الفصل الثامن عشر

اختبار الفرضيات
Test of Hypothesis

18.1. مقدمة Introduction

فى الأجزاء السابقة من هذا الكتاب أوضحنا أن الهدف الأساسى للإحصاء الاستدلالي Inferential Statistic هو استنتاج معلمات (خصائص) المجتمع Population Parameters من قيم إحصاءات العينة Samples Statistics، وكذلك أكدنا على أن القيم المستنتجة لن تكون قيم وحيدة وإنما ستكون مدى من القيم Interval، وعرفنا كيف نحدد هذا المدى لمعلمات المجتمع عن طريق فترة الثقة Confidence Interval، وفى هذا الجزء سنتناول موضوعا هاما آخر، وهو بحث صحة الافتراضات Test of Hypotheses التي تم فرضها لقيم معلمات المجتمع، أى التي تم حسابها والحصول عليها، وبعبارة أخرى هل القيم التي تم فرضها لمعلمات المجتمع اعتمادا على القيم التى تم حسابها لإحصاءات العينات صحيحة أم لا ؟

18.2. أنواع الفرضيات Types of Hypotheses

وفى أثناء إجراء اختبار الفرضيات نستخدم نوعين من الفرضية هما:

-1 فرضية العدم Null Hypotheses H_0

وهى الفرضية التي تُبنى على أمل أن يُتخذ قرارٌ بعدم صحتها ويعبر عنها بالرمز H_0، وهى فرضية تنص دائما على عدم وجود فروق فى النتائج، وبمعنى آخر أن المتغير المستقل لا يؤثر فى المتغير التابع، أو أنه لا يوجد فرق إحصائى بين خصائص العينة وخصائص المجتمع الذي سحبت منه هذه العينة، وفرضية العدم هي التي يتم اختبارها إحصائياً وهى التي يتم رفضها أو قبولها، وفى ضوء هذا الرفض أو القبول يتحدد الموقف من الفرضية البديلة، ففى حالة رفض فرضية العدم نقبل الفرضية البديلة، أما إذا لم نستطع رفض فرضية العدم بالطرق الإحصائية فإننا نقرر رفض الفرضية البديلة (مجدي عبد الكريم حبيب، 1996).

ومثال ذلك أن نفرض قيمة وحيدة لمعلمات المجتمع مثل $\mu = 5$ بناءا على دراسة العينات، فهذا معناه أن المتوسط يساوى خمسة فقط ولا يساوى أية قيمة أخرى.

2- الفرضية البديلة Alternative Hypotheses H_1

وهي الفرضية الناتجة عن رفض فرضية العدم H_0، ويعبر عنها بالرمز H_1، وهي فرضية تنص دائما على وجود فروق في النتائج ترجع إلى المتغير المستقل، أو أن خصائص العينة التي نقوم بدراستها لا تعبر عن خصائص المجتمع الذي سحبت منه.

ومثال ذلك أن يكون لقيم معلمات المجتمع قيما كثيرة ومتعددة مثل $5 < \mu < 8$، فهذا معناه أن المتوسط يحتوى عدد لا نهاني من القيم المحصورة بين خمسة وثمانية وليس قيمة محددة.

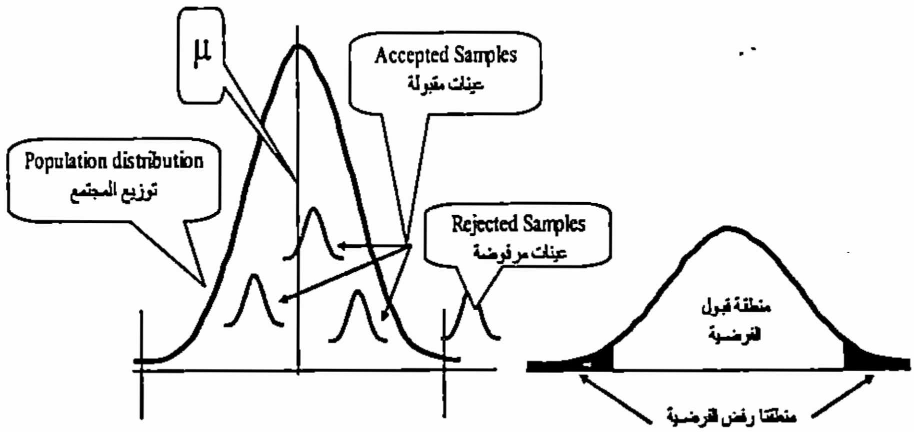
18.3. مفهوم اختبار الفرضيات Test of Hypotheses Concept

بعد هذا التقديم والتمهيد لموضوع إختبار الفرضيات Test of Hypothesis، نود أن نوضح المفهوم الذي بنى عليه هذا الموضوع، فنقول أنه طبقا ونظرية تقارب التوزيعات Central Limit Theory إذا كان لدينا مجتمعا يخضع للتوزيع الطبيعي Normal Population، وكانت قيمة متوسط هذا المجتمع $\mu = 57$، فإن توزيع متوسط العينات $\bar{x}$ Statistic Mean للعينات كبيرة الحجم سيكون قريبا جدا من منحنى التوزيع الطبيعي الذي متوسطه 57 وتباينه $s^2 = \frac{\sigma^2}{n}$، حيث n هي حجم العينة، وهذا يعني أنه بسحب عينة عشوائية Random Sample ذات حجم كبير من هذا المجتمع ، فإن هناك فرصا صغيرة جدا لوقوع وسطها الحسابي $\bar{x}$ على بعد أكثر من انحرافين معياريين 2σ أو ثلاثة انحرافات معيارية 3σ عن متوسط ذلك المجتمع، وهو في حالتنا 57، وطبقا وخواص منحنى التوزيع الطبيعي، فإن احتمال وقوع $\bar{x}$ ضمن انحرافين معياريين أي 2σ من المتوسط 57 هو 95.544% (يرجى مراجعة موضوع التوزيع الطبيعي في الفصل الثاني عشر) أي أن احتمال وقوع $\bar{x}$ على بعد أكثر من انحرافين معياريين سيكون صغيرا جدا ولا يصل إلى 5%، أما إذا حدث وحصلنا على قيمة متطرفة لمتوسط العينة Sample $\bar{x}$ Mean، بحيث كانت هذه القيمة على بعد أكثر من انحرافين معياريين من القيمة 57 فهناك احتمالين:

1. إما أن عارضا (حادثا) نادر الحدوث قد وقع، أى أننا اخترنا عينة من تلك العينات نادرة الحدوث وغير الممثلة للمجتمع ، وهذا خطأ في أخذ العينة. أو
2. أن متوسط المجتمع الذى سحبنا منه هذه العينة لا يساوى القيمة محل الاختبار والتي حددناها في فرضية العدم وهي $H_0 = 57$.

وسنستخدم الحالة الثانية (كون القيمة $\bar{x}$ بعيدة أكثر من انحرافين معياريين عن متوسط المجتمع وهى 57) كقاعدة لرفض فرضية العدم Null Hypotheses.

