บทที่ 2

สถิติทดสอบและผลงานวิจัยที่เกี่ยวข้อง

ในบทนี้จะกล่าวถึงรายละเอียดของสถิติทดสอบและคุณสมบัติของการแจกแจงที่ใช้ในการวิจัยครั้งนี้ รวมทั้งผลงานวิจัยที่เกี่ยวข้อง

ให้ $\underset{\sim}{X} = (x_{i1}, x_{i2}, x_{i3}, \cdots, x_{in_i})$ เป็นตัวอย่างสุ่มขนาด n_i จากประชากรที่ i ซึ่งมีความแปรปรวน σ_i^2 $(i=1,2,\ldots,k)$

ให้ $\bar{x}_i$ เป็นค่าเฉลี่ยตัวอย่างที่ i

s_i^2 เป็นค่าความแปรปรวนตัวอย่างที่ i

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

$$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

สมมติฐานในการทดสอบความเท่ากันของความแปรปรวนของประชากร k กลุ่ม คือ

$$H_o : \sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \cdots = \sigma_k^2$$

H_1 : ความแปรปรวนอย่างน้อย 1 คู่ไม่เท่ากัน

ในการวิจัยครั้งนี้ กำหนดจำนวนประชากรเท่ากับ 3,4 กลุ่ม (k=3,4) ตัวสถิติทดสอบที่ใช้ในการทดสอบมีรายละเอียดดังนี้

2.1 สถิติทดสอบบาร์ตเลต (Bartllet's test)

Bartlett ได้พัฒนาสถิติทดสอบนี้เมื่อปี ค.ศ. 1937 (Seber 1977:147) ซึ่งหาค่าสถิติดังนี้

$$B = \frac{\sum_{i=1}^{k}(n_i-1)\ \ell n\ s^2 - \sum_{i=1}^{k}(n_i-1)\ \ell n\ s_i^2}{1 + \frac{1}{3(k-1)}\left[\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^{k}(n_i-1)}\right]}$$

$$s_i^2 = \frac{1}{n_i - 1}\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2$$

$$s^2 = \frac{\sum_{i=1}^{k}(n_i-1)\ s_i^2}{\sum_{i=1}^{k}(n_i - 1)}$$

$i = 1,2,\ \ldots,\ k$

$j = 1,2,\ \ldots,\ n_i$

n_i แทนขนาดของกลุ่มตัวอย่างที่ i

k แทนจำนวนประชากร

s_i^2 แทนความแปรปรวนของกลุ่มตัวอย่างที่ i

s^2 แทนความแปรปรวนของกลุ่มตัวอย่างทั้ง k กลุ่ม

ค่าสถิติบาร์ตเลต B เขียนอยู่ในรูป ln (natural logarithm) ซึ่งมีฐานเท่ากับ e (e มีค่าประมาณ 2.718281) อาศัยคุณสมบัติการเปลี่ยนฐานของ logarithm

$$\ln x = \frac{\log x}{\log e} \qquad \text{โดยที่} \quad \log e \cong 0.4343$$

ดังนั้นสามารถเขียน B ในรูปลอการิทึมสามัญ (common logarithm) ได้ดังนี้

$$B = \frac{2.3026\left[\sum_{i=1}^{k}(n_i - 1)\log s^2 - \sum_{i=1}^{k}(n_i - 1)\log s_i^2\right]}{1 + \frac{1}{3(k-1)}\left[\sum_{i=1}^{k}\frac{1}{n_i - 1} - \frac{1}{\sum_{i=1}^{k}(n_i - 1)}\right]}$$

โดยที่ สถิติ B มีการแจกแจงแบบไคสแควร์ ซึ่งมีองศาความเป็นอิสระ (degree of freedom) เท่ากับ k-1

จะปฏิเสธสมมติฐานว่าง H_o เมื่อค่าสถิติ B ที่คำนวณได้มากกว่าค่าไคสแควร์จากตาราง ที่องศาความเป็นอิสระเท่ากับ k-1

2.2 สถิติทดสอบโอบรีน (O'Brien's test)

O'Brien,R.G.(1981)ได้พัฒนาสถิติทดสอบนี้โดยพัฒนาตามแนวความคิดของการวิเคราะห์ความแปรปรวนและการแปลงค่าสังเกต (transformation) x_{ij} เป็น

$$z_{ij} = \frac{[(w + n_i - 2)n_i(x_{ij} - \bar{x}_i) - ws_i^2(n_i - 1)]}{(n_i - 1)(n_i - 2)}$$

$i = 1,2,\ldots,k$

$j = 1,2,\ldots,n_i$

w เป็น weighting factor ซึ่ง O'Brien แนะนำ w = 0.5

s_i^2 เป็นค่าความแปรปรวนของกลุ่มตัวอย่างที่ i

แล้วนำไปสร้างตาราง ANOVA ดังนี้

ตารางที่ 2.1 แสดงตารางการวิเคราะห์ความแปรปรวน (ANOVA)

สาเหตุของความแปรปรวน	ผลบวกกำลังสอง (S.S)	ชั้นแห่งความเป็นอิสระ	ผลรวมกำลังสองเฉลี่ย	อัตราส่วน F
ระหว่างประชากร	$SS(B)=\sum_{i}^{k} n_i(z_i - \bar{z})^2$	$k-1$	$MS(B)=\frac{SS(B)}{k-1}$	$\frac{MS(B)}{MS(W)}$
ภายในประชากร	$SS(W)=\sum_{i}^{k}\sum_{j}^{n_i}(z_{ij}-\bar{z}_i)^2$	$\sum_{i}^{k}(n_i-1)$	$MS(W)=\frac{SS(W)}{\sum_{i}^{k}(n_i-1)}$	
รวม	$SS(T)=\sum_{i}^{k}\sum_{j}^{n_i}(z_{ij}-\bar{\bar{z}})^2$			

