บทที่ 2

ระเบียบวิธีที่ใช้ในการวิจัย

2.1 <u>สถิติที่ใช้พารามิเตอร์กับสถิติที่ไม่ใช้พารามิเตอร์</u> (Parametric Statistics and Nonparametric Statistics)

การอ้างอิงทางสถิติส่วนมากหรือเกือบทั้งหมดตั้งอยู่บนข้อสมมติที่ว่า ตัวอย่างที่สุ่มมานั้นมาจากประชากรที่เราทราบว่ามีลักษณะการแจกแจงอย่างใดอย่างหนึ่ง เพราะฉะนั้น ในทางทฤษฎีฟังก์ชันของการแจกแจงจึงขึ้นอยู่กับค่าของพารามิเตอร์ (Parameter) ค่าหนึ่งหรือหลายค่าของประชากรนั้น ๆ จึงเรียกวิธีการทางสถิติเพื่อกะประมาณและทดสอบสมมติฐานเกี่ยวกับค่าพารามิเตอร์ว่า เป็นสถิติที่ใช้พารามิเตอร์ หรือ Parametric Statistics ซึ่งเป็นการกะประมาณและทดสอบค่าพารามิเตอร์ที่ยังไม่ทราบค่า จากค่าที่เรียกว่าค่าสถิติที่คำนวณจากตัวอย่างที่เลือกมาจากประชากรนั้น ๆ โดยวิธีการสุ่ม

การใช้วิธีการทดสอบแบบใช้พารามิเตอร์ในการกะประมาณและทดสอบสมมติฐานเกี่ยวกับค่าของประชากร จะเป็นไปอย่างมีประสิทธิภาพจะต้องขึ้นอยู่กับเงื่อนไขสองประการ ประการแรก ลักษณะการแจกแจงของประชากรจะต้องเป็นแบบปกติ (Normality of Population) และประการที่สองค่าความแปรปรวนจะต้องคงที่ (Stability of Variance) ซึ่งในการทดสอบแบบ t-test และ F-test ในทางสถิติ ได้ตั้งข้อสมมติว่าลักษณะของประชากรนั้นๆ เป็นไปตามเงื่อนไขดังกล่าวทั้งสองประการ

ภายใต้สถานการณ์บางอย่าง หากลักษณะการแจกแจงของประชากรไม่เป็นแบบปกติอย่างเด่นชัด หรือผู้วิจัยไม่สามารถที่จะระบุลงไปอย่างแจ้งชัดว่าลักษณะการแจกแจงของประชากรเป็นแบบใด และไม่สามารถหาเหตุผลมาสนับสนุนการตั้งข้อสมมติว่าลักษณะการแจกแจงของประ-

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

ประกรเป็นไปตามเงื่อนไขทั้งสองประการดังกล่าวมาแล้วข้างต้น ในกรณีเช่นนี้อาจใช้วิธีการทางสถิติที่เรียกว่า วิธีการทดสอบแบบไม่ใช้พารามิเตอร์ หรือ Nonparametric Test หรือเรียกอีกนัยหนึ่งว่า Distribution-Free Test ซึ่งเป็นวิธีการทดสอบซึ่งไม่จำเป็นต้องระบุหรือกำหนดลักษณะการแจกแจงของข้อมูลในประชากรแต่อย่างใด เพียงแต่กำหนดว่าประชากรนั้นจะต้องมีฟังก์ชั่นของการแจกแจงเป็นแบบต่อเนื่อง (Continuous Probability Distribution Function) เท่านั้น แล้วใช้วิธีการทดสอบที่ได้มีผู้คิดค้นขึ้นมาโดยไม่ต้องกำหนดข้อสมมติเกี่ยวกับการแจกแจงของประชากรไว้อย่างเคร่งครัด

วิธีการทดสอบแบบไม่ใช้พารามิเตอร์นี้มีข้อดีหลายประการ ที่สำคัญที่สุดก็คือ นำไปใช้ประโยชน์ได้ง่าย เมื่อเปรียบเทียบกับวิธีการทดสอบแบบใช้พารามิเตอร์แล้ว จะเห็นได้ว่า วิธีการทดสอบแบบไม่ใช้พารามิเตอร์นี้ง่ายต่อการอธิบายและทำความเข้าใจ การคำนวณโดยทั่วไปก็ไม่ยุ่งยากอีกด้วย

2.2 <u>แผนการทดลองแบบสุ่มภายในบล๊อก</u> (Randomized Blocks Design, RBD)

แผนการทดลองแบบสุ่มภายในบล๊อกที่นำมาศึกษาวิจัยในครั้งนี้ จะพิจารณาถึงแผนการทดลองที่เป็นบล๊อกสมบูรณ์ ซึ่งหมายความว่าในแต่ละบล๊อกจะต้องมีครบทุกทรีทเมนท์ แผนการทดลองแบบนี้ ใช้สำหรับกรณีที่ผู้ทดลองสามารถแบ่งหน่วยทดลองออกเป็นกลุ่ม หรือประเภทได้โดยมีจุดประสงค์เพื่อให้หน่วยทดลองที่อยู่ภายในบล๊อกเดียวกัน มีลักษณะเหมือนกันหรือคล้ายคลึงกันมากที่สุด (Homogeneous) และหน่วยทดลองที่อยู่ต่างบล๊อกกันจะมีความแตกต่างกันมากที่สุด กล่าวคือให้หน่วยทดลองภายในบล๊อกเดียวกันมีความผันแปรน้อยกว่าความผันแปรระหว่างบล๊อกก่อนที่จะให้ทรีทเมนท์กับหน่วยทดลอง

