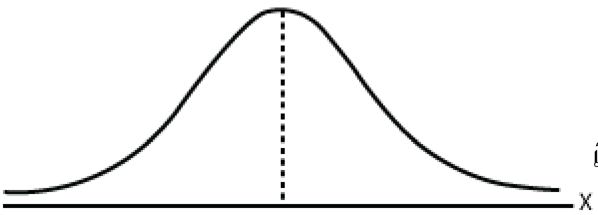
6. การแจกแจงแบบปกติ (Normal Distribution)

โดยทั่วไป การแจกแจงของตัวแปรที่ต่อเนื่องกันที่สำคัญที่สุด ได้แก่ การแจกแจงปกติ และทฤษฎีต่างๆในทางสถิติมักตั้งอยู่บนพื้นฐาน ของการแจกแจงแบบนี้ ลักษณะกราฟของการแจกแจงปกติ เรียกว่า เส้นโค้งปกติมีลักษณะเป็นรูประฆังที่สมมาตร มีความโค้งพอดี



เส้นโค้งปกติ The normal curve

ตัวแปรสุ่ม X ที่มีลักษณะเป็นการแจกแจงปกตินี้ เรียกว่า ตัวแปรสุ่มปกติ (Normal Random Variable)

มีรูปแบบคือ

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

โดยที่ μ,σ,π และ e ต่างเป็นค่าคงที่ (Constant)

เมื่อ $\pi = 3.14159$, $e = 2.71828$

ถ้า X เป็นตัวแปรสุ่มที่มีการแจกแจงปกติ มีมัชฌิมเลขคณิตเท่ากับ μ ความแปรปรวนเท่ากับ σ^2 เราใช้สัญลักษณ์ $X \sim N(\mu,\sigma^2)$ แทน เราสามารถหาค่าของ $f(x)$ ได้ทุกๆค่าของ X ที่เป็นเลขจำนวนจริง

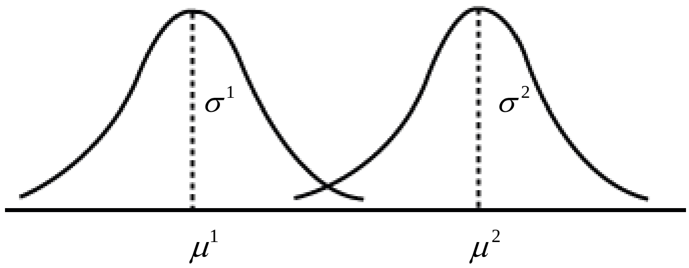
Example ถ้า X เป็นการแจกแจงปกติเป็น มีมัชฌิมเลขคณิตเป็น 50 และส่วนเบี่ยงเบนมาตรฐานเป็น 5 นั่นคือ X ~ N(50,25)

และ $$f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\cdot e^{-\frac{1}{2}\left(\frac{X-\mu}{\sigma}\right)^2} \quad ; \ -\infty < X < \infty$$

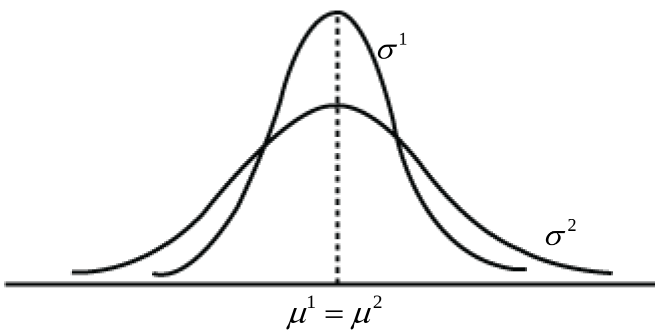
$$=\frac{1}{5\sqrt{2\pi}}\cdot e^{-\frac{1}{2}\left(\frac{X-50}{5}\right)^2} \quad ; \ -\infty < X < \infty$$

จากสมการนี้เราสามารถคำนวณค่า $f(x)$ ได้ทุกค่าของ X ที่เป็นเลขจำนวนจริง และนำมาวาดรูปเป็นรูปกราฟเส้นโค้งได้

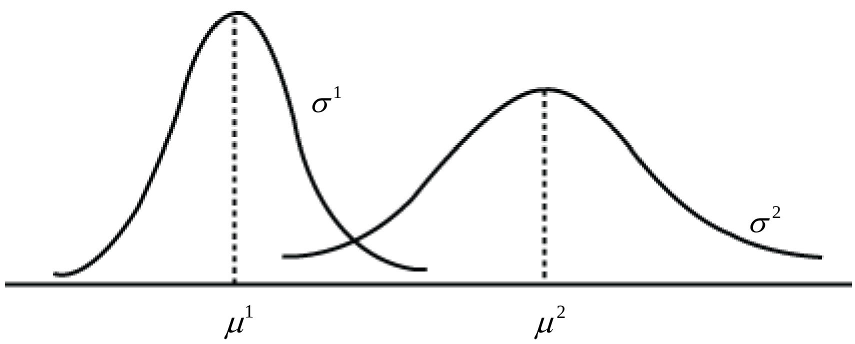
จากรูปต่อไปนี้แสดงถึงเส้นโค้งในลักษณะต่างๆ เช่น



