実験データー解析概論

— 統計学に基づく「よりよい推論」のために —

三中 信宏

MINAKA Nobuhiro

独立行政法人 農業環境技術研究所 生態系計測研究領域 上席研究員［生物統計学］

東京大学大学院 農学生命科学研究科 生物・環境工学専攻 教授［生態系計測学］

東京農業大学大学院 農学研究科 客員教授［応用昆虫学］

mailto:minaka@affrc.go.jp（メール）

http://twitter.com/leeswijzer/（ツイッター）

http://cse.niaes.affrc.go.jp/minaka/（ウェブサイト）

http://d.hatena.ne.jp/leeswijzer/（ブログ）

2013 年度・東京農業大学生物応用化学科（毎週金曜第 2 限）

三中信宏〈実験データー解析概論〉：2013 年 4 月〜 2013 年 9 月

東京農業大学世田谷キャンパス 431 番教室

1 要因乱塊法（Box 2）

［実験］播種密度条件（6 水準 a ～ f）を変えたイネ収量実験を 4 反復の乱塊法で実施した（IRRI）.

ブロック I

e	b
a	d
f	c

ブロック II

c	a
b	f
e	d

ブロック III

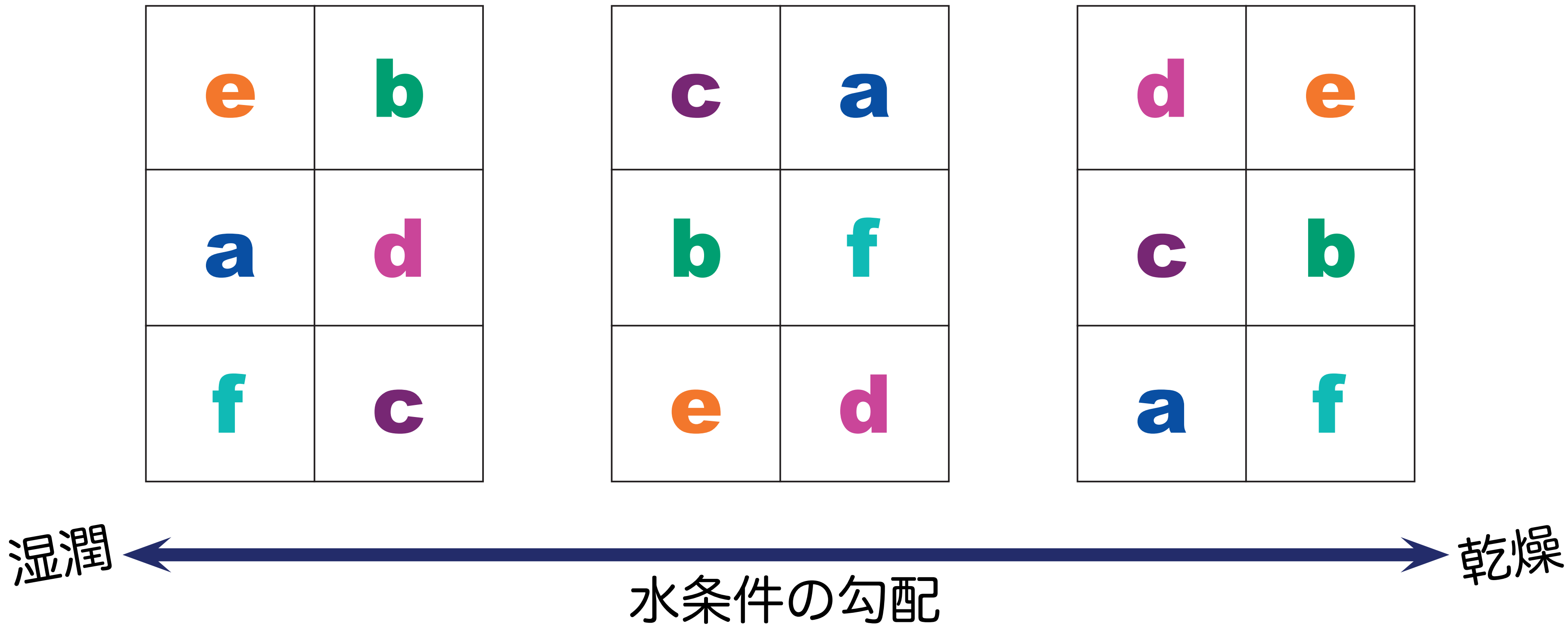
d	e
c	b
a	f

ブロック IV

f	d
e	c
b	a

1 要因乱塊法（Box 2）

なぜブロックをつくるのか？



背景要因の勾配に関する事前情報があるとき，その勾配に直行する方向にブロックを切ることにより，背景要因の影響をブロック効果としてコントロールできる．

1 要因乱塊法（Box 2）

乱塊法の線形モデル

総平均

データ

$$x_{ij} = \mu + \boxed{\alpha_i} + \boxed{\rho_j} + \boxed{\epsilon_{ij}} \qquad (\epsilon_{ij} \sim N(0, \sigma^2))$$

