แผนการทดลองพื้นฐาน:
แผนการทดลองแบบสุ่มตลอด

แผนการทดลองแบบสุ่มตลอด

แผนการทดลองแบบสุ่มตลอด หรือที่เรียกว่า Completely Randomized Design (CRD) เป็นหนึ่งในสามของแผนการทดลองขั้นพื้นฐาน เป็นแผนการทดลองที่มีลักษณะง่ายๆ และสะดวก จะใช้ได้ในกรณีที่หน่วยทดลองมีความสม่ำเสมอ คล้ายคลึงกันหรือเหมือนกันมากที่สุด ดังนั้นแผนการทดลองแบบนี้จึงเป็นการวิเคราะห์ความผันแปรทั้งหมดในการทดลองว่ามีสาเหตุเนื่องมาจากทรีทเมนต์แต่เพียงอย่างเดียว ในตำราบางเล่มจึงเรียกแผนการทดลองแบบสุ่มตลอดนี้ว่า One Way Classification

ในแผนการทดลองแบบนี้ เราจะใช้วิธีการจัดทรีทเมนต์ให้กับหน่วยทดลองที่มีความสม่ำเสมอกัน โดยวิธีการสุ่มดังได้กล่าวแล้วในบทที่ 3 ซึ่งทุกหน่วยของวัตถุทดลอง จะมีโอกาสเท่ากันที่จะได้รับทรีท-เมนต์ใดทรีทเมนต์หนึ่งก็ได้

แบบหุ่นของแผนการทดลองแบบสุ่มตลอด คือ

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

i คือ ทรีทเมนต์ที่ i ; i = 1, 2, ..., t

j คือ ซ้ำที่ j ในแต่ละทรีทเมนต์; j = 1, 2, ..., r

n คือ จำนวนค่าสังเกต (observation) ทั้งหมด = tr

T_1	T_2	•	•	•	•	•	T_t
Y_{11}	Y_{21}	•	•	•	•	•	Y_{t1}
Y_{12}	Y_{22}	•	•	•	•	•	Y_{t2}
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
Y_{1r_1}	Y_{2r_2}	•	•	•	•	•	Y_{tr_t}

จากแบบหุ่น เราก็ทราบได้ว่า ค่าสังเกต (Y_{ij}) แต่ละค่าเป็นผลมาจาก

1) ค่าเฉลี่ยทั่วๆ ไป (overall mean, μ) เป็นค่าที่ต้องมีอยู่แล้ว เป็นค่าคงที่ร่วมกันค่าหนึ่งก่อนการทดลอง

2) อิทธิพลของทรีทเมนต์ที่ได้รับ (treatment effect, α_i)

3) ความคลาดเคลื่อนของการทดลอง (experimental error, ε_{ij}) นั่นคือความแตกต่างของหน่วยทดลองที่ได้รับทรีทเมนต์เดียวกัน

$$Y_{11} = \mu + \alpha_i$$

$$Y_{21} = \mu + \alpha_2$$

จะเห็นว่า Y_{11} กับ Y_{21} แตกต่างกันเพราะอิทธิพลของทรีทเมนต์ (α_1 และ α_2)

ขณะเดียวกันในทรีทเมนต์เดียวกัน แต่ละซ้ำหรือแต่ละหน่วยทดลองก็จะแตกต่างกันด้วยเพราะตัวของวัตถุทดลองเอง

$$Y_{11} = \mu + \alpha_I + \varepsilon_{11}$$

$$Y_{12} = \mu + \alpha_I + \varepsilon_{12}$$

$$Y_{21} = \mu + \alpha_2 + \varepsilon_{21}$$

นั่นก็คือ Y_{11} กับ Y_{12} จะแตกต่างกัน เพราะตัววัตถุทดลองเอง (ε_{11} กับ ε_{12}) แต่ Y_{11} กับ Y_{21} หรือ Y_{12} กับ Y_{21} นั้นแตกต่างกันเพราะอิทธิพลของทรีทเมนต์ที่ได้รับต่างกัน (α_I กับ α_2) และความแตกต่างของวัตถุทดลองเอง (ε_{11} กับ ε_{21} หรือ ε_{12} กับ ε_{21})

จากแบบหุ่นของการทดลองแบบสุ่มตลอด เราก็จะได้

$$\varepsilon_{ij} = Y_{ij} - \mu - \alpha_i$$

ซึ่งค่าความคลาดเคลื่อน (ε_{ij}) หลายๆ ค่า เมื่อนำมาหาค่าเฉลี่ยก็จะได้ค่าเฉลี่ยเท่ากับ ศูนย์ และมีการกระจายเป็นแบบปกติ (normal distribution) และค่าความคลาดเคลื่อนแต่ละค่าจะเป็นอิสระต่อกัน (independent distribution) โดยค่าวาเรียนซ์ของความคลาดเคลื่อนเท่ากับ σ^2 นั่นก็คือทุกทรีทเมนต์จะต้องมีวาเรียนซ์เท่ากัน

$$\sigma_1^2 = \sigma_2^2 = \ldots\ldots\ldots\ldots\ldots\ldots = \sigma_t^2 = \sigma^2$$

การวิเคราะห์ข้อมูลจะทำได้ต่อเมื่อข้อมูลเป็นไปตามข้อกำหนด (assumption) ถ้าข้อมูลไม่เป็นไปตามข้อกำหนด คือ มีการกระจายไม่เป็นปกติ วาเรียนซ์ของแต่ละทรีทเมนต์ จะไม่เท่ากัน และไม่เท่ากับ σ^2 จะมีผลทำให้การทดสอบมีนัยสำคัญบ่อยครั้งเกินไป หรือไม่มีนัยสำคัญน้อยครั้งเกินไป ดังนั้นควรที่จะมีการทดสอบข้อมูลเสียก่อนว่าเป็นไปตามข้อกำหนดหรือไม่ การทดสอบก็จะเป็นการทดสอบว่า $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$ หรือไม่ โดยการใช้ค่า ไคร์สแควร์ (χ^2) เป็นค่าทดสอบ วิธีการที่ใช้ทดสอบนี้เรียกว่า Bartlett's test of homogeneity of variance ซึ่งจะกล่าวถึงรายละเอียดต่อไป เมื่อข้อมูลเป็นไปตามข้อ กำหนดก็ทำการวิเคราะห์ข้อมูลต่อไปได้

การวิเคราะห์วาเรียนซ์ของการทดลองแบบสุ่มตลอด

จากแบบหุ่นของแผนการทดลอง

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

$$Y_{ij} - \mu = \alpha_i + \varepsilon_{ij} \quad \text{(ความผันแปรทั้งหมด)}$$

$$(Y_{ij} - \mu)^2 = (\alpha_i + \varepsilon_{ij})^2$$

$$\sum_{i=1}^{t}\sum_{j=1}^{r}\left(Y_{ij}-\mu\right)^2 = \sum_{i=1}^{t}\sum_{j=1}^{r}\left(\alpha_i+\varepsilon_{ij}\right)^2$$

$$\Uparrow \qquad = \sum_{i=1}^{t}\sum_{j=1}^{r}\left(\alpha_i^2+\varepsilon_{ij}^2+2\alpha_i\varepsilon_{ij}\right)$$

$$\text{เรียกว่า total variation} = \sum_i\sum_j \alpha_i^2 + \sum_i\sum_j \varepsilon_{ij}^2 + \sum_i\sum_j 2\alpha_i\varepsilon_{ij}$$

$$\text{นั่นคือ total sum of square} = \sum_i\sum_j \alpha_i^2 + \sum_i\sum_j \varepsilon_{ij}^2$$

$$= \text{treatment variation} + \text{experimental error}$$

$$= \text{treatment sum of square} + \text{error sum of square}$$

ดังนั้น total SS = treatment SS + error SS

	T_1	T_2	•	•	•	•	•	T_t
	Y_{11}	Y_{21}	•	•	•	•	•	Y_{t1}
	Y_{12}	Y_{22}	•	•	•	•	•	Y_{t2}
	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•
	•	•	•	•	•	•	•	•
	Y_{1r}	Y_{2r}	•	•	•	•	•	Y_{tr}
Total	$Y_{1\bullet}$	$Y_{2\bullet}$	•	•	•	•	•	$Y_{t\bullet}$
mean	$\overline{Y}_{1\bullet}$	$\overline{Y}_{2\bullet}$	•	•	•	•	•	$\overline{Y}_{t\bullet}$