แผนการทดลองแบบสุ่มภายในบล๊อก มีแบบหุ่น (Model) และลักษณะของข้อมูลแสดงไว้ในตารางที่ 2.1 ดังนี้

แบบหุ่นสำหรับแผนการทดลองแบบสุ่มภายในบล๊อก

$$x_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} ; \quad i = 1,2,....,t$$
$$j = 1,2,....,r$$

โดยที่ x_{ij} หมายถึงค่าสังเกตจากหน่วยทดลองในบล๊อกที่ j ได้รับทรีทเมนท์ที่ i

μ หมายถึงค่าเฉลี่ยของประชากร

τ_i หมายถึงอิทธิพลของทรีทเมนท์ที่ i

β_j หมายถึงอิทธิพลของบล๊อกที่ j

ϵ_{ij} หมายถึงค่าความคลาดเคลื่อนของการทดลองจากหน่วยทดลองที่ (i,j)

ตารางที่ 2.1 ข้อมูลจากแผนการทดลองแบบสุ่มภายในบล๊อก

ทรีทเมนท์	บล๊อก j = 1,2,.....,r				รวม	ค่าเฉลี่ย
i = 1,2,....,t	1	2	.....	r	$(x_{i.})$	$(\bar{x}_{i.})$
1	x_{11}	x_{12}	.....	x_{1r}	$x_{1.}$	$\bar{x}_{1.}$
2	x_{21}	x_{22}	.....	x_{2r}	$x_{2.}$	$\bar{x}_{2.}$
.	.	.	.....	.	.	.
.	.	.	.....	.	.	.
.	.	.	.....	.	.	.
t	x_{t1}	x_{t2}	.....	x_{tr}	$x_{t.}$	$\bar{x}_{t.}$
รวม $(x_{.j})$	$x_{.1}$	$x_{.2}$	.....	$x_{.r}$	$x_{..}$	
ค่าเฉลี่ย $(\bar{x}_{.j})$	$\bar{x}_{.1}$	$\bar{x}_{.2}$	.....	$\bar{x}_{.r}$		$\bar{x}_{..}$

2.3 การวิเคราะห์ความแปรปรวนสำหรับแผนการทดลองแบบสุ่มภายในบล็อก (The Analysis of Variance for Randomized Blocks Design)

การวิเคราะห์ความแปรปรวน เป็นการวิเคราะห์ความผันแปรทั้งหมดของข้อมูลที่เกิดขึ้นจากการทดลองว่ามีสาเหตุมาจากอะไรบ้าง แล้ววัดความผันแปรของข้อมูลที่เกิดขึ้นจากสาเหตุออกมาในรูปของผลบวกกำลังสอง (Sum of Squares) เพื่อทดสอบต่อไป

การวิเคราะห์ความแปรปรวนสำหรับแผนการทดลองแบบสุ่มภายในบล็อกเพื่อทดสอบความแตกต่างระหว่างอิทธิพลของบล็อกและทรีทเมนท์ แสดงไว้ในตารางที่ 2.2 ดังนี้

ตารางที่ 2.2 การวิเคราะห์ความแปรปรวนของข้อมูลจากตารางที่ 2.1

SOV	df	SS	MS	F
Blocks	(r-1)	$[(\sum_j x^2_{.j})/t] - [\sum_{ij}\sum x_{ij})^2/rt]$	SSB/df = MSB	MSB/MSE
Treatments	(t-1)	$[(\sum_i x^2_{i.})/r] - [\sum_{ij}\sum x_{ij})^2/rt]$	SST/df = MST	MST/MSE
Error	(r-1)(t-1)	SS Total - SSB - SST = SSE	SSE/df = MSE	
Total	(rt-1)	$\sum_{ij}\sum x^2_{ij} - [(\sum_{ij}\sum x_{ij})^2/rt]$		

สมมติฐานหลักที่ต้องการทดสอบ (Null Hypothesis) คือ

H_{01} : ไม่มีความแตกต่างระหว่างอิทธิพลของบล็อก

H_{02} : ไม่มีความแตกต่างระหว่างอิทธิพลของทรีทเมนท์

ในการทดสอบ จะปฏิเสธสมมติฐานหลัก เมื่อค่า F จากการคำนวณมีค่ามากกว่าค่า F จากตารางการแจกแจงแบบ F (F Distribution) ที่ระดับความเป็นอิสระ (Degrees of Freedom) เท่ากับ (r-1),(r-1)(t-1) และ (t-1),(r-1)(t-1) ตามลำดับ

2.4 การทดสอบของฟรีดแมน (Friedman's Test for Two-Way Analysis of Variance)

การทดสอบของฟรีดแมน เป็นการทดสอบแบบไม่ใช้พารามิเตอร์กับข้อมูลที่มีความผันแปรเกิดขึ้น 2 ทาง คือจากบล็อกและทรีทเมนท์