จากตาราง ANOVA จะได้ตัวสถิติทดสอบโอบริน (OB) ดังนี้

$$OB = \frac{\sum_{i=1}^{k} n_i(\bar{\bar{z}}_i - \bar{z})^2/(k-1)}{\sum_{i=1}^{k}\sum_{j=1}^{n_i}(z_{ij} - \bar{z}_i)^2/\sum_{i=1}^{k}(n_i-1)}$$

$i = 1,2,...,k$

$j = 1,2,...,n_i$

ซึ่งมีการแจกแจงแบบ F และมีองศาความเป็นอิสระเท่ากับ $(k-1)$, $\sum_{i=1}^{k}(n_i-1)$

จะปฏิเสธสมมติฐานว่าง (H_o) เมื่อค่าสถิติ OB ที่คำนวณได้มีค่ามากกว่าหรือเท่ากับค่า F จากตารางที่องศาความเป็นอิสระเท่ากับ $(k-1), \sum_{i=1}^{k}(n_i-1)$

2.3 สถิติทดสอบสแควร์แรงค์ (Squared Rank Test)

Conover (1978) ได้เสนอการทดสอบที่ใช้ผลบวกกำลังสองของอันดับเพื่อทดสอบความเท่ากันของความแปรปรวนของประชากรซึ่งสามารถใช้ทดสอบความเท่ากันของความแปรปรวนของประชากรสองกลุ่ม หรือมากกว่าสองกลุ่มได้

ข้อตกลงเบื้องต้นของการทดสอบสแควร์แรงค์ (Conover 1980:240)

1 ข้อมูลประกอบด้วยกลุ่มตัวอย่าง k กลุ่ม สุ่มมาจากประชากร k กลุ่ม

2 กลุ่มตัวอย่างทั้งหมดเป็นอิสระต่อกัน

3 เสกลที่ใช้วัดข้อมูลอย่างน้อยเป็นเสกลแบบช่วง (interval scale)

สถิติทดสอบสแควร์แรงค์ซึ่งหาค่าสถิติ S ดังนี้

$$S = \frac{1}{D^2}\left(\sum_{i=1}^{k}\left(\frac{T_i^2}{n_i}\right) - N\bar{T}^2\right)$$

$$T_i = \sum_{j=1}^{n_i} R_j^2 \qquad ; \qquad \bar{T} = \frac{1}{N}\left(\sum_{i}^{k} T_i\right)$$

R_j = อันดับของ $|x_{ij} - \bar{x}_i|$

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

$$N = \sum_{i=1}^{k} n_i \qquad ; \qquad D^2 = \frac{1}{N-1}\left(\sum_{j=1}^{n_i} R_j^4 - N\bar{T}^2\right)$$

n_i แทนขนาดตัวอย่างของกลุ่มตัวอย่างที่ i

$i = 1,2,3, \ldots, k$

$j = 1,2,3, \ldots, n_i$

จะปฏิเสธสมมติฐานว่า (H_o) เมื่อค่าสถิติ S ที่คำนวณได้มากกว่าค่าไคสแควร์จากตาราง ที่องศาความเป็นอิสระเท่ากับ k-1

2.4 ตัวอย่างการคำนวณค่าสถิติทั้ง 3 สถิติทดสอบ

บริษัทผลิตผงซักฟอกแห่งหนึ่ง ทดลองใช้เครื่องจักร 3 เครื่อง A,B,C ในการบรรจุผงซักฟอกใส่กล่อง จากการสุ่มกล่องผงซักฟอกโดยการบรรจุของเครื่องจักรทั้งสามเครื่องมาเครื่องละ 10 กล่อง ในขบวนการผลิตครั้งหนึ่งเพื่อตรวจสอบว่า ความแปรปรวนของปริมาณผงซักฟอกที่บรรจุจากเครื่องจักรทั้งสามมีความแตกต่างกันหรือไม่ ปรากฏว่าน้ำหนักผงซักฟอก (กรัม)จากเครื่องจักรที่ใช้บรรจุแต่ละเครื่องมีดังนี้

A	B	C
220.2	220.4	218.2
220.8	220.6	220.3
220.1	220.6	218.4
219.7	216.0	219.5
221.3	221.6	221.0
218.5	219.6	218.7
219.7	218.1	221.9
220.5	219.4	218.8
218.7	223.3	221.4
220.4	222.8	221.3

2.4.1 วิธีการคำนวณค่าสถิติบาร์ตเลต

$$B = \frac{2.3026 \quad \sum_{i}^{k}(n_i-1)\log s^2 - \sum_{i}^{k}(n_i-1)\log s_i^2}{1 + \frac{1}{3(k-1)}\left[\sum_{i}^{k}\frac{1}{n_i-1} - \frac{1}{\sum_{i}^{k}(n_i-1)}\right]}$$

