แผนการทดลองขั้นพื้นฐาน: แผนการทดลองแบบสุ่มในบล๊อก

แผนการทดลองแบบสุ่มในบล๊อกอย่างสมบูรณ์

ในบทก่อนได้กล่าวถึง แผนการทดลองแบบสุ่มตลอด ซึ่งจะใช้ในกรณีที่สิ่งทดลองของเรา มีความสม่ำเสมอหรือคล้ายคลึงกัน ความผันแปรที่เกิดขึ้นก็จะมีเพียงอย่างเดียว คือ ความผันแปรของทรีทเมนต์ เมื่อสิ่งทดลองไม่มีความสม่ำเสมอ การใช้แผนการทดลองแบบสุ่มตลอดก็จะถูกต้องน้อยลง ในที่นี้จะขอกล่าวถึง แผนการทดลองแบบสุ่มในบล๊อก (Randomized Blocks Design, RBD) ซึ่งจะใช้เมื่อสิ่งทดลองมีความผันแปรหรือแตกต่างกันอยู่แล้วหนึ่งทางก่อนที่จะถูกกระทำโดยทรีทเมนต์ ดังเช่น การศึกษาการเจริญเติบโตของสุกร สุกรแต่ละพันธุ์ย่อมมีอัตราการเจริญเติบโตที่แตกต่างกัน หากในการศึกษาใช้สุกรหลายพันธุ์ด้วยกัน เราจึงควรจัดแยกกลุ่มสิ่งทดลองที่สม่ำเสมอ หรือสุกรที่เป็นพันธุ์เดียวกันอยู่กลุ่มเดียวกัน นั่นก็คือ การจัดสิ่งทดลองออกเป็นบล๊อก (block) โดยให้สิ่งทดลองภายในบล๊อกเดียวกันมีความสม่ำเสมอกัน แต่มีความแตกต่างกันระหว่างกลุ่มหรือบล๊อก

วิธีการสุ่ม

ตามที่กล่าวมาแล้ว ในการทดลองแบบสุ่มในบล๊อกจะมีสิ่งทดลองแตกต่างกันอยู่หนึ่งทาง ซึ่งเราจะต้องจัดสิ่งทดลองออกเป็นกลุ่ม หรือบล๊อก สิ่งทดลองในบล๊อกเดียวกันมีความสม่ำเสมอกัน แต่จะมีความแตกต่างกันระหว่างบล๊อก ดังนั้นเพื่อให้ความยุติธรรมกับทุกทรีทเมนต์ที่ใช้ทดลอง จึงต้องจัดให้แต่ละทรีทเมนต์ได้อยู่ในสภาพเดียวกัน อาจกล่าวได้ว่า ทุกทรีทเมนต์จะต้องปรากฏในแต่ละบล๊อกเท่ากัน การสุ่มของแผนการทดลองนี้จึงเป็น การสุ่มทรีทเมนต์ให้กับหน่วยทดลอง ที่อยู่ภายในบล๊อกแต่ละบล๊อก ซึ่งจำนวนหน่วยทดลองในแต่ละบล๊อกจะเท่ากับจำนวนทรีทเมนต์ เราจึงเรียกแผนการทดลองแบบนี้ว่าแผนการทดลองแบบสุ่มในบล๊อกอย่างสมบูรณ์ (Randomized Completely Blocks Design) เช่น การทดลองเปรียบเทียบอาหารสุกร 4 สูตร สุกรที่ใช้ทดลองมีอยู่ด้วยกัน 3 พันธุ์ จะเห็นว่าในการทดลองนี้สิ่งทดลองก็คือตัวสุกร ซึ่งจะมีความแตกต่างกันในเรื่องของพันธุ์ การทดลองจึงต้องจัดแบ่งกลุ่มของสิ่งทดลองออกเป็น 3 กลุ่ม หรือบล๊อกตามพันธุ์ สิ่งที่สำคัญที่สุดก็คือ สุกรที่ใช้ในการทดลองนี้ควรจะมีอายุ น้ำหนักเมื่อเริ่มต้นเท่ากัน และควรเป็นเพศเดียวกัน เพื่อให้ทุกทรีทเมนต์ปรากฏอยู่ในแต่ละบล๊อก การทดลองนี้จึงต้องใช้สุกร 4 ตัวต่อบล๊อก รวมเป็นใช้สุกรในการทดลอง 12 ตัว กรณีนี้ถือว่าสุกร 1 ตัว เป็นหนึ่งหน่วยทดลอง ทำการสุ่มทรีทเมนต์หรืออาหารทุกสูตรให้กับแต่ละบล๊อก ลักษณะการทดลองก็อาจเป็นดังนี้

สุกรพันธุ์ A (บล๊อกที่ 1)	สุกรพันธุ์ B (บล๊อกที่ 2)	สุกรพันธุ์ C (บล๊อกที่ 3)
สูตร 2	สูตร 3	สูตร 2
สูตร 1	สูตร 4	สูตร 3
สูตร 3	สูตร 2	สูตร 4
สูตร 4	สูตร 1	สูตร 1

ขนาดของบล๊อก (block size) เท่ากับ 4 เท่ากับจำนวนทรีทเมนต์

แบบหุ่น (Mathematical model)

