٨ المقارنة بين متوسطين

Comparison between two means

١- طريقة العينتين المستقلتين

٢- طريقة العينة المزدوجة

٣- طبيعة الفرق بين الطريقتين

فى كثير من التجارب يكون الفرق بين تأثير المعاملات وليس أثر المعاملات نفسها هو الأكثر أهمية بالنسبة للمجرب. فمثلاً إذا كان هناك دواء معين يقلل من الإصابة بمرض ما أو يؤدى إلى إحداث تغيير فسيولوجى معين كخفض لمستوى مادة معينة فى الدم أو لزيادة معدل الاستفادة من الغذاء ... الخ؛ فإن هناك طريقتين لإجراء مثل هذه المقارنات.

٨-١ طريقة العينتين المستقلتين Independent samples method

وتستخدم هذه الطريقة عندما يراد مقارنة الفرق بين متوسطى عشيرتين (معاملتين) كل منهما قد تم الحصول عليها بطريقة عشوائية من داخل العشيرة كما يحدث عند مقارنة نظامين لتغذية العجول حديثة الولادة أو لتسمين الحملان بعد فطامها. وهذا النوع من التجريب هو الأكثر استخداما فى التجارب حيث لا يوجد أساس لوجود تشابه بين فردين من أفراد العشيرة أكثر من بقية الأفراد (كما سيتضح فى الفصل التالى). وفى هذه الحالة فإن التجربة تتم عن طريق ما يعرف بمقارنة المجموعتين group comparison حيث توزع المعاملتان عشوائيا على أفراد العينة. ثم يقارن ويختبر الفرق فى النهاية بين متوسطى المجموعتين.

٨-٢ طريقة العينة المزدوجة Paired sample method

وفى هذا النوع من العينات تختار أزواج من الأفراد المتشابهة فى كثير من الجوانب مثال ذلك التوائم من نفس البطن والجنس فى الأرانب أو الأبقار أو نصفى ورقة أو ورقتان فى نبات واحد أو أصيصان فى نفس الموقع ... الخ.

وفى هذا المجال فإن أحد أفراد كل زوج يتلقى عشوائياً أحد المعاملتين ويتلقى الفرد الآخر المعاملة الأخرى.

وفى بعض الحالات التى تستخدم فيها هذه العينات بكثرة يستخدم نفس الفرد فى مرحلتين مختلفتين مثل قياس ضغط الدم للفرد قبل وبعد إعطائه دواء معيناً self-pairing والفرق بين القيم على أفراد الزوج الواحد يعبر عن الفرق بين المعاملات، وتؤخذ هذه الفروق على أنها المشاهدات فى التجربة، وتكون مجموعة الفروق بين أفراد الأزواج المختلفة مجموعة البيانات التى يتم تحليلها لاستخلاص المقارنة المطلوبة بين المعاملتين. وبذلك يكون عدد المشاهدات فى التجربة مساويا لعدد الفروق أى n وليس بعدد الأفراد كما هو الحال فى النوع الأول.

٨-٣ طبيعة الفرق بين الطريقتين

إن الفرق بين الطريقتين يتوقف على طبيعة التباين (V) للفرق بين متغيرين. فالتباين للفرق بين المتغيرين X_1، X_2 يتبع المعادلة العامة التالية:

$$V(X_1 - X_2) = V(X_1) + V(X_2) - 2\,Cov(X_1, X_2) \qquad (٨-١)$$

ففي حالة كون القيم مستقلة عن بعضها البعض أى أن X_1، X_2 متغيران عشوائيان. أى فى حالة العينات المستقلة فإن تباين الفرق يساوى مجموع التباين لكل من X_1، X_2 حيث إنه يمكن إثبات هذه العلاقة كالتالى:

$$V(X_1 - X_2) = E[(X_1 - X_2) - (\mu_1 - \mu_2)]^2 \qquad (٨-٢)$$

الرمز E فى المعادلة (٨-٢) هو اختصار لكلمة القيمة المتوقعة expected value وهى من الناحية الرياضية تعنى القيمة المتوقعة للكمية التى بين الأقواس، وبما أن القيمة المتوقعة لأى متغير هى متوسط العشيرة المأخوذ منها هذا المتغير، فإن E تعنى متوسط الكمية داخل القوس.