ในการทดสอบว่ามีความแตกต่างระหว่างอิทธิพลของบล็อก หรือทดสอบสมมติฐานหลักว่า

H_{01} : ไม่มีความแตกต่างระหว่างอิทธิพลของบล็อก

ข้อมูลที่นำมาทดสอบจัดไว้ในรูปของตารางสองทาง (Two-Way Table) โดยมีแถวตั้ง (Column) เป็นบล็อก ($j = 1,2, \ldots,r$) และแนวนอน (Row) เป็นทรีทเมนท์ ($i = 1,2,\ldots,t$) วิธีการทดสอบทำดังนี้

1) จัดอันดับ (Rank) ให้แก่ข้อมูลในแถวนอนแต่ละแถว

2) รวมค่าอันดับของแถวตั้งแต่ละแถว ให้เป็น R_j

R_j = ผลรวมของอันดับของข้อมูลในบล็อกที่ j

3) คำนวณค่าสถิติ

$$S_B = \left[\frac{12}{tr(r+1)} \sum_{j=1}^{r} R_j\right] - 3t(r+1)$$

4) จะปฏิเสธสมมติฐานหลัก เมื่อค่า S_B จากการคำนวณมีค่ามากกว่าหรือเท่ากับค่าสถิติ $s(\alpha,r,t)$ จากตารางที่แสดงไว้ในภาคผนวก และสำหรับกรณีที่ t มีค่ามาก ค่า S_B จะมีลักษณะการแจกแจงใกล้เคียงกับการแจกแจงแบบไคสแควร์ ที่ระดับความเป็นอิสระเท่ากับ r-1

ในทำนองเดียวกัน เมื่อต้องการทดสอบว่ามีความแตกต่างระหว่างอิทธิพลของทรีทเมนท์หรือไม่ ตารางสองทางดังกล่าว จะมีแถวตั้งเป็นทรีทเมนท์ และแถวนอนเป็นบล็อก แล้วดำเนินการทดสอบเช่นเดียวกับการทดสอบข้างต้น โดยคำนวณค่าสถิติดังนี้

$$S_T = \left[\frac{12}{rt(t+1)} \sum_{i=1}^{t} R_i^2 \right] - 3r(t+1)$$

สมมติฐานหลักที่ต้องการทดสอบดังนี้คือ

H_{02} : ไม่มีความแตกต่างระหว่างอิทธิพลของทรีทเมนท์

ในการทดสอบจะปฏิเสธสมมติฐานหลัก เมื่อค่า S_T จากการคำนวณมีค่ามากกว่าหรือเท่ากับค่าสถิติ $s(\alpha, t, r)$ และสำหรับกรณีที่ r มีค่ามาก ค่า S_T จะมีลักษณะการแจกแจงใกล้เคียงกับการแจกแจงแบบไคสแควร์ ที่ระดับความเป็นอิสระเท่ากับ t-1

2.5 <u>ข้อสมมติในการวิเคราะห์ความแปรปรวน</u> (Assumptions Underlying the Analysis of Variance)

การวิเคราะห์ความแปรปรวนเพื่อทดสอบอิทธิพลต่าง ๆ ที่มีอยู่ในการทดลอง มีข้อสมมติ 2 ประการ คือ

1) อิทธิพลของทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ เป็นแบบบวก (Additive)

2) ความคลาดเคลื่อนของการทดลอง เกิดขึ้นโดยสุ่มเป็นอิสระต่อกัน และมีลักษณะการแจกแจงแบบปกติ ที่มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวนเป็น σ^2

โดยทางปฏิบัตินั้น ผู้วิจัยไม่อาจแน่ใจว่าข้อมูลที่ได้จากการทดลองจะมีคุณสมบัติคล้อยตามข้อสมมติที่กำหนดไว้ทุกข้อไป และมีบ่อยครั้งที่ปรากฏว่าลักษณะของข้อมูลไม่ตรงกับข้อสมมติข้อใดข้อหนึ่งหรือทั้งสองข้อ ในกรณีนี้จำเป็นจะต้องแก้ไขข้อมูลก่อนทำการทดสอบด้วยการแปลงข้อมูล เพราะหากข้อมูลคลาดเคลื่อนจากข้อสมมติข้อใดข้อหนึ่งหรือทั้งสองข้อดังกล่าว อาจจะทำให้ผลการวิเคราะห์

ผิดพลาดไป อันจะทำให้ระดับนัยสำคัญที่ใช้ในการทดสอบ (Significance Level) และความไว (Sensitivity) ของค่า F เปลี่ยนแปลงไป สำหรับเหตุที่ข้อมูลไม่เป็นไปตามข้อสมมตินั้น สืบเนื่องมาจากสาเหตุดังต่อไปนี้

1) อิทธิพลของทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ มีลักษณะไม่เป็นแบบบวก (Non-additive) เช่นหากพบผลต่าง ๆ เกิดขึ้นแบบผลคูณแล้ว แสดงว่าความแตกต่างในระหว่างบล็อก (จากบล็อกหนึ่งไปยังอีกบล็อกหนึ่ง) จะแตกต่างกันระหว่างทรีทเมนท์ และมีความแตกต่างระหว่างทรีทเมนท์ไม่เหมือนกันในระหว่างบล็อก เพราะหากอิทธิพลทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ มีลักษณะเป็นแบบบวก ความแตกต่างระหว่างบล็อกจะเหมือนกันระหว่างทรีทเมนท์ และความแตกต่างระหว่างทรีทเมนท์ก็เหมือนกันในระหว่างบล็อก หากในกรณีที่พบอิทธิพลต่าง ๆ เกิดขึ้นแบบผลคูณนี้ สามารถแก้ไขได้ด้วยการหาค่าลอคการิทึมของข้อมูลเดิม จะเห็นว่าเมื่อเปลี่ยนเป็นค่าลอคการิทึมแล้วแบบหุ่นผลคูณก็กลายเป็นแบบบวก ข้อมูลใหม่นี้เรียกว่า ข้อมูลแปลง (Transformed Data) สำหรับข้อมูลลักษณะอื่น ๆ ที่ไม่เป็นแบบบวกอาจจะต้องใช้วิธีการแปลงข้อมูลแบบอื่น การแปลงข้อมูลนี้จะต้องทำให้ข้อมูลที่แปลงแล้วมีคุณสมบัติครบตามข้อสมมติจึงจะใช้ได้

