الفصل السادس

تحليل التباين باتجاه واحد

One-Way Analysis of Variance

$6-1$ مقدمة :

رأينا في بحث اختبار الفرضيات كيف يمكننا أن نقارن بين وسيطين أو انحرافين معيارين لعينتين مختلفتين بهدف معرفة ما إذا كان الفرق بينهما جوهريا أم عشوائيا، وذلك باستخدام التوزيع الطبيعي أو توزيع ستيودنت حسب حجم العينة.

ولكن عندما يكون لدينا عدة عينات ونريد أن نقارن بين أوساطها أو تبايناتها باستعمال الطرق السابقة في اختبار الفرضيات، سيكون لدينا عدد كبير من المقارنات تُحسب بطريقة المجموعات. فإذا كان لدينا 5 عينات وأردنا أن نقارن بين أوساطها الحسابية، يجب أن نقوم بعدد كبير من المقارنات حسب العلاقات الآتية:

$$C_n^k = \frac{n!}{k!(n-k)!}$$

$$C_5^2 = \frac{5!}{2!(5-2)!} = \frac{5\times 4\times 3!}{2\times 1\times 3!} = 10$$

إن عملية من هذا النوع باهظة التكاليف من حيث الوقت والجهد، إضافة إلى أن احتمال الحصول على جواب خاطئ سوف يزداد. فلو فرضنا أن درجة الثقة 95% فإن احتمال الحصول على جواب خاطئ في كل عملية هو 5%

واحتمال الحصول على جواب صحيح في كل هذه المقارنات عندما نجري 10 مقارنات هو جداء احتمال تحقق الحادث المرغوب بنفسه 10 مرات(حسب مبدأ استقلالية الحوادث)،ويكون لدينا:

- احتمال الحصول على جواب صحيح يساوي:

$$P(A) = (0.95)^{10} \approx 0.6$$

- واحتمال الحصول على جواب واحد خاطئ على الأقل يساوي:

$$P(\overline{A}) = 1 - P(A) = 1 - 0.6 = 0.4$$

لو أخذنا عينات عند مستوى دلالة $\alpha = 0.05$ عند كل اختبار، فإن:

- احتمال الوصول إلى قرار صحيح للاختبارات الخمسة عشرة يساوي:

$$P(A) = (0.95)^{15} \approx 0.46$$

- واحتمال الحصول على جواب واحد خاطئ على الأقل يساوي:

$$P(\overline{A}) = 1 - P(A) = 1 - 0.46 = 0.54$$

ومنه نستنتج أن احتمال الوصول إلى قرار غير سليم يكبر كلما كبر عدد العينات، لذلك يتطلب منا البحث عن طريقة إحصائية جيدة يمكن استخدامها لمثل هذه الحالات. وبالتالي فإن الاختبار المناسب هو اختبار F الذي نحصل عليه من تحليل التباين، والذي يُبنى على تفكيك التباين إلى عدة مركبات جزئية تشكل كل منها قياسا للتغير .

فقد تكون التغيرات داخلية ضمن كل مجتمع من المجتمعات المدروسة أو تكون تغيرات بين المجتمعات المدروسة . ومن أحد الطرق المهمة في التحليل هو تحليل التباين في اتجاه واحد ويدعى تحليل آحادي التصنيف، لأنه يتضمن عاملاً واحدا فقط، ويستخدم لدراسة الفروقات في الأوساط الحسابية لمشاهدات متغير ما تعرض لمؤثر خارجي، ويعتبر من أفضل الأساليب الإحصائية المستخدمة في التحليل عندما تتعلق الدراسة بمقارنة عدة متوسطات لمجتمعات ثم نحسب منها عينات بأحجام متساوية أو مختلفة.

6 – 2 أسس تحليل التباين:

وجدنا أنه عندما يكون لدينا عدد كبير من العينات، فمن غير الممكن استخدام الطرائق السابقة في الاستدلال الإحصائي بالاعتماد على التوزيع الطبيعي أو على توزيع ستيودنت في إجراء اختبارات من الأنواع التي تعرضنا لها في الفصول السابق. لهذا نلجأ في مثل هذه الحالات إلى طريقة تحليل التباين التي تتم وفق المراحل الآتية:

أولا– تحديد مجموع التباينات المختلفة:

1 –التباين الكلي ويمثل مجموع مربع انحرافات كل قيمة من قيم المشاهدات عن الوسط الحسابي العام، ويقيس مدى تشتت جميع قيم الظاهرة المدروسة حول وسطها العام.