$Y_{1\bullet}$ = ผลรวมของทุกค่าสังเกตในทรีทเมนต์ที่ 1

$$= Y_{11} + Y_{12} + \ldots\ldots\ldots\ldots\ldots + Y_{1r} = \sum_{j=1}^{r} Y_{ij}$$

หรือ $Y_{i\bullet}$ = ผลรวมของทุกค่าสังเกตในทรีทเมนต์ที่ i

$$= \sum_{j=1}^{r} Y_{ij} = \textbf{treatment total}$$

$\overline{Y}_{1\bullet}$ = **ผลเฉลี่ยของทรีทเมนต์ที่ 1**

$\overline{Y}_{i\bullet}$ = **ผลเฉลี่ยของทรีทเมนต์ที่ i** $= \dfrac{\mathbf{Y}_{i\bullet}}{\mathbf{r}}$

$Y_{\bullet\bullet}$ = ผลรวมของทุกค่าสังเกตในการทดลองนี้

$= Y_{1\bullet} + Y_{2\bullet} + Y_{3\bullet} + \ldots\ldots\ldots\ldots + Y_{t\bullet}$

$$= \sum_{i=1}^{t} Y_{i\bullet} = \sum_{i=1}^{t}\sum_{j=1}^{r} Y_{ij}$$

= grand total

$\overline{Y}_{\bullet\bullet}$ **= ผลเฉลี่ยร่วม (grand mean หรือ overall mean)**

$$= \frac{\mathbf{Y}_{\bullet\bullet}}{\mathbf{tr}}$$

จาก total SS = treatment SS + error SS

$$\therefore \text{ total SS} = \sum_{i=1}^{t}\sum_{j=1}^{r}\left(Y_{ij} - \overline{Y}_{\bullet\bullet}\right)^2$$

treatment SS = ความผันแปรระหว่างทรีทเมนต์ ซึ่งตัวแทนของแต่ละทรีทเมนต์ คือ $\overline{y}_{i\bullet}$ และตัวแทนของทั้งหมด คือ $\overline{y}_{\bullet\bullet}$

$$= r\sum_{i=1}^{t}\left(\overline{Y}_{i\bullet} - \overline{Y}_{\bullet\bullet}\right)^2$$

error SS = ความผันแปรระหว่างหน่วยทดลอง

$$= \sum_{i=1}^{t}\sum_{j=1}^{r}\left(Y_{ij}-\overline{Y}_{i\bullet}\right)^2$$

สูตรเหล่านี้ เรียกว่า definition formula ซึ่งไม่นิยมใช้ในทางปฏิบัติ แต่จะใช้ working formula แทน

จาก total SS $= \sum_{i=1}^{t}\sum_{j=1}^{r}\left(Y_{ij}-\overline{Y}_{\bullet\bullet}\right)^2$

$$= \sum_{i}\sum_{j}Y_{ij}^2 - \left(\sum_{i}\sum_{j}Y_{ij}\right)^2 \Big/ tr$$

ซึ่ง $\left(\sum_{i}\sum_{j}Y_{ij}\right)^2 \Big/ tr$ **เรียกว่า ค่า correction term (C.T.)**

นั่นก็คือ **total SS** $= \sum_{i}\sum_{j}Y_{ij}^2 - \text{C.T.}$

treatment SS $= \sum_{i}Y_{i\bullet}^2 \big/ r - \text{C.T.}$

error SS (SSE) = total SS - treatment SS

จำนวนความเป็นอิสระ (degree of freedom)

df ของ total = tr-1

df ของ treatment = t-1

df ของ error = tr-t = t(r-1)

หรือ = total df - treatment df

= (tr-1)-(t-1) = t(r-1)

mean square (MS) คือ ค่าวาเรียนซ์ที่ใช้วัดความผันแปรเฉลี่ยต่อหนึ่งหน่วยค่าสังเกต

ดังนั้น $$\text{MS} = \frac{\text{SS}}{\text{df}}$$

$$\therefore \text{ treatment mean square (MST)} = \frac{\text{SST}}{t-1}$$

$$\text{error mean square (MSE)} = \frac{\text{SSE}}{t(r-1)}$$

การวิเคราะห์วาเรียนซ์ (analysis of variance, ANOVA)

sources of variation	df	SS	MS	F - ratio
treatment	t – 1	SST	$\text{MST} = \frac{\text{SST}}{t-1}$	$\frac{\text{MST}}{\text{MSE}}$
error	t(r-1)	SSE	$\text{MSE} = \frac{\text{SSE}}{t(r-1)}$	
total	tr-1	total SS		

วิธีการทดสอบ

สมมุติฐาน สำหรับการทดสอบ คือ

H_0 : $\mu_1 = \mu_2 = \mu_3 = ... = \mu_t$

H_1 : อย่างน้อยมีค่าเฉลี่ย 2 ค่า ที่ไม่เท่ากัน

หรือ H_0 : ไม่มีอิทธิพลของทรีทเมนต์

H_1 : มีอิทธิพลของทรีทเมนต์

ทั้งค่า MST และ MSE ต่างก็เป็นค่าประมาณของ σ^2 ซึ่ง MST รวมกับ MSE จะไม่เท่ากับ total mean square เพราะค่า MSE เป็นค่าประมาณของ σ^2 ที่ไม่มีผลของทรีทเมนต์รวมอยู่ นั่นคือ MSE = $\hat{\sigma}^2 = S^2$ แต่ MST เป็นค่าประมาณของ σ^2 รวมกับผลของทรีทเมนต์อยู่ด้วย หรือ MST = $\hat{\sigma}^2 + r\hat{\sigma}_\alpha^2$

ถ้า H_0 เป็นจริงแล้วก็จะไม่มีผลของทรีทเมนต์ ซึ่งค่า $r\hat{\sigma}_\alpha^2$ จะเท่ากับ ศูนย์ ค่า MST ก็จะเท่ากับ MSE

ถ้า H_0 ไม่เป็นจริง ค่า $r\hat{\sigma}_\alpha^2$ จะไม่เท่ากับ ศูนย์ ค่า MST จะต้องมากกว่า MSE

$$F-ratio = \frac{MST}{MSE}$$

ดังนั้นเมื่อ H_0 เป็นจริง ค่า F - ratio ก็จะเข้าใกล้ 1 หรือเท่ากับ 1 แต่เมื่อใด F - ratio มากกว่า 1 ก็แสดงว่ามีอิทธิพลของทรีทเมนต์ เราก็จะปฏิเสธการรับ H_0

ในการทดสอบ เราจะนำค่า F ที่คำนวณได้ ไปเทียบกับค่า F ในตารางที่ *df* เท่ากับ t-1, t(r-1) เมื่อค่า F คำนวณมีค่าน้อยกว่า F ในตาราง เราก็จะยอมรับ H_0 แต่ถ้าค่า F คำนวณมากกว่า F ในตาราง เราก็ไม่ยอมรับ H_0

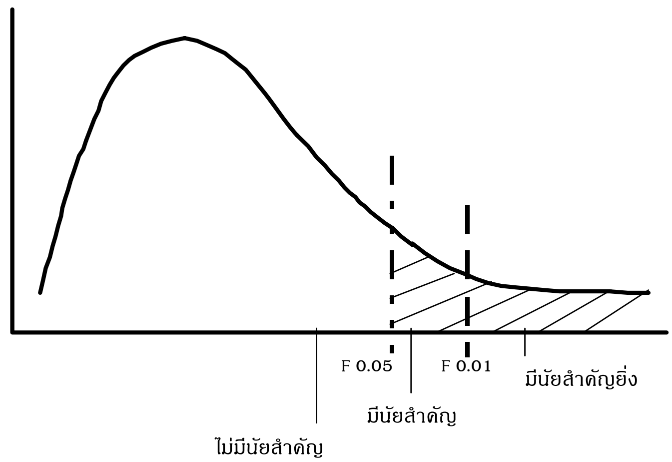
การสรุปผล

ถ้าการทดสอบใช้ค่า α = 0.05 แล้ว ปรากฏว่า ค่า F คำนวณมากกว่า F ตาราง เราก็สรุปว่า มีความแตกต่างกันระหว่างทรีทเมนต์อย่างมีนัยสำคัญ (ใส่เครื่องหมาย * หนึ่งอัน บนค่า F ในตารางวิเคราะห์วาเรียนซ์) แต่ถ้าใช้ α = 0.01 ก็จะสรุปว่ามีความแตกต่างอย่างมีนัยสำคัญยิ่ง ทั้งนี้เพราะ ค่า F ที่ 0.01 จะมีค่ามากกว่า F ที่ 0.05 (ใส่เครื่องหมาย ** สองอัน บนค่า F ในตารางวิเคราะห์วาเรียนซ์)

