محاضرات

مادة الاحصاء التربوي

**الدكتور عباس علي شلال**

**الفصل الأول**

**المتغيرات والعلاقة فيما بينها**

* المتغيرات وانواعها
* تدريبات شكل الانتشار
* تدريبات معامل ارتباط بيرسون
* تدريبات معامل ارتباط سبيرمان
* تدريبات معامل ارتباط كندال
* تدريبات معامل ارتباط بوينت باي سيريال
* تدريبات معامل ارتباط فاي
* تدريبات معامل الارتباط الجزئي

**المتغيرات وأنوعها**

يشير مصطلح المتغيرات Variables الى الخصائص التي يشترك فيها أفراد المجتمع الاحصائي ، ولكنها تختلف من فرد الى آخر ، فالعمر ، ودرجة الذكاء ، وطول القامة ، واللياقة البدنية ، والقدرة على القراءة ، والراتب الشهري أمثلة على المتغيرات ، وتتميز هذه المتغيرات بأنها قابلة للقياس الكمي وامكانية تحديد قيمة معينة لها ؛ إذن هي ظاهرة او حدث او خاصية لها قيماً تتغير من ظرف لآخر ، كما انها الوحدة الأساسية للتحليل الاحصائي.

ويمكن تصنيف المتغيرات الى أكثر من طريقة ولكل منها تفسير منطقي يرتكز عليه ومنها:

1. *المتغيرات الكمية ، والمتغيرات النوعية ..* يرتكز هذا التصنيف على مدلول القيمة الممثلة للخاصية المقاسة ، فاذا كانت هذه القيمة تشير الى مقدار ما في الفرد من خاصية مقارنة بافراد مجموعته فان هذه القيمة تحمل معنى كمي وان هذا المتغير هو متغير كمي ، اي ان المتغيرات الكمية هي التي تأخذ قيماً كمية أو عددية ومن امثلتها درجات الطلبة في مادة معينة والتي تتراوح بين (30-90) درجة ، والوزن بالكيلوغرام لعينة من الأطفال في الصف الاول الابتدائي .

أما المتغيرات النوعية فهي تلك المتغيرات التي نصف فيها مفردات او عناصر المجتمع الاحصائي او العينة في عدة مجموعات تشترك كل مجموعة في صفة معينة ، مثلا النوع (ذكور – اناث) او الحالة الاجتماعية (أعزب-متزوج) او نوع العمل (موظف-عمل حر-متقاعد).

1. *المتغيرات المتصلة (المستمرة) ، والمتغيرات المنفصلة(المتقطعة) ..* فالمتغيرات المتقطعة هي المتغيرات الكمية التي تأخذ قيماً عددية محددة صحيحة ولا تحتوي على قيم كسرية مثل عدد المدرسين في كل مدرسة ثانوية في العراق ، او عدد الطلبة في القاعات الدراسية المختلفة ، أو عدد أفراد الأسرة.

أما المتغيرات المتصلة فهي التي تأخذ قيماً عددية غير محددة ومن الممكن أن تحتوي قيماً كسرية ، كوحدات القياس –الكيلوغرام في الاوزان والشهور للأعمار- ودرجات الذكاء والتحصيل.

1. *المتغيرات المستقلة ، والمتغيرات التابعة ..* ان المتغير المستقل Independent Variable هو المتغير الذي يُحدث تغيراً في متغير آخر ويؤثر فيه وعادة ما يستخدم في البحوث التجريبية كما في "أثر طريقة تدريس حديثة على تحصيل الطلبة في إحدى المواد العلمية" فان طريقة التدريس الحديثة هي المتغير المستقل.

بينما المتغير التابع Dependent Variable فهو الذي يتعرض للتغير نتيجة لأثر المتغير المستقل كما في تحصيل الطلبة في المثال السابق.

ولا بد من الإشارة هنا الى ان المتغير مستقلا او تابعا هو امر نسبي وليس مطلقا ، فكون المتغير مستقلا او تابعا يتوقف على موقعه ومدى تأثيره أو تأثره بالعوامل الأخرى ، ويمكن القول ان المتغير الذي يكون مستقلا في حالة معينة قد يكون تابعا في حالات أخرى.

(البياتي واثناسيوس ، 1977 : 19)

**الارتباط ..**

الارتباط **Correlation** في معناه العلمي هو التغير الإقتراني ، أو بمعنى آخر هو النزعة الى اقتران التغير في ظاهرة بالتغير في ظاهرة أخرى ، كما في تغير طول عمود من الحديد تبعاً لتغير درجات الحرارة التي يتعرض لها ، فكلما زادت درجات الحرارة زاد تبعاً لذلك الطول ، وكلما بدأت تلك الزيادة بالتناقص تناقص تبعاً لذلك الطول ، أي ان تغير الطول يقترن بتغير الحرارة ؛ وأيضا مثل نقصان حجم قطعة الثلج تبعاً لزيادة درجات الحرارة ، فكلما زادت درجات الحرارة نقص حجم الثلج ، أي ان تغير حجم الثلج يقترن بتغير درجات الحرارة (السيد ،1971 : 289).

ويجب التنبيه الى ان الارتباط بين ظاهرتين متغيرتين ليس دليلاً على ان إحداهما نتيجة للأخرى ، وان التغير في واحدة تابع للتغير في الأخرى ولا ينشأ إلا بسببها ، بل هو يشير فقط الى احتمال وجود هذه العلاقة ، لان هذه العلاقة ما هي الا نوع خاص من أنواع العلاقات التي يدل الارتباط على وجودها ، وهذه الأنواع المختلفة للعلاقات تتمثل في:

* *حالة العلاقة السببية المباشرة* .. أي أن يكون أحد المتغيرين نتيجة مباشرة للمتغير الآخر كالعلاقة بين نظام معين للتعزيز وكفاءة التعلم.
* *حالة العلاقة السببية غير المباشرة* .. كأن يكون أحد المتغيرين سبباً غير مباشر للثاني يؤثر فيه بواسطة متغير ثالث ، كالعلاقة بين الطول والوزن في بحوث النمو ، فهذه العلاقة تنشأ عن متغير ثالث هو العمر الزمني أو الصحة الجسمية.
* *حالة أن يؤثر عامل واحد في المتغيرين معاً* .. وفي ذلك يكون كل من المتغيرين المرتبطين نتيجة عامل آخر ثالث مشترك بينهما يؤثر فيهما في وقت واحد فيكون التغيير في أحدهما مصحوباً بالتغير في الآخر ، كما في الارتباط بين أسعار سلعتين تمتلكها طبقة معينة من السكان فأن اسعارهما تكون مرهونة بالحالة الاقتصادية لهذه الطبقة ، او ارتباط السلعتين بأسعار النقل.
* *حالة أن تكون بعض العوامل مشتركة بين المتغيرين* .. ومن ذلك مثلا لو اختبرنا عددا من التلاميذ في مادتين مثل جغرافية العالم الاسلامي وتاريخ العالم الاسلامي فإننا نجد الارتباط شديدا بين درجة هاتين المادتين والسبب في ذلك ان الأداء في الاختبارين يوجد فيه بعض العوامل المشتركة.

