جامعة محمد بوضياف بالمسيلة

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

الاحصاء 03: دروس و امثلة محلولة باستخدام برنامج MINITAB

موجهة لطلبة الليسانس كل التخصصات

اعداد الاستاذ مصطفى جاب الله

السنة الجامعية: 2020-2021

مقدمة

الإحصاء هو مجموعة الأدوات والبحوث الرياضية المستخدمة لتحديد خصائص مجموعة من البيانات (كبيرة بشكل عام)..

الإحصاء هو نتاج التحليلات القائمة على استخدام الإحصائيات. يجمع هذا النشاط ثلاثة فروع رئيسية:

• جمع البيانات.

• معالجة البيانات المجمعة ، وتسمى أيضًا الإحصاءات الوصفية.

• تفسير البيانات ، ويسمى أيضًا الاستدلال الإحصائي ، والذي يعتمد على نظرية المسح والإحصاء الرياضي.

هذا التمييز لا يتمثل في تحديد عدة مجالات مانعة لتسرب الماء. في الواقع ، لا يمكن معالجة البيانات وتفسيرها إلا بعد جمعها. على العكس من ذلك ، تحدد الإحصائيات الرياضية قواعد وطرق جمع البيانات ، بحيث يمكن تفسيرها بشكل صحيح. و الغرض من الإحصاء هو استخراج المعلومات ذات الصلة من قائمة الأرقام التي يصعب تفسيرها بقراءة بسيطة. يتم استخدام مجموعتين رئيسيتين من الأساليب حسب الظروف. لا شيء يمنع استخدامها بالتوازي في مشكلة ملموسة ، لكن يجب ألا ننسى أنها تحل مشاكل ذات طبيعة مختلفة تمامًا. وفقًا للمصطلحات الكلاسيكية ، هذه هي الإحصاء الوصفي والإحصاء الرياضي. اليوم ، يبدو أن التعبيرات مثل تحليل البيانات والإحصاءات الاستنتاجية مفضلة ، وهو ما يبرره تقدم الأساليب المستخدمة في الحالة الأولى.

على سبيل المثال الدرجات الإجمالية في الامتحان. قد يكون من المثير للاهتمام استخلاص قيمة مركزية منها تعطي فكرة تركيبية عن مستوى الطلاب. يمكن استكمال ذلك بقيمة تشتت تقيس ، بطريقة معينة ، تجانس المجموعة. إذا أردنا معلومات أكثر دقة حول هذه النقطة الأخيرة ، فيمكننا إنشاء مدرج تكراري أو ، من وجهة نظر مختلفة قليلاً ، النظر في الفئات العشرية. قد تكون هذه المفاهيم مثيرة للاهتمام لإجراء مقارنات مع اختبارات مماثلة أجريت في السنوات السابقة أو في أماكن أخرى. هذه هي

المشاكل الأساسية لتحليل البيانات التي تتعلق بمجموعة سكانية محدودة. تتطلب مشاكل الإحصاء متعدد الأبعاد استخدام الجبر الخطي. بغض النظر عن طبيعة المشكلة ، أولية أم لا ، إنها مسألة التخفيضات الإحصائية للبيانات المعروفة التي لا يؤدي فيها إدخال الاحتمالات إلى تحسين المعلومات التي تم الحصول عليها. من المعقول تجميع هذه المفاهيم المختلفة: الإحصاء الوصفي للمفاهيم الأولية.

يحدث التغيير الجذري عندما لم تعد البيانات تعتبر معلومات كاملة يتم فك تشفيرها وفقًا لقواعد الجبر ولكن كمعلومات جزئية عن مجموعة أكبر من السكان ، تُعتبر عمومًا مجموعة لا حصر لها. للحث على معلومات عن السكان المجهولين ، من الضروري إدخال مفهوم قانون الاحتمالات. تشكل البيانات المعروفة في هذه الحالة تحقيقًا لعينة ، وهي مجموعة من المتغيرات العشوائية التي يفترض أنها مستقلة (انظر قانون الاحتمالات مع عدة متغيرات). تسمح نظرية الاحتمالات بعد ذلك ، من بين عمليات أخرى:

• لربط خصائص العينة بتلك التي تنسب إلى قانون الاحتمال ، غير معروف بكل صرامة ، فهو أخذ العينات ؛

• استنتاج معاكسات قانون الاحتمالات من المعلومات التي قدمتها العينة ، وهذا هو التقدير ؛

• تحديد فترة الثقة التي تقيس صحة التقدير؛

• لإجراء اختبارات الفرضيات ، وأكثرها استخدامًا هو اختبار كاي تربيع لقياس مدى ملاءمة قانون الاحتمالات المختار للعينة المستخدمة ؛

هذه المطبوعة تحوي ثلاثة فصول رئيسية وهي

- المعاينة وتدوير العينات
- التقدير الاحصائي
- اختبار الفرضيات

الفصل الاول : المعاينة وتدوير العينات sampling and bootstraping

الباحث عادة ما يجد نفسه أمام مجموعة كبيرة من البيانات الخام التي لا تفصح عن شئ على حين أنة مطالب باستخلاص حقائق علمية واضحة ومحددة من تلك البيانات سواء كانت بيانات مسوح اجتماعية شاملة . أو بالعينة أو بيانات تعدادات سكانية عندئذ يستطيع الباحث من خلال الإحصاء أن يغير من شكل البيانات بعد تصنيفها وتنظيمها وتلخيصها مستخدما فى ذلك الجانب الوصفى من الإحصاء حيث يمكنه أن يطبق هنا مجموعة من المقاييس الإحصائية التى لا تتعدى حد الوصف مثل مقاييس النزعة المركزية ومقاييس التشتت ومقاييس الارتباط والانحدار ... الخ ومن ثم يتبين لدينا أن الوظيفة الإحصائية الأولى للإحصاء هى توصيف البيانات المتاحة والخروج منها بمجموعة من المؤشرات والمعدلات الإحصائية الثانية : تتلخص في الاستدلال ، ففى مجال البحوث الاجتماعية ، عادة ما تستخدم العينة sample لتمثل المجتمع الذى سحبت منه ويرجع استخدام العينات في البحوث الاجتماعية إلى عدة أسباب لعل أهمها توفير الوقت ، والجهد ، والإمكانيات التى تجعل من المتعذر أحيانا وربما من المستحيل أحيانا أخرى دراسة المجتمع ككل . والعينة ببساطة هى جزء أو قطاع من المجتمع تم اختيارها على أساس إحصائي لكى تمثل المجتمع الذى هى جزء منة وهنا يكون دور الإحصاء هو الوصول إلى تقديرات واستدلالات عن المجتمع ككل من خلال المعلومات المتوفرة عن العينة التى تم سحبها من هذا المجتمع ، إذ إن جُل اهتمام الباحث ليس مجرد العينة المستخدمة في الدراسة بل المجتمع ككل ، باختصار فان الجانب الاستدلإلى من الإحصاء يهتم بتقدير معالم المجتمع Population Parameters فيما يتعلق بالظاهرة موضوع الدراسة مستخدما البيانات والمعلومات المتوفرة لدية عن العينة أو ما يسمى بـ Sample Statistics حول نفس الظاهرة في محاولة الوصول إلى تصميمات Generalizations عن مجتمع الدراسة.

الفصل التمهيدي مراجعة اساسيات الاحصاء الوصفي

1

- العرض البياني :

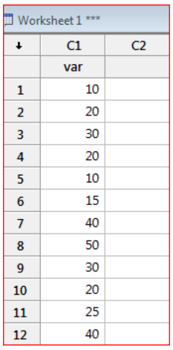
ونتطرق في هذا الجزء الى بعض التمثيلات البيانية التي تستخدم في وصف الظواهر الاحصائية الوصفية خاصة ،ومنها

1 - 1 المدرج التكراري Histogram

يتم تمثيل محتوي هذا الشكل بالاعتماد على الامر

Graph > Histogram...

مثال 1: بعد ادخال البيانات في ورقة العمل تحصلنا على مايلي

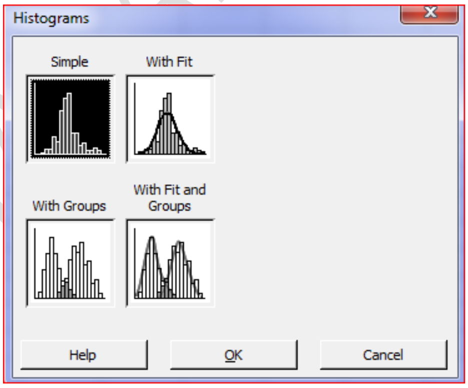

Worksheet 1 ***

↓	C1	C2
	var	
1	10	
2	20	
3	30	
4	20	
5	10	
6	15	
7	40	
8	50	
9	30	
10	20	
11	25	
12	40	

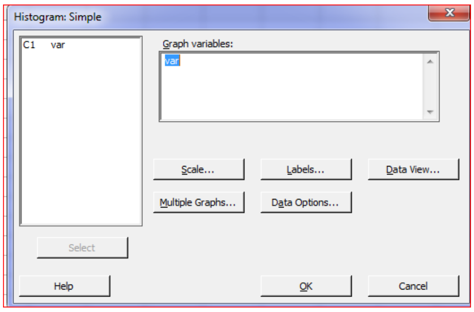
ثم انشاء التعليمة او الامر **...Histogram**

Graph>

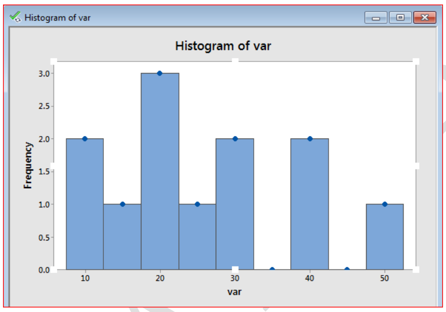
فيظهر المربع الحواري التالي



نختار احدى الحالات الاربع لتحديد شكل المدرج التكراري وذلك بالضغط OK ، و تتحصل عندئذ على المربع التالي

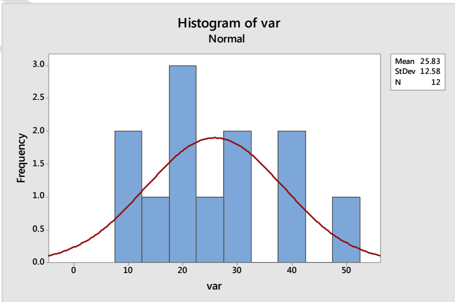


بعد ادخال var او c1 في حيز Graph variables ، ثم الضغط على OK



1 - 2 المضلع التكراري :

ويتم الحصول عليها مثل المدرج التكراري ولكن في المربع الحواري نختار ايقونة with fit فنضغط عليها



1 -3 الاعمدة البيانية :

لإنشاء اعمدة بيانية تستخدم خاصة في الاحصاء الوصفي للبيانات

مثال 2

من خلال الملاحق الاحصائية لتقارير بنك الجزائر توفرت لدينا بيانات حول الاستهلاك بنوعيه العمومي والخاص حسب الجدول

لسنتي 2015 2016

السنة	الاستهلاك الخاص	الاستهلاك العمومي
2015	6854	3613.4
2016	7446	3617.7

3617.7 و 7446 مليار دينار نريد ان نظهر هذه البيانات في شكل اعمدة بيانية باستخدام Minitab 17

الاستهلاك الخاص PC والاستهلاك الحكومي GC بحيث تكون الاعمدة الموالية لكل من C1 C2 هي المتغيرات في السطر الموالي مباشرة والسطر الثالث تسجل قيم المتغيرات ، من خلال الصفحة الموالية

وبعدها نستخدم الامر التالي

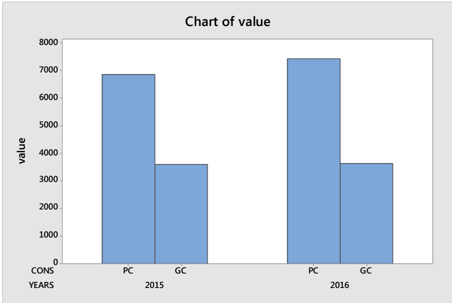
Graph>Bar Chart

المطلوب عرض البيانات بطريقة الاعمدة المتلاصقة

Worksheet 3 ***

↓	C1	C2-T	C3
	value	CONS	YEARS
1	6854.0	PC	2015
2	7446.0	PC	2016
3	3613.4	GC	2015
4	3617.7	GC	2016

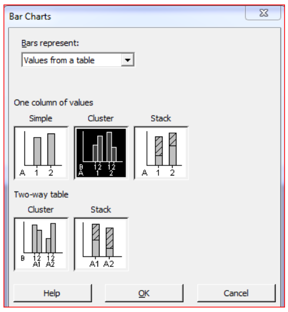
ادخال البيانات

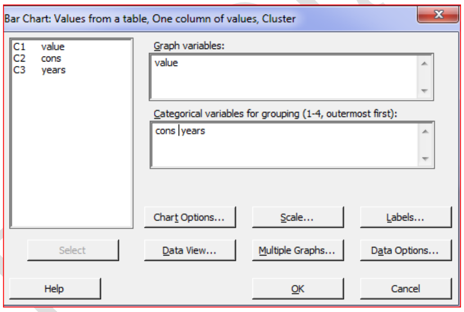


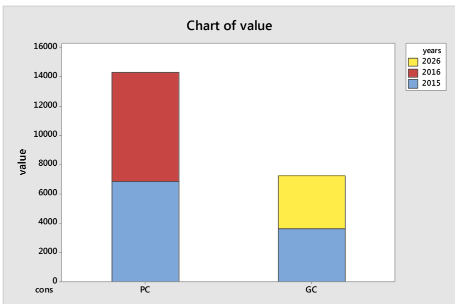
1 - 4 المستطيلات البيانية :

لناخذ المثال السابق ونقوم بالخطوات التالية

اختيار التعليمة **Graph>Bar Chart** ثم الحصول على المربع الحواري







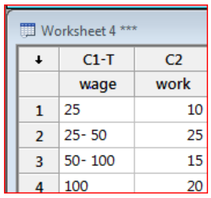
5-1 الدائرة النسبية : Pie Chart

مثال 3

باخذ المثال التالي في الجدول حول اجور العمال ب 10^2 دينار

الفئة الاجرية wage	اقل من 25	25- 50	50 -100	اكثر من 100
عدد العمال work	10	25	15	20

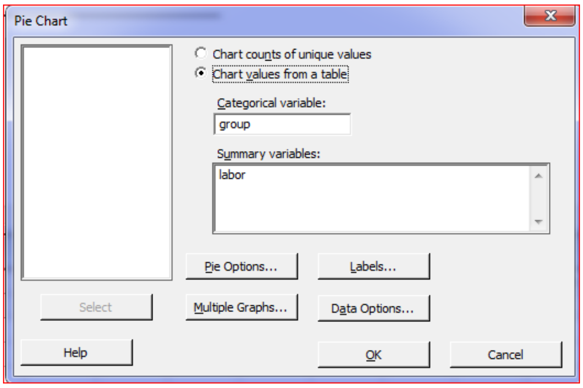
نقوم بنقل هذا الجدول الى ورقة عمل في برنامج Minitab



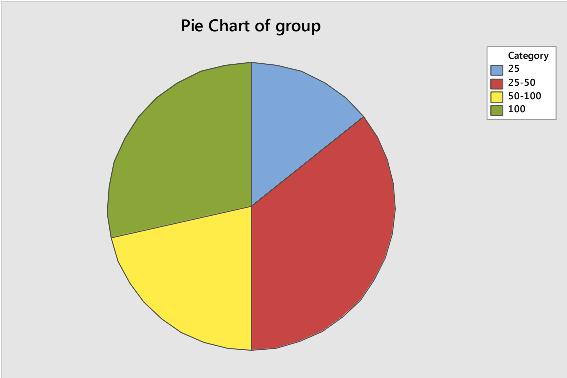
Worksheet 4 ***

↓	C1-T	C2
	wage	work
1	25	10
2	25- 50	25
3	50- 100	15
4	100	20

ثم ننشأ التعليمة Graph > Pie Chart ..... ونكون المربع الحواري التالي



وبعد الضغط على ok



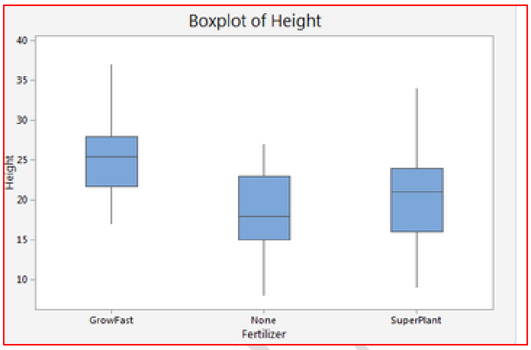
6-1 التمثيل البياني للمربع boxplot

يريد مصنع الأسمدة النباتية تطوير صيغة من الأسمدة التي تحقق أكبر زيادة في ارتفاع النباتات. لاختبار الصيغ الأسمدة ، يتم تصنيف ثلاث مجموعات من 50 شتلة متطابقة نسمي متغير السماد fertilizer ويتم تقسيمه الى ثلاث مجموعات مجموعة مراقبة مع عدم وجود الأسمدة نسميها Nonfertilizer ، ومجموعة مع الأسمدة المصنعة التي تنمو بسرعة تسمى ، GrowFast ، ومجموعة مع الأسمدة الممتازة المسماة SuperPlant من الشركة المصنعة المنافسة. بعد أن تكون النباتات في بيئة دافئة لمدة ثلاثة أشهر ، يتم قياس ارتفاع النباتات Height. ،كجزء من التحقيق الأولي ، يقوم العالم بإنشاء قطعة مربعة لارتفاعات النبات من المجموعات الثلاث لتقييم الاختلافات في نمو النبات بين النباتات التي لا تحتوي على سماد ، والنباتات التي تحتوي على سماد الشركة المصنعة ، والنباتات ذات السماد المنافس لها

الحل

نختار الامر الاتي Graphs > Boxplot > Single Y Variable: With Groups

في مربع Y variable نضع المتغير Height اما في مربع Group variable ندخل fertilizer ونضغط على ok



Summary Statistics

Fertilizer	N	Minimum	Q1	Median	Q3	Maximum	95% Median CI
GrowFast	50	17.0000	21.7500	25.5000	28.0000	37.0000	(23.0000, 27.0000)
None	50	8.0000	15.0000	18.0000	23.0000	27.0000	(17.0000, 20.0000)
SuperPlant	49	9.0000	16.0000	21.0000	24.0000	34.0000	(19.0000, 22.7856)

يوضح الرسم البياني أن GrowFast تسبب زيادة أكبر وأكثر ثباتًا في ارتفاع النبات.في حين ان النباتات التي تنمو بسرعة GrowFast تنتج أطول النباتات عموما. كما يزيد نوع الاسمدة الممتازة SuperPlant أيضًا من ارتفاع النبات ، لكن تباينه أكبر ، ولا يكون SuperPlant له تأثير إيجابي على نسبة كبيرة من الشتلات.

7-1 بيان السلسلة الزمنية Time series plot

نستخدم Time Series Plot للبحث عن أنماط في البيانات بمرور الزمن ، مثل الاتجاه العام أو المركبات الموسمية. يمكن أن تساعد بيانات السلاسل الزمنية على اختيار تحليل السلاسل الزمنية لنمذجة البيانات

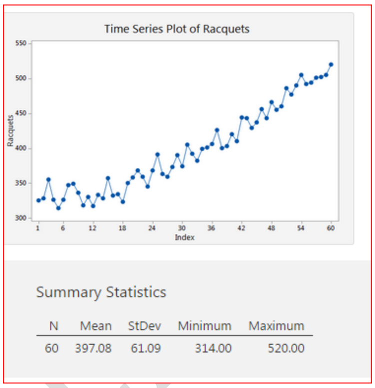
مثال

يريد محلل تسويق تقييم اتجاهات مبيعات مضارب التنس. يجمع المحلل بيانات المبيعات من السنوات الخمس الماضية للتنبؤ بمبيعات المنتج للأشهر الثلاثة القادمة. كجزء من التحقيق الأولي ، يقوم المحلل بإنشاء مخطط سلسلة زمنية لمعرفة كيف تغيرت المبيعات مع مرور الوقت.

الحل

نكتب الامر التالي

Graphs > Time Series Plot > Single Y Variable: Simple



تظهر سلسلة السلاسل الزمنية اتجاه عام صعودي واضح. قد يكون هناك أيضًا منحنى بسيط في البيانات ؛ يبدو أن الزيادة في قيم البيانات تتسارع بمرور الوقت

2 -الجداول التكرارية :

هذا النوع من الجداول يستخدم الجداول العادية ويحولها الى جداول تكرارية تستخدم في معالجة البيانات في ميدان الاحصاء الوصفي ، ويمكن تجزئة هذا النوع من الجداول الى جزأين

1-2 جداول تكرارية لمتغير واحد:

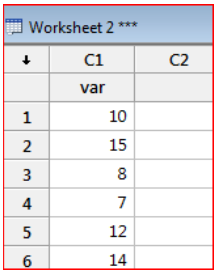
متال 4

بأخذ الجدول التالي

8	15	10
14	12	7

المطلوب اعداد جدول تكراري للبيانات التالية

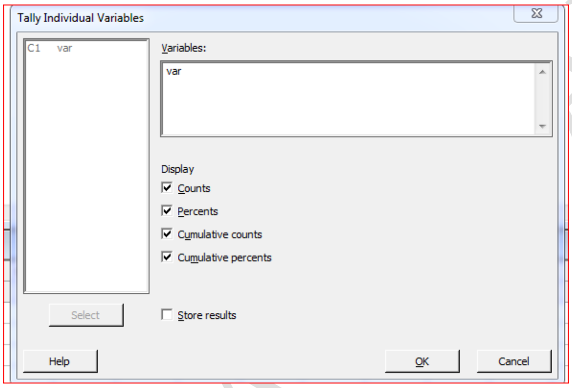
الحل : ندخل البيانات في ورقة العمل كما يلي



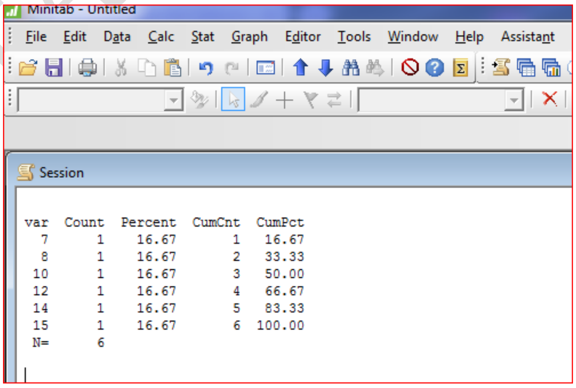
↓	C1	C2
	var	
1	10	
2	15	
3	8	
4	7	
5	12	
6	14	

ثم نكون الامر الاتي **Stat > Tables > Tally Individual Variables**

و نستخدم المربع الحواري التالي الذي يوضح التكرارات و التكرارات النسبية وكذا التكرارات المجمعة الصاعدة



وبالضغط على ok تتحصل على النتيجة



var	Count	Percent	CumCnt	CumPct
7	1	16.67	1	16.67
8	1	16.67	2	33.33
10	1	16.67	3	50.00
12	1	16.67	4	66.67
14	1	16.67	5	83.33
15	1	16.67	6	100.00
N=	6			

2-2 جداول تكرارية لأكثر من متغير:

في حالة دراسة اكثر من متغير نتبع التعليمة التالية

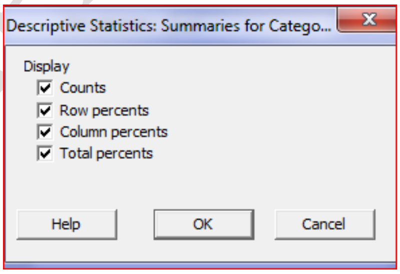
باتباع الامر التالي Stat > Tables > Descriptive Statistics

مثال 5 :بافتراض وجود نشاطين في السوق الصناعة ind والزراعة agr وكل نشاط له بعض المنتجات المحلية dom

واخرى مستورد imp و بعض منتجات النشاطين خاضعة للضريبة yes و الاخرى معفاة من الضرائب No بتحويل كل هذه المعطيات الى ورقة العمل نجد

↓	C1-T	C2-T	C3-T
	trad	act	tax
1	imp	ind	no
2	imp	agr	no
3	dom	ind	yes
4	imp	agr	yes
5	dom	ind	no
6	dom	ind	no
7	imp	agr	no
8	imp	agr	no
9	exp	ser	no
10	imp	agr	no

وباستخراج المربع الحواري التالي ثم الضغط على ok نجد



النتائج كما تبينها مخرجات برنامج Minitab

```
Results for tax = no
Rows: trad   Columns: act
agr  ind  All

dom      0    0    0
         0    0    0
         0    0    0
imp      1    1    2
        50   50  100
       100  100  100
        50   50  100
All      1    1    2
        50   50  100
       100  100  100
        50   50  100
Cell Contents:     Count
                   % of Row
                   % of Column
                   % of Total
Results for tax = yes
Rows: trad   Columns: act
agr  ind  All
dom      0    1    1
         0  100  100
         0  100   50
         0   50   50
imp      1    0    1
       100    0  100
       100    0   50
        50    0   50
All      1    1    2
        50   50  100
       100  100  100
        50   50  100
```

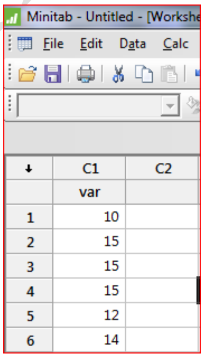
3-مقاييس النزعة المركزية   Central Tendency Measures :

نعني بمقاييس النزعة المركزية الوسط الحسابي ، الوسيط ، المنوال ، وباقي الاوساط التي تشبه الوسط الحسابي

3 - 1 مقاييس النزعة المركزية في حالة المعطيات غير المبوبة Ungrouped data :

مثال 6

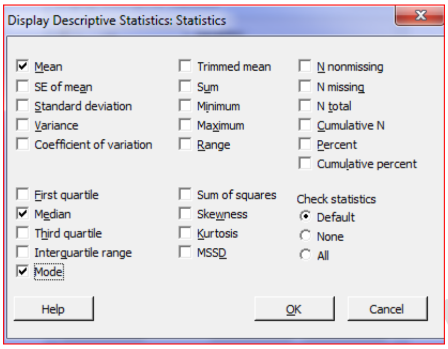
بتعديل المثال 4 وتحويله الى ورقة عمل MINITAB نجد



ثم نكتب التعليمة التالية

Stat> Basic Statistics > Display Descriptive Staistics ......

