(٥) تقدير التأثيرات باتباع طريقة المربعات الصغرى

(Least Squares Method)

ان الطريقة المُثلى لدراسة العلاقات الخطية التي توضحها النماذج الرياضية هي طريقة المربعات الصغرى والمبنية على أساس ان اوفق خط لمجموعة من المشاهدات هو الذى ينتج عنه ان مجموع مربع انحرافات قيم هذه المشاهدات عن الخط المستقيم المحسوب تكون اقل ما يمكن . ولقد سبق لنا ان ذكرنا ان المعادلة الرياضية لهذا التصميم والتي تحدد العلاقة الخطية بين المشاهدات هي :

$$y_{ij} = \mu + \tau_i + e_{ij}$$

ويمكن عن طريق استعمال التفاضل الوصول الى تقدير كل من μ و τ_i بحيث يصبح مجموع مربع انحرافات جميع المشاهدات عن الخط المفروض اقل ما يمكن ، وذلك باعتبار أن قيمة y_{ij} كدالة للمتغيرات المستقلة بالاضافة الى الخطأ العشوائي . وعلى ذلك فانه يلزم تقدير معاملات الانحدار الجزئية التالية :

$$\hat{\mu}, \hat{t}_1, \hat{t}_2, ..., \hat{t}_i, ... \hat{t}_t$$

ولكي يمكن الوصول الى حل معقول ، فانه لا بد من فرض الشرط بأن $\sum \tau_i = 0$ ويكون مجموع مربع الانحرافات الذي نريد بأن يكون أقل ما يمكن هو :

$$R = \sum_{ij} (y_{ij} - \hat{\mu} - \hat{\tau}_i)^2$$

والذي هو مجموع مربعات انحرافات الخطأ .

وبمفاضلة المعادلة R جزئيا بالنسبة لكل معامل انحدار $(\hat{\mu} \text{ and } \hat{\tau}_i)$ وبمساواة الناتج بصفر ، نحصل على

$$\frac{\partial R}{\partial \hat{\mu}} = -2 \sum_{ij} (y_{ij} - \hat{\mu} - \hat{t}_i) = 0$$

$$\frac{\partial R}{\partial \tau_i} = -2 \sum_{j} (y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0$$

$$\frac{\partial R}{\partial \tau_i} = -2 \sum_{j} (y_{ij} - \hat{\mu} - \hat{\tau}_i) = 0$$

ومن ذلك نحصل على عدد من المعادلات المستقلة Normal Equations عددها يساوي $t + 1$ وهي :

$$Y_{..} = rt\hat{\mu} + \sum_{i=1}^{t} \hat{\tau}_i$$

$$Y_{1.} = r\hat{\mu} + r\hat{\tau}_1$$

$$Y_{2.} = r\hat{\mu} + r\hat{\tau}_2$$

$$\vdots$$

$$Y_{i.} = r\hat{\mu} + r\hat{\tau}_i$$

وباستخدام الشرط المفروض سابقا بان $\Sigma\hat{\tau}_i = 0$ فان تقديرات التأثيرات باستخدام طريقة المربعات الصغرى تكون كما يلي :

$$\hat{\mu} = \frac{Y_{..}}{rt} = \bar{y}_{..}$$

$$\hat{\tau}_1 = \frac{Y_{1.} - r\hat{\mu}}{r} = \frac{Y_{1.}}{r} - \frac{Y_{..}}{rt} = \bar{y}_{1.} - \bar{y}_{..}$$

$$\hat{\tau}_2 = \frac{Y_{2.} - r\hat{\mu}}{r} = \frac{Y_{2.}}{r} - \frac{Y_{..}}{rt} = \bar{y}_{2.} - \bar{y}_{..}$$

$$\hat{\tau}_i = \frac{Y_{i.} - r\hat{\mu}}{r} = \frac{Y_{i.}}{r} - \frac{Y_{..}}{rt} = \bar{y}_{i.} - \bar{y}_{..}$$

(٦) تقدير مكونات التباين Estimation of Variance Components:

بالرجوع الى جدول تحليل التباين يتضح لنا أن التباين المتوقع للخطأ هو σ_ε^2 ويقدر بقيمة تباين الخطأ المحسوب من التجربة ، اي أن .. $\hat{\sigma}_\varepsilon^2 = MSe$

في حين ان التباين المتوقع بين المعاملات هو $\sigma_e^2 + r\sigma_\tau^2$ ويقدر بقيمة التباين بين المعاملات المحسوب من التجربة اي أن .. $\hat{\sigma}_\varepsilon^2 + r\hat{\sigma}_\tau^2 = MSt$

وعلى ذلك فان تباين تأثير المعاملات يمكن حسابه كما يلي :

$$\hat{\sigma}_{\tau}^2 = \frac{MSt - MSe}{r}$$

ويمكن اجراء بعض التقديرات الأخرى كما يـلي :

تباين أي مشاهدة في التجربة ..

$$S_{y_{ij}}^2 = MSe$$

الأنحراف القياسي (أو الخطأ القياسي) لأي مشاهدة ..

$$Sy_{ij} = \sqrt{S_{y_{ij}}^2} = \sqrt{MSe}$$

تباين متوسط أي معاملة ..

$$S_{\bar{y}_{i.}}^2 = \frac{MSe}{r}$$

الأنحراف القياسي (الخطأ القياسي) لمتوسط أي معامله ..

$$S\bar{y}_{i.} = \sqrt{\frac{MSe}{r}}$$

تباين الفرق بين متوسطي أي معاملتين $(\bar{y}_{i.} - \bar{y}_{i'.})$..

$$S^2_{(\bar{y}_{i.} - \bar{y}_{i'.})} = \frac{2MSe}{r}$$

الأنحراف القياسي (الخطأ القياسي) للفرق بين متوسطي أي معاملتين ...

$$S_{(\bar{y}_{i.} - \bar{y}_{i'.})} = \sqrt{\frac{2MSe}{r}}$$

تباين العلاقات الخطية بين متوسطات المعاملات يحسب أيضا باستخدام تباين الخطأ المقدر في جدول تحليل التباين ، فاذا فرضنا العلاقة الخطية التالية :

$$Z = C_1 \bar{y}_{1.} + C_2 \bar{y}_{2.} + \ldots\ldots\ldots C_t \bar{y}_{t.}$$

فان تباين هذه العلاقة الخطية يكون :

$$S_z^2 = \frac{MSe}{r}(C_1^2 + C_2^2 + \ldots\ldots\ldots + C_t^2)$$

ويكون انحرافها القياسي هو :

$$S_z = \sqrt{\frac{MSe}{r}} \sqrt{(C_1^2 + C_2^2 + \ldots\ldots\ldots + C_t^2)}$$

وأما معامل اختلاف التجربة فيحسب بمعرفة تباين الخطأ المقدر والمتوسط العام للتجربة كما يلي :

$$C.V = \frac{\sqrt{MSe}}{\bar{y}_{..}} \times 100$$

(٧) حدود الثقة حول المتوسطات والفروق بينها

Confidence Intervals Around the Means and their Differences

يكون من المرغوب احيانا وضع حدود للثقة حول متوسطات المعاملات أو حول الفرق بين متوسطي معاملتين وذلك بعد الانتهاء من تحليل التباين . ولكي يمكن حساب هذه الحدود لابد من تحديد الأحتمال المطلوب ، ومن المعروف أنه كلما زاد الاحتمال تباعدت حدود الثقة المحسوبة والتي نتوقع وجود المتوسط بينها بدرجة الاحتمال المفروض. ويستعمل توزيع t في حساب فترة الثقة كما يتضح من المعادلات التالية :

فمثلا نحسب حدود الثقة حول متوسط المعاملة i (أي معاملة) باحتمال ٩٥٪ ، نستخرج قيمة t من جدول توزيع t عند مستوى ٠,٠٥ وعند درجات حرية الخطأ ونحسب قيمة الخطأ القياسي لمتوسط اى معاملة $Sy_{i.}$ باستخدام قيمة MSe، ثم نحدد الحد الاعلى والحد الأدنى من حدود الثقة حول متوسط المعاملة كما يلي :