2 –توزيع التباين الكلي إلى عدة أجزاء بحسب تصنيف المعلومات وفق:

- معيار واحد، ويسمى " التصنيف الأحادي " .
- معيارين ، ويسمى " التصنيف الثنائي "
- ثلاث معايير فأكثر ، ويسمى " التصنيف المتعدد "

ثانيا– مقارنة التباينات المختلفة:

تتم مقارنة التباينات بالاعتماد على توزيع F لتقدير فيما إذا كانت الفروق بين مختلف الأوساط جوهري أم لا.

6 – 3 النموذج الرياضي العام لتحليل التباين أحادي التصنيف:

لنفرض أنه لدينا k عينة عشوائية بسيطة مستقلة عن بعضها بعضاً ذات أحجام $n_1, n_2, n_3, \ldots\ldots n_k$ ، حيث أن $n = \sum_{i=1}^{k} n_i$ محسوبة من k مجتمع موزعة توزعا طبيعيا أوساطها الحسابية $\mu_1, \mu_2, \mu_3, \ldots\ldots \mu_k$ غير معلومة وتبايناتها متساوية.

لنرمز بـ x_{ij} إلى القيم العددية المشاهدة والتي ترتيبها j من العينة المسحوبة من المجتمع i. حيث $(i = 1,2,3,\ldots k)$ و $(j = 1,2,3,\ldots n)$ ، ويعتبر x_{ij} متحول عشوائي وسطه الحسابي μ_1 وانحرافه المعياري σ .

يمكن التعبير عن النموذج النظري على اعتبار أن هذه المتحولات مستقلة وضمن الشروط السابقة، كالآتي:

$$x_{ij} = \mu_i + e_{ij} \qquad (1-6)$$

أو

$$e_{ij} = x_{ij} - \mu_i$$

حيث أن e_{ij} متحولات عشوائية مستقلة وموزعة توزعا طبيعيا ، وسطها الحسابي معدوم وتبايناتها متساوية.

إذا رمزنا بـ α_i إلى تأثير المجتمع i ، يمكننا أن نكتب:

$$\alpha_i = \mu_i - \mu \Rightarrow \mu_i = \mu + \alpha_i \qquad (2-6)$$

حيث أن μ_i يمثل الأوساط الحسابية لمختلف المجتمعات، و μ الوسط الحسابي العام. وبتعويض قيمة μ_i بما يعادلها في المعادلة $(1-6)$ ، يكون لدينا :

$$x_{ij} = \mu + \alpha_{ij} + e_{ij} \qquad (3-6)$$

وحيث أن $\sum_{i=1}^{k} \alpha_i = 0$ ، إذا كانت جميع العينات مسحوبة من مجتمع واحد لأن $E(\mu_i) = \mu$

وتكتب أيضا

$$x_{ij} - \mu = \alpha_{ij} + e_{ij} \qquad (4-6)$$

وحيث أن

$$\alpha_{ij} = \mu_i - \mu \qquad (5-6)$$

α_i تعبر عن انحرافات المتوسطات μ_i لمختلف المجتمعات عن الوسط الحسابي العام μ ، وحيث أن:

$$\mu = \frac{1}{n}\sum_{i=k}^{n} n_i \mu_i \qquad (6-6)$$

وبتعويض قيمة μ في العلاقة $(6-6)$ بما يعادلها في العلاقة $(6-5)$ ينتج لدينا :

$$\alpha_{ij} = \mu_i - \frac{1}{n}\sum_{i=1}^{k} n_i \mu$$

بتثقيل قيم α_i بالتكرارات n_i وبأخذ المجموع ، يكون لدينا :

$$\sum n_i \alpha_i = \sum_{i=1}^{k} n_i \mu_i - \frac{1}{n}\sum_{i=1}^{k} n_i \mu_i \sum_{i=1}^{k} n_i \qquad (7-6)$$

$$\sum_{i=1}^{k} n_i \alpha_i = 0$$

إن جميع الشروط السابقة يجب أن تتوفر في النموذج النظري ومن خلالها سيتم اختبار الفرضيات.

6 – 3 – 1 الاختبار:

إن اختبار الفرضيات بالاعتماد على أسس تحليل التباين يتم بعد دراسة النقاط الآتية :

1 –تحديد الفرضيات:

الفرضية الابتدائية: $H_0 : \alpha_I = 0$

الأوساط الحسابية لـ k مجتمعا متساوية حيث $(i = 1,2,3,\ldots k)$

الفرضية البديلة(على الأقل من أجل قيمة واحدة لـ i) $H_1 : \alpha_I \neq 0$

الأوساط الحسابية لـ k مجتمعا غير متساوية .