เส้นโค้งปกติที่มี $\mu_1 \neq \mu_2$ และ $\sigma_1 = \sigma_2$ จะมีลักษณะเหมือนกันทุกประการ ผิดแต่ตำแหน่งของจุดยอดของรูปทั้ง สองรูปอยู่ต่างกัน



เส้นโค้งปกติที่มี $\mu_1 = \mu_2$ และ $\sigma_1 < \sigma_2$ ลักษณะของเส้นทั้งสองจะแตกต่างกัน เส้นโค้งปกติที่มีส่วนเบี่ยงเบนมาตรฐานมากกว่าจะมีลักษณะที่แบน และลาดต่ำกว่าเส้นโค้งปกติที่มีส่วนเบี่ยงเบนมาตรฐานน้อย แต่ตำแหน่งของจุดยอดอยู่ที่เดียวกัน



เส้นโค้งปกติที่มี $\mu_1 \neq \mu_2$ และ $\sigma_1 \neq \sigma_2$ ตำแหน่งของจุดยอดอยู่ต่างกัน และความโค้งของเส้นปกติของทั้งสองรูปก็ต่างกัน

จะเห็นได้ว่า เส้นโค้งปกติจะมีลักษณะโด่งมาก หรือโด่งน้อย ขึ้นอยู่กับส่วนเบี่ยงเบนมาตรฐาน ถ้าส่วนเบี่ยงเบนมาตรฐานใหญ่ เส้นโค้งจะมีลักษณะโด่งน้อยกว่า เส้นโค้งที่มีมาตรฐานเล็ก

6.1 คุณสมบัติของเส้นโค้งปกติ

1. เส้นโค้งจะมีความถี่สูงอยู่ที่ค่ามัชฌิมเลขคณิต μ และจะกระจายลดน้อยลงไปทางค่าสูง และค่าต่ำอย่างสม่ำเสมอ

2. เส้นโค้งจะมีลักษณะสมมาตรที่ คงที่ μ ค่าออร์ดิเนตของจุด $\mathrm{X} = \mu \pm \sigma$ เมื่อ k เป็นค่าคงที่ จะเท่ากันเช่น ความสูงของเส้นโค้งที่จุด $\mathrm{X} = \mu + \sigma$ จะเท่ากับความสูงของเส้นโค้งที่จุด $\mathrm{X} = \mu - \sigma$ จากความสมมาตรของเส้นโค้งนี้เองจะได้ค่า มัชฌิมเลขคณิต มัธยฐาน และฐานนิยม มีค่าเท่ากับ (mean=median=mode)

3. เส้นโค้งปกติมีลักษณะเป็นรูประฆัง ปลายเส้นโค้งทั้งสองจะค่อยๆลาดลงสู่แกนนอน แต่ไม่มีโอกาสจะสัมผัสกับแกนนอน พื้นที่ภายใต้เส้นโค้งปกติที่มีคะแนนมาตรฐานมากกว่า $\mu + 3$ หรือมีคะแนนมาตรฐานน้อยกว่า $\mu - 3$ จะมีค่าน้อยมาก

6.2 ทฤษฎีที่สำคัญเกี่ยวกับการแจกแจงปกติ

ทฤษฎีที่เกี่ยวข้องกับการแจกแจงปกติได้แก่

Theorem1 : ถ้า X มีการแจกแจงปกติ ที่มีมัชฌิมเลขคณิตเป็น μ และส่วนเบี่ยงเบนมาตรฐานเป็น σ แล้ว $\mathrm{Y} = c\mathrm{X} + d$ ก็จะมีการแจกแจงปกติที่มีมัชฌิมเลขคณิตเป็น $c\mu + d$ และส่วนเบี่ยงเบนมาตรฐาน $c\sigma$

Theorem2 : ถ้า $X_1, X_2, ..., X_N$ เป็นตัวแปรสุ่มที่เป็นอิสระ ที่มีมัชฌิมเลขคณิตเป็น μ และส่วนเบี่ยงเบนมาตรฐานเป็น $\sigma_i (i = 1,2,...,N)$ แล้ว $S = X_1 + X_2 + ... + X_N$ ก็จะมีการแจกแจงปกติที่มีมัชฌิมเลขคณิตเป็น μ_1 และความแปรปรวน σ^2

Theorem3 : ถ้า X_1, X_2 เป็นตัวแปรที่เป็นอิสระต่อกันและมีการแจกแจงปกติ โดยมีมัชฌิมเลขคณิตเป็น μ_1 และส่วนเบี่ยงเบนมาตรฐานเป็น $\sigma_i (i = 1,2)$ แล้ว $D = X_1 - X_2$ ก็จะมีการแจกแจงปกติที่มีมัชฌิมเลขคณิตเป็น $\mu_1 - \mu_2$ และความแปรปรวนเป็น $\sigma_1^{\ 2} + \sigma_2^{\ 2}$