(ابو حطب وصادق ، 1991 : 245-246)

ويعرف **معامل الارتباط Correlation Coefficient** على انه العلاقة بين متغيرين او أكثر ، فنحتاج أحياناً الى معرفة العلاقة بين متغيرين (ظاهرتين) او أكثر ، ودرجة ارتباطهما او علاقة بعضهما ببعض ، اذ ان دراسة العلاقة بين ظاهرتين ومعرفة مقدار هذه العلاقة أمر مهم جدا في حياتنا اليومية والمستقبلية مثل العلاقة بين العمر والطول لمجموعة من الأطفال ، والعلاقة بين غياب الطالب عن المدرسة وتحصيله الأكاديمي.

**مؤشرات الارتباط** .. للكشف عن درجة الارتباط بين متغيرين يوجد مؤشران رئيسان هما شكل الانتشار ومعامل الارتباط.

ويعد معامل الارتباط أهمهما لأنه مؤشر كمي على قوة واتجاه العلاقة ، وتتراوح قيمة معامل الارتباط بين ( +1 --- 0 --- -1) ، حيث تدل القيمة المطلقة على قوة الارتباط ، وتدل الإشارة الى إتجاه العلاقة ، أي يمكن ان يكون الارتباط سالب (عكسي) ، او موجب (طردي).

وتدل القيمة (-1) على ان معامل الارتباط تام سالب وتقع جميع النقط على الخط المستقيم ، وهنا تقل قيم المتغير (س) بزيادة قيم المتغير (ص) او العكس ، وتدل القيمة (+1) على ان معامل الارتباط تام موجب ، وتقع جميع النقط على الخط المستقيم ، وتزيد قيم المتغير (س) بزيادة قيم المتغير (ص) او العكس ، أما القيمة (صفر) فانها تشير الى ان المتغيرين (س ، ص) مستقلان بعضهما عن بعض.

(علام ، 2003 : 268)

أما **شكل الانتشار Scatter Diagram** أو مخطط الانتشار Scatter plot فهو أداة بيانية مفيدة ، بل أمر أساس قبل البدء بأي تحليل للارتباط وذلك لمعرفة المنحى العام للبيانات ، وشكل الانتشار مخطط يتم فيه رسم كل درجة من درجات الشخص للمتغير الأول مقابل درجته على المتغير الثاني –*كما يمكن رسم درجته على متغير ثالث*- ويستعمل لعرض العلاقة بين متغيرين ، إذ يعطي فكرة سريعة عن العلاقة واتجاهها دون حساب معامل الارتباط ، ويمكن رسم خط الملائمة الأفضل Best Fit Line او ما يعرف بخط الانحدار لأجراء المقارنة المنظورة بين هذا الخط وبين النقاط حوله والتي تمثل تقاطع قيم المتغيرين موضوع الدراسة ، فكلما كانت مجموعة النقاط قريبة من هذا الخط كلما كانت العلاقة بين المتغيرين أقوى ، وكذلك كلما كانت هذه النقاط مبعثرة أكثر كانت العلاقة بين المتغيرين ضعيفة.

وسوف نستعرض مجموعة من الأمثلة توضح شكل الانتشار وتمثل العلاقة بيانياً وكيفية رسم مخطط الانتشار وأسلوب تطبيق ذلك في البرنامج الإحصائي spss :

مثال : لمعرفة علاقة متغير العمر بمتغير الوزن لـ (8) أطفال كانت بياناتهم كالآتي :

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| الاطفال | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| الوزن *بالكغم*  س | 4 | 5 | 5.5 | 6 | 7 | 8 | 8.5 | 9 |
| العمر *بالشهر*  ص | 2 | 4 | 6 | 8 | 10 | 12 | 14 | 16 |

ولمعرفة شكل الانتشار واستنتاج نوع العلاقة من الرسم نقوم بتمثيل البيانات بيانياً ، ونقوم بتعيين النقطة الخاصة بكل طفل وذلك بمعرفة القيمتين (س) ، و(ص) على المحورين السيني (X-axes) والصادي (Y-axes) ، فالطفل الأول الذي حصل على قيمة مقدارها (4) في متغير الوزن (س) ، وقيمة مقدارها (2) في متغير العمر (ص) تتحدد النقطة الخاصة به بواسطة اقامة عمودين احدهما على المحور السيني عند القيمة (س) ، والآخر على المحور الصادي عند القيمة (ص) ، ويشكل التقاء العمودين النقطة المطلوبة لذلك الطفل وهكذا لجميع القيم ، وبعد الانتهاء من تعيين النقاط كافة يمكن ان تتضح لنا نوعية العلاقة من خلال اتجاه النقط وشكل انتشارها وكما موضح في الشكل الآتي:

**ص**

**س**

10 11 12 13 14 15 16 9 8 7 6 5 4 3 2 1

1

2

3

5

7

8

10

11

12

6

9

4

ومن الشكل ومخطط الانتشار يتضح ان هناك علاقة موجبة بين متغيري الوزن والعمر للأطفال الثمانية.

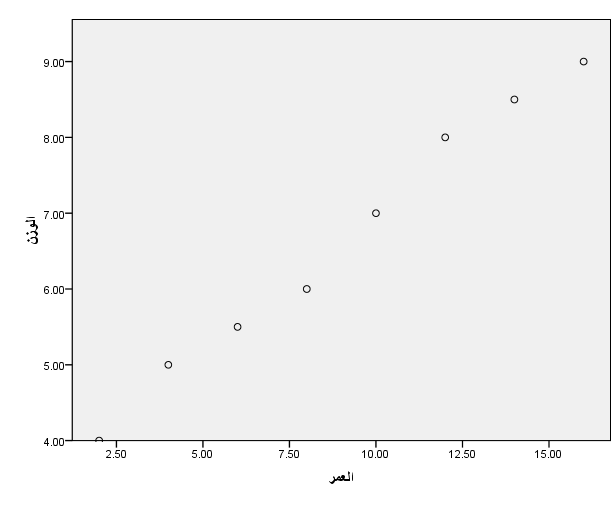
وعند تطبيق المثال السابق في البرنامج الإحصائي spss من أجل التعرف على مخطط الانتشار والعلاقة بين متغيري الوزن والعمر للأطفال ، ينبغي إتباع الآتي:

1-إدخال بيانات الأطفال الثمانية للمتغيرين بعد تسميتهما -حسب معطيات المثال او السؤال- في برنامج spss ثم النقر على Graphs في الخيارات الأساسية للبرنامج والتي تقع أعلى الواجهة الرئيسة للبرنامج ، ثم اختيار Legacy Dialogs ثم Interactive ثم الأمر line أو Scatterplot وكما موضح في الشكل.