ونضع المتغير var في المستطيل variables ونضغط على statistics وينشا لدينا المربع الحواري التالي الذي نختار منه على سبيل المثال الاوساط الثلاثة في الاحصاء الوصفي وهي المتوسط الحسابي الوسيط والمنوال (Mean ; Median ; Mode) ثم نضغط على ok



بالضغط على ok ثانية نحصل على صفحة خاصة بالنتائج

```
11/23/2019 10:06:17 AM

Descriptive Statistics: var

                                         N for
Variable    Mean  Median  Mode   Mode
var       13.500  14.500    15      3
```

نلاحظ قيم كل من المتوسط الوسيط والمنوال الذي تكررت ثلاث مرات

3 - 2- مقاييس النزعة المركزية في حالة البيانات المبوبة : Grouped data

مثال 7

بتعديل المثال 3 بصيغة البيانات المبوبة من الاعلى والأسفل وحساب الفئة المتوسطة للأجور xi ثم نقل المعطيات الى ورقة العمل نجد

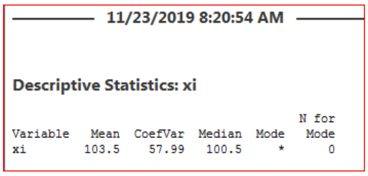
Worksheet 4 ***

↓	C1-T	C2	C3
	wage	work	xi
1	25 - 50	10	37.5
2	50- 100	25	75.5
3	100- 150	15	125.5
4	150- 200	20	175.5

و اختيار التعليمة

Stat> Basic Statistics > Display Descriptive Staistics

ونضع المتغير xi في المستطيل variables ونضغط على statistics وينشا لدينا المربع الحواري التالي الذي نختار منه على سبيل المثال الاوساط الثلاثة في الاحصاء الوصفي وهي المتوسط الحسابي الوسيط والمنوال (Mean ; Median ; Mode) ثم نضغط على ok



11/23/2019 8:20:54 AM

Descriptive Statistics: xi

Variable	Mean	CoefVar	Median	Mode	N for Mode
xi	103.5	57.99	100.5	*	0

نلاحظ ان قيمة المنوال غير موجودة في هذا التوزيع نظرا لعدم تكرر أي اجر

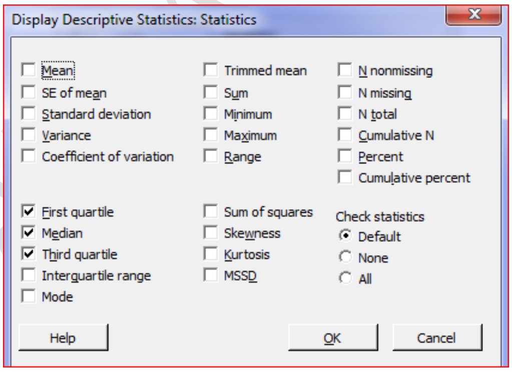
3-3 الربيعيات - (Quartiles)

كثيرا ما تتم تجزئة التوزيع الى اجزاء من اجل دراسة محتويات المتغيرات بعمق اكثر ، واذا كان الوسيط يقسم التوزيع الى قسمين فبامكاننا ان نقسمه الى اربعة اقسام وتسمى الربيعيات او الى عشرة وتسمى العشريات (deciles) او الى مئة قسم وتسمى المئينيات(percentiles)

وتعتبر هذه الاجزاء احدى مقاييس النزعة المركزية تتم بها اختبار مدى اعتدالية التوزيع

باستخدام المثال 7 دون تعديل وباتباع التعليمة **Stat> Basic Statistics > Display Descriptive Statistics**

وباختيار نافذة statistics من المربع الحواري مع الاشارة الى الربيعيات الثلاثة كما هو موضح



نلاحظ اننا اشرنا على كل من الربيعي الاول ، والثاني و الثالث وبالضغط على ok نجد

———— 11/23/20195

Descriptive Statistics: xi

```
Variable     Q1  Median     Q3
xi         47.0   100.5  163.0
```

ملاحظة : يمكن حساب الربيعيات في الحالتين سواء كانت المعطيات مبوبة او غير مبوبة حالها حال الوسيط

4 مقاييس التشتت Measures of dispersion

4-1 حالة البيانات غير البيانات المبوبة

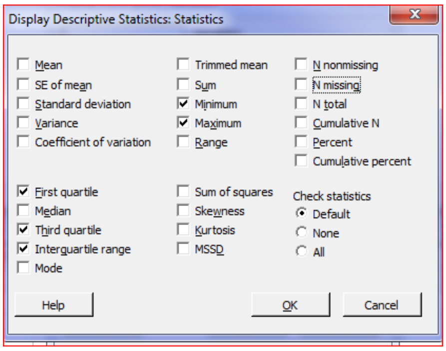
يمكن حساب المدى ، والمدى الربيعي التباين ، والانحراف المعياري بمعطيات فردية ، ولنحسب المدى والمدى الربيعي اولا ، الى جانب القيمة العظمى max ، والصغرى min في التوزيع بإعادة المثال 4 ، و استخراج ورقة عمل برنامج MINITAB ثم نستخدم الامر او التعليمة

Stat> Basic Statistics > Display Descriptive Statistics

Worksheet 2 ***

↓	C1	C2
	var	
1	10	
2	15	
3	8	
4	7	
5	12	
6	14	

نضع المتغير المستطيل variables ونضغط على statistics وينشا لدينا المربع الحواري التالي



الذي نختار منه على سبيل المثال المدى range نصف المدى الربيعي و الذي يمثل معامل الاختلاف للبيانات غير المبوبة (IQR) interquartile rang و يحسب بالصيغة الاتية

$$IQR = \frac{Q3 - Q1}{Q2} X100$$

علما ان الربيعي الثاني Q2 هو الوسيط Me

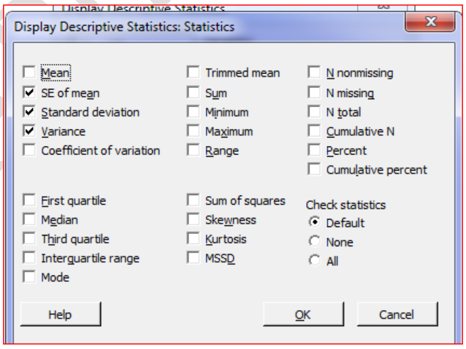
الى جانب القيمة القصوى Max القيمة الدنيا MIN للتوزيع كما يمكن حساب الربيعين الاول Q1 والثالث Q3 على حدى ويكمن الحصول على صفحة النتيجة التالية

11/23/2019 1:58:28 PM

Descriptive Statistics: var

Variable	Minimum	Q1	Q3	Maximum	Range	IQR
var	7.00	7.75	14.25	15.00	8.00	6.50

كما يمكن حساب مقاييس تشتت اخرى وبالمعطيات غير المبوبة دائما وهي التباين Variance والانحراف المعياري St dev و انحراف خطأ المتوسط SE Mean ، وذلك بكتابة المتغير var في المستطلع variables ونضغط على statistics وينشا لدينا المربع الحواري التالي الذي نختار منه المعايير سالفة الذكر



وبعد الضغط على ok نتحصل على صفحة النتائج

Descriptive Statistics: var

Variable	SE Mean	StDev	Variance
var	1.32	3.22	10.40

4-2 البيانات المبوبة :

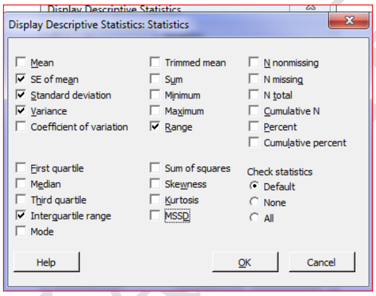
بالنفس الطريقة يمكن حساب المقاييس سالفة الذكر ، لكن في حالة القيم المبوبة هذه المرة ، وباستعراض المثال 7 في ورقة العمل

Worksheet 4 ***

↓	C1-T	C2	C3
	wage	work	xi
1	25 - 50	10	37.5
2	50- 100	25	75.5
3	100- 150	15	125.5
4	150- 200	20	175.5

يمكن عندئذ حساب بعض مقاييس التشتت ولنأخذ كل المقاييس التي تم التطرق اليها باستخدام التعليمة

Stat> Basic Statistics > Display Descriptive Statistics اولا ، ثم المربع الحواري الذي نختار مقاييس التشتت المراد تقديرها



ونضغط بعدئذ على فنتحصل على صفحة النتائج التالية

11/23/2019 1:58:28 PM

Descriptive Statistics: xi

```
Variable  SE Mean  StDev  Variance  Range    IQR
xi           30.0   60.0    3602.7  138.0  116.0
```

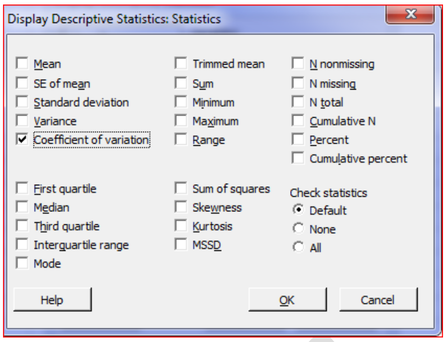
4-3 -معامل الاختلاف : Coefficient of Variance

هو احد مقاييس التشتت النسبية ، وهو نسبة مئوية بين الانحراف المعياري ، والمتوسط الحسابي ويحسب في البيانات المبوبة

$$CV = \frac{\delta}{\overline{X}} \times 100$$

نقيس بواسطته مدى تشتت التوزيع او تماسكه ، وكلما ابتعدت قيمته عن الواحد او مائة بالمائة كان التوزيع تشتتا ، والعكس صحيح

ويمكن حسابه باستخدام MINITAB بنفس الاسلوب مع مقاييس التشتت المطلق السابقة ، وفي المربع الحواري تتم الاشارة له



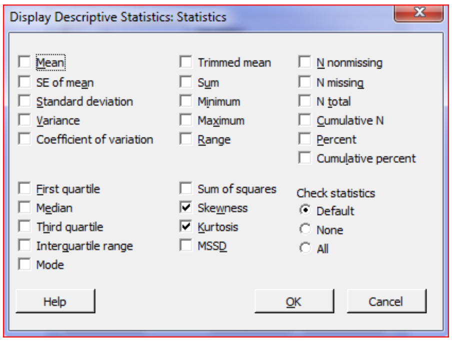
وبعدها بالضغط على ok

```
———————— 11/23/2019 1:58:20

Descriptive Statistics: xi

Variable  CoefVar
xi          57.99
```

4 - 4 مقاييس الشكل Skewness and Kurtosis :

يمكن حساب وايجاد كل من الالتواء skewness والتفرطح kurtosis كمايلي

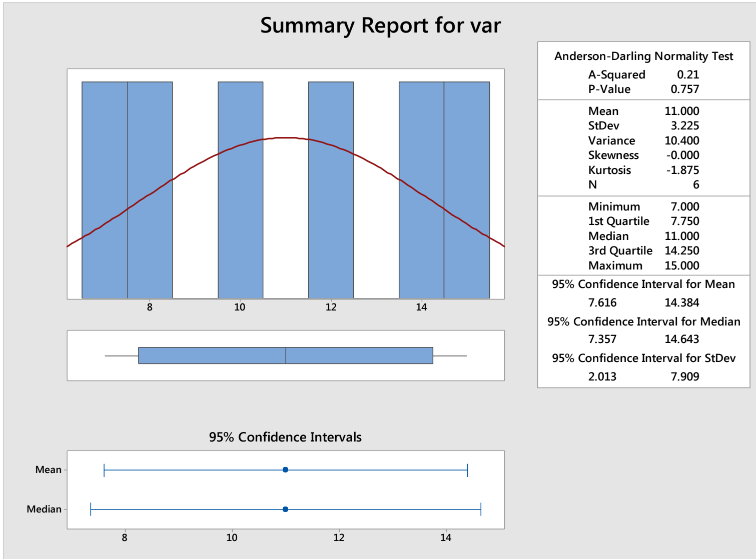
اولا حسابيا باستخدام نفس الخطوات في المقاييس السابقة المشار اليها ، وعند الوصول الى المربع الحواري يتم التأشير على هذين المقياسين كما هو موضح



ثم الضغط على ok كي نحصل على النتيجة

```
Descriptive Statistics: xi

Variable  Skewness  Kurtosis
xi            0.23     -1.51
```

ثانيا بيانيا ، وذلك بتغيير الامر او التعليمة لتصبح **STAT >Basic Statistics >Graphical SUMMRAY**



1-حجم العينة اللازم للتقدير Sample size for estimation

ان استخدام حجم العينة لتقدير عدد المشاهدات التي نحتاجها لتحقيق هامش خطأ معين لمجالات الثقة لمتوسط أو الانحراف المعياري أو التباين أو النسبة أو معدل بواسون على العكس ، يمكن تقدير هامش الخطأ استنادًا إلى حجم العينة الذي تخطط لاستخدامه. ويمكن تقدير حجم العينة بالصيغة التالية

$$n = \left[\frac{\sigma \times z_{score}}{MOE}\right]^2$$

حيث تمثل كل من σ الانحراف المعياري و z_{score} القيمة الحسابية للتوزيع الطبيعي اما MOE فهو هامش خطا التقدير

مثال

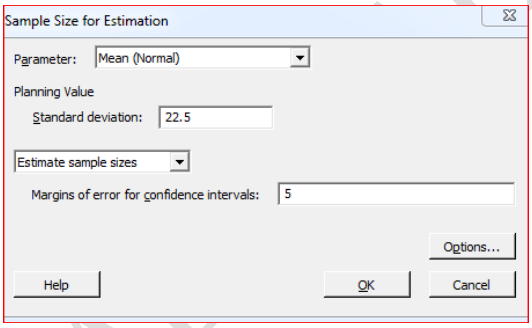
تستخدم شركة إلكترونيات حجم العينة لحساب التقدير قبل إجراء دراسة لتقدير متوسط الجهد لخط جديد من المقاومات المستخدمة في لوحات الدوائر. تريد الشركة أن تعرف حجم العينة المطلوب للحصول على هامش الخطأ 5 بناءً على الدراسات السابقة ، يبلغ الانحراف المعياري 22.5.

الحل

لحل هذا المثال باستخدام minitab نقوم باصدار الامر التالي

.Stat > Power and Sample Size > Sample Size for Estimation

فيظهر المربع الحواري التالي



نضع في مربع parameter : Mean (Normal) و Standard deviation =22.5 . Standard of error for confidence level=5

```
Sample Size for Estimation

Method

Parameter           Mean
Distribution        Normal
Standard deviation  22.5 (estimate)
Confidence level    95%
Confidence interval Two-sided

Results

  Margin  Sample
of Error    Size
       5      81
```

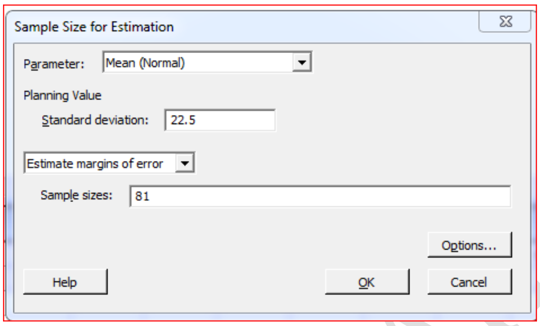
لتحقيق هامش خطأ قدره 5 عند تقدير متوسط الجهد للمقاومات ، يحتاج المحلل إلى جمع حجم عينة من 81.

ملاحظة لتقدير هامش خطأ التقدير (MOE) فاذا كانت العينة تتبع التوزيع الطبيعي فان هامش خطا التقدير يساوي

$$\pm MOE = Z_{score} \times \frac{\sigma}{\sqrt{n}}$$

يكفي ان نعود الى المربع الحواري السابق ونفعل ايقونة Estimate margin of error

وبالاعتماد على المثال نفسه نضع في ايقونة Sample sizes حجم العينة نفسه وليكن 81



وبالضغط على

```
Sample Size for Estimation

Method

Parameter            Mean
Distribution         Normal
Standard deviation   22.5 (estimate)
Confidence level     95%
Confidence interval  Two-sided

Results

  Margin  Sample
of Error    Size
       5      81
```

وهذا يعني ان هامش الخطأ اللازم لتقدير عينة بحجم 81 مشاهدة هو 5% اي 0,05

- 2- **حجم العينة لمجال معدل السماح**

يستخدم حجم العينة لمجال التسامح لفحص العلاقة بين حجم العينة والحد الأقصى للنسب المئوية المقبولة من المجتمع في المجال الزمني للتسامح. وبرنامج Minitab يعرض فترات لطريقتين. تطبق الطريقة العادية فقط عندما تتبع البيانات التوزيع الطبيعي. تنطبق الطريقة اللامعلمية على أي توزيع مستمر ، ولكنها أكثر تحفظًا من الطريقة العادية.

مثال

تريد مصنع للغسالات الكهربائية تحديد حجم عينة من الغسالات اللازمة لقياس لتحقيق أقصى النسب المئوية المقبولة من مجتمع انتاجها في الفترة من 96 ٪ و 97 ٪ لفاصل التسامح. يريد ادارة المصنع أيضًا معرفة الحد الأقصى للنسب المئوية المقبولة لأحجام العينات من 50 أو 100 غسالات. يمكن لها افتراض أن البيانات موزعة بشكل طبيعي.

الحل

2-1 لتحديد حجم العينة

يمكن ايجاد الحل اعتمادا على الامر الاتي

Stat > Power and Sample Size > Sample Size for Tolerance Intervals

نختار Calculate sample sizes

في المربع Minimum percentage of population in interval نختار 95

في المربع Maximum acceptable percentages of population in interval (*P) نختار 96 97 ثم ok

Sample Size for Tolerance Intervals

Method

Confidence level	95%
Minimum percentage of population in interval	95%
Probability the population coverage exceeds p*	0.05

Sample size for 95% Tolerance Interval

P*	Normal Method	Nonparametric Method	Achieved Confidence	Achieved Error Probability
96.000%	2480	4654	95.0%	0.049
97.000%	525	1036	95.1%	0.048

P* = Maximum acceptable percentage of population in interval
Achieved confidence and achieved error probability apply only to nonparametric method.

عندما تحدد ادارة المصنع أحجام العينة المستهدفة ، يقوم Minitab بحساب النسب المئوية القصوى المقبولة من السكان في الفترة الفاصلة. مع احتمال تجاوز التغطية السكانية p * تساوي 0.05 (5٪) ، تكون النسبة المئوية القصوى المقبولة للطريقة العادية هي 99.4015٪ عندما يكون حجم العينة 50. ، اما عندما يكون حجم العينة 100 ، تكون النسبة المئوية القصوى المقبولة هي 98.6914٪.

ملاحظة: إذا لم تستطع ادارة المصنع افتراض حالة التوزيع الطبيعي، فستكون النسب المئوية القصوى المقبولة من المجتمع أعلى باستخدام الطريقة غير المعلمية Nonparametric method

- 2-2 لتحديد حجم الخطأ المسموح

نختار الامر التالي

Stat > Power and Sample Size > Sample Size for Tolerance Intervals

نختار Calculate maximum acceptable percentages of population in interval (*p)

Minimum percentage of population in interval نختار 95

Sample sizes نضع 50 100 ثم ok

```
Sample Size for Tolerance Intervals

Method

Confidence level                                   95%
Minimum percentage of population in interval       95%
Probability the population coverage exceeds p*    0.05

Maximum Acceptable Percentages of Population for 95% Tolerance Interval

Sample    Normal  Nonparametric    Achieved  Achieved Error
  Size    Method         Method  Confidence    Probability
    50  99.4015%       99.2846%       72.1%          0.050
   100  98.6914%       99.6435%       96.3%          0.050

Achieved confidence and achieved error probability apply only to nonparametric
     method.
```

قد تقرر ادارة المصنع أن النسبة المئوية القصوى المقبولة مرتفعة للغاية وقد يعيد تشغيل التحليل باستخدام أحجام أكبر للعينات لتقليل النسبة المئوية القصوى المقبولة. على سبيل المثال ، يمكن للادارة تجربة 250 أو 400 غسالة. ومع ذلك ، فهي تعلم من التحليل الأول أن هناك حاجة إلى 525 غسالة على الأقل للحصول على احتمال بنسبة 5٪ ألا يحتوي مجال التسامح على أكثر من 97٪ من السكان ، مع افتراض التوزيع الطبيعي.

3 - اختبار فرضيات حجم العينة

1-3 في حالة عينة واحدة

1-3 -1 اختبار فرضية حجم عينة طبيعية 1 Sample Z test :

يتم استخدام قوة وحجم العينة تتوزع طبيعيا Sample Z Power and Sample Size لفحص العلاقة بين قوة ، وحجم العينة ، والفرق عندما تريد مقارنة متوسط مجتمع ما بالهدف أو القيمة المرجعية. تتطلب هذه الحسابات معرفة الانحراف المعياري للمجتمع . استخدم هذه الحسابات للأسباب التالية:

قبل أن تجمع بيانات لاختبار Z على عينة واحدة ، لضمان أن يكون للاختبار حجم عينة مناسب لتحقيق طاقة مقبولة

متال

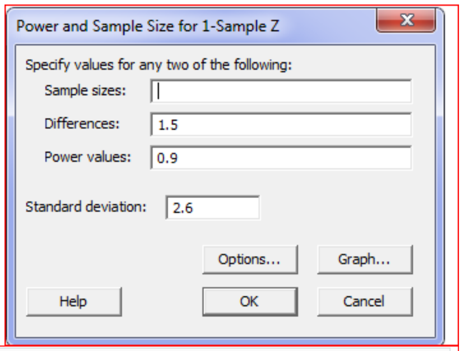
تريد إحدى الشركات التي تصنع الأغذية المصنعة تقييم نسبة الدهون في صلصة الشركة المعبأة في زجاجات. النسبة المعلن عنها هي 15 ٪. يقيس العالم نسبة الدهون في 20 عينة عشوائية. وجدت القياسات السابقة أن الانحراف المعياري للسكان هو 2.6 ٪.

قبل جمع البيانات لاختبار Z على عينة واحدة ، تستخدم الشركة حسابًا لحجم القوة والعينة لتحديد حجم العينة المطلوب للحصول على قدرة 0.9 ولاكتشاف وجود اختلاف قدره 1.5٪ أو أكبر.

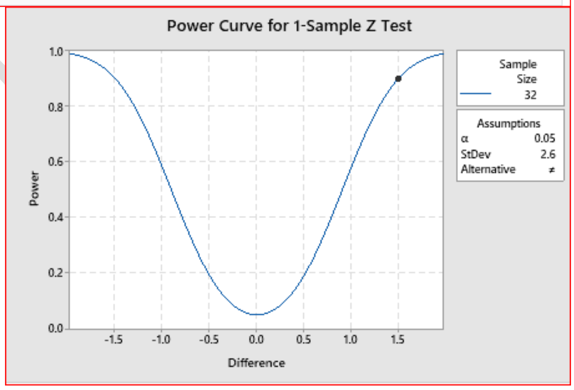
الحل : نكتب الامر الاتي

Stat > Power and Sample Size > 1-Sample Z

فيظهر المربع الحواري الاتي




```
Power and Sample Size

1-Sample Z Test

Testing mean = null (versus ≠ null)
Calculating power for mean = null + difference
α = 0.05  Assumed standard deviation = 2.6

Results

             Sample  Target
Difference     Size   Power  Actual Power
       1.5       32     0.9      0.903816
```



للكشف عن وجود اختلاف بنسبة 1.5 ٪ بقوة 0.9 ، يحتاج الشركةإلى جمع حجم عينة من 32. تحدد أن حجم العينة 32 معقول ، ونستمر في جمع البيانات

2-1-3 اختبا ر فرضية حجم عينة 1-Sample t test:

نستخدم هذا الاختبار لفحص العلاقة بين قوة الاختبار ، وحجم العينة ، والفرق عندما تريد مقارنة متوسط جماعة ما بالهدف أو القيمة المرجعية. لا تتطلب منا هذه الحسابات معرفة الانحراف المعياري للمجتمع . نستخدم ، قبل أن جمع بيانات من أجل اختبار t لعينة واحدة ، لضمان أن يكون للاختبار حجم عينة مناسب لتحقيق قدرة مقبولة على التقدير

مثال

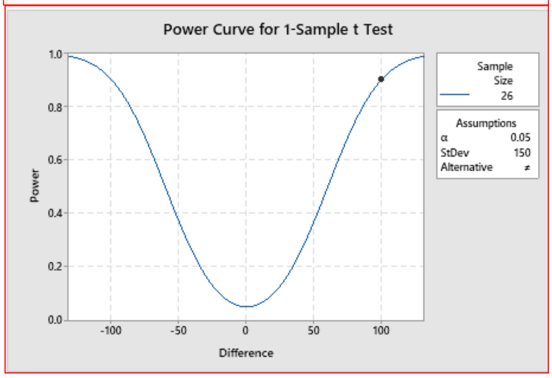
قبل جمع البيانات من أجل اختبار t لعينة واحدة ، يستخدم الخبير الاقتصادي حساب قوة الاختبار و حجم العينة لتحديد مدى حجم العينة التي يجب أن تكون للحصول على قوة اختبار قدرها 90٪ (0.9). أي فرق بقيمة 100 دولار على الأقل في أي من الاتجاهين يعتبر ذا معنى ويكون الانحراف المعياري المقدر هو 150 دولارًا.

الحل :

نقوم بكتابة التعليمة التالية Stat > Power and Sample Size > 1-Sample t.

ثم ندون المربع الحواري التالي ونضع فيه معطيات المثال

```
Power and Sample Size

1-Sample t Test

Testing mean = null (versus ≠ null)
Calculating power for mean = null + difference
α = 0.05  Assumed standard deviation = 150

Results

             Sample  Target
Difference     Size   Power  Actual Power
       100       26     0.9      0.904254
```



للكشف عن اختلاف قدره 100 بقوة 0.9 ، يحتاج الاقتصادي إلى جمع عينة من 26 ملاحظة. هذا حجم عينة يمكن الحصول عليه ، لذلك يستمر الخبير الاقتصادي في جمع البيانات واختبار t المكون من عينة واحدة.

3-1-3 اختبار فرضية حجم عينة للنسبة :

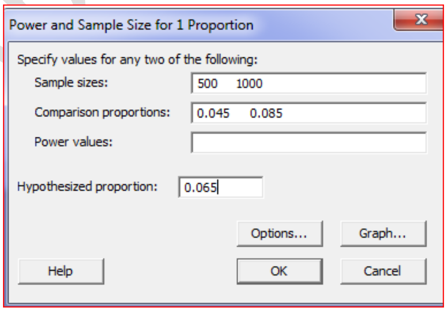
يستخدم هذا الاختبار لفحص العلاقة بين القوة وحجم العينة ونسبة المقارنة عندما تريد مقارنة نسبة المجتمع بالهدف أو القيمة المرجعية. تتطلب هذه الحسابات أن تحتوي البيانات على فئتين فقط ، مثل النجاح / الفشل. مثلا

مثال

يريد اخصائي تسويق تحديد ما إذا كانت الإعلانات المرسلة بالبريد التي تم إرسالها إلى عينة عشوائية من الأسر تؤدي إلى معدل استجابة يختلف عن المعدل الوطني البالغ 6.5٪ (القيمة المستهدفة). قبل جمع البيانات لاختبار نسبة 1 ، يستخدم المحلل حساب حجم العينة. يريد هذا الاخصائي تحديد مدى قوة الاختبار عندما يكون حجم العينة إما 500 أو 1000 ويمكن للاختبار اكتشاف نسبة مقارنة بين 4.5٪ و 8.5٪.

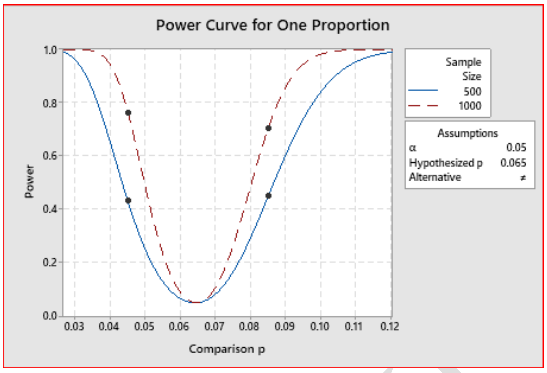
الحل

نختار التعليمة Power and Sample Size > 1 Proportion ، ونملء المربع الحواري بمعطيات المثال



فنحصل على النتيجة المبينة في الصفحة

```
Power and Sample Size

Test for One Proportion

Testing p = 0.065 (versus ≠ 0.065)
α = 0.05

Results

                  Sample
Comparison p        Size     Power
       0.045         500  0.431131
       0.045        1000  0.764259
       0.085         500  0.449114
       0.085        1000  0.703796


Power Curve for One Proportion
```



بواسطة حجم عينة من 500 مشاهدة ، سيكون للاختبار قدرة 0.431 و 0.449 للكشف عن نسبة المقارنة بين 0.045 و 0.085. مع حجم العينة 1000 ، سيكون للاختبار قدرة 0.764 و 0.704 للكشف عن نسبة مقارنة بين 0.045 و 0.085. وبالتالي سوف يقرر هذا الاخصائي أن 0.764 لا تكفي ، ويجمع حجم عينة أكبر من 1000. مشاهدة

4-1-3 قوة اختبار وحجم عينة ذات نسبة بواسونية Power and Sample Size for 1-Sample Poisson Rate

نستخدم القوة وحجم العينة لنسبة بواسون في عينة واحدة لفحص العلاقة بين القوة وحجم العينة ومعدل المقارنة عندما تريد مقارنة معدل حدوث مجتمع ما بالقيمة المستهدفة أو المرجعية. تتطلب هذه الحسابات أن تكون بياناتك مهمة لكل وحدة.

مثال :

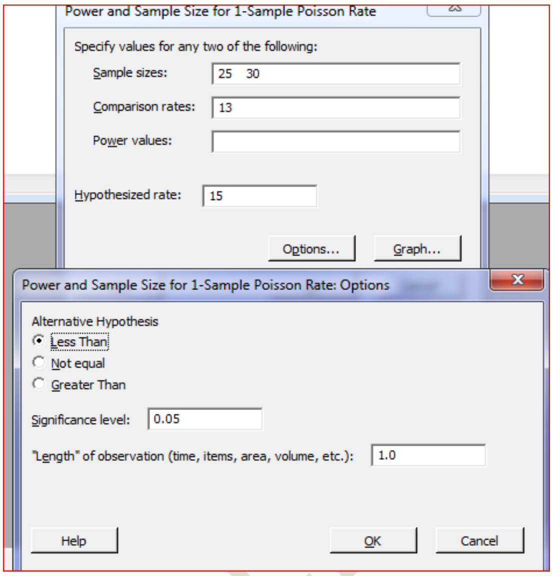
ترغب شركة تصنيع السيارات في تحديد ما إذا كان عدد العيوب الموجودة على أبواب السيارة قبل تجميع السيارات أقل بكثير من 15. قبل جمع البيانات لاختبار Poisson ذي العينة الواحدة ، تستخدم الشركة المصنعة حساب حجم العينة. يريد الصانع تحديد مدى قوة الاختبار عندما يكون حجم العينة 25 أو 30 وعندما يمكن للاختبار اكتشاف معدل مقارنة لا يقل عن 13.

