اختبارات T حول المتوسطات
T Tests about Means

مقدمة:

اختبارات الفروض هي إحدى أساليب الاستدلال الإحصائي حيث يتم من خلالها الحكم على مصداقية افتراضات الباحث حول قيمة المعالم الحقيقية للمجتمع. ويتم في عملية اختبار الفرضيات استخدام البيانات الإحصائية كأداة لإثبات صحة أو عدم صحة افتراضات الباحث. وسيتم خلال هذا الباب التعرف على الفرضية الإحصائية وإجراء اختبارات لفرضيات حول متوسطات المجتمع، بالإضافة إلى إنشاء فترات ثقة حول متوسطات المجتمع والتي تعطي معلومات أكثر عن المتوسطات. ويجب التأكيد هنا على أن صحة جميع الاختبارات مرتبطة بصحة الفرضية القائلة بأن البيانات تتبع التوزيع الطبيعي.

الفرضية الإحصائية:

تتلخص فكرة اختبار الفروض في أنه لدى الباحث اعتقاد أو قناعه حول مجتمع الدراسة ويرغب في تأكيد صحة هذه القناعة، لذلك يلجأ الباحث إلى جمع البيانات الإحصائية حول مجتمع الدراسة ومن ثم التأكد من أن البيانات الإحصائية تؤيد أو لا تؤيد فرضيته. وبذلك يمكن تعريف الفرض الإحصائي على أنه إدعاء أو اعتقاد حول معلمه من معالم المجتمع من قبل الباحث، ويعتقد الباحث أن هذا الإدعاء صحيح ما لم تثبت البيانات خلاف ذلك.

اختبار الفرضيات الإحصائية

يمكن تعريف اختبار الفرضيات بأنه مجموعة من العمليات التي يتم فيها استخدام نتائج العينة والنظرية الإحصائية للتأكد من أن الفرضية المطروحة هي عبارة صحيحة أم غير صحيحة. ويمكن تحديد مجموعة من الخطوات التي يجب إتباعها للقيام باختبار الفرضيات بطريقة علمية، وهي على النحو التالي:

1. تحديد فرضية العدم والفرضية البديلة

تتطلب عملية اختبار الفرضيات أن يكون لدى الباحث ادعاء أو فرضية تقبل الخطأ والصواب ويرغب اختبار صحة هذه الفرضية، وكما ذكرنا سابقاً، فإن هذا الادعاء مرتبطاً بقيمة معلمة من معالم المجتمع الإحصائي والتي تعتبر مجهولة مثل متوسط المجتمع. ويستخدم المصطلح فرضية العدم للدلالة على الحالة التي تشير إلى عدم صحة فرضية الباحث، أما الفرضية البديلة فهي المصطلح المستخدم للدلالة على ادعاء الباحث. ونظراً لأن فرضية العدم تناقض فرضية الباحث، فقد يكون لدى الباحث الرغبة بأن لا تؤيد البيانات الإحصائية فرضية العدم وبذلك تكون فرضية الباحث هي الفرضية الصحيحة. وكم ذكرنا سابقا فإن الفرضية البديلة تشير إلى قناعة الباحث حول أمر ما في مجتمع الدراسة، ويكون هدف الباحث هو استخدام البيانات الإحصائية للتأكد من صحة أو عدم صحة هذا الإدعاء أو القناعة. ويجب التأكيد هنا بأن رفض فرضية العدم وتأييد فرضية الباحث سيؤدي بالضرورة إلى إجراء ما أو إلى التوصية بعمل إجراء ما. وهذا الإجراء هو في العادة إجراء غير مرغوب فيه من قبل الباحث لأنه قد يؤدي إلى دفع تكاليف يمكن تجنبها في حالة عدم رفض فرضية العدم.

ولتوضيح فكرة فرضية العدم والفرضية البديلة، سنذكر فيما يلي مثالين يتم فيهما طرح فكرة اختبار فرضية حول معلم مجتمع، فعلى سبيل المثال، فقد قام مدير مراقبة الجودة في شركة ألبان (الباحث) وبناء على ملاحظة شخصية والتي تشير إلى وجود خلل في الكمية المعبأة من الحليب في خط إنتاج رقم 37. وبنا على هذه الملاحظة، فقد أفترض مدير مراقبة الجودة بأن متوسط كمية الحليب المعبأة في عبوات ذات 2 لتر تختلف عن 2000 ملل، وفي هذه الحالة يتم تكوين فرضية العدم (H_0) بأن متوسط كمية الحليب المعبأة هو 2000 ملل وتكون الفرضية البديلة (H_a) هو أن متوسط كمية الحليب المعبأة يختلف عن 2000 ملل. وتتم صياغة هذه الفرضية رياضياً على النحو التالي:

$$H_0 : \mu = 2000$$
$$H_a : \mu \neq 2000$$

وبناء على هذه الفرضية، فإن الباحث يقوم بجمع عينة من خط الإنتاج رقم 37 ثم يقيس كمية الحليب المعبأ في العبوات، وبناء على نتائج العينة فإن الباحث يستطيع رفض فرضية العدم وتأييد الفرضية البديلة أو العكس. إن الحصول على قيمة $\bar{X}$ قريبة من القيمة المفترضة لمتوسط المجتمع وهي 2000 يؤيد فرضية العدم، إما الحصول على $\bar{X}$ بعيدة بما فيه الكفاية عن القيمة المفترضة لمتوسط المجتمع وهي 2000 سيؤدي إلى رفض فرضية العدم وقبول الفرضية البديلة (فرضية الباحث). إن رفض فرضية العدم وقبول فرضية الباحث سيؤدي إلى قيام الباحث وهو مدير مراقبة الجودة في شركة الألبان إلى وقف إنتاج عبوات الحليب على خط إنتاج رقم 37 ثم القيام بصيانة الخط ليتم تصحيح مشاكل هذا الخط. وبناء على ذلك، فإن رفض فرضية العدم سيؤدي إلى تحمل الشركة تكاليف تصحيح مشاكل خط الإنتاج، بالإضافة إلى تكاليف أخرى ناتجة عن توقف الخط حتى يتم تصحيح المشكلة.