المعادلة (٨-٢) يمكن إعادة كتابتها كالتالى:

$$\begin{aligned} V(X_1 - X_2) &= E[(X_1 - X_2) - (\mu_1 - \mu_2)]^2 \\ &= E[(X_1 - \mu_1) - (X_2 - \mu_2)]^2 \\ &= E(X_1 - \mu_1)^2 + E(X_2 - \mu_2)^2 - 2E[(X_1 - \mu_1)(X_2 - \mu_2)] \end{aligned}$$

ومن التوقعات الرياضية فإن القيمة المتوقعة للحدين الأول والثانى من الجانب الأيمن للمعادلة السابقة عبارة عن:

$$E(X_2 - \mu_2)^2 = \sigma_2^2 \ ، \ E(X_1 - \mu_1)^2 = \sigma_1^2$$

أما الحد الثالث فهو عبارة عن:

$$E(X_1 - \mu_1)(X_2 - \mu_2) = Cov(X_1, X_2)$$

حيث Cov هى تعبير عن التغاير covariance بين المتغيرين X_1، X_2. وفى حالة كون X_1، X_2 مأخوذين عشوائياً من العشيرة، فإن الحد الثالث أى التغاير يصبح مساوياً للصفر ويصبح تباين الفرق بين X_1، X_2 يساوى مجموع التباين لكليهما. أى:

$$V(X_1 - X_2) = V(X_1) + V(X_2) \qquad (٨-٣)$$

فإذا كانا مأخوذين من نفس العشيرة (دون تأثير للمعاملة على التباين) فإنه فى هذه الحالة:

$$V(X_1 - X_2) = 2\sigma_X^2 \qquad (٨-٤)$$

وبتطبيق هذا المفهوم على الفرق بين متوسطين $(\overline{X}_1 - \overline{X}_2)$ وحيث إنه قد سبق توضيح أن تباين المتوسط للعينات التى عدد أفرادها n عبارة عن σ^2 / n فإنه ينتج أن تباين الفرق بين المتوسطين:

$$V(\overline{X}_1 - \overline{X}_2) = \frac{2\sigma^2}{n} \qquad (٨-٥)$$

وبناء على ما تقدم وفى حالة مقارنة الأزواج تكون قيمة $Cov(X_1, X_2)$ فى (٨-١) دائما موجبة وذلك لأنه كما ذكر سابقاً إذا كانت عملية الأزواج ناجحة فإن أفراد نفس الزوج تكون متشابهة، فلو كانت القيمة $(X_1 - \mu_1)$ موجبة فإن $(X_2 - \mu_2)$ أيضا تكون موجبة ويكون الناتج موجباً. أما إذا كانت قيمة $(X_1 - \mu_1)$ سالبة فإن قيمة $(X_2 - \mu_2)$ تتجه لتصبح سالبة أيضا، وبالتالى يصبح الناتج موجباً. ولذا فإن تباين الفرق بين متوسطين يكون أقل من مجموع التباين فى (٨-٣) حيث تصبح قيمة التغاير covariance موجبة ويكون نتيجة ما تقدم أن يقل حجم تباين متوسط الفروق عن تباين الفرق بين متوسطين مما يجعل التجربة أكثر حساسية فى كشف الفروق بين تأثير المعاملات.

مثال ٨-١

البيانات التالية تمثل الزيادة فى الوزن بالجرام بعد نهاية التجربة فى 10 أزواج من الأرانب مأخوذة من نفس البطن لمعاملتين غذائيتين مختلفتين فى نسبة البروتين فيهما:

رقم الزوج	عليقة ١	عليقة ٢	الفرق D
١	403	234	169
٢	117	130	-13
٣	234	182	52
٤	221	143	78
٥	91	78	13
٦	130	65	65
٧	260	221	39
٨	195	195	0
٩	104	91	13
١٠	156	143	13
المجموع	**1911**	**1482**	**429**
المتوسط	**191.1**	**148.2**	**42.9**

الحل:

$H_1 : \mu_d \neq 0$ ، $H_\circ : \mu_d = 0$

عدد الفروق: 10، مجموع مربع الفروق: $\sum D^2 = 43771$

مجموع مربع الفروق المصحح: $\sum d^2 = 43771 - \frac{(429)^2}{10} = 25366.9$

تباين الفروق: $S_d^2 = \frac{\sum d^2}{n-1} = \frac{25366.9}{9} = 2818.54$

الخطأ القياسي لمتوسط الفروق: $S_{\bar{d}} = \sqrt{S_d^2/n} = \sqrt{2818.54/10} = 16.79 \text{ g}$

