

## 7. การแจกแจงของตัวอย่างสุ่ม

1. | การแจกแจงของ  $x$
2. | การแจกแจงของ  $p$
3. | แบบฝึกหัด

### การแจกแจงของตัวอย่างสุ่ม

เมื่อเรากำหนดว่าจะสุ่มจากประชากรหนึ่งมา  $n$  จำนวน เมื่อได้ข้อมูลมา  $n$  จำนวน เราจะคำนวณค่าสถิติ เช่น หาค่าเฉลี่ย หาคความแปรปรวน หรือหาค่าสัดส่วน สมมติเราสนใจอยากทราบค่าเฉลี่ย จากตัวอย่างชุดหนึ่ง เราย่อมได้ค่าเฉลี่ยมาเพียงตัวเดียว จึงถือว่าเป็นค่าคงที่ไม่ใช่ตัวแปรเชิงสุ่ม แต่อย่าลืมว่า เรายังสามารถสุ่มตัวอย่างขนาดเดิมนี้อีกหลายชุด แต่ละชุดจะให้ค่าเฉลี่ยได้ชุดละ 1 ค่า เราจึงอยากทราบลักษณะการแจกแจงของค่าเฉลี่ยจากตัวอย่างขนาดเดียวกันนี้ ( $n$ ) ที่เป็นไปได้ทั้งหมด ว่ามีความสัมพันธ์กับประชากรที่ถูกสุ่มมาหรือไม่? อย่างไร?

สมมติประชากรของเราประกอบด้วยเด็ก 4 คน ซึ่งมีอายุ 1, 3, 5, 7 ตามลำดับ นั่นคือ

$$\begin{aligned}\mu &= (1 + 3 + 5 + 7)/4 = 16/4 = 4 \\ \sigma^2 &= \frac{\sum(X - \mu)^2}{N} = \frac{(1 - 4)^2 + (3 - 4)^2 + (5 - 4)^2 + (7 - 4)^2}{4} \\ &= \frac{(9 + 1 + 1 + 9)}{4} = \frac{20}{4} = 5 \\ \sigma &= \sqrt{5} = 2.326\end{aligned}$$

ถ้าสุ่มตัวอย่างเด็กมา 1 คน แต่เป็นการสุ่มแบบแทนที่ เด็กแต่ละคนจะมีโอกาสถูกเลือกเท่ากัน จะได้การแจกแจงความน่าจะเป็น ดังนี้

อายุ (x)	1	3	5	7	รวม
P(x)	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1.0

ถ้าสุ่มเด็กมา 2 คน แบบแทนที่ ต้องการทราบการแจกแจงความน่าจะเป็นของอายุเฉลี่ยจากตัวอย่างสุ่ม จะต้องหาตัวอย่างสุ่มที่เป็นไปได้ทั้งหมดก่อน ซึ่งจะมีทั้งหมด  $= 4 \times 4 = 16$  ชุด ดังนี้

ตัวอย่างที่	$X_1$	$X_2$	รวม	ค่าเฉลี่ย
1	1	1	2	1
2	1	3	4	2
3	1	5	6	3
4	1	7	8	4
5	3	1	4	2
6	3	3	6	3
7	3	5	8	4
8	3	7	10	5
9	5	1	6	3
10	5	3	8	4
11	5	5	10	5
12	5	7	12	6
13	7	1	8	4
14	7	3	10	5
15	7	5	12	6
16	7	7	14	7

ตารางที่ 7.1

แสดงกลุ่มอายุของเด็กที่สุ่ม  
มา 2 ค่าจากกลุ่มอายุ

{ 1 3 5 7 }

$N = 4, n = 2$

$\mu = 4, \sigma^2 = 5$

$\sigma = 2.326$

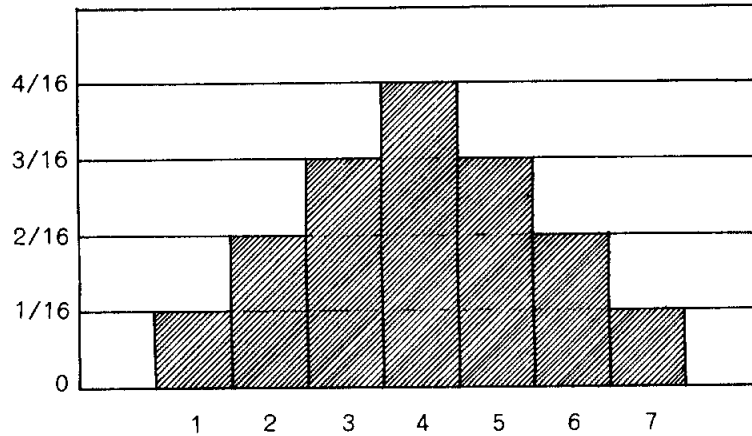
จากตารางที่ 7.1 จะมีค่าเฉลี่ยที่เป็นไปได้ 16 ตัว แต่ละตัวจะมีโอกาสเกิด  $= (\frac{1}{4})(\frac{1}{4}) = \frac{1}{16}$   
บางตัวก็เป็นค่าซ้ำกัน จึงมีค่าต่าง ๆ กันเพียง 7 ค่า เราจึงหาการแจกแจงความน่าจะเป็นได้ดังนี้