แบบหุ่นของแผนการทดลองแบบสุ่มในบล๊อกอย่างสมบูรณ์ ได้แก่

$$Y_{ij} = \mu + \beta_i + \tau_j + \varepsilon_{ij}$$

μ = ค่าเฉลี่ยร่วม

β_i = อิทธิพลของบล๊อกที่ i; i = 1, 2, ..., b

τ_j = อิทธิพลของทรีทเมนต์ที่ j; j = 1, 2, ..., t

ε_{ij} = ความคลาดเคลื่อนของการทดลอง; $\varepsilon_{ij} \sim nid(0, \sigma^2)$

การวิเคราะห์

บล๊อกที่	T_1	T_2	●	●	●	●	T_t	ผลรวมของบล๊อก
1	Y_{11}	Y_{12}	●	●	●	●	Y_{1t}	$Y_{1.}$
2	Y_{21}	Y_{22}	●	●	●	●	Y_{2t}	$Y_{2.}$
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
●	●	●	●	●	●	●	●	●
b	Y_{b1}	Y_{b2}	●	●	●	●	Y_{bt}	$Y_{b.}$
ผลรวมทรีทเมนต์	$Y_{.1}$	$Y_{.2}$	●	●	●	●	$Y_{.t}$	$Y_{..}$

block total = $Y_{i.}$

treatment total = $Y_{.j}$

grand total = $Y_{..}$

correction term (C.T.) = $\left(\sum Y_{ij}\right)^2 / n = \left(Y_{..}\right)^2 / bt$

$$\text{total SS} = \sum_{i,j}\left(Y_{ij} - \overline{Y}_{..}\right)^2$$

$$= \sum_{i,j} Y_{ij}^2 - C.T.$$

$$\text{block SS} = t\sum_{i}\left(\overline{Y}_{i\bullet}-\overline{Y}_{\bullet\bullet}\right)^2$$

$$= \sum_{i}\frac{Y_{i\bullet}^2}{t} - C.T.$$

$$\text{treatment SS} = b\sum_{j}\left(\overline{Y}_{\bullet j}-\overline{Y}_{\bullet\bullet}\right)^2$$

$$= \sum_{j}\frac{Y_{\bullet j}^2}{b} - C.T.$$

error SS = total SS - block SS - treatment SS

ตารางการวิเคราะห์วาเรียนซ์

source	*df*	SS	MS	F-ratio
block	b-1	SSB	MSB = SSB / b-1	MSB / MSE
treatment	t-1	SST	MST = SST / t-1	MST / MSE
error	(b-1)(t-1)	SSE	MSE = SSE /(b-1)(t-1)	
total	tb-1			

การเปรียบเทียบความแตกต่าง ระหว่างค่าเฉลี่ยของทรีทเมนต์ ทำได้เช่นเดียวกับแผนการทดลองแบบสุ่มตลอด

ตัวอย่างที่ 5.1 การเปรียบเทียบอายุเมื่อเริ่มเป็นสัดครั้งแรกของแม่สุกรสาว ซึ่งได้พบปะกับสุกรเพศผู้เมื่ออายุ 135, 165, 190 วัน กับพวกที่ไม่ได้พบปะสุกรเพศผู้เลย แม่สุกรที่ใช้ทดลองเป็นสุกรพันธุ์เดียวกัน โดยสุ่มสุกรสาวจากสุกร 5 ครอกๆ ละ 4 ตัว อายุเมื่อเริ่มเป็นสัดครั้งแรกของสุกรสาวเป็นดังนี้

บล๊อกที่ (ครอกที่)	พบปะสุกรเพศผู้เมื่ออายุ				ผลรวมบล๊อก
	135	165	190	ไม่ได้พบปะ	
1	174	183	257	212	826
2	164	207	198	201	770
3	169	297	257	277	1000
4	160	223	206	239	828
5	169	206	242	258	875
ผลรวมทรีทเมนต์	836	1116	1160	1187	4299

$$C.T. = \frac{(4299)^2}{20} = 924070.05$$

$$\text{total SS} = 174^2 + 164^2 + \ldots\ldots\ldots\ldots\ldots + 258^2 - C.T.$$

$$= 954667 - 924070.05 = 30596.95$$

$$\text{block SS} = \frac{826^2 + 770^2 + 1000^2 + 828^2 + 875^2}{4} - \text{C.T.}$$

$$= 931596.25 - 924070.05 = 7526.2$$

$$\text{treatment SS} = \frac{836^2 + 1116^2 + 1160^2 + 1187^2}{5} - \text{C.T.}$$

$$= 939784.2 - 924070.05 = 15714.15$$

error SS = 30596.95 - 7526.2 - 15714.15 = 7356.6

ตารางการวิเคราะห์วาเรียนซ์

source	*df*	SS	MS	F-ratio
block	5-1 = 4	7526.2	1881.55	3.069^{ns}
treatment	4-1 = 3	15714.15	5238.05	8.544^{**}
error	(5-1)(4-1) = 12	7356.6	613.05	
total	4×5 - 1 = 19	30596.95		

$$F_{0.05(4, 12)} = 3.26$$

$$F_{0.01(3, 12)} = 5.95$$

สรุปผล: สุกรที่ได้พบปะกับสุกรเพศผู้ครั้งแรกเมื่ออายุต่างกัน จะมีอายุเมื่อเริ่มเป็นสัดครั้งแรก แตกต่างกันอย่างมีนัยสำคัญยิ่ง แต่สุกรที่ต่างครอกกันจะมีอายุเมื่อเริ่มเป็นสัดครั้งแรก แตกต่างกันอย่างไม่มีนัยสำคัญ