أى أن بداية الموضوع يكون لدينا مجتمعاً ذا قيم كثيرة يمكن تمثيلها بمنحنى توزيع طبيعى Normal Distribution، ونقوم بحساب (بفرض) قيمة لمعلمات هذا المجتمع نرغب فى أن تكون صحيحة، فنحاول أخذ عينات ممثلة لهذا للمجتمع ومعبرة عنه، ومنها نحسب إحصاءات العينة Sample Statistics المناسبة، ونقارنها بالقيمة المفروضة لمعلمات المجتمع Population Parameter، فإذا كان الفرق بينهما أكبر من $\pm3\sigma$ ، نستطيع الحكم بأن القيمة المفروضة لمعلمات المجتمع غير صحيحة، أما إذا كانتا متساويتين، أو قريبتين من بعضهما بقيمة لا تزيد عن $\pm3\sigma$ ، فنستطيع الحكم بأن القيمة المفروضة لمعلمات المجتمع قيمة صحيحة.



شكل رقم 18- 1 العلاقة بين العينة والمجتمع

والشكل 18-1 يوضح مناطق رفض و قبول فرضية العدم، وهى المناطق الأبعد من ثلاثة انحرافات معيارية $\pm3\sigma$.

وحيث إنه من المعلوم أن لمعلمات المجتمع قيمة وحيدة حقيقية غير معلومة، وستظل غير معلومة إلا إذا قمنا بقياس كل عناصر المجتمع، وهذا كما أشرنا غير ممكن وغير عملي، وعلى ذلك فإن القيم التي سنقوم بحسابها هي قيم إحصاءات العينة ، والقيم التي سنقوم بمحاولة استنتاجها هي قيم معلمات المجتمع، وقبل أن نتخذ قراراً مبنياً على هذه القيم التي تم الحصول عليها، فيجب أولا أن نتأكد أنها قيم صحيحة.

فمثلا لو أعلنت إحدى شركات الأدوية أن متوسط وزن زجاجة دواء تقوم بإنتاجها هو 50 جراما، وأردنا اختبار صحة هذا الادعاء وهذا الفرض، فمن المنطقي أن نحاول إثبات أن متوسط مجموعة عينات يساوى 50 جراما بالضبط، فإذا صح الفرض نكون متأكدين أن متوسط الزجاجات (المجتمع) يساوى 50 جراما، وإذا لم نستطع إثبات صحة الفرض فنستنتج أن متوسط الزجاجات (المجتمع) لا يساوى 50 جراما، هذا بدلا من أن نجرى محاولات عديدة ولا نهائية لإثبات أن متوسط العينات لا يساوى 50، فنحاول إثبات أنه يساوى 49 أو 48.2 أو 51.3 أو 50.6 أو.... أو... إلى مالا نهاية.

والفرض الأول فى هذه الحالة وهو أن المتوسط يساوى 50 هو فرضية العدم Null Hypotheses ويعبر عنه بالمعادلة $H_0 = 50$ ، و الفرض الثاني فى هذه الحالة وهو أن المتوسط لا يساوى 50 هو الفرضية البديلة Alternative Hypotheses ويعبر عنه بالمعادلة $H_1 \neq 50$.

18.4. مستوى الدلالة α فى اختبار الفرضيات Level of Significance for Hypotheses

أشرنا إلى أن فرضية العدم التي تنص دائما على عدم وجود فروق فى النتائج، وبمعنى آخر أنها تشير إلى أن المتغير المستقل لا يؤثر فى المتغير التابع، أو أنه لا يوجد فرق بين خصائص العينة وخصائص المجتمع الذي سحبت منه هذه العينة، وعند اختبار هذه الفرضية بهدف قبولها أو رفضها، فإننا فى الحقيقة نختبر مدى صحة هذا الافتراض بحساب احتمال الحصول على هذه النتيجة عن طريق الصدفة، فإذا كان هذا الاحتمال مساويا أو أقل من مستوى احتمال معين فإننا نرفض فرضية العدم، ويسمى هذا الاحتمال بمستوى الدلالة ألفا α ، ومن الطبيعي فإن رفض فرضية العدم يؤدى إلى قبول الفرضية البديلة، وحينئذ نقول بأن النتائج ذات دلالة إحصائية، أما إذا حصلنا على مستوى احتمال آخر أكبر من مستوى الاحتمال -الدلالة – المحدد فإن استنتاجنا أننا أخفقنا فى رفض فرضية العدم ، وبالتالى عدم قبول الفرضية البديلة.

18.5. أنواع الأخطاء فى اختبار الفرضيات Test of Hypotheses Error Types

ذكرنا أننا نختبر فرضية العدم بهدف قبولها أو رفضها عند مستوى دلالة معين (درجة تأكد معينة)، وهذا يعنى أننا معرضون دائما للخطأ ما دام رفض الفرضية أو قبولها كان احتمالا ولم يكن يقينا مائة فى المائة، وبفرض أن مستوى الدلالة (نسبة احتمال) هو 5%، فمعنى ذلك أننا لو كررنا تجربة ما مائة مرة، فإن النتيجة التي نحصل عليها ستكون عن طريق الصدفة خمس مرات β ، لذا فإننا لا ندرى حقيقة إن كانت النتيجة التي حصلنا عليها برفضنا لفرضية العدم عند مستوى الدلالة $\alpha = 0.05$ واحدة من هذه المرات الخمسة أم لا، أى أننا لا ندرى هل وقعنا فى الخطأ أم لا؟ ، ولذا فإن معظم الاختبارات تتم عند مستوى دلالة $\alpha = 0.05$ أى بدرجة ثقة وبنسبة تأكد 95%.

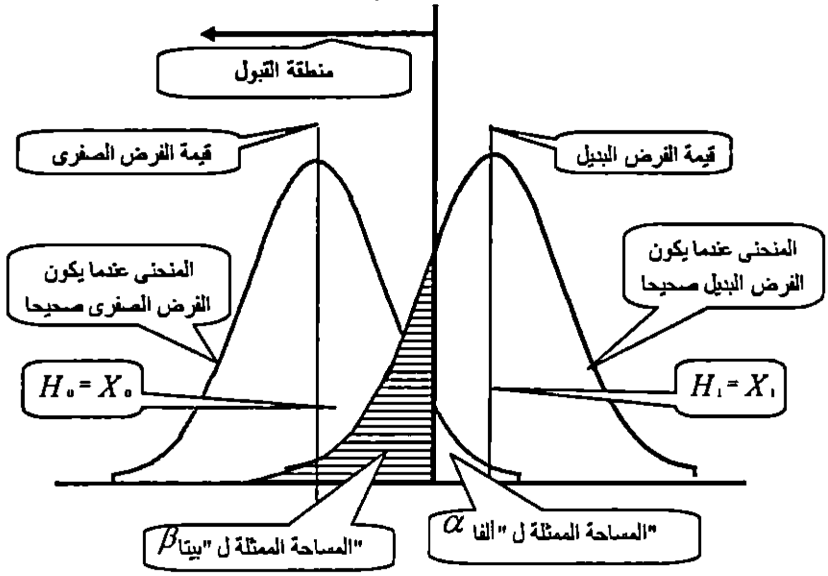
وفى الحقيقة يوجد نوعان من الخطأ عند اختبار الفرضيات وهما :