إلا أن السؤال المهم هنا هو كيفية تحديد نقاط الفصل بين منطقة الرفض والقبول لفرضية العدم؟ أو بعبارة أخرى، ما هي قيمة $\bar{X}$ التي تؤدي إلى رفض أو عدم رفض فرضية العدم؟ وللإجابة على هذا السؤال، فإننا نلجأ إلى النظرية الإحصائية لتحديد نقاط الفصل التي تحقق لنا معايير محددة ومنطقية.

وكمثال آخر، إذا كان لدى باحث قناعه بأن لديه قطعة نقود غير متوازنة أو متحيزة لأحد الجهتين (أرقام أو شعار)، لذلك فإن الباحث يعتقد بأن نسبة عدد مرات الحصول على شعار تختلف عن 0.5 ، لذلك يمكن صياغة هذه الفرضية على النحو التالي:

$$H_0 : \lambda = 0.5$$
$$H_a : \lambda \neq 0.5$$

حيث λ تمثل نسبة عدد مرات الحصول على شعار من العدد الكلي لرمي القطعة النقدية

فإذا تم رمي قطعة النقود مئة مره، فإن يمكن رفض فرضية العدم إذا كان عدد مرات الحصول على شعار قليل جدا أو كبير جدا، أو بمعنى آخر، فإنه عند الحصول على شعار 20 مره والحصول على أرقام 80 مره فإن هذه النتيجة لا تؤيد فرضية العدم، وكذلك فإنه عند الحصول على شعار 80 مره والحصول على أرقام 20 مره، فإن هذه النتيجة أيضاً لا تؤيد فرضية العدم. إن قناعة الباحث بأن لديه قطعة نقود غير متوازنة أدى إلى إجراء التجربة السابقة وذلك لجمع البيانات ثم تم استخدام نتائج التجربة لتأييد أو عدم تأييد فرضية الباحث. ومرة أخرى فإن السؤال

المهم هنا هو متى يمكن القول بأن قطعة النقود غير متحيزة، أو ما هي النتائج التي تؤدي إلى عدم رفض فرضية العدم؟ وما هي النتائج التي تؤدي إلى رفض فرضية العدم؟ وللإجابة على هذا السؤال المهم، يجب تحديد نقاط نستطيع من خلالها الفصل بين النتائج التي تؤدي إلى رفض فرضية العدم أو عدم رفض فرضية العدم. وفي الحقيقة فإننا سنلجأ إلى نتائج العينة واستخدام النظرية الإحصائية للإجابة على هذا السؤال وتحديد نقاط الفصل بين منطقة (مناطق) رفض فرضية العدم ومنطقة عدم رفض فرضية العدم.

2. *تحديد مستوى المعنوية*

ويمكن تعريف مستوى المعنوية بأنه احتمال رفض فرضية العدم في حين أن فرضية العدم صحيحة ويرمز له بالرمز α، ويطلق عادة على هذا الاحتمال بالخطأ من النوع الأول (Type I Error). أي أنه بناء على عينة الدراسة، فإنه يمكن رفض فرضية عدم صحيحة وذلك لأنه نتائج العينة كانت متطرفة إلى درجة أدت إلى رفض فرضية العدم، وهذه النتائج المتطرفة تحصل باحتمال α. وتسمى α بمستوى المعنوية في حين تسمى $(1-\alpha)$ بمستوى ثقة الباحث في صحة قرار رفض فرضية العدم، وبذلك فإنه كلما انخفضت ثقة الباحث في القرار كانت منطقة الرفض أكبر والعكس صحيح. وعادة ما تعكس قيمة α درجة تحفظ الباحث حول الدراسة، فتكون قيمة α تساوي 0.01 إذا كان الباحث متحفظ بدرجة كبيرة ولا يريد أن يتحقق خطأ رفض فرضية عدم صحيحة إلا باحتمال 0.01. وقد يخفف الباحث درجة التحفظ ويقرر بأن تكون قيمة α تساوي 0.05، وذلك لأن درجة التحفظ القصوى غير مطلوبة. وعادة ما يقوم الباحث بتحديد قيمة α بناء على مستوى الثقة المطلوبة في قرار رفض فرضية العدم والذي يرغب فيه الباحث. فعندما يكون مستوى الثقة $(1-\alpha)$ كبير فإن قيمة α تكون صغيرة وبذلك تكون منطقة الرفض صغيرة، وعندها فإن الباحث يكون على ثقة كبيرة في صحة قراره عندما يتم رفض فرضية العدم. وقد يرى البعض أنه من المناسب اعتبار أن قيمة α تساوي 0.001 هو الأنسب لاختبار الفرضية لأنها بذلك سترفع درجة ثقة الباحث في قراره، إلا أن وضع قيمة α صغير جداً يؤدي إلى زيادة احتمال عدم رفض فرضية عدم خاطئة وهو ما يسمى الخطأ من النوع الثاني ويرمز له بالرمز β (Type II Error)، كذلك فإن زيادة ثقة الباحث يؤدي إلى عدم الدقة في تحديد ما يسمى بفترات الثقة.

ويجب التأكيد على أن تحديد قيمة α تتم قبل إجراء الحسابات اللازمة لتحديد قيمة إحصائية الاختبار. وعلى الرغم من أنه يمكن السماح باستخدام أي قيمة لـ α، إلا أنه في كثير من الحالات يتم استخدام القيم (0.1، 0.05، 0.01) باعتبارها القيم المناسبة لـ α. إن العامل

الرئيس في عدم استخدام قيمة α أقل من 0.01 هو أن ذلك سيؤدي إلى ارتفاع قيمة β وهي الخطأ من النوع الثاني، لذا فإنه يجب المفاضلة بين الأمرين والسعي إلى تقليل كلا النوعين من الخطأ وتحديد القيمة المناسبة لـ α. إن تقليل قيمة α سيؤدي إلى عدم رفض فرضية العدم بصورة أكبر مما يجب، مما يؤدي إلى زيادة حالات عدم رفض فرضيات عدم خاطئة وهو الخطأ من النوع الثاني. ويجب التنويه هنا أنه يمكن تقليل قيمة α وقيمة β آنياً وذلك بزيادة حجم العينة.