第 i 水準
第 j ブロック

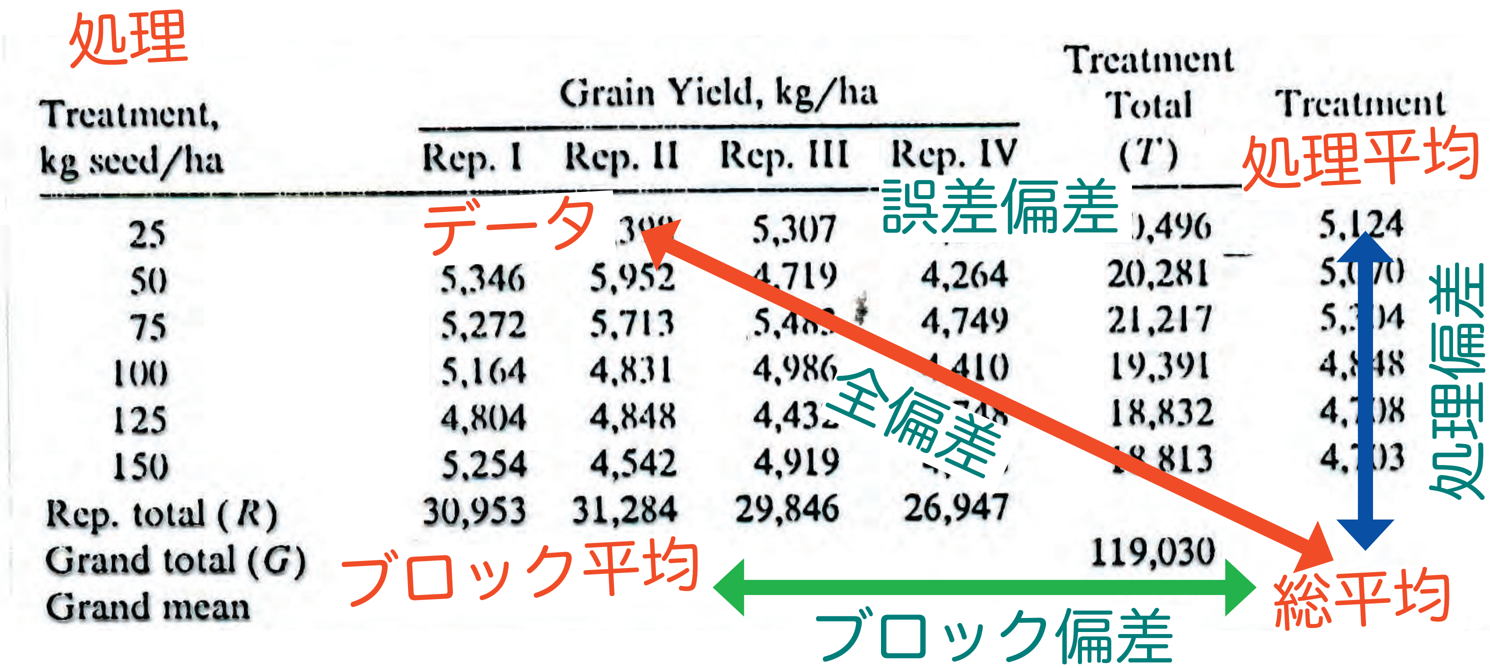
誤差の正規性（仮定）

処理効果　　ブロック効果　　誤差項

「ブロック効果」を明示的にモデルに組み込むことにより，背景要因がデータに与える影響を解明できる．

1 要因乱塊法（Box 2）

データの構造



Treatment, kg seed/ha	Grain Yield, kg/ha				Treatment Total (T)	Treatment
	Rep. I	Rep. II	Rep. III	Rep. IV		
25	[illegible]	[illegible]	5,307	[illegible]	[illegible]	5,124
50	5,346	5,952	4,719	4,264	20,281	[illegible]
75	5,272	5,713	[illegible]	4,749	21,217	[illegible]
100	5,164	4,831	4,986	[illegible]	19,391	[illegible]
125	4,804	4,848	[illegible]	[illegible]	18,832	[illegible]
150	5,254	4,542	4,919	[illegible]	[illegible]	[illegible]
Rep. total (R)	30,953	31,284	29,846	26,947		
Grand total (G)					119,030	
Grand mean						

「全偏差＝処理偏差＋ブロック偏差＋誤差偏差」と分割する

1 要因乱塊法（Box 2）

偏差の分割ならびにその後の手順

$$\begin{aligned}
\underbrace{x_{ij} - \overline{x}_{..}}_{\text{全偏差}} &= \underbrace{(\overline{x}_{i.} - \overline{x}_{..})}_{\text{処理偏差}} \\
&\quad + \underbrace{(\overline{x}_{.j} - \overline{x}_{..})}_{\text{ブロック偏差}} \\