18.6. الخطأ من النوع الأول Type I Error

ويحدث عندما نرفض فرضية العدم Null Hypotheses H_0 عندما تكون هذه الفرضية صحيحة ، أى عندما نقرر وجود اختلاف وتباين فى حين أنه لا يوجد أصلا، أو أن التباين الموجود ناتج عن الصدفة، أو أنه تباين الطبيعي Normal Variation، ولهذا النوع من الخطأ نسبة احتمال تسمى α، وهى تعبر عن نسبة احتمال رفض فرضية العدم H_0 بالرغم من صحته، وتتراوح قيمة مستوى الدلالة ألفا α بين 0.0001 إلى 0.1، أى من 0.01% حتى 10% طبقا ونوعية الدراسة وأهميتها والتكاليف المرتبطة بها، و الشائع منها هو 5%.

أ- الخطأ من النوع الثاني Type II Error

ويحدث عندما لا نستطيع رفض فرضية العدم Null Hypotheses H_0 عندما تكون هذه الفرضية غير صحيحة، (أى أنه يوجد تباين حقيقي فى حين أننا لا نستطيع إثبات وجود هذا التباين)، ولهذا النوع من الخطأ نسبة احتمال تسمى β، وهى تعبر عن نسبة احتمال عدم رفض فرضية العدم H_0 بالرغم من عدم صحته، وتتراوح قيمته بين 90% و 95%.



شكل رقم 18- 2 العلاقة بين ألفا ، وبيتا

ولتوضيح ذلك هندسيا كما بالشكل 18-2 نرسم منحنى التوزيع فى حالتين، الحالة الأولى عندما تكون فرضية العدم صحيحة، أى يكون التوزيع معتدلا طبيعيا ومتوسطه $H_0 = X_0$ ، والحالة الثانية عندما يكون الفرض البديل صحيحا، أى يكون التوزيع معتدلا طبيعيا ومتوسطه $H_1 = X_1$ ، ويتضح أن بعض العينات التى تنتمى إلى منحنى الفرض البديل يحمتل أن تقع متوسطاتها فى منطقة القبول لمنحنى فرض العدم ، ونسبة هذه العينات هى نسبة الجزء من توزيع الفرض البديل الذى يشترك مع توزيع فرضية العدم فى منطقة القبول، وهذه النسبة هى احتمال قبول فرضية العدم عندما يكون الفرض البديل هو الصحيح، أى هو الاحتمال بيتا β للخطأ من النوع الثانى.

والجدول 18-1 يوضح الفرق بين نوعى الخطأ والمعبر عنهما بألفا α وبيتا β.

جدول رقم 18-1 الفرق بين الخطأ من النوع الأول و الخطأ من النوع الثانى

		الفرضية الصفرية H_0 صحيحة	الفرضية الصفرية H_0 ليست صحيحة
القرار	لا توجد أدله لرفض الفرضية الصفرية (تقبل الفرضية الصفرية)	قرار صائب	خطأ من النوع الثانى ويعبر عنه بيتا Type II Error
	نرفض الفرضية الصفرية	خطأ من النوع الأول ويعبر عنه بالفا Type I Error	قرار صائب

والوقوع فى الخطأ من النوع الثاني أشد خطراً من الوقوع فى الخطأ من النوع الأول، إذ فى الخطأ من النوع الثاني Type II Error والذي يشار إلي احتماليته بالرمز بيتا β، يكون الفرض بالعدم خطأ ونقرر أنه صحيح، أى أنه توجد مشكلة حقيقية ولا نراها ولا نتخذ قراراً بحلها، بينما فى الخطأ من النوع الأول Type I Error والذي يشار إلى احتماليته بالرمز ألفا α، يكون فرض العدم صحيحا ونقرر أنه خطأ، أى أنه لا توجد مشكلة حقيقية ونقرروجود مشكلة ونتدخل فى العملية لمحاولة حلها، وهذا ما نطلق عليه العبث بالعملية Tempering، ويمكن توضيح ذلك من المثال التالي.

بفرض أن أحد مرضى القلب استشار طبيبه، فقام بعمل الفحوصات الطبية اللازمة، وكان فرض العدم Null Hypotheses H_0 هو عدم وجود مرض، وبالتالي كانت الفرضية البديلة H_1 Alternative Hypotheses هو وجود مرض، وعلى ذلك فإن α ستعبر عن احتمال أن تشير نتائج الفحوص إلى وجود مرض فى حين أنه لا يوجد مرض، وكذلك فإن β ستعبر عن احتمال أن تشير نتائج الفحوص إلى عدم وجود مرض فى حين أنه بالفعل يوجد مرض.

فى هذا المثال يتضح أن اتخاذ قرار بوجود مرض، وما يتبع ذلك من عناية أكبر بهذا المريض وإعطاءه دواءً بالرغم من عدم وجود مرض، وهو الخطأ من النوع الأول Type I Error معبراً عنه بالرمز ألفا α، يعتبر أقل خطورة من اتخاذ قرار بعدم وجود مرض، وما يتبع ذلك من إهمال للمريض وعدم إعطاءه دواءً بالرغم من وجود مرض حقيقى قد يودى بحياة المريض.

18.7. قوة الاختبار Power of The Test

قوة الاختبار هي احتمال تجنب الوقوع فى الخطأ من النوع الثانى، أى اتخاذ قرار بقبول فرضية العدم بالرغم من كونها خاطئة، وتتراوح قيمته بين الصفر والواحد، ويلاحظ أنه وثيق الصلة بالخطأ من النوع الثاني Type II Error فكلما زاد حجم هذا الخطأ قلت قوة الاختبار، ويعبر عنها بالقيمة $(\beta-1)$.

18.8. استخدامات اختبار الفرضيات Test of Hypotheses Usages

يستخدم هذا الاختبار فيما يلي:

1-الحكم إحصائيا بأن العينة المنتقاه تتبع المجتمع محل الدراسة ام لا؟ فإذا كان الفرق بين متوسط العينة ومتوسط المجتمع بسيطا (فى حدود $\pm 3\sigma$) فإنه يمكننا التنبؤ والحكم بأن العينة تتبع المجتمع محل الدراسة.

2-الحكم إحصائيا على مدى تحسن العملية نتيجة إجراء تعديلات أو نتيجة جهود تحسين من عدمه، فالزيادة الملحوظة فى متوسط مخرجات خط الإنتاج تعنى تحسن هذه العملية ، وأن النقص الملحوظ فى التباين يعنى كذلك تحسن تلك العملية، وهكذا.

