الفصل الثالث عشر

تحليل التباين

(١٣-١) مقدمة :

درسنا في الفصول السابقة كيفية إجراء اختبار إحصائي للفرق بين وسطين حسابيين أو اختبار تساوي وسطين حسابيين سواءً كان ذلك باستخدام التوزيع الطبيعي أو باستخدام توزيع t .

ولكن كثيراً ما يحدث في الحياة العملية أن يكون المطلوب إجراء اختبار إحصائي لتساوي أكثر من وسطين حسابيين في الوقت نفسه ، أي لتساوي ثلاثة متوسطات أو أكثر .

فمثلاً :

قد يكون المطلوب إجراء اختبار تساوي متوسطات درجات الطلاب في ثلاث شعب أو أكثر تدرس المقرر نفسه . أو اختبار ما إذا كانت ثلاثة أنواع من القمح أو أكثر تعطي المتوسط نفسه من المحصول ، أو ما إذا كانت عدة أبقار تعطي متوسط الإنتاج نفسه من الحليب ... الخ .

وهكذا نجد أن هذا الموضوع مهم جداً وله تطبيقات كثيرة في مختلف المجالات : في الطب ، في الزراعة ، في الصناعة ، في العلوم الاجتماعية أو الإنسانية ... الخ .

وفي الطريقة التقليدية التي ذكرناها في الفصول السابقة كان يتم اختبار كل

متوسطين على حده ، ولكن هذه الطريقة بالإضافة إلى أنها تحتاج حسابات كثيرة ، فإن احتمال الخطأ منها يكون كبيراً . فمثلاً : إذا كان المطلوب هو إجراء اختبار تساوي 5 متوسطات ، فكان يتم أخذ كل متوسطين على حده ، وبالتالي فإنه يتم إجراء الاختبار 10 مرات ($\binom{5}{2} = 10$) ، فإذا كانت درجة الثقة هي 0.95 (أو مستوى المعنوية هو 0.05) فإن احتمال أن يكون القرار صحيحاً لإجراء الاختبارات العشرة هو $(0.95)^{10}$ أي يساوي 0.599 أي احتمال الخطأ يساوي $1 - 0.599 = 0.401$ وهو احتمال كبير ، بل ويكبر أكثر كلما زاد عدد المتوسطات المراد اختبار تساويها . فإذا كان عدد المتوسطات – مثلاً – 8 متوسطات فإن عدد الاختبارات المطلوب إجراؤها لكل متوسطين على حدة هو 28 اختبار ($\binom{8}{2} = 28$) . وبالتالي فإن احتمال الخطأ في اتخاذ القرار في هذه الحالة يصبح $1-(0.95)^{28}$ أي يساوي 0.762 وهو احتمال كبير جداً مما يقلل كثيراً من كفاءة الطريقة التقليدية بأخذ كل متوسطين على حدة وإجراء الاختبار بينها ، علاوة على أنها طريقة مملة ومتعبة وتحتاج لعمليات حسابية كثيرة .

والأسلوب البديل والمناسب تماماً لإجراء اختبار واحد لكل المتوسطات معاً هو ما يسمى (تحليل التباين) والذي يعتمد على اختبار F أو توزيع F والذي يحلل التباين (أو مجموع المربعات) إلى مكوناته المختلفة .

وفي هذا الكتاب سنشرح بالتفصيل تحليل التباين في اتجاه واحد أو في حالة التصنيف الأحادي والذي يعتبر أبسط أنواع تحليل التباين والذي يتم تصنيف المشاهدات فيه إلى عدة مجموعات على أساس خاصية واحدة (أو عامل واحد)، ودراسة إمكانية وجود تأثير على المتغير من استخدام علاجات مختلفة أو معاملته بطرق مختلفة واختبار تأثير هذه المعاملات على المشاهدات ثم نوضح بإيجاز تحليل التباين في اتجاهين .

(١٣-٢) تحليل التباين في اتجاه واحد (أو التصنيف الأحادي) :

تحليل التباين في اتجاه واحد يعتبر أبسط الأنواع أو الحالات في تحليل التباين حيث يتم تصنيف المشاهدات إلى عدة مجموعات بناءً على خاصية واحدة أو متغير واحد ، ودراسة إمكانية وجود تأثير على المتغير من استخدام علاجات مختلفة ، أو معاملته بطرق مختلفة .

ولاستخدام تحليل التباين لإجراء اختبار تساوي عدة متوسطات نفترض ما يلي :

١- المجتمعات المدروسة مستقلة وعدد K .

٢- هذه المجتمعات تخضع جميعها للتوزيعات الطبيعية بمتوسطات μ_1 ، μ_2 ، .... ، μ_k .

٣- تباين هذه المجتمعات واحد ويساوي σ^2.

٤- العينات السحوبة من هذه المجتمعات عشوائية ومستقلة وحجم كل منها n (ويمكن أن تكون أحجام العينات مختلفة) .

والمطلوب : هو إجراء اختبار إحصائي لفرض العدم التالي :

$$Ho : \mu_1 = \mu_2 = ... = \mu_k$$

أي اختبار أن متوسطات المجتمعات التي عددها K كلها متساوية .

مقابل الفرض البديل H1 :

وجود متوسطين على الأقل غير متساويين أو أن المتوسطات غير متساوية .

فمثلاً :

إذا كان المطلوب هو اختبار ما إذا كانت عدة طرق مختلفة للتعليم تعطى المتوسط نفسه لدرجات الطلاب عند إجراء امتحان واحد لهم ، فإنه يتم أخذ عينات عشوائية من مجموعات الطلاب (والتي نفترض أنها مجتمعات طبيعية

مستقلة) والتي لها الظروف نفسها ، وأن يتم تدريس كل مجموعة بطريقة مختلفة.

فإذا فرضنا أن عدد طرق التدريس (أي عدد المجموعات أو المجتمعات) هو K وأن أحجام العينات العشوائية كلها متساوية وتساوي n ، وإذا رمزنا للمشاهدات (أي درجات الطلاب في هذا المثال) بالرمز Y_{ij} حيث i ترمز للمجتمع (أو الطريقة التعليمية أو ما يسمى المعاملة) ، j ترمز للطالب رقم j فإنه يمكن وضع المشاهدات لهذه التجربة (والتجارب الأخرى المشابهة) في الجدول التالي :

جدول رقم (١)

المجتمع أو المعاملة (أو طريقة التدريس)	المشاهدات (أو درجات الطلاب)			
	1	2	.....	n
1	Y_{11}	Y_{12}	....	Y_{1n}
2	Y_{21}	Y_{22}	....	Y_{2n}
.		.		
.		.		
.		.		
.		.		
.		.		
K	Y_{k1}	Y_{k2}	....	Y_{kn}

فالصف الأول من الجدول يمثل مشاهدات العينة الأولى والمسحوبة من المجتمع الأول (أو المعاملة الأولى) والذين يدرسون بالطريقة الأولى من طرق التعليم ، والصف الثاني يمثل مشاهدات العينة الثانية (أي درجات طلاب العينة الثانية) والذين يدرسون بالطريقة الثانية ، ... وهكذا .