6.3 พื้นที่ใต้เส้นโค้งปกติ

เนื่องจาก ผลรวมของการทดลองสุ่มต้องเท่ากับหนึ่ง และพื้นที่ใต้เส้นโค้งปกติทั้งหมดเหนือแกน X มีค่าเท่ากับหนึ่ง ดังนั้นในการหาความน่าจะเป็นของตัวแปรสุ่มปกติใดอาจใช้พื้นที่ภายใต้เส้นโค้งปกติเป็นตัวแทนของตัวแปรสุ่มนั้น

ดังนั้น พื้นที่ใต้เส้นโค้งย่อมขึ้นอยู่กับค่าของมัชฌิมเลขคณิต และส่วนเบียงเบนมาตรฐานของประชากรหมู่นั้นๆ สำหรับตัวแปรสุ่ม X ซึ่งเป็นแบบ $\mathrm{N} \sim (\mu, \sigma^2)$ จะเห็นได้ว่า

$$\mathrm{P}(\mathrm{X}_1 \le \mathrm{X} \le \mathrm{X}_2) = \int_{\mathrm{X}_2}^{\mathrm{X}_1} f(x)dx = \int_{\mathrm{X}_2}^{\mathrm{X}_1} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

เพื่อให้ง่ายในการหาพื้นที่ใต้เส้นโค้ง สำหรับทุกๆค่าของ μ และ σ ที่เปลี่ยนไปจึงมีการสร้างตารางมาตรฐานของพื้นที่ใต้เส้นปกติ (ดูตาราง Z) ซึ่งแสดงถึงพื้นที่ใต้เส้นโค้งระหว่าง $Z = 0$ ถึง $Z = 1$ โดยที่ Z เป็นคะแนนมาตรฐานที่มีการแจกแจงปกติ ซึ่งมีค่าเฉลี่ย (mean) เท่ากับ 0 และส่วนเบี่ยงเบนมาตรฐานเท่ากับ 1 ตารางนี้จะใช้ได้กับข้อมูลทั่วๆไป ดังนั้นก่อนที่จะใช้ตารางนี้ จึงจำเป็นต้องเปลี่ยนข้อมูลดิบที่ได้มาให้เป็นคะแนนมาตรฐานเสียก่อน โดยใช้สูตร

$$Z = \frac{\text{ส่วนเบี่ยงเบนจากมัชฌิมเลขคณิต}}{\text{ส่วนเบี่ยงเบนมาตรฐาน}}$$

นั่นคือ $Z = \frac{X - \mu}{\sigma_x} = \frac{Y - \mu_y}{\sigma_y} = \frac{P - \mu_p}{\sigma_p}$

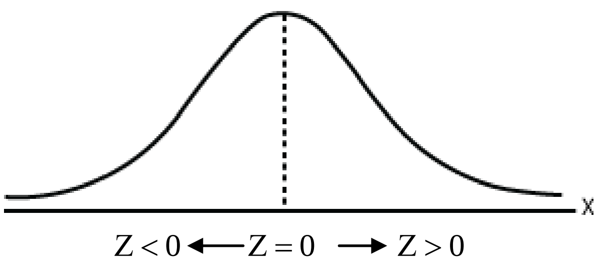
ซึ่งค่าเฉลี่ย (mean) และส่วนเบี่ยงเบนมาตรฐาน (standard deviation) ของ Z อาจหาได้ดังนี้ คือ

$$\begin{aligned}
\mu = E(Z) &= E\left(\frac{X - \mu}{\sigma}\right) = \frac{1}{\sigma} \cdot E[X - \mu] \\
&= \frac{1}{\sigma}[E(X) - \mu] \\
&= \frac{1}{\sigma}[\mu - \mu] = 0
\end{aligned}$$

และ

$$\begin{aligned}
\sigma_z^2 = Var(Z) &= Var\frac{(X - \mu)}{\sigma} = \frac{1}{\sigma^2}Var[X - \mu] \\
&= \frac{1}{\sigma^2}[Var(X) + Var(\mu)] \\
&= \frac{\sigma^2}{\sigma^2} = 1
\end{aligned}$$

<u>Definition</u> การแจกแจงปกติแบบมาตรฐาน (Standard Normal Distribution) คือ การแจกแจงของตัวแปรสุ่มปกติที่มีมัชฌิมเลขคณิต μ เป็นศูนย์ และส่วนเบี่ยงเบนมาตรฐาน σ เป็นหนึ่ง ใช้สัญลักษณ์ $Z \sim N(0,1)$



