บทที่ 2

ทฤษฎีและงานวิจัยที่เกี่ยวข้อง

งานวิจัยครั้งนี้เป็นการทดสอบความเท่ากันของความแปรปรวนของประชากร 2 กลุ่มเมื่อข้อมูลไม่มีการแจกแจงปรกติ ด้วยสถิติทดสอบเอฟ สถิติทดสอบแมน-วิทนีย์ สถิติทดสอบอย่างง่าย สถิติทดสอบอย่างง่ายด้วยบูทสแทรป สถิติทดสอบอย่างง่ายด้วยบูทสแทรปสองชั้น สถิติทดสอบคาเซลลา สถิติทดสอบคาเซลลาด้วยบูทสแทรป และสถิติทดสอบคาเซลลาด้วยบูทสแทรปแบบสองชั้น ในบทนี้จะอธิบาย การแจกแจงที่ใช้ในการศึกษา รายละเอียดของสถิติทดสอบทั้ง 8 วิธี และทฤษฎีที่เกี่ยวข้องกับงานวิจัย

2.1 การแจกแจงที่ใช้ในการศึกษา

2.1.1 การแจกแจงปรกติ

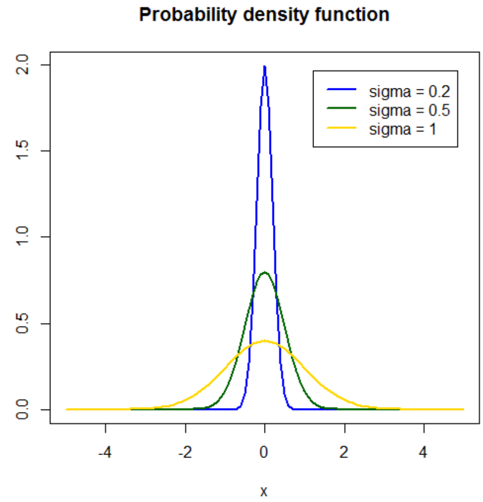
Abraham De Moire นักคณิตศาสตร์ชาวฝรั่งเศสเป็นผู้ค้นพบการแจกแจงปรกติ เมื่อปี ค.ศ. 1733 แต่ไม่เป็นที่รู้จักแพร่หลาย ต่อมา ค.ศ.1749-1855 Carl Guass นักคณิตศาสตร์และดาราศาสตร์ชาวเยอรมัน ค้นพบว่าการแจกแจงของความคลาดเคลื่อนในการวัดทางวิทยาศาสตร์กายภาพสามารถประมาณได้ใกล้เคียงโดยใช้โค้งปรกติ ซึ่งเรียกว่า โค้งปรกติความคลาดเคลื่อน และนำมาใช้กันอย่างแพร่หลาย หรือเรียกอีกอย่างหนึ่งว่า การแจกแจงเกาส์เซียน (Gaussian Distribution) โดยมีฟังก์ชันความน่าจะเป็นดังนี้ (สุชาดา เพ็ชรขาวเขียว, 2555)

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

มีค่าเฉลี่ยและความแปรปรวน ดังนี้

$E(X) = \mu$

$Var(X) = \sigma^2$



ภาพที่ 2.1 แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจงปรกติ

2.1.2 การแจกแจงล็อกปรกติ

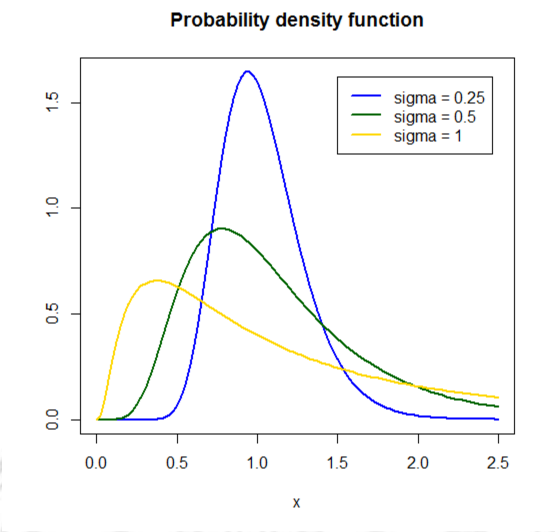
การแจกแจงล็อกปรกติเริ่มเป็นที่รู้จักตั้งแต่ปี ค.ศ. 1879 โดย Galton และ McAlister เป็นนักวิทยาศาสตร์กลุ่มแรกที่ใช้การแจกแจงดังกล่าว การแจกแจงล็อกปรกติถูกนำไปประยุกต์กับเหตุการณ์ต่าง ๆ ในปัจจุบันอย่างกว้างขวาง เช่น ใช้แทนการแจกแจงของปริมาณยาที่ให้ต่อครั้ง และในทางเศรษฐศาสตร์ได้นำการแจกแจงล็อกปรกติไปอธิบายการกระจายรายได้ของประชากร และการกระจายของเงินผลตอบแทนที่บริษัทต้องจ่ายให้กับลูกค้าผู้ทำประกันภัยในแต่ละปี ทางวิศกรรมนำไปอธิบายการกระจายความเร็วลมเป็นต้น โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้ (นภดล วันชัยชนะ, 2553)

$$f(x) = \frac{1}{x\phi\sqrt{2\pi}} e^{-\frac{(\ln x-\mu)^2}{2\phi^2}} \quad ; \quad x > 0$$