ตารางที่ 7.2 แสดงการแจกแจงตัวอย่างสุ่มของค่าเฉลี่ย  $\bar{x}$

$\bar{x}$	1	2	3	4	5	6	7	รวม
จำนวนกลุ่มตัวอย่าง	1	2	3	4	3	2	1	
ความน่าจะเป็น	1/16	2/16	3/16	4/16	3/16	2/16	1/16	1.0

### รูปที่ 7.1

แสดงการแจกแจงของ  $\bar{x}$   
ซึ่งมี  $N = 2$  จากกลุ่ม  
{1, 3, 5, 7}



นิยาม การแจกแจงของตัวอย่างสุ่มคือ การแจกแจงความน่าจะเป็นของค่าสถิติซึ่งได้จากตัวอย่างสุ่มทั้งหมดที่มีขนาดเท่ากัน และสุ่มจากประชากรเดียวกัน

### ค่าเฉลี่ยของ $\bar{X}$

เราจะหาค่าเฉลี่ยของ  $\bar{X}$  หรือ  $\mu_{\bar{x}}$  จากตารางที่ 7.2 ได้ ดังนี้

$$\begin{aligned}\mu_{\bar{x}} &= E(\bar{X}) = \sum \bar{x}_i P(\bar{x}_i) \\ &= 1(1/16) + 2(2/16) + 3(3/16) + 4(4/16) + 5(3/16) + 6(2/16) + 7(1/16) \\ &= \frac{1 + 4 + 9 + 16 + 15 + 12 + 7}{16} = \frac{64}{16} = 4 = \mu_x\end{aligned}$$

จึงสรุปได้ว่า  $E(\bar{X}) = E(X)$  หรือ  $\mu_{\bar{x}} = \mu_x$

ซึ่งอาจพิสูจน์ได้ ดังนี้

$$\begin{aligned}E(\bar{X}) &= E(\sum_{i=1}^n x_i/n) = \sum_{i=1}^n E(X_i)/n \\ &= \sum_{i=1}^n (\mu)/n = \frac{n\mu}{n} = \mu\end{aligned}$$

การหาค่าเบี่ยงเบนมาตรฐานของ  $\bar{X}$

$$\begin{aligned}\sigma_{\bar{x}}^2 &= E(\bar{X} - \mu_{\bar{x}})^2 \\ &= E(\bar{X}^2) - \mu_{\bar{x}}^2\end{aligned}$$

หาค่า  $E(\bar{X}^2)$  ดังนี้

$\bar{x}$	1	2	3	4	5	6	7
$P(\bar{x})$	$\frac{1}{16}$	$\frac{2}{16}$	$\frac{3}{16}$	$\frac{4}{16}$	$\frac{3}{16}$	$\frac{2}{16}$	$\frac{1}{16}$
$\bar{x}^2$	1	4	9	16	25	36	49

$$E(\bar{X}^2) = \sum \bar{x}^2 P(\bar{x})$$

$$= 1\left(\frac{1}{16}\right) + 4\left(\frac{2}{16}\right) + 9\left(\frac{3}{16}\right) + 16\left(\frac{4}{16}\right) + 25\left(\frac{3}{16}\right) + 36\left(\frac{2}{16}\right) + 49\left(\frac{1}{16}\right)$$

$$= 18.5$$

ดังนั้น

$$\sigma_{\bar{X}}^2 = 18.5 - 4^2 = 2.5 \quad \text{และ} \quad \sigma_{\bar{X}} = \sqrt{2.5} = 1.5811$$

พึงสังเกตว่า  $\sigma_x^2 = 5$

$$\frac{\sigma_x^2}{n} = \frac{5}{2} = 2.5 = \sigma_{\bar{X}}^2$$

จึงสรุปได้ว่า ความแปรปรวนของค่าเฉลี่ยเท่ากับความแปรปรวนของประชากรหารด้วยขนาดตัวอย่าง  
 สุ่ม เมื่อสุ่มจากประชากรแบบนับไม่ถ้วน (infinite population) หรือการสุ่มแบบมีการแทนที่

ดังนั้น  $\sigma_{\bar{x}} = \sigma/\sqrt{n} = 2.326/\sqrt{2} = 1.58$

**นิยาม** ค่าเบี่ยงเบนมาตรฐาน (standard deviation) ของค่าสถิติเรียกว่า ความคลาดเคลื่อน  
 มาตรฐาน (standard error)

เหตุที่เรียกว่าความ “คลาดเคลื่อน” หรือ “error” เนื่องจากในตัวอย่างขนาดเดียวกัน และสุ่มมาจากประชากรเดียวกัน จะมีค่าเบี่ยงเบนมาตรฐานแตกต่างกัน และแตกต่างจากค่าเบี่ยงเบนมาตรฐานของประชากรที่สุ่มตัวอย่างมาด้วย

. ลองพิจารณาความสัมพันธ์ระหว่างตัวอย่างสุ่มกับประชากรแบบนับถ้วน (finite population) จะใช้ตัวอย่างเดิมคือประชากรประกอบด้วยเด็ก 4 คน ซึ่งอายุ 1, 3, 5, 7 ถ้าสุ่มมา 2 คน แบบไม่แทนที่ จะได้ตัวอย่างทั้งหมด  $\binom{4}{2} = 6$  ชุด และมีค่าเฉลี่ยต่าง ๆ ดังนี้