3-الحكم إحصائيا بوجود تباين بين أداء أكثر من عملية أو مناوبة (وردية)، أو مصنع، أو خط إنتاج أو.....، و بيان ما إذا كان هذا التباين يرجع إلى عامل الصدفة وأن وجوده حتمي وناشئ عن الأسباب العامة للتغير Common Cause Variation وأنه فى حدود مدى التباين الطبيعي للعملية ، أم أنه اختلاف حقيقي نتيجة تباين حقيقي فى تلك العملية.

18.9. أنواع اختبارات الفرضيات Types of Hypotheses Tests

يوجد العديد من الطرق التي تستخدم لاختبار الفرضيات، يوضحها الجدول 18-2 :

جدول رقم 18- 2 الطرق المختلفة لاختبار الفرضيات

		Data type نوع البيانات	Hypotheses test اختبار الفرضيات	Purpose الهدف
averages المتوسطات	1	continuous مستمرة	t- test اختبار "تى"	Compare two groups averages مقارنة بين متوسطات مجموعتين
	2		Paired T- test اختبار "تى" المتزاوج	Compare two groups averages when data is matched مقارنة بين متوسطات مجموعتين متزاوجتين
	3		ANOVA تحليل التباين	Compare two or more groups averages مقارنة بين متوسطات مجموعتين أو اكثر
variances التباين	4		F test for equal variances الاختبار الفائى عند تساوى التباين	Compare two or more groups variances مقارنة بين تباين مجموعتين أو أكثر
proportions النسبة	5	discrete متقطعة	Chi-square test إختبار كاى تربيع	Compare two or more groups proportions (counts or frequencies) مقارنة بين نسبة مجموعتين أو أكثر

وفى هذا الجدول نلاحظ الآتي:

1-أن للثلاثة أنواع الأولى تتعرض لدراسة المتوسط ، وتستخدم للمقارنة بين متوسطات مجموعات البيانات المتصلة أو المستمرة Continuous Data.

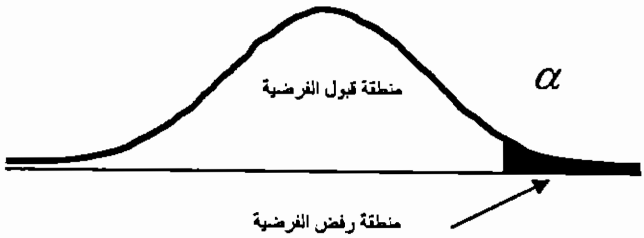
2-النوع الرابع ويتعرض لدراسة التباين ويستخدم للمقارنة بين مجموعة من البيانات المتصلة أو المستمرة Continuous Data.

3-للنوع الخامس ويتعرض لدراسة النسبة Proportion ويستخدم للمقارنة بين مجموعة من النسب ويستخدم مع البيانات من النوع المتقطع أو الوثاب Discrete.

18.10. تحديد مناطق رفض الفرضيات Rejection Area Determination

يعتمد موضوع تحديد منطقة الرفض على طبيعة العنصر أو المعامل محل الدراسة كما يلي:

1-تكون منطقة الرفض فى جهة واحدة من منحنى التوزيع الطبيعي، ويطلق عليه إختبار فى جهة واحدة One Sided Test، إذا كانت معلمات المجتمع المراد اختبارها أكبر من قيمة محددة، مثل متوسط الأعمار لا يقل عن 40 سنة ، وحينئذ تكون α بالكامل فى جهة واحدة كما فى شكل 18-3.



شكل رقم 18- 3 منطقة الرفض فى جهة واحدة

2-تكون منطقة الرفض فى جهة واحدة من منحنى التوزيع الطبيعي، ويطلق عليه إختبار فى جهة واحدة One Sided Test أيضا، ولكن معلمات المجتمع المراد اختبارها تكون أقل من قيمة محددة، مثل متوسط الأعمار لا يزيد عن 20 سنة ، وحينئذ تكون α بالكامل فى جهة واحدة أيضا كما فى شكل 18-4.

شكل رقم 18- 4 منطقة الرفض فى جهة واحدة

3-تكون منطقة الرفض فى جهتين من منحنى التوزيع الطبيعي، ويطلق عليه إختبار فى جهتين Two Sided Test، إذا كانت معلمات المجتمع المراد اختبارها محصورة بين قيمتين محددتين، مثل متوسط الأعمار لا يقل عن 40 سنة ولا يزيد عن 20 سنة ، وحينئذ تكون α فى جهتين ومساحة الرفض فى كلا الجهتين تساوى نصف مساحة الرفض الكلية كما فى شكل 18-5.

شكل رقم 18- 5 منطقة الرفض فى جهتين إختبار فى جهتين

18.11.شروط إجراء اختبارات الفرضيات Assumption For Testing Hypotheses

عند إجراء اختبارات الفرضيات يجب توخى الحذر الشديد والتأكد من الشروط التالية:

1-لا يتم إجراء اختبار المتوسط والتباين إلا على بيانات من النوع المستمر Continuous Data.

2-يجب أن تخضع البيانات المتاحة لمنحنى التوزيع الطبيعي Normally Distributed Data.

3-يجب أن تكون العينات عشوائية وغير متحيزة ومستقلة Random & Unbiased & Independent Samples.

4-عند إجراء الاختبار لمقارنة الاختلافات الموجودة في مجموعة من العمليات فلابد من توافر الشروط التالية:

- أن تكون العمليات متزنة Stable Processes.
- أن لا تحتوى العملية على أسباب خاصة Special Causes تؤدى إلى عدم الإتزان.
- أن تكون العينات معبرةً عن المجتمع ومن ثم عن العملية.

وحتمية تحقق الشروط السابقة قبل إجراء اختبار الفرضيات Test of Hypotheses هو ما يجعله أداة من الأدوات المتقدمة، إذ أن تحقق مثل هذه الشروط يستلزم أن تكون العملية محل الدراسة عملية متزنة ومنتظمة Standardized Process وهذا يحتاج إلى مجهود مضني و كبير.

18.12.خطوات إجراء اختبارات الفرضيات Test of Hypotheses Procedure

يستطيع المرء بشيء من التركيز جعل عملية إجراء اختبار الفرضيات عملية روتينية و بسيطة وسهلة، وذلك بإتباع الخطوات التالية:

1. افترض كلا من فرضية العدم H_0 وكذلك الفرضية البديلة H_1، وإذا كان الهدف هو اختبار ما إذا كانت معلمات المجتمع تساوى قيمة محددة أم لا، فدائما افترض أن H_0 يساوى القيمة المراد اختبارها، وأن H_1 لا يساوى القيمة المراد اختبارها.

2. حدد قيمة α، ثم حدد منطقة الرفض تحديدا دقيقا وكذلك حدود هذه المنطقة (القيم الحرجة).