&\quad + \underbrace{\{(x_{ij} - \overline{x}_{..}) - (\overline{x}_{i.} - \overline{x}_{..}) - (\overline{x}_{.j} - \overline{x}_{..})\}}_{\text{誤差偏差}} \\
&= \underbrace{(\overline{x}_{i.} - \overline{x}_{..})}_{\text{処理偏差}} + \underbrace{(\overline{x}_{.j} - \overline{x}_{..})}_{\text{ブロック偏差}} + \underbrace{(x_{ij} - \overline{x}_{i.} - \overline{x}_{.j} + \overline{x}_{..})}_{\text{誤差偏差}}
\end{aligned}$$

全偏差の分割に対応して．全平方和を各要因に対応する平方和に分割できる．さらに，平方和の自由度もまた並行的に分割される．

1 要因乱塊法（Box 2）

分散分析表（1）

Source of Variation	Degree of Freedom	Sum of Squares	Mean Square	Computed F^b	Tabular F 5%	Tabular F 1%
Replication	3	1,944,361	648,120			
Treatment	5	1,198,331	239,666	2.17^{ns}	2.90	4.56
Error	15	1,658,376	110,558			
Total	23	4,801,068				

2 要因乱塊法（Box 3）

［実験］窒素施肥量 N（5 水準）とイネ品種 V（3 水準）
でのイネ収量実験を 4 反復の乱塊法で実施した（IRRI）.

Nitrogen Level, kg/ha	Factorial Treatment Combination		
	6966 (V_1)	P1215936 (V_2)	Milfor 6(2) (V_3)
0(N_0)	N_0V_1	N_0V_2	N_0V_3
40(N_1)	N_1V_1	N_1V_2	N_1V_3
70(N_2)	N_2V_1	N_2V_2	N_2V_3
100(N_3)	N_3V_1	N_3V_2	N_3V_3
130(N_4)	N_4V_1	N_4V_2	N_4V_3

2 要因乱塊法（Box 3）

実験区配置

窒素（N）と品種（V）のすべての水準の組み合わせを，各ブロックにおいて無作為配置した.