S.O.V.	*df*	SS	MS	F-ratio
treatment	t-1	SST	MST	$\frac{MST}{MSE}^{**}$
error	t(r-1)	SSE	MSE	
total	tr-1			

แต่เมื่อ ค่า F คำนวณน้อยกว่า ค่า F ในตารางที่ $\alpha = 0.05$ เราก็จะสรุปว่า มีความแตกต่างระหว่างทรีทเมนต์อย่างไม่มีนัยสำคัญทางสถิติ (ใส่ ns ไว้บนค่า F ในตารางวิเคราะห์ว่าเรียนซ์)

F



ค่าประมาณของ σ^2 ก็คือ MSE (error mean square) ดังนั้นค่าความคลาดเคลื่อนมาตรฐานของค่าเฉลี่ยของทรีทเมนต์ (Standard deviation of treatment mean, $S_{\bar{Y}}$) และความคลาดเคลื่อนมาตรฐานของผลต่างระหว่างค่าเฉลี่ยของทรีทเมนต์(Standard deviation of the difference, $S_{\bar{Y}_1-\bar{Y}_2}$) จะหาได้ดังนี้

$$S_{\overline{Y}} = \sqrt{\frac{MSE}{r}}$$

$$S_{\overline{Y}_1-\overline{Y}_2} = \sqrt{\frac{2\,MSE}{r}}$$

สัมประสิทธิ์ความผันแปรของค่าเฉลี่ยของทรีทเมนต์วัดได้จาก

$$\text{Coefficient of variation (C.V.)} = \frac{S}{\overline{Y}} \times 100 \quad \%$$

$$= \frac{\sqrt{MSE}}{\overline{Y}} \times 100 \quad \%$$

ตัวอย่างที่ 4.1 **การศึกษาเปรียบเทียบการเลี้ยงสุกรขุนจากน้ำหนัก 60 กิโลกรัม จนถึงส่งขายตลาดที่น้ำหนัก 90 กิโลกรัม โดยใช้อาหารที่มีพลังงานที่ใช้ประโยชน์ได้ต่างกัน 3 ระดับ คือ 3150, 2800 และ 2600 กิโลแคลอรี่ต่อกิโลกรัมอาหาร ในการทดลองนี้ใช้สุกรพันธุ์ลาร์จไวท์ จำนวน 24 ตัว เป็นเพศผู้ตอน 12 ตัว เพศเมีย 12 ตัว ทำการสุ่มสุกรเข้าขังในคอกทดสอบ คอกละ 2 ตัว เป็นเพศผู้ตอน 1 ตัว และเพศเมีย 1 ตัว จากนั้นสุ่มอาหารทดลองทั้ง 3 สูตรให้กับสุกร โดยจะมีสุกรที่ได้รับอาหารทดลองสูตรเดียวกันอยู่สูตรละ 4 คอก ซึ่งแผนผังการทดลองมีดังนี้**

B 0.74	C 0.73	A 0.77	C 0.72
A 0.80	B 0.70	A 0.79	C 0.75
C 0.71	B 0.75	B 0.76	A 0.78

อัตราการเจริญเติบโต (กก. ต่อวัน) ของสุกร ในช่วงการทดลองเฉลี่ยแต่ละคอก มีดังนี้

คอกที่	พลังงาน (กิโลแคลอรี่ ต่อ กก.)			
	(A) 3150	(B) 2800	(C) 2600	
1	0.77	0.74	0.73	
2	0.80	0.70	0.72	
3	0.79	0.75	0.75	
4	0.78	0.76	0.71	
ผลรวม	3.14	2.95	2.91	9.0
เฉลี่ย	0.79	0.74	0.73	

การวิเคราะห์

$$C.T. = \left(\sum_i \sum_j Y_{ij} \right)^2 \Big/ tr$$

$$= (9.0)^2 \big/ 3 \times 4$$

$$= 6.75$$

total SS $= \sum_i \sum_j Y_{ij}^2 - C.T.$

$$= (0.77)^2 + (0.80)^2 + \ldots\ldots\ldots\ldots\ldots\ldots + (0.71)^2 - C.T.$$

$$= 6.761 - 6.75$$

$$= 0.011$$

treatment SS $= \sum_i (Y_{i\bullet})^2 \big/ r - C.T.$

$$= \frac{(3.14)^2 + (2.095)^2 + (2.91)^2}{4} - 6.75$$

$$= 6.757 - 6.75$$

$$= 0.007$$

error SS = total SS - treatment SS

= 0.011-0.007

= 0.004

การวิเคราะห์วาเรียนซ์

S.O.V.	*df*	SS	MS	F-ratio
treatment	3-1 = 2	0.007	0.0035	8.75**
error	3(4-1) = 9	0.004	0.0004	
total	12-1 = 11	0.011		

จากตาราง $F_{0.01(2, 9)} = 8.02$

สรุปผลการทดลอง: ระดับพลังงานในอาหารมีผล ทำให้อัตราการเจริญเติบโตของสุกรขุน แตกต่างกันอย่างมีนัยสำคัญยิ่ง (P<0.01)

แผนการทดลองแบบสุ่มตลอดที่มีจำนวนซ้ำไม่เท่ากัน

ที่กล่าวมาในตอนต้นนั้น จะเป็นแผนการทดลองแบบสุ่มตลอด ซึ่งมีจำนวนซ้ำ ของแต่ละทรีทเมนต์เท่ากัน ในบางครั้งเมื่อแรกเริ่มทดลอง เราอาจจะจัดให้ทุกทรีทเมนต์มีจำนวนซ้ำเท่ากันหมด แต่เมื่อทดลองไปเกิดมีสัตว์ทดลองตาย หรือจำเป็นต้องคัดออกจากการทดลองทำให้เราไม่สามารถบันทึกค่าสังเกตตัวนั้นได้ ก็จะทำให้เกิดกรณีที่ว่า ในแต่ละทรีทเมนต์มีจำนวนซ้ำไม่เท่ากัน เราเรียกว่าเป็นแผนการทดลองแบบสุ่มตลอดที่มีจำนวนซ้ำไม่เท่ากัน unequal replication (CRD) วิธีวิเคราะห์ข้อมูลก็จะยังคงเหมือนเดิม เพียงแต่จำนวนซ้ำของแต่ละทรีทเมนต์ไม่เท่ากันเท่านั้น

T_1	T_2	•	•	•	•	•	T_t
ตาย Y_{11}	Y_{21}	•	•	•	•	•	Y_{t1}
Y_{12}	Y_{22}	•	•	•	•	•	ตาย Y_{t2}
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
•	•	•	•	•	•	•	•
คัดออก Y_{11}	Y_{2r}	•	•	•	•	•	Y_{tr}

สมมติให้ r_i เป็นจำนวนซ้ำในทรีทเมนต์เมนต์ที่ i

นั่นก็คือ ทรีทเมนต์ที่ 1 มีจำนวนซ้ำ = r_1

ให้ n = จำนวนค่าสังเกตทั้งหมด

$$= r_1 + r_2 + r_3 + \dots\dots\dots + r_t = \sum_{i=1}^{t} r_i$$

ตารางการวิเคราะห์วาเรียนซ์

S.O.V.	*df*	SS	MS	F - ratio
treatment	t-1	SST	MST	$\frac{MST}{MSE}^{**}$
error	$\sum r_i$-t	SSE	MSE	
total	$\sum r_i$-1	total SS		

ตัวอย่างที่ 4.2 การศึกษาผลของอาหาร 3 สูตรที่มีต่อน้ำหนักไข่(กรัม) ของไก่ โดยใช้ไก่พันธุ์เดียวกัน อายุการไข่เท่ากัน จำนวน 150 ตัว แยกขังเป็น 15 คอก ๆ ละ 10 ตัว ดังนั้น อาหารแต่ละสูตรจะใช้เลี้ยงไก่ 5 คอก แต่ขณะทำการทดลองจำเป็นต้องคัดไก่บางคอกออก เนื่องจากป่วย ค่าเฉลี่ยน้ำหนักไข่ของไก่แต่ละคอกเป็นดังนี้