الحل

نكتب الامر التالي

Stat > Power and Sample Size > 1-Sample Poisson Rate

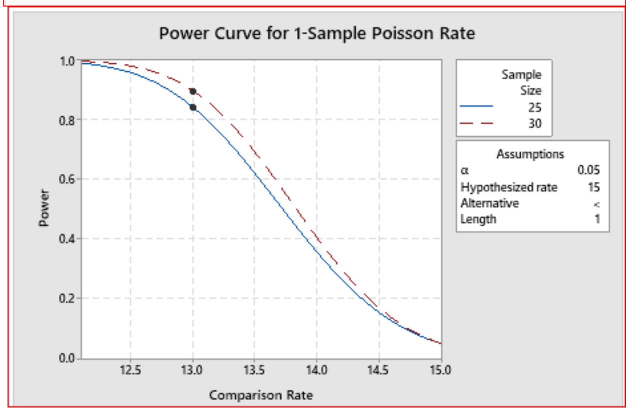
فيظهر لنا المربع الحواري ونقوم بتدوين معطيات المثال فيه



نلاحظ ان الفرضية البديلة هي Less Than اي اقل من 13

والنتيجة هي

```
Power and Sample Size

Test for 1-Sample Poisson Rate

Testing rate = 15 (versus < 15)
α = 0.05
"Length" of observation = 1


Comparison  Sample
      Rate    Size     Power
        13      25  0.842947
        13      30  0.898200
```



حتى نكشف أن معدل المقارنة هو 13 ، ستكون قوة الاختبار هي 0.843 عندما يكون حجم العينة 25 ، وقوة 0.898 عندما يكون حجم العينة 30. وتقرر الشركة المصنعة أن 0.843 هي كافية ، و بالتالي يعتبر حجم عينة من 25. مشاهدة كاف

5-1-3 قوة اختبار و حجم العينة للتباين Power and Sample Size for 1 Variance

نستخدم هذا الاختبار لفحص العلاقة بين قوة الاختبار وحجم العينة والنسبة التي نريد فيها مقارنة التباين أو الانحراف المعياري لمجتمع ما بالقيمة المستهدفة أو المرجعية

متال

نريد تقييم أداء الى تقطع الحزم التي يفترض أن يبلغ طولها 100 سم. يخطط مسير هذه الالة لإجراء اختبار التباين لتحديد تباينها.

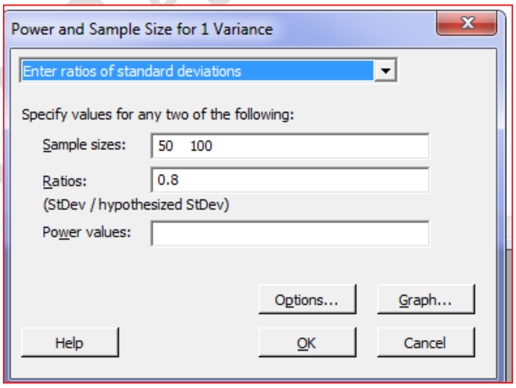
قبل جمع البيانات لاختبار التباين ، نستخدم حساب قوة وحجم العينة لتحديد مدى قوة الاختبار عندما تكون أحجام العينة 50 و 100 ويكتشف الاختبار نسبة 0.8 بين المقارنة و الانحرافات.المعيارية الافتراضية

الحل

نقوم بكتابة الامر الموافق

Stat > Power and Sample Size > 1 Variance

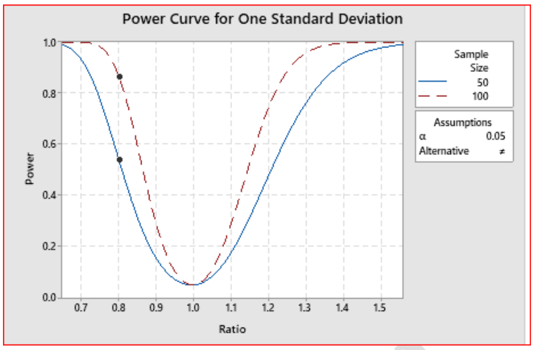
تم نستخرج المربع الحواري ، وندون فيه بيانات المثال



اماالنتيجة فهي

Power and Sample Size

```
Test for One Standard Deviation

Testing StDev = null (versus ≠ null)
Calculating power for (StDev / null) = ratio
α = 0.05

Results

        Sample
Ratio    Size     Power
  0.8      50  0.539065
  0.8     100  0.865153


Power Curve for One Standard Deviation
```



لاكتشاف نسبة 0.8 ، يمكن الحصول على قوة 0.539 بحجم عينة 50 مشاهدة وقوة اختبار 0.865 بحجم عينة 100 مشاهدة. للحصول على قوة اختبار كافية للكشف عن نسبة 0.8 ، ولذا يجب جمع 100 عينة

.

2-3 في حالة عينتين :

1-2-3 قوة اختبار وحجم عينتين مستقلتين Power and Sample Size for 2-Sample t :

نستخدم هذا الاختبار لفحص العلاقة بين قوة الاختبار وحجم العينة والفرق عندما تريد مقارنة الفرق بين مجتمعين طبيعيين مستقلين لكن الفرق بينهما يتبع قانون ستيودنت .

مثال

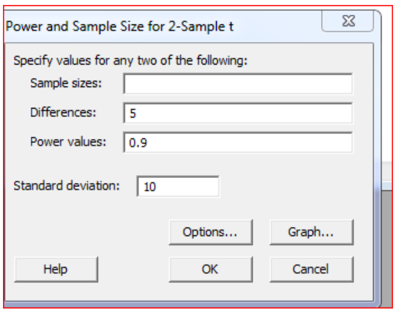
يريد مستشار الرعاية الصحية مقارنة معدلات رضا المرضى في مستشفيين. قبل جمع البيانات لاختبار عينة ، يستخدم الاستشاري حساب حجم العينة لتحديد حجم العينة المطلوبة للكشف عن اختلاف 5 مع احتمال يصل إلى 90 ٪ (قوة 0.9). تشير الدراسات السابقة إلى أن التقييمات لها انحراف معياري قدره 10.

الحل

نقوم بكتابة الامر التالي

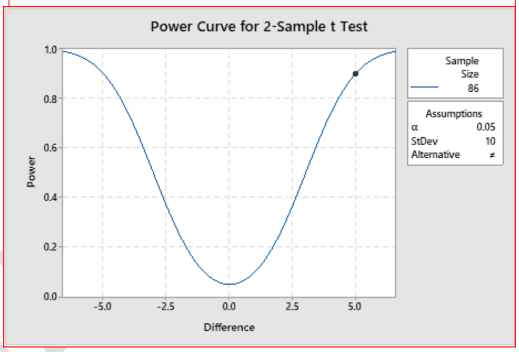
.Stat > Power and Sample Size > 2-Sample t

فنحصل على مربع حواري ندون فيه المعلومات



وتكون النتيجة هي

```
Power and Sample Size

2-Sample t Test

Testing mean 1 = mean 2 (versus ≠)
Calculating power for mean 1 = mean 2 + difference
α = 0.05  Assumed standard deviation = 10

Results

            Sample  Target
Difference    Size   Power  Actual Power
         5      86     0.9      0.903230

The sample size is for each group.


Power Curve for 2-Sample t Test
```



للكشف عن فرق 5 بقوة 0.9 ، يحتاج الاستشاري إلى جمع الحد الأدنى لحجم العينة وهو 86. نظرًا لأن قوة الاختبار المستهدفة البالغة 0.9 تؤدي إلى حجم عينة ليس عددًا صحيحًا ، يعرض Minitab أيضًا الطاقة (القدرة الفعلية)) لحجم العينة مدورة.

3-2-2 قوة اختبار وحجم عينتين مرتبطتين Power and Sample Size for Paired t :

نستخدم هذا الاختبار بين عينتين مرتبطتين في مجتمع واحد الطاقة، وحجم العينة الزوجين لفحص العلاقة بين الطاقة وحجم العينة والفرق عندما تريد مقارنة معلمات المجتمع الطبيعي بناءً على الملاحظات المزدوجة.

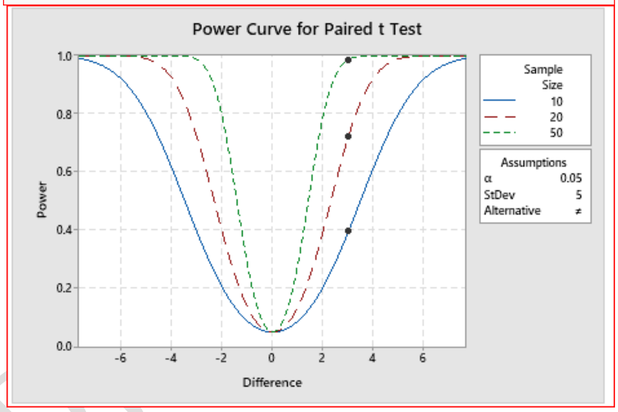
مثال

يريد مدير في مركز للياقة البدنية تحديد ما إذا كان برنامج انقاص الوزن فعال بحيث يوجد فرق لا يقل عن 3 أرطال. في دراسة سابقة ، قرر المدير أن الانحراف المعياري للفروق المزدوجة هو 5. قبل جمع البيانات لاختبار t بقيم مرتبطة ، يستخدم المدير حساب حجم وحجم العينة لتحديد مدى قوة الاختبار مع أحجام عينة مختلفة. وهي 10 20 50

الحل

ندون الامر Stat > Power and Sample Size > Paired t. ونملأ المربع الحواري

```
Power and Sample Size

Paired t Test

Testing mean paired difference = 0 (versus ≠ 0)
Calculating power for mean paired difference = difference
α = 0.05  Assumed standard deviation of paired differences = 5

Results

             Sample
Difference     Size     Power
         3       10  0.395918
         3       20  0.721005
         3       50  0.986031


Power Curve for Paired t Test
```



للكشف عن اختلاف قدره 3 أرطال في برنامج انقاص الوزن ، يمكن للمدير الحصول على قوة اختبار تبلغ حوالي 0.4 مع حجم عينة 10 ، وقوة اختبار حوالي 0.72 مع حجم عينة من 20 ، وقوة اختبار حوالي 0.99 بحجم العينة 50. لا يمنح حجم العينة 20 أو أقل الاختبار قوة كافية للكشف عن اختلاف قدره 3 ، وقد يمنح حجم العينة 50 مشاهدة قوة اختبار كبيرة جدًا.

3-2-3 قوة اختبار وحجم العينة لنسبتين Power and Sample Size for 2 Proportions :

نستخدم قوة الاختبار وحجم العينة لنسبتين لدراسة العلاقة بين القوة وحجم العينة ونسبة المقارنة عندما تريد مقارنة الفرق بين نسبي للمجتمع الاحصائي . تتطلب هذه الحسابات أن تحتوي البيانات على فئتين فقط ، مثل النجاح / الفشل مثلا وفي العينات الكبيرة فلن توزيع النسب عادة يقرب من توزيع ذي الحدين الى الطبيعي

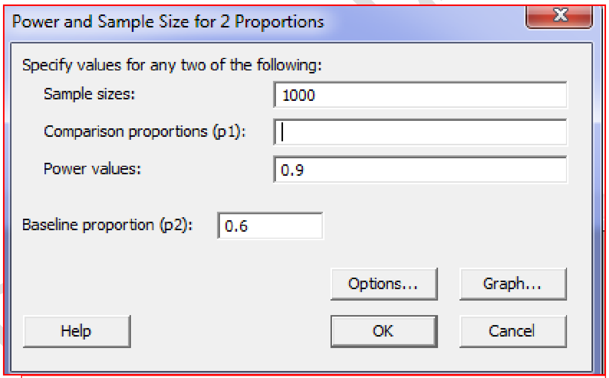
متال

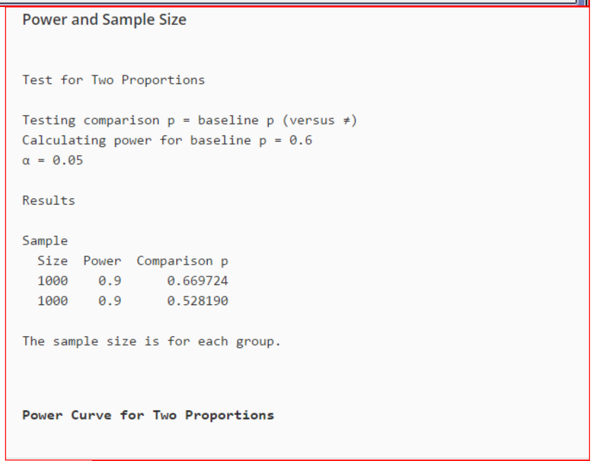
يريد مسؤول المساعدات المالية بالجامعة تحديد ما إذا كان من المرجح أن يحصل الطلاب على وظيفة في فصل الصيف. تشير نتائج دراسة سابقة إلى أن 60٪ من الطلاب يحصلون على وظيفة في فصل الصيف. قبل جمع البيانات لاختبار النسبتين ، يستخدم هذا المسؤول حساب حجم وحجم العينة

لتحديد مدى اختلاف الفرق الذي يمكن للاختبار اكتشافه عندما يكون حجم العينة 1000 وتكون الطاقة 0.9.

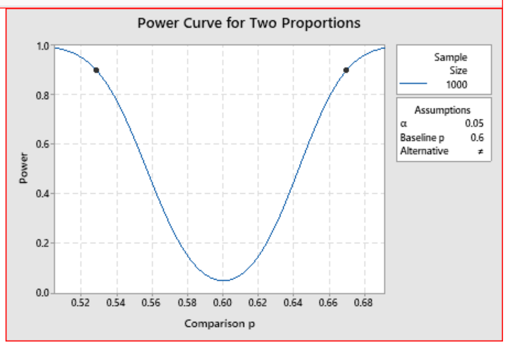
الحل

Stat > Power and Sample Size > 2 Proportions.





```
Power and Sample Size

Test for Two Proportions

Testing comparison p = baseline p (versus ≠)
Calculating power for baseline p = 0.6
α = 0.05

Results

Sample
  Size  Power  Comparison p
  1000    0.9      0.669724
  1000    0.9      0.528190

The sample size is for each group.


Power Curve for Two Proportions
```



بحجم العينة 1000 وقوة اختبار 0.9 ، يمكن للموظف اكتشاف الفرق بين نسب حوالي 7 ٪ في أي من الاتجاهين. هذا الاختلاف كافٍ ، لذلك يقوم بجمع البيانات لتحليل النسبتين.

4-2-3قوة واختبار حجم العينة لنسبتي بواسون Power and Sample Size for 2-Sample Poisson Rate :

نستخدم القوة وحجم العينة لنسبتي بواسون عينتين لفحص العلاقة بين القوة وحجم العينة ومعدل المقارنة عندما تريد مقارنة الفرق بين نسبتي مجتمعين بواسونيين

مثال

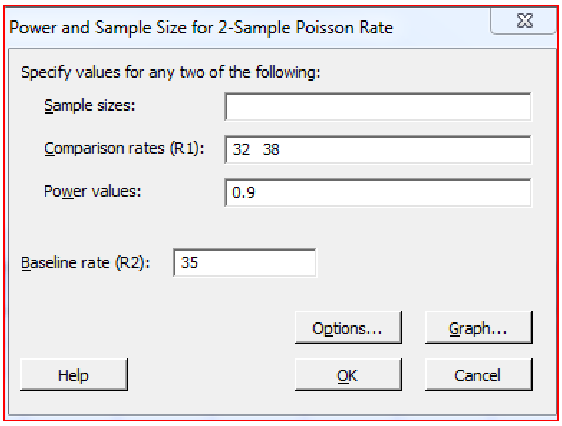
يريد مستشار السلامة المرورية مقارنة عدد السيارات في الساعة التي تسير في شارعين مختلفين. قبل جمع البيانات لاختبار معدل بواسون ذي عينتين ، يستخدم الخبير الاستشاري حساب حجم العينة. يريد الاستشاري تحديد حجم حجم العينة الذي يحتاجه الاختبار للحصول على قوة اختبار قدرها 0.9 واكتشاف معدل مقارنة 32 أو 38 (أي بفارق 3 عن معدل خط الأساس البالغ 35).

الحل

لحل هذا المثال نكتب الامر

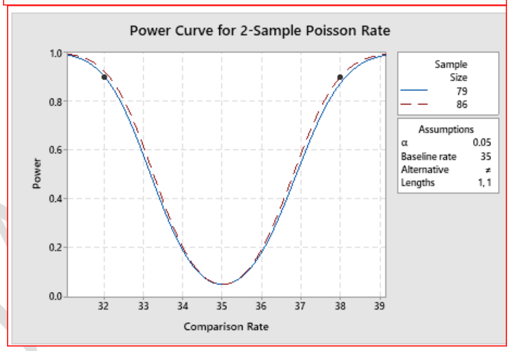
Power and Sample Size > 2-Sample Poisson Rate.

فيظهر لنا مربع حواري نكتب فيه معطيات المثال



وتكون صفحة النتائج هي

```
Power and Sample Size

Test for 2-Sample Poisson Rate

Testing comparison rate = baseline rate (versus ≠)
Calculating power for baseline rate = 35
α = 0.05
"Lengths" of observation for sample 1, sample 2 = 1, 1

Results

Comparison  Sample  Target
      Rate    Size   Power  Actual Power
        32      79     0.9      0.902793
        38      86     0.9      0.902550

The sample size is for each group.


Power Curve for 2-Sample Poisson Rate
```



للكشف عن معدل مقارنة 32 بقوة اختبار 0.9 ، يحتاج الاستشاري إلى حجم عينة 79. للكشف عن معدل مقارنة 38 بقيمة طاقة 0.9 ، يحتاج الاستشاري إلى حجم عينة من 86. يقرر المحلل لجمع حجم عينة من 86 لإعطاء الاختبار قوة لا تقل عن 0.9 لكل معدلات المقارنة.

5-2-3 قوة وحجم عينة لتباينين Power and Sample Size for 2 Variances :

نستخدم هذا الاختبار لفحص العلاقة بين قوة ، وحجم العينة ، ونسبة تباينين عندما تريد مقارنة النسبة بين تبايني مجتمعين أو انحرافات معيارية لهدف أو قيمة مرجعية.

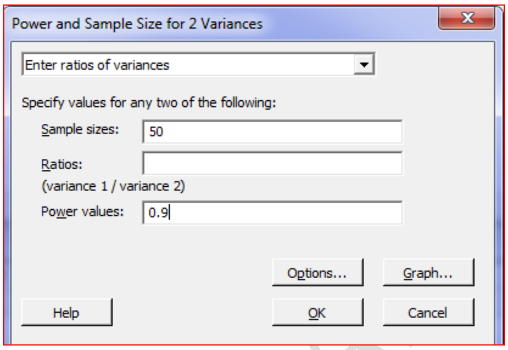
مثال

تريد شركة تصنيع السيارات مقارنة الفروق بين محركين مختلفين . قبل جمع البيانات لاختبار التباين ، يستخدم المدير حساب حجم وحجم العينة لتحديد النسبة التي يمكن اكتشافها عندما يكون حجم كلتا العيّنتين 50 وقوة الاختبار هي 0.9.

الحل

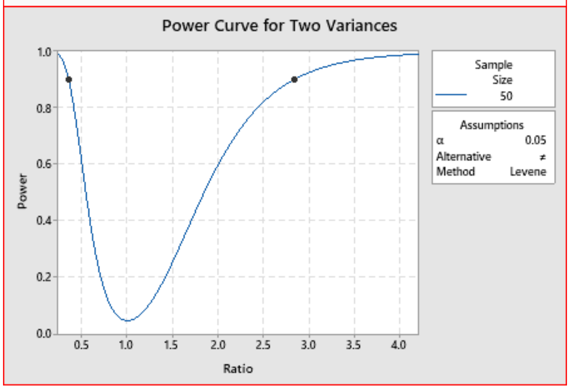
من اجل حل هذا المثال ببرنامج Minitab نكتب الامر التالي Power and Sample Size > 2 Variances.

ثم نكتب المعطيات التالية في المربع الحواري



ثم تظهر النتائج مبينة فيما يلي

```
Power and Sample Size

Test for Two Variances

Testing (variance 1 / variance 2) = 1 (versus ≠)
Calculating power for (variance 1 / variance 2) = ratio
α = 0.05
Method:  Levene's Test

Results

Sample
  Size  Power    Ratio
    50    0.9  2.83320
    50    0.9  0.35296

The sample size is for each group.


Power Curve for Two Variances
```



4 -اختبارات التعادل:

4-1 قوة وحجم عينة لاختبار تكافؤ عينة واحدة Power and Sample Size for 1-Sample Equivalence Test:

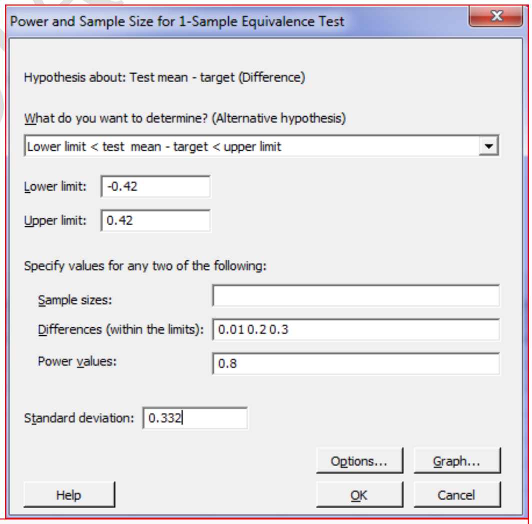
نستخدم   القوة وحجم العينة لاختبار تكافؤ عينة واحدة لفحص العلاقة بين القوة وحجم العينة والفرق عندما تريد تقييم التكافؤ بين متوسط المنتج أو العملية والقيمة المستهدفة. كما نستخدم هذه الاختبار لاختبار تكافؤ عينة واحدة ، للتأكد من أن حجم عينة كاف لتحقيق قوة   اختبار مقبولة بعد اختبار تكافؤ عينة واحدة ، لتحسين التصميم للدراسة القادمة

مثال

يريد مصنع التغليف اختبار طريقة جديدة لختم أكياس الوجبات الخفيفة. يجب أن تكون القوة المطلوبة لفتح الأكياس في حدود 10٪ من القيمة المستهدفة البالغة 4.2 نيوتن . قبل جمع البيانات لاختبار تكافؤ عينة واحدة ، يرى المصنع لتحديد حجم العينة يجب الحصول على قوة اختبار بنسبة 80 ٪ (0.8). من العينات السابقة ، يقدر المصنع ان الانحراف المعياري للمجتمع هو 0.332.

الحل

لحل هذا المثال نكتب الامر الاتي Stat > Power and Sample Size > Equivalence Tests > 1-Sample.
وندون المعلومات في المربع الحواري




```
Power and Sample Size

1-Sample Equivalence Test

Method

Power for difference:        Test mean - target
Null hypothesis:             Difference ≤ -0.42 or Difference ≥ 0.42
Alternative hypothesis:      -0.42 < Difference < 0.42
α level:                     0.05
Assumed standard deviation: 0.332

Results

             Sample  Target
Difference     Size   Power  Actual Power
       0.0        7     0.8      0.805075
       0.1        9     0.8      0.834590
       0.2       16     0.8      0.811465
       0.3       49     0.8      0.802154
```

إذا كان الفرق هو 0 (القوة المتوسطة في الهدف) ، فإن المصنع يحتاج إلى حجم عينة من 7 مشاهدات لتحقيق قوة اختبار قدرها 0.8. إذا كان المصنع يستخدم حجم عينة من 9 مشاهدات ، فإن قوة الاختبار تتجاوز 0.9 للفرق 0. عندما يكون الفرق أقرب إلى حد التكافؤ العلوي (0.42) ، يحتاج المصنع إلى حجم عينة أكبر لتحقيق نفس القدرة. على سبيل المثال ، لفرق قدره 0.3 ، يحتاج المهندس إلى حجم عينة يبلغ 49 للوصول إلى قوة قدرها 0.8. بالنسبة لأي حجم للعينة ، مع اقتراب الفرق من حد التكافؤ الأدنى أو حد التكافؤ العلوي ، تقل قوة الاختبار وتقترب من α ، وهو خطر المطالبة بالتعادل عندما لا يكون صحيحًا).

4 -2 قوة وحجم عينة لاختبار تكافؤ عينتين Power and Sample Size for 2-Sample Equivalence Test

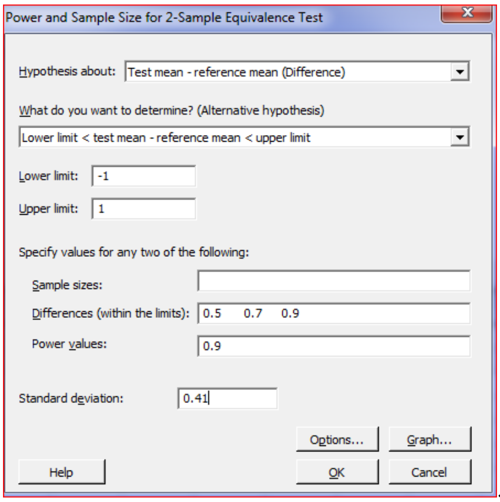
استخدم القوة وحجم العينة لاختبار العينة لفحص العلاقة بين القوة وحجم العينة والفرق عندما تريد تقييم التكافؤ بين متوسط الاختبار والمتوسط المرجعي للعينات المستقلة.

مثال

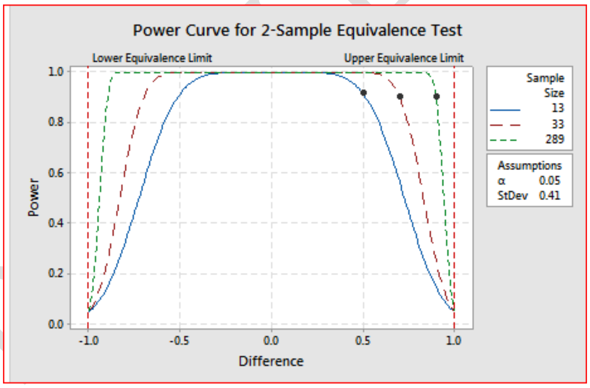
من اجل الجودة تحديد ما إذا كان متوسط كمية المكونات النشيطة في علامة تجارية عامة من مسكنات الألم هو ضمن 1 ملغ من المبلغ المتوسط تجارية. قبل جمع البيانات لاختبار تكافؤ -عينتين ، نستخدم حساب حجم العينة لتحديد الحجم الممكن للعينة للحصول على قوة 90٪ (0.9). من العينات السابقة ، ويقدر المحلل الانحراف المعياري للسكان هو 0.41.

الحل

Stat > Power and Sample Size > Equivalence Tests > 2-Sample



```
Power and Sample Size

2-Sample Equivalence Test

Method

Power for difference:        Test mean - reference mean
Null hypothesis:             Difference ≤ -1 or Difference ≥ 1
Alternative hypothesis:      -1 < Difference < 1
α level:                     0.05
Assumed standard deviation:  0.41

Results

            Sample  Target
Difference    Size   Power  Actual Power
       0.5      13     0.9      0.915407
       0.7      33     0.9      0.902461
       0.9     289     0.9      0.900360

The sample size is for each group.
```



إذا كان الفرق 0.5 ، فإن المحلل يحتاج إلى 13 ملاحظة في كل مجموعة لتحقيق قوة اختبار لا تقل عن 0.9. إذا كننا نستخدم حجم عينة يبلغ 13 ، فإن قوة الاختبار تبلغ 0.92 تقريبًا. عندما يكون الفرق أقرب إلى حد التكافؤ الأدنى (-1) أو حد التكافؤ العلوي (1) ، فإن المحلل يحتاج إلى حجم عينة أكبر لتحقيق نفس القوة. على سبيل المثال ، لفارق 0.9 ، كما نحتاج إلى حجم عينة لا يقل عن 289 ملاحظة في كل مجموعة لتحقيق قوة اختبار 0.9 .