3. اختيار إحصائية الاختبار

يمكن تعريف إحصائية الاختبار بأنها قيمة يتم حسابها من بيانات العينة وتستخدم هذه القيمة في اتخاذ قرار رفض أو عدم رفض فرضية العدم. ويجب التأكيد بأن إحصائية الاختبار تختلف من اختبار إلى آخر وذلك حسب نوع الفرضية. فعلى سبيل المثال، فإن الإحصائية التالية تستخدم لاختبار فرضيات حول متوسطات مجتمع واحد، وبناء على النظرية الإحصائية فإن هذه الإحصائية تتبع توزيع احتمالي يسمى توزيع (T) وهو أحد التوزيعات الاحتمالية والتي تشابه وبصورة كبيرة التوزيع الطبيعي.

$$t_v = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$$

حيث μ_0 هي القيمة المفترضة من قبل الباحث، S هو الانحراف المعياري للمتغير محل الدراسة والتي تم حسابها من بيانات العينة، n تمثل عدد المشاهدات للمتغير أو عدد حالات الدراسة، $(v = n - 1)$ هي درجات الحرية لتوزيع T. وبناء على قيمة α والتي تم تحديدها في الفقرة السابقة، يتم تحديد نقاط الفصل بين منطقة الرفض ومنطقة القبول على توزيع T بدرجات حرية (v)، ومن ثم يتم رفض فرضية العدم وقبول الفرضية البديلة عندما تقع قيمة t_v المحسوبة في منطقة الرفض، في حين يتم عدم رفض فرضية العدم ورفض الفرضية البديلة عندما يحدث عكس ما سبق.

4. تحديد قاعدة اتخاذ القرار

قاعدة اتخاذ القرار هي الحالة أو مجموعة من الحالات أو الشروط والتي في حالة تحققها سيتم رفض فرضية العدم، وفي حالة عدم تحققها فإنه سيتم عدم رفض فرضية العدم. وفي حالة اختبارات T حول المتوسطات فإنه يمكن القول بأن الباحث يستطيع رفض فرضية العدم إذا كان:

$$P\left(T \geq |t_v|\right) < \frac{\alpha}{2}$$ لاختبار ذو طرفين

و $$P\left(T \geq |t_v|\right) < \alpha$$ لاختبار ذو طرف واحد

وتعني العبارة الرياضية الأولى أنه إذا كان لدينا متغير عشوائي يتبع توزيع T وكان احتمال الحصول على قيمة للمتغير أكبر من أو تساوي القيمة المطلقة لـ t_v أقل من $\alpha/2$، فإن t_v تقع في منطقة الرفض ولذلك يتم رفض فرضية العدم لاختبار ذو طرفين. وبطريقة مشابهة يمكن إيضاح معنى العبارة الرياضية الثانية. وتسمى $P\left(T \geq |t_v|\right)$ بالقيمة الإحتمالية للإختبار أو P-Value.

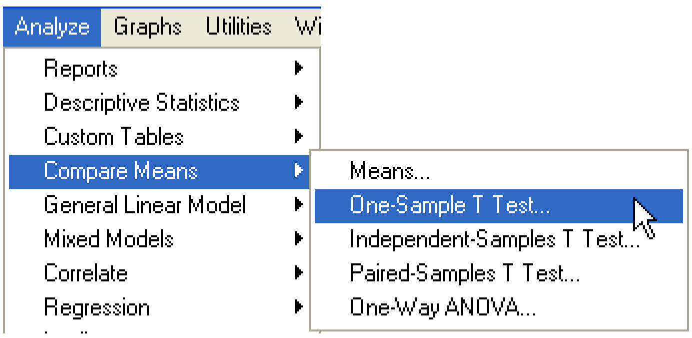
5. اتخاذ القرار

الخطوة الأخيرة في عملية اختبار الفروض هي اتخاذ القرار حول فرضية العدم وذلك بعد إجراء الحسابات اللازمة والتي تم ذكرها في الخطوات السابقة. وقد يكون من المناسب بعد اتخاذ القرار هو إيضاح دلالة هذا القرار من وجهة نظر الباحث.

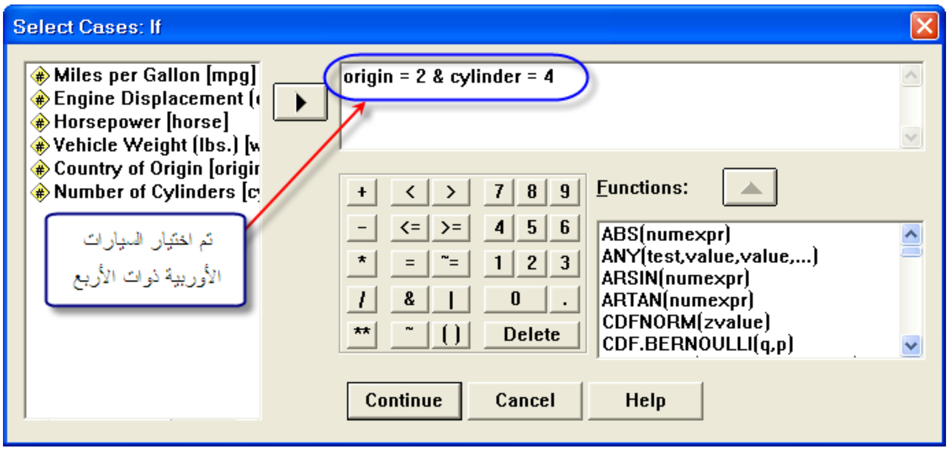
وسيتم فيما يلي اختبار العديد من الفرضيات حول بيانات السيارات باستخدام برنامج SPSS.
لنفرض أن لدى الباحث الرغبة في اختبار الفرضية التي تقول بأن قوة السيارات الأوروبية ذوات الأربع اسطوانات تختلف عن 100 حصان. ويمكن صياغة هذه الفرضية على النحو التالي:

$$H_0 : \mu = 100$$
$$H_a : \mu \neq 100$$

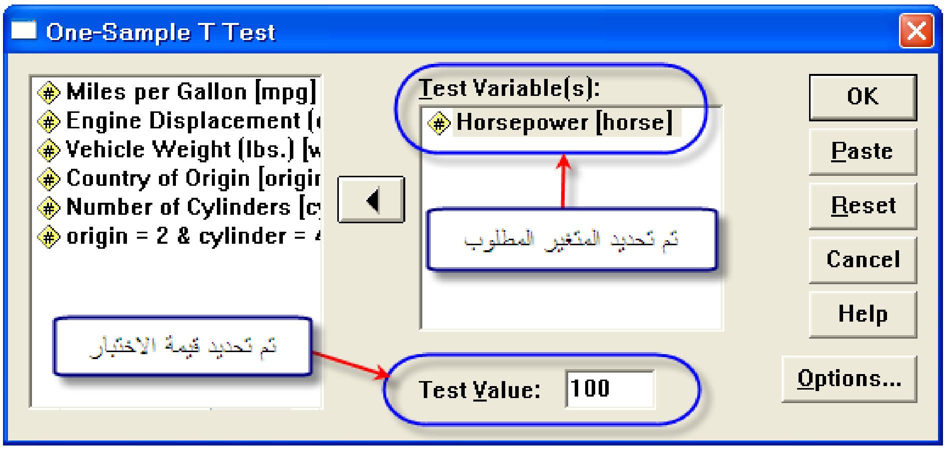
حيث يمثل μ متوسط قوة السيارات الأوروبية ذوات الأربع. ويتم اختبار الفرضية السابقة باستخدام الأمر



إلا أنه يجب اختيار مجموعة السيارات الأوروبية ذوات الأربع اسطوانات باستخدام الأمر Select Cases على النحو التالي:



وعندها سيتم استبعاد جميع السيارات من التحليل ماعدا السيارات المختارة. وعند اختيار الأمر One-Sample T Test يظهر مربع الحوار التالي.



وبالنقر على OK تظهر نتائج الاختبار على شاشة عارض النتائج.

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
Horsepower	61	78.31	18.219	2.333

One-Sample Test

	Test Value = 100					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Horsepower	-9.290	60	.000	-21.69	-26.35	-17.02

فترة الثقة للفرق بين الوسط و القيمة الأفتراضية

قيمة T المحسوبة

درجات الحرية

2(P-Value)

الفرق بين الوسط الحسابي والقيمة الافتراضية

ويوضح الجدول الأول بعض الإحصاءات الوصفية للعينة، في حين يوضح الجدول الثاني نتائج اختبار T. وتوضح النتائج أن قيمة T المحسوبة هي (9.298–) والتي تم حسابها على النحو التالي.

$$t_v = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} = \frac{78.31 - 100}{18.219/\sqrt{61}} = -9.298$$

أما P-Value فإنها تمثل احتمال أن تكون قيمة متغير يتبع توزيع T أكبر من القيمة المطلقة لـ T المحسوبة.

$$P\left(T \geq \left|-9.298\right|\right) = 0$$

ويتم مقارنة الـ P-Value بمستوى معنوية تساوي $\alpha/2$ وذلك لوجود منطقتي رفض تقعان على طرفي توزيع T، فإذا كانت الـ P-Value أقل من $\alpha/2$ ترفض فرضية العدم. لذلك فإن قرار الباحث في هذه الحالة هو رفض فرضية العدم وقبول الفرضية البديلة والتي تقول أن متوسط قوة محركات السيارات الأوروبية يختلف عن 100 حصان. ويمكن ملاحظة أنه إذا كان الوسط الحسابي يقل بكثير عن القيمة الافتراضية أو كان الوسط الحسابي أكبر بكثير عن القيمة الافتراضية، فإن هذا يعطي دلالة على أن بيانات العينة لا تؤيد التساوي بين وسط المجتمع والقيمة الافتراضية.

ويمكن استخدام النتائج السابقة للاختبار الفرضية التالية.

$$H_0 : \mu \geq 100$$
$$H_a : \mu < 100$$

وتختبر هذه الفرضية صحة الادعاء بأن قوة محركات السيارات الأوروبية ذوات الأربع اسطوانات تقل عن 100 حصان. وتبعاً لهذه الفرضية فإن منطقة الرفض لفرضية العدم تقع في المنطقة السالبة وهي المنطقة التي تقع في الطرف الأيسر من توزيع T، ويمكن في هذه الحالة الاستفادة من إشارة t_v، حيث تشير قيمة t_v السالبة مع قيمة الـ P-Value التي تساوي الصفر بأن t_v تقع في منطقة الرفض وبذلك فإن الاختبار يؤيد أن قيمة متوسط قوة المحركات للسيارات الأوروبية ذوات الأربع اسطوانات يقل بكثير عن 100 حصان. وكما تم ذكره سابقاُ فإنه إذا كان للاختبار منطقتي رفض في طرفي التوزيع فإنه يتم مقارنة قيمة الـ P-Value بقيمة $\alpha/2$ أما إذا كان للاختبار منطقة رفض واحدة فإنه يتم مقارنة الـ P-Value بقيمة α، ويكون القرار هو رفض فرضية العدم إذا كانت الـ P-Value أقل من α.