	อาหารสูตร			
	A	B	C	
	51	58	59	
	43	60	57	
	49	60	55	
		53	54	
		55		
total	143	286	225	654
mean	47.67	57.20	56.25	

$$C.T. = \left(\sum_{ij} Y_{ij}\right)^2 \Big/ n$$

$$= (654)^2 / 12 = 35,643$$

$$\text{total SS} = \sum_{ij} Y_{ij}^2 - C.T.$$

$$= 51^2 + 43^2 + \ldots\ldots\ldots\ldots + 54^2 - C.T.$$

$$= 35{,}920 - 35{,}643$$

$$= 277$$

error SS = 277-188.78 = 88.22

การวิเคราะห์วาเรียนซ์

S.O.V.	*df*	SS	MS	F-ratio
treatment	3-1 = 2	188.78	94.39	9.63**
error	11-2 = 9	88.22	9.80	
total	12-1 =11	277		

จากตาราง $F_{0.01(2, 9)} = 8.02$

สรุปผล: อาหารทั้ง 3 สูตรจะทำให้น้ำหนักไข่ของไก่แตกต่างกันอย่างมีนัยสำคัญยิ่ง (P<0.01)

การเปรียบเทียบค่าเฉลี่ยของทรีทเมนต์

ในการวิเคราะห์วาเรียนซ์ เราสรุปได้แต่เพียงว่า ทรีทเมนต์มีความแตกต่างกัน แต่ไม่สามารถบอกได้ว่า ทรีทเมนต์ใดบ้างที่แตกต่างกัน ดังนั้นหลังจากทำการวิเคราะห์วาเรียนซ์แล้วพบว่า ไม่มีนัยสำคัญเราก็จะหยุดเพียงแค่นั้น แต่ถ้ามีนัยสำคัญเราก็จะทำการทดสอบต่อไปว่า มีทรีทเมนต์คู่ใดบ้างที่แตกต่างกัน การเปรียบเทียบค่าเฉลี่ยทรีทเมนต์เป็นคู่ๆ จนหมดทุกคู่ที่จะเป็นไปได้ในการทดลองนั้น เราเรียกว่า multiple comparisons ซึ่งมีอยู่หลายวิธีด้วยกัน เช่น

Least significant difference (*lsd*)

เป็นค่าความแตกต่างที่น้อยที่สุดที่จะถือว่ามีนัยสำคัญ ถ้าคู่ใดคู่หนึ่งมีค่ามากกว่าค่า *lsd* ก็แสดงว่ามีนัยสำคัญ

$$\overline{Y}_1 - \overline{Y}_2 \rangle \, lsd \quad \text{แสดงว่ามีนัยสำคัญ}$$

วิธีการนี้บางครั้งอาจจะเรียกว่า unplanned comparisons เพราะไม่ได้มีการวางแผนไว้ล่วงหน้าว่ามีจุดประสงค์จะเปรียบเทียบอะไรกับอะไร วิธี *lsd* มีหลักเหมือนกับ t-test ทุกอย่าง

จำนวนคู่ของค่าเฉลี่ยทรีทเมนต์ที่จะเปรียบเทียบกันได้จะเท่ากับ

$$\frac{t!}{2!\,(t-2)!}$$

ถ้าการทดลองมี 3 ทรีทเมนต์ ค่าเฉลี่ยทรีทเมนต์ที่จะเปรียบเทียบกันได้ก็จะเท่ากับ

$$\frac{3!}{2!(3-2)!} = 3 \text{ คู่}$$

สมมุติฐานของการทดลองก็จะเป็น

$H_0: \mu_1 = \mu_2 ; \mu_1 = \mu_3; \ \mu_2 = \mu_3$

$H_A: \mu_1 \neq \mu_2 ; \mu_1 \neq \mu_3; \ \mu_2 \neq \mu_3$

เมื่อแยกทดสอบก็จะเป็น $H_0: \mu_1 = \mu_2$

$H_A: \mu_1 \neq \mu_2$

$$t = \frac{(\bar{Y}_1 - \bar{Y}_2) - (\mu_1 - \mu_2)}{S_{\bar{Y}_1 - \bar{Y}_2}}$$

$$= \frac{(\bar{Y}_1 - \bar{Y}_2)}{S_{\bar{Y}_1 - \bar{Y}_2}} \qquad \text{เมื่อ} \quad \mu_1 = \mu_2$$

$$= \frac{\bar{Y}_1 - \bar{Y}_2}{S_{\bar{d}}}$$

$$S^2_{\bar{Y}_1 - \bar{Y}_2} = S^2_{\bar{d}} = \frac{S_1^2}{n_1} + \frac{S_2^2}{n_2} = S_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

$$= \frac{2S_p^2}{r} \quad (\mathbf{n_1 = n_2 = r = \text{จำนวนซ้ำ}})$$

$$= \frac{2MSE}{r}$$

$$\therefore S_{\bar{d}} = \sqrt{\frac{2\,MSE}{r}}$$

เราจะไม่ยอมรับ H$_0$ เมื่อ $t \rangle t_{\frac{\alpha}{2}}$ **หรือ** $t \langle -t_{\frac{\alpha}{2}}$

หรือ $\frac{\overline{Y}_1 - \overline{Y}_2}{S_{\bar{d}}} \rangle t_{\frac{\alpha}{2}}$ หรือ $\frac{\overline{Y}_1 - \overline{Y}_2}{S_{\bar{d}}} \langle -t_{\frac{\alpha}{2}}$

นั่นคือ ไม่ยอมรับ H$_0$ เมื่อ

$\overline{Y}_1 - \overline{Y}_2 \rangle t_{\frac{\alpha}{2}} \bullet S_{\bar{d}}$ หรือ $\overline{Y}_1 - \overline{Y}_2 \langle -t_{\frac{\alpha}{2}} \bullet S_{\bar{d}}$

ซึ่ง $t_{\frac{\alpha}{2}} \bullet S_{\bar{d}}$ **ก็คือ ค่า** $lsd_{\frac{\alpha}{2}}$

ขั้นตอนการวิเคราะห์

1) จัดเรียงค่าเฉลี่ยของทรีทเมนต์จากค่าน้อยไปหาค่ามาก หรือจากค่ามากไปหาค่าน้อย

2) คำนวณหาจำนวนคู่ที่สามารถเปรียบเทียบได้ (comparisons) ทั้งหมด จาก

$$\binom{t}{2} = \frac{t!}{2!(t-2)!}$$

t คือ จำนวนค่าเฉลี่ยของทรีทเมนต์

2 คือ จำนวนที่เลือกมาโดยวิธี without replacement

3) คำนวณหา ค่า $lsd_{\frac{\alpha}{2}}$

$$lsd_{\frac{\alpha}{2}} = t_{\frac{\alpha}{2}} \bullet S_{\bar{d}}$$

$t_{\frac{\alpha}{2}}$ **เป็นค่า t จากตาราง t ที่ *df* เท่ากับ *df* ของความคลาดเคลื่อน หรือ t(r-1)**

$$S_{\bar{d}} = \sqrt{\frac{2\,MSE}{r}}$$

4) คำนวณผลต่างของค่าเฉลี่ยของทรีทเมนต์ทุกคู่ โดยเอาค่าเฉลี่ยที่มากตั้ง ลบด้วยค่าเฉลี่ยที่น้อย เช่น

$$\overline{Y}_3, \quad \overline{Y}_1, \quad \overline{Y}_2$$

$$\overline{Y}_2 - \overline{Y}_3 = A^*$$

$$\overline{Y}_2 - \overline{Y}_1 = B^{ns}$$

$$\overline{Y}_1 - \overline{Y}_3 = C^{ns}$$

นำผลต่างที่ได้ (A, B, C) ไปเทียบกับค่า $lsd_{\frac{\alpha}{2}}$ ในข้อ 3) ถ้าคู่ไหนน้อยกว่า $lsd_{\frac{\alpha}{2}}$ ก็ใส่ ns (non-significant) ถ้าคู่ไหนมากกว่า $lsd_{\frac{\alpha}{2}}$ ก็ใส่เครื่องหมาย * หนึ่งอัน เมื่อ α = 0.05 หรือใส่เครื่องหมาย ** เมื่อ α = 0.01