มีค่าเฉลี่ยและความแปรปรวน ดังนี้

$$E(X) = e^{\frac{\mu+\phi^2}{2}}$$

$$Var(X) = \left(e^{\phi^2} - 1\right) e^{2\mu+\phi^2}$$



ภาพที่ 2.2 แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจงล็อกปรกติ

2.1.3 การแจกแจงเอ็กซ์โปเนนเชียล

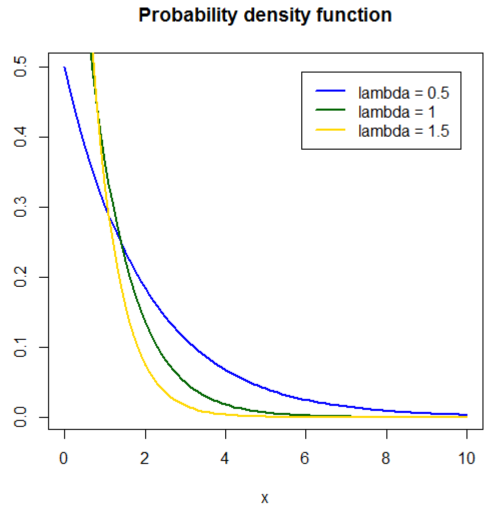
การแจกแจงเอ็กซ์โปเนนเชียลเป็นการแจกแจงแบบต่อเนื่องของตัวแปรสุ่มที่แสดงเวลาที่รอคอยจนกระทั่งเกิดเหตุการณ์แรกหรือเหตุการณ์ที่สนใจ เช่น ระยะเวลา (ตั้งแต่ปัจจุบัน) จนกระทั่งเกิดแผ่นดินไหว หรือเวลา (นาที) ที่ลูกค้ารอจนได้รับการบริการในธนาคารแห่งหนึ่ง หรืออายุการใช้งานของเครื่องใช้ไฟฟ้า โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้ (นภดล วันชัยชนะ, 2553)

$$f(x) = \lambda e^{-\lambda x} \qquad ; \qquad 0 < x < \infty,\ \lambda > 0$$

มีค่าเฉลี่ยและความแปรปรวน ดังนี้

$$E(X) = \frac{1}{\lambda}$$

$$Var(X) = \frac{1}{\lambda^2}$$



ภาพที่ 2.3 แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจงเอ็กซ์โปเนนเชียล

2.1.4 การแจกแจงไวบูล

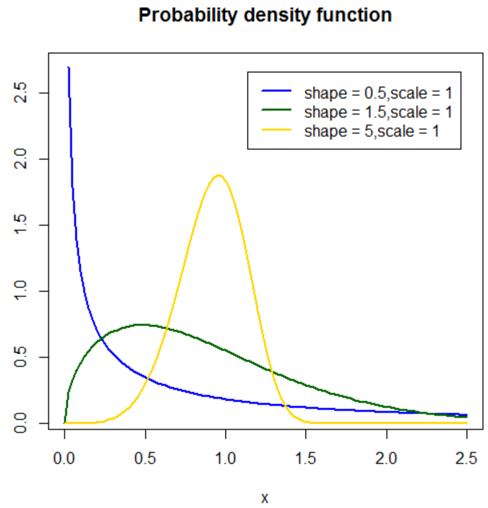
การแจกแจงไวบูลแสดงอายุการใช้งานของวัตถุสิ่งของโดยวัดเวลาตั้งแต่เริ่มต้น จนกระทั่งวัตถุนั้นเสีย เสื่อมสภาพ หรือเริ่มนับเมื่อปรากฏการณ์ทำนองเดียวกันได้เกิดขึ้นเสร็จสิ้นไปแล้วในเวลาดังกล่าว โดยจะนับติดต่อกันไปเรื่อย ๆ จนกระทั่งปรากฏการณ์ที่รอคอยนั้นจะเกิดขึ้นอีก การแจกแจงไวบูลเคยใช้อธิบายการเสื่อมสภาพของหลอดสูญญากาศ โดย Kao ในปี 1958 โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้ (ศรสวรรค์ บุญเพ็ญ และคณะ, 2558)

$$f(x) = \left(\frac{\gamma}{\beta}\right)\left(\frac{x}{\beta}\right)^{\gamma-1} e^{-(x/\beta)^{\gamma}} \quad ; \; 0 \le x < \infty, \gamma > 0, \beta > 0$$

มีค่าเฉลี่ยและความแปรปรวน ดังนี้

$$E(X) = \beta\Gamma\left(1+\frac{1}{\gamma}\right)$$

$$Var(X) = \beta^2\left[\Gamma\left(1+\frac{2}{\gamma}\right) - \left(\Gamma\left(1+\frac{1}{\gamma}\right)\right)^2\right]$$