فإن الصف الأخير يمثل مشاهدات العينة الأخيرة (أي درجات طلاب العينة الأخيرة رقم K) والذين يدرسون بالطريقة الأخيرة رقم K .

وهذه المشاهدات Y_{ij} يمكن كتابة كل منها كما يلي :

$$Y_{ij} = \mu_i + e_{ij} \longrightarrow (1)$$

حيث :

μ_i هو الوسط الحسابي للمجتمع i (أو الطريقة رقم i)

e_{ij} هو انحراف المشاهدة رقم j في العينة i عن الوسط الحسابي للمجتمع i

أي عن μ_i (أو هي تمثل الخطأ أو الفرق بين قيمة المشاهدة الفعلية وقيمة الوسط) .

وبما أن لدينا k من المجتمعات فإنه سيكون لدينا عدد k من المتوسطات هي μ_1 ، μ_2 ، ... ، μ_i ، μ_k ، فإن الوسط الحسابي العام لكل هذه المتوسطات هو :

$$\mu = \frac{\sum_{i=1}^{k} \mu_i}{k}$$

وبالتالي فإن متوسط كل من هذه المجتمعات يمكن كتابته بدلالة المتوسط العام كما يلي :

$$\mu i = \mu + \alpha i \longrightarrow (2)$$

حيث αi هي الفرق بين متوسط المجتمع i والمتوسط العام فهي تعبر عن تأثير المجتمع i (أو المعاملة i) .

وبالتالي فإنه تحت الشرط :

$$\sum_{i=1}^{K} \alpha_i = 0 \longrightarrow (3)$$

يمكن التعبير عن المشاهدة Y_{ij} بالنموذج التالي والذي يعتبر النموذج الأساس في هذه الحالة وذلك بدمج العلاقة (2) بالعلاقة (1) كما يلي :

$$Y_{ij} = \mu + \alpha i + e_{ij} \longrightarrow (4)$$

حيث يفترض هذا النموذج أن المشاهدة Y_{ij} تختلف عن المتوسط العام μ بقيمتين:

الأولى : هي تأثير المجتمع i أو المعاملة i بالقيمة $\propto i$.

الثانية : هي الخطأ العشوائي e_{ij} .

الفرضيات الأساسية التي يجب توفرها عن e_{ij} لتطبيق أسلوب تحليل التباين:

١- يمثل e_{ij} الخطأ العشوائي لكل i ، j وهي تخضع للتوزيع الطبيعي .

٢- e_{ij} مستقلة عن بعضها ووسطها الحسابي يساوي صفر أي أن :

$$E(e_{ij}) = 0 \qquad \forall \; i ، j$$

٣- تباين e_{ij} واحد ويساوي σ^2 .

وبناءً على هذه الفرضيات يصبح فرض العدم Ho بتساوي متوسطات المجتمعات التي عددها k مكافئاً لفرض العدم Ho ان جميع قيم $\propto i$ واحدة وتساوي صفر .

أي تكافئ فرض العدم :

$$Ho : \propto_1 = \propto_2 = \cdots = \propto i = \cdots = \propto_K = 0$$

مقابل الفرض البديل H_1 أن واحدة على الأقل من $\propto_i$ لا تساوي صفر .

والفكرة الأساسية في تحليل التباين هي تحليل التباين (الاختلاف) الكلي إلى مكوناته وهما التباين (الاختلاف) الذي يرجع إلى متوسطات الصفوف أو المعاملات ، والتباين (الاختلاف) الذي يرجع إلى الخطأ ، ثم قسمة التباين الأول على التباين الثاني (وهما تقديران مستقلان للتباين σ^2) فتحصل على قيمة F المحسوبة والتي تعتبر الإحصائية في هذه الحالة والتي لها توزيع F بدرجات حرية للبسط يساوي K - 1 (K تمثل عدد الصفوف أو المجتمعات) ودرجات حرية

للمقام تساوي (K . (n - 1] والتي تساوي درجات الحرية لمجموع المربعات الكلي nK - 1 مطروحاً منه درجات حرية المعاملات أو الصفوف [K - 1] .

وهذه الفكرة مبنية على أساس أن مجموع المربعات الكلي يساوي مجموع المربعات لمتوسطات الصفوف (أو المعاملات) مضافاً إليه مجموع مربعات الخطأ. أي أن :

مجموع المربعات الكلي = مجموع مربعات متوسطات الصفوف + مجموع مربعات الخطأ

ويمكن التعبير عن ذلك كما يلي :

$$SST = SSR + SSE \longrightarrow (5)$$

حيث SST هو المجموع الكلي ويأخذ الشكل التالي :

$$SST = \sum_{i=1}^{K} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{..})^2 \longrightarrow (6)$$

حيث $\bar{Y}_{..}$ هو الوسط الحسابي الإجمالي (أو الكلي) للمشاهدات . والذي يأخذ الشكل التالي :

$$\bar{Y}_{..} = \frac{\sum_{i=1}^{K} \sum_{j=1}^{n} Y_{ij}}{nK} \longrightarrow (7)$$

أي هو عبارة عن مجموع جميع المشاهدات المدروسة مقسوماً على عددها .

أي أن :

SST تمثل مجموع مربعات انحرافات المشاهدات عن الوسط الحسابي العام أو الكلي .

SSR هو مجموع مربعات انحرافات متوسطات الصفوف (أو المعاملات) عن الوسط الحسابي العام ويأخذ الشكل التالي (بافتراض أن حجم كل عينة يساوي n):

$$SSR = n\sum_{i=1}^{K} (\bar{Y}_{i.} - \bar{Y}_{..})^2 \longrightarrow (8)$$

حيث $\bar{y}_{i.}$ يمثل الوسط الحسابي للصف (أو المعاملة) i ويأخذ الشكل التالي:

$$\bar{y}_{i.} = \frac{\sum_{j=1}^{n} Y_{ij}}{n} \longrightarrow (9)$$

$$i = 1, 2, \ldots\ldots k$$

أي أن :

SSR تمثل مجموع مربعات انحرافات متوسطات الصفوف (أو المعاملات) عن الوسط الحسابي العام أو الكلي .

SSE هو مجموع مربعات الخطأ ويأخذ الشكل التالي :

$$SSE = \sum_{i=1}^{K} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{i.})^2 \longrightarrow (10)$$

أي أن :

SSE تمثل مجموع مربعات انحرافات المشاهدات عن الأوساط الحسابية للصفوف (أو المعاملات) .