5) การสรุปผล โดยการขีดเส้นใต้ค่าเฉลี่ยของทรีทเมนต์คู่ที่ไม่มีนัยสำคัญ เช่น

$$\underline{\bar{Y}_3 \quad \bar{Y}_1} \quad \bar{Y}_2$$
$$\bar{Y}_3 \quad \underline{\bar{Y}_1 \quad \bar{Y}_2}$$

หรือจะใช้ตัวอักษรกำกับข้างบนค่าเฉลี่ย โดยใช้อักษรเหมือนกันสำหรับค่าเฉลี่ยที่ไม่มีนัยสำคัญเช่น

$$\bar{Y}_3^{\text{ก}} \qquad \bar{Y}_1^{\text{กข}} \qquad \bar{Y}_2^{\text{ข}}$$

หมายความว่า $\bar{Y}_1$ กับ $\bar{Y}_2$ นั้นแตกต่างกันอย่างไม่มีนัยสำคัญ ขณะเดียวกัน $\bar{Y}_1$ กับ $\bar{Y}_3$ ก็แตกต่างกันอย่างไรไม่มีนัยสำคัญ แต่ $\bar{Y}_2$ กับ $\bar{Y}_3$ จะแตกต่างกันอย่างมีนัยสำคัญ

จากตัวอย่างที่ 4.1 ค่าเฉลี่ยอัตราการเจริญเติบโตของสุกร ที่ได้รับอาหารที่มีพลังงานต่างกัน 3 ระดับ คือ 0.79, 0.74 และ 0.73 กก. ต่อวัน ตามลำดับ และค่า MSE เท่ากับ 0.0004

จัดเรียง ค่าเฉลี่ยของทรีทเมนต์จากน้อยไปหามาก

$\overline{Y}_C$	$\overline{Y}_B$	$\overline{Y}_A$
0.73	**0.74**	**0.79**

จำนวนคู่ที่จะเปรียบเทียบได้ทั้งหมด คือ

$$\binom{3}{2} = \frac{3!}{2!(3-2)!}$$

$$= \frac{3\times2\times1}{2\times1(1)} = 3 \text{ คู่}$$

$$S_{\bar{d}} = \sqrt{\frac{2\,MSE}{r}}$$

$$= \sqrt{\frac{2(0.0004)}{4}}$$

$$= 0.014$$

$$t_{\frac{\alpha}{2}(\text{error } df)} = t_{\frac{0.01}{2}(9)} = 3.250$$

$$lsd_{\frac{\alpha}{2}} = lsd_{\frac{0.01}{2}}$$

$$= 3.25\times0.014$$

$$= 0.045$$

$$\bar{Y}_A - \bar{Y}_C = 0.79 - 0.73 = 0.06^{**} > 0.045$$
$$\bar{Y}_A - \bar{Y}_B = 0.79 - 0.74 = 0.05^{**} > 0.045$$
$$\bar{Y}_B - \bar{Y}_C = 0.74 - 0.73 = 0.01^{ns} < 0.045$$

สรุปผล: $\bar{Y}_C$ ก $\bar{Y}_B$ ก $\bar{Y}_A$ ข

0.73 **0.74** **0.79**

นั่นคือ สุกรที่ได้รับอาหารที่มีพลังงาน 3150 กิโลแคลอรี่ ต่อ กก. อาหาร มีอัตราการเพิ่มน้ำ-หนักตัวสูงกว่าพวกที่ได้รับพลังงาน 2800 และ 2600 กิโลแคลอรี่ ต่อ กก. อาหาร อย่างมีนัยสำคัญยิ่งทางสถิติ แต่พวกที่ได้รับพลังงาน 2800 และ 2600 กิโลแคลอรี่ ต่อ กก. อาหาร จะมีอัตราการเพิ่มน้ำหนักตัว ต่อวัน ไม่แตกต่างกันอย่างมีนัยสำคัญทางสถิติ (P>0.05)

Duncan's New Multiple Range Test (DMRT)

ในวิธี *lsd* นั้นเราไม่ได้คำนึงถึงจำนวนทรีทเมนต์ที่ใช้ในการทดสอบ แต่วิธี Duncan's จะคำนึงถึงจำนวนทรีทเมนต์ในการทดสอบ วิธีการวิเคราะห์จะเหมือนกับวิธี *lsd* แต่จะต่างกันที่ขั้นที่ 3 โดยจะหาค่า least significant ranges (LSR) แทนค่า *lsd*

$$LSR = SSR \bullet S_{\bar{Y}}$$

$$S_{\bar{Y}} = \sqrt{\frac{MSE}{r}}$$

SSR หรือ Significant Studentized Ranges ซึ่งหาได้จากตารางผนวกที่ 6 ค่าความเป็นอิสระ เท่ากับ *df* ของความคลาดเคลื่อน เช่น ถ้าจำนวนทรีทเมนต์เท่ากับ 3 จำนวนค่าเฉลี่ยของทรีทเมนต์ที่อยู่ในการทดสอบ (p) ก็จะมีได้ตั้งแต่ 2-3

จากตัวอย่างเดิม ค่าเฉลี่ยอัตราการเจริญเติบโตของสุกรที่ได้รับอาหารที่มีระดับพลังงานต่าง ๆ กันเรียงตามลำดับ จากค่าน้อยไปหาค่ามาก ได้แก่

อาหารสูตร	**C**	**B**	**A**
อัตราการเจริญเติบโตเฉลี่ย	**0.73**	**0.74**	**0.79 กก. ต่อวัน**

$$S_{\bar{Y}} = \sqrt{\frac{MSE}{r}} = \sqrt{\frac{0.0004}{4}} = 0.01$$

จากตารางค่า $SSR_{0.01}$ ที่ $df = 9$ จะมีค่าเฉลี่ยของทรีทเมนต์ที่อยู่ในการทดสอบ ตั้งแต่ 2-3

P	**2**	**3**
$SSR_{0.01}$	**4.60**	**4.86**
$LSR_{0.0}$	**4.60×0.01**	**4.86×0.01**
	= 0.046	**= 0.049**

การเปรียบเทียบความแตกต่าง: **เริ่มจากค่าเฉลี่ยสูงสุดกับค่าเฉลี่ยต่ำสุด เปรียบเทียบกับค่า LSR ที่สูงสุด ถัดมาก็ค่าเฉลี่ยสูงสุดกับค่าเฉลี่ยรองต่ำสุด และถัดขึ้นไปเรื่อย ๆ จนถึงกับรองสูงสุด โดยเทียบกับค่า LSR ที่รองสูงสุด ถัดไปเรื่อยๆ จนถึง LSR ต่ำสุด แล้วจึงเริ่มเปรียบเทียบค่าเฉลี่ยรองสูงสุดกับค่าเฉลี่ยต่ำสุด กับรองต่ำสุดขึ้นไปเรื่อย ๆ โดยเทียบกับ LSR รองสูงสุดจนถึงต่ำสุด ทำไปจนกระทั่งเป็นการเปรียบเทียบค่าเฉลี่ยรองต่ำสุดกับต่ำสุด โดยเทียบกับค่า LSR ที่ต่ำที่สุด ถ้าผลต่างระหว่างค่าเฉลี่ยคู่ใดมีค่ามากกว่า LSR ที่เปรียบเทียบ ก็หมายถึงมีนัยสำคัญ ซึ่งในที่นี้การเปรียบเทียบก็คือ**

$$\bar{Y}_A - \bar{Y}_C = 0.79 - 0.73 = 0.06^{**} \rangle\ 0.049$$
$$\bar{Y}_A - \bar{Y}_B = 0.79 - 0.74 = 0.05^{**} \rangle\ 0.046$$
$$\bar{Y}_B - \bar{Y}_C = 0.74 - 0.73 = 0.01^{ns} \langle\ 0.046$$

สรุปผล:

$\bar{Y}_C$ ก	$\bar{Y}_B$ ก	$\bar{Y}_A$ ข
0.73	**0.74**	**0.79**

นั่นคือ อัตราการเจริญเติบโตต่อวันของสุกรที่ได้รับอาหารพลังงาน 3150 กิโลแคลอรี่ ต่อ กก. อาหาร สูงกว่าอัตราการเจริญเติบโตของพวกที่ได้รับอาหารพลังงาน 2800 และ 2600 กิโลแคลอรี่ ต่อ กก. อาหาร อย่างมีนัยสำคัญยิ่งทางสถิติ (P<0.01) แต่ไม่มีความแตกต่างอย่างมีนัยสำคัญทางสถิติ