2- سوف تظهر نافذة الحوار Create Scatterplot الخاصة بشكل الانتشار ، فيتم –عن طريق الفأرة (الماوس)- سحب كل من المتغيرين وإدخاله الى الحقل المخصص له ليمثلا المحورين السيني والصادي في مخطط الانتشار وكما موضح في الشكل الآتي:



3- يتم النقر على زر ok فتظهر نافذة النتائج Output والتي تمثل مخطط الانتشار للعلاقة بين قيم الأطفال على متغيري العمر والوزن ، وكما موضح في الشكل.



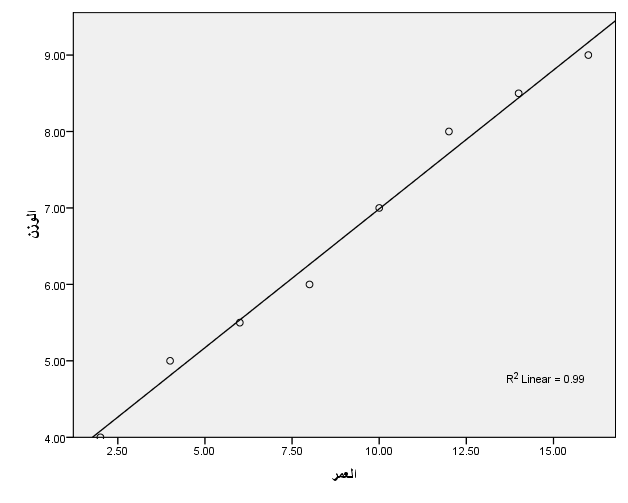
ولأجل اكتمال صورة مخطط الإنتشار وتحديد خط الملائمة الأفضل يتم النقر على الشكل نفسه فتظهر نافذة الحوار Chart Editor ومن ثم يتم النقر على الخيار Add Fit Line at Total كما في الشكل الآتي.

Add Fit Line at Total



فتظهر نافذة حوار جديدة بإسم Properties وفيها الاختيار الافتراضي هو Linear فيتم النقر على Close وكما موضح في الشكل.

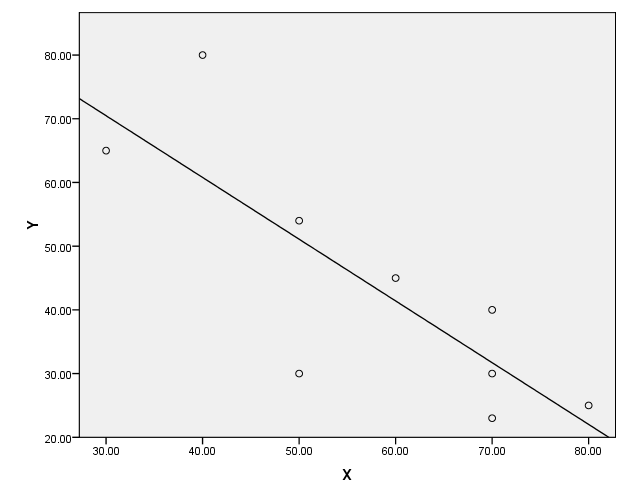
وبعدها يتغير شكل مخطط الانتشار الى النتيجة النهائية المطلوبة وكما موضح في الشكل.



ومن الشكل ، ومخطط الانتشار يتضح ان جميع النقط تقع قريبة من خط مستقيم وبالاتجاه الموجب مما يشير الى وجود علاقة قوية موجبة لكنها غير تامة.

مثال2: الجدول الآتي يبين درجات مجموعة من الطلبة في مادتي التاريخ والرياضيات.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| الطلبة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| التاريخ (X) | 80 | 70 | 70 | 50 | 60 | 70 | 50 | 30 | 40 |
| الرياضيات (Y) | 25 | 40 | 30 | 30 | 45 | 23 | 54 | 65 | 80 |

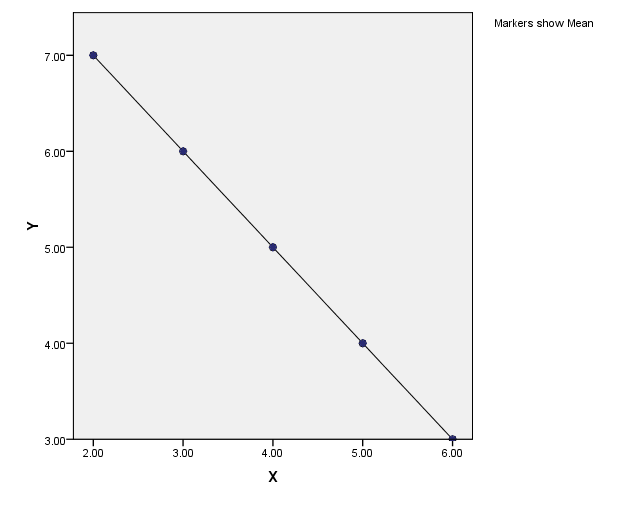
الحل :

ونلاحظ من الشكل ومخطط الانتشار ان جميع النقط تقع قريبة من الخط المستقيم وبالاتجاه السالب مما يشير الى وجود علاقة قوية عكسية.

مثال3: الجدول الآتي يبين درجات (5) تلاميذ في مادتي العلوم والرياضيات.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| التلاميذ | 1 | 2 | 3 | 4 | 5 |
| درجات مادة العلوم (X) | 2 | 3 | 4 | 5 | 6 |
| درجات مادة الرياضيات (Y) | 7 | 6 | 5 | 4 | 3 |

الحل:



ونلاحظ من الشكل ومخطط الانتشار ان جميع النقط تقع على خط مستقيم وبالاتجاه السالب مما يشير الى وجود علاقة تامة عكسية ( r= -1) .

مثال4: الجدول الآتي يبين درجات (4) تلاميذ في مادتي العلوم والرياضيات.