ومن المعادلات (5) ، (6) ، (8) ، (10) فإن مجموع المربعات الكلي SST يمكن أن يكتب بدلالة SSR ، SSE كما يلي :

$$SST = SSR + SSE$$

$$\Rightarrow \sum_{i=1}^{K} \sum_{j=1}^{n} (Y_{ij} - \bar{Y}_{..})^2 = n\sum (\bar{y}_{i.} - \bar{Y}_{..})^2 + \sum_{i=1}^{K} \sum_{j=1}^{n} (Y_{ij} - \bar{y}_{i.})^2$$

ملاحظة مهمة :

عند حساب كل من SST ، SSR ، SSE تستخدم صيغ أخرى أسهل في العمليات الحسابية كما يتم حساب SSE بطرح SSR من SST وبالتالي فإن الصيغ التالية هي التي تستخدم في العمليات الحسابية والتي تعطي النتائج نفسها :

$$SST = \sum_{i=1}^{K} \sum_{j=1}^{n} Y_{ij}^2 - n \, . \, k\bar{Y}_{..}^{\,2} \longrightarrow (11)$$

$$SSR = n\sum_{i=1}^{K} \bar{Y}_{i.}^2 - n \, . \, k\bar{Y}_{..}^{\,2} \longrightarrow (12)$$

$$SSE = SST - SSR \longrightarrow (13)$$

حيث إن SSR تمثل مجموع مربعات الصفوف ، ودرجات الحرية لها تساوي $(k - 1)$ أي عدد الصفوف مطروحاً منه واحد صحيح . وبالتالي فإن تقدير التباين للصفوف هو :

$$S_R^2 = \frac{SSR}{k - 1} \longrightarrow (14)$$

وحيث إن SSE يمثل مجموع مربعات الأخطاء (والذي يساوي مجموع المربعات الكلي مطروحاً منه مجموع مربعات الصفوف) فإن درجات الحرية له تساوي درجات الحرية الكلية مطروحاً منها درجات حرية الصفوف . أي أن درجات الحرية للخطأ D . E تساوي $k(n - 1)$ ويتم إيجادها كما يلي :

$$D.E = nk - 1 - (k - 1)$$

$$= nk - k = k(n - 1)$$

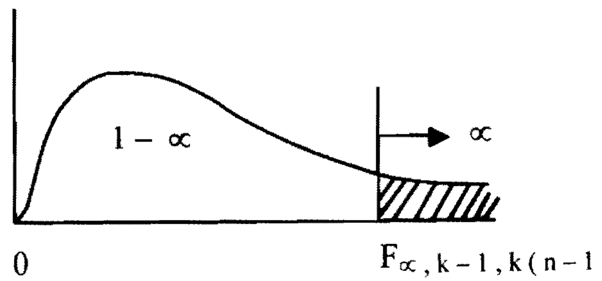
فيكون تقدير تباين الخطأ هو :

$$S_E^2 = \frac{SSE}{k(n-1)} \longrightarrow (15)$$

وبالتالي فإن الإحصائية المناسبة هي المقدار الحاصل بقسمة تباين الصفوف على تباين الخطأ ويكون له توزيع F بدرجات حرية للبسط $k - 1 = V_1$ وللمقام $k(n-1) = V_2$ وذلك بافتراض صحة فرض العدم Ho ، أي أن إحصائية الاختبار هي:

$$F = \frac{S_R^2}{S_E^2} \longrightarrow (16)$$

مع ملاحظة مهمة وهي استخدام اختبار الطرف الأيمن لتوزيع F كما يلي :



ويمكن تلخيص خطوات اختبار تساوي عدة متوسطات باستخدام تحليل التباين فيما يلي :

أولاً : تحديد فرض العدم كما يلي :

$$Ho : \mu_1 = \mu_2 = \ldots = \mu_K$$

ثانياً : تحديد الفرض البديل H_1 :

يوجد متوسطان على الأقل غير متساويين (أو أن المتوسطات غير متساوية).

ثالثاً : حساب إحصائية الاختبار (وتسمى F الحسابية) :

$$F = \frac{S_R^2}{S_E^2}$$

حيث يمكن تلخيص العمليات الحسابية لحساب الإحصائية F في جدول يسمى "جدول تحليل التباين الأحادي" والذي يسمى اختصاراً "جدول ANOVA" كما يلي :

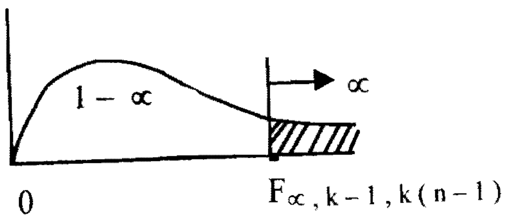
جدول رقم (٢)

(جدول تحليل التباين)

مصدر التغير أو التباين	مجموع المربعات (SS)	درجات الحرية	متوسط المربعات (التباين)	F الحسابية
المعاملات (أو الصفوف)	$SSR = n\sum_{i=1}^{K}\bar{Y}_{i.}^2 - nk\bar{Y}_{..}^2$	K − 1	$S_R^2 = \frac{SSR}{k-1}$	$F = \frac{S_R^2}{S_E^2}$
الخطأ	SSE = SST − SSR	k (n − 1)	$S_E^2 = \frac{SSE}{k(n-1)}$	
الكلي (المجموع)	$SST=\sum_{i=1}^{K}\sum_{j=1}^{n} Y_{ij}^2 - nk\bar{Y}_{..}^2$	nk − 1		

رابعاً : تعيين منطقتي القبول والرفض :

لاحظ أن منطقة الرفض هي المنطقة التي تقع على يمين F_{α, v_1, v_2} حيث ∝ هي مستوى المعنوية المطلوب .



خامساً : اتخاذ القرار الإحصائي :

فإذا كانت قيمة F الحسابية تقع في منطقة القبول (أي إذا كانت أصغر من أو تساوي قيمة F الجدولية) فإنه يتم قبول فرض العدم Ho بتساوي المتوسطات المختلفة. والعكس صحيح ، أي إذا وقعت F الحسابية في منطقة الرفض (أي إذا كانت أكبر من قيمة F الجدولية) فإنه يتم رفض Ho وقبول الفرض البديل H_1.

والمثال التالي يوضح بالتفصيل كيفية إجراء تحليل التباين والعمليات الحسابية اللازمة لذلك .

مثال (١) :

الجدول التالي يمثل نتائج طلاب أربع عينات عشوائية (حجم كل منها 5) مسحوبة من أربع شعب للمقرر نفسه (الدرجة من 10) يقوم بالتدريس لكل منها أستاذ مختلف . وبافتراض أن الشعب الأربع مستقلة عن بعضها ولها توزيع طبيعي بمتوسطات μ_1 ، μ_2 ، μ_3 ، μ_4 ، وتباين متساوي قدره σ^2 .

فإذا كان المطلوب هو اختبار فرض العدم $Ho : \mu_1 = \mu_2 = \mu_3 = \mu_4$ (أي أن المطلوب هو اختبار أن متوسطات درجات الطلاب في الشعب الأربع متساوية) . مقابل الفرض البديل H_1 أن متوسطي شعبتين على الأقل غير متساوي .

استخدم في الاختبار مستوى معنوية 5% ثم 1% .

جدول رقم (٣)

الشعب المختلفة	درجات الطلاب في العينات (المشاهدات)					مجموع الدرجات
الأولى	10	7	8	9	6	40
الثانية	6	6	4	10	9	35
الثالثة	8	7	8	5	8	36
الرابعة	7	8	9	4	9	37
المجموع الكلي						**148**

الحل :

أولاً : تحديد فرض العدم :

$$Ho : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

ثانياً : تحديد الفرض البديل :

يوجد متوسطان على الأقل غير متساويين : H_1 .

ثالثاً : حساب إحصائية الاختبار (F الحسابية) :

$$F = \frac{S_R^2}{S_E^2}$$

وتكون الحسابات لاستخراج قيمة F الحسابية كما يلي ، حيث :

$k = 4$ = عدد الشعب أو العينات (المجتمعات أو المعاملات) .