ภาพที่ 2.4 แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจงไวบูล

2.1.5 การแจกแจงโลจิสติก

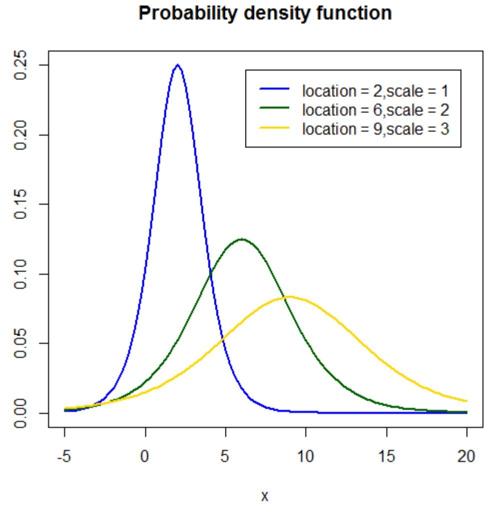
การแจกแจงโลจิสติก มาจากการศึกษาของ Pierre Francois Verhulst ค.ศ. 1804-1849 ผู้เชี่ยวชาญในวิทยาลัยการทหารของเบลเยียม โดยศึกษารูปแบบการเพิ่มขึ้นของประชากรของเบลเยียมก่อนปี ค.ศ. 1800 อธิบายการเพิ่มขึ้นของประชากรตามฟังก์ชันการแจกแจงความถี่ของการแจกแจงโลจิสติก การแจกแจงโลจิสติกมีลักษณะคล้ายการแจกแจงปรกติ แต่มีหางยาวกว่าการแจกแจงปรกติและมีความโด่งมากกว่า โดยมีฟังก์ชันความหนาแน่นน่าจะเป็นดังนี้ (สมประสงค์ สิทธิสมบัติ, 2550)

$$f(x) = \frac{e^{-(x-\mu)/\delta}}{\delta\left(1+e^{-(x-\mu)/\delta}\right)^2} \quad ; \quad -\infty < x < \infty\,,\ -\infty < \mu < \infty\,,\ \delta > 0$$

มีค่าเฉลี่ยและความแปรปรวน ดังนี้

$$E(X) = \mu$$

$$Var(X) = \frac{\delta^2\pi^2}{3}$$



ภาพที่ 2.5 แสดงฟังก์ชันความหนาแน่นน่าจะเป็นของการแจกแจงโลจิสติก

2.2 ตัวสถิติทดสอบความเท่ากันของความแปรปรวน

2.2.1 สถิติทดสอบเอฟ

สถิติทดสอบเอฟ ถูกค้นพบโดย Ronald Aylmer Fisher เป็นวิธีการทดสอบความเท่ากันของความแปรปรวนของประชากร 2 กลุ่มโดยตัวแปรสุ่มจากประชากรทั้ง 2 กลุ่มที่เป็นอิสระกันและมีการแจกแจงปรกติ โดยกำหนดให้ $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากรกลุ่มที่ 1 และกลุ่มที่ 2 และขนาดตัวอย่างเท่ากับ n_1 และ n_2 ตามลำดับ (ประชุม สุวัตถี, 2545)

$$
\begin{aligned}
F_c &= \frac{\chi^2_{\upsilon_1}/\upsilon_1}{\chi^2_{\upsilon_2}/\upsilon_2} \\
&= \frac{(n_1-1)S_1^2/\sigma_1^2}{n_1-1} \Bigg/ \frac{(n_2-1)S_2^2/\sigma_2^2}{n_2-1} \\
&= \frac{S_1^2}{S_2^2} \Bigg/ \frac{\sigma_1^2}{\sigma_2^2}
\end{aligned}
$$

ปฏิเสธสมมติฐานหลักเมื่อ $f_c < f_{\frac{\alpha}{2},\upsilon_1,\upsilon_2}$ หรือ $f_c > f_{1-\frac{\alpha}{2},\upsilon_1,\upsilon_2}$

โดย υ_2 และ υ_2 คือ องศาความเป็นอิสระของการแจกแจงเอฟ ซึ่งมีค่าเท่ากับ $n_1 - 1$ และ $n_2 - 1$ ตามลำดับ

2.2.2 สถิติทดสอบแมน - วิทนีย์

ในปี 1945 - 1947 Wilcox ได้เสนอสถิติทดสอบวิลคอกสัน-แมนวิทนีย์ (Wilcoxon-Mann-Whitney Test) หรือ สถิติทดสอบผลรวมของอันดับแมน - วิทนีย์ (The Mann-Whitney Ranked Sum Test) หรือบางครั้งเรียกว่า การทดสอบยู (U Test) ซึ่งเป็นการทดสอบที่ใช้ผลรวมของอันดับ โดยสถิติทดสอบแมน – วิทนีย์ เป็นสถิติไม่อิงพารามิเตอร์ที่มีสมบัติการทดสอบใกล้เคียงสถิติอิงพารามิเตอร์ ซึ่งมีประสิทธิภาพในการทดสอบสูง จะมีความไว (Sensitivity) ต่อการปฏิเสธสมมติฐานหลัก ขึ้นอยู่กับอัตราส่วนของขนาดตัวอย่างและความแปรปรวนของข้อมูล (Wilcox, 2012)