$\bar{x}_1 = 219.98$ $\quad s_1^2 = 0.77$

$\bar{x}_2 = 220.25$ $\quad s_2^2 = 4.67$

$\bar{x}_3 = 219.95$ $\quad s_3^2 = 1.99$

$\log s_1^2 = -0.1135$

$\log s_2^2 = 0.6693$

$\log s_3^2 = 0.2988$

$$s^2 = \frac{1}{\sum_{i}^{k} n_i - k}\left(\sum_{i}^{k}(n_i-1)s_i^2\right)$$

$s^2 = 2.476$

$\log s^2 = 0.3939$

$B = 19.6095$

2.4.2 วิธีการคำนวณค่าสถิติทดสอบโอบรีน

$$OB = \frac{\sum_{i}^{k} n_i (z_i - \bar{z})^2 / (k-1)}{\sum_{i}^{k}\sum_{j}^{n_i} (z_{ij} - \bar{z})^2 / \sum_{i}^{k} (n_i - 1)}$$

$$z_{ij} = \frac{[(w+n_i-2)n_i(x_{ij} - \bar{x}_i)^2 - ws_i^2(n_i-1)]}{(n_i-1)(n_i-2)}$$

$$s_i^2 = \frac{1}{n_i-1} \sum_{j}^{n_i} (x_{ij} - \bar{x}_i)^2 \quad ; \quad i = 1,2,\ldots,k$$

$$j = 1,2,\ldots,n_i$$

$$s^2 = \frac{\sum_{i}^{k} (n_i-1)\, s_i^2}{\sum_{i}^{k} (n_i-1)}$$

การคำนวณ

สำหรับตัวอย่างสุ่มชุดที่ 1 $\bar{x}_1 = 219.98$

$s_1^2 = 0.77$

สำหรับตัวอย่างสุ่มชุดที่ 2 $\bar{x}_2 = 220.25$

$s_2^2 = 4.67$

สำหรับตัวอย่างสุ่มชุดที่ 3 $\bar{x}_3 = 219.95$

$s_3^2 = 1.99$

ตารางที่ 2.2 แสดงค่าของตัวแปรใหม่ที่เกิดจากการแปลงโดยวิธีของโอบรีนสำหรับตัวอย่าง 3 ชุด เพื่อการคำนวณค่าสถิติโอบรีน (OB)

j	z_{1j}	z_{2j}	z_{3j}
1	0.004	-0.260	3.494
2	0.775	0.153	0.145
3	0.014	0.153	2.836
4	0.009	21.224	0.239
5	2.026	2.184	1.302
6	2.621	0.484	1.845
7	0.099	5.406	4.489
8	0.307	0.833	1.561
9	1.965	11.054	2.482
10	0.198	7.737	1.152

ค่าเฉลี่ยของตัวอย่างสุ่มแต่ละชุด

$\bar{z}_1 = 0.8108$

$\bar{z}_2 = 4.896$

$\bar{z}_3 = 2.54$

ค่าเฉลี่ยของตัวอย่างสุ่มรวมทั้งสามชุด

$\bar{z} = 2.587$

ตารางที่ 2.3 การหาค่า $(z_{ij} - \bar{z}_i)^2$

j	$(z_{1j}-\bar{z}_1)^2$	$(z_{2j}-\bar{z}_2)^2$	$(z_{3j}-\bar{z}_2)^2$
1	0.804	32.494	1.467
2	0.016	27.961	4.571
3	0.786	27.961	0.306
4	0.643	249.093	4.176
5	1.266	10.610	0.963
6	2.958	24.575	0.192
7	0.643	0.001	4.868
8	0.353	21.232	0.520
9	1.131	31.510	0.040
10	0.493	5.272	0.017
$\sum_i^k \sum_j^{n_i} (z_{ij}-\bar{z}_i)^2$	9.093	430.711	17.121

$$\frac{\sum_i^k \sum_j^{n_i} (z_{ij} - \bar{z}_i)^2}{\sum_i^k (n_i - 1)} = \frac{456.925}{27}$$

$$= 16.923$$

$$\frac{\sum_i^k n_i (z_i - \bar{z})}{(k-1)} = \{10(2.844 + 29.603 + 5.21)\}/(3-1)$$

$$= 188.2875$$

$$OB = 11.126$$

2.4.3 วิธีการคำนวณค่าสถิติทดสอบสแควร์แรงค์

$$S = \frac{1}{D^2}\left[\sum_{i=1}^{k}\left(\frac{T_i^2}{n_i}\right) - NT^2\right]$$

$$D^2 = \frac{1}{N-1}\left(\sum_{j=1}^{N} R_j^4 - N\bar{T}^2\right)$$

n_i แทนขนาดตัวอย่างของกลุ่มตัวอย่างที่ i

$$N = \sum_{i=1}^{k} n_i$$

R_j = อันดับของ $|x_{ij} - \bar{x}_i|$

$\bar{x}_i$ เป็นค่าเฉลี่ยของกลุ่มตัวอย่างที่ i

$$T_i = \sum_{j=1}^{n_i} R_j^2$$

$$\bar{T} = \frac{1}{N}\left(\sum_{i=1}^{k} T_i\right)$$

$i = 1,2,3, \ldots, k$

$j = 1,2,3, \ldots, n_i$

T_i เป็นผลบวกกำลังสองของอันดับของค่าสัมบูรณ์ของผลต่างระหว่างค่าสังเกตกับค่าเฉลี่ยของกลุ่มตัวอย่างที่ i