$n = 5$ = حجم كل من العينات الأربع .

$nk = 20$ = إجمالي عدد المشاهدات في العينات الأربع .

١- حساب المتوسط العام $\bar{Y}_{..}$:

$$\bar{Y}_{..} = \frac{40+35+36+37}{20} = \frac{148}{20}$$

$$\therefore \bar{Y}_{..} = \frac{148}{20} = 7.4 \quad \text{درجة}$$

٢- حساب متوسطات العينات للشعب (المختلفة) :

$$\bar{Y}_{1.} = \frac{40}{5} = 8 \qquad \bar{Y}_{2.} = \frac{35}{5} = 7$$

$$\bar{Y}_{3.} = \frac{36}{5} = 7.2 \qquad \bar{Y}_{4.} = \frac{37}{5} = 7.4$$

٣ - حساب مجموع المربعات الكلي SST كما يلي :

$$SST = \sum_{i=1}^{4} \sum_{j=1}^{5} Y_{ij}^2 - nk\bar{Y}_{..}^{\,2}$$

$$= (6^2 + 9^2 + 8^2 + 7^2 + 10^2 + 9^2 + 10^2 + 4^2 + 6^2 + 6^2 + 8^2 + 5^2 + 8^2 + 7^2 + 8^2 + 9^2 + 4^2 + 9^2 + 8^2 + 7^2) - [4 \times 5\,(7.4)^2]$$

$$\therefore SST = 1156 - 20\,(54.76) = 1156 - 1095.2$$

$$\therefore SST = 60.8$$

٤- حساب مجموع المربعات للصفوف (للعينات) SSR :

$$SSR = n\sum_{i=1}^{4} \bar{Y}_i^2 - nk\bar{Y}_{..}^2$$

$$= 5\,(8^2 + 7^2 + 7.2^2 + 7.4^2) - 4 \times 5\,(7.4)^2$$

$$= 5\,(64 + 49 + 51.84 + 54.76) - 20\,(54.76)$$

$$= 1098 - 1095.2$$

$$\therefore SSR = 2.8$$

٥- حساب مجموع مربعات الخطأ SSE :

$$SSE = SST - SSR = 60.8 - 2.8$$

$$\therefore SSE = 58$$

٦- حساب درجات الحرية كما يلي :

أ- درجات الحرية للمجموع (أو الكلي) :

$$D.T = nk - 1 = 4 \times 5 - 1 = 20 - 1 = \boxed{19.}$$

ب - درجات الحرية للصفوف :

$$D.R = k - 1 = 4 - 1 = \boxed{3.}$$

ج - درجات للخطأ :

$$D.E = k\,(n - 1) = 4\,(5 - 1) = 4 \times 4 = \boxed{16.}$$

٧ - حساب التباينات (S_R^2 ، S_E^2) :

$$S_R^2 = \frac{SSR}{k-1} \quad \Rightarrow \quad S_R^2 = \frac{2.8}{3} = 0.933$$

$$S_E^2 = \frac{SSE}{k(n-1)} \quad \Rightarrow \quad S_E^2 = \frac{58}{16} = 3.625$$

٨- حساب قيمة الإحصائية (F الحسابية) :

$$F = \frac{S_R^2}{S_E^2} \quad \Rightarrow \quad F = \frac{0.933}{3.625} = 0.257$$

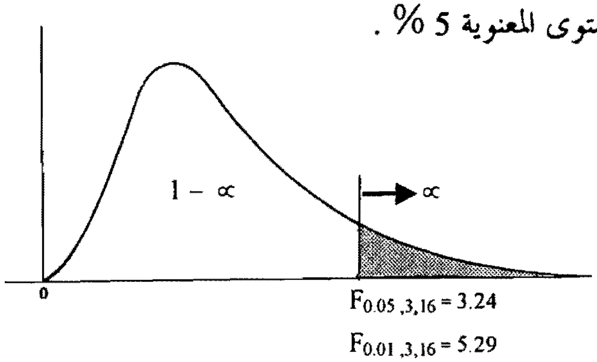
ويمكن تلخيص كل الخطوات السابقة لحساب قيمة الإحصائية F في جدول تحليل التباين التالي :

جدول رقم (٤)

مصدر التغير أو التباين	مجموع المربعات	درجات الحرية	متوسط المربعات (التباين)	F الحسابية
المعاملات (أو الصفوف)	SSR = 2.8	3	$S_R^2 = \frac{2.8}{3} = 0.93$	$F = \frac{S_R^2}{S_E^2} = \frac{0.93}{3.63} = 0.26$
الخطأ	SSE = 58	16	$S_E^2 = \frac{58}{16} = 3.63$	
الكلي (المجموع)	SST = 60.8	19		

رابعاً : تعيين منطقتي القبول والرفض :

لاحظ أن منطقة الرفض هي المنطقة التي تقع على يمين $F_{0.01\,,\,3\,,\,16} = 5.29$ وذلك في حالة مستوى معنوية 1 % ، وتقع على يمين $F_{0.05\,,\,3\,,\,16} = 3.24$ في حالة مستوى المعنوية 5 % .



خامساً : اتخاذ القرار الإحصائي :

وحيث إن قيمة F الحسابية تبلغ 0.26 تقع في منطقة القبول حيث إنها أصغر من قيمة F الجدولية سواءً عند 1% أو حتى 5% فإن القرار هو : قبول فرض العدم Ho بتساوي متوسطات درجات الطلاب في الشعب الأربع وذلك بمستوى معنوية 1% أو 5% .

ملاحظات مهمة :

١- يتضح من المثال البسيط السابق الجهد الكبير والعمليات الحسابية التي يحتاجها تحليل التباين ، لذلك يستخدم الحاسب الآلي لإجراء كل هذه الحسابات في زمن قصير جداً خاصة مع وجود برامج جاهزة كثيرة لتحليل التباين كما في SAS ، SPSS .

٢ - لاحظنا في الشرح –والمثال– السابق أننا افترضنا أن أحجام العينات متساوية وكل منها يساوي n . ولكن لا توجد أي مشكلة على الإطلاق إذا كانت أحجام العينات غير متساوية ، كل ما في الأمر هو مراعاة ذلك عند حساب مجاميع المربعات المختلفة والتي تأخذ الأشكال التالية في هذه الحالة :

مجاميع المربعات في حالة عدم تساوي العينات :

$$SST = \sum_{i=1}^{K} \sum_{j=1}^{n_i} Y_{ij}^2 - N\bar{Y}_{..}^2$$

$$SSR = \sum_{i=1}^{K} n_i \bar{Y}_{i.}^2 - N\bar{Y}_{..}^2$$

$$SSE = SST - SSR$$

حيث : $N = n_1 + n_1 + \ldots + n_i + n_k = \sum_{i=1}^{K} n_i$

حيث تشير n_i إلى حجم العينة المسحوبة من المجتمع (أو المعاملة) i .

والمثال التالي يوضح كيفية تحليل التباين في حالة عينات غير متساوية الحجم.