การทดสอบผลรวมของอันดับของแมน – วิทนีย์ กรณีใช้สำหรับทดสอบความแปรปรวนของตัวแปรสุ่มจากประชากรทั้ง 2 กลุ่มที่เป็นอิสระกัน โดยที่มาตรวัดของข้อมูลอย่างน้อยต้องอยู่ในมาตรเรียงลำดับขึ้นไป กำหนดให้ $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ เป็นข้อมูลกลุ่มที่ 1 และ 2 เมื่อขนาดตัวอย่างเท่ากับ n_1 และ n_2 ตามลำดับ โดยจะพิจารณาจากอันดับที่ให้แก่ข้อมูล โดยทำการให้อันดับที่ 1 สำหรับข้อมูลที่น้อยที่สุด และอันดับ 2 สำหรับข้อมูลที่มีค่ามากที่สุด จากนั้นก็สลับมาเป็นให้อันดับ 3 สำหรับข้อมูลที่มากรองลงมา และอันดับที่ 4 สำหรับข้อมูลที่น้อยรองลงมา ทำแบบนี้ไปจนถึงค่ากลางในอันดับที่ $n_1 + n_2$ และถ้ากรณีที่ข้อมูลมีค่าเท่ากันก็จะทำการเฉลี่ยอันดับให้กับข้อมูลที่เท่ากัน จากนั้นคำนวณสถิติทดสอบผลรวมของอันดับแมน – วิทนีย์ ดังนี้

$$W = S - \frac{n_1(n_1+1)}{2}$$

เมื่อ S คือ ผลรวมอันดับของกลุ่มที่ 1

n_1 คือ ขนาดตัวอย่างของกลุ่มที่ 1

จะปฏิเสธ H_0 เมื่อ $w < w_{\frac{\alpha}{2}, n_1}$ หรือ $w < w_{1-\frac{\alpha}{2}, n_1}$ โดย $w_{\frac{\alpha}{2}, n_1}$ และ $w < w_{1-\frac{\alpha}{2}, n_1}$ เป็นค่าที่เปิดได้จากตารางแมน – วิทนีย์ หรือ ปฏิเสธ H_0 เมื่อ p-value < α

กรณีขนาดตัวอย่างมากจะใช้ทฤษฎีขีดจำกัดเข้าสู่ส่วนกลาง โดยที่ $\mu_w = \frac{n_1 n_2}{2}$,

$$\sigma_w^2 = \frac{n_1 n_2}{2} + Z_p \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

2.2.3 สถิติทดสอบอย่างง่าย

วิธีอย่างง่ายมีหลักการคือ ภายใต้ $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงปรกติที่มีค่าเฉลี่ย μ และความแปรปรวนคือ σ^2 เมื่อ $\bar{X}$ และ S^2 เป็นตัวประมาณไม่เอนเอียงที่มีความแปรปรวนต่ำสุด ของ μ และ σ^2 ตามลำดับ (ประชุม สุวัตถี, 2545) สามารถคำนวณสถิติทดสอบได้ดังนี้

$$T_s = \frac{(S_1^2 - S_2^2) - (\sigma_1^2 - \sigma_2^2)}{\sqrt{\hat{V}_s(S_1^2) + \hat{V}_s(S_1^2)}}$$

(Cojbasic, 2012)

โดยที่ S_1^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 1

S_2^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 2

$\hat{V}_s(S_1^2)$ หมายถึง ค่าประมาณความแปรปรวนของ S_1^2 โดยวิธีอย่างง่าย

$\hat{V}_s(S_2^2)$ หมายถึง ค่าประมาณความแปรปรวนของ S_2^2 โดยวิธีอย่างง่าย

เมื่อ

$$S_1^2 = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i} - \bar{X}_1)^2$$

$$S_2^2 = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i} - \bar{X}_2)^2$$

เนื่องจาก $V(S^2) = \dfrac{2\sigma^4}{n-1}$ (Casella, 2001)

$$\hat{V}_s(S_1^2) = \frac{2S_1^4}{n_1 - 1}$$

$$\hat{V}_s(S_2^2) = \frac{2S_2^4}{n_2 - 1}$$

ปฏิเสธสมมติฐานหลักเมื่อ $t_s < t_{\frac{\alpha}{2}, \upsilon}$ หรือ $t_s > t_{1-\frac{\alpha}{2}, \upsilon}$

โดย υ คือองศาความเป็นอิสระของการแจกแจงที ซึ่งมีค่าเท่ากับ $n_1 + n_2 - 2$

2.2.4 สถิติทดสอบอย่างง่ายด้วยบูทสแทรป

ให้ $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากร ที่มีขนาดตัวอย่างเท่ากับ n_1 และ n_2 ตามลำดับ ทำการสุ่มตัวอย่างทีละ 1 ค่าแบบใส่คืน จำนวน n_1 และ n_2 ครั้ง จากชุดตัวอย่าง $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ จะได้ชุดตัวอย่างจากวิธีการบูทสแทรป คือ $X^*_{11}, X^*_{12}, \ldots, X^*_{1n_1}$ และ $X^*_{21}, X^*_{22}, \ldots, X^*_{2n_2}$

การทดสอบความเท่ากันของความแปรปรวนสำหรับประชากรสองกลุ่มโดยสถิติทดสอบอย่างง่ายด้วยบูทสแทรปคำนวณได้ดังนี้

$$T_{s^*} = \frac{(S_1^{*2} - S_2^{*2}) - (\sigma_1^2 - \sigma_2^2)}{\sqrt{\hat{V}_s(S_1^{*2}) + \hat{V}_s(S_2^{*2})}}$$

โดยที่ S_1^{*2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 1 โดยวิธีบูทสแทรป

S_2^{*2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 2 โดยวิธีบูทสแทรป

$\hat{V}_s(S_1^{*2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_1^{*2} โดยวิธีอย่างง่ายด้วยบูทสแทรป