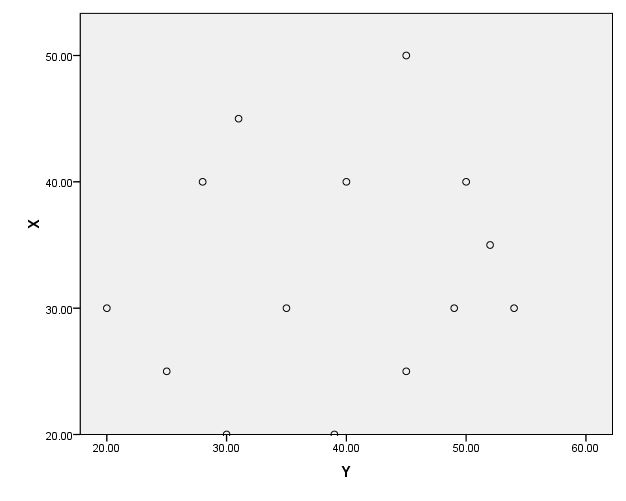
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| التلاميذ | 1 | 2 | 3 | 4 |
| درجات العلوم (X) | 2 | 3 | 5 | 6 |
| درجات الرياضيات (Y) | 2 | 3 | 5 | 6 |

الحل:

ومن الشكل ومخطط الانتشار يتضح ان جميع النقط تقع على خط مستقيم وبالاتجاه الموجب مما يشير الى وجود علاقة تامة موجبة ( r= +1) .

مثال5: الجدول الآتي يبين درجات مجموعة من الطلبة عددهم (14) طالباً في مادتي الرياضيات والتاريخ.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| الطلبة | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 |
| رياضيات  (X) | 20 | 30 | 30 | 40 | 50 | 25 | 40 | 45 | 30 | 35 | 40 | 25 | 20 | 30 |
| تاريخ  (Y) | 30 | 20 | 35 | 28 | 45 | 45 | 40 | 31 | 49 | 52 | 50 | 25 | 39 | 54 |

الحل:

ومن الشكل ومخطط الانتشار يتضح ان النقاط مبعثرة مما يدل على عدم وجود علاقة بين درجات مادتي الرياضيات والتاريخ (صفر = r).

نشاط1: أمامك جدول يوضح درجات مجموعة من التلاميذ في مادتي العلوم والرياضيات ، إرسم شكل انتشار الدرجات .. ثم بين طبيعة العلاقة فيما بينها:

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| العلوم | 9 | 10 | 5 | 5 | 4 | 9 | 9 | 7 | 3 | 5 |
| الرياضيات | 8 | 9 | 6 | 5 | 3 | 8 | 7 | 6 | 2 | 8 |

نشاط2: أمامك جدول يوضح درجات مجموعة من الطلبة في مادتي الرياضيات والتاريخ ، ارسم شكل انتشار الدرجات .. ثم بين طبيعة العلاقة فيما بينها:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| الرياضيات | 65 | 73 | 82 | 90 | 45 | 44 | 20 | 35 | 39 | 40 | 50 | 18 | 90 |
| التاريخ | 67 | 80 | 65 | 69 | 79 | 50 | 60 | 75 | 18 | 65 | 67 | 40 | 89 |

**الفصل الثاني**

**دلالة الفروق بين العينات**

* اختبار دلالة الفروق بين العينات
* تدريبات الاختبار التائي لعينة واحدة
* تدريبات الاختبار التائي للعينات المستقلة
* تدريبات حجم الأثر
* تدريبات الاختبار التائي للعينات المترابطة

**اختبار دلالة الفروق بين العينات-الاختبار التائي**

من المعروف ان إتخاذ أي قرار لا يتم إلا من خلال إختبارات الفروض الاحصائية التي تعتمد بدورها على الاحتمالات وتوزيعات المعاينة ، وهذا يؤكد الدور الذي تلعبه نظرية الاحتمالات في التنبؤ والتخطيط وإتخاذ القرارات فضلاً عن أهميتها في تقدير معالم المجتمع والتي تعد إحدى أهم إهتمامات الباحثين.

­­­

1. **الاختبار التائي لعينة واحدة )عينة ومجتمع)**

يختص النوع الاول وهو One sample t-test بإجراء إختبار الفرضيات Testing Hypotheses تدور حول معالم المجتمع المجهولة ، والاستدلال الاحصائي يتم باستخدام عينة عشوائية اختيرت من المجتمع وذلك لاستحالة التعامل مع المجتمع كافة.

مثال 1: باحث مهتم بفحص فرضية مفادها ان المدخنين يحتاجون لمعدل من ساعات نوم أقل من المعدل الذي يحتاجه الناس بصورة عامة والذي يقدر بـ (8) ساعات ولهذا الغرض قام الباحث بسؤال (16) مدخناَ ووجد أن معدل نومهم (7.5) ساعة وأن الانحراف المعياري لمعدل نوم أفراد العينة يساوي (1.5) ، قم بفحص الفرضية عند مستوى دلالة ().

الحل:

1- الفرضية الاحصائية

2- نحسب قيمة احصائي الاختبار وفق القانون

-1.333 = = t =

3**- القرار** : بما أن القيمة التائية المحسوبة والبالغة (-1.333) أكبر من القيمة التائية الحرجة والبالغة (2.131) عند درجة حرية (15) وبمستوى دلالة (0.05) ؛ تقبل الفرضية الصفرية أي أن الفرق بين معدل أفراد العينة وافراد المجتمع في ساعات النوم غير جوهري .

**ب- الاختبار التائي للعينات المستقلة – عينتين مستقلتين**

تصادفنا الكثير من المواقف التي نرغب فيها باجراء المقارنة بين مجتمعين ، كأن نقارن بين أداء الذكور بأداء الإناث ، أو أداء الذين درسوا بطريقة معينة كطريقة المناقشة مثلاً ؛ بأداء الذين درسوا بطريقة أخرى كطريقة المحاضرة مثلاً ؛ أو أداء الطلاب على مقياس معين قبل معالجة تجريبية وبعدها ... الخ.

وقبل القيام بذلك إحصائيا فانه لابد من التفريق بين نوعين من الحالات هما البيانات المستقلة Independent والبيانات غير المستقلة Dependent ، إذ ان لكل منهما أسلوبه الخاص في التحليل الإحصائي ، وكما يلي:

يقصد بالبيانات المستقلة تلك البيانات التي لا يوجد فيما بينها ارتباط ، ومن الأمثلة على ذلك أداء مجموعتين على اختبار تحصيلي كانتا قد درستا بطريقتين مختلفتين ، أو أداء مجموعة من الذكور وأخرى من الإناث على مقياس أداء طلاب من مستوى دراسي معين على مقياس للرأي مقابل أداء طلاب من مستوى دراسي أخر على نفس المقياس أو أداء مجموعة تجريبية على اختبار معين وأداء مجموعة ضابطة على نفس الاختبار ... الخ.

وفي جميع الحالات يكون وسط أداء العينة الأولى مستقلاً عن وسط العينة الثانية وتصاغ الفرضيات الاحصائية كما يلي :

الفرضية الصفرية: = :

الفرضية البديلة : :

أما درجات الحرية (df) فستكون () حيث () حجم العينة الأولى و() حجم العينة الثانية وسيكون الاختبار الإحصائي المناسب هو الاختبار التائي (t-test) .