ค่าเฉลี่ยของตัวอย่างสุ่มแต่ละกลุ่มเป็นดังนี้

$\bar{x}_1$ = 219.98

$\bar{x}_2$ = 220.25

$\bar{x}_3$ = 219.95

ตารางที่ 2.3 แสดงการเรียงอันดับของค่าสัมบูรณ์ของผลต่างระหว่างค่าสังเกตกับค่าเฉลี่ยของกลุ่มตัวอย่างเพื่อคำนวณค่าสถิติสแควร์แรงค์

j	$\lvert x_{1j} - \bar{x}_1 \rvert$	R_j	$\lvert x_{2j} - \bar{x}_2 \rvert$	R_j	$\lvert x_{3j} - \bar{x}_3 \rvert$	R_j
1	0.21	3	0.16	2	1.70	25
2	0.81	13	0.36	7.5	0.35	6
3	0.11	1	0.36	7.5	1.50	24
4	0.29	4.5	4.24	30	0.40	9
5	1.31	19	1.36	21	1.05	15
6	1.49	23	0.64	12	1.20	17
7	0.29	4.5	2.14	27	1.95	26
8	0.51	11	0.84	14	1.10	16
9	1.29	18	3.06	29	1.45	22
10	0.41	10	2.56	28	1.35	20

$$T_1 = 1654.5$$

$$T_2 = 4151.5$$

$$T_3 = 3648$$

$$\sum_{i=1}^{k} \frac{(T_i)^2}{n_i} = \frac{1654.5^2}{10} + \frac{4151^2}{10} + \frac{3648^2}{10}$$

$$= 348752.116$$

$$\bar{T} = \frac{1}{N} \sum_{i=1}^{k} T_i$$

$$= 315.1$$

$$\sum_{j=1}^{N} (R_j)^4 = 5273769$$

$$D^2 = \frac{1}{30-1} [5273769 - 30(315.1)^2]$$

$$S = \frac{[3328022 - 30(315.1)^2]}{D^2}$$

$$= 3.3947$$

เนื่องจากในการวิจัยครั้งนี้สนใจศึกษาถึงอำนาจการทดสอบของตัวสถิติทดสอบทั้ง 3 ตัว ข้างต้นเมื่อข้อมูลตัวอย่างสุ่มมาจากประชากรที่มีการแจกแจงปกติ และการแจกแจงที่ไม่ใช่การแจกแจงแบบปกติ ซึ่งได้แก่ การแจกแจงแบบเอกซ์โปเนนเชียล การแจกแจงแบบไวบูลล์ การแจกแจงแบบที ซึ่งรายละเอียดและคุณสมบัติต่างๆ เกี่ยวกับการแจกแจงดังกล่าว มีดังนี้

2.5 การแจกแจงแบบปกติ (Normal distribution)

การแจกแจงแบบปกติเป็นการแจกแจงแบบต่อเนื่องที่มีความสำคัญมากทั้งในทางสถิติประยุกต์และใช้ประโยชน์ในการประมาณค่าของประชากร และ การทดสอบสมมติฐานทางสถิติ (ศิริชัย กาญจนวาสี 2526:94)

อับราฮัม เดอร์มัวร์ (Abraham De Moire , 7667-1754) นักคณิตศาสตร์ชาวฝรั่งเศสเป็นผู้ค้นพบการแจกแจงแบบปกติ เมื่อปี ค.ศ. 1733 ต่อมา ปิแอร์ ลาปลาส (Pierre Laplace,1749-1827) นักคณิตศาสตร์ชาวฝรั่งเศส กับ คาร์ล เกาส์ (Carl Gauss,1777-4855) นักคณิตศาสตร์และดาราศาสตร์ชาวเยอรมัน ได้ค้นพบการแจกแจงปกติโดยไม่ทราบผลงานของเดอร์มัวร์ มาก่อนเลย ซึ่งพบว่าการแจกแจงของความคลาดเคลื่อนในการวัดทางวิทยาศาสตร์กายภาพสามารถประมาณได้อย่างใกล้เคียงโดยใช้โค้งปกติ ซึ่งเขาเรียกว่า "โค้งปกติของความคลาดเคลื่อน" (The Normal curve of Error) และถือได้ว่าเป็นกฎของความน่าจะเป็น (The Laws of chance) ผลงานของลาปลาส และ เกาส์ เป็นที่รู้จักกันแพร่หลายและถูกนำไปใช้ประโยชน์อย่างกว้างขวาง การแจกแจงแบบปกติสามารถอธิบายการแจกแจงของตัวแปรต่าง ๆ ในทางชีววิทยา จิตวิทยา และทางการศึกษา เพื่อเป็นเกียรติแก่บุคคลทั้งสองบางทีเรียกการแจกแจงแบบปกติว่า "การแจกแจงของลาปลาส" (Laplacian Distribution) หรือ "การแจกแจงของเกาส์" (Gaussian Distribution) (ศิริชัย กาญจนวาสี, 2526:94-95)

ฟังก์ชันความหนาแน่น (pdf.) ของการแจกแจงแบบปกติ คือ

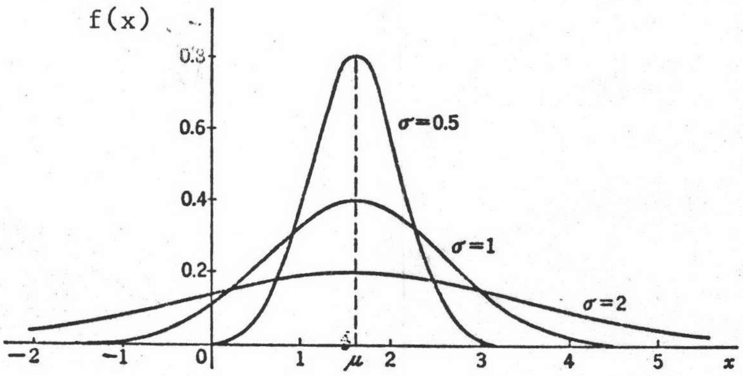
$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2} \right], \qquad -\infty < x < \infty$$