بالنسبة لأي حجم للعينة ، مع اقتراب الفرق من حد التكافؤ الأدنى أو حد التكافؤ العلوي ، تقل قوة الاختبار وتقترب من α (ألفا) ، وهو خطر المطالبة بالتعادل عندما لا يكون صحيحًا).

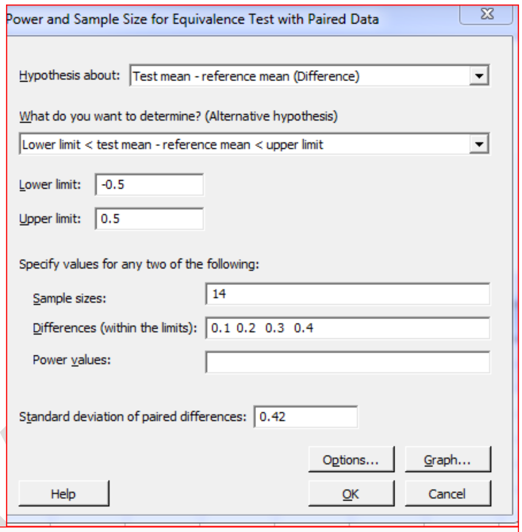
3-4 قوة وحجم عينة لاختبار تكافؤ عينتين مرتبطتين Power and Sample Size for Equivalence Test with Paired Data

يعتبر هذا الاختبار مهما اذا ما تعلق الامر بقوة وحجم العينة لاختبار التكافؤ مع البيانات المقترنة لفحص العلاقة بين القوة وحجم العينة والفرق عندما نريد تقييم التكافؤ بين متوسط الاختبار والمتوسط مرجعي باستخدام الملاحظات المزدوجة.

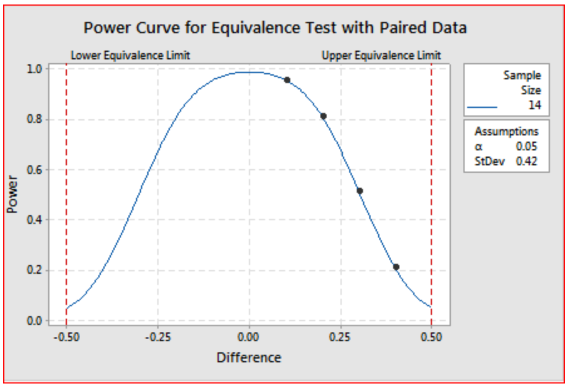
مثال

يريد مخبر لتركيب العدسات اللاصقة اختبار طريقة تركيب جديدة علما ان لدى المخبر 14 زبونا يوميا ،.لكي تكون الطريقة المقترحة فعالة ، يجب أن تكون زاوية الرؤية للطريقة الجديدة ضمن ±0.5 درجة لحساب حجم العينة لتحديد ما إذا كان حجم العينة البالغ 14 يوفر قوة اختبار كافية من العينات السابقة ، يقدر المخبر الانحراف المعياري للمجتمع هو 0.42.

الحل :




```
Power and Sample Size

Equivalence Test with Paired Data

Method

Power for difference:   Test mean - reference mean
Null hypothesis:        Difference ≤ -0.5 or Difference ≥ 0.5
Alternative hypothesis: -0.5 < Difference < 0.5
α level:                0.05

Assumed standard deviation of paired differences = 0.42

Results

             Sample
Difference     Size     Power
       0.1       14  0.957371
       0.2       14  0.811858
       0.3       14  0.517255
       0.4       14  0.211869
```



إذا كان الفرق هو 0.1 وكان المخبر يستخدم حجم عينة من 14 زوجًا من الملاحظات ، فإن قوة الاختبار أكبر من 0.9. إذا كان الفرق هو 0.2 وكان المخبر يستخدم حجم عينة مكونة من 14 زوجًا من الملاحظات ، عندئذ يكون للقدرة اختبار يزيد عن 0.8. ومع ذلك ، إذا كان الفرق هو 0.3 ، فالمخبر يستخدم حجم عينة مكونة من 14 زوجًا من الملاحظات ، عندئذٍ تبلغ قوة الاختبار 0.52 تقريبًا ، وهي نسبة غير كافية. عندما يكون الفرق أقرب إلى حد التكافؤ العلوي (0.5) ، تكون قوة الاختبار أقل. على سبيل المثال ، لفرق قدره 0.4 ، إذا استخدم المخبر حجم عينة مكونة من 14 زوجًا من الملاحظات ، عندئذٍ تبلغ قوة الاختبار 0.22 تقريبًا. بالنسبة لأي حجم للعينة ، مع اقتراب الفرق من حد التكافؤ الأدنى أو حد التكافؤ العلوي ، تقل قوة الاختبار وتقترب من α ، وهو خطر المطالبة بالتعادل عندما لا يكون صحيحًا

4 - تدوير العينات :

4- - 1تدوير متوسط عينة: Bootstrapping for 1-Sample Mean

التدوير او اعادة المعاينة (bootstrap)طريقة تقدر توزيع العينات من خلال أخذ عينات متعددة مع الاستبدال من عينة عشوائية واحدة. وتسمى هذه العينات المتكررة عينات. كل resample هو نفس حجم العينة الأصلية

مثال

يريد رياضي في السباقات نصف الطويلة تحديد وقت العدو لبعض السباقات داخل القاعة .

فيحصل على الوقت بالدقائق

16	8.8	9.2	14.8	8.2	10.9	15	9.1
8.8	12.5	9.2	8	7.7	9.2	15.9	15.2

المطلوب مامجال الثقة لتدوير متوسط العينات

الحل : بنقل معطيات هذا الجدول الى ورقة عمل من برنامج الاكسل مثلا MINITAB نسمي المتغير المدروس Time

C1
Time
10.9
15.0
11.9
8.8
8.2
14.8
9.2
8.8
16.0
15.2
15.9
9.2
9.2
7.7
8.0
12.5

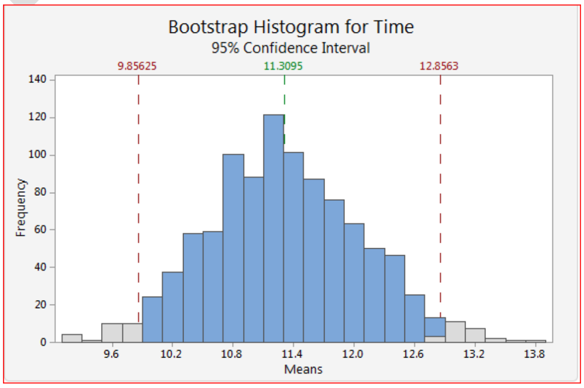
حتى نتمكن من حل هذا المثال نختار الامر التالي في MINITAB 19

Stat > Resampling > Bootstrapping for 1-Sample Mean

في ايقونة Options نختار 1 in Base for random number generator

تفسير النتائج :

يشير مجال الثقة CI 95٪ إلى أن العداء يمكن أن يكون واثقًا بنسبة 95٪ من أن متوسط وقت التفاعل يتراوح بين 9.9 دقيقة و 12.9 دقيقة تقريبًا. يوضح الرسم البياني أن توزيع bootstrap يتوزع طبيعياً ، لذلك يمكن لهذا العداء أن يثق في النتائج.



Observed Sample

Variable	N	Mean	StDev	Variance	Sum	Minimum	Median	Maximum
Time	16	11.3313	3.1149	9.7023	181.300	7.7000	10.0500	16.0000

Bootstrap Samples for Mean

Number of Resamples	Mean	StDev	95% CI for µ
1000	11.3095	0.7625	(9.8563, 12.8563)

2-4 تدوير متوسطي عينتين Bootstrapping for 2-Sample Means :

نستخدم تدوير متوسطي عينتين من اجل الحصول معاينة التوزيع للفرق بين مجتمعين مستقلين ، وتقدير مجال الثقة لهذا الفرق ، ولكي تكون المشاهدات مستقلة يجب الا تعتمد مشاهدة معينة على المشاهدات السابقة ، وعدم وجود هذه الاستقلالية قد يؤثر على صحة النتائج

مثال

يريد مستشار الرعاية الصحية تحديد الفرق في تقييمات رضا المرضى عن الخدمات الطبية المقدمة لهم وقد اجرى استبيانا مباشرا يجاوب فيه مرضى مستشفيين A وB عن مدى رضاهم فتحصل على النتائج التالية – نسبة الرضا مقيمة بالنسبة المئوية-

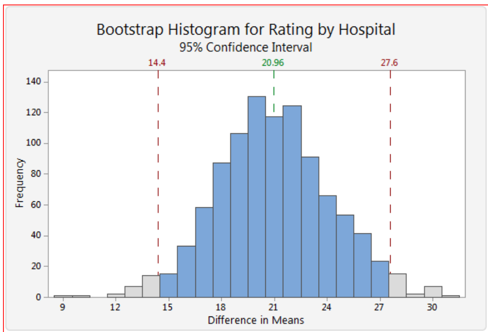
الحل

نرمز لمعدل رضا المريض في كل مستشفى تواليا rating H1و rating H2

C1	C2	C3
rating H1	rating H2	
81	89	
77	64	
75	35	
74	68	
86	69	
90	55	
62	37	
73	57	
91	42	
98	49	
81	59	
85	58	
77	65	
78	71	
83	67	
90	58	
78	63	
76	68	
71	55	
80	57	

نقوم بكتابة الامر التالي

Statistics > Resampling > Bootstrapping for 2-Sample Means



Observed Samples

Hospital	N	Mean	StDev	Variance	Minimum	Median	Maximum
A	20	80.300	8.183	66.958	62.000	79.000	98.000
B	20	59.300	12.431	154.537	35.000	58.500	89.000

Difference in Observed Means

Mean of A - Mean of B = 21

Bootstrap Samples for Difference in Means

Number of Resamples	Average	StDev	95% CI for Difference
1000	20.960	3.279	(14.400, 27.600)

يوضح الرسم البياني أن توزيع الى أن مستشار الخدمات الفندقية يكون CI 95٪ يشير والتوزيع واثقًا بنسبة 95٪ من أن الفرق بين متوسطي المجتمعين يعني أن معدل رضا النزلاء في الفندقين يبدو طبيعياً ، لذلك يمكن للمستشار الوثوق بالنتائج bootstrapيتراوح بين 14.4 و 27.6.

3- 4 تدوير نسبة عينة Bootstrapping for 1-Sample Proportion :

استخدم Bootstrapping لنسب عينة واحدة لاستكشاف توزيع أخذ العينات لنسبة عينة من البيانات ولتقدير مجال الثقة لنسبة المجتمع

متال

اذاكان عدد التجارب هو 200 وعدد الحوادث الممكنة هو124 استخدام تدوير نسبة عينة ل1000 مشاهدة

الحل

نقوم بكتابة الامرالتالي

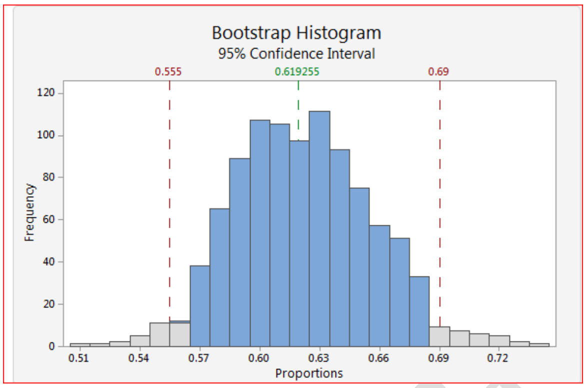
Statistics > Resampling > Bootstrapping for 1-Sample Proportion

ثم نختار summarized data

عدد النجاحات .In Number of events, enter 124

عدد المشاهدات .In Number of trials, enter 200

.On the Options tab, enter 1 in Base for random number generator



Observed Sample

N	Proportion
200	0.6200

Bootstrap Samples for Proportion

Number of Resamples	Average	95% CI for Proportion
1000	0.6193	(0.5550, 0.6900)

يشير 95٪ CI إ نه يمكن أن يكون واثقّين بنسبة 95٪ من أن نسبة المجتمع تتراوح بين 0.56 و 0.69 تقريبًا. يوضح الرسم البياني أن توزيع bootstrap يبدو طبيعياً ، بحيث يمكن الوثوق في النتائج.

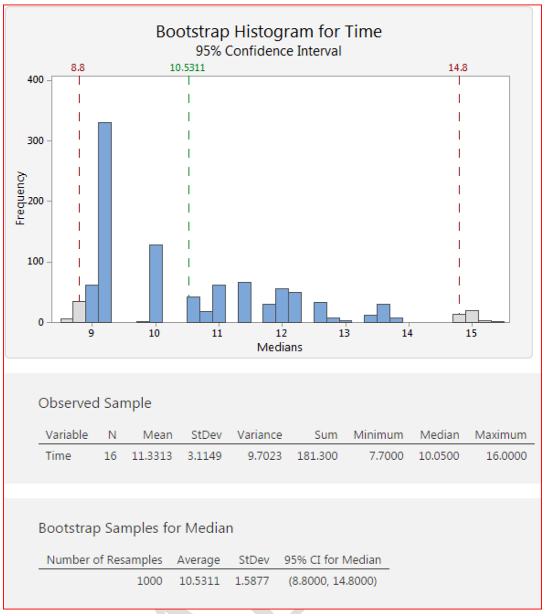
4-4 تدوير دوال العينة Bootstrapping for 1-Sample Function :

يتم استخدام Bootstrapping لاستكشاف توزيع العينات لإحصائية محددة لعينة من البيانات ولتقدير مجال الثقة لمعلمة المجتمع . يمكننا الاختيار من الإحصاءات التالية: الوسيط ، مجموع المربعات ، التباين ، الانحراف المعياري

متال : ناخذ مثال رقم 12 ونغير المطلوب للوسيط Median

الحل : بعد كتابة معطيات المثال في ورقة العمل نقوم بالامر التالي

Statistics > Resampling > Bootstrapping for 1-Sample Function



يشير CI 95٪ إلى أن الكيميائي يمكن أن يكون واثقًا بنسبة 95٪ من أن قيمة وقت رد الفعل وسيط المجتمع تتراوح بين حوالي 8.8 دقيقة و 14.8 دقيقة. ومع ذلك ، يوضح الرسم البياني أن الأشرطة متناثرة للغاية ، مما يجعل من الصعب رؤية توزيع العينات. لأن توزيع العينات غير واضح ، قد يكون فترة الثقة غير موثوق بها. تحتوي العينة الأصلية على 16 مشاهدة بيانات فقط. للحصول على مجال ثقة موثوق ، يجب على الكيميائي جمع عينة أكبر وإجراء التحليل مرة أخرى

5- اختبارات التوزيع العشوائي :

1-5 اختبار التوزيع العشوائي لمتوسط عينة واحدة Randomization Test for 1-Sample Mean:

يمكن استخدام اختبار التوزيع العشوائي لمتوسط عينة واحد لمقارنة متوسط المجتمع بالقيمة المستهدفة أو القيمة المرجعية.

لتاخذ بيانات المثال

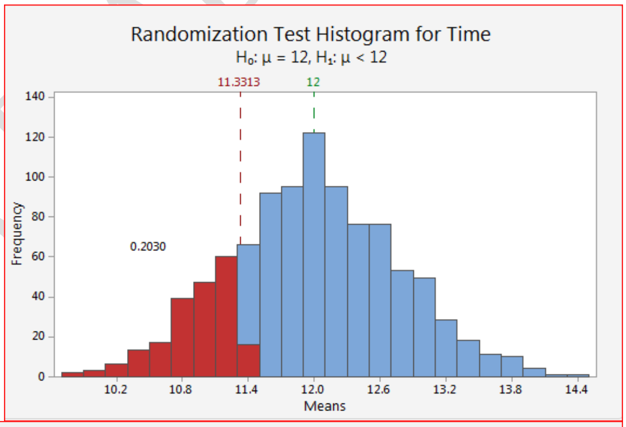
يريد الكيميائي لشركة صيدلانية تحديد ما إذا كان متوسط وقت رد الفعل لمضادات الحموضة المطورة حديثًا أقل من 12 دقيقة.

الحل تصبح الحل عبارة عن اختبار فرضيات

C1
Time
10.9
15.0
11.9
8.8
8.2
14.8
9.2
8.8
16.0
15.2
15.9
9.2
9.2
7.7
8.0
12.5

Statistics > Resampling > Randomization Test for 1-Sample Mean

في ايقونة In Sample نختار Time في ايقونة Hypothesized mea2 ندخل رقم 12 . في ايقونة Select Options نختار Alternative hypothesis, select Mean < hypothesized value . اما في مستطيل Base for random number generator 1 in ثم ok



Observed Sample

Variable	N	Mean	StDev	Variance	Sum	Minimum	Median	Maximum
Time	16	11.3313	3.1149	9.7023	181.300	7.7000	10.0500	16.0000

Randomization Test

Null hypothesis H_0: μ = 12
Alternative hypothesis H_1: μ < 12

Number of Resamples	Mean	StDev	P-Value
1000	11.9783	0.7625	0.2030

تنص الفرضية البديلة على أن متوسط زمن التفاعل أقل من 12 دقيقة. لأن القيمة p هي 0.203 ، وهي أكبر من مستوى المعنوية 0.05 ، يفشل الكيميائي في رفض الفرضية الصفرية ولا يمكنه أن يستنتج أن متوسط زمن التفاعل أقل من 12 دقيقة. كما يوضح الرسم البياني أن توزيع bootstrap يبدو طبيعياً ، لذلك يمكن للصيدلي أن يثق في النتائج.

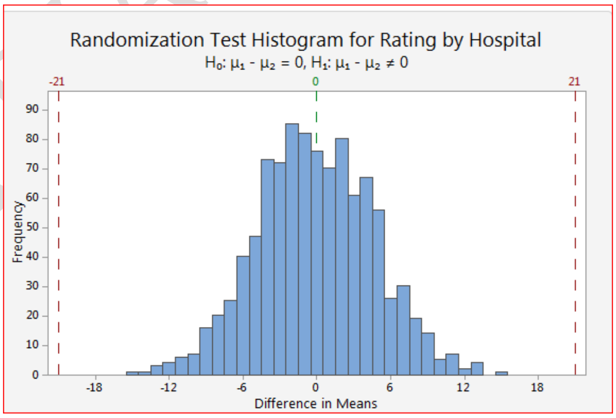
5-2- اختبار التوزيع العشوائي للمتوسط عينتين Randomization Test for 2-Sample Means

نستخدم اختبار التوزيع العشوائي لمتوسط عينتين لتحديد ما إذا كان المجتمعان مستقلتين. لكي تكون الملاحظات مستقلة ، لا تعتمد قيمة ملاحظة معينة على أي ملاحظة سابقة. إذا كانت ملاحظات غير مستقلة ، فقد لا تكون النتائج صالحة

مثال : نتطرق الى المثال رقم 14 ويتم تحويل مطلوبه في شكل اختبار فرضيات نحول ورقة العمل الى MINITAB ونكتب الامر التالي

Statistics > Resampling > Randomization Test for 2-Sample Means

من المربع الحواري نختار الايقونة كلتا العينتين في عمود واحد both samples in one column في مربع Samples نكتب المتغير rating ، اما في المربع Sample IDS نكتب hotel في ايقونة options نكتب in Base for random 1 number generator



Method

μ_1: mean of Rating when Hospital = A
μ_2: mean of Rating when Hospital = B
Difference: $\mu_1 - \mu_2$

Observed Samples

Hospital	N	Mean	StDev	Variance	Minimum	Median	Maximum
A	20	80.300	8.183	66.958	62.000	79.000	98.000
B	20	59.300	12.431	154.537	35.000	58.500	89.000

Difference in Observed Means

Mean of A - Mean of B = 21

Randomization Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$
Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

Number of Resamples	Average	StDev	P-Value
1000	-0.185	4.728	<0.0020

تنص الفرضية الصفرية على أن الفرق في تصنيف المريض بين المستشفيات يساوي 0. لأن القيمة p أقل من 0.002 ، وهو أقل من مستوى الأهمية البالغ 0.05 ، يرفض الاستشاري الفرضية الصفرية ويخلص إلى أن الفرق في تقييمات المرضى بين المستشفيات لا تساوي 0. يوضح الرسم البياني أن توزيع bootstrap يبدو طبيعياً ، لذلك يمكن للمستشار الوثوق في النتائج. الفرق في الوسائل المرصودة هو 21 ، مما يشير إلى أن المستشفى A لديه معدلات رضا المرضى أعلى من المستشفى B

3-5 اختبار التوزيع العشوائي للنسبة Randomization Test for 1-Sample Proportion

اختبار التوزيع العشوائي لنسبة P للعينة الواحدة لمقارنة نسبة المجتمعΠ بالقيمة المستهدفة أو القيمة المرجعية.

متال

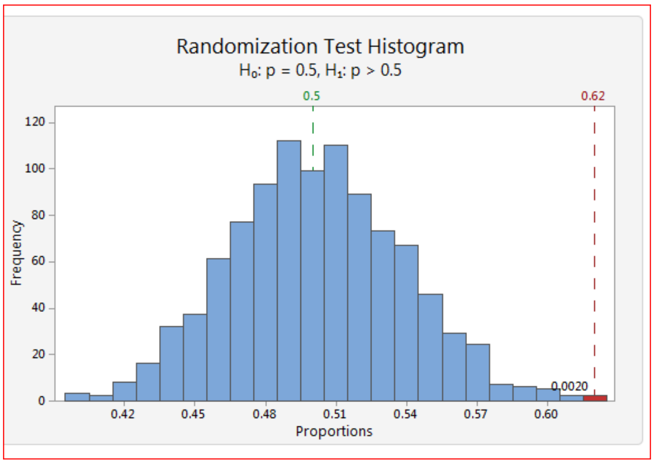
عندما تتم كتابة الامر التالي

Statistics > Resampling > Randomization Test for 1-Sample Proportion

يظهر لنا مربع حواري نختار Summarized ثم ندون المعلومات المفترضة الاتية events,
Number of trials, 200 .124

.Alternative hypothesis, select Proportion > hypothesized value

Base for random number generator نضع 1 ثم نضغط على ok



Observed Sample

N	Proportion
200	0.62

Randomization Test

Null hypothesis	H_0: p = 0.5
Alternative hypothesis	H_1: p > 0.5

Number of Resamples	Average	P-Value
1000	0.49942	0.0020

تنص الفرضية البديلة على أن نسبة أكبر من 0.5. نظرًا لأن قيمة p هي 0.002 ، وهي أقل من مستوى المعنوية 0.05 ، نرفض فرضية العدم ويخلص إلى أن النسبة أكبر من 0.5. يوضح الرسم البياني أن توزيع bootstrap يبدو طبيعياً ، بحيث يمكن الوثوق في النتائج

4 -5 اختبار القيم الشاذة او المتطرفة : Outlier Test

تعرف القيم المتطرفة على انها عبارة عن قيم البيانات التي تختلف بشكل كبير عن غالبية مجموعة من البيانات. تقع هذه القيم خارج اتجاه عام موجود في البيانات. إن الفحص الدقيق لمجموعة من البيانات للبحث عن القيم الخارجية يسبب بعض الصعوبة. سواء كنا بدراسة المجتمع الاحصائي او بصدد تقدير عينة احصائية

للقيم المتطرفة على أنها قيم غير عادية يمكن أن تؤدي إلى حدوث تغييرات سلبية في نتائج التحليل الإحصائي. مما سبق يتضح بأن القيمة المتطرفة هي قيمة تخرج عن النسق المميز لمجموعة البيانات بأن تطرف البيانات يعزى إما لأخطاء حسابيه أو أخطاء قراءه أو أخطاء تسجيل، .إلى أن القيم المتطرفة في مجموعة البيانات قد تظهر بسبب أن البيانات تعود إلى توزيعات غير ومتماثلة بمعنى قد يكون فيها التواء عال إما نحو اليمن أو نحو اليسار

ويعتمد هذا الاختبار على معيارين

1 - معيار Dixon

يحدد اختبار ديكسون ما إذا كانت القيمة القصوى للعينة هي القيم المتطرفة. يتضمن اختبار Dixon اختيارًا لإحصاءات الاختبار التي تسمح لك بتجاوز التأثيرات المخفية المحتملة للقيم المتطرفة الأخرى للعينة.

وهو على شكلين

أ - الصيغة الاحادية

$$r_{ij} = \frac{y_{i+1} - y_1}{y_{n-j} - y_1}, \; i = 1, 2; \; j = 0, 1, 2$$

ب -الصيغة الثنائية

$$r_{ij} = \max\left\{\frac{y_{i+1} - y_1}{y_{n-j} - y_1}, \frac{y_n - y_{n-i}}{y_n - y_{j+1}}\right\}, \; i = 1, 2; \; j = 0, 1, 2$$

حيث rij قيمة اختبار Dixon i=1,2 ; j=0,1,2   بينما   yi القيمة الاقل في العينة   n عدد مشاهدات العينة

2 - معيار Grubb :

أ - الصيغة الاحادية

إذا قمنا بإجراء هذا الاختبار لمعرفة ما إذا كانت قيمة أصغر البيانات هي القيم المتطرفة ، يتم حساب إحصاء الاختبار G على النحو التالي:

$$G = \frac{\bar{y} - y_1}{s}$$

اما إذا قمنا بإجراء هذا الاختبار لمعرفة ما إذا كانت القيمة الأكبر متطرفة ، فسيتم حساب G على النحو التالي:

$$G = \frac{y_n - \bar{y}}{s}$$

ب – الصيغة الثنائية

$$G = \max\left\{\frac{\bar{y} - y_1}{s}, \frac{y_n - \bar{y}}{s}\right\}$$

حيث

$\bar{y}$ هو متوسط العينة بينما yiالقيمة الادنى في العينة   S   هو الانحراف المعياري واخيرا n عدد مشاهدات العينة

مثال :

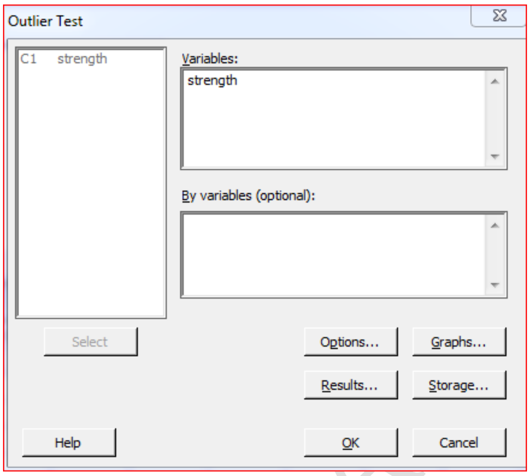
يقوم مهندس الجودة من شركة   لتصنيع المقاومات الكهربائية باختبار مقاومة عينة عشوائية. ويلاحظ القوة المطلوبة لانقطاع خيوط مجموعة من المنتجات التي تم اختيارها عشوائيا . يقوم بإنشاء رسم بياني للبيانات ويلاحظ أن إحدى القيم في العينة تبدو صغيرة بشكل غير طبيعي.

يقوم المهندس باختبار القيم الشاذة لتحديد ما إذا كانت القيمة الأصغر شاذة.