مثال1: لنفترض أن باحثاً أراد أن يعرف فاعلية أسلوب معين في التفكير الإبداعي للأطفال في مستوى السادس الابتدائي فعمل على توزيع (50) تلميذاً عشوائياً في مجموعتين ثم عين عشوائياً إحداهما لتكون مجموعة تجريبية والأخرى ضابطة ، وفي نهاية التجربة أعطيت المجموعتان اختبارا يقيس التفكير الإبداعي وكانت النتائج كما يلي:

|  |  |
| --- | --- |
| المجموعة التجريبية | المجموعة الضابطة |
| 25 = | 25= |
| 7.65 = | 6 = |
| 2.55 = | *= 2.43* |

هل تدل هذه البيانات على أن أداء المجموعة التجريبية كأن أفضل من أداء المجموعة الضابطة على مستوى () ؟

1- صياغة الفرضيات.

الفرضية الصفرية : = :

الفرضية البديلة : :

2- تحدد القيمة الحرجة للرفض أو القبول وبما أن مستوى الدلالة محدد في السؤال وهو ().

وأن درجة الحرية (df = 25 + 25 - 2 = 48) وبذلك تكون قيمة (t) الحرجة تساوي (2.009).

3-نستخدم القانون :

=t

2.43 =

4- نقارن القيمة التائية المحسوبة والبالغة (2.43) بالقيمة التائية الحرجة والبالغة (2.009) تحت درجة حرية (48) ومستوى دلالة (0.05) ، وبما أن القيمة المحسوبة أكبر من القيمة الجدولية ، إذن ترفض الفرضية الصفرية وتقبل الفرضية البديلة ، أي أن البيانات تدل على أن الذين يخضعون للبرنامج التدريبي يصبح أداؤهم في اختبار التفكير الإبداعي أفضل من الذين لا يخضعون للبرنامج التدريبي.

**الفصل الثالث**

**تحليل التباين الأحادي**

* الأساس المنطقي لتحليل التباين
* خطوات تحليل التباين
* تدريبات تدريبات تحليل التباين

**تحليل التباين الأحادي Analysis Of Variance**

لقد توصل فيشر R.A. Fisher من خلال الأبحاث الإحصائية التي قام بها الى طريقة للمقارنة بين مجموعات متعددة بطريق مباشر سميت بـ تحليل التباين.

وقد جاءت هذه الطريقة للمقارنة بين أكثر من مجموعتين ، إذ يستطيع الباحث أن يستخدم الاختبار (t) للمقارنة بين مجموعتين ، ولكن قد يقوم الباحث باستخدام أكثر من مجموعة في دراسته لذا فهو يحتاج للمقارنة بين هذه المجموعات ، ولو أنه أستخدم المقارنة بالاختبار التائي وكانت مجموعات الدراسة أربع مجموعات مثلاً ، فإننا سنحتاج عندئذ الى تطبيق سلسلة من الاختبارات التائية بين المجموعات وكالاتي:

1. أختبار (t) بين المجموعتين (1-2) .
2. أختبار (t ) بين المجموعتين (1-3).
3. أختبار (t ) بين المجموعتين (1-4).
4. أختبار (t ) بين المجموعتين (2-3).
5. أختبار (t ) بين المجموعتين (2-4).
6. أختبار (t ) بين المجموعتين (3-4).

وكلما زاد عدد المجموعات زادت أعداد الاختبارات المطلوبة من اختبار (t) ، وبصورة عامة إذا كان لدينا (n) من المتوسطات فاننا نحتاج الى توافق () مقارنة.

وهذا يجعل عملية إجراء المقارنات باستخدام (t) بغاية الصعوبة من الناحية العملية ، وكذلك إن استخدام اختبار (t) لإجراء مقارنات متعددة يضخم الخطأ من النوع الأول (α) مما يؤدي الى زيادة احتمال رفض ( ( وهي صحيحة ، لذا جاء أسلوب تحليل التباين ANOVA Analysis of variance.

ولاستعمال تحليل التباين ثلاث ميزات مهمة تميزه عن استعمال الاختبار التائي ، هي:

1. احتمالية أقل للوقوع في الخطأ من النوع الأول.
2. أكثر قوة عندما تكون قيمة (a)ثابتة.
3. إمكانية معرفة تأثير متغيرين مستقلين أو أكثر في آن واحد.

وتحليل التباين أنواع عدة منها تحليل التباين الأحادي الذي يعتمد على فحص دلالة الفروق بين ثلاثة متوسطات أو أكثر لمستويات المتغير المستقل ، وهو الأسلوب الذي نطبقه عندما يكون لدينا متغير مستقل واحد بمستويات عدة ، ونود معرفة أثره على متغير تابع ، كما في اختلاف الدافعية نحو العمل اليدوي باختلاف الجنس ، فالجنس هنا متغير مستقل بمستويين (ذكر ، أنثى).

وهناك أنواع أخرى منها تحليل التباين الثنائي 2- Way Analysis Of Variance الذي يستخدم عندما يكون لدينا متغيرين مستقلين مثل متغير النوع ومتغير المستوى الأكاديمي ونود قياس أثرهما على مستوى الدافعية نحو العمل اليدوي ، فالنوع بمستويين (ذكر ، أنثى) والمستوى الأكاديمي بثلاثة مستويات مثلا ً (دبلوم عالٍ ، ماجستير ، دكتوراه) ، وبذلك عند اجراء تحليل التباين الثنائي (2×3) نحتاج الى ست مجموعات كون متغير الجنس بمستويين ومتغير المستوى الأكاديمي بثلاثة مستويات لإجراء المقارنات التي تجيب على تساؤلات معينة حول أثر المتغير المستقل على المتغير التابع.

**الأساس المنطقي لتحليل التباين**

يقوم منطق تحليل التباين على أساس تجزئة التباين الكلي للمشاهدات الى جزأين ، الجزء الأول يتعلق بالتباين بين المجموعات والجزء الثاني التباين داخل المجموعات ، ولتحليل التباين بين المجموعات وداخل المجموعات لابد أن نعرف مصدر هذه التباينات ، فالتباينات بين المجموعات يكون مردها المعالجات Treatments والفروق الفردية Individual differences والأخطاء التجريبية ؛ والتباين داخل المجموعات أما ان يكون مردها الفروق الفردية او الأخطاء التجريبية ، وعليه فان نسبة التباين بين المجموعات الى التباين داخل المجموعات يمثل الإحصائي الذي يمكن ان نستخدمه لفحص الفرضيات حول أوساط أكثر من مجموعتين ويطلق عليه النسبة الفائية F-Ratio والتي تصاغ بالصورة:

F=

أي: F=

نلاحظ من صيغة المعادلة إن المعالجات هي التي تعمل على إيجاد فروق جوهرية بين المجموعات ، وسيتم مقارنة قيمة F المحسوبة بقيمة F الجدولية التي تعتمد قيمتها على نسبة التباين بين المجموعات الى التباين داخل المجموعات تخضع لتوزيع F .