مثال (٢) :

الجدول التالي يمثل عدد الأطفال في ثلاث عينات من الأسر من ثلاث مناطق بإحدى المدن . والمطلوب استخدام تحليل التباين واختبار F لمعرفة ما إذا كانت توجد فروق بين متوسطات أعداد الأطفال في المناطق الثلاث (متوسطات أحجام الأسر) وذلك بمستوى معنوية 5% (وذلك بافتراض أن عدد الأطفال بالأسر في المناطق المختلفة لها توزيعات طبيعية بتباين واحد يساوي $\sigma^2 > 0$) .

جدول رقم (٥)

المناطق	عدد الأطفال بالأسر (المشاهدات)					مجاميع الصفوف
الأولى			6	4	5	15
الثانية	5	4	7	8	6	30
الثالثة		3	6	5	7	21
المجموع الكلي						**66**

الحل :

أولاً : تحديد فرض العدم : متوسطات عدد الأطفال متساوي في المناطق الثلاث. أي أن : $Ho: \mu_1 = \mu_2 = \mu_3$

ثانياً : تحديد الفرض البديل :

H_1 : يوجد متوسطان على الأقل غير متساويين

ثالثاً : حساب إحصائية الاختبار : $F = \frac{S_R^2}{S_E^2}$

وتكون تنظيم الحسابات (بصورة مختصرة كما يلي) ، حيث :

عدد المناطق (أو المجتمعات أو المعاملات)= $k = 3$.

$N = n1 + n2 + n3 = 3 + 5 + 4 = 12$ ، $n_1 = 3$ ، $n_2 = 5$ ، $n_3 = 4$.

١– حساب المتوسط العام (الكلي) :

$$\overline{Y}_{..} = \frac{15+30+21}{12} = \frac{66}{12} = \boxed{5.5}$$

٢– حساب متوسطات الصفوف (أو العينات) :

$$\overline{Y}_{1.} = \frac{15}{3} = \cancel{3}\ 5 \qquad \overline{Y}_{2.} = \frac{30}{5} = 6 \qquad \overline{Y}_{3.} = \frac{21}{4} = 5.25$$

٣– حساب مجموع المربعات الكلي SST :

$$SST = \sum_{i=1}^{3} \sum_{j=1}^{ni} Y_{ij}^2 - N\overline{Y}_{..}^{\,2}$$

$$= (5^2 + 4^2 + 6^2 + 6^2 + 8^2 + 7^2 + 4^2 + 5^2 + 7^2 + 5^2 + 6^2 + 3^2) - 12\,(5.5)^2$$

$$= 386 - 12\,(30.25) = 386 - 363 = \boxed{23}$$

٤– حساب مجموع المربعات للصفوف (للعينات) SSR :

$$SSR = \sum_{i=1}^{3} n_i \overline{Y}_i^2 - N\overline{Y}^2_{..}$$

$$= [3\,(5)^2 + 5\,(6)^2 + 4\,(5.25)^2] - 12\,(5.5)^2$$

$$= [75 + 180 + 110.25] - 363 = 365.25 - 363 = \boxed{2.25}$$

٥– حساب مجموع مربعات الخطأ SSE :

$$SSE = SST - SSR$$

$$= 23 - 2.25 = \boxed{20.75}$$

٦– حساب درجات الحرية :

أ – درجات الحرية للمجموع (الكلي) :

$$D.T = N - 1 = 12 - 1 = \boxed{11}$$

ب – درجات الحرية للصفوف :

$$D.R = k - 1 = 3 - 1 = \boxed{2}$$

ج – درجات الحرية للخطأ :

$$D.E = D.T - D.R = 11 - 2 = \boxed{9}$$

٧ – حساب التباينات (S_R^2 ، S_E^2) :

$$S_R^2 = \frac{SSR}{k-1} = \frac{2.25}{2} = \boxed{1.13}$$

$$S_E^2 = \frac{SSE}{N-k} = \frac{20.75}{9} = \boxed{2.31}$$

٨ – حساب قيمة إحصائية الاختبار F :

$$F = \frac{S_R^2}{S_E^2} = \frac{1.13}{2.31} = \boxed{0.49}$$

ويكون جدول تحليل التباين كما يلي :

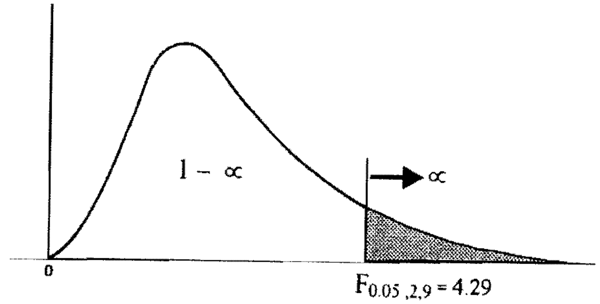
جدول رقم (٦)

مصدر التغير أو التباين	مجموع المربعات	درجات الحرية	متوسط المربعات (التباين)	F الحسابية
الصفوف	SSR = 2.25	2	$S_R^2 = \frac{2.25}{2} = 1.13$	$F = \frac{S_R^2}{S_E^2}$
الخطأ	SSE = 20.75	9	$S_E^2 = \frac{20.75}{9} = 2.31$	$= \frac{1.13}{2.31}$
الكلي	SST = 23	11		$= 0.49$

رابعاً : تعيين حدود منطقتي القبول والرفض :

لاحظ أن قيمة F المعيارية (الجدولية) هي : $F_{0.05,2,9} = 4.26$

فتكون منطقة الرفض على يمين القيمة 4.26 كما يظهر في الشكل التالي :



خامساً : اتخاذ القرار الإحصائي :

وحيث إن قيمة الإحصائية F المحسوبة (والتي تساوي 0.49) تقع في منطقة القبول (أقل من القيمة الجدولية) فإن القرار هو: "قبول فرض العدم Ho بتساوي متوسطات أحجام الأسر في المناطق الثلاث وذلك بمستوى معنوية 5%" .

(١٣-٣) تحليل التباين في اتجاهين :

بحثنا فيما سبق أبسط أنواع تحليل التباين وهو تحليل التباين في اتجاه واحد (التصنيف الأحادي) ، وسنتطرق الآن - باختصار - لتحليل التباين في اتجاهين (التصنيف الثنائي وفقاً لخاصيتين أو عاملين) .

فمثلاً إذا رغب أستاذ مقرر الإحصاء في كلية العلوم الإدارية معرفة تأثير المرجع أو الكتاب المعتمد على التحصيل العلمي لطلابه في أحد المقررات الدراسية ، ولنفترض أن هناك ثلاثة كتب مقترحة كمرجع لهذا المقرر ، وأن طلاب ذلك الأستاذ موزعون على أربعة تخصصات مختلفة (إدارة ، اقتصاد ، محاسبة ، قانون) . ويرغب أستاذ المقرر التخلص من (حذف) الاختلاف بين التحصيل العلمي للطلاب الناتج بسبب اختلاف التخصصات بحيث يتبقى الاختلاف الناتج بسبب اختلاف الكتب المستخدمة كمرجع لذلك المقرر فقط. ونسمي - عادة - أصناف الخاصية أو العامل الأول بالمعاملات أو المعالجات ، أما أصناف الخاصية أو العامل الثاني فتسمى بالقطاعات . وفي المثال السابق تعتبر المعاملات (جمع معاملة) هي الكتب المقترحة كمرجع للمقرر وعددها ثلاثة (k = 3) ، أما القطاعات فهي التخصصات المختلفة وعددها أربعة (n = 4) .