ويوفر الأمر One-Sample T Test إنشاء فترة ثقة للفرق بين متوسط المجتمع والقيمة الافتراضية بناء على بيانات العينة، وبذلك فإنه إذا كانت هذه الفترة لا تحتوي الصفر فإن ذلك

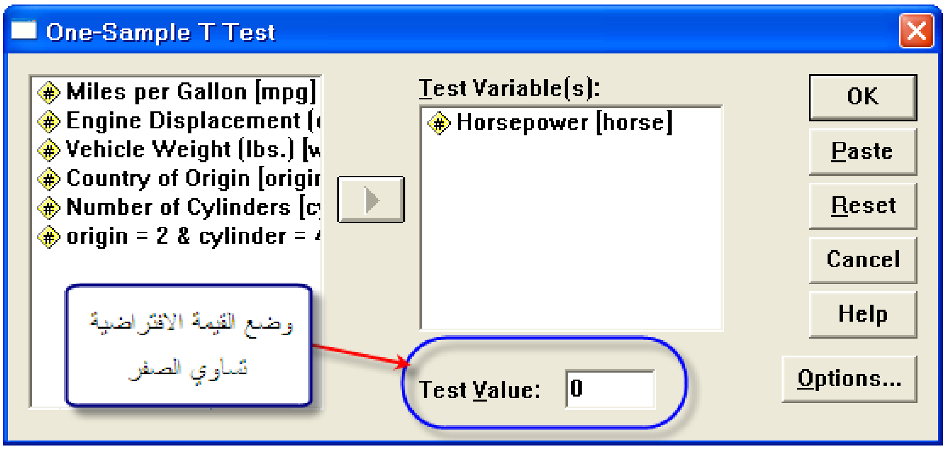
يدل على العينة تؤيد وجود فرق جوهري بين الوسط الحقيقي المجتمع والقيمة الافتراضية. فإذا كان قيم حدود الفترة سالبة، دل ذلك على أن متوسط المجتمع يقل عن القيمة الافتراضية، وإذا كان قيم حدود الفترة موجبة فإن متوسط المجتمع أكبر من القيمة الافتراضية. ويمكن صياغة هذه الفترة رياضياً على النحو التالي.

$$P\left(\left(\bar{X}-\mu_0\right)-t_{v,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\le(\mu-\mu_0)\le\left(\bar{X}-\mu_0\right)+t_{v,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\right)=1-\alpha$$

وتعني أن الباحث على ثقة بمقدار $(1-\alpha)$ في أن الفرق بين متوسط قوة محرك السيارات الأوروبية والقيمة الافتراضية يقع في الفترة المحسوبة.

ويمكن حساب فترة ثقة لوسط المجتمع وذلك بوضع القيمة الافتراضية تساوي صفر عند استخدام الأمر One-Sample T Test. والصيغة الرياضية لهذه الفترة هي:

$$\bar{X}-t_{v,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}\le\mu\le\bar{X}+t_{v,\frac{\alpha}{2}}\frac{S}{\sqrt{n}}$$



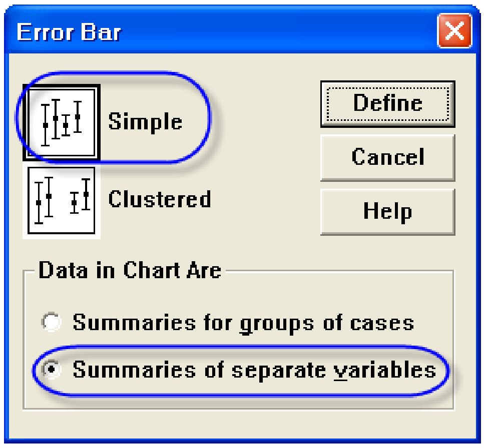
وبذلك تظهر النتائج على النحو التالي.

One-Sample Test

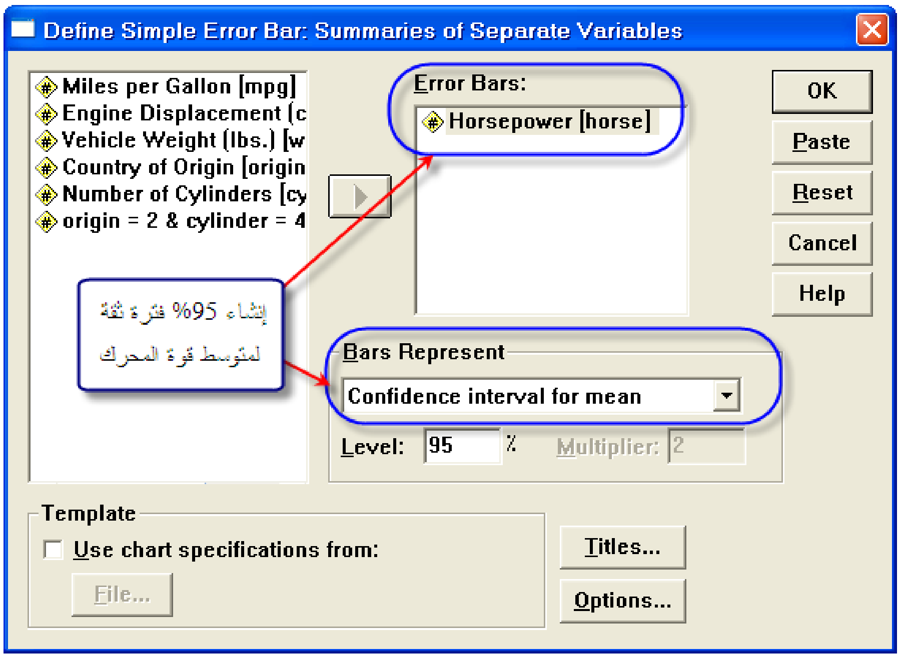
فترة ثقة حول وسط المجتمع

	Test Value = 0					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
Horsepower	33.572	60	.000	78.31	73.65	82.98

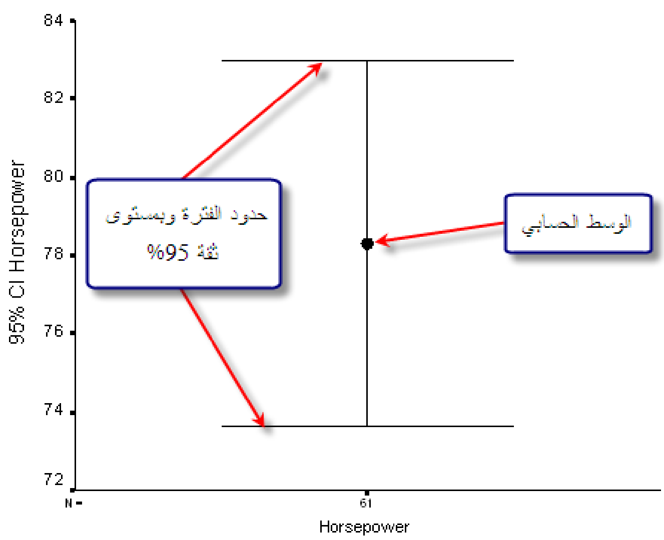
وبذلك فإن الباحث على ثقة بمقدار 95% بأن قوة المحرك للسيارات الأوروبية ذوات الأربع اسطوانات تقع (73.65، 82.98). ويمكن تمثيل الفترة السابقة بيانياً باختيار الأمر Error Bar من قائمة Graphs ليظهر مربع الحوار التالي.