(P>0.05) ระหว่างพวกที่ได้รับพลังงาน 2800 และ 2600 กิโลแคลอรี่ ต่อ กก. อาหาร

Orthogonal comparison

เป็นวิธีการที่นิยมใช้กับการทดลองที่มีการตั้งคำถามไว้ล่วงหน้าก่อนทำการทดสอบว่า หลังจากการวิเคราะห์วาเรียนซ์แล้ว ต้องการจะเปรียบเทียบความแตกต่างของทรีทเมนต์ใดบ้างจึงเป็น planned comparisons ซึ่งจำนวนคำถาม หรือคู่เปรียบเทียบ จะมีได้ไม่เกิน *df* ของทรีทเมนต์ โดยแต่ละคู่เปรียบเทียบจะมี *df* เท่ากับ 1 เมื่อมีการตั้งคู่เปรียบเทียบแล้วจะต้องจัดค่าสัมประสิทธิ์ ให้กับแต่ละทรีทเมนต์ในคู่เปรียบเทียบนั้น ผลรวมของค่าสัมประสิทธิ์ในแต่ละคู่เปรียบเทียบจะต้องเป็นศูนย์ นอกจากนี้แล้ว ค่าเปรียบเทียบที่เลือกนั้นจะต้องเป็นอิสระต่อกันด้วย

กำหนด $T_1, T_2, \ldots\ldots\ldots\ldots, T_t$ เป็นผลรวมของทรีทเมนต์

$$Q_l = \sum_{i=1}^{t} C_{li} T_i$$

Q_l คือ คู่เปรียบเทียบที่ *l* อาจจะเรียกเป็น comparison หรือ contrast ก็ได้

C_{li} คือ สัมประสิทธิ์เป็นเลขจำนวนเต็ม ที่ให้กับทรีทเมนต์ที่ i ในคู่เปรียบเทียบที่ l

$$\sum_{i=1}^{t} C_{li} = 0$$

ดังนั้น

$$Q_1 = \sum_{i=1}^{t} C_{1i} T_i$$

$$Q_2 = \sum_{i=1}^{t} C_{2i} T_i$$

Q_1 กับ Q_2 จะเป็นอิสระต่อกันเมื่อ

$$\sum_{i=1}^{t} C_{1i} C_{2i} = 0$$

การทดสอบความแตกต่างของค่าเฉลี่ยทรีทเมนต์ โดยวิธีนี้เป็นการแบ่งค่า treatment sum of square ออกเป็น sum of square ของแต่ละคู่เปรียบเทียบ

comparison sum of square $= SS(Q_l) = \dfrac{Q_l^2}{r\sum_{i=1}^{t} C_{li}^2}$

ซึ่งผลรวม sum of square ของทุกคู่เปรียบเทียบจะต้องเท่ากับ treatment sum of square

ตัวอย่างที่ 4.3 การศึกษาผลการเสริมวิตามิน บี12 ในอาหารสุกรได้น้ำหนักเพิ่มของสุกรต่อวัน เป็นกรัม ดังนี้

	ระดับวิตามิน บี12 (มก.)		
	(T_1)	(T_2)	(T_3)
	0	5	10
	0.472	0.689	0.739
	0.454	0.708	0.712
	0.313	0.699	0.699
รวม	1.239	2.096	2.150
เฉลี่ย	0.413	0.699	0.717

ตารางวิเคราะห์วาเรียนซ์

Source	*df*	SS	MS	F-ratio
treatment	2	0.174	0.087	29.00**
error	6	0.016	0.003	

ตั้งคู่เปรียบเทียบได้ดังนี้

1) พวกที่ไม่เสริมวิตามิน กับ พวกที่เสริมวิตามิน (T_1 vs T_2, T_3)

2) พวกเสริมวิตามินระดับ 5 มก. กับ 10 มก. (T_2 vs T_3)

	T_1	T_2	T_3	$\sum C_i^2$	Q	SS
Q_1: T_1 vs T_2, T_3:	-2	+1	+1	6	1.768	0.1736
Q_2: T_2 vs T_3 :	0	-1	+1	2	0.054	0.0004

จะเห็นว่า $$\sum C_{1i} = \sum C_{2i} = 0$$

และ $$\sum C_{1i} C_{2i} = 0$$

$$Q_1 = (-2)(1.239)+(1)(2.096)+(1)(2.15) = 1.768$$

$$Q_2 = (0)(1.239)+(-1)(2.096)+(1)(2.15) = 0.054$$

$$SS(Q_1) = \frac{Q_1^2}{r\sum_{i=1} C_{1i}^2} = \frac{(1.768)^2}{3(6)}$$

$$= 0.1736$$

$$SS(Q_2) = \frac{Q_2^2}{r\sum_{i=1} C_{2i}^2} = \frac{(0.054)^2}{3(2)}$$

$$= 0.0004$$

ตารางวิเคราะห์วาเรียนซ์

Source	*df*	SS	MS	F-ratio
treatment	2	0.174	0.087	29.00**
Q_1	1	0.1736	0.1736	57.87**
Q_2	1	0.0004	0.0004	0.13<1 ns
error	6	0.016	0.003	

สรุปผล: สุกรพวกที่ได้รับเสริมวิตามิน บี12 ในระดับต่างกันจะมีความแตกต่างอย่างไม่มีนัยสำคัญ แต่พวกที่ได้รับเสริมวิตามิน บี12 มีน้ำหนักเพิ่มต่อวัน มากกว่าพวกไม่ได้รับเสริมวิตามินอย่างมีนัยสำคัญยิ่ง (P<0.01)

แผนการทดลองแบบสุ่มตลอดเมื่อมีตัวอย่างย่อย

ในการวางแผนทดลองบางครั้งเราเก็บข้อมูลจากหน่วยทดลองได้มากกว่า 1 ค่า ซึ่งข้อมูลที่เก็บมานั้นจะได้มาจากหน่วยย่อยของหน่วยทดลอง เราเรียก หน่วยย่อยนี้ว่า Sampling units หรือ sub-samples เช่น ในการเก็บข้อมูลเลือดจากสัตว์เพื่อศึกษาฮีโมโกลบิน (haemoglobin) เนื่องจากการหาฮีโมโกลบินมักจะคลาดเคลื่อนได้ง่าย เราจึงเก็บตัวอย่างสามครั้งจากสัตว์ตัวเดียวกัน ซึ่งเป็นหน่วยทดลองอยู่แล้ว อาจจะเขียนเป็นแผนผังดังนี้

พันธุ์	ตัวที่	จำนวนครั้งในการหาฮีโมโกลบิน
A	1	1
		2
		3
	2	1
		2
		3
	3	1
		2
		3
B	1	1

		2
		3
	2	1
		2
		3
	3	1
		2
		3
C	1	1
		2
		3
	2	1
		2
		3
	3	1
		2
		3

จะเห็นว่า ข้อมูลที่เก็บในแต่ละครั้งนั้น จะเป็นตัวแทนที่อยู่ในหน่วยทดลอง (samples within experimental units) นั่นเอง หรือ อาจกล่าวได้ว่า ในหนึ่งหน่วยทดลองจะมีหน่วยย่อย 3 หน่วยด้วยกัน บางครั้งเราจะเรียกการทดลองที่มีหน่วยย่อยนี้ว่า Nested Classification หรือ Hierarchal Classification