การวิเคราะห์ข้อมูลที่มีลักษณะไม่เป็นแบบบวกโดยมิได้แก้ไขข้อมูลเสียก่อน จะมีผลทำให้ความแปรปรวนของแต่ละทรีทเมนท์ไม่เท่ากัน (Heterogeneity of Variance) ซึ่งหมายความว่า ส่วนประกอบของความแปรปรวนของความคลาดเคลื่อน (Components of Error Variance) ซึ่งประกอบด้วยความแปรปรวนจากค่าสังเกตต่าง ๆ จะไม่ให้ค่าประมาณความแปรปรวนที่ถูกต้อง และความแปรปรวนที่ได้นี้ เมื่อนำไปประมาณค่าช่วงเชื่อมั่น (Confidence interval) และทดสอบสมมติฐานของค่าพารามิเตอร์จะทำให้ผิดพลาดไป และอาจทำให้ระดับนัยสำคัญในการทดสอบผิดพลาดด้วย ส่วนค่า F นั้นจะมีผลกระทบเพียงเล็กน้อย

2) ถ้าผู้ทดลองจงใจเรียงทรีทเมนท์โดยวิธีใดวิธีหนึ่ง จะทำให้ผลลัพธ์ระหว่างทรีทเมนท์ที่ใกล้ ๆ กันคล้ายคลึงกันกว่าทรีทเมนท์ที่ห่างกัน ในกรณีนี้ความคลาดเคลื่อนที่เกิดขึ้นภายในแต่ละทรีทเมนท์จะไม่เป็นอิสระต่อกัน แต่จะมีความสัมพันธ์กัน ซึ่งจะทำให้การทดสอบนัยสำคัญผิดพลาดถ้าหาก

ไม่ได้แก้ไขเสียก่อน ในทางปฏิบัติจะจัดทรีทเมนท์ให้กับหน่วยทดลองโดยสุ่ม เพื่อทำให้ค่าความคลาดเคลื่อนเหล่านี้เป็นอิสระต่อกัน

3) ความคลาดเคลื่อนของการทดลองไม่มีลักษณะการแจกแจงแบบปกติ เช่น เอียงข้างใดข้างหนึ่ง หรือที่เรียกว่าเบ้ (Skew) ในกรณีนี้ ความแปรปรวนของความคลาดเคลื่อน (Error Variance) มักจะปรากฏเป็นฟังก์ชั่นของค่าเฉลี่ยของทรีทเมนท์ ซึ่งจะทำให้ความแปรปรวนของความคลาดเคลื่อนแตกต่างกัน ดังนี้ หากข้อมูลมีค่าความคลาดเคลื่อนที่มีลักษณะการแจกแจงแตกต่างไปจากแบบปกติแล้ว ก็จำเป็นจะต้องแก้ไขด้วยการแปลงข้อมูลเพื่อให้มีลักษณะการแจกแจงแบบปกติเสียก่อน แล้วทำการวิเคราะห์ความแปรปรวนของข้อมูลที่แปลงแล้ว และสรุปผลต่อไปได้

4) การทดลองบางอย่างอาจมีทรีทเมนท์บางทรีทเมนท์ที่แตกต่างไปจากทรีทเมนท์อื่นอย่างมาก ทำให้ความแปรปรวนของความคลาดเคลื่อนที่ได้จากแต่ละทรีทเมนท์มีความแปรปรวนไม่เท่ากัน ซึ่งหากค่าเฉลี่ยของทรีทเมนท์นั้นมีค่ามากกว่าทรีทเมนท์อื่นอย่างมาก ทั้งยังมีความผันแปรสูงมาก ทรีทเมนท์นี้อาจถูกตัดออกจากการวิเคราะห์เพื่อทำให้ข้อมูลชุดนี้เป็นไปตามข้อสมมติที่กำหนด

2.6 <u>การสร้างข้อมูลจากแผนการทดลองแบบสุ่มภายในบล็อกและไม่เป็นไปตามข้อสมมติสำหรับการวิเคราะห์ความแปรปรวน</u>

ในการวิจัยครั้งนี้ ผู้วิจัยสร้างข้อมูลจากแผนการทดลองแบบสุ่มภายในบล็อกโดยมีจุดประสงค์ให้ข้อมูลมีคุณสมบัติไม่เป็นไปตามข้อสมมติของการวิเคราะห์ความแปรปรวน โดยแบ่งข้อมูลออกเป็น 2 แบบ คือ