สำหรับการเปรียบเทียบค่าเฉลี่ยทรีทเมนต์ กระทำเช่นเดียวกับแผนการทดลองแบบสุ่มตลอด ซึ่งในที่นี้จะแสดงเฉพาะวิธี Duncan's New Multiple Range Test โดยใช้ตัวอย่างที่ 5.1 ซึ่งมีค่าเฉลี่ยของทรีทเมนต์ ดังนี้

ค่าเฉลี่ยอายุเมื่อเป็นสัดครั้งแรก (วัน) ของสุกรสาวที่พบปะสุกรเพศผู้เมื่ออายุ

ต่าง ๆ กัน

135 วัน (T_1)	165 วัน (T_2)	190 วัน (T_3)	ไม่ได้พบปะเลย (T_4)
167.2	223.2	232	237.4

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

$$= \sqrt{\frac{613.05}{5}}$$

$$= 11.07$$

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

จากตาราง ค่า $SSR_{0.01}$ ที่ $df = 12$ จะมีค่าเฉลี่ยของทรีทเมนต์ที่อยู่ในการทดสอบตั้งแต่ 2-4

p =	2	3	4
$SSR_{0.01}$	4.32	4.55	4.68
$LSR_{0.01}$	4.32×11.07 = 47.82	4.55×11.07 = 50.37	4.68×11.07 = 51.81

จัดเรียงค่าเฉลี่ยทรีทเมนต์จากน้อยไปมาก

(T_1)	(T_2)	(T_3)	(T_4)
167.2	223.2	232	237.4

การเปรียบเทียบก็คือ

$$\bar{Y}_4 - \bar{Y}_1 = 237.4-167.2 = 70.2^{**} \rangle 51.81$$
$$\bar{Y}_4 - \bar{Y}_2 = 237.4-223.2 = 14.2^{ns} \langle 50.37$$
$$\bar{Y}_4 - \bar{Y}_3 = 237.4-232 = 5.4^{ns} \langle 47.82$$
$$\bar{Y}_3 - \bar{Y}_1 = 232-167.2 = 64.8^{**} \rangle 50.37$$
$$\bar{Y}_3 - \bar{Y}_2 = 232-223.2 = 8.8^{ns} \langle 47.82$$
$$\bar{Y}_2 - \bar{Y}_1 = 223.2-167.2 = 56.0^{**} \rangle 47.82$$

สรุปผล:

$\bar{Y}_1$	$\bar{Y}_2$	$\bar{Y}_3$	$\bar{Y}_4$
167.2 ก	223.2 ข	232 ข	237.4 ข

นั่นคือ สุกรสาวที่ได้พบปะกับสุกรเพศผู้เมื่ออายุ 135 วัน จะมีอายุเป็นสัดครั้งแรกน้อยกว่าพวกที่ได้พบปะกับสุกรเพศผู้เมื่ออายุ 165, 190 วัน และพวกที่ไม่ได้พบปะกับสุกรเพศผู้เลย อย่างมีนัยสำคัญยิ่ง แต่สุกรสาวพวกที่พบปะสุกรเพศผู้เมื่ออายุ 165 และ 190 วัน กับพวกที่ไม่ได้พบปะสุกรเพศผู้จะมีอายุเมื่อเป็นสัดครั้งแรกแตกต่างกันอย่างไม่มีนัยสำคัญ ($p>0.05$)

การเปรียบเทียบประสิทธิภาพของแผนการทดลอง

ในการทดลองแบบสุ่มในบล๊อก เราจะทดสอบทรีทเมนต์เพียงอย่างเดียว ไม่จำเป็นต้องทดสอบบล๊อกก็ได้เพราะเราแบ่งบล๊อก เพื่อลดความผันแปรอันเนื่องมาจากบล๊อก โดยแบ่งความผันแปรเนื่องจากบล๊อกออกจากความผันแปรของความคลาดเคลื่อน ซึ่งแผนแบบสุ่มในบล๊อกจะมีความแม่นยำ ถ้าเราสามารถแยกความผันแปรอันเนื่องมาจากบล๊อกออกจากความผันแปรของความคลาดเคลื่อนได้มากที่สุดเท่าที่จะทำได้ โดยยอมเสียความเป็นอิสระ (*df*) ของความคลาดเคลื่อนให้ลดลง นับว่าเป็นผลที่คุ้มค่า ถ้ามีอิทธิพลของบล๊อกอย่างแท้จริง โดยเมื่อทดสอบบล๊อกแล้วมีความแตกต่างอย่างมีนัยสำคัญ การใช้แผนแบบสุ่มในบล๊อกก็นับว่าถูกต้องแล้ว แต่ถ้าทดสอบแล้ว มีความแตกต่างระหว่างบล๊อกอย่างไม่มีนัยสำคัญ นั่นก็คือ ไม่มีอิทธิพลของบล๊อกเลย การใช้แผนแบบสุ่มในบล๊อกก็จะทำให้เราเสีย *df* ของความคลาดเคลื่อนไปเปล่าๆ และการทดสอบยากขึ้นโดยใช่เหตุ