2 –حساب التباينات:

قبل البدء بحساب التباينات لابد من التعريف بـ :

الأوساط الحسابية لمختلف العينات ، وتعرف بالعلاقة الآتية :

$$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$$

الوسط الحسابي العام لجميع العينات، وتعرف بالعلاقة الآتية :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij} = \frac{1}{n}\sum_{i=1}^{k} n_i \bar{x}_i$$

كما ويمكننا التعبير عن النموذج المشاهد لتحليل التباين بـ:

$$(\bar{x}_{ij} - \bar{x}) = (\bar{x}_i - \bar{x}) + (x_{ij} - \bar{x}_i) \qquad (8-6)$$

ذلك يعني أن انحرافات القيم عن الوسط الحسابي العام تدعى بالتباين الكلي الذي ينقسم بدوره إلى جزأين، هما انحرافات الأوساط الحسابية للعينات عن الوسط الحسابي العام والذي نسميه بتباين المعاملات، بينما الانحرافات الموجودة داخل العينات نسميها بتباين الأخطاء.

وإذا قارنا العلاقة $(6-8)$ مع $(6-4)$ بعد كتابة هذه الأخيرة، بالاعتماد على العلاقتين $(6-5)$ و $(6-1)$ وفق الآتي:

$$(\bar{x}_{ij} - \mu) = (\mu_i - \mu) + (x_{ij} - \mu_i)$$

نلاحظ وجه التشابه بين النموذج النظري والنموذج المشاهد.

بتربيع العلاقة $(6-8)$ وبأخذ المجموع بالنسبة إلى جميع القيم المشاهدة، نحصل على معادلة تحليل التباين وفق الآتي:

$$\sum_{i=1}^{k}\sum_{j-1}^{n_i}(x_{ij} - \bar{x})^2 = \sum\sum(\bar{x}_i - \bar{x})^2 + \sum\sum(x_{ij} - \bar{x}_i)^2 + 2\sum\sum(x_{ij} - \bar{x})(x_i - \bar{x}) \qquad (9-6)$$

نلاحظ أن الحد الأخير يساوي الصفر لأن $2\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)(\bar{x}_i-\bar{x}) = 2\sum_{i=1}^{k}\left[(\bar{x}_i-\bar{x})\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)\right]$

وبملاحظة أن $\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i) = 0$ ، وذلك لأن مجموع انحرافات القيم حول وسطها الحسابي يساوي الصفر، وبالتالي نجد:

$$\sum_{i=1}^{k}\sum_{j-1}^{n_i}(x_{ij}-\bar{x})^2 = \sum\sum(\bar{x}_i-\bar{x})^2 + \sum\sum(x_{ij}-\bar{x}_i)^2 \qquad (10-6)$$

ويمكن أن تكتب أيضا على الشكل التالي :

$$\sum_{i=1}^{k}\sum_{j-1}^{n_i}(x_{ij}-\bar{x})^2 = \sum_{i=1}^{k} n_i(\bar{x}_i-\bar{x})^2 + \sum\sum(x_{ij}-\bar{x}_i)^2 \qquad (11-6)$$

فإذا رمزنا :

إلى الحد الموجود في يسار المعادلة الذي يعبر عن مجموع مربع الانحرافات عن الوسط الحسابي العام بـ SST والذي نسميه بالتباين الكلي، وإلى الحد الأول من يمين المعادلة الذي يعبر عن مجموع مربع الانحرافات بين العينات بـ SSB والذي نسميه بتباين المعاملات، والحد الثاني عن يمين المعادلة الذي يعبر عن مجموع مربعات الانحرافات داخل العينة بـ SSE والذي نسميه يتباين الأخطاء.

يمكننا عندها أن نكتب:

$$SST = SSB + SSE \qquad (12-6)$$

حساب التوقع الرياضي للتباينات:

لإيجاد التوقع الرياضي للتباين الكلي، نكتب:

$$E(SST) = E(SSB) + E(SSE) \qquad (13-6)$$

حيث أن :

$$SST = \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \bar{x})^2 \qquad (14-6)$$
$$= \sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \alpha_i - \bar{x})^2 + \sum\sum \alpha_i^2 + 2\sum_{i=1}^{k}\sum_{j=1}^{n_i}\alpha_i(x_{ij} - \alpha_i - \bar{x})$$
$$= \sum\sum(x_{ij} - \alpha_i - \bar{x})^2 + \sum_{i=1}^{k} n_i.\alpha_i^2$$