แบบที่ 1 ข้อมูลแบบหุ่นผลคูณ กำหนดให้ข้อมูลแต่ละชุดมีอิทธิพลของทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ มีลักษณะเป็นแบบผลคูณ และค่าสังเกตการณ์ของค่าความคลาดเคลื่อนของการทดลองเป็นค่าที่เกิดขึ้นโดยสุ่มเป็นอิสระต่อกัน และมีลักษณะการแจกแจงปกติที่มีค่าเฉลี่ยเป็นศูนย์ และมีความ

แปรปรวนเท่ากันเป็น σ^2

การสร้างข้อมูลจากค่าลอคการิทึมของความคลาดเคลื่อนนั้น เนื่องจากข้อมูลนี้เป็นแบบหุ่นผลคูณที่มีแบบหุ่นเป็น

$$X_{ij} = \mu\tau_i\beta_j\epsilon_{ij} \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (2.6.1)$$

จะเห็นว่าวิธีการแปลงข้อมูลให้เป็นค่าลอคการิทึม จะช่วยทำให้ข้อมูลชุดนี้มีอิทธิพลของทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ เป็นแบบบวกที่มีแบบหุ่นเป็น

$$\log X_{ij} = \log\mu + \log\tau_i + \log\beta_j + \log\epsilon_{ij} \quad \ldots\ldots \quad (2.6.2)$$

ภายใต้แบบหุ่นใน (2.6.2) ค่า $\log \epsilon_{ij}$ จะต้องเป็นไปตามข้อสมมติที่ว่า มีลักษณะการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์ และความแปรปรวนเป็น σ^2 ดังนั้นผู้วิจัยจึงสร้างค่าสังเกต X_{ij} โดยให้ค่าลอคการิทึมของค่าความคลาดเคลื่อนเป็นค่าสุ่ม และมีลักษณะการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวนเป็น σ^2

แบบที่ 2 ข้อมูลแบบหุ่นผลบวก กำหนดให้ข้อมูลแต่ละชุด มีอิทธิพลของทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ มีลักษณะเป็นแบบบวก และค่าความคลาดเคลื่อนของการทดลองเป็นค่าที่เกิดขึ้นโดยสุ่มเป็นอิสระต่อกัน และมีการแจกแจงแบบปกติที่มีค่าเฉลี่ยเป็นศูนย์ และมีค่าความแปรปรวนของแต่ละทรีทเมนท์ไม่เท่ากัน

ข้อมูลแบบแรกนั้น สร้างขึ้นจำนวน 112 ชุด โดยมีขนาดทรีทเมนท์และบล็อกต่าง ๆ กันตั้งแต่ 3 ถึง 9 ทรีทเมนท์ และตั้งแต่ 3 ถึง 9 บล็อก ส่วนข้อมูลแบบที่สอง สร้างขึ้นจำนวน 100 ชุด โดยมีขนาดทรีทเมนท์และบล็อกต่าง ๆ กัน ตั้งแต่ 5 ถึง 9 ทรีทเมนท์ และตั้งแต่ 3 ถึง 9 บล็อก

ตัวอย่างการสร้างข้อมูลแต่ละชุดทั้งสองแบบ โดยกำหนดให้มีขนาดทรีทเมนท์ = 5 และขนาดบล็อก = 3

1) การสร้างข้อมูลแบบผลคูณ

1.1 สุ่มค่าความคลาดเคลื่อนของการทดลอง (ϵ_{ij} ; i = 1,2,....,5 และ j = 1,2,3) ที่กำหนดให้ $\log(\epsilon_{ij})$ มีลักษณะการแจกแจงแบบปกติ มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวนเท่ากันเป็น σ^2 จากโปรแกรมมาตรฐาน "GAUSS"

1.2 สุ่มค่า μ, τ_i ; i = 1,2,...,5 และ β_j ; j = 1,2,3 จากตารางเลขสุ่ม แล้วสร้างค่าสังเกตของตัวอย่างข้อมูลจากแบบหุ่นผลคูณ

$$x_{ij} = \mu\tau_i\beta_j\epsilon_{ij} \quad ; \; i = 1,2,...,5 \quad j = 1,2,3$$

005038

ได้ค่าสังเกตจำนวน 15 ค่า ที่มีขนาดทรีทเมนท์ = 5 และขนาดบล็อก = 3 เป็นข้อมูล 1 ชุด

2) การสร้างข้อมูลแบบผลบวก

2.1 สุ่มค่าความคลาดเคลื่อนจากการทดลอง (ϵ_{ij} ; i = 1,2,...,5 และ j = 1,2,3) ที่มีการแจกแจงแบบปกติ มีค่าเฉลี่ยเป็นศูนย์ และมีความแปรปรวนต่างกัน จากโปรแกรมมาตรฐาน "GAUSS" เช่นกัน

2.2 สุ่มค่า μ, τ_i ; i = 1,2,...,5 β_j ; j = 1,2,3 จากตารางเลขสุ่ม แล้วสร้างค่าสังเกตตัวอย่างข้อมูลจากแบบหุ่นผลบวก

$$x_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \quad ; \; i = 1,2,...,5 \quad j = 1,2,3$$

ได้ค่าสังเกตจำนวน 15 ค่า ที่มีขนาดทรีทเมนท์ =5 และขนาดบล็อก =3 เป็นข้อมูล 1 ชุด

2.7 <u>การทดสอบแบบบวกของทูกี้</u> (Tukey's Test for Additivity)