ในการเปรียบเทียบว่า แผนการทดลองใดจะมีประสิทธิภาพดีกว่ากัน เราจะใช้ค่าประสิทธิภาพสัมพันธ์ (relative efficiency) เป็นค่าเปรียบเทียบ ซึ่งประสิทธิภาพจะขึ้นอยู่กับความคลาดเคลื่อน ถ้าความคลาดเคลื่อนน้อย ประสิทธิภาพจะสูง แต่ถ้าความคลาดเคลื่อนสูง ประสิทธิภาพจะต่ำลง

$$efficiency = \frac{1}{error}$$

สำหรับแผนการทดลองแบบสุ่มในบล๊อก เราจะเปรียบเทียบประสิทธิภาพ กับ แผนแบบสุ่มตลอด โดยการหาค่าประสิทธิภาพสัมพันธ์ระหว่างแผนแบบสุ่มในบล๊อก กับ แผนแบบสุ่มตลอด (rela-tive efficiency RBD to CRD)

$$R.E. = \frac{efficiency\ RBD}{efficiency\ CRD}$$

$$= \frac{\frac{1}{MSE \text{ ของ } RBD}}{\frac{1}{MSE \text{ ของ } CRD}}$$

$$= \frac{MSE \text{ ของ } CRD}{MSE \text{ ของ } RBD}$$

ตารางการวิเคราะห์วาเรียนซ์ของ RBD

source	*df*	MS
block	n_b	E_b
treatment	n_t	E_t
error	n_e	E_e

n_b = *df* ของ block E_b = mean square ของ block

n_t = *df* ของ treatment E_t = mean square ของ treatment

n_e = *df* ของ error E_e = mean square ของ error

ในการเปรียบเทียบแผนแบบสุ่มในบล๊อก กับแผนแบบสุ่มตลอดโดยใช้ค่าประสิทธิภาพสัมพันธ์ (R.E.) เราไม่สามารถทราบค่า mean square ของความคลาดเคลื่อนของแผนการทดลองแบบสุ่มตลอด ดังนั้น เราจึงต้องคำนวนประมาณค่า MSE ของแผน CRD โดยไม่ให้มีอคติด้วยสูตร

$$E_e\,(CRD) \;=\; \frac{n_b E_b + (n_t + n_e) E_e(RBD)}{n_b + n_t + n_e}$$

การใช้สูตร ประมาณค่า $E_e(CRD)$ นั้น ถ้า n_e มีค่าน้อยกว่า 20 แล้ว ค่า R.E. ที่ได้ จะยังไม่ถูกต้องนัก ทั้งนี้เนื่องมาจาก *df* ลดลงในการประมาณค่า MSE โดยวิธีการของแผนแบบสุ่มในบล๊อก เราจึงจำเป็นที่จะต้องปรับด้วยค่า precision factor ซึ่งมีค่าดังนี้

$$precision\ factor = \frac{(n_1 + 1)(n_2 + 3)}{(n_2 + 1)(n_1 + 3)}$$

ดังนั้น $$R.E.\ adjusted = \frac{E_e(CRD)}{E_e(RBD)} \times \frac{(n_1 + 1)(n_2 + 3)}{(n_2 + 1)(n_1 + 3)} \times 100$$

n_1 คือ *df* ของ error ใน RBD คือ n_e

n_1 คือ *df* ของ error ใน CRD คือ $n_b + n_e$

จากตัวอย่างที่ 5.1 ตารางการวิเคราะห์วาเรียนซ์ คือ

source	df	SS	MS	F-value
block	4	7536.2	1881.55	3.069^{ns}
treatment	3	15714.15	5238.05	8.544^{**}
error	12	7356.6	613.05	

ผลการทดสอบปรากฏว่า ความแตกต่างระหว่างบล๊อกไม่มีนัยสำคัญ การแบ่งบล๊อกทำให้ *df* ของความคลาดเคลื่อนลดลง ความแม่นยำก็จะลดลงด้วย ซึ่งถ้าเราใช้แผนแบบสุ่มตลอด ค่า *df* ของความคลาดเคลื่อนก็จะเพิ่มขึ้นเป็น 16 แต่เราก็ยังไม่รู้ว่าใช้แผนแบบสุ่มตลอดจะดีกว่าแผนแบบสุ่มในบล๊อกหรือไม่ เราจึงต้องเปรียบเทียบโดยใช้ค่า **R.E. (RBD to CRD)**

$$E_e(CRD) = \frac{n_b E_b + (n_t + n_e) E_e(RBD)}{n_b + n_t + n_e}$$

$$= \frac{4(1881.55) + (3+12)(613.05)}{4+3+12}$$

$$= \frac{7526.2 + 9195.75}{19}$$

$$= 880.10$$

$$\text{R.E. (RBD to CRD)} = \frac{E_e(CRD)}{E_e(RBD)} \times \frac{(n_1+1)(n_2+3)}{(n_2+1)(n_1+3)} \times 100$$

$$= \frac{880.10}{613.05} \times \frac{(12+1)(16+3)}{(16+1)(12+3)} \times 100$$

$$= \frac{880.10}{613.05} \times \frac{247}{255} \times 100$$

$$= 139\ \%$$

นั่นก็หมายความว่า ถ้าเราต้องการให้แผนการทดลองแบบสุ่มตลอดและแผนการทดลองแบบสุ่มในบล๊อก มีประสิทธิภาพเท่ากัน เราก็ต้องใช้จำนวนซ้ำ 139 ซ้ำสำหรับแผนแบบสุ่มตลอด และแผนแบบสุ่มในบล๊อกจะต้องใช้ 100 ซ้ำ