وذلك لأن الحد الأخير يساوي الصفر اعتمادا على العلاقة $(6-7)$ ، وبما أن x_{ij} هي القيم المشاهدة لـ n متحولا طبيعيا، أوساطها الحسابية تساوي $\mu_i = \mu + \alpha_i$ ، وتبايناها σ^2 ، فالانحرافات $(x_{ij} - \alpha_i)$ هي قيم مشاهدة لـ n متحولا طبيعيا لها الوسط الحسابي نفسه μ ، والتباين نفسه σ^2 ، كما أن متوسط القيم $(x_{ij} - \alpha_i)$ يساوي:

$$\frac{1}{n}\sum_{i=1}^{k}\sum_{j-1}^{n_i}(x_{ij} - \alpha_i) = \frac{1}{n}\sum_{i=1}^{k}\sum_{j=1}^{n_i}x_{ij} - \frac{1}{n}\sum_{i=1}^{k}n_i\alpha_i = \bar{x} \qquad (15-6)$$

وذلك نظرا إلى كون الحد الثاني من الفرق معدوم، بالاعتماد على العلاقة $(6-7)$.

إذا رمزنا بـ x_{ij} و $\bar{x}$ إلى المتحولات العشوائية الموافقة، وبالرجوع إلى العلاقة $(6-3)$ ، يمكننا إن نكتب:

$$E(x_{ij} - \alpha_i - \bar{x}) = E(\mu + e_{ij} - \bar{x})$$

وبما أن $E(e_{ij}) = 0$ ، لذلك نكتب:

$$E(x_{ij} - \alpha_i - \bar{x}) = E(\mu - \bar{x})$$

وبالاعتماد على مفاهيم العزوم الابتدائية والمركزية، يكون لدينا:

$$E(\mu-\bar{x})^2 = E(x_{ij}-\alpha_i-\bar{x})^2 = \sigma^2$$

أي أن:

$$E(x_{ij}-\alpha_i-\bar{x})^2 = \sigma^2 \qquad ,i=1,2,3..k \quad j=1,2,3..n_i$$

وبما أنه يمكن حساب قيمة كل هذه الانحرافات التي تساوي عددها $\sum_{i=1}^{k} n_i = n$ انطلاقا من عينة عشوائية بسيطة تشمل عددا من القيم $(x_{ij}-\alpha)$ يساوي n والمحسوبة من مجتمع إحصائي واحد تباينه σ^2 (حجم العينات غير متساو)، وبالتالي يمكننا أن نكتب:

$$E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\alpha_i-\bar{x})^2\right] = (n-1).\sigma^2 \qquad (16-6)$$

وبذلك نجد أن التوقع الرياضي للتباين الكلي يساوي:

$$E(SST) = E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij}-\alpha_i-\bar{x})^2\right] + E\left[\sum_{i=1}^{k} n_i\alpha_i^2\right]$$

أو

$$E(SST) = (n-1)\sigma^2 + \sum_{i=1}^{k} n_i\alpha_i^2 \qquad (17-6)$$

أما فيما يتعلق بمجموع انحرافات الأخطاء، يكون لدينا بالنسبة لكل عينة منفردة:

$$E\left[\sum_{j=1}^{n_i}(x_{ij}-\bar{x}_i)^2\right] = (n_i-1)\sigma^2 \qquad (18-6)$$

وإلى مجموع العينات :

$$E(SSE) = E\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}(x_{ij} - \alpha_i - \bar{x})^2\right]$$
$$= \sum_{i=1}^{k} E\left[\sum_{j=1}^{n_i}(x_{ij} - \bar{x}_i)^2\right] = \sum_{i=1}^{k}(n_i - 1)\sigma^2$$

أو

$$E(SSE) = (n-k)\sigma^2 \qquad (19-6)$$

بالتعويض في $(13-6)$ ، نجد :

$$E(SSB) = (k-1)\sigma^2 + \sum_{i=1}^{k} n_i.\alpha_i^2 \qquad (20-6)$$

بالاعتماد على قوانين التوزيعات الاحتمالية وبشكل خاص على توزيع χ^2 ، يمكننا أن نقول:

أن معامل σ^2 بالنسبة إلى مختلف القيم المتوقعة ، يمثل ما نسميه بدرجات الحرية، وبتقسيم مجموع مربعات الانحرافات على درجات الحرية المتعلقة ، نحصل على التباينات وفقا لما يلي:

– متوسط تباين المعاملات (بين العينات)، ويعطى بالعلاقة الآتية:

$$MSB = \frac{SSB}{k-1} \qquad (21-6)$$

متوسط تباين الأخطاء(داخل العينات)، ويعطى بالعلاقة الآتية:

$$MSE = \frac{SSE}{n-k} \qquad (22-6)$$

وبالتالي فإن قيمة مؤشر الاختبار $\widetilde{F}$ يعطى بالعلاقة :

$$\widetilde{F} = \frac{\frac{SSB}{k-1}}{\frac{SSE}{n-k}} = \frac{MSB}{MSE} \qquad (23-6)$$

والتي تخضع بتغيراتها إلى توزيع F (فيشر– سينديكور) بدرجات حرية في الصورة تساوي $(k-1)$ وبدرجات حرية في المخرج تساوي $(n-k)$ ، وهذه الكمية هي القيمة المقدرة للمتغير F ، ونرمز لها بـ $\widetilde{F}$.

نقارن هذه القيمة مع القيمة الحرجة (الجدولية) $F_{(1-\alpha,k-1,n-k)}$ ، فإذا كان $\widetilde{F} > F_{(1-\alpha,k-1,n-k)}$ فإننا نرفض الفرضية H_0 ونقبل الفرضية H_1 .

يمكن تلخيص ما تقدم في جدول تحليل التباين على الشكل الآتي:

قيمة مؤشر الاختبار	متوسط المربعات	مجموع مربع الانحرافات	درجات الحرية	البيان / مصدر التباين
$\widetilde{F} = \frac{MSB}{MSE}$	MSB	SSB	$k-1$	التباين بين المعاملات (بين العينات)
	MSE	SSE	$n-k$	تباين الأخطاء (داخل العينات)
		SST	$n-1$	التباين الكلي

ولتحديد عناصر جدول تحليل التباين وحساب قيمة $\widetilde{F}$ ، فإننا نحتاج إلى جدول الحسابات المساعد الآتي:

المجتمع / العينة	1	2	.....	k	المجموع
1	x_{11}	x_{12}	.....	x_{1k}	
2	x_{21}	x_{22}	.....	x_{2k}	
3	x_{31}	x_{32}	.....	x_{3k}	
.....	.....	.....	.....	.....	
n	x_{n1}	x_{n2}	.....	x_{nk}	
n_i	n_1	n_2	.....	n_k	$\sum_{i=1}^{k} n_i = n$
$x_i = \sum_{j=1}^{n_i} x_{ij}$	x_1	x_2	.....	x_k	$\sum_{i=1}^{k} x_i = X$
$\sum_{j=1}^{n_i} x_{ij}^2$	$\sum_{j=1}^{n_1} x_{1j}^2$	$\sum_{j=1}^{n_2} x_{2j}^2$	.....	$\sum_{j=1}^{n_k} x_{kj}^2$	$\sum_{i=1}^{k}\sum_{j=1}^{n_i} x_{ij}^2 = T$
$\frac{X_i^2}{n_i}$	$\frac{X_1^2}{n_1}$	$\frac{X_2^2}{n_2}$	.....	$\frac{X_k^2}{n_k}$	
$(SSE)_i$	$(SSE)_1$	$(SSE)_2$	.....	$(SSE)_k$	$\sum_{i=1}^{k} (SSE)_i = SSE$
$\sigma_i^2 = \frac{(SSE)_i}{(n_i - 1)}$	$\tilde{\sigma}_1^2$	$\tilde{\sigma}_2^2$	.....	$\tilde{\sigma}_k^2$	
$\bar{x}_i = \frac{x_i}{n_i}$	$\bar{x}_1$	$\bar{x}_2$	......	$\bar{x}_k$	

حيث أن:

$$\bar{x} = \frac{X}{N} \qquad (24-6)$$

و

$$(SSE)_i = \sum_{j=1}^{n_i} x_{ij}^2 - \frac{X_i^2}{n_i} \qquad (25-6)$$

و

$$SST = T - \frac{X^2}{n} \qquad (26-6)$$

مثال $(1-6)$:

قام احد منتجي المنظفات بعرض إنتاجه بثلاث طرق:

الطريقة الأولى: عبوات بلاستيكية – وقد تم عرض المنتج في 6 مخازن.

الطريقة الثاني: أكياس نايلون – وقد تم عرض المنتج في 4 مخازن.

الطريقة الثالثة: غير معبأة – وقد تم اختيارها بشكل عشوائي.