ตัวอย่าง	1	2	3	4	5	6
อายุเด็ก	1, 3	1, 5	1, 7	3, 5	3, 7	5, 7
อายุเฉลี่ย ( $\bar{x}$ )	2	3	4	4	5	6

จึงสร้างตารางการแจกแจงความน่าจะเป็นของค่าเฉลี่ย ได้ดังนี้

$\bar{x}$	2	3	4	5	6	รวม
$P(\bar{x})$	1/6	1/6	2/6	1/6	1/6	1.0

จึงหาค่าเฉลี่ยและความแปรปรวนของ  $\bar{X}$  ได้ดังนี้

$$\begin{aligned}
 E(\bar{X}) &= 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{2}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) \\
 &= \frac{24}{6} = 4 = \mu_{\bar{X}} \\
 \sigma_{\bar{X}}^2 &= E(\bar{X} - \mu_{\bar{X}})^2 \\
 &= E(X^2) - \mu_{\bar{X}}^2 \\
 E(\bar{X}^2) &= \sum \bar{x}^2 p(\bar{x}) \\
 &= 4\left(\frac{1}{6}\right) + 9\left(\frac{1}{6}\right) + 16\left(\frac{2}{6}\right) + 25\left(\frac{1}{6}\right) + 36\left(\frac{1}{6}\right) \\
 &= 106/6
 \end{aligned}$$

ดังนั้น

$$\begin{aligned}
 \sigma_{\bar{X}}^2 &= E(\bar{X}^2) - \mu_{\bar{X}}^2 \\
 &= \frac{106}{6} - 4^2 \\
 &= \frac{10}{6} = \frac{5}{3}
 \end{aligned}$$

จะเห็นว่า  $\sigma_{\bar{X}}^2$  ไม่เท่ากับ  $\sigma_{\frac{X}{n}}^2$  เหมือนการสุ่มแบบแทนที่

ความจริง  $\sigma_{\bar{X}}^2 = \frac{5}{3}$  คือ  $\frac{\sigma^2}{n} \left(\frac{N-n}{N-1}\right)$

$$\begin{aligned}
 \text{ลองแทนค่า } \frac{\sigma^2}{n} \left(\frac{N-n}{N-1}\right), \quad \sigma^2 = 5, N = 4, n = 2 \\
 = \frac{5}{2} \left(\frac{4-2}{4-1}\right) = \frac{5}{2} \left(\frac{2}{3}\right) = \frac{5}{3}
 \end{aligned}$$

เทอม  $\left(\frac{N-n}{N-1}\right)$  เรียกว่า finite population correction factor หรือ f p c

บางครั้ง  $\left(\frac{N-n}{N-1}\right)$  จะมีค่าใกล้ 1 มาก ดังนั้น เมื่อนำไปคูณ  $\frac{\sigma^2}{n}$  จึงไม่ทำให้ค่าเปลี่ยนแปลงมาก กรณีที่ fpc จะมีค่าใกล้ 1 คือ เมื่อขนาดตัวอย่าง  $n$  เล็กมากเมื่อเทียบกับ  $N$  จนทำให้การสุ่มแบบไม่มีการแทนที่ ไม่ต่างกับการแทนที่ นั่นคือการสุ่มจาก finite population ไม่ต่างกับ infinite

population หลักใหญ่ที่ใช้ปฏิบัติคือ ให้พิจารณาเทอม  $n/N$  ซึ่งเรียกว่า sampling fraction ถ้ามีค่าน้อยกว่า .05 ไม่ต้องใช้ fpc

### รูปร่างการแจกแจงของตัวอย่างสุ่ม

จากตัวอย่างเรื่องอายุของเด็ก ทำให้เราทราบว่า การแจกแจงของตัวอย่างสุ่มมีความสัมพันธ์กับประชากร คือ  $\mu_{\bar{x}} = \mu$  ไม่ว่าจะสุ่มจากประชากรที่มีรูปร่างแบบใดก็ตาม และ  $\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \left( \frac{N-n}{N-1} \right)$  เมื่อสุ่มจาก finite population หรือเป็นการสุ่มแบบไม่แทนที่ และ  $\left( \frac{N-n}{N-1} \right) = 1$  ถ้าสุ่มจาก infinite population หรือ สุ่มแบบแทนที่ แต่ยังไม่ได้กล่าวถึงรูปร่างการแจกแจงของ  $\bar{X}$  ว่ามีความสัมพันธ์กับรูปร่างของประชากรอย่างไร ทฤษฎี central limit theorem จะอธิบายความสัมพันธ์ดังกล่าวได้ ดังนี้

“สำหรับประชากรเกือบทั้งหมด คือจะเป็นประชากรแบบปกติหรือไม่ใช้ก็ตาม การแจกแจงของ  $\bar{x}$  จะมีการแจกแจงโดยประมาณแบบการแจกแจงปกติ เมื่อขนาดตัวอย่าง  $n$  มีค่าโตพอควร”