3. إذا كان حجم العينة أكثر من 30 استخدم جداول منحنى التوزيع الطبيعي المعياري Z، و إذا كان حجم العينة أقل من 30 استخدم جداول المنحنى t، وذلك لحساب حدود منطقة الرفض.

4. حدد إحصاءات العينة المناظرة لقيمة المجتمع المعطاة (فمثلا إذا كانت μ معطاة، فإننا سنحسب $\bar{x}$) واحسب قيمتها إما باستخدام برامج الكمبيوتر، أو يدويا من المعادلات الموضحة بالجدول 18-3 :

جدول رقم 18- 3 بعض معادلات اختبارات الفرضيات

One sample عينة واحدة	Mean المتوسط	Large sample size عينات كبيرة الحجم	$n \geq 30$	$Z = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \approx \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$
		Small sample size عينات صغيرة الحجم	$n < 30$	$t = \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \approx \frac{\bar{x}-\mu_0}{s/\sqrt{n}}$
	Proportion النسبة			$Z = \frac{\bar{p}-p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$
	Variance التباين			$\chi^2 = \frac{(1-n)S^2}{\sigma_0^2}$
Two samples عينتان	Difference between two means الفرق بين متوسطين	Large sample size عينات كبيرة الحجم	$n \geq 30$	$Z = \frac{\bar{x}-\bar{y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$
		Small sample size عينات صغيرة الحجم	$n < 30$	$t = \frac{\bar{x}-\bar{y}}{S * \sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$
	Difference between two Proportions الفرق بين نسبتين			$Z = \frac{\bar{p}_1-\bar{p}_2}{\sqrt{\frac{\bar{p}(1-\bar{p})}{n_1}+\frac{\bar{p}(1-\bar{p})}{n_2}}}$
	Comparison between two Variances الفرق بين تباينى مجتمعين			$\left(F\left[\frac{\alpha}{2};n_2-1,n_2-1\right],F\left[1-\frac{\alpha}{2};n_2-1,n_2-1\right]\right)$

5. إذا تم الحساب يدويا، قارن بين القيمة المحسوبة لإحصاءة العينة، والتي تم الحصول عليها من الخطوة رقم 4، وبين القيمة الحرجة التي تم الحصول عليها

من الخطوة رقم 2، ومنها اتخذ القرار المناسب، أى ارفض أو اقبل فرضية العدم.

6. أما إذا تم الحساب باستعمال الكمبيوتر فإننا نحدد قيمة المعامل "بى" P-Value والتي تمثل المساحة المحصورة بين القيمة الحرجة وتقع تحت نهاية المنحنى وتتراوح قيمتها بين الصفر والواحد الصحيح، فإذا كانت أقل من 0.05 كان ذلك دليلاً على وجود اختلاف إحصائى وتباين بين القيمة المفروضة والقيمة المحسوبة، أى أننا داخل منطقة الرفض، أى يمكننا رفض فرضية العدم ، وحينها يكون التعليق على هذه النتيجة هو"نرفض فرضية العدم حيث إن الاختلاف والتباين فى البيانات معنويٌ وذو قيمة حقيقية وملموسة We Reject the Null Hypotheses and the Difference is Statistically Significant"، أما إذا كانت القيمة بى P- Value أكبر من 0.05 كان ذلك دليلاً على عدم وجود اختلاف إحصائى بين القيمة المفروضة والقيمة المحسوبة (أى أننا داخل منطقة القبول وخارج منطقة الرفض)، أى لا يمكننا رفض فرضية العدم ، وحينها يكون التعليق على هذه النتيجة هو أنه "لا يوجد دليل قاطع لرفض فرضية العدم حيث إن التباين والاختلاف فى البيانات المتاحة ليس ذى معنوية أو ليس ملموسا "There are no Significant Evidences to Reject the Null Hypotheses and the Difference is Not "Statistically Significant

18.13. إجراء اختبارات الفرضيات بالكمبيوتر Testing Hypotheses With Computer

فى الصفحات القليلة السابقة تعمدنا شيء من الإسهاب فى الشرح والتوضيح بالمعادلات، ولكننا عمليا لن نلجأ إلى استخدام هذه المعادلات إلا قليلا، وإنما سيقوم بذلك العديد من برامج الكمبيوتر، فيمكن استخدام البرامج الإحصائية المتخصصة مثل المينيتاب Minitab (او Excel فى بعض الحالات) فى إجراء كافة الاختبارات الموضحة بجدول رقم 18-4:

جدول رقم 18 -4 الاختبارات التى يمكن اجراؤها بالكمبيوتر

Single Factor معامل واحد	ANOVA اختبار الأنوفا
Two Factor With Replication	
Two Factors Without Replication	
Paired Two Sample For Means	T-Test اختبار تى
Two Samples Assuming Equal Variances	
Two Samples Assuming Unequal Variances	
اختبار تباين عينتان	Two Sample F Test For Variance
الانحدار	Regression
اختبار متوسط عينتان	Z Test Two Sample For Means

18.14. أمثلة على اختبار الفرضيات Examples On Hypotheses Testing

مثال رقم 18- 1

قام فريق الجودة بإحدى الشركات بعمل دراسة لتحسين عملية إنتاج أحد أجزاء السيارة، والجزء الذي خضع للدراسة هو جزء إسطواني قطره الخارجي يساوى 5.25 سم، وكانت القيم المسجلة للإنتاج السابق للقطر تشير إلى أن متوسطه 5.23 سم، وبعد الدراسة وتطبيق بعض التعديلات، قام الفريق بإبلاغ الإدارة أنه قد حدث تحسن ملموس فى عملية الإنتاج، وبناءً على ذلك أمرت الإدارة بأخذ عينات من المنتج بعد التعديل وكانت القراءات الموضحة بالجدول 18-5.

جدول رقم 18-5 بيانات المثال 18-1

Diameter	5.26696	5.24524	5.24037	5.2588	5.25141	5.24321	5.23441	5.2328	5.25287	5.26026
Sample	1	2	3	4	5	6	7	8	9	10

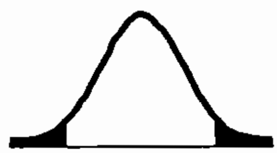
والمطلوب تحديد ما إذا كانت التعديلات التي أجريت على عملية الإنتاج قد أدت بالفعل إلى تحسن ملموس أم لا؟

الحل:

نطبق الخطوات العامة لاختبار الفرضيات كما يلي :

1-حيث إن القيمة المسجلة للإنتاج السابق للقطر هي 5.23 سم، وحيث إن هدف مشروع التحسين كان تغيير هذه القيمة إلى القيمة القياسية 5.25، فإننا سنفترض أن $H_0 = 5.23$ وكذلك نفترض أن $H_1 \neq 5.23$

2-نحدد قيمة $\alpha = 0.05$، ونحدد منطقة الرفض بانها إختبار من الجهتين Two Sided Test كما في شكل 18-6.



شكل رقم 18- 6 مناطق الرفض

3-حيث إن معلمات المجتمع المعطاة هي المتوسط فإننا سندرس متوسط العينة .