الحل : نقوم بكتابة التعليمة

Stat>Basic Statistics >Outlier Test

فيظهر المربع الحواري التالي



أي ندخل قيم المقاومة التي تحصل عليها المهندس من عينة المشاهدات في مربع Variables : strength
ثم نضغط على Options فنختار المربع الحواري التالي

نلاحظ هنا اننا اخترنا معيار Grubbs بينما تصاغ الفرضية الصفرية والبديلة كمايلي

Null hypothesis All data values come from the same normal population

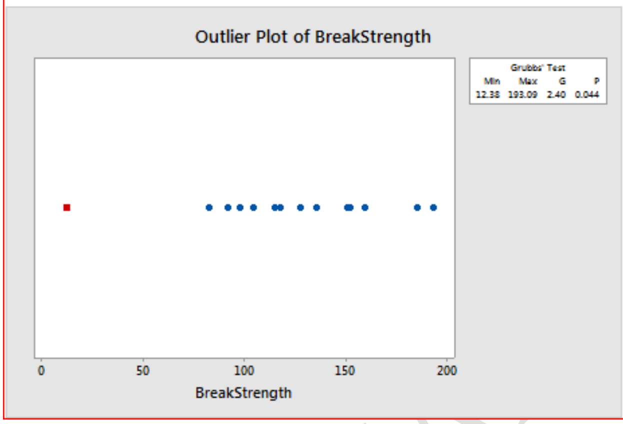
Alternative hypothesis Smallest data value is an outlier

Significance level $\alpha = 0.05$

فالفرضية البديلة تعني ان القيمة الصغرى هي القيمة المتطرفة

ثم نضغط على OK

```
Grubbs' Test

Variable       N   Mean  StDev   Min    Max    G      P
BreakStrength  14  123.4  46.3  12.4  193.1  2.40  0.044

Outlier

Variable       Row  Outlier
BreakStrength   10    12.38
```



متوسط العينة هو 123.4. أن أصغر قيمة بيانات ،هي 12.38 ، تشير إحصائية G والتي تساوي 2,4 وهي اقل من الخطأ المعياري للمتوسط والبالغ 46,3 كما تشير قيمة p إلى أنه إذا كانت جميع المشاهدات من نفس المجتمع الاحصائي الذين يتم توزيعهم بشكل طبيعي ، فإن احتمالية الحصول على قيمة دنيا صغيرة هي 0.044 فقط. لأن القيمة الاحتمالية 0.044 أقل من مستوى الدلالة (المشار إليه بـ α أو alpha) البالغ 0.05 ، يرفض المهندس الفرضية الصفرية ويستنتج أن أصغر قيمة هي أبعد. يتحقق المهندس ويكتشف أن الشخص الذي أدخل البيانات كتب بطريق الخطأ 12.38 بدلاً من 123.8. (خطا في الفاصلة)

6 -حساب التوزيع الاحتمالي لمتوسط عينة من مجتمع طبيعي

لحساب التوزيع الاحتمالي لمتوسطات تتبع التوزيع الطبيعي يوجد شرط أساسي و كاف، وهو معلومية تباين المجتمع الى توفر حجم عينة كاف يحدده الاحصائيون بأكثر من ثلاثين مشاهدة $n>30$

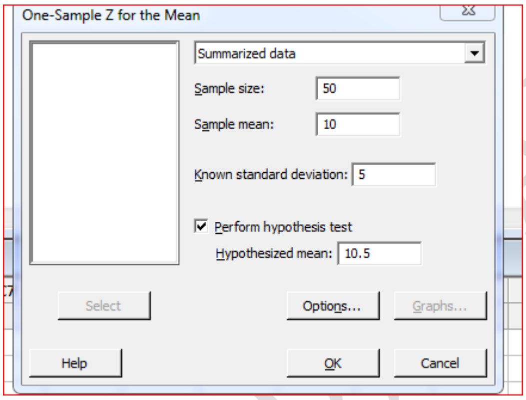
مثال 1 : اذا كانت لدينا عينة حجمها 50 و متوسط المجتمع هو 10 بانحراف معياري للمجتمع 5 المطلوب حساب احتمال متوسط العينة المسحوبة اقل من 10,5

الحل :

ومن اجل حساب احتمال متوسط عينة نقوم بمايلي
كتابة الامر الاتي

Stat>Basic Statistics>One sample Z

ويكون الحل بطريقتين اما عن طريق ورقة العمل worksheet مع ادخال البيانات كالمعتاد او بالطريقة الثانية التي سوف نختارها وهي ادخال المعلومات التي نحتاجها مباشرة دون استخدام ورقة عمل وهذا حسب المربع الحواري الاتي بعد كتابة الامر السابق



في حالة توفر معطيات عن المتوسط والتباين نختار Summarized data وندخل حجم العينة 50 مشاهدة اكبر من 30 ، الى جانب كل من متوسط العينة   10 والانحراف المعياري المعلوم (خاص بالمجتمع) 5

ونضغط بعدها على OK لتظهر نتائج هذه العملية

One-Sample Z

```
Test of μ = 10.5 vs ≠ 10.5
The assumed standard deviation = 5


 N    Mean  SE Mean         95% CI             Z      P
50  10.000    0.707  (8.614, 11.386)  -0.71  0.480
```

ويمكن تفسير النتيجة التالية ان احتمال ان يكون متوسط العينة المسحوبة اقل من 10,5 هو 0,48

تجدر الملاحظة ان هناك حالة اخرى لا يفرها برنامج المينيتاب وهي وجود التوزيع الاحتمالي لمتوسط عينة ذات توزيع طبيعي بحجم كبير $n>30$ مع تباين مجتمع مجهول في هذه الحالة فان التوزيع الاحتمالي لمتوسط العينة هو قريب من التوزيع الطبيعي

1-2 توزيع متوسط العينة لمجتمع مجهول التباين :

كثيرا ما تكون قيمة تباين المجتمع (مربع الانحراف المعياري) غير معلومة ، اضافة الى عدم توفر عدد كافي من مشاهدات العينة المسحوبة $n<30$، وعندئذ يتحول التوزيع الاحتمالي لمتوسط المجتمع من الطبيعي الى توزيع مقارب هو توزيع Student .t

مثال 2 :

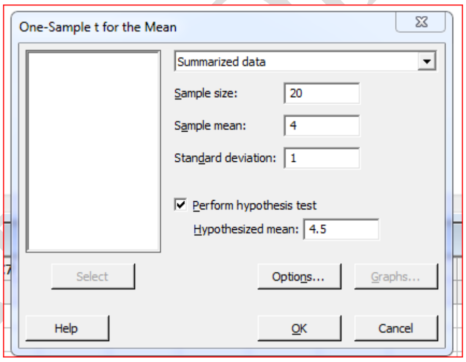
اذا كان حجم العينة في هذه المرة هو 20 ومتوسطها هو 4 وانحراف المعياري للمجتمع الاصلي هو 1

المطلوب : ما هو احتمال ان يكون متوسط العينة المسحوبة اكبر من 4,5

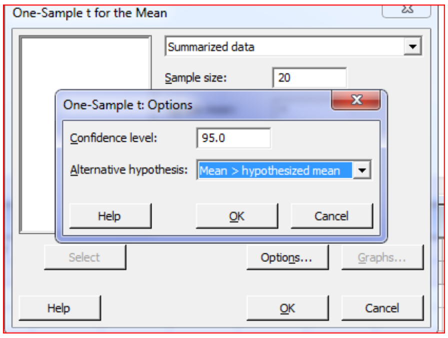
الحل

لتقدير متوسط مجتمع نقوم بما يلي كتابة الامر التالي **Stat>Basic Statistics>One sample t**

فيظهر لدينا المربع الحواري التالي



نلاحظ ان حجم العينة اقل من 30 مشاهدة ، والانحراف المعياري للمجتمع غير معلوم يتم تقديره بواسطة الانحراف المعياري للعينة و بالضغط على ايقونة options ينشا لدينا مربع حواري اخر فوق السابق



والنتيجة تكون في المربع التالي

One-Sample T

```
Test of μ = 4.5 vs > 4.5

 N   Mean  StDev  SE Mean  95% Lower Bound      T      P
20  4.000  1.000    0.224            3.613  -2.24  0.981
```

وهذا يعني ان احتمال ان يكون متوسط العينة المسحوبة والتابع لتوزيع Student هو 0,981

1 - 3 التوزيع الاحتمالي لفرق متوسطي مجتمعين :

كثيرا ما نجد بعض الحالات الخاصة بالمجتمعات المستقلة والمرتبطة من خلال حالة مجتمعين مستقلين ، او حالة مجتمع واحد غير متجانس

1 – 3 – 1 حالة مجتمعين مستقلين (غير مرتبطين)

وهي المجتمعات غير المتجانسة التي نقوم باجراء مقارنة بينها متل الاناث والذكور عمال قطاع الصناعة ونظرائهم في قطاع الزراعة ,,,, ، وكذا تقدير الفرق بين متوسطي مجتمعين من خلال الفرق بين متوسطي عينتين

متال 3

اذا كانت لدينا عينتان مسحوبتان من مجتمعين مستقلين وكان حجمهما على التوالي 35 و 40 وكان متوسطا المجتمعين الاصلين على التوالي 5 و6 في حين كان انحرافاهما المعياريين 3 و2 المطلوب ما هو التوزيع الاحتمالي للفرق بين متوسطي العينتين اكبر او يساوي 1

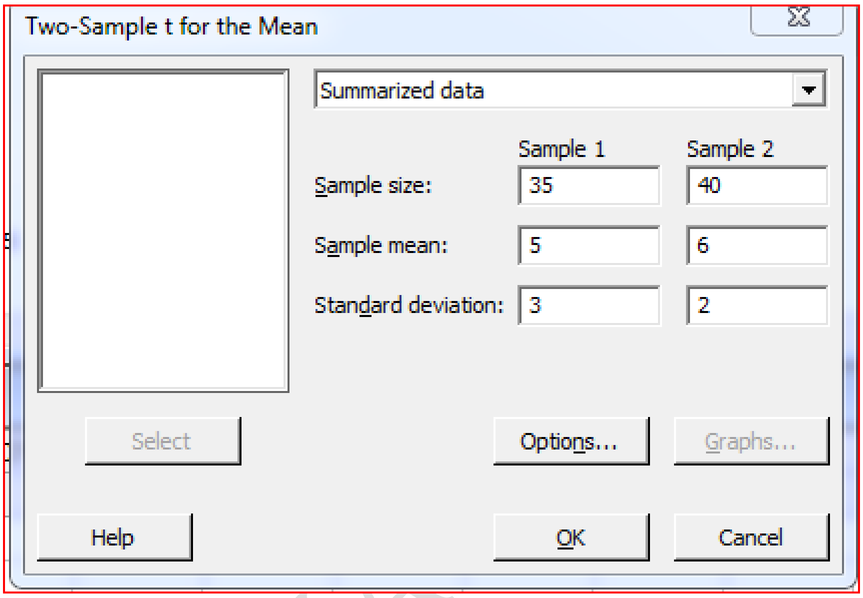
الحل

وبتطبيق هذا التقدير بالاستعانة ببرنامج MINITAB نقوم بكتابة الامر التالي

Stat>Basic Statistics>2- sample t....

لتسهيل الحل نختار Summarized Data

فيظهر المربع الحواري التالي



وبالضغط على نجدOK

Two-Sample T-Test and CI

```
Sample   N  Mean  StDev  SE Mean
1       35  5.00   3.00     0.51
2       40  6.00   2.00     0.32


Difference = μ (1) - μ (2)
Estimate for difference:  -1.000
95% upper bound for difference:  -0.030
T-Test of difference = 0 (vs <): T-Value = -1.72  P-Value = 0.045  DF = 73
```

هذا يعني ان التوزيع الاحتمالي للفرق بين متوسطي العينتين اكبر او يساوي 1 هو 0,045

1- 3- 2 مجتمع بعينتين مرتبطتين :

وهي الحالة الخاصة بمجتمع واحد تسحب منه عينتان غير مرتبطتين مثل الحالة الصحية للمرضى قبل ،وبعد تناول دواء جديد طرح في السوق ، او المستوى الدراسة لمجموعة من الطلبة قبل ، وبعد اجراء اختبار تحسين المستوى

مثال 4

قامت ادارة احدى الكليات بتحديث اساليب التدريس وأرادت معاينة وجود الفرق في درجات الطلبة بين اسلوب التدريس القديم والجديد على مجموعة من الطلبة حسب الجدول التالي

العلامات قبل تطبيق الاسلوب	06	10	04	08	11	14	10	12	17
العلامات بعد تطبيق الاسلوب	10	13	06	07	08	14	17	11	17

المطلوب

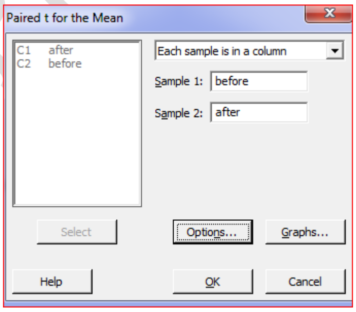
ايجاد التوزيع الاحتمالي للفرق بين متوسطي العينتين

الحل

نقوم بتحويل معطيات الجدول الى ورقة عمل MINITAB كما هو موضح

C1	C2
after	before
10	6
13	10
6	4
7	8
8	11
14	14
17	10
11	12
17	17

فيظهر لدينا مربع حواري نختار منه (each sample is in a colum) اي بيانات كل عينة في عمود خاص



وباتباع نفس اسلوب الامثلة السابقة نحصل على النتيجة التالية

```
11/26/2019 4:53:47 PM

Paired T-Test and CI: after, before

Paired T for after - before

             N   Mean  StDev  SE Mean
after        9  11.44   4.10     1.37
before       9  10.22   3.96     1.32
Difference   9   1.22   3.07     1.02

92% CI for mean difference: (-0.83, 3.28)
T-Test of mean difference = 10 (vs ≠ 10): T-Value = -8.57  P-Value = 0.000
```

اذن **احتمال وجود** فرق بين متوسطي علامات الطلبة قبل وبعد تحديث اسلوب التدريس هو 0

4-1 التوزيع الاحتمالي لعينة النسبة

النسبة عبارة عن توزيع لمجتمع او عينة تخضع لقانون ذي الحدين وتمثل قيمة محصورة بين الصفر والواحد ويمكن حسابها من العلاقة

$\Pi = \frac{\sum Xi}{N}$ للمجتمع

$P = \frac{\sum Xi}{n}$ للعينة

حيث تمثل P تقديرا ل Π نسبة النجاح اما $q = (1-P)$ هي نسبة الفشل ، كما يمكن حساب النسبة بانها عدد الحالات الممكنة على الحالات الكلية

بفرض ان X هو المتغير العشوائي الذي يعبر عن عدد مرات النجاح وكان احتمال النجاح هو P وكان n هو عدد مرات تكرار التجربة فإن دالة الكثافة الاحتمالية للمتغير العشوائي الذي يتبع توزيع ذي الحدين هي

$$P(X) = \binom{n}{x} P^x (1-P)^{n-x}$$

$$x = 0,1,2, \dots, n$$

حيث

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

ومؤشرات توزيع ذي الحدين هي

المتوسط	$n.p$
التباين	$n.p.q$
الانحراف المعياري	$\sqrt{n.p.q}$

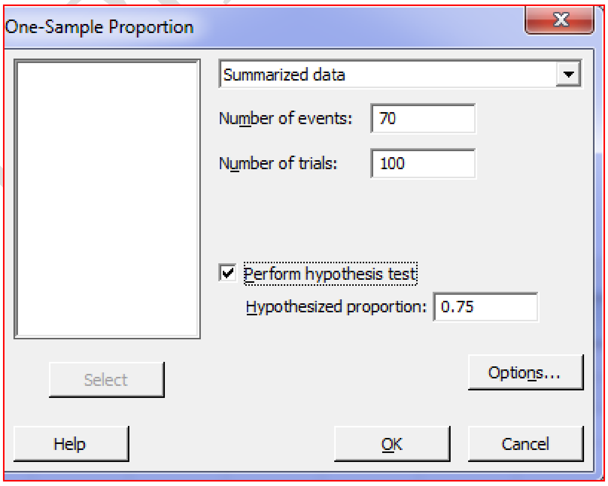
مثال 5 : اذا كان عدد سيارات الوزن الخفيف في احدي الحظائر 70 من مجموع السيارات المتواجدة في الحظيرة والبالغ عددها 100

المطلوب احسب التوزيع الاحتمالي لنسبة السيارات ذات الوزن الخفيف في هذه الحظيرة التى تزيد عن 75

الحل يمكن حل هذا المثال دون اللجوء الى ورقة عمل

Stat>Basic Statistics> 1 Proportion.......

نحصل على المربع الحواري التالي



نضع في ايقونة Number of events عدد السيارات ذات الوزن الخفيف 70، بينما في ايقونة Number of trials نكتب عدد السيارات المتواجدة في الحظيرة ونفعل ايقونة Perform Hypothesis test ونضع القيمة 75 ، ثم نضغط على options ، لنختار مستوى الثقة وفي ايقونة Method

نختار Normal approximation ، والسبب انه في العينات الكبيرة يقرب توزيع ذي الحدين الى التوزيع الطبيعي ، وذلك استنادا الى نظرية النهاية المركزية

ثم نضغط على ok فنحصل على النتيجة

Test and CI for One Proportion

```
Test of p = 0.75 vs p < 0.75


Sample   X    N  Sample p  95% Upper Bound  Z-Value  P-Value
1       70  100  0.700000         0.775377    -1.15    0.124

Using the normal approximation.
```

هذا يعني احتمال ان تكون النسبة اقل من 0,75 هو 0,124

5-1 التوزيع الاحتمالي لفرق نسبتين :

مثال 6 :

قامت احدى شركات التامين بدراسة استطلاعية حول نسبة التامين عندها فجمعت 600 ملف ضد الحريق وجدت انها عوضت 400 منهم كما احصت 700 ملف تامين ضد حوادث المرور عوضت 300 مؤمن منهم

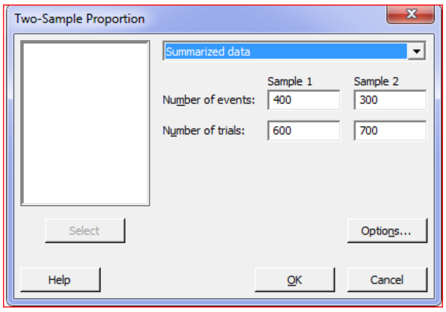
المطلوب حساب التوزيع الاحتمالي لفرق نسبتي عينتين اكبر او يساوي 0,1 اي 10 %

الحل

نقوم بكتابة التعليمة او الامر

Stat>Basic Statistics> 2 Proportions.......

فنكتب بيانات المثال في المربع الحواري كمايلي



وبالضغط على options

نجد ان هناك خياران وهما تقدير العينتين بشكل منفصل estimate the proportions separately او استخدام التقدير المجمع للنسبة use the pooled estimate of the proportion،

```
Test and CI for Two Proportions

Sample    X    N  Sample p
1       400  700  0.571429
        300  600  0.500000


Difference = p (1) - p (2)
Estimate for difference:  0.0714286
95% CI for difference:  (0.0171648, 0.125692)
Test for difference = 0 (vs ≠ 0):  Z = 2.58  P-Value = 0.010

Fisher's exact test: P-Value = 0.010
```

ان التوزيع الاحتمالي لفرق بين نسبتي المؤمنين على الحوادث المرور وضد الحريق ، اكبر من 10 % هو 0,01

تقدير تباين مجتمع طبيعي

اذا اخذت عينات عشوائية كل بحجم n وتباين عينة S^2 من مجتمع طبيعي متوسطه μ في حين ان فترة الثقة لتقدير التباين σ^2 هي

$$\boxed{\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},n-1)}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},n-1)}}}$$

مع العلم ان توزيع الاحتمالي لتباين المجتمع يتبع توزيع χ^2 بدرجة حرية υ =n-1

مثال 8 اذا توفرت لدينا قيم المتغير x في الجدول التالي

X	10	13	6	7	8	14	17	11	17

المطلوب ايجاد التوزيع الاحتمالي لتباين العينة S^2 =17 مع العلم ان تباين المجتمع =15.1 σ^2

الحل تنقل بيانات الجدول في ورقة عمل برنامج minitab

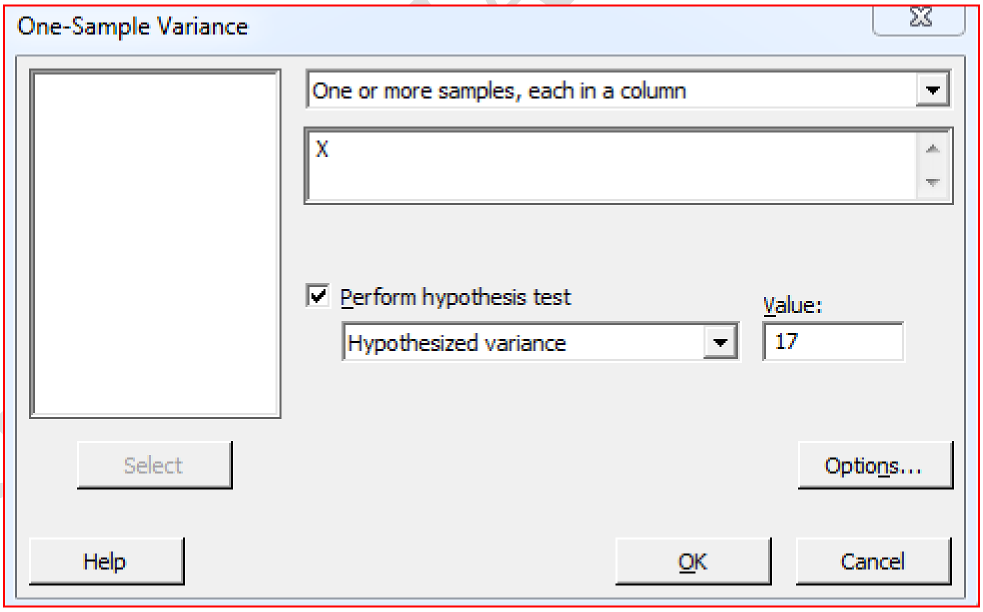
C1
X
10
13
6
7
8
14
17
11
17

ونكتب الامر الاتي

Stat>Basic Statistics>1 variance ....

فيظهر هذا المربع الحواري نختار منه one or more samples , each in **a** colum بمعنى عينة واحدة او اكثر في كل عمود

مع امكانية اختيار التباين ،او الانحراف المعياري حسب المطلوب ، او مدى توفر المعطيات ، مع الاشارة الى تفعيل ايقونة Perform Hypothesis Test



والحصول على النتائج التالية

```
Test and CI for One Variance: X

Method

Null hypothesis          σ-squared = 17
Alternative hypothesis   σ-squared ≠ 17

The chi-square method is only for the normal distribution.
The Bonett method is for any continuous distribution.


Statistics

Variable   N  StDev  Variance
X         12   3.88      15.1


95% Confidence Intervals

                            CI for        CI for
Variable  Method             StDev       Variance
X         Chi-Square  (2.75, 6.60)  (7.6, 43.5)
          Bonett      (2.88, 6.25)  (8.3, 39.1)


Tests

                           Test
Variable  Method      Statistic  DF  P-Value
X         Chi-Square       9.76  11    0.897
          Bonett              –   –    0.778
```


نشير الى ان التوزيع الاحتمالي التابع لقانون كاي تربيع هو 0,897

وصفحة النتائج تشير الى فترتي ثقة لكل من التباين ، والانحراف المعياري

1-8 تقدير النسبة بين تبايني مجتمعين طبيعيين

التوزيع الاحتمالي للنسبة بين تباينين من التوزيعات الهامة التي تبحث في تجانس المجتمعات

ويمكن حساب فترة الثقة لنسبة تباينين بالصيغة الاتية

$$\frac{S_2^2}{S_1^2}\frac{1}{F(\frac{\alpha}{2}, n_2-1, n_1-1)} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2}F(\frac{\alpha}{2}, n_1-1, n_2-1)$$

وهذه النسبة تتبع التوزيع الاحتمالي F Fisher

مثال 9

اذا قمنا بتعديل الجدول السابق في المثال 8 وأضفنا له عمودا اخر هو y

X	10	13	6	7	8	14	17	11	17
Y	6	10	4	8	11	14	10	12	17

و من خلال النتائج يتضح ان تبايني المجتمعين الاصليين هما على التوالي $\sigma_1^2 = 15.1$ $\sigma_2^2 = 13.47$

المطلوب جد احتمال النسبة بين تبايني العينتين S_1^2 / S_2^2 اقل او تساوي 1 علما انها تساوي 1,12 بين تبايني المجتمعين

الحل

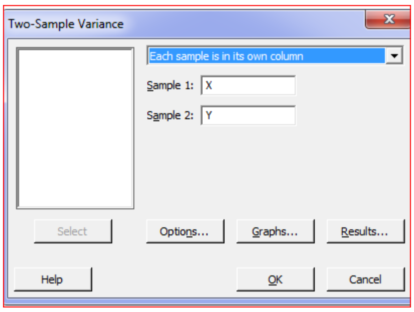
ننشئ اولا ورقة عمل في البرنامج كما يلي

C1	C2
X	Y
10	6
13	10
6	4
7	8
8	11
14	14
17	10
11	12
17	17

نقوم بكتابة الامر

Stat> Basic Statistics >2 Variances....

فيظهر المربع الحواري التالي نختار منه ايقونة Each sample in its own colum بمعنى كل عينة في عمودها الخاص



فنحصل على نتائج التقدير في الصفحة الموالية

```
Null hypothesis         Variance(X) / Variance(y) = 1
Alternative hypothesis  Variance(X) / Variance(y) ≠ 1
Significance level      α = 0.05

F method was used. This method is accurate for normal data only.


Statistics

                                       95% CI for
Variable   N  StDev  Variance          Variances
X         12  3.885    15.091   (7.573, 43.504)
y         12  3.671    13.477   (6.763, 38.852)

Ratio of standard deviations = 1.058
Ratio of variances = 1.120


95% Confidence Intervals

                                CI for
          CI for StDev         Variance
Method        Ratio              Ratio
F        (0.568, 1.972)  (0.322, 3.890)


Tests

                        Test
Method  DF1  DF2   Statistic  P-Value
F        11   11        1.12    0.855
```

وهذا يعني ان احتمال النسبة بين تبايني العينتين اقل من 1 والتي تتبع توزيع F- Fisher هو 0,855

الفصل الثاني : التقدير الاحصائي

مقدمة:

المقصود بالتقدير هو تقدير معالم المجتمع الإحصائي (أو التوزيع الاحتمالي) والتي غالباً ما تكون مجهولة ويكون المطلوب هو الحصول على تقديرات لها -عادة– من بيانات العينة فقد يكون المطلوب تقدير متوسط دخل الدولة، أو تقدير متوسط عمر الناخب... وغيرها.

وهناك نوعان (أو أسلوبان) للتقدير يسمى الأول تقدير النقطة (أو القيمة الواحدة)، ويسمى الثاني تقدير الفترة (أو فترة التقدير أو الثقة).

ففي حالة تقدير النقطة نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة. فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدولة نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة. وكمثال آخر لو أخذنا نسبة الناخبين في العينة الذين يؤيدون مرشحا معيناً كتقدير لهذه النسبة في المجتمع نكون حصلنا على تقدير نقطة للنسبة في مجتمع الناخبين

أما في تقدير الفترة أو فترة التقدير فنحصل على مدى Range أو فترة تتحدد بحدين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن فترة التقدير (أو تقدير الفترة) تحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات.