ทฤษฎีนี้ มีความสำคัญมาก เพราะเป็นรากฐานของการประมาณค่าและการทดสอบสมมุติฐานที่จะกล่าวในบทต่อไป คำว่า ขนาดตัวอย่างโตพอควร ปกติถือว่า ต้องไม่น้อยกว่า 30 และถือ  $n \geq 30$  เป็นค่าทั่ว ๆ ไปสำหรับประชากรทุกประเภท แต่ถ้าสุ่มจากประชากรปกติ หรือที่มีรูปร่างใกล้เคียงประชากรแบบปกติ อาจให้  $n$  มีค่าไม่ต้องโตมากก็ได้

จากทฤษฎีนี้ ทำให้เราสามารถหาความน่าจะเป็นของ  $\bar{x}$  ในรูปของ  $Z$  ได้ ดังนี้

$$Z = \frac{\bar{x} - \mu}{\sigma/\sqrt{n}}$$

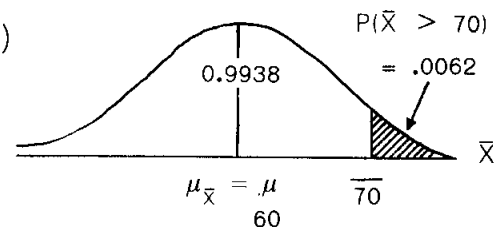
ตัวอย่าง 1 ถ้าคะแนนสอบของนักเรียนห้องหนึ่งมีการแจกแจงแบบปกติและมีค่าเฉลี่ย 60 คะแนน ความแปรปรวน 256 คะแนน<sup>2</sup> เมื่อสุ่มตัวอย่างมา 16 คน จงหาความน่าจะเป็นที่จะได้คะแนนเฉลี่ยสูงกว่า 70 คะแนน

กรณีนี้ ถือว่าประชากรเป็นแบบนับไม่ถ้วน เพราะไม่ทราบค่า  $N$  ดังนั้น

$$\mu_{\bar{x}} = \mu = 60 \text{ และ } \sigma_{\bar{x}}^2 = \sigma^2/n = 256/16 = 16$$

$$\text{และ } \sigma_{\bar{x}} = \sqrt{16} = 4 \text{ หรือจะใช้สูตร } \sigma_{\bar{x}} = \sigma/\sqrt{n} = 16/\sqrt{16} = 4$$

$$\begin{aligned} P(\bar{X} > 70) &= P\left(\frac{\bar{X} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} > \frac{70 - 60}{16/\sqrt{16}}\right) \\ &= P(Z > 2.5) \\ &= .0062 \end{aligned}$$

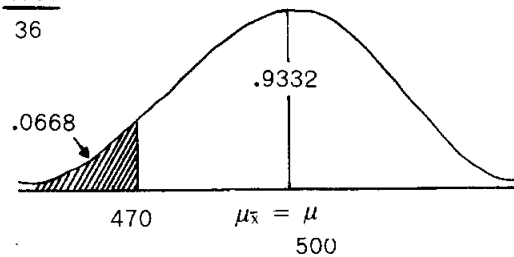


กรณีนี้ เป็นการสุ่มจากประชากรแบบปกติ แม้ว่าขนาดตัวอย่างจะไม่โตมาก การแจกแจงของ  $\bar{x}$  จะยังคงเป็นแบบปกติ แต่ถ้าประชากรไม่เป็นแบบปกติ  $n$  ต้องโตจึงจะทำให้  $\bar{x}$  มีการแจกแจงแบบปกติ

ตัวอย่าง 2 ให้  $X$  แทนค่าจ้างรายสัปดาห์ของพนักงานขาย ซึ่งไม่ทราบรูปแบบการแจกแจงของประชากรแต่คาดว่าไม่ใช่แบบปกติ ถ้าสุ่มพนักงานขายมา 36 คน แล้วจดค่าจ้างรายสัปดาห์ไว้ และสมมุติว่าค่าจ้างรายสัปดาห์มาจากประชากรที่มี ค่าเฉลี่ย  $\mu = 500$  และ  $\sigma^2 = 14,400$  จงหาความน่าจะเป็นที่  $\bar{x}$  จะน้อยกว่า 470

แม้ว่า  $X$  จะไม่มีการแจกแจงแบบปกติ แต่  $n = 36$  ถือว่าโต เพราะมากกว่า 30 จึงแปลงเป็นค่า  $Z$  ได้

$$\begin{aligned} P(\bar{x} < 470) &= P\left(\frac{\bar{x} - \mu_{\bar{x}}}{\sigma/\sqrt{n}} < \frac{470 - 500}{\sqrt{\frac{14700}{36}}}\right) \\ &= P(Z < -1.5) \\ &= .0668 \end{aligned}$$



### การแจกแจงตัวอย่างของ $p$

ในบทที่ 6 ได้กล่าวถึงตัวแปรเชิงสุ่มแบบทวินาม คือ  $X$  หมายถึง จำนวนครั้งของความสำเร็จจากการทดลองแบบเบอร์นูลลี  $n$  ครั้ง หรือหมายถึงความสำเร็จจากตัวอย่างสุ่มขนาด  $n$  จากประชากรแบบนับไม่ถ้วน  $X$  จะมีการแจกแจงแบบทวินามและมีฟังก์ชันน่าจะเป็น