σ^2 = ความแปรปรวนของประชากร

σ = ส่วนเบี่ยงเบนมาตรฐานของประชากร

μ = ค่าเฉลี่ยของประชากร

π = 3.14159



<u>รูปที่ 2.1</u> แสดงโค้งการแจกแจงแบบปกติ เมื่อค่าค่าเฉลี่ยเท่ากับ μ และค่าส่วนเบี่ยงเบนมาตรฐาน σ = 0.5, 1, 2

<u>2.5.1 คุณสมบัติของการแจกแจงแบบปกติ</u>

2.5.1.1 ลักษณะของโค้งเป็นรูประฆังคว่ำ (bell shaped) โค้งสมมาตรกับแกนตั้งที่ $x = \mu$

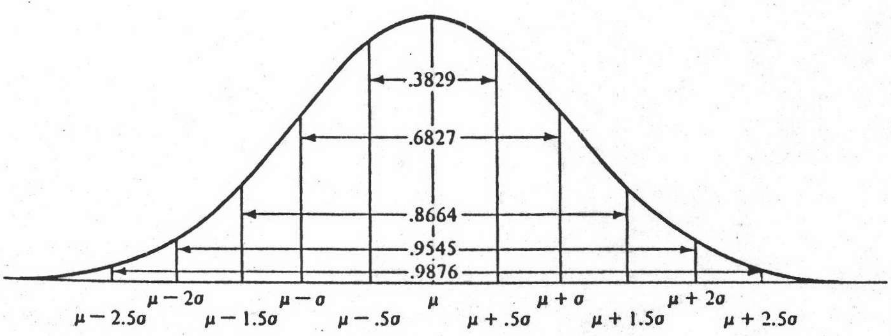
2.5.1.2 ปลายโค้งจะเข้าใกล้แกน X เมื่อ x มีค่าห่างจาก μ ออกไปแต่จะไม่ตัดแกน X

2.5.1.3 เป็นโค้งที่มีจุดสูงสุดเพียงจุดเดียว (unimodal) อยู่ที่ $x = \mu$ ซึ่งเป็นค่าฐานนิยม

2.5.1.4 ค่าเฉลี่ย ค่ามัธยฐาน และค่าฐานนิยมมีค่าเท่ากันอยู่ที่จุด $x = \mu$

2.5.1.5 มีค่าความโด่ง (Kurtosis) เท่ากับ 3 และค่าความเบ้ (Skewness) เท่ากับ 0

2.5.1.6 μ และ σ^2 เป็นค่าพารามิเตอร์ของการแจกแจงแบบปกติ โดย μ และ σ^2 จะเป็นตัวกำหนดตำแหน่งที่ตั้งของเส้นโค้งและลักษณะของเส้นโค้งว่าจะแบนหรือโด่งอย่างไร



<u>รูปที่ 2.2</u> แสดงพื้นที่โค้งของการแจกแจงแบบปกติ

2.5.1.7 พื้นที่ใต้โค้งปกติที่อยู่ระหว่าง $\mu \pm 1\sigma$, $\mu \pm 2\sigma$ และ $\mu \pm 3\sigma$ มีค่าเป็น 68.27 %, 95.45 %, และ 99.73 % ตามลำดับ

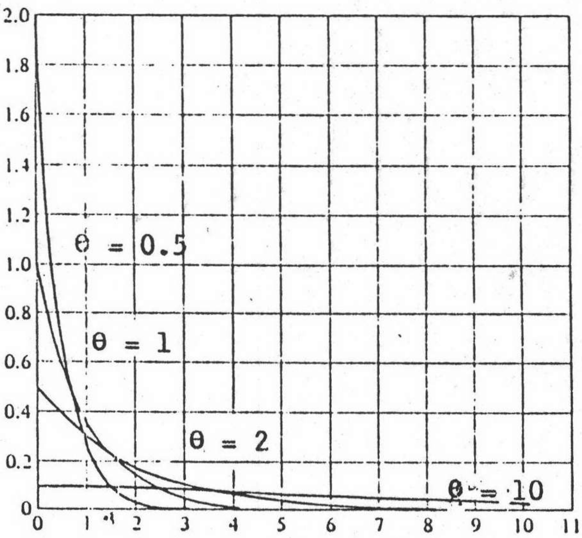
2.6 การแจกแจงเอกซ์โปเนนเชียล (Exponential distribution)

การแจกแจงเอกซ์โปเนนเชียลเป็นการแจกแจงของตัวแปรสุ่มที่แสดงระยะเวลาระหว่างเหตุการณ์ที่สนใจศึกษา เช่น ระยะเวลาที่รอคอยในการจ่ายเงินที่แผนกซุปเปอร์มาร์เก็ตของห้างสรรพสินค้า แสดงอายุการใช้งาน(Life time) ของวัตถุสิ่งของ เช่น อายุการใช้งานของเครื่องใช้ไฟฟ้า