หรือ ถ้าเราใช้จำนวนซ้ำเท่ากัน แผนแบบสุ่มในบล๊อกจะให้ความแม่นยำ 139 เปอร์เซ็นต์ แต่แผนแบบสุ่มตลอดให้ความแม่นยำ 100 เปอร์เซ็นต์

ค่าสังเกตสูญหาย

ในการทดลองบางครั้ง เราอาจจะไม่สามารถบันทึก ค่าสังเกตจากหน่วยทดลองบางหน่วย ทั้งนี้อาจจะเป็นเพราะสัตว์ทดลองป่วยหรือตายไป ค่าสังเกตที่เราไม่สามารถบันทึกได้นี้ เรียกว่า ค่าสังเกตสูญหาย (missing value) ในกรณีที่การทดลองใช้แผนแบบสุ่มตลอด เราก็สามารถวิเคราะห์ได้ถึงแม้จะมีค่าสังเกตสูญหาย โดยใช้การวิเคราะห์ของแผนการทดลองแบบสุ่มตลอดมีจำนวนซ้ำไม่เท่ากัน ดังได้กล่าวมาแล้วในบทก่อน แต่แผนการทดลองแบบสุ่มในบล๊อกนั้น จำเป็นที่จะต้องมีการประมาณค่าสังเกตที่หายไปก่อน เพื่อช่วยให้สะดวกในการวิเคราะห์ข้อมูลที่เหลือ ซึ่งการประมาณค่าสังเกตสูญหายแบ่งออกเป็น 2 กรณี ดังนี้

กรณีที่มีค่าสังเกตสูญหายไปหนึ่งค่า

การคำนวนประมาณค่าสังเกตที่หายไปเพียงค่าเดียว กระทำได้จาก

$$M = \frac{rB + tT - G}{(r-1)(t-1)}$$

เมื่อ M คือ ค่าสังเกตที่หายไป

B คิอ ผลรวมของบล๊อกที่มีค่าสังเกตสูญหาย

T คือ ผลรวมของทรีทเมนต์ที่มีค่าสังเกตสูญหาย

r คือ จำนวนบล๊อก

t คือ จำนวนทรีทเมนต์

G คือ ผลรวมของข้อมูลทั้งหมด

ค่าสังเกตที่ประมาณได้นั้น ไม่ใช่เป็นค่าสังเกตที่จะให้ข้อเท็จจริงเพิ่มเติม จะเป็นแต่เพียงหุ่นเชิด (dummy) ที่เราประมาณมาเพื่อให้การวิเคราะห์ดำเนินต่อไปได้ ซึ่งค่า sum of square ของทรีทเมนต์ จะเป็นค่าอคติที่สูงกว่าที่ควรจะเป็น (bias upward) จึงควรที่จะปรับเสียก่อนด้วยการลบด้วยค่า

$$\text{bias} = \frac{[B-(t-1)M]^2}{t(t-1)}$$

ค่า *df* ของความคลาดเคลื่อน กับ *df* ของ total จะต้องลดลงไป 1 ค่า เนื่องจากค่าสังเกตที่หายไปไม่ใช่ค่าสังเกตที่แท้จริง

ค่าความคลาดเคลื่อนมาตรฐานของความแตกต่างระหว่างค่าเฉลี่ยของทรีทเมนต์ที่มีค่าสูญหาย กับทรีทเมนต์อื่น ๆ เท่ากับ

$$S_{\bar{d}} = \sqrt{S^2\left[\frac{2}{r} + \frac{t}{r(r-1)(t-1)}\right]}$$

จากตัวอย่างที่ 5.1 ถ้าค่าสังเกตของทรีทเมนต์ที่ 1 ในบล๊อกที่ 3 หายไป การประมาณค่าสังเกตสูญหายกระทำดังนี้

บล๊อก	สุกรสาวพบปะสุกรเพศผู้เมื่ออายุ (วัน)				ผลรวมของบล๊อก
	135	165	190	ไม่ได้พบปะ	
1	174	183	257	212	826
2	164	207	198	201	770
3	-	297	257	277	831
4	160	223	206	239	828
5	169	206	242	258	875
ผลรวมทรีทเมนต์	667	1116	1160	1187	4130

$$M = \frac{rB + tT - G}{(r-1)(t-1)}$$

$$= \frac{5(831) + 4(667) - 4130}{(5-1)(4-1)}$$

$$= \frac{2693}{12}$$

$$= 224.42$$

นำค่าสังเกตที่ประมาณขึ้น (224.42) ใส่ลงในตารางข้อมูลแล้วทำการวิเคราะห์ต่อไป

$$C.T. = \frac{(4130+224.42)^2}{20}$$
$$= 948048.65$$

$$\text{total SS} = 174^2+164^2+224.42^2+.....+258^2-C.T.$$
$$= 976470.33-948048.65$$
$$= 28421.68$$

$$\text{block SS} = \frac{826^2+770^2+(831+224.42)^2+828^2+875^2}{4} - C.T.$$
$$= 960074.07 - 948048.65$$
$$= 12025.42$$

$$\text{treatment SS} = \frac{(667+224.42)^2+1116^2+1160^2+1187^2}{5} - C.T.$$
$$= 958930.92-948048.65$$
$$= 10882.72$$

error SS = 28421.68 - 12025.42 - 10882.27

= 5513.99

$$\text{bias} = \frac{[B-(t-1)M]^2}{t(t-1)}$$
$$= \frac{[831-(4-1)224.42]^2}{4(4-1)}$$
$$= \frac{24881.907}{12}$$
$$= 2073.49$$