	V_3N_2	V_2N_1	V_1N_4	V_1N_1	V_2N_3
Rep. I	V_3N_0	V_1N_3	V_3N_4	V_1N_2	V_3N_3
	V_2N_4	V_3N_1	V_2N_0	V_1N_0	V_2N_2

	V_2N_3	V_3N_3	V_1N_1	V_2N_0	V_2N_1
Rep. II	V_1N_3	V_3N_2	V_1N_2	V_1N_4	V_2N_4
	V_1N_0	V_3N_4	V_2N_2	V_3N_1	V_3N_0

	V_1N_1	V_3N_0	V_1N_0	V_3N_1	V_1N_4
Rep. III	V_2N_2	V_1N_2	V_1N_3	V_2N_4	V_3N_4
	V_2N_0	V_3N_2	V_2N_1	V_2N_3	V_3N_3

	V_1N_2	V_2N_2	V_2N_4	V_1N_0	V_2N_0
Rep. IV	V_1N_3	V_3N_1	V_1N_4	V_1N_1	V_2N_3
	V_3N_0	V_2N_1	V_3N_2	V_3N_3	V_3N_4

2 要因乱塊法（Box 3）

得られたデータ

Nitrogen Level, kg/ha	Grain Yield, t/ha Rep. I	Rep. II	Rep. III	Rep. IV	Treatment Total (T)
			V_1		
N_0	3.852	2.606	3.144	2.894	12.496
N_1	4.788	4.936	4.562	4.608	18.894
N_2	4.576	4.454	4.884	3.924	17.838
N_3	6.034	5.276	5.906	5.652	22.868
N_4	5.874	5.916	5.984	5.518	23.292
			V_2		
N_0	2.846	3.794	4.108	3.444	14.192
N_1	4.956	5.128	4.150	4.990	19.224
N_2	5.928	5.698	5.810	4.308	21.744
N_3	5.664	5.362	6.458	5.474	22.958
N_4	5.458	5.546	5.786	5.932	22.722
			V_3		
N_0	4.192	3.754	3.738	3.428	15.112
N_1	5.250	4.582	4.896	4.286	19.014
N_2	5.822	4.848	5.678	4.932	21.280
N_3	5.888	5.524	6.042	4.756	22.210
N_4	5.864	6.264	6.056	5.362	23.546
Rep. total (R)	76.992	73.688	77.202	69.508	
Grand total (G)					297.390

2 要因乱塊法（Box 3）

線形モデル

データ　総平均　処理効果

$$\underbrace{x_{ijk}}_{\text{データ}} = \underbrace{\mu}_{\text{総平均}} + \overbrace{\{\boxed{\alpha_i} + \boxed{\beta_j} + \boxed{(\alpha\beta)_{ij}}\}}^{\text{処理効果}} + \underbrace{\boxed{\rho_k}}_{\text{ブロック効果}} + \underbrace{\boxed{\epsilon_{ijk}}}_{\text{誤差項}} \quad \underline{(\epsilon_{ijk} \sim N(0, \sigma^2))}$$