ويرمز لدرجات الحرية (q , p df ) لكل من البسط والمقام على الترتيب.

أما فرضيات تحليل التباين فهي:

1. الفرضية الصفرية لاختبار تحليل التباين تعطي على الصورة:
2. الفرضية البديلة لاختبار تحليل التباين تعطى على الصورة:

فاذا كان القرار رفض فان الاستنتاج الذي يمكن أن نخرج به هو إن بعض قيم تختلف عن بعضها الآخر ولا نستطيع أن نتنبأ بأي المعلمات تختلف عن بعضها ، أما إذا قبلنا الفرضية البديلة فإن ذلك لا يعني إن جميع قيم يختلف بعضها عن بعض ، بل يشير قبولنا للفرضية البديلة اي أن هناك على الأقل معلمان مختلفان عن بعضهما.

**خطوات تحليل التباين**

لمعرفة دلالة الفروق الإحصائية بين المجموعات نتبع الخطوات الآتية:

1. إيجاد مجموع المربعات داخل المجموعات.
2. إيجاد مجموع المربعات بين المجموعات.
3. تحديد درجات الحرية .
4. استخراج النسبة الفائية للكشف عن الدلالة الاحصائية ، وذلك لمعرفة مدى تجانس واختلاف تلك المجموعات .

مثال (1): ................. للتبسيط والتوضيح سيكون المثال الأول ذو مجموعتين فقط ..

أجري إختبار للرياضيات على مجموعة من الطلبة وتألفت المجموعة الأولى من (5) طلاب والمجموعة الثانية من (5) طالبات ، والمطلوب هو حساب الدلالة الاحصائية للفروق القائمة بين درجات الطلاب والطالبات بطريقة تحليل القيم وفق البيانات الآتية :

|  |  |
| --- | --- |
| درجات الطلاب  (x ) | مربع درجات الطلاب () |
| 8 | 64 |
| 16 | 256 |
| 16 | 256 |
| 12 | 144 |
| 13 | 169 |
| **65** | **889** |

889 =

13==

4225 = =

|  |  |
| --- | --- |
| درجات الطالبات  (y ) | مربع درجات الطالبات ) |
| 8 | 64 |
| 3 | 9 |
| 9 | 81 |
| 8 | 64 |
| 7 | 49 |
| **35** | **267** |

267 =

6 = =

1225 = =

الخطوة(1):

الخطوة(2): حساب مجموع المربعات داخل المجموعات:

=

=

إذن: n =5 x 8.8 = 44

إذن: n =5 x 4.4 = 22

إذن: مجموع المربعات داخل المجموعات =44 + 22 = 66

الخطوة(3): حساب مجموع المربعات بين المجموعات:

nn

* نستخرج المتوسط الوزني (m ) =

=

إذن: (m) =10

13 – 10 = 3

7 – 10 = -3

وحيث أن مجموع المربعات بين المجموعات =

n + n

إذن : مجموع المربعات بين المجموعات =90

الخطوة (4): تحديد درجات الحرية:

أ- درجات الحرية داخل المجموعات =

5-1 +5 -1 =8

إذن : درجات الحرية داخل المجموعات = 8

ب: درجات الحرية داخل المجموعات = عدد المجموعات -1

= 2-1 =1

إذن: درجات الحرية بين المجموعات = 1

الخطوة(5): حساب التباين داخل المجموعات وبين المجموعات:

التباين داخل المجموعات ===8.25

التباين بين المجموعات = = = 90

الخطوة (6 ) : حساب النسبة الفائية :F

F = = =10.90

إذن النسبة الفائية = 10.90

الخطوة (7): الدلالة الاحصائية للنسبة الفائية :

بما أن درجات الحرية للتباين الاكبر =1

ودرجات الحرية للتباين الاصغر =8

نجد القيمة (F ) الجدولية عند مستوى دلالة (0.05 )=5.32

وعند مستوى دلالة (0.01 )=11.26

وبما أن (F) المحسوبة والبالغة 10.90))هي أكبر من القيمة الجدولية عند مستوى دلالة (0.05) إذن هناك فروق بين المجموعتين فهما مستقلتان وليس من أصل واحد ، لذا ترفض الفرضية الصفرية وتقبل الفرضية .

ولكن عند مستوى دلالة (0.01 ) نجد إن F = 10.90 وهي أقل من قيمة (F) الجدولية والبالغة (11.29) لذا تقبل الفرضية البديلة ، أي ان نسبة (F) دالة احصائياً عند مستوى دلالة 0.05 وغير دالة احصائياَ عند مستوى دلالة 0.01 .

الخطوة (8) ملخص تحليل التباين وكما يأتي:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| مصدر  التباين | درجات الحرية | مجموع المربعات | متوسط مجموع المربعات | F | الدلالة الاحصائية |
| بين المجموعات | 1 | 90 | 90 | 10.90 | دالة عند 0.05  وغير دالة عند 0.01 |
| داخل المجموعات | 8 | 66 | 8.25 |
| المجموع | 9 | 156 |

**الفصل الرابع**

**الاختبارات اللامعلمية – مربع كاي**

* الاحصاءات اللامعلمية لقياس الفروق بين عينتين
* اختبار مربع كاي للاستقلالية
* اختبار مربع كاي لحسن المطابقة
* تدريبات مربع كاي

**الاختبارات اللامعلمية Non-Parametric Tests**

في كثير من المجالات التطبيقية يكون توزيع المجتمع الذي سحبت منه العينة غير معروف ويرغب الباحث في اجراء استدلال احصائي حول معالم المجتمع ، وفي هذه الحالة يعتمد على التوزيع الفعلي للعينة Empirical Distribution لذلك يجب استعمال الاختبارات اللامعلمية.

وهناك مجموعة من المزايا التي تتمتع بها الاختبارات اللامعلمية منها:

* عند اجراء الاختبار اللامعلمي لانفترض معرفة معلومات عن توزيع المجتمع.
* الاختبارات اللامعلمية أسهل وأسرع في اختبارات الفروض ولذلك تستخدم كثيراً.
* في الاختبارات اللامعلمية لا نفترض معرفة قيمة المفردات ونكتفي بترتيب المفردات.
* الاختبارات اللامعلمية تجري دائماً لاختبار بعض الفروض الخاصة بمقاييس النزعة المركزية أو مقاييس التشتت ولكن الاختبارات اللامعلمية يمكن أن تجرى لدراسة أي خصائص في المجتمع حتى وان كانت هذه الخصائص وصفية .