∴ **treatment SS (adjusted) = 10882.27 - 2073.49 = 8808.78**

ตารางการวิเคราะห์วาเรียนซ์

source	*df*	SS	MS	F-ratio
block	4	12025.42		
treatment	3	8808.78	2936.26	5.86*
error	11	5513.99	501.27	
total	19-1 = 18	28421.68		

สรุปผล: สุกรสาวที่ได้พบปะสุกรเพศผู้เมื่ออายุต่างกันจะเริ่มเป็นสัดครั้งแรกเมื่ออายุต่างกัน(P<0.05)

กรณีที่มีค่าสังเกตสูญหายมากกว่า 1 ค่า

ในขั้นตอนแรกนี้ให้คำนวนค่าสังเกตที่สูญหายไปทุกค่าอย่างหยาบๆ โดยใช้สูตร

$$\frac{\overline{Y}_{i\bullet} + \overline{Y}_{\bullet j}}{2}$$

$\overline{Y}_{i\bullet}$ คือ ค่าเฉลี่ยของทรีทเมนต์ที่มีค่าสังเกตสูญหายไป

$\overline{Y}_{\bullet j}$ คือ ค่าเฉลี่ยของบล๊อกที่มีค่าสังเกตสูญหายไป

โดยยกเว้นไว้เพียงค่าเดียวให้คำนวนโดยใช้สูตร

$$M = \frac{rB+tT-G}{(r-1)(t-1)}$$

แล้วนำกลับไปใส่ไว้ในข้อมูลเพื่อคำนวนตัวอื่นอีก โดยใช้สูตร $M = \frac{rB+tT-G}{(r-1)(t-1)}$ ทำซ้ำกลับไปกลับมาอีกจนกว่า ทุกค่าที่ประมาณขึ้นมีค่าคงที่ จึงทำการวิเคราะห์วาเรียนซ์ต่อไปได้

ในกรณีนี้ ก็จะเกิดอคติที่สูงเกินกว่าที่ควรจะเป็น ค่าที่ใช้ปรับค่า Sum of square ของทรีทเมนต์จะมีหลายค่าด้วยกัน เราจะต้องนำมารวมกันแล้วหักออกจากค่า Sum of square ของทรีทเมนต์

$$\text{bias} = \frac{[B-(t-1)\,M_1]^2}{t(t-1)} + \frac{[B-(t-1)\,M_2]^2}{t(t-1)} + \ldots\ldots\ldots\ldots$$

ค่า *df* ของ error และ *df* ของ total ก็จะต้องลดลงตามจำนวนค่าสังเกตที่สูญหาย แต่ถ้ามีค่าสังเกตสูญหายมากจนทำให้ค่า *df* ของ error ต่ำมากๆ เช่น เป็น 0 หรือ 1 เราก็จะทดสอบความแตกต่างไม่ได้เลย

ตัวอย่างที่ 5.2 จากการทดลองเปรียบเทียบใช้อาหาร 6 สูตร เลี้ยงสุกรขุน โดยใช้สุกรทดลองทั้งหมด 4 พันธุ์ด้วยกัน เป็นสุกรที่มีอายุและน้ำหนักตัวใกล้เคียงกัน น้ำหนักตัวเพิ่มของสุกรเมื่อสิ้นสุดการทดลองมีดังต่อไปนี้

อาหาร สูตร	สุกรพันธุ์ A	B	C	D	ผลรวมของทรีทเมนต์
1	4.4	5.9	6.0	4.1	20.4
2	(M_1)	1.9	4.9	7.1	13.9
3	4.4	4.0	4.5	3.1	16.0
4	6.8	6.6	(M_2)	6.4	19.8
5	6.3	4.9	5.9	7.1	24.2
6	6.4	6.3	7.7	6.7	27.1
ผลรวมบล๊อก	28.3	29.6	29.0	34.5	121.4

วิธีทำ

คำนวนประมาณค่าที่หายไปอย่างหยาบ ๆ

$$M_1 = \frac{\bar{Y}_{2\bullet} + \bar{Y}_{\bullet 1}}{2}$$

$$= \frac{13.9/3 + 28.3/5}{2}$$

$$= 5.15$$

ประมาณค่า M_2 โดยใช้

$$M_2 = \frac{rB + tT - G}{(r-1)(t-1)}$$

$$= \frac{4(29.0) + 6(19.8) - (121.4 + 5.15)}{(4-1)(6-1)}$$

$$= \frac{108.25}{15}$$

$$= 7.22$$

ประมาณค่า M_1 ในรอบแรก

$$M_1 = \frac{rB + tT - G}{(r-1)(t-1)}$$

$$= \frac{4(28.3) + 6(13.9) - (121.4 + 7.22)}{(4-1)(6-1)}$$

$$= 4.53$$

ประมาณค่า M_2 รอบที่ 2

$$M_2 = \frac{4(29.0) + 6(19.8) - (121.4 + 4.53)}{(4-1)(6-1)}$$

$$= 7.26$$

ประมาณค่า M_1 รอบที่ 2

$$M_1 = \frac{4(28.3) + 6(13.9) - (121.4 + 7.26)}{(4-1)(6-1)}$$

$$= 4.53$$

ประมาณค่า M_2 รอบที่ 3

$$M_2 = \frac{4(29.0) + 6(19.8) - (121.4 + 4.53)}{(4-1)(6-1)}$$

$$= 7.26$$

เมื่อค่า M_1 และ M_2 ที่ประมาณเริ่มคงที่ เราก็แทนค่า M_1 และ M_2 ลงในตารางข้อมูล แล้ว ทำการวิเคราะห์วาเรียนซ์ต่อไปตามปกติ