$\hat{V}_s(S_2^{*2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_2^{*2} โดยวิธีอย่างง่ายด้วยบูทสแทรป

เมื่อ

$$S_1^{*2} = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X^*_{1i} - \bar{X}^*_1)^2$$

$$S_2^{*2} = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X^*_{2i} - \bar{X}^*_2)^2$$

$$\hat{V}_s(S_1^{*2}) = \frac{2S_1^{*4}}{n_1 - 1}$$

$$\hat{V}_s(S_2^{*2}) = \frac{2S_2^{*4}}{n_2 - 1}$$

ปฏิเสธสมมติฐานหลักเมื่อ $t_{s^*} < t_{\frac{\alpha}{2},\upsilon}$ หรือ $t_{s^*} > t_{1-\frac{\alpha}{2},\upsilon}$

โดย υ คือ องศาความเป็นอิสระของการแจกแจงที ซึ่งมีค่าเท่ากับ $n_1 + n_2 - 2$

$P-value = \dfrac{\sum_{b=1}^{B} \mathrm{I}_{(t_{1-\frac{\alpha}{2},\upsilon,\infty})}(|t_{s^*}|)}{B}$ ปฏิเสธสมมติฐานหลักเมื่อ $P-value < \alpha$

2.2.5 สถิติทดสอบอย่างง่ายด้วยบูทสแทรปสองชั้น

ให้ $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากร ที่มีขนาดตัวอย่างเท่ากับ n_1 และ n_2 ทำการสุ่มตัวอย่างแบบใส่คืนทีละ 1 ค่า จำนวน n_1 และ n_2 ครั้ง จากชุดตัวอย่าง $X_{11}, X_{12}, ..., X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ ตามลำดับ ดังนั้นชุดตัวอย่างจากวิธีการบูทสแทรปหนึ่งชั้น คือ $X_{11}^*, X_{12}^*, \ldots, X_{1n_1}^*$ และ $X_{21}^*, X_{22}^*, \ldots, X_{2n_2}^*$ จากนั้นทำการสุ่มตัวอย่างแบบใส่คืน n_1 และ n_2 ครั้ง จากชุดตัวอย่าง $X_{11}^*, X_{12}^*, \ldots, X_{1n_1}^*$ และ $X_{21}^*, X_{22}^*, \ldots, X_{2n_2}^*$ จะได้ชุดตัวอย่างจากวิธีการบูทสแทรปสองชั้น คือ $X_{11}^{**}, X_{12}^{**}, \ldots, X_{1n_1}^{**}$ และ $X_{21}^{**}, X_{22}^{**}, \ldots, X_{2n_2}^{**}$

การทดสอบความเท่ากันของความแปรปรวนสำหรับประชากรสองกลุ่มโดยสถิติทดสอบอย่างง่ายด้วยบูทสแทรปสองชั้นคำนวณได้ดังนี้

$$T_{s^{**}} = \frac{(S_1^{**2} - S_2^{**2}) - (\sigma_1^2 - \sigma_2^2)}{\sqrt{\hat{V}_s(S_1^{**2}) + \hat{V}_s(S_2^{**2})}}$$

โดยที่ S_1^{**2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 1 โดยวิธีบูทสแทรปสองชั้น

S_2^{**2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 2 โดยวิธีบูทสแทรปสองชั้น

$\hat{V}_s(S_1^{*2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_1^{**2} โดยวิธีอย่างง่ายด้วยบูทสแทรปสองชั้น

$\hat{V}_s(S_2^{*2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_2^{**2} โดยวิธีอย่างง่ายด้วยบูทสแทรปสองชั้น

เมื่อ

$$S_1^{**2} = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i}^{**} - \bar{X}_1^{**})^2$$

$$S_2^{**2} = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i}^{**} - \bar{X}_2^{**})^2$$

$$\hat{V}_s(S_1^{**2}) = \frac{2S_1^{**4}}{n_1 - 1}$$

$$\hat{V}_s(S_2^{**2}) = \frac{2S_2^{**4}}{n_2 - 1}$$

ปฏิเสธสมมติฐานหลักเมื่อ $t_{s^{**}} < t_{\frac{\alpha}{2}, \upsilon}$ หรือ $t_{s^{**}} > t_{1-\frac{\alpha}{2}, \upsilon}$

โดย υ คือองศาความเป็นอิสระของการแจกแจงที ซึ่งมีค่าเท่ากับ $n_1 + n_2 - 2$

$$P-value = \frac{\sum_{b=1}^{B} \mathrm{I}_{(t_{1-\frac{\alpha}{2},\upsilon,\infty})}(|t_{s^{**}}|)}{B}$$ ปฏิเสธสมมติฐานหลักเมื่อ $P-value<\alpha$

2.2.6 สถิติทดสอบคาเซลลา

วิธีนี้มีหลักการคือ ให้ $X_{11}, X_{12},..., X_{1n_1}$ และ $X_{21}, X_{22},\ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากรที่มีการแจกแจงปรกติ มีค่าเฉลี่ย μ และความแปรปรวน σ^2 โดยที่ $\bar{X}$ และ S^2 เป็นตัวประมาณ ของ μ และ σ^2 ตามลำดับ และคาเซลลา ได้แสดง $V(S^2)=\frac{1}{n}\left(\mu_4-\frac{n-3}{n-1}\sigma^4\right)$ เมื่อ $\mu_4 = E(X-\mu)^4$ (Casella, 2001) การทดสอบความเท่ากันของความแปรปรวนสำหรับประชากรสองกลุ่มโดยสถิติทดสอบคาเซลลาคำนวณได้ดังนี้

$$T_c = \frac{(S_1^2-S_2^2)-(\sigma_1^2-\sigma_2^2)}{\sqrt{\hat{V}_c(S_1^2)+\hat{V}_c(S_2^2)}}$$