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

ในเมื่อ  $p$  คือสัดส่วนความสำเร็จของประชากร ดังนั้นเพื่อหลีกเลี่ยงความสับสนในบทต่อไป จะใช้  $p$  แทน population proportion ซึ่งเป็นค่าพารามิเตอร์ที่ต้องทราบล่วงหน้าก่อนคำนวณความน่าจะเป็น และเมื่อได้กระทำการทดลองแล้ว และนับความสำเร็จได้  $x$  ครั้ง จาก  $n$  ครั้ง จะได้สัดส่วนความสำเร็จจากตัวอย่าง คือ  $p = x/n$ ,  $0 \leq p \leq 1$

$p$  จึงเป็นค่าสถิติ เพราะหาได้จากตัวอย่าง และต้องทำการทดลองสิ้นสุดแล้วจึงจะทราบค่า  $p$  เมื่อกำหนดขนาดตัวอย่างหรือจำนวนครั้ง  $= n$  ซึ่งเป็นค่าคงที่ เราอาจทำการทดลองซ้ำ ๆ กันได้



หลาย ๆ ครั้ง โดยใช้ขนาดตัวอย่างเท่าเดิมคือ  $n$  ดังนั้น เราจะได้ค่าสถิติ  $p =$  สัดส่วนความสำเร็จจากตัวอย่างสุ่มหลายค่า เราจึงอยากทราบลักษณะการแจกแจงของ  $p$  จากบทที่ 6 เราทราบว่า ค่าเฉลี่ยและความแปรปรวนของตัวแปรเชิงสุ่มทวินามคือ

$$E(X) = n \pi \quad \sigma_x^2 = n \pi (1 - \pi)$$

$$p = \frac{x}{n}$$

$$E(p) = E\left(\frac{X}{n}\right) = \frac{1}{n} E(x) = \frac{1}{n} (n \pi) = \pi$$

$$\sigma_p^2 = \sigma^2\left(\frac{X}{n}\right) = \frac{1}{n^2} \sigma_x^2 = \frac{n \pi (1 - \pi)}{n^2} = \frac{\pi (1 - \pi)}{n}$$

$$\sigma_p = \sqrt{\frac{\pi (1 - \pi)}{n}}$$

สรุป

การแจกแจงตัวอย่างสุ่ม ของ $X$	การแจกแจงตัวอย่างสุ่ม ของ $p$
$E(X) = n \pi$	$E(p) = \pi$
$\sigma_x^2 = n \pi (1 - \pi)$	$\sigma_p^2 = \pi (1 - \pi) / n$
$\sigma_x = \sqrt{n \pi (1 - \pi)}$	$\sigma_p = \sqrt{\frac{\pi (1 - \pi)}{n}}$

Central Limit Theorem

การแจกแจงตัวอย่างสุ่มของ  $p$  จะเบ้ ถ้า  $p \neq 0.5$  แต่ความเบ้จะลดลงเมื่อ  $n$  โตขึ้น ๆ เมื่อ  $n \geq 30$  จะมีรูปร่างใกล้เคียงกับโค้งปกติ จึงเป็นจริงตามทฤษฎีขีดจำกัดเข้าสู่ส่วนกลางที่ว่า

“การแจกแจงตัวอย่างของ  $x$  และ  $p$  จะเป็นแบบปกติโดยประมาณเมื่อขนาดตัวอย่าง  $n$  โดพอสมควร”

คำว่า “โพอสมควร” กว้างมาก เพราะ ถ้า  $\pi$  มีค่าใกล้เคียง 0.5 ขนาดตัวอย่าง ( $n$ ) ก็ไม่ต้องโตมาก เพราะการแจกแจงของ  $p$  มีลักษณะสมมาตร เมื่อนำโค้งปกติประมาณจึงมีความผิด

พลาดค่า แต่ ถ้า  $\pi$  มีค่าต่างจาก 0.5 มาก การแจกแจงจะเบ้ เมื่อประมาณด้วยโค้งปกติซึ่งมีลักษณะสมมาตรจะมีความผิดพลาดสูง ดังนั้น จึงต้องพิจารณาทั้งค่า  $n$  และ  $\pi$  กฎที่ใช้คือ เมื่อ  $n\pi \geq 5$  และ  $n(1 - \pi) \geq 5$  เมื่อประมาณด้วยโค้งปกติแล้วก็จะหาความน่าจะเป็นได้ในรูปพื้นที่ภายใต้โค้งปกติแต่ต้องแปลงค่า  $x$  หรือ  $p$  อยู่ในรูปค่า  $Z$  ก่อน ดังนี้

$$Z = \frac{X - E(X)}{\sigma_x} = \frac{X - n\pi}{\sqrt{n\pi(1-\pi)}} \quad \text{สำหรับแปลงค่า } X$$

และ

$$Z = \frac{p - E(p)}{\sigma_p} = \frac{p - \pi}{\sqrt{\frac{\pi(1-\pi)}{n}}} \quad \text{สำหรับแปลงค่า } p$$

