6.4 القيم المفقودة وطرق تقديرها

Missing Experimental Units and the procedure for its Estimate

في بعض تجارب القطاعات الكاملة العشوائية قد يحصل فقدان نتيجة وحدة تجريبية أو أكثر أو أن تكون قيم بعض الوحدات شاذة أو غير دقيقة فيقرر المجرب عدم استخدامها في التحليل وهناك كثير من هذه الحالات ولاسيما في التجارب الزراعية والحيوانية فقد تتلف بعض الوحدات الزراعية نتيجة العوارض الطبيعية كالأمطار والثلوج أو نتيجة أن الحشرات تقضي على محصول أحدى الوحدات (القطع) التجريبية . ولأجل أن تكون التجربة (المكرر أو القطاع) كاملة فلابد من القيام بتقدير القيمة (القيم) المفقودة.

هذا وأن أفضل قيمة مقدرة هي التي تجعل الفرق بين مجموع مربعات الخطأ التجريبي المستخرج بعد تقدير القيمة المفقودة والخطأ التجريبي في حالة عدم وجود قيمة مفقودة (قيمة حقيقية) أصغر ما يمكن ، ولقد وجد Allen and Wishart علاقة لتقدير قيمة واحدة مفقودة معتمدين على هذه الفكرة (تصغير مجموع مربعات الخطأ) ألا أن Yates طور هذه الفكرة ووجد أسلوب تتابعي (Iterative Procedure) لحساب عدة قيم مفقودة شرط أن تخصم من درجات حرية الخطأ درجة واحدة عند تقدير كل قيمة . وصيغة التقدير(16) إذا كانت y_{ij} مفقودة هي :

$$\hat{y}_{ij} = \frac{t\,y_{i.} + b\,y_{.j} - y_{..}}{(t-1)(b-1)} \quad \ldots (16)$$

هذا وأن عدد القطاعات $j=1,2,\ldots,b$ ، وعدد المعالجات $i=1,2,\ldots,t$.

القطاعات / اصناف القمح	1	2	3	4	$Y_{i.}$
A	42	40	38	40	160
B	80	85	89	90	344
C	55	60	58	50	223
D	40	50	45	42	177
E	67	73	66	70	276
$y_{.j}$	284	308	296	292	1180

وان:

$y_{.j}$:مجموع نتائج وحدات القطاع j الذي وقعت فيه نتيجة الوحدة مفقودة.

$y_{i.}$:مجموع نتائج نفس المعالجة i المفقودة والموجودة في بقية القطاعات.

$y_{..}$: المجموع العام لكل النتائج.

تطبيق (B)

لو فرضنا أن نتيجة المعالجة الثانية (B) ضمن القطاع الثاني بمعنى $y_{22}=85$ كانت قد فقدت , فيمكن ان نقوم بتقديرها عن طريق تطبيق الصيغة (16) حيث سيكون

$$y_{2.} = 259$$

$$y_{.2} = 223$$

$$y_{..} = 1095$$

$$\hat{y}_{22} = \frac{5(259) + 4(223) - 1095}{(5-1)(4-1)} = 91$$

فنضع هذه النتيجة المقدرة في موقعها ضمن البيانات ثم نقوم باجراء تحليل التباين وعند كتابة درجات الحرية فيتم تنقيص درجة حرية واحدة (لكل نتيجة مفقودة وتم تقديرها) من درجات حرية الخطأ وكذلك من الكلي بمعنى:

S.O.V	df	SS	MS	F
treatments	4			
blocks	3			
error	12-1=11			
Total	20-1-1=18			

وفي حالة تقدير أكثر من قيمة مفقودة فيتم تطبيق هذه الطريقة حيث يتم افتراض قيم أولية وقتية للقيم المفقودة ماعدا واحدة يتم تقديرها عن طريق الصيغة (16) السابقة ثم نهمل إحدى القيم الافتراضية الأولية الوقتية ونعيد تقديرها عن طريق تطبيق الصيغة (16) السابقة ثم نهمل القيمة الافتراضية الأخرى ونعيد تقديرها بتطبيق الصيغة (16) السابقة، وتستمر هذه العملية لتقدير جميع القيم بحسب الصيغة (16) وهكذا حتى تتم دورة كاملة ثم يعاد التقدير مرة ثانية وثالثة حتى تستقر النتائج للتقديرات (الفرق بين نتيجة وتقديرها السابق بحدود (0.001)) على هذا الأساس سميت بطريقة التعويض ألتتابعي .