وقد كانت المبيعات الشهرية في كل مخزن وحسب كل طريقة كما في الجدول الآتي:

المخازن	طرق عرض المنتج		
	1	2	3
1	50	45	55
2	55	35	30
3	40	40	60
4	35	40	45
5	60	–	60
6	66	–	55
7	–	–	65
8	–	–	70

والمطلوب اختبار تأثير الطرق الثلاث على مبيعات المنظفات إذا علمت المبيعات مقدرة بـ كغ وذلك عند $\alpha = 0.05$

الحل:

1 -نضع الفرضيات:

فرضية العدم H_0 : لا توجد فروق بين متوسطات مبيعات المنظفات حسب الطرق الثلاث .

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = 0$$

الفرضية البديلة H_1 : توجد فروق بين متوسطات مبيعات المنظفات حسب الطرق الثلاث .

على الأقل من أجل قيمة واحدة لـ i ، $i = 1,2,3$

$$H_1 : \alpha_i \neq 0$$

2 –إيجاد القيمة الفعلية لإحصائية الاختبار: نشكل الجدول المساعد الآتي:

المخازن i	طرق عرض المنتج j			المجموع
	1	2	3	
1	50	45	55	
2	55	35	30	
3	40	40	60	
4	35	40	45	
5	60	–	60	
6	66	–	55	
7	–	–	65	
8	–	–	70	
n_i	6	4	8	$\sum_{i=1}^{3} n_i = 18$
$x_i = \sum_{j=1}^{n_i} x_{ij}$	306	160	440	$\sum_{i=1}^{3} x_i = X = 906$
$\sum_{j=1}^{n_i} x_{ij}^2$	16306	6450	25300	$\sum_{i=1}^{3}\sum_{j=1}^{8} x_{ij}^2 = T = 48056$
$\frac{X_i^2}{n_i}$	15606	6400	24200	
$(SSE)_i$	700	50	1100	$\sum_{i=1}^{3} (SSE)_i = 1850$

ولحساب التباين الكلي نطبق العلاقة $(6\text{–}26)$ ، فنجد :

$$SST = 48056 - \frac{(906)^2}{18} = 2454$$

أما تباين الأخطاء (داخل العينات) نجده من الجدول المساعد ويساوي $SSE = 1850$.

بينما تباين المعاملات (بين العينات) نجده بتطبيق العلاقة $(6-13)$ ، فنجد:

$$SSB = 2454 - 1850 = 604$$

نحسب متوسط المربعات من العلاقات $(6-21)$ و $(6-22)$ ، كالآتي:

$$MSB = \frac{604}{2} = 302$$

$$MSE = \frac{1850}{15} = 123.33$$

وبالتالي نجد قيمة $\tilde{F}$ من العلاقة $(6-23)$ ، كالآتي:

$$\tilde{F} = \frac{302}{123.33} = 2.45$$

يمكن تلخيص ما تقدم في جدول تحليل التباين على الشكل الآتي:

البيان / مصدر التباين	درجات الحرية	مجموع مربع الانحرافات	متوسط المربعات	قيمة مؤشر الاختبار
التباين بين المعاملات (بين العينات)	2	$SSB = 604$	$MSB = 302$	$\widetilde{F} = 2.45$
تباين الأخطاء (داخل العينات)	15	$SSE = 1850$	$MSE = 123,33$	
التباين الكلي	17	$SST = 2454$		

3 -إيجاد القيمة الجدولية $\widetilde{F}_{0.05(2,15)} = 3.68$.

4 -المقارنة واتخاذ القرار:

نلاحظ أن القيمة الفعلية المحسوبة $\widetilde{F}$ هي أصغر من القيمة الجدولية $\widetilde{F}_{0.05(2,15)} = 3.68$ ، لذلك نقبل H_0 والقائلة بأنه لا توجد فروق جوهرية بين متوسطات المبيعات بالطرق الثلاث وبالتالي لا يوجد تأثير لكل طريقة على كمية المبيعات.

حالة خاصة:

إذا كانت العينات ذات أحجام متساوية $n_1 = n_2 = n_3 = ...= n_m$ ، فإن ذلك يؤدي إلى تبسيط العمليات الحسابية بصورة عامة والجدول الآتي يوضح ذلك:

المجتمع / العينة	1	2	.....	k	المجموع
1	x_{11}	x_{12}	.....	x_{1k}	
2	x_{21}	x_{22}	.....	x_{2k}	
3	x_{31}	x_{32}	.....	x_{3k}	
.....	.....	.....	.....	.....	
n	x_{n1}	x_{n2}	.....	x_{nk}	
n_i	n_1	n_2	.....	n_k	$\sum_{i=1}^{k} n_i = n$
$x_i = \sum_{j=1}^{n} x_{ij}$	x_1	x_2	.....	x_k	$\sum_{i=1}^{k} x_i = X$
$\sum_{j=1}^{n} x_{ij}^2$	$\sum_{j=1}^{n_1} x_{1j}^2$	$\sum_{j=1}^{n_2} x_{2j}^2$	.....	$\sum_{j=1}^{n_k} x_{kj}^2$	$\sum_{i=1}^{k}\sum_{j=1}^{n} x_{ij}^2 = T$
$\frac{X_i^2}{n_i}$	$\frac{X_1^2}{n_1}$	$\frac{X_2^2}{n_2}$	.....	$\frac{X_k^2}{n_k}$	
$(SSE)_i$	$(SSE)_1$	$(SSE)_2$	.....	$(SSE)_k$	$\sum_{i=1}^{k} (SSE)_i = SSE$

حيث أن:

$$(SSE)_i = \sum_{j=1}^{n} x_{ij}^2 - \frac{X_i^2}{n}$$

و

$$SST = T - \frac{X^2}{n_k}$$

يمكن تلخيص ما تقدم في جدول تحليل التباين على الشكل الآتي:

البيان / مصدر التباين	درجات الحرية	مجموع مربع الانحرافات	متوسط المربعات	قيمة مؤشر الاختبار
التباين بين المعاملات (بين العينات)	$k-1$	SSB	MSB	$\tilde{F} = \frac{MSB}{MSE}$
تباين الأخطاء (داخل العينات)	$k(n-1)$	SSE	MSE	
التباين الكلي	$k.n-1$	SST		

مثال $(6-2)$:

قام احد الباحثين بإجراء تجارب حول استخدام أربعة أنواع من الأسمدة وتأثيرها على مردود إنتاج القطن خلال ستة تجارب وكانت المعلومات التي حصل عليها، كالآتي:

i	j			
	1	2	3	4
1	50	48	42	64
2	52	56	48	56
3	53	60	50	55
4	60	62	50	54
5	49	50	47	48
6	54	54	51	47

والمطلوب اختبار تأثيرات الأسمدة على مردودية إنتاج القطن وذلك عند مستوى دلالة $\alpha = 0.05$.

الحل:

1 -نضع الفرضيات:

فرضية العدم H_0 : لا توجد فروق بين متوسطات الإنتاج وأن العينات الأربعة مأخوذة من مجتمع واحد .

$$H_0 : \alpha_1 = \alpha_2 = \alpha_3 = \alpha_4 = 0$$

الفرضية البديلة H_1 : توجد فروق بين متوسطات الإنتاج وأن العينات الأربعة مأخوذة من مجتمع واحد .

$$H_1 : \alpha_i \neq 0$$

على الأقل من أجل قيمة واحدة لـ i ، $i = 1,2,3,4$

2 –إيجاد القيمة الفعلية لإحصائية الاختبار :

لإيجاد القيمة الفعلية لمؤشر الاختبار $\tilde{F}$: نشكل الجدول المساعد الآتي:

القطن i	أنواع الأسمدة j				المجموع
	1	2	3	4	
1	50	48	42	64	
2	52	56	48	56	
3	53	60	50	55	
4	60	62	50	54	
5	49	50	47	48	
6	54	54	51	47	
n	6	6	6	6	$\sum_{i=1}^{4} n_i = 24$
$x_i = \sum_{j=1}^{n_i} x_{ij}$	318	330	288	324	$\sum_{i=1}^{3} x_i = X = 1260$
$\sum_{j=1}^{n_i} x_{ij}^2$	16930	18300	13878	17686	$\sum_{i=1}^{3}\sum_{j=1}^{8} x_{ij}^2 = T = 66794$
$\frac{X_i^2}{n_i}$	16854	18150	13824	17496	
$(SSE)_i$	76	150	54	190	$\sum_{i=1}^{3} (SSE)_i = 470$

ولحساب التباين الكلي نطبق العلاقة (6-26)، فنجد:

$$SST = 66794 - \frac{(1260)^2}{24} = 644$$

أما تباين الأخطاء (داخل العينات) نجده من الجدول المساعد ويساوي $SSE = 470$.