وبالتالى فإنه لاختبار معنوية هذه الفروق فيما إذا كانت تختلف عن قيمة الصفر أى أن فرض العدم هو: $H_\circ : \mu_d = 0$ وعليه فإن الاختبار للمعنوية يأخذ الصورة العامة وهى:

$$t = \frac{\bar{d} - \mu_d}{S_{\bar{d}}} \qquad (٨-٦)$$

وبالتالى تكون قيمة t المحسوبة: $t = \frac{42.9 - 0}{16.79} = 2.56$

قيمة t الجدولية (جدول ٤ ملحق أ) عند درجات حرية 9 ومستوى معنوية 0.05 هى $t_{(9,.05)} = 2.26$ ، وبالتالى يرفض فرض العدم، أى أن هناك فرقاً معنوياً بين العليقتين فى تأثيريهما على الزيادة فى الوزن.

ماذا يمكن أن يتحقق من زيادة فى حساسية التجربة عند اتباع طريقة الأزواج مقارنة بطريقة العينات المستقلة؟ من الواضح أنه يمكن الحصول على معلومات توضح مقارنة بين الطريقتين. فإذا فرض أنه فى جدول البيانات تم تحليل كل من الأفراد المنتمية إلى كل من المعاملتين على أساس أن كلا من العينتين تمثل أفراداً مستقلة لا ترتبط مع بعضها بأى تشابه، وبالتالى فإن تحليل التجربة سوف يتم على أساس حساب التباين المجمع pooled variance من العينتين بصفة مستقلة ثم اختبار الفرق بين متوسطى المعاملتين حسب فرض العدم بتساوى متوسطى المعاملتين وليس عدم وجود اختلاف لمتوسط الفروق عن الصفر كما فى الحالة الأولى ويكون التحليل بالتالى كما يلى:

العينة الأولى (العليقة ١)

المجموع: $\sum X_1 = 1911$ ، مجموع المربعات غير المصحح $\sum X_1^2 = 445653$

مجموع المربعات المصحح: $\sum x_1^2 = 445653 - \dfrac{(1911)^2}{10} = 80460.9$

التباين $S_1^2 = 80460.9/9 = 8940.1$

العينة الثانية (العليقة ٢)

المجموع: $\sum X_2 = 1482$ ، مجموع المربعات غير المصحح $\sum X_2^2 = 251134$

مجموع المربعات المصحح: $\sum x_2^2 = 251134 - \dfrac{(1482)^2}{10} = 31501.6$

التباين $S_1^2 = 31501.6/9 = 3500.18$

ولكن يمكن الحصول على قيمة واحدة للتباين تمثل التباين المرجح بدرجات الحرية ولكن يلزم أن يجرى أولاً اختبار التجانس test of homogeneity باستخدام اختبار F بأن يقسم التباين الأكبر على التباين الأصغر وتقارن القيمة الناتجة بقيمة F الجدولية (جدول ٥ ملحق أ) بدرجات حرية كل من البسط والمقام، فإذا كانت القيمة المحسوبة أقل من الجدولية فهذا يعنى أن التباينين متجانسان و عليه فإن اختبار التجانس:

$$F = 8940.1/3500.18 = 2.55$$

قيمة F الجدولية عند درجة حرية 9 لكل من البسط والمقام ومستوى معنوية 5% تساوى 3.18 وهى أكبر من F المحسوبة وبالتالى فإن التباينين متجانسان ويحسب التباين المجمع pooled variance كالتالى:

$$S_p^2 = \frac{S_1^2(n_1-1)+S_2^2(n_2-1)}{(n_1-1)+(n_2-1)}$$

$$= \frac{\sum x_1^2 + \sum x_2^2}{n_1+n_2-2} = \frac{80460.9+31501.6}{18} = 6220.14$$

وتباين الفرق بين متوسطين:

$$S^2_{(\overline{X}_1-\overline{X}_2)} = S_p^2\left(\frac{1}{n_1}+\frac{1}{n_2}\right) = 6220.14\left(\frac{1}{10}+\frac{1}{10}\right) = \frac{2(6220.14)}{10}$$