แบบหุ่นของการทดลองแบบสุ่มตลอดที่มีหน่วยย่อยได้แก่

$$Y_{ijk} = \mu + A_i + B_{ij} + \varepsilon_{ijl};$$

$$B_{ij} \sim N(0,\ \sigma_B^2),$$
$$\varepsilon_{ijk} \sim N(0,\ \sigma^2)$$

เมื่อ Y_{ijk} เป็นข้อมูลที่เก็บจากหน่วยย่อยที่ k ซ้ำๆ ที่ j ของทรีทเมนต์ที่ i

A_i เป็นอิทธิพลของทรีทเมนต์ที่ i; i = 1, 2, ..., t

B_{ij} เป็นความคลาดเคลื่อนที่เกิดจากหน่วยทดลอง; j = 1, 2, ..., r

ε_{ijk} เป็นความคลาดเคลื่อนที่เกิดจากหน่วยย่อย; k = 1,2, ...,

s

ตัวอย่างที่ 4.4 ในการเปรียบเทียบความสามารถของสุกรพันธุ์แท้ 3 พันธุ์ สุ่มแม่สุกรแต่ละพันธุ์มา 3พันธุ์ แล้วสุ่มลูกสุกรจากแม่พันธุ์แต่ละแม่มา 2 ตัว เพื่อหาอัตราการเจริญเติบโตต่อวันได้ผลดังนี้

พันธุ์	แม่พันธุ์	การเจริญเติบโตของลูกสุกรต่อวัน (กก.)		รวมของแม่พันธุ์	รวมของพันธุ์	รวมทั้งหมด
i=1,2,3	j=1,2,3	ตัวที่ 1	ตัวที่ 2	($Y_{ij.}$)	($Y_{i..}$)	($Y_{...}$)
A	1	1.25	1.38	2.63		
	2	1.31	1.33	2.64		
	3	1.29	1.36	2.65	7.92	

B	**1**	**1.01**	**1.07**	**2.08**		
	2	**1.22**	**1.24**	**2.46**		
	3	**1.05**	**1.02**	**2.07**	**6.61**	
C	**1**	**1.30**	**1.26**	**2.56**		
	2	**1.24**	**1.16**	**2.40**		
	3	**1.28**	**1.25**	**2.53**	**7.49**	**22.02**

$$C.T. = \frac{(22.02)^2}{18} = 26.94$$

total SS = $(1.25)^2 + (1.38)^2 + \dots\dots + (1.25)^2$ - C.T.

= 27.16 - 26.94

= 0.22

ความผันแปรระหว่างแม่พันธุ์ $= \sum\left(Y_{ij\bullet}\right)^2 / s - C.T.$

(among dams)

$$= \frac{(2.63)^2 + (2.64)^2 + \dots\dots\dots + (2.53)^2}{2} - C.T.$$

= 27.14 - 26.94

= 0.20

ความผันแปรระหว่างพันธุ์ $= \sum\left(Y_{i\bullet\bullet}\right)^2 / rs - C.T.$

(among breeds)

$$\frac{(7.92\ \ )^2 + (6.61\ \ )^2 + (7.49\ \ )^2}{3 \times 2} - C.T.$$

=

= 27.09 - 26.94

= 0.15

ความผันแปรระหว่างแม่พันธุ์ในพันธุ์เดียวกัน (among dams within breeds)

= ความผันแปรระหว่างแม่พันธุ์ - ความผันแปรระหว่างพันธุ์

= 0.20 - 0.15

= 0.05

ความคลาดเคลื่อนที่เกิดจากการสุ่มหน่วยย่อย (sampling error)

= total SS - ความผันแปรระหว่างแม่พันธุ์

= 0.22 - 0.20

= 0.02

ซึ่งความคลาดเคลื่อนที่เกิดจากการสุ่มหน่วยย่อย (sampling error) นั้นจะรวมอยู่ในความผันแปรระหว่างแม่พันธุ์ในพันธุ์เดียวกัน เพราะความผันแปรระหว่างแม่พันธุ์ จะรวมเอาความผันแปรระหว่างตัวลูกสุกรที่สุ่มมาเข้าไว้ด้วย เช่นเดียวกัน ความผันแปรระหว่างพันธุ์ ก็จะรวมเอาความผันแปรระหว่างตัวลูกสุกรที่สุ่มมาไว้ด้วย เพราะความผันแปรระหว่างพันธุ์ จะรวมเอาความผันแปรระหว่างแม่พันธุ์เดียวกันเข้าไว้ด้วย เพราะความแตกต่างระหว่างพันธุ์เราวัดจากความแตกต่างระหว่างแม่พันธุ์

สมมุติฐานการทดสอบ

1) H_0: $\sigma_A^2 = 0$ หรือไม่มีความแตกต่างระหว่างพันธุ์

.

$$F = \frac{MST}{MSE}$$

2) H_0: $\sigma_B^2 = 0$ หรือไม่มีความแตกต่างระหว่างแม่พันธุ์

$$F = \frac{MSE}{\text{Sampling error}}$$

การวิเคราะห์วาเรียนซ์

Sources	*df*	SS	MS	F-ratio	EMS
พันธุ์	(3-1)=2	0.15	0.075	9.375*	$\sigma^2 + s\sigma_B^2 + rs\sigma_A^2$
แม่พันธุ์ในพันธุ์เดียวกัน	3(3-1)=6	0.05	0.008	4.0*	$\sigma^2 + s\sigma_B^2$
ระหว่างลูก	3x3(2-1)=9	0.02	0.002		σ^2

EMS หมายถึง ค่า mean square ในทางทฤษฎี หรือเรียกว่า expected mean square เป็นผลรวมของค่าวาเรียนซ์ หรือความผันแปรของปัจจัยต่าง ๆ ในงานทดลอง

สรุปผล: การเจริญเติบโตของลูกสุกรจะมีความแตกต่างกันระหว่างพันธุ์อย่างมีนัยสำคัญ และยังมีความแตกต่างระหว่างแม่พันธุ์เดียวกันอย่างมีนัยสำคัญ

การทดสอบว่าวาเรียนซ์มีค่าเท่ากัน

การทดสอบว่าวาเรียนซ์มีค่าเท่ากัน หรือ เรียกว่า test of homogeneity of variances เป็นการทดสอบว่าข้อมูลที่ได้มาเป็นไปตามข้อกำหนดที่ตั้งไว้ว่าวาเรียนซ์ของแต่ละทรีทเมนต์ต้องเท่ากัน และเท่ากับวาเรียนซ์ของประชากร การทดสอบวาเรียนซ์สองค่าเท่ากัน อาจทำโดยใช้ F-test ดังได้กล่าวมาแล้วในการเปรียบเทียบประชากร 2 กลุ่ม แต่ในกรณีที่มีวาเรียนซ์มากกว่า 2 ค่าขึ้นไป การทดสอบไม่สามารถจะใช้ F-test ได้ แต่จะใช้ค่าไคร์สแควร์ (χ^2) เป็นค่าทดสอบ วิธีการนี้เรียกว่า Bartlett's test

สมมุติฐานสำหรับการทดสอบ ได้แก่

$$H_0: \sigma_1^2 = \sigma_2^2 = \ldots\ldots\ldots\ldots = \sigma_t^2$$

$$H_A: \sigma_1^2 \neq \sigma_2^2 \neq \ldots\ldots\ldots\ldots \neq \sigma_t^2$$

ค่าสถิติที่ใช้สำหรับทดสอบคือ

$$\chi^2 = 2.3026\left\{\left[\sum_{i=1}^{t}(n_i - 1)\right]\log \overline{S}^2 - \left[\sum_{i=1}^{t}(n_i - 1)\right]\log S_i^2\right\}$$

ค่า df = t-1

โดยมี corrected factor (C) = $1 + \frac{1}{3(k-1)}\left\{\sum \frac{1}{n_i - 1} - \frac{1}{\sum(n_i - 1)}\right\}$

k คือ จำนวนประชากร หรือทรีทเมนต์

$$\therefore \text{ จะได้ } \chi^2\text{corrected} = \frac{\chi^2}{C}$$

$$\overline{S}^2 = \text{ค่า pooled variance} = \frac{\sum(n_i-1)S_i^2}{\sum(n_i-1)}$$

แต่เนื่องจากค่า corrected factor มีค่ามากกว่า 1 เสมอ ดังนั้น ถ้าค่า χ^2 ไม่มีนัยสำคัญ ก็ไม่จำเป็นที่จะต้องคำนวณ χ^2corrected

ตัวอย่างที่ 4.5 จากตัวอย่างการศึกษาผลของอาหาร 3 สูตร ที่มีต่อน้ำหนักไข่ (กรัม) ของไก่ ค่าเฉลี่ยน้ำหนักไข่ของไก่เป็นดังนี้