หมายเหตุ ถ้าเปลี่ยนค่า  $x$  เป็น  $Z$  ต้องปรับความต่อเนื่องโดยนำ 0.5 ห่างจากค่าต่ำของ  $x$  และเอา 0.5 บวกเพิ่มค่าสูงของ  $x$

ตัวอย่าง 1 สุ่มใบสมัครเพื่อขอประกันมา 50 ใบ จงหาความน่าจะเป็นที่จะมีใบประกันที่สมบูรณ์อย่างน้อย 30 ใบ สมมติ  $\pi = 0.7$

$$p(x \geq 30) = \sum_{x=30}^{50} \binom{50}{x} (.7)^x (.3)^{50-x} \quad \text{คิดจากสูตรการแจกแจงที่แท้จริงคือแบบทวินาม ซึ่งยาก}$$

มาก โดย Central Limit Theorem จะประมาณโดยการแจกแจงแบบปกติได้ เพราะ

$$n\pi = 50(.7) = 35 > 5 \quad \text{และ} \quad n(1 - \pi) = 50(.3) = 15 > 5$$

$$E(x) = n\pi = 35, \quad \sigma_x = \sqrt{n\pi(1-\pi)} = \sqrt{50(.7)(.3)} = 3.24$$

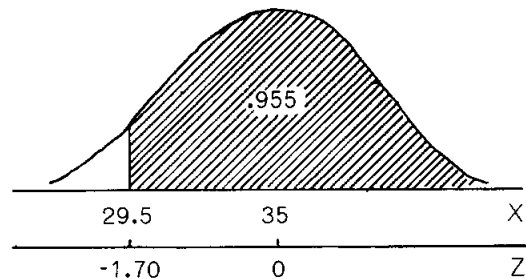
$P(x \geq 30)$  เมื่อปรับความต่อเนื่อง คือ  $P(X \geq 29.5)$

เปลี่ยน  $x$  ในเทอมของ  $Z$  จะได้

$$Z = \frac{29.5 - 35}{3.24} = -1.70$$

$$P(X \geq 29.5) = P(Z \geq -1.70)$$

$$= .955$$



ตัวอย่าง 2 ถ้าสุ่มมา 200 ใบ จงหาความน่าจะเป็นที่จะได้ค่า  $p$  ไม่ต่างจากค่าจริง ( $\pi = .7$ ) เกิน 5%

นั่นคือต้องการ  $P(.65 < p < .75)$

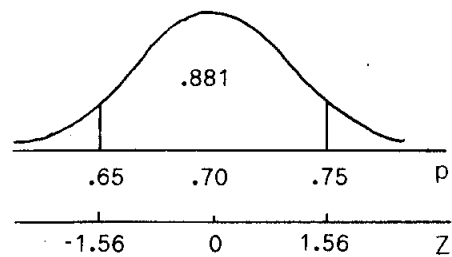
$$E(p) = \pi = 0.7, \quad \sigma_p = \sqrt{\frac{\pi(1-\pi)}{n}} = \sqrt{\frac{.7(.3)}{200}} = .032$$

เนื่องจาก  $n = 200$  นับว่าโตมาก จึงไม่ปรับความต่อเนื่อง เปลี่ยนค่า  $p$  เป็น  $Z$  ดังนี้

$$\text{เมื่อ } p = .65, Z = \frac{.65 - .70}{.032} = -1.56$$

$$\text{เมื่อ } p = .75, Z = \frac{.75 - .70}{.032} = 1.56$$

$$\begin{aligned} P(.65 < p < .75) &= P(-1.56 < Z < 1.56) \\ &= 0.881 \end{aligned}$$



### แบบฝึกหัด

- 7.1 เครื่องจักรบรรจุน้ำอัดลมขนาดแก้วละ 100 กรัม และมีค่าเบี่ยงเบนมาตรฐาน 15 กรัม ฝ่ายควบคุมคุณภาพได้สุ่มตัวอย่างมาจำนวนหนึ่งได้น้ำหนักเฉลี่ย 105 กรัม ฝ่ายควบคุมคุณภาพจึงสรุปว่า ตัวอย่างที่ได้ไม่ทำหน้าที่ "ตัวแทน" ที่ดี จงวิจารณ์
- 7.2 เทอม "error" ที่ใช้กับ "standard error" หมายถึงความคลาดเคลื่อนแบบใด?
- 7.3 สุ่มตัวอย่างมา 36 ค่าจากประชากรที่มีค่าเฉลี่ย 125 และความแปรปรวน 225 จงหา  
 (ก)  $P(\bar{X} > 127)$     (ข)  $P(\bar{X} > 130)$     (.2119, .0228)  
 ถ้าสุ่มมา 81 ค่า จงหา(ค)  $P(\bar{X} < 127)$  (ง)  $P(\bar{X} > 130)$     (.8849, .0013)
- 7.4 จะต้องสุ่มตัวอย่างจากประชากรที่มี  $\mu = 72, \sigma = 10$  ก็จำนวนจึงจะทำให้ความเชื่อมั่น 90% ที่จะได้ค่าเฉลี่ยของตัวอย่างสูงกว่า 70    (42)
- 7.5 จะต้องใช้ขนาดตัวอย่างเท่าใดเพื่อสุ่มจากประชากรปกติที่มี  $\mu = 250$  และ  $\sigma = 20$  และให้มีความน่าจะเป็น .95 ที่ค่าเฉลี่ยจากตัวอย่างจะตกอยู่ระหว่าง 240 - 260    (16)