وللاختبارات اللامعلمية مجموعة من العيوب منها:

* إذا كان المجتمع المسحوب منه العينة يتبع التوزيع المعتدل فان الاختبار اللامعلمي الذي يعطي نفس قوة الاختبار المعلمي يحتاج الى حجم عينة أكبر.
* عندما تكون قوة الاختبار اللامعلمي بالنسبة للاختبار المعلمي غير معروفة بالضبط فان حجم العينة اللازم لاجراء اختبار لا معلمي يغطي نفس قوة المعلمي لا يمكن تحديده بالضبط.

**الاحصاءات اللامعلمية لقياس الفروق بين عينتين**

تستخدم الاختبارات اللابارامترية للكشف عن دلالة الفروق بين متوسطي عينتين عندما :

1. لا تتوافر شروط استخدام الاختبار التائي (t-test) كأن تكون مفردات العينتين صغيرة.
2. عندما يكون توزيع أحد العينتين غير اعتدالي أو ملتوي بدرجة كبيرة.
3. عندما يكون تباين sample variance مختلف بصورة كبيرة عن بعضها.

وهنا يفضل إستخدام الرتب فضلاً عن القيم الاصلية في حساب دلالة الفروق بين متوسطي عينتين ، كما يمكن استخدام الرتب في حساب معامل الارتباط وقوة العلاقة بين متغيرين.

ولعل أهم الاختبارات اللابارامترية هو اختبار مربع كاي () ، والذي يستعمل بمقارنة قيمته المحسوبة (Observed) بقيمته النظرية (الجدولية) بدرجة حرية مقدارها (1) ، فاذا كانت القيمة المحسوبة تساوي أو أكبر من الجدولية فمعنى ذلك إن هناك ارتباط بين المتغير الاول والثاني (من خلال مقارنة الفروق بين القيم الملاحظة والمتوقعة) ، ومن ثم يمكن رفض الفرض الصفري ، أما إذا كانت () المحسوبة أقل من () الجدولية فلا وجود لهذه العلاقة أو أن هذين المتغيرين مستقلان عن بعضهما البعض.

كما ويستعمل () لاختبار مدى إتفاق توزيع القيم الملاحظة مع التوزيع المتوقع ، ويستخدم في حالة البيانات الاسمية.

أما في البحوث التربوية والنفسية فإنه يستعمل في حالة إيجاد الصدق الظاهري للمقياس أو الاختبار والذي يعد أحد مؤشرات صدق المحتوى ، كما يستعمل في البحوث التجريبية في حالة تكافؤ المجموعتين التجريبية والضابطة في متغيري (الشهادة والمهنة) مثلاً.

سنحاول عرض أنواع () وكما يلي:

أولاً: **إختبار مربع كاي للاستقلالية** **Chi Square test for Independence**

*أ- اختبار الدلالة الاحصائية لعينة واحدة*

مثال1: لنفرض اننا طبقنا اختباراً معيناً على مجموعة من التلاميذ في محاولة لمعرفة رأيهم في طريقة التدريس التي يتبعها المعلم معهم ، وكان عدد التلاميذ (40) تلميذا ً، ولنفرض ان منهم (28) تلميذاً قالوا بإنها طريقة جيدة ، أي يحبذون هذه الطريقة وان منهم (12) تلميذاً قالوا إنها طريقة غير مناسبة.

فهل هناك فروق دالة إحصائياٌ بين التلاميذ الذين يحبذون الطريقة والذين قالوا إنها غير مناسبة ، اختبر ذلك عند مستوى دلالة (0.05).

الحل:

1. صياغة الفرضيات:

E O :

حيث : (O) التكرار الملاحظ

(E) التكرار المتوقع

2. تحديد التكرار المتوقع من خلال تقسيم المجموعة الى نصفين أي (50%) من التلاميذ يحبذون طريقة و(50%) آخرون لا يحبذونها .

فأن التكرارات التي نتوقعها = =20

3. نطبق قانون مربع كاي () بالمعادلة التالية :

حيث أن:

: O التكرار الملاحظ

E : التكرار المتوقع

ومن خلال تطبيق القانون نحصل على ()

+

6.4 = + =

4. وبالرجوع الى جداول () بدرجة حرية (1) ومستوى دلالة (0.05) نجد أن قيمة () النظرية = 3.84

5. القرار: نقارن قيمة () المحسوبة مع القيمة النظرية لمربع كاي ، بما أن قيمة () المحسوبة والبالغة (6.4) وهي أكبر من قيمة () الجدولية (النظرية) والبالغة (3.84) لذا ترفض الفرضية الصفرية وتقبل الفرضية البديلة أي أن هناك فروقاً ذات دلالة إحصائية بين التكرارات الملاحظة والتكرارات المتوقعة.

(المعهد الوطني ، 2005 : 107- 108)

مثال2: أراد معلم أن يعرف اليوم الذي يفضله تلاميذه للذهاب الى المكتبة للمطالعة ، وكانت النتائج التالية:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| اليوم | السبت | الاحد | الاثنين | الثلاثاء | الاربعاء | الخميس |
| التكرار | 6 | 7 | 5 | 7 | 15 | 20 |

فهل هنلك فروقا دالة أحصائياً بين تفضيل التلاميذ للايام ، اختبر ذلك ؟

1. صياغة الفرضيات

E O :

2- التكرار المتوقع = = =10

3-

=

=

1.6 + 0.9 + 2.5 + 0.9 + 2.5 +10 = 18.4=

4 – القيمة الجدولية تحت درجة حرية (5) ومستوى دلالة (0.05) = 11.070

5 – القرار: بما أن قيمة المحسوبة والبالغة (18.4) أكبر من قيمة الجدولية والبالغة (11.070) ؛ إذن تقبل الفرضية البديلة وترفض الفرضية الصفرية ، أي يوجد فرق دال إحصائياً بين تفضيل التلاميذ ليوم الذهاب الى المكتبة.

*ب- إختبار الدلالة الاحصائية لعينتين مستقلتين*

إختبار مربع كاي () للإستقلالية يستعمل من قبل الباحثين للمقارنة بين عينتين مستقلتين كل منهما ذات بيانات اسمية ثنائية التصنيف ، وفيما اذا كانت نفس العينتين هما حقاً من نفس المجتمع أم لا ؟

وللتعرف على ذلك ندرج المثال الاتي:

مثال2: نفرض ان أحد الباحثين قام بإختبار عينة عشوائية من الطلبة الذكور والاناث في المرحلة المتوسطة وإستطلاع آرائهم بالفرع الذي يرغبون إختياره لمواصلة الدراسة في المرحلة الثانوية (علمي ، أدبي) وفق البيانات الاتية:

|  |  |  |  |
| --- | --- | --- | --- |
| الجنس التخصص | علمي | أدبي | مجموع |
| ذكور | 110 (A) | 90 (B) | 200 () |
| إناث | 40 (C) | 60 (D) | 100 () |
| محموع | 150() | 150 () | 300 (n) |

فهل هناك فروقاً دالة إحصائياً بين استجابات الذكور واستجابات الاناث ؟ إختبر ذلك عند مستوى دلالة (0.01).