$$C.T. = \frac{(121.4+4.53+7.26)^2}{24}$$

$$= 739.15$$

$$\text{total SS} = 4.4^2+4.53^2+\ldots\ldots\ldots\ldots+6.7^2-C.T.$$

$$= 787.97-739.15 = 48.82$$

$$\text{block SS} = \frac{32.83^2+29.6^2+36.26^2+34.5^2}{6}-C.T.$$

$$= 743.17-739.15 = 4.02$$

$$\text{treatment SS} = \frac{20.4^2+18.43^2+16.0^2+27.06^2+24.2^2+27.1^2}{4} - C.T.$$

$$= 766.03-739.15 = 26.88$$

error SS = total SS-block SS-treatment SS

= 48.82-4.02-26.88 = 17.92

$$\text{bias} = \frac{[B-(t-1)M_1]^2}{t(t-1)} + \frac{[B-(t-1)M_2]^2}{t(t-1)}$$

$$= \frac{[28.3-(6-1)4.53]^2}{6(6-1)} + \frac{[29.0-(6-1)7.26]^2}{6(6-1)}$$

$$= 1.06+1.78 = 2.84$$

treatment SS (adjusted) = 26.88-2.84 = 24.04

ตารางการวิเคราะห์วาเรียนซ์

source	*df*	SS	MS	F-ratio
block	3	4.02	1.34	
treatment	5	24.04	4.81	3.48*
error	13	17.92	1.38	
total	21	48.82		

$F_{0.05(5, 13)} = 3.02$

สรุปผล: **อาหารทดลองมีผลต่อการเพิ่มน้ำหนักตัวของสุกรอย่างมีนัยสำคัญทางสถิติ(P<0.05)**

การเปรียบเทียบค่าเฉลี่ยของทรีทเมนต์

ค่าความเบี่ยงเบนมาตรฐานของ ผลต่างระหว่างค่าเฉลี่ยทรีทเมนต์ที่มีค่าสังเกตสูญหายคำนวนได้จาก

$$S_{\bar{d}} = \sqrt{S^2\left(\frac{1}{r_i}+\frac{1}{r_j}\right)}$$

โดย r_i และ r_j เป็น effective number of replicates (ENR) ซึ่งคำนวนหาได้ด้วยกฎต่อไปนี้

1) จำนวนซ้ำของทรีทเมนต์ที่กำลังพิจารณาอยู่ให้นับเป็น 1 ถ้าในบล๊อกนั้นมีค่าสังเกตของทั้งสองทรีทเมนต์

2) จำนวนซ้ำของทรีทเมนต์ที่กำลังพิจารณาอยู่ให้นับเป็น $\frac{t-2}{t-1}$ ถ้าในบล๊อกนั้นอีกทรีทเมนต์หนึ่งเป็นค่าสังเกตสูญหาย

3) จำนวนซ้ำของทรีทเมนต์ที่กำลังพิจารณาอยู่ให้นับเป็น 0 ถ้าทรีทเมนต์นี้เป็นค่าสังเกตสูญหายในบล๊อกนั้น

4) จำนวน ENR จะเท่ากับจำนวนซ้ำที่คำนวนได้ในแต่ละบล๊อกรวมกัน

จากตัวอย่างที่ 5.2 ถ้าต้องการเปรียบเทียบความแตกต่างระหว่างค่าเฉลี่ยของทรีทเมนต์ที่ 2 กับ ทรีทเมนต์ที่ 4 คำนวนค่า ENR ได้ดังนี้

$$\text{ENR ของทรีทเมนต์ที่ 2} = 0+1+\frac{(6-2)}{(6-1)}+1$$
$$= 0+1+\frac{4}{5}+1$$
$$= 2.8$$

$$\text{ENR ของทรีทเมนต์ที่ 4} = \frac{4}{5}+1+0+1$$
$$= 2.8$$

$$\therefore\ S_{\bar{Y}_4-\bar{Y}_2} = \sqrt{S^2\left(\frac{1}{r_4}+\frac{1}{r_2}\right)}$$
$$= \sqrt{1.378\left(\frac{1}{2.8}+\frac{1}{2.8}\right)} = 0.99$$

ถ้าต้องการเปรียบเทียบค่าเฉลี่ยของทรีทเมนต์ที่ 2 กับ ทรีทเมนต์ที่ 5 ค่า ENR จะเป็น