- 7.6 บริษัทขายรถได้โดยถัวเฉลี่ยวันละ 10 คัน ด้วยความแปรปรวน  $35.5$  คัน<sup>2</sup> ถ้าสุ่มมา 7 วัน จงหาความน่าจะเป็น
- ก) ที่จะขายได้มากกว่า 16 คัน หรือน้อยกว่า 12 คัน (0.8171)
- ข) ขายได้น้อยกว่า 16 คัน และมากกว่า 5 คัน (0.9830)
- 7.7 โรงงานผลิตจักรยานได้ประมาณต้นทุนแรงงานโดยถัวเฉลี่ยต่อคัน = 50.2 บาท และมีค่าเบี่ยงเบนมาตรฐาน 1.2 บาท ถ้าสุ่มมา 20 คัน เพื่อคำนวณค่าต้นทุนแรงงาน จะกล่าวด้วยความเชื่อมั่น 98% ได้หรือไม่ว่าค่าเฉลี่ยแรงงานต่อคันจากตัวอย่างสุ่มจะอยู่ระหว่าง 49.5 และ 50.7 บาท
- 7.8 ถ้าสุ่มตัวอย่างขนาด  $n = 16$  จากประชากร  $N = 65, \mu = 12$ , และ  $\sigma = 2.1$  จงหาความน่าจะเป็นที่จะได้ค่าเฉลี่ยอยู่ระหว่าง 11.5 ถึง 12.5 (0.7234)
- 7.9 ถ้า  $N = 145, \mu = 120, \sigma = 15, n = 64$
- (ก) จงประมาณความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ย (1.41)
- (ข) จงหา  $P(122 < \bar{X} < 124)$  (0.0755)
- 7.10 ถ้า  $N = 120, \mu = 7.5, \sigma = 1.5$  จงประมาณความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยเมื่อใช้ขนาดตัวอย่างต่าง ๆ ดังนี้
- (ก)  $n = 9$       (ข)  $n = 25$       (ค)  $n = 49$
- 7.11 นักธุรกิจผู้หนึ่งมีร้านอาหาร 25 ร้าน ได้กำไรถัวเฉลี่ย 21,000 บาท ต่อร้านต่อปี และค่าเบี่ยงเบนมาตรฐาน 3,400 บาท จงหาความน่าจะเป็นที่ร้านอาหารที่สุ่มมา 5 ร้าน จะได้กำไรรวมกันไม่ถึง 100,000 บาท ต่อปี (0.2358)
- 7.12 ให้  $X$  คือจำนวนครั้งที่เครื่องจักรขัดข้องใน 1 วัน ในโรงงานหนึ่งซึ่งทราบว่า  $X$  มีการแจกแจงแบบปกติด้วยค่าเฉลี่ย  $\mu = 5$  และ  $\sigma^2 = 12$  ถ้าสุ่มตัวอย่างวันทำการมา 4 วัน เพื่อดูจำนวนครั้งที่เครื่องจักรขัดข้องในแต่ละวัน จงหา
- ก) จำนวนคาดหวังของเครื่องจักรขัดข้องต่อวัน หรือ  $E(\bar{x})$
- ข) ความแปรปรวนของ  $\bar{X}$  (5, 3.0)
- 7.13 กำหนดให้  $\{10, 12, 14, 16\}$  เป็นกลุ่มผลทดลองของ  $X$  ถ้าสุ่มเลขมา 2 จำนวนแบบแทนที่
- ก) จงหาค่าเฉลี่ย และความแปรปรวนของ  $\bar{X}$  (13, 5.0)
- ข) จงหาค่าเฉลี่ย และความแปรปรวนของ  $\bar{X}$  (13, 2.5)

- 7.14 ให้  $Y$  คือ จำนวนครั้งที่แม่บ้านคนหนึ่งไปจ่ายตลาดต่อ 1 เดือน และ  $Y$  มีการแจกแจงความน่าจะเป็นดังนี้

$Y$	0	1	2	3	4	5	รวม
$f(y)$	0.1	0.2	0.3	0.2	0.1	0.1	1.0

- ก) จงหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของ  $Y$  (2.3, 1.414)  
 จ) จงหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของ  $\bar{Y}$  ซึ่งเป็นค่าเฉลี่ยของแม่บ้านที่สุ่มมา 4 คน (2.3, 0.707)
- 7.15 ถ้า  $X$  แทนจำนวนครั้งที่ประธานบริษัทต้องไปประชุมต่างประเทศต่อ 1 เดือน และ  $X$  มีการแจกแจงความน่าจะเป็นดังนี้