ตัวแปรสุ่มต่อเนื่อง x ที่มีการแจกแจงเอกซ์โปเนนเชียลพารามิเตอร์ $\theta > 0$ จะมีฟังก์ชันความหนาแน่นในรูป

$$f(x) = \frac{1}{\theta} e^{-\frac{x}{\theta}} \qquad \text{เมื่อ } x > 0$$

$$= 0 \qquad \text{เมื่อ } x \text{ มีค่าอื่น ๆ}$$



รูปที่ 2.3 กราฟแสดงการแจกแจงแบบเอกซ์โปเนนเชียล

2.6.1 คุณสมบัติของการแจกแจงเอกซ์โปเนนเชียล

2.6.1.1 ค่าเฉลี่ยและความแปรปรวนของตัวแปรสุ่มที่มีการแจกแจงเอกซ์โปเนนเชียลจะขึ้นอยู่กับ พารามิเตอร์ θ โดยที่ ค่าเฉลี่ยเท่ากับ θ และความแปรปรวนเท่ากับ θ^2

2.6.1.2 ตัวแปรสุ่มเอกซ์โปเนนเชียล มีคุณสมบัติ Memoryless

นั่นคือ ถ้า x เป็นตัวแปรเกี่ยวกับอายุการใช้งานของเครื่องใช้

$$P(X \geq x+y \mid X > x) = P(X > y)$$

สมมติว่าอายุการใช้งานของหลอดไฟฟ้ายี่ห้อหนึ่ง มีการแจกแจงเอกซ์โปเนนเชียล ที่มีอายุการใช้งานเฉลี่ยเท่ากับ 1,000 ชั่วโมง

ให้ x แทนอายุการใช้งานของหลอดไฟฟ้า

$$P(X > x) = \int_x^{\infty} \frac{1}{1000} e^{-t/1000} \, dt$$

$$= e^{-x/1000}$$

ถ้าใช้หลอดไฟฟ้าหลอดหนึ่งมาแล้ว 500 ชั่วโมง

ความน่าจะเป็นที่หลอดไฟฟ้าหลอดนี้จะมีอายุการใช้งานต่อไปอีก อย่างน้อย 800 ชั่วโมง นั่นคือ

$$P(X \geq 1300 \mid X > 500) = \frac{P(X > 1300)}{P(X > 500)}$$

$$= \frac{e^{-(1300/1000)}}{e^{-(500/1000)}}$$

$$= e^{-(8/10)}$$

ซึ่งเท่ากับ $P(X > 800) = e^{-(8/10)}$

ดังนั้น $P(X > 1300 \mid X > 500) = P(X > 800)$

แสดงให้เห็นว่าความน่าจะเป็นที่หลอดไฟฟ้านั้นจะมีอายุการใช้งานอย่างน้อย 1,300 ชั่วโมง เมื่อถูกใช้มาแล้วอย่างน้อย 500 ชั่วโมง จะมีค่าเท่ากับความน่าจะเป็นที่หลอดไฟฟ้านี้จะมีอายุการใช้งานอย่างน้อย 800 ชั่วโมง นับตั้งแต่เริ่มใช้งาน

2.7 การแจกแจงแบบไวบูลล์ (Weibull Distribution)

Waloddi Weibull นักฟิสิกส์ชาวสวีเดนได้แนะนำการแจกแจงนี้เมื่อ ค.ศ.1939 เป็นการแจกแจงที่เกิดขึ้น เนื่องจากรูปแบบความเป็นจริงโดยทั่วไป สำหรับอายุการใช้งานของเครื่องจักรกลต่าง ๆ นอกจากนี้ยังมีประโยชน์มากในทางทฤษฎีความเชื่อถือได้ (reliability) โดยมีฟังก์ชันความหนาแน่นและลักษณะของการแจกแจงดังนี้

$$f(x) = \frac{\alpha}{\beta}\left(\frac{x-\nu}{\beta}\right)^{\alpha-1} \exp\left[-\left(\frac{x-\nu}{\beta}\right)^{\alpha}\right]; \quad x \geqslant \nu$$

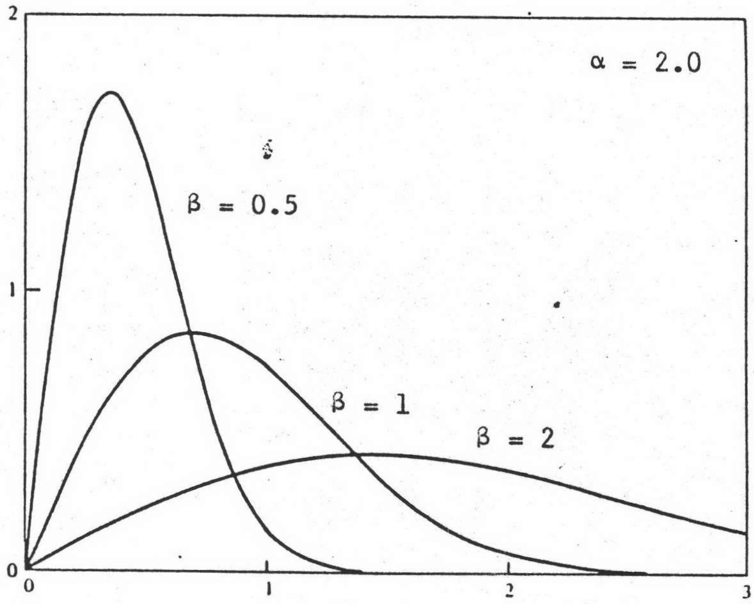
$$= 0 \; ; \qquad x < \nu$$

$$\alpha > 0; \quad \beta > 0 \; ; \quad \nu \geqslant 0$$

α เป็นพารามิเตอร์ที่แสดงรูปร่างของการแจกแจง (Shape parameter)