والخطأ القياسى للفرق بين المتوسطين: $S_{(\bar{x}_1-\bar{x}_2)} = \sqrt{\frac{2(6220.14)}{10}} = 35.27$

وبما أن فرض العدم بتساوى المعاملتين يكون فى هذه الحالة: $H_\circ : \mu_1 - \mu_2 = 0$ والفرض البديل $H_1 : \mu_1 - \mu_2 \neq 0$ فإنه لاختباره بواسطة اختبار t تكون صورته كالتالى:

$$t = \frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_{(\bar{x}_1-\bar{x}_2)}} \qquad (٧-٨)$$

وبالتالى تكون t المحسوبة: $t = \frac{191.1-148.2}{35.27} = 1.216$

وبالكشف عن معنوية هذا الفرق عند مستوى 0.05 ودرجات حرية $18 = (n_1+n_2-2)$، فإن قيمة t الجدولية 2.101، وبالتالى يصبح الفرق غير معنوى، أى أنه لا يمكن رفض فرض العدم بتساوى متوسطى المعاملتين فى تأثيرهما على الزيادة فى الوزن.

مثال ٨-٢

استخدام برنامج SAS لإجراء اختبار t فى أزواج للبيانات المذكورة فى مثال ٨-١

```
DATA PAIRED;
INPUT RATION1 RATION2 @@;
DIFF=RATION1-RATION2;
CARDS;
403 234 117 130 234 182 221 143 91 78
130 65 260 221 195 195 104 91 156 143
PROC MEANS N MEAN STD STDERR T PRT;
VAR RATION1 RATION2 DIFF;
TITLE 't-test in paired comparison';
Run;
```

نتيجة التحليل:

t-test in paired comparison

The MEANS Procedure

Variable	N	Mean	Std Dev	Std Error	t Value	Pr > |t|
RATION1	10	191.1000000	94.5521020	29.9000000	6.39	0.0001
RATION2	10	148.2000000	59.1623003	18.7087621	7.92	<.0001
DIFF	10	42.9000000	53.0899656	16.7885212	2.56	0.0309

مثال ٨-٣

استخدام برنامج SAS لإجراء اختبار t فى مجاميع للبيانات المذكورة فى مثال ٨-١

```
DATA GROUPS;
INPUT TRT GAIN @@;
CARDS;
1 403 1 117 1 234 1 221 1 91 1 130 1 260 1 195 1 104 1 156
2 234 2 130 2 182 2 143 2 78 2 65 2 221 2 195 2 91 2 143
PROC TTEST COCHRAN;
CLASS TRT;
VAR GAIN;
TITLE 't-test in GROUPS';
Run;
```

لاحظ أن اختبار COCHRAN يستخدم عند عدم تجانس التباين وذلك لتقدير قيمة مستوى المعنوية لاختبار t المقرب (Steel and Torrie, 1980).

نتيجة التحليل:

t-test in GROUPS

The TTEST Procedure

Variable	Method	Variances	DF	t Value	Pr > |t|
GAIN	Pooled	Equal	18	1.22	0.2396
GAIN	Satterthwaite	Unequal	15.1	1.22	0.2425
GAIN	Cochran	Unequal	9	1.22	0.2548

Equality of Variances

Variable	Method	Num DF	Den DF	F Value	Pr > F
GAIN	Folded F	9	9	2.55	0.1786

مما سبق يتضح أن إجراء عملية الإزواج قد أدى إلى تخفيض فى حجم الخطأ القياسى لاختبار الفرق بين المعاملتين ولقد نتج هذا عن استخدام أفراد متشابهة مع بعضها مبدئياً مما قلل من الاختلافات داخل الأزواج أى أن معامل الارتباط فيما بينهما مرتفعاً.

ويمكن القول إن تجارب الأزواج فى حالة نقص الاختلافات داخل الأزواج عن الاختلافات بين الأفراد العشوائية تؤدى إلى زيادة فى كفاءة التجربة للإحساس بالفروق بين المعاملات، أو بمعنى آخر أنه باستخدام العينات المستقلة كان يستوجب للحصول على نفس القدر من الكفاءة (مقدرة على أساس الخطأ القياسى للفرق بين المتوسطين) استخدام عدد أكبر من المشاهدات فى التجربة. ولكن فى نفس الوقت فإن درجات الحرية المستخدمة فى نظام العينات المستقلة يكون ضعف درجات الحرية المتاحة للخطأ فى حالة الأزواج، وأيضا قد يؤدى استخدام أزواج المشاهدات إلى الحصول على تباين أكبر للفرق بين المتوسطين إذا ما كان الارتباط موجباً بين أزواج المشاهدات التى تستخدم فى التجريب كما وأن إهمال الأزواج وإجراء التجارب كما لو كانت مستقلة مع وجود الارتباط يؤدى إلى فقد معلومات قيمة وزيادة فى الأعباء كان يمكن توفيرها.

