\left(\frac{1}{2}\right)} = 1.58$$

ในการคำนวณโดยใช้การแจกแจงปกติซึ่งเป็นการแจกแจงแบบต่อเนื่อง เราต้องทำให้ข้อมูลซึ่งเป็นข้อมูลที่ไม่ต่อเนื่องให้เป็นข้อมูลที่ต่อเนื่องด้วย จึงจะประมาณได้ใกล้เคียงขึ้น

นั่นคือ ให้ 1.5 แทน 2 และ 6.5 แทน 6

คะแนนมาตรฐานของ 1.5 คือ $z = \frac{1.5-5}{1.58} \approx -2.215$

คะแนนมาตรฐานของ 6.5 คือ $z = \frac{6.5-5}{1.58} \approx .9494$

พื้นที่ใต้โค้งระหว่าง $z = -2.215$ ถึง $z = .9494$ ประมาณ .8157 (ดูจากตาราง)

ดังนั้นความน่าจะเป็นที่จะได้หัว 2 ถึง 6 ครั้งจากการโยนเหรียญ 10 ครั้ง เป็น .8157

ถ้าใช้การแจกแจงแบบไบโนเมียล ก็จะได้ว่า

$$P(2 \leqslant X \leqslant 6) = f(2) + f(3) + f(4) + f(5) + f(6)$$

$$= \binom{10}{2}\left(\frac{1}{2}\right)^{10} + \binom{10}{3}\left(\frac{1}{2}\right)^{10} + \binom{10}{4}\left(\frac{1}{2}\right)^{10} + \binom{10}{5}\left(\frac{1}{2}\right)^{10} + \binom{10}{6}\left(\frac{1}{2}\right)^{10}$$

$$= \left(\frac{1}{2}\right)^{10}\left[\binom{10}{2} + \binom{10}{3} + \binom{10}{4} + \binom{10}{5} + \binom{10}{6}\right]$$

$$= \frac{1}{1024}(837)$$

$$\approx .8174$$

ซึ่งเป็นค่าที่แท้จริง จะเห็นได้ว่าไม่ต่างกันมากนัก

แบบฝึกหัด 7

1. จงแสดงว่าถ้าตัวแปรสุ่ม X มีฟังก์ชั่นความน่าจะเป็นแบบ Normal มีค่าเฉลี่ยเป็น μ ความแปรปรวนเป็น $\sigma^2 \neq 0$ แล้ว ตัวแปรสุ่ม $W = (X-\mu)/\sigma$ จะมีฟังก์ชั่นความน่าจะเป็นแบบ Normal มีค่าเฉลี่ยเป็น 0 ความแปรปรวนเป็น1
2. ถ้า X เป็นตัวแปรสุ่มที่มีฟังก์ชั่นความน่าจะเป็นแบบ Normal มีค่าเฉลี่ย 75 ความเบี่ยงเบนมาตรฐาน 10 จงหาความน่าจะเป็นที่ตัวแปรสุ่ม X จะมีค่าอยู่ระหว่าง 80 ถึง 90
3. ถ้า Y เป็นตัวแปรสุ่มที่มีฟังก์ชั่นความน่าจะเป็นแบบ Normal มีค่าเฉลี่ย μ ความแปรปรวน σ^2 จงหาค่า b ซึ่งความน่าจะเป็นที่ $(X-\mu) / \sigma$ จะน้อยกว่า b และมากกว่า$-$b เท่ากับ 0.9
4. จงหาพื้นที่ภายใต้โค้งปกติตั้งแต่ $Z_1 = 0.25$ ถึง $Z_2 = 0.95$
5. ถ้าความสูงของนักศึกษารามคำแหง 10,000 คน มีการแจกแจงปกติ มีค่าเฉลี่ย 150 ซ.ม. ความเบี่ยงเบนมาตรฐาน 15 ซ.ม. จะมีนักศึกษาที่มีความสูงอยู่ระหว่าง 110—155 ซ.ม. อยู่กี่คน และจะมีนักศึกษาที่มีความสูงเกิน 180 ซ.ม. กี่คน
6. ในการผลิตกระดาษถ้าหากกระดาษมีความหนาอยู่ระหว่าง 0.503 ม.ม. ถึง 0.642 ม.ม. ถือว่าใช้ได้ ถ้าหากว่าหนามากหรือน้อยกว่านี้ถือว่าไม่ได้มาตรฐาน ถ้าผลิตกระดาษออกมา 2500 แผ่น มีค่าเฉลี่ยความหนา 0.595 ม.ม. ความเบี่ยงเบนมาตรฐาน 0.05 ม.ม. จะมีกระดาษที่ไม่ได้มาตรฐานอยู่กี่แผ่น
7. จงหาความน่าจะเป็นที่จะได้ก้อย ระหว่าง 2—6 ครั้ง ในการโยนเหรียญ 20 ครั้ง โดยใช้การแจกแจงปกติเป็นค่าประมาณการแจกแจงไบโนเมียล

8. ในการสอบวิชา MR 371 มีข้อสอบ 10 ข้อ ให้คะแนนแบบ 0–1 ปรากฏว่าคะแนนเฉลี่ยเท่ากับ 8.0 ความเบี่ยงเบนมาตรฐานเท่ากับ 2.4 ถ้าหาก 40% ล่างจะต้องได้ F อยากทราบว่าคะแนนต่ำกว่าเท่าไรจึงจะได้ F

9. จงแสดงว่าการแจกแจงปกติมาตรฐานมีคุณสมบัติ

ก. สมมาตรกับแกน Y

ข. ไม่แปรไปกับ degrees of freedom

ค. ordinate ที่จุด mean เท่ากับ .3989

ง. พื้นที่ภายใต้โค้งเป็น 1