β เป็นพารามิเตอร์ที่แสดงถึงขนาดของการแจกแจง (Scale parameter)

ν เป็นพารามิเตอร์ที่แสดงตำแหน่งของการแจกแจง (Location parameter)



<u>รูปที่ 2.</u>4 การแจกแจงแบบไวบูลล์ เมื่อ $\alpha = 2.0$
และ $\beta = 0.5,\ 1,\ 2$

<u>2.7.1 คุณสมบัติของการแจกแจงแบบไวบูลล์</u>

2.7.1.1 โค้งมีลักษณะเปลี่ยนแปลงไปตามพารามิเตอร์ α เมื่อ $\alpha = 2$ โค้งจะมีลักษณะเบ้ขวา

2.7.1.2 ค่าเฉลี่ยของตัวแปรสุ่มไวบูลล์ x ขึ้นกับค่าพารามิเตอร์ α และ β และ ν

โดยที่ ค่าเฉลี่ย = $\beta\Gamma(1 + \frac{1}{\alpha}) + \nu$

2.7.1.3 ค่าความแปรปรวน = $\beta^2\{\Gamma(1 + \frac{2}{\alpha}) - [\Gamma(1 + \frac{1}{\alpha})]^2\}$

2.7.1.3 การแจกแจงไวบูลล์ที่มี $\alpha = 1$ และ β เป็นค่าใด ๆ ที่มากกว่า 0 และ $\nu = 0$ นั้นจะเหมือนกับการแจกแจงแบบเอกซ์โปเนนเชียลที่มีพารามิเตอร์เป็น $\theta = \frac{1}{\beta}$

2.8 การแจกแจงที (t-distribution)

ในการวิจัยที่กลุ่มตัวอย่างมีขนาดเล็กและไม่ทราบค่าความแปรปรวนของประชากร จะอาศัยตัวสถิติ $t = \frac{x - \mu}{s/\sqrt{n}}$ ซึ่งมีการแจกแจงแบบที ในการอนุมานทางสถิติ

William Sealy Gosset (1876-1937) นักสถิติชาวอังกฤษเสนอผลงานของเขาเกี่ยวกับการแจกแจงแบบที ในวารสาร Biometrica เรื่อง "The Probable Error of a Mean" โดยใช้นามแฝงว่า "Student"

การแจกแจงแบบที เกี่ยวข้องกับการแจกแจงแบบปกติมาตรฐานและการแจกแจงแบบแกมมา [1] ดังต่อไปนี้

ถ้า x เป็นตัวแปรสุ่มซึ่งมีการแจกแจงแบบปกติมาตรฐาน

v เป็นตัวแปรสุ่มซึ่งมีการแจกแจงแบบแกมมาด้วยพารามิเตอร์ $\theta = 1/2$

และ $r = n/2$ เมื่อ n เป็นจำนวนเต็มและ $n > 1$

$T = \frac{X}{\sqrt{V/n}}$ จะมีการแจกแจงแบบที มี degree of freedom = n

ซึ่งมีฟังก์ชันการแจกแจง

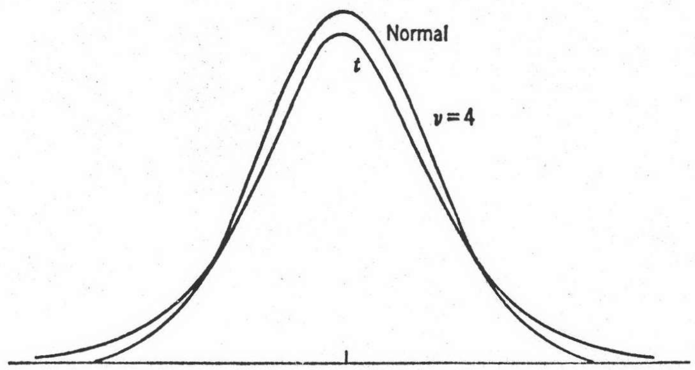
$$f(t) = \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})\ \sqrt{n\pi}\ (1 + \frac{t^2}{n})^{(\frac{n+1}{2})}} \quad , \quad -\infty < t < \infty$$

---

[1] X เป็นตัวแปรสุ่มที่มีการแจกแจงแบบแกมมา (Gamma distribution) ซึ่งมีพารามิเตอร์ $\Theta > 0, \ r > 0$ จะมีฟังก์ชันความหนาแน่นอยู่ในรูป

$$f(x) = \frac{\theta^r y^{r-1} e^{-\theta x}}{\Gamma(r)} \quad ; \quad x \geqslant 0$$

$$= 0 \quad ; x < 0$$



รูปที่ 2.5 การแจกแจงแบบที ที่ระดับองศาความเป็นอิสระ = 4

Karl Pearson (1857-1936) นักสถิติชาวอังกฤษได้แสดงว่าการแจกแจงของตัวแปรสุ่ม V ที่มีการแจกแจงแบบแกมมาด้วยพารามิเตอร์ $\theta=1/2$ และ $r = n/2$ เมื่อ n เป็นจำนวนเต็มบวก ก็คือการแจกแจงแบบไคสแควร์ ที่มี degree of freedom เท่ากับ n (Olkin,1980:392)