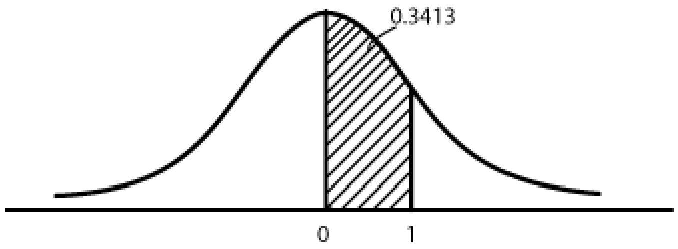
ที่จุด $X = \mu$, $Z = \frac{X - \mu}{\sigma} = 0$

ที่จุด $X > \mu$, $Z > 0$

ที่จุด $X < \mu$, $Z < 0$

โดยที่ส่วนโค้งสมมาตร (symmetry) กันที่จุด $X = \mu$ ฉะนั้น จึงไม่จำเป็นต้องสร้างตารางที่แสดงค่า Z เป็นลบไว้ เพราะข้างที่ Z เป็นบวกก็ใช้ได้สำหรับข้างที่ Z เป็นลบด้วย

ตารางที่สร้างขึ้น จะบอกค่าพื้นที่ใต้เส้นโค้งนับจาดจุดมัชฌิมเลขคณิต มายังจุดที่มีค่า Z ตรงที่กำหนดให้ สมมติว่าที่จุด Z มี $Z=1$ จากตารางอ่านพื้นที่ได้ $=0.3413$ พื้นที่นี้คือ พื้นที่ล้อมรอบด้วยเส้นโค้งแกนนอนและเส้นตรง $Z=1$ ดังรูป หรือ $P(0<Z<1)=0.3413$ นั่นเอง



นั่นคือ พื้นที่ทั้งหมดทางขวามือ $=$ พื้นที่ทั้งหมดทางซ้ายมือ $=0.5$

ฉะนั้น $P(Z\le 1) = 0.5+P(0<Z<1)$

$= 0.5+0.3413$

$= 0.8413$

และ $P(Z\ge 1) = 0.5-P(0<Z<1)$

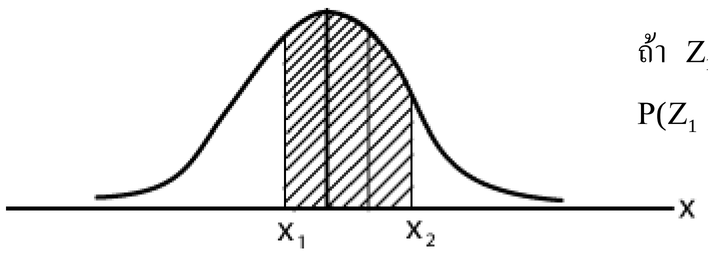
$= 0.5-0.3413$

$= 0.1587$

ความน่าจะเป็นของ X ที่มีค่าอยู่ระหว่าง $X=X_1$ และ $X=X_2$ โดยที่ $X_1<\mu$ และ $X_2>\mu$ จะมีค่าเท่ากับความน่าจะเป็นของ Z ที่ตกอยู่ระหว่าง $Z=Z_1$ และ $Z=Z_2$

โดยที่ $Z_1=\frac{X_1-\mu}{\sigma}$ และ $Z_2=\frac{X_2-\mu}{\sigma}$

$$
\begin{aligned}
P(X_1<X<X_2) &= P\left[\frac{X_1-\mu}{\sigma}<\frac{X-\mu}{\sigma}<\frac{X_2-\mu}{\sigma}\right] \\
&= P(Z_1<Z<Z_2) \\
&= P(Z_1<Z<0)+P(0<Z<Z_2)
\end{aligned}
$$



ถ้า $Z_1=-Z_2$

$P(Z_1<Z<Z_2)=2P(0<Z<Z_2)$

Example ถ้า X มีการแจกแจงปกติ ที่มีมัชฌิมเลขคณิตเป็น 3 และความแปรปรวนเป็น 4 จงหาความน่าจะเป็นที่ X จะมีค่าอยู่ระหว่าง 3 และ 5

วิธีทำ เพราะว่า $X \sim N(3,4)$

ต้องการหา $P(3 < X < 5)$

เมื่อ $X_1 = 3$, $Z_1 = \dfrac{X_1 - \mu}{\sigma}$, $= \dfrac{3-3}{2} = 0$

เมื่อ $X_2 = 5$, $Z_2 = \dfrac{X_2 - \mu}{\sigma}$, $= \dfrac{5-3}{2} = 1$

ดังนั้น $P(3 < X < 5) = P(0 < Z < 1)$

$=$ พื้นที่ระหว่าง $Z = 0$ และ $Z = 1$

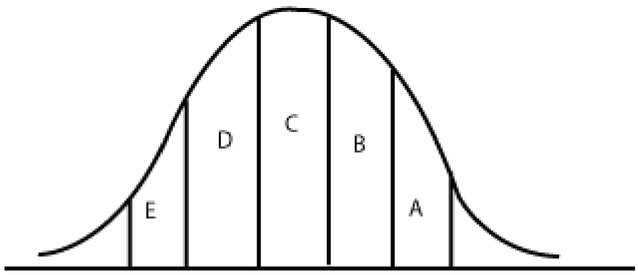
$= 0.3413$

6.4 การกำหนดระดับ (Grading)