โดยที่ S_1^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 1

S_2^2 หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 2

$\hat{V}_c(S_1^2)$ หมายถึง ค่าประมาณความแปรปรวนของ S_1^2 โดยวิธีคาเซลลา

$\hat{V}_c(S_2^2)$ หมายถึง ค่าประมาณความแปรปรวนของ S_2^2 โดยวิธีคาเซลลา

เมื่อ

$$S_1^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}(X_{1i}-\bar{X}_1)^2$$

$$S_2^2 = \frac{1}{n_2-1}\sum_{i=1}^{n_2}(X_{2i}-\bar{X}_2)^2$$

$$\hat{V}_c(S_1^2) = \frac{1}{n_1}\left(\hat{\mu}_{4(1)}-\frac{n_1-3}{n_1-1}S_1^4\right)$$

$$\hat{V}_c(S_2^2) = \frac{1}{n_2}\left(\hat{\mu}_{4(2)}-\frac{n_2-3}{n_2-1}S_2^4\right)$$

$$\hat{\mu}_{4(1)} = \frac{1}{n_1}\sum_{i=1}^{n_1}\left(X_{1i}-\bar{X}_1\right)^4$$

$$\hat{\mu}_{4(2)} = \frac{1}{n_2}\sum_{i=1}^{n_2}\left(X_{2i}-\bar{X}_2\right)^4$$

$$S_1^4 = \left(S_1^2\right)^2$$

$$S_2^4 = \left(S_2^2\right)^2$$

ปฏิเสธสมมติฐานหลักเมื่อ $t_c < t_{\frac{\alpha}{2},\upsilon}$ หรือ $t_c > t_{1-\frac{\alpha}{2},\upsilon}$

โดย υ คือองศาความเป็นอิสระของการแจกแจงที ซึ่งมีค่าเท่ากับ $n_1 + n_2 - 2$

2.2.7 สถิติทดสอบคาเซลลาด้วยบูทสแทรป

ให้ $X_{11}, X_{12}, \ldots, X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากร ที่มีขนาดตัวอย่างเท่ากับ n_1 และ n_2 ตามลำดับ ทำการสุ่มตัวอย่างทีละ 1 ค่าแบบใส่คืน จำนวน n_1 และ n_2 ครั้ง จากชุดตัวอย่าง $X_{11}, X_{12}, ..., X_{1n_1}$ และ $X_{21}, X_{22}, \ldots, X_{2n_2}$ จะได้ชุดตัวอย่างจากวิธีการบูทสแทรป คือ $X_{11}^*, X_{12}^*, ..., X_{1n_1}^*$ และ $X_{21}^*, X_{22}^*, \ldots, X_{2n_2}^*$

การทดสอบความเท่ากันของความแปรปรวนสำหรับประชากรสองกลุ่มโดยสถิติทดสอบคาเซลลาด้วยบูทสแทรปคำนวณได้ดังนี้

$$T_{C^*} = \frac{(S_1^{*2} - S_2^{*2}) - (\sigma_1^2 - \sigma_2^2)}{\sqrt{\hat{V}_c(S_1^{*2}) + \hat{V}_c(S_2^{*2})}}$$

โดยที่ S_1^{*2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 1 โดยวิธีบูทสแทรป

S_2^{*2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 2 โดยวิธีบูทสแทรป

$\hat{V}_c(S_1^{*2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_1^{*2} โดยวิธีคาเซลลาด้วยบูทสแทรป

$\hat{V}_c(S_2^{*2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_2^{*2} โดยวิธีคาเซลลาด้วยบูทสแทรป

เมื่อ

$$S_1^{*2} = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i}^* - \bar{X}_1^*)^2$$

$$S_2^{*2} = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i}^* - \bar{X}_2^*)^2$$

$$\hat{V}_c(S_1^{*2}) = \frac{1}{n_1}\left(\hat{\mu}_{4(1)}^* - \frac{n_1 - 3}{n_1 - 1}S_1^{*4}\right)$$

$$\hat{V}_c(S_2^{*2}) = \frac{1}{n_2}\left(\hat{\mu}_{4(2)}^* - \frac{n_1 - 3}{n_1 - 1}S_2^{*4}\right)$$

$$\hat{\mu}_{4(1)}^* = \frac{1}{n_1}\sum_{i=1}^{n_1}\left(X_{1i}^* - \bar{X}_1^*\right)^4$$

$$\hat{\mu}_{4(2)}^* = \frac{1}{n_2}\sum_{i=1}^{n_2}\left(X_{2i}^* - \bar{X}_2^*\right)^4$$

$$S_1^{*4} = \left(S_1^{*2}\right)^2$$
$$S_2^{*4} = \left(S_2^{*2}\right)^2$$

ปฏิเสธสมมติฐานหลักเมื่อ $t_{C^*} < t_{\frac{\alpha}{2},\upsilon}$ หรือ $t_{C^*} > t_{1-\frac{\alpha}{2},\upsilon}$