الحد الأدنى .. $\text{Lower Limit (L)} = \bar{y}_{i.} - t_{(0.05)}\, S\bar{y}_{i.}$

الحد الأعلى .. $\text{Upper Limit (U)} = \bar{y}_{i.} + t_{(0.05)}\, S\bar{y}_{i.}$

وليس معنى حساب الحدين السابقين اننا نعني بأن متوسط العشيرة-الذي تعتبر هذه المعاملة احدى عيناتها – تتراوح قيمته بين الحدين الأدنى والأعلى المقدرين ، فقيمة متوسط العشيرة ثابتة ولكننا نعني بأن قيمة متوسط العشيرة الثابتة تقع بين هذين الحدين في ٩٥٪ من الحالات . وطبعا اذا حسبنا هذه الحدود باحتمال أعلى (أي عند ٩٩٪) فان الحدين سيزدادان بُعداً (قيمة t عند ٠,٠١ أعلى منها عند ٠,٠٥) . وعلى العكس من ذلك فان مجال الثقة يضيق – أي يقترب الحدان الأدنى والأعلى كلما زاد حجم العينة أي زاد عدد التكرارات ، وذلك لأن قيمة $S\bar{y}_{i.}$ تقل بزيادة r كما أن قيمة t تقل أيضا كلما زادت درجات الحرية للخطأ .

وبالمثل يمكن حساب فترة أو مجال ثقة للفرق بين متوسطي معاملتين كما يلي :

$$(\bar{y}_{i.} - \bar{y}_{i.}') \pm t\, S_{(\bar{y}_{i.} - \bar{y}_{i.}')}$$

أي أن الحد الأدنى للفرق بين المتوسطين هو $(\bar{y}_{i.} - \bar{y}_{i.}') - tS_{(\bar{y}_{i.} - \bar{y}_{i.}')}$

والحد الأعلى هو $(\bar{y}_{i.} - \bar{y}_{i.}') + tS_{(\bar{y}_{i.} - \bar{y}_{i.}')}$

ولحساب هذين الحدين نحدد مستوى الاحتمال المطلوب ونستخرج قيمة t من جدول توزيعها عند درجات حرية الخطأ ثم نتم حساب المعادلة.

(٨) مثال محلول Example

الجدول (٣-٣) يوضح بيانات كمية الحاصل بالكغم للقطعة التجريبية في احدى تجارب القطن استخدم فيها اربع مستويات نيتروجين (وهي صفر ، ١٠ ، ٢٠ ، ٣٠ كغم N/دونم) لدراسة تأثير التسميد النيتروجيني على الصنف كوكر ١٠٠ ولت Cooker 100 wilt وذلك باستخدام التصميم العشوائي الكامل .

جدول (٣-٣) : بيانات حاصل القطن بالكغم / القطعة التجريبية من تجربة تأثير التسميد النيتروجيني على الصنف كوكر

المعاملات Treatments t_i	المشاهدات Observations y_{ij}				مجاميع المعاملات Treat.Totals $Y_{i.}$	متوسطات المعاملات Treat.Means $\bar{y}_i$
$N_0 = t_1$	24	52	45	55	234	46.80
$N_{10} = t_2$	92	115	64	66	370	74.00
$N_{20} = t_3$	98	100	45	95	422	84.40
$N_{30} = t_5$	88	140	172	157	729	145.80
					$Y_{..} = 1755$	$\bar{y}_{..} = 87.75$

١- النموذج الرياضي ..

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1,2,.....,4 \\ j = 1,2,....., \end{cases}$$

٢- خطوات التحليل واعداد جدول تحليل التباين (جدول ٣ - ٤) .