4-حيث إنه قد تم أخذ عينه لمتغير واحد فإننا سنستخدم إختبار عينة واحدة One Sample Test.

5-حيث إن حجم العينة أصغر من 30 عينه، فإننا سنستخدم الإختبار تى T-Test.

الحل اليدوي (سنعمد الى الحل اليدوى للتوضيح فقط ولكننا سنعتمد على برامج المينيتاب Minitab والإكسيل Excel فيما بعد) :

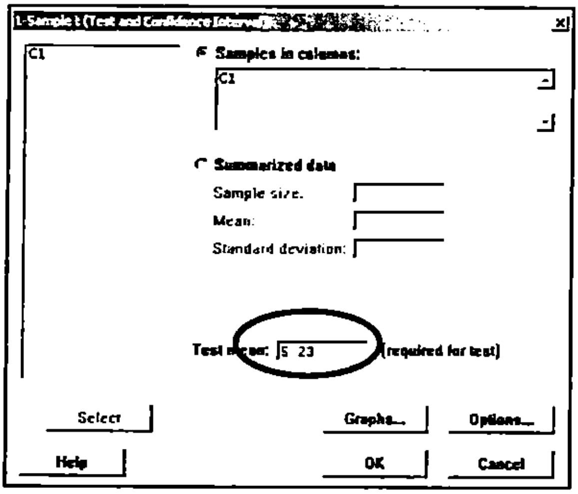
- من البيانات المعطاة،يمكننا حساب قيمة المتوسط=5.24863 سم، والانحراف المعياري=0.01136
- نحدد القيمة الحرجة من الجدول T في نهاية هذا الكتاب فنجد أنها تساوى 2.262.
- نحسب قيمة T من المعادلة

$$t = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \approx \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{5.24863 - 5.23}{0.01136 / \sqrt{10}} = 5.186024$$

وحيث إن T المحسوبة أكبر من قيمة القيمة الحرجـة ، إذن فهي تقع فـى منطقـة الرفض، أى أننا نرفض فرضية العدم ، أى أن المتوسط قد تباين وقد يرجع ذلك إلى جهود التحسين، ومـن الواضح أنه تحسن أدى إلى اقتراب المتوسط من Target=5.25.

الحل بالمينيتاب Minitab

ندخل البيانات المتاحة إلى المينيتـاب Minitab، ومـن القائمـة Stat>Basic Statistics > One Sample T Test



شكل رقم 18- 7 حل المثال 17-1 بالمينيتاب

يظهر الشـكل 18-7 ، ثـم نـذهب إلـى نافـذة النتـائج فتظهـر بعـض القـيم ومنهـا نجـد أن قيمـة المتوسط Mean=5.24863، وأن القيمة بى P=0.001، وهـى قيمـه أقـل مـن 0.05 أى أنهـا تقع فى منطقة الرفض، وبالتالي نرفض فرض العدم، ونقر بأنه قد حدث تباين للعملية ، ومـن الواضح أنه تحسن أدى إلى اقتراب المتوسط من Target=5.25

One-Sample T: C1

Test of mu = 5.23 vs not = 5.23

Variable	N	Mean	StDev	SE Mean	95% CI	T	P
C1	10	5.24863	0.01136	0.00359	(5.24050; 5.25676)	5.19	0.001

مثال رقم 18- 2

فى المثال السابق ماذا لو كان عدد القراءات التي تم تسجيلها خمسون قراءة كما فيجدول 6-18 :

جدول رقم 18-6 بيانات المثال 18-2

5.29526	5.25731	5.25896	5.27212	5.2249	5.26064	5.24432	5.2476	5.23671	5.25977
5.22149	5.27735	5.2075	5.24348	5.24007	5.22408	5.26996	5.25795	5.26591	5.24063
5.23892	5.25204	5.22508	5.27211	5.24665	5.27934	5.23501	5.2573	5.2326	5.25517
5.27799	5.27368	5.24421	5.26515	5.2578	5.24291	5.23678	5.25446	5.26788	5.23241
5.25382	5.24264	5.24875	5.23584	5.29011	5.2262	5.24879	5.27732	5.29461	5.26922

الحل بالجدول:

نطبق الخطوات العامة لاختبار الفرضيات كما يلي

1-حيث إن القيمة الحالية للقطر هي 5.23 سم، وحيث إن هدف مشروع التحسين كان تغيير هذه القيمة إلى القيمة القياسية 5.25، فإننا سنفترض أن $H_0 = 5.23$ وكذلك نفترض أن $H_1 \neq 5.23$

2-نحدد قيمة $\alpha = 0.05$ ، ونحدد منطقة الرفض بأنها إختبار من الجهتين Two Sided Test كما فى الشكل 18-8.

شكل رقم 18- 8 مناطق الرفض للمثال 18-2

3-حيث إن معلمات المجتمع المعطاة هي المتوسط فإننا سندرس متوسط العينة .

6-حيث إنه قد تم أخذ عينه لمتغير واحد فإننا سنستخدم إختبار عينة واحدة One Sample Test.

4-حيث إن حجم العينة أكبر من 30 عينه، فإننا سنستخدم الإختبار "زد" Z-Test

الحل :

- من البيانات المعطاة، نجد أن المتوسط=5.2528 سم، والانحراف المعياري=0.01975

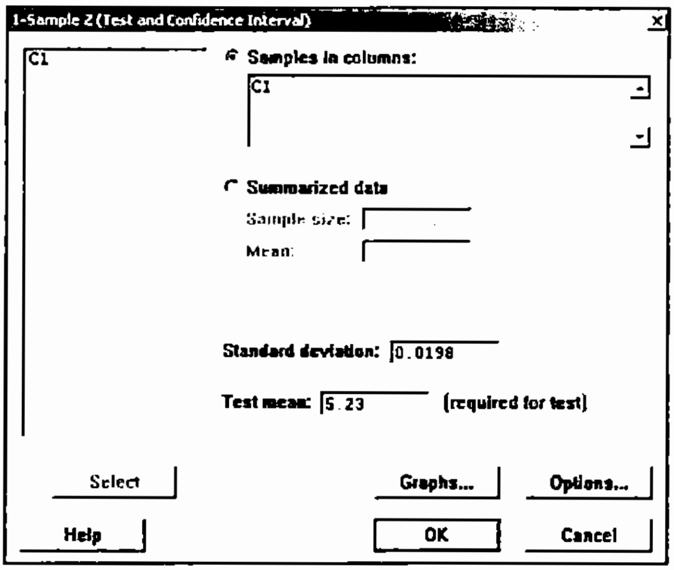
- نحدد القيمة الحرجة من الجدول Z فى نهاية الكتاب وهى 1.65 عند المساحة 0.975
- نحسب قيمة Z من المعادلة

$$Z = \frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \approx \frac{\bar{x} - \mu_0}{s / \sqrt{n}} = \frac{5.25282 - 5.23}{0.01975 / \sqrt{10}} = 3.65$$

وحيث إن Z المحسوبة أكبر من قيمة القيمة الحرجة ، إذن فهي تقع فى منطقة الرفض، أى أننا نرفض فرضية العدم ، أى أن المتوسط قد تغير، ونقر بأنه قد حدث تباين للعملية، ومن الواضح أنه تحسن أدى إلى اقتراب المتوسط من الهدف.