ENR ของทรีทเมนต์ที่ 2 = 0+1+1+1 = 3.0

$$\text{ENR ของทรีทเมนต์ที่ 5} = \frac{4}{5}+1+1+1 = 3.8$$

$$\therefore\ S_{\bar{Y}_5-\bar{Y}_2} = \sqrt{S^2\left(\frac{1}{r_5}+\frac{1}{r_2}\right)}$$

$$= \sqrt{1.378\left(\frac{1}{3.8}+\frac{1}{3.0}\right)} = 0.91$$

สำหรับการเปรียบเทียบค่าเฉลี่ยของทรีทเมนต์ที่เป็นคู่ปกติไม่มีค่าสังเกตุสูญหาย

$$S_{\bar{d}} = \sqrt{\frac{2\,MSE}{r}} = \sqrt{\frac{2(1.378)}{4}} = 0.83$$

ค่าเฉลี่ยของทรีทเมนต์เรียงลำดับจากน้อยไปหามาก ได้แก่

T_3	T_2	T_1	T_5	T_4	T_6
4.0	**4.63**	**5.1**	**6.05**	**6.6**	**6.78**

ค่า $t\frac{0.05}{2}(13) = 2.160$

$T_6 - T_3 = 6.78 - 4.0 = 2.78^{*} > 0.83 \times 2.16 = 1.79$

$T_6 - T_2 = 6.78 - 4.63 = 2.15^{*} > 0.91 \times 2.16 = 1.97$

$T_6 - T_1 = 6.78 - 5.1 = 1.68^{ns} < 0.83 \times 2.16 = 1.79$

$T_6 - T_5 = 6.78 - 6.05 = 0.73^{ns} < 0.83 \times 2.16 = 1.79$

$T_6 - T_4 = 6.78 - 6.6 = 0.18^{ns} < 0.91 \times 2.16 = 1.97$

$T_4 - T_3 = 6.6 - 4.0 = 2.6^{*} > 0.91 \times 2.16 = 1.97$

$T_4 - T_2 = 6.6 - 4.63 = 1.97^{ns} < 0.99 \times 2.16 = 2.14$

$T_4 - T_1 = 6.6 - 5.1 = 1.5^{ns} < 0.91 \times 2.16 = 1.97$

$T_4 - T_5 = 6.6 - 6.05 = 0.55^{ns} < 0.91 \times 2.16 = 1.97$

$T_5 - T_3 = 6.05 - 4.0 = 2.05^{*} > 0.83 \times 2.16 = 1.79$

$T_5 - T_2 = 6.05 - 4.63 = 1.42^{ns} < 0.91 \times 2.16 = 1.97$

$T_5 - T_1 = 6.05 - 5.1 = 0.95^{ns} < 0.83 \times 2.16 = 1.79$

$T_1 - T_3 = 5.1 - 4.0 = 1.1^{ns} < 0.83 \times 2.16 = 1.79$

$T_1 - T_2 = 5.1 - 4.63 = 0.47^{ns} < 0.91 \times 2.16 = 1.97$

$T_2 - T_3 = 4.63 - 4.0 = 0.63^{ns} < 0.91 \times 2.16 = 1.97$

สรุปผล:

T_1	T_2	T_3	T_4	T_5	T_6
$5.1^{กขค}$	$4.63^{กข}$	$4.0^{ก}$	$6.6^{ขค}$	$6.05^{ขค}$	$6.78^{ค}$

นั่นคือ น้ำหนักตัวเพิ่มของสุกรที่ได้รับอาหารสูตรที่ 3 จะแตกต่างกับ พวกที่ได้รับอาหารสูตรที่ 2 และ1 อย่างไม่มีนัยสำคัญทางสถิติ แต่จะแตกต่างกับพวกที่ได้รับอาหารสูตรที่ 5, 4 และ 6 อย่างมีนัยสำคัญทางสถิติ ขณะเดียวกันน้ำหนักตัวเพิ่มของสุกรที่ได้รับอาหารสูตรที่ 1, 2, 4 และ 5 ก็แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ และน้ำหนักตัวเพิ่มของสุกรที่ได้รับอาหารสูตร 1, 4, 5 และ6 ก็แตกต่างกันอย่างไม่มีนัยสำคัญทางสถิติ แต่สุกรที่ได้รับอาหารสูตร 2 จะมีน้ำหนักเพิ่มแตกต่างกับพวกที่ได้รับอาหารสูตร 6 อย่างมีนัยสำคัญทางสถิติ

ข้อดีและข้อเสียของแผนการทดลองแบบสุ่มในบล๊อก

ข้อดี

เป็นแผนการทดลองที่ให้ความแม่นยำสูงกว่าแผนการทดลองแบบสุ่มตลอด เมื่อสิ่งทดลองที่ใช้มีความผันแปรตามที่เราแบ่งอย่างแท้จริง

2) การวิเคราะห์ไม่ค่อยยุ่งยากมากนัก ถึงแม้จะเกิดมีค่าสังเกตสูญหาย เมื่อเปรียบเทียบกับการวิเคราะห์ของแผนการทดลองอื่นๆ ที่จะกล่าวถึงต่อไป

ข้อเสีย

1) เมื่อเปรียบเทียบกับแผนการทดลองแบบสุ่มตลอดแล้ว อาจจะมีประสิทธิภาพต่ำกว่าแผนแบบสุ่มตลอด โดยเฉพาะอย่างยิ่งเมื่อบล๊อกไม่มีอิทธิพลจริง ทั้งนี้เพราะ ค่า *df* ของความคลาดเคลื่อนจะต่ำกว่า เพราะถูกแบ่งไปให้กับบล๊อก

2) การวิเคราะห์จะยุ่งยากกว่าแผนการทดลองแบบสุ่มตลอด ถ้ามีค่าสังเกตสูญหายหลายค่า