ในการทดสอบว่าข้อมูลแต่ละชุดมีลักษณะเป็นไปตามข้อสมมติที่ว่าอิทธิพลของทรีทเมนท์

I 15482492

และสิ่งแวดล้อมอื่น ๆ มีลักษณะเป็นแบบบวกหรือไม่นั้น ทดสอบโดยวิธีการทดสอบแบบบวกของทูกี้ โดยแบ่งค่าผลบวกกำลังสองของค่าความคลาดเคลื่อนของการทดลอง (Error Sum of Squares ,SSE) ที่ได้จากวิเคราะห์ความแปรปรวนเป็น 2 ส่วน คือ ค่าผลบวกกำลังสองของค่าความไม่เป็นแบบบวก (Sum of Squares for Non-additivity, SSN) และค่าผลบวกกำลังสองของค่ารีเมนเดอร์ (Remainder Sum of Squares, SS (Remainder)) ซึ่งมีวิธีการคำนวณดังนี้

1) คำนวณค่าผลบวกกำลังสองของค่าความไม่เป็นแบบบวก (SSN) โดยมีขั้นตอนการคำนวณเป็น

1.1 คำนวณ $d_i = \bar{X}_{i.} - \bar{X}_{..}$; $i = 1,2,....,t$

และ $d_j = \bar{X}_{.j} - \bar{X}_{..}$; $j = 1,2,....,r$

1.2 คำนวณ $W_i = \sum_j X_{ij} d_j$

1.3 คำนวณ $N = \sum_i W_i d_i = \sum_{ij} \sum X_{ij} d_i d_j$

1.4 คำนวณ $SSN = \dfrac{N^2}{\sum_i d_i^2 \sum_j d_j^2}$

2) คำนวณค่าผลบวกกำลังสองของค่ารีเมนเดอร์ (SS (Remainder)) โดยที่

SS (Remainder) = SSE - SSN

ค่า SSN และ SS (Remainder) มีระดับความเป็นอิสระเท่ากับ 1 และ $(t-1)(r-1)-1$ ตามลำดับ การทดสอบแบบบวกของทูกี้ แสดงไว้ในตารางที่ 2.3 ดังนี้

ตารางที่ 2.3 การทดสอบแบบบวกของทูกี้ของข้อมูลจากแผนการทดลองแบบสุ่มภายในบล๊อก

SOV	df	SS	MS	F
Error	$(t-1)(r-1)$	SSE		
Nonadditivity	1	SSN	$\frac{SSN}{df}=MSN$	$\frac{MSN}{MS(Remainder)}$
Remainder	$(t-1)(r-1)-1$	SS(Remainder)	$\frac{SS(Remainder)}{df}$ $= MS(Remainder)$	

สมมติฐานหลักที่ต้องการทดสอบ คือ

H_0 : อิทธิพลของทรีทเมนท์และสิ่งแวดล้อมอื่น ๆ เป็นแบบบวก

ผลการทดสอบ จะปฏิเสธสมมติฐานหลัก เมื่อค่า F จากการคำนวณมีค่ามากกว่าค่า F จากตารางการแจกแจง F ที่ระดับความเป็นอิสระเท่ากับ 1 และ $(t-1)(r-1)-1$ ตามลำดับ

2.8 การทดสอบ Homogeneity of Variance

ในกรณีที่ต้องการทดสอบว่า ความแปรปรวนมากกว่า 2 ค่า มีความแตกต่างกันหรือไม่นั้น ทดสอบโดยวิธีการทดสอบของบาร์ตเล็ตต์ แต่สำหรับกรณีที่มีความแปรปรวนเพียงสองค่า อาจทำการทดสอบได้โดยวิธีการทดสอบแบบ F

ในที่นี้ ต้องการทดสอบว่า ทรีทเมนท์ต่าง ๆ ของข้อมูลแต่ละชุดนั้นมีค่าความแปรปรวนแตกต่างกันหรือไม่ ซึ่งทำได้โดยการคำนวณค่าความแปรปรวนของแต่ละทรีทเมนท์จากข้อมูลที่ต้องการทดสอบเป็น s_i^2 จำนวน t ค่า ซึ่งแต่ละค่ามีระดับความเป็นอิสระไม่เท่ากันคือ $df = r_i$; $i = 1,2,....,t$ (r_i เป็นจำนวนซ้ำในทรีทเมนท์ที่ i) คำนวณค่าสถิติแบบ χ^2 โดยใช้สูตรดังนี้

$$\chi_T^2 = \frac{2.3026\ [(\Sigma r_i)\{\log(\Sigma r_i s_i^2/_- r_i)\} - \Sigma r_i \log s_i^2]}{1 + [\{1/3(t-1)\}\{\Sigma(1/r_i) - (1/\Sigma r_i)\}]}$$

ค่า χ_T^2 ที่คำนวณได้นี้เป็นค่าประมาณของการแจกแจงไคสแควร์ ที่ระดับความเป็นอิสระเท่ากับ t-1

สำหรับกรณีที่ทรีทเมนท์แต่ละค่ามีระดับความเป็นอิสระเท่ากัน คือ df = r ค่า χ_T^2 จากสูตรข้างต้นเป็นดังนี้

$$\chi_T^2 = \frac{2.3026\ (r)\,[t\log(\Sigma s_i^2/t) - \Sigma \log s_i^2]}{1 + [(t+1)/3tr]}$$