โดย υ คือองศาความเป็นอิสระของการแจกแจงที ซึ่งมีค่าเท่ากับ $n_1 + n_2 - 2$

$$P-value = \frac{\sum_{b=1}^{B} \mathrm{I}_{(t_{1-\frac{\alpha}{2},\upsilon,\infty})}(|t_{c^*}|)}{B}$$ ปฏิเสธสมมติฐานหลักเมื่อ $P-value<\alpha$

2.2.8 สถิติทดสอบคาเซลลาด้วยบูทสแทรปสองชั้น

ให้ $X_{11}, X_{12},\ldots, X_{1n_1}$ และ $X_{21}, X_{22},\ldots, X_{2n_2}$ เป็นตัวอย่างสุ่มจากประชากร ที่มีขนาดตัวอย่างเท่ากับ n_1 และ n_2 ทำการสุ่มตัวอย่างแบบใส่คืนทีละ 1 ค่า จำนวน n_1 และ n_2 ครั้ง จากชุดตัวอย่าง $X_{11}, X_{12},..., X_{1n_1}$ และ $X_{21}, X_{22},\ldots, X_{2n_2}$ ตามลำดับ ดังนั้นชุดตัวอย่างจากวิธีการบูทสแทรปหนึ่งชั้น คือ $X_{11}^*, X_{12}^*,\ldots, X_{1n_1}^*$ และ $X_{21}^*, X_{22}^*,\ldots, X_{2n_2}^*$ จากนั้นทำการสุ่มตัวอย่างแบบใส่คืน n_1 และ n_2 ครั้ง จากชุดตัวอย่าง $X_{11}^*, X_{12}^*,\ldots, X_{1n_1}^*$ และ $X_{21}^*, X_{22}^*,\ldots, X_{2n_2}^*$ จะได้ชุดตัวอย่างจากวิธีการบูทสแทรปสองชั้น คือ $X_{11}^{**}, X_{12}^{**},\ldots, X_{1n_1}^{**}$ และ $X_{21}^{**}, X_{22}^{**},\ldots, X_{2n_2}^{**}$

การทดสอบความเท่ากันของความแปรปรวนสำหรับประชากรสองกลุ่มโดยสถิติทดสอบคาเซลลาด้วยบูทสแทรปสองชั้นคำนวณได้ดังนี้

$$T_{C^{**}} = \frac{(S_1^{**2} - S_2^{**2}) - (\sigma_1^2 - \sigma_2^2)}{\sqrt{\hat{V}_c(S_1^{**2}) + \hat{V}_c(S_2^{**2})}}$$

โดยที่ S_1^{**2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 1 โดยวิธีบูทสแทรปสองชั้น
S_2^{**2} หมายถึง ความแปรปรวนของกลุ่มตัวอย่างที่ 2 โดยวิธีบูทสแทรปสองชั้น
$\hat{V}_c(S_1^{**2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_1^{**2} โดยวิธีคาเซลลาด้วยบูทสแทรปสองชั้น
$\hat{V}_c(S_2^{**2})$ หมายถึง ค่าประมาณความแปรปรวนของ S_2^{**2} โดยวิธีคาเซลลาด้วยบูทสแทรปสองชั้น

เมื่อ

$$S_1^{**2} = \frac{1}{n_1 - 1}\sum_{i=1}^{n_1}(X_{1i}^{**} - \bar{X}_1^{**})^2$$

$$S_2^{**2} = \frac{1}{n_2 - 1}\sum_{i=1}^{n_2}(X_{2i}^{**} - \bar{X}_2^{**})^2$$

$$\hat{V}_c(S_1^{**2}) = \frac{1}{n_1}\left(\hat{\mu}_{4(1)}^{**} - \frac{n_1 - 3}{n_1 - 1}S_1^{**4}\right)$$

$$\hat{V}_c(S_2^{**2}) = \frac{1}{n_2}\left(\hat{\mu}_{4(2)}^{**} - \frac{n_2 - 3}{n_2 - 1}S_2^{**4}\right)$$

$$\hat{\mu}_{4(1)}^{**} = \frac{1}{n_1}\sum_{i=1}^{n_1}\left(X_{1i}^{**} - \bar{X}_1^{**}\right)^4$$

$$\hat{\mu}_{4(2)}^{**} = \frac{1}{n_2}\sum_{i=1}^{n_2}\left(X_{2i}^{**} - \bar{X}_2^{**}\right)^4$$

$$S_1^{**4} = \left(S_1^{**2}\right)^2$$

$$S_2^{**4} = \left(S_2^{**2}\right)^2$$

ปฏิเสธสมมติฐานหลักเมื่อ $t_{C^{**}} < t_{\frac{\alpha}{2},\upsilon}$ หรือ $t_{C^{**}} > t_{1-\frac{\alpha}{2},\upsilon}$

โดย υ คือองศาความเป็นอิสระของการแจกแจงที ซึ่งมีค่าเท่ากับ $n_1 + n_2 - 2$

$P - value = \dfrac{\sum_{b=1}^{B} \mathrm{I}_{(t_{1-\frac{\alpha}{2},\upsilon,\infty})}(|t_{c^{**}}|)}{B}$ ปฏิเสธสมมติฐานหลักเมื่อ $P - value < \alpha$