بينما تباين المعاملات (بين العينات) نجده بتطبيق العلاقة $(6-13)$ ، فنجد:

$$SSB = 644 - 470 = 174$$

نحسب متوسط المربعات من العلاقات $(6-21)$ و$(6-22)$ ، كالآتي:

$$MSB = \frac{174}{3} = 58$$

$$MSE = \frac{470}{20} = 23.5$$

وبالتالي نجد قيمة $\tilde{F}$ من العلاقة $(6-23)$ ، كالآتي:

$$\tilde{F} = \frac{58}{23.5} = 2.47$$

يمكن تلخيص ما تقدم في جدول تحليل التباين على الشكل الآتي:

البيان / مصدر التباين	درجات الحرية	مجموع مربع الانحرافات	متوسط المربعات	قيمة مؤشر الاختبار
التباين بين المعاملات (بين العينات)	3	$SSB = 174$	$MSB = 58$	$\tilde{F} = 2.47$
تباين الأخطاء (داخل العينات)	20	$SSE = 470$	$MSE = 23.5$	
التباين الكلي	23	$SST = 644$		

3 -إيجاد القيمة الجدولية $\tilde{F}_{0.05(3,20)} = 3.10$.

4 -المقارنة واتخاذ القرار :

نلاحظ أن القيمة الفعلية المحسوبة $\tilde{F}$ هي أصغر من القيمة الجدولية $\tilde{F}_{0.05(3,20)} = 3.10$ ، لذلك نقبل H_0 والقائلة بأنه لا توجد فروق جوهرية بين متوسطات الإنتاج ، وبالتالي لا يوجد فروق بين تأثيرات الأسمدة المتنوعة على إنتاج القطن .

تمارين عامة

1 –اختار احد الباحثين في كلية الزراعة ثلاث عينات من الخراف من إحدى المزارع لتربية الخراف وبشكل عشوائي ، ثم طبق على كل عينة نوع خاص من التغذية وبعد شهر وجد أن الزيادة الحاصلة في الوزن (بالكغ) كانت كالآتي:

ترتيب الخروف	1	2	3	4	5	6	7	8	9	10
العينة A	5	7	4	3	4	5	6	–	–	–
العينة B	3	4	4	3	5	4	6	5	–	–
العينة C	6	7	4	5	8	7	4	5	8	6

والمطلوب معرفة فيما إذا كان هناك تأثير لنوع التغذية على الزيادة في وزن الخراف مستخدما مستوى دلالة $\alpha = 0.01$ ثم $\alpha = 0.05$.

2 –لاختبار تأثير أربعة أنواع من التكنولوجيا A, B, C, D على انتاج القمح، فقد تم اختيار أربع قطع من الأرض متجانسة تماما. ثم تم تطبيق كل نوع من التكنولوجيا على قطعة أرض واحدة ولمدة أربع سنوات ، فكانت النتائج مقدرة (كغ/ الدونم) كالآتي:

إنتاج القمح i	نوع التكنولوجيا j			
	A	B	C	D
1	200	315	270	250
2	250	320	290	260
3	225	290	310	270
4	240	280	330	240

والمطلوب اختبار مدى تأثير التكنولوجيا المستخدمة على إنتاج القمح مستوى دلالة $\alpha = 0.01$.

3 –أربعة طلاب من كلية الاقتصاد أخضعوا لثلاثة اختبارات إحصائية، مفترضين أنها من نفس درجة الصعوبة، فكانت النتائج، كالآتي:

الطلاب i	الاختبارات j		
	I	II	III
1	10	12	13
2	9	12	12
3	10	11	11
4	10	11	12

فهل نستطيع ضمن مستوى دلالة $\alpha = 0.05$ قبول فرضية أن هؤلاء الطلاب نفس المستوى العلمي ؟

4 –أدعى أحد مقتني منظفات الغسيل أن منتجه يزيل جميع بقع زيت الشحوم والسيارات مهما كان نوع الماء .

وللتأكد من صحة أدعائه لجأ مركز حماية المستهلك إلى اختبار هذا المنظف في خمسة منازل مختلفة اختيرت عشوائيا وتتميز بماء (كلسية ، عذبة ، عادية)، وتم استخدام هذا المنظف على عدة عينات من الثياب المتماثلة في النوع والملوثة وزن كل عينة 100كغ ، وقد تبين إثرها أن هناك بعض الثياب التي لم يستطع المنظف إزالة الزيوت عنها وقد تم تحديد وزن الثياب مقدرا بالغرام التي مازلت ملوثة بعد الغسيل من كل عينة كالآتي:

العينة i	المياه j		
	كلسية	عادية	عذبة
1	5	4	4
2	3	7	–
3	2	8	1
4	10	3	3
5	6	2	2

والمطلوب اختبر ضمن مستوى دلالة $\alpha = 0.05$ وهل يمكن القول أن نوعية الماء لديها تأثير في المنظف ؟