1-تقدير المتوسطات

نعتمد في تقدير معالم المجتمع الاحصائي على نوعين من التقدير ، وهما

أ - التقدير النقطي ،والذي يمكننا من اختصار معلومات عن المجتمع ،او العينة في معلم واحد كالوسط الحسابي او التباين

ب -التقدير بفترة ثقة ، ونستعين فيه بمجال او مجموعة من القيم التي يأخذها معلم المجتمع او العينة اعتمادا على مستوى المعنوية

1-1 تقدير متوسطات من مجتمع طبيعي

لتقدير متوسطات تتبع التوزيع الطبيعي يوجد شرط أساسي و كاف، وهو معلومية تباين المجتمع الى توفر حجم عينة كاف يحدده الاحصائيون باكثر من ثلاثين مشاهدة $n>30$

مثال 1 : اذا كانت لدينا عينة حجمها 10 و متوسطها هو 10 بانحراف معياري للمجتمع 5

المطلوب تقدير متوسط المجتمع بدرجة ثقة 95%

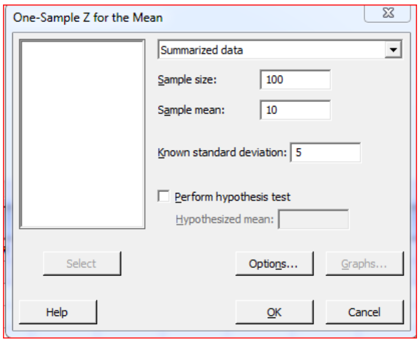
الحل :

ومن اجل تقدير متوسط عينة نقوم بمايلي

كتابة الامر الاتي

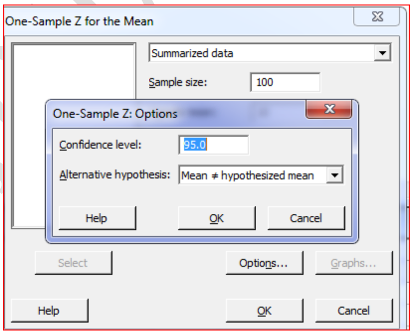
Stat>Basic Statistics>One sample Z

ويكون الحل بطريقتين اما عن طريق ورقة العمل worksheet مع ادخال البيانات كالمعتاد او بالطريقة الثانية التي سوف نختارها وهي ادخال المعلومات التي نحتاجها مباشرة دون استخدام ورقة عمل وهذا حسب المربع الحواري الاتي بعد كتابة الامر السابق



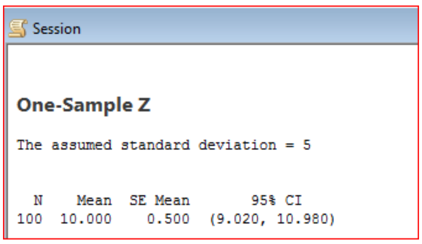
في حالة توفر معطيات عن المتوسط والتباين نختار Summarized data وندخل حجم العينة 100 مشاهدة اكبر من 30 ، الى جانب كل من متوسط العينة 10 والانحراف المعياري المعلوم (خاص بالمجتمع) 5

ونضغط بعدها على Options فيظهر مربع حواري اخر فوق المربع السابق مباشرة



والمستطيل المؤشر عليه باللون الازرق هو مستوى الثقة Confidence level ، وهو درجة الثقة المكونة لهذا المجال وتأخذ قيما قد تصل الى 99%

واشهرها قيم 90 % 95 % 99 % وبعد اختيار 0,95 ثم الضغط على ok نجد



ويمكن تفسير النتيجة التالية ان هناك درجة ثقة بنسبة 95 % بان قيمة متوسط المجتمع سوف تكون محصورة في المجال

C I : (9.02 ; 10.98)

2-1 تقدير متوسط مجتمع مجهول التباين :

كثيرا ما تكون قيمة تباين المجتمع (مربع الانحراف المعياري) غير معلومة ، اضافة الى عدم توفر عدد كافي من مشاهدات العينة المسحوبة $n<30$، وعندئذ يتحول التوزيع الاحتمالي لمتوسط المجتمع من الطبيعي الى توزيع مقارب هو توزيع Student .t

مثال 2 :

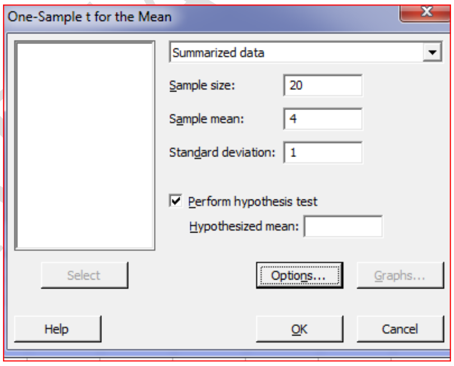
اذا كان حجم العينة في هذه المرة هو 20 ومتوسطها هو 4 وانحرافها المعياري هو 1

المطلوب تقدير متوسط المجتمع بدرجة ثقة 99%

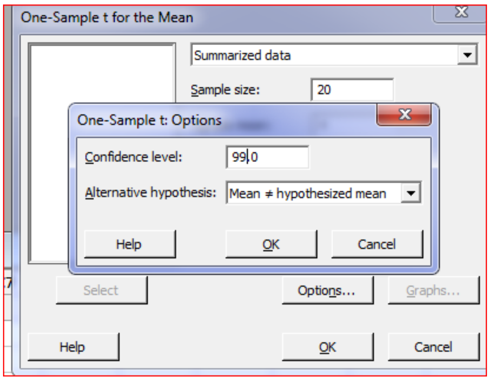
الحل

لتقدير متوسط مجتمع نقوم بما يلي كتابة الامر التالي **Stat>Basic Statistics>One sample t**

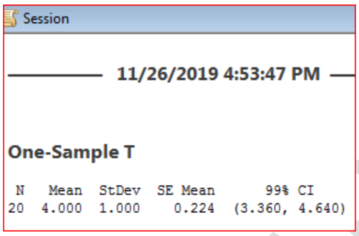
فيظهر لدينا المربع الحواري التالي



نلاحظ ان حجم العينة اقل من 30 مشاهدة ، والانحراف المعياري للمجتمع غير معلوم يتم تقديره بواسطة الانحراف المعياري للعينة و بالضغط على ايقونة options ينشا لدينا مربع حواري اخر فوق السابق



نلاحظ انه في مستطيل Confidence level قد رفعنا مستوى الثقة الى 99% وبهذا فالنتيجة سوف تكون



والنتيجة هي ان هناك مستوى ثقة بنسبة 99% ان مجال القيم المقدرة لمتوسط المجتمع هو CI : (3.36 ; 4.64)

1 - 3 تقدير فرق متوسطي مجتمعين :

كثيرا ما نجد بعض حالات التقدير الخاصة بالمجتمعات المستقلة والمرتبطة من خلال حالة مجتمعين مستقلين ، او حالة مجتمع واحد غير متجانس

1 – 3 – 1 حالة مجتمعين مستقلين (غير مرتبطين)

وهي المجتمعات غير المتجانسة التي نقوم باجراء مقارنة بينها متل الانات والذكور عمال قطاع الصناعة ونظرائهم في قطاع الزراعة ,,,, ، وكذا تقدير الفرق بين متوسطي مجتمعين من خلال الفرق بين متوسطي عينتين

متال 3

اذا كان لدينا عينتين مسحوبتين من مجتمعين مستقلين وكان متوسطا العينتين على التوالي 10 و15 اما حجمهما 25 و40 وكان انحرافاهما المعياريين 3 و5

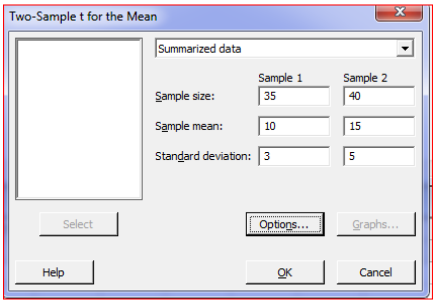
المطلوب تقدير الفرق بين متوسطي المجتمعين بمستوى ثقة 90%

الحل

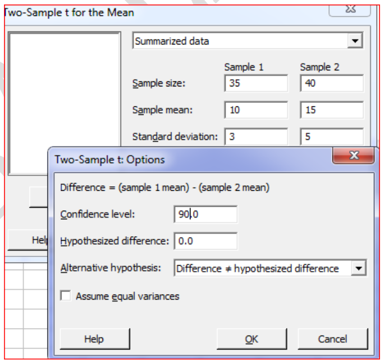
وبتطبيق هذا التقدير بالاستعانة ببرنامج MINITAB نقوم بكتابة الامر التالي

Stat>Basic Statistics>2- sample t....

فيظهر المربع الحواري التالي



وبالضغط على ايقونة options يظهر المربع الحواري الذي يوافق مستوى ثقة 90%



اذا ضغطنا على ok تظهر نتيجة التقدير لفترة الثقة

Two-Sample T-Test and CI

```
Sample   N   Mean  StDev  SE Mean
1       35  10.00   3.00     0.51
2       40  15.00   5.00     0.79


Difference = μ (1) - μ (2)
Estimate for difference:  -5.000
90% CI for difference:  (-6.567, -3.433)
T-Test of difference = 0 (vs ≠): T-Value = -5.32  P-Value = 0.000  DF = 65
```

اي ان مجال الثقة للفرق بين متوسطين هو (-6.567 ; -3.433) CI : بنسبة ثقة 90%

1-3-2 مجتمع بعينتين مرتبطتين :

وهي الحالة الخاصة بمجتمع واحد تسحب منه عينتان غير مرتبطتين مثل الحالة الصحية للمرضى قبل ،وبعد تناول دواء جديد طرح في السوق ، او المستوى الدراسة لمجموعة من الطلبة قبل ، وبعد اجراء اختبار تحسين المستوى

مثال 4

قامت ادارة احدى الكليات بتحديث اساليب التدريس واردات معاينة وجود الفرق في درجات الطلبة بين اسلوب التدريس القديم والجديد على مجموعة من الطلبة حسب الجدول التالي

العلامات قبل تطبيق الاسلوب	06	10	04	08	11	14	10	12	17
العلامات بعد تطبيق الاسلوب	10	13	06	07	08	14	17	11	17

المطلوب

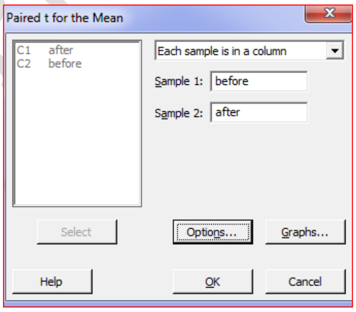
تقدير فترة الثقة للمتوسط بين علامات الطلاب بنسبة 92 %

الحل

نقوم بتحويل معطيات الجدول الى ورقة عمل MINITAB كما هو موضح

C1	C2
after	before
10	6
13	10
6	4
7	8
8	11
14	14
17	10
11	12
17	17

فيظهر لدينا مربع حواري نختار منه (each sample is in a colum) اي بيانات كل عينة في عمود خاص



وباتباع نفس اسلوب الامثلة السابقة نحصل على النتيجة التالية

```
————— 11/26/2019 4:53:47 PM —————

Paired T-Test and CI: after, before

Paired T for after - before

             N   Mean  StDev  SE Mean
after        9  11.44   4.10     1.37
before       9  10.22   3.96     1.32
Difference   9   1.22   3.07     1.02

92% CI for mean difference: (-0.83, 3.28)
T-Test of mean difference = 10 (vs ≠ 10): T-Value = -8.57  P-Value = 0.000
```

اذن فالفرق بين متوسطي علامات الطلبة قبل وبعد تحديث اسلوب التدريس يقع بين القيمتين (3,28 و -0,83)، وذلك بدرجة ثقة 92%

4-1 تقدير النسبة

النسبة عبارة عن توزيع لمجتمع او عينة تخضع لقانون ذي الحدين وتمثل قيمة محصورة بين الصفر والواحد ويمكن حسابها من العلاقة

$\Pi = \frac{\sum Xi}{N}$ للمجتمع

$P = \frac{\sum Xi}{n}$ للعينة

حيث تمثل P تقديرا ل Π نسبة النجاح اما $q = (1-P)$ هي نسبة الفشل ، كما يمكن حساب النسبة بانها عدد الحالات الممكنة على الحالات الكلية

بفرض ان X هو المتغير العشوائي الذي يعبر عن عدد مرات النجاح وكان احتمال النجاح هو P وكان n هو عدد مرات تكرار التجربة فإن دالة الكثافة الاحتمالية للمتغير العشوائي الذي يتبع توزيع ذي الحدين هي

$$P(X) = \binom{n}{x} P^{x}(1-P)^{n-x}$$

$$x = 0,1,2, \dots, n$$

حيث

$$\binom{n}{x} = \frac{n!}{x!\,(n-x)!}$$

ومؤشرات توزيع ذي الحدين هي

المتوسط	$n.p$
التباين	$n.p.q$
الانحراف المعياري	$\sqrt{n.p.q}$

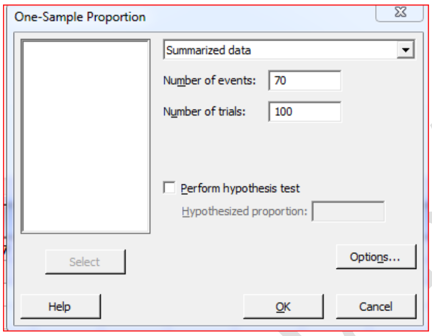
مثال 5 : اذا كان عدد سيارات الوزن الخفيف في احدي الحظائر 70 من مجموع السيارات المتواجدة في الحظيرة والبالغ عددها 100

المطلوب تقدير نسبة السيارات ذات الوزن الخفيف في هذه الحظيرة بدرجة ثقة 95 %

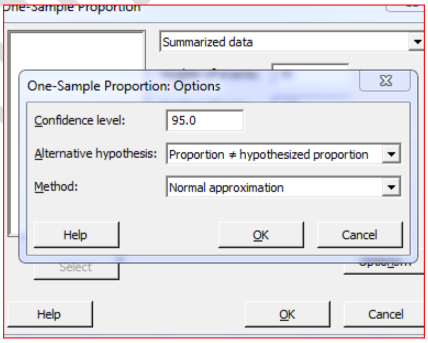
الحل يمكن حل هذا المثال دون اللجوء الى ورقة عمل

Stat>Basic Statistics> 1 Proportion.......

نحصل على المربع الحواري التالي



نضع في ايقونة Number of events عدد السيارات ذات الوزن الخفيف 70، بينما في ايقونة Number of trials نكتب عدد السيارات المتواجدة في الحظيرة ، ثم نضغط على options ، لنختار مستوى الثقة وفي ايقونة Method نختار Normal approximation ، والسبب انه في العينات الكبيرة يقرب توزيع ذي الحدين الى التوزيع الطبيعي ، وذلك استنادا الى نظرية النهاية المركزية



ثم نضغط على ok فنحصل على النتيجة

```
Welcome to Minitab, press F1 for help.

Test and CI for One Proportion

Sample   X    N  Sample p          95% CI
1       70  100  0.700000  (0.610183, 0.789817)

Using the normal approximation.
```

95% CI of P : (0.61 ; 0.78)

5-1 تقدير فرق نسبتين :

مثال 6 :

قامت احدى شركات التامين بدراسة استطلاعية حول نسبة التامين عندها فجمعت 600 ملف ضد الحريق وجدت انها عوضت 400 منهم كما احصت 700 ملف تامين ضد حوادث المرور عوضت 300 مؤمن منهم

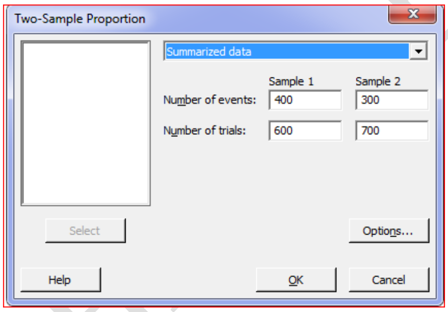
المطلوب تقدير الفرق بين النسبتين بدرجة ثقة 99%

الحل

نقوم بكتابة التعليمة او الامر

Stat>Basic Statistics> 2 Proportions…….

فنكتب بيانات المثال في المربع الحواري كمايلي



وبالضغط على options

نجد ان هناك خياران وهما تقدير العينتين بشكل منفصل estimate the proportions separately

او استخدام التقدير المجمع للنسبة use the pooled estimate of the proportion، الى جانب درجة الثقة 99%

```
———— 11/27/2019 3:39:

Test and CI for Two Proportions

Sample    X    N  Sample p
1       400  600  0.666667
2       300  700  0.428571


Difference = p (1) - p (2)
Estimate for difference:  0.238095
99% CI for difference:  (0.168968, 0.307223)
Test for difference = 0 (vs ≠ 0):  Z = 8.87  P-Value = 0.000

Fisher's exact test: P-Value = 0.000
```

ان الفرق بين نسبتي المؤمنين على الحوادث المرور وضد الحريق تقع بين 16,8% و 30,7 % ، وذلك بدرجة ثقة 99%

6-1 تقدير متوسط بواسوني

يعتبر قانون بواسون حالة خاصة لقانون ذي الحدين ، ويسمى قانون الحوادث النادرة أي عند تؤول n الى مالا نهاية ويكون حجم P صغيرا جدا

حوادث الطيران من الامثلة البارزة على توزيع بواسون وقانون Poisson نسبةعالم الرياضيات Siméon Denis Poisson) هو قانون احتمالي يستخدم لقياس عدد الأحداث التي تحدث في فترة زمنية معينة ، عندما تكون هذه الأحداث نادرة ومستقلة إلى حد ما. ويعطى قانون بواسون بالصيغة الاتية

$$P(X = k) = \frac{\lambda^{\kappa} e^{-\lambda}}{k!}$$

اين عبارة عن معلمة و X هو عدد الحوادث

ومعالم توزيع بواسون هي

المتوسط	λ
التباين	λ
الانحراف المعياري	$\sqrt{\lambda}$

مثال 7

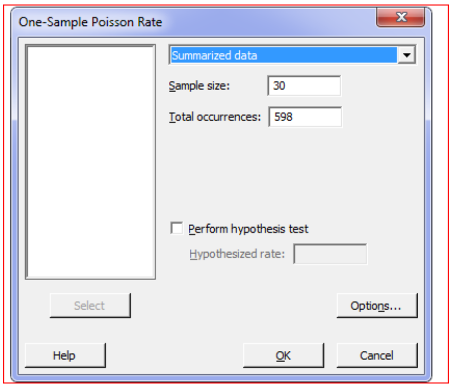
يريد مسؤول مراقبة الجودة في خدمة النقل الحضري تحسين رضا العملاء. لتقييم مستوى رضا العملاء ، يحسب المدير عدد المطالبات المستلمة لمدة 30 يومًا والتي بلغت 598

المطلوب تقدير متوسط هذه المطالبات خلال مدة شهر وبدرجة ثقة 95%

الحل

للقيام بحل هذا المثال نكتب الامر التالي

Stat>Basic Statistics> 1-Sample Poisson Rate………



وفي ايقونة options نختار درجة ثقة 95% الى جانب طريقة المقاربة الى التوزيع الطبيعي Normal approximation

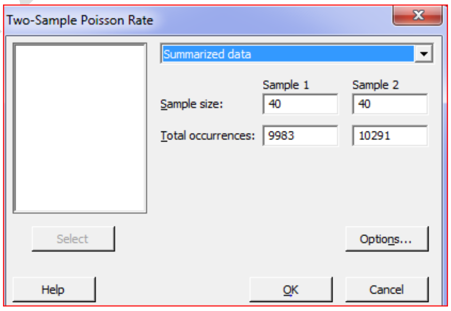
Confidence Interval for One-Sample Poisson Rate

```
         Total          Rate of
Sample  Occurrences   N  Occurrence       95% CI
     1          598  30     19.9333  (18.3357, 21.5310)

"Length" of observation = 1.
```

7-1 تقدير الفرق بين متوسطين بواسونيين

يريد محلل الخدمة البريدية مقارنة عدد زيارات العملاء في مكتبين للبريد. يحسب المحلل عدد العملاء الذين يدخلون كل مكتب لمدة 40 يوم عمل. حيت كانت عدد زيارات المكتب الاول 9983 في حين كانت عدد زيارات المكتب الثاني 10291

يقوم المحلل بإجراء تقدير فرق لمتوسطي بواسون لزيارات العملاء بين مكتبي البريد ،.بثقة 95%



```
Test and CI for Two-Sample Poisson Rates

               Total         Rate of
Sample  Occurrences   N  Occurrence
1              9983  40     249.575
2             10291  40     257.275


Difference = rate(1) - rate(2)
Estimate for difference: -7.7
95% CI for difference: (-14.6768, -0.723175)
Test for difference = 0 (vs ≠ 0): Z = -2.16 P-Value = 0.031

Exact Test: P-Value = 0.031
```


1 - 8 تقدير تباين مجتمع طبيعي

اذا اخذت عينات عشوائية كل بحجم n وتباين عينة S^2 من مجتمع طبيعي متوسطه μ في حين ان فترة الثقة لتقدير التباين σ^2 هي

$$\boxed{\frac{(n-1)S^2}{\chi^2_{(\frac{\alpha}{2},n-1)}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{(1-\frac{\alpha}{2},n-1)}}}$$

مع العلم ان توزيع الاحتمالي لتباين المجتمع يتبع توزيع χ^2 بدرجة حرية $\upsilon = n-1$

مثال 8 اذا توفرت لدينا قيم المتغير x في الجدول التالي

X	10	13	6	7	8	14	17	11	17

المطلوب تقدير تباين المجتمع بدرجة ثقة 95%

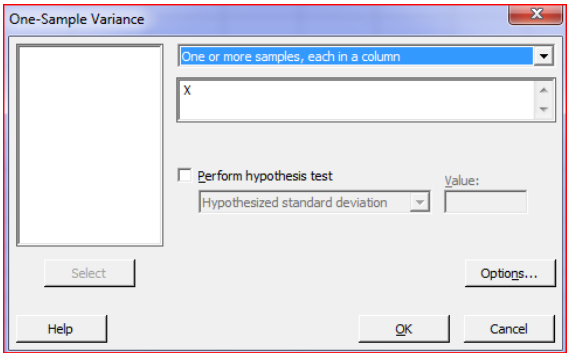
الحل تنقل بيانات الجدول في ورقة عمل برنامج minitab

C1
x
10
13
6
7
8
14
17
11
17

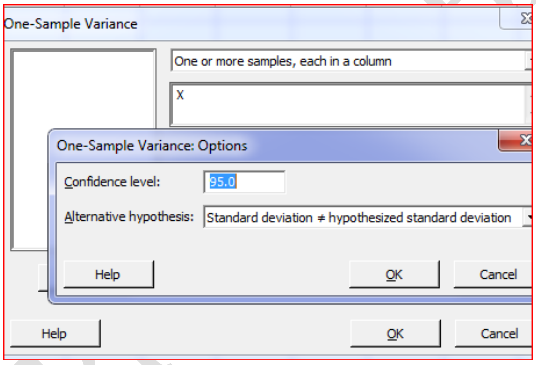
ونكتب الامر الاتي

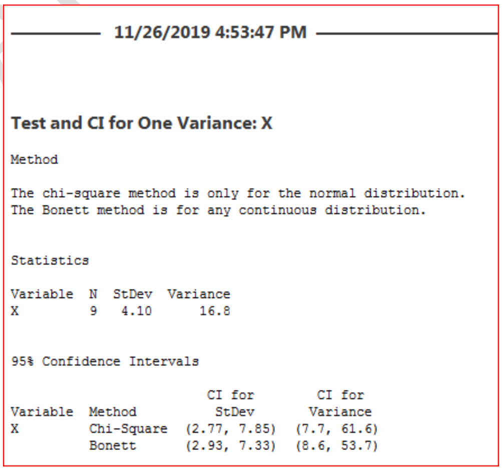
Stat>Basic Statistics>1 variance ....

فيظهر هذا المربع الحواري نختار منه one or more samples , each in a colum بمعنى عينة واحدة او اكثر في كل عمود



مع امكانية اختيار التباين ،او الانحراف المعياري حسب المطلوب ، او مدى توفر المعطيات ، مع الاشارة الى درجة الثقة المطلوبة





—— 11/26/2019 4:53:47 PM ——

Test and CI for One Variance: X

```
Method

The chi-square method is only for the normal distribution.
The Bonett method is for any continuous distribution.


Statistics

Variable  N  StDev  Variance
X         9   4.10      16.8


95% Confidence Intervals

                         CI for       CI for
Variable  Method         StDev        Variance
X         Chi-Square  (2.77, 7.85)  (7.7, 61.6)
          Bonett      (2.93, 7.33)  (8.6, 53.7)
```

وصفحة النتائج تشير الى فترتي ثقة لكل من التباين ، والانحراف المعياري

8-1 تقدير النسبة بين تبايني مجتمعين طبيعيين

التوزيع الاحتمالي للنسبة بين تباينين من التوزيعات الهامة التي تبحث في تجانس المجتمعات

ويمكن حساب فترة الثقة لنسبة تباينين بالصيغة الاتية

$$\frac{S_2^2}{S_1^2}\frac{1}{F(\frac{\alpha}{2},n_2-1,n_1-1)} < \frac{\sigma_2^2}{\sigma_1^2} < \frac{S_2^2}{S_1^2}F(\frac{\alpha}{2},n_1-1,n_2-1)$$

وهذه النسبة تتبع التوزيع الاحتمالي F Fisher

مثال 9

اذا قمنا بتعديل الجدول السابق في المثال 8 وأضفنا له عمودا اخر هو y

X	10	13	6	7	8	14	17	11	17
Y	6	10	4	8	11	14	10	12	17

المطلوب تقدير النسبة بين التباينين بدرجة ثقة

الحل

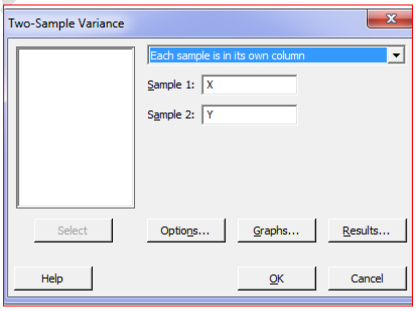
ننشئ اولا ورقة عمل في البرنامج كما يلي

C1	C2
X	Y
10	6
13	10
6	4
7	8
8	11
14	14
17	10
11	12
17	17

نقوم بكتابة الامر

Stat> Basic Statistics >2 Variances....

فيظهر المربع الحواري التالي نختار منه ايقونة Each sample in its own colum بمعنى كل عينة في عمودها الخاص



فنحصل على نتائج التقدير في الصفحة الموالية

```
Test and CI for Two Variances: X, Y

95% Confidence Intervals

                                CI for
              CI for StDev     Variance
Method           Ratio           Ratio
Bonett   (0.547, 2.331)   (0.300,  5.432)
Levene   (0.538, 3.188)   (0.289, 10.162)
```

نلاحظ ان هناك فترتي ثقة حسب معياري كل من Bonett و Levene ،وكلاهما يعتمدان على توزيع Fisher

1 - 9 تقدير وسيط مجتمع :

الوسيط هو احد مقاييس النزعة المركزية ، ويأتي الثاني في الاهمية بعد المتوسط الحسابي و يستخدم في البيانات المفتوحة او ذات القيم الشاذة والمتباعدة ، كما ان الوسيط توزيعه غير معلمي

ويمكن تقدير فترة ثقة لوسيط مجتمع بطريقتين او بالاعتماد على اختبارين وهما

1-9-1 Sign Test :

متال 10 : قامت احدى الشركات بإجراء سبر اراء عن مدى تقبل المستهلكين لمنتج كانت قد طرحته في السوق حديتا ، وبينت نتائج الاستبيان اعتمادا على سلم ليكارت الخماسي

5	موافق تماما	4	موافق	3	محايد	2	غير موافق	1	غير موافق تماما

وقد ظهرت نتائج الاستبيان كما يلي

3	2	1	1	5	4	3	1	5	4

المطلوب تقدير الوسيط بدرجة ثقة 93%

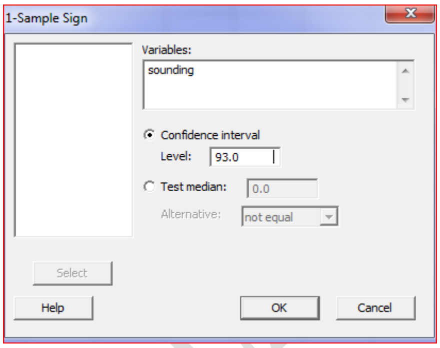
الحل نقوم اولا بنقل جدول نتائج الاستبيان الى ورقة العمل ونسمي المتغير sounding

↓	C1	C2
	sounding	
1	3	
2	2	
3	1	
4	1	
5	5	
6	4	
7	3	
8	1	
9	5	
10	4	

ثم تكتب الامر الاتي

Stat > Nonparametrics > 1-Sample Sign

لنحصل على المربع الحواري



وقد وضعنا فيه المتغير المدروس ومستوى الثقة اللازم للتقدير ، والنتيجة هي

```
Welcome to Minitab, press F1 for help.

Sign CI: sounding

Sign confidence interval for median

                                    Confidence
                        Achieved     Interval
           N  Median  Confidence  Lower  Upper  Position
sounding  10   3.000      0.8906  1.000  4.000         3
                          0.9300  1.000  4.169       NLI
                          0.9785  1.000  5.000         2
```

عند درجة ثقة اقل من المطلوبة اي 89,06% فان وسيط الاستبيان يقع بين القمتين 1 و4

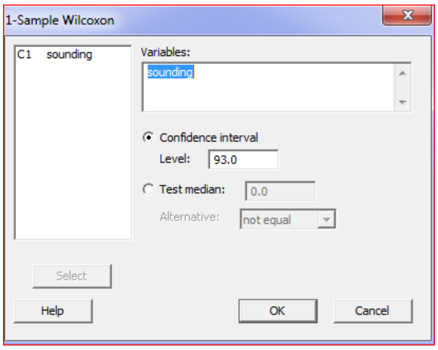
عند درجة الثقة المطلوبة اي 93% فان الوسيط يقع بين القيمتين 1 و 4,16

عند درجة الثقة اكبر من المطلوبة فان الوسيط يقع بين القيمتين 1 و 5

2-9-1 Wilcoxon test :

سنقوم بإجراء نفس معطيات المثال السابق ، وذلك بتطبيق التعليمة

Stat > Nonparametrics > 1-Sample Wilcoxon



Wilcoxon Signed Rank CI: sounding

```
                                           Confidence
                Estimated    Achieved       Interval
            N      Median  Confidence    Lower  Upper
sounding   10        3.00        93.3     2.00   4.00
```

اذن فوسيط الاستبيان (وسيط المجتمع) يقع بين القيمتين 2 و 4 بمستوى ثقة 93,3%

1 - 10 تقدير فرق وسيطين :

يمكن تقدير فيق وسيطين من خلال برنامج Minitab و ذلك بالاعتماد على اختبار غير معلمي هو Mann-Whitney

متال

قدمت للسوق علامتان تجاريتان و كانت المقارنة بينها من خلال وسيط عدد اشهر التعمير

وسيط العلامة التجارية 1	81	75	77	78
وسيط العلامة التجارية 2	77	74	70	71

نقدر فرق وسيطين بدرجة تقة 95%

الحل

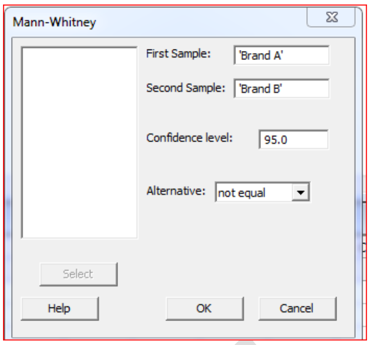
نقوم بنقل الجدول من برنامج الاكسل الى ورقة عمل Minitab

Worksheet 1 ***

↓	C1	C2	C3
	Brand A	Brand B	
1	81	77	
2	75	74	
3	77	70	
4	78	71	

ونكتب التعليمة التالية

Stat > Nonparametrics > Mann-Whitney. ثم ندون المعلومات في المربع الحواري



Method

η₁: median of Brand A

η₂: median of Brand B

Difference: η₁ - η₂

Descriptive Statistics

Sample	N	Median
Brand A	11	36.0
Brand B	10	37.6

Estimation for Difference

Difference	CI for Difference	Achieved Confidence
-1.85	(-3.0, -0.9)	95.52%

تشير نسبة 95.5 بالمائة CI إلى أن فرق وسيطي المجتمعين يقع ضمن المجال (0,9 - 3-)

الفصل الثالث اختبار الفرضيات : Hypothesis test

اختبار الفرضيات هو القاعدة التي تحدد ما إذا كان يجب قبول أو رفض مطالبة حول مجتمع ما بناءً على الأدلة المقدمة من عينة من البيانات.