الأخطاء المعيارية Standard Errors (SE)

حينما يحدث فقدان نتيجة أو نتائج مفقودة ولغرض إختبار معنوية الفرق بين متوسطي معالجتين بتطبيق إحدى طرق المقارنات المتعددة فإن صيغة احتساب الخطأ المعياري ستتأثر حيث أن احتساب الخطأ المعياري يمكن أن يكون وفقا لأحدى الحالات الثلاثة التالية :

B612

الحالة الأولى : لمقارنة أو لاختبار الفرق بين متوسطي معالجتين كلاهما

تحتوي قيمة مفقودة فإن صيغة الخطأ المعياري ستكون:

$$SE = \sqrt{\frac{2MSe}{r}}$$

حيث :

MSe: متوسط مربعات الخطأ في جدول تحليل التباين.

r : عدد التكرارات لكل معالجة (هنا يساوي عدد القطاعات).

الحالة الثانية : لمقارنة أو لاختبار الفرق بين متوسطي معالجتين أحداهما

تحتوي قيمة مفقودة والأخرى لا تتضمن قيمة مفقودة فإن

صيغة الخطأ المعياري ستكون :

$$SE = \sqrt{MSe[\frac{2}{r} + \frac{t}{r(r-1)(t-1)}]}$$

MSe: متوسط مربعات الخطأ في جدول تحليل التباين .

r : عدد التكرارات لكل معالجة (عدد القطاعات) .

t : عدد المعالجات .

الحالة الثالثة : لأغراض مقارنة أو اختبار الفرق بين متوسطي معالجتين كلاهما تحتوي قيمة مفقودة فإن صيغة الخطأ المعياري ستكون :

$$SE = \sqrt{MSe(\frac{1}{r_1} + \frac{1}{r_2})}$$

حيث أن :

r_1 :التكرار الفعال Effective Replication للمعالجة الأولى الداخلة في المقارنة أو الإختبار.

r_2 : التكرار الفعال للمعالجة الأخرى .

ويحسب التكرار الفعال كالآتي :

لنفرض أن لمخطط في أدناه هو لتجربة وفق RCBD وموزعة فيها المعالجات a , b , c , d , e كالآتي :

القطاعات			
1	2	3	4
a	--	a	d
b	a	d	b
d	b	e	e
c	e	--	c
e	d	c	a

من ملاحظتنا للجدول أعلاه نجد أن المعالجتين b , c مفقودتان ، وللمقارنة بين متوسطيهما نحسب التكرار الفعال لكل من المعالجتين كالآتي :

أولاً : لاستخراج التكرار الفعال للمعالجة b نأخذ القطاعات بالتسلسل ونتبع القواعد التالية :

1. إذا كانت المعالجتان b , c موجودتان في القطاع نعتبر التكرار مساو إلى 1 .
2. إذا كانت المعالجة b موجودة في القطاع والمعالجة c غير موجودة في القطاع نعتبر التكرار مساو إلى $\frac{t-2}{t-1}$.
3. إذا كانت المعالجة b غير موجودة فيكون التكرار مساو إلى الصفر .وعليه فالتكرار الفعال لـ b سيكون :

$$r_1 = r_b = 1 + \frac{5-2}{5-1} + 0 + 1 = 2.75$$

ثانيا : بالمثل لاستخراج التكرار الفعال للمعالجة c نتبع القواعد أعلاه وعليه فالتكرار الفعال لـ c سيكون:

$$r_2 = r_c = 1 + 0 + \frac{5-2}{5-1} + 1 = 2.75$$

7-4 الكفاية النسبية لتصميم القطاعات الكاملة العشوائية

Relative Efficiency (RE)

كما قلنا سابقا من المعروف إن دقة التجربة تزداد بزيادة عدد التكرارات أو القطاعات ولكن هنا من جهة أخرى ولأغراض ملاحظة ومتابعة دقة التجربة المنفذة في هذا التصميم نستعين أو نحسب ما يسمى

بكفاية هذا التصميم ولتقدير مدى فاعلية مواصفاته أي مدى فاعلية طريقة تجميع القطع التجريبية في مجموعات متجانسة (قطاعات) بهدف تصغير الخطأ التجريبي وبالتالي في زيادة دقة التجربة وخاصة إذا كان هناك أي شك في الأسس التي أتبعت في تجميع القطع (الوحدات) التجريبية، خصوصا إذا كانت طريقة التجميع فيها نوع من الصعوبة (شاقة)، إذ نقوم باحتساب أو تقدير ما يسمى بالكفاية النسبية لهذا التصميم مقارنة بالتصميم السابق الذي هو تصميم تام التعشية، فمن المعلومات المتوفرة في جدول تحليل التباين للتجربة المقامة بتصميم القطاعات الكاملة العشوائية نقوم بحساب تقدير متوسط مربعات الخطأ (تباين الخطأ) وكأن التجربة كانت قد أقيمت بتصميم تام التعشية وكالاتي :

$$\widehat{M}_{e(RCD)} = \frac{(df_b)(MS_b) + (df_t + df_{e(RCBD)})MS_{e(RCBD)}}{df_b + df_t + df_{e(RCBD)}} \quad \dots (17)$$

حيث أن :

$\widehat{M}_{e(RCD)}$: يمثل تقدير تباين الخطأ التجريبي على أساس أن التجربة مقامة بتصميم تام التعشية.