อาหารสูตร		
A	**B**	**C**
51	**58**	**59**
43	**60**	**57**

49	60	55
	53	54
	55	
รวม 143	286	225
เฉลี่ย 47.31	57.2	56.25
S^2 17.31	9.67	4.93

ต้องการทดสอบว่าวาเรียนซ์ของน้ำหนักไข่ของไก่ที่ได้รับอาหารทั้ง 3 สูตรแตกต่างกันหรือไม่

อาหาร	จำนวนซ้ำ	*df*	S^2	log S^2	(n-1)log S^2	1/n-1
A	3	2	17.31	1.2383	2.4766	0.5
B	5	4	9.67	0.9854	3.9416	0.25
C	4	3	4.93	0.6984	2.0784	0.33
รวม	12	9			8.4966	1.08

$$\overline{S}^2 = \frac{\sum(n_i-1)S_i^2}{\sum(n_i-1)}$$

$$= \frac{2(17.31)+4(9.67)+3(4.93)}{9} = 88.09$$

log $\overline{S}^2$ = log 88.09 = 1.9449

$$\chi^2 = 2.3026\ [9(1.9449) - 8.4966]$$

$$= 2.3026\ (17.5041 - 8.4966)$$

$$= 2.3026 \times 9.0075$$

$$= 20.74^{**}$$

จากตาราง χ^2 ที่ $df = 2$ ความน่าจะเป็น 5 เปอร์เซนต์ มีค่าเท่ากับ 5.99 และที่ความน่าจะเป็น 1 เปอร์เซนต์ มีค่าเท่ากับ 9.21

$$C = 1 + \frac{1}{3(k-1)}\left\{\sum \frac{1}{n_i - 1} - \frac{1}{\sum(n_i - 1)}\right\}$$

$$= 1 + \frac{1}{3(2)}\left(1.08 - \frac{1}{9}\right)$$

$$= 1 + 0.17(1.08 - 0.11)$$

$$= 1 + 0.165$$

$$= 1.165$$

$$\chi^2 \text{ corrected} = \frac{20.74}{1.165} = 17.80^{**}$$

สรุปผล: มีความแตกต่างระหว่างวาเรียนซ์ของน้ำหนักไข่ของไก่ที่ได้รับอาหารสูตรต่างๆ กัน อย่างมีนัยสำคัญยิ่ง

การแปลงข้อมูล

ก่อนการวิเคราะห์วาเรียนซ์เราจะต้องทดสอบว่าข้อมูลเป็นไปตามข้อกำหนดหรือไม่ ถ้าข้อมูลไม่เป็นไปตามข้อกำหนด เราก็จะวิเคราะห์วาเรียนซ์ไม่ได้ จำเป็นที่จะต้องแปลงข้อมูลเสียก่อน แต่โดยทั่วไปงานทดลองมักจะเป็นไปตามข้อกำหนด ยกเว้นบางประเภทเท่านั้นที่ไม่เป็นไปตามข้อกำหนดจะต้องทดสอบก่อน ดังเช่น ข้อมูลที่เป็นแบบไบโนเมียล (binomial data) เช่น ตายกับไม่ตาย ข้อมูลที่เป็นลักษณะจำนวนนับเป็นสัดส่วนหรือเป็นเปอร์เซนต์ ข้อมูลเกี่ยวกับการให้คะแนน เหล่านี้มักจะมีการกระจายไม่เป็นแบบปกติ เราอาจจะใช้วิธีอื่นที่ไม่คำนึงถึงการกระจายของข้อมูล สำหรับการเปรียบเทียบ เช่น non-parametric methods แต่มีประสิทธิภาพต่ำจึงไม่นิยมใช้ การแปลงข้อมูลมีวัตถุประสงค์ก็เพื่อ

1) ให้ข้อมูลที่แปลงแล้วมีการกระจายเป็นแบบปกติ หรือใกล้เคียงกับแบบปกติ

2) ค่าเฉลี่ยและวาเรียนซ์ของข้อมูลที่แปลงแล้ว เป็นอิสระต่อกัน ทำให้วาเรียนซ์ของทุกทรีทเมนต์ เท่ากันทั้งหมด ในบางครั้งการแปลงข้อมูล ก็อาจจะไม่มีผลตามวัตถุประสงค์ ผู้วิจัยจะต้องหาวิธีวิเคราะห์แบบอื่นๆ ที่ใช้สำหรับการตรวจสอบนัยสำคัญของการทดลอง การแปลงข้อมูลมีอยู่หลายวิธีด้วยกัน แต่ละวิธีก็จะเหมาะสมกับข้อมูลที่แตกต่างกันไปดังนี้

การแปลงเป็นค่าลอการิทึม (Logarithmic transformation, log Y)

การแปลงข้อมูลเป็นค่าลอการิทึม เป็นวิธีการแปลงข้อมูลที่มีประสิทธิภาพมากที่สุด ใช้ในกรณีที่ความเบี่ยงเบนมาตรฐานเป็นสัดส่วนกับ

ค่าเฉลี่ยของทรีทเมนต์ การแปลงเป็นลอการิทึมจะทำให้ทุกทรีทเมนต์มีวาเรียนซ์เท่ากัน และทำให้อิทธิพลต่าง ๆ ที่มีผลแบบคูณในข้อมูลเดิมเปลี่ยนเป็นมีผลแบบบวกในค่าของล๊อก ข้อมูลที่มีค่าเป็นลบไม่สามารถที่จะแปลงได้โดยวิธีนี้ แต่ถ้ามีค่าศูนย์อยู่ในข้อมูล การแปลงเป็นล๊อก จะทำให้ได้ค่าเป็น ลบอินฟินิตี้ (-α) การแปลงข้อมูลดังกล่าวนี้ จะทำได้โดยการบวกหนึ่งให้กับค่าสังเกตทุกค่าก่อนการแปลง นั่นก็คือ การแปลงแบบ log(Y+1) ลอการิทึมฐาน 10 จะง่ายที่สุดสำหรับการแปลงข้อมูล

การแปลงโดยใช้รากที่สอง (The square root transformation, $\sqrt{Y}$)

การแปลงข้อมูลโดยใช้รากที่สอง จะใช้กับข้อมูลที่มีค่าสังเกตมีค่าต่ำมาก เช่น การนับจำนวนเหตุการณ์ผิดปกติ การนับจำนวนแมลงในช่วงเวลาหนึ่ง เป็นต้น ข้อมูลแบบนี้มักจะมีการกระจายเป็นแบบ พัวซอง (Poisson distribution) ซึ่งจะมีค่าวาเรียนซ์เท่ากับค่าเฉลี่ย ข้อมูลแบบนี้สามารถทำให้เป็นแบบปกติได้ และให้วาเรียนซ์เป็นอิสระกับค่าเฉลี่ยได้ โดยการแปลงให้อยู่ในรูปของรากที่สอง การใช้ $\sqrt{Y+\frac{1}{2}}$ จะได้ผลดี โดยเฉพาะถ้ามีค่านับได้ต่ำกว่า 10

การแปลงข้อมูลโดยใช้รากที่สอง จะมีผลต่อรูปร่างลักษณะของการกระจาย ของความคลาดเคลื่อน และลักษณะอิทธิพลร่วมแบบบวก ถ้าบล๊อกและทรีทเมนต์มีอิทธิพลร่วมกันแบบบวกสะสมในข้อมูลเดิม เมื่อแปลงข้อมูลแล้วอิทธิพลเหล่านี้จะยังคงลักษณะบวกสะสมอยู่

การแปลงเป็นอาร์คไซน์ (arcsine or angular transformation, arcsine $\sqrt{Y}$)

การแปลงข้อมูลโดยใช้อาร์คไซน์ จะใช้กับข้อมูลที่มีการนับเป็นเปอร์เซนต์ หรือ สัดส่วนของตัวอย่างทั้งหมด เป็นข้อมูลที่มีการกระจายแบบไบโนเมียล (binomial distribution) มากกว่าที่จะเป็นแบบปกติ การแปลงข้อมูลชนิดนี้เป็นการหามุมวัดเป็นองศา ซึ่งมีค่าไซน์ (sine) เป็นรากที่สองของสัดส่วน เขียนเป็นรูปคณิตศาสตร์ได้ดังนี้ arcsine $\sqrt{Y}$ หรือ sine$^{-1}\sqrt{Y}$ ในตารางผนวกที่ 7 แสดงค่ามุมวัดเป็นองศาที่แปลงมาจากข้อมูลเดิมที่เป็นเปอร์เซนต์