ในการพิจารณาเปรียบเทียบคะแนนผลการสอบของนิสิตแต่ละคนในแต่ละกลุ่ม โดยการจัดตำแหน่งของคะแนน กำหนดกลุ่มของคะแนนเป็นพวกๆ และให้ระดับคะแนนในกลุ่มนั้น เป็นระดับคะแนน A,B,C,D หรือ E หรืออาจแบ่งย่อยเป็นระดับ A,B^+,B,C^+,C,D^+,D หรือ E

สมมติให้การกระจายของคะแนนเป็นแบบโค้งปกติ (ที่มีการกระจายเป็นรูประฆัง) และมีการให้ระดับคะแนนเป็น A,B,C,D หรือ E ขั้นแรกแบ่งการแจกแจงของโค้งปกตินี้เป็น 5 ส่วนเท่าๆกันตามแกนนอน (X) โดยที่ X แทนคะแนน โดยให้คะแนนสูงสุดของระดับ A มากกว่าค่าเฉลี่ยอยู่ 3 เท่าของส่วนเบี่ยงเบนมาตรฐาน $(\mu + 3\sigma)$ และคะแนนต่ำสุดของระดับ E น้อยกว่า ค่าเฉลี่ยอยู่ 3 เท่าของส่วนเบี่ยงเบนมาตรฐาน $(\mu - 3\sigma)$

นั่นคือ 99.74 % ของคะแนนทั้งหมดจะอยู่ภายใต้ในช่อง $\mu - 3\sigma$ และ $\mu + 3\sigma$ (คะแนนสูงสุดของระดับ A ถึงคะแนนต่ำสุดของระดับ E จะกลุ่มคะแนนทั้งหมดไว้ถึง 99.74 %) ดังรูป



$-3\sigma \quad -2\sigma \quad -1\sigma \quad 1\sigma \quad 2\sigma \quad 3\sigma$

พิสัยของแต่ละเกรด (คะแนนสูงสุด-คะแนนต่ำสุด) จะเท่ากับ 1.2 เท่าของส่วนเบี่ยงเบนมาตรฐาน (6 เท่าของส่วนเบี่ยงเบนมาตรฐาน ถูกแบ่งออกเป็น 5 ส่วนเท่าๆ กันตามแนวแกน X) นั่นคือ

ระดับ A	คะแนนสูงสุด	= $\mu + 3\sigma$
	คะแนนต่ำสุด	= $\mu + 1.8\sigma$
ระดับ B	คะแนนสูงสุด	= $\mu + 1.8\sigma$
	คะแนนต่ำสุด	= $\mu + 0.6\sigma$
ระดับ C	คะแนนสูงสุด	= $\mu + 0.6\sigma$
	คะแนนต่ำสุด	= $\mu - 0.6\sigma$
ระดับ D	คะแนนสูงสุด	= $\mu - 0.6\sigma$
	คะแนนต่ำสุด	= $\mu - 1.8\sigma$
ระดับ E	คะแนนสูงสุด	= $\mu - 1.8\sigma$
	คะแนนต่ำสุด	= $\mu - 3.0\sigma$

Example สมมติว่าคะแนนของการสอบวิชาหนึ่งมีการแจกแจงเป็นโค้งปกติที่มีมัชฌิมเลขคณิตเป็น 65 และส่วนเบี่ยงเบนมาตรฐานเป็น 15 จงหาคะแนนสูงสุดและคะแนนต่ำสุดของผู้ที่ได้คะแนนระดับ A,B,C,D หรือ E

ระดับ A คะแนนสูงสุด = $65+(3\times15)=110$

คะแนนต่ำสุด = $65+(1.8\times15)=92$

ระดับ B คะแนนสูงสุด = $65+(1.8\times15)=92$

คะแนนต่ำสุด = $65+(0.6\times15)=74$

ระดับ C คะแนนสูงสุด = $65+(0.6\times15)=74$

คะแนนต่ำสุด = $65-(0.6\times15)=56$

ระดับ D คะแนนสูงสุด = $65-(0.6\times15)=56$

คะแนนต่ำสุด = $65-(1.8\times15)=38$

ระดับ E คะแนนสูงสุด = $65-(1.8\times15)=38$

คะแนนต่ำสุด = $65-(3\times15)=20$

ในทางปฏิบัติ การกำหนดระดับของคะแนนอาจทำได้หลายวิธี เช่น ควรจะกำหนดร้อยละของผู้ที่ได้ระดับต่างๆ กันดังนี้