(df_b): درجة الحرية للقطاعات.

(MS_b): متوسط مربعات القطاعات.

$(df_b)(MS_b) = SS_b$: مجموع المربعات لبين القطاعات في جدول تحليل التباين.

df_t : درجة حرية المعالجات .

$df_{e(RCBD)}$: درجة حرية الخطأ في جدول تحليل التباين للتجربة المتفذة بالقطاعات.

$MS_{e(RCBD)}$: متوسط مربعات الخطأ (تباين الخطأ) لتجربة القطاعات الكاملة العشوائية (في جدول تحليل التباين) .

$df_b + df_t + df_{e(RCBD)} = df_{Total}$: درجة الحرية للكلي .

ولما كان عدد درجات الحرية للخطأ التجريبي يختلف باختلاف التصميم للتجربة لنفس المعالجات والتكرارات وعليه يجري تصحيح درجات الحرية أي يتم ترجيح كل من تباين الخطأ لتصميم القطاعات ولتصميم تام التعشية بدرجات الحرية المقابلة لهما وعليه فإن الكفاية النسبية تكون وفق الصيغة الآتية:

$$RE = \frac{(df_{e(RCBD)} + 1)(df_{e(RCD)} + 3)\,\hat{M}_{e(RCD)}}{(df_{e(RCBD)} + 3)(df_{e(RCD)} + 1)\,MS_{e(RCBD)}}(100\%) \quad \ldots (18)$$

حيث أن :

$$df_{e(RCD)} = df_{e(RCBD)} + df_b$$

إن التصحيح لاختلاف درجات الحرية والذي صيغته:

$$\frac{(df_{e(RCBD)} + 1)(df_{e(RCD)} + 3)}{(df_{e(RCBD)} + 3)(df_{e(RCD)} + 1)}$$

يكون له اثر قليل في التجارب الكبيرة (التي تضم عدد كبيرمن المعالجات ومن القطاعات) ولكنه يؤثر كثيرا في التجارب الصغيرة.

حينما تكون قيمة الكفاية النسبية أكبر من 100% فهذا يعني أن التجربة المنفذة فعلا وفق تصميم القطاعات الكاملة العشوائية لها كفاية نسبتها

$$RE\% - 100\% = \quad \%$$

و حينما تكون قيمة الكفاية النسبية أقل من 100% فهذا يعني أن التجربة لو كانت قد نفذت وفق تصميم تام التعشية لكان لها كفاية نسبتها

$$100\% - RE\% = \quad \%$$

تطبيق (9)

للتجربة أعلاه إحسب الكفاية النسبية لها بالمقارنة مع لو أنها كانت وفق RCD.

الحل : المعلومات في جدول تحليل التباين للتجربة هي:

S.O.V	df	SS	MS
blocks	3	60	20
treatments	4	5672.5	1418.125
Error	12	153.5	12.792
Total	19	5886	

نطبق الصيغة (17) لحساب الآتي:

$$\widehat{M}_{e(RCD)} = \frac{(df_b)(MS_b) + (df_t + df_{e(RCBD)})MS_{e(RCBD)}}{df_b + df_t + df_{e(RCBD)}}$$

$$= \frac{(3)(20) + (4 + 12)(12.792)}{3 + 4 + 12} = 13.93$$

وإن

$$df_{e(RCD)} = df_{e(RCBD)} + df_b = 12 + 3 = 15$$

وعليه نحسب الكفاية النسبية وفق الصيغة (18) وكالآتي :

$$\boldsymbol{RE = \frac{(df_{e(RCBD)} + 1)(df_{e(RCD)} + 3)\widehat{M}_{e(RCD)}}{(df_{e(RCBD)} + 3)(df_{e(RCD)} + 1)MS_{e(RCBD)}}(100\%)}$$

$$= \frac{(12 + 1)(15 + 3)(13.93)}{(12 + 3)(15 + 1)(12.792)}(100\%) = 106.17\%$$

بما أن النتيجة ظهرت أكبر من 100% فهذا يعني أن هذه التجربة المنفذة بتصميم القطاعات الكاملة العشوائية لها كفاية بالمقارنة مع لو أنها كانت قد نفذت وفق تصميم تام التعشية نسبتها

$$RE\% - 100\% = \quad 106.17\% - 100\% = 6.17\%$$