يفحص اختبار الفرضية فرضيتين متعارضتين حول مجتمع ما: فالفرضية . الصفرية هي البيان الذي يتم اختباره. عادةً ما تكون الفرضية الصفرية عبارة عن "لا تأثير" أو "لا فرق". الفرضية البديلة هي أن العبارة التي نريد أن نكون قادرٍين على استنتاجها صحيحة ، استنادًا إلى الأدلة المقدمة من نموذج البيانات. بناءً على نموذج البيانات ، يحدد الاختبار ما إذا كان سيتم رفض فرضية صفرية . يمكننا استخدام قيمة فيشر ، لاتخاذ قرار. مإذا كانت قيمة p أقل من مستوى االمعنوية (يُشار إليها بـ α أو alpha) ، وعندئذ نرفض الفرضية الصفرية .

تعريف : الفرضية: Hypothesis

هي ادعاء حول صحة شيء ما. وتنقسم إلى فرضية مبدئية (فرضية العدم H_O) والفرضية البديلة H_a .

الفرضية الصفرية (المبدئية) H_O (Null Hypothesis) :

هي الفرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، وخلاف ذلك نقبلها وتعني كلمة Nul انه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).

الفرضية البديلة (H_a) Alternative Hypothesis :

هي الفرضية التي يضعها الباحث كبديل عن فرضية العدم و نقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

أنواع اختبارات الفروض:

عندما نقبل الفرضية المبدئية فإننا نقبلها بنسبة دقة 90% أو 95% أو 99% أو غير ذلك وتسمى مستويات الثقة Significance Levels أي يوجد نسبة خطأ معين في قبولنا للفرضية المبدئية بمعنى أننا نقبل صحة الفرضية المبدئية وهي خاطئة وهذا الخطأ هو الخطأ α ويسمى مستوى المعنوية، أي إذا كان مستوى الثقة 95% (α -1) فان مستوى المعنوية α تساوي 5% وهي عبارة عن مساحة منطقة تحت منحنى التوزيع تمثل منطقة الرفض وتكون أما على صورة ذيل واحد جهة اليمين أو اليسار أو ذيلين متساويين في المساحة واحد جهة اليمين والثاني جهة اليسار.

تعريف اختبار الفروض في جانب واحد:

هو الاختبار الذي تبين فيه الفروض البديلة أن المعلمة للمجتمع اكبر أو اصغر من إحصائية العينة ، فهناك تحديد للاتجاه.

تعريف اختبار الفروض في جانبين(ذيلين):

هو الاختبار الذي لا تبين فيه الفرضية البديلة أن معلمة المجتمع أكبر أو أصغر من إحصائية العينة، بل مجرد أنها تختلف .

اختبارات العينة الواحدة :

الاختبارات المعلمية للعينة الواحدة

اختبار فرضيات متوسط حسابي لمجتمع

في حالة معلومية تباين المجتمع :

يمكن استخدام اختبار فرضيات متعلقة بمتوسط مجتمع ومقارنته بالقيمة المستهدفة أو قيمة مرجعية عندما تعرف الانحراف المعياري او التباين للمجتمع . باستخدام هذا الاختبار ، يمكنك القيام بما يلي:

نحدد ما إذا كان متوسط المجتمع يختلف عن المتوسط المفترض الذي تحدده.

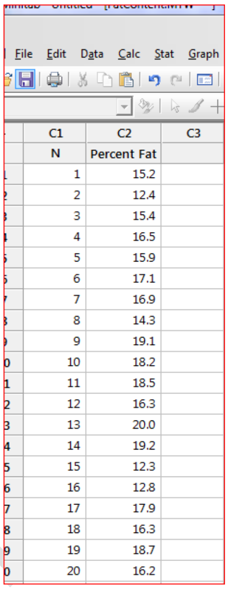
احسب مجموعة من القيم التي من المحتمل أن تشمل متوسط المجتمع .

مثال :

تريد شركة للصناعات الغذائية التقليل من النسبة المئوية للدهون في منتجاتها ، و الوصول الى نسبة مستهدفة معلن عنها هي 15% ودلك بعد سحب عينة من 20 منتج معطاة الجدول بانحراف معياري للمنتجات ككل (للمجتمع) 2,6%

الحل

باستخدام النسب المئوية للدهون التي تحتوي عليها المواد المنتجة وإدخالها ضمن ورقة minitab نجد



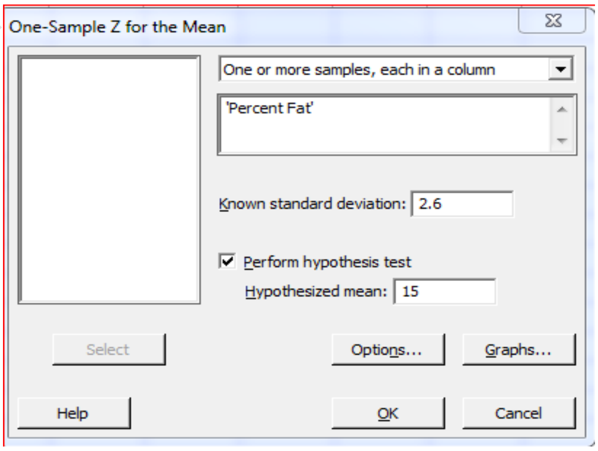
C1 N	C2 Percent Fat	C3
1	15.2	
2	12.4	
3	15.4	
4	16.5	
5	15.9	
6	17.1	
7	16.9	
8	14.3	
9	19.1	
10	18.2	
11	18.5	
12	16.3	
13	20.0	
14	19.2	
15	12.3	
16	12.8	
17	17.9	
18	16.3	
19	18.7	
20	16.2	

باستخدام الامر التالي

Stat >Basic Statistics>1-sample Z

نختار one or more samples, each in acolumn

ثم يكتب اسم المتغير تلقائيا وهو Percent Fat في المربع الحواري التالي



وبما اننا نعلم التباين او الانحراف المعياري للمجتمع نضع قيمته في مربع Know standard deviation (2,6)

مع تفعيل مربع perform hypothesis test ثم وضع قيمة المتوسط (15) ونضغط على ok

Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
20	16.4600	2.2582	0.5814	(15.3205, 17.5995)

μ: mean of Percent Fat

Known standard deviation = 2.6

Test

Null hypothesis	H_0: μ = 15
Alternative hypothesis	H_1: μ ≠ 15

Z-Value	P-Value
2.51	0.0120

. بما ان القيمة p هي 0.012 ، وهو أقل من مستوى المعنوية 0.05 ، اذن نرفض الفرضية الصفرية. تشير النتائج إلى أن النسبة المئوية للدهون تختلف عن 15 ٪ لان القيمة المحسوبة Z-value =2.51 اكبر من القيمة الحرجة او الجدولية Z-tab =1.96 وبالتالي القرار يكون في منطقة الرفض H1

في حالة مجهولية تباين المجتمع وكان حجم العينة صغيرا n<30

نستخدم اختبار فرضية متوسط المجتمع ولمقارنته بالقيمة المستهدفة أو القيمة المرجعية عندما لا تعرف الانحراف المعياري للمجتمع . ونقوم عندئذ بتقدير تباين او انحراف معياري للمجتمع بواسطة تباين العينة او الانحراف المعياري لها S مستخدمين توزيع Student بدلا من التوزيع الطبيعي لذا كان حجم العينة صغيرا كما حدده الاحصائيون ب n<30 مشاهدة

متال

يريد خبير اقتصادي تحديد ما إذا كانت تكلفة استهلاك الطاقة المنزلية للعائلات قد تغيرت عن العام السابق ، عندما كان متوسط التكلفة في الشهر 200 دولار. يقوم الاقتصادي باختبار 25 عائلة بشكل عشوائي ويسجل تكاليف الطاقة الخاصة بهم للعام الحالي. يقوم الخبير الاقتصادي بإجراء اختبار لعينة واحدة لتحديد ما إذا كانت تكلفة الطاقة الشهرية تختلف عن 200 دولار.

Family ID	Energy Cost
1	211
2	572
3	558
4	250
5	478
6	307
7	184
8	435
9	460
10	308
11	188
12	111
13	676
14	326
15	142
16	255
17	205
18	77
19	190
20	320
21	407
22	333
23	488

نقوم بكتابة الامر التالي stat > basic statiscs>1-sample t

Descriptive Statistics

N	Mean	StDev	SE Mean	95% CI for μ
25	330.56	154.18	30.84	(266.92, 394.20)

μ: mean of Energy Cost

Test

Null hypothesis $H_0: \mu = 200$
Alternative hypothesis $H_1: \mu \neq 200$

T-Value	P-Value
4.23	0.0003

تنص الفرضية الصفرية على أن متوسط تكاليف الطاقة للعام الحالي هو 200 دولار. لأن القيمة p هي 0.0003 ، وهي أقل من مستوى المعنوية البالغ 0.05 ، يرفض الاقتصادي الفرضية المعنوية ويخلص إلى أن متوسط تكلفة الطاقة المنزلية للعائلات يختلف عن 200 دولار. يشير CI 95٪ إلى أن متوسط المجتمع من المرجح أن يكون أكبر من 200 دولار

اختبار فرضيات تباين او انحراف معياري لمجتمع :

كما راينا في الفصول السابقة فان تقدير تباين σ^2 مجتمع طبيعي يخضع لتوزيع كاي تربيع غير انه تجدر الاشارة الى ان

طريقة Bonett صالحة لأي توزيع مستمر.

طريقة chi-square صالحة فقط للتوزيع الطبيعي

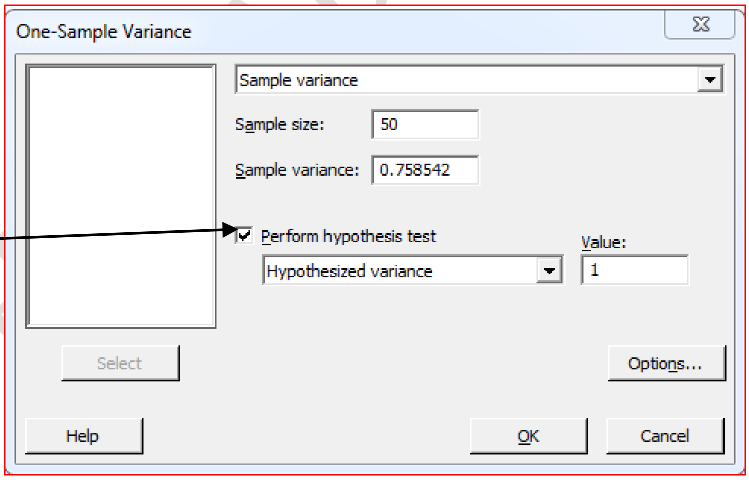
مثال

اذا كان حجم العينة n=50 بتباين $\sigma^2=0.758542$ نريد اختبار فرضية ان $\sigma^2= 1$

الحل :

نقوم بكتابة الامر الاتي Stat >Basic Statistics >1 Variance ونختار sample variance لان معطيات المثال متعلقة بالتباين وليس الانحراف المعياري ونقوم بتفعيل ايقونة sample variance بعد ان كتابة قيمة حجم العينة وتباين العينة

نفعل دائما هذه الايقونة كي نحصل على اختبار الفرضيات



Descriptive Statistics

N	StDev	Variance	95% CI for σ Bonett	95% CI for σ Chi-Square
50	0.870943	0.758542	(0.70409, 1.12130)	(0.72753, 1.08531)

Test

Null hypothesis $H_0: \sigma = 1$
Alternative hypothesis $H_1: \sigma \neq 1$

Method	Test Statistic	DF	P-Value
Bonett			0.2755
Chi-Square	37.17	49	0.2154

نظرًا لأنه لا يبدو أن البيانات تأتي من توزيع طبيعي ، فإننا نستخدم فترة الثقة لطريقة Bonett. تظهر فترة الثقة 95٪ أن النطاق المحتمل لانحراف المعيار للمجتمع هو حوالي 0.704 و 1.121. لأن القيمة p=0.2755 أكبر من 0.05 ، لا يمكن أن نستنتج ان تباين المجتمع يختلف عن 1.

ملاحظة :

نستخدم طريقة chi-square فقط إذا كنا متأكدين من أن البيانات تتبع توزيع طبيعي ، لأن أي انحراف صغير عن الحالة الطبيعية يمكن أن يؤثر بشكل كبير على نتائج طريقة chi-square.

اختبار فرضيات النسبة :

استخدام اختبار فرضيات متعلقة بنسبة تخضع لتوزيع ذي الحدين ولمقارنة النسبة بالقيمة المستهدفة أو القيمة المرجعية. باستخدام هذا الاختبار ، يمكننا القيام بما يلي:

تحديد ما إذا كانت نسبة المجتمع تختلف عن النسبة المفترضة التي تحددها.

حساب مجموعة من القيم التي من المحتمل أن تشمل نسبة المجتمع

مثال:

يريد محلل تسويق تحديد ما إذا كانت الإعلانات التي يتم إرسالها بالبريد لمنتج جديد تؤدي إلى معدل استجابة يختلف عن المتوسط الوطني. يتم اختيار عينة عشوائية من 1000 أسرة لتلقي الإعلانات. من بين 1000 أسرة تم أخذ عينات منها ، أجرت 87 عملية شراء بعد تلقي الإعلان.

يقوم المحلل بإجراء اختبار نسبة واحد لتحديد ما إذا كانت نسبة الأسر التي أجرت عملية شراء مختلفة عن المعدل الوطني البالغ 6.5٪.

الحل :

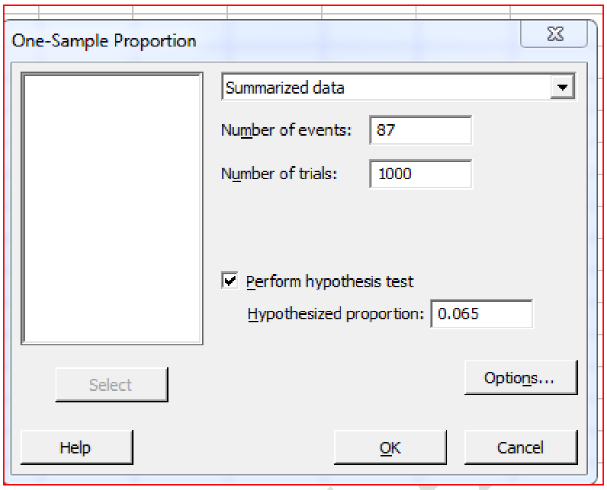
نقوم بكتابة الامر التالي

Stat> Basic Statistics > 1 Proportion....

وتماشيا مع معطيات المثال نختار Summarized data

Number of events : 87 Number of trials: 1000

وحتى نميز اختبار الفرضيات عن التقدير نقوم بتفعيل ايقونة perform hypothesis test ثم نضع قيمة النسبة المفترضة 6,5% والتي تعادل 0,065 كما هو موضح في المربع الحواري



وبالضغط على ok تظهر النتائج كمايلي

Method

p: event proportion
Exact method is used for this analysis.

Descriptive Statistics

N	Event	Sample p	95% CI for p
1000	87	0.087000	(0.070268, 0.106208)

Test

Null hypothesis $H_0: p = 0.065$
Alternative hypothesis $H_1: p \neq 0.065$

P-Value
0.0085

تنص الفرضية الصفرية على أن نسبة الأسر التي تقوم بعملية شراء تساوي 0.065. لأن القيمة p هي 0.0085 ، وهو أقل من مستوى المعنوية 0.05 ، يرفض المحلل الفرضية الصفرية . تشير النتائج إلى أن نسبة الأسر التي تقوم بعملية شراء مختلفة عن المعدل الوطني البالغ 6.5٪

اختبار فرضيات متعلقة بمتوسط بواسوني

يعتبر توزيع بواسون حالة خاصة من توزيع ذي الحدين كما اشرنا في الفصل الخامس نستخدم اختبار الفرضيات المتعلقة بمتوسط حسابي يتبع احصائيا توزيع بواسون 1 - Sample Poisson Rate واختبار متوسط المجتمع ومقارنته بالقيمة المستهدفة أو القيمة المرجعية. باستخدام

هذا التحليل ، يمكننا القيام بما يلي عندما يتم حساب بيانات لكل وحده نحدد ما إذا كان معدل الحدوث يختلف عن المعدل المفترض الذي نحدده.

نحسب مجموعة من القيم التي من المحتمل أن تتضمن متوسط المجتمع كما قد يكون التقريب ممكنا للتوزيع الطبيعي اذا كان حجم العينة كبيرا

متال

نستخدم نفس المثال المذكور انفا في الفصل الرابع الخاص بالتقدير مع شيء من التعديل

يريد مسؤول مراقبة الجودة في خدمة النقل الحضري تحسين رضا العملاء. لتقييم مستوى رضا العملاء ، يحسب المدير عدد المطالبات المستلمة لمدة 30 يومًا والتي بلغت 598 ، يقوم المدير باختبار معدل عينة بواسون في عينة واحدة لتحديد ما إذا كان متوسط المطالبات المستلمة في اليوم أكبر من 10

الحل : بافتراض ان عدد المطالبات المستلمة في اليوم هو λ نقوم اذن باختبار الفرضيتين التاليتين

$H_0: \lambda = 10$

$H_1: \lambda > 10$

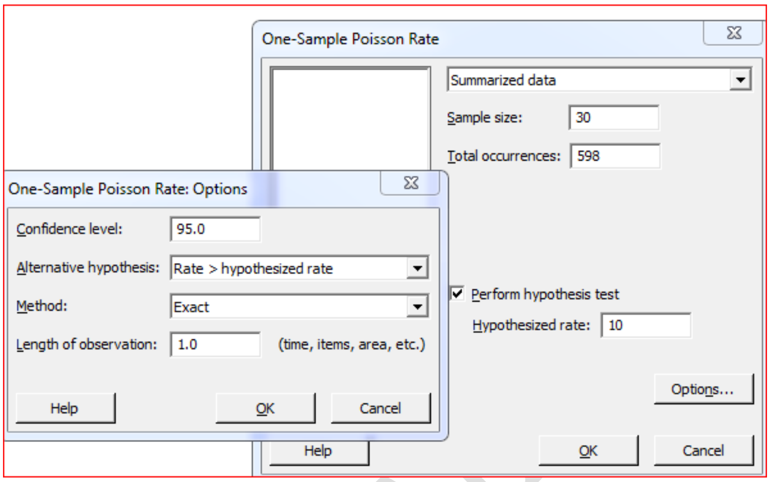
وباستخدام برنامج MINITAB لحل هذا المثال نقوم بكتابة الامر

Stat > Basic Statistics > 1- Sample Poison Rate....

نختار Summarized data لأننا لا نملك معلومات عن القيم مفصلة بل القيم معطاة بشكل تجميعي في ايقونة Sample Size نضع 30 وفي ايقونة Total occurrences نضع 598 ثم نفعل ايقونة Perform Hypothesis test ونكتب قيمة المتوسط الافتراضي في مربع Hypothesized rate تأكيدا على وجود اختبار للفرضية وبعد الضغط على Options بإمكاننا استخدام مجال الثقة عند 0,95 اما عند الفرضية البديلة فنختار

Rate >hypothezed rate وفي مثالنا هذا نختار Exact ولا نقرب هذا التوزيع الى التوزيع الطبيعي اي Normal approximation

وهذا كله موضح في المربع الحواري التالي



وبعد الضغط على OK نجد النتائج

```
Method

λ: Poisson rate of Number of complaints
Exact method is used for this analysis.

Descriptive Statistics

          Total                      95% Lower
 N  Occurrences  Sample Rate  Bound for λ
30          598      19.9333      18.6118

Test

Null hypothesis         H₀: λ = 10
Alternative hypothesis  H₁: λ > 10

P-Value
  0.000
```

تنص الفرضية الصفرية على أن معدل المطالبات هو اكبر من 10 في اليوم. نظرًا لأن القيمة الاحتمالية لـ 0.000 أقل من مستوى المعنوية البالغ 0.05 (المشار إليه بـ α أو alpha) ، يرفض المدير الفرضية الصفرية ويخلص إلى أن معدل المطالبات λ أكبر من 10 في اليوم.

اختبارات الفرضيات غير المعلميه للعينة الواحدة:

اختبار الاشارة للوسيط

نستخدم هذا الاختبار غير المعلمي لوسيط المجتمع ومقارنتها بالقيمة المستهدفة أو القيمة المرجعية. باستخدام هذا الاختبار ، يمكننا القيام بما يلي:

تحديد ما إذا كان وسيط المجتمع يختلف عن الوسيط المفترض الذي نحدده.

حساب مجموعة من القيم التي من المحتمل أن تشمل وسيط المجتمع ويعد هذا الاختبار بديلا لاختباري z –test و t-test

متال

نريد تحديد ما إذا كان محتوى الكروم الوسيطي في مجموعة من عينات الفولاذ المقاوم للصدأ يساوي 18٪ نختار بشكل عشوائي عينة من 12 مشاهدة ، ويقيس محتوى الكروم. نقوم باختبار الاشارة لعينة واحدة لتحديد ما إذا كان وسيط محتوى الكروم يختلف عن 18٪

الحل :

بعد ان نكتب نسب 12مشاهدة من مادة الكروم في ورقة عمل MINITAB نجد

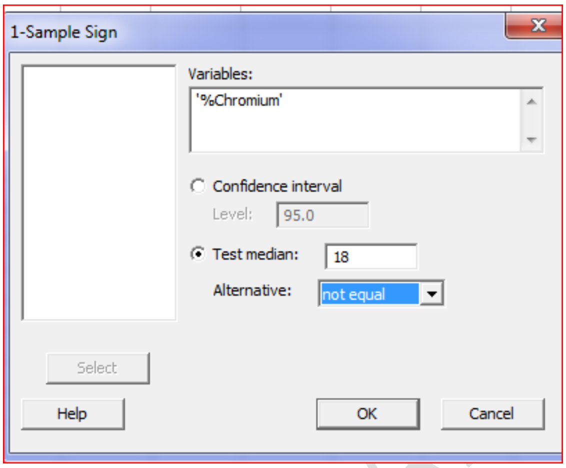
C1
%Chromium
17.4
17.8
17.6
18.1
17.6
19.0
16.9
17.5
17.8
17.4
24.6
25.9

ثم نقوم بكتابة الامر التالي

STATISTICS > One Sample > Sign

او

Stat > Nonparametrics > 1-sample Sign في حال توفر minitab 17 مثلا ويظهر المربع الحواري التالي



في خانة variables نضع نسبة الكروم %Chromium

ف نفعل خانة Test median ونضع فيها الوسيط الفرضي 18 بما ان شكل الفرضية هي

$H0: \eta = 18$

$H1: \eta \neq 18$

حيث η تمثل الوسيط فنضع في ايقونة Alternative : not equal وهذا يعني ان الفرضية البديلة لاتساوي 18

وتظهر النتائج كما يلي

Method

η: median of %Chromium

Descriptive Statistics

N	Median
12	17.7

95% Confidence Interval for η

CI for η	Achieved Confidence	Position
(17.5000, 18.1000)	85.40%	(4, 9)
(17.4263, 18.7632)	95.00%	Interpolation
(17.4000, 19.0000)	96.14%	(3, 10)

Test

Null hypothesis $H_0: \eta = 18$
Alternative hypothesis $H_1: \eta \neq 18$

Number < 18	Number = 18	Number > 18	P-Value
8	0	4	0.3877

تنص الفرضية الصفرية على أن النسبة الوسيطية لمحتوى الكروم تساوي 18٪. نظرًا لأن قيمة p تبلغ 0.39 تقريبًا ، وهي أكبر من مستوى المعنوية البالغ 0.05 ، لا يمكن رفض فرضية الصفرية . لا نمتلك أدلة كافية لاستنتاج أن محتوى الكروم الوسيط يختلف عن 18٪.

اختبار فرضيات الدورات a run test hypothsis

يستخدم هذا الاختبار في حال توفرت بيانات العينة ونريد معرفة هل هذه البيانات تم سحبها بطريقة عشوائية ام لا وفق الفرضيتين التاليتين

H0 : البياتا ت التي تم سحبها عشوائية

H1 : البياتات الت تم سحبها غير غشوائية

مثال

يريد باحث في إحدى شركات الأدوية مقارنة آثار ثلاثة علاجات على 35 مريضًا. يقوم الباحث بتخصيص علاج لكل مريض. يريد الباحث تحديد ما إذا كانت مهام العلاج عشوائية.

الحل

نقوم بكتابة الامر التالي

Stat > Nonparametrics > Runs Test

او

STATISTICS > One Sample > Runs Test في حالة توفر **Minitab19**

نقوم بتحميل معطيات المثال الاصلية في ورقة العمل

Descriptive Statistics

Number of Observations

N	K	≤ K	> K
35	2	23	12

K = sample mean

Test

Null hypothesis $\quad H_0$: The order of the data is random
Alternative hypothesis $\quad H_1$: The order of the data is not random

Number of Runs

Observed	Expected	P-Value
17	16.77	0.9304

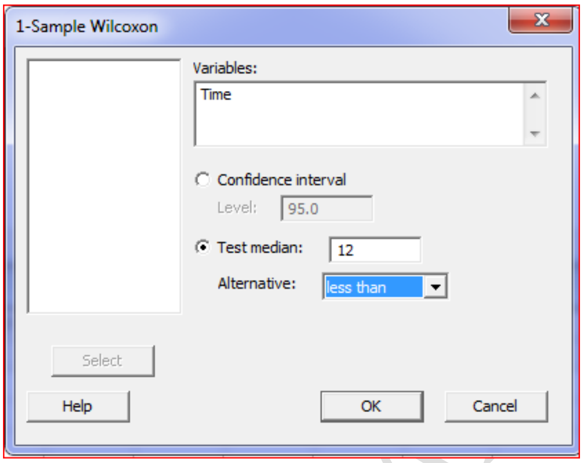
قيمة p هي 0.93 تقريبًا ، وهي أكبر من مستوى المعنوية البالغ 0.05. لا يمتلك الباحث أدلة كافية لاستنتاج أن البيانات ليست بترتيب عشوائي.