สมมติฐานหลักที่ต้องการทดสอบคือ

H_O : ค่าความแปรปรวนของแต่ละทรีทเมนท์ไม่แตกต่างกัน

จากผลการทดสอบ จะปฏิเสธสมมติฐานหลัก เมื่อค่า χ_T^2 จากการคำนวณมีค่ามากกว่าค่า χ^2 จากตารางการแจกแจงไคสแควร์ที่ระดับความเป็นอิสระเท่ากับ (t-1)

2.9 การแปลงข้อมูล (Transformation of Data)

ดังได้กล่าวมาแล้วว่า หากข้อมูลที่ได้จากการทดลองมีคุณสมบัติไม่เป็นไปตามข้อสมมติ ก็จะมีผลทำให้การทดสอบผิดพลาด ดังนั้น ก่อนที่จะทำการวิเคราะห์ความแปรปรวน จะต้องมีการแก้ไขด้วยการแปลงข้อมูลเพื่อให้เป็นไปตามข้อสมมติ แล้วทำการวิเคราะห์และสรุปผลจากข้อมูลที่แปลงแล้วนี้

การแปลงข้อมูลมีหลายวิธีให้เลือกใช้แล้วแต่ลักษณะของข้อมูล และเป็นที่ยอมรับโดยทั่วไปว่า การแปลงเป็นค่าลอคการิทึม (The Logarithmic Transformation) เป็นวิธีการแปลงข้อมูลที่สามารถขจัดปัญหาข้อมูลไม่เป็นไปตามข้อสมมติได้ดีกว่าวิธีอื่น ๆ

หอสมุดกลาง สถาบันวิทยบริการ จุฬาลงกรณ์มหาวิทยาลัย

วิธีการแปลงข้อมูลที่ใช้กันทั่วไปมีดังนี้

กำหนดให้ X เป็นค่าสังเกตที่ไม่เป็นไปตามข้อสมมติ (Original Data)

Y เป็นค่าสังเกตที่แปลงแล้ว (Transformed Data)

1) การแปลงเป็นค่าลอคการิทึม (The Logarithmic Transformation) มี 2 แบบคือ

1.1 $Y = \log(X)$

1.2 $Y = \log(X+1)$

การแปลงเป็นค่าลอคการิทึมแบบที่ 2 ได้ผลใกล้เคียงกับแบบที่ 1 แต่ในกรณีที่ค่าสังเกตมีค่าเป็นศูนย์ (0) การแปลงเป็นค่าลอคการิทึมแบบที่ 2 จะให้ผลดีกว่า

การแปลงเป็นค่าลอคการิทึมเหมาะสำหรับข้อมูลที่เป็นเลขจำนวนเต็มบวกและมีช่วงกว้าง ใช้ในกรณีที่ความแปรปรวนมีค่าเป็นสัดส่วนกับกำลังสองของค่าเฉลี่ยของทรีทเมนท์ หรือส่วนเบี่ยงเบนมาตรฐานเป็นสัดส่วนกับค่าเฉลี่ย การแปลงเป็นค่าลอคการิทึมจะมีผลทำให้ข้อมูลในแต่ละทรีทเมนท์มีความแปรปรวนเท่ากัน นอกจากนั้นการแปลงโดยวิธีนี้จะทำให้อิทธิพลต่าง ๆ ที่มีลักษณะเป็นแบบผลคูณในข้อมูลเดิมเปลี่ยนเป็นแบบบวกในมาตราลอคการิทึม

2) การแปลงเป็นค่าอาร์คไซน์ (Arcsine Transformation) เป็นวิธีการแปลงข้อมูลที่อยู่ในรูปของร้อยละหรืออัตราส่วนซึ่งมีลักษณะการแจกแจงแบบไบโนเมียล (Binomial Distribution) นั่นคือ ค่าความแปรปรวนเป็นฟังก์ชันของค่าเฉลี่ยนั่นเอง โดยการแปลงค่าสังเกตเป็นมุมวัดเป็นองศา การแปลงโดยวิธีนี้จะเปลี่ยนการแจกแจงแบบไบโนเมียลของค่าความคลาดเคลื่อนให้มีการแจกแจงเป็นแบบปกติ และมีความแปรปรวนคงที่โดยประมาณ คือ หากวัดมุมเป็นองศา ความแปรปรวนของค่าแปลงเป็นมุมจะเท่ากับ $821/n$ หากวัดมุมเป็นเรเดียน (Radian) ความแปรปรวนจะเท่ากับ $0.25/n$ โดยประมาณ ซึ่งค่า n นั้นเป็นตัวหารร่วมของค่าในข้อมูลเดิมที่คิดเป็นอัตราส่วน จะเห็นได้ว่า ข้อมูลที่คิดเป็นร้อยละก็เพื่อจะได้มีตัวหารร่วมเท่ากัน คือ 100 หรืออีกนัยหนึ่ง เพื่อให้จำนวนค่าสังเกตเท่ากันโดยคิดจากร้อย อย่างไรก็ตามในกรณีที่