ระดับ A 5%

ระดับ B^+ 10%

ระดับ B 15%

ระดับ C^+ 20%

ระดับ C 30%

ระดับ D^+ 10%

ระดับ D 8%

ระดับ E 2%

นั่นคือจะได้

ระดับ A - คะแนนต่ำสุด $=\mu+1.65\sigma$

ระดับ B^+ - คะแนนต่ำสุด $=\mu+1.04\sigma$

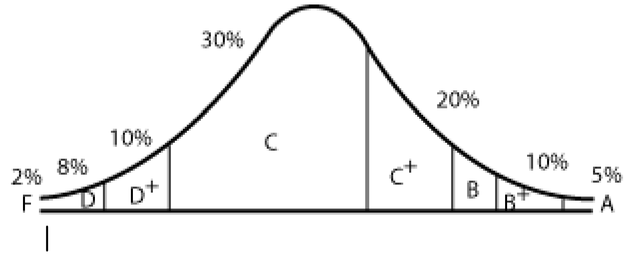
ระดับ B - คะแนนต่ำสุด $=\mu+0.52\sigma$

ระดับ C^+ - คะแนนต่ำสุด $=\mu+0\sigma$

ระดับ C - คะแนนต่ำสุด $=\mu-0.84\sigma$

ระดับ D^+ - คะแนนต่ำสุด $=\mu-1.28\sigma$

ระดับ D - คะแนนต่ำสุด $=\mu-2.05\sigma$



-2.05σ -1.28σ -0.84σ 0σ 0.52σ 1.04σ 1.65σ

6.5 การประมาณค่าของการแจกแจงทวินามโดยใช้การแจกแจงปกติ
(Normal Approximation to the Binomial Distribution)

เราทราบมาแล้วว่า การแจกแจงทวินามที่มี $\rho \to 0$ หรือ $q \to 0$ จะได้รูปเส้นโค้งเบ้เป็นทางใดทางหนึ่งถ้า $\rho \to 0$ ละจำนวนการทดลองมากพอ จะได้รูปเส้นโค้งที่มีลักษณะคล้ายคลึงกับการแจกแจงพัวซองเราจึงประมาณค่าของการแจกแจงทวินามโดยใช้การแจกแจงพัวซอง

ในกรณีที่การแจกแจงทวินาม มี $p \to \frac{1}{2}$ หรือ $q \to \frac{1}{2}$ เมื่อนำมาแสดงด้วยแท่งฮิสโตแกรมและเกลาเป็นเส้นโค้ง จะได้รูปเส้นที่ไม่เบ้ เป็นรูประฆังที่มีส่วนทั้งสองข้างสมมาตรกัน ดังนั้นเราอาจจะประมาณค่าของการแจกแจงทวินามด้วยการแจกแจงปกติได้ การประมาณค่าของการแจกแจงทวินามด้วยการแจกแจงปกติ จะใช้ได้ดีเมื่อ $p \to \frac{1}{2}$ หรือ q $\frac{1}{2}$ และ n มีค่ามาก และแม้ว่า n มีค่ามาก และแม้ว่า n จะไม่โตนัก แต่ถ้า p ไม่เข้าใกล้ 0 หรือ 1 จนเกินไป การประมาณค่าของการแจกแจงทวินามด้วยการแจกแจงปกติก็ยังให้ผลดีพอควร ประการสำคัญก็คือ ตัวแปรสุ่มของการแจกแจงทวินามเป็นชนิดไม่ต่อเนื่องกัน (Discrete Random Variable) เมื่อเราจะประมาณค่าด้วยการแจกแจงปกติ เราจึงต้องปรับค่าของตัวแปรสุ่มให้เป็นต่อเนื่องกัน ที่เรียกว่า continuity correction เสียก่อนโดยการพยายามขยายค่าออกไปให้ครอบคลุมค่าที่ต้องการ โดยทั่วไปเรามักจะนำค่า 0.5 ไปบวกหรือลบออกจากค่าที่ต้องการ ความถูกต้องในการใช้แจกแจงปกติประมาณค่าแจกแจงทวินามนั้นขึ้นอยู่กับค่าของ n เป็นสำคัญ เมื่อ n ยิ่งมีค่ามาก ค่าที่ประมาณได้จะยิ่งมีค่าใกล้เคียงกับค่าที่หาได้จากทวินามโดยตรงมากขึ้น แม้ว่า p จะไม่ใกล้กับ $\frac{1}{2}$ ซึ่งในการจะตัดสินใจว่าเราควรใช้การแจกแจงปกติ ประมาณค่าของความน่าจะเป็นที่ได้จากการแจกแจงทวินามหรือไม่นั้น เราอาจพิจารณาค่าของ np และ nq ถ้าค่าทั้งสองนี้เกิน 5 การประมาณค่าความน่าจะเป็นแบบทวินามโดยใช้การแจกแจงแบบปกติ ก็จะให้ผลดี

ตัวอย่างที่ 1 ความน่าจะเป็นที่นาย ก จะยิงปืนถูกเป้าในแต่ละนัดเป็น 0.4 ในการยิงปืนทั้งหมด 15 นัด จงหา