الحل :

1 – صياغة الفرضية:

E O

2 – نجد التكرار المتوقع من خلال الجدول وكالاتي:

التكرار المتوقع للخلية (A ) = = = 100

التكرار المتوقع للخلية (B ) = = = 100

التكرار المتوقع للخلية (C ) = = = 50

التكرار المتوقع للخلية (D ) = = = 50

3 – نطبق قانون مربع كاي:

=

إذن قيمة مربع كاي المحسوبة هي (6).

4 – نستخرج القيمة الجدولية (النظرية) لدرجة حرية (1) وعند مستوى دلالة (0.01) والتي بلغت (6.64).

5 – القرار: بما ان القيمة المحسوبة لمربع كاي والبالغة (6) أصغر من القيمة الجدولية والبالغة (6.64) ؛ إذن تقبل الفرضية الصفرية وترفض الفرضية البديلة ، أي إن البيانات التي حصل عليها الباحث تؤيد الفرضية الصفرية.

**الفصل الخامس**

**الاختبارات اللامعلمية – مان وتني**

* اختبار مان وتني لعينتين مستقلتين
* تدريبات اختبار مان وتني للعينات الصغيرة
* تدريبات مان وتني للعينات المتوسطة
* تدريبات مان وتني للعينات الكبيرة

**إختبار مان وتني (ي) لعينتين مستقلتين Mann Whitney Test (U)**

وهو من الاختبارات الإحصائية اللامعلمية للمقارنة بين العينات المستقلة ، ويعد من الأساليب الإحصائية التي شاع استخدامها في التحليلات الإحصائية بشكل كبير في السنوات القليلة الماضية ، ويستخدم للمقارنة بين عينتين مستقلتين عندما تكون البيانات عددية بطبيعتها ، وهو في الغالب يستخدم بدلاً عن الاختبار التائي.

واختبار مان وتني يستند الى أساس كون الدرجات الخاصة بمجموعتين متشابهتين مرتبة معاً وكأنها مجموعة واحدة ، فانه سيكون تمازج بين رتب المجموعتين ، ولكن إذا تفوقت أحدى المجموعتين على الأخرى فان معظم رتب المجموعة المتفوقة ستكون أعلى من المجموعة الدنيا ، لذا فان قيمة (U) تحسب بعد دمج رتب المجموعتين معاً ، ثم يحسب عدد الرتب الخاصة بالمجموعة العليا والتي تقع تحت رتب المجموعة الدنيا.

ويمكن استخدام اختبار (U) في حالة العينات الصغيرة جداً التي لا يتجاوز عدد أفرادها (8) كما يمكن استخدامه في حالة العينات ذات الأحجام المتوسطة 9)-20) وكذلك العينات التي يزيد عدد أفرادها عن (20).

لذلك يستخدم (U) بحسب واحدة من ثلاث طرائق وفقاً لحجم كل من العينتين التي تجري عليهما المقارنة وكما يلي:

**أ- اختبار مان وتني للعينات الصغيرة**

مثال: لنفرض أن احد الباحثين اختار عينتين عشوائيتين تتألف العينة الأولى من (5) أفراد وتتألف العينة الثانية من (3) أفراد ، ثم قام بتطبيق اختبار معين على أفراد العينتين ، ثم حصل على البيانات التالية:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| درجات العينة (A) | 10 | 12 | 13 | 18 | 21 |
| درجات العينة (B) | 9 | 14 | 15 |  |  |

المطلوب هو التحقق من وجود فرق ذو دلالة إحصائية بين درجات العينتين ؟

الحل:

الخطوة 1: صياغة الفرضيات.

1. الفرضية الصفرية ( ) : لا يوجد فرق ذو دلالة إحصائية بين درجات العينتين.
2. الفرضية البديلة ( ) : يوجد فرق ذو دلالة إحصائية بين درجات العينتين.

الخطوة 2: نفرض إن عدد أفراد العينة (A) =

نفرض إن عدد أفراد العينة (B) =

أي أن : 5*= ، 3*

الخطوة 3: نقوم بتنظيم جدول لدرجات العينتين معاً بشكل تصاعدي من الأصغر الى الأكبر وكما يأتي:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| الدرجات | 9 | 10 | 12 | 13 | 14 | 15 | 18 | 21 |
| العينة | B | A | A | A | B | B | A | A |

الخطوة 4: نستخرج (U) وذلك بحسب عدد الدرجات من المجموعة (A) والتي ترتيبها تحت أو يسبق كل درجة من درجات المجموعة (B) أي إن الدرجة الأولى (9) من المجموعة (B) لا تسبقها أية درجة من درجات المجموعة (A) ، أي أن هناك (صفر) من درجات المجموعة (A) أقل من الدرجة الأولى للمجموعة (B) ، أما الدرجة الثانية في المجموعة (B) فهي (14) نلاحظ إن هناك ثلاث درجات في المجموعة (A) دون (14) وهي (10 ، 12 ، 13) على التوالي ، كما إن هناك ثلاث درجات من المجموعة (A) تقع تحت الدرجة ( 15 ) من المجموعة (B).

إذن: 6= 3 +3 + صفر U =

وبنفس الطريقة يمكن حساب عدد المرات التي يكون فيها (B) أقل من (A) فيكون:

1+1+1+3+3 = 9

الخطوة 5: نلاحظ أن هناك قيمتان لـ مان وتني هي (6 ، 9) نختار القيمة الأصغر وهي (6) وبالرجوع الى جداول مان وتني لاستخراج القيمة النظرية من خلال الجدول المناسب لحجم العينة الأكبر وهو =5 ونلاحظ العمود الأول والذي يبدأ من الرقم (صفر–13 ) نلاحظ أن قيمة مان وتني المحسوبة والبالغة (6) ما يقابلها تحت =3 هي (0.393).

الخطوة 6: القرار: بما ان القيمة المحسوبة لـ مان وتني والبالغة (6) أكبر من القيمة النظرية والبالغة (0.393) إذن تقبل الفرضية الصفرية وترفض الفرضية البديلة ، وهذا يعني إن الفرق بين العينتين غير ذي دلالة إحصائية (توفيق، 1985 :156-159).