الحل بالمينيتاب Minitab ندخل البيانات المتاحة إلى Minitabl، ومن القائمة Stat>Basic Statistics > One Sample Z Test

فيظهر الشكل 18-9.



شكل رقم 18- 9 خطوة لحل المثال 18-2

نذهب إلى نافذة النتائج فتظهر بعض القيم، ومنها نجد أن قيمة المتوسط Mean=5.25282، وأن القيمة بى P=0.00 وهى قيمه أقل من 0.05 وبالتالي نرفض الفرضية الصفرية، ونقر بأنه قد حدث تباين للعملية ، ومن الواضح أنه تحسن أدى إلى اقتراب المتوسط من الهدف.

"We reject the null, and the difference is statistically significant"

```
One-Sample Z: C1

Test of mu = 5.23 vs not = 5.23
The assumed standard deviation = 0.0198

Variable   N    Mean     StDev   SE Mean        95% CI              Z      P
C1        50  5.25282  0.01975  0.00280  (5.24733; 5.25830)  8.15  0.000
```

مثال رقم 18- 3

سجلت القراءات التالية كما فى جدول 18-7 أداء مجموعتين من الموظفين الذين يقضون وقتا إضافيا فى العمل

جدول رقم 18-7 بيانات المثال 18-3

Group 1	5.25382	5.24264	5.24875	5.2358	5.29011	5.2262	5.24879	5.2773	5.29461	5.26922
Group 2	6.56882	6.61608	6.59907	6.5888	6.62624	6.60912	6.60632	6.6158	6.59436	6.5813

والمطلوب بحث ما إذا كان هناك اختلاف بين المجموعتين أم لا؟ مع اعتبار أن الانحراف المعياري للمجموعتين متساو.

الحل: نطبق الخطوات العامة لاختبار الفرضيات كما يلي

1-حيث إن البيانات المتاحة هي لمجموعتين، فإن فرضية العدم ستكون أن نفترض أن متوسط المجموعتين متساو، وستكون الفرضية البديلة أن نفترض أن متوسط المجموعتين غير متساو.

2-نحدد قيمة $\alpha = 0.05$، ونحدد منطقة الرفض باتها إختبار من الجهتين Two Sided Test

3-حيث إن معلمات المجتمع المعطاة هي المتوسط فإننا سندرس متوسط العينة .

4-حيث إنه قد تم أخذ عينه لمتغيرين فإننا سنستخدم Two Sample Test.

5-حيث إن حجم العينة أصغر من 30 عينه، فإننا سنستخدم الT-Test

الحل اليدوي : من البيانات المعطاة نحدد القيمة الحرجة من الجدول T فى نهاية الكتاب وهى 2.262

نحسب قيمة T من المعادلة

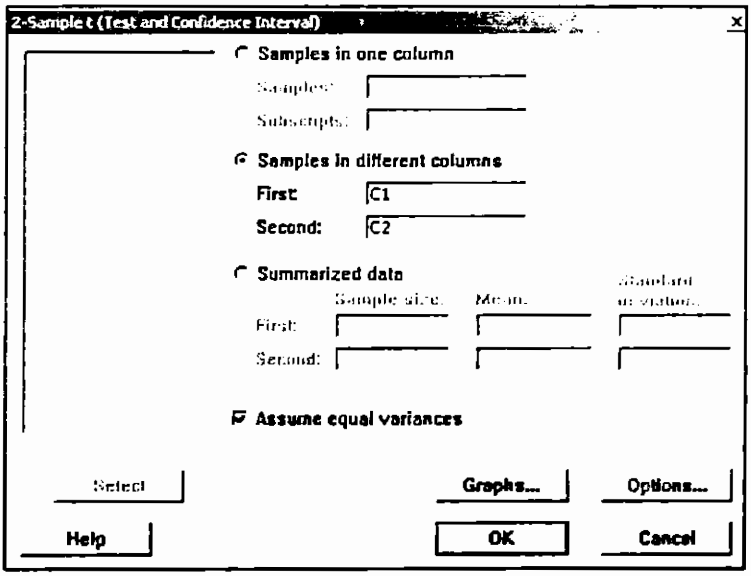
$$t = \frac{\bar{x} - \bar{y}}{S * \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

ونقارن هذه القيمة المحسوبة بالقيمة الحرجة ، فإذا وقعت في منطقة الرفض، فإننا نرفض الفرضية الصفرية، ويكون متوسطي المجموعتين مختلفين.

الحل بالمينيتاب Minitab

ندخل البيانات المتاحة إلى المينيتاب Minitab، ومن القائمة Stat>Basic Statistics > Two Sample T Test

يظهر الشكل 18-10.



شكل رقم 18- 10 خطوة لحل المثال 3-18

ثم نذهب إلى نافذة النتائج فتظهر القيم التالية، ومنها نجد أن القيمة بي P=0.00-وهي قيمه أقل من 0.05 وبالتالي نرفض الفرضية الصفرية، ونقر بأن متوسط المجموعتين غير متساويين.

"We Reject the Null Hypotheses، and the Difference is Statistically Significant"

```
Two-sample T for C1 vs C2

     N    Mean   StDev  SE Mean
C1  10  5.2587  0.0231   0.0073
C2  10  6.6006  0.0176   0.0056


Difference = mu (C1) - mu (C2)
Estimate for difference:  -1.34186
95% CI for difference:  (-1.36114; -1.32257)
T-Test of difference = 0 (vs not =): T-Value = -146.16  P-Value = 0.000  DF =
     18
Both use Pooled StDev = 0.0205
```

مثال رقم 18- 4

فى محاولة لإحدى شركات الصيانة تقليل الوقت اللازم لصيانة إحدى المعدات الهامة، قام فريق من مهندسي الصيانة المختصين بتعديل الإجراءات اللازمة للبدء فى تنفيذ أعمال صيانة هذه المعدة، وذلك بما لا يتعارض مع تعليمات الجهة المصنعة، وقام الفريق بتسجيل الوقت المستغرق بالطريقة القديمة، وكذلك الوقت المستغرق بالطريقة الجديدة بالدقائق وكانت ما بالجدول 18-8:

جدول رقم 18-8 بيانات المثال 18-4

trial	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Before	42.0	41.2	39.0	50.6	39.9	51.5	48.9	50.9	39.2	48.0	45.3	28.0	46.7	29.0	53.6
After	41.8	45.4	46.8	39.8	41.2	38.2	40.6	42.1	32.8	38.9	32.5	39.7	45.1	39.8	44.4

والمطلوب تقييم ما إذا كانت الطريقة الجديدة والمعدلة قد أحدثت انخفاضا ملحوظا فى زمن تنفيذ أعمال الصيانة أم لا؟ وهل نتبع هذه الطريقة الجديدة أم نتمسك بالطريقة القديمة؟

الحل:

نطبق الخطوات العامة لاختبار الفرضيات كما يلي

1-حيث إنه توجد قراءتين لنفس العملية فإن هذه البيانات تسمى Paired Data وحيث إن هدف مشروع التحسين كان تغيير متوسط الزمن اللازم لإجراء الصيانة، فإن فرضية العدم هى أن متوسط القراءات قبل التحسين يساوى متوسط القراءات بعد التحسين، وبالتالي تكون الفرضية البديلة هي أن متوسط القراءات قبل التحسين لايساوى متوسط القراءات بعد التحسين.