وحقيقة فإن التقسيم إلى قطاعات يعتبر طريقة جيدة وسهلة لمعرفة الاختلافات الناتجة عن كل عامل ، وبالتالي معرفة الأثر الحقيقي الناتج بسبب المعاملات (المعالجات) محل البحث والدراسة . فمثلاً في التجارب الزراعية (الحقلية) يتكون كل قطاع من قطع (أحواض) مربعة ومتجاورة من الأرض الزراعية بحيث يزرع فيها أنواع مختلفة من البذور (المعاملات) لمعرفة التفاوت بين هذه الأنواع من البذور . وفي مثال المراجع الدراسية السابق يكون لدينا ثلاث معاملات (معالجات) ، وأربعة قطاعات وبالتالي يكون لدينا 12 (4 × 3) خلية أو وحدة تجريبية . ونقوم بتوزيع المعاملات (الكتب) الثلاث بشكل عشوائي داخل كل قطاع (التخصصات وعددها أربعة) ، فإذا فرضنا أن الكتب المقترحة كمرجع للمادة هي R_1 ، R_2 ، R_3 فإن أحد التوزيعات العشوائية الممكنة موضح في الجدول التالي :

جدول رقم (٧)

القطاع (١) إدارة	القطاع (٢) اقتصاد	القطاع (٣) محاسبة	القطاع (٤) قانون
R_2	R_1	R_1	R_3
R_1	R_3	R_2	R_1
R_3	R_2	R_3	R_2

نلاحظ من جدول رقم (٧) أن كل معاملة تظهر في كل قطاع مرة واحدة، وأن جميع المعاملات (المعالجات) تظهر في كل قطاع ، كما أن توزيع المعاملات على القطاعات تم بطريقه عشوائية . فالجدول رقم (٧) يعطي تصميم بحسب القطاعات وبشكل كامل أو تام (يغطي جميع المعاملات والقطاعات) وأيضاً بشكل عشوائي (يتم توزيع المعاملات أو المعالجات على القطاعات بشكل عشوائي) ، ولذا يسمى هذا التصميم بـ (تصميم قطاعي عشوائي تام أو كامل) أو (تصميم القطاعات العشوائية التامة) ، وكذلك فإن أسلوب تحليل

التباين المستخدم في هذه الحالة يسمى (تحليل التباين في اتجاهين) حيث نقوم بتصنيف البيانات بحسب الظاهرتين أو الخاصتين المدروستين . وعموماً يفضل استخدام الجدول التالي رقم (٨) لعرض البيانات التي نحصل عليها بالتصميم القطاعي العشوائي التام .

جدول رقم (٨)

القطاعات	المعاملات (المعالجات) 1	2	k	المجموع	المتوسط
1	Y_{11}	Y_{12}	Y_{1k}	$T_{1.}$	$\bar{Y}_{1.}$
2	Y_{21}	Y_{22}	Y_{2k}	$T_{2.}$	$\bar{Y}_{2.}$
⋮		⋮		⋮	⋮
n	Y_{n1}	Y_{n2}	Y_{nk}	$T_{n.}$	$\bar{Y}_{n.}$
المجموع	$T_{.1}$	$T_{.2}$	$T_{.k}$		
المتوسط	$\bar{Y}_{.1}$	$\bar{Y}_{.2}$	$\bar{Y}_{.k}$		

ونوضح معنى الرموز الواردة في الجدول رقم (٨) فيا يلي :

عدد المعاملات (المعالجات) = k

عدد القطاعات = n

مجموع القطاع رقم i = $T_{i.} = \sum_{j=1}^{k} Y_{ij}$ حيث $i = 1, 2, \ldots, n$

مجموع المعاملة رقم j = $T_{.j} = \sum_{i=1}^{n} Y_{ij}$ حيث $j = 1, 2, \ldots, k$

متوسط القطاع رقم i = $\bar{Y}_{i.} = \dfrac{T_{i.}}{k}$ حيث $i = 1, 2, \ldots, n$

متوسط المعاملة رقم j = $\bar{Y}_{.j} = \dfrac{T_{i.}}{k}$ حيث $j = 1, 2, \ldots, k$

المجموع الكلي = $T_{..} = \sum_{j=1}^{k} T_{.j} = \sum_{i=1}^{n} T_{i.}$

لاحظ أن المجموع الكلي $(T_{..})$ يمكن الحصول عليه بجمع مجاميع الصفوف $(T_{i.})$ أو مجاميع الأعمدة $(T_{.j})$.

ويستخدم تحليل التباين في اتجاهين لاختبار فرض العدم (Ho) ، أن متوسطات المعاملات (المعالجات) متساوية أي أن :

$$Ho : \mu_1 = \mu_2 = \ldots\ldots = \mu_k$$

وفرض العدم (Ho) يعني عدم وجود فروقات معنوية بين المعاملات (المعالجات)، وأن الفروقات الموجودة ترجع إلى اختلاف القطاعات أو ترجع إلى العشوائية (فروقات طفيفة لا اعتبار لها) . فمثلاً في مثال المراجع الدراسية يكون فرض العدم كما يلي :

$$Ho : \mu_1 = \mu_2 = \mu_3$$

فإذا كان هذا الفرض صحيحاً ، نستنتج أن اختلاف المراجع (الكتب) الدراسية ليس له تأثير على التحصيل العلمي للطلاب ، وأن الفروقات الموجودة لدى الطلاب ترجع إلى اختلاف تخصصات الطلاب (القطاعات) أو ترجع إلى العشوائية .

أما الفرض البديل H_1 فهو :

H_1 : يوجد على الأقل متوسطان غير متساويين

فإذا رفضنا Ho ، نقبل H_1 وهذا يعني وجود فروقات معنوية بين المعاملات (المعالجات) ، وبالتالي فإن هذه المعاملات (المعالجات) لها تأثير معنوي على العينة المدروسة.

ونستخدم في تحليل التباين في اتجاهين النموذج الرياضي التالي :

$$Y_{ij} = \mu + \beta_i + \alpha_j + e_{ij}$$

حيث :

Y_{ij} = القيمة المشاهدة في الخلية C_{ij}

μ = المتوسط العام (مجهول)

β_i = يمثل تأثير القطاع i

α_j = يمثل تأثير المعاملة (المعالجة) j

e_{ij} = الخطأ العشوائي للمشاهدة Y_{ij}

ونفترض في هذا النموذج افتراضات إحصائية مشابهة للافتراضات التي ذكرناها في نموذج تحليل التباين في اتجاه واحد .