第 i 水準
第 j 水準
第 k ブロック

誤差の正規性（仮定）

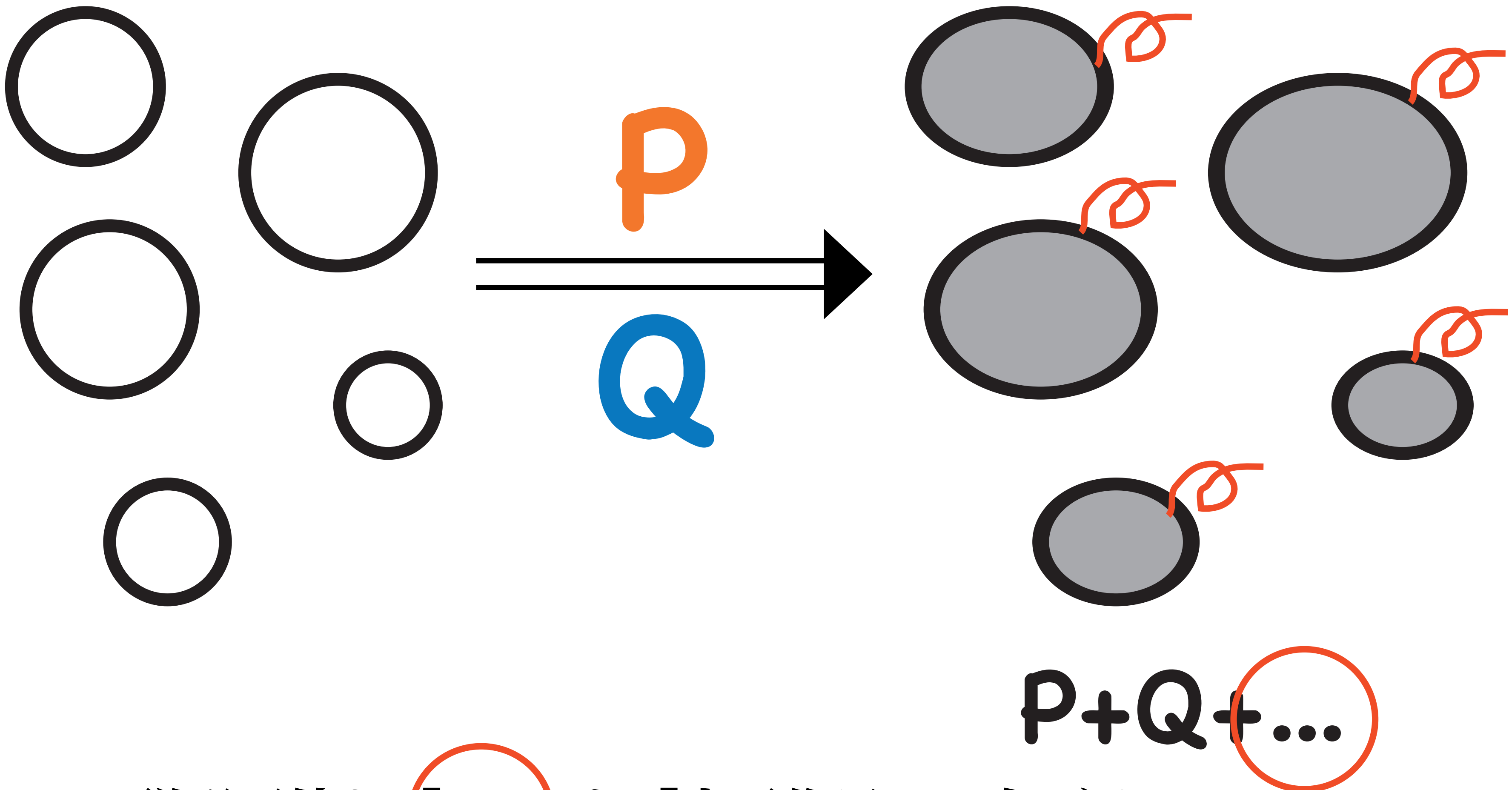
誤差項

ブロック効果

第 i 処理効果　第 j 処理効果　交互作用効果

複数の要因を含む実験計画では，要因間の「交互作用」をモデルに組み込む必要がある．

要因間の交互作用とは何か？



説明不能な「…」を「交互作用」と名づける.