$x$	0	1	2	3	รวม
$P(x)$	0.3	0.4	0.2	0.1	1.0

- ก) จงหาค่าเฉลี่ยและความแปรปรวนของ  $x$  (1.1, 0.89)  
 ข) จงหาค่าเฉลี่ยและความแปรปรวนของ  $\bar{x}$  ซึ่งเป็นค่าเฉลี่ยของการสุ่มมา 9 เดือน (1.1, .09889)
- 7.16 ถ้าเชือกเส้นหนึ่งมีความทนทานเฉลี่ย 2,000 ปอนด์ และมีค่าเบี่ยงเบนมาตรฐาน 50 ปอนด์ ให้  $\bar{x}$  เป็นค่าเฉลี่ยจากตัวอย่างเชือกที่สุ่มมา 100 เส้น จงหาค่าคาดหวัง และความคลาดเคลื่อนมาตรฐานของ  $\bar{x}$  (2,000, 5)
- 7.17 ถ้า  $X$  เป็นอายุการใช้งานของยางรถยนต์ ซึ่งมีค่าเฉลี่ย 30,000 ไมล์ และค่าเบี่ยงเบนมาตรฐาน 200 ไมล์ ถ้าสุ่มตัวอย่างมา 16 อัน จงหาค่าคาดหวังและความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยของตัวอย่างสุ่ม (30,000, 50)
- 7.18 ถ้าน้ำหนักของลูกวัวแรกเกิดมีการแจกแจงแบบปกติ ด้วยค่าเฉลี่ย 50 ปอนด์ และค่าเบี่ยงเบนมาตรฐาน 4 ปอนด์
- ก) ถ้าสุ่มลูกวัวมา 1 ตัว จงหาความน่าจะเป็นที่จะมีน้ำหนักต่ำกว่า 45 ปอนด์ (.1056)  
 ข) ถ้าสุ่มมา 25 ตัว จงหาความน่าจะเป็นที่จะได้น้ำหนักเฉลี่ยสูงกว่า 51 ปอนด์ (.1056)  
 ค) จากข้อ (ข) จงหาโอกาสที่จะได้น้ำหนักเฉลี่ยระหว่าง 49 ถึง 51 ปอนด์ (.7888)

- 7.19 กลัวยหอมมีความยาวเฉลี่ย 8 นิ้ว และค่าเบี่ยงเบนมาตรฐาน 1.44 นิ้ว สุ่มกลัวยหอมจากแปลงทดลองมา 36 ผล จงหาความน่าจะเป็นที่จะได้ความยาวเฉลี่ย 8.3 นิ้วขึ้นไป (.1056)
- 7.20 ข้อใดต่อไปนี้จะใช้ประมาณด้วยการแจกแจงแบบปกติตามกฎที่เรียนมา  
 (ก)  $n = 100, \pi = .01$       (ค)  $n = 50, \pi = .95$   
 (ข)  $n = 100, \pi = .10$       (ง)  $n = 20, \pi = .40$
- 7.21 ความน่าจะเป็นที่ลูกค้าของร้านสรรพสินค้าแห่งหนึ่งจะซื้อสินค้า = 0.40 ถ้าสุ่มลูกค้ามา 20 ราย จงหาความน่าจะเป็นที่จะมีผู้ซื้อสินค้าน้อยกว่า 6 ราย โดยใช้โค้งปกติประมาณ และควรปรับความต่อเนื่อง (.2467)
- 7.22 องค์การโทรศัพท์ต้องการทราบเปอร์เซ็นต์เสาขั้วรถที่จะต้องเปลี่ยนภายใน 5 ปี สมมติ  $\pi = .35$  และทำการสุ่มตัวอย่าง 500 หน่วย  
 (ก) จงหาค่าเฉลี่ยและค่าเบี่ยงเบนมาตรฐานของการแจกแจงตัวอย่างสุ่มของ  $p$   
 (ข) จงหาความน่าจะเป็นที่  $p$  จะมีค่าต่างจาก  $\pi$  ไม่เกิน 5%  
 (ค) จงหาความน่าจะเป็นในข้อ (ข) โดยให้  $\pi = .50$ , และ  $\pi = .20$  ท่านได้แนวความคิดอะไรบ้างไหม? (.02, .9876, .9750, .9948)

**การแจกแจงตัวอย่างของผลบวกและผลต่างของตัวสถิติ**

สำหรับตัวอย่างที่เป็นอิสระกัน 2 ชุด

$$\mu(\bar{x}_1 - \bar{x}_2) = \mu_1 - \mu_2, \quad \sigma(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\mu(\bar{x}_1 + \bar{x}_2) = \mu_1 + \mu_2, \quad \sigma(\bar{x}_1 + \bar{x}_2) = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

$$\mu(p_1 - p_2) = \pi_1 - \pi_2, \quad \sigma(p_1 - p_2) = \sqrt{\pi_1(1-\pi_1)/n_1 + \pi_2(1-\pi_2)/n_2}$$

$$\mu(p_1 + p_2) = \pi_1 + \pi_2, \quad \sigma(p_1 + p_2) = \sqrt{\pi_1(1-\pi_1)/n_1 + \pi_2(1-\pi_2)/n_2}$$

การแปลงให้เป็น Standard Normal

ใช้  $Z = (\hat{\theta} - \mu_{\hat{\theta}}) / \sigma_{\hat{\theta}}$ ,  $\hat{\theta}$  คือ  $(\bar{x}_1 - \bar{x}_2)$  หรือ  $(\bar{x}_1 + \bar{x}_2)$  หรือ  $(p_1 - p_2)$   
 หรือ  $(p_1 + p_2)$