حساب معامل التصحيح (C) :

$$C = \frac{(Y_{..})^2}{tr} = \frac{(1755)^2}{(4)(5)} = 154001.25$$

حساب مجموع المربعات الكلية (SST) :

$$SST = \Sigma y_{ij}^2 - C$$

$$= (24)^2 + (52)^2 + ... + (157)^2 - 154001.25 = 37793.7$$

حساب مجموع مربعات المعاملات (SSt) :

$$SSt = \frac{\Sigma Y_{i.}^2}{r} - C$$

$$= \frac{(234)^2 + (370)^2 + (422)^2 + (729)^2}{5} - 154001.25$$

$$= 26234.95$$

حساب مجموع مربعات الخطأ (SSe) :

$$SSe = SST - SSt$$

$$= 37793.75 - 26234.95 = 11558.80$$

جدول (٣ – ٤) . جدول تحليل التباين لبيانات الجدول (٣ – ٣) .

S. O. V.	d.f	S.S	M.S	Cal.F	Tab. F	
					.05	.01
Treatments	3	26234.95	8744.98	12.11	3.24	5.29
Error	16	11558.80	722.42			
Tatal	19	37793.75				

٣- تقدير التأثيرات وحساب المقاييس الأحصائية التي سبق التعرف عليها ..

$$\hat{\mu} = \bar{y}_{..} = \frac{1755}{20} = 87.75$$

$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}$$

$$\hat{\tau}_2 = \bar{y}_{2.} - \bar{y}_{..} = 74.0 - 87.75 = -13.75$$

$$\hat{\varepsilon}_{ij} = y_{ij} - \bar{y}_{i.}$$

$$\hat{\varepsilon}_{34} = y_{34} - \bar{y}_{3.} = 45.0 - 84.4 = -39.4$$

$$\hat{\sigma}^2_{\varepsilon} = MSe = 722.42$$

$$\hat{\sigma}_t^2 = \frac{MSt - MSe}{r} = \frac{8744.98 - 722.42}{5} = 1604.51$$

$$S_{y_{ij}} = \sqrt{MSe} = \sqrt{722.42} = 26.9$$

$$S_{\bar{y}_{i.}} = \sqrt{\frac{MSe}{r}} = \sqrt{\frac{772.42}{5}} = 12.0$$

$$S_{(\bar{y}_{i.} - \bar{y}_{i'.})} = \sqrt{\frac{2MSe}{r}} = \sqrt{\frac{2(722.42)}{5}} = 17.0$$

$$C.V. = \frac{\sqrt{MSe}}{\bar{y}_{..}} \times 100 = \frac{\sqrt{722.42}}{87.75} \times 100 = 32\%$$

(٩) بعد تحليل التباين .. التمييز بين المتوسطات

After ANOVA ... Comparisons of Means

اذا كنا بعد تحليل التباين واجراء اختبار F قد قررنا وجود اختلافات معنوية بين المعاملات أي بين متوسطات المعاملات فانه لمن الطبيعي أن تبرز بعض الأسئلة .. فمثلا .. بين أي متوسطات توجد هذه الاختلافات ، هل متوسط المعاملة الأولى يختلف عن متوسط المعاملة الثانية ؟ ومتوسط المعاملة الثالثة عن الرابعة ؟ أم متوسط المعاملتين الأولى والثانية مثلا يختلف عن متوسط المعاملتين الثالثة والرابعة – اذا كان عندنا اربع معاملات – ؟ .. وهكذا .. وهناك اتجاهان للاجابة على مثل هذه الاسئلة التى تثار حول المتوسطات – يختلفان فى الوقت الذى يتم فيه اختيار المقارنات – التى تبدو مهمة فى نظر الباحث – بين المتوسطات .. أي هل تم تحديدها قبل اجراء التجربة أم بعد جمع البيانات ؟

(١) الاختبارات المحددة قبل اجراء التجربة

Tests determined before experimentation

أ– المقارنات المستقلة Orthogonal Comparisons

في معظم التجارب التي تشتمل على عدة معاملات ، سيهتم الباحث قطعا ببعض المقارنات الخاصة والمحددة بين متوسطات المعاملات ، الا أنه يجب اختيار وتحديد هذه المقارنات التى سيوجه اليها اهتمامه قبل بدء تنفيذ التجربة وعليه فان اجراء المقارنات