وكما أوضحنا خلال شرحنا لتحليل التباين في اتجاه واحد ، فإنه يمكن وبالأسلوب نفسه الحصول على جدول تحليل التباين (جدول Anova) كما يلي :

جدول رقم (٩)

مصدر التغير أو التباين	مجموع المربعات	درجات الحرية	متوسط المربعات (التباين)	F الحسابية
المعاملات (المعالجات)	SSR	k - 1	$S_R^2 = \frac{SSR}{k-1}$	$F = \frac{S_R^2}{S_E^2}$
القطاعات	SSB	n - 1	$S_B^2 = \frac{SSB}{n-1}$	
الخطأ	SSE	(n - 1) (k - 1)	$S_E^2 = \frac{SSE}{(n-1)(k-1)}$	
المجموع (الكلي)	SST	nk - 1	———	

حيث :

$$SSR = \frac{1}{n}\sum_{j=1}^{k} T_{.j}^2 - nk\overline{Y}_{..}^2$$

$$SSB = \frac{1}{k}\sum_{i=1}^{n} T_{i.}^2 - nk\overline{Y}_{..}^2$$

$$SST = \sum_{j=1}^{k}\sum_{i=1}^{n} Y_{ij}^2 - nk\overline{Y}_{..}^2$$

$$SSE = SST - SSR - SSB$$

والعمود الأخير في الجدول رقم (٩) يعطي قيمة F الحسابية وبمقارنة قيمة F الحسابية مع F الجدولية نقبل أو نرفض Ho . وعموماً فإن خطوات اختبار تساوي متوسطات ($Ho : \mu_1 = \mu_2 = ... = \mu_k$) باستخدام تحليل التباين في اتجاهين مشابهة للخطوات التي درسناها في تحليل التباين في اتجاه واحد مع بعض الاختلافات اليسيرة في طريقة حساب إحصائية الاختبار (F الحسابية) ، ودرجات الحرية ، ونوضح ذلك في المثال التالي .

مثال (٣) :

إشارة إلى مثال المراجع الدراسية المعتمدة (المذكورة في مقدمة الحديث عن تحليل التباين في اتجاهين) ، فإن الجدول التالي يعطي درجات الطلاب (الدرجة الكبرى 100) التي حصل عليها أستاذ الإحصاء بعد إجراء التجربة على مجموعة من طلابه .

جدول رقم (١٠)

التخصصات	المراجع أو الكتب المعتمدة			المجموع	المتوسط
	R_1	R_2	R_3		
إدارة (M_1)	65	70	90	$T_{1.} = 225$	$\overline{Y}_{1.} = 75$
اقتصاد (M_2)	80	75	85	$T_{2.} = 240$	$\overline{Y}_{2.} = 80$
محاسبة (M_3)	75	65	85	$T_{3.} = 225$	$\overline{Y}_{3.} = 75$
قانون (M_4)	60	50	70	$T_{4.} = 180$	$\overline{Y}_{4.} = 60$
المجموع	$T_{.1} = 280$	$T_{.2} = 260$	$T_{.3} = 330$	$T_{..} = 870$	———
المتوسط	$\overline{Y}_{.1} = 70$	$\overline{Y}_{.2} = 65$	$\overline{Y}_{.3} = 82.5$	———	$\overline{Y}_{..} = 72.5$

وبافتراض أن درجات الطلاب المعروضة في الجدول رقم (١٠) تمثل عينات عشوائية مستقلة ذات حجم 1 مسحوبة من 12 مجتمع إحصائي (4 تخصصات × 3 مراجع = 12) ، وبافتراض أن هذه المجتمعات الإحصائية لها توزيع طبيعي بمتوسط μ_{ij} (حيث : $i = 1 , 2 , 3 , 4$ ، $j = 1 , 2 , 3$) وتباين متساوي قدره $\sigma^2 > 0$.

والمطلوب اختبار فرض العدم : $Ho : \mu_1 = \mu_2 = \mu_3$

مقابل الفرض البديل : يوجد متوسطان على الأقل غير متساويين : H_1

أي أن المطلوب هو اختبار الفرض بأن متوسطات المراجع (المعالجات) متساوية مقابل الفرض بأن متوسطي مرجعين (معالجتين) على الأقل غير متساوي . استخدم مستوى معنوية 1 % ثم 10 % .

أولاً : تحديد فرض العدم :

$$Ho : \mu_1 = \mu_2 = \mu_3$$

ثانياً : تحديد الفرض البديل :

يوجد متوسطان على الأقل غير متساويين : H_1 .

ثالثاً : حساب إحصائية الاختبار (F الحسابية)

عدد المراجع المعتمدة (المعاملات أو المعالجات) = k = 3

عدد التخصصات (القطاعات) = n = 4

إجمالي عدد المشاهدات = n . k = 12

$$SST = \sum_{j=1}^{k} \sum_{i=1}^{n} Y_{ij}^2 - nk\overline{Y}_{..}^2$$

$$\Rightarrow SST = (65^2 + 70^2 + 90^2 + 80^2 + 75^2 + 85^2 + 75^2 + 65^2 + 85^2 + 60^2 + 50^2 + 70^2) - (4 \times 3 \times 72.5^2)$$

$$\Rightarrow SST = 64550 - 63075 = 1475$$

$$، SSR = \frac{1}{n}\sum_{j=1}^{k} T_{.j}^2 - nk\overline{Y}_{..}^2$$

$$\Rightarrow SSR = \frac{1}{4}(280^2 + 260^2 + 330^2) - 63075$$

$$\Rightarrow SSR = 63725 - 63075 = 650$$

$$، SSB = \frac{1}{k}\sum_{i=1}^{n} T_{i.}^2 - nk\overline{Y}_{..}^2$$

$$\Rightarrow SSB = \frac{1}{3}(225^2 + 240^2 + 225^2 + 180^2) - 63075$$

$$\Rightarrow SSB = 63750 - 63075 = 675$$

$$، SSE = SST - SSR - SSB$$

$$\Rightarrow SSE = 1475 - 650 - 675 = 150$$

$$S_R^2 = \frac{SSR}{k-1} = \frac{650}{2} = 325 \; , \; S_B^2 = \frac{SSB}{n-1} = \frac{675}{3} = 225$$

$$, S_E^2 = \frac{SSE}{(n-1)(k-1)} = \frac{150}{6} = 25$$

$$F = \frac{S_R^2}{S_E^2} = \frac{325}{25} = 13$$

ويمكن تلخيص حسابات الحصول على قيمة F الحسابية في جدول تحليل التباين التالي:

جدول رقم (١١)

مصدر التغير أو التباين	مجموع المربعات	درجات الحرية	متوسط المربعات (التباين)	F الحسابية
المراجع (المعالجات)	SSR = 650	k - 1 = 2	$S_R^2 = 325$	$F = \frac{325}{25} = 13$
التخصصات (القطاعات)	SSB = 675	n - 1 = 3	$S_B^2 = 225$	
الخطأ	SSE = 150	(n - 1) (k - 1) = 6	$S_E^2 = 25$	
المجموع (الكلي)	SST=1475	nk - 1 = 11	ــــــــ	

رابعاً : تحديد منطقتي القبول والرفض :

أ– عند مستوى معنوية 1 % ($\propto = 0.01$) :

منطقة الرفض تقع على يمين F الجدولية وهي :

$$F_{\propto, k-1, (n-1)(k-1)} = F_{0.01, 2, 6} = 10.92$$

ب– عند مستوى معنوية 10 % ($\propto = 0.10$) :

منطقة الرفض تقع على يمين F الجدولية وهي :

$$F_{0.10, 2, 6} = 3.46$$

خامساً : اتخاذ القرار الإحصائي :

أ– عند مستوى معنوية 1 % ($\propto = 0.01$) :

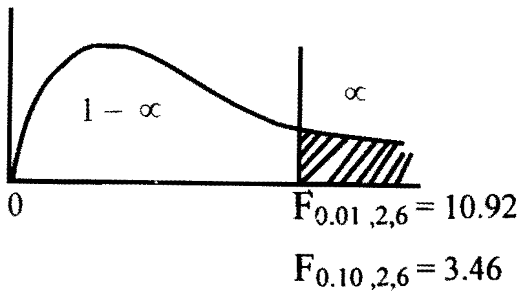
$$\because F = 13 > 10.92 = F_{0.01, 2, 6}$$

∴ نرفض Ho ونقبل H_1 ، ونقول أن متوسط درجات الطلاب للمراجع الثلاثة غير متساوي وذلك بمستوى معنوية 1 % ، وهذا يعني أن المرجع الدراسي المعتمد لمقرر الإحصاء له تأثير معنوي على التحصيل العلمي للطلاب .