実験区配置の「ネスト」と「クロス」

通常の乱塊法配置

要因A（水準 a_1 と a_2）と要因B（水準 b_1 ～ b_3）に関する3反復の乱塊法を実施するとき，通常は各ブロック内で無作為化配置する．

a_2b_1	a_1b_3
a_1b_2	a_2b_1
a_1b_3	a_2b_2

ブロックⅠ

a_1b_3	a_1b_2
a_2b_2	a_2b_1
a_1b_1	a_2b_3

ブロックⅡ

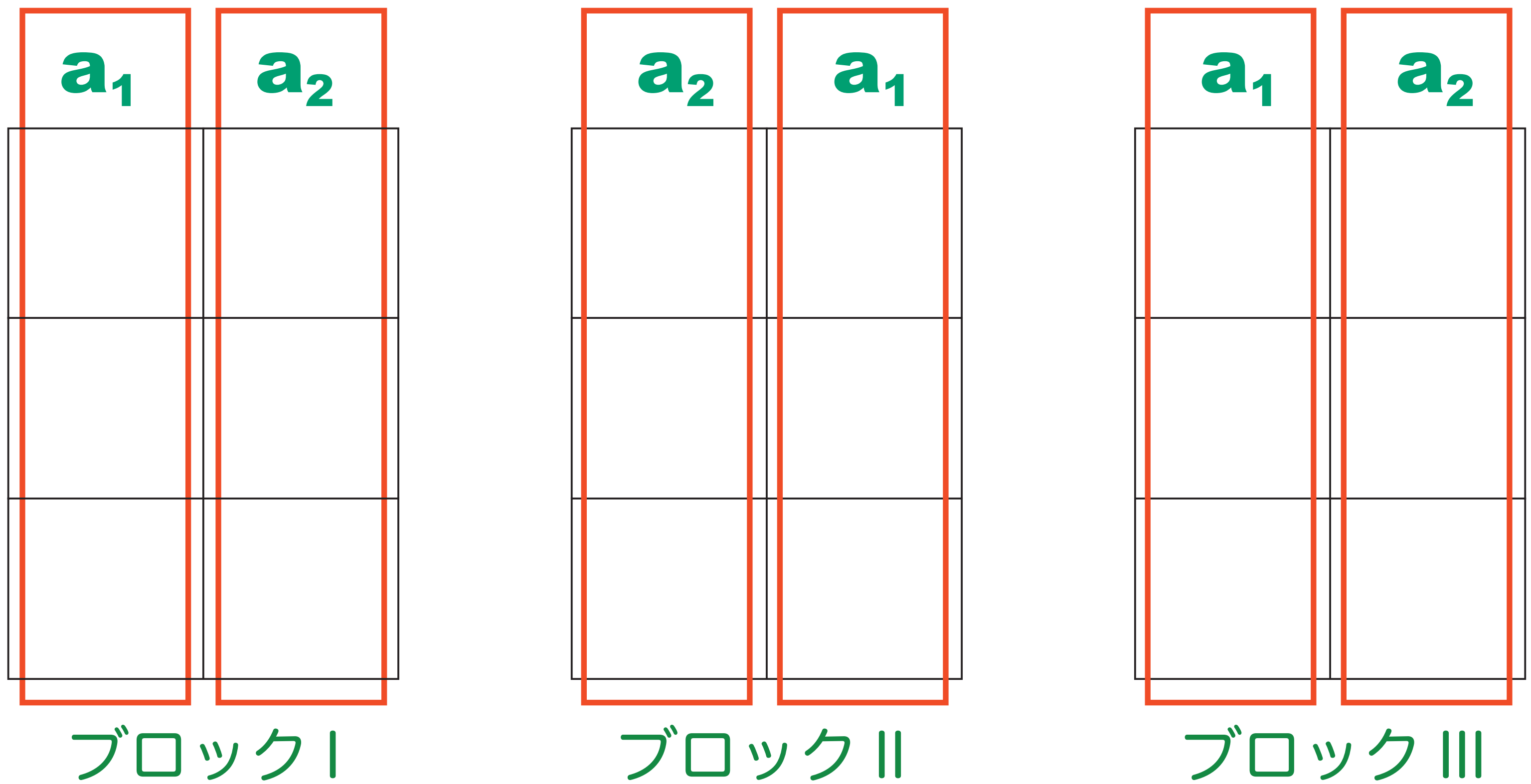
a_1b_2	a_2b_1
a_2b_3	a_1b_2
a_2b_1	a_1b_3