وبالنقر على Define يظهر مربع الحوار التالي.



وبالنقر على OK تظهر النتائج على شاشة عارض النتائج.



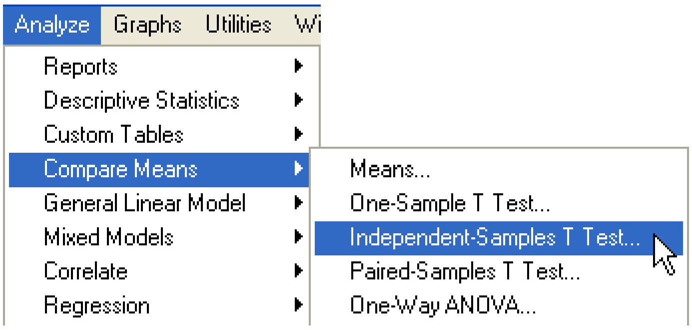
وبذلك فإن فترة الثقة تعطي معلومات أشمل عن الفترة التي من الممكن أن يكون فيها متوسط قوة محركات السيارات الأوروبية حيث أن الباحث يكون على ثقة بمقدار $(1-\alpha)$ بأنه يستطيع رفض أي فرضية عدم تدعي أن متوسط المجتمع يقع خارج فترة الثقة المحسوبة.

اختبارات T حول متوسطات مجتمعين مستقلين

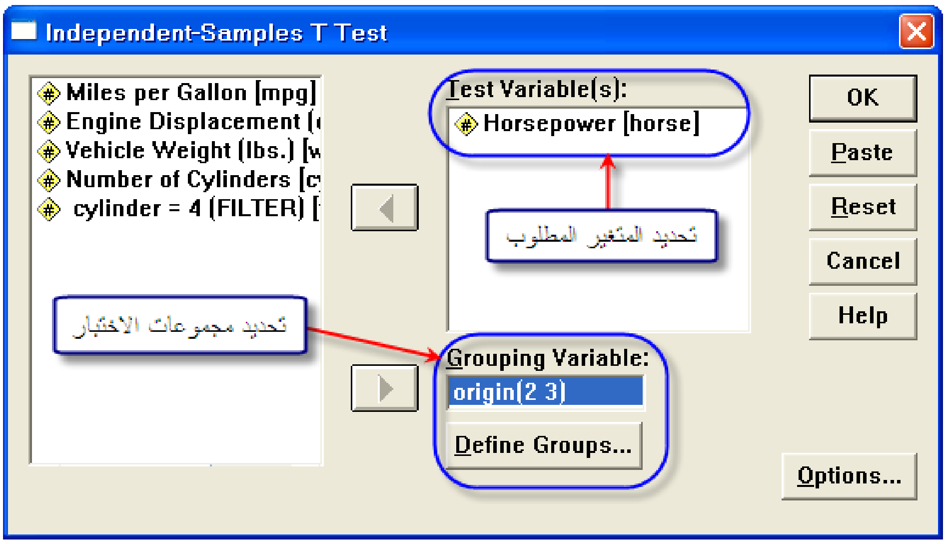
كانت الفرضيات السابقة تعنى باختبار حول وسط المجتمع لمتغير واحد، وفيما يلي ستكون الفرضية تعنى باختبار باختلاف متوسطات متغير في مجتمعين مستقلين. لنفرض أن لدينا الرغبة في اختبار اختلاف متوسط قوة محركات السيارات الأوروبية ذوات الأربع اسطوانات مع متوسط قوة محركات السيارات اليابانية ذوات الأربع اسطوانات، ويمكن صياغة هذه الفرضية على النحو التالي.

$$H_0 : \mu_2 = \mu_3$$
$$H_a : \mu_2 \neq \mu_3$$

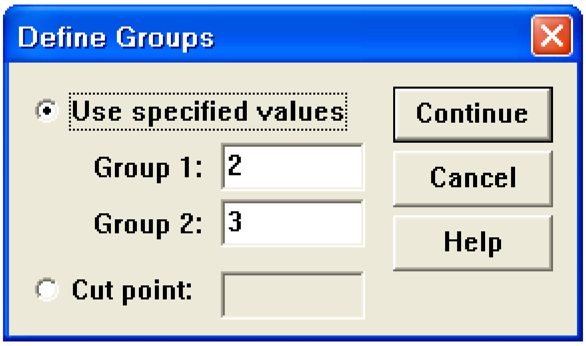
وتشير فرضية العدم إلى تساوي متوسطات قوة محركات السيارات اليابانية والأوروبية حيث μ_2 ترمز لمتوسط قوة محركات السيارات الأوروبية و μ_3 ترمز إلى متوسط قوة محركات السيارات اليابانية. أما الفرضية البديلة فإنها تشير إلى عدم تساوي المتوسطات. ويمكن إجراء الاختبار بواسطة الأمر



إلا أنه يجب التأكد قبل إجراء الاختبار من اختيار مجموعة السيارات ذوات الأربع اسطوانات بواسطة الأمر Select Cases من قائمة Data. وعند اختيار الأمر السابق، يظهر مربع الحوار التالي.