ตัวหารร่วมมีค่าเท่ากันเพียงโดยประมาณการแปลงแบบนี้ก็ยังใช้ได้

3) การแปลงเป็นค่ารากที่สอง (Square root Transformation) มี 3 แบบคือ

$$3.1 \quad Y = \sqrt{X}$$

$$3.2 \quad Y = \sqrt{X+1}$$

$$3.3 \quad Y = \sqrt{X}+\sqrt{X+1}$$

การแปลงเป็นค่ารากที่สองใช้ได้กับข้อมูลที่เป็นร้อยละและมีค่าอยู่ในช่วงแคบ ๆ เช่น ร้อยละ 0-20 หรือร้อยละ 80-100 ดังนั้นจึงอาจแปลงค่าเดิมเป็นร้อยละเสียก่อน และหากค่าเดิมอยู่ในช่วงร้อยละ 80-100 จะต้องนำมาหักลบออกจาก 100 ก่อนทำการแปลง และในกรณีที่ข้อมูลมีค่าสังเกตบางค่าต่ำมาก โดยเฉพาะบางค่าเท่ากับศูนย์ (0) การแปลงเป็นค่ารากที่สองแบบที่ 2 และ 3 จะได้ผลที่ดีกว่าแบบที่ 1

การแปลงโดยวิธีนี้ ใช้กับข้อมูลที่มีค่าสังเกตที่มีค่าต่ำมาก ซึ่งมีการแจกแจงแบบปัวซอง (Poisson Distribution) เมื่อแปลงข้อมูลแล้วจะช่วยให้ค่าเฉลี่ยและความแปรปรวนเป็นอิสระจากกัน เนื่องจากการแจกแจงแบบปัวซองมีค่าเฉลี่ยและความแปรปรวนเท่ากัน และความแปรปรวนของข้อมูลแปลงเป็นค่ารากที่สองแล้ว จะมีค่าใกล้เคียงกับ 0.25 นอกจากนี้ การแปลงโดยวิธีนี้ใช้ได้กับกรณีที่ค่าสังเกตเป็นค่าที่ได้จากการนับ และความแปรปรวนของค่าสังเกตเป็นสัดส่วนกับค่าเฉลี่ยของค่าสังเกต คือ $\sigma^2 = k\bar{x}$ เช่น สำหรับการแจกแจงแบบปัวซอง ค่า k จะเท่ากับ 1 การแปลงค่าเฉลี่ยของทรีทเมนท์กลับเป็นข้อมูลเดิม กระทำได้โดยยกกำลังสองของค่าเฉลี่ยของข้อมูลแปลง แต่อย่างไรก็ตาม ค่าเฉลี่ยแปลงกลับจะมีค่าต่ำกว่าค่าเฉลี่ยเดิมเล็กน้อย เพราะค่าเฉลี่ยของข้อมูลแปลงเป็นค่ารากที่สองจะมีค่าน้อยกว่ารากที่สองของค่าเฉลี่ยข้อมูลเดิม การปรับแบบคร่าว ๆ อาจกระทำได้โดยการบวกค่าเฉลี่ยความคลาดเคลื่อนกำลังสอง (Mean Squares Error) จากการวิเคราะห์ข้อมูลแปลงให้กับค่าเฉลี่ยแปลงกลับแล้วของแต่ละทรีทเมนท์

การแปลงข้อมูลเป็นค่ารากที่สอง จะมีผลต่อลักษณะของการแจกแจงของค่าความคลาดเคลื่อนของการทดลอง ตลอดจนลักษณะการรวมอิทธิพลต่าง ๆ ที่เป็นแบบบวก สมมติว่า จาก

ข้อมูลเดิมอิทธิพลของทรีทเมนท์และบล็อกกระทำร่วมกันแบบบวกอยู่แล้ว เมื่อนำมาแปลงแล้วอิทธิพลเหล่านี้จะยังคงลักษณะบวกอยู่เช่นเดิม

2.10 การทดสอบสัดส่วนของผลการทดสอบที่ตรงกัน

ภายหลังจากการแปลงข้อมูลที่ไม่เป็นไปตามข้อสมมติแล้ว นำข้อมูลหลังจากการแปลงมาทดสอบความแตกต่างระหว่างอิทธิพลบล็อกและทรีทเมนท์ด้วยวิธีการวิเคราะห์ความแปรปรวน แล้วนำผลการทดสอบมาเปรียบเทียบกับผลการทดสอบข้อมูลที่ยังไม่ได้แปลง เพื่อทดสอบความแตกต่างดังกล่าวด้วยวิธีการทดสอบของฟรีดแมนว่าผลที่ได้จากการทดลองโดยวิธีทั้งสองนี้ ให้ผลการทดสอบที่ตรงกันด้วยสัดส่วนเท่าใด

การทดสอบสัดส่วนของผลการทดสอบที่ตรงกันนี้ ใช้ทดสอบโดย Z-test ดังนี้

$$Z = \frac{P - P_O}{\sqrt{\frac{P_O(1-P_O)}{n}}}$$

เมื่อ P เป็นสัดส่วนของผลการทดสอบที่ตรงกัน เมื่อใช้วิธีการทดสอบทั้งสองวิธี

P_O เป็นสัดส่วนของผลการสอบที่คาดว่าตรงกัน เมื่อใช้วิธีการทดสอบทั้งสองวิธี

และ n เป็นจำนวนข้อมูลที่ใช้ในการทดสอบ

สมมติฐานที่ต้องการทดสอบคือ

$$H_O : P = P_O$$

$$H_1 : P < P_O$$

ผลการทดสอบ จะปฏิเสธสมมติฐานหลัก เมื่อค่า Z ที่ได้จากการคำนวณมีค่าน้อยกว่า -1.28, -1.65 และ -2.33 ณ ระดับนัยสำคัญ 0.10, 0.05 และ 0.01 ตามลำดับ