ก. ความน่าจะเป็นที่นาย ก. จะยิงปืนถูกเป้า 4 นัด

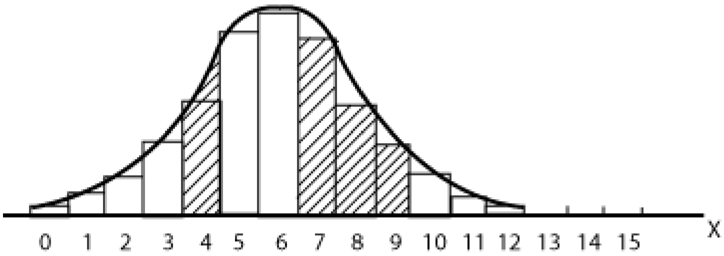
ข. ความน่าจะเป็นที่นาย ก. จะยิงปืนถูกเป้า ตั้งแต่ 7 นัด ถึง 9 นัด

วิธีทำ การทดลองนี้เป็นการทดลองแบบทวินามที่มี p = 0.4 , n=5

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

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

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



ให้ x เป็นจำนวนครั้งที่ยิงถูกเป้า

$\therefore$ X อาจจะมีค่าเป็น 0,1,2,…,15

พื้นที่ที่แรเงาคือพื้นที่ที่ต้องการหา

ก) $P(X = 4) \quad = {}^{15}C_4 (0.4)^4 (0.6)^{11}$

$$= 0.1268$$

ประมาณค่าโดยการใช้การแจกแจงปกติ

$$P(X = 4) \quad = {}^{15}C_4 (0.4)^4 (0.6)^{11} \simeq P(3.5 \le X \le 4.5)$$

$$= P\left(\frac{3.5-6}{1.9} \le \frac{X-\mu}{\sigma} \le \frac{4.5-6}{1.9}\right)$$

$$= P(-1.316 \le Z \le -0.789)$$

$$= P(-1.316 \le Z \le 0) - P(-0.789 \le Z \le 0)$$

$$= 0.4062 - 0.2852$$

$$= 0.1210$$

ข) $P(7 \le X \le 9) \ = \sum_{X=7}^{9} {}^{15}C_X (0.4)^X (0.6)^{15-X}$

$$= \sum_{X=0}^{9} {}^{15}C_X (0.4)^X (0.6)^{15-X} - \sum_{X=0}^{6} {}^{15}C_X (0.4)^X (0.6)^{15-X}$$

= 0.9662-0.6098

= 0.3546

ประมาณค่าโดยการใช้การแจกแจงปกติ

$$P(7 \le X \le 9) \ \simeq P\left(\frac{6.5-6}{1.9} \le \frac{X-\mu}{\sigma} \le \frac{9.5-6}{1.9}\right)$$

$$= P(0.263 \le Z \le 1.842)$$

$$= P(0 \le Z \le 1.842) - (0 \le Z \le 0.263)$$

= 0.4672-0.1036 = 0.3636

จะเห็นได้ว่า ทั้ง ก. และ ข. การประมาณค่าของการแจกแจงทวินามโดยใช้การแจกแจงปกติให้ผลลัพท์ใกล้เคียงกันมาก

ตัวอย่างที่ 2 ทอดลูกเต๋า 1 ลูก 120 ครั้ง จงหาความน่าจะเป็นที่ลูกเต๋าจะขึ้นหน้า 6 น้อยกว่า 15 ครั้ง

วิธีทำ ให้ x เป็นจำนวนครั้งที่ลูกเต๋าขึ้นหน้า 6

X มีการแจกแจงทวินาม ที่มี $\mathrm{p} = \frac{1}{6}, q = \frac{5}{6}, n = 120$

$$\mu = nq \quad = 120\left(\frac{1}{6}\right) \quad = 20$$

$$\sigma^2 = npq \quad = 120\left(\frac{1}{6}\right)\left(\frac{5}{6}\right) \quad = 16.67$$

$$\therefore \quad s = \sqrt{\sigma^2} \quad = \sqrt{16.67} \quad = 4.08$$

$$\therefore \quad \mathrm{P}(\mathrm{X} < 15) = \sum_{x=7}^{14} {}^{120}C_X \left(\frac{1}{6}\right)\mathrm{X}\left(\frac{5}{6}\right)120 - \mathrm{X}$$

$$= \mathrm{P}\left(\frac{\mathrm{X} - \mu}{\sigma} \leq \frac{14.5 - 20}{4.08}\right)$$

$$= \mathrm{P}(\mathrm{Z} \leq -1.35)$$

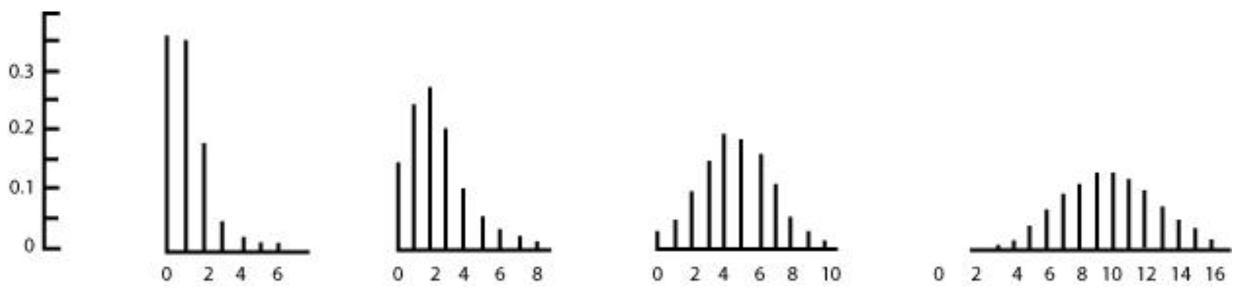
$$= 0.5 - \mathrm{P}(-1.35 \leq \mathrm{Z} \leq 0)$$