ブロックⅢ

実験区配置の「ネスト」と「クロス」

分割区法（split-plot）による配置

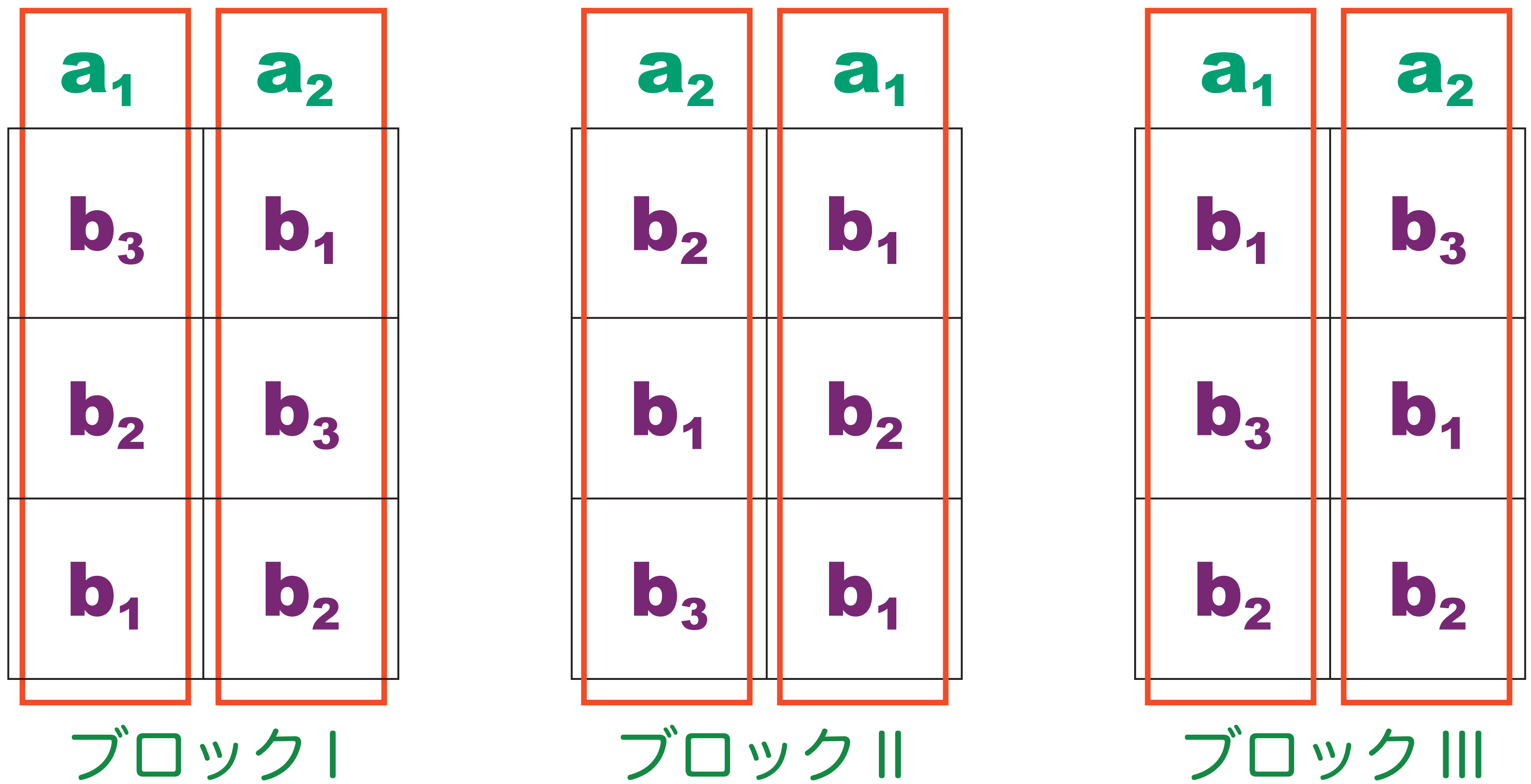
分割区法は，最初に無作為に割付ける一次要因（A）の分割区の中で，あとで無作為に割付ける二次要因 B をネストさせる．



実験区配置の「ネスト」と「クロス」

分割区法（split-plot）による配置

分割区法は，最初に無作為に割付ける一次要因（A）の分割区の中で，あとで無作為に割付ける二次要因 B をネストさせる．



実験区配置の「ネスト」と「クロス」

分割区法（split-plot）の線形モデル

総平均　ブロック効果

誤差項１　誤差項２

$$\text{データ}\quad x_{ijk} = \mu + \rho_k + \underbrace{\{\alpha_i + \epsilon^{(1)}_{ik}\}}_{\text{一次要因分析}} + \underbrace{\{\beta_j + (\alpha\beta)_{ij} + \epsilon^{(2)}_{ijk}\}}_{\text{二次要因分析}}$$

第 i 水準
第 j 水準
第 k ブロック

分割区法の線形モデルは，一次要因と二次要因（および交互作用）で別々の誤差項を指定し，それぞれの分散分析は対応する誤差項を基準にして実施される．