وبناء على ما تم تحديده في مربع الحوار السابق، فإنه سوف يتم مقارنة متوسط المجموعة 2 وهي السيارات الأوروبية مع متوسط المجموعة 3 وهي السيارات اليابانية، أي أنه سوف يتم التعامل مع بيانات السيارات الأوروبية كمتغير مستقل عن المتغير الذي يمثل بيانات السيارات اليابانية.. ويجب دائما تحديد مجموعات الاختبار وذلك النقر على زر Define Groups ثم تحديد القيمة التي ترمز للمجموعة، حيث تمثل كل قيمة مستوى من مستويات متغير تصنيفي. ويجب التأكيد هنا على أن الرموز المستخدمة لتصنيف المجموعات تظهر حسب نفس ترتيبها في فرضية العدم.



وبالنقر على زر Continue ثم OK تظهر النتائج على شاشة عارض النتائج.

Independent Samples Test

		Levene's Test for Equality of Variances		t-test for Equality of Means						
									95% Confidence Interval of the Difference	
		F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Horsepower	Equal variances assumed	1.684	.197	.97	128	.336	2.73	2.831	-2.869	8.333
	Equal variances not assumed			.95	112	.344	2.73	2.877	-2.968	8.431

اختبار ليفين لتساوي تباين المجموعتين

اختبار T بافتراض تساوي التباين

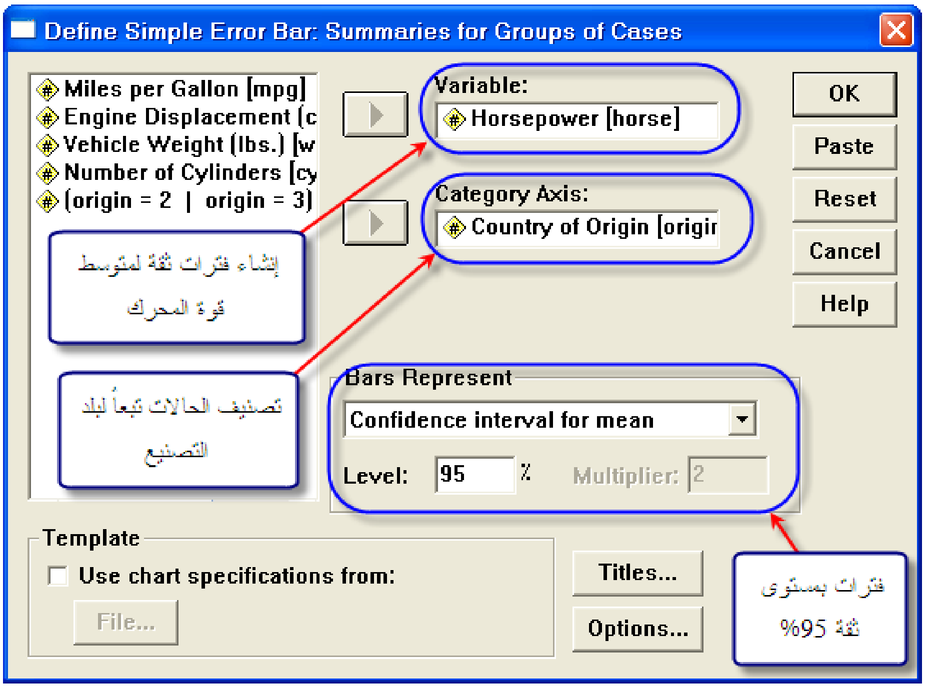
اختبار T التقريبي بافتراض عدم تساوي التباين

وتحتوي النتائج على اختبارين، الأول يعنى باختبار اختلاف تباين المجموعتين والثاني يعنى باختبار T لاختلاف متوسطات المجموعتين. ويمكن صياغة فرضية الاختبار الأول على النحو التالي:

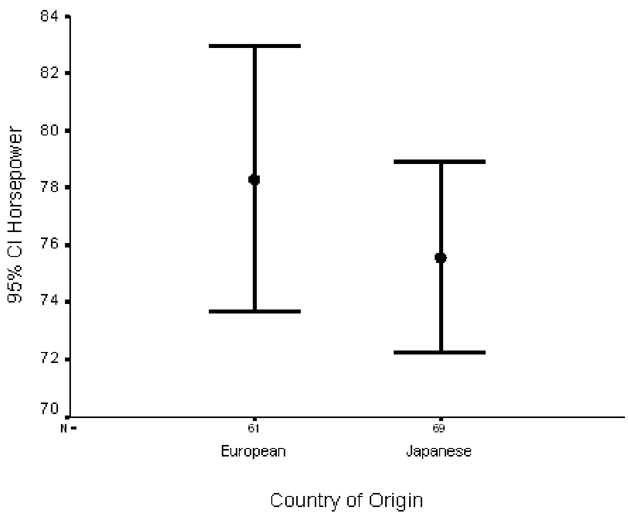
$$H_0: \sigma_1^2 = \sigma_2^2$$
$$H_a: \sigma_1^2 \neq \sigma_2^2$$

وبناء على نتائج اختبار ليفين، فإنه لا يمكن رفض فرضية العدم والتي تشير إلى تساوي تباين المجموعتين، وبذلك فإنه يمكن استخدام نتائج اختبار T لاختبار تساوي متوسطات المجموعتين. وفي الحالة التي يتم فيها رفض فرضية العدم لتساوي التباين، فإنه لا يمكن اعتماد نتائج اختبار T ويمكن استخدام اختبار T التقريبي والذي يفترض عدم تساوي التباين. ويمكن ملاحظة أن

نتائج اختبار T بافتراض تساوي التباين ونتائج اختبار T التقريبي بافتراض عدم تساوي التباين متقاربة جداً، لذلك قد يكون من المناسب دائماً استخدام اختبار T التقريبي حيث أن النتائج ستكون متقاربة بالمقارنة باختبار T وذلك في حالة تساوي تباين، وسيكون هو الاختبار المناسب في حالة عدم تساوي التباين. وبهدف المقارنة بيانياُ بين متوسطات المجموعتين، فإنه يمكن استخدام الأمر Error Bar من قائمة Graphs لعرض فترات ثقة للمتوسطات. وباختيار الرسم من نوع Simple ثم اختيار Summaries for groups of cases ثم النقر على زر Define يظهر مربع الحوار التالي.