اختبار فرضيات الوسيط ويلكوكسون للعينة الواحدة :

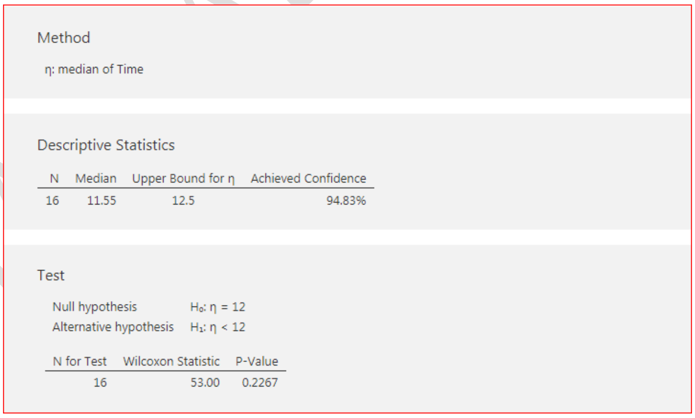
يريد الكيميائي لشركة صيدلانية اختبار ما إذا كان وسيط وقت رد الفعل لمضادات الحموضة المطورة حديثًا أقل من 12 دقيقة. يقيس الكيميائي زمن التفاعل لمدة 16 عينة من مضادات الحموضة

STATISTICS > One Sample > Wilcoxon

C1
Time
10.9
15.0
11.9
8.8
8.2
14.8
9.2
8.8
16.0
15.2
15.9
9.2
9.2
7.7
8.0
12.5



وبعد الضغط على OK تظهر النتيجة


Method

η: median of Time

Descriptive Statistics

N	Median	Upper Bound for η	Achieved Confidence
16	11.55	12.5	94.83%

Test

Null hypothesis H₀: η = 12
Alternative hypothesis H₁: η < 12

N for Test	Wilcoxon Statistic	P-Value
16	53.00	0.2267

تنص الفرضية الصفرية على أن وسيط زمن رد الفعل هو 12 دقيقة. نظرًا لأن القيمة p تقارب 0.23 ، وهو أكبر من مستوى المعنوية البالغ 0.05 ، فإن الكيميائي يفشل في رفض الفرضية الصفرية ولا يمكنه أن يستنتج أن وسيط زمن التفاعل أقل من 12 دقيقة.

اختبار فرضيات عينتين

اختبارات معلمية لعينتين

اختبار فرضية الفرق بين متوسطين لعينتين مستقلتين :

مثال

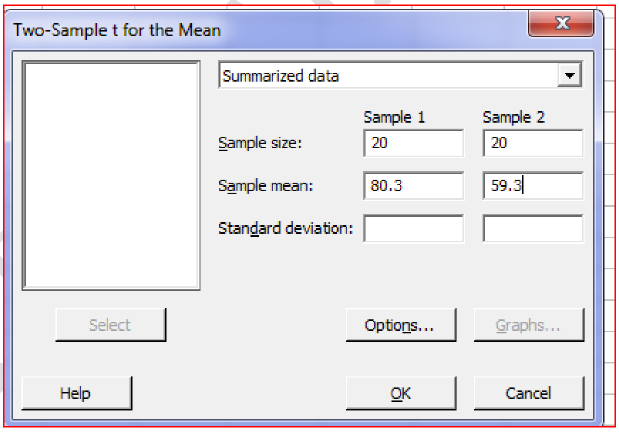
نريد مستشار الرعاية الصحية مقارنة معدلات رضا المرضى في مستشفيين A.و B تم سحب عينتين من 20 مريضا من كل مستشفى فوجد ان متوسط الرضا في المستشفين 80,3 % للمستشفى A و 59,3% للمستشفى B فهل نستطيع ان يحكم بان هناك اختلاف بين تصنيفات المستشفيين عند 5%

الحل

نقوم باختيار الامر التالي

Stat > Basic Statistics >2-samples t

فيظهر لنا مربع حواري نحتار منه Summarized data تماشيا مع معطيات المثال ثم نقوم بوضع المعطيات في ايقوناتها الخاصة بها



نلاحظ ان الانحرافين المعياريين للمجتمعين مجهولان وان حجم العينتين اقل من 30 نضغط على ok

Method

μ₁: mean of Rating when Hospital = A

μ₂: mean of Rating when Hospital = B

Difference: $\mu_1 - \mu_2$

Equal variances are not assumed for this analysis.

Descriptive Statistics: Rating

Hospital	N	Mean	StDev	SE Mean
A	20	80.300	8.183	1.830
B	20	59.300	12.431	2.780

Estimation for Difference

Difference	95% CI for Difference
21.000	(14.221, 27.779)

Test

Null hypothesis $H_0: \mu_1 - \mu_2 = 0$

Alternative hypothesis $H_1: \mu_1 - \mu_2 \neq 0$

T-Value	DF	P-Value
6.31	32	<0.0001

تنص الفرضية الصفرية على أن الفرق بين التصنيفات الخاصة بالمستشفيين هو 0. لأن القيمة p أقل من 0.0001 ، وهو أقل من مستوى المعنوية البالغ 0.05 ، كما ان قيمة T-value البالغة 6,31 هي اقل من القيمة الجدولية اذن يرفض الاستشاري الفرضية الصفرية ويخلص إلى أن التصنيفات الخاصة بالمستشفيين المستشفيات تختلف.

اختبار فرضيات الفرق بين متوسطين لعينتين مرتبطتين :

مثال

يريد اختصاصي فيزيولوجيا تحديد ما إذا كان برنامج تشغيل معين له تأثير على معدل ضربات القلب أثناء الراحة. تم قياس معدل ضربات القلب من20 شخصا تم اختيارهم عشوائيا. ثم تم وضع الأشخاص في البرنامج قيد التشغيل وتم القياس مرة أخرى . (قبل وبعد البرنامج) يجري الأخصائي اختبار t مقترنًا لتحديد ما إذا كانت معدلات ضربات القلب تختلف قبل وبعد تشغيل البرنامج

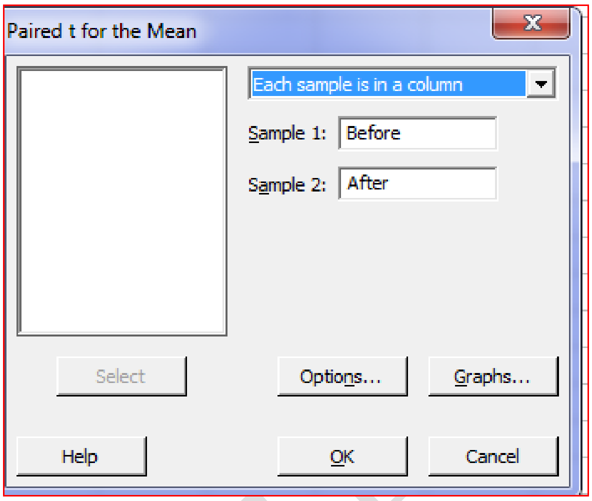
الحل

نقوم بتحميل نتائج الاختبار المتوصل عليها مباشرة في ورقة عمل Minitab فنحصل على مايلي

C1 Before	C2 After	C3 Difference	C4
68	67	1	
76	77	-1	
74	74	0	
71	74	-3	
71	69	2	
72	70	2	
75	71	4	
83	77	6	
75	71	4	
74	74	0	
76	73	3	
77	68	9	
78	71	7	
75	72	3	
75	77	-2	
84	80	4	
77	74	3	
69	73	-4	
75	72	3	
65	62	3	

ثم نقوم بكتابة الامر الاتي

يظهر المربع الحواري الاتي Stat>Basic Statistics > Paired t....



نضغط على ok فتظهر النتائج التالية

Descriptive Statistics

Sample	N	Mean	StDev	SE Mean
Before	20	74.500	4.513	1.009
After	20	72.300	4.054	0.906

Estimation for Paired Difference

Mean	StDev	SE Mean	95% CI for µ{sub}d{#sub}
2.2000	3.2541	0.7276	(0.6770, 3.7230)

µ{sub}d{#sub}: mean of (Before - After)

Test

Null hypothesis H₀: µ{sub}d{#sub} = 0
Alternative hypothesis H₁: µ{sub}d{#sub} ≠ 0

T-Value	P-Value
3.02	0.0070

تنص الفرضية الصفرية على أن الفارق المتوسط في أوقات الجري هو 0. لأن القيمة p هي 0.007 ، وهو أقل من مستوى المعنوية البالغ 0.05 ، يرفض الفيزيولوجي الفرضية الصفرية ، ويخلص إلى أن هناك فرقًا بين معدل ضربات القلب من الموضوعات اختبار قبل وبعد البرنامج قيد التشغيل.

اختبار فرضيات لنسبة تبايني مجتمعين :

يمكن استخدام فرضيات متعلقة بالنسبة بين تبايني مجتمعين للقيام بما يلي يلي:

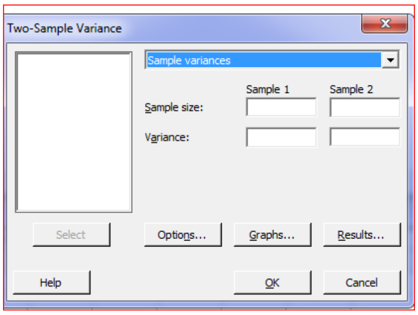
تحديد ما إذا كانت الاختلافات أو الانحرافات المعيارية لمجموعتين تختلف.

حساب مجموعة من القيم التي من المحتمل أن تشمل نسبة المجتمع من التباينات أو الانحرافات المعيارية للمجموعتين.

متال

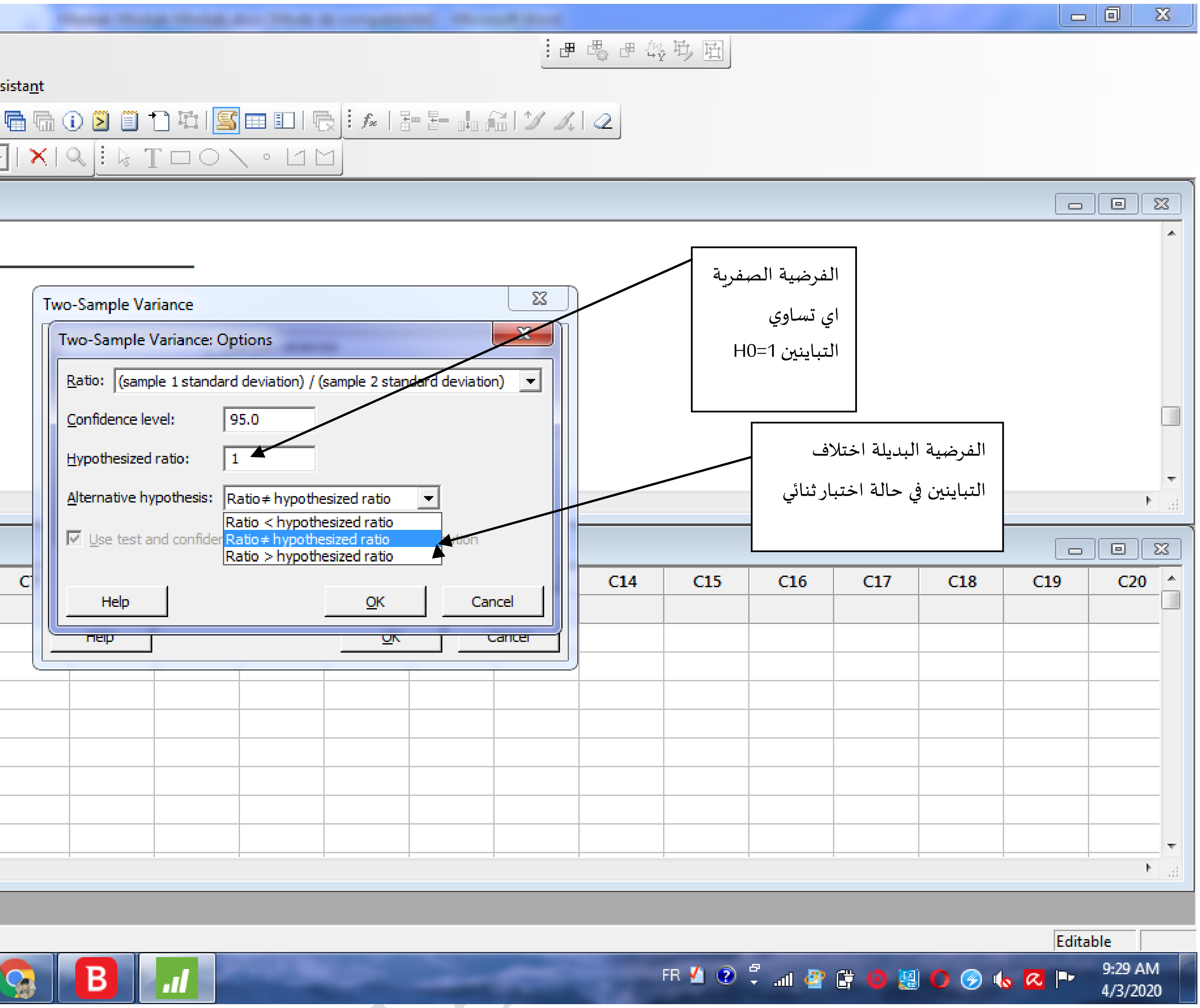
بتفعيل الامر التالي Stat >Basic Statistics> 2 Variances ثم الضغط على OK

يمكننا افتراض امثلة مباشرة في المربع الحواري التالي مع الاشارة الى ان هذه الطريقة تمكننا من استخدام التباين او الانحراف المعياري بحسب توفر المعطيات

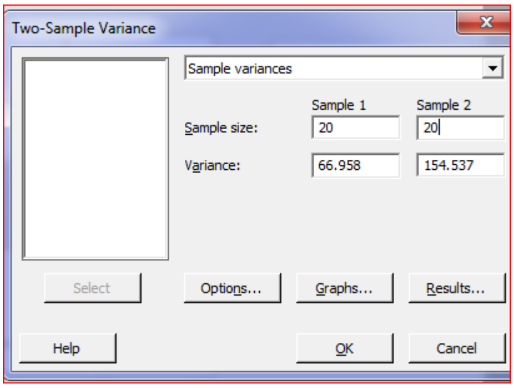


نلاحظ اننا اخترنا تباينين لعينتين مختلفتين

وهنا يمكن اجراء الفرضيات المقترحة وذلك بالنقر على ايقونة Options



وبوضع القيم المفترضة التالي بعد الرجوع الى المربع الحواري التالي



ثم الضغط على   OK من جديد

Descriptive Statistics: Rating

Hospital	N	StDev	Variance	95% CI for σ
A	20	8.1828	66.958	(5.8931, 12.5966)
B	20	12.4313	154.537	(8.6927, 19.7093)

Ratio of Standard Deviations

Estimated Ratio	95% CI for Ratio Bonett	95% CI for Ratio Levene
0.658241	(0.37247, 1.21499)	(0.37779, 1.29619)

Test

Null hypothesis $H_0: \sigma_1 / \sigma_2 = 1$
Alternative hypothesis $H_1: \sigma_1 / \sigma_2 \neq 1$

Method	Test Statistic	DF1	DF2	P-Value
Bonett	2.09			0.1485
Levene	1.60	1	38	0.2141

. كما هو معلو م في الاحصاء الاستدلالي فان النسبة بين تبايني مجتمعين تتبع توزيع F الاحصائي وهذا حسب معياري كل من Bonett و Levene

اختبار فرق نسبتي مجتمعين

يمكن استخدام اختبار الفرضيات المتعلقة بفرق نسبتين للقيام بما يلي:

تحدد ما إذا كانت نسبتا مجتمعين مختلفين. او لا

نحسب مجموعة من القيم التي من المحتمل أن تتضمن الفرق بين نسبتي المجتمعين وإحصائيا يتبع هذا الفرق التوزيع الطبيعي وذلك استنادا الى نظرية النهاية المركزية خاصة في حالة المجتمعات الكبيرة مع العلم ان توزيع النسب في العموم يتبع توزيع ذي الحدين

مثال

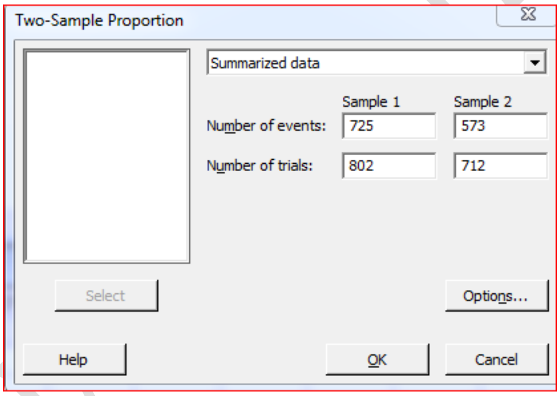
يقوم مكتب دراسات خاص بإحصاء التشغيل وفرص الحصول على وظائف بأخذ عينات من الطلاب الجامعيين المتخرجين حديثا لتحديد ما إذا كان الطلاب الذكور أو الإناث أكثر احتمالا للحصول على وظيفة. من بين 802 طالبًا تم تخرجه ، حصل منهم 725 على وظيفة ، و من بين 712 طالبة تخرجن حصلت 573 طالبة على وظيفة.

هل يمكن ان يقبل المركز فرضية عدم وجود فرق جوهري بين نسبتي الطلبة المتخرجين والطالبات المتخرجات

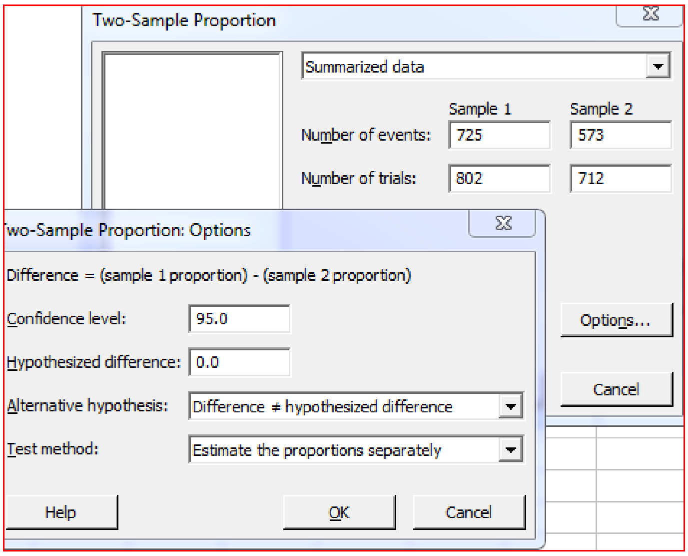
الحل : بعد وضع الفرضيتين HO الفرق بين النسبتين متساوي اي p1-p2=0

في حين ان الفرضية البديلة H1 تنص على عدم تساوي الفرق بين النسبتين اي p1-p2≠**0**
ومن ثم نكتب الامر الاتي في برنامج MINITAB

Stat >Basic Statistics 2- Proportions ثم نختار Summarized data وهذا يوافق معطيات المثال لاننا نملك معلومات عن حجم كل عينة وكذا النسبتين وهذا يوضحه المربع الحواري الاتي



نلاحظ اننا وضعنا في العينة الاولى حجم العينة الخاص بالطلبة اي Number of trials و نسبة الطلبة الحاصلين على وظيفة Number of events وبنفس الطريقة يمكن التعامل مع العينة الثانية الخاصة بالطالبات وبعدها نضغط على Options



نلاحظ اننا اخترنا مستوى الثقة 95 % اي 0.95 Confidence level

وفي ايقونة الفرضية الصفرية او فرضية الفرق اخترنا عدم وجود فرق Hypothesized difference 0.0

و في ايقونة الفرضية البديلة اخترنا اختلاف الفرق Alternative Hypothesis : Deference ≠hepothesis difference

وبعد الضغط على OK

Descriptive Statistics

Sample	N	Event	Sample p
Sample 1	802	725	0.903990
Sample 2	712	573	0.804775

Estimation for Difference

Difference	95% CI for Difference
0.099215	(0.063671, 0.134759)

Test

Null hypothesis $H_0: p_1 - p_2 = 0$
Alternative hypothesis $H_1: p_1 - p_2 \neq 0$

Method	Z-Value	P-Value
Fisher's exact		<0.0001
Normal approximation	5.47	<0.0001

تنص الفرضية الصفرية على أن الفرق في نسبة الطلاب الذكور ونسبة الطالبات اللواتي يحصلن على وظيفة هو 0. لأن القيمة الاحتمالية أقل من 0.0001 ، وهي أقل من مستوى الدلالة البالغ 0.05 (نسبة الثقة 0,95)، لهذا فان مكتب التشغيل يرفض الفرضية الصفرية. تشير النتائج إلى وجود فرق بين نسبة الطلاب الذكور الذين يحصلون على وظيفة ونسبة الطالبات اللواتي يحصلن على وظيفة

اختبار فرضيات متعلقة بفرق متوسطين بواسونين

يريد محلل الخدمات البريدية مقارنة عدد زيارات العملاء في فرعين لمركز البريد. يحسب المحلل عدد العملاء الذين يدخلون كل فرع لمدة 40 يوم عمل. حيت كانت عدد زيارات الفرع الاول 9983 في حين كانت عدد زيارات الفرع الثاني 10291

يقوم المحلل بإجراء اختبار الفرق بين متوسطين بواسونين في عينتين لتحديد ما إذا كان المعدل اليومي لزيارات العملاء يختلف بين فرعي مكتب البريد.

الحل

ادا رمزنا ل λ_i متوسط عدد عملاء كل فرع في اليوم الواحد

اي اختبار الفرضيتين

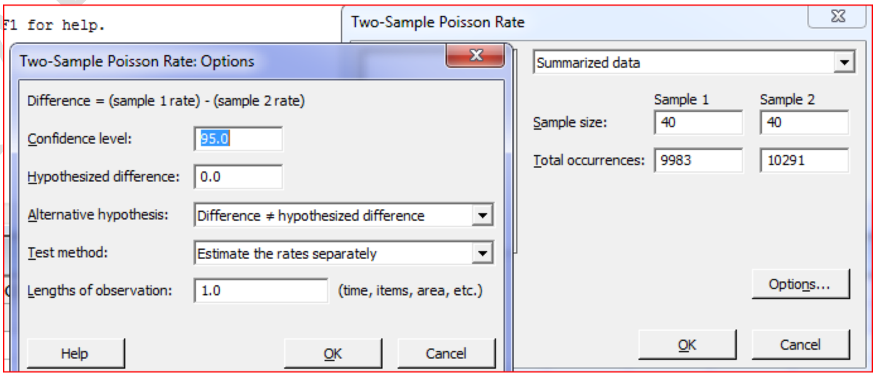
$$H_0: \lambda_1 - \lambda_2 = 0$$
$$H_1: \lambda_1 - \lambda_2 \neq 0$$

نقوم بكتابة الامر التالي

Stat >Basic Statistics> 2-Sample Poisson Rate

نختار Summarized data ونكتب قيمة حجم وحوادث كل العينة ثم نضغط على ايقونة options

ونؤشر على مستوى الثقة بالإضافة الى الفرضية الصفرية والبديلة فيظهر لنا هذان المربعان الحواريان



ثم نضغط على OK من جديد لتظهر لنا النتيجة

```
Method

λ₁: Poisson rate of Branch A
λ₂: Poisson rate of Branch B
Difference: λ₁ - λ₂

Descriptive Statistics

                     Total
Sample      N  Occurrences  Sample Rate
Branch A  40         9983      249.575
Branch B  40        10291      257.275

Estimation for Difference

 Estimated        95% CI for
Difference        Difference
      -7.7  (-14.6768, -0.723175)

Test

Null hypothesis         H₀: λ₁ - λ₂ = 0
Alternative hypothesis  H₁: λ₁ - λ₂ ≠ 0

Method                Z-Value  P-Value
Exact                            0.031
Normal approximation    -2.16    0.031
```

تنص الفرضية الصفرية على أن الفرق في المعدل اليومي لزيارات العملاء بين فرعي مكتب البريد هو 0. نظرًا لأن القيمة الاحتمالية 0.031 أقل من مستوى المعنوية (المشار إليه بـ α أو ألفا) البالغ 0.05 ، اذن يرفض المحلل الفرضية الصفرية ويخلص إلى أن المعدل اليومي لزيارات العملاء يختلف بين فرعي مكتب البريد. ،كما يشير مؤشر الثقة 95٪ إلى أنه من المرجح أن يحصل الفرع ب على معدل زيارات عملاء أعلى من الفرع أ

اختبار الفرضيات اللامعلمية لعينتين

اختبار فرضية متعلقة بفرق وسيطين

يمكن اختبار فيق وسيطين من خلال برنامج Minitab و ذلك بالاعتماد على اختبار غير معلمي هو Mann-Whitney

متال نفرض نفس المتال

قدمت للسوق علامتان تجاريتان و كانت المقارنة بينها من خلال وسيط عدد اشهر التعمير لكل علامة

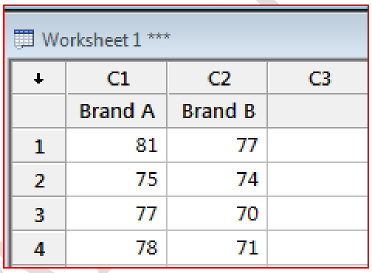
وسيط العلامة التجارية	81	75	77	78

				1
71	70	74	77	وسيط العلامة التجارية 2

فهل يوجد فرق بين وسيطين اشهر التعمير

الحل

نقوم بنقل الجدول من برنامج الاكسل الى ورقة عمل Minitab


Worksheet 1 ***

↓	C1	C2	C3
	Brand A	Brand B	
1	81	77	
2	75	74	
3	77	70	
4	78	71	

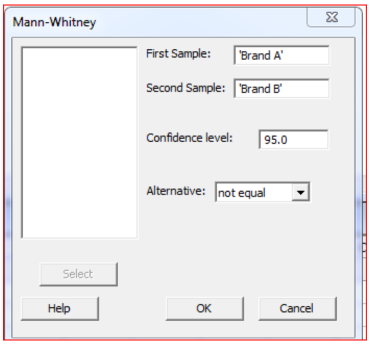
اذا رمزنا لوسيط كل علامة تجارية بالرمز η ثم نختبر الفرضيتين التاليتين حول الفرق بين وسيطين

$$\eta_1 - \eta_2 = 0$$

$$\eta_1 - \eta_2 \neq 0$$

ونكتب التعليمة التالية

Stat > Nonparametrics > Mann-Whitney. ثم ندون المعلومات في المربع الحواري



Mann-Whitney: Brand A, Brand B

```
Method

η1: median of Brand A
η2: median of Brand B
Difference: η1 - η2

Descriptive Statistics

 Sample   N  Median
Brand A  11    36.0
Brand B  10    37.6

Estimation for Difference

                CI for       Achieved
Difference   Difference    Confidence
     -1.85  (-3,    -0.9)       95.52%

Test

Null hypothesis         H0: η1 - η2 = 0
Alternative hypothesis  H1: η1 - η2 ≠ 0

Method                  W-Value  P-Value
Not adjusted for ties     76.50    0.002
Adjusted for ties         76.50    0.002
```

تنص الفرضية الصفرية على أن الفرق في وسيط عدد الأشهر بين العلامتين التجاريتين هو 0. ولأن القيمة p هي 0.0019 ، وهو أقل من مستوى المعنوية وهو 0.05 ، نرفض الفرضية الصفرية. ونؤكد إلى الفرق بين وسيطي عدد الأشهر بين العلامتين التجاريتين ليس صفرًا.

المراجع :

باللغة العربية :

اسامة امين ربيع سليمان (2007) دليل الباحثين في التحليل الاحصائي للبيانات باستخدام برنامج MINITAB كلية التجارة جامعة المتوفية جمهورية مصر العربية

باللغة الاجنبية :

Amechi H. Igweze & Harrison Etaga(2011) STATISTICAL ANALYSES With Excel, Minitab and SPSS file:///C:/Users/BLACK%20HOLE/Downloads/StatisticalAnalyses.pdf

D. B. Rorabacher (1991), "Statistical Treatment for Rejection of Deviant Values: Critical Values of Dixon Q Parameter and

Detection", Journal of Statistical Software, vol. 16, No. 3, pages 1-9.

E. P. King (1953), "On Some Procedures for the Rejection of Suspected Data", *Journal of the American Statistical Association*, vol.

G. C. McBane (2006), "Programs to Compute Distribution Functions and Critical Values for Extreme Value Ratios for Outlier

Jan-Eric Englund (2011) MINITAB a primer release 16, Swedish university of agricultural sciences

Jerry Garcia (2007) Six Sigma Statistics with Excel and Minitab , Mc Graw Hill

Nadarajah Ramesh(2009) The role of Minitab in teaching and learning statistics , MSOR Connections

Vol 9 No 3 August – October 2009

Micheal Evans(2009) MINITAB Manual for introduction to the practice of statistics , university of Toronto , Canada

SHONDA KUIPER and Jeffrey Sklar (2013) PRACTICING STATISTICS: GUIDED INVESTIGATIONS FOR THE SECOND course Publishing as Pearson, 75 Arlington Street, Boston, MA 02116. Subrange Ratios at the 95 percent Confidence Level", *Analytic Chemistry*,

83, 2, 139-146.

W. J. Dixon (1951), "Ratios Involving Extreme Values", *Annals of Mathematical Statistics*, 22(1), 68-78.