ب– عند مستوى معنوية 10 % :

$$\because F = 13 > 3.46 = F_{0.10, 2, 6}$$

∴ نرفض Ho ونقبل H_1 وذلك بمستوى معنوية 10 % ($\propto = 0.10$).



(لاحظ أننا رفضنا Ho عند مستوى 1 % لذا فإن زيادة مستوى المعنوية إلى 10 % لن يؤثر على قرار الرفض لأن منطقة الرفض ستكبر وبالتالي فإن رفض Ho سيكون أكثر تأكيداً، أما لو نقصنا مستوى المعنوية إلى 0.5 % مثلاً، فهنا قد يتغير القرار ونقبل Ho ويمكن للطالب أن يجرب ذلك بنفسه) .

تمارين الفصل الثالث عشر :

السؤال الأول :

الجدول التالي يظهر درجات طلاب ثلاث شعب في أحد المقررات (لمدرسين مختلفين) :

المدرس	درجات الطلاب								
الأول	70	92	87	60	64	81	78	57	69
الثاني	75	86	50	63	54	88	70	65	48
الثالث	85	92	77	66	52	40	80	64	70

والمطلوب :

اختبار ما إذا كانت متوسطات درجات الطلاب في الشعب الثلاث تساوية أم لا . وذلك باستخدام مستوى المعنوية 5% (بافتراض أن درجات طلاب تتبع التوزيع الطبيعي ولها التباين نفسه) .

سؤال الثاني :

الجدول التالي يظهر أعمار ثلاث مجموعات من العمال اختيرت بطريقة شوائية من ثلاثة مصانع مختلفة :

صنع	أعمار العمال									
أول	35	28	33	40	38	42	50	26	36	30
ثاني	40	37	31	25	39	52	43			
الث	41	23	56	32	35					

والمطلوب :

اختبار ما إذا كانت متوسطات أعمار العمال في المصانع الثلاثة متساوية أم وذلك بمستوى المعنوية 1% (بافتراض أن أعمار العمال في المصانع تتبع زيع الطبيعي ولها التباين نفسه) .

السؤال الثالث :

لاختبار مدى فاعلية ثلاثة أنواع من السماد على محصول القمح أخذت 8 قطع من الأرض متجانسة تماماً واستعمل لكل منها نوع من ا لأسمدة A ، B ، C وكانت النتائج كما في الجدول التالي :

نوع السماد	الإنتاج من القمح					
A	420	395	380	402	385	400
B	430	410	390	412	398	408
C	380	366	402	375	406	372

والمطلوب :

اختبار تساوي فاعلية الأسمدة الثلاثة بمستوى المعنوية 5% (بافتراض أن المجتمعات الثلاثة لها توزيعات طبيعية ولها التباين نفسه) .

السؤال الرابع :

الجدول التالي يعطي درجات الطلاب في أحد المقررات باستخدام أربعة مراجع مختلفة (R_1 ، R_2 ، R_3 ، R_4) في تخصصات الإدارة والاقتصاد والمحاسبة:

القسم	المراجع المختلفة			
	R_1	R_2	R_3	R_4
الإدارة	72	75	84	92
الاقتصاد	82	78	90	91
المحاسبة	77	80	83	95

وبافتراض أن درجات الطلاب تمثل عينات عشوائية مستقلة مسحوبة من مجتمعات طبيعية بمتوسطات μ_i ولها التباين نفسه فإن المطلوب اختبار فرض العدم :

$$Ho : \mu_1 = \mu_2 = \mu_3 = \mu_4$$

أي أن متوسط المراجع أو المعالجات متساوية مقابل الفرض البديل .

يوجد متوسطان على الأقل غير متساويين : H_1

السؤال الخامس :

في السؤال الرابع السابق اختبر الفرض العدمي أن متوسطات الأقسام الثلاثة متساوية مقابل الفرض البديل أن متوسطين على الأقل غير متساويين .

السؤال السادس :

إذا كان عدد الأطفال في ثلاث عينات من الأسر مسحوبة من ثلاث مناطق مختلفة ، وكل عينة تتكون من خمس اسر ، كما في الجدول التالي :

المنطقة	عدد الأطفال				
أ	8	6	10	7	5
ب	6	4	3	5	4
ج	5	7	4	5	3

استخدم تحليل التباين واختبار F لاختبار ما إذا كانت متوسطات أعداد الأطفال بالأسر في المناطق الثلاث متساوية أم لا ، وذلك بمستوى معنوية 5%.

السؤال السابع :

إذا كان يقوم بالتدريس لثلاث شعب لأحد المقررات ثلاثة مدرسين مختلفين ، وكانت درجات الطلاب في الامتحانات النهائية لكل مدرس كما في الجدول التالي :

المدرس	درجات الطلاب									
الأول	60	70	85	90	55	92	86	74	78	50
الثاني	75	82	93	95	72	60	65	80		
الثالث	80	95	96	70	84	90				

اختبر ما إذا كانت متوسطات الدرجات للشعب الثلاث متساوية أم لا ، وذلك بمستوى معنوية 5 % .

السؤال الثامن :

الجدول التالي يمثل الزيادة في وزن 15 خروف تعيش في الظروف نفسها بعد إعطاء كل 5 منها (اختيرت عشوائياً) نظاماً غذائياً مختلفاً ، وذلك بعد عدة أسابيع (الزيادة في الوزن مقاسه بالكيلو غرام) :

نظام الغذاء	الزيادة في وزن الخراف بالكيلو				
A	5	4	3	4	2
B	6	4	4	5	3
C	4	5	2	3	2

اختبر ما إذا كانت متوسطات الزيادة في الوزن متساوية للأنظمة الغذائية المختلفة ، استخدم بمستوى معنوية 1 % .

السؤال التاسع :

الجدول التالي يمثل أسعار ثلاث عينات من الشقق في ثلاث مناطق مختلفة (الأسعار بآلاف بالريالات) :

المنطقة	أسعار الشقق بآلاف الريالات					
الأولى	60	80	90	65	75	70
الثانية	100	95	85	90		
الثالثة	50	55	60	70	75	

فهل توجد اختلافات بين متوسطات أسعار الشقق في المناطق الثلاث ؟