2.8.1 คุณสมบัติของการแจกแจงแบบที

2.8.1.1 กราฟของการแจกแจงแบบทีมีลักษณะสมมาตร ณ จุดที่ $t = 0$

2.8.1.2 ค่าเฉลี่ยเท่ากับ 0

2.8.1.3 มีค่าความแปรปรวน $= n/(n-2)$ เมื่อ $n > 2$

2.8.1.4 ถ้า n มีค่ามากการแจกแจงใกล้เคียงกับการแจกแจงแบบปกติมาตรฐาน

2.9 ผลงานวิจัยที่เกี่ยวข้อง

Games(1972:887-909) ศึกษาเกี่ยวกับการทดสอบความเท่ากันของความแปรปรวนของประชากรสามประชากร โดยใช้สถิติทดสอบบาร์ตเลต (Bartlett's test) สถิติทดสอบฮาร์ตเลย์ (Hartley's test) สถิติทดสอบคอครัน (Cochran's test) สถิติทดสอบ L-A สถิติทดสอบ L-X^2 เมื่อใช้กลุ่มตัวอย่างขนาด 6 และ 18 เมื่อประชากรมีการแจกแจงแบบปกติและแบบเบ้ พบว่า ในกรณีกลุ่มตัวอย่างที่มีขนาด 6 ถ้าประชากรมีการแจกแจงแบบปกติและความแปรปรวนไม่เท่ากันทั้งหมด สถิติทดสอบบาร์ตเลต จะมีอำนาจทดสอบสูงที่สุด และถ้าประชากรมีการแจกแจงแบบเบ้เพียงเล็กน้อยก็จะไม่มีผลต่ออำนาจการทดสอบของสถิติทดสอบ

Layard (1973:195-198) ศึกษาความแกร่งของสถิติทดสอบความเท่ากันของความแปรปรวนของประชากรสี่ประชากร กรณีขนาดตัวอย่าง 10 และ 25 โดยศึกษาตัวสถิติทดสอบบาร์ตเลต (Bartlett's test) สถิติทดสอบ χ^2 (Layard Chi Square) สถิติทดสอบ Box (Box's test) สถิติทดสอบแจคไนฟ์ (Jacknife's test) ซึ่งกำหนดรูปแบบการแจกแจงของประชากรเป็นแบบปกติ การแจกแจงแบบยูนิฟอร์ม (Uniform distribution) การแจกแจงแบบดับเบิ้ลเอกซ์โปเนนเชียล (Double Exponential distribution)โดยใช้อัตราส่วนของความแปรปรวนเท่ากับ 1:1:1:1, 1:1:2:2, 1:2:3:4 และ 1:1:4:4 พบว่าสถิติทดสอบบาร์ตเลตมีอำนาจการทดสอบสูงกว่าสถิติทดสอบอื่น ๆ เมื่อประชากรมีการแจกแจงแบบปกติ และเมื่อประชากรมีการแจกแจงไม่ใช่แบบปกติ แล้วสถิติทดสอบแต่ละตัวมีอำนาจการทดสอบไม่แตกต่างกัน

Seber (1977:147-149) กล่าวว่า ในการทดสอบความเท่ากันของความแปรปรวนของประชากรที่มากกว่าสองประชากรนั้นผู้วิจัยมักเลือกใช้สถิติทดสอบ Bartlet's test , Hartley's test หรือ Cochran's test และพบว่า Bartlett's test มีอำนาจการทดสอบสูงกว่าสถิติทดสอบอีกสองตัว ภายใต้ข้อตกลงเบื้องต้นว่าการแจกแจงของประชากรต้องมีการแจกแจงแบบปกติ (normality assumption)

Conover,Johnson and Johnson (1981:351-361) ได้ศึกษาเปรียบเทียบอำนาจการทดสอบของสถิติทดสอบความเท่ากันของความแปรปรวนซึ่งเป็นแบบ parametric testและ

nonparametric test จำนวน 56 วิธี (procedure) พบว่าสถิติทดสอบส่วนมากจะมีความไว (sensitive)ต่อรูปแบบการแจกแจงของประชากร ยกเว้นสถิติทดสอบของ Brown-Forsythe ซึ่งมีความสามารถในการควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้แม้ว่าการแจกแจงของประชากรจะเปลี่ยนแปลงไป

Olejnik and Algina (1987:45-61) ได้ศึกษาเปรียบเทียบค่าความน่าจะเป็นของความคลาดเคลื่อนประเภทที่ 1 และค่าอำนาจการทดสอบของสถิติทดสอบความเท่ากันของความแปรปรวนสำหรับสองประชากร โดยใช้สถิติทดสอบ Brown-Forsythe สถิติทดสอบ O'Brien สถิติทดสอบ Klotz สถิติทดสอบ Siegel-Tukey กรณีที่ประชากรมีการแจกแจงแบบปกติและการแจกแจงที่ไม่ใช่แบบปกติ ปรากฏว่า สถิติทดสอบ O'Brien สามารถควบคุมความคลาดเคลื่อนประเภทที่ 1 ได้ทั้งกรณีที่มีการแจกแจงแบบปกติ และการแจกแจงที่ไม่ใช่แบบปกติ โดยใช้เกณฑ์ของ Cochran และมีอำนาจการทดสอบสูงกว่าสถิติทดสอบอื่น ๆ