$$= 0.5 - 4.115 = 0.0885$$

6.6 การประมาณค่าของการแจกแจงพัวซองโดยใช้การแจกแจงปกติ

(Normal Approximation to the Poisson Distribution)

พิจารณาลักษณะของการแจกแจงพัวซองเมื่อ λ มีค่าต่างๆกัน



$\lambda = 1$ $\lambda = 2$ $\lambda = 5$ $\lambda = 10$

Poisson distribution for selected values of λ

ซึ่งจะเห็นว่า เมื่อ $\lambda \geq 10$ การแจกแจงปกตินั้นจะลักษณะเป็นเส้นโค้งปกติ และเนื่องจากในการแจกแจงแบบพัวซอง $p \to 0$ หรือ $q \to 0$ ดังนั้นการที่ $\lambda \geq 10$ ก็แสดงว่า n ต้องมีค่ามก และ เมื่อ n มีค่ามาก เราก็จะสามารถประมาณค่าของการแจกแจงพัวซองด้วยการแจกแจงปกติ เราจึงต้องปรับค่าเช่นเดียวกันกับการประมาณค่าความน่าจะเป็นแบบทวินามที่กล่าวแล้ว

ตัวอย่างที่ 1 บริษัทประกันชีวิตแห่งหนึ่งโดยเฉลี่ยแล้วในแต่ละวันจะมีผู้โทรศัพท์เข้ามา 400 ครั้งจงหา

(ก) ความน่าจะเป็นที่จะมีผู้โทรศัพท์เข้ามา 450 ครั้ง ในแต่ละวัน

(ข) ความน่าจะเป็นที่จะมีผู้โทรศัพท์ไม่น้อยกว่า 450 ครั้งในแต่ละวัน

(ค) ความน่าจะเป็นที่จะมีผู้โทรศัพท์ไม่เกิน 380 ครั้งในแต่ละวัน

วิธีทำ ให้ X เป็นจำนวนครั้งที่จะมีผู้โทรศัพท์เข้ามาในแต่ละวัน

$X \sim P(400,400)$

$\therefore \quad \mu = 400, \sigma = 20$

ก) $P(X = 450)$

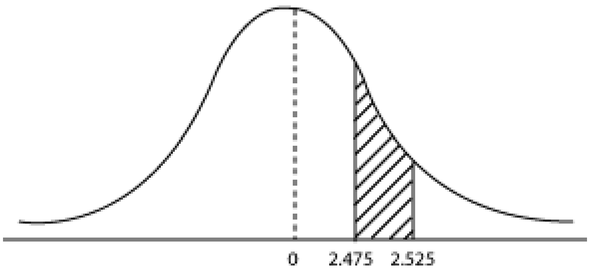
$= P(449.5 \le X \le 450.5)$

$= P\left(\frac{449.5 - 400}{20} \le \frac{X - \mu}{\sigma} \le \frac{450.5 - 400}{20}\right)$

$= P(2.475 \le Z \le 2.525)$

$= P(0 \le Z \le 2.525) - P(0 \le Z \le 2.475)$

= 0.4942-0.4933=0.009



ข) $P(X \ge 450)$

$= P(X \le 449.5)$

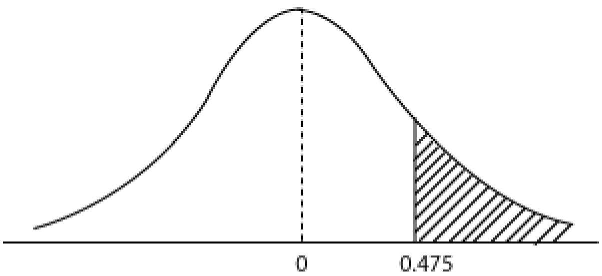
$= P\left(\frac{X - \mu}{\sigma} \ge \frac{449.5 - 400}{20}\right)$

$= P(Z \ge 2.475)$

= 0.5-0.4933

= 0.4942-0.4933 =0.009

= 0.0067



ค) $P(X \le 380)$

$= P(X \le 380.5)$

$= P\left(\frac{X - \mu}{\sigma} \le \frac{380.5 - 400}{20}\right)$

$= P(Z \le 0.975)$

$= 0.5 - P(-0.975 \le Z \le 0)$

= 0.5-0.33525

= 0.16475