2.3 เอกสารและงานวิจัยที่เกี่ยวข้อง

ปิยวรรณ ถือแก้ว (2552) ศึกษาการเปรียบเทียบวิธีทดสอบความเท่ากันของความแปรปรวนเมื่อประชากรมีการแจกแจงสมมาตรและไม่สมมาตร จำนวน 4 กลุ่ม และ 5 กลุ่ม โดยกลุ่มตัวอย่างมีขนาดเท่ากัน และไม่เท่ากัน 5 วิธี ผลการศึกษาพบว่าวิธีของ Pardo (1997) เป็นวิธีทดสอบที่สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ และมีอำนาจการทดสอบสูงสุดสำหรับกลุ่มตัวอย่างที่มีขนาดเท่ากัน แต่ถ้ากลุ่มตัวอย่างมีขนาดไม่เท่ากัน วิธีทดสอบ O'bren เป็นวิธีทดสอบที่สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ และมีอำนาจการทดสอบสูงสุด

นภดล วันชัยชนะ (2553) ได้ทำการเปรียบเทียบอำนาจการทดสอบของสถิติที่ไม่อิงพารามิเตอร์ในการทดสอบความเท่ากันของความแปรปรวนระหว่างประชากร 2 กลุ่ม เมื่อข้อมูลมาจากการแจกแจงล็อกปรกติและการแจกแจงเอ็กซ์โปเนนเชียล ผลการศึกษาพบว่าสถิติทดสอบ Brunner และ Munsel ให้ค่าอำนาจการทดสอบสูง และสามารถควบคุมความน่าจะเป็นของความคลาดเคลื่อนแบบที่ 1 ได้เกือบทุกกรณีที่ศึกษา และให้ค่าอำนาจการทดสอบสูงสุด

Cojbasic และ Tomovic (2007) ได้นำวิธีการบูทสแทรปมาประยุกต์ใช้กับวิธีการแปลงของ Hall และ Johnson เพื่อประมาณค่าแบบช่วงของความแปรปรวนผลการศึกษาพบว่าวิธีการของ Johnson ให้ค่าความน่าจะเป็นของการคลอบคลุมพารามิเตอร์แคบกว่าทุกวิธีเมื่อขนาดตัวอย่างมีขนาดเล็ก และวิธีการของ Hall ให้ค่าความน่าจะเป็นของการคลอบคลุมพารามิเตอร์แคบกว่าทุกวิธีเมื่อขนาดตัวอย่างมีขนาดใหญ่

Cojbasic และคณะ (2012) ศึกษาวิธีการทดสอบความแปรปรวนจากตัวอย่าง 1 กลุ่ม และความแปรปรวนจากตัวอย่าง 2 กลุ่ม ผลการศึกษาพบว่า วิธีบูทสแทรปทีรวมกับวิธีแปลงของ Hall เหมาะสมสำหรับกรณีที่ขนาดตัวอย่างมีขนาดเล็ก เมื่อประชากรมีการแจกแจงแบบไวบูลล์ และการแจงแจงแบบเอ็กซ์โพเนนเชียล

Junyong และ Dohwan (2012) ศึกษาการทดสอบความเท่ากันของความแปรปรวนในการวิเคราะห์ความแปรปรวน ในตัวอย่างขนาดใหญ่ผลการศึกษาพบว่า วิธีการของ Edword สามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้เหมาะสม และมีอำนาจการทดสอบสูงกว่าวิธีอื่นทุกกรณีที่ศึกษา

Panichkitkosolkol (2013) ศึกษาการประมาณค่าแบบช่วงของความแปรปรวนแบบทางเดียวของการแจกแจงที่มีลักษณะแบบเบ้โดยการปรับวิธีบูทสแทรปทีชั้นเดียว และแบบสองชั้นพบว่าการปรับวิธีบูทสแทรปแบบสองชั้นเหมาะสมสำหรับการประมาณค่าแบบช่วงด้านเดียวทางด้านขวา และยังสามารถใช้วิธีบูทสแทรปแก้ปัญหาการประมาณช่วงความเชื่อมั่นเมื่อข้อมูลไม่มีการแจกแจงปรกติ

Chen และคณะ (2015) ได้ทำการทดสอบความเท่ากันของความแปรปรวนของประชากร 2 กลุ่ม ด้วยสถิติทดสอบเอฟ สถิติทดสอบแจ็คไนฟ์เอ็มพิริคอลไลค์ลิฮูด และสถิติทดสอบเลอวีน (Levene) ภายใต้การแจกแจงปรกติ การแจกแจงเอ็กซ์โพเนนเชียล และการแจกแจงไวบูล ผลการศึกษาพบว่าเมื่อข้อมูลมากจากการแจกแจงปรกติสถิติทดสอบเอฟ และสถิติทดสอบเลอวีนสามารถควบคุมความคลาดเคลื่อนแบบที่ 1 ได้ทุกกรณีที่ศึกษา แต่สถิติทดสอบแจ็คไนฟ์เอ็มพิริคอลไลค์ลิฮูดมีอำนาจการทดสอบสูงกว่าสถิติทดสอบเลอวีน และเมื่อขนาดตัวอย่างใหญ่สถิติทดสอบแจ็คไนฟ์เอ็มพิริคอลไลค์ลิฮูดมีความแกร่ง (Robust) กว่าสถิติทดสอบเลอวีน และสถิติทดสอบเอฟ