2-نحدد قيمة $\alpha = 0.05$ ، ونحدد منطقة الرفض بأنها Two Sided Test

3-حيث إن معلمة المجتمع المعطاة هي المتوسط فإننا نختار دراسة متوسط العينة .

4-حيث إنه قد تم أخذ عينه لمتغير واحد بقيمتين فإننا سنستخدم Paired T Test.

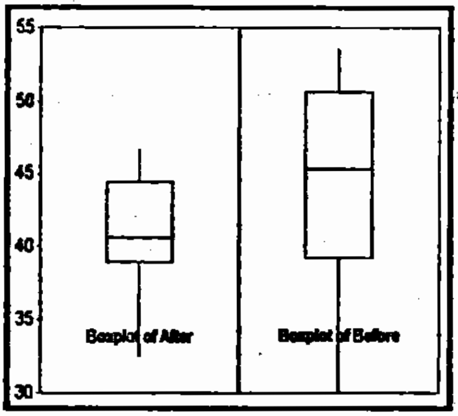
5-حيث إن حجم العينة أصغر من 30 عينه، فإننا سنستخدم الT-Test

الحل اليدوي من البيانات المعطاة، نجد أن هناك اختلاف فى متوسط الطريقتين فالمتوسط قبل التعديل يساوى 40.6 دقيقة، و المتوسط بعد التعديل يساوى 43.6 دقيقة.

Descriptive Statistics: After; Before

Variable	Mean	Minimum	Median	Maximum	Range
After	40.60	32.47	40.62	46.75	14.28
Before	43.60	28.03	45.33	53.61	25.58

برسم مخطط الصندوق Box Plot كما فى شكل 18-11 للمتوسطين نجد أن هناك اختلاف فى متوسط الطريقتين، وأن متوسط الزمن بعد التعديل أقل من متوسط الزمن قبل التعديل، كذلك نجد أن التباين قد قل.



شكل رقم 18- 11 مخطط الصندوق للمثال 18-4

الحل بالمينيتاب Minitab

ندخل البيانات المتاحة إلى المينيتاب Minitab، ومن القائمة < Stat>Basic Statistics Paired T Test

ثم نذهب إلى نافذة النتائج فتظهر القيم التالية، ومنها نجد أن القيمة بى P=0.194 وهى قيمه أكبر من 0.05، وبالتالي لا نستطيع رفض الفرضية الصفرية، وأنه لا يوجد دليل على أن

هناك تباين فى المتوسط قبل وبعد التعديل، ويكون التعليق "There is no evidences to say there is statistically significant difference between the two groups"

```
Paired T for After - Before

              N      Mean     StDev   SE Mean
After        15   40.6004    4.1168    1.0630
Before       15   43.5964    7.7543    2.0022
Difference   15  -2.99592   8.51131   2.19761


95% CI for mean difference: (-7.70933; 1.71749)
T-Test of mean difference = 0 (vs not = 0): T-Value = -1.36  P-Value = 0.194
```

كذلك نجد أن فترة الثقة Confidence Interval تحتوى على قيمة سالبة وقيمة موجبة، وهذا دليل على أن قيمة المعامل "بى" P Value ستكون أكبر من 0.05، ودليل على أنه لا يوجد اختلاف بين المتوسطات.

الحل بالإكسيل Excel

ندخل البيانات المتاحة إلى Excel، ومن القائمة Tools>Data Analysis> Paired Two Samples For Mean

ثم نذهب إلى نافذة النتائج فتظهر القيم التالية، ومنها نجد أن القيمة بى P=0.097 وهى قيمه أكبر من 0.05، وبالتالي لا نستطيع رفض الفرضية الصفرية، وأنه لا يوجد دليل على أن هناك تباين فى المتوسط قبل وبعد التعديل، ويكون التعليق "There is no evidences to reject the null hypothesis (to say there is statistically significant difference between the two groups)"

جدول رقم 18- 9 نتيجة حل المثال 17-4 بالأكسيل

t-Test: Paired Two Sample for Means		
	After	*Before*
Mean	40.60042	43.59635333
Variance	16.94820063	60.12976812
Observations	15	15
Pearson Correlation	0.072602303	
Hypothesized Mean Difference	0	
df	14	
t Stat	-1.363266759	
P(T<=t) one-tail	0.097160813	
t Critical one-tail	1.761310115	
P(T<=t) two-tail	0.194321626	
t Critical two-tail	2.144786681	

تعليق على المثال الرابع

يلاحظ أن طريقة المعالجة اليدوية أوضحت أن هناك تباين فى المتوسط بينما طريقة الميناب Minitab، وExcel تشير إلى عدم وجود دليل إحصائى لحدوث تباين للمتوسط فما تفسير ذلك؟

يفسر ذلك بأن التباين الذي حدث – برغم أنه تغير- إلا أنه قد يكون فى حدود التباين الطبيعي للعملية، أو أن حجم العينة كان أقل من أن يكشف أو يظهر هذا التغير، وهذا يبين أهمية أن يكون حجم العينة كبيراً، وأهمية أن يتم الاختبار بأكثر من طريقة كما أشرنا إلى ذلك فيما قبل.

أيضا يجب الانتباه إلى أنه بالرغم من أن الاختبار قد يشير إلى عدم وجود اختلاف إحصائي – كما فى حالتنا- إلا أننا متيقنين من أن هناك تغير، صحيح أن هذا التباين قد يكون غير ملموس إحصائيا، إلا أنه من وجهة نظر أصحاب العمليات تباين مهم ويجب أخذه فى الاعتبار، فالتعديل الذي تم من وجهة النظر الإحصائية لم يؤد إلى تباين ملموس، إلا أننا نلحظ أن هذا التعديل لم يؤدى فقط إلى تقليل زمن الصيانة، بل أدى كذلك إلى تقليل التباين الموجود فى عملية الصيانة ويتضح ذلك من مخطط الصندوق والشنب، ومن قيمة الانحراف المعيارى بعد التعديل، وهنا سيكون قرار الإدارة هو تبنى التعديل وتفعيلة على هذه المعدة، وإجراء مزيد من التجارب قبل تعميمه على المعدات المشابهه.