وبتحديد المتغير المطلوب وتحديد المتغير التصنيفي ومستوى الثقة ثم النقر على OK، تظهر النتائج التالية.



ويلاحظ التداخل الكبير بين فترات الثقة، وهذه إشارة إلى تساوي متوسط قوة المحركات للسيارات الأوروبية مع متوسط قوة المحركات للسيارات اليابانية.

اختبار T للبيانات المزدوجة Paired-Sample T Test

يستخدم اختبار T للبيانات المزدوجة لاختبار الاختلاف بين متغيرين كميين، إلا أن الاختلاف هنا عن الاختبارات السابقة لـ T هو أن المشاهدات في المتغيرين تكون لنفس الحالة. فعلى سبيل المثال يمكن قياس ضغط الدم لدى مجموعة من الأفراد، ثم تم توفير علاج لجميع الأفراد، ومن بعدها يتم قياس ضغط الدم بعد فترة زمنية، وبذلك سيتم الحصول على مشاهدتين لكل حالة، مشاهدة قبل العلاج ومشاهدة بعد العلاج. لذلك فإن المتغير الأول يمثل قياس ضغط الدم قبل العلاج، في حين يمثل المتغير الثاني قياس ضغط الدم بعد العلاج. ويقوم اختبار T للبيانات المزدوجة باختبار عدم وجود اختلاف بين متوسطات المتغيرين، أو بعبارة أخرى فإن الاختبار يختبر إذا كان متوسط الفرق بين المشاهدات المتقابلة لا يختلف عن الصفر. لذلك فإن اختبار T للبيانات المزدوجة يقوم بحساب الفروقات بين المتغيرين وذلك بطرح أحد المتغيرين من الآخر ثم استخدام النتائج لاختبار أن متوسط الفروقات يختلف عن الصفر. لنفرض أن لدينا بيانات عن 15 مراجع لأحد المستشفيات وقد تم قياس ضغط الدم قبل تقديم علاج ثم بعد فترة من تقديم العلاج. وسوف نستخدم هذه البيانات الافتراضية للقيام باختبار T للبيانات المزدوجة.

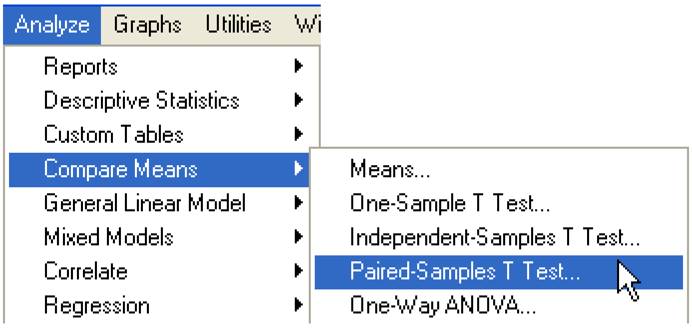
الجدول التالي يوضح البيانات الافتراضية.

ضغط الدم بعد العلاج After	ضغط الدم قبل العلاج Before
5.30	5.40
5.60	5.40
6.50	6.20
6.10	5.80
6.20	6.30
5.80	5.60
5.70	5.80
5.70	5.30
6.20	5.80
6.30	6.10
5.70	5.60
5.80	5.40
5.90	5.80
6.20	5.90
6.30	6.40

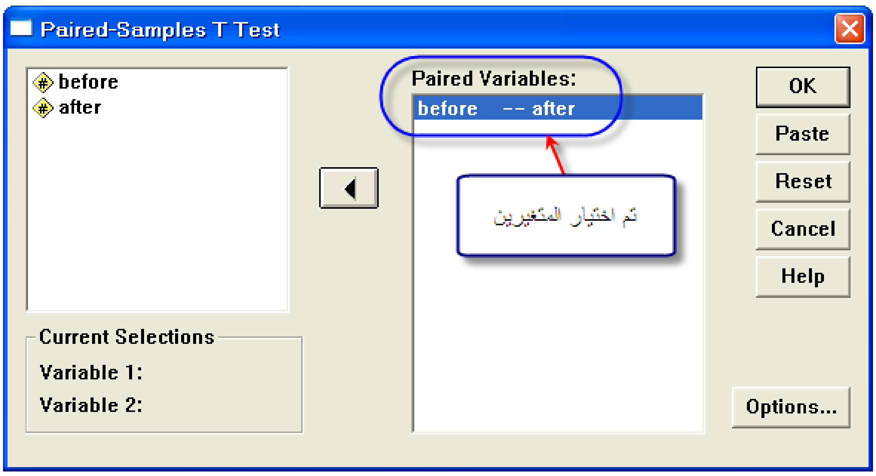
وبعد إدخال البيانات في برنامج SPSS واستخدام أسماء المتغيرات before و after للمتغيرين على التوالي، سيتم اختبار تأثير العلاج على المراجعين وذلك باختبار الفرضية التالية:

$$H_0 : \mu_B = \mu_A$$
$$H_a : \mu_B \neq \mu_A$$

وبذلك فإن فرضية العدم تدعي أن متوسط قياس ضغط الدم قبل العلاج مساوي لمتوسط قياس ضغط الدم بعد العلاج، في حين أن الفرضية البديلة تشير إلى اختلاف المتوسطات، مما يعني أن العلاج له تأثير على ضغط الدم للمراجعين. وباستخدام الأمر



يظهر مربع الحوار التالي.



وبالنقر على زر OK تظهر النتائج التالية.

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	BEFORE	5.9533	15	.33138	.08556
	AFTER	5.7867	15	.34614	.08937

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	BEFORE & AFTER	15	.841	.000

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	BEFORE - AFTER	.1667	.19149	.04944	.0606	.2727	3.371	14	.005

وتشير النتائج إلى أن P-Value=0.0025، لذلك نرفض فرضية العدم لصالح الفرضية البديلة، مما يشير إلى أن العلاج كان له تأثير ايجابي في خفض مستوى ضغط الدم لدى المراجعين.