صندوق ٨-١

يستخدم اختبار t بطريقة العينتين المستقلتين إذا لم يكن التشابه بين الفرد في معاملة وفرد في المعاملة الأخرى أكثر من التشابه بين الأفراد في العينة عموماً.

بينما يستخدم اختبار t بطريقة الأزواج إذا كان التشابه بين فرد في معامله وفرد في المعاملة الأخرى أكبر من التشابه بين الأفراد عموماً.

وبالإشارة إلى اختبار التشابه السابق ذكره إذا كانت F المحسوبة معنوية عند مستوى المعنوية المتعارف عليه فى مجالات الزراعة المختلفة وهو 5% فإنه بالتالى لا يمكن استخراج قيمة متجانسة للتباين ولكن تكون خطوات التحليل كما يلى:

بعد حساب المتوسط والتباين لكل من المجموعتين واختبار التجانس والذى أدى إلى رفض الفرض القائل بأن التباين متجانس فإن:

$$t = \frac{(\overline{X}_1 - \overline{X}_2) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \qquad (٨-٨)$$

تقارن قيمة t المحسوبة هذه بقيمة t المعدلة والتى تستخرج من جدول t (جدول ٤ ملحق أ) بعد حساب قيمة جديدة لدرجات الحرية تسمى effective df وهى مأخوذة عن Satterthwaite (1946) حيث:

$$\text{effective df} = \frac{(S_1^2/n_1 + S_2^2/n_2)^2}{(S_1^2/n_1)^2/(n_1-1) + (S_2^2/n_2)^2/(n_2-1)} \qquad (٩-٨)$$

ثم يستكمل التحليل الإحصائى كما سبق.

تمارين الباب الثامن

٨-١ التالى هى درجات 8 طلاب التى حصلوا عليها فى مادتين مختلفتين، هل درجات الطلاب متساوية للمادتين ؟

رقم الطالب	المادة الأولى	المادة الثانية
١	42	84
٢	41	28
٣	13	34
٤	99	69
٥	80	38
٦	88	28
٧	30	5
٨	100	80

٨-٢ لمقارنة الدرجات التى تحصل عليها الطلاب فى مادتين مختلفتين، أخذ 10 طلاب بطريقة عشوائية وأعطوا اختبار المادة الأولى، كما أخذ 10 طلاب آخرين وأعطوا اختبار المادة الثانية. وكانت درجات الطلاب كما يلى

المادة الأولى		المادة الثانية	
رقم الطالب	الدرجة	رقم الطالب	الدرجة
١	75	١	34
٢	35	٢	47
٣	73	٣	99
٤	49	٤	92
٥	50	٥	20
٦	100	٦	78
٧	41	٧	84
٨	17	٨	27
٩	94	٩	80
١٠	56	١٠	87

هل الدرجات متساوية فى المادتين ؟

٨-٣ البيانات التالية تمثل الزيادة الكلية في الوزن بالكيلوجرام في ذكور الحملان من الولادة التوأمية والمغذاة على عليقتين أ ، ب

رقم زوج التوائم	عليقة أ	عليقة ب
١	14.0	14.4
٢	15.2	13.7
٣	15.7	14.7
٤	14.2	12.8
٥	14.8	14.2
٦	14.9	14.3
٧	14.1	13.5
٨	13.0	13.7
٩	14.6	13.1
١٠	16.8	11.9

والمطلوب اختبار فرض العدم بتساوى كمية الزيادة الكلية فى الوزن فى العليقتين.

٨-٤ اختبر فرض العدم بأنه لا يوجد فروق بين متوسطى العليقتين فى التمرين السابق بغض النظر عن العلاقة بين فردى نفس الزوج ثم قارن النتائج المتحصل عليها مع نتائج التمرين السابق.